

## RESEARCH ARTICLE

# An investigation on the relationship between the Hurst exponent and the predictability of a rainfall time series

Sivapragasam Chandrasekaran<sup>1</sup> | Saravanan Poomalai<sup>1</sup>  | Balamurali Saminathan<sup>2</sup> | Sumila Suthanthiravel<sup>1</sup> | Keerthi Sundaram<sup>1</sup> | Farjana Farveen Abdul Hakkim<sup>1</sup>

<sup>1</sup>Center for Water Technology, Department of Civil Engineering, Kalasalingam Academy of Research Education, Virudhunagar, India

<sup>2</sup>Department of Computer Applications, Kalasalingam Academy of Research Education, Virudhunagar, India

**Correspondence**

Saravanan Poomalai, Center for Water Technology, Department of Civil Engineering, Kalasalingam Academy of Research Education, Krishnankoil, Srivilliputtur (taluk), Virudhunagar (Dist), Tamilnadu, India.  
Email: p.saravanan.me.1986@gmail.com

Rainfall prediction is a very challenging task due to its dependence on many meteorological parameters. Because of the complex nature of rainfall, the uncertainty associated with its predictability continues to be an issue in rainfall forecasting. The Hurst exponent is considered as a measure of persistence and it is believed that if a time series has persistence (as reflected by a Hurst exponent value greater than 0.5) it is also predictable. However, very limited studies have been carried out to demonstrate this hypothesis. This study, through experimental work on hypothetical data as well as real data, demonstrates that the Hurst exponent can be taken as an indicator for predictability provided that the exponent values at “smaller levels” of the time series are also significantly greater than 0.5 together with the Hurst exponent of the overall time series. It is also shown that it is better to predict the “similarity” aspect associated with a time series (and derive the predicted rainfall) than to predict the rainfall directly.

**KEYWORDS**

artificial neural network, fractals, Hurst exponent, persistence, rainfall predictability

## 1 | INTRODUCTION

Since rainfall directly affects the flood and drought of a country, it is a very important meteorological parameter to predict. Due to the dependence on other meteorological parameters (temperature, wind speed, relative humidity and pressure), the prediction of rainfall becomes difficult. As a measure to address this issue, a review of previously reported work indicates that researchers have resorted to different approaches to preprocess the inputs or include additional information which can aid in prediction. For instance, Ahmadi *et al.* (2009) examined the relationship between monsoon rainfall in Iran and a large scale climate signal such as sea surface pressure, sea level pressure difference over the effective region and sea level temperature, to forecast the monthly rainfall. A fuzzy rule model was developed for predicting a 6 month scale using statistical models. Hung *et al.* (2009) applied an artificial neural network (ANN) model for predicting the real time hourly rainfall using a combination

of meteorological parameters such as relative humidity, air pressure, wet bulb temperature and cloudiness as well as rainfall at surrounding stations and the rainfall at the point of forecasting. Based on the model performance, a sensitivity analysis was done to identify the dominant model input parameters. In addition, the input preprocessing method, spectrum analysis method, mutual information technique and wavelet analysis have also seen applications in rainfall prediction (Wu *et al.*, 2010; Nourani *et al.*, 2011; Goyal, 2014; Babel *et al.*, 2015).

Researchers have also reported forecasting accuracy in terms of error measures other than the root mean square error (RMSE) and the correlation co-efficient in order to represent the predictive ability of the models proposed better. Kumarasiri and Sonnadara (2008), for instance, proposed short term and long term rainfall forecasting models reporting prediction in terms of success rates within an error limit. Similarly, Olsson *et al.* (2004), instead of predicting actual rainfall intensities directly, used a neural network to determine

rainfall occurrence and rainfall intensity during rainy periods, and after categorizing rainfall into intensity categories these categories were predicted. From the reported work on rainfall forecasting, it is also observed that different researchers have proposed different approaches to rainfall prediction and no single method or technique can be accepted as universal as far as rainfall forecasting is concerned. For instance, Kar *et al.* (2012) proposed a multi-model ensemble scheme for predicting monthly rainfall in India during July. They observed that, whereas skills of rainfall prediction over eastern parts of India were significantly improved, the same conclusions were found to be invalid for some other regions because rainfall pattern, intensity and amount are affected by many complex parameters such as sea surface temperature and sea level pressure in the Pacific and Indian Oceans (resulting in the occurrence of the El Niño Southern Oscillation) and the Indian Ocean Dipole (Houston, 2006; Ashok *et al.*, 2007).

From the foregoing discussion, it can be concluded that before attempting rainfall prediction with statistical methods it is better to ascertain first whether the given rainfall time series is predictable or not. Researchers have used many methods to analyse the complexity of a time series including the Hurst exponent, deterministic index and spectral clustering (Rehman and El-Gebeily, 2009; Tatli, 2014; Tatli and Dalfes, 2016). Since the Hurst exponent cannot be used as a prognostic tool, the present study uses the Hurst exponent as a diagnostic tool for investigating the complexity of a time series. A time series may contain both short term memory and long term memory associated with the event. So, correct identification of the memory effect will aid in identifying the pattern associated with the given time series which when modelled can lead to reliable forecasting accuracy. While short term memory can be estimated using an autocorrelation function (ACF) (and consequently relevant antecedent rainfalls can be considered as inputs), the identification of long term memory poses certain difficulties particularly when dealing with limited time series records. Conventionally, researchers have reported estimation of the Hurst exponent as an indicator of long term memory and hence persistence of a prevailing trend or pattern in a given time series. A Hurst exponent value greater than 0.5 indicates good persistence of the time series while a value less than 0.5 indicates anti-persistence. A Hurst exponent of 0.5 indicates Brownian motion behaviour. Work by researchers such as Tatli (2015) indicates that the term “persistence” may also be considered as a criterion to be applied as a predictability measure. Many earlier researchers also inferred similar conclusions (Rangarajan and Sant, 1997, 2004; Kalauzi *et al.*, 2009; Rehman and Siddiqi, 2009; Salomao *et al.*, 2009; Li *et al.*, 2012). Although such conclusions are reported, to the authors’ knowledge only a few studies have actually been done to demonstrate the Hurst exponent as a measure of predictability. For instance, Peyghami and Khanduzi (2012)

reported very good accuracy in one lead forecast using an ANN for automotive price time series data with a higher Hurst exponent. Similarly, in rainfall prediction studies, the work reported by Khalili *et al.* (2016) is the only study which shows a good one lead prediction of a monthly time series with a higher Hurst exponent. Since only limited work has been reported on the relationship between rainfall predictability and the Hurst exponent, it is necessary to have a more detailed understanding of this relationship. This study investigates whether rainfall time series with a Hurst exponent greater than 0.5 is always predictable or not. It is proposed to employ the rescaled range ( $R/S$ ) analysis method for estimation of the Hurst exponent. Further, since ANN is very extensively used in studies related to forecasting of hydrological variables including rainfall, the same has been accepted in this study for one lead time forecast (Chattopadhyay and Chattopadhyay, 2008; Aksoy and Dahamsheh, 2009; Babel *et al.*, 2015). The rest of the paper is organized as follows. First the methodology used in this study is briefly discussed together with the data used for the analysis. The next section is devoted to a discussion of the results. In order to investigate the relationship between the Hurst exponent and the predictability of the rainfall time series, it is desired to conduct the experiments first for a set of hypothetical rainfall time series and then for the real data. Finally conclusions are arrived at.

## 2 | METHODS AND DATA

### 2.1 | Rescaled range ( $R/S$ ) analysis method

Several techniques/methods are preferred by various researchers for estimating the Hurst exponent of which the  $R/S$  analysis method is commonly used (Taqqu *et al.*, 1995; Kendzioriski *et al.*, 1999; Feng and Zhou, 2013; Tatli, 2015). In the work reported by Rao and Bhattacharya (1999), the robustness of the  $R/S$  method in comparison to the fractional Gaussian noise model is established for estimation of the Hurst exponent and is recommended for use. Hence, the  $R/S$  analysis method was adopted in this study. For a non-stationary time series,  $R/S$  analysis might give a Hurst exponent greater than 1 (Setty and Sharma, 2015), in which case the detrended fluctuation analysis method can be considered.

$R/S$  analysis is a means of characterizing the time series, the operation of which is summarized as follows. Splitting time series into many shorter series is the first step. Hurst (1951) proposed the following five general equations for  $R/S$  analysis which can be used for any time series and is not restricted to series in Brownian motion alone:

$$\left(\frac{R}{S}\right)_s = k s^H \quad (1)$$

where  $k$  is a constant and  $s$  is the length of each of the shorter time series;  $1 \leq s \leq N$ ,  $N$  being the entire length of

the time series.  $R$  is the range of the time series and  $S$  is the standard deviation.

The range of each size is calculated as:

$$R = \max(z_1, z_2, z_3, \dots, z_s) - \min(z_1, z_2, z_3, \dots, z_s) \quad s = 1, 2, \dots, N \quad (2)$$

$z_s$  is the cumulative series estimated as:

$$z_s = \sum_{i=1}^s y_i \quad (3)$$

where  $s = 1, 2, \dots, N$ .  $y_s$  is the adjusted time series estimated by subtracting the sample mean from each of the shorter time series as:

$$y_s = x_s - \bar{x} \quad (4)$$

where  $s = 1, 2, \dots, N$  and

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad (5)$$

The Hurst exponent is estimated as the slope of the line plotted between  $(R/S)_s$  and  $s$  on a log–log scale.

## 2.2 | Artificial neural network

An ANN can model a complex, nonlinear time series without assuming a relationship between the input and output variables. Since the Hurst exponent indicates nonlinearity in terms of the predictability of a time series, it is proposed to use an ANN because of its ability to model complex nonlinear processes. The choice of an ANN in this study compared to nonlinear multiple regression is justified in view of the reported success of ANNs compared to the latter in various applications (Srivastava and Tripathi, 2012; Mahmoodi and Naderi, 2016; Vinoth *et al.*, 2016; Yildirim *et al.*, 2017).

The ANN is a statistical regression tool motivated by biological neural networks which are used in machine learning. These networks correspond to a system of interconnected “neurons.” The neurons are linked through connections or communication channels to transmit numeric data encoded in different ways. Each neuron functions only when data are received through the connections. The iterative learning process, characteristic of neural networks, adjusts the data cases each time. An appropriate ANN architecture has to be designed for the given set of inputs and outputs in the input and output layers respectively and by selecting the hidden neurons. The data series has to be divided into training, testing and validation, and the best combination is arrived at only by trial and error. The stopping criterion has to be set based on the performance of the network in the test set or any other suitable criterion such as the total number of epochs. As the patterns are introduced to the network, the input signals are transferred to output signals by means of activation functions such as sigmoid, hyperbolic and Gaussian. The correct choice of activation

function is also done through trial and error. Many different types of training algorithms are available with various characteristics and performances, among which the back propagation algorithm is most commonly used. This algorithm updates the weights of the network by adjusting the error between observed and predicted, finally leading to a trained network after repeating this process a sufficient number of times (Haykin, 1999; ASCE Task Committee, 2000a; 2000b; Nourani *et al.*, 2011; Ajmera and Goyal, 2012; Siva-pragasam *et al.*, 2014). Many researchers have used an ANN as a “forecasting model” in the field of atmospheric sciences and meteorology (Gheiby *et al.*, 2003; Tapiador *et al.*, 2004; Chattopadhyay and Chattopadhyay, 2008; Dahamsheh and Aksoy, 2009; Babel *et al.*, 2015; Collins and Tissot, 2015; Valipour, 2016; Modarres *et al.*, 2018).

The performance of the forecast is assessed by the RMSE goodness-of-fit measure defined by the following equation:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \{(X_m)_i - (X_s)_i\}^2} \quad (6)$$

where  $X$  is the variable that is being forecasted; the subscripts  $s$  and  $m$  represent the simulated and measured values, respectively, and the total number of training records is  $n$ .

### 2.2.1 | A note on ANN training

Of the total available data, about 64% is used for training, 21% for testing and 15% for validation for the case studies considered (on average), although the optimal combination may vary from model to model. The multilayer feedforward back propagation algorithm is used. The ANN training is terminated when the mean square error for the test data is minimum. The simplest architecture consisting of one input layer, one hidden layer and one output layer is adopted for all the case studies, with the optimal number of hidden neurons arrived at by trial and error in each case. A more complex architecture is not preferred due to the limited dataset available for model training which might pose serious problems in convergence. Apart from the selection of architecture, the number of epochs, the choice of activation functions and the learning rate are also suitably refined to arrive at the best model for each of the case studies considered.

## 2.3 | Data description

Two sets of hypothetical data of rainfall time series are generated. The first hypothetical monthly rainfall time series (HS1) is generated randomly between a rainfall range of 0–100 mm for the first half and 100–200 mm for the second half and is shown in Figure 1. The second hypothetical rainfall time series (HS2) is generated randomly between a rainfall range of 100–400 mm and is also shown in Figure 1. The first series is designed to indicate a strong increasing

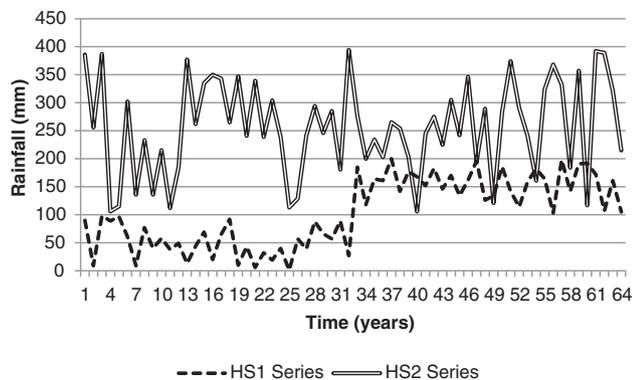


FIGURE 1 Hypothetical rainfall time series HS1 and HS2

trend and the second series a weak trend. This is confirmed by trend analysis which indicated that the first hypothetical time series has a statistically significant positive trend while the second series indicated a positive trend which is not statistically significant.

For the real time data, monthly rainfall records from Cherrapunji were collected from the Indian Water Portal website for the period between 1901 and 1964. Cherrapunji is a subdivisional town in the East Khasi Hills district in the Indian state of Meghalaya and is commonly known to be the wettest part on the Earth. It is located at  $25.30^{\circ}$  N  $91.70^{\circ}$  E with an average elevation of 1,484 m, average temperature of  $11.5^{\circ}$  C and an estimated average annual precipitation of 11,777 mm. The rainfall time series for June and September is shown in Figure 2.

### 3 | RESULTS AND DISCUSSION

#### 3.1 | Hypothetical rainfall time series

First the hypothetical time series HS1 and HS2 are considered which do not represent any actual physical area. The ACF is calculated for both series. A typical ACF equation is shown in Equation 7 based on Box and Jenkins (1976). The lag  $k$  ACF defined for the given measurement of  $x_1, x_2, \dots, x_n$  at time  $t_1, t_2, \dots, t_n$  is:

$$\rho_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (7)$$

Here it is assumed that the observations are equally spaced and hence the time variable  $t$  is not used.

The results of the ACF for both series indicate the absence of short term memory in HS2. The inclusion of antecedent rainfall in the ANN model building for HS1 did not improve the prediction. The Hurst exponents are estimated for both the time series HS1 and HS2 with values of 0.96 and 0.76 respectively. Since the Hurst exponent is significantly higher than 0.5 in both the cases indicating strong persistence, it is expected that the one lead forecast will also be good. ANN models are developed for HS1 and HS2 with

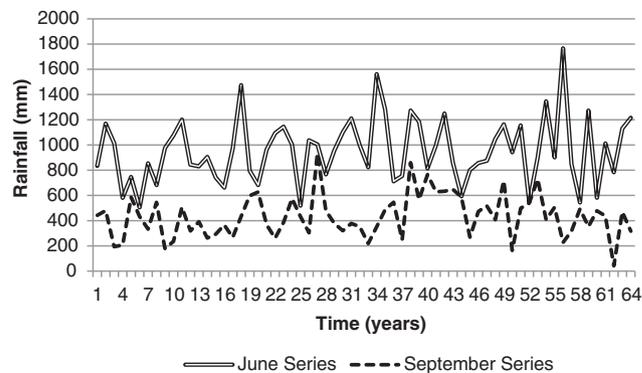


FIGURE 2 Rainfall time series for June and September

40 data for training, 13 for testing and the remaining 10 data for validation (as the optimal combination). The optimal architecture was found to be 1:7:1 for the HS1 using logistic functions whereas it was found to be 1:9:1 with the hyperbolic tangent function for HS2. The input and output can be related functionally as:

$$P_{t+1} = f(P_t) \quad (8)$$

where  $P_t$  is the precipitation at time  $t$  and  $P_{t+1}$  is the precipitation at time  $t + 1$ .

The one lead forecast for the validation set for HS1 and HS2 is shown in Table 1 and plotted in Figures 3 and 4. As seen from Figures 3 and 4, the prediction for HS2 is almost a failure compared to HS1 which has an RMSE of 53.92 mm. Whilst a reasonably good prediction accuracy for HS1 can be attributed to the high Hurst exponent (0.96), with a Hurst exponent of 0.76 for HS2 the prediction is not expected to fail.

A more detailed analysis was carried out to understand the reason for this failure by estimating the Hurst exponent by considering the time series at different levels. The first level consists of the entire data series with a size of 64. The second level consists of a data size of 32 (which is obtained by dividing the entire series into two halves). The third level consists of a data size of 16 (obtained by dividing the entire

TABLE 1 Comparison of Hurst exponent and RMSE value for both the hypothetical rainfall series

Parameters	Dataset	
	HS1	HS2
Hurst exponent values		
$H$ (full series)	0.96	0.76
$H_1$	1	0.51
$H_2$	0.84	0.86
$H_{T1}$	0.98	0.35
$H_{T2}$	0.83	0.95
$H_{T3}$	0.94	0.9
RMSE (mm) for prediction without SF		
Training	34.18	79.39
Testing	24.01	67.28
Validation	53.92	97.01

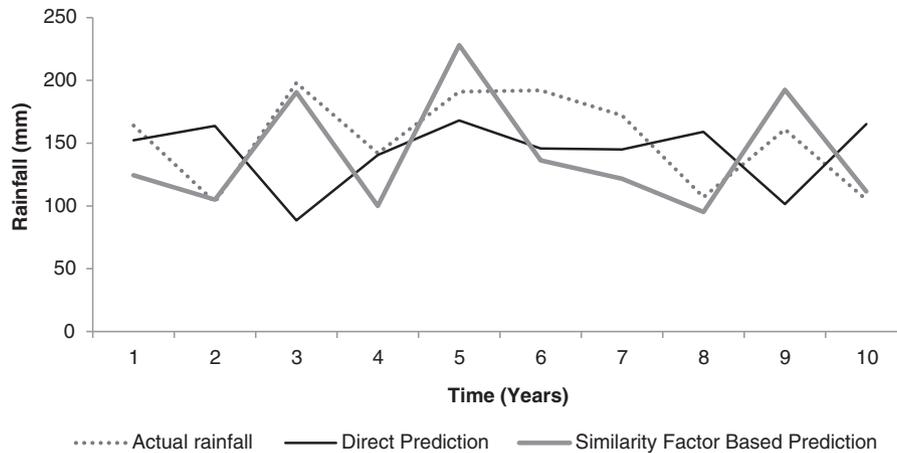


FIGURE 3 Comparison of rainfall prediction for HS1

series into four equal parts). Similarly, the data size can be obtained for the fourth to sixth levels. The Hurst exponent is estimated for the first three levels ( $H_1$ ) and the next three levels ( $H_2$ ) by plotting  $\log R/S$  against the  $\log$  of data size. The Hurst exponent is also estimated for the first two levels ( $H_{T1}$ ), the next two levels ( $H_{T2}$ ) and the last two levels ( $H_{T3}$ ) of data size (i.e. one-third of the series is considered at a time). This is tabulated for both the hypothetical rainfall series in Table 1. An interesting understanding emerges from a closer look at all the levels of Hurst exponents in Table 1. It is seen that, while for the HS1 the Hurst exponents at all levels are greater than 0.5, for HS2  $H_1$  is 0.51 and  $H_{T1}$  is 0.35. These differences reflect an understanding of the time series about its self-similar statistical structure. As noted by Tatli (2015), the Hurst exponent can be used to detect statistical self-similarity when a fractal structure indicates randomness at the local level and deterministic behaviour at the global level. Thus, for the prediction to be good, it is desired that not only the overall Hurst exponent is significantly greater than 0.5 but also the Hurst exponents at different levels should reflect the same tendency (as indicated in Table 1 for the HS1). In such cases, it is expected that there will be some pattern in the time series (or self-similarity) and hence the chances of prediction increase. With this inference, an attempt is made to improve the prediction of the

HS1 by finding the relationship between input and output data. The one lead ahead rainfall is divided by the current rainfall to get the similarity factor (SF), which is taken as the output instead of the actual output for the training and testing set as represented in the functional form:

$$SF_{t+1} = f(P_t) \tag{9}$$

where  $SF_{t+1}$  is the similarity factor at time  $t + 1$ . These data are now trained using the ANN and validated for the validation set. The RMSE is found to be 34 mm. Figure 3 shows the predicted values when rainfall is directly predicted and when it is predicted using the SF. As seen from the figure, the actual rainfall output and the predicted rainfall output match closely when the SF is used in the training process. The prediction accuracy, however, might be considerably improved if more data were used for the ANN training. A similar attempt is made with HS2 to improve the prediction. As seen from Figure 4, there is only marginal improvement in the predicted rainfall. This confirms our earlier inference that the prediction of a time series can be ensured more confidently if both the overall Hurst exponent and the Hurst exponents at different levels of the time series are greater than 0.5. It should be noted that since the rainfall time series invariably has large variations (due to the random nature of the rainfall event), it will be more useful to divide the entire

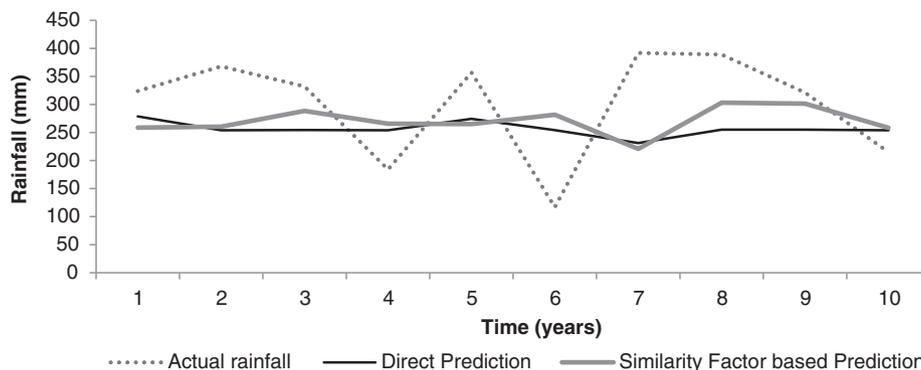
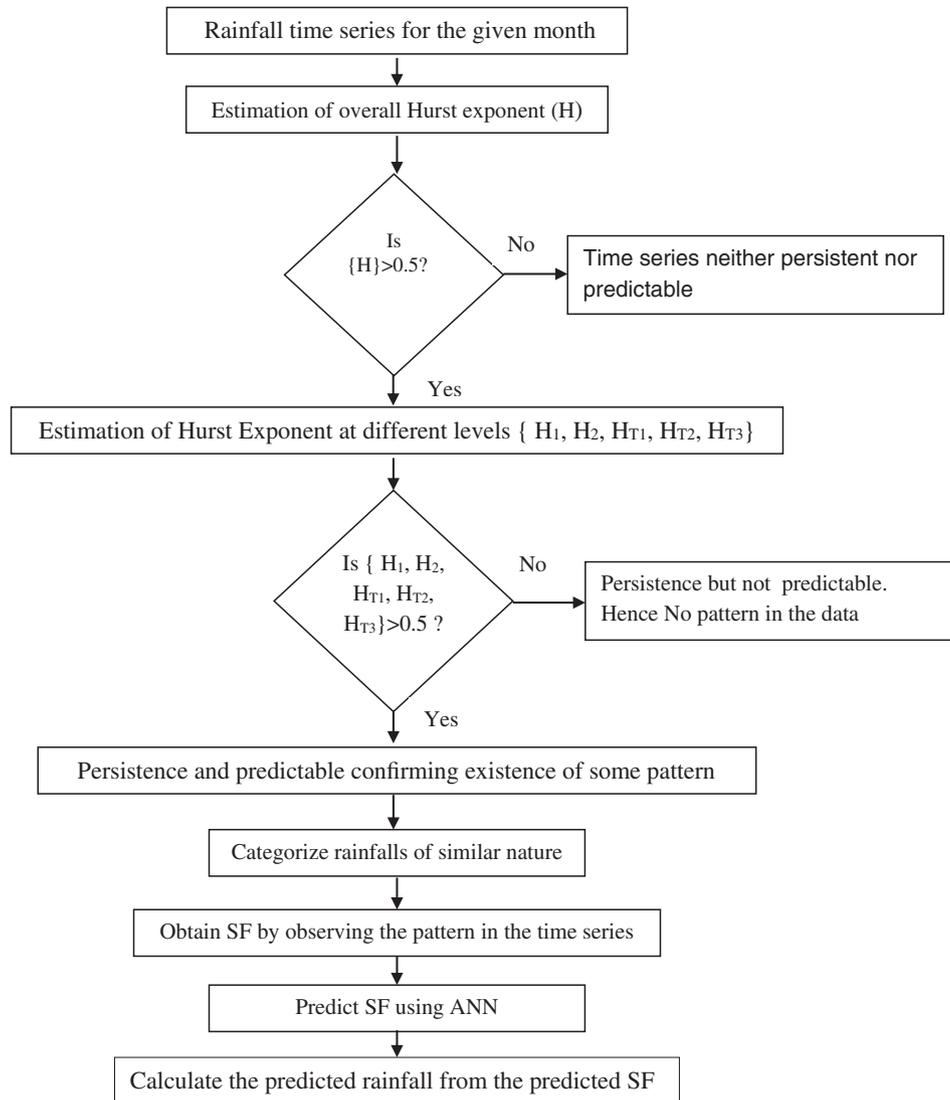


FIGURE 4 Comparison of rainfall prediction for HS2

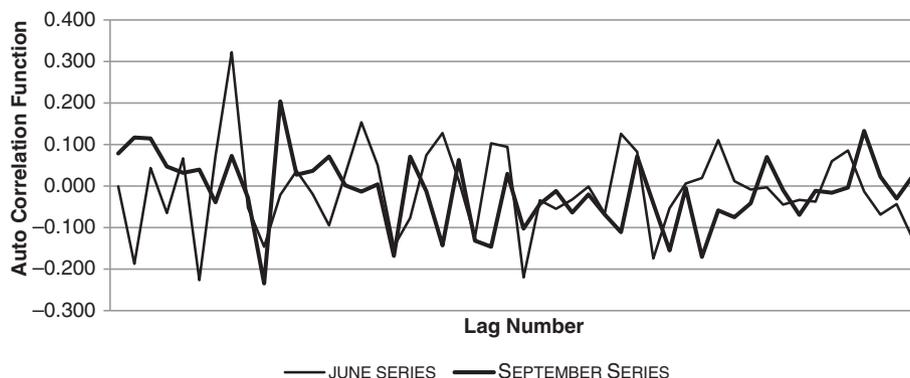


**FIGURE 5** Proposed methodology to assess the predictability of the rainfall time series to improve their forecast using Hurst exponents, similarity factors (SF) and artificial neural networks (ANN)

time series into more than one class with rainfall prediction done separately for each class.

Based on this analysis, the following generalized methodology (Figure 5) can be proposed for checking the predictability of a rainfall time series for a given month.

- The overall Hurst exponent for the times series has to be estimated along with Hurst exponents at different levels (obtained by dividing the time series into equal halves at each subsequent step). If the overall Hurst exponent as well as the Hurst exponents at different levels are



**FIGURE 6** Autocorrelation function for the June and September series

TABLE 2 Hurst exponent value of Cherrapunji rainfall for all months

Dataset	$H$ (over all)	$H_1$	$H_2$	$H_{T1}$	$H_{T2}$	$H_{T3}$
January	0.65	0.36	0.94	0.29	0.66	1.00
February	0.78	0.93	0.92	0.88	0.35	0.96
March	0.75	0.53	0.85	0.35	0.92	1.00
April	0.76	0.63	0.96	0.61	0.68	1.00
May	0.79	0.75	0.95	0.75	0.66	1.00
June	0.65	0.49	0.91	0.56	0.60	1.00
July	0.67	0.55	0.87	0.51	0.60	1.00
August	0.69	0.48	0.87	0.29	0.72	1.00
September	0.81	0.80	0.92	0.74	0.68	1.00
October	0.66	0.37	0.89	0.04	0.69	1.00
November	0.74	0.50	0.89	0.27	0.82	1.00
December	0.72	0.53	0.89	0.47	0.76	1.00

significantly higher than 0.5, the predictability of the time series is higher.

- The rainfall time series is classified to more than one class in order to categorize rainfalls of a similar nature in one category. In the case of the presence of self-similarity in the time series, this may help improve the predictability.
- Instead of using the actual rainfall in the output (for ANN training), the SF can be used. From the predicted SF, the predicted rainfall can be calculated.

### 3.2 | Cherrapunji rainfall time series

The proposed generalized methodology is applied to the real rainfall data from Cherrapunji which represents an area characterized by very high rainfall. The physical framework of the study area affects the Hurst exponent (Rangarajan and Sant, 1997). First the ACFs are estimated for June and September and indicate the absence of short

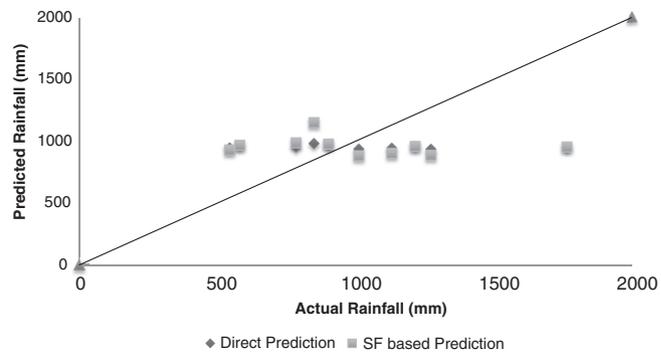


FIGURE 7 Comparison of rainfall prediction for June

term memory in the time series (Figure 6). The overall Hurst exponent and the Hurst exponents at different levels for 64 years of rainfall data for all the 12 months are shown in Table 2. It is seen that, although the overall Hurst exponent is significantly higher than 0.5 for all months except the month of September, for other months not all the Hurst exponents at different levels are greater than 0.5, indicating that the predictability of rainfall is less. An attempt is made to check the predictability for the time series of June and September.

The rainfall time series for June has a Hurst exponent of 0.65 whereas  $H_1$  and  $H_{T1}$  are 0.49 and 0.56 respectively. Hence, the predictability of this time series is expected to be more uncertain. This was investigated through the ANN training. The number of hidden neurons was found to be seven with the logistic function used as the activation function for the optimal architecture. A comparison of the actual and predicted values for the month is shown in Figures 7 and 8. As seen from Figure 7, the ANN has failed in the one lead prediction under both conditions, i.e. with and without the use of the SF. The ANN was also trained with many other alternative architectures and other combinations of

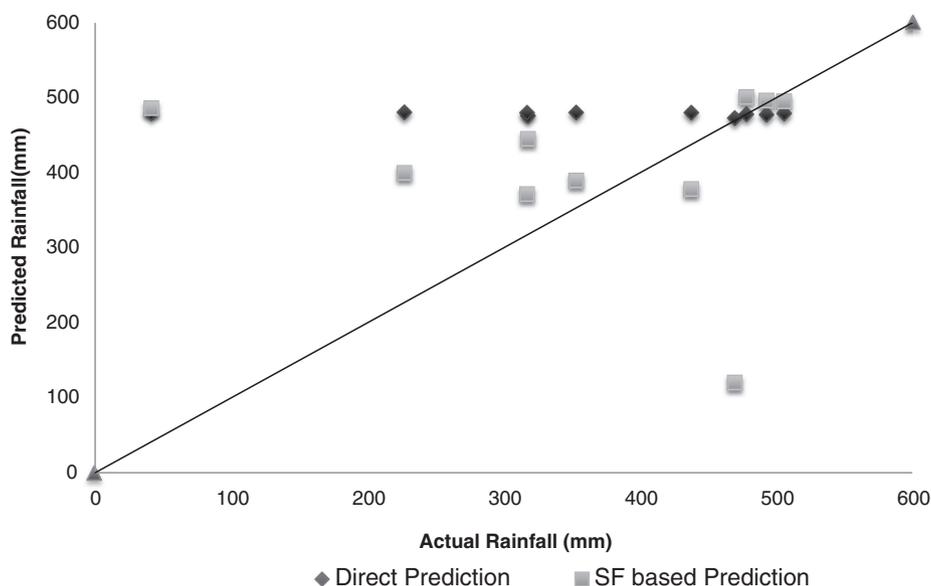


FIGURE 8 Comparison of rainfall prediction for September

training, testing and validation data but it did not improve the prediction.

For September, the Hurst exponent was found to be 0.81. The Hurst exponents at all the levels were also found to be significantly higher than 0.5 indicating the predictability of the time series. ANN training was attempted with an optimal architecture consisting of eight hidden neurons with the sine function as the activation function. The time series was split into two classes using a Kohonen network and each sub-series was trained separately with the SF in the output. Figure 8 shows the comparison of actual output and predicted output for the validation set from the two different ANN trainings. Out of 10 data in the validation set, the rainfall prediction for six data has improved with the inclusion of the SF. Further, as seen from the figure, the SF based prediction follows the pattern of the actual output although differing in the values in some places. It is clearly seen that the rainfall of September is more predictable.

#### 4 | CONCLUSIONS

The prediction of rainfall time series is a typically challenging problem. If the series has a strong autocorrelation function, the prediction should first be attempted with the necessary antecedent rainfall values. For a Hurst exponent greater than 0.5, the fractal behaviour of the time series will show the presence of long term memory and indicate that the series will be persistent. This, however, does not necessarily indicate the predictability of the series. To ensure predictability, the Hurst exponent should be significantly greater than 1 not only for the overall time series but also for the sublevels of the time series. In the absence of this, the fractal self-similarity characteristics of the time series will be less prominent leading to more uncertainty in the prediction. It is also seen that the self-similarity can be represented in the model training in terms of a similarity factor, the prediction of which can be carried out instead of directly predicting the rainfall values. The effect of self-similarity can be further ensured by classifying the time series of a given month into two or more categories and constructing an individual model for training. In this study the Kohonen neural network was used for classifying the time series, but any other appropriate classification algorithm can also be used.

As a future scope of the study, it is recommended that more detailed studies are carried out with data representing the study area featuring arid, semi-arid, monsoon or rainforest, desert, oceanic or terrestrial regions. The fractal behaviour of the rainfall time series can also be ascertained using other indices, such as spectral clustering and deterministic index in addition to estimation of the Hurst exponent.

#### CONFLICTS OF INTEREST

The authors have no conflicts of interest to disclose.

#### ORCID

Saravanan Poomalai  <https://orcid.org/0000-0002-1400-0701>

#### REFERENCES

- Ahmadi A, Karamouz M, Nazif S, Noori N. 2009. Long-lead forecasting of monthly rainfall using large scale climate signals and statistical disaggregation models. In *World Environmental and Water Resources Congress 2009: Great Rivers*. pp. 1–10. [https://doi.org/10.1061/41036\(342\)501](https://doi.org/10.1061/41036(342)501).
- Ajmera, T.K. and Goyal, M.K. (2012) Development of stage–discharge rating curve using model tree and neural networks: an application to Peachtree Creek in Atlanta. *Expert Systems with Applications*, 39, 5702–5710.
- Aksoy, H. and Dahamsheh, A. (2009) Artificial neural network models for forecasting monthly precipitation in Jordan. *Stochastic Environmental Research and Risk Assessment*, 23, 917–931.
- ASCE Task Committee. (2000a) Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of Hydrologic Engineering*, 5, 115–123.
- ASCE Task Committee. (2000b) Artificial neural networks in hydrology. II: Hydrologic applications. *Journal of Hydrologic Engineering*, 5, 124–137.
- Ashok, K., Behera, S.K., Rao, S.A., Weng, H. and Yamagata, T. (2007) El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research*, 1, 112.
- Babel, M.S., Badgujar, G.B. and Shinde, V.R. (2015) Using the mutual information technique to select explanatory variables in artificial neural networks for rainfall forecasting. *Meteorological Applications*, 22, 610–616.
- Box, G.E.P. and Jenkins, G. (1976) *Time Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day Inc.
- Chattopadhyay, S. and Chattopadhyay, G. (2008) Comparative study among different neural net learning algorithms applied to rainfall time series. *Meteorological Applications*, 15, 273–280.
- Collins, W. and Tissot, P. (2015) An artificial neural network model to predict thunderstorms within 400 km<sup>2</sup> South Texas domains. *Meteorological Applications*, 22, 650–665.
- Dahamsheh, A. and Aksoy, H. (2009) Artificial neural network models for forecasting intermittent monthly precipitation in arid regions. *Meteorological Applications*, 16, 325–337.
- Feng, L. and Zhou, J. (2013) Trend predictions in water resources using rescaled range (R/S) analysis. *Environment and Earth Science*, 68, 2359–2363.
- Gheiby, A., Sen, P.N., Puranik, D.M. and Karekar, R.N. (2003) Thunderstorm identification from AMSU-B data using an artificial neural network. *Meteorological Applications*, 10, 329–336.
- Goyal, M.K. (2014) Monthly rainfall prediction using wavelet regression and neural network: an analysis of 1901–2002 data, Assam, India. *Theoretical and Applied Climatology*, 118, 25–34.
- Haykin, S. (1999) *Neural Network: A Comprehensive Foundation*, 2nd edition. PTR Upper Saddle River, NJ: Prentice Hall.
- Houston, J. (2006) Variability of precipitation in the Atacama Desert: its causes and hydrological impact. *International Journal of Climatology*, 26, 2181–2198.
- Hung, N.Q., Babel, M.S., Weesakul, S. and Tripathi, N.K. (2009) An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology and Earth System Sciences*, 13, 1413–1425.
- Hurst, H.E. (1951) Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770–808.
- Kalauzi, A., Cukic, M., Millán, H., Bonafoni, S. and Biondi, R. (2009) Comparison of fractal dimension oscillations and trends of rainfall data from Pastaza Province, Ecuador and Veneto, Italy. *Atmospheric Research*, 93, 673–679.
- Kar, S.C., Acharya, N., Mohanty, U.C. and Kulkarni, M.A. (2012) Skill of monthly rainfall forecasts over India using multi-model ensemble schemes. *International Journal of Climatology*, 32, 1271–1286.
- Kendzioriski, C.M., Bassingthwaight, J.B. and Tonellato, P.J. (1999) Evaluating maximum likelihood estimation methods to determine the Hurst coefficient. *Physica A*, 273, 439–451.
- Khalili, N., Khodashenas, S.R., Davary, K., Baygi, M.M. and Karimaldini, F. (2016) Prediction of rainfall using artificial neural networks for synoptic station of Mashhad: a case study. *Arabian Journal of Geosciences*, 9, 624.
- Kumarasiri, A.D. and Sonnadara, U.J. (2008) Performance of an artificial neural network on forecasting the daily occurrence and annual depth of rainfall at a tropical site. *Hydrological Processes*, 22, 3535–3542.

- Li, M., Xia, J. and Meng, D.J. (2012) DFA based predictability indices analysis of climatic dynamics in Beijing area, China. *Advanced Materials Research*, 382, 60–64.
- Mahmoodi, M. and Naderi, A. (2016) Applicability of artificial neural network and nonlinear regression to predict mechanical properties of equal channel angular rolled Al5083 sheets. *Latin American Journal of Solids and Structures*, 13, 1515–1525.
- Modarres, R., Ghadami, M., Naderi, S. and Naderi, M. (2018) Future extreme rainfall change projections in the north of Iran. *Meteorological Applications*, 25, 40–48.
- Nourani, V., Komasi, M. and Alami, M.T. (2011) Hybrid wavelet–genetic programming approach to optimize ANN modeling of rainfall-runoff process. *Journal of Hydrologic Engineering*, 17, 724–741.
- Olsson, J., Uvo, C.B., Jinno, K., Kawamura, A., Nishiyama, K., Koreeda, N., Nakashima, T. and Morita, O. (2004) Neural networks for rainfall forecasting by atmospheric downscaling. *Journal of Hydrologic Engineering*, 9, 1–2.
- Peyghami, M.R. and Khanduzi, R. (2012) Predictability and forecasting automotive price based on a hybrid train algorithm of MLP neural network. *Neural Computing and Applications*, 21, 125–132.
- Rangarajan, G. and Sant, D.A. (1997) A climate predictability index and its applications. *Geophysical Research Letters*, 24, 1239–1242.
- Rangarajan, G. and Sant, D.A. (2004) Fractal dimensional analysis of Indian climatic dynamics. *Chaos, Solitons & Fractals*, 19, 285–291.
- Rao, A.R. and Bhattacharya, D. (1999) Comparison of Hurst exponent estimates in hydrometeorological time series. *Journal of Hydrologic Engineering*, 4, 225–231.
- Rehman, S. and El-Gebeily, M. (2009) A study of Saudi climatic parameters using climatic predictability indices. *Chaos, Solitons & Fractals*, 41, 105–1069.
- Rehman, S. and Siddiqi, A.H. (2009) Wavelet based Hurst exponent and fractal dimensional analysis of Saudi climatic dynamics. *Chaos, Solitons & Fractals*, 40, 1081–1090.
- Salomao, L.R., Campanha, J.R. and Gupta, H.M. (2009) Rescaled range analysis of pluviometric records in Sao Paulo state, Brazil. *Theoretical and Applied Climatology*, 95, 83–89.
- Setty, V.A. and Sharma, A.S. (2015) Characterizing detrended fluctuation analysis of multifractional Brownian motion. *Physica A Statistical Mechanics and its Applications*, 419, 698–706.
- Sivapragasam, C., Vanitha, S., Muttill, N., Suganya, K., Suji, S., Selvi, M.T., Selvi, R. and Sudha, S.J. (2014) Monthly flow forecast for Mississippi River basin using artificial neural networks. *Neural Computing and Applications*, 24, 1785–1793.
- Srivastava, S. and Tripathi, K.C. (2012) Artificial neural network and non-linear regression: a comparative study. *International Journal of Scientific and Research Publications*, 2, 740–744.
- Tapiador, F.J., Kidd, C., Hsu, K. and Marzano, F. (2004) Neural networks in satellite rainfall estimation. *Meteorological Applications*, 11, 83–91.
- Taqqu, M.S., Teverovsky, V. and Willinger, W. (1995) Estimators for long-range dependence: an empirical study. *Fractals*, 3, 785–798.
- Tatli, H. (2014) Statistical complexity in daily precipitation of NCEP/NCAR reanalysis over the Mediterranean Basin. *International Journal of Climatology*, 34, 155–161.
- Tatli, H. (2015) Detecting persistence of meteorological drought via the Hurst exponent. *Meteorological Applications*, 22, 763–769.
- Tatli, H. and Dalfes, H.N. (2016) Defining Holdridge's life zones over Turkey. *International Journal of Climatology*, 36, 3864–3872.
- Valipour, M. (2016) Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. *Meteorological Applications*, 23, 91–100.
- Vinoth, B., Rajarathian, A. and Manju Bargavi, S.K. (2016) Nonlinear regression and artificial neural network based model for forecasting Paddy (*Oryza sativa*) production in Tamil Nadu. *IOSR-JMCA*, 3, 01–06.
- Wu, C.L., Chau, K.W. and Fan, C. (2010) Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. *Journal of Hydrology*, 389, 146–167.
- Yildirim, K., Ogut, H. and Ulcay, Y. (2017) Comparing the prediction capabilities of artificial neural network (ANN) and nonlinear regression models in pet-Poy yarn characteristics and optimization of yarn production conditions. *Journal of Engineered Fibers and Fabrics*, 12. <https://doi.org/10.1177/155892501701200302>.

**How to cite this article:** Chandrasekaran S, Poomalai S, Saminathan B, Suthanthiravel S, Sundaram K, Abdul Hakkim FF. An investigation on the relationship between the Hurst exponent and the predictability of a rainfall time series. *Meteorol Appl.* 2019;26:511–519. <https://doi.org/10.1002/met.1784>