Chapter 15
Comparison of Probability Distributions for Extreme Value Analysis and Predicting Monthly Rainfall Pattern Using Bayesian Regularized ANN

N. Vivekanandan, Shivam Singh, and Manish Kumar Goyal

Abstract The estimation of rainfall for a given return period is of utmost importance for the planning and design of minor and major hydraulic structures. This can be accomplished using an extreme value analysis (EVA) of rainfall, which involves fitting a series of annual 1-day maximum rainfall data to probability distributions including the 2-parameter normal, 2-parameter log normal, Pearson type 3, log Pearson type 3, extreme value type 1 (EV1), and generalized extreme value (GEV). The method of moments (MoM), maximum likelihood method (MLM), and L-moments (LMO) are used to determine the distributional parameters depending on the intended applications and the variable under consideration. The six probability distributions used in the EVA of rainfall for the Afzalpur, Aland, and Kalaburagi sites are adequately fitted when measured quantitatively by the goodness-of-fit and diagnostic tests (chi-square and Kolmogorov-Smirnov) and qualitatively by the fitted curves of the estimated rainfall. According to the study's findings, the GEV (LMO) is the most suited among the six distributions tested in EVA for estimating rainfall for Afzalpur and Kalaburagi, while the EV1 (MLM) is more suited for Aland. An artificial neural network with Bayesian regularization has been implemented to model and predict monthly rainfall patterns from all three sites. The model performance analysis shows significant correlation of approximately 0.95 for the training dataset and 0.40 for validation.

Keywords Chi-square · D-index · Extreme value type 1 · Generalized extreme value · Kolmogorov-Smirnov · L-moments · Maximum likelihood method · Rainfall · Artificial neural network · Bayesian regularization

N. Vivekanandan (✉)
Central Water and Power Research Station, Pune, Maharashtra, India
e-mail: vivek.n@cwprs.gov.in

S. Singh · M. K. Goyal
Department of Civil Engineering, Indian Institute of Technology Indore, Indore, Madhya Pradesh, India

15.1   Introduction

Assessment of extreme rainfall for a desired return period is of utmost importance for the planning and design of minor and major hydraulic structures, viz., dams, bridges, barrages, and storm water drainage systems. Such information can be applied to the planning and design of water resources projects linked to reservoir design, river bank protection works, soil and water conservation, etc., as well as to the prevention of floods and droughts (CWC 2010). In the post-commissioning stage, where it is necessary to analyze the risk of hydraulic structures failing, extreme rainfall occurrences are crucial (Baratti et al. 2012). Extreme rainfall with a desirable return duration is used, depending on the design life of the structure. Extreme value analysis (EVA), which involves fitting a probability distribution to the series of annual 1-day maximum rainfall (AMR) data, can be used to achieve this (Abida and Ellouze 2008). The robust forecasts of precipitation patterns at a significant lead time can be of great importance in managing water resources as well as mitigating the risk associated with prolonged and flash flooding (Das et al. 2020; Floods 2019; Goswami et al. 2018). Numerical weather prediction models are conventionally used to forecast these extremes by mathematically modeling the governing physical processes (Bauer et al. 2015; Dueben et al. 2021). Being inspired by the learning capability of neural networks, an advancement in machine learning from nonlinear complex data (Pisner and Schnyer 2020; Scher and Messori 2018), we have incorporated artificial neural networks (ANNs) to model and predict the monthly rainfall of all three considered stations with Bayesian regularization to overcome the issue of overfitting in the data.

15.2   Literature Review

For analyzing the rainfall data, a variety of probability distributions from the normal, gamma, and extreme value families of distributions will be provided (Arvind et al. 2017; Esberto, 2018; Sasireka et al. 2019). The distributions that are most frequently applied and employed in EVA are the two-parameter normal (N2) and log normal (LN2), Pearson type 3 (P3), log Pearson type 3 (LP3), extreme value type 1 (EV1), and generalized extreme value (GEV) distributions. From this, it is clear that the N2 and LN2 belong to the family of normal distributions, the P3 and LP3 to the family of gamma distributions, and the EV1 and GEV to the family of extreme value distributions (Bhuyan et al. 2010; Mujere 2011; Olumide et al. 2013; Haberlandt and Radtke 2014; Sharma and Sharma 2019; Singh et al. n.d.; Tank et al. 2021). The parameter estimation algorithms method of moments (MoM), maximum likelihood method (MLM), and L-moments (LMO) are used to determine the distribution's parameters based on the planned applications and the variate under consideration (Acar et al. 2008; Malekinezhad et al. 2011; Vivekanandan 2020).

Using Gumbel (also known as EV1), LN2, and LP3, AlHassoun (2011) conducted a study on establishing an empirical formula to estimate rainfall intensity in the Riyadh region. He came to the conclusion that among the three distributions examined for the assessment of rainfall intensity, the LP3 provides greater accuracy. In order to calculate the extreme rainfall depths at several rain gauge stations in southeast United Kingdom, Esteves (2013) used the EV1 distribution. For the purpose of creating intensity-duration-frequency curves for seven divisions in Bangladesh, Rasel and Hossain (2015) used the EV1 distribution. In the Bamenda Mountain region of Cameroon, Afungang and Bateira (2016) used the EV1 distribution to estimate the maximum quantity of rainfall for various times. Eight probability distributions were used by Baghel et al. (2019) to analyze the frequency of daily maximum rainfall data in the Udaipur district. For the Anakapalli, Atchutapuram, Kasimkota, and Parvada sites, Vivekanandan and Srishailam (2020) examined the MoM and MLM estimators of the EV1, LN2, and LP3 distributions utilized in the EVA of rainfall. The question of which distribution model best fits a given collection of data comes frequently when multiple probability distributions are used in the EVA of rainfall. Both quantitative and qualitative analyses may be able to provide an answer, and the conclusions are quantifiable and trustworthy. The effectiveness of fitting the selected probability distributions is assessed quantitatively using chi-square (χ^2) and Kolmogorov-Smirnov (KS), and diagnostic (viz., D-index) tests, as well as qualitatively using fitted curves for the estimated extreme rainfall. The methods used in the EVA of rainfall and the evaluation of EVA results using GoF and diagnostic tests are succinctly discussed with an example, and the outcomes of the study are reported in the paper.

There have been numerous attempts to predict rainfall using different physical, empirical, and physio-empirical models. In many parts of the world, these models have been used to predict rainfall on an annual, seasonal, monthly, and daily scale (Bauer et al. 2015; Goyal et al. 2018; Nardi et al. 2018). Physical models are created taking into account the synoptic climate and the physical processes by which various climatic variables interact to produce rainfall in a location. In order to anticipate rainfall in physical models, numerical models are created to describe ingrained physical processes (Goyal and Ojha 2010). These models require data and knowledge on numerous land-ocean-atmospheric variables and are typically quite complicated. In spite of this, physical models frequently fail to provide accurate rainfall predictions, because they frequently use crude approximations of complicated physical events (Coles 2001; Su et al. 2012). Physical models have been replaced with empirical models based on statistical techniques. These models are typically created with a specific climate variable in mind, such as rainfall prediction, based on the connection defined by statistics between that variable and its antecedent variables. Statistical models are always region-specific; thus, the model created using the statistical relationship between local rainfall and other atmospheric factors is only applicable for the region (Goyal et al. 2012; Katz 2013). These models are also referred to as empirical models for this reason. In most instances, empirical models are more accurate than physical models and are easier to construct and apply. The primary flaw of empirical models is their total reliance on the historical data utilized

to construct them. As a result, they are unable to replicate rainfall in an unknown environment. For instance, empirical models cannot accurately predict the abrupt variations in rainfall that the region has experienced recently and is expected to experience in the future since the models were not constructed with such a wide range of data (Hewitson et al. 2014). In this circumstance, physical models are used as a forecasting option. The development of physical-empirical models, in which empirical models are based on the variables physically responsible for the region's climate, has recently received attention in an effort to address the limitations of both types of modeling approaches (Vergés et al. 2016).

Regression analysis, including linear and higher-order polynomial (nonlinear) regressions, is typically used to create statistical models (Chen and Zhang 2022; Su et al. 2012). However, the relationship between rainfall and the climatic variables that cause it is frequently enigmatic and cannot be captured using commonly used statistical approaches. Complex regression analysis is frequently created using machine learning (ML) methods (Hinge et al. 2018; Sharma and Goyal 2020). Climate extremes in changing climate are getting worse, and robust prediction of these extremes is one of the most essential aspects in managing water resources and disaster management (Poonia et al. 2021a, b). As a result, the focus of meteorologists has been shifted to the development of machine learning (ML)-based physical-empirical forecasting models in recent years. ANNs are one of the numerous possibilities explored and presented in the literature for prediction. They are a flexible and rich concept that may be used to tackle difficulties with clustering, time series, and function approximation in addition to classification tasks (Rautela et al. 2022). The adaptability of ANNs encouraged researchers to look into their suitability for classification and regression problems (Vu et al. 2019). Studies reveal that ANNs and other artificial intelligence techniques are capable of outperforming conventional statistical techniques. The Bayesian assessment of an ANN for prediction indicates that the performance of the neural architecture is critical to its configuration, because it strongly influences the estimate efficacy of the framework (Burden and Winkler 2008). To avoid over fitting, however, in this study we concentrate on Bayesian regularization (BR) of the ANN employed for predicting monthly rainfall (Okut 2016). The huge amount of complex nonlinear meteorological data from various sources necessitate great care to avoid over fitting ANN algorithm (Ye et al. 2021). We have implemented BR-ANN to explore the nonlinear relationships associated with monthly rainfall at three stations and predict it for the next.

15.3 Methodology

The cumulative distribution function (CDF) and quantile estimator of six distributions (viz., N2, LN2, P3, LP3, EV1, and GEV) adopted in EVA is presented in Table 15.1. The empirical equations involved in determining the MoM, MLM, and

Table 15.1 CDF with quantile estimator six probability distributions

Distribution	CDF (F(x) or F(y))	Quantile estimator x(T)
N2 ($\mu(x)$, $\sigma(x)$)	$F(x) = \varphi\left(\frac{x - \mu(x)}{\sigma(x)}\right)$	$x(T) = \mu(x) + K(T)\sigma(x)$
LN2 ($\mu(y)$, $\sigma(y)$)	$F(y) = \varphi\left(\frac{y - \mu(y)}{\sigma(y)}\right)$ wherein y = ln(x)	$x(T) = \exp(\mu(y) + K(T)\sigma(y))$
P3 (ξ, α, β)	$F(x) = \begin{cases} G\left(\beta, \frac{x - \xi}{\alpha}\right) & , \alpha > 0 \\ 1 - G\left(\beta, \frac{x - \xi}{\alpha}\right) & , \alpha < 0 \end{cases}$	No explicit expression of the quantile function is available
LP3 (ξ, α, β)	$F(x) = \begin{cases} G\left(\beta, \frac{\ln(x) - \xi}{\alpha}\right) & , \alpha > 0 \\ 1 - G\left(\beta, \frac{\ln(x) - \xi}{\alpha}\right) & , \alpha < 0 \end{cases}$	No explicit expression of the quantile function is available
EV1 (ξ, α)	$F(x) = e^{-e^{-\left(\frac{x - \xi}{\alpha}\right)}},\ \alpha > 0$	$x(T) = \xi + \alpha[-\ln(-\ln(F(x)))]$
GEV (ξ, α, β)	$F(x) = e^{-\left(1 - \frac{\beta(x - \xi)}{\alpha}\right)1/\beta},\ \alpha > 0, \beta > 0$	$x(T) = \xi + \frac{\alpha\left[1 - (-\ln(F(x)))^{\beta}\right]}{\beta}$
Wherein		
ξ	:	Location parameter
α	:	Scale parameter
β	:	Shape parameter
T	:	Return period (in year)
F(x)	:	CDF of a variable x (i.e., AMR)
K(T)	:	Frequency factor of a return period (T) corresponding to the coefficient of skewness (CS), say CS = 0.0 for N2 and LN2
x(T)	:	Estimated extreme rainfall for a return period (T)
G(...)	:	Incomplete gamma integral
ϕ(...)	:	CDF of standard normal distribution
$\mu(x)$	:	Average of observed data (x)
$\sigma(x)$	:	Standard deviation of observed data (x)
$\mu(y)$	:	Average of log transformed (y = ln(x)) series of observed data
$\sigma(y)$	:	Standard deviation of log transformed series of observed data
A relation between F(x) and T is given by F(x) = 1–1/T		

LMO estimators of the distributions are presented in Table 15.2 (Rao and Hamed 2000).

15.3.1 MoM of P3 Distribution

The MoM estimators of P3 distribution (Bobee and Askhar 1991) can be determined by solving the system of equations, which is given as below:

Table 15.2 Determination of parameters of six distributions using MoM, MLM, and LMO

Distribution	Parameters of the distribution		
	MoM	MLM	LMO
N2	$\mu(x) = \frac{1}{N}\sum_{i=1}^{N}x(i) - \frac{1}{2N}\sum_{i=1}^{N}x(i)^2 + \frac{1}{2}\left(\frac{1}{N}\sum_{i=1}^{N}x(i)\right)^2$ $\sigma(x) = \left[\left(\frac{1}{N}\sum_{i=1}^{N}x(i)^2\right) - 2\left(\frac{1}{N}\sum_{i=1}^{N}x(i)\right)^2\right]^{1/2}$	$\mu(x) = \frac{1}{N}\sum_{i=1}^{N}x(i)$ $\sigma(x) = \left(\frac{1}{N}\sum_{i=1}^{N}(x(i)-\mu(x))^2\right)^{1/2}$	$\lambda_1 = \mu(x)$ $\lambda_2 = \sigma(x)/\sqrt{N}$
LN2	$\mu(y) = \frac{1}{N}\sum_{i=1}^{N}y(i) - \frac{1}{2N}\sum_{i=1}^{N}y(i)^2 + \frac{1}{2}\left(\frac{1}{N}\sum_{i=1}^{N}y(i)\right)^2$ $\sigma(y) = \left[\frac{1}{N}\sum_{i=1}^{N}y(i)^2 - 2\left(\frac{1}{N}\sum_{i=1}^{N}y(i)\right)^2\right]^{1/2}$	$\mu(y) = \frac{1}{N}\sum_{i=1}^{N}y(i)$ where $y(i)= \ln(x(i))$ $\sigma(y) = \left(\frac{1}{N}\sum_{i=1}^{N}(y(i)-\mu(y))^2\right)^{1/2}$	$\lambda_1 = \mu(y)$ $\lambda_2 = \sigma(y)/\sqrt{N}$
P3	Can be determined by solving (Eqs. 15.1 and 15.2)	Can be determined by solving Eq. (15.3)	No explicit expression is available
LP3	Can be determined by solving (Eqs. 15.1 and 15.2) by replacing x(i) with ln(x(i))	Can be determined by solving Eq. (15.3) by replacing x(i) with ln(x(i))	No explicit expression is available
EV1	$\xi = \mu(x) - (0.5772157)\alpha$ $\alpha = \left(\frac{\sqrt{6}}{\pi}\right)\sigma(x)$	$\xi = -\alpha\ln\left[\sum_{i=1}^{N}\exp(-x(i)/\alpha)/N\right]$ $\alpha = \mu(x) - \left[\sum_{i=1}^{N}x(i)\exp(-x(i)/\alpha)/\alpha \Big/ \sum_{i=1}^{N}\exp(-x(i)/\alpha)\right]$	$\xi = \lambda_1 - (0.5772157)\alpha$ $\alpha = \frac{\lambda_2}{\ln(2)}$
GEV	$\mu(x) = \xi + \frac{\alpha(1-\Gamma(1+\beta))}{\beta}$ $\sigma(x) = \frac{\alpha}{\beta}\left(\Gamma(1+2\beta) - \Gamma(1+\beta)^2\right)^{1/2}$ $\quad$ $CS = (sign\,\beta)\frac{\Gamma(1+3\beta)+3\Gamma(1+\beta)\Gamma(1+2\beta)-2\Gamma^3(1+\beta)}{\{\Gamma(1+\beta)-\Gamma^2(1+\beta)\}^{1/2}}$	Can be determined by using the modified Newton-Raphson algorithm	$z = (2/(3+\tau_3)) - (\ln(2)/\ln(3))$; $\tau_3 = (2(1 - 3^{-\beta})/(1 - 2^{-\beta}) - 3$; $\alpha = (\lambda_2\beta)/(1 - 2^{-\beta})\Gamma(1+\beta)$; $\xi = \lambda_1 + (\alpha(\Gamma(1+\beta) - 1)/\beta)$; $\beta = 7.817740z + 2.930462z^2 + 13.641492z^3 + 17.206675z^4$

$$M'_r = \frac{\exp(r\xi)}{(1-r\alpha)^\beta} \tag{15.1}$$

where in 1-r$\alpha > 0$, $r = 1, 2, 3$

$$\mu(x) = M'_1, \quad \sigma(x)^2 = M'_2 - \left(M'_1\right)^2 \text{and} \quad \gamma(x) = M'_3 - 3M'_2M'_1 + 2\left(M'_1\right)^3 \tag{15.2}$$

where M'_r is the r$^{\text{th}}$ moment of x about the origin and $\mu(x)$ is the coefficient of skewness of the observed data.

15.3.2   MLM of P3 Distribution

The MLM estimators of P3 distribution (Bobee and Askhar 1991) can be determined by solving the following system of equations:

$$\left.\begin{array}{l} \sum_{i=1}^{N} (x(i) - \xi) = N\alpha\beta \\[2mm] \sum_{i=1}^{N} \frac{1}{\alpha}(x(i) - \xi) = N\psi(\beta) \\[2mm] \sum_{i=1}^{N} \frac{1}{(x(i) - \xi)} = \frac{N}{\alpha(\beta - 1)} \end{array}\right\} \tag{15.3}$$

Here, $\psi(\beta)$ is the digamma function of estimator of the scale parameter (β).

   In Table 15.2, λ_1, λ_2, and λ_3 are the first, second, and third, respectively, LMOs (Hosking 1990) that can be determined in Eq. (15.4), which is given as below:

$$\lambda_1 = b_0, \quad \lambda_2 = 2b_1 - b_0, \quad \lambda_3 = 6b_2 - 6b_1 + b_0 \text{ and } \tau_3 = \lambda_3/\lambda_2 \tag{15.4}$$

wherein λ_{r+1} is the r + 1th LMO (Hosking and Wallis 1993), which is defined by:

$$\lambda_{r+1} = \sum_{k=0}^{r} \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!} b_k \tag{15.5}$$

wherein b_k is an unbiased estimator (Saf 2009; Gubareva and Gartsman 2010) and given by:

$$b_k = N^{-1} \sum_{i=k+1}^{N} \frac{(i-1)(i-2).....(i-k)}{(N-1)(N-2).....(N-k)} x(i) \tag{15.6}$$

where x(i) is the observed data of i$^{\text{th}}$ sample and N is the total number of samples.

15.3.3 Goodness-of-Fit Tests

GoF tests are essential for checking the adequacy of probability distributions to the AMR series in rainfall estimation. Out of a number GoF tests available, the widely accepted GoF tests are χ^2 and KS (Zhang 2002), which are used in the study.

χ^2 test statistic is defined by:

$$\chi^2 = \sum_{j=1}^{NC} \frac{\left(O_j(x) - E_j(x)\right)^2}{E_j(x)} \tag{15.7}$$

where $O_j(x)$ is the observed frequency value of x for j^{th} class, $E_j(x)$ is the expected frequency value of x for j^{th} class, and NC is the number of frequency classes (Charles Annis 2009). The rejection region of χ^2 statistic at the desired significance level (η) is given by $\chi^2_C \geq \chi^2_{1-\eta,NC-m-1}$. Here, m denotes the number of parameters of the distribution, and χ^2_C is the computed value of χ^2 statistic by the probability distribution.

KS test statistic is defined by:

$$KS = \underset{i=1}{\overset{N}{Max}} |F_e(x(i)) - F_c(x(i))| \tag{15.8}$$

where x(i) is the observed data for i^{th} sample, $F_e(x(i)) = r/(N+1)$ is the empirical CDF of x(i) of ith sample, "r" is the rank assigned to sample values arranged in ascending order (i.e., x(1) < x(2) < .....x(N)), and $F_c(x(i))$ is the computed CDF of x(i) of ith sample.

Test criteria: If the computed values of GoF tests statistic given by the distribution are less than that of the theoretical values at the desired level of significance, then the distribution is considered to be acceptable for EVA at that level.

15.3.4 Diagnostic Test

Sometimes the GoF test results would not offer a conclusive inference, thereby posing a bottleneck for the user in selecting the suitable distribution for the application. In such cases, a diagnostic test in adoption to GoF is applied for making inference. The selection of the most suitable distribution is performed through the D-index test (United States Water Resources Council (USWRC) 1981), which is defined as:

$$\mathrm{D-index} = (1/\mu(x)) \sum_{i=1}^{6} |x(i) - x(i)^*| \qquad (15.9)$$

Here, $x(i)$ ($i = 1$ to 6) and $x(i)^*$ are the six highest observed and the corresponding estimated values of ith sample. The probability distribution having the least D-index is considered as a better suited for rainfall estimation.

15.3.5 Bayesian Regularized Artificial Neural Network

BR-ANN (Burden and Winkler 2008; Okut 2016) is a more robust variant of ANNs than the standard ANN (Sasireka et al. 2019). This robustness of BR-ANN is obtained through the BR of the ANN parameters. A popular error function (E_D) of ANN is as follows:

$$E_D(D|w, M) = \sum_{i=1}^{n} \left(t_i - \widehat{t_i}\right)^2 \qquad (15.10)$$

where w denotes weight, M denotes ANN structure, n denotes size of training data, t_i is the ith target output, and $\widehat{t_i}$ is the ith model (BR-ANN) output.

Regularizing ANN with the Bayesian technique aids in optimizing the ANN parameter by utilizing prior ANN parameter values. In order to do this, an additional term (E_w) is added to the BR-ANN's target function as follows:

$$E_D(D|w, M) = \sum_{i=1}^{n} \left(t_i - \widehat{t_i}\right)^2 + E_w \qquad (15.11)$$

In order to improve generalization and gradual conversion, E_w is employed to compensate the unrealistic weights. The function is minimized using an optimization technique based on gradients:

$$F = \beta \; E_D(D|w, M) = \sum_{i=1}^{n} \left(t_i - \widehat{t_i}\right)^2 + \alpha \; E_w(w|M) \qquad (15.12)$$

where the hyper-parameters that need to be optimized are represented by α and β and $E_w(w|M)$ represent sum of square of the ANN architecture. BR-ANN is an efficient predictive model, because it can uncover theoretically complex input-output relationships.

15.4   Application

In this paper, a study on the comparison of six probability distributions (viz., N2, LN2, P3, LP3, EV1, and GEV) adopted in the EVA of rainfall is carried out. MoM, MLM, and LMO determine the distribution's characteristics, which are also employed in the estimation of rainfall. The daily rainfall data recorded at the Afzalpur, Kalaburagi, and Aland locations from 1970 to 2018 and 1970 to 2017, respectively, are used. From the daily rainfall data, the AMR series is taken out and used for EVA CWPRS (2021) do not have data for the intermittent period, according to a review of the daily rainfall statistics. Additionally, it is highlighted that the observed rainfall at Kalaburagi, which was 1.5 mm in 1993 and 15.6 mm in 2015, is inconsistent and was not taken into account while analyzing the data. However, the data for the missing years are ignored and not taken into account in EVA because of the significance of the hydrological extremes. The AMR descriptive data are provided in Table 15.3. The index map shown in Fig. 15.1 shows the locations of the rain gauge stations taken into consideration for the investigation.

From the descriptive statistics (Table 15.3), it is noted that the higher order moments (CS and CK) of the AMR series behave differently for Afzalpur when compared to the values for Aland and Kalaburagi. The CV of the AMR series of Afzalpur, Aland, and Kalaburagi varies between about 27% and 67%, as shown in Table 15.3.

Table 15.3 Descriptive statistics of the observed AMR

Site	Average (mm)	SD (mm)	CS	CK	Minimum (mm)	Maximum (mm)
Afzalpur	79.1	52.6	5.061	30.614	37.1	400.3
Aland	83.9	35.6	0.935	1.368	30.1	190.0
Kalaburagi	81.7	22.3	0.367	−0.474	41.6	129.0

SD standard deviation, *CS* coefficient of skewness, *CK* coefficient of kurtosis
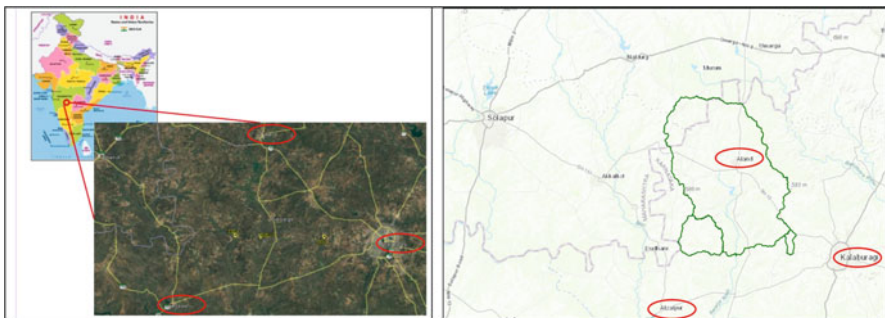

Fig. 15.1   Index map of the study area with locations of rain gauge stations

15.5   Results and Discussion

By applying the procedures as described above, a computer code was developed and used in the EVA of rainfall. The code computes the (1) parameters of N2, LN2, P3, LP3, EV1, and GEV (using MoM, MLM, and LMO) distributions; (2) extreme rainfall estimates for different return periods; and (3) GoF test statistic and D-index values.

15.5.1   Estimation of Extreme Rainfall

The MoM, MLM and LMO were used to determine the characterisitcs of N2, LN2, P3,LP3, EV1 and GEV distributions adopted in EVA, wherever applicable. These parameters are used in the following study. Tables 15.4, 15.5, and 15.6 provide estimates of the 1-day maximum rainfall for Afzalpur, Aland, and Kalaburagi for various return periods. The EVA and GoF test results of P3 (LMO) and LP3 (LMO) are not shown in Tables 15.4, 15.5, 15.6, and 15.7 due to the absence of LMO in P3 and LP3 distributions. From the EVA results, it can be seen that, when compared to the values of other distributions for the return periods ranging from 20 to 1000 years, the LP3 (MLM) offered higher estimates for Afzalpur and Aland and the EV1 (MLM) for Kalaburagi.

15.5.2   Analysis of Results Based on GoF Tests

Six distributions were used to compute the GoF test values for the AMR series of Afzalpur, Aland, and Kalaburagi. The results are shown in Table 15.7. In the current study, the number of frequency classes (NC) is taken into account to be six, and as a result, the degree of freedom (NC-m-1) is taken into account to be two for distributions with three parameters (m), namely, P3, LP3, and GEV, and three for distributions with two parameters (m), namely, N2, LN2, and EV1, when computing the two statistic values. According to the degree of freedom, the theoretical values at the 5% level of significance are observed to be 5.99 for P3, LP3, and GEV and 7.815 for N2, LN2, and EV1. Likewise, the theoretical values of the KS statistic at 5% level of significance with reference to the number of samples considered in EVA are observed as 0.196 for Afzalpur, 0.203 for Aland, and 0.200 for Kalaburagi. From the GoF test results, some of the observations drawn from the study were summarized and presented below:

- χ^2 test results didn't support the use of MoM, MLM, and LMO estimators of five distributions (viz., LN2, P3, LP3, EV1, and GEV) for the EVA of rainfall for Afzalpur and Aland.

Table 15.4 Estimated 1-day maximum rainfall for different return periods by six probability distributions for Afzalpur

Distribution	Method	Estimated 1-day maximum rainfall (mm) for a return period (in year)									
		2	5	10	20	25	50	100	200	500	1000
N2	MoM	79.1	123.3	146.4	165.5	171.1	187.0	201.4	214.5	230.4	241.5
	MLM	79.1	122.8	145.7	164.6	170.1	185.9	200.1	213.1	228.8	239.8
	LMO	79.1	107.5	122.3	134.6	138.1	148.4	157.6	166.0	176.2	183.3
LN2	MoM	65.8	109.5	143.0	178.1	189.9	228.1	269.1	312.9	375.7	427.2
	MLM	71.4	99.6	118.5	136.7	142.6	160.7	178.9	197.4	222.4	241.9
	LMO	71.4	97.2	114.1	130.3	135.4	151.3	167.1	183.1	204.4	218.4
P3	MoM	59.3	81.6	120.2	171.7	190.3	252.5	320.0	391.3	489.8	566.7
	MLM	60.8	83.7	123.2	176.0	195.0	258.8	328.0	401.1	502.0	580.9
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LP3	MoM	64.3	93.2	121.2	156.6	169.9	218.2	279.5	357.2	492.9	627.8
	MLM	80.2	111.0	144.1	188.2	205.3	269.5	354.6	467.6	675.4	893.1
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
EV1	MoM	70.4	116.9	147.7	177.2	186.6	215.4	244.0	272.6	310.2	338.7
	MLM	72.2	98.8	116.5	133.4	138.8	155.3	171.7	188.1	209.7	226.0
	LMO	73.3	104.4	125.0	144.8	151.1	170.4	189.6	208.7	233.9	252.9
GEV	MoM	66.0	102.9	133.5	168.5	180.9	223.9	274.3	333.7	429.0	516.4
	MLM	68.1	95.2	117.6	143.2	152.3	183.5	220.1	263.1	331.9	394.8
	LMO	66.4	93.8	119.3	151.2	163.2	207.1	263.2	335.3	463.0	592.0

Table 15.5 Estimated 1-day maximum rainfall for different return periods by six probability distributions for Aland

Distribution	Method	Estimated 1-day maximum rainfall (mm) for a return period (in year)									
		2	5	10	20	25	50	100	200	500	1000
N2	MoM	83.9	113.8	129.5	142.4	146.2	157.0	166.7	175.5	186.3	193.8
	MLM	83.9	113.5	129.0	141.8	145.5	156.1	165.7	174.5	185.1	192.6
	LMO	83.9	113.0	128.2	140.8	144.5	155.0	164.4	173.0	183.5	190.8
LN2	MoM	77.2	108.8	130.1	150.8	157.4	178.0	198.9	220.1	248.9	271.4
	MLM	76.9	110.0	132.6	154.8	161.9	184.2	206.9	230.1	261.7	286.4
	LMO	76.9	110.7	134.0	156.9	164.2	187.3	210.8	234.9	267.8	289.7
P3	MoM	78.4	111.1	131.6	150.2	156.0	173.3	189.9	206.0	226.6	241.8
	MLM	77.8	110.8	131.8	151.2	157.2	175.4	192.9	209.9	231.8	248.0
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LP3	MoM	76.9	110.4	133.4	156.0	163.3	186.1	209.2	233.0	265.4	316.9
	MLM	75.9	111.9	135.2	158.1	165.4	188.5	212.0	236.0	268.8	321.0
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
EV1	MoM	78.1	109.5	130.3	150.3	156.6	176.2	195.5	214.8	240.3	259.5
	MLM	78.0	109.8	130.9	151.1	157.5	177.2	196.8	216.3	242.0	261.5
	LMO	78.0	109.9	131.0	151.3	157.7	177.5	197.2	216.8	242.6	262.2
GEV	MoM	78.7	110.5	130.9	149.9	155.8	173.7	191.1	207.9	229.5	245.3
	MLM	78.3	109.8	130.3	149.7	155.8	174.4	192.6	210.6	233.9	251.3
	LMO	79.5	111.2	130.9	148.8	154.2	170.6	186.0	200.6	218.7	231.7

Table 15.6 Estimated 1-day maximum rainfall for different return periods by six probability distributions for Kalaburagi

Distribution	Method	Estimated 1-day maximum rainfall (mm) for a return period (in year)									
		2	5	10	20	25	50	100	200	500	1000
N2	MoM	81.7	100.4	110.2	118.3	120.7	127.4	133.5	139.1	145.8	150.5
	MLM	81.7	100.2	109.9	117.9	120.3	126.9	132.9	138.4	145.1	149.8
	LMO	81.7	100.7	110.7	118.9	121.3	128.2	134.3	140.0	146.8	151.6
LN2	MoM	78.8	98.7	111.1	122.4	126.0	136.6	147.0	157.1	170.4	180.4
	MLM	78.7	99.3	112.1	123.9	127.6	138.8	149.6	160.3	174.2	184.7
	LMO	78.7	99.9	113.3	125.6	129.4	141.1	152.4	163.6	178.3	190.5
P3	MoM	80.3	99.9	111.0	120.5	123.4	131.7	139.4	146.7	155.8	162.3
	MLM	80.1	99.6	110.7	120.4	123.3	131.8	139.7	147.2	156.6	163.4
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LP3	MoM	77.9	99.2	113.2	126.7	131.0	137.8	145.5	152.8	161.9	168.4
	MLM	78.7	99.5	112.5	124.5	128.2	139.5	150.5	161.4	175.5	186.2
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
EV1	MoM	78.0	97.7	110.8	123.3	127.2	139.5	151.6	163.7	179.7	191.7
	MLM	78.1	99.8	114.2	128.0	132.4	145.9	159.3	172.7	190.3	203.6
	LMO	77.8	98.7	112.5	125.8	130.0	142.9	155.8	168.6	185.5	198.3
GEV	MoM	80.1	100.1	111.5	121.1	123.9	131.9	139.0	145.2	152.4	157.2
	MLM	82.5	103.2	114.8	124.7	127.6	135.9	143.1	149.6	157.0	161.9
	LMO	79.6	100.2	112.2	122.6	125.7	134.7	142.9	150.3	159.1	165.0

Table 15.7 Computed and values of GoF and diagnostic tests statistic by six probability distributions for Afzalpur, Aland, and Kalaburagi

Distribution	Method	Afzalpur			Aland			Kalaburagi		
		χ^2	KS	D-index	χ^2	KS	D-index	χ^2	KS	D-index
N2	MoM	37.250	0.234	4.828	5.267	0.088	1.866	1.739	0.075	0.452
	MLM	25.250	0.218	4.793	6.734	0.091	1.426	1.739	0.075	0.479
	LMO	13.250	0.201	3.714	7.200	0.094	1.871	1.739	0.076	0.413
LN2	MoM	22.250	0.202	4.510	9.000	0.114	1.813	0.696	0.051	0.387
	MLM	23.250	0.205	3.379	10.100	0.120	1.799	0.698	0.053	0.340
	LMO	24.250	0.208	3.381	11.200	0.126	1.812	0.700	0.055	0.320
P3	MoM	36.250	0.277	3.200	11.133	0.095	1.810	0.957	0.074	0.375
	MLM	37.250	0.278	3.280	11.135	0.098	1.077	0.975	0.078	0.390
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA
LP3	MoM	8.250	0.118	3.287	8.525	0.105	1.898	2.783	0.069	0.301
	MLM	9.125	0.120	4.335	8.750	0.108	1.955	2.825	0.071	0.323
	LMO	NA	NA	NA	NA	NA	NA	NA	NA	NA
EV1	MoM	29.750	0.189	4.818	8.467	0.108	1.813	3.565	0.075	0.415
	MLM	30.250	0.192	3.404	8.725	0.110	1.030	3.675	0.078	0.362
	LMO	8.750	0.137	3.708	9.800	0.103	1.810	3.043	0.048	0.369
GEV	MoM	12.250	0.171	3.971	10.333	0.101	1.813	0.957	0.044	0.346
	MLM	7.250	0.125	3.234	10.333	0.100	1.075	0.827	0.060	0.383
	LMO	6.250	0.084	3.216	10.333	0.098	1.816	0.696	0.075	0.319

Computed values of GoF (viz., χ^2 and KS) and diagnostic (viz., D-index) tests

- χ^2 test results supported the use of MoM, MLM, and LMO estimators of all six distributions adopted in EVA of rainfall for Kalaburagi.
- KS test results didn't support the use of MoM, MLM, and LMO estimators of N2, LN2, and P3 distributions for EVA of rainfall for Afzalpur.
- KS test results confirmed the applicability of MoM, MLM, and LMO estimators of all six distributions adopted in the EVA of rainfall for Aland and Kalaburagi.

15.5.3   Analysis of Results Based on Diagnostic Test

In addition to the GoF test, the D-index was used to determine which of the six distributions in the EVA model best fit the criteria for estimating rainfall. These values were calculated using the N2, LN2, P3, LP3, EV1, and GEV distributions and are shown in Table 15.7. Based on the results of the diagnostic tests, it can be deduced that the D-index values of P3 (MoM) for Afzalpur, EV1 (MLM) for Aland, and LP3 (MoM) for Kalaburagi are less than those of other distributions used in the EVA.

15.5.4   Selection of Probability Distribution

Based on the EVA results from the diagnostic tests and GoF quantitative assessment, it was determined that the analysis produced conflicting inferences, necessitating qualitative evaluation. As a result, the best fit for rainfall estimates was again evaluated using fitted curves of the estimated severe rainfall along with D-index values, and a decision was taken as a result.

- According to the results of the diagnostic tests, EVA could be performed using P3 (MoM) for Afzalpur, EV1 (MLM) for Aland, and LP3 (MoM) for Kalaburagi.
- The MoM estimators of the distributions, however, are frequently less precise than MLM and LMO, as were previously mentioned. As a result, while choosing the best fit for estimating rainfall in Afzalpur and Kalaburagi, the D-index values obtained from P3 (MoM) and LP3 (MoM) are not taken into account.
- In light of the foregoing, it is determined that the D-index value of GEV (LMO) is the second subsequent minimum for Afzalpur and Kalaburagi after excluding the D-index values obtained from MoM of P3 and LP3 distributions from the selection.
- The GEV (LMO) is more matched among the six distributions chosen in EVA for rainfall estimation for Afzalpur and Kalaburagi, whereas EV1 (MLM) is better suited for Aland, according to the qualitative assessment (plots of EVA results) of the outcomes. Figure 15.2 shows the plots of the estimated 1-day maximum rainfall with 95% confidence limits based on the chosen distribution and actual AMR data for the Afzalpur, Aland, and Kalaburagi sites.
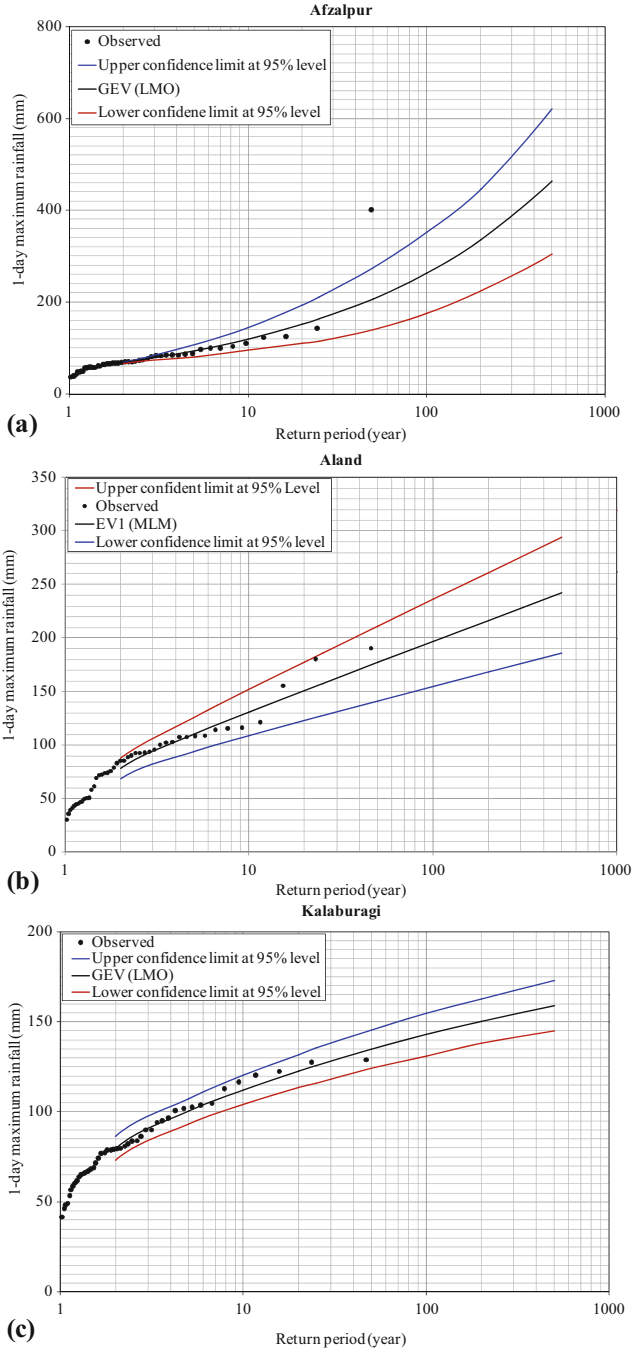
Fig. 15.2 Plots of estimated 1-day maximum rainfall by the selected distribution and observed AMR data for Afzalpur, Aland, and Kalaburagi

15.5.5   Efficiency Analysis of BR-ANN

Since a neural network demands comparatively more data to recognize the probable pattern (correlation) associated in a time series, the monthly rainfall dataset of all three sites, Afzalpur, Aland, and Kalaburagi, have been taken for 64 years, starting from January 1951 to December 2014. We have distributed this data for all three sites in the percentage of 70:30 for training and validation (testing) purposes to avoid overfitting in the performance evaluation of the predicting capability of ANN.

The performance of the ANN model with the BR technique used in predicting the targeted output has been expressed for training, validation, and the collective monthly rainfall data of all three sites, as shown in Fig. 15.3 (Afzalpur), Fig. 15.4 (Aland), and Fig. 15.5 (Kalaburagi). The efficiency of predicting outputs and targets through fit and errors has also been presented collectively for training and testing for all three sites. The model shows a high correlation between output and targeted input monthly rainfall data while training for all three stations. Since the model was trained for the training dataset, it has a higher correlation of approximately 0.95 between
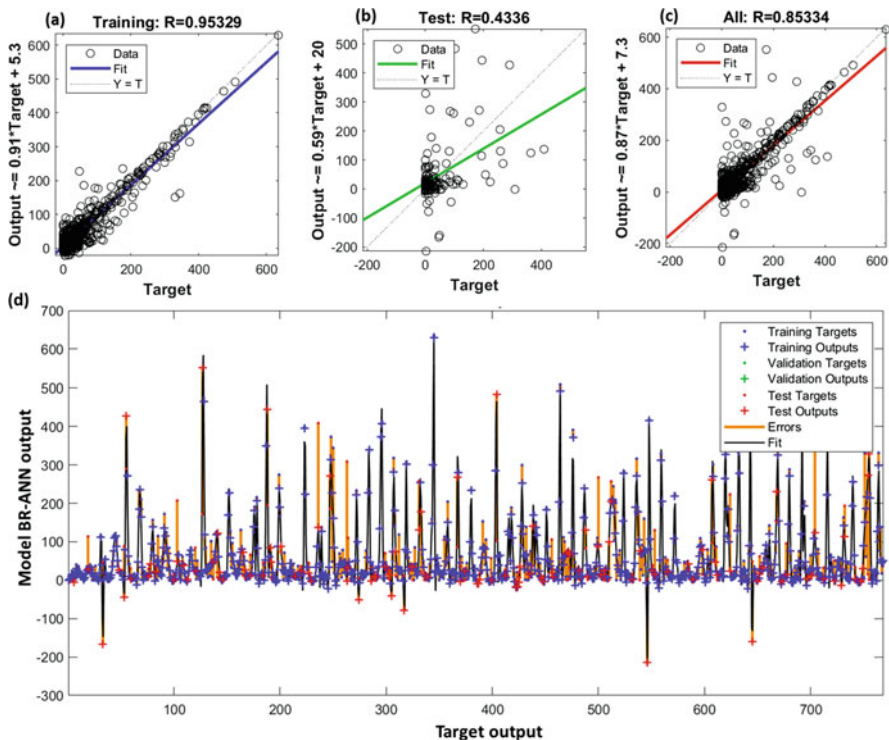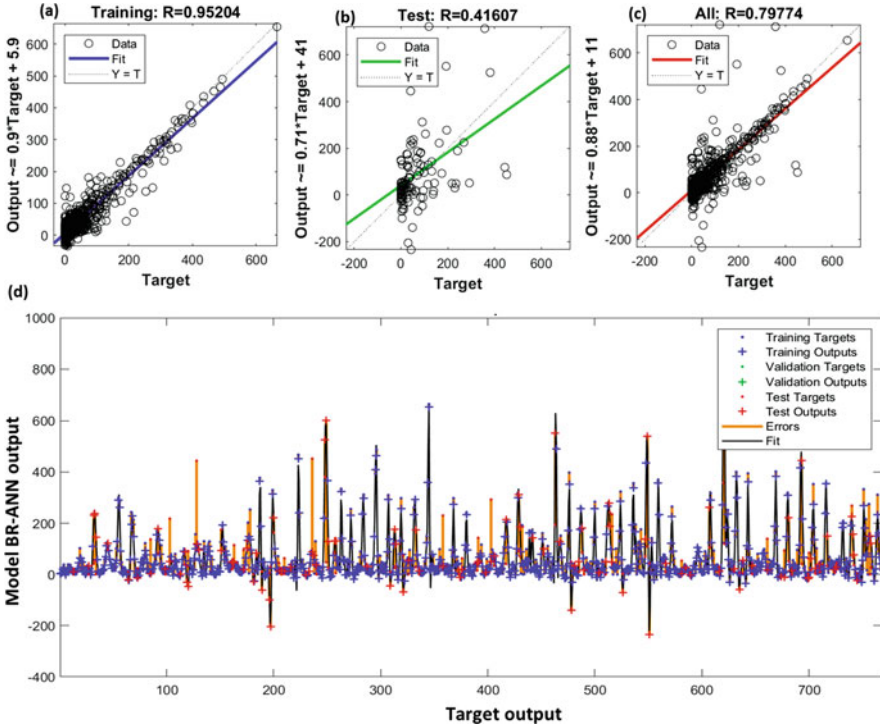

Fig. 15.3 Scatter plot showing correlation between the outputs of BR-ANN and input monthly data (target) of (**a**) training dataset; (**b**) testing dataset; and (**c**) all dataset of Afzalpur; (**d**) time series plot of predicted monthly rainfall using BR-ANN
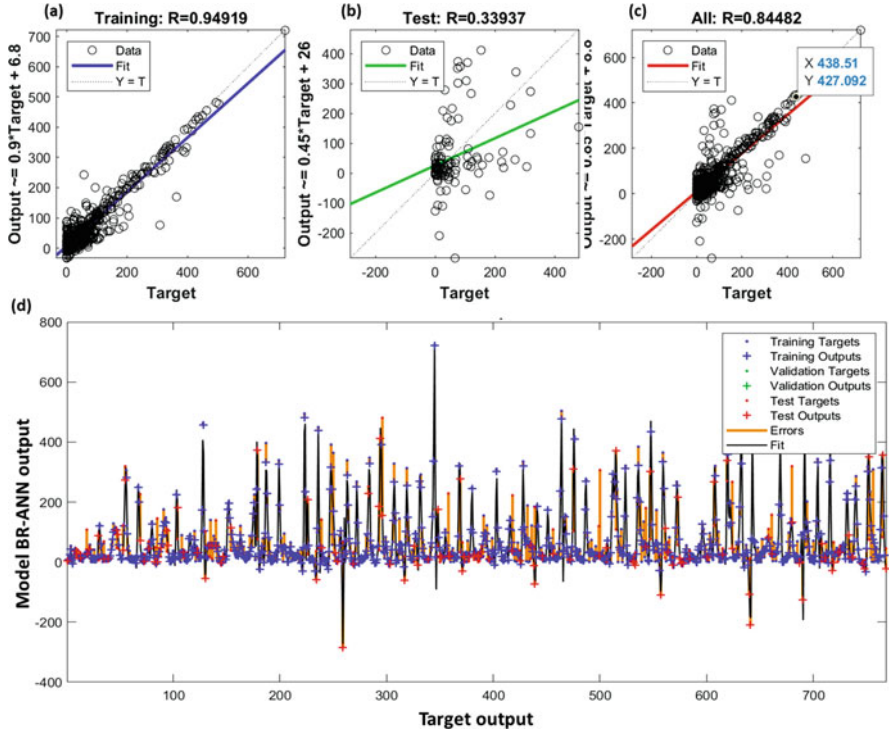
Fig. 15.4 Scatter plot showing correlation between the outputs of BR-ANN and input monthly data (target) of (**a**) training datasets; (**b**) testing dataset; and (**c**) all the dataset of Aland; (**d**) time series plot of predicted monthly rainfall using BR-ANN

model output and target, whereas it gets reduced to approximately 0.40 in the case of validation (testing).

15.5.6 Conclusions

This study compared the MoM, MLM, and LMO estimators of six probability distributions (N2, LN2, P3, LP3, EV1, and GEV) used in the EVA of rainfall for Afzalpur, Aland, and Kalaburagi with the aim of identifying the best distribution for rainfall estimation through quantitative (GoF tests using χ^2 and KS, and diagnostic test using D-index) and qualitative (fitted curves of the estimated rainfall) assessments. The monthly rainfall patterns of all three stations have also been proposed to model and predict using an ANN with Bayesian regularization technique. The inferences made from the study were condensed and are given below based on the EVA data and BR-ANN predictions:

Fig. 15.5 Scatter plot showing correlation between the outputs of BR-ANN and input monthly data (target) of (**a**) training datasets; (**b**) testing dataset; and (**c**) all the dataset of Kalaburagi; (**d**) time series plot of predicted monthly rainfall using BR-ANN

- χ^2 test results didn't support the use of all six distributions for EVA of rainfall for Afzalpur. For Aland, the χ^2 test result indicates the N2 distribution is acceptable for EVA.
- χ^2 test results supported the use of all six distributions for EVA of rainfall for Kalaburagi.
- KS test results didn't support the use of N2, LN2, and P3 distributions for EVA of rainfall for Afzalpur.
- KS test results confirmed the applicability of all six distributions for EVA of rainfall for Aland and Kalaburagi.
- The GEV (LMO) is the superior choice among the six distributions chosen in EVA for rainfall estimation for Afzalpur and Kalaburagi, while EV1 (MLM) is better suited for Aland, according to the qualitative assessment (plots of EVA findings) of the outcomes, which was weighed with D-index values.
- The BR-ANN model has shown high correlation (approximately 0.95) in predicted output values and targeted input data at monthly scale from all three sites while training the datasets.

- The proposed model has shown comparatively lesser correlation between output and target values in validation, which means the model can be further optimized by hyper tuning the Bayesian regularization parameters.

By considering the data length (i.e., 48 years for Afzalpur, 45 years for Aland, and 46 years for Kalaburagi) of the AMR series available for the study, the study suggested that the estimated rainfall for the return period beyond 200 years may be cautiously used due to uncertainty in the higher-order return periods while designing the hydraulic structures in the respective sites. Some other machine learning algorithms, such as support vector machine (regression), Bayesian linear regression, etc., and some deep learning models, such as recurrent neural network, long short-term memory, etc., can also be explored to model nonstationary and complex rainfall patterns in a comparative analysis.

References

Abida H, Ellouze M (2008) Probability distribution of flood flows in Tunisia. Hydrol Earth Syst Sci 12(3):703–714. https://doi.org/10.5194/hess-12-703-2008

Acar R, Celik S, Senocak S (2008) Rainfall intensity-duration-frequency (IDF) model using an artificial neural network approach. J Sci Ind Res 67(3):198–202

Afungang R, Bateira C (2016) Statistical modelling of extreme rainfall, return periods and associated hazards in the Bamenda Mountain, NW. Geogr Spat Plan J 1(9):5–19. https://doi.org/10.17127/got/2016.9.001

AlHassoun SA (2011) Developing empirical formulae to estimate rainfall intensity in Riyadh region. J King Saud Univ Eng Sci 23(1):81–88. https://doi.org/10.1016/j.jksues.2011.03.003

Arvind G, Kumar PA, Girish Karthi S, Suribabu CR (2017) Statistical analysis of 30 years rainfall data: a case study. Proceedings of IOP Conference Series: Earth and Environmental Science. 80, 012067. https://doi.org/10.1088/1755-1315/80/1/012067

Baghel H, Mittal HK, Singh PK, Yadav KK, Jain S (2019) Frequency analysis of rainfall data using probability distribution models. Int J Curr Microbiol Appl Sci 8(6):1390–1396. https://doi.org/10.20546/ijcmas.2019.806.168

Baratti E, Montanari A, Castellarin A, Salinas JL, Viglione A, Bezzi A (2012) Estimating the flood frequency distribution at seasonal and annual time scales. Hydrol Earth Syst Sci 16(12):4651–4660. https://doi.org/10.5194/hess-16-4651-2012

Bauer P, Thorpe A, Brunet G (2015) The quiet revolution of numerical weather prediction. Nature 525:47. https://doi.org/10.1038/nature14956

Bhuyan A, Borah M, Kumar R (2010) Regional flood frequency analysis of north-bank of the river Brahmaputra by using LH-moments. Water Resour Manag 24(9):1779–1790. https://doi.org/10.1007/s11269-009-9524-0

Bobee B, Askhar F (1991) The gamma family and derived distributions applied in hydrology. Water Resources Publications, Littleton, CO

Burden F, Winkler D (2008) Bayesian regularization of neural networks, pp 23–42. https://doi.org/10.1007/978-1-60327-101-1_3

Central Water Commission (2010) Development of hydrological design aids (surface water): State of the Art Report (Report No.2009097/WR/REP-02). Central Water Commission, New Delhi

Charles Annis PE 2009 Goodness-of-fit tests for statistical distributions

Chen, Y., & Zhang, N. (2022). Optimal subsampling for large sample ridge regression

Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London. https://doi.org/10.1007/978-1-4471-3675-0

CWPRS (2021) Desk Studies on Estimation of Flood at Various Railway Bridges on Daund-Kalaburagi Line Doubling Project for Railway Vikas Nigam Limited, Mumbai. Technical Report No. 5940, July 2021, Pune

Das J, Jha S, Goyal MK (2020) On the relationship of climatic and monsoon teleconnections with monthly precipitation over meteorologically homogenous regions in India: Wavelet & global coherence approaches. Atmos Res 238:104889. https://doi.org/10.1016/j.atmosres.2020.104889

Dueben PD, Bauer P, Adams S (2021) Deep learning to improve weather predictions. In: Deep learning for the earth sciences. Wiley, Hoboken, NJ, pp 204–217. https://doi.org/10.1002/9781119646181.ch14

Esberto MDP (2018) Probability distribution fitting of rainfall patterns in Philippine regions for effective risk management. Environ Ecol Res 6(3):178–186. https://doi.org/10.13189/eer.2018.060305

Esteves LS (2013) Consequences to flood management of using different probability distributions to estimate extreme rainfall. J Environ Manag 115(1):98–105. https://doi.org/10.1016/j.jenvman.2012.11.013

Floods E (2019) Extreme floods and droughts under future climate scenarios, pp 1–5

Goswami UP, Hazra B, Goyal MK (2018) Copula-based probabilistic characterization of precipitation extremes over North Sikkim Himalaya. Atmos Res 212:273–284. https://doi.org/10.1016/j.atmosres.2018.05.019

Goyal MK, Ojha CSP (2010) Evaluation of various linear regression methods for downscaling of mean monthly precipitation in arid Pichola watershed. Nat Res 01(01):11–18. https://doi.org/10.4236/nr.2010.11002

Goyal MK, Ojha CSP, Burn DH (2012) Nonparametric statistical downscaling of temperature, precipitation, and evaporation in a semiarid region in India. J Hydrol Eng 17(5):615–627. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000479

Goyal MK, Panchariya VK, Sharma A, Singh V (2018) Comparative assessment of SWAT model performance in two distinct catchments under various DEM scenarios of varying resolution, sources and resampling methods. Water Resour Manag 32(2):805–825. https://doi.org/10.1007/s11269-017-1840-1

Gubareva TS, Gartsman BI (2010) Estimating distribution parameters of extreme hydrometeorological characteristics by L-moments method. Water Resour 37(4):437–445. https://doi.org/10.1134/S0097807810040020

Haberlandt U, Radtke I (2014) Hydrological model calibration for derived flood frequency analysis using stochastic rainfall and probability distributions of peak flows. Hydrol Earth Syst Sci 18(1):353–365. https://doi.org/10.15488/602

Hewitson BC, Daron J, Crane RG, Zermoglio MF, Jack C (2014) Interrogating empirical-statistical downscaling. Clim Chang 122:539. https://doi.org/10.1007/s10584-013-1021-z

Hinge G, Surampalli RY, Goyal MK (2018) Prediction of soil organic carbon stock using digital mapping approach in humid India. Environ Earth Sci 77(5):172. https://doi.org/10.1007/s12665-018-7374-x

Hosking JRM (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. J R Stat Soc 52(1):105–124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

Hosking JRM, Wallis JR (1993) Some statistics useful in regional frequency analysis. Water Resour Res 29:271–281. https://doi.org/10.1029/92WR01980

Katz RW (2013) Statistical methods for nonstationary extremes. In: Extremes in a changing climate. Springer, Dordrecht, pp 15–37. https://doi.org/10.1007/978-94-007-4479-0_2

Malekinezhad H, Nachtnebel HP, Klik A (2011) Regionalization approach for extreme flood analysis using L-moments. Agric Sci Technol 13:1183–1196

Mujere N (2011) Flood frequency analysis using the Gumbel distribution. J Comput Sci Eng 3(7): 2774–2778

Nardi KM, Barnes EA, Ralph FM (2018) Assessment of numerical weather prediction model reforecasts of the occurrence, intensity, and location of atmospheric rivers along the West Coast of North America. Mon Weather Rev 146(10):3343–3362. https://doi.org/10.1175/MWR-D-18-0060.1

Okut H (2016) Bayesian regularized neural networks for small n big p data. In: Artificial neural networks—models and applications. InTech, Lahore. https://doi.org/10.5772/63256

Olumide BA, Saidu M, Oluwasesan A (2013) Evaluation of best fit probability distribution models for the prediction of rainfall—runoff volume (case study Tagwai dam, Minna-Nigeria). Int J Eng Technol 3(2):94–98

Pisner DA, Schnyer DM (2020) Chapter 6—support vector machine. In: Mechelli A, Vieira SBT-ML (eds) Machine learning. Academic Press, Cambridge, MA, pp 101–121. https://doi.org/10.1016/B978-0-12-815739-8.00006-7

Poonia V, Goyal MK, Gupta BB, Gupta AK, Jha S, Das J (2021a) Drought occurrence in different river basins of India and blockchain technology based framework for disaster management. J Clean Prod 312:127737. https://doi.org/10.1016/j.jclepro.2021.127737

Poonia V, Jha S, Goyal MK (2021b) Copula based analysis of meteorological, hydrological and agricultural drought characteristics across Indian river basins. Int J Climatol 41(9):4637–4652. https://doi.org/10.1002/joc.7091

Rao AR, Hamed KH (2000) Flood frequency analysis. CRC Publications, New York, NY

Rasel M, Hossain SM (2015) Development of rainfall intensity duration frequency equations and curves for seven divisions in Bangladesh. Int J Sci Eng Res 6(5):96–101

Rautela KS, Kumar D, Gandhi BGR, Kumar A, Dubey AK (2022) Application of ANNs for the modeling of streamflow, sediment transport, and erosion rate of a high-altitude river system in Western Himalaya, Uttarakhand. RBRH 27:e22. https://doi.org/10.1590/2318-0331.272220220045

Saf B (2009) Regional flood frequency analysis using L-moments for the west mediterranean region of Turkey. Water Resour Manag 23(3):531–551. https://doi.org/10.1007/s11269-008-9287-z

Sasireka K, Suribabu CR, Neelakantan TR (2019) Extreme rainfall return periods using Gumbel and gamma distributions. Int J Recent Technol Eng 8(4):27–29. https://doi.org/10.35940/ijrte.D1007.1284S219

Scher S, Messori G (2018) Predicting weather forecast uncertainty with machine learning. Q J R Meteorol Soc 144(717):2830–2841. https://doi.org/10.1002/qj.3410

Sharma A, Goyal MK (2020) Assessment of drought trend and variability in India using wavelet transform. Hydrol Sci J 65(9):1539–1554. https://doi.org/10.1080/02626667.2020.1754422

Sharma NK, Sharma S (2019) Frequency analysis of rainfall data of Dharamshala region. Int J Sci Res 8(2):886–892. https://doi.org/10.21275/ART20195211

Singh S, Goyal M, Jha S (n.d.) Role of large-scale climate oscillations in precipitation extremes associated with atmospheric rivers: nonstationary framework. Hydrol Sci J 68:395. https://doi.org/10.1080/02626667.2022.2159412

Su X, Yan X, Tsai C-L (2012) Linear regression. Wiley Interdiscip Rev Comput Stat 4(3):275–294. https://doi.org/10.1002/wics.1198

Tank G, Dongre P, Obi Reddy GP, Sen P (2021) Rainfall trend analysis—a review. Int Res J Eng Technol 8(4):4028–4030

United States Water Resources Council (USWRC) (1981) Guidelines for determining flood flow frequency. In: Bulletin 17A. U.S Geological Survey, Washington, D.C

Vergés A, Doropoulos C, Malcolm HA, Skye M, Garcia-Pizá M, Marzinelli EM, Vila-Concejo A (2016) Long-term empirical evidence of ocean warming leading to tropicalization of fish

communities, increased herbivory, and loss of kelp. Proc Natl Acad Sci 113:13791. https://doi.org/10.1073/pnas.1610725113

Vivekanandan N (2020) A comparative study on Gumbel and LP3 probability distributions for estimation of extreme rainfall. Int J Water Res Eng 6(1):21–33

Vivekanandan N, Srishailam C (2020) Selection of best fit probability distribution for extreme value analysis of rainfall. Water Energy Int 63(10):13–19

Vu HL, Ng KTW, Bolingbroke D (2019) Time-lagged effects of weekly climatic and socio-economic factors on ANN municipal yard waste prediction models. Waste Manag 84:129–140. https://doi.org/10.1016/j.wasman.2018.11.038

Ye L, Jabbar SF, Abdul Zahra MM, Tan ML (2021) Bayesian regularized neural network model development for predicting daily rainfall from sea level pressure data: investigation on solving complex hydrology problem. Complexity 2021:1–14. https://doi.org/10.1155/2021/6631564

Zhang J (2002) Powerful goodness-of-fit tests based on the likelihood ratio. J R Stat Soc 64(2):281–294. https://doi.org/10.1111/1467-9868.00337