



A new method based on artificial bee colony programming for the regional standardized intensity–duration–frequency relationship

Beyza Görkemli¹ · Hatice Citakoglu² · Tefaruk Haktanir³ · Dervis Karaboga^{1,4}

Received: 14 July 2021 / Accepted: 15 December 2021 / Published online: 28 January 2022
© Saudi Society for Geosciences 2021

Abstract

The aim of this study was to develop new analytical models to define the regional rainfall intensity – duration – frequency (IDF) relationship for the Inland Anatolia Region, which is determined discretely by the L-Moments method at many values of rainfall durations and frequencies. First, the parameters of each one of nine commonly used empirical equations were calibrated to provide the best possible definition of the IDF relationship for the Inland Anatolia Region. Next, analytical models best fitted to the IDF relationship of the L-Moments method were generated by the artificial bee colony programming (ABCP) approach, such that a combination of nine different sets were simulated, taking into account three cost functions and three maximum depths. Mean absolute error, root mean square error, mean square error, Nash–Sutcliffe efficiency coefficient, Willmott’s refined index, performance index, and coefficient of determination were computed to assess the accuracies of the empirical equations and of the ABCP models. These criteria revealed that the ABCP models defined the IDF relationship better than the empirical equations over the entire range of frequencies from 2 to 10,000 years. The accuracy of the empirical equations is much worse than the ABCP model, especially for frequencies smaller than 2000 years. Finally, Kruskal – Wallis tests were applied on all of the IDF relationships given by (1) the L-Moments method, (2) the empirical equations, and (3) the ABCP approach. These results indicated that the numerical values of these three models were from the same population.

Keywords Regional intensity – duration – frequency relationship · Artificial bee colony programming (ABCP)

Communicated by Broder J. Merkel.

✉ Hatice Citakoglu
hcitakoglu@erciyes.edu.tr

Beyza Görkemli
bgorkemli@erciyes.edu.tr

Tefaruk Haktanir
thaktanir@nny.edu.tr

Dervis Karaboga
karaboga@erciyes.edu.tr

- ¹ Computer Engineering Department, Engineering Faculty, Erciyes University, Kayseri, Turkey
- ² Civil Engineering Department, Engineering Faculty, Erciyes University, Kayseri, Turkey
- ³ Civil Engineering Department, Nuh Naci Yazgan University, Engineering Faculty, Kayseri, Turkey
- ⁴ Erciyes University Big data and Artificial Intelligence Research and Application Center, Kayseri, Turkey

Introduction

An extreme rainfall of a definite duration has the greatest magnitude among many rainfalls of the same duration that occur in a single year. Hence, a gauged series of extreme rainfall of a specific duration contains as many numerical values of rainfall as the record length in years. Peculiar to a geographical location, the intensity of an extreme rainfall event (I) as a function of its duration (D) and its statistical frequency (F) is depicted by the intensity–duration–frequency (IDF) relationship, which is given in either a tabular or a diagrammatic form. The IDF relationship is needed to design various hydraulic structures, such as storm drainages, combined sewage systems of residential areas, and flood spillways of dams. The IDF relationship in a diagrammatic form consists of as many curves as the frequencies taken into consideration. Each curve presents the intensity of rainfall in mm/h units pouring onto a geographical point over a certain time period in minutes, having an average frequency in years, which is also called the average return period. Hence,

the IDF relationship is a significant tool for the planning, design, and risk assessment stages in water resources engineering. The numerical values of an IDF relationship are determined by statistical frequency analyses applied on a significantly long series of gauged extreme rainfall records. The magnitude of the intensity of rainfall is inversely proportional to its duration and directly proportional to the frequency. Highly intense rainfalls are of short duration and the intensity gradually increases with increasing frequency (F). Statistical frequency analysis using suitable probability distributions is the common tool to relate the magnitude of extreme rainfalls to their frequencies (e.g., Bernard 1932; Hershfield 1961; Chow 1988; Koutsoyiannis et al. 1998; Dupont and Allen 2006; Asikoglu and Benzedden 2014).

In hydrometric studies, streamflow and rainfall are measured as point data in places where gauging stations are available. However, it is impossible to measure hydrologic data at every point on earth. Hence, hydraulic structures are not always optimally designed because of missing point measurements at their locations. Because hydrologic data are point-specific, they cannot directly be used in places away from the gauging stations, and they should be adapted to the points where they are needed. Similarly, determining the spatial distribution of rainfall over a specific drainage area from available point measurements is a common practice of utmost importance.

As a pioneering contribution, Bernard (1932) developed an empirical equation in the form: $i = (c_1 \cdot T^{c_2})/t^{c_3}$ for the IDF relationship, where, c_1 , c_2 , and c_3 are coefficients peculiar to a geographical region. Since then, various analytical expressions have been presented for different regions of the world. Froehlich (1995) presented rainfall intensity–duration equations for durations of 1 through 24 h for four geographical regions delineated by the United States Weather Bureau. He determined optimum parameter values for each one of four equations and gave the ratio of t –h rainfall depth to 1-h rainfall depth for various frequencies. Bartual and Schneider (2001) applied frequency analyses on a series of many extreme rainfall events in the Alicante province of Spain, gauged between the years 1925 through 1992, by using the general extreme values (GEV) distribution for standard durations from 2 to 240 min, and developed nine different empirical equations for the IDF relationship. As a result of their work, they noted that three-parameter expressions were more meaningful for frequencies between 5 and 500 years. Yu et al. (2004) developed a regional IDF relationship for ungauged locations based on the scaling theory combining the Gumbel distribution and partial simple scaling hypothesis. Nhat et al. (2006) presented curves for the IDF relationship for the Yodo River basin and derived a generalized IDF formula. AlHassoun (2011) developed an empirical formula to estimate the rainfall intensity in the Riyadh region of Saudi Arabia and showed

a good match between the Gumbel method and the other analytical methods. Jaleel and Farawn (2013) obtained the IDF relationship for the Basrah province of Iraq by using Gumbel distribution and noted that maximum rainfall intensities exhibited high variability in short periods. Dourte et al. (2013) developed IDF relationships for the Indian peninsula and investigated the effects of variations in rainfalls on flows and groundwater replenishment. They noted that the updated IDF relationships revealed a significant change in rainfall characteristics as compared to the old relationships and reported increased rainfall intensities for all rainfall durations and frequencies. They concluded that severe rainfalls resulted in decreased groundwater replenishment and increased flows. Asikoglu and Benzedden (2014) used the two-parameter lognormal and Gumbel distributions in frequency analyses and obtained two types of IDF functions for the Aegean region of Turkey, based on the simple generalization procedure and the robust estimation procedure. These researchers compared the root mean squared error (RMSE) values of both procedures and reported that the former procedure yielded better rainfall intensity values than the latter. Rasel and Islam (2015) conducted a frequency analysis by applying the Gumbel and Log – Pearson – 3 distributions to derive the IDF relationship for Bangladesh by using the data of extreme rainfalls measured between 1974 and 2014. They calculated the IDF equation parameters for 2, 5, 10, 25, 50, and 100-year frequencies by non-linear multiple regression and reported high correlation coefficients for the resultant IDF equations. Liuzzo and Freni (2015) used climate change scenarios for the Sicilia Province and investigated the IDF relationships of extreme rainfalls for durations of 1, 3, 6, 12, and 24 h. They first conducted a trend analysis on all extreme rainfalls by the Mann–Kendal method and determined decreasing–increasing trends. Next, they used the outcomes of the trend analyses to devise an empirical equation for a climate scenario–dependent IDF relationship. They noticed that decreasing and increasing trends influenced the intensity of extreme rainfalls with frequency. Hamaamin (2017) determined an IDF relationship for the Sulaimani province in Iraq, and achieved a determination coefficient of almost 1 for rainfall intensity values in the empirical equation, and hence indicated that the resultant empirical equation represented the actual values. There are several other research studies conducted in different regions of the world for generations of empirical rainfall intensity estimation equations and IDF curves (Bougadis and Adamowski 2006, Nhat et al. 2007, Trevor and Guillermo 2008, Omotosho and Oluwafemi 2009, Olofintoye et al. 2009, El–Sayed, 2011, Vivekanandan 2012, Antigha and Ogarekpe 2013, Al–anazi and El–sebaie, 2013, Chang et al. 2013, García–Marín et al. 2013, Wang et al. 2014, Akinsanola and Ogunjobi 2014).

Karahan et al. (2007) employed a genetic algorithm (GA) approach for rainfall intensity estimations and indicated that the GA method employing mean square error as the cost function could reliably be used to determine the IDF relationship. They also reported a good match for the measured and estimated values and applied this method for the IDF relationships in the GAP region of Turkey. These researchers stated that the GA method was an efficient tool to determine the mathematical model coefficients enabling it to represent the available data successfully (Karahan et al. 2008). Başakın et al. (2021a) developed an equation with the help of genetic programming (GP) using the information of the IDF relationships computed as the outcome of statistical frequency analyses to calculate intensities of extreme rainfalls for the Kayseri, Nevşehir, Niğde, and Yozgat provinces. They also used one of the empirical equations available in the literature for comparison with their model. They optimized the parameters of the empirical equation with the particle swarm optimization (PSO) scheme. In the end, they concluded that the equation derived by GP had higher accuracy than the equation obtained with PSO.

It is a known fact that algorithms such as GA and PSO can be employed to find the optimal values of the parameters of a predetermined mathematical expression. In order to obtain the best estimation accuracy, first a best model must be chosen and next the best values for the parameters of that model must be determined. Hence, with this objective in mind, the artificial bee colony programming (ABCP) was used in this study to obtain a model for the IDF relationship with high estimation accuracy. ABCP was introduced in 2012 by Karaboga et al. (2012) as a versatile means of machine learning. Golafshani and Ashour (2016) employed the ABCP method to predict the elasticity modulus of self-compacting concrete. Golafshani and Behnood (2018) applied the ABCP approach for predicting the elasticity modulus of recycled aggregate concrete. Boudardara and Gorkemli (2018) proposed the ABCP method in a robotic path planning problem. They later presented a new version of ABCP (Boudardara and Gorkemli 2020). Arslan and Ozturk (2019) provided an ABCP image descriptor for multi class texture classification. Prior to that, they used ABCP as a new tool for cancer data classification (Arslan and Ozturk 2018). Gorkemli and Karaboga (2019) proposed three new versions of ABCP to improve the convergence performance of the algorithm. Hara et al. (2018) proposed a modified ABCP method using semantic control crossover. Boudouaoui and Habbi (2018) made some structural modifications to the ABCP method. Arslan and Ozturk (2019) introduced a feature selection method based on ABCP for high-dimensional symbolic regression problems.

ABCP is a novel, evolutionary, and GP (Koza 1992)-like automatic programming and machine learning technique. It is based on the artificial bee colony (ABC) optimization

algorithm (Karaboga 2005, 2010; Karaboga and Basturk 2007), which is inspired from the foraging behavior of honeybee swarms while the GP is an extension of GA. ABC is one of the most popular and widely used swarm intelligence based optimization algorithms for solving various types of problems from a wide spectrum of disciplines (Karaboga 2010, Bansal et al. 2013, Karaboga et al. 2014, Agarwal and Yadav 2019, Pooja and Shirmal 2020). Therefore, in this study, the ABCP model is proposed as a new and efficient approach to represent the numerical values of rainfall intensities over a wide range of frequencies from 2 to 10,000 years analytically. This range of frequencies resulted as the outcome of comprehensive regional statistical frequency analyses by the L-Moments method using 14 standard-duration-recorded series of extreme rainfalls carried out by Haktanir et al (2016). In the first phase of the study, nine different empirical equations available in literature for the IDF relationship were tested with the purpose of how well they represented the IDF relationship of the Inland Anatolia Region, as obtained by the L-Moments method. In the second phase, equations were developed by the ABCP approach for the same regionalized IDF relationship. In the last phase of the study, the performances of the nine empirical equations and the equations produced by the ABCP method were compared from the standpoint of their accuracies in representing the regionalized IDF relationship developed by the L-Moments method. These three stages are concisely explained in the following sections.

Hence, the main objective of this study was to quantitatively depict the regional IDF relationship peculiar to the Inland Anatolia Region, as determined by the outcome of comprehensive frequency analyses by the L-Moments method. This was done in a succinct analytical way to be obtained by the ABCP approach and to compare its accuracy against the accuracies of already used empirical equations, whose parameters were optimized based on the numerical values given by the L-Moments method.

Material and method

Research site

Seven geographical regions of Turkey and the positions of 31 meteorological stations in the Inland Anatolia Region, each having at least 30 years of gauged data, are shown in Fig. 1. Prior to this study, a regional frequency analysis by the L-Moments method was completed for the Inland Anatolia Region of Turkey by Haktanir et al. (2016). This study used the gauged series of successively increasing 14 standard – duration extreme rainfalls (AMR) with durations of 5, 10, 15, and 30 min, and 1, 2, 3, 4, 5, 6, 8, 12, 18, and 24 h at the 31 stations.

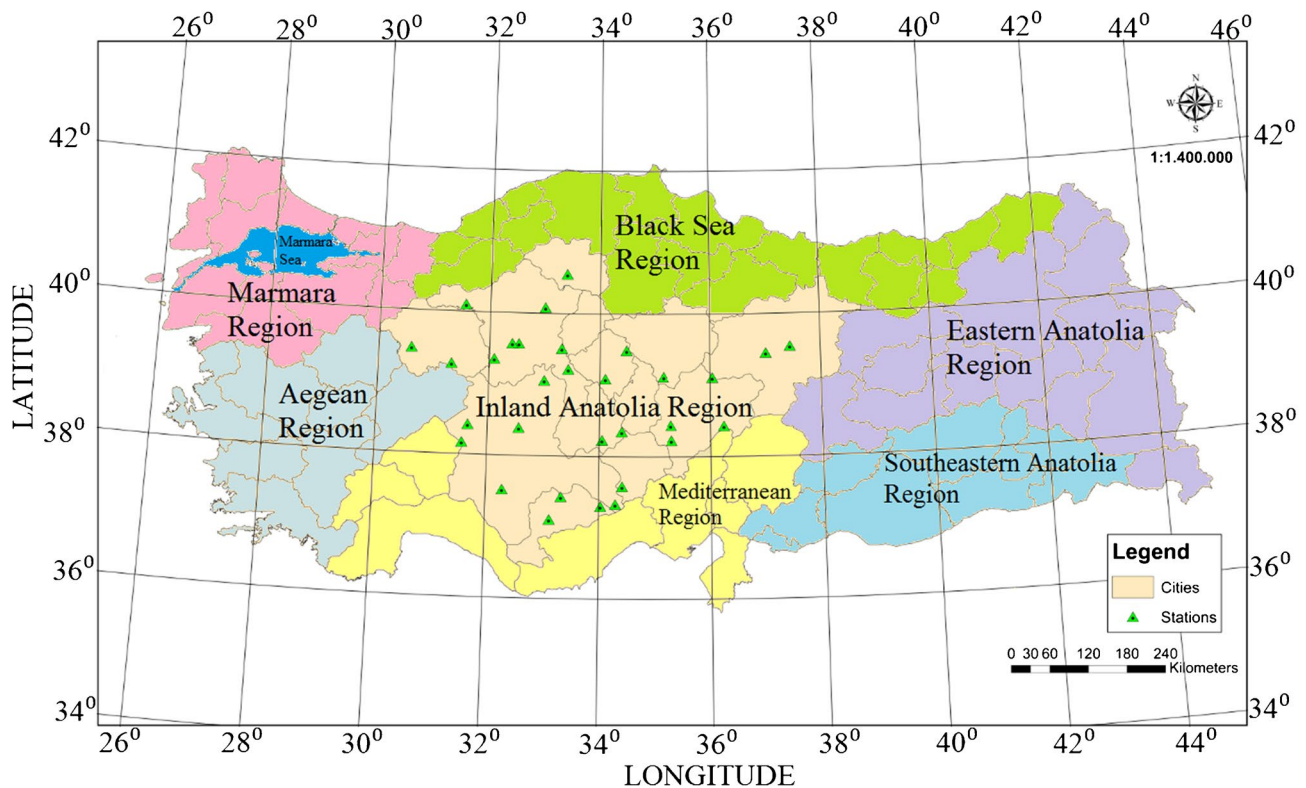


Fig. 1 Seven geographical regions of Turkey and locations of the meteorological stations in the Inland Anatolia Region, whose gauged data of 14 standard–duration extreme rainfalls series, were used in

determining the regional intensity–duration–frequency relationship for the Inland Anatolia Region by the L-Moments method

Including the year 2010, all rainfall data were obtained from the General Directorate of Meteorology. Although there were more than 31 stations in this region, Haktanir et al. (2016) did not include those having record lengths shorter than 30 years in the frequency analyses, whose ultimate product was the regional IDF relationship consisting of many intensity values for 14 rainfall durations and for many frequencies ranging between 2 and 10,000 years. In the current study, a best-fit single mathematical model was searched for the regional IDF relationship obtained by Haktanir et al. (2016) as the outcome of comprehensive frequency analyses, based on a modified version of the L-Moments method.

The northern sections of the Inland Anatolia Region are surrounded by mountains reaching altitudes of 3000 m, which generally extend parallel to the Black Sea shoreline. The southern sections of the Inland Anatolia Region are surrounded by mountains with altitudes up to 3500 m, and they mostly lie parallel to the Mediterranean Sea coastline. The Inland Anatolia Region is a large plateau bordered by these rows of mountains in the North and South. Erciyes and Suphan are two independent volcanic mountains in

this region. The mountain rows in the north and the south extend into the Eastern Anatolia Region where they intersect. The mountains in the western sections of the Inland Anatolia Region lie mostly perpendicular to the Aegean Sea coast. In brief, the Inland Anatolia Region constitutes a landlocked portion of Turkey and experiences a dominant terrestrial climate. Forests and woods sparsely exist in this region because of the harsh climate with less than average precipitation and steppes are common plant covers. There is less vegetable cultivation in this region compared to the other regions because of insufficient rainfalls and prevailing droughts. Yet, the major amount of precipitation occurs in spring, and severe storms produce considerable rainfalls of durations up to 24 h. Mostly cold-resistant food plants are cultivated in the Inland Anatolia Region. Potatoes, rye, green lentils, apples, and pears are the most common agricultural products.

Frequency analyses by L-Moments method

Statistical frequency analysis is developed to allow practitioners to make probabilistic predictions in the future.

In addition, it is used to quantify the relationship between the magnitude and the probability of either exceedance or non-exceedance of hydrologic random variables. Commonly, frequency analyses in hydrology are applied for flood peaks and extreme rainfalls. The frequency analysis is done using a suitable probability distribution, and usually, one of three-parameter probability distributions of generalized normal (GNO), which is also known as the three-parameter log-normal, Pearson-3 (PE3), general extreme values (GEV), generalized logistic (GLO), and generalized Pareto (GPA) are used for this purpose. Estimation of the magnitudes of the parameters of a probability distribution using the observed series is an important stage of the frequency analysis and the conventional methods of moments and of maximum-likelihood are commonly used for this purpose. The L-Moments method has become recently popular because of its various merits.

The L-Moments method, which was originally put forth by Hosking (1990), and Hosking and Wallis (1997), is especially suitable for regional frequency analyses. The L-Moments of a random variable are some linear combinations of its probability-weighted moments, which were originally defined and presented to the statistics literature by Greenwood et al (1979). The L-Coefficients defined, as ratios of the L-Moments, which take values within -1 and $+1$, are able to depict the overall peculiarities of probability distributions, such as mean, variance, and skewness. There exist analytical relationships among the distribution parameters and the L-Coefficients peculiar to any probability distribution. Therefore, the parameters are computed using the estimates of the L-Coefficients calculated using the available recorded series. First, that particular three-parameter distribution most suitable to represent a homogeneous region from the frequency analysis standpoint is determined, based on the weighted average of the L-Coefficients computed for each station where gauged data exists. Ultimately, the regional growth curve is obtained, which is a curve relating the magnitude of the normalized random variable, whose mean is 1.0, to its probability of non-exceedance (P_{nex}). Growth curves usually are extended up to an average return period of 10,000 years. Aside from the regional growth curve, a meaningful regression equation relating the mean value of the random variable to a few explanatory variables, which are relevant geographical and meteorological characteristics of the homogeneous region, is obtained. The value taken from the growth curve is multiplied by the mean given by the regression equation to obtain the magnitude of the hydrologic variable at any geographical location in the region. The problem with 14 successively increasing standard – duration extreme rainfalls is more

involved than a single hydrologic variable, like an annual flood peak. First, there are 14 growth curves. Second, more importantly, the parameters of each probability distribution suitable for each standard-duration extreme rainfalls series must be modified, so as to satisfy the divergence criterion, which emanates from the principle of conservation of mass (Haktanir et al. 2010). Detailed information about the L-Moments method to be applied to successively increasing standard – duration extreme rainfalls series can be found in Haktanir et al. (2016).

Artificial bee colony programming (ABCP) method

Since ABCP is based on the ABC algorithm, its basic steps are similar to those of ABC. However, while ABC can be employed to find optimal values for the parameters of a model, ABCP can be used to optimize the model and its parameters altogether (Karaboga et al. 2012; Gorkemli and Karaboga 2019). The basic steps of the ABCP are given below.

- Initialization phase.
- Repeat.
- Employed Bee Phase.
- Onlooker Bee Phase.
- Memorize the best solution found so far.
- Scout Bee Phase.
- Until (Termination criteria are satisfied).

In ABCP, there are three foraging bee groups, namely, employed bees, onlooker bees, and scout bees, as in ABC. An employed bee has a food source in her mind and she leaves the hive to exploit it. When she returns to the hive, she performs some dances to give information about her food source to the onlookers. The onlooker bees wait in the hive and watch the employed bees' dances. Considering these dances, an onlooker bee selects a food source and leaves the hive to exploit it. Every time an employed bee or onlooker bee goes to a source to exploit it, first, she searches for a better food source around the current one. If she finds a better food source, she starts to exploit it; otherwise, she continues to exploit the current source. In the colony of ABCP, the number of the employed bees is taken to be equal to the number of the onlooker bees. Since every employed bee has one food source, the number of the employed bees also equals the number of the food sources in the population. If the food source of an employed bee is exhausted, then this bee becomes a scout and randomly finds a new food source. In a cycle or iteration of the optimization process, the employed bee, onlooker bee, and scout bee phases are run orderly.

In ABCP, a food source refers to a tree structure based solution. A solution is given in tree form in Fig. 2 as an example. The mathematical formulation obtained from this

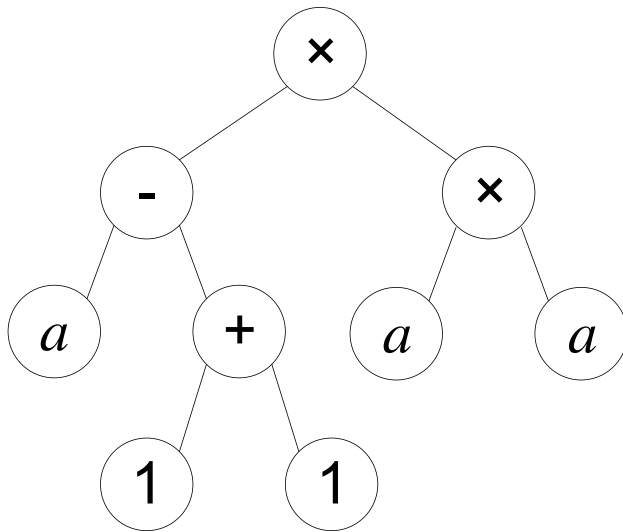


Fig. 2 An exemplary solution by the ABCP approach in tree form

solution is presented in Eq. 1, where a and $f(a)$ refer to the independent and dependent variables, respectively. In a solution tree, there can be function nodes or terminal nodes. Function nodes can include some operators, like arithmetic functions, trigonometric functions, and logic operators. These nodes must have one or more operands. Terminal nodes cannot have any operands or any child nodes since they can only include some constants and input variables.

$$f(a) = (a - (1 + 1)) \times (a \times a) \tag{1}$$

Algorithm 1 Steps of the information sharing mechanism of the ABCP scheme.

In order to determine the quality of the solutions from their costs, Eq. 2 was used in this study, as was done by Karaboga and Ozturk (2012) and Gorkemli and Karaboga (2019),

$$fit(x_i) = \frac{1}{1 + f(x_1)} \tag{2}$$

where, $fit(x_i)$ is the quality of the solution x_i and $f(x_i)$ is the cost function value of this solution. After the employed bee phase, a probability of being selected by an onlooker is assigned to each food source according to its quality using Eq. 3 below.

$$P_1 = \frac{\alpha \times fit(x_i)}{fit(x_{best})} + (1 - \alpha) \tag{3}$$

where α is a parameter assuming values between 0 and 1. In this study, 0.9 is taken, as the value of α . x_{best} is the best solution in the population.

Steps of the ABCP approach are given in Algorithm 2 below. The flowchart of the ABCP method is given in Fig. 3.

- Select a solution x_k from the population randomly ($i \neq k$).
- Randomly select a sub tree, n_x , from the current solution, x_i , so that, its root node will include a function with the probability, P_{ip} ($P_{ip} = 0.9$) and a terminal with the probability $1 - P_{ip}$.
- Randomly select a sub tree, n_k , from the current solution, x_k , so that, its root node will include a function with the probability P_{ip} and a terminal with the probability $1 - P_{ip}$.
- Produce the candidate, v_i , by replacing the sub tree, n_x , with the sub tree, n_k in x_i .

Although its basic optimization structure is similar to the ABC algorithm, some mechanisms cannot be directly used in ABCP, since the solutions are in tree form. For producing a candidate, v_i , from the current solution, x_i , (the i th solution in the population), and the procedure summarized in Algorithm 1 below are used in the employed bee and onlooker bee phases.

Algorithm 2 Basic steps of the ABCP scheme.

For more details about the ABCP method, interested researchers can read the publications by Karaboga et al. (2012) and Gorkemli and Karaboga (2019).

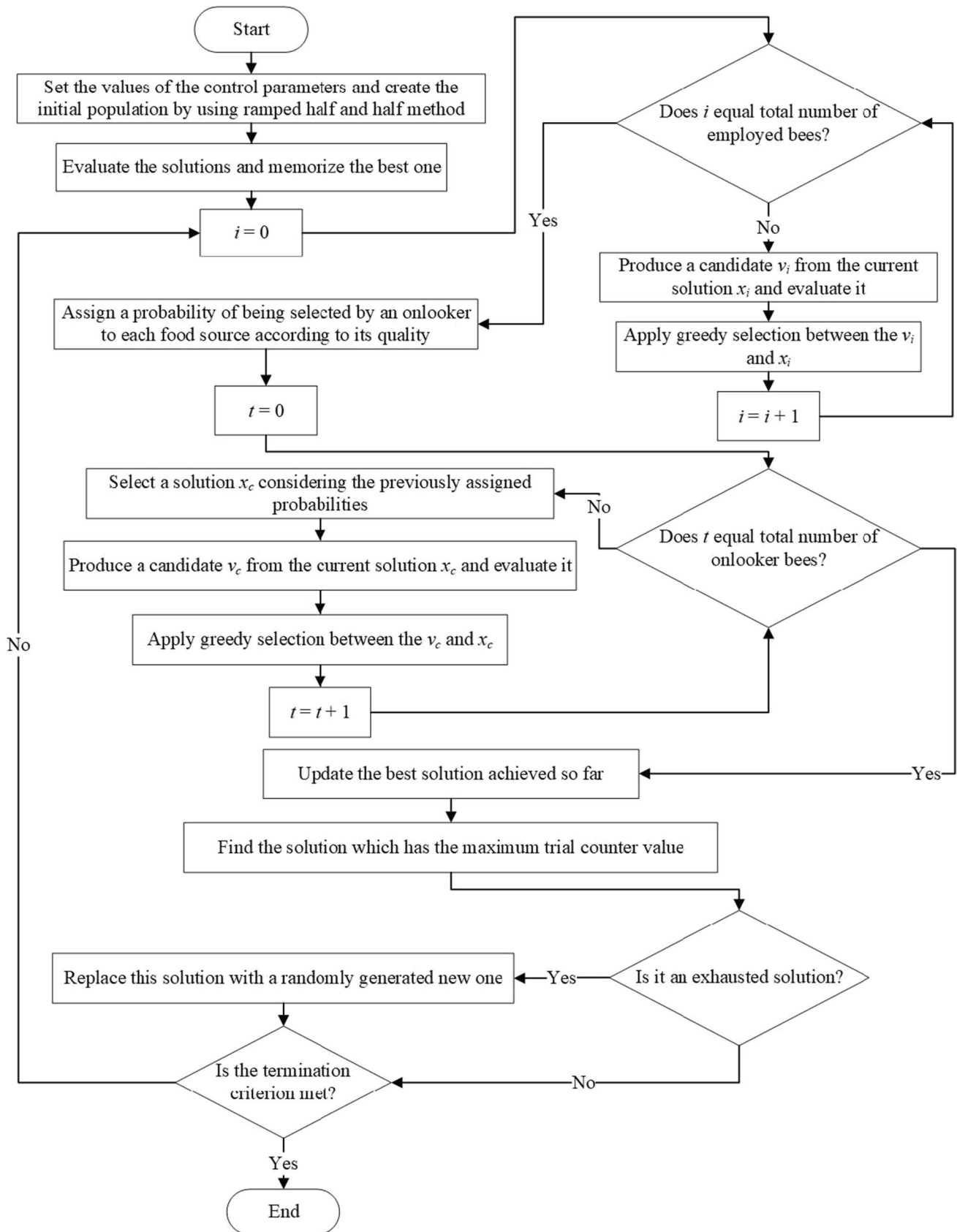


Fig. 3 Flowchart of the ABCP method

- Set the values of the control parameters, like colony size, limit, maximum number of cycles, maximum initial depth of the trees, maximum depth of the trees, etc.
- Create the initial population of food sources (solutions) using the ramped half and half method (avoid any duplications in the population). Set the trial counter value of each solution as 0.
- Evaluate the solutions and memorize the best one.
- counter=0.

REPEAT

Employed Bee Phase:

- For each employed bee
 - Produce a candidate v_i from the current solution x_i using the information sharing mechanism (Algorithm 1) and evaluate it.
 - Apply greedy selection between v_i and x_i . If v_i is selected, set the trial counter of the related solution as 0. Otherwise, increase the value of the trial counter.
- Assign a probability of being selected by an onlooker to each food source according to its quality using Equation 2.

Onlooker Bee Phase:

- For each onlooker bee
 - Select a solution x_c considering the previously assigned probabilities.
 - Produce a candidate v_c from the current solution x_i , using the information sharing mechanism (Algorithm 1) and evaluate it.
 - Apply greedy selection between v_c and x_c . If v_c is selected, set the trial counter of the related solution as 0. Otherwise, increase the value of this trial counter.

- Update the best solution achieved so far.

Scout Bee Phase:

- Find the solution which has the maximum trial counter value in the population. Determine if this solution is exhausted or not by comparing its trial counter value to limit. If it is exhausted, then randomly create a new solution using the grow method (avoid any duplications in the population) and replace it with the exhausted one. Set the trial counter value of new solution as 0.
- counter = counter + 1.

UNTIL (counter = the maximum number of cycles)

Empirical rainfall intensity – duration – frequency equations

IDF equations are the empirical formulas defining the relationship between the average intensity of an extreme rainfall as the dependent variable and rainfall duration and frequency as the independent variables. The ranges of rainfall duration and frequency are 5–1440 min and 2–10,000 years, respectively, in most countries. When the relevant publications to date were examined, it was seen that such empirical formulas for IDF could be analytically generalized by Eq. 4 below.

$$i(t, T) = \frac{A(T)}{B(t)} \tag{4}$$

The most common equations for the nominator of Eq. 4 are as follows:

$$A(T) = a \cdot T^b \tag{5}$$

(Sherman, 1931; Bernard, 1932; Koutsoyionnis, 1994).

$$A(T) = a + b \cdot \ln(T) \tag{6}$$

(Koutsoyionnis, 1998)

$$A(T) = a + b \cdot [\ln(\ln T)] \tag{7}$$

(Nhat et al. 2007).

While, the most common equations for the denominator of Eq. 4 are as follows:

$$B(t) = t^c \tag{8}$$

(Sherman, 1931; Bernard, 1932)

$$B(t) = (t^c + d) \tag{9}$$

(Nhat et al. 2007)

$$B(t) = (t + d)^c \tag{10}$$

(Koutsoyionnis, 1998).

Combining $A(T)$ and $B(t)$, Lopcu (2007) presented the nine expressions given below, which were used in this study.

$$i = (a \cdot T^b)/(t + c)^d \tag{11}$$

$$i = (a \cdot T^b)/t^c \tag{12}$$

$$i = (a \cdot T^b)/(t^c + d) \tag{13}$$

$$i = (a + b \cdot \ln T)/t^c \tag{14}$$

$$i = (a + b \cdot \ln T)/(t^c + d) \tag{15}$$

$$i = (a + b \cdot \ln T)/(t + c)^d \tag{16}$$

$$i = (a + b \cdot [\ln(\ln T)])/(t + c)^d \tag{17}$$

$$i = (a + b \cdot [\ln(\ln T)])/t^c \tag{18}$$

$$i = (a + b \cdot [\ln(\ln T)])/(t^c + d) \tag{19}$$

In these equations, i is the standard rainfall intensity (l/ min); T is the frequency (years); t is the rainfall duration (min); and a , b , c , and d are the coefficients.

Table 1 The numerical values of the coefficients of the nine empirical equations for the IDF relationship of the Inland Anatolia Region

Equations	a	b	c	d
Equation 11	53.2515294	0.1467572	-0.1035154	0.8248525
Equation 12	92.2926127	0.05807680	0.84735567	—
Equation 13	54.1774186	0.14665879	0.82946426	-0.0272955
Equation 14	29.270894	18.689827	0.83403453	—
Equation 15	28.4540421	18.1675984	0.82675173	-0.0647524
Equation 16	28.4634851	18.1736242	-0.0990213	0.8264291
Equation 17	33.5020426	62.3925318	-0.1056031	0.8332015
Equation 18	34.5309787	64.3097618	0.84145318	—
Equation 19	33.4623956	62.3177325	0.83331679	-0.0722126

Table 2 The error statistics of the nine empirical equations for representing the IDF relationship obtained by the modified L-Moments method

Equations	MSE	RMSE	MAE	MARE	NSE	PI	WI	R ²
Equation 11	1.049	1.024	0.603	50.082	0.989	0.088	0.956	0.990
Equation 12	7.657	2.767	1.504	74.752	0.923	0.242	0.891	0.927
Equation 13	1.049	1.024	0.602	49.009	0.989	0.088	0.956	0.990
Equation 14	0.435	0.659	0.452	40.494	0.996	0.057	0.967	0.996
Equation 15	0.434	0.659	0.455	41.962	0.996	0.057	0.967	0.996
Equation 16	0.435	0.659	0.452	40.494	0.996	0.057	0.967	0.996
Equation 17	2.385	1.544	0.924	48.197	0.976	0.133	0.933	0.976
Equation 18	2.385	1.544	0.924	48.194	0.976	0.133	0.933	0.976
Equation 19	2.384	1.544	0.927	49.544	0.976	0.133	0.933	0.976

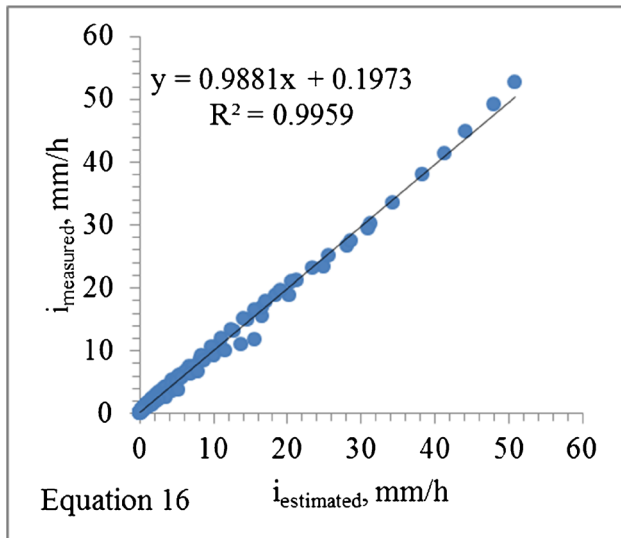


Fig. 4 Scatter plot of Eq. 16, which was the best among the nine empirical equations for fitting the IDF relationship obtained by the L-Moments method for the Inland Anatolia Region

Implementations and results

Rainfall intensity–duration–frequency relationship

The IDF relationship used in this study has been developed

by a modified version of the L-Moments method taking into consideration the ‘divergence criterion’ using the extreme rainfalls data observed at 31 meteorological gauging stations in Inland Anatolia Region. The final IDF relationship is a long table comprising values of intensity against so many durations and frequencies, which is also presented in the form of plots of so many points in a log–log scale graph. The summary of this study by Haktanir et al. (2016) can be phrased as follows:

- Frequency analyses by the L-Moments method was separately conducted for 14 extreme rainfalls series with sequential durations from 5 min to 24 h.
- At the very beginning, for each station, all of the 14 extreme rainfalls series were normalized by dividing each recorded rainfall value by the sample mean of that particular duration.
- The incompatibility test was applied to all extreme rainfalls series of 31 stations from 5-min to 24-h durations of each station and it was checked if the incompatibility criterion (D statistic) was satisfied.
- The homogeneity analyses were applied to all of the 14 extreme rainfalls series of 31 stations, and it was concluded that the Inland Anatolia Region was a single homogeneous region for all of 31 × 14 extreme rainfalls series.
- The parameters of the GLO, GEV, GNO, PE3, and GPA distributions were computed using the average L-Coef-

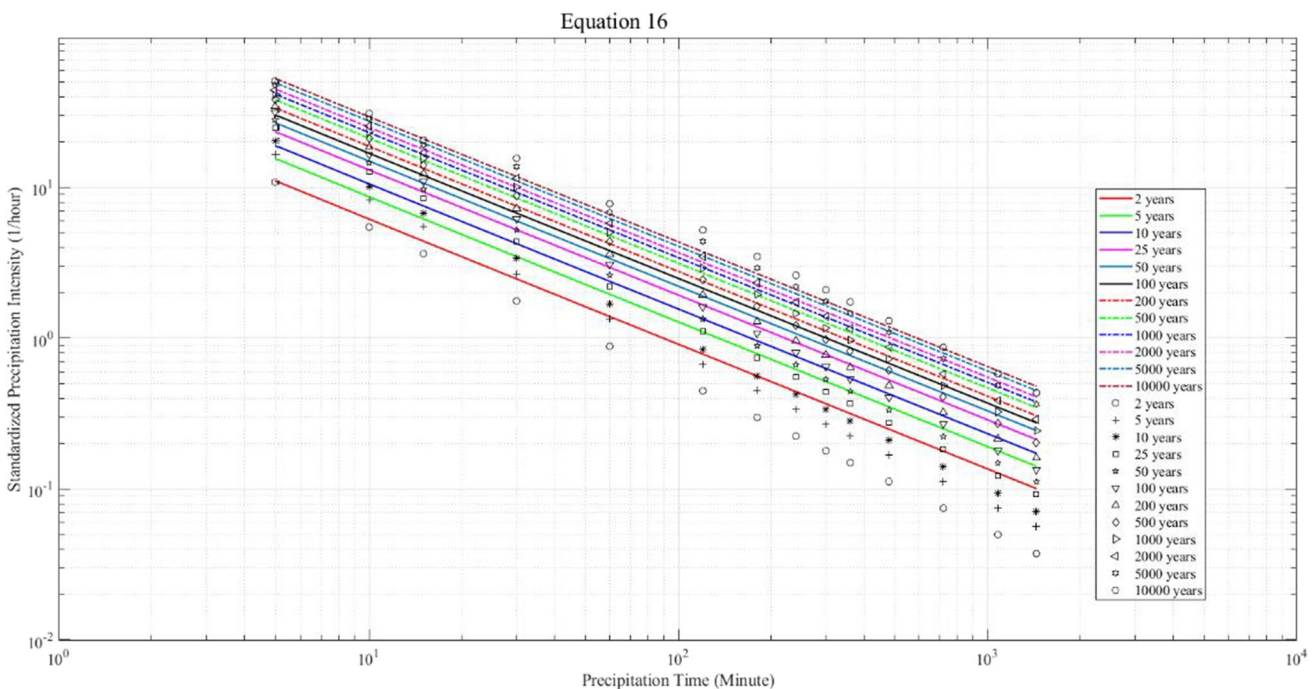


Fig. 5 Graphical presentation of the numerical values obtained by the L-Moments method and the curves yielded by Eq. 16 with its optimized coefficients for the IDF relationship for Inland Anatolia Region for frequencies from 2 to 10,000 years

Table 3 Values of the parameters of the ABCP approach as applied to the IDF relationship of the Inland Anatolia Region

Parameters	Values
Colony size	500
Number of assessments	500,000
Limit	500
Function set	+ , - , × , / (protected version), pow, exp, log (protected version)
Terminal set	<i>t, T, R</i>
Initial tree depth	2, 3, 6
Maximum tree depth	2, 3, 6

ficients. Following 500 synthetic simulations, the Z^{DIST} statistics were computed for the GLO, GEV, GNO, PE3, and GPA distributions.

- Because they had the smallest Z_{DIST} statistics among the others, the most suitable distributions were: PE3 distribution for the 5 min, the GNO distribution for the 10 min and 15 min, the GEV distribution for the 30-min and 60-min series, and the GLO distribution for all of the other extreme rainfalls series between 120 and 1440 min.
- The magnitudes of the coefficients of the significant explanatory variables of the multiple regression equation for the means of extreme rainfalls, which turned out to be rainfall duration, longitude, annual average precipitation, and annual average relative humidity, were computed. The average values yielded by the regression equation were multiplied by the magnitudes given by the growth curves of all of the extreme rainfalls series. Thus, the numerical values of the regional IDF relationship were obtained.

Empirical IDF equations and search for the best one for the regionalized IDF relationship

The numerical values of the regional IDF relationship obtained by Haktanir et al. (2016), as summarized above,

were used in checking the accuracy of the common nine empirical equations to represent the regionalized relationship. The coefficients of Eqs. 11 through 19 were determined to define the regionalized relationship in the best way possible by its analytical form. The best numerical values of the *a, b, c,* and *d* coefficients in Eqs. 11 – 19 were determined so that each one of these equations come as close to the regional IDF relationship for the Inland Anatolia Region as possible with the Excel Solver tool aid. Next, to test the estimation accuracy of the equations with the best values for the coefficients, the mean square error (MSE) and determination coefficient (R^2) were used as error statistics. For the Inland Anatolia Region, the best values of the coefficients of the nine equations and the error statistics are given in Tables 1 and 2. Although some of the equations were more successful, the resultant IDF relationships defined by these nine expressions yielded fairly accurate outcomes. Yet, Eq. 16, having the smallest MSE (0.4343) and the greatest R^2 (0.995), was the best among them. The scatter plot of the rainfall intensities computed by Eq. 16 against the generalized rainfall intensities obtained by the frequency analyses by the L-Moments method for the Inland Anatolia Region is presented in Fig. 4.

At first glance, the rainfall intensities obtained by the regional frequency analyses and those computed by Eq. 16 seem to be close to each other. Considering the R^2 values given in Table 2, Eqs. 11 and 12 yielded close outcomes, while Eqs. 13, 14, and 15 provided less accurate estimates. Yet, as seen in Table 2 and Fig. 4, Eq. 16 yielded intensities closest to those given by the regional IDF relationship. However, when the IDF relationship was plotted on a graph, in Fig. 5, it was visually observed that this empirical equation, Eq. 16, was not so successful in representing the overall IDF relationship. It can be observed in Fig. 5 that the IDF values obtained by the frequency analyses and those given by Eq. 16 do not reveal close conformity for all of the ranges of the frequencies. Equation 16 yielded successful outcomes for the frequencies of 2000, 5000, and 10,000 years, but was remarkably unsuccessful for shorter frequencies. The overall

Table 4 Summaries of the MSE values obtained in 30 runs by the ABCP method, with three different maximum depths and three different cost functions

Maximum depth	Cost function	Mean	Standard deviation	The best	The worst
2	Equation 20	49.1913557	24.1202094	3.673946	117.179494
	Equation 21	175.1015573	336.1776537	16.983126	1601.242286
	Equation 22	281.7737057	329.8247712	17.287928	1230.662705
3	Equation 20	19.4484865	18.4608345	0.919986	65.99849
	Equation 21	675.6272232	2319.0614622	3.894141	12,980.019825
	Equation 22	255.0361107	325.8483690	4.748996	1149.669268
6	Equation 20	1.1877460	1.5837633	0.212856	8.734532
	Equation 21	3.8883458	4.4935622	0.272695	16.74901
	Equation 22	2.2836237	3.0458527	0.300312	15.402084

Table 5 The analytical models obtained by the ABCP approach, with three different cost functions (CF) and three different maximum depths (MD)

Model	CF	MD
Model 1	$\frac{\log(T)}{a \cdot t}$	1
Model 2	$\frac{T^a}{b \cdot t}$	2
Model 3	$\frac{\log(T)}{a \cdot t}$	3
Model 4	$\frac{1}{a} \left(\frac{\log(T)}{t} + \frac{\log(t)}{T} \right)$	1
Model 5	$\frac{\log(T)}{a \cdot \log(b^t)}$	2
Model 6	$\frac{T^a \cdot t^c}{c \cdot b \cdot t}$	3
Model 7	$\frac{a}{T^{2a} \cdot t^2} - \left(\frac{\log(T) + b^c}{d} b \cdot t \right) + \frac{\log(b \cdot t)}{T^e} / t \cdot T^a$	1
Model 8	$\frac{\log(\log(T) - T)^{\exp(\exp(a))}}{\log(t^d) + \log(b^c) + t} - \frac{\log((d/T)/T)^{\log(b^c)}}{t+a}$	2
Model 9	$\frac{\log\left(\left(2 \cdot \frac{T}{a}\right)^{\log\left(\frac{T}{a}\right)}\right)}{t} + \frac{\log\left(\left((T \cdot t) + \frac{t}{a}\right)^{\log(T)}\right)}{(t \cdot \log(T))^b + \left(\frac{t+b}{T}\right)^b}$	3

result is that even the best of these nine empirical equations available in the relevant literature for the IDF relationship is not able to define the regionalized IDF relationship accurately.

Determination of the best-fit expression by the ABCP method for the regional IDF relationship

Other analytical expressions were developed by the ABCP approach, which would comply with the IDF relationship determined by the L-Moments method for the Inland Anatolia Region. For this purpose, experimental studies were conducted with different tree depths and different cost functions of the ABCP method. Since models with different structures could be achieved by using different error functions, three different cost functions, which are Eqs. 20, 21, and 22 below, were taken into consideration, and the ABCP scheme was run with the parameter values given in Table 3.

$$f_i = \sum_{j=1}^N |g_j - t_j| \quad (1st\ cost\ function) \tag{20}$$

$$f_i = \sum_{j=1}^N \left| \log_{10}(g_j + 1) - \log_{10}(t_j + 1) \right| \quad (2nd\ cost\ function) \tag{21}$$

$$f_i = \sqrt{\frac{1}{N} \cdot \sum_{j=1}^N (\log_{10}(g_j + 1) - \log_{10}(t_j + 1))^2} \quad (3rd\ cost\ function) \tag{22}$$

In these equations, g_j represents the rainfall intensity obtained by the ABCP method, t_j represents the rainfall intensity obtained by the L-Moments method (the target

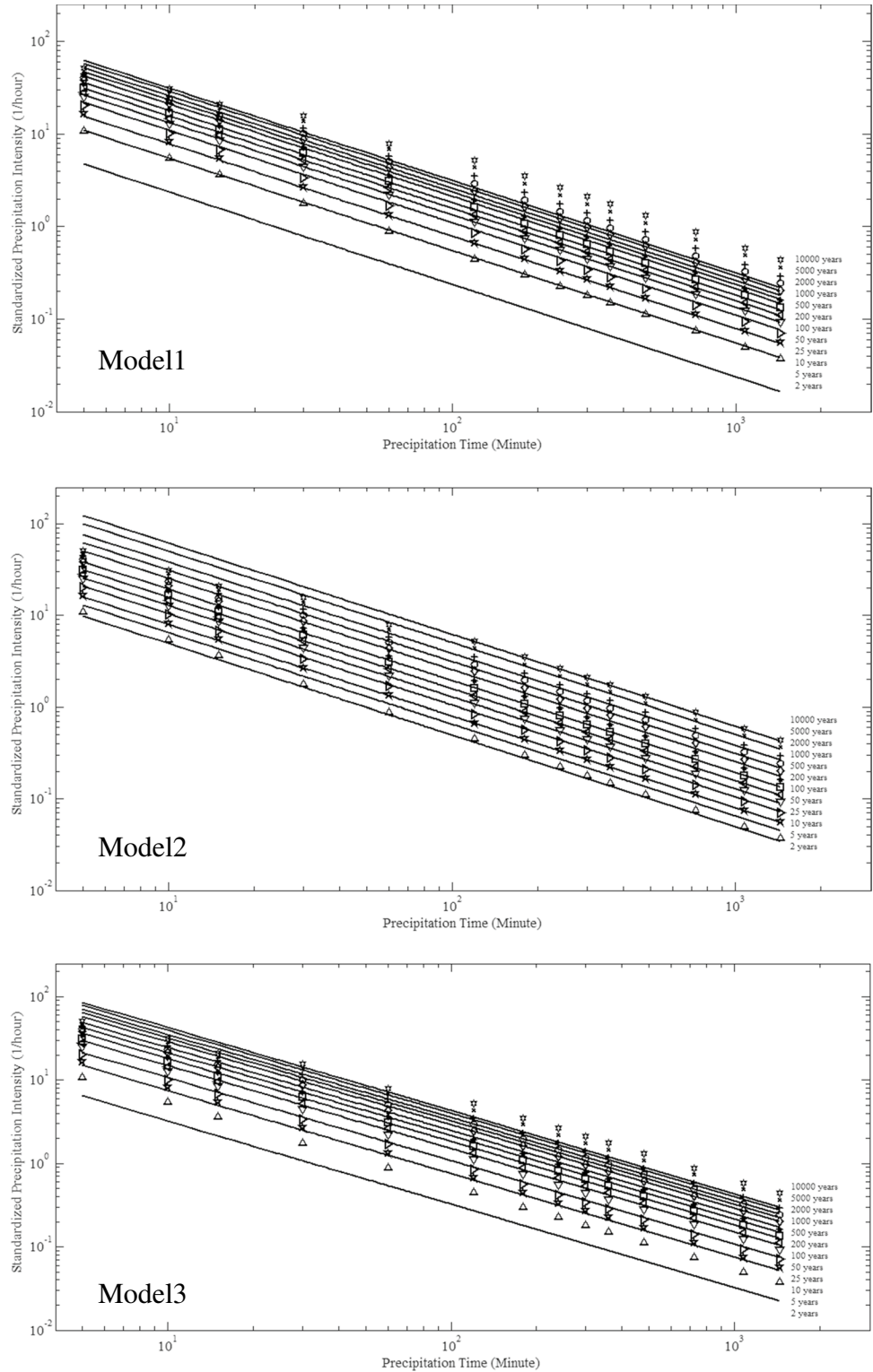
Table 6 The numerical values of the coefficients of nine ABCP-based models presented in Table 5

Models	a	b	c	d	e
Model 1	-0.0292625152642199				
Model 2	0.296428124092719	-0.024820394359911			
Model 3	0.0214301119658304				
Model 4	-0.0419094995790671				
Model 5	-0.33381610239568	-0.968115260809714			
Model 6	0.24165744206014	-0.548901953524399	0.0359635842200199		
Model 7	-0.284778959716102	0.105890298777209	-0.0972491507871306	0.334895958814256	-0.210784611855999
Model 8	-0.0678627933691548	-0.232998991959262	0.373578949539726	0.0814660671546432	
Model 9	0.00238452339655937	-0.380376258576464			

Table 7 The error statistics of nine ABCP-based models for representing the IDF relationship obtained by the modified L-Moments method for the Inland Anatolia Region

Models	MSE	RMSE	MAE	MARE	NSE	PI	WI	R ²
Model 1	3.674	1.917	0.903	26.934	0.963	0.166	0.934	0.977
Model 2	73.026	8.546	2.313	18.386	0.265	0.756	0.832	0.889
Model 3	31.069	5.574	1.957	18.356	0.687	0.482	0.858	0.977
Model 4	1.019	1.009	0.623	43.144	0.99	0.087	0.955	0.991
Model 5	3.894	1.973	0.779	14.989	0.961	0.170	0.943	0.989
Model 6	33.797	5.814	1.536	11.285	0.66	0.507	0.888	0.940
Model 7	0.213	0.461	0.209	6.182	0.998	0.040	0.985	0.998
Model 8	2.635	1.623	0.407	4.245	0.973	0.140	0.970	0.987
Model 9	0.316	0.562	0.276	7.83	0.997	0.048	0.980	0.997

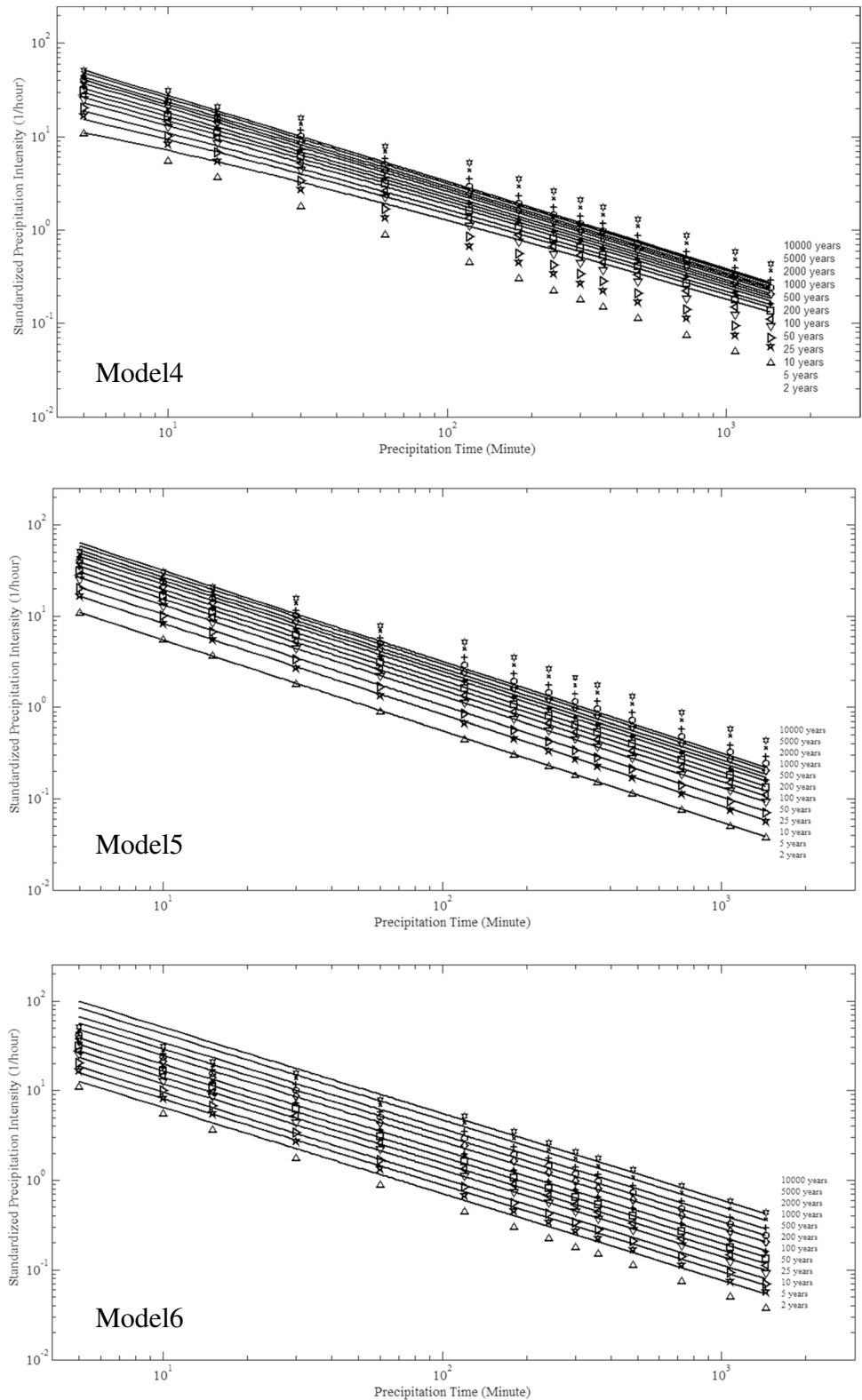
Fig. 6 Graphical presentation of the numerical values obtained by the L-Moments method and the curves yielded by the ABCP model with maximum depth of 2 for the IDF relationship for the Inland Anatolia Region for frequencies from 2 to 10,000 years



value), and N equals 168 (14×12) in this study. The first expression (Eq. 20) defines the model quality based on the sum of absolute values of differences between the target value (t_j) of all cases and the model outcome (g_j) (SAE: sum of absolute errors). Equation 20 is one of the most used cost

functions in evolutionary computation based automatic programming applications (Koza 1992; Karaboga et al. 2012; Gorkemli and Karaboga 2019). In the other expressions (Eqs. 21 and 22), the target and model outcomes are assessed through logarithmic errors. Because the IDF relationship

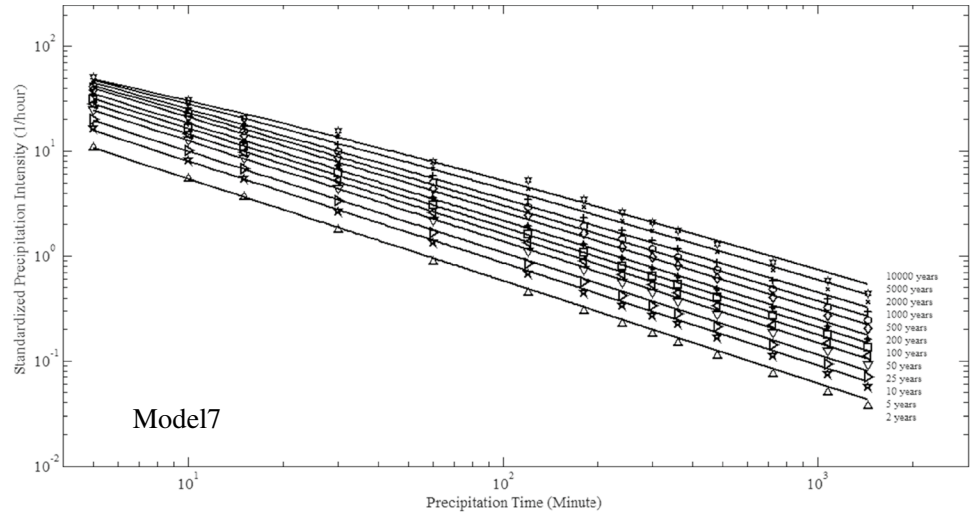
Fig. 7 Graphical presentation of the numerical values obtained by the L-Moments method and the curves yielded by the ABCP model with maximum depth of 3 for the IDF relationship for the Inland Anatolia Region for frequencies from 2 to 10,000 years



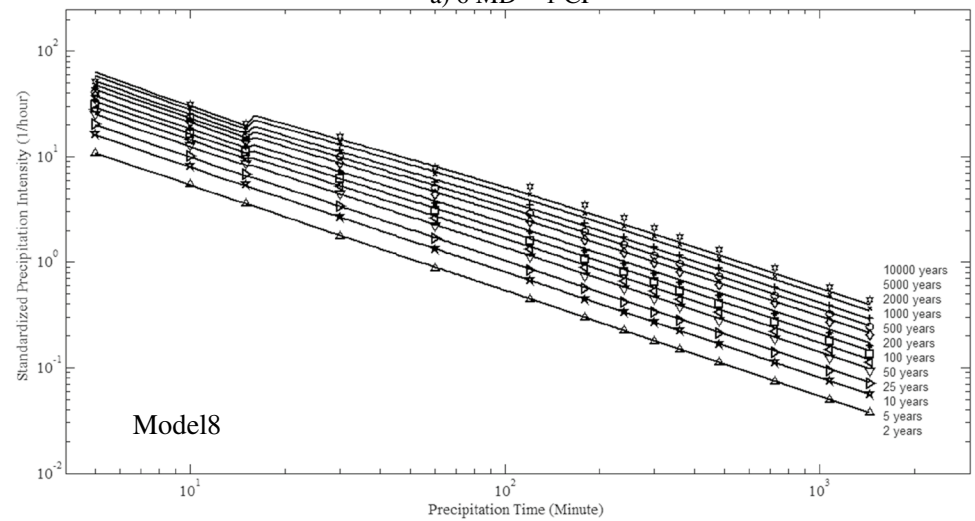
is always represented on graphs of logarithmic scales, the logarithmic error-based cost functions were considered in this study. While Eq. 21 was obtained by using a logarithmic function in Eq. 20, Eq. 22 provides a cost function that

was formed by using root mean squared logarithmic error (RMSLE). In the cost function expressed by Eq. 22, large errors assume greater weights than the small errors.

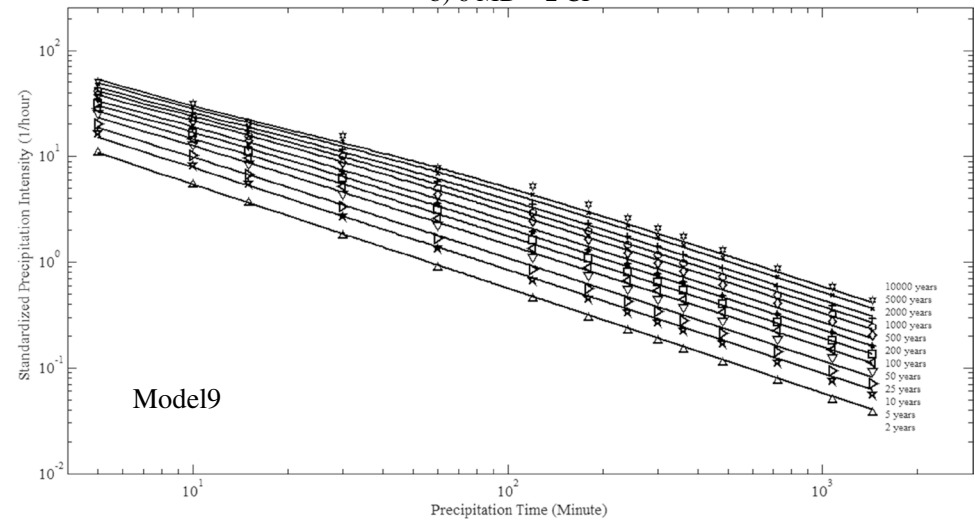
Fig. 8 Graphical presentation of the numerical values obtained by the L-Moments method and the curves yielded by the ABCP model, with maximum depth of 6 for the IDF relationship for the Inland Anatolia Region for frequencies from 2 to 10,000 years



a) 6 MD – 1 CF



b) 6 MD – 2 CF



c) 6 MD – 3 CF

Table 8 The p -values of the null hypothesis (H_0) of the Kruskal–Wallis test at 95% significance level and the decisions reached based on the magnitudes of the p -values

Equations by ABCP	p value	* H_0	Empirical models	p value	* H_0
Equation 11	0.0801	Reject	Model 1	0.1001	Reject
Equation 12	0.0918	Reject	Model 2	0.5577	Reject
Equation 13	0.0871	Reject	Model 3	0.8812	Reject
Equation 14	0.1349	Reject	Model 4	0.4504	Reject
Equation 15	0.1193	Reject	Model 5	0.9812	Reject
Equation 16	0.1349	Reject	Model 6	0.8883	Reject
Equation 17	0.2557	Reject	Model 7	0.8857	Reject
Equation 18	0.2557	Reject	Model 8	0.4213	Reject
Equation 19	0.2294	Reject	Model 9	0.8040	Reject

H_0 : There is a significant difference between predicted and measured values.

Parameter settings: As given in Table 3, the colony size was taken as 500 and a total of 500,000 cost function assessments were considered as the termination criterion of an ABCP run. The t of the terminal set denotes the rainfall duration in minutes and T designates the frequency in years. R is a generated random number uniformly distributed within the interval: $(-1, 1)$. The “log” of the function set is a protected natural logarithm function. The function returns zero if the value to be processed is equal to zero; otherwise, the natural logarithm of the absolute value of the number to be processed is returned. Similarly, the protected version of the “/” function was used. In this function, if the denominator is equal to zero, the function returns 1; otherwise, the division operation is performed. Our studies were conducted by ABCP for simpler relationships with smaller tree depths. Therefore, initial and maximum tree depths were taken as equal to each other and their values were set as 2, 3, and 6.

The MSE and R^2 error statistics were used to evaluate the accuracy of estimations by the ABCP approach, as done for the previously tried empirical equations. The mean, the standard deviation, the best, and the worst MSE error values obtained after 30 independent runs for each maximum depth of ABCP are given in Table 4. As seen in Table 4, the lowest MSE values for all depths were obtained by the first cost function. For structures with a maximum depth of 6, the MSE values of all three cost functions were close to each other. In experiments with the maximum depths of 2 and 3, the differences between the results of the 2nd and the 3rd cost functions and the 1st cost function increased. Generally, greater error values are observed in less deep trees than in deeper trees. However, the models with low depths may have simpler structures. The best models (solutions) obtained by ABCP for different maximum depths and different cost functions are presented in Table 5. The models given in Table 5 were obtained by ABCP. The models used

in this study, given in Table, were simplified in order for the reader to understand them more easily. As seen in Table 6, Model 1 and Model 3 are similar in structure and have different coefficients. The reason for the coefficients in these two models being different stems from the fact that the cost functions are different. For the Inland Anatolia Region, the coefficients of the equations of the nine ABCP models and the error statistics are given in Tables 6 and 7. As seen in Table 6, these nine ABCP models for the Inland Anatolia Region have a different number of coefficients.

The best models of each cost function and the MSE and R^2 values of these models are given in Table 6. Since rainfalls cannot be negative, in this study, the absolute values of the ABCP models were used when calculating the model outcomes. As seen from Tables 6 and 7, Model 7 of the 1st cost function, with a maximum depth of 6 yielded the lowest MSE (0.2129) and the greatest R^2 (0.9979) values. The graphical presentations of the IDF relationships obtained by the ABCP method by three different cost functions and three different maximum depths of all models are given in Figs. 6, 7, and 8. The outcomes of the ABCP model for three different cost functions with a maximum depth of 2 and 3 are presented in Figs. 6 and 7. As seen in Figs. 6 and 7, the IDF values obtained by the L-Moments method had a poor fit to those yielded by any of the ABCP models with a maximum depth of 2 and 3. The ABCP model outcomes for three different cost functions with a maximum depth of 6 are presented in Fig. 8. As seen in this figure, the IDF values of the L-Moments method overlapped with the IDF values computed by this version of the ABCP models. As observed in Figs. 6, 7, and 8, the ABCP models with a maximum depth of 6 were more successful than the ABCP models with 2 and 3 maximum depths. According to Fig. 8, the estimations by Model 7 of the 1st cost function were greater than the L-Moments values for small frequencies like 2 and 5 years. Model 7 also yielded small deviations for the 10,000-year frequency. In Model 8, high estimation outcomes resulted for 3 h and longer rainfall durations for 5000- and 10,000-year frequencies. The results of Model 9 of the third cost function deviated from the L-Moments values for mid-term frequencies. Besides the graphical presentations of the results of Models 7, 8, and 9, the MSE and R^2 values presented in Table 7 were investigated. As seen in Table 7, Model 7 of the third cost function, with a maximum depth of 6 had the highest R^2 and the lowest MSE values. Hence, it can be stated that Model 7 yielded outcomes closest to the target values among nine ABCP models. Yet, Models 8 and 9 did not have remarkably high MSE values; thus, they can be used as an alternative to Model 7.

In the last phase of the study, the best model obtained by the ABCP method was compared with the nine empirical equations considered herein. Among the nine empirical equations, Eq. 16 was determined to be the most appropriate

one. Among the nine different models developed as different versions of the ABCP method, Model 7 was identified as the best model. Despite high R^2 values by Eq. 16, it was visually observed in graphical form that that equation did not yield outcomes close to the actual values. When the results of Eq. 16 and Model 7 were compared with each other, it was observed that according to the IDF curves presented in Figs. 5 and 8, Model 7 produced by the ABCP method was more successful than empirical Eq. 16. Although the empirical equations available in literature were analytically simple expressions, they did not match the results of the L-Moments method, which is clearly noticeable to the naked eye when presented in graphical forms.

At the end of the study, in addition to the classical performance metrics, the nonparametric Kruskal – Wallis tests were also applied to all of the quantitative values given first by the regional frequency analyses by the L-Moments method and next by either those by the empirical equations or by those equations derived by the ABCP approach. The purpose of these tests was to check whether the predicted and actual values come from the same space (same population) (Başakın et al. 2020, 2021b; Özger et al 2020; Citakoglu 2021). The results of the Kruskal – Wallis tests are given in Table 8. As seen in Table 8, all models, the empirical equations, and the ABCP-produced ones passed the Kruskal – Wallis test at 95% significance level. Thus, the results of the empirical equations and the ABCP models come from the same population as those issued by the more accurate L-Moments approach. However, the ABCP method yields better accuracy than any of the empirical equations as summarized heretofore.

Conclusions

In this study, an automatic programming method based on artificial bee colony programming (ABCP) was used to find the optimal model to define the (standardized rainfall intensity) – (rainfall duration) – (frequency) (IDF) relationship obtained for the Inland Anatolia Region of Turkey by the L-Moments method. Furthermore, an analytical expression was developed, which is more accurate than any of the empirical equations available in literature for the same purpose. Initially, the IDF relationship was investigated using each one of nine different empirical formulas. The comparison studies on these nine empirical equations with optimized values for their coefficients revealed that the one expressed by Eq. 16 showed the relatively best performance for the Inland Anatolia Region, for representing the IDF relationship (Table 1). However, as shown in Fig. 4, Eq. 16 exhibited an unsatisfactory performance for frequencies smaller than 2000 years and a seemingly high coefficient of determination was misleading. Then, ABCP, with three different cost

functions and three different maximum depths, was used and new analytical expressions were generated for the present problem. The resultant best model obtained by ABCP was compared with the nine empirical formulas and the former was found to define the IDF relationship more accurately, as determined by the L-Moments method for the Inland Anatolia Region.

In analyses done by the ABCP method, by using three cost functions and three maximum depths, nine different models were produced for the relationship of rainfall intensity as the dependent variable and rainfall duration and frequency as the explanatory variables. The same comparison studies as the ones applied to those nine empirical equations revealed that Model 7, having the 1st cost function with a maximum depth of 6, yielded the lowest MSE and the highest R^2 values. Hence, the ABCP method, recommended as an alternative technique for determining a single analytical expression for the IDF relationship, yielded highly accurate outcomes for the Inland Anatolia Region and Model 7, with the smallest MSE was appreciably better than the best empirical model.

In conclusion, the ABCP-based models can be used elsewhere in the world and will most probably quantify the regional IDF relationship more accurately than any of the conventionally used empirical formulas for places without rainfall measurements or with gauging records that are too short.

Conflicts of interest

The authors declare no competing interests.

Ethics approval

The authors paid attention to ethical rules in the study. There was no violation of ethics.

Consent for publication

If this study is accepted, it can be published in the Arabian Journal of Geosciences.

Data availability Climatic data and hydrometric data were provided by the General Directorate of State Meteorological Affairs (MGM) and General Directorate of State Hydraulics Works (DSİ).

Code availability The codes were written and the analyses were done by Beyza Gorkemli. Licensed Matlab provided by Erciyes University was used for some analyses.

References

- Agarwal SK, Yadav SA Comprehensive Survey on Artificial Bee Colony Algorithm as a Frontier in Swarm Intelligence. In: Hu YC., Tiwari S., Mishra K., Trivedi M. (eds) *Ambient Communications and Computer Systems. Advances in Intelligent Systems and Computing*, vol 904. Springer, Singapore. https://doi.org/10.1007/978-981-13-5934-7_12, 2019.
- Akinsanola AA, Ogunjobi KO (2014) Analysis of rainfall and temperature variability over Nigeria. *Glob J Human Soc Sci: Geography Environ GeoSci* 14(3):10–28
- Al-anazi K, El-sebaie I (2013) Development of Intensity-Duration-Frequency Relationships for Abha City in Saudi Arabia. *Int J Comput Eng Res* 3(10):58–65
- AlHassoun SA (2011) Developing an empirical formulae to estimate rainfall intensity in Riyadh region. *J King Saud Univ Eng Sci* 23:81–88
- Antigha RE, Ogarekpe NM (2013) Development of Intensity Duration Frequency Curves for Calabar Metropolis, South-South, Nigeria. *Int J Eng Sci (THE IJES)* 2(3):39–42
- Arslan S, Ozturk C, Artificial Bee Colony Programming Descriptor for Multi-Class Texture Classification. *Appl Sci*. 9:9, Article number: 1930, 2019.
- Arslan S, Ozturk C (2018) Artificial Bee Colony Programming for Feature Selected Cancer Data Classification. *Int J Sci Technol Res* 4(7):75–84
- Arslan S, Ozturk C (2019) Multi Hive Artificial Bee Colony Programming for high dimensional symbolic regression with feature selection. *Appl Soft Comput* 78:515–527
- Asian Pacific Friend. Rainfall Intensity Duration Frequency (IDF) Analysis for the Asia Pacific Region. HP-VII | Technical Documents in Hydrology | No. 2 Regional Steering Committee for Southeast Asia and the Pacific UNESCO Office, Jakarta, 2008.
- Asikoglu OL, Benzedem E (2014) Simple generalization approach for intensity-duration-frequency relationships. *Hydrol Process* 28:1114–1123
- Bansal JC, Sharma H, Jadon SS (2013) Artificial bee colony algorithm: a survey. *Int J Adv Intell Paradigms* 5(1–2):123–159
- Bartual RG, Schneider M (2001) Estimating Maximum Expected Short-Duration Rainfall Intensities from Extreme Convective Storms. *Phys Chem Earth Part B* 26:675–681
- Başakın E, Ekmekcioğlu Ö, Özger M, Çelik A (2020) Prediction of Turkey Wheat Yield by Wavelet Fuzzy Time Series and Gray Prediction Methods. *Turk J Agric Res* 7(3):246–252. <https://doi.org/10.19159/tutad.685342>
- Başakın EE, Ekmekcioğlu Ö, Özger M, & Citakoglu H (2021a). Determination of Intensity-Duration-Frequency Relation by Particle Swarm Optimization and Genetic Programming. In II. International Applied Statistics Conference (UYIK-2021), Tokat, Turkey (pp. 1–8).
- Başakın EE, Ekmekcioğlu Ö, Çitakoğlu H, & Özger M (2021b). A new insight to the wind speed forecasting: robust multi-stage ensemble soft computing approach based on pre-processing uncertainty assessment. *Neural Computing and Applications*, 1–30.
- Bernard, M. M. Formulas for Rainfall Intensities of Long Duration, 96: 592–624, 1932.
- Boudardara F, Gorkemli B. 2018 Application of artificial bee colony programming to two trails of the artificial ant problem, in: Proceedings of the 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT 2018), Ankara, Turkey, Oct 19–21, <https://doi.org/10.1109/ISMSIT.2018.8567048>
- Boudardara F, Gorkemli B (2020) Solving artificial ant problem using two artificial bee colony programming versions. *Appl Intell*. <https://doi.org/10.1007/s10489-020-01741-0>
- Boudouaoui Y, Habbi H Scaled artificial bee colony programming, in: 2018 International Conference on Applied Smart Systems (ICASS), Medea, Algeria, Nov 24–25, 2018.
- Bougadis J, Adamowski K (2006) Scaling model of a rainfall intensity-duration-frequency relationship. *Hydrol Process* 20:3747–3757
- Chang KB, Lai SH, Faridah O (2013) Rain IDF: automated derivation of rainfall intensity-duration-frequency relationship from annual maxima and partial duration series. *J Hydroinf* 15(4):1224–1233
- Chen CI (1983) Rainfall Intensity-Duration-Frequency Formulas. *J Hydraul Eng* 109(12):1603–1621
- Chow VT. *Handbook of Applied Hydrology*. McGraw-Hill Book, 1988.
- Citakoglu H (2021) Comparison of multiple learning artificial intelligence models for estimation of long-term monthly temperatures in Turkey. *Arab J Geosci* 14:2131. <https://doi.org/10.1007/s12517-021-08484-3>
- Dourte D, Shukla S, Singh P, Haman D (2013) Rainfall Intensity-Duration-Frequency Relationships for Andhra Pradesh, India: Changing Rainfall Patterns and Implications for Runoff and Groundwater Recharge. *J Hydrol Eng* 18(3):324–330
- Dupont BS, Allen DL 2006 Establishment of Intensity-Duration-Frequency Curves for Rainfall in the Monsoon Area of Vietnam. Kentucky Transportation Center, College of Engineer, University of Kentucky in corporation with US Department of Transportation
- El-Sayed EAH (2011) Generation of rainfall intensity duration frequency curves for ungauged sites. *Nile Basin Water Sci Eng J* 4(1):112–124
- Froehlich DC (1995) Long-Duration-Rainfall Intensity Equations. *J Irrig Drain Eng* 121:248–252
- García-Marín AP, Ayuso-Muñoz JL, Jiménez-Hornero FJ, Estévez J (2013) Selecting the best IDF model by using the multi fractal approach. *Hydrol Processes* 27:433–443
- García-Marín AP, Ayuso-Muñoz JL, Jiménez-Hornero FJ, and Estévez J. Selecting the best IDF model by using the multifractal approach. *Hydrological Processes*, 27, 433–443, 2011)
- Golafshani EM, Ashour A (2016) Prediction of self-compacting concrete elastic modulus using two symbolic regression techniques. *Autom Constr* 64:7–19
- Golafshani EM, Behnood A (2018) Automatic regression methods for formulation of elastic modulus of recycled aggregate concrete. *Appl Soft Comput* 64:377–400
- Gorkemli B, Karaboga D (2019) A quick semantic artificial bee colony programming (qsABCP) for symbolic regression. *Inf Sci* 502:346–362
- Greenwood JA, Landwehr JM, Matalas NC, Wallis JR (1979) Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour Res* 15(5):1049–1054
- Haktanir T, Cobaner M, Kisi O (2010) Frequency analyses of annual extreme rainfall series from 5 minutes to 24 hours. *Hydrol Process* 24:3574–3588
- Haktanir T, Citakoglu H, Seckin N (2016) Regional frequency analysis of successive-duration annual maximum rainfalls by L-Moments method. *Hydrol Sci J* 61(4):647–668
- Hamaamin YA (2017) Developing of rainfall intensity-duration-frequency model for sulaimani city. *J Homepage* 19(3–4):93–102
- Hara A, Kushida J, Takemoto R, Takahama T 2018 Artificial Bee Colony Programming Using Semantic Control Crossover, in: 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, Oct 07–10, 2018, <https://doi.org/10.1109/SMC.2018.00043>
- Hershfield DM 1961 Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years, United States Weather Bureau Technical Paper, 40, Washington DC

- Hosking JRM (1990) L-Moments: Analysis and estimation of distributions using linear combinations of order statistics. *J R Stat Soc Ser B* 52(2):105–124
- Hosking JRM. and Wallis JR 1997 Regional Frequency Analysis. Cambridge University Pres.,
- Jaleel LA, Farawn MA (2013) Developing Rainfall Intensity-Duration-Frequency Relationship for Basrah City. *Kufa J Eng* 5(1):105–112
- Karaboga D (2005) An Idea Based on Honey Bee Swarm for Numerical Optimization, Technical Report–TR06. Erciyes University, Engineering Faculty, Computer Engineering Department, Turkey
- Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J Global Optim* 39(3):459–171
- Karaboga D, Ozturk C, Karaboga N, Gorkemli B (2012) Artificial bee colony programming for symbolic regression. *Inf Sci* 209:1–15
- Karaboga D, Gorkemli B, Ozturk C, Karaboga N (2014) A comprehensive survey: artificial bee colony (ABC) algorithm and applications. *Artif Intell Rev* 42(1):21–57
- Karaboga D. 2010 Artificial bee colony algorithm, www.scholarpedia.org/article/Artificial_bee_colony_algorithm. *Scholarpedia*, 5(3), 6915–6915
- Karahan H, Ceylan H, Ayvaz MT (2007) Predicting rainfall intensity using a genetic algorithm approach. *Hydrol Process* 21:470–475
- Karahan H, Ayvaz MT, Gurarslan G (2008) Determination of Intensity-Duration-Frequency Relationship by Genetic Algorithm: Case Study of GAP. *Teknik Dergi* 19(2):4393–4407
- Koutsoyiannis D (1994) A Stochastic Disaggregation Method for Design Storm and Flood Synthesis. *J Hydrol* 156:193–225
- Koutsoyiannis D, Kozonis D, Manetas A (1998) A mathematical framework for studying rainfall intensity–duration–frequency relationships. *J Hydrol* 206:118–135
- Koza JR. 1992 Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge, MA, USA, 819p
- Liuzzo L, Freni G (2015) Analysis of Extreme Rainfall Trends in Sicily for the Evaluation of Depth–Duration–Frequency Curves in Climate Change Scenarios. *J Hydrol Eng* 20(12):04015036
- Lopcu Y 2007 Modeling the Intensity–Duration–Frequency Relationships of Annual Maximum Storms. MSc Thesis, in Civil Engineering, Dokuz Eylul University, Izmir, Turkey
- Nhat LM, Tachikawa Y, Takara K (2006) Establishment of intensity-duration-frequency curves for precipitation in the monsoon area of Vietnam. *Annals of Disas Prev Res Inst* 93–103
- Nhat LM, Tachikawa Y, Sayama T, Takara K 2007 Regional rainfall intensity–duration–frequency relationships for ungauged catchments based on scaling properties. *Annals of Disas. Prev. Res. Inst., Kyoto Univ., No. 50 B*
- Olofintoye OO, Sule BF, Salami AW (2009) Best-fit Probability distribution model for peak daily rainfall of selected Cities in Nigeria. *New York Sci J* 2(3):1–12
- Omotosho TV, and Oluwafemi CO 2009 One-minute rain rate distribution in Nigeria derived from TRMM satellite data. *Journal of Atmospheric and Solar–Terrestrial Physics*, Vol. 71, No. 5, pp. 625–633
- Özger M, Başakın EE, Ekmekcioğlu Ö, Hacısüleyman V (2020) Comparison of wavelet and empirical mode decomposition hybrid models in drought prediction. *Comput Electron Agric* 179:105851. <https://doi.org/10.1016/j.compag.2020.105851>
- Pooja, Shirmal G. 2020 Innovative Review on Artificial Bee Colony Algorithm and Its Variants. In: Sharma H., Govindan K., Poonia R., Kumar S., El-Medany W. (eds) *Advances in Computing and Intelligent Systems. Algorithms for Intelligent Systems*. Springer, Singapore. https://doi.org/10.1007/978-981-15-0222-4_14
- Rasel M, Islam M (2015) Generation of rainfall intensity-duration-frequency relationship for north-western region in Bangladesh. *J Environ Sci, Toxicol Food Technol (IOSR–JESTFT)* 9(9):41–47
- Sherman CW (1931) Frequency and Intensity of Excessive Rainfall At Boston. *Mass Trans Pap* 95:951–960
- Trevor D, and Guillermo Q (2008) Rainfall intensity duration frequency (IDF) analysis for Asia Pacific Region. International Hydrological Programme, IHP-VII
- Vijayagopal P, Vivekanandan N, Kannan S (2013) Assessing adequacy of probability distribution for development of IDF relationships for Mandla and Jabalpur. *Int J Sci Res Rev* 2(3):99–114
- Vivekanandan N (2012) Probabilistic Modelling of Hourly Rainfall Data for Development of Intensity-Duration-Frequency Relationships. *Bonfring Int J Data Min* 2(4):16
- Vivekanandan N (2013) Analysis of hourly rainfall data for the development of IDF relationships using the order statistics approach of probability distributions. *Int J Manag Sci Eng Manag* 8(4):283–291
- Wang X, Huang G, Liu J (2014) Projected increases in intensity and frequency of rainfall extremes through a regional climate modeling approach. *J Geophys Res Atmos* 119(23):13271–13286
- Yu PS, Yang TC, Lin CS (2004) Regional rainfall intensity formulas based on scaling property of rainfall. *J Hydrol* 295(1–4):108–123