

Investigation of Applicability of Artificial Bee Colony Algorithm on Rainfall Intensity Duration Frequency Equations

Utku Zeybekoglu^{1*}, Asli Ulke Keskin²

¹ Construction Department, Vocational School of Higher Education, Sinop University, 57200, Boyabat, Sinop, Turkey

² Civil Engineering Department, Faculty of Engineering, Ondokuz Mayıs University, 55270, Atakum, Samsun, Turkey

* Corresponding author's e-mail: utkuz@sinop.edu.tr

ABSTRACT

The main objective of this study was to investigate the applicability and efficiency of an artificial bee colony optimization algorithm to determine two statistical-based rainfall intensity duration frequency equations' weighting parameters. For this aim, the annual maximum rainfall records were obtained from seven meteorological stations of seven geographic regions in Turkey. It was observed that the Artificial Bee Colony algorithm, which is an alternative technique for solving the rainfall intensity duration frequency equations, gives very good results in selected seven meteorological stations.

Keywords: rainfall intensity, artificial bee colony algorithm, optimization, heuristic optimization methods.

INTRODUCTION

Nowadays, extreme events such as floods and droughts occur more frequently due to the spatial and temporal changes of precipitation regimes as a result of global warming. Therefore, the efficient use of water resources becomes increasingly important. The main input of water resources is rainfall; knowing such properties as amount, duration, intensity, spatial and temporal change is required for planning, design, construction, and operation studies of different sectors such as water resources, agriculture, urbanization, drainage and flood control (Karahana et al., 2008).

The relationship between the duration, intensity, and recurrence of rainfall is called the Intensity Duration Frequency (RIDF) relationship. This relation can be expressed in a mathematical or statistical form, as a function of the parameters that change according to the location and time, called rainfall duration (t) and recurrence period (T) and weight parameters (w_1, \dots, w_n) (Chen, 1983; Koutsoyiannis et al., 1998; Karahana et al., 2007; Karahana, 2019). The 5-parameter RIDF

equations based on Gumbel and Exponential distribution are widely used (Koutsoyiannis et al., 1998). These equations are given in Eq. (1) and Eq. (2), respectively.

$$I = w_1 \frac{w_2 - \ln \left[-\ln \left(1 - \frac{1}{T} \right) \right]}{(w_3 + t^{w_4})^{w_5}} \quad (1)$$

$$I = \frac{w_1(w_2 + \ln(T))}{(w_3 + t^{w_4})^{w_5}} \quad (2)$$

Several empirical and numerical methods, especially heuristic optimization algorithms (Genetic Algorithm, Particle Swarm Optimization, Differential Evolution Algorithm), have been developed and used widely to estimate the rainfall intensities and RIDF equation parameters (Aron et al., 1987; Garcia-Bartual and Schneider, 2001; Karahana et al., 2007; Raiford et al., 2007; Karahana et al., 2008; Karahana, 2012; Elsebaie, 2012; Chebbi et al., 2013; Awadallah, 2013; Manzano-Agugliaro et al., 2014; Guclu et al., 2016; Noor et al., 2018; Karahana 2019).

This study, is aimed at determining the parameters of the RIDF equations by the Artificial Bee Colony (ABC) algorithm and compare their performance. Moreover, the ABC algorithm was applied for the first time to the RIDF equations.

STUDY AREA AND METHODOLOGY

Study area

The Annual Maximum Rainfall (AMR) series was used from seven meteorological stations in the seven geographical regions from Turkey. The recorded data were aggregated based on the duration of 5, 10, 15, to 1440 min. The AMR data were obtained from the Turkish State Meteorological Service. The geographical locations of the stations are shown in Figure 1 and other features are given in Table 1.

Artificial bee colony algorithm

One of the heuristic optimization algorithms is the Artificial Bee Colony Algorithm, inspired by the real nectar search behavior of bees developed by Karaboga (2005). The algorithm was developed by observing the behavior of real bees

while finding nectar sources and sharing the amount of this source with other bees in the hive. The basic ABC algorithm has only three control parameters: the number of food sources (CS), the maximum cycle number (MCN), and the parameter “limit.” If the total number of successive unsuccessful trials exceeds the “limit” parameter, this source is assumed to be exhausted. The effort to find the best resource in the algorithm continues in every cycle. The algorithm ends when it reaches the maximum cycle number. Detailed information on the ABC algorithm is given in Karaboga 2005; Karaboga and Basturk, 2007; Karaboga and Akay, 2011; Ozturk et al., 2014; Akay and Karaboga, 2015.

RESULTS AND DISCUSSION

This study aimed at determining the coefficients of rainfall-intensity-duration-frequency equations based on the Gumbel and Exponential distribution of the stations selected using the artificial bee colony algorithm. In order to create optimization model inputs, a frequency analysis is required first. The empirical return period T_{iq} for each rainfall intensity values can be obtained from the raw data by using the Gringorten formula Eq. (3) (Cunnane; 1978; Karahan et al., 2007).

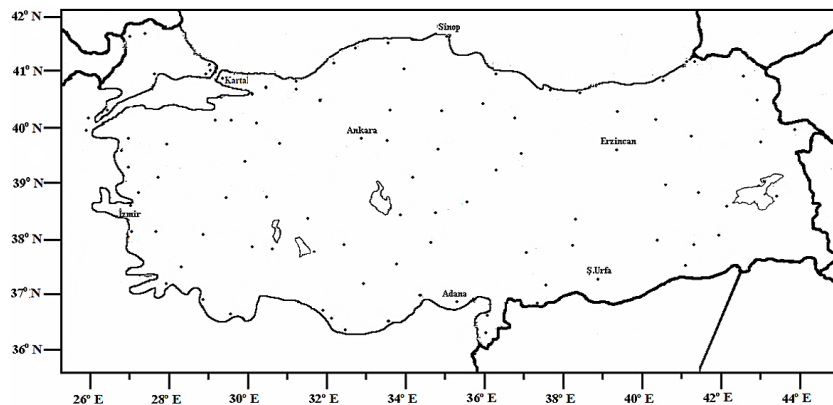


Figure 1. Geographical distribution of meteorological stations

Table 1. The detailed information of meteorology stations in the study

Station	Obs. Period	Altitude(m)	Latitude(N)	Longitude(E)	Geographical Region
Adana	1944–2015	23	37.0041	35.3443	Mediterranean
Ankara	1940–2015	891	39.9727	32.8637	Inner Anatolia
Erzinca	1957–2015	1216	39.7523	39.4868	Eastern Anatolia
Izmir	1938–2015	29	38.3949	27.0819	Aegean
Kartal (Istanbul)	1974–2015	18	40.9113	29.1558	Marmara
Sinop	1965–2015	32	42.0299	35.1545	Black Sea
Sanliurfa	1959–2015	550	37.1608	38.7863	Southeastern Anatolia

$$T_{jq} = \frac{m_j + 0.12}{L - 0.44} j = 1, 2, 3, \dots, 14; q = 1, 2, 3, \dots, L \quad (3)$$

As a result of the frequency analysis of the stations in the study area, a matrix consisting of three columns with the return period (*T*), duration of rainfall (*t*) and rainfall intensity (*l*; *mm/min*).

The weight coefficients of the RIDF equations were determined by running the ABC algorithm until the maximum cycle number was reached, minimizing the Mean Square Error (MSE) selected as the fitness function. The parameters were used in the optimization process with ABC in Table 2.

The weighting coefficients of RIDF equations were determined by means of the ABC algorithm. Coefficients and performances of each RIDF equation were given in tables 3–4, and tables 5–6, respectively. In the related tables, MSE is the mean square error; MAE is the mean absolute errors; R represents the correlation coefficient, E indicates the model efficiency coefficient and *t* denotes the duration by seconds.

Table 2. Parameters of ABC Algorithm used in this study

Number of food sources (CS)	P*10
Limit	0.5*CS*P
Maximum cycle number (MCN)	2000

As can be seen from Tables 5 and 6, the Gumbel distribution based RIDF equations shows the best performances for all meteorological stations according to the fitness function. It was determined that this situation was the same for other evaluation criteria (MAE, R and E) except time (*t*).

The results of the Exponential based equations fitness function ranged from 0.001004 to 0.008463 and the average value was 0.003062. The ranges of the other criteria were found from 0.013246 to 0.038782 for MAE, from 0.988332 to 0.995946 for R and from 0.91348 to 0.941784 for E.

The results for the Gumbel distribution-based equations gave better results than the Exponential based equations. MSE, MAE, R, and E were found for the Gumbel distribution-based equation ranges from 0.000357 to 0.004172, from 0.010559 to 0.028503, and from 0.993346 to 0.996996 and from 0.927013 to 0.951115, respectively.

Scatter diagrams between the observed values and calculated values of the Exponential and Gumbel based RIDF equation by parameters for each station are presented in Figure 2 and Figure 3, respectively.

It is seen that the scatter diagrams of stations prepared using equations based on the Gumbel distribution are more compatible than the diagrams of the Exponential equations. Therefore, the statistical results are determined to be compatible with the graphical results.

Table 3. The weight coefficients of the equations based on the Exponential distribution

Station	w ₁	w ₂	w ₃	w ₄	w ₅
Adana	6.487564	1.298562	9.244116	0.979849	0.818207
Ankara	4.49454	0.888162	3.065257	0.830196	1
Erzincan	1.760129	1.164605	2.944518	1	0.74789
Izmir	3.112518	1.472447	2.191233	0.745064	0.947887
Kartal (Istanbul)	6.037774	0.683225	6.755473	0.991928	0.792676
Sinop	6.339297	0.895435	4.167488	0.812492	0.999137
Sanliurfa	2.639467	0.992535	3.75791	1	0.738918

Table 4. The weight coefficients of the equations based on the Gumbel distribution

Station	w ₁	w ₂	w ₃	w ₄	w ₅
Adana	6.013926	2.496147	4.974424	0.74603	1.12078
Ankara	3.689889	1.95545	3.001928	0.780893	1.092558
Erzincan	1.30138	2.20998	3.507338	1.065964	0.696794
Izmir	3.267853	2.559104	1.446883	0.504307	1.475799
Kartal (Istanbul)	4.347613	1.702137	5	0.917564	0.858455
Sinop	5.7652	2.01151	2.361319	0.577106	1.443282
Sanliurfa	1.8954	1.975675	5	1.103499	0.660307

Table 5. Performances for the equations based on the Exponential distribution

Station	t	MSE	MAE	R	E
Adana	692.7882	0.008463	0.038782	0.988478	0.91348
Ankara	500.1653	0.004394	0.020364	0.988332	0.929473
Erzincan	422.7719	0.001004	0.013246	0.992118	0.918738
Izmir	536.6802	0.003063	0.022143	0.994012	0.939241
Kartal	270.1458	0.0014	0.01474	0.995946	0.939091
Sinop	349.2746	0.001935	0.017149	0.99569	0.941784
Sanliurfa	402.1657	0.001178	0.013418	0.994607	0.936797
Minimum	270.1458	0.001004	0.013246	0.988332	0.91348
Maximum	692.7882	0.008463	0.038782	0.995946	0.941784
Average	453.4274	0.003062	0.019978	0.992741	0.931229

Table 6. Performances for the equations based on the Gumbel distribution

Station	t	MSE	MAE	R	E
Adana	838.1269	0.004172	0.028503	0.994085	0.936413
Ankara	613.2153	0.00228	0.015851	0.993346	0.945105
Erzincan	546.1715	0.000357	0.010559	0.996996	0.935225
Izmir	750.1329	0.001998	0.017816	0.996094	0.951115
Kartal	313.6158	0.001712	0.017663	0.995213	0.927013
Sinop	415.2171	0.001351	0.017277	0.996886	0.941351
Sanliurfa	475.1302	0.001175	0.011719	0.995203	0.9448
Minimum	313.6158	0.000357	0.010559	0.993346	0.927013
Maximum	838.1269	0.004172	0.028503	0.996996	0.951115
Average	564.5157	0.001864	0.017055	0.995403	0.940146

CONCLUSION

In this study, a solution algorithm to determine the rainfall intensity duration frequency relationship using the Artificial Bee Colony optimization algorithm was applied to 7 meteorological stations located in 7 different regions in Turkey. The rainfall intensity duration frequency relationship was evaluated using statistical formulations with 5 parameters based on the exponential distribution and Gumbel distribution. The mean square error (MSE), which identified the relationship between the observed and calculated rainfall intensities in the optimization process, was defined as the objective function, and the authors aimed to minimize this value in 2000 iteration steps.

The weight coefficients of the RIDF equations were determined and it was found out that the equation based on the Gumbel distribution was more suitable, even though they have the same number of parameters. This situation can be clearly seen from the scatter diagrams of the relevant stations. The average MSE value for the Gumbel distribution is 0.001864 and this value

for Exponential is 0.003062. Other evaluation criteria called MAE, R and E were calculated as 0.017055, 0.995403 and 0.940146 for the Gumbel distribution, respectively. On the other hand, these criteria were found 0.019978, 0.992741 and 0.931229 for the Exponential distribution, respectively. When evaluated in terms of processing time (t), it was determined that the equations based on the low-performing the Exponential distribution gave quicker results. This situation is thought to stem from the structure of the equation. The results showed that an artificial bee colony optimization algorithm can be used successfully in the RIDF problem.

In further studies, it is planned to apply various rainfall intensity duration frequency relationships in different geographic regions by using different heuristic optimization algorithms. With these studies, it will be possible to compare the performances of heuristic optimization algorithms on the RIDF problem. At the same time, the RIDF equations that best adapt to the study areas will be determined.

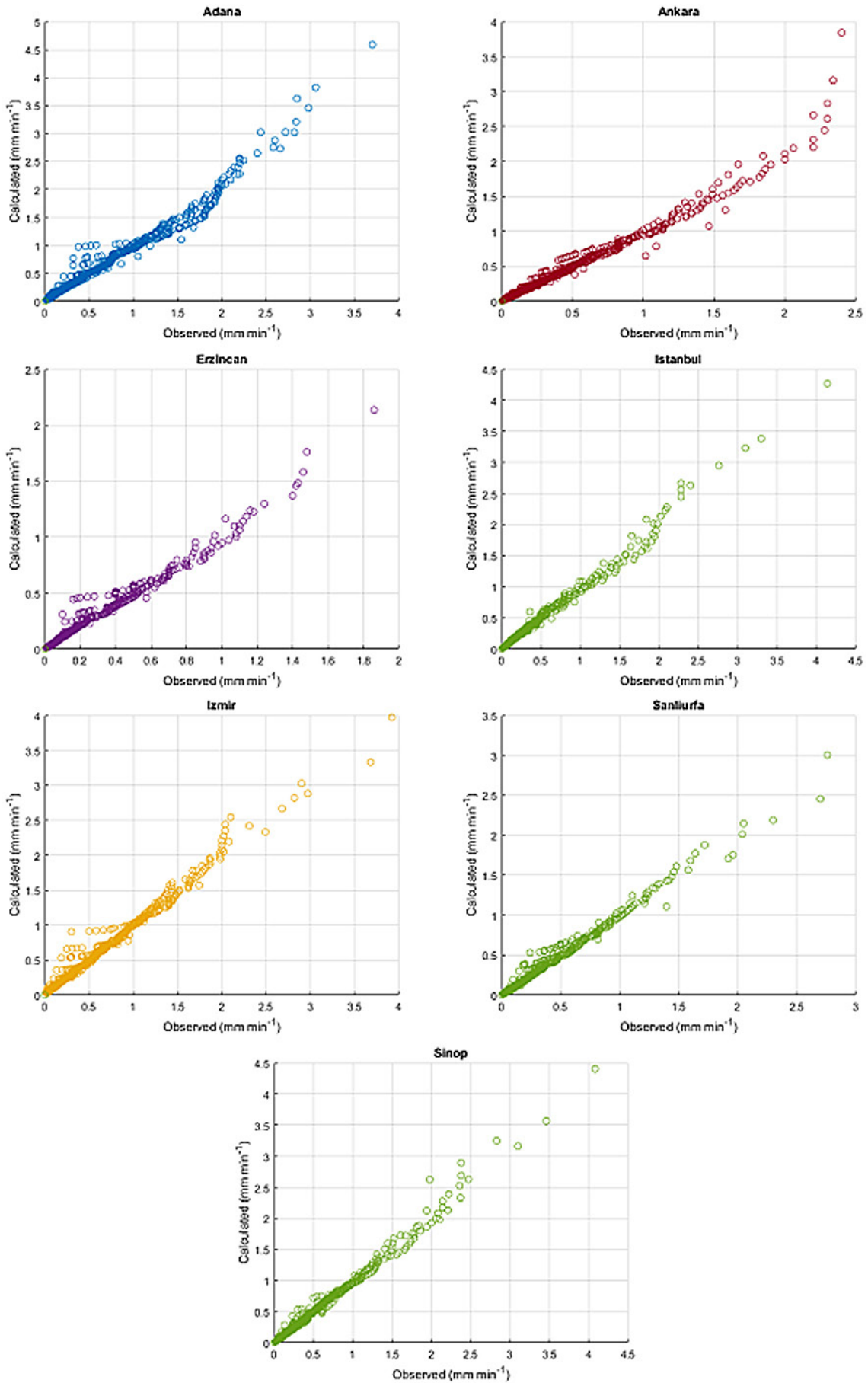


Figure 2. Scatter diagram of results of the Exponential distribution based equations

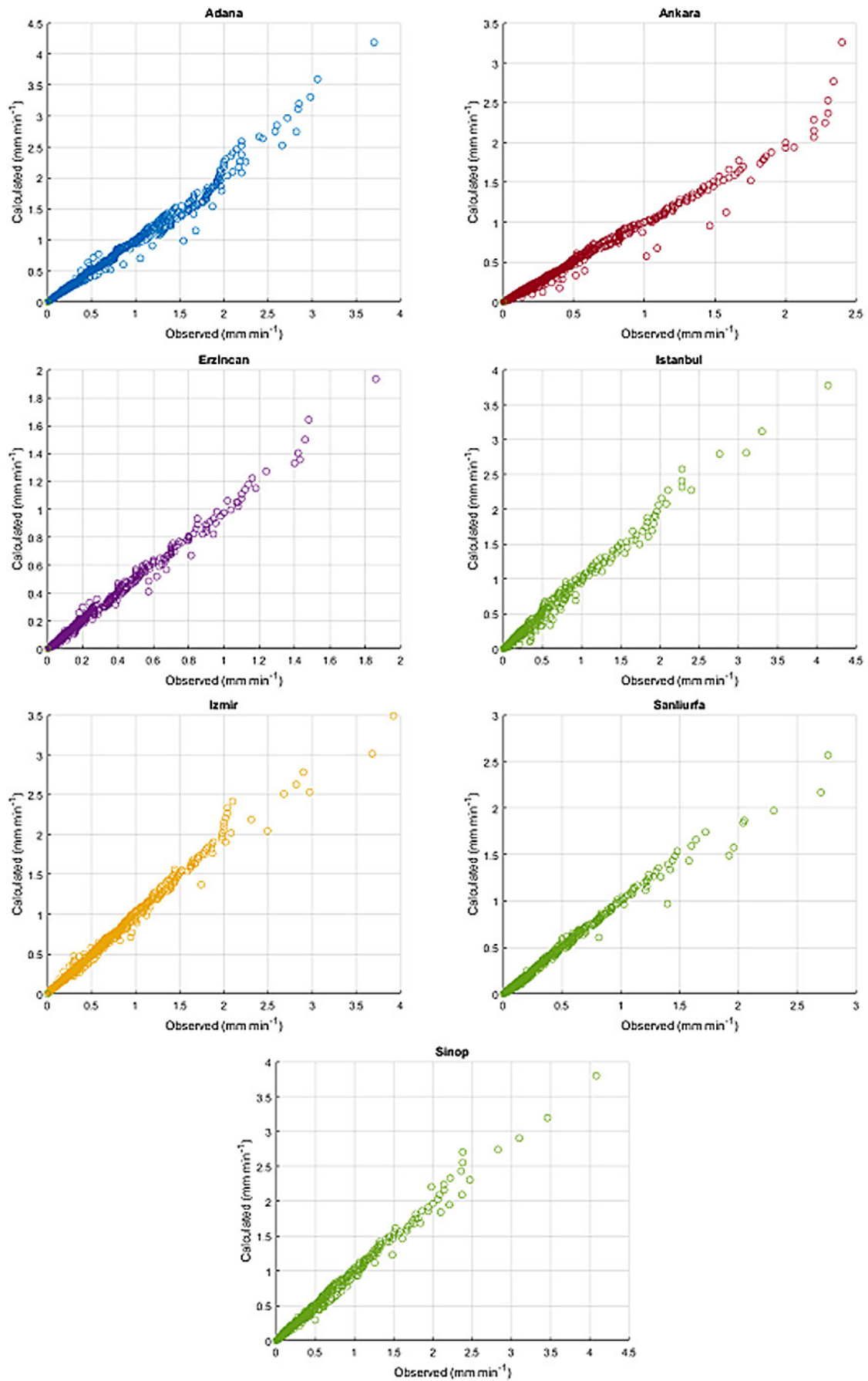


Figure 3. Scatter diagram of results of the Gumbel distribution based equations

REFERENCES

1. Akay B., Karaboga D. 2015. A survey on the applications of Artificial Bee Colony in signal image and video processing. *Signal, Image And Video Processing*, 9(4), 967–990.
2. Aron G., Wall D.J., White E.L., Dunn C.N. 1987. Regional rainfall intensity–duration–frequency curves for Pennsylvania. *Water Resour. Bull.*, 23(3), 479–485.
3. Awadallah A.G. 2013. regional intensity–duration–frequency curves for Jeddah region, Saudi Arabia using ordinary and l-moments approaches. *J. Flood Risk Manage.*, 8(3), 195–207.
4. Chebbi A., Bargaoui Z.K., Da Conceição Cunha M. 2013. Development of a method of robust rain gauge network optimization based on intensity-duration-frequency results. *Hydrol. Earth Syst. Sci.*, 17, 4259–4268
5. Chen C.I. 1983. Rainfall intensity-duration-frequency formulas. *J. Hydraul. Eng.*, 109(12), 1603–1621.
6. Cunnane C. 1978. Unbiased plotting positions: A review. *Journal of Hydrology*, 37, 205–222.
7. Elsebaie I.H. 2012. Developing rainfall intensity–duration–frequency relationship for two regions in Saudi Arabia. *Journal Of King Saud University-Engineering Sciences*, 24(2), 131–140.
8. Garcia-Bartual R., Schneider M. 2001. Estimating maximum expected short-duration rainfall intensities from extreme convective storms. *Physics And Chemistry Of The Earth, Part B: Hydrology, Oceans, And Atmosphere*, 26(9), 675–681.
9. Guclu Y.S., Sisman E., Yelegen M.O. 2016. Climate change and frequency–intensity–duration (FID) curves for Florya station, Istanbul. *Journal Of Flood Risk Management*, 11(1), 403–418.
10. Karaboga D. 2005. An idea based on honey bee swarm for numerical optimization. Technical Report Tr06, Erciyes University, Engineering Faculty, Computer Engineering Department.
11. Karaboga D., Akay B. 2011. A Modified Artificial Bee Colony (ABC) algorithm for constrained optimization problems. *Applied Soft Computing*, 11(3), 3021–3031.
12. Karaboga D., Basturk B. 2007. A powerful and efficient algorithm for numerical function optimization: Artificial Bee Colony (ABC) algorithm. *Journal of Global Optimization* 39(3), 459–471.
13. Karahan H. 2012. Determining rainfall intensity duration frequency relationship using Particle Swarm Optimization. *Ksce Journal Of Civil Engineering*, 16(4), 667–675.
14. Karahan H. 2019. Determination of homogeneous sub-regions by using intensity-duration-frequency relationships and cluster analysis: An application for the Aegean region. *Pamukkale University Journal Of Engineering Sciences*, 25(8), 998–1013.
15. Karahan H., Ayvaz M.T., Gurarslan G. 2008. determination of intensity duration frequency relationship by Genetic Algorithm: Case study of Gap. *IMO Technical Journal*, 19(2), 4393–4407.
16. Karahan H., Ceylan H., Ayvaz M.T. 2007. Predicting rainfall intensity using a Genetic Algorithm approach. *Hydrological Processes*, 21, 470–475.
17. Koutsoyiannis D., Kozonis D., Manetas A. 1998. A mathematical framework for studying rainfall intensity duration frequency relationships. *Journal Of Hydrology*, 206(1–2), 118–135.
18. Manzano-Agugliaro F., Zapata-Sierra A., Rubí J.F., Hernández-Escobedo Q. 2014. Assessment of methods to obtain IDF curves for Mexico. *Water Technology And Sciences*, 5(3), 149–158.
19. Noor M., Ismail T., Chung E.S., Shahid S., Sung J.H. 2018. Uncertainty in rainfall intensity duration frequency curves of peninsular Malaysia under changing climate scenarios. *Water*, 10, 1750–1774.
20. Ozturk C., Hancer E., Karaboga D. 2014. Improved clustering criterion for image clustering with Artificial Bee Colony Algorithm. *Pattern Anal. Appl.*, 18(3), 587–599.
21. Raiford J.P., Aziz N.M., Khan A.A., Powell D.N. 2007. Rainfall depth-duration-frequency relationships for South Carolina, North Carolina, and Georgia. *American Journal Of Environmental Science*, 3(2), 78–84.