Analyzing trend and developing an empirical formula for estimating rainfall intensity: a case study Eskisehir city, Turkey.

Sinan Jasim Hadi[1,*], and *Mustafa* TOMBUL[1]

[1]Department of Civil Engineering, Faculty of Engineering, Anadolu University Eskisehir, Turkey

Abstract. Trend and stationarity analysis conducted for the city center of Eskisehir province, Turkey. The Intensity- Duration – Frequency (IDF) curves and formulas for the estimation of the design rainfall intensity have been developed. Through the application of four different frequency methods, the IDF has been developed for a record of 64 years. The functions employed are; Normal, Log – Normal, Gumbel, and Log –Pearson Type III. A formula for each function developed by determining the parameters with the help of a non-linear regression method. The most appropriate fit of probability distribution determined by goodness of fit test i.e. Chi – Square. The results of the study revealed no evidence of significant trend or nonstationarity in any duration. With an exception to the Normal function which exceeded the critical region in several durations, all the functions found to show high fitting. LPT III identified as the best fit. The coefficient of correlation for the developed formulas found high in all the functions except for LPT III. The study endorsed that the estimation of rainfall intensity for any storm duration and return period can be done for Eskisehir using LPT III and Gumbel formulas as they are evaluated as best fitting.

1 Introduction

The frequency and intensity of extreme rainfall events are expected to increase in the future for the regions that already have intense and frequent events, whereas for the regions having less events are expected to witnessed clear decreases in the amount of rainfall [1]. Stationarity and trend analyses are considered as good tools for investigating the historical change in the time series data especially in hydrological observations. Stationarity has been taken into consideration in many studies as a result of it being one of the main assumptions of analysing the frequency of extreme rainfalls [2]. Therefore, conforming stationarity or detecting nonstationarity is crucial for extreme rainfall observations [3, 4]. There are several studies which have been conducted around the world in order to examine stationarity/nonstationarity and trend analysis of extreme rainfall data [5-7].

Management and planning of water resources includes the determination of required discharge capacity of channels, pumping station capacity and planning out the design and

---

[*] Corresponding author: sinan.jasim@yahoo.com

building of sewage and storm systems. This management can be eased by using statistical methods that use extreme rainfall data for carrying out the assessment of water resource management. These methods, particularly the IDF curves, can play an important role in: reducing the loss of property and life by judging and assessing hazards, the damage that can occur, and the preventive methods that need to be implemented [8]. The definition of Intensity – Duration – Frequency (IDF) curve of rainfall is given by [9, 10] as the illustration of the amount of rainfall falling on a watershed for a given period of time in the form of a graph.

The mathematical relationship among Return period T (which is also known as frequency referring to the annual frequency of exceedance f), Rainfall intensity I and Storm duration d can easily be found from the IDF curves [9, 11, 12]. Numerous studies have been conducted around the world to develop formulas from the designed IDF curves such as [8, 9, 13-15]. Several studies of developing such a formula have been, also, conducted in several cities in Turkey. For example, [16-21].

The presence of a huge mountain range, different surface shapes and three seas in the surrounding of Turkey has resulted in a varying climate all over the country. In order to get a clear understanding of that climate variation, each of the provinces must be studied. Eskisehir is one of these provinces where neither the historical changes in the extreme rainfall events nor the construction of IDF have been studied. The objective of this study to analyze the trend of the annual maximum rainfall of Eskisehir and develop IDF curves using several distribution functions followed by deriving formula can be used for estimating the rainfall intensity for any return period and duration.

2 Study area and data collection

Eskisehir is a north-western province of Turkey lying in between Istanbul (cultural and touristic capital of Turkey) and Ankara (administrative capital of Turkey); it is 233 km away from Ankara and 324 km from Istanbul. The location of the study area, shown in Figure 1, is between 39°44'00"N and 39°49'00"N latitudes and 30°27'00"E and 30°36'00"E longitudes. The study area has an area of 190 km2 and is inhabited by 755,427 people. Only the centre of Eskisehir city which can be represented by the used rainfall station is chosen.

Data was collected for the chosen metrological station located in the centre of Eskisehir from General Department of Meteorology - Ministry of Forest and Water Affairs, Tukrey. Data from 1940 to 2010 except for the period between 1950 and 1956 (7 years' duration) was recorded at the station. The recorded extreme rainfall storms have the duration of 5, 10, 15, 30, 60, 120, 180, 240, 300, 360, 480, 720, 1080, 1440 min.

3 Trend analysis and nonstationarity

It is important to detect the trend in any time series data in order to identify whether there is an increase or decrease during the historical observations. In this study, three trend tests were implemented: Mann-Kendall test (MK), Cox and Stuart test, and Sen's Slope test. Null hypothesis (H0) of MK test is that no monotonic trend is present whereas the alternative hypothesis (Ha) is that Monotonic trend is present. P value for every duration was calculated and compared with the confidence level. In the same manner, Cox and Stuart test has the same hypotheses (i.e. H0: No monotonic trend, and Ha: Monotonic trend is present) and the same P value comparison. On the contrast, the statistic test, was calculated for Sen's slope test in order to be compared with the confidence interval.

Trend analysis is implemented to detect whether there is a general increase or decrease in the time series data. However, this increase or decrease not always ensures the existence

of nonstationarity in the series. Therefore, further analysis was conducted to detect whether nonstationarity is present. In this study, three nonstationarity tests were implemented: Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test, Phillips–Perron (PP) test, and Augmented Dickey–Fuller (ADF) test. Null hypothesis (H0) of KPSS test is that data are stationary and the alternative hypothesis (Ha) is that nonstationarity is present. On the other hand, the hypotheses of PP and ADF tests are the opposite (i.e. H0: Nonstationary, and Ha: Stationary).
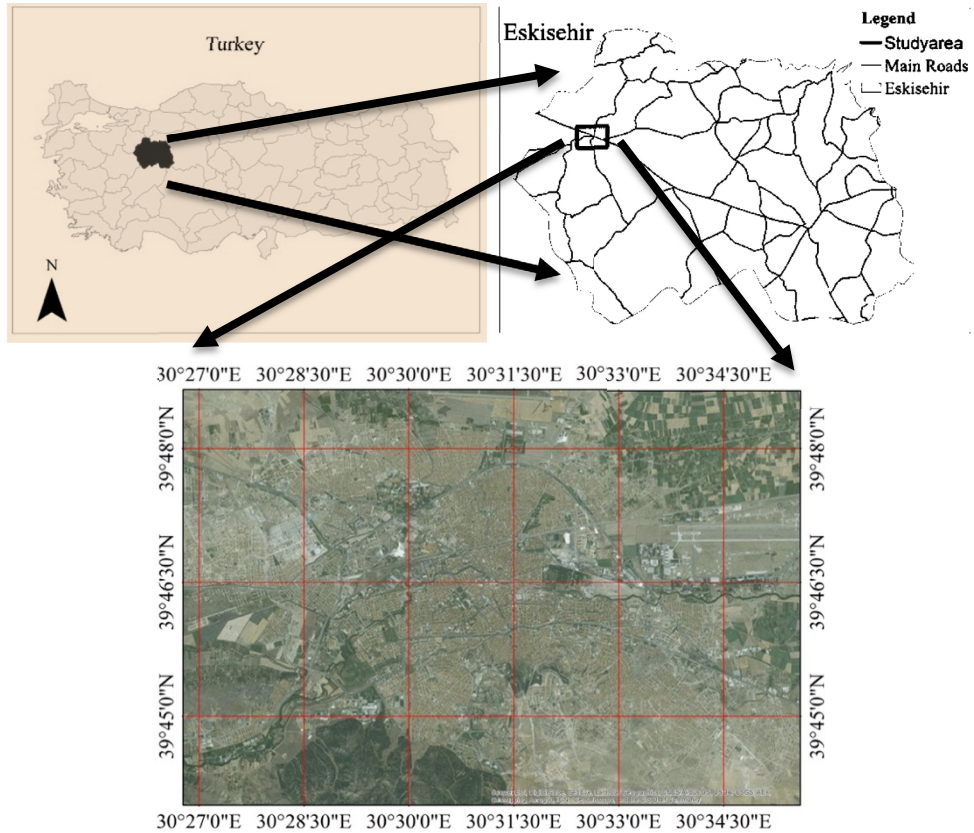

Fig. 1. Map of the study area (i.e. city centre of Eskisehir province represented by one rainfall station, Turkey) extracted by ArcGIS 10 using administrative and satellite imagery base maps and the main roads digitized from the StreetMap of ArcGIS online [22].

4 Development of IDF curves

The IDF curves are developed through a three steps procedure [23, 24]. Firstly, the records are fit into any of the probability distribution functions for a specific storm duration, for instance 10 minutes, 20 minutes, 30 minutes, 1 hour, … 24 hours. Then, the probability distribution function selected in the earlier step is used for determining rainfall intensities for a specific storm duration and one of the return periods that have already been defined (e.g. 2, 5, 10, 50, 100 years, etc.). Lastly, two approaches can be used: (a) an equation or mathematical relation is developed relating rainfall intensities, return periods and storm durations, (b) a graphical illustration of the rainfall intensities determined earlier for specific storm durations and each of the defined return periods. IDFs for four functions; Normal, Log-Normal, Gumbel, and Log-Pearson; were developed and shown on Figure 2.

5 Fitting of probability distribution function

There are some of the theoretical probability distribution functions are used commonly and were applied in various areas around the globe: GEV (Generalized Extreme Value Distribution), Gumbel (Extreme Value Distribution type I), Normal Distribution, Log-Normal Distribution, Pearson, Log Pearson Type III Distribution, etc. See [8, 9, 17, 23, 25]. In this study, four common probability distribution functions were used for the purpose of developing the IDF curves. These functions are: Normal, Log-Normal, Gumbel, and Log-Pearson. The goodness of fit test; Chi-Square test was used for testing the best fitting function.

6 Derivation of empirical IDF formula

The maximum rainfall intensities are calculated as dependent variable and the return period and the storm duration as independent variables using any of the four IDF curves developed earlier [8, 9, 12, 25]. A power-law relation as in Eq. (1) can be used for this purpose due to the advantage of having the intensity I_T being dependent on the return period T_r and storm duration t_d separately.

$$\mathbf{I_T} = \frac{fn\,(T_r)}{fn\,(t_d)} \tag{1}$$

The function $fn(T_r)$, according to [8, 23, 24] can be given as:

$$\mathbf{fn(T_r)} = a\,(T_r)^m \tag{2}$$

According to [23, 25], the storm duration $fn(t_d)$ can be given by:

$$\mathbf{fn(t_d)} = (t_d + b)^e \tag{3}$$

Placing Eq. (2) and (3) in this Eq. (1), following power-law equation is obtained:

$$\mathbf{I_T} = \frac{a\,(T_r)^m}{(t_d + b)^e} \tag{4}$$

In the equation, a, m, b and e represent empirical parameters and their values are to be determined. To calculate these parameters, the least squares method used and an equation for every distribution function was developed.

7 Results and discussion

P values of MK and Cox and Stuart tests for all the durations that shown in Table 1 are higher than any of the confidence levels (0.01, 0.05, and 0.1) which means we fail to reject the null hypothesis that data has no trend. Also, as can be seen in Table 1, all the Sen's slope test values are within the confidence interval which, hence, leads to the same decision that we fail to reject the null hypothesis that no monotonic trend is present. Therefore, there is no evidence that there is a significant trend in the observed extreme rainfall events.

Nonstationarity analyses results are shown in Table 2. The critical values for every test and confidence level are shown, also, in order to ease the inference. The result showed that even if the test values compared with the 0.1 confidence level value which is 0.347, they still all inside the confidence interval. So, we fail to reject the null hypothesis that the series are stationary. PP and ADF tests are left sided test. In specific, the least critical value correspond to 0.01 confidence level for PP test is -4.138 and for ADF test is -4.040. By comparing the test statistics for each duration with these values, it can be obviously seen that all the values are less which means that for both tests the null hypothesis is rejected. Due to the rejection

of the null hypothesis and as the null hypothesis in these two tests is that series are nonstationary, it is concluded that there is no evidence the data is nonstationary.

Table 1. Test statistics values for each duration of three Trend analysis tests: Mann-Kendall (MK), Cox and Stuart, and Sen's Slope.

Test	5 m	10 m	15 m	30 m	1 h	2 h	3 h	4 h	5 h	6 h	8 h	12 h	18 h	24 h
MK Test	0.17	0.14	0.13	0.23	0.38	0.32	0.27	0.16	0.29	0.39	0.16	0.19	0.51	0.94
Cox and Stuart	0.99	0.97	0.99	0.97	0.78	0.88	0.5	0.88	0.78	0.78	0.78	0.78	0.5	0.22
Sen's Slope	**0.04**	**0.06**	**0.09**	**0.09**	**0.06**	**0.04**	**0.06**	**0.07**	**0.05**	**0.05**	**0.07**	**0.06**	**0.04**	**0.003**
Sen's Slope Test	0.02	0.02	0.02	0.06	0.08	0.07	0.06	0.05	0.06	0.08	0.05	0.05	0.10	0.16
Confidence Interval (Upper and lower value)	**0.09**	**0.14**	**0.19**	**0.24**	**0.22**	**0.18**	**0.17**	**0.18**	**0.17**	**0.18**	**0.18**	**0.17**	**0.17**	**0.14**
Confidence levels for MK and Cox and Stuart test are: 0.1, 0.05, 0.01 Confidence Level for Sen's Slope is: 0.05 Bold numbers are negative values														

Table 2. Test statistics values for each duration of three Nonstationarity analysis tests.

Test	5 m	10 m	15 m	30 m	1 h	2 h	3 h	4 h	5 h	6 h	8 h	12 h	18 h	24 h
KPSS Test	0.25	0.31	0.32	0.2	0.09	0.06	0.09	0.05	0.05	0.04	0.03	0.03	0.04	0.09
PP Test	**5.98**	**6.98**	**6.72**	**6.99**	**7.41**	**7.75**	**7.51**	**8.13**	**7.93**	**7.92**	**7.17**	**7.11**	**7.84**	**7.06**
ADF Test	**5.92**	**6.95**	**6.71**	**6.99**	**7.41**	**7.75**	**7.51**	**8.13**	**7.93**	**7.92**	**7.17**	**7.11**	**7.84**	**7.06**
C. Level	KPSS	0.1	0.05	0.01	PP Test	0.1	0.05	0.01	ADF Test	0.1	0.05	0.01		
Critical value		0.35	0.46	0.74		**3.17**	**3.49**	**4.14**		**3.15**	**3.45**	**4.04**		
Bold numbers are negative values														

Kwiatkowski–Phillips–Schmidt–Shin (KPSS), Phillips–Perron (PP), and Augmented Dickey–Fuller (ADF).

According to the IDFs seen in Figure 2, the intensity and return period in all the functions are directly proportional to one another, i.e. if the return period increases the intensity was also seen increasing for a specific storm duration. It was also seen that rainfall intensity and storm duration are inversely proportional for the four functions. The intensity was seen changing dramatically for a certain return period until duration reaches 8-10 hours and after that it starts changing gradually. Also, the figure show that Log-Normal function has the highest intensity values for high return periods and values deduced from rest of the functions were close to each other.

To identify which function is the best fitting among the used distribution functions, the goodness of fit tests must be implemented. In this study, Chi-Square test was utilized. Two significance levels (i.e. $\alpha = 0.01$ and $\alpha = 0.05$) were used. The results of Chi-Square test are shown in Table 3. Based on the result in the table, all the distributions were considered to be well fitting the data as the test values are not significant in all the storm durations with an exception of Normal function which has significant result in several durations: 3, 6, 8 hours at the significance level $\alpha = 0.05$ and 12 hours at $\alpha = 0.01$.
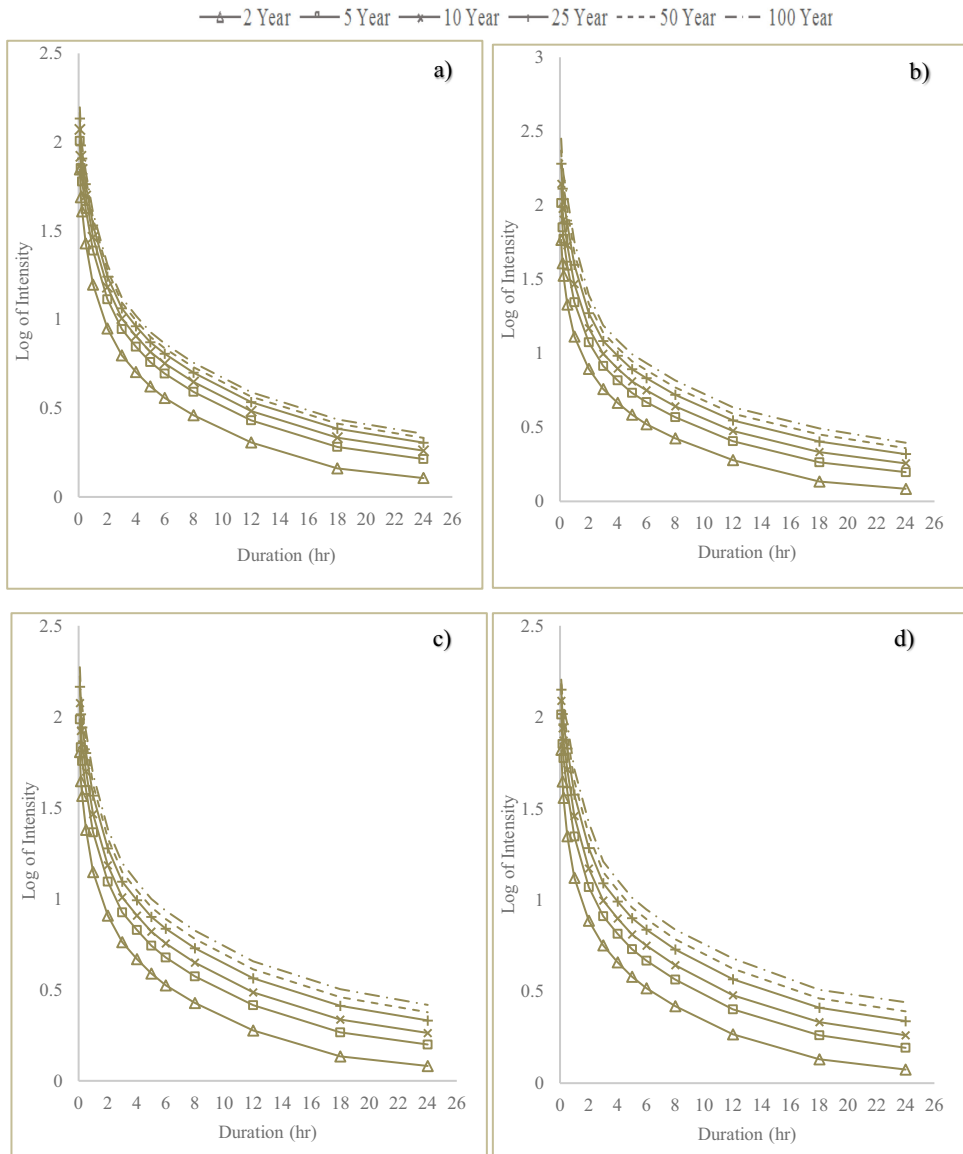
Fig. 2. Fitted Intensity- Duration – Frequency (IDF) Curve using extreme rainfall events for 14 storm durations (i.e. 5, 10, 15, 30 minutes and 1, 2, 3, 4, 5, 6, 8, 12, 18, 24 hour) and 6 return periods (i.e. 2, 5, 10, 50, 100 years) for four functions: a) Normal, b) Log-Normal, c) Gumbel, and d) Log-Pearson.

Table 3. the result of goodness of fit test (Chi – Square) of annual maximum rainfall.

Duration	Normal	LogNormal	Gumbel	Log - Pearson III
5 m	9.28	3.22	8.34	4.83
10 m	1.06	8.25	4.92	2.94
15 m	3.98	3.74	7.57	2.15
30 m	3.49	2.6	5.36	5.36
1 h	7.39	4.25	3.97	4.01
2 h	9.61	4.02	4.66	5.05

3 h	11.89**	4.15	8.56	2.45
4 h	8.91	6.83	10.8	7.8
5 h	9.62	8.7	10.3	1.85
6 h	12.14**	6.8	5.4	8.86
8 h	12.35 **	5.02	6.67	3.97
12 h	15.49 *	4.93	6.87	7.16
18 h	7.08	2.63	2.71	3.77
24 h	4.57	2.27	1.44	2.25
When $\alpha = 0.01$, degree of freedom = 5, the critical region is > 15.086				
When $\alpha = 0.05$, degree of freedom = 5, the critical region is > 11.07				
* X^2 significant at $\alpha = 0.01$				
** X^2 significant at $\alpha = 0.05$				

Table 4 lists out the parameters along with their obtained values and the developed formulas. The parameters set of each function formula was worked out using least squares method which implies the obtainment of accurate parameters. Determined parameters were substituted in Eq. (4) in order to obtain a formula for every function that can estimate the intensity for any return period and storm duration.

In the same table, coefficients of correlation were also listed for all the formulas of the four functions. This coefficient indicates the relation between the actual values and predicted values that were obtained by the developed formula. The highest correlation values obtained belong to the formulas of Log-Normal and Gumbel functions whereas the lowest value is for the formula of LPT III function.

Table 4. calculated parameters and coefficient of correlation for every function used in the derived formulas.

Function/Parameter	a	b	e	m	Derived Equation	Correlation
Normal	19.10	0.14	0.87	0.18	$I_T = \dfrac{19.10\,(T_r)^{0.18}}{(t_d + 0.14)^{0.87}}$	0.995
Log - Normal	13.70	0.14	0.97	0.34	$I_T = \dfrac{13.70\,(T_r)^{0.34}}{(t_d + 0.14)^{0.97}}$	0.996
Gumbel	16.93	0.14	0.88	0.24	$I_T = \dfrac{16.93\,(T_r)^{0.24}}{(t_d + 0.14)^{0.88}}$	0.996
Log - Pearson	22.83	0.28	1.01	0.21	$I_T = \dfrac{22.83\,(T_r)^{0.21}}{(t_d + 0.28)^{1.01}}$	0.992

8 Conclusion

Eskisehir is one of Turkey provinces requires its climate to be studied. The trend of extreme rainfall events for Eskisehir was investigated for the period between 1940 and 2010 (with an exception of the years between 1950 and 1956) using three tests: Mann-Kendall test (MK), Cox and Stuart test, and Sen's Slope test. There has been no significant trend found throughout the studied period. According to these results, the argument that areas with intense and frequent extreme storms are more likely to witness an increase while on the other hand, the areas which have less intense and frequent storms are expected to witness a decrease is proved, as Eskisehir considered as an area that has moderate extreme rainfall events. This could be due to the location of the study area in the middle of Turkey where the effect of the Mediterranean and Black sea climate is not present and therefore concluding that the precipitation of Eskisehir has not been affected by the climate change.

Stationarity of the extreme rainfall events for all the 14 storm durations examined using several tests: Kwiatkowski–Phillips–Schmidt–Shin (KPSS), Phillips–Perron (PP), and

Augmented Dickey–Fuller (ADF) tests. Log – Normal, Log –Pearson Type III, Normal, and Gumbel functions used to fit IDF curves with 6 return periods (i.e. 2, 5, 10, 50, 100 years) and 14 storm durations. Formulas representing these curves developed for the reason that can be used for estimating rainfall intensity rather than the curves. Therefore, it can be stated that the developed formulas for the study area of the Log-Pearson III and the Gumbel function, with preference given to the Gumbel function, can be used in order to estimate the rainfall intensity for any return period and storm duration.

In general, the use of formula proved to be a good alternative for getting the values from a graph i.e. IDF curve in order to obtain design storm. In this study, only the city center of Eskisehir (i.e. the urban area) with one rainfall station was studied while the consideration of studying the area with a regional effect by using many stations is highly recommended. Also, the effect of the climate change along with effect of the seas and mountains, surrounding Turkey and Eskisehir respectively, on the study area must be studied in detail.

References

1. IPCC, *Climate Change 2007: The Physical Science Basis. Contribution ofWorking Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge and NewYork: Cambridge University Press, (2007).
2. A. G. Yilmaz, M. A. Imteaz, and B. J. C. Perera, Investigation of non-stationarity of extreme rainfalls and spatial variability of rainfall intensity–frequency–duration relationships: a case study of Victoria, Australia, Int. J. Climatol., (2016).
3. E. A. Rosenberg, P. W. Keys, D. B. Booth, D. Hartley, J. Burkey, A. C. Steinemann, and D. P. Lettenmaier, Precipitation extremes and the impacts of climate change on stormwater infrastructure in Washington State, Clim. Change, **102**, 1–2, pp. 319–349, (2010).
4. a. G. Yilmaz and B. J. C. Perera, Extreme Rainfall Non-Stationarity Investigation and Intensity-Frequency-Duration Relationship, J. Hydrol. Eng., (2013).
5. S. Shahid, Trends in extreme rainfall events of Bangladesh, Theor. Appl. Climatol., **104**, 3, pp. 489–499, (2011).
6. H. Shang, J. Yan, M. Gebremichael, and S. M. Ayalew, Trend analysis of extreme precipitation in the Northwestern Highlands of Ethiopia with a case study of Debre Markos, Hydrol. Earth Syst. Sci., **15**, 6, pp. 1937–1944, (2011).
7. E. M. Douglas, M. Asce, and C. A. Fairbank, Is Precipitation in Northern New England Becoming More Extreme ? Statistical Analysis of Extreme Rainfall in Massachusetts , New Hampshire , and Maine and Updated Estimates of the 100-Year Storm, Most, **16**, 3, pp. 203–218, (2011).
8. S. A. AlHassoun, Developing an empirical formulae to estimate rainfall intensity in Riyadh region, J. King Saud Univ. - Eng. Sci., **23**, 2, pp. 81–88, (2011).
9. I. H. Elsebaie, Developing rainfall intensity–duration–frequency relationship for two regions in Saudi Arabia, J. King Saud Univ. - Eng. Sci., **24**, 2, pp. 131–140, (2012).
10. B. S. Dupont and D. L. Allen, Revision of the rainfall intensity duration curves for the commonwealth of kentucky., USA, (2000).
11. D. Koutsoyiannis, *On the appropriateness of the gumbel distribution in modelling extreme rainfall*, in *Proceedings of the ESF LESC Exploratory Workshop, European Science Foundation, National Resea*, (2003), pp. 303–319.
12. D. Koutsoyiannis, D. Kozonis, and A. Manetas, A mathematical framework for studying rainfall intensity-duration-frequency relationships, J. Hydrol., **206**, 1–2, pp.

118–135, (1998).

13. V. R. Baghirathan and E. M. Shaw, Rainfall depth-duration-frequency studies for Sri Lanka, J. Hydrol., **37**, 3–4, pp. 223–239, (1978).

14. A. Gert, D. J. Wall, E. L. White, and C. N. Dunn, Regional Rainfall intensity-duration-function curves for Pennsylvania, Water Resour. Bull, **23**, pp. 479–486, (1987).

15. C. Chen, Rainfall Intensity-Duration-Frequency Formulas, ASCE J. Hydraul. Eng., **109**, 12, pp. 1603–1621, (1983).

16. R. Acar, S. Celik, and S. Senocak, Rainfall Intensity-Duration - Frequency Model using an artificial neural network approach, J. Sci. Ind. Res., **67**, pp. 198–202, (2008).

17. R. Acar and S. Senocak, *Modelling of Short Duration Rainfall (SDR) Intensity Equations for Ankara, Turkey*, in *Balwois 2008*, (2008), May, pp. 1–9.

18. S. Senocak and R. Acar, Modelling of Short Duration Rainfall (SDR) Intensity Equations for Erzurum, Turkey, J. Eng. Sci., **13**, 1, pp. 75–80, (2007).

19. S. Senocak and R. Acar, Modelling of short-duration rainfall intensity equations for the Agean region of Turkey, FRESENIUS Environ. Bull., **16**, 9b, (2007).

20. O. L. Asikoglu and Ertugrul Benzeden, Simple generalization approach for intensity–duration–frequency relationships, Hydrol. Process., **28**, pp. 1114–1123, (2014).

21. T. Haktanir, M. Cobaner, and O. Kisi, Frequency analyses of annual extreme rainfall series from 5 min to 24 h, Hydrol. Process., **24**, 24, pp. 3574–3588, (2010).

22. ESRI, ArcGIS online, 2016. [Online]. Available: https://www.arcgis.com/home/index.html. [Accessed: 12-Jan-2016].

23. W. T. Chow, *Handbook of Applied Hydrology*. McGraw-Hill, (1988).

24. V. P. Singh, *Elementary hydrology*. New Jersey: Prentice Hall, (1992).

25. L. Nhat, Y. Tachikawa, and K. Takara, Establishment of intensity-duration-frequency curves for precipitation in the monsoon area of Vietnam, Annu. Dis. Prev. Res. Inst, 49B, pp. 93–103, (2006).