ANALYSIS AND MODELLING OF RAINFALL EVENTS

IOANNIDIS K., KARAGRIGORIOU A. and LEKKAS D.F.

Analysis and Simulation of Environmental Systems (ASES), University of the Aegean
E-mail: sasm13005@sas.aegean.gr

ABSTRACT

The purpose of this study is the statistical analysis of rainfall events to explore patterns and dependencies that would allow the generalization in cases of missing or truncated data. More specifically, in this paper we estimate intensity–duration–frequency (IDF) curves, which are widely used to model rainfall. We use data from Eresos meteorological station and estimate the parameters of the model for fixed return periods. Sensitivity analysis is conducted to check whether the estimates are optimal. Finally, a more general model is applied that allows for simultaneous modelling of rainfall duration, intensity and frequency (via return periods).

Keywords: rainfall modelling, intensity-duration-frequency curves, sensitivity analysis, Monte Carlo simulation

1. Introduction

It is critical for hydrological design and water engineering to develop modelsthat explain the pattern under which rainfall events occur and accurately predict thefuture inflow of rainfalls. More specifically, it is necessary to find relationshipsthat connect rainfall intensity and duration while simultaneously taking into accountthe frequency of occurrence. Surprisingly, a model proposed almost a century ago byBernard (1932) is still one of the most widely used for rainfall modeling. A moregeneral formula for the same purpose was proposed and thoroughly examinedmore recently by Koutsoyiannis *et al*. (1998).

The tools used for the modelling of such characteristics are curves commonly known as intensity–rainfall–frequency (IDF) curves. An Intensity-Duration-Frequency curve (IDF curve) is a graphical representation of the probability that a rainfall of given duration and intensity will occur. The parameters that make up the axes of the graph are:

- Rainfall Duration (in hrs)
- Rainfall Intensity (in mms/hr)
- Rainfall Frequency (Return Period) (in yrs)

Data comes from Eresos (Lesvos, Greece) meteorological station over a period of 3 years (17.11.2009 - 16.12.2012) and consists of 234 rainfall events. For each event the following variables are recorded:

- (R) Total rainfall amount (in mms)
- (D) Event duration (in hrs)
- (AI) Average intensity (in mm/hrs)

To have a preliminary picture of the data, we provide in Table 1 their statistical characteristics.

The return period (T_r) is defined as $T_r=\frac{1}{1-F(r)}$, where F is the cumulative distribution function of the total rainfall amount. For every return period we want to find an IDF curve that describes the relation between duration (D) and intensity (I) and satisfies the following equation:

$$I=\frac{c}{(D+a)^b} \qquad (1)$$

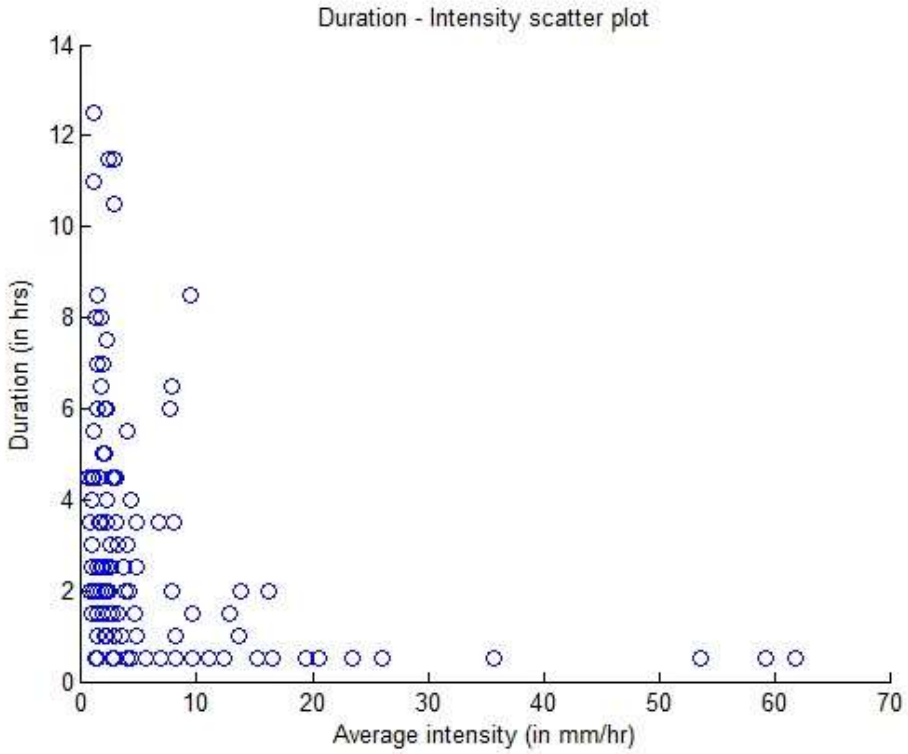
where a, b and c are parameters to be estimated. It is common for parameters a and b to have similar values for every return period in a given location. In such cases, one can apply a more general model that simultaneously connects rainfall duration, intensity and frequency (via return periods). The general model, which includes the return period (T_r) and an additional parameter K, is defined as:

$$I = \frac{K T_r^d}{(D+a)^b} \quad (2)$$

Table 1: Descriptive statistics of rainfall events

Statistic	R	D	AI
Maximum	79.80	12.50	61.8
Minimum	0.60	0.50	0.62
Mean	5.87	1.89	4.21
Variance	86.68	5.85	57.85
Standard Deviation	9.31	2.41	7.61
1st Quartile	0.70	0.50	1.40
Median	2.10	0.50	1.69
3rd Quartile	6.90	2.38	3.55

Since it is natural for rainfall duration and intensity to be inversely proportional we choose to apply the proposed model which obviously satisfies this property. The scatter plot for the Eresos data set clearly confirms inverse proportionality of duration and intensity.



Graph 1: Duration-Intensity scatter plot

Similar analyses have been conducted in many geographical areas and have resulted in a global appreciation of the model. One can refer to Stern and Coe (1984), Lee (2005) and Rao and Kao (2006) for related results.

2. Model estimation

We applied model (1) for 4 different return periods, that is for rainfalls that appear once every 2, 5, 10 and 20 years. The estimates along with the 95% confidence intervals for the parameters are presented in the following table.

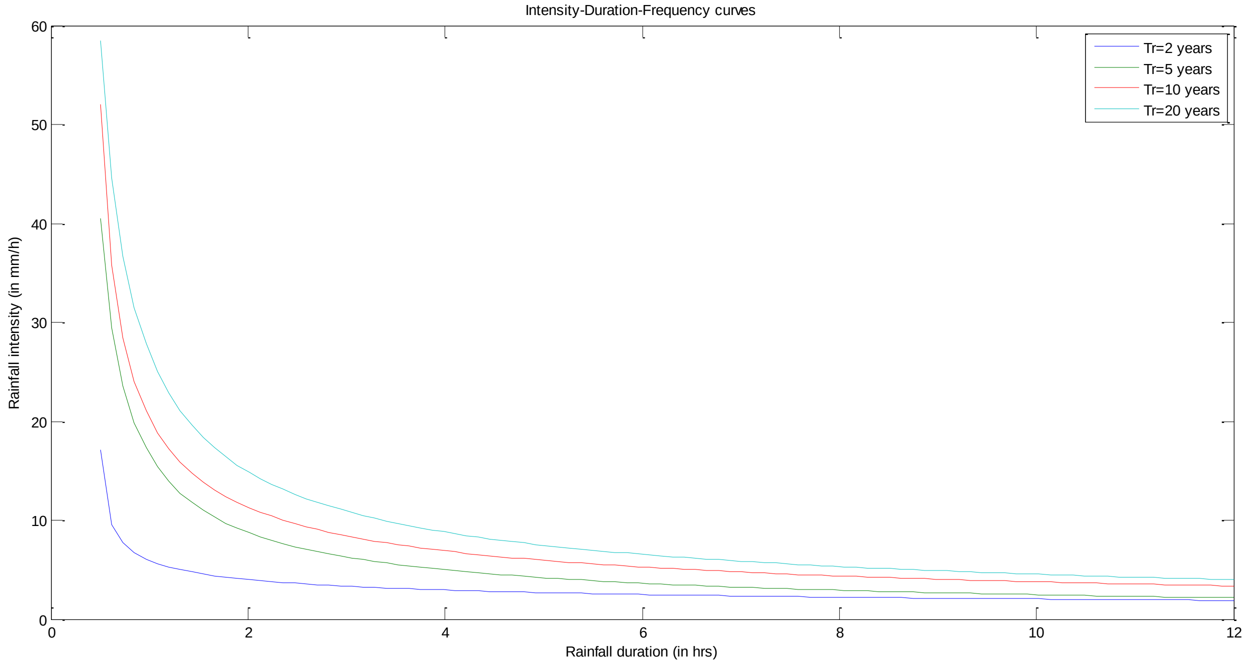
Table 2: Parameter estimates with 95% confidence intervals

T_r	2	5	10	20
a	-0.47	-0.29	-0.36	-0.26
95% CI	(-0.70, -0.24)	(-0.91, 0.33)	(-0.82, 0.10)	(-0.67, 0.16)
b	0.37	0.73	0.62	0.69
95% CI	(-0.31, 1.06)	(-0.06, 1.52)	(-0.04, 1.29)	(0.27, 1.11)
c	4.69	12.98	15.39	21.86
95% CI	(1.60, 7.79)	(0.39, 25.57)	(2.89, 27.89)	(8.56, 35.15)

The resulting IDF curves in an equation form are:

- Return period T_r=2 years $I=\frac{4.69}{(D-0.47)^{0.37}}$
- Return period T_r=5 years $I=\frac{12.98}{(D-0.29)^{0.73}}$
- Return period T_r=10 years $I=\frac{15.39}{(D-0.36)^{0.62}}$
- Return period T_r=20 year $I=\frac{21.86}{(D-0.26)^{0.69}}$s

As mentioned in the previous section, an IDF curve is a graphical method with 3 axes, duration, intensity and return period. The above equations in a common plot provide this graphical picture.



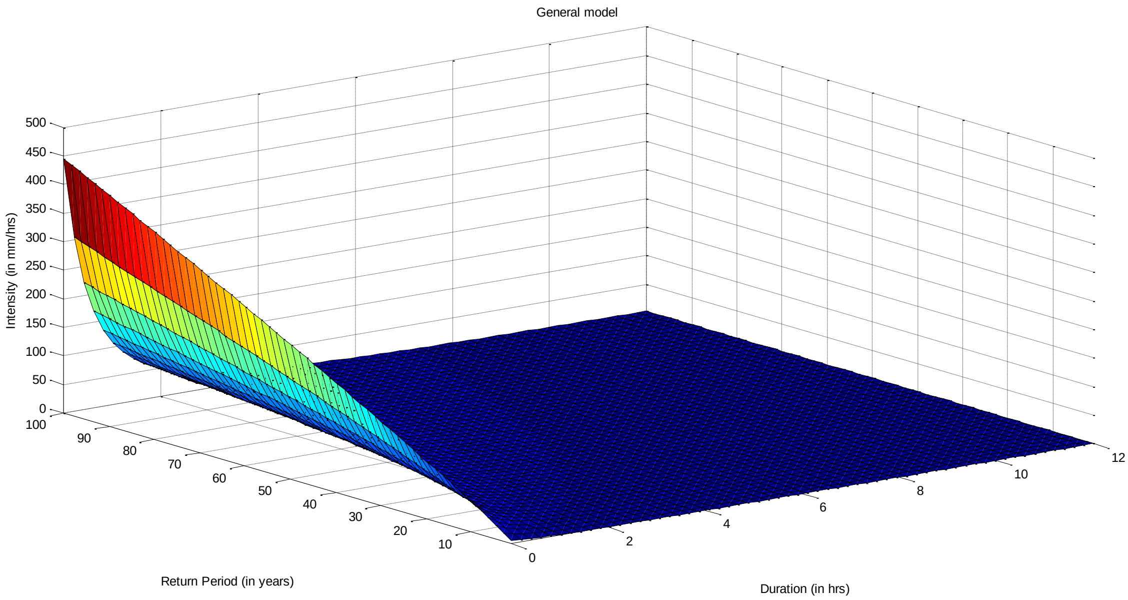
Graph 2: IDF curves for T_r = 2, 5, 10 and 20 years

Since estimates for a and b were close, we applied the general model, given in (2), which includes the return period (T_r) and the extra parameter K and found the following results:

Table 3: Estimates and confidence intervals for general model

Parameter	Estimate	95% CI
a	0.9325	(0.3229, 1.542)
b	1.921	(1.449, 2.393)
K	5.893	(0.1919, 11.59)
d	0.91	(0.8813, 0.9387)

A 3D plot was created that graphically depicts the above equation.



Graph 3: IDF curve with duration, intensity and return period

3. Sensitivity analysis

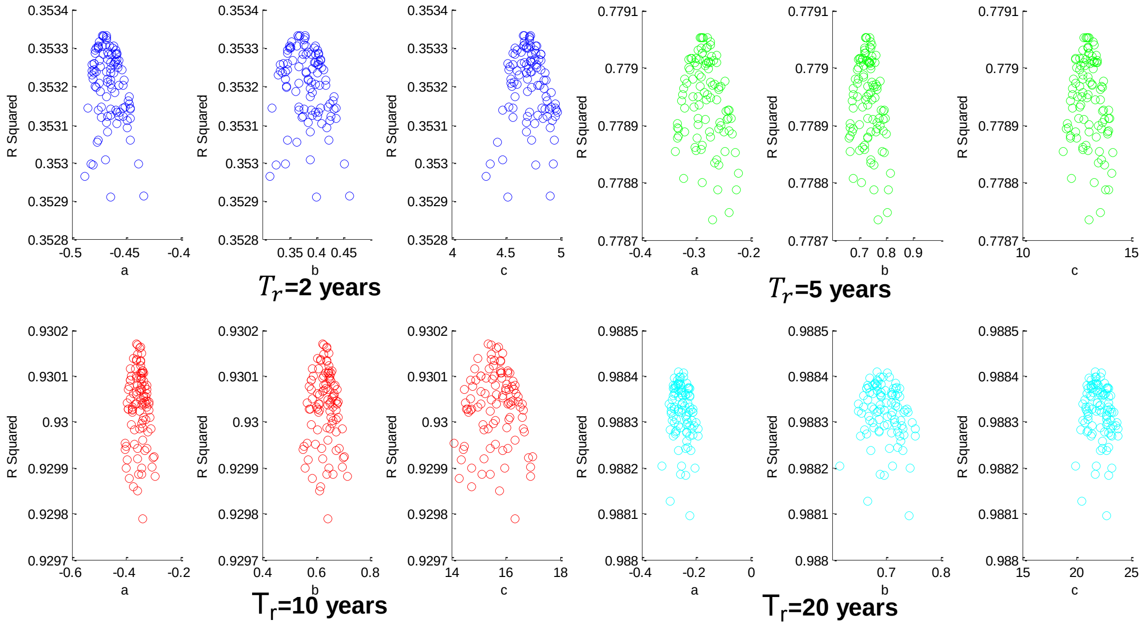
In non-linear models, there is a chance that estimation methods can result in sub-optimal estimates of the parameters. This can happen when the objective function we try to maximize (here R^2) gets stuck in a local optima or is not well defined for the selected model. To check whether the set of parameters is the optimal (the global maxima has been reached), an analysis using a Monte Carlo type method was conducted to investigate parameters' behaviour.

For each return period T_r, we create 10000 random samples (a_i,b_i,c_i), where a_i,b_i,c_i~Uniform for different choices of intervals and each time we keep the parameter-vector with the largest R^2. The above procedure was repeated 100 times. Results, as shown in the following graph, validated the optimality of the parameter estimates.

4. Conclusions

We applied a model that has been extensively used in the literature to model rainfall, IDF curves. The parameters of the model were estimated and graphical representations of the IDF curves were created. Sensitivity analysis was conducted on the parameters to check for their optimality and the estimates were found optimal. This was an initial part of ongoing research with data from only one meteorological station and can be further explored in two directions:

- Spatial generalization of the model by using data from nearby stations in the island of Lesvos and possibly the broader North-Eastern Aegean area.
- Temporal generalization of the model by comparing predicted future data with future inflows of actual data to validate the accuracy of the model to predict rainfall behavior.



Graph 4: R^2 as a function of model parameters

REFERENCES

1. Bernard, M. (1932), Formulas for rainfall intensities of long duration, Transactions ofthe American Society of Civil Engineers, 96, 592–624
2. Koutsoyannis D. *et al*. (1998), A mathematical framework for studying rainfall intensity-duration-frequency relationships, Journal of Hydrology, 206, 118–135
3. Lee, C.-Y. (2005), Application of Rainfall Frequency Analysis on studying rainfall distribution characteristics of Chia-Nan plain area in southern Taiwan, Crop, Environment & Bioinformatics, 2, 31-38
4. Rao, A.R and Kao, S.C. (2006), Statistical Analysis of Indiana rainfall data, Joint Transportation Research Program, Final Report
5. Stern, R.D. and Coe, R. (1984), A model fitting analysis of daily rainfall data, Journal of the Royal Statistical Society. Series A (General), 147, 1-34.