

The impact of the time series resolution on the reliability of the maximum precipitation models

Bartosz Kaźmierczak^{1,*}, Katarzyna Wartalska¹, and Marcin Wdowikowski²

¹Wrocław University of Science and Technology, Faculty of Environmental Engineering, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

²Institute of Meteorology and Water Management – National Research Institute, Podleśna Street 61, 01-673 Warsaw, Poland

Abstract. In the paper, two maximum precipitation models for Legnica were developed. For this purpose archival pluviographic records from the time span 1961–2010 were used. The first model was developed on the basis of rainfall data of durations ranging from 5 minutes to 6 days. In the second model, rainfall data with durations ranging from 10 min to 6 days were used, and precipitation amounts for 5 minutes duration were extrapolated. Generalized exponential distribution was used to develop the models. Both models were compared with measurement data using relative residual mean square error.

1 Introduction

The basic form of quantitative rainfall description are models of dependence: intensity I (in mm/min), unit intensity q (in $\text{dm}^3/\text{s}\cdot\text{ha}$) or the height h (in mm) of precipitation on its duration t and the probability of exceedance p [1].

The relationship between precipitation intensity (unit intensity or height) and its duration is most often presented in terms of intensity-duration-frequency (IDF) or depth-duration-frequency (DDF) curves, for different probabilities p of precipitation occurrence. These curves represent the hyperbolic family given by the general equation [1, 2]:

$$I = \frac{a}{(t + b)^n} + c \quad (1)$$

in which: a , b , c and n are the empirical coefficients, dependent on the occurrence probability of a given rainfall and on the climatic and physiographic factors of the catchment. A series of homogeneous observation for decades is required for their determination [1].

Authoritative for designing of areas drainage systems are both short-term precipitation with a high-intensity and long-lasting precipitation with a significant territorial scope and high efficiency. Recommended by the European standard PN-EN 752:2008 modeling of existing or newly designed sewerage systems encounters in Poland a barrier of a lack of

* Corresponding author: bartosz.kazmierczak@pwr.edu.pl

access by designers to appropriate and reliable rainfall databases. At the entrance to hydrodynamic models are most commonly required storm hyetographs with a 5 minutes time resolution, such as Euler's type II model precipitation. The basis for the development of model precipitation are the models of maximum rainfall in the form of IDF or DDF curves [3–10].

In Poland, access to the precipitation source data is a matter for the Institute of Meteorology and Water Management (IMWM), the owner of the largest number of meteorological stations in the country.

By 2010, precipitation in the IMWM were recorded on paper pluviographs, which involved a very time-consuming development of measurement data. However, due to the number of the stations and the length of the measurement series (often decades) paper pluviographic strips are an extremely valuable research material on precipitation in Poland. With appropriate commitment, information about rainfall height can be read from the paper strips with a resolution time of 5 minutes. Based on these form of analogue but continuous records of rainfall events, many models of maximum precipitation have been developed and currently applied in engineering practice [1, 11, 12].

Since 2007, the IMWM began to record precipitation by digital rain gauges, initially in parallel with traditional devices. Time series are currently recorded with a 10-minute time resolution. The assumptive resolution from the point of view of urban hydrology is a matter of concern because in the sewage systems designing and modeling it is needed information about shorter precipitation, lasting 5 minutes [2].

The aim of the study is to evaluate the impact of time series resolution on the reliability of the maximum precipitation models. The analysis, were conducted using two maximum precipitation models based on archival pluviographs from Legnica from the time span 1961–2010. The first model was based on rainfall data of durations ranging from 5 minutes to 6 days. In the second model, rainfall data with durations ranging from 10 minutes to 6 days were used, and the amounts of precipitation with the duration of 5 minutes were extrapolated. Both models were compared with measurement data.

2 Materials and methods

Archived pluviographs from meteorological station of IMWM in Legnica from the time span 1961–2010 were used as research data. Measuring station in Legnica, as part of a national measurement and observation network at hydrological and meteorological service, is a synoptic station which is participating in the international weather monitoring program (Weather World Watch) as part of the World Meteorological Organization (WMO), of which Poland is a member. The station building is located on the south-eastern outskirts of the city of Legnica, at an elevation of 122 m above the sea level. The predominant land use in both the municipality and rural area around the station are fields and wasteland [12].

In order to determine the relationship between the amount of rainfall from duration and probability of exceedance $h(t,p)$, there must be done a selection of data on which the relationship will be developed. Elaborating archival pluviographs authors limited period of analysis to months from May to October (V–X) [1, 11, 12].

For the purpose of this paper, based on the Peak-Over-Threshold (POT) method [13–19] there were isolated from the tested 50-years period top 50 maximum amounts (h , mm) of rainfall for each of the following rainfall durations: $t = 5, 10, 20, 30, 40, 50, 60, 90, 120, 180, 360, 720, 1080, 1440, 2160, 2880, 4320, 5760, 7200$ and 8640 minutes.

In the first place, the top 50 amounts of rainfall were ordered decreasing. Then there were successively assigned to it the empirical probability of exceedance according to (1) from $p = 0.020$ (for the highest value) to $p = 0.980$ (for the lowest value) [1, 11]:

$$p(m, N) = \frac{m}{N + 1} \tag{2}$$

where m is the sequence number within a decreasing ordered string of the number of N . The amount of rainfall recorded for selected values of empirical probability are shown in Table 1.

Table 1. The amount of rainfall (h , mm) recorded for selected values of empirical probability ($m = 1, 5, 10, 25$ and 50).

t , min	p				
	0.020	0.098	0.196	0.490	0.980
5	16.2	11.2	9.8	7.6	4.6
10	23.5	17.5	14.1	10.9	8.3
20	31.7	25.1	20.1	14.5	11.3
30	32.0	26.9	24.5	16.0	12.5
40	37.4	28.2	25.2	17.6	13.0
50	39.8	31.0	25.2	18.2	13.7
60	40.6	31.6	25.9	19.9	14.2
90	41.0	36.2	29.4	21.6	16.5
120	49.6	39.6	32.2	22.8	18.2
180	57.6	40.7	34.5	26.3	20.0

t , min	p				
	0.020	0.098	0.196	0.490	0.980
360	57.7	47.2	40.9	30.6	23.9
720	74.9	51.8	44.6	35.8	28.6
1080	77.3	57.7	51.6	41.3	32.4
1440	77.3	66.3	57.7	46.4	35.7
2160	114.8	77.3	61.1	49.4	38.7
2880	129.3	97.8	74.3	53.0	41.3
4320	143.1	97.8	77.6	57.4	41.3
5760	157.0	116.0	83.9	61.9	47.1
7200	158.5	121.1	86.0	69.9	51.7
8640	167.9	132.3	91.9	71.4	53.7

To describe the measurement data generalized exponential distribution (GED) was used. Likelihood function of this distribution describes an equation:

$$\ln L(\alpha, \lambda, \gamma) = n \ln \alpha + n \ln \lambda - \sum_{i=1}^n \lambda(x_i - \gamma) + (\alpha - 1) \sum_{i=1}^n \ln(1 - e^{-(x_i - \gamma)\lambda}) \tag{3}$$

where α , λ and μ are parameters of the distribution. Parameters can be determined by the maximization of the likelihood function or by solving the system of normal equations [20–24]:

$$\begin{cases} \frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \ln(1 - e^{-(x_i - \gamma)\lambda}) = 0 \\ \frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n (x_i - \gamma) + (\alpha - 1) \sum_{i=1}^n \frac{(x_i - \gamma)e^{-(x_i - \gamma)\lambda}}{1 - e^{-(x_i - \gamma)\lambda}} = 0 \\ \frac{\partial L}{\partial \gamma} = n\lambda - \lambda(\alpha - 1) \sum_{i=1}^n \frac{e^{-(x_i - \gamma)\lambda}}{1 - e^{-(x_i - \gamma)\lambda}} = 0 \end{cases} \tag{4}$$

Quantiles of a random variable for the GED described by the equation

$$h = \gamma - \frac{1}{\lambda} \ln \left(1 - (1 - p)^{\frac{1}{\alpha}} \right) \tag{5}$$

The coincidence of theoretical distributions with measured data was examined using the Anderson-Darling test for statistics [25, 26]:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) [\ln F(X_i) + \ln(1 - F(X_{n-i+1}))] \tag{6}$$

where: x_i – i -th value in the decreasing ordered random sample, $F(x)$ – cumulative distribution function for the theoretical distribution. The null hypothesis H_0 (when the measurement data were suitable for tested theoretical distribution), were taken on a significance level of 0.05 if the A^2 test statistic was less than the critical value A_{kr}^2 . The alternative hypothesis was taken otherwise. The critical values can be read from the statistical tables or obtained by Monte Carlo method [25, 27]. For GED distribution and $N = 50$ critical value $A_{kr}^2 = 0.723$.

Relative residual mean square error (R_{RMSE}) was used to evaluate the aptitude of investigated distributions and to describe the measurement data

$$R_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{h_{l,i} - h_{m,i}}{h_{m,i}} \right)^2} \cdot 100\% \tag{7}$$

where: h_l – the theoretical amount of rainfall (mm), h_m – amount of rainfall from measurements (mm).

3 Results

Calculation results of particular parameters of GED were presented in Table 2. The parameter estimates were determined by numerical maximization of the log-likelihood function (3). The calculations were carried out for each of 20 durations of maximum precipitation amounts analyzed in the paper. GED distribution fulfils the compliance criterion A^2 for each of the 20 analyzed rainfall durations. There were also calculated relative residual mean square error statistics, covering the entire range of data – all 20 durations. In this case $R_{RMSE} = 3.3\%$.

Table 2. Calculation results of parameters of GED.

t , min	α	λ , 1/mm	γ , mm	A^2	R_{RMSE}
5	2.457	0.434	4.3	0.178	2.90
10	0.785	0.242	8.3	0.261	2.84
20	0.596	0.144	11.3	0.491	3.94
30	0.731	0.146	12.5	0.319	4.32
40	0.861	0.143	13.0	0.139	2.57
50	0.937	0.143	13.7	0.197	2.36
60	1.081	0.145	14.2	0.311	2.53
90	0.831	0.120	16.5	0.422	3.64
120	0.901	0.124	18.2	0.405	2.75
180	0.835	0.106	20.0	0.176	2.45
360	0.870	0.097	23.9	0.292	3.37
720	0.886	0.088	28.6	0.416	2.44
1080	0.910	0.084	32.4	0.255	2.35
1440	0.874	0.070	35.7	0.722	3.78
2160	0.879	0.059	38.7	0.474	3.13
2880	0.790	0.043	41.3	0.166	3.14
4320	0.924	0.044	41.3	0.424	3.53
5760	0.823	0.035	47.1	0.370	3.76
7200	0.633	0.031	51.7	0.396	3.94
8640	0.762	0.032	53.7	0.268	4.30

Following the equation (5) and the parameters listed in Table 2, the precipitation amount with any exceedance probability and selected rainfall duration ranging from 5 to 8640 minutes can be calculated. In this respect, it should be noted that the current 10-minute time step of rainfall registration will not allow in the future to estimate parameters for precipitation with the duration of 5 minutes. So these parameters will have to be extrapolated.

The parameters λ and γ have a clearly marked trend, while the parameter α is deprived of a trend. Since there is no dependency trend $\alpha(t)$, the mean value of $\bar{\alpha} = 0.837$ was assumed in the calculations.

Based on the calculated GED distribution parameters there were prepared plots (Fig. 1) showing their dependence on the rainfall duration (from 10 to 8640 minutes).

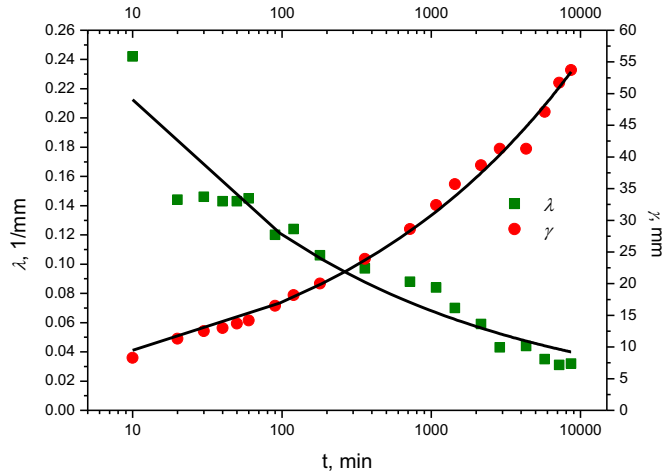


Fig. 1. The dependence of the parameters λ and γ of the rainfall duration.

The relationship of parameters λ and γ of the rainfall duration are described as functions:

$$\lambda = 0.376 t^{-0.247} \tag{8}$$

$$\gamma = 5.261 t^{0.256} \tag{9}$$

for the coefficient of determination of 0.929 and 0.994, respectively.

Finally, a model describing the dependence of the amount of rainfall on its duration and a specified exceedance probability, based on the quantiles of GED distribution (5), takes the form of:

$$h = 5.261 t^{0.256} - 2,660 t^{0.247} \ln(1 - (1 - p)^{1.195}) \tag{10}$$

Based on the obtained formula, precipitation amounts for the 5 min duration and the exceedance probability $p = 0.020, 0.098, 0.196, 0.490$ and 0.980 were computed. The received results were compared with the measurement data (Table 1) and the precipitation amount calculated using formula (5) with the parameters from Table 2. The results are shown in Table 3.

Table 3. Calculation results of parameters of GED.

p	$h_m, \text{ mm}$	$h_p, \text{ mm}$ (5)	R_{RMSE}	$h_p, \text{ mm}$ (10)	R_{RMSE}
0.020	16.2	15.4	4.94%	22.8	40.74%
0.098	11.2	11.6	3.57%	16.5	47.32%
0.196	9.8	10.0	2.04%	13.8	40.82%
0.490	7.6	7.6	0.00%	10.3	35.53%
0.980	4.6	4.8	4.35%	8.0	73.91%

There were also calculated relative residual mean square error statistics, covering the entire range of data ($t = 5$ min). In this case $R_{RMSE} = 2.9\%$ for h calculated by (5) and $R_{RMSE} = 41.5\%$ for h calculated by (10). The fit quality of the equations for rainfall data from Legnica is shown in the $h-h$ plot (Fig. 2).

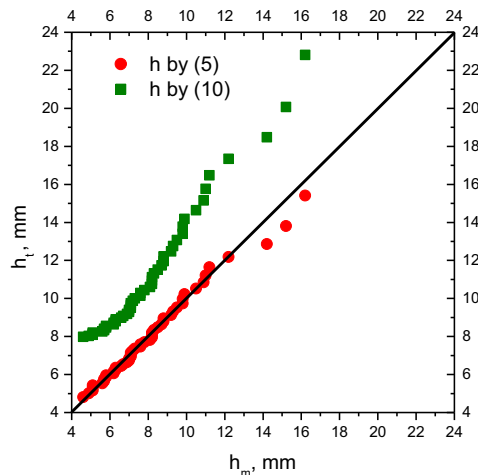


Fig. 2. The h - h plot for the GED distribution.

The analysis of the results clearly indicates that the extrapolation of the GED distribution parameters does not produce acceptable results. Calculated according to (10) rainfall heights of a $t = 5$ min duration do not correlate with the results of measurements.

4 Conclusions

In this paper, two maximum precipitation models for Legnica were developed. For this purpose archival pluviographic records from the time span 1961–2010 were used. The first model was developed on the basis of rainfall data of durations ranging from 5 minutes to 6 days. In the second model, rainfall data with durations ranging from 10 min to 6 days were used, and precipitation amounts for 5 minutes duration were extrapolated. Generalized exponential distribution (GED) was used to develop the models. Both models were compared with measurement data using relative residual mean square error. The conducted analysis allowed to draw the following conclusions:

- The maximum precipitation model (5) with the parameters shown in Table 2 accurately maps the measurement results - $R_{RMSE} = 3.3\%$. In particular, for rainfall with a duration of $t = 5$ min, the differences between the predicted and measured precipitation amounts at the level of $R_{RMSE} = 2.9\%$ were noted.
- In the case of using the model (10) to determine precipitation heights with a duration of $t = 5$ min (parameters extrapolation) obtained results were not correlated with the measurement results - $R_{RMSE} = 41.5\%$. Therefore, this method should be regarded as inappropriate for estimating precipitation amounts for the duration of $t = 5$ min.
- Precipitation amounts with a duration of $t = 5$ min can be estimated with some approximation based on the precipitation heights data for the duration of $t = 10$ min. The analysis of measurement data indicates that during the period considered, the maximum precipitation with a duration of $t = 5$ min represented on average 68% of maximum rainfall heights with a duration of $t = 10$ min. However, this estimation should be treated only indicatively, since this variability ranged in the case of the largest 50 precipitations from 55% to 75%. In order to draw up reliable conclusions, further research in this field involving other measurement stations is needed.
- In the light of the obtained results, it is requested to increase the temporal resolution of recorded by IMWM rainfall for 5-minute or shorter intervals.

The work was realized within the allocation No. 0401/0069/16 awarded for Faculty of Environmental Engineering Wrocław University of Science and Technology by Ministry of Science and Higher Education in years 2016–2017.

References

1. A. Kotowski, B. Kaźmierczak, A. Dancewicz, *The modeling of precipitations for the dimensioning of sewer systems* (Polish Academy of Sciences, Warsaw, 2010)
2. A. Kotowski, *The principles of safe dimensioning of sewer systems* (Seidel-Przywecki, Warsaw, 2015)
3. E. Burszta-Adamiak, M. Mrowiec, *Water Sci Technol* **68**, 1 (2013)
4. A. Kotowski, K. Wartalska, M. Nowakowska, *Ochrona Środowiska* **38**, 1 (2016)
5. B. Kowalska, D. Kowalski, G. Łagód, M. Widomski, *Modelling of Hydraulics and Pollutants Transport in Sewer Systems with Exemplary Calculations in SWMM* (Lublin University of Technology, Lublin, 2013)
6. S. Y. Park, K. W. Lee, I. H. Park, S. R. Ha, *Desalination* **226**, 1–3 (2008)
7. T. G. Schmitt, *Kommentar zum Arbeitsblatt A 118 Hydraulische Bemessung und Nachweis von Entwässerungssystemen* (DWA, Hennef, 2000)
8. T. G. Schmitt, M. Thomas, *KA – Wasserwirtschaft, Abwasser, Abfall* **47**, 1 (2000)
9. D. Słyś, A. Stec, *Environment Protection Engineering* **38**, 4 (2012)
10. A. Wałęga, G. Kaczor, B. Stęplewski, *Pol J Environ Stud* **25**, 5 (2016)
11. E. Bogdanowicz, J. Stachý, *Maximum rainfall in Poland. Design characteristics* (The Publishing House of the Institute of Meteorology and Water Management, Warsaw, 1998)
12. B. Kaźmierczak, M. Wdowikowski, *Periodica Polytechnica Civil Engineering* **60**, 2 (2016)
13. W. Jakubowski, *Meteorology Hydrology and Water Management* **3**, 2 (2015)
14. A. Mailhota, S. Lachance-Cloutierb, G. Talbota, A.C. Favrec, *J Hydrol* **476**, 7 (2013)
15. S. Coles, *An introduction to statistical modeling of extreme values* (Springer Series in Statistics, Springer-Verlag, London, 2001)
16. C. Onyutha, *Journal of Urban and Environmental Engineering* **6**, 2 (2012)
17. W. L. Shinyie, N. Ismail, A. A. Jemain, *Water Resour Manag* **28**, 11 (2014)
18. S. Wi, B. J. Valdés, S. Steinschneider, T. W. Kim, *Stoch Env Res Risk A* **30**, 2 (2016)
19. A. G. Yilmaz, I. Hossain, B. J. C. Perera, *Hydrol Earth Syst Sc* **18** (2014)
20. R. D. Gupta, D. Kundu, *J Stat Plan Infer* **137**, 11 (2007)
21. R. D. Gupta, D. Kundu, *J Stat Comput Sim*, **69** (2001)
22. B. Kaźmierczak, A. Kotowski, *J Hydrol* **525** (2015)
23. F. Laio, G. Di Baldassarre, A. Montanari, *Water Resour Res* **45**, 7 (2009)
24. S. H. Lee, S. J. Maeng, *Irrigat Drain* **52** (2003)
25. R. D'Agostino, M. A. Stephens, *Goodness of Fit Techniques* (Marcel Dekker, New York, 1986)
26. S. Hongjoon, J. Younghun, J. Changsam, H. Jun-Haeng, *Stoch Env Res Risk A* **26**, 1 (2012)
27. A. S. Hassan, *InterStat Electronic J* (2005)