Union of Mathematicians of Macedonia - ARMAGANKA

IX SEMINAR OF DIFFERENTIAL EQUATIONS AND ANALYSIS

and

1st CONGRESS OF DIFFERENTIAL EQUATIONS, MATHEMATICAL ANALYSIS AND APPLICATIONS

CODEMA 2020

Proceedings of the CODEMA 2020
Зборник на трудови од CODEMA 2020

Skopje, 2021

DETERMINATION OF A FLOOD WAVE PROPAGATION CAUSED BY HIGH INTENSITY RAINFALLS USING PROBABILITY TECHINQUES

ISBN 978-608-4904-09-0 UDC: 519.248:556.166

Mihail Naumovski, Daniel Velinov, Zoran Misajleski

Abstract. As the most unexpected occurrence in hydrotechnical engineering, high intensity rainfalls can lead to a significant increase in water level, creating a flood wave propagation in the catchment area. Therefore, it is necessary to predict the scenario that can be critical in terms of creating a high – performed hazard, caused by the flood wave propagation, which influence can affect not only on the catchment area but also on the human beings nearby. In that case, as better the hazard is assessed, the better will the consequences be treated.

Even upon the receipitation of high intensity rainfall data base, from the hydrometeorological station, it processing can be approached. The data base is obtained as a pluviogram in every tenth minute, from which can be created a hydrogram of the maximum rainfall intensity, whose values are given in time intervals not less than the pluviograph time. So, this kind of hydrograph is the base of determination of the maximum value of the rainfall intensity that can contribute to the creation of the flood wave propagation. Also, it is important that the high intensity rainfall data base is given for a period of tenth minutes pluviograms, measured at least of a year. Hence, the created hydrogram is the base of defining the frequency of occurrence of the maximum high intensity rainfall value, which is inversely proportional by the measured period. In this way, it can be determined the variables which will be used for creating a distribution of the probability density function. The probability density function is mostly based on a Gumbel distribution, so the results are the best possible simulated. The variables are based on the parameters of a time interval sequences, which refer to the measured period.

In this case, the benefit using probability density function is not just the determination of the value of maximum intensity rainfall which causes flood wave propagation, but also is the determination of the frequency of its occurrence. This helps to correct hazard assessment in order to design buildings in hydrotechnical engineering, which are permanent and lasting, serving the surroundings, not destroying the environment.

1. INTRODUCTION

Nowadays, engineers in the domain of hydrotechnical engineering are being more aware of determinating and managing a problem, such as high intensity rainfalls, than in the past. This is due to the contemporary investigations and technologies, developed on treating such an occurrence with a stochastic character, which can cause a catastrophical scenario followed by large-scale consequences. Nevertheless there are many circumstances that can undoubtedly, as in the engineering word is said, affect to the genetic code of the buildings, which are build just for the benefit of the mankind and the environment. And when once the genetic code is destroyed, than every idea of quality hydrotechnical engineering fades away.

The stochastic character of the high intensity rainfalls is generally due to the probability of their occurrence. Namely, there are no ways to decisively predict when and in which amount of can the high intensity rainfalls appear. This may cause distortion in the flow of watercourses and rivers, followed by increasing in water level. Such a phenomenon can inevitably form a flood wave propagation, which poses a huge threat as it is said in the abstract. For that purpose, scientific research activity till now is based on determinating the flood wave propagation, by processing the high intensity rainfalls data base, measured nearby the location in question. This location in question refers to the catchment area followed by the watercourse or the river in which the flood wave propagation is made.

2. Conditions and Measurements

Not always in the surroundings of the location in question can a hydrometeorological station be found. Therefore, it is necessary to obtain a clear representation of the location, more specifically to the distance from the nearest hydrometeorological station. If there are more of them the relevant measurements for the high intensity rainfalls data base will be taken from all of them, but before the processing, they will be optimized and fused into one. This fusion is always made of one of well know methods in hydrotechnical engineering, especially in hydrology: *Method of Tiesen* or *Method of arithmetic mean*. Some investigations show that both methods give approximate results, but the *Method of arithmetic mean* has greater degree of confidentiality in showing the real condition of the high intensity rainfalls data base.

Figure 1. Construction of the automatical pluviograph.

In every hydrometeorological station, there are not just instruments that measure one parameter, as precipitation which is directly connected with the high intensity rainfalls, but also parameters that affect on it such as: humidity, windiness, the number of sunniest days in the year and flora and fauna nearby. The instrument which measures the precipitation is called pluviograph. There are two types of pluviographs. The traditional one, which is being used by many of the engineers from the old school and the automatical one, which is used nowadays from the new aged school engineers. In this paper it is going to be presented the automatical pluviograph, because of its widely application not just in the world, but also in our country. As it is shown in the Figure 1, the construction of the pluviograph consists of a concrete foundation, metal casing – mostly as an aluminum casing, because of its endurance of environmental aggression, in which is installed the digital device and the top construction. The digital device is connected with the funnel at the top of the construction, in which the precipitation is collecting. The digital device is tuned at a minimum time interval of ten minute, so the pluviogram could not be registered at a smaller time interval. This means that when the high intensity data base is ready for processing, it should be put a point in time, where the measurement has been stared. It is quite important to distinguish the unevenness between the started point of the measurement and the started point of the precipitation. Mostly the started point of measurement is matched with the started point of the precipitation, at the time of 00:00 when initially the day starts. As it is shown in Figure 2., the digital device takes many circumstances which can affect on the intensity of the rainfalls, so the device may register the clear image of the outside precipitation scenario.

Figure 2. Digital device of the pluviograph. (Source: https://www.pce-instruments.com)

It is also important that the pluviograph must be isolate in a radius 4-5 meters around. That is because of the digital device, which works on signals. Namely, if there is another digital device around the signals of both devices can be mixed up and the measured data base will be incorrect and inappropriate for further processing. For proper functioning the pluviograph should be subjected to a regular inspection and maintenance.

3. Data Base Processing And Statistical Analysis

After the registered high intensity rainfalls data base is been given, the processing can be started. At first it is necessary to process the pluviograms in a table by ordering them in columns, which are represented as time interval columns. It means that the registered precipitation can be analyzed in a time frame from the smallest time interval of 10 minutes, to the largest of a day. This table has to be more organized in the detail, such as the Table 1, where are separated just the monthly maximum precipitation in a period of at least three years (the maximum period is not limited).

Table 1: Monthly maximum rainfall for characteristic time interval. (Source: [1])

Year	Month (n)	10’	20’	40’	60’	90’	120’	180’	300’	720’	1440’
2016	01	3.0	3.0	3.4	4.2	5.0	5.6	5.8	7.6	13.0	14.0
	02	1.8	2.4	3.4	3.8	4.2	4.4	5.2	5.4	5.7	7.8
	03	4.6	8.0	12.4	15.8	20.8	22.8	23.6	23.6	25.8	40.4
	04	4.0	4.8	6.4	8.0	10.0	10.2	10.8	11.8	17.0	17.6
	05	5.8	7.0	7.8	9.2	11.2	13.6	16.4	20.8	34.2	40.2
	06	7.6	8.4	10.4	12.4	12.6	13.8	18.2	22.8	22.8	22.8
	07	1.4	2.6	2.8	2.8	2.8	3.0	3.0	3.0	3.0	3.0
	08	16.0	18.6	19.2	19.2	19.2	19.2	19.4	19.8	19.8	19.8
	09	16.4	22.6	31.2	34.6	35.6	35.6	36.4	37.6	37.6	37.6
	10	2.2	2.6	3.2	3.6	4.2	4.6	6.6	8.6	9.4	9.8
	11	3.4	4.8	5.4	6.0	7.0	7.8	8.0	8.2	15.4	18.0
	12	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
2017	01	0.4	0.8	1.2	1.6	1.8	2.0	2.4	2.6	3.0	3.2
	02	0.8	1.2	1.6	2.2	2.4	2.4	2.4	3.0	3.6	3.6
	03	1.2	1.2	2.2	2.8	3.2	3.2	3.8	4.2	5.2	6.2
	04	2.6	4.8	6.0	7.2	9.0	11.6	13.4	16.8	16.8	19.0
	05	13.8	15.4	29.6	32.0	32.0	46.4	52.4	57.0	57.6	59.0
	06	3.6	4.6	4.8	6.6	8.4	8.4	10.4	10.4	10.4	20.2
	07	1.6	1.8	1.8	2.6	3.6	4.4	5.0	5.0	6.4	7.4
	08	2.0	3.0	4.2	5.0	6.6	7.6	7.2	9.4	10.0	10.0
	09	1.6	1.6	1.6	2.4	3.4	4.6	6.0	6.6	8.2	8.8
	10	1.4	2.2	4.0	4.6	5.2	6.0	6.8	8.0	10.2	10.2
	11	1.4	2.6	4.6	6.4	8.6	9.4	9.6	10.6	17.4	25.8
	12	6.2	6.2	7.6	9.4	11.0	12.8	14.0	17.6	22.4	37.6

2018	01	1.8	3.2	6.2	8.6	11.8	13.2	13.6	13.6	13.8	13.8
	02	1.8	2.2	2.4	3.4	4.2	5.2	6.6	8.8	11.6	16.0
	03	4.8	5.4	6.2	6.4	6.8	7.0	7.0	8.6	9.0	9.6
	04	2.0	3.4	4.6	5.4	6.6	6.6	6.6	6.8	7.4	7.4
	05	2.4	3.8	6.2	6.6	9.4	9.8	10.0	11.4	13.0	13.0
	06	8.6	12.8	13.2	13.2	13.4	13.6	13.8	14.2	14.2	17.4
	07	3.6	4.4	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	08	4.6	5.6	7.4	7.8	8.0	8.0	8.0	10.2	10.6	10.6
	09	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	10	0.4	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
	11	4.2	5.6	6.4	7.0	8.6	11.0	15.6	18.6	22.4	26.4
	12	1.0	1.8	3.0	3.6	4.4	5.4	7.8	8.6	9.4	11.0

For further more detailed examination it is necessary to analyze the time interval columns as sequences, for which we should calculate some necessary statisctical parameters, in order to obtain more precise and vivid picture for the problem under investigation. In sequel, we recall some basic terms and equations from descriptive statistics. The arithmetic mean of the data sequence $x_1, x_2, ..., x_{n-1}, x_n$ is given by

$$\bar{x} = \frac{x_1 + x_2 + ... + x_{n-1} + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i ,$$

and its standard deviation is given by

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \text{ , for } n \geq 30 .$$

The standard deviation and the variance s^2 is common measure of spread about the mean as center. The standard deviation s is zero, when there is no spread and gets larger as the spread increases. The mean and standard deviation are good descriptions for symmetric distributions without outliers. Additionally, concerning the standard deviation, when $n < 30$, we use the following formula

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} ,$$

measuring more accurately the standard deviation (spread).

Next, for better analysis of the data sequences, we will use scale parameter (depends on standard deviation), given by:

$$\alpha = \frac{1,282}{s} ,$$

and location parameter (depends on both, arithmetic mean and standard deviation) given by:

$$\beta = \bar{x} - 0,45 \cdot s .$$

Concerning the seeking for a correlation among two variables we stress the following notations and relations. If we think that a variable x may explain or even cause changes in another variable y, we call x an explanatory variable and y a response variable.

Let we have data of an explanatory variable x and a response variable y for n.

The correlation measures the direction and strength of the linear relationship between two quantitative variables. It will be abbreviated by r. Suppose that we have data on variables x and y for n individuals. The values for the first individual are (x_1, y_1), the values for the second individual are (x_2, y_2) and so on. The arithmetic mean and standard deviation of these data sequences are $\overline{x}$ and s_x for the data sequence $x_1, x_2, ..., x_n$, $\overline{y}$ and s_y for the data sequence $y_1, y_2, ..., y_n$. The correlation r between data sequences x and y is:

$$r = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{s_x} \right) \cdot \left(\frac{y_i - \overline{y}}{s_y} \right).$$

The least-squares regression line is given with the equation $y = a + bx$, with a slope:

$$b = r \frac{s_y}{s_x},$$

where s_x, s_y are standard deviations of the variable x, y, respectively and r is their correlation.

The interception a is calculated by

$$a = \overline{y} - b\overline{x},$$

where $\overline{x}$, $\overline{y}$ are means of the variables x, y, respectively.

With this regression line we obtain the predicted response y for any x.

There is a close connection between correlation and the slope of the least-squares regression line. The slope and the correlation always have the same sign.

The correlation r describes the strength of a straight-line relationship. This description takes a specific form: the square of the correlation, r^2, is the fraction of the variation in the values of y that is explained by the least-squares regression of y on x. The last one can be briefly written as:

$$r^2 = \frac{\text{var}\,iation\ \ in\ \ y\ as\ x\ \ pulls\ it\ along\ the\ line}{total\ \text{var}\,iation\ in\ observed\ values\ of\ \ y}.$$

From the last one, it is clear that we can always find a regression line for any relationship between two quantitative variables, but the usefulness of the line for prediction depends on the strength of the linear relationship. Hence, r^2 is almost as important as the equation of the regression line.

Analysing the quantity of the coefficient of correlation (r^2) we can determine the strength of the regression, adopted in hydrology. We have the following grades of the strength of the linear regression: 1) $R^2 < 0,3$ there isn't any dependence; 2) if $0,3 \le R^2 < 0,5$ there is some dependence; 3) if $0,5 \le R^2 < 0,7$ there is mild dependence; 4) if $0,7 \le R^2 < 0,9$ there is strong dependence; 5) if $R^2 \ge 0,9$ there is very strong dependence.

For comprehensive approach on this statistical ideas, the interested reader may consult [5] and [7]-[11].

4. Analysis Using Probability Techniques

We start this section with the following table, in which the basic statistical parameters are summarized, in each specific considered time interval.

Table 2. Statistical parameters (Source: [1])

Statistical Parameters	10’	20’	40’	60’	90’	120’	180’	300’	720’	1440’
$\bar{x}$	3.84	4.94	6.56	7.51	8.52	9.57	10.60	11.86	13.66	15.92
σ	4.03	4.98	6.94	7.50	7.71	9.21	10.03	10.81	11.42	13.12
α	0.32	0.26	0.18	0.17	0.17	0.14	0.13	0.12	0.11	0.10
β	3.36	4.48	6.08	7.05	8.11	9.13	10.17	11.45	13.28	15.54

Using regression analysis on the previously presented sequences of data, organized into certain time intervals, we can obtain the following plot given on Figure 3. The following important results (not presented of the Figure 3), concerning the regression lines, can be stated:

- For $10'$, the equation for linear regression is given by
$$y = 2,4464x + 18,677,$$
with strength of the linear regression $r^2 = 0,8709$.
- For $20'$, the equation for linear regression is given by
$$y = 2,4115x + 20,138,$$
with strength of the linear regression $r^2 = 0,865$.
- For $40'$, the equation for linear regression is given by
$$y = 2,2798x + 23,593,$$
with strength of the linear regression $r^2 = 0,8452$.
- For $60'$, the equation for linear regression is given by

$$y = 2{,}3318x + 22{,}59,$$

with strength of the linear regression $r^2 = 0{,}8635$.

- For $90'$, the equation for linear regression is given by

$$y = 2{,}483x + 19{,}725,$$

with strength of the linear regression $r^2 = 0{,}8951$.

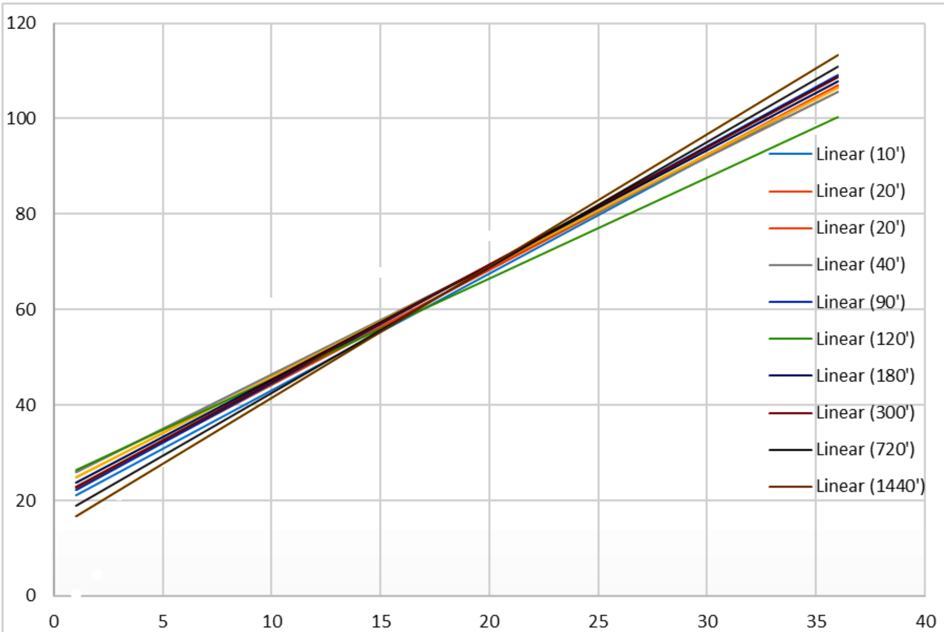

Figure 3. Regression analysis of months and maximum rainfalls for the specific time intervals

- For $120'$, the equation for linear regression is given by

$$y = 2{,}11x + 24{,}311,$$

with strength of the linear regression $r^2 = 0{,}6133$.

- For $180'$, the equation for linear regression is given by

$$y = 2{,}4024x + 31{,}322,$$

with strength of the linear regression $r^2 = 0{,}8942$.

- For $300'$, the equation for linear regression is given by

$$y = 2{,}4574x + 20{,}407,$$

with strength of the linear regression $r^2 = 0{,}8979$.

- For $720'$, the equation for linear regression is given by

$$y = 2{,}6272x + 16{,}363,$$

with strength of the linear regression $r^2 = 0,9306$.

- For $1440'$, the equation for linear regression is given by
$$y = 2,7582x + 13,926,$$
with strength of the linear regression $r^2 = 0,918$.

We can conclude that in all obtained linear regressions (except in the case of $120'$), the strength of the linear regression is strong, i.e. there exists strong dependence between considered parameters. Also, we can deduce that the spread of the values is not quite big. An interesting fact that affects on a weak correlation at $120'$ linear regressions, is the error that occurs as a consequence of the digital deviced sleep time. Namely, every two hours per a day the digital device has to sleep not more than a minute, so the sleep time affects on a very small part of the data base not to be registered. This is most pronounced in $120'$ linear regressions, because it is about a punctual part of the time when the sleep time begins. Such a phenomenon does not affect on the larger time intervals, because of the significant increase of the data base.

Next, statistical parameters for certain time period (in months) and certain time intervals, or more precisely, a monthly maximum rainfall [mm] for characteristic probabilities is given.

Table 3. Statistical parameters (Source: [1])

p(z) (%)	**Months (n)**	10’	20’	40’	60’	90’	120’	180’	300’	720’	1440’
50	2	4.51	5.91	8.09	9.19	10.32	11.76	13.05	14.55	16.55	19.29
20	5	8.10	10.78	14.61	15.83	17.58	20.02	22.10	24.50	26.97	32.06
10	10	11.34	13.33	18.60	21.38	22.45	27.13	29.94	32.89	33.61	38.64
4	25	13.62	17.22	25.68	28.79	29.49	33.89	35.84	39.55	45.83	52.16
2	50	15.72	20.05	28.58	31.25	31.84	38.28	44.09	49.38	52.94	57.47
1	100		22.45	30.02	34.05		43.90	49.45	54.29	56.49	

On Figure 4, probability distributions of monthly maximum rainfall data i.e. the regression lines for p(z)-monthly maximum rainfalls for specific time intervals, for each considered time interval, are given.

The following facts, can be stated, concerning the obtained linear regressions. We have obtained the following:

- For $10'$, the equation for linear regression is given by
$$y = -0,214x + 14,399,$$
with strength of the linear regression $r^2 = 0,8911$.
- For $20'$, the equation for linear regression is given by
$$y = -0,2986x + 19,286,$$
with strength of the linear regression $r^2 = 0,8294$.

- For $40'$, the equation for linear regression is given by
$$y = -0,4253x + 27,097,$$
with strength of the linear regression $r^2 = 0,8511$.
- For $60'$, the equation for linear regression is given by
$$y = -0,4748x + 30,3,$$
with strength of the linear regression $r^2 = 0,8488$.

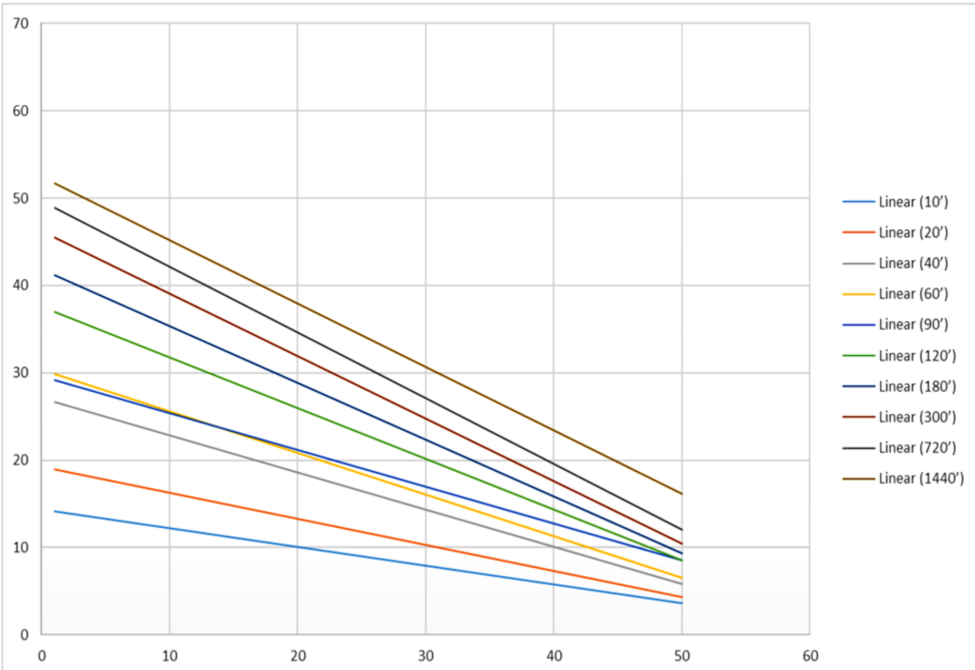

Figure 4. Probability distributions of monthly maximum rainfall data

- For $90'$, the equation for linear regression is given by
$$y = -0,42x + 29,56,$$
with strength of the linear regression $r^2 = 0,881$.

- For $120'$, the equation for linear regression is given by
$$y = -0,5806x + 37,581,$$
with strength of the linear regression $r^2 = 0,8315$.
- For $180'$, the equation for linear regression is given by
$$y = -0,6507x + 41,847,$$
with strength of the linear regression $r^2 = 0,8048$.
- For $300'$, the equation for linear regression is given by

$$y = -0,7173x + 46,262,$$

with strength of the linear regression $r^2 = 0,8026$.

- For $720'$, the equation for linear regression is given by

$$y = -0,7508x + 49,618,$$

with strength of the linear regression $r^2 = 0,8101$.

- For $1440'$, the equation for linear regression is given by

$$y = -0,7255x + 52,402,$$

with strength of the linear regression $r^2 = 0,8564$.

We can conclude that in all obtained linear regressions the strength of the linear regression is strong, i.e. there exists strong dependence between considered parameters. Such a representation of the high intensity rainfall data base, alludes to clear picture of determinating the maximum value of the precipitation that can cause a wave propagation in the catchment area. In essence, that was the purpose of this paper, not only to determinate a flood wave propagation caused by intensity rainfalls, but also to make it with a high degree of confidentiality.

5. CONCLUSIONS

Undoubtedly, the connection between engineering and mathematics can create models that can be essential in determinating processes with very small probability of occurrence. Therefore, with this paper we want to show the strength of a model that can predict a cathastrophical scenario caused by flood wave propagation, using a probability technique in order to provide protection and safety. And as an engineer challenge: The more unexpected, the better!

References

[1] B. Susinov, M. Naumovski, J. Josifovski: *Hydrological analysis of high intensity rainfalls over Topilnica tailing dam*, WMHE 2019 (16th International Symposium, Skopje, Macedonia, 2019.

[2] *PCE-FWS 20 Water Station*, www. pce-instruments.com, 05.05.2019.

[3] Z. A. Ivkovic, Matematicka statistika, Naucna knjiga, Beograd, 1982.

[4] С. Јадровски, Примена на статистички методи во хидрологијата, Магистерска работа, Градежен факултет, Скопје, 2018.

[5] Р. Малчески, В. *Малческа, Теорија на веројатност,* Армаганка, Скопје, 2019.

[6] M. Marriott, *Civil Engineering Hydraulics*, John Wiley & Sons, Inc, Singapore, 2016.

[7] J. H. McDonald, K. W. Dunn, Statistical tests for measures of colocalization in biological microscopy. Journal of Microscopy 252 (2013), 295-302.
[8] B. H. McCardle, Lines, models, and errors: Regression in the field. Limnology and Oceanography 48 (2003), 1363-1366.
[9] D. S. Moore, The Basic Practice of statistics, W.H. Freedman and Company, New York, 2010.
[10] I. Pavlic, Statisticka teorija i primjena, Tehnicka knjiga, Zagreb, 1977.
[11] A. Papoulis, Probability, Random variables and Stochastic Processes, International Student Edition, New York, 1965.
[12] Ц. Поповска, В. Ѓешовска, *Хидрологија: теорија со решени задачи*, Градежен факултет, Скопје, 2012.
[13] Ц. Поповска, С. Крстиќ, *Прирачник за реставрација на реки*, UNDP, Скопје, 2010.
[14] S. Prohaska, V. Ristić, Hidrologija kroz teoriju i praksu, Univerzitet u Beogradu, Rudarsko-geološki fakultet, Beograd, 1996.
[15] S. Prohaska, Hidrologija I, Univerzitet u Beogradu, Rudarsko-geološki fakultet, Beograd, 2006.
[16] S. Prohaska, Hidrologija II, Univerzitet u Beogradu, Rudarsko-geološki fakultet, Beograd, 2006.
[17] N. Kannan, J. P. Keating, R. L. Mason, A comparison of classical and inverse estimators in the calibration problem. Communications in Statistics: Theory and Methods 36 (2007), 83-95.
[18] T. Lwin, J. S. Maritz, An analysis of the linear-calibration controversy from the perspective of compound estimation. Technometrics 24 (1982), 235-242.
[19] Z. Shkoklevski, B. Todorovski B, *Intensity rainfalls in Republic of Macedonia*, Ss. Cyril and Methodius University, Faculty of Civil Engineering, Skopje, 1993.

Faculty of Civil Engineering, Ss Cyril and Methodius University, Skopje, Macedonia
E-mail address: naumovskimgf@outlook.com
Faculty of Civil Engineering, Ss Cyril and Methodius University, Skopje, Macedonia
E-mail address: velinovd@gf.ukim.edu.mk
Faculty of Civil Engineering, Ss Cyril and Methodius University, Skopje, Macedonia
E-mail address: misajleski@gf.ukim.edu.mk

Publisher
Union od Mathematicians of Macedonia – ARMAGANKA

Editor in chief
Prof. d-r Aleksa Malcheski

CIP - Каталогизација во публикација
Национална и универзитетска библиотека "Св. Климент Охридски", Скопје

51(082)

PROCEEDINGS of CODEMA 2020 = Зборник на трудови од CODEMA 2020. -
Skopje : Armaganka, 2021. - 375 стр. : илустр. ; 25 см

Текст на мак. и англ. јазик. - Фусноти кон текстот. - Библиографија кон трудовите

ISBN 978-608-4904-09-0

а) Математика -- Зборници

COBISS.MK-ID 53570309