FORECASTING RELATIONS OF THE FLASH FLOODS MAXIMUM DISCHARGES ON THE BASIS OF THE CORRELATION WITH THE RELEASING FACTORS

Rodica Mic, Ciprian Corbuş

National Institute of Hydrology and Water Management, Bucureşti, Romania,
e-mail: rodica.mic@hidro.ro; corbus@hidro.ro

Abstract: The lag-time to the rainfall, in case of small areas basins (S<100÷200km²) is from about 15 minutes to a few hours. Therefore, the afferent time for the flash floods forecasting is very short, so that this activity should be concentrated especially on the danger warning system.

In this situation, in order to assure real time warnings, it is not recommended to use the hydrological modelling; on the contrary it is preferable to rapidly assess the precipitation values supplied by the automatic stations, corroborated with the radar information and the comparing of these precipitation values with different critical thresholds of the precipitation associated to flooding level, pre-established.

The main purpose of this work is to increase the warning time in case of flash floods in torrential regime basins, by elaborating a simple methodology of inundation warnings based on the correlation between the peak discharges of the flash floods with causal factors.

The elements that stand on the basis of the maximum discharge forecast relations are: the precipitation, the duration of the rain, the soil previous humidity and the base discharge.

Keywords: flash floods, hydrological forecast

SCHEITELABFLÜSSE PROGNOSE RELATIONEN VON DER STURZFLUTEN AUFGRUND DER MIT DEN AUSLÖSENDEN FAKTOREN KORRELATION

Kurzfassung: Die Verzögerungszeit des Regens, im Falle der kleinen Zuflussgebiete (S<100÷200km²), variiert von 15 Minuten bis einige Stunden. Folglich, die notwendige Zeit der Sturzflutensprognose ist sehr kurz und darum muss sich diese Aktiwität besonders auf dem Warnungssystem der von den Sturzfluten representiert Gefahr konzentrieren.

Wegen dieser Situation, für rechtzeitige Warnungenssicherungen, empfielt es sich, nicht Flussmodell zu benutzen, sondern ist es vorzuziehen, eine schnelle Bewertung der Niederschlägeswerte, von den automatischen Stationen geliefert, mit der Radarsinformation korreliert und diese Werte mit den Werten der verschiedenen kritischen Niederschlägsschwelle, mit früheren festgelegten Schutzwassersspiegel zuziehend, gleichzusetzen.

Der Arbeitshauptzweck ist die Vergrösserung der Antizipationszeit, im Falle der Sturzfluten, auf Einzugsgebiete mit strömenden Regime, durch die Ausarbeitung einer einfache Methodologie für die Ausgabe der Überflutungenswarnungen, die auf der Scheitelabflüsseskorrelation der Sturzfluten mit den auslösenden Faktoren beruht.

Die Grundeselemente der Prognosesrelationen sind: die Niederschläge, die Regenzeit, die vorherige Grundesfeutigkeit und den Basisabfluss.

Schlüsselworte: Sturzfluten, hydrologische Prognose.

Introduction

The flash floods represent one of the most frequent natural disasters in many countries.

These floods have an extremely rapid evolution and constitute an unexpected element for the inhabitants of the flooded area. In addition, the flooding produced by these floods is difficult to be evaluated, because of the space and time variability of the rainstorm.

From the operational forecast point of view the flash floods are characterized by a very short lag time. Moreover, in case of flash floods, the forecasting activity is concentrated especially on the danger warning system.

A hydrological forecast in real time supposes both the cognition of the hydro-meteorological situation from the basin as exactly as possible and the existence of some rapid methodologies for estimating the maximum discharges at the reaching of some precipitation thresholds *(Georgakakos, 1996)*.

1. Utilized data

For this study were chosen 8 sub-basins *(figure 1)*, from the drainage basin of Crisuri, with a rapid flow regime and a small basin surface (S<100÷200km^2).

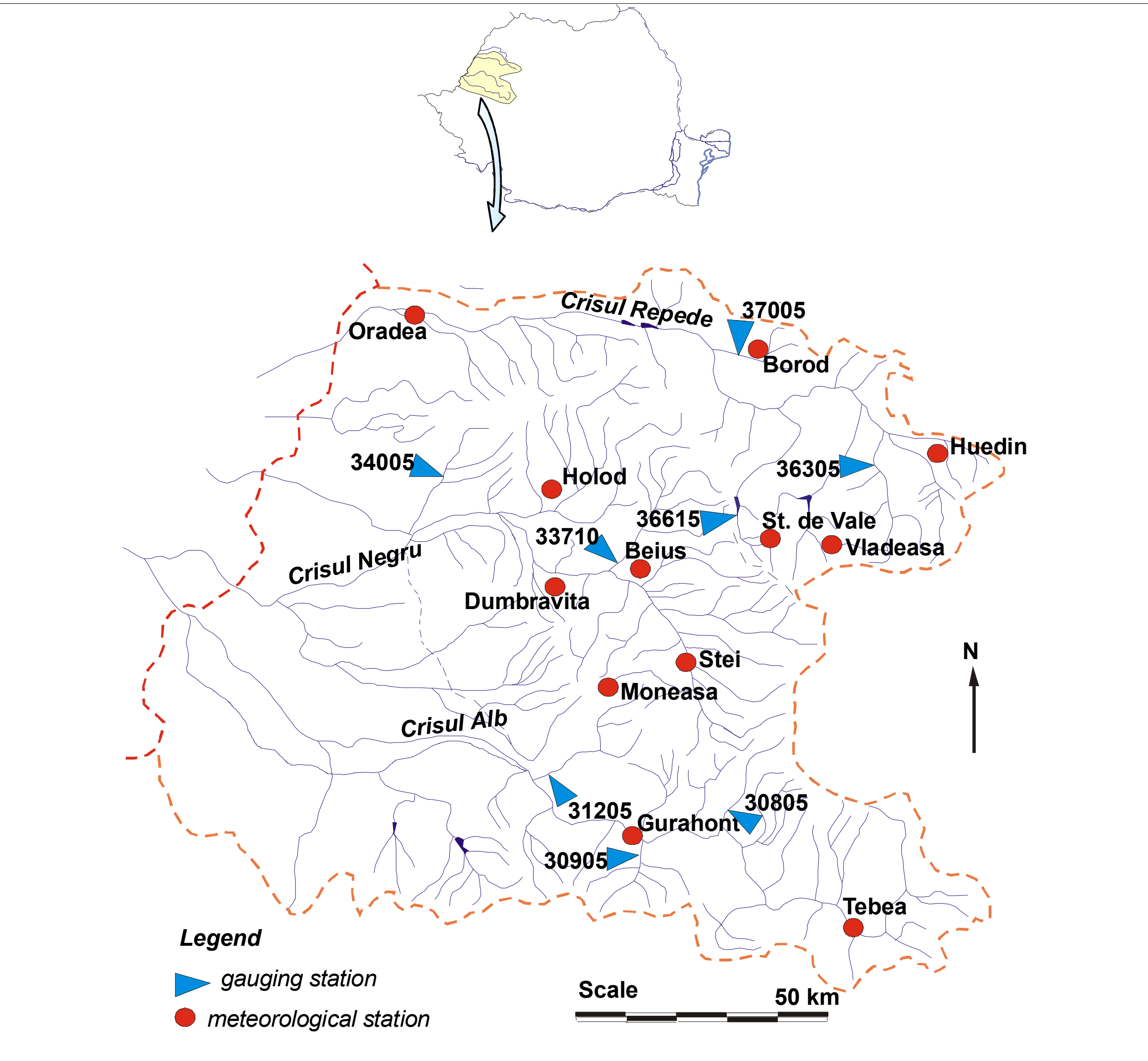

Figure 1. Drainage basin of Crişuri
Selected gauging stations and meteorological stations

In table *1* are presented for the selected gauging stations, the surface and the mean altitude of the basins, and also the mean increase time of the floods.

There were also selected some meteorological stations situated on the drainage basins or close to them and considered to be representative from precipitation on basin point of view.

Table 1. Selected hydrometrical stations

River – Gauging station	Code	S (km^2)	Hm (m)	t_c (h)
V. de Lazuri - VÂRFURILE	30805	78.6	648	15
Sighişoara - BRAZI	30905	120	455	6
Sebiş - SEBIŞ	31205	209	432	16
V. Roşie - POCOLA	33710	267	427	17
V. Nouă - HUSASĂU DE TINCA	34005	136	217	16
Henţ - MORLACA HENŢ	36305	209	1014	20
Iad - LEŞU AMONTE	36615	51	1088	12
Borod - TOPA de CRIS	37005	115.5	484	6

1.1. Rain-flow events selection

In a first stage there were chosen all floods whose maximum discharges has values over or close to the discharge corresponding to the alarm level.

If, after this first criterion, the events are much fewer, there were also selected floods with lower maximum discharges, so that for each section are analysed about 20-30 rainfall-runoff events.

In order to determine as exactly as possible the runoff volume, when the rain-flow events were selected, it was also taken into consideration the fact that the selected floods must be complete from hydrometrical point of view.

To assess the duration of the precipitation, another criterion taken into consideration was that the selected floods were produced during the functioning of the recording rain gage at the meteorological stations.

Table 2 presents the total number of analysed events and also the number of events used for the calibration and validation of the computed relationships of the maximum discharges.

Table 2. Number of the analysed rain-flow events

River – Gauging station	Code	S (km^2)	No. of events		
			Total	Calibration	Validation
V. Satului - BUCEŞ	30205	100	12	8	4
V. de Lazuri - VÂRFURILE	30805	78.6	21	15	6
Sighişoara - BRAZI	30905	120	23	16	7
Sebiş - SEBIŞ	31205	209	33	15	8
V. Roşie - POCOLA	33710	267	17	12	5
V. Nouă - HUSASĂU DE TINCA	34005	136	20	14	6
Henţ - MORLACA HENŢ	36305	209	27	18	9
Iad - LEŞU AMONTE	36615	51	20	14	6
Borod - TOPA de CRIS	37005	115.5	21	14	7

1.2. Elements used in the estimated relations of the maximum discharge

The elements that are considered for the maximum discharge forecasting relations of the flood Q_{max} *(m^3/s)* are:

- Precipitation depth, h_p *(mm)*;
- Rain duration, D_p *(h)*;
- Antecedent soil moisture, U_i *(mm);*
- Base runoff, Q_b *(m^3/s).*

The precipitated depth h_p *(mm)* was considered for one or more meteorological stations, considered to be representative for each analysed basin.

The precipitation duration D_p *(h)*, which releases the flood, was estimated after the simultaneous analysis of flood and hourly distribution of the rain registered to the meteorological station, considered being significant for each analysed basin.

The antecedent soil moisture U_i *(mm)*, namely the soil humidity at the moment of the flood release, was computed by a balance method, which estimates that the water quantity remained in the soil during the given time period is the difference between: the quantity of mean precipitation on basin, depth of runoff and evapotranspiration.

Computing formula of the antecedent soil moisture *(Simota & Mic, 1993)* is the following:

$$U_i = (1-\alpha)\cdot\left(\sum_{i=1}^{10} P_i\right) - N\cdot E \tag{1}$$

in which: α - the flow coefficient; E - the evapotranspiration; P_i - the daily mean precipitation on basin; N - the number of days without precipitation.

The α coefficient is considered being a function of $\sum P_i$, and E depends on the month when the flood occurs.

2. Analysis of the link between the maximum discharge and the determinant characteristic elements

On the basis of the selected data, for each of the 8 hydrometrical stations was analysed the dependence of the maximum discharges of the flood waves (Q_t), on different characteristics of the generating rain, antecedent soil moisture and base runoff and which one of these characteristics has the biggest influence upon the maximum discharge.

A measure of the "goodness of fit" is the correlation coefficient. To explain the meaning of this measure, it is define the standard deviation (S_t), which quantifies the spread of the data around the mean *(Dauphin, 1992)*:

$$S_t = \sum_{i=1}^{N}(\bar{y} - y_i)^2 \tag{2}$$

where the average of the data points ($\bar{y}$) is simply given by :

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{3}$$

The quantity S_t considers the spread around a constant line (the mean) as opposed to the spread around the regression model. This is the uncertainty of the dependent variable prior to regression. It is also define the deviation from the fitting curve as:

$$S_r = \sum_{i=1}^{N}(y_i - f(x_i))^2 \tag{4}$$

This quantity likewise measures the spread of the points around the fitting function. Thus, the improvement (or error reduction) due to describing the data in terms of a regression model can be quantified by subtracting the two quantities. Because the magnitude of the quantity is dependent on the scale of the data, this difference is normalized to yield:

$$r \equiv \sqrt{\frac{S_t - S_r}{S_t}} \tag{5}$$

where: r is defined as the correlation coefficient.

As the regression model better describes the data, the correlation coefficient will approach unity.

By the partial correlations computation it has been ascertained that, in the most cases, the maximum discharge of the flood waves does not depend only on a single considered, but on two or more *(table 3)*.

Table 3. Dependence of the maximum discharge Q_t on different characteristics of the generating rain (H_p, D_p), antecedent soil moisture (P_a) and base runoff (Q_b)

Independ. variable	Computed formula	r
Valea de Lazuri - VÂRFURILE		
H_p	$Q_t = 1 / (0.20081999 - 0.03994878 * ln(H_p))$	0.831
D_p	$Q_t = 16.035533 (1 - exp(-0.35520912 D_p)$	0.537
P_a	$Q_t = 11.623875 + 0.15781238 P_a$	0.192
Q_b	$Q_t = 14.470705 * Q_b^{(0.31548859 / Qb)}$	0.402
Sighişoara - BRAZI		
H_p	$Q_t = 19.829622 * 1.0184108^{Hp}$	0.287
D_p	$Q_t = 7.9322475 * D_p^{(6.5661409 / Dp)}$	0.426
P_a	$Q_t = 1 / (-0.0023043848 P_a + 0.08395441)$	0.886
Q_b	$Q_t = 345.08176 * exp((-(3.8208357 - Q_b)^2) / (2 * 0.68735398^2))$	0.865
Sebiş - SEBIŞ		
H_p	$Q_t = 15.760759 + 0.98242748 H_p$	0.832
D_p	$Q_t = 27.606309 + 1.4790027 D_p$	0.575
P_a	$Q_t = 40.514192 + 34.266878 * cos(0.96753018 P_a + -1.6473246)$	0.726
Q_b	$Q_t = 58.15321 / (1 + 58.15321 * exp(-3.0520454 Q_b))$	0.542
Valea Roşie - POCOLA		
H_p	$Q_t = -0.36216661 (-133.45852 - exp(0.084141265 H_p))$	0.817
D_p	$Q_t = 1 / (0.024795616 - 0.0040890931 * ln(D_p))$	0.250
P_a	$Q_t = (58.431333 - 0.72767901 P_a) / (1 - 0.0095076735 P_a -0.0004077003 P_a^2)$	0.884
Q_b	$Q_t = -110.68565 + 118.85963 Q_b - 25.613628 Q_b^2 + 1.778826 Q_b^3$	0.911
Valea Nouă - HUSASĂU DE TINCA		
H_p	$Q_t = -26.330923 + 2.7997453 H_p - 0.04625257 H_p^2 + 0.00027075744 H_p^3$	0.963
D_p	$Q_t = (-9.3474119 + 3.7717792 D_p) / (1 - 0.31962447 D_p + 0.030853629 D_p^2)$	0.522
P_a	$Q_t = 41.199927 (1 - exp(-0.060052158 P_a)$	0.302
Q_b	$Q_t = 24.231073 + 23.473742 * cos(3.3710656 Q_b + 2.5396367)$	0.581
Henț - MORLACA HENȚ		
H_p	$Q_t = 15.841018 + 0.40476923 H_p$	0.452
D_p	$Q_t = 35.223221 - 0.38839101 D_p$	0.263
P_a	$Q_t = 18.33618 + 0.81163033 P_a$	0.489
Q_b	$Q_t = (19.226409 - 2.3489978 Q_b) / (1 - 0.17812628 Q_b + 0.006727063 Q_b^2)$	0.673
Iad - LEŞU AMONTE		
H_p	$Q_t = 6.5574791 + 0.39856892 H_p$	0.899
D_p	$Q_t = (14.789763 - 0.44716952 D_p) / (1 - 0.063804084 D_p + 0.0011103085 D_p^2)$	0.524
P_a	$Q_t = 43.609575 - 6.283673 P_a + 0.41637128 P_a^2 - 0.0078025015 P_a^3$	0.759
Q_b	$Q_t = 23.414143 + 7.0990081 * cos(13.894476 Q_b - 9.7706268)$	0.444
Borod - TOPA de CRIS		
H_p	$Q_t = 92.125577 * exp((-(75.834427 - H_p)^2)/(2 * 30.826601^2))$	0.780
D_p	$Q_t = 44.810415 - 3.2527198 D_p + 0.14934257 D_p^2$	0.535
P_a	$Q_t = 47.679318 - 1.3626913 P_a - 0.18866449 P_a^2 + 0.0093902641 P_a^3$	0.693
Q_b	$Q_t = 24.948018 Q_b^{(0.44379541 Qb)}$	0.742

That is why it was tried to be used the linear multiple regressions for estimating the maximum discharge depending on the characteristics of the rain (h_p, D_p) and on the initial state of the basin (P_a, Q_b).

3. Computation of the maximum discharge by multiple linear regressions

Starting from the relation between the flood maximum discharge and the characteristic elements that generate the flood, which takes the following form:

$$Q_{max} = e^{\alpha_0} \cdot C_1^{\alpha_1} \cdot C_2^{\alpha_2} \cdot ..... \cdot C_p^{\alpha_p} \cdot \varepsilon^{\varepsilon} \tag{6}$$

where: α_1 , α_2 , $...\alpha_m$ are the regression coefficients, α_0 is a constant and ε represents the error of estimation or the residue.

The equation *(2)* is transformed logarithmically and the estimated value of Q_{max}, noted $\hat{Q}_{max}$, are the following form:

$$log\,\hat{Q}_{max} = \alpha_0 + \alpha_1\, log\, C_1 + \alpha_2\, log\, C_2 ... \alpha_p\, log\, C_p + \varepsilon \tag{7}$$

For every section, the computation of the regressions coefficients α_0 , α_1 , α_2 , $...\alpha_m$ was made using about 2/3 of the analysed floods, and the rest were used for validation of the obtained computing relations.

The determination of the regression coefficients was based on minimizing the relation *:*

$$\sum_{i=1}^{n} \left(\hat{Q}_{max} - Q_{max}\right)^2 = min \tag{8}$$

where: n is the number of events used in computations.

3.1. Checking the pertinence of each independent variables

The checking of the correlation that exists, between the different precipitation characteristics and those of the flood produced' by them was made by computing the R^2 correlation coefficient.

In order to establish the contribution of each Xj independent variable, there were computed the partial correlation coefficients, which express the intensity of the link between the dependent variable one of the independent variables then when it is eliminated the influence of the other independent variables upon them *(Duband, 1975).*

Taking as example the partial correlation between Y and X_1 independent variable, the residual error from the relation of defining the multiple linear regressions takes the following forms:

$$\begin{aligned} \xi &= Y - (a_2 X_2 + a_3 X_3 + ... + a_m X_m + a_0) \\ \psi &= X_1 - (b_2 X_2 + b_3 X_3 + ... + b_m X_m + b_0) \end{aligned} \tag{9}$$

a and b coefficients were computed by the least-squares method.

The partial correlation coefficient between Y and X_1 variable will take the following form:

$$r_{Y,X_1,X_2,...Xm} = \frac{\sum_{i=1}^{n} \xi_{1i} \cdot \psi_{1i}}{\sqrt{\sum_{i=1}^{n} \xi_{1i}^2} \cdot \sqrt{\sum_{i=1}^{n} \psi_{1i}^2}} \tag{10}$$

In order to know if an independent variable X_j is or is not significant, in other words, if has a real contribution in estimation of the Y variable, Student test could be made upon the partial correlation coefficients.

Having a dependent variable defined through m independent variables, a θ partial coefficient can be considered, from which there were eliminated the influences of m-1 variables, whose probability density can be expressed as follows *[Michel, 1982]:*

$$g_n(\theta) = \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma\left(\frac{n-m}{2}\right)}{\Gamma\frac{(n-m-1)}{2}} \cdot (1-\theta^2)^{\left(\frac{n-m-3}{2}\right)} \tag{11}$$

The transformed variable $(\sqrt{n-m-1})\cdot\theta/(\sqrt{1-\theta^2})$ follows a Student law with (*n-m-1*) degree of liberty. Therefore, if t_p is the value corresponding to the *p%* probability, there are *p* chances from 100 to obtain $|t|\cdot > t_p$ or $\theta > t_p \Big/ \sqrt{t_p^2 + n - m - 1}$.

Considering *1%* possibility of error, for having only one chance from 100 to mistake, considering that it is a link between Y and X_1, X_2,….X_m variables, it is necessary that: $\theta > t \ / \sqrt{t^2 + n - m - 1}$ with $|t| = 0.99$. If *(n-m-1)* >10, it will be taken as limit value: $1/(1+\dfrac{(n-m)^2}{6.7\cdot(n-m+2)^2})$

It has been ascertained that, on the whole of the analysed sections, the four characteristics are all significant *(table 4)*.

Table 4. Partial correlation coefficients

River – Gauging station	*Independent variables*	r_{Y,X_1,X_2,X^3,X_4}	*Limit value*
Valea de Lazuri - VÂRFURILE	h_p D_p P_a Q_b	0.876 -0.633 0.257 0.004	0.33
Sighişoara - BRAZI	h_p D_p P_a Q_b	0.844 -0.820 0.771 -0.629	0.301
Sebiş - SEBIS	h_p D_p P_a Q_b	0.908 -0.203 0.746 0.585	0.301
Valea Roşie - POCOLA	h_p D_p P_a Q_b	0.898 -0.426 0.330 0.902	0.385
Valea Noua - HUSASAU DE TINCA	h_p D_p P_a Q_b	0.970 -0.174 0.373 0.180	0.346
Henț - MORLACA HENT	h_p D_p P_a Q_b	0.817 -0.737 0.067 0.665	0.256
Iad - LESU AMONTE	h_p D_p P_a Q_b	0.969 -0.815 -0.351 0.614	0.346
Borod - TOPA DE CRIS	h_p D_p P_a Q_b	0.890 -0.552 0.390 0.733	0.33

3.2. Computation of the regression coefficients

As a result of an optimisation computation there were obtained the values of the regressions coefficients with the following form:

$$Q_{max} = e^{\alpha_0} \cdot H_p^{\ \alpha_1} \cdot D_p^{\ \alpha_2} \cdot P_a^{\ \alpha_3} \cdot Q_b^{\ \alpha_4} \qquad (12)$$

The values of these coefficients computed for the analysed 8 sub-basins are presented in *table 5.*

Table 5. The regression coefficients established for the computation of the maximum discharge

River – Gauging station	*Meteorological station used*	α_0	α_1	α_2	α_3	α_4
Valea de Lazuri - VÂRFURILE	*GURAHONȚ*	0.811	0.772	-0.26	0.02	0.1
Sighişoara - BRAZI	*GURAHONȚ*	0.217	1.384	-0.910	0.120	1.111
Sebiş - SEBIŞ	*ŞTEI*	0.956	0.707	-0.1	0.152	0.433
Valea Roşie - POCOLA	*HOLOD*	-0.714	1.074	-0.2	0.02	1.065
Valea Nouă - HUSASĂU DE TINCA	*HOLOD*	-1.16	1.248	-0.16	0.122	0.276
Henț - MORLACA HENȚ	*VLĂDEASA*	-0.09	1.005	-0.4	0.03	0.53
Iad - LEŞU AMONTE	*STÂNA DE VALE*	0.007	1.097	-0.42	0.04	0.016
Borod - ȚOPA DE CRIŞ	*BOROD*	-1.6	1.549	-0.38	0.09	0.753

3.3. Checking the relations of computation for the maximum discharge

For emphasizing the fact that on the whole of the analysed floods, the maximum discharges computed by the regression method are close to those observed, it was computed the relative root mean square deviation, at each analysed section, after the following relation:

$$rRMSE(\%) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{(Q_{max}^{obs})_i - (Q_{max}^{calc})_i}{(Q_{max}^{obs})_i}\right)^2} \cdot 100 \qquad (13)$$

where: *n* represents the number of analysed floods.

The figure 2 shows that the relative root mean square error between the computed and observed values is situated in acceptable limits, both for the floods utilized for the computation of the regressions coefficients and for the floods used for validation.

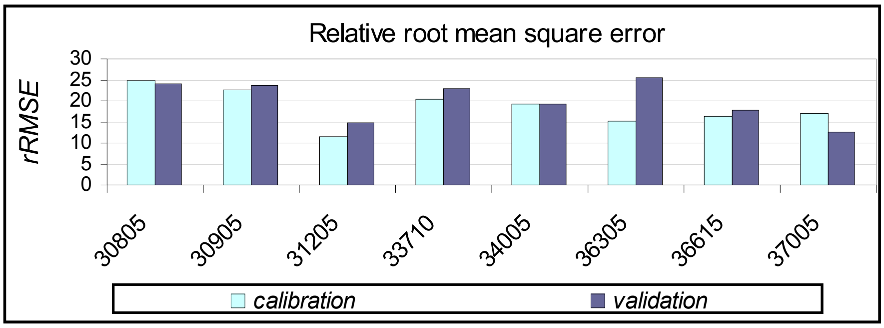

Figure 2. The relative mean square deviation between the observed and computed maximum discharge

Conclusions

The relations obtained through the method presented above could be used for forecasting the maximum discharges in case of flash floods.

According to the checking made for a part of the analysed floods events, others than those used in computations, the error of estimation of these maximum discharges is small than 20-25%.

In order to use these computation relations in the operative forecasting has been created an automatic computation sheet using the EXCEL program. For the running of the EXCEL sheet it is necessary the daily updating of the mean precipitation on basin, during 10 days before the flood occurring, and the introduction of the precipitation depth $h_p(mm)$, the rain duration $D_p(h)$ and the *base runoff* $Q_b(m^3/s)$.

Referents

Dauphin J.C. (1992) IDAMS - the Reference Manual, UNESCO
Duband D. (1975) Ridge regression, Note interne, EDF-DTG, France
Georgakakos K. (1996) *On the Design of National, Real_time Warning Systems with Capability for Site-Specific, Flash Flood Forecasts* - Iowa Institute of Hydraulic Research
Michel C. (1982) *Perfectionnement en Hydrologie* - Cemagref Antony
Simota M., Mic R. (1993) *Evaluation des seuils critiques de précipitation qui declenchent l'alerte hydrologique ŕ l'interieur d'un systeme informatique d'annonce et de prevision des crues.* - Rencontre Hydrologique Franco-Roumaines, Tulcea 1993
*** (1994) *Guide to hydrological practices -* World Meteorological Organization