BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iaşi
Tomul LX (LXIV), Fasc. 3, 2014
Secţia
CONSTRUCŢII. ARHITECTURĂ

A PARTICULAR FORECASTING CASE FOR THE MONTHLY FLOW RATES OF THE PRUT RIVER

BY

DORIN COTIUŞCĂ ZAUCĂ*, RALUCA-CATRINEL GIURMA-HANDLEY and PETRU CERCEL

"Gheorghe Asachi" Technical University of Iaşi
Faculty of Hydrotechnics, Geodesy and Environmental Engineering

Received: June 11, 2014
Accepted for publication: June 30, 2014

Abstract. The paper presents a forecasting method for the averaged flow rates applicable to the Prut River using multiple linear regressions. The dependent variable, the river flow is calculated as a function of monthly average temperatures recorded for 30 years at two weather stations in the river basin. The method could be useful in the current context of climate change.

Key words: monthly flow; forecast; correlations.

1. Introduction

The activity operating any yearly water accumulation requires that the responsible staff take decisions regarding the use of discharge installations. These will allow the evacuation of flow rates based on consumer demands and the monthly inflow debits.

As regards the necessary flow for consumption, it is known with precision.

---

*Corresponding author: *e-mail*: cotiusca@tuiasi.ro

The most delicate problem remains the evaluation of the incoming flow into the reservoir. Underestimating this flow imposes spills water exceeding the transport capacity of the riverbed, and thus the appearance of the flood risk. Overestimating the flow rate, leads to the inability to meet the needs of consumers. Therefore the decision is largely based on the experience of the personnel responsible for this area.

2. Presentation of the Method

Given the above considerations we tried to look if there is a link between average monthly temperatures recorded in the Prut River Basin (from two weather stations) and monthly average flows recorded in a particular section of the river Prut. This link would seem logical though it is impossible to find a mathematical relationship to establish with precision. Thus we chose the solution of a statistical analysis by multiple linear regression of a set of values recorded over a period of 30 years. It is known that generally the means to study the regression of variables (dependent) according to one or more of explanatory variables (independent) in order to estimate (predict) the median or the mean value thereof based on the known values of the explanatory variables (Giurma *et al*., 1990).

Or mathematically:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + ... + \beta_k X_{ki} + e_i, \ (i=1,2,3,....,N), \quad (1)$$

where β_1 is the intersection, $\beta_2 ,...,\beta_k$ – the partial regression coefficients, u – the stochastic distortion term and i – its corresponding observation, N – population size, Y – the dependent variable and X – the independent variable.

$$\left(\mathbf{X}'\mathbf{X}\right)\hat{\beta} = \mathbf{X}'\mathbf{y} , \quad (2)$$

where: $\mathbf{X}$ is the matrix of independent variables, $\mathbf{X}'$ – its transpose, $\hat{\beta}$ – the column vector of the coefficients (k elements), $\mathbf{y}$ – the column matrix of dependent variables (N elements).

The average monthly flows were analyzed as depending on the average monthly temperatures recorded during the previous 12 months at two weather stations in the basin (Avrămeni and Dorohoi). Database of monthly average temperatures for the 30 studied years was build using FoxPro language. To create variables containing temperatures in the 12 preceding current months, the following program was drawn:

```
set talk off
dimension c(50,15)
use temperat
go top
i=1
do while .not. eof()
 j=2
 do while j<14
 f=field(j)
 c(i,j-1)=&f
 j=j+1
 enddo
f=field(1)
c(i,14)=&f
skip
i=i+1
enddo
clear
store space(2) to l
@2,10 say 'ultima luna a
intervalului este (nr.): ' get l
read
nf='l9-'+l
copy stru to &nf
        use &nf
b=val(l)+1
cor=1
if b>12
cor=0
endif
m=1
k=cor+1
clear
do while k<i
append blank
 m=1
 cor=0
 do while m<13
 if m>val(l)
 cor=1
 endif
 f=field(m+1)
 repla &f with c(k-cor,m)
 m=m+1
 enddo
f=field(1)
repla &f with c(k,14)
k=k+1
          enddo
```

3. Case Study: the Forecast of an Average Monthly Flow on Prut River

To determine the regression coefficients we used a specialized program for statistical calculations. The variables obtained with the previous program can be exported in text mode, depending on the compatibility requirements of such software.

The following variables were created: deb – dependent variable containing the mean values of the monthly flows measured on the Prut river, upstream the reservoir for a period of 30 years; dm – independent variable containing the average flow recorded in the same section in the last 12 months (they cover over 30 years); Pm – independent variable containing the average rainfall in the last 12 months; $t1_{i-j}$ – independent variables containing the values of monthly average temperatures recorded in the *j* month at *i* weather station (to ensure some compliance with IMH coding for weather stations, the

Avrameni weather station has no value while Dorohoi weather station has the value of 9).

It should be emphasized that the range of average values includes values registered until the *j* month, previous to analyzed month. Index *j* refers to the month of a calendar year (ex.: $j = 1$ for January, $j = 2$ for February, etc.).

The solution with a larger number of explanatory variables was chosen to improve the coefficient of determination R^2, knowing that:

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2}. \tag{3}$$

So, as seen in the relationship, increasing the number of explanatory variables, R^2 will always increase. Analyzing the results of the calculations presented by computer as a spreadsheet, it is found that R^2 test indicates a significant relationship between the explanatory variables: average flows, average rainfall, average temperature and the dependent ones: the average monthly flow. Regressions could be improved by eliminating insignificant variables at *t* test. Since such tests, performed in some cases, did not yield to significantly better results, and finally the forecasted values are good, the operation was cancel. Besides, the fact that, in the analyzed case, the values of the explanatory variables recorded in some month of a year could have a higher or lower share in the forecasting of monthly average flows was not, for now, the subject of this research.

Since the method was used to make forecasts in the field of variation of explanatory variables (especially), the fact that collinearity exists or not between them has little importance (Hesel & Hirsch, 1992).

For catchment area of the Prut River, the resulting data are relevant and could be taken into account in setting the operating rules of the Stânca-Costeşti reservoir, located downstream the considered section.

An example of running the regressions and charts including the predicted and measured flows in one of the studied years are presented below. (Table 1).

The following notations were used: **d*i*** – measured flow during the analyzed month; **d*i*p** = the predicted flows using the regression (flow rate for the month analyzed was unknown).

The meanings of the main notations of the listing are: ss – the sum of squares; ms – the average of the squares; *R*-squared – the determination coefficient; *t* = t test.

Table 1
Listing of the Regression for June

Source	SS	df	MS	
				Number of obs= 28
Model	46977.4268	4	11744.3567	F(4, 23) = 2.34
Residual	115342.235	23	5014.87977	Prob > F = 0.0851
				R-squared = 0.6894
Total	162319.662	27	6011.83932	Adj R-squared = 0.1658
				Root MSE = 70.816

d6	Coef .	Std. Err.	*t*	P>|*t*|	[95% Conf. Interval]	
dm6	1.290361	.5524576	2.336	0.029	.147515	2.433206
t9m	1.404978	23.10533	0.061	0.952	-46.39204	49.20199
tm	-.9548473	21.27318	-0.045	0.965	-44.96176	43.05207
pm	-3.584868	1.60076	-2.239	0.035	-6.896292	.2734435
_cons	187.9733	208.9695	0.900	0.378	-244.3131	620.2596

A comparative sample chart of measured and predicted average flows, as shown in Fig. 1, confirm the correlation.

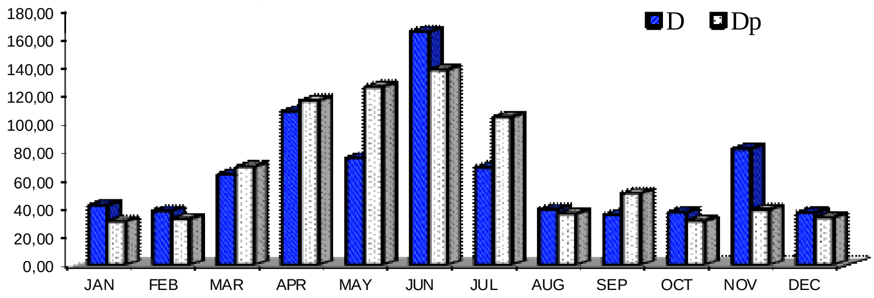

Fig. 1 – The chart of forecasted flows for the year under review.

4. Results and Conclusions

For the analyzed catchment area it was found that there is a good enough correlation between the sequences containing monthly average rainfall, average flows and averages temperature recorded in the 12 months preceding the analyzed month and the values of current month average flow. It appears that flow rate predictions are closest to the recorded value in December to April and August to October.

The values of the forecasted flows using the proposed method could be used in a "dynamic calculation" of operating charts in order to make available additional water volumes for the utilities downstream of the accumulation lake Stânca-Costeşti. Correlations could be improved if using data from more representative areas of the basin, for example from Moldova and Ukraine. For this study, however, these were the only data to which we had access.

REFERENCES

Giurma I., Drobot R., Antohi C., *Hidrologie*. IPI 1987, 1990.
Gujami D., *Basic Econometric*. McGraw-Hill Inc., 1978.
Crețu G., *Economia apelor*. Edit. Did. Pedag., Bucureşti, 1976.
Cotiuşcă Zaucă D., *Contribuții la modernizarea exploatării lacurilor de acumulare*. Ph. D. Diss., "Gh. Asachi" Techn. Univ., Iaşi, 1990.
Hesel D., Hirsch R., *Statistical Methods in Water Resources*, Elsevier Sci. Publ., New York, 1992, 305.

UN CAZ PARTICULAR DE PROGNOZĂ A DEBITELOR LUNARE PE RÂUL PRUT

(Rezumat)

Se prezintă o metodă de prognoză pentru debitele medii aplicabile râulu Prut, folosind mai multe regresii liniare. Variabila dependentă, debitul râului este calculată în funcție de temperaturile medii lunare înregistrate timp de 30 de ani, la două stații meteorologice din bazinul hidrografic. Metoda ar putea fi utilă în contextul actual al schimbărilor climatice.