On the variability of a river water flow under seasonal conditions: Case study

C A Mocanu-Vargancsik\textsuperscript{1,3} and A Barbulescu\textsuperscript{1,2,3}

\textsuperscript{1}Doctoral School, Technical University of Civil Engineering, Bucharest, Romania
\textsuperscript{2}Ovidius University of Constanta, Constanta, Romania

E-mails: cristina_mocanu06@yahoo.com; alinadumitriu@yahoo.com

Abstract. Precipitation regime and anthropogenic activities are the main factors that influence the river flow along the time. To assure the flow regulation and the flood mitigation some structural measures must be implemented. Building dams represents one of these measures. Their impact on the stream flow and climate change must be studied for avoiding negative effects on the environment. Therefore, our study focuses on the variability of the Buzau River (in Romania) water flow and its seasonal variation. The analysis is done for the period before and after the moment when the Siriu Dam (Romania) was built. The stationarity and the trend existence are analysed and the results are compared by statistical methods.

1. Introduction
The consequences and the impact on rivers’ discharges due to the climate change, hydrological constructions, and the anthropogenic activities require a study over long periods of time. If the drawn conclusions lead to environment’s damages, then ways are suggested for reducing them, as rational exploitation, and protection. A method to control and reduce floods is dam construction that also advantages the power generation, irrigation, and control, industrial and domestic supply [1]. In temperate zones, floods happen mostly in spring due snow-melting and rains, summer and autumn due frontal rains or heavy rain showers. In Romania, during summer and spring occur up to 40%-50% of overall annual floods [2]. The high peaks and short duration are two of their seasonal characteristics. Buzau River is one of the most important rivers in Romania, whose monthly mean discharges have been analyzed in papers as [3-7]. Some of authors divided the same time series in two sub-series - before and after Siriu Dam’s start year of operating [6]. This article adds to the image on the river’s variability discharge the annual and seasonal studies, before and after the Siriu Dam construction: 1955-1984 and 1985-2010.

2. Data and methodology

2.1. Data
The used data have been provided by the National Institute of Hydrology and Water Management Bucharest (INMH) and consists in the mean daily discharges collected at two hydrometric stations (hs), namely Nehoiu and Basca. The first one is located on the Buzau River and the latter on Basca, one its tributaries. The data spans on 56 years, from 1\textsuperscript{st} January 1955 to 31\textsuperscript{st} December 2010. The Buzau River’s catchment area is situated in the Carpathians’ Curvature, with a rich drainage area that includes 102 tributaries, lying on $5264 \text{ m}^2$ [8] (figure 1). The Buzau’s multiannual mean discharge
varies from 1 m³s⁻¹ on its first kilometres from the source to 25 m³s⁻¹ towards mid-length until the junction with the Siret River.

![Figure 1. Buzau river catchment (map provided by INMH).](image)

The Basca tributary flows into Buzau at a point upstream Nehoiu station, such that Nehoiu hs records the joined flows from Buzau River downstream the Siriu Dam and Basca River (figure 1). Situated at the latitude of 45°26'32” and the longitude of 26°19'13”, Busca Roziliei hs has the following morphometric and hydrological characteristics: the river length from the source to the hydrometric station is 17.1 km, the average slope of the river from the source to the hydrometric station of 1.66°, the basin area associated to the hydrometric station of 778 km², the average elevation of the basin at the hydrometric station – 1108 m, the average multi-annual fluid flow 12.6 m³s⁻¹. On the other hand, Nehoiu hs is situated at the latitude of 45°25'29” and the longitude of 26°18'27”. Its morphometric and hydrological characteristics are: river length from the source of 71.5 km, the average slope of the river from the source of 0.73°, the areal area of 1572 km², the average elevation of the areal at the station of 1020 m, the multi-annual mean of the flow of 21.7 m³s⁻¹ [9].

2.2. Methodology

This article takes the annual and seasonal series, so the following steps will be performed for both hs. They consist in detection of the change points using the CUSUM method, test of existence of a monotonic trend, using the Mann-Kendall and the seasonal Mann-Kendall trend tests. In case of a monotonic linear trend, the slope computation is performed by the Sen’s method. As mentioned above, the time series has been divided in two subseries (01.01.1955 - 31.12.1984 and 01.01.1985 - 31.12.2010), corresponding to pre- and post-Siriu dam’s construction. An ARIMA model has also been built for the seasonal mean series at Nehoiu.

2.2.1. CUSUM method. CUSUM (CUmulative SUM) is a type of control chart good at detecting shifts away from the target, usually mean or standard deviation [10]. The plotted points are cumulative sum of the deviations. Any fluctuation upwards or downwards proves the process has shifted. In our case the mean is our reference, and the CPA software [11] provides the cumulative sums $S_i$ as:

$$S_0 = 0, S_i = S_{i-1} + (x_i - \bar{x}), \ i = 1,n,$$

where $x_i$, $i = 1,n$ represent the registered data and $\bar{x} = (\sum_{i=1}^{n} x_i)/n$ is the sample average. The cumulative sum ends at zero because the average is subtracted from each value [11].
2.2.2. Mann-Kendall and seasonal Mann-Kendall tests. The non-parametric Mann-Kendall (MK) for trend detection tests the null hypothesis $H_0$ of a monotonic trend against the alternative hypothesis, $H_1$, that assumes non-existence of a such trend. The data entries are $x_i$ ($i = 1, n$), and the test statistic $S$ is computed by [12,13]:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sign}(x_j - x_i)$$

(2)

where

$$\text{sign}(x_j - x_i) = \begin{cases} 
1 & \text{if } x_j - x_i > 0 \\
0 & \text{if } x_j - x_i = 0 \\
-1 & \text{if } x_j - x_i < 0 
\end{cases}$$

(3)

If $n \geq 10$, the statistic $S$ is approximately normally distributed and has the mean zero $E(S) = 0$ and the variance ($\sigma^2$) given by as:

$$\sigma^2 = \frac{1}{18} [n(n-1)(2n+5) - \sum t_i(t_i - 1)(2t_i + 5)]$$

(4)

where $t_i$ denotes the number of ties to extent $i$ and the terms in the sum exist only the data series contains tied values. The standard test statistic $z_S$ is calculated as per relation:

$$z_S = \begin{cases} 
\frac{S - 1}{\sigma} & \text{if } S > 0 \\
0 & \text{if } S = 0 \\
\frac{S - 1}{\sigma} & \text{if } S < 0 
\end{cases}$$

(5)

The conclusion of a trend existence is drawn if $|z_S| > z_{\alpha/2}$, case when the null hypothesis is rejected. With the p-value approach, the null hypothesis can be rejected if $p < 0.05$.

The Seasonal Mann-Kendall test (SMK) analyzes if there is any monotonic trend in seasonal data. SMK works as MK test for each season, looking over the data from the same period [14]. For the corresponding period, the individual mean and variance are added up and then equation (5) is applied to find the test statistic.

Combined with MK trend test, the Sen’slope estimator fits a linear trend of the data [15]. It considers that the estimation of the real slope is the median of all values of the paired data:

$$m_i = \frac{x_j - x_k}{j - k}, i = 1, 2, ..., n, j > k$$

(6)

The method is robust to the outliers existence, more competitive than the linear regression for skewed and heteroskedastic data [16].

2.2.3. ARIMA model. A sequence of uncorrelated random variables with zero mean and constant variance $\sigma^2$ is called a white noise.

Let us define the following operators by:

$$B(X_i) = X_{i-1},$$

(7)

$$\Phi(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p, \varphi_p \neq 0,$$

(8)

$$\Theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q, \theta_q \neq 0,$$

(9)
\[ \Delta^d (X_t) = (1 - B)^d X_t. \] (10)

The process \((X_t)\) is said to be ARIMA(p, d, q) process if \(\Phi(B)\Delta^d X_t = \Theta(B)\varepsilon_t\) where the absolute values of the roots of \(\Phi\) and \(\Theta\) are greater than 1, \((\varepsilon_t)\) is a white noise, \(p\) is the order of the autoregressive terms given by equation (8), \(d\) is the number of non-seasonal differences needed for stationarity given by (10), and \(q\) is the number of lagged forecast errors in the prediction equation given by equation (9) [17].

The Ljung-Box and Box-Pierce tests are used to validate the ARIMA(p, d, q) models for a time series. These tests are used to check the null hypothesis that the residuals are independent and identically distributed. The test statistics is

\[ Q = n(n + 2) \sum_{k=1}^{m} \frac{\hat{r}_k^2}{n-k} \] (11)

where \(n\) is the time series length, \(\hat{r}_k\) is autocorrelation at lag \(k\), and \(m\) is the number of lags being tested. For the significance level \(\alpha\), the null hypothesis is rejected if

\[ Q > \chi^2_{1-\alpha, h} \] (12)

where \(\chi^2_{1-\alpha, h}\) is the \(\alpha\)-quantile of the chi-squared distribution table with \(h\) degrees of freedom, \(h = m - p - q\) (\(p\) and \(q\) are the parameters from ARMA(p, q) or ARIMA(p, d, q) model) [18].

The Box-Pierce test uses the statistics

\[ Q_{BP} = n \sum_{k=1}^{m} \hat{r}_k^2, \] (13)

and the same critical region described in equation (12).

The advantage of the distribution the Ljung–Box statistic is the closeness to a \(\chi^2_h\) than the one for the Box-Pierce statistic, including small samples [18,19].

3. Results and discussion

Figure 2 displays the annual mean discharges at Nehoiu hs and Basca hs as bi-dimensional time series together with a dotted vertical line that divides the timeline into two periods, corresponding to the periods before and after the Siriu Dam’s construction. At first glance, from figure 2 it could be noticed that the values of both series are in accordance as their graphical representations have quite the same shape with a small shift, becoming closer after the 1985. One can also remark the numerous peaks and the apparent lack of a trend – fact that has to be confirmed by statistical methods. By CUSUM method, change points have been detected only for the Nehoiu hs (figures 3(a) and 3(c)). Basca series has no change point (figure 3(b)).

Figure 2. Annual mean discharges’ series at Nehoiu hs (blue) and Basca Roziliei (red)
Figure 3. CUSUM charts for (a) Annual mean discharges’ series at Nehoiu (b) Autumn mean discharges at Nehoiu.

For the annual series, all the breakpoints occurred in four consecutive years in mid 70’s. These years correspond to heavy floods in Romania, floods that affected dramatically the population, the economic life, forcing a lot of people to relocate. The autumn series has only two breakpoints, both occurring before 1985, the year of dam’s construction, proving the dam’s importance (table 1). Apparently, the contribution for the breakpoints comes mostly from the autumn as some years coincide.

Table 1. Change points detected by CUSUM methods for Nehoiu series

| Series  | CUSUM                  |
|---------|------------------------|
| Annual  | 1974, 1975, 1976, 1977 |
| Winter  | no                     |
| Spring  | no                     |
| Summer  | no                     |
| Autumn  | 1973, 1975             |

But knowing the temperate climate behavior, all the seasons bring their contribution to the annual means. In spring snow and ice melts and in summer fall heavy rains, increasing the river’s level and discharge.

Due the outcomes, from now on, the study focuses on Nehoiu hs, being interested in evolution of the mean discharges for subs series divided by the changes points and the dam operation start. The Mann-Kendall test has been used and, in case of an existing trend, the slope has been computed by Sen’s method. The results are presented in table 2, for different significance levels (column 4). Remarkable is the fact that, except for Summer, the seasonal means have increased after 1985, almost with the same slope. Figures 4 and 5 illustrate graphically the outcomes.

Table 2. Mann-Kendall trend test and Sen’s slope for Nehoiu series

| Season | Period | Significance level | Sen’s slope | Comments         |
|--------|--------|--------------------|-------------|------------------|
| Annual | 1955-74| 7%                 | 0.424       | Increasing trend |
| Annual | 1985-2010| 5%              | 0.407       | Increasing trend |
| Autumn | 1985-2010| 5%              | 0.425       | Increasing trend |
| Winter | 1985-2010| 5%              | 0.435       | Increasing trend |
| Spring | 1985-2010| 8%              | 0.541       | Increasing trend |
An ARIMA model for the mean seasonal series completes the analysis of the evolution of the discharges’ series. The model’s validation requires a diagnostic check of the residuals that should form a white noise process [17]. R-software performs accurately this diagnostic and selects the parameters with significant coefficients and a good fit.

The model obtained using auto.arima function is autoregressive of order five – AR (5), with the coefficients and the standard errors (SE) given in table 3.

| Coefficients | AR1  | AR2  | AR3  | AR4  | AR5  | Mean   |
|--------------|------|------|------|------|------|--------|
| SE           | 0.1489 | -0.1682 | 0.0496 | 0.3977 | -0.1057 | 21.9113  |
|              | 0.0675 | 0.0620 | 0.0632 | 0.0623 | 1.2219 |

The p-value corresponding to the Ljung-Box-Pierce statistic test is 0.8127. Therefore the null hypothesis on the residual can’t be rejected. For the proposed model, the mean error is -0.00079, the mean absolute error is 9.669, and the mean squared error 12.431. Based on this model, the prediction for the next 12 seasons has been performed and graphical representation of the data series together with the forecast are presented in figure 6. The blue line represents the forecast and the grey zone surrounding the forecast – the confidence interval.
4. Conclusions

A river flow exhibits a high degree of temporal variability and creates numerous challenges in its study. This difficulty involves complex methods and models. Based on availability and accuracy of local measurements, statistical outcomes - especially in forecasting - should be used just to better understand the possible trend evolution, not a certain future fact.

The above study proves that the Siriu dam’s construction on Buzau River had a good impact on the discharges in smoothing the flow, but not reducing it. Using the study data, a model has been built and a prediction of the future discharge has been done in order to manage the impact on the humans’ life and activity, and also for the environment protection. The last five years brought in Buzau River a real unexpected heavy rain that led to floods with disastrous consequences as lost human lives and animals, relocations, agricultural damages. These serious reasons obliged in 2015 the Romanian water authorities to adopt, implement, and monitor The Flood Management Plan for safety measures in the Buzau-Ialomita hydrographic areal [9].

References
[1] Anders Brant S 2000 Classification of geomorphological effects downstream of dams Catena 40 375-401
[2] Counties of the Socialist Republic of Romania available at: https://www.scribd.com/doc/154175542/Judetele-Romaniei-Socialiste (in Romanian).
[3] Chendes V 2011 Water resources in curvature subcarpathians Geospatial Assessments (Bucureşti: Editura Academiei Romane) (in Romanian)
[4] Minea S I 2011 The rivers from Buzău River Catchment Hydrographical and hydrological Constructions (Buzau: Editura Alpha MDN) (in Romanian)
[5] Minea G and Barbulescu A 2014 Statistical assessing of hydrological alteration of Buzau River Induced by Siriu Dam Forum Geografic XIII 50-8
[6] Barbulescu A 2017 Statistical assessment and model for a river flow under variable conditions available at: https://cest.gnest.org/sites/default/files/presentation_file_list/cest2017_00715_poster_paper.pdf
[7] Mocanu-Vargancsik C A and Barbulescu A 2018 Analysis on Variability of Buzau River Monthly Discharges Ovidius University Annals Series: Civil Engineering 20 51-6
[8] Balteanu D and Stefanescu I 1992 Subcarpathians of Buzau, Geography of Romania, IV, Peripheral Regions (Editura Academiei Romane) pp 292-301 (in Romanian)
[9] Water Basin Administration (ABA) Buzau-Ialomita. Risk Management Plan at Floods 2015 http://www.inhga.ro/documents/10184/121027/8+PMRI+Buzau-Ialomita.pdf/c1f09e05-ab2d-4e6b-ac47-17923a7acea5 (in Romanian)
[10] Page E S 1954 Continuous inspection scheme Biometrika 41 100-15
[11] Taylor W 2002 Change-point analyzer, ASA available at: http://www.variation.com/lib
Mann H B 1945 Nonparametric tests against trend *Econometrica* **13** 245-50

Kendall M G 1975 *Rank Correlation Methods* (New York: Hafner)

Hirsch R M, Slack J R and Smith R A 1982 Techniques of trend analysis for monthly water quality data *Water Resour. Res.* **18** 107-21

Sen P K 1968 Estimates of the regression coefficient based on Kendall’s tau *J. Am. Stat. Assoc.* **63** 1379-89

Wilcox R R 2001 *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy* (Springer, New York, USA)

Barbulescu A 2016 *Studies on Time Series. Applications in Environmental Sciences* (Springer Nature, Switzerland)

Ljung G M and Box G E P 1978 On a measure of a lack of fit in time series models *Biometrika* **65** 297-303

Box G E P and Pierce D A 1970 Distribution of residual autocorrelations in autoregressive-integrated moving average time series models *J. Amer. Statisticl Assoc.* **65** 1509-26