J. Hydrol. Hydromech., 65, 2017, 2, 165–174
DOI: 10.1515/johh-2017-0012

# Annual and seasonal discharge prediction in the middle Danube River basin based on a modified TIPS (Tendency, Intermittency, Periodicity, Stochasticity) methodology

Milan Stojković[1*], Jasna Plavšić[2], Stevan Prohaska[1]

[1] Jaroslav Černi Institute for the Development of Water Resources, Jaroslava Černog 80, 11226 Belgrade, Serbia.
[2] University of Belgrade, Faculty of Civil Engineering, Bulevar Kralja Aleksandra 73, 11000 Belgrade, Serbia.
[*] Corresponding author. E-mail: milan.stojkovic@jcerni.co.rs

**Abstract:** The short-term predictions of annual and seasonal discharge derived by a modified TIPS (Tendency, Intermittency, Periodicity and Stochasticity) methodology are presented in this paper. The TIPS method (Yevjevich, 1984) is modified in such a way that annual time scale is used instead of daily. The reason of extracting a seasonal component from discharge time series represents an attempt to identify the long-term stochastic behaviour. The methodology is applied for modelling annual discharges at six gauging stations in the middle Danube River basin using the observed data in the common period from 1931 to 2012. The model performance measures suggest that the modelled time series are matched reasonably well. The model is then used for the short-time predictions for three annual step ahead (2013–2015). The annual discharge predictions of larger river basins for moderate hydrological conditions show reasonable matching with records expressed as the relative error from –8% to +3%. Irrespective of this, wet and dry periods for the aforementioned river basins show significant departures from annual observations. Also, the smaller river basins display greater deviations up to 26% of the observed annual discharges, whereas the accuracy of annual predictions do not strictly depend on the prevailing hydrological conditions.

**Keywords:** Stochastic modelling; Annual and seasonal hydrological predictions; TIPS method; The middle Danube River basin.

## INTRODUCTION

River discharge rates reflect the inherent characteristics of a river basin and are a result of the interaction between the following factors: geophysical processes over a relatively large area, physiographic characteristics of the basin, and water-related human activities. The relation between river discharge and these factors is highly complex, whereas monitoring of hydro-meteorological processes is somewhat limited in both space and time.

The stochastic autoregressive-moving-average models (ARMA) are based on autocorrelation and random nature of time series. They are extensively used for hydrological predictions conducted in the latter half of the 20th century, both to model and to predict annual and periodic time series. These approaches are useful tools for modelling the time series based on the very feature of the time series whose elements are related in the Markov chain (Aksoy and Bayazit, 2000). The ARMA model is developed over two time periods (Salas et al., 1980). The first began in 1960 and the notable works include those of Thomas and Fiering (1962) and Yevjevich (1963). During that period, the standard procedure for assessing model parameters is founded upon the method of moments. The second period of development, since 1970, is motivated by the landmark book by Box et al. (2008). This period is characterized by improved methods for assessing the autoregressive model parameters. However, a constraint of the ARMA model remains evident in its applications to seasonal time series. To overcome the shortfall, a modification included the first and second order differencing of time series (Salas et al., 1980). Yevjevich (1984) introduced the TIPS method (Tendency, Intermittency, Periodicity, Stochasticity). This method decomposes the time series into the deterministic and the stochastic component based on the spectral analysis and ARMA models. In the same manner, Pekarova and Pekar (2006) developed the stochastic model for long-term prediction which consists of the following components: the harmonic component, the autoregressive component, the regressive component connected with the NAO (North Atlantic Oscillation) phenomenon and a random component. As in the case with previous model, the low and high frequencies components are also included in the methodology for stochastic simulation of annual discharges proposed by Stojković et al. (2015). This methodology aims to reduce uncertainty caused by short observations. For this reason, the random time series is generated by the single bootstrap model to provide sufficient sample needed for water resources planning.

Furthermore, the stochastic models based on the Box-Jenkins linear regression exhibit short-memory and their autocorrelation function decreases rapidly with the time lag (Koutsoyiannis, 2000). These models are unable to adopt for modelling hydrological time series which possess the long-memory. This special behavior of hydrologic time series is examined by Hurst (1951) when it conducts the long-term water storage study of the Nile River. His discovery is known as the Hurst phenomenon, based on the tendency of dry years to be grouped within long dry periods, and wet years to be grouped within long wet periods, respectively. At the same time with the ARMA models, Mandelbrot (1965) developed the other class of models such as the fractional Gaussian noise (FGN). The FGN model exhibits long run statistical dependence and has the form required for self-similar process. Regardless, the short-term properties, as shown by the correlations between nearly successive annual values, are preserved in modelled time series. In the same vein, Koutsoyiannis (2000) proposed a generalized mathematical framework for stochastic simulation and forecast of hydrologic series incorporating the short-memory (ARMA) and long-memory (FGN) model. In the next period, Efstratiadis et al. (2014) presented a robust three-level multivariate scheme for stochastic simulation of the correlated processes.

The goal of this paper is to introduce a methodology for determining seasonal and annual hydrological predictions based on a statistical pattern of long observations. However, there are a number of stochastic models with a finer time discretisation proposed for an hourly, daily or monthly forecast. One of them is the TIPS method (Yevjevich, 1984) aimed at determining a forecast of daily hydrological series under the assumption of time series stationarity, whereby it is not possible to predict hydrological series in the long-run, needed for water resources management strategies. This issue is addressed in the paper proposing a modification of the TIPS method in a way that could be used for assessing discharges several years ahead. Because of this, annual and seasonal time series are used instead of daily discharges. The seasonal cycle is removed to capture the long-term statistic characteristics needed to assess the low-frequency component of discharge series, which defines shifting of multi-annual wet and dry periods. Beside this, the residuals are used to assess the short-term statistic characteristics that constitute the high-frequency component. This component shows the correlations among successive annual values of hydrological series. The rest of time-series modelling is a random part which has to be an independent time series with characteristics such as the Gaussian process (white noise). The proposed method is applied to hydrologic records of six gauging stations in the middle Danube River basin.

## METHOD

### Description of the modified TIPS model

Main components of time series in the TIPS approach introduced by Yevjevich (1984) are the trend, the periodic component, the stochastic component and the random time series. The TIPS approach is generally aimed for daily hydrologic time series for which the seasonal cycle or the intra-annual distribution plays the major role. In addition to the principal frequency of 365 days in daily series, or the frequency of 12 months in monthly series, hydrologic series can also contain periodicities with low frequencies on different multi-annual or multi-decadal scales. When modelling hydrologic time series with seasonal or annual time step, detection of the low-frequency periodicities becomes more important. To make a distinction between the periodicities on the sub-annual scale and the long-term scale, we refer to the latter as the macro-periodicity or the long-term periodicity.

In order to apply the TIPS approach to annual and seasonal hydrologic time series, we have introduced a modification of the TIPS approach described in the sequel. The time series are decomposed by extracting the deterministic component $Q_{DET}$ and the stochastic component $Q_{STOCH}$. The deterministic component is further decomposed into the trend and the periodic components. The decomposition principle of the modified TIPS method is described by the following equation:

$$Q(t) = Q_{DET}(t) + Q_{STOH}(t) + \varepsilon(t) = \left[Q_T(t) + Q_P(t)\right] + Q_{STOH}(t) + \varepsilon(t), \tag{1}$$

where $Q(t)$ is the time series of annual or seasonal river discharges at time step $t$, $Q_{DET}(t)$ is the deterministic part comprised of the long-term trend component $Q_T(t)$ and the macro-periodic component $Q_P(t)$, $Q_{STOCH}(t)$ is the stochastic component, and $\varepsilon(t)$ is the random time series representing the error term. This approach is based upon the TIPS method (Yevjevich, 1984), whereas it is modified in such a way that annual and seasonal discharge time series are used instead of daily time series. The time series are composed of annual and seasonal discharges which do not contain a seasonal component. In the case of seasonal discharges, sub-series comprised of the same seasons are used. The reason for removing the seasonal component is an attempt to identify the long-term stochastic behaviour. The long-term pattern presents the altering wet and dry perennial intervals with the length for European region of approximately 30 years (Labat, 2006; Pekarova and Pekar, 2006; Pekarova et al., 2006; Stojković et al., 2015). Nevertheless, the short-term changes are successive shifting annual discharges in a few years (Fendeková et al., 2014).

The annual discharge time series presents a self-similar process which is correlated at long and short time scale. The goals of the modified TIPS model preserve the (1) long-term and (2) short-term statistic characteristics of annual and seasonal discharges. Also, the randomness of hydrological process is represented in error term which is the independent time series. In this regard, the deterministic component tends to preserve (1) multi-annual statistic characteristics which is closely connected to the Hurst phenomenon. The (2) short-term statistic characteristics consist of the correlations among nearly successive annual values. Furthermore, the detrended time series were checked for long-term periodicity. The LOESS (locally weighted scatterplot smoothing) technique is applied to better assess the macroperiodic component. The long-term macroperiodic is modelled by the spectral analysis. The residuals of the time series represent the stochastic component characterized by high-frequency periodicity. The stochastic component is modelled by applying the ARMA models, founded upon the conventional method of Box et al. (2008). The last segment of the time series is the modelling error, which should have the characteristics of a random time series with zero mathematical expectation and constant variance.

### Modelling the deterministic component

In order to model the deterministic component, the trend is identified first and then removed from the discharge time series to proceed with identification of the macro-periodic component. Presence of a monotonic trend in the annual and seasonal discharges $Q$ in the modified TIPS approach is tested using the nonparametric Mann-Kendall (MK) test at the significance level $\alpha$.

If the monotonic trend is significant at the significance level $\alpha$, the linear trend slope is determined by using the Theil-Sen's slope estimator $\beta$, which is defined as the median of all pairwise slopes in time series (Sen, 1968):

$$\beta = Median\left[\frac{Q_j - Q_i}{j - i}\right], \quad \text{for all} \quad i < j; \tag{2}$$

where $Q_i$ and $Q_j$ are the annual or seasonal discharges at time steps $t_i$ and $t_j$. The linear trend is then formulated as:

$$Q_T(t) = \alpha + \beta t \tag{3}$$

where $\alpha$ is the intercept in the linear equation.

The detrended series $Q'$ is obtained by removing $Q_T$ from the original series $Q$:

$$Q'(t) = Q(t) - Q_T(t). \tag{4}$$

Identification of the periodic component begins with smoothing the detrended series $Q'$ by the LOESS method. The spectral analysis is then used to determine the long-term harmonics of smoothed annual and seasonal discharges. Significant long-term harmonics are identified by the Fisher's statistic at significance level $\alpha = 0.05$. The Fisher's statistic $g_1$ is based on the ratio of a variance explained by the largest harmonic $(c_1^2 / 2)$ and the total observed variance $\sigma^2$ (Yevjevich, 1972):

$$g_1 = \frac{c_1^2}{2 \cdot \sigma^2}. \tag{5}$$

Once the largest harmonic is determined as a statistically significant one, the statistical significance of the following harmonics sorted in descending order is examined. The Fisher's statistic ($g_i$) for the $i^{\text{th}}$ harmonic is defined as follows:

$$g_i = \frac{c_i^2}{2 \cdot \sigma^2 - \sum_{i=1}^{i-1} c_i^2}. \tag{6}$$

The critical value of the test statistic $g_{cr}$ is determined in following equation (Yevjevich, 1972):

$$g_{cr} = 1 - \left( \frac{0.05}{N} \right)^{\frac{1}{N-1}} \tag{7}$$

where $N$ is the sample size of $Q$. Once significant harmonics are determined, the macroperiodic component $Q_P$ is modelled by using the cosine and sine waves for significant frequencies (Stojković et al., 2015).

**Modelling the stochastic component**

For identification of the stochastic component $Q_{STOCH}$, the complete deterministic component $Q_{DET}$ is subtracted from the time series $Q$ to provide the residuals $Q''$:

$$Q''(t) = Q(t) - Q_{DET}(t) = Q(t) - (Q_T(t) + Q_P(t)). \tag{8}$$

Modelling the stochastic components in time series has to be performed on the transformed series in order to arrive at normally distributed model error term/residuals (Hipel and McLeod, 1994; Salas et al., 1980). The input series $Q''$ for modelling the stochastic component have zero mean and standard deviation $\sigma_{Q''}$. In this study, some of the input series $Q''$ exhibited substantial skew and needed some transformation for reducing these skews. The transformation applied is based on the normal scores of the empirical cumulative distribution function (ECDF). The transformed series are obtained as the inverse normal distribution of the ECDF, $x_t = \Phi^{-1}(p_i)$, where $\Phi$ denotes cumulative normal distribution function and $p_i$ is the value of ECDF for the $i^{\text{th}}$ element in the ordered series $Q''$. Once the stochastic component model is identified for $x_t$, it is transformed back into the $Q_{\text{STOCH}} = x_t \cdot \sigma_{Q''}$.

The stochastic component is modelled with the linear stochastic model based on the traditional Box and Jenkins approach. A general ARMA model ($p$, $q$) has the following form (Box et al., 2008):

$$\begin{aligned} x_t &= \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + \varepsilon_t \\ &- \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \end{aligned} \tag{9}$$

where $x_t$ is the normalized time series at time step $t$, $\varepsilon_t$ is the independent random series, $\phi_1, \phi_2, \ldots, \phi_p$ are the parameters of the AR($p$) model, and $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the MA($q$) model.

Development of the stochastic model consists of four steps: (1) identification, (2) estimation, (3) selection, and (4) verification. In the first step, identification of the ARMA model is conducted by analysing the autocorrelation function (ACF) and the partial autocorrelation function (PACF). In the second step, the AR($p$) parameters $\phi_p$ are estimated by solving the Yule-Walker equations, while the MA($q$) parameters $\theta_q$ are determined by using the covariance function of the $x_t$ series (Box et al., 2008). In the third step, the preferred model is selected on the basis of the minimum value of the Akaike Information Criterion (AIC) as the model performance measure (Salas et al., 1980). Finally, the selected model is verified in terms of the model error ($\varepsilon_t$) by testing its serial correlation. In the ideal case, the model error $\varepsilon_t$ represents the white noise if it fulfils the conditions of having zero mean $E(\varepsilon_t) = 0$, constant variance $\sigma^2(\varepsilon_t) = \text{const}$ and covariance function $C(\varepsilon_t) = 0$ for time lag greater than zero. The most important feature of the model error is that it is a time-independent random variable and that it does not contain any autocorrelation including hidden cyclical patterns. Independence of the model error $\varepsilon_t$ is verified by using the portmanteau test and the Box-Ljung test (Salas et al., 1980):

$$W_{PM} = N \sum_{t=1}^{j} r_\varepsilon^2(t), \tag{10}$$

$$W_{BL} = N(N+2) \sum_{t=1}^{j} \frac{r_\varepsilon^2(t)}{N-t}, \tag{11}$$

where $W_{PM}$ and $W_{BL}$ are test statistics of the portmanteau test and the Box-Ljung test respectively, $r_\varepsilon$ is the autocorrelation coefficient of random time series $\varepsilon_t$ at time step $t$, $N$ is the simple size and the sums are evaluated with autocorrelations up to lag $j$, which depends on the series length and the orders $p$ and $q$ of the ARMA model: $j = N/10 + p + q$ (Salas et al., 1980). Both $W_{PM}$ and $W_{BL}$ are $\chi^2$-distributed with degrees of freedom ($\nu$). The null hypothesis that the model error is random is rejected at the significance level $\alpha$ if $W_{PM}$ and $W_{BL} > \chi^2(1-\alpha; \nu)$ while the number of degrees of freedom is equal to $N/10$.

**Model evaluation**

The efficiency of the modified TIPS model is tested by using the Nash-Sutcliffe efficiency (NSE) and the root mean square error (RMSE) to standard deviation ratio (RSR) as the performance measures (Moriasi et al., 2007). The NSE is a well-known measure that determines the relative magnitude of the error variance compared to the observed data variance:

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{N} \left[ Q(t) - Q_m(t) \right]^2}{\sum_{t=1}^{N} \left[ Q(t) - \bar{Q} \right]^2}. \tag{12}$$

where $Q$ is the observed discharge, $Q_m$ is the simulated discharge, and $\bar{Q}$ is the mean discharge for the total number of observations $N$. A value of NSE equal to 1 indicates perfect fit between the model and the observed data.

The RMSE is also one of the commonly used model performance measures. In addition, Moriasi et al. (2007) suggests

RSR as a measure defined as the ratio between the RMSE and the standard deviation $\sigma_Q$ of the observed data:

$$\mathrm{RSR}=\frac{\mathrm{RMSE}}{\sigma_Q}=\frac{\sqrt{\sum_{t=1}^{N}\left[Q(t)-Q_m(t)\right]^2}}{\sqrt{\sum_{t=1}^{N}\left[Q(t)-\bar{Q}\right]^2}}. \quad (13)$$

This ratio provides a scaled, non-dimensional measure of the RMSE, while the RSR with the value of 0 indicates a perfect model fit.

**Short-term predictions**

The short-term predictions of the annual and seasonal river discharges are based on the assumption that the underlying processes consist of oscillations with both low and high frequency. The low frequency component corresponds to multi-decadal variability and the high frequency component represents successive variation of annual discharges in a smaller number of years. The simulations performed to obtain the short-term runoff predictions therefore represent simulation of the small-scale variability in the next $l$ future annual steps needed for planning in water resources.

The short-term discharge predictions are obtained by extrapolating the deterministic model components and by providing forecasts of the stochastic component. The future values of the trend $Q_T$ are extrapolated by using the equation (3). Given that the linear trend determines the tendency of the entire time series, it is justifiable to extrapolate the trend component over a shorter horizon of $l$ time steps. Also, long-term trend extrapolation is not justified since it would imply non-stationarity that would have to be supported by very long observed series or by proving that a reason for non-stationarities exists at the given locations. The macroperiodic component $Q_P$ of the annual or seasonal discharges is also extrapolated according to the observed long-term periodicity model.

The forecast of the stochastic component is obtained by the ARMA model given in the equation (9) and by applying the minimum mean square error method which provides the confidence intervals of the forecast. It assumes that the best forecast of time series $x_t$ for the lead time ($N+l$) is (Box et al., 2008):

$$\tilde{x}_t(N+l)=\psi_{N+l}\,\varepsilon_t+\psi_{N+l+1}\varepsilon_{t-1}+\psi_{N+l+2}\varepsilon_{t-2}+\dots \quad (14)$$

where the forecast weights $\psi_{N+l}$, $\psi_{N+l+1}$, ... are determined by the minimum mean square error method and $\varepsilon_{t-1}$, $\varepsilon_{t-1}$, ... are terms with zero mean $E(\varepsilon_t)=0$ and constant variance $\sigma_\varepsilon^2=\mathrm{const.}$

The mean square error of the unbiased forecast is shown as follows:

$$E\left[x_{N+1}-\tilde{x}_t(N+l)\right]^2=\left(1+\psi_1^2+\dots+\psi_{N+l-1}^2\right)\sigma_\varepsilon^2 \quad (15)$$

where $x_{N+1}$ and $\tilde{x}_t(N+l)$ are the observed value of time series and its forecast for the lead time $N+l$, respectively.

The variance of the forecast error for the lead time $N+l$ is expressed as:

$$\sigma^2(\varepsilon_t)=\left(1+\psi_1^2+\psi_2^2\dots+\psi_{N+l-1}^2\right)\sigma_\varepsilon^2 \quad (16)$$

while the confidence intervals are determined as follows:

$$\pm z_{\alpha/2}\left[1+\sum_{t=N}^{N+l}\psi_t^2\right]^{1/2}\left[\sigma^2(\varepsilon_t)\right],\quad t=N+1,N+2,\dots \quad (17)$$

where $z_{\alpha/2}$ is the standard normal variate for the $\alpha/2$ probability of exceedance, $\sigma^2(\varepsilon_t)$ is the variance of the error term which is used as forecasting error, $\psi_t$ is the forecast weights, and $N$ is a sample size.

## DATA

The study is conducted at six hydrologic stations situated in the middle Danube River basin shown in Figure 1 and listed in Table 1, including two stations on the Danube River and one station on the Sava River, the Tisza River, the Velika Morava River and the Lim River. The discharges over different seasons (winter-$Q_{WIN}$, spring-$Q_{SPR}$, summer-$Q_{SUM}$, autumn-$Q_{AUT}$) and annual discharges are analysed in the study. The annual and seasonal hydrologic time series are obtained from the Hydrometeorological Service of Serbia. The analyses are carried out for a synchronous period at all stations from 1931 to 2012.

## RESULTS

The annual discharge series at six locations in the middle Danube River basin, analysed in this paper, are presented in Figure 2. The linear trend component and the long-term periodic component for each annual series are also plotted in Figure 2.


**Fig. 1.** The Danube River basin with location of hydrologic stations used in the study.

**Table 1.** Hydrologic stations in the middle Danube River basin used in the study with basin areas *F*, mean annual discharges *Q*, mean annual specific yields *q*, standard deviation *S* and skew of annual discharges.

| River | Hydrologic station | $F$(km$^2$) | $Q$ (m$^3$/s) | $q$ (l/s/km$^2$) | $S$ (m$^3$/s) | skew |
|---|---|---|---|---|---|---|
| Danube | Orsova | 576232 | 5478 | 9.5 | 959 | 0.598 |
| Danube | Bogojevo | 251593 | 2879 | 11.4 | 541 | 0.704 |
| Tisza | Senta | 141715 | 799 | 5.6 | 254 | 0.881 |
| Sava | Sremska Mitrovica | 87996 | 1532 | 17.4 | 311 | 0.471 |
| Velika Morava | Lubičevski Most | 37320 | 229 | 6.1 | 73.2 | 0.989 |
| Lim | Prijepolje | 3160 | 76.5 | 24.2 | 18.3 | 0.581 |

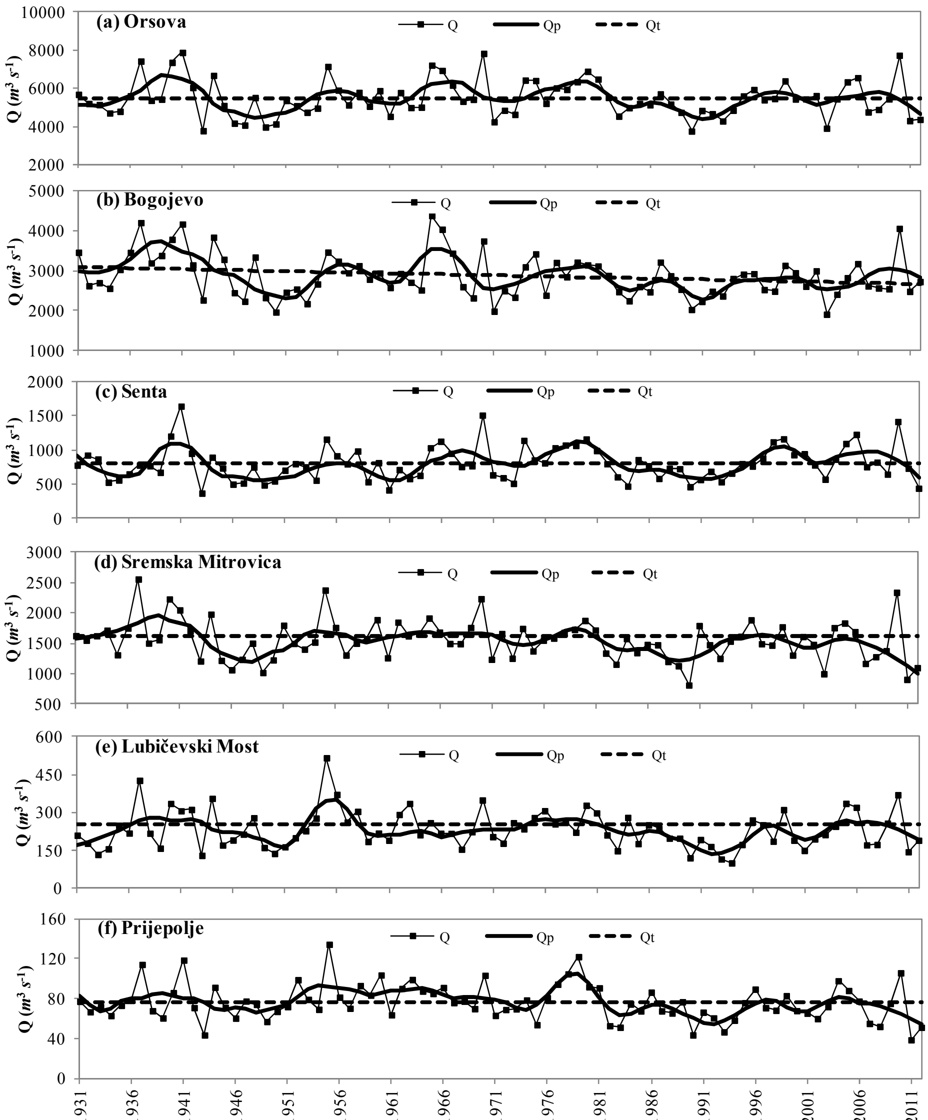


**Fig. 2.** Annual discharges at six stations of the middle Danube basin ($Q$ – observed annual discharges, $Q_T$ – linear trend, $Q_P$ – long-term periodic component).

The trend of annual and seasonal discharges is tested by using the MK test at significant level of $\alpha$ = 0.10. The Kendall's statistic is used to obtain the test statistic $z_s$ which follows the standard normal distribution. The results for the test statistic $z_s$ of the considered annual and seasonal time series are shown in Table 2.

The results shown in Table 2 indicate that the annual time series do not show a significant trend at the 10% significance level ($|z_s| < 1.645$), except for the decreasing trend at the Bogojevo station located at the Danube River. It can be seen from Table 2 that the summer discharges at Bogojevo (the Danube River) and Prijepolje (the Lim River) exhibit a significant decreasing trend, while the winter discharges at Senta (the Tisza River) have a significant increasing trend. The results do not show any pattern in trends at different stations, but this is expected having in mind that the stations analysed are not geographically close and have different hydrologic regimes.

The significant linear trend $Q_T$ is removed from the time series $Q$ and it produces detrended time series $Q'$ (equation 4). For the time series which do not show a significant trend, the mean discharge for the observed period is removed from $Q$. Furthermore, the time series $Q'$ is smoothed to extract the long-term periodic component. Smoothing of the time series is performed by applying the local-regression LOESS technique.

**Table 2.** The values of the test statistic $z_s$ (MK test) for the annual (ANN) and seasonal (WIN, SPR, SUM, AUT) discharge series at six stations in the middle Danube basin.

| Station | Mann-Kendall test statistic $z_S$ | | | | |
|---|---|---|---|---|---|
| | ANN | WIN | SPR | SUM | AUT |
| Orsova | 0.264 | –0.369 | 0.712 | 0.403 | –0.417 |
| Bogojevo | 1.870 | –0.162 | 1.367 | 2.184 | 1.239 |
| Senta | –1.418 | –2.078 | –0.272 | –0.599 | –1.063 |
| S. Mitrovica | 1.458 | 1.278 | 1.518 | 0.379 | 1.357 |
| Ljubičevski Most | 0.457 | 0.762 | 0.740 | –0.628 | –0.486 |
| Prijepolje | 1.550 | 1.012 | 0.749 | 1.936 | 0.689 |

**Table 3.** The results of modelling the stochastic component in the annual discharge series and testing the model error term for independence: the AR model order $p$, Akaike Information Criterion AIC, Box-Ljung statistic $W_{BL}$ and portmanteau statistic $W_{PM}$ .

| Station | $p$ | AIC | $W_{PM}$ | $W_{BL}$ |
|---|---|---|---|---|
| Orsova | 6 | –113.0 | 13.80 | 13.06 |
| Bogojevo | 6 | –80.5 | 11.87 | 14.56 |
| S. Mitrovica | 6 | –74.5 | 13.40 | 14.57 |
| Senta | 7 | –83.0 | 10.62 | 11.51 |
| Ljubičevski Most | 7 | –116.9 | 14.67 | 15.26 |
| Prijepolje | 7 | –97.1 | 12.43 | 13.55 |

The length of the smoothing window is from 7 to 13 members, depending on the stochastic characteristics of the time series.

Once detrending and smoothing of time series $Q'$ are conducted, the periodograms are estimated by means of the Fourier transformation. The annual periodograms of the observed and the smoothed detrended annual time series $Q'$ for stations considered are depicted in Figure 3. The smoothed detrended series allow identifying low frequency harmonics, since the high frequency ones are filtered out from the original detrended series. The significant harmonics are identified by testing the Fisher's statistic at significance level $\alpha = 0.05$. The long-term harmonics are used to formulate the macroperiodic component as a sum of statistically significant waves. By subtracting the total deterministic component (the linear trend and the long-term periodicity) from the original annual and seasonal time series $Q$ (equation 8), the residuals $Q''$ are determined and normalized to obtain time series $x_t$ that are used for identification of the stochastic component. In order to choose the appropriate model of the stochastic component, ACF and PACF of the normalized series $x_t$ for the stations considered are determined (Figure 4).

The decision whether to use the AR($p$) or ARMA($p$, $q$) model is made on the basis of the shape of the autocorrelation and the corresponded partial autocorrelation functions (Figure 4). Because of the wavy shape of the ACF, the order $p$ of the autoregressive part should be higher and the order $q$ of the moving average part should be lower. The ARMA model parameters are estimated for different model orders and the best model is selected depending on the values of the AIC. The AIC values suggest that the best fit is achieved for the AR(6) and AR(7) models, as it is shown in Table 3. Table 3 also shows the test results for randomness of the error term by applying the portmanteau test and the Box-Ljung test.

The critical value according to the chi-squared distribution with the sample size $N$ = 82 at significant level $\alpha$ = 0.05 is $\chi^2_{1-\alpha}$ = 15.79. In the case of annual discharges at the stations analysed, the test statistics $W_{BL}$ and $W_{PM}$ indicate lower-than-critical value $\chi^2_{1-\alpha}$ for the significant level $\alpha$ = 0.05. Thus, the modelling errors $\varepsilon_t$ of the annual discharges are the random series. In the case of seasonal discharges, the modelling error $\varepsilon_t$ is also a random time series without a significant serial correlation at significant level $\alpha$ = 0.05.

Having demonstrated the randomness of annual and seasonal error terms, their distribution functions are considered. Actually, it is assumed that the annual and seasonal modelling error is normally distributed examined by the Jarque-Bera test. Results suggest that the error terms of annual discharges belong to normal distribution at significant level $\alpha$ = 0.05. In the case of seasonal discharges, the error terms do not strictly follow a normal distribution since 6 out of 24 seasonal samples are heavily tailed at the same significant level.

Once the deterministic and stochastic components of discharge time series are determined, they are aggregated at the same time step to obtain the modelled annual and seasonal

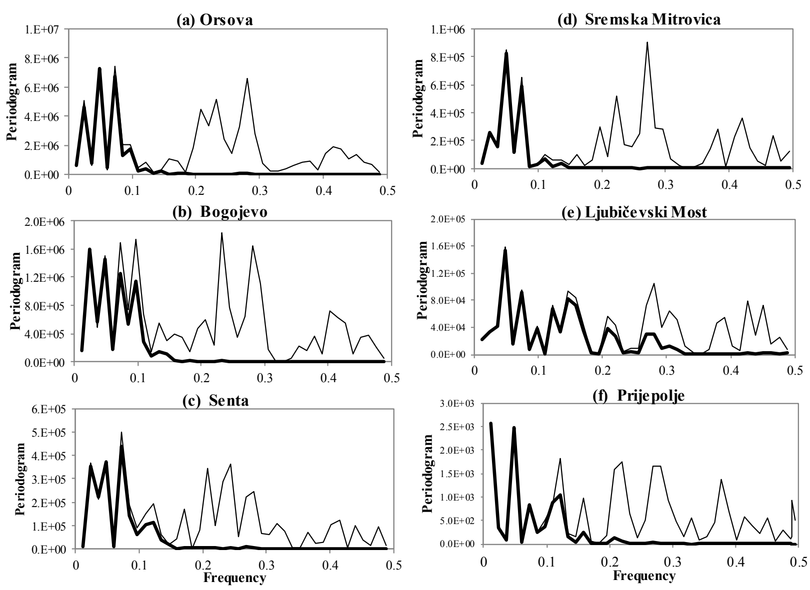


**Fig. 3.** Periodograms of the observed (thin lines) and LOESS-smoothed (thick lines) detrended annual discharges $Q'$ at six stations in the middle Danube basin.

**Table 4.** Short-term predictions of annual and seasonal discharges (m$^3$/s) with confidence intervals of 95% at the stations in the middle Danube River basin for the lead time (2013–2015).

| Year | Station | $Q_{ANN}$ | $Q_{WIN}$ | $Q_{SPR}$ | $Q_{SUM}$ | $Q_{AUT}$ |
|---|---|---|---|---|---|---|
| 2013 | Orsova | 6116±743 | 7279±1569 | 8500±1439 | 4852±1188 | 3834±1520 |
| | Bogojevo | 3496±136 | 2815±329 | 4333±303 | 3646±416 | 3188±710 |
| | Senta | 737±37 | 664±49 | 500±55 | 943±43 | 844±52 |
| | S. Mitrovica | 1844±132 | 2059±62 | 1706±76 | 1919±48 | 1680±84 |
| | Ljubičevski M. | 202±25 | 307±35 | 278±36 | 69±18 | 152±31 |
| | Prijepolje | 74.3±11 | 102±14 | 96±18 | 49±10 | 50±22 |
| 2014 | Orsova | 5173±793 | 4888±2803 | 6764±3372 | 4695±2173 | 4349±3562 |
| | Bogojevo | 2826±214 | 2769±520 | 3277±804 | 3259±1102 | 1999±2206 |
| | Senta | 778±68 | 826±105 | 789±115 | 801±183 | 697±122 |
| | S. Mitrovica | 1441±138 | 2086±141 | 1643±207 | 687±97 | 1348±243 |
| | Ljubičevski M. | 235±32 | 255±85 | 472±104 | 94±27 | 118±91 |
| | Prijepolje | 63.3±15 | 62±28 | 78±43 | 16±21 | 97±66 |
| 2015 | Orsova | 5532±1068 | 6681±5548 | 7582±7303 | 4810±3888 | 3054±2657 |
| | Bogojevo | 2712±385 | 2098±882 | 2828±2204 | 3141±2250 | 2780±2294 |
| | Senta | 501±150 | 444±242 | 621±261 | 495±208 | 443±313 |
| | S. Mitrovica | 1337±142 | 1234±344 | 1872±630 | 729±265 | 1513±608 |
| | Ljubičevski M. | 242±58 | 349±228 | 396±311 | 131±47 | 94±76 |
| | Prijepolje | 73.8±16 | 65±60 | 169±114 | 25±19 | 35±26 |

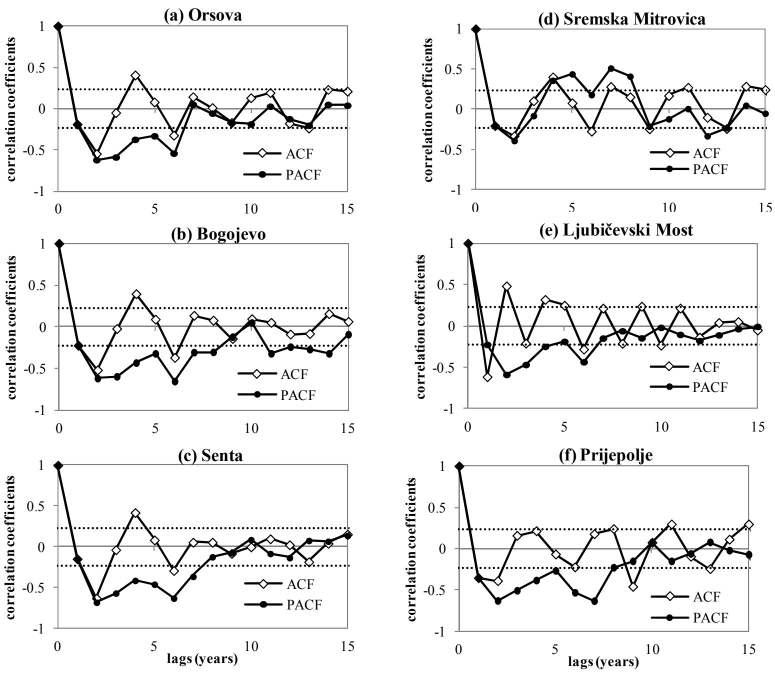


**Fig. 4.** Autocorrelation (ACF) and partial autocorrelation (PACF) functions of the annual normalized residuals $x_t$ with the 95% confidence intervals (dashed lines) at six stations in the middle Danube basin.

discharges at the stations considered. The aggregation of the modelled components is conducted in accordance with the equation (1). The parameters of the modelled component are estimated in the observed period (1931–2012), while the annual and seasonal predictions are determined for three annual step ahead (2013–2015). The short-term predictions of the annual and seasonal discharges are constituted by summarizing all predicted components from the equation (1). At the same way as for the observed period, the short-term annual predictions are aggregated from the seasonal discharge predictions. The obtained annual and seasonal discharge predictions at six stations in the middle Danube basin with their confidence intervals of 95% are shown in Table 4.

As could be seen from Table 4, the short-term annual prediction of the analysed stations suggests an increase of discharges in 2013 compared to the observed period (Table 1). The great river basins such as the Danube River (Orsova, Bogojevo) and the Sava River (Sremska Mitrovica) show a significant increment in the range of 12–22%, while the smallest river stream (the Lim River at Prijepolje) indicates a slight increase of 3%. The Tisza River and the Great Morava River suggest an increase of annual discharges equal to 8% and 11%, respectively. Regarding the annual discharge predictions in 2014 and 2015, the results from Table 4 indicate that the predicted discharges are approximately equal to the mean annual values varying in the range from –5% to +6%. Irrespective of this, the short-term annual predictions in 2015 for the Tisza River (Senta) and the Sava River (Sremska Mitrovica) imply an overall reduction of annual values by 34% and 13%, respectively. A decrement of annual discharges by 17% is also expected for the Lim River (Prijepolje) in 2014 compared to the observed values (1931–2012).

**DISCUSSION**

Stochastic simulation of annual discharges is founded upon assumption that time series can be divided into following components: deterministic, stochastic and random (Yevjevich, 1984). This assumption is used to extract the long-term oscillations of annual discharges for the main course of the Danube River implying that these fluctuations have occurred synchronously for the sites considered (Stojković et al., 2012). Moreover, the stochastic characteristics of the Danube River, the Sava River, the Tisza River and the Great Morava River are examined by using the Hurst exponent as a measure of the process with a long-term memory (Stojković et al., 2014). It is shown that these hydrological series have a Hurst exponent greater than 0.5, indicating that the autocorrelation has to exist at the long lag, i.e. that from a statistical point of view, these time series must be those with a long-term memory. Such behaviour is modelled on a multi-annual time scale by two low-frequency components: multi-annual trend of different sub-series and long-term periodicity (Stojković et al., 2014). These components represent irregular perennial changes of the hydrological process, which are equivalent to the simple scaling behaviour of variability over time scale (Koutsoyiannis, 2003).

The modified TIPS methodology is applied to the selected stations (Table 1) in the middle Danube River basin, using the aforementioned assumption of time series decomposition. The stations are chosen according to long and reliable records needed for capturing their stochastic behaviour. Also, the temporal changes of the considered time series could be predicted in a more precise manner, despite small river basins with "uncontrolled" behaviour. In this research, a long-memory characteristic of annual and seasonal discharge series, closely related to the Hurst phenomenon, is considered by the low-frequency (macro-periodic) component which defines altering of perennial wet and dry periods. Apart from this, the high-frequency (stochastic) component describes a short-term memory of hydrological series expressed as autocorrelation among several time series members. The rest of time series modelling is the random term which displays the randomness of the hydrological process.

Model performance is evaluated in terms of NSE and RSR for the period from 1931 to 2012. The results suggest that the agreement between the observed and modelled annual and seasonal time series is reasonably good. It should be noted that the seasonal models demonstrate better performance than annual ones for most time series. Moreover, the annual discharges averaged from the modelled seasonal discharges exhibit a better agreement with records than in the case of the modelled annual time series. It suggests that the reduction of time discretisation provides better model performance at annual level. Such a result implies that finer resolution is capable for capturing the statistical properties of time series better than a coarser one.

However, the values of NSE and RSR indicate that the fit between corresponding discharges is good and very good according to the guidance proposed by Moriasi et al. (2007). The best fit of annual discharges is achieved for the Velika Morava River at Ljubičevski Most and for the Danube at Bogojevo with NSE = 0.943, RSR = 0.096 and NSE = 0.924, RSR = 0.077, respectively. Nevertheless, the lowest matching between the modelled and the observed annual discharges is in the case of the Sava River at Sremska Mitrovica (NSE = 0.919, RSR = 0.081). The modified TIPS model for the Danube River at Orsova, the Tisza River at Senta and the Lim River at Prijepolje shows satisfactory model performance at annual time scale (NSE = 0.887–0.905, RSR = 0.070–0.112). Also, the modelled seasonal discharges fairly mimic the observed ones since the model efficiency parameters expressed as NSE and RSR are in the range of 0.771–0.953 and 0.062–0.154 for the considered stations, respectively.

Considering that the modified TIPS model is constituted by the separately modelled components, it is necessary to determine the contribution of each component in the observed annual and seasonal discharges. As a measure of their contribution a ratio of the componential variance and a total observed variance is used. Therefore, the annual and seasonal linear trend explains lower part of annual and seasonal variance (0–6.3%). Despite the trend component, the macroperiodical and stochastic component have a substantial share in a total variance in the range of 27.9–51.3% and 29.6–39.6%, respectively. A rest of time series modelling (error term) are represented by the share from 11.6% to 36.2%.

However, a reasonably good agreement of modelled and observed time series suggests that the proposed methodology can be used for short-term predictions. The predictions of annual and seasonal discharges for the analysed stations are given in Table 4. A comparison between the predicted annual discharges (aggregated from the seasonal values) and the observed ones at six stations in the middle Danube basin from 2013 to 2015 is shown in Figure 5.

The annual predictions for the lead time ($N$+3) are compared to the observations in the period 2013–2015 (Figure 5). The relative annual errors for both moderate hydrological conditions and the larger river basins indicate reasonable departures from records. Therefore, the relative error for the Danube River at Orsova and Bogojevo (Figures 5a, 5b), the Sava River at Sremska Mitrovica (Figure 5d) and the Tisza River (Figure 5c) are in the range from –3% to –2%, from –6% to –3%, from –8% to +3% and from –5% to –0.1%, respectively. Despite of this, the substantial departure of the annual forecast for larger rivers could be seen for wet and dry periods. For instance, the observed annual discharges in 2014 exceed mean annual values of the Danube River (Orsova), the Sava River (Sremska Mitrovica) and the Tisza River (Senta) by +16%, +56% and –36%, respectively. Such hydrological conditions lead to underestimation or overestimation of the short-term discharge predictions. The relative annual errors in 2014 for the Danube River (Orsova), the Sava River (Sremska Mitrovica) and the Tisza River (Senta) are –24%, –35% and +57%, respectively. The exception to this behaviour is the Danube River at Bogojevo, since the annual prediction in 2014 fairly fits to the observed values (Figure 5b). This could be attributed to the fact that the observed annual discharge in 2014 exceeds mean annual value for only 5%.

The departures from the observed annual values from 2013 to 2015 for the Great Morava River at Ljubičevski Most and the Lim River at Prijepolje are in the range from –26% to +23% and from –17% to +16%, respectively (Figures 5e, 5f). Such results indicate that annual forecast for moderate hydrological conditions is more accurate in the case of larger rivers (the Danube River, the Sava River, the Tisza River) than it is for the Great Morava River and the Lim River. As opposed to the large river basins, the relative errors for smaller ones do not indicate any pattern between predicted and observed discharges for different hydrological conditions (wet, moderate or dry).

In the contrast with the annual time scale, the short-term seasonal predictions show greater deviation compared to the records in the period 2013–2015 (Table 4). The predictions over the winter and spring seasons have a significantly better agreement with the observed seasonal discharges than it is the case

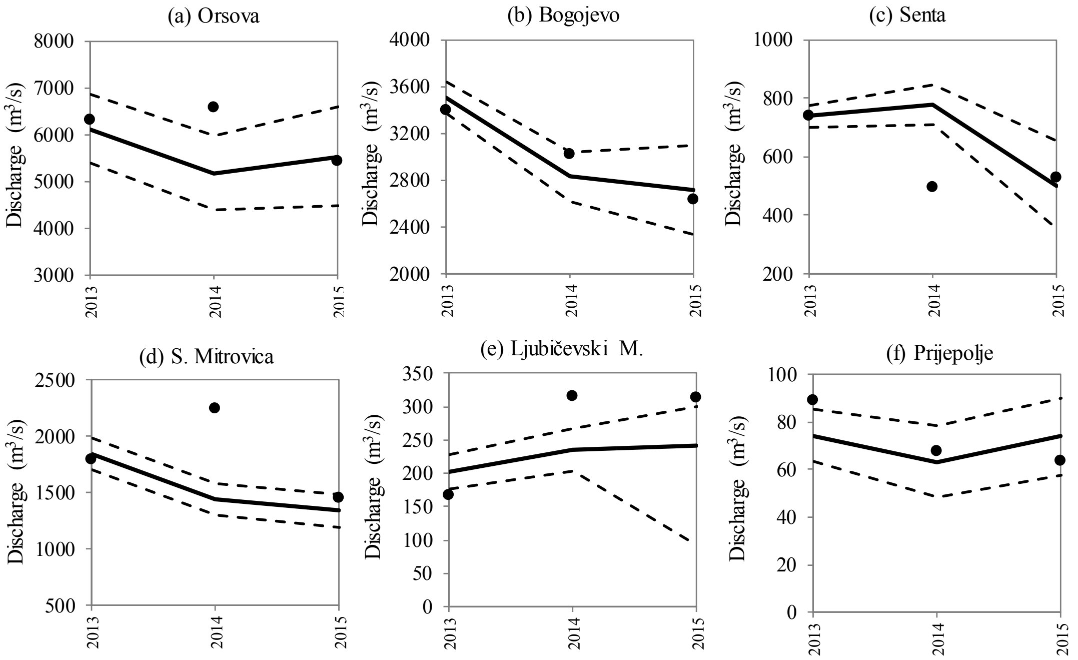


**Fig. 5.** Predicted annual discharges (bold line) with confidence intervals of 95% (dashed line) and annual observations (bold marker) at six stations in the middle Danube basin from 2013 to 2015.

with the other ones. Actually, this behaviour can be attributed to a greater share of stochastic component in discharges over the summer and autumn season. Once the forecast errors of seasonal predictions are summed over the year, the deviation from the observed discharges at annual level becomes fairly reduced. It should note the same behaviour of seasonal time series is shown during the observed period (1931–2012), since the modelled annual discharges constituted by the seasonal times series have a better agreement than the modelled discharge at annual scale.

The aforementioned results imply that the forecast error of great river basins are lower than it is the case with smaller domestic rivers that forecast has a substantial departure from records. This results from the fact that the smaller basins have higher runoff variability than the larger ones. It has been suggested that the larger basins possess a more pronounced multiannual variability due to greater underground retention (Wanga et al., 2014), and their prediction can, therefore, be predicted in a more precise manner. One should note that the proposed methodology for short-term predictions fails to mimic annual discharges for wet or dry hydrological conditions duo to the fact that the stochastic component is capable for modelling merely linear dependence among time series members. Hence, the unexplained variance of annual discharges of high and low annual values could be diminished by utilising the techniques capable for reproducing non-linear structure of hydrological series such as artificial neural networks (Kostić et al., 2016).

It should be noted that the proposed methodology is suitable for the annual time scale, to support information on discharge for the next several years. The use of a more precise time scale requires major modification of the proposed methodology in order to deal with increasing skewness of discharge time series at a finer time step. Also, a modified TIPS methodology cannot be implemented for flood prediction, but it may be upgraded for monthly discharge forecasts. In this manner, a seasonal cyclical component will be introduced and modelled by the wavelet analysis (Cengiz, 2011), able to deal with seasonal nonstationarity.

## CONCLUSION

In order to predict hydrologic processes, annual and seasonal discharges need to be represented by a stochastic model. Therefore, a modified TIPS methodology is proposed for deriving short-term hydrological prediction of hydrological time series with a long memory. The modification involves the use of annual or seasonal discharges, excluding the seasonal periodicity. The aim of the modification is reflected in the determination of the multiple-year hydrologic pattern. It presents an essential departure from the original TIPS method on two grounds: (1) the research methodology is improved in several ways to preserve the characteristics of hydrological series in the long-run, (2) the application of the proposed methodology is an original attempt to assess annual and seasonal discharge predictions for gauging stations in Serbia as a part of the middle Danube River basin. The research methodology proposes forecasting of the low and high frequency components in such a way that the observed discharge pattern is preserved in the predictions. Therefore, extrapolation for predicting the deterministic component for three years ahead is proposed, followed by forecasting of the stochastic component founded upon the Box and Jenkins methodology. The rest of time series modelling is then used to provide the confidence interval of the discharge predictions.

The study is concentrated on the middle Danube River basin for the time interval from 1931 to 2012. The statistically significant linear trend is extracted from the discharge time series. The residuals of annual or seasonal discharges include a long-term periodicity; large periods are extracted by the LOESS method, and the Fourier transform is applied to model the macroperiodic component. The stochastic component is modelled by the AR($p$) models. The model efficiency is demonstrated by using the performance estimators NSE and RSR estimated in the range 0.888–0.943 and 0.070–0.112, respectively. These estimators suggest that modelled annual and seasonal discharges have achieved good matching. Also, it is shown that modelled annual discharges derived by seasonal discharges have better agreement with the records than that at annual time scale. This fact suggests that a reduced time discretisation leads to a better model performance. Moreover, the modelling errors in the case of annual discharges belong to normal distribution, whereas few seasonal samples are heavily tailed.

The purpose of the proposed methodology is to improve management efficiency in the middle Danube River basin and to provide the short-term water resources management strategies. Therefore, the modified TIPS methodology is constituted to be used for prediction of annual and seasonal discharges for next three years (2013–2015). A comparison between the observed and predicted time series implies that large basins exhibit lower relative errors, whereas smaller ones have shown a greater departure for observations. Also, it is found that, in the case of the larger river basins, the prevailing hydrological conditions have a significant influence on the prediction accuracy. It seems that the modified TIPS methodology fails to predict annual discharges of larger rivers in wet and dry periods, since the stochastic component is able to model only linear dependence of time series. Furthermore, incorporating the non-stationary seasonal component into the proposed methodology can lead to a reliable monthly discharge forecast for the considered hydrological series.

*Acknowledgement.* The research presented in this paper is funded by Serbia's Ministry of Education and Science as part of an ongoing scientific project titled "Assessment of Climate Change Impact on Water Resources in Serbia" (TR-37005, 2010-2016).

**REFERENCES**

Aksoy, H., Bayazit, M., 2000. A model for daily flows of intermittent streams. Hydrological Processes, 14, 1725–1744.

Box., G.E.P., Jenkins, G.M., Reinsel, G.C., 2008. Time Series Analysis, Forecasting and Control. Fourth Edition. John Wiley & Sons, INC., Publication. USA.

Efstratiadis, A., Dialynas, Y.G., Kozanis, S., Koutsoyiannis, D., 2014. A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence. Environmental Modeling and Software, 62, 139–152.

Cengiz, T.M., 2011. Periodic structures of great lakes levels using wavelet analysis. J. Hydrol. Hydromech., 59, 24–35.

Fendeková, M., Pekárová, P., Fendek, M., Pekár, J., Škoda, P., 2014. Global drivers effect in multi-annual variability of runoff. J. Hydrol. Hydromech., 62, 169–176.

Hipel, K.W., McLeod, A.I., 1994. Time Series Modelling of Water Resources and Environmental Systems. Elsevier, Amsterdam, The Netherlands.

Hurst, H., 1951. Long term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers, 6, 770–799.

Kostić, S., Stojković, M., Prohaska, S., 2016. Hydrological flow rate estimation using artificial neural networks: model development and potential applications. Applied Mathematics and Computation. DOI: 10.1016/j.amc.2016.07.014.

Koutsoyiannis, D., 2000. A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. Water Resources Research, 36, 1519–1533. DOI: 10.1029/2000WR900044.

Koutsoyiannis, D., 2003. Climate change, the Hurst phenomenon and hydrological statistics. Hydrol. Sci. J., 48, 3–24.

Labat, D., 2006. Oscillations in land surface hydrological cycle. Earth and Planetary Science Letters, 242, 143–154.

Mandelbrot, B.B., 1965. Une class de processus stochastiques nomothetiquesa mothetiques a soi: Application a la loi climatologiqu de H. E. Hurst. [A class of stochastic homogeneous processes: Application to the climatological law of H. E. Hurst]. C. R. Hebd. Seances Acad. Sci., 260, 3284–3277. (In French.)

Moriasi, D.N., Arnold, J.G., van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50, 885–900.

Pekarova, P., Pekar, J., 2006. Long-term discharge prediction for the Turnu Severin station (the Danube) using a linear autoregressive model. Hydrological Processes, 20, 1217–1228. DOI: 10.1002/hyp.5939.

Pekarova, P., Miklanek, P., Pekar, J., 2006. Long-term trends and runoff fluctuations of European rivers. IAHS Publ. 308. IAHS Press, Wallingford, pp. 520–525.

Salas, J.D., Delleur, J.W., Yevjevich, V., Lane, W.L., 1980. Applied Modeling of Hydrologic Time Series, Water Resources Publications. Littleton. Colorado. USA. 484.

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association, 63, 1379–1389.

Stojković., M., Plavšić, J., Prohaska, S., 2012. Stohastička analiza serija srednje godišnjih proticaja na stanicama na Dunavu. [Stochastic analysis of mean annual flow time series for the sites of the Danube River]. 16. Savetovanje SDHI i SDH. Donji Milanovac. Serbia. (In Serbian.)

Stojković, M., Prohaska, S., Plavšić, J., 2014. Internal stochastic structure of annual discharge time series of Serbia´s large Rivers. Journal of Serbian Water Pollution Control Society "Water Research and Management", 4, 3–13.

Stojković., M., Prohaska, S., Plavšić, J., 2015. Stochastic structure of annual discharges of large European rivers. J. Hydrol. Hydromech., 63, 63–70.

Thomas., H.A., Fiering, M.B., 1962. Mathematical Synthesis of Streamflow Sequences for the Snalysis of River Basin by Simulation. In Design of Water Resources Systems. Harvard University Press, Cambridge, Massachusetts.

Yevjevich, V., 1963. Fluctuation of wet and dry years - Part 1. Research data assembly and mathematical models. Hydrology paper 1. Colorado State University, Fort Collins, Colorado, USA.

Yevjevich, V., 1972. Stochastic Processes in Hydrology. Water Resources Publications, Fort Collins, Colorado, USA.

Yevjevich, V., 1984. Structure of Daily Hydrologic Series. Water Resources Publications. Water Resources Publications, Fort Collins, Colorado, USA.

Wanga, H., Sankarasubramanianb, A., Ranjithanb, R.S., 2014. Understanding the low-frequency variability in hydroclimatic attributes over the southeastern US. Journal of Hydrology, 521, 170–181.

Received 20 January 2016
Accepted 22 August 2016