Seasonality and multiannual variability of floods: case study from Eastern Poland

Grzegorz Siwek

Abstract. Nowadays, under increasing climate change effects on the environment, we can observe increasing number of extreme phenomena, including meteorological and hydrological ones. One of such phenomena are floods. The objective of this article is the assessment of basic flood characteristics seasonality in the annual distribution. Analysis were performed based on time series of daily flow values recorded in the years 1951–2014 in three gauging stations located on rivers in Easter Poland, in upper Wieprz catchment. Floods were defined according to TLM algorithm and were assumed to be all cases of flow occurrence exceeding 10% read from FDC (flow duration curve) ($Q_{10}$). Seasonality was analysed using Markham’s Seasonality Index and Period of Seasonal Concentration, analysis of autocorrelation function (ACF) as well as proposed by the author Seasonal Winter Floods Index. The distribution of floods during year indicates one flood season in year which occurs in the spring.

1 Introduction

Most of hydrological, and other natural, phenomena demonstrate seasonal and multiannual variability. It is also observed in series of characteristics connected to water surplus. Floods are usually caused by intensified water supply to rivers inducted by heavy or long-lasting rains, spring snowmelt or sea storms [1-3]. Therefore seasonal and multiannual analysis of extreme hydrological events is relevant and useful from practical and scientific point of view.

1.1 Study Area

Study area was located in Eastern Poland in the south part of Lubelskie Voivodship (Fig. 1). Author focused on upper Wieprz river basin, closed by gauging station in Krasnystaw. There were also analysed 2 subcatchments – Wieprz river up to Zwierzyniec and Łabuńka river up to Krzak gauging station. Basic characteristic of gauged catchments were presented in table 1.

As presented in table 1, analysed catchments has different environmental conditions what affects the process of formation and course of floods on local rivers.

1.2 Data

All the calculations presented in this paper were made based on daily discharge series in 3 gauging profiles in hydrological years 1951-2014 (from 1.11.1951 to 31.10.2014). Data has been obtained from Institute of Meteorology and Water Management (IMGW).

Table 1. Chosen characteristics of gauged catchments.

<table>
<thead>
<tr>
<th>River</th>
<th>Wieprz</th>
<th>Łabuńka</th>
<th>Wieprz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gauging station</td>
<td>Krasnystaw</td>
<td>Krzak</td>
<td>Zwierzyniec</td>
</tr>
<tr>
<td>Area [km²]</td>
<td>3010</td>
<td>427</td>
<td>394</td>
</tr>
<tr>
<td>Mean height [m a.s.l.]</td>
<td>250</td>
<td>233</td>
<td>291</td>
</tr>
<tr>
<td>Mean slope [°]</td>
<td>4.7</td>
<td>2.8</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Fig. 1. Localisation of study area and gauging stations.
2 Methodology

2.1 Flood definition

The issue of water surplus was investigated and analysed in hydrological literature for ages. Floods were defined in different ways by various methods and authors. Due to problem in delimitation genetic beginning of flood, there was established threshold level method (TLM) that define flood as a period of flows that is over established level [2,4-9]. Such a definition allows computation a comparable time series analysis but it prevents recognition of flood formation conditions and it’s genetic characteristics [5].

The determination of threshold level may be based on hydrological, statistical or economic criteria [2,10]. There are many different values of threshold level used. Some hydrologists use characteristic flows [2,5,11-13] while others establish threshold level values as flows at certain probability of occurrence [9,14-17]. According to recent research conducted in upper Wieprz river basin, the most suitable criterion, for environmental analysis, is value of flow occurrence exceeding 10% read from flow duration curve (FDC) \((Q_{10})\) [9,18]. To estimate flood volume there was applied threshold level method (TLM). The surface area limited by the hydrograph and threshold flow level was assumed to be the flood volume. For each gauging station time series there were calculated two flood parameters: number of days with flood in each month and monthly sum of flood volume.

2.2 Seasonal Winter Flood Index (SWFI)

In order to obtain general information on the distribution of floods in the half-years of the hydrological year, a seasonal coefficient of floods developed in the winter half-year (SWFI) was proposed, expressed with following equation (1):

\[
SWFI = \frac{V_{FSW}}{V_{FAP} - 100\%}
\]

where:

- \(V_{FSW}\) – sum of flood volumes in winter half year (XI-IV);
- \(V_{FAP}\) – annual sum of flood volumes.

The coefficient takes values from 0 to 100%. Values close to 100% mean that there is practically no floods of the summer half-year, while close to 50% - seasonal equilibrium of floods developed in both seasons. An increase in the value of the coefficient above 50% indicates the predominance of floods in the winter half-year (September-June), while its decline means the dominance of the volume of floods in the summer half-year (V-X). This coefficient was adopted on the model of the coefficient proposed by E. Tomaszewski [19-20].

2.3 Markham seasonality indices

In 1970, C. G. Markham proposed a tool for assessing the many-year variability of precipitation in the United States [21-22]. Markham’s seasonality measures are also used in the seasonality studies of hydrological variables such as river runoff [10,23-25], physicochemical properties of surface and underground water [26], fluctuations of the groundwater [27], changes in the volume of deficits in low wages [19,28], and less changes in the volume of floods [18].

In this method, seasonality analysis consists of two indicators based on angular measures: IS - seasonality index and PSC - index of the time of concentration. Both measures, characterizing the seasonal distribution of the phenomenon, are based on 12 vectors presenting the monthly values of the investigated variable and the time of its occurrence. When constructing vectors, it is assumed that the value of the analysed feature in a given month is represented by a vector \((\mathbf{r}_m)\) which length is proportional to the value of this variable. The angle of inclination of individual monthly vectors \((\alpha_m)\) depends on the location of the centre of the month in question relative to the beginning of the hydrological year (1 XI) and is expressed by the formula (2):

\[
\alpha_m = \frac{360 \cdot L_m}{365}
\]

where:

- \(\alpha_m\) – the angle of inclination of the \(\mathbf{r}_m\) vector of monthly values [°];
- \(L_m\) – number of days between the beginning of the hydrological year and the middle of a given month.

As a result, there are 12 vectors created. From these vectors there is obtained resultant vector \(\mathbf{R}_S\) with the module \(|\mathbf{R}_S|\) and the direction \(\omega\) (Fig. 2). Then the value of the seasonality index \((IS)\) is calculated, that is the quotient of the resultant length of the vector \(|\mathbf{R}_S|\) by the sum of the length of the partial vectors \(|\mathbf{r}_{Si}|\):

\[
IS = \frac{|\mathbf{R}_S|}{\sum_{i=1}^{12} |\mathbf{r}_{Si}|} \cdot 100\%
\]

where:

- \(IS\) – Seasonality Index;
- \(|\mathbf{R}_S|\) – the length of the resultant vector from the \(\mathbf{r}_m\) vectors;
- \(|\mathbf{r}_{Si}|\) – the length of the vector corresponding to the value of the feature in the \(i\)-th month of the hydrological year.

The IS characteristics describe the degree of seasonality of occurrence of the examined feature and assume values between 0-100%. An increase in the value of the index means an increase in the seasonality of the occurrence of the phenomenon. The IS value equal to 0% can occur in two cases: firstly, with perfectly even distribution of the analysed feature (all 12 vectors have the same length) or with the concentration (exclusive occurrence) of the phenomenon in two opposite months.
of the year (shifted by half a year - e.g. March and September).

![Markham's seasonality coefficients construction procedure](image)

**Fig. 2.** Markham’s seasonality coefficients construction procedure. Explanations: \(X_3, X_{33},..., X_k\) - vectors corresponding to the sums of flood volumes in a given hydrological month; \(R_X\) - accidental vector for monthly vectors; \(\alpha_M\) - the inclination angle of the given month (e.g. November) from the beginning of the hydrological year; \(\omega\) - angle of incident vector (\(R_X\)).

The second measure of seasonality proposed by Markham is the Period of Seasonal Concentration (PSC). It is defined as the angle of inclination of the resultant vector \(R_X - \omega \) (Fig. 2). This measure indicates the time (expressed in days relative to the beginning of the hydrological year) of the concentration of the "mass" of the phenomenon and is calculated according to the equation (4):

\[
PSC = \arctg \left( \frac{\sum_{i=1}^{n} |r_{si}| \cos \alpha_{si}}{\sum_{i=1}^{n} |r_{si}| \sin \alpha_{si}} \right) \times \frac{365}{360}
\]

where:
- \(PSC\) – Period of Seasonal Concentration;
- \(|r_{si}|\) – the length of the vector corresponding to the value of the feature in the \(i\)-th month of the hydrological year;
- \(\alpha_{si}\) – the angle of inclination of the \(r_{si}\) vector of monthly values [°].

Despite the fact that \(PSC\) is expressed in days from the beginning of the hydrological year and indicates the time of concentration of the studied phenomenon, it should not be equated with the month in which the highest values of the variable occur, because it is the resultant value from all monthly values [27].

**2.4 Autocorrelation function (ACF)**

One of the statistical methods of seasonality analysis during the year, as well as long-term variability, is the analysis of the correlation (dependence) of selected flood characteristics between consecutive months of the hydrological year or hydrological years (long-term variability). In addition, the analysis of the autocorrelation function (ACF) of sequences of selected variables in particular months or hydrological years (including volume and duration of floods) is also carried out.

Autocorrelation is defined as the correlation of a series of \(X_t\) data with itself, but shifted by an observer (so-called delay or offset). It indicates the degree to which the next element of the sequence depends on the preceding ones. The autocorrelation function (ACF) assigns \(k\) (delay) to the autocorrelation coefficient between the variables \(X_t\) and \(X_{t+k}\) [29] as follows (5):

\[
R_k = \frac{1}{n-k} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})\; ACF = \frac{R_k}{r_1}
\]

The estimator of the autocorrelation function is expressed by the formula (6):

\[
S_n = \frac{1}{n} \sum_{i=1}^{n} sgn(X_i - X_{i+1})
\]

**2.5 Mann-Kendall nonparametric test**

The non-parametric Mann-Kendall test on the monotonic trend verifies the \(H_0\) hypothesis that there is no trend of the time series (variables \(X_i\), \(i = 1, 2, ..., n\) are independent and have identical distributions) against the alternative hypothesis about the occurrence of the trend. The hypothesis is verified according to statistics defined as [29-36] (7):

\[
S = \sum_{i=1}^{n-k} sgn(X_i - X_{i+1})
\]

The test statistic \(S\) with the truth of the null hypothesis \((H_0)\) has a normal distribution with an average of 0 and the variance described by the formula (8):

\[
\sigma^2 \approx \frac{1}{18} n(n-1)(2n-5)
\]

The above equality occurs in the absence of repeating elements of the series (the formula can be used when there are at most several of them). However, if there are a lot of repeating elements, the formula should be completed with an amendment (9):

\[
\sigma^2 \approx \frac{1}{18} n(n-1)(2n-5) \sum_{i=1}^{k} \frac{t_i(t-1)(2t+5)}{}
\]

In case \(n \geq 10\), the test statistic is expressed by following equation (10) and approximately fits to the normal distribution \(N (0, 1)\):

\[
U = \frac{S - sgn(S)}{\sigma_S}
\]

The null hypothesis of no trend is rejected at the significance level \(\alpha\) if \(|U| > u_{\alpha/2}\) (where \(u_{\alpha/2}\) is a quantile of order \(1-\alpha/2\) in the distribution \(N (0, 1)\)) or the value of statistics \(p_v\) is smaller than the assumed significance level \(\alpha\). If the \(U\) statistic is positive, then the trend of the analysed time series is growing (when negative – decreasing).
The basic assumption of the Mann-Kendall test for the presence of the trend is the independence and randomness of the data, which means that there is no autocorrelation in the time series. If this assumption is not met - variance is underestimated. In this case, the correction of variance introduced by Bayley and Hammersley should be applied [32] (it is calculated only for data with significant partial autocorrelation) (11):

\[ a^2 = 1 + \frac{2}{n(n-1)(n-2)} \sum_{t=1}^{n-1} \frac{(n-t-1)(n-1-2t)p(t)}{n(t+1)} \] (11)

In this paper statistical significance was assumed at the level of \(\alpha=5\%\). In the case of statistical significance of the results obtained at a different level, appropriate explanations are provided. The statistical calculations presented in this paper have been made in the MS Excel 365 ProPlus software (with XLSTAT extension), Statistica 12, Gretl 2018 and the R environment using the appropriate analytical packages.

3 Results

Among the many different floods characteristics, one of the most important information is the moment of their occurrence. Thanks to the analysis of seasonal dynamics, it is possible to determine in which period of the year floods occur most often, when their concentration takes place, and when they are practically non-existent. The dynamics of floods can be developed on an annual and long-term basis. In the first case, information is obtained on the distribution of floods during the year and its individual half-year (winter - summer). In the second one, we can get an information about changes in the basin's water resources and their changes in the form of occurrence of wet and dry years cycles.

3.1 Seasonality

Basic information on the flood distribution in the winter and summer half-year was obtained on the basis of the Seasonal Winter Flood volume coefficient developed in the winter half-year (SWFI). The highest share of floods formed in the winter half-year (73.7\%) occurred in Krasnystaw, while the lowest in Zwierzyniec (65.0\%). For Krzak gauging station, the value of the coefficient was 69.6\%. From the values obtained in the analysis, it appears that in all analysed catchments the prevalence of floods developed in the winter half-year (winter - summer). In the second one, we can get an information about changes in the basin's water resources and their changes in the form of occurrence of wet and dry years cycles.

Figure 3 graphically present Markham’s seasonality coefficients in upper Wieprz catchment standardised by subcatchment area.

In Krzak, in addition to the largest spring freshets (March-April), there are more floods with relatively large volumes and evenly distributed in particular months (especially May-June, September-October and January-February). The role of the occurrence of smaller but frequent floods in the summer and autumn (rain) and winter (slight thaw) is evident.

In Zwierzyniec, the extremely seasonal distribution of freshets is more visible. Higher significance (compared to Łabunka in Krzak) is achieved by spring floods, most often created from meltwater, however rainfall freshets are still noticeable (especially in autumn).

In Krasnystaw, on the other hand, spring (snowmelt) floods with large volumes strongly dominate, with a decidedly smaller number and scale of summer and autumn rainfall freshets. It is observed due to the geological structure of the catchment, where rocks with good permeability predominate. The majority of rainwater before it reaches the watercourses soaks up, and therefore small rainfalls observed in sub-basins are no longer noticeable at the level of the entire upper Wieprz river catchment.

The second measure of seasonality proposed by Markham – Period od Seasonal Concentration (PSC) - expressed in the number of days from the beginning of the hydrological year, has values from 141.4 days (from the beginning of hydrological year) in Zwierzyniec to 149.7 days in Krasnystaw (Fig. 3). In Krzak PSC reached the value of 146.3 days. In all analysed water gauging sections, the turn of the seasonal seasonality vectors indicated the second half of March (3rd decade). The earliest concentration time of freshets was observed in Zwierzyniec, and at the latest in Krasnystaw. This diversity is caused by the distribution of occurrence and...
size of floods, as well as their genesis. Large, thawing spring surges are much more common in Krasnystaw; they usually last longer, and start a few days later than in other profiles. In Zwierzyniec and Krzak, PSC adopted lower values also due to the occurrence of a higher number of freshets in the autumn (October-November), which contributes to the shift of the accidental vector towards the beginning of the hydrological year.

The analysis of the autocorrelation of the monthly time series of volumes and the duration of floods from the years 1951-2014 made it possible to determine the seasonality of freshets occurrence. The obtained results allow to conclude that the characteristics of floods set at the level of Q_{10%} in the analysed gauging stations show seasonal occurrence. In all profiles, statistically significant autocorrelations were found at the level of p=5% at shifts of 1-2 and 12 months. Values of autocorrelation coefficients reached relatively low values (maximum below 0.4). There is no autocorrelation relationship between flood volumes in delays from 3-4 to 10-11 months, what indicates the existence of one main flooding season during the year. A clear annual rhythm can be observed, however, the investigated catchments react in different ways to seasonal changes. Values of autocorrelation coefficients in all catchments change in a smooth way, and in some of delays they take on slightly negative values. Negative values of autocorrelation coefficients occurred in the largest number of delays in Krasnystaw (6 consecutive months).

Such a distribution of autocorrelation coefficients may suggest the functioning of different aquiferous horizons. Short shifts of statistically significant coefficients suggest the occurrence of local, suspended levels that are replenished during freshets in a relatively short time. Statistically significant autocorrelation coefficients occurring in 12-month shifts may, however, suggest the share of main aquiferous horizons which supply and replenishment occur in a much longer time than in the case of suspended ones.

3.2 Multiannual variability

Analysis of the long-term variability of floods and analysis of the variability of their parameters allows to study the dynamics of water resources in catchments. In years 1951-2014, in the analysed series, 178 to 338 floods occurred, however their number, duration and size were very diverse. In every gauging station there were years without floods. Most of them were observed in Krasnystaw (10 years), while the lowest in Krzak (3 years). Therefore, the most years with a flood (61) occurred in Krzak, and the least in Krasnystaw (54). The longest string of years without floods occurred in Krasnystaw in the years 1990-1992 and lasted three years. However, in the period 1990-1997, there were as many as 5 years without floods, and in 1993 only two small floods were observed. It was one of the driest periods in the 64-year study.

In the figure below (Fig. 4), the distribution of the number of days with floods in hydrological years 1951-2014 has been presented. Most days with floods were recorded in Zwierzyniec (203) in 1981, when large flows lasted from the beginning of October to the middle of May. In Krzak, the most days with freshet (over 150) occurred in 1967, 2000 and 2011, while in Krasnystaw in 1967, 1981, 2000, 2010 and 2011.

Analysing the distribution of the number of days with flood it is possible to determine wet periods, in which larger amounts of flooding days were observed. Two such periods can be determined: 1966-1982 and 1998-2014. They are separated by more dry periods when freshets are practically absent or only single, short episodes are observed. These periods, like the abundant periods in floods, are similar for the analysed stations. They occurred in the years 1951-1965 and 1983-1997. It is worth paying attention to a certain cyclicality of the alternating occurrence of dry and humid periods, which last in the studied catchments about 15 years (from 14 to 16).

The diversity of the distribution of freshets duration in years 1951-2014 in analysed cross-sections may be related to the occurrence of local flood events, caused by increased atmospheric supply in a small spatial scale (i.e. the emergence of torrential rain or cloudburst). Such local floods, whether on quite small watercourses (an example may be the episode in Piaski Szlacheckie) [37-39], or on the tributaries of Wieprz river, often significantly flatten and remain unnoticed in the further course of the river.

![Fig. 4. The annual and long-term courses of number of days with flood in Krasnystaw (A), Krzak (B) and Zwierzyniec (C).](https://example.com/figure4.png)
million m$^3$). It is worth noting that there are years in the gauging stations in Krzak and Zwierzyniec, when the annual volume of freshets in one profile is almost twice as high as in the second, despite the similar catchment area. This testifies to the high impact of environmental conditions (especially relief, coverage and land use) and the way of atmospheric supply.

Table 2 presents the results of the non-parametric Mann-Kendall test on the monotonic trend including the correction for autocorrelations. As results from the obtained test statistic values, in the analysed 64-year measurement strata of 3 water gauge sections, there are no statistically significant trends both in terms of volume and duration determined by the threshold flow at the level of $Q_{10\%}$.

<table>
<thead>
<tr>
<th>Gauging station, river</th>
<th>Shapiro-Wilk test</th>
<th>Mann-Kendall test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_1$</td>
<td>$U$</td>
</tr>
<tr>
<td>Flood volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krasnystaw, Wieprz</td>
<td>$&lt;0.0001$</td>
<td>0.4180</td>
</tr>
<tr>
<td>Zwierzyniec, Wieprz</td>
<td>$&lt;0.0001$</td>
<td>0.5041</td>
</tr>
<tr>
<td>Krzak, Łabańska</td>
<td>$&lt;0.0001$</td>
<td>0.2086</td>
</tr>
<tr>
<td>Flood duration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krasnystaw, Wieprz</td>
<td>$&lt;0.0001$</td>
<td>0.6797</td>
</tr>
<tr>
<td>Zwierzyniec, Wieprz</td>
<td>$&lt;0.0001$</td>
<td>0.9220</td>
</tr>
<tr>
<td>Krzak, Łabańska</td>
<td>$&lt;0.0001$</td>
<td>1.3102</td>
</tr>
</tbody>
</table>

4 Summary

Seasonal analyses of the occurrence of high water rises in the Wieprz River catchment in the years 1951-2014 showed that:

- In the research area, the floods of winter half-year prevail (most often from February to April); however, there was no period during the year when floods were completely absent.

- Concentration of the "mass" of flood volume falls on the last decade of March, in Zwierzyniec at the earliest, while in Krasnystaw, at the latest among the analysed cross-sections.

- In studied rivers, there is a clear seasonal freshets occurrence rhythm, while the distribution of them during the year in individual profiles is different. The seasonal schedule is clearly visible in Krasnystaw, where large spring thaw floods dominate. The least clear seasonality is shown in Krzak gauging station, where, in addition to the largest spring floods, the role of frequent, but smaller

in volume, freshets in summer, autumn and winter is evident. In Zwierzyniec, on the other hand, in addition to spring floods, the occurrence of freshets in the autumn season is slightly marked.

- Analysis of Markham's seasonality measures could be helpful in determining the time of concentration and distribution of the phenomenon throughout the year.

- Time series of the volume and duration of freshets in the analysed cross-sections do not meet the normality condition (according to the Shapiro-Wilk test). Therefore, a nonparametric test should be used to identify the trend in these series. The study of volume sequences and the duration of floods using the Mann-Kendall test on the monotonic trend did not show the existence of statistically significant trends, at $p = 0.05$.

References

1. Z. Mikulski, Zarys hydrografii Polski (PWN, Warszawa, 1963)