REGIONALIZATION METHODS FOR LOW FLOW ESTIMATION IN UNGAUGED CATCHMENTS – A REVIEW

Agnieszka Cupak✉

Department of Sanitary Engineering and Water Management, University of Agriculture in Krakow, al. Adama Mickiewicza 24/28, 30-059 Kraków

ABSTRACT

Aim of the study

The purpose of the paper is to look at the topic of estimation of the low flow characteristics at ungauged catchments in a balanced and systematic way, to provide a recent documentation of the subject and trace the most recent trends in low flow regionalization. The review focuses on methods, which are based on diverse methods that is the statistical, seasonal approach which are commonly used for the estimation of low flow characteristics in ungauged sites and which can be adopted in Polish conditions.

Material and methods

The paper examines the variety of literature sources published primarily during the last two decades. The most popular approaches to regionalization of watersheds are described: methods of residuals pattern approach, multivariate statistics such as cluster analysis, Classification and Regression Tree (CART) models or seasonality of low flows.

Results and conclusions

The described regional regression approach for low flow estimation in ungauged catchments, which correlates low flow indices and catchment characteristics, is the most widely used method. The methods were chosen because they differ in terms of methodological approach to using statistical methods. Also, the methods described in the article are "universal", which means that they can be adopted for different regions, countries, including Poland; they are interdisciplinary, and can be adopted for low flow calculation as well as average or flood flows.

Keywords: low flow, regionalization, regression, ungauged catchment

INTRODUCTION

Low flow is a natural phenomenon occurring seasonally and is an integral component of the flow regime (Vezza et al., 2010). WMO (1974) defines low flow as "the flow of water in a stream during prolonged dry weather". Low flows are normally derived from groundwater discharge and are influenced by a number of aspects, mainly geological, hydrological and climactic factors (air temperature and precipitation) (Mandal and Cunnane, 2009; Števková et al., 2012; Kozek, 2018). Information about low flow is essential for water management and planning in connection with the optimal use of water resources including environmental flow requirements, particularly the SNQ, which according to Polish legislation (Ustawa... 2017) is essential information for water-legal permit. In the aforementioned Act there is no specified method of SNQ calculation and such methodology will have to be developed by 31 December 2021 by National

✉e-mail: agnieszka.cupak@urk.edu.pl

Water Holding Polish Waters). Other key areas for which this is important include the quality and quantity of water for wildlife, water supply and irrigation systems, hydropower operation or recreation (Laaha, 2002; Laaha and Blöschl, 2006a; Vezza et al., 2010; Tomaszewski, 2018).

Whereas there exist sites under long-term observation (so that low flow characteristics can be estimated with the use of direct – statistical methods) still, in many cases streamflow information is not available or is insufficient in terms of quality and quantity (Pruski et al., 2015; Arai et al., 2012). A lot of effort was made to establish low flow procedures and to evaluate methods for the estimation of low flow parameters at ungauged sites (Schreiber and Demuth, 1997). In the literature, two different basic methods are presented, the stochastic and the deterministic approach (Smakhtin, 2001; Engeland and Hisdal, 2009). In the deterministic approach, a precipitation-runoff model is used to generate a continuous streamflow time series at ungauged sites from which the desired streamflow statistic can be extracted. In the stochastic approach, the streamflow statistics at ungauged sites are conditioned by the streamflow statistics at a gauged site, using either catchment descriptors or spatial distance as similarity measure (Engeland and Hisdal, 2009). In such cases, regionalization techniques can be used to evaluate low flow characteristics based on information from catchments where streamflow data have been collected (Demuth and Young, 2004; Laaha and Blöschl, 2006a). This type of low flow estimation is the most popular method in hydrology.

The regionalization techniques are widely used in i.e. Austria, Italy, Great Britain, Germany, Norway, Czech Republic, Turkey, and Slovakia. In Europe, the first statistical procedures for low flows were developed for catchments in the United Kingdom. The most extensive study to regionalize low flows was carried out in the framework of the FRIEND project (1985–1988) for the region of Western Europe (Gustard et al., 1989; Schreiber and Demuth, 1997). In this project following models were developed: global statistical models for the entire study area and detailed regional statistical models covering single countries or clusters of countries or single regions only (Gustard and Gross, 1989; Demuth, 1989; Schreiber and Demuth, 1997). Also the national procedure for low flow

estimation depending on available data was made in Austria. In Poland there is no generally established procedure for this issue. So, there is a need to develop procedures or methods, which can be applied in Polish conditions. In Poland, in the case of controlled catchments, statistical methods based on different type of distribution of extreme value are used, i.e. Gumbel method for low flow calculation (Byczkowski, 1972). For uncontrolled catchments, empirical formulas are used. There are few formulas for low and average flow calculation, but they were developed based on hydrological data for the years 1950–1980, for example Punzet formula (Punzet, 1981). At present, there is a need to verify or update these formulas, especially that actual hydrological data sequences are much longer at present (Wałęga et al., 2014). An assumption of the regionalization is that the variability in space of flow characteristics can be explained by drainage basins characteristics (Clarke, 1977; Vogel and Kroll, 1992; Mazvimavi et al., 2005). The regionalisation of streamflow characteristics in general is based on the premise that catchments with similar climate, geology, topography, vegetation and soils would normally have similar streamflow responses, for example, in terms of unit runoff from the catchment area, average monthly flow distribution, duration of certain flow periods, etc. (Smakthin, 2001). The regionalisation of low flow characteristics is usually based on regression between the low flow characteristics of interest and catchment characteristics that are available for ungauged sites (Vogel and Kroll, 1992; Gustard et al., 1992; Schreiber and Demuth, 1997; Skop and Loaiciga, 1998; Ziernicka-Wojtaszek and Kaczor, 2013; Laaha and Blöschl, 2006a).

In some instances it is clear how to group studying catchments into regions, but very often this procedure is not so obvious. The definition of regions may be accomplished by using convenient boundaries based on geographic, administrative, or physiographic assumptions. However, the regions thus obtained may not always be 'sufficiently' homogeneous from the hydrological point of view. Even two adjacent river catchments may have rather different topography, soils or other local anomalies. A homogeneous region may therefore be viewed as a collection of catchments, which are similar in terms of catchment hydrological response, but not necessarily geographically contigu-

ous (Smakthin, 2001; Laaha and Blöschl, 2006a). Shu and Burn (2003) suggested that geographically close catchments are not necessarily homogeneous in terms of hydrological response. In their case study in Great Britain, homogeneous spatial clustering patterns were found within a 62.5 km radius of the local clustering centre. Burn and Boorman (1993) assigned donor catchments based on a similarity measure of physiographic catchment characteristics (Parajka et al., 2005). Also, with classification techniques such as cluster analysis, it is not necessary for the catchments to form a given group, but it is necessary for them to be geographically contiguous (Nathan, 1990).

A number of methods for identifying homogenous regions have been proposed in the literature in the context of low flow regionalisation. All of these methods use low flow data, which are under consideration, and most of them use catchment characteristics as well (Laaha and Blöschl, 2006a).

The purpose of the paper is to look at the topic of estimation of the low flow characteristics at ungauged catchments in a balanced and systematic way, to provide a recent documentation of the subject, and trace the most recent trends in low flow regionalization. The review focuses on methods, which are based on diverse methods – that is the statistical, seasonal approach – which are commonly used for the estimation of low flow characteristics in ungauged sites, and which can be adopted in Polish conditions. In the area of Poland, the climate – and most distinctly, the temperature and precipitation – are driven by altitude. Generally, the higher the altitude, the colder and wetter the climate; with the exception of Suwałki region in the north part of Poland. The mean annual temperature varies between 6–8.5°C. The warmest part is the Silesian Lowland, with the mean annual temperature about 8.5°C and the Sandomierz basin (8°C). The winter is the coldest season when the mean annual temperature drops to 0.8°C in the Kasprowy Wierch Mountain. The mean annual precipitation in Poland is about 600 mm, while the highest – about 1400 mm – is recorded in the south part of Poland (mountain catchments) and the lowest – about 500 mm – in the central part of the country (lowland catchments). Also, Poland is diverse in terms of physical and geographical parameters. The mean altitude of Poland is 173 m with a range from 1.8 m.b.s.l. (the point near Raczki Elbląskie village)

to 2499 m.a.s.l. (Rysy Mountain in the High Tatras). In the area of Poland, there are 4 main morphogenetic zones: Young Carpathians with mountain valleys, Sudety with the Uplands: Silesia-Krakow and Małopolska, Lublin-Lviv, The Saxon-Lusatian Lowlands and the Central Poland Lowlands, and the South Baltic Coastland Lakelands (www.encyklopedia.pwn.pl). The paper examines the variety of literature sources published primarily during the last two decades.

REGIONAL REGRESSION APPROACH

In hydrology, the most popular approaches to regionalization of watersheds include: methods of residuals pattern approach, multivariate statistics such as cluster analysis, Classification and Regression Tree (CART) models or seasonality of low flow (see: Table 1). Regionalization methods are applied in searching for hydrological similarity between catchments by examining the attributes that describe their geomorphology, land cover, climate, and soils, for low flow characteristics estimation in ungauged catchments (Ahuja, 2012; Vezza et al., 2010). Regionalisation techniques can be applied in estimating low flow characteristics under consideration, in order to obtain a predictive model for low flows in ungauged catchments as preparatory reference for the environmental flows assessment (Vezza et al., 2010).

The methods of low flow regionalisation, which are described in the article, were chosen because they differ in terms of methodological approach to using statistical methods. This is to say that in cluster analysis, the similarity (or dissimilarity) of objects is taken into consideration, as well as seasonality of the analysed phenomena using circular statistics. These methods are "universal", which means that they can be applied to different regions, countries, including Poland; they are interdisciplinary, and thus can be adopted in low flow calculation as well as average or flood flows – the first analysis of seasonality of flows was conducted by Burn (1997) for flood seasonality analysis, and then it was adopted for low flow. Cluster analysis is another popular tool in hydrology, but it can also be applied to other discipline, such as medicine or economy. Another aspect is that its use in other countries seems to be promising, and can be a useful tool for low flow estimation in uncontrolled catchments.

Table 1. Description of regionalization methods used for low flow characteristics' regionalization

Regionali- zation method	Low flow characteristics which can be established	Parameters that can be used for regionalization	Application of the method	Advantages and disadvantages of the method
Residual pattern approach	the low flow characteristics, which are dictated by the engineering or water planning of particular country (Smakthin 2001) i.e. Q_{50}; Q_{75}; Q_{90}; Q_{95}; 7Q10; 7Q2; 30Q10; 30Q; MAM(10) etc.	Climatic and physiographic characteristics: catchment area; precipitation; geology; land use; soils; stream network density; topographic elevation i.e.: altitude of the stream gauge; maximum altitude; mean altitude; range of altitude; topographic slope; topographic slope i.e.: mean slope; moderate slope;	Vezza et al., 2010; Laaha and Blöschl, 2006a	Advantages: effective identification of regional differences; Disadvantages: the initial model may be far from correct and the shape of the regions may then be artefacts of an inadequate model, regional regression model can have little physical significance (Laaha and Blöschl, 2006a)
Classifi- cation and regression tree	as above	as above	Vezza et al., 2010; Laaha and Blöschl, 2006a	Advantages: the CRT is nonparametric, is invariant to monotone transformations of its independent variables, easily handles outliers and noisy data isolating them in a separate node, and the trees obtained are readily interpretable (Breiman et al., 1984; Vezza et. al., 2010); small trees are readily interpretable (Laaha and Blöschl, 2006a); Disadvantages: the method consists of having unstable results with the modification of the learning sample (the structure of the tree may change when models are refitted for subsets of the data (Vezza et. al., 2010); big trees are difficult to interpret, there is a lack of smoothness and there are potential problems with overfitting the data; a method for pruning the tree is needed (Laaha and Blöschl, 2006a)
Cluster analysis	as above	as above	Vezza et al., 2010; Števková et al., 2012; Laaha and Blöschl, 2006a; Nathan and McMahon, 1992; Kahya and Demirel, 2007; Ahuja, 2012; Cupak, 2013; Cupak, 2017	Advantages: one of the most frequently used methods in hydrology; few clustering methods can be used; the data is grouped in such a way that the data points in one cluster are very similar and objects in different clusters are quite distinct; in partitional clustering the number of clusters need to be assumed; K-means method is effective for grouping large number of data (Rao and Srinivas, 2008)

Table 1. cont.

Regionali-zation method	Low flow characteristics which can be established	Parameters that can be used for regionalization	Application of the method	Advantages and disadvantages of the method
				Disadvantages: depending of clustering methods different groups can be obtained; there is no clear formal criterion for determining clusters; in partitional clustering the number of clusters need to be assumed; increase in the number of catchments resulted in segregation of a collection of sites that are highly heterogeneous; K-means method is sensitive to presence of errors (Rao and Srinivas, 2003, 2008); requires the selection of variables that are used to define similarity/ dissimilarity for the catchments (Burn, 1997)
Seasona-lity	as above	as above	Vezza et al., 2010; Laaha and Blöschl, 2006a; Laaha and Blöschl, 2006b; Kohnová et al., 2009; Engeland and Hisdal, 2009	Advantages: can be a useful indicator of catchment similarity in terms of hydrological processes; can separate of summer and winter low flows Disadvantages: the SR gives the least information about low flows seasonality; the synoptic interpretation of seasonality measures from a large number of catchments could be difficult because of increasing number of parameters, the SH required some classification technique to exploit the higher information content (Laaha and Blöschl, 2006b)

RESIDUAL PATTERN APPROACH (RPA)

In the residual pattern approach (RPA), residuals from an initial, global regression model between flow characteristic and catchment characteristics are plotted, from which geographically contiguous regions are obtained by manual generalisation on a map (Hayes, 1992; Aschwanden and Kan, 1999a; Laaha and Blöschl, 2006a).

The residual pattern approach to catchment grouping consists of three steps (Vezza et al., 2010):
1. Perform stepwise regression to obtain a global regression model;
2. Plot the residuals from the global regression model in the geographic space;
3. If residual patterns are apparent, delineate contiguous regions of similar sign and magnitude of residuals.

The delineation of sub-regions in Switzerland was based on the residual pattern approach (Hayes, 1992). A map that combines low flow measurement of gauged catchments and low flow estimates of selected ungauged catchments was compiled and published in the Hydrological Atlas of Switzerland (Aschwanden and Kan, 1999b). Based on residual pattern approach, Laaha and Blöschl (2006a) classified Austria's catchment into two main units. The first one consists of flatland and hilly terrain, with the low magnitude of residuals. The second unit consists of the Alpine catchments and the Molassezone in the North (Laaha and Blöschl, 2006a). Also, Vezza et al. (2010) used RPA in the regionalization of low flows in North-Western Italy. As a result of its application, three regions were identified, which are: the South-Western Alps, characterized by not extremely high mountains and nivo-pluvial streamflow regimes; the Valle d'Aosta Region (including

Orco and Stura di Lanzo rivers) with higher mountains and presence of glaciers; and the Northern-Eastern very wet Sesia and Toce watersheds (Vezza et al., 2010).

CLASSIFICATION AND REGRESSION TREE (CART)

The Classification and Regression Tree (CART) approach is a classification method, which uses historical data to construct the decision trees. The CART was used for the first time in low flows regionalization by Laaha and Blöschl (Vezza et al., 2010). Also, Vezza et al. (2010) used CART in their study of low flow regionalization in Italy. For building decision trees, CART splits a learning sample (low flows and catchment characteristics) by using an algorithm known as binary recursive partitioning (Breiman et al., 1984; Laaha and Blöschl, 2006a; Vezza et al., 2010). By this algorithm, groups of catchments are subsequently subdivided by binary conditions, starting from the most important catchment characteristics and proceeding to the less important ones (Laaha and Blöschl, 2006a). The tree stops growing when each terminal node consists of one single observation, but this leads to overfitting. To avoid such overfitting, trees need to be pruned back, and the optimal number of nodes is best determined by an independent validation data set (Laaha and Blöschl, 2006a; Vezza et al., 2010). This procedure, called 10-fold cross-validation, is based on minimizing the average prediction error. The CART algorithm has a very good performance in catchments if one is interested in finding groups that are most distinct in terms of both catchment characteristics and low flow catchment response (Laaha and Blöschl, 2006a; Vezza et al., 2010).

The regression tree approach to catchment grouping consists of the following steps (Laaha and Blöschl, 2006a):
1. Perform transformation to normality.
2. Fit an initial regression tree to the data.
3. Determine the optimal tree size by 10-fold cross -validation.
4. Prune the initial tree back to the tree size derived in point 3.

Looking at the regression model fitted to the Apennine-Mediterranean region identified by the CART, the only relevant predictor variable was the maximum altitude (H_{max}), which was positively correlated to low flow characteristics. The CART groups together only the small catchments located in the highlands and moors in Valle d'Aosta, and it identifies a number of small high-elevation catchments in which the intense drought period during winter has the soil freezing processes as the driving force (Vezza et al., 2010). In Laaha and Blöschl study (2006a), the use of regression tree divided Austria into seven regions. The regression tree grouping performed the second best among other used methods (They used 4 different catchment regionalization methods).

CLUSTER ANALYSIS

Clustering is a process by which a set of feature vectors is divided into clusters or groups such that the feature vectors within a cluster are as similar as possible, whereas the feature vectors of different cluster are as dissimilar as possible. Cluster analysis has been used in many studies of low flow regionalization e.g. Gottschalk (1985), Rao and Srinivas (2006a,b), Lin and Chen (2006), Laaha and Blöschl (2006a), Kahya and Demirel (2007), Števková et al. (2012), Cupak (2017) and Cupak et al. (2017).

Since clustering is the grouping of similar instances/ objects, some sort of measure that can determine whether two objects are similar or dissimilar is required. There are two main types of measures used to estimate this relation: distance measures and similarity measures. Distance measures are used most often (see: Table 2).

Clustering algorithms can be broadly classified into two categories: hierarchical and partitional clustering (see: Fig. 1).

In hierarchical algorithms, two categories can be distinguished: agglomerative and divisive. In the agglomerative category, the following methods can be used (see: Fig. 1): complete and single linkage, pair group average, weighted pair group average, pair-group method using the centroid average, weighted pair-group method using the centroid average, and Ward's method (Cupak, 2013). In this type of clustering, the number of clusters is given as the result. The disadvantage of these methods is that there is no clear formal criterion for determining clusters. Therefore, a chart illustrating agglomeration course can be used as a criterion for stopping the agglomeration process and final verification of the objects (catchments) (Cu-

pak et al., 2017). The partitional clustering procedure, unlike the hierarchical procedure, requires an initial assumption about the number of clusters and cluster centres (Rao and Srinivas, 2006b). Partitional methods include the K-means, K-medoids, and K-modes methods (Rao and Srinivas, 2008).

Table 2. The measures used for distance evaluation between clusters centres (Rao and Srinivas, 2008)

Distance measure	Equation
Euclidean distance	$\sqrt{\sum_{k=1}^{n}(x_{ik} - x_{jk})^2}$
Mahalonobis distance	$\sqrt{(x_i - x_j)^T \, \Sigma^{-1}(x_i - x_j)}$
Manhattan distance	$\sum_{k=1}^{n}\left|x_{ik} - x_{jk}\right|$
Chebychev distance	$\max_{1 \le k \le n}\left|x_{ik} - x_{jk}\right|$
Minkowski distance	$\left(\sum_{k=1}^{n}\left|x_{ik} - x_{jk}\right|v^t\right)^{\frac{1}{t}}$

where: n – number of attributes; x_{ik} – attribute k of x_i vector function for cluster 1; x_{jk} – attribute k of x_j vector function for cluster 2; T – transpose matrix, Σ – covariance matrix; t – order of Minkowski distance

In hydrology and climatology, K-means and Ward's method are the most commonly used for low flow regionalization. Rao and Srinivas (2006b) used cluster analysis on the attributes and flow records from 245 gauging stations in Indiana. They used hierarchical and non-hierarchical analyses (Števková et al., 2012). Kahya and Demirel (2007) used hierarchical clustering (single, complete and Ward methods) to create regional types of monthly low flows in Turkey. They found that the Ward's method with Euclidean distance was more effective in the production of homogenous clusters compared to single linkage and complete linkage methods. Laaha and Blösch (2006a) and Vezza et al. (2010) in their studies used a combination of different distance measure and linkage methods for a range of numbers of clusters, and found that Ward's method
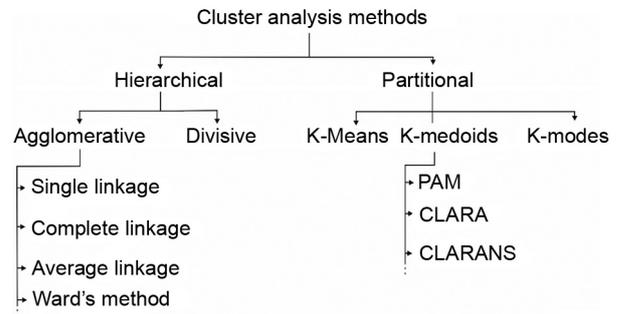
Fig. 1. Cluster analysis methods (Rao and Srinivas, 2008)

and Euclidean distance metric was preferable to other methods. The same conclusion about the best clustering method of grouping (Ward's method) was drawn by Cupak (2013), who used different hierarchical methods for 34 catchments' grouping located in the Upper Vistula basin (Poland), based on hydrological, meteorological and physiographical catchment parameters. Also, Cupak et al. (2017) used Ward's method for the determination of hydrologically homogeneous regions with low flows.

In the Ward's algorithm (Ward, 1963) the distance between clusters is estimated by the analysis of variance, therefore, it helps to minimize the sum of squared deviations of the feature vectors from the centroid of their respective cluster, as in formula (1) below (Rao and Srinivas, 2006b).

$$W = \sum_{k=1}^{K}\sum_{j=1}^{m}\sum_{i=1}^{N_k}\left(y_{ij}^{k} - y_j^{k}\right)^2 \qquad (1)$$

The Ward's algorithm starts with a singleton cluster. At this point the cluster centers are the same as feature centers. Therefore, the value of the objective function is zero. At each step in the analysis, union of every possible pair of clusters is considered, and two clusters whose fusion results in the smallest increase in W are merged. The change in the value of the objective function due to the merger depends only on the correlation between the two merged clusters and not on the correlations with other cluster (Rao and Srinivas, 2006b). Ward's algorithm is effective in recovering cluster structure, and it tends to form spherical clusters of equal size. This characteristic of the Ward's algorithm makes it useful in the identification of ho-

mogeneous regions for regionalization (Hosking and Wallis, 1997; Rao and Srinivas, 2006b).

The author used Ward's method in Poland for 15 selected catchments located in the upper Vistula basin, and 2 clusters were selected (see: Fig. 2). The correlation was examined between specific low flow – q_{95} and physiographic and meteorological parameter of catchments. Correlation and regression relationships were determined separately for each cluster. In the first cluster, comprising a small number of catchments, only the correlations between q_{95} and individual independent variables were determined. The catchments of this cluster were devoid of cambisols and gravel-based eutric cambisols, and so they were not accounted for in the correlation model. No significant correlations were found for the significance level of 0.05. However, when the significance level was increased to 0.10, a significant correlation was found for the percentage of the area covered by fluvisols and arable lands (Cupak et al., 2017).

In the second cluster, comprising 10 catchments, the multiple regression model formula appeared as follows (Cupak et al., 2017):

$$q_{95} = 1.726 + 19.063 \cdot I - 0.052 \cdot S_M + 0.0083 \cdot P - 0.0076 \cdot H_{Me} + 0.0044 \cdot P_z$$

In the formula, the parameters significant for the significance level of 0.05 included mean catchment slope (I), presence of fluvisols (S_M), precipitation (P), and median altitude of the catchment (H_{Me}).

The K-means method is the best known of the non-hierarchical clustering methods. Burn (1989) used the K-means clustering algorithm to determine appropriate grouping of a network of streamflow gauging stations in southern Manitoba, Canada. Lecce (2000) also used the K-means method to examine spatial variations in the timing of flooding in the southeastern United States (Lin and Chen, 2006). Burn and Goel (2000) applied the K-means algorithm to site characteristics (catchment area, length and slope of the main stream of river) of a collection of catchments in India to derive regions for flood frequency analysis (Rao and Srinivas, 2008). In India, Ahuja (2012) used K-means method for regionalization data of Godaravi catchments. In Poland, Kowalczak (1986) used non-hierarchical cluster analysis for catchments classification for the Upper Noteć basin (Poland). Also, Cupak (2017) used this algorithm for the grouping of catchments, located in the Upper Vistula basin, with respect to low flow characteristic.

K-means algorithm is based on minimisation of the objective function, and it breaks down the set of data into clusters. It has the properties of dividing and clas-
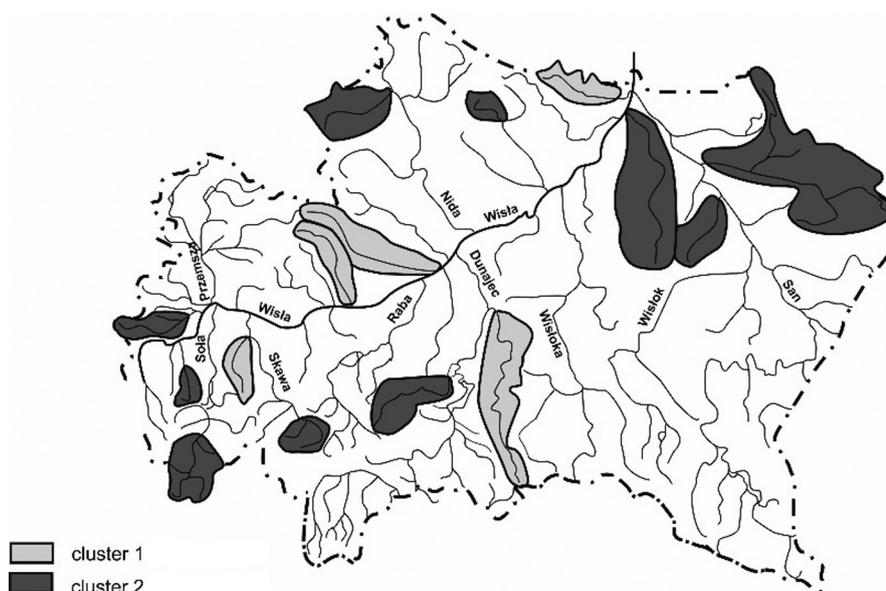
Fig. 2. Grouping of catchments located in the Upper Vistula basin with the use of Ward's method (Cupak et al., 2017)

sifying l variable numbers and N numbers of data of the X set into the K number of clusters. Similarity of the elements in the given cluster should be as high as possible, and specific clusters should maximally differ from each other (Dikbas et al., 2013). This method is effective for grouping large sets of data with numerical attributes. However, there are some limitations to this method in the breaking down of the data into categories. The method is also sensitive to presence of errors (Rao and Srinivas, 2008). In this method, the procedure is as follows: at first, the objects (catchments) are assigned to clusters at random, and then iteratively they are moved between clusters, in order to minimize intra-group changeability and maximize intergroup changeability (Cupak, 2017). The formula (2) below describes iteration of the function with variables k, j, i.

$$F = \sum_{k=1}^{K} \sum_{j=1}^{n} \sum_{i=1}^{N_k} d^2 \left(x_{ij}^k - x_{.j}^k \right) \qquad (2)$$

where:
- K – the number of clusters,
- N_k – the number of functions of vectors in the cluster,
- x_{ij}^k – the scale of the value of the function of the vector assigned to the cluster k,

$x_{.j}^k$ – the average value of the feature j for the cluster k, calculated from the formula (3) below:

$$x_{.j}^k = \sum_{i=1}^{N_k} x_{ij}^k \cdot \frac{1}{N_k} \qquad (3)$$

Thus, the calculation procedure is repeated until variability of the features within the given cluster is as low as possible.

In the author's research conducted in the area of the Upper Vistula basin, the correlation was examined between specific low flow – q_{95} and physiographic and meteorological parameters of the catchments. The use of K-means method facilitated the identification of two groups of catchments sharing similar values of low flows. The identified clusters comprised catchments similar in terms of unit runoff, watercourse length, mean precipitation, median altitude, mean catchment slope, watercourse staff gauge zero, the area covered by coniferous forests, arable lands, and soils. In the K-means method, the procedure of assigning catchments to the clusters was started with two clusters, and it finished with five. The best fit was obtained in the instance of choosing two clusters. In the first cluster, there were five catchments; and in the second, another eleven (see: Fig. 3) (Cupak, 2017).

Fig. 3. Grouping of catchments located in the Upper Vistula basin with the use of K-Means method (Cupak, 2017)

The catchments of this cluster were devoid of grassland (U_z), and so they were not accounted for in the correlation model. No significant correlations were found for the significance level of 0.05. Even though in case of such parameters as median catchment altitude and soils (fluvisols and eutric cambisols) the correlation was weak (correlation coefficient under 0.3), it was decided to show both the correlation model and the correlation coefficient. In the second cluster, including 14 catchments, a multiple regression model was formulated as follows:

$$q_{95} = -1.096 + 0.016 \cdot U_z + 0.052 \cdot L_i - 0.02 \cdot S_p + \\ + 0.29 \cdot H_{Me} + 0.532 \cdot T - 106.0 \cdot i - 0.013 \cdot P + \\ + 0.036 \cdot S_{Bw}$$

In the above equation, at a significance level of 0.05, only coniferous forests (L_i) were statically important (Cupak, 2017).

At present, more detailed research results are available covering a larger number of catchments located within the Vistula basin, and obtained with use other methods of grouping.

SEASONALITY METHODS

As shown in the Austrian study, the seasonality indices approach has a potential in identifying homogenous groups of catchments using differences in seasonal low processes between the catchments (Laaha and Blöschl, 2006a; Vezza et al., 2010). Engeland and Hisdal (2009) compared a regional regression model based on seasonality with a regional rainfall-runoff model (HBV) in Norway for regionalizing low flows, and they found that the first one performs better than the second. A regional analysis of low flows in small catchments in Austria was carried out by Laaha (2006) and Laaha and Blöschl (2006a, b). Schreiber and Demuth (1997) analysed seasonality of mean annual 10-day minimum MAM(10) of total discharges measured in 169 catchments in southwest Germany. The results indicated typical low flow occurrence from September to October for large parts of the study area, apart from the Pre-Alps (Voralpen region), which are dominated by winter low flows (January and February). The differences of low flow seasonality were found to depend mainly on catchment altitude. In Switzerland, Aschwanden and

Kan (1999) investigated the seasonal distribution of Q_{95} for 143 headwater catchments. They found two different typical seasonal distributions of low flows, which – as in case of Germany – depended on catchment altitude. In alpine catchments, low flows occur from November to March. In the hilly landscapes of Mittelland and Jura, low flows may occur during the whole year, but clearly most frequently during summer and autumn. In the region of Vermont and New Hampshire, Dingman and Lawlor (1995) stated that, annual 7-day minimum flows usually occur in late summer or early fall in response to regional climatic patterns, but they occur in some years during late winter in the more northern and high-elevation streams. The mean time of occurrence for annual 7-day minimum flows is in August (Vermont and the Connecticut River basin), in September (the Saco River basin), and in August or September in the rest of New Hampshire. But at the highest elevations, annual 7-day minimum flows occur in February (Laaha and Blöschl, 2007). Jokiel and Stanisławczyk (2016) analysed the changes and the multiannual variability of the seasonal structure of river runoff in Poland. Also, the seasonality of median monthly discharge in selected Carpathian rivers of the upper Vistula basin was analysed by Wałęga and Młyński (2017).

The most often used seasonality parameters in case of low flow characteristics are: the seasonality ratio (SR), a cyclic seasonality index (SI), and seasonality histograms (SHs).

The seasonality ratio (SR). Summer and winter low flows are subject to important differences in the underlying hydrological processes. Daily discharge time-series need to been stratified into summer and winter discharge. From winter and summer discharge time-series, characteristic values for summer low flows q_s and winter low flows q_w are calculated for each catchment. The SR is calculated using formula (4) below (Laaha and Blöschl, 2006b):

$$SR = \frac{q_s}{q_w} \qquad (4)$$

Values of SR > 1 indicate the presence of a winter low flow regime, whereas values of SR < 1 indicate the presence of a summer low flow regime.

The seasonality index (SI). This index is similar to Burn (1997) and Laaha and Blöschl (2006b), and it

represents the seasonal distribution of low flow occurrence. The index is based on two parameters, θ and r, which are calculated from the Julian dates of all days of the observation period, when discharges are equal or below low flow characteristic. The mean day of occurrence θ is measured in radians and reflects the average seasonality of low flows. Parameter r is the dimensionless mean of the days of occurrence. The parameter describes the variability of low flow seasonality, which varies between 0 (low seasonality) and 1 (strong seasonality) (Laaha 2006a, b). Flows lower than low flow characteristics under consideration are identified by the day of the occurrence D_j, which can be described by the following equation (Burn, 1997; Laaha and Blöschl, 2006b):

$$\theta = \frac{D_j \cdot 2\pi}{365}$$

The Cartesian coordinates x_θ and y_θ of the total of the n single days is calculated as: $x_\theta = \frac{1}{n} \sum_j \cos(\theta_j)$ and $y_\theta = \frac{1}{n} \sum_j \sin(\theta_j)$. Then, the angle of the mean directional vector is defined as: $\theta = \arctan\left(\frac{y_\theta}{x_\theta}\right)$ 1st and 4th quadrant: $x > 0$ and $\theta = \arctan\left(\frac{y_\theta}{x_\theta}\right) + \pi$ 2nd and 3rd quadrant: $x < 0$. By transforming the mean angle to a Julian date, the mean day of occurrence is determined: $D = \theta \cdot \frac{365}{2\pi}$. Finally, the seasonal concentration index r is calculated as: $r = \sqrt{x_\theta^2 + y_\theta^2}$. As the pooling variable, the seasonal concentration index and the X, Y coordinates of the mean day of the occurrence of low flows are used. Seasonality indices can then be displayed on a vector map, which gives a synoptic representation of the mean day of occurrence and the intensity of seasonality for a large number of catchments (Laaha and Bloschl, 2006b).

The seasonality histogram (SH). The SH allows a more detailed description of seasonal distribution of low flows than the SI (Laaha and Bloschl, 2006b). The description is based on Julian dates of all days of the observation period when discharges are equal or below low flow characteristic. Histograms based on monthly classes are plotted from these data. The SH illustrates the occurrence of low flow in each month, and provides supplemental information to the SI. It illustrates which months are affected by low flow, and it provides a good representation of the shape of the seasonal distribution – multimodal or skewed distribution (Laaha and Bloschl, 2006b).

CONSTRUCTION OF REGRESSION MODEL

Regional regression is performed by building a multi-regressive model, which relates the Q_{LF} (dependent variable) to morphoclimatic descriptors (independent variables) with the view to selecting the most influential descriptors for low flows regionalization – see formula (5) below (Vezza et al., 2010).

$$Q_{LF} = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + ... + \beta_{p-1} \cdot x_{p-1} \qquad (5)$$

where:
   x_i – morphoclimatic parameters of the catchment,
   β_i – regression coefficient.

To establish a regression relationship, the stepwise regression approach is commonly used, when the model is derived one step (i.e. one independent variable) at a time (Smakthin, 2001; Laaha and Blöschl, 2006a; Vezza et al., 2010). The streamflow data used should represent natural flow conditions in the catchments. This means that the approach will most probably not work or will be misleading if flow regimes analysed are continually changing under man-induced impacts (Smakthin, 2001). Basin and climate characteristics (independent variables) that are the most commonly related to low-flow indices include: catchment area, mean annual precipitation, channel and/or catchment slope, stream frequency and/or density, percentage of lakes and forested areas, various soil and geology indices, length of the main stream, catchment shape and watershed perimeter, and mean catchment elevation (Smakthin, 2001).

CONCLUSIONS

The paper focuses on the subject of the methods of the low flow characteristics estimation in ungauged sites, as in Poland there is no generally established procedure for this purpose. Controlled sites are not a problem, where statistical methods based on various types of distribution of extreme value are used, i.e. Gumbel

method for low flow calculation (Byczkowski, 1972). But for uncontrolled catchments there is a need for the verification or updating of empirical formulas, especially since the currently used formulas for low and average flow calculation had been developed based on hydrological data for the years 1950–1980 (Wałęga et al., 2014). The paper considers the variety of the research, as well as national procedures that have been applied in other countries, and the methods that can be adopted in Polish conditions. Information about low flows is essential for the quality and quantity of water for wildlife, for water supply and irrigation systems, for hydropower operation or recreation, and for water management and planning, such as optimal use of water resources calculation (Laaha, 2002; Laaha and Blöschl, 2006a; Vezza et al., 2010; Tomaszewski, 2018). This very important issue concerns in particular the environmental flow requirements, especially the SNQ, which according to Polish legislation (Ustawa... 2017) is essential information in the context of water-legal permit. Still, the aforementioned Act does not specify the method of SNQ calculation.

The described regional regression approach for low flow estimation in ungauged catchments, which correlates low flow indices and catchment characteristics, is the most widely used method. The methods were chosen because they differ in terms of methodological approach to using statistical methods. Namely, in cluster analysis, the similarity (or dissimilarity) of objects is taken into consideration, as well as seasonality of the analysed phenomena using circular statistics. Also, the methods described in the article are "universal", which means that they can be adopted for different regions, countries, including Poland; they are interdisciplinary, and can be adopted for low flow calculation as well as average or flood flows – the first analysis of seasonality of flows was made by Burn (1997) for flood seasonality analysis, and then was adopted to low flows. Also cluster analysis is a popular tool in hydrology, but it can also be applied in other disciplines, such as medicine. In fact, the first use of cluster analysis was in economy.

The methods described herein, and their use in other countries, seem interesting and can be a useful tool for low flow estimation in uncontrolled catchments. However, based on subject literature alone, it is difficult to point which method will be the best and should be used

for Polish catchments, as the best classification method is site dependent and cannot be chosen a priori. And so, the best results of regionalization in case of Austrian catchments were produced by the method of seasonality of low flows, while in Italy, regression and classification tree proved to be the best. Also, in order to decide which grouping methods perform the best, the same data set needs to be taken to analysis. Despite the significant amount of specialist knowledge, which has been accumulated with the view to low flow characteristics estimation, the calculation of this parameter in ungauged catchments seems to be limited. This is probably the result of fewer studies in the field of low flow compared to the studies of floods. On the other hand, there is an increasing need for the evaluation of low flow characteristics, especially in small, ungauged catchments. Although the paper has examined a number of recent literature sources, it is impossible to include in the review all the relevant publications that have been published during the last two decades. It is therefore possible that some aspects has been overlooked, and that could be a scope for any future study and discussion in the context of low flow estimation in ungauged catchments.

REFERENCES

Ahuja, S. (2012). Regionalization of River Basins Using Cluster Ensemble. Journal of Water Resource and Protection, 4, 560–566.

Arai, F.K., Pereira, S.B., Goncalves, G.G.G. (2012). Characterization of water availability in a hydrographic basin. Engenharia Agricola, Jaboticabal. 32, 3, 591–601.

Aschwanden, H., Kan, C. (1999a). Le débit d'étiage Q347--Etat de la question. Communications hydrologiques, 27, Servies hydrol. et geol. National, Berne.

Aschwanden, H., Kan, C. (1999b). Niedrgwasser-Grundlagen zur Bestmmung der Abflussmenge Q347. Tafel 5.8 in: Hydrologischer Atlas der Schwez, 3. Lieferung. Landeshydrologie und geologie, Berm, Switzerland.

Breiman, L., Friedman, J.H., Olshen, R., Stone, C.J. (1984). Classification and regression trees. Belmont: Wadsworth International Group.

Burn, D.H. (1989). Cluster analysis as applied to regional flood frequency – analysis. Journal of Hydrology, 104 (1–4), 345–361.

Burn, D.H. (1997). Catchments Similarity for Regional Flood Frequency Analysis Using Seasonality Measures. Journal of Hydrology, 202(1), 212–230.

Burn, D.H., Boorman, D.B. (1993). Estimation of hydrological parameters at ungauged catchments, J. Hydrol. 143, 429–454.

Burn, D. H., Goel, N. K. (2000). The formation of groups for regional flood frequency analysis. Hydrol. Sci. J., 45(1), 97–112.

Byczkowski, A. (1972). Hydrologiczne podstawy projektowania budowli wodno-melioracyjnych. Warszawa: Państwowe Wydawnictwo Rolnicze i Leśne.

Clarke, R.T. (1977). A review of research on methods for extrapolation of data and scientific findings from representative and experimental basins. Technical Documents in Hydrology, Paris: UNESCO.

Cupak, A. (2013). Regionalization of catchments with use hierarchical cluster analysis methods. Teka Kom. Ochr. Kszt. Środ. Przyr. – OL PAN, 10, 5–13.

Cupak, A. (2017). Initial results of nonhierarchical cluster methods use for low flow grouping. Journal of Ecological Engineering, 18(2), 44–50.

Cupak, A., Wałega, A., Michalec, B. (2017). Cluster analysis in determination of hydrologically homogeneous regions with low flow. Acta Sci. Pol. Formatio Circumiectus 16(1), 53–63.

Demuth, S. (1989). Small research basin studies. In: Regimes from Experimental and Network Data (FRIEND), vol. I Hydrological Analysis, ed. By A. Gustard, L. Roald, S. Demuth, H. Lumadjeng and R. Gross, 141–186. Institute of Hydrology, Wallingford, UK.

Demuth, S., Young, A.E. (2004). Regionalisation procedures. In: Tallaksen L.M., van Lanen H.A.J. (Eds.), Hydrological Droughts: Processes and Estimation Methods for Streamflow and Groundwater Developments in Water Science, 48, Elsevier.

Dikbas, F. Firat, M. Koc, C.A., Gungor, M. (2013). Defining homogenous regions for streamflow processes In Turkey Rusing k – means clustering method. Civil engineering, 38, 1313–1319.

Engeland, K., Hisdal, H. (2009). A comparison of low flow estimates in ungauged catchments using regional regression and HBV-Model. Water Resour Manage 23(12), 2567–2586.

Gottschalk, L. (1985). Hydrological regionalization of Sweden. Hydrological Sciences Journal 30, 1, 65–83.

Gustard, A., Bullock, A., Dixon, J.M. (1992). Low flow estimation in the United Kingdom. Institute of Hydrology. Report No. 108.

Gustard, A., Gross, R. (1989). Low flow regime of Northern and Western Europe. In: FRIENDS in Hydrology (Proc. 1st Int. FRIEND Symp., Bolkesjø, Norway, April 1989), 205–212, IAHS Publ. no. 187.

Gustard, A., Roald, L.A., Demuth, S., Lumadjeng, H.,S., Gross, R., Arnell, R. (1989). Flow regime from Experimental Network Data, vol. 1 Hydrological Studies. Wallingford, UK: Institure Of Hdrology.

Hayes, D.C. (1992). Low flow characteristics of streams in Virginia. US Geological Survey, Water Supply Paper 2374.

Hosking, J.R.M., Wallis, J.R. (1997). Regional frequency analysis: an approach based on L-moments. Cambridge University Press, New York, USA.

Jokiel, P., Stanisławczyk, B. (2016). Zmiany i wieloletnia zmienność sezonowości przepływu wybranych rzek Polski. Prace Geograficzne 144, 9–33.

Kahya, E., Demirel, M.C. (2007). A comparison of low-flow clustering methods: streamflow grouping. Journal of Engineering and Applied Sciences 2(3), 524–530.

Kohnová, S., Hlavčová, K., Szlogay, J., Števková, A. (2009). Seasonality analysis of the occurrence of low flows in Slovakia. International Symposium on Water Management and Hydraulic Engineering, Ohrid, Macedonia 1–5 September, 711–720.

Kowalczak, P. (1986). Metoda typologii hydrograficznej niehierarchiczną analizą skupień (na przykładzie dorzecza górnej Noteci) In: Niektóre problemy metodyczne w hydrologii pod red. Z. Mikulskiego. Dokumentacja Geograficzna. Instytut Geografii i Przestrzennego Zagospodarowania, 2, PAN, 38049.

Kozek, M. (2018). Spatial variability of low flows in the upper Warta river catchment. Acta Sci. Pol. Formatio Circumiectus, 17 (3), 67–76.

Laaha, G. (2002). Modelling summer and winter droughts as a basis for estimating river low flow. In: FRIEND 2002 – Regional Hydrology. IHAS Publ. 274, 289–295.

Laaha, G. (2006). Process based regionalisation of low flows. Wien, Wiener Mitteilungen, 8–53.

Laaha, G., Blöschl, G. (2006a). A comparison of low flow regionalisation methods- catchment grouping, Journal of Hydrology 323, 193–214.

Laaha, G., Blöschl, G., (2006b). Seasonality indices for regionalizing low flows. Hydrol. Process. 20, 3851––3878.

Laaha, G., Blöschl, G. (2007). A national low flow estimation procedure for Austria. Hydrological Science Journal, 52(4), 625–644.

Lecce, S.A. (2000). Spatial variations in the timing of annual floods in the southeastern United States. Journal of Hydrology, 235, 151–169.

Lin, G.W., Chen, L.H. (2006). Identification of homogeneous regions for regional frequency analysis using self-organizing map. Journal of Hydrology, 324,1–9.

Mandal U., Cunnane, C. (2009). Low-flow prediction for unguaged river catchments in Ireland. Irish National Hydrology Seminar 2009.

Mazvimavi, D., Meijerink, A.M.J., Savenije, H.H.G., Stein, A. (2005). Prediction of flow characteristics using multiple regression and neural networks: A case study in Zimbabwe, Physics and Chemistry

Nathan, R.J. (1990). Low flow hydrology: Application of a system approach, PhD, University of Melbourne, June.

Nathan, R.J., McMahon, T.A. (1992). Estimating low flow characteristics in unguaged catchments. Water Resources Management 6, 85–100.

Parajka, J., Merz, R., Bloschl, G. (2005). A comparison of regionalisation methods for catchment model parameters. Hydrol. Earth Sys. Sci. Discuss 2, 509–542.

Pruski, F.F., Rodriguez, del R.G., Nunes, A.A., Pruski, P.L., Singh, V.P. (2015). Low-flow estimates in regions of extrapolation of the regionalization equations: a new concept. Journal of the Brazilian Association of Agricultural Engineering, 35, 5, 806–816.

Punzet, J. (1981). Empiryczny system ocen charakterystycznych przepływów rzek i potoków w karpackiej części Dorzecza Wisły. Wiadomości IMGW, 1 – 2, 31–39.

Rao, A.R., Srinivas, V.V. (2003). Some problems in regionalization of watersheds. Water Resources Systems-Water Availability and Global Change, IAHS Publ., 280, 301–308.

Rao, A.R., Srinivas, V.V. (2006a). Regionalization of watersheds by fuzzy cluster analysis. Journal of Hydrology, 31(1–4), 37–56.

Rao, A.R., Srinivas, V.V (2006b). Regionalization of watersheds by hybrid cluster analysis. Journal of Hydrology, 31(1–4), 57–79.

Rao, A.R., Srinivas, V.V. (2008). Regionalization of Watersheds. An approach based on cluster analysis, Springer.

Riggs, H.C. (1990). Estimating flow characteristics at ungauged sites. In: Regionalisation in Hydrology. Proceeding of the Ljubljana Symposium, April 1990, IAHS Publication No 191.

Schreiber, P., Demuth, S. (1997). Regionalization of low flows in southwest Germany. Hydrol. Sci. J. 42 (6), 845.

Shu, Ch., Burn, D.H. (2003). Spatial patterns of homogeneous pooling groups for flood frequency analysis. Hydrol. Sci. J. 48, 4, 601–618.

Skop, E., Loaiciga, H.A. (1998). Investigating catchment hydrology and low-flow characteristics using GIS. Nordic Hydrol. 29(2), 105–128.

Smakhtin, V.U. (2001). Low flow hydrology: a review. Journal of Hydrology 240,147–186.

Števková, A., Sabo, M., Kohnová, S. (2012). Pooling of low flow regimes using cluster and principal component analysis. Slovak Journal of Civil Engineering, XX, 2, 19–27.

Tomaszewski, E. (2018). Drought streamflow deficits assessment, applying constant and variable threshold levels, as illustrate with the example of selected catchments in the Vistula river basin. Acta Sci. Pol. Formatio Curcumiectus 17(3), 205–216.

Ustawa Prawo wodne z dnia 20 lipca 2017 roku. Dz. U. 2017 poz. 1566.

Vezza, P., Compoglio, C., Rosso, M., Viglione, A. (2010). Low Flows Regionalization in North-Western Italy, Water Resours. Manage 24, 4049–4074.

Vogel, R.M., Kroll, C.N. (1992). Regional geohydrologic-geomorphic relationships for the estimation of low flow statistics. Water Resources Research 28(2), 2451–2458.

Wałęga, A., Młyński, D. (2017). Seasonality Of Median Monthly Discharge In Selected Carpathian Rivers Of The Upper Vistula Basin. Carpathian Journal of Earth and Environmental Sciences, July 2017, 12,2, 617–628.

Wałęga, A., Młyński, D., Kokoszka, R. (2014). Weryfikacja wybranych metod empirycznych do obliczania przepływów minimalnych i średnich w zlewniach dorzecza Dunajca. Infrastruktura i Ekologia Terenów Wiejskich Nr II/3, 825–837.

Ward, Jr. J.H. (1963). Regional flood frequency analysis II. Multivariate classification of drainage basins in Britain. Hydrological Sciences Journal, 31(3), 335–346.

WMO: World Meteorological Organization. (1974). International glossary of hydrology, WMO, Geneva.

www.encyklopedia.pwn.pl.

Ziernicka-Wojtaszek, A., Kaczor, G. (2013). Wysokość i natężenie opadów atmosferycznych w Krakowie i okolicach podczas powodzi w okresie maj-czerwiec 2010. Acta Sci. Pol., Formatio Circumiectus 12(2), 143–151.

PRZEGLĄD METOD WYKORZYSTYWANYCH DO REGIONALIZACJI CHARAKTERYSTYK PRZEPŁYWU NISKIEGO W ZLEWNIACH NIEKONTROLOWANYCH

ABSTRAKT

Cel pracy

Celem artykułu jest przedstawienie metod szacowania charakterystyk przepływów niskich w zlewniach nie-kontrolowanych, przegląd aktualnej literatury tematu oraz najnowszych trendów regionalizacji przepływów niskich. W artykule skupiono się na metodach opierających się na różnorodnym podejściu, między innymi statystycznym, które mogą być zaadoptowane do warunków polskich, a także sezonowości występowania przepływu niskiego.

Materiał i metody

W pracy powołano się na różnorodne źródła literaturowe, skupiając się przede wszystkim na tych opubli-kowanych w ciągu ostatnich dwóch dekad. Scharakteryzowano najpopularniejsze podejścia regionalizacji zlewni, takie jak: metody oparte na analizie reszt, statystyki wielowymiarowe: analiza skupień, modele drzew klasyfikacyjnych i regresyjnych (CART) czy sezonowości przepływów niskich.

Wyniki i wnioski

Opisane w pracy podejście oparte na regresji regionalnej jest najczęściej stosowaną metodą do szacowania przepływów niskich w zlewniach niekontrolowanych, wykorzystujące zależność pomiędzy charakterystyką przepływu niskiego i parametrami zlewni. Regiony hydrologicznie homogeniczne można wyodrębnić wyko-rzystując w tym celu metody statystyczne. Te opisane w pracy zostały wybrane ze względu na różnice w po-dejściu metodologicznym. Dodatkowo, są one „uniwersalne" co oznacza, że mogą być stosowane w różnych regionach, krajach, w tym w Polsce; mają charakter interdyscyplinarny i mogą być stosowane do obliczeń przepływów niskich, jak również średnich lub powodziowych.

Słowa kluczowe: przepływ niski, regionalizacja, regresja, zlewnia niekontrolowana