



# Regionalization of geographical space according to selected topographic factors in reference to spatial distribution of precipitation: application of artificial neural networks in GIS

Joanna Bac-Bronowicz<sup>1</sup> · Piotr Grzempowski<sup>1</sup>

Received: 1 November 2017 / Accepted: 4 September 2018 / Published online: 12 September 2018  
© The Author(s) 2018

## Abstract

In this paper, the regionalization of geographical space according to selected topographic factors and the spatial distribution of precipitation is discussed. The model takes into account qualitative and quantitative data describing the conditions associated with the studied precipitation. In the modelling, data mining methods including data clustering methods for agglomeration and artificial neural networks for classification have been used. The reason for their use was the classification of the area due to conditions related to precipitation, the distinguishing of similar areas and the delimitation of the propagation of the phenomenon or transition zones. To realize the research aims, professional software for data management, spatial data analysis, mathematical calculations and data mining have been used. The result of the research was a model of the classes representing areas with specific conditions affecting the phenomenon, transition zones between classes and areas with conditions other than those in the surroundings of the measuring stations, which are not classified in any of the classes. Classification results indicate the boundaries of the areas in which we can model the values measured at stations, the transition zones of possible discontinuous change and areas in which the phenomenon should not be modelled due to significantly different conditions from those in the neighbourhoods of measuring stations. Unclassified areas are also potential locations for new measuring stations.

**Keywords** Artificial neural network · Regionalization of geographical space · Multi-criteria classification · Precipitation regions · Topographical factors

## Introduction

Due to global development in the research studying changes in the environment, the number of measurement stations that monitor specific global parameters has increased. The most common examples are automated observation networks monitoring the parameters of climate or air pollution such as gas, solid or liquid substances. Decisions about leaving stations where they are or building new stations are usually preceded by various analyses that consider time and space

distribution of the studied phenomenon. Developing measurement stations usually involves an optimization problem in terms of designing a minimum number of measurement stations whilst maintaining the necessity of correctly describing and modelling the phenomenon in the given area, the aim of which is to provide values for the studied phenomenon as continuous data and to determine their spatial distribution (Al-Zahrani and Husain 1998; Goovaerts 1998; Changhyoun et al. 2014; Shafiei et al. 2014).

The most common way is to interpolate values between points using statistical methods or mathematical functions. Information about conditions accompanying the modelled phenomenon, such as a natural boundary for abrupt changes in values, may not always be used as function parameters of the statistical model, and the model of the phenomenon in the form of isolines averages the abrupt values of neighbouring classes.

The requirement to standardize and integrate data according to the Infrastructure for Spatial Information in the

✉ Piotr Grzempowski  
piotr.grzempowski@pwr.edu.pl

Joanna Bac-Bronowicz  
joanna.bac-bronowicz@pwr.edu.pl

<sup>1</sup> Geodesy and Geoinformatics Unit, Faculty of Geoengineering, Mining and Geology, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

European Community (INSPIRE) directive increased access to acquired spatial data and facilitated analysis. The constant development of remote sensing methods made it possible to acquire more information about the environment and its changes more frequently. The information acquired using new geoinformation methods results in new possibilities, but it requires a new approach to developing methods that would allow frequent actualization of the previous model in such a way that the 'knowledge' acquired using the model is not removed but supplemented. The widespread availability of digital weather (states of atmosphere described by physical quantities such as temperature, precipitation, and atmospheric pressure) databases and climate (characteristic atmospheric conditions determined on the basis of at least 30 years of observations in a specific area) databases has resulted in the frequent use of meteorological information in spatial analyses that are used extensively in geographic information systems (GISs) equipped with spatial estimation procedures. Interest in the possibility of including high-resolution spatial information about precipitation in GIS systems is the result of the demand generated by the fields of science and economics. It is believed that the greatest influence of future climatic changes on society will come from the changes in the distribution and variability of precipitation (McAvaney et al. 2001; Trenberth et al. 2003; Meehl et al. 2005; Covey et al. 2014).

By analysing qualitative and quantitative attributes and their spatial analyses, we can assign thematic data from a given area to measurement stations. Assigning data is the first step on the way to building a model. It is not always possible to describe relationships between accompanying phenomena and the dependent variable using a simple mathematical model, and finding interdependences involves the process of data exploration. Finding knowledge in data is a multi-step process and allows the selection of many methods in subsequent stages of building the relationship between conditions and the studied parameter.

The aim of the created model is to indicate regions with specific outstanding conditions that influence the studied phenomenon without estimating the quantitative influence of factors on that phenomenon and to determine the degree of trust in the obtained results to indicate areas of incomplete or uncertain information. In studying natural phenomena, cartographic presentations of areas with outstanding features or outstanding values of phenomena and their scales is a natural way of reducing the complexity of the problem to indicate significant differentiation of the characteristics of the phenomena. The applied method of regionalization is consistent with general idea of the first law of geography (Tobler 1970, 1979).

Distinguishing areas with similar geographical (severe climate, natural resources), economic (earning facility, profits in agriculture production), national and demographic

(population density) features is a natural way of classifying geographic space. Geographic regionalization derives from the needs of scientific research and making various decisions, such as localization decisions. In many instances of synthetic regionalization, it is assumed that areas with similar topographic features have similar natural conditions and can cause similar reactions of environment, for example, to the change of plans and funding rules of area spatial development, causes of disasters, etc. The division into natural units became the basis for the construction of data models for various GISs and then for modelling environmental conditions, which results in models of phenomena. The examples consist of at least several hundred climate models created in the past 10 years, for instance general circulation model/global climate models (Climateprediction.net 2015) or the elaborations that verify them (Dai 2005; Hamilton and Ohfuchi 2007; Covey et al. 2014).

The objective of this paper is to develop a method for regionalizing geographical space according to selected topographic factors and the spatial distribution of precipitation.

## Methodology

Developing a model of regionalization involves several steps of data processing, which include:

- identification of topographic conditions accompanying precipitation and building a spatial data model,
- grouping of measurement stations with taking into account average precipitations,
- develop a dependency model based on the relationship between the characteristics of the terrain and the class of precipitation.

## Topographic conditions to separate regions

Methods for the determination of topographic indicators, value evaluation and classification in the informative environment are important parts of geographical analyses. The notion of attributes (indicators) of topographic factors has been used in the literature for a long time in relation to the group of features characterizing terrain topography, which mainly includes terrain, durability, permeability, slope, height, exposition and land cover (Cressie 1991; Li et al. 2005; Zhang et al. 2013). Some of the most commonly used topographic determinants are geographic location; average absolute and relative height; maximum, minimum and average declines; curvature profiles; paths of waterfall on slopes (Manoj et al. 2004; Mutasem et al. 2013); distance from rivers, roads and built-up areas and the course of morphological barriers. There have been many attempts

at the objective classification of areas using topographical indicators in modelling many natural phenomena as model parameters.

It was assumed that the relationship between precipitation distribution and topography has been confirmed based on many years of research conducted by climatologists and on consultations with them. Many studies confirm the existence of the relationship between average precipitation values and topographic conditions. Many authors describe local impacts of various topographic factors on the spatial distribution of precipitation, but the complex relationship between topography and precipitation distribution has not been completely elucidated. Basista and Bell (1994), pointing to much earlier studies and to their own, claim that interactions between precipitation and exposition, slope and exposure together and separately on both regional and continental scales are unknown; however, local interactions exist and have been verified. Verified local relationships between topographic conditions in the neighbourhood of meteorological stations and daily, monthly and annual average precipitation are discussed in many studies. The above-mentioned work (Basista and Bell 1994) is quoted in most articles that consider precipitation dependency on terrain.

Studies on the relationship between precipitation and topography conducted in Japan and of so-called PT (Precipitation Terrain) combination in single mountains (Yoshiharu et al. 2007) showed that a PT combination designated as a ‘Gaussian functional relationship’ (GRIM) can be found on the windward side of a mountain, and more precipitation is generated on the windward side when the wind speed is less.

It was concluded that an increase in precipitation became stronger towards the leeward side together with an increase in wind speed. Torrential rainfall in southwestern Germany and eastern France (Kunz and Kottmeier 2006) pointed to the existence of several mechanisms that determine the spatial distribution of precipitation over complex terrain, such as the influence of a different variation of the model on the results of a precipitation distribution model. They confirmed the most significant relationship between increased intensity of rainfall above and below the mountain peaks. In the literature appearing since the beginning of the 20th century, we see many references confirming the dependence of rainfall and topography. The influence of topography is clearly visible in hilly and mountainous regions. The above relationships were analysed using statistical methods mainly used for quantitative data. It has been confirmed many times that qualitative data about land cover, physiographic units or watersheds are useful for indicating or distinguishing precipitation regions (Prudhomme and Reed 1999; Bac-Bronowicz 2005, 2007; Gouvas et al. 2009; Myoung-Jin et al. 2010).

## Spatial data model

The data model was built using a GIS and contains spatial and descriptive data characterizing the modelled phenomenon and the accompanying factors and conditions. Spatial data are stored in the vector and raster models, which together with descriptive data are the inputs for the construction of the final models. Precipitation stations and precipitation averages form a layer (object class) with point-type geometry. Accompanying conditions were recorded in the form of vector layers with area-type geometry in the natural limits of their range. Thematic layers contain quantitative and qualitative data.

For the construction of a classification model, the integration of data from thematic layers (feature classes) in one class of objects with a fixed reference unit is crucial. In this context, one of the systems of basic units called TEMKART (Podlacha 1986) constructed from fields of approximately 1 km<sup>2</sup> (between 0.981 and 1.022 km<sup>2</sup>) was used. The size of the basic field used to collect data in a thematic database is consistent with the dimensions used in international studies for Europe (Roekaerts 2002; Noirfalise 2007; Metzger et al. 2005; ETC/BD 2006). Various combinations of the sizes of basic fields and their applications are presented in the table for climatic networks and data sets in The Tyndall Centre for Climate Change Research (Mitchell et al. 2004). The subject has also been analysed often in Poland for the distribution of climate parameters (Stach 2010). In the results of attribute and spatial analyses, data from different thematic layers were integrated in the elementary fields of the raster model grid. The data characterizing precipitation and related conditions include the following thematic layers:

- measurement stations,
- altitude - DTM,
- declines of the terrain,
- directions of terrain slope (exhibition),
- land cover (land use),
- physiogeographic unit,
- river catchment areas.

Quantitative data applied in the classification model are mainly data that can be measured and presented in the form of natural or actual numbers, characterizing topographic conditions frequently discussed in the literature in modelling the spatial distribution of precipitation (“**Topographic conditions to separate regions**”). Qualitative data entered into the classification model represents the diversity of the characteristics of the test area because of the landform and land use, which also affects the diversity of the precipitation phenomenon. Introducing information about the catchment belonging to the model assigns information about the elementary fields belonging to the zone. The boundaries of the

catchment run mostly along the ridges or lines of changes of the directions of slopes and cover different forms of terrain at the same time, indicating the area of common catchment. Catchments are natural dividing zones that can be used to locate phenomena instead of the coordinate system. The precipitation station coordinates and the elementary fields have not been entered into the classification model in order not to create a close relationship between the position and the value of precipitation. Types of data and geometric types of thematic layers are shown in Table 1.

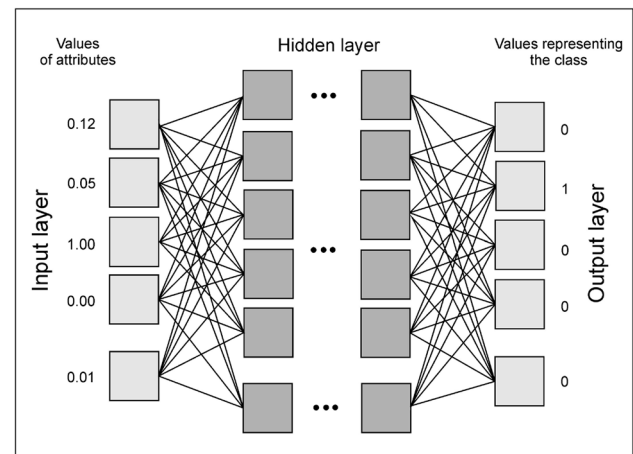
## Grouping stations

The stage that precedes the creation of the classification model of the area is the analysis of average precipitation at precipitation stations. Measuring stations are grouped into classes on the basis of similarity between the average values within four time periods. To divide stations into groups, hierarchical grouping was used, in which a tree structure is formed by recursive combining of existing groups. Agglomeration methods assume that each observation is a separate group, and in subsequent iterations, two groups (focus) that are closest to each other are combined into a new group (Sreejesh et al. 2014; Borcard et al. 2011). The Euclidean distance was accepted as the measure of similarity, while the distance between the groups was determined by average combination. In this method, a predetermined number of classes does not exist. The division into groups is made based on the analysis of the chart showing the distance between individual objects and groups. The results obtained by that method are used to make decisions about the division. The final boundaries of classes are based on a classification tree, analysis of the average values of the characteristics in the group and expertise.

## Applying artificial neural networks to the classification

To conduct classification, multilayer, non-linear feedforward artificial neural networks were used (Bishop 1995). The created model is the model of assigning features that describe areas to predicted classes of precipitation. The idea of using artificial neural networks is presented in Fig. 1.

The input layer is formed by neurons, and figures representing quantitative and qualitative attributes are entered onto this layer. The output layer consists of neurons that represent the classes of precipitation. The number of input neurons depends on the number of numerical attributes (1 feature = 1 neuron) and the number of qualitative attributes, as well as the number of bits necessary to record a feature combination in the form of a binary number (1 feature = the number of cells needed to record a feature in the zero-one



**Fig. 1** The idea of using artificial neural networks to classify areas according to chosen conditions accompanying precipitation

**Table 1** List of thematic layers, data format and geometrical types

Layers	Data format	Geometry	Resolution	Data type
Measurement stations	Vector	Point	-	Quantitative
NMT	GRID	Elementary field in the form of square	1 × 1 km	Quantitative
Terrain slopes	GRID	Elementary field in the form of square	1 × 1 km	Quantitative
Directions of terrain slopes	GRID	Elementary field in the form of square	1 × 1 km	Qualitative
Land cover	Vector	Area	-	Qualitative
Physio-geographic units	Vector	Area	-	Qualitative
River catchments	Vector	Area	-	Qualitative
Integrative layer TEMKART	GRID	Elementary field in the form of square	1 × 1 km	Quantitative and qualitative



format =  $n$  neurons). In the hidden layer, the hyperbolic tangent activation function was used (Bridle 1989; Vogl et al. 1988):

$$O_j = F(I_j) = \frac{1}{1 + e^{-I_j}} \quad (1)$$

On the output layer, the softmax activation function was used:

$$O_j = F(I_j) = \frac{e^{I_j}}{\sum_k e^{I_k}} \quad (2)$$

where:

$I_j$  - input value,  
 $O_j$  - output value.

Each neuron represents one class. The output neuron having the highest value indicates the class. The softmax function enables us to determine the probability that a recognized object belongs to the class (Bridle 1989). The value of function is normalized in such a way that the sum of all values of output neurons is 1. Normalized values can be interpreted as the probability of belonging to the class.

The number of output neurons have to match the number of separated groups of precipitation stations. Each output neuron is assigned to the specific group of stations. Process of artificial neural network learning requires sets of data with known patterns that contain factors which are entered onto the input layers and known response of the network entered onto the output layer. The value "1" is entered onto neuron representing given group of stations and "0" onto other neurons.

In the learning process, the weights of connections between input neurons that characterize conditions around the station and output neurons representing precipitation classes are determined. After the learning process, the neural network classifies elementary fields outside the surroundings of precipitation measurement stations, indicating similarity to the conditions existing in selected precipitation classes. The training set has been randomly but proportionally divided into three subsets for each class: the training set (70%), the validation set (15%) and the testing set (15%). The training set is involved in the process of changing weights of connections in subsequent iterations, the validation set is involved in the process of weight changes and in testing the correctness of solutions and the testing set is used only to evaluate the correctness of recognition. The iteration process ends when the minimal error is obtained and at the moment of obtaining convergence of charts of listed teaching sets. It should be emphasized that in the learning process and recognition, the same number of features describing elementary fields must be entered

onto the input layer. In the learning process, we know the input attributes of elementary fields and the network answers (belonging to precipitation class), while during recognition only the input attributes take place. The number of features and the way of data scaling must be equal in the teaching and recognition processes.

## Specification of test area

The selected test area for the construction of a classification model is in southwestern Poland. It is characterized by varied topography, including lowlands, uplands and mountains. The terrain is closely related to the geological and tectonic construction. To the north, the area is surrounded by moraine hills that are about 300 m above sea level. To the south of the hills is a wide valley of the Oder River running northwest to southeast, where the terrain falls to approximately 115 m above sea level (Badura et al. 2004). Further south, the terrain rises gently to the Sudetic Marginal Fault, which is the boundary between the Fore-Sudetic Block and the Sudetes Mountains. The fault creates a clear morphologic edge showing height differences of approximately 250 m. The Sudetes Mountains also run northwest to southeast. The average height of the mountains ranges from 600 to 1400 m, with peaks up to 1602 m. The terrain directly affects agroclimate conditions. Due to the parallel arrangement of physical-geographical units in Lower Silesia (dominant course) the air masses flowing from the Atlantic Ocean to northeastern Europe and Scandinavia affect its climate and weather conditions. Much less frequently there are masses of warm air from the Azores or from the direction of the Mediterranean. The dominant masses cause mild winters and relatively cool summer months in that part of Europe (Dubicki 2002; Soczyńska et al. 1997). A special feature is frequent but short periods of cloudy weather with rainfall. In winter, mild southwestern winds are characteristic, while in summer there are northwest winds that lower the temperatures in Lower Silesia. Long-term highs are common weather conditions. They are also connected with subtropical air and are characterized by long periods of beautiful sunny weather that is often windless. Most often, highs appear in May and September. The coexistence of marine and continental climate features, as well as the occasional influx of Arctic air and tropical masses, results in great volatility of the weather throughout the year. The atmospheric pressure decreases with altitude and transparency of the atmosphere increases, temperatures decrease and precipitation increases according to relief and exposition. Cold patches and local wind systems are formed.

Data about atmospheric precipitation at measurement points are taken from *The Atlas of Precipitation in Poland 1891–1930* (Wiszniewski 1953) and from yearbooks

published by the Institute of Meteorology and Water Management. Average sums of precipitation for May and June, July, July and August and for the vegetation period were entered into the database. Listed periods are particularly important for agriculture because of the needs of specific crops (Bac-Bronowicz 2010). Values were introduced for:

- 460 stations from 1881–1930,
- 250 stations from 1948–1980,
- 45 stations from 1981–2013.

The height of the terrain, slope and exposure were calculated as the average of 16 samples in a square with sides of 1 km based on the DTM, with an accuracy corresponding to the scale of 1:50,000. Information about land cover was obtained from the database created under the CORINA Land Cover program (CLC 2006). Elementary fields belonging to the physio-geographic units were determined based on Kondracki and Walczak's elaboration, with boundaries deriving from the analysis of the shape and morphology of the terrain (Pawlak et al. 2008; Kondracki 2000). Using this type of unit allows morphological barriers, which are the main factors in the distribution of climate characteristics, to be distinguished. Using the regular basic fields, a regular continuous model for the whole studied area was obtained. Adopting those units as basic ones in climate modelling boosts the credibility of the results because they are frequently tested material. Detailed information about the course of the catchment borders was obtained from the HYDRO database (HYDRO base and maps; hydrography on a scale of 1:50,000) at the Geodesy and Cartography Documentation Centre in Wrocław.

## Results

### Grouping precipitation stations into classes according to average values of precipitation

Two variants of multi-criteria grouping were prepared. They differ in terms of the set of attributes describing the precipitation stations. In the first variant, the following features were accepted:

- average precipitation during the growing season from April to September,
- average precipitation between May and June,
- average precipitation between July and August and
- average precipitation in July.

The average precipitation from many years of observations was analysed, which eliminates the influence of randomness.

In the second variant, the height of terrain in the location of precipitation stations was added. Based on obtained charts and on expertise, the following groups within individual variants were separated (Tables 2, 3).

The first option provided only division with regard to the sum of average precipitation from listed periods, but in Table 2 for the variant I also the altitude ranges in each class are presented.

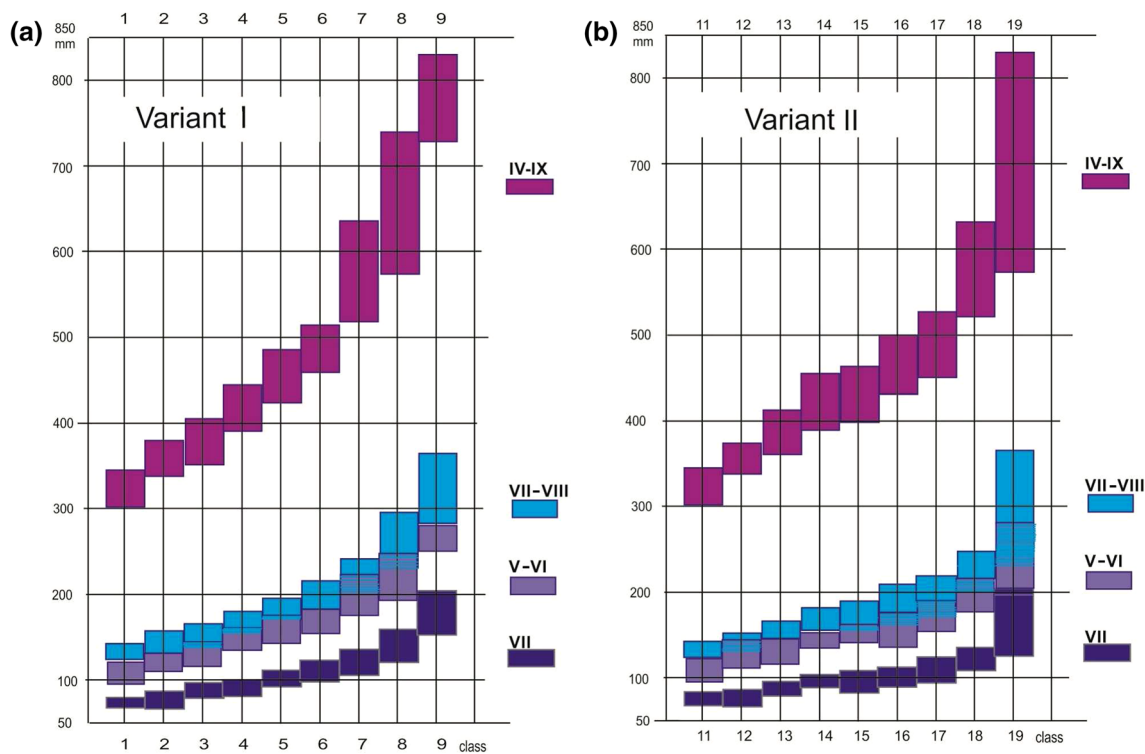
Adding the altitude in the second variant diversified groups with similar values of precipitation but with different altitude conditions. The intervals of average precipitation in individual classes are also presented graphically in Fig. 2a. Because of multi-criteria division, which takes into account several features at the same time, intervals of individual classes (values of average precipitation sums) overlap partially.

**Table 2** The average values of precipitation in separate groups—variant I

No.	IV–IX (mm)		V–VI (mm)		VII–VIII (mm)		VII (mm)		H a.s.l. (m)		Number of stations
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	
GI-1	303	345	95	122	124	142	68	80	28	146	44
GI-2	338	380	111	132	132	158	67	88	45	195	78
GI-3	352	405	117	144	139	167	79	95	62	373	42
GI-4	390	445	136	161	156	180	82	100	82	604	79
GI-5	426	485	145	175	171	196	93	112	172	670	60
GI-6	458	516	156	183	183	216	97	124	190	780	52
GI-7	519	634	177	223	200	242	107	135	296	800	33
GI-8	575	741	192	247	231	294	121	160	280	1603	29
GI-9	728	830	251	280	282	365	155	204	416	1490	10

**Table 3** The average values of precipitation in separate groups—variant II

No.	IV–IX (mm)		V–VI (mm)		VII–VIII (mm)		VII (mm)		H a.s.l. (m)		Number of stations
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	
GII-1	303	346	95	122	124	143	68	84	28	146	47
GII-2	338	373	113	144	132	152	67	87	45	186	65
GII-3	359	413	117	146	147	167	79	96	85	240	72
GII-4	389	455	135	154	157	182	89	104	335	470	18
GII-5	399	461	141	162	158	189	82	109	172	320	67
GII-6	430	500	136	175	161	209	89	112	380	780	18
GII-7	451	528	154	189	171	218	93	124	190	565	81
GII-8	522	631	177	215	200	248	109	136	280	800	37
GII-9	575	830	196	280	231	365	128	204	300	1603	31



**Fig. 2** The chart of average precipitation at measurement stations; the division into groups **a** variant I, **b** variant II

Separated groups were organized according to the increase in minimum values. In the first version, together with the increase in minimum sums of average precipitation, the minimum altitude in groups increases. At the same time, the difference between extreme values of altitude intervals also usually increases (Fig. 2a). In the second version, because of introducing the attribute describing altitude, the precipitation stations were separated into stations where the values of average precipitation are similar but differ when it comes to altitude (Fig. 2b). Examples are groups GII-4 and GII-5, which have a similar range of average precipitation

but form separate groups due to differences in height. In group GII-9, there are stations with high amounts of average precipitation that are so large that they dominated the altitude differences.

The division in both versions refers to the terrain correlated with the geology of the area. Groups GI-1 - GI-3 and GII-1 - GII-3 appear mainly in lowland areas with little differentiation of altitude. Groups GI-4 and GI-5 and GII-4 and GII-5 appear mainly on the borders between upland and lowland areas and mountainous terrain. Groups GI-6, GI-7, GI-8 and GI-9 and GII-6, GII-7, GII-8 and GII-9 are

generally in mountainous areas where there is great diversity of both precipitation and altitude. In the set of those groups, groups GI-8, GI-9 and GII-9 should be especially distinguished as classes with a significant minimum value of average precipitation and with a great span in the interval of their values. The classes separated in both variants should not be compared directly. The first division aimed at separating groups based on precipitation, and the second one separated groups with similar values of precipitation but with different altitudes. These divisions prove that precipitation does not depend directly on altitude but on the shape of the terrain around the station. The examples are stations located at tectonic faults with clear morphological edges that have similar values of precipitation to stations located within a few kilometres on the wing of a raised fault.

### **Applying artificial neural networks to the classification of areas according to chosen conditions accompanying precipitation**

The structure of network consists of an input layer containing 66 neurons, a hidden layer containing 66 neurons and an output layer containing 9 neurons. Different structures of networks, activation functions and algorithms used in such tasks were tested (Bishop 1995; Bridle 1989). The choice of activation function on the output layer and the learning algorithm were made considering the criterion of the best adjustment to data measured by the correctness of recognition of patterns after the learning process. Calculations were made in the MATLAB program using the Neural Network Toolbox 8.2 packet. A scaled conjugate algorithm for fast supervised learning (Moller 1993) was used to learn artificial neural networks. The learning process consists of several thousand examples where both the previously listed qualitative and quantitative attributes of elementary fields and the precipitation class the closest to the precipitation station are known. The patterns have been chosen according to the following criteria:

- an elementary field cannot be further than 3.5 km from a measurement station,
- an elementary field should be located in the same physiographic unit as the precipitation measurement station,
- an elementary field belongs to the same catchment as the precipitation measurement station.

In the process of artificial neural network learning a large set of patterns containing data for the input and output layers is needed. Because the number of measurement stations is relatively small (max. 400), similar conditions in a neighbourhood of the precipitation station are assumed to increase the number of patterns. Each next basic field near the station is treated as a pattern of a given group of stations. Radius of the neighbourhood does not come directly from the mathematical dependencies but from the knowledge and

experience gained from the researched area. The distance of 3.5 kilometers selected in the research area is a result of the spatial distribution of the measurement stations. The minimum distance between stations is 7 km. During establishment of the neighborhoods it was assumed that they should not overlap. Lengthening distance from the station increases the probability of significant changes of the local factors adopted for the model building. The neighborhoods can have different radius for each station. On the plains, where there are smaller variations in relief, the radius may be larger, whereas in the mountains, the radius should be smaller. The basic principle of the pattern creation is the choice of representative neighborhoods for a given group of precipitation measurement stations. Choosing a neighborhood from a different physiographic unit than the location of a station may introduce errors in the process of artificial neural network learning and cause incorrect classification of data. Each catchment area is characterized by specific local conditions that have influence on the precipitation. The catchment is a potential precipitation region, so the choice of neighborhood from different catchment can have the same effect as choosing the incorrect physiographic unit.

The surroundings of stations (elementary fields) that are representative for the *i*-th separated class have been chosen. Patterns were indicated using existing knowledge of precipitation in that part of Poland. Patterns should be chosen carefully so that they can represent conditions from the precipitation class around the station. At the same time, they should include possible differentiation within that class. The mentioned criteria assume that the studied phenomenon is not isotropic and that it changes depending on topographic factors that usually change directionally. The created model of the spatial distribution of the phenomenon is not an isotropic model (changeability of distribution depending on direction in assumed coordinate system) in its classical meaning because the dependency model does not include coordinates of the elementary fields' localization. Only the features mentioned in previous chapters are entered into the input layer.

Figure 3 presents an example of indicating elementary fields in the surroundings of a precipitation measurement station, taking into account the distance from the station and localization in the same physiographic unit. Indicating several elementary fields in the stations surroundings introduces to the learning process possible differentiation of numerical attributes values for a given precipitation class.

The chosen examples of classification models are presented below.

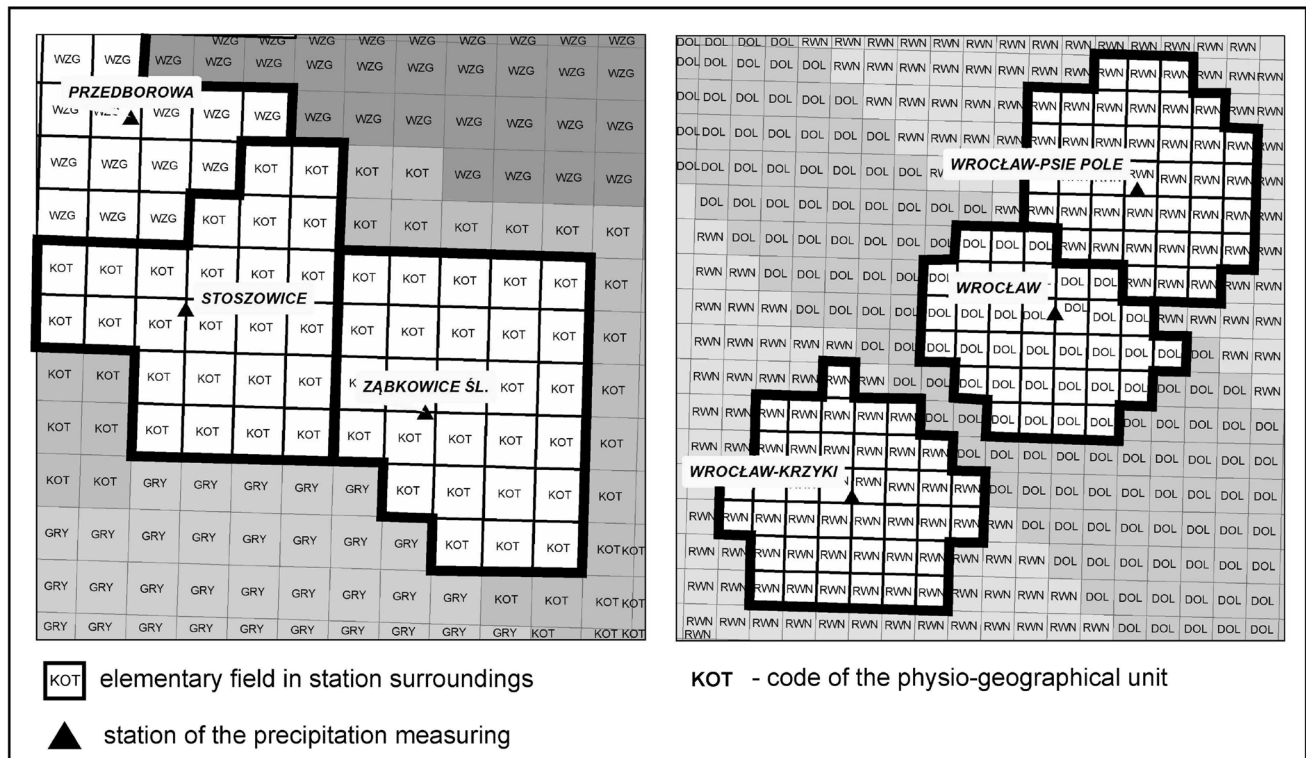


Fig. 3 The example of the choice of elementary fields in the surroundings of a precipitation measurement station for learning process

Table 4 The number of elementary fields included in the process of the neural network training—option I

Class	Number of patterns	Percentage (%)	Number of stations in the class	Percentage (%)
GI-1	306	5.1	9	3.8
GI-2	1374	22.9	40	16.7
GI-3	795	13.2	24	10.0
GI-4	1432	23.9	55	22.9
GI-5	663	11.1	36	15.0
GI-6	586	9.8	28	11.6
GI-7	475	7.9	25	10.4
GI-8	285	4.8	17	7.1
GI-9	79	1.3	6	2.5
	$\Sigma = 5995$	$\Sigma = 100\%$	$\Sigma = 240$	$\Sigma = 100\%$

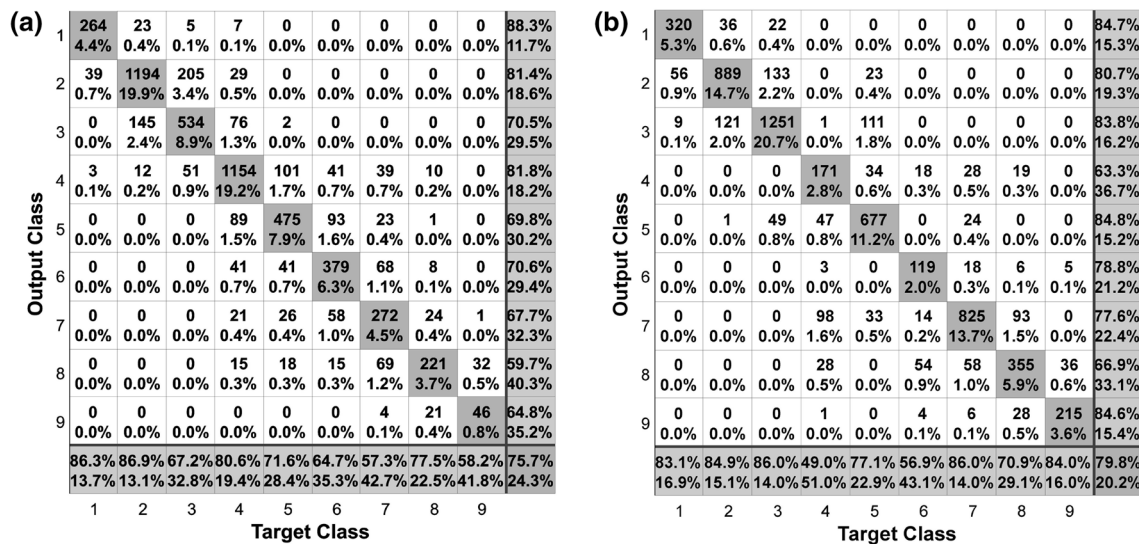
**Option I—without taking altitude into account during multi-criteria classification of patterns**

The number of patterns used to learn artificial neural networks for individual classes is presented in Table 4. Their number in individual classes is proportional to the number of stations in the precipitation class. To learn networks, the training set including 5995 model fields with 9 assigned model classes was used. The greatest number has patterns from classes GI-2 and GI-4 and result from the number of measurement stations (respectively 16.7% and 22.9% in

total) and the area ranges of that concentration. The smallest number of patterns is from class GI-9, which includes the surroundings of stations located in high mountains. The spatial extent of the land forms in the studied area is relatively small in relation to other ones; therefore, the number of measurement stations is small and comprises just 2.5% of the total number of stations in the studied area.

The horizontal column labels in Fig. 4 provide numbers of classes (target class) indicated in the training set, while in the columns there are numbers of correctly (diagonal) and incorrectly (outside the diagonal) classified cases (output





**Fig. 4** The results of classification of the training set elements after the learning process (target class—class indicated in training set, output class—class recognized by the neural network after learning process): **a** option I, **b** option II

class) after the learning process. After the learning process, the neural network can include patterns in other precipitation classes than that indicated at the beginning because it generalizes the acquired knowledge. Based on chart in Fig. 4a, we can predict the migration of patterns between classes and indicate neighbourhoods and similarity of classes. For example, in class GI-6, the network recognized a few cases from classes GI-4, GI-5, GI-7 and GI-8, which should be interpreted as spatial localization of class GI-6 near the mentioned classes and the possibility of transition zones occurring between them. In the process of learning network, the compatibility between target classes and output classes at the level of 76% was obtained. Incorrect recognition of patterns after learning was 24%. However, the result should not be considered negative. Usually during recognition, the network mistakes one neighbouring class (meaning spatial neighbourhood of localization of the class) in relation to

the expected value. The example is class GI-4, where elementary fields are included in neighbouring classes GI-3 and GI-5 (correspondingly 1.3% and 1.5%). It should be emphasized that the model does not include localization of elementary fields but only of the features describing them and that accompany precipitation. Spatial differentiation is characterized only by belonging to a catchment. Classification to not spatially neighbouring classes is just 0.1%. The result confirms the assumption that the borders of the climatic phenomena are not clear and that transitional zones between classes can appear (alternative assessment).

**Option II—taking into account altitude during multi-criteria classification of patterns**

In option II, we create the model of classification of elementary fields indicating weights of connections between the

**Table 5** The number of elementary fields included in the process of the neural network training—option II

Class	Number of patterns	Percentage (%)	Number of stations in the class	Percentage (%)
GII-1	385	6.4	11	4.6
GII-2	1047	17.3	31	12.9
GII-3	1455	24.1	43	17.9
GII-4	349	5.8	16	6.7
GII-5	878	14.5	33	13.8
GII-6	209	3.5	10	4.2
GII-7	959	15.9	49	20.4
GII-8	501	8.3	27	11.2
GII-9	256	4.2	20	8.3
	Σ=6039	Σ=100%	Σ=240	Σ=100%

conditions in the surroundings of measurement stations and precipitation classes grouped according to the altitude in their localization. The number of patterns selected in individual classes is presented in Table 5. There were 6039 patterns involved in the learning process. Taking into account altitude in the localization of the measurement station during the process of grouping caused us to include measurement stations in various classes with similar precipitation values or to include the same group of measurement stations according to similar altitude but greater differences between precipitation values.

In each option, the choice of patterns took place automatically at first, according to selected criteria, and then it was verified by expertise. In some cases, patterns were added or removed from the training set, which is why there are small differences in the number of patterns in options I and II.

The greatest number of patterns is in classes GII-3 and GII-7, which is the result of the number of precipitation measurement stations (respectively, 17.9% and 20.4% of the whole number) and the spatial extension of that concentration. The smallest number of patterns is in class GII-6, which includes the surroundings of stations located in mountain valleys.

The degree of correct answers provided by the network is similar to the one obtained in option I and is 79.8%. The smallest recognition took place in classes GII-4 and GII-6, and some elementary fields from those classes were recognized as belonging to GII-5 and GII-7 (instead of GII-4) and GII-7 (instead of GII-6).

### Classification of elementary fields

After the process of teaching, the artificial neural network specified weights of connections between neurons. To classify elementary fields, the sets of data that characterize those fields (in the form of sets and formats determined at the

beginning of the learning process) are entered onto the input layer. The selection of classes is made by selecting the neuron (corresponding to the given class, determined during the phase of creating the structure of the network and during the learning process) with the highest value of activation.

At the same time, due to the use of the softmax activation function, we obtain the probability of accepting (choice) the class. The result is not always unequivocal if the probabilities obtained on the output layer are below the assumed threshold or a few neurons obtain similar values. To summarize the results, the rules of classification and division into certain, uncertain or alternative results should be elaborated.

In practice, the activation values of at most three neurons with the highest value obtained on the output layer were analysed (main class and two alternative classes). To determine classes, the maximum value  $p_1$  with respect to the thresholds and the next two highest values  $p_2$  and  $p_3$  received on the output neurons (in respect to one another) are analysed. Calculated relationships present relative differentiation of values of belonging to the classes. Indicating alternative class shows that in the given area, there are features characteristic of other precipitation classes equal to the probability of belonging to another class. The lack of selecting an alternative class means that there is no dominating features characteristic of the defined classes. The rules used to visualize the results of classification are presented in Table 6. For each  $i$ -th set of data that characterize the conditions in the basic field, the calculated activation function values on nine output neurons are received.

The neuron with the highest activation function value indicates the class in which the elementary field should be included. Because of probable, theoretical range of output values between 0 and 1, which corresponds to the probability of inclusion in the class, the thresholds that allow us to categorize results as certain or uncertain or to indicate an alternative class have been determined.

**Table 6** Decision table—summary of the classification results

Criterion I	Criterion II	Criterion III	Classification characteristics	Description on the map
$p_1 \geq 0.7$	–	–	Choice of one class.Certain selection of the base class	Dependable
$0.5 \leq p_1 < 0.7$	$\frac{p_1}{p_2} \geq 2.5$	–	Choice of one class.Certain selection of the base class	Dependable
$0.5 \leq p_1 < 0.7$	$\frac{p_1}{p_2} < 2.5$	$\frac{p_2}{p_3} \geq 2.5$	Choice of less certain base class and the choice of alternative class. The probability of belonging to the other class determined by $p_2$ is relatively high	Dispensable
$0.5 \leq p_1 < 0.7$	$\frac{p_1}{p_2} < 2.5$	$\frac{p_2}{p_3} < 2.5$	Choice of less certain base class lack of alternative class.The probability of belonging to the other class determined by $p_2$ and $p_3$ is relatively high. Lack of belonging to one alternative class	Less reliable
$p_1 < 0.5$	–	–	Lack of possibility to choose base class. Uncertain belonging	Unreliable

$p_1$  the highest value of activation for output neuron for  $i$ -th set of input data for the class assigned to neuron

$p_2$  value of activation for the next output neuron less than  $p_1$  for  $i$ -th set of input data

$p_3$  value of activation for next output neuron less than  $p_2$  for  $i$ -th set of input data

It was assumed that:

- If one neuron on the output layer gives  $p_1 \geq 0.7$ , option I is fulfilled and the case is considered as certainly belonging to the class represented by the neuron
- If one neuron on the output layer gives  $p_1 < 0.7$  and  $p_1 \geq 0.5$ , then option II is fulfilled. If the activation value of the next output neuron is 2.5 times less than  $p_1$ , then the case is also considered as certainly belonging to the class represented by the neuron.
- If one neuron on the output layer gives  $p_1 < 0.7$  and  $p_1 \geq 0.5$  and a  $\frac{p_2}{p_3} < 2.5$  appears, the alternative class 1 can be indicated if the condition III,  $\frac{p_2}{p_3} \geq 2.5$ , is met. In practice, it means that the probability of indicating the correct class is more than or equal to 0.5. However, the second neuron also obtains a relatively high value less than 0.5, while the third one, when it comes to value, which represents the class, obtains a significantly different probability from the second neuron; thus, the alternative class 1 can be distinguished.
- If the relationship  $\frac{p_2}{p_3} < 2.5$ , then the differences between output values are so small that it is difficult to indicate one alternative class.

The expected value at the output layer corresponding to the recognized class should be close to 1. However, in practice, values of 0.9–1 are obtained for a small number of

cases. Due to the possible similarity of the set of classified features to other classes and the activation of other neurons, obtaining a probability of belonging to a class at level above 0.9 is difficult to achieve. The range of the correct classification determined experimentally is from 0.6 to 0.8. The accepted threshold of 0.7 is the center of the interval. For the assumed threshold value of  $p_1 = 0.7$ , the highest achieved value of  $p_2$  was 0.28, so the  $p_1$  value was approximately 2.5 times higher than the  $p_2$  value. Assuming this proportion with a decreased  $p_1$  value to 0.5, it was assumed that the  $p_2$  value should be higher than the  $\frac{p_1}{2.5}$  value to be able to indicate the alternative class.

The analysed area includes 20525 basic fields (the area of each is 1 km<sup>2</sup>) and the fields that were the examples in the learning process. Recognition was conducted using two models of neural networks with determined (during the learning process) weights of connections for options I and II. The examples of graphic interpretation of received results of classification, taking into account criteria from Table 6 divided into certain, uncertain results and probable indication of alternative class, are presented in Figs. 5, 6.

## Discussion

The application of the proposed method for classification model construction makes it possible to indicate areas with features similar to indicated patterns with probability,

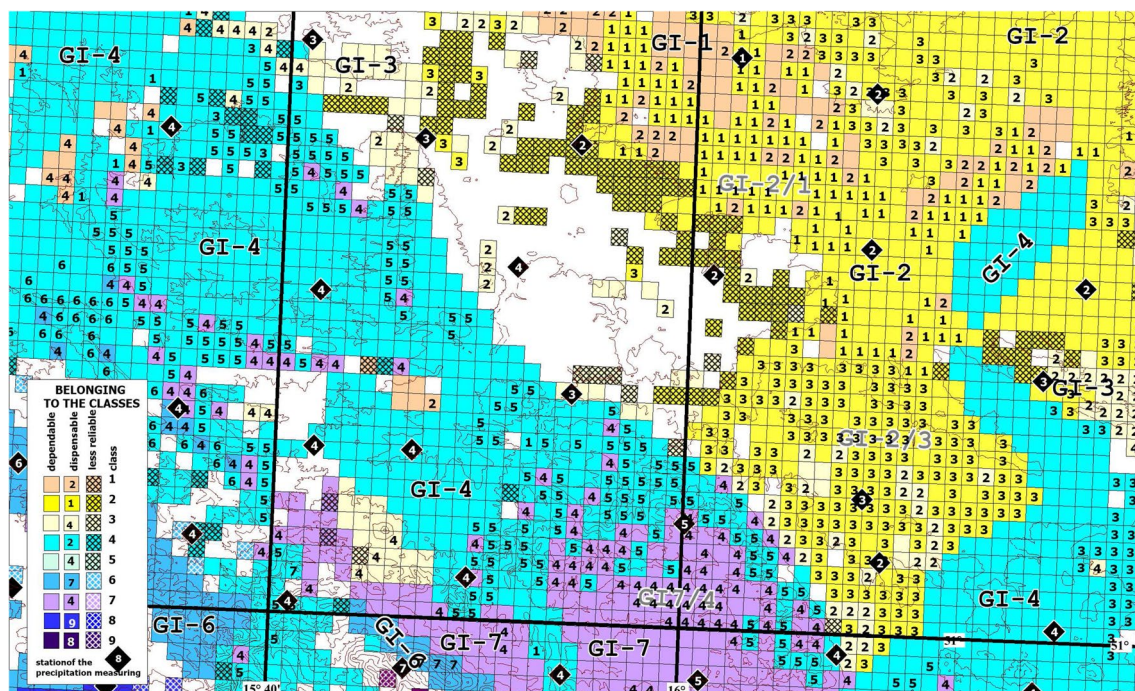


Fig. 5 A fragment of the upland area with graphic presentation of classification results using artificial neural networks from option I



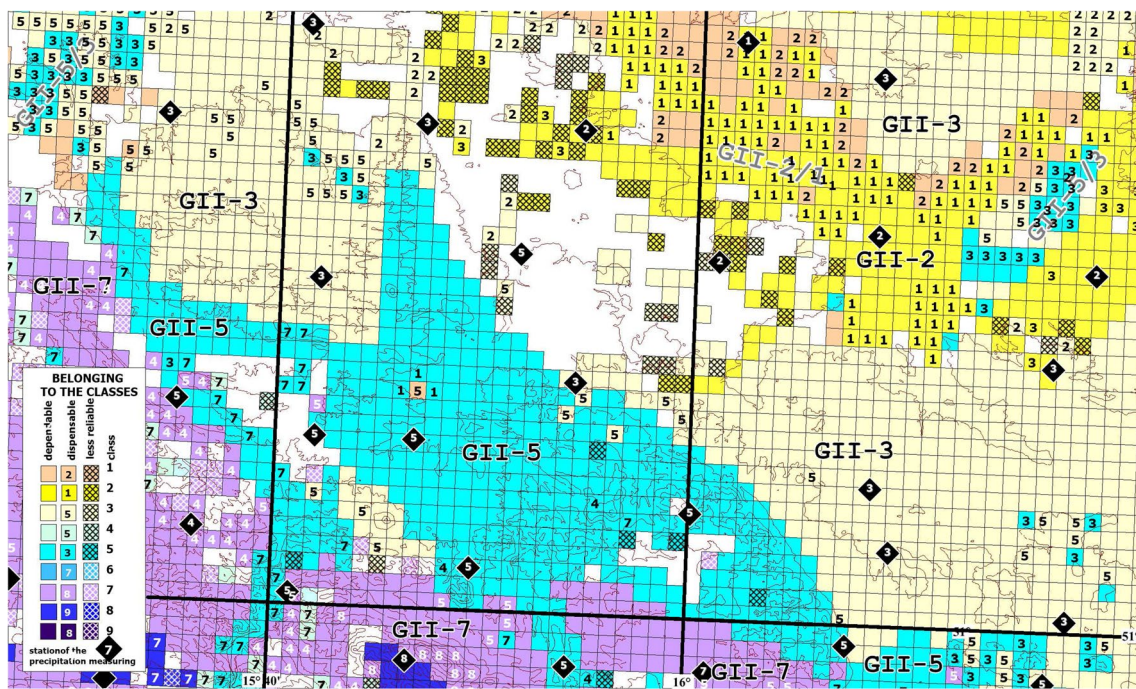
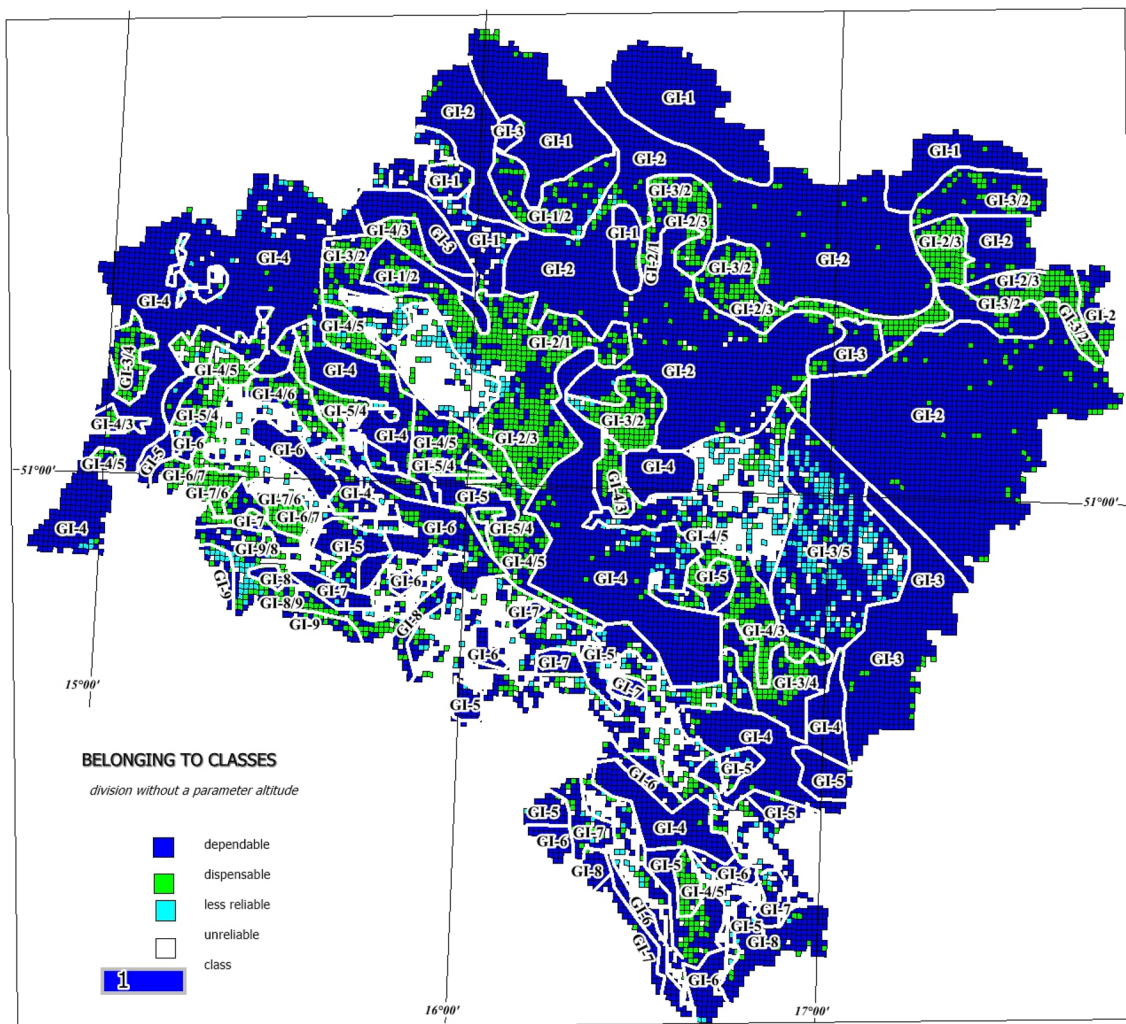


Fig. 6 A fragment of the upland area with graphic presentation of classification results using artificial neural networks from option II

which allows inclusion in the class or the indication of areas for which we cannot indicate predominant (certain) class belonging. Elementary fields that belong to a certain class should be interpreted as areas similar to the surroundings of measurement stations in the class. There is high probability of the occurrence (in elementary fields) of precipitation values similar to those at measurement stations. Elementary fields for which the probability of belonging to the class is higher than the accepted threshold are the areas where the interpolation models of the phenomenon can be used. Belonging to a less certain class should be interpreted as the possibility of indicating one class that exceeds the threshold of probability of 0.5, but the probability of belonging to the alternative class is relatively high ( $p_1/p_2$ ) or the probability of belonging to other classes is relatively high ( $p_1/p_2$  and  $p_2/p_3$ ). Thus, there are two classes, each of which does not have a high enough probability of becoming the alternative class. There are usually two or three classes in the border zone, where one of them always prevails and there is a similarity to one or two classes. Lack of class recognition on the probability level less than 0.5 should be interpreted as lack of dominant similarity to the conditions that exist in one of the separated pattern classes. In practice, there is usually a transition/border zone of several zones of studied phenomenon or a lack of measurement stations in places with a defined set of features. Unrecognized zones are potential areas to locate new measurement stations to measure the studied phenomenon. Including the height

of the measurement station in the division into classes changes both the classification of measurement stations and the classification of elementary fields. Separate classes in options I and II should not be compared directly. Division into two options proves the correctness of accepting parameters connected with topography as factors supporting the explanation of spatial distribution of precipitation for further studies. Using knowledge of the neural network according to option I, we mainly obtain separate areas because of the studied phenomenon. In the second option, in the step of stations grouping, an additional differentiation because of height is visible. In the classification of elementary fields, there is a noticeable division of one class from option I into two separate but similar classes in terms of precipitation values but different in terms of terrain. In other case, the amalgamation of classes with little variation in precipitation into one common class with similar terrain was observed. The created model of area classification allows us to identify areas similar to the surroundings of measurement stations and to determine with high probability the boundaries of precipitation classes, transitional zones and unclassified areas. An example of the graphic visualization of such zones is presented in Figs. 7, 8.

Option I presents the boundaries of precipitation classes for which the precipitation value is a dominating feature of division. The boundaries change depending on altitude and terrain, but those are not the dominating factors. Unclassified fields show lack of patterns (a result of bad design of station



**Fig. 7** Indication of precipitation classes' boundaries, transitional zones and unclassified areas—option I

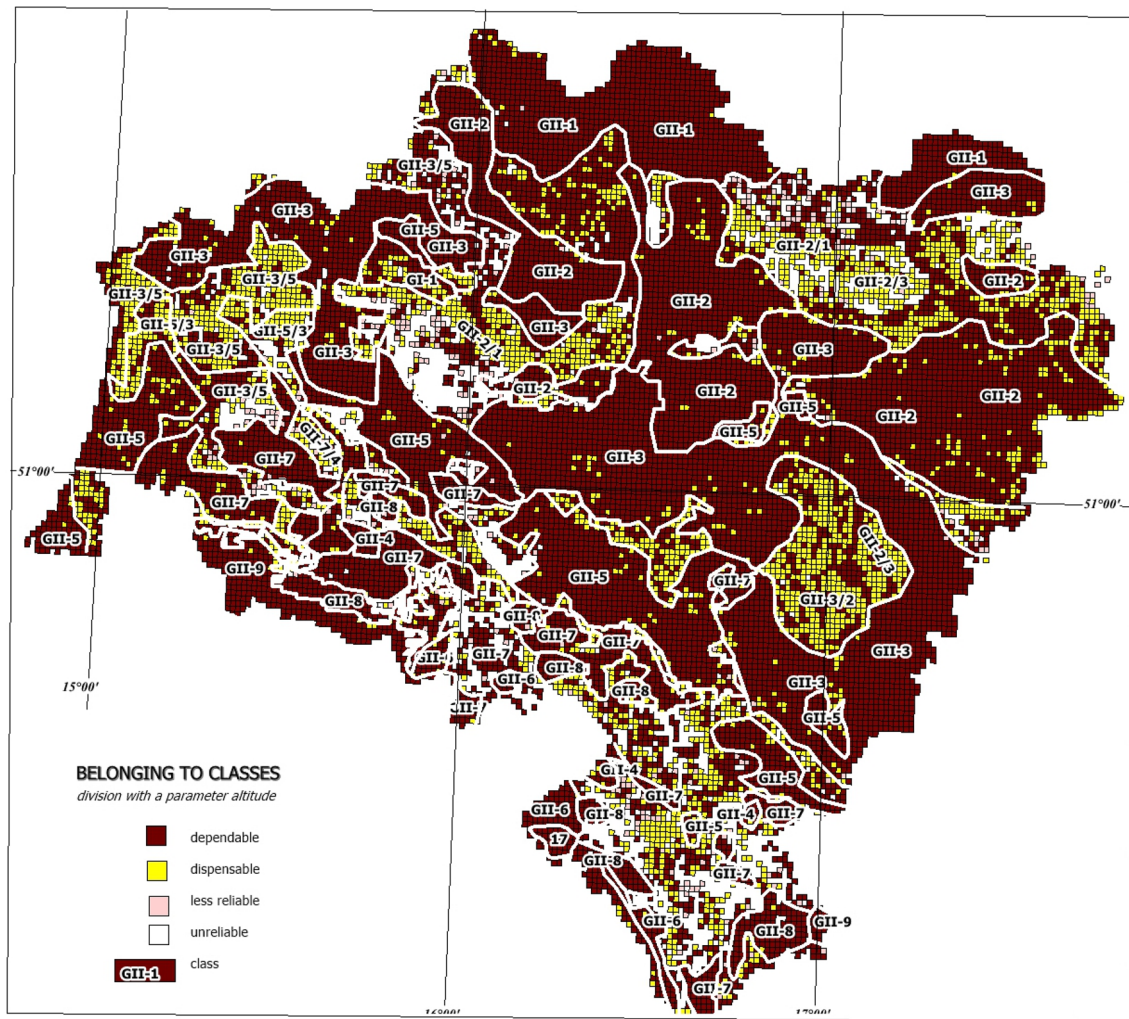
location) for precipitation classes near the tops of hills and mountains and in valleys away from rivers. Option II presents the borders boundaries of precipitation classes, where altitude determines the division of precipitation classes. The division and classification refer to the geological structure in the tested area directly connected with the terrain. Transitional zones mainly refer to tops of hills and hills without measurement stations.

## Conclusion

Thematic layers connected with the distribution of precipitation are important elements of spatial databases in all environmental studies. This article presents different attitudes towards building thematic layers that concern precipitation in relation to widely used geostatistical modelling and interpolation functions. The constructed model does

not directly show the numerical values of precipitation but presents the interval of possible values corresponding to separated classes. The result of classification is given with the probability of belonging to class. For accepted elementary units of area, the possibility of having a few features of pattern classes was assumed; in practice, the most frequently recognized is one main class and possibly an alternative class, or there is a lack of recognition if the probability of belonging is below the accepted threshold. The result of classifying elementary fields that are above the accepted threshold to a certain class is interpreted as a significant similarity to the pattern conditions in the localization of measurement stations and the high probability of the occurrence of similar precipitation in terms of value intervals corresponding to the recognized class. The alternative class appears most often in the transitional zone between areas with certain selection of base class. In reality, the alternative class can correspond to the areas





**Fig. 8** Indication of precipitation classes' boundaries, transitional zones and not classified areas—option II

with precipitation and conditions similar to two or more classes. Not belonging to the class may mean the lack of measurement stations in the area with the conditions that appear in the classified elementary filed or incorrect measurement data at measurement stations. Because of the presented features, the model can be applied to realize the following tasks:

- spatial analysis of phenomenon distribution in relation to topographic factors accompanying the studied phenomenon without determining the mathematical function of the relationship between them,
- spatial analysis of phenomenon distribution taking into account attributive data describing factors accompanying the studied phenomenon,
- indicating borders of phenomenon's spread and transitional zones on the basis of point measurements,
- indicating areas without patterns, that is, measurement stations in the areas with features not similar to classified areas.

The presented model of classification can be used if there is a large amount of data and the possibility of creating a training set that includes full differentiation of features describing conditions and a full range of precipitation values in a given area. The result of classification should be treated as an introduction to advanced models created together with specialists, such as geostatistical modelling within the borders of designated classes. Hierarchical identification of ordered zones, estimated and marked in the database as basic, alternative or unrecognized classes, allows the user to maintain subjectivity when making decisions during the modelling of the studied climate parameter. GIS-based multicriteria decision analysis (GIS-MCDA)

arbitrate many spatial decision problems (Malczewski 2006). One of the most remarkable features of the GIS-MCDA approaches is the wide range of decision. Dividing zones of uncertainty indicate areas for which the decision has to be preceded by additional analyses of values that influence the chosen climate parameter. New technologies and the development of knowledge allow for the rationalization of many research processes of not well-recognized phenomena. The example is the attempt to use qualitative data to elaborate climatic regions using neural networks. The result of that work is receiving the compatibility of area delimitation with the perspective of increasing the detailed indication of dividing zones if including other data into training sets, which explains the distribution of the phenomena. The proposed way of classifying areas and indicating similarities to the surroundings of measurement points can be also implemented in studies on phenomena measurements in real time, daily, monthly and decade averages, and from many-years average measurements. Each classification model should be preceded by an analysis of the data influencing or accompanying the studied phenomenon. Indicating zones with high probability for which the values of natural parameters have been determined is significant for interdisciplinary studies when specialists have to cooperate closely. Reliable information allows for effective cooperation. Data about climatic parameters are part of informative systems concerning the management of natural resources and following the current course of phenomena for early warning against natural disasters. The construction of probability maps and indicating the categories of climatic data reliability allow us to distinguish areas for which values can be questionable (for instance, the risk of precipitation occurrence and the strategy for agriculture development, flood cover, etc.). It seems that it is sometimes better to have information about the lack of data in some areas of studied terrain than to use interpolated data that do not correspond to reality and on that basis make wrong decisions (Mitas and Mitasova 1988, 1999; Jiang and Eastman 2000; Eastman 1999; Store and Kangas 2001; Zhang and Goodchild 2002; Bac-Bronowicz 2004; Nas et al. 2010; Huang et al. 2011).

**Acknowledgements** This work was financed as a research project at the Faculty of Geoengineering, Mining and Geology (No. S30134, No. S40043). Calculations were carried out at the Wrocław Centre for Networking and Supercomputing <http://www.wcss.wroc.pl> (Grant No. 336).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Al-Zahrani M, Husain T (1998) An algorithm for designing a precipitation network in the south-western region of Arabia. *J Hydrol* 205:205–206
- Bac-Bronowicz J (2004) Technique for constructing continuous meteorological fields from point measurements using incorporated effects of topography. In: 10th EC-GI and GIS Workshop, ESDI: The State of the Art, Warsaw, <http://www.ec-gis.org/Workshops/10ec-gis/papers/poster.bac-bronowicz.pdf>. Accessed Nov 2016
- Bac-Bronowicz J (2005) Topographic factors as a usable tool for correct modeling of continuous features indicated in points. In: Proceedings of the 22th international cartographic conference 'Mapping Approaches into a Changing World', International Cartographic Association, [http://www.cartesia.org/geodoc/icc2005/pdf/oral/TEMA8/JOANNA\\_BAC-BRONOWICZ.pdf](http://www.cartesia.org/geodoc/icc2005/pdf/oral/TEMA8/JOANNA_BAC-BRONOWICZ.pdf). Accessed Nov 2016
- Bac-Bronowicz J (2007) GIS as a useful tool for spatial distribution of climate parameters observed in point over relatively long period of time. GIS development. In: Conference map Asia, International Cartographic Association, Kuala Lumpur
- Bac-Bronowicz J (2010) Cartographic presentation of spatio-temporal information related to point data on the basis of historical data from atlases of precipitation. In: Livieratos E, Gartner G (eds) Proceedings of 5th international workshop on digital approaches in cartographic heritage. Vienna University of Technology, Vienna
- Badura J, Przybylski B, Zuchiewicz W (2004) Cainozoic evolution of Lower Silesia, SW Poland: a new interpretation in the light of sub-Cainozoic and sub-Quaternary topography. *Acta Geodyn Geomater* 1(3):7–29
- Basist A, Bell G (1994) Statistical relationship between topography and precipitation patterns. *J Clim* 7:1305–1315
- Bishop C (1995) Neural networks for pattern recognition. Clarendon Press, Oxford
- Borcard D, Gillet F, Legendre P (2011) Cluster analysis. Springer, New York, pp 53–114. [https://doi.org/10.1007/978-1-4419-7976-6\\_4](https://doi.org/10.1007/978-1-4419-7976-6_4)
- Bridle J (1989) Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In: Fogelman-Soulie F, Hérault J (eds) Neuro-computing: algorithms, architectures. Springer-Verlag, New York
- Changhyoun L, Kyungrock P, Yosang L (2014) Optimal sampling network for monitoring the representative water quality of an entire reservoir on the basis of information theory. *J Water Clim Change* 5(2):151–162
- CLC (2006) CORINE land cover (CLC). Główny Inspektorat Ochrony Środowiska, <http://clc.gios.gov.pl/index.php?IdCss=0>
- Climateprediction.net (2015) GCMs-global climate models. [http://www.ipcc-data.org/guidelines/pages/gcm\\_guide.html](http://www.ipcc-data.org/guidelines/pages/gcm_guide.html)
- Covey C, Dai A, Lindzen R, Marsh D (2014) Atmospheric tides in the latest generation of climate models. *J Atmos Sci* 71:1905–1913. <https://doi.org/10.1175/JAS-D-13-0358.1>
- Cressie N (1991) Statistic of spatial data. Wiley, New York
- Dai A (2005) Precipitation characteristics in eighteen coupled climate models. *J Clim* 19(18):4605–4630
- Dubicki A (ed) (2002) Zasoby Wodne w dorzeczu górnej i środkowej Odry w warunkach suszy. Atlasy i monografie, IMGW Warszawa (in Polish)
- Eastman JR (1999) Multi-criteria evaluation and GIS. In: Longley P, Goodchild M, Maguire D, Rhind D (eds) Geographical information systems. Wiley, New York, pp 493–502
- ETC/BD (2006) The indicative map of European biogeographical regions: methodology and development. Tech. rep., European Topic Centre On Biological Diversity, <http://www.eea.europa.eu>

- [a.eu/data-and-maps/data/biogeographical-regions-europe-2005/methodology-description-pdf-format/methodology-description-pdf-format/at\\_download/file](#)
- Goovaerts P (1998) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J Hydrol* 228:113–129
- Gouvas M, Sakellariou N, Xystrakis F (2009) The relationship between altitude of meteorological stations and average monthly and annual precipitation. *Stud Geophys Geod* 53(4):557–570. <https://doi.org/10.1007/s11200-009-0039-1>
- Hamilton K, Ohfuchi W (eds) (2007) High resolution numerical modelling of the atmosphere and ocean. Springer, New York
- Huang IB, Keisler J, Linkov I (2011) Multi-criteria decision analysis in environmental sciences: ten years of applications and trends. *Sci Total Environ* 409(19):3578–3594. <https://doi.org/10.1016/j.scitotenv.2011.06.022>
- Jiang H, Eastman JR (2000) Application of fuzzy measures in multi-criteria evaluation in GIS. *Int J Geogr Inf Sci* 14(2):173–184
- Kondracki J (2000) Polish regional geography (Geografia regionalna Polski). PWN, Warszawa (in Polish)
- Kunz M, Kottmeier C (2006) Orographic enhancement of precipitation over low mountain ranges. Part II: simulations of heavy precipitation events over southwest Germany. *J Appl Meteorol Climatol* 45(8):1041–1055
- Li Z, Zhu Q, Gold C (2005) Digital terrain modelling: principles and methodology. CRC Press, Boca Raton
- Malczewski J (2006) GIS-based multicriteria decision analysis: a survey of the literature. *Int J Geogr Inf Sci* 20(7):703–726
- Manoj KJ, Umesh CK, Kittur GRR (2004) A GIS based distributed rainfall-runoff model. *J Hydrol* 299(1):107–135
- McAvaney B, Covey C, Joussaume S, Kattsov V, Kitoh A, Ogana W, Pitman A, Weaver A, Wood R, Zhao ZC, et al (2001) Model evaluation. In: Climate change 2001: the scientific basis. Contribution of WG1 to the third assessment report of the IPCC (TAR), Cambridge University Press, Cambridge, pp 471–523
- Meehl GA, Washington WM, Collins WD, Arblaster JM, Aixue HU, Buja LE, Strand WG, Haiyan T (2005) How much more global warming and sea level rise? *Science* 307:1769–1772. <https://doi.org/10.1126/science>
- Metzger M, Bunce R, Jongman R, Mucher C, Watkins J (2005) A climatic stratification of the environment of Europe. *Glob Ecol Biogeogr* 14(6):549–563. <https://doi.org/10.1111/j.1466-822X.2005.00190.x>
- Mitas L, Mitasova H (1988) General variational approach to the interpolation problem. *Comput Math Appl* 16:983–992
- Mitas L, Mitasova H (1999) Spatial interpolation. In: Longley P, Goodchild M, Maguire D, Rhind D (eds) Geographical information systems, 2nd edn. Wiley, Hoboken, pp 481–492
- Mitchell D, Carter T, Jones P, Hulme M, New M (2004) A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100). [http://ipcc-data.org/docs/tyndall\\_working\\_paper\\_s\\_wp55.pdf](http://ipcc-data.org/docs/tyndall_working_paper_s_wp55.pdf). Accessed Nov 2016
- Moller M (1993) A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw* 6:525–533
- Mutasem SA, Umi KN, Lea TT, Nor AMI, Mohammad SAB (2013) Determination of important topographic factors for landslide mapping analysis using MPL network. *Sci World J*. <https://doi.org/10.1155/2013/415023>
- Myoung-Jin U, Hyeseon Y, Woncheol C, Jun-Haeng H (2010) Analysis of orographic precipitation on Jeju-Island using regional frequency analysis and regression. *Water Resour Manag* 24:1461–1487. <https://doi.org/10.1007/s11269-009-9509-z>
- Nas B, Cay T, Iscan F, Berkay A (2010) Selection of MSW landfill site for Konya, Turkey using GIS and multi-criteria evaluation. *Environ Monit Assess* 160(1–4):491–500. <https://doi.org/10.1007/s10661-008-0713-8>
- Noirfalise A (2007) Map of the natural vegetation of the member countries of the European Community and of the Council of Europe, scale map 1:3,000,000, second edition. Council Of Europe, Commission Of the European Communities, <https://publications.europa.eu/en/publication-detail/-/publication/465907c2-e754-47e2-9bbd-01d92be17cb2/language-en>. Accessed Sep 2018
- Pawlak W, Anioł-Kwiatkowska J, Pawlak J, Nowak-Ferdus E, Migoń P, Malicka A, Marciniak A, Żak W (2008) Atlas of Lower and Opole Silesia. In: Collection of maps, Wrocław University, Polish Academy of Sciences - Wrocław Branch, printed maps (in Polish)
- Podlacha K (1986) Kartograficzny system TEMKART dla komputerowego sporządzania map tematycznych. *Prace Instytutu Geodezji i Kartografii* XXXIII(2):3–18 (77) (in Polish)
- Prudhomme C, Reed D (1999) Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. *Int J Climatol* 19:1337–1356
- Roekaerts M (2002) The biogeographical regions map of Europe - basic principles of its creation and overview of its development. Tech. rep., European Environment Agency, [http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-2001/methodology-basic-principles-of-the-biogeographical-regions-map-creation-and-overview-of-its-development/methodology-basic-principles-of-the-biogeographical-regions-map-creation-and-overview-of-its-development/at\\_download/file](http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-2001/methodology-basic-principles-of-the-biogeographical-regions-map-creation-and-overview-of-its-development/methodology-basic-principles-of-the-biogeographical-regions-map-creation-and-overview-of-its-development/at_download/file). Accessed Sep 2017
- Shafiei M, Ghahraman B, Saghafian B, Gharari Pande S S, Davary K (2014) Assessment of rain-gauge networks using a probabilistic GIS based approach. *Hydrol Res* 45(4–5):551–562. <https://doi.org/10.2166/nh.2013.042>
- Soczyńska U, Nowicka B, Somorowska U, Kupczyk E, Suligowski R (1997) Prediction of design storms and floods. *Int Assoc Hydrol Sci* 246:297–304
- Sreejesh S, Mohapatra S, Anusree M (2014) Cluster analysis. Springer International Publishing, Cham, pp 229–244. [https://doi.org/10.1007/978-3-319-00539-3\\_10](https://doi.org/10.1007/978-3-319-00539-3_10)
- Stach A (2010) Spatial estimation of precipitation (ocena przestrzenna opadów). In: Zwoliński Z (ed) GIS - woda w środowisku, Bogucki Wydawnictwo Naukowe, pp 23–50, (in Polish)
- Store R, Kangas A (2001) Integrating spatial multi-criteria evaluation and expert knowledge for GIS-based habitat suitability modelling. *Landsc Urban Plan* 14(2):79–93
- Tobler W (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46:234–240
- Tobler W (1979) Cellular geography. In: Gale S, Olsson G (eds) Philosophy in geography. Reidel, Dordrecht, pp 379–386
- Trenberth K, Dai A, Rasmussen R, Parsons D (2003) The changing character of precipitation. *Bull Am Meteorol Soc* 84(9):1205–1217
- Vogl T, Mangis J, Rigler A, Zink W, Alkon D (1988) Accelerating the convergence of the backpropagation method. *Biol Cybern* 59:257–263
- Wisniewski W (1953) Atlas opadów atmosferycznych w Polsce 1891–1930. PIHM, Komunikacyjne, printed maps (in Polish)
- Yoshiharu S, Miyata S, Nakakita E, Hasebe M (2007) Numerical approach on the mechanism of precipitation-topography relationship in mountainous complex terrain. [http://www.mssanz.org.au/MODSIM07/papers/39\\_s55/numericalApproach\\_s55\\_Yoshiharu\\_.pdf](http://www.mssanz.org.au/MODSIM07/papers/39_s55/numericalApproach_s55_Yoshiharu_.pdf). Accessed Nov 2015
- Zhang J, Goodchild MF (2002) Uncertainty in geographical information. Taylor and Francis, New York
- Zhang Z, Gang H, Ni J (2013) Effects of topographical and edaphic factors on the distribution of plant communities in two subtropical karst forests, southwestern China. *J Mt Sci* 10(1):95–104. <https://doi.org/10.1007/s11629-013-2429-7>