Tourism Management 75 (2019) 509–521

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman

Clustering nature-based tourists by activity. Social, economic and spatial dimensions

Marta Derek[a,*], Edyta Woźniak[b], Sylwia Kulczyk[a]

[a] University of Warsaw, Faculty of Geography and Regional Studies, Department of Tourism Geography and Recreation, Krakowskie Przedmieście 30, 00-927, Warsaw, Poland

[b] Space Research Centre of the Polish Academy of Sciences, Earth Observation Department, Bartycka 18a, 00-716, Warsaw, Poland

ARTICLE INFO

Keywords:
Cluster analysis
Spatial analysis
Nature-based destination
Outdoor activities
Travel cost method
Mapping
The Great Masurian Lakes
Poland

ABSTRACT

This study develops a typology of tourists visiting a nature-based destination as a function of outdoor activities. It then identifies social, economic and spatial differences between them. A cluster analysis is applied to survey data collected in the Great Masurian Lakes region of Poland. Six clusters were identified: angling sailors, non-angling sailors, cyclists, anglers, water recreationists and passive tourists. The results show that these groups differ not only in terms of their social and demographic features, or travel patterns. Spatial analyses identify important differences in spatial behaviours, while the Travel Cost Method highlights major differences in the economic value of nature. This new methodological framework shows that cluster analyses, combined with spatial and economic analyses, may be an extremely useful tool for planners and destination managers. The results contribute to the discussion on tourists' use of a nature-based destination.

1. Introduction

The past twenty years have seen a boom in tourism to natural areas (Newsome, Moore, & Dowling, 2012). This is linked to growing urbanization, which not only disconnects people from their natural environment but also fails to meet their basic need for contact with nature. In their search for the benefits that are associated with this environment, such as improved mental and physical health (Bell, Tyrväinen, Sievänen, Pröbstl, & Simpson, 2007), people tend to spend their free time in nature, which is understood here not only as protected areas, but also rural and suburban areas, or even urban parks (Buijs, Pedroli, & Luginbühl, 2006; Casado-Arzuaga, Onaindia, Madariaga, & Verburg, 2013). This, however, does not leave the natural environment untouched, and responsible tourism planning and management is urgently needed. This is a big challenge, as trends in nature-based tourism are changing fast; in particular, as Fredman and Tyrväinen (2010) note, it is increasingly becoming both specialized and diversified. The more heterogenic it is, the more difficult it is to manage, as the needs of different groups of tourists are likely to interfere. This is especially challenging when spatial conflicts arise among tourists who undertake different outdoor activities (Sæþórsdóttir, 2010) (e.g. angling, which needs peace and quiet, and shore recreation or swimming, which can be noisy). According to Beeco and Brown (2013), understanding the spatial context of social data is needed to maintain both the quality of the experience for visitors and the proper protection of resources.

In Poland, the Great Masurian Lakes region is one area where nature-based tourism has been developing. Although the region is not designated as a national park, the mosaic of lakes and forests, together with the traditional rural landscape, have been attracting tourists for many years. In the past few years the number of tourists has increased significantly – by 35% between 2010 and 2015 (Derek, Woźniak, & Kulczyk, 2017) and, therefore, management of the area must take into account both environmental threats and the interests of various groups of tourists.

This study develops a typology of tourists visiting the Great Masurian Lakes as a function of the activities they undertake. In a second step, we identify differences between these types in terms of travel patterns and spatial behaviours, and the economic value given to nature. This is achieved using a cluster analysis based on survey data. Following previous work (e.g. Konu, Laukkanen, & Komppula, 2011; Mehmetoglu, 2007) segmentation was activity-based and characterized by the demographic features of tourists and their travel patterns. Although these aspects are very important, it is also essential to understand the spatial distribution of clusters within the area, and (especially in a nature-based destination) the value different types of tourists attach to nature. In this paper, we aim to address both of these aspects using spatial analysis methods and the Individual Travel Cost Method. Identifying differences is especially useful for planning tourism within the perimeter of the destination, as well as for developing diverse marketing strategies for tourism development.

* Corresponding author.
E-mail addresses: m.derek@uw.edu.pl (M. Derek), ewozniak@cbk.waw.pl (E. Woźniak), skulczyk@uw.edu.pl (S. Kulczyk).

https://doi.org/10.1016/j.tourman.2019.06.014
Received 8 August 2018; Received in revised form 10 March 2019; Accepted 23 June 2019
Available online 02 July 2019

2. Literature review

According to Fredman and Tyrväinen (2010), nature-based tourism is most often associated with leisure activities that take place in natural areas, and its key components are visitors and their experience of (or in) nature. These experiences may differ and involve various outdoor activities. Valentine (1992) proposed three types of relationships between tourists and nature: activities dependent on nature (e.g. birdwatching), activities enhanced by nature (e.g. camping) and activities for which the natural setting is incidental (e.g. swimming). Derek et al. (2017) introduced a Nature Cruciality Index, which expresses the relation between stated preferences regarding the natural setting and a range of outdoor activities. Three groups of activities are distinguished: activities for which nature is crucial (e.g. sailing); neutral (e.g. cycling), or not crucial (e.g. walking).

As tourists' relationships with nature are reflected in the outdoor activities they undertake, it is crucial to know where, within a nature-based destination, these activities take place. There is a wide range of options for gathering such data. In particular, methods that involve advanced technical equipment, such as mobile-tracking technologies (Beeco, Hallo, & Brownlee, 2014; Birenboim & Shoval, 2016) or infrared sensors and camera monitoring (Fairfax, Dowling, & Neldner, 2014) are helpful in acquiring data on tourists' mobility. They have the advantage of high accuracy and objectivity (Isaacson, Shoval, Wahl, Oswald, & Auslander, 2016). However, the cost of equipment and its technical limitations (e.g. battery life) mean that these methods are typically applied in smaller (preferably closed) areas such as national parks or city centres (Kajala, 2007; Shoval & Isaacson, 2007). Consequently, traditional, non-mechanized methods continue to be used.

One of the most frequently-used tool in this respect is the survey (e.g. De Aranzabal, Schmitz, & Pineda, 2009; Moscardo, Pearce, & Morrison, 2001), where respondents are asked how they spend their time in a destination. As data collected in this way describe a wide variety of individual behaviours, there is a need to generalize them in order to get a more precise overview. As Ditton, Goodale, and Johnsen (1975, p. 293) point out, "since recreation patterns are so individualized, it is impractical to examine individuals separately. Instead, individuals can be aggregated into groups within which recreation patterns are similar". Cluster analysis is often used to achieve this. The method is widely used in marketing research, as companies need information about broad classes of buyers who differ in terms of their product needs or buying responses. The main variables used to segment consumer markets are geographic (e.g. size of city), demographic (e.g. age), psychographic (e.g. lifestyle) and behavioural (e.g. attitude to the product) (Kotler, Bowen, & Makens, 2014).

In tourism research, cluster analysis is used to segment visitors. As tourists come from different areas, have different lifestyles and personalities, and travel in different ways, they choose different destinations, means of transport, accommodation and activities. These characteristics, according to Konu et al. (2011), make segmentation a useful way to identify distinct groups. Various variables have been used in the context of segmentation in nature-based destinations: tourists' perceptions of expected benefits (Palacio & McCool, 1997), ecotourists' behaviours (Weaver & Lawton, 2002, 2005), tourists' motivations (Beh & Bruyere, 2007), their motivations and behaviours (Perera, Vlosky, & Wahala, 2012), their attitude to facilities (Ryan & Sterling, 2001), etc.

However, another option is activity-based segmentation. This can be particularly useful in the context of tourism management and planning (Ditton et al., 1975). As Smith, Tuffin, Taplin, Moore, and Tonge (2014, 15) state, in the context of areas of high natural value; in order to successfully manage visitor use, "managers need to know the diversity of experiences that are sought in order to determine the diversity of settings in which these experiences can be undertaken". This diversity of experiences, expressed by activities, can help in defining different types of nature-based tourists (Mehmetoglu, 2007) which, in turn, is important for planning purposes.

Activities have been applied as a variable in many earlier studies, often mixed with other criteria. Hvenegaard and Dearden (1998), for example, segmented tourists visiting a national park in Thailand based on activities, motivations, respondent-self descriptions and researchers' assessments of respondents' dominant activities. They identified five clusters that distinguished respondents in terms of their attitude to nature and found that they only slightly differed with respect to their willingness to pay for nature conservation.

A segmentation of tourists visiting Austria, undertaken by Dolnicar and Leisch (2003), included leisure activities and travel motivations. Here, respondents were asked to indicate activities that they engaged in 'often', 'sometimes' or 'never' during their winter holiday. A bagged clustering framework was implemented, based on behavioural (seven types) and psychographic (five types) segmentations. As a result, five vacation styles were identified. The study concluded that the combination of activities and motivations rendered a more holistic perspective of the tourist and avoided running a cluster analysis on too many variables.

Trip activities, together with tourists' motivations were also a basis of a segmentation of nature tourists carried out by Mehmetoglu (2007). He identified three groups of respondents: 'culture and pleasure activity oriented', 'nature activity oriented' and 'low-activity oriented'. Activities not directly connected to nature (such as dining in restaurants or visiting museums) were also included. Clusters were distinguished based on demographics, trip features, and trip motivations. Unlike Dolnicar and Leisch (2003), Mehmetoglu (2007) concluded that in nature-based tourism research, motivation and activity dimensions should be treated separately, and that activity-based segmentation better served the needs of the tourism industry. Similarly, Moscardo et al. (2001) tested the usefulness of several criteria in segmenting tourists and found that activities were the most useful and efficient.

None of these studies, however, used spatial analyses in order to show the distribution of clusters, or verified the economic value of different types of tourists. All of the earlier work has identified activities undertaken by tourists for each study site as a whole, while any internal variety in the spatial distribution of different groups of tourists has been neglected. The same can be said for environmental value – to the best of our knowledge, no previous work has examined the relation with tourists' activities. In this paper, we do not simply report the results of a cluster analysis. Instead, we aim to contribute to the discussion of tourists' use of a nature-based destination by analysing the spatial distribution of tourist accommodation within each cluster, and showing that different clusters attach different economic value to nature. As far as we know, this is the first attempt to combine cluster analysis with spatial analyses, and cluster analysis with the economic value given to nature.

3. Methodology

3.1. Study site

The Great Masurian Lakes region is located in north-east Poland and covers 2811 km^2. The area consists of 10 municipalities, including seven small and medium-sized towns (Fig. 1). With its numerous lakes, forests, fields and meadows, it is widely perceived as a nature-based destination. A broad range of outdoor activities that are more-or-less active, and land- and water-based can be undertaken, almost exclusively in summer. The area is particularly interesting for sailors as it is the only place in Poland (and one of only a few in Europe) with such a long inshore sailing area: the primary route is 110 km long, with numerous secondary waterways. This means that sailors can travel for many days, mooring in marinas or offshore (using their yachts as accommodation).

3.2. Data collection

We carried out an on-site survey of tourists. As Wall Reinius (2011) points out, questionnaires are widely used for collecting data on tourists in natural areas. The survey was conducted in nine towns or villages throughout the area, i.e. in every municipality in the area (see Fig. 1).

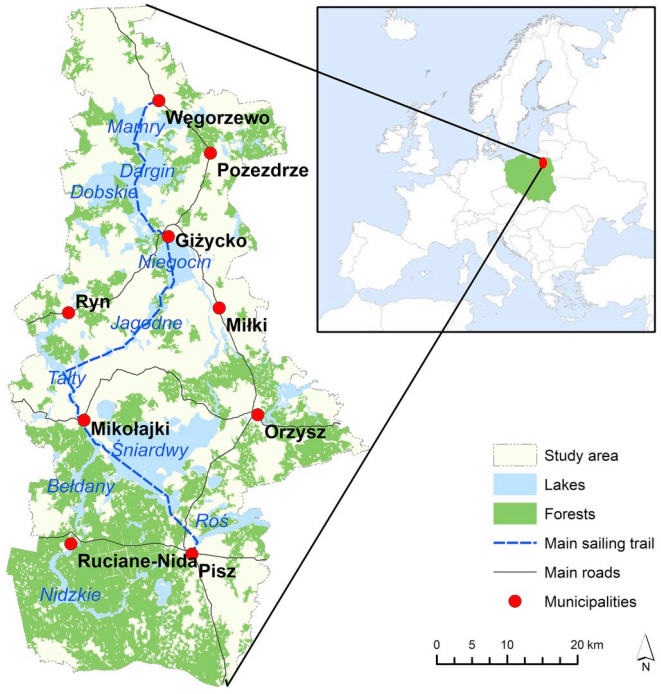

Fig. 1. Study area.

Specifically, it was undertaken in front of convenience stores or small supermarkets (if there were any) as we assumed that every tourist (from sailors to leisure tourists) has to purchase food and drink from time to time. We deliberately chose not to collect data in tourist centres, tourist accommodation facilities, marinas, or any other tourist establishments, in order to avoid any particular activity having an undue influence. For example, more sailors or others interested in water-based activities would be found in marinas or near the shores of a lake, while more leisure tourists would be found in hotels, etc. The aim was to obtain a sample that was as reliable as possible in terms of the representativeness of the activities undertaken by tourists.

The survey was carried out over two weekends and two different working days during the summer of 2014, in the morning, during the day and in the evening, simultaneously at all nine locations. It was conducted in Polish as, according to official statistics, domestic clientele constitutes the main group of visitors in the study area (Derek et al., 2017). Non-eligible respondents were filtered out (e.g. local residents or people whose stay in the region was unconnected to tourism or leisure). A pilot survey was run over a weekend in June. Fifty tourists in the four biggest towns were asked to complete the survey, to ensure that the questionnaire was complete, clear, and reliable. This resulted in the introduction of some changes. The questionnaire was broken down into several sections: general questions about the visit (such as the reason for travel, length of stay); respondents' socio-demographic background; detailed questions about their accommodation; and very detailed questions about what outdoor recreational activities were undertaken during their stay.

Respondents could choose from seventeen outdoor recreational activities. This list was developed from a literature review and the authors' experience of tourism in the region. It included eight water-based activities (bathing in a lake or a river, sailing, fishing, using a boat or pedalo, kayaking, motor boating, windsurfing, scuba diving) and nine land-based activities (walking, picking mushrooms or wild fruit, horse riding, cycling, sunbathing or picnicking, visiting natural attractions, observing wild animals, outdoor motorsports, paragliding). Respondents were also asked to indicate if they had participated in any other activities that were not listed (nothing significant was reported). They were asked to indicate if, and how often, they had performed a given activity during their stay. Responses were recorded on a five-point ordinal scale, where: 1 = no, 2 = rarely, 3 = from time to time, 4 = often, 5 = very often, and 9 = difficult to say (coded as missing values).

A total of 516 interviews were conducted with tourists using a paper-and-pencil questionnaire. Fourteen were excluded from the analysis because of incomplete responses, leaving a total of 502. Data were entered into the Statistical Package for the Social Sciences (SPSS) software to calculate clusters and carry out statistical analyses. ArcGIS 10.5 software was used to map where clusters stayed and analyse their spatial distribution.

Data on location of accommodation facilities and the number of beds were collected from the municipalities and local tourist organisations, and, where necessary, supplemented by phone interviews.

3.3. Data analysis

3.3.1. Clustering procedure

An exploratory cluster analysis (Hair, Black, Babin, & Anderson, 2014) was performed in order to produce activity-based segments. The aim was to distinguish clusters of visitors by the type and frequency of their nature-based recreation activities.

A hierarchical clustering procedure was adopted based on Ward (1963). This procedure creates a treelike structure and generates a complete set of solutions, ranging from all one-member clusters to the one-cluster solution where all observations are in a single cluster. The researcher can observe and control the process and judge how many clusters should be retained, rather than deciding the number of clusters in advance. Ward's method is appropriate when the researcher wants clustering patterns to reflect clusters of more-or-less equal size (Hair et al., 2014) and the dataset does not include outliers (Mooi & Sarstedt, 2011). It provides a useful indication of the number of groups to be used for operational purposes.

As recommended in earlier work (Hair et al., 2014), the squared Euclidean distance was used as the distance measure. The ordinal scale was treated as if it was a ratio scale. As Mooi and Sarstedt (2011) note, market researchers usually treat ordinal data as metric data to calculate distance metrics by assuming that scale steps are equidistant. This approach is also widely used in tourism studies (e.g. Konu et al., 2011; Mehmetoglu, 2007; Weaver & Lawton, 2005).

A univariate analysis of variance (ANOVA) identified any statistically significant differences between clusters in terms of variables. F-statistics indicated which factors best differentiated segments, while p-values indicated the level of significance.

In order to define and name clusters we calculated the mean for every activity within each cluster and compared this to the mean for the whole sample. This captured the key characteristics of each cluster and made it possible to name it as a function of the activities that were undertaken most frequently by its members.

3.3.2. Demographic analyses

Regarding demographic characteristics, comparisons between clusters included: gender, age, educational level, place of residence, length of stay, visitor type, number of people in the group, and respondents travelling with children.

3.3.3. Individual Travel Cost Method

The Individual Travel Cost Method measured the economic value that each cluster placed on nature. In this method, the value of a given good is estimated based on how much people pay (or are willing to pay) to access and 'use' it (Czajkowski, Giergiczny, Kronenberg, & Tryjanowski, 2014). Assessing the monetary value of non-market goods is a helpful management tool, as it makes it possible to compare various types of goods and use scenarios. It is also an important marketing tool, because monetary value is easy to understand for non-specialists and therefore helps to make the vague notion of the value of nature real.

Our analysis included the following elements (see Kulczyk, Derek, & Woźniak, 2016 for more details):

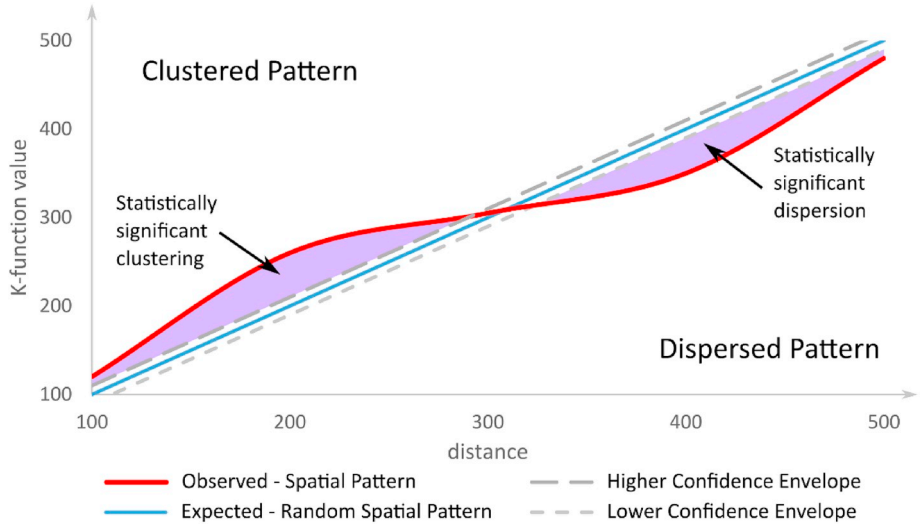

Fig. 2. Interpretation of the Multi-Distance Spatial Cluster Analysis based on Ripley's K-function.

- The cost of the journey: calculated as the sum of the cost of travel distance (distance travelled from the place of residence to the destination multiplied by the price of fuel) and cost of travel time (journey time multiplied by the cost of journey time, described as a percentage of the mean wage in Poland). Travel distances and times were calculated using Google Maps based on data provided by surveyed individuals;
- Accommodation cost: the mean price of accommodation per activity type multiplied by the respondent's length of stay. In the case of sailors who used their yacht as accommodation, marina fees were used;
- Use cost. The cost of activities undertaken by tourists estimated on the basis of the cost of equipment rental.

All of the above costs were calculated in Polish zloty, recalculated into euros and summed to obtain individual, per tourist, travel costs. Next, means were calculated for each cluster.

3.3.4. Spatial analysis

Spatial analysis of distribution of clusters' members was done based on their places of accommodation. All places of accommodation reported in the survey, and their spatial distribution, were mapped. The clusters' distribution was also compared to the location of all accommodation facilities in the study area.

Clusters were evaluated by analysing their geographic distribution and spatial patterns. The same procedure was used regarding the supply of accommodation facilities. The following measures were used:

- Directional Distribution (Standard Deviational Ellipse) (Chew, 1966; Mitchell, 2005; Wang, Wenzhong, & Zelang, 2015) summarizes the geographic distribution of features: central tendency, dispersion, and directional trends (rotation). The *x* and *y* axes of an ellipse are calculated separately as a standard distance in the *x* and *y* direction from the mean centre which is the average *x* and *y* coordinate of all the features in the study area. The measure was used, for example, in monitoring urban activity using mobile phone–based sensor data (Ahas et al., 2015), or in analysing the dispersion of historic monuments (Jażdżewska, 2018).
- Average Nearest Neighbour (Ebdon, 1985; Mitchell, 2005) is used in order to measure the degree of spatial clustering. It was used, for example, in order to identify clustering patterns of restaurant location in Hamilton (Prayag, Landré, & Ryan, 2012), or to analyse clustering of tourist attractions in Macedonia (Vasiliadis & Kobotis, 1999). It measures the Euclidean distance between each point and the location of its nearest neighbour. Next, the average distance is calculated on the basis of all partial distances. This is compared to the average for a hypothetical random distribution. The average nearest neighbour ratio index is established by dividing the observed distance by the hypothetical one. The pattern of the analysed dataset is said to be spatially clustered if the index < 1 or dispersed if the index > 1. Values around 1 indicate a random distribution. The confidence level indicates the percentage of cases with the pattern indicated by the nearest neighbour ratio.
- Multi-Distance Spatial Cluster Analysis is based on Ripley's K-function (Bailey & Gatrell, 1995; Mitchell, 2005) and is used to quantify the degree of spatial dependence (see for example clustering of fast food restaurants in Chicago; Austin et al., 2005). It determines whether the clustering of features is statistically significant or dispersed over a range of Euclidean distances from the geometric centres of classes. In our case, five ranges were taken into account. The confidence envelope, which determines statistical significance, was established using nine permutations to obtain a graph that shows the expected, random pattern and the observed pattern. The dataset is spatially clustered if the curve of the observed pattern is above the expected one, otherwise it indicates dispersion (Fig. 2).

The Standard Deviational Ellipse method was applied to identify the location of accommodation where tourists stayed. The latter two methods aim to indicate the spatial pattern of tourists constituting each cluster.

4. Results

4.1. Sample profile

The sample profile included slightly more men, while the majority of visitors were aged between 25 and 44. Respondents predominantly came from large towns and cities, and almost half possessed a university degree. Regular visitors, defined as respondents who had visited the region more than five times (Konu et al., 2011) dominated, while there were very few first-time visitors (see 'Sample profile' in Table 2 for details).

4.2. Accommodation facilities in the study area

Spatial distribution of all accommodation facilities is presented in Fig. 3. Vast majority of them is concentrated in the central-western part of the study area. Eastern part of the Great Masurian Lakes offers very limited accommodation opportunities. The concentration of accommodation places is reflected by Standard Deviational Ellipse. It covers

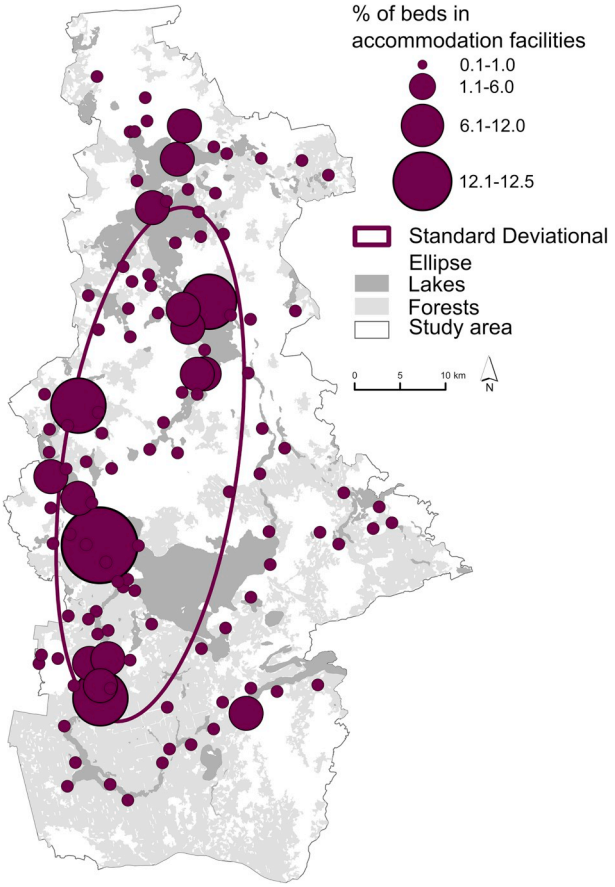

Fig. 3. Share of beds in accommodation facilities and their spatial distribution in the study area.

only 31% of the area.

Average Nearest Neighbour analysis shows that accommodation places are spatially aggregated to a large extent (ratio = 0.000096, confidence level 99%). 37% of beds are located in 4 main localities. Multi-Distance Spatial Cluster Analysis based on Ripley's K Function shows that accommodation places are aggregated in all distance ranges (Fig. 4).

4.3. Defining clusters

As we aimed to create clusters for all 17 recreational activities, initially all were considered as variables. However, of these, five did not usefully differentiate the sample and, following Hair et al. (2014), were excluded (horse-riding, windsurfing, outdoor motorsports, scuba-diving and paragliding). The first cluster analysis (containing two to six clusters) and ANOVA found that one of the remaining variables (visiting natural attractions) was not significantly different in five- and six-cluster solutions (the best options for interpretation) and, following Weaver and Lawton (2005), was also discarded. The cluster analysis was then re-run using the remaining 11 variables. Two-, three-, four-, five-, six- and seven-group solutions were analysed, and the six-group solution was chosen because it differentiated most clearly between clusters.

Means were calculated for each activity within each cluster and compared to means for the whole sample. This found that results for *walking* and *sunbathing or picnicking* were very similar and therefore did not help to distinguish the two groups (Table 1). Fig. 5 shows that two clusters were labelled *sailors* (sailing was almost never recorded in other clusters). The two sailing clusters were distinguished at a third level by angling. The first sailing sub-group were labelled *angling sailors* (Table 1), while the remainder could not be differentiated by a particular activity and were simply named *non-angling sailors*.

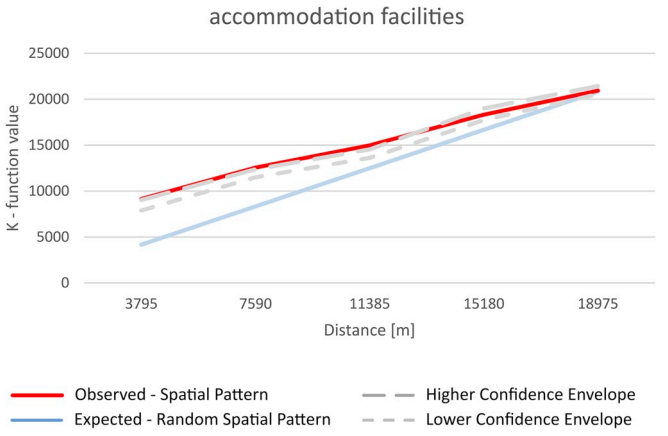

Fig. 4. Spatial distribution of accommodation facilities for distance ranges using Ripley's K-function.

Non-sailors consisted of four groups (Fig. 5) that, at a first level, could be distinguished by their level of activity. *Cyclists* also participated in bathing and sunbathing (and more rarely kayaking and using a boat or pedalo), but mostly they cycled. *Anglers* made up a second group. The third group was the biggest. This cluster engaged in sunbathing or picnicking, walking, or bathing in a lake or river. Consequently, they were called *water recreationists*. Finally, the last (smallest) cluster comprised the least-active respondents, who were labelled *passive tourists*. This was the only cluster whose members did not bath in a lake or river. The only activities undertaken by this group were, from time to time, walking and sunbathing or picnicking (Table 1, Fig. 5).

The ANOVA analysis showed that sailing and angling (F-value > 400), cycling (F-value > 200; p-value 0.000) and bathing in a lake or river (F-value > 80) were the activities that differentiated between groups to the greatest extent, while the remaining activities did not distinguish clusters.

4.4. Differences between clusters

4.4.1. Social characteristics

The most male-dominated cluster was *angling sailors*, followed by *non-angling sailors* and *anglers* (Table 2). The other three clusters were not dominated by gender. The only significant difference was found for passive tourists where there was > 10% more women than men. Regarding age, *cyclists* contained the largest number of young people (aged 18–24), while no other significant difference was found (max difference ~6%). *Angling sailors* was heavily dominated by 25–34 year-olds (~20% more than the next age group) as was *anglers*, although by a smaller margin. Most passive tourists were aged 35–44. Although most respondents aged over 64 (presumably pensioners) are in the passive tourist group, they remain in the minority. Respondents with a university degree dominated among sailors and passive tourists (Table 2).

Respondents from villages and small towns (less than 20 000 inhabitants) dominated among anglers and passive tourists. Respondents from the largest cities dominated the four other clusters, especially *non-angling sailors* (66.6% live in cities with more than 100 000 inhabitants) and *cyclists* (61.4%). Surprisingly, *angling sailors* could be distinguished from *non-angling sailors* in this respect: the former living mostly in smaller towns and cities. Only 14.3% of the angling group lived in cities with more than 500 000 inhabitants, compared to 35.6% of non-angling sailors.

In almost every cluster the dominate length of stay was 4–7 days. It appears that in summer, the typical stay is up to one week. A significant exception is *angling sailors*, where 84.7% of the group spends more than one week, while a significant percentage stays more than two weeks (Table 2).

Regular guests dominate in every cluster, but especially the two sailing clusters. More than 94% of these groups was visiting the region for at least the sixth time. One reason is that the area is unique in

Table 1
Means of cluster variables and ANOVA results.

Variable (activity)	Overall mean (502; 100%)	Cluster means						ANOVA results	
		Angling sailors (91; 18.1%)	Non- angling sailors (87; 17.3%)	Cyclists (70, 13.9%)	Anglers (73, 14.5%)	Water recreationists (123, 24.5%)	Passive tourists (58, 11.6%)	F-value	p-value
Sailing	2.3	4.6	4.1	1.3	1.2	1.1	1.3	418.14	0.000
Motorboating	1.2	1.5	1.0	1.3	1.1	1.3	1.1	6.96	0.000
Kayaking	1.7	1.4	1.2	2.2	2.0	1.9	1.7	10.76	0.000
Using a boat or a pedalo	1.9	1.4	1.3	2.0	2.4	2.1	1.9	12.75	0.000
Bathing in a lake or river	4.1	4.6	4.2	4.2	4.4	4.6	1.7	84.96	0.000
Angling	2.1	4.3	1.1	1.4	3.7	1.1	1.1	419.30	0.000
Walking	3.7	3.1	3.8	4.0	3.8	4.1	3.4	9.09	0.000
Cycling	1.8	1.2	1.1	4.3	2.3	1.1	1.3	213.88	0.000
Picking mushrooms or wild fruit	1.3	1.2	1.3	1.6	1.6	1.1	1.1	8.24	0.000
Observing wild animals	1.2	1.2	1.1	1.4	1.3	1.2	1.2	2.49	0.031
Sunbathing or picnicking	3.5	3.2	3.4	3.6	3.7	3.8	3.2	4.06	0.000

Note: 1 = never; 2 = rarely; 3 = from time to time; 4 = often; 5 = very often.

Poland. First-time visitors constituted more than 10% of respondents in every other cluster, with maximum of 17% in the case of the cyclists.

Very few tourists travelled alone. In most clusters, tourists travelled as a group of 3–4, with *angling sailors* as the most 'sociable' group. Only *cyclists* differed, preferring to travel as a couple.

Fig. 6 shows that respondents varied in terms of their type of accommodation (see Table A in the Appendix for more detailed data). While participants were free to choose more than one type, 85–90% of members of every cluster (apart from sailors) chose only one. Most sailors preferred to stay on their boat, which was moored offshore. This type of accommodation was chosen at least once by nearly 80% of *angling sailors* and 46% of *non-angling sailors*. While 60% of the first group also used marina facilities at least once, this was far less popular in the latter group (32%). Staying with friends and relatives was the first choice of *passive tourists*, followed by staying in a hotel (19%) or bed and breakfast accommodation. Hotel accommodation was less popular with all other clusters, although it should be noted that hotels are not very numerous in the region. Members of the three other clusters (especially *anglers*) picked bed and breakfast accommodation the most often, followed by visiting friends and relatives and staying in a second home. Tents were most popular among *cyclists* and *water recreationists*.

Note: Percentages do not total 100% as respondents could select more than one type of accommodation.

4.4.2. Individual Travel Cost Method

The two clusters that value nature the most are *sailors*, especially *angling sailors*, whose per-person travel costs are estimated at an average of 325 euros. This is because the cost of renting a yacht is very high (use cost) and angling sailors stay for the longest time among all clusters (see section 4.4.1). On the other hand, sailors rarely pay for accommodation, as they often sleep on their yachts that are moored offshore.

Clusters that spend the least money are *water recreationists* (107 euros) and *passive tourists* (114 euros). This may be explained by the fact that neither cluster participates in a specialized activity, and the activities they do undertake the most often (walking, sunbathing, or even bathing in a lake or a river) can be done almost anywhere. They also do not spend money on facilities (use costs are estimated at 8 euros in both cases).

Travel costs clearly divide clusters into two groups: *sailors*, whose travel costs are derived from use costs (renting yachts) and others whose costs are mainly associated with accommodation. Journey costs are least diverse among clusters (the maximum is around 20%). *Anglers*, *angling sailors* and *sailors* have the highest costs, as they travel the longest distances to their destination (Fig. 7) (see Table B and Table C in the Appendix for detailed data).

4.4.3. Spatial behaviours

Fig. 8 shows the spatial distribution of each cluster. The percentage of cluster members staying in a given place of accommodation and standard deviational ellipses are shown. Most tourists in each cluster congregate in towns located near the large lakes along the main sailing route. There are, however, some differences.

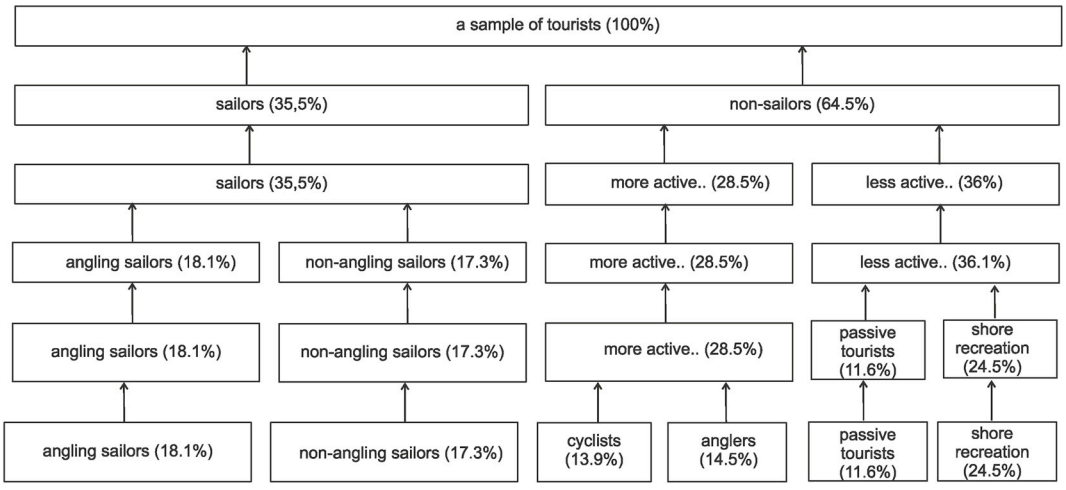

Fig. 5. Cluster analysis solutions ranging from two to six clusters.

Table 2
Clusters as a function of demographics and trip features (%).

Characteristics	Sample profile	Clusters					
		Angling sailors	Non-angling sailors	Cyclists	Anglers	Water recreationists	Passive tourists
Gender							
Male	56.4	74.7	60.9	48.6	58.9	48.0	44.8
Female	42.6	25.3	39.1	51.4	37.0	50.4	55.2
Missing values	1.0				4.1	1.6	
Age							
18–24 years	12.5	7.7	10.3	22.9	12.3	13.0	10.3
25–34 years	29.1	47.3	27.6	20.0	27.4	23.6	27.6
35–44 years	26.5	27.5	28.7	20.0	20.5	26.8	36.2
45–54 years	13.9	9.9	8.0	14.3	21.9	17.1	12.1
55–64 years	12.9	4.4	19.5	18.6	13.7	13.8	6.9
> 64 years	4.2	1.1	4.6	4.3	4.1	4.9	6.9
Missing values	0.8	2.2	1.1			0.8	
Educational level							
Basic	2.8	1.1	1.1	2.9	5.5	3.3	3.4
Professional training	8.4	6.6	2.3	10.0	11.0	9.8	12.1
High school	40.2	33.0	39.1	42.9	45.2	45.5	32.8
University degree	48.2	59.3	56.3	44.3	38.4	40.7	51.7
Missing values	0.4		1.1			0.8	
Place of residence							
Village	8.2	8.8	4.6	8.6	5.5	9.8	12.1
Town < 19 999	12.4	8.8	9.2	8.6	21.9	10.6	19.0
Town 20 000–99 000	24.7	31.9	18.4	21.4	27.4	23.6	25.9
City 100 000–499 999	28.1	36.3	31.0	30	19.2	26.0	24.1
City ≥ 500 000	25.3	14.3	35.6	31.4	23.3	27.6	17.2
Abroad*	1.4		1.1		2.7	2.4	1.7
Length of stay							
1 day	0.2	0	0	0	0	0.8	0
2–3 days	10.2	1.1	13.8	11.4	6.8	13.0	15.5
4–7 days	37.6	14.3	39.1	37.1	47.9	45.5	43.1
8–14 days	37.5	64.8	34.5	27.1	34.2	29.3	32.8
15–30 days	9.2	19.9	4.6	12.9	4.1	7.3	5.2
> 30 days	4.8	0	5.7	11.4	6.8	4.1	1.7
Don't know	0.6	0	2.3	0	0	0	1.7
Visitor type							
First-time visitor	8.8	1.1	2.3	17.1	11.0	10.6	13.8
Repeat visitor (2–5)	4.0	2.2	3.4	7.1	2.7	4.1	5.2
Regular visitor (> 6)	86.7	94.5	94.3	74.3	86.3	85.4	81.0
Missing values	0.6	2.2		1.4			
Number of people you are travelling with							
Alone (1)	4.6	2.2	5.7	4.3	6.8	2.4	8.6
2	26.5	9.9	32.2	42.9	21.9	29.3	24.1
3–4	45.0	64.8	40.2	27.1	50.7	42.3	41.4
5–6	15.0	15.4	10.3	12.9	11.0	21.1	15.5
More than 6	5.8	3.3	9.2	5.7	9.6	4.1	3.4
Missing values	3.2	4.4	2.3	7.1	0	0.8	6.9
Travelling with children							
Yes	23.5	24.2	20.7	14.3	26.0	30.9	19.0
No	76.5	75.8	79.3	85.7	74.0	69.1	81.0

* Only Polish-speaking tourists were surveyed. Seven lived abroad: 4 in Germany, 1 in Austria, 1 in Belgium and 1 in Denmark.
Note: Errors in percentages are due to rounding.

Cyclists stay close to the largest lakes, exploring the area along the north–south axis. Their dispersion along the east–west axis is the smallest among all groups. The semi-minor axis of cyclists' standard deviational ellipse constitutes only 28% of the semi-major axis. *Water recreationists* are present in all of main locations almost equally. Members of this cluster are also accommodated in towns that are not on the main sailing route. This group has the largest range, and their standard deviational ellipse occupies 46% of the study area. *Passive tourists* congregate principally in towns in the central and northern part of the study area, while they are far less numerous in villages. They only explore 36% of the study area. *Anglers*, in turn, use small lakes much more intensively than other groups and are mostly accommodated off the main tourist trail (except for the two main tourist hubs: Mikołajki and Giżycko). *Anglers* and *non-angling sailors* have similar standard deviational ellipses. The disproportion between semi-axes is smallest, the semi-minor axis comprising around 43% of the major axis, indicating that they are widespread over the region. *Angling sailors* are the only group who are mainly accommodated on lakes or in villages. They principally explore the largest interconnected lakes and tend to stay in the northern part of the sailing route, which is less crowded. At the same time, their range of exploration is limited (36% of the study area).

Average Nearest Neighbour analysis shows that all groups are spatially aggregated, while *angling* and *non-angling sailors* are most aggregated (ratio = 0.26, confidence level 99%). This is likely to be due to the environment: sailing space is principally limited to interconnected large lakes. The ratio of the other four groups is around 0.46 (confidence level 99%), which also indicates a high level of spatial aggregation within clusters. It is much higher than the ratio of all the beds in accommodation facilities, which means that the accommodation facilities are much more concentrated in the study area than the clusters' members.

A deeper analysis of tourists' distribution uses a Multi-Distance Spatial Cluster Analysis based on Ripley's K Function. This finds two spatial behaviours. *Angling sailors* is the only group that is aggregated in all distance ranges, indicating that wherever they are in the study area

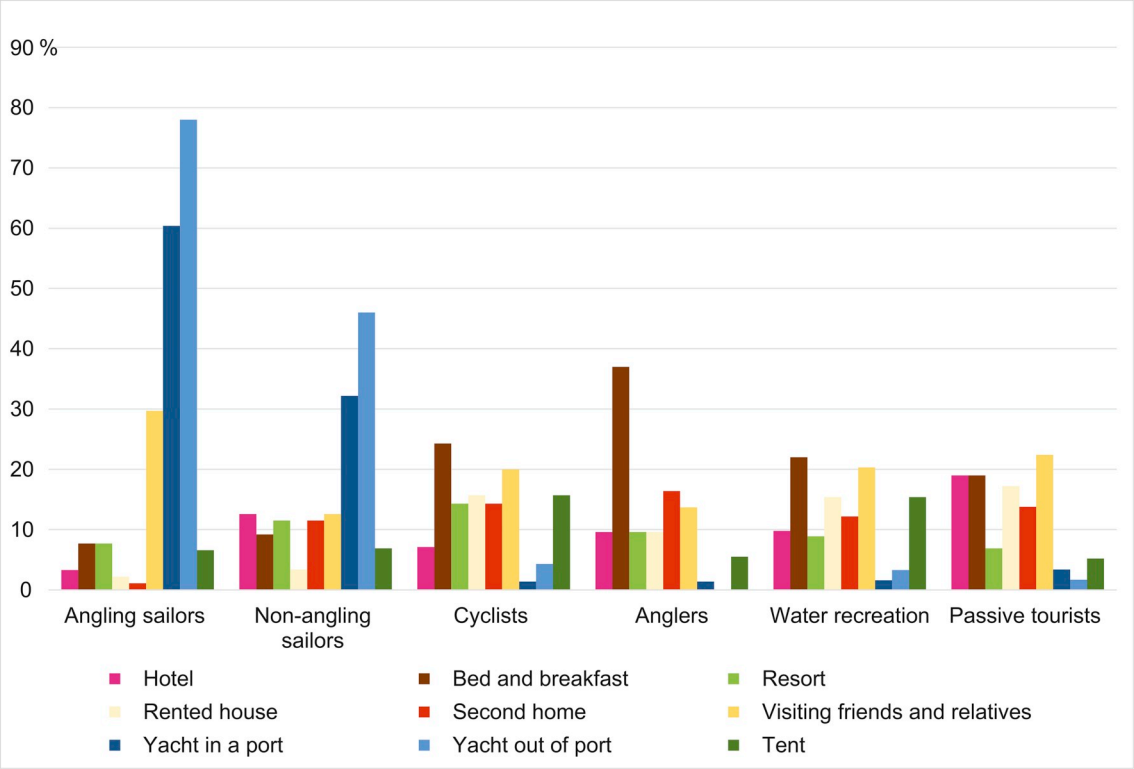

Fig. 6. Respondents' accommodation in the Great Masurian Lakes as a function of cluster (n = 502).

they tend to stay in groups. This is similar to the accommodation facilities, which are also aggregated in all distance ranges. For the other five clusters, tourists are aggregated up to a certain distance from the geometric centre of the class, ranging from 10 km (*non-angling sailors*) to 14 km (*anglers*). Finally, Fig. 9 shows that the spatial behaviour of the latter groups of tourists is random.

4.4.4. Overall cluster characteristics

Overall, clusters differ in terms of their socio-demographic characteristics, travel costs and spatial behaviour. Table 3 presents the main features of each group.

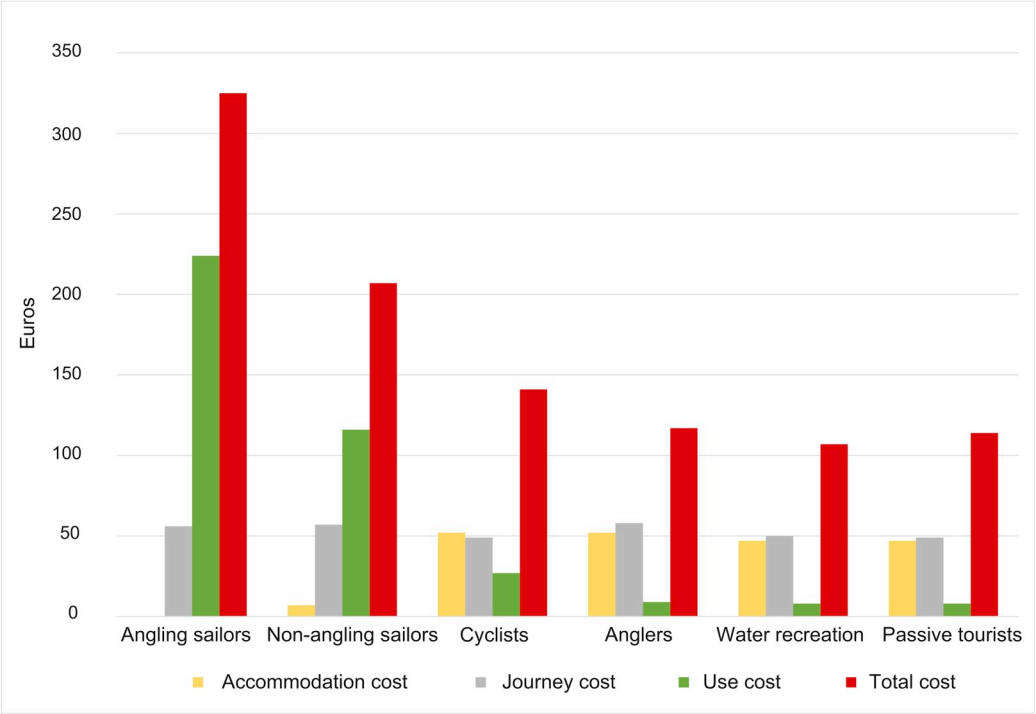

Fig. 7. Individual travel costs by type of cost and cluster (median values in euros).

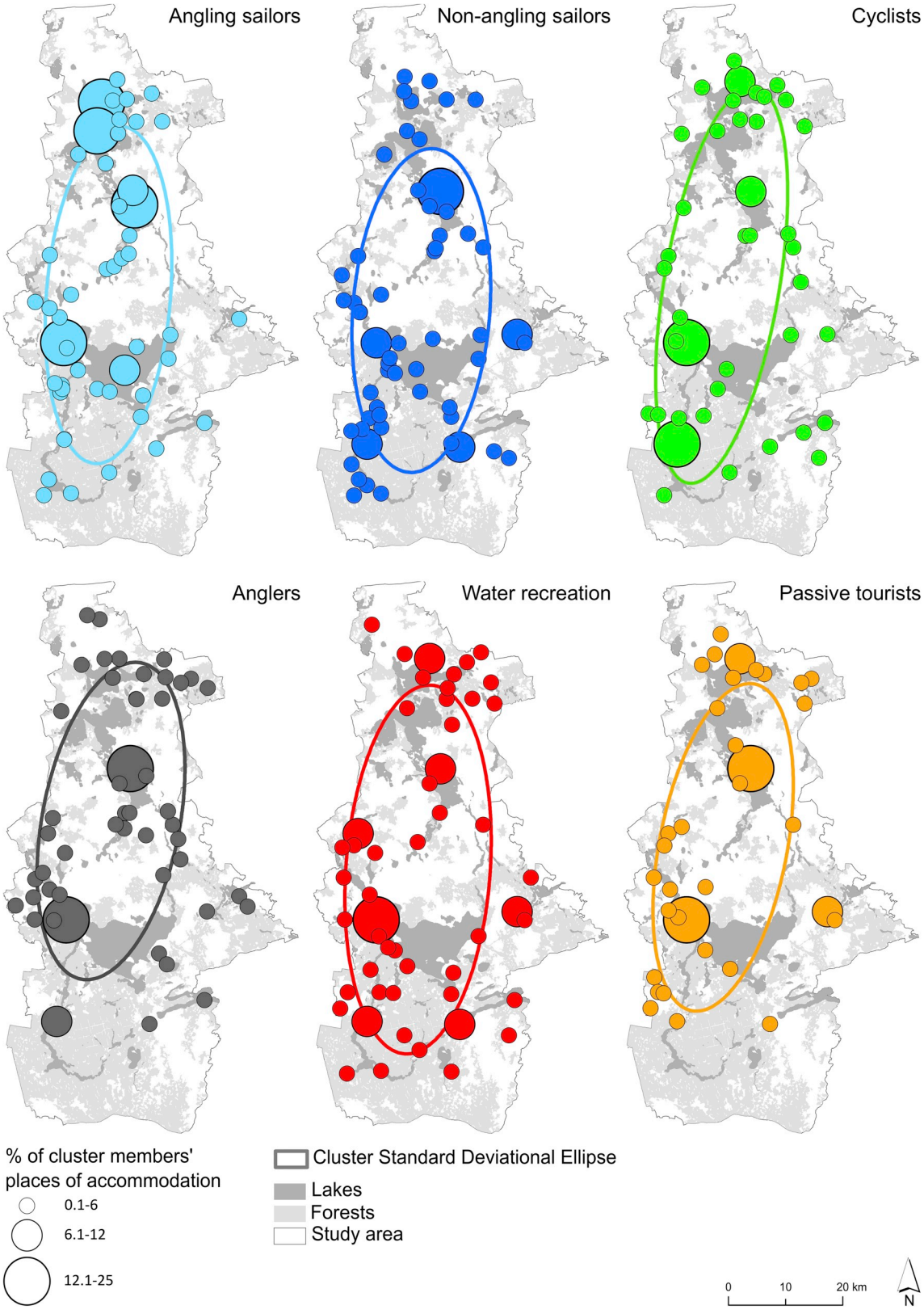

Fig. 8. Distribution of places of accommodation by cluster.

5. Discussion

This study finds that tourists in a nature-based destination can be divided into six groups based on the activities they undertake. Although, obviously, this particular combination of clusters is destination-specific in terms of its environmental and social features, similarities with other groups of users, performing the same activities in other destinations can be found. The most noticeable differences are observed for sailors. Both angling and non-angling sailors are relatively young: most are aged 25–44. This can be compared to the Netherlands,

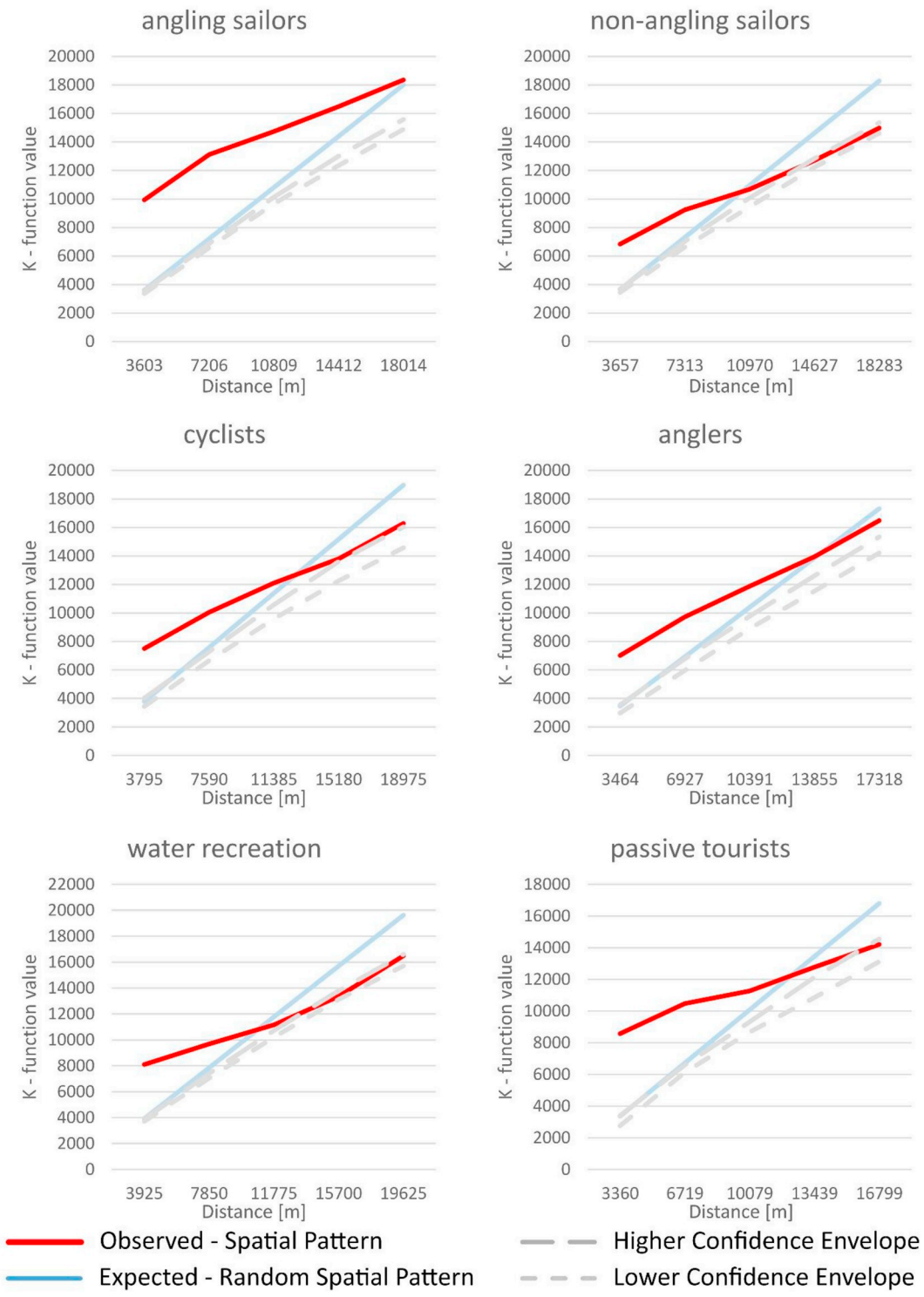

Fig. 9. Spatial distribution of tourists in clusters for distance ranges using Ripley's K-function.

for example, where sailors are significantly older. Goossen (2006) found that almost half of skippers are over 50 years old. In Poland, there are more people on board a yacht (usually three or more) than in the Netherlands (one or two people). Moreover, sailors in Poland stay on board for a shorter period of time (one or two weeks on average), while Dutch skippers tend to make trips lasting an average of 21 days (Goossen, 2006). These differences are economically conditioned: a typical Polish sailor cannot afford to own a yacht. Therefore, yacht rentals are numerous and, consequently, sailing is more affordable than in many other countries, notably for younger and less well-off people.

Many cluster segmentations in this domain have identified a group of passive tourists. Whether they focus on downhill skiing (Konu et al., 2011), rural tourism (Park & Yoon, 2009) or water-based recreation (Ditton et al., 1975), passive tourists do not participate in a wide variety of activities. Our results are consistent with earlier work. In the case of rural tourism in Korea, Park and Yoon (2009) note that passive tourists participate in more classic activities (such as relaxing in nature) and have relatively low expectations. As in our study, they are well educated. Mehmetoglu (2007) identified a 'low-activity oriented' group, whose characteristics are similar to our passive group.

Table 3
The main and the most distinctive characteristics of each of the clusters.

Cluster	Characteristics
Angling sailors	Mostly men with a university degree, aged 25–34. They sail in groups of 3–4 and sleep on yachts moored offshore. They stay in the region for the longest period of time and spend the most money (mainly use costs). They are the only group to cluster spatially on lakes and villages and are the only group that never disperses, even far from the cluster's geographical centre. They are almost exclusively regular visitors.
Non-angling sailors	Hold a university degree, and are from big cities. They spend less time in the area than angling sailors, sail much more often in pairs, and sleep in onshore accommodation a bit more often than other sailors. They cover a large distance to reach their destination. They are highly spatially clustered, mainly in towns. They disperse over a relatively short distance of 10 km. They are almost exclusively regular visitors.
Cyclists	Slightly more women than men, mainly travelling in pairs, with no children. More people aged under 24 and over 55 than average. They have relatively high accommodation costs and sleep in tents or resorts more often than any other group, although bed and breakfast and staying with friends and relatives are most common. Many are first time visitors. They tend to explore the main north–south axis.
Anglers	Most do not have a university degree. They come from small and medium-sized towns, stay 4–7 days in groups of 3–4. They sleep in bed and breakfast accommodation and second homes. They have relatively high accommodation costs. They cover the largest distance to reach the destination. They use small lakes much more intensively than other groups.
Water recreationists	Most do not have a university degree, and stay for 4–7 days. The most children-oriented group. Most stay for a maximum of one week. Their travel costs are the lowest. They are equally present in all of the main locations, forming small and medium-sized clusters that are distributed across the whole study area.
Passive tourists	The most female-dominated group. Aged 35–44 with a university degree. More than average live in small towns and villages. Most stay for a maximum of one week. More visit alone than any other group. Most stay in a hotel or rent a house. Very low individual travel costs. They congregate principally in towns in the central and northern part of the study area.

Our approach shares the main limitation of management-oriented segmentation, in that it tends to only be valid for case studies and unsuited to generalization (Arnegger, Woltering, & Job, 2010). Obviously, the activities undertaken by tourists in the Great Masurian Lakes and the resulting clusters are destination-specific and cannot be automatically generalized to other nature-based destinations, even if it would be reasonable to expect similar behaviours in other lake areas. However, even if the results presented here may only be important for the planning and management of this particular destination, the methods used in this paper are universal.

Another limitation relates to how data was collected regarded places of accommodation and activities. We asked respondents to name a place where they slept during their stay and indicate how often they undertook the listed activities. In the first case, data could have been gathered using tracking methods to improve accuracy. In the second case, data could have been gathered in the form of precise answers (a few times a day, once a day, once a week, etc.) rather than a Likert scale (very often, […], never). Future research could consider these suggestions in order to improve the spatial analysis data.

6. Implications

Although our work has some limitations, the methodology here presented may have important implications for planning and management in nature-based destinations. Our findings show that tourists who perform different activities differ from each other not only in terms of education, age or length of stay, but also in terms of their spatial behaviour and the amount of money they spend on their visit. This knowledge may have very important implications, since different activities have different use requirements, and the needs of tourists engaged in one activity may be at odds with tourists engaged in another. Consequently, it may help to avoid spatial conflicts between different groups of users, as well as avoid environmental overuse.

Our results may help to choose an optimal tourism development model. Knowing the amount of money different groups of users spend on travel, local stakeholders can decide what kind of tourism they want to support and identify their key market. For example, knowing that sailors spend most money and stay in the region for the longest time may influence the decision to support investment in sailing facilities, even if it is inconsistent with the development of other activities.

Second, the definition of key markets can be extremely useful for planning marketing strategies. Managers can differentiate their offer based on cluster profiles. For example, water recreationists often travel with children and are spread across the whole area (both small and large lakes). Therefore, a possible marketing strategy could aim to attract more of them to places with facilities for children, which avoid conflicts with other clusters (such as anglers).

Third, knowing the spatial behaviour of different groups of users may be helpful for planners and tourism businesses, who are interested in identifying the best sites for investment. Examples include places where visitors undertake certain activities but no appropriate facilities are available, such as equipment rental.

Last but not least, the results of this study may be implemented by planners in order to protect the natural environment. Identifying places visited by different groups of tourists may be extremely helpful in managing tourism. Tourists can be discouraged from visiting the most ecologically vulnerable and valuable places and/or they can be encouraged to visit other places where the natural environment is more resilient. Other uses of the results of this study are to educate visitors in the most frequented and environmentally valuable places, or simply imposing environmental regulations. This is especially important in the case of groups who spend their time in places with very limited facilities. Taking the case of sailors, our findings show that many moor their yacht offshore; therefore, education is needed to make them aware of which behaviours are harmful for the environment and why. This is especially challenging, as the results show that angling sailors are the only group that is spatially clustered in each setting. Influencing their behaviour should, however, be easy, given that our findings show that most of them have a high level of education.

7. Conclusions

Activity-based clustering may be a useful and efficient way of developing a typology of tourists (Moscardo et al., 2001). As the outdoor activities they undertake reflect their relationship with nature, clustering tourists by activity in a nature-based destination seems to be a particularly useful way to find out how they use it. Such knowledge may not only serve the tourism industry (Mehmetoglu, 2007), but also contribute to our knowledge of the human–nature relationship. Our study shows that the method works well in practice, as significant differences between clusters were identified.

The continuous growth of tourism in natural areas requires an ongoing search for new methods and instruments to help in assessing how tourists use and value nature. This knowledge is crucial in improving the planning and management of nature-based destinations. Our research not only presents an activity-based segmentation, but also goes one step further by adding spatial analyses and the Travel Cost Method. It demonstrates that combining all of these methods can do more than just segment tourists into relatively homogenous, activity-based groups. Like many other studies, the framework makes it possible to characterize each group by its social and demographic features and travel patterns. However, clusters can also be described by their spatial behaviour and the economic value they give to nature. The findings show

that cluster analysis, which is principally a marketing tool used by companies to segment visitors, may be also an extremely useful tool for planners and destination managers. In this way, the study takes a holistic approach to the analysis of tourism demand in a nature-based destination, and, as such, contributes to the discussion on tourists' use of a nature-based destination.

Declarations of interest

None.

Author contribution

Marta Derek is the chief investigator. She contributed to the design of the survey, performed cluster analysis, and wrote the majority of the article with comments and input from all authors. She also contributed to building the literature review.

Edyta Woźniak contributed critically during the entire study, performed all of the spatial analyses, and contributed to the design of the economic valuation based on the Travel Cost Method. She also helped to write the methods section.

Sylwia Kulczyk contributed to building the literature review, to the design of the survey and to the design of the economic valuation based on the Travel Cost Method.

All Authors discussed the subsequent drafts of the paper.

Acknowledgements

This work was supported by the Narodowe Centrum Nauki, Poland, under grant No. DEC-2012/07/B/HS4/00306 entitled 'The use of ecosystem services concept for the sustainable tourism management – an example of lake areas'.

Appendix

Table A
Respondents' accommodation in the Great Masurian Lakes by clusters (n = 502)

	Sample profile	Angling sailors	Non-angling sailors	Cyclists	Anglers	Water recreationists	Passive tourists
Type of accommodation							
Hotel	9.8	3.3	12.6	7.1	9.6	9.8	19.0
Bed and breakfast	19.3	7.7	9.2	24.3	37.0	22.0	19.0
Resort	9.8	7.7	11.5	14.3	9.6	8.9	6.9
Rented house	10.4	2.2	3.4	15.7	9.6	15.4	17.2
Second home	11.2	1.1	11.5	14.3	16.4	12.2	13.8
Visiting friends and relatives	19.9	29.7	12.6	20.0	13.7	20.3	22.4
Yacht in a port	17.7	60.4	32.2	1.4	1.4	1.6	3.4
Yacht out of port	23.7	78.0	46.0	4.3	0	3.3	1.7
Tent	9.8	6.6	6.9	15.7	5.5	15.4	5.2
Others	0.4	1.1	1.1	1.4	2.7	0	1.7

Note: errors in percentage total due to a possibility of choosing more than one answer.

Table B
Individual travel costs by a type of a cost and by clusters (median values in Euros)

	Angling sailors	Non-angling sailors	Cyclists	Anglers	Water recreationists	Passive tourists
Total cost	325	207	137	117	107	114
Accommodation cost	0	7	52	52	47	47
Journey cost	56	55	49	58	50	49
Use cost	224	112	27	9	8	8

Table C
Distance between and time spent to travel from a place of residence to a place of stay by clusters

	n	Distance (in km; median value)*	Time of travel (in minutes; median values)*
Angling sailors	91	294,5	272
Non-angling sailors	87	311	265
Cyclists	70	265,5	223
Anglers	73	312	267
Water recreationists	123	265	237
Passive tourists	58	266	239

* Estimations using Google maps.

References

Ahas, R., Aasa, A., Yuan, Y., Raubal, M., Smoreda, Z., Liu, Y., et al. (2015). Everyday space–time geographies: Using mobile phone-based sensor data to monitor urban activity in harbin, paris, and tallinn. *International Journal of Geographical Information Science, 29*(11), 2017–2039. https://doi.org/10.1080/13658816.2015.1063151.

Arnegger, J., Woltering, M., & Job, H. (2010). Toward a product-based typology for nature-based tourism: A conceptual framework. *Journal of Sustainable Tourism, 18*(7), 915–928. https://doi.org/10.1080/09669582.2010.485680.

Austin, S. B., Melly, S. J., Sanchez, B. N., Patel, A., Buka, S., & Gortmaker, S. L. (2005). Clustering of fast-food restaurants around schools: A novel application of spatial statistics to the study of food environments. *American Journal of Public Health, 95*(9), 1575–1581.

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis.* Essex: Longman Scientific & Technical.

Beeco, J. A., & Brown, G. (2013). Integrating space, spatial tools, and spatial analysis into the human dimensions of parks and outdoor recreation. *Applied Geography, 38*,

76–85.
Beeco, J. A., Hallo, J. C., & Brownlee, M. T. (2014). GPS visitor tracking and recreation suitability mapping: Tools for understanding and managing visitor use. *Landscape and Urban Planning, 127*, 136–145.
Beh, A., & Bruyere, B. L. (2007). Segmentation by visitor motivation in three Kenyan national reserves. *Tourism Management, 28*(6), 1464–1471.
Bell, S., Tyrväinen, L., Sievänen, T., Pröbstl, U., & Simpson, M. (2007). Outdoor recreation and nature tourism: A european perspective. *Living Reviews in Landscape Research, 1*(2), 1–46.
Birenboim, A., & Shoval, N. (2016). Mobility research in the age of the smartphone. *Annals of the Association of American Geographers, 106*(2), 283–291.
Buijs, A. E., Pedroli, B., & Luginbühl, Y. (2006). From hiking through farmland to farming in a leisure landscape: Changing social perceptions of the European landscape. *Landscape Ecology, 21*(3 SPEC. ISS.), 375–389. http://doi.org/10.1007/s10980-005-5223-2.
Casado-Arzuaga, I., Onaindia, M., Madariaga, I., & Verburg, P. H. (2013). Mapping recreation and aesthetic value of ecosystems in the Bilbao Metropolitan Greenbelt (northern Spain) to support landscape planning. *Landscape Ecology, 29*(8), 1393–1405. http://doi.org/10.1007/s10980-013-9945-2.
Chew, V. (1966). Confidence, prediction, and tolerance regions for the multivariate normal distribution. *Journal of the American Statistical Association, 61*(315), 605–617.
Czajkowski, M., Giergiczny, M., Kronenberg, J., & Tryjanowski, P. (2014). The economic recreational value of a white stork nesting colony: A case of 'stork village' in Poland. *Tourism Management, 40*, 352–360. https://doi.org/10.1016/j.tourman.2013.07.009.
De Aranzabal, I., Schmitz, M. F., & Pineda, F. D. (2009). Integrating landscape analysis and planning: A multiscale approach for oriented management of tourist recreation. *Environmental Management, 44*(5), 938–951.
Derek, M., Woźniak, E., & Kulczyk, S. (2017). Tourism in a nature-based destination: The human versus the ecological perspectives. *Tourism Geographies, 19*(4), 548–574.
Ditton, R. B., Goodale, T. L., & Johnsen, P. K. (1975). A cluster Analysis of activity, frequency, and environment variables to identify water-based recreation types. *Journal of Leisure Research, 7*(4), 282–295.
Dolnicar, S., & Leisch, F. (2003). Winter tourist segments in Austria: Identifying stable vacation styles using bagged clustering techniques. *Journal of Travel Research, 41*(3), 281–292.
Ebdon, D. (1985). *Statistics in geography*. Wiley–Blackwell.
Fairfax, R. J., Dowling, R. M., & Neldner, V. J. (2014). The use of infrared sensors and digital cameras for documenting visitor use patterns: A case study from D'aguilar national park, south-east queensland, Australia. *Current Issues in Tourism, 17*(1), 72–83.
Fredman, P., & Tyrväinen, L. (2010). Frontiers in nature-based tourism. *Scandinavian Journal of Hospitality and Tourism, 10*(3), 177–189.
Goossen, M. (2006). Lake tourism in The Netherlands. In C. M. Hall, & T. Härkönen (Eds.). *Lake tourism. An integrated approach to lacustrine tourism Systems* (pp. 119–130). Channel View Publications.
Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Harlow: Pearson Education Limited.
Hvenegaard, G. T., & Dearden, P. (1998). Ecotourism versus tourism in a Thai national park. *Annals of Tourism Research, 25*(3), 700–720.
Isaacson, M., Shoval, N., Wahl, H. W., Oswald, F., & Auslander, G. (2016). Compliance and data quality in GPS-based studies. *Transportation, 43*(1), 25–36.
Jażdżewska, I. (2018). The use of centrographic measures in analysing the dispersion of historic factories, villas and palaces in łódź, Poland. *Folia Geographica, 60*(1), 50–61.
Kajala, L. (2007). *Visitor monitoring in nature areas: A manual based on experiences from the nordic and baltic countries.* Nordic Council of Ministers.
Konu, H., Laukkanen, T., & Komppula, R. (2011). Using ski destination choice criteria to segment Finnish ski resort customers. *Tourism Management, 32*(5), 1096–1105.
Kotler, P., Bowen, J. T., & Makens, J. C. (2014). *Marketing for hospitality and tourism.* Pearson Education.
Kulczyk, S., Derek, M., & Woźniak, E. (2016). How much is the "wonder of nature" worth? The valuation of tourism in the great masurian lakes using travel cost method. *Ekonomia i Środowisko, 4*(59), 235–249.
Mehmetoglu, M. (2007). Typologising nature-based tourists by activity—theoretical and practical implications. *Tourism Management, 28*(3), 651–660.
Mitchell, A. (2005). *The ESRI guide to GIS analysis, Vol. 2*. ESRI Press.
Mooi, E., & Sarstedt, M. (2011). *A concise guide to market research. The process, data, and methods using IBM SPSS statistics.* Berlin, Heidelberg: Springer Verlag.
Moscardo, G., Pearce, P., & Morrison, A. (2001). Evaluating different bases for market segmentation: A comparison of geographic origin versus activity participation for generating tourist market segments. *Journal of Travel & Tourism Marketing, 10*(1), 29–49.
Newsome, D., Moore, S. A., & Dowling, R. K. (2012). *Natural area tourism: Ecology, impacts and management, Vol. 58*. Channel View Publications.
Palacio, V., & McCool, S. F. (1997). Identifying ecotourists in Belize through benefit segmentation: A preliminary analysis. *Journal of Sustainable Tourism, 5*(3), 234–243.
Park, D. B., & Yoon, Y. S. (2009). Segmentation by motivation in rural tourism: A Korean case study. *Tourism Management, 30*(1), 99–108.
Perera, P., Vlosky, R. P., & Wahala, S. B. (2012). Motivational and behavioral profiling of visitors to forest-based recreational destinations in Sri Lanka. *Asia Pacific Journal of Tourism Research, 17*(4), 451–467.
Prayag, G., Landré, M., & Ryan, R. (2012). Restaurant location in Hamilton, New Zealand: Clustering patterns from 1996 to 2008. *International Journal of Contemporary Hospitality Management, 24*(3), 430–450. https://doi.org/10.1108/09596111211217897.
Ryan, C., & Sterling, L. (2001). Visitors to litchfield national park, Australia: A typology based on behaviours. *Journal of Sustainable Tourism, 9*(1), 61–75. https://doi.org/10.1080/09669580108667389.
Shoval, N., & Isaacson, M. (2007). Tracking tourists in the digital age. *Annals of Tourism Research, 34*(1), 141–159.
Smith, A. J., Tuffin, M., Taplin, R. H., Moore, S. A., & Tonge, J. (2014). Visitor segmentation for a park system using research and managerial judgement. *Journal of Ecotourism, 13*(2–3), 93–109. https://doi.org/10.1080/14724049.2014.963112.
Sæþórsdóttir, A. D. (2010). Planning nature tourism in Iceland based on tourist attitudes. *Tourism Geographies, 12*(1), 25–52. http://doi.org/10.1080/14616680903493639.
Valentine, P. (1992). *Nature-based tourism.* Belhaven Press.
Vasiliadis, C. A., & Kobotis, A. (1999). Spatial analysis – an application of nearest–neighbor analysis to tourism locations in Macedonia. *Tourism Management, 20*(1), 141–148.
Wall Reinius, S. (2011). Researching tourists in the outdoors. Challenges and experiences from protected areas in Sweden. In C. M. Hall (Ed.). *Fieldwork in tourism. Methods, issues and reflections* (pp. 232–239). Routledge.
Wang, B., Wenzhong, S., & Zelang, M. (2015). Confidence analysis of standard deviational ellipse and its extension into higher dimensional euclidean space. *PLoS One, 10*(3), e0118537. https://doi.org/10.1371/journal.pone.0118537.
Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236–244.
Weaver, D. B., & Lawton, L. J. (2002). Overnight ecotourist market segmentation in the Gold Coast hinterland of Australia. *Journal of Travel Research, 40*(3), 270–280.
Weaver, D., & Lawton, L. (2005). Using cluster Analysis to segment a sample of Australian ecotourists. In B. W. Ritchie, P. Burns, & C. Palmer (Eds.). *Tourism research methods. Integrating theory with practice* (pp. 211–220). Wallingford: CABI Publishing.

Marta Derek (Ph.D.) is a geographer, assistant professor in the Faculty of Geography and Regional Studies, University of Warsaw, Poland. Her research and teaching interests include tourism planning and development, tourism and local development, and tourism and regeneration, with a focus on urban areas. Her current research projects focus on the ecosystem services concept in tourism and recreation as well as gastronomy in urban space.

Edyta Woźniak (Ph.D.) is a geographer, assistant professor at the Space Research Centre of the Polish Academy of Sciences, Warsaw, Poland. She is specialized in Geographic Information Systems and Remote Sensing. Her research interests focus on the use of satellite images and spatial analysis for assessment of ecosystem services, natural hazards and land cover changes.

Sylwia Kulczyk (Ph.D.) is an assistant professor in the Faculty of Geography and Regional Studies, University of Warsaw, Poland. She specializes in tourism geography and landscape ecology, and has conducted field studies in various landscape zones. Her research concerns human–nature relations, with a focus on tourism. She links her scientific work with practice through the promotion of sustainable tourism and work with local administrations on development strategies.