THE USAGE OF GEOTAGGED SOCIAL MEDIA PHOTOS AS A METHOD TO MONITOR THE TOURISM ACTIVITY IN THE DANUBE DELTA

Anca - Teodora Bulai[1]
Cristina Zaharof[2]
Octavian Groza[1]
[1] Alexandru Ioan Cuza University of Iaşi, Department of Geography, **Romania**
[2] Cadastre, Geodesy and Cartography Firm, Tulcea, **Romania**

ABSTRACT

The analysis of geotagged photos from social media platforms is one of the latest forms of obtaining spatial information that can be used by scientist and especially by geographers. This new form of data, named crowd sourced geographic information has been used in studies as a monitoring method of the number of visitors in natural parks or to calculate the intensity with which there are frequented some tourist routes.

This paper proposes a methodology that uses geotagged photos to monitor the tourist flows in the Danube Delta based on the information available on Google Earth geobrowser in the form of photos from different locations inside the analyzed area. At the same time we collected information about the content of the photos, which has given us information about the preferences of tourists regarding the landscape.

The method that we used to visualize the patterns created by the location of the photos is Kernel Density Estimator (KDE). This method advantage is that it only needs the location of the photos and it shows if there are clusters/ hot-spots in the distribution and at the same time it predicts the density of the photos starting from existing data. Another method that we used to confirm what we observed through KDE is a grid that covers the analyzed area and we noticed that the grid boxes with the greatest number of photos are overlapping with the areas of high density.

The value of this paper is given by the fact that not only shows an image of the areas that tourists prefer to go to, it explores a complementary method to estimate the number of tourists, beside the number of overnight stays that is available only for the conventional accommodation units, and so it contributes to spatially adjust the tourists attendance in the deltaic space.

Keywords: geotagged photos, volunteered geographic information, Kernel Density Estimator

INTRODUCTION

Volunteered Geographic Information (VGI) is considered to be a significant innovation because ordinary citizens became sensors that send information due to the usage of location-based devices. The advantages of VGI are related to its accessibility and the large amount of information [1], the diversity of fields of application and the experiential and perceptional nature of the content [2]. One of the main sources of VGI is location based social media because it is a source of data which contains spatial and temporal information associated to a specific content like photos or text. Social media is considered

to be a passive source of crowdsourced geographic information because its purpose is different than spatial planning. In this case, the user gets the role of sensor without him knowing, which is why compared to the collaborative mapping, collecting spatial information from social media is considered to be less intrusive, since it does not require extraordinary efforts from the users [3], [4, [5]. Because of its passive nature social media data can capture a different vision on the activities that citizens prefer compared to active data collection campaigns [1]. There are still reservation regarding the use of social media as a complementary data source for spatial planning since it presents only a selective representation of reality, which is why is recommended to compare it to the classical data sources such as census data and surveys [1]. Traditionally, data regarding the number of visitors from an area is collected through surveys. In the countries where are implemented advanced monitoring systems, surveys are carried out systematically, but the impediment is that they are time consuming and expensive, which is why data from social media is about to be considered complementary to classical surveys [6]. Many researchers have already used social media data in tourism studies to quantify the number of visitors and the tourism revenues and to establish visitation patterns of tourists inside natural areas [7], [8], [9], [10]; social media is also used to determine tourist's preferences regarding landscapes [11].

METHODOLOGY

The dataset is composed of 12740 photographs from 7322 locations in the analyzed area, collected manually from the Google Earth geobrowser, the initial source of them being the social media platform Panoramio. The time interval in which the photographs were taken covers a period of 12 years, from 2005 to 2017. The dataset contains the location of the photographs and information about their content (whether it surprises the type of household, the water scenery etc.) which was filled in when the dataset was created.

The dataset permitted us to calculate the density of photographs based on their location using the Kernel Density Tool from the ArcGIS 10.4.1 software. The Kernel Density tool calculates the density of features in a neighborhood around those features. This method was used in finding the density of houses, crime reports, roads or utility lines influencing urbanized areas or wildlife habitat [12]. In this case, the photographs are considered point events that show the presence of tourists in the analyzed area. One of the most popular methods for analyzing points event distribution is Kernel Density Estimator because it is easy to understand and implement [12], especially since some software already have tools that calculate it. In this case, higher values of the density show an increased interest in that area; Kisilevich assumed that the quantitative spatial density of pictures shared online is a measure of spatial preferences [13]. Kernel density estimator can be used on datasets containing only the coordinates of the points but has also the option to weight the features depending on their importance. In this case, the features were weighted by the number of photographs taken in the same location. Because the location of the photographs is not equally distributed in the analyzed area, and there are spatial outliers, it is recommended to use the default value for the bandwidth.

The second method we used is a grid that covers the analyzed area and has the role to verify the accuracy of the results obtained previously and to count the number of pictures from each category of landscape observed in order to see tourists preferences. The grid was created using the Fishnet tool from Data Management toolset, and each cell has an

area of 5 square kilometers. The dimension of the grid cell was selected so that the differences in the spatial distribution can be easily observed. If the grid cell was bigger, the results could not capture the differences in the area. The grid was also used to classify the deltaic space based on the preferences of social media users. In this sense, for each grid cell was counted the number of photos from each category* and the maximum value gives the type of the cell (Figure 1).

*Category	ID	built space	ships	birds	aquatic scenary	boats	fishermen	terrestrial landscapes	tents and caravans
No. of photos	1	15	3	55	20	8	3	5	0
	2	0	0	0	0	0	0	0	0
	3	0	0	35	26	5	2	0	0
	4	0	2	10	15	6	2	0	0

Figure 3. Classification method of tourist's preferences for the photographs at grid level

RESULTS AND DISCUSSION

A first result consists of the localization of the photographs inside the analyzed area. As seen in figure 1, the geotagged photographs are concentrated in some areas, and that can be explained by the existence of accommodation units in the localities or by the location of the navigable ways, that offers better accessibility for tourists. Of course, as observed before many photographs are concentrated on the main branches of the Delta and at the limit between the delta and the sea. Because one of the goals of this paper is to prove that geotagged photographs are a complementary method to estimate tourist flows, we compared the existing statistical data with this cartogram. In figure 3, we represented the number of incoming tourists registered in the accommodation units, in the analyzed period of time, 2005-2016. The data is aggregated on NUTS3 level (communes). We observed that the communes represented in the graph, meaning that they have registered tourists across the entire time period, have also a great density of geotagged photographs. What is interesting is the fact that there are some communes that don't appear on the graph but have a great density of points. This means that the tourists, who have taken photos in those areas, were not registered. This is the case of Pardina, Ceatalchioi, Chilia Veche and C.A.Rosetti. If it were to trust the statistical data for those localities, it would appear that in the area there is no touristic activity. This can be explained by the fact that, even though in the area there are no accommodation units, there is a program of daily trips, so that is how tourists get there. There are three explanations for the concentration of geotagged photographs. The first one is the proximity of the accommodation units. There are some areas of great density situated in the localities that have many accommodation units, as the case of Sfântu Gheorghe, Crişan, Mila 23 and Murighiol. High densities of photos are observed across the navigable routes inside the Delta and on the main lakes. This is possible because there are a number of touristic routes that the locals promote as

day trips. Other areas with great density of geotagged photographs are some points of interest in the Danube Delta, which everyone who comes here wants to see (such as Letea Forest). There are even trips specially created for photographers.
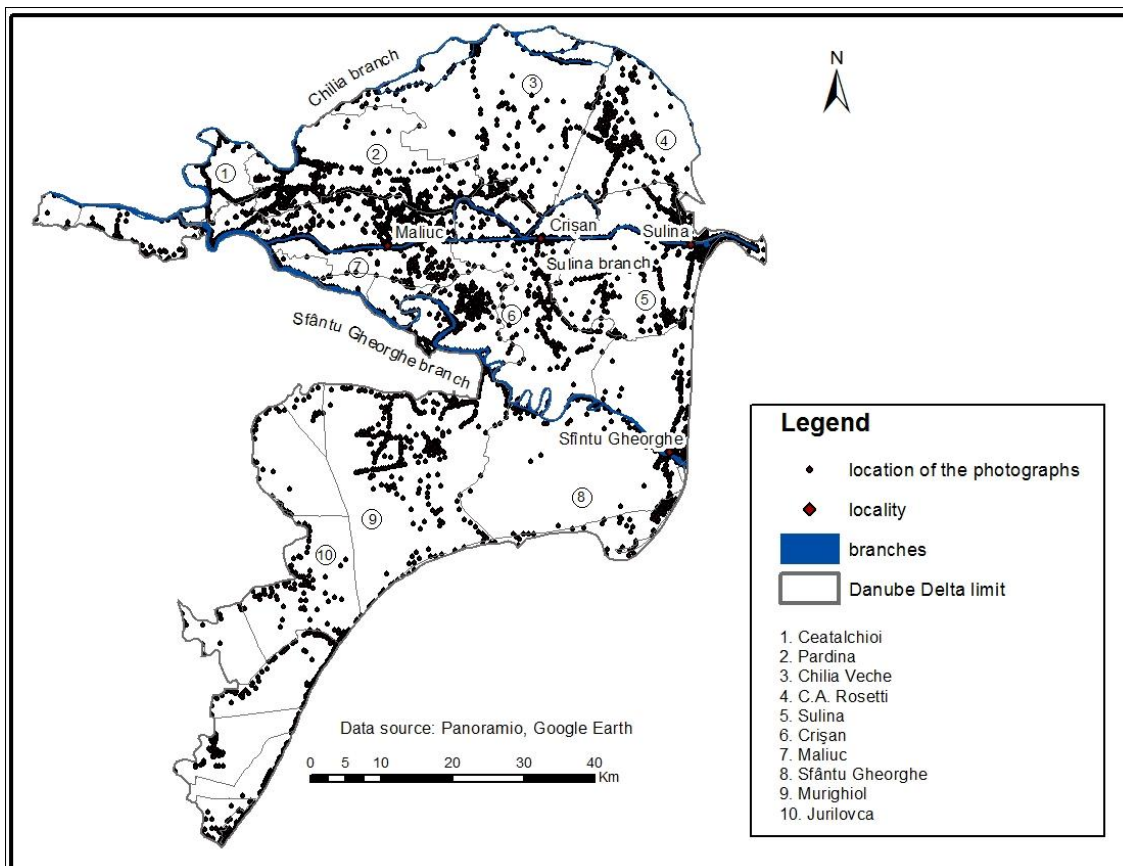

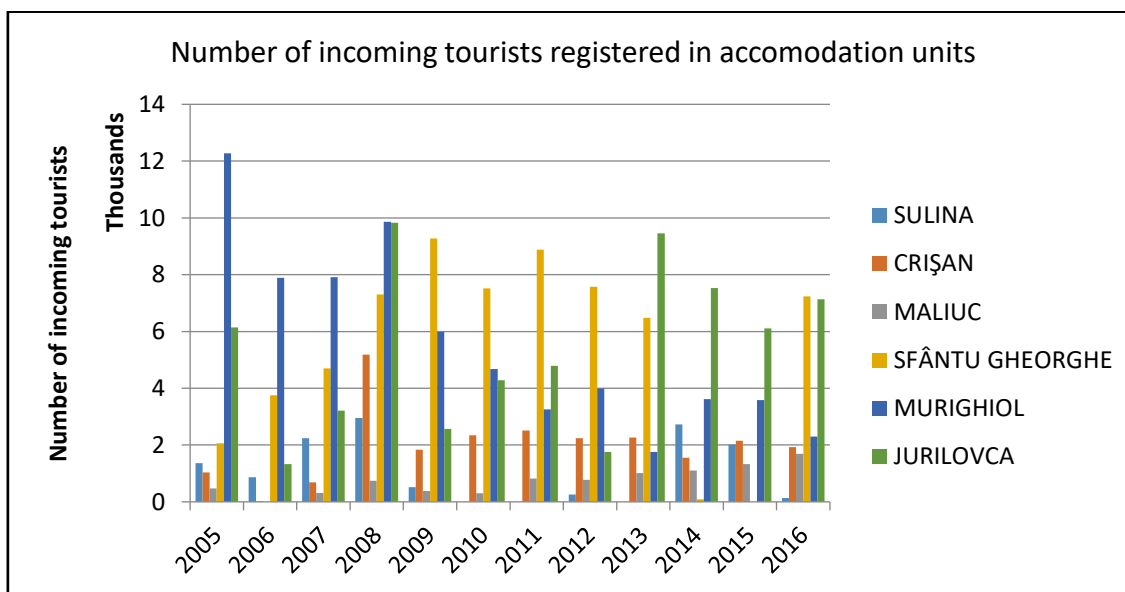**Figure 4.** Geolocated photographs in the period 2005-2016

Figure 5. Number of tourists registered from 2005 to 2016

The cartogram presented in figure 4 represents the kernel density of the geotagged photographs, and this map was necessary because some of the photographs are

overlapping, so the density is higher in those areas. This map shows a clearer view on the patterns already outlined in figure 2, the highest density being recorded in the areas well known and that register a great number of incoming tourists, on one hand, due to a well-preserved image (such as Sfântu Gheorghe and Sulina) and on the other due to new investments in accommodation units as is the case of Crişan and Murighiol.
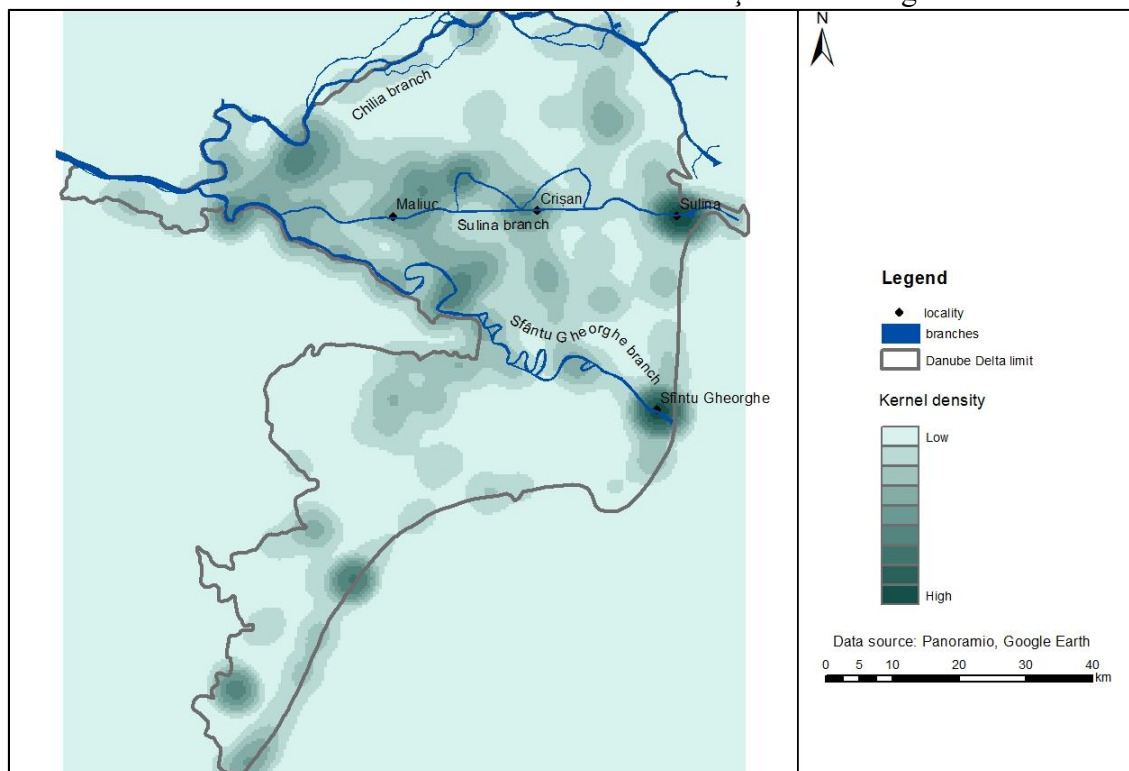

Figure 6. Kernel density of the geolocated photographs

Figure 5 is also a representation of the number of geotagged photographs, but this time they are aggregated in a grid with the dimension of the cell of 5 km square. If kernel density was an approximation of the real values, in this map, each of the grid cells is recording the exact number of photos. As seen below, the patterns observed earlier are maintaining. The highest number of geotagged photographs is registered in Sfântul Gheorghe and Sulina followed by Mila 23, Crişan and Murighiol. If these high values can be explained based on the proximity of the accommodation units, there are areas where the number of photographs can be explained by the fact that tourists are traveling and exploring the navigable channels and the well-known lakes in order to benefit from the aquatic scenery. Another explanation is that people are interested in some specific landscapes and places, as is the case of Letea Forest. This area is very popular among the tourists, so the locals organize daily trips. Figure 6 is a detailed image of the patterns created by the photographs in some areas. In figure 6.2 the pattern created by the photographs shows the areas of origins of the tourists: some of them are coming to the forest from Crişan and Mila 23 and others are coming from Sulina. The fact that these 2 locations are closer to Letea Forest is another advantage of the tourists: they can take boat trips in the surrounding area, come to see the forest, and in the night they can return at the hotels and benefit from the comfort they offer.
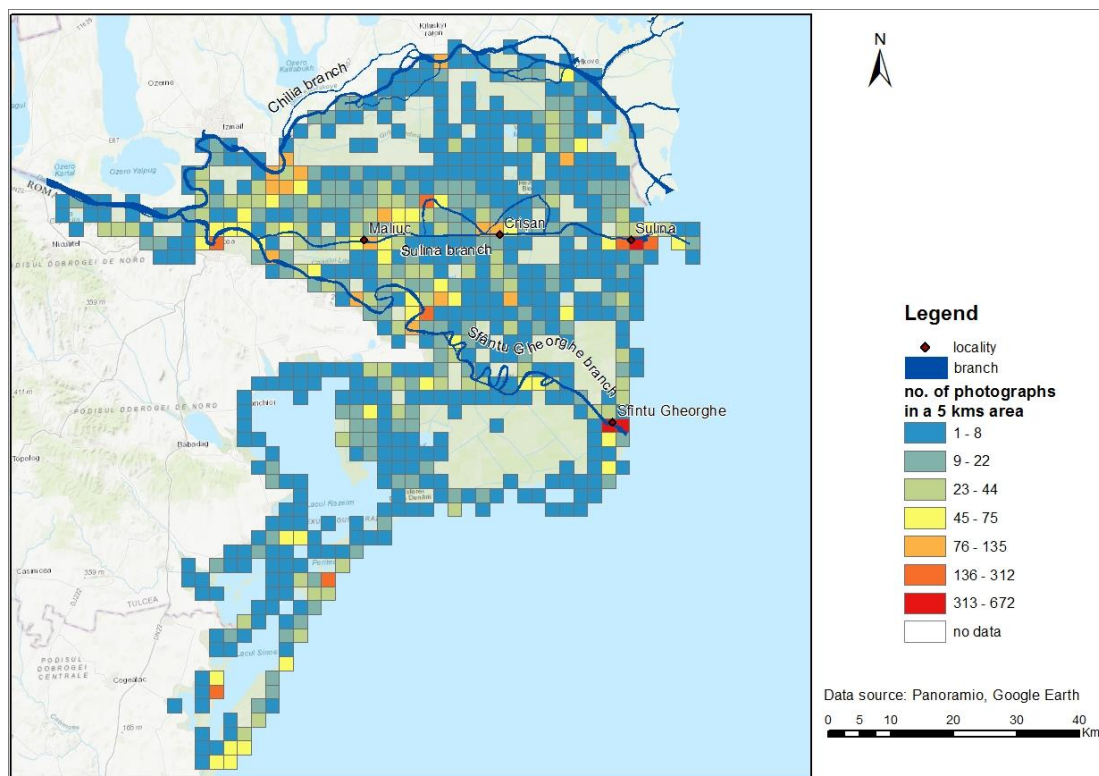
Figure 7. Number of photograph aggregated on a grid with the area of a cell of 5 km square

In figure 6.2 is presented another situation of clustered photographs. This time the proximity of Chilia branch and the channels that connect the lakes make this area captivating and accessible for tourists. The last example shows an alignment of geotagged photographs on the shoreline, and the interest in this area comes from the beauty of the Delta entering the sea: on one side the interest is at the sea and on the other at the lagoons. The specifics of this area are given by the beaches Vadu and Corbu, which are less affected by tourist facilities; the tourists that come here are camping in the attempt to not destroy the natural environment. This affirmation is confirmed by the cartogram presented in figure 7, which shows the dominant type of photograph for each cell of the grid, and the only photographs that captured tents and caravans are located in this area. The preferences of tourists regarding the landscape they photograph are concentrated on aquatic scenery (but this category includes specific vegetation, lakes covered by lotus flowers, sunset and sunrise etc.), but there are also some clustered areas where the element captured in the photographs are the birds. Another type of landscape photographed is generally named terrestrial landscape and includes pictures from the villages (local housing and surroundings) or from places of interest like forests or hills. The elements that could not be missed by the cameras are the fishermen boats since fishing is an important activity of the locals.
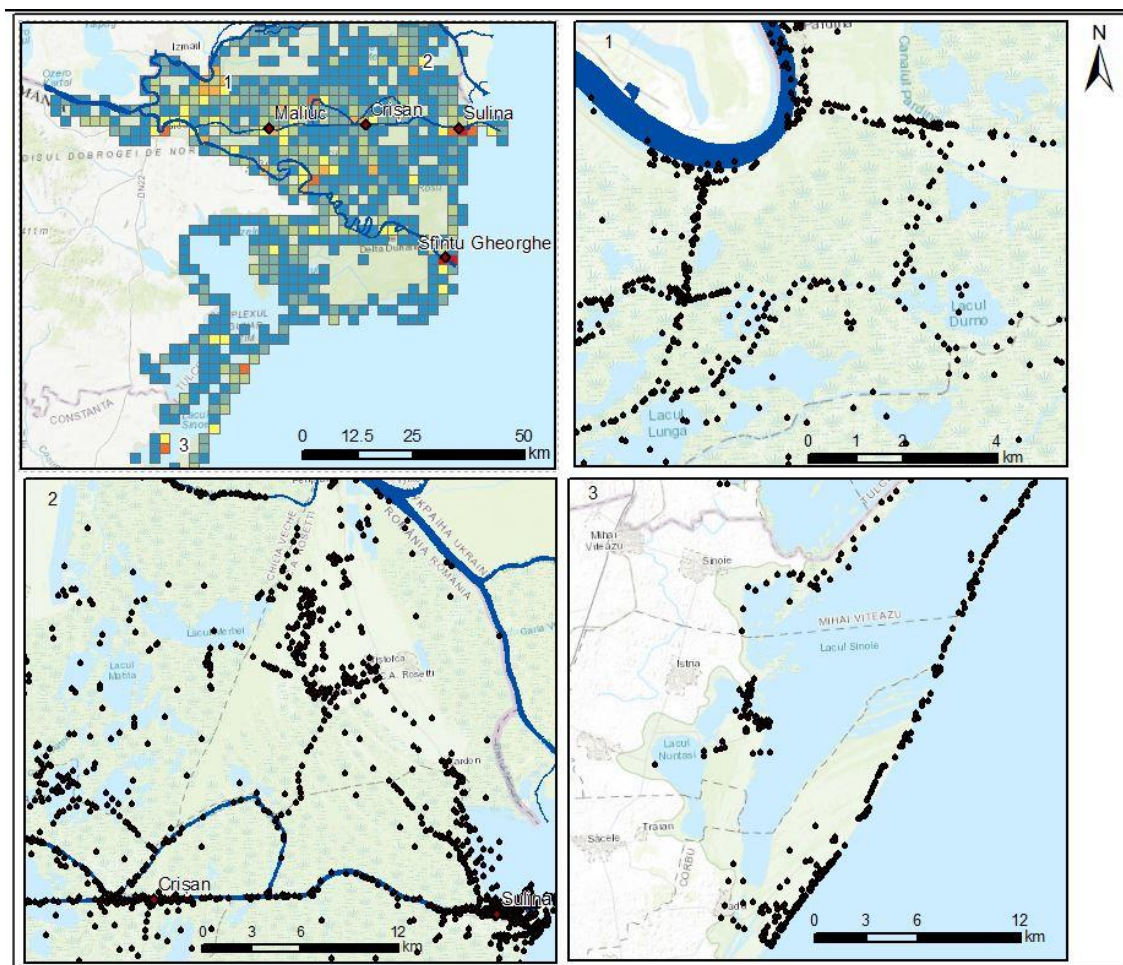
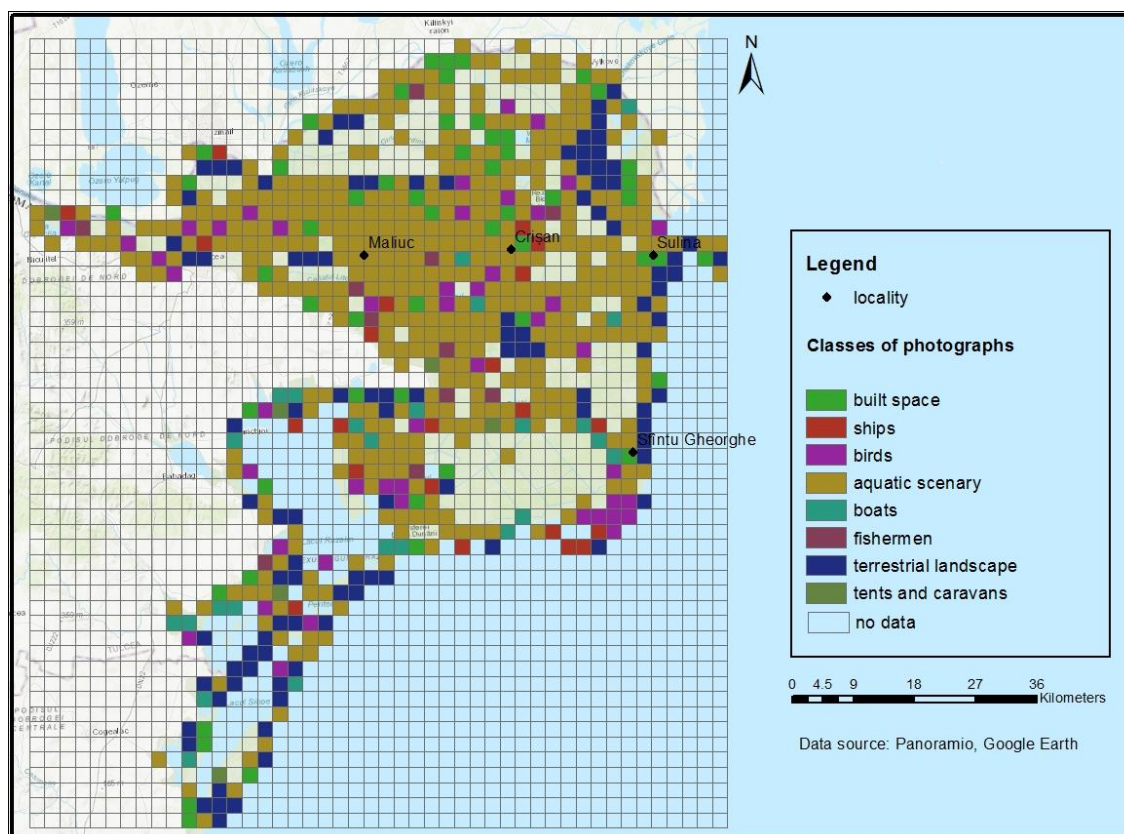**Figure 8.** Detailed situation for some areas

Figure 9. Classes of photographs in the analyzed area

CONCLUSIONS:

The usage of the geotagged photographs are of great help in monitoring the activity of tourists especially in large areas like the Danube Delta where there are no other methods of estimating the number of people who come to visit except for the incoming logs from the accommodation units. The advantage is that the location of the photographs provides the spatial information necessary in establishing the patterns created by tourists and the quantity of shared information offers an image of the locations they want to see. This can be helpful for the locals in order to create touristic packages adapted for their guests and for the local authorities, by monitoring tourists' flows and preventing them from entering in the areas strictly protected.

REFERENCES:

[1] Heikinheimo V & Di Minin E & Tenkanen H & Hausmann A & Erkkonen J & Toivonen T, User-Generated Geographic Information for Visitor Monitoring in a National Park: A comparison of Social Media Data and Visitor Survey, International Journal of Geo-Information, 6, 85, 2017. http://dx.doi.org/10.3390/ijgi6030085

[2] Capineri C, The nature of Volunteered Geographic Information, European Handbook of Crowdsourced Geographic Information, UK, Pp 15-33, 2016. http://dx.doi.org/10.5334/bax

[3] See L & Mooney P & Foody G & Bastin L & Comber A & Estima J & Fritz S & Kerle N & Jiang B & Laasko M et al., Crowdsourcing, Citizen Science or Volunteered Geographic

Information? The Current State of Crowdsourced Geographic Information, International Journal of Geo-Information, 5, 55, 2016 http://dx.doi.org/10.3390/ijgi5050055

[4] Goodchild M. F, Citizens as sensors: The world of volunteered geography, GeoJournal, 69, Pp 211-221, 2007. https://doi.org/10.1007/s10708-007-9111-y

[5] Newman G & Wiggings A & Crall A & Graham E & Newman S & Crowston K, The future of citizen science: Emerging technologies and shifting paradigms, Frontiers in Ecology and the Environment, 10, Pp 298-304, 2012. http://dx.doi.org/10.1890/110294

[6] Di Minin E & Tenkanen H & Toivonen T, Prospects and challenges for social media data in conservation science, Frontiers in Environmental Science, 3, 2015. https://doi.org/10.3389/fenvs.2015.00063

[7] Levin N & Kark S & Crandall D, Where have all the people gone? Enhancing global conservation using night lights and social media, Ecological Applications, 25, Pp 2153-2167, 2015. http://dx.doi.org/10.1890/15-0113.1

[8] Wood S.A & Guerry A.D & Silver J.M & Lacayo M, Using social media to quantify nature-based tourism and recreation, Scientific Reports, 3, 2976, 2013. : http://dx.doi.org/10.1038/srep02976

[9] Sonter L.J & Watson K.B & Wood S.A & Ricketts T.H, Spatial and Temporal Dynamics and Value of Nature – Based Recreation, Estimated via Social Media, Plos One, 11, 2016. https://doi.org/10.1371/journal.pone.0162372

[10] Orsi F & Geneletti D, Using geotagged photographs and GIS Analysis to estimate visitor flows in natural areas, Journal for Nature Conservation, 21, 5, Pp 359-368, 2013. https://doi.org/10.1016/j.jnc.2013.03.001

[11] Van Zanten B.T & Van Berkel D.B & Meentemeyer R.K & Smith J.W & Tieskens K.F & Verburg P.H, Continental – scale quantification of landscape values using social-media data, Proceedings of the National Academy of Sciences of the United States of America, 113, Pp 12974-12979, 2016. https://dx.doi.org/10.1073/pnas.1614158113

[12] Silverman, B. W. (1986). Density estimation for statistics and data analysis. London: Chapman Hall.

[13] Kisilevich, S.& M. Krstajec & D. Keim & N. Andrienko & G. Andrienko. Event-based analysis of people's activities and behaviour using Flickr and Panoramio geotagged photo collections. [book auth.] S. Bertschi, J. Counsell, M. Dastbaz (eds.) E. Banissi. Information visualisation. Los Alamitos : IEEE Computer Society, 2010.