A Spatial Analysis of Tourism Activity in Romania

Daniela Luminita Constantin, Adriana Elena Reveiu (Dardala)
Bucharest University of Economic Studies

Abstract. Location is a key concept in tourism sector analysis, given the dependence of this activity on the natural, built, cultural and social characteristics of a certain territory. As a result, the tourist zoning is an important instrument for delimiting tourist areas in accordance with multiple criteria, so as to lay the foundations for finding the most suitable solutions of turning to good account the resources in this field. The modern approaches proposed in this paper use a series of analytical tools that combine GIS and spatial agglomeration analysis based techniques. They can be also employed in order to examine and explain the differences between tourist zones (and sub-zones) in terms of economic and social results and thus to suggest realistic ways to improve the efficiency and effectiveness of tourist activities in various geographical areas.

In the described context this paper proposes an interdisciplinary perspective (spatial statistics and Geographical Information Systems) for analysing the tourism activity in Romania, mainly aiming to identify the agglomerations of companies acting in this industry and assess their performance and contribution to the economic development of the corresponding regions. It also intends to contribute to a better understanding of the way in which tourism related business activities develop, in order to enhance appropriate support networks. Territorial and spatial statistics, as well as GIS based analyses are applied, using data about all companies acting in tourism industry in Romania provided by the National Authority for Tourism as well as data from the Environmental Systems Research Institute (ESRI).

Keywords: tourism activity, geostatistics, GIS based analysis, spatial autocorrelation, Moran’s index
JEL Classification: C19, C88, L83, R12
Introduction

The Travel and Tourism Competitiveness Report 2013 issued by the World Economic Forum indicates for Romania a direct contribution of tourism of 1.5% to GDP and 2.3% to the total employment. If total effects (direct and indirect) are taken into consideration, the contribution is higher, namely 4.7% to GDP and 5.3% to total employment (WEF, 2013). It confirms tourism’s capacity to generate important income and employment multiplier effects through the activity of traditional service providers and industry suppliers.

However, Romania’s tourism competitiveness is far behind its big potential: the same report shows that it ranks only 68 out of 140 countries considered, almost all other countries from Central and Eastern Europe displaying better ranks. If the tourism competitiveness pillars are examined, Romania presents competitive advantages with regard to tourism infrastructure units – rank 34), health and hygiene (rank 54), environmental sustainability (rank 58), ICT infrastructure (rank 59), safety and security (rank 63). The drawbacks are recorded in ground transport infrastructure (rank 109), prioritization in travel and tourism (rank 103), air transport infrastructure (rank 93), price competitiveness (rank 84), etc.

As a result, various EU funded programmes for 2014-2020 incorporate priorities regarding tourism development: the Regional Operational Programme, the Economic Competitiveness Programme, the National Programme for Rural Development, their denominator being the regional dimension of tourism development. From this viewpoint Romania is characterized by a relatively well-balanced spatial distribution of its natural and cultural-historic landscapes, making it possible to address tourism as a solution for boosting the development of lagging behind regions (Constantin and Mitrut, 2009).

Based on these overall considerations this paper proposes the use of geographical information system (GIS) tools and spatial statistical models in order to investigate the spatial associations of territorial units with significant tourism activity, by examining the spatial relationships between accommodation companies/units (hotels, motels, pensions, etc.) and foodservice companies/units (restaurants, fast food chains, cafés, etc.). It aims to reveal their distribution and resulted spatial agglomerations as a background for rational decisions regarding the support that will be offered to the most relevant tourism destinations as well as the measures meant to enhance collaborative networks in the tourism activity-based agglomerations.

A brief literature survey shows that GIS have been used to develop a lot of applications for tourism, seeking to analyse the regional specific information (Poslad et al., 2001). The approaches
employed are spatial decision support applications and spatial statistics support applications. The former propose GIS based solutions specifically designed to identify spatial relationships to integrate tourism specific information like tourist characteristics (Lau and McKercher, 2006), landscape elements and tourist locations (Brown, 2006), temporal–spatial behavior (Shoval et al., 2011), and the images added to these locations (Gaughan et al., 2009).

One the research mainstreams identified in the literature is the empirical analysis of the distribution of tourism-related activities, such as selected attractions, supporting facilities and accommodation in general (Pearce, 1995).

When presence of dependence in spatial and temporal data is examined, the classical assumption of classic statistical models (e.g. OLS) is violated, the use of special techniques being required. Usually the dependence testing is done by means of autocorrelation analysis. Autocorrelation is “the cross-correlation of a signal with itself” (Cheng et al., 2014, p.1176) and, in case of spatial data, it can be measured using an index, most frequently the Moran index.

As far as the explanations for location choice and spatial distribution of companies providing accommodation and foodservices are concerned, the main research approach employed in recent studies consists of regression methods, based on classical economic theory (Zhang et al., 2012 and Yang et al., 2014). Usually the explanatory variables used in the regression models are relating to labour, culture, capital, and policy characteristics.

As pointed out by Albert et al. (2014) and Seul (2015), tourism related activities in accommodation and foodservices compete with neighbours of similar quality, rather than those who are differentiated in term of quality. The research results suggest that accommodation companies are usually highly clustered, in order to obtain benefits from agglomeration effects. When there is a particular interest in the identification of local clusters (hot spots) of cases the phenomenon is named local heterogeneity (Haining R., 2014).

In the described context, as the first step of a larger exploratory research, this paper aims to identify – in methodology terms - the significant spatial associations of territorial units in terms of two relevant indicators for tourist activity.

**Research Methodology and Support Data**

The paper first introduces the spatial characteristics of accommodation companies and of foodservice companies acting in Romania. The spatial relationships are then analysed with a set of spatial statistics and GIS based models.
Frequency maps, spatial autocorrelation approach, global and local spatial autocorrelation testing are used to identify the nature of spatial distribution of tourism activity performed in Romania.

A. Data Sets

In order to perform the proposed analysis, three data sets are employed. Two data sets contain public data about all companies acting in tourism industry in Romania in December 2014. The first data set includes information about 7157 classified foodservice companies and the second data set contains information about 10007 classified accommodation companies. The data source for both data sets is the Romanian National Authority for Tourism. Both data sets are processed in order to aggregate data at LAU-2 level\(^1\).

For analysing the distribution of economic activity developed in tourism sector, the lowest level of aggregation for data about companies (LAU-2) is envisaged. Even if the statistical data about economic activity in tourism is not available at this aggregation level, a higher level of aggregation for this kind of analysis is not appropriate. Because the tourism activity is deeply influenced by the environmental features this sort of analysis performed at county, regional or macro-regional level of aggregation for economic data could not be considered appropriate.

Another data set employed in this research contains spatial data about Romania - data provided by Environmental Systems Research Institute (ESRI).

Both economic and spatial data are integrated and stored in a spatial database - a geodatabase - and are managed by means of Geographic Information System (ArcGIS), from ESRI.

B. Spatial Distribution of Companies Acting in Tourism Sector

First of all, aggregate data about accommodation and foodservice companies are distributed in geographic territory, according to the number of companies acting in the specified field.

C. Spatial Autocorrelation Analysis

---

\(^1\) Local Administrative Units
In order to identify significant spatial associations of LAUs in terms of number of accommodation and foodservice companies the spatial autocorrelation analysis (univariate and bivariate) is envisaged.

There are two types of spatial autocorrelations: positive and negative. Positive spatial autocorrelation occurs when LAUs with high or low values of a variable tend to group together (“spatial clusters”), and negative spatial autocorrelation appear when LAUs with high values are surrounded by LAUs with low values or vice-versa (Anselin, 2002; Dominicis, 2007; Goschin, 2015).

Spatial autocorrelation can be interpreted in various ways. For example, it can be seen as self-correlation which appears in 2-D space. Unlike the traditional Pearson correlation coefficient, which measures the co-viability of paired values in two variables, spatial autocorrelation measures “correlation among paired values of a single variable based on relative spatial locations” (Griffith and Chun., 2014, pp.1478-1479). As it concentrates on a tendency among values of a variable based on their spatial closeness, spatial autocorrelation “is measured within the combinatorial context of all possible pairs of observed values for a given variable where corresponding weights that are determined by spatial closeness identify the pairings of interest” (Griffith and Chun, 2014, p.1479).

Another interpretation of spatial autocorrelation is as a map pattern. Regional science operates with datasets of individual observations post-stratified by geographical unit such as census blocks/block groups, county boundaries, county borders. When such areal units are used, the choropleth mapping of a variable portrays a pattern over space. A tendency towards similarity or dissimilarity for neighbouring values on such a map can be directly taken as spatial autocorrelation. Whereas large clusters of similar values on the map indicate positive spatial autocorrelation, when the tendency is for values to be dissimilar compared to those of their neighbours it can be interpreted as negative spatial autocorrelation (Griffith and Chun, 2014).

Various studies in regional science attempted to numerically quantify the spatial autocorrelation. The most frequently used quantitative measure of spatial autocorrelation is the Moran coefficient, as the analogous of the Pearson’s correlation coefficient. (Griffith and Chun, 2014). In addition, various local indicators of spatial association have been proposed, such as a local variant of Moran’s coefficient, Getis and Ord’s Gj and Gj statistics, which show to what extent high and low values are clustered together.
From the available statistics, as index designed to measure spatial autocorrelation, the local Moran's I has been chosen.

Local Moran's I has been used in order to detect the local agglomerations of companies providing accommodation and foodservices for Romania:

\[ I_i = z_i \sum_j W_{ij} z_j \]

where:
- \( z_i \) and \( z_j \) are standardized scores of attribute values for administrative unit \( i \) and \( j \);
- \( j \) is among the identified neighbourhood of \( i \), according to the weights matrix \( w_{ij} \) (Anselin, 1995).

Local Moran statistic, together with local Gamma statistic, local Geary statistics, Moran scatterplot, etc. are relevant examples of Local Indicators of Spatial Association (LISA), which are employed in order to assess the spatial association at a location (Cheng et al., 2014), making it possible to identify local spatial clusters and to assess local instability. LISA is the most frequently employed technique for exploratory spatial data analysis (ESDA), applications being found in regional science, spatial econometrics, social sciences, etc. (Symanzik, 2014).

When the ESDA techniques are discussed in the GIS context, the aim is to explore the spatial nature of the envisaged data. These techniques can be grouped into techniques based on the neighbourhood view of spatial association (e.g. Moran scatterplots and LISA statistics) and techniques based on the distance view of spatial association (e.g. lagged scatterplots, variogram-cloud plots) (Anselin, 1995).

In order to identify significant spatial associations of LAUs for each of the two variables the significance map has been created (allowing the identifying of locations with significant local Moran statistic).

Besides the spatial auto-correlation for a given variable (number of accommodation units or number of foodservice units), cross-corellation between one variable and another has been been also analysed. In this case the bivariate spatial autocorrelation analysis has been applied, using the bivariate Local Moran’s I:

\[ I_{kl}^i = z_k^i \sum_j w_{ij} z_l^i \]

where \( k \) and \( l \) are the two variables considered (Anselin et al., 2002).
Thus, the types of spatial autocorrelation (positive and negative) for the two data sets could be identified for both cases – univariate and, respectively, bivariate analysis.

Further on, the classes of spatial associations for each of the two types of autocorrelation have been highlighted (two classes for each type) by means of different colour codes.

**Results**

*Spatial distribution of companies acting in tourism sector of Romania*

Most of the accommodation companies are distributed in Romania’s mountain areas and Black Sea region, as is presented in Figure 1. The most important localities are Bucuresti, Eforie, Costinesti, Brasov, Busteni, Constanta, Mangalia, Moieciu, Bran, Baile Felix, Predeal, Sibiu, Cluj-Napoca, Sinaia, Timisoara and Navodari.

![Figure 1. Spatial distribution of accommodation companies](image)

*Source:* authors’ construction, with ArcGIS software using data provided by Romanian National Authority for Tourism

The top of localities with the largest number of companies acting in servicefood industry shows the following ranking: Constanta, Cluj-Napoca, Arad, Mangalia, Bucuresti, Brasov, Timisoara, Predeal, Sinaia. The distribution of all companies acting in public food sector is presented in Figure 2.
Spatial Autocorrelation Analysis

A. Univariate spatial correlation

The map of locations that have a significant Moran statistic (for p-values below 0.05 and 0.01) corresponding to the “number of accommodation companies” variable is presented in Figure 3.
The significance map shows the locations with a significant Local Moran statistic, by using different shades of green depending on the p-value\(^2\).

In the first category, for \( p = 0.05 \) are included 214 LAUs (light green), and in the second category, for \( p = 0.01 \) are 774 LAUs (deep green). For the rest of LAUs, the local Moran statistic is not significant.

For the significant associations of LAUs in the case of “number of accommodation companies” variable the map in Figure 4 offers the interpretation of this univariate Local Moran, exploring the type of autocorrelation and the category of spatial association. For the considered variables, all 988 LAUs are included in positive autocorrelation agglomerations, two categories of spatial associations being distinguished: the first (in red), with 198 LAUs, indicates “high-high” similarity based spatial clusters (agglomerations) (each LAU with a high value is surrounded by neighbours with high values too); the second (in blue), with 790 LAUs, indicates “low-low” similarity based spatial associations, with small number of accommodation units. This configuration is a confirmation of the first category incorporating the most important tourist areas in Romania (Black Sea, Delta of Danube, Prahova Valley, Bucovina, etc.), which indicates a “natural clusterization” as a response to the natural and, in some cases, historic and cultural environment advantages rather than the result of a clearly targeted tourism-support policy.

---

\(^2\) P-value is associated with the risk of rejecting the H0 (null hypothesis: random spatial variance or, in other words, there are not spatial associations between neighbour territorial units).
B. Bivariate spatial correlation

Moran’s scatterplot for spatial correlation between the number of accommodation companies and the number of foodservice companies is presented in Figure 5.

The value of Global Moran’s I is 0.07. This value is positive, but very low. Considering this, one can conclude that the pattern indicates a positive spatial autocorrelation between the number of accommodation companies and the number of foodservice companies; though, as the value of Global Moran’s I is very small, the spatial autocorrelation it is not present in all country. Nevertheless, the calculation and analysis of Local Moran’s I (LISA) is recommended in order to see whether there are local spatial associations.

![Figure 5](image)

Figure 5. Moran’s scatter plot for bivariate spatial correlation between the number of accommodation companies and the number of foodservice companies

Source: authors’ construction, using GeoDa software tool and data provided by Romanian National Authority for Tourism

The map of locations with significant bivariate spatial correlation between the number of accommodation companies and the number of foodservice companies (significant bivariate Local Moran statistic), for p-values below 0.05 and 0.01 is presented in Figure 6. In the first category, for
p = 0.05 are included 164 LAUs, and in the second category, for p = 0.01 are 87 LAUs. Large areas with significant bivariate Local Moran statistic are inside the following counties: Brasov, Bucharest, Constanta Timis, Arad, Cluj, Maramures, Sibiu, Iasi, Bihor, Tulcea and Suceava, many of them including the most important tourist attractions in Romania.

![Figure 6. Moran significance map for the cross-correlation between the number of accommodation companies and number of foodservice companies](image)

**Source:** authors’ construction, using GeoDa software tool and data provided by Romanian National Authority for Tourism

The map with significant spatial associations of LAUs for the cross-correlation between the number of accommodation companies and the number of foodservice companies is presented in Figure 7. It reveals two classes (categories) of positive spatial correlations („high-high” and „low-low”) and two classes of negative spatial correlation („high-low” and „low-high”). Usually the spatial associations corresponding to positive correlations are named „spatial clusters” whereas those corresponding to negative correlations are associated with the „spatial outlier” notion (Anselin et al., 2001).

---

3 This notion is employed in a different acception compared to that of “industrial cluster” (or “tourist cluster”) as defined by M. Porter. However, the existence of spatial associations for high values of the two variables can be the first sign of existence of such clusters. Further investigation would be necessary in order to see their stage of development: incipient, pure agglomerations or mature clusters.
The four colour codes used for representing the four classes (categories) of significant associations of LAUs are as follows:

- **dark red** - for representing LAUs with large number of accommodation companies surrounded by neighbourouring LAUs with large number of foodservice companies too;

- **dark blue** - for representing LAUs with a small number of accommodation companies and surrounded by neighboring LAUs with small number of foodservice companies too;

- **pink** - for LAUs with a large number of accommodation companies (“high outlier in”), but surrounded by neighbouring LAUSs with low number of foodservice companies (“low neighbours in”);

- **light blue** - for LAUs where there is a small number of accommodation companies (“low outlier in”), but surrounded by LAUs with a large number of foodservice companies (“high neighbours in”).

**Figure 7.** Categories of significant spatial associations of LAUs for the cross-correlation between the number of accommodation companies and the number of foodservice companies

**Source:** our own construction, using GeoDa software tool, and data provided by Romanian National Authority for Tourism
Considering the frequency and the significance of each class of spatial association, two categories are of a particular interest for policy purposes, namely:

1. **dark red (“high – high”)** spatial clusters, indicating the traditional, well-developed tourist areas such as Black Sea area, Danube Delta, Prahova Valley (mountain tourism), Maramures (traditional village/rural and mountain tourism), Bucovina (traditional village/rural and ecumenical tourism), Valcea and Harghita counties (balnear and mountain tourism), Sibiu area (mountain, traditional village/rural and cultural tourism), Cluj-Napoca area (cultural tourism). They can be considered *functional tourist areas* – interpreted as differentiated geographical areas characterised by “a concentration of uses, activities and visitation related to tourism”, which incorporates “clear references to varied elements of natural space – area, concentration, soil usage, visitation and frontiers” (Panosso Netto and Trigo, 2015, p. 66, with reference to Haylar et al., 2008) In these areas investment support is necessary in order to boost their competitiveness not only in national but also in international context via increased quality and diversification of the provided services. At the same time the policy-makers’ attention should be directed to actions able to develop working tourist clusters, with strong organisation of the inter-firm relations as well as advanced networks between all significant local actors.

2. **light blue spatial associations (“low – high”),** where LAUs with small number of accommodation units are surrounded by LAUs with big number of foodservice units. This is the case of the metropolitan areas of the big cities (e.g. Bucharest, Constanta, Timisoara, Iasi, Cluj-Napoca, Galati, Oradea), suggesting the interest of their inhabitants – based on high income and living standards – to dine out in attractive natural areas. In such cases the efforts must be concentrated on providing good access infrastructure combined with rational land use in order to preserve the natural, green areas surrounding big cities.

In addition, considering the well-balanced distribution of natural and cultural-historic landscapes in Romania, adequate actions are recommended in order to create and promote new tourist destinations especially in the lagging behind regions, where, so far, the map does not indicate significant spatial associations of LAUs in terms of number of accommodation and foodservice units.
Conclusions and Further Developments

The performed analysis can be considered helpful from theoretical point of view, based on its capability to improve the methodologies for examining the relationships between companies acting in tourism activity and the landscape elements, for conceptualizing and identifying functional tourism areas. It is helpful for practitioners too, as it provides useful information for selecting sites for new businesses in tourism industry, as well as for policy-makers indicating those tourist areas where additional support for their development could be beneficial.

However, this exploratory research should be seen just as a first step of a larger inquiry, able to offer a broader view on the spatial associations with relevant tourist activity. To this end, further investigation would envisage a wider range of indicators characterizing tourism development (e.g. number of beds in tourist accommodation units, number of accommodation units by quality class, number of arrivals, number of overnight stays, etc.), as well as indicators regarding the social-economic development level, the access to transportation infrastructure and so on. Also, the robustness of findings needs to be considered, so as to check whether the results are stable in time.

Another future direction of investigation points at the internal features of the ‘spatial clusters’ in the meaning derived from the interpretation of the univariate and bivariate Local Moran statistic: in other words, to what extent these significant spatial associations (and agglomerations of firms) exhibit the characteristics of tourist clusters, as clusters “á la Porter”, i.e. geographic concentrations of tourist resources and attractions, related infrastructure, equipment and service firms and other supporting sectors and administrative institutions with integrated and coordinated activities (Kirschner, 2015). And, even if this paper cannot provide the empirical evidence necessary for establishing the stage of development, the simple existence of tourist clusters in the “dark red” areas may also suggest the other side of the coin, namely competition relationships, which can be a source of increasing the quality of tourist services. Such relationships would be also interesting to be explored.

In methodological terms, as mentioned by van Herwijnen et al. (2004), the success in applying GIS techniques for local or regional planning is closely related to the responses to requirements such as the meeting of scientific credibility standards (.i.e. very good links between GIS and spatial statistics, etc.) and the provision of customized products for scientific analysis. In such a context the local indicators of spatial association – local Moran statistic included – are seen
as useful instruments for identifying local spatial clusters and for assessing the influence of a single location on the corresponding global statistics (Symanzik, 2014). They can be also employed for highlighting influential points in a regression framework, representing a further direction of investigation for our research.

**Acknowledgement.** The exchange of ideas on the paper topic with our colleague, Professor Zizi Goschin, has contributed to widening the scope of our investigation. Her support is gratefully acknowledged.

**References**


