Treasure Hunting in the 21st century: A Decade of Geocaching in Portugal

Teresa Santos, Ricardo N. Mendes, António M. Rodrigues, Sérgio Freire

e-GEO, Faculdade de Ciências Sociais e Humanas, FCSH, Universidade Nova de Lisboa,

Lisboa, Portugal

teresasantos@fcsh.unl.pt
rnmendes@fcsh.unl.pt
amrodrigues@fcsh.unl.pt
sfreire@fcsh.unl.pt

Abstract: The present study looks at geocaching, a popular location-based mobile game, where the goal is to use a Global Navigation Satellite System (GNSS), usually the Global Positioning System (GPS) to hide and seek containers placed anywhere in the field. People who engage in this activity, the *geocachers*, constitute a geographically distributed community that makes use of mobile and Web 2.0 technologies to coordinate and document their activities. Consequently, this treasure-hunting game, besides being a ludic activity, associated with a strong social networking element, also promotes new ways of exploring, interacting and communicating experiences and perceptions of the geographical environment where the activity occurs.

The majority of existing literature analyzes geocaching from a social point of view, and little reference is made to the geographical context of this activity. The aim of this study is to fill that gap and thus characterize the phenomenon in terms of its temporal and spatial distribution. Observation instruments are proposed based on motorization indexes built from available data attributes. Such attributes reveal behaviors and patterns of geocachers (individuals) and geocaches (objects).

The methodology is based on spatial data analysis; this can play an important role in exploring social phenomena that have a strong geographic component. Through the analysis of the freely available dataset that is voluntarily maintained by people engaged in the geocaching activity, a new dimension is explored: the spatial dimension. When, where and why this activity occurs was used as the framework for the analysis in this paper. The final output is an overall picture of the geocaching activity in mainland Portugal in this decade.

In a later stage, environmental characteristics are used as possible explanations for observed patterns. It is shown, using spatial model specifications, that a small number of regressors are able to highlight important characteristics in the data. A final discussion underlines the potential of geocaching to encourage social interaction, and promote cultural and natural heritage; in short, it has some paramount attributes of an economically sound and sustainable sector.

Keywords: geocaches, geocaching, GPS, Web 2.0, spatial analyses

1. Introduction

The Web 2.0 has brought a huge change in human interactions in the first decade of the 21st century. One of the recent social phenomena of Web 2.0 is geocaching. This 10 year-old activity brings together a treasure hunting game that: 1) looks like a true sport/open air activity, 2) involves hi-tech gadgets (from handled GPS units to fancy smartphones or tablets), and 3) includes the Internet! Altogether, geocaching is a true grownups' playground, where individuals, groups of friends or entire families can interact. The goal of geocaching is to find a hidden container, named *geocache*, in a public place and then sign the logbook to record the visit. The caches' information is published in the official geocaching website (www.geocaching.com), and it includes a GPS position, along with some clues and most often information regarding the place where the cache is hidden. Anyone can register for free, and the only requirements are a GNSS unit or a GNSS-enabled device. Usually, but not always, the log is later registered in the official website and experiences are shared with other geocachers.

Created in 2000, geocaching reached Portugal in February 2001. Ten years and 18,500 geocaches later, almost 2,000,000 logs were done, bringing together a community of over 17,000 registered individuals just in Portugal.

Geocaching can be seen as a touristic activity. Many caches are placed in heritage sites, and have detailed description and photos of the monument/site in their webpage. Furthermore, numerous logs are posted by foreigners. The most visited cache in Portugal has been found 1577 times and is located in *Centro Cultural de Belém*, followed by a cache located in *Parque das Nações* (both in Lisbon), with 1492 findings. The two places are well known touristic spots. In this way, geocaching is a volunteer activity that can also be used as an attraction factor for travelers to visit certain destinations, combining holiday time with the enjoyment of local history, culture and landscape. Several touristic projects that include geocache trails were already developed in the United States of America and Canada, among other countries.

There are several types of caches (Traditional, Multi, Mystery, Event, Letterbox, Virtual, Webcam, Wherigo™ and Earthcache), of different sizes (from micro, or nano-caches to large, virtual or unknown). Normally each cache contains a logbook or logsheet, and geocachers are invited to exchange personal objects or symbolic gifts that travel from cache to cache. Some of these gifts, called *travel bugs,* can even be traceable. The event and mega-event type, are a particular social phenomenon. These caches require a large group of people (up to 500 geocachers or more) to be present in a specific day for the event; afterwards, the cache is archived. Mystery or virtual caches may involve puzzles or brainteasers. Other caches, when hidden relatively close or in one particular trek or trail, are named power trail. Whereigo™ are particular caches that use a toolset for creating and playing GPS-enabled adventures. These toolsets are already included in many GPS units. Earthcaches are dedicated to earth sciences or special geological features. Most geocaches are totally free for the owner to hide in any public place; others, like Earthcaches, should follow specific rules in order to be officially published.

Although geocaching promotes the exploration and contact with nature and outdoor activity, it can also result in high pressure on more fragile environments. In fact, geocaching is a volunteer activity that is not yet subject to any type of control, and its implications and effects on, for example, protected areas recommend a deeper discussion about such outdoor activities. In Portugal, several such situations have been detected during this analysis: 8 caches (with over 500 logs – 1 per week) have been hidden in the *total protection area* of the Arrábida Natural Park, 50 km south of Lisbon. Other caches hidden in the vicinity of nesting colonies of endangered species were also identified and reported to local authorities. Newsome et al (2011) present a study on the environmental impacts of adventure racing in Australia in protected areas. These include soil erosion, loss of vegetation or alteration of feeding patterns of local fauna, among others. Also, human waste can play a negative roll on these habitats. Monz et al (2010) look at outdoor recreation as an agent of ecological change with the potential to affect soil, vegetation, wildlife, and water quality. The authors conclude that the current research on environmental recreation needs to be extended and include, besides local impact indicators, aspects of spatial and temporal analysis that will allow improving and bring the study to the landscape level. As a response to these critics, Geocaching.com attempts to introduce some ethics and rules on the geocaching activity by profiling the participants as virtual defenders of the environment, and providing guidelines on how to navigate in nature without harming the environment (Gram-Hansen, 2009).

In the literature, Geocaching has been studied as a social phenomenon (Gram-Hansen, 2009), that can be seen as an informal learning tool (Clough, 2010), or as a deviant behavior (Hawley, 2010), along with its practice and motivations (O'Hara, 2008). However, besides being a social activity, geocaching is also a location-based experience. The study of the territory and its influence in this outdoor activity requires therefore the application of spatial analysis' methodologies.

In this study, the temporal and spatial dissemination of geocaches in Portugal is analyzed. Key questions emerged: “Are caches distributed the same way as natural, anthropic and social attributes?”, “As an outdoor activity, are there relatively more caches hidden in non-urban areas?”, “Which are the regions with more/less geocaches?”, “What is the visiting frequency and intensity per region?”, “Which land-uses relate to the location where most popular caches are placed?”, “Are the caches’ difficulty levels related with the landscape?”. In sum, all these questions help to examine what are the spatial characteristics of the geocaches that make them attractable.

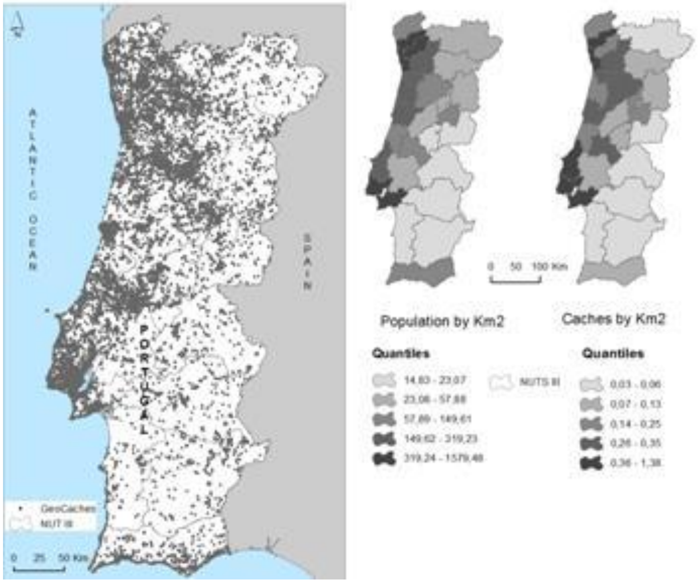
Spatial data analysis can play an important role in studying social phenomena that have a strong geographic component like geocaching. Through the analysis of the freely-available dataset, web-forums and web pages that are voluntarily maintained by geocachers, a new dimension is explored: the spatial dimension, allowing new insights regarding the relationships between people and the surrounding urban and rural environment.

2. Study area and data sets

The area selected to study the geocaching activity is mainland Portugal (Figure 1). The country is divided in several administrative levels. According to the Nomenclature of Territorial Units for Statistics (NUTS), mainland Portugal comprises 28 NUTS III regions. Portuguese landscape and climate invite to outdoor activities, such as geocaching. Portugal has a very diverse landscape, varying from mountains in the center to flat plains in the south, and from deep green valleys with vineyards in the north, to beaches in the coast. Population density is higher near the coastline, especially in the metropolitan areas of Oporto and Lisbon. Landscape becomes more rural as we move to the interior. The forest area covers about 38% of the territory, with agriculture land and farming being the second most significant land use. Portugal has a Mediterranean climate and is one of the European countries with the highest levels of annual solar radiation, a factor which favors outdoor activities.

Data for this study was collected from www.geopt.org (one of the Portuguese geocaching forums) on February 2012. The geocaches’ subset includes 18026 caches, hidden between February, 2nd, 2001 and February, 8th, 2012. In order to make the analyses more feasible, only caches located on the mainland were considered. For the temporal analysis, the caches with no visiting activity or misplaced caches (with coordinates outside of the study area) where removed from the data set. The total number of caches equals 17291. For the geographical analysis, archived caches (caches disabled by the owner), were also deleted from the database. The final database for geographic analysis consists in 13553 operational caches. This set allows evaluating geocaching for a specific time-stamp. The attributes included in the data set are cache name, cache owner, location, difficulty, terrain, cache type, status, size, hidden date, average size of logs, total number of logs, number of findings, no-findings, notes, photos, traceable items and votes.

In order to analyze the relation between landscape and geocaching activity, two maps were used. The Land Use / Land Cover Map of Continental Portugal (*Cartografia de Ocupação do Solo de Portugal Continental,* IGP 2011) is available for the whole country for 2007, with a minimum mapping unit of 1 ha, and 2 levels of thematic detail. The Natura 2000 network is composed of areas of community importance for the conservation of habitats and species, in which human activities must be compatible with the preservation of those natural values (EC, 2012). For the study of population distribution, the census tracts for 2011 (preliminary results) were used (INE, 2011).



CODE	NUTS III	Area (km2)	Population in 2011	Nº GeoCaches
101	MINHO-LIMA	2219	244836	557
102	CÁVADO	1246	410149	442
103	AVE	1246	511737	301
104	GRANDE PORTO	815	1287276	578
105	TÂMEGA	2620	550469	785
106	ENTRE DOURO E VOUGA	861	274859	278
107	DOURO	4108	205902	488
108	ALTO TRÁS-OS-MONTES	8172	204381	470
201	BAIXO VOUGA	1804	390840	317
202	BAIXO MONDEGO	2063	332306	587
203	PINHAL LITORAL	1744	260924	420
204	PINHAL INTERIOR NORTE	2616	131371	205
205	DÃO-LAFÕES	3489	277216	911
206	PINHAL INTERIOR SUL	1905	40705	178
207	SERRA DA ESTRELA	868	43737	119
208	BEIRA INTERIOR NORTE	4063	104403	250
209	BEIRA INTERIOR SUL	3748	75026	203
210	COVA DA BEIRA	1375	87869	238
301	OESTE	2220	362523	773
302	GRANDE LISBOA	1377	2042326	1898
303	PENÍNSULA DE SETÚBAL	1625	779373	642
304	MÉDIO TEJO	2306	220660	799
305	LEZÍRIA DO TEJO	4275	247449	544
401	ALENTEJO LITORAL	5309	97895	166
402	ALTO ALENTEJO	6249	118352	266
403	ALENTEJO CENTRAL	7229	166802	259
404	BAIXO ALENTEJO	8543	126692	283
501	ALGARVE	4997	451006	596
	TOTAL	89092	10047083	13553

Figure 1. Study area and data set used for studying the geocaching phenomena at NUTSIII level

3. Methodology

Geocaching is a social networking location-based game. In the present study, a Geographic Information System (GIS) was used to model and analyze the spatial distribution of this activity in Portugal. In order to take into account the neighborhood effect of geographical data, Exploratory Spatial Data Analysis (ESDA) and spatial statistics are used (Lloyd, 2011; Olaya, 2012). The former are particularly important in the present context, since cross spatial variation is studied in order to evaluate the relation between location of the caches and the social and natural context of the areas where they are placed. Hence, geographical datasets are aggregated for particular regions, which raise common analytical questions related to lattice data (Rogerson, 2001).

Geographical phenomena are generally point-based. This means that associated to most actions, there is a pair of geographical coordinates which pin-point their location relative to a particular model of the Earth. When such data is aggregated and summarized for specific areas, there is a loss of information. Spatial stationarity must be assumed since intra-regional variations are not known (Lloyd, 2011). Resulting mis-interpretations are generally related to ecological fallacy (Freedman, 2001). The other common issue which should be taken into account when dealing with lattice datasets is that aggregates are highly dependent on the shape of each spatial unit; this is what is known as the *Modifiable Areal Unit Problem* (MAUP) In most cases, this may not be controlled, but should always be acknowledged (Rogerson, 2001).

Assuming spatial phenomena occurs in a continuous - although irregular – surface, it is a too strong assumption to analyze observations as independent (Anselin, 1988; Fingleton, 1999). Exploratory Spatial Data Analysis (ESDA) has the merit of making space endogenous. Neighborhood relations are quantified and serve as weights in the construction of spatially lagged variable transformations. As an illustration, for a given variable *X*, its i^{th} observation (U_i^{lag}) is given by the weighted average of *X* observed in the *i*'s *k* set of neighbors. Formally:

$$U_i^{lag} = \sum_{j=1}^{k} w_{ij} x_j ,$$

where w_{ij} represent the neighborhood between spatial units *i* and *j*.

Spatial Lag transformations have various uses, which range from making endogenous spatial effects in a regression – allowing estimated coefficients to be BLUE (Best Linear Unbiased Estimators) (Anselin, 1988), to the smoothing of series, facilitating the identification of geographical clusters. Their importance increases with the level of global or partial spatial autocorrelation (Getis, 2007). Spatial autocorrelation may be generally understood as the existing co-variation between observations along any given spatial surface. This is in general terms the result of spatial dependence and heterogeneity.

In the present article, spatial autocorrelation is taken into account in two stages: first, spatial lags for variables considered of significant importance are used to visually analyze patterns at the municipality level. Second, when testing the functional relationships which may explain the distribution of the geo-caching phenomenon, spatial dependence is made endogenous. In this latter case, spatial autoregressive and spatial error forms (SAR and SEM) are estimated using maximum likelihood. In the SAR specifications, the spatial lag of the dependent variable is used as a regressor; in the SEM specifications, spatial dependence is assumed to be captured in the error term. Formally:

SAR: $y = \rho W y + X\beta + \epsilon$,

SEM: $y = X\beta + u; u = \lambda W u + \epsilon$,

where $Y$is the dependent variable, Xa matrix of regressors, βits regressor, W a spatial weights matrix $\rho$and $\lambda$the spatial autocorrelation coefficients, and ϵ white noise. A set of appropriate tests will be conducted in order to infer on robustness (R^2, Log-likelihood, Schwarz and Akaike Information Criteria) and test the existence of heteroscedasticity (Breusch-Pagan test). Finally, a Likelihood-Ratio statistic will infer on spatial dependence in the series.

In order to explain the distribution of caches over Mainland Portugal, the gross number will not be used, since this is obviously related to two mass measures associated with municipalities: physical area and resident population. Hence, two density metrics are used: number of caches per resident (C[pop]) and per population density (c[dpop]). As regressors, two sets of variables are used: one related to geographical position in relation to centripetal nodes, and another related to land-use.

In terms of geographical position, two measures of peripherality are used; one quantifies position and mass in relation to the coast (*p[coast]*) and the other in relation to the two main Portuguese urban nodes, Lisbon and Oporto (*p[lp]*). The metrics used are taken from Rodrigues (2001). The reasoning for the use of these two measures is the role of the Atlantic coast as a decisive pooling factor which helps to explain the existing coastal/interior divide (Ferrão, 2002); also, the Lisbon and Oporto metropolitan areas concentrate most of the country's economic activity which make these two nodes the most important pooling centers of the country.

The second set of regressors tries to explain the distribution of caches as being partially dictated by land-use. The main hypothesis is that users may show preference towards urban areas (which would be partially explained by their easier accessibility) or alternatively, towards less densely populated parts of the study-area. With this in mind, the Land Use / Land Cover Map of Continental Portugal was used. For each municipality, the proportion of urban area in respect to total area was used as a regressor (COS[urb]). The same method was used to calculate the proportion of green area (COS[green]).The exact land-use classes were chosen based on the preliminary exploratory analysis presented below. Finally, Natura 2000 network was used as a final measure of landscape with natural interest. Again, this was weighted by total area for each municipality (NAT).

4. Results

4.1 When: temporal analysis of geocaching activity in Portugal

Although geocaching in Portugal has started in 2001, it took 5 years for the activity to take off. After 2006 it has truly exploded, reaching almost 18,000 caches spread all over the country (Figure 2). The mean growth rate from 2001 to 2011 (2012 was not used in this analysis) is 213% per year. This rate includes all status of caches: active, “needs maintenance” (caches that are momentarily not available) and archived.

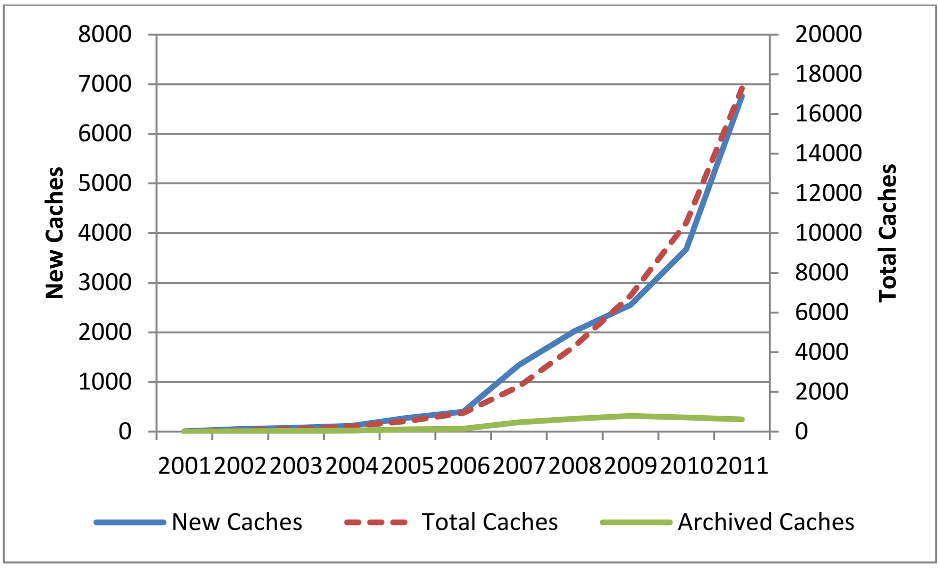
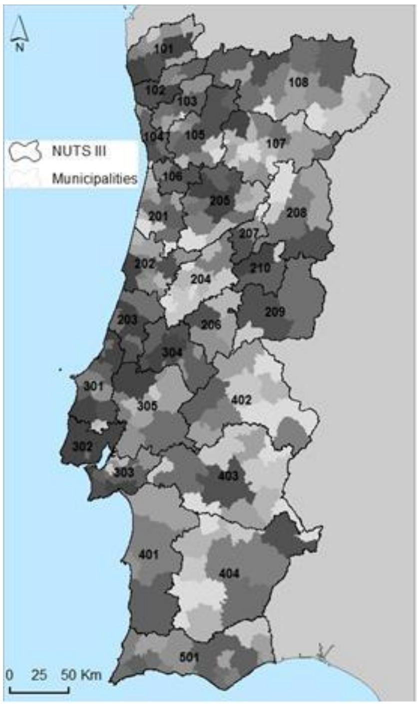


Figure 2. Yearly evolution of geocaches hidden in Portugal.

4.2 Where: spatial distribution of geocaches in Portugal

The spatial characterization of caches in the study area was performed by overlaying the caches location with administrative and thematic maps and summarizing their totals.

The spatial distribution of caches in Continental Portugal by NUTS III shows that there is a strong presence along the coast. The Great Lisbon area has the highest percentage of the total caches (14%) followed by the Dão-Lafões area, with 6,7% (Figure 3). The area with the lowest amount of caches is Serra da Estrela, the highest mountain range in continental Portugal, with nearly 1%.



		GeoCaches	
CODE	NUT III	Number	Percentage
101	MINHO-LIMA	557	4,11
102	CÁVADO	442	3,26
103	AVE	301	2,22
104	GRANDE PORTO	578	4,26
105	TÂMEGA	785	5,79
106	ENTRE DOURO E VOUGA	278	2,05
107	DOURO	488	3,60
108	ALTO TRÁS-OS-MONTES	470	3,47
201	BAIXO VOUGA	317	2,34
202	BAIXO MONDEGO	587	4,33
203	PINHAL LITORAL	420	3,10
204	PINHAL INTERIOR NORTE	205	1,51
205	DÃO-LAFÕES	911	6,72
206	PINHAL INTERIOR SUL	178	1,31
207	SERRA DA ESTRELA	119	0,88
208	BEIRA INTERIOR NORTE	250	1,84
209	BEIRA INTERIOR SUL	203	1,50
210	COVA DA BEIRA	238	1,76
301	OESTE	773	5,70
302	GRANDE LISBOA	1898	14,00
303	PENÍNSULA DE SETÚBAL	642	4,74
304	MÉDIO TEJO	799	5,90
305	LEZÍRIA DO TEJO	544	4,01
401	ALENTEJO LITORAL	166	1,22
402	ALTO ALENTEJO	266	1,96
403	ALENTEJO CENTRAL	259	1,91
404	BAIXO ALENTEJO	283	2,09
501	ALGARVE	596	4,40
	TOTAL	13553	100,00

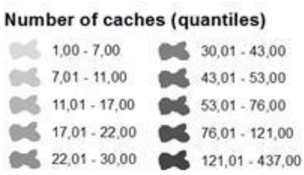


Figure 3. Distribution of Geocaches (NUTSIII and municipalities)

The spatial distribution of geocaches shows some clear patterns. Figure 4 shows: (1) the distribution of the number of caches weighted by resident population, (2) by total area and (3) by population density. The original point-dataset was aggregated at the municipalities' level in order to analyze the cross-variation taking into account the three mass indicators just mentioned. The resulting densities were spatially weighted using a row-standardized binary contiguity spatial weights matrix. The result is a set of spatial lagged variables.

When caches are weighted by resident population, what is observed is a strong corridor running from the coastal area just north of Lisbon, following the north-east direction (northern mountain system Montejunto-Estrela). When weighted by area, concentration along the coast is clear, justified by the increasing surface area of municipalities as we move inland.

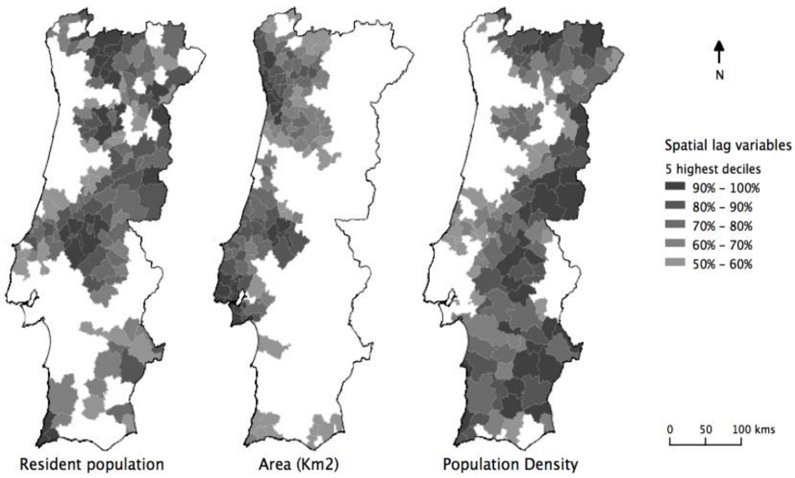


Figure 4. Density of geocaches

Regarding land cover, caches are placed preferably in natural or semi-natural areas, covered with scrubs and/or herbaceous vegetation (22,86%), followed by urban areas (19,50%) and forests (18,85%) (Table 1). This confirms that geocaching is mostly an outdoor activity conducted outside of urban centers. Although most people live in artificial areas, most of the caches are placed in natural environments. Based on this analysis, two indicators were constructed and used as regressors: classes 1.1 ("Urban fabric"), on one hand, 3.1 ("Forests") and 3.2 ("Scrub and/or herbaceous vegetation associations") on the other, are used and their area per municipality calculated representing the dichotomy between urban and natural landscapes (represented by COS[urb] and COS[green]).

Table 1. Distribution of geocaches by land use / land cover class

Land Use / Land Cover Map of 2007	Geocaches Number	Geocaches Percentage
1.1 Urban fabric	2643	19,50
1.2 Industrial, commercial and transport units	710	5,24
1.3 Mine, dump and construction sites	132	0,97
1.4 Green urban areas, sports and leisure facilities	228	1,68
2.1 Arable land	1210	8,93
2.2 Permanent crops	606	4,47
2.3 Pastures	142	1,05
2.4 Heterogeneous agricultural areas	1157	8,54
3.1 Forests	2555	18,85
3.2 Scrub and/or herbaceous vegetation associations	3098	22,86
3.3 Open spaces with little or no vegetation	407	3,00
4.1 Inland wetlands	17	0,13
4.2 Maritime wetlands	56	0,41
5.1 Inland waters	291	2,15
5.2 Marine waters	301	2,22

4.3 Why: environmental conditioning

As described above, besides exploring the distribution of the phenomenon, the goal of the study was to explain the observed patterns. Tables 2 and 3 show the econometric results from the spatial lag (SAR) and the spatial error specifications (SEM). The variables used were first standardized, subtracting the mean and dividing by two times the standard deviation, as proposed by Andrew Gelman (2007). When the number of caches weighted by resident population is used as the dependent variable (Model 1), peripherality in relation to Lisbon and Oporto (p[lp]) is significant with a negative sign in the estimated coefficient. COS[urb] is also significant with an opposite sign.

Table 2. Estimation results (SAR specification)

	Model 1	Model 2	Model 3
Spatial lag	0.0050 (0.8409)	-0.002 (0.0174)	-0.007 (0.0000)
Constant	0.3762 (0.8409)	0.1250 (0.8609)	0.3858 (0.7061)
p[coast]	-0.090 (0.3833)	0.0942 (0.1170)	-0.183 (0.0007)
p[lp]	-0.370 (0.0025)	-0.019 (0.7764)	-
COS[urb]	0.3021 (0.0000)	0.0703 (0.0973)	0.1553 (0.0029)
COS[green]	-0.032 (0.6024)	0.3840 (0.0000)	-
NAT	0.1456 (0.0205)	0.6318 (0.0000)	0.5278 (0.0000)
R-squared	0.2821	0.7664	0.5122
Log likelihood	-159.385	0.4654	-105.906
Akaike info criterion	332.77	13.0691	221.813
Schwarz criterion	358.163	38.4625	239.951
Breusch-Pagan test	9.5777 (0.0881)	370.0447 (0.0000)	250.1606 (0.0000)
Likelihood Ratio Test	19.6414 (0.0000)	4.6072 (0.0000)	29.4284 (0.0000)

This shows some tendency of caches to be distributed near urbanized areas. On the other hand, natural heritage is also positively rated (positive significant coefficient associated with NAT).

When the total number of caches is weighted by population density (Model 2), results are similar, yet with some notable differences. First, robustness increases; this is shown by a drastic increase in the R-squared statistic, smaller absolute value of the Log likelihood function and Akaike Information Criteria and Schwarz Criterium. The Likelihood Ratio test also corroborates these results. Second, results of the Breusch-Pagan test indicate that heteroscedasticity is no longer an issue. In relation to the estimated coefficients, COS[green] and NAT are both significant with a positive sign, peripherality is no longer relevant whilst COS[urb] is only significant at the 90% level. In order to get rid of any multicollinearity, one peripherality measure together with COS[green] were dropped and the model estimated one last time.

The results are clear and quite attractive: location of the caches is stronger near the coast. Natural heritage (NAT) is particularly important with by far the highest coefficient. Yet, although caches tend to be located within municipalities with attractive landscapes, they generally tend to be in or near urban areas. Finally, Model 3 excludes the measure of peripherality with the least significance (p[lp]), as both are highly correlated. For the same reason, COS[green] is also not included. The estimation results confirm previous inferences. Although the percentage of variance explained is smaller (an expected result given the smaller number of regressors), the unexplained variance is captured by the autocorrelation coefficient. Results from the SEM specification (table 3) confirm the SAR estimation results.

Table 3. Estimation results (SEM specification)

	Model 1	Model 2	Model 3
Constant	0.0003 (0.9938)	0.0049 (0.8692)	-0.009 (0.8172)
p[coast]	-0.087 (0.5778)	-0.110 (0.2945)	-0.235 (0.0014)
p[lp]	-0.581 (0.0007)	-0.181 (0.1202)	-
COS[urb]	0.3123 (0.0000)	0.1785 (0.0015)	0.1486 (0.0070)
COS[green]	-0.077 (0.2704)	0.3807 (0.0000)	-
NAT	0.2428 (0.0003)	0.3870 (0.0000)	0.6062 (0.0000)
Spatial lag	0.5093 (0.0000)	0.3703 (0.0000)	0.5239 (0.0000)
R-squared	0.3335	0.604542	0.551
Log likelihood	-153.021513	-76.362685	-98.682475
Akaike info criterion	318.043	164.725	205.365
Schwarz criterion	339.809	186.491	219.875
Breusch-Pagan test	6.9773 (0.2223)	513.2465 (0.0000)	261.6574 (0.0000)
Likelihood Ratio Test	32.3682 (0.0000)	8.0456 (0.0000)	43.8764 (0.0000)

5. Conclusions

This study intended to analyze the national scenario of the geocaching activity in Portugal, in terms of temporal and spatial distribution. The results allowed inferring the evolution of geocaching in mainland Portugal; as well as the places where the activity is most prevalent and provided clues to what are the caches' characteristics that make them more popular than others.

Due to the lack of literature, this research proposes a methodology and some indexes for monitoring the geocaching phenomenon, allowing for comparison with other countries and other time periods. Results for the Portuguese dataset show that geo-cachers prefer locations with significant natural heritage sites although, on average, they tend not to travel far for their treasure-hunting activities. Cross-variation with other datasets should provide important and highly innovative insights about the traveling behavior of individuals during their "adventures into the wild".

Future studies should include other factors that will help to fully understand the phenomena. Local and regional trends are surely affected by other aspects like individual motivations, expectations and perceptions; social networking or physical aspects of places where caches are hidden (landscape, scenic views, cultural heritage, natural phenomena, and so on). Other variables like the presence or absence of structured supply of outdoor activities such as hiking and mountain biking trails, climbing or speleology, bird watching and other soft nature activities may also explain what seems to be the huge success of geocaching in Portugal. Naturally those assumptions require testing and confirmation. Yet, care should be taken when adding specific variables with specific local trends, as these normally add non-stochastic (deterministic) trends in the residuals.

References

Anselin, L. (1988). *Spatial Econometrics and Methods*, Kluwer Academic Publishers.

Clough, C. (2010). "Geolearners: Location-Based Informal Learning with Mobile and Social Technologies". *IEEE Transactions on Learning Technologies*, Vol. 3, No. 1, pp 33-44.

EC – European Commission (2012). Natura 2000 network. http://ec.europa.eu/environment/nature/natura2000/index_en.htm

Ferrão, J. (2002). "Portugal, três geografias em recompilação: Espacialidades, mapas cognitivos e identidades territoriais". *Lusotopie*, Vol. 2, pp.151-158.

Fingleton, B. (1999). "Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration". *Journal of Regional Science*, Vol. 39, No.1, pp.1-19.

Freedman, D.A. (2001). “Ecological Inference and the Ecological Fallacy”. In Neil J. Smelser and Paul B. Baltes (eds), *International Encyclopedia of the Social & Behavioral Sciences*. Vol.6, pp. 4027-4030.

Getis, A. (2007). "Reflections on spatial autocorrelation". *Regional Science and Urban Economics*, Vol.37, pp.491-496.

Gelman, A, (2007). “Scaling regression inputs by dividing by two standard deviations”. *Statistics in Medicine*, Vol. 27, Wiley Online Library, pp: 2865-2873.

Gram-Hansen, L.B. (2009). “Geocaching in a persuasive perspective.” Proceedings of the 4th International Conference on Persuasive Technology, April 26-29, Claremont, California, USA.

Hawley, F.F. (2010). “Agon and Ecstasy: Transgression, Transformation, and Transcendence in Competitive Geocaching”. *Deviant Behavior*, Vol. 31, No. 3, pp 225-250.

IGP – Instituto Geográfico Português. (2011). Cartografia de Ocupação do Solo de Portugal Continental para 2007, [online] http://www.igeo.pt/produtos/CEGIG/Cos2007.htm.

INE - Instituto Nacional de Estatística, I.P (2011). "Censos 2011 – Resultados Provisórios". INE, Lisboa.

Lloyd, C. R. (2011). *Local Models for Spatial Analysis*, 2nd Edition. CRC Press, Boca Raton.

Monz, C.A., Cole, D. N., Leung, Y.F., Marion, J.L. (2010). “Sustaining visitor use in protected areas: Future opportunities in recreation ecology research based on the USA experience”. *Environmental Management*, Vol. 45, pp 551-562.

Newsome, D., Lacroix, C., Pickering, C. (2011). “Adventure Racing Events in Australia: context, assessment and implications for protected area management”, *Australian Geographer*, Vol. 42, No.4, pp 403-418.

O’Hara, K. (2008). “Understanding Geocaching Practices and Motivations”. CHI 2008 Proceedings On the Move, April 5-10, Florence, Italy.

Olaya, V. (2012). Sistemas de Información Geográfica. (http://sextante.googlecode.com/files/Libro_SIG.pdf)

Rodrigues, A. (2001). "The State, Market and Accessibility Triangle: Insights Into the Manufacturing Landscape of Northern Portugal". PhD. Thesis, Faculty of Urban and Regional Studies, University of Reading.

Rogerson, P. (2001). *Statistical Methods for Geography*. SAGE Publications.