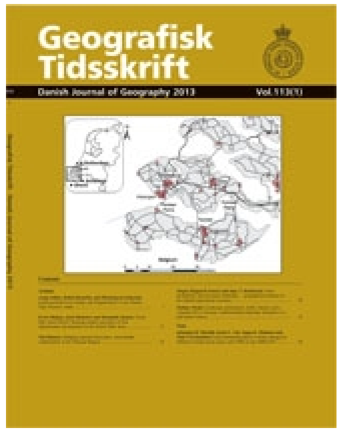

Visitor monitoring in protected areas: an approach to Natura 2000 sites using big VGI

Journal:	*Geografisk Tidsskrift-Danish Journal of Geography*
Manuscript ID	RDGS-2018-0037.R4
Manuscript Type:	Original Article
Keywords:	volunteered geographic information, webshare services, visitor monitoring, protected areas, Natura 2000 network, protected areas management

Abstract:	This paper analyses visitor monitoring and the public use of protected areas using volunteered geographic information (VGI) as a source of big data and, as the object of study, the Ebro Delta Natura 2000 site (west Mediterranean basin). Over 5,000 voluntarily recorded tracks spread over 10 years have been analysed, showing a predominance of cycling activities on trails of up to 50 km in length. Using cluster analysis and a hotspot approach, we found that the highest intensities of use are concentrated in less than 1% of the area studied, while a high dispersion of track start/finish points suggests low but scattered pressure all over the site. Spatial analysis reveals a number of spatial interactions, including the use of official trails, trespassing on reserves and potential nuisance to birdlife. This information is useful for site managers and helps them design sustainable practices. By applying the same analysis using VGI and data mining to other Natura 2000 sites, comparisons can be made and thus provide valuable assessment regarding visitor monitoring and public use patterns for the largest network of protected areas in the world.

Visitor monitoring in protected areas: an approach to Natura 2000 sites using big VGI

Joan Jurado Rota

Department of Geography, Universitat Rovira i Virgili, Faculty of Tourism and Geography, c/ Joanot Martorell, 15, 43480 Vila-seca, Spain. Orcid:0000-0002-0255-0001

María Yolanda Pérez Albert

Department of Geography, Universitat Rovira i Virgili, Faculty of Tourism and Geography, c/ Joanot Martorell, 15, 43480 Vila-seca, Spain. Orcid: 0000-0003-1634-4986

David Serrano Giné

Department of Geography, Universitat Rovira i Virgili, Faculty of Tourism and Geography, c/ Joanot Martorell, 15, 43480 Vila-seca, Spain Orcid: 0000-0003-1676-7110

Corresponding author: David Serrano Giné, email: david.serrano@urv.cat

Visitor monitoring in protected areas: an approach to Natura 2000 sites using big VGI

This paper analyses visitor monitoring and the public use of protected areas using volunteered geographic information (VGI) as a source of big data and, as the object of study, the Ebro Delta Natura 2000 site (west Mediterranean basin). Over 5,000 voluntarily recorded tracks spread over 10 years have been analysed, showing a predominance of cycling activities on trails of up to 50 km in length. Using cluster analysis and a hotspot approach, we found that the highest intensities of use are concentrated in less than 1% of the area studied, while a high dispersion of track start/finish points suggests low but scattered pressure all over the site. Spatial analysis reveals a number of spatial interactions, including the use of official trails, trespassing on reserves and potential nuisance to birdlife. This information is useful for site managers and helps them design sustainable practices. By applying the same analysis using VGI and data mining to other Natura 2000 sites, comparisons can be made and thus provide valuable assessment regarding visitor monitoring and public use patterns for the largest network of protected areas in the world.

Keywords: volunteered geographic information; webshare services; visitor monitoring; protected areas; Natura 2000 network; protected areas management

1. Introduction

The Natura 2000 network covers around 18% of the European Union's land area and almost 6% of its marine territory. It is considered to be the largest network of protected areas in the world (European Commission, 2018). Like any other protected areas, Natura 2000 sites have to deal with issues relating to planning and management (Cabalar, Pazos, Armas, & Macía, 2011; Manolache, Ciocanea, Rozylowicz, & Nita, 2017) resulting from public use and frequentation (Beeco & Brown, 2013; Campelo & Nogueira, 2016).

Visitor monitoring gathers information on visitor numbers and movements in protected areas and is of the utmost importance for various management and planning tasks ranging from environmental assessments to scheduling services and calculating budgets (Wolf, Hagenloch, & Croft, 2012). Studying visitor flows is essential for recreation policies, tourist impact assessment, the identification and avoidance of conflicts, and the analysis of frequentation trends and demands, not to mention the protection of endangered species, biodiversity, fragile ecosystems and especially sensitive sites within protected areas (Cessford & Muhar, 2003; Beeco & Brown, 2013; Tolvanen & Kangas, 2016). Indeed, visitor monitoring in some Natura 2000 sites has been a crucial tool for understanding public use and suggesting management improvements (Meijles, de Bakker, Groote, & Barske, 2014; Campelo & Nogueira, 2016; Santos, Nogueira & Vasco, 2016). Understandably, the greater the amount of data on visitor monitoring, the greater the knowledge and the more useful its analysis (Garbe, 2010; Beeco & Brown, 2013).

This paper focuses on visitor monitoring and the public use of protected areas using volunteered geographic information (VGI) as a source of big data. The Ebro Delta, a marine and terrestrial Natura 2000 site, was chosen as the object on which to perform data mining and spatial analysis using a geographic information system (GIS). The study seeks to answer three research questions: (1) What kind of visitors and outdoor activities can be tracked using VGI?, (2) What is the intensity and spatial distribution of public use in a marine and terrestrial protected area? and, (3) What implications can be derived for management practices?

2. Background

The spatial behaviour of visitors in protected areas is known as intra-site flow, and together with temporal distribution provides accurate, useful information for conservation management (Cessford & Muhar, 2003; Beeco & Brown, 2013).

Intra-site flow analysis has traditionally been undertaken using interviews, surveys, direct observation, fee permits and/or entry records (Barros, Pickering, & Gudes, 2015; Ankre, Fredman, & Lindhagen, 2016). For some time now, however, geographic technologies have also been used, thereby making available a wide range of worthwhile sources of data and procedures (Goodchild, 2007; Kitchin, 2013) including global positioning system (GPS), GIS, VGI and social media data (Orellana, Bregt, Ligtenberg, & Wachowicz, 2012; Wolf, 2012; D'Antonio, Monz, Newman, Lawson, & Taff, 2013).

2.1. GNSS-based tracking procedures

Global navigation satellite systems (GNSS) such as GPS include mobile phones, personal digital assistants and other mobile electronic devices. They provide an easy way of gathering datasets and supply a large amount of information revealing visitor movements and spatial patterns (Cessford & Muhar, 2003).

Orellana et al. (2012) tracked 372 visitor movements in a Natura 2000 site in the Netherlands and defined spatial patterns such as itineraries and pauses in their hikes, apparently because they were in a spot that interested them. They also studied the generalized sequence in which places of interest are visited, such as visitor centres and rest areas, regardless of the hike chosen, thus revealing shared features among visitors. Other studies conducted in Dutch Natura 2000 sites defined user profiles according to their spatial patterns. For instance, by tracking 138 hikers, Meijles et al. (2014) found and defined visitor profiles such as hikers with children, revealing that people rambling in protected areas rarely leave the trails.

The main question that any GNSS tracking procedure seeks to answer is what kind of spatial pattern a visitor traces within an area, and related to that, what kind of spatial behaviour they present. Beeco, Hallo, English & Giumetti (2013) used GNSS techniques along with field work to find and measure the relationship between intensity of use and recreational impacts, highlighting the relationship between user type and trail impacts. Using social science data they incorporated the opinions of 267 questionnaire respondents into cartography and generated suitability models for recreational activities, thereby creating a useful tool for environmental managers (Beeco, Hallo, & Brownlee, 2014). D'Antonio et al. (2013) also integrated GNSS and social science data to understand visitors' judgements on ecological impacts. By tracking visitors' hikes and using 408 surveys, they provided the authorities for a protected area with a useful tool for management decisions that takes into account the visitors' experience and their exposure to impaired areas. Anke et al. (2016) and other studies (e.g. Wolf et al., 2012; Tolvanen & Kangas, 2016) analysed visitor movements in protected areas and stressed the importance and usefulness of monitoring hikers so as to reduce the risk of conflicts, improve conservation, target appropriate users and, in short, manage public use efficiently.

A thorough analysis of visitor movements in protected areas reveals that tracking procedures require a good deal of time and effort. The use of GNSS devices involves finding, convincing and training potential users, while questionnaires and surveys are very time-consuming and are often seen as encroaching on visitors' time. Some visitors refuse to participate in studies on the grounds that they do not want to waste their free time, they feel suspicious when asked for their mobile number for research purposes, or they are simply uneasy about being tracked (Taczanowska, Muhar, & Brandenburg, 2008). Other issues include unnatural behaviour deriving from the fact that they know they are being monitored, the changing of hike patterns to return GNSS devices, an effort not to violate regulations (such as walking off trail), privacy concerns, type of visitor studied, particular scenarios for each study and so on and so forth (O'Connor, Zerger, & Itami, 2005; Taczanowska et al., 2008; Xiao-Ting & Bi-Hu; 2012; Wolf et al., 2012; Meijles et al., 2014).

As an alternative and complement to these procedures, researchers are envisaging new methodologies and sources of information for exploring massive amounts of data that, while minimizing the biases inherent to GNSS tracking, make it possible to extract spatial patterns for individuals and groups (Vitolo, Elkhatib, Reusser, Macleod, & Buytaert, 2015).

2.2. Procedures based on social participation and VGI

Procedures based on social participation and VGI include public participation geographic information systems (PPGIS), other map-based survey approaches and VGI obtained from altruistic collaborations, social networks and online sports sharing services, which transform users from being data consumers into data producers (Goodchild, 2007; Garbe, 2010; Norman and Pickering, 2017). One of the most valuable benefits of these procedures is that they provide massive amounts of data where hardly any or none were previously available (Kitchin, 2013).

For some time now PPGIS have been developed to gather information massively from ordinary users to be used for planning and management purposes (Beeco & Brown, 2013; Palacios & Pérez, 2015). PPGIS invite participants to share information about perceived values or attributes, and the subsequent analysis of this information enables activities to be identified and visitor distribution and behaviour inferred. Wolf, Wohlfart, Brown, & Bartolomé (2015) mapped the spatial distribution of visitor activities such as cycling and then added location-specific measures to improve the visitor experience, also including management implications. In their study, a user of the PPGIS software was not only able to ask questions about trails but also add locations and propose actions for improving the experience for other users. Over 600 participants built up a geodatabase showing the location and use of trails, the time needed to complete them, their perception and tips for future visitors, thus creating a useful tool for managers to assess user flow and improve the experience.

VGI comes from the selfless collaboration of countless users who register with GNSS devices and share the tracks of their outdoor activities on sport websites. Many websites and apps enable people to track, share, rate and comment on trails (Norman and Pickering, 2017). Some of the most popular are Endomondo, GPSies, MapMyRide, Runtastic, Sports-tracker, Strava and Wikiloc, and GNSS loggers and smartphones are used to enter movement data easily and cheaply (Campelo & Nogueira, 2016).

Nogueira, Silva, Grilo, Rosalno, & Silva (2012) analysed 750 valid tracks for a Portuguese Natura 2000 site recorded on GPSies over two years, and then used spatial analysis to reveal the most used trails and the location and characteristics of other trails including shortcuts, illegal paths and trails encroaching onto private property, making this piece of research a useful tool for managers. VGI also proved to be helpful in detecting conflicts among users, as demonstrated by Santos et al. (2016) when analysing runners and cyclists on over 500 tracks recorded on GPSies between 2007 and 2015, and to check whether visitors used the official network of trails or informal/illegal ones. This could help prevent environmental impacts, as pointed out by Campelo & Nogueira (2016), who explored over 600 tracks from GPSies and Wikiloc websites. VGI has also been used to monitor user behaviour and analyse visitor movements in relation to natural and cultural heritage, as in the case of Mínguez, Calle, & García (2016), who worked on Spanish Natura 2000 sites by studying over 3,000 tracks from the Wikiloc website. Finally, VGI has even been also used to analyse pedestrian transport and trail networks in protected areas (Márquez-Pérez, Vallejo-Villalta, & Álvarez-Francoso, 2017).

3. Area studied

The Ebro Delta is the biggest wetland on the Iberian Mediterranean coast and one of the most important in the western Mediterranean basin (Figure 1). It extends over 320 km^2 and, although it is mainly covered with rice fields, its landscape is remarkably rich and includes marine and terrestrial habitats such as marsh, meadows, sand plains, lagoons, ponds, riverine forests, human settlements and marine ecosystems. Vegetation and wildlife, mainly birds, are especially important. The Ebro Delta's natural and cultural values are safeguarded under various overlapping protected designations – including Natura 2000, Ramsar, UNESCO Man and the Biosphere Programme and Natural Park –

which cover 475 km^2, 350 km^2 of which are on the sea, and include 6 reserves where human activity is strictly controlled. The managers of the site encourage environmental education, visits and eco-friendly outdoor activities such as bird-watching, cycling and kayaking.

Over 150,000 visitors a year take in the sights, in addition to people living in the vicinity (Generalitat de Catalunya, 2017a). Visitors spend time in the interpretation centres, go to viewing points, go bird-watching, hike, cycle, go horse-riding or kayaking, visit the beach and picnic in designated and non-designated areas. Ten official trails make it easy to discover the Ebro Delta on foot, bike or horseback and stretch for over 220 km, linking a number of onsite facilities and amenities (Table 1). Problems have recently been detected by managers regarding the flow of visitors in certain areas or seasons, access to restricted areas, and illegal activities such as disturbing birds during the breeding season. As a result, intra-site flow management is needed along with new sustainable practices (Generalitat de Catalunya, 2017b).

4. Analysis

The data were analysed following a four-step procedure: (1) selection of criteria for choosing the geoserver and download of tracks, (2) webscraping, (3) data cleansing and construction of geodatabase, and (4) statistical and spatial analysis.

4.1. Data source

The Wikiloc website was chosen as the VGI data source for this study because it is the most popular and active outdoor activities geoserver in the region. It is a free location-sharing online server that started operating in 2006 and that one decade later had over 3.5 million users all over the world, sharing more than 8.5 million outdoor trails and 14.5 million photographs (https://www.wikiloc.com/). Since 2008 Wikiloc has run on Google, and live tracking can be uploaded using apps for IOS and Android devices.

It has been used as a data source in a number of VGI studies (e.g. Mínguez et al., 2015; Márquez-Pérez et al., 2017). As Campelo & Nogueira (2016) point out, the use of one web platform or another depends on its popularity among users, which varies according to the region and the emphasis of each website. Public availability, cost and profile of the platform are other important aspects to consider (Norman and Pickering, 2017). Notwithstanding, the more tracks available in an online location data server, the better the quality of the analysis (Garbe, 2010).

4.2. Webscraping

A demarcation line (37*36 km) was drawn around the area studied in order to select all the tracks in the Ebro Delta, including those generated within the area and those crossing the demarcation boundaries. The resulting territory embraces marine and terrestrial areas and includes not only the protected sites, but the entire Ebro Delta region, since it is understood that visitors to the area do not restrict their experience within the boundaries of protected sites (Generalitat de Catalunya, 2017b). On this subject, Campelo & Nogueira (2016) noticed that, although most tracks do in fact remain within the boundaries, sections of others extend beyond them. Other studies (e.g. Santos et al., 2016) also show that the spatial extent of the trails usually stretches further than the protected areas.

A shapefile polygon was therefore used to delineate the reference framework, allowing the massive download of tracks recorded on Wikiloc over the 10 years from 2006 to 2016. During webscraping, a number of spatial attributes were selected and kept for each feature, including an identifier and X, Y coordinates stored in a .kml extension. Thematic attributes were also considered, including an identifier, the name of the track, the name and identification of the user who uploaded it, the type of outdoor activity, distance, slope, perceived difficulty, number of visits, number of downloads, date of performance, date of upload, description, number of linked photographs, number of comments, type of device used to access the track, and other web aspects (Table 2).

4.3. Geodatabase

A geodatabase was set up using ArcGIS 10.2 software following a three-step process.

First of all, the .kml files were merged and exported to the .shp extension to be used in the ArcGis suite. Secondly, the resulting geodatabase was cleansed and systematized. The commonest error involved hundred-kilometre straight tracks that crossed the area studied from south-west to north-east. Five errors fell within this category. No duplicate tracks were recorded and no topological errors encountered. Finally, the resulting database was linked to the spatial extent of the data via a simple linking operation.

4.4. Data analysis

The data analysis was undertaken considering the statistical and spatial elements of the tracks.

The statistical analysis included (1) characterization, (2) outline, and (3) temporal evolution. A general characterization of the data helped us obtain an overall understanding of the tracks, checking issues such as perceived difficulty, length, slope and the existence of circular routes. By outlining the tracks, a series of maps was produced showing the spatial distribution of the tracks and their typology (i.e. on foot, by bike, by car, by boat and so on). The temporal evolution took into account Wikiloc usage on a monthly and yearly basis, distinguishing between the origins of the uploaded data (i.e. GPS or smartphone).

Spatial analysis was performed by mapping the start and finish points of the tracks and the path they took. This was done using binary analysis that involved overlapping the tracks with a square grid of 100 m, following a procedure previously developed on other Natura 2000 sites (Campelo & Nogueira, 2016; Santos et al., 2016). Blocks of 100 m per side were specially selected because of the extent of the area studied and the type of terrain, plus the fact that there was a massive amount of data to analyse. The result is a 100 m side grid showing sections of the area that are crossed by a track and sections that are not. A spatial analysis of this grid allowed the start and finish points for each track to be extracted, and a tool was applied to count the number of tracks crossing each cell. Various spatial analysis operations and queries were later performed to monitor visitor movements, assess the intensity and distribution of public use, and point out management issues. Some spatial statistics were also performed. On the one hand, Anselin Local Moran's I algorithm was used to identify spatial clusters with high or low values; this tool is useful because it points out outliers with no statistical meaning. On the other hand, Getis-Ord Gi* statistic was calculated to detect high and low spatial clusters, revealing hot and cold spots; this algorithm is particularly interesting in spatial analysis because it analyzes features within their context, proportionally comparing the local sum for a feature and its neighbours to the sum of all features, identifying statistically significant scores (Mitchell, 2005). For this statistic, a fixed distance band was chosen in a moving window automatically computed so each feature analyzed had at least one neighbour in a systematically defined Euclidean distance. This fixed distance was preferred to other spatial relationships because it better computes the entire studied population and fits properly in polarized data (Mitchell, 2005).

5. Results

Visitor monitoring was undertaken by considering the statistical and spatial analysis.

The statistical analysis included (1) a quantitative approach, (2) characterization, (3) seasonal evolution, and (4) intensity of use. The spatial analysis included (1) a spatial patterns approach, (2) cluster analysis, (3) hotspot description, and (4) geographic distribution analysis.

5.1. Track growth

Over the ten years from 2006 to 2016, the tracks shared on Wikiloc totalled 5,323 covering 111,117.4 km. The growth in the number of separate tracks uploaded has been continuous, starting from the 9 recorded during the first year to the 1,506 added in 2016. Tracks were recorded using either GPS loggers or smartphones, but with the passing of time the use of GPS devices has decreased. In fact 100% of the tracks were recorded on GPS at the beginning of the period, but only 19.39% by the end (Figure 2).

5.2. Track characteristics

Most of the tracks (60%) are circular, i.e. the start and finish points are the same. As regards difficulty, nearly 64% are perceived as “moderate”, and “easy” and “moderate” routes together account for 96.19%. Just under 5% are perceived as “difficult” or intended for "experts".

Nearly a third of the tracks (31.69%) are between 25 and 50 km long, while about 25% are either between 0 and 10 km, or between 10 and 25 km. Tracks of between 50 and 100 km are less common (14.73%) (Table 3).

Most of the tracks are intended to be used by cyclists, since 60% of them are labelled "mountainbiking", "cycling" or "bicycle touring". A third of the preferred outdoor activities are carried out on foot, with "hiking", "running" and "walking" being the tags used to describe them. Just 5% involve motor vehicles, sailing ("kayaking-canoeing", "sailing", "swimming") or non-sporting activities ("bird-watching"). Outdoor activities strictly recorded as involving water are few in number (Figure 3) and mostly concentrated near the shoreline or on beaches that are interlinked.

5.3. Seasonal distribution

Track-sharing is steady throughout the year, but lower values can be found in winter time (November, 298) and higher values in summer (August, 622). Cycling activities are clearly predominant, although other activities follow the same pattern (Figure 4).

5.4. Intensity of use

Our approach to intensity of use involved analysing the start and finish points of shared tracks along with their frequency.

Start and finish points are clustered in a few 100 m plots, meaning less than 1.5% of the blocks covering the Ebro Delta. About half (48.43%) of these plots contain the start or finish point for just one track, one fifth (20.17%) for 6 or more tracks, and just 2.5% for 30 or more tracks. Out of all plots (2048), 6 of them are start/finish point for more than 90 tracks, and one of them for 157. Data are clustered towards lower values (start/ending point for one track) and a table of probabilities (rate of occurrence) shows that any random cell has a 0.02% probability of hosting more than 29 start/finish points. A Standard Deviation calculus points 10.60 start/finish tracks (mean 4.93), and scores are highly skewed to the left. This must be understood as an unfixed pressure on the environment, as most of the plots are not start/finish point for any track, and just a few of them concentrate the bulk of the start or end of the tracks.

Start/finish points are clustered in towns such as Amposta, Camarles, l'Aldea and Sant Carles de la Ràpita, tourist beach resorts (e.g. Riumar), information centres (e.g. Casa de Fusta, Bassa de les Olles) and landmarks such as lighthouses (e.g. Fangar). However, cluster analysis (Getis-Ord Gi*) has also revealed an unequal distribution in the area and, hence, an unfixed pressure on the environment (Figure 5). Main clusters appear around towns, tourist resorts, tourist attractions and amenities, but they are higher and more compact in towns, and smaller and scattered in tourist attractions. Besides, clusters in the southern part of the Delta are used far more than clusters in the northern part and, remarkably, 28.71% of clusters in the Delta as a whole have no statistical significance (Figure 6).

Track distribution covers a quarter (25.73%) of the 100 m plots layering the area studied, appearing in a highly unequal spatial pattern from a thematic point of view, as the number of tracks crossing a 100 m plot varies notably. Nearly four-fifths (77.45%) of the cells are crossed by fewer than 36 tracks, which means that 0.70% of all recorded trails and 14.83% of the plots contain between 37 and 130 tracks. Trails are highly clustered in a few cells, since up to 349 tracks cross 5.75% of the cells, and less than 1 in 100 plots (0.98%) concentrate up to 634 trails. From a spatial point of view, this distribution is highly unequal, and a cluster approach (Anselin Local Moran's I) draws a very well-defined pattern of four main routes, in which 0.77% of the cells show high clustered values and 99% non-significant values. Of these non-significant cells, some are used only occasionally, as indicated by the hotspot analysis (Getis-Ord Gi*) (Figure 7). The four main routes account for the highest number of tracks. The first connects the town of Sant Carles de la Ràpita to Amposta and the hinterland, the second links Amposta to the ocean by a towpath alongside the river Ebro and divides into two along the right and left banks after the town of Deltebre, and the third and fourth follow the southern and northern shorelines of the delta. Finally, the linear-directional mean of the tracks indicates a movement from the sea to the hinterland in a south-westerly direction.

5.5. Spatial interactions

Spatial interactions were analysed focusing on the relationship between Wikiloc tracks and items or details that were significant for Ebro Delta managers, such as whether user tracks coincided with official trails, penetration into protected areas, the most frequently visible type of land cover, and potential nuisance to birdlife.

Spatial analysis reveals that over 50% of the tracks shared substantially cover (>75%) any of the official trails suggested by the site managers. Official trails 8, 9, 2, and 1 (Table 1) are fully identified with the highest intensity of use, while 3, 4, and 5, which are circular routes, show higher intensity on the leg of the path alongside the coast and lower intensity away from it. Many other tracks follow the paved and unpaved roads and paths on the site and frequently join up with official trails around landmarks such as viewpoints (e.g. Zigurat) and interpretation centres (e.g. Casa de Fusta, both on official trail 2).

Nearly a quarter of the tracks (23.58%) cross protected areas, both marine and terrestrial. Considering only terrestrial ecosystems, 39.06% of the Ebro Delta is protected in some way and contains 26,195.42 km of the tracks. There is a big difference between track density (and potential spatial interactions) inside and outside protected sites, since the relative number of tracks outside the protected sites is more than twice that inside them (rate index = 125.41 and 60.26 and, respectively).

Some sites are of important natural interest and human activity is strictly controlled, with only certain management practices and scientific research allowed, always with special permission. Rambling in these areas is strictly prohibited. Almost a thousand (933) tracks or sections of track amounting to almost 2,700 km have been recorded as encroaching across reserve boundaries, which means that 10.22% of the tracks cross protected areas. Paths are up to 15 km in length, with 164 of them being over 5 km. Most follow the perimeter of the site, but in 52 cases cross it. In the Fangar reserve, tracks go through dune fields that are clearly roped off, and in the Banya reserve they cross a fence and penetrate over 10 km into the site (Figure 8).

As far as land cover is concerned, more than nine out of ten tracks go across rice fields for most of their length, since this is the predominant land cover, followed by beaches and forests. Considering only protected areas, the most frequently visible types of land cover are swamp, beaches, dune fields or meadows. Although the most frequently visible type of land cover is rich and variegated, one particular land cover, be it rice fields or swamps, usually predominates.

The managers of the site are especially interested in discovering which tracks cause the bigger nuisance to birdlife, since birds are one of the site's most important natural values. By combining nesting and breeding areas with tracks and seasonal records, it has been found that visitors do not bother flamingo colonies, not even during the breeding season, which mostly coincides with the summer (Figure 9). Besides flamingos, other species analysed include the purple heron (*Ardea purpurea*), Audouin's gull (*Ichthyaetus audouinii*) and the night heron (*Nycticorax nycticorax*), along with others to be found in Annex 1 of the European Commission's Birds Directive.

6. Discussion

This study combines three main goals: visitor and outdoor characterization using VGI, the intensity and spatial distribution of public use, and possible threats and opportunities related to visitor monitoring.

6.1. Visitor and outdoor characterization using VGI

The analysis gathered over 5,000 tracks shared by users on Wikiloc in the course of 10 years. This is a great deal of data over a long time span and allows a fairly accurate characterization of the kind of user and the type of activity performed in the Ebro Delta.

Most visitors to the site take part in cycling activities, followed by hiking. Almost all users tag their routes as "moderate" or "easy", even though trails of up to 50 km in length are the most common. No doubt the flat terrain of the delta facilitates these kinds of activities. Other studies in mountainous areas (e.g. Santos et al., 2016) also show a prevalence of cycling over hiking, although the distance cycled is notably lower. Just a few tracks are tagged as to be covered on horseback or sailing, in spite of the fact that there are two official trails designed to be followed on horseback and or canoeing/kayaking. Personal observation and manager's opinion suggest that this might be explained by the types of visitor who are less inclined to record and upload their activities on webshare services.

As regards seasonality, users of the area are spread throughout the year, although summer time is popular, since the delta is a tourist destination. Investigations carried out in places less popular with tourists show other seasonal patterns, which may be linked to other types of visitor (Mínguez et al., 2015). A number of studies point out the importance of considering types of user in protected areas (e.g. Cessford and Muhar, 2003; Becco et al., 2013), while others (Barros et al., 2015) stress the relationship with management practices. In our study we also find different types of users and activities, but their typologies are more related to the use of space (i.e. kind of preferred outdoor activity, seasonality and so on) than to particular behaviours or individual attitudes, which are difficult to infer.

We believe that the main contribution of VGI to visitor monitoring and the public use of protected areas involves defining trends. As Garbe (2010) notes, no other type of information source provides such amounts of data for location movement, not only on a spatial basis but also temporally. We cannot state that VGI provides exhaustive data for tracking visitor movements in protected areas because not all visitors record or share their tracks. However, we are absolutely convinced of their representativeness, especially since we have been analysing long-term series of data (10 years). VGI advantages and disadvantages have been widely debated (Kitchin, 2013) and include representativeness, engagement and cost (Goodchild, 2007; Norman and Pickering, 2017) as well as bias (Garbe, 2010). A key concept, nonetheless, is that VGI data sources provide massive amounts of data where scarcely any or none were previously available (Kitchin 2013).

By using big data sources, gigantic amounts of data can be managed and long-term periods thoroughly analysed. The possibilities of VGI in tracking movements are still at an early stage, but are understood to be a valuable tool in visitor monitoring (e.g. Mínguez et al., 2015; Santos et al., 2016). As Campelo & Nogueira (2016) point out, VGI is free of privacy issues as it is freely given by users, and the role of the user as a creator of data yields rewarding results (Goodchild, 2007; Vitolo et al., 2015). In our experience, the most beneficial aspect is the savings in time and money, as no GNSS loggers have to be provided and individually analysed. Other opportunities have been summarized by Kitchin (2013) and include the huge volume of data, the high speed with which it is obtained, its variety, exhaustive character, fine-grained resolution, relational nature and flexibility, and importantly, the fact that it provides large data sets where information was previously lacking.

6.2. Intensity and spatial distribution of public use

Public use is spread all over the area studied, but is spatially concentrated along four main routes across the delta that link the principal towns, tourist attractions and landmarks, and given structure by the towpath along the river. Although paths spread out in all directions, only a few tracks have statistical significance since the intensity of use is concentrated in less than 1% of the overall area. This means high pressure in just a few spots, but also a lower, scattered intensity of use in many others. Studies on public use in protected areas show patterns that are difficult to generalize about because they vary from one site to another. Orellana et al. (2012) reported high clusters of users in attractions and amenities, stating that users spend much more time in the places they find most interesting or that catch their eye. Similarly, Meijles et al. (2014) revealed that, even if some paths seem not to be used at all, users spread throughout the site but in low densities and big clusters of visitors could only be easily identified in particular areas such as car parks. Meanwhile Campelo et al. (2016) point out increasing intensity levels around tourist attractions as well as car parks, and low levels of user intensity at distant points, and Barros et al. (2015) reveal an even more dramatic concentration of users to the point where only 2% of the area studied was being directly used by visitors. Mínguez et al. (2015) sum up intensity of use according to three variables: the existing trail network, preferred land cover and type of terrain. Other research (i.e. Meijles et al., 2014) argue that land cover is not a significant variable and stresses the importance of user types. For the Ebro Delta we find that landmarks, outstanding natural sites, tourist attractions and amenities including car parks are the most important items drawing visitors. But the spatial distribution of public use is not related only to the location of these items but also and in particular to the network of trails that link them, of which the towpaths along the river Ebro are the most important. This is a crucial difference compared to some other investigations that explore visitor movements in protected areas because, even though public use is scattered, most of the movements analysed follow a spatial pattern defined by a network of paths that in turn is defined by the location of items that exert an attraction, with one natural feature – the river Ebro – being the most important. Other items include landmarks, facilities and the seaside, given that the official trails are more crowded on the coast than in the interior.

As regards land cover, the prevalence of rice fields might not be considered a factor of attraction within the delta, or at least not to protected sites, where they are almost absent. On the contrary, the land cover itself might be appealing to visitors to the Ebro

Delta, and this could partly explain the extensive sprawl of tracks that could be understood as discovery trails. Research combining individual responses, such as that undertaken by Orellana et al. (2012) and Beeco et al. (2014), and VGI, such as that carried out by Palacios & Pérez (2015), could provide greater insight into collective preferences and visitor patterns.

Apart from the pressure on the environment deriving from the use of paths, the number and distribution of their start/finish points are also important, since they bring additional impacts such as unregulated car parks, litter and noise to the environment. This is even more important bearing in mind that most of the tracks start on the coastal side (the linear-directional mean points to a south-westerly movement).This information is useful for site managers as they can promote park & ride car parks near towns, thus preventing the appearance of unregulated car parks and car crowding near the coast. A considerable proportion of the start/finish points receive a very low number of users/tracks, meaning that the pressure, although lower, is widely spread all over the site. A third of the cells have no statistical significance, so their importance as regards the tracks is irrelevant but their impact on the environment might not be. Wolf et al. (2012) monitored usage levels within protected areas and noticed that accessibility impact not only involves the number of users but also the quality of users, since some visitors will be more careful than others. Recreation Ecology points no linear relationship between number of visitors and impacts but between type of users and type of disturbance (Leung and Marion, 1999; Baker and Leberg, 2018).

Managers of the site can use the distribution and intensity of the start/finish points to avoid or mitigate impacts. They can also divert the flow of users or modify their perception of the site and eventually improve their experience, as noted by D'Antonio et al. (2013) when reporting impact perception spatially gathered near sites of attraction.

6.3. Possible threats and opportunities related to visitor monitoring

A small but meaningful number of tracks encroach onto special protected sites, sometimes penetrating up to 15 km, with others deliberately criss-crossing the area. Meijles et al. (2014) observed that only 2% of visitors leave the paths, whereas Santos et al. (2016) raise this value to 64% for some users, depending on type. User behaviour and civic-mindedness are usually given as explanations for these figures (Meijles et al., 2014), but we believe that other factors such as topography, type of environment, information on the site and similar elements should also be considered. It is difficult to estimate the real impact of trespassers on the environment as sometimes any disturbance whatsoever may be enough to affect a fragile ecosystem such as dune fields. But by performing spatial analysis between tracks and areas of interest for birdlife, some interesting results can be obtained. The implications of this are of great interest to managers, and users can also benefit from an environmentally friendly experience if they are duly informed. Some protected areas such as Natura 2000 sites frequently face management challenges but lack tools for coordination and planning (Cabalar et al., 2011; Manolache et al., 2017), and therefore visitor monitoring could be extremely useful.

Visitor levels can be approached and monitored in a number of ways, with suitability mapping proving to be particularly fruitful of social data are also considered (Beeco et al., 2014). Managers of the site may be interested in mapping fragility and vulnerability to visitors as a tool for assessing tourist impact in order to prevent or minimize environmental damage or to introduce suitable visiting practices (e.g. Ankre et al., 2016).

Other applications include landscape perception and public use (Palacios & Pérez, 2015), especially taking into account perceived land cover and visitor judgments (D'Antonio et al., 2013; Mínguez et al., 2015).

7. Conclusions

Outdoor recreational activities and visitor movements are a crucial issue in the management of protected areas. Using big VGI as a data source, this study has characterized these activities and the implications they have for management in the Ebro Delta Natura 2000 site.

More than 5,000 tracks voluntarily recorded in the course of 10 years and totalling over 100,000 km were analysed. The sustained growth of shared tracks reveals a predomination of cycling activities, perceived as easy or moderate and frequently covering up to 50 km in length. Hiking activities are also important, while driving, sailing and bird-watching are much less common, with direct observation suggesting that they may be undervalued due to user type.

Spatial analysis using GIS reveals an unequal intensity of use all over the area studied. Although almost the entire site is used for recreational purposes, the most highly used tracks are concentrated in less than 1% of the area, and a hotspot approach reveals high pressure in certain locations. However, there is high dispersion found in the start/finish points of tracks that coincide only in part with the most popular routes, meaning lower but scattered intensity of use all over the delta. Cluster analysis shows that up to a third of the 100 m^2 start/finish plots have no statistical significance, regardless of the environmental impact of visitors. This information is valuable for site managers, enabling them to undertake actions to prevent or minimize damage. The analysis of other useful spatial interactions includes an assessment of users trespassing onto specially protected sites, most frequently visible type of land cover, and potential damage to areas of interest to birdlife.

VGI data has been shown to be extremely useful for assessing visitor activities and movements in open spaces and promises to be an invaluable source of information for protected areas, especially those like Natura 2000 sites where management practices are not always well defined. Massive VGI data and GIS procedures can be used by managers or local authorities to detect conflicts between users and introduce sustainable practices. Several sportive web shared services offer useful VGI data with enormous potential; the choice of one platform or another depends on their popularity in a given region, user's profile and emphasis of the service, as well as cost of data. Natura 2000 sites can be compared if the same tools and indicators are used at each. A comparative analysis of visitor monitoring across all Natura 2000 sites could be undertaken and valuable information on public use obtained at little cost. The resulting analysis would provide helpful assessment of visitor monitoring and public use patterns, and management assessments for the largest network of protected areas in the world. Finally, by combining spatially-based procedures with social data and particular behaviour/attitudes, visitor judgements and landscape perception, an analysis of public use practices can be undertaken.

Acknowledgements

This work was supported by the study is part of the Spanish Ministry of Economy and Competitiveness (AEI / FEDER, UE) under Grant SPATIUM (contract number CSO2014-52721-P), Grant CHORA (contract number CSO2017-82411-P), and the GRATET Research Group funded by the Catalan Government under code 2009-SG744.

References

Ankre, R., Fredman, P., & Lindhagen, A. (2016). Managers' experiences of visitor monitoring in Swedish outdoor recreational areas. *Journal of Outdoor Recreation and Tourism*, 14, 35-40.

Baker, A.D., Leberg, P.L. (2018). Impacts of human recreation on carnivores in protected areas. *PLoS ONE*, 13, 4, 1-21.

Barros, A., Pickering, C., & Gudes, O. (2015). Desktop analysis of potential impacts of visitor use: A case study for the highest park in the Southern Hemisphere. *Journal of Environmental Management*, 150, 179-195.

Beeco, J.A., & Brown, G. (2013). Integrating space, patial tools, and spatial analysis into the human dimensions of parks and outdoor recreation. *Applied Geography*, 38, 76-85.

Beeco, A. J., Hallo, J.C., English, W., & Giumetti, G.W. (2013). The importance of spatial nested data in understanding the relationship between visitor use and landscape impacts. *Applied Geography*, 45, 147-157.

Beeco, A.J., Hallo, J.C., & Brownlee, M.T.J. (2014). Visitor tracking and recreation suitability mapping: tools for understanding and managing visitor use. *Landscape and urban Planning*, 127, 136-145.

Cabalar, M., Pazos, M., Armas, F.J, & Macía, X.C. (2011). The Natura 2000 network in Spain and its lack of protection. *European Journal of Geography*, 1, 1-11.

Campelo, M.B. & Nogueira, R.M. (2016). Comparing webshare services to assess Mountain bike use in protected areas. *Journal of Outdoor Recreation and Tourism*, 15, 82-88.

Cessford, G., & Muhar, A. (2003). Monitoring options for visitor numbers in national parks and natural areas. *Journal for Nature Conservation*, 11, 240-250.

D'Antonio, A., Monz, C., Newman, P., Lawson, S., &Taff, D. (2013). Enhancing the utility of visitor impact assessment in parks and protected areas: A combined social-ecological approach. *Journal of Environmental Management*, 124,72-81.

European Commission. (2018). Environment. Retrieved from http://ec.europa.eu/environment/nature/natura2000/index_en.htm

Garbe, M. (2010). Online location data: analysis of publicly available trajectory data sources. In Cap, C. (Ed.) *Baltic conference learning networks*, (pp. 1-8). Kaunas: University Rostock.

Goodchild, M. (2007). Citizens as a voluntary sensors: spatial data infraestructure in the world of web 2.0. *International journal of spatial data infraestructures research*, 2, 24 32.

Kitchin, R. (2013). Big Data and human geography: opportunities, challenges and risks. *Dialogues in Human Geography*, 3 (3), 262-267.

Leung, Y-F., Marion, J.L. (1999). Characterizing backcountry camping impacts in great Smoky Mountains National Park, USA. *Journal of Environmental Management*, 57, 193-203.

Manolache, S., Ciocanea, C.M., Rozylowicz, L. & Nita, A. (2017). Natura 2000 in Romania-A decade of governance challenges. *European Journal of Geography*, 8 (2), 24-34.

Márquez-Pérez, J., Vallejo-Villalta, I., & Álvarez-Francoso, J.I. (2017). Estimated travel time for walking trails in natural areas. *Geografisk Tidsskrift - Danish Journal of Geography*, 117 (1), 53-62.

Meijles, E., de Bakker, M., Groote, P. & Barske, R. (2014). Analysing hiker movement patters using GPS data: Implications for park management. *Computers, Environment and urban Systems*, 47, 44-57.

Mínguez García, M.C., Calle Vaquero, M., & García Hernández, M. (2015). Nuevas lecturas territoriales de las actividades turístico-deportivas: huella digital y SIG [New spatial points of view on tourist-sportive activities: digital footprint and GIS]. In de la Riva, J. R., Ibarra, P., Montorio, R., & Rodriguez, M. *Análisis espacial y representación geográfica: innovación y aplicación* (pp. 759-768). Zaragoza: Universidad de Zaragoza-Asociación de Geógrafos Españoles.

Nogueira, R., Silva A., Grilo, C., Rosalino, L.M., & Silva, C.P. (2012). MTB monitoring in Arrábida Natural Park, Portugal. In *6th international conference on monitoring and management of visitors in recreational and protected areas: outdoor recreation in change – current knowledge and future challenges* (pp. 32-33).

Bodenkultur University, Stockholm: Institute for Landscape Architecture and Landscape Management.

Norman, P., Pickering, C. (2017). Using volunteered geographic information to assess park visitation: Comparing three on-line platforms. *Applied Geography*, 89, 163-172.

O'Connor, A., Zerger, A., & Itami, B. (2005). Geo-temporal tracking and analysis of tourist movement. *Mathematics and Computers in Simulation*, 69, 135-150.

Orellana, D., Bregt, A.K., Ligtenberg, A., & Wachowicz, M. (2012). Exploring visitor movement patterns in natural recreational areas. *Tourism Management*, 33, 672-682.

Palacios Buendía, A.V., & Pérez Albert, M.Y. (2015). Propuesta metodológica para la evaluación de la percepción social del potencial turístico del paisaje y del uso público en espacios naturales protegidos. El papel de los sistemas de información geográfica de participación pública [Methodological proposal to evaluate social perception on landscape touristy potential and public use in protected areas. The role of geographic information systems in public participation]. *International Conference on Regional Science* (pp. 1-21). Reus: XLI Reunión de Estudios Regionales.

Generalitat de Catalunya (2017). *Pla de Gestió del Parc Natural del Delta de l'Ebre* [Management plan for Ebro Delta Natural Park]. Retrieved from http://parcsnaturals.gencat.cat/ca/delta-ebre (Accessed 2018-11-05).

Generalitat de Catalunya (2017). *Xarxa de parcs naturals. Parc Natural del Delta de l'Ebre* [Natural parks network. Ebro Delta Natural Park]. Retrieved from http://parcsnaturals.gencat.cat/ca/delta-ebre (Accessed 2018-11-05).

Santos, T., Nogueira Mendes, R., & Vasco, A.R. (2016). Recreational activities in urban parks: Spatial interaction among users. *Journal of Outdoor Recreation and Tourism*, 15, 1-9.

Taczanowska, K., Muhar, A., & Brandenburg, C. (2008). Potential and limitations of GPS tracking for monitoring spatial and temporal aspects of visitor behaviour in recreational areas, (pp. 451-455). In Raschi, A. & Trampetti, S. (Eds.) *Management for Protection and Sustainable Development*. Montecatini: Conciglio Nationale della Ricerche.

Tolvanen, A., & Kangas, K. (2016). Tourism, biodiversity and protected areas. Review from northern Fennoscandia. *Journal of Environmental Management*, 169, 58-66.

Vitolo, C., Elkhatib, Y., Reusser, D., Macleod, C.J.A., & Buytaert, W. (2015). Web technologies for environmental Big Data. Environmental Modelling & Software, 63, 185-198.

Wolf, I.D., Hagenloch, G., & Croft, D.B. (2012). Visitor monitoring along roads and hiking trails: How to determine usage levels in tourist sites. *Tourism Management*, 33, 16-28.

Wolf, I.D., Wohlfart, T., Brown, G., & Bartolomé, A. (2015). The use of public participation GIS (PPGIS) for park visitor management: a case study of mountain biking. *Tourism Management*, 51, 112-130.

Xiao-Ting, H., & Bi-Hu, W. (2012). Intra-attraction tourist spatial-temporal behaviour patterns. *Tourism Geographies*, 14, 4, 625-645.

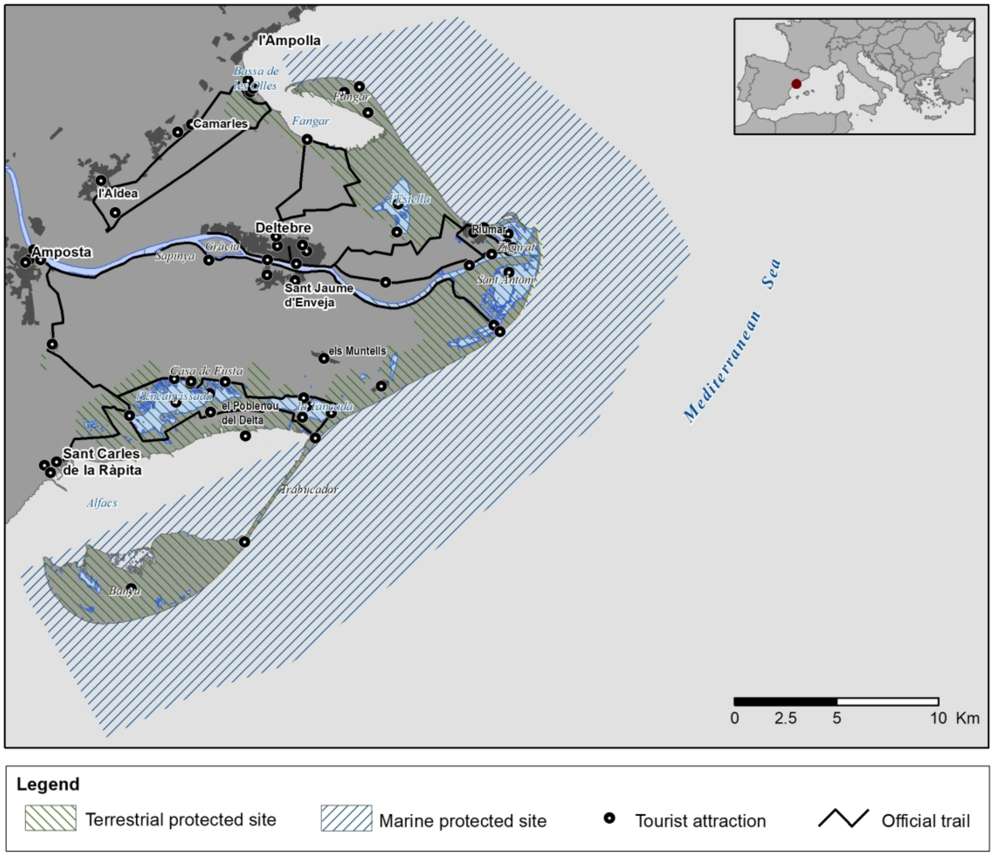

Figure 1. Study area: Ebro Delta.

196x169mm (300 x 300 DPI)

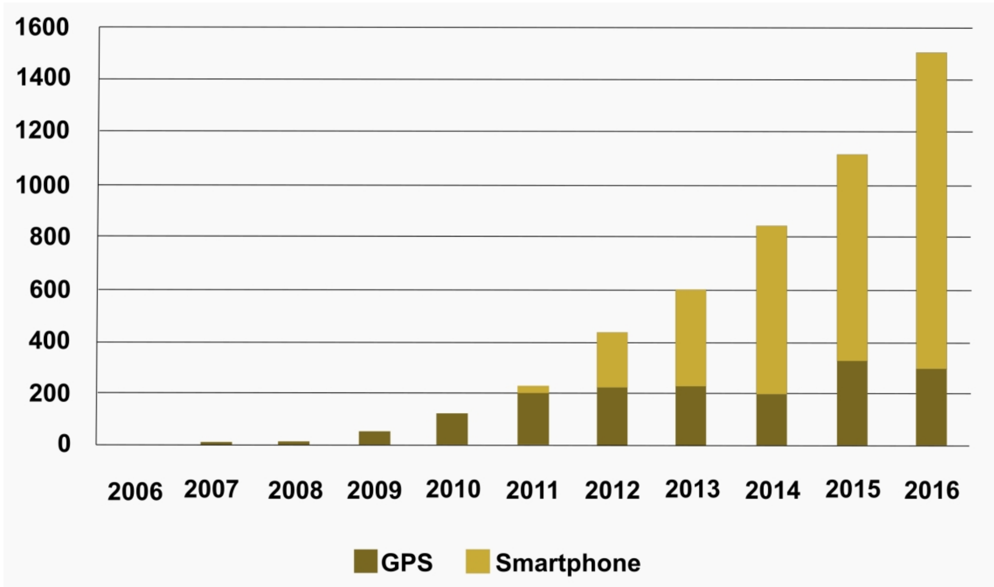

Figure 2. Tracks uploaded to Wikiloc and data entry device, 2006-2016.

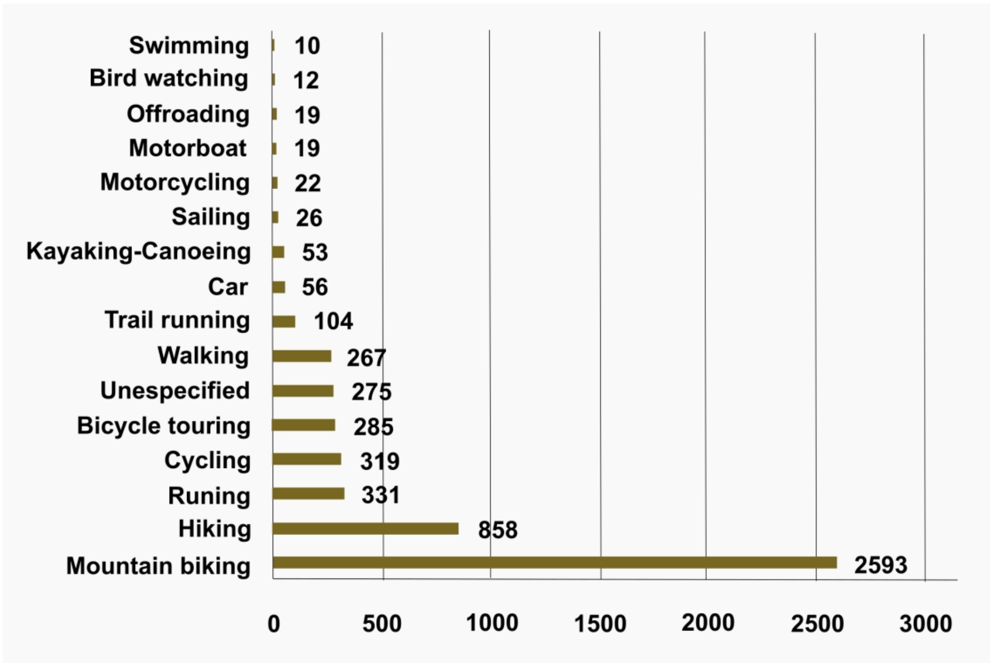

Figure 3. Tracks characteristics: intended activity or purpose (grouped by cycling, hiking or other activity).

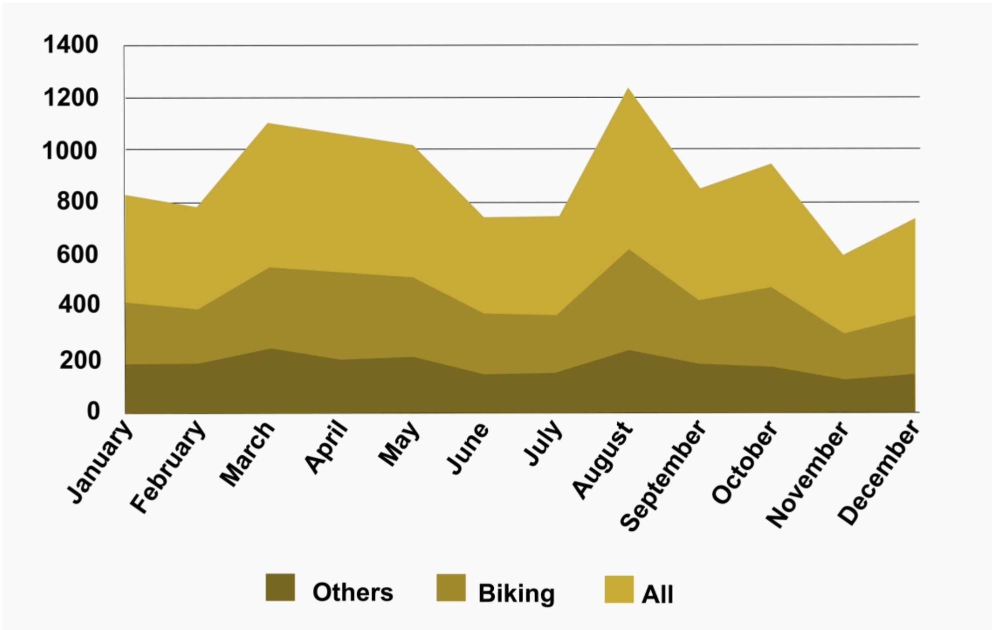

Figure 4. Tracks seasonal distribution.

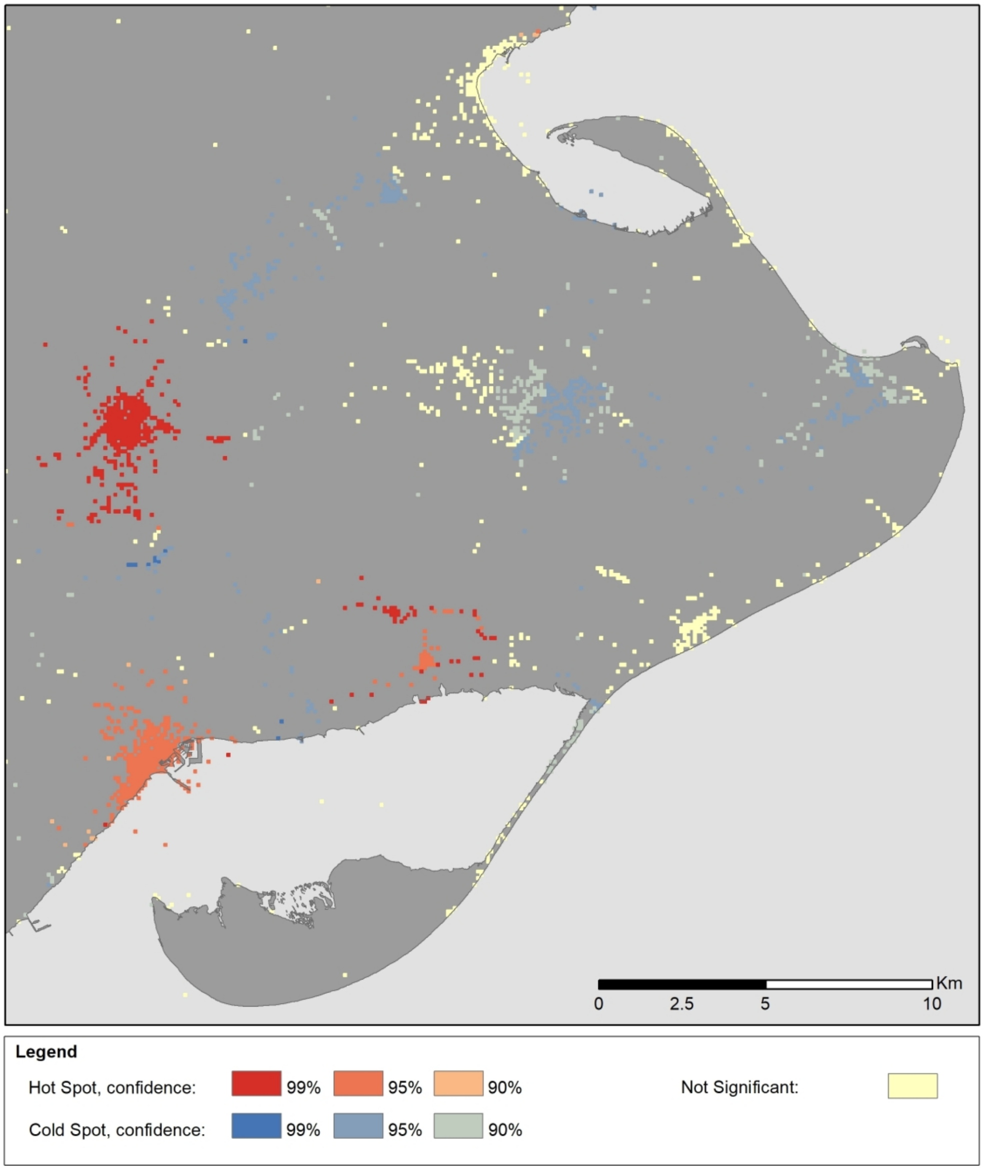

Figure 5. Start/finish plots. A hotspot analysis reveals start/finish plots less frequented on the northern side than the southern side of the Delta. For more detail see Figure 6.

200x237mm (300 x 300 DPI)

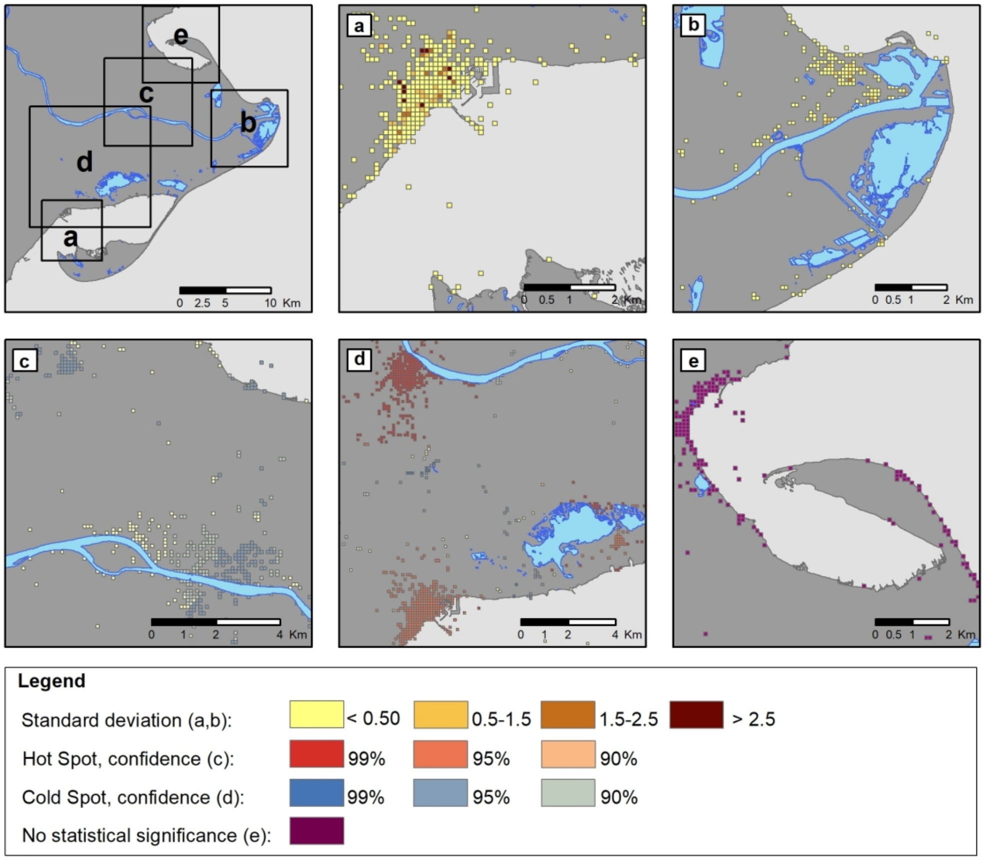

Figure 6. Start/finish plots. Standard deviation is clustered at lower values, while high values are concentrated in just a few plots (a, Sant Carles de la Ràpita, and b, Riumar). A hotspot analysis reveals start/finish plots less frequented on the northern (c) side than the southern side of the Delta (d), and up to a third of plots with no statistical value (e).

182x160mm (300 x 300 DPI)

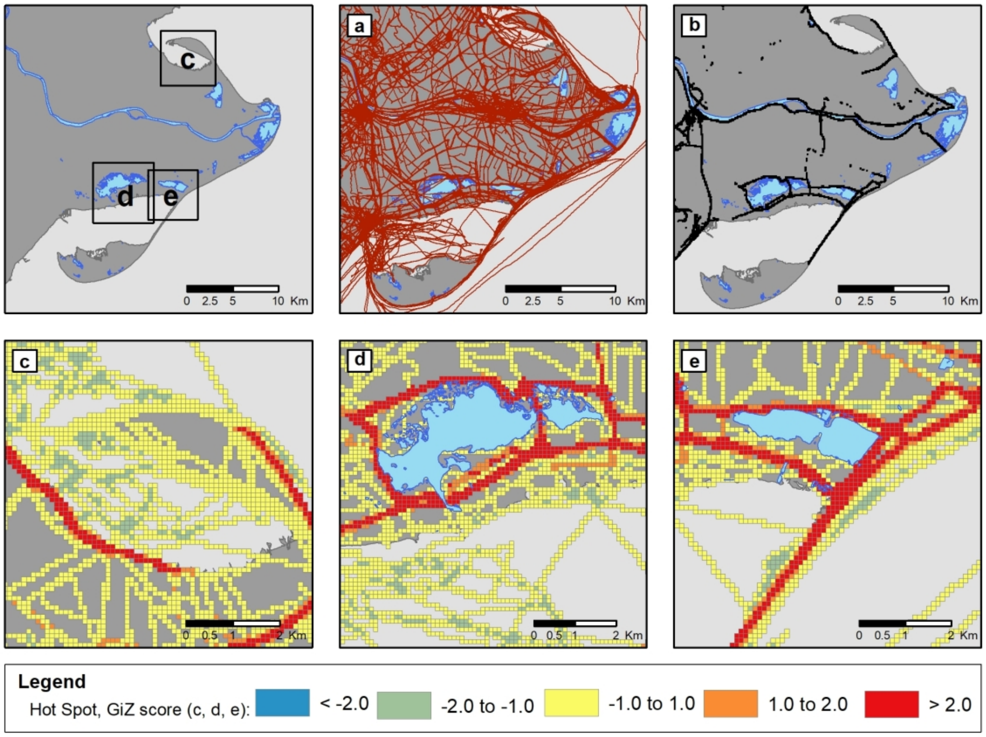

Figure 7. Over 5,000 tracks cover the area studied (a), but less than 1% have significance in a cluster analysis (b). A hotspot approach confirms extremely concentrated scores (c, northern shoreline, d, the Encanyissada Lagoon, and e, the Tancada Lagoon and southern shoreline).

182x138mm (300 x 300 DPI)

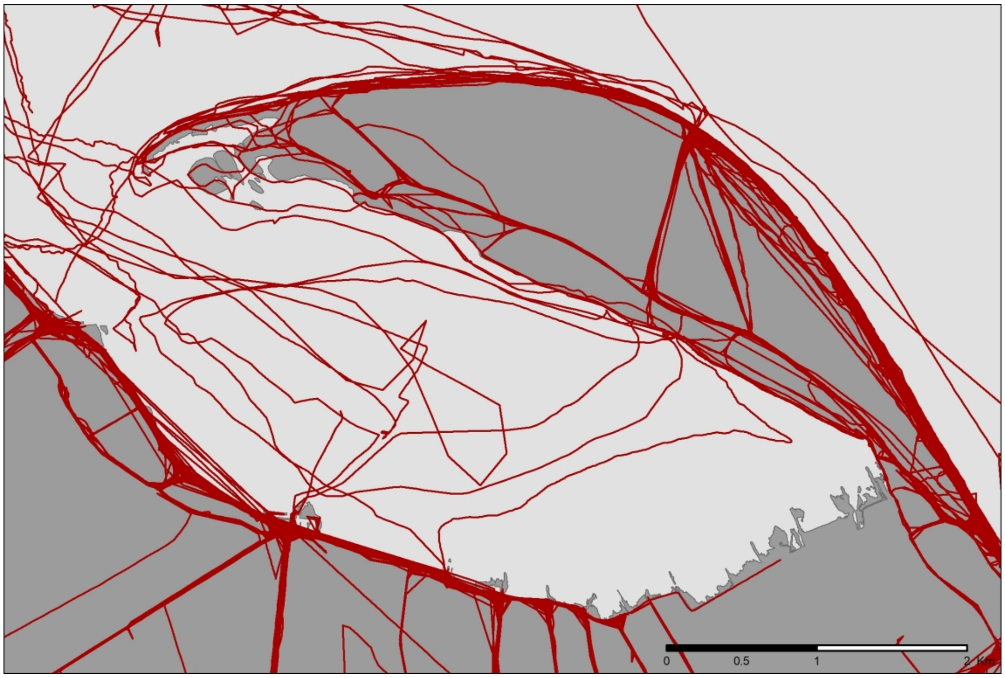

Figure 8. Some tracks encroach onto special protected areas, such as the Fangar Peninsula.

179x120mm (300 x 300 DPI)

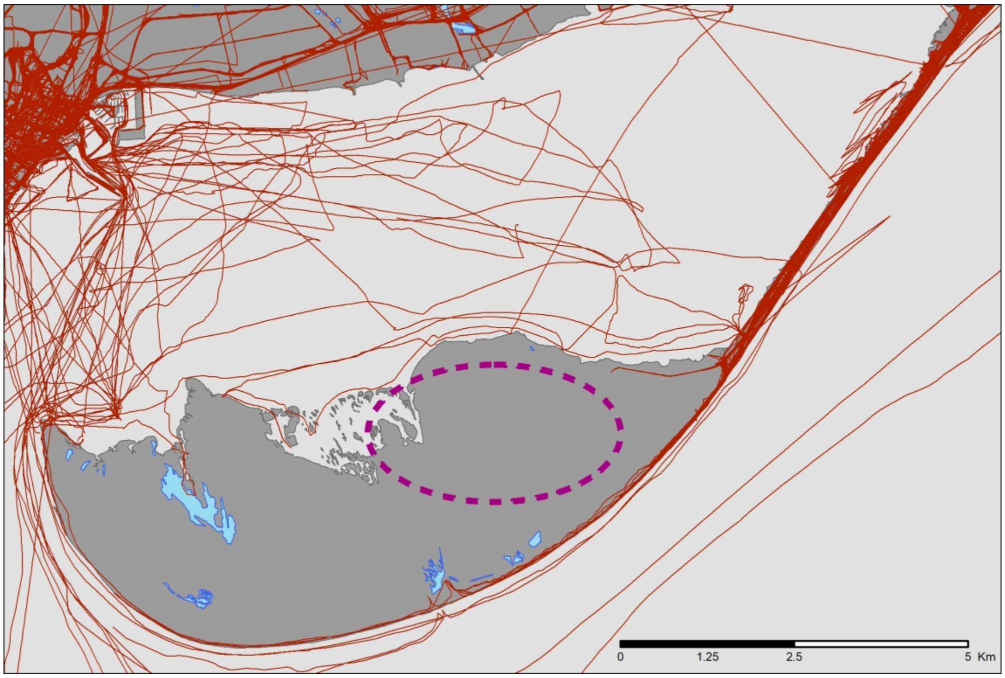

Figure 9. Tracks encroaching onto the Banya Peninsula, a reserve where flamingos breed. The breeding area approximately matches the dashed area.

179x120mm (300 x 300 DPI)

	Name	Difficulty	User	Km	Type
1	Ponds path	easy	foot/bike/horse	26.00	circular
2	Ecomuseum-Ebro mouth	easy	foot/bike/horse	32.00	circular
3	Ecomuseum-Fangar bay	easy	bike	17.00	circular
4	Aldea-Camarles-Ampolla	easy	bike	24.30	circular
5	Family path	easy	foot/bike	6.9/9.8	circular
6	Amposta-Ullals-Casa de Fusta	easy	foot/bike	27.40	circular
7	Sant Carles de la Ràpita-Casa de Fusta	easy	foot/bike	24.4/25.8	one way/circular
8	Towpath	moderate	foot/bike	43.20	one way
9	Riverine path	easy	foot/bike	12.80	one way
10	Bassa de les Olles	easy	foot/bike	3.59	one way/circular

Table 1. Official trails in the Ebro Delta

id	name	near	tr	uid	uom	uomslo	distanc	slope	skill	lon	lat	author	url_trac	trail_ra	viewed	downlo
370277	Un tom	Ampos	0.8689	17220	km	m	94,65	27	2	0.5819	40.709.	joanbtt	https://	84	6552	340
818158	Platja d	Riumar ,	0.7869	46408	km	m	28,97	14	1	0.8422	40.727.	Germin	https://	78	1735	122
133468	160517	L'Amp	0.7883	10646	km	m	55,57	29	1	0.6993	40.800.	GERC	https://	77	528	24
time_m	coord	upload	upl_y	upl_m	upl_d	upl_we	record	record	num_p	loop	track_t	elevati	elevati	elevati	avg_rat	num_ra
380	2605	08/12/2	2012	12	8	6	2012	12	27	Yes	Mounta	17	27	27	4.84	2
189	1136	03/11/2	2014	11	3	1	2014	11	45	Yes	Mounta	9	14	14	5	1
290	1589	17/05/2	2016	5	17	2	2016	5	42	Yes	Mounta	6	28	28	5	1

Table 2. Items considered in the geodatabase derived from webscraping.

		Circular		Perceived difficulty				Distance (km)				
		yes	no	easy	moderate	difficult	expert	0-10	10-25	25-50	50-100	>100
Number	5,323	3,186	2,137	3,413	181	16	6	1,409	1,320	1,687	784	123
Percentage	100	59.85	40.15	64.12	3.40	0.30	0.11	26.47	24.80	31.69	14.73	2.31

Table 3. Track characteristics: type, difficulty and length