land

MDPI

Article

Can We Foresee Landscape Interest? Maximum Entropy Applied to Social Media Photographs: A Case Study in Madrid

Nicolas Marine [1,*], Cecilia Arnaiz-Schmitz [2], Luis Santos-Cid [2] and María F. Schmitz [2]

[1] Department of Architectural Composition, Escuela Tecnica Superior de Arquitectura de Madrid, Universidad Politécnica de Madrid, 28040 Madrid, Spain
[2] Department of Biodiversity, Ecology and Evolution, Universidad Complutense de Madrid, 28040 Madrid, Spain; caschmitz@ucm.es (C.A.-S.); lusant02@ucm.es (L.S.-C.); ma296@ucm.es (M.F.S.)
* Correspondence: nicolas.marine@upm.es

Abstract: Cultural Ecosystem Services (CES) are undervalued and poorly understood compared to other types of ecosystem services. The sociocultural preferences of the different actors who enjoy a landscape are intangible aspects of a complex evaluation. Landscape photographs available on social media have opened up the possibility of quantifying landscape values and ecosystem services that were previously difficult to measure. Thus, a new research methodology has been developed based on the spatial distribution of geotagged photographs that, based on probabilistic models, allows us to estimate the potential of the landscape to provide CES. This study tests the effectiveness of predictive models from MaxEnt, a software based on a machine learning technique called the maximum entropy approach, as tools for land management and for detecting CES hot spots. From a sample of photographs obtained from the Panoramio network, taken between 2007 and 2008 in the Lozoya Valley in Madrid (Central Spain), we have developed a predictive model of the future and compared it with the photographs available on the social network between 2009 and 2015. The results highlight a low correspondence between the prediction of the supply of CES and its real demand, which indicates that MaxEnt is not a sufficiently useful predictive tool in complex and changing landscapes such as the one studied here.

Keywords: cultural ecosystem services; social media; geotagged photographs; maximum entropy models; MaxEnt

Citation: Marine, N.; Arnaiz-Schmitz, C.; Santos-Cid, L.; Schmitz, M.F. Can We Foresee Landscape Interest? Maximum Entropy Applied to Social Media Photographs: A Case Study in Madrid. *Land* **2022**, *11*, 715. https://doi.org/10.3390/land11050715

Academic Editor: Hannes Palang

Received: 20 April 2022
Accepted: 8 May 2022
Published: 10 May 2022

1. Introduction

In the last ten years, there has been an increase in studies using social media geotagged photos to analyze both people's perception of their lived environment and their behavior in it [1–3]. These photographs, and their accompanying information, have opened up the possibility of quantifying landscape values that have been difficult to measure until now, especially those related to cultural ecosystem services (CES). CES can be defined as "non-material benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation and aesthetic experiences" [4] (p. 58).

Despite their relevance, CES evaluation remains disregarded and poorly understood in comparison with the other material ecosystem services [5]. The socio-cultural preferences of the various stakeholders who enjoy a landscape are intangible and challenging to assess. Consequently, researchers usually resort to interviews, questionnaires, participatory mapping and focus groups [6]. Yet, the growth of social media's users and, particularly, platforms where people post geotagged photographs have provided us with a large amount of data on landscape perception. Numerous methods have unfolded to spatially define CES, as well as to characterize and to visualize them [7]. The more common tend to bring together quantitative and qualitative analysis, frequently combining land cover maps with image cluster analysis and automated image recognition [8–13].

Recently, a new line of research has been opened that consists of determining areas that can potentially provide CES from the current distribution of geotagged photos. For this purpose, distribution modeling software is used to identify degrees of significance of different environmental variables in relation to the photographs and to identify potential hotspot areas of CES. For example, well-known models, such as the Integrated Valuation of Ecosystem Services and Tradeoffs (InVEST), have developed specific applications to predict recreation and tourism hotspots and future patterns of use from social network photographs [14]. It is more common, however, to use the open-source software MaxEnt in this operation, a maximum entropy modeling software oriented, in principle, to biologists to predict the distribution of species from current presence data. The use of MaxEnt in relation to social appreciation of the environment is not new. The SolVES (Social Values for Ecosystem Services) program, developed by the U.S. Department of the Interior, already incorporates the use of maximal entropy modeling software to cross-reference social values (aesthetic, biodiversity or recreational) with explanatory environmental variables [15]. By crossing these data with environmental variables, it provides a map of suitable habitats where these species may be present [16]. In the present case, the species presence layer is replaced by that of georeferenced photographs to calculate the probability of a photograph being taken in a certain place. In short, the potential supply of certain recreational services.

In fact, so-called affective computing has been making use of maximum entropy models also in close relation to social networks and CES [17–19]. As a result, there has been a growing number of investigations using georeferenced photographs and MaxEnt to determine the potential supply of CES in a given geography. The advantage offered by this software is that the presence data are sufficient to model the potential distribution. Furthermore, by means of jackknife resampling, it us allows to establish the degrees of relationship between the presence of the photographs and the different environmental variables [20]. Specific applications of the study of geotagged social media photos using MaxEnt are varied, both in terms of task and procedure. For example, Richards and Friess [11] use it to study coastal mangrove forest habitats and potential visitor interest based on distances to access points, communication infrastructure and viewpoints. Yoshimura and Hiura [21] apply it to the island of Hokkaido by comparing an area of demand based on viewsheds using geotagged photos as a viewpoint and a supply provided by cross-referencing data with MaxEnt. Clemente et al. [22] apply it to a natural park in Portugal and study proximity indices at different variables, assuming that the greater the distance, the lower the attractiveness of a biophysical component or infrastructure. More recent studies have increased the scale of application to the whole European continent [23] or have added the location of historical and cultural sites to the environmental variables, which brings a heritage reading to the potential interest of the landscape [24].

These references coincide in that they obtain the georeferenced photographs from the Flickr social network, mainly because of the ease with which they can be downloaded from its API and because of the data it contains. The number of photos used in each study varies greatly, from 250 to almost 7 million. This depends on the size of the study area, which, as can be seen, is also very varied, ranging from small, protected areas to entire continents. The photos are usually classified according to different CES or according to the elements photographed, either manually or automatically (using, for example, Google Cloud Vision). The environmental variables with which they are crossed are generally the same: land use and land cover, geomorphology and co-communication infrastructures. Although other variables such as heritage assets are often added. The final result is highly dependent on these variables.

The references cited above use MaxEnt as a CES supply potentiality tool, but do not question the effectiveness of this software itself. In general, they all validate the quality of the models using the area under the receiver operating characteristic curve as a parameter [20]. If this area, or AUC, is close to 1, the prediction is considered perfect, and if it is below 0.5, the model prediction is considered to be as good as random. Therefore, this method is validated by a result provided by the software itself and not on the basis of external checks. A review of the literature shows that, although several of the references include in the discussion a critique of social media photographs as univocal representations of the population's interest [12,22,24], they do not criticize MaxEnt as a mechanism for predicting their potential distribution.

Based on these arguments, this paper takes as its starting point a criticism that The Natural Capital Project [14] had already made of predictive models: that they require assuming that people's responses to environmental variables will not change over time. That is, the use of MaxEnt presumes that, in the future, people will continue to be attracted or repelled by the same factors as today, and even that these factors will remain unchanged over time. A question arises here: how valid is MaxEnt really in predicting potential interest in specific CES? To answer it, this research consists of comparing the actual evolution of social media geotagged photos with the prediction that MaxEnt would have made based on passed information. To do this, we use the photographs uploaded to the Panoramio social media network between 2006 and 2015 in the Lozoya Valley, a complex landscape north of Madrid. The case study has been selected because it combines natural and cultural values, because it is of great tourist interest and because previous work has shown changes in the valuation of CES in recent decades [25].

Here, we question the validity of MaxEnt as a predictive tool for landscapes as complex as those of the Lozoya Valley. The objectives of the study are the following:

i. To determine the validity of social media geotagged photos as a basis, and of MaxEnt as a tool, for predictive studies of future landscape users' behavior.
ii. To determine the differences between the quantification of the future spatial distribution of geotagged photographs and their real qualitative changes.
iii. To propose a comprehensive approach to the actual complexity of the photographs uploaded by users to social networks to assess future CES interest.

2. Materials and Methods

2.1. Research Area

The study area, selected within the Lozoya Valley, covers approximately 776 km^2, belonging to 25 of the 30 municipalities that constitute the valley. The Lozoya Valley (Figure 1), is located in the northern Sierra de Guadarrama in the Lozoya river basin (Madrid region, Spain). The area's main connection with the rest of the region is through a single highway (A-1 Route). The Lozoya Valley has its highest elevation at Peñalara peak (2428 m a.s.l.) and the lowest elevation in the adjoining area of the Lozoya rivers (100 m a.s.l.) and includes 30 municipalities. The Lozoya Valley is a heterogeneous landscape with forest, settlements, water bodies such as reservoirs and a mosaic of traditional land uses containing pastures, meadows, hedgerows, ash groves and riparian forests, all of which are well preserved in most cases [26]. Over the centuries, the valley has come under different land uses and rural activities which have shaped the landscape and their traditional way of living along the centuries that have resulted in a region of great socio-ecological value. Currently, this heritage landscape is under several categories of protection.

Thus, the Lozoya Valley is within the boundaries of the Sierra de Guadarrama National Park (established in 2013) and the Sierra del Rincón Biosphere Reserve (2005), and it also belongs to the European network of protected sites Natura 2000. Several areas inside the valley also fall under other types of recognitions: (i) The Montejo de la Sierra beech forest (Natural Site of National Interest, 1974, and subsequently UNESCO's World Heritage Site, 2017); (ii) The Monastery of El Paular (Historical-Artistic Monument of Spain, 1876), and (iii) the Neanderthals Valley (Cultural Interest Asset, 2004). Recently, the High Valley of the Lozoya River has been proposed as a model of Heritage Cultural Landscape to UNESCO by the Spanish National Plan of Cultural Landscape [27]. Because of this, the area is a touristic hotspot appreciated by visitors to the Madrid region [26]. However, recent studies have shown that increasing tourism and conflicting management legislation caused a rurality loss and an urban sprawl throughout the territory, transforming the ancient agropastoral landscape into a wilderness [28]. The comparison of surveys conducted with visitors in

2007 and 2017 shows that these changes have had an impact on their way of understanding the landscape of the Lozoya Valley: from valuing the cultural components more, to now valuing the "naturalistic" ones more [26].

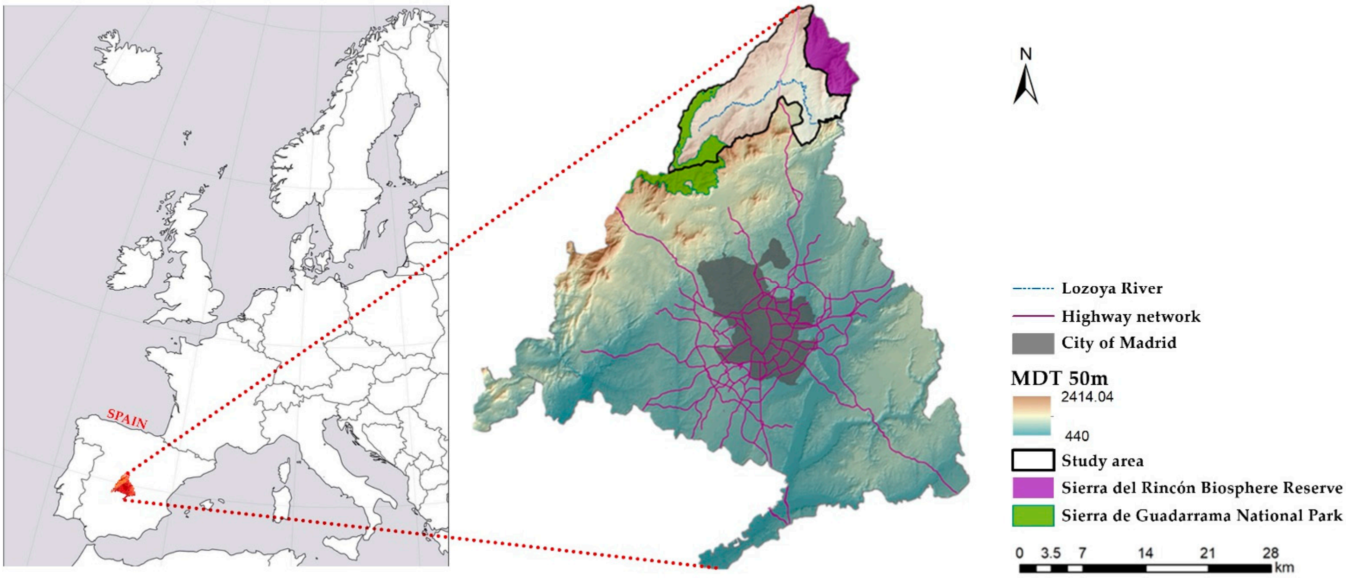

Figure 1. Map of the Madrid Region with the Lozoya Valley area of study in the north.

2.2. Materials

The basis of the study is the georeferenced photographs uploaded by users to the Panoramio website between 2007 and 2015. Panoramio was a website specialized in sharing georeferenced audiovisual material accessible as a layer in Google Earth and Google Maps. It was active between 2005 and 2016, when it closed, although the layer on Google Earth was available until January 2018. Panoramio is very similar to Flickr but, rather than a social network per se, it is considered to be a means of sharing photos and videos by users [24]. Sometimes both Panoramio and Flickr have been used to study the CES of a place [29] and have been found to provide similar patterns of landscape values, at least on the European continent [30]. Panoramio photos, as opposed to Flickr photos, are usually obtained manually [31], although they can also be obtained through APIs [32]. In the case of this article, the professional services of a company were hired to bulk download all the photographs uploaded to Panoramio in the Region of Madrid between 2007 and 2015.

As a result of this operation, a list of 54,956 photographs was obtained, half of which were located in the metropolitan area of Madrid. A selection of those located in the municipalities that make up the study area reduced this number to 3192 photographs located in the Lozoya Valley. Given that the references of studies of this type first make a classification of the photographs based on the elements that appear in them, only those photographs that could still be located on-line were taken for the sample. Before Panoramio disappeared completely, the Mapio website (https://mapio.net/, accessed on 7 May 2022) made an extensive transfer of its collection. Some of which can still be located today. An automated search made it possible to locate and download a total of 1728 photographs that serve as a starting point for the study. The attributes associated with the photos include the username, upload date, title and hashtags provided by the user, and the location of the photo. The biggest difference between the photos refers to the years in which they were uploaded (Figure 2). Subtracting 91 that do not contain a date, most of the photos in the sample were uploaded in 2007 (142), 2008 (752) and 2009 (442). The year with the least number of photos is 2014 (18).

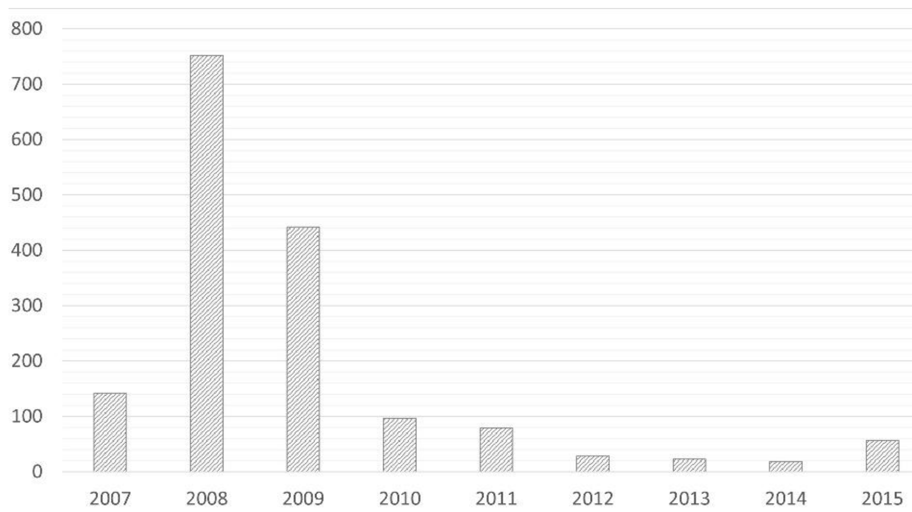

Figure 2. Number of photographs per year in the Panoramio sample.

2.3. Methodology

This research compares the prediction of potential demand of CES in an area with the actual demand. To do so, we use social media photos instead of species occurrences in a species distribution modeling software for understanding the spatial distribution of people preferences for CES. Actual demand is measured based on photo density. To correlate both datasets this methodology followed a four-step process: (1) database preparation and filtering, (2) variable selection and MaxEnt modeling, (3) modeling of the actual demand, and (4) elaboration of a correlation matrix.

2.3.1. Georeferenced Database of Social Media Photographs

First, the images are classified according to a series of categories linked to CES and the most represented elements (Figure 3). The classification is done manually and by two different researchers, and the differences are then compared by a third one. The method thus follows other manual classification methods used in similar studies [11,22]. In this way, each photograph is assigned one of the categories in Table 1.

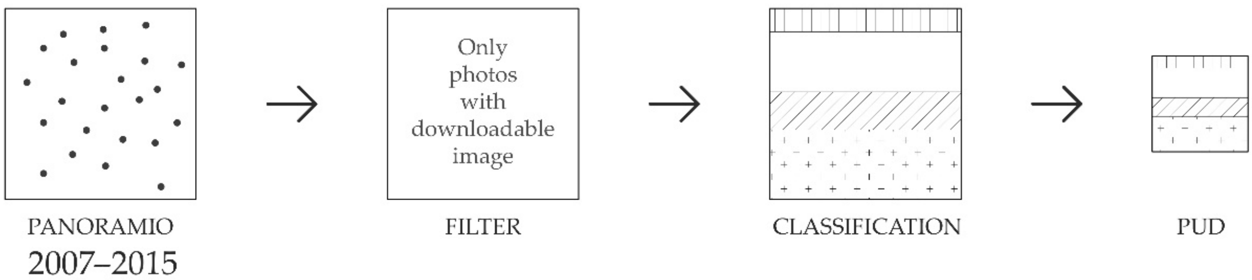

Figure 3. Diagram of photograph database processing and sampling.

Table 1. Photographs categories.

Category	Description
Natural system	Majority presence of flora and fauna in a wild state.
Urban system	Majority presence of architectural and urban elements
Rural system	Majority presence of agrosilvopastoral elements.
Water bodies	Majority presence of aquatic elements, very common in the zone.
Recreational activities	Presence of people engaged in sports or walking activities
Cultural activities	Presence of museums, monuments, food or typical products.

The categories are divided into systems (natural or naturalistic tendency, urban tendency and rural tendency) and activities linked to the CES (recreational and cultural). The presence of bodies of water was added as a category in its own right after checking the frequency with which they appeared in the photographs. The Lozoya Valley is characterized by a hydraulic system that combines rivers and artificial elements such as reservoirs. Since it is difficult to differentiate whether a user is photographing one of these bodies of water on the basis of whether it is considered natural or artificial, the presence of water is taken as a characteristic aesthetic value of the valley.

Once classified, a sieve is applied to the collection of images to remove unwanted tendencies. This is common in research related to social network photographs, since there are usually users who upload several photos on the same day, which can generate biased deductions. For this purpose, the PUD (Per User Day) method is used, a form of screening that avoids this problem based on randomly selecting, from the initial sample, one photo per user per day [7]. Given that there are also methodologies that consider this type of reduction in the original sample to be negative [33], and that this paper is oriented towards a critique of an established method, the analysis is carried out with two samples: the original, with 1728 photographs, and the PUD, with 709 photographs.

2.3.2. Data Processing and Confrontation

Once the photographs have been classified and screened, we proceed to their analysis using MaxEnt. To do this, we divided the sample of photographs into two periods: one used to develop the predictive model (Base Demand Sample) and the other with the actual distribution of the photographs (Actual Demand Sample). For the first period, photographs from 2007 and 2008 are selected, which account for approximately half of the photographs. For the second period, we take the photographs between 2009 and 2015. The purpose is to test how close the MaxEnt would have been to determine CES demand in the future. This division of the database makes it possible to develop a predictive model from the 2007–2008 data and check whether it matches the actual demand up to 2015 (Figure 4).

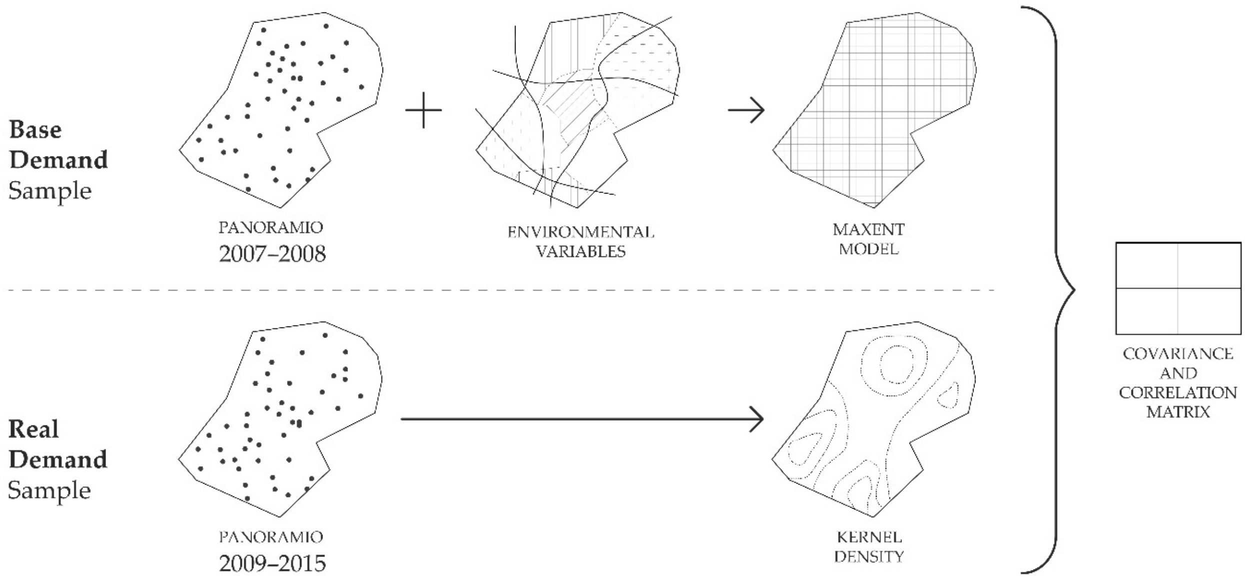

Figure 4. Modeling and correlation diagram.

Therefore, we use as species the point data from the Base Demand Sample and transfer it to MaxEnt 3.4.4 [20]. To this, we add a series of environmental variables: (1) 2006 land cover; (2) altitude; (3) average atmospheric temperature; (4) distance to roads; and (5) distance to cultural assets. Except for the last one, the data come from open sources (Table 2). All the variables are continuous, and the proximity indices have been calculated using the Kernel tool in ArcMap 10.5.1. The MaxEnt model provides an ACISS file that will be used for comparison. In addition, it provides a measure of goodness of fit that quantifies how closely the model is concentrated around occurrences.

Table 2. Environmental variables, source and processing.

Variable	Source	Processing
Land Cover	Corine Land Cover 2006 (https://land.copernicus.eu/pan-european/corine-land-cover, accessed on 7 May 2022)	Unificación de categorías
Altitude	SDI of Spain (https://www.idee.es/, accessed on 7 May 2022)	MDT as downloaded
Average atmospheric temperature	SDI of Spain	Kernel from medium temperature (station points)
Distance to roads	SDI of Spain	Kernel from road lines
Distance to cultural assets	Madrid Heritage Information System [34]	Kernel from cultural asset points

On the other hand, we study the real evolution of the photographs from the Actual Demand Sample (Panoramio photos uploaded between 2009–2015). Here, we conducted a photograph points density study using the Kernel tool in ArcMap 10.5.1. With these two layers (MaxEnt prediction from 2007–2008 data and density of photographs between 2009–2015) we performed a multivariate analysis using the Band Collection Statistics tool (ArcMap 10.5.1 Spatial Analyst). This tool allows us to confront the variation of two or more overlapping rasters. When requested, it computes covariance and correlation matrices. The result matrix presents the variances of all raster bands along the diagonal from the upper left to lower right and covariances between all raster bands in the remaining entries.

In our case, the final correlation matrix provides a correlation coefficient between the MaxEnt layer (prediction) and the density layer (actual demand). The proximity of this coefficient to 1 indicates that both layers vary similarly. That is, the potential and actual demand intentions not only coincide at certain points but are distributed equally. Therefore, we take the proximity of this coefficient to 1 as the measure of success of the MaxEnt predictive model. The process shown in Figure 4 is carried out with both the original sample and the PUD sample. From the one that is closest to 1 in the correlation matrix, we check the correlation of each of the categories.

3. Results

3.1. Photograph Samples

The classification of the photographs reveals large differences between categories in both the original and the unbiased or PUD sample (Table 3). In the former, photographs tend to be of elements related to natural systems (455) and bodies of water (426) followed by photographs related to urban environments (378). Photographs related to cultural activities are the least present (68). As for the PUD sample, water bodies are the most photographed (195), although closely matched by natural systems (190). After these categories, recreational activities (123) and urban systems (128) are closely matched. Cultural activities are again the least represented (29).

Table 3. Number of photographs in each sample and category.

Sample	Natural System	Urban System	Rural System	Water Bodies	Recreational Activities	Cultural Activities
Original sample	455	378	114	426	282	68
PUD sample	190	128	42	195	123	29

Percentage comparison of the two samples reveals very little difference between them (Figure 5). The representation of natural systems changes by only one percentage point when removing the bias of the original sample, and water bodies vary from 25% to

28%. Cultural activities, recreational activities and rural systems barely vary. The greatest variation from one sample to another occurs among the photos of urban systems (from 22% in the original sample to 18% in the PUD).

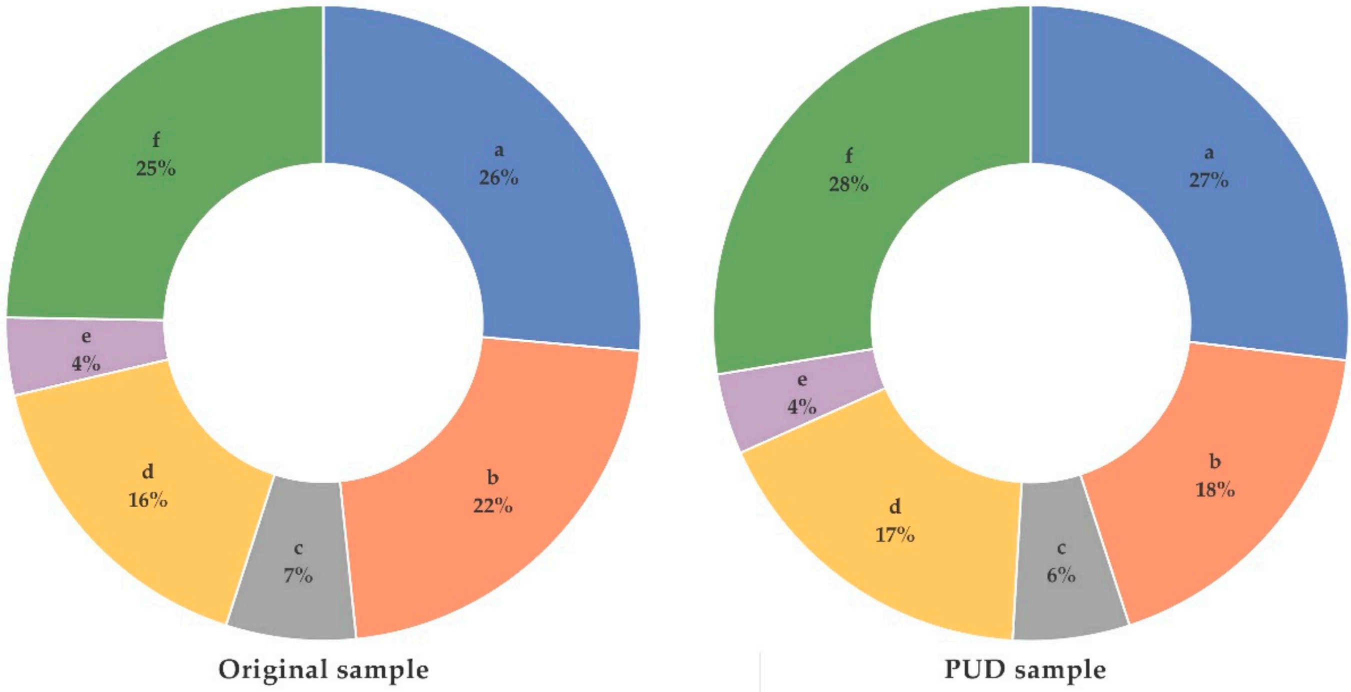

Figure 5. Percentage of each category in each sample: (**a**) natural systems, (**b**) urban systems, (**c**) rural systems, (**d**) recreational activities, (**e**) cultural activities, and (**f**) water bodies.

3.2. MaxEnt Modeling

Although there is little variation between the classes of photographs, the MaxEnt models performed with each sample do change more (Figure 6). The potential demand model from the original sample predicts a higher intensity in the southwest and northeast extremes. In between, the potential is also intensified by the different settlements located along the valley. The roads connecting them are also highlighted as points of potential interest. When the model is performed from the PUD sample, however, it places more emphasis on the southwestern part of the study area. The rest is shown as an area of low potentiality, except for the northeast zone, which is more intense than the center, but does not compensate for the more intense areas. The AUC, the standard measure of model reliability, is 0.755 in the case of the original sample and 0.908 in the case of the PUD sample. In both cases, it is much higher than 0.5, indicating reliability of the model. In the PUD sample, it is very close to 1.

The jackknife values also reveal several differences (Figure 7). In the MaxEnt model made from the original sample, the CES demand potential depends on the distance to roads and land uses. This corresponds to the map itself, where the surroundings of towns and infrastructure for road traffic are highlighted. The other variables (altitude, proximity to cultural assets and temperature) do not seem to have much influence on the prediction. In contrast, the model performed on the PUD sample shows a strong dependence on altitude and temperature. If we take into account that distance to roads and land use are the least influential variables, we observe that the change from the original sample to the PUD sample provides opposite models in MaxEnt.

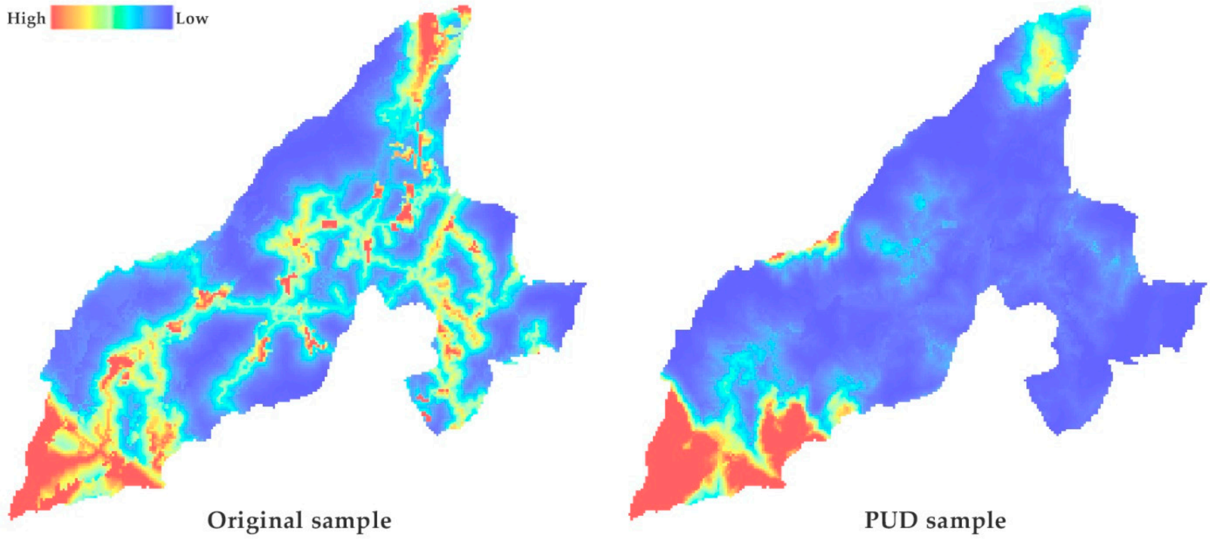

Figure 6. MaxEnt models for each sample.

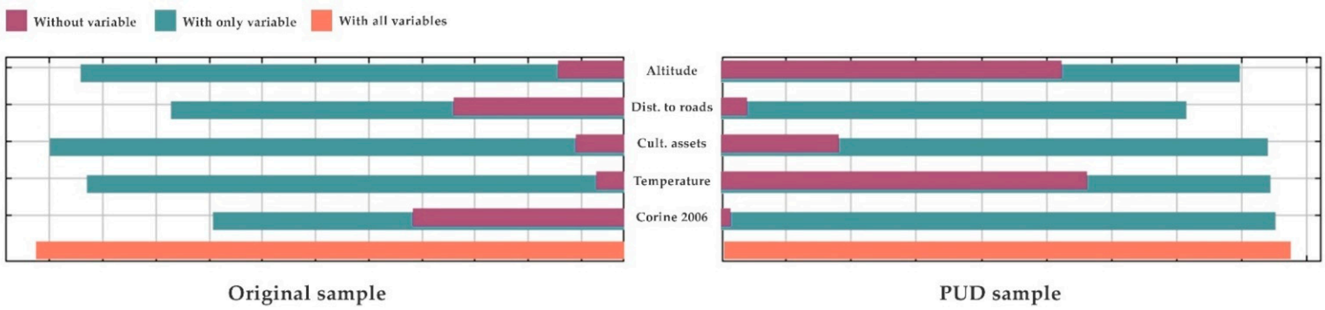

Figure 7. Jackknife test applied to environmental values for each of the samples.

3.3. Actual Demand and Correlation

The photo density from the actual demand provides similar results for both the original sample and the PUD (Figure 8). In both cases, there is a high concentration of photographs in the southwest area, although there is a proportionate distribution of lower concentrations throughout the rest of the Lozoya Valley. It is also noteworthy that in both samples, the concentration in the southeastern area is significant. The correlation matrix between these maps and their corresponding MaxEnt models gives a value of 0.30940 in the case of the original sample and 0.435 in the case of the PUD sample. That is, in both samples there is a low correlation (<0.5) between the actual demand and the potential demand determined by the MaxEnt. In other words, if MaxEnt had determined a potential evolution in 2008 from the information available in the Panoramio network, it would not have provided a model corresponding to the evolution of the photographs that actually took place.

Since the closest correlation occurs in the photographs of the PUD sample, we study the correlation of the photographs on the basis of the different categories (Table 4). It is observed here that the highest correspondence occurs in natural systems (65%) and outdoor activities (64%), being the only ones with a correspondence in the variation of more than 50%. Urban systems, however, maintain a negative correlation, meaning that the actual demand has more variation in intensity than the MaxEnt model.

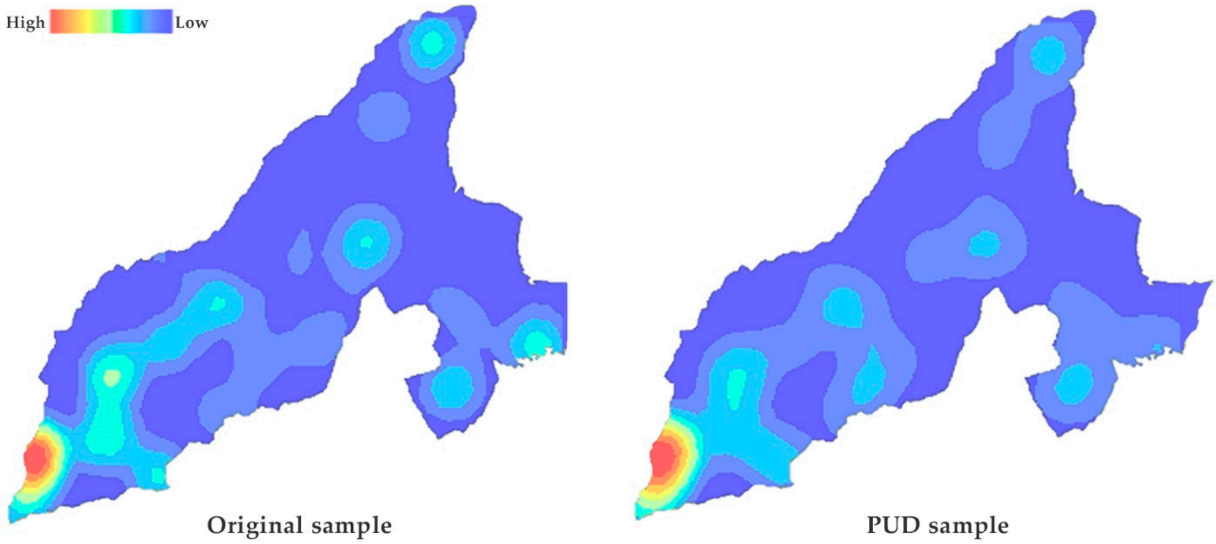

Figure 8. Density of photographs taken between 2009–2015 in both samples.

Table 4. Correlation between MaxEnt model and actual demand by category in PUD sample.

PUD Sample	Natural System	Urban System	Rural System	Water Bodies	Recreational Activities	Cultural Activities
Correlation Base Demand and Actual Demand	0.651	−0.07	0.204	0.246	0.642	0.244

4. Discussion

This study tests the effectiveness of MaxEnt predictive models, based on a machine-learning technique called maximum entropy approach, as a tool for land management to detect hot spots of CES. With this objective, we have evaluated this software through a set of photographs categorized according to their content in two different periods of time. From the photographs taken in the study area during the first period, the potential of the landscape for recreational use was estimated using MaxEnt. This estimate was correlated with the photographs of the second period, obtaining a series of results that are discussed below.

4.1. Photography Samples

The article presents the comparison of a predictive model of CES demand with the actual evolution of such demand. The demand is identified here with the georeferenced photographs uploaded to the Panoramio network by various users between 2007 and 2015. Several studies use MaxEnt to estimate the predicted future demand from social network photographs [1]. In contrast to the articles cited above, here we use photographs from the Panoramio network instead of Flickr, as both have been identified as similar [24]. In addition, some research defends the use of Panoramio over Flickr because it is a better measure of the aesthetic value of a place, since its contents were more focused on landscape and environment [35]. This is evident in the low number of photographs in our own sample devoted to cultural topics such as food, monuments or ethnographic elements. However, our sample of photographs reflects a high number of views of natural, urban and aquatic landscapes.

The fact that both the original sample of photographs and the PUD maintain a similar percentage classification by category indicates a certain consistency in the type of information uploaded by users to this type of network. This means that, at least thematically, the screening of photographs on the basis of user and day does not have an influence. However, the development of the MaxEnt model with one sample and with another provides opposite results. This means that spatially the PUD does have an impact. This allows us to refine certain explanations of this screening method [7], as it subtracts spatial bias, but not thematic bias. On the other hand, most of the photographs focus on natural or aquatic elements. This, in line with certain criticisms of samples made up of georeferenced photographs, could mean that, regardless of the type of sample, it will always be biased by an interest in photographing non-anthropized landscapes [12]. Specifically, in the study area, this is greatly influenced by the higher presence of photographs in the protected park areas (National Park in the southwest and Biosphere Re-reserve in the northeast).

4.2. MaxEnt Models

Neither of the two MaxEnt models developed bore any resemblance to the actual evolution of the distribution of the photographs except where a greater number of samples were concentrated. In spite of this, both models had AUC parameters above 0.5 and even close to 1. Interestingly, in most of the literature studied, the outcome of the models is closely linked to the environmental variables that are incorporated into the MaxEnt [11,21,24]. In our case, however, the dependence of the variables has changed greatly depending on the sample type. This contradicts the claim that MaxEnt results do not depend on the point sample size [24]. In reality, this is only true if the sample, when reduced, maintains a similar spatial distribution.

As can be seen in the two models developed (Figure 5), the PUD has reduced the number of photographs taken on the roads and settlements and, therefore, these factors are no longer important for the calculation of the prediction. What this shows is that MaxEnt is a program that is heavily influenced by high sample concentrations. A place, however small it may be in relation to the rest, will be decisive in the model if a very high number of points are located there. This is why some studies incorporate a percentage of points randomly distributed over the studied area [21]. The dependence of MaxEnt on the type of sample treatment is decisive. There are research methods that intentionally do not want to reduce the photographs to PUD, if, for example, researchers want to study the widest possible variety of images [33]. This means that different methods of spatial study from social media photographs would get different results in MaxEnt, since they would treat their samples differently.

Finally, several studies have shown certain changes in the public's interest in the values of the Lozoya Valley, with a tendency to value its wild aspects more highly [26,28]. Hence, the greatest overlap in correlation occurs in natural systems and recreational activities. However, the MaxEnt model predicts that there will be interest in infrastructure and settlements in the original sample. This model, therefore, errs in that it lacks sufficient complexity to adapt to changes in population interest. This is consistent with the criticism made by the authors of the InVEST model of predictive models in general [14]. In the case of the PUD sample, where MaxEnt predicts interest in locations at a certain altitude and temperature, the model falls short in its prediction, since it is based on a much smaller concentration of photographs.

5. Conclusions

In this article, we questioned the validity of MaxEnt as a predictive tool for landscapes as complex as those of the Lozoya Valley. The following conclusions can be drawn in line with our objectives:

i. Photographs from social networks are valid for predictive modeling as long as they are at sites that remain unchanged over time. If the configuration of the sites or the interest of the people changes, a present sample is invalid for determining future interest. On the other hand, MaxEnt is a program that allows us to determine with some accuracy to which spatial variables a certain sample of photographs is related, but it is very dependent on the concentration of these photographs. From the same sample, treated differently, it is possible to obtain models that are absolutely opposite. Comparison with the real evolution of the distribution of photographs shows that in complex and changing landscapes, MaxEnt is not useful as a predictive tool.

ii. There is a difference between the quantification of the future spatial distribution of geotagged photographs and their qualitative changes. MaxEnt establishes locations of potential interest of the photographs independently of the photographed element. The correspondence with the actual evolution of the photographs varies greatly depending on each category. For some categories, the model is closer in its prediction, but for others, the prediction is opposite to the actual evolution.

iii. This paper opens a comprehensive approach to the actual complexity of the photographs uploaded by users to social networks to assess future CES interest. Studies using MaxEnt to model potential demand can use other testers besides the AUC. For example, they can run the predictive model with a portion of the sample and use the most current photographs to check how accurate it is. They can also run the predictive models on a year-by-year basis, adjusting it according to the actual evolution of the photographs in the following year. Taking into account that most of the studies use time ranges of five or more years, this would allow us to establish rectification coefficients from one year to another to improve a global predictive model.

Author Contributions: Conceptualization, N.M. and C.A.-S.; methodology, N.M. and C.A.-S.; software, C.A.-S. and L.S.-C.; formal analysis, N.M., C.A.-S. and L.S.-C.; investigation, N.M., C.A.-S. and M.F.S.; writing—original draft preparation, N.M., C.A.-S., L.S.-C. and M.F.S.; writing—review and editing, N.M., C.A.-S. and M.F.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project LABPA-CM: CONTEMPORARY CRITERIA, METHODS AND TECHNIQUES FOR LANDSCAPE KNOWLEDGE AND CONSERVATION (H2019/HUM-5692), funded by the European Social Fund and the Madrid regional government.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank Ivan Nuñez Espinosa for his valuable help in image categorization.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bubalo, M.; van Zanten, B.T.; Verburg, P.H. Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landsc. Urban Plan.* **2019**, *184*, 101–111. [CrossRef]
2. Toivonen, T.; Heikinheimo, V.; Fink, C.; Hausmann, A.; Hiippala, T.; Järv, O.; Tenkanen, H.; Di Minin, E. Social media data for conservation science: A methodological overview. *Biol. Conserv.* **2019**, *233*, 298–315. [CrossRef]
3. Karasov, O.; Heremans, S.; Külvik, M.; Domnich, A.; Chervanyov, I. On How Crowdsourced Data and Landscape Organisation Metrics Can Facilitate the Mapping of Cultural Ecosystem Services: An Estonian Case Study. *Land* **2020**, *9*, 158. [CrossRef]
4. Leemans, R.; De Groot, R.S. *Millennium Ecosystem Assessment: Ecosystems and Human Well-Being: A Framework for Assessment*; Island Press: Washington, DC, USA, 2003.
5. Cheng, X.; Van Damme, S.; Li, L.; Uyttenhove, P. Evaluation of cultural ecosystem services: A review of methods. *Ecosyst. Serv.* **2019**, *37*, 100925. [CrossRef]
6. Milcu, A.; Ioana, A.; Hanspach, J.; Abson, D.; Fischer, J. Cultural ecosystem services: A literature review and prospects for future research. *Ecol. Soc.* **2013**, *18*, 44. [CrossRef]
7. Zhang, H.; Huang, R.; Zhang, Y.; Buhalis, D. Cultural ecosystem services evaluation using geolocated social media data: A review. *Tour. Geogr.* **2020**, 1–23. [CrossRef]
8. Ghermandi, A.; Camacho-Valdez, V.; Trejo-Espinosa, H. Social media-based analysis of cultural ecosystem services and heritage tourism in a coastal region of Mexico. *Tour. Manag.* **2020**, *77*, 104002. [CrossRef]
9. Dunkel, A. Visualizing the perceived environment using crowdsourced photo geodata. *Landsc. Urban Plan.* **2015**, *142*, 173–186. [CrossRef]

10. Oteros-Rozas, E.; Martín-López, B.; Fagerholm, N.; Bieling, C.; Plieninger, T. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecol. Ind.* **2018**, *94*, 74–86. [CrossRef]
11. Richards, D.R.; Friess, D.A. A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. *Ecol. Ind.* **2015**, *53*, 187–195. [CrossRef]
12. Richards, D.R.; Tunçer, B.; Tunçer, B. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecos. Serv.* **2018**, *31*, 318–325. [CrossRef]
13. Abousaleh, F.S.; Cheng, W.H.; Yu, N.H.; Tsao, Y. Multimodal Deep Learning Framework for Image Popularity Prediction on Social Media. *IEEE Trans. Cogn. Dev. Syst.* **2020**, *13*, 679–692. [CrossRef]
14. Sharp, R.; Tallis, H.T.; Ricketts, T.; Guerry, A.D.; Wood, S.A.; Chaplin-Kramer, R.; Nelson, E.; Ennaanay, D.; Wolny, S.; Olwero, N.; et al. *InVEST User's Guide*; The Natural Capital Project: Stanford, CA, USA, 2014.
15. Sherrouse, B.C.; Semmens, D.J. *Social Values for Ecosystem Services, Version 4.0 (SolVES 4.0)—Documentation and User Manual*; US Department of the Interior, US Geological Survey: Washington, DC, USA, 2020.
16. Elith, J.; Phillips, S.J.; Hastie, T.; Dudík, M.; Chee, Y.E.; Yates, C.J. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **2011**, *17*, 43–57. [CrossRef]
17. Poria, S.; Cambria, E.; Bajpai, R.; Hussain, A. A review of affective computing: From unimodal analysis to multimodal fusion. *Inf. Fusion* **2017**, *37*, 98–125. [CrossRef]
18. Li, Y.; Fei, T.; Huang, Y.; Li, J.; Li, X.; Zhang, F.; Kang, Y.; Wu, G. Emotional habitat: Mapping the global geographic distribution of human emotion with physical environmental factors using a species distribution model. *Int. J. Geo. Inf. Sci.* **2021**, *35*, 227–249. [CrossRef]
19. He, S.; Su, Y.; Shahtahmassebi, A.R.; Huang, L.; Zhou, M.; Gan, M.; Deng, J.; Zhao, G.; Wang, K. Assessing and mapping cultural ecosystem services supply, demand and flow of farmlands in the Hangzhou metropolitan area, China. *Sci. Total Environ.* **2019**, *692*, 756–768. [CrossRef]
20. Phillips, S.J. *A Brief Tutorial on Maxent*; AT&T Research, American Museum of Natural History: New York, NY, USA, 2017; Available online: http://biodiversityinformatics.amnh.org/open_source/maxent/ (accessed on 9 May 2022).
21. Yoshimura, N.; Hiura, T. Demand and supply of cultural ecosystem services: Use of geotagged photos to map the aesthetic value of landscapes in Hokkaido. *Ecosyst. Serv.* **2017**, *24*, 68–78. [CrossRef]
22. Clemente, P.; Calvache, M.; Antunes, P.; Santos, R.; Cerdeira, J.O.; Martins, M.J. Combining social media photographs and species distribution models to map cultural ecosystem services: The case of a Natural Park in Portugal. *Ecol. Ind.* **2019**, *96*, 59–68. [CrossRef]
23. Long, P.R.; Nogué, S.; Benz, D.; Willis, K.J. Devising a method to remotely model and map the distribution of natural landscapes in Europe with the greatest recreational amenity value (cultural services). *Front. Biogeogr.* **2021**, *13*, 1. [CrossRef]
24. Arslan, E.S.; Örücü, Ö.K. MaxEnt modelling of the potential distribution areas of cultural ecosystem services using social media data and GIS. *Environ. Dev. Sustain.* **2021**, *23*, 2655–2667. [CrossRef]
25. Schmitz, M.F.; De Aranzabal, I.; Pineda, F.D. Spatial analysis of visitors preferences in the outdoor recreational niche of Mediterranean cultural landscapes. *Environ. Conserv.* **2007**, *34*, 300–312. [CrossRef]
26. Arnaiz-Schmitz, C.; Santos, L.; Herrero-Jáuregui, C.; Díaz, P.; Pineda, F.D.; Schmitz, M.F. Rural Tourism: Crossroads between nature, socio-ecological decoupling and urban sprawl. *WIT Trans. Ecol. Environ.* **2018**, *227*, 1–9.
27. Cruz, L.; Carrión, A. (Eds.) *Cien Paisajes Culturales en España*; Ministerio de Educación, Cultura y Deporte: Madrid, Spain, 2015.
28. Sarmiento-Mateos, P.; Arnaiz-Schmitz, C.; Herrero-Jáuregui, C.; Pineda, F.D.; Schmitz, M.F. Designing Protected Areas for Social–Ecological Sustainability: Effectiveness of Management Guidelines for Preserving Cultural Landscapes. *Sustainability* **2019**, *11*, 2871. [CrossRef]
29. Gülçin, D. Predicting visual aesthetic preferences of landscapes near historical sites by fluency theory using social media data and GIS. *Int. J. Geogr. Geogr. Educ. (IGGE)* **2021**, *43*, 265–277. [CrossRef]
30. Van Zanten, B.T.; Van Berkel, D.B.; Meentemeyer, R.K.; Smith, J.W.; Tieskens, K.F.; Verburg, P.H. Continental-scale quantification of landscape values using social media data. *Proc. Natl. Acad. Sci. USA* **2016**, *13*, 12974–12979. [CrossRef] [PubMed]
31. Orsi, F.; Geneletti, D. Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *J. Nat. Conserv.* **2013**, *21*, 359–368. [CrossRef]
32. Lieskovský, J.; Rusňák, T.; Klimantová, A.; Izsóff, M.; Gašparovičová, P. Appreciation of landscape aesthetic values in Slovakia assessed by social media photographs. *Open Geosci.* **2017**, *9*, 593–599. [CrossRef]
33. Donaire, J.A.; Camprubí, R.; Galí, N. Tourist clusters from Flickr travel photography. *Tour. Manag. Perspect.* **2014**, *11*, 26–33. [CrossRef]
34. Bermudez, J. Sistemas de información geográfica y Administración pública: El sistema de información de Patrimonio Histórico inmueble de la Comunidad de Madrid. In *Manual de Tecnologías de la Información Geográfica Aplicadas a la Arqueología*; en Minguez, M.C., Capdevila, E., Eds.; Comunidad de Madrid, Museo Arqueológico Regional: Madrid, Spain, 2016; pp. 399–424.
35. Casalegno, S.; Inger, R.; DeSilvey, C.; Gaston, K.J. Spatial Covariance between Aesthetic Value & Other Ecosystem Services. *PLoS ONE* **2013**, *8*, e68437.