AESTIMUM 74, Giugno 2019: 43-62

Rural environment and landscape quality: an evaluation model integrating social media analysis and geostatistics techniques

Veronica Alampi Sottini, Elena Barbierato, Iacopo Bernetti*, Irene Capecchi, Sara Fabbrizzi, Silvio Menghini

University of Florence GESAAF Piazzale delle Cascine 18, I 50144, Firenze (Italy)

*E-mail: *iacopo.bernetti@unifi.it*

Keywords: *Cultural Ecosystem Services, Geographical Weighted Regression, Big Data*

JEL: *C14, Q15*

The use of geo-tagged photographs seems to be a promising alternative for assessing the scenic beauty of the agricultural landscape compared to the traditional investigation based on expert and perceptual approaches. The aim of this study is integrating the cumulative viewshed calculated from geotagged photo metadata publicly shared on Flickr with raster data on geomorphology, historical sites, and the natural environment, using landscape ecology metrics and Geographically Weighted Regression modelling. Crowdsourced data provided empirical assessments of the covariates associated with visitor distribution, highlighting how changes in infrastructure, crops and environmental factors can affect visitor's use. This information can help researchers, managers, and public planners to develop projects, plans and guidelines to increase the visual quality of the agricultural landscape.

1. Introduction

Humans benefit from the many services that rural ecosystems deliver whether it is food supply, clean water regulation or inspiration invoked by a beautiful landscape. The Millennium Ecosystem Assessment (MA, 2003) in the early 2000s popularized this concept as "ecosystem services". The main reference for ecosystem services assessment in public policies for rural landscapes remains the ecosystem services cascade model defined by de Groot (2006). It classifies ecosystem services into four classes, identifying for each class the ecosystem functions relevant for human needs: regulating or regulation services, supporting or habitat services, provisioning or production and cultural ecosystem services (CES). The Millennium Ecosystem Assessment (MA, 2003) defined "cultural ecosystem services" as the nonmaterial benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation, and aesthetic experiences.

In Europe, many agricultural landscapes are hot spots in the provision of CES (Pinto-Correia *et al.*, 2006; Stenseke, 2009). These agricultural landscapes are often referred to as cultural landscapes, which are generally defined as landscapes managed by traditional agricultural techniques, locally and historically adapted, by familiar and/or subsistence methods (IEEP, 2007). They often contribute to a unique aesthetic character and support a co-produced human-ecological system.

Over the past twenty years, much attention has been paid to maintaining spatial and economic synergies between ecosystem functions in rural areas in the

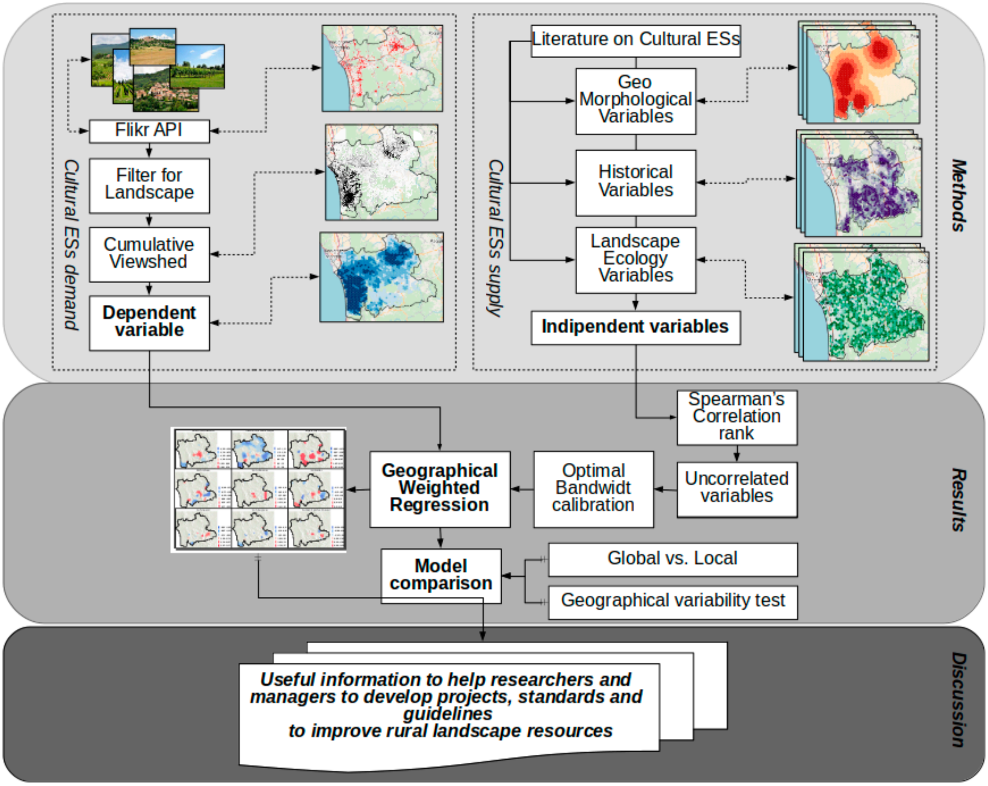
DOI: 10.13128/aestim-7379

context of development planning. The promotion of tourism based on territorial characters and traditions is increasingly a winning strategy (Van Berkel and Verburg, 2011) as it allows the generation of income outside the intensification of agricultural production and promotes the conservation of rural landscape features (Buijs *et al.*, 2006). Tourism attractions are related to people's perception of aesthetic beauty, cultural heritage, spirituality and inspiration (Brown, 2006). These characteristics are non-material benefits related to land management and therefore not exclusive. Failure to provide sufficient incentives to maintain cultural landscapes can result in loss and/or degradation (Swinton *et al.*, 2007). The quantification of cultural services provided by landscapes can therefore help to understand the options for future development that maintain and develop tourism resources. Values that emerge from cultural services are often estimated using stated preferences (e.g., van Berkel and Verburg, 2013; Plieningeretal, 2013). Moreover, a difficult in spatialisation of monetary values with proper detail (resolution) is highlighted in literature (Carvalho-Ribeiro *et al.*, 2016). To cope with this troubles a series of alternative methods in respect to economic analysis have been applied to quantify CES (see Fontana *et al.*, 2013; Nahuelhual *et al.*, 2013; Brown & Fagerholm, 2015; Saarikoski *et al.*, 2016; Rovai *et al.*, 2016; Pastorella *et al.*, 2017; Dunford *et al.*, 2018). The above researches have the merit of having laid the foundations for CES analysis allowing for subjectivity evaluation in participative processes.

Many studies use crowd-sourced images in the analysis of CES, and we can group them into two categories. The first group focuses on the spatial and temporal information of photos (Casalegno *et al.*, 2013; Keeler *et al.*, 2015; Gliozzo *et al.*, 2016; Tieskens *et al.*, 2017). The emphasis of these studies was on the location and the users who took and uploaded the photos. The Integrated Valuation recreation model of Ecosystem Services and Tradeoffs (InVEST) applies the concept of photo-user-days (Redhead *et al.*, 2016), which considers the total number of days the users took photos (at least one photo from a user) in each mapping (Wood *et al.*, 2013). The InVEST recreation model started to be applied to several CES analyses (Keeler *et al.*, 2015; Sonter *et al.*, 2016). The second group of studies aims to correlate the landscape context and the biophysical settings with the positions of georeferenced photos (Pastur *et al.*, 2016; Tenerelli *et al.*, 2016; van Zanten *et al.*, 2016; Oteros-Rozas *et al.*, 2017), using geostatistical analysis methods derived from biology, such as the Maximum entropy models (MaxEnt). The researchers applied MaxEnt model to manage visitor impacts on natural resources, including human-nature interactions (Braunisch *et al.*, 2011), and off-piste recreational behaviour prediction (Coppes and Braunisch, 2013; Westcott and Andrew, 2015; Richards and Friess, 2015). The authors implemented MaxEnt model to the estimate CES correlating the locations of Flickr geo-referenced photos with the environmental characteristics of the territory (Yoshimura and Hiura, 2017; Walden-Schreiner, *et al.*, 2018). However, the models highlighted have two critical limits in the assessment of the visual quality of complex cultural rural landscapes.

On the one hand, the approaches based on the probabilistic models (MaxEnt and Negative Bernoulli distribution) consider only the territorial characteristics that occur in a single location or close to its spatial proximity. On the other hand,

Figure 1. Flow-chart of the work.



the entire surrounding landscape influences photographic recovery (Van Berkel *et al.*, 2018). In this regard, the calculation of the views is potentially useful to capture the perception of the landscape.

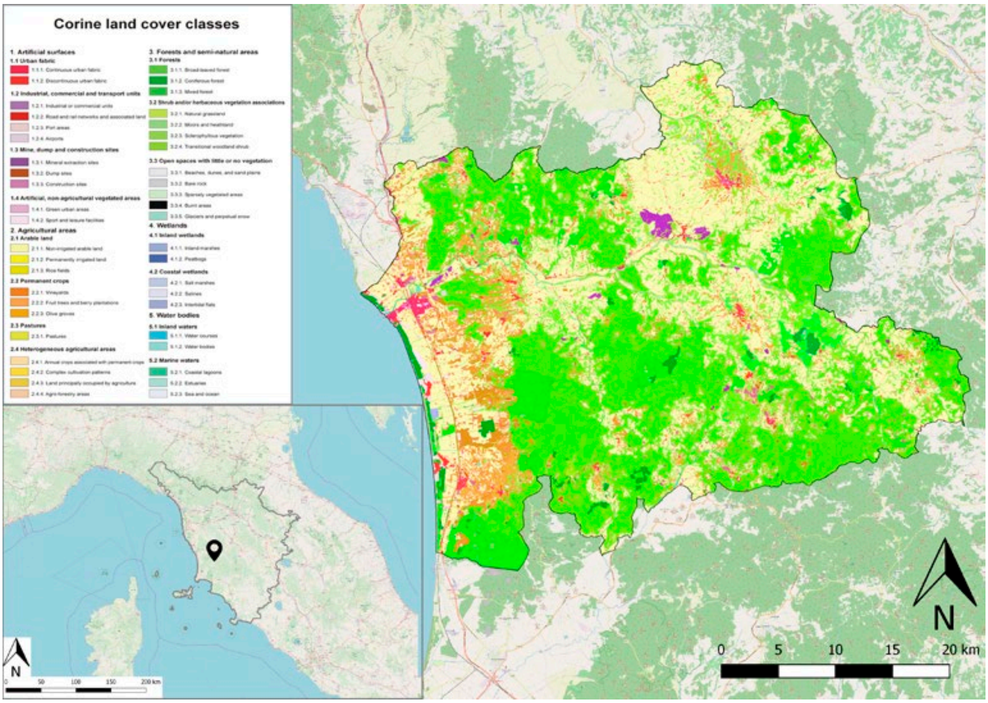
Moreover, the hypothesis at the basis of the two approaches is that the statistical relationship between explanatory variables of landscape quality and concentration of shared photos is constant in space. In complex landscapes, it seems reasonable to assume that there may be intrinsic differences regarding space that occur in terms of spatially variable parameters. In both cases, it seems preferable to use geostatistical techniques to describe and map these spatial variations as an exploratory tool to develop a better understanding of the relationships studied.

The aim of this paper is integrating the geotagged photo metadata publicly shared on Flickr with raster data on geomorphology, historic sites and the natural environment, using landscape ecology indexes and Geographically Weighted Regression (GWR) modelling. Figure 1 shows the workflow of the approach.

2. Study area

The study area is located on the river basin of the Cecina River, located along the coast of Livorno and Pisa. Forest and crops make up the landscape. Today, the coastal strip is characterised by prevalent agriculture of plains (with arable crops and horticultural crops) and hills (with olive groves, promiscuous crops and specialised vineyards), and by widespread and concentrated urbanisation, particularly relevant in some places dedicated to summer tourism. Although it is a context of high anthropization, the coastal territory shows significant naturalistic areas of value linked to the presence of humid areas and back-dunal woods, on the one side, and continuous sandy coastal system of dune habitats and natural pine groves of domestic pine, on the other. Agro-forest-pastoral landscapes of high naturalistic value, crossed by the course of the Cecina River and by a dense hydrographic network, dominate the internal hilly territory. Vast sclerophyllous and broad-leaved thermophile woods alternate with traditional agricultural landscapes. On one of the hills lies the historic city of Volterra, surrounded by beautiful scenic hills characterised by extensive agriculture (arable crops). About 50,000 inhabitants live in Val di Cecina. The area covers more than 200,000 hectares, 43% of which is forest and 35% arable land. Figure 2 shows the study area.

Figure 2. Study area.



3. Methods

3.1 Demand for cultural ecosystem services

In our research, the geotagged photos were queried from the Flickr Application Programming Interface using the statistical software program R. The raw database contained about 35,000 localizations of photos taken in the period 2005-2017. The pictures containing in the tags the "agriculture", "rural landscape", "vineyard", "olive", "grassland", and the related words were filtered. Finally, specific filters were applied to avoid distortions due to photos repeated many times in a single location by a single photographer. The final database counted 11,296 photographic points. The analysis of the spatial distribution of the Cultural ESs application was carried out through the following elaborations.

As a proxy for the demand for Cultural ESs, we develop an index using cumulative viewsheds calculated from photographing positions. Visibility analysis is increasingly applied by landscape planners as well, being useful as a decision support system, since it deals with the best possible spatial arrangement of land uses and it assesses the visual impact of given features in the landscape (e.g., Bell, 2001; Bryan, 2003; Hernández *et al.*, 2004; Palmer and Hoffman, 2001). Perhaps the most popular concept used to explore visual space in a landscape has been the cumulative viewshed (Wheatley 1995; Ramos and Pastor, 2012), sometimes called total viewshed or intrinsic viewshed (Franch-Pardo, Cancer-Pomar and Napoletano, 2017). In general, cumulative viewsheds are created by repeatedly calculating the viewshed from various viewpoint locations and then adding them up one at a time using map algebra, in order to produce a single image. We defined and calculated each viewshed using a digital elevation model (DEM) of 10 m from a height of 165 cm and within a maximum radius of 5 km (Willemen *et al.*, 2008; Chesnokova *et al.*, 2017). The single viewsheds were added together to obtain a cumulative viewshed. The result was transferred into a hexagonal grid theme with a cell size of 1 km, with visibility attributes assigned to each cell. We chose the hexagonal grid because of its topological and geometric properties (Feick and Robertson, 2015). The maps of the indicators, such as the cumulative viewshed, were sampled using a hexagonal grid with a 1-kilometre side, resulting in 1,444 statistical observations.

3.2 Potential supply of cultural ecosystem services

It is possible to map the potential supply of CES by analysing the relationship between the demand area and its environmental factors as the demand map shows the visitors' aesthetic preferences.

The analysis of the relationships between the visual quality of the landscape and its structural properties is an active area of research in the field of environmental perception. The following visual quality indicators were selected, and, according to Ode, Tveit and Fry (2008), divided into five conceptual categories:

1. indicators of complexity: number of different land covers per view, Shannon index.
2. indicators of naturalness: percentage area, edge density, and number of patches of natural and semi-natural vegetation; percentage area, edge density, and number of patches of water bodies, Shannon index, number of patches, landscape shape index;
3. indicators of historicity: distance from historic villages; distance from historical roads;
4. indicators of coherence: percentage area, edge density, and number of patches of vineyards; percentage area, edge density, and number of patches of the olive grove; percentage area, edge density, and number of patches of arable land;
5. indicators of visual scale: elevation, the standard deviation of elevation, the range of elevation.

The indicators at points 1, 2 and 4 were calculated at landscape level using the Fragstats software. According to the standards legend Corine Land Cover level 2, we calculated the indicators of naturalness and complexity for each land use class. The indicator at point 3 derives from historical territorial geodatabases of the Tuscany Region. Finally the indicators at point 5 derive from our elaboration using the DEM of Tuscany Region. The initial set results to be composed of 78 explanatory variables.

To estimate the spatial distribution of the potential supply of Cultural SEs, a Geographically Weighted Regression model was used with the cumulative viewshed as the dependent variable and the potential offer indicators as independent variables.

3.3 Geographically Weighted Regression model for cultural ecosystem services

To investigate the presence of spatial variability in the relationships between the dependent variable (cumulative viewshed) and the explanatory variables (potential supply of CES), we implemented a spatial statistical approach using Geographically Weighted Regression (GWR) (Fotheringham *et al.*, 2002). Classical statistical methods, such as multivariate regression, assume that the same relationship occurs everywhere in space and, thus, they generate a global average value valid for the entire data set, even though, in reality, it can not be valid anywhere. Geographical methods can capture spatial variability, which is one of the main attributes able to explain local differences, and can solve the problem linked to one global average value by calibrating in each position a separate model that considers only the data of the neighbourhood closest to the point of analysis. Moreover, the data are weighted according to their geographical distance from each local regression point so that the closer they are to the point of analysis the more important they are. The result is a set of local models, one for each point, that capture any spatial variability in the relationships.

The first "law" of geography states that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). This is the key concept of spatial data analysis and is related to the concept of spatial correlation.

GWR is a local spatial statistical technique used to analyse and map spatial non-stationarity, i.e., the measurement of relationships among variables that may differ at different locations. Unlike conventional regression, which produces a single regression equation to summarize global relationships among the explanatory and dependent variables, GWR provides a calibration of separate regression equations for each observation of dataset, consisting of a dependent (response) variable y and a set of k independent (explanatory) variables x_k, k=1 … m, and of n observations with known geographical coordinates. Each equation is calibrated using a different weighting of the observations contained in the dataset. The equation for a typical GWR model is (Fotheringham *et al.*, 2001, Fotheringham *et al.*, 1998):

$$y_i(u) = \beta_{0i}(u,v) + \beta_{1i}(u,v)x_{1i} + \dots + \beta_{mi}(u,v)x_{mi}$$

As GWR generally (but not necessarily) assumes that Tobler's first law is verified to a given dataset, the calibration of the GWR model requires a decision regarding the size of the subset of n observations to be included in the neighbourhood of the predicted values. This is referred to as the bandwidth size for estimating the local regression parameters (Brunsdon *et al.*, 1998). Thus, the weighting scheme is that the values near to point i have more influence in the estimated regression values than values located far away from that same point (Fotheringham *et al.*, 2001). In this study we adopt the Gaussian kernel type that weights continuously and gradually decreases from the centre of the kernel but never reaches zero. The kernel shape is defined by the following equation, which takes into account only the *nth* nearest neighbours:

$$w_{ij} = \exp\frac{-d_{ij}^2}{b^2}$$

where i is the regression point index; j is the locational index; w_{ij} is the weight value of observation at location j for estimating the coefficient at location i; d_{ij} is the Euclidean distance between i and j; b is a bandwidth size defined by a distance metric measure.

Bandwidths for GWR models can be user-specified or found via some automated (e.g., cross-validation) procedure provided some objective function exists. Different methods are proposed to define the finest bandwidth value or the appropriate value of n (Hurvich *et al.*, 1998; Akaike, 1974; Fotheringham *et al.*, 2003).

Many studies have applied GWR in human and political geography (Mansley and Demšar, 2015; Brunsdon *et al.*, 1996; Fotheringham *et al.*, 2013), as well as in physical geography and ecology (Atkinson *et al.*, 2003; Clement *et al.*, 2009; Harris *et al.*, 2010; Jetz *et al.*, 2005), proving the suitability of this tool to provide an explanatory approach in spatially varying relationships (Páez *et al.*, 2011). For the evaluation of CES, Tenerelli *et al.* (2016) used a GWR method to study the relationship between the geo-tagged images account and the landscape settings, whose spatial variation may affect the cultural service. Schirpke *et al.* (2018) used a GWR model to analyse how spatial and temporal patterns correlate spatially explicit indicators and crowd-sourced information from social media. The estimation of the GWR models was carried out through the GWmodel library of the statistical program R (Gollini *et al.*, 2013; Lu *et al.*, 2013). Fotheringham and Park (2018) investigates both spatial and temporal elements of the apartment pricing process by modelling the determinants of apartment prices. Riccioli *et al.* (2018) analysed and tested the spatial non-stationarity of the relationship between ungulates and human activities.

The GWR approach uses a moving window weighting technique, where localised models are at target locations. Here, for a single model in a specific target location, we weight all neighbouring observations according to a certain distance-decay kernel function and then locally apply the model to the weighted data. The bandwidth controls the size of the window over which this localised model might apply. A fundamental element in GW modelling is the spatial weighting function (Fotheringham *et al.*, 2002) that quantifies (or sets) the spatial relationship or spatial dependency between the observed variables. There are three critical elements in structuring this weighting system: (i) the type of distance, (ii) the kernel function and (iii) its bandwidth. According to Gollini *et al.* (2013), we adopted the Euclidean distance with a bi-square kernel. Having the data set organised on a regular hexagonal tessellation, we set an adaptive kernel bandwidth that to include the N hexagons closest to the observation/calibration hex. When an objective function exists (e.g., when the model can predict it), we can find an optimal bandwidth, using cross-validation and related approaches. We can find an optimum kernel bandwidth for GW regression by minimising some diagnostic models of adaptation, such as a leave-one-out cross-validation (CV) score (Bowman, 1984), which represents the accuracy of the model prediction; or the Akaike Information Criterion (AIC) (Akaike, 1973), which represents the parsimony of the model (i.e., a compromise between prediction accuracy and complexity). Once we calibrated our local model, we evaluated the spatial variability in the relationships through a visual representation of the parameter estimate surfaces. The surfaces were cross-mapped with the local t-values for each parameter estimate to identify areas where the relationships are significant. We also mapped the local percentage of explained deviance to identify areas where the model is performing better (percentage of explained deviance higher than the average) or worse, and we relate these patterns with the most significant local parameter estimates. Finally, we tested the spatial distribution of the local and global residuals both through visual representation and using Moran's I measure of spatial autocorrelation. The level of spatial autocorrelation can be investigated visually by mapping the standardised residuals for both models as well as calculating measures of spatial autocorrelation, such as Moran's I (Goodchild, 1986; Moran, 1950).

4. Results

The first step in the GWR procedure was to test the multicollinearity between the variables using Spearman's correlation rank. We kept all the variables as they showed a Spearman's correlation lower than 0.7. In the end, we considered a final set of 9 variables. Figure 3 shows the map of the explanatory variable (cumulative viewsheds) and Figure 4 the 6 maps of the independent variables.

Table 1 shows the results for the global Generalized Last Squares (GLS) model. The results suggest that all parameter estimates are significant except the patch richness value. The explained deviation is only about 41%, with an AICc coefficient of 17,389. The model significance is assessed by the F-Statistic. The F-Statistic is trustworthy only when the Koenker's studentized Breusch-Pagan (KBP) statistic is not statistically significant (Breusch and Pagan, 1979; Koenker, 1981). In this case, the KBP statistic is significant (*cfr.* Tab. 1). Furthermore, the KBP statistic determines whether the explanatory variables in the model have a consistent relationship to the dependent variable, both in geographic space and in data space. When the model is consistent in geographic space, the spatial processes represented by the explanatory variables behave the same everywhere in the study area (the processes are stationary). When the model is consistent in data space, the variation in the relationship between predicted values and each explanatory vari-

Figure 3. Maps of cumulative viewsheds (explanatory variable).

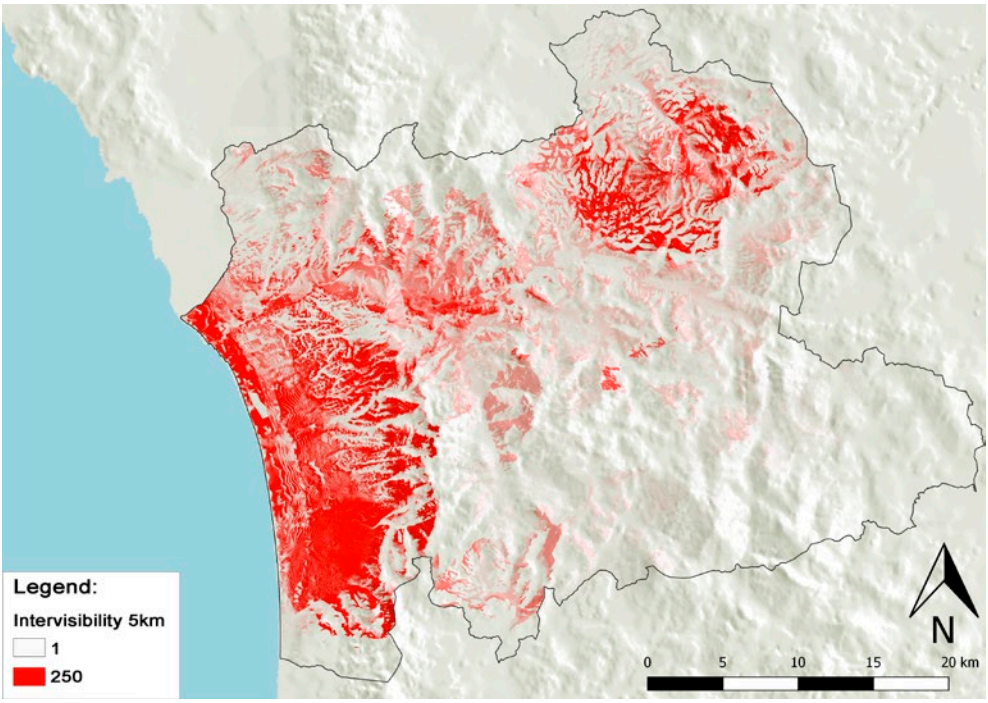
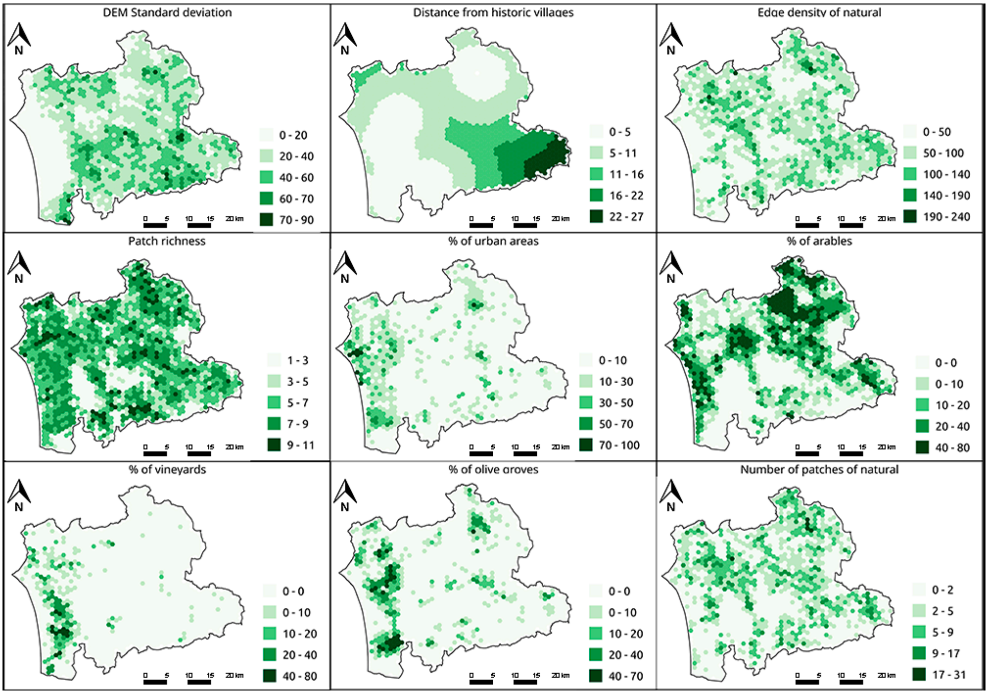


Figure 4. Maps of independents variables.



able does not change with changes in explanatory variable magnitudes (there is no heteroscedasticity in the model). We performs the Breusch-Pagan test for heteroskedasticity on the least squares fit of the spatial models using the procedure bptest.sarlm of the statistical program R (Bivand *et al.*, 2018). The significance of the KBP statistic indicates heteroscedasticity and/or non-stationarity of the model; this model is, therefore, a good candidate for Geographically Weighted Regression analysis.

In the next step, we first built an entirely local GWR model. The result of the bandwidth optimization suggested an optimal bandwidth of 86 cells (i.e. for each of the 1,444 cells, a local model was calibrated using data from the nearest 86 cells). The adaptation of the model was much improved compared to the local model (Table 3) with an average 78.6% of deviance explained (i.e. a significant increase from the global model) and with an AICc of 15.773. The improvement in the quality of the model from global to local shows that there is indeed a spatial variability in the data and that it is essential to unravel it.

According to Lu *et al.* (2015), we performed a model specification exercise to find an independent variables subset for our GW regression. To support this procedure, we implemented a pseudo stepwise procedure, going in a forward direction. The following four steps, where the results are displayed using plots with the AICc values of each model, describe this procedure:

Table 1. Generalized Last Square model.

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	
Intercept	164.6	12.25	13.432	< 2e-16	***
DEM standard deviation	-1.07	0.2207	-4.847	.000001390000	***
Distance from hystoric village	-0.005395	0.0004906	-10.998	< 2e-16	***
Edge density of natural areas	-0.3562	0.08907	-4	.000066700000	***
Patch richness	-2.231	1.57	-1.421	.155000000000	
Percent of urban areas	1.851	0.4493	4.119	.000040300000	***
Percent of arables	0.6574	0.1298	5.064	.000000463000	***
Percent of vineyards	5.74	0.424	13.536	< 2e-16	***
Percent of olive grow	2.023	0.3581	5.648	.000000019500	***
Number of natural patches	-6.67	1.083	-6.159	.000000000948	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 99.3 on 1434 degrees of freedom
Multiple R-squared: 0.4221
Adjusted R-squared: 0.4185
F-statistic: 116.4 on 9 and 1434 DF, p-value: < 2.2e-16
Diagnostic information
Residual sum of squares: 14139488
Sigma(hat): 99.02258
AIC: 17389.26
AICc: 17389.44
Koenker (BP) Statistic 39.543, df = 9, p-value = 9.194e-06

1. Calibration of all possible bivariate geographically weighted regressions by sequential regression of a single independent variable to the dependent variable.
2. Detection of the best performing model that produces the minimum AICc, and permanent incorporation of the corresponding independent variable in subsequent models.
3. Sequential introduction of a variable of the remaining group of independent variables for the creation of new models with the independent variables permanently included, and determination of the following permanently included variable from the best fitting model that has the minimum AICc.
4. Reiteration of step 3 until the model includes permanently all independent variables.

These steps were performed using the package GWmodel of the statistical software R (Lu e al, 2014). Figure 5 shows a circle view of the 45 geographically weighted regressions (numbered 1 to 45) that result from the stepwise procedure.

In the figure, the dependent variable is located in the center of the chart and the independent variables are represented as nodes differentiated by shapes and color. The first independent variable permanently included is "distance from historic villages", the second one is "edge density of naturals", the third one is "per-

Table 2. Results of Geographically Weigthed Regression model.

	Min.	1st Qu.	Median	3rd Qu.	Max.
Intercept	-170.47	11.225	68.895	204.27	981.7019
DEM standard deviation	-92.992	-0.38785	0.011412	0.29064	8.81
Distance from hystoric village	-0.15695	-0.016151	-0.0052386	-0.00028583	0.0765
Edge density of natural areas	-4.4125	-0.17754	-0.0033543	0.052348	2.7686
Patch richness	-28.059	-1.8661	-0.086415	1.3978	54.715
Percent of urban areas	-27.009	-0.085966	0.6723	2.3271	27.3817
Percent of arables	-5.9141	-0.077748	0.02702	0.58779	14.0546
Percent of vineyards	-33.376	-0.5068	0.074384	1.6036	22.7161
Percent of olive grow	-13.962	-0.43454	-0.015489	0.47527	17.3402
Number of natural patches	-25.306	-2.1949	-0.18176	0.30492	21.8281

AICc : 16274.47
AIC: 15773.28
R-square value: 0.8462371
Adjusted R-square value: 0.786029

centage of arable land" and the last one is "numbers of patches". Moreover, figure 5 shows the corresponding AICc values for the same fits. The two graphs together explain the model performance when we introduce an increasing number of variables. As can be expected, AICc values continue to fall until all independent variables are included. The results suggest that it is worth continuing with all eight independent variables.

To interpret the spatial relationships resulting from GWR, we represented the local parameter estimate surfaces, and we analysed the spatial distribution of local coefficients and their relative significance levels (Figure 6 and 7).

In general, the parameters are not significant in the south-east area of the territory under study, characterised by low photo density (see also Figure 3). We notice that there are two distinct areas. In the north-west area (the area around the city of Volterra), the standard deviation of the elevations, the distance from historic villages, the percentage of olive groves, the density of margins from natural areas and the percentage of arable land are significant. In the East area, close to the coast, the DEM standard deviation, the distance from the historic villages, the margins density of the natural areas, the percentage of area affected by arable land, vineyards and olive groves and the number of natural patches are significant on a vast area. About the signs of the coefficient, the distance from the historic villages and the standard deviation of the DEM are both negative in the two areas characterized by the highest concentration of photos. For the dependent variables of landscape ecology instead, the signs of the coefficient are different in the two areas. The perception of the landscape of Volterra is positively correlated to the percentage of olive groves, and the edge density of natural areas, while it is

Figure 5. Model view of the stepwise specification procedure.

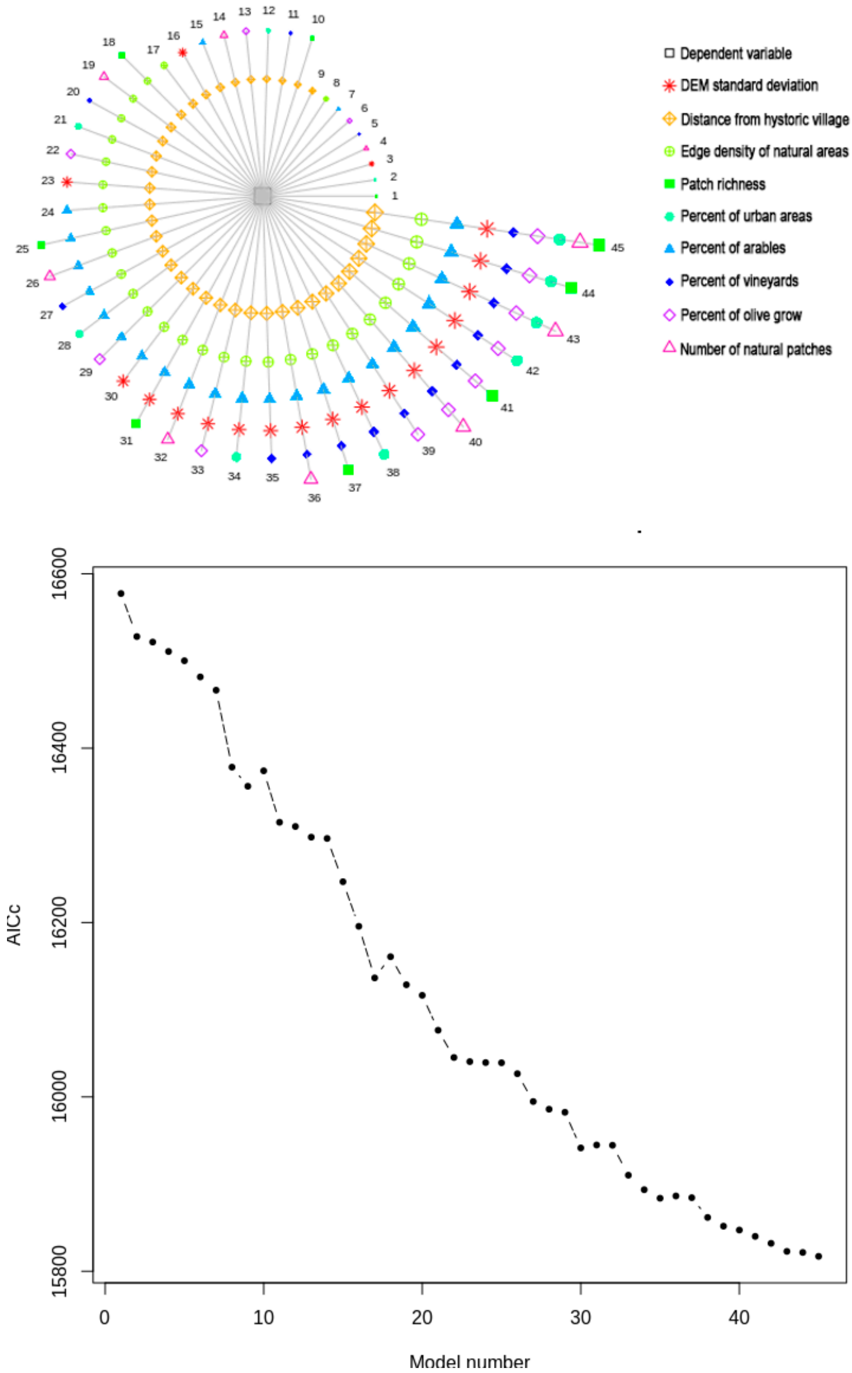


Figure 6. Maps of spatial distribution of local coefficients.

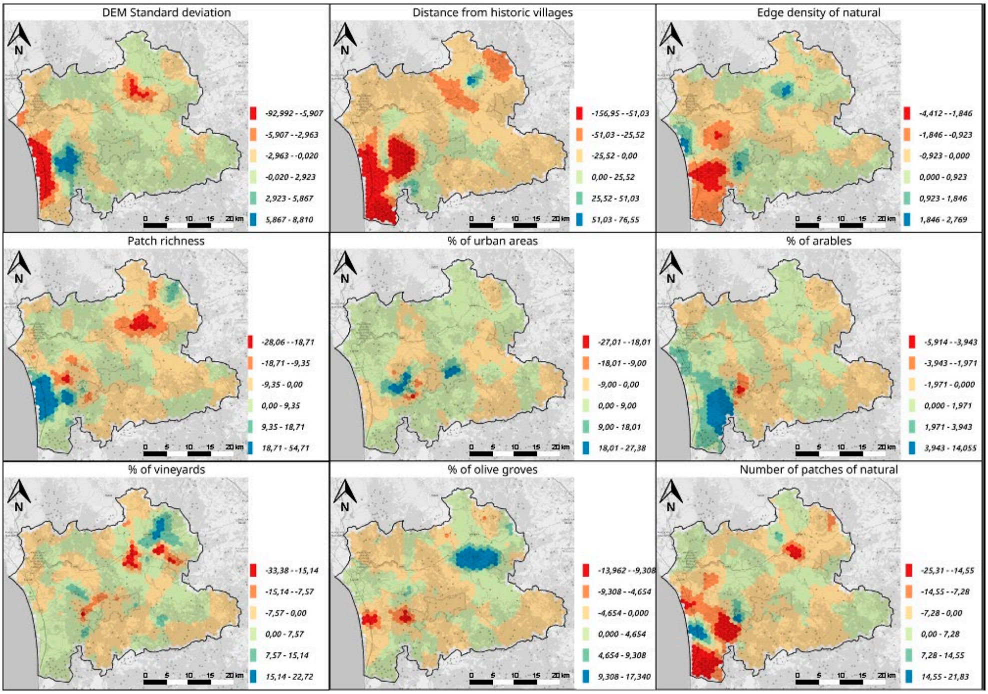


Figure 7. Maps of spatial distribution of significance levels.

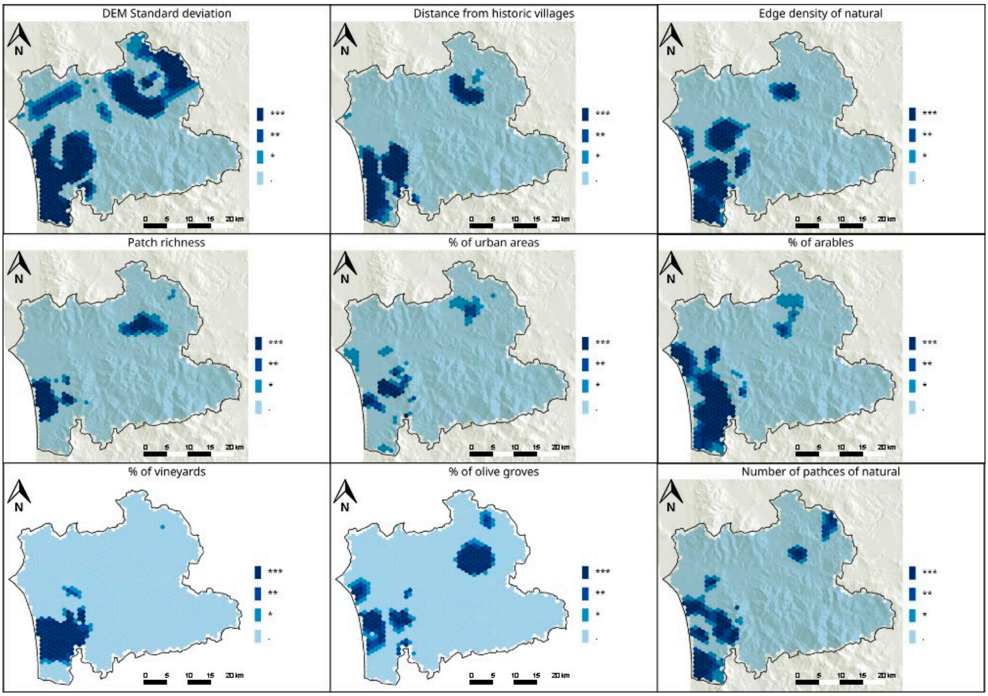
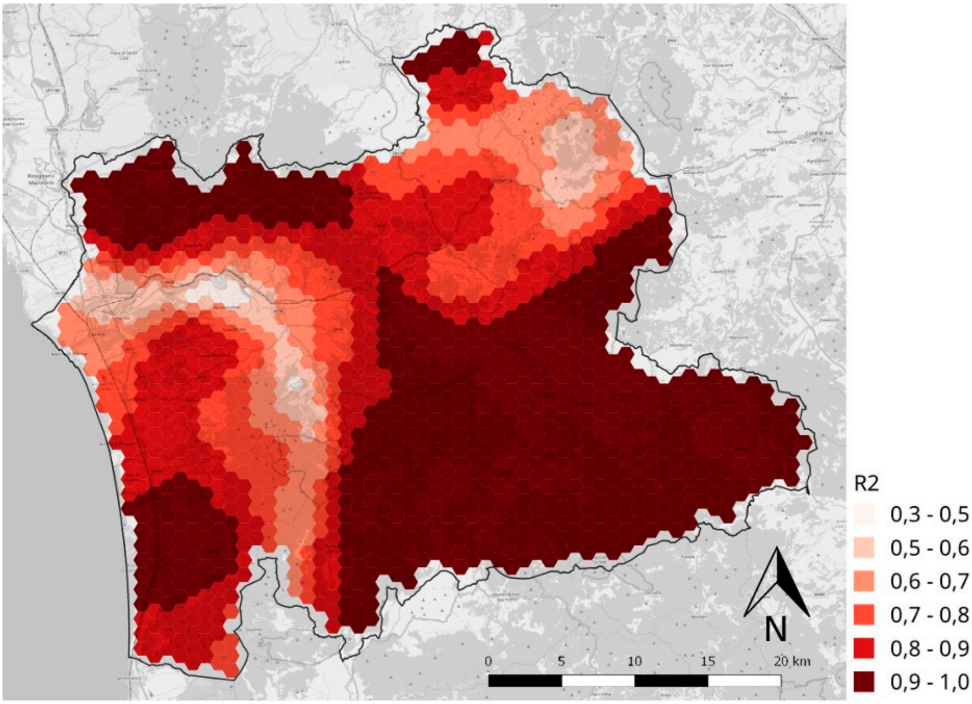


Figure 8. Map of explained deviance.



negatively correlated to the patch richness and the percentage of vineyards. In the area near the coast, the perception of the landscape is positively correlated to the patch richness, to the percentage of arable land and it is inversely proportional to the density of margins and the number of patches of natural areas. In general terms, therefore, the GWR highlights the presence of highly differentiated areas relating to the appreciation of the characteristics of the landscape.

To analyse the local variability of the relationships between the photo counting and the explanatory variables, we mapped the local percentage of explained deviance. Figure 8 shows the explained deviance, highlighting that it is everywhere higher than in the global model.

5. Discussion and conclusion

The implemented models confirmed the importance of agricultural cultivations for the value of the landscape and allowed to obtain a spatial evaluation of the consistency of the externalities produced by agriculture, with obvious benefits for the choices of territorial government and rural development.

Furthermore, Flickr provides a free, up-to-date, and high spatial and temporal resolution information source. However, as our analyses revealed, each crowdsourced database has limitations in terms of spatial data quality and sampling bias. The results of the spatial analysis of the photographic series indicate specific models of visit preferences and how the perception of the agricultural landscape is influenced both by the complementary characteristics of the rural landscape and by the agronomic choices at different scales of analysis. The spatial distribution of visit preferences provides an indicator of the social benefits of agriculture, allowing a local analysis of the areas providing services and addressing the lack of quantitative indicators.

Our explanatory analysis allows the identification of areas of interest in which land use planning and management strategies of the agricultural ecosystem should take into account the actual provision of non-material benefits related to the landscape. The analysis performed supports setting landscape planning priorities by providing an understanding of how changes in specific environmental settings can influence the supply of landscape in certain areas. Therefore, the proposed method represents a significant first step in informing stakeholders and policymakers about priority areas. A further improvement of this study is to conduct interviews and surveys with questionnaires to visitors. It would allow us to evaluate the benefits and the different values relating to the landscape. Validating these data sources and addressing uncertainty in data deriving from social media represents an important area of future research as as it is necessary before crowdsourced data achieves acceptance for use in protected area planning and management, and for quantifying and qualifying the characteristics and values of cultural ecosystem services in rural areas.

References

Akaike H (1973). “Information Theory and an Extension of the Maximum Likelihood Principle.” In BN Petrov, F Csaki (eds.), 2nd Symposium on Information Theory, pp. 267–281. Akademiai Kiado, Budapest.

Akaike H. (1974). A new look at the statistical model identification. In Selected Papers of Hirotugu Akaike (pp. 215-222). Springer, New York, NY.

Atkinson P.M., German, S.E., Sear, D.A., Clark, M.J., 2003. Exploring the relations between riverbank erosion and geomorphological controls using geographically weighted logistic regression. *Geogr. Anal.* 35(1): 58-82.

Bell S. (2001). Landscape pattern, perception and visualisation in the visual management of forests. *Landscape and Urban planning* 54(1-4): 201-211.

Bivand R.S. and Wong D.W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST* 27(3): 716-748. doi.org/10.1007/s11749-018-0599-x

Bowman A (1984). An Alternative Method of Cross-Validation for the Smoothing of Density Estimates. *Biometrika* 71: 353-360.

Braunisch V., Patthey P. and Arlettaz R. (2011). Spatially explicit modeling of conflict zones between wildlife and snow sports: prioritizing areas for winter refuges. *Ecological Applications* 21(3): 955-967.

Brown G, Fagerholm N. (2015). Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. *Ecosystem Services* 13: 119-133.

Brown G. 2006. Mapping landscape values and development preferences: a method for tourism and residential development planning. *Int. J. Tourism Res.* 8: 101-113.

Brunsdon C., Fotheringham A.S. and Charlton M.E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis* 28(4): 281-298.
Brunsdon C., Fotheringham S. and Charlton M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)* 47(3): 431-443.
Bryan B.A. (2003). Physical environmental modeling, visualization and query for supporting landscape planning decisions. *Landscape and urban planning* 65(4): 237-259.
Breusch T.S. and Pagan A.R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society* 1287-1294.
Buijs A.E., Pedroli B. and Luginbühl Y. (2006). From hiking through farmland to farming in a leisure landscape: changing social perceptions of the European landscape. *Landscape ecology* 21(3): 375-389.
Campello R.J., Moulavi D., Zimek A., Sander J. (2015). Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 10(1): 5.
Carvalho-Ribeiro S., Ramos I.L., Madeira L., Barroso F., Menezes H. and Correia T.P. (2013). Is land cover an important asset for addressing the subjective landscape dimensions? *Land Use Policy* 35: 50-60.
Casalegno S., Inger R., DeSilvey C. and Gaston K.J. (2013). Spatial covariance between aesthetic value and other ecosystem services. *PloS one* 8(6): e68437.
Chesnokova O., Nowak M. and Purves R.S. (2017). *A crowdsourced model of landscape preference*. In LIPIcs-Leibniz International Proceedings in Informatics (Vol. 86). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
Clement F., Orange D., Williams M., Mulley C., Epprecht M. (2009). Drivers of afforestation in Northern Vietnam: assessing local variations using geographically weighted regression. *Appl. Geogr*. 29(4): 561-576,
Coppes J. and Braunisch V. (2013). Managing visitors in nature areas: where do they leave the trails? A spatial model. *Wildlife biology* 19(1): 1-11.
De Groot R. (2006). Function-analysis and valuation as a tool to assess land use conflicts in planning for sustainable, multi-functional landscapes. *Landscape and urban planning* 75(3-4): 175-186.
Dunford R., Harrison P., Smith A., Dick J., Barton D.N., Martin-Lopez B., Kelemen E., Jacobs S., Saarikoski H., Turkelboom F., Verheyden W., Hauck J., Antunes P., Aszalós R., Badea O., Baró F., Berry P., Carvalho L., Conte G., Czúcz B., Garcia Blanco G., Howard D., Giuca R., Gomez-Baggethun E., Grizetti B., Izakovicova Z., Kopperoinen L., Langemeyer J., Luque S., Lapola D.M., Martinez-Pastur G., Mukhopadhyay R., Roy S.B., Niemelä J., Norton L., Ochieng J., Odee D., Palomo I., Pinho P., Priess J., Rusch G., Saarela S.R., Santos R., van der Wal J.T., Vadineanu A., Vári A., Woods H., Yli-Pelkonen V. (2013). Integrating methods forecosystem service assessment: Experiences from real world situations. *Ecosystem Services* 29: 499-514
Feick R. and Robertson C. (2015). A multi-scale approach to exploring urban places in geotagged photographs. *Computers, Environment and Urban Systems* 53: 96-109.
Fontana V., Radtke A., Bossi Fedrigotti V., Tappeiner U., Tasser E., Zerbe S., Buchholz T. (2013). Comparing land-use alternatives: Using the ecosystem services concept to define a multi-criteria decision analysis. *Ecological Economics* 93: 128-136.
Fotheringham A.S. and Park B. (2018). Localized spatiotemporal effects in the determinants of property prices: A case study of Seoul. *Applied Spatial Analysis and Policy* 11(3): 581-598.
Fotheringham A.S., Brunsdon C. and Charlton M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.
Fotheringham A.S., Charlton M.E. and Brunsdon C. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and planning A* 30(11): 1905-1927.
Fotheringham A.S., Charlton M.E. and Brunsdon C. (2001). Spatial variations in school performance: a local analysis using geographically weighted regression. *Geographical and Environmental Modelling* 5(1): 43-66.

Fotheringham A.S., Brunsdon C. and Charlton M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley, Chichester, England/Hoboken, NJ, USA.

Fotheringham A.S., Kelly M.H. and Charlton M. (2013). The demographic impacts of the Irish famine: towards a greater geographical understanding: the demographic impacts of the Irish famine. *Trans. Inst. Br. Geogr*. 38(2): 221-237.

Franch-Pardo I., Cancer-Pomar L. and Napoletano B. M. (2017). Visibility analysis and landscape evaluation in Martin river cultural park (Aragon, Spain) integrating biophysical and visual units. *Journal of Maps* 13(2): 415-424. DOI: 10.1080/17445647.2017.1319881

Gliozzo G., Pettorelli N. and Haklay M. (2016). Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK. *Ecology and Society* 21(3).

Gollini I., Lu B., Charlton M., Brunsdon C. and Harris P. (2013). GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models. *arXiv preprint* arXiv:1306.0413.

Goodchild M.F. (1986). *Spatial Autocorrelation*. Geo Books, Norwich.

Harris P., Fotheringham A.S., Juggins S. (2010). Robust geographically weighted regression: a technique for quantifying spatial relationships between freshwater acidification critical loads and catchment attributes. *Ann. Assoc. Am. Geogr*. 100(2): 286-306.

Hernández J., Garcıa L. and Ayuga F. (2004). Assessment of the visual impact made on the landscape by new buildings: a methodology for site selection. *Landscape and Urban Planning* 68(1): 15-28.

Hurvich C.M., Simonoff J.S. and Tsai C.L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 60(2): 271-293.

IEEP (Institute for European Environmental Policy) (2007). *Final Report on the Study of HNV Indicators for Evaluation*. European Commission, DG Agriculture, Brussels.

Jetz W., Rahbek C. and Lichstein J.W. (2005). Local and global approaches to spatial data analysis in ecology. *Global Ecol. Biogeogr*. 14(1): 97-98.

Keeler B.L., Wood S.A., Polasky S., Kling C., Filstrup C.T. and Downing J.A. (2015). Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes. *Frontiers in Ecology and the Environment* 13(2): 76-81.

Koenker R. (1981). A note on studentizing a test for heteroscedasticity. *Journal of Econometrics* 17(1): 107-112.

Levin N., Lechner A.M. and Brown G. (2017). An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied geography* 79: 115-126.

Li D., Zhou X. and Wang M. (2018). Analyzing and visualizing the spatial interactions between tourists and locals: A Flickr study in ten US cities. *Cities*.

Lu B., Harris P., Charlton M., Brunsdon C., Nakaya T. and Gollini I. (2015). Package 'Gwmodel'.

Lu B., Harris P., Gollini I., Charlton M. and Brunsdon C. (2013). GWmodel: an R package for exploring spatial heterogeneity. GISRUK 2013, 3-5.

MA, 2003. *Millennium Ecosystem Assessment, Ecosystems and Human Well-being: A Framework for Assessment*. Island Press, Washington.

Mansley E. and Demšar U. (2015). Space matters: Geographic variability of electoral turnout determinants in the 2012 London mayoral election. *Electoral Studies* 40: 322-334.

Moran P.A.P. (1950). Notes on continuous stochastic phenomena. *Biometrika* 37: 17-23.

Nahuelhual L., Carmona A., Lozada P., Jaramillo A. and Aguayo M. (2013). Mapping recreation and ecotourism as a cultural ecosystem service: An application at the local level in Southern Chile. *Applied Geography* 40: 71-82.

Ode Å., Tveit M.S. and Fry G. (2008). Capturing landscape visual character using indicators: touching base with landscape aesthetic theory. *Landscape research* 33(1): 89-117.

Oteros-Rozas E., Martín-López B., Fagerholm N., Bieling C. and Plieninger T. (2018). Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators* 94: 74-86.

Páez A., Farber S., Wheeler D. (2011). A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. *Environ. Plan. A* 43(12): 2992-3010.

Palmer J.F. and Hoffman R.E. (2001). Rating reliability and representation validity in scenic landscape assessments. *Landscape and urban planning* 54(1): 149-161.

Pastorella F., Giacovelli G., De Meo I., Paletto A. (2017). People's preferences for Alpine forest landscapes: Results of an internet-based survey. *Journal of Forest Research* 22(1): 36-43.

Pastur G.M., Peri P.L., Lencinas M.V., García-Llorente M. and Martín-López B. (2016). Spatial patterns of cultural ecosystem services provision in Southern Patagonia. *Landscape Ecology* 31(2): 383-399.

Pinto-Correia T., Gustavsson R. and Pirnat J. (2006). Bridging the gap between centrally defined policies and local decisions–Towards more sensitive and creative rural landscape management. *Landscape ecology* 21(3): 333-346.

Ramos B.M. and Pastor I.O. (2012). Mapping the visual landscape quality in Europe using physical attributes. *Journal of Maps*, 8(1), 56-61.

Redhead J.W., Stratford C., Sharps K., Jones L., Ziv G., Clarke D., … and Bullock J. M. (2016). Empirical validation of the InVEST water yield ecosystem service model at a national scale. *Science of the Total Environment* 569: 1418-1426.

Riccioli F., Boncinelli F., Fratini R. and Casini L. (2018). Geographical Relationship between Ungulates, Human Pressure and Territory. *Applied Spatial Analysis and Policy* 1-24.

Richards D.R. and Friess D.A. (2015). A rapid indicator of cultural ecosystem service usage at a fine spatial scale: content analysis of social media photographs. *Ecological Indicators* 53: 187-195.

Rovai M., Andreoli M., Gorelli G., Jussila H. (2016). A DSS model for the governance of sustainable rural landscape: A first application to the cultural landscape of Orcia Valley (Tuscany, Italy). *Land Use Policy* 56: 217-237.

Saarikoski H., Mustajoki J., Barton D.N., Geneletti D., Langemeyer J., Gomez-Baggethun E., Marttunen M., Antunes P., Keune H. (2016). Multi-Criteria Decision Analysis and Cost-Benefit Analysis: Comparing alternative frameworks for integrated valuation of ecosystem services. *Ecosystem Services* 22B: 238-249.

Schirpke U., Meisch C., Marsoner T. and Tappeiner U. (2018). Revealing spatial and temporal patterns of outdoor recreation in the European Alps and their surroundings. *Ecosystem services* 31: 336-350.

Sonter L.J., Watson K.B., Wood S.A. and Ricketts T.H. (2016). Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLoS one* 11(9): e0162372.

Stenseke M. (2009). Local participation in cultural landscape maintenance: Lessons from Sweden. *Land Use Policy* 26(2): 214-223.

Swinton S.M., Lupi F., Robertson G.P., Hamilton S.K., 2007. Ecosystem services and agriculture: cultivating agricultural ecosystems for diverse benefits. *Ecol. Econ*. 64: 245-252.

Tenerelli P., Demšar U. and Luque S. (2016). Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. *Ecological Indicators* 64: 237-248.

Tieskens K.F., Schulp C.J., Levers C., Lieskovský J., Kuemmerle T., Plieninger T. and Verburg P.H. (2017). Characterizing European cultural landscapes: Accounting for structure, management intensity and value of agricultural and forest landscapes. *Land use policy* 62: 29-39.

Tobler W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46: 234-240.

van Berkel D.B. and Verburg P.H. (2011). Sensitising rural policy: Assessing spatial variation in rural development options for Europe. *Land Use Policy* 28(3): 447-459.

van Zanten B.T., Van Berkel D.B., Meentemeyer R.K., Smith J.W., Tieskens K.F. and Verburg P.H. (2016). Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences* 113(46): 12974-12979.

Walden-Schreiner C., Leung Y.F. and Tateosian L. (2018). Digital footprints: Incorporating crowdsourced geographic information for protected area management. *Applied Geography* 90: 44-54.

Walden-Schreiner C., Leung Y.F. and Tateosian L. (2018). Digital footprints: Incorporating crowd-sourced geographic information for protected area management. *Applied Geography* 90: 44-54.

Welling S.H., Refsgaard H.H., Brockhoff P.B. and Clemmensen L.H. (2016). Forest floor visualizations of random forests. *arXiv preprint* arXiv:1605.09196.

Westcott F. and Andrew M.E. (2015). Spatial and environmental patterns of off-road vehicle recreation in a semi-arid woodland. *Applied Geography* 62: 97-106.

Wheatley D. (1995). *Cumulative viewshed analysis: A GIS-based method for investigating intervisibility, and its archaeological application*. In G. Lock and Z. Stancic (Eds.), Archaeology and geographical information systems (pp. 171-186). London: Taylor and Francis.

Willemen L., Verburg P.H., Hein L. and van Mensvoort M.E. (2008). Spatial characterization of landscape functions. *Landscape and Urban Planning* 88(1): 34-43.

Wood S.A., Guerry A.D., Silver J.M. and Lacayo M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific reports* 3: 2976.

Yoshimura N. and Hiura T. (2017). Demand and supply of cultural ecosystem services: Use of Geo-tagged photos to map the aesthetic value of landscapes in Hokkaido. *Ecosystem Services* 24: 68-78.