Article

The Distribution of Rural Accommodation in Extremadura, Spain-between the Randomness and the Suitability Achieved by Means of Regression Models (OLS vs. GWR)

José-Manuel Sánchez-Martín ^{1,*} , José-Luis Gurría-Gascón ² and Juan-Ignacio Rengifo-Gallego ²

¹ Faculty of Business, Finance and Tourism, University of Extremadura, 10071 Cáceres, Spain

² Faculty of Letters, University of Extremadura, 10071 Cáceres, Spain; jlgurria@unex.es (J.-L.G.-G.); irengifo@unex.es (J.-I.R.-G.)

* Correspondence: jmsanche@unex.es

Received: 21 May 2020; Accepted: 8 June 2020; Published: 10 June 2020

Abstract: There are multiple types of regression, the essential task of which is the obtaining of models which, starting from a set of regressive values, are capable of finding explanations for the variability of a dependent. However, in many cases, the territorial criterion is not considered to be a noteworthy factor of analysis, owing to which this deficiency has encouraged the arising of spatial statistics. Nevertheless, given the variety of regressions, it is not clear which can best be adapted to the analysis of tourism. In this sector, when the supply of accommodation is analysed, it is understood that it must be strongly related to the presence of resources, owing to which it has been taken as an example of an application between two differentiated regression techniques: ordinary least squares (OLS) and geographically weighted regression (GWR), with the objective of determining which of the two is best adapted to this type of analysis. The model has been drawn up based on various methods, although it has been shown that it is more efficient to resort to the declared preferences of the rural tourist, with the starting point being a survey made of the tourists. These aspects have been taken as independent variables with the aim of explaining the distribution of accommodation establishments. The results obtained show that the configuration of the spatial relations between the variable included in the model encourages the explanation of the latter, owing to which GWR is much more suitable than OLS, even when a system as complex as the distribution of accommodation establishments is analysed. Likewise, it is noteworthy that the distribution of accommodation does not also follow the guidelines marked by demand; far from it, it appears that in some areas, it is of a random nature.

Keywords: geographically weighted regression; ordinary least squares regression; rural accommodation; rural tourist; Extremadura

1. Introduction

Tourism has developed greatly in many parts of the world. It has become an essential industry in the achieving of the socioeconomic development of a multitude of places [1]. It is for this reason that numerous studies analyse the activity from various perspectives. These include those studying the economy of tourism in various countries [2–4], those which concentrate on the analysis of the tourism satellite account [5–7], or those which analyse specific products and aspects such as sun and beach tourism [8], rural tourism [9], and cultural tourism [10].

There is a considerable diversity of themes analysed from the point of view of tourism, although all need data to implement them [11]. Until recent years, the information was concentrated on economic,

heritage, social, and territorial aspects and naturally, also, all those linked to the characteristics of the tourist or the accommodation. This was abundant information although its amount was limited. Nowadays, thanks to the continuous progress of technology, information has increased exponentially. Indeed, big data synthesises the huge volume of tourist information which exists and can be used to design and administer smart tourist destinations [12] in order to predict tourist demand [13,14] or even analyse tourist behaviour [15]. Indeed, many studies examine the use of big data [16] and data mining [17]. Nevertheless, recent studies are also appearing which put forward the need to resort to small data [18] and stress the need to use only the necessary information.

Despite the disparity of approaches and subjects, it is clear that tourism needs more and more information and also the application of techniques allowing its processing and comprehension. Nevertheless, it continues to be a phenomenon which is difficult to analyse as it is characterised by a variety of both themes and spatiotemporal factors [19].

This information needs to be correctly processed by means of a multitude of statistical techniques [20], which are criticised on occasion [21], owing to their concept. These techniques may be of various types, and numerous studies use qualitative analyses applied to very varied themes. Among them, those which analyse social networks as distribution channels [22], the impact of tourism on residents [23], the image projected by a destination [24] or accommodation [25], and even the very internationalisation of quality standards [26] stand out. Others resort to quantitative analyses to ascertain the effect of seasonality and the rating of the accommodation on the price [27] or the characteristics of the tourists [28], among other subjects.

The literature includes intense epistemological debates on qualitative research compared with quantitative research applied to the analysis of tourism [29]. At the same time, studies have appeared which detail the research methods most frequently used on this subject [30–32]. In spite of this, the use of quantitative techniques clearly predominates. These include regression analyses in their multiple variants. Numerous studies allude to structural equation modelling [33], quantile regression [34,35], multiple linear regression [36–38], and even attribute a key role to the geostatistical analyst [25,39–41]. Among the latter, geographically weighted regression is becoming more and more important, owing to the influence of the surrounding area on tourist parameters [42].

Two clear trends are observed by those who advocate the use of quantitative analyses applied to tourism. On the one hand, there are those who favour the direct application of quantitative statistical techniques using a variable number of numerical parameters. On the other, we have those who defend the decisive role of territory in statistical analyses. In other words, two ways of approaching the statistical analyses of data can be found, which are the use of quantitative and exclusively numerical methods [43,44] and the use of geostatistics, with which the spatial criterion becomes important [41] in statistical calculations qualified by the territorial nature.

It is clear that the use of quantitative techniques has great advantages as it facilitates the analysis of considerable volumes of data and is capable of obtaining models and even predicting estimated values from these models. There is no doubt that this is one of the major contributions of multiple regression together with the discovery of outliers [45]. However, they do not always capture the essence of tourist activities, which on numerous occasions qualify all relationships, owing to the importance of territory. It should indeed be stressed that the role of geographical space in tourism is a vital and inherent part of the same [46–48].

The use of multiple regression to establish models and calculate the necessary supply of accommodation according to the tourist potential of a space is not new [49]. After having analysed a large amount of bibliography on the location of tourist accommodation, some authors concluded that the most logical course of action is to use geographical information systems [50,51]. The idea is that the location of the supply can be explained by using them; this can be complemented with the application of statistical techniques.

The putting of faith in this type of software is due to the conceptualisation of tourism as going beyond purely economic considerations in which territory and spatial relations acquire considerable

importance [52]. Many studies carried out decades ago [53] stressed the role of spatial analysis and remain valid today [54]. In consequence, both geography and territory science give tourism a different meaning, complementing it and providing it with a key tool: the geographical information system. With this contribution, together with progress in hardware and software, we have an efficient instrument for performing a tourist analysis based on territorial support itself, albeit without renouncing the quantitative analysis.

The need for including territory in tourist analyses arises as a consequence of the fact that tourist resources, accommodation, and complementary services, together with the tourists themselves, travel movements, etc., can be represented on the territory [55]. If this is the case, an obvious circumstance which is not questioned by the literature, it would be logical to think that it is necessary to adapt analytical techniques to this situation. It is likewise necessary to mention the importance of the criterion of proximity in tourism, as nearby spaces affect and are modified by this activity, influencing and altering strictly statistical models [56,57].

Studies applying geostatistical techniques to facilitate an understanding of tourism in its most varied facets have currently become relevant. In this sense, faith is placed in analysing patterns or mapping clusters according to the formulations of Getis–Ord [58,59] or Moran [60,61]. Despite the multiple applications of both groups of geostatistical techniques, when modelling spatial relationships are used, it is necessary to resort to geographically weighted regression [62].

The literature consulted confirmed that the application of geostatistical techniques is increasing in tourism studies. It is thus clear that the importance of the territory is vital if the results are to be interpreted correctly. In order to do so, the aim of the paper is to apply and compare the results obtained by means of ordinary least squares regression (OLS) and geographically weighted regression (GWR) in the distribution of the supply of rural accommodation in Extremadura, Spain.

If on a technical level there are numerous options for analysing tourism, when specific aspects of the latter are studied, significant problems also arise. In the case chosen as a subject of study, rural accommodation establishments in a specific territorial context, doubts always arise as to their appropriate location [63]. At the same time, we have tried to determine their optimum location [51,55]. Despite the considerable efforts which have been made, only in few cases have satisfactory explanations been obtained for the location of rural accommodation establishments. This is due to the fact that their location does not always follow a logical coherence which would imply resorting to areas with the most favourable characteristics or which are adapted to the preferences of the demand. Indeed, at times, their distribution appears to be random [46] as they are not always located in the areas with the greatest capacity for attracting tourists. These facts would appear to be corroborated by observation of the preferences of the demand for this type of establishment and the location of both the accommodation and tourist attractions. This imbalance may constitute a major hindrance for establishments and in consequence may affect their profitability.

The result of the circumstances mentioned is the need for detecting the main factors which attract tourists who stay at rural accommodation establishments and comparing them with the distribution of the lodging capacity they offer. In order to do so, various techniques are used, both statistical and geostatistical, which if appropriate allow the verification of the most significant models or guidelines for the implementation of rural accommodation establishments.

The study takes as a starting point the following hypotheses: (1) the distribution of rural accommodation establishments is not always due to criteria of suitability as the implementation of an establishment depends on human decisions which are at times variable and random; (2) it is necessary to delimit correctly the independent variables which may explain the location of the establishments; in order to do this, it is necessary to find out the opinion of tourists; (3) as the territorial component is not taken into account, OLS obtains worse results than GWR in the general regression model; (4) the local regression models obtained by GWR reveal their suitability on allowing the obtaining of a differentiated equation for each entity analysed; (5) it is possible to detect outliers in the territory which must be studied specifically; (6) predictive calculations on the lodging capacity of rural accommodation may

serve as a basis for planning action when this is necessary as there will not always be a balance between actual values and those calculated by the models built.

2. Materials and Methods

2.1. Study Area and Procedure

The study area chosen is Extremadura, Spain, an eminently rural region with a surface area of 41,634 km². This territory contains numerous tourist resources of a cultural and natural character which serve to attract tourists. Likewise, it has a considerable lodging capacity if its nature as an inland destination is taken into account. Despite this, the distribution of accommodation establishments in the region is unequal, although it does follow certain guidelines depending on the type of accommodation [41,46,64,65] (Figure 1).

Figure 1. Location of the study area.

In 2019, a total of 1768 accommodation establishments were registered with a maximum capacity of 42,214 beds (45,957 if extra beds are included) which have a variable distribution depending on the type of accommodation [66] (Table 1).

Table 1. Main tourist parameters of Extremadura (2019).

Type of Establishment	Establishments	Beds	Beds Including Extra Beds
Hotels	433	18,373	19,994
Rural accommodation	862	8933	11,055
Tourist apartment	396	3015	3015
Hostels	45	1827	1827
Campsites	32	10,066	10,066
Total	1768	42,214	45,957

In the specific case of the accommodation establishments considered as representing rural tourism, according to Decree 65/2015 which classifies their type [67], there has been an increase out of all proportion in recent decades [63]. This increase, which has been uncoordinated and poorly planned, has given rise to a clear imbalance between the supply and the demand. This has meant that during the recent economic crisis, some establishments have closed down [65]. It is evident that the reason for this needs to be understood. According to the literature focussing on the study area, apart from the economic situation, the most frequently heard explanation is the location of these establishments in areas lacking the attractions preferred by the demand [41,56,63,65].

This type of accommodation is subjected to other strains. These include few travellers and few overnight stays in comparison with other surrounding areas, low occupation levels, and shorter average stays [68] (Table 2).

Table 2. Main parameters of rural tourism establishments (2019).

Autonomous Region	Beds *	Travellers	Overnight Stays	Average Stay	Occupation (%)
Andalusia	21,642	395,723	1,358,967	3.43	18.06
Aragón	9644	194,182	581,749	3.00	17.03
Principality of Asturias	15,221	302,765	973,077	3.21	19.26
Balearic Islands	10,470	329,479	1,249,695	3.79	46.55
Canary Islands	3199	78,420	366,049	4.67	29.22
Cantabria	8784	263,220	625,328	2.38	24.56
Castilla y León	33,113	829,451	1,825,065	2.20	15.88
Castilla-La Mancha	14,283	313,030	688,244	2.20	13.05
Cataluña	18,790	496,748	1,244,328	2.50	19.03
Valencia	10,414	186,649	467,731	2.51	12.49
Extremadura	7758	224,908	490,949	2.18	17.64
Galicia	6496	228,449	425,974	1.86	19.81
Madrid	3753	169,950	294,231	1.73	21.76
Murcia	2406	56,850	125,553	2.21	13.97
Navarra	6821	136,901	449,619	3.28	19.89
Basque Country	4322	165,247	415,703	2.52	28.19
La Rioja	1168	38,958	87,777	2.25	21.09

* Data for August 2019.

2.2. Materials

The materials used for carrying out this research are of two types: alphanumeric and cartographical. Both have been implemented on a GIS, which has allowed the application of the techniques necessary for demonstrating the initial hypotheses.

The cartography used to perform the analysis is that of the National Geographic Institute (*Instituto Geográfico Nacional*, IGN) [69], which operates under the Creative Commons CC-BY 4.0 International licence. The scale of reference is 1:100,000 and the resolution is 20 m, which is sufficient to approach with guarantees to the type of analysis required by the research. It gives a vision of the territory which aims to describe the appearance and the details of the surface area together with the geographical objects found on it, and whether these are natural or a product of human activity. It allows the making of geographical and alphanumeric enquiries. Its class structure of geographical objects is based on the phenomena representing the scale mentioned with simple geometries (dot, line, and area) [70].

The alphanumeric information comes from various sources. That referring to tourist accommodation establishments comes from the Register of Tourist Companies of Extremadura [66]. It has been updated to 31 December 2019. This alphanumeric database has been conveniently georeferenced for its subsequent implementation in a geographical information system.

On the other hand, the information on the declared preferences of rural tourists regarding tourist resources comes from a direct survey. It contains data which allow the extraction of their sociodemographic profiles, the type of accommodation where they stay, the variety of tourism they practise, and naturally the elements they value in choosing their destination (Table 3). This survey was completed by 710 tourists in 2015 and was repeated by 140 in 2019 with the aim of contrasting the changes. As there are no significant differences between both years, neither in the preferences of tourists nor in the sociodemographic aspects, the entire sample was used (850 surveys). This is a reliable survey, since with 95% confidence the sampling error is 3.36%. From it, the tourist resources preferred by those staying in the study area have been discovered.

Table 3. Aspects included in the survey completed by the tourists.

Questions	Answers
Sex	Male/Female
Age	Year of birth
Studies	No studies/Primary studies/Secondary/University
Income	Up to EUR 20,000/EUR 20,000 to 30,000/Over EUR 30,000
Place of residence	Town or City
Place of overnight stay	Town/Towns
Type of tourism practised	Rural/Cultural/Rural and cultural
Accommodation type chosen	Tourist hostel/Tourist apartment/Campsite
	House of one's own or of friends or family/
	Holiday cottage/Hostelry/Hostel or guest house
	Spa hotel/1 to 3 star hotel/4 or 5 star hotel
Duration of the stay	Rural hotel/Not stated/Other accommodation
	No. of days
Activities carried out	Cultural visits/rural tourism/birdwatching/enjoying gastronomy/visiting caves or geological formations/using bathing areas
Daily expense per person	<EUR 25 /EUR 25 to 50/EUR 50 to 75 > EUR 75
Noteworthy reasons for choosing the destination	Bathing area/rivers or reservoirs/natural areas/type of vegetation/accessibility/cultural resources/climate/altitude/other (open reply)
Assessment of the destination	1 to 10
Observations	Free opinion

2.3. Research Procedure

Among the resources finally selected to explain the distribution of rural accommodation establishments, the criterion of distance was taken into account. To calculate it, we resorted on occasion to network analyst tools and at other times to the Euclidean distance taken from the resource to the capital of the municipality (Table 4).

Table 4. Alphanumerical variables and how they were obtained.

Variable	How Obtained
Mountain	Euclidian distance
Sierra	Euclidian distance
Bathing area	Shortest distance by road
Reservoir	Shortest distance by road
Monfragüe National Park	Shortest distance by road
Garganta de los Infiernos Nature Reserve	Shortest distance by road

Table 4. Cont.

Variable	How Obtained
Villuercas-Ibores- Jara Geopark	Shortest distance by road
Tajo-Internacional Nature Reserve	Shortest distance by road
Historical Ensemble	Shortest distance by road
Interpretation Centre	Shortest distance by road
Distance to highway	Shortest distance by road
Resident population	Direct allocation to the centre

Owing to the problems faced by rural accommodation establishments, the need for finding out whether they are located in optimum areas by establishing regression models is put forward. At a methodological level, the process followed is summarised in 5 stages (Figure 2).

Figure 2. The methodological diagram reflects the stages followed in the research.

The first stage consisted of the obtaining of information on the supply of accommodation, tourist resources, the territory analysed, and tourists. With these data, a database was designed which was implemented in a GIS in the second stage. This software, together with the statistical analysis of the database, allowed us to obtain the tourist potential of each of the municipalities analysed. The tourism potential is obtained considering the tourism resources that tourists prefer. The third stage consisted of determining the independent variables which explain the location of rural accommodation establishments and providing two differentiated models. The first of them is obtained from ordinary least squares regression (OLS) and the second by means of geographically weighted regression (GWR) in which the criterion of proximity is considered. The fourth stage concentrates on the analysis of the results obtained, while in the fifth, both models are discussed.

2.4. Exploratory Regression (OLS)

Tourist activities have a privileged position in the geographical space in which they are carried out. In a large proportion of cases, they are conditioned by a series of factors with a certain spatial continuity. The tourist resources present in the territory play their role of attraction in the centre where they are to be found, although they also benefit nearby areas; from this, it can be deduced that proximity may have a direct effect on the configuration of the tourist space [52].

It is likewise evident that resources may be superimposed in the territory analysed, which may give rise to an incorrect interpretation of the existing relationships. After assessing the advantages and disadvantages of various techniques which allow the detecting of strong correlations between the predicting variables, we decided to carry out exploratory regressions. By this, a set of models on which to decide is achieved.

In order to carry out these exploratory regressions, we decided to take as regressive variables the 12 which may influence the location of rural accommodation establishments (Table 4) and as a dependent variable the lodging capacity in rural accommodation. Both types of variables were allocated to each of the 388 population centres located in the area of analysis.

The application of the exploratory regressions was dependent on compliance with the following requirements:

- Adjusted R squared $\geq 0.3/\geq 0.2$;
- Variance inflation factor (VIF) ≤ 3 . This value reflects how much redundancy (multicollinearity) among the model explanatory variables can be tolerated. When the VIF (variance inflation factor) value is higher than about 7.5, multicollinearity may make a model unstable; consequently 7.5 is the default value here;
- Coefficient p -value < 0.05 . The selected value is 0.05, indicating passing models which will only contain explanatory variables when their coefficients are statistically at a 95 percent confidence level (p -values smaller than 0.05).

The viable models obtained are analysed to select that with the best requirements for the application of OLS and GWR regressions.

The basis of exploratory regression is ordinary least squares regression. This is used frequently in tourism studies either as a single technique or together with others [71,72]. Its main task consists of predicting the values attained by the dependent variable in accordance with the regressive variables which make up the model.

There tends to be considerable discussion on the final configuration of this model as to the possible problems of multicollinearity [73,74]. Nevertheless, more and more references are appearing which stress the important role played by remainders in the explanation of the same [75]. Some authors even recommend checking whether homoskedasticity or heteroskedasticity exists to validate the model, starting from the covariance matrix [76].

Although this technique is frequently used, the possible problems that may derive from the choice of an erroneous model should always be considered. The most frequent among them are perhaps the omission of explanatory variables, the existence of nonlinear relations, the presence of outliers, multicollinearity, the inconsistent variance in the remainders, etc. Despite this, almost all these disadvantages tend to have different solutions, among which the revision of the model stands out.

For the purpose of this article, ordinary least squares regression is carried out by means of the analysis module integrated in the spatial statistics tools included in the ArcGIS v.10.5 software.

The ordinary least squares method uses the following Equation (1) [77]:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + [\varepsilon] \quad (1)$$

in which:

- y_i = the value observed for the dependent variable at point i ;
- β_0 = the interception point y (constant value);
- β_n = the regression coefficient or slope of the explanatory variable n at point i ;
- x_n = the value of the variable n at point i ;
- $[\varepsilon]$ = the error of the regression equation.

For its application the following requirements must be complied with [77,78]:

- The model must be linear: this can be analysed by scatter plots;

- The data used must not depend on any external factor;
- The explanatory variables must not be related to each other;
- The explanatory variables must have an insignificant error when measured;
- The remainders must come to 0;
- The remainders must have a homogeneous variance and follow a normal distribution.

The definitive ordinary least squares regression model was subject to different tests to confirm its reliability. Among them, the F-statistic, Wald, Koenker, and Jarque–Bera statistics tests stand out.

2.5. Geographically Weighted Regression (GWR)

Ordinary least squares regression does not include in its formulation the criterion of proximity, which may compromise the models obtained in specific systems such as that of tourism. In a system as complex as the latter, it is inevitable to resort to the proximity analysis to explain certain facts. The presence of certain tourist resources thus favours not only the areas in which they are located but also those nearby, which shows the importance of proximity [79]. The need for including distance as a key parameter has been demonstrated by numerous publications which use geostatistical analyses [41,58,80–82].

In this context, spatially weighted regression emerges as a viable alternative for building models in which spatial relationships are analysed. This variety of regression allows the construction of a model by means of explanatory variables in the same way as OLS. However, it goes further as it specifies whether the kernel is constructed as a fixed distance, or whether it is allowed to vary in extent as a function of feature density [77]. Likewise, it allows specifying how the extension of the kernel can be determined, whether by means of the corrected Akaike’s information criterion (AICc), cross validation (CV), or bandwidth parameter.

This regression builds a different equation for each entity on including the dependent and explanatory variables of the entities in the bandwidth of each destination entity.

The following Equation (2) should be used:

$$y_i = \beta_{i0} + \sum_{k=1}^m \beta_{ik} x_{ik} + \varepsilon_i \quad (2)$$

in which y_i is the dependent variable at location i ; x_{ik} is the k th independent variable at location i ; m is the number of independent variables; β_{i0} is the intercept parameter at location i ; β_{ik} is the local regression coefficient for the k th independent variable at location i ; and ε_i is the random error at location i [48] (p. 2).

GWR allows coefficients to vary continuously over the study area, and a set of coefficients can be estimated at any location—typically on a grid so that a coefficient surface can be visualised and interrogated for relationship heterogeneity. GWR makes a point-wise calibration concerning a “bump of influence” around each regression point where nearer observations have more influence on estimating the local set of coefficients than observations farther away [47]. GWR measures the inherent relationships around each regression point i , where each set of regression coefficients is estimated by weighted least squares.

To calculate it, the following Equation (3) should be used:

$$\hat{\beta}_i = (X^T W_i X)^{-1} X^T W_i y \quad (3)$$

in which X is the matrix of the independent variables with a column of 1 s for the intercept; y is the dependent variable vector; $\hat{\beta}_i = (\beta_{i0}, \dots, \beta_{im})^T$ is the vector of $m + 1$ local regression coefficients; and W_i is the diagonal matrix denoting the geographical weighting of each observed data for regression point i [48] (pp. 2–3).

3. Results

3.1. Basis of the Regression Model: Activities Carried Out by Rural Tourists

The items used to determine the activities carried out by tourists staying in a rural accommodation establishment indicate which aspects are most attractive to them (Table 5). A detailed analysis shows that they carry out activities related to generic rural tourism, although they also make cultural visits. They declare that they practise tourism in rivers, gorges, or reservoirs and gastronomy. Finally, albeit less significantly, they also declare that they watch birds and visit mines or geological formations or practise sport. Meanwhile, the remainder of activities go unnoticed.

Table 5. Activities and motivations.

Activities/Motivations	N	%
Rural tourism	1267	28.82
Cultural visits	1108	25.20
Tourism in rivers, gorges, or reservoirs	636	14.46
Gastronomy	635	14.44
Birdwatching	245	5.57
Visiting mines, caves, or geological formations	149	3.39
Practising sport	143	3.25
Stargazing	69	1.57
Visiting wine cellars	61	1.39
Participating in events (congresses or meetings)	39	0.89
Visiting scenarios of films or TV series	18	0.41
Learning Spanish	14	0.32
Hunting	13	0.30

If the activities preferred by tourists staying in rural accommodation establishments are considered, it is possible to determine the factors which most attract them. From this, it can be deduced that it is a good way of understanding which should be the parameters used to analyse the implementation of rural accommodation establishments. These parameters include some which are highly characteristic of certain territories, although there are others which can be enjoyed over a large part of Extremadura, as is the case with gastronomy. The most noteworthy are:

- Cultural heritage. For this reason, proximity to historical ensembles has been chosen as a regressive variable; these ensembles also include World Heritage cities located outside the rural context;
- Proximity to water resources. As a consequence of their importance, bathing areas and reservoirs have been selected as factors justifying the presence of a rural accommodation establishment. We have also included the distances to two protected landscapes (the Garganta de los Infiernos Nature Reserve and the Tajo-Internacional Nature Reserve). The former is a well-known bathing area and the latter offers river cruises which take tourists even as far as Portugal;
- Birdwatching. In this sense, we have included the distance to Monfragüe National Park, the main centre of attraction for ornithological tourism;
- Visiting mines, caves, or geological formations. Owing to its importance, the distance to the Villuercas-Ibores-Jara Geopark has been considered.

The replies obtained in each of the types of resources analysed justify the fact that the main attractions sought by tourists staying at rural establishments act as regressive variables. It is logical to think that the areas containing a large proportion of these attractions are the most appropriate for maintaining this type of accommodation and will have more advantages if they are near a highway as a consequence of their better access.

The tourists themselves define a logical model for the distribution of rural accommodation establishments, which is at variance with the actual situation (Figure 3a,b). It can be observed that

there is a natural preference for certain areas although the offer of these accommodation establishments follows a markedly different pattern [63]. It can also be appreciated that there are certain differences depending on the population living in each municipality, especially if this is compared with the distribution of other types of accommodation [64].

Figure 3. Distribution of rural accommodation establishments and main (a) and other (b) regressive variables.

The distribution of rural accommodation and the presence of the resources which support the activities carried out by tourists closely coincide in certain areas. In this sense, there is an important relationship between mountain areas and the presence of bathing areas; the latter represent the highest concentrations of the lodging capacity in rural accommodation establishments. This situation is understandable if it is taken into account that summer is the time of the year when most rural tourists are received [64,65]. It can likewise be observed that in the proximity of the sierras, numerous beds in rural accommodation establishments are also concentrated, although in many cases they do not coincide with areas prepared for bathing with the subsequent loss of attraction for summer tourism. Moreover, there are also accommodation establishments in areas lacking the main attractions sought by the demand (Table 6).

This most inconsistent situation gives rise to three major groups of rural accommodation establishments. The first of these, located in mountain areas and with the presence of bathing areas, has the best competitive advantages for attracting tourists. The second also has a certain capacity of attraction, although it is not so competitive during the long summer period characteristic of Extremadura, where temperatures often exceed 35 degrees and on occasion reach 40 degrees [83]. Finally, the third, which is unequally distributed over the territory, is that near Special Bird Protection Areas (*Zonas de Especial Protección para Aves*, ZEPAs). This distribution, which is more or less satisfactory, is faithfully reflected in its capacity for capturing tourists according to the time of the year. This is corroborated when the areas located in the north of Extremadura, a mountain region with natural

bathing areas, attract a large number of tourists and overnight stays during the summer. Meanwhile, other areas have greater potential during the rest of the year [84].

Table 6. Tourist figures in rural accommodation establishments according to types of landscape.

Tourist Territories	Beds	Travellers	Overnight Stays	Average Stay	Annual Occupation (%)	Type of Landscape
Valle del Jerte and La Vera	2595	46,546	100,383	2.16	31.36	Mountain
Sierra de Gata, Las Hurdes, and Valle del Alagón	1438	30,163	71,981	2.39	21.91	Mountain
Valle del Ambroz and Tierras de Granadilla	921	37,810	89,137	2.36	46.97	Mountain
Alqueva, Sierra Suroeste, and Tentudía	505	10,618	20,288	1.94	25.41	Sierra
Trujillo, Miajadas, and Montánchez	671	19,427	40,484	2.08	25.51	Peneplain
Monfragüe Biosphere Reserve	491	5900	12,816	2.17	27.30	Protected space
Vegas del Guadiana	239	10,775	20,819	1.93	25.96	River valley

It is obvious that the essential attractions sought by the demand are mountain or sierra areas, bathing areas, and some natural spaces, although one should always stress the enormous territorial variability regarding the lodging capacity. For this reason, despite the fact that there are clear determinants which explain the presence of the rural accommodation establishments available in specific areas, it is evident that it would be viable at least in theory to model their distribution on the presence or proximity of specific factors. Nevertheless, certain conditions can be observed among some of these factors, which means that it is as well to improve these models by using specific techniques which make it possible to eliminate the effects of collinearity.

3.2. Exploratory Regression Model

Taking into account the contrasts between the distribution of accommodation establishments and the presence of the main factors mentioned by the demand staying in rural establishments, we chose to carry out an exploratory regression. Its main advantage lies in the fact that it includes all possible combinations of the regressive variables which attempt to explain the number of rural accommodation beds in each population centre.

This type of regression also allows the use of complementary criteria to select variable models in addition to value R.

After its application, a considerable variety of models were found, of which three are worthy of note. The first takes into account all the variables which may act as regressors, while the second has been improved, taking into account exclusively the variables complying with some reliability requirements. In contrast, the third model uses as regressive variables the factors which can be deduced from the activities carried out by the type of tourist analysed (Table 7).

To obtain the exploratory regression mode, slightly different criteria were used (Table 8). In the first place, an adjusted R-squared minimum value exceeding 0.3 was required; this value was not attained in any of the 4095 combinations analysed. Likewise, a p -value of <0.05 was required; this condition was complied with in 399 cases. A VIF value of <3.00 was also required; this condition was complied with by 1695 models. On the contrary, no possible combination appeared if a p -value (JB) of >0.05 was considered or a spatial autocorrelation (SA) of >0.20 . In contrast, the second model saw the adjusted R-squared value lowered to figures above 0.2, which was achieved in two models. Meanwhile, a p -value of <0.05 was provided by 49 models, and a VIF value of <3.00 obtained 63 viable results. Nevertheless, although greater permissiveness is possible with a p -value (JB) or a spatial autocorrelation (SA) of >0.1 , no viable model is achieved. Finally, the third model presents 511 tests. A total of 17.81% of these comply with the minimum adjusted R-squared value considered. In turn, they also comply with the

p -value and VIF value criteria in between 29.75% and 43.64% of cases. However, the robustness of the models is compromised if we should resort to a p -value (JB) or a spatial autocorrelation (SA).

Table 7. Regressive variables applicable.

Distance to:	Model 1 (Not Improved)	Model 2 (Improved)	Model 3 (Improved)
Historical Ensemble	X	X	X
Interpretation Centre	X		
Highway	X	X	X
Reservoir	X	X	X
Bathing Area	X	X	X
Monfragüe National Park	X		X
Garganta de los Infiernos Nature Reserve	X		X
Mountains	X	X	X
Sierras	X	X	
Villuercas- Ibores-Jara Geopark	X		X
Tajo-Internacional Nature Reserve	X		X
Population in 2019	X		

Dependent variable: Beds in rural accommodation establishments.

Table 8. Criteria used in exploratory regression.

	Search Criterion	Cutoff	Trials	# Passed	% Passed
MODEL 1	Min Adjusted R-Squared	>0.30	4095	0	0
	Max Coefficient p -value	<0.05	4095	399	9.74
	Max VIF Value	<3.00	4095	1695	41.39
	Min Jarque-Bera p -value	>0.05	4095	0	0
	Min Spatial Autocorrelation p -value	>0.20	34	0	0
MODEL 2	Min Adjusted R-Squared	>0.20	63	2	3.17
	Max Coefficient p -value	<0.05	63	49	77.78
	Max VIF Value	<3.00	63	63	100
	Min Jarque-Bera p -value	>0.1	63	0	0
	Min Spatial Autocorrelation p -value	>0.1	16	0	0
MODEL 3	Min Adjusted R-Squared	>0.20	511	91	17.81
	Max Coefficient p -value	<0.05	511	152	29.75
	Max VIF Value	<3.00	511	223	43.64
	Min Jarque-Bera p -value	>0.1	63	0	0
	Min Spatial Autocorrelation p -value	>0.1	16	0	0

These first results show how difficult it is to explain the distribution of the beds available in rural accommodation establishments in a logical manner. It is also as well to point out that there is a degree of territorial concentration of the same in some places, which undoubtedly affects the fact that there is a similar concentration in the remainders obtained with these regression tests [55].

The significance of the variables considered in each model certifies the consistency of each of them (Table 9). It can thus be observed that while Model 1 attaches much importance to most variables, Model 2 distinguishes them better. In this sense, it is revealed that the relief (mountains and sierras) and the water resources (bathing areas and reservoirs) reach significant values; cultural attractions and the distance from highways are also noteworthy. Model 3 shows similarities with Model 1 in the variables which have been used jointly.

This analysis demonstrates that Model 2 is more accurate as it selects fewer variables, but all of them are highly significant from a statistical point of view. Moreover, they agree with the main attractions which tourists emphasise. This same model reflects the importance of being near mountain and bathing areas, as their negative significance is 100%, which is somewhat similar to the case of

historical ensembles. This is in addition to the fact that there is a positive relationship with sierras, reservoirs, and the distance from highways.

Table 9. Summary of variable significance.

Distance to:	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	% Significant			% Negative			% Positive		
Historical Ensemble	99.46	100.00	98.83	100.00	100.00	100.00	0.00	0.00	0.00
Interpretation Centre	98.54			100.00			0.00		
Highway	92.14	93.75	89.84	0.00	0.00	0.00	100.00	100.00	100.00
Reservoir	88.62	78.12	91.80	0.00	0.00	0.00	100.00	100.00	100.00
Bathing Area	83.54	100.00	87.11	98.44	100.00	100.00	1.56	0.00	0.00
Monfragüe National Park	70.41		77.73	26.46		26.17	73.54		73.83
Garganta de los Infiernos Nature Reserve	63.62		73.05	90.77		92.19	9.23		7.81
Mountains	51.22	96.88	55.08	79.15	100.00	84.77	20.85	0.00	15.23
Sierras	34.47	81.25		5.86	0.00		94.14	100.00	
Villuercas-Ibores-Jara Geopark	31.59		39.45	43.65		38.67	56.35		61.33
Tajo-Internacional Nature Reserve	31.15		41.42	27.00		22.27	73.00		77.73
Population in 2019	3.17			35.84			64.16		

This initial analysis reflects the possibility of finding two well differentiated types of areas. On the one hand, we have mountain areas clearly linked to bathing areas, and on the other, areas far from these spaces containing reservoirs and protected areas.

Despite the interest of the analysis of statistical significance, the fact that multicollinearity is one of the most serious problems that may arise in the application of multiple regression cannot be ignored. For this reason, its analysis is essential if we are to determine the goodness of the regression models constructed (Table 10). It is noteworthy that Model 1 has three variables with unacceptable VIF values, which gives rise to their participation in numerous models with multicollinearity (violations). In contrast, Model 2 represents precisely the opposite as all the variables included have very low VIF values, especially if it is taken into account that the maximum limit established is 7.5. Clearly, having low VIF values means that no violation by collinearity occurs. For its part, Model 3 is not robust as there are three variables with a VIF much higher than the value statistically permitted (7.5).

Table 10. Summary of multicollinearity.

Distance to:	Model 1		Model 2		Model 3	
	VIF	Violations	VIF	Violations	VIF	Violations
Historical Ensemble	50.45	1664	1.49	0	1.94	0
Interpretation Centre	2.55	0				
Highway	2.35	0	1.46	0	2.83	0
Reservoir	1.37	0	1.22	0	1.33	0
Bathing Area	61.55	1664	1.43	0	2	0
Monfragüe National Park	66.05	1632			57.15	208
Garganta de los Infiernos Nature Reserve	5.62	490			42.97	200
Mountains	6.12	601	1.90	0	23.07	200
Sierras	2.01	0	1.86	0		
Villuercas-Ibores-Jara Geopark	1.33	0			4.79	50
Tajo-Internacional Nature Reserve	3.18	64			4.49	56
Population in 2019	1.07	0				

Table 11. Assessment of models.

Method	MODEL 1		MODEL 2		MODEL 3	
	Parameters	Independent Variables	Parameters	Independent Variables	Parameters	Independent Variables
Highest Adjusted R-Squared	AdjR ² : 0.24	(−) Mountain ***	AdjR ² : 0.21	(−) Mountain ***	AdjR ² : 0.22	(−) Bathing area ***
	AICc: 3850	(−) Bathing area ***	AICc: 3864	(−) Bathing area ***	AICc: 3860	(+) Reservoir ***
	JB: 0.0	(+) Reservoir ***	JB: 0.0	(+) Reservoir ***	JB: 0.0	(−) Garganta de los Infiernos NR ***
	K(BP):0.0	(−) Historical Ensemble ***	K(BP):0.0	(−) Historical Ensemble ***	K(BP):0.0	(−) Historical Ensemble ***
	VIF: 1.40 SA: 0.0	(−) Interpretation Centre *** (+) Highway ***	VIF: 1.40 SA: 0.0	(+) Highway ***	VIF: 1.45 SA: 0.0	(+) Highway ***
Residual Normality (JB)	JB: 0.0		JB: 0.0		JB: 0.0	
	AdjR ² : 0.086		AdjR ² : 0.086		AdjR ² : 0.008	
	AICc: 3918		AICc: 3918		AICc: 3949	
	K(BP): 0.000004	(+) Sierra ***	K(BP): 0.000004	(+) Sierra ***	K(BP): 0.000004	(+) Reservoir**
	VIF: 1.00 SA: 0.0		VIF: 1.00 SA: 0.0		VIF: 0.13 SA: 0.0	
Residual Spatial Autocorrelation (SA)	SA: 0.0465	(+) Sierra	SA: 0.00006	(−) Mountain ***	SA: 0.0465	(−) Bathing area ***
	AdjR ² : 0.25	(−) Bathing area **	AdjR ² : 0.21	(−) Bathing area ***	AdjR ² : 0.23	(+) Reservoir ***
	AICc: 3849	(+) Reservoir ***	AICc: 3864	(+) Reservoir ***	AICc: 1391	(+) Monfragüe NP
	JB: 0.0	(+) Monfragüe NP	JB: 0.0	(−) Historical Ensemble ***	JB: 0.0	(−) Garganta de los Infiernos NR **
	K(BP):0.0	(−) Garganta de los Infiernos NR **	K(BP):0.000005	(+) Highway ***	K(BP):0.0	(−) Villuercas-Ibores-Jara Geopark
	VIF: 60.62	(−) Villuercas- Ibores-Jara Geopark	VIF: 1.40		VIF: 53,39	(−) Tajo Internacional NR
		(−) Tajo Internacional NR				(−) Historical Ensemble ***
		(−) Historical Ensemble *** (−) Interpretation Centre *** (+) Highway ***				(+) Highway ***

Note: AdjR²: adjusted R-squared; VIF: max variance inflation factor; AICc: Akaike's information criterion; SA: global Moran's I *p*-value; JB: Jarque–Bera *p*-value; model variable sign: (+/−); K(BP): Koenker (BP) statistic *p*-value; model variable significance: (** = 0.05; *** = 0.01).

The assessment of the models obtained indicates that Model 2 has the best conditions for explaining the presence of the supply of rural accommodation (Table 11). Indeed, the most viable model constructed consists of the main attractions required by the rural tourist. Nevertheless, it should be taken into account that this model attains an R^2 of 0.21, although it is also in keeping with the SA. To this, what must also be added is the fact that the variables selected by the model itself are significant at a level of 0.01. From this, it can be deduced that it is unlikely that the result is due to chance.

The above analyses show that there is a model which can provide an explanation according to the activities carried out by tourists, although it provides a reduced explanation of the dependent variable. For this reason, it is difficult to find a logical explanation for the distribution of the supply of rural accommodation.

3.3. Ordinary Least Squares Regression (OLS)

The results given by OLS in Model 2 include interesting data because the VIF has a lower value in all the regressive variables. This shows that there is no collinearity between the variables. Likewise, it confirms that the coefficients obtained by each of these variables provide an explanation of the model, although three stand out negatively (historical ensemble, bathing area, and mountain) (Table 12). This means that the centres nearer these attractions have a higher number of beds, which is in keeping with the activities carried out by the tourists staying in rural accommodation establishments, at least during the summer. Nevertheless, the model is completed with two other positive coefficients (reservoir and highway), from which it can be deduced that the further a population centre is from highways and reservoirs, the greater the capacity for rural accommodation. All this is faithfully reflected in the current situation, as shown in Figure 3. On the other hand, both probability and robust probability are significant, from which it can be deduced that the possibility of the coefficient being essentially zero is also low.

Table 12. Summary of ordinary least squares (OLS) results—model variables and diagnostics.

Variable	Intercept	Mountain	Bathing Area	Reservoir	Historical Ensemble	Highway
Coefficient (a)	38.08944	−0.00014	−0.000402	0.000381	−0.000636	0.000525
StdError	4.869659	0.000029	0.000107	0.000135	0.000118	0.000089
t-Statistic	7.821788	−4.792055	−3.767674	2.818674	−5.409349	5.886486
Probability (b)	0.000000 *	0.000003 *	0.000201 *	0.005075 *	0.000000 *	0.000000 *
Robust_SE	5.811659	0.000027	0.000092	0.000143	0.000131	0.000098
Robust_t	6.55397	−5.129564	−4.378485	2.662947	−4.83881	5.328641
Robust_Pr (b)	0.000000 *	0.000001 *	0.000019 *	0.008069 *	0.000003 *	0.000000 *
VIF (c)	—	1.299913	1.355292	1.195813	1.403449	1.339888

* An asterisk next to a number indicates a statistically significant p -value ($p < 0.01$); (a) Coefficient: this represents the strength and type of relationship between each explanatory variable and the dependent variable; (b) Probability and Robust Probability (Robust_Pr); (c) variance inflation factor (VIF).

Despite the fact that the regressive variables correspond to the attractions valued by tourists who stay in rural accommodation establishments, a full diagnosis of the model must be carried out (Table 13). It is from this analysis that the first signs of complexity arise. This is due to the fact that an explanation can only be found for 21% of the distribution of beds in rural accommodation establishments. Moreover, the statistics of Koenker and de Wald are significant. Owing to this, OLS reflects that the lodging capacity is not strongly linked to the presence of the essential attractions for rural demand. This may mean that the location of these establishments does not follow a logical criterion but rather other factors which have little to do with the attraction capacity and the tourist potential of Extremadura.

Table 13. OLS diagnostics.

Input Features:	Population Centres	Dependent Variable:	Rural Accommodation
Number of Observations:	388	Akaike's Information Criterion (AICc)	3864.24795
Multiple R-Squared	0.223397	Adjusted R-Squared	0.213232
Joint F-Statistic	21.977147	Prob (>F), -5382 degrees of freedom:	0.000000 *
Joint Wald Statistic	43.364549	Prob (>chi-squared) -5 degrees of freedom:	0.000000 *
Koenker (BP) Statistic	32.437577	Prob (>chi-squared) -5 degrees of freedom:	0.000005 *
Jarque-Bera Statistic	3271.23973	Prob (>chi-squared), -2 degrees of freedom:	0.000000 *

* An asterisk next to a number indicates a statistically significant p -value ($p < 0.01$).

Both the low level of explanation of the model and the negative assessments of the main tests used are a clear sign that there is no obvious trend to explain the distribution of the beds of accommodation establishments. This peculiarity is due to the fact that setting up an establishment of this type is a personal decision that does not always follow a logical criterion.

The graphic analysis of the histograms showing the distribution of the regressive variables and the linear regressions obtained by means of a scatter plot demonstrate that there are no atypical values conditioning the relationship and the tendency (Figure 4). However, considerable bias can be observed in the number of beds in rural accommodation establishments with a clear predominance of lower values. Indeed, of the 388 municipalities analysed, 141 have no beds and only 57 population centres have over 50 beds.

Figure 4. Variable distributions and relationships.

The analysis of the standardised results reflects a Gaussian distribution of frequencies and the positive bias is minimal (Figure 5a). This is corroborated when resorting to the spatial autocorrelation index obtained by means of Moran's I (Figure 5b), the value of which attains 0.063515 to certify a random distribution of the remainders. Moreover, the territorial distribution of the same reveals no concentrations that may lead us to suspect the omission of a specific regressor (Figure 5c).

3.4. Geographically Weighted Regression

The application of OLS has shown that it is difficult to obtain a satisfactory global regression model. This is essentially due to the unsatisfactory explanation of the dependent variable and the non-compliance of some indexes which check the suitability of the set of regressive variables. Other parameters however exist which suggest the possible applicability of the same, although they show ambiguities. For this reason, it is a case of discovering whether any changes occur by resorting to the conceptualisation of the territorial aspect beyond its own geographical representation. The need

thus arises to contrast this model by GWR with the essential objective of revealing whether any other territorial pattern exists which has escaped analysis.

The conceptualisation of the model applied by GWR is similar with the aim of being able to contrast the results and confirm which technique is best fitted to the situation deriving from the territorial distribution of rural establishments. The decision was taken to use the same regressive variables and naturally maintain the lodging capacity in rural accommodation establishments as a dependent variable. Nevertheless, this type of regression allows the adding of a variable which acts as spatial weighting for each individual entity. For this reason, the number of rural accommodation establishments has been included as a calibration factor of the model.

Figure 5. Analysis of the remainders: (a) Gaussian distribution of standardised remainders, (b) Moran's I spatial autocorrelation; (c) territorial distribution of the standardised remainders.

From then on, different combinations have been applied to select the kernel of the function and the most suitable method used to calculate its extent. In the case of the selection of the kernel, it was decided to use both fixed and adaptive parameters. The former covers a fixed distance (metres) while the latter considers a variable distance depending on a specific number of neighbours. In both cases, it is calculated automatically in the case of the corrected Akaike's information criterion (AICc) and cross validation (CV) and selected by the authors in the case of the bandwidth parameter.

A synthesis of the results obtained shows that, overall, they are better than those obtained by OLS (Table 14).

An analysis of the different methods used reflects that AICc and CV provide superior performances to bandwidth parameter, both resorting to fixed kernels and to adaptive kernels. This is due to the fact that the condition number does not exceed the critical value of 30. Even so, the most effective result is achieved by means of the application of AICc with fixed kernels. In this case, the automatic distance calculation indicates that the models are optimised by using a bandwidth of 63.627 metres. It can likewise be observed that the difference between the R^2 and the adjusted R^2 is not very high and that

the proportion of variance of the variable explained is reasonable if the complexity of the model is taken into account.

Table 14. Synthesis of geographically weighted regression (GWR) models obtained.

	Method	Corrected Akaike's Information Criterion (AICc)	Cross Validation (CV)	Bandwidth Parameter
Kernel Fixed	Bandwidth (metres)	63,627.4	2,677,852.7	15,000
	Residual Squares (a)	2,815,576.7	3,759,062.1	1,248,124.3
	Effective Number (b)	28.0	6.0	31.7
	Sigma (c)	88.4	99.2	123.9
	AICc (d)	4598.9	4676.7	1443.2
	R ² (e)	0.475	0.299	0.665
	R ² Adjusted (f)	0.436	0.290	0.538
	Condition number (g)	<30	<30	>30
Kernel Adaptive	Neighbours	158	388	15
	Residual Squares (a)	2,572,646.7	3,531,296.9	33,308.1
	Effective Number (b)	35.2	12.3	7.7
	Sigma (c)	85.4	96.9	159.4
	AICc (d)	4579.1	4661.0	−32,474.5
	R ² (e)	0.520	0.342	0.966
	R ² Adjusted (f)	0.474	0.322	0.791
	Condition number (g)	<30	<30	>30

(a): The sum of the residual squares obtained by the model; (b): this value reflects a balance between the variance of the adjusted values and the influence on the coefficient estimates; (c): the square root of the standardised residual sum of squares in which the latter is divided between the effective degrees of freedom of the residual; (d): this is a measurement of the performance of the model and is useful for comparing different regression models; (e): a measurement of the goodness of the adjustment; (f): the calculations for the adjusted square *r*-value standardise the numerator and denominator by means of the degrees of freedom; (g): this assesses the local multicollinearity.

The territorial distribution of the local regression models reveals that the R² follows certain patterns (Figure 6a). In this sense, the best adjustments are obtained in the district of Alcántara and its vicinity where they are over 40%, although there are other areas in which 30% is exceeded such as the Villuerca-Ibores-Jara Geopark. In contrast, the adjustments of the areas of the north of the province of Cáceres or the south-centre of the province of Badajoz are lower than 20%.

The territorial analysis of the remainders obtained by the model shows a negative imbalance in those areas with the highest number of beds (Figure 6b). In the north of the province of Cáceres therefore, the model estimates a higher lodging capacity than that existing in a large proportion of the centres. This is due to the considerable potential for development which said areas still have as a consequence of the superimposition of the resources preferred by the demand from tourists. In contrast, in these areas, there are also centres in which the model estimates a number of beds lower than the actual number to coincide with the centres with smaller and more aged populations and with a very limited capacity for enterprise. In the same way, other areas which are very well delimited geographically can be observed in which there are centres with a considerable imbalance, both positive and negative, between the values calculated by the model and those actually found (Table 15).

Figure 6. Local R^2 (a) and residual (b).

Table 15. Extreme results of positive and negative residuals (detail).

Centre	Beds Observed	Beds Calculated	Residuals
Navaconcejo	298	155	143
Hervás	253	139	114
Jarandilla de la Vera	254	140	114
Gargantilla	8	126	−117
Talaveruela de la Vera	10	152	−141
Viandar de la Vera	0	150	−150

The distribution of the values observed and those predicted by the model (Figure 7a,b) corroborates certain guidelines. On the one hand, it is clear that the lodging capacity of rural accommodation establishments is concentrated in very few areas. In them, there are population centres with many beds in this kind of accommodation also intermingling with others in which the number of beds is more limited or even non-existent. On the other hand, when the distribution of the number of beds estimated by the regression model is analysed, it can be observed that they are concentrated in specific areas such as the north-east and other well-defined areas.

The comparison of the beds available in rural accommodation establishments with those calculated by the model reveals that the areas where the latter best operates maintain a certain balance between both variables (Figure 8). When the situation of other areas in which the operation of the model is less adequate is analysed, the results are similar. Nevertheless, these adjustments are acceptable when R^2 exceeds 0.3. In contrast, the opposite is true in those areas of Extremadura in which rural tourism is more developed, such as the northern district, as a consequence of having less satisfactory adjustments. This scenario is completed with the existence of numerous population centres which have no beds available, although they are allocated some by the model.

Figure 7. Beds in rural accommodation establishments (a) and predicted values (b).

An analysis of the results obtained by means of GWR shows the complexity of trying to explain the distribution of rural accommodation establishments by tourist preferences, which reveals that the implementation of the same has not followed an objective criterion. Far from it, in certain places, their location follows an almost random criterion which may perhaps be explained by other variables which have nothing to do with the tourism and business logic that should govern any tourist business.

Figure 8. Graphic distribution of the beds observed vs. the beds calculated.

4. Discussion

Regression models aim to explain the relationship between independent variables and the dependent variable, although in cases in which human decisions or behaviour intervene, it is always difficult to find strong causal relationships [75]. This is clear in the case analysed in which we aim to explain whether the location of rural accommodation establishments and consequently the presence of this type of supply follows logical criteria or depends on personal inconsistent decisions, as was found in some studies [64].

In order to obtain probable explanations which reveal the incidence of a set of explanatory variables on an explained variable, one generically resorts to the application of regressions of multiple types. For this reason, taking into account the literature analysed, we decided to use ordinary least squares regression which is one of the techniques most frequently used by scientists [85–88], although given the latter's limitation when contextualising spatial relationships, geographically weighed regression is

becoming more and more acceptable in analyses of tourism [89,90]. The essential objective of using these two types of analysis is to determine which of the two best adapts to the creation of models explaining the location of the supply.

In both cases, it has been found that it is very difficult to achieve a valid model which complies with the statistical requirements necessary so that predictions can be made according to the same. In this sense, it is clear that GWR obtains better results than OLS (hypothesis 3), although both types of regression show that the distribution of rural accommodation establishments is not ideal; this means that they also refute other techniques such as grouping analysis [57], hot spot analysis, or cluster and outlier analysis [41].

The difficulty in defining a suitable model to explain the distribution of rural accommodation establishments is confirmed when the preferences of the demand for rural tourism in the area studied are taken into account. For this reason, it can be inferred that their location depends on a statistically complex environment in which decisions are made on many occasions, owing to a personal criterion supported by expansive tourist policies which sometimes lack rigour. These policies pursued the profusion of the supply to the detriment of other necessary options. We can thus understand declarations calling for the need to implement a supply of quality tourist products which are capable of making destinations more attractive. Among these products, the following stand out: dehesas as an essential resource to support currently non-existent agrotourism [91], bathing areas and reservoirs [92], protected areas [93], the historical heritage [94], and very specific segments such as hunting [95].

The demonstration that the distribution of the pool of rural accommodation establishments is not ideal, as there are areas of this supply which do not have the necessary attractions in the opinion of the demand, makes it clear that it is necessary to delimit satisfactorily the explanatory variables which can reliably define and delimit the location of this type of supply (hypothesis 1). In order to do so, it is recommendable to use the criterion of the demand of the tourist type analysed. In this sense, the carrying out of surveys in which the tourist expounds the activities he/she prefers assumes that the criteria established are much more realistic and effective at the end of the day than any other method. It should not be forgotten that when the tourist indicates his/her preferences, he converts them into rigorous variables which would explain the presence of accommodation establishments (hypothesis 2). As a result, it is considered that above all the main attractions that must act as explanatory variables are mountain areas, bathing areas, and historical-artistic ensembles, locations which are also in areas far from highways and reservoirs. The simple superimposition of these locations on the supply of rural accommodation establishments shows that some of the locations in which they are available do not have this kind of attraction. This undoubtedly means that this location does not follow logical criteria, which explains that in these areas lacking the tourist attractions required by tourists, the occupation level is appreciably lower than in other areas better adapted to the preferences of tourists.

This fact is revealed by both techniques, although it is obvious that spatially weighted regression is much more efficient as it is capable of determining the areas to which the regression models are better adapted. In this respect, it is as well to remember that this type of regression is capable of generating a different equation for each of the locations selected (hypothesis 4). Owing to this, its analytical superiority is understood. It is interesting that there are areas in which the adjustment of the models can be improved; they coincide with areas in which the attractions required by tourists do not abound. Moreover, the models themselves described define some outliers, imbalances in short, owing to the existence of a specific criterion such as the presence of a low demographic volume (hypothesis 5). In this sense, it is understood that small population centres, normally with less than 500 inhabitants, lack the endogenous capacity for enterprise and also suffer from a scarcity of services and tourist equipment.

By taking into account the aspects listed, it can be affirmed that the distribution of rural accommodation establishments does not correspond as it should to the presence of the lodging preferred by tourists. In its turn, it is also noteworthy that ordinary least squares regression is not the most effective technique for explaining this distribution, not even if its aim is to provide a generic

model. In contrast, spatially weighted regression operates better as it explains a higher percentage of the variance and also defines areas in which the models constructed are more effective.

The analyses carried out reveal that GWR is much superior to OLS, even when analysing models based on variable behaviour such as that of human decisions, far from being adapted to physical laws or following coherent guidelines, and predicting the supply of accommodation. This is because although both types of regression obtain general models, only GWR allows local models adapted to each individual case (hypothesis 6). Moreover, it has another consubstantial advantage when tackling spatial statistics, i.e., the consideration of multiple types of spatial relationships which is particularly useful for tourism. In this discipline, it has been shown on numerous occasions that proximity is an essential element as is also in a sense the first law of geography, which indicates that interactions between places are inversely proportional to the cost of the distance between them [79].

The very complexity of the model analysed leads us to think that this type of technique is very useful, even in erratic systems such as that analysed, and facilitates decision-making at least in the places better adapted to the models designed. Nevertheless, both the global and local models defined by GWR can be improved from a statistical perspective by adding another type of explanatory variable, although when this possibility was explored with the inclusion of the most emblematic protected natural spaces of the study area, it was found that statistical consistency fell. Due to this, the results given here were obtained after multiple improvements and discarding the most problematical models, seeking the best relationship between a logical explanation and one adapted to the preferences of tourists and statistical robustness.

The applicability of this research allows both the local and regional administration to define a tourist policy conducive to the search for alternatives in the designing of tourist products in the most problematical areas and to stress the value of other attractions of these territories which are not mentioned by tourists. Indeed, it can be observed that many areas, in which the supply is not adapted to the resources required by visitors, contain other elements which may be of interest such as protected natural spaces. In them, specific products could be created, orientated towards nature tourism, ornithological tourism, or even agrotourism, as many spaces feature the typical landscape of the southwest of the Iberian Peninsula, the “dehesa”.

All this opens a new line of analysis based on finding explanations for the suitable location of investments, in this case, illustrated with the supply of rural accommodation in a specific context, Extremadura. This line would improve the model, perhaps including more attractions so that the tourist can define it better.

5. Conclusions

The analysis of the distribution of the supply of rural accommodation in Extremadura, using two types of regression widely accepted by the scientific community, leads to the following conclusions:

The first shows that the location of the supply of rural accommodation is random over a large part of the territory, as is shown by the fact that there is no correspondence between the resources required by the tourist and those present in said areas. Fortunately, the centres in which there are fewer tourist attractions according to the demand tend not to have many beds. In contrast, where this correspondence does exist, an important pool of lodging is detected. In line with the above, it is shown that the north of the province of Cáceres, recognised by the INE as an outstanding area for rural tourism at a national level, together with other areas with the highest relief in Extremadura, does have the necessary attractions and moreover a considerable number of rural accommodation beds.

The second is that the most important attractions for the demand from rural tourists who stay in this type of accommodation are varied, although those predominating are altitude, bathing areas, and historical-artistic ensembles. It is indeed observed that the distance to these elements has a negative relation with the number of beds available. In other words, the nearer the centres are to them, the greater their capacity for rural accommodation.

The third reveals that the application of OLS does not provide a convincing explanation for the model defined by tourists, despite the fact that it is a global statistic, on defining a single equation to explain the relationship between the independent variables and the dependent variable. This is due to the fact that the variability of circumstances reflected in the territory conceals different realities and cannot therefore be synthesised in a single equation.

The fourth is that GWR gives better results for both the global model and the local models which it is capable of creating. Moreover, it allows the distinguishing of spaces in which the model is more reliable than others in which the level of adjustment is appreciably lower. As has been mentioned, the reason for this is that different situations require a differentiated analysis, which this regression allows as it has not only a global outlet but also a local one, calculating a specific case for each case analysed.

The fifth shows that the configuration of spatial relations is necessary to explain models in which territorial proximity is vital. In this case, tourist activities are a clear exponent as proximity to resources is a clear sign which can explain the presence of the supply of accommodation.

The sixth and last conclusion refers to the need to continue with a line of research into the application of spatial statistics to all those systems in which the territory can be considered as one more variable to analyse. For this reason, when both regression models have been compared, the immediate conclusion has been the goodness of GWR compared with the poor results of OLS.

Author Contributions: Conceptualization: J.-M.S.-M., J.-I.R.-G., and J.-L.G.-G.; methodology: J.-M.S.-M.; formal analysis: J.-M.S.-M. and J.-L.G.-G.; investigation: J.-M.S.-M.; resources: J.-M.S.-M.; data curation: J.-M.S.-M., and J.-I.R.-G.; writing—original draft preparation: J.-M.S.-M., J.-I.R.-G., and J.-L.G.-G.; writing—review and editing: J.-M.S.-M. and J.-L.G.-G.; supervision: J.-M.S.-M. and J.-L.G.-G.; project administration: J.-M.S.-M.; funding acquisition: J.-M.S.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by DISEÑO Y ELABORACIÓN DE PRODUCTOS ESTRATÉGICOS DIFERENCIADOS PARA LA POTENCIACIÓN DEL TURISMO RURAL EN EXTREMADURA. DE LA DETECCIÓN DE PROBLEMAS A LA PROPUESTA DE SOLUCIONES BASADAS EN CRITERIOS GEOESTADÍSTICOS, grant number PRI-IB16040 and the APC was funded by the Consejería de Economía e Infraestructuras de la Junta de Extremadura (the branch of the regional government that covers economy and infrastructure) and by the European Regional Development Fund (ERDF).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Jackson, J. Developing regional tourism in China: The potential for activating business clusters in a socialist market economy. *Tour. Manag.* **2006**, *27*, 695–706. [[CrossRef](#)]
2. Rogerson, C.M. Regional tourism in South Africa: A case of ‘mass tourism of the South’. *GeoJournal* **2004**, *60*, 229–237. [[CrossRef](#)]
3. Braun, P. Regional tourism networks: The nexus between ict diffusion and change in Australia. *Inf. Technol. Tour.* **2003**, *6*, 231–243. [[CrossRef](#)]
4. Ateljevic, J. Tourism entrepreneurship and regional development: Example from New Zealand. *Int. J. Entrep. Behav. Res.* **2009**, *15*, 282–308. [[CrossRef](#)]
5. Frechtling, D.C. The tourism satellite account: Foundations, progress and issues. *Tour. Manag.* **1999**, *20*, 163–170. [[CrossRef](#)]
6. Smeral, E. Tourism satellite accounts: A critical assessment. *J. Travel Res.* **2006**, *45*, 92–98. [[CrossRef](#)]
7. Xu, C.; Jones, C.; Munday, M. Tourism inward investment and regional economic development effects: Perspectives from tourism satellite accounts. *Reg. Stud.* **2019**, 1–12. [[CrossRef](#)]
8. Perna, F.; Custódio, M.J.; Oliveira, V. Tourism destination competitiveness: An application model for the south of Portugal versus the Mediterranean region of Spain: Competitivtour. *Tour. Manag. Stud.* **2018**, *14*, 19–29. [[CrossRef](#)]

9. Sánchez-Rivero, M.; Martín, J.M.S.; Gallego, J.I.R. Methodological approach for assessing the potential of a rural tourism destination: An application in the province of Cáceres (Spain). *Curr. Issues Tour.* **2014**, *19*, 1084–1102. [\[CrossRef\]](#)
10. Ebejer, J. Urban heritage and cultural tourism development: A case study of Valletta's role in Malta's tourism. *J. Tour. Cult. Chang.* **2018**, *17*, 306–320. [\[CrossRef\]](#)
11. Li, J.; Xu, L.; Tang, L.; Wang, S.; Li, L. Big data in tourism research: A literature review. *Tour. Manag.* **2018**, *68*, 301–323. [\[CrossRef\]](#)
12. Xiang, Z.; Fesenmaier, D.R. Big data analytics, tourism design and smart tourism. In *Co-Creation and Well-Being in Tourism*; Springer Science and Business Media LLC: Cambridge, UK, 2016; pp. 299–307.
13. Song, H.; Liu, H. Predicting tourist demand using big data. In *Co-Creation and Well-Being in Tourism*; Springer Science and Business Media LLC: Cambridge, UK, 2016; pp. 13–29.
14. Yang, Y.; Pan, B.; Song, H. Predicting hotel demand using destination marketing organization's web traffic data. *J. Travel Res.* **2013**, *53*, 433–447. [\[CrossRef\]](#)
15. Miah, S.J.; Vu, H.Q.; Gammack, J.; McGrath, M. A big data analytics method for tourist behaviour analysis. *Inf. Manag.* **2017**, *54*, 771–785. [\[CrossRef\]](#)
16. Alaei, A.; Becken, S.; Stantic, B. Sentiment analysis in tourism: Capitalizing on big data. *J. Travel Res.* **2017**, *58*, 175–191. [\[CrossRef\]](#)
17. Gu, Z.; Zhang, Y.; Chen, Y.; Chang, X. Analysis of attraction features of tourism destinations in a mega-city based on check-in data mining—A case study of Shenzhen, China. *ISPRS Int. J. Geo Inf.* **2016**, *5*, 210. [\[CrossRef\]](#)
18. Xu, F.; Nash, N.; Whitmarsh, L. Big data or small data? A methodological review of sustainable tourism. *J. Sustain. Tour.* **2019**, *28*, 144–163. [\[CrossRef\]](#)
19. Yang, Y.; Zhang, H. Spatial-temporal forecasting of tourism demand. *Ann. Tour. Res.* **2019**, *75*, 106–119. [\[CrossRef\]](#)
20. Reid, L.J.; Andereck, K.L. Statistical analyses use in tourism research. *J. Travel Res.* **1989**, *28*, 21–24. [\[CrossRef\]](#)
21. Nunkoo, R. The routledge handbook of tourism impacts. In *The Routledge Handbook of Tourism Impacts*; Routledge: London, UK, 2019; p. 14.
22. Chatzigeorgiou, C.; Christou, E. Adoption of social media as distribution channels in tourism marketing: A qualitative analysis of consumers' experiences. *J. Tour. Herit. Serv. Mark.* **2020**, *6*, 25–32.
23. Dai, L.; Wang, S.; Xu, J.; Wan, L.; Wu, B. Qualitative analysis of residents' perceptions of tourism impacts on historic districts: A case study of Nanluoguxiang in Beijing, China. *J. Asian Arch. Build. Eng.* **2017**, *16*, 107–114. [\[CrossRef\]](#)
24. Mano, A.; Costa, R. Imagem projetada de Portugal como destino turístico: Análise qualitativa do portal oficial de promoção turística. *Rev. Tur. Desenvol.* **2018**, *29*, 65–83.
25. Martín, J.M.S.; Delgado, L.M.M.; Gallego, J.I.R. La reputación online de los alojamientos rurales en Extremadura desde una óptica geoestadística. *BAGE* **2019**, *82*, 1–36. [\[CrossRef\]](#)
26. Tarí, J.J.; Molina-Azorin, J.F.; Pereira-Moliner, J.; López-Gamero, M.D. The internalization of a sectorial standard for quality management: A qualitative analysis in tourism. In *Chinese Acquisitions in Developed Countries*; Springer: Cham, Switzerland, 2017; pp. 97–108.
27. Wang, X.; Sun, J.; Wen, H. Tourism seasonality, online user rating and hotel price: A quantitative approach based on the hedonic price model. *Int. J. Hosp. Manag.* **2019**, *79*, 140–147. [\[CrossRef\]](#)
28. Beerli, A.; Martín, J.D. Tourists' characteristics and the perceived image of tourist destinations: A quantitative analysis—A case study of Lanzarote, Spain. *Tour. Manag.* **2004**, *25*, 623–636. [\[CrossRef\]](#)
29. Walle, A.H. Quantitative versus qualitative tourism research. *Ann. Tour. Res.* **1997**, *24*, 524–536. [\[CrossRef\]](#)
30. Nunkoo, R. Handbook of research methods for tourism and hospitality management. In *Handbook of Research Methods for Tourism and Hospitality Management*; Edward Elgar Publishing: Cheltenham, UK, 2018; pp. 3–23.
31. Hall, C.M. Quantitative and qualitative content analysis. In *Handbook of Research Methods in Tourism and Hospitality*; Edward Elgar Publishing: Cheltenham, UK, 2018; pp. 395–406.
32. Dwyer, L.; Gill, A.; Seetaram, N. *Handbook of Research Methods in Tourism: Quantitative and Qualitative Approaches*; Edward Elgar Publishing: Cheltenham, UK, 2012; p. 528.
33. Nunkoo, R.; Ramkissoon, H. Structural equation modelling and regression analysis in tourism research. *Curr. Issues Tour.* **2012**, *15*, 777–802. [\[CrossRef\]](#)

34. Lv, Z.; Xu, T. A panel data quantile regression analysis of the impact of corruption on tourism. *Curr. Issues Tour.* **2016**, *20*, 1–14. [\[CrossRef\]](#)
35. Santeramo, F.G.; Morelli, M. Modelling tourism flows through gravity models: A quantile regression approach. *Curr. Issues Tour.* **2015**, *19*, 1–7. [\[CrossRef\]](#)
36. Ferreira, F.A.; Castro, C. The impact of terrorism and corruption on tourism in Turkey: A regression analysis. *J. Comput. Methods Sci. Eng.* **2019**, 1–8. [\[CrossRef\]](#)
37. Kim, N.; Song, H.; Pyun, J.H. The relationship among tourism, poverty, and economic development in developing countries. *Tour. Econ.* **2016**, *22*, 1174–1190. [\[CrossRef\]](#)
38. Dogru, T.; Sirakaya-Turk, E. Engines of tourism's growth: An examination of efficacy of shift-share regression analysis in South Carolina. *Tour. Manag.* **2017**, *58*, 205–214. [\[CrossRef\]](#)
39. Gutiérrez, J.; García-Palomares, J.C.; Romanillos, G.; Salas-Olmedo, M.H. The eruption of Airbnb in tourist cities: Comparing spatial patterns of hotels and peer-to-peer accommodation in Barcelona. *Tour. Manag.* **2017**, *62*, 278–291. [\[CrossRef\]](#)
40. Sarrión-Gavilán, M.D.; Márquez, M.D.B.; Mora-Rangel, E.O. Spatial distribution of tourism supply in Andalusia. *Tour. Manag. Perspect.* **2015**, *15*, 29–45. [\[CrossRef\]](#)
41. Martín, J.M.S.; Gallego, J.I.R.; Blas-Morato, R. Hot spot analysis versus cluster and outlier analysis: An enquiry into the grouping of rural accommodation in Extremadura (Spain). *ISPRS Int. J. Geo Inf.* **2019**, *8*, 176. [\[CrossRef\]](#)
42. Xu, F.; Hu, M.; La, L.; Wang, J.; Huang, C. The influence of neighbourhood environment on Airbnb: A geographically weighed regression analysis. *Tour. Geogr.* **2019**, *22*, 192–209. [\[CrossRef\]](#)
43. Baggio, R.; Klobas, J. *Quantitative Methods in Tourism*; Channel View Publications, Ltd.: Bristol, UK, 2017.
44. Rezaei, S. Quantitative methods, applications, and trends in Asian tourism research. In *Quantitative Tourism Research in Asia*; Springer: Singapore, 2018; pp. 1–10.
45. Exteberría, J. *Regresión Múltiple*; La Muralla: Madrid, Spain, 2007; p. 160.
46. Martín, J.M.S.; Rivero, M.S.; Gallego, J.I.R. Analysis of the balance between tourist potential and supply of rural tourism accommodation using spatial statistical techniques. An application to the province of Cáceres (Spain). *Cuad. Tur.* **2017**, *39*, 699–702.
47. Fotheringham, A.S.; Charlton, M.; Brunsdon, C. Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environ. Plan. A* **1998**, *30*, 1905–1927. [\[CrossRef\]](#)
48. Lu, B.; Charlton, M.; Harris, P.; Fotheringham, A.S. Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 660–681. [\[CrossRef\]](#)
49. Smith, S.L. Room for rooms: A procedure for the estimation of potential expansion of tourist accommodations. *J. Travel Res.* **1977**, *15*, 26–29. [\[CrossRef\]](#)
50. Yang, Y.; Luo, H.; Law, R. Theoretical, empirical, and operational models in hotel location research. *Int. J. Hosp. Manag.* **2014**, *36*, 209–220. [\[CrossRef\]](#)
51. Sánchez, J.M. El sistema de información geográfica como herramienta de análisis turístico. Una aplicación para la localización idónea de alojamientos rurales en la provincia de Cáceres mediante análisis multicriterio. *Rev. Estudios Turíst.* **2009**, *182*, 71–94.
52. Sánchez, M.; Sánchez, J.M. Nuevas técnicas de investigación aplicadas al turismo: La estadística espacial. In M. Figuerola & C. Martín. *Modelos De Gestión E Innovación En Turismo*; Civitas: Madrid, Spain, 2019; pp. 109–132.
53. Goodrich, J.N. Differences in perceived similarity of tourism regions: A spatial analysis. *J. Travel Res.* **1977**, *16*, 10–13. [\[CrossRef\]](#)
54. Wilson, J. The routledge handbook of tourism geographies. In *The Routledge Handbook of Tourism Geographies*; Routledge: London, UK, 2012; pp. 163–173.
55. Martín, J.M.S.; Rivero, M.S.; Gallego, J.I.R. La evaluación del potencial para el desarrollo del turismo rural. Aplicación metodológica sobre la provincia de Cáceres. *Geofocus* **2013**, *13*, 99–130.
56. Sánchez, J.M.; Pérez, M.N.; Jurado, J.C.; Granados, M.M. Detección de áreas óptimas para la implantación de alojamientos rurales en Extremadura. Una aplicación SIG. *Lurralde* **1999**, *22*, 367–384.
57. Martín, J.M.S.; Gallego, J.I.R.; Sánchez-Rivero, M. Caracterización espacial del turismo en extremadura mediante análisis de agrupamiento (grouping analysis). Un ensayo técnico. *GeoFocus* **2017**, *19*, 207–235. [\[CrossRef\]](#)

58. Getis, A.; Ord, J.K. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **2010**, *24*, 189–206. [CrossRef]
59. Ord, J.K.; Getis, A. Local spatial autocorrelation statistics: Distributional issues and an application. *Geogr. Anal.* **2010**, *27*, 286–306. [CrossRef]
60. Griffith, D.A. Teaching spatial autocorrelation by simulation. *J. Geogr. High. Educ.* **1987**, *11*, 143–153. [CrossRef]
61. Anselin, L. A local indicator of multivariate spatial association: Extending geary's c. *Geogr. Anal.* **2018**, *51*, 133–150. [CrossRef]
62. Fotheringham, A.S.; Brunsdon, C.; Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*; John Wiley & Sons: New York, NY, USA, 2002; p. 284.
63. Díaz, M.V. Tills y otros depósitos relacionados con la dinámica glaciar en la Sierra de Xistral: Interpretación de litofacies y reconstrucción paleoglaciaria. *BAGE* **2018**, *76*, 276–302. [CrossRef]
64. Gallego, J.I.R.; Martín, J.M.S. Análisis de la distribución territorial de los alojamientos rurales y convencionales en los núcleos rurales de Extremadura. *An. Geogr. Univ. Complut.* **2019**, *39*, 463–484. [CrossRef]
65. Sánchez, J.M.; Rengifo, J.I. Evolución del sector turístico en la Extremadura del siglo XXI: Auge, crisis y recuperación. *Lurralde* **2019**, *42*, 19–50.
66. Junta de Extremadura. Extremadura Turismo. Available online: <https://turismoextremadura.com/> (accessed on 12 January 2020).
67. DOE Núm. 74 Decreto 65/2015, de 14 de Abril, Por El Que Se Establece la Ordenación y Sistema de Clasificación de Los Alojamientos de Turismo Rural de la Comunidad Autónoma de Extremadura; DOE: Washington, DC, USA, 2015.
68. Instituto Nacional de Estadística (INE). Available online: <https://www.ine.es/daco/daco42/ocuptr/eotr0019.xlsx> (accessed on 15 January 2020).
69. Instituto Geográfico Nacional (IGN). Centro de Descargas. Available online: <http://centrodedescargas.cnig.es/CentroDescargas/> (accessed on 19 January 2020).
70. Instituto Geográfico Nacional (IGN). Base Topográfica Nacional 1:100.000 (BTN100). Available online: <http://www.ign.es/web/resources/docs/IGNCnig/CBG%20-%20BTN100.pdf> (accessed on 19 January 2020).
71. Santos, C.A.S.M.; Vieira, J.A.C. An analysis of visitors' expenditures in a tourist destination: OLS, quantile regression and instrumental variable estimators. *Tour. Econ.* **2012**, *18*, 555–576. [CrossRef]
72. Brida, J.G.; Scuderi, R. Determinants of tourist expenditure: A review of microeconomic models. *Tour. Manag. Perspect.* **2013**, *6*, 28–40. [CrossRef]
73. Wilcox, R.R. Robust regression: Testing global hypotheses about the slopes when there is multicollinearity or heteroscedasticity. *Br. J. Math. Stat. Psychol.* **2018**, *72*, 355–369. [CrossRef]
74. Wilcox, R.R. Multicollinearity and ridge regression: Results on type I errors, power and heteroscedasticity. *J. Appl. Stat.* **2018**, *46*, 946–957. [CrossRef]
75. Sánchez, J.M. La matriz de correlación lineal en climatología. Los riesgos interpretativos: Su reducción o eliminación. *Estudios Geogr.* **1995**, *219*, 411–434.
76. Meloun, M.; Militký, J. Detection of single influential points in OLS regression model building. *Anal. Chim. Acta* **2001**, *439*, 169–191. [CrossRef]
77. ESRI. Available online: <https://doc.arcgis.com/es/insights/analyze/regression-analysis.htm> (accessed on 19 February 2020).
78. Berry, W. *Understanding Regression Assumptions*; SAGE Publications: Ventura, CA, USA, 1993.
79. Tobler, W.R. A computer movie simulating urban growth in the detroit region. *Econ. Geogr.* **1970**, *46*, 234. [CrossRef]
80. Anselin, L. Local indicators of spatial association-lisa. *Geogr. Anal.* **2010**, *27*, 93–115. [CrossRef]
81. Tu, J.; Xia, Z.-G. Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation. *Sci. Total Environ.* **2008**, *407*, 358–378. [CrossRef] [PubMed]
82. De Toro, P.; Nocca, F.; Renna, A.; Sepe, L. Real estate market dynamics in the city of Naples: An integration of a multi-criteria decision analysis and geographical information system. *Sustainability* **2020**, *12*, 1211. [CrossRef]
83. Sánchez, J.M. *Propuesta Metodológica Para La Generación De Información Climática En La Provincia De Cáceres*; Fundicot-Extremadura Universidad de Extremadura: Cáceres, Spain, 1995.

84. Sánchez, M. Memoria Turística de Extremadura Por Territorios. Año 2018. Available online: https://www.turismoextremadura.com/.content/observatorio/2019/EstudiosYMemoriasAnuales/Memoria_turistica_por_territorios_2018.pdf (accessed on 12 December 2019).
85. Thrane, C. Students' summer tourism: Determinants of length of stay (LOS). *Tour. Manag.* **2016**, *54*, 178–184. [[CrossRef](#)]
86. Wang, Y.; Davidson, M.C. A review of micro-analyses of tourist expenditure. *Curr. Issues Tour.* **2010**, *13*, 507–524. [[CrossRef](#)]
87. Wu, L.; Zhang, J.; Fujiwara, A. Tourism participation and expenditure behaviour: Analysis using a scobit based discrete–continuous choice model. *Ann. Tour. Res.* **2013**, *40*, 1–17. [[CrossRef](#)]
88. Zheng, B.; Zhang, Y. Household expenditures for leisure tourism in the USA, 1996 and 2006. *Int. J. Tour. Res.* **2011**, *15*, 197–208. [[CrossRef](#)]
89. Jin, C.; Xu, J.; Huang, Z. Spatiotemporal analysis of regional tourism development: A semiparametric geographically weighted regression model approach. *Habitat Int.* **2019**, *87*, 1–10. [[CrossRef](#)]
90. Ahmadi, M.; Baaghde, M.; Roudbari, A.D.; Asadi, M. Modeling the role of topography on the potential of tourism climate in Iran. *Model. Earth Syst. Environ.* **2018**, *4*, 13–25. [[CrossRef](#)]
91. Sánchez-Martín, J.M.; Blas-Morato, R.; Rengifo-Gallego, J.I. The dehesas of extremadura, spain: A potential for socio-economic development based on agritourism activities. *Forests* **2019**, *10*, 620. [[CrossRef](#)]
92. Martín, J.M.S.; Sánchez-Rivero, M.; Gallego, J.I.R. Water as a tourist resource in extremadura: Assessment of its attraction capacity and approximation to the tourist profile. *Sustainability* **2020**, *12*, 1659. [[CrossRef](#)]
93. Martín, J.M.S.; Gallego, J.I.R.; Sánchez-Rivero, M. Protected areas as a center of attraction for visits from world heritage cities: Extremadura (Spain). *Land* **2020**, *9*, 47. [[CrossRef](#)]
94. Martín, J.M.S.; Gurriá-Gascón, J.-L.; García-Berzosa, M.-J. The cultural heritage and the shaping of tourist itineraries in rural areas: The case of historical ensembles of Extremadura, Spain. *ISPRS Int. J. Geo Inf.* **2020**, *9*, 200. [[CrossRef](#)]
95. Delgado, L.M.M.; Gallego, J.I.R.; Martín, J.M.S. Hunting tourism as a possible development tool in protected areas of Extremadura, Spain. *Land* **2020**, *9*, 86. [[CrossRef](#)]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).