

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)

José Manuel Sánchez Martín

Marcelino Sánchez Rivero

Juan Ignacio Rengifo Gallego

Universidad de Extremadura

jmsanche@unex.es, sanriver@unex.es, irengifo@unex.es

Rural lodgings are often located in natural environments but, contrary to what would be expected, they show a very low occupancy rate. This circumstance is explained not only by the adverse economic situation, but also by over-supply, an inconvenient location, unfavorable weather conditions, insufficient or inadequate promotion, a mismatch between the tourist potential of the territory and the supply of accommodation, etc.

In the latter case, one would expect some positive relationship between the tourism potential of a rural tourist destination and its accommodation supply. However, this is not always the case, given that, sometimes, there is a high concentration of accommodation places in areas with low tourist potential while there are other areas with great tourist potential but insufficient accommodation.

In order to analyze these mismatches, it is not possible to use conventional statistical models, because they do not usually consider the spatial dimension of tourism phenomenon. This paper proposes the use of statistical models based on spatial parameterization, which have been rarely used in the scientific literature. The incorporation of this spatial component is fully justified by the fact that tourism, like most economic activities, has a strong spatial nature as both supply and tourist attractions are distributed over a territory that shapes the tourist space. The spatial distribution of certain components that make up the tourist potential follows, in most of the cases, a particular pattern.

In fact, one of the major current problems in the conventional analysis of tourism is that it does not consider the full extent of its spatial character. A tourist destination cannot be analyzed in isolation, but rather it must be examined taking into consideration the influence that neighboring destinations can exert on it, or vice versa.

In order to prove the usefulness of the spatial models proposed in this paper, the supply of rural accommodation in the province of Cáceres (Spain) will be analyzed in combination with the tourism potential of its municipalities. The province of Cáceres, which in the mid-90s was a peripheral frontier territory, away from the main tourism centers at national level, began to slowly and gradually integrate into the tourist market, with the exception of some northern mountain regions which showed a more accelerated process.

Moreover, the recent development of rural tourism in the province has led to an almost spontaneous implementation of accommodation supply in many municipalities. This situation has caused an increase in the supply of rural accommodation that affects many population centers, even though they did not have the attractions mainly demanded by rural tourists, which reveals a mismatch between the tourist potential and the accommodation supply of this territory. Furthermore, the rapid and poorly planned growth of rural tourism in this province favors the lack of implementation of an appropriate tourism policy aimed at overcoming the absence of key attractions through the generation of tourist products that depend on other resources that can be found in this territory. Another mistake in tourism planning has been to base the development of rural tourism exclusively on the creation of new rural accommodation units, without taking into account the adequacy between tourist potential and the current accommodation supply or the generation of differentiated tourist products. Thus, some mismatches have occurred. On the one hand, there is rural accommodation supply in areas with a limited tourist potential with respect to the main attractions. On the other hand, it is observed that accommodation supply continues to increase in those areas with greatest tourist potential, although occupancy rates are not the best to ensure the optimal performance of the tourist activity. This situation may endanger the economic viability of some rural lodgings in the province.

From a methodological point of view, this paper proposes the use of a geographically weighted regression (GWR) model, which is very innovative in the territorial (spatial) analysis of tourism. This type of regression model estimates a linear model whose parameters take into account the spatial relationships among the units analyzed. This fact represents one of the main strengths of the GWR model and differentiates it from other types of multiple regressions.

Specifically, the GWR model provides a local model for the dependent variable based on the independent ones. That is, it calculates an equation for each entity being analyzed, unlike what happens with other regressions, in which a global model is obtained. This property represents a significant advance over other types of regressions, making it possible to go from a global overview to a local analysis of the problem, thus providing a geographical analysis in greater detail.

For the empirical analysis presented in this paper, different exploratory regressions were considered in order to obtain a suitable model and compare it to other model proposed by the authors. In the original configuration of the model, 19 parameters were considered to try to explain, in the best possible way, the distribution of rural accommodation places in the municipalities of the province of Cáceres. In order to make a choice among the different regression models finally selected, criteria such as the highest goodness of fit, a low inflation variance factor and the existence of random residuals were taken into account.

As an alternative to the four conventional regression models finally selected, it is proposed an extended model based on empirical knowledge of the tourism system and

tourists' opinion (REC model). This extended model shows a slightly lower explanation of variance, but it is considered to be a much more realistic model.

Analysis of the results obtained with the proposed regression model indicates that there are some well-defined attractors, which act as genuine regressors when the presence of rural accommodation units is explained (altimetry, biodiversity). Moreover, there are other independent variables that could be considered to be a consequence of these tourist activities (activity businesses, information offices and visitor centers), which could also serve to explain the volume of available accommodation units or the level of development achieved.

The model based on the empirical knowledge of rural tourism in the area analyzed and on the opinion of the demand (REC model) is a robust model that will allow us to apply a geographically weighted regression (GWR) model. The lower explanation of the total variance in this model is justified by the fact that accommodation units have not always been built in the most appropriate places, that is, the most advisable regarding its tourist potential. On the contrary, there are other establishments located in places that do not meet demand's requirements. This situation leads to the appearance of certain mismatches. At the same time, the model confirmed that there are places with interesting attractors to practice rural tourism which, however, lack accommodation or have an insufficient number of accommodation places. In addition, activity businesses, tourist information and visitor centers are of fundamental importance in the REC model and represent aspects that provide services to rural tourists.

Before estimating the GWR model, it is necessary to calculate Moran's I spatial autocorrelation statistic from residuals. This statistic aims to test the spatial randomness of residuals and thus avoid the omission of some important predictor variables with a territorial explanation. Another objective of I statistic is to ensure the applicability of a spatial regression to the available data. To do this, it should be determined whether the dependent variable (rural accommodation places) is aggregated, regular or completely randomly distributed. After calculation of this I statistic, it has been possible to confirm the spatial randomness of residuals.

Results of GWR model estimation reveal a R2 value of 72% and a R2 adjusted value of 71%. Therefore, it is concluded that the estimated model is robust for the available data and has a remarkable contribution to the explanation of the variance of rural accommodation places.

Another interesting result is the presence of certain mismatches between estimated values and observed values of rural accommodation places. The two possible causes of these mismatches are the following:

- a) The exclusion of explanatory variables, although it should be noted that some independent variables were discarded due to collinearity problems. Even including these variables in the model, the goodness of fit did not increase in a significant way.
- b) The model detects that accommodation places do not fit the real tourist potential of the territory. This fact is especially significant considering that there are population centers where the concentration of places has been very high (more than 400 places). Meanwhile, other centers, with a high tourist potential, have a low supply of rural accommodation places. This situation results in a loss of explanation in the estimated model.

In order to establish the alignment between the observed volume of places in rural accommodations and the values that should correspond in terms of their attractiveness, this paper proposes a classification of municipalities in the province of Cáceres into three different groups:

- a) Population centers with an adjusted supply of rural accommodation places (residuals less than ± 25 and with presence of lodgings). 138 centers grouped into 46 municipalities are included in this group. In this first group, it is possible to identify two different types of population centers. On the one hand, centers without rural lodgings (71 centers) and where the predicted value exceeds 20 places compared with the observed value. These centers do not have a high tourist potential, mainly because most of them could be placed in the “low” category of tourist potential. On the other hand, there are 67 centers that have rural lodgings and where it is usual that the predicted values are higher than the observed ones (but the difference is less than 25).
- b) Population centers with a saturated supply of rural accommodation places (residuals above 25). This over-supply occurs when the estimated values are lower than the actual (observed) places. This second group comprises 38 population centers grouped into 36 municipalities. These centers have the advantage of being entities where the rural activity is fully established. They are located in the northern part of the province, the most interesting one for practicing rural tourism.
- c) Population centers with an insufficient supply of rural accommodation places (with no places and with predicted values higher than 50 places). Population centers where the model estimate a number of places higher than the observed (or actual) number are included in this final group. The group is made up of 100 population centers grouped into 45 municipalities. It is in this third group where growth of lodgings supply has more possibilities of success. This growth could continue until the actual supply is equal to the estimated supply according to its tourist potential.

Finally, the main conclusions in this paper, both from a technical and from an applied (empirical) point of view, can be listed as follows.

First of all, it is necessary to previously debug any regression model in order to avoid that the variables included show strong correlations among them, although the elimination of collinearity could imply a loss in the ability to explain the variance of the dependent variable.

Secondly, application of a test of spatial autocorrelation through Moran's I statistic on residuals obtained by ordinary least squares regression is necessary. The aim of this test is to prove that mismatches between predicted and observed values are randomly distributed and that they are not concentrated on the territory.

The third conclusion is that the results obtained in a geographically weighted regression can be used to test the adequacy of the model to each of the geographical entities analyzed, enabling thus local analyses and surpassing the simple overall analysis of conventional regression models.

Finally, the last conclusion is that, from the comparison between the observed values and the predicted values by the model, it is possible to define different groups of population centers. It is also possible to identify which of these groups have greater possibilities of growth in the rural accommodation places supply, when considering the tourist potential of each population center.