Article

A Spatial Approach for Modeling Amphibian Road-Kills: Comparison of Regression Techniques

Diana Sousa-Guedes [1,*], **Marc Franch** [1,2] **and Neftalí Sillero** [1]

1 Centro de Investigação em Ciências Geo-Espaciais (CICGE), Faculdade de Ciências da Universidade do Porto, Alameda do Monte da Virgem, 4430-146 Vila Nova de Gaia, Portugal; apoarmatu@gmail.com (M.F.); neftali.sillero@gmail.com (N.S.)
2 PECAT Research Group, Departament de Ciències Ambientals, Universitat de Girona, 17003 Girona, Spain
* Correspondence: dianasguedes@gmail.com

Abstract: Road networks are the main source of mortality for many species. Amphibians, which are in global decline, are the most road-killed fauna group, due to their activity patterns and preferred habitats. Many different methodologies have been applied in modeling the relationship between environment and road-kills events, such as logistic regression. Here, we compared the performance of five regression techniques to relate amphibians' road-kill frequency to environmental variables. For this, we surveyed three country roads in northern Portugal in search of road-killed amphibians. To explain the presence of road-kills, we selected a set of environmental variables important for the presence of amphibians and the occurrence of road-kills. We compared the performances of five modeling techniques: (i) generalized linear models, (ii) generalized additive models, (iii) random forest, (iv) boosted regression trees, and (v) geographically weighted regression. The boosted regression trees and geographically weighted regression techniques performed the best, with a percentage of deviance explained between 61.8% and 76.6% and between 55.3% and 66.7%, respectively. Moreover, the geographically weighted regression showed a great advantage over the other techniques, as it allows mapping local parameter coefficients as well as local model performance (pseudo-R^2). The results suggest that geographically weighted regression is a useful tool for road-kill modeling, as well as to better visualize and map the spatial variability of the models.

Keywords: ecological niche modeling; geographically weighted regression; logistic regression; road ecology; spatial regression; wildlife-vehicle collision

1. Introduction

Road networks fragment landscapes globally [1,2]. These linear infrastructures affect wildlife by promoting barrier effects, changes in distribution patterns or direct mortality [3–5]. Amphibians are the most road-killed fauna group, due to their activity patterns and preferred habitats [6–9]. Beyond that, amphibians do not perceive the risks associated with roads [10], with some species being more affected than others, such as the ones with higher vagility [6] or with the tendency to paralyze when facing a moving vehicle [11]. At a global level, amphibians are in critical decline [12], and road density and traffic expansion are major facilitators [13,14]. For instance, in a study in Southern Portugal, amphibians constituted up to 70% of the road-killed vertebrates found during the rainy season [15]. To mitigate road mortality, it is crucial to understand the factors influencing the occurrence and spatial location of road-kills, which are usually clustered in time and space, not occurring randomly [16]. Animal road-kills levels depend on many different factors, namely road characteristics, weather conditions, surrounding vegetation, proximity to water points, traffic, and species ecological traits [7,9,16–22]. Species abundance is also highly important for the determination of road-kill patterns [23,24], yet this type of data is very difficult to obtain, namely in amphibians [7,9].

Road-kills are frequently very difficult to model, due to the excessive number of zeros and the clustered distribution of mortality events [25]. Many different methodologies have been applied to model the relationship between environmental features and the number of road-kills, such as logistic regression and ecological niche modeling [9,16,22,26,27]. These analyses often involve non-spatial modeling methods such as generalized linear models (e.g., [28]). However, important variables frequently have spatial correlation, exhibiting non-stationarity along the road [29,30]. Therefore, modeling methods should include spatial relationships within the study area: indeed, most conventional methods do not include the spatial location of road-kills [29]. These non-spatial methods assume that each variable has one fixed coefficient for the entire geographical area [31] and generally only produce average and global parameter estimates [32]. Thus, spatial regression methods such as geographically weighted regression (GWR) are a good alternative. This technique may resolve issues with spatial heterogeneity and non-stationarity in modeling [29,33]. GWR has other advantages over other non-spatial modeling methodologies [31,32,34]: allows different relationships between the road-kills and environmental variables to vary within the study area, and provides local parameter estimates in addition to global estimates [33]. This method has been proved effective to forecast human accident impacts [35,36] and is increasingly being applied in ecological studies [29,34,37–40], but little in road ecology.

In this study we compared the performances of five regression techniques in modeling amphibian road-kills: (i) generalized linear models (GLM), (ii) generalized additive models (GAM), (iii) random forest (RF), (iv) boosted regression trees (BRT), and (v) geographically weighted regression (GWR). Specifically, the objectives of this study were: (1) to address the relationship between amphibian road-kills and environmental variables; (2) to evaluate the performance of the five techniques using model predicted values and residuals, and (3) to spatially visualize the performance of the models in terms of the distribution of positive and negative residuals. We highlighted the main steps of the proposed research in Figure 1.
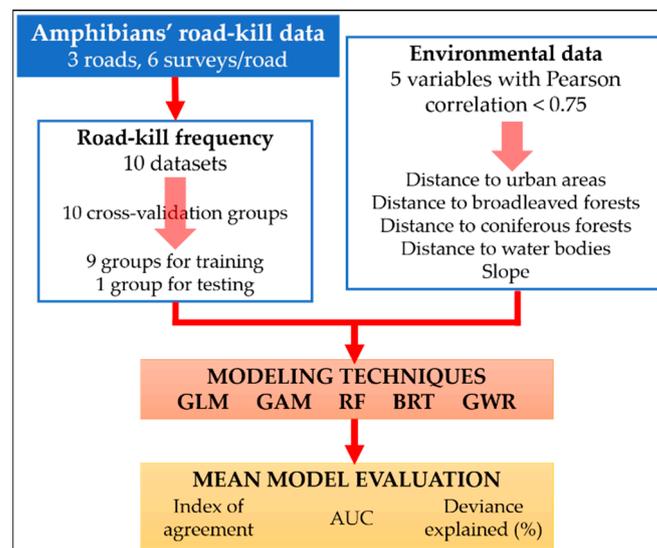

Figure 1. Main steps of the research process for comparing the performance of five regression techniques in modeling amphibian road-kills. GLM: generalized linear models; GAM: generalized additive models; RF: random forest; BRT: boosted regression trees; GWR: geographically weighted regression; AUC: area under the curve.

There is still no scientific consensus on what modeling method may better support conservation actions, namely, to prioritize road segments for the establishment of mitigation measures or to organize more intensive monitoring on those segments [26]. Our study shows an alternative technique for modeling road-kills, as well as for exploring the data along the roads. To our knowledge, this is the first study aiming to use spatial regression analysis to model the frequency of animal road-kills.

2. Materials and Methods

2.1. Road-Kills' Data

We collected data on amphibian road-kills in northern Portugal (Braga district, Figure 2), between February and June 2014. This region has a high diversity of amphibians [41,42]. However, it is a very populated region with a large and dense road network. The landscape is dominated by eucalyptus and pine forests, agricultural lands, and small natural/broadleaved forests. We selected three two-way country roads, with 27 to 40 km in length and moderate traffic. These types of country roads host the highest levels of amphibian mortality [43].
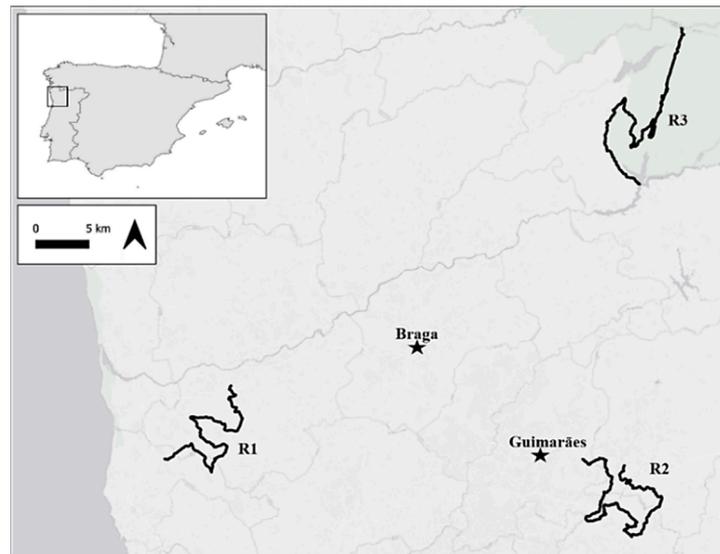

Figure 2. Study area with the three surveyed roads (R1, R2 and R3) in northern Portugal.

We conducted six surveys by car per road (totaling 18) in search of road-killed amphibians, at a maximum speed of 20 km/h [7,8,44]. We sampled early in the morning after rainy nights with temperatures not lower than 6–8 °C [7,8,45]. We recorded amphibians' positions with a 2–3 m accuracy GPS receiver, removing all carcasses from the road afterwards to avoid attracting scavengers [46]. For each road, we calculated the abundance kilometer index (AKI), which expresses the ratio of the total number of road-killed amphibians observed in a road by the total length of the road [47].

2.2. Environmental Data

We divided each surveyed road into sections of 250 m based on species' dispersal data [45,48,49]. We assigned 11 environmental variables to each road segment (Table 1). Variables were selected considering the habitat preferences of amphibians that facilitate the occurrence of road-kills [6–9].

We tested for correlation and excluded variables with a Pearson correlation higher than 0.75. When pairs of variables were correlated, we chose the variable more ecologically meaningful [16]. We selected five environmental variables: distance to urban areas, distance to broadleaved forests, distance to coniferous forests, distance to water bodies, and slope. Environmental variables were prepared using QGIS 2.14.1.

Table 1. Description, units, and source of the eleven environmental variables.

Variable	Description	Unit	Source
Distance to urban areas	Linear distance to the closest urban and artificial surface	m	Corine Land Cover (v2012)
Distance to irrigated land	Linear distance to the closest permanently irrigated land	m	Corine Land Cover (v2012)
Distance to agricultural areas	Linear distance to the closest agricultural areas, mainly composed of vineyards, pastures, annual crops, and complex cultivation patterns	m	Corine Land Cover (v2012)
Distance to broadleaved forests	Linear distance to the closest broad-leaved forest, mainly composed of trees, some shrubs, and bushes	m	Corine Land Cover (v2012)
Distance to coniferous forests	Linear distance to the closest coniferous forest, mainly composed of trees, shrubs, and bushes	m	Corine Land Cover (v2012)
Distance to mixed forests	Linear distance to the closest mixed forest	m	Corine Land Cover (v2012)
Distance to shrubs	Linear distance to scrub or herbaceous vegetation associations, mainly composed of natural grasslands, moors and heathland, and transitional woodland shrub	m	Corine Land Cover (v2012)
Distance to open areas	Linear distance to open spaces with little or no vegetation, composed of bare rocks, sparsely vegetated areas, and burnt areas	m	Corine Land Cover (v2012)
Distance to water bodies	Linear distance to the closest water body	m	Corine Land Cover (v2012)
Fires	Frequency of forest fires between 1975 and 2013	-	ICNF, Territórios ardidos 1975–2013
Slope	Slope range (*maximum–minimum*)	-	Sistema Nacional de Informação Geográfica—DGT

2.3. Modeling Techniques

We counted the number of amphibian road-kills per road section to create a database of road-kill frequency. We performed a 10-fold cross-validation by generating ten different datasets in order to avoid an excessive number of absences. In each dataset, we divided the segments with or without presences into ten groups: nine groups were used to train the model and one for testing it. In this way, we obtained ten different datasets with nine training sections without road-kills and nine with road-kills, and two test groups, one without road-kills and one with road-kills. For the absences, we randomly selected different road segments without presences for each dataset [22], limiting the number of absence segments to the same number of presences. This methodology is useful to avoid the zero-inflation problem [50]. The ten datasets included road segments of all the three surveyed roads. Results were the average of the ten model replicates.

We used five presence-absence regression techniques: generalized linear model (GLM), generalized additive model (GAM), random forest (RF), boosted regression tree (BRT), and geographically weighted regression (GWR). The first four techniques are non-spatial regression analyses: GLM, a parametric generalization of ordinary linear regression allowing error distributions other than a normal distribution [51]; GAM, a more flexible and non-parametric method where the linear predictor depends on smooth functions of predictor variables [51]; RF, an ensemble machine learning method for classification and regression that operates by constructing decision trees [52]; and BRT, a machine learning that can model functions and interactions between variables without making assumptions about the shape of fitted functions and is immune to the effects of outliers and inclusion of irrelevant variables [53]. Both GLM and GAM are extensions of linear models: GLM can be used when the features do not follow normal distributions; GAM can be used when the relationships between variables are not linear [51]. RF consists of multiple individual decision trees: it splits the data multiple times by bootstrap according to a threshold, creating different subsets of the dataset [52]. Both RF and BRT fit many decision trees to improve model accuracy, but in RF each occurrence has an equal probability of being

selected in subsequent samples while in BRT the input data are weighted in subsequent trees [52,53]. Thus, BRT combines decision trees with boosting methods, with the previously poorly modeled data having a higher probability of being selected in the new tree [53]. Non-spatial regression analyses assume that each explanatory response is constant across geographic space (i.e., there is one regression coefficient for each variable, for the entire study area), exhibiting spatial stationarity [34]. However, environmental features may vary across geographic space (spatial non-stationarity) which may lead to inaccurate model predictions [54]. This spatial heterogeneity often results from the interaction between different environmental factors [54]. On the other hand, spatial analysis techniques such as the GWR allow the spatial variation of the relationships between independent and dependent variables [33]. The GWR is a spatial local form of linear regression, allowing the measurement and mapping of local models. This method fits an ordinary least-squares regression around a regression point, in which the data closer to this point is weighted more heavily [33,55]. In all methods, we used the road-kill frequencies as the response variable and the five environmental variables as explanatory variables, with a Gaussian distribution for the non-spatial analyses, when applicable.

For GAMs, we set a maximum of four splines (k = 4), in order to limit the complexity of the smoothers describing the explanatory variables and to prevent over-fitted models. For BRT models, we set the parameters based on the small sample size: learning rate (lr) as 0.001, tree complexity (tc) as 3, and bag fraction as 0.5 [53]. The learning rate determines the contribution of each tree to the growing model, the tree complexity determines the complexity of variables interactions, and the bag fraction controls the stochasticity [53]. We optimized the parameters in order to reach a minimum of 1000 trees for each model [53,56]. For GWR, a different bandwidth for each target location is calculated. A moving window with a predetermined bandwidth and a spatial weighting kernel function allows the estimation of the intercept and coefficients for the explanatory variables of each road segment. Each observation within the local window is weighted based on its proximity to the center of that window [34]. We determined the bandwidth using the adaptive golden section search method [33] based on the Akaike information criterion (AICc, with a correction for finite sample sizes). We selected the Poisson model type and the Adaptive Bisquare spatial kernel type, which is suitable for unevenly distributed points [57].

To assess model performance, we used three cross-validation criteria to compare fitted with observed values [58]: the index of agreement (a refined Willmott's index of agreement [59]), the area under the receiver operator curve (AUC), and the percentage of deviance explained. Index of agreement is an error estimator that varies from 0 to 1 (1 means a perfect prediction) [59]. The AUC measures the ability of each model to correctly discriminate between presences and absences of road-kills, ranging between 0 and 1. The AUC is equal to the probability that a randomly chosen presence occurrence will be ranked higher than a randomly chosen absence [60]. The percentage of deviance explained was calculated as follows: (*null deviance − residual deviance*)/*null deviance*. We also compared parameter estimates and significances of the average model in each technique [29].

The spatial distribution of mean model residuals was assessed and displayed for the two techniques with the best model performance. We measured residuals as the difference between observed and fitted values $(y - \hat{y})$. We standardized the residuals on a scale of −1 to 1, in order to compare the values between techniques. Regarding the GWR results, we also mapped the local pseudo-R^2, indicating model performance in specific locations, and the significant parameter coefficients of the environmental variables [29]. In order to map only the significant estimates, we applied a significant threshold of 95% to mask out points where the relationship between road-kills and environmental variables were not significant (coefficient points with p-value > 0.05 were excluded) [61].

The GWR analysis was implemented in both GWR4.0 and MGWR 2.2 [57]. All the other non-spatial modeling techniques and statistical analyses were carried out using R 3.4.4 [62] using the packages mgcv, randomForest and dismo. Maps were displayed with QGIS 2.14.1.

3. Results

We recorded a total of 343 road-killed amphibians in the three study roads (Table 2). The abundance kilometer index (AKI) for each road and survey ranged between 0.00 and 2.52 amphibians/km, with a mean AKI for all surveys and roads of 0.56 amphibians/km.

Table 2. Selected road locations, number of road segments (and total segments length), number of road segments with road-kills, number of road-killed amphibians, and mean computed abundance kilometer index (AKI; number of road-killed amphibians per kilometer).

ROAD	Location	Road Segments	Segments with Road-Kills	Road-Kills	Mean AKI (ind/km)
R1	Barcelos	111 (27.70 km)	17	89	0.54
R2	Guimarães	143 (35.70 km)	42	171	0.80
R3	Gerês	157 (39.10 km)	24	83	0.35
TOTAL =		411 (102.50 km)	83	343	MEAN = 0.56

According to the selected criteria for evaluating model performance, the two regression techniques performing worse were the generalized linear models (GLM) and random forest (RF), followed by the generalized additive models (GAM) (Table 3). The boosted regression trees (BRT) and geographically weighted regression (GWR) performed the best and close to each other. The GLM models explained between 21.0% and 27.0% of deviance while the GAMs had a deviance explained between 41.7% and 53.1%. The BRT models performed well overall, with a deviance explained between 61.8% and 76.6%. Models obtained with the GWR technique had also good performance, with the analysis explaining between 55.3% and 66.7% of deviance.

Table 3. Mean model performances using three criteria: mean index of agreement (Wilmott's D test), mean AUC and mean percentage of deviance explained calculated for each of the five techniques (GLM, GAM, RF, BRT, GWR). Between brackets, the lowest and highest value obtained in the ten models with each method. In bold, the two best-classified methods in each criterion.

Methods	Index of Agreement	AUC	Deviance Explained (%)
GLM	0.567 (0.554–0.586)	0.812 (0.787–0.831)	22.9 (21.0–27.0)
GAM	0.635 (0.618–0.655)	0.843 (0.822–0.871)	48.4 (41.7–53.1)
RF	0.593 (0.577–0.610)	0.817 (0.794–0.840)	-
BRT	**0.755** (0.699–0.792)	**0.933** (0.900–0.961)	**70.8** (61.8–76.6)
GWR	**0.700** (0.675–0.719)	**0.881** (0.860–0.902)	**61.9** (55.3–66.7)

In GLMs, all variables were highly significant (Supplementary Materials, Table S1). In GAMs, almost all variables were also highly significant, and the response curves were slightly negative for all variables (Supplementary Materials, Table S2 and Figure S1). In RF and BRT, the most important variable was the slope (in nine models with both methods) (Supplementary Materials, Tables S3 and S4). In BRT models the response curves were not clear, so we cannot state whether there were positive or negative relations. In GWR, the most significant variable was the distance to coniferous forests (significant negative relation in nine average models) (Supplementary Materials, Table S5). Moreover, in GWR, all variables had a significant local coefficient in more than one road segment.

Regarding the spatial distribution of mean residuals, the GWR showed overall better spatial patterns over the BRT (with higher correct predictions, (i.e., residuals very close to zero) (Figure 3). The BRT created 46.3% of negative residuals, 19.1% of correct predictions and 37.0% of positive residuals; the GWR created 43.2% of negative residuals, 26.5% of correct predictions and 32.7% of positive residuals.
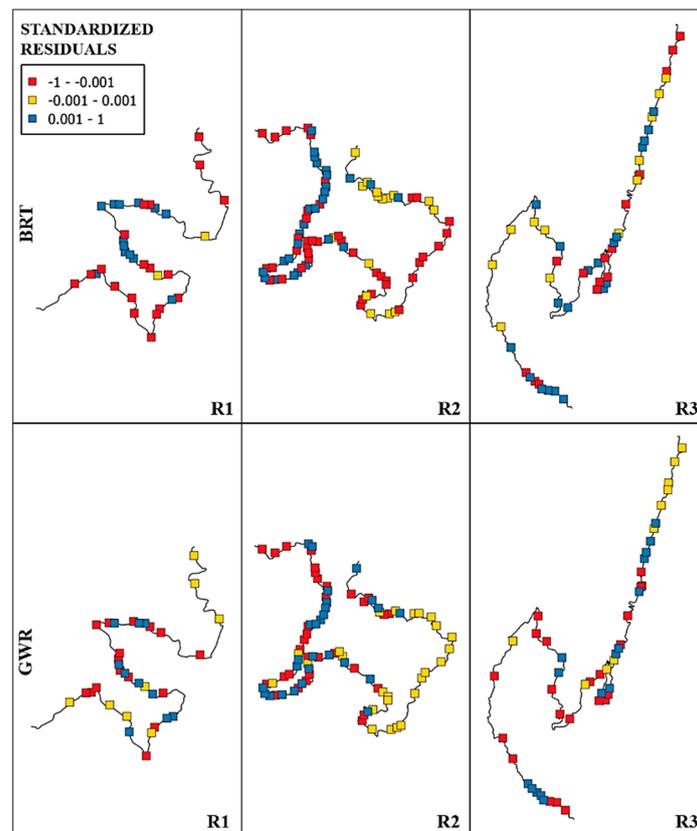
Figure 3. Spatial distribution of the mean model standardized residuals in each surveyed road of the two techniques with the best performance: BRT and GWR. Red squares are the locations with negative residuals, yellow squares are the locations with correct predictions (residuals close to zero) and blue squares the locations with positive residuals.

GWR performance varied over the study area, as indicated by the map of local pseudo-R^2, with a mean value of 0.58. (Figure 4). The roads R1 and R2 presented more constant pseudo-R^2 values along the roads, while the road R3 had greater variability between road segments.

The spatial distribution of significant coefficients of GWR models showed that the relationship between environmental variables and road-kills also changed with the location. Despite the distance to urban areas and distance to broadleaved forests of the average model being almost always non-significant, in some locations of the study roads, these variables significantly influenced the models. For instance, the variable distance to urban areas had a significant positive influence on roads R1 and R2 and a significant negative influence on road R3 (Figure 5). The opposite happens with the variable distance to broadleaved forests, which demonstrates a significant negative relationship in roads R1 and R2 and significant positive relation in roads R3 (Figure 5).
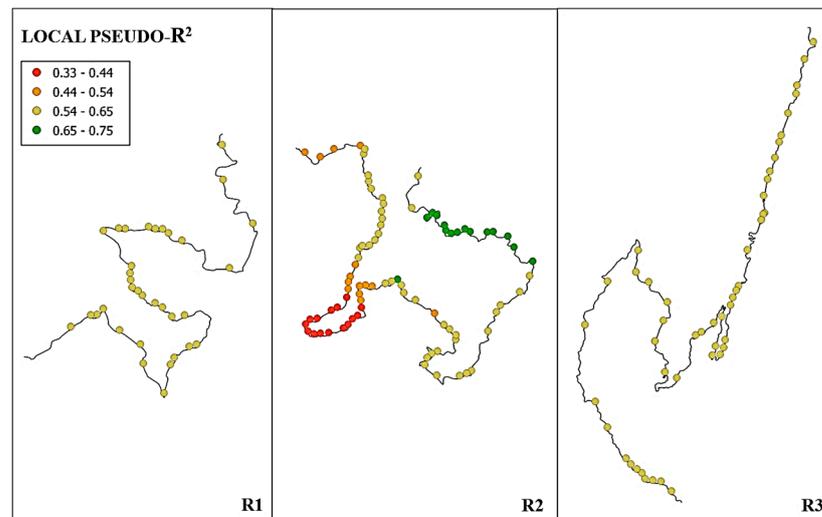
Figure 4. Local pseudo-R^2 of the GWR analysis in each surveyed road. In a gradient from red to green, the green points represent higher values of the local pseudo-R^2.
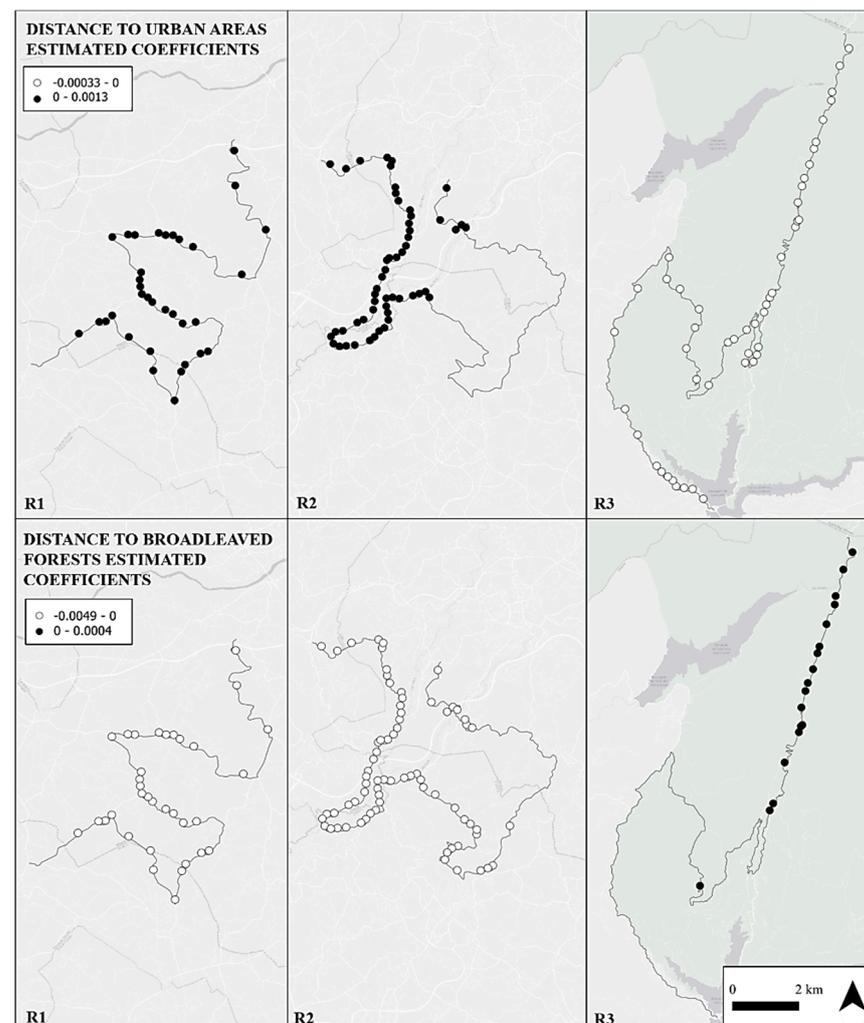

Figure 5. Estimated significant coefficients ($p < 0.05$) of two variables obtained with the GWR technique in the three roads: distance to urban areas (at the **top**) and distance to broadleaved forests (at the **bottom**). White points represent the locations with negative coefficients and black points represent the locations with positive coefficients.

4. Discussion

Our results highlight the difficulty of modelling the frequency of road-kills, with overall great variability between models and modeling methods. The BRT and GWR techniques showed better model performance and predictive power in comparison to the other tested techniques. There is no single technique that will perform best in all circumstances [63]: each method has its advantages that are useful in specific cases. In the case of modeling road-kills, we believe that both BRT and GWR are good approaches and should be tested. When choosing a non-spatial regression technique (such as BRT), it is of great benefit to explore and visualize the data using GWR. BRT provides a useful table of the relative influence of each variable, but not for the coefficients and significance of variables. Instead, this technique provides a response curve of each environmental variable, which is in our case of difficult interpretation. The BRT performed better in the three performance criteria used, yet the GWR provided a higher percentage of correct predictions (26.5% of residuals very close to zero). Both techniques created higher percentages of negative residuals in comparison to positive residuals, meaning that models overestimated more often than underestimated the number of road-kills.

All techniques provided similar results between each model replica, with small performance variability (e.g., deviance explained of BRT between 61.8% and 76.6% or index of agreement of GLM between 0.554 and 0.586) which could mean that the obtained results are trustworthy. Regarding the two techniques that performed poorest, the RF has some advantages over GLM, such as higher adaptability to evolve with the data [64], but lacks several important information, as it does not provide a measure of deviance. RF (as well as BRT) can be used as modeling or exploratory tools for searching the most important variables. Similar to BRT, GAM provides a response curve of each variable included in the model, and also the significance of each variable, but does not provide a coefficient. Most of the response curves of our GAM models were slightly negative but almost flat, which hinder the visualization of a clear relationship between road-kills and environmental variables. Moreover, GAM models are often over-fitted, due to the introduction of smoothing splines to fit non-linear relationships [65,66]. All the non-spatial methods we tested here have been proved useful for the modeling of road-kills in previous studies. For instance, Sillero et al. [22] efficiently addressed the influence of landscape factors on amphibian road-kill levels in Slovenia using GLM; Wright et al. [67] used GAM to characterize the seasonal trends in hedgehog roadkill levels; Grilo et al. [68] predicted road-kill rates of mammals and birds in Europe using RF; and Ascensão et al. [56] used BRT to relate road mortality levels of mammals with environmental variables in Brazil.

Spatial approaches (GWR) have several advantages over non-spatial analyses: GWR helps in the visualization of spatial variability within the study area, allows the evaluation of spatially clustered data [34,69], and allows mapping local parameter coefficients as well as local model diagnostics (obtained through local pseudo-R^2) to visualize and interpret the spatial non-stationarity. Other ecological studies found a better performance of GWR in comparison with least squares regression [34,40], GLM [39,70], and GAM [39,61,66,70]. Mellin et al. [55] suggested the use of local spatial weights implemented in other modeling methods such as boosted regression trees for ecological modeling. In GWR, the ability to map parameter coefficients allows the identification of missing values or factors associated with non-stationarity [54]. Some variables might be important or might have a certain effect in some regions and not in others. For instance, in our models, we noticed that the variable distance to urban areas had positive local coefficients in roads R1 and R2 (more urbanized areas, closer to the coast) and negative coefficients in road R3 (a greener area, close to a national park; Figure 5). The variable distance to broadleaved forests showed the opposite effect (Figure 5). This result is interesting, in light of the effect of landscape urbanization in road-kill rates [24]: in more urbanized regions, the amphibians might get road-killed more often closer to natural forests while in more natural regions, amphibians might get road-killed more often closer to urban areas. GWR is also useful to explore and visualize the data [34]. We noticed that the road R2 got the poorest explanatory

performance (some locations with low local pseudo-R^2; Figure 3) but also some points with the highest performance. This may be revealing some underlying biased effect in this road or that some important explanatory variable is missing. Moreover, using the information provided by the distribution of residuals and the local pseudo-R^2 values, we may be able to identify the areas that are not providing trustworthy results and further investigate the reasons.

North Portugal is a densely fragmented, urbanized, and populated region [71]. Consequently, we registered a mean number of 0.56 road-killed amphibians per kilometer. This is a very low value considering that in Southern Portugal the mean abundance of road-killed amphibians is around 4.8 ind/km [72] but corresponds to similar levels of road-kills found in a previous study in northern Portugal (Matos et al. [9]: approximately 0.45 ind/km for the surveys taking place between February and May 2011). Another study in Southern Portugal registered a very different value of road-killed amphibians per kilometer (Mestre et al. [21]: 0.14 ind/km); however, this study took place over two years and covered almost all the seasonal variability (between September and May). The density of road-killed amphibians we obtained in our study in north Portugal is similar as well to the one described for the Salamanca region, Spain (0.23 ind/km), which also took place during spring [7]. Amphibian populations near our study roads have not been quantified; thus, we do not know how road-kill intensity is really affecting those populations. These populations may have been extinguished near roads with great urbanization, suffering over long-time high rates of road-kills [24]. Although the number of road-killed animals on roads is highly variable within roads, seasons, and days [26,73], the low number of road-killed amphibians found in this study is somewhat alarming.

5. Conclusions

In summary, our results suggest that GWR is a useful tool for modeling road-kills, as well as to better explore, visualize, and map the spatial variability of the models. We recommend using GWR for modeling road-kill events, as their driving factors are not stationary across space. An extension of this comparison between techniques to model amphibian road-kills would be desirable, namely including a larger dataset and including other variables related to traffic and species abundance. Finally, our study will contribute to better understand the influence of non-stationary environmental variables in the occurrence of road-kills, which might help to better apply mitigation measures.

References

1. Goosem, M. Fragmentation impacts caused by roads through rainforests. *Curr. Sci.* **2007**, *93*, 1587–1595.
2. Ibisch, P.; Hoffmann, M.T.; Kreft, S.; Pe'er, G.; Kati, V.; Biber-Freudenberger, L.; DellaSala, D.A.; Vale, M.M.; Hobson, P.R.; Selva, N. A global map of roadless areas and their conservation status. *Science* **2016**, *354*, 1423–1427. [CrossRef]
3. Fahrig, L. Effects of Habitat Fragmentation on Biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **2003**, *34*, 487–515. [CrossRef]
4. Seiler, A. The Toll of the Automobile: Wildlife and Roads in Sweden. Ph.D. Thesis, Swedish University of Agricultural Sciences, Uppsala, Alnarp, Skara, Sweden, 2003.
5. Ewers, R.M.; Didham, R.K. Confounding factors in the detection of species responses to habitat fragmentation. *Biol. Rev.* **2006**, *81*, 117–142. [CrossRef] [PubMed]
6. Carr, L.W.; Fahrig, L. Effect of road traffic on two amphibian species of differing vagility. *Conserv. Biol.* **2001**, *15*, 1071–1078. [CrossRef]
7. Sillero, N. Amphibian mortality levels on Spanish country roads: Descriptive and spatial analysis. *Amphib. Reptil.* **2008**, *29*, 337–347. [CrossRef]
8. Garriga, N.; Santos, X.; Montori, A.; Richter-Boix, A.; Franch, M.; Llorente, G.A. Are protected areas truly protected? The impact of road traffic on vertebrate fauna. *Biodivers. Conserv.* **2012**, *21*, 2761–2774. [CrossRef]
9. Matos, C.; Sillero, N.; Argana, E. Spatial analysis of Amphibian road mortality levels in northern Portugal country roads. *Amphib. Reptil.* **2012**, *33*, 469–483. [CrossRef]
10. Gibbs, J.P.; Shriver, W.G. Can road mortality limit populations of pool-breeding amphibians? *Wetl. Ecol. Manag.* **2005**, *13*, 281–289. [CrossRef]
11. Lima, S.L.; Blackwell, B.F.; DeVault, T.L.; Fernández-Juricic, E. Animal reactions to oncoming vehicles: A conceptual review. *Biol. Rev.* **2015**, *90*, 60–76. [CrossRef] [PubMed]
12. Alford, R.A.; Richards, S.J. Global amphibian declines: A problem in applied ecology. *Annu. Rev. Ecol. Syst.* **1999**, *30*, 133–165. [CrossRef]
13. Stuart, S.N.; Chanson, J.S.; Cox, N.A.; Young, B.E.; Rodrigues, A.S.; Fischman, D.L.; Waller, R.W. Status and trends of amphibian declines and extinctions worldwide. *Science* **2004**, *306*, 1783–1786. [CrossRef] [PubMed]
14. Meijer, J.R.; Huijbregts, M.A.J.; Schotten, K.C.G.J.; Schipper, A.M. Global patterns of current and future road infrastructure. *Environ. Res. Lett.* **2018**, *13*, 064006. [CrossRef]
15. Carvalho, F.; Mira, A. Comparing annual vertebrate road kills over two time periods, 9 years apart: A case study in Mediterranean farmland. *Eur. J. Wildl. Res.* **2011**, *57*, 157–174. [CrossRef]
16. Clevenger, A.P.; Chruszcz, B.; Gunson, K.E. Spatial patterns and factors influencing small vetebrate fauna road-kill aggregations. *Biodivers. Conserv.* **2003**, *109*, 15–26.
17. Seiler, A.; Helldin, J.; Seiler, C. Road mortality in Swedish mammals: Results of a drivers' questionnaire. *Wildl. Biol.* **2004**, *10*, 183–191. [CrossRef]
18. Seiler, A. Predicting locations of moose-vehicle collisions in Sweden. *J. Appl. Ecol.* **2005**, *2*, 371–382. [CrossRef]
19. Coelho, I.P.; Teixeira, F.Z.; Colombo, P.; Coelho, A.V.; Kindel, A. Anuran road-kills neighboring a peri-urban reserve in the Atlantic Forest, Brazil. *J. Environ. Manag.* **2012**, *112*, 17–26. [CrossRef]
20. Bennett, V.J. Effects of Road Density and Pattern on the Conservation of Species and Biodiversity. *Curr. Landsc. Ecol. Rep.* **2017**, *2*, 1–11. [CrossRef]
21. Mestre, F.; Lopes, H.; Pinto, T.; Sousa, L.G.; Mira, A.; Santos, S.M. Bad moon rising? The influence of the lunar cycle on amphibian roadkills. *Eur. J. Wildl. Res.* **2019**, *65*, 58. [CrossRef]
22. Sillero, N.; Poboljsaj, K.; Lesnik, A.; Salamun, A. Influence of landscape factors on amphibian roadkills at the national level. *Diversity* **2019**, *11*, 13. [CrossRef]
23. D'Amico, M.; Román, J.; de los Reyes, L.; Revilla, E. Vertebrate road-kill patterns in Mediterranean habitats: Who, when and where. *Biol. Conserv.* **2015**, *191*, 234–242. [CrossRef]
24. Ascensão, F.; Kindel, A.; Zimmermann Teixeira, F.; Barrientos, R.; D'Amico, M.; Borda-de-Água, L.; Pereira, H.M. Beware that the lack of wildlife mortality records can mask a serious impact of linear infrastructures. *Glob. Ecol. Conserv.* **2019**, *19*, e00661. [CrossRef]
25. Hasan, M.T.; Sneddon, G.; Ma, R. Modeling binomial amphibian roadkill data in distance sampling while accounting for zero-inflation, serial correlation and varying cluster sizes simultaneously. *Environ. Ecol. Stat.* **2017**, *24*, 201–217. [CrossRef]
26. Lin, Y.P.; Anthony, J.; Lin, W.C.; Lien, W.Y.; Petway, J.R.; Lin, T.E. Spatiotemporal identification of roadkill probability and systematic conservation planning. *Landsc. Ecol.* **2019**, *34*, 717–735. [CrossRef]
27. Petrovan, S.O.; Vale, C.G.; Sillero, N. Using citizen science in road surveys for large-scale amphibian monitoring: Are biased data representative for species distribution? *Biodivers. Conserv.* **2020**, *29*, 1767–1781. [CrossRef]

28. Grilo, C.; Ascensão, F.; Santos-Reis, M.; Bissonette, J.A. Do well-connected landscapes promote road-related mortality? *Eur. J. Wildl. Res.* **2011**, *57*, 707–716. [CrossRef]

29. Sheehan, K.R.; Strager, M.P.; Welsh, S.A. Advantages of geographically weighted regression for modelling benthic substrate in two greater Yellowstone ecosystem streams. *Eviron. Model. Assess.* **2013**, *18*, 209–219. [CrossRef]

30. Graham, H.; Elith, J.; Hijmans, J.; Guisan, A.; Townsend Peterson, A.; Loiselle, B.A.; NCEAS Predicting Species Distributions Working Group. The influence of spatial errors in species occurrence data used in distribution models. *J. Appl. Ecol.* **2008**, *45*, 239–247. [CrossRef]

31. Ali, K.; Partridge, D.; Olfert, M.R. Can geographically weighted regressions improve regional analysis and policy making? *Int. Reg. Sci. Rev.* **2007**, *30*, 300–329. [CrossRef]

32. Gao, J.; Li, S. Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using geographically weighted regression. *Appl. Geogr.* **2011**, *31*, 292–302. [CrossRef]

33. Fotheringham, A.S.; Brunsdon, C.; Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*; John Wiley & Sons Inc.: Chichester, UK, 2002.

34. Kala, A.K.; Tiwari, C.; Mikler, A.R.; Atkinson, S.F. A comparison of least squares regression and geographically weighted regression modeling of West Nile virus risk based on environmental parameters. *PeerJ* **2017**, *5*, e3070. [CrossRef]

35. Erdogan, S. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *J. Saf. Res.* **2009**, *40*, 341–351. [CrossRef] [PubMed]

36. Zheng, L.; Robinson, R.M.; Khattak, A.; Wang, X. All Accidents are Not Equal: Using Geographically Weighted Regressions Models to Assess and Forecast Accident Impacts. In Proceedings of the 3rd International Conference on Road Safety and Simulation, Indianapolis, IN, USA, 14–16 September 2011.

37. Austin, M. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecol. Model.* **2007**, *200*, 1–19. [CrossRef]

38. Kupfer, J.A.; Farris, C.A. Incorporating spatial non-stationarity of regression coefficients into predictive vegetation models. *Landsc. Ecol.* **2007**, *22*, 837–852. [CrossRef]

39. Nogués-Bravo, D. Comparing regression methods to predict species richness patterns. *Web Ecol.* **2009**, *9*, 58–67. [CrossRef]

40. Ye, X.; Yu, X.; Wang, T. Investigating spatial non-stationary environmental effects on the distribution of giant pandas in the Qinling Mountains, China. *Glob. Ecol. Conserv.* **2020**, *21*, e00894. [CrossRef]

41. Loureiro, A.; Ferrand, N.; Carretero, M.A.; Paulo, O. *Atlas dos Anfíbios e Répteis de Portugal*, 1st ed.; Esfera do Caos: Lisbon, Portugal, 2010.

42. Ribeiro, R.; Torres, J.; Gomes, V.; Carretero, M.A.; Sillero, N.; Llorente, G.A. Unsuspected richness near home: New herpetological records in Porto Metropolitan Area (NW Portugal). *Bol. De La Asoc. Herpetol. Esp.* **2010**, *21*, 27–30.

43. Orlowski, G. Spatial distribution and seasonal pattern in road mortality of the common toad Bufo bufo in an agricultural landscape of south-western Poland. *Amphib. Reptil.* **2007**, *28*, 25–31. [CrossRef]

44. Glista, D.J.; DeVault, T.L.; DeWoody, J.A. Vertebrate road mortality predominantly impacts amphibians. *Herpetol. Conserv. Biol.* **2008**, *3*, 77–87.

45. Santos, X.; Llorente, G.A.; Montori, A.; Carretero, M.A.; Franch, M.; Garriga, N.; Richter-Boix, A. Evaluating factors affecting amphibian mortality on roads: The case of the Common Toad *Bufo bufo*, near a breeding place. *Anim. Biodivers. Conserv.* **2007**, *30*, 97–104.

46. Santos, S.M.; Carvalho, F.; Mira, A. How long do the dead survive on the road? Carcass persistence probability and implications for road-kill monitoring surveys. *PLoS ONE* **2011**, *6*, e25383. [CrossRef] [PubMed]

47. Preatoni, D.G.; Tattoni, C.; Bisi, F.; Masseroni, E.; D'Acunto, D.; Lunardi, S.; Grimod, I.; Martinoli, A.; Tosi, G. Open source evaluation of kilometric indexes of abundance. *Ecol. Inform.* **2012**, *7*, 35–40. [CrossRef]

48. Semlitsch, R.D.; Bodie, J.R. Biological criteria for buffer zones around wetlands and riparian habitats for amphibians and reptiles. *Conserv. Biol.* **2003**, *17*, 1219–1228. [CrossRef]

49. Smith, A.M.; Green, D.M. Dispersal and the metapopulation paradigm in amphibian ecology and conservation: Are all amphibian populations metapopulations? *Ecography* **2005**, *28*, 110–128. [CrossRef]

50. Guisan, A.; Zimmermann, N.E. Predictive habitat distribution models in ecology. *Ecol. Model.* **2000**, *135*, 147–186. [CrossRef]

51. Field, A.; Miles, J.; Field, Z. *Discovering Statistics Using R*; Sage Publications Ltd.: London, UK, 2012.

52. Breiman, L. Random forest. *Mach. Learn.* **1999**, *45*, 1–35.

53. Elith, J.; Leathwick, J.R.; Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **2008**, *77*, 802–813. [CrossRef]

54. Miller, J.A. Species distribution models spatial autocorrelation and non-stationarity. *Prog. Phys. Geogr.* **2012**, *36*, 681692. [CrossRef]

55. Mellin, C.; Mengersen, K.; Bradshaw, C.J.A.; Caley, M.J. Generalizing the use of geographical weights in biodiversity modelling. *Glob. Ecol. Biogeogr.* **2014**, *23*, 1314–1323. [CrossRef]

56. Ascensão, F.; Desbiez, A.L.J.; Medici, E.P.; Bager, A. Spatial patterns of road mortality of medium-large mammals in Mato Grosso do Sul, Brazil. *Wildl. Res.* **2017**, *44*, 135–146. [CrossRef]

57. Oshan, T.M.; Li, Z.; Kang, W.; Wolf, L.J.; Fotheringham, A.S. mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 269. [CrossRef]

58. Elith, J.; Graham, C.H.; Anderson, P.R.; Dudík, M.; Ferrier, S.; Guisan, A.; JHijmans, R.; Huettmann, F.; RLeathwick, J.; Lehmann, A.; et al. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **2006**, *29*, 129–151. [CrossRef]

59. Willmott, C.J.; Robeson, S.M.; Matsuura, K. A refined index of model performance. *Int. J. Climatol.* **2011**, *32*, 2088–2094. [CrossRef]
60. Zou, K.H.; O'Malley, A.J.; Mauri, L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* **2007**, *115*, 654–657. [CrossRef]
61. Windle, M.J.S.; Rose, G.A.; Devillers, R.; Fortin, M.-J. Exploring spatial non-stationarity of fisheries survey data using geographically weighted regression (GWR): An example from the Northwest Atlantic. *ICES J. Mar. Sci.* **2010**, *67*, 145–154. [CrossRef]
62. Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.
63. Qiao, H.; Soberón, J.; Peterson, A.T. No silver bullets in correlative ecological niche modelling: Insights from testing among many potential algorithms for niche estimation. *Methods Ecol. Evol.* **2015**, *6*, 1126–1136. [CrossRef]
64. Lopatin, J.; Dolos, K.; Hernández, H.J.; Galleguillos, M.; Fassnacht, F.E. Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. *Remote Sens. Environ.* **2016**, *173*, 200–210. [CrossRef]
65. Gu, C.; Wahba, G. Minimizing GCV/GML scores with multiple smoothing parameters via the Newton method. *SIAM J. Sci. Stat. Comput.* **1991**, *12*, 383–398. [CrossRef]
66. Li, M.; Zhang, C.; Xu, B.; Xue, Y.; Ren, Y. A comparison of GAM and GWR in modelling spatial distribution of Japanese mantis shrimp (Oratosquilla oratoria) in coastal waters. *Estuar. Coast. Shelf Sci.* **2020**, *244*, 106928. [CrossRef]
67. Wright, P.G.R.; Coomber, F.G.; Bellamy, C.C.; Perkins, S.E.; Mathews, F. Predicting hedgehog mortality risks on British roads using habitat suitability modelling. *PeerJ* **2020**, *7*, e8154. [CrossRef] [PubMed]
68. Grilo, C.; Koroleva, E.; Andrášik, R.; Bíl, M.; González-Suárez, M. Roadkill risk and population vulnerability in European birds and mammals. *Front. Ecol. Environ.* **2020**, *18*, 323–328. [CrossRef]
69. Goovaerts, P. Geostatistical analysis of health data: State-of-the-art and perspectives. In *GeoENV VI—Geostatistics for Environmental Applications (Quantitative Geology and Geostatistics)*; Soares, A., Pereira, M.J., Dimitrakopoulos, R., Eds.; Springer: Dordrecht, The Netherlands, 2008; Volume 15.
70. Osborne, P.E.; Foody, G.M.; Suárez-Seoane, S. Non-stationarity and local approaches to modelling the distributions of wildlife. *Divers. Distrib.* **2007**, *13*, 313–323. [CrossRef]
71. INE. *Estatísticas dos Transportes e Comunicações 2019—Instituto Nacional de Estatística (I.N.E)*; INE: Lisboa, Portugal, 2020; p. 106.
72. Baptista, N. Amphibian Roadkills: Hotspot Analysis and Locations of Amphibian Underpasses Using Gorelick's Index. Master's Thesis, University of Évora, Largo dos Colegiais, Portugal, 2006.
73. Garriga, N.; Franch, M.; Santos, X.; Montori, A.; Llorente, G.A. Seasonal variation in vertebrate traffic casualties and its implications for mitigation measures. *Landsc. Urban Plan.* **2017**, *157*, 36–44. [CrossRef]