HERPETOLOGICAL JOURNAL, Vol. 6, pp. 43-47 (1996)

DISTRIBUTION OF SCHREIBER'S GREEN LIZARD (*LACERTA SCHREIBERI*) IN PORTUGAL: A PREDICTIVE MODEL

JOSÉ C. BRITO, FAUSTO BRITO E ABREU, OCTÁVIO S. PAULO, HUMBERTO D. ROSA AND EDUARDO G. CRESPO

Dept. Zoologia e Antropologia, Faculdade de Ciências da Universidade de Lisboa, P-1700 Lisboa, Portugal

The geographic distribution of the endemic Iberian lizard *Lacerta schreiberi* in Portugal was determined through extensive field surveys. Subsequently, a logistic regression model which predicts the probability of occurrence, based on environmental variables, was developed. We found that *L. schreiberi* is more widely distributed than previously thought, through most of central/northern Portugal, including the coastal zone and extending into low altitude zones. New isolated populations were also detected and the area occupied by three previously known southern isolates was enlarged. The model indicates that the distribution of *L. schreiberi* is largely explained by environmental parameters such as insolation, evapotranspiration, rain, humidity and soil-drainage. Values of probability of occurrence greater than 0.50, as determined by our model, correspond with the actual presence of the species.

INTRODUCTION

Lacerta schreiberi is a medium-sized lizard (adult snout-vent-length 117-120 mm), endemic to the Iberian Peninsula. It exhibits a pronounced sexual dimorphism, the adult males being smaller than the females and having an intense blue coloration on the head during the mating season (Barbadillo, 1987). It usually occupies mountain regions (up to 2100 m) and their surroundings, but may also be found at low altitudes if suitable conditions occur (Marco & Pollo, 1993). *L. schreiberi* inhabits relatively high humidity locations, mostly water courses, with dense vegetation cover. It is a good climber of stone walls and bushes (Brito, 1994), and feeds mostly on Coleoptera, Formicidae and Diptera (Marco & Pérez-Mellado, 1988).

L. schreiberi has a marked Atlantic distribution, occurring in Spain in Galicia, Cantábria, Northern Castilla-León, Central Range Mountains (Sistema Central) and also in several isolated populations in the south (De La Riva, 1987; Marco & Pollo, 1993) (Fig. 1). In Portugal, it occurs in the northern half of the country and in three known isolated populations in the south: Serra de S. Mamede, Serra de Sintra and Serra de Monchique (Malkmus, 1981; Crespo & Oliveira, 1989; Malkmus, 1995) (Fig. 1).

Our work was conducted as part of a project on the conservation of *L. schreiberi* in Portugal, under guidelines of the European Union LIFE Program. The two aims were (1) to describe through field surveys the distribution of the species in Portugal, and (2) to derive a predictive model for the distribution, using logistic regression, in order to evaluate the probability of occurrence of the species in any given area.

Logistic regression is a tool for analysing the effects of one or several independent variables, discrete or continuous, on one dependent polychotomous variable (e.g., presence/absence). It has been used recently for modelling wildlife distributions (Walker, 1990) and abundance (Gates, Gibbons, Lack & Fuller, 1994). A logistic regression has the form: $\pi(x) = e^{g(x)} / (1 + e^{g(x)})$, where e is the base of the Naperian logarithm and $\pi(x)$ is the probability of occurrence of the species, ranging from 0 to 1. The value of g(x) is obtained by a regression equation of the form: $g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p$ where β_0 is a constant and $\beta_1 ... \beta_p$ are the coefficients of the $x_1 ... x_p$ independent variables. A multivariate logistic regression model, based on environmental variables, was used to predict the probability of occurrence of *L. schreiberi* in all of the Portuguese UTM (Universal Transverse of Mercator grid) 10x10 km cells.

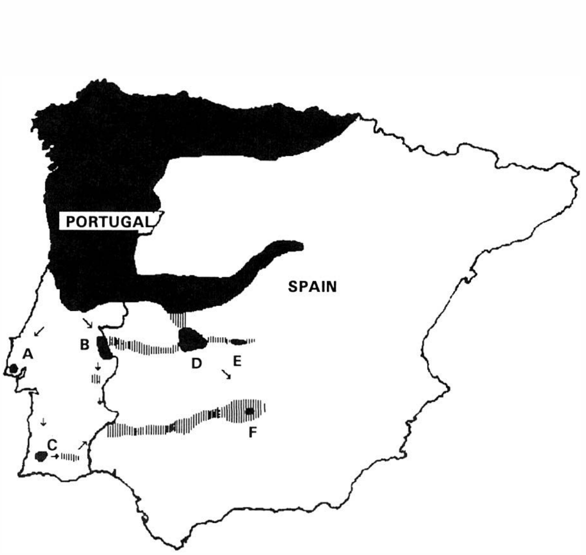

FIG. 1. Previously known distribution of *L. schreiberi* in the Iberian Peninsula. A, Serra de Sintra; B, Serra de S. Mamede; C, Serra de Monchique; D, Sierra de las Villuercas; E, Montes de Toledo; F, Sierra de San Andrés. Shaded: known area; vertical hatching: possible occurrences; arrows: possible ancient connection passages (Modified from De La Riva, 1987).

MATERIAL AND METHODS

DISTRIBUTION

The field sampling strategy for surveys was based on two criteria: first, the previously known distribution according to Malkmus (1981, 1995) and Crespo & Oliveira (1989), but instead of reconfirming this distribution we opted for trying to delimit the areas in which the lizard was present. Second, we sought to determine the potential areas of occurrence, which included habitats where the environmental characteristics, considered important by Marco & Pollo (1993) (e.g. annual rainfall above 800 mm), are favourable for *L. schreiberi*. These characteristics were obtained from the Portuguese Environmental Atlas (C.N.A., 1983). The field work was conducted between March and September 1994, and covered most of the northern/central Portuguese territory and also the southern regions where isolates are known or where areas of potential occurrence exist, in a total area of approximately 31000 km^2. In total, 470 rivers and streams were sampled. At each, two observers diverged in opposite directions from a common starting point, and walked for about 20 min., visually detecting lizards within the water margins. The known information on the distribution was compiled in a database and represented through occurrence points in a UTM 10x10 km grid, which was later updated with our results.

LOGISTIC REGRESSION

A matrix was built with information on presence/absence of the species (dependent variable) in 422 UTM 10x10 km cells. Each cell was assigned values for the following independent variables, thought to be relevant, and grouped in classes (Appendix 1): number of hours of sun per year (*insolation*); mean daily temperature (*temperature*); total rain in mm per year (*rain*), relative humidity at 9.00 G.M.T. (*humidity*); amount of water in the streams (*soil drainage*); amount of water returned to the atmosphere (*evapotranspiration*); altitude; type of soil (*soil*); phytoedapho-climatic information (*ecology*); number of humans/km^2 (*human population density*).

From the initial set of cells we chose two random subsamples. One subsample, with 362 cells, was used to build and assess the fit of the model. The other, the validation subsample, with 60 cells, was used only to evaluate the accuracy of the predictions of the final model. To determine the regression model, we used the software Egret (1991) and followed the procedure of building and evaluating multivariate logistic regressions suggested by Hosmer & Lemeshow (1989).

UNIVARIATE ANALYSIS

Univariate logistic regression analysis indicates whether there is any association between each environmental variable and the binary dependent variable. The selection of the relevant independent variables was performed using the maximum likelihood ratio test (*G* test) and the Wald test (Hosmer & Lemeshow, 1989). Any variable for which the Wald test had a $P < 0.25$ was considered for inclusion in the multivariate model.

MULTIVARIATE ANALYSIS

Once the variables were selected by univariate analysis, they were included in a preliminary model and in a backward elimination procedure those with $P > 0.25$ and those whose Odds ratio estimation (95% confidence) included the value 1 were eliminated. The Odds ratio is a measure of association which approximates how much more likely (or unlikely) it is for the outcome to be present among those with x=1 than among those with x=0 (Hosmer & Lemeshow, 1989). For example, if y denotes the presence/absence of a species and if x denotes a binary independent variable, then an Odds ratio of 2 indicates that the species occurs twice as often among the places with the variable equal to one than among those with value zero. In continuous variables we have to develop a method for point and interval estimation for an arbitrary change of c units in the variable. The Odds ratio for a change of c units of the variable x is obtained through: $\Psi(c)=\exp(c\beta_1)$, being β_1 the coefficient for the variable x. Hence, variables are rejected when their Odds ratio estimation (95% confidence) include the value 1, which denotes a low level of association between the variable and the presence/absence of the species (Hosmer & Lemeshow, 1989).

LINEARITY

Continuous variables were checked for linearity, by two methods. Adding the Box-Tidwell transformation (x.ln x) to the model, resulting in evidence of non-linearity if the values of the P and G tests for this new variable were significant. The other method involved fitting the square and log transformations of each variable and performing a univariate analysis with these new variables. If the G test or the Wald tests were significant the transformed variable was included in the model and used in subsequent analyses (see Hosmer & Lemeshow, 1989; and Gates *et al.*, 1994, for technical details).

CONFOUNDED VARIABLES AND INTERACTIONS

Any confounded variables were assessed by adding all the initial variables, one after the other, to the model and simultaneously checking the effect of the added variable on the estimated coefficients of the variables that were already in the model, and on the G test. Confounded effects were significant when an increment of the G test value or its significance was observed. Interactions were added to the model, derived from the previous analysis, by creating a new variable which is equal to the product of the value of two variables. Again, interactions were significant if an increment of the *G* test value or its significance was observed, and if so they were included in the model (see Hosmer & Lemeshow (1989) for technical details).

ASSESSING THE FIT OF THE MODEL

To assess the fit of the model we used the Pearson chi-square and a classification table (Hosmer & Lemeshow, 1989). The Pearson chi-square was calculated as indicated below:

$$X^2=\sum_{j=1}^{n} r(y_j,\pi_j)^2$$

$$\text{where } r(y_j,\pi_j) = \frac{(y_j - \pi_j)}{(\pi_j(1 - \pi_j))^{0.5}}$$

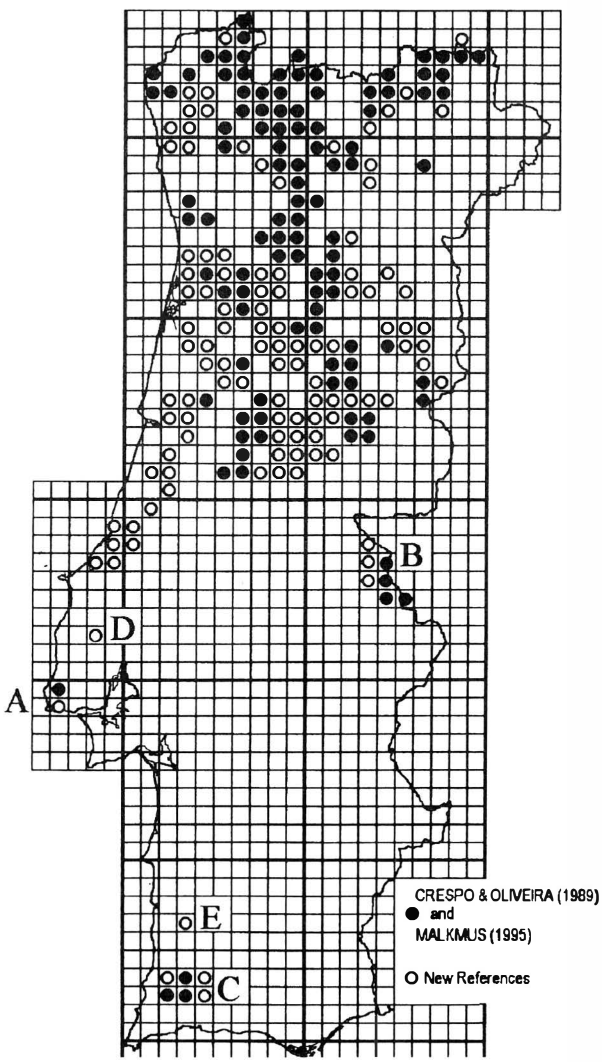

FIG. 2. Updating of the distribution area of *L. schreiberi* in Portugal with new sites detected in this work. A, Serra de Sintra; B, Serra de S. Mamede; C, Serra de Monchique; D, Serra de Montejunto; E, Serra do Cercal (Grid: UTM 10 x 10 km).

The classification table considered the observed value of the dependent variable, the presence or absence of the species, and the predicted value of presence or absence as a function of the estimated probability of occurrence based on the multivariate model. To generate the table, it was necessary to set a cut-off point, the probability value above which we consider that the species is present. We follow the recommended cut-off point of 0.30 suggested by Walker (1990) for this kind of analysis.

RESULTS AND DISCUSSION

DISTRIBUTION

Fig. 2 presents the new sites of occurrence found during our field work together with those previously known. Expansion of the geographic distribution is evident in three cases. First, there is an extension to all the littoral strips, to the north of the river Tagus, including zones of high human pressure (urbanization

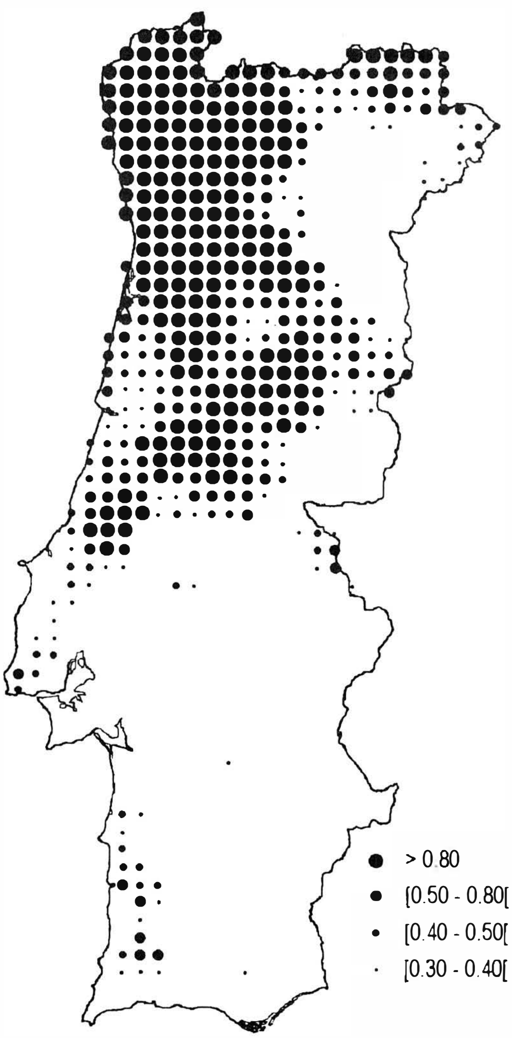

FIG. 3. Probability of occurrence of *L. schreiberi* in Portugal following the model defined (each circle represents the probability of a UTM 10 x 10 km cell).

and industries). In this coastal area, the presence of *L. schreiberi* is frequent in marshes where the vegetation cover is of reed (*Arundo donax*) and slender reed (*Phragmites* sp.) (Brito *et al.*, 1994). This habitat has not been previously considered as suitable for this species.

Second, for central/northern Portugal, the occurrence of *L. schreiberi* was only known for areas of relatively high altitude (cf. Crespo & Oliveira, 1989; Malkmus, 1995). Our work has shown that the distribution is continuous for this region rather than restricted to mountains, with an exception for part of the north-eastern region and the more interior part of the River Douro. These exceptions are probably due to climatic constraints (see below).

Thirdly, the area occupied by the known isolated populations in the south (Sintra, S. Mamede and Monchique) has proved to be wider than previously reported (Malkmus 1981, 1995; Crespo & Oliveira, 1989). In addition, two new isolates were detected: near Serra de Montejunto, in the region of Estremadura, and in Serra do Cercal, Alentejo (Fig. 2). Although the possibility of a connection between the Sintra and Montejunto isolates with the northern area was considered, later field work could not confirm this. The reduced number of individuals detected in these two new isolates, and also in the isolate of Serra de Sintra, leads us to believe that these are very small populations and may be facing extinction.

The general enlargement of the distribution of *L. schreiberi* in Portugal is, in our view, the result of a lack of extensive sampling in the past, and not to a subsequent expansion of the species.

LOGISTIC REGRESSION

Although six models were produced, we only present one because this explains significantly more variation than all the others. For this model, in the univariate analysis we eliminated the variables *population density* and *soil*, and in the multivariate analysis, the variables *temperature*, *altitude* and *ecology* were also eliminated. No confounded effects were detected but interactions appeared between *insolation* and *evapotranspiration* and between *rain* and *humidity*. Although *temperature* and *humidity* showed evidence of non-linearity, the inclusion of the transformed variables resulted in a regression model with less statistical power. The resulting regression equation is as follows:

$$g(x) = -5.001 - 0.04028\ (insolation \times evapotranspiration) + 0.104\ (rain \times humidity) + 0.7023\ (soil\ draining)$$

The model agrees closely with the observed values. The Pearson chi-square ($X^2_{0.05,356}$ = 374.97) and the classification table (75% of presences were classified as presence; 94% of absences were classified as absence and 82% of the cells were correctly classified) suggests that the model is a good predictor of the distribution of *L. schreiberi*. The validation sample gave inferior chi-square values and percentages ($X^2_{0.001,54}$= 83.89, 56% of presences were classified as presence; 96% of absence were classified as absence and 72% of the cells were correctly classified). The model explains absence of the species much better than presence. The smaller value obtained by the "presence classification" is due to the fact that the model only considers climatic variables as determinants for the distribution. In reality, climate is not the only factor determining *L. schreiberi* distribution. For example, other habitat characteristics may be unsuitable in areas which are climatically ideal for the species. This is especially evident in the littoral areas (cf. Fig. 3 with Fig. 2) where the climatic conditions are favourable but the habitat is widely disturbed, so although the model produces a high score for these areas *L. schreiberi* is frequently absent. Future research should aim at improving the model by incorporating other relevant variables, such as descriptors of habitat quality. The results could also be improved either by changing the scale to a more precise one or by using different modelling approaches, namely by using different sets of random cells in order to deal with spatial autocorrelation.

The probability of occurrence of *L. schreiberi* in Portugal in presented in Fig. 3. The null probability for the inner Douro River and north-eastern region confirms the climatic unsuitability of this area for this species. Above the 0.50 probability of occurrence, the presence of the species is frequent (cf. Fig. 2). The map of probabilities also gives some clues for further research on this species.

ACKNOWLEDGEMENTS

This research was supported by the "LIFE" European Union Program and by Instituto da Conservação da Natureza, Portugal. We thank all the colleagues that helped in the field work, but special thanks are due to Constança Camilo-Alves, Luis Carneiro, Pedro Leitão, Pedro Segurado and Rui Cabanita. Special thanks also to Nelson Raimundo and João Pargana for data on Serra de S. Mamede. We thank R. Malkmus for assent unpublished data and A. Marco for his bibliography on the species. We also thank to Dr. Ilik Saccheri for his useful comments on early versions of this manuscript, and to the comments of an anonymous referee.

REFERENCES

Barbadillo, L. J. (1987). *La Guia de INCAFO de los Anfibios y Reptiles de la Peninsula Iberica, Islas Baleares y Camarias.* Madrid: Incafo S. A..

Brito, J. C. (1994). *Areas Potenciais para a Conservação do Lagarto-de-Água, Lacerta schreiberi, em Portugal.* Degree Thesis. Dept. Zoologia e Antropologia, Faculdade de Ciências Universidade de Lisboa.

Brito, J.C., Pinto, I., Rosário, I., Pombo, I., Monteiro, J. L., Brito e Abreu, F., Paulo O. S., Rosa, H. D.; Crespo, E. G. (1994). New populations of *Lacerta schreiberi* found in the Portuguese Estremadura. Abstract *III Congreso Luso-Español de Herpetologia*, Badajoz, 19-23/09/1994.

C.N.A. (1983). *Atlas do Ambiente*. (Lisboa): Direcçãao-Geral do Ambiente, Ministério do Ambiente e dos Recursos Naturais.

Crespo, E. G., Oliveira, M. E., (1989). *Atlas da Distribuição dos Anfibios e Répteis de Portugal Continental*. (Lisboa): S.N.P.R.C.N.

De La Riva, I. (1987). Zoogeografia de *Lacerta schreiberi* Bedriaga, 1878. *Rev. Esp. Herpetologia* **2**, 49-69.

Egret, (1991). EGRET - Statistical analysis package. Statistics and Epidemiological Research Corporation, Seattle, WA 98105, USA.

Gates, S., Gibbons, D. W., Lack, P. C. Fuller, R. J. (1994). Declining farmland bird species: modelling geographical pattems of abundance in Britain. In: *Large-scale ecology and conservation biology*, 111-128. Edwards, P. J., May, R. M., Webb, N. R. (Eds.). London: Blackwell Scientific Publications.

Hosmer, Jr., D.W., Lemeshow, S. (1989). *Applied Logistic Regression*. New York: John Wiley & Sons.

Malkmus, R. (1981). Zur verbbreitung der Iberischen Smaragdeidechse *Lacerta schreiberi* BEDRIAGA 1878 in Portugal sudlich des 40' breitengrades. *Nachr. Natur. Mus. Aschaffenburg* **89**, 60-74.

Malkmus, R. (1995). *Die Amphibien und Reptilien Portugals, Madeiras und der Azoren*. Magdeburg: Westarp Wissenschaften.

Marco, A., Pérez-Mellado, V. (1988). Alimentación de *Lacerta schreiberi* Bedriaga,1878 (Sauria: Lacertidae) en el Sistema Central. *Rev. Esp. Herpetologia* **3**, 133-141.

Marco, A., Pollo, C. P. (1993). Analisis biogeografico de la distribucion del lagarto verdinegro *(Lacerta schreiberi* Bedriaga, 1878). *Ecologia* **7**, 457-466.

Walker, P.A. (1990). Modelling wildlife distributions using a geographic information system: kangaroos in relation to climate. *J. Biogeog.* **17**, 279-289.

Accepted: 8.11.95

APPENDIX 1. Classes of the environmental variables used in the regression equation

class	insolation (hrs/yr)	evapotranspiration (mm/yr)	rain (mm/yr)	humidity (%)	soil drainage (mm)
1	<1900	<350	<400	<65	<25
2	1900-2000	350-400	400-500	65-70	25-50
3	2000-2100	400-450	500-600	70-75	50-100
4	2100-2200	450-500	600-700	75-80	100-150
5	2200-2300	500-600	700-800	80-85	150-200
6	2300-2400	600-700	800-1000	>85	200-300
7	2400-2500	700-800	1000-1200	-	300-400
8	2500-2600	>800	1200-1400	-	400-600
9	2600-2700	-	1400-1600	-	600-800
10	2700-2800	-	1600-2000	-	800-1000
11	2800-2900	-	2000-2400	-	1000-1400
12	2900-3000	-	2400-2800	-	1400-1800
13	3000-3100	-	>2800	-	1800-2200
14	>3100	-	-	-	>2200