Hystrix, (n.s.) 7 (1-2) (1995): 209-218 Proc. II It. Symp. on Carnivores

A PREDICTIVE MODEL OF THE EFFECT OF ENVIRONMENTAL FACTORS ON THE OCCURRENCE OF OTTERS *(LUTRA LUTRA* L.) IN HUNGARY

ILDIKÓ KEMENES* & ANDRÁS DEMETER**

Hungarian Natural History Museum, Budapest
**Zoological and Botanical Garden of the City of Budapest Budapest, Állatkerti krt. 6-12, Hungary, 11-1371*
*** Hungarian Academy of Sciences, Budapest, Nádor u. 7. Hungary, H-1361*

ABSTRACT – A survey of the distribution of otters (*Lutra lutra* L.) in Hungary revealed that this species is common in most parts of the country where there appear to be suitable aquatic habitats. However, there were a large number of apparently "good" habitats where no otters were found. On the other hand, in some places where, based on a qualitative assessment, otters should not have been present, we still found signs of them. The only strictly and consistently limiting factor was heavy chemical pollution of the water which could not be assayed during the survey but was analysed based on data provided by the water authorities. These observations led us to employ a quantitative method which takes into account 3 scalable and 5 non-scalable variables of the environment and their relationships which might influence the occurrence of otters. The technique was based on a non-parametric multiple regression method specifically developed for use on PCs. This so called logistic regression model is useful for investigating the relationships between a binary dependent variable and a set of categorical independent variables. We recorded the presence (1) or absence (0) of signs of otters as well as the water depth, steepness of the bank, density of the bank vegetation and the presence or absence of various disturbance factors, such as agricultural use of the water bank, obvious signs of pollution of the water, etc., at 369 sites in Hungary. The three former environmental variables were scaled, whereas the disturbance factors were each assigned a value of either 0 or 1 (0 = absent, 1 = present). The analysis has shown that this method can be used to characterise particular combinations of factors at which otters were most likely to occur and even predictions can be made on the probability of finding otters at particular places with a known combination of these environmental factors. Besides its theoretical importance, this method is a very useful tool to pinpoint sites of possible significance for the purposes of conservation of existing or potential otter habitats.

Key words: *Lutra lutra*, Distribution, Environmental factors, Logistic regression, Habitats, Conservation, Hungary.

RIASSUNTO – *Modello predittivo dell'effetto dei fattori ambientali sulla presenza della Lontra* (Lutra lutra *L.) in Ungheria* – Un'indagine sulla distribuzione della lontra (*Lutra lutra* L.) in Ungheria ha evidenziato che la specie è comune in buona parte del paese dove l'ambiente acquatico appare adatto. C'era comunque un largo numero di habitat apparentemente in buone condizioni dove la presenza della specie non era accertata. Al contrario, alcuni siti, che sulla base di una valutazione qualitativa sarebbero stati negativi per la lontra, rivelavano segni di presenza del mustelide. L'unico fattore che limitava fortemente la presenza della specie era l'elevato inquinamento chimico delle acque, come accertato dai dati raccolti dalle autorità competenti. Quanto sopra esposto ci ha portato all'impiego di un metodo quantitativo che tenesse conto delle variabili ambientali (3 valutate secondo una scala di punteggio e 5 sulla base della loro presenza/assenza) che potevano influenzare la presenza della specie. Il modello di analisi utilizzato, chiamato regressione logistica, è utile per esaminarc le relazioni tra una variabile dipcndente binaria e un gruppo di variabili indipendenti espresse secondo categoric. Per 369 siti esaminati in diversi corpi idrici è stata accertata la presenza (valore 1) o l'assenza (valore 0) della lontra e sono state stimate le seguenti variabili: profondith dell'acqua, pendcnza delle rive, copertura vegetale riparia e la prescnza o assenza di vari fattori legati allc attività produttive, quali, ad esempio, l'uso agricolo dei terreni a ridosso delle rivc dei corpi idrici e la presenza visibile di inquinamento. Le prime tre variabili erano valutate secondo una scala di puntcggio, lc altre invece avevano valore 1 quando prcsenti, 0 quando assenti (v. metodi). I risultati ottenuti hanno evidenziato chc il modello di analisi dei dati può esserc utilizzato per caratterizzare combinazioni di fattori ambientali chc molto probabilmente influenzano la prcscnza della lontra e per valutare la probabilità di ritrovamento del mustelide in siti caratterizzati da una combinazionc nota di fattori. Tale modello, oltre ad essere importante dal punto di vista tcorico, è particolarmente utile per individuarc e proteggere ambienti realmente o potcnzialmente favorevoli alla lontra.

Parole chiave: *Lutra lutra*, Distribuzione, Fattori ambientali, Rcgrcssione logistica, Habitat, Conservazione, Unghcria.

INTRODUCTION

A survey of 369 sites in aquatic habitats in Hungary revealed that more than 50% of them were still inhabited by otters (*Lutra lutra* L.) (Kemenes, 1991). However, the distribution of the otter population was not even: there were places where the population density seemed to be high, while at other areas there was no sign of the presence of otters, despite the fact that these were apparently "good" habitats, as judged by qualitative assessments.

In earlier studies to detect the major influencing factors on the occurrence of otters, a variety of habitat characteristics were recorded and their importance investigated. Pollution of the water proved to be the most limiting factor: the presence of heavy metal ions or PCBs and other pollutants was always correlated with the absence or decline of otter populations in the study areas (Mason, 1989). In non-polluted habitats a variety of other factors, such as the type of thc bank vegetation (Macdonald and Mason, 1985) or the depth of the water (Glimmerveen and Ouwerkerk, 1984) seemed to have influence on the distribution of otters. Most of these studies, however, focused on the effect of single environmental factors, whereas under natural circumstances the occurrence of otters is more likely to be determined by a combination of several characteristics of the habitat.

In the present work, to obtain a more realistic picture of the combined effect of a variety of environmental factors on the distribution of otters in Hungary, we employed a quantitative multivariate method. This takes into account 3 scalable and 5 non-scalable variables of the environment and their relationships which might influence the occurrence of otters. By using this method we tried to answer the following questions:

i) What are the most limiting factors in unpolluted habitats?

ii) What are the characteristics of the most suitable habitats for otters?

iii) What is the probability of finding otters at a particular place with a known combination of a set of environmental factors?

iv) How can this model be used in practical conservation?

METHODS

ESTABLISHING THE PRESENCE OF OTTERS IN AQUATIC HABITATS

During a nation-wide survey between 1988 and 1990, 369 sites at ponds, lakes and rivers were visited in Hungary. A standard field survey technique, described by Chapman and Chapman (1982) and Macdonald (1983) was used to establish the prcscncc of ottcrs at each study site. The basic unit area used during the survey was the 100 km^2 of the National Grid of Hungary. Representative survey points at all the major streams and big lakes and at most of the smaller streams and ponds were chosen. Survey points were frequently located at bridges, piers etc. that provided good access to the water. In areas with few roads or tracks, other features were chosen, such as lake inflows/outflows and stream junctions, to enable future surveyors to locate the same sites with precision. Only spraints and footprints wcrc acccptcd as proof of thc otters' presence as these are well distinguished from the similar signs of any other species. At each survey point a maximum of 600 meters of watercourse was checked. Within this distance, the search was terminated as soon as the first spraint or footprint was found and the area was considered as inhabited by otters and a score of 1 was assigned to thc sitc. On thc othcr hand, if no signs of otter were found within 600 mctcrs, thc arca was considered as uninhabited by otters, and accordingly, a score of 0 was assigned to the site.

THE ENVIRONMENTAL VARIABLES USED IN THE STUDY

At each study site (n = 369) three scalablc cnvironmcntal variablcs were recorded and scaled as shown in Tablc 1. The presence (1) or absence (0)of the following 5 non-scalable parameters was also recorded:

1. any kind of disturbance of human or natural origin (0 rated only if nonc of thc factors dctailcd below was present)
2. agricultural use of the land surrounding the study area
3. activities related to animal husbandry in the immediate vicinity of the study area (grazing etc.)
4. vicinity of towns
5. other factors (c.g. visiblc watcr pollution, blocked or temporarily dry waterway etc.).

It has to be noted here that whereas there was only one condition, namely the lack of any disturbance, when a score of 0 disturbance was recorded at a sitc, thcrc wcrc many combinations of factors (each of thcm cithcr present (1) or absent (0) at a particular site) which resulted in a general disturbancc score of I (one or more factors present).

Tab. 1 – Scalable environmental variables recorded during the survey of the distribution of otters in Hungary.

PARAMETER	SCALE
- water depth	I = 0 - 30 cm
	2 = 30 - 100 cm
	3 = over 100 cni
- density of the bank vegetation	0 = no vcgctation
	I = sparse vegetation
	2 = dense vcgctation with thin patches
	3 = dense vegetation covering large areas
- steepness of the water bank	0 = plain (@ 0")
	I = gently sloping (<45°)
	2 = steep (>45°)

LOGISTIC REGUESSION ANALYSIS OF THE EFFECT OF ENVIRONMENTAL VARIABLES ON THE OCCURRENCE OF OTTERS

There arc a variety of multivariate statistical techniques that can be used to predict a dependent variable from a set of independent variables. However, the most commonly used multivariate analysis tcchniques pose difficulties when the dependent variable can have only two values - like in our case, the presence (1) or absence (0) of otters at a particular site. In this case we cannot assume that the distribution of errors is normal, therefore the assumptions necessary for hypothesis testing in regression analysis, the most commonly used multivariate method, arc violated. Moreover, the categorical independent variables do not satisfy the assumption of multivariate normality either. This extreme case also violates the assumptions necessary for a linear discriminant analysis.

In our survey it was impossible to give an accurate, scaled estimate of the number of otters living in an area, therefore the dependent variable has only two values: 0 = otters absent, 1 = otters present. The distribution of the independent variables also differ from normal, and they can only be interpreted on an ordinal scale. For estimating the probability that an event, depending on several ordinarily scaled, categorical variables, occurs, the logistic regression model can be used. A detailed description of this technique for use on PCs as part of the SPSS+ statistical package is given by Norusis (1990). We used this technique to analyse the occurrence of otters depending on several categorical environmental variables, such as the depth of the water, steepness of the water bank, density of the bank vegetation and a variety of disturbance factors.

In logistic regression we directly estimate the probability of an event (in our case, otter living in an area) occurring. For several independent variables the model can be written as

$$Prob\ (otter) = \frac{1}{1 + e^{-z}}$$

where e is the base of the natural logarithms and z is the linear combination of the independent variables (B_1 to B_p)

The probability estimates will always be between 0 and 1, regardless of the value of z. In general, if the estimated probability of otter occurring was less than 0.5, we predicted that otters would not be present. If the probability was greater than 0.5, we predicted that otters would be present.

The logistic model can be rewritten in terms of the odds of otters occurring (i.e. the ratio of the probability that they will occur to the probability that they will not).

$$\frac{Prob\ (otter)}{Prob\ (no\ otter)} = e^{B0+B1X1+......+BpXp}$$

Then e raised to the power B_i is the factor by which the odds of finding otters change when the ,th independent variable increases by one unit. If B_i is positive this factor will be greater than 1, which means that the odds are increased; if B_i is negative the factor will be less than 1, which means that the odds are decreased. When B_i is 0 the factor equals 1, which leaves the odds unchanged.

There arc various ways to assess whether or not the model fits the data. In our study classification tables were used to compare our predictions to the observed outcomes. Histograms were also prepared to reveal the distribution of the estimated probabilities.

For the examination of the adequacy of our model we examined residuals to find outliners. The residual obtained by this method for variable i is the difference between the observed probability of the event and the predicted probability (Pred. Prob.) of the event based on the model. For example, if the model predicts the occurrence of otters to be 0.80 at a site where they actually occur, the residual is 1-0.80 = 0.20. The standardized residual *(Zi)* is the residual divided by an estimate of its standard deviation.

$$Z_i = \frac{Residual_i}{\sqrt{(Pred.Prob._i)\,(1 - Pred.\ Prob._i)}}$$

In our present study, standard residuum values greater than 2 indicated data that did not fit the prediction of the model. When such data were found, we always investigated the possible effect of factors others than the ones we used in our analysis.

RESULTS

Figure 1 shows the distribution of otter populations in Hungary. The data used to create this distribution map was then compared with predictions of the logistic regression model.

Table 2 summarizes the parameter estimates for the logistic regression model, based on the data collected during the nation-wide survey of the Hungarian otter populations.

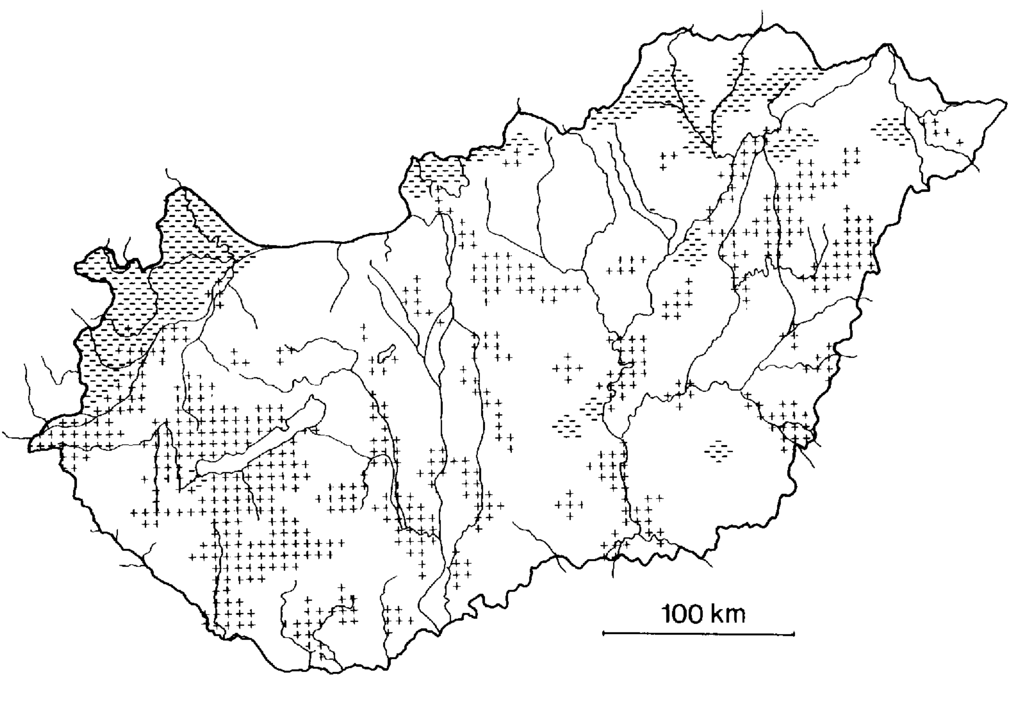

Fig. 1 – Distribution of otters in Hungary. The map was constructed by analysing the findings at 369 survey sites and marking areas where otters were found with "+" and those where otters were not found with "-".

Tab. 2 – Parameter estimates for the logistic regression model, based on the data collected during the nation-wide survey of the Hungarian otter populations.
B: logistic regression coefficients estimated from the data, S.E.: standard error for the logistic regression coefficient, Sig: level of significance (Wald statistic). **R:** partial correlation coefficient of individual variable. Exp(B): the factor by which the odds of finding otters change when the independent variable increases by one unit. For B and R positive values indicate that as the variable increases in value, so does the likelihood of finding otters at a particular site. Parameter estimates for variables with a B not significantly greater than 0 (Wald statistic, Sig > 0.05) are not shown. For *a* list of these variables see Methods and Table 1.

VARIABLE	B	S.E.	SIG	R	EXP(B)
Water depth	0.85	0.20	0.00	0.21	2.35
Vegetation	1.05	0.17	0.00	0.33	2.87
Disturbing factors					
Agriculture	0.73	0.29	0.01	0.11	2.07
Other	-3.22	0.60	0.00	-0.28	0.04
Constant	-3.15	0.49	0.00		

Given these coefficients, the linear combination of the independent environmental variables can be written as $Z = 0.85WD+1.05V+0.73A-3.22O-3.15$ (WD = water depth, **V** = density of vegetation, **A** = agricultural use of area, **O** = other, general disturbing factors (blocked watercourse etc., see Table 1) 3.15 = constant).

The positive logistic regression coefficient (B) values in Table **2** indicate that increasing water depth, density of bank vegetation and general agricultural use increase the probability of finding otters, whereas the negative B value indicates that an increase in the other disturbing factors have a detrimental effect on this probability. However, the degrees to which a unit change in these independent variables changes the odds of finding otters are not equal: the odds of finding otters are much more resistant to a unit change in the other disturbing factors (Exp(B) = 0.04) than to a unit change in any of the other factors. In other words, decreasing the density of the bank vegetation or water depth by one unit will have a much more detrimental effect on the chances of finding otters than an increase by one unit of the other disturbing factors. It is also clear, that the other variables investigated, such as the steepness of the water bank, the lack of disturbing factors in general, the vicinity of towns or activity related to animal husbandry all have only insignificant effects on the probability of finding otters at a particular site, as the logistic regression coefficients associated with them were not significantly different from 0 (Wald statistic, Sig > 0.05).

For the assessment of the goodness of the fit of the model, a classification table was used (Tab. 3). Table 3 shows that nearly **77%** of the survey sites without otters were predicted by the model not to have otters, whereas over 81% of the sites where signs of otters were found were in fact predicted by the model to have otters. Overall, the model correctly predicted the actual result in **79%** of the cases.

Tab. 3 – Assessment of the fitness of the logistic regression model on the data obtained from the national otter survey in Hungary. The table shows that, i.e., 135 sites without otters were predicted by the model not to have otters, whereas at 41 sites where the model predicted occurrence, otters were not found. Therefore, of the sites with no signs of otter, 76.70% were correctly classified by the model.

		Predicted occurrence (number of cases when otters were found)		Percent correct
		0	1	
Observed occurrence	0	135	41	76.70%
	1	36	157	81.35%

This classification table, however, does not reveal the distribution of estimated probabilities for sites in the two groups (i.e. where otters were and where were not found). For each predicted group, Tab. 3 only shows whether or not the estimated probability is greater than 0.5. This is why it is necessary to show the histogram of estimated probabilities as well (Fig. 2). If we have a model that successfully distinguishes the two groups, the cases for whom the event has occurred (signs of otters have been found) should be to the right of 0.5, while those cases who have not had the event (no signs of otter have been found) should be to the left of 0.5. The more the two groups cluster at their respective ends of the plot, the better. This was indeed the case, as the majority of cases where signs of otters were not found, clustered between the 0 and 50% probability values of finding them, whereas the majority of the cases where signs of otters were found, clustered between the 50 and 100% probability values predicted by the model.

A: signs of otter not found

B: signs of otter found

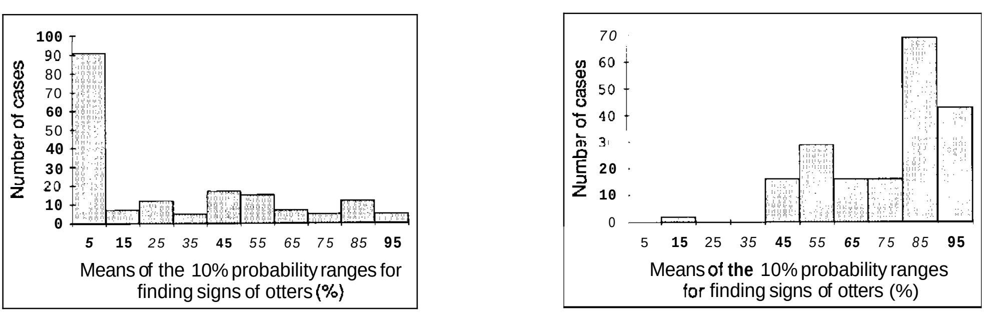

Fig. 2 – Estimated probabilities of finding signs of otters at sites where such signs were not found (A) and where they were actually found (B).

Finally, to examine the adequacy of our model, we looked at the residual, which is the difference between the predicted and observed probabilities of the occurrence of otters. By using this diagnostic method the sites for which the model does not fit well can be identified. We found that in the overwhelming majority of the cases the model did fit the data well. However, we found four cases where despite the 100% negative prediction of the model, the presence of otter was observed. In one of these cases a single spraint was found near to a village, where the bank vegetation was sparse and the water seemed to be very polluted. However, this site could have been visited only once by chance, most probably by a young individual, but according to the criteria of the survey, the place had to be recorded as inhabited by otters. At two other sites the fields were freshly fertilized and we actually found traces of the fertilizer in the water, therefore the predicted probability of finding otters at this place was 0. In these cases the otter spraints had probably been deposited before the fertilization. The fourth place was a trout farm. Here again, all the circumstances were against the presence of otters, i.e. the concrete waterbanks surrounded by fence were considered by the model as unsuitable for otters. However, the survey showed that otters regularly visit the place, probably attracted by the easy fish prey.

We found six sites where according to the model, the presence of otters had a 100% likelihood, yet no signs of otters were discovered. However, four of these six places proved to be situated at waterways that were identified as the most polluted waters in Hungary. The remaining two were situated by the river Tisza, where only the waterbanks were searched but the many small islands in the river were not.

DISCUSSION

According to the logistic regression model, the depth of the water and the density of the bank vegetation are the two factors (of the 8 considered in present study) that have the greatest influence on the occurrence of otters at particular sites. This was also the conclusion from previous studies focusing on either the effect of the depth of water (Glimmerween and Ouwerkerk, 1984) or the effect of the density of the bank vegetation (Macdonald and Mason, 1985) and our multivariate analysis confirmed these observations. However, a surprising new finding of the present work is that cultivation of the land around the aquatic habitat appears to have a positive effect on the occurrence of otters. In our opinion, this is due to the fact that in such areas the irrigation canals and natural watercourses are kept in a good condition in order to supply sufficient water for irrigation purposes. Although in these canals the availability of food may not be very good, they can provide connections between larger water courses, and therefore are frequently used by otters. Also, apart from short-term pollution waves caused spraying pesticides on the crop, etc., in these rural areas there are no chemical and other industrial plants which would be the source of a more acute pollution. Another interesting prediction from the analysis is that the presence of various disturbance factors, such as the vicinity of towns, or even household rubbish floating in the water have no or very little effect on the occurrence of otters. More importantly, if we compare the degree by which a unit change in the water depth or vegetation density and in the disturbing factors change the odds of finding otters, we see that the former two have a much more profound effect. This means that otter populations are surprisingly tolerant to disturbance caused by human activity but extremely sensitive to the deterioration of the vegetation or reduction of water depth. From a practical point of view this means that it is not necessary to create isolated safe havens for otters trying to exclude all human disturbance but it is absolutely essential to try to preserve the bank vegetation and ensure that the water depth in aquatic habitats is sufficient to support otters. A dense bank vegetation provides cover for the otters, whereas the depth of water is mainly important for maintaining good prey animal populations.

If we look at the predictive value of the model, we can conclude that it is very good, because it correctly predicts the presence or absence of otters in around 80% of the cases. The estimated probability of occurrence histograms revealed, that in the vast majority of the cases where the model predicted a low probability of occurrence (between 0 and 50%) no signs of otters were found, whereas in the majority of cases where the predicted probability of occurrence was high (50 to 100 %), signs of otters were indeed found. It is becoming increasingly important to predict the Occurrence of otters by habitat characteristics alone (see Dubuc et al., 1990) as an alternative to laborious and sometimes impossible field surveys. Areas with an optimal set of environmental variables can then be explored in detail and their suitability as otter habitats established. This method may have an importance in reintroduction projects as well as in the protection of already existing otter habitats.

Since Mason (1989) named pollution of the water as the single most limiting factor for otter populations, where the habitats seemed to be physically suitable but no otters were found, chemical pollution must have played a part. This was borne out in most cases (4 out of 6) when we checked water pollution data retrospectively. This emphasizes that results from field surveys should always be compared with water quality data when trying to build models on the effect of environmental variables. The problem is more complex when signs of otters are found at apparently uninhabitable sites. The most common cause of this "error" is a recent rapid change in the quality of the environment. In sheltered places, such as under bridges, it is often difficult to estimate the time when a footprint was made by a passing otter or when a spraint was deposited; the environment may have changed between this and the time of the survey. However, this kind of discrepancy was only found in four cases out of 369.

On the whole we conclude that the logistic regression analysis is an extremely useful tool in assessing the effects of multiple environmental factors and find the variables that are most likely to affect the otter populations. If the effect is positive, the optimum state of that variable should be preserved, if it is negative, it should be eliminated or reduced. This method will help conservationists find the best ways of maintaining suitable habitats for otters and other endangered species.

REFERENCES

CHAPMAN, P.J. & L.L. CHAPMAN. 1982. Otter survey of Ireland. The Vincent Wildlife Trust report, London, 40 pp.

DUBUC, L. J., KRHON W.B. & R.B. OWEN. 1990. Predicting occurrence of river otters (*Lutra canadensis*) by habitat on Mount Desert Island, Maine. J. Wildl. Manag. , 54: 594-599.

GLIMMERVEEN, U. & E.J. OUWERKERK. 1984. Habitat use of marine otters; an analysis of otter activity and its relation to environmental factors in a coastal area on Shetland. Inst. of Therrestr. Ecol., Banchory, Scotland, 44 pp.

KEMENES, I. 1991. Otter distribution, status and conservation problems in Hungary. Otters; IUCN OSG Bull., 6: 20-24.

MACDONALD, S.M. 1983. The status of the otter (*Lutra lutra*) in the British Isles. Mammal Rev., 13: 11-23.

MACDONALD, S.M. & C.F. MASON. 1985. Otters, their habitat and conservation in northeast Greece. Biol. Conserv., 31: 191-210.

MASON, C.F., 1989. Water pollution and otter distribution: a review. Lutra, 32: 97-131.

NORUSIS, M. J. 1990. SPSS/PC+ Advanced Statistics TM 4.0 for the IBM PC/XT/AT and Ps/2. SPSS Inc. Chicago. 427pp.