INSUBRIA UNIVERSITY

PhD School Analysis, Protection and Management of Biodiversity

XXIII Cycle

Data infrastructures and spatial models for biodiversity assessment and analysis: applications to vertebrate communities

Tutor: Prof. Guido TOSI

PhD student:
Monica CARRO

2010

Contents

Abstract 1

1. Introduction 7

2. Study areas 15

2.1. Brescia plain (small spatial scale) 15

2.2. Mount Meru Region - Tanzania (medium scale) 18

2.2.1. Climate 19

2.2.2. Geology and soil 20

2.2.3. Water sources 20

2.2.4. Vegetation and land cover characteristics 21

2.2.5. Agriculture and livestock keeping 26

2.3. Lombardy Region (large spatial scale) 26

2.3.1. Geology, Geomorphology and Hydrography 27

2.3.2. Vegetation and land cover characteristics 28

2.3.3. Climate 29

2.3.4. Demography 30

3. Material and methodology 33
3.1. Datawarehouse for wildlife data handling . . . 33
3.1.1. Spatial Database and Database Management System (DBMS) . . 33
3.1.2. Web-database applications . . . 34
3.2. Statistical, mathematical and GIS methods for wildlife distribution assessment . . . 35
3.2.1. Generalized Linear Models (GLM) . . . 35
3.2.2. Model Evaluation and Classification . . . 40
3.3. Software tools . . . 43
3.3.1. PostgreSQL and PostGIS . . . 44
3.3.2. GRASS . . . 45
3.3.3. R . . . 46
3.3.4. LAMP system . . . 47
3.4. Automated mapping process: SOS . . . 47

4. Datasets 53
4.1. Brescia plain . . . 53
4.2. Mount Meru region - Tanzania . . . 60
4.3. Lombardy region . . . 67

5. Results 73
5.1. Spatial database . . . 73
5.2. Spatial models . . . 77
5.2.1. Brescia Plain (small scale) . . . 77
5.2.2. Mount Meru region - Tanzania (medium scale) . . . 79
5.2.3. Lombardy region (big scale) . . . 89

6. Conclusion 95

Acknowledgements 99

Bibliography **100**

A. Results of LRA **111**

A.0.4. Mount Meru region - Tanzania 111

A.0.5. Lombardy region 176

B. SOS scripts **243**

B.0.6. SOS scripts for statistical and spatial modeling 243

B.0.7. SOS scripts for geoprocessing 273

Abstract

In conservation biology the computation of biodiversity maps, based on statistical models is a central concern. These maps, produced with objective and repeatable methods are an essential tool for conservation and monitoring programs as well as for landuse planning.
Since the computation of biodiversity maps requires complex and time consuming procedures for data processing and analysis, it is necessary to design methods for homogeneous, scalable and repeatable data management and analysis.
Moreover, the huge volume of data used in ecological modelling requires suitable software architectures to store, analyze, retrieve and distribute information in order to support research and management actions in due time.

First of all we developed an analysis system (SOS - Species Open Spreader) providing statistical and mathematical models to predict species distribution in relation to a set of predictive environmental and geographical variables
The system is composed of a module for data input/output toward and from the GIS and of a package of scripts for the application of different modelling techniques. At present, three statistical techniques are integrated in SOS: Logistic Regression Analysis (LRA), Environmental Niche Factor Analysis (ENFA) and flexible Discriminant Analysis with method BRUTO. Furthermore, two empirical spatial methods of analysis are available within SOS: Habitat Suitability Index (HSI) and Spatial Overlay.
The system is designed to work with the GIS (Geographical Information System) soft-

ware GRASS and the statistical environment R, coupled together through the SP-GRASS6 library. Three different outputs are expected: text and graphical outputs with statistical results and suitability maps.

Second, we tested the use of spatial Database Management Systems (Spatial DBMS) to handle wildlife and socio-economic data and we developed a web database application to provide facilities for database access. The information system was built for the Meru district (Tanzania) in the context of an Italian cooperation project of land use planning in Maasai rural areas.
We tested two different solutions: SpatiaLite and PostgreSQL-PostGIS; they both offer advanced technical facilities and spatial extensions to analyze spatial data. SpatiaLite is a new solution and offers the main advantages to consist of a unique file and to present a user-friendly interface, which make it the best solution for many applications. in spite of this we used PostgreSQL-PostGIS since it represents a well-established information system supported by libraries for web applications development.

We applied SOS to three case studies at different spatial scale: Brescia plain (small scale), Mount Meru region - Tanzania (medium scale) and Lombardy region (big scale) in order to produce maps of species potential distribution and biodiversity maps for planning and management.
We applied logistic regression analyses to compute models and ROC analysis for classification performance evaluation. The automation of processes through SOS gave us the possibility to build models for a large number of vertebrate species. The analysis produced very reliable results at middle and big scale while regression methods did not converge at small scale. This is probably due to habitat homogeneity and to the use of environmental variables with an insufficient level of detail.
The potential distribution and biodiversity maps produced also had in all cases an applicative use in fact we used mammal species models computed for Mt. Meru region to

produce a map of biodiversity within the area: this map represents an informative base for land use planning at village level within a cooperation project for Maasai economic development and environmental redemption.

Amphibians and reptiles models, computed for Lombardy, represent a good informative base for planning management actions in the region.

1. Introduction

Species distribution models are becoming an essential tool in Ecology and Environmental Management research. These predictive modelling techniques are common in numerous areas such as Biogeography and Evolution, Conservation Biology and Climate Change Research.

In a recent review by Guisan and Thuiller (2005) the history of species distribution models is revisited and described as having three major phases. First of all, models relied on statistical quantification of species-environment relationship based on empirical data. Then an expert-based spatial modelling phase occurred, without statistical or empirical treatment. In the meantime, the advancements and diffusion of Geographic Information Systems (GIS) has supported spatial analysis in ecological studies and finally a spatially explicit statistical and empirical modelling of species distribution framework has been defined. The development on predictive modelling was largely supported by the advances and interralation of statistics, information and computational science, improving the understanding of complex ecological systems.

As for most habitats and taxonomic groups detailed species distribution maps are unavailable, the prediction of species spatial distribution is a central concern in ecology. Moreover, policies for preserving global species richness and the assessments on species conservation status depend on the knowledge of temporal and spatial distributions (IUCN, 2001; Green et al., 2005). The predictive modelling importance comes from the ability to estimate species distribution (species data) by relating it with some environmental predictors (Elith et al., 2006; Araujo and Guisan, 2006; Meynard and Quinn, 2007) and producing maps of potentially suitable habitat distributions (Brotons et al., 2004; Olivier and Wotherspoon, 2006; Rodriguez et al., 2007).

In the last decade the availability of spatial data ready to be used as predictors of species distribution has dramatically increased (e.g. remotely sensed data, digital elevation models, etc). Environmental predictors are any environmental variable from Geographic Information System (GIS) or other mapped data; in Meynard and Quinn (2007) they appear divided in three categories:

1. resource variables - describe consumed matter or energy;
2. direct gradients - having a direct relationship to the species physiology;
3. indirect gradients - having a strong relationship with other direct gradients or resource variables and are easily measured.

Species data are field observations, therefore they rely on several issues that may affect the quality of the data:

- species biology: organism characteristics may diminish species detectability (e.g. mobility, inconspicuousness) or determine different species prevalence.
- Planning of the census: monitoring efforts are limited in space, time, taxonomic coverage, and are altered by the variability among observers and habitat types.
- Data storage: species data usually have errors and biases due to an unsystematic method of manner of accumulating samples; for instance, data collected at different geographic scales.

Researchers often have to work with sample data that result from complete lack of standardization in surveying and in data storage (Zaniewski et al., 2002; Elith et al., 2006; Rodriguez et al., 2007; Meynard and Quinn, 2007) and so the main issue affecting model performance seems to be species data quality (Zaniewski et al., 2002; Stockwell and Peterson, 2002; Brotons et al., 2004). Moreover, the modelling process will begin with the perspective of low prevalence of species and lower performance of the models.

There are, in broad a sense, two groups of methods to generate habitat suitability maps (Brotons et al., 2004; Tsoar et al., 2007). Apart from obvious differences in the statistical and computational backgrounds, they differ in the kind of data they require. The first group of methods needs presence-absence data to generate predictions. This group includes popular statistical approaches such as Generalized Linear Models (GLM), Generalized Additive Models (GAM), Classification and Regression Tree analysis as well as Artificial Neural Networks (ANN).

The other set of methods includes Ecological Niche Factor Analysis (ENFA; Hirzel et al. 2001), Environmental envelopes (BIOCLIM, DOMAIN, HABITAT; Walker and Cocks 1991), Genetic Algorithms (GARP; Stockwell and Peters 1999) and it uses sets of observed occurrences without sites of observed absences (hereafter called presence-only data). The latter methods were developed in order to allow the creation of distribution maps as well, even if starting from "incomplete data", that either had inadequate absence information or had none. In fact, as often stressed in the literature (Hirzel et al., 2002; Engler et al., 2004; Elith et al., 2006; Gibson et al., 2007), is that absence data may not be inferred certainly. This is particularly true for rare species, for those species which do not occupy the entire suitable area available to them and, as said before, for species which are difficult to detect.

Methods that use presence-only data usually do not surpass the performance of techniques that employ higher quality (presence-absence) data, but they have been found to model particularly well those data sets (Zaniewski et al., 2002; Brotons et al., 2004; Engler et al., 2004).

Thus, some modelling techniques, such as GLM and GAM, are being improved in order to use pseudo-absences data (i.e. random generated points) for model construction (Engler et al., 2004; Olivier and Wotherspoon, 2006; Gibson et al., 2007; Tsoar et al., 2007). These enhanced techniques are classified as presence-only methods too, because there is no use of real absence data. The way in which pseudo-absences are generated is particularly important because it may exert influence upon the final quality of the models (Zaniewski et al., 2002).

The simpler and quickest way to generate pseudo absences is to choose them totally at random over the entire study area (Stockwell and Peterson, 2002). It is important to remark that, however, this approach can generate absences in areas that correspond, in fact, to suitable areas.

For this reason other methods have been tested to reduce this problem. Gibson et al. (2007) uses GLM with case-weighting to reduce the effective sample size of randomly selected pseudo-absences for modelling a rare parrot species in Western Australia. Another way was proposed by Zaniewski et al. (2002), for instance, in the case study of New Zealand ferns, by choosing the absences using an environmental weighted random sample. They have created a habitat suitability map for all the fern species using GAM with totally random pseudoabsences; only after that another set of generated absences is selected, these ones being proportional to the predictions made by the first habitat suitability map fitting GAM models. A very interesting proposal is the one by Engler et al. (2004) which combines the strengths of ENFA and GLM; just like the one before, this is a two-step procedure but the first habitat suitability map is made with ENFA.

The model evaluation phase is extremely important to assess the accuracy of the predictions. Two types of prediction errors can be detected from presence-absence models: false positives (type I) and false negatives (type II, Fielding and Bell, 1997; Ottaviani et al., 2004). False positive errors are unavoidable, as they occur when the model predicts a suitable habitat where the target organism does not exists, and not all suitable habitats are occupied or reported for the said organism. False negatives correspond to points predicted as unsuitable habitat while the species indeed exist (Ottaviani et al., 2004). These errors may be caused by incomplete surveys or scarsely detectable species.

As discussed before, the reliability of absences data depends on several factors; moreover most of the available data banks have vague location details. Presence-absence models and presence-only models can provide the same kind of outputs (habitat suitability maps), however usually they cannot use the same evaluation method. This is one of the reasons for enhanced presence-absence models development, because presence only methods cannot contrast predictions with absence locations (Zaniewski et al., 2002; Hirzel et al., 2006).

In this research work, the modelling process uses a typical presence-absence method fitted with pseudo-absences, the Generalized Linear Model and specifically Logistic Regression Analysis (LRA). Modelling was applied to three different case studies characterized by different scale and different species datasets.
The first one is a small scale case study and is applied to an area in the Brescia plain (Northern Italy) characterized by deep habitat modifications; the area is investigated through monitoring programs in order to assess environmental impacts. Species data are represented by field data for different taxa (Amphibians, Birds and Mammals).
The second case study (medium spatial scale) is located in the North of Tanzania, in the Mount Meru ecosystem, that comprises mountain forest covering volcanic cones and semi-arid vegetation. Species data came from field work and refer to Mammals; models are used for land use planning within a cooperation project for sustainable development.
The third is a big spatial scale case study regarding the whole Lombardy region (Northern Italy) and is targeted to model the distribution of Reptiles and Amphibians using data from distribution atlases.

The application of analytical techniques for wildlife spatial modelling is a consequence of the present availability of large datasets that contribute to improve the explanatory power of ecological complexity models and thus ecosystem management. On the other hand this huge and heterogeneous volume of data brings the challenge of managing and analyzing huge volumes of data (Cagnacci et al., 2008). This drives the conceptual definition of a suitable software architecture that can be developed with specific tools on different platforms; the main requirements and needs are summarized in Urbano et al. (2008):

1. data scalability: to handle large amount of data consistently, persistent and very large data storage capability is needed.
2. Long-term storage for data reuse: data must be consistently stored in the long term, independently from a specific application, to permit data reuse for different studies.
3. Periodic and automatic data acquisition: this requires automated procedures to receive, review and store data from different inputs.
4. Efficient data retrieval: fast data search and retrieval tools are needed to support efficient data analysis and management.
5. Management of spatial information: spatial data require retrieval, manipolation and management tools specific to spatial domains.
6. Global spatial and time reference: they are needed to handle studies with regional or global perspectives.
7. Heterogeneity of applications: it requires a software architecture that supports the integration of different software tools for specific task-oriented applications.
8. Easy implementation of new algorithms: it is imporant to support implementation and customization of new algorithms to test, or even apply new analytical techniques.
9. Integration of different data sources: spatial and non spatial datasets should be correctly managed and efficiently integrated into a comprehensive data structure.

This allows the correct analysis of data derived from different sources: remote sensing, environmental and socio-economic databases, wildlife-related data etc.

10. Multi-user support: several users might need to access data simultaneously, both locally and remotely, with different access privileges (Wong et al., 2007).
11. Data sharing: this requires adherence to standard data formats, definition of metadata and methods for data storage and management that, in turn, guarantee interoperability.
12. Data dissemination: this requires the integration of specific tools to make data accessible (e.g. Data Web interfaces and Web-GIS tools).
13. Cost-effectiveness: the cost-effectiveness of software tools is an important accessibility factor for institutions with limited financial resources that can be applied to production and analysis of data instead of data handling.

Advanced information systems currently developed to manage wildlife data are based on relational or object-relational database management system (DBMS), with dedicated spatial tools (Urbano et al., 2010). From the technical point of view DBMS include software architectures and tools that completely meet the requirement of DBMS for wildlife ecological studies. Technical facilities include: storage capacity, backup and recovery, data integrity and consistency, automation of processes, data retrieval performance, reduced data redundancy, client/server architecture, multi-user environment, data security, and standards compliance.

In addition DBMS are increasingly provided with spatial extensions (spatial DBMS): this gives the opportunity to store and manipulate native data types and spatial reference systems and to integrate spatial objects with standard DBMS data types. Moreover, in a spatial database, geometric data are effectively manipulated with spatial indexes and spatial extensions of the standard structured query language.

Finally the cost-effectiveness requirement can be fulfilled using open source software that includes spatial DBMS, libraries, desktop GIS, Web database packeges and Web-GIS packages.

In this study we tested the use of Spatial database Management Systems (DBMS) to handle wildlife data. The application aimed the creation of an informative system to store, analyze and retrieve spatial and non spatial data, addressed to Institutions for socioeconomical and conservation planning.

This research study was focused on three major points:

1. applicability of Generalized Linear Models at different scales;
2. automatization of statistical analysis and spatial processing;
3. use of spatial DBMS to handle wildlife data.

To understand to which extent a chosen model could give reliable species spatial predictions at different scales was a objective of this study. The goal of producing biodiversity maps is to detect their possible use as decision support systems on conservation and management actions. We aimed at seeing whether it was possible to perform the modelling processes in a expeditious manner by starting from existing datasets and available GIS variables, in order to produce outputs readily exploitable for management purposes. Finally another goal was to find solutions for data warehousing and data processing in order to standardize data storage and to automate computational processes.

2. Study areas

2.1. Brescia plain (small spatial scale)

The study area corresponds to a 3 km buffer along the planned 29 km highway Ospitaletto-Montichiari, South-West of Brescia. The highway will connect the A4 highway, at Ospitaletto resort, with the airport of Brescia-Montichiari; it is an important infrastructure for transportation, serving as belt-way to Brescia city, that partly follows the existing roadway Provinciale 19. The construction of the highway started in September 2008 and will be finished at the end of 2012.
The study area extends for 87 km^2, is completely flat (about 50 m a.s.l) and falls within the Brescia plain, known as Bassa Bresciana. Economy is based on agriculture, mainly single-crop farming (maize) even though industry is extensively present.

The climate is typical of the Po Plain, with long and humid summer and cold and foggy winter. Rainfall is abundant and concentrate in the months of March-May and October-November. Winter usually lasts from November to the end of March and is characterized by short rainfalls; snowfalls can occur.

The Brescia plain is crossed by numerous streams and artificial channels for irrigation; Mella and Garza are the main rivers in the study area. The Mella river is 96 km long,

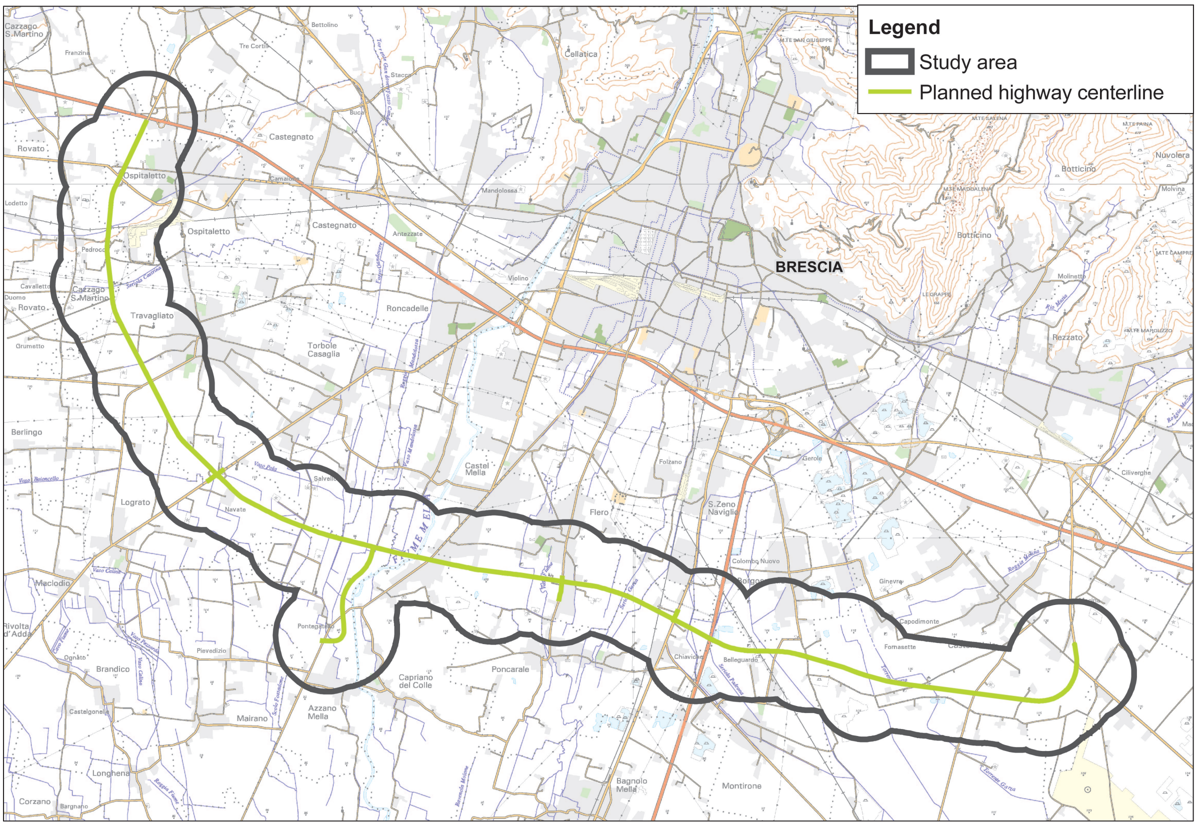


Figure 2.1.: Study area in the Brescia plain corresponding to a 3 km buffer along the projected highway Ospitaletto-Montichiari

it rises at the Maniva pass, run through the Trompia Valley and the city of Brescia and flows into the Oglio river in the Cremona province. The Mella river is an important resource for irrigation, since it provides water through numerous channels. Unfortunately its water, starting from Brescia, is characterized by high levels of pollution. The Garza torrent rises between Lumezzane and Agnosine, runs through the Bertone Valley, the Garza Valley, the low Trompia Valley and the city of Brescia, partly open and partly underground; then it flows in the plain of southern Brescia and then gets lost in Ghedi. A second branch of the river, called Garzetta, comes from Brescia and flows into the Mella. The area is characterized by resurgences, typical water springs of the Po plain, with characteristic flora and fauna. Water has a constant temperature of 9-10° C in winter and 12-14° C in summer and is widely used for irrigation. The main resourgences in the area are: Vaso Mandolossa, Vaso Orso and Vaso Fiume. All these rivers are characterized by physical, chemical and biotic modifications, underlined by surveys (Istituto Oikos, 2008). In general the quality index for rivers as evaluated by the Fluvial Functionality Index - IFF is of poor quality (Istituto Oikos, 2008) because of the characteristic of the Brescia plain, with extended cultivations and widespread, even if limited, urban areas that decrease river banks quality. There are also some riverlines treats enclosed between artificial banks with consequent deep habitat modifications. The aquatic habitat of all the rivers in the area is seriously degraded with negative consequences on the stability and balance of aquatic biocoenoses. Some of the rivers are over-exploited for irrigation and thus completely dry in summer. Waters are characterized by a quite elevated degree of pollution: as a matter of fact the assessment based on the Extended Biotic Index ranges between "polluted environment" to "deeply polluted environment". (Istituto Oikos, 2008) Aquatic vegetation was found only in rivers characterized by the presence of water throughout the whole year; species are typical of rivers in flat lands with slow water flows, although in some cases species indicating the presence of organic pollution have been found.

The study area is mainly interested by monocultures and industrial settlements with limited natural areas along rivers. In such a compromised landscape, hedge rows between fields, small woods and single trees play an important role in the ecosystem. A study on vegetation in the area (Oikos, 2008) registered some valuable tree entities, mainly single trees with, only a few groups of trees.
The species with highest number of samples is *Quercus robur*. This species, once common, is now very rare in plain woods but persists with isolated specimen of big dimension at roads and water channels margins. *Populus nigra* is the second species represented in the "valuable" samples, *Ulmus minor* is the most represented autochthonous species and is always present in good vegetative conditions. *Celtis australis* and *Alnus glutinosa*

were also registered with a few samples. The phytosociological analysis of the arboreal layer underlined the dominance of *Platanus hybrida*, cultivated and then made wild in the whole Italian territory. The area does not show peculiarities in vegetation, apart from the presence of communities of the order *Fagetalia sylvaticae*, which are relicts of ancient plain woods and of the order *Alnetalia glutinosae* and *Salicion albae*, typical units of wetland. The rest of vegetation is characterized by common and exotic species such as *Robinia pseudoacacia* that originally came from North America.

2.2. Mount Meru Region - Tanzania (medium scale)

Mount Meru Ecosystem is located in Arusha region, northern Tanzania. The region is well known for its unique wildlife resources and serves as the center for tourists in the northern circuit. This study was conducted on an area that belongs to different administrations:

- the Mount Meru Ecosystem, partly inside and partly outside Arusha National Park, covering the northern part of Arumeru district,
- the eastern part of Longido district,
- a small portion of Siha district of Kilimanjaro region.

The study area (Figure 2.2) covers a surface of 784 km^2 and, excepted Arusha National Park, a large part of it falls in village communal land where people and wildlife coexist. The entire study area is a plain lying trough between three important high land marks. On the East side is the well known volcanic mountain Kilimanjaro (5895 m a.s.l), on the South-west Mount Meru, a recent volcano (4566 m a.s.l), and on the north Mount Longido (2629 m a.s.l.). It borders Amboseli National Park to the further North at the border between Tanzania and Kenya. Thus the area is basically considered as part of

the much broader Meru-Kilimanjaro and Amboseli ecosystems.
At the center of the study area is located the village of Uwiro, which is particularly important because the cartographic outputs of this study will be used as an instrument supporting decision processes in landuse planning for this village.

2.2.1. Climate

The main factors for climate determination have been identified as rainfall and temperature. Mount Meru ecosystem has a bimodal rainfall type, with a marked dry season. The short rains start in November-December and the long rains occur in March-May with a peak in April. There is a prolonged dry period from June to October, and a short dry season in January-February, which does not occur in all years. The short rains differ both in intensity and distribution from one year to another, but the long rains are more regular. The beginning of the rainy seasons is variable; sometimes the short rains start in October and end in January.
The higher altitude slopes receive an average of 800-1500 mm of rainfall per year; these comprise the southern, western and southeastern slopes of Mount Meru (Hedberg, 1951 cited in Bolick, 1974). On the northeastern and northern slopes, in the low-lying trough where the study was conducted, the rainfall is highly affected by the presence of the two high mountains Meru and Kilimanjaro. Mount Meru has higher rainfall on its southern, southeastern and western slopes, whereas the northern slope is on the leeward side, hence with low rainfall. Moreover, Kilimanjaro has higher rainfall on the eastern side while the western side, where the study area lies, receives much less water. The northern Mount Meru area is semi arid and arid, receiving an annual average rainfall of 300-600 mm. Although the rainfall amount is very low in the area, its intensity is high resulting into considerable erosion in several zones, forming deep-cutting gullies (Kidane, 1974).
There is no much data recorded on temperature, but data from Arusha regional integrated development plan of 1981 show that the average minimum temperature is 15-18° C and maximum average temperature is 27-30° C .

2.2.2. Geology and soil

The geological characteristics of the study area are highly influenced by the volcanic activities of the Great East African Rift Valley. The high numbers of volcanic activities within the Rift Valley resulted into high numbers of volcanic cones.

The distribution of soil types is strongly related to the geology and terrain of the area. Soil drainage varies from well drained to poorly drained.

On higher slopes near Mount Meru, soils are typically volcanic in origin and derived from the volcanic rocks and ashes hurled during the eruption of Meru and Kilimanjaro (Lasan, 1971; Bolick, 1974). With little rainfall and scanty vegetation cover, the weathering process of the volcanic rocks has been slow, and, as a result, a large part is characterized by shallow soils, and in some areas bare rocks dominate the ground. In the swamps and depressions soils are alkaline in nature dominated by leached soluble materials being transported from higher slopes. (Kidane, 1974; Kiunsi, 1993).

2.2.3. Water sources

Water availability is the main problem in the whole area except the part adjacent to Arusha National Park boundary that is close to Mount Meru slopes. In the past two main rivers, Ngare Nanyuki and Ngare Nairobi flowed all over from the Meru-Kilimanjaro basin, that is a part of the great Pangani basin, to the Amboseli basin.

With high water demands from an increasing population and change in land use in which the majority is now adopting small-scale irrigation farming in the upstream area, the flow of the two rivers has been reduced dramatically. Ngare Nanyuki river no longer flows further north. Also the other important river, Ngare Nairobi, that once was flowing up to Ngasurai hill and to Nesiwandeti plains, by 1997 had only sporadic flow to Tinga Tinga village (Poole & Reuling, 1997). The highly reduced water flow today in both rivers is due to high demand of agriculture activities upstream.

An increase of irrigation in the villages of Olkung'wado, Uwiro and the newly emerged irrigated fields at Ngabobo village contributeed to the high reduction of the water volume, and currently the flow of Ngare Nanyuki is only up to Ngabobo village and within the NARCO ranch where some people have encroached it for farming.

Sections of Mkuru, Engutukoit, and Losinoni villages experienced severe water problems with no river flowing into their lands. They totally depend on seasonal rivers that are only available in the wet season. This problem of water scarcity has adversely impacted people's life and properties.

With the cut off of the Maasai furrow water flow, there has been an increase in human-wildlife conflicts that in most cases are caused by animals searching for water. This applies particularly to elephants. Severe conflicts occur repeatedly at Ngereiyani and Tinga Tinga, particularly during the dry season when no water is available in the plains. Nowadays much fewer species are seen in the plains during the dry season, as most of them move close to water points of Amboseli, resulting in high herbivore pressure in the ecosystem (Poole & Reuling, 1997). Other species move into bush lands in areas of Mkuru, Uwiro and Ngabobo.

Other water sources available in the area are boreholes, that are very few and scattered in the area, man made dams, artificial ponds, springs, and water that is retained in the big gullies after rains. Other seasonal swamps and man made dams remain the major source for both wildlife and livestock in the whole area, particularly in the Maasai lands of Losinoni, Mkuru and Engutukuit.

2.2.4. Vegetation and land cover characteristics

The East African vegetation is highly connected to the local geological formations. Most of East Africa is covered by savannah and scrubs of different types, while the mountain zones harbour a series of different vegetation zones (Kiunsi, 1993).

A brief description of the vegetation of the study area has been derived from work done by Istituto Oikos in the preparation of the Mount Meru Conservation Project Land cover (MMCP, 2004). Additional information was added from Kidane (1974), and

Kiunsi (1993).

The most common vegetation types of the study area are closed forest, wooded shrub land, bushed grassland, open grass land-shrub savannah, and swamps.

Within Mount Meru ecosystem, the mountain forest is found in areas covering sections of Arusha National Park in the high altitudes from 1700 to 3300 m a.s.l (Hecky, 1971, cited in Bolick, 1974). The mountain forest is well developed on the southern, western, and southeastern slopes and, to a lesser extent, on the higher slope in the North.

The lower section of the park is covered by secondary forest vegetation or dry mountain forest as explained by Vesey-Fitzgerald (1974). Closed forest of typical dry mountain forest types is found dominating the high altitudes of Longido Mountain, serving as the water catchment reservoir.

The remaining part of the area is covered by semi arid vegetation, the typical vegetation type of the study area (Kiunsi, 1993). Vegetation cover is poor, dominated by shrubs, thickets, open woodlands and grassland of typical savannah lands. Bushes are the most dominating land cover, which is characterized by scattered trees and shrubs. Dominant species include *Acacia mellifera*, *Acacia nubica* and *Sericocomopsis hildebrandtii*, *Acacia drepanolobium* and some *Acacia tortilis*.

The wooded shrub land type is also known by others as wooded grassland (Kidane, 1974). In this type a mixture of trees and shrubs is found. In most cases grass cover is poor, and the main dominant tree species are *Acacia tortilis*, *Acacia drepanolobium*, *Salvadora persica*, *Balanites aegyptiaca* and *Commiphora spp.* It is common on the foot slopes of mountains, both in the northern Meru and Longido sections.

The open grassland-shrub savannah is defined by its open grassland and low tree cover. It is a mixture of different vegetation types: sparse trees, shrubs, and open grasses that constitute a large part of it. This vegetation type is dominant all over the area. The plains characterized by this vegetation are always overstocked. In the wet season they are the major grazing and calving points for wildlife, particularly grazers. The shallow volcanic soil favours the growth of the grasses.

Swamps and seasonally flooded areas are scattered in the zone, and can be found in many different places in the village lands where the terrain allows water to settle.

The main permanent swamps are found at Olkung'wado and Uwiro villages. Other seasonal swamps can be found all over the place and they are always used as grazing areas in the dry season for most of the livestock.

Swamps and riverine vegetation is much more pronounced at Olkung'wado and Uwiro villages. These swamps are permanently wet throughout the year and several springs are found in the swamps that supply water for both people and livestock.

The main dominating species along the river are *Acacia xanthophloea* and *Ficus spp.* Irrigated crop fields are found scattered all along the river.

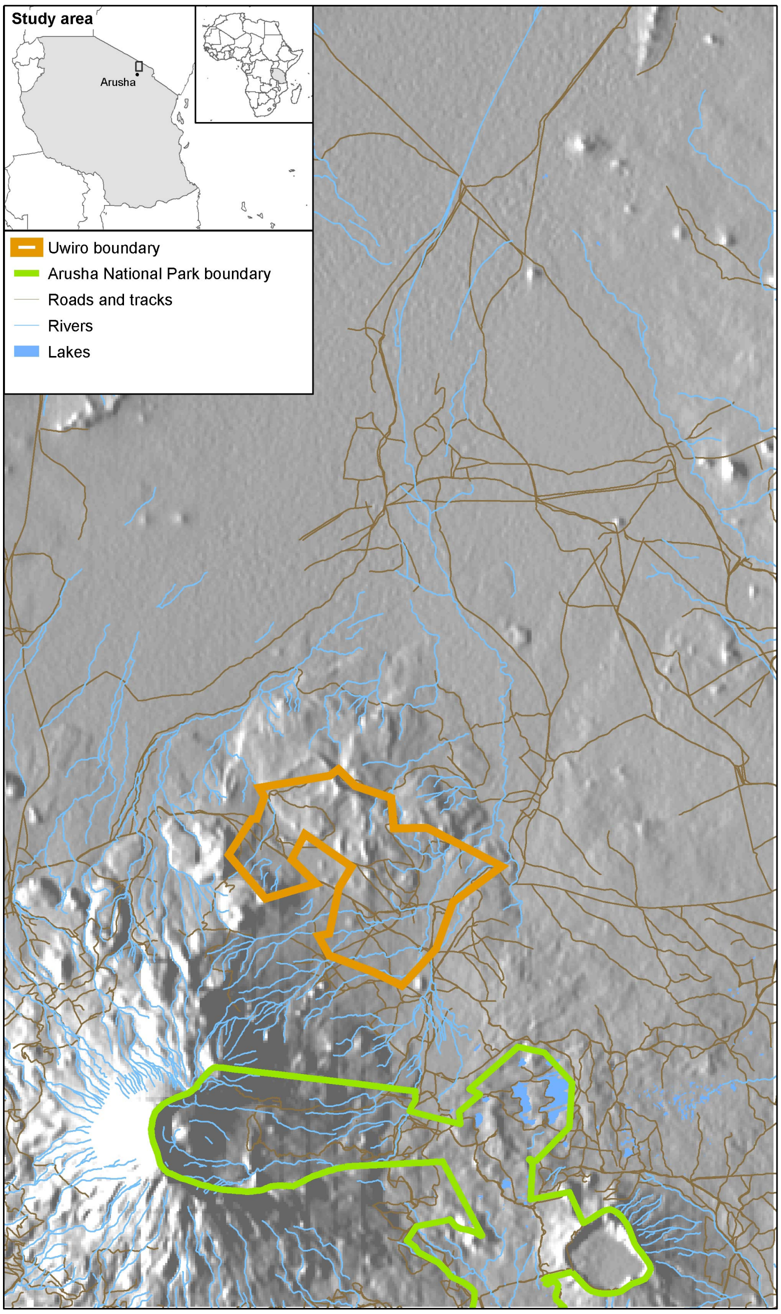


Figure 2.2.: Mount Meru region - Arusha - Tanzania

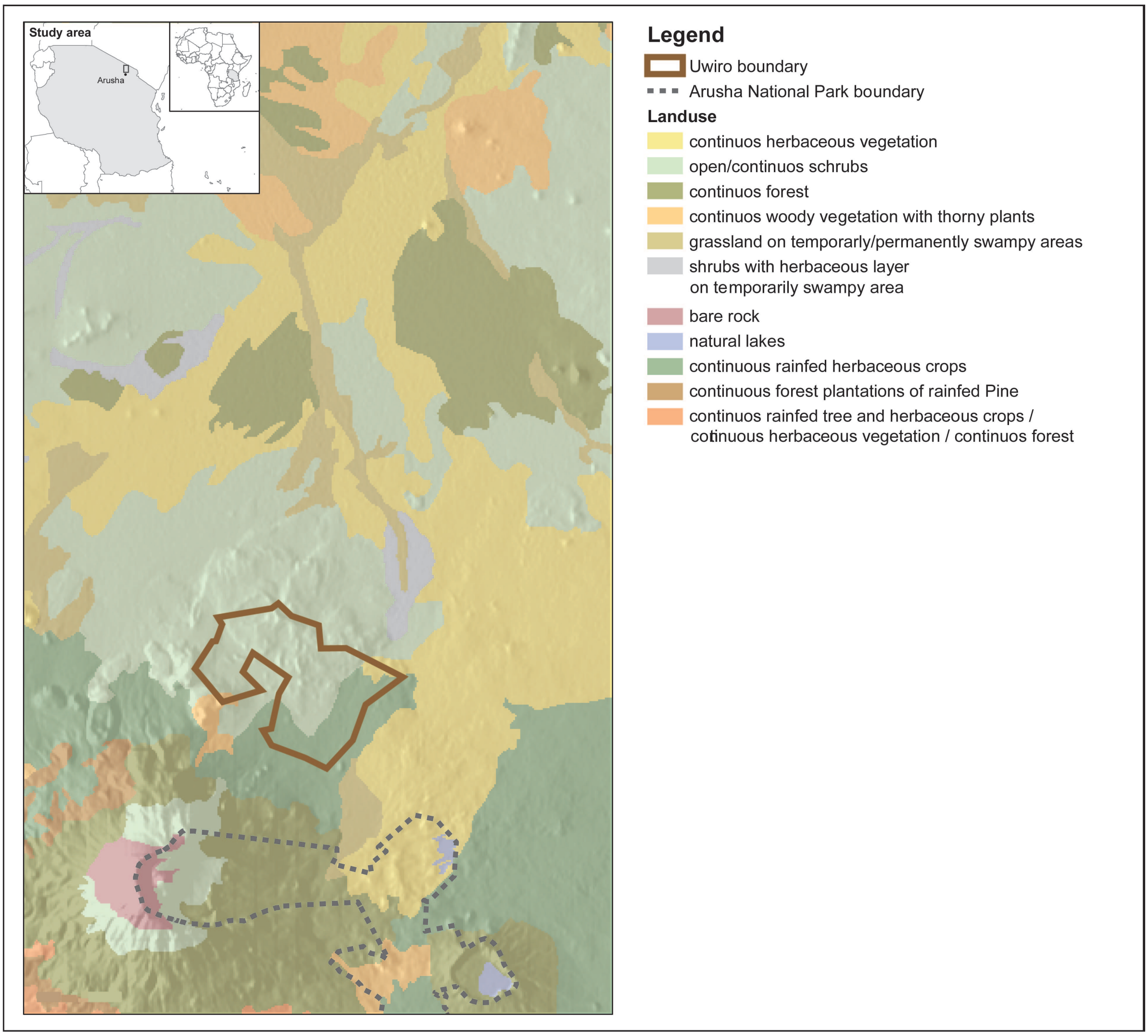


Figure 2.3.: Vegetation and land cover characteristics in Mount Meru region - Arusha - Tanzania

2.2.5. Agriculture and livestock keeping

Agriculture in this study is considered as small scale farming, basically for food crops. It is the main economic activity among the Meru and Waarusha communities. The community practices small-scale farming, mainly at household level.

The main food crops cultivated are maize, beans, and tomatoes; mostly they depend on the rainfall seasons of the year. The traditional cash crops are banana and coffee, cultivated on a small scale. Recently, a change in rainfall patterns, which are unpredictable, and a high demand for tomatoes, onions and Irish potatoes from the neighbouring regions lead the communities to modify their traditional agricultural system and introduce irrigation schemes.

There are several increasing farms, increasing in numbers, around Olkung'wado, Uwiro and Kisimiri chini, with several irrigation channels from springs and Ngare Nanyuki river supplying water to the farms. In the rest of the Maasai communities, despite their efforts to try to change their lifestyle into agro-pastoralism, the soils are very shallow with insufficient rains, which causes low harvests and sometimes no harvest at all.

Livestock keeping remains the key economic activity among the Maasai communities. They are the key livestock keepers in the area: both Waarusha and Meru communities also practice livestock keeping but with smaller numbers of heads. Modes of grazing patterns within Maasai communities as well as cattle movement patterns are done in different ways, based on pasture availability and the season of the year.

2.3. Lombardy Region (large spatial scale)

Lombardy is a northern Italian region formed by twelve provinces: Bergamo, Brescia, Como, Cremona, Pavia, Lecco, Lodi, Monza, Mantova, Milan, Sondrio and Varese. This 23859 km^2 region includes all the italian geological structures: Alpine formation at the

North, Apennines in the South-West (Oltrepo Pavese) and the Po plain in the center and South of the region. Approximately 20% of this region corresponds to protected areas, including regional parks, natural reserves and other areas of conservational interest.

2.3.1. Geology, Geomorphology and Hydrography

Lombardy, according to geomorphological and vegetation differences, can be subdivided in three areas: the Alpine zone, the Prealpine zone and the Plain (Prigioni et al., 2001). The Alpine zone in the Lombardy region is located at the center of the Alpine arc and is constituted on the East by Retiche Alps range with the major glaciers and peaks reaching 4000 m a.s.l., and on the West by the Lepontine Alps with their numerous pikes reaching elevations in the 3000 m a.s.l. range.

The line that joins the Como lake, the Mezzola lake, and through the Chiavenna Plain and the S. Giacomo Valley heads North to the Spluga pass represents the boundary between Lepontine and Retiche Alps (Prigioni et al., 2001).

The Alpine zone is characterized by a clear glacial geomorphology, which dates back to the Quaternary Age, while today water is the main agent modeling valleys and slopes.

In Lombardy the Alps present a central crystalline band between two norther and souther calcareous bands.

The Lombard Prealps are located at the foot of the Lepontine Alps, with calcareous prevalence, containing the great lakes: Verbano (Maggiore), Ceresio (Lugano), Lario (Como) and Benaco (Garda). The ridges hit the 2500 m a.s.l. elevation range and the relief exhibits a heterogeneous aspect due to the geologic variability (Prigioni et al., 2001).

The Po Plain can be differentiated between high plains and low plains. The high plain lays on the North of the superior limit of resurgences and it is characterized by acid soils, poor humus content and reduced number of rivers.

The main rivers that represent the boundaries of this area (Ticino and Adda) flow in canyons and for this reason they are hardly exploited for irrigation. The lower plain lays between the river Po and the rivers Ticino and Mincio and it is a water rich zone.

2.3.2. Vegetation and land cover characteristics

The Alpine vegetation changes as the elevation varies. Generally, starting from higher elevations and descending towards valleys bottom, several different typical vegetations are found.

On summit, where there is snow for most part of the year, we find *Dryas octopetala* and *Carex firma*. The following horizon is characterized by grasslands with Gramineae, Ciperaceae and other herbaceous vegetation.

The introduction of cattle resulted in a transition towards secondary grasslands leading to proliferation of *Nardus stricta* (Prigioni et al., 2001).

At lower heights begins the Subalpine area, that is the area of transition between the shrubby phase and arboreous phase. This zone shows different phenologies such as brushwoods, moorlands and the typical trees are Conifers. In the arboreous environment the coniferous forest is composed mainly by *Picea excelsea*, *Larix decidua* and *Pinus cembra*. The broadleaf horizon is mostly composed by *Quercus petraea*, *Tilia cordata*, *Acer campestre* and *Corylus avellana* (Prigioni et al., 2001). The long streams found in this region lead to forest and scrub specialized formations.

The submountain layer is a zone of expansion for *Quercus pubescens*, *Quercus petraea*, *Tilia cordata* and *Carpinus betulus* (Prigioni et al., 2001). The frequent and abundant rainfall favours the development of the vegetation. The dominant species are, in higher ranges *Fagus sylvatica*, and in lower ones *Castanea sativa*. There are several floristic endemisms that are typical of the Prealps (Prigioni et al., 2001).

The natural cover of the high plain is constituted by the moorland, but nowadays the landscape is dominated by intensive mais, wheat and clover cultivations and industrial and residential settlements. The lower plain has been used for cereal cultivation since ages. In the past, the Po Plain was covered by broadleaved forest, shrubs, grasses and marshes. Nowadays small island of natural vegetation can be found mainly along the Ticino river and partly along other rivers. Natural vegetation along rivers is mainly constituted by willow shrubs and trees (*Salix alba*), poplar and alder.

Apart from rivers banks, natural plain woods are constituted by *Quercus robur* and *Carpinus betulus.* Cultivated rows of *Populus canadensis* are widespread and represent one of the few elements of diversification of the agricultural landscape (Prigioni et al., 2001).

In the last forty years an increasing percentage of land has been used for industrial and housing settlements due to high increase in demography and industry, mainly near the metropolitan area of Milan, Bergamo and Brescia. Mountain regions showed deep abandon of agriculture and livestock keeping (Prigioni et al., 2001).

2.3.3. Climate

Lombardy lies in a temperate climatic zone with a continental component.

The Alps and the Apennines are meteorological barriers which dictate the thermal and rainfall regimes in the region. This leads to several microclimates typical of different circumscribed areas (Prigioni et al., 2001).

The lakes zone shows sublitoral temperate climate, the plains and the great valleys are characterized by a subcontinental temperate one, and the Prealps and the Alps, depending on the height, exhibit a cool temperate, a cold temperate, or a glacial climate. The latter can be found in areas higher than the 3000 m above sea level (Prigioni et al., 2001).

The average temperature in January (Winter) is between 0° C and 3° C , the average temperature for the month of July (Summer) is 22-24° C . The highest precipitation in the region occur in the Prealps (more than 2000 mm) while the lowest values are registered along the Po river (less then 750 mm). Rainfall in the Alps reaches its maximum in Summer and its minimum in Winter. (Prigioni et al., 2001).

For the Prealps and plains there are two maximum precipitation peaks, during Spring and Autumn, and two minimum precipitation peaks, in Summer and in Winter (Prigioni et al., 2001).

2.3.4. Demography

The population resident in Lombardy amount to about 9600000 and represent the 16.2% of Italian population. The popolution growed constantly since 1961 from 3.5 millions to 9.5 millions in 2006.

Plain and hills represent 60% of the regional territory and house 90% of the population. Since Sondrio province is located in the Alpine zone all the population lives in the mountain but it's concentrated in the valleys.

Other provinces that present mountain territory are Bergamo, Brescia, Como, Varese (Alps and Prealps) and Pavia (Appennine). Only a quarter of the population of Bergamo, Brescia and Como is resident in mountain ares and the percentage further drops in the Varese and Pavia provinces.

The plain is characterized by areas with very high demographic density and less populated areas (Prigioni et al., 2001).

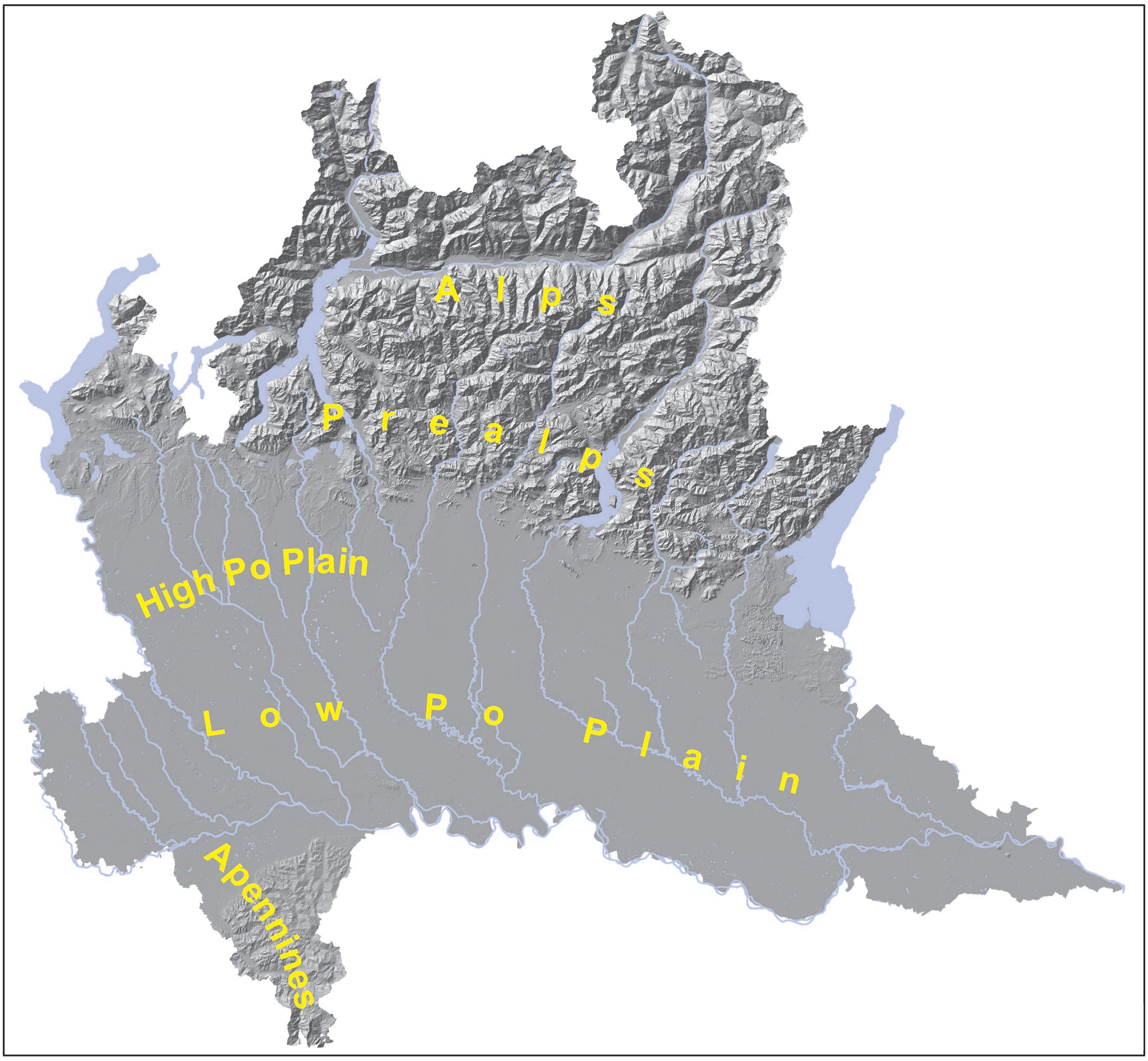


Figure 2.4.: Lombardy region

3. Material and methodology

3.1. Datawarehouse for wildlife data handling

In this study we tested the use of spatial DBMS as datawarehouse for wildlife projects. Moreover, we developed a web-database application for data input and output. The information system was based on PostgreSQL-PostGIS while the architecture of the web-database application was based on the LAMP system.

3.1.1. Spatial Database and Database Management System (DBMS)

A database is an integrated collection of data records and files based on a data model. At the moment the relational data model is the predominant choice in most applications. The relational data model has a basic structure, the relation, which has a fixed structure for data storage (Codd, 1969). Data are stored in relations in an efficient way and can be matched by using common characteristics found within the relations, so that the resulting groups of data are organized and are much easier to understand by most people.

A DBMS is a software package that controls the creation, maintenance and the use of a database; it provides facilities to extract information in a query language and to specify the logical organization inside a database and to access and use the information within it. It also gives the possibility to control data access, enforce data integrity, manage concurrency, and restore the database from backups. Moreover a DBMS also provides the ability to logically present database information to users.

A Spatial DBMS is a spatial extension of normal DBMS, representing a powerful tool for geospatial data handling, giving the advantage of storing spatial and non-spatial data in a single environment with high efficiency in query building, spatial analysis and data viewing. Spatial DBMS support the geometry data-type and through spatial index are able to access very quickly geometric data (Shekhar, 2003).

3.1.2. Web-database applications

Users access databases in order to input, manipulate and retrieve data. The language used to communicate with the database is the Structured Query Language (SQL) but non specialized users can access the database through dedicated forms.

Data dissemination among institutions is important to support management decision and requires user-friendly interfaces for data access. Web database applications represent a technical solution for the connection to the database through a web server and give the opportunity to develop custom application for multi-users access.

Web database application are based on the so called three tier model (Eckerson, 1995): at the base of the application is the database tier consisting of the database management system that creates manages and query the database.

Built on top of the database tier is the middleware tier consisting of the web server with all the scripts that translate the request from the web browser to the DBMS and, on the other side, handle the data flowf from the DBMS to the web browser. On top is the client tier, which is usually a web browser software that interacts the other applications. In this three tier model, the web provides the protocol and the network that connects the client and the middle tier of the application.

Hypertext transfer protocol (HTTP) is one component that binds together the three component. The system is realized with a stack of technologies called LAMP, acronym for Linux (operating system), Apache HTTP Server for the web server, MySQL for the database software and PHP for the scripting of the middleware tier. The combination of software included in a LAMP package may vary, in our system the database tier was based on PostgreSQL instead of MySQL.

3.2. Statistical, mathematical and GIS methods for wildlife distribution assessment

One of the three main goals of this research project was the application of GLM, and specifically Logistic Regression Analysis (LRA) to three different case studies characterized by different scale and different species datasets.
The modelling process uses a typical presence-absence method fitted with pseudo-absences

3.2.1. Generalized Linear Models (GLM)

GLM are an extension of the classical multiple regression technique, allowing non-linearity in the data (Guisan et al., 2002). This regression tool is widely used mainly because of its ability to deal with the variety of distributions that describe ecological data (normal, binomial, Poisson, exponential, gamma), and also because it harmoniously fits with practices commonly used in linear modeling and analysis of variance (ANOVA) (McCullagh and Nelder, 1989; Guisan et al., 2002).
This predictive modelling methodology has been largely tested, and proved to be robust in a number of independent situations (Elith et al., 2006). GLM assumes a relationship (called link function) between the mean of the response variable and the linear combination of the explanatory variables (McCullagh and Nelder, 1989; Guisan et al., 2002). This model does not force results into unnatural scales (which would cause under and overestimations), and allows non-linearity and non-constant variance structures in the data. In this study, regressions were implemented as generalized linear models with a binomial error distribution. For this type of distribution the expected value varies be-

tween 0 and 1 and a link function should satisfy the condition that it maps the interval (0,1) on to the whole real line (Mc-Cullagh and Nelder, 1989).
There is a wide choice of link functions available (for example, the logistic, the probit, the complementary log-log function); in this study the logistic link function was chosen to compute the model because the response variable was limited to values between 0 and 1.

Logistic regression analysis (LRA) (Hosmer & Lemeshow, 2000) extends the techniques of multiple regression analysis to research situations in which the outcome dependent variable is categorical. In practice, situations involving categorical outcomes are quite common. In the setting of evaluating the potential distribution of species, for example, predictions may be made for the dichotomous outcome of presence/absence. Extensions of the LRA technique outcome are also available.
The fundamental model underlying multiple regression analysis (MRA) posits that a continuous outcome variable is, in theory, a linear combination of a set of predictors and error. Thus, for an outcome variable, Y, and a set of p predictor variables, $X_1, ..., X_p$, the MRA model is of the form:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon = \alpha + \sum_{j=1}^{n} \beta_j X_j + \varepsilon$$

where α is the intercept, β_j is a multiple (partial) regression coefficient and ε is the error of prediction. If error is omitted, the resulting model represents the expected, or predicted, value of Y:

$$E(Y|X_1, ..., X_n) = Y' = \alpha + \sum_{j=1}^{n} \beta_j X_j + \varepsilon$$

Since $Y = Y' + \varepsilon$, each observed score, Y, is made up of an expected, or predictable component, Y', that is a function of the predictor variables $X_1, ..., X_n$, and an error, or unpredictable component, ε, that represents error of measurement and/or error in the selection of the model.

The MRA model summarized above is applicable when the outcome variable, Y, is continuous, but it is not appropriate for situations in which Y is categorical.

The model for logistic regression analysis assumes that the outcome variable, Y, is categorical (e.g., dichotomous), but LRA does not model this outcome variable directly. Rather, LRA is based on probabilities associated with the values of Y.

In most cases Y is assumed to be dichotomous, taking on values of 1 (i.e., the positive outcome) and 0 (i.e., the negative outcome). In theory, the hypothetical proportion of cases for which $Y = 1$ is defined as p = P(Y =1). Then, the theoretical proportion of cases for which $Y = 0$ is $1 - p = P(Y = 0)$. In the absence of other information, p is estimated by the sample proportion of cases for which $Y = 1$. However, in the regression context, it is assumed that there is a set of predictor variables, $X_1, ..., X_n$, that are related to Y and, therefore, provide additional information for predicting Y. For theoretical, mathematical reasons, LRA is based on a linear model for the natural logarithm of the odds (i.e., the log-odds) in favor of $Y = 1$:

$$\log_e\left[\frac{P(Y=1|X_1,...,X_n)}{1-P(Y=1|X_1,...,X_n)}\right] = \log_e\left[\frac{\pi}{1-\pi}\right] =$$

$$= \alpha + \beta_1 X_1 + ... + \beta_n X_n + \varepsilon = \alpha + \sum_{j=1}^{n} \beta_j X_j + \varepsilon$$

In the LRA model, P is a conditional probability of the form

$$P(Y = 1|X_1, ..., X_n)$$

so it is assumed that "success" is more or less likely depending on combinations of values of the predictor variables.

The log-odds, as defined above is also known as the logit transformation of P and the analytical approach described here is sometimes known as logit analysis.

The LRA model above is identical to the MRA model except that the log-odds in favor of $Y = 1$ replaces the expected value of Y.

There are two basic reasons underlying the development of the model above. First, probabilities and odds obey multiplicative, rather than additive, rules. However, taking the logarithm of the odds allows for the simpler, additive model since logarithms convert multiplication into addition. Second, there is a (relatively) simple exponential transformation for converting log-odds back to probability. In particular, the inverse transformation is the logistic function of the form:

$$P(Y=1|X_1,...,X_n)=\frac{e^{\alpha+\sum_{j=1}^{n}\beta_j X_j}}{1+e^{\alpha+\sum_{j=1}^{n}\beta_j X_j}}$$

Due to the mathematical relationship, $e^a/(1+e^a)=1/(1+e^{-a})$, the logistic function for LRA is sometimes presented in the form:

$$P(Y=1|X_1,...,X_n)=\frac{1}{1+e^{-\alpha-\sum_{j=1}^{n}\beta_j X_j}}$$

Due to the mathematical relation, $1-e^a/(1+e^a)=1/(1+e^a)$, the probability for a 0 response is:

$$P(Y=0|X_1,...,X_n)=1-P(Y=1|X_1,...,X_n)=\frac{1}{1+e^{\alpha+\sum_{j=1}^{n}\beta_j X_j}}$$

As in MRA, there are two important stages in the analysis of data. First, estimates for the parameters in the model must be obtained and, second, some determination must be made of how well the model actually fits the observed data.

In MRA, the parameter estimates are obtained using the least-squares principle and assessment of fit is based on significance tests for the regression coefficients as well as on the interpretation of the multiple correlation coefficient. The parameters that must be estimated from the available data are the constant and the logistic regression coefficients.

Because of the nature of the model, estimation is based on the maximum likelihood principle rather than on the least-squares principle. The process of finding maximum likelihood estimates is somewhat more complicated than the corresponding minimization procedure in MRA for finding least-square estimates. However, the general approach involves establishing initial guesses for the unknown parameters and then continuously adjusting these estimates until the maximum value of likelihood is found. This iterative solution procedure is available in popular statistical software.

As McCullagh and Nelder (1989) proved, the logistic function (or logit function) has an important advantage over the alternative transformations in that it is suited for the analysis of data collected retrospectively.

The logit transformation of the probability of presence-absence (p) produced linear function according to the equation (McCullagh and Nelder, 1989; Manel et al., 2001):

$$logit(p) = \log \frac{p}{1-p}$$

The modelling process requires the finding of one or more appropriate parsimonious sets of explanatory variables. In order to obtain a parsimonious model, the variables included should have a detectable effect on the response. Should the sum of a single variable reduce the residual mean square to, for instance, one third of its original value, then it surely is a variable to be included in the model, and if the addition causes no reduction then it is to be excluded (McCullagh and Nelder, 1989).

Problems appear when a decision for intermediate ambiguous cases is needed. This could be suppressed with the use of computational skills. There are three methods for selecting predictors through a stepwise procedure:

1. Forward selection, in which the best unselected variable satisfying the selection criterion is added until no further candidate variables remain;
2. Backward selection, that starts with the complete set of variables and eliminates the worst, one by one, until all remaining variables are necessary;
3. Both selection, which combines the two previous procedures, following backward elimination by forward selection until both fail to change the model.

We used a "both" selection procedure combining the forward and backward methods. The best combination of predictors was selected using the Akaike's Information Criterion

(AIC) (Akaike, 1980). The criterion is based on the minimum choice of model parameters to form a tradeoff between the fit of the model (which lowers the sum of squared residuals) and the model's complexity.

LRA requires not only presence, but also absence data. As absence data were not available in the original data sets, pseudo-absences were generated (Keating and Cherry, 2004). The number of pseudo-absence points was 30% more than real presences; pseudo-absences were then combined with the real presences into a single presence-absence dataset. LRA was fitted within R software (R Development Core Team, 2007).

As the fitted model describes the probability of the species presence occurrence, we had to determine the threshold to distinguish between the binomial response of absence and presence. The choice was based on the Receiver Operating Characteristic (ROC) analysis.

3.2.2. Model Evaluation and Classification

In order to estimate classification performance for the GLM and ENFA models we performed the Receiver Operating Characteristic (ROC) analysis.

In ROC analysis, specificity corresponds to the number of true absences on the overall number of absences in the sample while sensitivity reflects the true positive fraction (Pearce and Ferrier, 2000).

The ROC plot is obtained by plotting all correctly predicted presences, divided by the total number of presences on the y axis, versus the false positive fraction (1-specificity) for all available thresholds on the x axis (Fielding and Bell, 1997).

In this way the ROC curve reflects a compromise between the sensitivity and the false positive proportion as the decision threshold is varied. Moreover these fraction values are independent of the prevalence of a species because they are expressed as a proportion of all locations with a given observation point (Pearce and Ferrier, 2000). In this way, the Area Under the (ROC) Curve (AUC) is an important index, as it provides a measure of overall accuracy that is not dependent on a particular threshold or species prevalence

(Pearce and Ferrier, 2000).

The AUC ranges between 0.0 and 1.0, where 1.0 indicates perfect discrimination between true positive and false positive, 0.5 shows that the predictive discrimination is near the random guess and values under 0.5 indicate performance worse than random (Elith et al., 2006).

The AUC values can be interpreted as the probability for a random selection from the positive group to have a higher predictive value than a random selection from the negative group. For the model predictions to be considered valid, AUC scores should be over 0,5 and the ROC plot curve should have a smoothed shape. An elbow shaped or an angular curve indicate over-fitting of the data. The graph in figure 3.1 shows three ROC curves representing excellent, good, and worthless results plotted on the same graph.

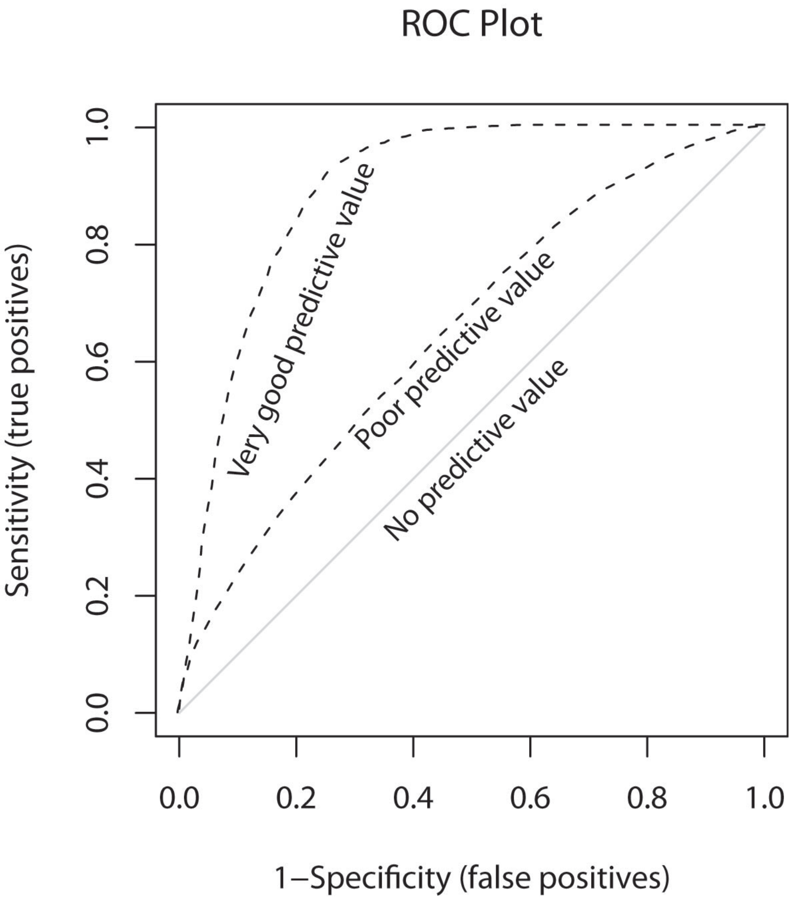


Figure 3.1.: Comparison of ROC curves

The ROC statistic can also be used to identify an appropriate threshold value for a given application. In our case, for methods producing probability values or continuous indexes that have to be translated into a binomial response we set a threshold value by determining a point in the ROC space. The continuous maps obtained from the GLM were reclassified using the threshold value based on the method specificity=sensitivity. This led to the conversion of continuous map values into binomial ones describing absence and presence of the species. This evaluation was also performed with R software using the package Presence Absence (Freeman and Moisen, 2008).

3.3. Software tools

Data management, processing and analysis were performed using exclusively Free and Open Source Software (FOSS). These software are freely distributed under GNU General Public License often called the GNU GPL for short.

The GNU GPL is a free, copyleft license for software and other kinds of works. It is intended to guarantee the freedom to share and change all versions of a program to make sure it remains free software for all its users. (Stallman, 1989 and Free Software Foundation, 2007)

Free software, refers to freedom, not price. The GPL is designed to ensure the freedom to distribute copies of free software (or charge for them), availability of source code, possibility to change the software or use pieces of it in new free programs.

The use of FOSS software is made easier by a huge and worldwide community of users and developers connected through mailing list and forums.

Since a huge part of this study had applications for institution with limited economic resources, even for those residing in developing country, the possibility to use and distribute the software without charging for it was an important factor. We used the following FOSS software:

- PostGIS/PostgreSQL - spatial DBMS

- GRASS - Geographical Information System (GIS)
- R - statical analysis
- LAMP - web server

3.3.1. PostgreSQL and PostGIS

PostgreSQL, or simply Postgres, is an object-relational database management system (ORDBMS); it evolved from the Ingres project at University of California, Berkeley in 1985 as a post-Ingres project to address the problems with contemporary database systems that had become increasingly clear during the early 1980s. In 1996, the project was renamed to PostgreSQL to reflect its support for SQL.

The first PostgreSQL release formed version 6.0 in January 1997. Since then, the software has been maintained by a group of database developers and volunteers around the world, coordinating via the Internet. Postgres provides a variety of features for data management such as functions, indexes, Multi-Version Concurrency Control (MVCC), triggers, rules and a huge variety of data types are supported. Open source front-end software like psql or pgAdmin is used to administrate the database. (PostgreSQL Global Development Group, 1996-2010)

PostGIS is the spatial extension of PostgreSQL: it's an open source software program that adds support for geographic objects to the PostgreSQL object-relational database. PostGIS includes all the geometry types specified by the Open Geospatial Consortium, spatial operators for measurements and spatial operations and spatial predicates for determining the interaction of geometries. It also provides indexes for high performance queries. The PostGIS implementation is based on light-weight geometries and optimized indexes. The first version was released in 2001 by Refractions Research under the GNU General Public License. A stable 1.0 version was released on April 19, 2005.

At the beginning of this study we concentrated on the spatial DBMS SpatiaLite since we were addressed to light solutions with possibility to transfer the architecture to other systems for management and planning purposes. In fact SpatiaLite is based on the well known DBMS SQLite, it supports all the SQL and OGC-SFS specifications even if it's very light, simple in structure and without any complexity.

Moreover SpatiaLite consists of a unique cross-platform file that can be easily transferred with all his geographical content. Despite of all this advantages we had to abandon this solution because of the difficulties that we found integrating this DBMS in the web-database application. In fact SpatiaLite is still little supported by PEAR libraries that are fundamental in PHP scripts for client-database connection.

3.3.2. GRASS

GRASS is the acronym for Geographic Resources Analysis Support System; it's a Geographical Information System free software, licensed under the GNU GPL and distributed with versions for different platforms. GRASS was born in the early 80's as a project of the United States Army (U.S. Army Corp of Engineering Research Laboratory - USA)(GRASS Development Team, 1999-2010).

The development was held by using in particular the C language and UNIX-like operating system reference.

In 1996 the U.S. Army took the decision to abandon the development of GRASS; users were encouraged to migrate to commercial systems while the latest version of GRASS (4.1) remained in the public domain. (GRASS Development Team, 1999-2010)

At the end of 1997, after more than a year, a new international team was organized for continuing development. In 1999 the GRASS Development Team (GDT) decided to grant GRASS (5.0b) the GPL license. Currently, the software development center is based in San Michele All'Adige (Trento, Italy) at the Fondazione Edmund Mach, but it uses mostly volunteers from all over the world as contributors. (GRASS Development Team, 1999-2010)

Grass is now a powerful GIS used for geospatial data management and analysis, image processing, graphics/map production, spatial modeling, and visualization. GRASS is

diffused worldwide for academic and commercial applications and is also used by government agencies (NASA, NOAA, USDA, DLR, CSIRO, National Park Service of USA, U.S. Census Bureau, USGS and JRC). (GRASS Development Team, 1999-2010)
The last versions were upgraded in function dealing with the management of the topology of two and three-dimensional vector data and attributes data handling.
Moreover, the introduction of a graphical interface contributed to increase the number of users, especially those who were familiar with commercial desktop GIS. One of the advantages of GRASS is that it can analyze data in a not-interactive way through scripts coded in the popular bash script language. GRASS can me coupled with R to perform statistical analyses on geographic data in GRASS native format; the package required is spgrass6 (Bivand,2009)

3.3.3. R

R is a language and environment for statistical computing and graphics licensed under the GNU GPL and is available for different platforms. Its architecture is similar to the S language so that R can be considered as a different implementation of S. (R, Core Development Team, 2010)
R provides a wide variety of statistical technique such as linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering. It also provides graphical techniques, and is highly extensible. (R, Core Development Team, 2010)
R is distributed with a command line interface and requires a strong knowledge of R language. In order to simplify the interaction with the software, graphical interfaces are available, running either under linux or windows (Tinn-R, RKward, ESS). As in GRASS, it is possible to use R in a non-interactive manner using scripts that contain ordered commands in R language.

3.3.4. LAMP system

LAMP is an acronym for a package of free and open source software, originally coined from the first letters of Linux (operating system), Apache HTTP Server, MySQL (database software), and PHP. These are the principal components to build a viable general purpose web server.

The precise combination of software included in a LAMP package may vary, especially with respect to the web scripting software. Even though the original authors of these programs did not design them all to work specifically with each other, the development philosophy and tool sets are shared and were developed in close conjunction.

The software combination has become popular because it is free of cost, open source, and therefore easily adaptable, and also thanks to the ubiquity of its components which are bundled with the most current Linux distributions.

When a DBMS web server and a srcripting language are used in combination they represent a solution stack of technologies that support application servers.

3.4. Automated mapping process: SOS

During this study we developed an analysis system for automatic models computing in order to make the modeling processes efficient, repeatable and prompt.

Biodiversity (species richness) maps were based on a large number of species. Moreover the modeling process required long and iterative procedures based on a large number of predictable variables. For all these reasons the computation of models required automated computing.

We prepared a package of scripts for GRASS and R, integrated into an analysis system that we called Species Open Spreader (SOS); the system allows automatic computation of spatial statistical model for a great number of species.

The structure of the analysis system is planned to perform geoprocessing for data preparation in GRASS and then to run different modeling techniques in R.

SOS is built up of a common module for input-output and plug-in scripts for statistical analysis, production of results, and preparation of outputs. SOS is designed to be extensible and perform a large number of statistical models; at the moment three model are already available: Logistic Regression Analysis, Environment Niche Factor Analysis (ENFA) and the flexible Discriminant Analysis with method BRUTO.

Moreover SOS is able to compute Habitat Suitability Index (HSI), a model based on the opinion of experts, and spatial overlay: these two methods are useful when dataset are not suitable for statistical analysis.

The statistical analysis is completely automated: it is controlled through a configuration file and is performed by calling a unique R script (`SPECIES.SPATIAL.MODELS.R` - Appendix B). R is launched from the GRASS environment and koupled to GRASS through the library spgrass6. The configuration file is a text file containing the following items:

- name of the species for which models are required;
- reference to the environmental variables;
- name of the statistical method to apply;
- outputs required.

The script `SPECIES.SPATIAL.MODELS.R` reads the configuration file and subsequently calls other scripts to perform the required statistical analysis and to produce outputs. The SOS suite for model computation is composed of the scripts listed in table 3.1 and reported in the Appendix B.

Table 3.1.: List of scripts for model computation available in SOS.

Name	**Description**
`SPECIES.SPATIAL.MODELS.R`	main R script, reads the configuration file and subsequently calls the required scripts
`CALC.LOGISTC.R`	script to compute LRA.
`CALC.ENFA.R`	script to compute ENFA.
`CALC.BRUTO.R`	script to compute BRUTO analysis.
`CALC.OVERLAY.sh`	script to compute spatial overlay.
`CALC.HSI.sh`	script to compute HSI.
`OUT.LOGISTIC.R`	script to produce outputs of LRA analysis.
`OUT.ENFA.R`	script to produce outputs for ENFA analysis.
`OUT.BRUTO.R`	script to produce outputs for BRUTO analysis.

`CALC.OVERLAY.sh` and `CALC.HSI` are called directly in GRASS, since they are based on spatial processing and do not require statistical procedures. For statistical analysis SOS produces three kinds of outputs:

- text files containing the results of statistical analysis in R format;
- images in .eps format containing graphs from statistical analysis results;
- maps of potential distribution, computed with the results of the statistical analysis.

Depending on the statistical model, maps either are generated in R as output of the statistical analysis (ascii file) or in GRASS, running a bash script automatically generated by R.

Outputs are stored inside GRASS mapsets and more precisely in a folder called SOS.rs containing the subfolders maps, text and plots.

SOS is also provided with GRASS scripts for common spatial data processing task used in preparation of datasets. This part of SOS is not completely automated but scripts are called by the user depending on the characteristics of the available data and on the dataset required for the statistical analysis.

The SOS suite of GRASS scripts is listed in table 3.2 and is presented in Appendix B.

Table 3.2.: List of GRASS scripts available in SOS

Name	**Description**
`V.WHATRAST.MANY.SH`	prepares a dataset containing the value of predictive variable at each location of presence and absence points.
`V.PRESABS.MANY.SH`	prepares a dataset of presence and absence points from real observations.
`V.PRESABS.MANY.TRANSECTS.SH`	prepares a dataset of presence and absence points from real observations along transects.
`V.PRESABS.MANY.GRID.SH`	prepares a dataset of presence and absence points from wildlife atlas grids.

Name	**Description**
`R.CATDIST.MANY.SH`	computes distance maps for vector environmental variables in order to produce continuous raster datasets.

4. Datasets

4.1. Brescia plain

The goal of this application was the computation of a synthetic map as the result of a monitoring program for environmental quality evaluation. The synthetic map of ecosystem value was based on the assessment of wildlife and vegetation values through field data processing in SOS. Wildlife value map was obtained through the computation of a weighted species richness map based on potential distribution maps.
The monitoring program was addressed to three taxonomic groups: Amphibians, Birds and Mammals: 88 species were registered of which 3 Amphibians, 76 Birds and 9 Mammals. Table 4.1 lists the species registered in the monitoring program with the number of observations for each species.
Birds were monitored with the point count technique (Bibby et al., 2000). Since census took place in a flat region, we assumed a 200 m contact distance, thus we enlarged the original dataset by computing 50 grid ordered points for each contact point, inside a 200 m buffer.

Table 4.1.: List of monitored species in the Brescia plain area

Class	**Scientific name**	**Common name**
Amphib.	*Rana synkl. esculenta*	edible frog
Amphib.	*Rana dalmatina*	agile froge
Amphib.	*Bufo viridis*	green toad
Birds	*Phalacrocorax carbo*	cormorant
Birds	*Ixobrychus minutus*	little bittern
Birds	*Nycticorax nycticorax*	night heron
Birds	*Ardeola ralloides*	squacco heron
Birds	*Egretta garzetta*	little egret
Birds	*Ardea cinerea*	grey heron
Birds	*Ardea purpurea*	purple heron
Birds	*Anas platyrhynchos*	mallard
Birds	*Buteo buteo*	buzzard
Birds	*Falco subbuteo*	hobby
Birds	*Falco tinnunculus*	kestrel
Birds	*Phasianus colchicus*	pheasant
Birds	*Perdix perdix*	grey partridge
Birds	*Coturnix coturnix*	quail
Birds	*Gallinula chloropu*	moorhen
Birds	*Fulica atra*	coot
Birds	*Rallus aquaticus*	water rail
Birds	*Himantopus himantopus*	blach-winged stilt

Class	**Scientific name**	**Common name**
Birds	*Charadrius dubius*	little ringed plover
Birds	*Actitis hypoleucos*	common sandpiper
Birds	*Tringa nebularia*	greenshank
Birds	*Larus ridibundus*	black-headed gull
Birds	*Larus michahellis*	yellow-legged gull
Birds	*Columba palumbus*	woodpigeon
Birds	*Streptopelia decaocto*	collared dove
Birds	*Streptopelia turtur*	turtle dove
Birds	*Cuculus canorus*	common cuckoo
Birds	*Athene noctua*	little owl
Birds	*Strix aluco*	sttrix aluco
Birds	*Tyto alba*	barn owl
Birds	*Apus apus*	swift
Birds	*Alcedo atthis*	kingfisher
Birds	*Upupa epops*	hoopoe
Birds	*Dendrocopos major*	great spotted woodpecker
Birds	*Picus viridis*	green woodpecker
Birds	*Jynx torquilla*	wryneck
Birds	*Alauda arvensis*	skylard
Birds	*Hirundo rustica*	swallow
Birds	*Delichon urbica*	house martin
Birds	*Motacilla alba*	white wagtail
Birds	*Motacilla cinerea*	grey wagtail
Birds	*Motacilla flava*	yellow wagtail
Birds	*Prunella modularis*	dunnock
Birds	*Troglodytes troglodytes*	wren

Class	**Scientific name**	**Common name**
Birds	*Erithacus rubecula*	robin
Birds	*Luscinia megarhynchos*	nightingale
Birds	*Phoenicurus phoenicurus*	redstart
Birds	*Phoenicurus ochruros*	black redstart
Birds	*Saxicola rubetra*	whinchat
Birds	*Saxicola torquata*	stonechat
Birds	*Turdus merula*	blackbird
Birds	*Turdus philomelos*	song thrusch
Birds	*Cettia cetti*	cetti's warbler
Birds	*Acrocephalus scirpaceus*	reed warbler
Birds	*Acrocephalus palustris*	marsh warbler
Birds	*Hippolais polyglotta*	melodious warbler
Birds	*Phylloscopus collybita*	chiffchaff
Birds	*Sylvia communis*	whitethroat
Birds	*Sylvia atricapilla*	blackcap
Birds	*Regulus regulus*	goldcrest
Birds	*Muscicapa striata*	spotted flycatcher
Birds	*Remiz pendulinus*	penduline tit
Birds	*Aegithalos caudatus*	long-tailed tit
Birds	*Parus major*	great tit
Birds	*Parus caeruleus*	blue tit
Birds	*Sitta europaea*	nuthatch
Birds	*Lanus excubitor*	great grey shrike
Birds	*Lanus collurio*	red-backed shrike
Birds	*Sturnus vulgaris*	starling
Birds	*Oriolus oriolus*	golden oriole

Class	**Scientific name**	**Common name**
Birds	*Pica pica*	magpie
Birds	*Corvus corone cornix*	hooded crow
Birds	*Passer domesticus*	italian sparrow
Birds	*Passer montanus*	tree sparrow
Birds	*Fringilla coelebs*	chaffinch
Birds	*Carduelis carduelis*	goldfinch
Birds	*Carduelis chloris*	greenfinch
Birds	*Serinus serinus*	serin
Birds	*Emberiza schoeniclus*	reed bunting
Birds	*Miliaria calandra*	corn bunting
Mammals	*Erinaceus europaeus*	western hedgehog
Mammals	*Talpa europaea*	european mole
Mammals	*Lepus europaeus*	brown hare
Mammals	*Arvicola terrestris*	water vole
Mammals	*Rattus norvegicus*	brown rat
Mammals	*Myocastor coypus*	coypu
Mammals	*Vulpes vulpes*	red fox
Mammals	*Meles meles*	badger
Mammals	*Martes foina*	beech marten

In order to obtain potential distribution maps for each species we first applied Logistic Regression Analysis. The analysis was base on the set of environmental variables listed in table 4.2.

Table 4.2.: Environmental variables used for LRA in the Brescia plain case study

Variable name	**Description**	**Source**
EV-CWED	contrast weighted edge density	computed
EV-LAKE	distance from lakes	CTR
EV-RVR1	distance from rivers (primary hydrographic net)	CTR
EV-RVR2	distance from rivers (secondary hydrographic net)	CTR
EV-ROAD	weighted distance from roads	CTR
EV-H103	broad-leaved forest	DUSAF
EV-H104	riparian vegetation	DUSAF
EV-H105	coniferous forest	DUSAF
EV-H106	mixed forest	DUSAF
EV-H107	recent reforestation	DUSAF
EV-H108	orchards and vegetables gardens	DUSAF
EV-H109	tree crops	DUSAF
EV-H110	marshy vegetation	DUSAF
EV-H111	debris vegetation	DUSAF
EV-H112	riverbed vegetation	DUSAF
EV-H113	schrubs	DUSAF
EV-H114	meadows and pastures	DUSAF
EV-H115	herbaceous crops	DUSAF
EV-H116	herbaceous and tree crops	DUSAF
EV-H117	protected crops	DUSAF
EV-H120	dumps	DUSAF

Variable name	**Description**	**Source**
EV-H121	continuous urban areas	DUSAF
EV-H122	open urban areas	DUSAF
EV-H123	farm building	DUSAF

Contrast weighted edge density (CWED) is a landscape metric computed with the GRASS r.li commands package for landscape structure analysis, it calculates the edge density between patch types specified as follows:

$$CWED : \frac{\sum_{k=1}^{m} e_{ik}d_{ik}}{AREA} 10000$$

with

- k: attribute;
- m: number of non-null attributed in the sampling area;
- e_ik: total length of edge in landscape between patch types i and k;
- d_ik: dissimilarity (edge contrast weight) between patch types i and k;
- $Area$:total landscape area.

CWED was computed for different groups of species with similar ecological characteristics. We also used layers from the Regional Technical Map (CTR) (Regione Lombardia, 1998), describing hidrography, and communication ways.
The variables describing habitat types were generated from an informative base called DUSAF - Destinazione d'Uso dei Suoli Agricoli e Forestali (Agriculture and Forest Soil usage destination). DUSAF was created in 2000 by the Agriculture Department of Lombardy Region and by ERSAF (Ente Regionale per i servizi all'agricoltura e alle foreste) to bring off a detailed database of soil usage in the region. It was based on photo interpretation of a 85 cm resolution digital color orthophoto from the IT2000 aerial survey.

4.2. Mount Meru region - Tanzania

In order to produce maps of species potential distribution and biodiversity maps in the Mount Meru region we applied Logistic regression analysis. The dataset used for the analysis consisted of presence/absence wildlife data whereas the set of environmental variables consisted of raster layers describing landuse, climate, landscape and hydrography. Landuse characteristics were derived from Africover (FAO, 2003), a Food and Agriculture Organization (FAO) project which produced a digital georeferenced database on land cover and a geographic referential for the whole of Africa at a scale of 1:200000 (www.africover.org). In order to obtain continuous data, raster maps of distance from each habitat type were computed.

Climatic variables were derived from the BIOCLIM project (Hijmans and Jarvis, 2005), a bioclimatic prediction system which uses surrogate terms (bioclimatic parameters) derived from mean monthly climate estimates to approximate energy and water balances at a given location. The present version can produce up to 35 bioclimatic parameters based on the climate variables maximum temperature, minimum temperature, rainfall, solar radiation and pan evaporation.

We used the SRTM 90 digital elevation model (CGIAR, 2004) from which aspect slope layers were computed. Other unpublished topographic layers regarding hydrography, location of villages and roads were collected by Istituto Oikos from previous research projects.

We also computed landscape indexes with the GRASS r.li commands package for landscape structure analysis. We produced four maps based on the indexes of dominance's diversity, Shannon diversity, edge density, patch density and mean patch size. The dominance's diversity index is computed as:

$$Dominance : \ln(m) + \sum_{1=1}^{m} p_i \ln(p_i)$$

with:

- m: number of different patch type

- $_i$: patch type
- p_i: proportion of the landscape occupied by patch type i

The Shannon diversity index is given by the formula:

$$Shannon : -\sum_{1=1}^{m} p_i \ln\left(p_i\right)$$

The edgedensity index is computed as:

$$Edgedensity : \frac{\sum_{k=1}^{m} \sum_{i=1}^{n} e_i k}{AREA} 10000$$

were

- k: patch type
- m: number of patch type
- n: number of edge segment of patch type k
- $e_i k$:total length of edge in landscape involving patch type k
- $AREA$: total landscape area

The patchdensity index is given by the fraction:

$$Patchdensity = \frac{N}{AREA} 1000000$$

were N is the total number of patches.

Table A.1 shows the list of environmental variables used in this study.

Table 4.3.: Environmental variables for spatial models in Mount Meru region

Variable name	**Description**
EV-CL08	mean temperature of wettest quarter
EV-CL09	mean temperature of driest quarter
EV-CL16	precipitation of wettest quarter
EV-CL17	precipitation of driest quarter
EV-DOMI	dominance's diversity index
EV-SHAN	Shannon's diversity index
EV-EDDE	edge density index
EV-PATC	patch density index
EV-LAKE	distance from lakes
EV-RVRS	distance from rivers
EV-ROAD	distance from roads
EV-SWPA	distance from swamps
EV-VLLG	distance from villages
EV-TASP	digital elevation model - aspect
EV-TSLP	digital elevation model - slope
EV-H101	continuous herbaceous vegetation
EV-H105	continuous herbaceous vegetation with scrubs
EV-H108	continuous herbaceous vegetation with isolated rainfed crop
EV-H113	continuous herbaceous vegetation with sparse trees and scrubs
EV-H115	continuous herbaceous vegetation with sparse scrubs

Variable name	**Description**
EV-H117	continuous herbaceous vegetation with isolated rainfed crop
EV-H122	open shrubs with sparse trees
EV-H126	open shrubs with sparse trees and with isolated fields of rainfed crop
EV-H127	open shrubs
EV-H131	open shrubs with isolated field of rainfed crop
EV-H132	open shrubs with combination of rainfed tree and herbaceous crops
EV-H134	continuous shrubs with herbaceous layers with rainfed tree and herbaceous crops
EV-H141	continuous shrubs with herbaceous layers/herbaceous vegetation with sparse trees and shrubs
EV-H143	open shrubs with sparse trees/continuous herbaceous vegetation with sparse trees and shrubs
EV-H145	open shrubs with sparse trees/isolated field of rainfed crop
EV-H146	continuous trees forest with scrubs
EV-H154	continuous broadleaved deciduous forest with herbaceous layer and sparse scrubs/urban area
EV-H162	continuous tree forest with shrubs
EV-H165	continuous broadleaved deciduous forest with shrubs
EV-H166	continuous broadleaved deciduous forest with shrubs/grassland

Variable name	**Description**
EV-H170	continuous woody vegetation with thorny plants
EV-H174	grassland with sparse shrubs on temporarily swampy area
EV-H175	grassland on permanently flooded area
EV-H177	grassland/woody vegetation with herbaceous layer on temporarily swampy area
EV-H180	shrubs with herbaceous layer on temporarily swampy area
EV-H182	shrubs with herbaceous layer on temporarily swampy/grassland on temporarily swampy area
EV-H190	bare rock
EV-H193	natural lakes
EV-H196	continuous rainfed herbaceous crops/continuous herbaceous vegetation with sparse trees and shrubs
EV-H202	continuous rainfed herbaceous crops/continuous herbaceous vegetation with sparse trees and shrubs
EV-H204	continuous rainfed herbaceous crops
EV-H205	continuous rainfed herbaceous crops/continuous herbaceous vegetation
EV-H207	continuous rainfed herbaceous crops/continuous herbaceous vegetation with sparse trees and shrubs
EV-H211	continuous rainfed herbaceous crops/combination of rainfed tree and shrubs crops
EV-H214	continuous forest plantations of rainfed Pine
EV-H216	continuous combination of rainfed tree and herbaceous crops/continuous herbaceous vegetation

Variable name	Description
EV-H218	continuous rainfed tree and herbaceous crops/continuous broadleaved deciduous forest with shrubs

Wildlife data for the Mount Meru Region consisted of real observations along foot and road transects for 39 mammal species. These data were collected from Istituto Oikos for previous projects (Ntalwila, 2007).

The total number of transects is 31 and covered the whole study area.

Using SOS we prepared a dataset of presence and absence points for each species, in which presences corresponded to real observation while absences were randomly generated. The number of absence points exceeded the number of presence points by 30%. Random absence data for each specie were generated along transects were no observation occurred.

The presence/absence dataset was prepared using the SOS script for GRASS `V.PRESABS.MANY.TRANSECTS.sh` (Appendix B).

Table 4.4.: List of species involved in the computation of potential distribution maps

Order	Scientific name	Common name	N.
Primates	*Cercopithecus mitis*	blue monkey	544
	Papio cynocephalus	yellow baboon	715
	Colobus guereza	mantled guereza	473
Rodentia	*Hystrix cristata*	crested porcupine	12
Lagomorpha	*Lepus saxatilis*	scrub hare	749

Order	Scientific name	Common name	N.
	Pronolagus rupestris	Smith's red rock hare	765
Carnivora	*Canis aureus*	golden jackal	219
	Canis mesomelas	black-backed jackal	250
	Lycaon pictus	wildog	14
	Otocyon megalotis	bat-eared fox	180
	Acynonix jubatus	cheetah	3
	Felis silvestrys	wildcat	3
	Crocuta crocuta	spotted hyena	102
	Hyaena hyaena	hyena	3
	Civettictis civetta	african civet	2496
Proboscida	*Loxodonta africana*	african elephant	2398
Perissodactyla	*Equus burchelli*	zebra	1253
Hyracoidae	*Procavia capensis*	rock hyrax	146
Tubulidentata	*Orycteropus afer*	aardvark	2
Hyracoidae	*Procavia capensis*	rock hyrax	146
Artiodactyla	*Phacochoerus africanus*	warthdog	253
	Giraffa camelopardalis	giraffe	655
	Aepyceros melampus	impala	173

Order	**Scientific name**	**Common name**	**N.**
	Connochaetes taurinus	wildbeest	447
	Gazella granti	Grant's gazelle	669
	Gazella thomsoni	Thomson gazelle	1295
	Litocranius walleri	gerenuk	138
	Madoqua kirkii	Kirk's dik-dik	3046
	Raphicerus campestris	steinbuck	333
	Syncerus caffer	african buffalo	691
	Tragelaphus imberbis	lesserkudu	450
	Tragelaphus scriptus	bushbuck	902
	Cephalophus harveyi	harvey's red duiker	434
	Oryx gazella	orix	3
	Kobus ellipsiprymnus	water buck	183

Dataset of species presence/absence was then processed in SOS to upload the values of environmental variables at each point location
(SOS GRASS script `V.WHATRAST.MANY.sh` in appendix B).
Finally the dataset was passed to the software R for the statistical analysis.

4.3. Lombardy region

The production of maps describing species potential and biodiversity was based on LRA. The dataset required for the analysis consisted of presence/absence wildlife data and environmental variables which described landuse, climate, and hydrography.
Landuse characteristics were generated from the informative base DUSAF (Regione Lombardia & ERSAF, 2000) and, since we wanted to use continuous data, distance maps for

each habitat type were computed.

In order to model climatic variables we looked for a network of data that covered the whole region. Complete dataset are available within the project "Regional Impact of Climatic Change in Lombardy Water Resources: Modelling and applications (RICLIC-WARM)" held by the University of Milan-Bicocca but, since we were not allowed to access these data, climatic variables were derived from the BIOCLIM project (Hijmans and Jarvis, 2005).

Relief was modeled using the raster layer at 20 m resolution produced by Direzione Generale Territorio e Urbanistica of Lombardy (Regione Lombardia, 2008).

On the other and the variable describing the road network was derived from the regional technical map of Lombardy with 10 m resolution 9Regione Lombardia, 1998).

Table 4.5.: Environmental variables used for LRA in Lombardy case study

Variable name	**Description**	**Source**
EV-DHYD	distance from lakes and rivers	CTR
EV-DINF	weighted distance from roads and railroads	CTR
EV-H100	snowfields and glaciers	DUSAF
EV-H103	broad-leaved forest	DUSAF
EV-H104	riparian vegetation	DUSAF
EV-H105	coniferous forest	DUSAF
EV-H106	mixed forest	DUSAF
EV-H107	recent reforestation	DUSAF
EV-H108	orchards and vegetable gardens	DUSAF
EV-H109	tree crops	DUSAF

Variable name	**Description**	**Source**
EV-H110	marshy vegetation	DUSAF
EV-H111	debris vegetation	DUSAF
EV-H112	riverbed vegetation	DUSAF
EV-H113	shrubs	DUSAF
EV-H114	meadows and pastures	DUSAF
EV-H115	herbaceous crops	DUSAF
EV-H116	herbaceous and tree crops	DUSAF
EV-H117	protected crops	DUSAF
EV-H119	debris and rock	DUSAF
EV-H120	dumps	DUSAF
EV-H121	continuous urban areas	DUSAF
EV-H122	open urban areas	DUSAF
EV-H123	farm building	DUSAF
EV-H126	airports	
EV-DTM	digital elevation model-elevation	CTR
EV-ASPECT	digital elevation model-aspect	CTR
EV-SLOPE	digital elevation model-slope	CTR
EV-CL15	Precipitation Seasonality(Coefficient of Variation)	
EV-CL18	Precipitation of warmest quarter	CTR
EV-CL01	Annual mean temperature	CTR
EV-CL07	Temperature annual range	CTR

The modeling process was applied to Reptiles and Amphibians; data came from the Atlas of Amphibians and Reptiles of Lombardy (Societas Herpetologica Italica sezione Lombardia - 2000).

Table 4.6 shows the list of species considered in this study. Point data do not represent real observation but refer to grid cells of 10 km^2. The dataset for each species was enlarged in order to be representative of the average ecological conditions of the cell area: for each presence point, 100 random presence points were generated within the cell of presence.

Since all the data set consisted of presence observations only, we computed random absence points in the whole region. The number of absence points exceeded the number of presence points by 30%.

This data treatment produced large datsets, up to thousands of points. The presence/absence dataset was prepared using the SOS script for GRASS `V.PRESABS.MANY.sh` (Appendix B)

Table 4.6.: List of species involved in the computation of potential distribution maps

Famiglia	**Scientific name**	**Common name**
Caudata	*Salamandra atra*	alpine salamander
Caudata	*Salamandra salamandra*	fire salamander
Caudata	*Salamandrina terdigitata*	spectacled salamander
Caudata	*Triturus alpestris*	alpine newt
Caudata	*Triturus carnifex*	italian crested newt
Caudata	*Triturus vulgaris*	smooth newt
Caudata	*Speleomantes strinatii*	Strinati's cave salamander
Anura	*Bombina variegata*	yellow bellied toad
Anura	*Pelobates fuscus*	spadefoot toad

Famiglia	**Scientific name**	**Common name**
Anura	*Bufo bufo*	common toad
Anura	*Bufo viridis*	green toad
Anura	*Hyla intermedia*	italian tree frog
Anura	*Rana dalmatina*	agile frog
Anura	*Rana italica*	italian stream frog
Anura	*Rana latastei*	italian agile frog
Anura	*Rana synk. esculenta*	edible frog
Anura	*Rana temporaria*	common frog
Cheloniidae	*Emys orbicularis*	european pond terrapin
Cheloniidae	*Trachemys scripta*	red-eared terrapin
Squamata(subord. Sauria)	*Tarentola mauritanica*	european leaf-toed gecko
Squamata(subord. Sauria)	*Anguis fragilis*	slow worm
Squamata(subord. Sauria)	*Lacerta bilineata*	western green lizard
Squamata(subord. Sauria)	*Podarcis muralis*	common wall lizard
Squamata(subord. Sauria)	*Zootoca vivipara*	viviparous lizard
Squamata(subord. Sauria)	*Chalcides chalcides*	three-toed skink
Squamata(subord. Serpentes)	*Coluber viridiflavus*	western whip snake
Squamata(subord. Serpentes)	*Coronella austriaca*	smooth snake

Famiglia	**Scientific name**	**Common name**
Squamata(subord. Serpentes)	*Coronella girondica*	southern smooth snake
Squamata(subord. Serpentes)	*Elaphe longissima*	aesculapian snake
Squamata(subord. Serpentes)	*Natrix maura*	viperine snake
Squamata(subord. Serpentes)	*Natrix natrix*	grass snake
Squamata(subord. Serpentes)	*Natrix tessellata*	dice snake
Squamata(subord. Serpentes)	*Vipera aspis*	asp viper
Squamata(subord. Serpentes)	*Vipera berus*	adder

Dataset of species presence/absence was then processed in SOS to upload the values of environmental variables at each point location (SOS GRASS script `V.WHATRAST.MANY.sh` in appendix B). Finally the dataset was passed to the software R for the statistical analysis.

5. Results

5.1. Spatial database

This application was developed in the context of our collaboration with the NGOs Istituto Oikos and Oikos East Africa within the project "Economic development and environmental redemption in Maasai pastoral areas of Arumeru District, Tanzania".

In order to support the District in the realization of Land Use plans at village level we created a spatial database to handle data about wildlife, population, agriculture and livestock. Socio-economic data were spatially related to subvillages while wildlife information consisted of field data and were collected along transect.

This database has been produced to become one of the instrument for the Planning Office of the Arumeru District Council for landuse planning in rural villages. The final users were not database experts, but rather district officers with only basic education in computer science, for this reason an user friendly solution with dedicated forms for database access was a central need to ensure effective use of the information system for data entry and retrieval to support management decisions.

We built an application upon a PostgreSQL/PostGIS spatial database; the basic entity of the database was the subvillage for which data about agriculture and livestock were collected every three months. A special interest of the administration was dedicated to the results of modernization programs that involved farmers and livestock keepers. Moreover statistical socioeconomic data were prepared by the government every year.

Wildlife data were not spatially related to administrative items but rather were spread

Figure 5.1.: Meru District officers taking part to database training

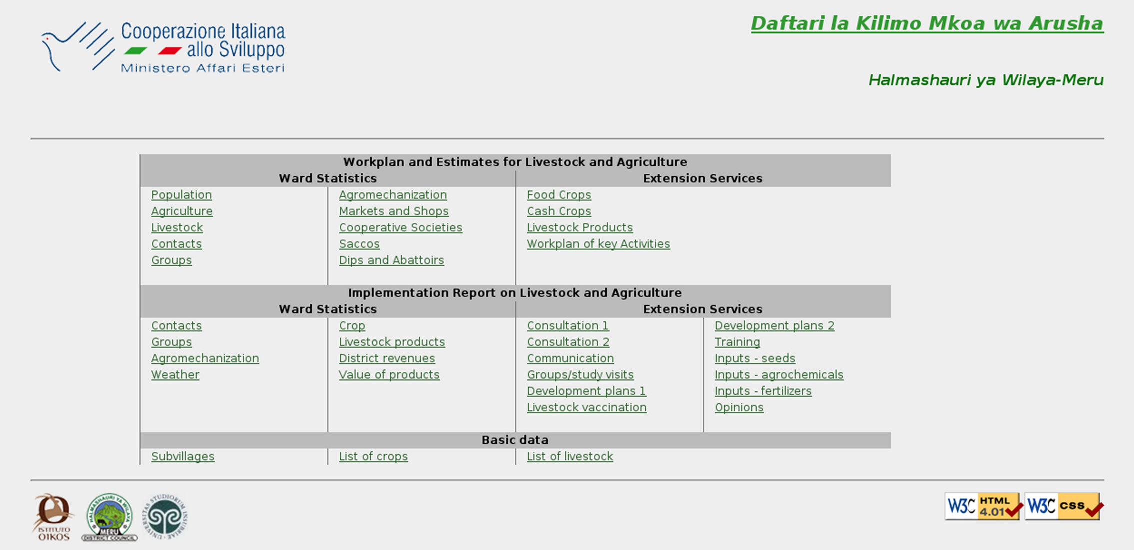
upon the entire study area along transects.

Different geometries and different topics co-exist in the spatial database. The database was composed of 96 relations and to each table corresponds a view. The web-database application consistsed of an Apache server web, 166 scripts in PHP language and 95 templates in HTML language.

Templates designed the appearance of web pages while PHP scripts enabled the connection between the client browser and the DBMS. PHP scripts acted like translators between the HTML language understood by the web browser and the SQL language understood by the DBMS. From the web side the database was accessed through an index page containing a list of items, where each item connected to a lister showing all the records of a single view of the database. From the lister it was possible to access to a form showing the details of each record for visualization and editing and to enter new data. Data of each view could be easily exported in Data Base IV (.dbf) format

The preparation of the system required a big effort in order to create PHP scripts for

database connection, the main difficulty being represented by the correct use of PEAR libraries objects types to access the different elements of the database.
As an advantage, once the system has been projected for one database, it is easy to adapt it to many other systems. On the other hand the system interface was very simple and after a five days training Arumeru District officers were able to enter, view, edit and retrieve data collected in the information system.



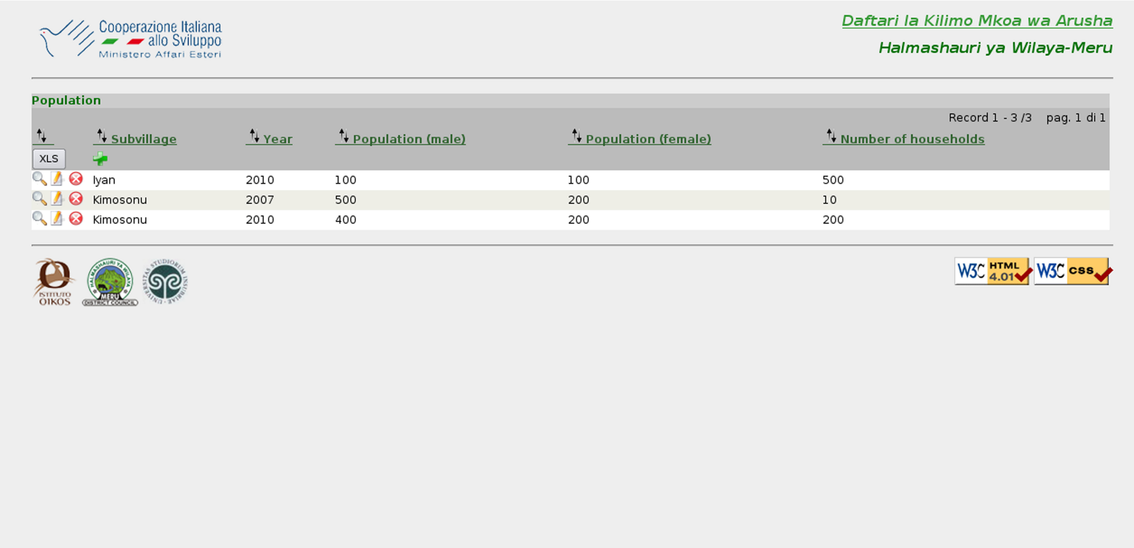
Cooperazione Italiana allo Sviluppo
Ministero Affari Esteri

Daftari la Kilimo Mkoa wa Arusha

Halmashauri ya Wilaya-Meru

Workplan and Estimates for Livestock and Agriculture			
Ward Statistics		**Extension Services**	
Population Agriculture Livestock Contacts Groups	Agromechanization Markets and Shops Cooperative Societies Saccos Dips and Abattoirs	Food Crops Cash Crops Livestock Products Workplan of key Activities	
Implementation Report on Livestock and Agriculture			
Ward Statistics		**Extension Services**	
Contacts Groups Agromechanization Weather	Crop Livestock products District revenues Value of products	Consultation 1 Consultation 2 Communication Groups/study visits Development plans 1 Livestock vaccination	Development plans 2 Training Inputs - seeds Inputs - agrochemicals Inputs - fertilizers Opinions
Basic data			
Subvillages	List of crops	List of livestock	

Figure 5.2.: Web side database access: the index page

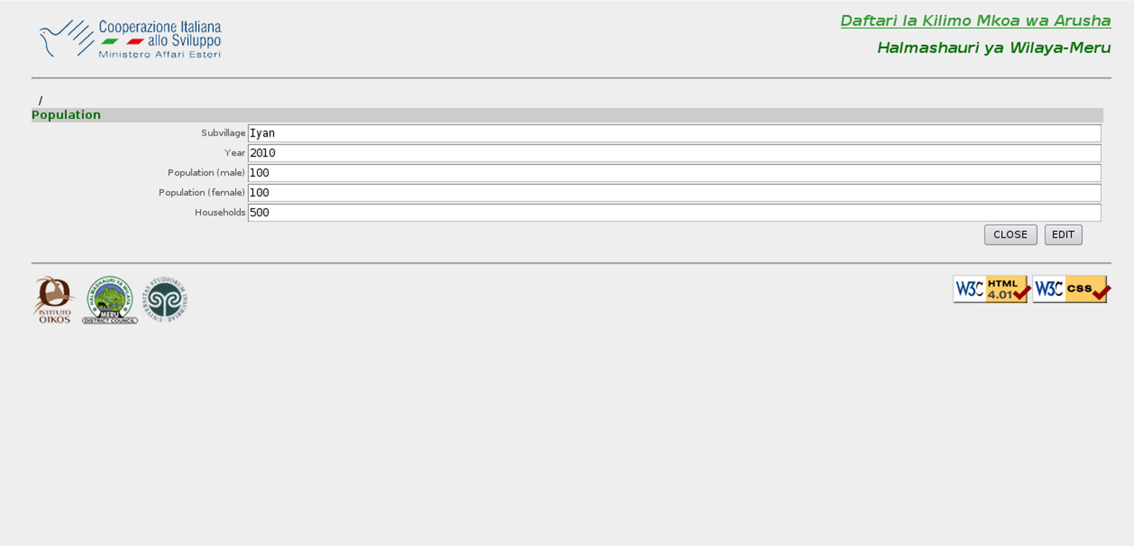


Population

Record 1 - 3 /3 pag. 1 di 1

Subvillage	Year	Population (male)	Population (female)	Number of households
Iyan	2010	100	100	500
Kimosonu	2007	500	200	10
Kimosonu	2010	400	200	200

Figure 5.3.: Web side database access: the table data lister



/

Population

Subvillage	Iyan
Year	2010
Population (male)	100
Population (female)	100
Households	500

Figure 5.4.: Web side database access: the data form of a single record

5.2. Spatial models

5.2.1. Brescia Plain (small scale)

The final aim of this study was the monitoring of the effects of landuse changes on wildlife and vegetation. Surveys and modeling regarded 88 species of which 3Amphibians, 76 Birds and 9 Mammals.

In order to model species potential distribution we first applied LRA, but the statistical analysis did not produce any acceptable results: stepwise selection of parameters didn't converge and ROC analysis showed overfitting of the data.

Models were then based on simple spatial overlay procedures based on GIS. The analysis was based on the spatial overlay of an informative base describing habitats and species presence points. The method expects the reclassification of land use map in the sense that each habitat type is given a value equal to 1 if one or more presence points fall within its boundary.

For each species the technique outputs a dichotomous map of potential distribution where 1 indicates presence and 0 indicates absence.

Habitat suitability models computed for each species were weighted in order to consider priorities; the species weight in conservation is established, for Lombardy, by a Regional law (Delibera della Giunta Regionale n.4345 del 20 Aprile 2001).

The analysis was carried out in SOS using the script `CALC.OVERLAY.sh`. The weighted sum of 88 habitat suitability maps produced the wildlife value map showed in Figure 5.5.

Survey also regarded vegetation: field surveys produced phytosociological data, floristic lists and lists of valuable trees that were processed together with a land cover layer with the spatial overlay method (Figure 5.6.)

The sum of the two maps of wildlife and vegetation value produced a synthesis map of ecosystem value showed in Figure 5.7.

Figure 5.5.: Wildlife value map of the Brescia plain

These maps refer to the *ante operam* stage of the monitoring program; in the next years new models will be computed according to new surveys and results will be compared in order to assess the environmental impact of the highway building site (*in opera* survey) and of the highway employment (*post operam* survey).

The main advantage of using the SOS system in this study is that it provided a coded methodology that can be reproduced for subsequent analysis in order to obtain comparable results. Moreover scripting assures automated model computation for repetitive procedures.

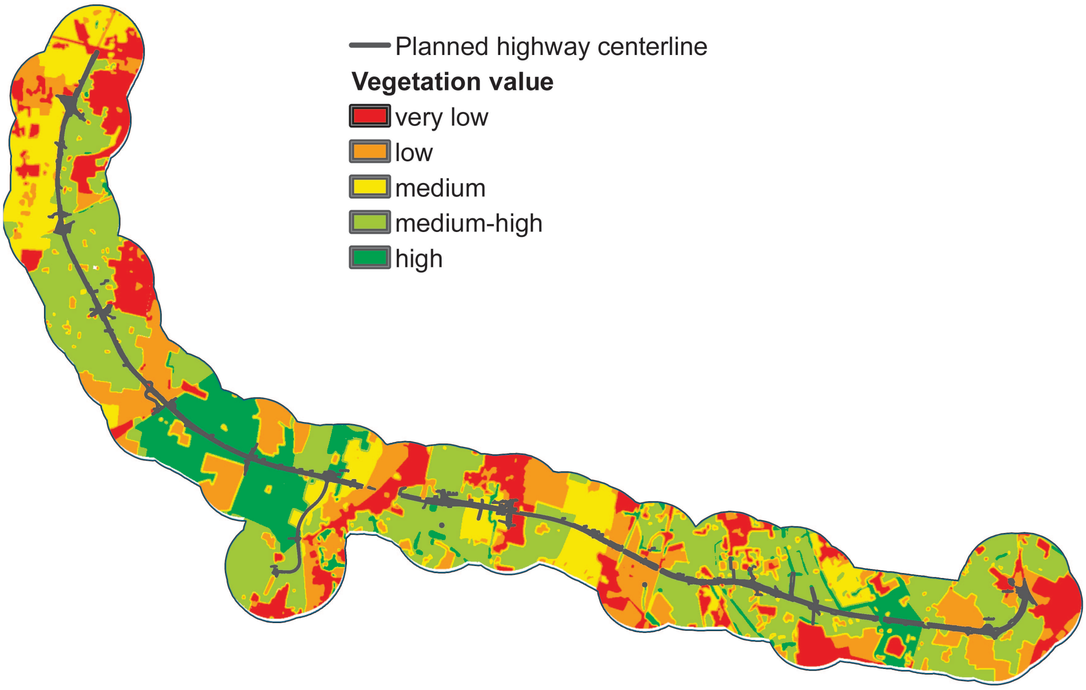


Figure 5.6.: Vegatation value map of the Brescia plain

5.2.2. Mount Meru region - Tanzania (medium scale)

This study was directed to mammal biodiversity assessment for land use planning. Potential distribution models for 35 mammals species were computed through LRA.

All the environmental variables listed in table A.1 have been selected in the models of species distribution, but with different frequencies.

Table 5.1 lists, for each environmental variable, the percentage of species models for which it was selected. The average percentage of selection of the variable was 36.7; the least selected variable was EV-TSLP, the variable related to the digital terrain model describing slopes, while the most selected (57.6%) was EV-H122, the habitat variable related to open schrubs with sparse trees. In addition three variables were retained for over 50% of the species: continuous trees forest with shrubs (EV-H146), shrubs with

Figure 5.7.: Ecosystem value map of the Brescia plain

herbaceous layer on temporarily swampy area (EV-H180) and natural lakes (EV-H193). The model for *Felys silvestrys* (wildcat) retained the minimum number of variables, only 4, while *Gazella granti*selected 80% of the set of variables. Other species selecting over the 70% were *Giraffa camelopardalis*, *Lepus saxatilis* and *Madoqua kirkii*. The mean number of variables retained by species models was 24.5.

Table 5.1.: selection of environmental variables in LRA

Environmental variable	**Percentage of selection**
EV-CL08	24.2

Environmental variable	**Percentage of selection**
EV-CL09	33.3
EV-CL16	33.3
EV-CL17	33.3
EV-DOMI	39.4
EV-SHAN	18.2
EV-EDDE	33.3
EV-PATC	24.2
EV-LAKE	39.4
EV-RVRS	33.3
EV-ROAD	48.5
EV-SWPA	39.4
EV-VLLG	36.4
EV-TASP	27.3
EV-TSLP	18.2
EV-H101	33.3
EV-H105	36.4
EV-H108	36.4
EV-H113	42.4
EV-H115	45.5
EV-H117	39.4
EV-H122	57.6
EV-H126	42.4
EV-H127	36.4
EV-H131	42.4
EV-H132	45.5
EV-H134	39.4

Environmental variable	Percentage of selection
EV-H141	30.3
EV-H143	39.4
EV-H145	45.5
EV-H146	51.5
EV-H154	36.4
EV-H162	36.4
EV-H165	33.3
EV-H166	33.3
EV-H170	36.4
EV-H174	48.5
EV-H175	27.3
EV-H177	30.3
EV-H180	51.5
EV-H182	36.4
EV-H190	39.4
EV-H193	51.5
EV-H196	24.2
EV-H202	36.4
EV-H204	36.4
EV-H205	30.3
EV-H207	24.2
EV-H211	33.3
EV-H214	45.5
EV-H216	30.3
EV-H218	42.4

AUC values of the ROC analysis computed for all the species are listed in table 5.2. All the species obtained AUC scores superior to 0.5 and thus models are considered better than random guess.
Six species (*Acynonix jubatus, Lycaon pictus, Felis silvestrys, Hyaena hyaena, Orycteropus afer, Oryx gazella* and *Hystrix cristata*) obtained AUC values equal to 1 thus meaning over-fitting of the data. This is a typical result when running LRA with a reduced number of presence data and, as a matter of fact these species were characterized by a number of observation comprised between 2 (*Orycteropus afer*) and 14 (*Lycaon pictus*). These six models were rejected.
The rest of the species obtained AUC scores comprised between 0.62 and 0.99, stating the reliability of computed models.

Table 5.2.: ROC plot AUC values for LRA

Scientific name	**AUC**	**model acceptance**
Canis aureus	0.762 ±0.002	*
Canis mesomelas	0.772 ±0.009	*
Lycaon pictus	1	
Otocyon megalotis	0.725 ±0.003	*
Acynonix jubatus	1	
Felis silvestrys	1	
Crocuta crocuta	0.653 ±0.001	*
Hyaena hyaena	1	
Civettictis civetta	0.953 ±0.003	*
Equus burchelli	0.932 ±0.008	*
Orycteropus afer	1	
Lepus saxatilis	0.830 ±0.003	*

Scientific name	**AUC**	**model acceptance**
Pronolagus rupestris	0.834 ±0.007	*
Cercopithecus mitis	0.803 ±0.002	*
Papio cynocephalus	0.820 ±0.002	*
Colobus guereza	0.793 ±0.007	*
Procavia capensis	0.770 ±0.002	*
Phacochoerus africanus	0.752 ±0.003	*
Giraffa camelopardalis	0.890 ±0.003	*
Aepyceros melampus	0.730 ±0.001	*
Connochaetes taurinus	0.797 ±0.002	*
Gazella granti	0.870 ±0.003	*
Gazella thomsoni	0.930 ±0.002	*
Litocranius walleri	0.713 ±0.004	*
Madoqua kirkii	0.922 ±0.006	*
Raphicerus campestris	0.780 ±0.003	*
Syncerus caffer	0.765 ±0.001	*
Tragelaphus imberbis	0.682 ±0.002	*
Tragelaphus scriptus	0.710 ±0.003	*
Cephalophus harveyi	0.740 ±0.006	*
Oryx gazella	1	
Kobus ellipsiprymnus	0.643 ±0.002	*
Loxodonta africana	0.995 ±0.001	*
Hystrix cristata	1	

Prevalence is an important factor determining the possibility to compute LRA models: species with very low prevalence obtained AUC scores equal to 1 thus indicating over-fitting of the model due to insufficient data sample. The minimum sample size resulting in an acceptable model is 102 and refers to the spotted hyena (*Crocuta crocuta*), for which an AUC value of 0.65 was obtained.

Regression coefficients of LRA, ROC plots and final maps for each specie are shown in appendix A.

We produced a species richness map resulting from the sum of 28 potential distribution maps. We selected only those species for which the statistical analysis produced an acceptable result in terms of AUC ($0.5 < AUC < 1$); species considered for the computation of species richness are pointed by the symbol * in table 5.2.

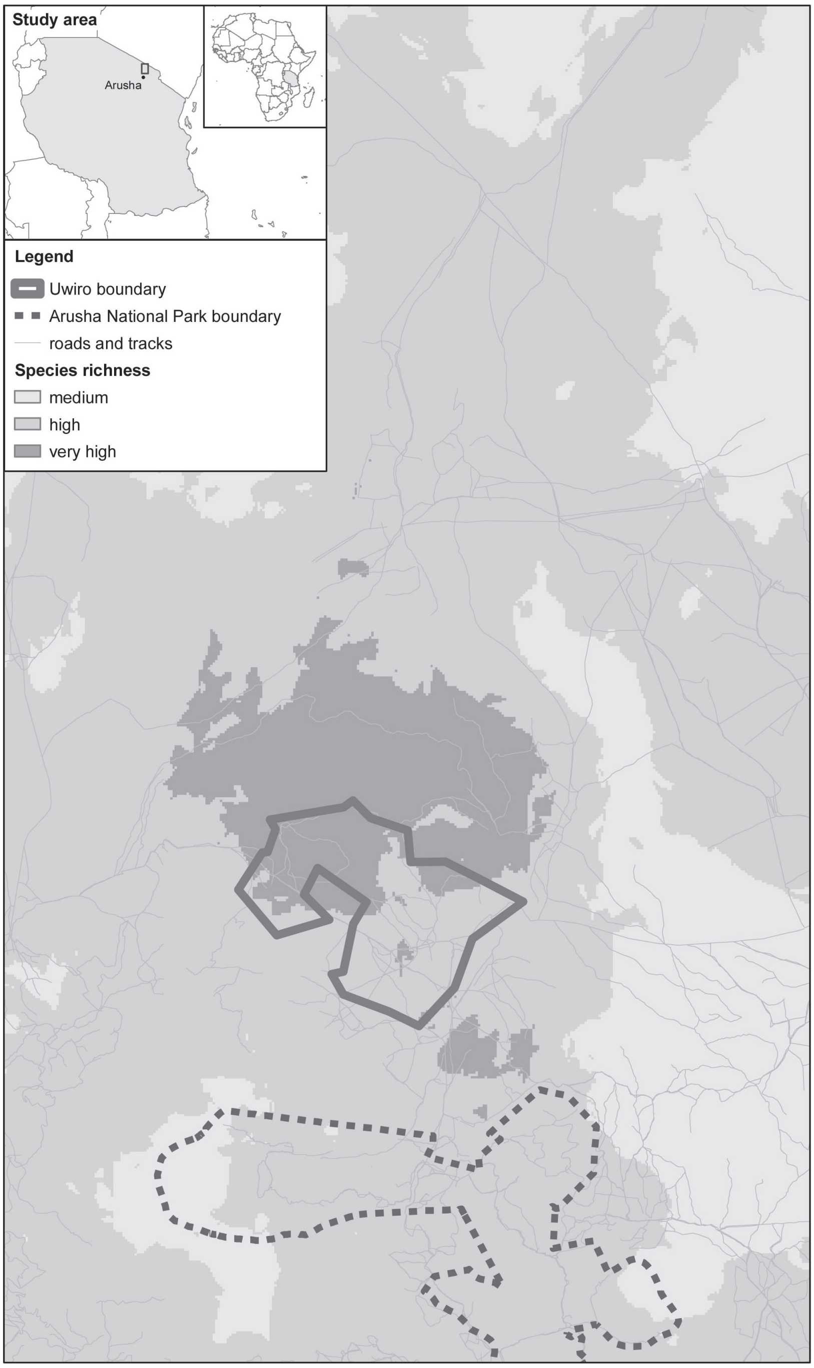


Figure 5.8.: Species richness computed from species potential distribution maps(n = 28)

Biodiversity was classified into 3 classes with the Jenks Natural Breaks Classification (Jenks, 1967); the area characterized by highest biodiversity corresponds to the central part of the study area. Graph of figure 5.9 shows the distribution of habitat types in each biodiversity class: open scrubs (habitat 101), with an extension of 11105 ha, are the most extensive and representative habitat in class 3 (very high biodiversity), but it's also the most common habitat in class 1 (medium biodiversity) and it's extensively present in class 2 (high biodiversity).

The other habitats present in class 3 are: continuous herbaceous vegetation-100 (762 ha), shrubs with herbaceous layer on temporarily swampy area-105 (596 ha), grassland on temporarily/permanently swampy area-104 (449 ha), continuous forest-102 (49 ha) and continuous rainfed herbaceous crops (36)-108.

All these habitats are also present, with wider extension, in class 1 and 2. Three habitats present in class 1 and 2 are not present in class 3, these are: bare rocks (329 ha), natural lakes (325),continuous forest plantations of rainfed Pine (2320 ha).

One more habitat is present in class 2 and not in class 3: continuos rainfed tree and herbaceous crops/continuous herbaceous vegetation (9424 ha) Comparing class 3 and class 1, the major difference in habitat extension regards habitat 100 (continuous herbaceous vegetation), 102 (continuos forest) and 108 (continuous rainfed herbaceous crops). The comparison of class 3 and class 2 underlines major differences in habitat extension for habitat 102 (continuos forest).

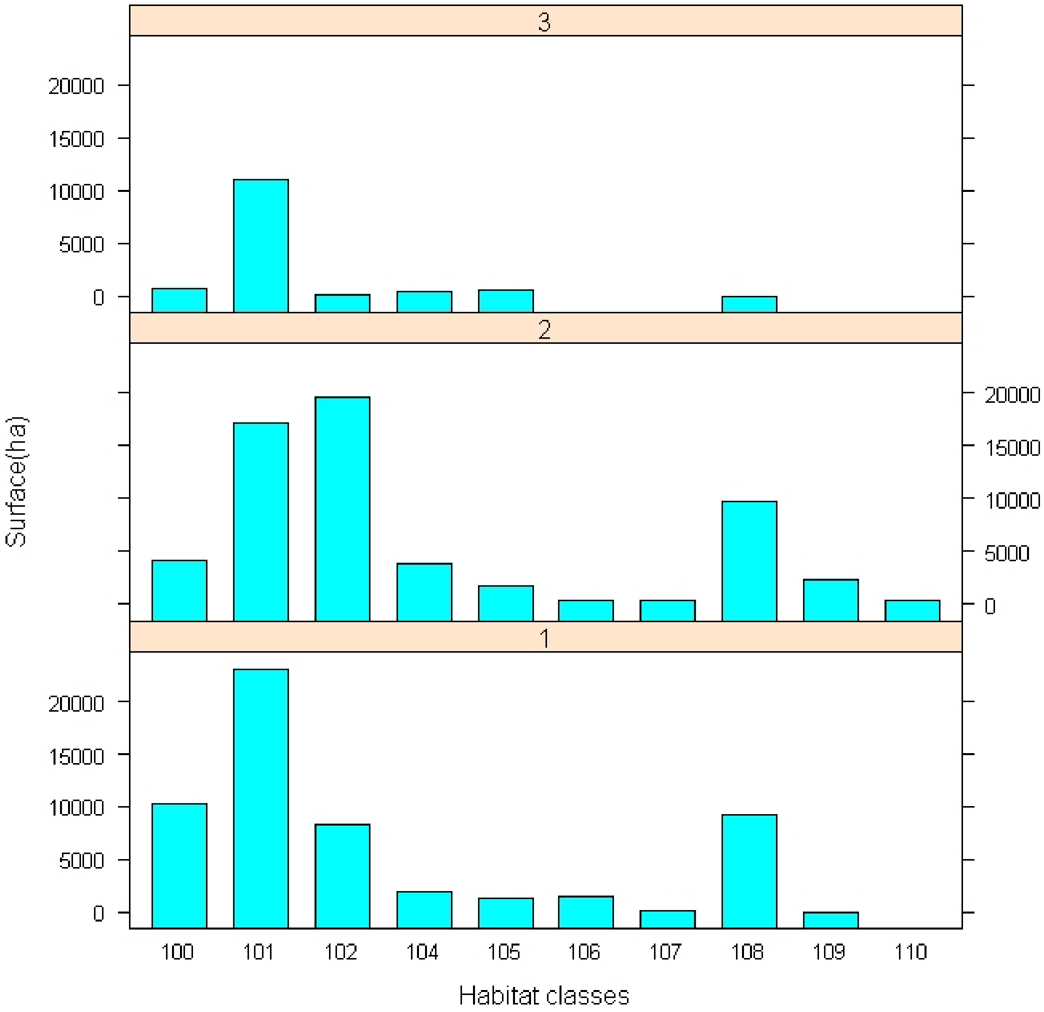


Figure 5.9.: Distribution of habitat types among classes of biodiversity.

The village of Uwiro falls completely in a territory with high (58%) and very high (42%) biodiversity value.

The north-western part of Uwiro village has very high biodiversity. This area corresponds to the subvillage of Mkuru and to the western sector of Kiamakata and Iyan subvillages. Table 5.3 shows, for each subvillage, land distribution into biodiversity classes: the Subvillages of Mkuru, Kiamakata and Iyan show the highest percentage of territory with very high biodiversity ($\simeq 60\%$).

Table 5.3.: Distribution of subvillage territory into biodiversisity classes in Uwiro Village

Subvillage name	**Biodiversity class**		
	Very High	**High**	**Medium**
Iyan	57%	43%	0%
Kiamakata	57%	43%	0%
Kimosonu	29%	71%	0%
Mkuru	60%	40%	0%
Nkuuny	3%	97%	0%
Total	42%	58%	0%

5.2.3. Lombardy region (big scale)

We applied LRA model to 34 species of Amphibians and Reptiles using 31 environmental variables as predictors. All the environmental variables listed in table A.1 have been selected in the models of species distribution, but with different frequencies.

Table 5.4 lists, for each environmental variable, the percentage of species models for which it was selected. The average percentage of selection of each variable was 63.7; the least selected variable was EV-SLOPE, the variable describing slopes, while the most selected (92%) was EV-H100, the habitat variable describing the distribution of snowfields and glaciers.

In addition three other variables were retained for over 80% of the species: broad-leaved forest (EV-H103), herbaceous and tree crops (EV-H116) and debris and rock (EV-H119). The model for the species *Speleomantes strinatii* retained the minimum number of variables, only 10, while 70% of the species selected more the 60% of the variables set. Selected variables resulted significant in most of the models.

Table 5.4.: selection of environmental variables in LRA

Environmental variable	**Percentage of selection**
EV-DHYD	60
EV-DINF	44
EV-H100	92
EV-H103	80
EV-H104	60
EV-H105	68
EV-H106	72
EV-H107	76
EV-H108	56
EV-H109	72
EV-H110	72
EV-H111	72
EV-H112	72
EV-H113	40
EV-H114	64
EV-H115	60
EV-H116	64

Environmental variable	**Percentage of selection**
EV-H117	80
EV-H119	84
EV-H120	52
EV-H121	48
EV-H122	52
EV-H123	60
EV-H126	60
EV-DTM	60
EV-ASPECT	44
EV-SLOPE	16
EV-CL15	16
EV-CL18	80
EV-CL01	68
EV-CL07	68

AUC values of the ROC analysis are listed in table 5.5. All the species obtained AUC scores superior to 0.5, therefore models are considered better than random guessing. None of the species had an AUC equal to 1 (meaning over-fitting of the data); this is probably a consequence of using large datasets of presence and absence. All the species obtained AUC scores comprised between 0.65 and 0.99, stating the reliability of computed models.

Table 5.5.: ROC plot AUC values for LRA

Scientific name	**AUC**
Salamandra atra	0.802131 ±0.003

Scientific name	**AUC**
Salamandra salamandra	0.761 ±0.004
Salamandrina terdigitata	0.780 ±0.003
Triturus alpestris	0.871 ±0.004
Triturus carnifex	0.646 ±0.004
Triturus vulgaris	0.762 ±0.004
Speleomantes strinatii	0.849 ±0.021
Bombina variegata	0.821 ±0.003
Pelobates fuscus	0.731 ±0.007
Bufo bufo	0.764 ±0.003
Bufo viridis	0.773 ±0.003
Hyla intermedia	0.739 ±0.004
Rana dalmatina	0.839 ±0.003
Rana italica	0.949 ±0.003
Rana latastei	0.807 ±0.003
Rana synk. esculenta	0.663 ±0.003
Rana temporaria	0.937 ±0.004
Emys orbicularis	0.723 ±0.004
Trachemys scripta	0.688 ±0.004
Anguis fragilis	0.738 ±0.003
Lacerta bilineata	0.759 ±0.004
Podarcis muralis	0.663 ±0.004
Zootoca vivipara	0.734 ±0.001
Chalcides chalcides	0.964 ±0.002
Coluber viridiflavus	0.679 ±0.004
Coronella austriaca	0.804 ±0.004
Coronella girondica	0.994 ±0.001

Scientific name	AUC
Elaphe longissima	0.767 ±0.003
Natrix maura	0.974 ±0.002
Natrix natrix	0.912 ±0.002
Natrix tessellata	0.695 ±0.004
Vipera aspis	0.785 ±0.003
Vipera berus	0.925 ±0.002

6. Conclusion

This project had three research aims:

1. explore and evaluate the applicability of Generalized Linear Models at different scales;
2. automation of statistical analysis and spatial processing;
3. use of spatial DBMS to handle wildlife data.

GLM were applied with good results at middle and big scale while at small scale it was not possible to compute models. The applicability of GLM at different scale seems to be related especially to habitat variability.

The small scale study area, corresponding to a 87 km^2 area in the Brescia plain, is characterized by a monotonous environment represented by monocultures and reduced natural landscape; in this conditions the stepwise analysis was not able to converge and to find the predictable variables for the regression analysis.

A focal point seems to be represented by the informative layers used to model the environment. In lack of high resolution land use maps built expressly for this application through interpretation of detailed aerial photographs, the modeling process was rather based on the informative base DUSAF which was produced at 1:25000 scale, and thus less detailed.

The analysis of reduced size areas requires the use of descriptors with small scale details in order to detect differences in landscape and environment. Unfortunately, the need for speditive analysis within low budget projects often forces the use of cartographic layers built at regional level. The application of GLM to the medium size study area of the Mount Meru ecosystem (784 km^2) produced good results in terms of the capacity to discriminate presence and absence areas for each species.

AUC scores were comprised between 0.62 and 0.99 while only six models on thirty-four were rejected since they obtained AUC values equal to 1, that is, were affected by overfitting of the data. These six species were characterized by very low prevalence (medium sample size equal to 2.66).

In this geographical context the size of the area is large enough to show habitat variability and thus, at this scale of analysis, the informative base used as predictor dataset was suitable to model species distribution. We used cartographic layers produced at a scale of 1:200000 Africover, a FAO project.

Finally the application of GLM to the large scale case study of Lombardy region produced very good results. It was possible to model potential distribution for all the 33 species of Amphibians and Reptiles analyzed; ROC analysis succeeded in discriminating presence and absence habitats for all the specie: AUC values were comprised between 0.65 and 0.99.

Since it was not possible to apply GLM to the Brescia plain case study, an empirical method based on spatial overlay was applied. A routine was written in order to synthetize potential distribution of 107 species in a unique map of wildlife value.

Species distribution models are an essential tool for decision support, especially in environmental evaluation: if the dataset does not allow the application of a statistical model, an empirical model based on real data and on repeatable and standardized methods represent a good substitute to take account of wildlife in monitoring projects.

The application of GLM to the Mount Meru and Lombardy case studies resulted in maps of species potential distribution suitable for planning and management purposes. Twenty-nine models of mammals distribution in the Mount Meru area were synthetized in a biodiversity map that was used to analyze wildlife distribution within the Uwiro

village which underwent a process of landuse planning.
The Thirty-three species models computed for Lombardy represent a cartographic base useful for future planning and management projects within the region; atlases vector data format are not directly usable for environment evaluation models, while species distribution models in raster datasets can be processed in map algebra computational processes.

The analysis system we built (SOS - Species Open Spreader) gave us the possibility to compute models for a large number of species for different applications. It would not have been possible to conduct an interactive process due to the large number of species and to to the large number of carthographic layers representing the dependent variables dataset. Once the dataset is prepared, SOS allows the automated computation of statistical and empirical models and also offers routines for data preparation in GIS environment.
At the moment datasets preparation in SOS is not completely automated but still requires a high degree of interaction. This is due in particular to the high heterogeneity of input species data: field data along transects, occasional observations, grids of presence in atlases etc. Automation is more advanced as regards the preparation of environmental variables datasets. In the next years it will be important to improve automated geoprocessing of datasets.

Finally we tested the use Spatial DBMS, to handle wildlife data. Either SpatiaLite or PostegreSQL-PostGIS represent optimal solutions to store heterogeneous wildlife data in a standardized repository. They both offer SQL tools for data query, import and output and represented a good solution for spatial data visualization.
SpatiaLite is a very light and simple structure software solution consisting in a unique cross-platform file easy to transfer, but is still lacking support to properly interface with LAMP

In this research project one of the aims was the development of a web application with forms for data input and output. This was a focal point in the research in consideration of technology transfer to users not familiar to DBMS and SQL.
The web application was built with the LAMP system. The need for integration of the spatial DBMS into a web application induced us to use PostegreSQL-PostGIS. The reason is that the integration of SpatiaLite in the web-database application was difficult to achieve as SpatiaLite is still little supported by PEAR libraries that are fundamental in PHP scripts for client-database connection.
The design and code development of the web-database applications gave us the possibility to achieve very good results in terms of usability of spatial database. A spatial database and a web database application for the management of wildlife and socio-economic data in the Meru District (Tanzania) was developed. After a short course, officers of the Meru District were able to correctly enter and retrieve data.

Every part of this research project was developed using Free and Open Source Software. All the research issues had implications in professional conservation actions.
The possibility to develop code for automated spatial analysis and flexible dataware-house makes FOSS an indispensable instrument for the development of conservation projects not only in academic environments but also for professional applications.

Acknowledgements

A sincere thank
to Damiano Preatoni, for trusting me, for always answering my questions and to be a constant reference;
to Istituto Oikos Onlus and Oikos East Africa, in particular Guido Tosi, Rossella Rossi and Caterina Carugati, for the great experience in Tanzania and to the African staff for the warm hospitality;
to the development teams of the Open Source Software I used in my project: your work has an inestimable value;
to all the UAGRA group, for the moments of work and enjoyment we shared;
to my sisters Nadia and Cristina, for the English revision.
Thanks to my family, for their support and their patient wait.

Bibliography

[1] Akaike, H. (1980). *Likelihood and the Bayes procedure, Bayesian Statistics*. University Press.

[2] Araujo, M. B. and Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal Of Biogeography*, 33(10):1677Ű1688.

[3] Barbosa, A. M., Real, R., and Vargas, J. M. (2010). Use of coarse-resolution models of species' distributions to guide local conservation inferences. *CONSERVATION BIOLOGY*, 24(5):1378–1387.

[4] Basille, M., Calenge, C., Marboutin, ., Andersen, R., and Gaillard, J. (2008). Assessing habitat selection using multivariate statistics: Some refinements of the ecological-niche factor analysis. *Ecological Modelling*, 211(1-2):233–240.

[5] Báez, J., Olivero, J., Peteiro, C., Ferri-Yáñez, F., Garcia-Soto, C., and Real, R. (2010). Macro-environmental modelling of the current distribution of <i>undaria pinnatifida</i> (laminariales, ochrophyta) in northern iberia. *Biological Invasions*, 12:2131–2139. 10.1007/s10530-009-9614-1.

[6] Bibby, C. J., Burgess, N. D., Hill, D. A., and Mustoe, S. H. (2000). *Bird census etchniques*. Academic Press, London, UK.

[7] Bivand, R. (2009). *spgrass6: Interface between GRASS6 and R.*, r packege version 0.6-9 edition.

[8] Bolick, M. R. (1974). A vegetation history of the mount meru lahar, tanzania. Master's thesis, A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Zoology in the Graduate School of Arts and Science of Duke University.

[9] Brotons, L., Thuiller, W., Araújo, M., and Hirzel, A. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27(4):437–448.

[10] Cagnacci, F. and Urbano, F. (2008). Managing wildlife: A spatial information system for gps collars data. *Environmental Modelling & Software*, 23(7):957–959.

[11] Calenge, C. (2006). The package adehabita for the r software: tool for the analysis of space and habitat use by animals. *Ecological Modelling*, 197:1035.

[12] Calenge, C. (2008). A general framework for the statistical exploration of the ecological niche. *Journal of theoretical biology*, 252.4:674–685.

[13] Cavallini, P. and Tufto, J. (2005). Should wildlife biologist use free software? *Wildlife Biology*, 11:67–76.

[14] CGIAR, C. f. S. I. C.-C. (2004). Cgiar - csi srtm 90m dem digital elevation database. http://srtm.csi.cgiar.org.

[15] Cianfrani, C.and Le Lay, G., Hirzel, A. H., and Loy, A. (2010). Do habitat suitability models reliably predict the recovery areas of threatened species? *Journal of Applied Ecology*, 47:421 – 430.

[16] Codd, E. (1969). Derivability, redundancy, and consistency of relations stored in large data banks. Research report, IBM.

[17] Eckerson, W. W. (1995). Three tier client/server architecture: Achieving scalability, performance, and efficiency in client server applications. *Open Information Systems*, 10:3(20).

[18] Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M., Peterson, A. T., Phillips, S. J., Richardson, K., Scachetti-Pereira, R., Schapire, R. E., Soberon, J., Williams, S., Wisz, M. S., and Zimmermann, N. E. (2006). Novel methods improve prediction of speciesŠ distributions from occurrence data. *Ecography,*, 29(2):129Ű151.

[19] Engler, R., Guisan, A., and Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41(2):263–274.

[20] FAO (2003). Africover. http://www.africover.org/.

[21] Fielding, A. H. and Bell, J. F. (1997a). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1):38Ű49.

[22] Fielding, A. H. and Bell, J. F. (1997b). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1):38Ű49.

[23] Freeman, E. (2007). *PresenceAbsence: An R Package for Presence-Absence Model Evluation.* USDA Forest Service, Rocky Mountain Research Station, 507 25th street, Ogden, UT, USA.

[24] Freeman, E. A. and Moisen, G. (2008). Presenceabsence: An r package for presence absence analysis. *Journal of Statistical Software*, 23(11):1Ű31.

[25] Furieri, A. (2009). Spatialite 2.3.1, a complete spatial dbms in a nutshell. http://www.gaia-gis.it/spatialite/.

[26] Gibson, L., Barrett, B., and Burbidge, A. (2007). Dealing with uncertain absences in habitat modelling: a case study of a rare ground-dwelling parrot. *Diversity And Distributions*, 13(6):704Ũ713.

[27] GRASS, D. T. (2007). *GRASS 6.2 Users Manual.* ITC-irst, Trento, Italy.

[28] Green, J. L., Hastings, A., Arzberger, P., Ayala, F. J., Cottingham, K. L., Cuddington, K., Davis, F., Dunne, J. A., Fortin, M. J., Gerber, L., , and Neubert, M. (2005). Complexity in ecology and conservation: Mathematical, statistical, and computational challenges. *Bioscience*, 55(6):501Ũ510.

[29] Group, P. G. D. (1996-2010). www.postgresql.org.

[Group] Group, T. P. G. D. Postgresql version 8.3.5. http://www,postgresql.org/.

[31] Guisan, A. and Thuiller, W. (2005). Predicting species distribution: o?ering more than simple habitat models. *Ecology Letters*, 8(9):993Ũ1009.

[32] Hijmans, R.J., S. C. J. P.-P. J. and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25:1965–1978.

[33] Hirzel, A. H., Hausser, J., Chessel, D., , and Perrin, N. (2002). Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology*, 83(7):2027Ũ2036.

[34] Hirzel, A. H., Helfer, V., , and Metral, F. (2001). Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, 145(2-3):111Ũ12.

[35] Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., and Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, 199(2):142Ũ 152.

[36] Hosmer, D. W. & Lemeshow, S. (2000). *Applied Logistic Regression.*

[37] Hu, J. and Jiang, Z. (2010). Predicting the potential distribution of the endangered przewalski's gazelle. *Journal of Zoology*, 282.

[38] IUCN (2001). *IUCN Red list Categories and Criteria.* Species Survival Commission, World Conservation Union (IUCN), Gland, Switzerland and Cambridge UK, version 3.1 edition.

[39] Jenks, G. F. (1967). The data model concept in statistical mapping. *International Yearbook of Cartography*, 7:186–190.

[40] Jepsen, J., Topping, C., Odderskær, P., and Andersen, P. (2005). Evaluating consequences of land-use strategies on wildlife populations using multiple-species predictive scenarios. *Agriculture, Ecosystems & Environment*, 105(4):581 – 594.

[41] Kidane, A. (1974). Interaction of livestock, game and vegetation in longido controlled hunting area. Master's thesis, Mweka Wildlife Management College.

[42] Kiunsi, R. (1993). Land cover change in the longido game controlled area and their possible effect on the function of the amboseli biosphere reserve. Master's thesis, International Institute for Aerospace survey and earth Science, Netherlands. Unpublished MasterŠs thesis.

[43] Lasan, A. (1971). Formulation of management plan for the longido controlled area, ngaserai. Master's thesis, Diploma theis submitted for the fulfillment of an award of diploma in Wildlife Management, Mweka.

[44] Li, J. (2008). Biodiversity and conservation. *Biodiversity and conservation*, 17(13):3079–3095.

[45] Lipsitz, S. R., Laird, N. M., Brennan, T. A., and Parzen, M. (2001). Estimating the kappa-coefficient from a selected sample. *The Statistician*, 50(4):407–416.

[46] Liu, C. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3):385–393.

[47] Luoto, M. (2000). Modelling of rare plant species richness by landscape variables in an agriculture area in finland. *Plant Ecology*, 149(2):157–168.

[48] Manel, S., Williams, H. C., and Ormerod, S. J. (2001a). Evaluating presence-absence models in ecology: The need to account for prevalence. *The Journal of Applied Ecology*, 38(5):921–931.

[49] Manel, S., Williams, H. C., and Ormerod, S. J. (2001b). Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal Of Applied Ecology*, 38(5):921Ű931.

[50] Martinez, J. C., Coll, E., and Irigoyen, J. (2005). Spatial analysis using opengis specifications: 'simple features sql'. *WSEAS Transactions on Information Science and Applications*, 2(4):409–15.

[51] McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models -Monographs on Statistics and Applied Probability.* Chapman & Hall, London, UK, 2nd edition edition.

[52] Meynard, C. N. and Quinn, J. F. (2007). Predicting species distributions: a critical comparison of the most common statistical models using arti?cial species. *Predicting species distributions: a critical comparison of the most common statistical models using arti?cial species.*, 34(8):1455Ű1469.

[53] Millington, J., Walters, M., Matonis, M. S., and Liu, J. (2010). Effects of local and regional landscape characteristics on wildlife distribution across managed forests. *Forest Ecology and Management*, 259(6):1102 – 1110.

[54] Mörtberg, U., Balfors, B., and Knol, W. (2007). Landscape ecological assessment:

A tool for integrating biodiversity issues in strategic environmental assessment and planning. *Journal of Environmental Management*, 82(4):457 – 470.

[55] Newbold, T., Reader, T., El-Gabbas, A., Berg, W., Shohdi, W. M., Zalat, S., El Din, S., and Gilbert, F. (2010). Testing the accuracy of species distribution models using species records from a new field survey. *Oikos*, 119:1326 – 1334.

[56] Ntalwila, J. (2007). *Wildlife diversity and their interaction with people for sustainable use in semi arid areas of northern Mount Meru Ecosystem (Tanzania).* PhD thesis, Analysis, Protection and Management of Biodiversity, XX Course.

[57] Oikos, I. (2004). Mmcp (mount meru conservation project) 2004. long term wildlife monitoring in arusha national park, tanzania. Technical report. A report submitted by Istituto Oikos for the European Commission, 109 pp.

[58] Oikos, I. (2008). Raccordo autostradale ospitaletto-montichiari-piano di monitoraggio ambientale (p.m.a.). monitoraggio ante operam. componente vegetazione, flora, fauna e ecosistemi. Technical report.

[59] Olivier, F. and Wotherspoon, S. J. (2006). Modelling habitat selection using presence-only data: Case study of a colonial hollow nesting bird, the snow petrel. *Ecological Modelling*, 195(3-4):187Ű204.

[60] Ottaviani, D., Lasinio, G. J., , and Boitani, L. (2004). Two statistical methods to validate habitat suitability models using presence-only data. *Ecological Modelling*, 179(4):417Ű443.

[61] Pearce, J. and Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133(3):225Ű245.

[pgAdmin Development Team] pgAdmin Development Team. pgadmin postrgresql tools (pgadmin). version 1.8.4. http://www.pgadmin.org/.

[63] Poole, J. and Reuling, M. (1997). A survey of elephant and other wildlife of the west kilimanjaro basin, tanzania. 57 pp. Technical report.

[64] Prigioni, C., Cantini, M., and Zilio, A. (2001). *Atlante dei Mammiferi della Lombardia.*

[65] R, C. D. T. (2010). *R: A Language and Environment for Statistical Computing.* Vienna, Austria.

[66] Rebelo, H. and Jones, G. (2010). Ground validation of presence-only modelling with rare species: a case study on barbastelles barbastella barbastellus (chiroptera: Vespertilionidae). *JOURNAL OF APPLIED ECOLOGY*, 47(2):410–420.

[Refraction Research Inc.] Refraction Research Inc., Victoria, B. C.-C. Postgis version 1.4. htto://postgis.refractions.net/.

[68] RIichter, R. and Ford, R. (1994). An object-oriented characterization of spatial ecosystem information. *Mathematical And Computer Modelling*, 20(8):17–29.

[69] Rodriguez, J. P., Brotons, L., Bustamante, J., , and Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions*, 13(3):243Ũ251.

[Shekhar and Chawla] Shekhar, S. and Chawla, S. *Spatial Databases: a Tour.* Prentice Hall,, Upper Saddle River,. New Jersey, USA.

[71] Shekhar, S. and Chawla, S. (2003). *Spatial Databases: A Tour.*

[72] Sindaco, R., Doria, G.and Razzetti, E., and Bernini, F. E. (2006). *Atlante degli Anfibi e dei Rettili d'Italia / Atlas of Italian Amphibians and Reptiles.* Societas Herpetologica Italica sezione Lombardia (2000).

[73] Stankowski, P. A. and Parker, W. (2010). Species distribution modelling: Does one size fit all? a phytogeographic analysis of salix in ontario. *Ecological Modelling*, 221(13-14):1655 – 1664.

[74] Stockwell, D. and Peters, D. (1999). The garp modelling system: problems and solutions to automated spatial prediction. *International Journal Of Geographical Information Science*, 13(2):143Ű158.

[75] Stockwell, D. R. B. and Peterson, A. T. (2002). E?ects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148(1):1Ű13.

[76] Team, G. D. Geographic resources analysis support system (grass) version 6.4.0. http://grass.itc.it/.

[77] Team, Q. D. Quantum gis (qgis). version 1.5.0. http://www.qgis.org/.

[78] Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., and Kadmon, R. (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*, 13(4):397–405.

[79] Urbano, F., Cagnacci, F., Calenge, C., Dettki, H., Cameron, A., and M., N. (2010). Wildlife tracking data management: a new vision. *Phil. Trans. R Soc B*, 365:2177–2185.

[80] Vaughan, I. (2005). The continuing challenges of testing species distribution models. *The Journal of applied ecology*, 42(4):720–730.

[81] Vesey-Fitzgerald, D. (1974). Utilisation of the grazing resources by buffaloes in the arusha national park, tanzania. *E. Afr. Wildl. J*, 12:107–134.

[82] Vigorita, V. & Cucè, L. a. c. d. (2008). *La fauna selavatica in Lombardia - rapporto 2008 su distribuzione, abbondanza e stato di conservazione di uccelli e mammiferi.* Regione Lombardia - Direzione Generale Agricoltura.

[83] Walker, P. A. and Cocks, K. D. (1991). Habitat: A procedure for modelling a disjoint environmental envelope for a plant or animal species. *Global Ecology and Biogeography Letters*, 1:108Ű118.

[84] Wong, I., Bloo, R., McNicol, D., Fong, P., Russel, R., and Chen, X. (2007). Species at risk: data and knoledge management within the wildspace decison support system. *Environmental Modelling and Software*, 22:423–430.

[85] Wong, I., McNicol, D., Fong, P., Fillman, D., Neysmith, J., and Russel, R. (2003). The wildspace decison support system. *Environmental Modelling and Software*, 18:521–530.

[86] Zaniewski, A. E., Lehmann, A., , and Overton, J. M. C. (2002). Predicting species spatial distributions using presence-only data: a case study of native new zealand ferns. *Ecological Modelling*, 157(2-3):261Ű280.

[87] Zeilhofer, P., Neto, P. S. A., Maja, W. Y., and Vecchiato, D. A. (2009). A web-based, component-oriented application for spatial modelling of habitat suitability of mosquito vectors. *International Journal Of Digital Earth*, 2(4):327–342.

A. Results of LRA

A.0.4. Mount Meru region - Tanzania

Table A.1.: Environmental variables for spatial models in Mount Meru region

Variable name	**Description**
ev-cl08	mean temperature of wettest quarter
ev-cl09	mean temperature of driest quarter
ev-cl16	precipitation of wettest quarter
ev-cl17	precipitation of driest quarter
ev-domi	dominance's diversity index
ev-shan	Shannon's diversity index
ev-edde	edge density index
ev-patc	patch density index
ev-lake	distance from lakes
ev-rvrs	distance from rivers
ev-road	distance from roads
ev-swpa	distance from swamps
ev-vllg	distance from villages
ev-tasp	digital elevation model - aspect
ev-tslp	digital elevation model - slope
ev-h101	continuous herbaceous vegetation
ev-h105	continuous herbaceous vegetation with scrubs
ev-h108	continuous herbaceous vegetation with isolated rainfed crop
ev-h113	continuous herbaceous vegetation with sparse trees and scrubs
ev-h115	continuous herbaceous vegetation with sparse scrubs
ev-h117	continuous herbaceous vegetation with isolated rainfed crop
ev-h122	open shrubs with sparse trees
ev-h126	open shrubs with sparse trees and with isolated fields of rainfed crop
ev-h127	open shrubs
ev-h131	open shrubs with isolated field of rainfed crop
ev-h132	open shrubs with combination of rainfed tree and herbaceous crops
ev-h134	continuous shrubs with herbaceous layers with rainfed tree and herbaceous crops

Table A.1.: (Continued) Environmental variables for spatial models in Mount Meru region

Variable name	**Description**
ev-h141	continuous shrubs with herbaceous layers/herbaceous vegetation with sparse trees and shrubs
ev-h143	open shrubs with sparse trees/continuous herbaceous vegetation with sparse trees and shrubs
ev-h145	open shrubs with sparse trees/isolated field of rainfed crop
ev-h146	continuous trees forest with scrubs
ev-h154	continuous broadleaved deciduous forest with herbaceous layer and sparse scrubs/urban area
ev-h162	continuous tree forest with shrubs
ev-h165	continuous broadleaved deciduous forest with shrubs
ev-h166	continuous broadleaved deciduous forest with shrubs/grassland
ev-h170	continuous woody vegetation with thorny plants
ev-h174	grassland with sparse shrubs on temporarily swampy area
ev-h175	grassland on permanently flooded area
ev-h177	grassland/woody vegetation with herbaceous layer on temporarily swampy area
ev-h180	shrubs with herbaceous layer on temporarily swampy area
ev-h182	shrubs with herbaceous layer on temporarily swampy/grassland on temporarily swampy area
ev-h190	bare rock
ev-h193	natural lakes
ev-h196	continuous rainfed herbaceous crops/continuous herbaceous vegetation with sparse trees and shrubs
ev-h202	continuous rainfed herbaceous crops/continuous herbaceous vegetation with sparse trees and shrubs
ev-h204	continuous rainfed herbaceous crops
ev-h205	continuous rainfed herbaceous crops/continuous herbaceous vegetation
ev-h207	continuous rainfed herbaceous crops/continuous herbaceous vegetation with sparse trees and shrubs
ev-h211	continuous rainfed herbaceous crops/combination of rainfed tree and shrubs crops
ev-h214	continuous forest plantations of rainfed Pine
ev-h216	continuous combination of rainfed tree and herbaceous crops/continuous herbaceous vegetation
ev-h218	continuous rainfed tree and herbaceous crops/continuous broadleaved deciduous forest with shrubs

Canis aureus

Table A.2.: selected environmental variables and computed coefficients from stepwise analysis for *Canis aureus* (golden jackal)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$1.918e+05$	$1.830e+05$	1.048	0.295
ev-cl09	$-4.316e+02$	$4.211e+02$	-1.025	0.305
ev-cl17	$-5.140e+02$	$4.992e+02$	-1.030	0.303
ev-edde	$-1.083e+01$	$1.012e+01$	-1.070	0.284
ev-patc	$1.623e+04$	$1.555e+04$	1.043	0.297
ev-lake	$-2.273e+00$	$2.187e+00$	-1.040	0.299
ev-rvrs	$1.747e-01$	$4.423e+00$	0.040	0.968
ev-swpa	$-1.161e-01$	$3.385e+00$	-0.034	0.973
ev-vllg	$-1.028e+00$	$1.375e+00$	-0.748	0.455
ev-h104	$1.226e+02$	$1.108e+02$	1.107	0.268
ev-h106	$2.194e+02$	$2.199e+02$	0.998	0.318
ev-h108	$8.709e+00$	$2.859e+02$	0.030	0.976
ev-h115	$1.823e+02$	$1.757e+02$	1.038	0.299
ev-h123	$3.243e+02$	$4.700e+02$	0.690	0.490
ev-h127	$7.197e+01$	$9.122e+01$	0.789	0.430
ev-h132	$1.283e+02$	$1.195e+02$	1.074	0.283
ev-h141	$3.271e+01$	$4.148e+02$	0.079	0.937
ev-h143	$-2.125e+02$	$1.927e+02$	-1.103	0.270
ev-h146	$-2.377e+02$	$2.583e+02$	-0.920	0.357
ev-h158	$3.128e+02$	$2.850e+02$	1.098	0.272
ev-h164	$-8.911e+01$	$1.150e+03$	-0.078	0.938
ev-h168	$7.924e+01$	$9.368e+02$	0.085	0.933
ev-h169	$6.762e+02$	$8.594e+02$	0.787	0.431
ev-h174	$-7.343e+01$	$7.689e+01$	-0.955	0.340
ev-h180	$-1.185e+02$	$1.308e+02$	-0.906	0.365
ev-h182	$1.226e-01$	$5.087e-02$	2.410	0.015
ev-h183	$-1.357e+02$	$3.314e+02$	-0.410	0.682
ev-h184	$4.895e-01$	$9.281e-02$	0.274	0.346
ev-h175	$1.035e+02$	$1.145e+02$	0.904	0.366
ev-h177	$-2.705e+01$	$3.363e+02$	-0.080	0.936
ev-h193	$-5.233e+02$	$5.830e+02$	-0.897	0.369
ev-h202	$-1.125e-01$	$5.482e-02$	-0.053	0.683
ev-h205	$7.217e+02$	$9.247e+02$	0.780	0.435
ev-h207	$-3.238e-01$	$6.224e-02$	0.569	0.972
ev-h214	$1.182e+00$	$1.446e-01$	0.177	0.916
ev-h216	$3.202e-01$	$9.737e-02$	3.289	0.734
ev-h218	$3.935e+00$	$6.252e+02$	0.006	0.995

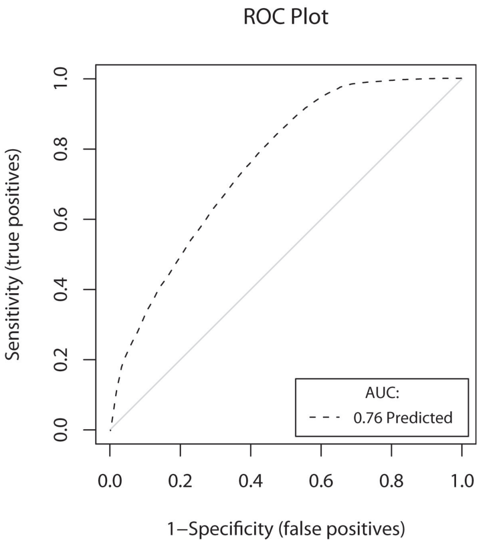


Figure A.1.: LRA for *Canis aureus* (golden jackal) ROC plot

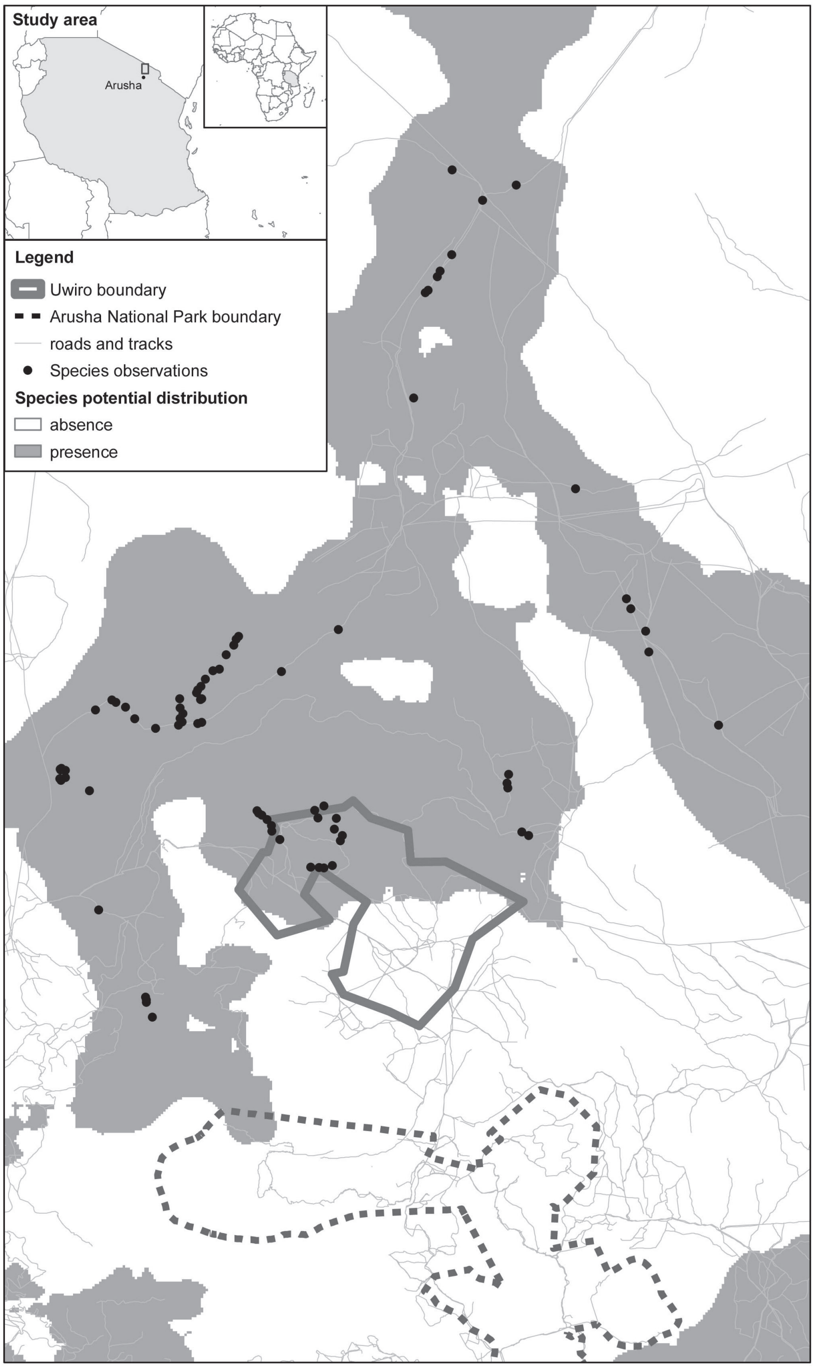


Figure A.2.: *Canis aureus* (golden jackal) potential distribution

Canis mesomelas (black-backed jackal)

Table A.3.: selected environmental variables and computed coefficients from stepwise analysis for *Canis mesomelas* (black-backed jackal)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	4.241e + 05	4.423e + 05	0.959	0.338
ev-cl04	2.953e + 01	3.498e + 01	0.844	0.399
ev-cl09	-5.969e + 02	5.891e + 02	-1.013	0.311
ev-cl15	5.145e + 02	5.922e + 02	0.869	0.385
ev-cl17	-3.542e + 02	5.328e + 02	-0.665	0.506
ev-domi	-1.215e + 03	1.300e + 03	-0.934	0.350
ev-lake	-7.017e − 01	8.634e − 01	-0.813	0.416
ev-road	-1.086e + 00	1.056e + 00	-1.028	0.304
ev-rvrs	1.151e + 00	1.173e + 00	0.981	0.326
ev-vllg	-8.992e − 01	9.364e − 01	-0.960	0.337
ev-tasp	9.908e − 01	1.671e + 00	0.593	0.553
ev-tslp	1.327e + 02	1.286e + 02	1.032	0.302
ev-h101	-6.319e + 01	6.077e + 01	-1.040	0.298
ev-h104	-5.803e + 01	7.189e + 01	-0.807	0.420
ev-h105	-5.536e + 01	5.466e + 01	-1.013	0.311
ev-h106	3.539e + 01	8.087e + 01	0.438	0.662
ev-h108	-1.175e + 02	1.222e + 02	-0.962	0.336
ev-h113	-6.838e + 01	6.985e + 01	-0.979	0.328
ev-h115	8.247e + 01	8.673e + 01	0.951	0.342
ev-h117	-5.042e + 01	5.213e + 01	-0.967	0.333
ev-h122	1.177e + 02	1.234e + 02	0.954	0.340
ev-h126	-9.098e + 01	1.242e + 02	-0.733	0.464
ev-h128	-1.465e + 02	1.455e + 02	-1.007	0.314
ev-h131	2.511e + 01	3.242e + 01	0.775	0.439
ev-h132	1.403e + 02	1.362e + 02	1.030	0.303
ev-h140	-2.704e + 02	2.853e + 02	-0.948	0.343
ev-h142	-5.595e + 01	7.458e + 01	-0.750	0.453
ev-h143	-9.911e + 01	1.304e + 02	-0.760	0.447
ev-h145	4.758e + 01	8.414e + 01	0.566	0.572
ev-h146	-8.860e + 01	1.037e + 02	-0.854	0.393
ev-h158	-9.093e + 01	1.235e + 02	-0.736	0.462
ev-h162	1.337e + 02	1.602e + 02	0.835	0.404
ev-h164	-1.243e + 02	2.551e + 02	-0.487	0.626
ev-h165	7.882e + 01	1.054e + 02	0.748	0.455
ev-h166	-1.152e + 02	1.402e + 02	-0.822	0.411
ev-h167	-3.249e + 02	3.794e + 02	-0.856	0.392
ev-h169	-6.750e + 01	1.161e + 02	-0.581	0.561
ev-h170	2.080e + 02	2.270e + 02	0.916	0.360
ev-h176	1.245e + 02	1.375e + 02	0.905	0.365
ev-h180	-1.125e + 02	1.072e + 02	-1.050	0.294
ev-h182	8.474e + 01	7.836e + 01	1.081	0.280
ev-h183	-4.220e + 02	4.987e + 02	-0.846	0.397

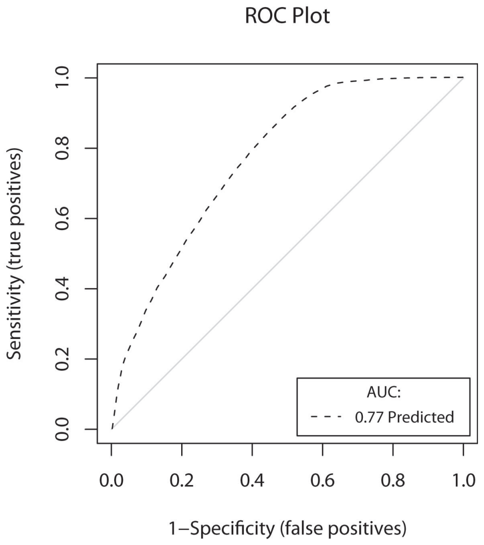


Figure A.3.: LRA for *Canis mesomelas* (black-backed jackal) ROC plot

Table A.3.: (continued) selected environmental variables and computed coefficients from stepwise analysis for *Canis mesomelas* (black-backed jackal)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
ev-h184	$3.469e+02$	$3.421e+02$	1.014	0.310
ev-h190	$5.746e+02$	$5.569e+02$	1.032	0.302
ev-h193	$-3.085e+02$	$3.745e+02$	-0.824	0.410
ev-h202	$-1.356e+02$	$1.343e+02$	-1.010	0.312
ev-h207	$-1.012e+02$	$9.298e+01$	-1.089	0.276
ev-h214	$-6.325e+02$	$6.152e+02$	-1.028	0.304
ev-h218	$-4.152e+02$	$3.967e+02$	-1.047	0.295

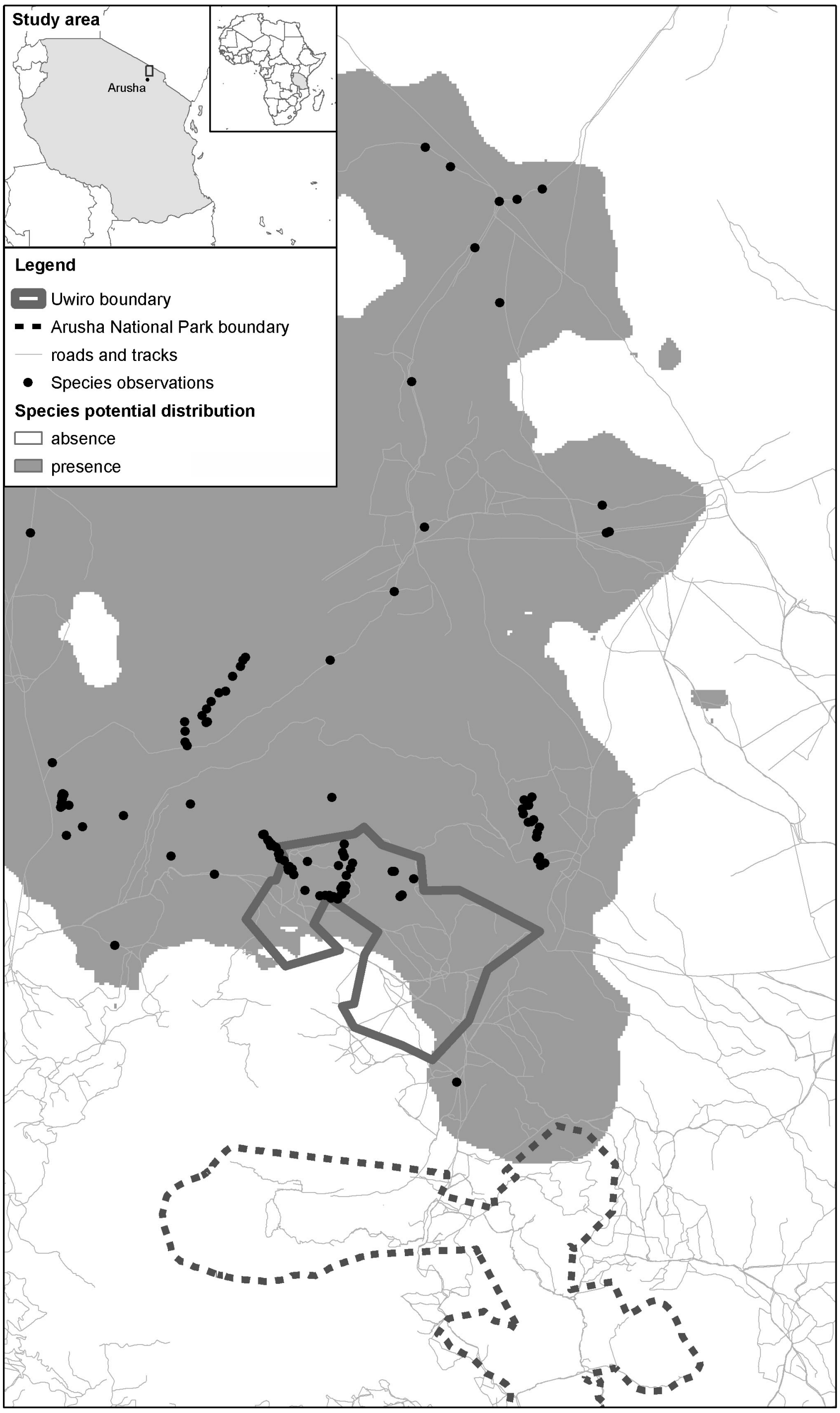


Figure A.4.: *Canis mesomelas* (black-backed jackal) potential distribution

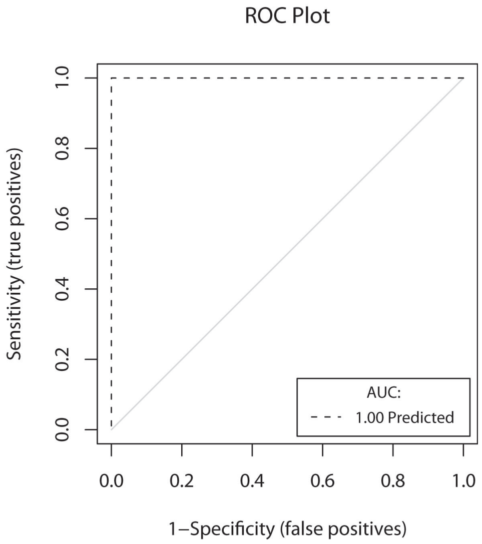


Figure A.5.: LRA for *Lycaon pictus* (wildog) ROC plot

Lycaon pictus (wildog)

Table A.4.: selected environmental variables and computed coefficients from stepwise analysis for *Lycaon pictus* (wildog)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	30550.353	116939.876	0.261	0.794
ev-cl09	-20.056	79.919	-0.251	0.802
ev-h106	-17.329	81.211	-0.213	0.831
ev-h108	-8.054	36.206	-0.222	0.824
ev-h126	9.199	36.235	0.254	0.800
ev-h146	-32.359	163.664	-0.198	0.843
ev-h156	-43.354	175.628	-0.247	0.805
ev-h174	-29.120	110.035	-0.265	0.791
ev-h177	-35.200	171.702	-0.205	0.838
ev-h190	28.782	109.534	0.263	0.793
ev-h193	24.428	196.258	0.124	0.901
ev-h211	-61.725	241.722	-0.255	0.798

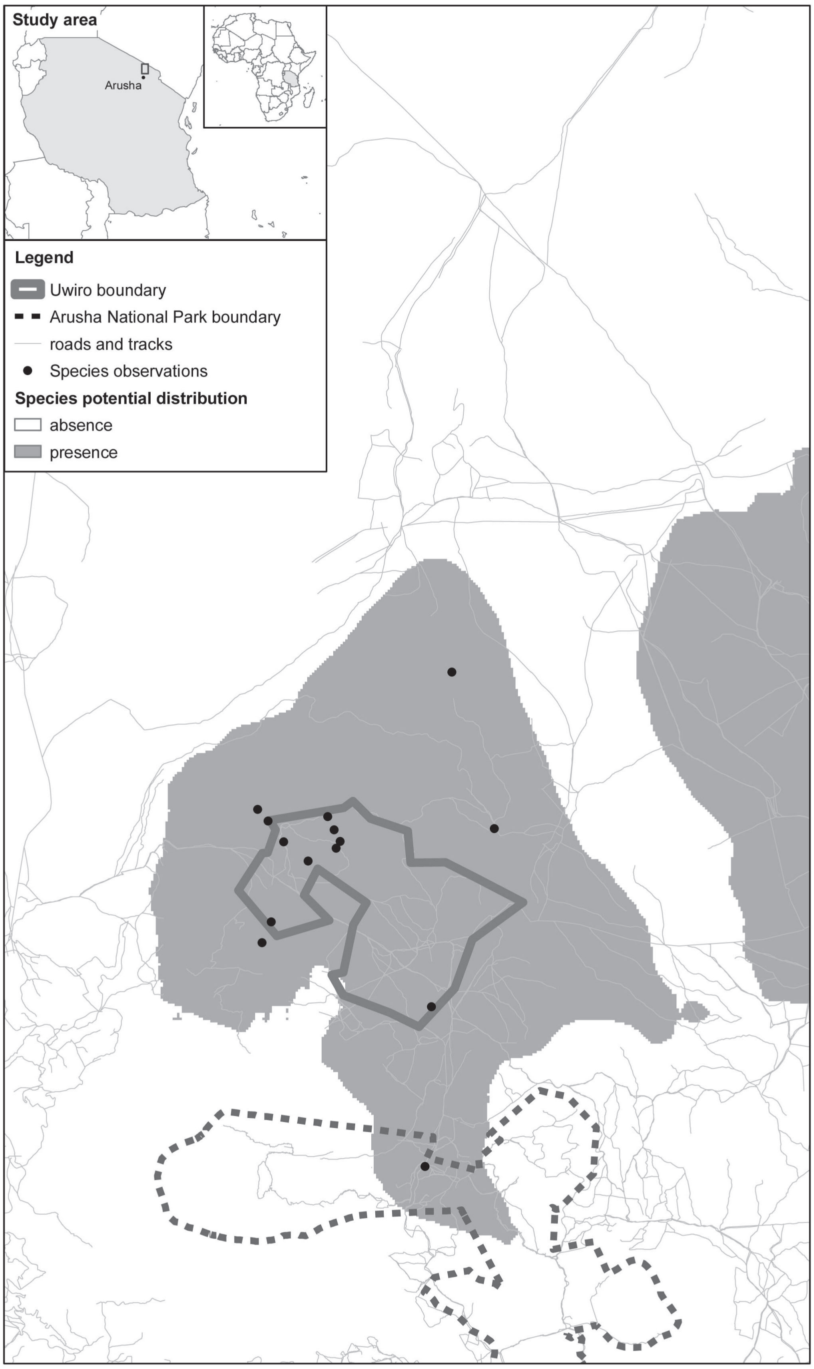


Figure A.6.: *Lycaon pictus* (wildog) potential distribution

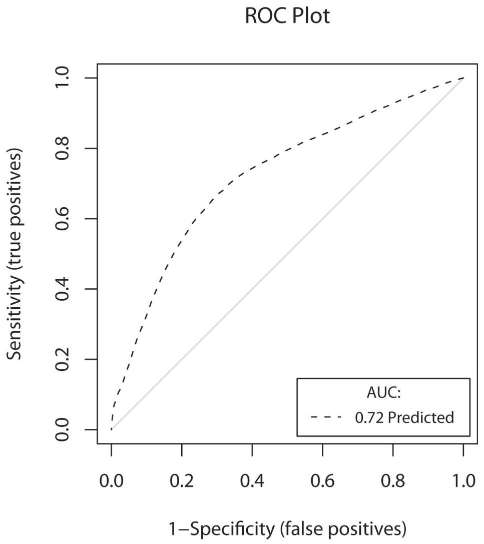


Figure A.7.: LRA for *Otocyon megalotis* (bat-eared fox) ROC plot

Otocyon megalotis (bat-eared fox)

Table A.5.: selected environmental variables and computed coefficients from stepwise analysis for *Otocyon megalotis* (bat-eared fox)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$7.066e+03$	$5.693e+04$	0.124	0.901
$\mathrm{ev}_e dde$	$-6.286e-01$	$5.436e+00$	-0.116	0.908
$\mathrm{ev}_r oad$	$1.020e-01$	$9.169e-01$	0.111	0.911
$\mathrm{ev}_r vrs$	$9.041e-02$	$8.285e-01$	0.109	0.913
$\mathrm{ev}_h 117$	$2.022e+01$	$1.624e+02$	0.125	0.901
$\mathrm{ev}_h 122$	$9.952e+00$	$1.127e+02$	0.088	0.930
$\mathrm{ev}_h 168$	$2.645e+01$	$2.313e+02$	0.114	0.909
$\mathrm{ev}_h 169$	$-5.231e+01$	$4.338e+02$	-0.121	0.904
$\mathrm{ev}_h 180$	$-1.133e+01$	$9.058e+01$	-0.125	0.900
$\mathrm{ev}_h 214$	$-2.232e+01$	$1.940e+02$	-0.115	0.908

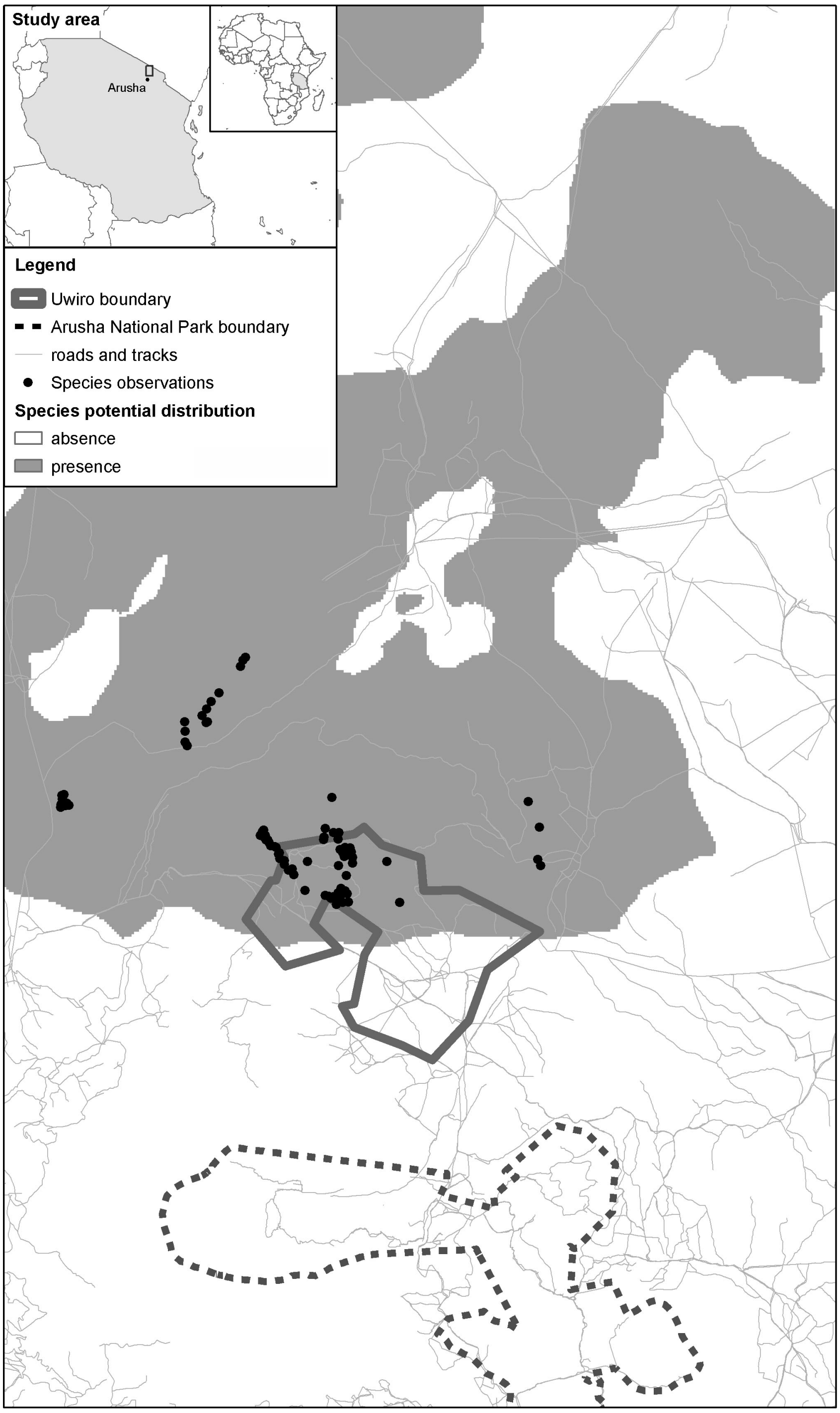


Figure A.8.: *Otocyon megalotis* (bat-eared fox) potential distribution

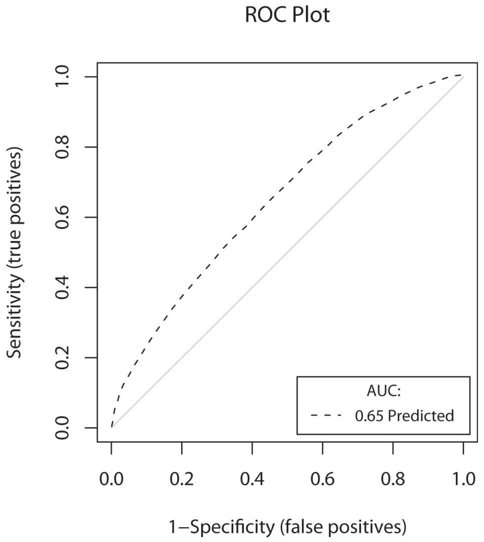


Figure A.9.: LRA for *Crocuta crocuta* (spotted hyena) ROC plot

Crocuta crocuta (spotted hyena)

Table A.6.: selected environmental variables and computed coefficients from stepwise analysis for *Crocuta crocuta* (spotted hyena)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	$-1.576e+04$	$4.380e+05$	-0.036	0.971
ev-rvrs	$1.747e-01$	$4.423e+00$	0.040	0.968
ev-swpa	$-1.161e-01$	$3.385e+00$	-0.034	0.973
ev-tasp	$4.383e-01$	$1.303e+01$	0.034	0.973
ev-h106	$-1.648e+01$	$3.623e+02$	-0.045	0.964
ev-h108	$8.709e+00$	$2.859e+02$	0.030	0.976
ev-h132	$1.107e+01$	$6.256e+02$	0.018	0.986
ev-h146	$2.529e+01$	$6.299e+02$	0.040	0.968
ev-h156	$-6.320e+01$	$1.087e+03$	-0.058	0.954
ev-h162	$4.782e+01$	$9.185e+02$	0.052	0.958
ev-h167	$6.623e+01$	$1.286e+03$	0.051	0.959
ev-h170	$1.757e+01$	$3.033e+02$	0.058	0.954
ev-h175	$-2.004e+01$	$3.840e+02$	-0.052	0.958
ev-h183	$-5.495e+01$	$9.622e+02$	-0.057	0.954
ev-h207	$2.298e+01$	$5.762e+02$	0.040	0.968
ev-h218	$3.935e+00$	$6.252e+02$	0.006	0.995

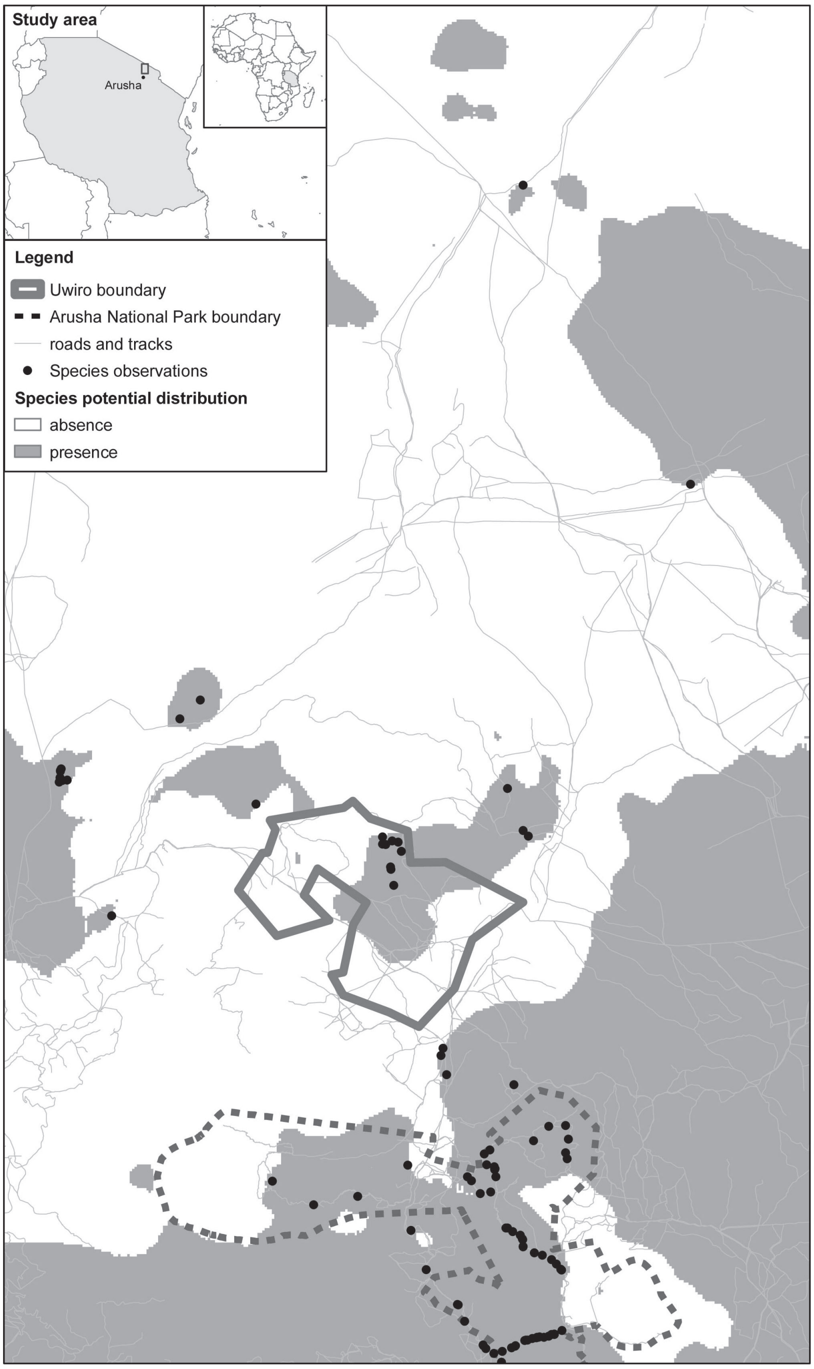


Figure A.10.: *Crocuta crocuta* (spotted hyena) potential distribution

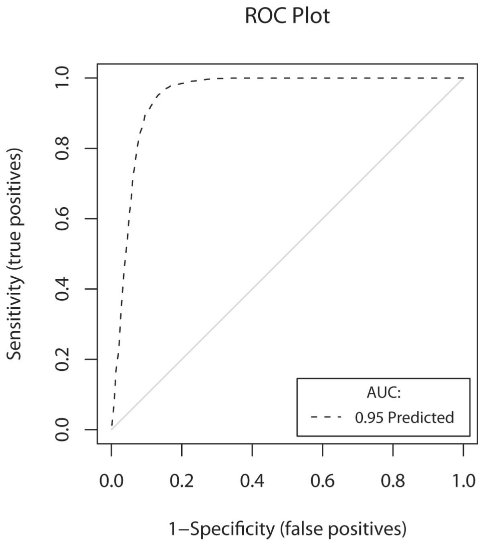


Figure A.11.: LRA for *Civettictis civetta* (african civet) ROC plot

Civettictis civetta (african civet)

Table A.7.: selected environmental variables and computed coefficients from stepwise analysis for *Civettictis civetta* (african civet)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$-4.871e+00$	$2.679e+00$	-1.819	0.068963
ev-cl09	$7.979e-03$	$5.279e-03$	1.511	0.130718
ev-domi	$4.191e-01$	$1.695e-01$	2.473	0.013394
ev-swpa	$-6.312e-05$	$2.162e-05$	-2.920	0.003501
ev-h106	$6.102e-03$	$2.785e-03$	2.191	0.028482
ev-h122	$-6.596e-03$	$2.604e-03$	-2.533	0.011314
ev-h132	$-8.927e-03$	$1.845e-03$	-4.838	1.31e-06
ev-h134	$-9.617e-03$	$3.041e-03$	-3.162	0.001567
ev-h156	$-1.183e-02$	$2.969e-03$	-3.985	6.74e-05
ev-h178	$-1.654e-02$	$4.160e-03$	-3.976	7.01e-05
ev-h216	$-1.164e-02$	$4.161e-03$	-2.797	0.005151

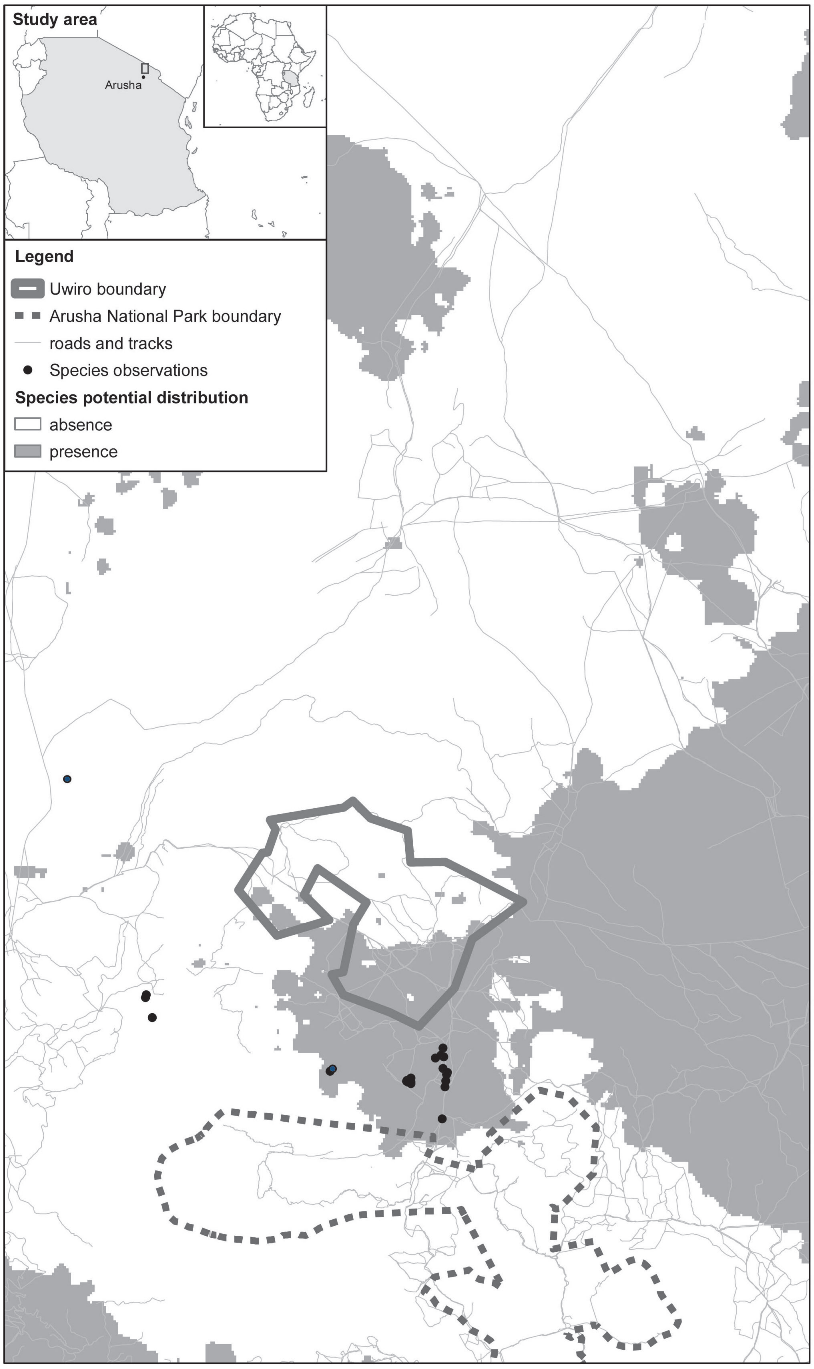


Figure A.12.: *Civettictis civetta* (african civet) potential distribution

Lepus saxatilis (scrub hare)

Table A.8.: selected environmental variables and computed coefficients from stepwise analysis for *Lepus saxatilis* (scrub hare)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	1.411e + 05	1.176e + 05	1.199	0.230
ev-cl04	2.790e + 01	72.195e + 01	1.271	0.204
ev-cl09	2.193e + 02	71.704e + 02	1.287	0.198
ev-cl15	-4.132e + 02	73.200e + 02	-1.291	0.197
ev-cl17	2.410e + 02	71.888e + 02	1.277	0.202
ev-domi	-2.892e + 03	72.260e + 03	-1.280	0.201
ev-shan	-6.016e + 03	74.735e + 03	-1.271	0.204
ev-edde	-5.903e + 00	74.590e + 00	-1.286	0.198
ev-patc	1.926e + 04	71.528e + 04	1.261	0.207
ev-lake	-4.622e − 01	73.637e − 01	-1.271	0.204
ev-road	-1.686e − 01	71.427e − 01	-1.182	0.237
ev-swpa	9.473e − 01	77.574e − 01	1.251	0.211
ev-vllg	-3.410e − 01	72.873e − 01	-1.187	0.235
ev-tasp	1.062e + 00	71.184e + 00	0.896	0.370
ev-tslp	2.881e + 01	72.375e + 01	1.213	0.225
ev-h101	-8.098e + 01	76.288e + 01	-1.288	0.198
ev-h104	8.135e + 01	76.380e + 01	1.275	0.202
ev-h105	1.020e + 02	78.001e + 01	1.275	0.202
ev-h106	6.209e + 02	74.802e + 02	1.293	0.196
ev-h107	-3.474e + 02	72.784e + 02	-1.248	0.212
ev-h108	-2.222e + 02	71.723e + 02	-1.290	0.197
ev-h110	-2.247e + 02	71.811e + 02	-1.241	0.215
ev-h113	-1.018e + 02	78.060e + 01	-1.263	0.207
ev-h115	8.017e + 00	71.698e + 01	0.472	0.637
ev-h117	-2.741e + 01	72.197e + 01	-1.248	0.212
ev-h122	8.343e + 01	76.505e + 01	1.283	0.200
ev-h123	1.249e + 02	79.902e + 01	1.261	0.207
ev-h126	-5.784e + 01	74.500e + 01	-1.285	0.199
ev-h127	1.271e + 02	71.038e + 02	1.225	0.221
ev-h128	6.225e + 01	74.940e + 01	1.260	0.208
ev-h131	1.552e + 02	71.212e + 02	1.281	0.200
ev-h132	3.737e + 01	73.146e + 01	1.188	0.235
ev-h134	-1.970e + 02	71.529e + 02	-1.288	0.198
ev-h140	7.951e + 01	76.647e + 01	1.196	0.232
ev-h141	6.283e + 01	76.727e + 01	0.934	0.350
ev-h142	-2.308e + 02	71.804e + 02	-1.280	0.201
ev-h143	4.176e + 02	73.261e + 02	1.280	0.200
ev-h145	4.062e + 02	73.172e + 02	1.280	0.200
ev-h146	3.013e + 01	72.477e + 01	1.217	0.224
ev-h165	1.348e + 02	71.130e + 02	1.193	0.233
ev-h167	6.964e + 01	75.993e + 01	1.162	0.245
ev-h170	8.553e + 01	76.819e + 01	1.254	0.210

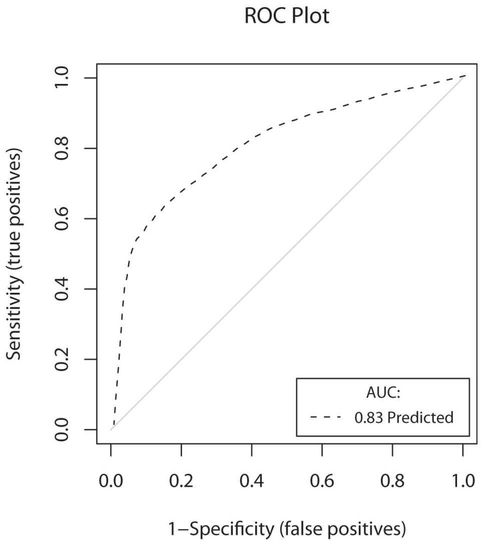


Figure A.13.: LRA for *Lepus saxatilis* (scrub hare) ROC plot

Table A.8.: selected environmental variables and computed coefficients from stepwise analysis for *Lepus saxatilis* (scrub hare)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
ev-h174	$7.174e+01$	$75.600e+01$	1.281	0.200
ev-h176	$-8.908e+01$	$76.997e+01$	-1.273	0.203
ev-h178	$-2.006e+02$	$71.568e+02$	-1.280	0.201
ev-h180	$-1.461e+02$	$71.133e+02$	-1.290	0.197
ev-h183	$2.662e+02$	$72.064e+02$	1.290	0.197
ev-h184	$-4.088e+02$	$73.222e+02$	-1.269	0.204
ev-h190	$-2.558e+02$	$71.968e+02$	-1.299	0.194
ev-h193	$-1.199e+02$	$79.483e+01$	-1.264	0.206
ev-h207	$-5.080e+02$	$73.932e+02$	-1.292	0.196
ev-h211	$-2.016e+02$	$71.604e+02$	-1.257	0.209
ev-h214	$-1.871e+02$	$71.482e+02$	-1.262	0.207
ev-h218	$-6.025e+01$	$76.502e+01$	-0.927	0.354

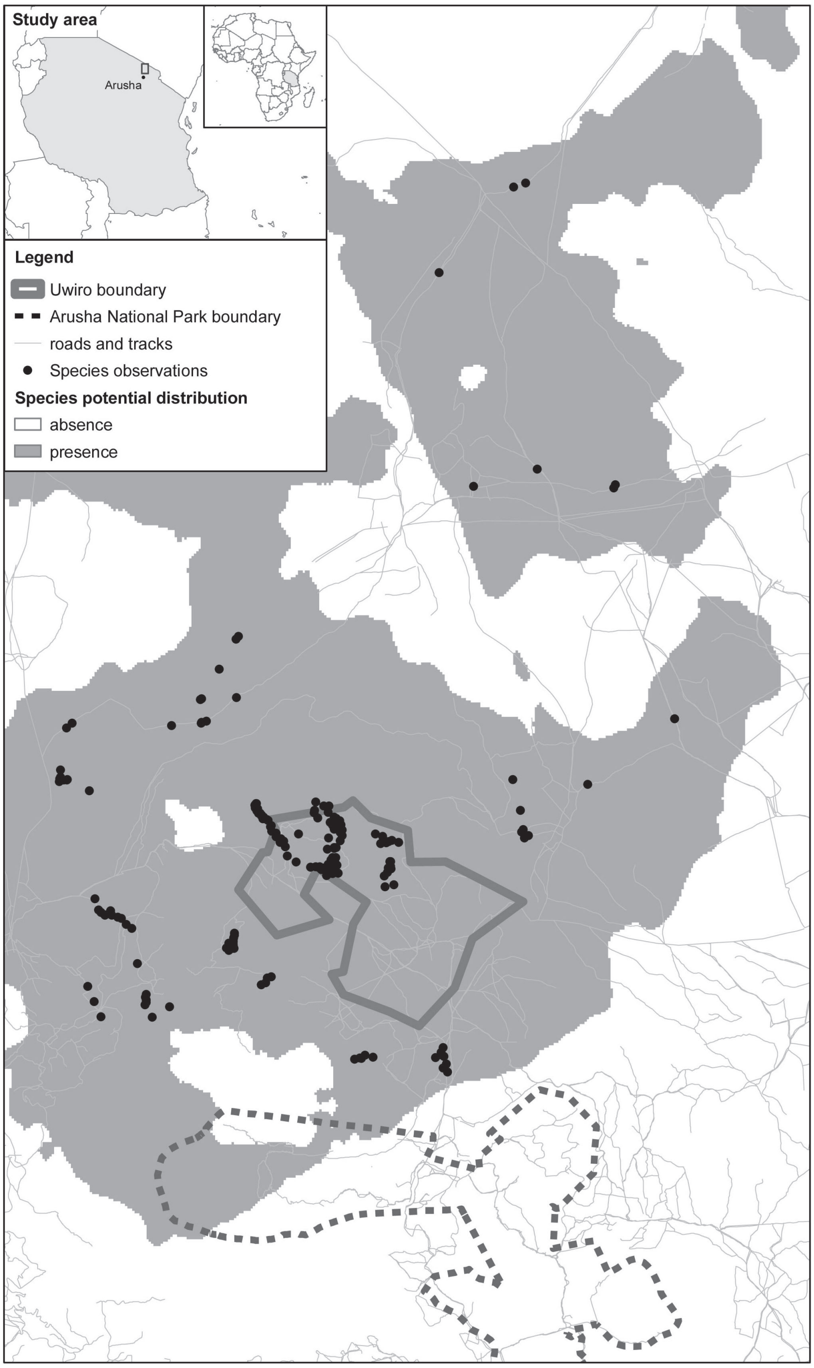


Figure A.14.: *Lepus saxatilis* (scrub hare) potential distribution

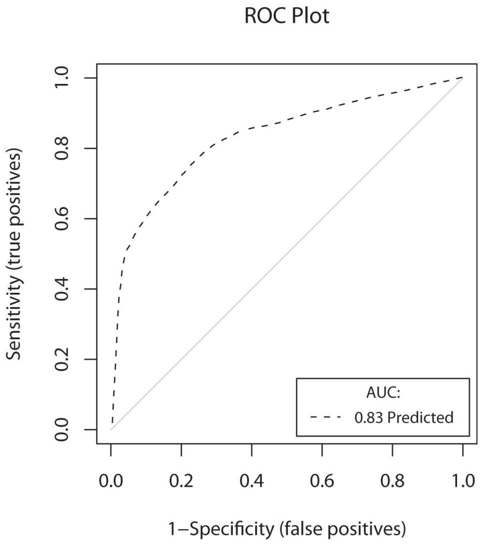


Figure A.15.: LRA for *Pronolagus rupestris* (Smith's red rock hare) ROC plot

***Pronolagus rupestris* (Smith's red rock hare)**

Table A.9.: selected environmental variables and computed coefficients from stepwise analysis for *Pronolagus rupestris* (Smith's red rock hare)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$3.721e+06$	$3.254e+07$	0.114	0.909
ev-cl15	$-4.766e+03$	$4.160e+04$	0.252	0.801
ev-cl17	$-8.319e+03$	$7.294e+04$	0.271	0.786
ev-patc	$3.087e+05$	$2.712e+06$	0.285	0.775
ev-lake	$-6.957e+00$	$6.088e+01$	-0.024	0.981
ev-road	$-3.136e+01$	$2.750e+02$	-0.086	0.932
ev-tasp	$1.875e+01$	$1.641e+02$	0.019	0.985
ev-h107	$-2.627e+03$	$2.292e+04$	0.017	0.986
ev-h122	$3.975e+03$	$3.479e+04$	-0.055	0.956
ev-h131	$-8.210e+02$	$7.190e+03$	0.042	0.967
ev-h134	$2.540e+03$	$2.227e+04$	-0.114	0.909
ev-h146	$-3.095e+03$	$2.707e+04$	0.311	0.756
ev-h164	$-1.770e+03$	$1.551e+04$	0.260	0.795
ev-h168	$1.318e+03$	$1.153e+04$	-0.126	0.899
ev-h178	$3.926e+03$	$3.438e+04$	-0.123	0.902
ev-h182	$-2.518e+03$	$2.203e+04$	0.127	0.899
ev-h211	$3.198e+03$	$2.801e+04$	-0.172	0.864
ev-h216	$3.068e+03$	$2.693e+04$	-0.215	0.830

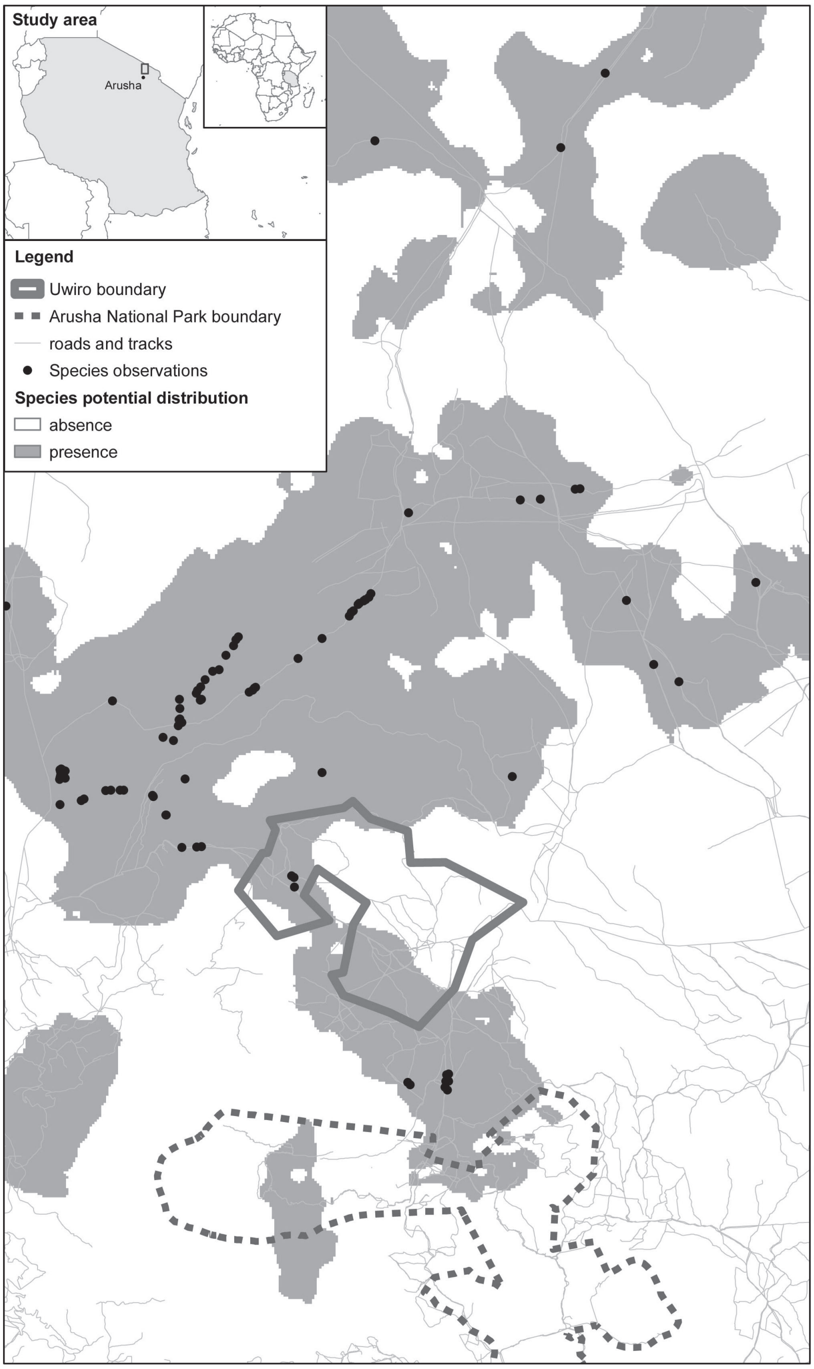


Figure A.16.: *Pronolagus rupestris* (Smith's red rock hare) potential distribution

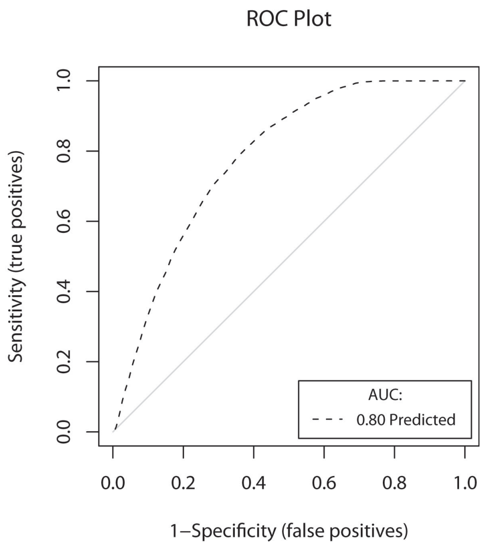


Figure A.17.: LRA for *Cercopithecus mitis* (blue monkey) ROC plot

Cercopithecus mitis (blue monkey)

Table A.10.: selected environmental variables and computed coefficients from stepwise analysis for *Cercopithecus mitis* (blue monkey)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$1.148e+04$	$1.868e+05$	0.061	0.951
ev-domi	$5.717e+02$	$9.139e+03$	0.063	0.950
ev-road	-3.529$e-01$	$5.797e+00$	-0.061	0.951
ev-h104	-1.238$e+01$	$2.030e+02$	-0.061	0.951
ev-h117	-2.491$e+01$	$3.348e+02$	-0.074	0.941
ev-h141	-2.392$e+01$	2.925e$+02$	-0.082	0.935
ev-h142	-3.733$e+01$	$4.879e+02$	-0.077	0.939
ev-h143	$3.042e+01$	$3.639e+02$	0.084	0.933
ev-h158	-1.368$e+01$	$1.842e+02$	-0.074	0.941
ev-h170	$1.879e+01$	$2.506e+02$	0.075	0.940
ev-h180	$1.032e+01$	$1.749e+02$	0.059	0.953
ev-h214	-2.348$e+01$	$3.905e+02$	-0.060	0.952

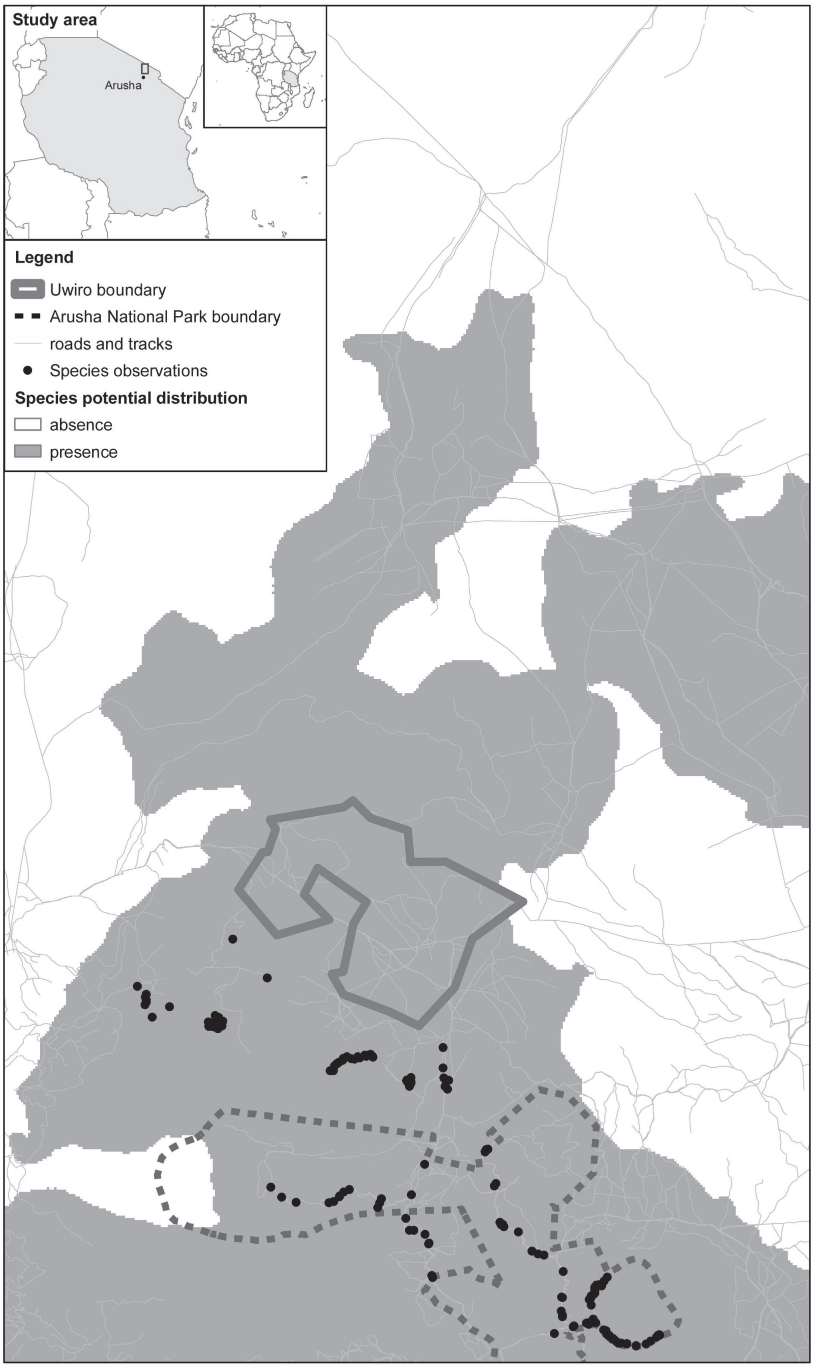


Figure A.18.: *Cercopithecus mitis* (blue monkey) potential distribution

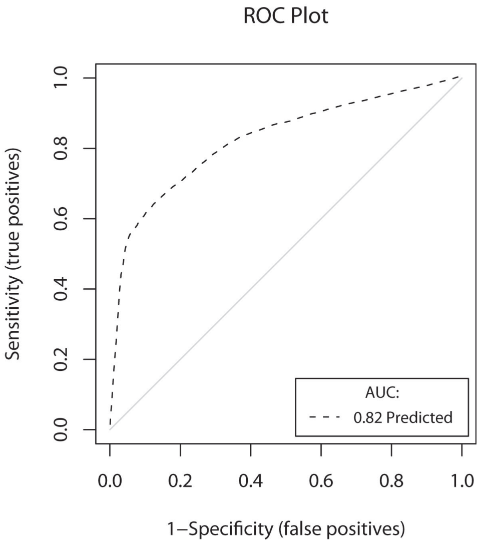


Figure A.19.: LRA for *Papio cynocephalus* (yellow baboon) ROC plot

Papio cynocephalus **(yellow baboon)**

Table A.11.: selected environmental variables and computed coefficients from stepwise analysis for *Papio cynocephalus* (yellow baboon)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	2.956e + 04	7.968e + 05	0.037	0.970
ev-cl15	1.052e + 02	2.042e + 03	0.052	0.959
ev-edde	2.323e + 00	1.044e + 02	0.022	0.982
ev-lake	1.778e − 01	9.730e + 00	0.018	0.985
ev-swpa	2.808e − 01	3.205e + 01	0.009	0.993
ev-h104	-1.865e + 01	8.276e + 02	-0.023	0.982
ev-h105	-6.777e + 01	2.867e + 03	-0.024	0.981
ev-h106	-5.380e + 01	6.278e + 02	-0.086	0.932
ev-h113	4.671e + 01	9.759e + 02	0.048	0.962
ev-h122	3.562e + 01	1.858e + 03	0.019	0.985
ev-h123	-6.224e + 01	1.121e + 03	-0.056	0.956
ev-h127	9.287e + 01	4.949e + 03	0.019	0.985
ev-h131	2.042e + 01	1.188e + 03	0.017	0.986
ev-h162	-1.089e + 02	5.206e + 03	-0.021	0.983
ev-h169	-9.793e + 01	3.312e + 03	-0.030	0.976
ev-h174	-3.896e + 01	2.562e + 03	-0.015	0.988
ev-h175	-3.504e + 01	6.363e + 02	-0.055	0.956
ev-h177	-1.724e + 01	1.747e + 03	-0.010	0.992
ev-h182	-8.474e + 01	3.190e + 03	-0.027	0.979
ev-h183	1.456e + 02	3.489e + 03	0.042	0.967

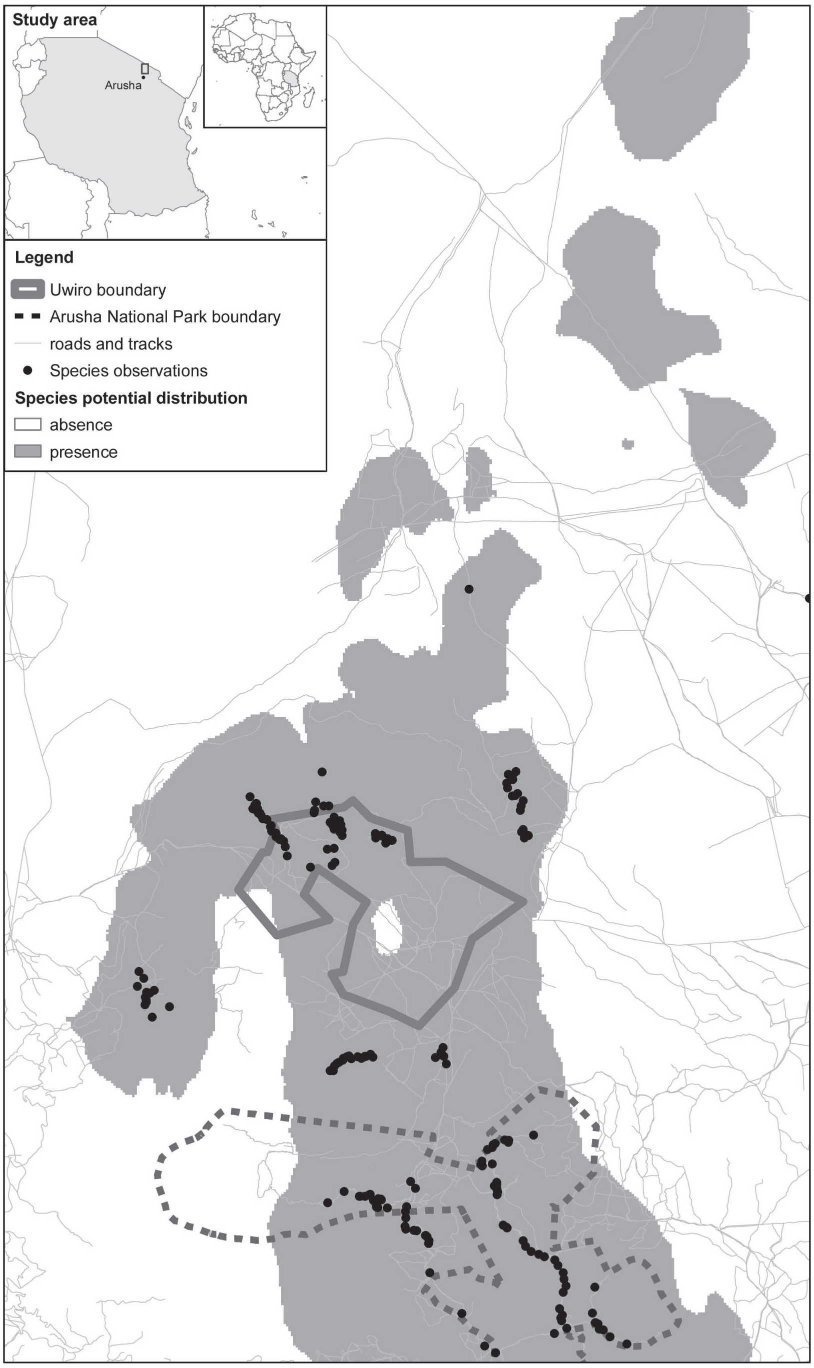


Figure A.20.: *Papio cynocephalus* (yellow baboon) potential distribution

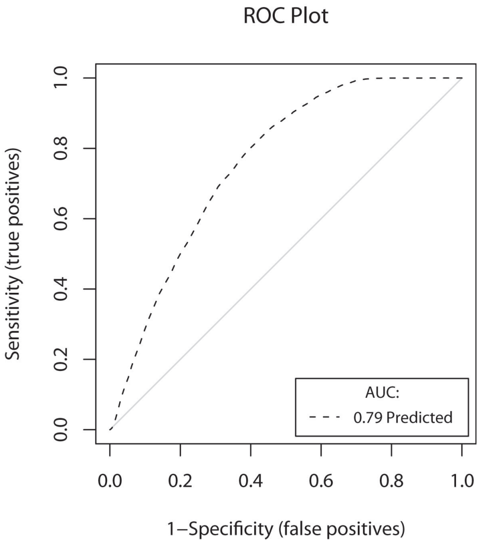


Figure A.21.: LRA for *Colobus guereza* (mantled guereza) ROC plot

Colobus guereza (mantled guereza)

Table A.12.: selected environmental variables and computed coefficients from stepwise analysis for *Colobus guereza* (mantled guereza)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	27098.0741	77401.9803	0.350	0.726
ev-shan	-1499.0778	4157.6777	-0.361	0.718
ev-road	-0.4483	1.2221	-0.367	0.714
ev-vllg	0.1887	0.5158	0.366	0.714
ev-h123	-49.2945	139.2956	-0.354	0.723
ev-h134	-51.4382	140.1831	-0.367	0.714
ev-h142	-50.1845	138.5773	-0.362	0.717
ev-h145	-94.3995	255.6411	-0.369	0.712
ev-h146	-80.5670	219.6753	-0.367	0.714
ev-h178	-71.3159	193.3506	-0.369	0.712
ev-h183	122.5455	329.7377	0.372	0.710

Figure A.22.: *Colobus guereza* (mantled guereza) potential distribution

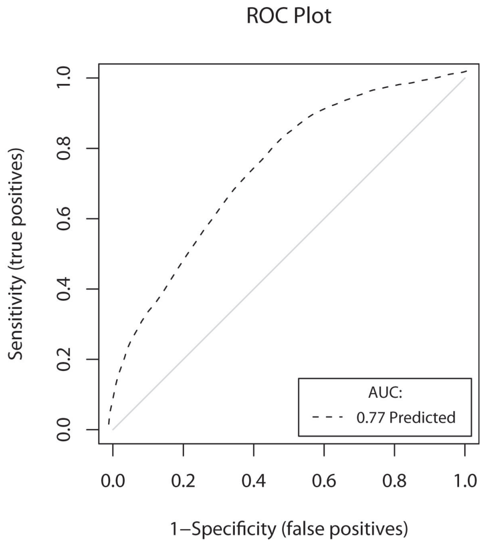


Figure A.23.: LRA for *Procavia capensis* (rock hyrax) ROC plot

Procavia capensis (rock hyrax)

Table A.13.: selected environmental variables and computed coefficients from stepwise analysis for *Procavia capensis* (rock hyrax)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	5212.4283	80858.4668	0.064	0.949
ev-domi	-183.0723	3404.9062	-0.054	0.957
ev-road	0.1137	2.0657	0.055	0.956
ev-h113	2.6780	80.8673	0.033	0.974
ev-h122	-5.7782	95.1944	-0.061	0.952
ev-h165	-10.9536	170.9874	-0.064	0.949
ev-h166	6.1611	141.8470	0.043	0.965
ev-h176	-7.4995	132.3412	-0.057	0.955
ev-h211	-7.7416	119.2485	-0.065	0.948

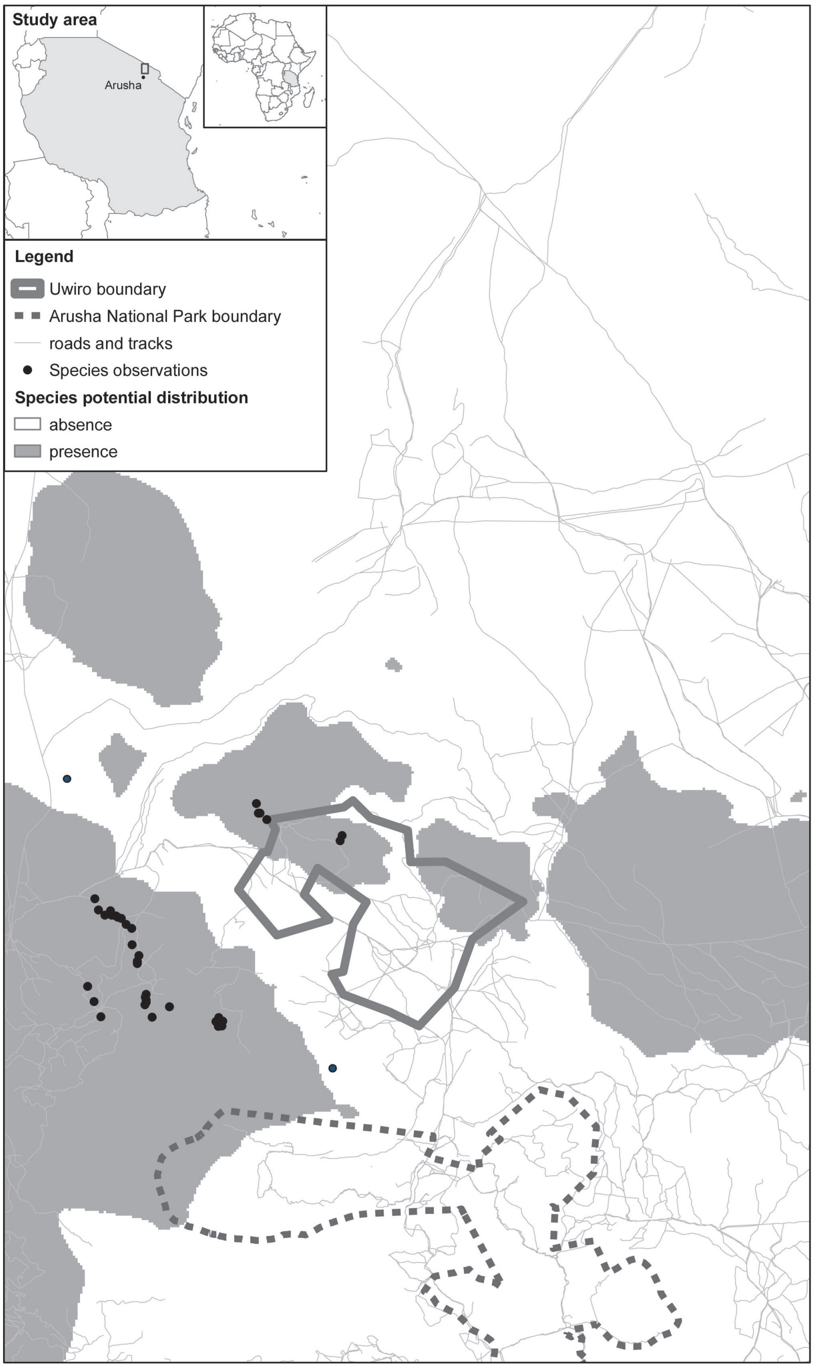


Figure A.24.: *Procavia capensis* (rock hyrax) potential distribution

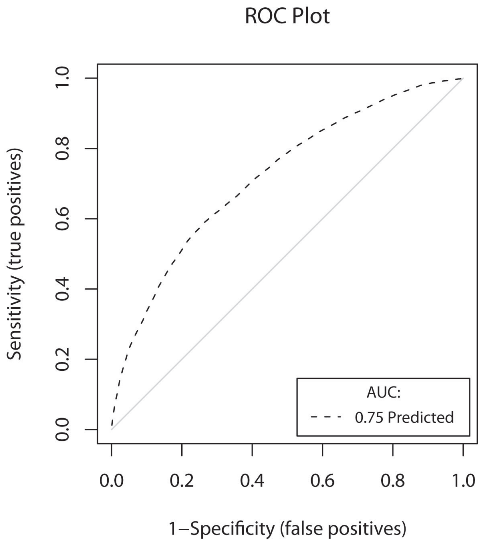


Figure A.25.: *Phacochoerus africanus* (warthdog) ROC plot

Phacochoerus africanus (warthdog)

Table A.14.: selected environmental variables and computed coefficients from stepwise analysis for *Phacochoerus africanus* (warthdog)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	-2342.23	309985.21	-0.008	0.994
ev-h107	67.13	1697.75	0.040	0.968
ev-h132	25.88	927.85	0.028	0.978
ev-h158	-30.00	1144.65	-0.026	0.979
ev-h164	-46.68	1424.77	-0.033	0.974
ev-h167	59.98	1790.15	0.034	0.973
ev-h170	-11.54	349.97	-0.033	0.974
ev-h175	-21.81	713.60	-0.031	0.976
ev-h178	32.82	1662.91	0.020	0.984
ev-h180	-19.55	539.74	-0.036	0.971
ev-h183	-72.71	2671.67	-0.027	0.978
ev-h190	29.62	1250.72	0.024	0.981
ev-h202	26.47	819.47	0.032	0.974
ev-h207	-28.43	1108.87	-0.026	0.980
ev-h218	-44.38	2016.10	-0.022	0.982

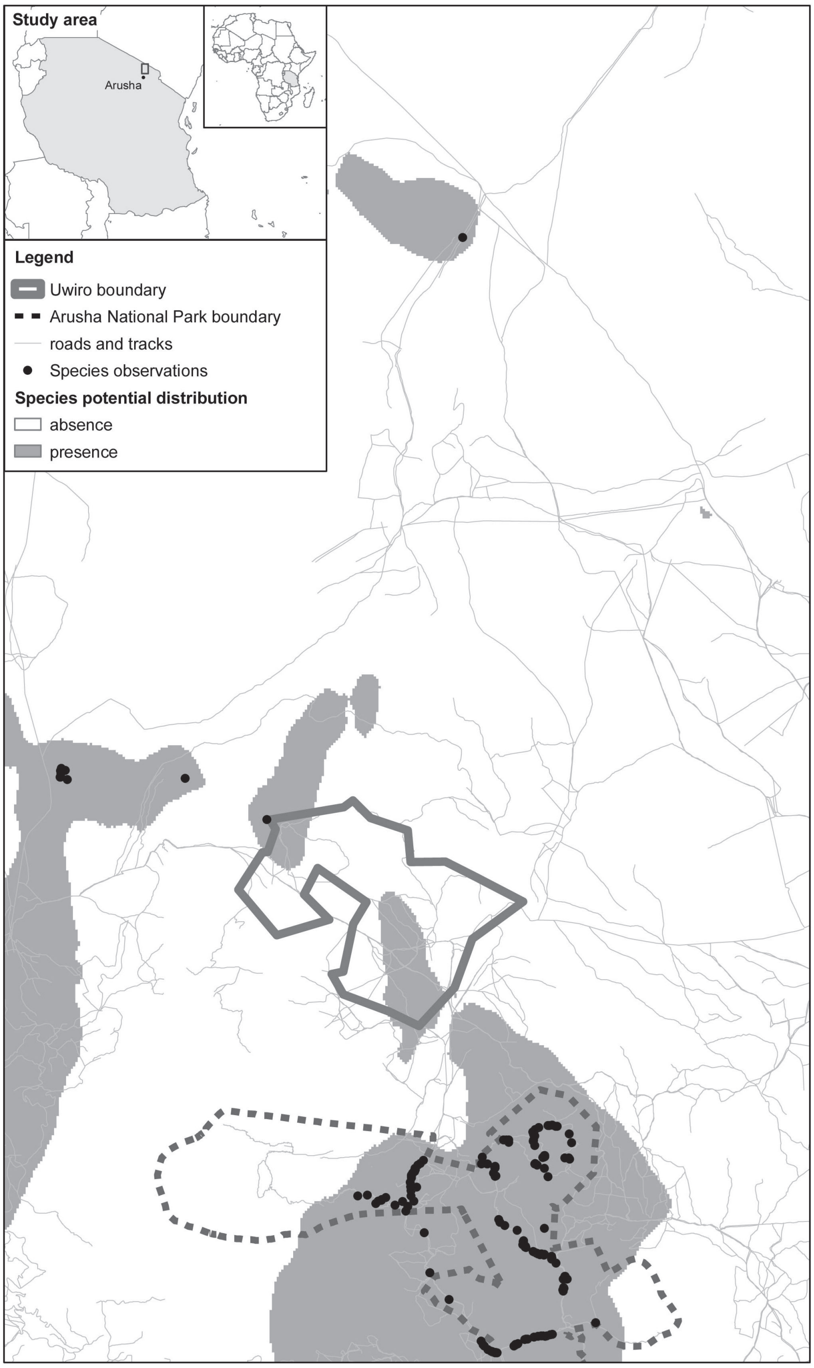


Figure A.26.: LRA for *Phacochoerus africanus* (warthdog) potential distribution

Giraffa camelopardalis (giraffe)

Table A.15.: selected environmental variables and computed coefficients from stepwise analysis for *Giraffa camelopardalis* (giraffe)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	4.126e+02	2.172e+02	1.899	0.057525
ev-cl09	1.086e+00	2.589e−01	4.195	2.73e-05
ev-cl15	1.218e+00	6.384e−01	1.908	0.056371
ev-domi	-1.134e+01	3.558e+00	-3.186	0.001442
ev-mpsi	-1.449e−06	4.811e−07	-3.012	0.002591
ev-lake	1.893e−03	7.344e−04	2.577	0.009970
ev-road	-7.156e−03	1.401e−03	-5.108	3.25e-07
ev-rvrs	2.898e−03	7.137e−04	4.061	4.90e-05
ev-swpa	2.002e−03	6.415e−04	3.121	0.001803
ev-vllg	1.873e−03	4.516e−04	4.149	3.35e-05
ev-tasp	7.354e−03	2.892e−03	2.543	0.010977
ev-tslp	-6.563e−01	2.549e−01	-2.575	0.010021
ev-h101	2.934e+01	1.247e+03	0.024	0.981224
ev-h104	-3.819e−01	1.100e−01	-3.473	0.000515
ev-h106	6.513e−01	1.344e−01	4.844	1.27e-06
ev-h107	-7.360e−01	2.461e−01	-2.990	0.002787
ev-h108	-4.666e−01	9.874e−02	-4.726	2.29e-06
ev-h113	1.370e−01	6.684e−02	2.050	0.040375
ev-h115	2.917e−01	9.662e−02	3.019	0.002535
ev-h117	-5.804e−01	1.302e−01	-4.456	8.35e-06
ev-h122	4.084e−01	1.304e−01	3.132	0.001738
ev-h126	-3.607e−01	9.221e−02	-3.912	9.16e-05
ev-h127	4.796e−01	1.763e−01	2.720	0.006532
ev-h128	1.789e−01	7.402e−02	2.417	0.015668
ev-h134	-2.194e−01	6.661e−02	-3.294	0.000987
ev-h140	-1.321e+00	2.208e−01	-5.983	2.18e-09
ev-h141	-1.704e+00	3.546e−01	-4.805	1.55e-06
ev-h143	3.892e−01	1.675e−01	2.324	0.020119
ev-h145	4.635e−01	1.375e−01	3.372	0.000746
ev-h146	-2.704e−01	1.074e−01	-2.517	0.011827
ev-h156	-8.627e−01	2.581e−01	-3.343	0.000829
ev-h158	3.099e−01	1.394e−01	2.223	0.026240
ev-h164	7.922e−01	2.536e−01	3.124	0.001783
ev-h165	5.000e−01	1.602e−01	3.121	0.001803
ev-h167	7.652e−01	2.651e−01	2.886	0.003898
ev-h168	3.451e−01	1.854e−01	1.861	0.062756
ev-h170	5.846e−01	1.225e−01	4.771	1.83e-06
ev-h174	-5.716e−01	1.150e−01	-4.972	6.61e-07
ev-h175	-2.907e+01	1.247e+03	-0.023	0.981393
ev-h177	-1.122e+00	2.312e−01	-4.854	1.21e-06
ev-h178	-7.300e−01	1.636e−01	-4.461	8.14e-06
ev-h182	-6.454e−01	1.587e−01	-4.066	4.79e-05

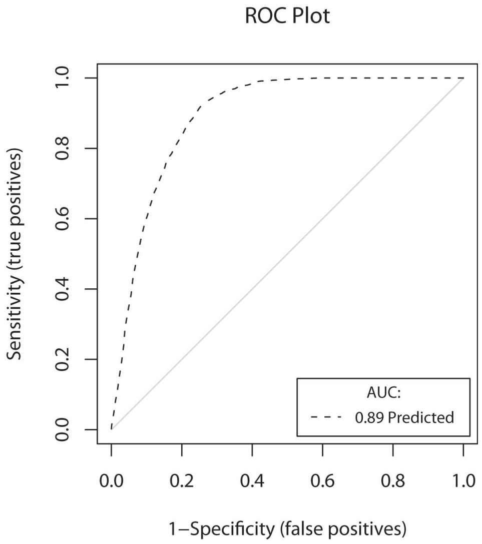


Figure A.27.: LRA for *Giraffa camelopardalis* (giraffe) ROC plot

Table A.15.: (continued) selected environmental variables and computed coefficients from stepwise analysis for *Giraffa camelopardalis* (giraffe)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
ev-h190	-6.164$e-01$	2.148$e-01$	-2.869	0.004119
ev-h193	-1.456$e-01$	8.601$e-02$	-1.693	0.090487
ev-h202	1.935$e+00$	3.250$e-01$	5.952	2.65e-09
ev-h205	1.056$e+00$	2.287$e-01$	4.617	3.90e-06
ev-h211	-3.285$e-01$	1.509$e-01$	-2.177	0.029478
ev-h218	-8.653$e-01$	4.046$e-01$	-2.139	0.032450
ev-h180	-9.743$e-02$	6.394$e-02$	-1.524	0.127533

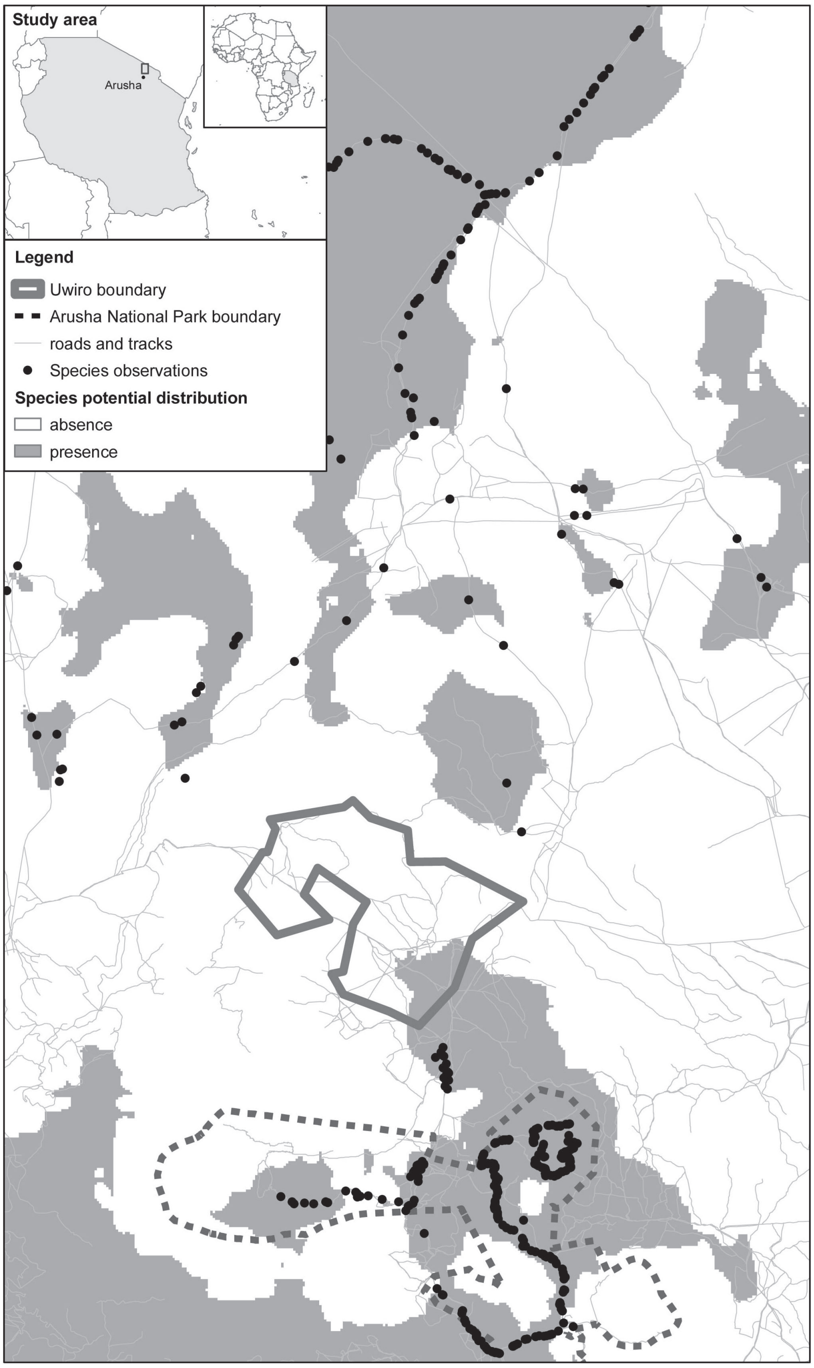


Figure A.28.: *Giraffa camelopardalis* (giraffe) potential distribution

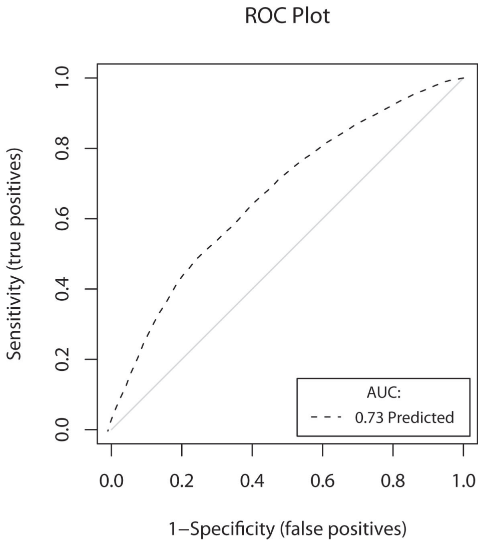


Figure A.29.: LRA for *Aepyceros melampus* (impala) ROC plot

Aepyceros melampus (impala)

Table A.16.: selected environmental variables and computed coefficients from stepwise analysis for *Aepyceros melampus* (impala)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	-14504.134	42211.323	-0.344	0.731
ev-cl17	-581.588	1574.567	-0.369	0.712
ev-swpa	1.009	2.717	0.371	0.710
ev-h104	-123.113	333.827	-0.369	0.712
ev-h115	-75.751	204.015	-0.371	0.710
ev-h122	166.355	451.791	0.368	0.713
ev-h123	105.363	286.392	0.368	0.713
ev-h126	86.776	234.983	0.369	0.712
ev-h140	22.567	60.445	0.373	0.709
ev-h142	-346.102	939.978	-0.368	0.713
ev-h145	141.878	382.805	0.371	0.711
ev-h169	-387.029	1047.026	-0.370	0.712
ev-h170	200.245	540.285	0.371	0.711
ev-h176	209.733	567.851	0.369	0.712
ev-h183	346.995	941.664	0.368	0.713
ev-h190	450.395	1225.006	0.368	0.713
ev-h214	-765.932	2082.276	-0.368	0.713
ev-h218	-81.751	222.008	-0.368	0.713

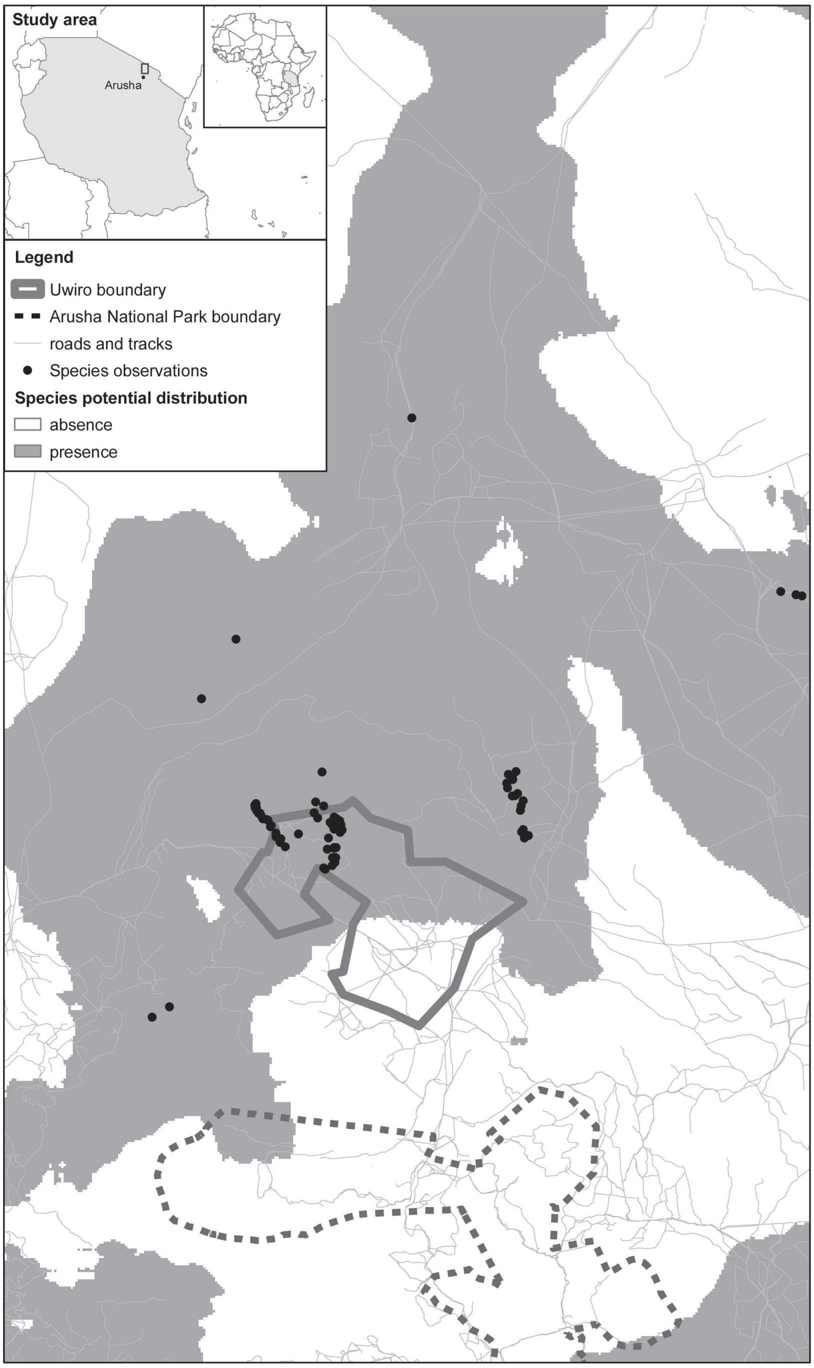


Figure A.30.: *Aepyceros melampus* (impala) potential distribution

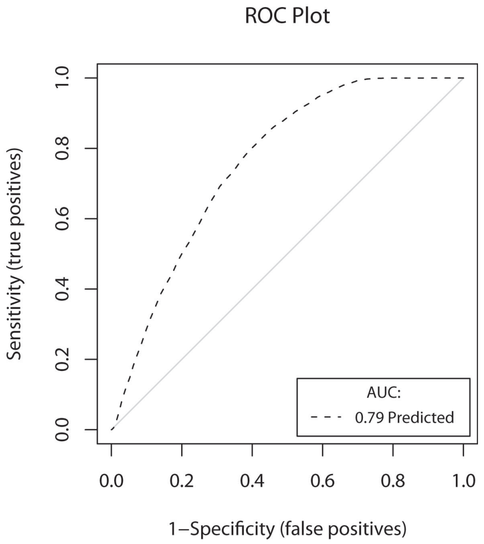


Figure A.31.: LRA for *Connochaetes taurinus* (wildbeest) ROC plot

Connochaetes taurinus (wildbeest)

Table A.17.: selected environmental variables and computed coefficients from stepwise analysis for *Connochaetes taurinus* (wildbeest)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	$1.216e+05$	$7.634e+05$	0.159	0.873
ev-cl04	$-1.374e+01$	$7.536e+01$	-0.182	0.855
ev-lake	$-7.703e-01$	$4.557e+00$	-0.169	0.866
ev-road	$-1.240e+00$	$6.833e+00$	-0.181	0.856
ev-h104	$1.066e+02$	$4.845e+02$	0.220	0.826
ev-h117	$-2.674e+01$	$1.165e+02$	-0.229	0.819
ev-h131	$-1.152e+02$	$7.542e+02$	-0.153	0.879
ev-h132	$1.714e+01$	$6.114e+01$	0.280	0.779
ev-h146	$-8.084e+01$	$4.907e+02$	-0.165	0.869
ev-h158	$6.331e+01$	$2.149e+02$	0.295	0.768
ev-h169	$-9.951e+01$	$3.885e+02$	-0.256	0.798
ev-h193	$-3.021e+01$	$1.008e+02$	-0.300	0.764
ev-h205	$1.255e+02$	$9.739e+02$	0.129	0.897

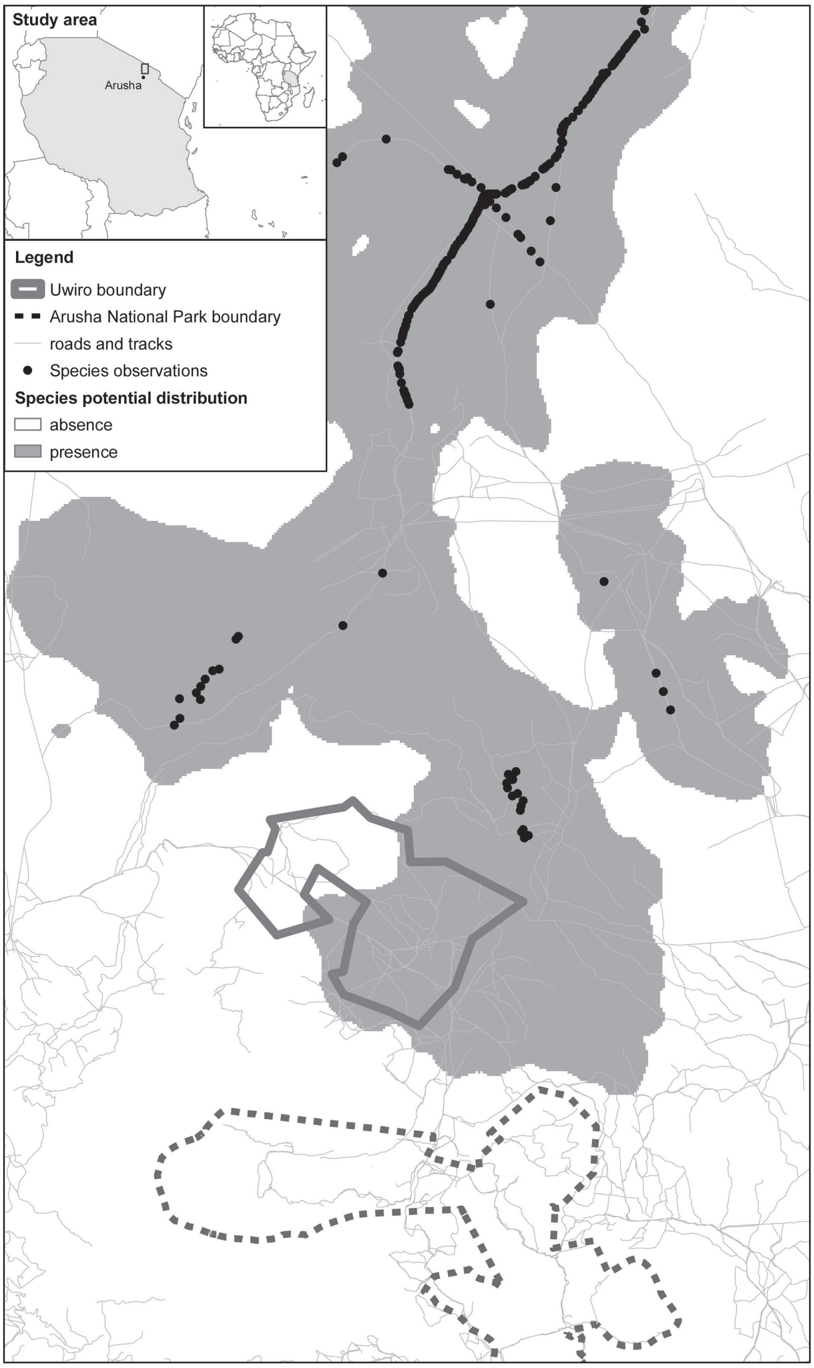


Figure A.32.: *Connochaetes taurinus* (wildbeest) potential distribution

Gazella granti (Grant's gazelle)

Table A.18.: selected environmental variables and computed coefficients from stepwise analysis for *Gazella granti* (Grant's gazelle)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$8.144e+04$	$3.237e+05$	0.252	0.801
ev-cl15	$4.608e+02$	$1.700e+03$	0.271	0.786
ev-cl17	$2.903e+02$	$1.018e+03$	0.285	0.775
ev-domi	$-2.656e+03$	$9.710e+03$	-0.274	0.784
ev-edde	$3.215e+00$	$1.189e+01$	0.270	0.787
ev-patc	$-1.203e+04$	$7.035e+04$	-0.171	0.864
ev-mpsi	$2.553e-04$	$1.007e-03$	0.254	0.800
ev-road	$-5.536e-01$	$2.685e+00$	-0.206	0.837
ev-rvrs	$5.101e-01$	$1.765e+00$	0.289	0.773
ev-vllg	$9.039e-01$	$2.978e+00$	0.304	0.761
ev-tasp	$-1.859e+00$	$7.756e+00$	-0.240	0.811
ev-h101	$-4.677e+01$	$3.027e+02$	-0.155	0.877
ev-h104	$-4.270e+01$	$1.193e+02$	-0.358	0.720
ev-h105	$-1.225e+02$	$5.020e+02$	-0.244	0.807
ev-h106	$2.630e+02$	$1.128e+03$	0.233	0.816
ev-h107	$1.398e+02$	$4.441e+02$	0.315	0.753
ev-h110	$2.889e+03$	$2.201e+04$	0.131	0.896
ev-h113	$-7.000e+01$	$2.962e+02$	-0.236	0.813
ev-h115	$2.437e+02$	$9.385e+02$	0.260	0.795
ev-h117	$-3.861e+01$	$2.784e+02$	-0.139	0.890
ev-h122	$1.928e+02$	$7.005e+02$	0.275	0.783
ev-h123	$-6.941e+01$	$2.612e+02$	-0.266	0.790
ev-h126	$-8.697e+01$	$3.234e+02$	-0.269	0.788
ev-h127	$-1.610e+02$	$7.414e+02$	-0.217	0.828
ev-h128	$-1.100e+02$	$4.957e+02$	-0.222	0.824
ev-h131	$-6.136e+01$	$2.749e+02$	-0.223	0.823
ev-h132	$1.981e+02$	$8.394e+02$	0.236	0.813
ev-h134	$-1.707e+02$	$6.696e+02$	-0.255	0.799
ev-h140	$-1.231e+02$	$5.250e+02$	-0.234	0.815
ev-h141	$-2.876e+03$	$2.192e+04$	-0.131	0.896
ev-h142	$-2.485e+02$	$9.598e+02$	-0.259	0.796
ev-h143	$-1.844e+02$	$8.725e+02$	-0.211	0.833
ev-h145	$-2.221e+02$	$8.651e+02$	-0.257	0.797
ev-h146	$8.847e+01$	$3.354e+02$	0.264	0.792
ev-h156	$-3.244e+02$	$1.555e+03$	-0.209	0.835
ev-h162	$2.720e+02$	$1.216e+03$	0.224	0.823
ev-h164	$1.160e+02$	$5.669e+02$	0.205	0.838
ev-h165	$1.807e+02$	$8.282e+02$	0.218	0.827
ev-h166	$-2.205e+02$	$8.755e+02$	-0.252	0.801
ev-h168	$5.607e+02$	$2.110e+03$	0.266	0.790
ev-h169	$-1.163e+02$	$4.115e+02$	-0.283	0.777
ev-h170	$-4.203e+01$	$2.156e+02$	-0.195	0.845

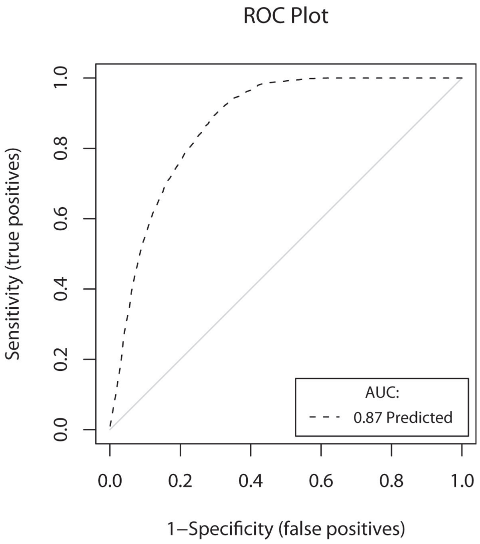


Figure A.33.: LRA for *Gazella granti* (Grant's gazelle) ROC plot

Table A.18.: (continued) selected environmental variables and computed coefficients from stepwise analysis for *Gazella granti* (Grant's gazelle)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
ev-h174	$4.551e+01$	$2.541e+02$	0.179	0.858
ev-h175	$1.425e+02$	$6.862e+02$	0.208	0.836
ev-h176	$-1.124e+02$	$3.072e+02$	-0.366	0.714
ev-h177	$1.272e+02$	$5.865e+02$	0.217	0.828
ev-h178	$-6.056e+01$	$1.994e+02$	-0.304	0.761
ev-h180	$2.252e+02$	$8.339e+02$	0.270	0.787
ev-h182	$7.035e+01$	$2.207e+02$	0.319	0.750
ev-h183	$-3.637e+02$	$1.202e+03$	-0.303	0.762
ev-h184	$1.669e+02$	$7.616e+02$	0.219	0.827
ev-h190	$-8.324e+01$	$2.934e+02$	-0.284	0.777
ev-h193	$-2.313e+02$	$8.901e+02$	-0.260	0.795
ev-h205	$3.475e+02$	$1.298e+03$	0.268	0.789
ev-h207	$-2.344e+02$	$9.731e+02$	-0.241	0.810
ev-h214	$-1.302e+02$	$6.987e+02$	-0.186	0.852
ev-h216	$2.090e+02$	$7.554e+02$	0.277	0.782
ev-h218	$-4.517e+02$	$1.763e+03$	-0.256	0.798

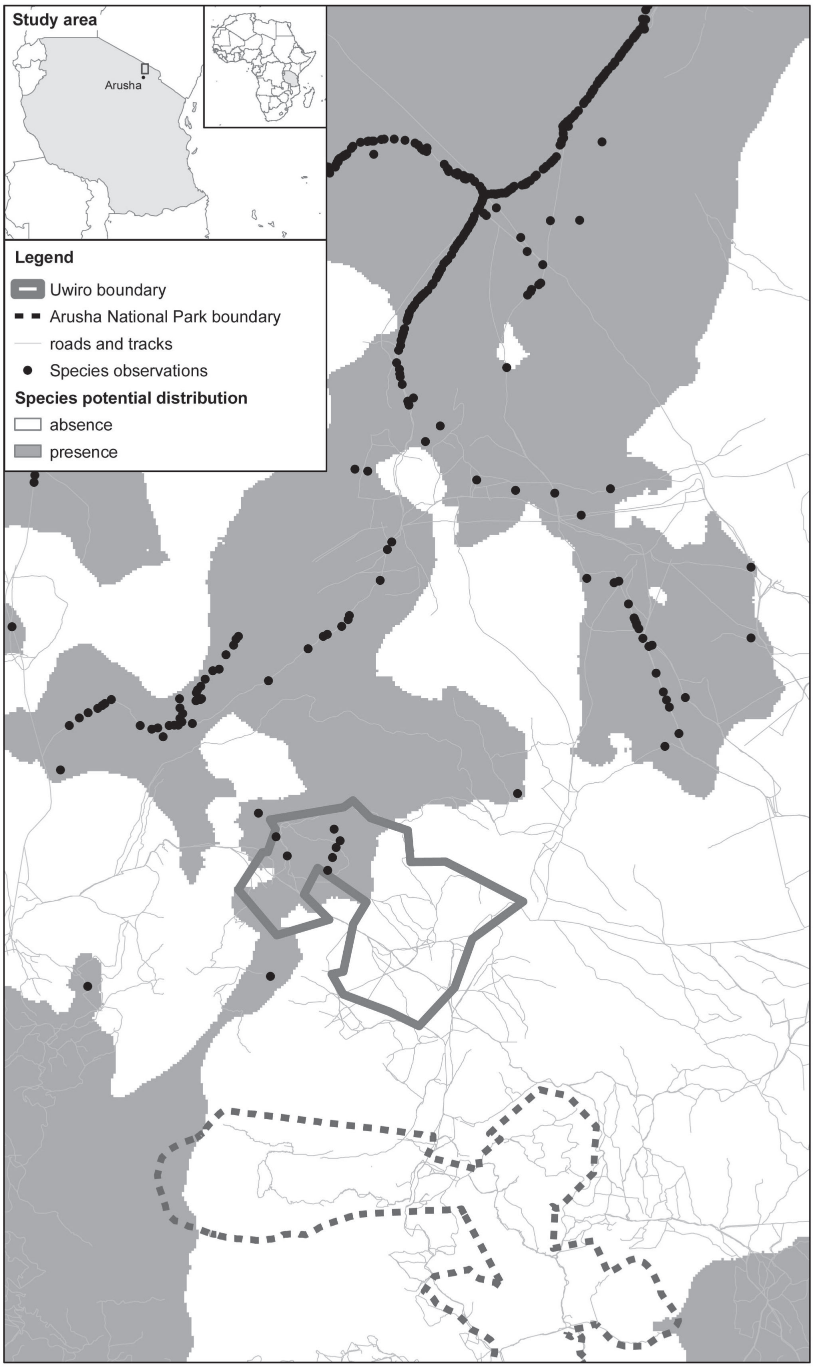


Figure A.34.: *Gazella granti* (Grant's gazelle) potential distribution

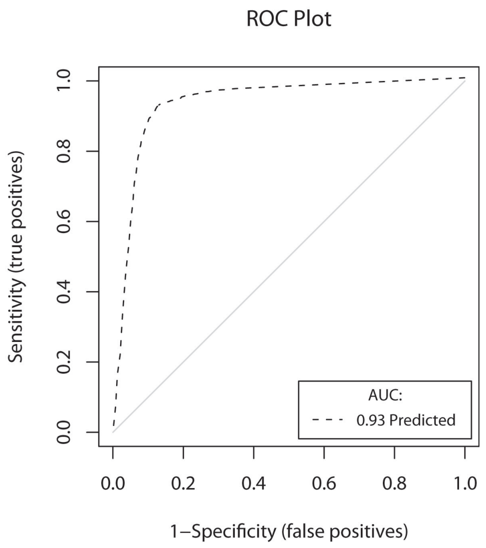


Figure A.35.: LRA for *Gazella thomsonii* (Thomson gazelle) ROC plot

Gazella thomsonii (Thomson gazelle)

Table A.19.: selected environmental variables and computed coefficients from stepwise analysis for *Gazella thomsonii* (Thomson gazelle)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	3.156e+04	1.144e+06	0.028	0.978
ev-shan	-7.863e+02	3.665e+04	-0.021	0.983
ev-mpsi	8.154e−05	2.040e−03	0.040	0.968
ev-rvrs	7.352e−02	1.239e+01	0.006	0.995
ev-vllg	-7.689e−02	3.547e+00	-0.022	0.983
ev-h105	2.743e+01	1.635e+03	0.017	0.987
ev-h113	-2.567e+01	1.558e+03	-0.016	0.987
ev-h117	-1.832e+01	5.034e+02	-0.036	0.971
ev-h122	5.193e+01	1.211e+03	0.043	0.966
ev-h127	-5.324e+01	2.210e+03	-0.024	0.981
ev-h131	-1.608e+01	3.610e+02	-0.045	0.964
ev-h140	2.638e+01	1.240e+03	0.021	0.983
ev-h146	1.528e+01	5.305e+02	0.029	0.977
ev-h162	6.056e+01	1.768e+03	0.034	0.973
ev-h167	-7.128e+01	1.375e+03	-0.052	0.959
ev-h174	-3.273e+01	1.551e+03	-0.021	0.983
ev-h193	-3.896e+01	1.619e+03	-0.024	0.981
ev-h202	3.932e+01	1.698e+03	0.023	0.982

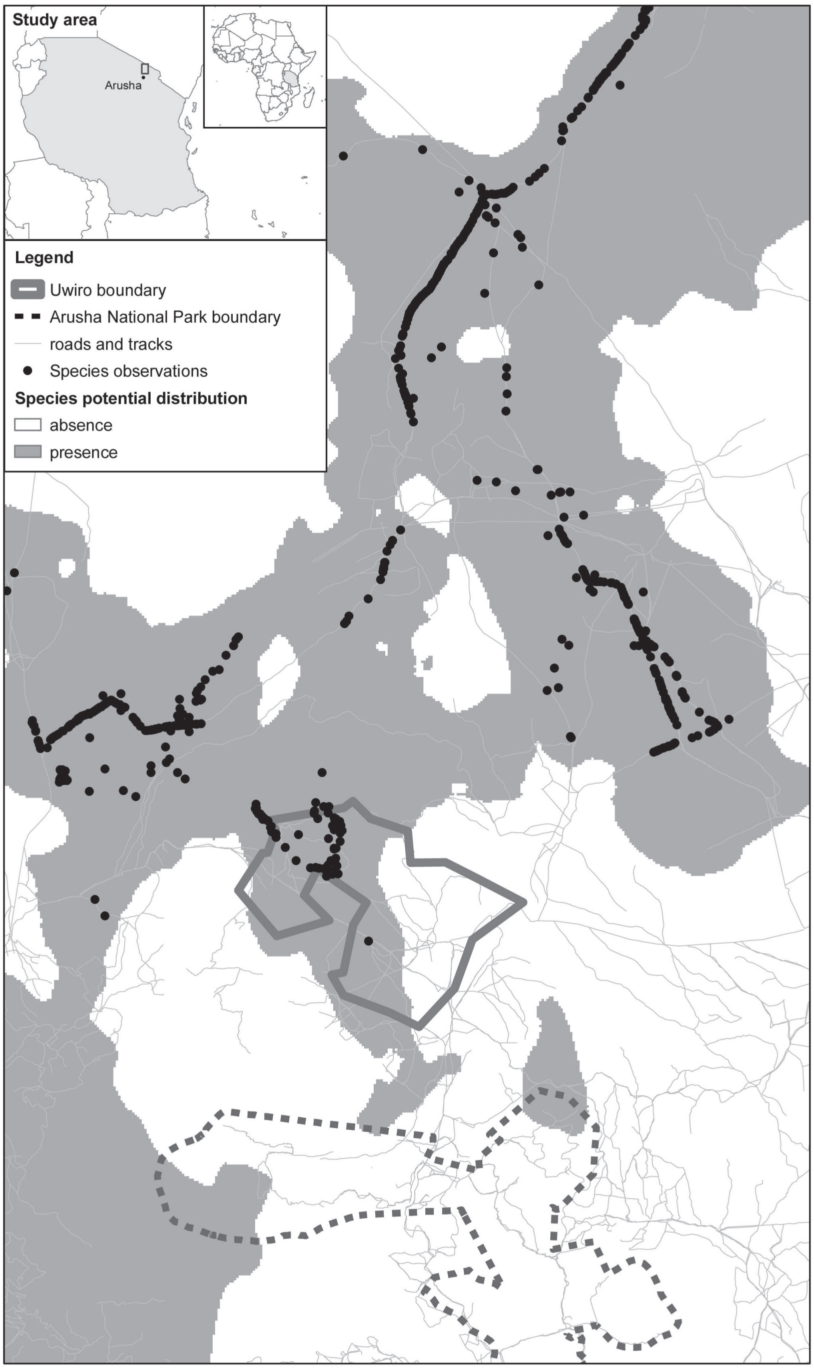


Figure A.36.: *Gazella thomsonii* (Thomson gazelle) potential distribution

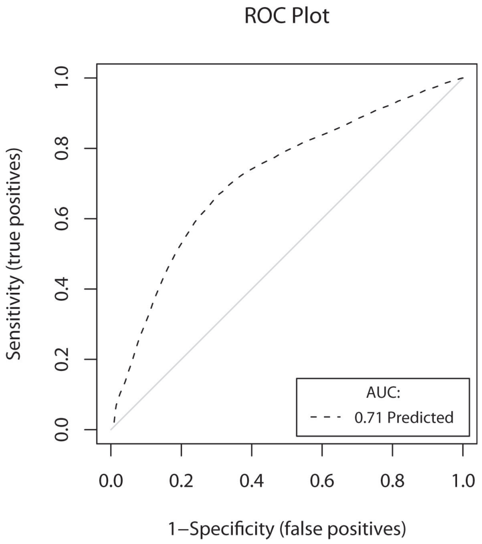


Figure A.37.: LRA for *Litocranius walleri* (gerenuk) ROC plot

Litocranius walleri (gerenuk)

Table A.20.: selected environmental variables and computed coefficients from stepwise analysis for *Litocranius walleri* (gerenuk)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	28890.4402	73437.3610	0.393	0.694
ev-cl17	-430.1806	1388.4208	-0.310	0.757
ev-edde	-0.8873	2.6924	-0.330	0.742
ev-lake	-0.6066	1.2743	-0.476	0.634
ev-vllg	-0.9585	1.9546	-0.490	0.624
ev-h101	-79.8951	5147.0425	-0.016	0.988
ev-h108	-43.9618	120.3239	-0.365	0.715
ev-h110	61.1317	243.3591	0.251	0.802
ev-h127	-53.6930	159.5818	-0.336	0.737
ev-h132	25.6294	59.7809	0.429	0.668
ev-h162	96.3214	199.3937	0.483	0.629
ev-h166	-10.7185	25.3334	-0.423	0.672
ev-h170	-82.5812	168.3068	-0.491	0.624
ev-h180	-87.3043	185.8965	-0.470	0.639
ev-h183	56.5884	146.3579	0.387	0.699
ev-h193	47.8303	99.1150	0.483	0.629
ev-h207	-55.0920	114.3827	-0.482	0.630

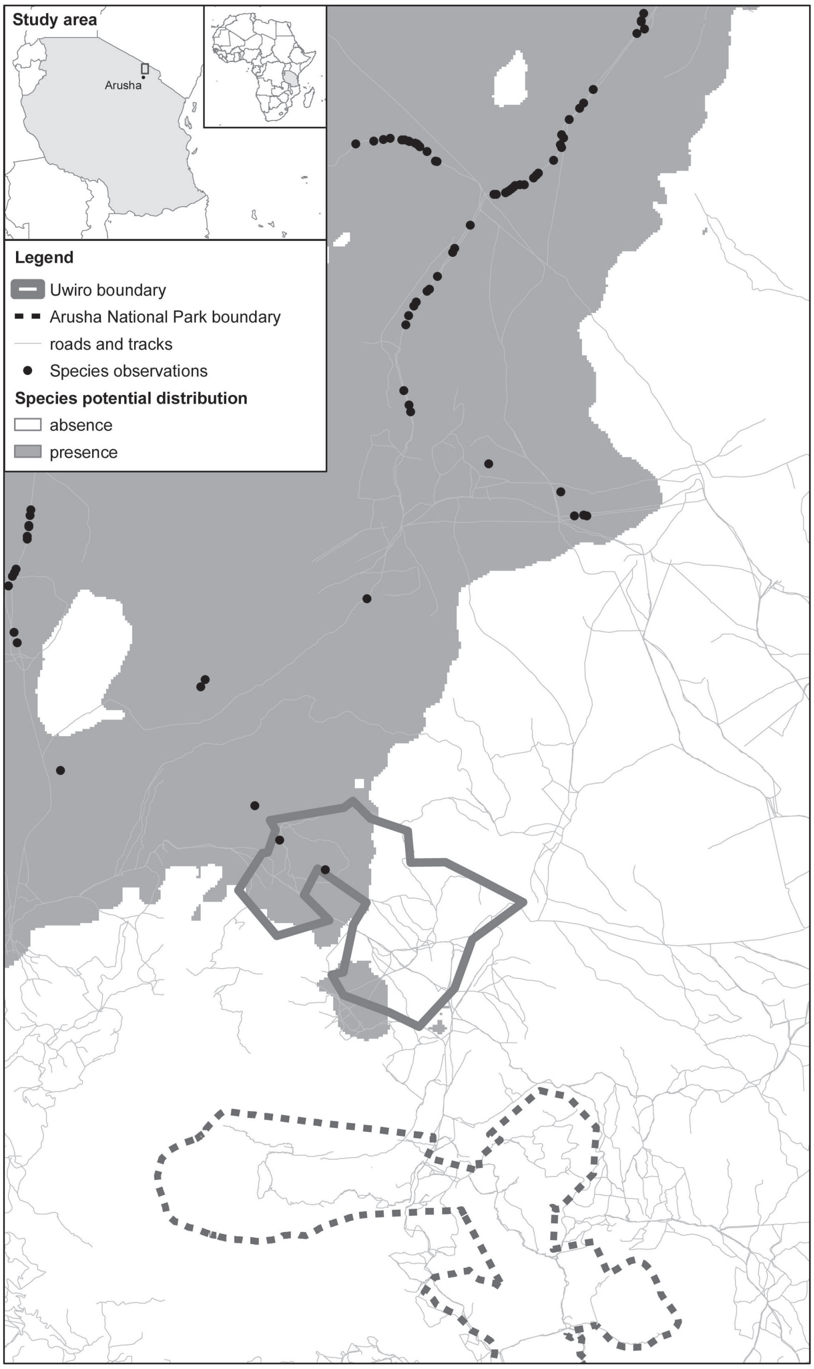


Figure A.38.: *Litocranius walleri* (gerenuk) potential distribution

Madoqua kirkii (gerenuk)

Table A.21.: selected environmental variables and computed coefficients from stepwise analysis for *Madoqua kirkii* (gerenuk)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	3.880e + 04	2.063e + 05	0.188	0.851
ev-cl04	1.802e + 01	5.346e + 01	0.337	0.736
ev-cl09	4.446e + 02	1.398e + 03	0.318	0.750
ev-cl15	6.404e + 02	1.723e + 03	0.372	0.710
ev-cl17	-2.160e + 02	7.145e + 02	-0.302	0.762
ev-domi	-1.842e + 03	7.125e + 03	-0.258	0.796
ev-edde	3.104e + 00	1.118e + 01	0.278	0.781
ev-patc	-4.459e + 04	1.460e + 05	-0.306	0.760
ev-mpsi	-9.276e − 04	2.935e − 03	-0.316	0.752
ev-road	-1.331e + 00	4.039e + 00	-0.330	0.742
ev-rvrs	2.612e − 01	9.398e − 01	0.278	0.781
ev-swpa	9.614e − 01	3.215e + 00	0.299	0.765
ev-vllg	-1.762e − 01	5.354e − 01	-0.329	0.742
ev-tasp	1.988e + 00	6.056e + 00	0.328	0.743
ev-tslp	1.732e + 02	5.144e + 02	0.337	0.736
ev-h105	-4.809e + 01	1.640e + 02	-0.293	0.769
ev-h106	3.956e + 01	1.236e + 02	0.320	0.749
ev-h107	-5.202e + 02	1.548e + 03	-0.336	0.737
ev-h108	-6.767e + 01	2.164e + 02	-0.313	0.755
ev-h113	-5.778e + 01	1.767e + 02	-0.327	0.744
ev-h115	-7.018e + 00	3.234e + 01	-0.217	0.828
ev-h117	-2.778e + 02	8.585e + 02	-0.324	0.746
ev-h123	1.102e + 02	3.113e + 02	0.354	0.723
ev-h126	-1.279e + 02	4.059e + 02	-0.315	0.753
ev-h127	2.590e + 02	8.599e + 02	0.301	0.763
ev-h134	-5.342e + 01	1.465e + 02	-0.365	0.715
ev-h140	-8.717e + 01	2.606e + 02	-0.334	0.738
ev-h141	-5.879e + 02	1.742e + 03	-0.337	0.736
ev-h143	1.474e + 02	4.538e + 02	0.325	0.745
ev-h145	1.653e + 02	5.540e + 02	0.298	0.765
ev-h146	-9.617e + 01	2.754e + 02	-0.349	0.727
ev-h156	-2.502e + 02	7.236e + 02	-0.346	0.729
ev-h158	3.366e + 02	1.010e + 03	0.333	0.739
ev-h162	7.674e + 01	2.394e + 02	0.321	0.749
ev-h164	1.422e + 02	4.333e + 02	0.328	0.743
ev-h165	1.275e + 02	3.914e + 02	0.326	0.745
ev-h166	-2.257e + 01	1.458e + 02	-0.155	0.877
ev-h168	2.168e + 02	6.515e + 02	0.333	0.739
ev-h170	9.079e + 01	2.748e + 02	0.330	0.741
ev-h177	8.532e + 01	2.549e + 02	0.335	0.738
ev-h178	4.143e + 01	1.745e + 02	0.237	0.812
ev-h180	-3.095e + 01	1.033e + 02	-0.300	0.764

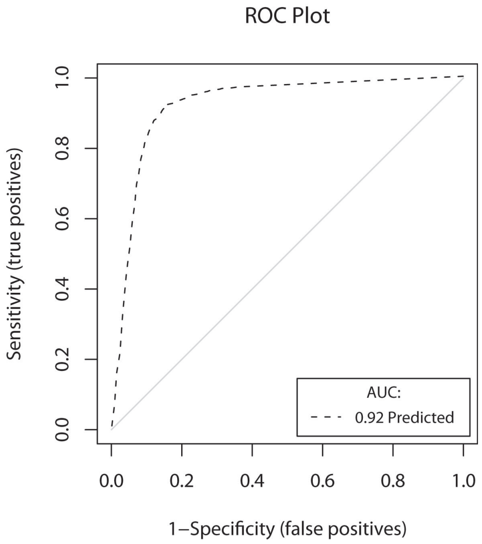


Figure A.39.: LRA for *Madoqua kirkii* (gerenuk) ROC plot

Table A.21.: (continued) selected environmental variables and computed coefficients from stepwise analysis for *Madoqua kirkii* (gerenuk)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
ev-h182	-2.726$e+02$	8.633$e+02$	-0.316	0.752
ev-h183	-1.216$e+02$	4.191$e+02$	-0.290	0.772
ev-h190	6.655$e+01$	3.273$e+02$	0.203	0.839
ev-h202	1.018$e+02$	2.904$e+02$	0.351	0.726
ev-h205	1.621$e+02$	5.210$e+02$	0.311	0.756
ev-h211	-1.197$e+02$	4.275$e+02$	-0.280	0.779
ev-h214	-1.968$e+02$	8.037$e+02$	-0.245	0.807
ev-h216	2.761$e+02$	1.062$e+03$	0.260	0.795
ev-h218	-3.502$e+02$	1.081$e+03$	-0.324	0.746

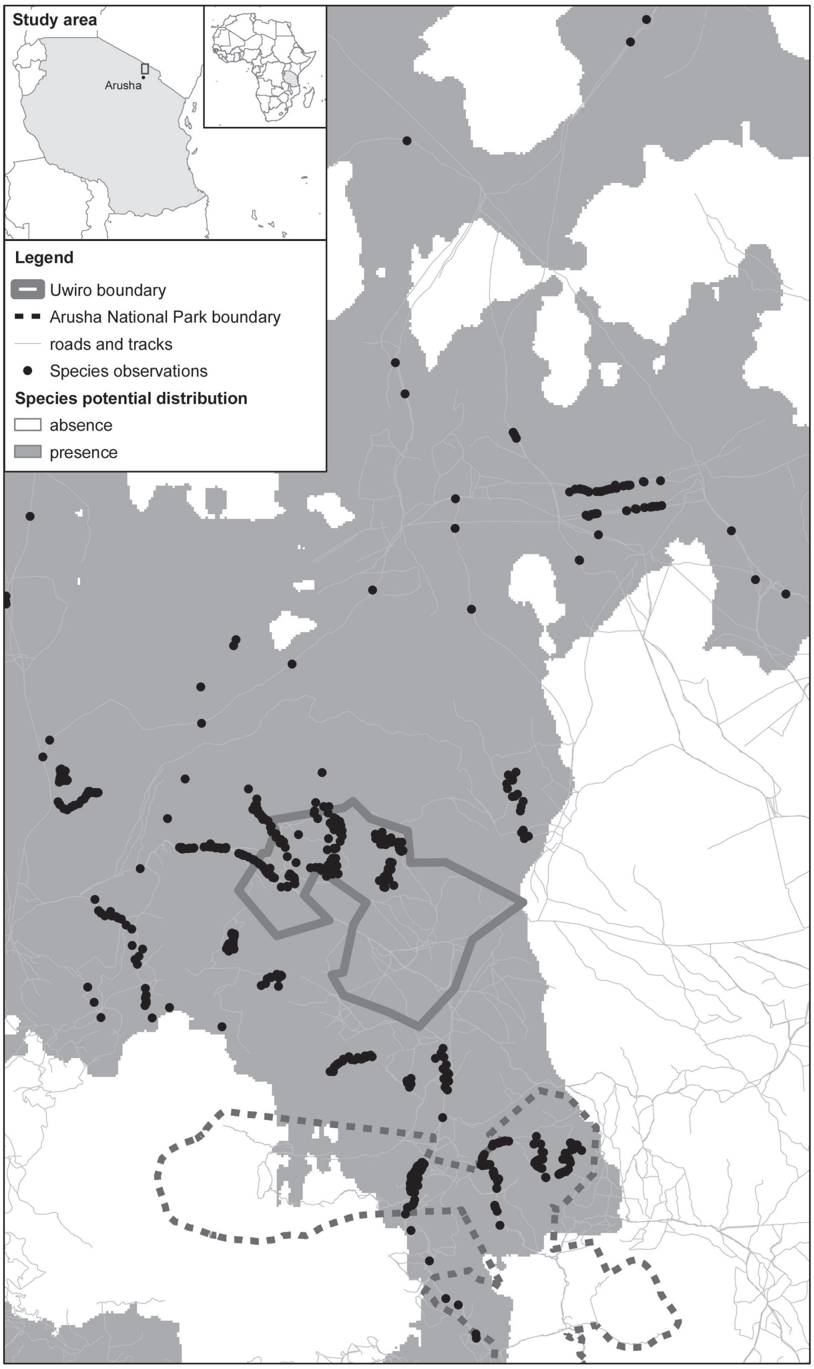


Figure A.40.: *Madoqua kirkii* (gerenuk) potential distribution

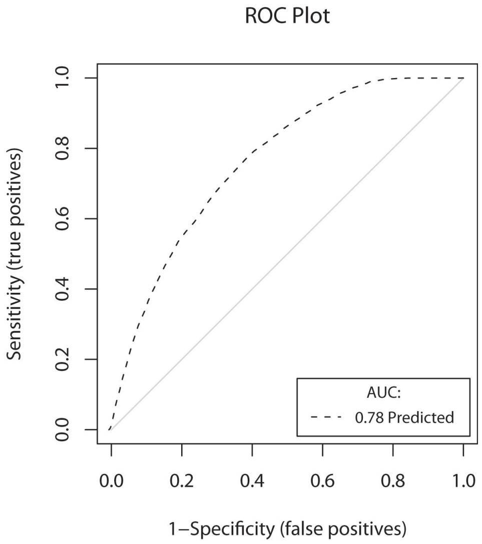


Figure A.41.: LRA for *Raphicerus campestris* (steinbuck) ROC plot

Raphicerus campestris (steinbuck)

Table A.22.: selected environmental variables and computed coefficients from stepwise analysis for *Raphicerus campestris* (steinbuck)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$2.049e+05$	$74.450e+05$	0.460	0.645
ev-cl09	$3.629e+02$	$7.877e+02$	0.461	0.645
ev-domi	-$4.997e+03$	$1.079e+04$	-0.463	0.643
ev-road	-$1.858e+00$	$4.059e+00$	-0.458	0.647
ev-rvrs	$8.792e-01$	$1.946e+00$	0.452	0.651
ev-vllg	-$1.231e+00$	$2.678e+00$	-0.460	0.646
ev-h101	-$1.149e+02$	$2.544e+02$	-0.452	0.651
ev-h115	$2.113e+02$	$4.595e+02$	0.460	0.646
ev-h122	$1.817e+02$	$3.964e+02$	0.458	0.647
ev-h134	$1.053e+02$	$2.311e+02$	0.456	0.649
ev-h145	$1.890e+02$	$4.100e+02$	0.461	0.645
ev-h167	-$5.309e+02$	$1.153e+03$	-0.461	0.645
ev-h174	-$3.112e+02$	$6.759e+02$	-0.460	0.645
ev-h190	-$7.881e+02$	$1.710e+03$	-0.461	0.645
ev-h193	-$2.214e+02$	$4.832e+02$	-0.458	0.647
ev-h202	$2.755e+02$	$5.976e+02$	0.461	0.645
ev-h207	-$3.188e+02$	$6.951e+02$	-0.459	0.646
ev-h211	-$3.304e+02$	$7.150e+02$	-0.462	0.644
ev-h214	$5.613e+02$	$1.216e+03$	0.461	0.644
ev-h216	$1.676e+02$	$3.663e+02$	0.458	0.647

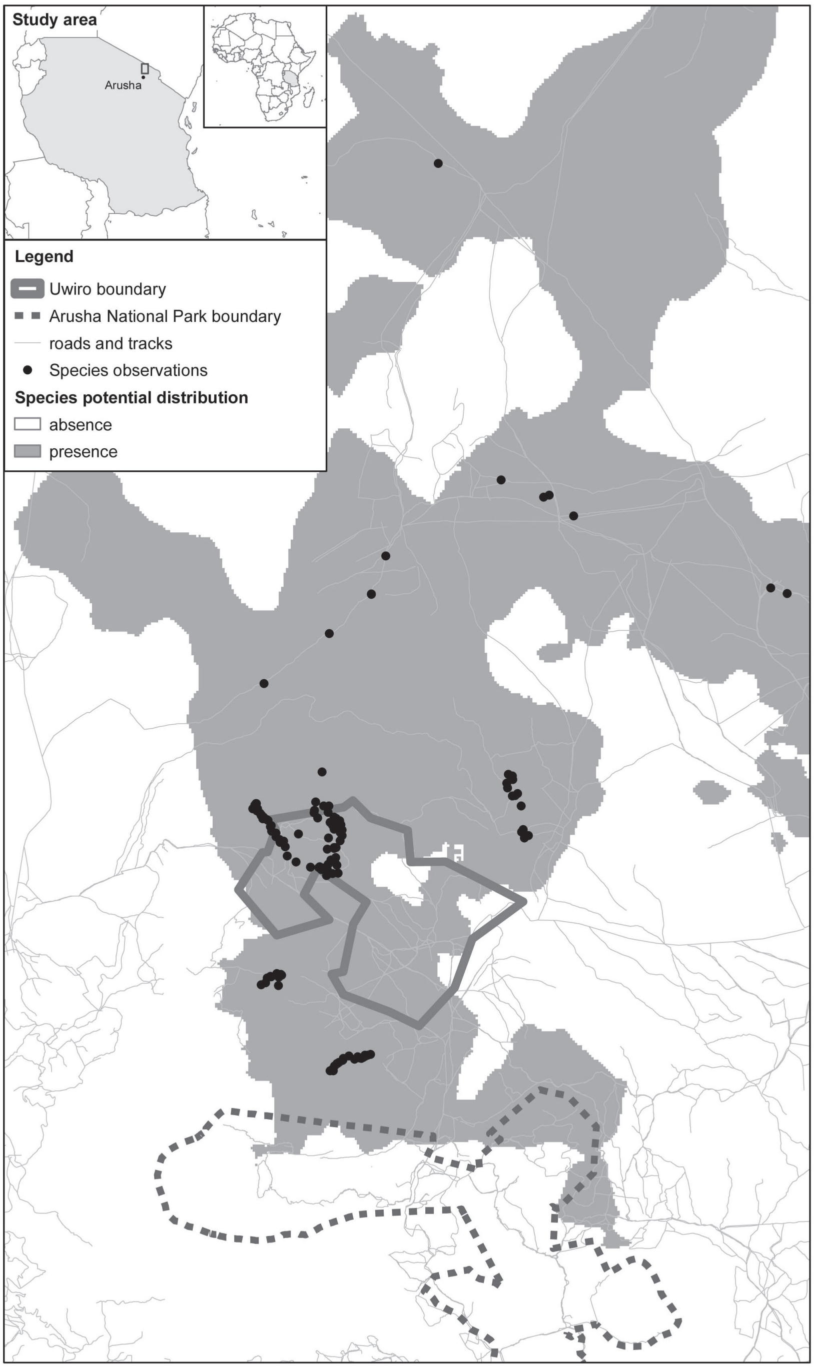


Figure A.42.: *Raphicerus campestris* (steinbuck) potential distribution

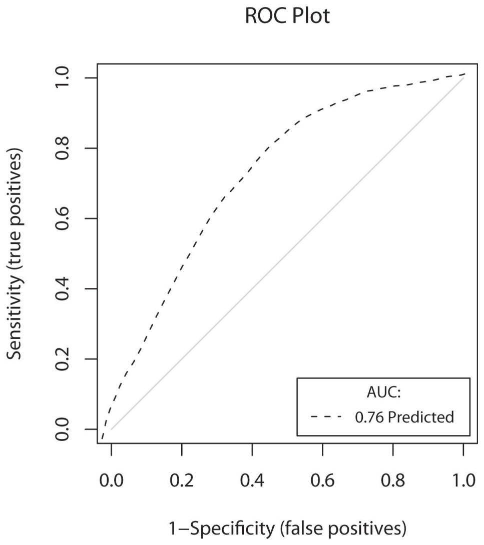


Figure A.43.: LRA for *Syncerus caffer* (african buffalo) ROC plot

Syncerus caffer (african buffalo)

Table A.23.: selected environmental variables and computed coefficients from stepwise analysis for *Syncerus caffer* (african buffalo)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	4920.393	179.793	0.027	0.978
ev-cl17	-11.063	642.290	-0.017	0.986
ev-h117	-1.628	29.157	-0.056	0.955
ev-h132	-6.453	221.015	-0.029	0.977
ev-h158	-18.869	316.635	-0.060	0.952
ev-h168	16.269	188.291	0.086	0.931
ev-h177	-4.326	78.446	-0.055	0.956
ev-h193	-3.441	156.950	-0.022	0.983
ev-h211	-2.912	119.198	-0.024	0.981
ev-h214	-4.887	221.804	-0.022	0.982

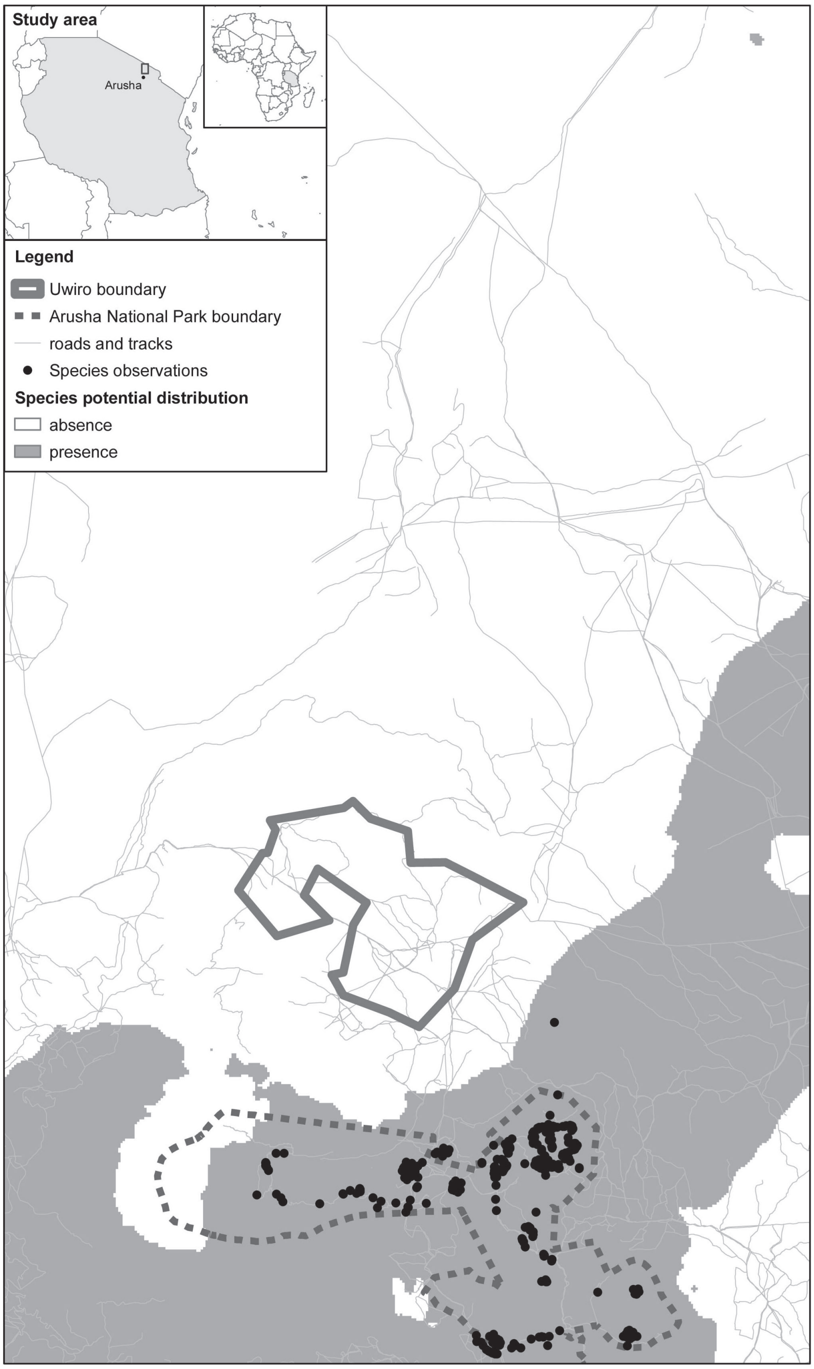


Figure A.44.: *Syncerus caffer* (african buffalo) potential distribution

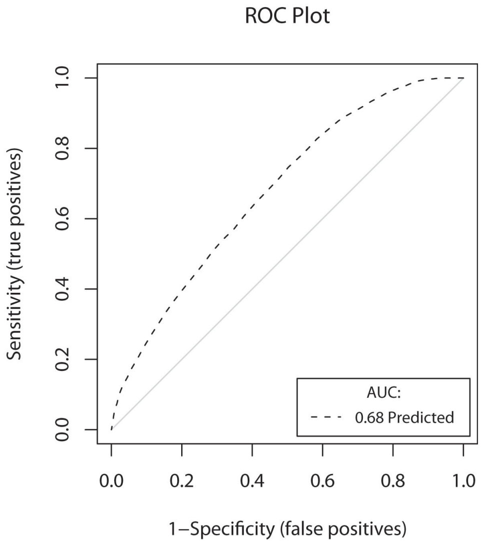


Figure A.45.: LRA for *Tragelaphus imberbis* (lesserkudu) ROC plot

Tragelaphus imberbis (lesserkudu)

Table A.24.: selected environmental variables and computed coefficients from stepwise analysis for *Tragelaphus imberbis* (lesserkudu)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	-1.874$e+05$	3.185$e+05$	-0.588	0.556
ev-cl04	2.966$e+01$	5.150$e+01$	0.576	0.565
ev-shan	-3.744$e+03$	6.334$e+03$	-0.591	0.554
ev-lake	-2.010$e+00$	3.436$e+00$	-0.585	0.559
ev-h107	2.915$e+02$	4.897$e+02$	0.595	0.552
ev-h115	-2.674$e+01$	4.570$e+01$	-0.585	0.558
ev-h131	5.742$e+01$	9.815$e+01$	0.585	0.559
ev-h158	-1.780$e+02$	3.038$e+02$	-0.586	0.558
ev-h166	-2.427$e+02$	4.151$e+02$	-0.585	0.559
ev-h167	-2.492$e+02$	4.274$e+02$	-0.583	0.560
ev-h168	-3.057$e+02$	5.256$e+02$	-0.582	0.561
ev-h178	1.859$e+02$	3.217$e+02$	0.578	0.563
ev-h180	-1.289$e+02$	2.203$e+02$	-0.585	0.558
ev-h193	-1.586$e+02$	2.684$e+02$	-0.591	0.554
ev-h214	-4.471$e+02$	7.565$e+02$	-0.591	0.555
ev-h218	1.954$e+02$	3.359$e+02$	0.582	0.561

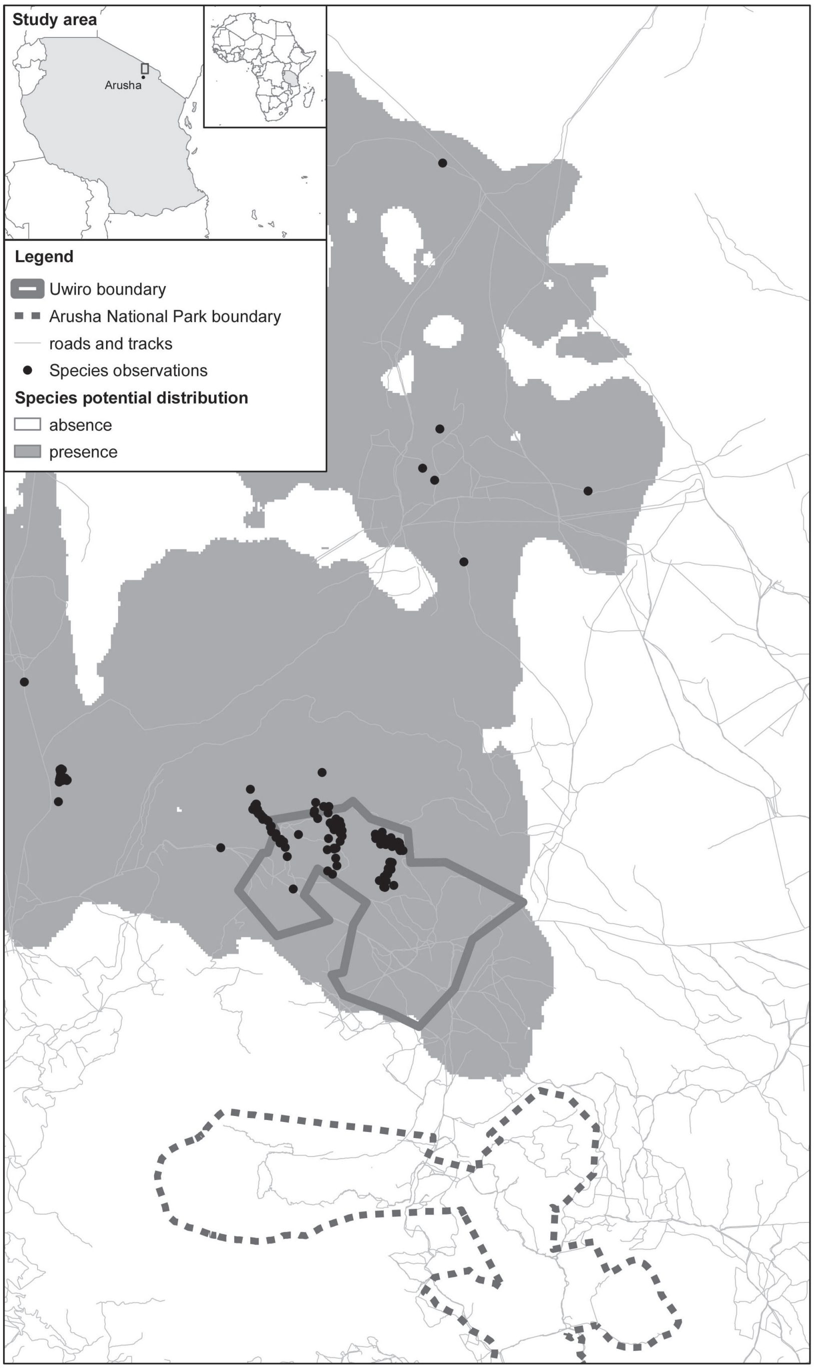


Figure A.46.: *Tragelaphus imberbis* (lesserkudu) potential distribution

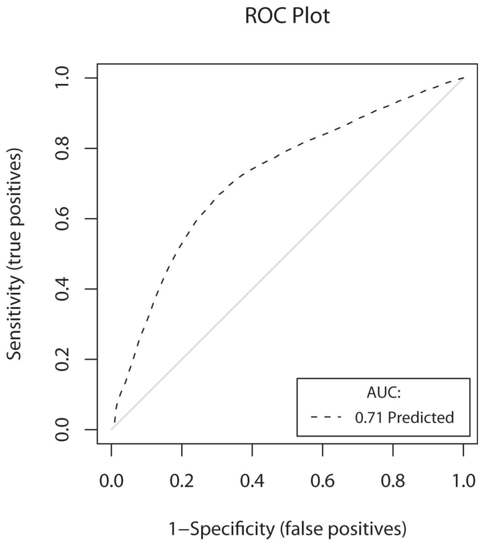


Figure A.47.: LRA for *Tragelaphus scriptus* (bushbuck) ROC plot

Tragelaphus scriptus (bushbuck)

Table A.25.: selected environmental variables and computed coefficients from stepwise analysis for *Tragelaphus scriptus* (bushbuck)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	$5.738e+03$	$1.677e+05$	0.034	0.973
ev-cl15	$-1.084e+02$	$1.118e+03$	-0.097	0.923
ev-edde	$1.057e+00$	$2.447e+01$	0.043	0.966
ev-lake	$1.477e-01$	$3.058e+00$	0.048	0.961
ev-h101	$-1.447e+01$	$2.523e+02$	-0.057	0.954
ev-h108	$1.841e+01$	$3.042e+02$	0.061	0.952
ev-h123	$-2.582e+01$	$5.068e+02$	-0.051	0.959
ev-h134	$4.451e+01$	$4.792e+02$	0.093	0.926
ev-h164	$-8.911e+01$	$1.150e+03$	-0.078	0.938
ev-h168	$7.924e+01$	$9.368e+02$	0.085	0.933
ev-h174	$1.848e+01$	$2.884e+02$	0.064	0.949
ev-h207	$-4.262e+01$	$5.591e+02$	-0.076	0.939
ev-h211	$-6.331e+00$	$2.005e+02$	-0.032	0.975
ev-h214	$-6.871e+01$	$8.204e+02$	-0.084	0.933
ev-h216	$4.842e+01$	$5.489e+02$	0.088	0.930

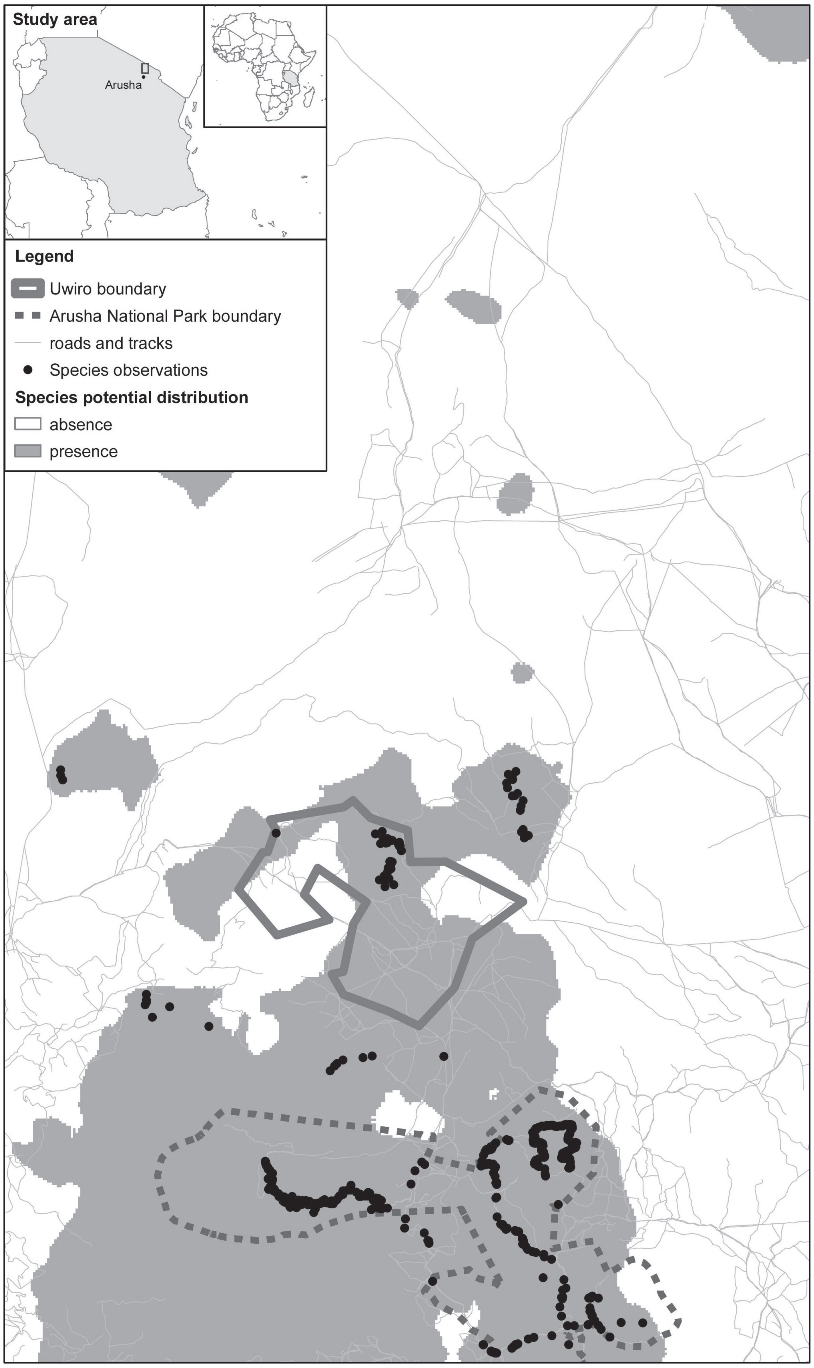


Figure A.48.: *Tragelaphus scriptus* (bushbuck) potential distribution

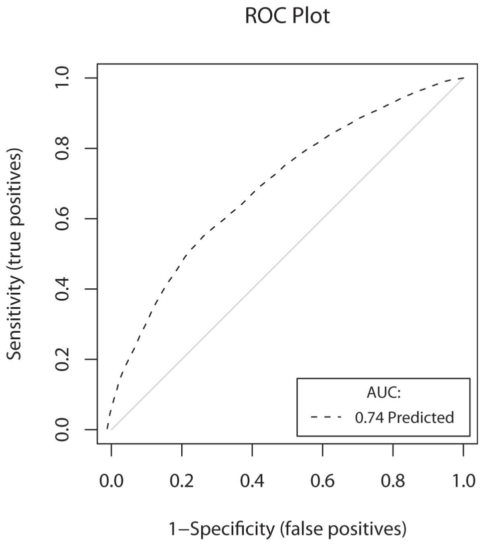


Figure A.49.: LRA for *Cephalophus harveyi* (Harvey's red duiker) ROC plot

Cephalophus harveyi (Harvey's red duiker)

Table A.26.: selected environmental variables and computed coefficients from stepwise analysis for *Cephalophus harveyi* (Harvey's red duiker)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	-9.602$e+03$	1.580$e+05$	-0.061	0.952
ev-edde	5.176$e-01$	1.364$e+01$	0.038	0.970
ev-tslp	1.217$e+01$	2.306$e+02$	0.053	0.958
ev-h105	-3.214$e+01$	5.761$e+02$	-0.056	0.956
ev-h110	8.651$e+01$	1.017$e+04$	0.009	0.993
ev-h113	1.054$e+01$	2.145$e+02$	0.049	0.961
ev-h128	-2.585$e+01$	4.249$e+02$	-0.061	0.951
ev-h141	-1.087$e+02$	1.021$e+04$	-0.011	0.992
ev-h142	-1.198$e+01$	2.301$e+02$	-0.052	0.958
ev-h143	-3.287$e+01$	5.303$e+02$	-0.062	0.951
ev-h145	-8.877$e+00$	1.623$e+02$	-0.055	0.956
ev-h146	-3.175$e+01$	5.453$e+02$	-0.058	0.954
ev-h165	-4.316$e+01$	6.794$e+02$	-0.064	0.949
ev-h174	-7.323$e+00$	1.595$e+02$	-0.046	0.963
ev-h183	4.421$e+01$	6.894$e+02$	0.064	0.949
ev-h184	4.300$e+01$	7.749$e+02$	0.055	0.956
ev-h190	2.570$e+01$	3.986$e+02$	0.064	0.949
ev-h193	3.490$e+01$	6.046$e+02$	0.058	0.954
ev-h202	1.015$e+01$	1.803$e+02$	0.056	0.955
ev-h211	1.320$e+01$	2.246$e+02$	0.059	0.953

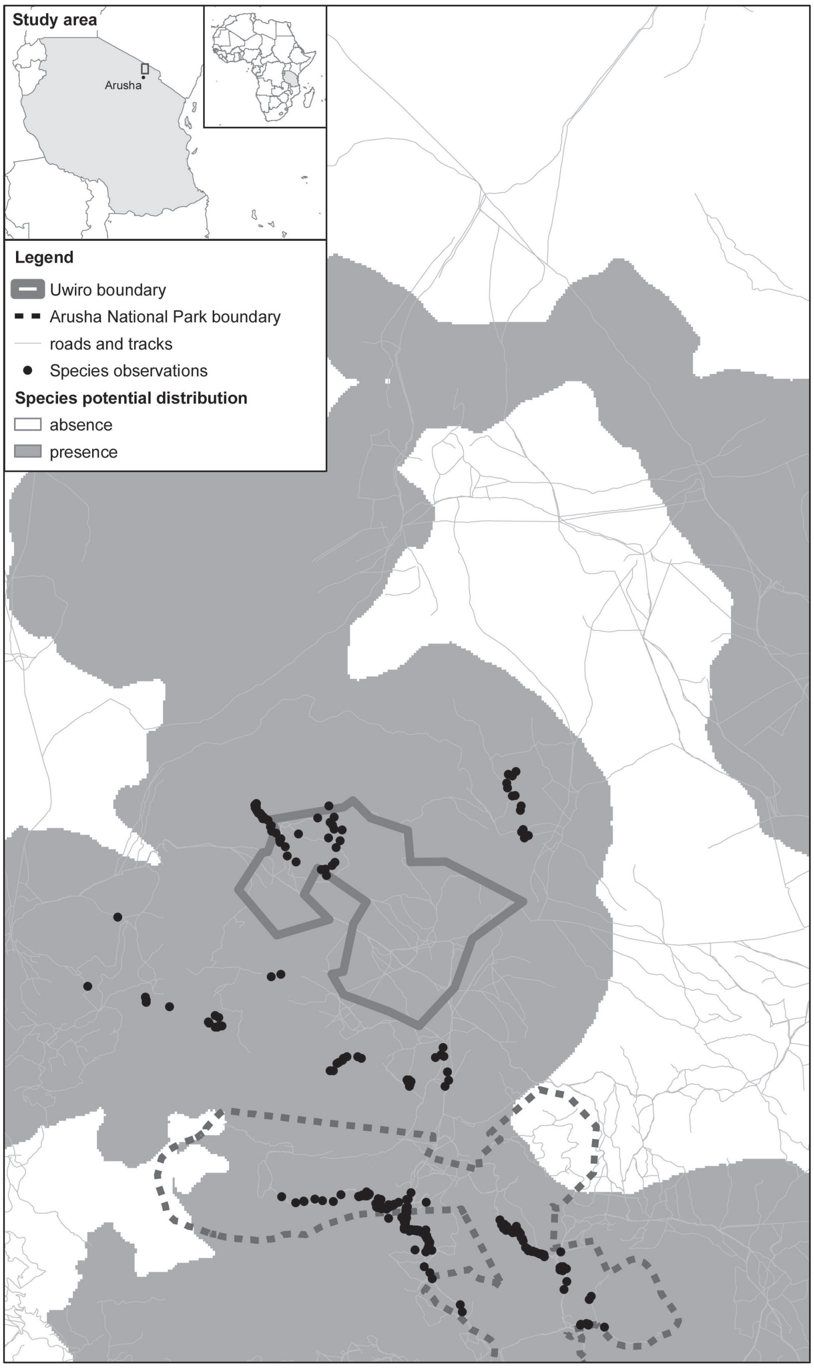


Figure A.50.: *Cephalophus harveyi* (Harvey's red duiker) potential distribution

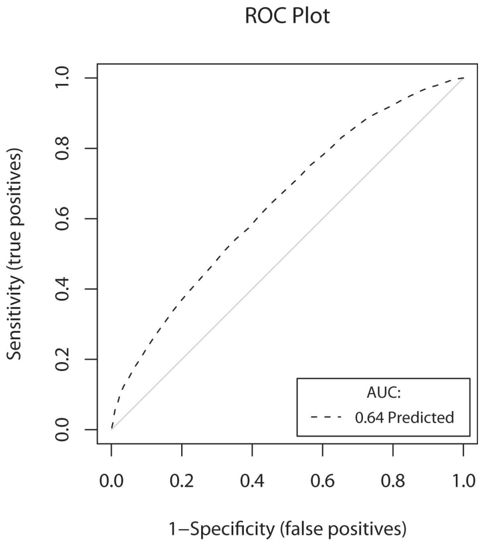


Figure A.51.: LRA for *Kobus ellipsiprymnus* (waterbuck) ROC plot

Kobus ellipsiprymnus (waterbuck)

Table A.27.: selected environmental variables and computed coefficients from stepwise analysis for *Kobus ellipsiprymnus* (waterbuck)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	5495.3817	28506.5001	0.193	0.847
ev-cl09	-9.9757	57.9183	-0.172	0.863
ev-road	-0.3691	2.9067	-0.127	0.899
ev-h106	5.6266	53.4116	0.105	0.916
ev-h117	-10.4703	72.3244	-0.145	0.885
ev-h126	-7.6000	44.3052	-0.172	0.864
ev-h193	-10.9940	51.1660	-0.215	0.830
ev-h214	-13.0086	66.0502	-0.197	0.844

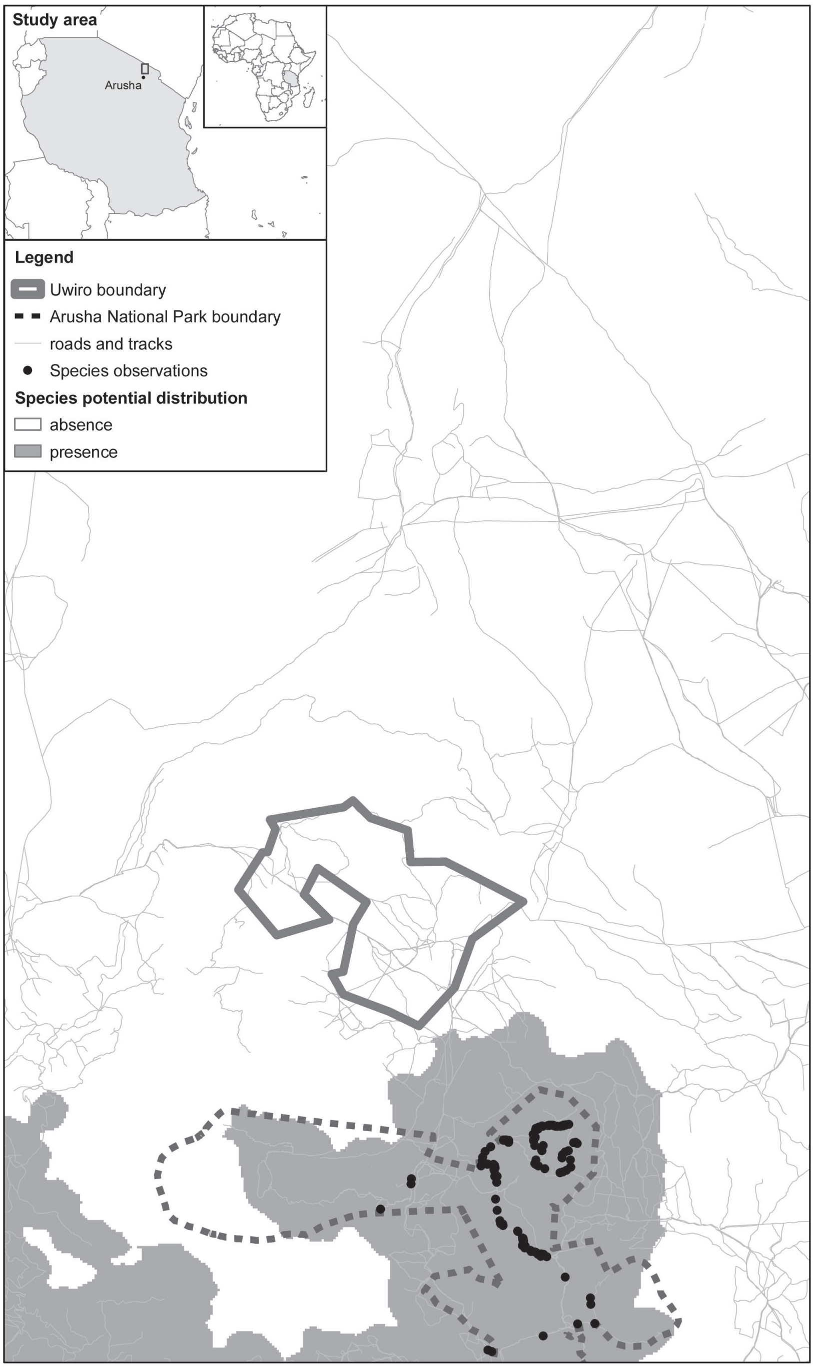


Figure A.52.: *Kobus ellipsiprymnus* (waterbuck) potential distribution

Loxodonta africana (african elephant)

Table A.28.: selected environmental variables and computed coefficients from stepwise analysis for *Loxodonta africana* (african elephant)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$-5.990e+02$	$8.610e+01$	-6.956	$3.49e-12$
ev-cl04	$3.320e-02$	$1.196e-02$	2.777	0.005492
ev-cl15	$8.738e-01$	$2.944e-01$	2.968	0.002999
ev-cl17	$-4.589e-01$	$1.622e-01$	-2.829	0.004673
ev-domi	$-8.355e+00$	$1.970e+00$	-4.241	$2.23e-05$
ev-patc	$7.220e+01$	$1.320e+01$	5.471	$4.47e-08$
ev-road	$-3.081e-03$	$5.283e-04$	-5.832	$5.48e-09$
ev-rvrs	$-1.114e-03$	$3.879e-04$	-2.872	0.004073
ev-swpa	$8.480e-04$	$2.710e-04$	3.129	0.001754
ev-vllg	$5.219e-04$	$2.084e-04$	2.504	0.012274
ev-tslp	$-2.497e-01$	$8.491e-02$	-2.941	0.003272
ev-h101	$1.460e-01$	$2.825e-02$	5.166	$2.39e-07$
ev-h104	$-2.107e-01$	$3.651e-02$	-5.771	$7.89e-094$
ev-h105	$-3.397e-01$	$5.024e-02$	-6.762	$1.36e-11$
ev-h106	$-2.977e-01$	$5.237e-02$	-5.685	$1.31e-08$
ev-h110	$9.595e-01$	$2.796e-01$	3.431	0.000600
ev-h113	$-8.743e-02$	$2.918e-02$	-2.996	0.002731
ev-h117	$-1.225e-01$	$3.871e-02$	-3.165	0.001549
ev-h122	$-2.138e-01$	$5.242e-02$	-4.079	$4.53e-05$
ev-h123	$1.686e-01$	$5.383e-02$	3.133	0.001733
ev-h127	$-2.353e-01$	$5.721e-02$	-4.112	$3.91e-05$
ev-h128	$-2.288e-01$	$3.690e-02$	-6.202	$5.57e-10$
ev-h132	$3.084e-01$	$5.530e-02$	5.577	$2.45e-08$
ev-h134	$-1.877e-01$	$3.927e-02$	-4.781	$1.74e-06$
ev-h140	$3.402e-01$	$5.839e-02$	5.827	$5.66e-09$
ev-h141	$-7.155e-01$	$2.720e-01$	-2.631	0.008518
ev-h142	$2.695e-01$	$5.279e-02$	5.105	$3.31e-07$
ev-h143	$4.714e-01$	$1.160e-01$	4.063	$4.84e-05$
ev-h145	$1.578e-01$	$5.424e-02$	2.910	0.003620
ev-h146	$2.387e-01$	$4.696e-02$	5.082	$3.73e-07$
ev-h162	$1.117e-01$	$4.882e-02$	2.288	0.022136
ev-h164	$-1.928e-01$	$7.810e-02$	-2.468	0.013577
ev-h165	$-1.861e-01$	$4.935e-02$	-3.771	0.000163
ev-h166	$2.935e-01$	$5.519e-02$	5.318	$1.05e-07$
ev-h167	$2.698e-01$	$9.404e-02$	2.869	0.004124
ev-h168	$-5.512e-01$	$9.826e-02$	-5.609	$2.03e-08$
ev-h174	$-1.984e-01$	$3.826e-02$	-5.186	$2.15e-07$
ev-h176	$2.583e-01$	$5.747e-02$	4.493	$7.01e-06$
ev-h177	$1.664e-01$	$7.396e-02$	2.249	0.024497
ev-h180	$-2.907e-01$	$4.720e-02$	-6.158	$7.34e-10$
ev-h182	$1.226e-01$	$5.087e-02$	2.410	0.015944
ev-h184	$4.895e-01$	$9.281e-02$	5.274	$1.34e-07$

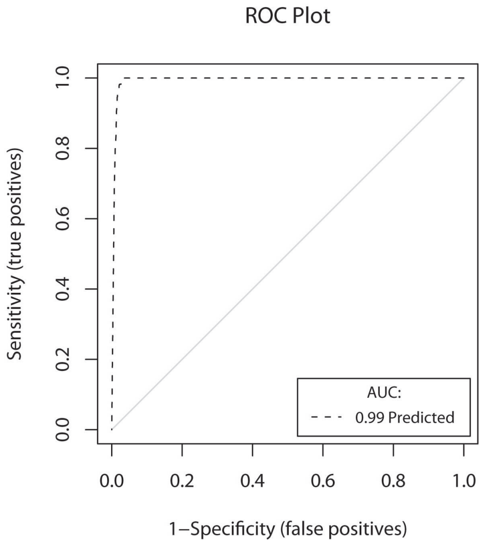


Figure A.53.: LRA for *Loxodonta africana* (african elephant) ROC plot

Table A.28.: (continued) selected environmental variables and computed coefficients from stepwise analysis for *Loxodonta africana* (african elephant)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
ev-h193	$6.066e-01$	$6.896e-02$	8.796	$< 2e-16$
ev-h202	-$1.125e-01$	$5.482e-02$	-2.053	0.040083
ev-h205	-$7.893e-01$	$9.557e-02$	-8.259	$< 2e-16$
ev-h207	-$3.238e-01$	$6.224e-02$	-5.202	$1.97e-07$
ev-h214	$1.182e+00$	$1.446e-01$	8.177	$2.91e-16$
ev-h216	$3.202e-01$	$9.737e-02$	3.289	0.001007
ev-h218	-$1.917e-01$	$9.845e-02$	-1.947	0.051563

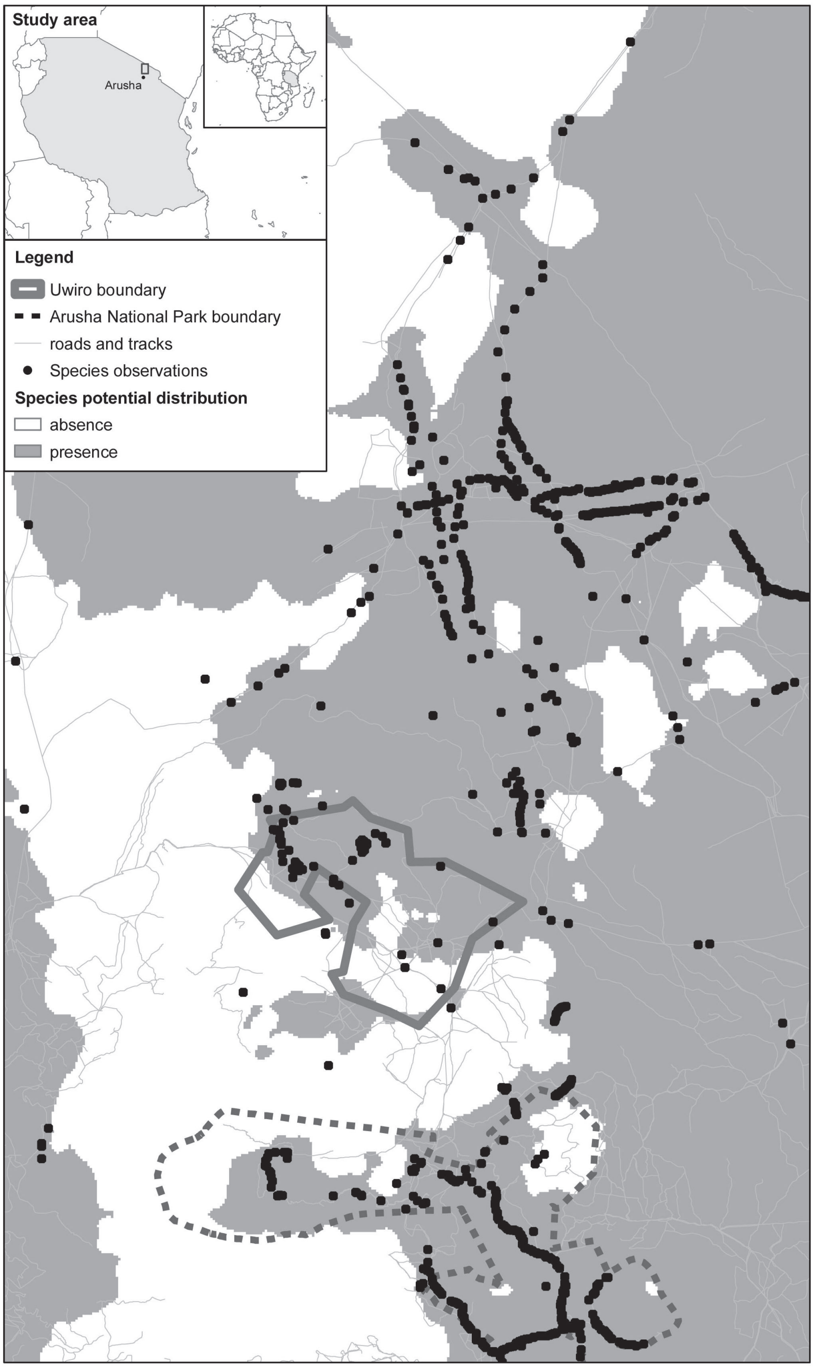


Figure A.54.: *Loxodonta africana* (african elephant) potential distribution

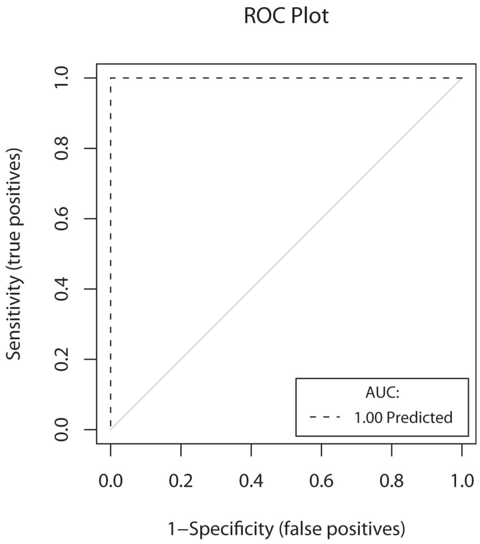


Figure A.55.: LRA for *Hystrix cristata* (crested porcupine) ROC plot

Hystrix cristata (crested porcupine)

Table A.29.: selected environmental variables and computed coefficients from stepwise analysis for *Hystrix cristata* (crested porcupine)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	$2.290e+04$	$1.422e+05$	0.161	0.872
ev-cl09	$-5.561e+01$	$3.435e+02$	-0.162	0.871
ev-vllg	$-1.042e-01$	$6.730e-01$	-0.155	0.877
ev-h108	$-2.129e+01$	$1.320e+02$	-0.161	0.872
ev-h115	$1.403e+01$	$8.930e+01$	0.157	0.875
ev-h132	$1.647e+01$	$1.018e+02$	0.162	0.871
ev-h156	$-3.621e+01$	$2.252e+02$	-0.161	0.872
ev-h180	$-3.424e+01$	$2.127e+02$	-0.161	0.872
ev-h182	$2.326e+01$	$1.457e+02$	0.160	0.873
ev-h216	$-2.801e+01$	$1.751e+02$	-0.160	0.873

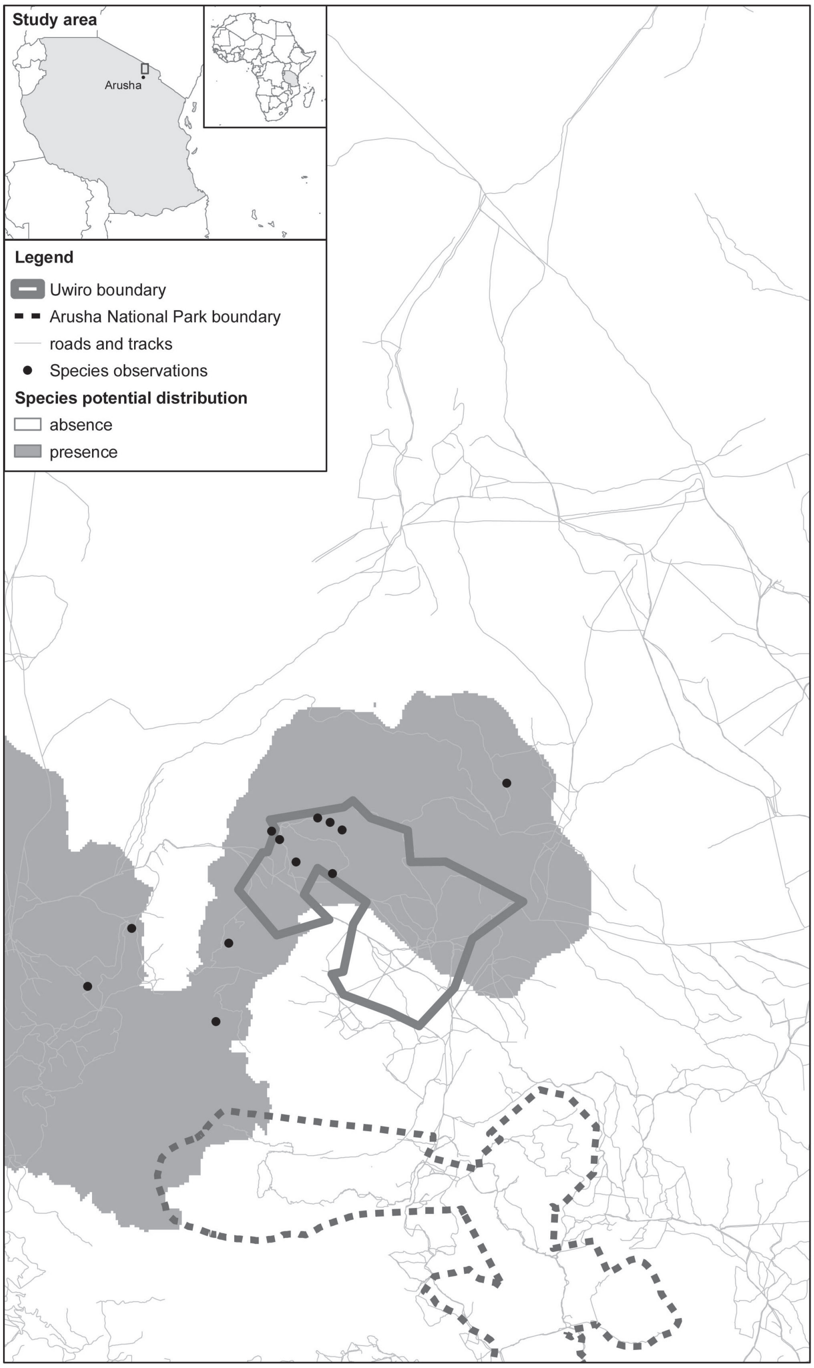


Figure A.56.: *Hystrix cristata* (crested porcupine) potential distribution

A.0.5. Lombardy region

Table A.30.: Environmental variables used for LRA in Lombardy case study

Variable name	Description	Source
EV-DHYD	distance from lakes and rivers	CTR
EV-DINF	weighted distance from roads and railroads	CTR
EV-D100	snowfields and glaciers	DUSAF
EV-D103	broad-leaved forest	DUSAF
EV-D104	riparian vegetation	DUSAF
EV-D105	coniferous forest	DUSAF
EV-D106	mixed forest	DUSAF
EV-D107	recent reforestation	DUSAF
EV-D108	orchards and vegetable gardens	DUSAF
EV-D109	tree crops	DUSAF
EV-D110	marshy vegetation	DUSAF
EV-D111	debris vegetation	DUSAF
EV-D112	riverbed vegetation	DUSAF
EV-D113	shrubs	DUSAF
EV-D114	meadows and pastures	DUSAF
EV-D115	herbaceous crops	DUSAF
EV-D116	herbaceous and tree crops	DUSAF
EV-D117	protected crops	DUSAF
EV-D119	debris and rock	DUSAF
EV-D120	dumps	DUSAF
EV-D121	continuous urban areas	DUSAF
EV-D122	open urban areas	DUSAF
EV-D123	farm building	DUSAF
EV-D126	airports	
EV-DTM	digital elevation model-elevation	CTR
EV-ASPECT	digital elevation model-aspect	CTR
EV-SLOPE	digital elevation model-slope	CTR
EV-CL15	Precipitation Seasonality(Coefficient of Variation)	
EV-CL18	Precipitation of warmest quarter	CTR
EV-CL01	Annual mean temperature	CTR
EV-CL07	Temperature annual range	CTR

Salamandra atra (alpine salamander)

Table A.31.: selected environmental variables and computed coefficients from stepwise analysis for *Salamandra atra* (alpine salamander)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>|z|$)**
Intercept	$1.138e+00$	$7.849e-01$	1.450	0.147074
EV-CL18	-3.167$e-03$	$1.282e-03$	-2.470	0.013522
EV-CL01	$6.527e-03$	$3.423e-03$	1.907	0.056564
EV-D100	-8.235$e-05$	$6.622e-06$	-12.436	$< 2e-16$
EV-D103	-8.216$e-05$	$2.160e-05$	-3.804	0.000142
EV-D104	$3.035e-05$	$1.327e-05$	2.287	0.022180
EV-D105	-7.195$e-05$	$2.199e-05$	-3.272	0.001069
EV-D106	-7.733$e-05$	$2.161e-05$	-3.579	0.000345
EV-D107	$3.153e-05$	$6.706e-06$	4.702	2.57e-06
EV-D108	$1.046e-04$	$9.967e-06$	10.492	$< 2e-16$
EV-D109	$3.255e-05$	$6.420e-06$	5.070	3.98e-07
EV-D110	$3.212e-05$	$7.567e-06$	4.245	2.19e-05
EV-D112	-6.590$e-05$	$1.213e-05$	-5.432	5.56e-08
EV-D117	$7.511e-06$	$4.072e-06$	1.844	0.065118
EV-D118	-7.202$e-05$	$6.273e-06$	-11.481	$< 2e-16$
EV-D119	-1.812$e-04$	$2.854e-05$	-6.351	2.14e-10
EV-D120	$3.944e-05$	$1.524e-05$	2.588	0.009658
EV-D122	-1.522$e-04$	$2.728e-05$	-5.579	2.42e-08
EV-D123	-1.348$e-04$	$1.079e-05$	-12.495	$< 2e-16$
EV-D126	-8.067$e-06$	$3.264e-06$	-2.472	0.013452
EV-DHYD	-7.430$e-05$	$4.265e-05$	-1.742	0.081473
EV-DINF	$1.669e-04$	$2.937e-05$	5.680	1.35e-08
EV-DTM	$5.550e-04$	$1.826e-04$	3.040	0.002367

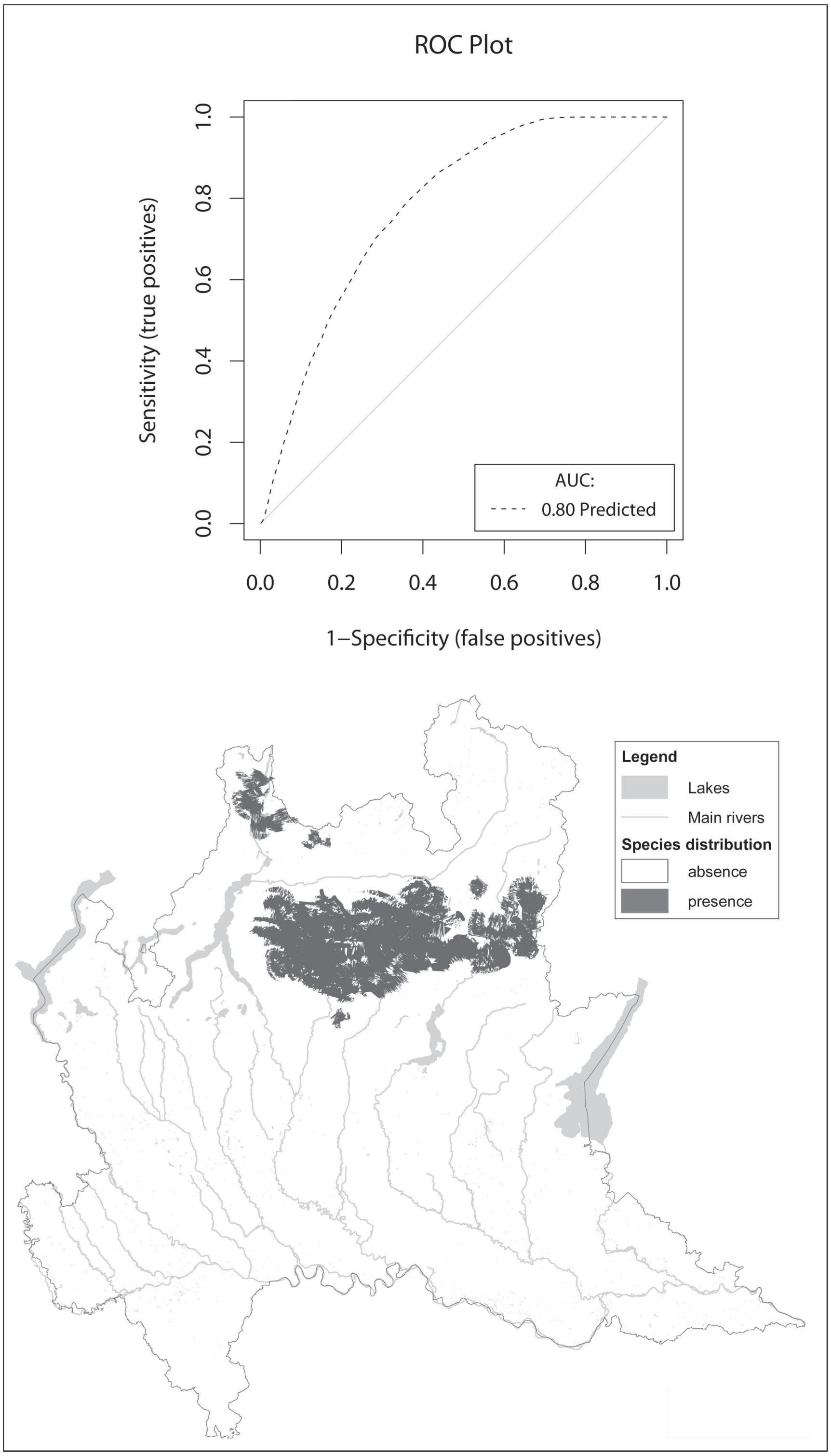


Figure A.57.: LRA for *Salamandra atra* (alpine salamander) ROC plot and potential distribution

Salamandra salamandra (fire salamander)

Table A.32.: selected environmental variables and computed coefficients from stepwise analysis for *Salamandra salamandra* (fire salamander)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	-5.909$e+00$	3.245$e-01$	-18.208	$< 2e-16$
EV-ASPECT	-4.718$e-04$	1.843$e-04$	-2.561	0.010446
EV-CL15	6.338$e-02$	6.565$e-03$	9.654	$< 2e-16$
EV-CL18	1.355$e-02$	6.530$e-04$	20.753	$< 2e-16$
EV-D100	6.878$e-06$	1.473$e-06$	4.670	3.01$e-06$
EV-D103	-7.056$e-05$	1.385$e-05$	-5.093	3.52e-07
EV-D105	3.612$e-05$	4.541$e-06$	7.956	1.78e-15
EV-D107	-1.701$e-05$	4.442$e-06$	-3.829	0.000129
EV-D110	-1.636$e-05$	5.050$e-06$	-3.240	0.001195
EV-D111	-3.377$e-05$	4.122$e-06$	-8.192	2.58e-16
EV-D112	-1.162$e-05$	7.542$e-06$	-1.541	0.123332
EV-D114	-1.276$e-04$	3.053$e-05$	-4.180	2.92e-05
EV-D115	-2.691$e-05$	9.601$e-06$	-2.803	0.005060
EV-D116	-2.322$e-05$	7.271$e-06$	-3.193	0.001408
EV-D117	2.833$e-05$	1.478$e-06$	19.162	$< 2e-16$
EV-D122	3.818$e-05$	2.701$e-05$	1.414	0.157420
EV-D123	3.421$e-05$	9.920$e-06$	3.448	0.000564
EV-D126	9.482$e-06$	1.316$e-06$	7.205	5.79e-13
EV-DHYD	-5.605$e-05$	1.652$e-05$	-3.393	0.000691
EV-DTM	-1.055$e-03$	8.915$e-05$	-11.829	$< 2e-16$

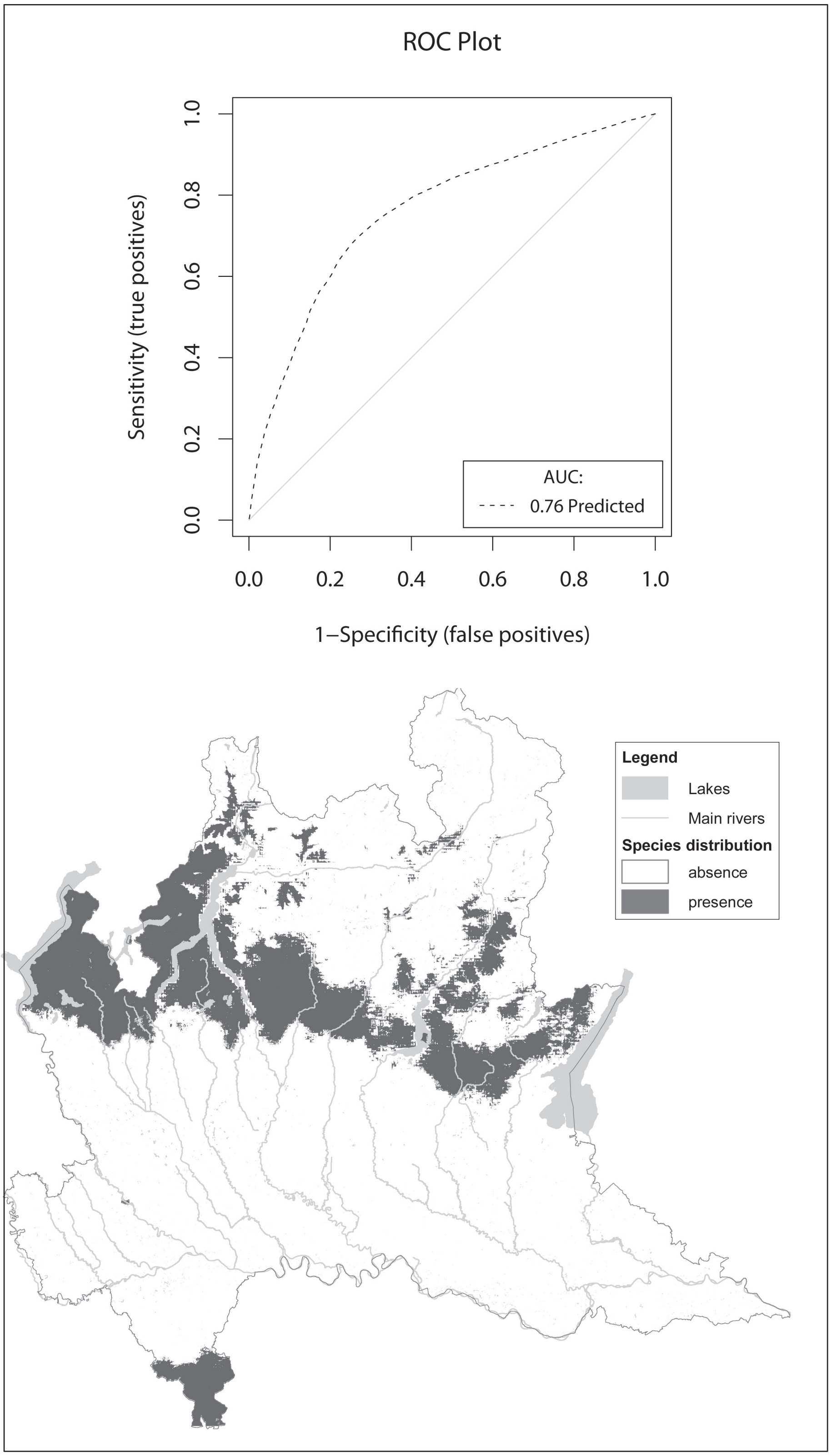


Figure A.58.: LRA for *Salamandra salamandra* (fire salamander) ROC plot and potential distribution

Salamandrina terdigitata (spectacled salamander)

Table A.33.: selected environmental variables and computed coefficients from stepwise analysis for *Salamandrina terdigitata* (spectacled salamander)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
Intercept	-5.909e+00	3.245e-01	-18.208	0.000439
EV-CL15	2.338e-02	3.575e-03	8.632	< 2e-16
EV-CL18	1.365e-03	5.530e-03	16.754	0.321332
EV-D100	4.278e-05	1.323e-06	32.632	2.21e-05
EV-D103	3.056e-05	2.385e-06	-6.021	3.23e-06
EV-D105	5.412e-04	3.431e-06	9.932	1.73e-15
EV-D108	-1.321e-07	5.442e-05	-2.549	0.000159
EV-D110	-0.766e-03	3.050e-05	-2.240	0.131195
EV-D111	-3.217e-05	2.342e-06	-9.134	2.58e-16
EV-D112	-3.162e-05	8.542e-05	-3.543	0.128332
EV-D113	-1.956e-04	4.563e-05	-5.240	6.32e-05
EV-D115	-1.691e-04	2.601e-07	-3.803	0.125060
EV-D116	-2.322e-05	6.221e-06	-4.123	0.001578
EV-D119	3.833e-04	2.443e-05	21.104	< 2e-16
EV-D121	3.818e-05	1.7651e-05	4.414	0.238420
EV-D123	3.421e-06	10.932e-06	5.486	0.001564
EV-D126	5.482e-06	1.323e-06	3.205	4.29e-13
EV-DHYD	-7.605e-02	5.612e-05	-3.243	0.100691
EV-DINF	1.645e-03	1.976e-06	6.320	1.29e-09
EV-DTM	-4.055e-03	8.921e-06	-11.839	< 2e-15

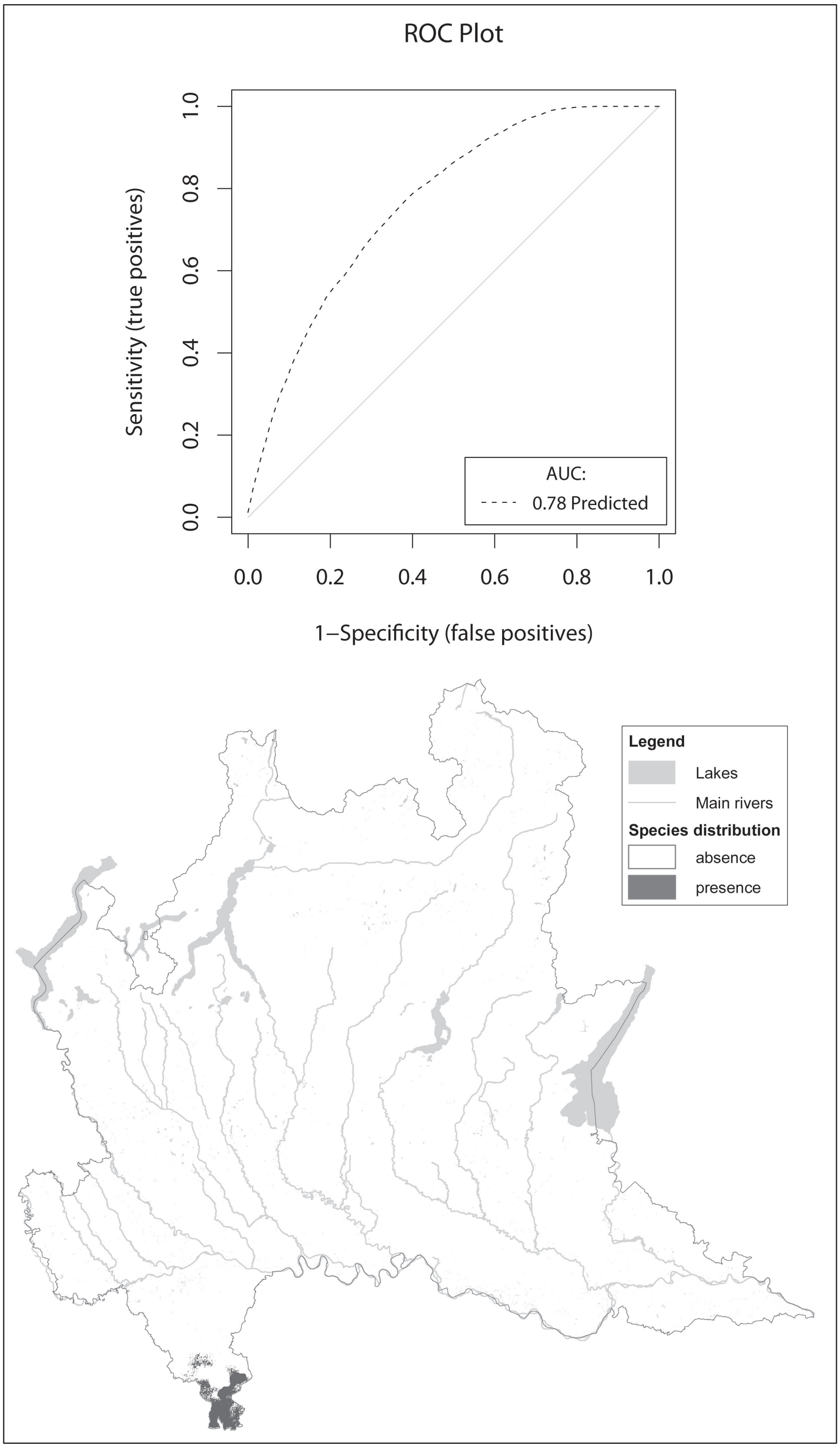


Figure A.59.: LRA for *Salamandrina terdigitata* (spectacled salamander) ROC plot and potential distribution

Triturus alpestris **(alpine newt)**

Table A.34.: selected environmental variables and computed coefficients from stepwise analysis for *Triturus alpestris* (alpine newt)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$-2.021e+01$	$1.79e+00$	-3.732	$3.53e-08$
EV-ASPECT	$2.423e-04$	$1.559e-04$	2.196	0.022872
EV-CL15	$4.237e-02$	$1.349e-02$	2.683	0.000431
EV-CL18	$6.121e-03$	$9.877e-04$	6.198	$5.73e-3$
EV-CL01	$2.256e-02$	$5.656e-03$	3.936	$3.28e-05$
EV-CL07	$1.327e-02$	$3.277e-03$	5.179	$2.34e-04$
EV-D100	$2.431e-05$	$2.082e-06$	12.673	$< 2e-16$
EV-D103	$-7.249e-05$	$1.981e-05$	-3.660	0.000252
EV-D105	$4.544e-05$	$4.434e-06$	11.127	$< 2e-16$
EV-D106	$-4.130e-05$	$3.923e-06$	-10.389	$< 2e-16$
EV-D107	$-3.445e-05$	$3.923e-06$	-9.927	$< 2e-16$
EV-D109	$-4.112e-05$	$9.543e-06$	-9.817	1.76e-06
EV-D111	$-1.608e-05$	$3.076e-06$	-5.227	1.54e-07
EV-D112	$-2.584e-05$	$7.075e-06$	-3.652	0.000230
EV-D114	$7.767e-06$	$2.297e-05$	3.447	0.000437
EV-D116	$-3.898e-05$	$8.826e-06$	-3.416	1.00e-05
EV-D117	$2.216e-05$	$2.803e-06$	12.287	$< 2e-12$
EV-D121	$-1.177e-04$	$4.034e-05$	-2.881	0.008968
EV-D123	$-7.857e-05$	$1.748e-05$	-4.517	$4.23e-02$
EV-D127	$-8.162e-05$	$2.443e-05$	-3.784	$1.71e-06$
EV-DYDR	$-2.839e-05$	$1.245e-05$	-2.279	0.034658
EV-DINF	$5.189e-04$	$5.422e-05$	5.881	$3.07e-08$
EV-DTM	$1.189e-03$	$3.145e-04$	3.212	0.000461

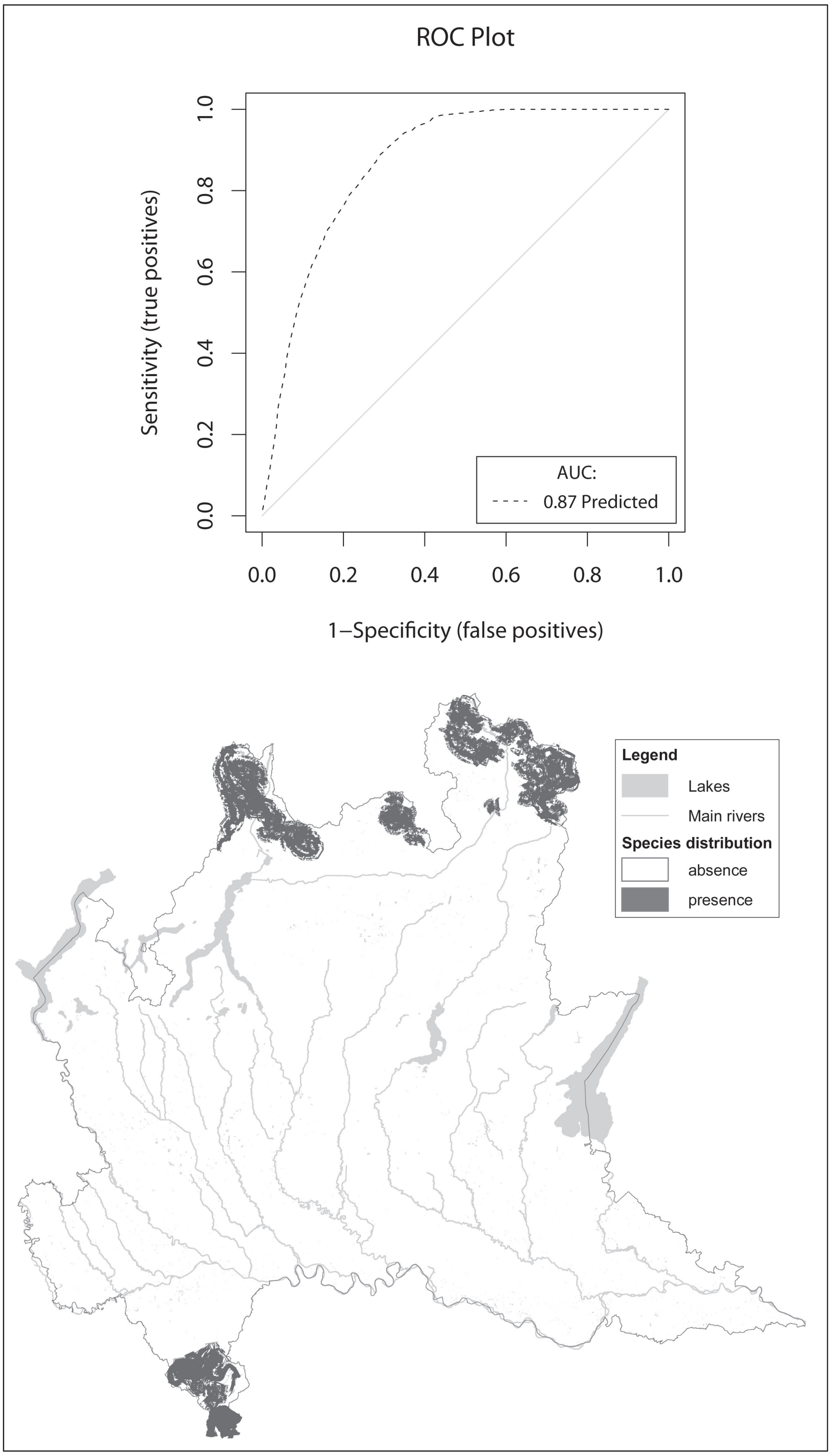


Figure A.60.: LRA for *Triturus alpestris* (alpine newt) ROC plot and potential distribution

Triturus carnifex (italian crested newt)

Table A.35.: selected environmental variables and computed coefficients from stepwise analysis for *Triturus carnifex* (italian crested newt)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$-2.021e+01$	$1.79e+00$	-3.732	$3.53e-08$
EV-ASPECT	$2.423e-04$	$1.559e-04$	2.196	0.022872
EV-CL15	$4.237e-02$	$1.349e-02$	2.683	0.000431
EV-CL18	$6.121e-03$	$9.877e-04$	6.198	$5.73e-3$
EV-CL01	$2.256e-02$	$5.656e-03$	3.936	$3.28e-05$
EV-CL07	$1.327e-02$	$3.277e-03$	5.179	$2.34e-04$
EV-D100	$2.431e-05$	$2.082e-06$	12.673	$< 2e-16$
EV-D105	$5.412e-04$	$3.431e-06$	9.932	$1.73e-15$
EV-D108	$-1.321e-07$	$5.442e-05$	-2.549	0.000159
EV-D110	$-0.766e-03$	$3.050e-05$	-2.240	0.131195
EV-D111	$-3.217e-05$	$2.342e-06$	-9.134	$2.58e-16$
EV-D112	$-3.162e-05$	$8.542e-05$	-3.543	0.128332
EV-D113	$-1.956e-04$	$4.563e-05$	-5.240	$6.32e-05$
EV-D114	$7.767e-06$	$2.297e-05$	3.447	0.000437
EV-D116	$-3.898e-05$	$8.826e-06$	-3.416	1.00e-05
EV-D117	$2.216e-05$	$2.803e-06$	12.287	$< 2e-12$
EV-D121	$-1.177e-04$	$4.034e-05$	-2.881	0.008968
EV-D123	$-7.857e-05$	$1.748e-05$	-4.517	$4.23e-02$
EV-D126	$-8.067e-06$	$3.264e-06$	-2.472	0.013452
EV-DHYD	$-7.430e-05$	$4.265e-05$	-1.742	0.081473
EV-DINF	$1.669e-04$	$2.937e-05$	5.680	1.35e-08
EV-DTM	$5.550e-04$	$1.826e-04$	3.040	0.002367

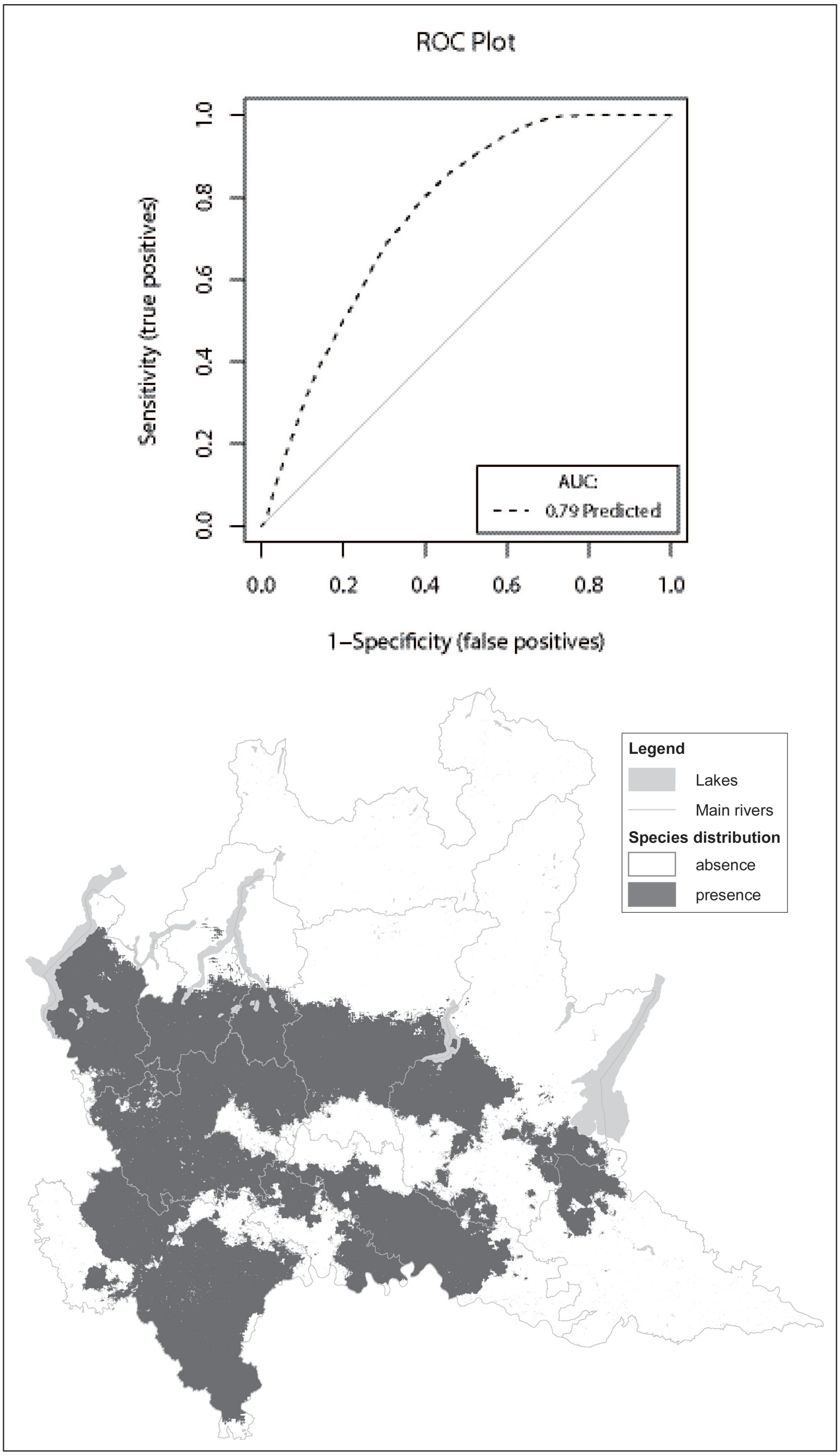


Figure A.61.: LRA for *Triturus carnifex* (italian crested newt) ROC plot and potential distribution

Triturus vulgaris (smooth newt)

Table A.36.: selected environmental variables and computed coefficients from stepwise analysis for *Triturus vulgaris* (smooth newt)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
Intercept	$-1.071e+01$	$1.79e+00$	-5.973	$2.33e-09$
EV-ASPECT	$3.423e-04$	$1.559e-04$	2.196	0.028092
EV-CL15	$4.967e-02$	$1.349e-02$	3.683	0.000231
EV-CL18	$6.121e-03$	$9.877e-04$	6.198	$5.73e-10$
EV-CL01	$2.226e-02$	$5.656e-03$	3.936	$8.28e-05$
EV-CL07	$1.787e-02$	$4.277e-03$	4.179	$2.93e-05$
EV-D100	$2.431e-05$	$2.082e-06$	11.673	$< 2e-16$
EV-D103	$-7.249e-05$	$1.981e-05$	-3.660	0.000252
EV-D105	$4.934e-05$	$4.434e-06$	11.127	$< 2e-16$
EV-D106	$-4.130e-05$	$3.975e-06$	-10.389	$< 2e-16$
EV-D107	$-3.895e-05$	$3.923e-06$	-9.927	$< 2e-16$
EV-D109	$-4.112e-05$	$8.536e-06$	-4.817	1.46e-06
EV-D111	$-1.608e-05$	$3.076e-06$	-5.227	1.72e-07
EV-D112	$-2.584e-05$	$7.075e-06$	-3.652	0.000260
EV-D114	$7.917e-05$	$2.297e-05$	3.447	0.000567
EV-D116	$-3.898e-05$	$8.826e-06$	-4.416	1.00e-05
EV-D117	$2.216e-05$	$1.803e-06$	12.287	$< 2e-16$
EV-D121	$-1.177e-04$	$4.084e-05$	-2.881	0.003968
EV-D123	$-7.897e-05$	$1.748e-05$	-4.517	$6.28e-06$
EV-D127	$-7.162e-05$	$1.497e-05$	-4.784	$1.72e-06$
EV-DYDR	$-2.839e-05$	$1.245e-05$	-2.279	0.022658
EV-DINF	$3.189e-04$	$5.422e-05$	5.881	$4.07e-09$
EV-DTM	$1.102e-03$	$3.145e-04$	3.502	0.000461

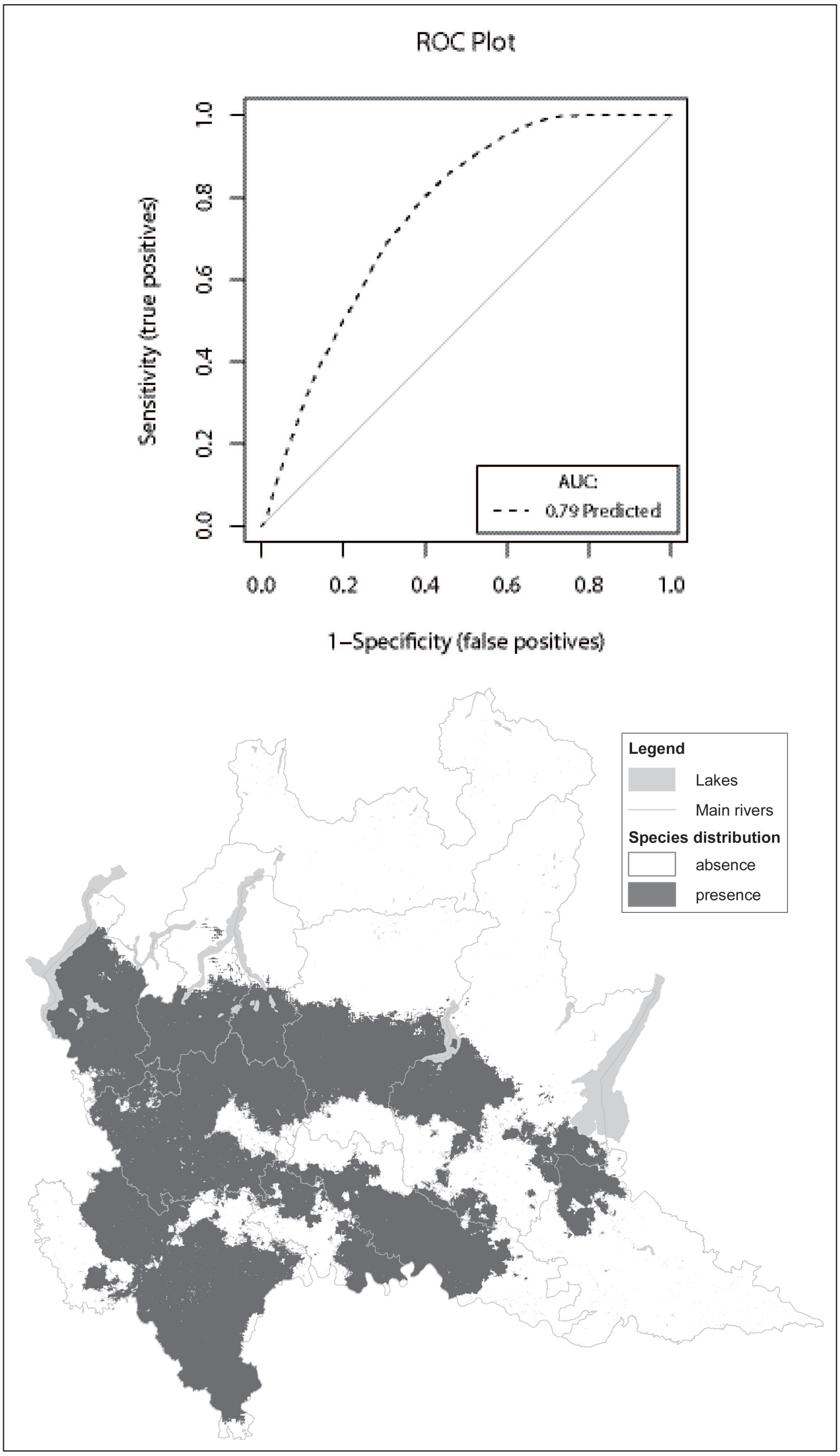


Figure A.62.: LRA for *Triturus vulgaris* (smooth newt) ROC plot and potential distribution

Speleomantes strinatii (Strinati's cave salamander)

Table A.37.: selected environmental variables and computed coefficients from stepwise analysis for *Speleomantes strinatii* (Strinati's cave salamander)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	-1.570$e+01$	3.871$e+00$	-4.056	4.99$e-05$
EV-CL07	4.552$e-02$	1.240$e-02$	3.670	0.000243
EV-D100	2.913$e-05$	5.226$e-06$	5.575	2.47$e-08$
EV-D104	1.146$e-04$	6.477$e-05$	1.769	0.076865
EV-D105	5.473$e-05$	3.816$e-05$	1.434	0.151488
EV-D111	-4.555$e-05$	3.001$e-05$	-1.518	0.129034
EV-D119	-8.022$e-05$	3.652$e-05$	-2.197	0.028047
EV-D121	3.369$e-04$	1.565$e-04$	2.152	0.031364
EV-D123	-1.215$e-04$	5.767$e-05$	-2.106	0.035183
EV-D126	3.378$e-05$	1.272$e-05$	2.655	0.007923
EV-DTM	1.583$e-02$	3.629$e-03$	4.361	1.29$e-05$

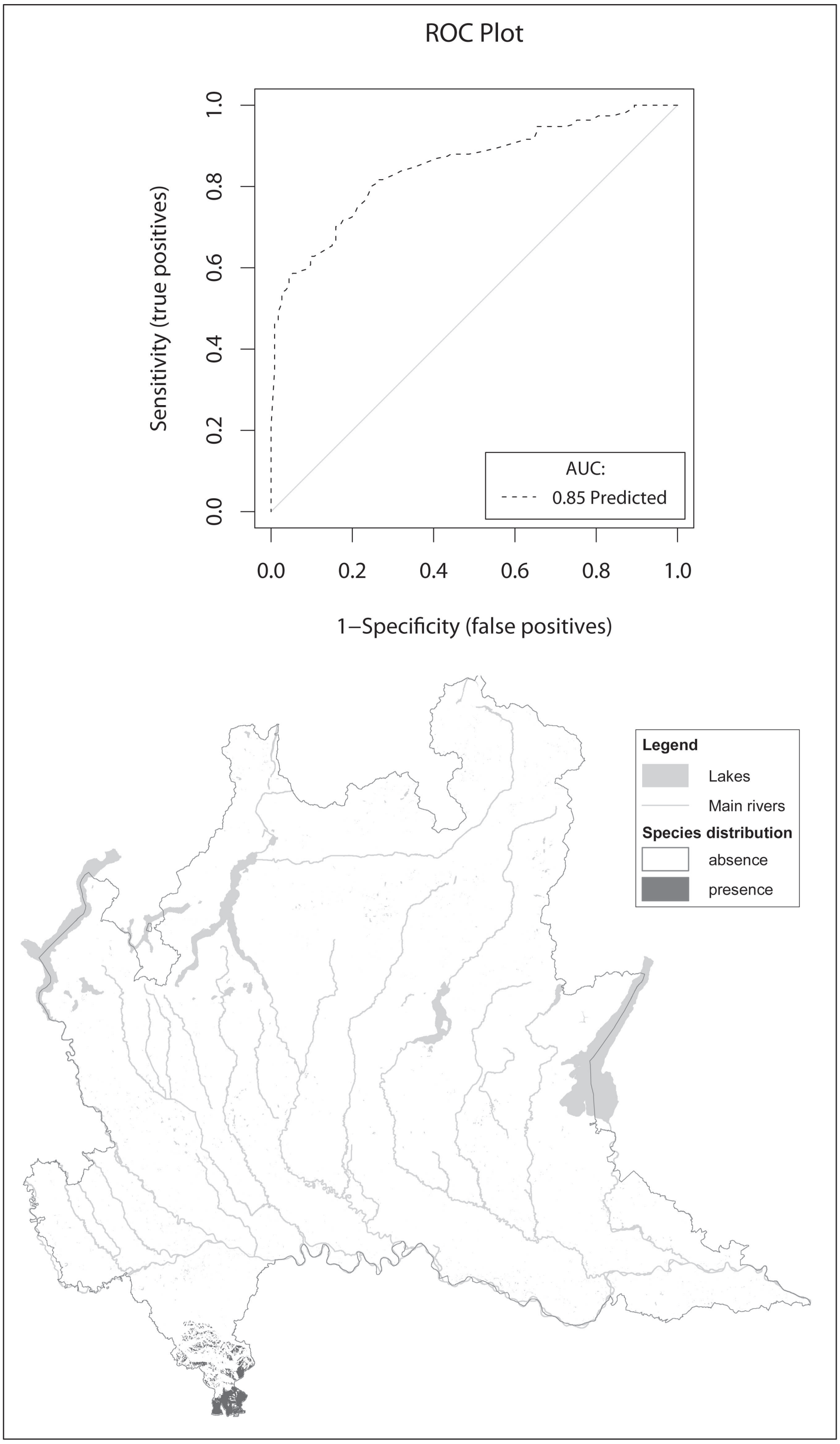


Figure A.63.: LRA for *Speleomantes strinatii* (Strinati's cave salamander) ROC plot and potential distribution

Bombina variegata (yellow bellied toad)

Table A.38.: selected environmental variables and computed coefficients from stepwise analysis for *Bombina variegata* (yellow bellied toad)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> $\lvert$z$\rvert$)**
Intercept	$-3.735e+00$	$1.587e+00$	-2.354	0.018568
EV-CL15	$4.534e-02$	$7.734e-03$	5.863	$4.54e-09$
EV-CL18	$5.049e-03$	$1.099e-03$	4.596	$4.32e-06$
EV-CL07	$6.680e-03$	$4.128e-03$	1.618	0.105599
EV-D100	$-1.277e-05$	$2.362e-06$	-5.408	$6.37e-08$
EV-D103	$-6.400e-05$	$1.519e-05$	-4.214	$2.51e-05$
EV-D105	$4.422e-05$	$7.608e-06$	5.812	$6.19e-09$
EV-D106	$2.193e-05$	$5.740e-06$	3.821	0.000133
EV-D107	$-1.873e-05$	$5.181e-06$	-3.615	0.000301
EV-D108	$-4.192e-05$	$1.216e-05$	-3.447	0.000566
EV-D109	$-2.178e-05$	$6.271e-06$	-3.474	0.000514
EV-D110	$8.571e-05$	$5.842e-06$	14.670	$< 2e-16$
EV-D111	$-1.515e-05$	$5.127e-06$	-2.956	0.003121
EV-D112	$-1.803e-05$	$8.728e-06$	-2.066	0.038789
EV-D114	$-8.930e-05$	$3.444e-05$	-2.593	0.009518
EV-D115	$3.608e-05$	$1.050e-05$	3.438	0.000586
EV-D116	$1.057e-04$	$9.273e-06$	11.397	$< 2e-16$
EV-D117	$7.780e-06$	$1.815e-06$	4.287	$1.81e-05$
EV-D119	$-3.537e-05$	$6.394e-06$	-5.531	$3.18e-08$
EV-D120	$-4.819e-05$	$1.755e-05$	-2.746	0.006028
EV-D121	$5.902e-05$	$2.943e-05$	2.005	0.044929
EV-D124	$6.290e-05$	$1.008e-05$	6.237	$4.46e-10$
EV-D126	$-3.903e-06$	$1.712e-06$	-2.280	0.022623
EV-DHYD	$-7.090e-05$	$1.912e-05$	-3.709	0.000208
EV-DTM	$-5.903e-04$	$1.703e-04$	-3.466	0.000528

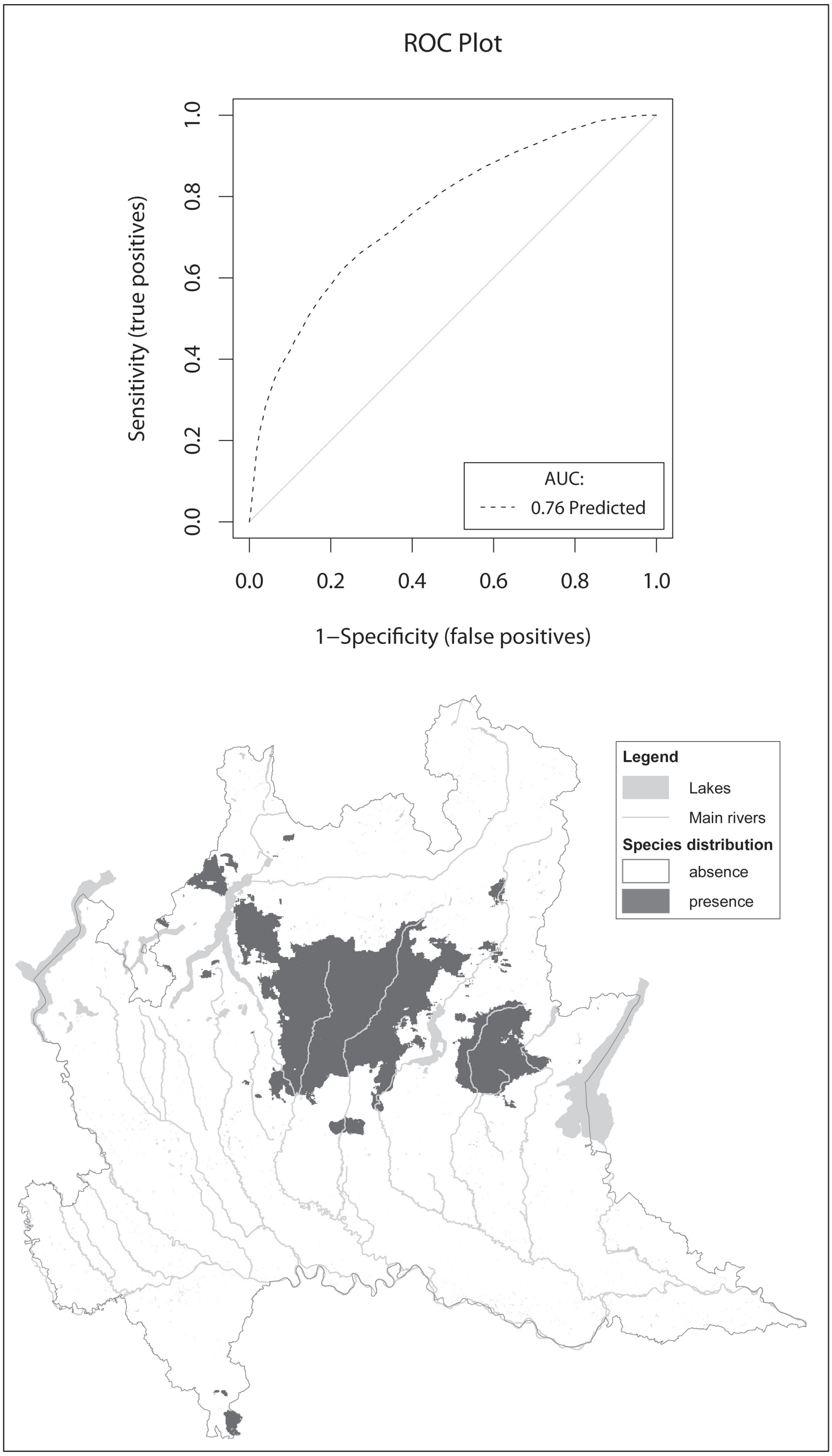


Figure A.64.: LRA for *Bombina variegata* (yellow bellied toad) ROC plot and potential distribution

***Pelobates fuscus* (spadefoot toad)**

Table A.39.: selected environmental variables and computed coefficients from stepwise analysis for *Pelobates fuscus* (spadefoot toad)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept-8.58$e+00$	1.291$e+00$	-6.647	2.99$e-11$	
EV-CL15	4.880$e-02$	6.927$e-03$	7.045	1.86$e-12$
EV-CL18	7.926$e-03$	8.861$e-04$	8.945	$< 2e-16$
EV-CL01	-1.506$e-02$	3.950$e-03$	-3.813	0.000137
EV-CL07	2.266$e-02$	3.879$e-03$	5.843	5.12$e-09$
EV-D100	1.694$e-05$	1.632$e-06$	10.380	$< 2e-16$
EV-D103	-2.138$e-04$	1.567$e-05$	-13.648	$< 2e-16$
EV-D104	5.040$e-05$	1.018$e-05$	4.952	7.34$e-07$
EV-D105	1.498$e-05$	4.237$e-06$	3.536	0.000406
EV-D107	-1.673$e-05$	3.437$e-06$	-4.867	1.13$e-06$
EV-D108	2.904$e-05$	9.277$e-06$	3.131	0.001745
EV-D109	1.972$e-05$	4.973$e-06$	3.965	7.32$e-05$
EV-D110	-2.925$e-05$	4.684$e-06$	-6.245	4.24$e-10$
EV-D111	-9.023$e-06$	3.332$e-06$	-2.708	0.006763
EV-D112	1.773$e-05$	6.443$e-06$	2.752	0.005924
EV-D113	1.008$e-04$	2.704$e-05$	3.726	0.000194
EV-D114	-7.069$e-05$	2.529$e-05$	-2.796	0.005178
EV-D115	-1.208$e-04$	9.994$e-06$	-12.086	$< 2e-16$
EV-D116	2.254$e-05$	7.319$e-06$	3.079	0.002075
EV-D117	1.353$e-05$	1.299$e-06$	10.415	$< 2e-16$
EV-D119	-1.821$e-05$	4.311$e-06$	-4.223	2.41$e-05$
EV-D121	-4.047$e-05$	2.537$e-05$	-1.595	0.110652
EV-D126	-5.635$e-06$	1.388$e-06$	-4.060	4.90$e-05$
EV-DTM	-3.350$e-04$	1.849$e-04$	-1.812	0.069981

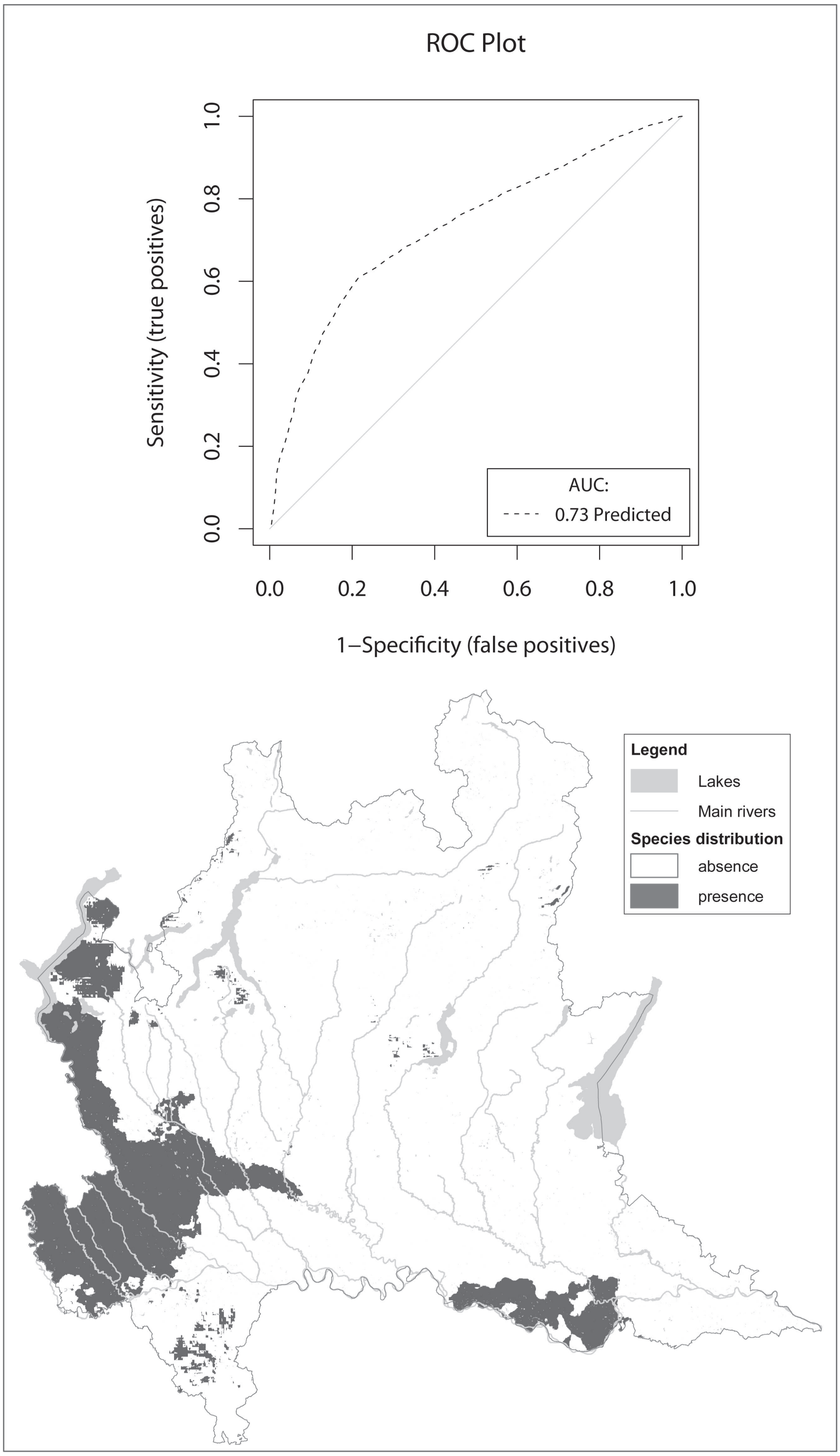


Figure A.65.: LRA for *Pelobates fuscus* (spadefoot toad) ROC plot and potential distribution

Bufo bufo (common toad)

Table A.40.: selected environmental variables and computed coefficients from stepwise analysis for *Bufo bufo* (common toad)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	-8.58e+00	1.291e+00	-6.647	2.99e-11
EV-CL15	4.880e-02	6.927e-03	7.045	1.86e-12
EV-CL18	7.926e-03	8.861e-04	8.945	< 2e-16
EV-CL01	-1.506e-02	3.950e-03	-3.813	0.000137
EV-CL07	2.266e-02	3.879e-03	5.843	5.12e-09
EV-D100	1.694e-05	1.632e-06	10.380	< 2e-16
EV-D103	-2.138e-04	1.567e-05	-13.648	< 2e-16
EV-D104	5.040e-05	1.018e-05	4.952	7.34e-07
EV-D105	1.498e-05	4.237e-06	3.536	0.000406
EV-D107	-1.673e-05	3.437e-06	-4.867	1.13e-06
EV-D108	2.904e-05	9.277e-06	3.131	0.001745
EV-D109	1.972e-05	4.973e-06	3.965	7.32e-05
EV-D110	-2.925e-05	4.684e-06	-6.245	4.24e-10
EV-D111	-9.023e-06	3.332e-06	-2.708	0.006763
EV-D112	1.773e-05	6.443e-06	2.752	0.005924
EV-D113	1.008e-04	2.704e-05	3.726	0.000194
EV-D114	-7.069e-05	2.529e-05	-2.796	0.005178
EV-D115	-1.208e-04	9.994e-06	-12.086	< 2e-16
EV-D116	2.254e-05	7.319e-06	3.079	0.002075
EV-D117	1.353e-05	1.299e-06	10.415	< 2e-16
EV-D119	-1.821e-05	4.311e-06	-4.223	2.41e-05
EV-D121	-4.047e-05	2.537e-05	-1.595	0.110652
EV-D126	-5.635e-06	1.388e-06	-4.060	4.90e-05
EV-DTM	-3.350e-04	1.849e-04	-1.812	0.069981

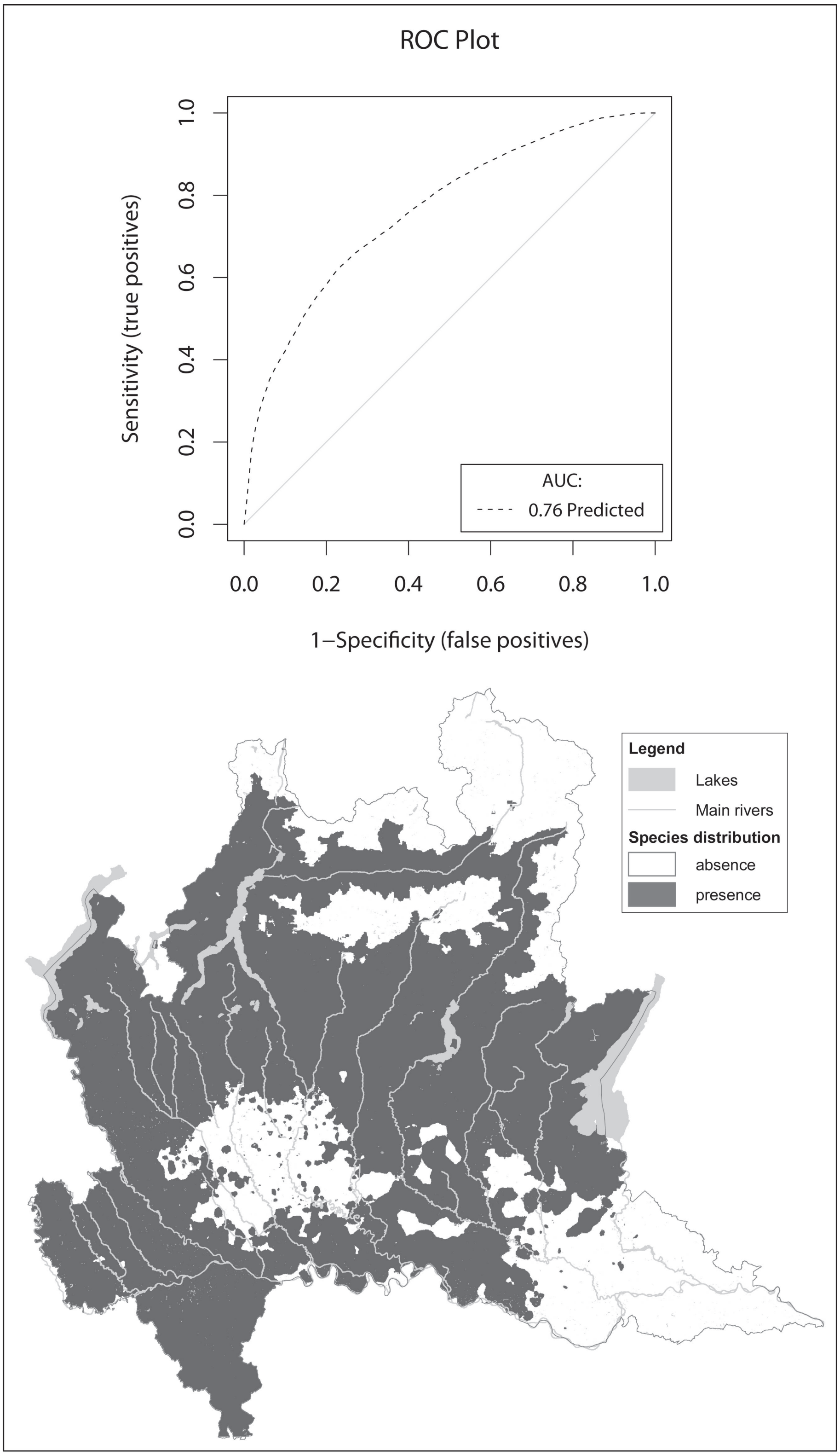


Figure A.66.: LRA for *Bufo bufo* (common toad) ROC plot and potential distribution

Bufo viridis (green toad)

Table A.41.: selected environmental variables and computed coefficients from stepwise analysis for *Bufo viridis* (green toad)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$7.283e+01$	$1.505e+01$	4.841	$1.29e-06$
EV-ASPECT	$2.246e-04$	$1.537e-04$	1.461	0.143958
EV-CL15	$7.567e-02$	$1.390e-02$	5.444	$5.20e-08$
EV-CL01	$4.983e-02$	$5.784e-03$	8.616	$< 2e-16$
EV-CL07	$-1.673e-02$	$3.354e-03$	-4.988	$6.11e-07$
EV-D100	$-6.572e-06$	$3.435e-06$	-1.913	0.055701
EV-D103	$-3.536e-05$	$1.620e-05$	-2.183	0.029060
EV-D107	$-3.807e-05$	$3.013e-06$	12.632	$< 2e-16$
EV-D108	$-7.609e-05$	$1.137e-05$	-6.694	$2.17e-11$
EV-D109	$1.013e-04$	$8.907e-06$	11.374	$< 2e-16$
EV-D110	$-1.497e-05$	$5.479e-06$	-2.732	0.006294
EV-D114	$-5.079e-05$	$2.176e-05$	-2.334	0.019583
EV-D115	$-1.204e-04$	$2.102e-05$	-5.730	$1.01e-08$
EV-D116	$2.160e-05$	$9.496e-06$	2.274	0.022957
EV-D119	$2.319e-05$	$1.959e-06$	11.840	$< 2e-16$
EV-D122	$1.292e-04$	$3.561e-05$	3.629	0.000284
EV-D123	$-7.906e-05$	$1.808e-05$	-4.373	$1.22e-05$
EV-D124	$-6.714e-05$	$9.959e-06$	-6.742	$1.56e-11$
EV-D125	$-3.348e-06$	$2.066e-06$	-1.621	0.105074
EV-D126	$-1.300e-05$	$1.199e-06$	10.846	$< 2e-16$
EV-D127	$-6.808e-05$	$1.553e-05$	-4.385	$1.16e-05$
EV-DHYD	$-7.792e-05$	$1.155e-05$	-6.748	$1.50e-11$
EV-DTM	$1.309e-03$	$3.636e-04$	3.601	0.000318

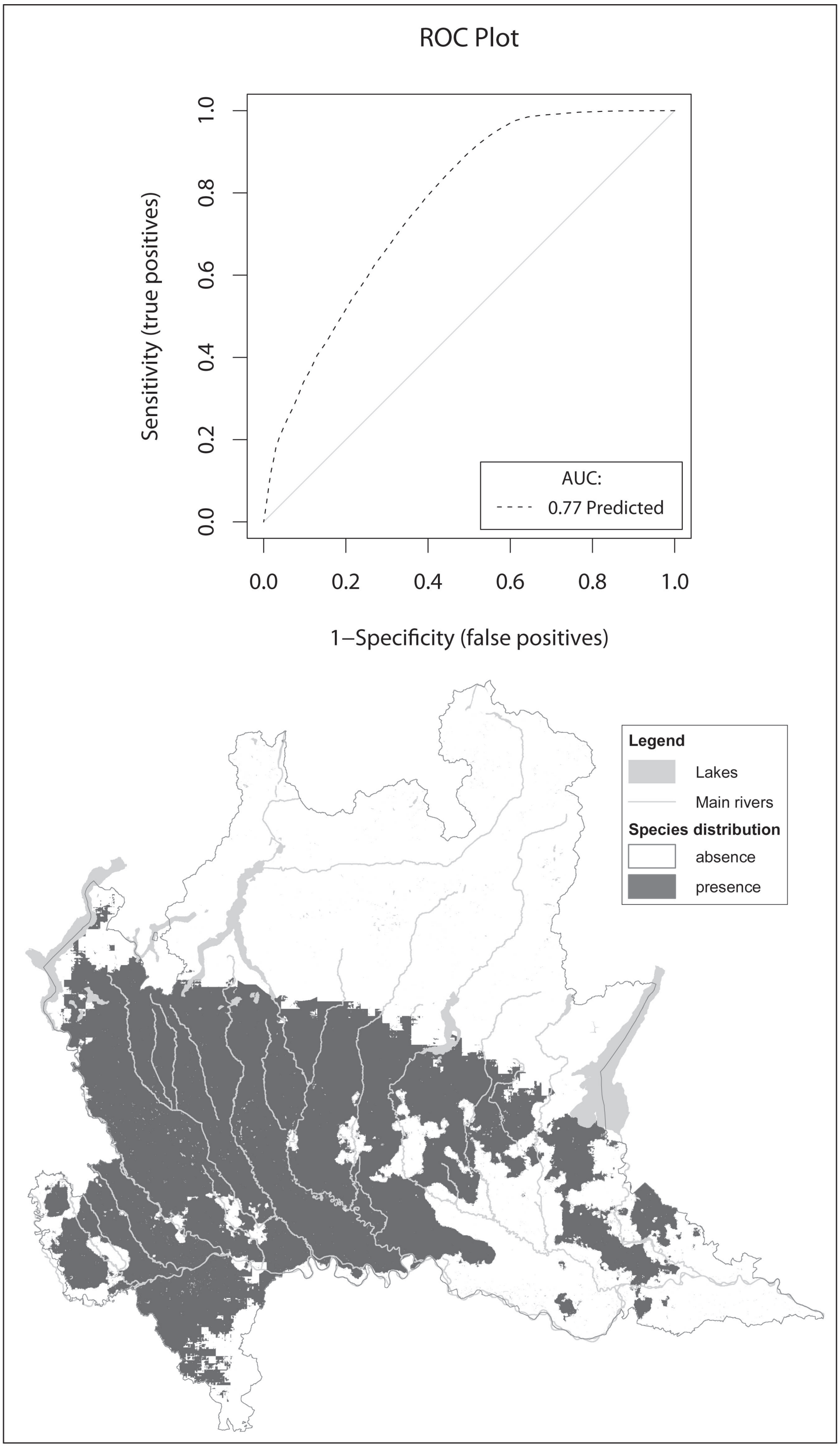


Figure A.67.: LRA for *Bufo viridis* (green toad) ROC plot and potential distribution

Hyla intermedia (italian tree frog)

Table A.42.: selected environmental variables and computed coefficients from stepwise analysis for *Hyla intermedia* (italian tree frog)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$-8.992e+00$	$1.447e+00$	-6.216	$5.10e-10$
EV-ASPECT	$2.148e-04$	$1.479e-04$	1.452	0.146415
EV-CL15	$2.036e-02$	$9.159e-03$	2.223	0.026204
EV-CL18	$1.328e-02$	$9.856e-04$	13.478	$< 2e-16$
EV-CL01	$-6.071e-03$	$3.257e-03$	-1.864	0.062288
EV-CL07	$2.115e-02$	$4.588e-03$	4.609	$4.05e-06$
EV-D100	$1.293e-05$	$1.798e-06$	7.193	$6.33e-13$
EV-D103	$-1.273e-04$	$1.533e-05$	-8.300	$< 2e-16$
EV-D105	$3.896e-05$	$5.009e-06$	7.778	$7.40e-15$
EV-D106	$1.019e-05$	$3.613e-06$	2.822	0.004777
EV-D107	$8.350e-06$	$3.346e-06$	2.496	0.012567
EV-D108	$-3.722e-05$	$1.061e-05$	-3.507	0.000454
EV-D109	$-1.627e-05$	$7.034e-06$	-2.314	0.020689
EV-D110	$-3.036e-05$	$5.146e-06$	-5.900	$3.63e-09$
EV-D111	$-2.137e-05$	$2.972e-06$	-7.189	$6.51e-13$
EV-D115	$-5.861e-05$	$1.487e-05$	-3.941	$8.11e-05$
EV-D117	$-3.970e-06$	$1.507e-06$	-2.635	0.008425
EV-D118	$1.356e-05$	$3.566e-06$	3.803	0.000143
EV-D119	$-8.738e-06$	$4.498e-06$	-1.943	0.052044
EV-D120	$-8.530e-05$	$1.803e-05$	-4.730	$2.25e-06$
EV-D122	$-5.787e-05$	$3.572e-05$	-1.620	0.105215
EV-D123	$-1.195e-04$	$1.575e-05$	-7.589	$3.23e-14$
EV-D126	$-2.821e-06$	$1.166e-06$	-2.419	0.015555
EV-DHYD	$-2.852e-05$	$1.150e-05$	-2.480	0.013144
EV-DINF	$2.424e-04$	$4.092e-05$	5.923	$3.16e-09$

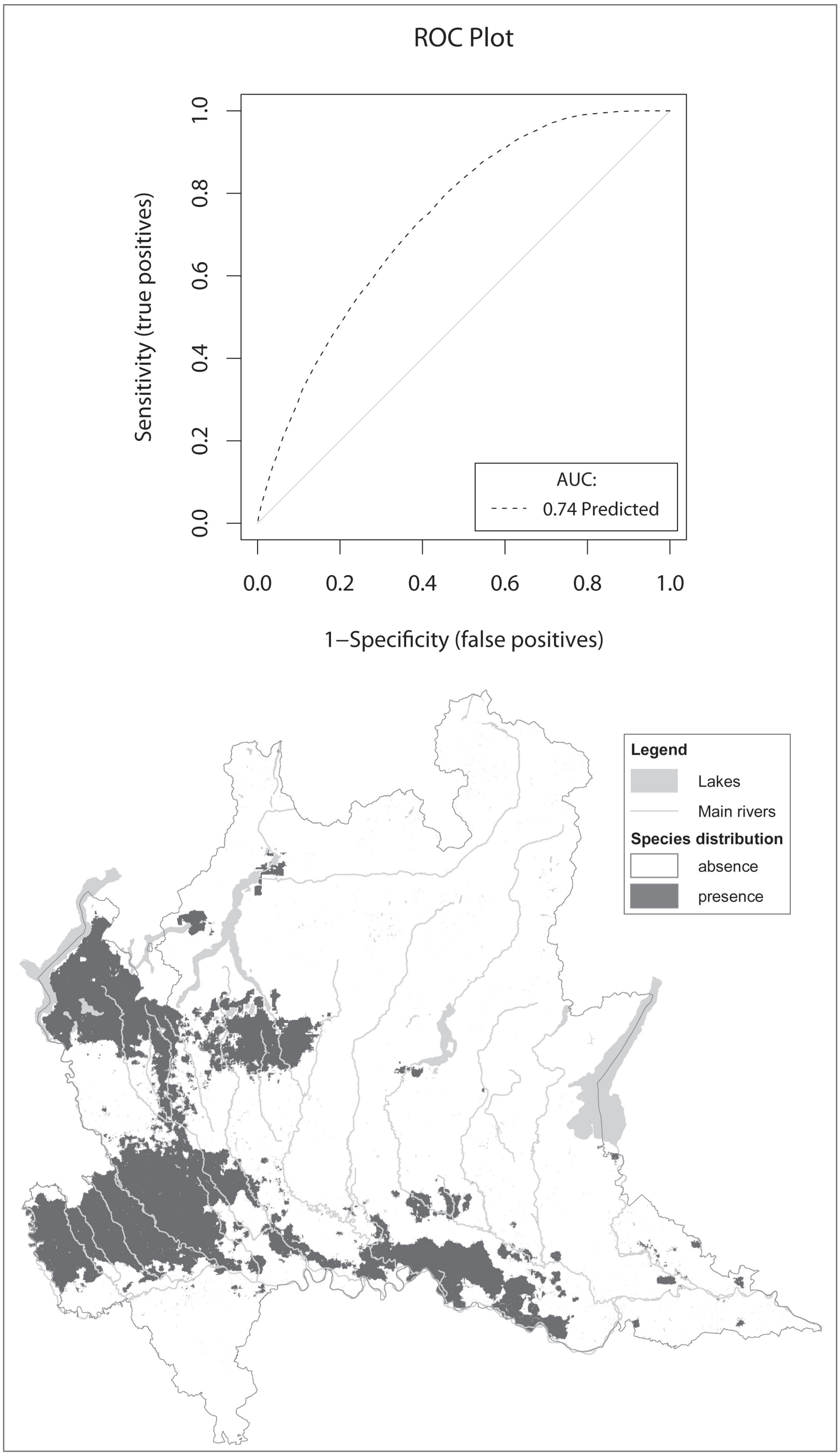


Figure A.68.: LRA for *Hyla intermedia* (italian tree frog) ROC plot and potential distribution

Rana dalmatina (agile frog)

Table A.43.: selected environmental variables and computed coefficients from stepwise analysis for *Rana dalmatina* (agile frog)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	$\mathbf{Pr}(>\lvert\mathbf{z}\rvert)$
Intercept	$2.899e+02$	$9.920e+00$	29.221	$< 2e-16$
EV-CL18	$2.664e-02$	$1.063e-03$	25.068	$< 2e-16$
EV-CL01	$1.544e-02$	$4.717e-03$	3.274	0.001059
EV-D103	$-2.179e-04$	$2.373e-05$	-9.183	$< 2e-16$
EV-D104	$4.360e-05$	$1.157e-05$	3.769	0.000164
EV-D105	$-3.697e-05$	$4.595e-06$	-8.045	$8.65e-16$
EV-D106	$-2.942e-05$	$4.571e-06$	-6.436	$1.23e-10$
EV-D107	$-1.189e-05$	$4.288e-06$	-2.772	0.005580
EV-D108	$3.245e-05$	$1.230e-05$	2.638	0.008341
EV-D110	$-2.913e-05$	$5.395e-06$	-5.400	$6.68e-08$
EV-D111	$-5.423e-05$	$3.755e-06$	14.443	$< 2e-16$
EV-D112	$4.724e-05$	$6.606e-06$	7.152	$8.56e-13$
EV-D113	$-4.423e-05$	$3.043e-05$	-1.454	0.146017
EV-D114	$-2.780e-04$	$3.195e-05$	-8.700	$< 2e-16$
EV-D115	$6.052e-05$	$1.348e-05$	4.489	$7.15e-06$
EV-D116	$2.323e-05$	$9.417e-06$	2.467	0.013637
EV-D117	$2.596e-05$	$1.606e-06$	16.169	$< 2e-16$
EV-D119	$6.697e-05$	$3.953e-06$	16.940	$< 2e-16$
EV-D120	$-1.141e-04$	$1.931e-05$	-5.906	$3.52e-09$
EV-D121	$8.949e-05$	$4.540e-05$	1.971	0.048682
EV-D122	$-7.875e-05$	$4.638e-05$	-1.698	0.089497
EV-D123	$-6.844e-05$	$1.337e-05$	-5.119	$3.07e-07$
EV-D124	$-2.898e-05$	$8.721e-06$	-3.323	0.000892
EV-D126	$1.572e-05$	$1.540e-06$	10.209	$< 2e-16$
EV-DHYD	$-4.871e-05$	$1.442e-05$	-3.376	0.000734
EV-DINF	$1.796e-04$	$4.968e-05$	3.616	0.000300
EV-DTM	$-4.437e-04$	$2.441e-04$	-1.818	0.069061

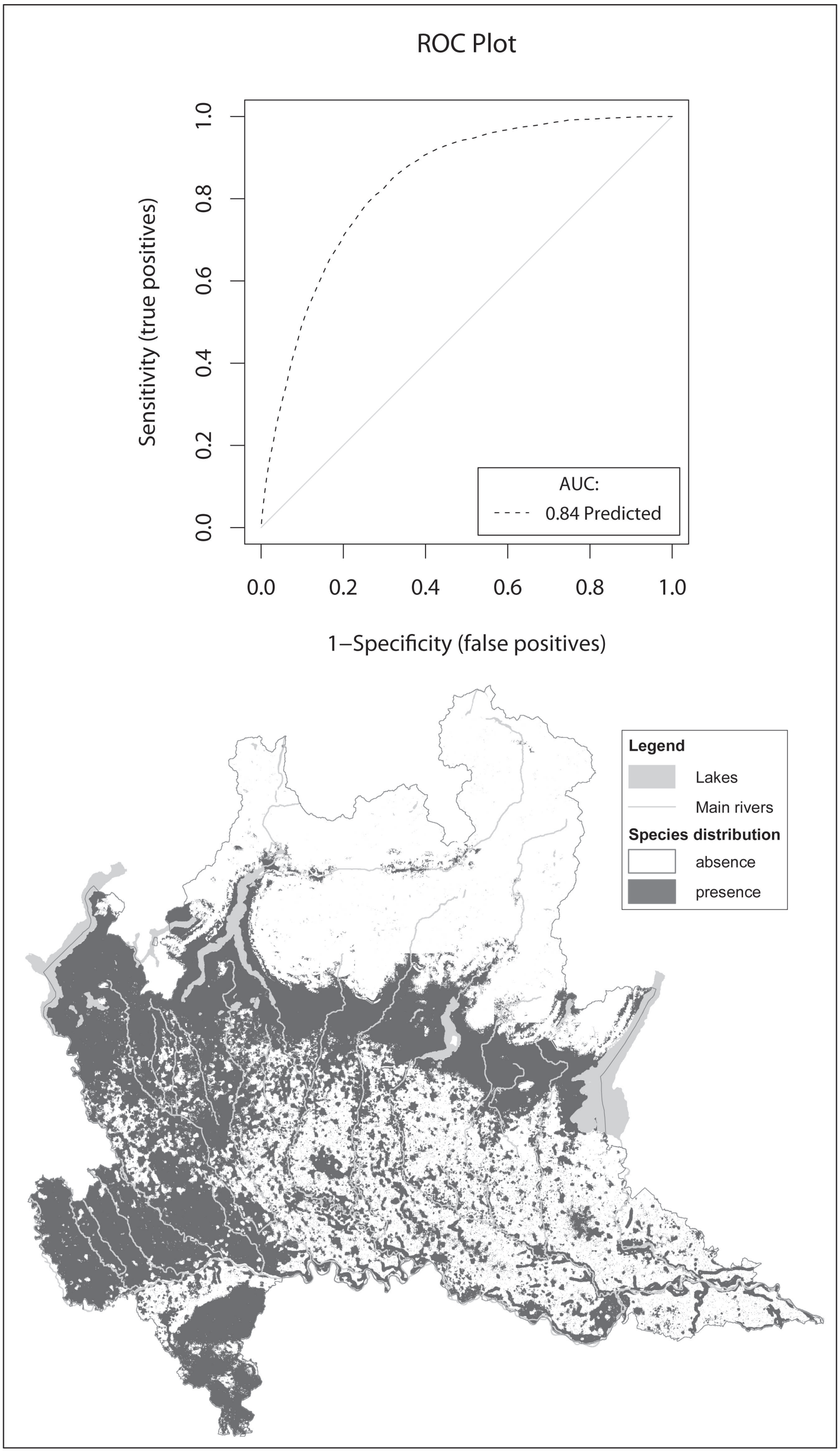


Figure A.69.: LRA for *Rana dalmatina* (agile frog) ROC plot and potential distribution

Rana italica (italian stream frog)

Table A.44.: selected environmental variables and computed coefficients from stepwise analysis for *Rana italica* (italian stream frog)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$3.892e+05$	$10.320e+01$	43.231	0.032029
EV-CL15	$1.634e-03$	$1.064e-04$	22.063	$< 2e-10$
EV-CL01	$0.744e-05$	$4.562e-03$	3.274	0.001069
EV-D103	$-5.199e-05$	$2.213e-06$	-10.146	$< 2e-08$
EV-D105	$-2.117e-04$	$5.321e-06$	-7.031	$8.62e-16$
EV-D106	$-2.452e-05$	$2.591e-06$	-4.472	$1.21e-09$
EV-D107	$-0.189e-04$	$3.299e-06$	-2.734	0.003281
EV-D108	$2.244e-04$	$1.330e-05$	1.639	0.009541
EV-D112	$6.324e-04$	$6.756e-06$	8.322	$9.26e-10$
EV-D113	$-8.432e-04$	$3.052e-05$	-1.824	0.146017
EV-D114	$-3.780e-04$	$6.215e-05$	-9.320	$< 2e-16$
EV-D115	$4.021e-05$	$1.343e-06$	3.491	$7.10e-06$
EV-D116	$2.431e-05$	$10.417e-06$	5.467	0.012397
EV-D117	$2.666e-05$	$1.606e-06$	16.169	$< 2e-10$
EV-D122	$-9.325e-04$	$4.868e-03$	-9.327	0.089670
EV-D123	$-7.873e-04$	$0.347e-05$	-3.132	$3.12e-08$
EV-D124	$-2.898e-05$	$8.721e-06$	-3.323	0.000892
EV-D126	$0.902e-04$	$1.546e-05$	09.232	$< 2e-10$
EV-DHYD	$-9.531e-05$	$0.922e-05$	-4.426	0.000823
EV-DINF	$1.926e-04$	$5.934e-04$	2.624	0.000323
EV-DTM	$-5.432e-05$	$2.381e-04$	-0.832	0.167490

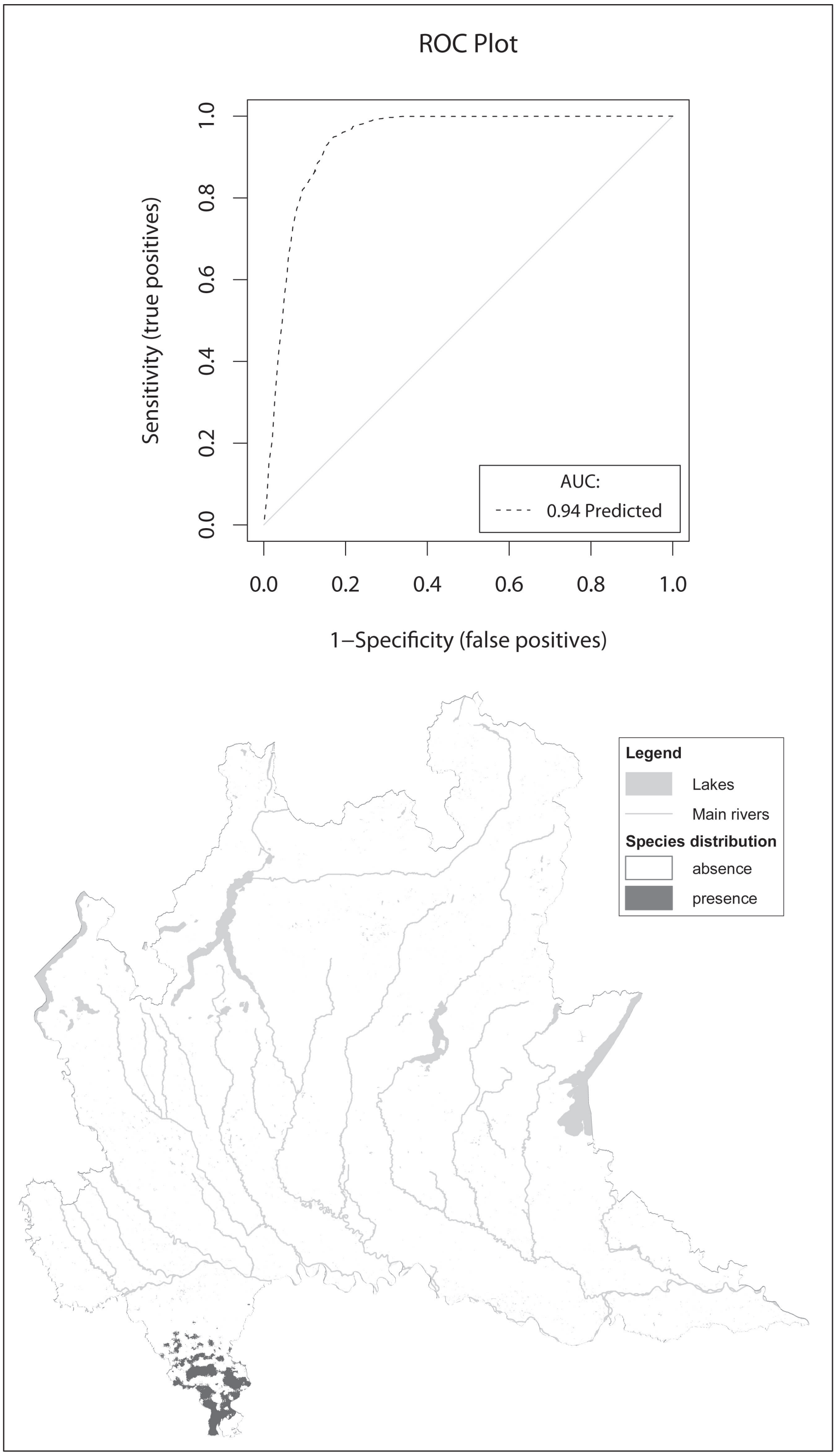


Figure A.70.: LRA for *Rana italica* (italian stream frog) ROC plot and potential distribution

Rana latastei (italian agile frog)

Table A.45.: selected environmental variables and computed coefficients from stepwise analysis for *Rana latastei* (italian agile frog)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	3.365e + 02	2.202e + 01	15.284	< 2e − 16
EV-CL15	2.740e − 02	1.849e − 02	1.482	0.138351
EV-CL18	4.091e − 02	1.997e − 03	20.484	< 2e − 16
EV-CL01	2.236e − 02	6.780e − 03	3.298	0.000975
EV-CL07	6.458e − 02	5.814e − 03	11.107	< 2e − 16
EV-D100	-2.312e − 05	3.347e − 06	-6.908	4.91e-12
EV-D103	-1.622e − 04	1.882e − 05	-8.619	< 2e − 16
EV-D106	-9.611e − 06	3.861e − 06	-2.490	0.012792
EV-D107	-4.507e − 05	3.944e − 06	-11.427	< 2e − 16
EV-D108	7.327e − 05	1.215e − 05	6.030	1.64e-09
EV-D109	-1.965e − 04	1.412e − 05	-13.918	< 2e − 16
EV-D110	-1.035e − 04	5.876e − 06	-17.610	< 2e − 16
EV-D111	1.897e − 05	3.480e − 06	5.453	4.97e − 08
EV-D112	2.905e − 05	7.447e − 06	3.901	9.59e − 05
EV-D114	4.013e − 05	2.188e − 05	1.834	0.066691
EV-D115	-1.112e − 04	2.731e − 05	-4.071	4.68e − 05
EV-D116	1.211e − 04	9.580e − 06	12.639	< 2e − 16
EV-D117	2.217e − 05	2.344e − 06	9.455	< 2e − 16
EV-D119	2.061e − 05	4.572e − 06	4.508	6.54e − 06
EV-D120	5.352e − 05	2.335e − 05	2.292	0.021912
EV-D121	8.237e − 05	4.775e − 05	1.725	0.084541
EV-D122	-2.290e − 04	4.771e − 05	-4.800	1.59e − 06
EV-D123	-1.019e − 04	2.641e − 05	-3.859	0.000114
EV-D124	-5.081e − 05	1.245e − 05	-4.082	4.47e − 05
EV-D125	-3.124e − 05	2.845e − 06	-10.981	< 2e − 16
EV-D126	-2.781e − 06	1.612e − 06	-1.725	0.084593
EV-DHID	1.897e − 05	1.170e − 05	1.621	0.104949
EV-DINF	3.042e − 04	6.997e − 05	4.348	1.37e − 05
EV-DTM	1.432e − 03	4.110e − 04	3.483	0.000496

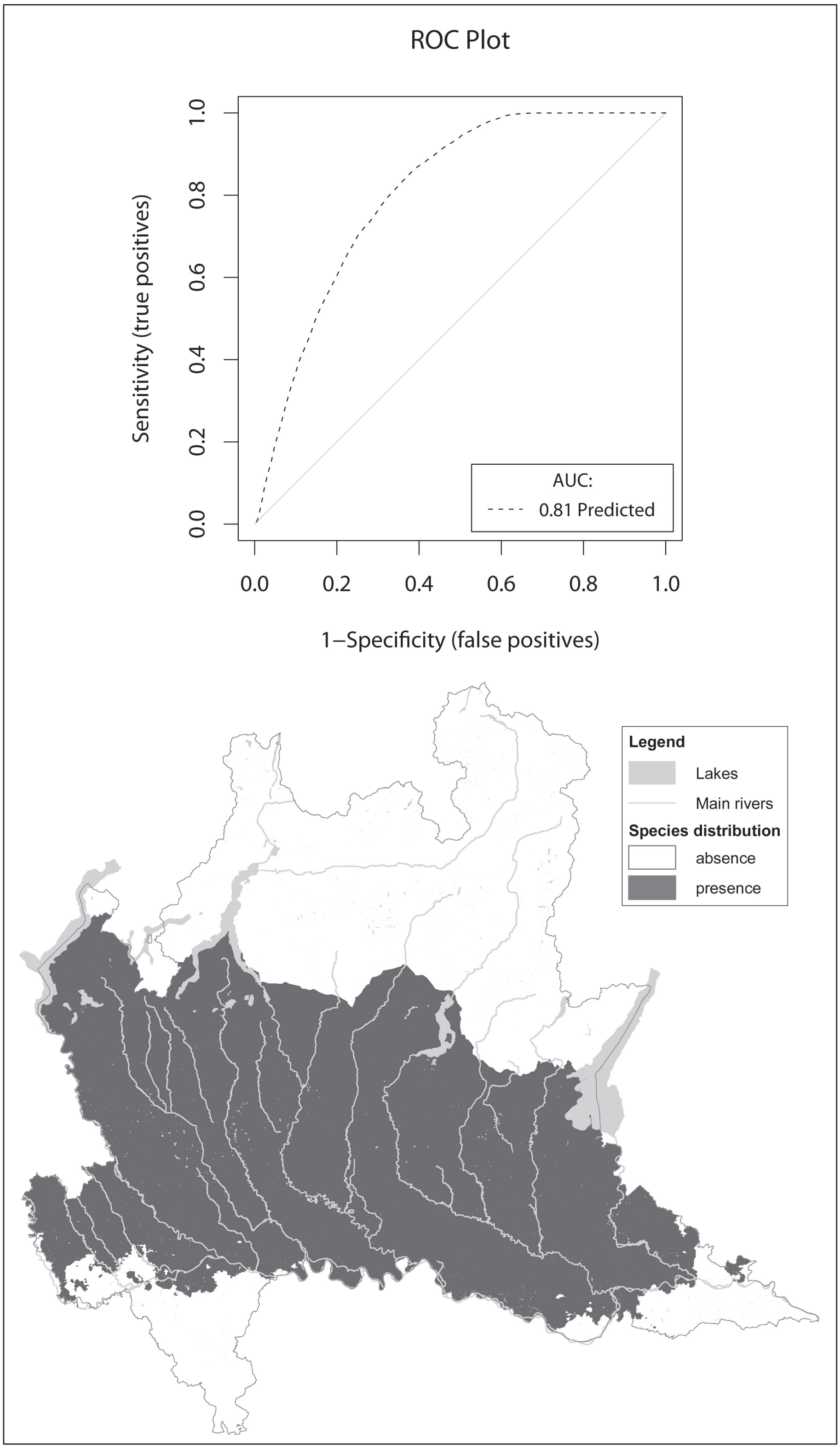


Figure A.71.: LRA for *Rana latastei* (italian agile frog) ROC plot and potential distribution

Rana synk. esculenta (edible frog)

Table A.46.: selected environmental variables and computed coefficients from stepwise analysis for *Rana synk. esculenta* (edible frog)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>|$z$|$)**
Intercept	-2.503$e-03$	2.854$e-05$	0.011	0.992
EV-CL18	-4.466$e-03$	5.868$e-04$	-7.612	2.71$e-14$
EV-CL01	1.502$e-02$	4.041$e-03$	3.716	0.000202
EV-D100	2.913$e-05$	5.226$e-06$	5.575	2.47$e-08$
EV-D104	1.146$e-04$	6.477$e-05$	1.769	0.076865
EV-D105	5.473$e-05$	3.816$e-05$	1.434	0.151488
EV-D106	-1.666$e-04$	1.409$e-05$	-11.818	$<2e-16$
EV-D108	-1.627$e-04$	5.332$e-05$	-3.052	0.002273
EV-D109	1.704$e-04$	2.969$e-05$	5.741	9.40$e-09$
EV-D112	9.065$e-05$	2.398$e-05$	3.780	0.000157
EV-D113	-7.134$e-04$	1.067$e-04$	-6.684	2.32$e-11$
EV-D117	7.511e-06	4.072$e-06$	1.844	0.065118
EV-D119	-1.812$e-04$	2.854$e-05$	-6.351	2.14$e-10$
EV-D120	3.944$e-05$	1.524$e-05$	2.588	0.009658
EV-D122	-1.522$e-04$	2.728$e-05$	-5.579	2.42$e-08$
EV-D126	-1.855$e-05$	1.118$e-06$	-16.595	$<2e-16$
EV-DHYD	-2.331$e-05$	1.138$e-05$	-2.048	0.040555
EV-DINF	5.284$e-05$	2.784$e-05$	1.898	0.057685
EV-DTM	-3.304$e-04$	1.938$e-04$	-1.705	0.088208

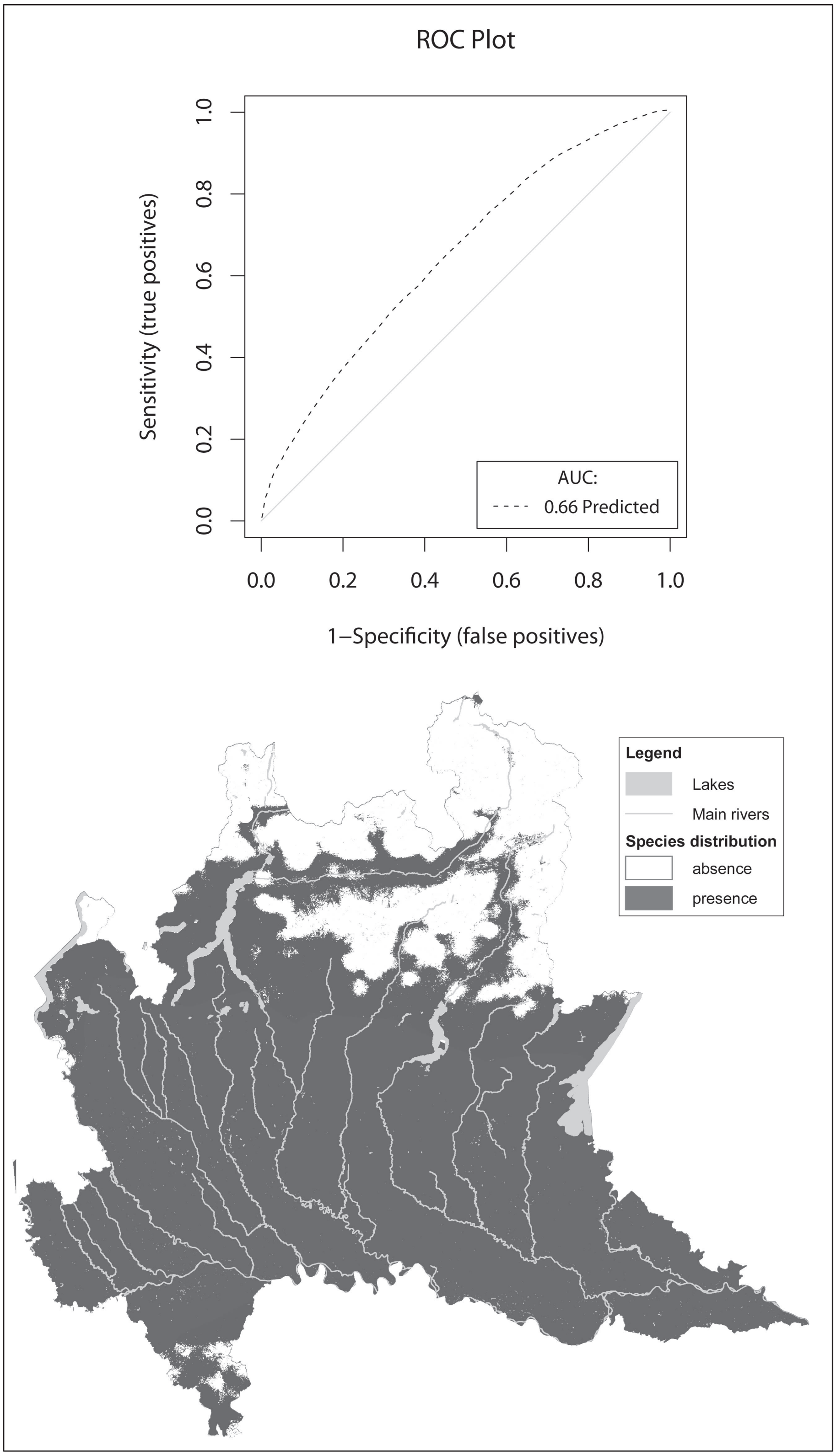


Figure A.72.: LRA for *Rana synk. esculenta* (edible frog) ROC plot and potential distribution

Rana temporaria (common frog)

Table A.47.: selected environmental variables and computed coefficients from stepwise analysis for *Rana temporaria* (common frog)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>|\mathbf{z}|$)**
Intercept	4.021$e+01$	5.255$e+00$	7.653	1.97e-14
EV$_a spect$	-4.718$e-04$	1.843$e-04$	-2.561	0.010446
EV-CL15	6.338$e-02$	6.565$e-03$	9.654	$< 2e-16$
EV-CL18	1.355$e-02$	6.530$e-04$	20.753	$< 2e-16$
EV-D100	6.878$e-06$	1.473$e-06$	4.670	3.01$e-06$
EV-D103	-7.056$e-05$	1.385$e-05$	-5.093	3.52$e-07$
EV-D104	1.146$e-04$	6.477$e-05$	1.769	0.076865
EV-D105	5.473$e-05$	3.816$e-05$	1.434	0.151488
EV-D106	2.021$e-04$	3.086$e-05$	6.550	5.75$e-11$
EV-D107	1.358$e-04$	1.848$e-05$	7.347	2.02$e-13$
EV-D108	1.367$e-04$	4.870$e-05$	2.808	0.004991
EV-D109	-1.451$e-04$	3.966$e-05$	-3.659	0.000253
EV-D110	9.785$e-05$	2.902$e-05$	3.372	0.000747
EV-D111	-2.175$e-04$	1.690$e-05$	-12.869	$< 2e-16$
EV-D112	9.446$e-05$	3.160$e-05$	2.989	0.002797
EV-D124	-4.722$e-05$	8.965$e-06$	-5.268	1.38$e-07$
EV-D125	-1.207$e-05$	2.031$e-06$	-5.945	2.76$e-09$
EV-D126	-2.821$e-06$	1.166$e-06$	-2.419	0.015555
EV-DHYD	-2.852$e-05$	1.150$e-05$	-2.480	0.013144
EV-DINF	2.424$e-04$	4.092$e-05$	5.923	3.16$e-09$

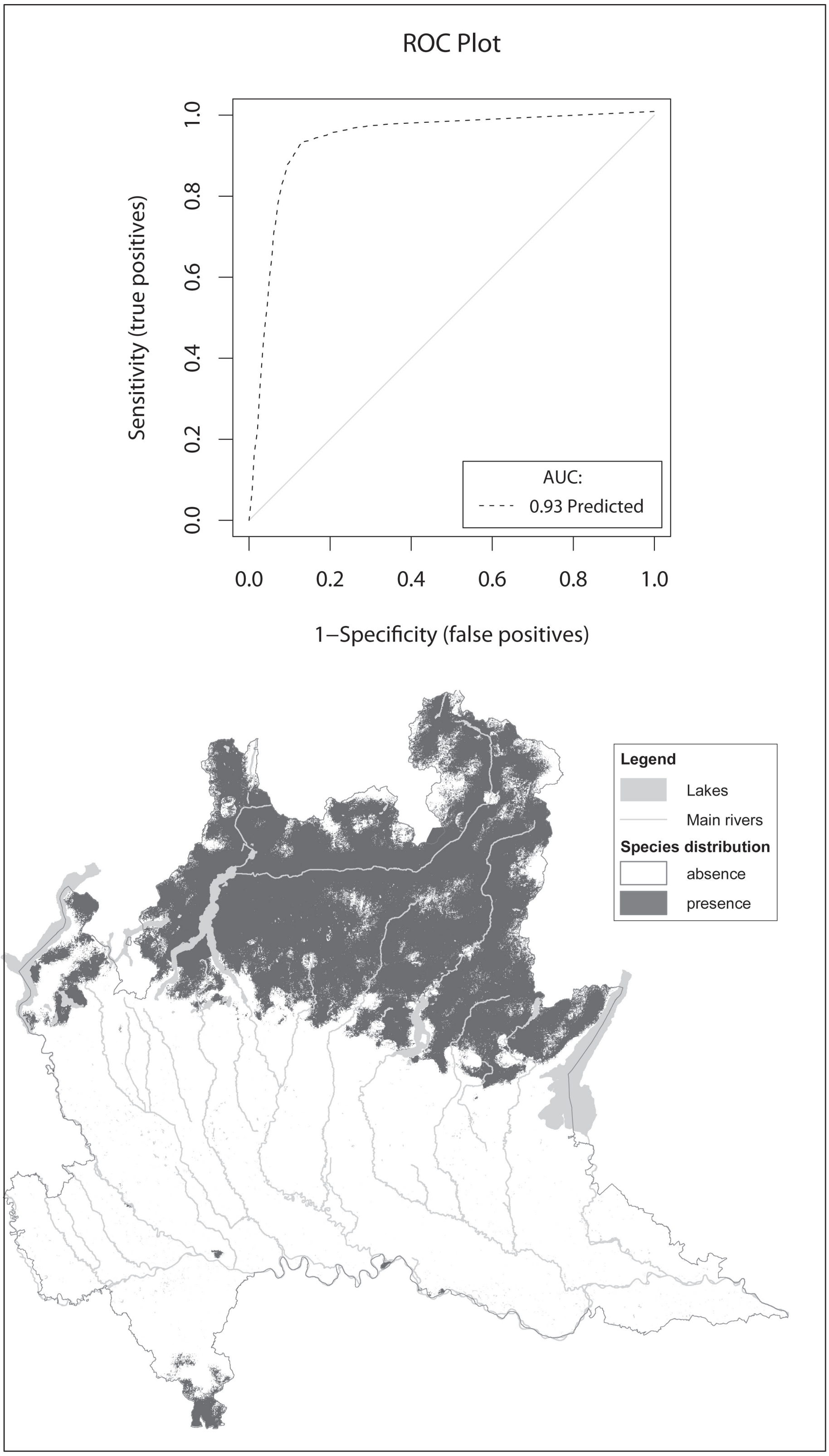


Figure A.73.: LRA for *Rana temporaria* (common frog) ROC plot and potential distribution

Emys orbicularis (european pond terrapin)

Table A.48.: selected environmental variables and computed coefficients from stepwise analysis for *Emys orbicularis* (european pond terrapin)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	-1.397e + 00	1.394e + 00	-1.002	0.316207
EV-CL15	5.180e - 02	6.925e - 03	7.480	7.41e - 14
EV-CL18	3.327e - 03	9.866e - 04	3.373	0.000744
EV-CL01	9.549e - 03	3.594e - 03	2.657	0.007887
EV-CL07	-7.895e - 03	4.711e - 03	-1.676	0.093758
EV-D100	1.795e - 05	2.047e - 06	8.768	< 2e - 16
EV-D103	-1.919e - 05	1.266e - 05	-1.516	0.129585
EV-D105	6.705e - 05	5.656e - 06	11.853	< 2e - 16
EV-D106	-6.225e - 06	4.163e - 06	-1.495	0.134820
EV-D107	-1.663e - 05	3.652e - 06	-4.553	5.30e - 06
EV-D108	3.582e - 05	1.041e - 05	3.443	0.000576
EV-D109	3.995e - 05	5.569e - 06	7.173	7.32e - 13
EV-D110	-3.783e - 05	4.960e - 06	-7.627	2.40e - 14
EV-D111	-2.121e - 05	3.317e - 06	-6.395	1.60e - 10
EV-D112	-2.656e - 05	7.277e - 06	-3.649	0.000263
EV-D113	-5.759e - 05	2.717e - 05	-2.120	0.034032
EV-D114	-9.613e - 05	2.439e - 05	-3.942	8.09e - 05
EV-D115	-2.591e - 05	9.771e - 06	-2.651	0.008017
EV-D116	-1.979e - 05	8.770e - 06	-2.257	0.024016
EV-D117	4.418e - 06	1.634e - 06	2.704	0.006861
EV-D119	-1.346e - 05	4.768e - 06	-2.822	0.004768
EV-D120	3.540e - 05	1.497e - 05	2.364	0.018059
EV-D121	6.148e - 05	2.601e - 05	2.364	0.018099
EV-D123	2.315e - 05	1.028e - 05	2.253	0.024286
EV-D126	-7.154e - 06	1.324e - 06	-5.405	6.49e - 08
EV-DHYD	-6.874e - 05	1.343e - 05	-5.118	3.09e - 07
EV-DINF	-5.883e - 05	3.315e - 05	-1.775	0.075946

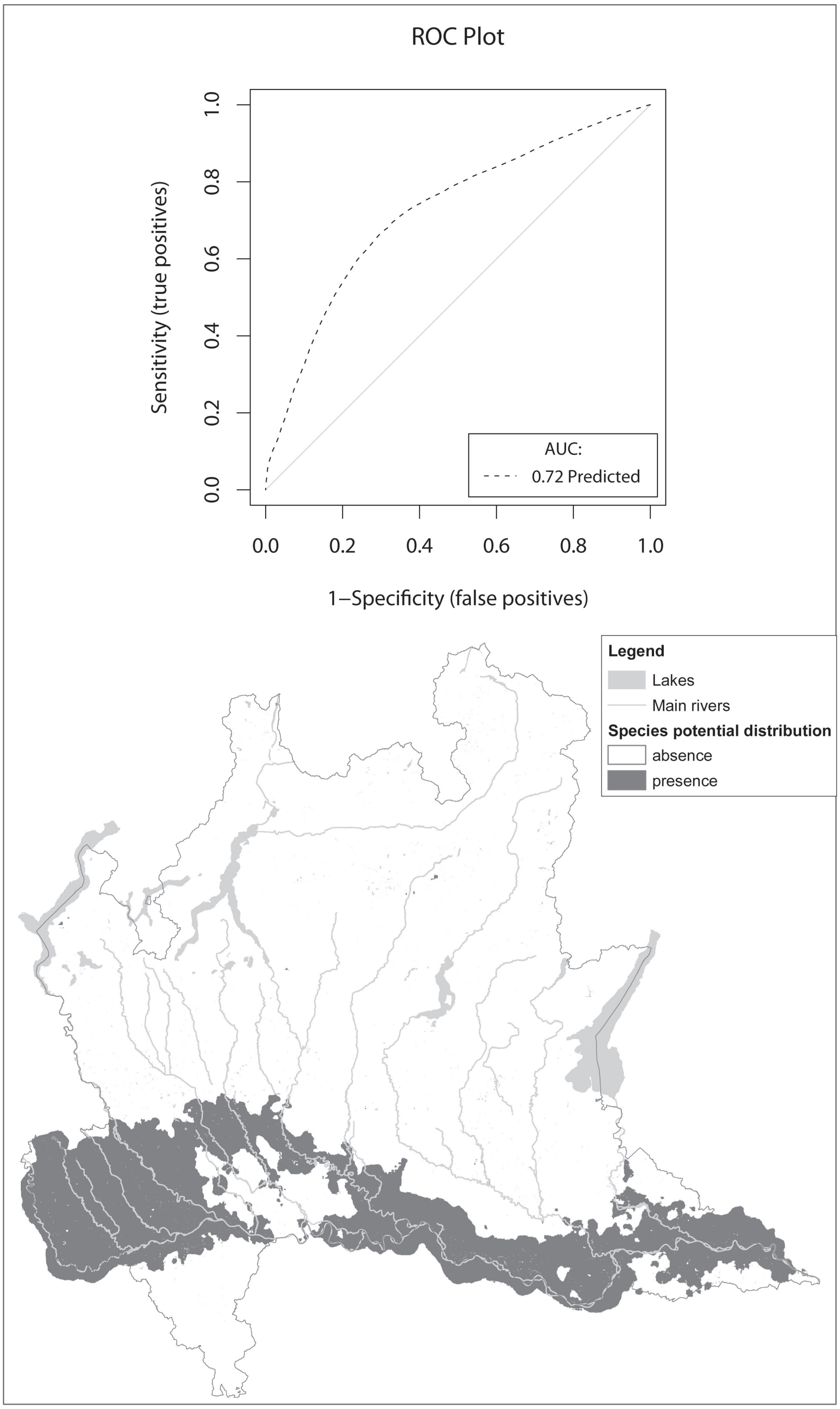


Figure A.74.: LRA for *Emys orbicularis* (european pond terrapin) ROC plot and potential distribution

Trachemys scripta (red eared terrapin)

Table A.49.: selected environmental variables and computed coefficients from stepwise analysis for *Trachemys scripta* (red eared terrapin)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$8.080e+00$	$9.753e-01$	8.285	$< 2e-16$
EV-ASPECT	$-3.148e-04$	$1.444e-04$	-2.180	0.029272
EV-CL15	$1.496e-02$	$5.904e-03$	2.533	0.011297
EV-CL18	$-4.466e-03$	$5.868e-04$	-7.612	$2.71e-14$
EV-CL01	$1.502e-02$	$4.041e-03$	3.716	0.000202
EV-CL07	$-2.841e-02$	$3.260e-03$	-8.715	$< 2e-16$
EV-D105	$4.926e-05$	$4.553e-06$	10.819	$< 2e-16$
EV-D106	$-3.671e-05$	$3.266e-06$	-11.239	$< 2e-16$
EV-D107	$-8.257e-06$	$3.269e-06$	-2.526	0.011553
EV-D109	$2.119e-05$	$4.449e-06$	4.762	1.92e-06
EV-D111	$-3.966e-05$	$3.205e-06$	-12.373	$< 2e-16$
EV-D112	$-1.227e-05$	$6.409e-06$	-1.915	0.055556
EV-D114	$-1.310e-04$	$2.116e-05$	-6.189	$6.06e-10$
EV-D116	$-2.871e-05$	$6.014e-06$	-4.774	$1.80e-06$
EV-D117	$1.061e-05$	$1.391e-06$	7.628	$2.39e-14$
EV-D119	$3.966e-05$	$3.642e-06$	10.888	$< 2e-16$
EV-D120	$2.094e-05$	$1.270e-05$	1.649	0.099211
EV-D123	$4.088e-05$	$7.453e-06$	5.485	4.13e-08
EV-D126	$-1.855e-05$	$1.118e-06$	-16.595	$< 2e-16$
EV-DHYD	$-2.331e-05$	$1.138e-05$	-2.048	0.040555
EV-DINF	$5.284e-05$	$2.784e-05$	1.898	0.057685
EV-DTM	$-3.304e-04$	$1.938e-04$	-1.705	0.088208

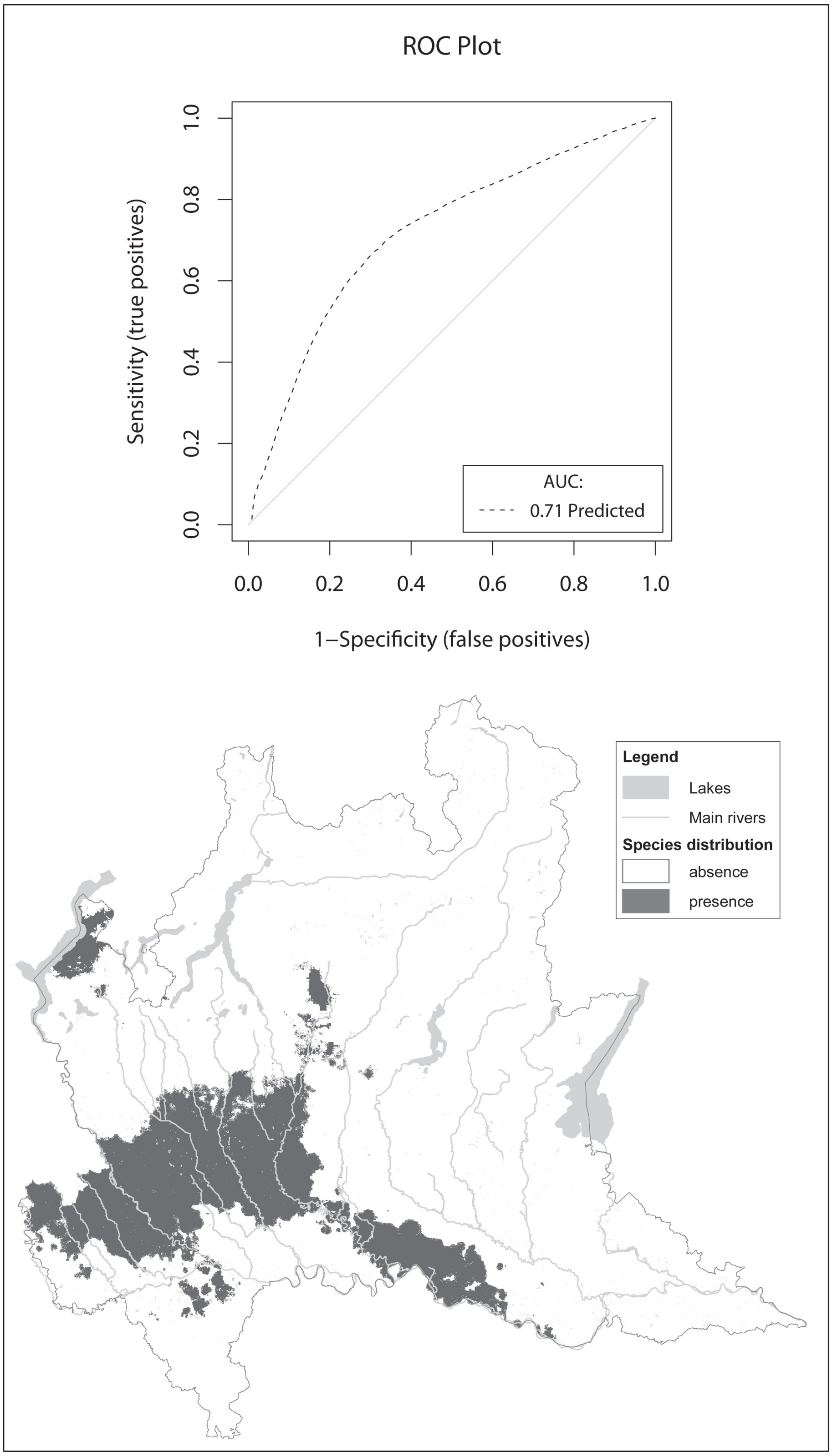


Figure A.75.: LRA for *Trachemys scripta* (red eared terrapin) ROC plot and potential distribution

Anguis fragilis (slow worm)

Table A.50.: selected environmental variables and computed coefficients from stepwise analysis for *Anguis fragilis* (slow worm)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\lvert z\rvert$)**
Intercept	$4.021e+01$	$5.255e+00$	7.653	$1.97e-14$
EV-CL15	-$4.219e-02$	$3.623e-03$	-11.645	$< 2e-16$
EV-CL18	-$7.187e-02$	$1.396e-02$	-5.150	$2.60e-07$
EV-D100	-$6.907e-05$	$7.706e-06$	-8.964	$< 2e-16$
EV-D103	$5.617e-04$	$6.193e-05$	9.070	$< 2e-16$
EV-D104	-$1.343e-04$	$4.889e-05$	-2.747	0.006007
EV-D106	-$1.666e-04$	$1.409e-05$	-11.818	$< 2e-16$
EV-D108	-$1.627e-04$	$5.332e-05$	-3.052	0.002273
EV-D109	$1.704e-04$	$2.969e-05$	5.741	$9.40e-09$
EV-D112	$9.065e-05$	$2.398e-05$	3.780	0.000157
EV-D113	-$7.134e-04$	$1.067e-04$	-6.684	$2.32e-11$
EV-D114	-$4.004e-04$	$9.488e-05$	-4.220	$2.44e-05$
EV-D115	$3.523e-04$	$7.611e-05$	4.629	$3.68e-06$
EV-D116	-$6.508e-04$	$5.125e-05$	-12.699	$< 2e-16$
EV-D119	$9.867e-05$	$8.573e-06$	11.509	$< 2e-16$
EV-D120	$3.328e-04$	$6.258e-05$	5.317	$1.05e-07$
EV-D122	$6.696e-04$	$1.192e-04$	5.617	$1.94e-08$
EV-D123	$5.674e-04$	$5.380e-05$	10.547	$< 2e-16$
EV-D124	-$2.579e-04$	$4.077e-05$	-6.325	$2.54e-10$
EV-D125	-$9.200e-05$	$8.329e-06$	-11.047	$< 2e-16$
EV-D127	-$6.235e-04$	$6.430e-05$	-9.696	$< 2e-16$
EV-DHYD	-$3.402e-04$	$4.936e-05$	-6.893	$5.45e-12$
EV-DTM	-$2.101e-03$	$5.397e-04$	-3.893	$9.89e-05$

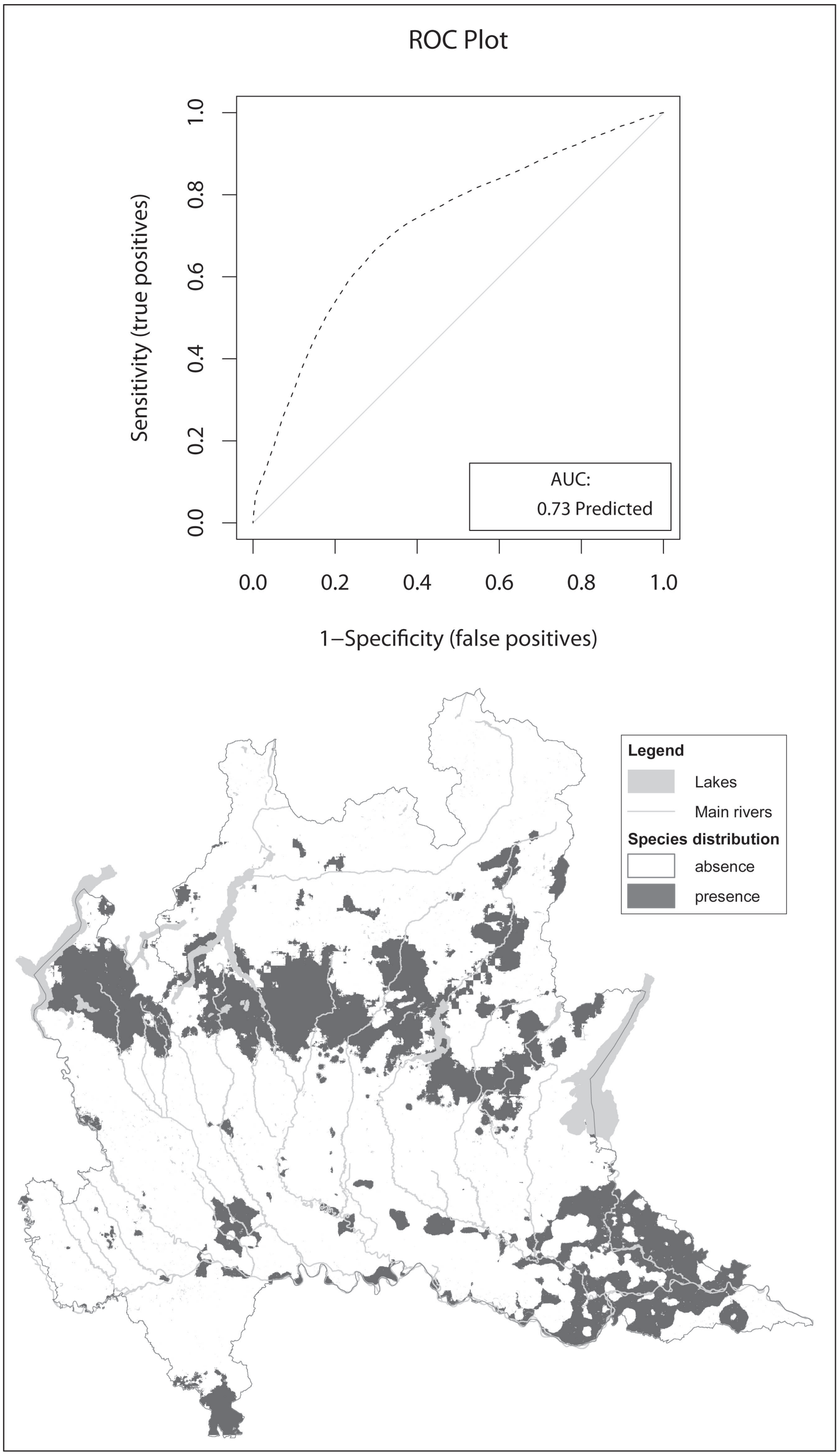


Figure A.76.: *Anguis fragilis* (slow worm) ROC plot and potential distribution

Lacerta bilineata (western green lizard)

Table A.51.: selected environmental variables and computed coefficients from stepwise analysis for *Lacerta bilineata* (western green lizard)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
Intercept	$1.924e+03$	$1.960e+02$	9.815	$< 2e-16$
EV-ASPECT	-$1.309e-03$	$5.052e-04$	-2.591	0.009582
EV-CL15	-$1.560e-01$	$5.744e-02$	-2.716	0.006606
EV-CL18	$1.366e-01$	$2.801e-02$	4.877	$1.08e-06$
EV-CL07	-$2.666e-01$	$2.553e-02$	-10.446	$< 2e-16$
EV-D100	-$2.741e-04$	$3.728e-05$	-7.354	$1.93e-13$
EV-D103	$5.617e-04$	$6.193e-05$	9.070	$< 2e-16$
EV-D104	-$1.343e-04$	$4.889e-05$	-2.747	0.006007
EV-D106	-$1.666e-04$	$1.409e-05$	-11.818	$< 2e-16$
EV-D108	-$1.627e-04$	$5.332e-05$	-3.052	0.002273
EV-D109	$1.704e-04$	$2.969e-05$	5.741	$9.40e-09$
EV-D112	$9.065e-05$	$2.398e-05$	3.780	0.000157
EV-D113	-$7.134e-04$	$1.067e-04$	-6.684	$2.32e-11$
EV-D114	-$4.004e-04$	$9.488e-05$	-4.220	$2.44e-05$
EV-D115	$3.523e-04$	$7.611e-05$	4.629	$3.68e-06$
EV-D121	$8.237e-05$	$4.775e-05$	1.725	0.084541
EV-D122	-$2.290e-04$	$4.771e-05$	-4.800	$1.59e-06$
EV-D123	-$1.019e-04$	$2.641e-05$	-3.859	0.000114
EV-D124	-$5.081e-05$	$1.245e-05$	-4.082	4.47e-05
EV-D125	-$3.124e-05$	$2.845e-06$	-10.981	$< 2e-16$
EV-D126	-$2.781e-06$	$1.612e-06$	-1.725	0.084593
EV-DHYD	$1.897e-05$	$1.170e-05$	1.621	0.104949
EV-DINF	$3.042e-04$	$6.997e-05$	4.348	$1.37e-05$
EV-DTM	$1.432e-03$	$4.110e-04$	3.483	0.000496

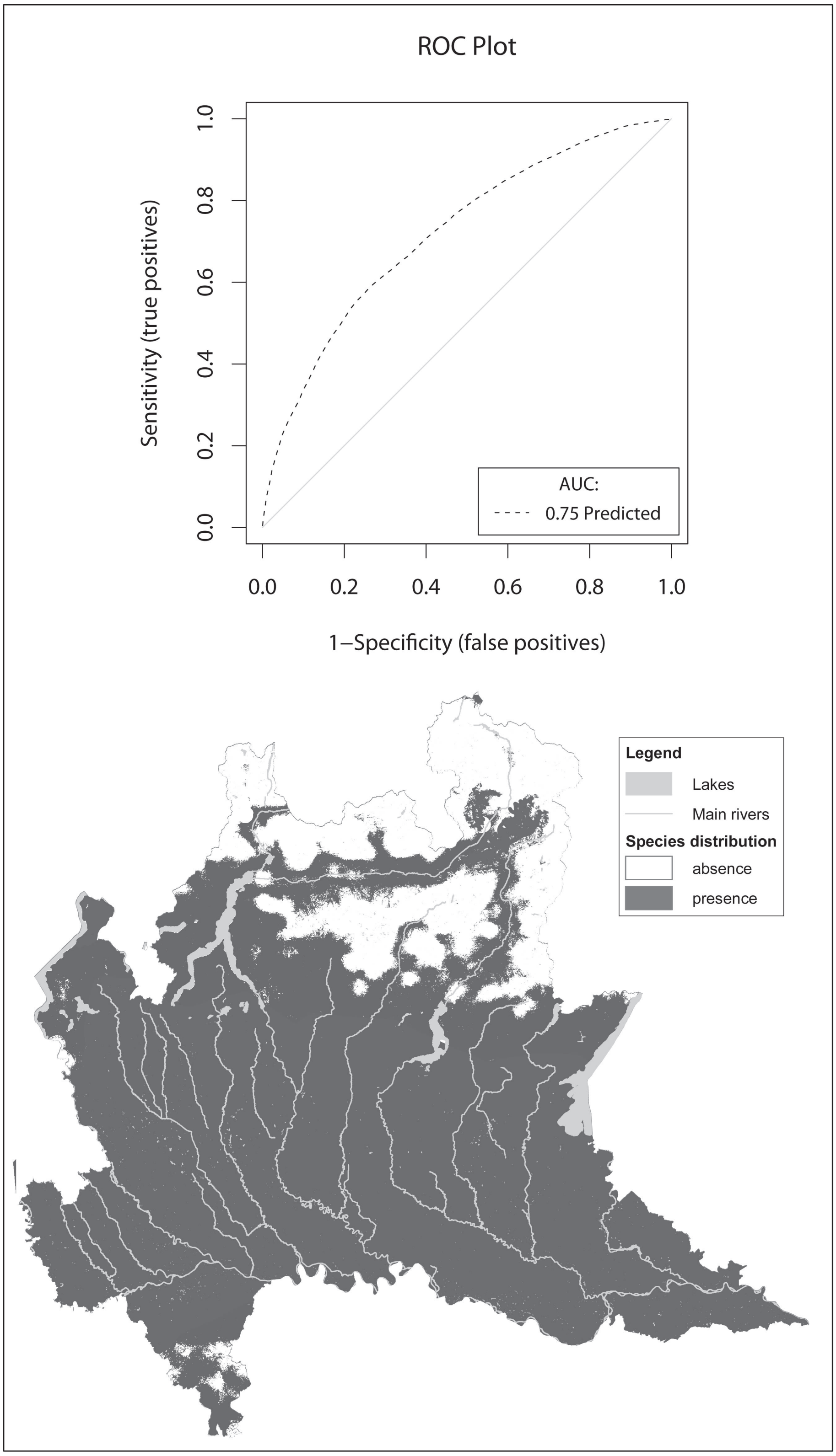


Figure A.77.: LRA for *Lacerta bilineata* (western green lizard) ROC plot and potential distribution

Podarcis muralis (common wall lizard)

Table A.52.: selected environmental variables and computed coefficients from stepwise analysis for *Podarcis muralis* (common wall lizard)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$8.239e-01$	$1.150e+00$	0.716	0.473698
EV-CL15	$9.924e-03$	$6.352e-03$	1.562	0.118233
EV-CL18	$1.739e-03$	$8.017e-04$	2.169	0.030071
EV-CL01	$6.443e-03$	$2.884e-03$	2.234	0.025487
EV-CL07	$-6.202e-03$	$3.824e-03$	-1.622	0.104772
EV-D100	$-8.259e-06$	$1.639e-06$	-5.040	$4.65e-07$
EV-D103	$-1.045e-04$	$1.292e-05$	-8.090	$5.97e-16$
EV-D104	$1.565e-05$	$1.013e-05$	1.546	0.122135
EV-D105	$2.509e-05$	$4.125e-06$	6.082	$1.19e-09$
EV-D106	$-7.233e-06$	$3.286e-06$	-2.201	0.027716
EV-D107	$-2.567e-05$	$3.131e-06$	-8.197	$2.46e-16$
EV-D109	$1.700e-05$	$4.866e-06$	3.494	0.000476
EV-D110	$-2.561e-05$	$4.381e-06$	-5.846	5.04e-09
EV-D112	$1.031e-05$	$6.065e-06$	1.700	0.089193
EV-D114	$-5.356e-05$	$2.067e-05$	-2.591	0.009571
EV-D115	$-8.683e-05$	$8.843e-06$	-9.819	$< 2e-16$
EV-D117	$1.291e-05$	$1.236e-06$	10.449	$< 2e-16$
EV-D119	$2.052e-05$	$3.813e-06$	5.383	$7.33e-08$
EV-D120	$-3.785e-05$	$1.408e-05$	-2.688	0.007180
EV-D121	$-1.205e-04$	$2.751e-05$	-4.380	$1.19e-05$
EV-D123	$1.862e-05$	$9.339e-06$	1.994	0.046196
EV-D126	$-3.930e-06$	$1.059e-06$	-3.709	0.000208
EV-DINF	$1.387e-04$	$3.215e-05$	4.314	$1.60e-05$

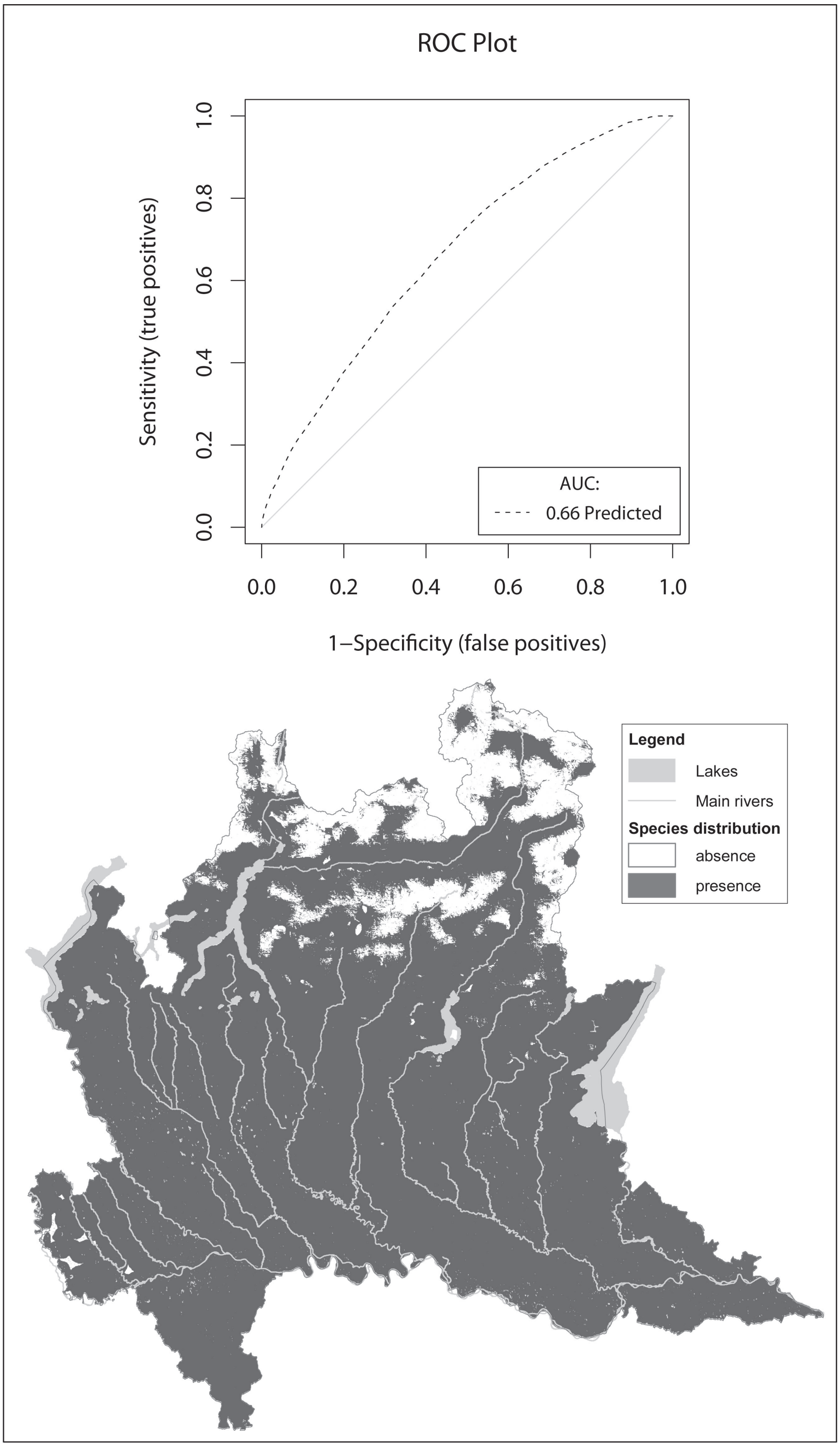


Figure A.78.: LRA for *Podarcis muralis* (common wall lizard) ROC plot and potential distribution

Zootoca vivipara (viviparous lizard)

Table A.53.: selected environmental variables and computed coefficients from stepwise analysis for *Zootoca vivipara* (viviparous lizard)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
Intercept	$1.138e+00$	$7.849e-01$	1.450	0.147074
EV-CL18	-$3.167e-03$	$1.282e-03$	-2.470	0.013522
EV-CL01	$6.527e-03$	$3.423e-03$	1.907	0.056564
EV-D100	-$8.235e-05$	$6.622e-06$	-12.436	$< 2e-16$
EV-D103	-$8.216e-05$	$2.160e-05$	-3.804	0.000142
EV-D104	$3.035e-05$	$1.327e-05$	2.287	0.022180
EV-D105	-$7.195e-05$	$2.199e-05$	-3.272	0.001069
EV-D106	-$4.130e-05$	$3.975e-06$	-10.389	$< 2e-16$
EV-D107	-$3.895e-05$	$3.923e-06$	-9.927	$< 2e-16$
EV-D109	-$4.112e-05$	$8.536e-06$	-4.817	$1.46e-06$
EV-D111	-$1.608e-05$	$3.076e-06$	-5.227	$1.72e-07$
EV-D112	-$2.584e-05$	$7.075e-06$	-3.652	0.000260
EV-D113	-$1.744e-04$	$5.520e-05$	-3.160	0.001580
EV-D114	-$3.184e-04$	$5.177e-05$	-6.151	$7.69e-10$
EV-D116	$1.682e-04$	$2.184e-05$	7.703	$1.33e-14$
EV-D117	$5.334e-05$	$4.936e-06$	10.806	$< 2e-16$
EV-D119	$1.038e-04$	$1.287e-05$	8.059	$7.67e-16$
EV-D121	$3.369e-04$	$1.565e-04$	2.152	0.031364
EV-D123	-$1.215e-04$	$5.767e-05$	-2.106	0.035183
EV-D125	$3.378e-05$	$1.272e-05$	2.655	0.007923
EV-DTM	$1.583e-02$	$3.629e-03$	4.361	$1.29e-05$

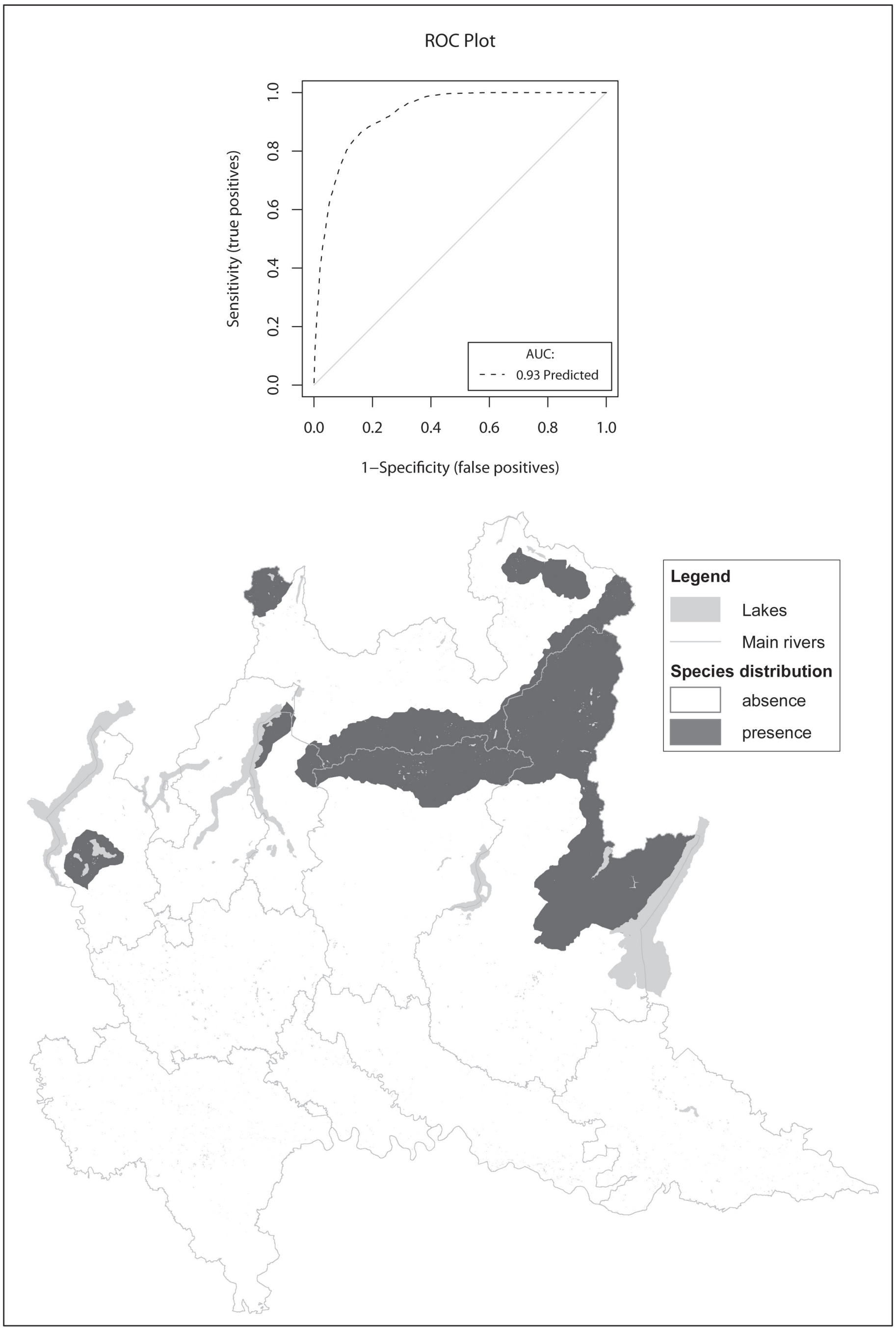


Figure A.79.: LRA for *Zootoca vivipara* (viviparous lizard)ROC plot and potential distribution

Chalcides chalcides (three-toed skink)

Table A.54.: selected environmental variables and computed coefficients from stepwise analysis for *Chalcides chalcides* (three-toed skink)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$1.411e+02$	$3.327e+01$	4.241	$2.23e-05$
EV-CL15	-$2.208e+00$	$4.054e-01$	-5.445	$5.18e-08$
EV-CL18	-$4.226e-01$	$7.023e-02$	-6.017	$1.78e-09$
EV-CL01	-$5.281e-01$	$1.006e-01$	-5.248	$1.54e-07$
EV-CL07	$2.368e-01$	$8.279e-02$	2.860	0.004230
EV-D101	-$3.649e-04$	$1.268e-04$	-2.878	0.003997
EV-D102	$2.557e-04$	$1.298e-04$	1.970	0.048865
EV-D103	$5.901e-04$	$2.362e-04$	2.499	0.012470
EV-D105	$4.926e-05$	$4.553e-06$	10.819	$< 2e-16$
EV-D106	-$3.671e-05$	$3.266e-06$	-11.239	$< 2e-16$
EV-D107	-$8.257e-06$	$3.269e-06$	-2.526	0.011553
EV-D109	$2.119e-05$	$4.449e-06$	4.762	$1.92e-06$
EV-D111	-$3.966e-05$	$3.205e-06$	-12.373	$< 2e-16$
EV-D112	-$1.227e-05$	$6.409e-06$	-1.915	0.055556
EV-D114	-$1.310e-04$	$2.116e-05$	-6.189	$6.06e-10$
EV-D116	-$2.871e-05$	$6.014e-06$	-4.774	$1.80e-06$
EV-D117	$1.061e-05$	$1.391e-06$	7.628	$2.39e-14$
EV-D118	$1.399e-05$	$3.245e-06$	4.311	$1.62e-05$
EV-D119	$3.966e-05$	$3.642e-06$	10.888	$< 2e-16$
EV-D120	$2.094e-05$	$1.270e-05$	1.649	0.099211
EV-D127	$3.843e-05$	$8.971e-06$	4.284	$1.83e-05$
EV-DHYD	-$8.686e-05$	$1.545e-05$	-5.622	$1.88e-08$
EV-DINF	$7.956e-05$	$3.581e-05$	2.222	0.026295
EV-SLOPE	-$7.029e-04$	$4.018e-04$	-1.749	0.080222

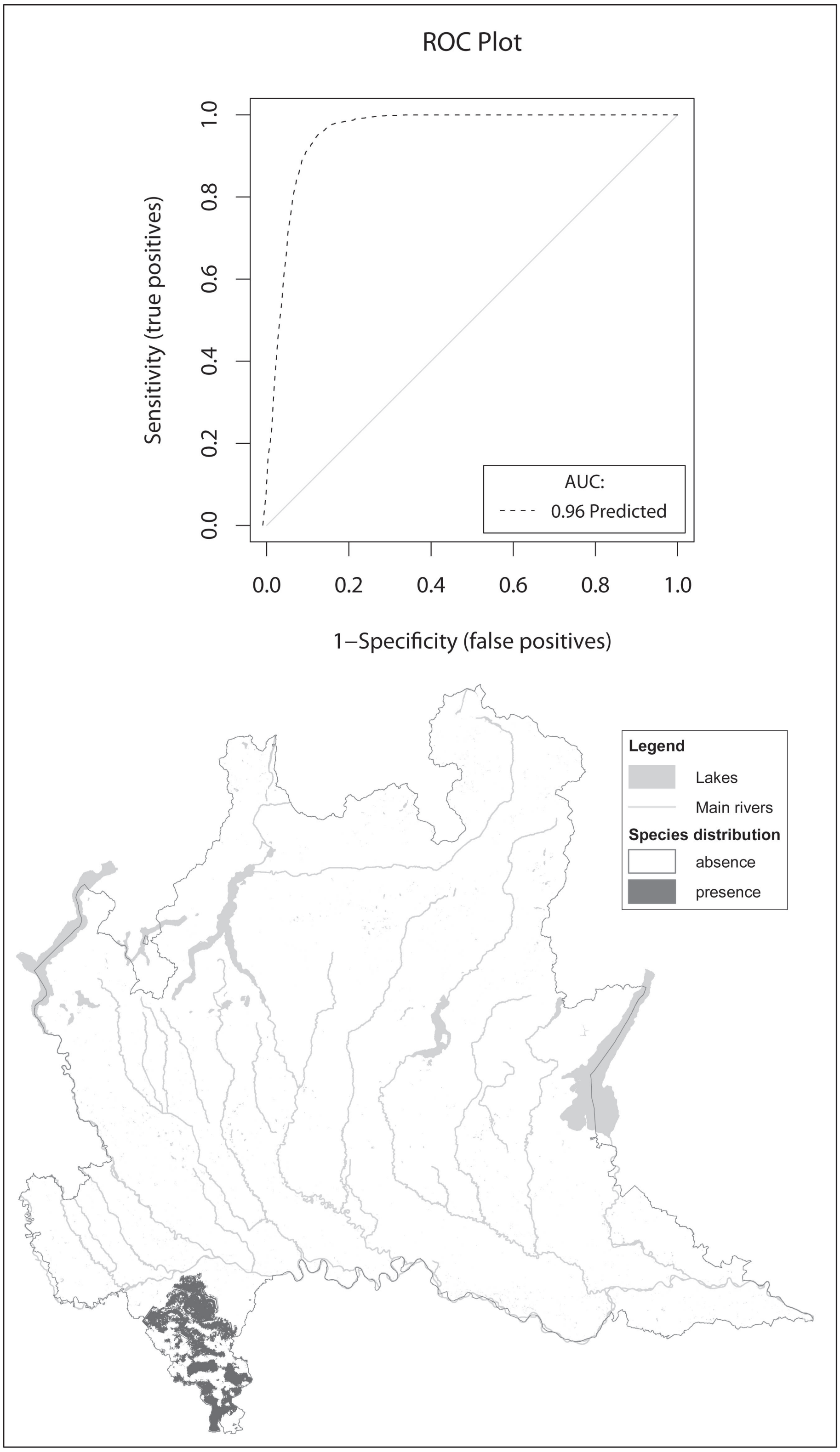


Figure A.80.: LRA for *Chalcides chalcides* (three-toed skink) ROC plot and potential distribution

Coluber viridiflavus (western whip snake)

Table A.55.: selected environmental variables and computed coefficients from stepwise analysis for *Coluber viridiflavus* (western whip snake)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
Intercept	-5.922$e-01$	5.837$e-01$	-1.015	0.310265
EV-CL15	4.210$e-02$	7.193$e-03$	5.853	4.82$e-09$
EV-CL18	4.447$e-03$	5.497$e-04$	8.089	6.00$e-16$
EV-CL01	-7.729$e-03$	3.561$e-03$	-2.171	0.029957
EV-D100	3.347$e-06$	1.310$e-06$	2.556	0.010597
EV-D101	4.242$e-05$	7.629$e-06$	5.560	2.69$e-08$
EV-D102	1.153$e-05$	5.507$e-06$	2.093	0.036356
EV-D103	-1.690$e-04$	1.474$e-05$	11.466	$< 2e-16$
EV-D104	4.206$e-05$	1.042$e-05$	4.038	5.40$e-05$
EV-D105	4.896$e-05$	4.258$e-06$	11.500	$< 2e-16$
EV-D106	-1.862$e-05$	3.358$e-06$	-5.547	2.91$e-08$
EV-D107	-1.042$e-05$	3.205$e-06$	-3.251	0.001150
EV-D108	-4.682$e-05$	9.769$e-06$	-4.793	1.64$e-06$
EV-D109	1.831$e-05$	5.417$e-06$	3.379	0.000727
EV-D110	-2.678$e-05$	4.593$e-06$	-5.831	5.53$e-09$
EV-D111	-2.578$e-05$	2.922$e-06$	-8.824	$< 2e-16$
EV-D112	-2.045$e-05$	6.452$e-06$	-3.169	0.001529
EV-D113	6.806$e-05$	2.375$e-05$	2.865	0.004169
EV-D115	-1.170$e-04$	1.212$e-05$	-9.653	$< 2e-16$
EV-D116	-2.009$e-05$	7.458$e-06$	-2.694	0.007067
EV-D117	7.346$e-06$	1.272$e-06$	5.773	7.77$e-09$
EV-D119	7.720$e-06$	3.494$e-06$	2.209	0.027141
EV-D120	-7.606$e-05$	1.587$e-05$	-4.792	1.65$e-06$
EV-D121	-5.776$e-05$	2.801$e-05$	-2.062	0.039201
EV-D126	-8.922$e-06$	1.104$e-06$	-8.079	6.53$e-16$
EV-D127	2.518$e-05$	9.022$e-06$	2.791	0.005248
EV-DHYD	-2.659$e-05$	1.121$e-05$	-2.372	0.017711
EV-SLOPE	-5.119$e-04$	2.007$e-04$	-2.551	0.010740
EV-DTM	-6.676$e-04$	4.074$e-04$	-1.639	0.101291

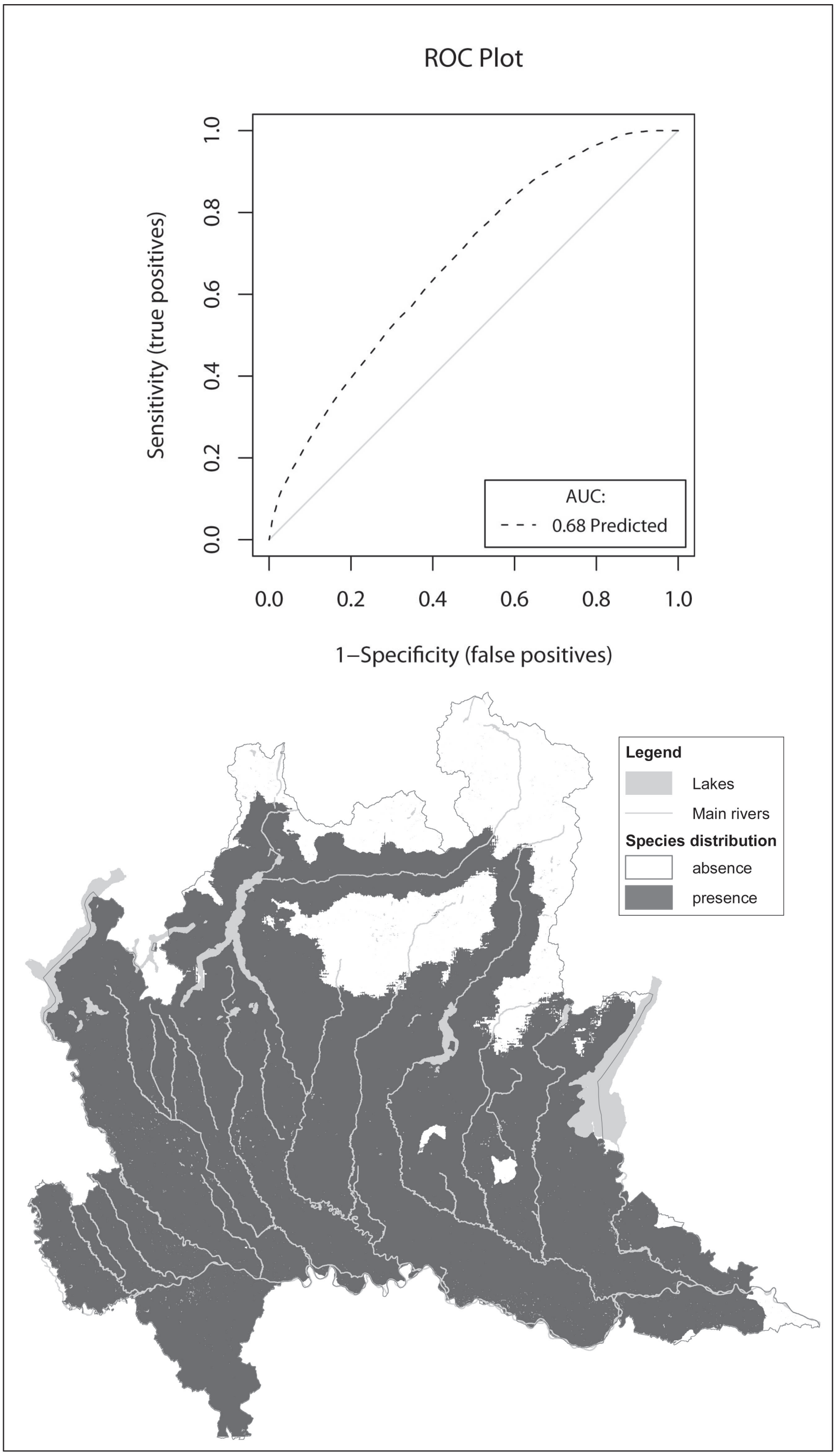


Figure A.81.: LRA for *Coluber viridiflavus* (western whip snake) ROC plot and potential distribution

Coronella austriaca (smooth snake)

Table A.56.: selected environmental variables and computed coefficients from stepwise analysis for *Coronella austriaca* (smooth snake)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$4.021e+01$	$5.255e+00$	7.653	$1.97e-14$
EV-CL15	$-4.219e-02$	$3.623e-03$	11.645	$< 2e-16$
EV-CL18	$-7.187e-02$	$1.396e-02$	-5.150	$2.60e-07$
EV-D100	$-6.907e-05$	$7.706e-06$	-8.964	$< 2e-16$
EV-D104	$1.509e-04$	$3.461e-05$	4.359	$1.31e-05$
EV-D105	$1.777e-04$	$2.890e-05$	6.150	$7.74e-10$
EV-D106	$5.037e-05$	$1.042e-05$	4.834	$1.34e-06$
EV-D108	$-1.683e-04$	$3.799e-05$	-4.429	$9.46e-06$
EV-D109	$-1.663e-04$	$4.948e-05$	-3.360	0.000779
EV-D110	$-1.136e-04$	$1.798e-05$	-6.316	$2.68e-10$
EV-D113	$1.569e-04$	$2.367e-05$	6.630	$3.36e-11$
EV-D115	$-3.037e-05$	$1.060e-05$	-2.865	0.004164
EV-D116	$-6.004e-05$	$7.452e-06$	-8.057	$7.85e-16$
EV-D117	$5.636e-06$	$1.298e-06$	4.343	$1.41e-05$
EV-D119	$8.833e-05$	$3.312e-06$	26.671	$< 2e-16$
EV-D123	$5.674e-04$	$5.380e-05$	10.547	$< 2e-16$
EV-D124	$-2.579e-04$	$4.077e-05$	-6.325	$2.54e-10$
EV-D125	$-9.200e-05$	$8.329e-06$	11.047	$< 2e-16$
EV-D126	$-5.993e-06$	$1.038e-06$	-5.772	$7.81e-09$
EV-DHYD	$-3.425e-05$	$1.145e-05$	-2.992	0.00277
EV-DTM	$-6.109e-04$	$3.481e-04$	-1.755	0.07927

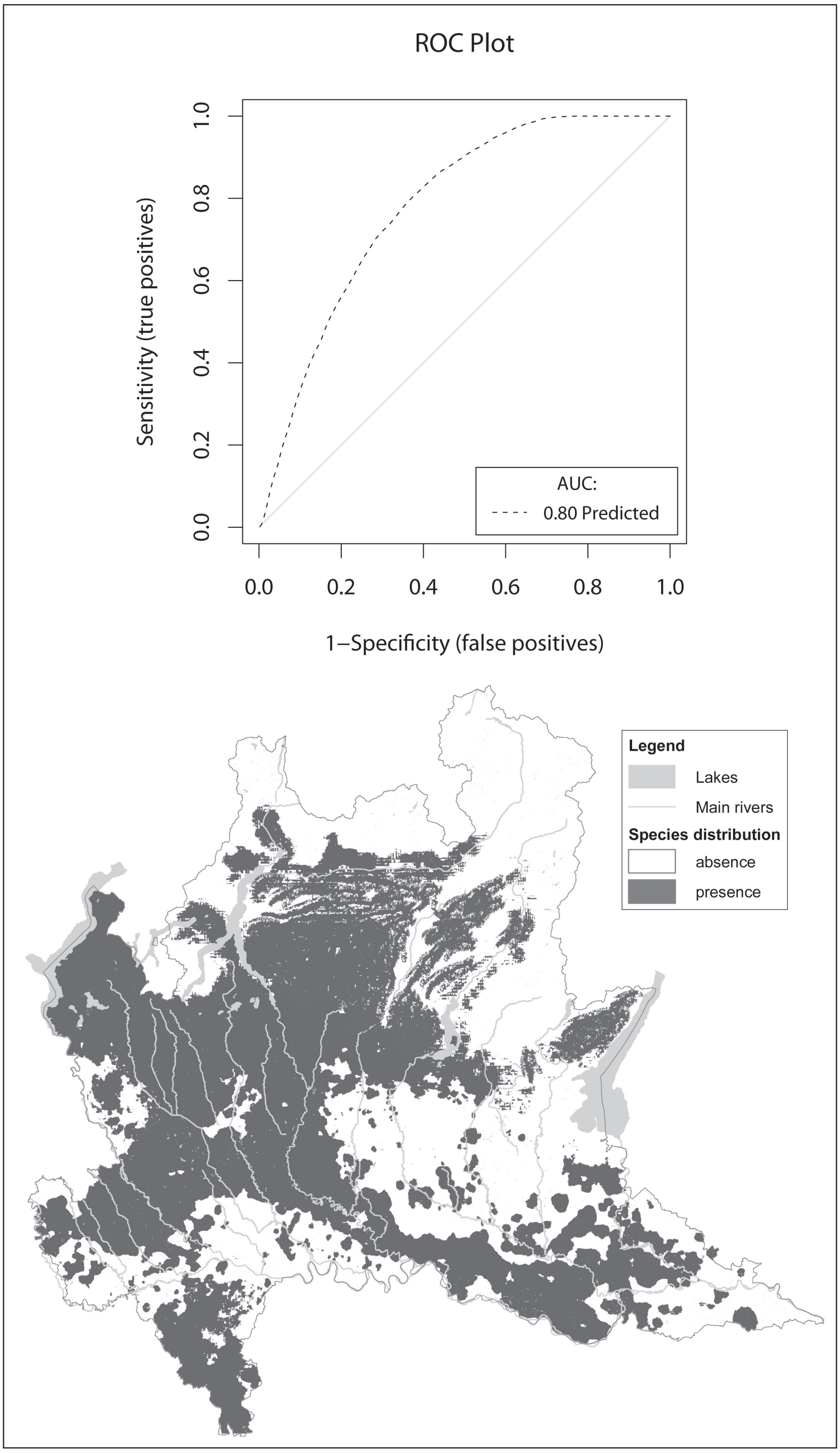


Figure A.82.: LRA for *Coronella austriaca* (smooth snake) ROC plot and potential distribution

Coronella girondica (southern smooth snake)

Table A.57.: selected environmental variables and computed coefficients from stepwise analysis for *Coronella girondica* (southern smooth snake)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\lvert z\rvert$)**
Intercept	$1.411e+02$	$3.327e+01$	4.241	$2.23e-05$
EV-CL15	$-2.208e+00$	$4.054e-01$	-5.445	$5.18e-08$
EV-CL18	$-4.226e-01$	$7.023e-02$	-6.017	$1.78e-09$
EV-CL07	$-5.281e-01$	$1.006e-01$	-5.248	$1.54e-07$
EV-CL01	$2.368e-01$	$8.279e-02$	2.860	0.004230
EV-D103	$5.901e-04$	$2.362e-04$	2.499	0.012470
EV-D104	$6.511e-04$	$1.941e-04$	3.355	0.000793
EV-D106	$-6.284e-04$	$1.220e-04$	-5.149	$2.62e-07$
EV-D107	$-1.852e-04$	$7.800e-05$	-2.374	0.017603
EV-D110	$2.885e-04$	$9.959e-05$	2.897	0.003767
EV-D111	$8.173e-04$	$2.218e-04$	3.685	0.000229
EV-D112	$-1.045e-03$	$1.793e-04$	-5.826	$5.67e-09$
EV-D113	$5.494e-04$	$3.095e-04$	1.775	0.075871
EV-D117	$1.282e-04$	$6.001e-05$	2.136	0.032641
EV-D119	$-8.997e-04$	$2.354e-04$	-3.822	0.000133
EV-D122	$9.402e-04$	$4.791e-04$	1.962	0.049708
EV-D123	$-4.376e-04$	$2.578e-04$	-1.697	0.089676
EV-D126	$-2.445e-04$	$5.761e-05$	-4.243	$2.20e-05$
EV-DINF	$-1.975e-03$	$7.509e-04$	-2.631	0.008518
EV-DHYD	$-3.156e-02$	$1.087e-02$	-2.904	0.003683

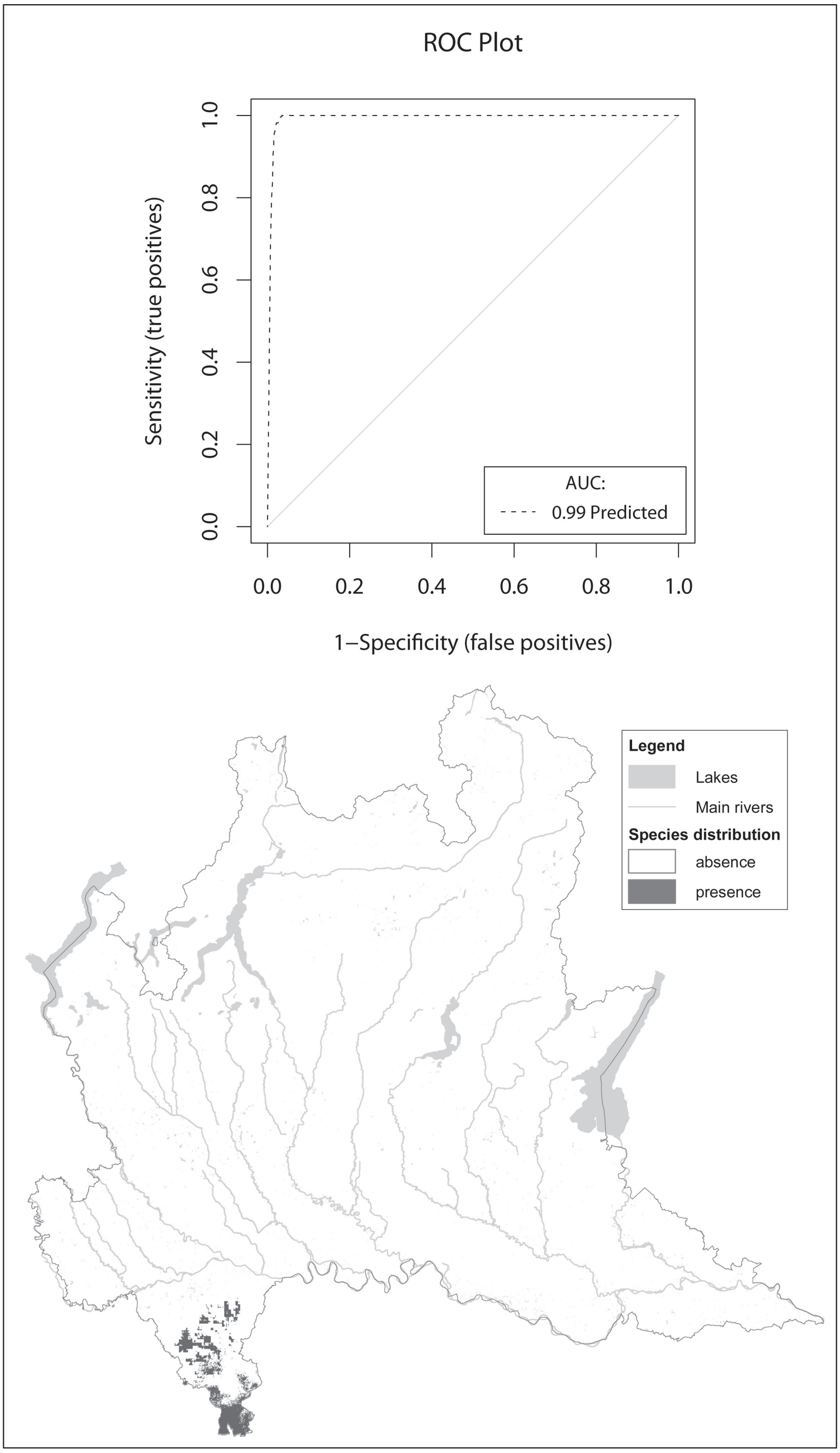


Figure A.83.: LRA for *Coronella girondica* (southern smooth snake) ROC plot and potential distribution

Elaphe longissima (aesculapian snake)

Table A.58.: selected environmental variables and computed coefficients from stepwise analysis for *Elaphe longissima* (aesculapian snake)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	$3.841e+00$	$1.216e+00$	3.159	0.001581
EV-ASPECT	$2.959e-04$	$1.510e-04$	1.959	0.050080
EV-CL15	$2.661e-02$	$7.260e-03$	3.666	0.000246
EV-CL18	$2.196e-03$	$7.940e-04$	2.766	0.005672
EV-CL01	$2.765e-02$	$3.981e-03$	6.946	$3.75e-12$
EV-CL07	$-2.442e-02$	$3.737e-03$	-6.533	$6.47e-11$
EV-D100	$2.769e-06$	$1.452e-06$	1.907	0.056480
EV-D103	$-2.067e-04$	$1.985e-05$	-10.415	$< 2e-16$
EV-D104	$-4.389e-05$	$1.059e-05$	-4.143	$3.43e-05$
EV-D106	$-1.940e-05$	$3.612e-06$	-5.372	$7.78e-08$
EV-D107	$-1.257e-05$	$3.898e-06$	-3.225	0.001260
EV-D108	$-2.502e-05$	$9.266e-06$	-2.700	0.006926
EV-D109	$1.822e-05$	$4.911e-06$	3.709	0.000208
EV-D110	$-4.165e-05$	$4.769e-06$	-8.734	$< 2e-16$
EV-D111	$-8.022e-05$	$3.024e-06$	-26.525	$< 2e-16$
EV-D112	$1.646e-05$	$6.127e-06$	2.687	0.007207
EV-D114	$-1.503e-04$	$2.820e-05$	-5.332	$9.71e-08$
EV-D115	$-5.859e-05$	$1.088e-05$	-5.385	$7.23e-08$
EV-D117	$1.030e-05$	$1.352e-06$	7.613	$2.67e-14$
EV-D119	$6.012e-05$	$3.666e-06$	16.401	$< 2e-16$
EV-D127	$4.910e-05$	$9.020e-06$	5.444	$5.20e-08$
EV-DINF	$-7.812e-05$	$3.146e-05$	-2.483	0.013033
EV-DTM	$3.995e-04$	$1.871e-04$	2.136	0.032693

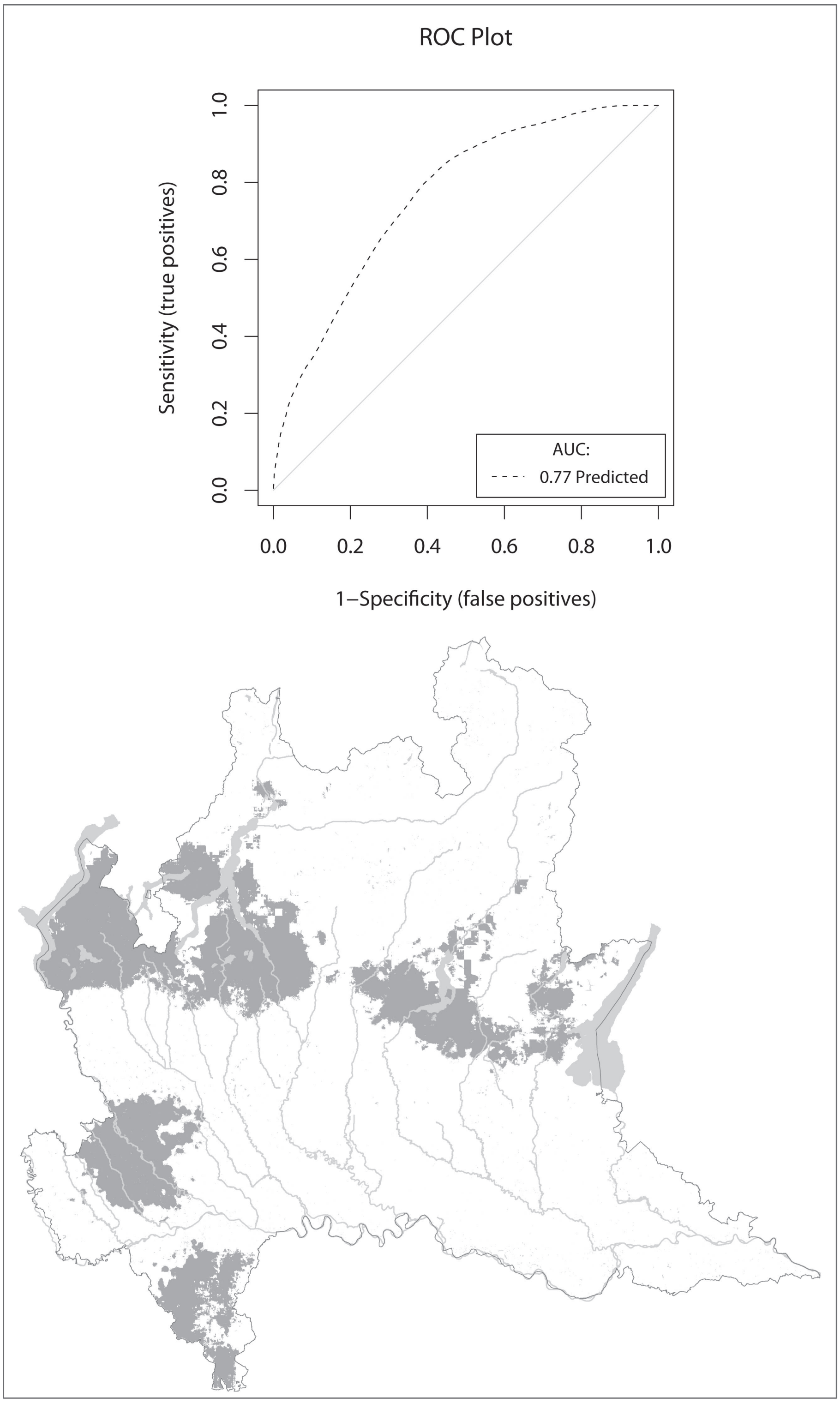


Figure A.84.: LRA for *Elaphe longissima* (aesculapian snake) ROC plot and potential distribution

Natrix maura (viperine snake)

Table A.59.: selected environmental variables and computed coefficients from stepwise analysis for *Natrix maura* (viperine snake)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\lvert z\rvert$)**
Intercept	$1.924e+03$	$1.960e+02$	9.815	$< 2e-16$
EV-ASPECT	$-1.309e-03$	$5.052e-04$	-2.591	0.009582
EV-CL15	$-1.560e-01$	$5.744e-02$	-2.716	0.006606
EV-CL18	$1.366e-01$	$2.801e-02$	4.877	$1.08e-06$
EV-CL01	$-2.666e-01$	$2.553e-02$	-10.446	$<< 2e-16$
EV-D100	$-2.741e-04$	$3.728e-05$	-7.354	$1.93e-13$
EV-D103	$-2.944e-04$	$7.443e-05$	-3.955	$7.65e-05$
EV-D105	$-2.964e-04$	$4.651e-05$	-6.372	$1.86e-10$
EV-D106	$2.021e-04$	$3.086e-05$	6.550	$5.75e-11$
EV-D107	$1.358e-04$	$1.848e-05$	7.347	$2.02e-13$
EV-D108	$1.367e-04$	$4.870e-05$	2.808	0.004991
EV-D109	$-1.451e-04$	$3.966e-05$	-3.659	0.000253
EV-D110	$9.785e-05$	$2.902e-05$	3.372	0.000747
EV-D111	$-2.175e-04$	$1.690e-05$	-12.869	$< 2e-16$
EV-D112	$9.446e-05$	$3.160e-05$	2.989	0.002797
EV-D113	$3.440e-04$	$1.113e-04$	3.091	0.001992
EV-D114	$-2.932e-04$	$9.313e-05$	-3.148	0.001643
EV-D115	$4.100e-04$	$1.045e-04$	3.923	$8.76e-05$
EV-D119	$1.379e-04$	$4.111e-05$	3.354	0.000795
EV-D120	$-2.730e-04$	$7.513e-05$	-3.633	0.000280
EV-D121	$-3.230e-04$	$1.295e-04$	-2.495	0.012606
EV-D122	$4.691e-04$	$1.421e-04$	3.302	0.000960
EV-D126	$2.791e-05$	$1.025e-05$	2.722	0.006481
EV-DINF	$-4.375e-04$	$2.391e-04$	-1.830	0.067283
EV-DTM	$-2.717e-03$	$1.380e-03$	-1.969	0.048953

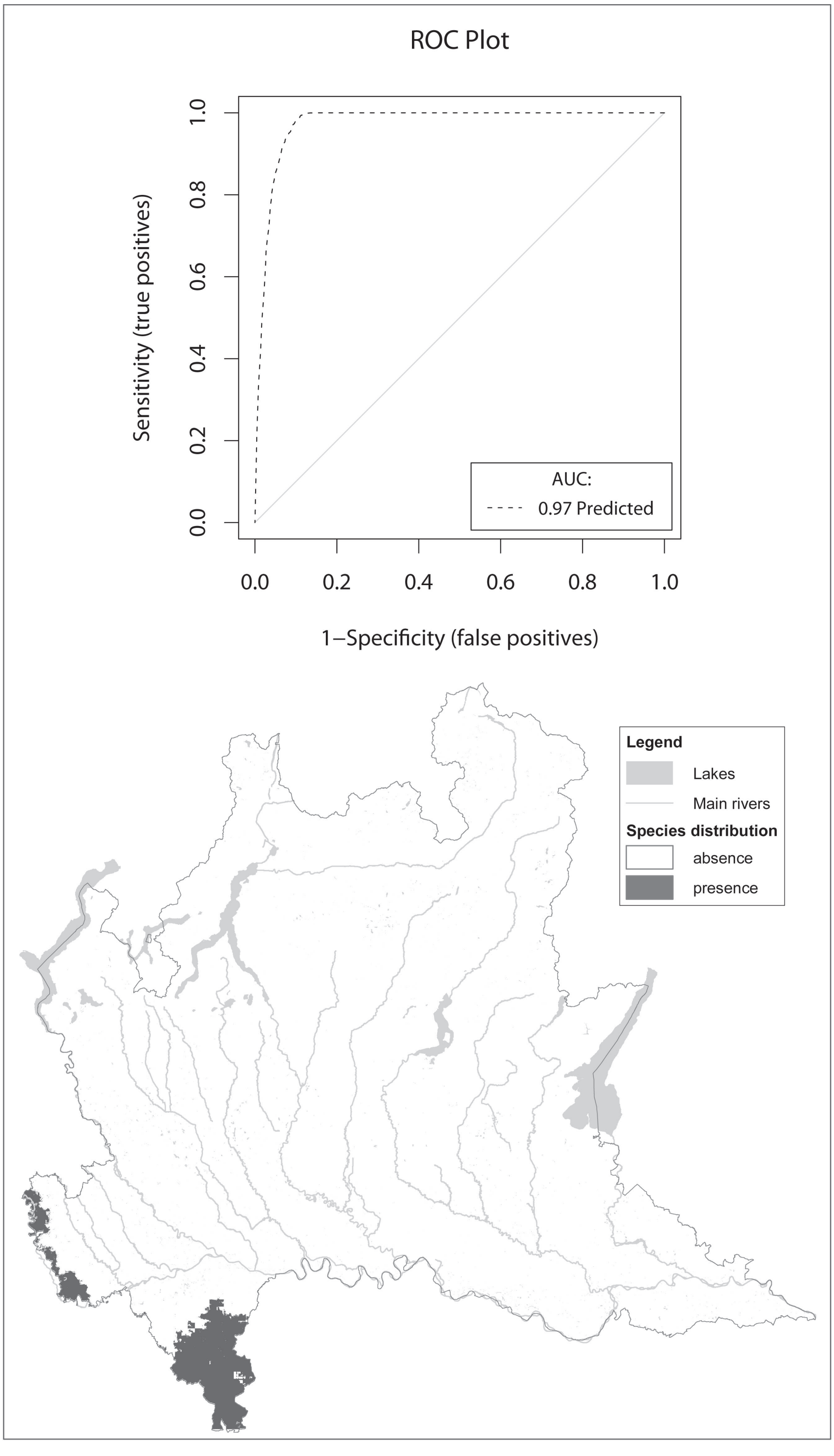


Figure A.85.: LRA for *Natrix maura* (viperine snake) ROC plot and potential distribution

Natrix natrix (grass snake)

Table A.60.: selected environmental variables and computed coefficients from stepwise analysis for *Natrix natrix* (grass snake)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
Intercept	$-8.584e+00$	$1.291e+00$	-6.647	$2.99e-11$
EV-CL15	$4.880e-02$	$6.927e-03$	7.045	$1.86e-12$
EV-CL18	$7.926e-03$	$8.861e-04$	8.945	$< 2e-16$
EV-CL01	$-1.506e-02$	$3.950e-03$	-3.813	0.000137
EV-CL07	$2.266e-02$	$3.879e-03$	5.843	$5.12e-09$
EV-D100	$1.694e-05$	$1.632e-06$	10.380	$< 2e-16$
EV-D103	$-1.492e-05$	$8.781e-06$	-1.699	0.08924
EV-D104	$1.516e-05$	$8.631e-06$	1.757	0.07895
EV-D105	$3.594e-05$	$4.610e-06$	7.797	$6.35e-15$
EV-D106	$-8.821e-06$	$3.332e-06$	-2.647	0.00812
EV-D108	$2.904e-05$	$9.277e-06$	3.131	0.001745
EV-D109	$1.972e-05$	$4.973e-06$	3.965	$7.32e-05$
EV-D110	$-2.925e-05$	$4.684e-06$	-6.245	$4.24e-10$
EV-D111	$-9.023e-06$	$3.332e-06$	-2.708	0.006763
EV-D112	$1.773e-05$	$6.443e-06$	2.752	0.005924
EV-D113	$1.008e-04$	$2.704e-05$	3.726	0.000194
EV-D114	$-1.310e-04$	$2.116e-05$	-6.189	$6.06e-10$
EV-D116	$-2.871e-05$	$6.014e-06$	-4.774	$1.80e-06$
EV-D117	$1.061e-05$	$1.391e-06$	7.628	$2.39e-14$
EV-D119	$3.966e-05$	$3.642e-06$	10.888	$< 2e-16$
EV-D120	$2.094e-05$	$1.270e-05$	1.649	0.099211
EV-D121	$-4.047e-05$	$2.537e-05$	-1.595	0.110652
EV-DHYD	$7.622e-03$	$1.343e-03$	5.674	$1.40e-08$
EV-DTM	$-2.749e-02$	$6.483e-03$	-4.240	$2.24e-05$

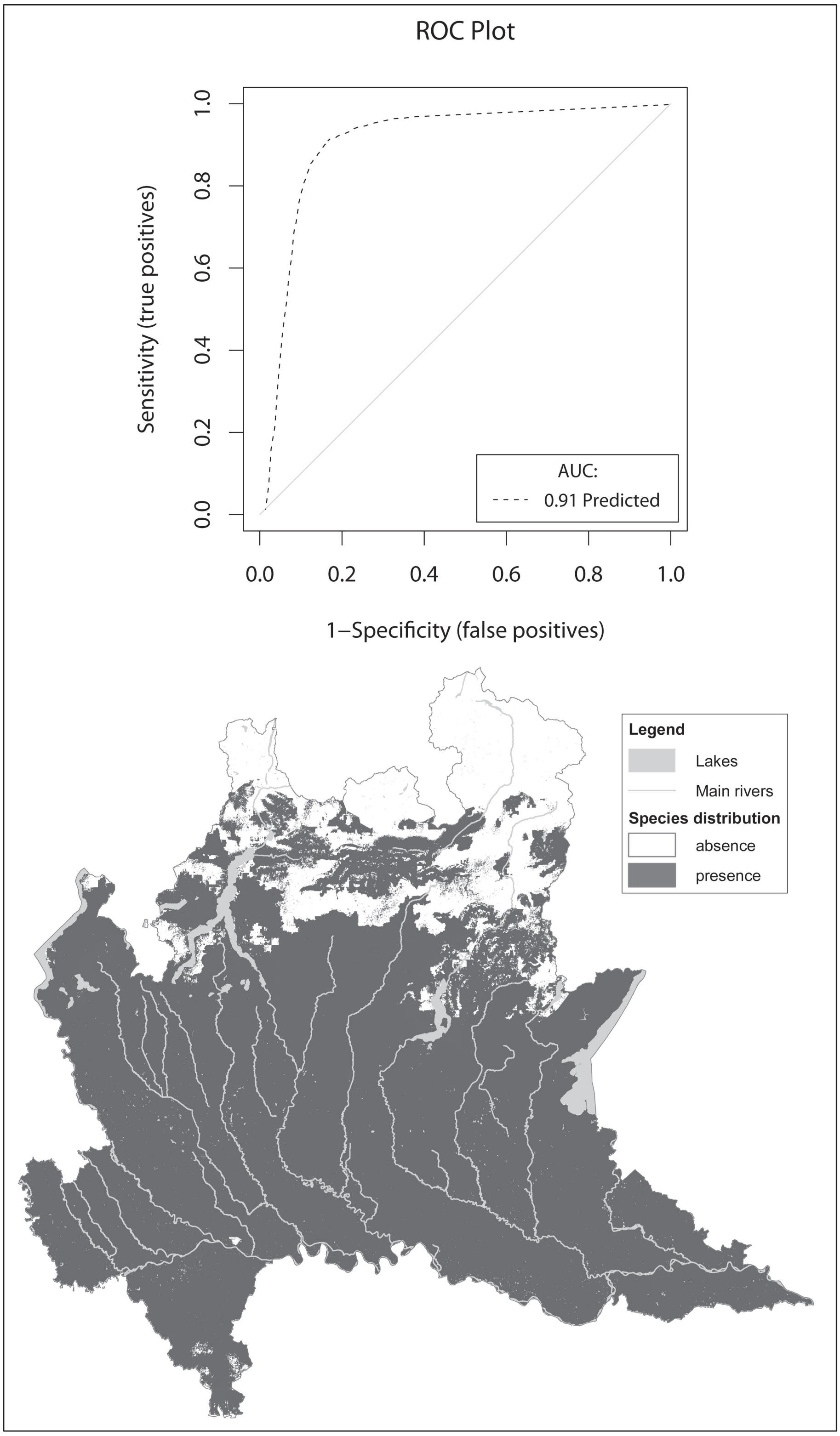


Figure A.86.: LRA for *Natrix natrix* (grass snake) ROC plot and potential distribution

Natrix tessellata (dice snake)

Table A.61.: selected environmental variables and computed coefficients from stepwise analysis for *Natrix tessellata* (dice snake)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr($>\|z\|$)**
Intercept	$9.828e+01$	$1.023e+01$	9.608	$< 2e-16$
EV-ASPECT	$-2.421e-04$	$1.422e-04$	-1.703	0.088526
EV-CL15	$2.156e-02$	$6.310e-03$	3.417	0.000634
EV-CL18	$1.651e-02$	$2.616e-03$	6.312	$2.76e-10$
EV-CL01	$-1.656e-02$	$2.757e-03$	-6.006	$1.90e-09$
EV-D100	$-3.228e-05$	$2.177e-06$	-14.829	$< 2e-16$
EV-D103	$-1.037e-04$	$1.409e-05$	-7.362	$1.82e-13$
EV-D104	$-1.092e-04$	$1.105e-05$	-9.887	$< 2e-16$
EV-D106	$-2.551e-05$	$3.007e-06$	-8.482	$< 2e-16$
EV-D107	$1.206e-05$	$3.192e-06$	3.779	0.000158
EV-D108	$1.924e-05$	$9.402e-06$	2.046	0.040747
EV-D109	$4.533e-05$	$5.494e-06$	8.251	$< 2e-16$
EV-D110	$-6.473e-05$	$4.768e-06$	-13.576	$< 2e-16$
EV-D111	$-5.907e-05$	$2.885e-06$	-20.475	$< 2e-16$
EV-D112	$-3.434e-05$	$7.009e-06$	-4.900	$9.60e-07$
EV-D113	$1.569e-04$	$2.367e-05$	6.630	$3.36e-11$
EV-D115	$-3.037e-05$	$1.060e-05$	-2.865	0.004164
EV-D116	$-6.004e-05$	$7.452e-06$	-8.057	$7.85e-16$
EV-D117	$5.636e-06$	$1.298e-06$	4.343	$1.41e-05$
EV-D119	$8.833e-05$	$3.312e-06$	26.671	$< 2e-16$
EV-D120	$-8.670e-05$	$1.533e-05$	-5.655	$1.56e-08$
EV-D122	$-5.304e-05$	$2.682e-05$	-1.977	0.048014
EV-D126	$3.499e-06$	$1.209e-06$	2.894	0.003805
EV-DTM	$-6.538e-04$	$3.965e-04$	-1.649	0.099145

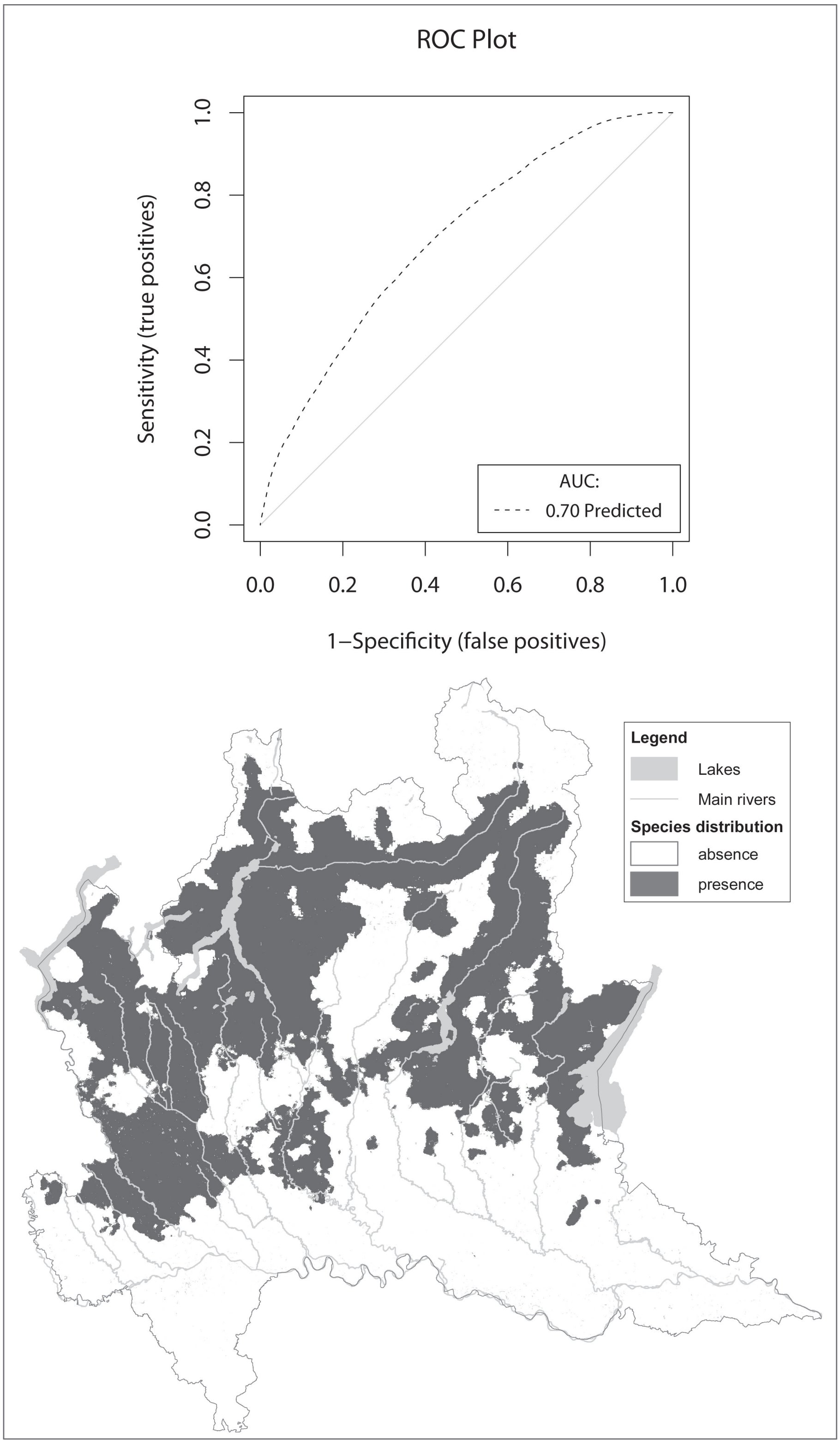


Figure A.87.: LRA for *Natrix tessellata* (dice snake) ROC plot and potential distribution

Vipera aspis (asp viper)

Table A.62.: selected environmental variables and computed coefficients from stepwise analysis for *Vipera aspis* (asp viper)

Regressor	Estimated coeff.	Std. Error	z value	Pr(> |z|)
Intercept	-3.314e + 00	9.506e − 01	-3.486	0.000491
EV-SPECT	-3.266e − 04	1.592e − 04	-2.052	0.040137
EV-CL15	-2.504e − 02	6.975e − 03	-3.590	0.000331
EV-CL18	8.381e − 03	8.321e − 04	10.072	< 2e − 16
EV-CL01	4.034e − 03	2.147e − 03	1.878	0.060327
EV-D100	8.058e − 06	1.479e − 06	5.449	5.06e-08
EV-D103	-2.969e − 04	1.713e − 05	-17.328	< 2e − 16
EV-D104	2.136e − 05	1.100e − 05	1.941	0.052285
EV-D105	2.912e − 05	4.951e − 06	5.883	4.03e − 09
EV-D106	-5.233e − 05	4.704e − 06	-11.124	< 2e − 16
EV-D108	9.128e − 05	1.037e − 05	8.801	< 2e − 16
EV-D109	1.869e − 05	5.394e − 06	3.465	0.000530
EV-D110	-7.911e − 05	4.974e − 06	-15.904	< 2e − 16
EV-D111	-1.086e − 04	3.548e − 06	-30.609	< 2e − 16
EV-D112	1.395e − 05	7.052e − 06	1.978	0.047976
EV-D113	9.919e − 05	2.950e − 05	3.362	0.000773
EV-D114	-2.676e − 04	3.193e − 05	-8.381	< 2e − 16
EV-D115	-8.908e − 05	1.005e − 05	-8.860	< 2e − 16
EV-D116	1.595e − 05	7.893e − 06	2.021	0.043308
EV-D117	1.069e − 05	1.318e − 06	8.111	5.03e − 16
EV-D119	8.311e − 05	4.031e − 06	20.618	< 2e − 16
EV-D120	-9.909e − 05	1.562e − 05	-6.345	2.22e-10
EV-D121	-8.606e − 05	3.246e − 05	-2.651	0.008027
EV-D122	-1.237e − 04	3.387e − 05	-3.651	0.000261
EV-D123	3.799e − 05	9.492e − 06	4.003	6.26e − 05
EV-D126	2.383e − 05	1.523e − 06	15.641	< 2e − 16
EV-DYDR	-8.686e − 05	1.545e − 05	-5.622	1.88e − 08
EV-DINF	7.956e − 05	3.581e − 05	2.222	0.026295

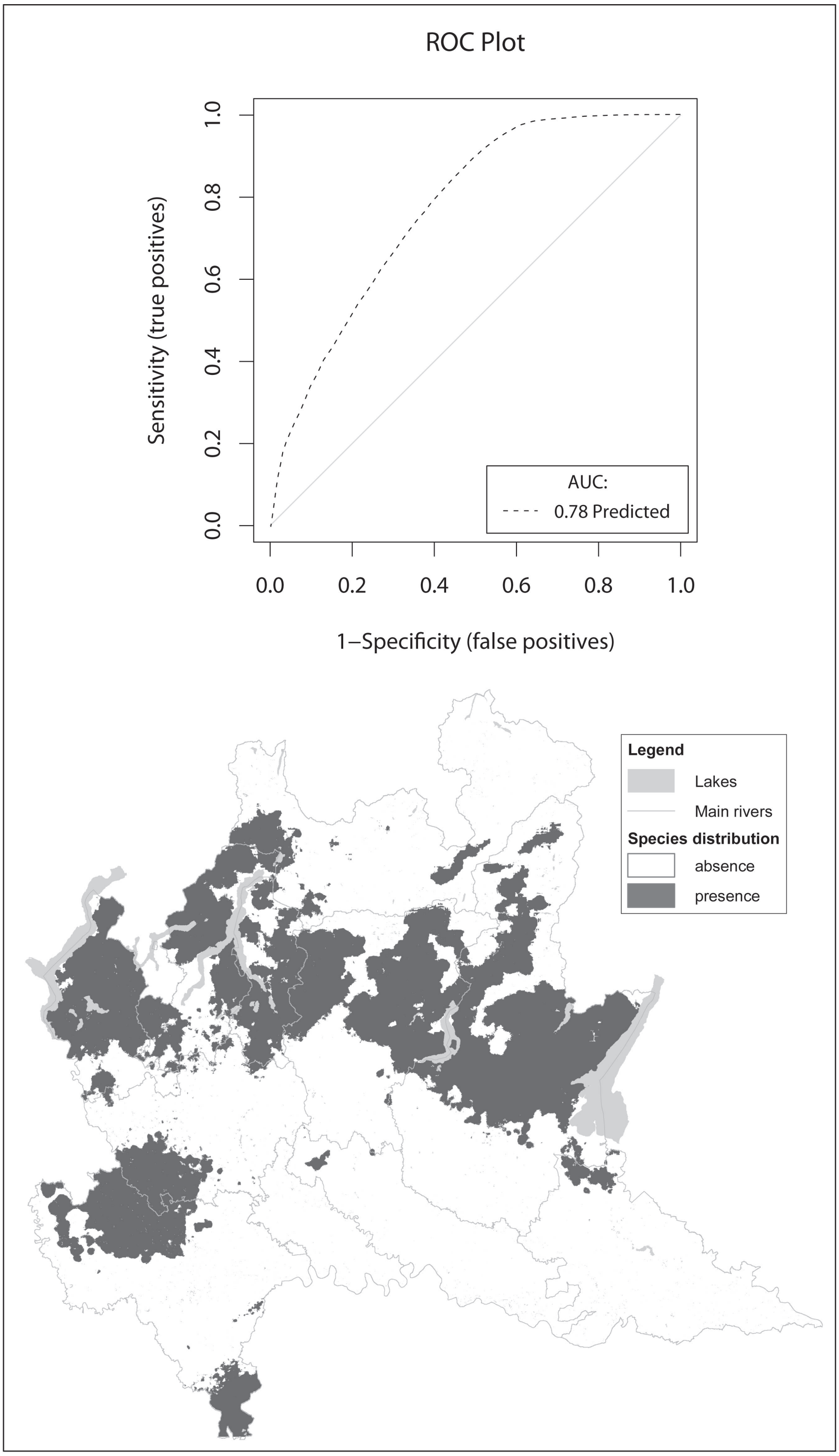


Figure A.88.: LRA for *Vipera aspis* (asp viper) ROC plot and potential distribution

Vipera berus (adder)

Table A.63.: selected environmental variables and computed coefficients from stepwise analysis for *Vipera berus* (adder)

Regressor	**Estimated coeff.**	**Std. Error**	**z value**	**Pr(> |z|)**
Intercept	1.138e+00	7.849e-01	1.450	0.147074
EV-CL15	-3.167e-03	1.282e-03	-2.470	0.013522
EV-CL18	6.527e-03	3.423e-03	1.907	0.056564
EV-D100	-8.235e-05	6.622e-06	-12.436	< 2e-16
EV-D103	-8.216e-05	2.160e-05	-3.804	0.000142
EV-D104	3.035e-05	1.327e-05	2.287	0.022180
EV-D105	-7.195e-05	2.199e-05	-3.272	0.001069
EV-D106	-7.733e-05	2.161e-05	-3.579	0.000345
EV-D107	3.153e-05	6.706e-06	4.702	2.57e-06
EV-D108	1.046e-04	9.967e-06	10.492	< 2e-16
EV-D109	3.255e-05	6.420e-06	5.070	3.98e-07
EV-D110	3.212e-05	7.567e-06	4.245	2.19e-05
EV-D112	-6.590e-05	1.213e-05	-5.432	5.56e-08
EV-D117	7.511e-06	4.072e-06	1.844	0.065118
EV-D119	-1.812e-04	2.854e-05	-6.351	2.14e-10
EV-D120	3.944e-05	1.524e-05	2.588	0.009658
EV-D122	-1.522e-04	2.728e-05	-5.579	2.42e-08
EV-D123	-1.348e-04	1.079e-05	-12.495	< 2e-16
EV-D124	1.850e-05	1.072e-05	1.726	0.084300
EV-D125	2.670e-05	4.743e-06	5.630	1.80e-08
EV-D126	-8.067e-06	3.264e-06	-2.472	0.013452
EV-D127	6.731e-05	1.039e-05	6.479	9.24e-11
EV-DHYD	-7.430e-05	4.265e-05	-1.742	0.081473
EV-DINF	1.669e-04	2.937e-05	5.680	1.35e-08
EV-DTM	5.550e-04	1.826e-04	3.040	0.002367

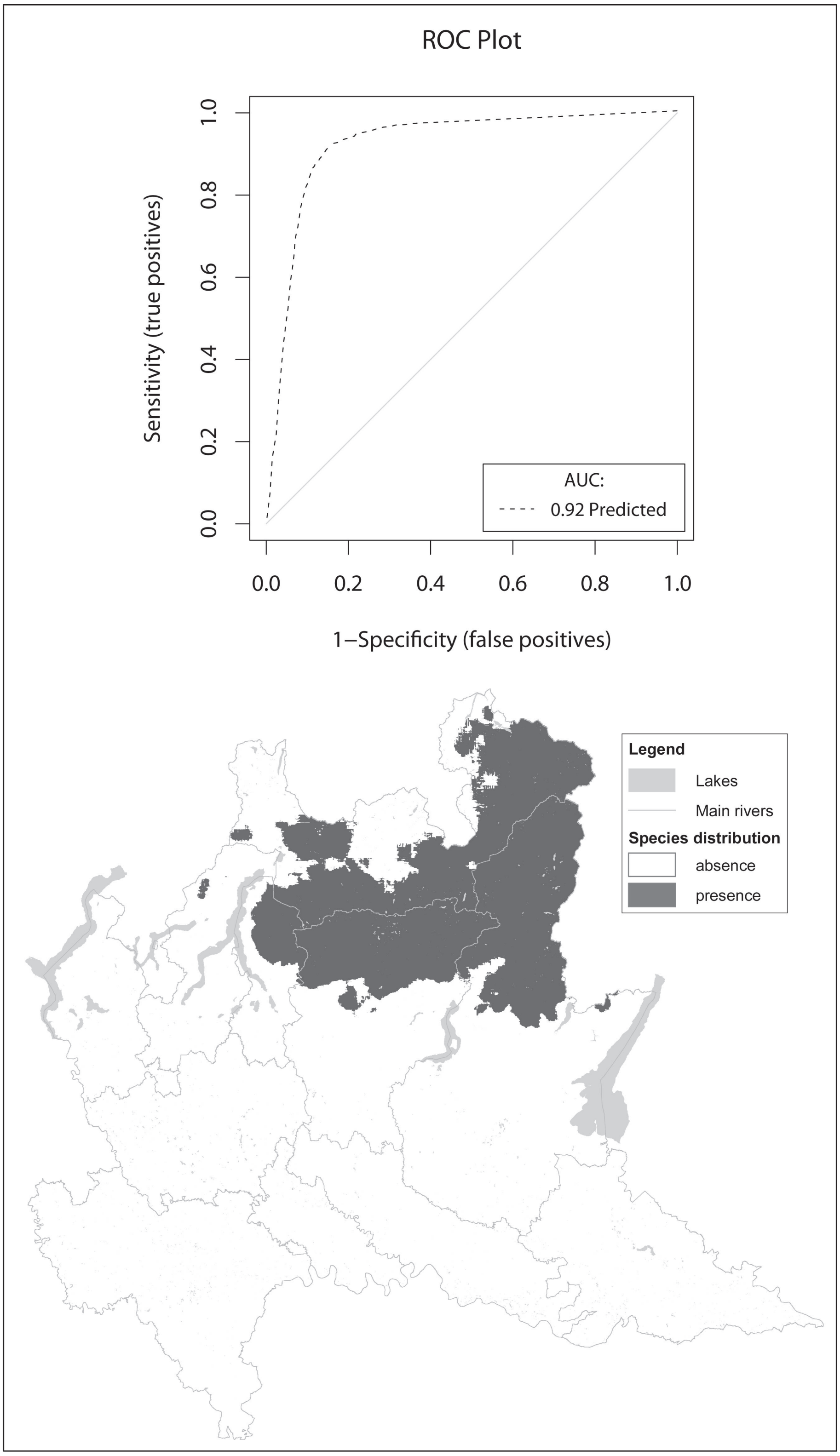


Figure A.89.: LRA for *Vipera berus* (adder) ROC plot and potential distribution

B. SOS scripts

B.0.6. SOS scripts for statistical and spatial modeling

SPECIES.SPATIAL.MODELS.R

```
##########################################################
# Basic script to launch functions for modeling species  #
# presence through different techniques:                 #
# 1)Generalized Linear Models (GLM)                      #
# 2)Environment Niche factor Analysis (ENFA)             #
# 3)Flexible Discriminant Analysis with method BRUTO     #
#                                                        #
# scripts sourced for the analysis:                      #
# 1) CALC_LOGISTIC.R computes logistic regression        #
#    analysis                                            #
# 2) CALC_ENFA.R  computes ENFA                          #
# 3) CALC_BRUTO.R computes flexible discriminant         #
#    analysis with method bruto                          #
# 4) OUT_LOGISTIC_GRASS.R  outputs results of logistic   #
#    regression analysis                                 #
# 5) OUT_ENFA.R outputs results of ENFA                  #
# 6) OUT_BRUTO.R outputs results of flexible             #
#    discriminant analysis with method bruto             #
# 7) MAKE_GRASS_SH.R creates a bash script for GRASS     #
#    in order to compute GRASS raster layer of predicted #
#    distribution. If ROC analysis is performed, the     #
#    scripts also compute classified rasters.            #
#                                                        #
# requires FOREIGN, ADEHABITAT, PRESENCEABSENCE, MDA     #
##########################################################
# version 1.0                                            #
# created 01/06/2008                                     #
# revision history:                                      #
#                                                        #
#                                                        #
##########################################################
#                                                        #
#  Copyright (C) 2008  Monica Carro                      #
#                                                        #
#  This program is free software; you can redistribute   #
#  it and/or modify it under the terms of the GNU        #
#  General Public License as published by the Free       #
#  Software Foundation; either version 2 of the          #
#  License or (at your option) any later version.        #
```

```
#                                                      #
#  This program is distributed in the hope that it     #
#  will be useful, but WITHOUT ANY WARRANTY; without   #
#  even the implied warranty of MERCHANTABILITY or     #
#  FITNESS FOR A PARTICULAR PURPOSE.  See the GNU      #
#  General Public License for more details.            #
#                                                      #
#  A copy of the GNU General Public License is availabl #
#  at http://www.gnu.org/licenses/gpl.txt, or can be   #
#  requested to the Free Software Foundation, Inc., 51 #
#  Franklin Street, Fifth Floor, Boston, MA 02110-1301 #
#  USA.                                                #
#                                                      #
########################################################
#DATA
#data for GLM analysis
#1) GRASS point vector layers provided with attribute
#table. The table will be red directly by R through
#spgrass6. The table should contain the value of each
#environmental variable at each point location and should
#be prepared with the GRASS script v.whatrast.sh.
#data for ENFA analysis
#1) shapefile(s) or table(s) in dbf file formatof presence
#   or presence/absence point data . The name of the
#   shapefile/dbf table must be the same of the keyword
#   referring to the specie (es. ruprup.dbf)
#   IMPORTANT: shapefile must contain a column 'pres'
#   specifyng the nature of the point (0= absence,
#   1=presence), a column with the X coordinate and a
#   column with the Y coordinate.
#2) grids of environmental variables and grid-pack(s):
#   ascii files listing omogeneous grids according to
#   the radius of the focal function applied.
#data for FDA-BRUTO analysis
#1) GRASS point vector layers provided with attribute
#table. The table will be red directly by R through
#spgrass6 The table should contain the value of each
#environmental variable at each point location and
#should be prepared with the GRASS script v.whatrast.sh.
#2) GRASS point vector layer resulting from the GRASS
#script r.whatrast.many.sh. Points should cover all the
#study area as a grid at the resolution set for the GIS
#analysis.The table completely describes environmental
#conditions in the whole study area
#CONFIGURATION FILE
#The computation and output of presence models requires a
#configuration file, in ASCII format; each row of the
#list contains, separeted by ",":
#1) keyword referring to the specie for wich presence
#   models are computed;
```

```
#2) name of the grid-pack (followed by the extension of
#   the file)containg environmental variables for the
#   analysis environmental;
#3) Keyword(s)of the modeling technique(s) choosed for
#   presence prediction. If more than 1 model is required,
#   keywords must be separated by space.
#4) Keyword(s) of the output(s) required at the and of the
#   analysis. If more than 1 model is required, keywords
#   must be separated by space.
#KEYWORDS:
#1) species are indicated with tre three first letter of
#   the genera name, followed by the three first letters
#   of the specie name. For example the specie
#   "Rupicapra rupicapra" is referred to as "ruprup"
#2) keyword for ENFA model: enfa
#   keyword for GLM model: logistic
# keyword for FDA-BRUTO model: bruto
#3) the following keyword can be put in the list according
#   to the kind of model applied and to the output required.
#   It should be noticed that models performance is
#   estimated in both cases with the Reicer Operating
#   Characteristics (ROC) analysis. Three kinds of outputs
#   are available for each analysis: presence prediction
#   map, graphical plots in eps format, text outputs in
#   txt format.
#   "enfa_grid": keyword to output presence grid from ENFA
#                analysis (output as asciigrid to import in
#                GRASS GIS)
#   "logistic_grid" keyword to output presence grid from
#                   GLM analysis (output as script in aml
#                   language to run in ArcInfo GIS to
#                   compute final grid)
#  "bruto_grid": keyword to output presence grid from ENFA
#                 analysis (output as asciigrid to import in
#                 GRASS GIS)
#   "enfa_summary": keyword to output text summary of ENFA
#                   analysis
#   "logistic_summary": keyword to output text summary of the
#                       results of GLM analysis
#   "bruto_summary": keyword to output text summary of the
#                    results of FDA-BRUTO analysis
#   "logistic_AIC": keyword to output Akaike's Information
#                   Criterion
#   "enfa_hist":keyword to output ENFA histogram of
#               marginality and specialisation
#   "enfa_plotsummary": keyword to output plots that
#                       summarize ENFA analysis results
#   "enfa_ROCauc": keyword to output the Area Under the
#                  Curve (AUC) value from ROC analysis for
#                  ENFA
```

```
#   "logistic_ROCauc": keyword to output the Area Under the
#                      Curve (AUC) value from ROC analysis
#                      for GLM analysis
#   "enfa_ROCsummary": keyword to output summary results
#                      from ROC analysis for ENFA
#   "logistic_ROCsummary": keyword to output summary
#                          results from ROC analysis for
#                          ENFA model
#   "enfa_ROCsummary": keyword to output summary results
#                      from ROC analysis for GLM model
#   "enfa_ROCcutoffs": keyword to output cutoffs from ROC
#                      analysis for ENFA model
#   "logistic_ROCcutoffs": keyword to output cutoffs from
#                          ROC analysis for GLM model
#   "enfa_ROCplot": keywords to output plot of ROC curve
#                   for ENFA model
#   "logistic_ROCplot": keywords to output plot of ROC
#                       curve for GLM model
#  "bruto_contingency": keywords to output contingency
#                        table for FDA-BRUTO analysis
#
#path to data (comprise configuration file)
data_path <- ".../SOS.rc"
#path to R script
script_path <- ".../SOS/R"
#paths for outputs (text and plots)
output_plots_path <- ".../SOS.rs/plots"
output_text_path <- ".../SOS.rs/text"
output_maps_path <- ".../SOS.rs/maps"
#name of configuration file
config_file <- "..."

setwd(data_path)
require(foreign)
#opens the configuration file and reads the table, it
#assigns names to columns
tbl_conf <- read.table(config_file, header=FALSE, sep=",",
            as.is=TRUE, col.names=c("specie","grid_pack",
            "model","output"))
#creates a list to accumulate results for each model and for
#each specie
results <- list()
#For each row of the tables reads values from each column
#("specie", 'grid_pack", "models" and "output")
for (row in 1:nrow(tbl_conf)) {
  rec <- tbl_conf[row,]
  print(paste(" now doing",rec$specie))
  extract the current specie
  specie <- rec$specie
  print(specie)
```

```
# extracts the keyword corresponding to models and outputs
models <- strsplit(rec$model, split=" ")[[1]]
outputs <- strsplit(rec$output, split=" ")[[1]]
print(models)
print(outputs)
#creates a list to accumulate results from each model, for
#the current specie
results.specie <- list()
# if the keyword "logistic" is found in the outputs string,
# load scripts for logistic regression analysis
if("logistic" %in% models) {
  print ("make logistic")
  source(paste(script_path,"CALC_LOGISTIC1.R",sep='/'))
  source(paste(script_path,"OUT_LOGISTIC_GRASS.R",sep='/'))
  logistic.index <- agrep("logistic", outputs,
                  ignore.case=TRUE)
# extracts and print keywords corresponding only to
  logistic outputs
  print(paste("output =", outputs[logistic.index],
        sep = " "))
  logistic.outputs <- outputs[logistic.index]
# the function "make_glm" (in "CALC_LOGISTIC.R") is
  called, which performs
# logistic regression analysis; results are accumulated in
  the list
# "results.specie" under the element "logistic"
  results.specie[["logistic"]] <- make_glm(specie)
# the function "out_glm" (in "OUT_LOGISTIC.R") is called,
# which prepares outputs from the results, following the
#keywords
  out.glm(specie, logistic.outputs,output_plots_path,
  output_text_path)
  }
# if the keyword "ENFA" is found in the outputs string,
  load packages and scripts for ENFA
if("enfa" %in% models){
  print ("make enfa")
  require(adehabitat)
  source(paste(script_path,"CALC_ENFA.R",sep='/'))
  source(paste(script_path,"OUT_ENFA.R",sep='/'))
  enfa.index <- agrep("enfa", outputs, ignore.case=TRUE)
  #extracts and print keywords corresponding only to enfa
  #outputs
  print(paste("output =", outputs[enfa.index], sep = " "))
  enfa.outputs <- outputs[enfa.index]
  # extract the name of the grid-pack of environmental
  #variables for the current specie
  gridpack<- rec$grid_pack
  # the function "make.enfa" (in "CALC_ENFA.R") is called,
  # which performs ENFA; results are accumulated in the
```

```
    # list "results.specie" under the element "enfa"
    results.specie[["enfa"]] <- make.enfa(specie, gridpack)
    # the function "out_enfa" (in "CALC_ENFA.R") is called
    # which prepares outputs from the results, following
    #the keywords
    out.enfa(specie, enfa.outputs)
    }
  # if the keyword "BRUTO" is found in the outputs string, load
  # packages and scripts for BRUTO
  if("bruto" %in% models){
    print ("make bruto")
    require(mda)
    source(paste(script_path,"CALC_BRUTO.R",sep='/'))
source(paste(script_path,"OUT_BRUTO.R",sep='/'))
bruto.index <- agrep("bruto", outputs,
ignore.case=TRUE)
    # extracts and print keywords corresponding only to bruto
    # outputs
    print(paste("output =",outputs[bruto.index], sep = " "))
    bruto.outputs <- outputs[bruto.index]
    # the function "make_bruto" (in "CALC_BRUTO.R") is
    #called, which performs
  # flexible discriminant analysis; results are accumulated
  # in the list "results.specie" under the element "bruto"
    results.specie[["bruto"]] <- make_bruto(specie)
  # the function "out_bruto" (in "OUT_BRUTO.R") is called,
  # which prepares outputs from the results, following the
  # keywords.
    out.bruto(specie, bruto.outputs,output_plots_path,
    output_text_path,output_maps_path)
    }
  # results from each model for the current specie are
  # accumulated in the final list "results" in the element
  # called with the keyword of the current specie
  results[[specie]] <- results.specie
  # the temporary list to accumulate results for the current
  # specie is deleted since results are transferred to the
  # list
  # "results"
  rm(results.specie)
  }
```

CALC.LOGISTIC.R

```
###########################################################
###########################################################
# Script to compute Logistic regression analysis(GLM)    #
#                                                         #
# requires FOREIGN                                        #
# has to be sourced by SPECIES_SPATIAL_MODELS.R          #
```

```
##########################################################
# version 1.0                                            #
# created 1/06/2008                                      #
# revision history:                                      #
#                                                        #
##########################################################
#                                                        #
# Copyright (C) 2008  Monica Carro                       #
#                                                        #
# This program is free software; you can redistribute it #
# and/or modify it under the terms of the GNU General    #
# Public License as published by the Free Software       #
# Foundation; either version 2 of the License, or (at    #
# your option) any later version.                        #
#                                                        #
# This program is distributed in the hope that it will   #
# be useful, but WITHOUT ANY WARRANTY; without even the  #
# implied warranty of MERCHANTABILITY or FITNESS FOR A   #
# PARTICULAR PURPOSE.  See the GNU General Public        #
# License for more details.                              #
#                                                        #
# A copy of the GNU General Public License is available  #
# at http://www.gnu.org/licenses/gpl.txt, or can be      #
# requested to the Free Software Foundation, Inc.,       #
# 51 Franklin Street, Fifth Floor,  Boston,              #
# MA 02110-1301 USA.                                     #
#                                                        #
##########################################################
# Defines function performing logistic regression analysis
make_glm <- function (specie){
# prepares a list to accumulate results
result <- list()
require(foreign)
require(spgrass6)
library(spgrass6)
data <- readVECT6(specie, plugin=FALSE)
#print(data)
#attach(data)
presence <- names(data)[2]
print(presence)
# column names corrisponding to environmental variables are
# collapsed with + in
# order to build a formula
environment <- paste(names(data)[5:(ncol(data))],
 collapse="+")
print(environment)
#detach(data)

# buldS a formula object based on extracted columns names
glm.formula <- as.formula(paste(presence,"~",environment,
```

```
 sep=""))
print(glm.formula)
# calles the "glm" function  to fit generalized linear
# models, specified by giving a symbolic description of
# the linear predictor
logit.out <- glm(glm.formula,family=binomial(link="logit"),
 data)
print('logit made')
# calles the "step" function which performs stepwise
# selection of the best combination of predictors, based
# on  AUC (Akaike's Information Criterion)
step.out <- step(logit.out, direction = "both", trace = 1,
  keep = NULL, steps = 1000000, k = 2)
# the analysis is completed, results from "glm" and "step"
# functions are copied in the list "results" respectevly
# in the elements "glm" and "step"
result[["glm"]] <- logit.out
result[["step"]] <- step.out
# the list "result" is returned invisibly (and so copied
# to the list "results.specie" in the element "logistic")
invisible(result)
}
```

CALC.ENFA.R

```
##########################################################
# Script for Environment Niche factor Analysis(ENFA)     #
#                                                        #
# requires FOREIGN, ADEHABITAT                           #
# has to be sourced by SPECIES.SPATIAL.MODELS.R #        #
##########################################################
# version 1.0                                            #
# created 15/07/2008                                     #
# revision history:                                      #
#                                                        #
##########################################################
#                                                        #
# Copyright (C) 2008  Monica Carro                       #
#                                                        #
# This program is free software; you can redistribute it #
# and/or modify it under the terms of the GNU General    #
# Public License as published by the Free Software       #
# Foundation; either version 2 of the License, or (at    #
# your option) any later version.                        #
#                                                        #
# This program is distributed in the hope that it will   #
# be useful, but WITHOUT ANY WARRANTY; without even the  #
# implied warranty of MERCHANTABILITY or FITNESS FOR A   #
# PARTICULAR PURPOSE.  See the GNU General Public        #
# License for more details.                              #
```

```
#                                                        #
# A copy of the GNU General Public License is available  #
# at http://www.gnu.org/licenses/gpl.txt, or can be      #
# requested to the Free Software Foundation, Inc.,       #
# 51 Franklin Street, Fifth Floor,  Boston,              #
# MA 02110-1301 USA.                                     #
#                                                        #
##########################################################
# Definition of function performing ENFA
make.enfa <- function(specie,gridpack){
# prepares a list to accumulate ENFA results
result <- list()
# prepares a list to accumulate imported ascii files
asc.list <- list()

# definition of some useful function; reduce.pca performs
# PCA and keeps variabless the number of input
# environmental variables
reduce.pca <- function(pcadata, tresh_enter=0.5,
 verbose=TRUE) {
  ret <- list()
  # do a pca on 1st axis only
  dudi <- dudi.pca(pcadata, scannf=FALSE, nf=1)
  # get variance explanied by 1st axis (for each component
  # the formula to compute
  # the percentage of explained variance is (eig * 100)/Nvar
  # where eig is the eigenvalue for the firts component
  # and Nvar is the number
  # of variables (We know from the principal components
  # theory that
  # 1)the number of variables equals the number of
  #   components
  # 2)Since the variance of each component is equal to 1
  #   the total variance 1* NComp(=Nvar))
  ret[['varexp']] <- (dudi$eig[1] * 100 / dudi$rank)
  if (verbose) print(paste("  explained variance:",
    ret[['varexp']]))
  # drop uncorrelated vars; this is done by keeping only
  # variables whose
  # coordinates of the 1st component (dudi$co$Comp1) are
  # above the treshold entered (tresh_enter)
  ret[['keep']] <- row.names(subset(dudi$co,
   abs(Comp1) >= tresh_enter))
  if (verbose) print(paste("  will keep:", ret[['keep']]))
  # reduces data entered for PCA by selecting only
  # correlated variables
  ret[['tab']] <- subset(pcadata,select=ret[['keep']])
  # returns list with 1)percentage of explained variance
  # from the 1st axis,
  # 3)string with the names of kept variables;
```

```
  # 4)reduction of dataframe pcadata$tab
  # containig only kept variables. PCA will be performed on
  # this new selected dataframe
  invisible(ret)
}

# step.pca is a function that calls iteratively reduce.pca,
# until explained variance doesn't get better than + 1%
step.pca <- function(pcadata) {
# pcadata is the result of data2enfa, namely is data2enfa$tab
  # some constants
  tresh_enter <- 0.5
  # stores variable names
  vars <- names(pcadata)
  #print(vars)
  p.old <- reduce.pca(pcadata, tresh_enter)
  p.new <- p.old
  v.delta <- 100
  while (v.delta >= 1) {
    p.new <- reduce.pca(p.old[['tab']], tresh_enter)
    v.delta <- p.new[['varexp']] - p.old[['varexp']]
    p.old <- p.new
  }
  invisible(p.new)
}
# reads grid-pack list of environmentale variables and
# creates the corresponding dataframe
grids.list <- read.table(gridpack, header=FALSE,
 col.names=c("GRID"), as.is=TRUE)
# imports ESRI ArcInfo ascii grids of environmental variables
# listed in each row of the dataframe
for (row in 1:nrow(grids.list)) {
  asc.list[[grids.list[row,]]]
   <- import.asc(paste(grids.list[row,], "asc", sep = "."))
  }
# prepares data for ENFA: converts a list of matrices of
# class asc into a data frame of class kasc.
maps <- as.kasc(asc.list)
# from 58 to 61 do not run:
#for (var in 1:length(maps)) {
# max <- max(maps[,var], na.rm = TRUE)
#  maps[,var] <- maps[,var] / max * 100
#  }
# removes the list of matrices of class asc; don't need it
# anymore
rm(asc.list)
gc()
# reads .dbf tables of Presence/Absence data
#(directly fron point ESRI shapefile)
# IMPORTANT: shapefile must contain a column 'pres'
```

```
# specifyng the nature of the point
# (0= absence, 1=presence), a column with the X coordinate
# and a column with the Y coordinate
data <- read.dbf(paste(paste("smp",specie,sep="_"),
 "dbf", sep="."))
# keeps only presence points since ENFA uses only presence
# data
locs <- subset(data, PRES == 1, select = c(xcoor, ycoor))
rm(data)
gc()

# prepares dat for ENFA, based on kasc maps and presence
# coordinates
dataenfa <- data2enfa(maps, locs)
print("prepared data for enfa")
# do the "stepwise" selection of correlated variables by
# calling function step.pca (and
# function reduce.pca) until the maximum percentage of
# explained variance of the
# first axis is reached
s <- step.pca(dataenfa$tab)
gc(verbose = TRUE)
# based on the kept variables from function reduce.pca
#(s$keep) rebuild a reduced gridpack
grids.list <- data.frame(s$keep,stringsAsFactors = FALSE)
names(grids.list) <- "GRID"

# imports teh reduced number of ascii grids
asc.list <- list()
for (row in 1:nrow(grids.list)) {
  asc.list[[grids.list[row,]]]
   <- import.asc(paste(grids.list[row,], "asc", sep = "."))
}
# re-builds kasc dataframe from a reduced list of matrices
# of class asc
maps <- as.kasc(asc.list)
rm(asc.list)
gc()

# prepares data for ENFA (reduced dataset)
dataenfa <- data2enfa(maps, locs)
print("prepared data for enfa")
gc()

# ENFA step 1: do a PCA
datapca <- dudi.pca(dataenfa$tab, scannf=FALSE, nf=2)
gc()
# ENFA step 2: do ENFA
enfa.out <- enfa(datapca, dataenfa$pr, scannf=FALSE)
print("enfa computed")
```

```
gc()
# prepares data for output; converts kasc object into
 #dataframe
# (necessary to produce output prediction grid)
maps.df <- kasc2df(maps)

# copies results and processed data to the list result
result[["output_enfa"]] <- enfa.out
result[["PCA"]] <- datapca
result[["maps"]] <- maps
result[["maps_index"]] <- maps.df$index
result[["locs"]] <- locs
result[["C1_Varexp"]] <- s$varexp
result[["KeptVariables"]] <- s$keep
# the list "result" is returned invisibly (and so copied
# to the list
# "results.specie" in the element "ENFA")
invisible(result)
}
```

CALC.BRUTO.R

```
##########################################################
# Script to compute FLEXIBLE DISCRIMINANT ANALYSIS with  #
# method "BRUTO"                                         #
#                                                        #
# requires FOREIGN                                       #
# has to be sourced by SPECIES.SPATIAL.MODELS.R          #
##########################################################
# version 1.0                                            #
# created 1/06/2008                                      #
# revision history:                                      #
#                                                        #
##########################################################
#                                                        #
#  Copyright (C) 2008  Monica Carro                      #
#                                                        #
#  This program is free software; you can redistribute   #
#  it and/or modify it under the terms of the GNU        #
#  General Public License as published by the Free       #
#  Software Foundation; either version 2 of the          #
#  License, or (at your option) any later version.       #
#                                                        #
#  This program is distributed in the hope that it will  #
#  be useful, but WITHOUT ANY WARRANTY; without even the #
#  implied warranty of merchantability or fitness for a  #
#  paricular purpose.  See the GNU General Public        #
#  License for more details.                             #
#                                                        #
#  A copy of the GNU General Public License is available #
```

```
#  at http://www.gnu.org/licenses/gpl.txt, or can be     #
#  requested to the Free Software Foundation, Inc., 51   #
#  Franklin Street, Fifth Floor, Boston,                 #
#  MA 02110-1301 USA                                     #
#                                                        #
##########################################################

# Defines function performing logistic regression analysis
make_bruto <- function (specie){
# prepares a list to accumulate results
result <- list()
require(foreign)
# opens the .dbf table resulting from the GRASS SOS script
# "v.whatrast.many.sh".
# The name of this file is built with the keyword of the
# specie, preceeded by  "smp_". (for example
# smp_ruprup.dbf)
# The script samples a GRID pack, starting from a point
# layer of presence ad absence points.
# This part of the analysis which prepares date for bruto
# analysis is performed in GRASS GIS.
specie.smp <- paste(paste("smp",specie,sep="_"),"dbf", sep=".")
data.specie <- read.dbf(specie.smp)
print(specie.smp)
# opens dbf table resulting from GRASS SOS script
# r.whatrast.many.sh which samples a list of specified grid
# at each centroid of the grid cells. This part of the
# analysis which prepares date for bruto analysis is
# performed in GRASS GIS.
data.area <- read.dbf(paste("smp_area.dbf"))
print("red data.area")
# joins data.area and specie.area data in order to copy
# presence attributes to data.area

# train.data <-data.frame()
# for(r in nrow(data.specie)){
# row.specie <- data.specie[r, ]
# first.sel <- data.specie[r,5]
# first.col <- names(data.area)[5]
# sel.area <- data.area[data.area$first.col==first.sel, ]
# for(c in names(data)[5:length(data.specie)]){
# val.specie <- data.specie[r,data.specie$c]
# sel.area <- sel.area[sel.area$c==val.specie, ]
# }
# train.data[r, ] <- sel.area
#}


# extraxts fields for PCA analysis
data.pca <- data.area
```

```
# col<- as.factor(names(data.specie[5:(length(data.specie)-1)]))
# data.pca <- subset(data.pca, select=col)


# extracts rows without NA values
funApply = function(x) {
 drop = apply(is.na(x), 1, any)
 x[!drop, ]
}
data.pca <- funApply(data.pca)
print("NA values removed")
cat<-data.pca$CAT
data.pca<-data.pca[-1]

# Performs PCA, first on the first axis, then considers
# axis
# with eigenvalues >1
pca <- dudi.pca(data.pca,scannf=FALSE, nf=1)
eig <- pca$eig
eig.keep <- subset(eig, eig>1)
axes.keep <- length(eig.keep)
pca <- dudi.pca(data.pca,scannf=FALSE, nf=axes.keep)
print("PCA done")

# keep scores from PCA analysis and uses them as new
# variables
# (linear combination of old variables
scores.pca <- pca$l1
scores.pca$CAT<- cat
print("built new dataframe with scores values")

for (c in data.specie$GRID_CAT) {
scores.pca[scores.pca$CAT==c, "PRES"]
 <- data.specie[data.specie$GRID_CAT==c, "PRES"]
}

presence <- names(scores.pca)[length(scores.pca)]
scores <- paste(names(scores.pca)[1:(length(scores.pca)-2)],
         collapse="+")
print(presence)
print(scores)
#builds formula for fda
fda.formula = as.formula(paste(presence,"~",scores,sep=""))
print(fda.formula)

#Creates new training setaxes.keep
train.data <- subset(scores.pca, PRES != "NA")
#performs fda with method bruto on training set
bruto.fda <- fda(fda.formula,data=train.data,method=bruto)
print("fda computed")
```

```
bruto.fitted <- predict(bruto.fda,scores.pca)
print("predict done")
bruto.confusion <- confusion(bruto.fda,train.data)
print("confusion table computed")
result[["bruto.fda"]] <- bruto.fda
result[["bruto.fitted"]] <- bruto.fitted
result[["bruto.confusion"]] <-bruto.confusion
print("calc_bruto done")
invisible(result)
}
```

CALC.HSI.sh

```
#!/bin/sh
##########################################################
#                                                        #
# MODULE: SOS.HSI.SH v.1.1 for GRASS 6.4                 #
# PACKAGE: SOS - Species Open Spreader                   #
# AUTHOR(S): Damiano G. Preatoni Monica Carro            #
# PURPOSE: Calculates an Habitat Suitability Index (HSI) #
#          raster map, given some environmental variable #
#          rasters along with their reclassification     #
#          rules.                                        #
# COPYRIGHT: (C) 2009 by D.G. Preatoni & Monica carro    #
#             and the GRASS Development Team.            #
#            This program is free software under the GNU #
#            General Public License (>=v2). Read the file#
#            COPYING that comes with GRASS for details.  #
#                                                        #
##########################################################
#                                                        #
# Copyright (C) 2008  Monica Carro                       #
#                                                        #
# This program is free software; you can redistribute it #
# and/or modify it under the terms of the GNU General    #
# Public License as published by the Free Software       #
# Foundation; either version 2 of the License, or (at    #
# your option) any later version.                        #
#                                                        #
# This program is distributed in the hope that it will   #
# be useful, but WITHOUT ANY WARRANTY; without even the  #
# implied warranty of MERCHANTABILITY or FITNESS FOR A   #
# PARTICULAR PURPOSE.  See the GNU General Public        #
# License for more details.                              #
#                                                        #
# A copy of the GNU General Public License is available  #
# at http://www.gnu.org/licenses/gpl.txt, or can be      #
# requested to the Free Software Foundation, Inc.,       #
# 51 Franklin Street, Fifth Floor,  Boston,              #
# MA 02110-1301 USA.                                     #
```

```
#                                                          #
############################################################
%module
%description: r.SOS.rules.sh
%end
%option
%key: output
%type: string
%required: yes
%gisprompt: new,cell,raster
%description: Name of the output habitat suitability
 raster
%end
%option
%key: rastlist
%type: string
%gisprompt: old_file,file,input
%description: Text file with the list of rasters to export
%required: yes
%end
%option
%key: method
%answer: mult
%description: Method used to calculate final HSI
 (multiplicative or additive)
%type: string
%options: mult,add
%gisprompt: new_file,file,input
%end


if  [ -z "$GISBASE" ]
then
echo ""
echo "You must be in GRASS GIS to run this program"
echo ""
exit 1
fi

if   [ "$1" != "@ARGS_PARSED@" ]
then
exec g.parser "$0" "$@"
fi

# Check for required arguments
if [ -z "$GIS_OPT_RASTLIST" ]
then
echo "ERROR: Missing list of raster!"
echo "Please enter a valid text file"
exit 1
```

```
fi

eval `g.gisenv`
: ${GISBASE?} ${GISDBASE?} ${LOCATION_NAME?} ${MAPSET?}

SOS_RC_DIR=$GISDBASE/$LOCATION_NAME/$MAPSET/SOS.rc

MAPCALC_STRING="("
NUM_VARS=0
FIRST=TRUE

# recode input rasters as partial HSI maps
while read LINE; do
  RULES=$SOS_RC_DIR/$GIS_OPT_OUTPUT.$LINE.rules
  g.message -v message="recoding $LINE using  $RULES"
OUT_RCL="tmp${LINE}_rcl"
# use r.recode, since r.reclass wants integer rasters
# only and we need -1-[0..1]
r.recode input=$LINE output=$OUT_RCL rules=$RULES
  --overwrite # --quiet
# reclass "-1" cells to NULL
r.null map=$OUT_RCL setnull=-1 --quiet
r.colors map=$OUT_RCL color=grey1.0 --quiet
NUM_VARS=$(($NUM_VARS+1))
if [ $FIRST = "TRUE" ]; then
MAPCALC_STRING="$MAPCALC_STRING $OUT_RCL"
FIRST=FALSE
else
    case "$GIS_OPT_METHOD" in
      "mult" )
        MAPCALC_STRING="$MAPCALC_STRING * $OUT_RCL"
        ;;
      "add" )
        MAPCALC_STRING="$MAPCALC_STRING + $OUT_RCL"
        ;;
    esac
fi
done < $GIS_OPT_RASTLIST
case "$GIS_OPT_METHOD" in
  "mult" )
    # assemble partial HSI maps using a productory + nth
    # rooth. use 1.0 and not 1 (which implies integer
    # truncation)
    MAPCALC_STRING="$MAPCALC_STRING) ^ (1.0/$NUM_VARS)"
    ;;
  "add" )
    # assemble partial HSI maps using averaging
    MAPCALC_STRING="$MAPCALC_STRING) / $NUM_VARS"
    ;;
esac
```

```
g.message -v message="about to calculate $MAPCALC_STRING"
r.mapcalc $GIS_OPT_OUTPUT="float($MAPCALC_STRING)"
r.colors map=$GIS_OPT_OUTPUT color=ryg --quiet
g.mremove rast="tmp*_rcl" -f --quiet
exit 0
## EOF ##
# kate: encoding utf-8; syntax bash; space-indent on;
# indent-width 2; kate: word-wrap-column 80;
# word-wrap-marker on; word-wrap-marker-color magenta;
# kate auto-brackets on; indent-mode cstyle;
```

CALC.OVERLAY.SH

```
#!/bin/sh
##########################################################
# MODULE: CALC.OVERLAY.sh v.1.0 for GRASS 6.3          #
# AUTHOR(S): M.Carro (monica.carro@libero.it)           #
# PURPOSE: Computes overlay between a map of use of soil#
#          and species presence points. The method      #
#          expects the reclassification of landuse map  #
#          in the sense that each habitat type is given #
#          a value equal to one if one or more presence #
#          points falls within his boundary. For each   #
#          species the technique outputs a dichotomous  #
#          map of potential distribution where 1        #
#          indicates presence and 0 indicates absence   #
# COPYRIGHT: (C) 2008 by Monica carro and by the GRASS  #
#           Development Team/Monica Carro               #
#                                                       #
##########################################################
#                                                       #
# Copyright (C) 2008  Monica Carro                      #
#                                                       #
# This program is free software; you can redistribute it #
# and/or modify it under the terms of the GNU General    #
# Public License as published by the Free Software      #
# Foundation; either version 2 of the License, or (at    #
# your option) any later version.                       #
#                                                       #
# This program is distributed in the hope that it will   #
# be useful, but WITHOUT ANY WARRANTY; without even the  #
# implied warranty of MERCHANTABILITY or FITNESS FOR A   #
# PARTICULAR PURPOSE.  See the GNU General Public       #
# License for more details.                             #
#                                                       #
# A copy of the GNU General Public License is available  #
# at http://www.gnu.org/licenses/gpl.txt, or can be      #
# requested to the Free Software Foundation, Inc.,      #
# 51 Franklin Street, Fifth Floor,  Boston,             #
# MA 02110-1301 USA.                                    #
```

```
#                                                        #
##########################################################
%Module
%  description: fau.model
%End
%option
% key: configfile
% type: string
% gisprompt: old_file,file,input
% description: name of configuration file
% required : yes
%end
%option
% key: datapath
% type: string
% gisprompt: old_file,file,input
% description: path to directory containing shapefiles
% required : yes
%end
%option
% key: usosuolo
% type: string
% gisprompt: old_file,file,input
% description: name of input soil use map
% required : yes
%end
%option
% key: spplist
% type: string
% gisprompt: old_file,file,input
% description: name of list containing species in
  directives habitat and birds
% required : yes
%end


if  [ -z "$GISBASE" ]
then
echo ""
echo "You must be in GRASS GIS to run this program"
echo ""
exit 1
fi

if   [ "$1" != "@ARGS_PARSED@" ]
then
exec g.parser "$0" "$@"
fi

# Check for required arguments
```

```
if [ -z "$GIS_OPT_CONFIGFILE" ]
then
echo "ERROR: Missing configuration file!"
echo "Please enter a valid text file"
exit 1
fi

# Check for required arguments
if [ -z "$GIS_OPT_DATAPATH" ]
then
echo "ERROR: Missing data path!"
echo "Please enter a valid path to directory
      containing shapefile"
exit 1
fi

# Check for required arguments
if [ -z "$GIS_OPT_USOSUOLO" ]
then
echo "ERROR: Missing input soil use map!"
echo "Please enter name of a valid raster map"
exit 1
fi

# Check for required arguments
if [ -z "$GIS_OPT_SPPLIST" ]
then
echo "ERROR: Missing input list of prioritaries species!"
echo "Please enter a valid text file"
exit 1
fi



eval ‘g.gisenv‘
: ${GISBASE?} ${GISDBASE?} ${LOCATION_NAME?} ${MAPSET?}

FAU_VAL=fau_val
r.mapcalc $FAU_VAL=0

while read LINE
  #extracts variables from configuration file end prints
  #them on screen
  do
  echo "$LINE" > configline.txt
  SPECIE=‘cut --delimiter=, --fields=1 configline.txt‘
  echo now doing specie "$SPECIE"
  PRS_MAP=$SPECIE"_prs"
  OUTPUTMAP=$SPECIE"_ovl"
  echo outputmap is "$OUTPUTMAP"
```

```
  v.in.ogr -o -r dsn=$GIS_OPT_DATAPATH output=$SPECIE"_prs"
   layer=$SPECIE type=point --overwrite
  v.db.droptable -f map=$PRS_MAP
  db.connect driver=dbf database='$GISDBASE/$LOCATION_NAME/
   $MAPSET/dbf/'
  v.db.addtable map=$PRS_MAP 'columns=DUSAFLAB int'
  v.what.rast vector=$PRS_MAP raster=$GIS_OPT_USOSUOLO
   layer=1 column="DUSAFLAB"
  CATLIST=‘v.db.select -c map=$SPECIE"_prs" layer=1
   column=DUSAFLAB | awk -F" " '{print $1}'‘
  echo "$CATLIST"
  r.mapcalc $SPECIE=0
  for C in $CATLIST
    do
      r.mapcalc "$SPECIE"_c"=if($GIS_OPT_USOSUOLO==$C,1,0)"
      r.mapcalc "$SPECIE=$SPECIE+$SPECIE"_c""
    done
  g.remove rast=$SPECIE"_c"
  r.mapcalc "$SPECIE=if($SPECIE>1,1,$SPECIE)"
  while read LINE
    do
    echo "$LINE" > spp_cons.txt
    S=‘cut --delimiter=, --fields=1 spp_cons.txt‘
    W=‘cut --delimiter=, --fields=2 spp_cons.txt‘
      case $S in
        $SPECIE) r.mapcalc "$SPECIE=$SPECIE*$W";;
        *) echo -n "$SPECIE non Ãĺ specie di interesse
         conservazionistico";;
      esac
    done < $GIS_OPT_SPPLIST
  r.mapcalc $FAU_VAL=$FAU_VAL+$SPECIE
  #v.select ainput=$GIS_OPT_USOSUOLO atype=area alayer=1
   binput=$SPECIE btype=point blayer=1 output=$OUTPUTMAP
   operator=overlap --overwrite
  done < $GIS_OPT_CONFIGFILE
exit 0
```

OUT.LOGISTIC.R

```
##########################################################
# Script to output required results from                 #
#LOGISTIC REGRESSION ANALYSIS - calling MAKE_AML to      #
# compute output map                                     #
# if required. (ARCINFO version of OUT_LOGISTIC_GRASS)   #
#                                                        #
# requires PRESENCEABSENCE                               #
# this script has to be sourced by CALC.LOGISTIC.R       #
##########################################################
# version 1.0                                            #
```

```
# created 20/06/2008                                     #
# revision history:                                      #
#                                                        #
##########################################################
#                                                        #
# Copyright (C) 2008  Monica Carro                       #
#                                                        #
# This program is free software; you can redistribute it #
# and/or modify it under the terms of the GNU General    #
# Public License as published by the Free Software       #
# Foundation; either version 2 of the License, or (at    #
# your option) any later version.                        #
#                                                        #
# This program is distributed in the hope that it will   #
# be useful, but WITHOUT ANY WARRANTY; without even the  #
# implied warranty of MERCHANTABILITY or FITNESS FOR A   #
# PARTICULAR PURPOSE.  See the GNU General Public        #
# License for more details.                              #
#                                                        #
# A copy of the GNU General Public License is available  #
# at http://www.gnu.org/licenses/gpl.txt, or can be      #
# requested to the Free Software Foundation, Inc.,       #
# 51 Franklin Street, Fifth Floor,  Boston,              #
# MA 02110-1301 USA.                                     #
#                                                        #
##########################################################
ps.options(width=5, height=5, horizontal=TRUE,
 onefile=FALSE, paper="special")

out.glm <- function(specie, logistic.outputs) {
# defines an output text file where results will be copied
output_file = paste(output_text_path,
 paste(specie, "glm_output.txt", sep="_"), sep="/")
print(logistic.outputs)
# if the keyword "logistic_grid" is found in
# logistic.outputs, than calles "MAKE_GRASS.R" to prepare
# GRASS script to compute the final presence/absence GRID
if("logistic_grid" %in% logistic.outputs) {
  source(paste(script_path,"MAKE_AML.R",sep='/'))
  make.aml(results.specie, specie)
  }
#diverts output to text file ouput_file
sink(file = output_file, split = TRUE)
cat("*************************************************\n")
cat(paste("\n*** OUTPUT OF LOGISTIC MODEL FOR",
 specie, "***", sep = " "))
# following keyword instruction for outputs...
# ...summarise results of GLM analysis
if("logistic_summary" %in% logistic.outputs ) {
  cat("\n    ***************** SUMMARY *****************")
```

```
  print(summary(results.specie$logistic$step))
  }
#...extract AIC parameter from GLM analysis results
if("logistic_AIC" %in% logistic.outputs ) {
  cat("\n  ******************* AIC *******************\n")
  print(AIC(results.specie$logistic$step))
  }
# if the keyword ROC comapares in logistic.outputs,
# then data are prepared for ROC analysis and
# Presence Absence package is required
ROC.output <- agrep("ROC", logistic.outputs,
 ignore.case=FALSE, value=TRUE)
if (length(ROC.output)> 0) {
  require(PresenceAbsence)
  # prepares dataframe for ROC analysis, containing a
  # number of rows equal to the number of rows of the
  # input specie.smp table which corresponds to the
  # number of presence/absence points for the current
  # specie. The dataframe
  # contains 3 columns: ID, Observed values,
  # Predicted values
  pa.preds <- data.frame(ID = seq(1:length
   (results.specie$logistic$step$fitted.values)),
    Observed = results.specie$logistic$step$model[,1],
     Predicted = results.specie$logistic$step$fitted.values)
  # following keyword instruction for outputs...
  #...summarize results from ROC analysis
  if("logistic_ROCsummary" %in% ROC.output) {
  cat("\n  *************** ROC SUMMARY ***************\n")
  print(presence.absence.summary(pa.preds))
   }
  #...extracts AUC from ROC analysis
  if("logistic_ROCauc" %in% ROC.output) {
  cat("\n  ***************** ROC AUC *****************\n")
  print(auc(pa.preds))
    }
  #...extracts cutoffs from ROC analysis
  if("logistic_ROCcutoffs" %in% ROC.output) {
  cat("\n  ***************** CUTOFFS *****************\n")
  print(optimal.thresholds(pa.preds))
   }
  # ends diversion to output text file
  sink()
  # ...creates a ROC plot and prints AUC for the model
  if("logistic_ROCplot" %in% ROC.output) {
# ROC plot is copied to output eps file
#(for example "ruprup_glm_rocplot.eps")
postscript(file=paste(output_plots_path,
 paste(specie,"glm_rocplot.eps",sep="_"),
  sep="/"))
```

```
    cat("plot done\n")
    auc.roc.plot(pa.preds)
    dev.off
    }
  }
}
```

MAKE.GRASS.SH.R

```
##########################################################
# Script to create a bash script file to be run into     #
# GRASS GIS. GRASS will produce some raster layers       #
# representing the predicted distribution;               #
# if ROC anlysis is required, then presence/absence      #
# maps of the species are calculated, using optimal      #
# tresholds                                              #
#                                                        #
#                                                        #
# requires: PRESENCEABSENCE                              #
# this script has to be sourced by OUT_LOGISTIC_GRASS.R  #
##########################################################
# version 1.0                                            #
# created 20/09/2008                                     #
# revision history:                                      #
#                                                        #
##########################################################
#                                                        #
#  Copyright (C) 2008  Monica Carro                      #
#                                                        #
#  This program is free software; you can redistribute   #
#  it and/or modify it under the terms of the GNU        #
#  General Public License as published by the Free       #
#  Software Foundation; either version 2 of the          #
#  License, or (at your option) any later version.       #
#                                                        #
#  This program is distributed in the hope that it will  #
#  be useful, but WITHOUT ANY WARRANTY; without even the #
#  implied warranty of merchantability or fitness for a  #
#  paricular purpose.  See the GNU General Public        #
#  License for more details.                             #
#                                                        #
#  A copy of the GNU General Public License is available #
#  at http://www.gnu.org/licenses/gpl.txt, or can be     #
#  requested to the Free Software Foundation, Inc., 51   #
#  Franklin Street, Fifth Floor, Boston,                 #
#  MA 02110-1301 USA                                     #
#                                                        #
##########################################################
make.GRASS <- function(results.specie, cutoffs, specie,
 extension="sh") {
```

```
if (is.data.frame(cutoffs) == TRUE) {
   # Here I choose coefficients 2 and 6 coming from ROC
   # as threshold cutoffs. You can add more declaring
   # them in variable A. The script will create the
   # threshold layers named cut+ [cut coefficient]+
   #[specie] A <- c(2,6)

   # This loop creates a text file with reclassification
   # rules
   # used later by r.reclass
   for (i in seq(1, length(A), by = 1)) {
     output_file = paste(output_text_path,
      paste(specie,A[i],"recrule.txt",sep=""),sep="/")
     sink(file = output_file, append = FALSE,
      type = "output",split=TRUE)
     cut <- cutoffs$Predicted[A]
     cut <- cut*10
     cut <- cat(cut[i],"00000",sep="")
     cat("0 thru ",cut[i],"00000"," = 1 absence","\n",
      sep="")
     cat(cut[i],"00000"," thru 1000000 = 2 presence",
      "\n", sep="")
     cat("end","\n",sep="")
     sink()
     }
   }
# get coefficents
c <- data.frame(c=coef(results.specie$logistic$step))
# coef[1] is the constant, called intercept
row.names(c)[1] <- "intercept"
# initialise output on file & screen
sink(file = paste(output_maps_path,paste(specie,
 extension,sep="."),sep="/"), append = FALSE,
  type = "output",split=TRUE)
  cat("#!/bin/bash\n")
  cat("##################################","\n",sep="")
  cat("# This script is auto generated by R script
  #MAKE_GRASS_SH.R","\n",sep="")
  cat("# The output are raster layers. They are the
  #result of ","\n",sep="")
  cat("# the logistic regression and ROCR ","\n",sp="")
  cat("##################################","\n",sep="")
# variables processing loop
for (i in seq(2, nrow(c), by = 1)) {
  beta <- c$c[i] # get coefficient
  alpha <- row.names(c)[i] # get grid name
  cat("r.mapcalc ",specie,alpha,"=",alpha,
   "*",beta,"\n",sep="")
  }
# Here the loop to calculate the result of logistic
```

```
  # function begins
  cat("r.mapcalc ",specie,"totfunc=",specie,
   row.names(c)[2],"\n",sep="")
  if (nrow(c)> 2) {
    for (i in seq(3, nrow(c), by = 1)) {
      alpha <- row.names(c)[i]
      cat("r.mapcalc ",specie,"totfunc=",specie,
       "totfunc+", specie,alpha,"\n",sep="")
      }
    }
  beta <- c$c[1] #get intercept
  cat("r.mapcalc ",specie,"totfunc= ",specie,"totfunc+",
   beta, "\n",sep="")
  cat("r.rescale input=",specie,"totfunc output=",specie,
   "totfunc1 to=-100,100","\n",sep="")
  cat("r.mapcalc ","'",sep="")
  cat(specie,"expfunc= exp(",specie,"totfunc1)",sep="")
  cat("'","\n")
  cat("r.mapcalc ","'",sep="")
  cat(specie,"pred= ",specie,"expfunc/(1+",specie,
   "expfunc)", Sep="")
  cat("'","\n")
  cat("r.colors map=",specie,"pred color=ryg","\n",sep="")
  # Now some cleaning....
  cat("g.rename rast=",specie,"pred,glm_",specie,"
   --overwrite", "\n",sep="")
  cat("g.rename rast=",specie,"totfunc,tmp_",specie,
   "totfunc --overwrite","\n",sep="")
  cat("g.mremove -f rast=",specie,"*","\n",sep="")
  cat("r.mapcalc glm_",specie,"=glm_",specie,"*1000000",
   "\n",sep="")
  if (is.data.frame(cutoffs) == TRUE) {
    for (i in seq(1, length(A), by = 1)) {
      cat("r.reclass input=glm_",specie," output=glm",A[i],"_",
       specie," rules=",specie,A[i],"recrule.txt","\n",sep="")
      cat("r.mapcalc glm",A[i],"_",specie,"=glm",A[i],"_",specie,
       "\n",sep="")
      }
    }
  cat("exit 128","\n",sep="")
  sink()
}

## -- end of file --
```

OUT.BRUTO.R

```
##########################################################
# Script to output required results from                 #
# FLEXIBLE DISCRIMINANT ANALYSIS with method "BRUTO"      #
```

```
#                                                        #
##########################################################
# version 1.0                                            #
# created 20/06/2008                                     #
# revision history:                                      #
#                                                        #
##########################################################
#                                                        #
# Copyright (C) 2008  Monica Carro                       #
#                                                        #
# This program is free software; you can redistribute it #
# and/or modify it under the terms of the GNU General    #
# Public License as published by the Free Software       #
# Foundation; either version 2 of the License, or (at    #
# your option) any later version.                        #
#                                                        #
# This program is distributed in the hope that it will   #
# be useful, but WITHOUT ANY WARRANTY; without even the  #
# implied warranty of MERCHANTABILITY or FITNESS FOR A   #
# PARTICULAR PURPOSE.  See the GNU General Public        #
# License for more details.                              #
#                                                        #
# A copy of the GNU General Public License is available  #
# at http://www.gnu.org/licenses/gpl.txt, or can be      #
# requested to the Free Software Foundation, Inc.,       #
# 51 Franklin Street, Fifth Floor,  Boston,              #
# MA 02110-1301 USA.                                     #
#                                                        #
##########################################################
# defines a function that prepares ouputs of bruto
# analysis; input
# parameters from "PRESENCE_MODELS.R" are specie and
#bruto.outputs,
out.bruto <- function(specie, bruto.outputs) {
# defines an output text file where results will be copied
output_file = paste(output_text_path, paste(specie,
 "bruto_output.txt", sep="_"), sep="/")
print(bruto.outputs)
# if the keyword "logistic_grid" is found in
# logistic.outputs, than calles "MAKE_AML.R" to prepare
# ArcInfo aml script to compute final presence/absence
# GRID
if("bruto_grid" %in% bruto.outputs) {
  output.asc <- paste(output_maps_path,
   paste(specie, "bruto.asc", sep="_"), sep="/")
  export.asc(results.specie$bruto.fitted, output.asc)
print("grid map computed")
}
#diverts output to text file ouput_file
sink(file = output_file, split = TRUE)
```

```
cat("************************************************\n")
cat(paste("\n * OUTPUT OF BRUTO MODEL FOR", specie, "*",
 sep = " "))
# following keyword instruction for outputs...
# ...summarise results of FDA analysis, method bruto
if("bruto_summary" %in% bruto.outputs ) {
  cat("\n   ***************** SUMMARY *****************")
  print(results.specie$bruto$bruto.fda)
  }
#...prints contingency table
if("bruto_contingency" %in% bruto.outputs ) {
  cat("\n ************* CONFUSION TABLE *************\n")
  print(results.specie$bruto$bruto.confusion)
  }
  sink()
print("outputs computed")
}
```

OUT.ENFA.R

```
##########################################################
# Script to output required results from                 #
# Environment Niche factor Analysis (ENFA)               #
#                                                        #
# requires FOREIGN, PRESENCEABSENCE                      #
#          this script has to be sourced by CALC_ENFA.R  #
##########################################################
# version 1.0                                            #
# created 20/06/2008                                     #
# revision history:                                      #
#                                                        #
##########################################################
#                                                        #
# Copyright (C) 2008  Monica Carro                       #
#                                                        #
# This program is free software; you can redistribute it #
# and/or modify it under the terms of the GNU General    #
# Public License as published by the Free Software       #
# Foundation; either version 2 of the License, or (at    #
# your option) any later version.                        #
#                                                        #
# This program is distributed in the hope that it will   #
# be useful, but WITHOUT ANY WARRANTY; without even the  #
# implied warranty of MERCHANTABILITY or FITNESS FOR A   #
# PARTICULAR PURPOSE.  See the GNU General Public        #
# License for more details.                              #
#                                                        #
# A copy of the GNU General Public License is available  #
# at http://www.gnu.org/licenses/gpl.txt, or can be      #
# requested to the Free Software Foundation, Inc.,       #
```

```
# 51 Franklin Street, Fifth Floor,  Boston,              #
# MA 02110-1301 USA.                                     #
#                                                        #
##########################################################

# defines a function that prepares ouputs of ENFA; input
# parameters from "PRESENCE_MODELS.R" are specie and
# enfa.outputs
out.enfa <- function(specie, enfa.outputs) {
# defines an output text file where results will be copied
output_file = paste(specie, "enfa_output.txt", sep="_")
print(enfa.outputs)
# if the keyword "ENFA grid" is found in enfa.outputs, than
# a raster map of class kasc is produced through
# predict.enfa, based on: 1)enfa result 2)integer vector
# giving the position of the rows of df in the returned
# kasc; 3)an object of class kasc
if("enfa_grid" %in% enfa.outputs) {
  output.asc <- paste(specie, "enfa.asc", sep="_")
  habsuit.pre <- predict(results.specie$enfa$output_enfa,
    results.specie$enfa$maps_index,
    results.specie$enfa$maps)
  habsuit.pre <- sqrt(habsuit.pre)
  habsuit.pre <- (1 - ( habsuit.pre /
   max(habsuit.pre, na.rm = TRUE)))
  #exports raster map of class asc into ESRI ErcInfo ASCII
  #raster file
  export.asc(habsuit.pre, output.asc)
  }
#diverts output to text file ouput_file
sink(file = output_file, split = TRUE)
# following keyword instruction for outputs...
# ...summarise results of ENFA analysis (kept variables
# and percentage of explained
# variance from 1st axis after variables reduction)
if("enfa_summary" %in% enfa.outputs) {
  cat("\n**************  ENFA SUMMARY ***************\n")
  print(paste("1st AXIS EXPLAINED VARIANCE:",
   results.specie$enfa$C1_Varexp, sep = " "))
  print(paste("KEPT VARIABLES: ",
   paste(results.specie$enfa$KeptVariables,
    collapse= "+")))
  }
# if the keyword ROC comapares in enfa.outputs, then data
# are prepared for # ROC analysis and Presence Absence
# package is required
ROC.output <- agrep("ROC", enfa.outputs, ignore.case=FALSE,
 value=TRUE)
if (length(ROC.output)> 0) {
  require(PresenceAbsence)
```

```
# prepares dataframe for ROC analysis, containing a number
# of rows equal to the number cells of output prediction
# grids where falls at leat 1 presence pred.locs <- list()
# counts the number of points in each pixel of the raster
# prediction map of asc.
cp <- count.points(results.specie$enfa$locs, habsuit.pre)
# puts result of count.points and the raster prediction map
# in a temporary list
# of matrices of class asc
pred.locs[["cp"]] <- cp
pred.locs[["pred"]] <- habsuit.pre
# in order to convert the list of matrices into dataframe
# it must be converted
# into kasc object firt and than to dataframe through
#kasc2df
pred.locs.kasc <- as.kasc(pred.locs)
pred.locs.df <- kasc2df(pred.locs.kasc)
# now results are available in the correct form for ROC
# analysis; The dataframe
# contains 3 columns: ID, Observed values(wich correspond
# to the number of counted presence points), Predicted
# values (wich correspond to the value of
# the prediction enfa output matrice)
pa.preds <- data.frame(ID=pred.locs.df$index,
 Observed=pred.locs.df$tab$cp,
  Predicted=pred.locs.df$tab$pred)
# Since ROC analysis admits only 0 and 1
# (absence or presence)as possible values
# for Observed in the input dataframes and values
# between 0 and 1 for Predicted...
# ...first, value 1 (presence) is assigned to
# locations with more than 1 presence point
pa.preds[pa.preds$Observed > 0, ]$Observed <- 1
# ...second, "Predicted" values are trasformed in
# order to be in the range 0-1
pa.preds$Predicted <- sqrt(pa.preds$Predicted)
pa.preds$Predicted <- (1-(pa.preds$Predicted /
 max(pa.preds$Predicted)))
# following keyword instruction for outputs...
#...summarize results from ROC analysis
if("enfa_ROCsummary" %in% ROC.output) {
  cat("\n *************** ROC SUMMARY ***************\n")
  print(presence.absence.summary(pa.preds))
  }
#...extracts AUC from ROC analysis
if("enfa_ROCauc" %in% ROC.output) {
  cat("\n ***************** ROC AUC *****************\n")
  print(auc(pa.preds))
  }
#...extracts cutoffs from ROC analysis
```

```
  if("enfa_ROCcutoffs" %in% ROC.output) {
    cat("\n ***************** CUTOFFS *****************\n")
    print(optimal.thresholds(pa.preds))
    }
  }
  # ends diversion to output text file
  sink()
  # ...creates a ROC plot and prints AUC for the model
  if("enfa_ROCplot" %in% ROC.output) {
    cat("plot done\n")
    auc.roc.plot(pa.preds)
    # ROC plot is copied to output eps file (for example
    #"ruprup_enfa_rocplot.eps")
    dev.copy2eps(file=paste(output_plots_path,paste(
     specie,"enfa_rocplot.eps",sep="_"),sep="/"))
    }
# ...creates a graphical summary of enfa results
# (eigenvalues barplot, correlation circle
# scatter diagram)
if("enfa_plotsummary" %in% enfa.outputs) {
  cat("graph done\n")
  opar <-par(mfrow=c(2,2))
  barplot(results.specie$enfa$PCA$eig)
  s.corcircle(results.specie$enfa$PCA$co)
  scatter(results.specie$enfa$output_enfa)
  dev.copy2eps(file=paste(specie,"enfa_graphsummary.eps",
   sep="_"))
  }
# ...creates histograms of marginality and specialisation
#  axis from ENFA results
if("enfa_hist" %in% enfa.outputs) {
  cat("hist done\n")
  hist(results.specie$enfa$output_enfa)
  dev.copy2eps(file=paste(specie,"enfa_hist.eps",sep="_"))
  }
}
```

B.0.7. SOS scripts for geoprocessing

V.WHATRAST.MANY.SH

```
#!/bin/sh
##########################################################
#                                                        #
# MODULE: V.WHATRAST.MANY.SH v.1.0 for GRASS 6.3         #
#                  #
# AUTHOR(S): M.Carro (monica.carro@libero.it)            #
#                  #
# PURPOSE: Given a text configuration file with a list  #
```

```
#           of vector point layers and for each vector   #
#           the corresponding text file containing the   #
#           list of raster layers, updates the attributes#
#           table of each vector point layer with values #
#           of the raster layers listed in the text file #
#                                                        #
# COPYRIGHT:(C) 2008 by the GRASS Development Team       #
# This program is free software under the GNU  #
#           General Public License (>=v2). Read the file #
#           COPYING that comes with GRASS for details.  #
#                      #
##########################################################
%Module
%  description: v.whatrast.many
%End
%option
% key: configfile
% type: string
% gisprompt: old_file,file,input
% description: Text file with the list of vector input
 layers and text file of rasters list
% required : yes
%end
%option
% key: datapath
% type: string
% gisprompt: old_file,file,input
% description: path to directory containing configuration
 file and text file with raster list
% required : yes
%end


if  [ -z "$GISBASE" ]
then
echo ""
echo "You must be in GRASS GIS to run this program"
echo ""
exit 1
fi

if   [ "$1" != "@ARGS_PARSED@" ]
then
exec g.parser "$0" "$@"
fi

# Check for required arguments
if [ -z "$GIS_OPT_CONFIGFILE" ]
then
echo "ERROR: Missing configuration file!"
```

```
echo "Please enter a valid text file"
exit 1
fi

# Check for required arguments
if [ -z "$GIS_OPT_DATAPATH" ]
then
echo "ERROR: Missing data path!"
echo "Please enter a valid path to data"
exit 1
fi


eval ‘g.gisenv‘
: ${GISBASE?} ${GISDBASE?} ${LOCATION_NAME?} ${MAPSET?}


#db.tables -p driver=sqlite database=$GISDBASE/
 $LOCATION_NAME/$MAPSET/sqlite.db
#TABLELIST=‘db.tables -p driver=sqlite database=$GISDBASE/
 $LOCATION_NAME/$MAPSET/sqlite.db‘
#echo "$TABLELIST" > tablelist.txt

#Iteration on each point vector map listed in the
#configuration file

while read LINE
  #extracts variables from configuration file end prints
  # them on screen
  do
  echo "$LINE" > configline.txt
  SPECIE=‘cut --delimiter=, --fields=1 configline.txt‘
  echo now doing specie "$SPECIE"
  OUTPUTTABLE=$SPECIE
  echo output table is "$OUTPUTTABLE"
  RASTERLIST=‘cut --delimiter=, --fields=2 configline.txt‘
  echo "using raste pack "$RASTERLIST""

  # Check if table of attributes already exists
  while read RIGA
  do
   if [ "$RIGA" = "$OUTPUTTABLE" ]
   then
      echo "drop table "$OUTPUTTABLE"" | db.execute
      echo table dropped
   fi
  done < tablelist.txt

  #Creates and populates table from vector points
```

```
  #sets database connection
  db.connect driver=sqlite database=’$GISDBASE/
  $LOCATION_NAME/$MAPSET/sqlite.db’
  #copy vector point attribute table to output table in
  #order to preserve attributes
  db.copy from_driver=sqlite from_database=$GISDBASE/
   $LOCATION_NAME/$MAPSET/sqlite.db from_table=$SPECIE
    to_driver=sqlite to_database=$GISDBASE/$LOCATION_NAME/
     $MAPSET/sqlite.db to_table=$OUTPUTTABLE
  # sets connection for the vector point map to output
  #table
  v.db.connect -o map="$SPECIE" table="$OUTPUTTABLE"
   key=CAT
   # populates table from vector features with categories
  v.to.db map="$SPECIE" type=point layer=1 qlayer=1 option=cat
  # add two columns to store points coordinates
  v.db.addcol map="$SPECIE" columns="xcoor double precision"
  v.db.addcol map="$SPECIE" columns="ycoor double precision"
  #populates table with coordinates
  v.to.db map="$SPECIE" type=point layer=1 qlayer=1
   option=coor column=xcoor,ycoor

  #updates tables with rster values
  while read LINE
    do
    echo "ALTER TABLE "$SPECIE" ADD COLUMN $LINE
     double precision" | db.execute
    v.what.rast vect="$SPECIE" rast="$LINE" col="$LINE"
    done < "$GIS_OPT_DATAPATH""/""$RASTERLIST"

    echo "Table $OUTPUTTABLE updated"

  done < $GIS_OPT_CONFIGFILE

rm -f tablelist.txt
exit 0
```

V.PRESABS.MANY.SH

```
#!/bin/sh
##########################################################
#                                                        #
# MODULE:   V.PRESABS.MANY.SH v.1.1 for GRASS 6.3        #
# AUTHOR(S):M.Carro (monica.carro@uninsubria.it)         #
# PURPOSE:  Creates presence/absence maps from point     #
#           vectors of real observations.                #
# COPYRIGHT:(C) 2008 by the GRASS Development Team       #
#                                                        #
# This program is free software under the GNU            #
```

```
# General Public License (>=v2)                          #
##########################################################

%Module
%  description: v.presabs.many
%End
%option
%key: specieslist
%type: string
%gisprompt: old_file,file,input
%description: Text file of species list
%required : yes
%end


if  [ -z "$GISBASE" ]
then
  echo ""
  echo "You must be in GRASS GIS to run this program"
  echo ""
  exit 1
fi

if   [ "$1" != "@ARGS_PARSED@" ]
then
  exec g.parser "$0" "$@"
fi

# Check for required arguments
if [ -z "$GIS_OPT_SPECIESLIST" ]
then
  echo "ERROR: Missing name of text file containing list
   of species!"
  echo "Please enter valid text file"
  exit 1
fi


SPECIESLIST=‘cat $GIS_OPT_SPECIESLIST‘

for SPECIE in $SPECIESLIST
do
  PRS_MAP=$SPECIE"_prs"
  ABS_MAP=$SPECIE"_abs"
  v.db.addcol map=$SPECIE ’columns=PRES int’
  v.db.update map=$SPECIE layer=1 column=PRES value=1
  echo "create table $SPECIE"_ori" as select *
   from $SPECIE;" | db.execute
  echo "create table $PRS_MAP as select CAT, PRES
   from $SPECIE;" | db.execute
```

```
  g.rename vect=$SPECIE,$PRS_MAP
  v.db.connect -d map=$PRS_MAP table=$SPECIE
  v.db.connect map=$PRS_MAP table=$PRS_MAP
  db.droptable -f table=$SPECIE
  NUM_PRS_POINTS=‘v.info map=$PRS_MAP | grep -F "Number of
    points" | awk -F" " '{print $5}'‘
  NUM_ABS_POINTS=$(echo "scale=2; $NUM_PRS_POINTS +
   $NUM_PRS_POINTS/100*30" | bc)
  v.random output=$ABS_MAP n=$NUM_ABS_POINTS
  v.db.addtable map=$ABS_MAP 'columns=PRES int'
  v.db.update map=$ABS_MAP layer=1 column=PRES value=0
  v.patch -e input=$PRS_MAP,$ABS_MAP output=$SPECIE
  g.remove vect=$PRS_MAP
  g.remove vect=$ABS_MAP
done
```

V.PRESABS.MANY.TRANSECTS.SH

```
#!/bin/sh
##########################################################
#                                                        #
# MODULE: V.PRESABS.MANY.TRANSECTS.SH v.1.0 for GRASS6.3 #
# AUTHOR(S):M.Carro (monica.carro@uninsubria.it)         #
# PURPOSE:  Creates presence/absence maps from point     #
#           vectors showing observation along transects  #
# COPYRIGHT:(C) 2008 by the GRASS Development Team       #
#                                                        #
# This program is free software under the GNU General    #
# Public License (>=v2)                                  #
                                                         #
##########################################################

%Module
%  description: v.presabs.many
%End
%option
%key: specieslist
%type: string
%gisprompt: old_file,file,input
%description: Text file of species list
%required : yes
%end
%option
%key: transects
%type: string
%gisprompt: old_file,file,input
%description: vector file with transects
%required : yes
%end
```

```
if  [ -z "$GISBASE" ]
then
  echo ""
  echo "You must be in GRASS GIS to run this program"
  echo ""
  exit 1
fi

if   [ "$1" != "@ARGS_PARSED@" ]
then
  exec g.parser "$0" "$@"
fi

# Check for required arguments
if [ -z "$GIS_OPT_SPECIESLIST" ]
then
  echo "ERROR: Missing name of text file containing list
   of species!"
  echo "Please enter valid text file"
  exit 1
fi

# Check for required arguments
if [ -z "$GIS_OPT_TRANSECTS" ]
then
  echo "ERROR: Missing name of map of transect!"
  echo "Please enter valid vector of lines"
  exit 1
fi


SPECIESLIST=‘cat $GIS_OPT_SPECIESLIST‘
TRS_BUF="transects_buf"
#v.buffer --overwrite input=$GIS_OPT_TRANSECTS
 output=$TRS_BUF type=line distance=1000
v.random output="random_abs" n=10000 --overwrite
v.category -g input="random_abs" output="random_abs1"
 option=sum type=point cat=100000 --overwrite
g.rename vect="random_abs1","random_abs" --overwrite

for SPECIE in $SPECIESLIST
do
  PRS_MAP=$SPECIE"_prs"
  ABS_MAP=$SPECIE"_abs"
  PRS_BUF=$SPECIE"_prs_buf"
  ABS_BUF=$SPECIE"_abs_buf"
  #v.extract input=$SPECIE"_ori" output=$SPECIE
   where='PRES="1"' --overwrite
  g.copy vect=$SPECIE,$SPECIE"_ori" --overwrite
```

```
  g.rename vect=$SPECIE,$PRS_MAP --overwrite
  db.droptable -f table=$PRS_MAP
  v.db.connect -d map=$PRS_MAP
  v.db.addtable map=$PRS_MAP 'columns=cat
   integer,PRES integer'
  v.db.update map=$PRS_MAP layer=1 column=PRES value=1
  #v.buffer --overwrite input=$PRS_MAP output=$PRS_BUF
   type=point distance=1500 minordistance=1500
  v.overlay --overwrite ainput=$TRS_BUF binput=$PRS_BUF
   output=$ABS_BUF operator=not
  v.select ainput="random_abs" atype=point binput=$ABS_BUF
   btype=area output=$ABS_MAP --overwrite
  v.db.addtable map=$ABS_MAP 'columns=PRES int'
  v.db.update map=$ABS_MAP layer=1 column=PRES value=0
  v.patch -e input=$PRS_MAP,$ABS_MAP output=$SPECIE --overwrite
  g.remove vect=$PRS_MAP
  g.remove vect=$ABS_MAP
  #g.remove vect=$PRS_BUF
  g.remove vect=$ABS_BUF
done
```

V.PRESABS.MANY.GRID.SH

```
#!/bin/sh
##########################################################
#                                                        #
# MODULE:   V.PRESABS.MANY.GRID v.1.1 for GRASS 6.3      #
# AUTHOR(S):M.Carro (monica.carro@uninsubria.it)         #
# PURPOSE:  Creates presence/absence maps from vector    #
#           grids of atlases                             #
# COPYRIGHT:(C) 2008 by the GRASS Development Team       #
#                                                        #
#This program is free software under the GNU General     #
#Public License (>=v2)                                   #
#Read the file COPYING that comes with GRASS for details #
##########################################################

%Module
%  description: v.presabs.many
%End
%option
%key: specieslist
%type: string
%gisprompt: old_file,file,input
%description: Text file of species list
%required : yes
%end
%option
%key: inputvectorgrid
```

```
%type: string
%gisprompt: old_file,file,input
%description: name of vector grid with presence/absence
 data
%required : yes
%end
%option
%key: inputvector
%type: string
%gisprompt: old_file,file,input
%description: name of vector with 0 points to use for
 patch command
%required : yes
%end

if  [ -z "$GISBASE" ]
then
  echo ""
  echo "You must be in GRASS GIS to run this program"
  echo
  exit 1
fi

if   [ "$1" != "@ARGS_PARSED@" ]
then
  exec g.parser "$0" "$@"
fi

# Check for required arguments
if [ -z "$GIS_OPT_SPECIESLIST" ]
then
  echo "ERROR: Missing name of text file containing list
    of species!"
  echo "Please enter valid text file"
  exit 1
fi

# Check for required arguments
if [ -z "$GIS_OPT_INPUTVECTORGRID" ]
then
  echo "ERROR: Missing name of vector grid!"
  echo "Please enter of a valid vector"
  exit 1
fi

# Check for required arguments
if [ -z "$GIS_OPT_INPUTVECTOR" ]
then
  echo "ERROR: Missing name of vector grid!"
  echo "Please enter of a valid vector"
```

```
  exit 1
fi


SPECIESLIST=‘cat $GIS_OPT_SPECIESLIST‘
GRID=$GIS_OPT_INPUTVECTORGRID

for SPECIE in $SPECIESLIST
do
  echo "now doing specie $SPECIE"
  PRS_MAP=$SPECIE"_prs"
  ABS_MAP=$SPECIE"_abs"
  v.db.droptable -f map=$GIS_OPT_INPUTVECTOR layer=1
  v.db.addtable map=$GIS_OPT_INPUTVECTOR ’columns=PRES int’
  echo "names written"
  g.copy vect=$GIS_OPT_INPUTVECTOR,$PRS_MAP
  g.copy vect=$GIS_OPT_INPUTVECTOR,$ABS_MAP
  SPECIEGRID="grid_"$SPECIE
  v.extract --overwrite input=$GRID output=$SPECIEGRID
   where="$SPECIE > ’0’"
  echo "$SPECIE extracted!"
  v.db.select -c map=$SPECIEGRID > extract.txt
  cut --delimiter="|" --fields="1" extract.txt > cats.txt
  echo cats.txt
  while read ROW
  do
    v.extract input=$SPECIEGRID output=$SPECIE"_cat"
     list=$ROW
    g.region -p vect=$SPECIE"_cat"
    NUM_PRS_POINTS1=‘v.db.select -c map=$SPECIE"_cat"
     columns=$SPECIE‘
    NUM_PRS_POINTS=$(echo "scale=2;
     $NUM_PRS_POINTS1 * 100"| bc)
    NUM_ABS_POINTS=$(echo "scale=2;
     $NUM_PRS_POINTS + $NUM_PRS_POINTS/100*30" | bc)
    v.random --overwrite output=$SPECIE"_random"
     n=$NUM_PRS_POINTS
    v.db.addtable map=$SPECIE"_random" ’columns=cat
     integer,PRES integer’
    v.db.update map=$SPECIE"_random" layer=1
     column=PRES value=1
    v.patch -a -e --overwrite input=$SPECIE"_random"
     output=$PRS_MAP
    g.region -p vect=$GRID
    echo "now region is" $GRID
    v.random --overwrite output=$SPECIE"_random"
     n=$NUM_ABS_POINTS
    v.db.addtable map=$SPECIE"_random" ’columns=cat integer,
     PRES integer’
    v.db.update map=$SPECIE"_random" layer=1 column=PRES
```

```
     value=0
    v.patch -a -e --overwrite input=$SPECIE"_random"
     output=$ABS_MAP
    g.remove vect=$SPECIE"_random"
    g.remove vect=$SPECIE"_cat"
  done < cats.txt
  v.patch -e --overwrite input=$PRS_MAP,$ABS_MAP
   output=$SPECIE
  g.remove vect=$PRS_MAP
  g.remove vect=$ABS_MAP
  g.remove vect=$SPECIEGRID
done
```

R.CATDIST.MANY.SH

```
#!/bin/sh
##########################################################
#                                                        #
# MODULE:     r.catdist.many v.1.0 for GRASS 6.3         #
# AUTHOR(S):  M.Carro (monica.carro@uninsubria.it);      #
# PURPOSE:    Creates a raster layer for each category   #
#             of the input raster layer and then         #
#             computes shortest distance surfaces from   #
#             raster polygons for each layer             #
#                                                        #
# COPYRIGHT:  (C) 2008 by the GRASS Development Team     #
#             This program is free software under the    #
#             GNU General Public                         #
#                                                        #
##########################################################

%Module
%description: r.catdist.many
%End
%option
%key: inputrast
%type: string
%gisprompt: raster
%description: Input raster layer
%required : yes
%end

if  [ -z "$GISBASE" ]
then
echo ""
echo "You must be in GRASS GIS to run this program"
echo ""
exit 1
```

```
fi

if   [ "$1" != "@ARGS_PARSED@" ]
then
exec g.parser "$0" "$@"
fi

# Check for required arguments
if [ -z "$GIS_OPT_INPUTRAST" ]
then
echo "ERROR: Missing raster layer!"
echo "Please enter a raster layer in the field Raster
 layer"
exit 1
fi

#creates a raster map with all values=1 as cost weights in
#order to compute shortest distance maps
g.remove rast=AREA_ONE
r.mapcalc 'AREA_ONE =1'

#extracts category values from raster input map
CATLIST=`r.category map=$GIS_OPT_INPUTRAST | awk -F" " '{print $1}'`
echo "categories copied"

#for each category of the input raster map
for C in $CATLIST
do
#defines name of first output raster
OUTRASTCAT="h"$C
g.remove rast=$OUTRASTCAT
#creates a raster with category value and nodata elsewhere
r.mapcalc "$OUTRASTCAT=if($GIS_OPT_INPUTRAST==$C,
 $GIS_OPT_INPUTRAST,null())"
#defines name of second output raster
OUTRASTDIS="ev_d"$C
        OUTRASTDISC="ev_dc"$C
        g.remove rast=$OUTRASTDISC
        g.remove rast=$OUTRASTDIS
#computes shortest distance map
r.cost -k input="AREA_ONE" output="$OUTRASTDISC"
 start_rast="$OUTRASTCAT"
        r.mapcalc "$OUTRASTDIS=$OUTRASTDISC *
         (ewres()+nsres())/2"
        g.remove rast=$OUTRASTCAT
        g.remove rast=$OUTRASTDISC
done
#removes weights map for the computation of cost surfaces
g.remove rast=AREA_ONE
```