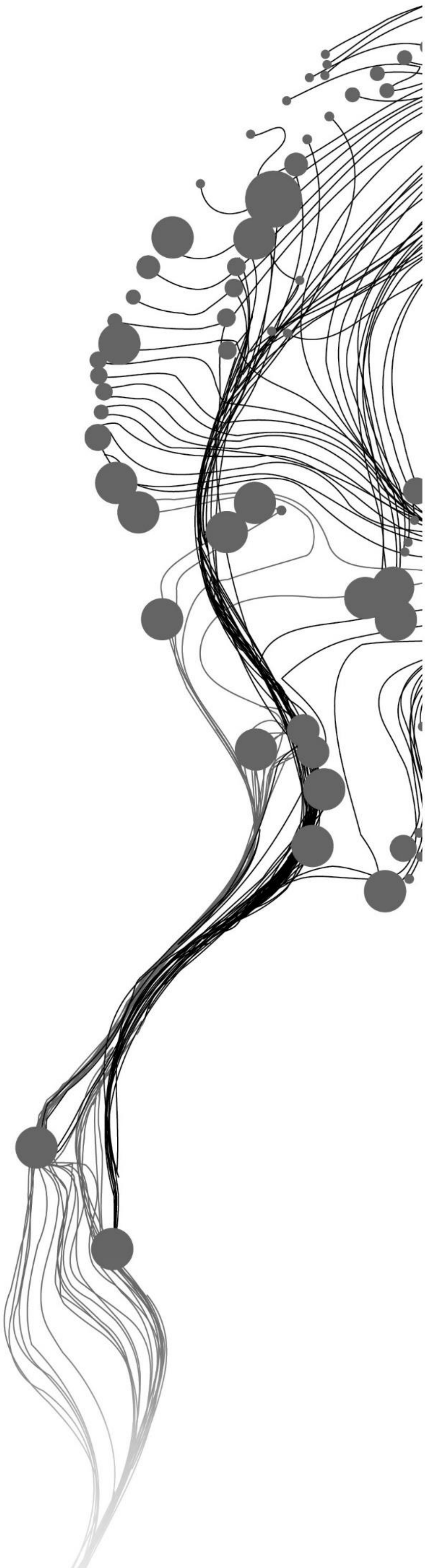
ADDRESSING THE IMPORTANCE OF THREAT MAPS AND RESOLUTION IN MODELLING THE DISTRIBUTION OF THE EUROPEAN GROUND SQUIRREL

MIRIAM DEL ROSARIO VÁZQUEZ PÉREZ

June, 2020

SUPERVISORS:

Dr P. Nyktas
Drs M. Looijen



ADDRESSING THE IMPORTANCE OF THREAT MAPS AND RESOLUTION IN MODELLING THE DISTRIBUTION OF THE EUROPEAN GROUND SQUIRREL

MIRIAM DEL ROSARIO VÁZQUEZ PÉREZ
Enschede, The Netherlands, June, 2020

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialisation: Natural Resources Management

SUPERVISORS:
Dr P. Nyktas
Drs J. Looijen

THESIS ASSESSMENT BOARD:
Dr A.G.B. Toxopeus (Chair)
Dr. G.H. Griffiths (External Examiner, Department of Geography and Environmental Science & School of Agriculture Policy & Development, University of Reading)

ABSTRACT

Protected Areas are not isolated from its surroundings; its management must consider an integrated approach beyond its borders. Threats can direct or indirect damage to the values that the Protected Area should keep. Direct threats can be addressed by the local management of the Protected Area, but indirect threats are more challenging, as the source is not easily traceable. Unfortunately, mapping threats is not extended practice as data availability is a recurrent issue. Since "maps are an invaluable tool to inform and assist decision-making" regardless of how a map was created; a manager will use the available data (Lecours, 2017). As conservation efforts do not always specify standards or guidelines for the production of maps, results can vary dramatically. Another issue is that the final product does not always have feedback from the manager of the resource. The consequences making conservation decisions based on a misinterpretation of maps include loss of trust in science, conservation plans or management plans that do not reach their objectives, financial costs and temporal costs (Guisan et al., 2013; Lecours, 2017).

Using species modelling and GLM and linking the environmental variables and the geo-representation of the threats the Author mapped the probable presence given burrow characteristics of the European Ground Squirrel (*Spermophilus citellus*) in an administrative region context in the Region of Central Macedonia in Greece, considering environmental variables and its reported threats and the effect of fitting the probable presence at different pixel resolutions: 30 x 30, 100 x 100, 500 x 500 and 1000 x 1000 in meters.

Mapping the probable presence with only environmental variables at different resolutions did not change the importance in the environmental variables altitude, slope, and soil parent material Fluvial clays, silts and loams. Trying to use the land cover was unsuccessful as by upscaling the pixel resolution, the category in which the presence points was originally shifted to another one in the new resolution.

The Author tried a simple approach for the mapping of the threats and linking them to geo-representation the using, density, distance and Boolean. Although including the threats did increase the accuracy of the models at all resolutions, no threat was consistent on being significant at all resolutions. Each resolution had a different mixture of threats that combined with the environmental variables explained the probable presence or absence of the species. The threats that contributed the most to the models were the roads.

For the ground squirrel in Central Macedonia, the model with the resolution that has more significant variables and threats, and reflects the most considerable improvement on accuracy by considering the threats, and as well as 2 of 3 of the best accuracy assessments is the 500 resolution. This does however not mean that all the European ground squirrel range of probable habitat should be mapped at this resolution nor that other related species should be mapped this way. It indicates that under the reviewed factors in the study area, these are the relationships.

This analysis suggests that at different pixel resolutions in the same extent, it possible to model the presence of the European Ground Squirrel or European Souslik (*Spermophilus citellus*) given its burrow environmental characteristics. The models were developed by implying that the effect will be the same in the study area, although we know that this generalization is not a reflection of the real world, is a helpful approximation. Finally, mapping threats is essential to conservation efforts, and further research has to be done to translate the concept into a map. And this way to implement efficient science-based natural resources management actions.

ACKNOWLEDGEMENTS

I would such as to express my gratitude to my supervisors, this thesis would not have been possible without their contributions. There was a language barrier, but everything worked out at the end. To my first supervisor, Dr P. Nyktas, for providing valuable data and for your willingness to help. As well, to my second supervisor Drs J. Looijen for all her reassurance, advice and steady mindset during this thesis process. Also, to the board chair Dr A.G. Berth Toxopeus, for his on-point brief commentary to improve my work.

I thank the ITC faculty, for providing the opportunity and the means to expand my knowledge. In particular to the Course Director Dr Raymond G. Nijmeijer for all his guidance and for always keeping his office door open for a talk.

I am grateful to the Mexican Science and Technology Council CONACYT, for believing in my potential and awarding me the scholarship that made all this possible. To my friend and teacher back in Mexico, Dr Alfonso Cuevas Jimenez, for all his wisdom and encouragement to keep learning and improving as a professional.

Last but not least, to my family, old and new friends; the ones that I found on the road and the ones that stay with me till the end, may our paths cross again. Thank you, this has been the weirdest travel yet.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Conservation and Protected areas

The International Union for Conservation of Nature (IUCN) defines a *Protected Areas* as an "area of land and sea mainly dedicated to the protection and maintenance of biological diversity, and of natural and associated cultural resources, and managed through legal or other effective means" (IUCN, Stolton, Shadie, & Dudley, 2008). Protected Areas are created to safeguard its values reflects the understanding, priorities and knowledge at the time of the environment and the people and nature feedbacks. Generally speaking, a value is something that is held in high regard. These spaces are chosen to preserve: specific flora, fauna, habitats, landforms, hydrology, cultural relevant locations, or even unique human-nature relationships to mention some. Today Protected Areas have become the main strategies for Conservation in situ as they eliminate, minimise or reduce human pressures within its borders (UNEP-WMCM, IUCN, & and NGS, 2018).

Protected Areas are cornerstones in which regional conservancy strategies are built up (Margules & Pressey, 2000). Transversally of a policy is a way to coordinate actions within public policy and align the decisions. The concept of managing whole landscapes is of recent creation, but it has already incorporated into public policy. However, as each management body has a different scope of action, reaching an agreement that please all the involved is a significant challenge. By being able to counter the threats from different fronts, risk mapping allows us to observe from a broader perspective and to coordinate and align different to policies and actions as part of the governability of territory and then how it impacts the Protected Area (Leverington, Hockings, & Lemos Costa, 2008; UNEP-WCMC and IUCN, 2019).

The designation of a Protected Area is not a guarantee of protection, but the first step in a management cycle that aids the safekeeping of the values. The efficiency of the Protected Areas depends on the management of its values and threats, and to be able to do that, first you have to know where they are. In the environmental decision-making process the characterisation or context setting of the system is the first step in an iterative cycle, that will go from identifying the suitable actions to planning, then implementation and monitoring, and then again to review if the setting has changed. Different management related tools use maps into aid the decision making process such as the identification of priority locations for Conservation, the General Management Plan with its zoning schema, Operational Plan, conservation goals, to mention some (Leverington et al., 2010; Stankey, Clark, & Bormann, 2005).

Since Protected Areas are subsets of a broader ecosystem, alterations outside its borders will have repercussions inside (Hansen & DeFries, 2007). As the Protected Area management team faces direct and indirect threats at the same time, it must be able to come up with actions to overcome more than one and coordinate its actions with other management bodies (Leverington et al., 2010; Wilson, Sleeter, Sherba, & Cameron, 2015).

Accurate data for the designation and management of the Protected Area values is essential; this way enabling different management bodies, stakeholders and other actors to reach an agreement. The United Nations Development Programme and the Convention on Biological Diversity acknowledge mapping Protected Areas threats as a component on the identification as part of the actions to support climate resilience and action, maintain ecosystem services and strength management planning for sustainable livelihoods (Ervin, Sekhran, Gidda, Vergeichik, & Mee, 2010; UNEP-WMCM et al., 2018). The Convention on Biological Biodiversity (CBD) recognises that threats to Protected Areas compromise its conservation goals. Goal 1.5 highlights that to overcome the threats, Protected areas must identify and assess critical threats in local and regional scales. Furthermore, from this develop strategies to mitigate the threats (CBD, 2004).

Despite this, the data on spatial and temporal Distribution of human-related threats to biodiversity is not always available, or the resolution does not allow to scale down the threat and focus the actions of policies (Joppa et al., 2016). The Habitat Directive requires that the Protected Areas within the Natura 2000 Network compile and report the threats to the species and habitats then generate actions, monitor its results and report them, such as the management cycle. These reports are made by the site, region, Country and Biogeographical area (EEC, 1992).

1.2. Threats

In Conservation, according to Salafsky et al., (2008), *threats* are "The proximate human activities or processes that have caused, are causing, or may cause the destruction, degradation, and/or impairment of biodiversity targets", threats are a "degraded condition or symptom of the target that results from a direct threat". In a traditional conservancy one species focused, one threat is linear to one species wellness, ecosystem dynamics are nonlinear but a mixture of feedback loops (Scheffer, Carpenter, Foley, Folke, & Walker, 2001). Threats can direct or indirect damage to the values that the Protected Area should keep. Direct threats can be addressed by the local management of the Protected Area, but indirect threats are more challenging, as the source is not easily traceable to (Bonebrake et al., 2019).

Human activities require space and resources, although not all of them will produce the same damage to the Protected Areas species and habitats. Different activities will generate a variety of effects that will become a threat to the values inside the Protected Areas, such as unsustainable hunting, recreational activities related to tourism, promoting or suppressing fires, illegal logging to mention some. Threats also originate from the outside, mining, quarrying and oil drilling, fragmentation caused by road construction and other utility lines, pollution of various kinds, invasive species, land-use changes (Schulze et al., 2018).

On the mapping of threats and linking them to geo-representation some authors have developed methodologies to map threats to protected areas (Mcpherson et al., 2008) but as these methodologies are highly dependent on expert opinion techniques, other approaches using, density, distance and Boolean can be used.

Mapping threats is part of the Structured Decision-Making framework; it improves transparency and the application of multiple objective decision making for environmental management and public policy problems such as the implementation of Programs and Actions for monitoring and managing of Protected Areas (Gregory et al., 2012). Unfortunately, mapping threats is not extended practice as data availability is a recurrent issue; "maps are an invaluable tool to inform and assist decision-making" regardless of how a map was created; a manager will use the available data (Lecours, 2017).

Threats are species-specific, and its effects depend on its spatial characteristics, as reviewed by Worboys et al. (2006) and Salafasky et al. (2008), threats are divided by contact in direct and indirect, identifying different scales local, regional and international for mapping. The IUCN uses Salafasky et al. (2008) threat definition in its Unified Classification of Direct Threats this classification it is used in the European Union to map and measure threats and pressures (IUCN-CMP, 2018).

Protected Areas are not isolated from its surroundings but a component in a coherent and broader conservation landscape, its management must consider an integrated approach beyond its borders. When making conservation decisions, the environmental and social context must be taken in account, including the threats, not as the only factor but as part of broader strategies this way the most effective mitigation action is chosen. Threat maps represent process on which Conservation values are at risk by connecting possible biodiversity outcomes to threatening process. Also, when threat mapping is combined with a transparent and replicable, considering constraints, consequences of the actions and the uncertainty of the method, is possible to obtain defendable conservation decisions as part of the Structured Decision-Making process (Tulloch et al., 2015).

1.3. Modelling the distribution of species

The distribution of the species is determined by a combination of biotic and abiotic interactions that are specific to a location and to the species, these combinations are -roughly- the species niche. The niche may be best described by a scale for one species as that is the species ecology and a completely different for another one, although they share the same space, they live in different geographic dimensions (Milne, 1995). *Potential distribution* the range that is environmentally suitable to maintain the population given the topography, and climatic conditions and *realized distribution*, the area where the species actually roams that includes other biotic interactions and pressures to predict the presence of a species (Guisan & Zimmermann, 2000). Furthermore, the *home rage* is the area of the movement were individuals from the species move, for ground squirrel is the area of movement from their own burrows (Ramos-Lara et al., 2014).

Species presence data comes from different sources: census, herbarium collections, scientific literature and official reports to mention some. Presences refer to suitability but also to the realized niche. On the topic of species distribution models Lobo et al., (2010) remarks the conceptual difficulties of trying to approach the potential distribution via modelling only with presences, as those presences are the realized distribution. He reviews that absences are highly informative but naïve to approach as most are if they even exist- are part of non-standardized surveys or protocols. He also remarks that threats can be transcribed as contingent absences since they refer to areas that are within the fundamental range prevent the presence of the species. However, tough absences provide more information than the lack of suitability of an area; true absences are difficult to obtain and so pseudo absences that cover the full range of the study have to be created (Lobo, Jiménez-Valverde, & Hortal, 2010).

Species distribution models tie to more intricate ecology concepts such as potential distribution and realised distribution to different geographic scales. Species distribution models are used in Conservation to nurture different types of analysis such as species decline, conservation management plans and report conservation action, to mention some (Stephanie, Ceri, & S.J., 2001).

General linear models (Glm) are a family of different models that generate a linear response to the explanatory variables. Glm as mechanistic models capable of predicting response from hypothesised causal relations. This hypothesis links the environmental variables in a continuous environment to a binary relationship. In species distribution models, the environmental variables are the continuous or categorical inputs, and the binary relationship is the presence (1) and the absence (0). The hypothesis is the possible explanation or inference of the presence or the absence given habitat conditions, or a constrain, that can be a threat to the species or other biotic or abiotic factors. For species distribution modelling Glm is widely used as they are straightforward (Guisan & Zimmermann, 2000).

1.4. Pixel size, scale and resolution

Nowadays, technology allows us to emulate natural patterns, process and phenomena in maps to great detail. However, not all phenomena can or should be mapped at the same minimum resolution. Scale and patterns are defined by smallest grain, cell or pixel understood given the "resolution at which patterns are measured, perceived or presented" (Johnson, 1980). Species react to orders of the selection of habitats that goes from geographic range, home range and local niche furthermore the response of the species to an environmental variable is different between orders and scales. The problem of pattern and scale is a central issue in ecology as different landscape mechanism have different effects at different scales (Levin, 1992)

The effects of the extent and pixel resolution on species distribution models have been reviewed in the literature for specific species by diverse authors: Connor et al., (2019) mentions is commonplace to select the grain size of a study are a priori based on expert opinion or in the data available for the modelling. He reviewed for the giant panda possible presence, model accuracy and sensitivity are affected by the extend and pixel size from 30m to 1000m. Song, Kim, Lee, Lee, & Jeon, (2013) increased the pixel size at the same time that increased the extent of the study area to maintain the scale relationship. He reviewed how the loss in accuracy in fine pixel sizes could be overcome by increasing the extent and suggested further explicit investigation on the effects of pixel size on model predictions. Given a minimum initial pixel size was 30m per 30m till 5km 5x 5km, expanding the extend size on habitat suitability analysis improved the accuracy of the prediction for the Korean water deer.

Umetsu, Paul Metzger, & Pardini, (2008) studied the response of modelling small forest mammals at different scales highlighting it is relevance for analysing the influence of landscape structure given that "despite its small size, small mammals perceive landscape features at large scale" and so are sensitive to landscape structure and so reviewed its response to landscape fragmentation within the species range of movement in 50, 100, 200 and 400m.

Haby et al., (2018) reviewed the sensitivity and response to landscape scales of small mammals were reviewed as a way to improve the performance and transferability of species distribution models by using local topography and land use characteristics. By including into the model biotic and abiotic factors that directly influenced the species niche. He also warns when using data to extrapolate the realised niche from one region to another or to expand the area of the study area, especially when relating it to climate data. A recurrent issue that these Authors faced is that when reviewing the possible presence of a species, they all used only environmental data and restricted the incursion of threats into the model reviewed.

1.5. Problem statement

Some Authors argue about the vast range of techniques, input data and scales may be misused-or misinterpreted by the conservation efforts "maps of the same area built using different methods and data may provide dissimilar representations, thus providing different information and possibly leading to different decisions" (Lecours, 2017). As conservation efforts do not always specify standards or guidelines for the production of maps, results can vary dramatically. Any manager or stakeholder has to understand before planning and implementing actions, the accuracy and uncertainties of the available data (Host et al., 2019). Another issue is that the final product does not always have feedback from the manager of the resource. The consequences making conservation decisions based on a misinterpretation of maps include loss of trust in science, conservation plans or management plans that do not reach their objectives, financial costs and temporal costs (Guisan et al., 2013; Lecours, 2017).

The Natura 2000 in the European Union is an ecological network of protected areas, within it is member states it oversees Sites of Community Importance (SCIs) and Special Protection Areas (SPAs), that have been chosen to safe keep key species and habitats (92/43/ECC). The Network has an umbrella effect as it benefits all species groups, even the ones which Natura 2000 areas were not specifically designated (Jones-Walters et al., 2016). For the Protected Areas in the Network is mandatory to report the List of pressures and threats and conservation measures with specific guidance on the use of distinct pressure and measure codes. Likewise, it is required to identify the human activities by land uses that can influence or jeopardise the conservation status of the key habitats and species in the Protected Area by implementing programs and positive actions. The list has to be updated every six years and be available to the public in the Standard Form (92/43/ECC). Neither the report nor the Standard Form includes mapping the threats.

The European Union reports the presence of its biodiversity at 10 x10 km resolution; this coarse grid works to set and report policy at a European or national scale. Sales-Luís, Bissonette, & Santos-Reis, (2012) mentions that the over usage of this scale for the Mediterranean otters overestimates local trends and mislead local Protected Areas management, and could jeopardise the success of local actions. So, how to map a species for management and conservation efforts accurately? The Autor will map a species considering threats in a set region of interest modifying the size if the minimum unit pixel resolution and measure the accuracy of the models (Figure 1).

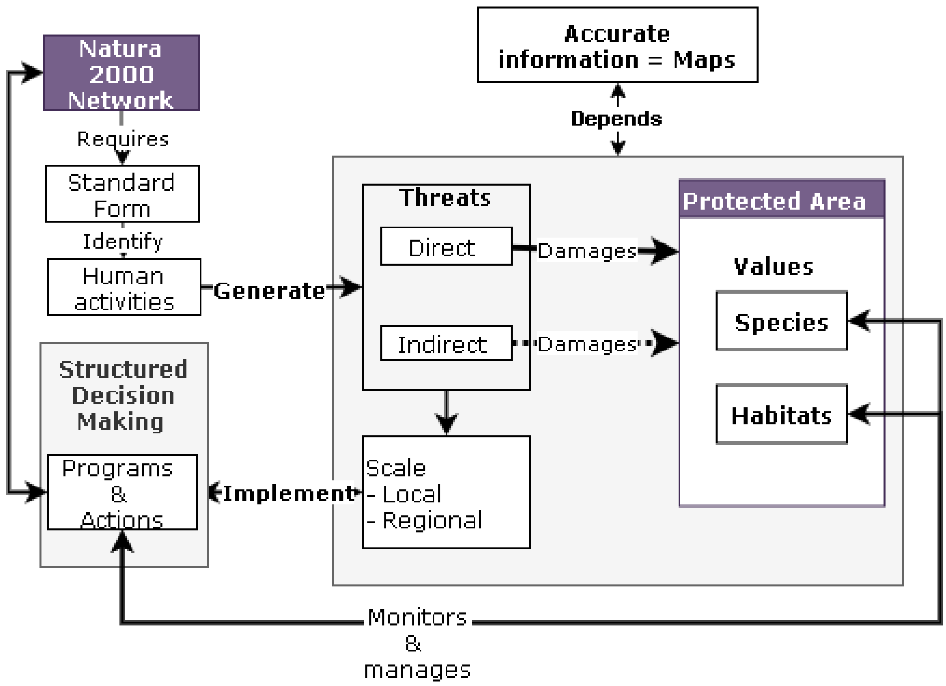


Figure 1 Conceptual framework

1.6. European ground squirrel ecology

The European Ground Squirrel or European Souslik (*Spermophilus citellus*) is a small mammal of the Rodentia family, is endemic to central and south-eastern Europe, in altitudes of 0-2,000m. Greece is part of the southern range distribution limit. This small-sized mammal has a length between 18 and 24 cm lives in small colonies. The squirrel is diurnal with an obligatory invernation period. The European Ground Squirrel not social but burrows in family groups and does not go far away from the burrow, as its daily movements go from 60 till 100m, only during reproduction season the males would travel more, this is the home range of movements. It prefers grassland habitats, seminatural and artificial, with a preference to grasslands with permanent short-stalk vegetation (10-20 cm) with light and well-drained soils, and medium aeration that allows them to build their burrows tunnel systems. Additionally, its presence has been linked to the total yearly precipitation and temperature (Ramos-Lara et al., 2014).

This priority species is listed by the Red List as vulnerable (IUCN, 2010). The European Union, on its action plan for the species, identifies significant threats for this species: agriculture, habitat fragmentation, urban development, poor water and grassland management. In the Mediterranean region, it is conservation status unfavourable, as historically the species was treated as a plague for the primary production sector. The main threats are agriculture, habitat fragmentation, urban development, poor water and grassland management (Janák, Marhoul, & Mateju, 2013).

The European Ground Squirrel in Central Macedonia has coexisted with the traditional primary production systems since in these areas they could find grassland and small landscapes feature to construct its burrows. As the agriculture practices shift into more intensive models, its habitat has decreased isolating the species population and cornering it to the vegetation in the borders of crops (Youlatos et al., 2007).

1.7. Study Area

Greece is a heterogeneous environment. Localised in the middle of the Europe Mediterranean biogeographical region on the Balkan Peninsula, its unique weather and rough landscape highlight it as an ecological hotspot. In the terrestrial animal kingdom, it has 115 recorded mammal species, 446 birds, 22 amphibian and 64 reptiles to mention some (IUCN, 2019) Dimopoulos, Bergmeier, & Fischer, in (2006) reports 85 habitat types occurring within the total number of Natura 2000 and 30 relevant to the nation habitats. Additionally, the country hosts 6,600 taxa of vascular plants from among them 22% are endemic.

The administrative region of Central Macedonia (Κεντρική Μακεδονία, Kentriki Makedonia) is located at the northern continental portion of Greece. The region has a rough terrain, most of is an extension is included within the Central Macedonia basin. The Axios river flows down from the hilly mountains that reach altitudes that range from 2400m to 1800 m, into the valley with to the Thessaloni plains 750 m to 250 m, into the river delta (250m to 0m) then discharge in the Bay of Thermaikos. The weather is the Mediterranean with a mean annual air temperature vary between 9 °C and 17.5 °C, and annual rainfall ranges from 400 mm to 1,300 mm (EU WISE, 2019; INWEB, 2004).

Figure 2 shows the landcover in the study area. The vegetation in the upper part includes coniferous forest, broadleaf forest, mixed forest, transitional woodland shrubs and natural grasslands and sclerophyllous vegetation. As altitude decreases, it changes into a system of river estuaries, swamps, lagoons and salt marshes. Intertwined on the natural landscape, the agriculture land cover is spread in the form of Non-irrigated arable land, permanently irrigated land, rice fields, vineyards, fruit trees and berry plantations, olive groves, pastures, annual crops associated with permanent crops and intricate cultivation patterns (Copernicus, 2018; Vokou, Giannakou, Kontaxi, & Vareltzidou, 2016).

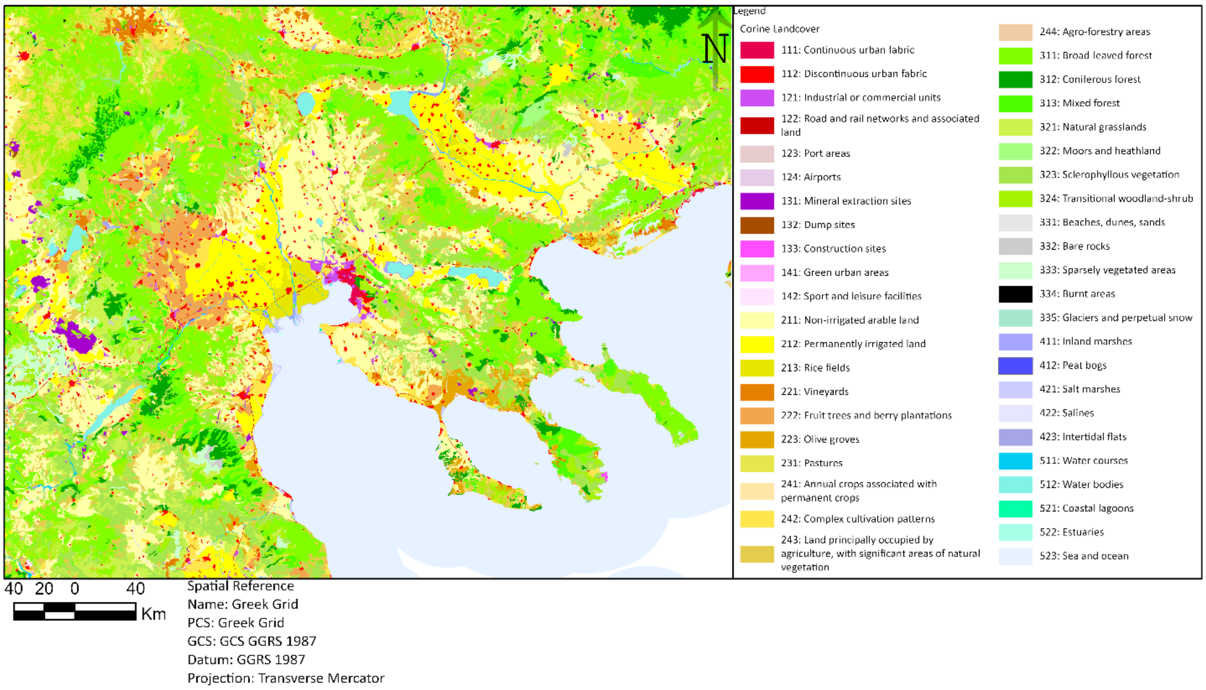


Figure 2 Corine Land cover map of the study area.

Central Macedonia population is the second-largest in the country. Almost 1.9 million inhabitants are distributed on Thessaloniki, Serres, Katerini, Veroia, Giannitsa, Kilkis, Naousa and Edessa. The main income comes from the primary production; this region contributes with 26% in gross value added to the primary sector for the country, other activities such as industry and manufacturing are on the rise (Synergassia regional Partnership & Enterprise Greece, 2018). A renewed interest in recent years on developing the region has fuelled different changes in the landscape such as the construction of roads, the intensification of agriculture and the construction of dams, and river channelling for agriculture these activities happening both inside and outside Protected Areas (Maragou & Mantziou, 2000).

The region host 73 Protected Areas included the Natura 2000 network (Figure 3). The Protected Areas in the Natura 2000 Network that include the European souslik (*Spermophilus citellus*) in its conservation objectives are Sites of Community Importance on the Habitats directive Delta Axiou - Loudia -Aliakmona – Evryteri Periochi - Axioupoli (GR1220002) and, Limnes Volvi Kai Lagkada – Evryteri Periochi (GR1220001).

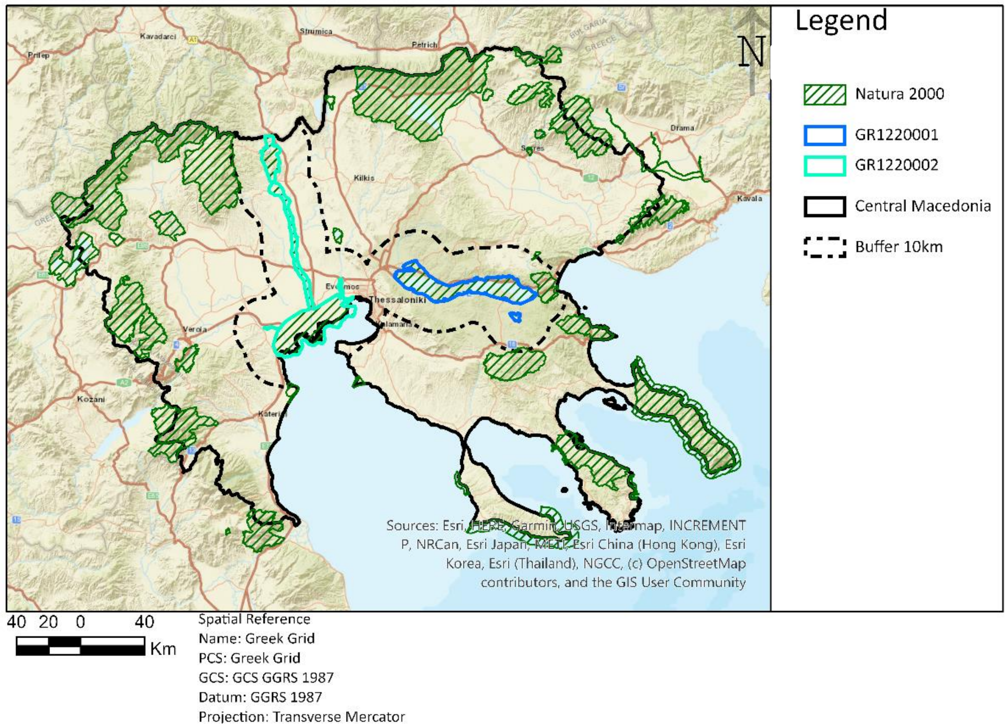


Figure 3 Protected Areas of the Natura 2000 within the Central Macedonia Region in green. Central Macedonia in continuous black line is, the blue and teal polygons are the Protected Areas, the dotted black line is the 10km buffer area around the selected Protected Areas.

1.8. Objectives and research questions

The main objective of the present study is to address the influence of threats and resolution in modelling the distribution of European ground squirrel in the Region of Central Macedonia in Greece.

1.8.1. Subobjectives and research questions

I. To review the environmental factors and threats linked to the distribution of the ground squirrel.

I.a. Which environmental variables can be linked to the distribution of the squirrel?

I.b. What are the threats to the squirrel and how they vary in space?

II. To model the distribution of the ground squirrel and assess model outcomes with and without the use of threat maps

II.a. How accurate can squirrel's distribution be modelled based on environmental variables?

II.b. What is the effect of threats on the squirrel distribution model?

III. To compare the modelled distribution of the ground squirrel at different input resolutions

III.a. What is the effect of different resolutions on the accuracy of the squirrel distribution model using the same environmental variables?

III.b. Does the importance of environmental factors and threats change with changing resolution?

IV. To compare the distribution of the squirrel and its threats inside and outside the Protected Areas.

IV.a. Is there a difference in the distribution of the squirrel inside and outside the Protected Areas?

IV.b. Are spatial patterns of threats the same inside and outside of the Protected Areas?

2. RESEARCH DESIGN AND RESEARCH METHODS

2.1. General Methodology

The methods employed in this study begin with the review the ecological requirements for the ground squirrel to relate them to environmental variables and spatial data; The next step to review the reported threats and generate maps from the threats. Subsequently, Species Distribution Models were generated with the selected environmental variables and mapped threats. Then the accuracy of the models was calculated. After this, the models were fitted at different resolutions. Finally, the accuracy of the models and the distribution of the species inside and outside of the Protected Areas were compared. The flow chart of the methods followed is presented in Figure 4.

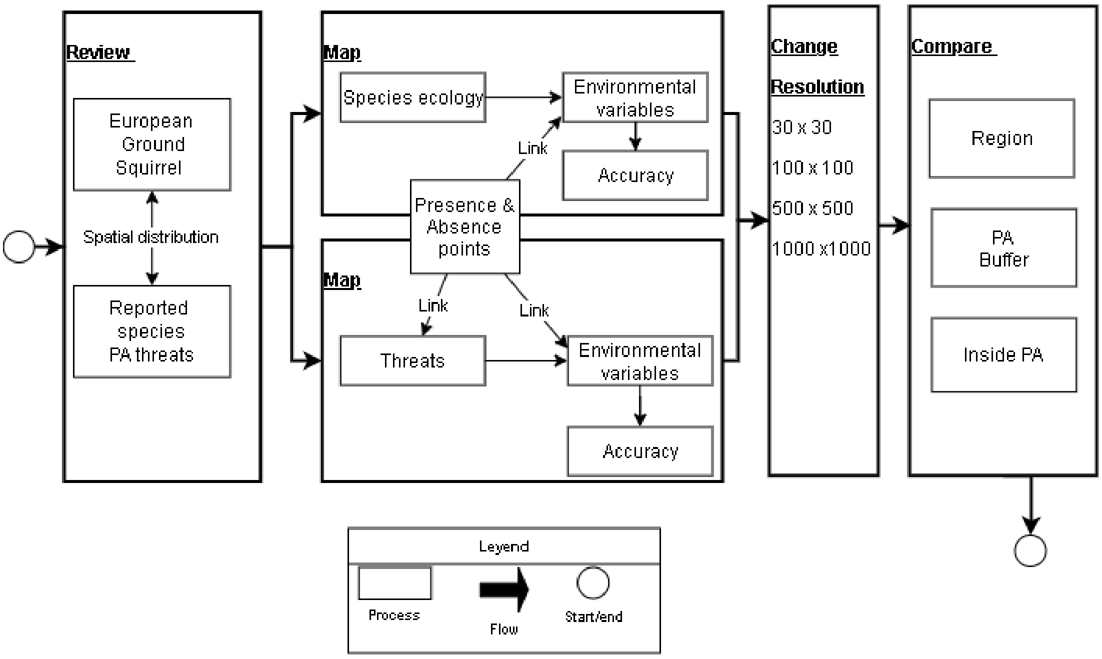


Figure 4 General methodology flowchart

2.2. Explanatory variables data

The environmental variables were collected based on the review the ecology of the species and as well on the data available (Table1). Climate is one of the major constrains for the distribution of species (Hooper et al., 2005). The first group of environmental variables the bioclimatic variables Annual Mean Temperature and Annual Precipitation. Since this is a global product interpolated from on the ground stations with the SRTM global digital elevation model (Fick and Hijmans, 2017) more local data is required. Climate and topography are deeply linked, the products derived from the Elevation Digital Model derived from SRTM, ASTER GDEM and Russian topographic maps are altitude, slope and aspect in continuous raster generated for all the European Union (European Environment Agency, 2016). This is supplemented by reviewing the soil dominant parent material (Panagos et al., 2012) and since both of the protected areas include water bodies the water and wetness thematic product that shows the occurrence of water assessed separately of the vegetation only in the physical sense and not limited by the land cover (EEA, 2015). Additionally, since the area has prevalent agriculture land cover, the Corine Land cover of 2018 will be used. This is a consistent thematic land cover product with 3-level hierarchical classification nomenclature system generated from satellite classification imagery and verified by the member state. The land cover map homogeneous landscape with more than 75% of the pattern that has the characteristics of a preselected class, we used the first and third level of the hierarchical classification (Corine, 2010) (See ANNEX 2 for the classes).

Table 1 Environmental variables selected to describe the species ecology

	Variable	Data type	Unit	Original resolution	Projection
Climate	Altitude	Continuous	m	25 m x 25 m	ETRS89-LAEA (EPSG code 3035) Ellipsoid GRS80 Vertical datum EVRS2000 with geoid EGG08
	Aspect	Continuous	degree		
	Northness	Continuous	degree		
	Eastness	Continuous	degree		
	Slope	Continuous	degree		
	Annual Mean Temperature	Continuous	°C	1 km x 1km 30 seconds	UTM WGS84
	Annual Precipitation	Continuous	mm		
Land cover	Corine land cover 2018	Categorical	code	Minimum Mapping Unit 5 ha	ETRS89-LAEA (EPSG code 3035) Ellipsoid GRS80 Vertical datum EVRS2000 with geoid EGG08
Soil	Soil parent material	Categorical	code	1 km x 1km	ETRS89-LAEA (EPSG code 3035) Ellipsoid GRS80 Vertical datum EVRS2000 with geoid EGG08
Water	Water and wetness	Categorical	code	20m x 20m	EPSG:3035 (ETRS89, LAEA)

2.3. Environmental variables pre-processing

The environmental variables were generated for the 30 x 30m resolution. The pre-processing included:

- **Corine land cover and Combined Corine land cover.** Using the third level of the nomenclature, the shapefile was rasterized to the initial resolution of 30m x 30m. Additionally, since this data is categorical, it was necessary to generate a Boolean attribute table to test if its variables could be included in the model and to rasterise each one. From this file also a combined Corine land cover Boolean raster was created including the third level of the classification that had the majority of the presences given the home range of movement of the species.

- **Sea and water bodies mask.** A Boolean raster mask was created by dividing the first level of nomenclature of the Corine land cover into two, aggregating all classes that wherefrom terrain, (1 artificial surface, 2 agriculture areas, 3 forest and seminatural areas, 4 wetlands) and the ones that had water: 5, Water bodies. The terrain was given the value of 1 and water bodies the value 0.

- **Distance from inland water, Distance from the sea, Distance from all water.** Euclidean distance raster were created with the first and third level of nomenclature of the Corine land cover. For distance to inland water the categories watercourses, estuaries, water bodies and coastal lagoons. For the distance from the sea, the category sea and ocean. Finally, the distance from all water used all the main group waterbodies.

- **Altitude.** From the digital elevation model, by using raster algebra first, 1m altitude was sum then, the product of the sum was multiplied per the Sea and water bodies mask—this way masking the values from the Sea to be able to distinguish the lowlands.

- **Slope.** The conversion of digital number (DN) into degrees off horizontal of the surface tangent was done following the provider specifications:

  slope[degrees] =float (acos (DN/250.0) * 180.0/! PI)

  Additionally, by using raster algebra first, 1m altitude was sum then, the product of the sum was multiplied per the Sea, and water bodies mask, this way masking the values from the Sea to be able to distinguish the lowlands.

- **Aspect.** The orientation of the pixel. A trigonometric transformation of aspect data, conversion to northness and eastness

  $$\text{northness} = \cos(\text{aspect})$$
  $$\text{eastness} = \sin(\text{aspect})$$

The software used for the pre-processing of the environmental variables is ArcGIS Pro 2.5.1 fitting, accuracy assessment and creation of the predictions is R studio Version 1.2.5042. The packages used are Glm (Hastie & Tay, 2020) and Sdm (Naimi & Araújo, 2016). The total list of categorical and continuous variables is in ANNEX 2.

2.4. Species Dataset

The presence points originate from data on species occurrences that were collected and provided by experts during the 3rd National Report on the implementation of the EU Habitats Directive 92/43/EC. Data consists of confirmed and accurate data found in literature complemented by field observations that took place in the summer seasons of 2013-14. The original database included 349 observation for Spermophilus citellus, to be able to measure the accuracy of the model, aleatory absence points were created in a rate of 3:1, so 1047 absence points were added into the database. The absence points were located using the third classification level of the Corine land cover, using the classes that according to a literature review and the frequency of occurrence of the presence points, were less likely to have presences.

The new database that included the presences and absences was randomly subset into two files, one with 70% of the observations to fit the model and the other with 30% to measure the accuracy. The environmental variables and the representation of the threats were stacked, and the values were extracted into the absence and presence points. The point data were further review for consistency, and ninety-eight points were removed as they were falling into the sea part of the map. Since there were no records of how the presence points were taken, it was decided not to relocate them manually. It was deemed relevant to remove them as the final result was going to be a prediction on a map and to minimise the false positives. This procedure was repeated with the test reserved points for the accuracy assessment.

2.5. Threat mapping

In the Natura 2000 Network in Central Macedonia, one of the major threats reported for the European ground squirrel is the reduction or loss of specific habitat features having high occurrence inside and outside the Protected Areas Limnes Volvi Kai Lagkada Evryteri Periochi (GR1220002)and Delta Axiou Loudia Aliakmona Evryteri Periochi Axioupoli (GR1220001). This last Protected Area hosts one of the most important colonies in the country. With an estimated population between 100-300 adults, is roughly one- a quarter of the registered European ground squirrel burrows. As the area contains pastures in an excellent state of preservation; these grass are getting scarce in the region, local farmers mention that the European Ground Squirrel used to be much more common (N2K GR1220002 data form, 2019).

The threats were all initially geocoded to a common 30x30m resolution. The threats that were reviewed in this work are:

- Roads and motorways,
- Discontinuous urbanization
- Reduction or loss of specific habitat features
- Anthropogenic reduction of habitat connectivity

In ANNEX3 all the threats, pressures and activities with impacts on the site reported for the study area.

2.5.1.1. Materials and data for the threats

The threats were created based on the review threats reported for the species and the relevant data available (Table 2). Threats were geocoded to the same extent, resolution and projection as the environmental factors explained in the previous section.

Table 2 Threats and the original data selected to generate the maps.

<table>
<tr><th>Threat</th><th>Geo representation of the threat</th><th>Original data</th><th>Data type</th><th>Original resolution</th><th>Projection</th></tr>
<tr><td rowspan="4">E01.02 Discontinuous urbanization

J03.01 Reduction or loss of specific habitat features

J03.02 Anthropogenic reduction of habitat connectivity</td><td>Area with changes of land cover on the last 20 years</td><td>Land cover change for the periods 1990-2000, 2000-2006, 2006-2012, 2012-2018.</td><td rowspan="2">Categorical</td><td rowspan="2">Minimum Mapping Unit 5 ha</td><td rowspan="2">ETRS89-LAEA (EPSG code 3035) Ellipsoid GRS80 Vertical datum EVRS2000 with geoid EGG08</td></tr>
<tr><td>Continuous urban fabric
Discontinuous urban fabric</td><td>Corine land cover 2018</td></tr>
<tr><td>Total population</td><td>GEOSTAT population grid from. Total population from 2011 Census</td><td>Continuous</td><td>1 km x 1km</td><td>ETRS89 LAEA</td></tr>
<tr><td>Total area cover density of small woody features
Distance from small woody features</td><td>Small woody features</td><td>Continuous</td><td>5 m x 5 m</td><td>EPSG:3035 (ETRS89, LAEA)</td></tr>
<tr><td>D01.02 Roads, motorways</td><td>Density of are covered by all roads or communication routes</td><td>OpenStreetMap database from Geofabrick</td><td>Shapefile</td><td></td><td>WGS84 EPSG:4326</td></tr>
</table>

	Distance from railways Distance from roads of 5m, 10m, 15m and 20m width.				

The threats were generated for the 30 x 30m resolution, using either distance from the origin, density or Boolean (have or not have the threat). The threats reviewed are:

Roads and motorways. Roads are one of the main landscape human-made structures that fragmentates the landscape (Fahrig & Rytwinski, 2009). They are closely linked to human activities, promote land use changes and so the reduction of specific habitat features. The data used to map these threats originates from the OpenStreetMap database from Geofabrick gathers both official and non-official mapped features aggregated in attributes; it updates daily its records(Ramm et al., 2019).

The development of the threat representation include to divide the roads per class given the width in meters, roads with less than 5 m were not included. The with was generated by a rough average using Google Earth (ANNEX1), then rasterised into 5 x 5 m. Then, the number of pixels was extracted into the 30x30 raster to calculate the density; this was generated for each road and railway. The distance to roads and railways was calculated from the original shapefile raster file by Euclidian distance from the feature.

Discontinuous urbanization

Discontinuous and continuous urbanizations are human settlements with different characteristics, mainly the total population and productive activities. Discontinuous urbanization are usually isolated, non-dense populated areas surrounded by mainly agriculture and natural lands. Continuous urbanizations are densely populated settlements that can combine residential, commercial activities, among others. The Continuous urban fabric and Discontinuous urban fabric form the Corine land cover will be used.

For the development of the threat representation, it will be using the GEOSTAT population shapefile grid and the total population from the 2011 Census, linked by Nomenclature of Territorial Units, NUTS-3 level code: small regions for specific diagnoses a continuous raster was created. For the areas that have no population, the value was set to cero.

Reduction or loss of specific habitat features and Anthropogenic reduction of habitat connectivity

Both threats are closely related, but as connectivity requires firs to known the local habitat they were reviewed together. From the Corine Land cover product described in section 2.3.1 for the periods 1990-2000, 2000-2006, 2006-2012, 2012-2018 the Land cover change is one sub-product from the classification from those periods. This land cover change is the only threat that considers the past and is included because once a land cover is changing the original relationship with the species is either lost or modified. Since the area has a dominant agriculture main land cover using an agriculture classification as a threat is not practical, but a second option is to use the density of small woody features and the distance from the woody features. The small woody features product from 2015 is a non-validated product from that harmonizes information on small landscape woody features such as hedgerows, as well as patches. These features are relevant in farm and agriculture lands as wildlife uses it. The GEOSTAT population grid par to the European Union statistical framework from to consolidate local and regional data within its Nomenclature of Territorial Units NUTS-3 level code: small regions for specific diagnoses, it's been combined with the Total population from 2011 Census (Eurostat, 2011).

The development of the threat representation includes the use of Land cover change subproduct; a Boolean layer was created for the areas that had any land cover change from the last 20 years. In addition, the density of small woody features in percentage will be created by extracting the total covered area in the base 30 x30m raster between the total area of the raster. The distance to small woody features was calculated from the original 5m raster file by Euclidian distance from the feature.

The software used for the pre-processing of the threats is ArcGIS Pro 2.5.1 fitting, accuracy assessment, and creation of the predictions is R studio Version 1.2.5042. The packages used are Glm(Hastie & Tay, 2020) and Sdm (Naimi & Araújo, 2016).

2.6. Ground squirrel distribution modelling

The significance of each variable to be included in the model was defined by reviewing:

Boxplot of the environmental variables spread

The boxplot graph is a statistical aid to describe the spread of the data between the range of its max and minimum values; it shows within its quartiles the percentage of observations distribution0 (Manly, 2009).

Pearson correlation

For the continuous variables, Pearson correlation is used to measure the degree of variation within a paired variable that can be explained by the other, so how much X change can be explained by Y. Correlation describes a linear relationship with values between +1 (positive) and -1 (negative). Correlation as a graphic also aids to identify possible collinearity to detect correlation issues. The strongest correlation will be 1 and no correlation 0 (Manly, 2009).

Frecuency if Presence

For the categorical variables, the categorical relation was defined as a boolean, 1 being have and 0 not have; Frequency is the total of occurrences of a categorical variable within a dataset; then frequency if presence is the total of the presence point that had the presence within the categorical variable (Manly, 2009).

2.6.1. Model fitting with environmental variables

The linear regression describes a relationship between a dependent (Y) and an explanatory variable (X) with the straight-line equation (Y= β0 + β1 X) and to add multiple explanatory variables ($Y= \beta_0 + \beta_1 X_1 + \beta_2 Y_2 + \ldots + \beta_p X_p$). To be able to predict a presence (1) or an absence (0) to combine the linear regression with a logit link function with the logarithm of the odds ratio (Π) is the probability of something happening against the probability that it will not happen:

$Y = b_0 + b_1X_1$ Becomes $f(X_1) = b_0 + b_1X_{1\ldots n}$ and $f(x) = \ln(\Pi) = \ln(Px / 1\text{-}Px)$

By adding the logit function is possible to stretch the probability and then to combine multiple explanatory variables. The straight forward simplicity of the Glm allows it to combine several different explanatory variables and so is widely used in species modelling.

$$P_x = \frac{e^{b_0+b_1x_1+b_2x_2}}{1+e^{b_0+b_1x_1+b_2x_2}} = \frac{1}{1+e^{-(b_0+b_1x_1+b_2x_2)}}$$

Modelling in the R environment Glm fit is the models via penalized maximum likelihood and extrapolates the values over a continuous raster (Hastie & Tay, 2020).

Furthermore, to reduce the environmental variables to the minimum significative ones that explained the probable distribution:

Significance values Pr |z|
A statistical hypothesis of one-tailed end to accept or reject how significant a value that is normally distributed to include in a model given the z-score. The p-value given for the z score is interpreted in a null hypothesis as if there is enough evidence to accept or to reject the statement. If the value of $P|z| >$ significant value then, there is evidence to reject the hypothesis if it is lower then there is not enough evidence to reject the hypothesis, so it is accepted (Manly, 2009). The value defined to be significant was '.' 0.1.

Variance inflator factor (VIF)
This factor provides information about multicollinearity issues, as it aids to detect variables that repeat information within the model (if that is not the desired case). It provides an index that measures how much the variance increases with a given set of predictors.

Stepwise regression and Akaike information criteria (AIC)
A regression estimates how close are the predicted values to the original data given a specific criterion. Stepwise regression for a model with multiple variables consists of removing variables in backward and forward direction of a model until the AIC gets to the lowest value. AIC is an estimator given the maximum likelihood of how good a model is fitted, and so is used as a goodness indicator of fit by minimising the number of explanatory variables of a model. AIC is used to reduce possible collinearity problems caused by including in the model variables that describe the same phenomena. Is not, however, the first filter to choose the environmental explanatory variables but a statistic test to improve the performance of the model (Velasco & González-Salazar, 2019).

2.6.2. Model accuracy assessment

In area-based studies, the spatial autocorrelation of the variables neighbouring sample unit tends to possess similar characteristics is a potential problem. Some traditional methods to measure the accuracy and error of a predictive model are: Sensitivity is the conditional probability that case X is correctly classified $p(X_{Alg} / X_{true})$, specificity is the inverse $p(X_{true} /X_{Alg})$, and negative predictive power that assess the probability that a case is not X, $p (X_{False}/X_{Alg})$. The receiver operating characteristics ROC is obtained by plotting all sensitivity values (true positive fraction) on the y-axis against their equivalent (1-specificity) values (false positive fraction). The area under the ROC function (AUC) is a threshold independent accuracy test that plotting all the "possible thresholds avoid the need for a selection of a single threshold" (Fielding & Bell, 1997), providing a single measure of overall accuracy and performance. Threshold independency is relevant because by setting an a-prior threshold of what will be considered presence, the model could be biased from the beginning.

ROC AUC is a widely used accuracy method, Jiménez-Valverde, (2012) warns about the use of ROC AUC for predicting presences since in practice given that it penalises the models that estimate potential Distribution and favours the realised Distribution, to overcome this issue the Author includes the Kappa and TSS.

Kappa is a robust indicator to compare presence-absence models performance. It can be generated for threshold dependant and threshold independent. Kappa statistics, also called proportion of specific agreement, ranges from -1 to +1, where 1+indicates perfect agreement and values of zero or less indicate a performance no better than random. This way to determine if the predictor is better than the chance, as

Kappa is the proportion of specific agreement 30 (Fielding & Bell, 1997). Stephanie et al., (2001) mentions that for ecology it could be used values for kappa of (0.0-0.4) indicate slight to fair model performance (0.4-0.6) moderate, (0.6-0.8) substantial and (0.8-1) almost, perfect. Kappa includes both omission and commission errors; it is sensitive to prevalence (total of true positives plus the total of false positives between the total predictions) the Author will supplement the analysis reviewing the True skill statistics (TSS), also known as the Hanssen-Kuippers discriminant, is defined as (sensitivity + specificity -1) (Allouche, Tsoar, & Kadmon, 2006).

2.6.3. Link the distribution of the threats to the species to the environmental variables

All the generated threats were reviewed following the same procedure as the one with the variables before; For continuous threats review the boxplot of the threats and the Pearson correlation. And, for the frequency of the categorical threat of the species points.

2.6.4. Model fitting with environmental variables and threats

The fitting of the model with threats included using all the environmental variables before the stepwise regression, then adding all the threats. To define significance to the model given the values Pr $|z|$, Variance inflator factor (VIF)and Stepwise regression and Akaike information criteria (AIC).

2.6.5. Accuracy assessment of the model fitted with variables and threats

To be able to compare the model fitted only with variables and the model that included variables and threats, the same accuracy assessment method will be used the ROC AUC, TSS and Kappa.

2.7. Modelling at different resolutions

The initial dataset on 30x30m resolution, after mapping the environmental variables and the threats, it was necessary to project and resample all the resulting data. The pre-processing of the data included, all the environmental layers were projected to GGRS87, EPSG:2100. A polygon for extracting all data s was created to include the Administrative region of Central Greece, covering an extension of 69,523 km^2, then the polygon was converted into different pixel resolutions (Table 3).

Table 3 Bounds of the study area polygon to extract all geographic data.

ID	X	Y
1	223318.473	4600284.866
2	545548.875	4599427.869
3	544975.062	4383673.785
4	222744.660	4384530.782

From this polygon four raster's were created with pixel sizes of 30m x 30m, 100m x 100m, 500m x 500m and 1000m x 1000m, these rasters were used to extract and project by nearest neighbour the environmental layers to the different pixel sizes, using 30x 30m to 100, then from 30 to 500 and from 30 to 1000. The pre-processing of the raster and shapefiles was done in an ArcGIS geodatabase environment then exported to TIFF.

2.8. Comparison of the modelled probable presence in the Study Area

In order to compare the effects of PAs on the various modelled outputs, three sub-zones within the study area were delimited compare the probable presence in of the ground squirrel and the threat's distribution. By the use of the boxplot to explore and compare the differences in the predicted presence within the three sub-zones. Additionally, the Protected Areas of the Natura 2000 Network in the Central Macedonia Region were reviewed for the probable presence within its limits by comparing their maximum, minimum and mean of the modelled prediction.

Since threats have the spatial characteristic of being direct or indirect, a direct threat will be anything inside the Protected Area that is reducing he probable presence and as well anything in the adjacent zone to the Protected Area. And an indirect threat-The region of central Macedonia. This way, the three subzones to review (Figure 5) are:

- **Central Macedonia:** Administrative region.

- **10km buffer:** A buffer of 10km of the area around the Protected Areas. The Protected Areas have a linear distance between each other of 15 km, and so a buffer area surrounding them of 10km could encompass interactions.

- **Protected Areas**: Includes the Protected Ares Delta Axiou Loudia Aliakmona Evryteri Periochi Axioupoli (GR1220002) and Limnes Volvi Kai Lagkada Evryteri Periochi (GR1220001). These Protected Areas are only ones that, in the region report the Ground Squirrel presence within its borders.

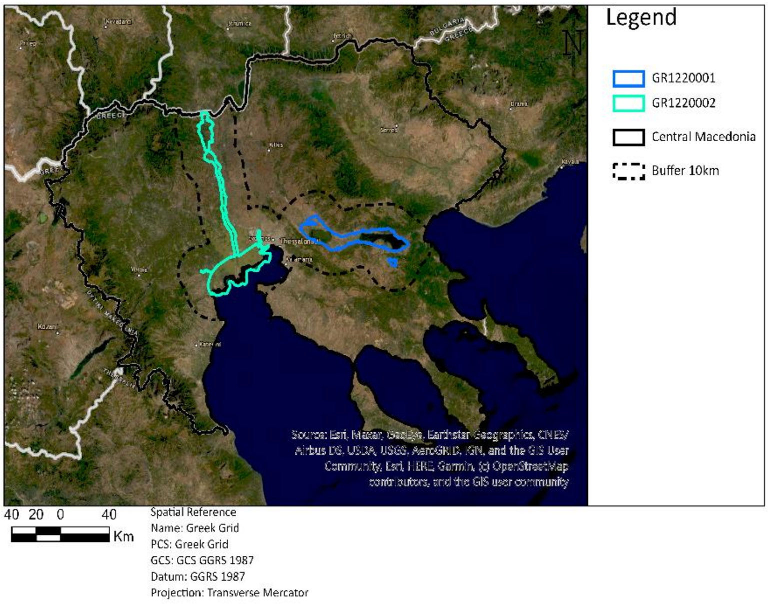


Figure 5 Zones to assess. The black continuous line is the Central Macedonia area, the blue and teal polygons are the Protected Areas, the dotted black line is the 10km buffer area around the Protected Areas.

3. RESULTS

3.1. Environmental variables linked to the distribution of the squirrel

A first a home range review of the land cover in the vicinity of the observations of the European ground squirrel in Central Macedonia, shows the majority are found in salt marshes and rice fields. Given the movements of the ground squirrel in the study area, the possible different land cover that could have the species include a variety of land covers that go from agriculture to natural land cover (Figure 6). The presences points are clustered inside the Protected Areas GR1220001 (20.4%), GR1220002 (68.5%), GR1220012 (0.5%), GR1250004 (0.9%) with some exceptions at the Border of Central Macedonia and a group between the Thessaloniki city and the Airport.

	Home range class count		
Corine landcover class	**50 m**	**80 m**	**100 m**
Discontinuous urban fabric	5	5	8
Industrial or commercial units	4	4	4
Road and rail networks and associated land	6	7	8
Airports	2	2	2
Non-irrigated arable land	33	32	28
Permanently irrigated land	21	22	23
Rice fields	40	43	66
Fruit trees and berry plantations		1	1
Mineral extraction sites	7	7	7
Airports	5	5	4
Natural grasslands	1	1	1
Transitional woodland-shrub	1	1	1
Beaches, dunes, sands			1
Inland marshes	27	24	1
Salt marshes	75	73	73
Salines	1	1	1

Figure 6 Home range, given movement of the Ground squirrel, generated from buffers from the presence points overlay on the Corine Land cover.

Then by reviewing the environmental variables with the resolution of 30x30, the environmental variables indicate that the ground squirrel presences are located in areas that have an altitude from -2.1 to 833m, mainly flat areas with a slope range from 0° to 14°. On a variety of land covers, such as Non-irrigated arable land (17%), permanently irrigated land had (8%), rice fields (18%), inland marshes with (13%) and Salt marshes (31%). The combined land cover category 44% of the total presences. Areas with no water, gather 50% of all presences. Then, zones with fluvial clays, silts and loams Fluvial clays, silts and loams, concentrated 94% of all presences.

Moreover, the spread of the observation showed that the continuous variables Annual Precipitation, Annual Mean Temperature and altitude had a strong correlation between the pairs. To avoid collinearity issues, it was decided to fit the model with slope and altitude. One of the main reasons was because slope and altitude had a smaller original pixel size of 30x30 against the 1x1km of the climatic data. This was reinforced by reviewed as the variance inflation factors (VIF), as the variables with altitude, slope and annual precipitation and annual mean temperature showed values of 15.5 and 8, at the same time altitude, had 11.36 and slope 1.95. After removing annual precipitation and annual mean temperature, the values for altitude and slope became 2.6 and 1.97 (See ANNEX 4 for correlation and boxplots).

3.2. Threats to the squirrel and their spatial pattern

3.2.1. Roads and motorways

Four distance from roads and one distance from the railways were generated. The low values distance from roads that indicate the roads are close, from the road of 5m and distance to the road of 10m are spread in almost all of the territory. On the contrary distance from railways and distance from roads of 15m and 20m has more differences and the low and high values are more spread and distinguishable. The density of transportation as shown in Figure 7 combines all the previously mentioned transport routes in one raster displaying the location of the roads lower densities are located in natural areas while medium densities in agricultures land and high densities in population centres (See ANNEX 5 for all threat maps).

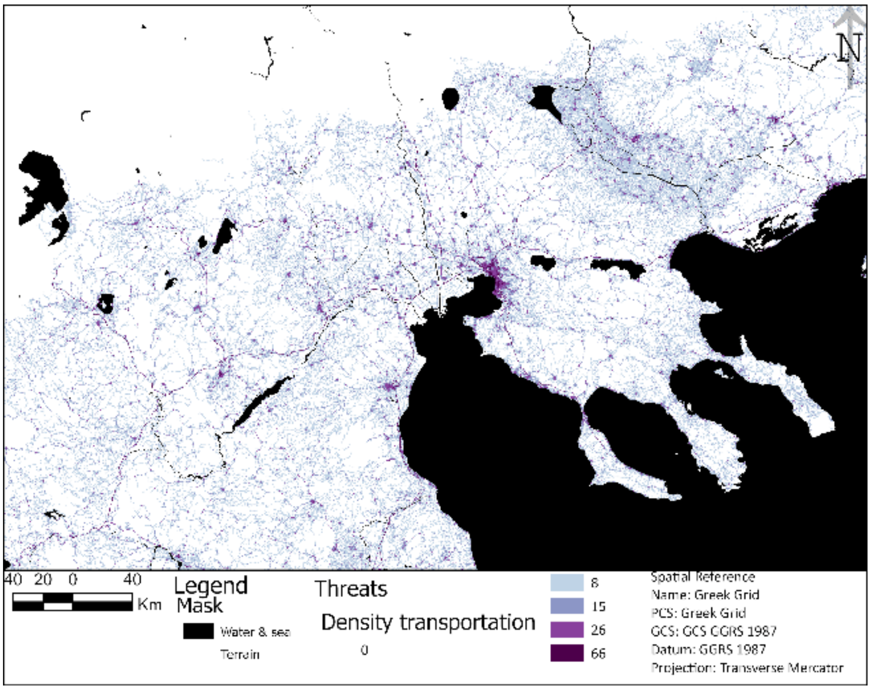


Figure 7 Density of roads. No density on white and highest value on purple.

3.2.2. Discontinuous urbanization

These threats were clustered at the population centres, particularly on the Thessaloniki city areas at the centre and east of the study area. The population centres, and so the density of roads and total population threats follow the same spread. As well for continuous and discontinuous urban fabric.

3.2.3. Reduction or loss of specific habitat features and Anthropogenic reduction of habitat connectivity

The density of small woody features is higher outside of the areas of the continuous urban fabric and discontinuous urban fabric; it is also higher closer to inland water bodies. As it was the case with distance from roads of 5m and 10m, the distance from the woody features has its lower values spread in almost all of the study area; this is because the original threat is present in all the territory extension. Changes within the land cover in the past 20 years were present in all the study area with a tendency either to change the land cover to natural or agriculture, before 2012, then after this period, change into agriculture land cover decreased and was surpassed by a change into artificial. These areas have experimented changes in its land cover since 1990 (figure 8 and 9).

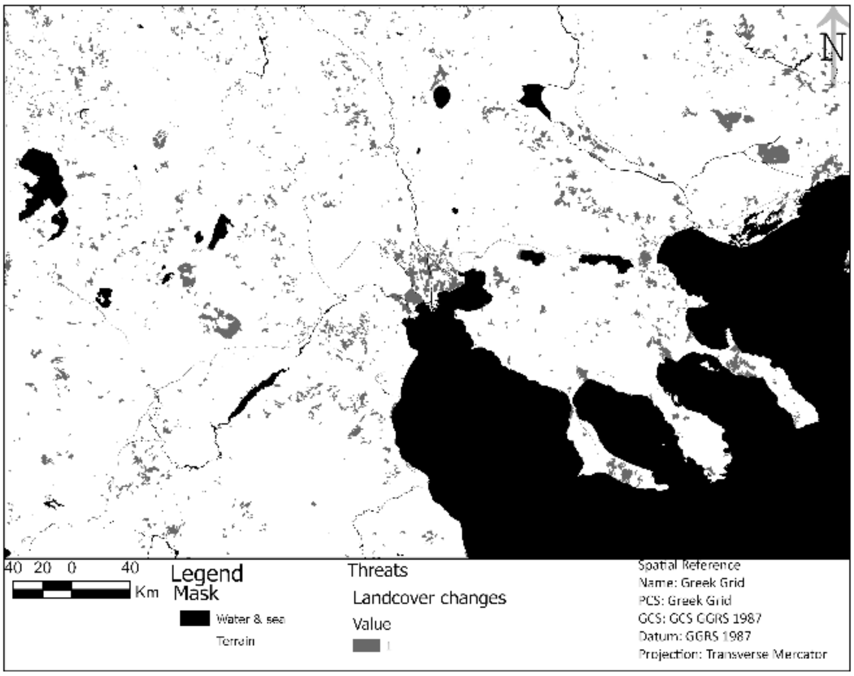


Figure 8 Land cover changes in the past 20 years in grey

	AREA km²			
	1990-2000	2000-2006	2006-2012	2012-2018
AGRICULTURE	3.84	1.24	0.44	0.2
ARTIFICIAL	1.03	1.16	0.74	0.5
NATURAL	5.86	3.60	3.70	2.4
WATER	0.28	0.05	0.09	0.3
WET	0.12	0.00	0.00	0

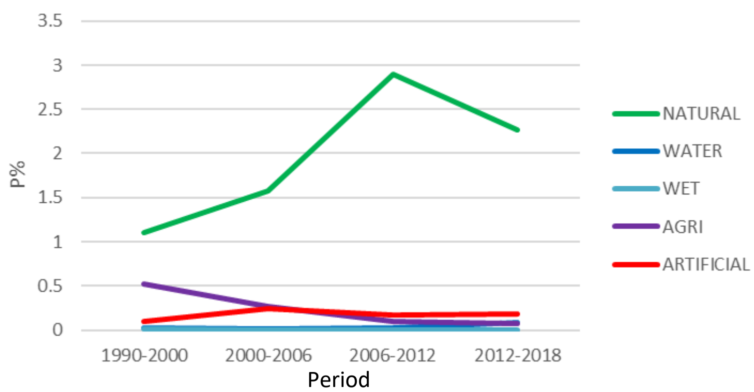


Figure 9 (left) Total area with land cover change for the main category of land cover. (right) The trend for the odds of a change into a land cover.

Then by reviewing the threats with the resolution of 30x30, 19% of the presence points fell into a land cover category that had some major category change within the last 20 years, this may mean that the species in some areas are living in zones that changed the land cover. And from the review of the home range, we can infer that these areas are rice fields or salt ponds (Figure 10). All the distance to roads had high correlation values between the pairs. None was removed as this is a characteristic of the structure of the road network. Additionally, the VIF also was reviewed for the threats, and the highest value was 5.7 and 5.9 for distance from roads of 5m width and distance from roads of 10m width since the value was lower than then they were conserved. The categorical threats continuous urban and discontinuous urban fabric had some presences, this was unexpected, but still, they were selected to fit the model (See ANNEX 4 for all correlation plots).

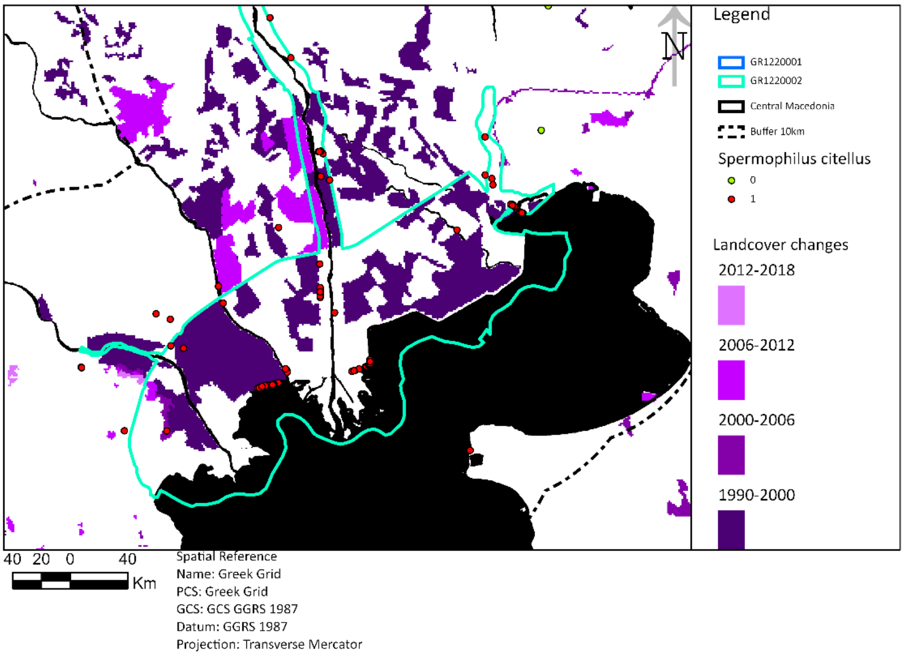


Figure 10 Protected Area GR1220001 land cover changes; the red dots shows the *Spermophilus citellus* presences records.

3.3. Modelling the distribution of the squirrel, with and without threats

The results of the modelling of the ground squirrel are presented in two sections. Firstly, the model is fitted with only the environmental explanatory variables, and secondly, the threats are added to assess their usefulness in predictions.

3.3.1. Species distribution of the squirrel based on environmental variables

Given the data reviewed on the previous section, the first fitted model included: Altitude, Slope, combined land cover category, Continuous urban fabric, Discontinuous urban fabric, Non-irrigated arable land, Permanently irrigated land, Rice fields, Inland marshes, Salt marshes, Fluvial clays, silts and loams, No water /No wet area.

The selected variables were tested for significance given a normal distribution, so the presences and absences were tested for Pr|z| to be at least of 0.5; Furthermore, at this level of confidence, there is no evidence to reject altitude, slope nor fluvial clays, silts and loams from the model. A further stepwise regression maintained the same variables, but the confidence level changed. All the other categorical variables had strong evidence to reject from the model (Table 4).

Table 4 Pr |z| values for the environmental variables, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). Grey cells are variables that the stepwise rejected.

	30		step (30)	
	Pr(>|z|)		Pr(>|z|)	
(Intercept)	0.88537		0.19397	
Altitude	0.00211	**	0.00131	**
slope	0.00338	**	0.00324	**
Combined land cover category	0.99429		9.88E-01	
Continuous urban fabric	NA			
Discontinuous urban fabric	0.9976		0.99758	
Non-irrigated arable land	0.99921			
Permanently irrigated land	0.99962			
Rice fields	NA			
Inland marshes	0.99518		9.95E-01	
Salt marshes	0.99224		9.92E-01	
Fluvial clays, silts and loams	0.0485	*	5.36E-02	.
No water nor No wet area	0.3805			
	AIC: 131.84		**AIC: 126.64**	

3.3.2. Accuracy assessment environmental variables

The accuracy assessment with the independent data showed kappa superior to 0.5, that means that the prediction is better than an aleatory fit. The model fitted only with environmental variables had an accuracy of AUC of 0.939, TSS0.817 and Kappa 0.793 Since these results are close to 1, they have good predictive power (See ANNEX 4 for ROC AUC plots). The Relative Variable Importance Correlation showed that altitude and slope explained 76% of the model and so the relationship between these variables had more contribution to the model than the environmental variable soil Fluvial clays, silts and loams.

3.3.3. The effect of threats on the squirrel distribution model

By adding the threats to the first model fitted, and performing the stepwise regression to identify the combination that is significant to the model, the environmental variables were reduced to the slope and no water. Moreover, the threats that were significant for the ground squirrel were Distance from roads of 15m width and Distance from roads of 20m width (Table 5). Even the tough density of all transports had high positive correlation value; it did not reach the significance level to be included in the final model.

Table 5 Pr |z| values for the environmental variables and threats, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). Grey cells are variables that the stepwise rejected.

	30 th Pr(>|z|)	step (30 th) Pr(>|z|)
(Intercept)	0.0448 *	0.0051 **
Altitude	0.3766	
Slope	0.0782 .	0.00501 **
Combined land cover category	0.9971	0.98849
Continuous urban fabric	NA	
Discontinuous urban fabric	0.9995	
Non-irrigated arable land	1	
Permanently irrigated land	0.9997	
Rice fields	NA	
Inland marshes	0.9985	0.99907
Salt marshes	0.9969	0.99786
Fluvial clays, silts and loams	0.7178	
No water nor No wet area	0.1784	0.03058 *
Area with changes of land cover on the last 20 years	0.455	0.18051
Density of are covered by all roads or communication routes	0.2679	0.00307 **
Total area cover density of small woody features	0.3794	
Distance from railways	0.0833 .	0.07563 .
Distance from roads of 5m width	0.315	
Distance from roads of 10m width	0.7759	
Distance from roads of 15m width	0.1624	0.01851 *
Distance from roads of 20m width	0.0866 .	0.01231 *
Distance from small woody features	0.2767	0.11593
Total population	0.7883	
	AIC: 131.84	**AIC: 126.64**

3.3.4. Accuracy assessment environmental variables and threats

Since the model fitted only with environmental variables had good predictive power, there was little room improvement. Including the threats increased the AUC but decreased the TSS and kappa. Still, the values for the accuracy assessment were close to 1 (See ANNEX 4 for ROC AUC plots). Two of the three accuracy assessment test showed a slight decrease in the accuracy of the model. Adding the threats, the increased the AUC0.951, but TSS and kappa decrease to 0.774 and 0.764. Furthermore, the Relative Variable Importance Correlation showed that distance of roads of 20m width was the threat that contributed the most to the model than for the variables it was the slope.

3.4. Relationship to Protected Areas

The previous section accuracy tests indicate a good fit for both models, thought the resulted modelled area differs greatly. Both models agree in the areas where the probability of not having the habitat-to burrow characteristics are but do not agree on the areas with a high probability of having the species (Figure 11). Moreover, the model with threats showed better accuracy according to the AUC; the final modelled area also made more notorious that the sample points presences location favours Distance from roads of 15m width and Distance from roads of 20m width, the linear model is producing not only restrictions but also creates areas with high probability inside the areas that were constrained.

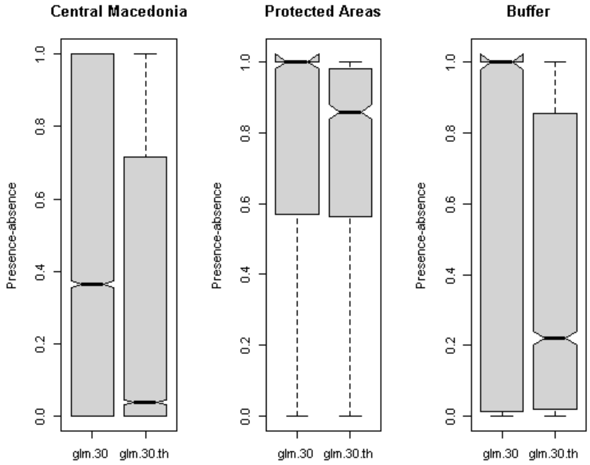


Figure 11 Frequency and probabilities based in model outputs (glm.30 model fitted with variables; glm.th model fitted including threats) for the area of Central Macedonia, Inside the Protected Areas and in a Buffer area of 10km around the Protected Areas.

On the Central Macedonia area, the spread of the predictions for the model fitted only with environmental variables favours the presences. By adding the threats, there is a marked reduction in the predicted areas with a high probability of having the ground squirrel. At the Protected Area zone, both models agree on a high probability of presence, again adding the threats reduce the area predicted with high probability inside the Protected Area. Finally, the area inside the Buffer of 10km around the Protected Areas area has the major difference between the models fitted with and without the threats, as the probable presence decreases the most (Figure 11).

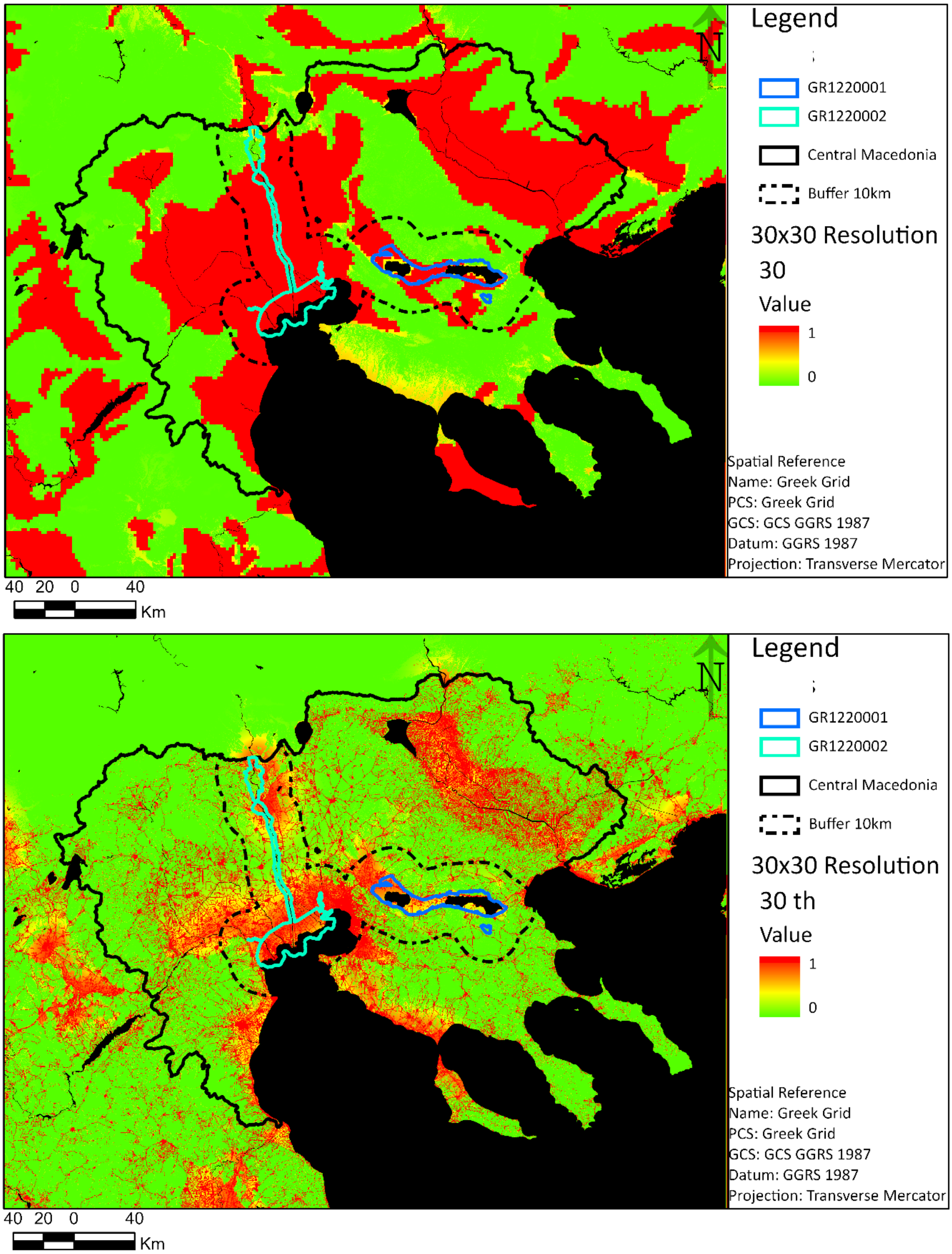


Figure 12 (top) Glm model fitted with only variables and (bottom) Glm fitted with environmental variables and threats. Max value in red, indicate high habitat probability, the minimum value in green indicate low habitat probability.

3.5. Assessment of the distribution of the squirrel at different resolutions

3.5.1. Importance of environmental variables and threats in relation to resolution

For the continuous variables, the high correlation values for the variables altitude and slope pointed them as relevant to the model at all resolutions. Both had at 30 resolution had high positive correlation value, at other resolutions, the value slightly decreased, but maintained the positive relationship. At the same time, aspect, easterness and northness had low correlation values, and the positive or negative relationship was different in each resolution. The threats were more consistent on the negative or positive relationship, thought two threats had a notorious decrease in correlation value was the density of all transport and distance to roads of 20m. Except for the density of all transportation, that had low negative correlated to altitude, the high altitude had less density of roads (ANNEX 4).

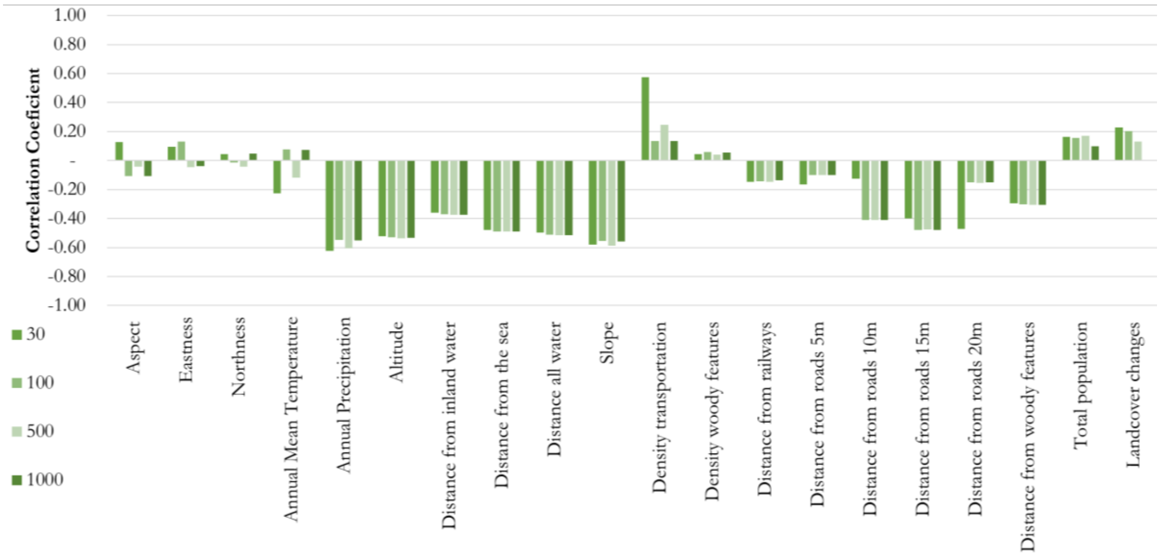


Figure 13 Comparation for all resolution correlation coefficient for the presences and absences with the continuous environmental variable sand the threats.

The categorical variables of the soil parent material and water and wetness were more consistent than the Corine land cover for the frequency if presence. At all resolution showed the category fluvial clays, silts and loams and the category no water nor no wet area to gather the presences. These two categories are relevant to the ecology to the species to have the burrow characteristics, as by combining them with the correct land cover, we can obtain an approximation to the probable niche of the species.

Unfortunately, the Corine land cover frequency if presence per category was different at each resolution. Even though the land cover shows a preference for the salt marshes, Non-irrigated arable land and Permanently irrigated land at all resolutions, this did not translate into high significance when fitting the models, neither once the model was fitted improved the relative variable importance Correlation. By changing the resolution, the presences at the categories Inland marshes and Rice fields that were originally at the 30 resolution were absorbed into other categories (Figure8) in the study area the presences fell into the border of the land cover categories. So, by changing the pixel resolution, the class was modified. This was also relevant for the combined land cover category and influence at the significance of the variables when fitting the model.

3.5.2. Fitting the models

When fitting the models with the environmental variables, consistently were significant for all the resolutions were altitude, slope, and soil parent material Fluvial clays, silts and loams. The combined categorical variable was significant at 100 and 1000. The Corine land cover Permanently irrigated land was significant only at the 1000 resolution. The no water nor no wet area category was not significant at the resolution 30 but was significant at all other resolutions (Figure15).

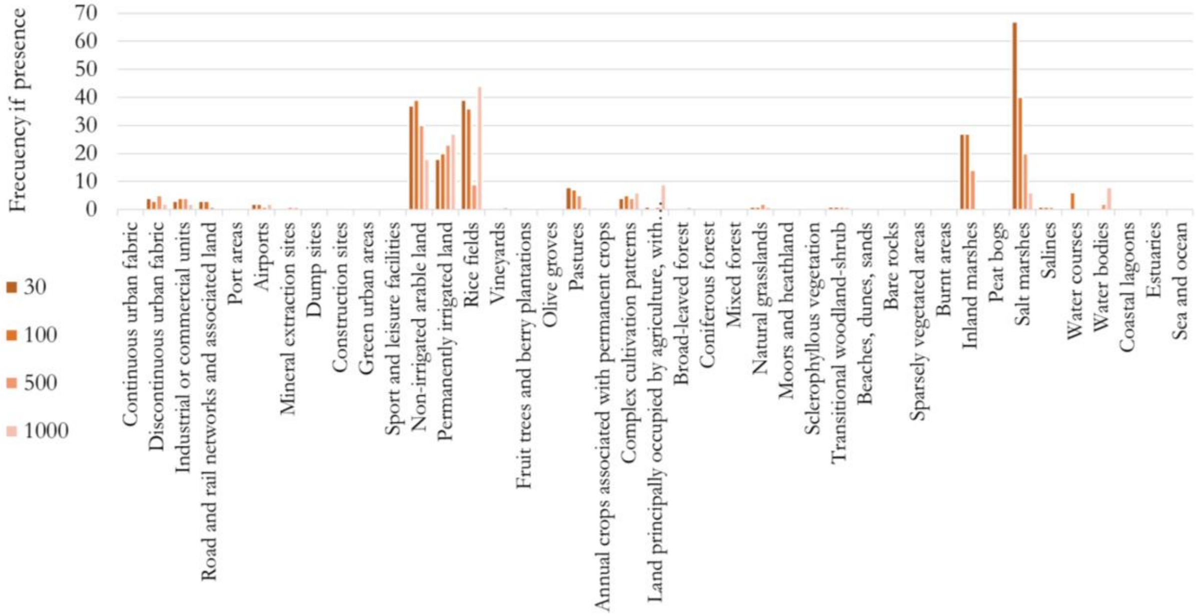


Figure 14 Corine Land cover frequency if present at all resolutions

By adding the threats into the models, the environmental variables that were significant in all resolutions were slope and no water nor no wet area. This time for soil parent material fluvial clays, silts and loams and altitude there was evidence to reject at 30 resolution. Moreover, the stepwise regression did not choose it to reduce the AIC it from the model. Nevertheless, these environmental variables were both significant at 100, 500 and 1000. These environmental variables also had high contribution into the Relative Variable Importance Correlation (Figure 15).

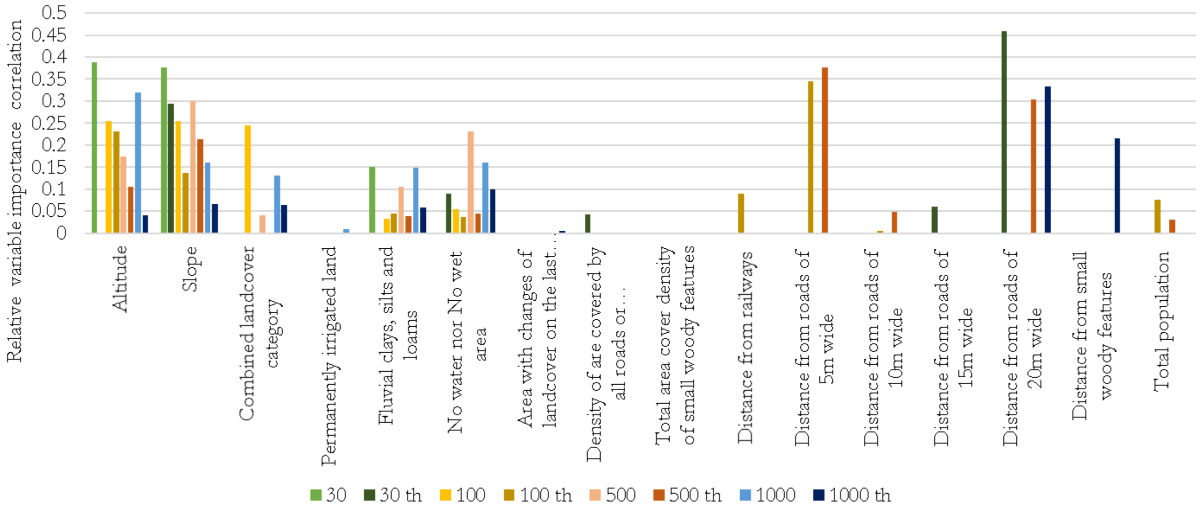


Figure 15 Relative Variable Importance of environmental variables and threats for the fitted models

No threat was significant at all resolutions. The resolution that had more significant threats was 500 next was 100, and finally, 1000 and 30 had the same number of significant threats. The threat that recurrently was significant was the distance to roads of 10m width. All other threats changed its significance between resolutions not being consistent with the results. As well, the relative variable importance correlation was different in all resolutions. This means that as the pixel resolution changes the model will also change and it will not be fitted with the same environmental variables nor threats.

3.5.3. Accuracy assessment

All model's prediction were better than an aleatory fit had. Since all models have good predictive power, to determine which resolution a more in detail review of the accuracy test is needed. The ROC curve displays the sensitivity (the proportion of correctly predicted presences) and the specificity (proportion of correctly predicted absences) generating the AUC. AUC represents a ratio between sensitivity and specificity. All models had an AUC bigger than 0.8 for the test independent data. Furthermore, the 30-resolution model fitted with only environmental variables had the best accuracy, as this resolution was the original.

The lowest accuracy for the models fitted only with environmental variables was at 500 resolution. Then this model also had more improvement by adding threats than the other models. 500 resolution also had the highest count of threats and variables that were significant to fit the model. Also, when performing the accuracy assessment among the models, the differences between the training data and test data AUC were small, but present, as the resolution 500 and 1000 had a lower difference between the accuracy of the training and independent data. The model that performed the best in accuracy after the 500 resolution was the 1000 resolution (Figure 16).

	AUC	TSS	Kappa
30	0.939	0.817	0.793
30 th	0.951	0.774	0.764
100	0.859	0.654	0.619
100 th	0.934	0.826	0.762
500	0.843	0.63	0.501
500 th	0.958	0.87	0.837
1000	0.886	0.674	0.674
1000 th	0.969	0.805	0.781

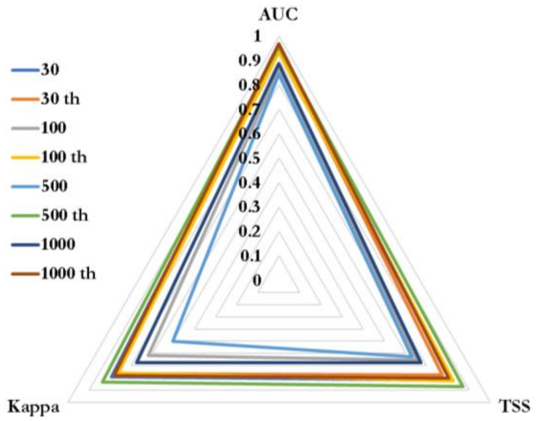


Figure 16 Accuracy assessment of all models. AUC, TSS and Kappa.

If we consider Kappa as the percentage of agreement between the sensitivity and specificity derived from a confusion matrix, then the model fitted at 500 resolution with threats is the one that outperforms all other models. But as Kappa is dependent on prevalence (total true positives plus the total of false positives between total predictions), but the TSS is not dependent on this issue; again, the 500 resolution outperforms. Finally, the ROC-AUC generates at all thresholds the accuracy of the model; in this case, the best performance is for the 1000th model.

3.6. Spatial patterns of threats in relation to the distribution of the species and the protected areas

No matter the resolution of the model all predicted more probable presence inside the Protected Area than within the 10 km buffer zone or the Region. It is noteworthy that the 500-resolution model with threats was the one that predicted less probable presence in all areas.

The results show that including the threats did reduce in the same areas the probable presence, thought the intensity of that change and improvement was different. Both resolution 30 and resolution 500 showed a marked decrease of probable presence in the north-central part of the Protected Area GR1220002. Figures 11, 18, 19 and 20 indicate how different models with high accuracy but different environmental variables and threats will generate a variety of results. When interpreting the spread is necessary to take into account how the model was produced. When adding threats to the model, they had more relative importance than when considering the environmental variables only. This means that the threats had more weight into the final modelled map (Figure 17).

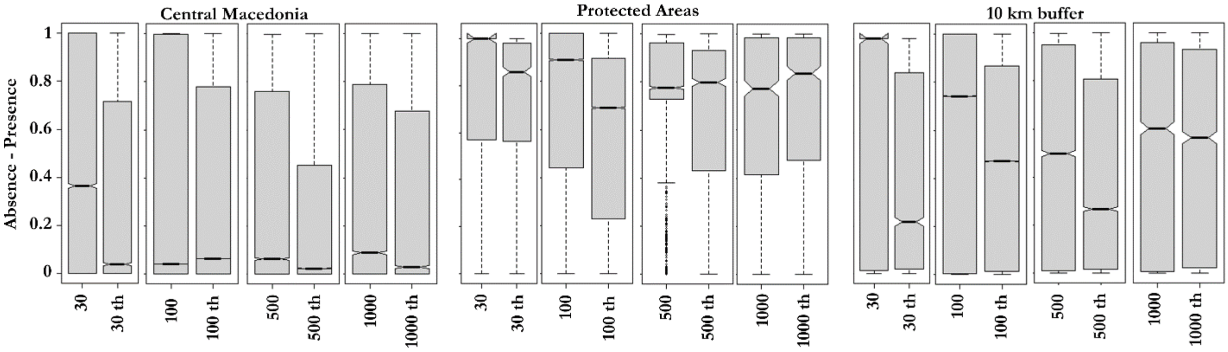


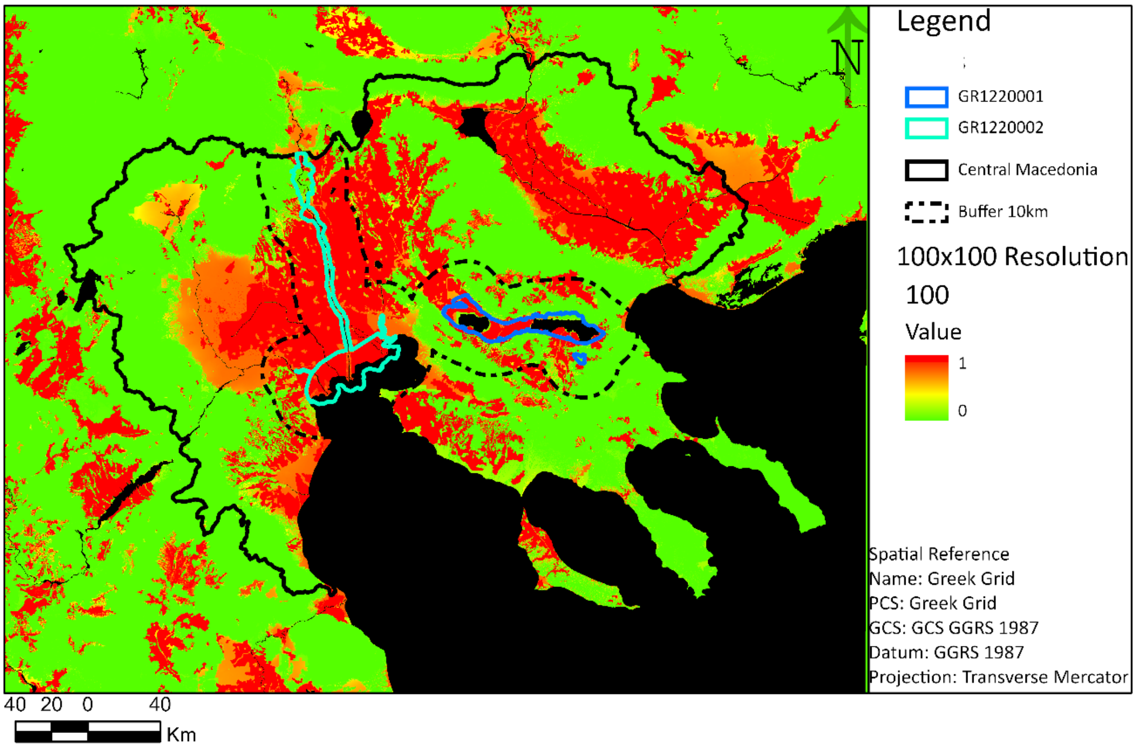
Figure 17 Boxplot modelled probable presence for the zones inside Central Macedonia Area, in order 30x30 resolution, 100x100 resolution, 500x500 resolution and 1000x1000 resolution.

Since the Central Greece Region is part of the geographic range of the ground squirrel in all its territory, other Natura 2000 Protected Areas within could also have the species. The Protected Areas that reported within its data form are GR1220001 and GR1220002. These Protected Areas were consistent with the prediction of having burrow characteristics. Furthermore, GR1250004 and GR1220012 have presence points of the species and high probable presence, but they do not report the ground squirrel within its objectives.

Several Protected Areas had not presence points, nor it reported the species, but according to the model, showed probable presence. The main issue is to differentiate which ones have that result because it is an error of the model or there are other factors such as local threats, other biotic interactions or a reporting mistake. The Protected Areas, GR1270008, GR1270010, GR1270009 and GR1270015 are coastal ones; they had probable presence because of the border effect of the pixels within the model.

Also, the Protected Areas GR1250004, GR1220012 and GR1260003 consistently at all resolutions had a high probable presence and do not have presence points neither report is within its conservation objectives this last one had all its area with a high probability of having the species. Nevertheless, it also has a small settlement and its adjacent to and discontinuous urban fabric area. Other Protected Areas such as GR1230001, GR1230002 had their probable presence, more reduced by the threats. With the model fitted only with environmental variables showed probable presence, but by including the threats, the probable presence decreased considerably. Finally, the Protected Area GR1230001 does not mention in their data form neither have presence points. However, their mean is similar to the protected areas in the region that does have the species, so it would be useful to do fieldwork to validate the local population of the species (See ANNEX 6).

Resolution 100 x 100



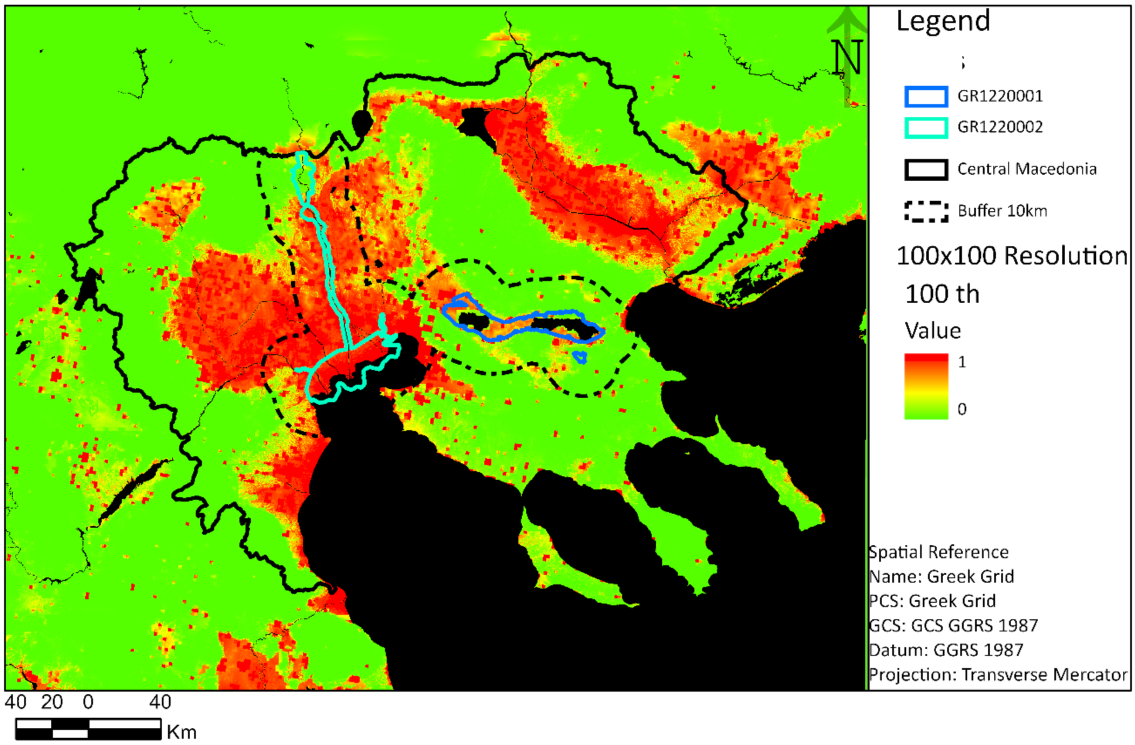
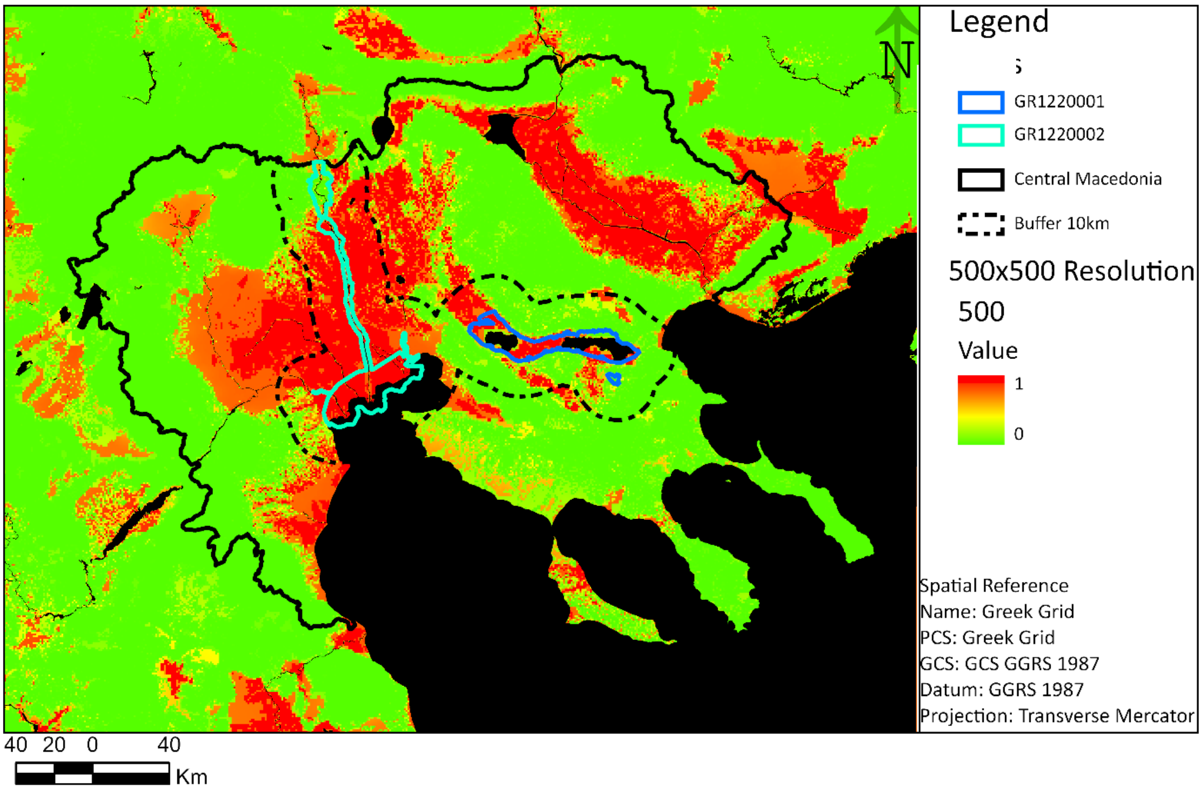


Figure 18 Glm model fitted with only variables(top) and (bottom) Glm fitted with environmental variables and threats. Max value in red, indicate high habitat probability, the minimum value in green indicate low habitat probability.

Resolution 500 x 500



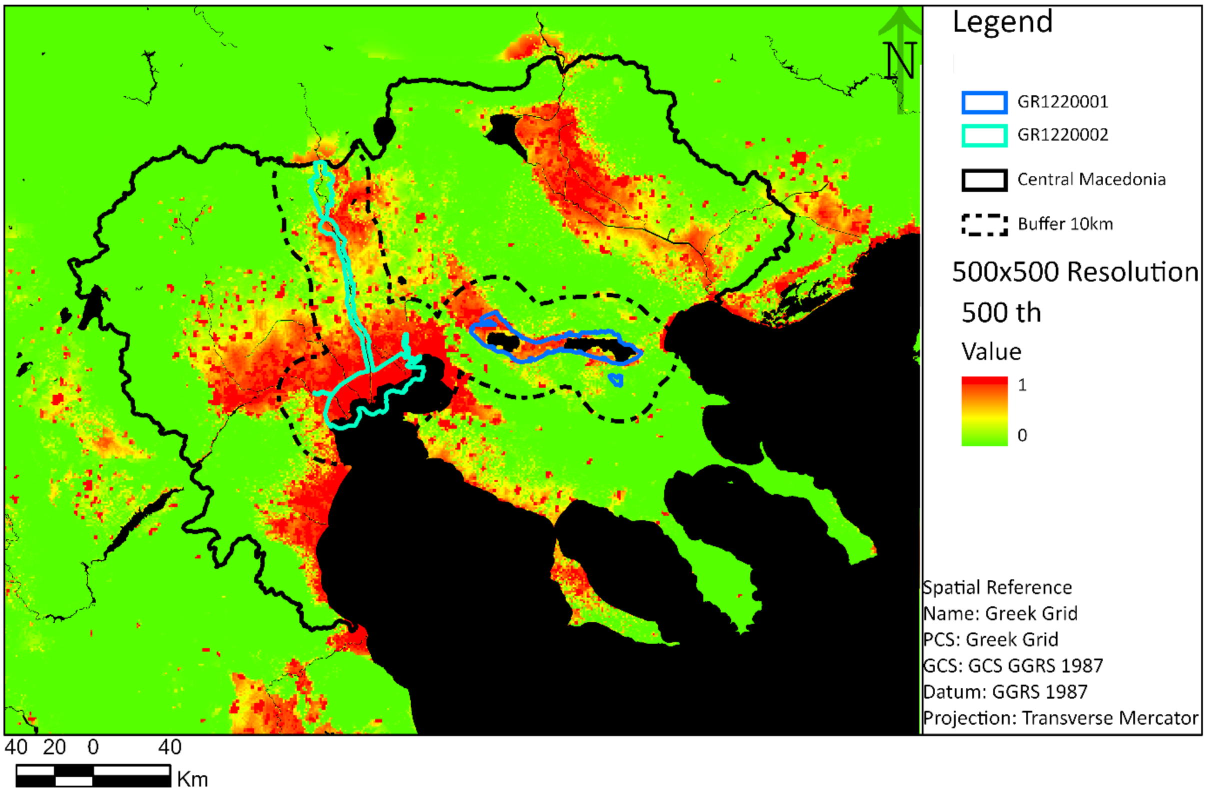
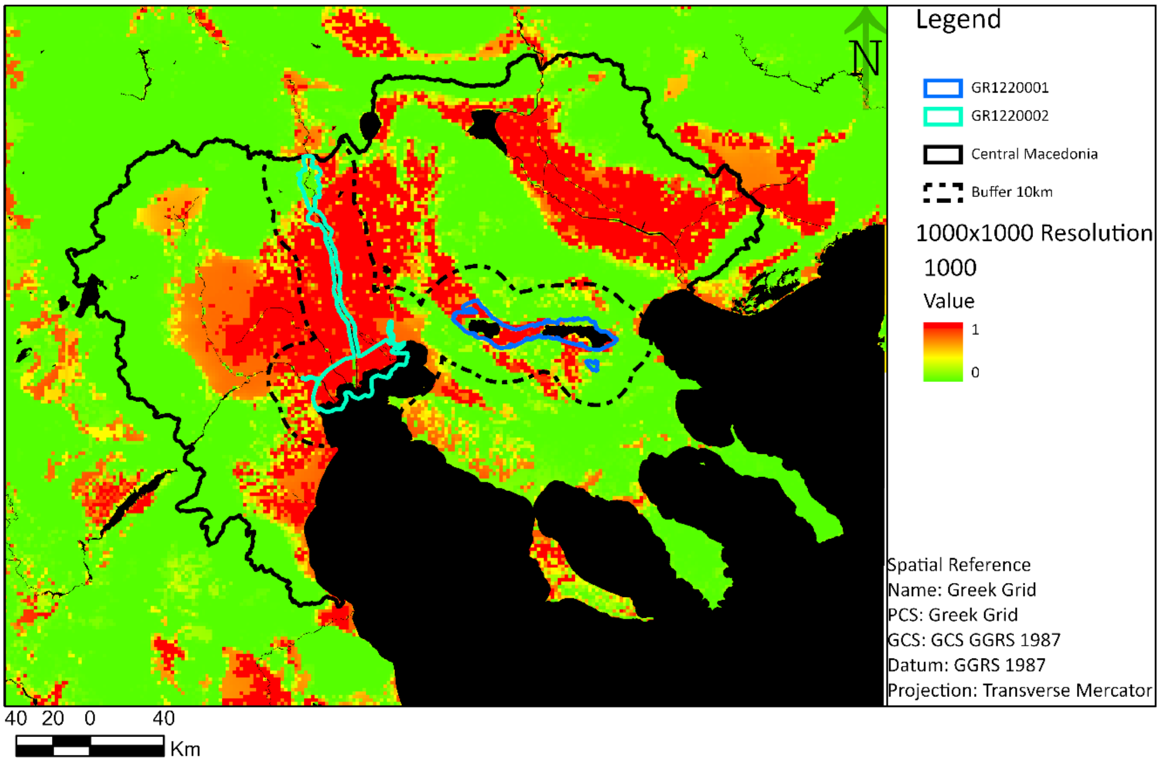


Figure 19 Glm model fitted with only variables(top) and (bottom) Glm fitted with environmental variables and threats. Max value in red, indicate high habitat probability, the minimum value in green indicate low habitat probability

Resolution 1000 x 1000



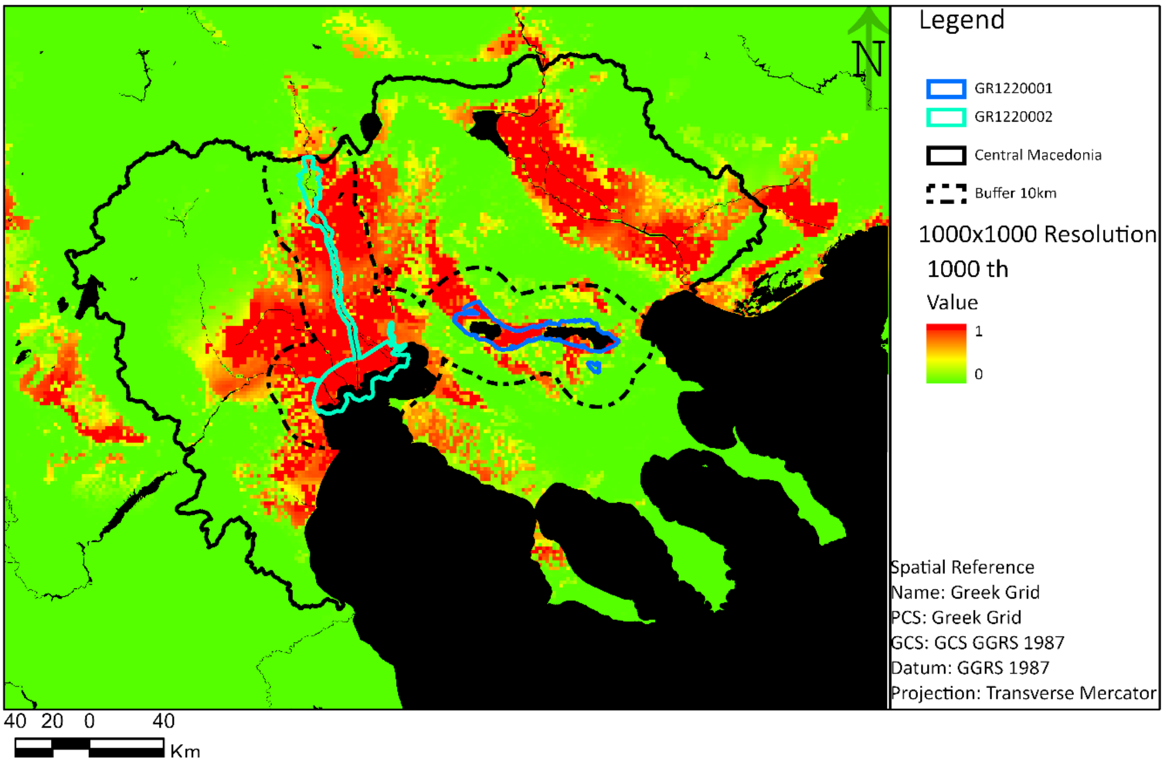


Figure 20 (Top) Glm fitted with only variables and (bottom) Glm fitted with environmental variables and threats. Max value in red, indicate high habitat probability, the minimum value in green indicate low habitat probability.

4. DISCUSSION

The results from the previous section provide arguments to support that the 500X 500 resolution can be used to explain probable presence given the environmental variables and threats in the Central Macedonia Region. Also, the results suggest that coarser the pixel the mean of the probability of having the characteristics within the Protected Area decreased for the models with threats. Then if we crossed this with the accuracy assessment information, the highest accuracy models with treats predicted lower presence overall in all Protected Areas in contrast with the models fitted with only environmental variables that predicted a high probability of presence but with slyly less accurate results.

Umetsu, Paul Metzger, & Pardini, (2008) reported: "small mammals perceive landscape features at large scale". This Autor argument that small mammals react strongly to landscape quality and structure given its home range, and so a fragmented landscape may require higher pixel size and extend to detect those. However, the present work is not modelling connectivity either fragmentation and so what Umetsu, Paul Metzger, & Pardini, (2008) the concept of the response of the species to a coarser pixel size remains.

Data availability was a recurrent problem; A weaknesses of the model is the spatial information with which it was created. The diverse original pixel resolution, temporality and coverage are of the environmental variables also could be introducing noise into the model, in particular the soil parent material raster, this being the oldest and coarser data. A national or regional soil database could reveal other relationships for this species.

Even though the obtained results have good accuracy, the model still has room for improvement. Some possible improvements include by presenting it to the protected area managers to verify the output and also to provide feedback on the areas that the model predicted to have the characteristics required for the European Ground Squirrel, and so to produce absence points with expert knowledge.

Also, the analysis can be improved by adding a high spatial resolution map of the vegetation cover to fit the model instead of the Corine land cover. The land cover classification groups areas with at least 75% of the same pattern, this creates the idea the species lives inside the crops or the salt marshes, and this is not the case. Since the area is cover by agriculture land, the grasses that the ground squirrel require do exist but as vegetation bordering the crops and mixed on farmland, not as a Corine land cover category. So, the grassland and border crops relationship analysis may improve the predictions. An even further step could include pastoral and farming actions and so approach a more realistic realized niche distribution for the species in the Region.

On the mapping of threats and linking them to geo-representation some authors have developed methodologies to map threats to protected areas (Mcpherson et al., 2008) but as these methodologies are highly dependent on expert opinion techniques, the Author tried a more simple approach using, density, distance and Boolean.

The main threats that had significance in all the models were the roads. It has to be mentioned the possible effect of roadside bias within the database used to fit the models. Having a road to access the area where a species is can increase the observation records because the chance of being seen is boosted (Ronen Kadmon, 2004). Most of the database presences were near a road or on the road of 5m and 10m. On this topic, Fahrig & Rytwinski, (2009) mentions that for small mammals, roads have a small positive relation to none, but the fitted models show a different result for the ground squirrel, this species response more like medium-sized mammals with a negative effect. Furthermore, the logistic model converts all relationships into linear responses, in the place where the road is the model gave probable prediction instead of generating an absence.

5. CONCLUSIONS AND RECOMMENDATIONS

The analysis suggests that at different pixel resolutions in the same extent, it possible to model the presence of the European Ground Squirrel or European Souslik (*Spermophilus citellus*) given its burrow environmental characteristics. The environmental variables were consistently explaining the presence and absence of the species. Although including threats did increase the accuracy of the models at all resolutions, no threat was consistent on being significant at all resolutions.

The most important contribution of this work is to answer the question What is the appropriate resolution to map a species for Protected Area management? For the ground squirrel in Central Macedonia, the resolution that has more significant variables and threats, and reflects the most considerable improvement on accuracy by considering the threats, and as well as 2 of 3 of the best accuracy assessments is the 500 resolution. This does however not mean that all the European ground squirrel range of probable habitat should be mapped at this resolution nor that other related species should be mapped this way. It indicates that under the reviewed factors in the study area, these are the relationships.

On the transferability o the model. Since the developed models are straightforward, they could be used as a first step into developing of ecological corridors, evaluation of reintroduction of species, population density calculation, connectivity and fragmentation analysis, to mention some. Moreover, the Author does not omit to mention that a more comprehensive study of the European Ground Squirrel in the same study area may produce different results as the absence data was not reviewed with local experts nor it proceeds from an official source. Furthermore, that the roads of 5m or 20m width has influence over the possible presence of the species does not mean that the other road sizes do not have an environmental impact as the models were not developed to measure this effect and so, it should not be used in that context.

The models were developed by implying that the effect will be the same in the study area, although we know that this generalization is not a reflection of the real world, is a helpful approximation. Finally, mapping threats is essential to conservation efforts, and further research has to be done to translate the concept into a map. And this way to implement efficient science-based natural resources management actions.

LIST OF REFERENCES

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*(6), 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Bonebrake, T. C., Guo, F., Dingle, C., Baker, D. M., Kitching, R. L., & Ashton, L. A. (2019). Integrating Proximal and Horizon Threats to Biodiversity for Conservation. *Trends in Ecology & Evolution*, *34*(9), 781–788. https://doi.org/10.1016/j.tree.2019.04.001

Connor, T., Viña, A., Winkler, J. A., Hull, V., Tang, Y., Shortridge, A., … Liu, J. (2019). Interactive spatial scale effects on species distribution modeling: The case of the giant panda. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-50953-z

Copernicus. (2018). CORINE Land Cover — Copernicus Land Monitoring Service. Retrieved April 12, 2019, from https://land.copernicus.eu/pan-european/corine-land-cover

Dimopoulos, P., Bergmeier, E., & Fischer, P. (2006). Natura 2000 habitat types of greece evaluated in the light of distribution, threat and responsibility. *Biology and Environment*, *106*(3), 175–187. https://doi.org/10.3318/BIOE.2006.106.3.175

EEC. 92/43/EEC Habitats Directive, 29 § (1992). European Commission. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:01992L0043-20130701&from=EN

Ervin, J., Sekhran, N., Gidda, A. D. S., Vergeichik, M., & Mee, J. (2010). *Protected Areas for the 21 st Century : Lessons from UNDP/GEF's Portfolio*. New York: United Nations Development Programme and Montrea: Convention on Biological Diversity: new York: United nations Development Programme and montreal: Convention on Biological Diversity.

EU WISE. (2019). Implementation of River basin management plans - Greece - Environment - European Commission. Retrieved April 11, 2020, from https://ec.europa.eu/environment/water/participation/map_mc/countries/greece_en.htm

Eurostat. (2011). GEOSTAT - Eurostat. Retrieved June 21, 2020, from https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat

Fahrig, L., & Rytwinski, T. (2009). Effects of roads on animal abundance: An empirical review and synthesis. Ecology and Society, 14(1), 21. *Ecology and Society*, *14*(1). Retrieved from http://http//ecologyand society.org/vol14/jss1/art21

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assesment of prediction errors in conservation pressence/absence models. *Environmental Conservation*, *24*(1), 38–49. https://doi.org/10.1017/S0376892900021214

Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., & Ohlson, D. (2012). *Structured Decision Making. Structured Decision Making*. Oxford, UK: Wiley-Blackwell. https://doi.org/10.1002/9781444398557

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., … Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424–1435. https://doi.org/10.1111/ele.12189

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Haby, N. A., Delean, S., & Brook, B. W. (2018). Improving performance and transferability of small mammal species distribution models. *Transactions of the Royal Society of South Australia*, *142*(2), 143–161. https://doi.org/10.1080/03721426.2018.1513770

Hansen, A. J., & DeFries, R. (2007). Ecological mechanisms linking protected areas to surrounding lands. *Ecological Applications*, *17*(4), 974–988. https://doi.org/10.1890/05-1098

Hastie, T., & Tay, K. (2020). Glm family functions in glmnet 4 . 0, 1–7.

Host, G. E., Kovalenko, K. E., Brown, T. N., Ciborowski, J. J. H., & Johnson, L. B. (2019). Risk-based classification and interactive map of watersheds contributing anthropogenic stress to Laurentian Great Lakes coastal ecosystems. *Journal of Great Lakes Research*, *45*(3), 609–618. https://doi.org/10.1016/j.jglr.2019.03.008

INWEB. (2004). Vardar/Axios River Sub-basin. Retrieved April 10, 2020, from http://www.inweb.gr/workshops2/sub_basins/10_Axios.html

IUCN-CMP. (2018). Unified Classifcation of Direct Threats Version 3.2, 1–20. Retrieved from http://s3.amazonaws.com/iucnredlist-

newcms/staging/public/attachments/3127/dec_2012_guidance_threats_classification_scheme.pdf
IUCN. (2010). Spermophilus citellus (European Souslik). Retrieved June 16, 2020, from https://www.iucnredlist.org/species/20472/9203833
IUCN. (2019). Greece | IUCN. Retrieved June 11, 2019, from https://www.iucn.org/regions/europe/resources/country-focus/greece
IUCN, Stolton, S., Shadie, P., & Dudley, N. (2008). *Guidelines for Applying Protected Area Management Categories*. Gland, Switzerland. Retrieved from www.iucn.org/pa_guidelines
Janák, M., Marhoul, P., & Mateju, J. (2013). *Action Plan for the Conservation of the European Ground Squirrel Spermophilus citellus in the European Union List of contributors*. European Commission.
Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, *21*(4), 498–507. https://doi.org/10.1111/j.1466-8238.2011.00683.x
Johnson, D. H. (1980). The Comparison of Usage and Availability Measurements for Evaluating Resource Preference. *Ecology*, *61*(1), 65–71. https://doi.org/10.2307/1937156
Jones-Walters, L., Gillings, S., Groen, T., Hennekens, S., Noble, D., Huskens, K., … van der Sluis, T. (2016). The “Umbrella Effect” of the Natura 2000 network An assessment of species inside and outside the European Natura 2000 protected area network Executive Summary. https://doi.org/10.18174/385796
Joppa, L. N., O’Connor, B., Visconti, P., Smith, C., Geldmann, J., Hoffmann, M., … Minin, E. Di. (2016). Filling in biodiversity threat gaps. *Science*, *352*(6292), 416–418. https://doi.org/10.1126/science.aaf3565
Lecours, V. (2017). On the use of maps and models in conservation and resource management (Warning: Results may vary). *Frontiers in Marine Science*, *4*(SEP), 1–18. https://doi.org/10.3389/fmars.2017.00288
Leverington, F., Costa, K. L., Courrau, J., Pavese, H., Nolte, C., Marr, M., … Hockings, M. (2010). *Methods Review of Management effectiveness evaluation in protected areas – a global study. Second edition*. The University of Queensland Brisbane. AUSTRALIA.
Leverington, F., Hockings, M., & Lemos Costa, K. (2008). Management effectiveness evaluation in protected areas: Report for the project “Global study into management effectiveness evaluation of protected areas,” 74.
Levin, S. (1992). The Problem of Pattern and Scale in Ecology : The Robert H . MacArthur Award Lecture Author (s): Simon A . Levin Published by : Ecological Society of America. *Source: Ecology*, *73*(6), 1943–1967. Retrieved from http://www.columbia.edu/cu/e3bgrads/JC/Levin_1992_Ecology.pdf
Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, *33*(1), 103–114. https://doi.org/10.1111/j.1600-0587.2009.06039.x
Manly, B. F. J. (2009). *Statistics for Environmental Science and Management*. (R. Smith, Ed.) (Second). Boca Raton FL: Chapman & Hall/CRC Taylor & Francis Group.
Maragou, P., & Mantziou, D. (2000). Assessment of the Greek Ramsar wetlands. *Assessment*.
Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature*, *405*(6783), 243–253. https://doi.org/10.1038/35012251
Mcpherson, M., Schill, S., Raber, G., John, K., Zenny, N., Thurlow, K., & Sutton, A. H. (2008). GIS-based Modeling of Environmental Risk Surfaces (ERS) for Conservation Planning in Jamaica. *Journal of Conservation Planning*, *4*, 60–89.
Milne, B. T. (1995). Aplications of Fractal Geometry in Wildlife Biology. In John A. Bissonette (Ed.), *Wildlife and Landscape Ecology Effects of pattern and Scale* (pp. 32–69). New York: Springer-Verlag.
N2K GR1220002 dataforms. (2019). Retrieved May 15, 2020, from http://natura2000.eea.europa.eu/Natura2000/SDF.aspx?site=GR1220002
Naimi, B., & Araújo, M. B. (2016). Sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, *39*(4), 368–375. https://doi.org/10.1111/ecog.01881
Ramm, F., Names, I., Files, S. S., Catalogue, F., Features, P., Features, N., … Fields, A. (2019). OpenStreetMap Data in Layered GIS Format, 1–20.
Ramos-Lara, N., Koprowski, J. L., Kryštufek, B., & Hoffmann, I. E. (2014). Spermophilus citellus (Rodentia: Sciuridae). *Mammalian Species*, *913*(913), 71–87. https://doi.org/10.1644/913.1
Ronen Kadmon, O. F. and A. D. (2004). Effect of Roadside Bias on the Accuracy of Predictive Maps Produced by Bioclimatic Models Author (s): Ronen Kadmon , Oren Farber and Avinoam Danin

Published by : Wiley on behalf of the Ecological Society of America Stable URL : https://www.jstor.org/sta. *Ecological Applications*, *14*(2), 401–413.
Salafsky, N., Salzer, D., Stattersfield, A. J., Hilton-Taylor, C., Neugarten, R., Butchart, S. H. M., … Wilkie, D. (2008). A standard lexicon for biodiversity conservation: Unified classifications of threats and actions. *Conservation Biology*, *22*(4), 897–911. https://doi.org/10.1111/j.1523-1739.2008.00937.x
Sales-Luís, T., Bissonette, J. A., & Santos-Reis, M. (2012). Conservation of Mediterranean otters: The influence of map scale resolution. *Biodiversity and Conservation*, *21*(8), 2061–2073. https://doi.org/10.1007/s10531-012-0297-z
Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, *413*(6856), 591–596. https://doi.org/10.1038/35098000
Schulze, K., Knights, K., Coad, L., Geldmann, J., Leverington, F., Eassom, A., … Burgess, N. D. (2018). An assessment of threats to terrestrial protected areas. *Conservation Letters*, *11*(3), 1–10. https://doi.org/10.1111/conl.12435
Secretariat of the Convention on Biological Diversity. (2004). *Programme of Work on Protected Areas (CBD Programmes of Work)*. Retrieved from https://www.cbd.int/doc/publications/pa-text-en.pdf
Song, W., Kim, E., Lee, D., Lee, M., & Jeon, S. W. (2013). The sensitivity of species distribution modeling to scale differences. *Ecological Modelling*, *248*, 113–118. https://doi.org/10.1016/j.ecolmodel.2012.09.012
Stankey, G. H., Clark, R. N., & Bormann, B. T. (2005). Adaptive management of natural resources: Theory, concepts, and management institutions. *USDA Forest Service - General Technical Report PNW*, (654), 1–73.
Stephanie, M., Ceri, W. H., & S.J., O. (2001). Evaluating presence – absence models in ecology : the need to account for prevalence. *Journal of Applied Ecology*, *38*, 921–931. https://doi.org/10.1080/09613210110101185
Synergassia regional Parthership, & Enterprise Greece. (2018). *Enterprise Greece, invest and trade. Region of Central Macedonia investment profile. Initiatives.*
Tulloch, V. J. D., Tulloch, A. I. T., Visconti, P., Halpern, B. S., Watson, J. E. M., Evans, M. C., … Possingham, H. P. (2015). Why do We map threats? Linking threat mapping with actions to make better conservation decisions. *Frontiers in Ecology and the Environment*, *13*(2), 91–99. https://doi.org/10.1890/140022
Umetsu, F., Paul Metzger, J., & Pardini, R. (2008). Importance of estimating matrix quality for modeling species distribution in complex tropical landscapes: a test with Atlantic forest small mammals. *Ecography*, *0*(0), 080304020349105–0. https://doi.org/10.1111/j.2008.0906-7590.05302.x
UNEP-WCMC and IUCN. (2019). Protected Planet: The Global Database on Protected Areas Management Effectiveness (GD-PAME). Retrieved July 24, 2019, from https://protectedplanet.net/c/protected-areas-management-effectiveness-pame
UNEP-WMCM, IUCN, & and NGS. (2018). *Protected Planet Report 2018*. Cambridge UK; Gland, Switzerland; and Washington, D.C., USA.
Velasco, J. A., & González-Salazar, C. (2019). Akaike information criterion should not be a "test" of geographical prediction accuracy in ecological niche modelling. *Ecological Informatics*, *51*(October 2018), 25–32. https://doi.org/10.1016/j.ecoinf.2019.02.005
Vokou, D., Giannakou, U., Kontaxi, C., & Vareltzidou, S. (2016). *The wetland book. II, Distribution, description, and conservation Axios, Aliakmon and Gallikos Delta Complex, Northern Greece,.*
Wilson, T., Sleeter, B., Sherba, J., & Cameron, D. (2015). Land-use impacts on water resources and protected areas: applications of state-and-transition simulation modeling of future scenarios. *AIMS Environmental Science*, *2*(2), 282–301. https://doi.org/10.3934/environsci.2015.2.282
Youlatos, D., Boutsis, Y., Pantis, J. D., & Hadjicharalambous, H. (2007). Activity patterns of European ground squirrels (Spermophilus citellus) in a cultivated field in northern Greece. *Mammalia*, *71*(4), 183–186. https://doi.org/10.1515/MAMM.2007.030

ANNEX

ANNEX 1: Roads Groups

class	Width (m)	Group into
cycleway	2	No included
steps	2	
track_grade3	2	
track_grade4	2	
track_grade5	2	
path	3	
bridleway	5	Distance from roads of 5m width
footway	5	
living_street	5	
pedestrian	5	
residential	5	
service	5	
tertiary	5	
tertiary_link	5	
track	5	
track_grade1	5	
track_grade2	5	
unclassified	5	
unknown	5	
motorway_link	10	Distance from roads of 10m width
primary	10	
primary_link	10	
secondary	10	
secondary_link	10	
trunk	15	Distance from roads of 15m width
trunk_link	15	
motorway	20	Distance from roads of 20m width
Railway	20	Distance from railways way

ANNEX 2: Categorical and continuous variables and Categorical and continuous threats

Corine Land cover

Main group	Abbreviation	Description
Artificial surfaces	**vcl 111**	Continuous urban fabric
	vcl 112	Discontinuous urban fabric
	vcl 121	Industrial or commercial units
	vcl 122	Road and rail networks and associated land
	vcl 123	Port areas
	vcl 124	Airports
	vcl 131	Mineral extraction sites
	vcl 132	Dump sites
	vcl 133	Construction sites
	vcl 141	Green urban areas
	vcl 142	Sport and leisure facilities
Forest and semi natural areas	**vcl 311**	Broad-leaved forest
	vcl 312	Coniferous forest
	vcl 313	Mixed forest
	vcl 321	Natural grasslands
	vcl 322	Moors and heathland
	vcl 323	Sclerophyllous vegetation
	vcl 324	Transitional woodland-shrub
	vcl 331	Beaches, dunes, sands
	vcl 332	Bare rocks
	vcl 333	Sparsely vegetated areas
Agricultural areas	**vcl 211**	Non-irrigated arable land
	vcl 212	Permanently irrigated land
	vcl 213	Rice fields
	vcl 221	Vineyards
	vcl 222	Fruit trees and berry plantations
	vcl 223	Olive groves
	vcl 231	Pastures
	vcl 241	Annual crops associated with permanent crops
	vcl 242	Complex cultivation patterns
	vcl 243	Land principally occupied by agriculture, with significant areas of natural vegetation
Wetlands	**vcl 334**	Burnt areas
	vcl 411	Inland marshes
	vcl 412	Peat bogs
	vcl 421	Salt marshes
	vcl 422	Salines
Water bodies	**vcl 511**	Watercourses
	vcl 512	Water bodies
	vcl 521	Coastal lagoons

Main group	Abbreviation	Description
	vcl 522	Estuaries
	vcl 523	Sea and ocean

PARMADO, Soil parent material

Abbreviation	Description
vpar 0	No information
vpar 1000	Consolidated-clastic-sedimentary rocks
vpar 1210	Sandstone
vpar 1211	calcareous sandstone
vpar 1300	Peelite, lutite or argillite
vpar 1410	Flysch
vpar 2100	calcareous rocks
vpar 2110	Limestone
vpar 2111	Hard limestone
vpar 3100	Acid to intermediate plutonic rocks
vpar 3110	Granite
vpar 3200	Basic plutonic rocks
vpar 3400	Acid to intermediate volcanic rocks
vpar 4200	Acid regional metamorphic rocks
vpar 4240	Gneiss
vpar 5000	Unconsolidated deposit is (alluvium, weathering residuum and slope deposits)
vpar 5400	Fluvial clays, silts and loams
vpar 5411	Terrace clay and silt
vpar 5500	Lake deposits
vpar 5610	Residual loam
vpar 5612	clayey loam
vpar 7100	Loess

WAW. Water and wetness

Abbreviation	Description
vW 0	No water nor No wet area
vW 1	Permanent water
vW 2	Temporary water
vW 3	Permanently wet areas (wetness)
vW 4	Temporary wet area (wetness)
vW 255	Unclassifiable (No image, clouds, shadows, snow or ice)

Continuous variables

Abbreviation	Description
vasp	Aspect
vasp_east	From aspect eastness

vasp_north	From aspect northness
vbio_1	Annual Mean Temperature
vbio_12	Total Annual Precipitation
vdem	Elevation
vdem_m	Elevation sea values masked
vdis_inwa	Distance from inland water bodies
vdist_sea	Distance from the coastal line
vdist_water_m	Distance from inland water bodies and the coastal line
vslope	Slope

Categorical threats

Abbreviation	Description
t_changes	Area that has had a land cover use change from 1990 till 2018, Boolean.
vcl 111	Continuous urban fabric
vcl 112	Discontinuous urban fabric

Continuous threats

Abbreviation	Description
tden_alltransp	Density of area covered by all roads or communication routes in percentage (%)
tden_woody	Total area cover density of small woody features (m^2)
tdis_rail	Distance from railways in (m)
tdis_roads_5m	Distance from road of 5m width in (m)
tdis_roads_10m	Distance from road of 10m width in (m)
tdis_roads_15m	Distance from road of 15m width in (m)
tdis_roads_20m	Distance from road of 20m width in (m)
tdis_woody	Distance from small woody features in (m)
tpop_density	Total population
tpop_density_m	Total population without n/a values

ANNEX 3: Threats, pressures and activities with impacts on the site reported for the Protected Areas for the *Spermophilus citellus*

Rank: H = high, M = medium, L = low

Occurrence: i = inside, o = outside, b = both

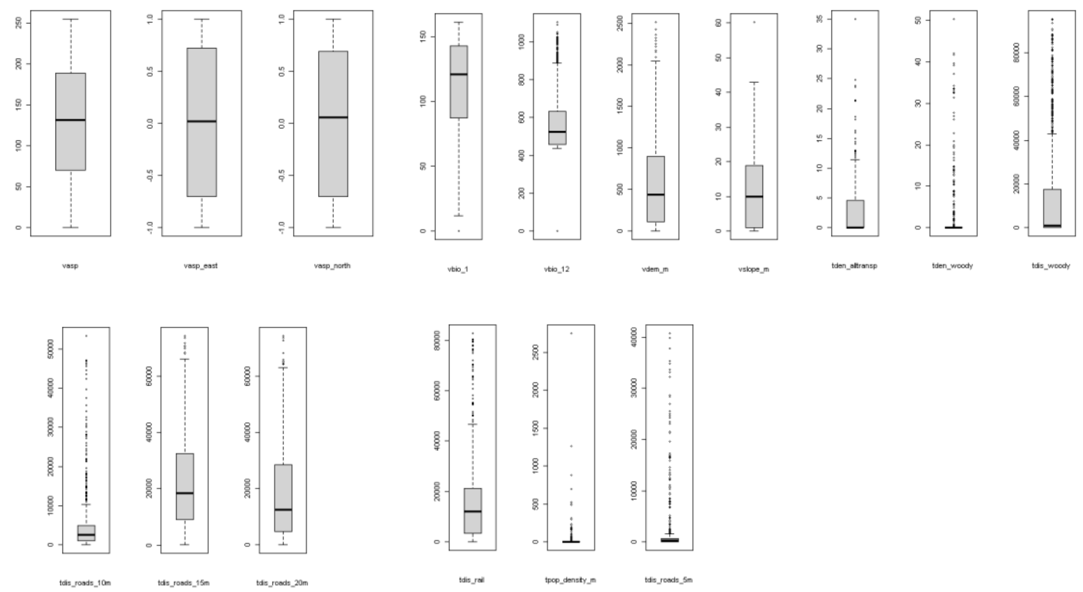
		GR1220001									GR1220002								
		H			L			M			H			L			M		
CODE	**DESCRIPTION**	**b**	**i**	**o**	**b**	**i**	**o**	**b**	**i**	**o**	**b**	**i**	**o**	**b**	**i**	**o**	**b**	**i**	**o**
A04	Grazing		1			1						1			1				
A04.01	Intensive grazing	1			1						1			1					
A04.01.01	Intensive cattle grazing				1									1					
A04.01.02	Intensive sheep grazing				1	1								1	1				
A04.02	Non-intensive grazing							1									1		
A04.03	Abandonment of pastoral systems, lack of grazing	1									1								
A08	Fertilisation			1									1						
B02.01	Forest replanting					1									1				
B02.02	Forestry clearance					1									1				
C01.01	Sand and gravel extraction						1									1			
C01.01.01	Sand and gravel quarries					1									1				
C03.03	Wind energy production																		
D01.02	Roads, motorways	1			1	1					1			1	1				
D01.05	Bridge, viaduct					1									1				
D01.06	Tunnel					1									1				
D02.01	Electricity and phone lines					1									1				
D03.02	Shipping lanes								1									1	
D03.02.02	Passenger ferry lanes (high speed)				1									1					
D05	Improved access to the site	1									1								
E01.02	Discontinuous urbanisation			1									1						
F02.01	Professional passive fishing	1						1			1						1		
F02.01.01	Potting					1									1				

F02.01.03	Demersal longlining							1									1		
F02.01.04	Pelagic longlining							1									1		
F02.02	Professional active fishing	1									1								
F02.03	Leisure fishing				1	1	1	1						1	1	1	1		
F03.01	Hunting	1							1		1							1	
F03.02.03	Trapping, poisoning, poaching							1									1		
F04.02	Collection (fungi, lichen, berries etc.)							1									1		
F05	Illegal taking/ removal of marine fauna	1									1								
F05.04	Poaching							1	1								1	1	
G01	Outdoor sports and leisure activities, recreational activities	1									1								
G01.01	Nautical sports					1									1				
G01.01.0 1	Motorized nautical sports	1									1								
G01.03	Motorised vehicles		1			1						1			1				
G01.04.0 3	Recreational cave visit is							1									1		
G02.08	Camping and caravans					1									1				
G03	Interpretative centres		1									1							
G05	Other human intrusions and disturbances		1						1			1						1	
G05.01	Trampling, overuse					1									1				
G05.11	Death or injury by collision							1									1		
H01	Pollution to surface waters (limnic & terrestrial, marine & brackish)				1									1					
H01.05	Diffuse pollution to surface waters due to agricultural and forestry activities	1									1								
H01.08	Diffuse pollution to surface waters due to household sewage and waste waters	1									1								
H03.01	Oil spills in the sea				1									1					
H03.03	Marine macro-pollution (i.e. Plastic bags, styrofoam)	1			1			1			1			1			1		
H06.01	Noise nuisance, noise pollution	1							1		1							1	
I01	Invasive non-native species	1			1				1		1			1				1	
J01	Fire and fire suppression					1			1	1					1			1	1
J02.03	Canalisation & water deviation									1									1

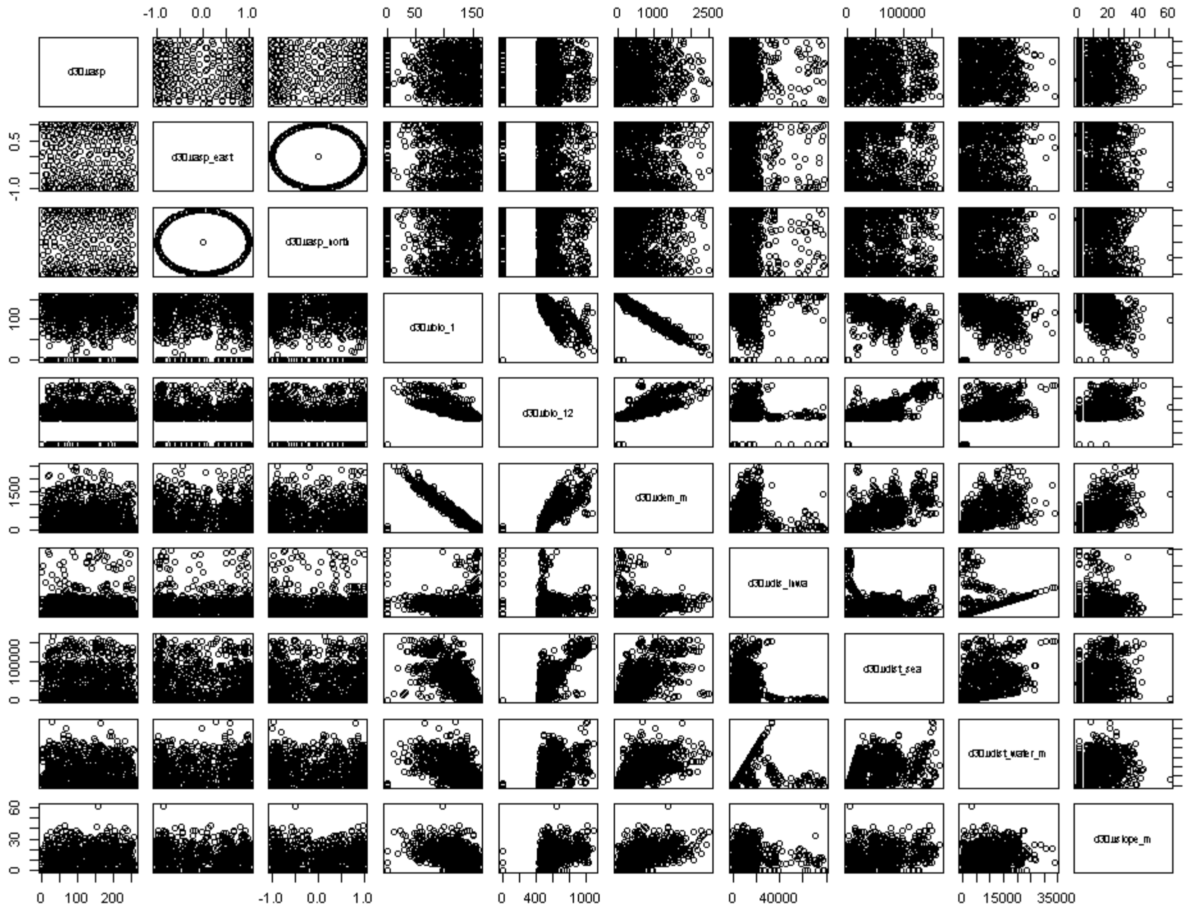
J02.03.02	Canalisation								1									1	
J02.05	Modification of hydrographic functioning, general		1									1							
J02.05.05	Small hydropower projects, weirs							1									1		
J02.06.01	Surface water abstractions for agriculture								1									1	
J03.01	Reduction or loss of specific habitat features	1	1								1	1							
J03.02	Anthropogenic reduction of habitat connectivity				1					1				1					1
K01.01	Erosion	1									1								
K01.02	Silting up									1									1
K02.01	Species composition change (succession)	1						1			1						1		
K02.03	Eutrophication (natural)					1		1	1						1		1	1	
K03	Interspecific faunal relations	1									1								
K03.06	Antagonism with domestic animals							1									1		
L04	Avalanche					1									1				
L09	Fire (natural)									1									1
M01.02	Droughts and less precipitations				1									1					

ANNEX 4: Graphs and tables: Boxplot, correlation, importance variables and threats, ROC-AUC plots.

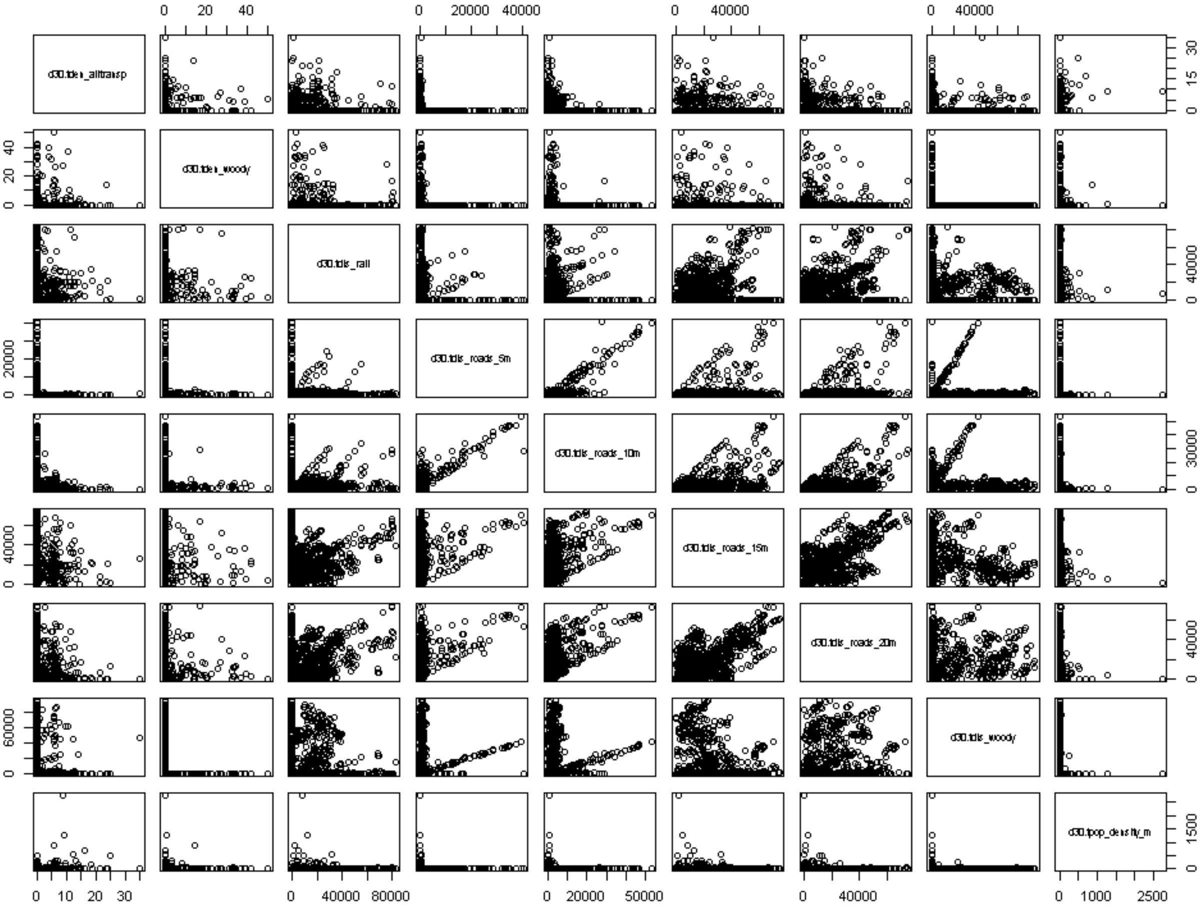
30 x 30 Resolution



Variables



Threats



Importance variables

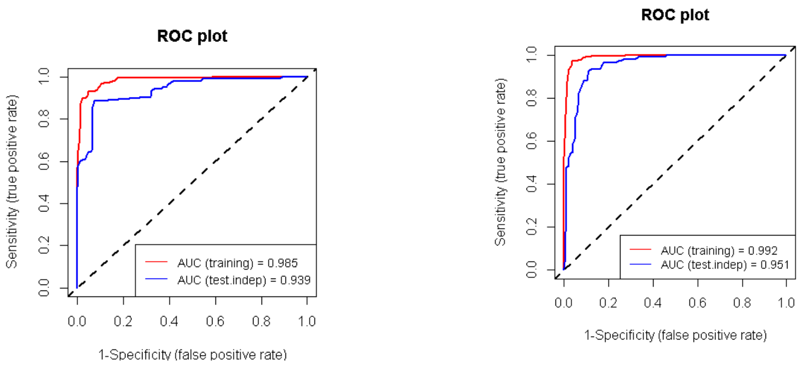
Pr |z| values for the environmental variables, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). Grey cells are variables that the stepwise rejected.

	30		step (30)	
	Pr(>|z|)		Pr(>|z|)	
(Intercept)	0.88537		0.19397	
Altitude	0.00211	**	0.00131	**
slope	0.00338	**	0.00324	**
Combined land cover category	0.99429		9.88E-01	
Continuous urban fabric	NA			
Discontinuous urban fabric	0.9976		0.99758	
Non-irrigated arable land	0.99921			
Permanently irrigated land	0.99962			
Rice fields	NA			
Inland marshes	0.99518		9.95E-01	
Salt marshes	0.99224		9.92E-01	
Fluvial clays, silts and loams	0.0485	*	5.36E-02	.
No water nor No wet area	0.3805			
	AIC: 131.84		**AIC: 126.64**	

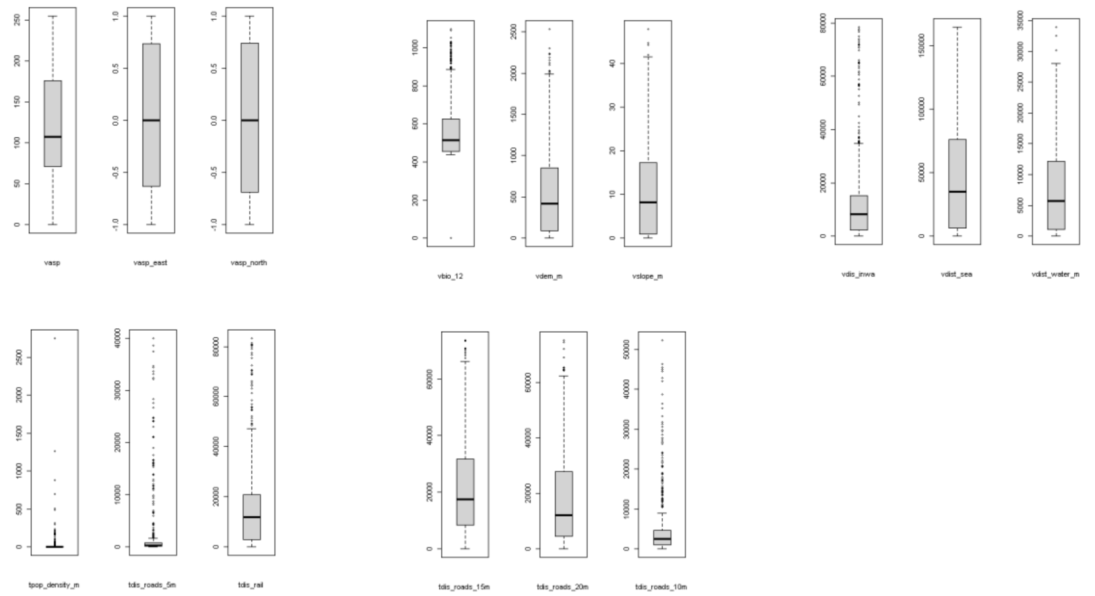
Pr |z| values for the environmental variables and threats, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1). Grey cells are variables that the stepwise rejected.

	30 th Pr(>|z|)	step (30 th) Pr(>|z|)
(Intercept)	0.0448 *	0.0051 **
Altitude	0.3766	
Slope	0.0782 .	0.00501 **
Combined land cover category	0.9971	0.98849
Continuous urban fabric	NA	
Discontinuous urban fabric	0.9995	
Non-irrigated arable land	1	
Permanently irrigated land	0.9997	
Rice fields	NA	
Inland marshes	0.9985	0.99907
Salt marshes	0.9969	0.99786
Fluvial clays, silts and loams	0.7178	
No water nor No wet area	0.1784	0.03058 *
Area with changes of land cover on the last 20 years	0.455	0.18051
Density of are covered by all roads or communication routes	0.2679	0.00307 **
Total area cover density of small woody features	0.3794	
Distance from railways	0.0833 .	0.07563 .
Distance from roads of 5m width	0.315	
Distance from roads of 10m width	0.7759	
Distance from roads of 15m width	0.1624	0.01851 *
Distance from roads of 20m width	0.0866 .	0.01231 *
Distance from small woody features	0.2767	0.11593
Total population	0.7883	
	AIC: 131.84	**AIC: 126.64**

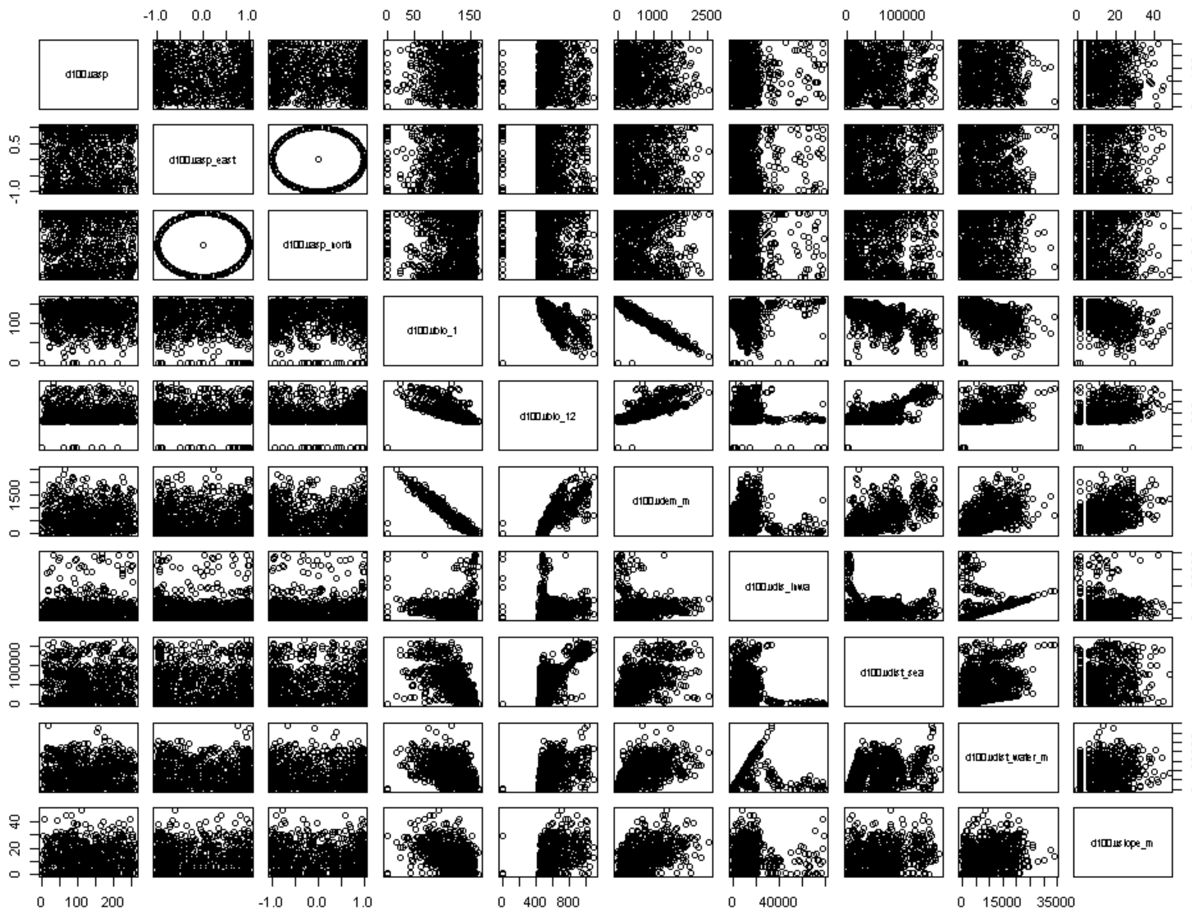
Accuracy ROC AUC plots



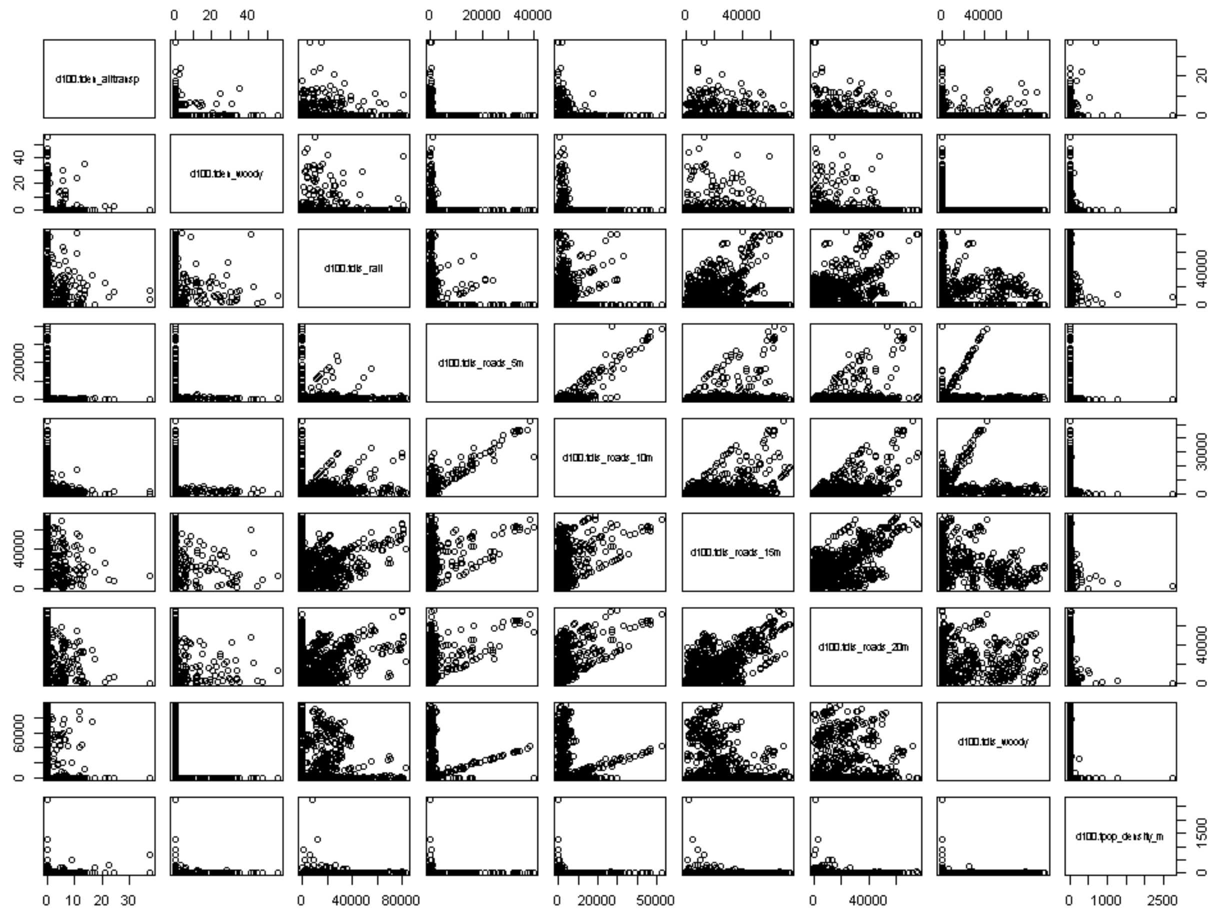
(left) ROC curve model fitted with environmental variables (right) ROC curve model fitted with environmental variables and threats at 30 resolution.


vasp
vasp_east
vasp_north
vbio_12
vdem_m
vslope_m
vdis_inwa
vdist_sea
vdist_water_m
tpop_density_m
tdis_roads_5m
tdis_rail
tdis_roads_15m
tdis_roads_20m
tdis_roads_10m

Variables



Threats



Importance variables

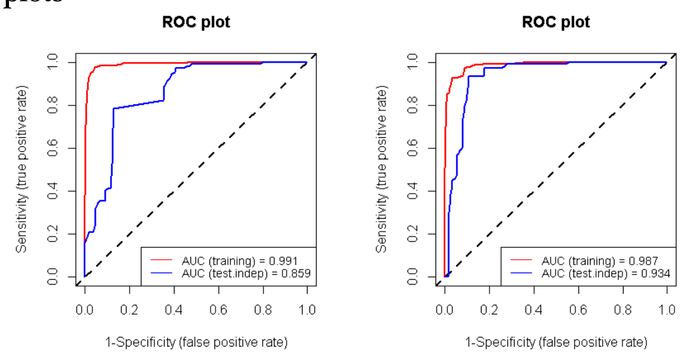
Pr |z| values for the environmental variables, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1). Grey cells are variables that the stepwise rejected.

	100		step (100)	
	Pr(>|z|)		Pr(>|z|)	
(Intercept)	0.528687		0.522064	
Attitude	0.000159	***	0.000166	***
Slope	0.000601	***	0.000613	***
Combined land cover category	0.991302		5.83E-08	***
Continuous urban fabric	NA			
Discontinuous urban fabric	0.996805		0.996811	
Non-irrigated arable land	0.995123			
Permanently irrigated land	0.998889			
Rice fields	NA			
Inland marshes	0.992581		0.992602	
Salt marshes	0.990156		0.990331	
Fluvial clays, silts and loams	0.0000065	***	0.00000581	***
No water nor No wet area	0.087223	.	0.086535	.
	AIC: 201.79		**AIC: 197.97**	

Pr |z| values for the environmental variables and threats, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). Grey cells are variables that the stepwise rejected.

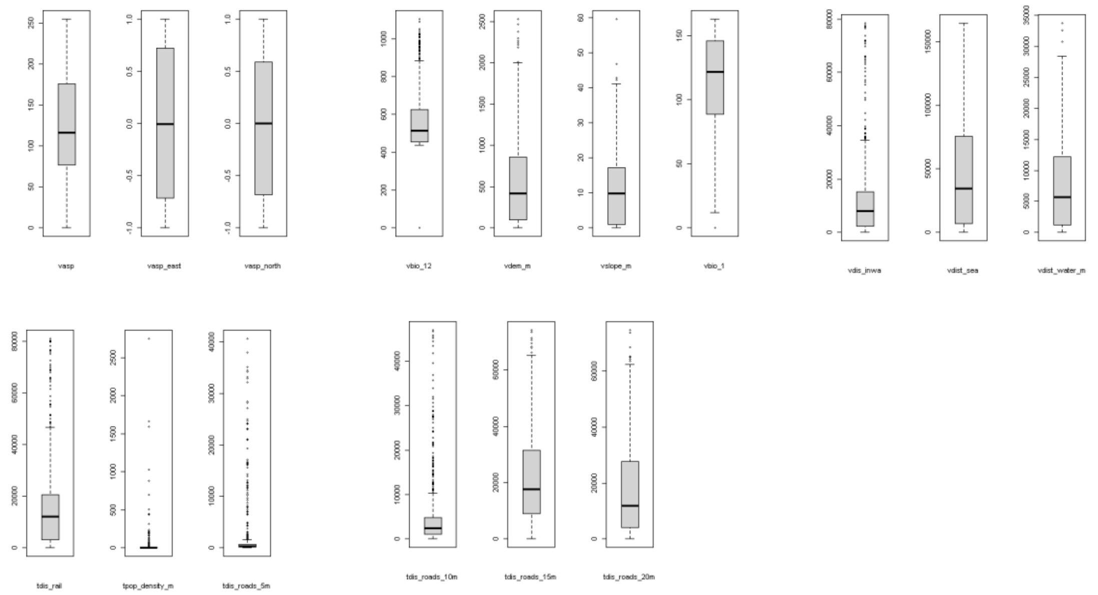
	100 th Pr(>|z|)		step (100 th) Pr(>|z|)	
(Intercept)	0.351896		0.484789	
Altitude	0.000135	***	0.000172	***
Slope	0.002336	**	0.001713	**
Combined land cover category	0.994614		0.985678	
Continuous urban fabric	NA			
Discontinuous urban fabric	0.998154			
Non-irrigated arable land	0.997237		0.989958	
Permanently irrigated land	0.997632			
Rice fields	NA			
Inland marshes	0.99234		0.992268	
Salt marshes	0.993655		0.993908	
Fluvial clays, silts and loams	0.070978	.	0.015596	*
No water nor No wet area	0.086878	.	0.090464	.
Area with changes of land cover on the last 20 years	0.244901			
Density of are covered by all roads or communication routes	0.032664	*	0.016351	*
Total area cover density of small woody features	0.135534		0.126774	
Distance from railways	0.090119	.	0.061102	.
Distance from roads of 5m width	0.00849	**	0.006946	**
Distance from roads of 10m width	0.0000199	***	0.0000293	***
Distance from roads of 15m width	0.119067		0.134382	
Distance from roads of 20m width	0.110575		1.10E-01	
Distance from small woody features	0.209642		1.99E-01	
Total population	0.00049	***	0.000262	***
	AIC: 148.6		**AIC: 144.2**	

Accuracy ROC AUC plots

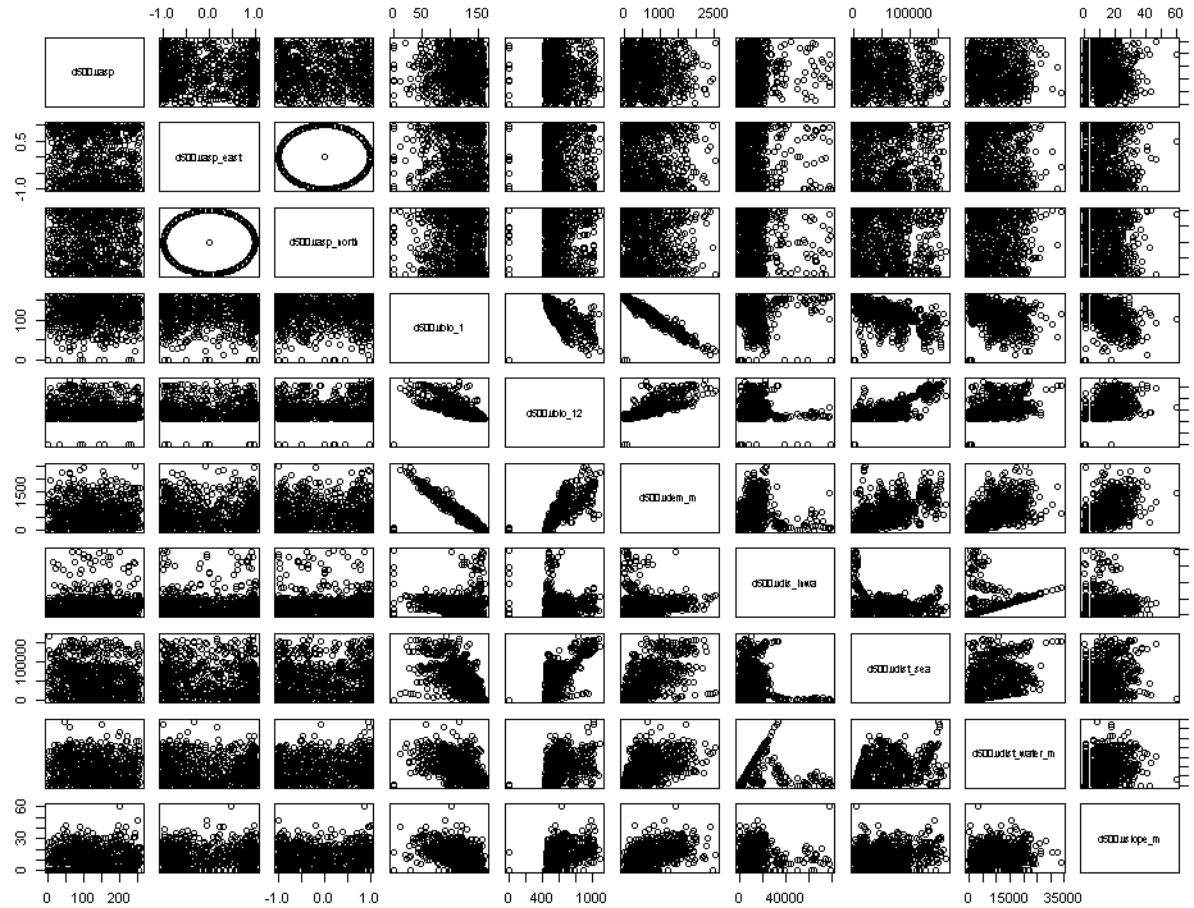


(left) ROC curve model fitted with environmental variables (right) ROC curve model fitted with environmental variables and threats at resolution 100

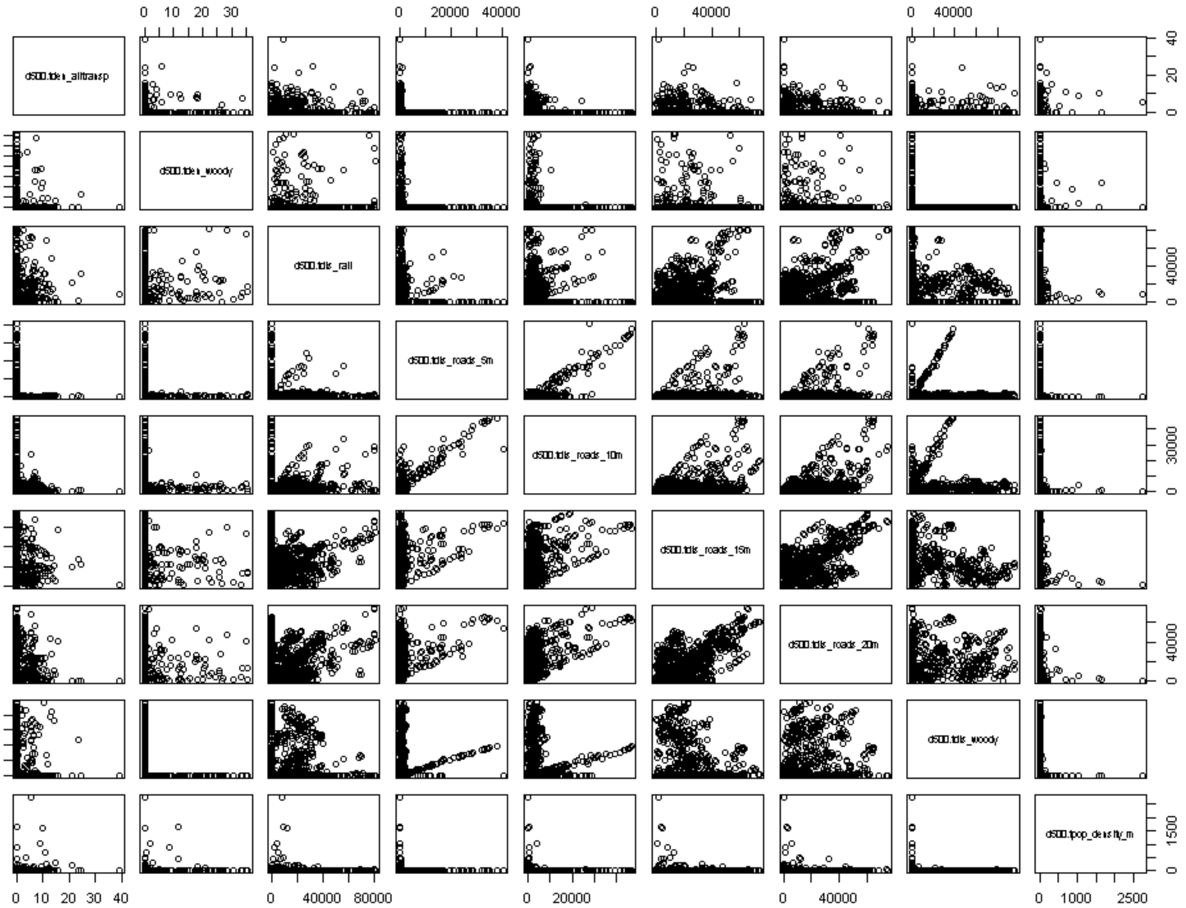
500 x 500 Resolution



Variables



Threats



Importance variables

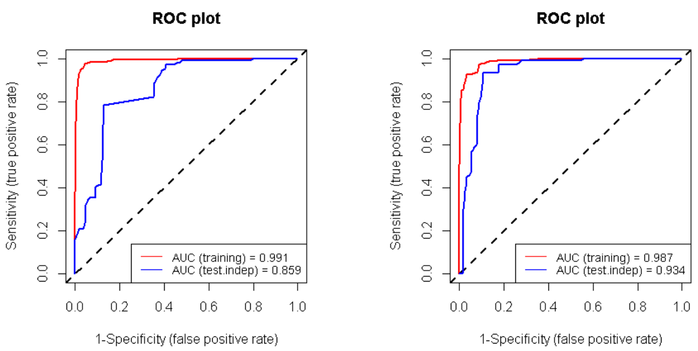
Pr |z| values for the environmental variables, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). Grey cells are variables that the stepwise rejected.

	500		step(500)	
	Pr(>|z|)		Pr(>|z|)	
(Intercept)	0.0000398	***	0.0001	***
Altitude	0.000802	***	0.000179	***
Slope	3.13E-08	***	0.0000035	***
Combined land cover category	0.992571		9.17E-01	
Continuous urban fabric	NA		NA	
Discontinuous urban fabric	0.292168		3.59E-01	
Non-irrigated arable land	0.993436		9.17E-01	
Permanently irrigated land	0.993778		9.17E-01	
Rice fields	NA		9.17E-01	
Inland marshes	0.039668	*		
Salt marshes	0.988179		1.00E+00	
Fluvial clays, silts and loams	1.09E-08	***	6.47E-09	***
No water nor No wet area	1.33E-06	***	3.72E-06	***
	309.42		**316.48**	

(left) Pr |z| values for the environmental variables and threats, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). Grey cells are variables that the stepwise rejected.

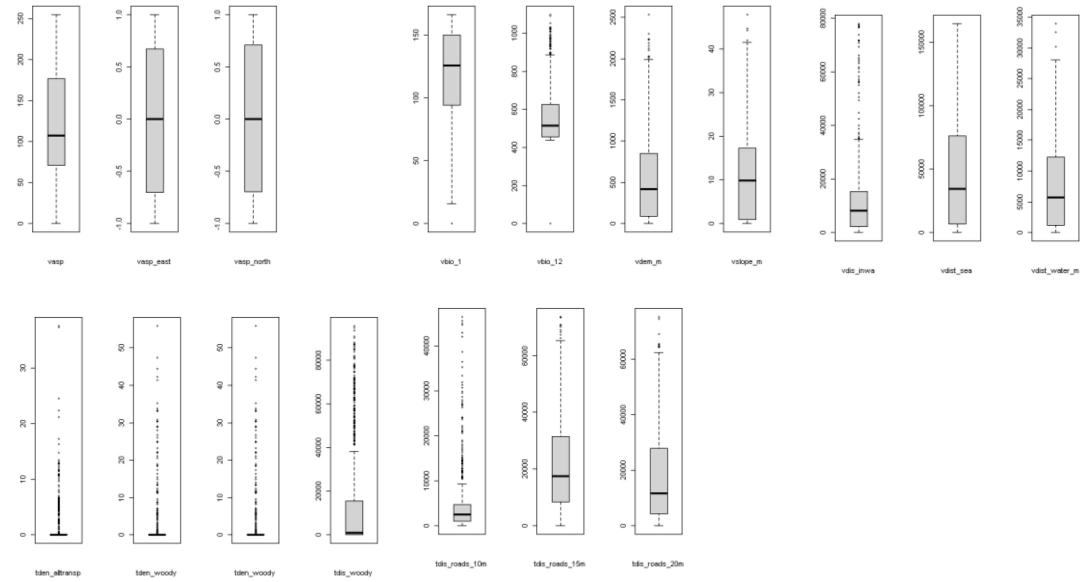
	500 th $\Pr(>\lvert z\rvert)$		**step (500 th)** $\Pr(>\lvert z\rvert)$	
(Intercept)	0.640723		0.697526	
Altitude	0.004062	**	0.001421	**
Slope	0.0094	**	0.004237	**
Combined land cover category	0.99263		0.000289	***
Continuous urban fabric	NA			
Discontinuous urban fabric	0.303116		0.232922	
Non-irrigated arable land	0.993798		0.161521	
Permanently irrigated land	0.994512			
Rice fields	NA			
Inland marshes	0.648365		0.646533	
Salt marshes	0.987437		0.987376	
Fluvial clays, silts and loams	0.000643	***	0.00035	***
No water nor No wet area	0.07596	.	0.063085	.
Area with changes of land cover on the last 20 years	0.029771	*	0.022585	*
Density of are covered by all roads or communication routes	0.724894			
Total area cover density of small woody features	0.076859	.	0.063654	.
Distance from railways	0.541307			
Distance from roads of 5m width	0.056222	.	0.052134	.
Distance from roads of 10m width	5.53E-08	***	3.39E-08	***
Distance from roads of 15m width	0.209475			
Distance from roads of 20m width	0.0000116	***	0.000002	***
Distance from small woody features	0.165501		0.161017	
Total population	0.000358	***	0.000203	***
	AIC: 225.85		**AIC: 219.66**	

Accuracy ROC AUC plots

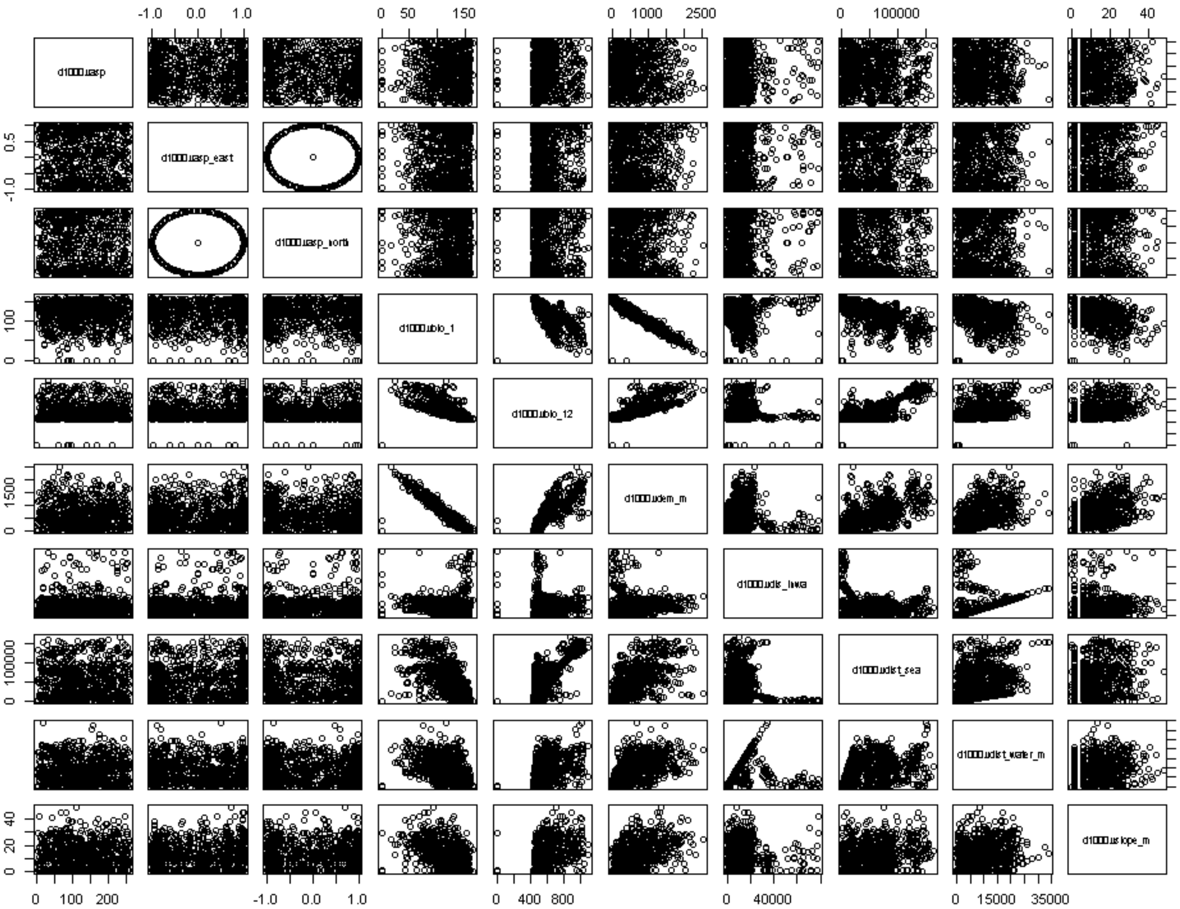


(left) ROC curve model fitted with environmental variables (right) ROC curve model fitted with environmental variables and threats at resolution 500

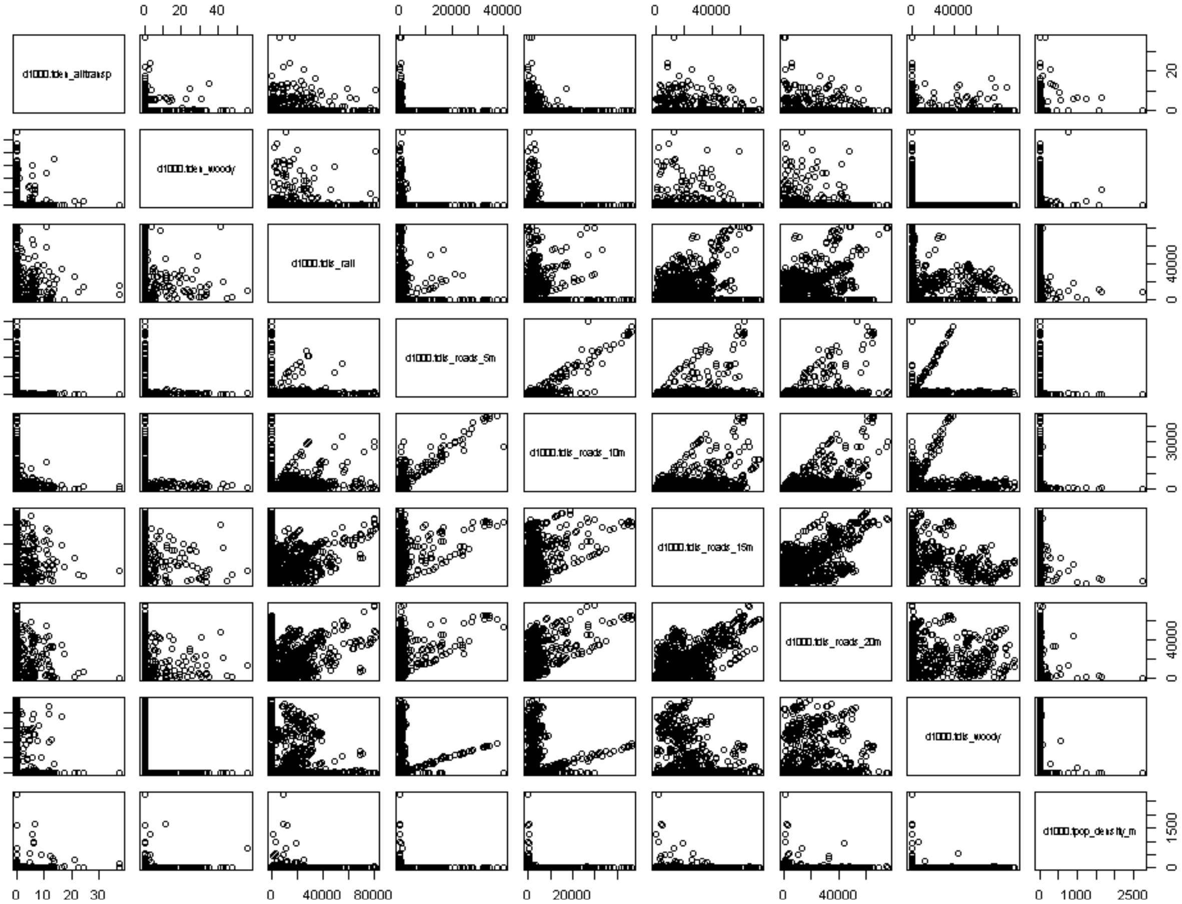
1000 x 1000 Resolution



Variables



Threats



Importance variables

Pr |z| values for the environmental variables, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1) Grey cells are variables that the stepwise rejected.

	1000 Pr(>|z|)		step (1000) Pr(>|z|)	
(Intercept)	0.600635		0.625863	
Altitude	5.64E-08	***	1.41E-08	***
Slope	0.0000777	***	0.0000826	***
Combined land cover category	0.983637		3.23E-09	***
Continuous urban fabric	NA			
Discontinuous urban fabric	0.924861			
Non-irrigated arable land	0.98694			
Permanently irrigated land	0.985577		0.056104	.
Rice fields	NA			
Inland marshes	0.997653		0.986244	
Salt marshes	0.994744			
Fluvial clays, silts and loams	1.02E-14	***	2.61E-15	***
No water nor No wet area	0.000141	***	0.000102	***
	AIC: 297.05		**AIC: 292.93**	

Pr |z| values for the environmental variables and threats, before and after the stepwise regression. (Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). Grey cells are variables that the stepwise rejected.

	1000 th Pr(>|z|)		step (1000 th) Pr(>|z|)	
(Intercept)	0.997238		0.890304	
Altitude	0.012279	*	0.006815	**
Slope	0.03137	*	0.02284	*
Combined land cover category	0.990191		4.35E-08	***
Continuous urban fabric	NA			
Discontinuous urban fabric	0.522514			
Non-irrigated arable land	0.992186			
Permanently irrigated land	0.992032			
Rice fields	NA			
Inland marshes	0.998399		0.984723	
Salt marshes	0.996636			
Fluvial clays, silts and loams	2.74E-09	***	1E-09	***
No water nor No wet area	0.01803	*	0.020317	*
Area with changes of land cover on the last 20 years	0.028997	*	0.035133	*
Density of are covered by all roads or communication routes	0.62539			
Total area cover density of small woody features	0.97394			
Distance from railways	0.149929		0.147887	
Distance from roads of 5m width	0.118583		0.106378	
Distance from roads of 10m width	0.001138	**	0.001042	**
Distance from roads of 15m width	0.134964		0.108484	
Distance from roads of 20m width	0.000151	***	0.000144	***
Distance from small woody features	0.039847	*	0.0364	*
Total population	0.492277			
	AIC: 244.49		AIC: 232.99	

Accuracy ROC AUC plots

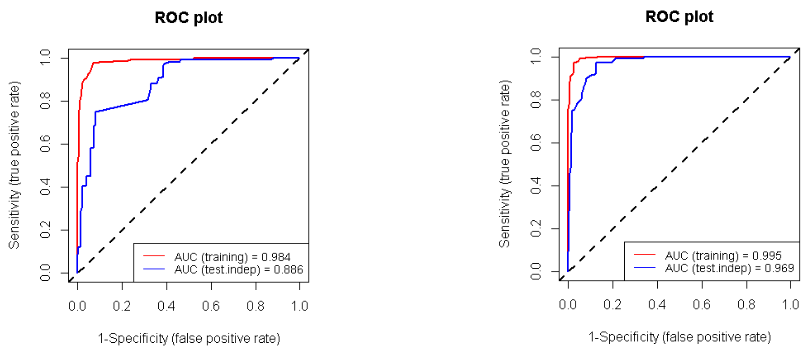
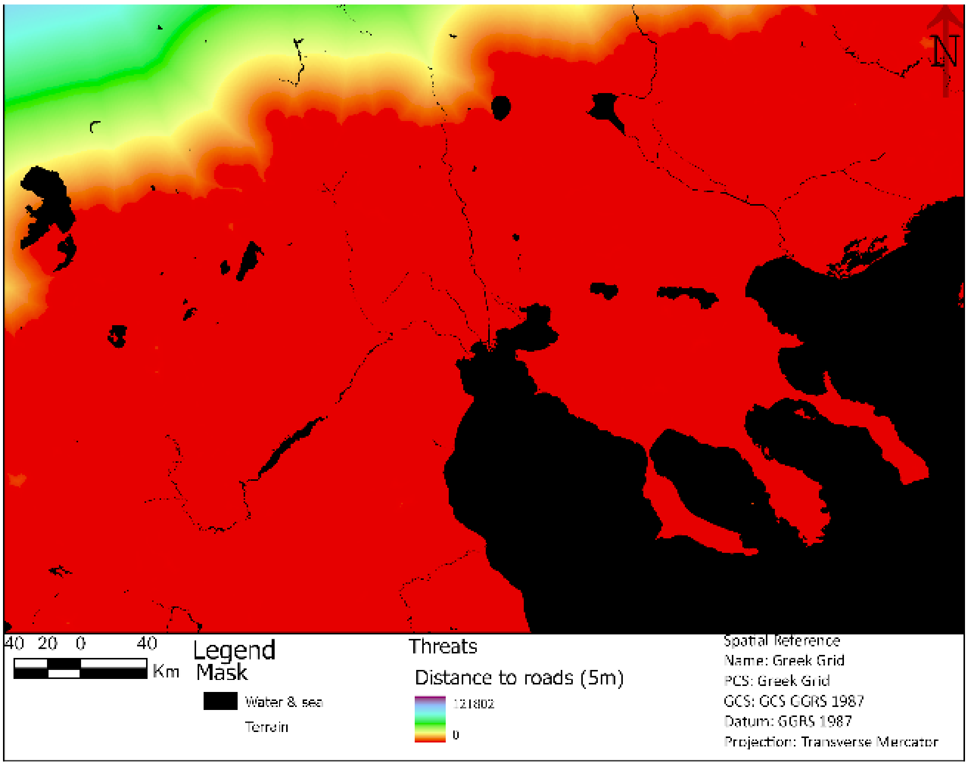
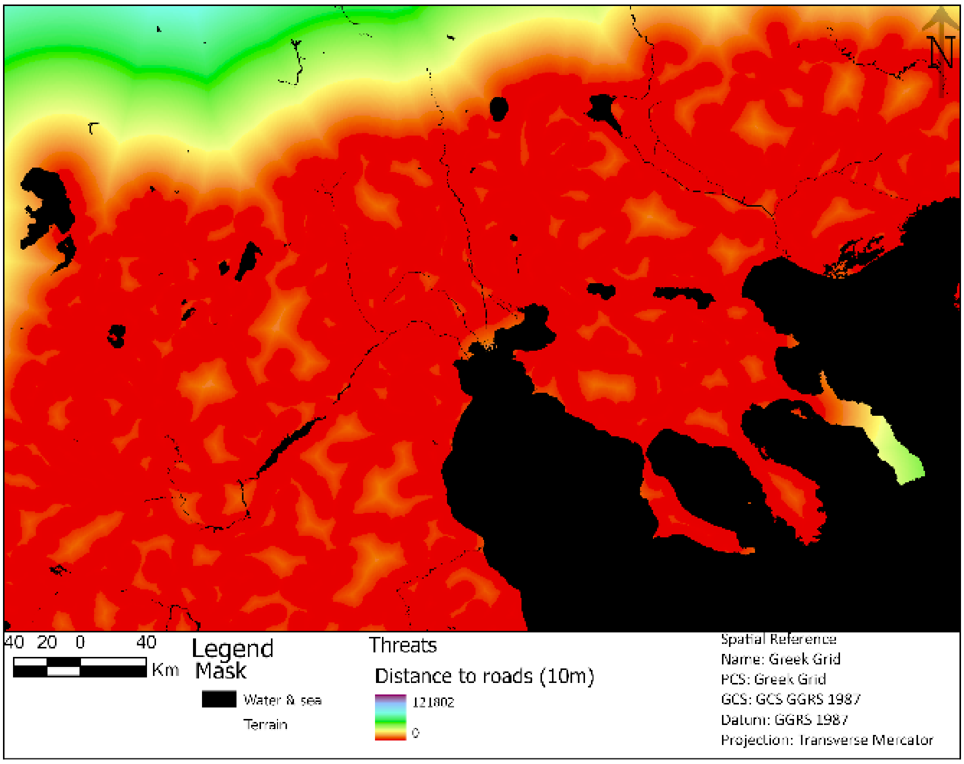
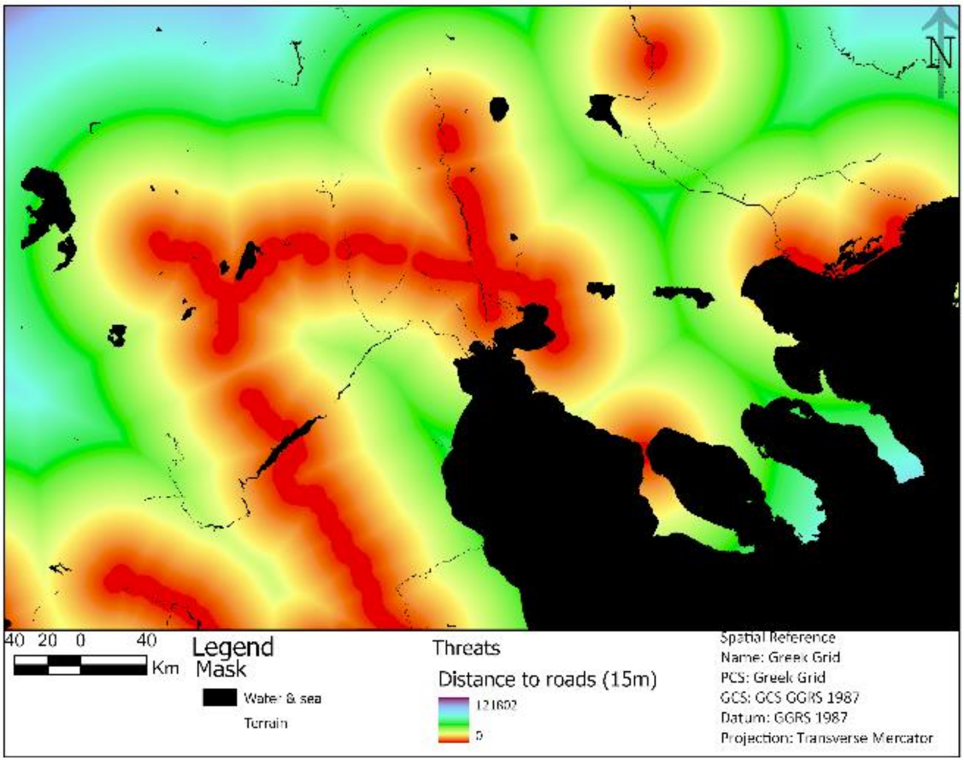


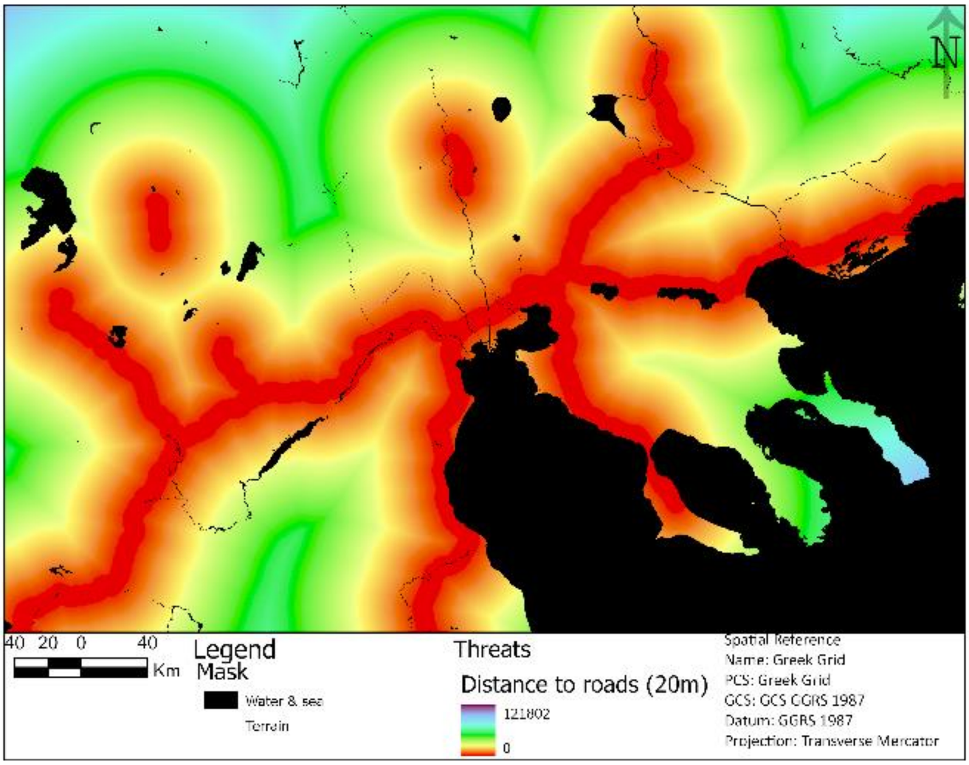
Figure 21 (left) ROC curve model fitted with environmental variables (right) ROC curve model fitted with environmental variables and threats at resolution 1000.

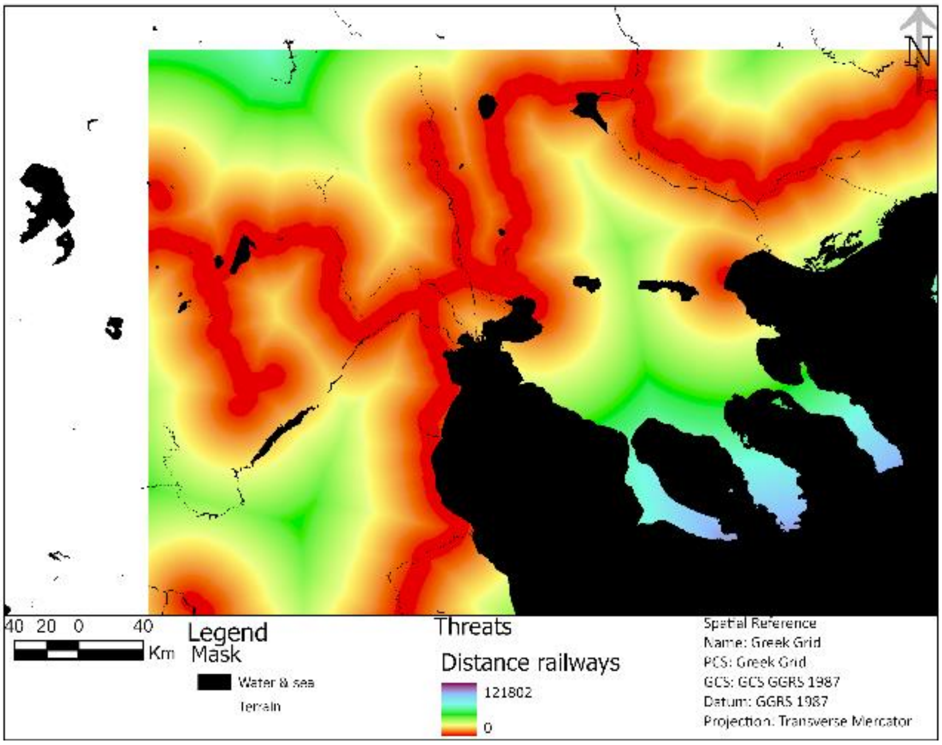
ANNEX 5: Threat Maps

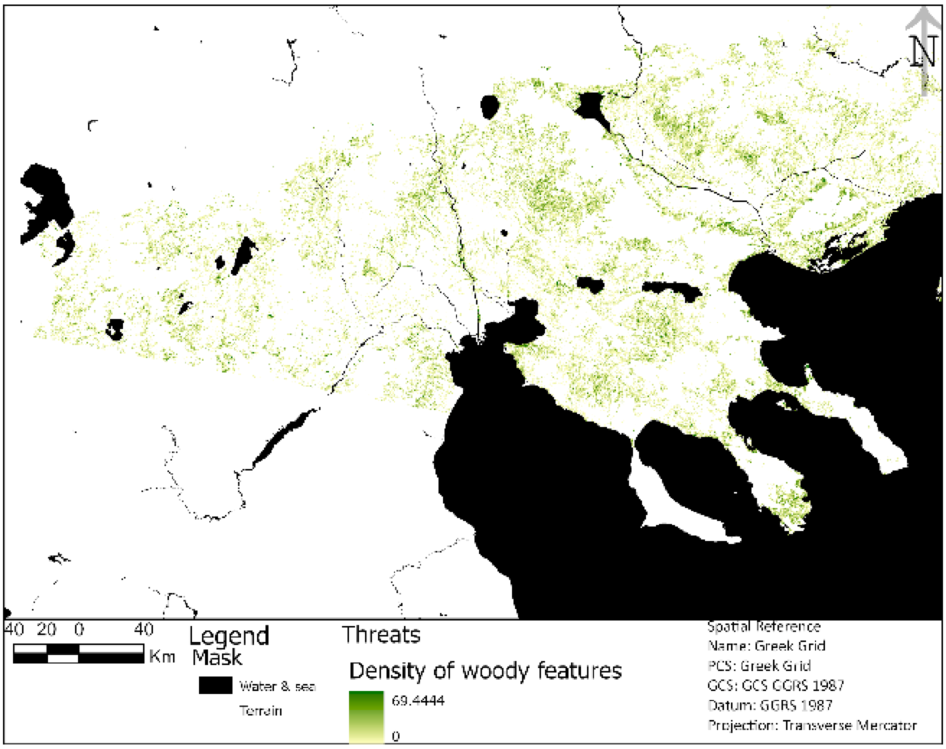


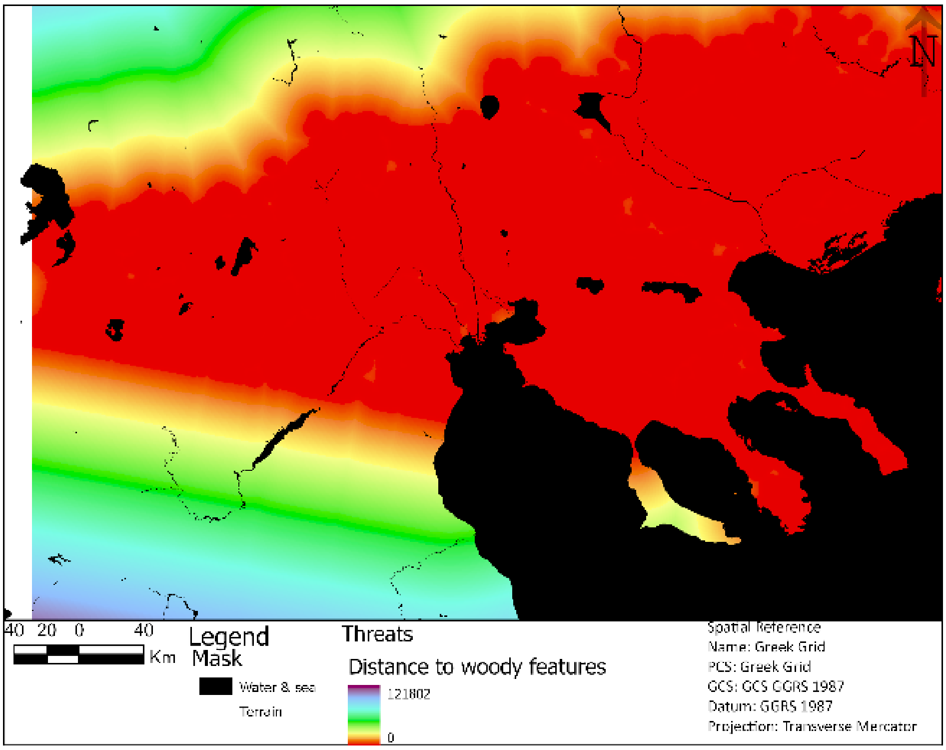


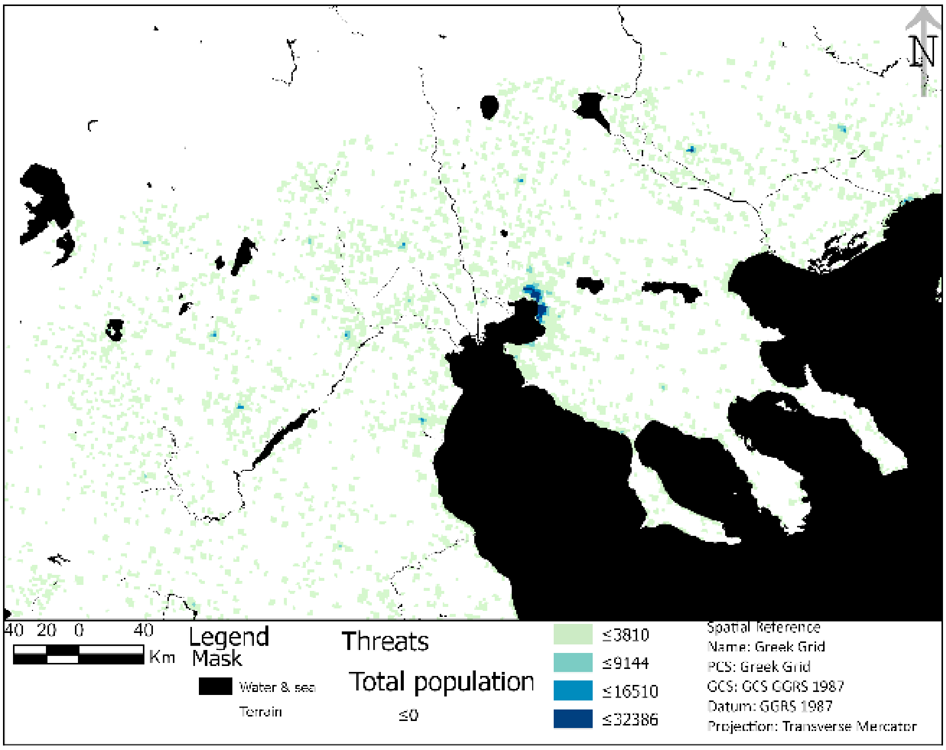

N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Distance to roads (15m)
121802
0
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator

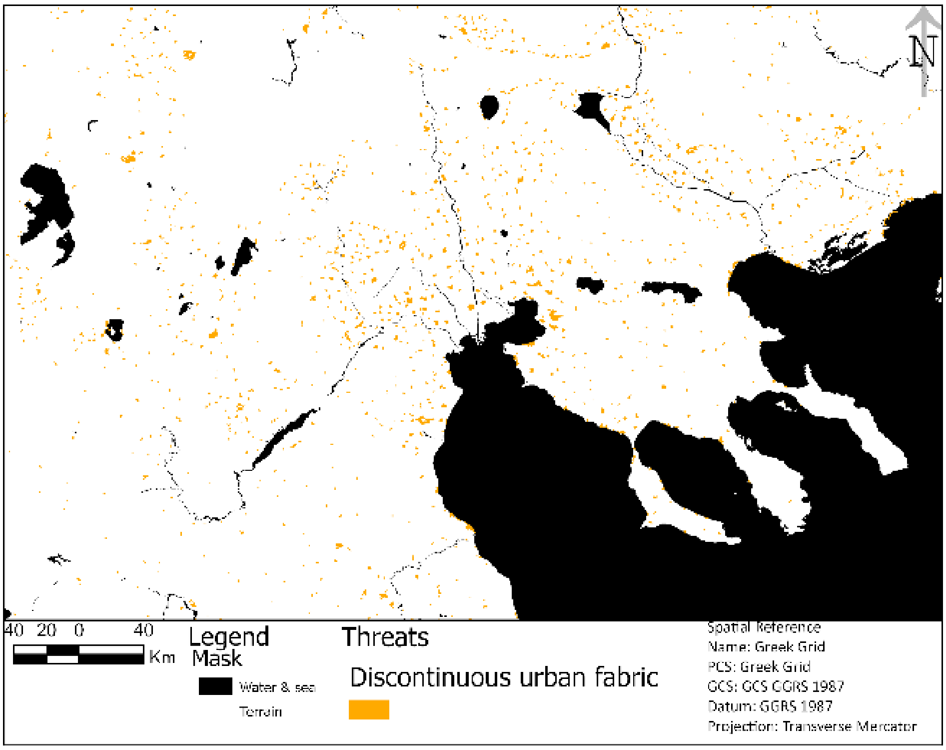

N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Distance to roads (20m)
121802
0
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator

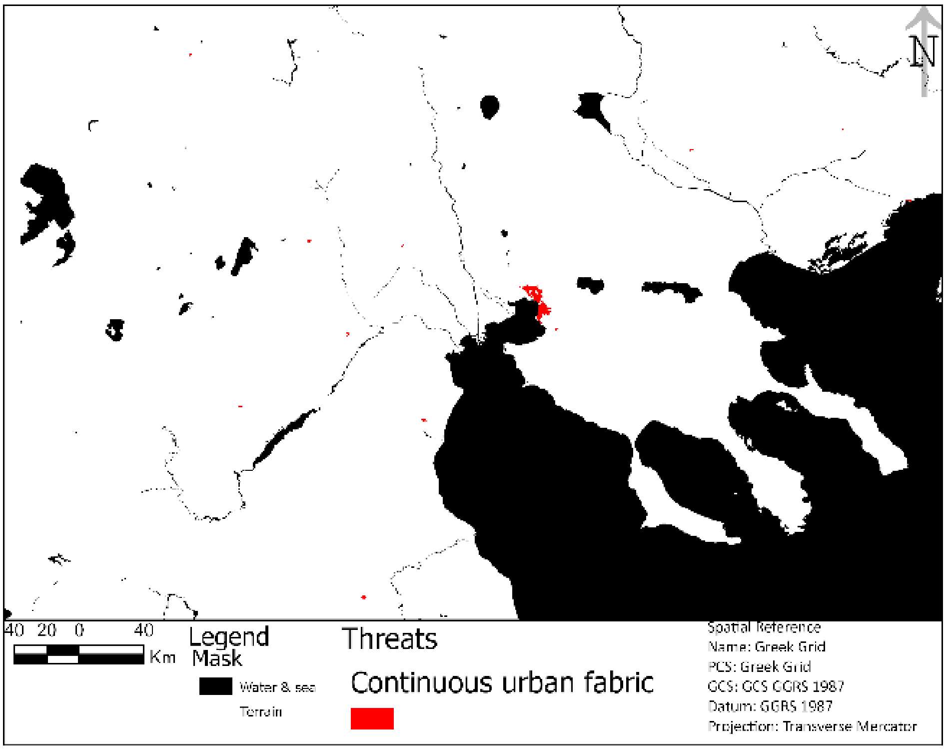

N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Distance railways
121802
0
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator


N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Density of woody features
69.4444
0
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator


N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Distance to woody features
121802
0
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator


N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Total population
≤0
≤3810
≤9144
≤16510
≤32386
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator


N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Discontinuous urban fabric
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator


N
40 20 0 40
Km
Legend
Mask
Water & sea
Terrain
Threats
Continuous urban fabric
Spatial Reference
Name: Greek Grid
PCS: Greek Grid
GCS: GCS GGRS 1987
Datum: GGRS 1987
Projection: Transverse Mercator

ANNEX 6: Modelled Probability inside Protected Areas in the Natura 2000 Network Maximum, Minimum and Mean. The colour scale goes from green for the lowest value (0) yellow for the middle (0.5) and red for the highest value (1).

	MAX								MIN								MEAN							
SITECODE	**30**	**30th**	**100**	**100th**	**500**	**500th**	**1000**	**1000th**	**30**	**30th**	**100**	**100th**	**500**	**500th**	**1000**	**1000th**	**30**	**30th**	**100**	**100th**	**500**	**500th**	**1000**	**1000th**
GR1150005	1.00	1.00	1.00	1.00	0.96	0.86	0.94	0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.10	0.04	0.05	0.04	0.02	0.03	0.01
GR1210001	1.00	1.00	0.98	0.18	0.20	0.04	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.09	0.00	0.00	0.00	0.00	0.00	0.00
GR1210002	1.00	1.00	0.99	1.00	0.82	0.93	0.42	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.32	0.04	0.05	0.04	0.09	0.03	0.06
GR1220001	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.99**	**1.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.70**	**0.71**	**0.65**	**0.40**	**0.71**	**0.57**	**0.58**	**0.66**
GR1220002	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.81**	**0.77**	**0.76**	**0.73**	**0.78**	**0.74**	**0.71**	**0.75**
GR1220003	1.00	1.00	1.00	1.00	0.78	1.00	0.75	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.29	0.10	0.12	0.09	0.17	0.09	0.13
GR1220012	**0.61**	**1.00**	**1.00**	**0.90**	**0.76**	**0.77**	**0.81**	**0.75**	**0.11**	**0.03**	**0.03**	**0.12**	**0.19**	**0.47**	**0.10**	**0.13**	**0.56**	**0.67**	**0.38**	**0.48**	**0.51**	**0.65**	**0.37**	**0.35**
GR1230001	1.00	1.00	1.00	0.93	0.96	0.69	0.99	0.95	0.55	0.00	0.16	0.24	0.59	0.12	0.00	0.12	0.82	0.45	0.74	0.68	0.85	0.40	0.68	0.62
GR1230002	1.00	1.00	1.00	0.90	0.94	0.13	0.91	0.66	0.05	0.00	0.08	0.13	0.17	0.01	0.06	0.02	0.50	0.29	0.67	0.42	0.60	0.05	0.41	0.22
GR1240001	1.00	1.00	0.68	0.93	0.49	0.00	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00
GR1240002	1.00	1.00	1.00	0.87	0.67	0.03	0.58	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.01	0.01	0.01	0.00	0.02	0.00
GR1240003	1.00	1.00	1.00	0.97	0.14	0.03	0.17	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.01	0.00	0.00	0.00	0.00	0.00
GR1240004	0.12	1.00	0.04	0.91	0.37	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.08	0.01	0.11	0.05	0.00	0.00	0.00
GR1240005	1.00	1.00	1.00	1.00	0.94	0.94	0.73	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.08	0.14	0.18	0.14	0.02	0.16	0.00
GR1250002	1.00	1.00	1.00	1.00	0.93	0.95	0.98	0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.14	0.02	0.01	0.01	0.01	0.02	0.00
GR1250003	0.02	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00
GR1250004	**1.00**	**1.00**	**1.00**	**0.97**	**1.00**	**0.97**	**0.96**	**0.95**	**0.50**	**0.35**	**0.20**	**0.12**	**0.69**	**0.22**	**0.46**	**0.36**	**0.91**	**0.85**	**0.79**	**0.78**	**0.90**	**0.79**	**0.76**	**0.73**
GR1260001	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.45	0.31	0.25	0.27	0.28	0.15	0.28	0.21
GR1260003	1.00	1.00	1.00	1.00	0.96	0.96	0.99	0.94	1.00	0.00	0.00	0.17	0.01	0.01	0.95	0.92	1.00	0.69	0.86	0.90	0.82	0.67	0.96	0.93
GR1260004	1.00	1.00	1.00	0.88	0.77	0.05	0.96	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.03	0.01	0.01	0.01	0.00	0.02	0.00
GR1260005	1.00	1.00	0.90	0.06	0.02	0.01	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.01	0.01	0.00	0.00	0.00	0.00	0.00
GR1260007	1.00	1.00	1.00	1.00	0.71	0.54	0.91	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.12	0.03	0.03	0.02	0.02	0.05	0.06
GR1270001	1.00	1.00	0.99	1.00	0.26	0.97	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.09	0.02	0.01	0.00	0.00	0.00	0.00
GR1270002	1.00	1.00	1.00	0.64	0.78	0.09	0.79	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.10	0.06	0.02	0.08	0.00	0.05	0.00
GR1270003	0.61	1.00	1.00	1.00	0.76	0.99	0.46	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.07	0.01	0.14	0.02	0.04	0.02	0.00
GR1270005	1.00	1.00	0.50	0.74	0.76	0.16	0.12	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.13	0.01	0.02	0.02	0.01	0.01	0.00
GR1270007	1.00	1.00	0.89	0.53	0.76	0.04	0.46	0.03	0.50	0.00	0.20	0.04	0.25	0.01	0.19	0.01	0.62	0.18	0.50	0.09	0.73	0.02	0.40	0.02

GR1270008	1.00	1.00	0.92	0.93	0.83	0.72	0.96	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.00	0.61	0.36	0.50	0.07	0.75	0.08	0.46	0.00
GR1270009	0.61	1.00	0.50	1.00	0.76	0.00	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.61	0.05	0.48	0.10	0.72	0.00	0.41	0.00
GR1270010	1.00	1.00	0.92	0.82	0.76	0.92	0.46	0.20	0.61	0.41	0.04	0.06	0.00	0.01	0.41	0.01	0.61	0.96	0.50	0.22	0.74	0.72	0.46	0.09
GR1270015	0.61	1.00	0.50	1.00	0.76	1.00	0.46	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61	0.04	0.49	0.76	0.74	0.40	0.44	0.03
GR1340004	1.00	1.00	1.00	1.00	0.86	1.00	0.83	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.35	0.06	0.07	0.19	0.02	0.05	0.01
GR1420001	1.00	1.00	1.00	1.00	0.96	0.96	0.99	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.18	0.01	0.04	0.02	0.03	0.05	0.00