Survey tools to predict and monitor the presence of brown bears (*Ursus arctos*) in Greece: a case study integrating spatial analysis and invertebrate-derived DNA

Angeliki Savvantoglou

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

Centre for Research in Biosciences, Department of Applied Sciences,
Faculty of Health and Applied Sciences, University of the West of England, Bristol

September 2022

Author's Declaration

As far as possible, where permission is required, I confirm that I have gained third party copyright permission for all relevant material and that I have attributed copyright permissions within the thesis where required.

All artwork included in this thesis (bear illustrations) was created by A. Savvantoglou.

Abstract

Increasing populations of brown bear (*Ursus arctos*) in Greece have resulted in recolonisation of areas previously unoccupied for decades. Large carnivore recolonisation often threatens hard-established human-wildlife coexistence efforts, therefore the need to monitor and predict bear presence and movement becomes more apparent as the animals make use of corridors existing between their current range and other suitable habitats. The challenge of carrying out surveys in infrequently used areas, such as wildlife corridors, is compounded by limited funds for wildlife conservation. This calls for the exploration of alternative monitoring techniques that are more cost and time-efficient than the standard methods such as scat surveys. This study explored the use of habitat suitability modelling techniques in mapping brown bear ecological networks in Greece throughout their current range and sites of potential future recolonisation. Using these suitability maps as a guide, the study utilised an innovative non-invasive genetic monitoring technique, invertebrate-derived DNA (iDNA), to survey the species in the field. A single-species targeted qPCR approach was used and method development experiments were conducted to form a protocol to optimise its performance in the field. Next, iDNA surveys combined with scat surveys were conducted to model bear distribution in northern Greece using occupancy modelling. These models describe the probability of bear detection in a landscape that incorporates a core brown bear habitat, a presumed corridor and a recently recolonised area. A review of the laboratory experiments and a comparison of the resulting occupancy models indicated that iDNA can effectively monitor the presence of a species as well as be subsequently used in occupancy modelling analyses. iDNA can be seen as an effective and complementary method of assessing brown bear distribution to inform conservation strategies and has the potential to assist with the conservation monitoring of other bear species.

Acknowledgments:

It is humbling to think how much of this journey was made possible because of the work and help of other people. Today, I get to type these names here, before the pages of my work, and I am filled with gratitude.

I want to thank *Dr Mark Steer*, who calling my Director of Studies would be too narrow of a title to describe his role in this work. The joy your work gives you will forever inspire the way I approach research and the way I teach. Thank you for the trust, enthusiasm, creative ideas and honesty. Your kindness will never be taken for granted. You have helped me find net-throwing spiders in Madagascar, shown me what academia could be like and have given me the strength and tools to follow what is important to me. I will forever be grateful.

Likewise, *Dr David Fernandez* and *Dr Joel Allainguillaume* for being the other two pillars of this support system. Thank you for trusting me, even when I didn't know how to use a pipette properly. Thank you for standing behind mad ideas and adding to them instead of taking away. I am so grateful to have worked with scientists like you.

Dr Yorgos Mertzanis and the NGO Callisto for introducing me to bear research and always having a place for me. Yorgos, you have been instrumental to my career and my most beloved bear mentor. Looking forward to collaborating in Greek bear country.

The International Organisation for Bear Research and Management, both in welcoming a small undergraduate to a big conference in Alaska and in keeping me in its large family ever since. I feel so grateful to have connected with this most supportive bear family. Thank you for the enthusiasm, inspiration, financial support, and trust in my work, even though it was more about flies than it was about bears sometimes.

Kostas and Alina, my parents. Their names belong here more than anyone else's. Their love for nature, kindness and joy of life lives though me. By being the people they were, they continue to shape my life, even in their absence. I owe this journey to my father and the way he lived his life. When good people pass on, they leave a sense of debt to those left behind, like we owe them the same goodness back that they freely gave when they were around. In the spirit of honouring his good life, a lot of my father's friends stepped up to support me and my siblings. A group of them decided that they want to help me fund my undergraduate studies in the UK, following what they knew I loved most. I was able to write this thesis thanks to them. *Maurizio, Kostas, Katerina, Mina, Vasilis, Irini, Panagiotis, Lisette, Antonia, Nina, Sofia, Sofia, Panagiotis, Andreas, Achilleas, and some that will remain anonymous here*. I am so humbled to be adding your names in the first pages of my thesis. Thank you for giving me my dream back.

Speaking of dreams, *Samantha and Ermis*, my siblings, thank you for always following yours with the force of a hurricane and inspiring me to do the same. You are both unstoppable and a constant reminder that nothing is impossible. Thank you for remaining the calm source of strength and humour in my life. You are so important. Thank you for the parcels of Greek coffee and handmade bowls that brought home to me when I needed it the most. And thank you for looking after my bear-sized dog while I am here, doing my part for the real bears.

The PGR community, especially *Laura, Eva, Josie, Josh, Buffy, Marina, Savanna and Dan*. No one stands in solidarity to a PhD student as well as another PhD student and I am so glad to have gained some of my most significant friends in this group. Thank you for all the good times in the office and the better times outside, planning our future while hiking up a hill or running across a forest. Thank you for the joy and your good hearts.

Laura and Buffy, for learning with me and being so generous with their knowledge and time. Laura, I would still be trying to figure out DNA extractions without your help. Thank you for the hours you spent teaching and sharing with me your knowledge and hard work. This would have never been possible without your guidance in the lab and I hope I can one day return the favour. Buffy, the iDNA world would be so lonely without you and I am so glad we have more things in common than just fly-crashing. So happy to be your friend and lab partner.

All my beloved people who supported me in their own way and all those who came bearing bear puns and memes. Thank you for being curious and for thinking of me when you see bear scat - it felt so supportive to know you cared.

Finally, my *beloved bears*. I am so fascinated by their ability to be such forces of change wherever they are, from the ecosystems they inhabit to the people they entertain with fat bear week. A force that always reminded me, in the hardest of times, why I did this. Bears are important and I feel privileged to have been in their company for the last few years. I may never leave.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	*Definition*
AIC	Aikake Information Criterion
AUC	Area Under the Curve
BSA	Bovine Serum Albumin
CI	Confidence Interval
DMSO	Dimethyl Sulfoxide
eDNA	Environmental DNA
GAM	Generalised Additive Model
GLM	General Linear Model
HSI	Habitat Suitability Index
HSM	Habitat Suitability Model/Modelling
iDNA	Invertebrate-derived DNA
LN	Liquid Nitrogen
LOD	Limit of Detection
MSE	Mean Square Error
mtDNA	Mitochondrial DNA
nDNA	Nucleic DNA
NGO	Non – Governmental Organisation
PCR	Polymerase Chain Reaction
PI	Predictor Importance
qPCR	Quantitative Polymerase Chain Reaction
SD	Standard Deviation
SDM	Species Distribution Model/Modelling
SE	Standard Error
ψ	Probability of occupancy

Figure 1. Brown bear surrounded by mosquitoes and other invertebrates. Photo © Drew Hamilton, used with permission.

Chapter 1: Introduction

The case study outlined in this thesis presents a protocol for the survey design, iDNA field sampling and occupancy modelling of single-species monitoring efforts.

1. Introduction

1.1 Monitoring for conservation

Habitat loss and fragmentation have been identified as major factors to species extinction and biodiversity loss (Fahrig, 1997; Groombridge and Jenkins, 2000; Wiegand, Revilla and Moloney, 2005). The urge to tackle the current species extinction events at a global or continental scale have been reflected in the efforts to prioritise conservation investments (Wilson et al., 2011) and inform conservation strategies (Henle et al., 2013; Rondinini, Rodrigues and Boitani, 2011). Realistically, all the work conducted to prioritise conservation and allocate funds where outcomes will be the most beneficial relies on monitoring efforts of wildlife and their habitats. Fundamentally, conservation starts with effective monitoring: understanding the presence and status of a species within an area over time (Long et al., 2012; Robinson et al., 2018; Stem et al., 2005; Tanentzap, Walker and Theo Stephens, 2017). Another component providing essential information to conservation planning is deciphering the study area's capacity to provide suitable and connected habitats for the species in question. These two aspects of conservation, therefore, provide the backbone of every strategy and are integral to the success of the project.

1.2 Non-invasive genetic monitoring

As a response to this conservation crisis, advances in survey techniques for wildlife and habitat monitoring have risen to meet the growing demand for species conservation. Non-invasive techniques have been on the forefront of this effort due to their ability to minimise disturbance of the target species. Here, the term 'non-invasive' is used to describe any method of collecting data that does "*not require animals to be directly observed or handled by the surveyor*"(Long et al., 2012). Due to this very broad description of the term, non-invasive techniques vary greatly in a practical sense in terms of the means used to collect data on a species, from scat surveys and remote

camera observations to hair trapping and scent detection using trained dogs (Beebe, Howell and Bennett, 2016; Browne, Stafford and Fordham, 2015; O'Connell, Nichols and Karanth, 2011; Phoebus et al., 2020). It is important to note that, even though these techniques do not require researchers to directly capture, handle, or observe the species, non-invasive techniques can still have a negative impact on the behaviour of the target animal. For example, tigers (*Panthera tigris*) have been observed avoiding remote camera stations as a result of camera trap flashes (Wegge, Pokheral and Jnawali, 2004) while other such instances of neophobia can occur in many species, like foxes (Culpeo fox (Pseudalopex culpaeus) and Grey fox (P. griseus)) (Travaini et al., 2013) and coyotes (Canis latrans) (Harris and Knowlton, 2011). Overall, however, non-invasive methods are all designed to minimise the impact of data collection on the target species and the ecosystems they inhabit. To add to that, these techniques are often described as more ethical, as the target species is not disturbed during data collection (Bekoff and Jamieson, 2019; Solberg et al., 2006).

Due to this focus on collecting data without requiring contact with the species itself, non-invasive techniques may lend themselves well to monitoring rare or elusive species (Solberg et al., 2006; Thompson, 2004). A branch of non-invasive monitoring has been proven particularly successful in detecting rare and elusive species has been non-invasive genetic sampling. Methods that belong to this branch of monitoring aim at remotely collecting genetic material from the target species. A review of the commonly used methods of non-invasive genetic monitoring in Long et al. (2012) outlines hair trapping and faecal sampling as the two main sources of genetic material for *Ursus arctos horribilis* (Grizzly Bears) and *U. americanus* (American Black Bears). *Ursus* species have long been in the forefront of these methods, from their developmental stages. In fact, the first non-invasive genetic monitoring efforts focused on brown bears (*Ursus arctos*) and their declining populations in the Pyrenees (Taberlet and Bouvet, 1992; Taberlet et al., 1997) and Northern Italy (Höss et al., 1992). The genetic material extracted from these samples (hair and scat) can confirm a species' presence in a habitat as well as draw information on lineages and genetically distinct populations or allow for individual identification. When it comes to the latter, nuclear DNA (nDNA) is used, paired with microsatellite markers that allow for these differentiations of DNA between individuals. However, there are a number of challenges associated with genetic

monitoring methods that rely on nDNA and the sample collection that can introduce bias to our understanding of the population structure and distribution of the target species. Literature related to non-invasive genetic monitoring has highlighted that bias introduced in the survey design (see Phoebus et al. (2020) for an example in biases introduced with brown bear scat sampling along forestry road networks) as well as climatic factors contributing to the degradation of nDNA (Farrell, Roman and Sunquist, 2000; Goossens et al., 2000; Paetkau, 2003) can significantly hinder monitoring efforts. Moreover, sex biases have been repeatedly recorded in both hair and scat collections (Bellemain et al., 2005; Boulanger et al., 2008; Phoebus et al., 2020; Pylidis et al., 2021).

When the challenges associated with these methods are controlled or accounted for, however, non-invasive genetic monitoring can be very effective in providing important insights into the population structure and distribution of a species. Since the introduction of non-invasive genetic sampling using hair samples from bears (Taberlet and Bouvet, 1992; Taberlet et al., 1997), this field has grown to encompass many alternative sources of DNA and collection methods. More recently, novel non-invasive genetic monitoring techniques targeting traces of DNA that wildlife leaves in its environment, labelled environmental DNA or eDNA, have been developed and used in species monitoring. Environmental DNA is increasingly used in both single-species and biodiversity monitoring because eDNA does not rely on collecting species-specific samples often underpinning the more conventional methods (Barnes and Turner, 2015; Ruppert, Kline and Rahman, 2019). Instead, by collecting eDNA samples in an area, any organism that has been in contact with the sampling material (i.e., water, soil, air) may be detected. This, in turn, gives the researcher the option to perform a targeted, single-species analysis or follow a metabarcoding approach to explore the diversity of species in the sample (Harper et al., 2018; Thomsen and Willerslev, 2015). The potential of eDNA techniques is being increasingly explored, in both species-specific methods and via next-generation sequencing techniques (Gold et al., 2021; Harper et al., 2018; Marshall, Vanderploeg and Chaganti, 2021; Thomsen and Willerslev, 2015; Wang et al., 2021) and special consideration is given to its ability to monitor rare or elusive species.

1.3 Invertebrate-derived DNA

Under the general umbrella of eDNA lies a method called iDNA, or invertebrate-derived DNA (Ruppert, Kline and Rahman, 2019; Schnell et al., 2015). Similar to eDNA techniques, iDNA seeks out external sources of target species' DNA, but in the case of iDNA the target species' genetic material is extracted from invertebrates. This technique has been widely used for monitoring invertebrates as vectors of pathogens such as Leishmania parasites (sand flies as vectors, e.g. Haouas et al., (2007)), malaria (mosquitoes, e.g. Scott et al., (2006)) and many other disease vectors (Kent, Norris and Feinstone, 2005). More recently, the technique was adapted in ecological monitoring and a number of invertebrate families, such as flies (Bohmann, Schnell and Gilbert, 2013; Calvignac-Spencer et al., 2013), mosquitoes (Kent, Norris and Feinstone, 2005; Townzen, Brower and Judd, 2008), ticks (Gariepy et al., 2012) and leeches (Schnell et al., 2015). Using invertebrates as a sampling unit means that the detection of the animal in question relies on being able to capture the focal invertebrate group instead of searching for a direct source of genetic material from the target species instead. Therefore, the method does not rely on the surveyor being able to detect the target species themselves, but instead indirectly monitors its presence through invertebrates that have come into contact with it. The effectiveness of the chosen invertebrate sampling method in the field, and the ability of the focal group of invertebrates themselves to feed on the target species or its scat, become important factors that influence the success of iDNA field sampling.

This approach of gathering information about a species by tracking traces of their DNA in their environment (eDNA and iDNA techniques) has introduced new avenues of monitoring rare or elusive species. However, this twice-removed (e.g. leeches, mosquitoes, horseflies) or three times-removed (e.g. flies, dung beetles) method of collecting genetic material means that the DNA within the iDNA samples has undergone considerable degradation (Lee, Sing and Wilson, 2015). For this reason, combined with the fact that each cell carries a single copy of nDNA, efforts following a microsatellite approach have shown negligible success rates (1%) (Schubert et al., 2015). Contrastingly, mitochondrial DNA (mtDNA) is found in multiple copies in the cells and is proven to be more resistant to degradation than nDNA while still carrying information distinct enough to determine identity between most vertebrate species (Delisle and Strobeck, 2002;

Johns and Avise, 1998; Townzen, Brower and Judd, 2008). Hence, the majority of eDNA and iDNA studies focus on mtDNA, a large proportion of which use universal primers, like universal mitochondrial 16S markers, following a metabarcoding approach (e.g. Calvignac-Spencer et al., 2013). While metabarcoding is invaluable for biodiversity monitoring, studies have shown that this approach can miss out on rare species within the samples. For example, Schubert et al. (2015), revealed a threefold increase in species-specific detection than when using a nonspecific PCR assay (metabarcoding). Real-time PCR, using species-specific primers and probes, can apply greater sensitivity to the sample analysis since the PCR analysis focuses on only finding and amplifying the target species' DNA (Melero et al., 2011; Schmittgen et al., 2000). Additionally, the costs and data processing time involved in a metabarcoding analysis in comparison to a real-time PCR (qPCR) approach are vastly higher, limiting the number of samples processed (Harper et al., 2018). Taking this into consideration, a single-species approach using iDNA samples could be more suitable at detecting rare or elusive species and even monitor species in areas they are known to use less frequently, such as stepping-stone and linkage (corridor) habitats.

1.4 Habitat Suitability Modelling

In order to conduct surveys that can reliably feed information into monitoring efforts, the planning stages of where these surveys should take place can be key (Anderson and Gonzalez, 2011). Ensuring that the effort, funds and time spent in the field are allocated effectively can drastically increase the amount and quality of data collected (Legg and Nagy, 2006; Long, 2008; Thompson, 2004). Habitat suitability analysis tools can provide the groundwork for focusing survey efforts to sites of interest by taking the species' use of the landscape into consideration. Habitat suitability models (HSM) calculate the likelihood of an area being a suitable habitat for the target species. Broadly, HSMs encompass many different spatial analysis approaches, but the common goal is to assess the probability of each cell within the study area in fulfilling the ecological requirements of the species in question (Anderson and Gonzalez, 2011). Information on these ecological requirements can be yielded from previous field monitoring of that area or neighbouring habitats, or by applying current understanding of the species' ecological requirements from literature and expert knowledge (Anderson and Gonzalez, 2011; Araújo and Guisan, 2006; Majka, Beier and Jenness, 2007; Segurado and Araújo, 2004).

The output from these models can serve to highlight the areas where surveying could take place depending on the aims of the project. This is especially useful for identifying areas that animals are expected to use less frequently but are essential for the long-term survival of a species, such as stepping-stones and corridors (Correa Ayram et al., 2016; Schaffer-Smith, Swenson and Boveda-Penalba, 2016). Similarly, HSM can help locate survey sites for populations where very little is known about, rare or elusive species, and areas that have never been previously monitored. In such cases where monitoring data alone is insufficient to guide the survey design, HSMs can extrapolate that data to generate predictions of suitability across the landscape. As funds for conservation are limited in many parts of the world, project planning that aims to reduce the overall cost and duration of monitoring without sacrificing data yield will be key.

1.5 Occupancy modelling

Furthermore, in the case of rare and elusive species or in monitoring animals with expected low use from the target species, detection in the field will always be reduced. It is thus important that the survey methods, once the most appropriate study area is selected, are sensitive to these reductions in the likelihood of detection. Occupancy modelling is often used to tackle this issue of imperfect detections, as this approach employs survey repeats to calculate not only the probability of an animal being on site, but also the probability of detecting it if it is present on site (Mackenzie et al., 2002; Mackenzie and Royle, 2005). Occupancy modelling works under the assumption that the detection probability will be less than 1 (perfect detection, where the animal is always recorded when present on site). Similar to HSM techniques, occupancy modelling can take into account the landscape and anthropogenic influences on the animal's use of the habitat, but it also considers survey-specific information (Comte and Grenouillet, 2013; Mackenzie et al., 2002; Mackenzie and Royle, 2005). This can include any record taken during sampling that is unique to that sampling occasion and could affect the detection of the target species, such as the weather, time, surveyor, and other factors, depending on the survey type (Strimas-Mackey et al., 2020). Since occupancy modelling is based on undertaking repeated survey - thus increasing likelihood of detection if the animal is present in a particular area - it has proven especially useful when trying to understand the distribution of rare and elusive species (Perkins-Taylor and Frey, 2020; Peterman, Crawford and Kuhns, 2013; Thompson, 2004).

This thesis uses a case study to showcase the use of these three techniques: habitat suitability modelling; non-invasive genetic monitoring using iDNA; and occupancy modelling, as tools in conservation planning. The aim underpinning this project is to demonstrate the use of tools that can reduce the cost, time and effort in the field, without sacrificing on data yield and quality. Furthermore, non-invasive monitoring using aquatic eDNA paired with occupancy modelling has been used extensively (Dorazio and Erickson, 2018; da Silva Neto et al., 2020; Strickland and Roberts, 2019), but, to my knowledge, this is the first case study to generate an occupancy model using terrestrial iDNA data, although the need to explore it using case studies has been previously highlighted (Gogarten et al., 2020; Schnell et al., 2015). I present iDNA as a tool for single species monitoring and show its strengths and disadvantages as a stand-alone survey technique compared to a complementary monitoring approach.

1.6 The case study

To successfully illustrate the methods outlined above, a few basic requirements were taken into consideration in selecting the study species and location. Firstly, to illustrate the benefits of using HSMs to select a suitable study area, the study species needed to be a wide-ranging animal the distribution of which is well-studied. This would allow for a comparison of the models generated with our already existing knowledge of the species' range, to ensure that the models act as a suitable guide for the field. Secondly, to assess the effectiveness of iDNA as a single species monitoring tool, it was essential to select the target species for which other, comparable survey methods are present and used frequently to detect its presence. Additionally, to be able to assess the sensitivity of iDNA in detecting rare occurrences, the target species needed to be rare or elusive, or, alternatively, monitored across a habitat it is not expected to use frequently.

Brown bears (*Ursus arctos*) fulfil all these requirements as their presence and ecological preferences have been well-monitored across much of their extensive range. Specifically, this case study focuses on the Eurasian brown bear (*U. arctos arctos*) and its distribution in Greece. Bears, as most large carnivores, are wide-ranging animals largely

dependent on large territories, which makes them especially vulnerable to habitat loss and fragmentation. In the nineteenth to early mid-twentieth century, habitat loss paired with hunting in Europe caused large declines and local extinctions in all four large European carnivore species: Eurasian brown bear, grey wolf (*Canis lupus lupus*), Eurasian lynx (*Lynx lynx*) and wolverine (*Gulo gulo*)(Boitani and Linnell, 2015; Chapron et al., 2014). Due to large pan-European conservation initiatives, changes in legislation (inc. Bern Convention of 1982, Habitats Directive of 1992) and local conservation efforts, the majority of large carnivore populations are now recovering, with some showing habitat expansion and even recolonisation of previous ranges (Boitani and Linnell, 2015; Chapron et al., 2014; Kaczensky et al., 2021, 2012a). This range expansion, realised despite ongoing habitat fragmentation as a result of human development, poses a new challenge in the conservation of large carnivores. Long-term conservation strategies are focusing on protecting large expanses of core habitats as well as improving stepping stones and corridors to prevent population bottlenecks (Bennet, 1999; Hendry et al., 2003). This focus on protecting habitat linkage zones has been in the forefront of bear conservation strategies in Europe, in an effort to tackle the effects of habitat fragmentation and preserve genetic diversity (Karamanlidis et al., 2012; Mateo Sánchez, Cushman and Saura, 2014; Mateo-Sánchez et al., 2015b, 2015a; Pylidis et al., 2021).

Bears, amongst other large carnivores, are habitat generalists and often labelled as 'umbrella species' (Dai et al., 2021; Mateo Sánchez, Cushman and Saura, 2014; Wang et al., 2018). Having such broad habitat requirements and large territories, they are often the first animals to be affected by fragmentation, but it also means that by focusing conservation efforts on the habitats they inhabit a large diversity of other species are also protected (Crespo-Gascón et al., 2019; Linnell, Swenson and Andersen, 2000). Additionally, brown bears display a well-studied tolerance to human disturbance, making them more likely to use patches of land between suitable habitats and even become habituated (Mattson, 1990; Mertzanis et al., 2005). Furthermore, the range and ecological requirements of bears, and especially brown bears, have been studied extensively across their global distribution (Brooke et al., 2014), including research focusing specifically on *U. arctos arctos*, providing a good foundation for the HSM analysis. To add to that, a recent study of the distribution of brown bear across Europe was conducted by collating data from each county's monitoring efforts, allowing for a

baseline of distribution that nationally-focused HSMs could be compared against (Kaczensky et al., 2021). Finally, bears are monitored using a variety of methods, from invasive techniques such as telemetry, to non-invasive methods such as trail cameras, hair trapping and scat surveys. Here, scat surveys were chosen as the most comparable method to iDNA monitoring since they can both be conducted during the same sampling session.

1.7 Study area

Since single species monitoring using iDNA is a novel method, brown bears are ideal as the case study species because detection in areas of high population density was highly probable using standard methods and thus expected to yield results with iDNA as well. However, another aim of this project was to test the use of iDNA in monitoring rare or elusive species. Even though brown bears are not often put in this category of animals, rare detection can still occur when monitoring areas outside core populations. Corridors and stepping stone habitats are used more infrequently in search of food and mates, creating a mosaic of occupancy across the landscape. This study focused on the brown bear range in Greece and the first part of the project, the HSM analysis, was conducted to look for these areas that displayed a variety of habitat suitabilities and suggested a gradient of habitat use.

U. arctos is the most abundant large carnivore in Europe, present in 22 countries and numbering up to 17,000 individuals across ten main populations: Scandinavian, Karelian, Baltic, Carpathian, Dinaric-Pindos, Eastern Balkan, Alpine, Central Apennine, Cantabrian, and Pyrenean (Boitani and Linnell, 2015; Kaczensky et al., 2021, 2012a). Greece hosts the southernmost range of brown bears in Europe, with two genetically distinct populations, the western and eastern population (Mertzanis et al., 2008; Pylidis et al., 2021). Both populations are thought to be growing and brown bears in Greece have seen a population increase over the last 50 years, recently estimated between 350-400 (Karamanlidis et al., 2015) and 500 individuals (Pylidis et al., 2021) across the country. The western population forms the southernmost portion of the Dinaric-Pindos population, stretching from Slovenia to Greece. The Eastern population belongs to the East Balkan population and is mostly confined within the Rhodope Mountain range and surrounding habitats (Kaczensky et al., 2012b; Pylidis et al., 2021). The two populations

were thought to be geographically fragmented, with the closest stretch between them in North Macedonia, but a recent study from Pylidis et al. (2021) showed evidence of individuals with admixed ancestry, suggesting a small genetic linkage between the two core populations. A connectivity pilot study conducted by Savvantoglou (2016) highlighted potential linkage zones between the two core areas as well as future sites of natural recolonisation.

As bear populations are recovering in Greece, it is important to monitor those corridors and understand the movement of bears across the landscape. Habitat suitability analyses allow for a closer look at the landscape's capacity to host viable bear populations and highlight potential areas of connectivity and recolonisation as the population continues to increase (Majka et al., 2007). With regard to monitoring bears using iDNA, the suitability models can be used to highlight sections of their distribution in Greece that offer instances across the gradient of habitat use, from core areas of high bear density to corridors and recolonised sites with more rare occurrences.

1.8 Study aims

The aims of this thesis therefore are to provide a robust assessment of iDNA as a tool for elucidating the distribution and habitat preferences of large mammals, using the *U. arctos arctos* as a case study. In the next chapter I use and evaluate multiple habitat suitability modelling techniques to define a suitable test landscape for evaluating iDNA, these provide a baseline against which I assess the efficacy of iDNA based distribution modelling. In chapter 3 I describe the development of the iDNA assay and chapter 4 illustrates how iDNA can perform to provide the data which underlies spatially-specific occupancy models for bear populations. Finally, in chapter 5, I discuss the potential and limitations of iDNA as a tool for use in conservation projects both those concerned with bears and more broadly.

Chapter 2: Look for the (Greek) bear necessities: habitat suitability assessments for brown bears (*Ursus arctos*) across their range in Greece

Introduction

Conservation in fragmented habitats: fighting the odds

In recent years, large European carnivore populations have been stable and even increasing populations in the majority of their range (Chapron et al., 2014; Kaczensky et al., 2012a). European populations of Eurasian brown bear (*Ursus arctos arctos*), grey wolf (*Canis lupus lupus*), Eurasian lynx (*Lynx lynx*) and wolverine (*Gulo gulo*) are increasing in numbers and expanding their range, colonising habitats from which they have been absent for decades (Boitani and Linnell, 2015; Chapron et al., 2014; Kaczensky et al., 2012a). Nevertheless, carnivore population growth often results in range expansion of these animals into areas more densely populated by humans. Conservation efforts focus on promoting human-wildlife coexistence, ensuring a peaceful integration of these animals as they expand their range near settlements and farmlands. Most European landscapes are a complex mosaic of suitable habitats and inhabited areas (including urban and rural settlement areas, farming and complex road networks), making it a conservation priority to study the relationship species have with their landscape (Psaralexi et al., 2022; Langen et al., 2017). Habitat Suitability Models (HSM) provide insights into the relationship a species has with its surroundings and predict areas the species is most likely to use. The evaluation of species' habitats is essential for conservation and this necessity is reflected in the increasingly high presence of GIS analyses in species reports and action plans (e.g. Kaczensky et al., 2012b; Mateo Sánchez, Cushman and Saura, 2014; Mertzanis, Psaroudas and Karamanlidis, 2020; Posillico et al., 2004). This chapter assesses European brown bear (*Ursus arctos arctos*) habitats in Greece by comparing two different approaches to suitability modelling to obtain a more comprehensive picture of the suitability mosaic across their range and potential areas of future recolonisation.

Habitat Suitability Modelling

Advances in remote sensing software and the development of open-source spatial analysis tools has meant that habitat suitability modelling and corridor studies have become a common tool in species conservation (Araújo et al., 2019; Elith and Graham, 2009). Software and specialised packages specifically created for modelling species distribution and habitat suitability (e.g. MaxEnt, ArcGIS Pro, CorridorDesigner, ENMeval in R) and wildlife corridors (e.g. The Circuitscape Project, CorridorDesigner, Life Mapper, Conefore, Unicor, etc; reviewed by Correa Ayram *et al.*, 2015) are becoming increasingly accessible, open source, and, in some cases, require minimal training. In this study I used a combination of R packages, a maximum entropy method (e.g. MaxEnt; (Phillips et al., 2006)), and ArcGIS software to create a collection of suitability models and highlight the importance of different environmental variables in predicting suitable habitat and potential corridors for brown bears in Greece. In this chapter I identify the ecological requirements of brown bears as highlighted in the literature that are relevant to the study area and test their significance in predicting brown bear habitat suitability in Greece.

Habitat suitability modelling software makes predictions by assessing the relationship between the input data and different environmental variables. These modelling tools can be broadly categorised into two types according to the input data they require. The first type uses occurrence data to reveal a species' ecological requirements and then inspects the landscape to find areas where those characteristics are present. The second type relies instead on the prior knowledge the species' ecological requirements, drawing information from literature and expert knowledge to draw a habitat scoring system. For the purposes of this study, I will hereafter use the terms ***Species Presence*** (SP) and ***Habitat Suitability Index*** (HSI) to describe type one and two respectively.

Species Presence models:

Species Presence (SP) model require an occurrence data set (presence-only) of the target species collected via sign surveys, GPS locations of telemetry collars, sightings, camera trap observations, etc. These data are combined with a set of environmental layers that are assumed to affect the species distribution. Large-scale habitat suitability modelling often

relies on environmental variables such as elevation, topography, climate and land cover or soil types, as well as anthropogenic variables, especially when it comes to small-scale and/or finer resolution modelling (Bellamy, Scott and Altringham, 2013; Chauvier et al., 2021; Condro et al., 2021; van Gils et al., 2014, 2012; Lennon, Greenwood and Turner, 2000; Rödder et al., 2021) . These environmental layers are tested for correlations to avoid over-prediction biases (Brown, Bennett and French, 2017; Elith and Graham, 2009). Once the best candidate layers are identified, they are tested against the occurrence data and a habitat suitability model is created, informed by the location the species was present in. The process is straight-forward and, given an occurrence dataset that is representative of the species' behavioural ecology, the resulting model can provide a very good understanding the study area's suitability for the target species.

Species Presence models are based on the principle that the probability of a species occurring in a cell is a function of the abiotic environmental variables a in that cell (Anderson et al., 2011; Elith et al., 2006). In other words, an SP model identifies cells in the study area that have similar characteristics to localities where the species has been observed. A variety of techniques are available for Species Distribution Model (SDM) studies, reviewed in detailed in Elith and Graham (2009) and Elith (2006). This study utilised one of the most prevalent species distribution modelling tools, MaxEnt, an ecological niche modelling software developed by Phillips, Anderson and Schapire, (2006), which has been shown to outperform other ecological niche modelling tools (Hijmans and Graham, 2006; Su, Bista and Li, 2021) with evidence of it being suitable even given a small set of occurrence records (Dudik and Phillips, 2008).

A benefit of using the MaxEnt algorithm in species distribution studies is that MaxEnt can introduce transparency to the selection of the most valuable predictors for species distribution, thus allowing to removal of less important variables from the analysis and creation of more parsimonious models, shown to be more resistant to over-fitting (Anderson and Gonzalez, 2011). In fact, given an *a priori* selection of initial variables thought to predict the species' distribution and a subsequent test of collinearity between those predictors allows for a small set of non-collinear predictors to be used for each SDM. Additionally, further removal of predictors guided by MaxEnt results can advise a hierarchical procedure for

making *a posteriori* decisions on removing predictors from the analysis based on their contribution to the test's gain (Elith and Graham, 2009; van Gils et al., 2014). The SP modelling work outlined in this chapter followed a similar step-wise assessment of the predictor variables focusing on the annual distribution of brown bears in Greece. The distribution models were focused on three broad categories of variables: topographic, land cover and human impact, and highlighted the importance of each of these categories in describing bear distribution in Greece.

Habitat Suitability Index models

One of the more challenging aspects of creating distribution models using occurrence data is that presence records can often draw a less representative picture of a species' habitat preferences. Although studies have shown that Species Presence models can perform a reliable analysis with a small amount of occurrence data (Elith et al., 2006; Hernandez et al., 2006, 2008), the data itself can be highly biased. Elusive or rare species, or even the lack of funds, can potentially prevent the collection of sufficient or unbiased data for the generation of robust models (Thompson, 2004). In these cases, a different approach, based on expert knowledge may be more effective, whereby the species presence data used in the previous method is replaced by our current understanding of the species' behavioural ecology. Habitat models are thus derived by analysing the knowledge of specialists of the likelihood of a specific set of circumstances serving as suitable habitat. The collection of expert knowledge involves rigorous literature reviews and the advice of experts who have observed the species in the field. The understanding of a species' habitat preferences is collated to create a suitability index whereby variables are evaluated and ranked according to their suitability for the target species.

HSI models essentially replace the information drawn from the occurrence data points collected in the field with prior knowledge about the species' ecological behaviour. The analysis is carried out by manually inputting these ecological preferences into the environmental variable analysis and creating a suitability scale for each of the variables. The final model combines those variables into one suitability model that depicts what expert opinion and previous literature has found to be a suitable condition for the target species.

Method comparisons

HSMs can be presented as a binary prediction landscape, showing all values above a certain suitability threshold as 'suitable' and values below that threshold as 'unsuitable', or a gradient suitability landscape, showing a range of different suitabilities across the study area (Majka, Beier and Jenness, 2007). Both types of HSM maps are used in this chapter, as the binary ones are more suitable in model comparisons, while the gradient maps reveal information on less suitable areas that may still act as corridors or stepping-stones (Majka, Beier and Jenness, 2007; Walker and Craighead, 1997). Binary models have been more common in literature for *U. arctos* in Europe (Chapron et al., 2014; Kaczensky et al., 2021, 2012b) as they are a very clear way to depict suitable sites, flag conservation sensitive areas and prioritise habitat management. However, binary models fail to show the dynamic change in habitat use by the species across the gradient of suitability. While it is important to understand the location of core habitats that are clearly revealed in binary models, the more nuanced understanding of habitat use gets lost. HSMs showing the gradient suitability can reveal a lot more information about the use of the entire landscape by the animal, ranging from completely avoiding an area to successfully breeding in it (Majka, Beier and Jenness, 2007). Furthermore, the maps essentially show the permeability of the landscape, allowing for the creation of corridor predictions and potentially, showing how the animals can move between core areas. Species action plans looking at the long-term conservation of species in fragmented landscapes are often focused on preventing genetic bottlenecks and creating or maintaining linkages between core habitats (Clevenger, 2012; Loro et al., 2015; Walker and Craighead, 1997), highlighting the necessity for habitat suitability modelling efforts.

Analysis resolution

A common practice in habitat suitability modelling is to resample the environmental variable dataset to the largest cell size present (e.g. Gastón et al., 2017; van Gils et al., 2014). However, corridor modelling and management regimes could benefit from the preservation of the more fine-scale variables, as literature suggests that cell size can influence the relationships between variables and affect the resulting models (van Gils et al., 2012; Gotelli, 2003). Studies looking into the response of bears to climatic variables (e.g. Bojarska and Selva, 2012; van Gils et al., 2014) reveal strong correlations with elevation and vegetation cover layers, thus

making it difficult to decipher the direct impact of climatic variables to bear distribution at a finer scale. For that reason, climatic layers were excluded from the models described in this chapter.

Chapter aims

This chapter follows basic principles for collecting and preparing appropriate datasets for habitat suitability modelling to generate suitability predictions for the European brown bear's range in Greece, following guidance from similar studies in the sub-Mediterranean climate. The two above-mentioned approaches are used to generate a series of HSMs and compare their outputs. The HSMs focus on creating parsimonious models using variables available online so that the methodology presented here can be easily transferable to other species. Using bears as a case study, this chapter presents a workflow that focuses on:

1. Creating suitability predictions using two different HSM approaches in both a binary (suitable/unsuitable) and a suitability gradient format.
2. Comparing the resulting models to understand areas of agreement and divergence.
3. Comparing the resulting models with the current known distribution of the target species as presented in Kaczensky et al. (2021).
4. Using gradient maps to provide a more detailed understanding the distribution of the species within the study area.

Ultimately, the work outlined in this chapter presents a series of habitat suitability assessments to create a more comprehensive image of predicted bear distribution and habitat use in Greece as well as to identify suitable habitats across the Greek mainland for brown bears in their current range and in areas of future colonisation. However, the focus was on creating a methodology that can be easily transferable to a species where less is known about, whether there is a gap in literature on the specie's ecology, or a lack of records data. The SP modelling methodology is focused on using methods and selection processes that limit the *a priori* decisions based on understanding the animal's ecology, while the HSI analysis presents a detailed example of utilising expert knowledge to gather information on habitat suitability.

Methods

Study area

The habitat suitability analysis in this study covered the Greek brown bear range as described in Kaczensky et al. (2021) (Figure 2). The study area also includes small parts of neighbouring Balkan countries (Albania, North Macedonia, and Bulgaria) in the north. The study area totalled 328,340 km^2.

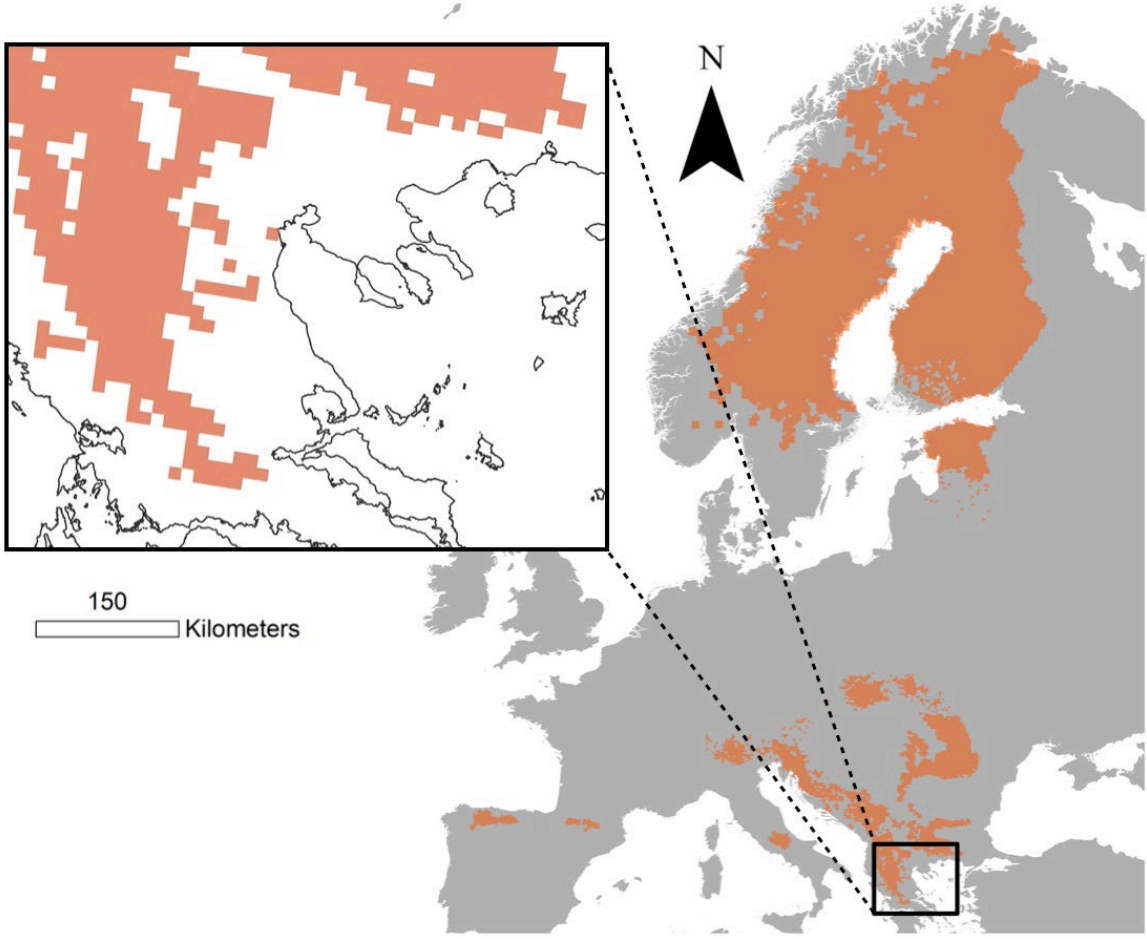


Figure 2. Distribution of bears the period 2012-2016 as described by Kaczensky et al, (2021) for the IUCN assessment of population figures for the European brown bear. Dataset available under 'COO 1.0 Universal (CCO 1.0) Public Domain Dedication' licence.

Modelling Variables

The habitat suitability models were created using the same variables to allow for a direct comparison between the two resulting models. To maintain our efforts to use this method as a preliminary, no-cost evaluation of the habitat, the candidate variables used were either freely available from an online source or created using a pre-existing layer and further spatial analysis. A list of the variables and their source, type and resolution is provided Table 1.

Table 1. Environmental layers used for the HSMs, categorised by their broader groups and listing their source, data type and raster resolution.

Predictor group	Predictor	Data type	Source	Resolution (metres)
Topography	**Elevation**	Continuous	EU DEM, Copernicus, European Environmental Agency (EEA)	30 x 30
	Topography	Categorical	Derived from elevation raster using the Corridor Designer toolbox on ArcMap 10.7	30 x 30
Land Cover	**Land cover**	Categorical	Corine Land Cover (CLC) 2018 - Copernicus, European Environmental Agency (EEA)	100 x 100
Human Impact	**Distance from urban areas**	Continuous	Euclidean distance from CLC 2018 classes 1.1.1 - 1.2.1	30 x 30
	Distance from roads	Continuous	Euclidean distance form major roads derived from OpenStreet Map (data licensed under the Open Data Commons Open Database Licence (ODdL)	30 x 30

Topography

Topography variables were comprised of two datasets: elevation and topographic position. The elevation raster (EU-DEM v1.1, tiles E50N20 and E50N10, European Environment Agency) was used to create a variable reflecting bear preferences of the topographic composition of the terrain. The Topographic Position raster was created using the Corridor Designer toolbox for ArcMap10.7 to split the landscape into four categories: *ridgetops*, *flat/gentle slopes*, *steep* slopes and *canyon bottoms* (Majka, Beier and Jenness, 2007).

Land cover

Corine Land Cover (CLC; EEA, 2018) is a European-wide dataset mapping the extend of 44 land cover classes. It was chosen for this study to allow for a uniform cover of the study area, taking into account that parts of it fall in neighbouring European countries.

Human Impact

The *'Distance from urban areas'* raster was created by measuring the Euclidian Distance (ArcMap 10.7) of the Urban Fabric category (CLC 1.1.1-1.1.2) and Industrial, Commercial and Transport Units (CLC 1.2.1-1.2.4) of the Corine Land Cover 2018 dataset. These urban classes were selected to represent an urban raster with the most important anthropogenic disturbances in the land cover mosaic. Other artificial sites represented by the categories 'Mine, dump and construction sites' (CLC 1.3) and 'Artificial, non-agricultural vegetated areas' (CLC 1.4) were not included in the urban area mosaic due to their less consistent (for example, ski resorts which are often only active in Greece between December and March that coincides with the bear hibernation period in the south of Europe (Swenson et al., 2007)) use or reported use by bears (see (Elfström et al., 2014a) for example of bear use of dump sites).

Road network datasets for the four countries within the study area were derived from OpenStreetMap (data licensed under the Open Data Commons Open Database Licence (ODdL)) and the major roads (motorway, primary, trunk and associated links) were used to a create '*distance from major roads'* raster (Euclidean distance tool, ArcMap 10.7).

Raster resolution

All variables were resampled to the same cell size, creating a raster dataset *min* which was resampled to match the highest resolution on the set of variables (elevation raster, cell size: 30 x 30 m).

Managing correlated environmental variables

In order to maintain an easily transferable SDM methodology, using bears as a case study, the decisions on what variables to include when correlations were present were not made on an '*a priori'* selection based on the species' habitat requirements. Instead, to account for the method transferability to a much less studied species, the variable selection was advised by a collinearity analysis by looking at the Variance Inflation Factor (VIF). The analysis was ran on

the R 'usdm' tool (Naimi et al., 2014), using a VIF threshold of 3 as described by Bellamy, Scott and Altringham, (2013). This analysis goes through a stepwise variable removal by looking at the linear correlations between pairs of variables above the VIF threshold and excludes the most highly collinear variables (Naimi et al., 2014).

Species Presence model

Bear presence data

Telemetry GPS tracking data were provided by the Environmental Organization for Wildlife and Nature, Callisto. The dataset comprised of 77,211 GPS locations collected using GPS/GSM (FOLLOWIT) radio-collars that were fitted on a total of 27 bears between 2003 and 2013.

A number of steps were taken to reduce any bias associated with the data. Firstly, in the two datasets where the quality of the GPS signal was recorded (2D or 3D), only the 3D points were kept, following findings by Lewis and colleagues (2007) which demonstrated that removing all 2D data from their GPS collar dataset resulted in the most accurate dataset. Secondly, all remaining data (77,211 GPS points) was merged into one shapefile, 169 location duplicates were removed (77,042 points remaining) and '*thinned*' using the 'Spatial Rarefy Occurrence Data for SDMs' tool in the SDM Toolbox 2.4(Brown, Bennett and French, 2017). To account for the complexity of the terrain, the GPS data was thinned using a topographic heterogeneity raster. This raster was generated from the elevation raster (European Digital Elevation Model (EU-DEM), version 1.1, European Environment Agency (EEA)) using the tool 'Calculate Topographic Heterogeneity' which illustrates the heterogeneity gradient of the landscape by comparing differences in the values of neighbouring cells (Figure 3). The rarefying tool on SDM toolbox calculates a distance matrix and uses a non-random process to systemically remove the closest cluster points first, then re-evaluates that distance matrix until all points are removed at the specified search distance. When a heterogeneity raster is present, the tool uses the heterogeneity gradient to split the rarefying distances into a chosen number of groups (here, three; 90m, 180m and 270m) that reflect the complexity of the terrain. A smaller rarefying distance is used in areas where the complexity is high, resulting in more occurrence points being retained in those more heterogenous areas. Contrastingly, where the

terrain is more homogenous, a larger rarefying distance is used, discarding a larger amount of occurrence points in those areas. The thinning distance of 90 m was chosen by averaging the minimum estimated home range for females and minimum estimated home range for male brown bears (Kaczensky et al., 2003; Kanellopoulos et al., 2006; Mertzanis et al., 2005).

The species presence data thinning process using the topographic heterogeneity raster retained 4,459 occurrence points, minimising the likelihood of carrying over bias associated with autocorrelation and over-thinning data in more heterogeneous regions of the landscape.

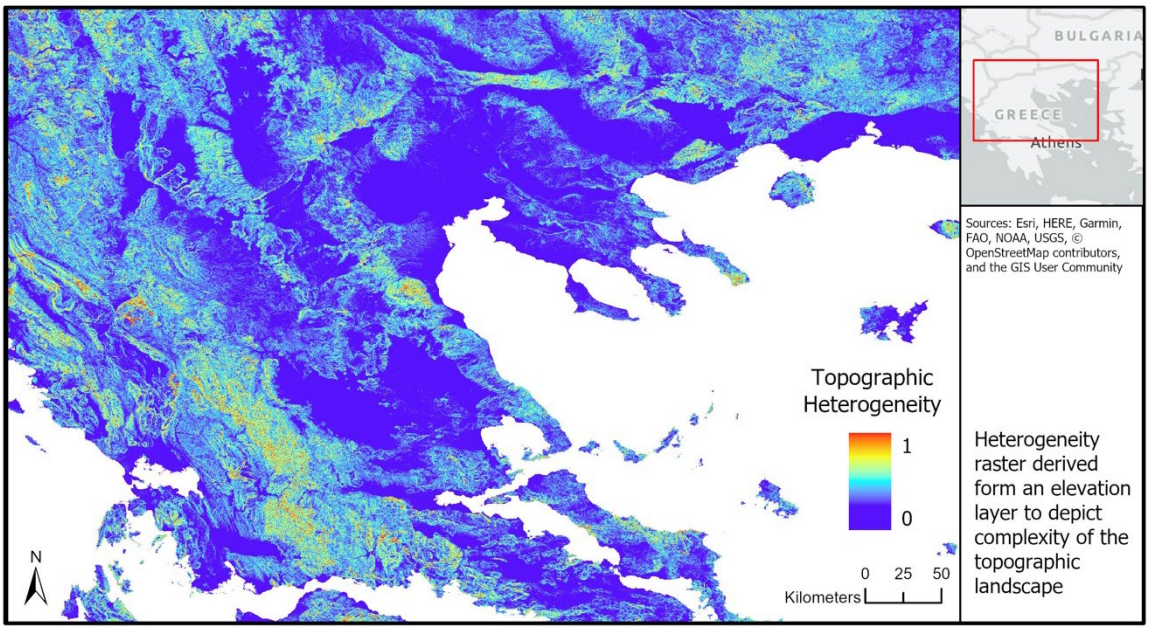


Figure 3. Topographic heterogeneity raster across the study area, used to spatially rarefy the occurrence data. Areas of high heterogeneity retained more occurrence points, while larger thinning distances were used for the more homogenous areas, resulting in further reduction of occurrences where the landscape is less complex.

Sampling Bias files for SP models:

To minimise bias related to the background data (pseudoabsence points) selection, a bias model for each model resolution (max and min) was generated using the Kernel Density analysis in the R package 'GISTools' (Pérez-Goya U et al., 2020). The function uses the occurrence data to create a two-dimensional kernel density raster and is used to constrict the MaxEnt data analysis to areas within a given radius (here, 30km) of the presence points.

The final analysis for each SP MaxEnt model was carried out in R and the parameters used for each model were selected using the package ENMeval (Muscarella et al., 2014). This runs the MaxEnt analysis across the entire range of parameter options within the algorithm. The analysis compares the Aikake Information Criterion (AIC) values to point out to the model with the lowest AIC value (Warren and Seifert, 2011), revealing the parameters that will create the model with the best fit and predictive ability. Once the most suitable parameters were identified for each model, the MaxEnt analysis was carried out using maxent.jar (Phillips, Anderson and Schapire, 2006; Phillips and Dudík, 2008) in R Studio using the 'Dismo' package (Robert J. Hijmans et al., 2020). Each test included a Jackknife analysis of test gain for each variable, to highlight any variables in the model the removal of which would result in a larger test gain. If such variables were identified, they were removed from the analysis to arrive at the most parsimonious model. An overview of the MaxEnt analysis workflow is shown on Figure 4.

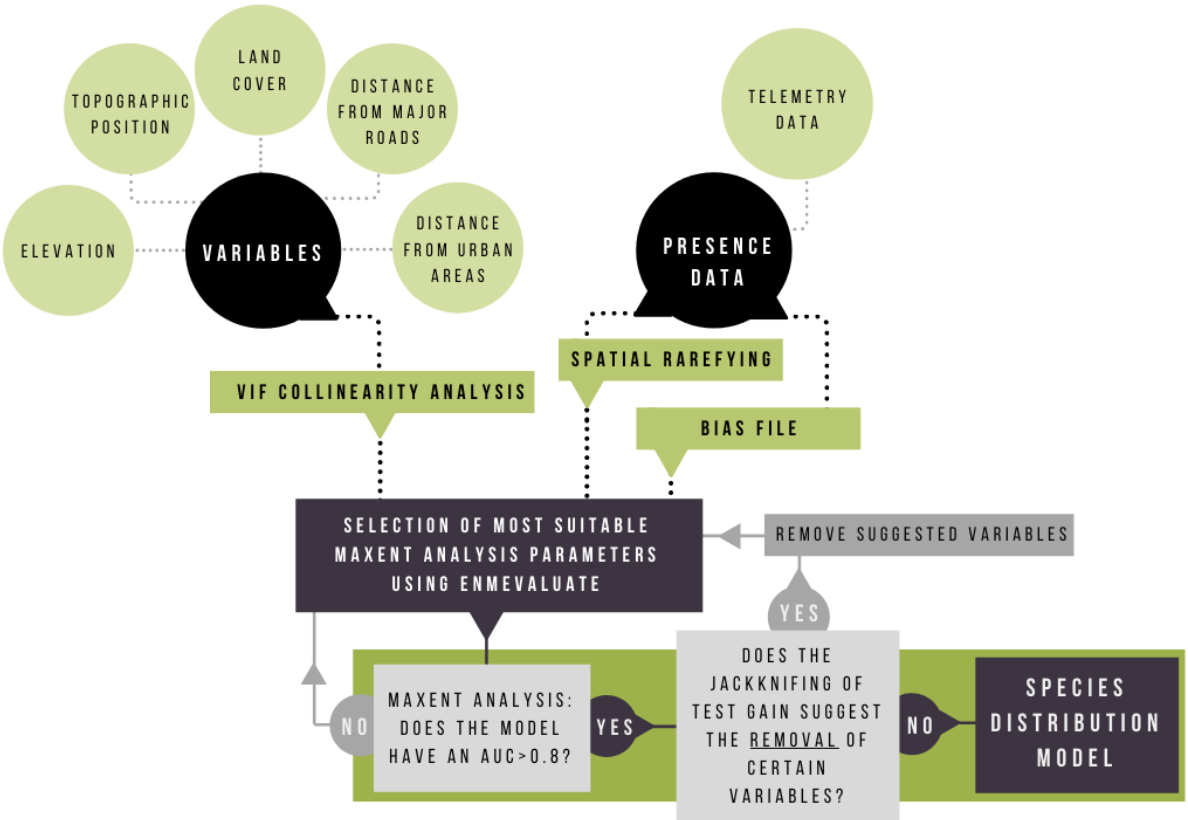


Figure 4. MaxEnt modelling workflow describing the process of selecting the most suitable MaxEnt parameter settings and eliminating variables that do not contribute, or contribute negatively to the model, to finally arrive to the most parsimonious species distribution model.

Habitat Suitability Index models

The HSI models were generated based on the guidelines of the GIS tool CorridorDesigner, developed for the creation of habitat suitability assessments and corridor models (Majka *et al.* 2007) and operated within ArcGIS v.10.4.1 (ESRI, 2014) and performed using the Suitability Modeller in ArcGIS Pro v2.7 (Esri 2021).

This process relies on an *a priori* index of a habitat suitability scores for each of the variables (see (Majka, Beier and Jenness, 2007). These scores ranged from 0, representing unsuitable habitat, to 100 being the highest suitability score (Figure 5), and were determined using the literature review from Savvantoglou et al. (2017) updated with the addition of recent literature.

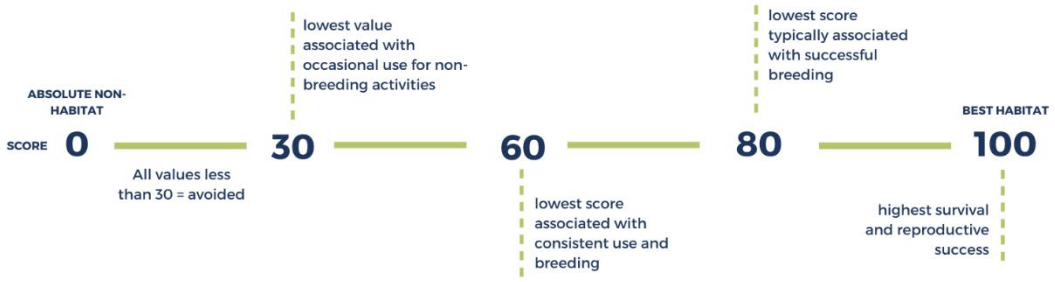


Figure 5. Biological interpretation of habitat suitability scores as defined by the CorridorDesigner team (Majka, Jenness and Beier, 2007).

The variables and justification of suitability scores is detailed in Table 2 and 3. For categorical variables (Topographic Position Index and Corine Land Cover) this was done by allocating a separate score to each variable category. Continuous variables were reclassified and converted to categorical variables to match zones within that variable's range that have been identified by literature as important thresholds for bear use.

Table 2. Habitat suitability values given to the CORINE layers and literature used to make the scoring decisions (Almpanidou et al., 2014; Bartoń et al., 2019; Brody and Pelton, 1989; Chapron et al., 2014; van Gils et al., 2014; Güthlin et al., 2011; Kanellopoulos et al., 2006; Kusak and Huber, 1998; Mertzanis, 1994; Mertzanis et al., 2006, 2008, 2005; Savvantoglou et al., 2017; Whiteman et al., 2017).

CLC category	CLC code	CORINE Layer	Suitability score	Justification
1. Artificial surfaces	1.1-1.4.2	Urban areas (codes 1.1 to 1.4.2)	0	Areas of high human activity
2. Agricultural areas	2.1.1	Non-irrigated arable land	50	Occasionally used for feeding, but not suitable for breeding
	2.1.2	Permanently irrigated land	30	Occasional feeding areas, but not suitable for breeding.
	2.1.3	Rice fields	10	Not suitable
	2.2.1	Vineyards	30	Potentially occasional feeding areas, but not suitable for breeding.
	2.2.2	Fruit trees and berry plantations	30	Very good source of food, but plantations associated with intensive agriculture. Occasional visits. Not suitable for breeding
	2.2.3	Olive groves	0	Unsuitable food source. Not suitable for breeding
	2.3.1	Pastures	50	Occasionally used for feeding, but not suitable for breeding
	2.4.1	Annual crops associated with permanent crops	30	Occasionally used for feeding, but not suitable for breeding
	2.4.2	Complex cultivation	30	Occasionally used for feeding, but not suitable for breeding
	2.4.3	Land principally occupied by agriculture, with significant areas of natural vegetation	60	Frequent use for feeding, possible breeding potential
3. Forest and seminatural areas	3.1.1	Broad-leaved forest	100	Consistent use for feeding and breeding
	3.1.2	Coniferous forest	80	Consistent use for feeding and breeding
	3.1.3	Mixed forest	90	Consistent use for feeding and breeding
	3.2.1	Natural grassland	60	Frequent use for feeding, possible breeding potential
	3.2.2	Moors and heathland	60	Frequent use for feeding, possible breeding potential
	3.2.3	Sclerophyllous vegetation	60	Frequent use for feeding, possible breeding potential
	3.2.4	Transitional woodland shrub	50	Occasionally used for feeding, but not suitable for breeding
	3.3.1	Beaches, dunes, and sand plains	0	Not suitable
	3.3.2	Bare rock	80	Consistent use for feeding and breeding
	3.3.3	Sparsely vegetated areas	20	Bears might cross in search for good habitat types
	3.3.4	Burnt areas	0	Not suitable
4. Wetlands	4.1.1	Inland marshes	20	Bears might cross in search for good habitat types

5. Water bodies	5.1.1	Water courses	0	To prevent the tool from drawing suitable habitat on river surfaces
	5.1.2	Water bodies	0	As 5.1.1
	Other	Other land cover types from CLC	0	Missing CLC habitat types were not present in the study area

Table 3. Habitat suitability category scores of elevation zones, topography type and distance form major roads and urban areas and literature used to make the scoring decisions (Bartoń et al., 2019; Brody and Pelton, 1989; Elfström, Swenson and Ball, 2008; van Gils et al., 2014; Kanellopoulos et al., 2006; Kusak and Huber, 1998; Lewis et al., 2011; Mattson, Knight and Blanchard, 1987; Mertzanis et al., 2006, 2011, 2005; Roever, Boyce and Stenhouse, 2008; Savvantoglou, 2015; Whiteman et al., 2017).

Layer	**Classifications**	**Suitability score**	**Justification**
Elevation (m)	0 - 400	20	Rarely used in search of food
	400 - 800	60	Frequent use for feeding, possible breeding potential
	800 - 1700	100	Ideal altitudinal range
	1700-2889	50	Tree line ends and alpine vegetation begins. Occasionally used for feeding
Topographic Position (slope)	Canyon bottom	40	Based on observations due to lack of data on this type of landscape structure in Greece
	Flat-gentle slope	80	Less energetically costly. Consistent use for feeding and breeding
	Steep slope	50	Use for commuting between places. Narrow field of vision and energetically costly. Occasionally used for feeding, not suitable for breeding.
	Ridge top	90	Best field of vision. Path of least resistance. Consistent use for feeding and breeding.
Distance from major roads (m)	0 - 100	30	Mostly avoided, unless crossing.
	100 - 500	60	Frequent use for feeding, potential breeding
	500 and above	100	Ideal minimum distance
Distance from urban areas (m)	0 - 500	50	Infrequent feeding. Not suitable for breeding.
	500 - 1500	60	Frequent use for feeding, potential breeding
	1500 and above	100	Ideal minimum distance

The variables and allocated suitability scores were used to generate HSIs using a Weighted Suitability analysis in ArcGIS Pro. This analysis accounts for the different level of significance the input variables have on the species' suitability by allocating weights to each of the variables. Two different weight models were selected in this step: a weight allocation based on expert knowledge, referred to here as 'HSI_EK' and a weight allocation informed by the species presence analysis, referred to as 'HSI_PI' and given by the Permutation Importance of its corresponding MaxEnt model of the 'min' resolution models. Figure 6 illustrates the basic methodology workflow of the EK model. Six different models were generated using the HSI approach to reflect their corresponding SP models (Figure 7).

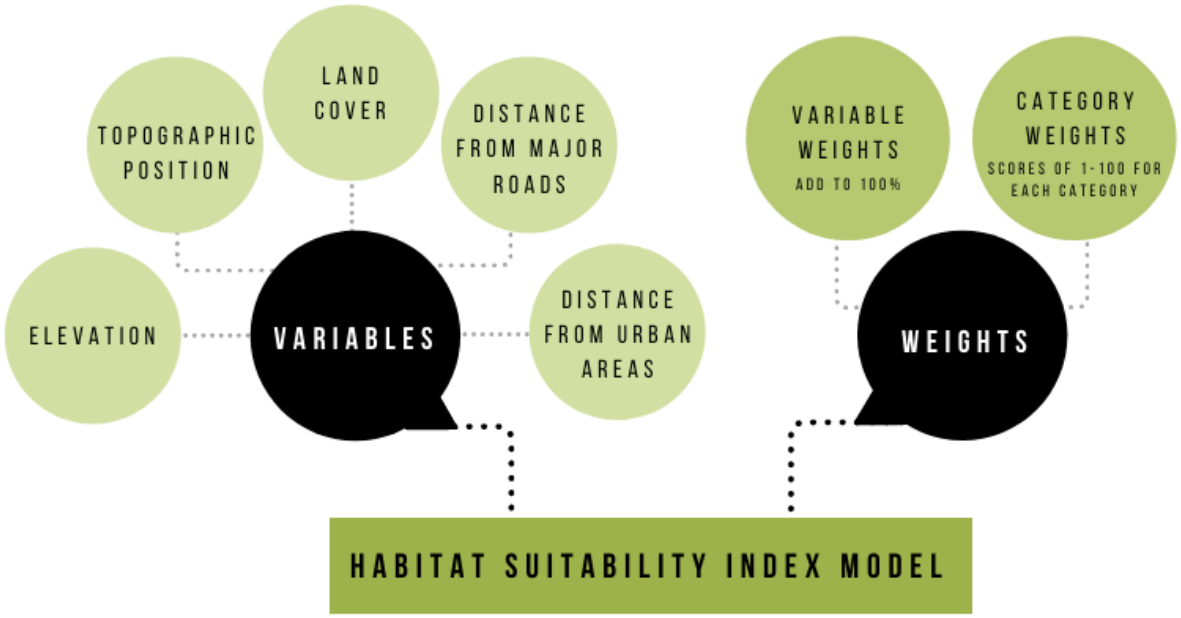


Figure 6. Flowchart of the Habitat Suitability Index modelling process, using ArcGIS Pro.

Resulting models and comparisons

A total of three SP models and six HSI models were generated (Figure 7).

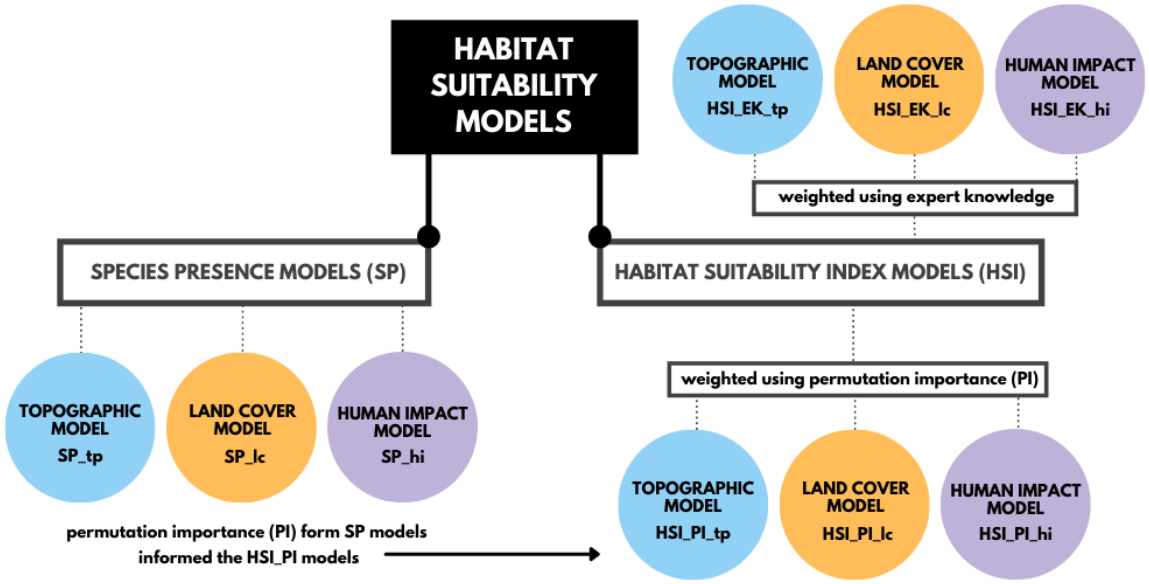


Figure 7. Summary of the habitat suitability models and names. The colours correspond to the groups of variables used for each model. The model codes will be used throughout the chapter.

All resulting models (SP and EK) were reclassified in ArcGIS Pro to match the biological interpretation of habitat suitability scores as described by (Majka, Beier and Jenness, 2007). The suitable patches (areas of suitability ≥ 50 used as default suitability threshold by the (Majka, Beier and Jenness, 2007)) were extracted and converted to polygons to calculate the suitable area predicted by each model. The corresponding suitable patches were compared by looking at the area cover (percent of study area covered by suitable patches predicted by each), suitable patch overlap (percent of study area predicted as suitable by both models), as well as the area covered solely by each of the models (percent of study area predicted as suitable by model a but not model b) (Figure 8). These four metrics were used to compare the model pairs.

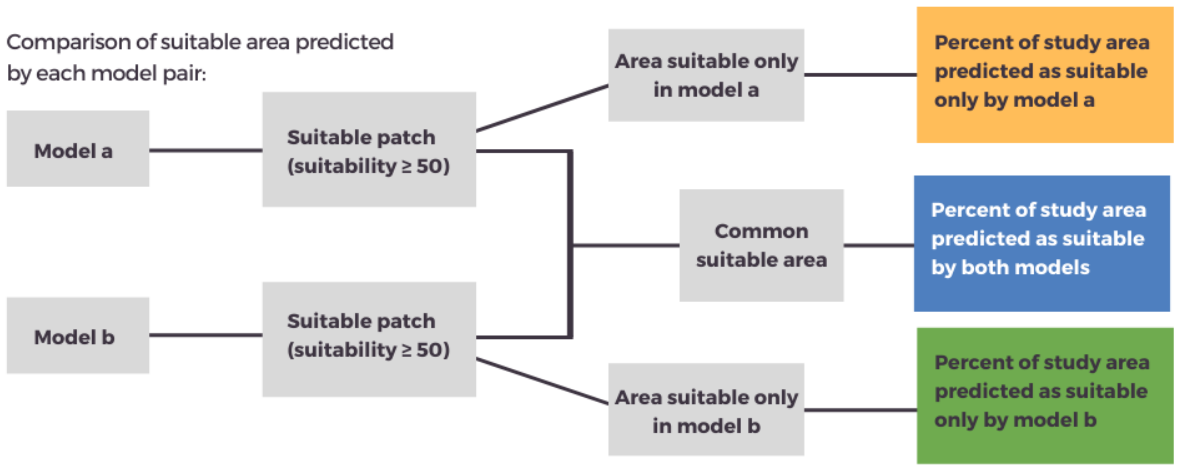


Figure 8. A flowchart of the model analysis outlining the process following the model generation whereby the area covered by suitable habitat was calculated.

In order to test the predictions against an existing dataset, the models were compared against the known bear distribution as described by Mertzanis in the European Report for the Status, Management and Distribution of Large Carnivores (Kaczensky et al., 2012b) using the updated distribution shapefiles for the period 2012-2016 provided by (Kaczensky et al., 2021) for the IUCN assessment of population figures for European brown bears. The dataset was available under 'CC0 1.0 Universal (CC0 1.0) Public Domain Dedication' licence and will be hereafter referred to as the 'IUCN dataset'. Each of the models was compared with the IUCN dataset to reveal the areas where the two datasets overlapped (*percent overlap*), where the model in question predicted suitable areas beyond the extent of the IUCN dataset, and the areas where known distribution in the IUCN was not predicted as suitable by the model.

Final models - Gradient Habitat Suitability models

Finally, gradient models were used to present the suitability as associated with different uses of the habitat by bears (Figure 5) across the field site, highlighting the suitable patches, potential linkage areas and stepping stones (results of this described in Savvantoglou et al. 2017), as well as areas of potential future recolonisation. The skeleton overview of the full methodology of this chapter is outlined in Figure 9.

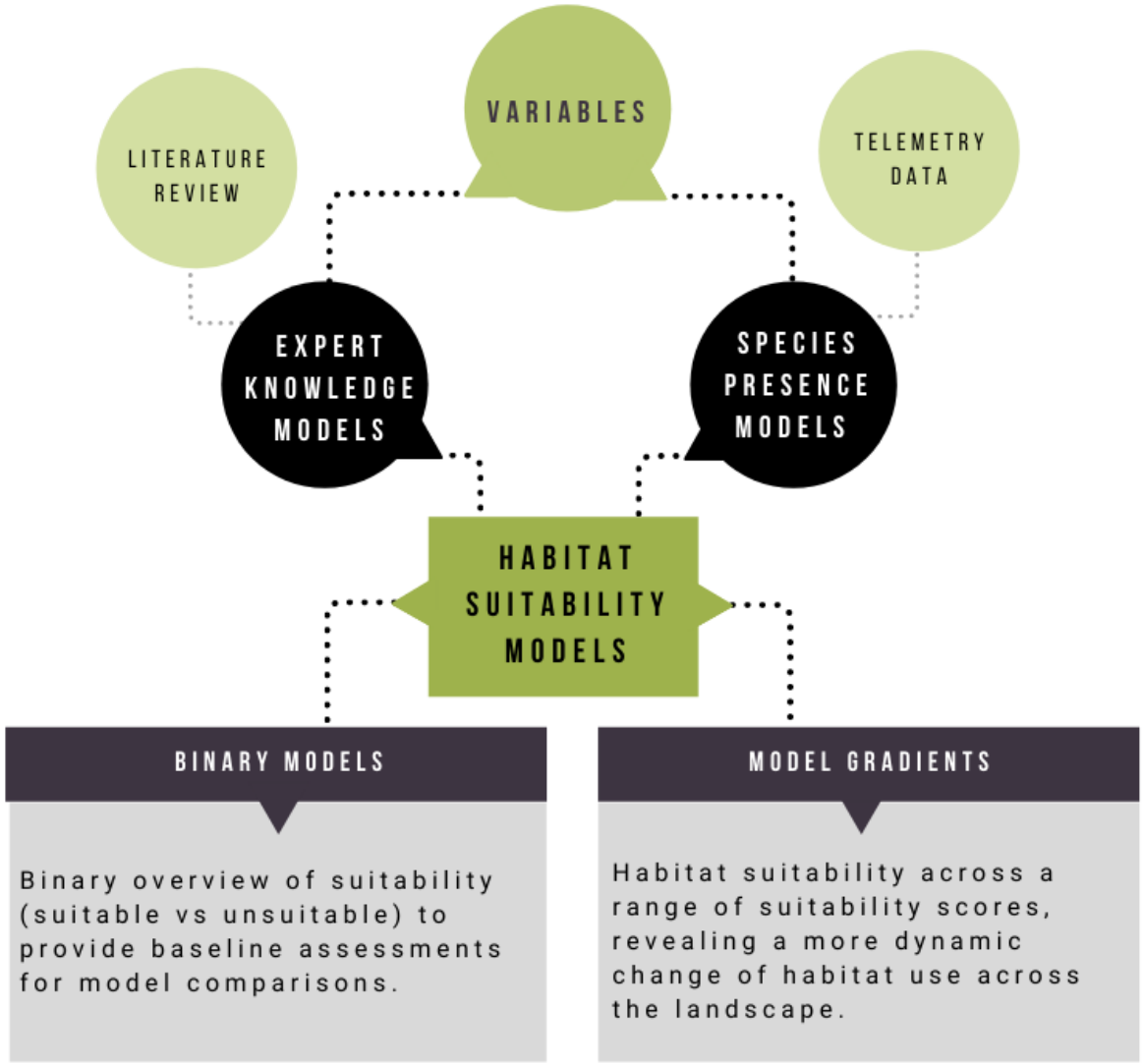


Figure 9. Flowchart of the methodology skeleton of this chapter.

Results

Species Presence Models

Following the resampling of the variable dataset to the finest cell size, the variables were tested for collinearity, aiming to remove one of two variables in a pair when their Variance Inflation Factor (VIF) was greater than three, to reduce over-prediction in the model. The VIF did not suggest any evidence of collinearity, so all variable predictors were considered.

Following the collinearity analysis, the variables were used to run an ENMevaluate analysis, which informed the final model. The variables used for each model, as well as the feature types and regularisation multipliers used in each model specified by the ENVevaluate analysis,

are outlined in Table 4. Due to the extent of the study area (328,340 km^2) and the fine resolution of the variables, these models were not able to run a full ENMevaluate analysis. Instead, a smaller sample area was used (850km^2) to inform the models' most suitable feature types and regularisation multiplier values. The resulting Training AUC of the final MaxEnt analysis for each model is shown in Table 4.

Table 4. Overview of the models, variables used and ENMevaluate best model AICc, feature type and regularisation multiplier values, along with the best model's Training AUC after the MaxEnt analysis. Due to the fine resolution of the variables, the regularisation values and feature types for these models were determined by running the ENMevaluate analysis in a smaller sample region within the study area due to processing limitations of running ENMevaluate on the full dataset.

Model type	Variables used	Model	Feature types	Regularisation multiplier	Training AUC
Topography	• Elevation • Topography	SP_tp	LQHPT	0.5	0.73
Land cover	• Elevation • Topography • Land cover	SP_lc	LQHPT	0.5	0.75
Human impact	• Elevation • Topography • Land cover • Distance from urban areas • Distance from roads	SP_hi	LQHPT	0.5	0.80

Binary Habitat Suitability models

The three highest contributing variables for each binary model along with the permutation importance and shape of the response curve are listed in the Appendix (A.2. MaxEnt model outputs for the Species Presence models) along with the MaxEnt model raster output and model results as well as the Sensitivity vs 1-Specificity and Omission graphs, Jackknife analysis of mode gain, and the response curves for each SP model. The 'min' model permutation importance values used to inform the HSI permutation importance-weighted models (HSI_PI) as well as the model weights informed by literature and expert knowledge are shown in Figure 10. The HSI_PI models relied heavily in elevation, while land cover was considered the most significant variable in the HSI_EK models, with the elevation weight in the PI models being as high as three times the value of the expert knowledge models.

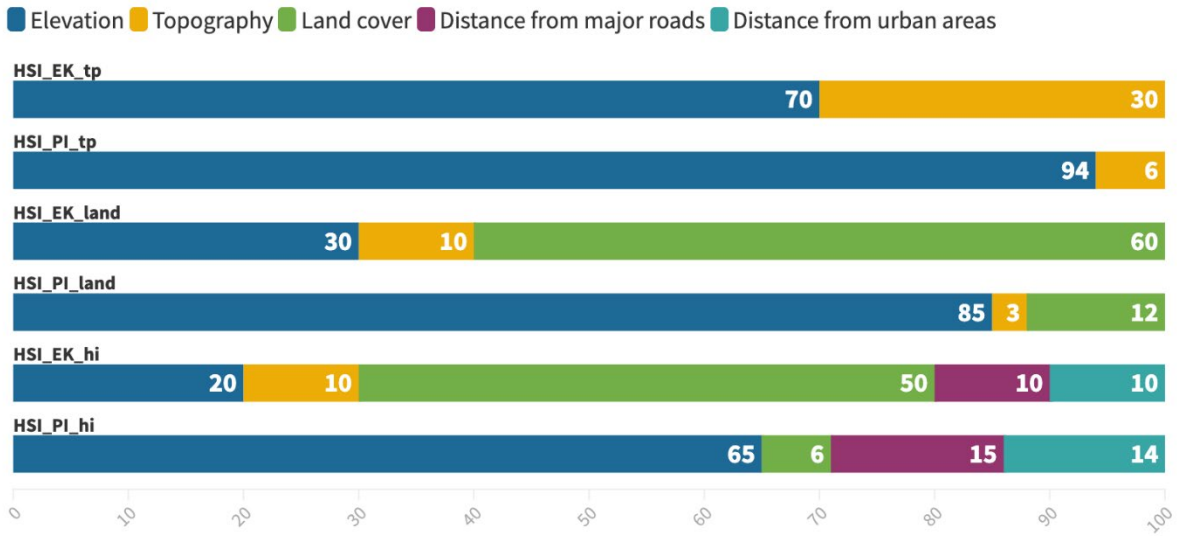


Figure 10. Weights used for the Expert Knowledge models. HIS_EK: Expert knowledge models advised by the literature. HSI_PI models informed by the results of the Species Presence models' Permutation Importance values, while 'tp' refers to topographic layers, 'land' to the addition of the land cover variable to the models, and 'hi' to the addition of human impact variables (distance from major roads and distance from urban areas).

The binary models were compared to reveal the differences in each model's coverage and model overlap. Firstly, the Expert Knowledge models comparison with the Permutation Importance models (HSI_EK vs HSI_PI) produced the most similar predictions across the model types, with the Topographic models predicting nearly the exact same suitability patches (Figure 11). The Expert Knowledge models were compared to the Species Presence models with the high-resolution data (HSI_EK vs SP), suggesting a smaller overlap (from under 60% for the topographic models to just over 22% for the Human Impact models (Figure 11). Overall, all model type predictions were consistent in generating more similar predictions when less layers were included in the models, with the overlap decreasing as the layers were added to the analysis (Figure 11). A full set of comparison results found in the Appendix, section A.3. Binary model comparisons.

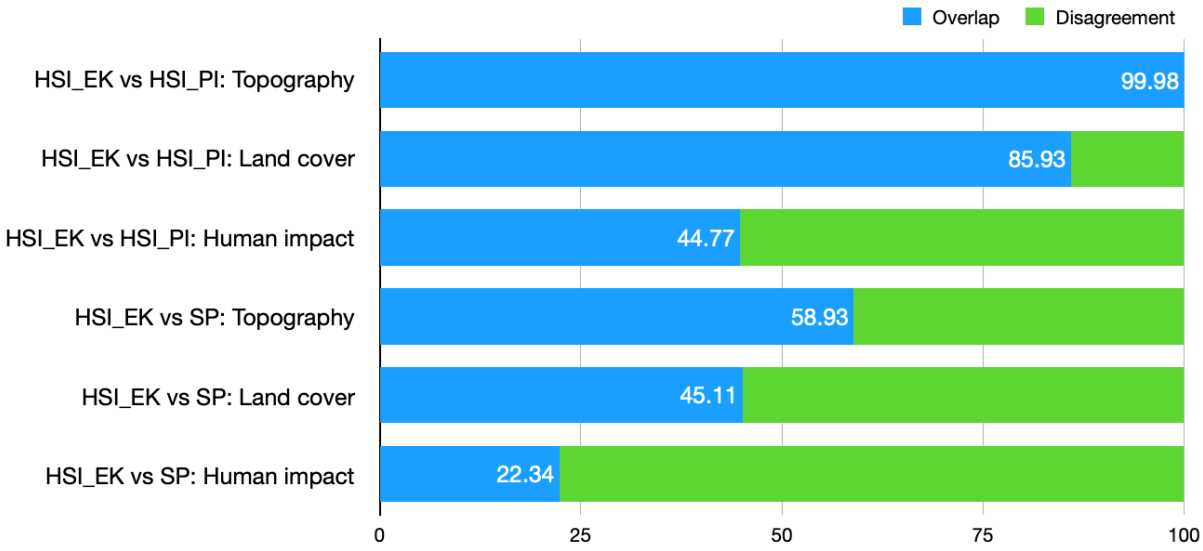


Figure 11. Proportion of overlap (blue) and disagreement (green) between binary model pairs. ***HSI_EK vs HSI_PI****: comparison of HSI models created advised by the Species Presence model's Permutation Importance Values and ones created using Expert Knowledge.* ***HSI_EK vs SP****: comparison of HSI models advised by Expert Knowledge and high-resolution Species Presence models.*

Finally, the last comparisons aligned each model binary output to the known distribution in Greece, as described in the IUCN assessment of population figures for European brown bears (Kaczensky et al., 2021). The percent of overlap between the IUCN dataset and the model prediction in question as well as the areas covered only by one of the layers are shown in Figure 12. Additionally, in all model types, some of the predictions fell outside the IUCN dataset extent, with the SP_max adding the least amount of suitable areas outside the extent, and the HSI models (PI models followed by EK) predicting the largest areas of suitability outside the IUCN cover models and percent agreement with IUCN data in Figure 12 (area predicted only by the model in green). Table 15 detailing the results found in the Appendix, section A.3. Binary model comparisons.

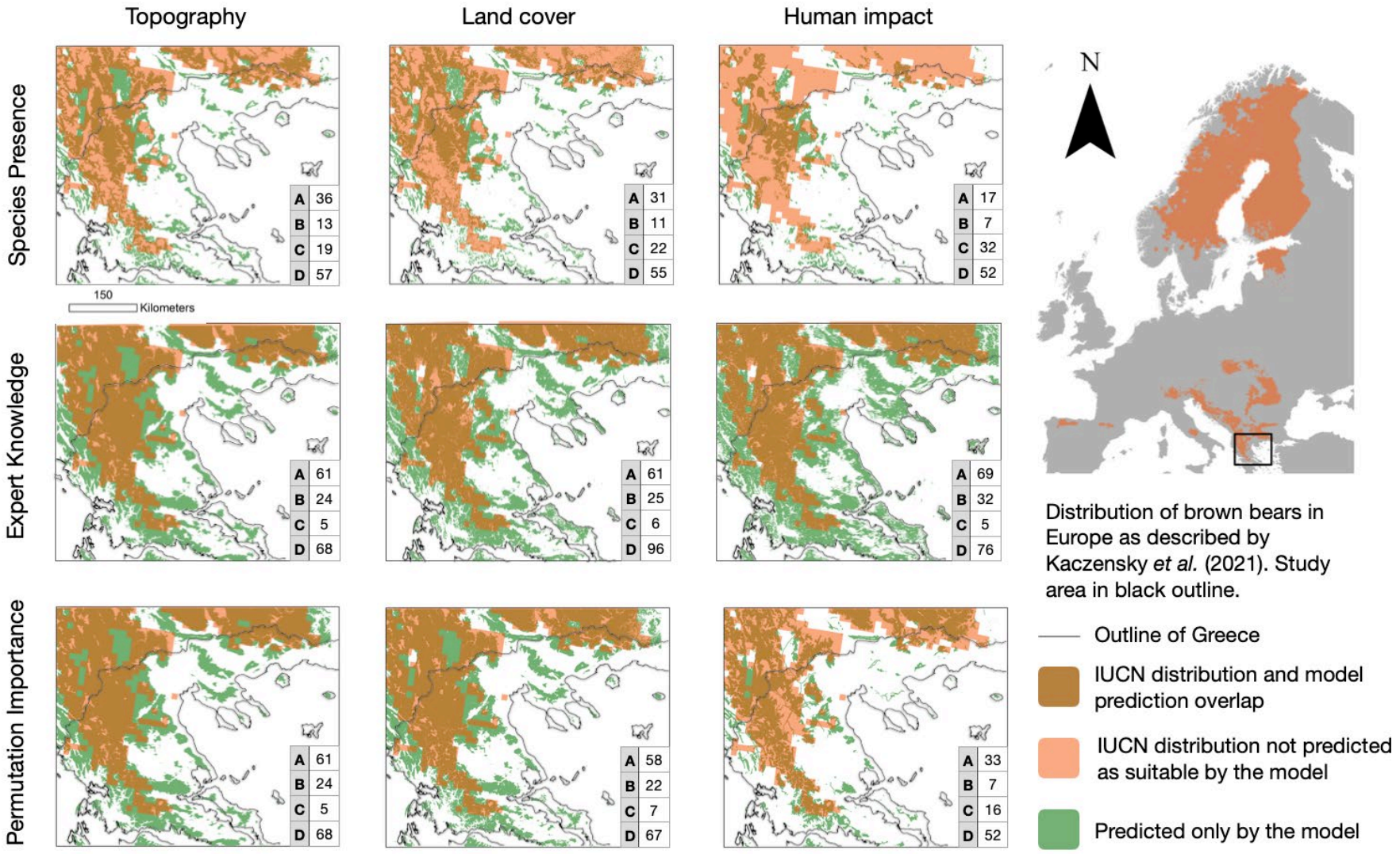


Figure 12. Binary suitability models (suitability ≥ 50) (in green) projected under the European brown bear distribution dataset (Kaczensky et al., 2021, available under a 'CC0 1.0 Universal (CC0 1.0) Public Domain Dedication' licence) (in orange). Brown showing where the two predictions overlap. Tables detail the model's performance in predicting suitable areas for bears in mainland Greece in comparison to the IUCN dataset (percent cover of study area by IUCN dataset = 44%), with A: percentage of Study Area (SA) covered by the model; B: percentage of SA covered only by the model; C: percentage of study area covered only by the IUCN dataset; and D: percentage of study area covered by both models.

Gradient Habitat Suitability models

The suitability gradient models show a larger degree of difference compared to the binary models. The gradients revealed by the Expert Knowledge models (HSI_EK) as well as those generated by the Species Presence models (SP) describe a more complex landscape of suitability (Figure 13 and Figure 15). The models created using the HSI method, but advised by the SP Permutation Importance values, presented a more simple landscape (Figure 14). The maps presented here are those of the Land Cover models that were created using elevation, topographic position and land cover type as adding human impact variables in the analysis created more inconsistent results across the three model types (large drop in percent overlap between the models) (Figure 11 and Appendix, Table 14).

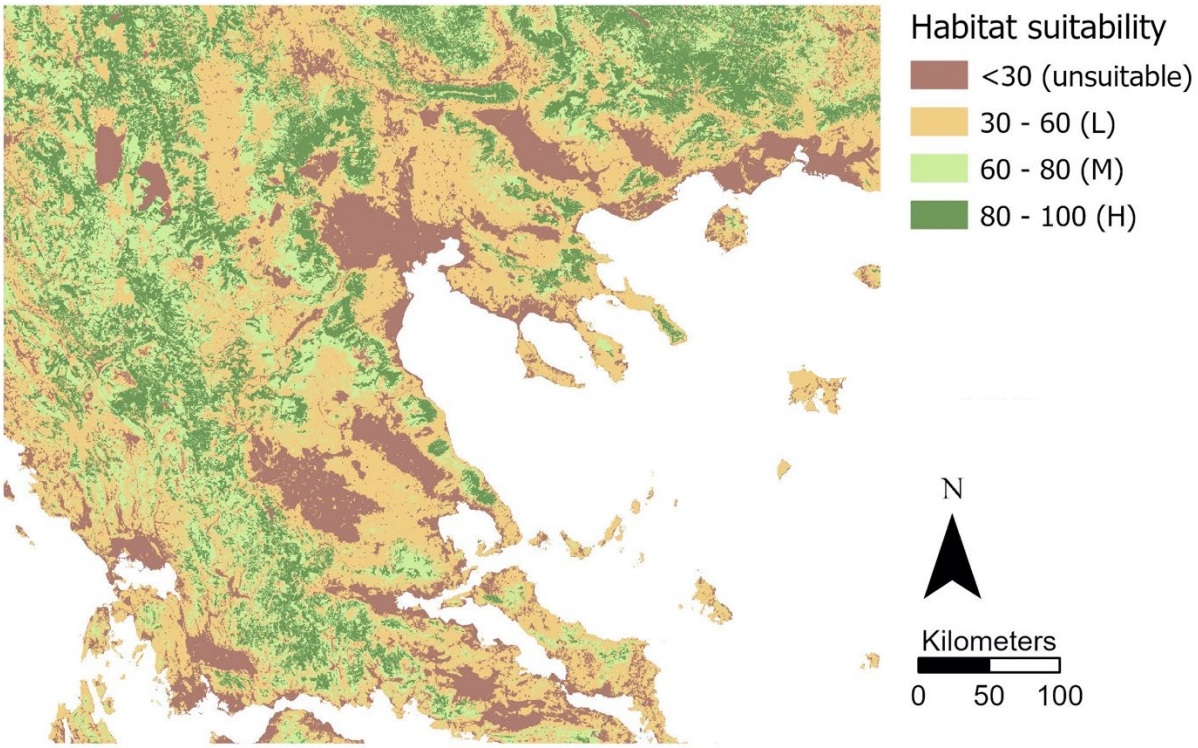


Figure 13. Habitat suitability gradient as predicted by the Expert Knowledge model using topographic and land cover layers (HSI_EK_lc).

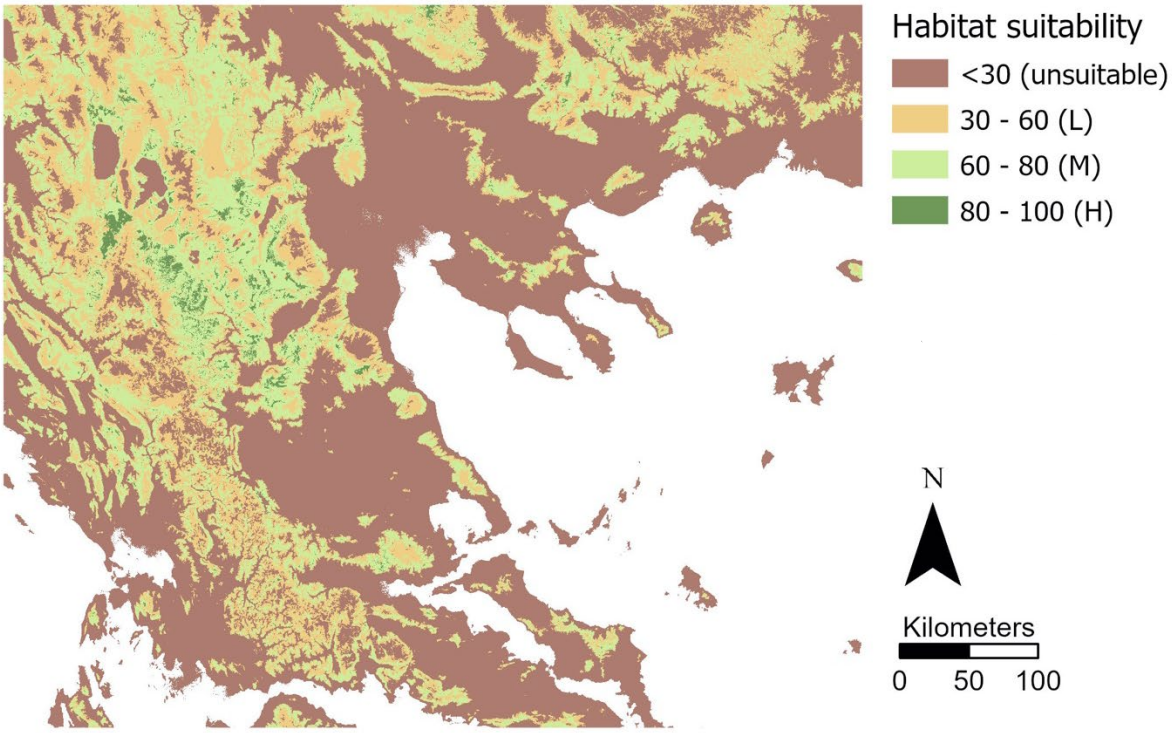


Figure 14. Habitat suitability gradient as predicted by the Species Distribution model advised using topographic and land cover layers (SPlc).

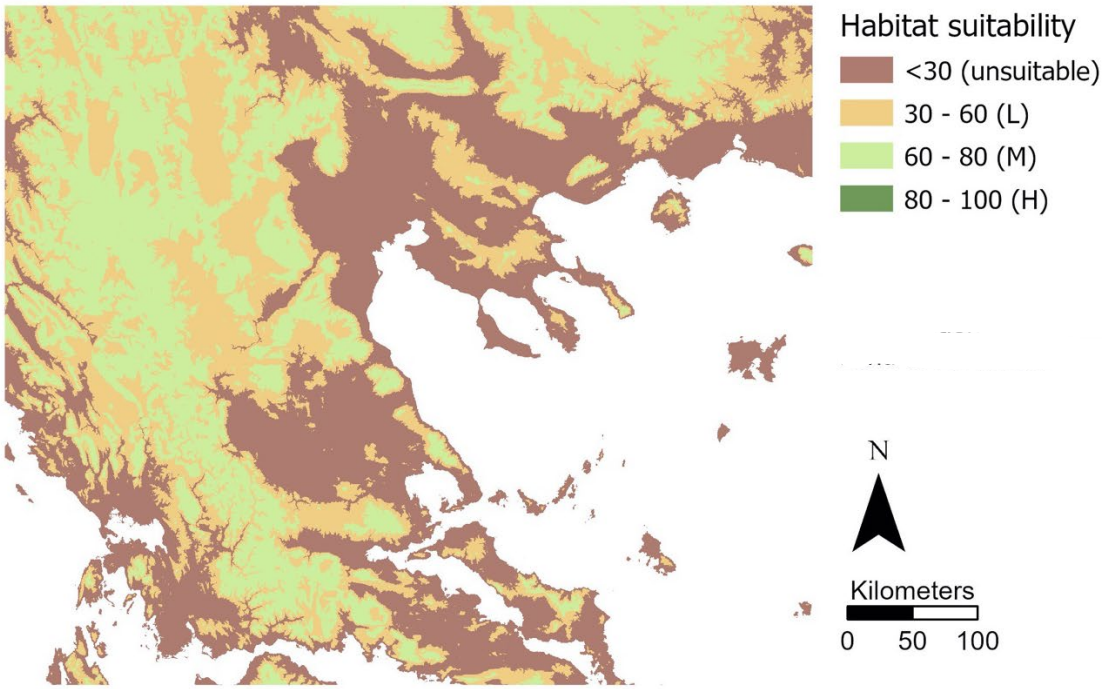


Figure 15. Habitat suitability gradient as predicted by the Expert Knowledge method, but weighting the topographic and land cover layers as advised by the Permutation Importance (HSI_PI_lc).

Moreover, the HSIs here reveal areas suitable to bears that are either not yet colonised by the species or where bears have been reported but are not currently considered areas of

permanent presence (Mertzanis, Psaroudas and Karamanlidis, 2020). These areas are highlighted in a binary model in Figure 16 using a model average of the Expert Knowledge land cover model (HSI_EK_lc) and the Species Presence land cover model (SP_lc) and overlaid on the IUCN dataset to highlight the areas outside the known distribution.

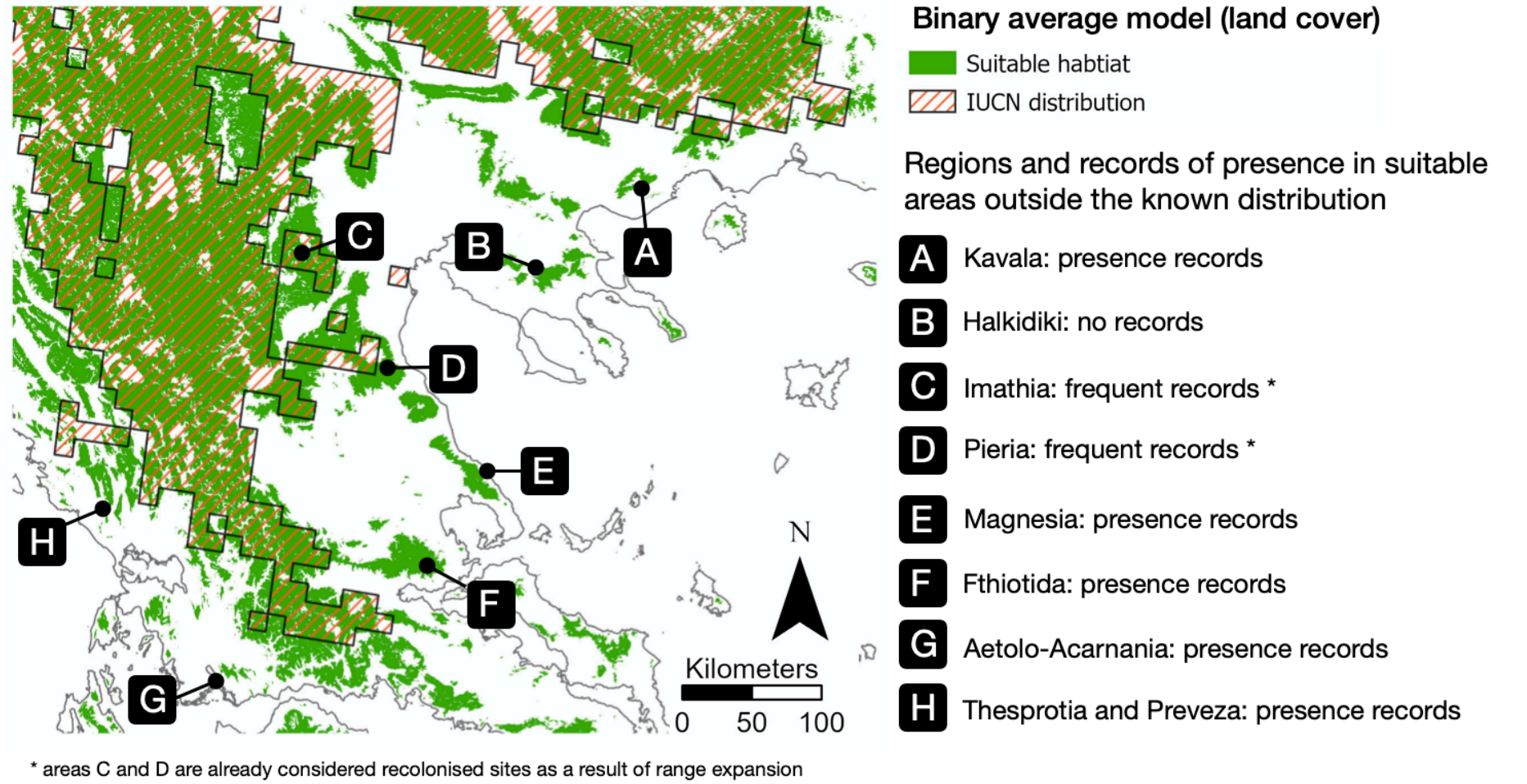


Figure 16. Binary average suitability of the HSI_EK and SP land cover models, showing regions and records of presence in suitable areas outside the known bear distribution as depicted by the IUCN dataset (Kaczensky et al., 2021)*. Records of presence as described in the National Action Plan for Ursus arctos* (Mertzanis, Psaroudas and Karamanlidis, 2020)*, pointing at the areas of current reestablishment (frequent records) and occasional presence records (presence records) highlighting areas of future colonisation.*

Discussion

The work here aimed to obtain a reliable representation of habitat suitability for Greek bears as predicted by two broad suitability modelling techniques. The immediate purpose of the suitability models, and subsequent analysis of habitat connectivity was to advise the fieldwork conducted in the following chapters. Our second aim was to use bears as a case study species for comparing the predictions of species distribution and habitat suitability techniques that are transferable across species. We believe that this case study provides evidence that these two broad categories for suitability modelling, when optimised with the best available data (set of variables, presence data, available literature) can create comparable predictions, suitable to guide future field studies and advise management. Thus, we believe that this case study can encourage the use of suitability modelling under scenarios where data is lacking, such as in:

1. Cases where less is known about a species, thus using Species Presence methods to explore suitability.
2. Cases where a robust national-level dataset is not present, or study areas with no prior monitoring efforts, but where more is known about the study animal, therefore using expert knowledge methods to predict suitability.

In this case, brown bears present a suitable candidate for comparing different modelling techniques as, in the cases of both modelling types (species presence and expert knowledge as the two broad categories), an abundance of information was available. The large dataset of telemetry records was able to represent the complexity of habitat use by bears in the Western core population. Additionally, brown bear ecological preferences and tolerance to human disturbance have been studied extensively in Europe and around the world, providing a very robust body of literature describing their response to the variables used in this analysis. In fact, brown bears are one of the most studied species in published literature, third only to red foxes, *Vulpes vulpes*, and wolves, *Canis lupus* (Brooke et al., 2014).

Having a comprehensive understanding of a species' ecological preferences means that it was possible for us to assess the outputs of the Species Presence models by comparing the species' response to each environmental variable against our current understanding

of bear behaviour. Finally, the current distribution of bears in Greece is well-monitored, which made it possible to test the model predictions by comparing them with the known bear distribution as described by Kaczensky *et al.* (2021). Therefore, the assumption here was that using bears as a case study species meant that each model could be generated with a large degree of confidence in location data integrity and ecological preferences, while also having access to a distribution dataset with which to compare the outputs.

Review of the comparisons between model pairs

Overall, when looking at the entire set of pairwise model comparisons, the mean overlap (35.7%) exceeds the mean disagreement (24.3%) between the pairs, with very comparable medians (52% for the overlap and 48% for the disagreement), suggesting large differences between some of the models. The inconsistent predictions between the pairs highlight the sensitivity of these models to their input data and parameters used. Bias associated with each of the approaches used to predict suitable habitat in this study. In their review of existing literature and creation of a set of '*Standards for Distribution Models in Biodiversity Assessments'*, (Araújo et al., 2019) outline the best practice standards for SDMs. Each model's quality is assessed by the four main aspects of the SDM analysis: 1. Response variables; 2. Predictor variables; 3. Model building; and 4. Model evaluation. To maintain consistency in the interpretation and evaluation of the work carried out in this chapter, these four aspects will be discussed here.

Response variables

In SP models, telemetry data can introduce autocorrelation in the model due to the spatial and temporal proximity of the datapoints. Additionally, differences in behaviour between sexes or even individuals of the same species can introduce bias in the dataset. As an example, female bears (with or without cubs) tend to have restricted home ranges, while adult and subadult male bears establish larger territories that greatly exceed those of females (Gau et al., 2004; Kanellopoulos et al., 2006; Karamanlidis, Kopatz and de Gabriel Hernando, 2021; Mertzanis et al., 2005). Furthermore, individual behaviour might not necessarily reflect that of the species as a whole as differences in behaviour exist between individual bears (Hertel et al., 2019; Hertel, Swenson and Bischof, 2017). To reduce spatial dependencies between data points, the occurrence dataset is often

rarefied or *thinned*, creating a dataset with less occurrence data points in exchange for a reduction in autocorrelation (Boria et al., 2014; Veloz, 2009). Here, species occurrence records were collected in the form of telemetry data from GPS collars. The accuracy of the telemetry data was estimated to have a mean error of 30m (Giannakopoulos, Akriotis and Mertzanis, 2011; Mertzanis et al., 2011) and the spatial thinning of the dataset minimised bias concerning spatial autocorrelation (Aiello-Lammens et al., 2015; Brown, 2014). The SDM Toolbox (Brown, 2014; Brown, Bennett and French, 2017) provides two spatial rarefying solutions to reduce autocorrelation: data can be thinned either by removing occurrences using a set distance or by gradually filtering it using a heterogeneity raster (climatic, topographic of habitat heterogeneity). This latter rarefying method ensures more data is preserved in areas of high heterogeneity, thus better representing the movement of the species through the landscape. The occurrence dataset for this study was thinned using a topographic heterogeneity raster to reflect the elevation changes in the species' range.

Another issue with the occurrence data is that SP models can also be prone to bias associated with background point selection (also referred to as pseudoabsences). Background data in MaxEnt analysis is a set of randomly selected points (raster cells) in the study area that are used to train the model. In cases where background data falls too far away from the clusters of occurrence data, those cells can be falsely interpreted as areas of low suitability. This can easily occur for locations where the habitat is environmentally suitable, but the species has not colonised it yet or surveys failed to locate it in the area. Selecting pseudoabsence points from such areas might result in a misleading model (Anderson and Gonzalez, 2011; Barbet-Massin et al., 2012). The way background data sampling from across the study area could bias the model is that all the bears fitted with a telemetry collars in this dataset were situated in the western core population of brown bears in Greece, resulting in a dataset biased to the environmental conditions in that region of Greece. MaxEnt incorporates 'bias files' in the analysis, ensuring that background points are selected preferentially from areas of higher sampling density (Anderson and Raza, 2010; Phillips et al., 2009). Here, all models were run using bias files, restricting the background point collection using a kernel density layer of the occurrence points (Bellamy, Scott and Altringham, 2013; Brown, 2014).

When it came to the Expert Knowledge models, the challenge was to extract and assess the expert understanding and published work on a species' habitat preferences that often varies depending on the aim, scale, season of the study and years of observations, which makes deciphering a representative suitability score particularly challenging. Furthermore, what might be suitable for a bear in one particular area might not be suitable in another as other underlying variables could be the driver for its selection or avoidance (as an example, hibernation season is different for bears in the north of Europe to the south (Swenson et al., 2007), suggesting that seasonal models would need to take great care in accounting for climatic differences across the species' range). These inconsistencies in literature regarding habitat selection were initially managed by informing the models with studies from Europe and, when possible, from Greece or countries with similar climate (Mediterranean and Balkan countries). Given the large body of literature for brown bears, the resulting Expert Knowledge models were more consistent with what we know about bear distribution in Greece.

Overall, we recognise the bias introduced by an occurrence dataset that does not represent the entire study area (Species Presence models) and potential inconsistencies/gaps in literature that may have resulted in inaccurate estimations of suitability scores (Expert Knowledge models) and believe that the measures taken were appropriate to create the most accurate models with the available data.

Predictor Variables

In the case of the predictor variable selection, variables already evidenced by previous literature as important factors for habitat selection by brown bears were chosen. The variables were tested for collinearity to ensure that no overprediction is introduced to the model. We are confident that the steps we took to select and prepare the variables for the study, coupled with the step-wise addition of variables (topographic, topographic + land cover, and topographic + land cover + human impact) produced a good example of the SDM best practice standards (Araújo et al., 2019).

Model building

Both model types utilised commonly used methods to generate their prediction (MaxEnt for Species Present models and Weighted Suitability for Expert Knowledge models). As

discussed in the methods section, measures were taken to ensure that the best parameters for the MaxEnt models were used for each Species Presence model. The Expert Knowledge models were created using a standardised method in ArcGIS Pro and the inclusion of a set of EK models informed by the Permutation Importance values allowed for an alternative analysis whereby the suitability scores and weights were informed by the MaxEnt output for their corresponding Species Presence (SP) models. Rather than suggesting that a single model generated here displayed the most accurate prediction, we present a variety of models, all of which were created with the best practice standards in mind and showcasing these - often stark - differences between outputs.

Evaluation

Following the protocol from (Bellamy, Scott and Altringham, 2013), evaluations for the SP models were conducted by sample bootstrapping, allowing for ten iterations of each model to run, resulting in a final average model. The final model's AUC was then used to demonstrate the model's fit. As an additional way to evaluate the model outputs, both the large body of literature on bear ecology and the existence of a European-wide distribution dataset served as tools to assess each model's predictions. The Species Presence models were assessed for their ability to reflect the patterns of habitat use by bears described in published literature.

The EK evaluation is a little more complicated as the analysis is completely guided by the researcher's selection of suitability scores and variable weights. Therefore, careful consideration of every decision taking in steps 1-3 will ensure a more reliable model based on a robust set of responses of the species to various environmental and anthropogenic factors.

Additionally, the fact that multiple models were generated in this study allowed for an evaluation that incorporated comparisons between the model outputs. The pair-wise comparisons conducted in this study, as well as each model's relationship with the IUCN dataset add another level of evaluation often missing from traditional SDM studies. It is worth noting that this European-wide dataset collated presence data from the period of 2012 to 2016 to create a distribution map at a 10x10km resolution (for more

information, see Kaczensky *et al.* (2021, 2012a). Therefore, the IUCN dataset represents bear distribution across Europe in a much more coarse scale than the models in this study and some of the variance between the models created here and the European distribution dataset might be down to scale differences. As an overview of the best practice standards (Araújo et al., 2019) we are confident that the different datasets used in this study were processed appropriately to minimise bias in the data and methods used and create accurate predictions relative to the data and method used for each model.

Species Distribution Model considerations

Studies exploring the most appropriate spatial scale for SDM analysis have highlighted the importance of considering this factor carefully, ensuring that the selected scale represents the species and goals of the analysis (Addicott et al., 1987; Araújo et al., 2019). Taking this into consideration, along with the fact that bears are large-ranging animals, the scale used here reflects not only the minimum resolution of the variable dataset, but also closely resembles the lower limit of female brown bear home range in Europe (Cirovic et al., 2015; Kanellopoulos et al., 2006; Mertzanis et al., 2005; Swenson, Sandegren and Soderberg, 1998). Therefore, we were confident that the spatial scale was appropriate for the study in terms of representing the species movement in the landscape in a fine scale and maintaining the landscape heterogeneity as represented by the DEM dataset (elevation). However, we would like to caution that, in the case of fine resolution models, computing power is a great limiting factor in data processing, adding a significant amount of time in the MaxEnt analysis.

Comparison of Expert Knowledge and Species Distribution models

Due to the different techniques used to generate the models, is hard to draw a definitive conclusion on the performance of the two methods. To account for this, the variable weights and suitability scores for each of the models were extracted from variable response curves and the permutation importance values of each Species Presence model. The resulting models (HSI_PI), generated in ArcGIS using the same method as the EK models, were thus more suitable for a direct comparison. Indeed, once the values

from the SP models were used to inform HSI_PI models, the outputs of the binary models resembled each other with as much as 99.98% overlap. This model similarity suggests that the differences in suitability scores and model weights in the HSI_PI models to the literature-derived information did not influence the model outputs to an extent that the predictions of the binary suitable habitat were altered. The results of this comparison indicate that the EK models was adequate in predicting suitable habitat in a very similar way to models informed by species response curves and the variable permutation importance generated with a SP model.

However, it is also important to acknowledge the difference in the size and extent of suitable patches with the HSI method compared to the SP models. The HSI analysis uses purely landscape-specific data and no presence/pseudo-absence data are required, which allows for the analysis to be carried out evenly across the study area. Admittedly, presence data in SP models can improve the reliability of the design (Correa Ayram *et al.*, 2015; Mateo-Sanchez *et al.*, 2015), but it can also restrict the prediction by biasing the analysis with autocorrelated data and non-representative individuals of the species. The HSI method, even though it perhaps presents a more speculative approach to SDM and introduces a large amount of overprediction in comparison to the SP models, it evenly assesses the study area unbiased of where the species has been found previously. Arguably, some of these areas of overprediction falling outside the known bear distribution in Greece fulfil the ecological requirements set by the model and therefore could be suitable areas for future bear colonisation (discussed further in (Savvantoglou et al., 2017).

Habitat gradient models

The gradient models revealed a much more comprehensive overview of the habitat suitability for brown bears across Greece. While binary models are more straight-forward to compare across the different methods, they do not portray the dynamic change of suitability across the landscape. Using a more diverse gradient of suitability ranges, it is possible to start making some assumptions on the type of use of that habitat by the animal in question. As outlined in Figure 5, high suitability scores are associated with consistent use and breeding, while lower scores can indicate habitats of used occasionally for activities such as feeding (Majka, Beier and Jenness, 2007).

Furthermore, this type of visualisation offers a better understanding of the areas that fall under the binary suitability threshold that could act as stepping-stones or corridors. Work done as a part of this project and described in Savvantoglou et al. (2017) reveals these potential corridors linking core areas and other smaller, newly recolonised zones. This is especially important, given the fact that conservation work has been carried out in Greece to improve connectivity within core areas dissected by major roads (Karamanlidis et al., 2012; Mertzanis et al., 2011). Understanding connectivity could be improved between core areas and predicting the future movements of bears into areas not currently colonised could drastically benefit bear conservation in Greece.

The results of our gradient maps, especially those of the Species Presence models (SP) and Expert Knowledge models (HSI_EK) provide a good estimate of the suitability mosaic across Greece. Some areas highlighted in Figure 16 as potential future recolonisation areas have already had confirmed sightings of bears (Mertzanis, Psaroudas and Karamanlidis, 2020), proving that movement between these areas is not only possible, but it is already underway. This is another evidence that both modelling methods were effective at predicting suitability across the landscape at a level refined enough to provide valuable information for future management plans. We believe that the work presented here can provide a baseline suitability mosaic to advise monitoring in areas outside the core habitats, where less is known about the presence of bears and use of habitats. Surveys aiming to create a better on-the-ground understanding of bear occupancy in these areas could contribute to a much more comprehensive overview of bear distribution across Greece.

The work outlined in this chapter explores the suitability mosaic across mainland Greece, as well as zones that could serve as linkages, and future colonisation areas. Suitability studies for bears in Greece have looked at suitability across smaller study areas such as the north-eastern part of the Pindos Mountain Range (Almpanidou et al., 2014), or have looked at a more coarse scale across in European-wide studies (Kaczensky et al., 2021, e.g. 2012a). The most recent study looking at bear distribution across Greece was included in the National Management Plan for *Ursus arctos* report for the LIFE-IP 4 NATURA (LIFE16 IPE/GR/000002) programme (Mertzanis, Psaroudas and Karamanlidis, 2020). The study outlines the areas of range expansion between the years

2000 and 2015 as well as areas of permanent areas and areas of recent and future recolonisation (Mertzanis, Psaroudas and Karamanlidis, 2020). Here I present work complimentary to this report by exploring multiple methods for modelling suitable habitats and creating a fine-scale suitability gradient for *U. arctos* in Greece. The inclusion of the neighbouring countries in the north allows for the exploration of suitability and potential connectivity assessments in a cross-country basis so as to not falsely restrict dispersal to country borders.

Additionally, while not included in this chapter, work conducted during this project carried out an initial exploration of the study area via an expert knowledge model as well as a subsequent habitat patch and corridor analysis and is described in (Savvantoglou et al., 2017). The results of that work assisted in revealing areas of dispersal potential and selecting a suitable study area for the work in following chapters. The corridor analysis case study was peer-reviewed and published in the 18th Hellenic Forestry Congress and has been appended in the Research Outputs section (Savvantoglou et al., 2017).

Conclusion and thoughts on application of results

The work conducted in this chapter allows for an analysis of our study area and an exploration in the habitat preferences of bears across their Greek range. This is, to my knowledge, the first species distribution modelling study for the Greek range of brown bears, incorporating differences in modelling outputs and datasets used. Our aim, for the purpose of this project, was to carry out habitat assessments that would aid the planning stage of the field surveys described later in thesis. However, the large body of literature in suitability modelling techniques prompted a more in-depth analysis of the study area, revealing a multitude of models and showcasing the sensitivity of these methods to changes in input data and parameters. Admittedly, it is important to acknowledge the large variation in predictions across the models and urge that these methods are be studied with much caution.

However, I strongly believe that habitat suitability models are invaluable at the planning stage of conservation efforts when it is essential to focus on modelling techniques that could be generated with a thorough understanding of the focal species' ecology and behaviour or species occurrence records available from previous studies. In the case of EK models, their power lies in their simplicity and in providing a system of rapid examination of a large study area without the existence of prior field surveys can be ideal for the initial steps of a study.

This study recognises that bears are an umbrella species and the more accurate the suitability predictions are for them, the more they will be able to envelop other species they share their habitats with. Ultimately, we hope that these models can be subsequently used to drive conservation management plans for bears and the wider community of organisms within current and future bear range in Greece.

Chapter 3: Developing an iDNA method for monitoring brown bears

3.1 Introduction

3.1.1 Single-species monitoring using iDNA samples

The use of non-invasive survey techniques that allow for the detection of a target species without the need to directly observe or trap the animal has radically changed the way we monitor species in the wild (Long et al., 2012). Techniques such as camera trapping and sign surveys have been used to monitor animals while minimising the risk of harming the animal or altering its natural behaviour (Long et al., 2012). Non-invasive monitoring often replaces other survey techniques that might cause stress or increase the mortality risk for vulnerable or declining populations (Bekoff and Jamieson, 2019). With the emergence of non-invasive genetic monitoring, the ability to collect traces of DNA the animal leaves behind, the field has expanded to techniques even further removed from acquiring data directly from the focal species. Moreover, these techniques have been instrumental in surveying rare or elusive species that other methods fail to detect (Thompson, 2004). Genetic techniques for monitoring wildlife using non-invasive sampling methods, such as hair trapping and faecal sampling, have been widely used to survey mammals since their introduction in 1992 (Höss et al., 1992; Taberlet and Bouvet, 1992). These first non-invasive genetic monitoring studies focused on bear monitoring in Europe but the field quickly expanded to make encompass a multitude of techniques and species, all working under the premise that the focal species is not directly targeted for the collection of samples, rather data collection in the field focuses on the traces of DNA they leave behind.

A novel non-invasive genetic monitoring tool for rare and elusive species targets traces of these animals' genetic material by sampling invertebrates in their environment. The technique, known as invertebrate-derived DNA or iDNA, targets invertebrate groups to

trace the DNA of target species they came into contact with (Bohmann, Schnell and Gilbert, 2013; Lynggaard et al., 2019). Invertebrate-derived DNA lies under the general umbrella of environmental DNA (eDNA) monitoring, describing survey techniques that allow for a collection of samples not directly derived from the target species (Ruppert, Kline and Rahman, 2019). These can lend themselves well for both single-species surveys as well as monitoring broad species assemblages (Harper et al., 2018; Ruppert, Kline and Rahman, 2019). Next-generation sequencing has allowed for the assessment of species assemblages in single eDNA samples and is becoming more accessible as the analysis costs are reduced. This has resulted in eDNA studies tending to gravitate towards using a metabarcoding approach, providing researchers and conservationists with snapshots of the area's diversity of the group of species in question (including, but not limited to, vertebrates (Ji et al., 2020), mammals (Gogarten et al., 2020; Lee et al., 2016) and amphibians (Harper et al., 2018)). However, when looking at the assemblages in eDNA samples, the analysis might miss rarer species' DNA amongst the more frequent species found in the samples due to the differences in initial copy numbers in the sample (Harper et al., 2018; Ruppert, Kline and Rahman, 2019). The more common species' DNA in the samples increases exponentially during PCR, making it more difficult to detect the rare species in the samples. Single species detection with species-specific primers, using real-time PCR, has been shown to be more sensitive in detecting rare species, where the amount of genetic material within the sample is negligible (Harper et al., 2018; Lacoursière-Roussel et al., 2016; Schneider et al., 2016). This same principle applies when processing iDNA samples, therefore suggesting that single species approaches could be more effective at detecting rare and elusive species' DNA in the invertebrate samples.

Another reason to explore the use of the single-species PCR approach is related to the costs and expertise of running the laboratory analysis and subsequent data. Metabarcoding is associated with higher costs when the analysis is all carried out externally, or a very specialised skillset to process the raw sequenced data when the data analysis is done *in situ*. In their study, comparing qPCR against a metabarcoding approach for 380 samples, Harper et al. (2018) found that the time it took to processes the samples was very similar for both methods, but metabarcoding increased the sample analysis cost by 12%. The per-sample cost of metabarcoding can increase with smaller sample sizes, therefore for small-scale and low budget efforts looking at rare or

elusive species, a single-species qPCR is a more suitable approach. Therefore, in species monitoring, metabarcoding is a very powerful tool when looking at species assemblages or looking for the occurrence of multiple, potentially unknown species. However, for targeting a specific species, a qPCR approach is more efficient and sensitive. This chapter focused on developing the use of iDNA in single species monitoring. Specifically, the work outlined describes a series of experiments that form a single-species iDNA lab protocol for monitoring brown bears (*Ursus arctos*) using flies.

Invertebrates have been extensively utilised for tracking human disease vectors (e.g. tsetse fly (Steuber, Abdel-Rady and Clausen, 2005) and mosquito (Beier, 2003; Kiszewski et al., 2004)). Following the principles of tracking the DNA of a different organism in invertebrates, iDNA has subsequently been developed for ecological monitoring for a number of invertebrate groups and families, such as flies (Lee, 2016; Bohmann *et al*., 2013; Schubert *et al.*, 2006), mosquitoes (Townzen *et al.*, 2008; Kent and Norris, 2005), ticks (Gariepy *et al.*, 2012) and leeches (Schnell *et al.*, 2015). Simply, the novel use of iDNA in wildlife monitoring focusses on isolating traces of target species' DNA from the digestive tracts of invertebrates (Schnell et al., 2015). Prior to the use of iDNA for mammal monitoring, studies using end-point PCR and qPCR have looked at invertebrate predation by examining the predator's gut contents for prey DNA (e.g. Ma et al., 2005). More recently, like in the case of eDNA, published methods have largely concentrated on a metagenomic approach which attempts to reveal biodiversity in the sample, rather on species-specific monitoring (e.g. Abrams et al., 2018; Bohmann, Schnell and Gilbert, 2013), with the exception of a few studies that that looked at single-species detection as a small part of the work (Drinkwater et al., 2021; Schubert et al., 2015). This chapter aims to explore this particular method by developing a rapid, species-specific iDNA method for monitoring specific species in the wild, using brown bears as a case study. This includes the use of already established iDNA methods and their adaptation to single-species detection using real-time PCR (qPCR).

Non-invasive genetic monitoring has often been successful for species living at low densities because the sample collection does not rely on detecting the animal itself, but merely on sampling traces of DNA the species has left in its environment (Long et al., 2012). A valuable aspect of eDNA/iDNA monitoring methods is that they utilise generic

genetic sampling methods instead of relying on collecting samples using methods unique to the target species. The species identification is implemented at a later stage, using species-specific or universal primers (single-species and metabarcoding approach respectively). One of the biggest challenges in most non-invasive genetic monitoring techniques, however, is the sample quality collected in the field. Due to limited sample collection possibilities, surveyor skills and the potential degradation of genetic material when exposed to extreme environmental conditions, detection or individual identification is not always successful (Barnes et al., 2014; Thompson, 2004). Furthermore, in the case of iDNA studies that focus on groups of invertebrates that feed on scat, the target species genetic material is twice removed from its original host – from the target species to the scat and from the scat to the invertebrate. Here, the genetic material on the scat itself becomes degraded (Thuo et al., 2019; Vynne et al., 2012) and, once ingested by the invertebrate, these DNA fragments become further degraded as they are digested (Lee, Sing and Wilson, 2015). Therefore, iDNA genetic material is expected to be degraded and broken down into small fragments, more so that in most other genetic non-invasive methods. As an example, hair and scat samples are often used to reliably amplify microsatellite fragments from nucleic DNA or mitochondrial DNA (mtDNA) (common method for phylogenetic and lineage studies in bears (Bellemain et al., 2005; Pylidis et al., 2021; Taberlet and Bouvet, 1992)), but similar studies in primate species using fly iDNA found only 1% of mammal mtDNA in the sample was of good enough quality to genotype (Schubert et al., 2015).

The benefit of using mitochondrial markers in monitoring studies and DNA barcoding is that there are more initial copies of mtDNA in the sample (Hebert et al., 2003; Nelson, Wallman and Dowton, 2007; Ruppert, Kline and Rahman, 2019). Compared to a nuclear locus, like a microsatellite, which would have 2 copies per cells, a mitochondrial locus would have between 1000-2000 mitochondria per cell, and between 2-10 mitochondrial genome per mitochondria (Chinnery and Hudson, 2013; Shadel, 2008; Wiesner, Rüegg and Morano, 1992). As DNA degrades, it is more likely that a mtDNA fragment of sufficient length that can be used as a template for PCR will be present in the sample. eDNA/iDNA studies, therefore, focus on targeting mitochondrial DNA, to increase the chances of amplification, given a larger initial copy number. Moreover, mtDNA is maternally passed to the offspring, allowing for a minimal variation between individuals

of the same species (Barnes and Turner, 2015; Hutchison et al., 1974). Although differences in the mitochondrial DNA are more limited, research involving mitochondrial DNA has shown that it can reveal phylogenetic patterns between populations and has been used to show historical linkages and differentiate between genetically distinct populations (Johns and Avise, 1998; Meyer, 1994; Yu et al., 2007).

Another way to mitigate the issue of DNA degradation is to target short sequence amplicons which have been proven to be persistent enough with extraorganismal DNA to recover DNA from paleontological samples (Hofreiter et al., 2001). To increase the chances of detection, species-specific eDNA studies target fragments short enough to increase the chances of them being present in the sample, but enough to enable the use of fluorescent probe for PCR amplification detection that would maintain species specific detection. Advised by protocols from eDNA studies, the targeted amplicons usually range between 62 base pairs (bp) (Foote et al., 2012) and 650 bp (Deiner et al., 2015; Egan et al., 2013) long, with primer pairs of 18-30 nucleotides (nt) long and a probe that ranges between 20 and 30 nt (Hajibabaei et al., 2006; Langlois et al., 2021; Piggott, 2016). A vast majority of eDNA studies focus on amplicons of 150 bp or less (Piggott, 2016). Expecting similar degrees of DNA degradation, this study targeted fragments between 80 bp and 150 bp long. This size is a good compromise that allows for the existence of species-specific differences in the mtDNA fragment and the addition of a probe in the qPCR analysis, while also accounting for the fragmentation of genetic material prior to and after its ingestion by the invertebrate.

In comparing the different sampling techniques between eDNA and iDNA, the physical samples taken with iDNA can be much smaller in size and covering a larger area of the study site. With iDNA, invertebrates concentrate the DNA themselves, compared to eDNA where concentration of DNA is achieved by filtering larger volumes of water/air or processing of larger volumes of soil. Furthermore, water sampling for eDNA, depending on the flow of the water body, the samples can represent a very large or much smaller area (Hunter et al., 2019; Milhau et al., 2021; Stoeckle et al., 2017). Similarly, with iDNA the sample coverage depends on the group of invertebrates used and dispersal behaviour of that invertebrate as that widely varies across the target invertebrate groups in iDNA studies (Calvignac-Spencer et al., 2013; Cutajar and Rowley,

2020; Ji et al., 2020; Lynggaard et al., 2021). Moreover, each invertebrate is a small sampling unit and thus individual iDNA samples are often processed separately (Schnell et al., 2018). In studies where individual invertebrate extractions are not necessary, it is more cost-effective to pool iDNA samples together. The final part of the laboratory optimisation examines the threshold in sample pooling where target species detection is compromised. The results can inform sample collection and pooling for iDNA studies and provide an inside into the sensitivity of qPCR in detecting the target species in very low concentrations.

3.1.2 Chapter aims

This chapter aims to explore the potential of using invertebrate-derived DNA (iDNA) as an alternative or complementary survey method for single-species monitoring, using *U. arctos* as a case study species. I describe the development, testing and optimisation of a real-time PCR iDNA assay *in silico*, *in vitro* and in laboratory and field trials. The field application demonstrates the technique's capacity to monitor a rare species in the wild and the effectiveness of iDNA in single-species detection.

Specifically, the chapter aims to:

1) To refine and test the iDNA method for single-species detection using brown bears as a case study target species.
2) To test the qPCR sensitivity in detecting target species DNA in diluted samples.
3) To estimate the period within which target species DNA remains amplifiable in a fly's digestive system.
4) To utilise field samples to explore the effectiveness of the method as a monitoring tool.

3.2 Methods

3.2.1 DNA extraction

DNA extraction from arthropods can be challenging due to their hard exoskeletons. This protective covering made of chitin can be more difficult to break down in the cell lysis stage of the extraction, therefore studies often homogenise the samples with cryogenic grinding using liquid nitrogen (Asghar et al., 2015; Post, Flook and Millest, 1993). Here, liquid nitrogen was used to cryogrind the iDNA samples into a fine powder. The samples were then processed using Quick-DNA Miniprep (D3024), a spin column-based kit, as per manufacturer's instructions (Zymo Research). The kit is designed to extract genetic material from up to 25 mg of total sample weight per extraction, so using cryogrinding allowed for the samples to be well homogenised prior to transferring the specified amount into a spin column. To ensure the DNA extraction method yield was maximised, end-point PCR (PCRBIO HS Taq Mix, PCR Biosystems) was performed in variations of the manufacturer's instructions, including the elongation of the lysis and elusion stages of the extraction. End-point PCR using brown bear-specific primers was also performed on positive pilot field samples to confirm the increase in DNA yield when the samples were ground up prior to the extraction using liquid nitrogen (cryogrinding).

3.2.2 In vitro testing and optimisation

Design of primer and probe in silico and in vitro

Primers were designed and tested for the detection of brown bear DNA targeting regions between 80-120bp in the mitochondrial cytochrome b (cytb) gene. Firstly, a consensus sequence was create using the software MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms (Kumar, Stecher, Li, Knyaz, and Tamura 2018) by aligning complete cytb *U. arctos* FASTA sequences from NCBI (https://www.ncbi.nlm.nih.gov/nuccore). The resulting alignment was copied into the NCBI Primer Blast tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast) to create ten candidate primer sequences. Three of these primer pairs were selected based on their amplicon length and low self-complementarity (no larger than a value of 5.0 globally or only 3.0 at 3' end). Additionally, a 500 bp sequence covering all potential primers was selected and ordered as a custom synthetic gene from IDT (gBlocks® Gene Fragments,

Integrated DNA Technologies, Inc.; full sequence included in the Appendix B.1. Synthetic DNA sequence section) to use as a reference positive control sample for the primer selection and as standard for the qPCR analysis. The selection of the final set of primers was made using end-point PCR (1.2% agarose TAE gel) for a visual examination of amplification products. The final primer pair was further examined using a melt curve qPCR analysis (qPCRBIO SyGreen Mix, PCR Biosystems) which tested the that the primer pair to ensure the amplification of a single amplicon, as outlined in the MIQE protocol (Bustin et al., 2009). The primer and assays were also tested against large carnivores in the area (Eurasian wolf, *Canis lupus lupus,* and Eurasian lynx, *Lynx lynx*) to ensure they do not match any region in candidate non target species, and are not likely to generate false positives.

Additionally, a probe within the selected amplicon was designed using the IDT Oligo Analyser™ Tool (https://eu.idtdna.com/pages/tools/oligoanalyzer) and was ordered as a custom PrimeTime qPCR Probe® (5' 6-FAM™/ ZEN™/3' IB®FQ) (Integrated DNA Technologies, Inc.). For the qPCR master-mix, qPCRBIO Probe Mix was used (PCR Biosystems), following manufacturer's protocol. The final primer/probe set targeted a 112 bp region in the brown bear mtDNA *cytb* gene and is detailed in Table 5.

Table 5. Amplified fragment and primer and probe sequences targeting a 112bp region in the brown bear mtDNA ctyb gene. The position of the primers and probe is underlined in the amplified sequence.

Amplified sequence

5'-
CATCGGTCACCCACATTTGCCGAGACGTTCACTACGGGTGAGTTATCCGATATGTACATGCAAATGGAGCCTCC
ATCTTCTTTATCTGCCTATTTATGCACGTAGGACGGGG-3'

Forward primer	**Reverse primer**	**Probe**
5'-CATCGGTCACCCACATTTGC-3'	5'-TTTATGCACGTAGGACGGGG-3'	5'-TCCGATATGTACATGCAAATGGAGCCTCCATCT-3'

The Mastermix for all qPCRs performed during this project consisted of 10 µl 2x qPCRBIO Probe Mix, 0.8 µl each of forward and reverse primer (10 µM), 0.4 µl of probe (10 µM), 0.4µl BSA (2 ng μl^{-1}), 5.6µl PCR grade H2O and 2 µl template DNA. The following conditions were programmed into the StepOne instrument: 1 cycle of 95°C for 2 minutes, and 40 cycles of 95°C for 5 seconds and 60°C for 20 seconds.

In addition to testing the primers and probe using the synthetic DNA fragment, DNA was extracted from three European brown bear faecal samples using the ZYMO Quick-DNA Miniprep kit according to manufacturer's instructions. Similarly, positive control fly samples from the lab-reared flies (extracted using ZYMO Quick-DNA Miniprep kit) were used to validate the amplification of brown bear DNA using the primers and probe set.

Further examination of the performance of the primer pair included an end-point PCR optimum temperature test and limit of detection analysis for the primers and probe set using synthetic DNA, as suggested in the MIQE protocol (Bustin et al., 2009). The Limit of Detection (LOD) was examined at an LOD_6 and $LOD_{95\%}$ level, to examine the sensitivity of the selected assay. The LOD is instrumental in understanding the lower threshold of copy numbers present in the sample sufficient for amplification(Bustin et al., 2009). Here, the LOD numbers explored looked at amplification with complete success (LOD_6) and with a reasonable certainty at a 95% probability ($LOD_{95\%}$) (Bustin et al., 2009; Weldon et al., 2020). The LOD_6 was calculated following protocol from published literature (Burns and Valdivia, 2008; Weldon et al., 2020), while the $LOD_{95\%}$ was computed in R using the POD script (Boenn, 2020).

Due to the nature of the samples, the extracted product of environmental DNA samples (eDNA, iDNA, airDNA) is expected carry over PCR inhibitors from the original sample and various studies have explored ways to mitigate inhibition from sample preservation to extraction and PCR preparation (Goldberg et al., 2016; Kumar, Eble and Gaither, 2020). To decrease inhibition in the samples, bovine serum albumin (BSA) and dimethyl sulfoxide (DMSO) are often introduced to the final PCR mastermix (Farell and Alexandre, 2012). Previous eDNA studies have used both reagents to improve DNA amplification (Díaz et al., 2020; Weldon et al., 2020; Wong, Nakao and Hyodo, 2020). This chapter explores the capacity and optimum quantity of BSA and DMSO in reducing inhibition iDNA samples, ensuring that the target species detection is not hindered by the presence of PCR inhibitors. To test this, BSA and DMSO, often used to boost amplification in environmental DNA samples (Díaz et al., 2020; Farell and Alexandre, 2012; Wong, Nakao and Hyodo, 2020), were tested for their added efficiency in improving DNA amplification in iDNA samples. A series of twelve master-mixes were prepared, using the two reagents

in different concentrations and maintaining the volume of qPCRBIO Probe Mix, primers and probe as per the manufacturer's protocol (PCR Biosystems), but altering the amount of water added to the reaction to account for the added BSA/DMSO (Table 6). The concentration of BSA used was 2mg/µl.

Table 6. Concentrations of all components of the twelve master-mixes (15µl reactions) for testing the benefits of adding BSA and DMSO in the qPCR master-mix for iDNA samples.

Master-mix	**qPCRBIO Probe (µl)**	**Primers (µl)**	**Probe (µl)**	**Water (µl)**	**BSA (µl)**	**DMSO (µl)**
Protocol	7.5	1.2	0.3	4.5	0	0
BSA (2%)	7.5	1.2	0.3	4.2	0.3	0
BSA (6%)	7.5	1.2	0.3	3.6	0.9	0
BSA (10%)	7.5	1.2	0.3	3	1.5	0
BSA (3%) + DMSO (1.25%)	7.5	1.2	0.3	3.675	0.45	0.375
BSA (3%) + DMSO (2.5%)	7.5	1.2	0.3	3.3	0.45	0.75
BSA (3%) + DMSO (5%)	7.5	1.2	0.3	2.55	0.45	1.5
DMSO (1.25%)	7.5	1.2	0.3	4.125	0	0.375
DMSO (2.5%)	7.5	1.2	0.3	3.75	0	0.75
DMSO (5%)	7.5	1.2	0.3	3.375	0	1.125
DMSO (7.5%)	7.5	1.2	0.3	2.25	0	2.25
DMSO (10%)	7.5	1.2	0.3	1.5	0	3

Control fly culture in the lab

Bluebottles (*Calliphora vomitoria*) were reared in fly rearing boxes in a laboratory setting. The flies were purchased as casters (pupa) and immediately placed in rearing boxes and raised following recommendations in Erzinçlioğlu (1996). The adult flies were kept with a constant supply of water and sugar by providing water-soaked cotton wool in a petri dish, and an additional petri dish with a water and sugar-soaked cotton wool (Erzincliogl u, 1996). Providing the flies with sugar and water as the sole source of food

ensured the absence of any other mammalian DNA in their intestinal tract prior to their exposure to target species DNA.

3.2.3 DNA retention in the fly gut

A time series experiment was carried out to examine the time period in which amplifiable target species DNA could be present in the sample. The experiment was carried out using the lab-reared colony described above and the protocol for this experiment was adapted from Lee (Lee, 2016). As positive control, three European brown bear scat samples were provided by the Welsh Mountain Zoo (Material Transfer Agreement reference: CE-17-0542). Following protocol from Lee (2016) all food sources were removed 24 hours prior to the experiments. Thirty flies were then removed from each colony and preserved in -20^{o}C to be used as a negative control (no target species DNA present).

Following the 24 hours past the removal of the sugar from the fly boxes, 20 g of bear faecal sample was introduced as the new source of food into each colony of flies. The blowflies were observed during that period to ensure they fed on the scat. As outlined in Figure 17, the fly colonies were provided with a brown bear scat sample for the duration of four hours. Subsequently, the scat samples were removed and replaced with sugar and water to provide a constant food source for the duration of the experiment. Ten flies from each rearing box were selected at random at the following intervals: 0h, +4h, +8h, +12h, +24h, +48h, +72h, +96h, +120h, +144h, +168h, following the removal of the scat samples. The fly samples from each colony were stored in -20^{o}C until further sample processing (cryogenic grinding and DNA extraction).

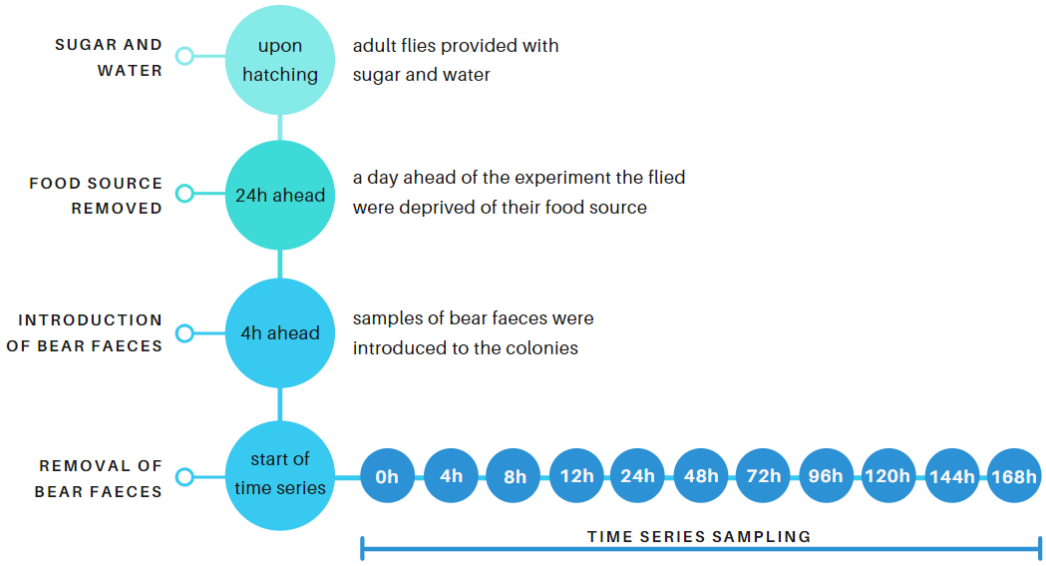


Figure 17. Time series experiment to examine the persistence of target species DNA in iDNA samples.

All flies were euthanised with 96% ethanol and preserved at -20°C (Post, Flook and Millest, 1993). The experiment was repeated twice using two different starting colonies, with each experiment using a total of 420 flies (flies per experiment run/rearing cage: 140 (30 negative controls and 10 post-feeding samples * 11 time points). The extraction and qPCR analysis follow the methodology described above. Experiments were carried out in May and June 2019. The welfare and ethical aspects of this experiment were reviewed by the University of the West of England's Animal Welfare and Ethics Committee (AWEC) (Ethics Approval document reference number: R34).

3.2.4 Target species DNA dilution in pooling of fly samples

A dilution experiment was carried out to test the reliability and limits of sample pooling in detecting rare DNA fragments in iDNA samples. Confirmed positive fly samples (T_0) from the three lab-reared colonies were pooled and macerated using liquid nitrogen. The process was repeated for control negative flies collected from the three colonies before the introduction of bear scat as food source. A dilution series was prepared for DNA extraction. Accounting for the 25 mg dry sample weight limit of the DNA extraction kit, eight dilutions were prepared, starting with a pure positive sample (pooled sample of T_0 positive flies), to a 1:50 positive to negative fly sample (Figure 18). Three extractions were performed for each dilution to test the repeatability in detecting small

traces of bear DNA in pooled iDNA samples. The qPCR included three replicates of each extraction, resulting in nine replicates for each dilution (3 sample replicates x 3 extraction replicates = 9 replicates for each dilution factor).

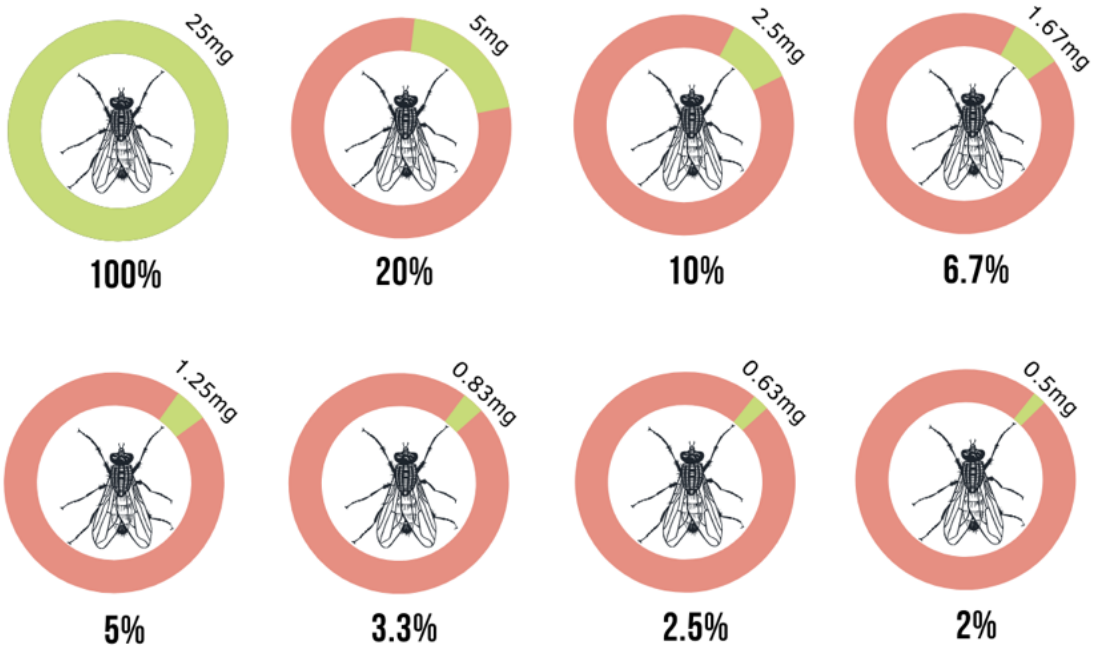


Figure 18. Fly pooling experiment combinations (green=positive bear DNA samples, red=negative control). Percentage of positive control in total sample and amount of positive control of ground material per extraction (total amount per sample: 25mg).

The qPCR was performed following the tested protocol from the sections above, using the qPCRBIO Probe Mix and synthetic bear gene as reaction standards. This experiment tested the presence of bear DNA in the samples and quantified the amount present in each dilution. A Binomial Linear Model looked for the indication of a dilution threshold after which the amplification of target species DNA could no longer be reliably traced in the sample.

3.2.5 Field experiments

Study area

The previous chapter focused on creating a picture of the habitat suitability and brown bear distribution in Greece. The Expert Knowledge models were the first habitat suitability assessments generated in this study and thus constituted the backbone of fieldwork survey planning. Specifically, a model taking topography, land use and human

impact into consideration was used here to guide field survey planning and select a smaller case study area within the greater analysis area to carry out the field experiments. Using the suitability map as a guide reflecting the suitability mosaic across the country, a smaller study site was drawn to include areas of low-medium-high suitability (Figure 19). Looking at our current understanding of brown bear distribution across Greece, the study site incorporates two core areas, a long-established population on the left, the Mt. Verno (left) habitat patch, as well as the relatively recently re-established Mt. Voras (right) habitat patch (population previously persecuted but recovering in the last 100 years) (Kaczensky et al., 2012b; Mertzanis, 1994). The population of Mt. Verno is well-monitored and belongs to the western brown bear population, the Greek part of the Dinaric-Pindos biological population (Pylidis et al., 2021; Tsaparis et al., 2014). The Mt. Voras population has been much less frequently monitored, with evidence of a small population of bears (Mertzanis, 1994; Savvantoglou, 2015). Both core areas are mountainous with the largest proportion of the land covered by deciduous, mixed and evergreen forests, while the connecting area is predominantly covered by grazing land and farmland with scattered villages and industrial activity, such as quarries.

Sample collection

The study area was split into 2 km x 2 km grid squares and ten grids from each predicted suitability type (high-medium-low, N = 30) were randomly selected for the field surveys (Figure 4). The cell size (2 x 2 km) was selected to reflect the average distance covered by a blowfly within which amplifiable DNA could still be tracked in their digestive tracts (Lee, 2016). The sampling sites were surveyed three times, with each sampling season taking place between August and September in the years 2017-2019, collecting a total of ninety iDNA samples.

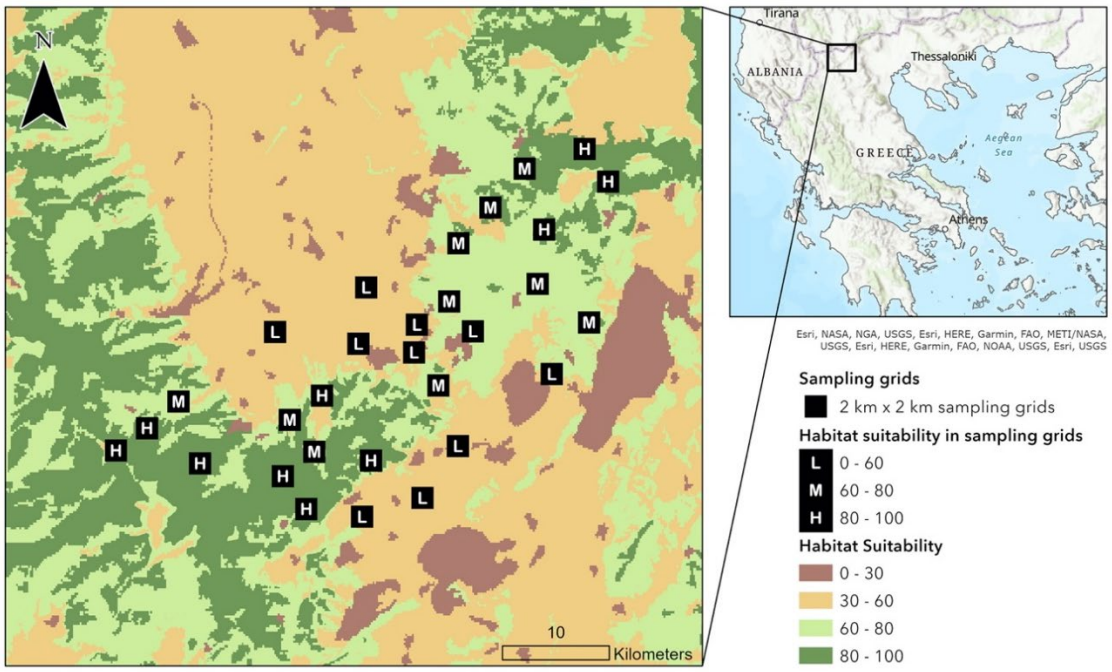


Figure 19. Field survey locations showing the sampling grids and corresponding mean suitability within each grid. The habitat suitability model was created using topographic, land use and human impact variables after an extensive literature review of U. arctos ecological requirements, detailed in Chapter 2.

Flies were caught using a modified mosquito net (Figure 20) and commercial fly bait (RedTop Flycatchers). The traps were active for one hour per sample point whereupon the traps were closed and all flies present in the trap were collected. Sets of 20-50 flies were transferred into 96% ethanol for one hour (Schubert et al., 2015). These were subsequently drained of the ethanol and sample tubes were filled with silica beads and then stored and transported back to the UK in fine silica beads (Post, Flook and Millest, 1993). The samples were transferred back to the University of the West of England and analysed for the presence of brown bear mtDNA. The samples varied in the number of flies caught in the fly traps, from just four individuals to over three hundred within a single sampling session. On sites where less than 40 flies were collected, the second sample was comprised of the remaining flies left after the first twenty were pooled together to allow for the processing of two separate iDNA extractions from each sampling session.

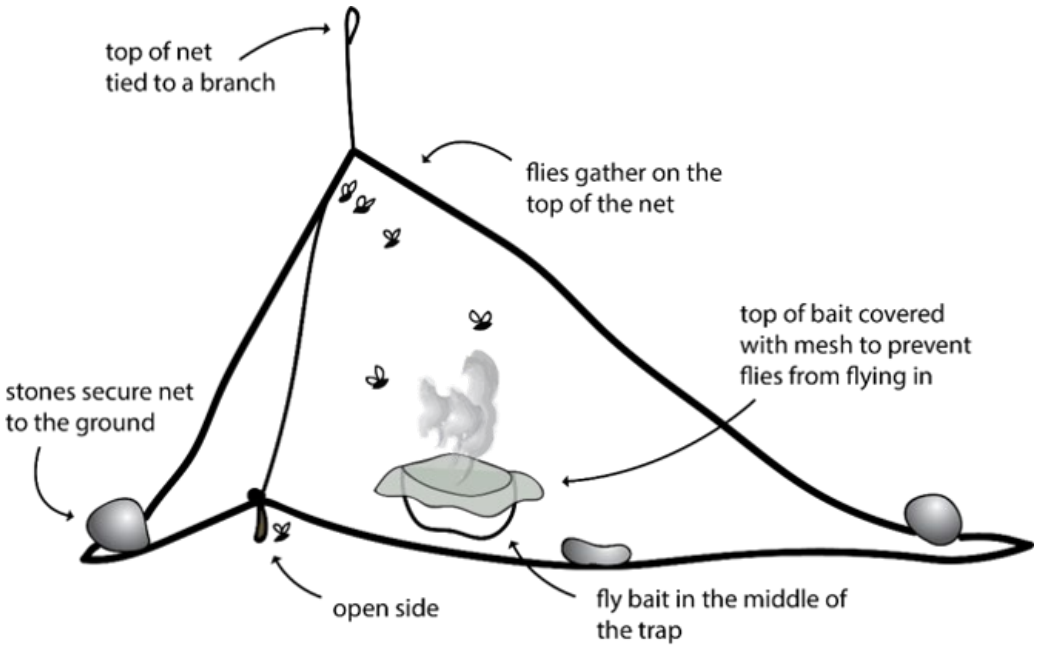


Figure 20. Mosquito net and bait setup, used for the iDNA sampling.

Field sample processing

Samples were processed following the techniques determined by the laboratory optimisation and two extractions of each sample pool were conducted to increase detection chances (30 sampling stations, 3 survey repeats, 2 extraction repeats for each survey). Each extraction was processed with six qPCR repeats (after Ficetola et al., 2015) and, given the fact that there were two extraction replicates for each sample, each survey sample was run under qPCR 12 times. A GLM analysis was performed to compare the effect of the number of positive amplifications in the CT values and DNA concentration in the samples.

Finally, 35% of the positive amplifications were Sanger sequenced in both directions. The resulting sequences were analysed using the 'sangerseqR' in R Studio (Hill and Demarest, 2014) and checked against the NCBI database to confirm that the primer/probe set was targeting the desired amplicon.

3.2.6 Data analysis

All qPCR analyses were performed in a StepOne™ Real Time PCR System (Applied Biosystems) instrument and the results were exported from the StepOne™ (Applied Biosystems) software as excel files. All data was analysed in R Studio (R Core Team, 2021) using 'stats', 'LOD' and 'sangerseqR' as the main data analysis packages. An overview of the methodology skeleton outlined in this chapter is shown on Figure 21.

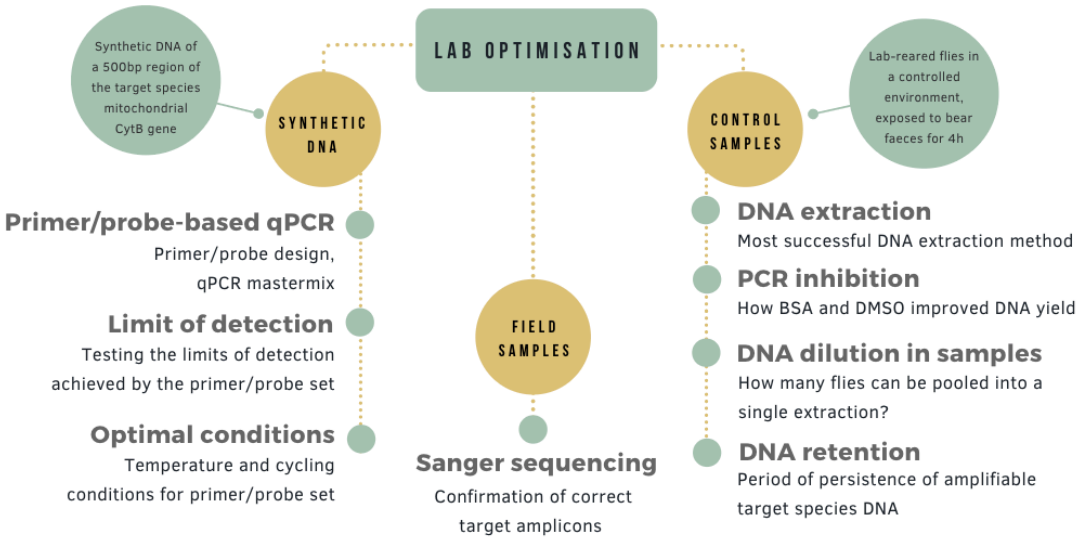


Figure 21. Overview of laboratory optimisation and field sample analysis

3.3 Results

3.3.1 DNA extractions and in vitro testing of primers and probe

A series of pilot experiments were carried out to optimise the DNA extraction, primer performance and qPCR setup. Firstly, it was confirmed that cryogrinding using liquid nitrogen in combinations with the ZYMO Quick-DNA Miniprep Kit (D3024) gave the most consistent results in DNA extraction in comparison to no pre-processing of samples or other DNA extraction kits. End-point agarose gels and a qPCR melting curve confirmed that the amplicon of the primer pair was a single product and the recommended annealing temperature in the qPCRBIO Probe Mix (PCRBiosystems, UK) protocol performed well for the primer pair and probe, so no changes in the qPCR reaction setup were made.

The dataset from the LOD qPCR for this specific primer/probe set suggested a drop in amplification in DNA concentrations under 1.02 x 10^{-8} ng μl^{-1}, suggesting that this concentration represents the LOD_6. The $LOD_{95\%}$ was computed in R using the POD R package (Boenn, 2020) as $LOD_{95\%}$ 2.15 x 10^{-9} ng μl^{-1} [95% CI 1.49 x 10^{-9} - 2.05 x 10^{-9}]. As expected, a linear relationship between the DNA concentration of the standards and the CT values, suggesting that the smaller the DNA concentration in the sample, the longer it takes for the amplified DNA to reach the amplification threshold (Figure 22). A Generalised Linear Model (GLM) was performed in R Studio (R Core Team, 2021) looking at the power (r-squared) of DNA concentration to explain the change in CT values. The DNA concentration datasets across all experiments was transformed logarithmically (log(concentration)) to reduce residuals and normalise the data (Sokal and Rohlf, 2009)). The GLM confirmed a negative relationship whereby the increase in DNA concentration decreases the CT values by 3.21% (R-squared: 0.9818; $F_{1,46}$ = 2531, p <0.001).

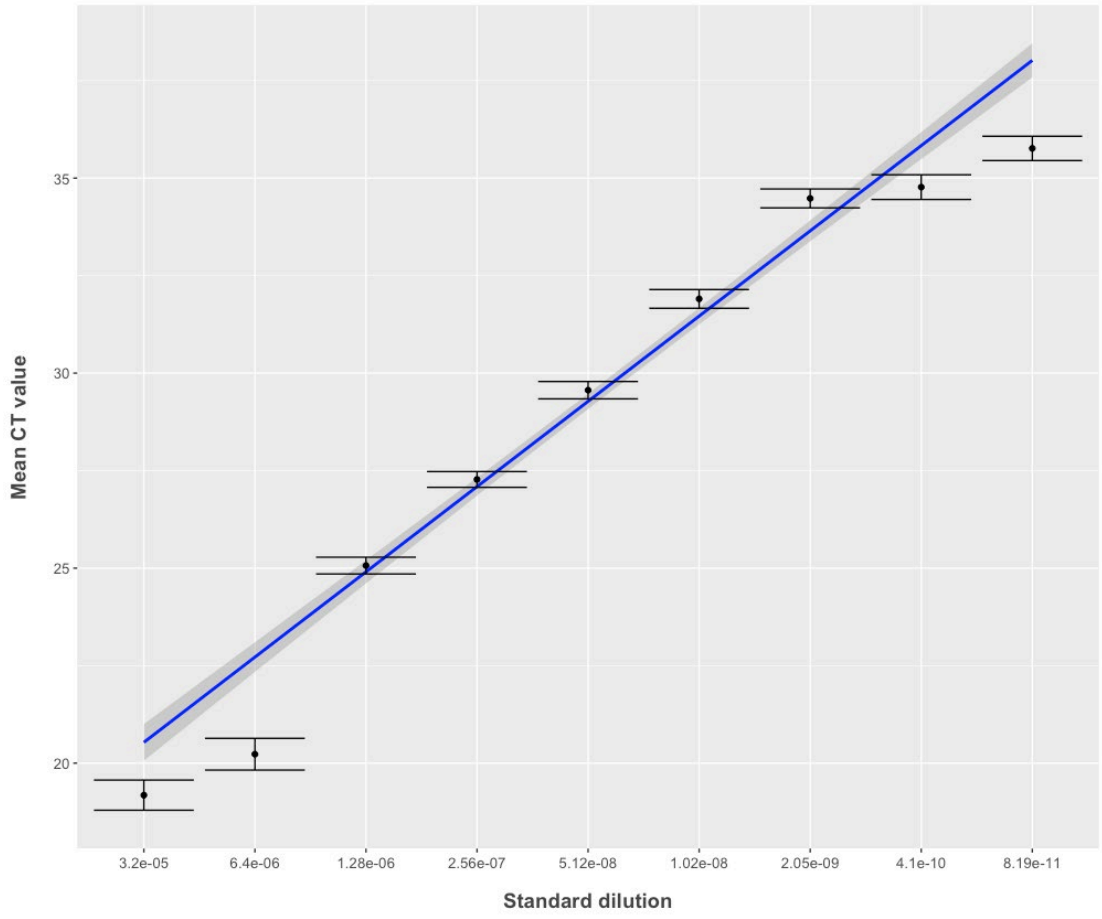


Figure 22. Increase in CT values as standards become more dilute. Blue line: fit of the GLM; grey shaded area: 95% confidence limits of the model.

3.3.2 BSA/DMSO and final qPCR Mastermix

The presence of BSA and DMSO in the qPCR mastermix did not follow the expected beneficial impact on the reaction. In the case of these iDNA samples, DMSO, used on its own or in combination with BSA, increased the CT values and in some cases completely blocked amplification. BSA, even though it showed a very slight increase in CT values at the higher concentrations, it marginally decreased the CT value at its lowest concentration (2% BSA) (Table 7). BSA (2%) was added in the final Mastermix, as it is also proven by other studies to positively impact the reactions (Díaz et al., 2020; Weldon et al., 2020; Wong, Nakao and Hyodo, 2020).

Table 7. BSA and DMSO effect on qPCR amplification.

Sample	C_T Median	C_T SD	Amplification (out of 6)
BSA (2%)	29.23	0.44	6
BSA (6%)	29.42	0.30	5
BSA (10%)	29.53	0.35	6
BSA (3%) + DMSO (1.25%)	29.86	0.46	6
BSA (3%) + DMSO (2.5%)	29.83	0.75	6
BSA (3%) + DMSO (5%)	29.79	0.49	1
DMSO (1.25%)	30.30	0.00	6
DMSO (2.5%)	29.15	0.26	6
DMSO (5%)	29.90	0.49	6
DMSO (7.5%)	0	0	0
DMSO (10%)	0	0	0
Protocol control	29.26	0.31	6

The following manufacturer's instructions for a 20µl reaction using the qPCRBIO Probe Mix (PCR Biosystems), with the addition of 2% BSA. The final Mastermix is detailed in the Methods section.

3.3.3 Target species DNA dilution in pooling of fly samples

The dilution series was examined both by the number of amplifications each dilution produced, as well as the amount of DNA that was amplified (in ng μl^{-1}). The samples that were not diluted (T_0 positive controls) were all amplified, while the dilutions from 1:5

(20% positive control in samples) to 1:20 (5% positive control in sample) had all similar success rates (6-7 amplifications out of 9 repeats). The amplification success was negligible at the 1:50 dilution samples (2% positive control in the sample), with only 3 out of the 9 samples successfully amplifying (note: these three amplifications were from three separate extractions of that dilution, so the 'per sample' amplification for the 2% samples were 1 out of three). A binary logistic regression showed a significant relationship of lower amplification with the more dilute the samples (β (estimate) = 2.5471, SE = 1.0415; Wald = 2.445, p = 0.0145; Figure 23).

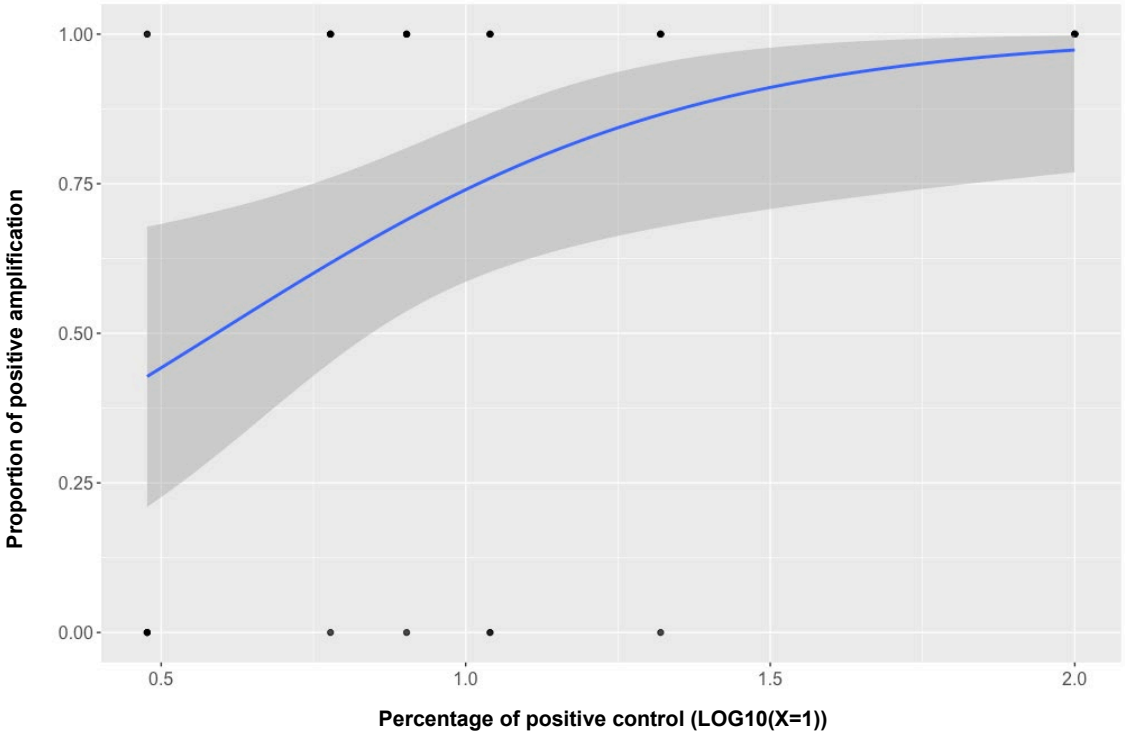


Figure 23. The effect of diluting an iDNA positive control sample. Percentage of positive control transformed using a logarithmic formula (Log10(X+1)), showing the more diluted samples on the left. Blue line: fit of the binomial logistic regression curve; grey shaded area: 95% confidence limits of the model. .

3.3.4 DNA retention in the fly gut

The Binomial Linear Model revealed a pattern of decrease in detection probability with time, reflected also in the decrease of DNA concentration and increasingly larger mean CT values as time passed. A significant positive relationship was suggested with time post-feeding (β (estimate) = -0.05285, SE = 0.01087, Wald = -4.862, $p < 0.001$)) and CT_{mean} (β (estimate) = -0. 1139, SE = 0.0208, Wald = 5.476, $p < 0.001$)) values, as well as a negative relationship between detections and DNA concentration (β (estimate) = -

0.2982, SE = 0.1127, Wald = -2.646, p < 0.01)). The regression line fitted on the data illustrates that quick drop in detection probability within the first 24h (Figure 24).

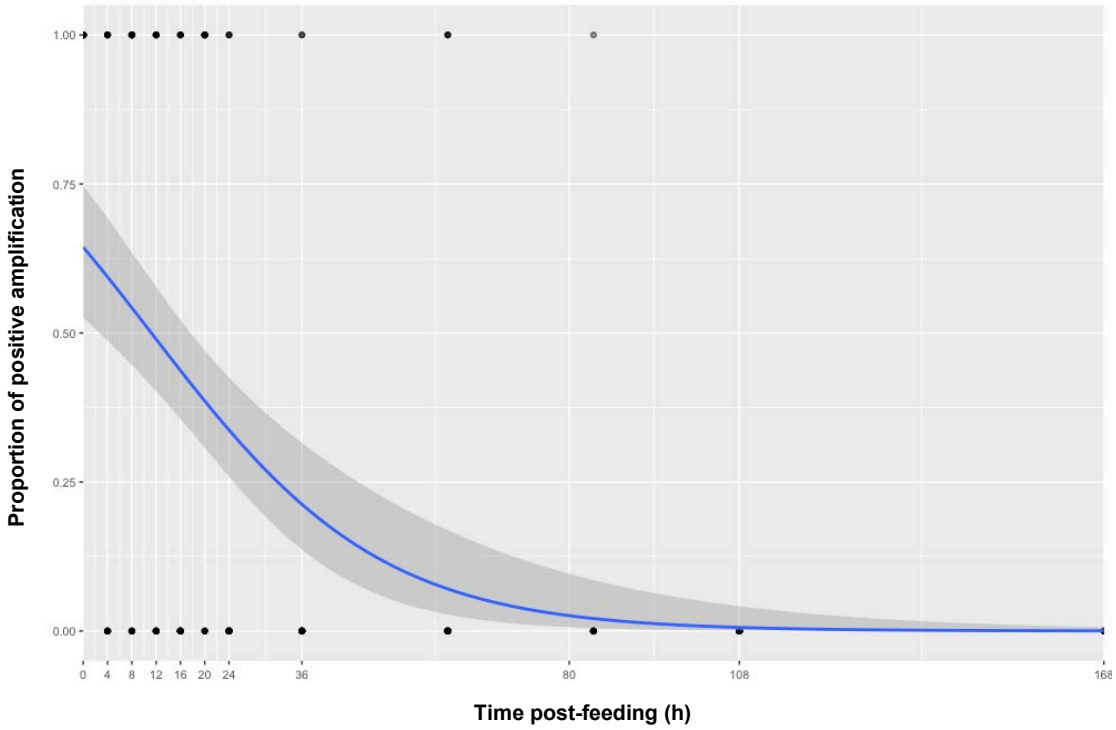


Figure 24. Reduction in target species' DNA amplification success as a result of time post-feeding the fly colonies with a source of target species DNA. Blue line: binomial logistic regression curve; grey shaded area: 95% confidence limits of the model.

A GLM examined the relationship between time points and the LOG concentration of DNA in the samples revealed a pattern of rapid decay in the amplifiable DNA found in the sample (Estimate -1.627, Std. Error: 3.193E-08, t value: -0.510, Pr(>|t|): 0.6122)). The CT values also reflected this pattern (Estimate 2.778, Std. Error: 1.427, t value: 1.946, Pr(>|t|): 0.0561).

Furthermore, the impact that time had on the CT values of the samples evaluated with a one-way ANOVA test suggested significant differences between the time points (ANOVA, F(9, 55) = 3.53, p = 0.002, η^2_g = 0.37). Tukey post-hoc analyses revealed significant differences between 0h and time points 8h (p = 0.04), 24h (p = 0.026), 60h (p = 0.045) (Figure 25). A GLM regression line illustrating the increase and eventual plateau in CT values as time post feeding passed is shown on Figure 26. Finally, a regression line showing the effect of time on DNA amplification (Figure 25) depicts the sharp drop in detection within the first 24h post-feeding. After period of 20h, detection dropped to

67%, with detection dropping below 20% shortly after. However, mtDNA was still detected (a single amplification in 3 experimental repeats of 6 qPCR repeats each (1 out of 18)) after 84 hours.

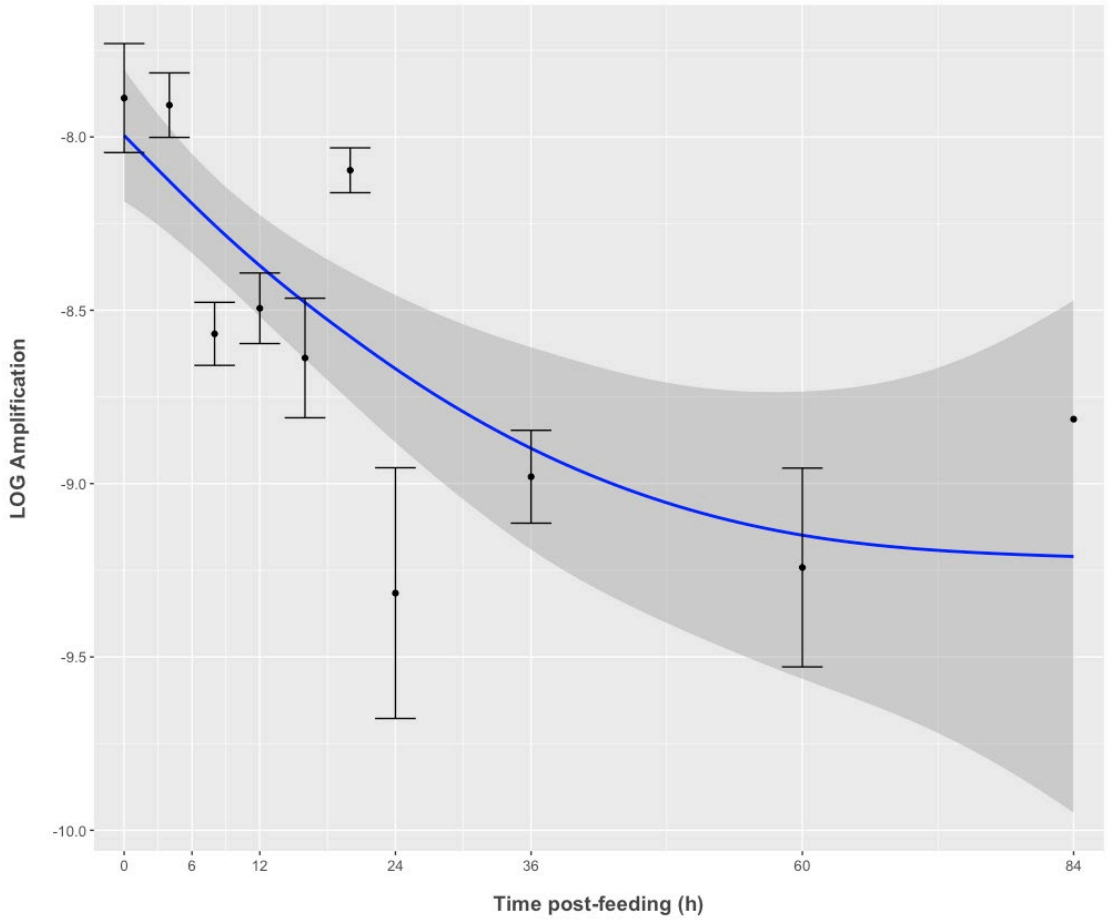


Figure 25. Mean DNA concentration and SD bars for each time point where target DNA was detected, looking at the persistence period of amplifiable brown bear mtDNA in control fly colonies. Blue line: fit of 95% CI; grey shaded area: 95% confidence limits of the model.

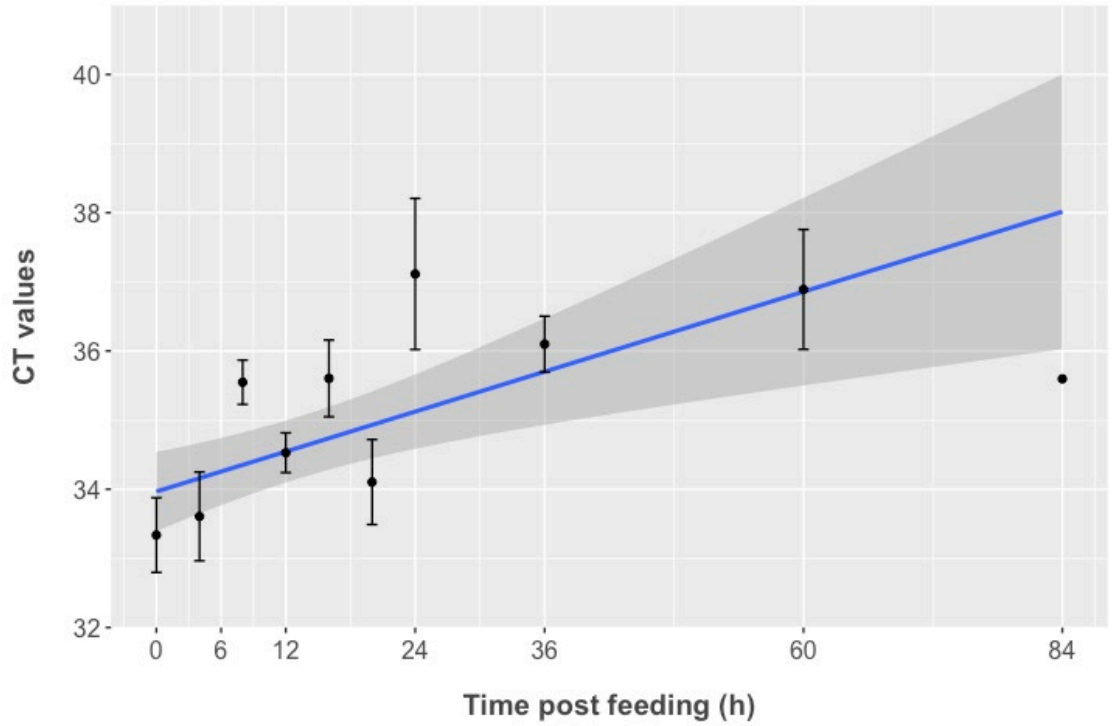


Figure 26. Mean CT value and SD bars for each time point where target DNA was detected, looking at the persistence period of amplifiable brown bear mtDNA in control fly colonies. Blue line: fit of 95% CI; grey shaded area: 95% confidence limits of the model.

3.3.5 Field data analysis

In total, *U. arctos* DNA was detected in 28 out of the 90 field samples with both extraction repeats detecting target species' DNA on over half of the occasions (16 out of 28). In total, adding a second extraction replicate added six more detections (three in the first season, one in the second and two in the last one), increasing overall detections by 21%. Bears were detected at seventeen of the thirty sample locations over the three seasons (Figure 27).

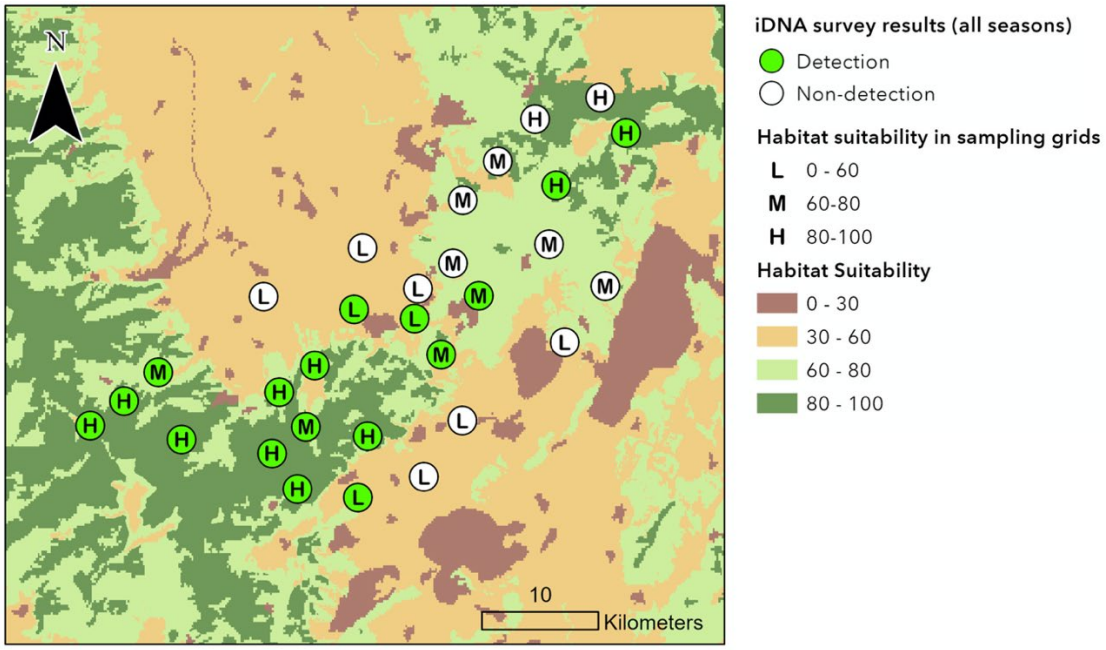


Figure 27. Field iDNA survey results from all seasons. The letters on the survey results points describe the habitat suitability score within the sampling grid. The habitat suitability model was created using topographic, land use and human impact variables after an extensive literature review of U. arctos ecological requirements, detailed in Chapter 2.

An observed positive relationship between the number of positive amplifications and DNA concentration was not statistically significant. A significant negative relationship was observed between the number of positive amplifications in the qPCR replicates and C_T value (GLM: $F_{2,41} = 22.81$, $p < 0.001$, $r^2 = 0.5036$; Figure 28).

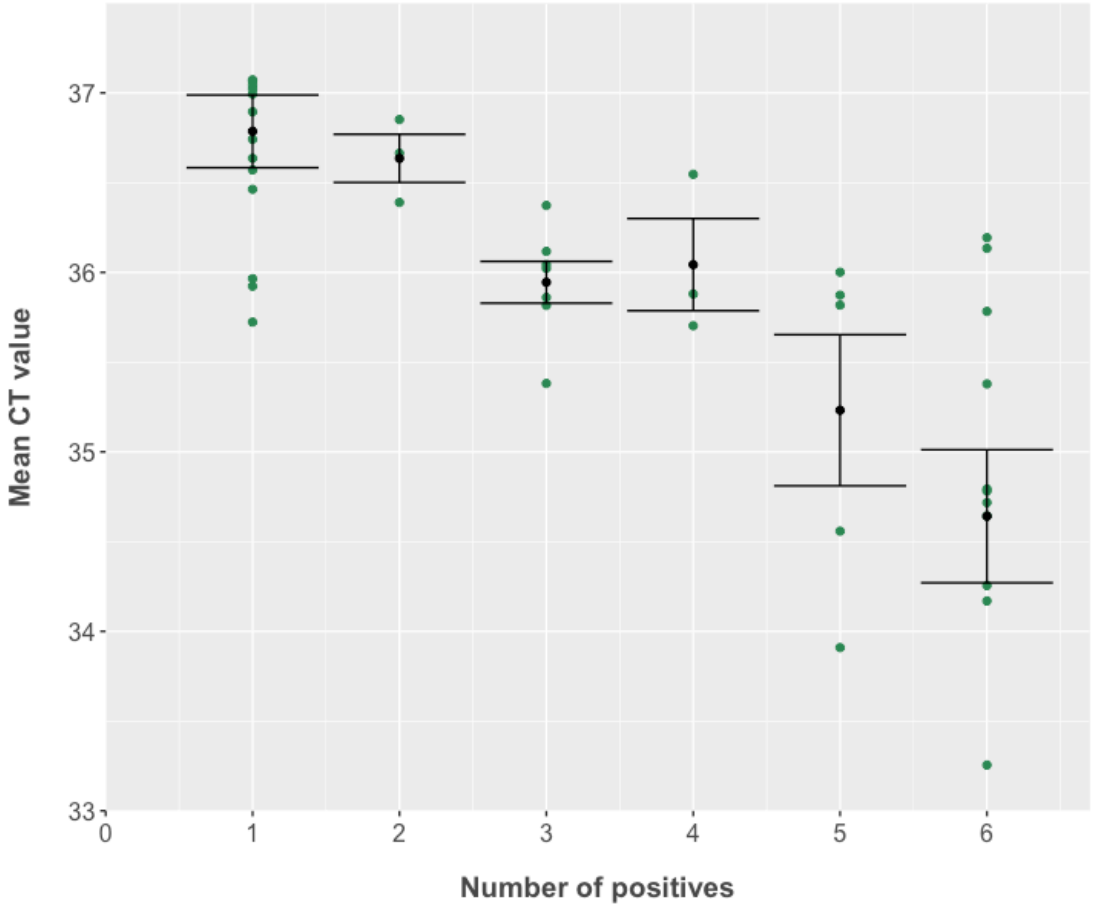


Figure 28. Response of CT values to the number of qPCR amplifications (out of six qPCR replicates per sample). The more positive amplifications within a sample, the lower the mean CT value. Blue line: fit of the GLM; grey shaded area: 95% confidence limits of the model.

3.3.6 Sanger sequencing

Twelve samples (over 35% of positive amplifications in all three seasons and extraction replicates) were sent for confirmatory sequencing. A BLAST (Morgulis et al., 2008; Zhang et al., 2004) of the primary sequence confirmed a 98% match to *U. arctos* (E value range: 3e-09 to 1e-07, complete list in the B.2. Sanger results BLAST section in the Appendix).

3.4 Discussion

This chapter demonstrates the development of an iDNA-based method for single-species detection and presents a cost-efficient protocol for the collection, storage, processing and analysis of the invertebrate samples. The results will be discussed below in detail, but a summary of the protocol developed here is presented in Figure 29 (note: carrion fly dispersal distance estimates as per Lee (2016)).

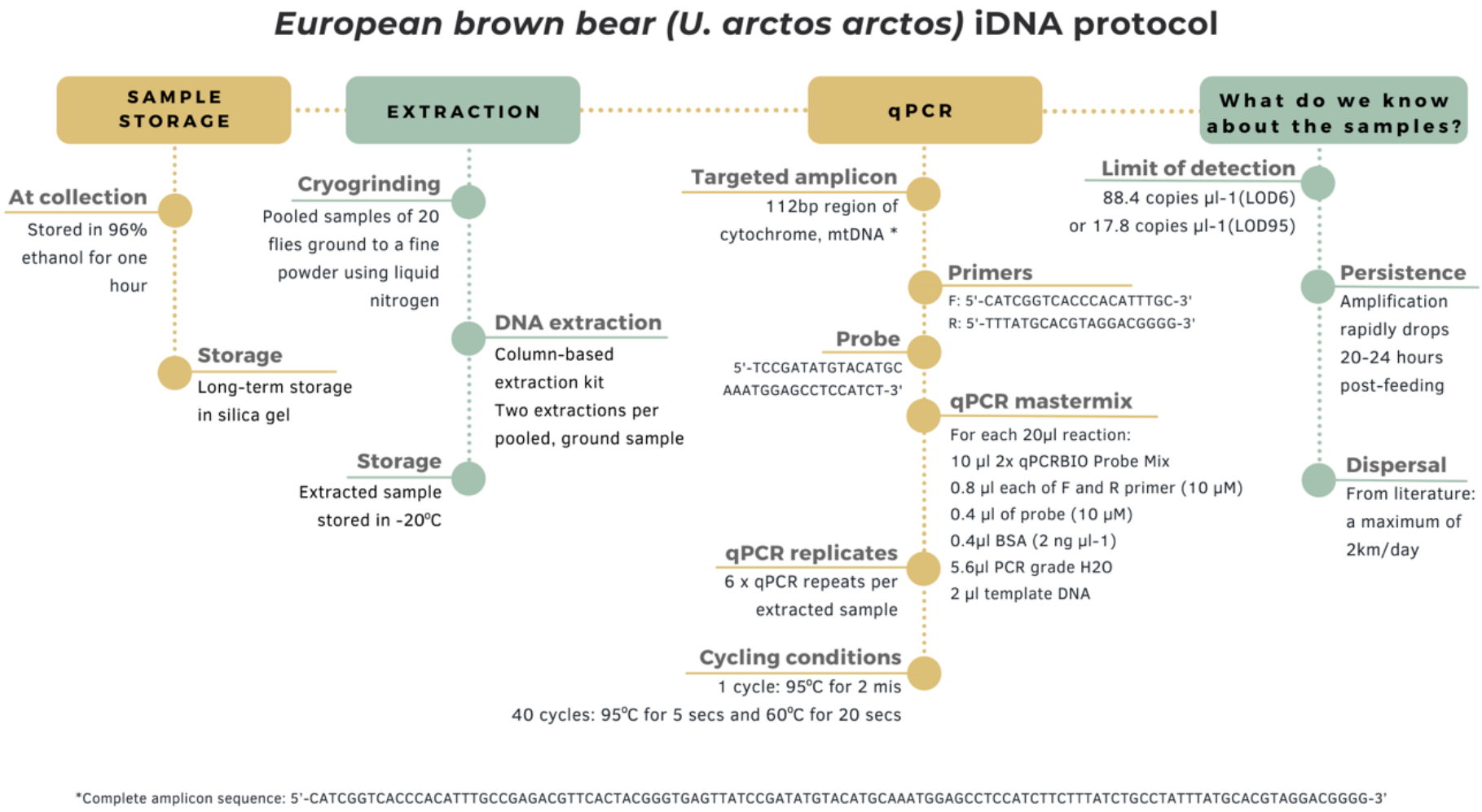


Figure 29. Protocol developed as a result of this work. Fly dispersal distances for carrion flies as seen in literature (Lee, 2016).

The laboratory experiments outlined here demonstrate the sensitivity, as well as some of the limitations, of iDNA as a monitoring method. This chapter is, to our knowledge, the first study to focus solely on single-species detection using iDNA, both in the refinement of the method in a laboratory setting, as well as field surveys. The introduction of this method opens new possibilities for bear monitoring, complementary to the commonly used methods, such as hair trapping and scat surveying. Monitoring bear species can be very challenging in different parts of the world, especially when it comes to bears in the tropics due to climatic conditions and the more elusive behaviour of tropical bear species (Tee et al., 2016; Tee et al., 2020). We believe that this study provides evidence of the potential of iDNA as a complimentary sampling technique for bear monitoring. The supporting laboratory experiments have allowed for the refinement of this method and understanding of the advantages and limitations involved. The aim of this study is also to serve as a template for the design of sampling and laboratory analysis of iDNA using flies for monitoring bears located in different part of the world, or targeting other invertebrate groups in efforts to detect rare or elusive animals.

3.4.1 Dung flies as a target invertebrate group

Flies (here predominantly, but not solely, represented by Sarcophagidae, Calliphoridae, Muscidae) were chosen as the most suitable invertebrate group for creating a transferable protocol based on a variety of criteria, such as distribution across the specific study site; ease of capture; and the ability to carry out control laboratory experiments. Firstly, they were abundant at the field site in comparison to other invertebrates present in the area such as mosquitoes and dung beetles. Fly trapping using mosquito nets and canopy traps was simple and efficient in terms of costs, time and surveyor effort (Calvignac-Spencer et al., 2013; Hoffmann et al., 2018). Additionally, climate differences in elevation, terrain and land cover type seemed to restrict the distribution of some of the other invertebrates but flies were found throughout the study area. Moreover, flies were straight forward to rear in a laboratory environment (Erzinçlioğlu, 1996) which allowed for the fine-tuning of the method prior to the field sampling and field sample processing, following similar efforts for iDNA monitoring in the tropics (Lee, 2016; Lee, Sing and Wilson, 2015). Finally, in regards to field collections and the potential use of this data for further analysis, such as occupancy modelling, fly

dispersal rates and preservation of DNA studies (here, as well as in published literature, e.g. Lee, Sing and Wilson, 2015; Lee, 2016) indicate that a closure assumption for the iDNA samples can be maintained. Contrastingly, Schnell et al. (2015) explore the use of leeches in yielding results for occupancy modelling and highlight the potential violation of population closure of such efforts due to the extent of time within which target species DNA remains amplifiable in leeches as blood meal.

In the case of flies used for monitoring mammals, the genetic material, in most cases, is derived by target species' faecal samples present on site. Flies of the Calliphoridae, Sarcophagidae and Muscidae families (mostly targeted in this study, although the field sampling did not select against other families to avoid bias associated with the potential presence of specialists) are most often targeted by iDNA studies (Calvignac-Spencer et al., 2013; Gogarten et al., 2020; Hoffmann et al., 2018; Schubert et al., 2015).

To understand the state of DNA fragmentation expected in the samples, it was assumed here that most of the DNA found in the samples would be twice removed from the bear it originated from (from the bear to the scat and from the scat to the fly). Indeed, the predicted larger degree of fragmentation within the sample was reflected in the DNA persistence experiment, both in this study and in literature (Drinkwater et al., 2021; Lee, 2016; Lee, Sing and Wilson, 2015), making it essential for this study to focus on short amplicons. A controlled experiment in the tropics using mtDNA mini-barcode target (205bp) found that the period within which mtDNA in blowfly guts remains amplifiable for is 24-96h, with the rate of detection using end-point PCR dropping to 22% for the latter (Lee, Sing and Wilson, 2015). Our replication of this experiment in a temperate climate using qPCR revealed a similar persistence period but lower rates of detection. It is worth considering a number of reasons that might have contributed to this lower success rate, including a heat wave with an uncharacteristic temperature spike during the experiment (over 30^{o}C); and the age and condition of the scat samples used to feed the lab-reared flies. The high temperatures in eDNA studies have shown to increase the rate of DNA degradation (Kasai et al., 2020) and it is possible that the uncharacteristically high temperature in the laboratory had similar effects on this experiment. The scat was collected fresh by the Welsh Mountain Zoo staff, but freeze-thawing the samples during transportation might have contributed to a level of DNA degradation prior to ingestion.

Furthermore, bacterial activities could start again once the samples are defrosted, fragmenting the DNA even further (Fouhy et al., 2015; Pérez-Burillo et al., 2021). In any case, this experiment demonstrates a drop in amplification success between the 20h and 24h time points, from over 44% (8 amplification out of 18 positive samples) to just under 17% (3 amplifications out of 18 positive samples), suggesting that the target species mtDNA in fly samples is most likely no older than 24 hours.

3.4.2 Brown bears as a target species

In selecting the most appropriate study species for single-species monitoring, this study considered a number of factors that would contribute the most amount of information. Firstly, bears were chosen because of their elusive nature, often avoiding the presence of humans and thus more challenging to detect. Moreover, bears roam large ranges and are territorial animals which, in this study, translates to a decreased rare of detection due a low species density within a given study area. However, bear presence is often monitored with the presence of scat, so iDNA monitoring could be a suitable alternative. Another factor considered here was to select a species with a clear genetic differentiation at an mtDNA level that did not compromise the reliability of the results. Wolves (*Canis lupus*), otherwise a suitable candidate due to their distribution and relative abundance, were not considered due to the significant mtDNA overlap with domestic dogs (*C. lupus familiaris*). Finally, this study aimed to explore the sensitivity and limits of iDNA in single-species monitoring, so a range of habitats where the animal was expected to use at different levels was desired. Using the habitat suitability modelling outlined in the previous chapter, we were able to locate areas with expected high/low/medium use advised by the suitability score. The study area, described in more detail in the following chapter, comprised of a gradient of suitability, allowing the testing of this method to be tested for its effectiveness in detecting bears in areas of low to high use. We believe that the results of the field study clearly demonstrate that iDNA was effective at detecting bears in the field and, to our knowledge, introducing this method as the first study solely focusing on single-species detection.

3.4.3 Target DNA dilution and its repercussions on its detection using qPCR

Single-species qPCR approaches in the literature have shown improved detection compared to other methods, such as end-point PCR or metabarcoding (Carvalho et al., 2021; Furlan et al., 2016). This may be especially useful when monitoring rare or elusive species. In this case, bears were considered rare, due to their low numbers and territorial nature (Dahle, Støen and Swenson, 2006; Mertzanis et al., 2005; Penteriani et al., 2018; Swenson, Sandegren and Soderberg, 1998; Tsaparis et al., 2014). This study explored the sensitivity of a single-species PCR approach by looking at the limit of detection for the chosen amplicon and understanding the dilution threshold for pooling samples together. The results of the LOD experiments highlight the sensitivity of this method, able to consistently detect (LOD_6) the mtDNA amplicon in concentrations as low as 1.02×10^{-8} ng μl^{-1} (88.4 copies μl^{-1}), with $LOD_{95} = 2.15 \times 10^{-9}$ ng μl^{-1} (17.8 copies μl^{-1}). Limit of detection in eDNA has been extensively studied and reported (e.g., Brys et al., 2021; Guan et al., 2019; Mauvisseau et al., 2019; Weldon et al., 2020) but, to our knowledge, this is the first iDNA study detailing LOD values and the effect of initial sample concentration to the detection of the target species. Results from eDNA LOD studies tend to be lower than the values reported here, with eDNA methods being sensitive enough to detect a 1-4 copies μl^{-1} (Guan et al., 2019; Weldon et al., 2020). The difference in detection limits here could be a result increased PCR inhibition (by comparison to eDNA samples) due to the nature of the iDNA samples. When extracted, the iDNA sample carries DNA not just from the target species (if present in the sample), but also the host's genetic material, any microorganism (such as gut microbiota) the invertebrate is a host of (Aksoy et al., 2014; Gupta et al., 2014), and the genetic material of non-targeted species the fly fed on (Calvignac-Spencer et al., 2013; Rodgers et al., 2017). The results in samples that are challenged by environmental inhibitors as well as the fact that, in the species-specific approach, the majority of genetic material found in the sample is not targeted by the primer- probe set.

Moreover, sample pooling in iDNA sample analysis has been used in various studies, but the effect of pooling in the detecting of the target species is generally not reported. Our efforts with this experiment focused on understanding the level of sample pooling of a single invertebrate group that allows for a reduction of extraction and qPCR costs without compromising the detection of the target species. The success rates of the 1:20

dilution suggested that two extraction replicates would be sufficient to detect the fragment if present in the pooled sample. Moreover, this protocol was able to detect the amplicon even in a 1:49 dilution (1 out of 3 times for each of the 3 extraction repeats). The 1:19 and 1:4 dilutions had a similar success rate, highlighting the sensitivity of this method in detecting the target fragment when present in the sample. Informed by the dilution results from this study and following methodologies from eDNA studies using 6 sample replicates for the qPCR analysis, we are confident that a pooled sample of 20 flies analysed in six qPCR sample replicates would be able to detect our target fragment if present in the sample.

Overall, this experiment contributed to understanding the threshold after which pooling fly samples for iDNA monitoring could affect detection in a single-species approach. To our knowledge no such information has been previously published for iDNA analysis so, given budget and time restrictions of this project, this experiment was instrumental in maximising detection probability while minimising the time and costs of sample analysis.

3.4.4 Fly dispersal and the assumptions we can make on the spatial resolution of positive samples

When working with iDNA samples it is important to take into consideration the dispersal behaviour of the invertebrate group the study focusses on to understand the coverage of each sampling station. An experiment conducted in the tropics studying the maximum dispersal distance of flies to draw conclusions on how far an iDNA sample might have originated from (Lee, 2016). In an attempt to replicate this study for a temperate region, an experiment was conducted during the field control sampling and fly trapping tests that took place in the initial stages of this project. The experiment was unsuccessful, failing to recapture any of the marked individuals. Therefore, this study followed published literature, suggesting daily dispersal within a 2 km radius (Lee, 2016). Combined with this study's findings on the persistence period of amplifiable mammal mtDNA in flies, the spatial resolution of the sampling grid was chosen to ensure, to the best of our efforts, closure between sampling stations.

3.4.5 Field samples

One of the advantages of iDNA is that each of the invertebrates collected is a single sampling unit, potentially carrying the target species' genetic material into the sampling station. Instead of the duration and field costs of more commonly used methods, such as camera trapping and scat surveys, or of filtering litres of water or processing large amounts of soil as with increasingly used eDNA methods, invertebrates constitute a small, concentrated sampling unit. In that sense, iDNA has the ability to collect large amounts of data, decreasing the amount of surveyor effort and bias associated with it. The field samples collected during this study were able to demonstrate that iDNA single-species monitoring is effective for rare or elusive species, but also highlighted some of the limitations of this technique. Repeated field sampling lends itself to distribution assessments, such as occupancy modelling, highlighted in previous iDNA work (Gogarten et al., 2020; Schnell et al., 2015). The use and effectiveness of occupancy modelling with iDNA data will be explored in the next chapter but it is important to study the benefits of repeat sampling in detection success. The mean DNA concentration of the field samples (1.67×10^{-9} ng μl^{-1}) falls below the LOD_6 value, but very close to $LOD_{95\%}$ (only 4.8×10^{-10} below the suggested $LOD_{95\%}$). It is possible, therefore, that positive samples show a slight decrease in amplification success. This means that the very small concentration of often very fragmented mtDNA in the samples might not always be successfully amplified to an amount the qPCR machine can detect reliably. It is important to keep in mind the optimal quality of the synthetic DNA used to create the standard curves and compute the LOD, as well as the meaning of the LOD values. The LOD simply shows levels of reliable detection (100% with LOD_6 or 95% with $LOD_{95\%}$), but the experiments have shown that this protocol (primer pair, probe, qPCR setup) can detect the amplicon in concentrations as low as 8.19×10^{-11} ng μl^{-1} with just over 30% success (18 out of 57 detections in the standard dilutions experiment). We have demonstrated that detecting bears in the field with this protocol is possible and, advised by the standard dilution experiment, the mean concentration in positive field samples has a detection success rate of just under 95% ($LOD_{95\%}$: 2.15×10^{-9} ng μl^{-1}).

As observed with the dilution experiment, pooling samples using cryogrinding was effective at detecting target species at 20 flies per pooled sample. To account for the significantly more degraded target species' DNA expected in the field, two extractions

were carried out for each field sample, under the hypothesis that a second extraction would increase detection instances. Indeed, the second extractions increased detection by 21% (6 more amplifications), suggesting that, where possible, extraction repeats as well as PCR replicates increase detection likelihood. Additionally, the sites where bears were detected changed throughout the three sampling seasons, demonstrating the efficacy of repeat site visits at increasing detections over the study area. In this case study, the number of detections increased by 37.5% after the second site visit and an additional 54.5% after the third sampling repeat. Accumulatively, the second and third repeats added an additional 9 site detections, increasing the overall site detections by 112.5%.

I believe that the issue of detecting false negatives resulting from the decreased amplification rate due to sample quality (reflected in the low mean DNA concentration in the field samples) and loss of data that could have resulted from performing a single survey repeat or a single extraction per pooled sample, was addressed by the various levels of replication (field repeat sampling, repeat DNA extractions and six qPCR replicates) aiming to increase the chances of detecting the target species when mtDNA is in very low concentrations. The lowest concentrations in the field iDNA samples were just over 1.6×10^{-10} ng μl^{-1} and the highest concentrations over 2.7×10^{-8} ng μl^{-1}, demonstrating the notable differences in mtDNA concentration between samples. Our results on repeat sampling for iDNA highlighted the importance of multiple site visits as well as DNA extraction repeats to increase yield and account for false negatives during sample analysis. Repeat sampling in iDNA surveys shows very promising results for the detection of rare and elusive species. In cases where the DNA is expected to be degraded, the results show that small DNA concentrations will result in smaller numbers of amplifications within the qPCR replicates, emphasising the need for repeat surveys and repeat DNA extractions. Our results demonstrate how sensitive this method is at tracing bears in the field, while also bringing attention to approaches that could drastically improve detection yield.

3.4.6 The potential of target species individual identification using iDNA

Overall, this chapter explored a number of approaches aiming to increase detection and optimise the laboratory analysis of iDNA samples. The laboratory experiments, with data

collected in the field, provides a comprehensive insight to the collection, processing and laboratory analysis of iDNA samples. The methods utilised were all centred around techniques and equipment available in smaller laboratory facilities and associated with lower costs relative to next-generation approaches. Exploring iDNA as a single-species detection approach to this depth is a novel study and we believe that this chapter contributes to our understanding of the advantages and limitations of iDNA as a species-specific monitoring tool. The results of this chapter can serve as the skeleton of an adaptable protocol for other iDNA single-species studies.

Chapter 4: Modelling brown bear occupancy in Northern Greece: a comparison of iDNA and scat sampling

4.1 Introduction

4.1.1 Considering imperfect detection

Imperfect detection during field sampling sessions, especially in the case of rare or elusive animals or where the habitats are used less intensely, often leads to incorrect interpretations of a species' ecological requirements and habitat use (Mackenzie et al., 2002). Occupancy modelling was developed to address this issue, by accounting for imperfect detections on repeat surveys. Occupancy models examine both the probability of a species inhabiting that area and the probability of detecting it if it is indeed present on site (Mackenzie et al., 2002). Furthermore, false absences are very common in the case rare or elusive animals or areas of low use, as most monitoring methods fail to detect the species when it is present in the area (Karanth et al., 2011; Keane et al., 2012; Strickland and Roberts, 2019). To tackle that, occupancy models require repeat surveys, increasing the number of observations per site. Revisiting sites multiple times increases the chances of detecting the species, thus reducing the likelihood of recording false absences.

To compare this method to the Species Distribution Modelling (SDM) method used in Chapter 2, occupancy modelling is often referred to as the realised niche, or the suitable habitat occupied by the species (Braschler et al., 2020; McGeoch and Gaston, 2002). SDMs, on the other hand, reveal a species' fundamental niche, identifying areas suitable for the species, but with no assumption that the species is present in the area. Therefore, theoretically, when detectability is close to "1", an SDM (assuming it is based on the correct underlying data) and an occupancy model should present the same results. However, most monitoring efforts are not effective to that extent, resulting in imperfect detection, false absence records and inaccurate predictions of a species' realised niche (Gu and Swihart, 2004; Mackenzie et al., 2002). In fact, considering that field surveys rely on a variety of factors related, but not limited to, the effectiveness of the specific method used for detecting the target species, surveyor experience and

training, climatic conditions, and the nature and abundance of the species in question, imperfect detection is to be expected in most monitoring efforts (Long, 2008; Thompson, 2004). It is, therefore, recognised that detectability is most likely imperfect (probability of detection < 1), thus, if we were to compare the resulting occupancy model (repeat visits) and SDM (presence-only dataset) of the same survey efforts, we would expect to see differences, highlighting the importance of accounting for imperfect detection and false absences in low-use habitats. In the case of rare or elusive species, especially, the need for multiple site visits and an analysis that takes imperfect detection into consideration becomes more apparent (Karanth et al., 2011; Keane et al., 2012).

In the case of species monitoring in wildlife corridors, imperfect detection could occur as a result of low habitat use and the relatively low number of individuals using the corridor at any given time. Furthermore, corridors are often less suitable habitats due to the fact that they are, by definition, relatively small 'linkage zone' spaces that improve functional connectivity between core areas, but often associated with higher levels of disturbance (Bennet, 1999; Lawton, 2010; Servheen, Waller and Sandstrom, 2001). In the case of brown bears, home range size and movement into less suitable areas can be the result of a number of reasons, such as sex, body mass and size, age, food availability, and population density, as outlined in (Dahle, Støen and Swenson, 2006). Male brown bears are territorial, with large home ranges that envelop those of several females who tend to establish smaller home ranges that are near or within their natal areas. In contrast, sub-adult males disperse in search of new home ranges, travelling distances of up to 90 km (Dahle, Støen and Swenson, 2006; Swenson, Sandegren and Soderberg, 1998). Additionally, to the potential use of corridors as pathways to other core areas by dispersing subadult males, these less suitable areas are serving as linkage zones to food sources and mates. The tolerance of bears to human disturbance indicates that bears are likely to opportunistically utilise the broader landscape (Elfström et al., 2014b). With the above in mind, we can expect that surveys in corridors would have a higher rate of false absences, and assume that an SDM generated from a single-survey data may not provide a reliable representation of these low-use areas. Due to the nature of the study area, occupancy modelling offers an alternative way of looking into bear distribution throughout this complex landscape, while taking detectability into account by performing repeat site visits.

4.1.2 Effective survey methods

Several non-invasive methods commonly used of detecting brown bears are carried out to monitor bears in Europe. Track and sign surveys aiming at detecting signs of bear presence, like tracks, scat, fur and damage on trees and electricity poles, as well as camera trap surveys are used interchangeably depending on the purpose of the study (Kaczensky et al., 2012b, 2012a). However, such monitoring efforts require a team of trained surveyors and call for a large effort in the field, often centred around the forestry network. In fact, scat surveys are frequently conducted on forestry roads and paths to increase visibility, ensure access and minimise disturbance. Scat is often recorded on trails and forestry roads (or 'low-volume roads' as described by (Chruszcz et al., 2003)) as bears and other large carnivores will repeatedly use specific routes in search of food and/or for leaving territorial markers (scat, scrapes, etc.) for conspecifics (Henschel, 2015; O'Brien, Kinnaird and Wibisono, 2003; Phoebus et al., 2020). In the case of road transects, it is important to consider that roadside observations can give rise to observation bias (Austin et al., 2000; Keller and Scallan, 1999), resulting in the model considering proximity to road as an important factor when it comes to predicting the species' distribution (see (Kadmon, Farber and Danin, 2004) as an example of roadside bias effects on predicting species distribution).

The capacity of occupancy models to account for imperfect detection in an analysis can create more accurate distribution models (Comte and Grenouillet, 2013; MacKenzie et al., 2017). Species Distribution Models, such as those created in chapter 2, use presence-only data and all the information accumulated for the model generation comes from our understanding of the species' preferences (Elith et al., 2006). Occupancy, on the other hand, adds the element of absence, allowing for the model to take the ecological settings in absence areas into account too. Additionally, occupancy modelling takes into account factors that could have affected the detection of the species, by performing multiple visits on the same site (repeat surveys) which SDMs do not account for (Mackenzie et al., 2002). Furthermore, occupancy models also provide information on factors affecting species detection in the field (observation covariates), including, but not limited to, survey-effort information (such as transect length, number of observers, time of survey, duration, etc.), biotic and physiological interactions, which MaxEnt

models are not able to incorporate (Guisan and Thuiller, 2005). As an example, (Peterman, Crawford and Kuhns, 2013) compared SDMs created on MaxEnt with occupancy modelling and found that the latter was able to reveal that the distribution of *Ambystoma jeffersonianum*, a threatened North American salamander species, was impacted by the presence of fish in potential breeding ponds. Such inclusion of biotic factors demonstrates another strength of occupancy modelling, even when a small sample size is considered. Overall, my decision to move forward with designing the field surveys as a smaller set of repeat surveys was made under the assumption that, based on all the above, occupancy modelling would be more suitable for processing data derived from rare or elusive species, areas of infrequent use and monitoring methods that are very sensitive and more successful in repeat survey scenarios.

This chapter explores the potential of using detections from the iDNA surveys to build a robust spatially-explicit occupancy model when compared to scat surveys. Gogarten and colleagues (2020) looked at iDNA as a means of monitoring biodiversity compared to camera traps. Their results suggested that iDNA can be especially useful as an effective complementary method, especially in terms of detecting smaller animals that camera traps fail to detect. As this study is solely focused on monitoring brown bears, it presents a good case for examining whether this complementarity is reflected when looking at single-species detection. If this is the case, it might be expected that merging the survey datasets for each season would improve the occupancy model.

Considering that iDNA is a novel monitoring method that has been promoted as a way of increasing time efficiency in the field and cost-reduction, scat surveys were chosen as the closest comparable and more commonly used monitoring method (see report by (Kaczensky et al., 2012b) highlighting the use of scat surveys for carnivore monitoring in many European countries). Bear scat is relatively easy to identify, which reduces the need for highly trained surveyors, as demonstrated by (Bellemain et al., 2005) who collaborated with hunters for the collection of scat samples.

4.2.3 Chapter aims

This chapter aimed to answer the following questions:

1. How does brown bear (*U. arctos*) detection using iDNA compare to a standard survey method: scat surveys?
2. Can a single-species iDNA survey protocol be used to generate occupancy models?
3. How do iDNA-derived occupancy models compare to those derived from scat surveys when predicting the distribution of *U. arctos* across a landscape of varying habitat suitability?

It is important to note that, to date, this is the only study to date that uses single-species iDNA detection data to produce habitat models and draw conclusions on the distribution of the target species. The results will contribute to the larger body of work concerning iDNA and highlight its use in occupancy modelling and potential to inform species conservation strategies.

4.2 Methods

4.2.1 Study area

As outlined in the previous chapter, this case study focused on a smaller study area, the Mt. Voras – Mt. Verno connectivity section, as described in Chapter 3, within the greater study site (see Chapter 2). The study area incorporated two core bear habitats and the less suitable area between them, part of which was presumed to function as a movement corridor (see Savvantoglou et al., 2017; here, Figure 30. The area was chosen due to the gradient suitability across the landscape, as predicted by the Expert Knowledge model that incorporated the area's topography, land use and human impact, making a good case study for the effectiveness of the two survey methods in habitats of predicted high, medium and low use (Figure 30).

4.2.2 Sample collection

A 2 x 2 km grid was laid over the area and ten grids from each predicted broad suitability range (high-medium-low, N = 30) were selected randomly and sampled for both scat and iDNA. The sampling season took place between August and September 2017 and

was repeated in the years 2018 and 2019, resulting in three repeat surveys for both methods in each of each of the 30 grids. By focusing monitoring efforts outside the bear breeding season (April-June) when males disperse further in search of females, we ensured that bear movement range is within normal limits (Clevenger, Purroy and Pelton, 1990; Piédallu et al., 2017).

Scat transects and fly sampling was conducted simultaneously within each survey cell on each sampling occasion. Both survey techniques were performed between the hours of 9am and 5pm by the same team, carried out during the same seasons and simillar weather conditions, and, in the case of recording the presence of scat, following the same road transects across the three sampling seasons. Scat detections were recorded out on road transects by inspecting the available forestry road network found within the sampling grid and observations were recorded using the sample's GPS location. The iDNA sampling targeted flies and were carried out using one sampling station in the middle of each sampling grid, repeated once per field season. The fly sampling, preservation and subsequent analysis details are outlined in Chapter 3. Finally, a *Combined* dataset was created by collating the results of the two surveys, indicating presence where either scat transects or iDNA sampling were successful at detecting bears.

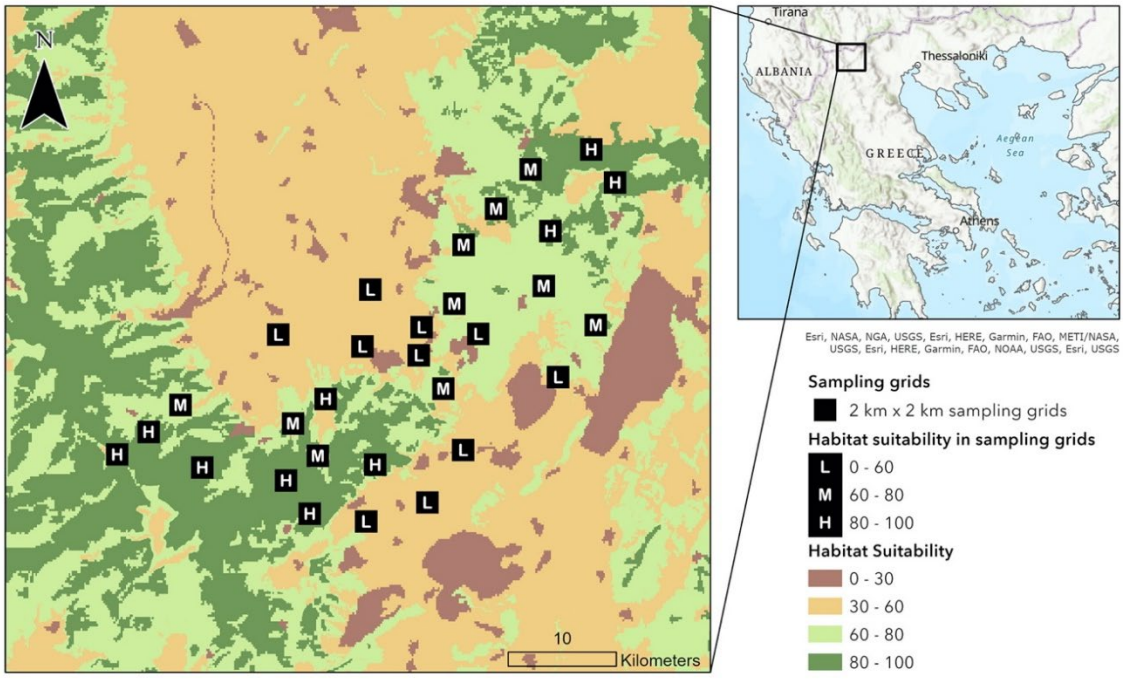


Figure 30. Study site and sampling grids on a habitat suitability model (Expert Knowledge model, see Chapter 2). The letters in the black sampling grids describe the habitat suitability score within the sampling grid. Figure taken from Chapter 3.

4.2.3 Environmental Parameters

To allow for a direct comparison, the same set of environmental and habitat covariates were used across all subsequent modelling scenarios. The set of predictors considered for the species distribution modelling described in Chapter 2 was used to explore the most descriptive combination of predictors for performing an occupancy analysis on the survey data (Table 8). Following the resolution of the survey grid squares, a raster with a cell size of 2 x 2 km was created, covering the entire study area. Additionally, the cells of this raster were used to create a dataset where values from each variable were given to their corresponding cell. For continuous variables (elevation, distance from major roads and distance from urban areas), the mean and mean standard deviation (SD) of the values within each sampling cell (2 x 2 km) were calculated for the original variable raster cells (resolution: 30 x 30 m). In categorical habitat covariates (land cover and topography) the categories were split to individual datasets, showing the percentage cover of each category within each sampling cell. The percentage cover was calculated in proportion to the other categories present in that cell.

To allow for uncontrollable inter-annual differences, a 'survey year' dataset was added to the analysis to allow for such differences across the monitoring years to influence detection. Additionally, guided by each survey dataset's predictor importance, some of the environmental and habitat variables mentioned above were included as survey covariates. Specifically, percentage of canyons was included in two models (iDNA and combined surveys) and the SD distance from roads was included in the iDNA surveys. Moreover, transect length was considered as a factor that could influence detection in the scat and combined surveys models.

Table 8. Model predictors: groups, sources and resolution of the categorical and continuous variables used to predict bear occupancy.

Predictor group	Predictor	Data type	Source	Resolution (metres)
Habitat and topography	**Land cover**	Categorical	Corine Land Cover (CLC) 2018 - Copernicus, European Environmental Agency (EEA)	100
	Elevation	Continuous	EU DEM, Copernicus, European Environmental Agency (EEA)	30
	Topography	Categorical	Derived from elevation raster using the Corridor Designer toolbox on ArcMap 10.7	30
Anthropogenic influences	**Distance from urban areas**	Continuous	Euclidean distance from CLC 2018 classes 1.1.1 - 1.2.1	30
	Distance from roads	Continuous	Euclidean distance form major roads derived from OpenStreet Map (open source data, licensed under the Open Data Commons Open Database Licence (ODdL)	30
Survey specific	**Survey year**	Categorical	Field data	n/a
	Transect length	Continuous	Field data	n/a

4.2.4 Data Analysis

Three encounter rate predictions were generated independently using the scat survey and iDNA survey datasets, as well as the Combined dataset, using the R package ranger

(Wright and Ziegler, 2015). Additionally, occupancy modelling was performed on the three datasets independently using the R package unmarked (Fiske and Chandler, 2011; 2019; see Figure 31 for a schematic of the workflow). Data preparation and spatial modelling was performed in R version 4.0.4 (R Core Team, 2021). To account for spatial autocorrelation within transects, each set of scat observations within a grid was considered a single detection record. Similarly, the results of replicate iDNA extractions and qPCR repeats were treated as a positive record when one of more qPCR sample repeats amplified. For the *Combined* model, the data was pooled together, and any evidence of species presence was a positive presence data.

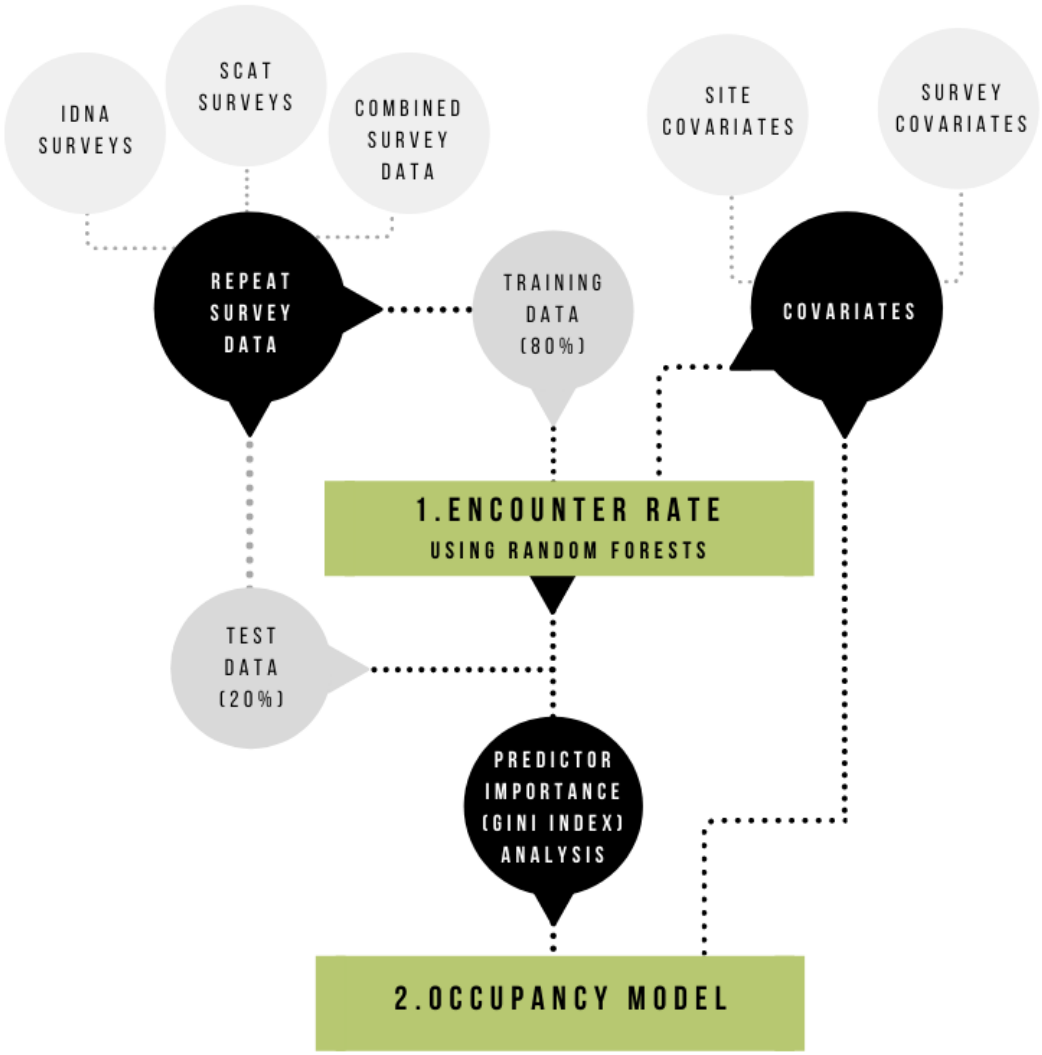


Figure 31. Schematic of the workflow for this chapter, outlining the basic steps taken to create a robust set of occupancy models using the iDNA, scat survey and combined survey datasets.

4.2.5 Evaluating brown bear encounter rate

In order to create robust, spatially-explicit occupancy models it is important that data is pre-assessed to provide statistically meaningful inputs into the occupancy analysis (Guillera-Arroita et al., 2015; Strimas-Mackey et al., 2020). This was achieved using a Random Forest (RF) approach, a machine learning algorithm process used in classification analyses. In this case, the RF algorithm was used to classify detections and non-detections of bears using the two different survey methods. Firstly, the Encounter Rate model showed the variation in detectability and occurrence as a joint framework (in contrast to occupancy models that look into these two processes separately). Secondly, a Predictor Importance (PI) rank was generated, in the form of an average Gini Index analysis, showing which covariates are most influential in the model. Gini Index, a statistical analysis originally used in economics, measures the distribution of values across a dataset (Gini, 1921). In the case of PI, the Gini Index ranges from 0, where all predictors are equally important, to 1, where only one predictor contributes to the variation in data (Chen and Liaw, 2004; Strimas-Mackey et al., 2020). Even though RF models do not take imperfect detection into account, including effort covariates in the RF analysis results in measurements of encounter rate proportional to occupancy, therefore this PI serves as indicator best predictors for occupancy modelling (Guillera-Arroita et al., 2015; Strimas-Mackey et al., 2020). In the case of more elusive species, these RF encounter rate measurements are expected to be lower than the occupancy values (Strimas-Mackey et al., 2020), but the predictor importance from the RF analysis will still highlight the covariates with the largest contribution to the model. In view of the number of covariates considered in this study, the PI analysis here was key in guiding unbiased decisions of the most important predictors for the generation of occupancy models.

Models were generated individually for each survey type (scat and iDNA) and by combining the two survey datasets (Combined dataset). In order to stay consistent with the field survey grid system, a study area raster with a cell size of 2 x 2 km^2 was used to develop the encounter rate and occupancy models. Each of the three survey datasets (scat surveys, iDNA surveys and combined surveys) was randomly split into 80% to be

used for training the model and 20% to test the model's predictive performance once the model was fit to the dataset following the protocol of (Strimas-Mackey et al., 2020).

The probability of encountering the species (joint detectability and occurrence) was calculated using balanced random forest analysis, a machine learning approach using R package ranger, (Wright and Ziegler, 2015) which models detection/non-detection of the species in relation to the environmental and habitat covariates. The balanced random forest analysis resamples the data using bootstrapping, so that on each generated 'tree' there is an equal amount of detection and non-detection data (Robinson, Ruiz-Gutierrez and Fink, 2018; Strimas-Mackey et al., 2020).

The balanced random forest approach is ideal for accounting for rare detections, but it can introduce bias due to the selective subsampling methods used. To account for that, following protocol from (Strimas-Mackey et al., 2020), the model results were calibrated by predicting the encounter rate for each sampling grid in the training set and fitting a binomial Generalised Additive Model (GAM) using the real observed encounter rate as the response variable and the predicted encounter rate as the predictor variable. GAM's for each model were generated using the R package scam (Pya and Wood, 2014) by applying monotone increasing P-splines ('bs="mpi"') to account for the *a priori* assumption that high real-observed encounter rate values will correspond to high estimated-observed encounter rate values.

The calibrated and uncalibrated models were assessed by examining the values of mean square error (MSE) opting for the prediction model with the smallest MSE value and therefore smallest error between observed and predicted values using the test data. The values of sensitivity and specificity will allow for an examination of how accurate the model is at predicting encounter rate when the animal is present (sensitivity), and how accurate the model is at correctly predicting absence where they animal is not present (specificity). Finally Cohen's κ, which looks at the level of agreement between two predictions (while also allowing for the agreement to be the product of chance) and Area Under the Curve (AUC), which measures the probability of the model's ability to make predictions that are better than random. The predictors that best describe bear encounters were selected using the Predictor Importance values (Gini Index approach),

which describe the changes in model accuracy when a predictor is removed from the algorithm. Each of the predictors was tested and ranked for their importance against a training subsample of the dataset. Finally, the calibrated random forest analysis was used to generate a prediction model of the brown bear encounter rate for each of the three methods.

4.2.6. Occupancy modelling

The occupancy models were generated using the R package unmarked (Fiske and Chandler, 2019, 2011) using covariates with predictor importance value larger than or equal to 0.8 (PI ≥ 0.8). As an additional measure, the predictors were tested for correlation (R package 'stats', v.4.0.4) and removed from the analysis when a correlation larger than or equal to 0.8 was present between predictor pairs. Further removal of predictors from the occupancy analysis was suggested by the Unmarked script when the model did not converge, suggesting the use of less covariates. A goodness-of-fit test (MacKenzie and Bailey, 2004) was applied to the resulting occupancy models with a one thousand bootstrap sample simulations (nsim = 1000), comparing the observed and expected frequencies of detection observations.

The function *dredge* was used to generate all possible covariate combinations from the original unmarked frame dataset as well as a model assuming constant occupancy (ψ(.)) and constant detectability (p(.) across the study site. The models were compared using ΔAICc values which show the difference between the corrected Akaike Information Criterion (AICc) between the best fit model and each of the other models generated by the *dredge* function. Where it was not clear that a single model was the most likely candidate to explain the data (i.e., ΔAICc ≤ 2.5), a model average prediction was generated (Strimas-Mackey et al., 2020). Finally, the model-averaged prediction was rasterised to generate a prediction surface using a 2 x 2 km template raster. Response curves were generated using the model-averaged predictions, highlighted the direct relationship between occupancy (ψ) or detection (p) and each of the predictors depending on whether they were used as site or observation covariates, respectively. The three resulting occupancy models (iDNA, scat and combined data) will be referred to as 'optimised models' in the rest of the text.

In addition to these models, a further three models were produced by adding all the site covariates and all the observation covariates used in the above models, thus creating a consistent covariate dataset for all three models. These three models will be referred to as 'comparison models' in the rest of the text.

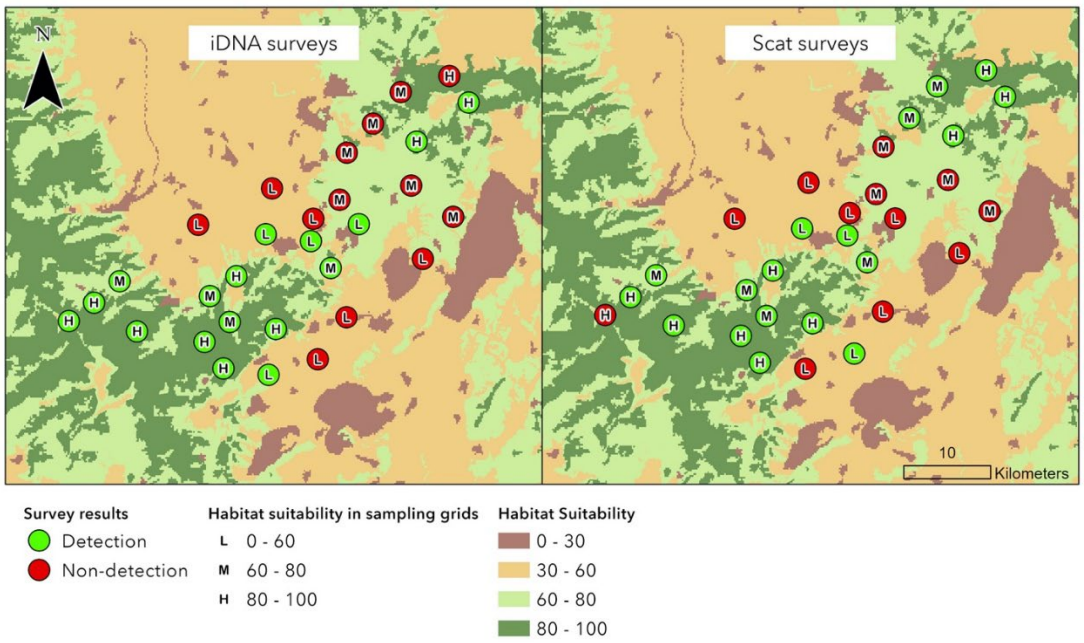
Finally, as a means of comparing the models, the Probability of Occupancy raster produced from the scat survey comparison model was subtracted from the iDNA model in ArcGIS Pro (v 2.8, Esri), revealing the areas where the models were the most different.

Ethical approval was granted by the UWE Animal Welfare and Ethics Sub-committee, reference number: R34.

4.3 Results

4.3.1 Presence data

A total of 30 out of 90 scat surveys (33.3%) detected scats and 27 out of 90 iDNA samples (30%) were amplified, indicating the presence of brown bear mtDNA within flies captured from that sample cell (Figure 32). The combined dataset increased the number of positive observations to 45 out of 90, with 12 surveys detecting bears in both iDNA sampling and scat transects simultaneously (13.3% of all surveys, 26.6% of all positive surveys; Figure 32). The spatially thinned scat transect detections and results from the laboratory analysis of iDNA samples formed the final datasets used for the spatial modelling. In terms of the relationship between detection and suitability, both methods had nine detections in high suitability areas, four in medium suitability areas for iDNA and six for scat survey data. Finally, four detections were recorded in low suitability areas with iDNA sampling and three with scat transects (Figure 32).



All Seasons		iDNA	
		0	1
Scat	0	45	15
	1	18	12

Season 1		iDNA	
		0	1
Scat	0	13	3
	1	9	5

Season 2		iDNA	
		0	1
Scat	0	17	2
	1	5	6

Season 3		iDNA	
		0	1
Scat	0	15	10
	1	4	1

Figure 32. Survey results and comparison matrices of field detections for all seasons as well as for each season for scat and iDNA detections. Survey results shown on top of a habitat suitability model and the overall suitability within the sampling grid is indicated as Low/Medium/High in the letters within the survey results.

4.3.2 Model predictors

The predictor selection for the optimised models was initially made using the Gini Index analysis for each of the three survey datasets (Table 9). These top predictors were also checked for correlation (Appendix, section C1. Correlation of top predictors).

Table 9. Gini Index results highlighting the importance of the top predictors (PI ≥ 0.8) for each of the three datasets. All data on land cover type (in green) and topography (in blue) covariates are represented here as their percentage cover (relative to the other layers from their respective categories) within a cell. Covariates associated with distance from roads and urban areas (in yellow) and elevation (in orange) were analysed as the mean and mean SD of the distance of each pixel within the sampling cell. Survey covariates in purple.

	Predictor Importance (PI)		
Predictors	iDNA surveys	Scat surveys	Combined surveys
Broadleaved forest	2.22	1.09	3.50
Flat/gentle slopes	0.87	1.22	2.32
Transect length	< 0.8	1.10	2.11
Natural grassland	< 0.8	1.30	1.63
Mean elevation	0.93	1.27	1.60
Steep slopes	0.83	0.94	1.50
Canyons	0.97	0.88	1.49
Mean distance from major roads SD	< 0.8	< 0.8	1.35
Mean distance from major roads	1.17	< 0.8	1.28
Transitional woodland	< 0.8	0.94	1.22
Mean elevation SD	< 0.8	0.87	1.08
Mean distance form urban areas	< 0.8	< 0.8	1.02
Non-irrigated arable land	1.22	< 0.8	0.89
Ridgetops	< 0.8	< 0.8	0.88
Mean distance form urban areas SD	< 0.8	< 0.8	0.81
Survey year	< 0.8	1.30	< 0.8

4.3.3 Encounter rate

The random forest calibration showed a small improvement on the mean squared error after calibration (Figure 45 in Appendix), while all other metrics remained constant. The calibrated models were used to generate an encounter rate model for each of the datasets. The resulting encounter models were rasterised to create a mosaic of encounter rate (a joint measurement of detectability and occurrence) for each of the survey methods (Figure 33).

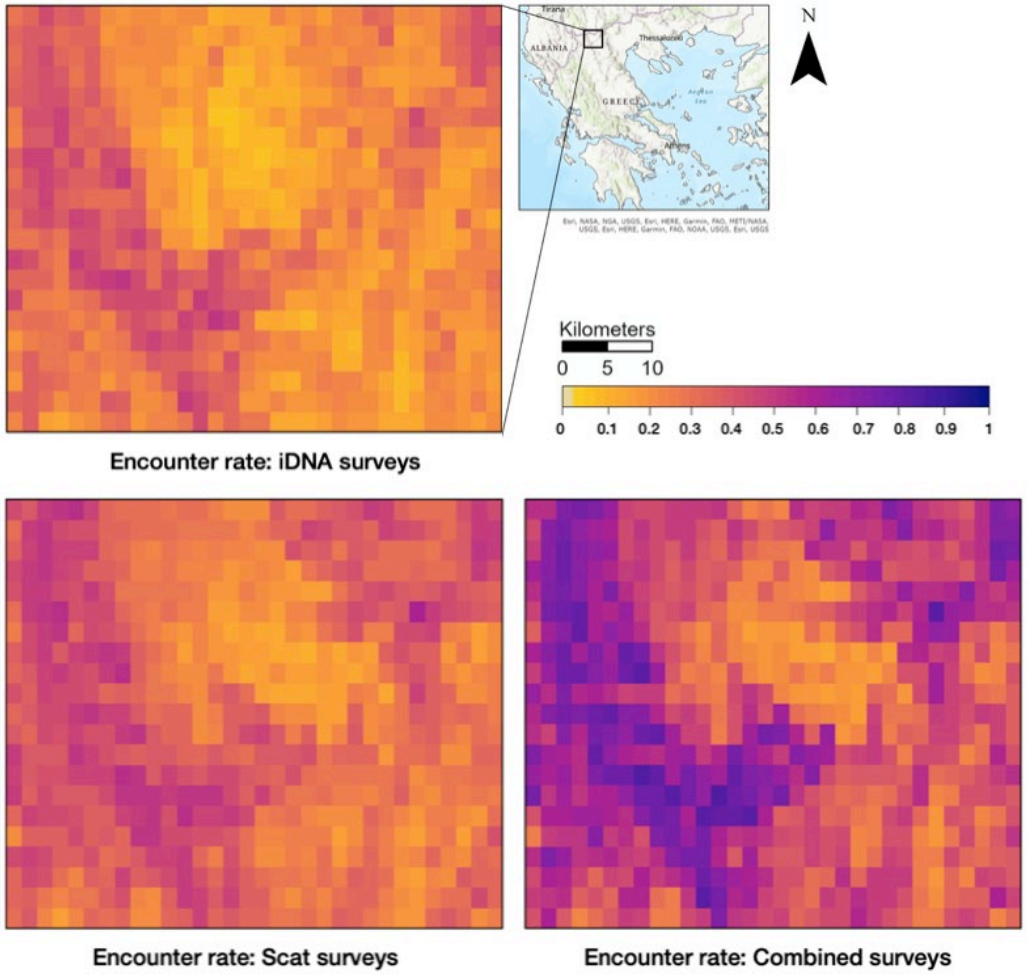


Figure 33. Encounter rate models of iDNA (top left), scat survey (bottom left) and combined dataset (bottom right).

4.3.4 Occupancy modelling

Occupancy model selection – optimised models

Each data frame of observations of the three survey combinations and the four predictors were converted to an *unmarked* object to inform a single-season occupancy model. The goodness-of-fit p-value for all three optimised models was larger than 0.1,

suggesting there was no reason to consider lack of fit, but the estimate of ĉ value for the scat data (0.23) showed some evidence of overdispersion (Table 10).

Table 10. Model evaluation metrics for the optimised occupancy models (AIC and Goodness of Fit) and model coefficients (in logit-scale) for occupancy (ψ) and detection (p) for each model. Note that habitat type and topography covariates show the percentage of land covered by that habitat/topography type within each sampling cell.

			iDNA	Scat	Combined
Model AIC			106.67	102.48	96.31
Goodness of fit		χ^2	5.36	1.19	3.86
		p	0.50	0.97	0.68
		$\hat{c}$	0.90	0.23	0.69
Model-averaged coefficients (logit-scale)	ψ	.	-2.71	16.51	12.87
		Mean elevation	-0.12	-13.30	-
		Broadleaved forests	4.64	-	-
		Natural grassland	-	-22.98	-16.09
		Flat/gentle slopes	-	-16.06	-13.54
	p	.	21.74	2.10	-0.49
		Canyons	10.44	-	8.61
		Distance from major roads	-17.69	-	-
		Survey year	-	-1.01	-
		Transect length	-	-	0.33

The function dredge resulted in a set of candidate models that were subsequently ranked by their AICc values, with the lowest AICc model and smallest delta AICc values being the ones considered to fit the data more suitably. The models with a ΔAICc value ≤ 2.5 were averaged to generate a single model-averaged prediction (see Table 11 for contributing models) with six models contributing to the iDNA, two to the scat and four to the combined model-averaged predictions.

Table 11. Model dredge showing the best fit models (ΔAICc≤ 2.5) contributing to the model-averaged predictions and model coefficients (in logit-scale) for occupancy (psi) and detection (p) for each model. Note that habitat type and topography covariates show the percentage of land covered by that habitat/topography type within each sampling cell.

iDNA model dredge - models with ΔAICc ≤ 2.5

Model	ψ (.)	ψ (mean elevation)	ψ (broadleaved forests)	*p* (.)	*p* (mean distance from major roads)	*p* (canyons)	df	AICc	ΔAICc	weight
9	-0.30	NA	3.73	-0.17	NA	NA	3.00	107.26	0.00	0.20
8	-14.73	21.74	NA	0.12	-22.59	13.31	5.00	108.14	0.88	0.13
12	-0.11	NA	7.14	0.28	-19.46	10.33	5.00	108.23	0.97	0.12
10	-0.33	NA	5.20	0.39	-8.53	NA	4.00	108.80	1.54	0.09
11	-0.11	NA	2.63	-0.61	NA	6.47	4.00	108.86	1.60	0.09
3	1.24	NA	NA	-1.11	NA	10.42	3.00	109.23	1.97	0.07

Scat model dredge - models with ΔAICc ≤ 2.5

Model number	ψ (.)	ψ (mean elevation)	ψ (natural grasslands)	ψ (flat/gentle slopes)	*p* (.)	*p* (survey year)	df	AICc	ΔAICc	weight
386	12.40	NA	-22.11	-13.81	2.10	-1.01	5.00	92.11	0.00	0.17
418	30.07	-13.30	-25.86	-23.52	2.10	-1.01	6.00	94.50	2.39	0.05

Combined data model dredge - models with ΔAICc ≤ 2.5

Model number	ψ (.)	ψ (flat/gentle slopes)	ψ (natural grasslands)	*p* (.)	*p* (transect length)	*p* (canyons)	df	AICc	ΔAICc	weight
46	14.32	-14.90	-17.70	-2.00	0.37	9.55	6.00	96.36	0.00	0.18
41	11.80	-12.55	-14.88	0.89	NA	NA	4.00	96.82	0.46	0.14
45	12.08	-12.84	-15.17	-0.65	0.26	NA	5.00	97.19	0.82	0.12
42	12.73	-13.36	-16.03	0.37	NA	7.04	5.00	97.40	1.03	0.11

The model-averaged response curves highlighted the relationship of each variable to the model's occupancy or detection predictions (Figure 34, Figure 35 and Figure 36).

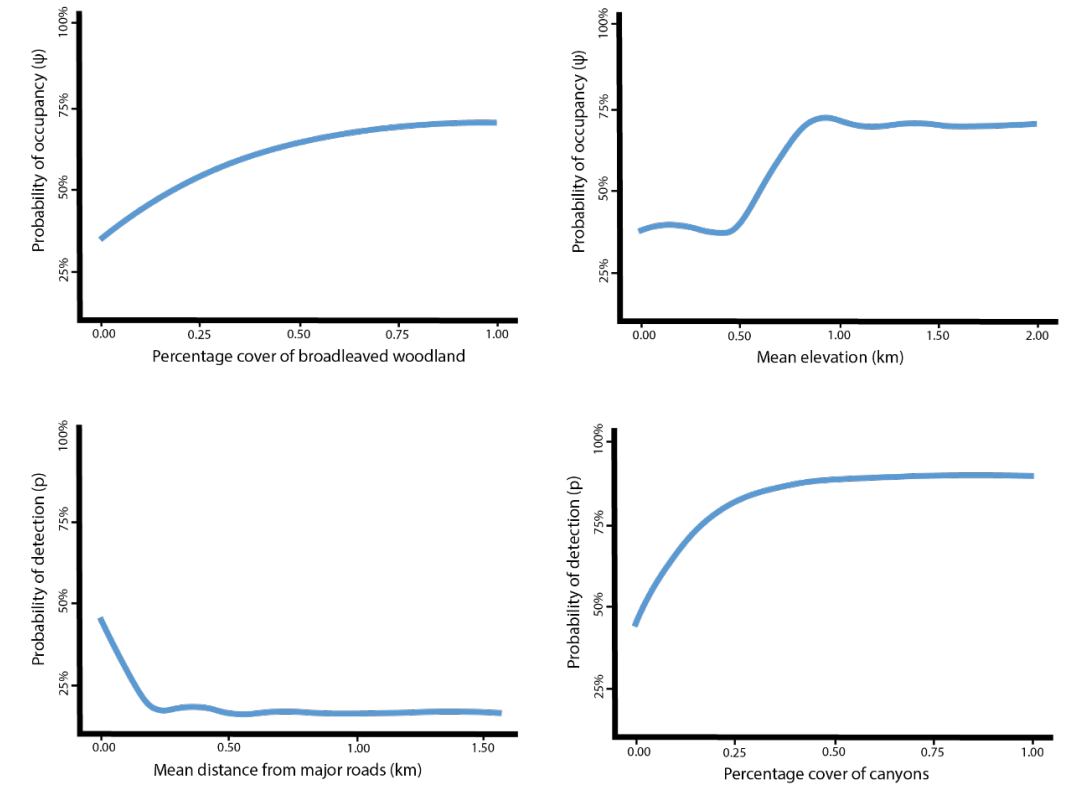


Figure 34. Response curves for iDNA model-averaged predictions with 95% CI.

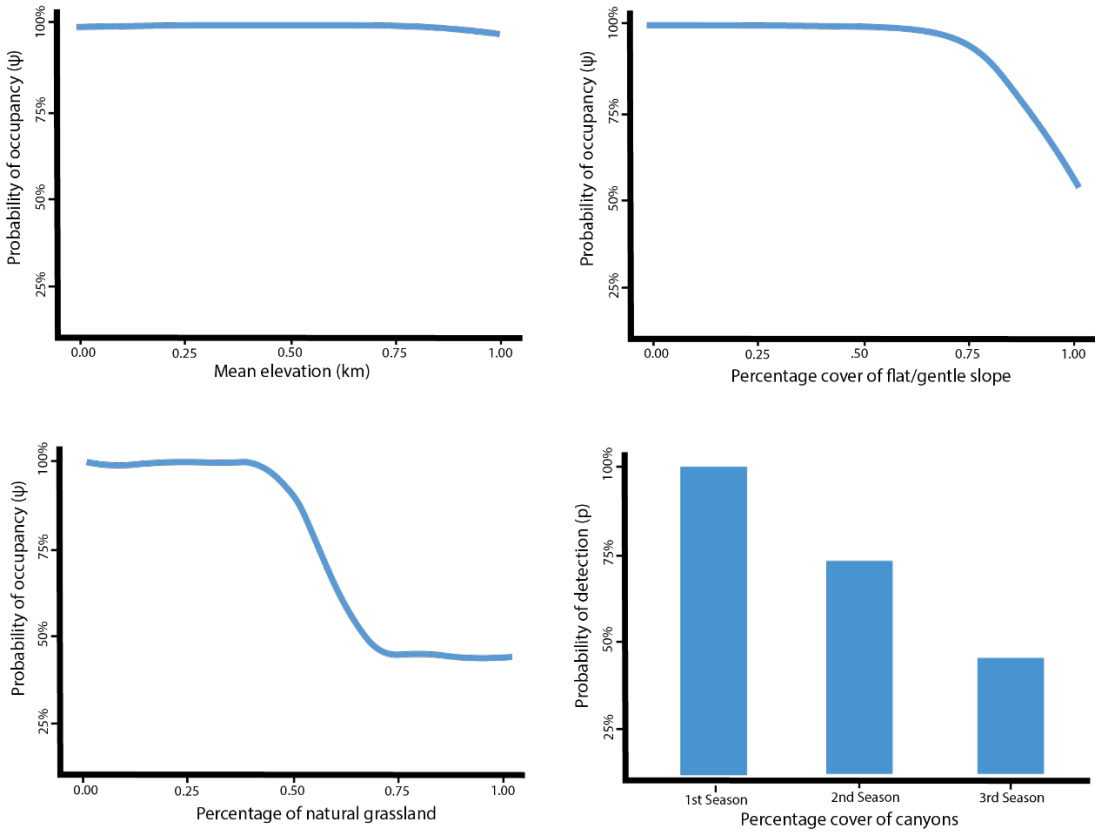


Figure 35. response curves for scat model-averaged predictions with 95% CI.

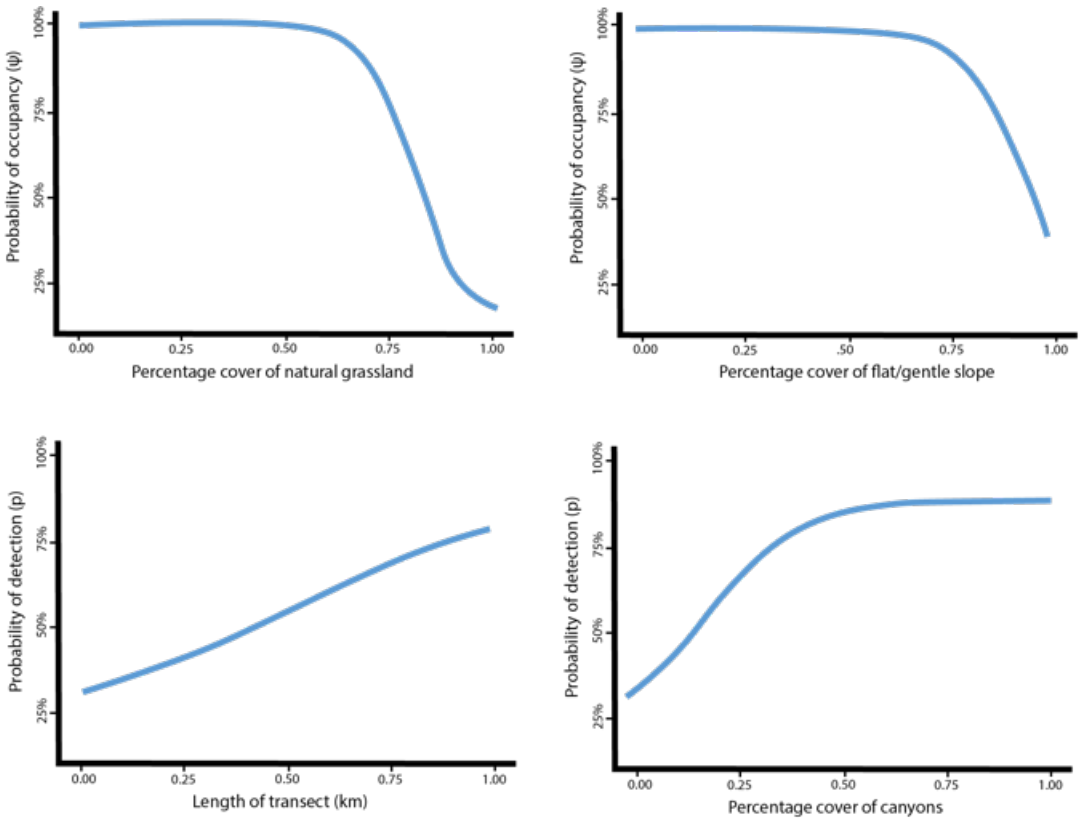


Figure 36. GAM response curves for combined data model-averaged predictions with 95% CI.

Finally, the comparison models were generated using seven out of the nine covariates used across the three optimised models. The percentage cover of broadleaved forests was excluded from the analysis, because the models did not converge when it was included in the analysis, while transect length was removed, because it was not associated with the collection of iDNA data. The results revealed that detection is influenced by different sets of observation covariates across the models. Specifically, the iDNA model-averaged predictions utilised the percentage of canyons, mean distance from major roads and survey year; while the scat model used only the percentage of canyons and survey year, and the combined surveys used the mean distance from roads and percentage of canyon cover to predict detection (Table 12).

Table 12. Model evaluation metrics for the comparison occupancy models (AIC and Goodness of Fit) and model coefficients (in logit-scale) for occupancy (psi) and detection (p) for each model. Occupancy estimates were not calculated for the combined model's mean elevation ($psi_{combined}$(mean elevation) because the dredge analysis did not highlight any models using mean elevation good enough to include in the model-averaged prediction. Note that habitat type and topography covariates show the percentage of land cover.

			Comparison iDNA model	Comparison scat model	Comparison combined model
Model AIC			100.54	93.87	94.10
Goodness of fit		χ^2	4.33	0.90	4.06
		p	0.51	0.98	0.48
		$\hat{c}$	0.87	0.19	0.89
Model-averaged coefficients (logit-scale)	ψ	.	17.27	15.92	11.58
		flat/gentle slopes	-18.88	-15.79	-12.36
		natural grassland	-20.41	23.10	-14.55
		mean elevation	-	-13.30	NA
	p	.	-0.03	2.04	1.37
		mean distance from major roads	-13.71	-	-23.27
		canyons	8.93	3.86	16.12
		survey year	0.30	-1.01	-

Optimised occupancy models

The model-averaged probability of occupancy and standard error (occupancy uncertainty) were extrapolated for the entire study area, creating rasters that depict occupancy across the site. The comparison model predictions are seen in Figure 37 and the optimised models can be found in the Appendix (C2. Optimised occupancy models).

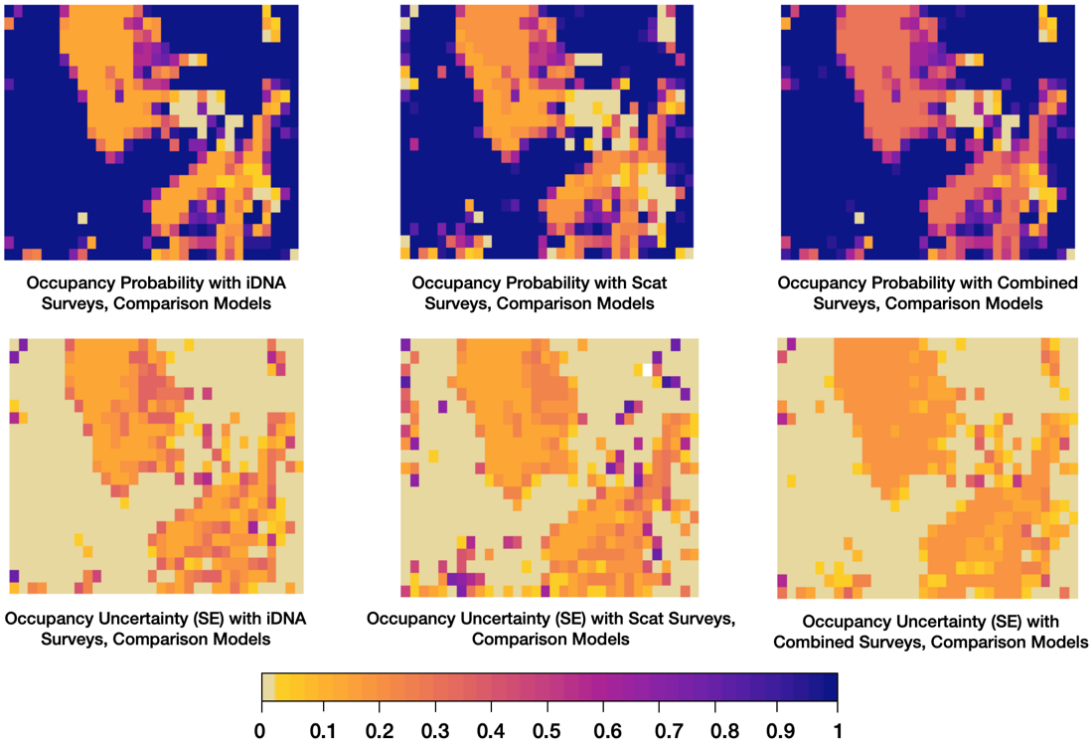


Figure 37. Comparison models: Occupancy probability and uncertainty for iDNA (left), scat survey (middle) and combined dataset (right).

Finally, the comparison between the iDNA and scat survey models highlighted the differences between the two models, pointing to areas within core habitats displaying high probability of occupancy with the scat survey dataset. These models appear to highlight the areas of higher frequency of bear use and display a larger degree of uncertainty in the edges between high and low occupancy. Contrastingly, the corridor and some edge habitats showed higher probability of occupancy with iDNA data, matched with smaller degrees of uncertainty in those areas. Slightly larger SE values were present in the iDNA model in the two large patches of mostly unsuitable land between the two core areas. The Combined model appeared to reduce the levels of uncertainty and display a model that still indicates the presence of a corridor connecting the two core areas (Figure 38).

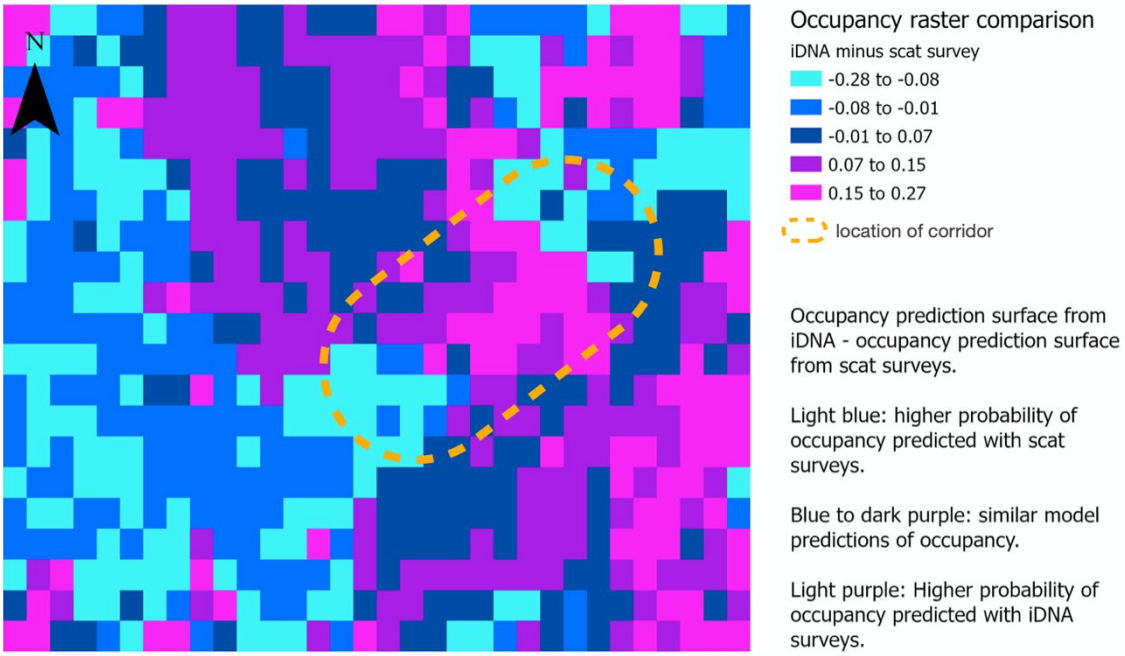


Figure 38. Raster comparison of the probability of occupancy predicted by iDNA data and scat survey data.

4.4 Discussion

4.4.1 Overview

The data presented in this chapter indicates that iDNA is a robust and complimentary survey tool for monitoring brown bears; that it may be more effective at identifying areas of low usage or population densities and that it provides data which can underpin the development of spatially explicit occupancy models. As the development of iDNA continues in different ecological survey scenarios, a small number of studies have compared it with more commonly used monitoring methods. (Gogarten et al., 2020) compared iDNA detection rates with camera trapping, finding that the methods are more effective at collecting information on species assemblages when used in combination. Their results suggest that the overlap of species detected using both methods ranged between 6%-43%, with iDNA being more successful at detecting smaller-bodied species those observed with camera traps (Gogarten et al., 2020). To date, however, published iDNA case studies primarily focus on multiple species detection using a metagenomic approach (Abrams et al., 2019; Hoffmann et al., 2018; Schnell et al., 2015; Schubert et al., 2015). The potential to use iDNA-derived data in occupancy modelling was proposed by previous literature (Baerholm Schnell et al.,

2015; Gogarten et al., 2020), suggesting that repeated sampling, combined with the understanding of dispersal distances and target species DNA persistence in or on the invertebrates is accounted for. To our knowledge, this is the first study which utilizes iDNA data in the development of spatially explicit occupancy models to draw conclusions about the distribution of a single species across a given geographical area. The data presented in this chapter demonstrates the use of iDNA in a statistically robust assessment of single-species, single-season occupancy.

4.4.2 iDNA as an effective single-species monitoring technique

As advised by (Gogarten et al., 2020), the estimation of fly dispersal and DNA persistence period in fly guts was necessary for this analysis, to ensure that the iDNA sampling sites were 'closed' for the duration of the surveys and no changes to occupancy occurred during sampling. In her PhD thesis, Lee described that blowflies disperse at a maximum of 1km per day and that mtDNA amplifications from iDNA samples significantly dropped within 24h (Lee, Sing and Wilson, 2015). Lee's findings, alongside the results from this project's controlled DNA retention experiment (Chapter 3), suggest that mtDNA of sufficient quantity and quality persists for a period of 24h in blowfly guts (Lee, Sing and Wilson, 2015). Furthermore, studies looking at blowfly dispersal distances in temperate and subtropical climates suggest a dispersal rate of 100-2400m per day (Braack and Retief, 1986; Lee et al., 2016; Tsuda et al., 2009; Wall, 1998). By sampling in the middle of a 2km x 2km grid, the distance between the two sampling points (if those grid squares happen to be adjacent) was 2km, which corresponded to the period of estimated time by which the majority of target species mtDNA in the iDNA sample would no longer be amplifiable. Using this information, I estimated that each iDNA sampling session was a separate event that recovered iDNA from within the boundaries of a "closed" sampling grid. Similarly, the scat data was spatially thinned to follow the same principle and account for multiple records of scat along the transects within each sampling grid.

It is evident that the two surveys complemented each other; there were many occasions when one of them detected bear presence when the other failed to do so (Figure 32). In fact, when looking at the detection data from all seasons, only 26.6% of the detections were recorded by both methods, while 33.3% were exclusively made by iDNA surveys and 40% were only recorded with scat surveys. Out of the 90 sampling events (30 sites

x 3 seasons), 67 sampling event records were consistent for detections/non detections between the two survey types (non detection by both methods = 45 sampling events, detection by both methods = 12 sampling events). Out of the thirty sampling sites across all seasons, four revealed presence of bears only in the scat dataset, while three revealed presence only in the iDNA dataset. Those detection differences came from a single repeat (one season's successful detection), with the exception of grid 14 and grid 29 where scat were recorded in two separate seasons (two repeats). The fact that the majority of the detection differences come from one of the survey repeats, could indicate areas of less frequent use by bears, decreasing the likelihood of detecting the species at the time of the survey. The results from this study add to the work of (Gogarten et al., 2020), expanding the body of research on using iDNA as a complementary monitoring method. This chapter introduces results that support the complementarity of iDNA with more commonly-used monitoring techniques on single-species detection.

4.4.3. Predictor importance

Each of the three models revealed a slightly different rank of predictor importance as a result of the Gini Index (GI) analysis (Table 9). The percentage cover of broadleaved forests was the strongest predictor of both the iDNA and the combined dataset and third most important predictor for the scat surveys. The other two habitat predictors that were highlighted by the GI analysis were percentage cover of natural grassland (scat and combined surveys), percentage cover of transitional woodland (scat and combined surveys), and percentage cover of non-irrigate arable land (iDNA and combined surveys). The results here are consistent with findings from van Gils et al. (2014) who showed that the summer/autumn distribution is better predicted by habitat types such as broadleaved woodland, abandoned farmland, meadows, pastures and residual forest patches. Particularly in the case of broadleaved woodland, the findings reflect habitat use in Spain were the percentage of deciduous forests cover was found to be a strong predictor of suitability for bears (Clevenger, Purroy and Campos, 1997; Clevenger, Purroy and Pelton, 1992; Güthlin et al., 2011). Indeed, the percentage broadleaved cover indicated a positive relationship to the probability of occupancy (Figure 34). Similarly, percentage cover of natural grassland was an important predictor of bear

presence, but here we saw a drop in occupancy probability as the percentage cover became greater, which points to the assumption that bears prefer these areas in landscapes where they are mixed with areas provide canopy cover (van Gils et al., 2014; Posillico et al., 2004). Mean elevation was one of the important predictors of the probability of occupancy for the scat and iDNA models. In the scat model, the probability of occupancy was predicted consistently high across elevations, while the iDNA model revealed an increase in occupancy probability as elevation increases, with a small drop and plateau after 1000m elevation. This could be a result of temperature differences in high elevation areas that made for a different composition of blowfly community, but it is consistent with bibliography for European brown bear altitude preferences in the sub-Mediterranean climate (Clevenger, Purroy and Campos, 1997; Falcucci et al., 2009; van Gils et al., 2014; Posillico et al., 2004). Additionally, the percentage cover of flat/gentle slopes showed negatively impacts the probability of occupancy in the scat and combined modes, indicating that bears tend to prefer steeper slopes instead of large extents of flat areas (Falcucci et al., 2009; van Gils et al., 2014; Güthlin et al., 2011).

My results suggest that detection probability for the iDNA and combined data models is positively affected by the percentage cover of canyon formations. Due to the fact that this specifically affected detection in the iDNA data which in turn influenced the combined model, this could be due to environmental reasons, such as scent dispersal. Nottingham, Johnson and Pelton (1973) present an example of the effects of scent dispersal in the detection of raccoons (*Procyon lotor*) on bottomland versus upland locations of scent stations. The same effect could be present in this study, with the scent of the fly bait essentially becoming contained by the canyon structures, allowing for the flies to detect and travel to it more effectively. Additionally, Macleod and Donnelly (1958) studied changes in the distribution of five blowfly species and demonstrated that flies often seek refuge in more sheltered areas (outlined as 'topographically conditioned shelter'). Canyons, forests and other 'shelter' type covariates used here could also explain some of the variation in detection. Additionally, detection was greater nearer major roads, which was unexpected from our understanding of bears. The negative impact of major roads, as well as other urban structures and human activity, to bear distribution has been covered thoroughly in Europe (Bartoń et al., 2019; van Gils et al., 2014; Güthlin et al., 2011) and North America (Brody and Pelton, 1989; Lewis et al.,

2011; Van Manen et al., 2012; Mattson, Knight and Blanchard, 1987). My results suggest that iDNA detection quickly drops after a mean distance from major roads of 250m and reaches a plateau at approximately 15% detection probability.

Transect length was found to be a good predictor for combined survey detection, confirming that the longer the forestry road network within the sampling grid, the more likely it is to detect bears. Finally, survey year was a strong predictor for the scat probability of detection, especially highlighting the lack of scat detected during the last season (2019). This reveals that, in the case of the scat surveys, it is likely that predictors outside the scope of this study were important for detecting bear presence, such as differences in the food resources and habitat usage within this particular geographical area across the three years, or other factors that were not measured here.

4.4.4. Encounter rate and detection probability

When assessing a binary presence-absence model, we can look at the model's sensitivity as the rate to which a probability of detection relates to a true field detection (true positive) and its specificity as the model's ability to match low detection probability values to non-detection events in the field (true absence). An ideal survey, with a detection probability of 1, would assume not only that the survey method is 100% successful at finding the species if it is there, but also 100% confident in confirming that the species is absent when it is not detected. However, this is not a realistic concept when it comes to most species monitoring efforts and assuming perfect detection can introduce bias and lead to underestimations of a species' spatial distribution (Mackenzie et al., 2002; Thompson, 2004).

This model's predictors were selected using a training dataset consisting of 80% of the total observation data (n=90 observations for each survey type) for each survey method. The rest of the data (test data) was kept aside and used for the validation of the random forest model that calculated the bear encounter rate. For each of the survey types the calibrated encounter models generated by the relatively small amount of training data produced models with an improved mean square error (MSE), showing the mean

squared difference between field encounter and estimated encounter but had no effect on the other model metrics (sensitivity, specificity, AUC and the Cohen's κ).

The calibrated scat model had the smallest mean squared difference between field and estimated encounter and highest AUC value, but the sensitivity and Cohen's κ suggested a moderate agreement between true and estimated encounter. Contrastingly, the iDNA model had a higher MSE and lower AUC, but the sensitivity and Kappa values were the highest across all models, suggesting substantial agreement between true and estimated encounter rate. Finally, the combined model had the lowest specificity value. This drop in specificity relative to the individual models suggests that the combined model is less likely to correctly identify true absences than the individual survey models. This could be explained by the wider range of values that go into the combined model when the data from the scat and iDNA detections are pooled. It is possible that where one survey was more effective in detecting bears than the other, especially when there was only one detection out of the six observations (three repeats for scat surveys and three repeats for iDNA sampling), the detection inconsistencies contribute to a model that is less effective in confirming true absences. This assumption is also supported by the combined model's MSE and Cohen's K values that present the highest and lowest values for those two parameters respectively, suggesting a moderate agreement between the model's true and estimated encounter rate. All three models are powerful in accurately predicting areas of species presence (sensitivity range = 0.667-0.833), as well as true absence and thus correctly identifying areas of species occurrence (specificity range = 0.636- 0.889) (Strimas-Mackey et al., 2020).

Visually, the three models result in very similar encounter rate predictions, with the combination model predictably showing a higher rate of encounter across the landscape since the effort the combined dataset comprises of more detection observations. The resemblance of the encounter models adds to our argument that iDNA is a robust monitoring method and, ultimately, the differences in the values of the encounter prediction metrics highlight the fact that the two survey methods were able to complement each other in detecting bears.

4.4.5 Occupancy modelling

Here, similar to the encounter rate models, the occupancy models (both optimised and comparison models) generated using the three datasets predicted a very similar landscape of bear distribution across the study area (Figure 37 and Appendix C2. Optimised occupancy models). The large difference in predictions seen in the optimised iDNA model was evidently introduced by the inclusion of the percentage of broadleaved woodland predictor and creating a model that predicted a much lower probability of bear occupancy along the study area, matched with very high standard error values. Once that predictor was removed, in the comparison models, the analysis generated a probability mosaic that matched that of the other two models. The discussion around this study's occupancy models will focus on the comparison models hereafter, to allow for a direct comparison of the outputs.

Interestingly, the iDNA and combined survey models describe a more continuous predicted corridor while still highlighting the complexity of habitat use, as confirmed by the rarity of observations across the two surveys and three seasons. The iDNA data here was able to illustrate a more permeable landscape, with more evidence of predicted corridor use as well as some higher predictions in core area edges and less suitable habitats (as seen in Figure 38). In contrast, the scat survey model, even though very similar, does not reveal a change in permeability to reflect the fact that presences at four sites were only detected with scat surveys. The ability of the iDNA model to reflect the more rare sightings and corridor use, combined with the high standard errors in the scat model's probability of occupancy, results in a combined model with an overall higher probability of occurrence and a lower standard error across the study site.

Occupancy modelling emerged out of the need to tackle imperfect detections in the field, which becomes especially important when it comes to understanding the distribution of rare or elusive species (Dorazio and Erickson, 2018; Elith et al., 2006; Keane et al., 2012). This chapter highlights the effectiveness of occupancy modelling in studying areas of low use, such as corridors and stepping stone habitats. The challenge in looking at corridors and other areas of low use is that these areas are often monitored to prove their use, but they are unlikely to be permanently inhabited by the species in

question. By definition, these habitats are areas the species uses to cross between core areas, therefore a drop in detection probability similar to rare or elusive species monitoring efforts can be expected, making occupancy modelling a more suitable method of estimating spatial distribution of a species. This study demonstrates the use of occupancy modelling in monitoring predicted corridors using brown bears in Greece as a case study. More importantly, it highlights the use of iDNA in this effort by comparing it with the occupancy model produced by scat survey observations.

The model created using iDNA detections demonstrates a marginally more connected network of bear habitats and illustrates the pattern of bear use of the corridor as seen in the telemetry data (see Chapter 2). In SDMs sampling bias can be introduced in the form of false absences by selecting background points away from the sampling area and thus risking training the model in discounting suitable areas where no detections were recorded (Elith and Leathwick, 2009; Fourcade et al., 2014). This case study demonstrates how, in addition to false absences, occupancy models generated from small datasets can be influenced by detections in 'atypical' areas, breaking the distribution pattern of the entire survey dataset by a single observation of a single repeat. The three occurrences that only iDNA was effective in detecting, two of which located in habitats outside core areas, meant that the resulting model further highlights the corridor existing between the two core areas and emphasises the corridor use by bears (see Figure 32). The possibility that bear detection in these areas was a result of flies dispersing from core to corridor areas instead of bear use of the corridor areas was considered. However, the grid size chosen along with the experiments of target DNA persistence in these iDNA samples suggests that the presence of bears in the sampling grid. Moreover, if flies were dispersing from core areas, that dispersal would be equal to corridor and non-corridor areas, suggesting that iDNA would have detected bear presence in other areas outside the corridor zone. The fact that these detections were confined to the areas that were predicted as suitable supports the argument that the detections are a real effect of bear presence.

Contrastingly, the fact that scat surveys located observations in core areas that iDNA failed to detect generated a model that seems to be slightly more restricted to the core areas (differences illustrated in Figure 38). Scat surveys showed a higher resilience to

error in core areas, but higher error values in the margins between high and low occurrence probability. In both surveys a pattern whereby the number of detections increases with predicted suitability emerged, indicates that both methods could be used to ground-truth suitability predictions effectively. This decrease in detections as areas become less suitable for bears shows that the methods appear to be sensitive to changes in habitat use.

4.4.5 Conclusion

The results of the occupancy modelling analysis are very encouraging for the future of iDNA monitoring in ecological network studies. Using iDNA data to make species distribution inferences is novel and this chapter is the first work describing it and comparing it to a more commonly used method. The results of this study demonstrate that iDNA can robustly predict occupancy and detection given a set of repeat surveys. It is important to reiterate that the results presented in this chapter do not, by any means, suggest that iDNA could solely replace other monitoring efforts. Instead, we hoped to unveil the benefits and disadvantages of this method and the compatibility of using iDNA in conjunction with other survey methods. The models above present a very interesting case study, illustrating the advantages and pitfalls of using iDNA in collecting species records for spatial analyses.

Furthermore, this work was carried out to explore the potential of using this novel method to monitor wildlife corridors and areas of expected low use. The case study demonstrates the use of iDNA as a single-species monitoring method for exploring the existence and use of wildlife corridors for the species in question. The results suggest that iDNA can be effective in detecting the species in areas of infrequent use. Caution must be taken in interpreting those results, as models that incorporate observations from less suitable habitats are more likely to have increased levels of over-prediction and uncertainty in occupancy predictions. We believe that the occurrence data and resulting occupancy models show some evidence of the strength of iDNA in detecting target species in areas of low use and rare occurrences.

Given the extended body of literature around bear habitat preferences and distribution predictors, we were able to confirm that our results concerning the iDNA encounter rate

and occupancy modelling reflected current understanding of bear ecology. The iDNA models and response curves were able to show the expected patterns of bear distribution, confirmed by similar responses in the scat model results. This consistency in the model outputs validates our efforts to create an accurate model for bears and the hypothesis that iDNA can be a successful tool not only for monitoring, but also for the spatial analysis of the species' distribution and habitat preferences. More importantly, this uniformity in model outputs suggests that this method could be a more robust way to assess the occupancy of a species that less is known about by implementing repeat survey sampling using iDNA and a supplementary survey method. By introducing a second survey technique, alternative to the iDNA, our results demonstrate that it is possible to further validate the results where the understanding of what to expect for the species' ecological preferences lacks.

Chapter 5: Discussion

Successful conservation projects are centred around species monitoring, from the very early stages of getting a grasp of the distribution, and population numbers and trends of a species in an area, to evaluating management action (Legg and Nagy, 2006; Nichols and Williams, 2006; Stem et al., 2005). Optimising the collection of data in the field has been in the forefront of new habitat and species monitoring advancements, with emerging technology being adapted for species surveys, such as the use of drones (Shewring and Vafidis, 2021; Vafidis et al., 2021; Wich and Koh, 2018), bioacoustic and chemical sensors (Bardeli et al., 2010; Calupca, Fristrup and Clark, 2000; Larsson and Svensson, 2009; Salamon et al., 2016; Svensson et al., 2011) and environmental DNA monitoring techniques (iDNA eDNA airDNA) (Abrams et al., 2018; Bohmann et al., 2014; Bohmann, Schnell and Gilbert, 2013; Calvignac-Spencer et al., 2013; Lynggaard et al., 2022; Ruppert, Kline and Rahman, 2019). Following such efforts, the overall work undertaken in this project has focused on exploring the use of cost- and time-efficient tools that can be used in conservation planning and species monitoring. The three broad techniques I present here describe a complete monitoring project, from the initial planning stages and understanding the distribution have habitat use for a species (habitat suitability modelling), to conducting field surveys (iDNA monitoring) and using repeat survey results to estimate the distribution of the species across the study area (occupancy modelling) (Figure 39).

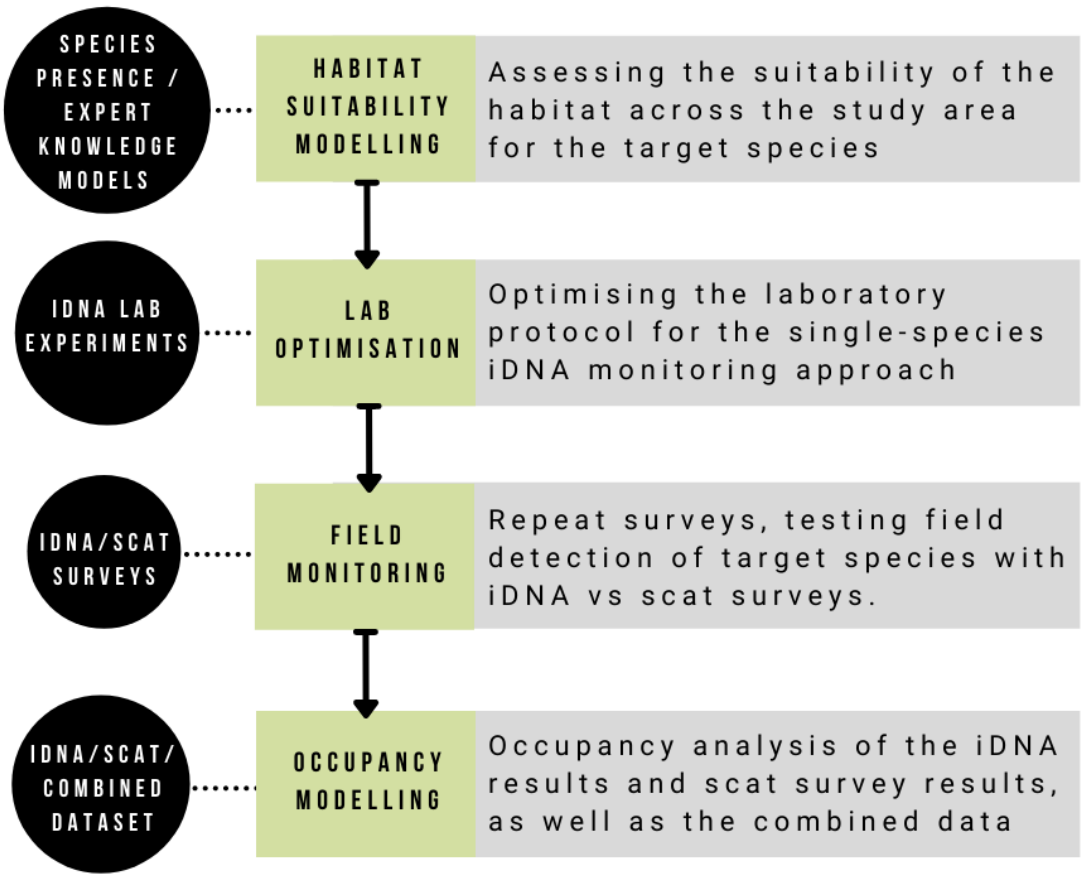


Figure 39. Overall workflow and project skeleton, detailing the outcomes of each part of the project.

The protocol outlined here takes into consideration cases where knowledge of the species ecology is lacking, or existing field data is sparse, ensuring that the model is easily transferable to species that are more elusive or less studied. The two methods used to create the Habitat Suitability Models (HSM) were tailored to these two scenarios of missing data. Both models were able to create distribution scenarios that follow the known target species' distribution, while also considering sites currently unoccupied that fulfil the same habitat requirements. Within the context of creating a monitoring protocol, the HSM study was conducted to reveal areas that would optimise the survey results and consequent understanding of the iDNA limits of detecting the target species in areas of low, medium and high use. The HSM analysis assisted, at the planning stages of the monitoring part of the project, in finding a study area that fulfilled all these requirements in terms of habitat type, while also incorporating a potential corridor, core habitats (expected high use) and unsuitable areas. The HSM analysis was instrumental to the selection of a study area that incorporated all the suitability gradient to allow for

a successful test of the iDNA method as well as provide evidence of corridor use by the species.

Keeping in mind the transferability of the method and focus on species that are more challenging to detect, the study focused on single-species monitoring that has proven in eDNA and iDNA studies to be more effective than metabarcoding at detecting target species when they are rare (Harper et al., 2018; Schubert et al., 2015). I believe that the protocol presented here could be used in this format, with small adjustments, to adapt and use in species monitoring of other target animals or in other climates. In fact, the targeted qPCR iDNA protocol developed here was adapted and used to collect and analyse iDNA samples from sun bears (*Helarctos malayanus*) in Cambodia in collaboration with the Free the Bears non-profit charity and funded by the International Association for Bear Research and Management Conservation and Research Grant. The pilot results from sun bear sanctuaries confirmed presence of sun bear DNA in fly samples and more research is currently underway to analyse and compare the targeted qPCR results with metabarcoding results from field samples (data pending). We believe that the single-species approach used in this project has great potential for monitoring efforts of rare or elusive species such as the sun bear and other, more cryptic species in a variety of environments. The ease of adaptation of this method using different traps, bait types and groups of invertebrates has been demonstrated in literature, even in the use of mixed invertebrate group samples (Bohmann, Schnell and Gilbert, 2013; Calvignac-Spencer et al., 2013; Lynggaard et al., 2019).

As iDNA continues to develop as a monitoring tool, studies looking at the nuanced information, such as limit of detection tolerances, optimising the qPCR reagents and DNA extraction techniques, and revealing the detection tolerances for each group of invertebrates (persistence period of amplifiable target species DNA) will be essential to move the field forward. Understanding such details will help create a more comprehensive body of research in the use of iDNA for wildlife monitoring, showcase the advantages and sensitivities of the method, and highlight some of the limitations. Although individual identification has not had a sufficient rate of success in the past (Schubert et al., 2015), further research could focus on adapting this iDNA protocol for population structure studies by changing the target gene. The mitochondrial D-loop

region can reveal genetic diversity in phylogeographic and demographic variation within populations (Jia et al., 2007; Mereu et al., 2019; Osman, Yonezawa and Nishibori, 2016; Wu et al., 2015). Adapting this iDNA protocol to target and sequence the mitochondrial D-loop region could introduce some variation between individuals and would greatly broaden the scope of this monitoring method.

Finally, a comparison survey method, repeat surveys and occupancy modelling ensured that the efforts to optimise this technique would be tested in the field and put under a modelling analysis that accounts for small sample sizes and rare detections. A model average was created using the Expert Knowledge model and corresponding Species Presence model (both made using elevation, topography and land cover variables), to illustrate the results of creating a combined prediction of these two different methods (Figure 40). The results of the occupancy modelling analysis very closely resemble those of the Habitat Suitability Index model that incorporated elevation, topography and land cover data (Figure 40). Overall, the creation of the HSM models was significant in the study site selection, but it is the occupancy modelling that explicitly focused on bear distribution in this area and therefore the occupancy results can be subsequently used to ground-truth the HSMs. The results indicate that the Expert Knowledge model at this particular study area was more effective at predicting the distribution of the target species and describing the use of the habitat by the animal.

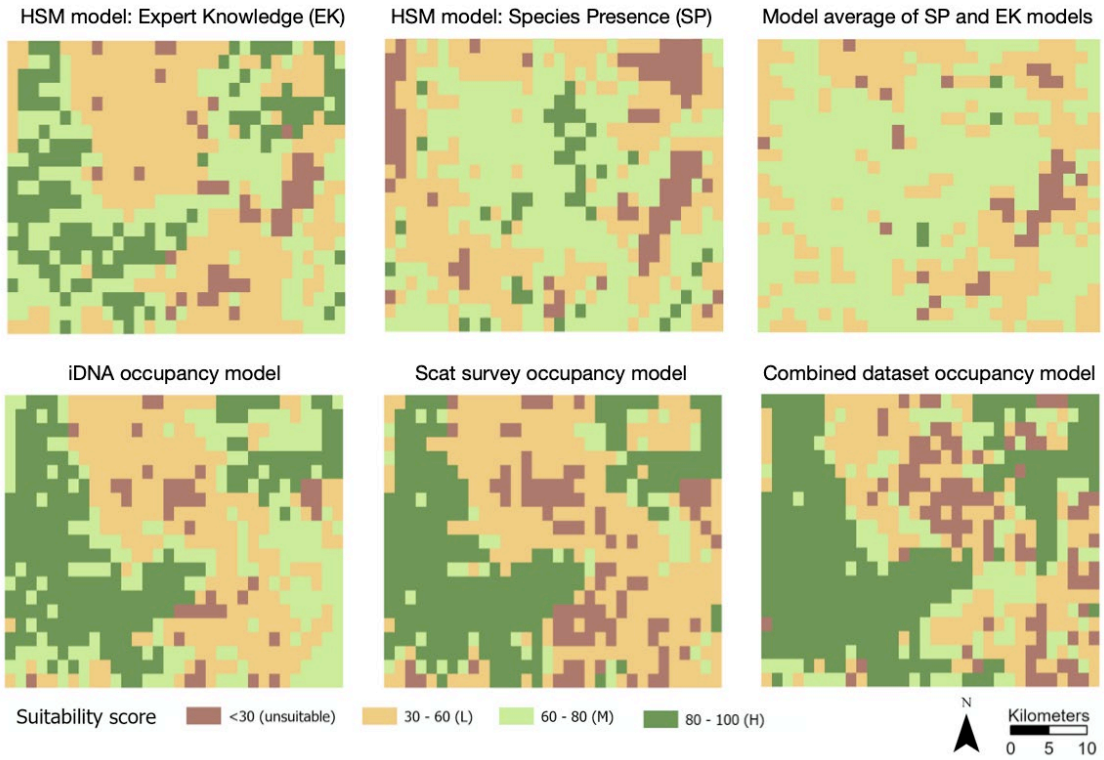


Figure 40. Model predictions of suitability (top) compared with the occupancy modelling outputs. The Expert Knowledge and Species Presence models were averaged to create a model that combines the two very different methods to test the potential that these two could be used in combination.

In regards to the impact of this study to the target species, I believe that this project has potential to impact bear conservation planning in Greece, while also introducing iDNA monitoring as a new method to survey bears. This study demonstrates the advantages and challenges in using iDNA for single species monitoring and highlights excellent potential for expanding species monitoring efforts within noninvasive survey techniques. I believe that iDNA can be used as a complementary method to other survey techniques, as this study demonstrates its ability to detect bears in areas where scat surveys failed to yield results. Furthermore, the spatial analysis carried out presents the first focused attempt to map the entire distribution of bears in Greece and potential areas of future colonisation. This information will be used to understand how best to conserve the areas that are already important for bears, while also improving and maintaining linkages between them. As bears continue to increase in numbers and expand their range, the suitable but currently not inhabited areas on the HSMs could serve as guides to where bears will move towards as they recolonise the land. This information is invaluable for large carnivore conservation efforts to prevent conflict and raise awareness in these communities that have not interacted with bears before. In

fact, human-bear conflict continues to be the largest conservation threat to bears in Greece and the costs related to all damages caused by bears amount to over £119,000 annually (Kaczensky et al., 2012b), making the understanding of future bear movement essential to their long-term conservation.

This work was conducted with the aim to serve as a baseline case study for the planning, implementation, analysis and interpretation of iDNA surveys, and so I hope that this thesis can be used as a protocol for monitoring efforts that incorporate single species iDNA monitoring.

Figure 41. Brown bear surrounded by mosquitoes and other invertebrates. Photo taken by Drew Hamilton and used with his permission.

References:

Abrams, J. et al. (2018) Shifting up a gear with iDNA: from mammal detection events to standardized surveys. *bioRxiv*. [online].

Abrams, J.F. et al. (2019) Shifting up a gear with iDNA: From mammal detection events to standardised surveys. *Journal of Applied Ecology*. [online]. 56 (7), Blackwell Publishing Ltd, pp.1637–1648.

Addicott, J.F. et al. (1987) Ecological Neighborhoods: Scaling Environmental Patterns. *Source: Oikos*. 49 (3), pp.340–346.

Aiello-Lammens, M.E. et al. (2015) spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*. [online].

Aksoy, E. et al. (2014) Analysis of multiple tsetse fly populations in Uganda reveals limited diversity and species-specific gut microbiota. *Applied and Environmental Microbiology*. [online]. 80 (14), American Society for Microbiology, pp.4301–4312. Available from: https://journals.asm.org/doi/abs/10.1128/AEM.00079-14 [Accessed 21 January 2022].

Almpanidou, V. et al. (2014) Providing insights on habitat connectivity for male brown bears: A combination of habitat suitability and landscape graph-based models. *Ecological Modelling*. [online]. 286.

Anderson, R.P. et al. (2011) Ecological Niches and Geographic Distributions (MPB-49). *Ecological Niches and Geographic Distributions (MPB-49)*. [online]. Princeton University Press.

Anderson, R.P. and Gonzalez, I. (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*. [online]. 222 (15), Elsevier, pp.2796–2811.

Anderson, R.P. and Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: Preliminary tests with montane rodents (genus Nephelomys) in Venezuela. *Journal of Biogeography*. [online].

Araújo, M.B. et al. (2019) Standards for distribution models in biodiversity assessments. *Science Advances*. [online]. 5 (1), American Association for the Advancement of Science, p.eaat4858.

Araújo, M.B. and Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*. [online]. 33 (10), John Wiley & Sons, Ltd, pp.1677–1688. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2699.2006.01584.x [Accessed 20 September 2021].

Asghar, U. et al. (2015) DNA Extraction from Insects by Using Different Techniques: A Review. *Advances in Entomology*. [online]. 03 (04), Scientific Research Publishing, pp.132–138. Available from: http://file.scirp.org/Html/ [Accessed 31 October 2021].

Austin, J.E. et al. (2000) Effects of roadside transect width on waterflowl and wetland estimates. *WETLANDS*. 20 (4), pp.660–670.

Baerholm Schnell, I. et al. (2015) Landscape requirements of a primate population in a human-dominated environment. [online]. Available from: https://frontiersinzoology.biomedcentral.com/track/pdf/10.1186/s12983-015-0115-z?site=frontiersinzoology.biomedcentral.com [Accessed 23 January 2018].

Barbet-Massin, M. et al. (2012) Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*. [online].

Bardeli, R. et al. (2010) Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. *Pattern Recognition Letters*. [online]. 31 (12), North-Holland, pp.1524–1534.

Barnes, M.A. et al. (2014) Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science and Technology*. [online]. 48 (3), pp.1819–1827.

Barnes, M.A. and Turner, C.R. (2015) The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*. [online]. 17 (17), pp.1–17.

Bartoń, K.A. et al. (2019) Bears without borders: Long-distance movement in human-dominated landscapes. *Global Ecology and Conservation*. [online]. 17, Elsevier, p.e00541.

Beebe, S.C., Howell, T.J. and Bennett, P.C. (2016) Using Scent Detection Dogs in Conservation Settings: A Review of Scientific Literature Regarding Their Selection. *Frontiers in Veterinary Science*. [online]. 3.

Beier, J.C. (2003) Malaria parasite development in mosquitoes. *http://dx.doi.org/10.1146/annurev.ento.43.1.519*. [online]. 43, Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA , pp.519–543. Available

from: https://www.annualreviews.org/doi/abs/10.1146/annurev.ento.43.1.519 [Accessed 5 December 2021].

Bekoff, M. and Jamieson, D. (2019) Ethics and the study of carnivores: Doing science while respecting animalsIn: Gittleman, J.L. (ed.) *Carnivore Behavior, Ecology, and Evolution*. [online]. Temple University Press, 15–45.

Bellamy, C., Scott, C. and Altringham, J. (2013) Multiscale, presence-only habitat suitability models: fine-resolution maps for eight bat species. *Journal of Applied Ecology*. [online]. 50 (4), John Wiley & Sons, Ltd, pp.892–901. Available from: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.12117 [Accessed 24 August 2021].

Bellemain, E. et al. (2005) Estimating Population Size of Elusive Animals with DNA from Hunter-Collected Feces: Four Methods for Brown Bears. *Conservation Biology*. [online]. 19 (1), John Wiley & Sons, Ltd, pp.150–161. Available from: https://conbio.onlinelibrary.wiley.com/doi/full/10.1111/j.1523-1739.2005.00549.x [Accessed 8 July 2021].

Bennet, A.F. (1999) *Linkages in the landscape: the role of corridors and connectivity in wildlife conservation . Iucn. IUCN*. [online]. Available from: https://books.google.co.uk/books?hl=en&lr=&id=hkkXEpcWOOgC&oi=fnd&pg=PR9&dq=Linkages+in+the+landscape:+the+role+of+corridors+and+connectivity+in+wildlife+conservation.&ots=12UAbRY6Sr&sig=Vqe6ZroIpjO3OzcqXM0r_5iW5Pc#v=onepage&q=Linkages in the landscape%3A [Accessed 11 July 2021].

Boenn, M. (2020) *POD: Probability of Detection for Qualitative PCR Methods* [online]. Available from: https://cran.r-project.org/package=POD.

Bohmann, K. et al. (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*. [online]. 29 (6), Elsevier Current Trends, pp.358–367.

Bohmann, K., Schnell, I.B.B. and Gilbert, M.T.P. (2013) When bugs reveal biodiversity. *Molecular Ecology*. [online]. 22 (4), John Wiley & Sons, Ltd, pp.909–911. Available from: http://doi.wiley.com/10.1111/mec.12221 [Accessed 23 January 2018].

Boitani, L. and Linnell, J.D.C. (2015) *Bringing Large Mammals Back: Large Carnivores in Europe Rewilding European Landscapes*. [online]. Cham, 67–84. Available from: http://link.springer.com/10.1007/978-3-319-12039-3_4 [Accessed 27 June 2017].

Bojarska, K. and Selva, N. (2012) Spatial patterns in brown bear Ursus arctos diet: The role of geographical and environmental factors. *Mammal Review*. [online]. 42 (2), Blackwell Publishing Ltd, pp.120–143.

Boria, R.A. et al. (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*. [online].

Boulanger, J. et al. (2008) Multiple data sources improve DNA-based mark-recapture population estimates of grizzly bears. *Ecological Applications*. [online]. 18 (3), John Wiley & Sons, Ltd, pp.577–589.

Braack, L.E. and Retief, P.F. (1986) Dispersal, density and habitat preference of the blow-flies Chrysomyia albiceps (Wd.) and Chrysomyia marginalis (Wd.) (Diptera: Calliphoridae). *The Onderstepoort journal of veterinary research*. 53 (1), pp.13–8.

Braschler, B. et al. (2020) Realised rather than fundamental thermal niches predict site occupancy: Implications for climate change forecasting. *Journal of Animal Ecology*. [online]. 89 (12), Blackwell Publishing Ltd, pp.2863–2875. Available from: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2656.13358 [Accessed 27 June 2021].

Brody, A.J. and Pelton, M.R. (1989) Effects of roads on black bear movements in western North Carolina. *Wildlife Society Bulletin (1973-2006)*. [online]. 17 (1), pp.5–10.

Brooke, Z.M. et al. (2014) Correlates of Research Effort in Carnivores: Body Size, Range Size and Diet Matter. *PLOS ONE*. [online]. 9 (4), Public Library of Science, p.e93195.

Brown, J.L. (2014) SDMtoolbox: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*. [online].

Brown, J.L., Bennett, J.R. and French, C.M. (2017) SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ*. [online].

Browne, C.M., Stafford, K.J. and Fordham, R.A. (2015) The detection and identification of tuatara and gecko scents by dogs. *Journal of Veterinary Behavior: Clinical Applications and Research*. [online]. 10 (6).

Brys, R. et al. (2021) Reliable eDNA detection and quantification of the European weather loach (Misgurnus fossilis). *Journal of Fish Biology*. [online]. 98 (2), Blackwell Publishing Ltd, pp.399–414.

Burns, M. and Valdivia, H. (2008) Modelling the limit of detection in real-time quantitative PCR. *European Food Research and Technology*. [online].

Bustin, S.A. et al. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*. [online]. 55 (4), Oxford Academic, pp.611–622. Available from: https://academic.oup.com/clinchem/article/55/4/611/5631762 [Accessed 6 December 2020].

Calupca, T.A., Fristrup, K.M. and Clark, C.W. (2000) A compact digital recording system for autonomous bioacoustic monitoring. *The Journal of the Acoustical Society of America*. [online]. 108 (5), Acoustical Society of AmericaASA, p.2582. Available from: https://asa.scitation.org/doi/abs/10.1121/1.4743595 [Accessed 18 January 2022].

Calvignac-Spencer, S. et al. (2013) An invertebrate stomach's view on vertebrate ecology: Certain invertebrates could be used as "vertebrate samplers" and deliver DNA-based information on many aspects of vertebrate ecology. *BioEssays*. [online]. 35 (11), pp.1004–1013.

Carvalho, C. et al. (2021) Efficiency of eDNA and iDNA in assessing vertebrate diversity and its abundance. *Authorea Preprints*. [online]. Authorea. Available from: https://www.authorea.com/doi/full/10.22541/au.161777019.96557508?commit=24882b9ec11c924510d7c0dcd45debcd26fc371e [Accessed 30 October 2021].

Chapron, G. et al. (2014) Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science*. [online]. 346 (6216), pp.1517–1519. Available from: http://www.sciencemag.org/lookup/doi/10.1126/science.1257553 [Accessed 22 January 2018].

Chauvier, Y. et al. (2021) Influence of climate, soil, and land cover on plant species distribution in the European Alps. *Ecological Monographs*. [online]. 91 (2), John Wiley & Sons, Ltd, p.e01433. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ecm.1433 [Accessed 16 January 2022].

Chen, C. and Liaw, A. (2004) Using Random Forest to Learn Imbalanced Data. *University of California, Berkeley*. 110 (1–12), p.24.

Chinnery, P.F. and Hudson, G. (2013) Mitochondrial genetics. *British Medical Bulletin*. [online]. 106 (1), Oxford University Press, p.135. Available from: /pmc/articles/PMC3675899/ [Accessed 22 March 2022].

Chruszcz, B. et al. (2003) Relationships among grizzly bears, highways, and habitat in the Banff-Bow Valley, Alberta, Canada. [online].

Cirovic, D. et al. (2015) Home range, movements, and activity patterns of a brown bear in Serbia. *Ursus*. [online]. 26 (2), pp.79–85.

Clevenger, A.P. (2012) Mitigating Continental-Scale Bottlenecks: How Small-Scale Highway Mitigation Has Large-Scale Impacts. *Ecological Restoration*. [online]. 30 (4), University of Wisconsin Press, pp.300–307. Available from: http://er.uwpress.org/content/30/4/300 [Accessed 17 January 2022].

Clevenger, A.P., Purroy, F.J. and Campos, M.A. (1997) Habitat assessment of a relict brown bear Ursus arctos population in northern Spain. *Biological Conservation*. [online]. 80 (1), Elsevier, pp.17–22.

Clevenger, A.P., Purroy, F.J. and Pelton, M.R. (1992) Brown bear (Ursus arctos L.) habitat use in the Cantabrian Mountains, Spain. *Mammalia*. [online]. 56 (2), De Gruyter, pp.203–214.

Clevenger, A.P., Purroy, F.J. and Pelton, M.R. (1990) Movement and Activity Patterns of a European Brown Bear in the Cantabrian Mountains, Spain. *Bears: Their Biology and Management*. 8, pp.205–211.

Comte, L. and Grenouillet, G. (2013) Species distribution modelling and imperfect detection: comparing occupancy versus consensus methods. *Diversity and Distributions*. [online]. 19 (8), John Wiley & Sons, Ltd, pp.996–1007. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/ddi.12078 [Accessed 31 July 2021].

Condro, A.A. et al. (2021) Predicting Hotspots and Prioritizing Protected Areas for Endangered Primate Species in Indonesia under Changing Climate. *Biology 2021, Vol. 10, Page 154*. [online]. 10 (2), Multidisciplinary Digital Publishing Institute, p.154. Available from: https://www.mdpi.com/2079-7737/10/2/154/htm [Accessed 16 January 2022].

Correa Ayram, C.A. et al. (2016) Habitat connectivity in biodiversity conservation: A review of recent studies and applications. *Progress in Physical Geography*. [online].

Crespo-Gascón, S. et al. (2019) The role of the spectacled bear (Tremarctos ornatus) as an umbrella species for Andean ecoregions. *Wildlife Research*. [online]. 46 (2), CSIRO PUBLISHING, pp.176–183. Available from: https://www.publish.csiro.au/wr/WR18056 [Accessed 17 October 2021].

Cutajar, T.P. and Rowley, J.J.L. (2020) Surveying frogs from the bellies of their parasites: Invertebrate-derived DNA as a novel survey method for frogs. *Global Ecology and Conservation*. [online]. 22, Elsevier, p.e00978.

Dahle, B., Støen, O.-G. and Swenson, J.E. (2006) Factors Influencing Home-Range Size in Subadult Brown Bears. *Journal of Mammalogy*. [online]. 87 (5), Oxford Academic, pp.859–865. Available from: https://academic.oup.com/jmammal/article/87/5/859/898771 [Accessed 11 July 2021].

Dai, Y. et al. (2021) Climate and land use changes shift the distribution and dispersal of two umbrella species in the Hindu Kush Himalayan region. *Science of The Total Environment*. [online]. 777, Elsevier, p.146207.

Deiner, K. et al. (2015) Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation*. [online]. 183, Elsevier, pp.53–63.

Delisle, I. and Strobeck, C. (2002) Conserved Primers for Rapid Sequencing of the Complete Mitochondrial Genome from Carnivores, Applied to Three Species of Bears. *Molecular Biology and Evolution*. [online]. 19 (3), pp.357–361. Available from: http://academic.oup.com/mbe/article/19/3/357/981077 [Accessed 10 May 2019].

Díaz, C. et al. (2020) Aquatic suspended particulate matter as source of eDNA for fish metabarcoding. *Scientific Reports 2020 10:1*. [online]. 10 (1), Nature Publishing Group, pp.1–8. Available from: https://www.nature.com/articles/s41598-020-71238-w [Accessed 31 October 2021].

Dorazio, R.M. and Erickson, R.A. (2018) ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology Resources*. [online]. 18 (2), Blackwell Publishing Ltd, pp.368–380. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/1755-0998.12735 [Accessed 27 June 2021].

Drinkwater, R. et al. (2021) Dung beetles as samplers of mammals in Malaysian Borneo—a test of high throughput metabarcoding of iDNA. *PeerJ*. [online]. 9, PeerJ Inc., p.e11897. Available from: https://peerj.com/articles/11897 [Accessed 30 October 2021].

Dudik, M. and Phillips, S.J. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*.

Egan, S.P. et al. (2013) Rapid Invasive Species Detection by Combining Environmental DNA with Light Transmission Spectroscopy. *Conservation Letters*. [online]. 6 (6), John Wiley & Sons, Ltd, pp.402–409. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/conl.12017 [Accessed 31 October 2021].

Elfström, M. et al. (2014a) Do Scandinavian brown bears approach settlements to obtain high-quality food? *Biological Conservation*. [online]. 178, Elsevier, pp.128–135.

Elfström, M. et al. (2014b) Ultimate and proximate mechanisms underlying the occurrence of bears close to human settlements: review and management implications. *Mammal Review*. [online]. 44 (1), John Wiley & Sons, Ltd, pp.5–18. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2907.2012.00223.x [Accessed 12 July 2021].

Elfström, M., Swenson, J.E. and Ball, J.P. (2008) Selection of denning habitats by Scandinavian brown bears Ursus arctos. *Wildlife Biology*. [online]. 14 (2), pp.176–187. Available from: https://www.researchgate.net/publication/232664101 [Accessed 16 January 2022].

Elith, J. et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*. [online]. 29 (2), John Wiley & Sons, Ltd, pp.129–151. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.2006.0906-7590.04596.x [Accessed 31 July 2021].

Elith, J. and Graham, C.H. (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*. [online]. 32.

Elith, J. and Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*. [online].

Erzinçlioğlu, Zakaria. (1996) *Blowflies* [online]. Published for the Company of Biologists by the Richmond Pub. Co. Available from: https://www.pemberleybooks.com/product/blowflies/1898/ [Accessed 28 March 2019].

Fahrig, L. (1997) Relative Effects of Habitat Loss and Fragmentation on Population Extinction. *Source: The Journal of Wildlife Management*. 61 (3), pp.603–610.

Falcucci, A. et al. (2009) Assessing habitat quality for conservation using an integrated occurrence-mortality model. *Journal of Applied Ecology*. [online]. 46 (3), John Wiley & Sons, Ltd, pp.600–609. Available from: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2664.2009.01634.x [Accessed 31 July 2021].

Farell, E.M. and Alexandre, G. (2012) *Bovine serum albumin further enhances the effects of organic solvents on increased yield of polymerase chain reaction of GC-rich templates* [online]. Available from: http://www.biomedcentral.com/1756-0500/5/257 [Accessed 15 April 2019].

Farrell, L.E., Roman, J. and Sunquist, M.E. (2000) Dietary separation of sympatric carnivores identified by molecular analysis of scats. *Molecular Ecology*. [online]. 9 (10), John Wiley & Sons, Ltd, pp.1583–1590.

Ficetola, G.F. et al. (2015) Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*. [online]. 15 (3), Blackwell Publishing Ltd, pp.543–556.

Fiske, I. and Chandler, R. (2019) *Overview of Unmarked: An R Package for the Analysis of Data from Unmarked Animals*

Fiske, I.J. and Chandler, R.B. (2011) Unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance. *Article in Journal of Statistical Software*. [online]. 43.

Foote, A.D. et al. (2012) Investigating the Potential Use of Environmental DNA (eDNA) for Genetic Monitoring of Marine Mammals. *PLOS ONE*. [online]. 7 (8), Public Library of Science, p.e41781. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0041781 [Accessed 31 October 2021].

Fouhy, F. et al. (2015) The Effects of Freezing on Faecal Microbiota as Determined Using MiSeq Sequencing and Culture-Based Investigations. *PLOS ONE*. [online]. 10 (3), Public Library of Science, p.e0119355. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0119355 [Accessed 22 March 2022].

Fourcade, Y. et al. (2014) Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods

for Correcting Sampling Bias. Valentine, J.F. (ed.) *PLoS ONE*. [online]. 9 (5), Public Library of Science, p.e97122.

Furlan, E.M. et al. (2016) A framework for estimating the sensitivity of eDNA surveys. *Molecular Ecology Resources*. [online]. 16 (3), John Wiley & Sons, Ltd (10.1111), pp.641–654. Available from: http://doi.wiley.com/10.1111/1755-0998.12483 [Accessed 23 January 2019].

Gariepy, T.D. et al. (2012) Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks. *Molecular Ecology Resources*. [online]. 12 (4), pp.646–652.

Gastón, A. et al. (2017) Species' habitat use inferred from environmental variables at multiple scales: How much we gain from high-resolution vegetation data? *International Journal of Applied Earth Observation and Geoinformation*. [online]. 55, pp.1–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0303243416301775 [Accessed 4 February 2019].

Gau, R.J. et al. (2004) Movements of Subadult Male Grizzly Bears, Ursus arctos, in the Central Canadian Arctic. *The Canadian Field-Naturalist*. [online]. 118 (2), Canadian Field-Naturalists' Club, pp.239–242.

Giannakopoulos, Al., Akriotis, T. and Mertzanis, Y. (2011) Spatio-temporal interactions in relation. *Submitted*.

van Gils, H. et al. (2012) Fine resolution distribution modelling of endemics in Majella National Park, Central Italy. *Plant Biosystems*. [online]. 146, Taylor & Francis, pp.276–287. Available from: https://www.tandfonline.com/doi/abs/10.1080/11263504.2012.685194 [Accessed 2 August 2021].

van Gils, H. et al. (2014) Where the bears roam in Majella National Park, Italy. *Journal for Nature Conservation*. [online]. 22 (1), Urban & Fischer, pp.23–34. Available from: https://www.sciencedirect.com/science/article/pii/S1617138113000782 [Accessed 14 March 2018].

Gini, C. (1921) Measurement of Inequality of Incomes. *The Economic Journal*. [online]. 31 (121), Oxford University Press (OUP), p.124.

Gogarten, J.F. et al. (2020) Fly-derived DNA and camera traps are complementary tools for assessing mammalian biodiversity. *Environmental DNA*. [online]. 2 (1), Wiley, pp.63–76.

Gold, Z. et al. (2021) eDNA metabarcoding as a biomonitoring tool for marine protected areas. *PLOS ONE*. [online]. 16 (2), Public Library of Science, p.e0238557. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0238557 [Accessed 11 January 2022].

Goldberg, C.S. et al. (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*. [online]. 7 (11), John Wiley & Sons, Ltd, pp.1299–1307. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.12595 [Accessed 5 December 2021].

Goossens, B. et al. (2000) A multi-samples, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. *Conservation Genetics*. 1, pp.157–162.

Gotelli, N.J. (2003) Predicting Species Occurrences: Issues of Accuracy and Scale. *The Auk*. [online].

Groombridge, B. and Jenkins, M.D. (2000) *Global Biodiversity: Earth's living resources in the 21st century.* World Conservation Press.

Gu, W. and Swihart, R.K. (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. *Biological Conservation*. [online]. 116 (2), Elsevier BV, pp.195–203.

Guan, X. et al. (2019) Environmental DNA (eDNA) Assays for Invasive Populations of Black Carp in North America. *Transactions of the American Fisheries Society*. [online]. 148 (6), John Wiley & Sons, Ltd, pp.1043–1055. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/tafs.10195 [Accessed 19 January 2022].

Guillera-Arroita, G. et al. (2015) Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*. [online]. 24 (3), Blackwell Publishing Ltd, pp.276–292. Available from: http://wileyonlinelibrary.com/journal/geb [Accessed 7 June 2021].

Guisan, A. and Thuiller, W. (2005) Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*. [online]. 8 (9), pp.993–1009.

Gupta, A.K. et al. (2014) Molecular phylogenetic profiling of gut-associated bacteria in larvae and adults of flesh flies. *Medical and Veterinary Entomology*. [online]. 28 (4), John Wiley & Sons, Ltd, pp.345–354. Available from:

https://onlinelibrary.wiley.com/doi/full/10.1111/mve.12054 [Accessed 21 January 2022].

Güthlin, D. et al. (2011) Estimating habitat suitability and potential population size for brown bears in the Eastern Alps. *Biological Conservation*. [online]. 144 (5), Elsevier, pp.1733–1741.

Hajibabaei, M. et al. (2006) A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes*. [online]. 6 (4), John Wiley & Sons, Ltd, pp.959–964. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1471-8286.2006.01470.x [Accessed 24 October 2021].

Haouas, N. et al. (2007) Development of a molecular tool for the identification of Leishmania reservoir hosts by blood meal analysis in the insect vectors. *American Journal of Tropical Medicine and Hygiene*. 77 (6), [Lawrence, Kan., etc., Allen Press, etc.], pp.1054–1059.

Harper, L.R. et al. (2018) Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (Triturus cristatus). *Ecology and Evolution*. [online]. 8 (12), John Wiley & Sons, Ltd, pp.6330–6341. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.4013 [Accessed 26 September 2021].

Harris, C.E. and Knowlton, F.F. (2011) Differential responses of coyotes to novel stimuli in familiar and unfamiliar settings. *https://doi.org/10.1139/z01-163*. [online]. 79 (11), NRC Research Press Ottawa, Canada , pp.2005–2013.

Hebert, P.D.N. et al. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*. [online]. 270 (1512), The Royal Society, p.313. Available from: /pmc/articles/PMC1691236/ [Accessed 22 March 2022].

Hendry, 15 et al. (2003) European corridors: Strategies for corridor development for target species. *Conserving Natura Rivers Ecology Series Water Science & Technology Nature and Environment Series Council of Europe. Strasbourg Veldgids Libellen. KNNV-Uitgeverij*. [online]. 29 (23), pp.273–280. Available from: https://www.ecnc.org/uploads/2012/10/2004-european-corridors.pdf [Accessed 19 January 2018].

Henle, K. et al. (2013) Priorities for biodiversity monitoring in Europe: A review of supranational policies and a novel scheme for integrative prioritization. *Ecological Indicators*. [online]. 33, Elsevier, pp.5–18.

Henschel, P. and J.R. (2015) Leopards in African Rainforests - Survey and Monitoring Techniques. Wildlife Conservation Society.

Hernandez, P.A. et al. (2008) Predicting species distributions in poorly-studied landscapes. *Biodiversity and Conservation*. [online].

Hernandez, P.A. et al. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*. [online].

Hertel, A.G. et al. (2019) Don't poke the bear: using tracking data to quantify behavioural syndromes in elusive wildlife. *Animal Behaviour*. [online]. 147, Academic Press, pp.91–104.

Hertel, A.G., Swenson, J.E. and Bischof, R. (2017) Behavioral Ecology A case for considering individual variation in diel activity patterns. *Behavioral Ecology*. [online]. 28 (6), pp.1524–1531. Available from: https://academic.oup.com/beheco/article-abstract/28/6/1524/4110358 [Accessed 20 September 2021].

Hijmans, R.J. and Graham, C.H. (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*. [online]. 12 (12), John Wiley & Sons, Ltd, pp.2272–2281. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2486.2006.01256.x [Accessed 16 January 2022].

Hill, J.T. and Demarest, B. (2014) *Tools for Sanger Sequencing Data in R* [online]. 1–16. Available from: http://home. [Accessed 22 November 2021].

Hoffmann, C. et al. (2018) Blow flies as urban wildlife sensors. *Molecular Ecology Resources*. [online]. 18 (3), John Wiley & Sons, Ltd (10.1111), pp.502–510. Available from: http://doi.wiley.com/10.1111/1755-0998.12754 [Accessed 3 April 2019].

Hofreiter, M. et al. (2001) Ancient DNA. *Nature Reviews Genetics 2001 2:5*. [online]. 2 (5), Nature Publishing Group, pp.353–359. Available from: https://www.nature.com/articles/35072071 [Accessed 30 October 2021].

Höss, M. et al. (1992) Excrement analysis by PCR. *Nature 1992 359:6392*. [online]. 359 (6392), Nature Publishing Group, pp.199–199. Available from: https://www.nature.com/articles/359199a0 [Accessed 26 September 2021].

Hunter, M.E. et al. (2019) Improving eDNA yield and inhibitor reduction through increased water volumes and multi-filter isolation techniques. *Scientific Reports 2019 9:1*. [online]. 9 (1), Nature Publishing Group, pp.1–9. Available from: https://www.nature.com/articles/s41598-019-40977-w [Accessed 19 January 2022].

Hutchison, C.A. et al. (1974) Maternal inheritance of mammalian mitochondrial DNA. *Nature 1974 251:5475*. [online]. 251 (5475), Nature Publishing Group, pp.536–538. Available from: https://www.nature.com/articles/251536a0 [Accessed 30 October 2021].

Ji, Y. et al. (2020) Large-scale Quantification of Vertebrate Biodiversity in Ailaoshan Nature Reserve from Leech iDNA PQ H3A2K6 Canada. *bioRxiv*.

Jia, S. et al. (2007) Genetic Variation of Mitochondrial D-loop Region and Evolution Analysis in Some Chinese Cattle Breeds. *Journal of Genetics and Genomics*. [online]. 34 (6), pp.510–518. Available from: www.jgenetgenomics.org [Accessed 26 March 2022].

Johns, G.C. and Avise, J.C. (1998) A Comparative Summary of Genetic Distances in the Vertebrates from the Mitochondrial Cytochrome b Gene. *Mol. Biol. Evol*. 15 (11), pp.1481–1490.

Kaczensky, P. et al. (2021) *Distribution of large carnivores in Europe 2012 - 2016: Distribution maps for Brown bear, Eurasian lynx, Grey wolf, and Wolverine Dryad, Dataset*. [online]. Available from: https://datadryad.org/stash/dataset/doi:10.5061/dryad.pc866t1p3 [Accessed 5 September 2021].

Kaczensky, P. et al. (2012a) *Status, management and distribution of large carnivores - bear, lynx, wolf & wolverine- in Europe- Part 1 European Commission*.

Kaczensky, P. et al. (2012b) Status, management and distribution of large carnivores–bear, lynx, wolf & wolverine–in Europe. Part 2. *Status of large carnivores in Europe*. (December), p.72.

Kaczensky, P. et al. (2003) The impact of high speed, high volume traffic axes on brown bears in Slovenia. *Biological Conservation*. [online]. 111 (2), Elsevier, pp.191–204.

Kadmon, R., Farber, O. and Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*. [online]. 14 (2), Ecological Society of America, pp.401–413.

Kanellopoulos, N. et al. (2006) Selective habitat use by brown (Ursus arctos L.) in northern Pindos, Greece. *Journal of Biological Research*. (5), pp.23–33.

Karamanlidis, A.A. et al. (2015) Evidence of a large carnivore population recovery: Counting bears in Greece. *Journal for Nature Conservation*. [online]. 27, pp.10–17. Available from: http://ac.els-cdn.com.ezproxy.uwe.ac.uk/S1617138115300030/1-s2.0-S1617138115300030-main.pdf?_tid=62496904-3e15-11e7-b6db-

00000aab0f01&acdnat=1495364902_7861e0450295bb0797325e8c50a96690 [Accessed 21 May 2017].

Karamanlidis, A.A. et al. (2012) Genetic diversity, structure, and size of an endangered brown bear population threatened by highway construction in the Pindos Mountains, Greece. *European Journal of Wildlife Research*. [online]. 58 (3), pp.511–522.

Karamanlidis, A.A., Kopatz, A. and de Gabriel Hernando, M. (2021) Dispersal patterns of a recovering brown bear (Ursus arctos) population in a human-dominated landscape. *Journal of Mammalogy*. [online]. 102 (2), Oxford Academic, pp.494–503. Available from: https://academic.oup.com/jmammal/article/102/2/494/6124575 [Accessed 16 January 2022].

Karanth, K.U. et al. (2011) Monitoring carnivore populations at the landscape scale: Occupancy modelling of tigers from sign surveys. *Journal of Applied Ecology*. [online]. 48 (4), John Wiley & Sons, Ltd, pp.1048–1056. Available from: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2664.2011.02002.x [Accessed 27 June 2021].

Kasai, A. et al. (2020) The effect of temperature on environmental DNA degradation of Japanese eel. [online]. 86, pp.465–471. Available from: https://doi.org/10.1007/s12562-020-01409-1 [Accessed 4 December 2021].

Keane, A. et al. (2012) The potential of occupancy modelling as a tool for monitoring wild primate populations. *Animal Conservation*. [online]. 15 (5), John Wiley & Sons, Ltd, pp.457–465. Available from: https://zslpublications.onlinelibrary.wiley.com/doi/full/10.1111/j.1469-1795.2012.00575.x [Accessed 27 June 2021].

Keller, C.M.E. and Scallan, J.T. (1999) Potential roadside biases due to habitat changes along breeding bird survey routes. *The Condor*. 101, pp.50–57.

Kent, R.J., Norris, D.E. and Feinstone, W.H. (2005) Identification of mammalian blood meals in mosquitoes by a multiplexed polymerase chain reaction targeting cytochrome B. *The American Society of Tropical Medicine and Hygiene*. 73 (2), pp.336–342.

Kiszewski, A. et al. (2004) A global index representing the stability of malaria transmission. *The American Journal of Tropical Medicine and Hygiene*. [online]. 70 (5), American Society of Tropical Medicine and Hygiene, pp.486–498. Available from: https://www.ajtmh.org/view/journals/tpmd/70/5/article-p486.xml [Accessed 5 December 2021].

Kumar, G., Eble, J.E. and Gaither, M.R. (2020) A practical guide to sample preservation and pre-PCR processing of aquatic environmental DNA. *Molecular Ecology Resources*. [online]. 20 (1), Blackwell Publishing Ltd, pp.29–39.

Kusak, J. and Huber, D. (1998) Brown Bear Habitat Quality in Gorski Kotar, Croatia. 10, pp.281–291.

Lacoursière-Roussel, A. et al. (2016) Improving herpetological surveys in eastern North America using the environmental DNA method. *https://doi.org/10.1139/gen-2015-0218*. [online]. 59 (11), NRC Research Press, pp.991–1007. Available from: https://cdnsciencepub.com/doi/abs/10.1139/gen-2015-0218%40gen-iblf.issue01 [Accessed 30 October 2021].

Langen, T. et al. (2017) Importance of Roadless Areas for the European Conservation Network. *Frontiers in Ecology and Evolution | www.frontiersin.org*. [online]. 5 (2). Available from: www.frontiersin.org [Accessed 16 January 2022].

Langlois, V.S. et al. (2021) The need for robust qPCR-based eDNA detection assays in environmental monitoring and species inventories. *Environmental DNA*. [online]. 3 (3), John Wiley & Sons, Ltd, pp.519–527. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/edn3.164 [Accessed 24 October 2021].

Larsson, M.C. and Svensson, G.P. (2009) Pheromone Monitoring of Rare and Threatened Insects: Exploiting a Pheromone–Kairomone System to Estimate Prey and Predator Abundance. *Conservation Biology*. [online]. 23 (6), John Wiley & Sons, Ltd, pp.1516–1525. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1523-1739.2009.01263.x [Accessed 18 January 2022].

Lawton, J. (2010) Making Space for Nature: A review of England's Wildlife Sites and Ecological Networks. *Defra*. pp.1–107.

Lee, P.S. (2016) A blowfly-derived dna approach to assess diversity of tropical mammals, PhD Thesis.

Lee, P.S. et al. (2016) Field calibration of blowfly-derived DNA against traditional methods for assessing mammal diversity in tropical forests. Adamowicz, S. (ed.) *Genome*. [online]. 59 (11), pp.1008–1022. Available from: http://www.nrcresearchpress.com/doi/10.1139/gen-2015-0193 [Accessed 4 February 2019].

Lee, P.S., Sing, K.-W. and Wilson, J.-J. (2015) Reading Mammal Diversity from Flies: The Persistence Period of Amplifiable Mammal mtDNA in Blowfly Guts (Chrysomya megacephala) and a New DNA Mini-Barcode Target. Mans, B.J. (ed.) *PLOS ONE*. [online]. 10 (4), p.e0123871. Available from: http://dx.plos.org/10.1371/journal.pone.0123871 [Accessed 14 May 2017].

Legg, C.J. and Nagy, L. (2006) Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management*. [online]. 78 (2), Academic Press, pp.194–199.

Lennon, J.J., Greenwood, J.J.D. and Turner, J.R.G. (2000) Bird diversity and enviromnental gradients in Britain: A test of the species-energy hypothesis. *Journal of Animal Ecology*. [online]. 69 (4), pp.581–598.

Lewis, J.S. et al. (2007) Effects of habitat on GPS collar performance: using data screening to reduce location error. *Journal of Applied Ecology*. [online]. 44 (3), John Wiley & Sons, Ltd, pp.663–671.

Lewis, J.S. et al. (2011) Identifying habitat characteristics to predict highway crossing areas for black bears within a human-modified landscape. *Landscape and Urban Planning*. [online]. 101 (2), Elsevier, pp.99–107.

Lim Tee, T. et al. (2016) Individual Identification in Sun Bears: Testing Methods to Capture Chest Marks and Hair Samples *25 (3)*. 25–26.

Linnell, J.D.C., Swenson, J.E. and Andersen, R. (2000) Conservation of biodiversity in Scandinavian boreal forests: large carnivores as flagships, umbrellas, indicators, or keystones? *Biodiversity and Conservation*. [online]. 9 (7), Kluwer Academic Publishers, pp.857–868. Available from: http://link.springer.com/10.1023/A:1008969104618 [Accessed 19 January 2018].

Long, R.A. et al. (2012) *Noninvasive Survey Methods for Carnivores* Long, R.A. et al. (eds.) [online]. Washington, Island Press. Available from: https://www.nhbs.com/noninvasive-survey-methods-for-carnivores-book.

Long, R.A. (2008) *Noninvasive survey methods for carnivores* Island Press.

Loro, M. et al. (2015) Ecological connectivity analysis to reduce the barrier effect of roads. An innovative graph-theory approach to define wildlife corridors with multiple paths and without bottlenecks. *Landscape and Urban Planning*. [online]. 139, Elsevier, pp.149–162.

Lynggaard, C. et al. (2022) Airborne environmental DNA for terrestrial vertebrate community monitoring. *Current Biology*. [online]. Cell Press. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0960982221016900 [Accessed 18 January 2022].

Lynggaard, C. et al. (2021) The potential of aquatic bloodfeeding and nonbloodfeeding leeches as a tool for iDNA characterisation. *Molecular Ecology Resources*. [online]. 22 (2), John Wiley & Sons, Ltd, pp.539–553. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/1755-0998.13486 [Accessed 28 October 2021].

Lynggaard, C. et al. (2019) Vertebrate diversity revealed by metabarcoding of bulk arthropod samples from tropical forests. *Environmental DNA*. [online]. 1 (4), John Wiley & Sons, Ltd, pp.329–341. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/edn3.34 [Accessed 28 October 2021].

Ma, J. et al. (2005) A DNA marker to identify predation of Plutella xylostella (Lep., Plutellidae) by Nabis kinbergii (Hem., Nabidae) and Lycosa sp. (Aranaea, Lycosidae). *Journal of Applied Entomology*. [online]. 129 (6), John Wiley & Sons, Ltd, pp.330–335. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1439-0418.2005.00977.x [Accessed 22 March 2022].

Mackenzie, D.I. et al. (2002) *ESTIMATING SITE OCCUPANCY RATES WHEN DETECTION PROBABILITIES ARE LESS THAN ONE Ecology*. [online]. 83 (8).

MacKenzie, D.I. et al. (2017) *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence* [online]. Elsevier. Available from: https://books.google.co.uk/books?hl=en&lr=&id=hs2cBAAAQBAJ&oi=fnd&pg=PP1&dq=Occupancy+Estimation+and+Modeling:+Inferring+Patterns+and+Dynamics+of+Species+Occurrence&ots=-YagUbfyuU&sig=rUpGqAKmN9iPqd7pQzj0A9GFmiw#v=onepage&q=Occupancy Estimation and Modeli [Accessed 22 July 2021].

MacKenzie, D.I. and Bailey, L.L. (2004) Assessing the fit of site-occupancy models. *Journal of Agricultural, Biological, and Environmental Statistics*. [online]. 9 (3), Springer, pp.300–318.

Mackenzie, D.I. and Royle, J.A. (2005) *Designing occupancy studies: General advice and allocating survey effort Journal of Applied Ecology*. [online]. 42 (6), 1105–1114.

Macleod, J. and Donnelly, J. (1958) Local Distribution and Dispersal Paths of Blowflies in Hill Country. *Source: Journal of Animal Ecology*. [online]. 27 (2), pp.349–374. Available from: https://www.jstor.org/stable/2244 [Accessed 31 July 2021].

Majka, D., Beier, P. and Jenness, J. (2007) CorridorDesigner ArcGIS Toolbox Tutorial. *Environmental Research, Development and Education for the New Economy (ERDENE)*.

Van Manen, F.T. et al. (2012) Short-term impacts of a 4-Lane highway on american black bears in Eastern North Carolina. *Wildlife Monographs*. [online]. 181 (181), John Wiley & Sons, Ltd, pp.1–35.

Marshall, N.T., Vanderploeg, H.A. and Chaganti, S.R. (2021) Environmental (e)RNA advances the reliability of eDNA by predicting its age. *Scientific Reports 2021 11:1*. [online]. 11 (1), Nature Publishing Group, pp.1–11. Available from: https://www.nature.com/articles/s41598-021-82205-4 [Accessed 11 January 2022].

Mateo Sánchez, M.C., Cushman, S.A. and Saura, S. (2014) Scale dependence in habitat selection: the case of the endangered brown bear (Ursus arctos) in the Cantabrian Range (NW Spain). *International Journal of Geographical Information Science*. [online].

Mateo-Sánchez, M.C. et al. (2015a) A comparative framework to infer landscape effects on population genetic structure: are habitat suitability models effective in explaining gene flow? *Landscape Ecology*. [online]. 30 (8), pp.1405–1420. Available from: http://link.springer.com/10.1007/s10980-015-0194-4 [Accessed 11 January 2018].

Mateo-Sánchez, M.C. et al. (2015b) Estimating effective landscape distances and movement corridors: Comparison of habitat and genetic data. *Ecosphere*. [online]. 6 (4).

Mattson, D.J. (1990) Human Impacts on Bear Habitat Use. *Bears: Their Biology and Management*. 8, pp.33–56.

Mattson, D.J., Knight, R.R. and Blanchard, B.M. (1987) The Effects of Developments and Primary Roads on Grizzly Bear Habitat Use in Yellowstone National Park, Wyoming. *Bears: Their Biology and Management*. [online]. 7, JSTOR, p.259.

Mauvisseau, Q. et al. (2019) Influence of accuracy, repeatability and detection probability in the reliability of species-specific eDNA based approaches. *Scientific Reports 2019 9:1*. [online]. 9 (1), Nature Publishing Group, pp.1–10. Available from: https://www.nature.com/articles/s41598-018-37001-y [Accessed 26 December 2021].

McGeoch, M.A. and Gaston, K.J. (2002) *Occupancy frequency distributions: Patterns, artefacts and mechanisms Biological Reviews of the Cambridge Philosophical Society*. [online]. 77 (3), Cambridge Philosophical Society, 311–331. Available from:

https://onlinelibrary.wiley.com/doi/full/10.1017/S1464793101005887 [Accessed 27 June 2021].

Melero, B. et al. (2011) Comparison between conventional and qPCR methods for enumerating Campylobacter jejuni in a poultry processing plant. *Food Microbiology*. [online]. 28 (7), Academic Press, pp.1353–1358.

Mereu, P. et al. (2019) Mitochondrial D-loop Sequence Variability in Three Native Insular Griffon Vulture (Gyps fulvus) Populations from the Mediterranean Basin. *BioMed Research International*. [online]. 2019, Hindawi Limited.

Mertzanis, G. et al. (2006) *Bear Habitat Suitability in Relation to Habitat Types of European Interest in NE Pindos Mountain Range, Greece*In: *Sustainable Management and Development of Mountainous and Island Areas*.

Mertzanis, G. (1994) *Brown Bear in Greece: Distribution, Present Status-Ecology of a Northern Pindus Subpopulation*In: *International Conference on Bear Research and Management*. [online]. 9(1): 187-197. Available from: https://www.bearbiology.com/fileadmin/tpl/Downloads/URSUS/Vol_9/Mertzanis_Vol_9.pdf [Accessed 31 January 2018].

Mertzanis, G. et al. (2008) *Brown bear (Ursus arctos L.) habitat use patterns in two regions of northern Pindos, Greece - Management implications*In: *Journal of Natural History*. [online]. 42 (5–8), Taylor & Francis , 301–315. Available from: https://www.tandfonline.com/action/journalInformation?journalCode=tnah20 [Accessed 31 January 2018].

Mertzanis, G. et al. (2011) *Telemetry as a Tool to Study Spatial Behaviour and Patterns of Brown Bears as Affected by the Newly Constructed Egnatia Highway – N. Pindos - Greece* [online]. Available from: http://cdn.intechweb.org/pdfs/21102.pdf [Accessed 31 January 2018].

Mertzanis, Y. et al. (2005) *Movements, activity patterns and home range of a female brown bear (Ursus arctos, L.) in the Rodopi Mountain Range, Greece Belg. J. Zool*. 135 (2).

Mertzanis, Y., Psaroudas, Sp. and Karamanlidis, A.A. (2020) *Εθνικό Σχέδιο Δράσης για το Ursus arctos Παραδοτέο Δράσης A.1* [online]. Available from: https://docplayer.gr/195901437-Ethniko-shedio-drasis-gia-to-ursus-arctos-paradoteo-drasis-a-1.html [Accessed 16 January 2022].

Meyer, A. (1994) Shortcomings of the cytochrome b gene as a molecular marker. *Trends in Ecology and Evolution*. [online]. 9 (8), Elsevier Current Trends, pp.278–280.

Milhau, T. et al. (2021) Seasonal dynamics of riverine fish communities using eDNA. *Journal of Fish Biology*. [online]. 98 (2), John Wiley & Sons, Ltd, pp.387–398. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/jfb.14190 [Accessed 19 January 2022].

Morgulis, A. et al. (2008) Database indexing for production MegaBLAST searches. *Bioinformatics*. [online]. 24 (16), Oxford Academic, pp.1757–1764. Available from: https://academic.oup.com/bioinformatics/article/24/16/1757/202524 [Accessed 21 November 2021].

Muscarella, R. et al. (2014) ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models . *Methods in Ecology and Evolution*. [online].

Naimi, B. et al. (2014) Where is positional uncertainty a problem for species distribution modelling? *Ecography*. [online]. 37 (2), pp.191–203.

Nelson, L.A., Wallman, J.F. and Dowton, M. (2007) Using COI barcodes to identify forensically and medically important blowflies. *Medical and Veterinary Entomology*. [online].

Nichols, J.D. and Williams, B.K. (2006) Monitoring for conservation. *Trends in Ecology & Evolution*. [online]. 21 (12), Elsevier Current Trends, pp.668–673.

Nottingham, B.G., Johnson, K.G. and Pelton, M.R. (1973) Evaluation of Scent-Station Surveys to Monitor Raccoon Density. *Bulletin*. 17 (1), pp.29–35.

O'Brien, T.G., Kinnaird, M.F. and Wibisono, H.T. (2003) Crouching tigers, hidden prey: Sumatran tiger and prey populations in a tropical forest landscape. *Animal Conservation*. [online]. 6 (2), John Wiley & Sons, Ltd, pp.131–139.

O'Connell, A.F., Nichols, J.D. and Karanth, U. (2011) *Camera Traps in Animal Ecology Camera Traps in Animal Ecology: Methods and Analyses*. [online].

Osman, S.A.M., Yonezawa, T. and Nishibori, M. (2016) Origin and genetic diversity of Egyptian native chickens based on complete sequence of mitochondrial DNA D-loop region. *Poultry Science*. [online]. 95 (6), Oxford University Press, pp.1248–1256. Available from: http://dx.doi.org/10.3382/ps/pew029 [Accessed 26 March 2022].

Paetkau, D. (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology*. [online]. 12 (6), John Wiley & Sons, Ltd, pp.1375–1387.

Penteriani, V. et al. (2018) Density and reproductive characteristics of female brown bears in the Cantabrian Mountains, NW Spain. *European Zoological Journal*. [online]. 85 (1).

Pérez-Burillo, S. et al. (2021) Effect of freezing on gut microbiota composition and functionality for in vitro fermentation experiments. *Nutrients*. [online]. 13 (7), MDPI AG. Available from: /pmc/articles/PMC8308218/ [Accessed 22 March 2022].

Pérez-Goya U et al. (2020) *GISTools*

Perkins-Taylor, I.E. and Frey, J.K. (2020) Predicting the distribution of a rare chipmunk (Neotamias quadrivittatus oscuraensis): comparing MaxEnt and occupancy models. *Journal of Mammalogy*. [online]. 101 (4), pp.1035–1048. Available from: https://academic.oup.com/jmammal/article/101/4/1035/5858053 [Accessed 31 July 2021].

Peterman, W.E., Crawford, J.A. and Kuhns, A.R. (2013) Using species distribution and occupancy modeling to guide survey efforts and assess species status. *Journal for Nature Conservation*. [online]. 21 (2), Urban & Fischer, pp.114–121.

Phillips, S.J. et al. (2009) Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications*. [online].

Phillips, S.J., Anderson, R.P. and Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*. [online].

Phillips, S.J. and Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation - Phillips - 2008 - Ecography - Wiley Online Library. *Ecography*.

Phoebus, I. et al. (2020) Comparison of grizzly bear hair-snag and scat sampling along roads to inform wildlife population monitoring. *Wildlife Biology*. [online]. 2020 (3), Nordic Council for Wildlife Research.

Piédallu, B. et al. (2017) Determinants and patterns of habitat use by the brown bear Ursus arctos in the French Pyrenees revealed by occupancy modelling *ORYX*. [online].

Piggott, M.P. (2016) Evaluating the effects of laboratory protocols on eDNA detection probability for an endangered freshwater fish. *Ecology and Evolution*. [online]. 6 (9), John Wiley & Sons, Ltd, pp.2739–2750. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.2083 [Accessed 31 October 2021].

Posillico, M. et al. (2004) A habitat model for brown bear conservation and land use planning in the central Apennines. *Biological Conservation*. [online]. 118 (2), Elsevier, pp.141–150.

Post, R.J., Flook, P.K. and Millest, A.L. (1993) Methods for the preservation of insects for DNA studies. *Biochemical Systematics and Ecology*. [online]. 21 (1), Pergamon, pp.85–92.

Psaralexi, M., Lazarina, M., Mertzanis, Y., Michaelidou, D.E. and Sgardelis, S., 2022. Exploring 15 years of brown bear (Ursus arctos)-vehicle collisions in northwestern Greece. Nature Conservation, 47, pp.105-119.

Pya, N. and Wood, S.N. (2014) Shape constrained additive models. *Statistics and Computing*. [online]. 25 (3), Kluwer Academic Publishers, pp.543–559.

Pylidis, C. et al. (2021) Multisource noninvasive genetics of brown bears (Ursus arctos) in Greece reveals a highly structured population and a new matrilineal contact zone in southern Europe. *Ecology and Evolution*. [online]. 11 (11), John Wiley & Sons, Ltd, pp.6427–6443. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.7493 [Accessed 9 July 2021].

Robert J. Hijmans et al. (2020) *Dismo*

Robinson, N.M. et al. (2018) How to ensure threatened species monitoring leads to threatened species conservation. *Ecological Management & Restoration*. [online]. 19 (3), John Wiley & Sons, Ltd, pp.222–229. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/emr.12335 [Accessed 11 January 2022].

Robinson, O.J., Ruiz-Gutierrez, V. and Fink, D. (2018) Correcting for bias in distribution modelling for rare species using citizen science data. *Diversity and Distributions*. [online]. 24 (4), Blackwell Publishing Ltd, pp.460–472.

Rödder, D. et al. (2021) Climate change drives mountain butterflies towards the summits. *Scientific Reports 2021 11:1*. [online]. 11 (1), Nature Publishing Group, pp.1–12. Available from: https://www.nature.com/articles/s41598-021-93826-0 [Accessed 16 January 2022].

Rodgers, T.W. et al. (2017) Carrion fly-derived DNA metabarcoding is an effective tool for mammal surveys: Evidence from a known tropical mammal community. *Molecular Ecology Resources*. [online]. 17 (6).

Roever, C.L., Boyce, M.S. and Stenhouse, G.B. (2008) Grizzly bears and forestry: II: Grizzly bear habitat selection and conflicts with road placement. *Forest Ecology and Management*. [online]. 256 (6), Elsevier, pp.1262–1269.

Rondinini, C., Rodrigues, A.S.L. and Boitani, L. (2011) The key elements of a comprehensive global mammal conservation strategy. *Philosophical Transactions of the Royal Society B: Biological Sciences*. [online]. 366 (1578), The Royal Society , pp.2591–2597. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2011.0111 [Accessed 26 September 2021].

Ruppert, K.M., Kline, R.J. and Rahman, M.S. (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*. [online]. 17, Elsevier, p.e00547.

Salamon, J. et al. (2016) Towards the Automatic Classification of Avian Flight Calls for Bioacoustic Monitoring. *PLOS ONE*. [online]. 11 (11), Public Library of Science, p.e0166866. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0166866 [Accessed 18 January 2022].

Savvantoglou, A. et al. (2017) *A GIS approach to identifying connectivity potential between brown bear (Ursus arctos) habitat in northern Greece*In: *18th Hellenic Forestry Congress & International Workshop*. [online]. 1422–1432. Available from: https://www.forestry.gr/assets/pdf/conf/FULL CONGRESS PROCEEDINGS HFS 2017 WORKSHOP 2(2) with ISBN.pdf.

Savvantoglou, A. (2015) *A pilot study: The distribution of the brown bear (Ursus arctos) metapopulation of Mt. Voras, N. Greece. An estimation of the human-bear conflicts and risks. Placement report, BSc.*

Schaffer-Smith, D., Swenson, J.J. and Boveda-Penalba, A. (2016) Rapid conservation assessment for endangered species using habitat connectivity models. *Environmental Conservation C Foundation for Environmental Conservation*. [online]. 43 (3), pp.221–230. Available from: http://dx.doi. [Accessed 22 April 2017].

Schmittgen, T.D. et al. (2000) Quantitative Reverse Transcription–Polymerase Chain Reaction to Study mRNA Decay: Comparison of Endpoint and Real-Time Methods. *Analytical Biochemistry*. [online]. 285 (2), Academic Press, pp.194–204.

Schneider, J. et al. (2016) Detection of Invasive Mosquito Vectors Using Environmental DNA (eDNA) from Water Samples. *PLOS ONE*. [online]. 11 (9), Public Library of Science, p.e0162493. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0162493 [Accessed 30 October 2021].

Schnell, I.B. et al. (2018) Debugging diversity – a pan-continental exploration of the potential of terrestrial blood-feeding leeches as a vertebrate monitoring tool. *Molecular Ecology Resources*. [online]. 18 (6), John Wiley & Sons, Ltd, pp.1282–1298. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/1755-0998.12912 [Accessed 28 October 2021].

Schnell, I.B. et al. (2015) iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool – prospects, pitfalls and avenues to be developed. *Frontiers in Zoology*. [online]. 12 (1), BioMed Central, p.24. Available from: http://www.frontiersinzoology.com/content/12/1/24 [Accessed 31 January 2018].

Schubert, G. et al. (2015) Targeted detection of mammalian species using carrion fly-derived DNA. *Molecular Ecology Resources*. [online]. 15 (2), John Wiley & Sons, Ltd, pp.285–294. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/1755-0998.12306 [Accessed 26 March 2022].

Scott, T.W. et al. (2006) DNA profiling of human blood in anophelines from lowland and highland sites in western Kenya. *The American journal of tropical medicine and hygiene*. 75 (2), American Society of Tropical Medicine and Hygiene, pp.231–237.

Segurado, P. and Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*. [online]. 31 (10).

Servheen, C., Waller, J.S. and Sandstrom, P. (2001) *Identification and management of linkage zones for grizzly bears between the large blocks of public land in the Northern Rocky Mountains*In: Irwin CL, G.P. and McDermott KP (eds.) *Proceedings of the 2001 International Conference on Ecology and Transportation,*. [online]. UC Davis: Road Ecology Center, 161–179. Available from: https://escholarship.org/uc/item/9kr1w8fp [Accessed 11 July 2021].

Shadel, G.S. (2008) Expression and Maintenance of Mitochondrial DNA : New Insights into Human Disease Pathology. *The American Journal of Pathology*. [online]. 172 (6), American Society for Investigative Pathology, p.1445. Available from: /pmc/articles/PMC2408405/ [Accessed 22 March 2022].

Shewring, M.P. and Vafidis, J.O. (2021) Using UAV-mounted thermal cameras to detect the presence of nesting nightjar in upland clear-fell: A case study in South Wales, UK. *Ecological Solutions and Evidence*. [online]. 2 (1), John Wiley & Sons, Ltd, p.e12052. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/2688-8319.12052 [Accessed 18 January 2022].

da Silva Neto, J.G. et al. (2020) Integrating species distribution and occupancy modeling to study hellbender (Cryptobranchus alleganiensis) occurrence based on eDNA surveys. *Biological Conservation*. [online]. 251.

Sokal, R.R. and Rohlf, F.J. (2009) *Introduction to Biostatistics: Second Edition New York*.

Solberg, K.H. et al. (2006) An evaluation of field and non-invasive genetic methods to estimate brown bear (Ursus arctos) population size. *Biological Conservation*. [online]. 128 (2), Elsevier, pp.158–168.

Stem, C. et al. (2005) Monitoring and Evaluation in Conservation: a Review of Trends and Approaches. *Conservation Biology*. [online]. 19 (2), John Wiley & Sons, Ltd, pp.295–309. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1523-1739.2005.00594.x [Accessed 11 January 2022].

Steuber, S., Abdel-Rady, A. and Clausen, P.-H. (2005) PCR-RFLP analysis: a promising technique for host species identification of blood meals from tsetse flies (Diptera: Glossinidae). *Parasitology Research*. [online]. 97 (3), pp.247–254. Available from: http://link.springer.com/10.1007/s00436-005-1410-y [Accessed 31 January 2018].

Stoeckle, B.C. et al. (2017) A systematic approach to evaluate the influence of environmental conditions on eDNA detection success in aquatic ecosystems. *PLOS ONE*. [online]. 12 (12), Public Library of Science, p.e0189119. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0189119 [Accessed 19 January 2022].

Strickland, G.J. and Roberts, J.H. (2019) Utility of eDNA and occupancy models for monitoring an endangered fish across diverse riverine habitats. *Hydrobiologia*. [online]. 826 (1), Springer International Publishing, pp.129–144. Available from: https://link.springer.com/article/10.1007/s10750-018-3723-8 [Accessed 19 June 2021].

Strimas-Mackey, M. et al. (2020) Best Practices for Using eBird Data v1.0. [online]. Available from: https://zenodo.org/record/3620739 [Accessed 28 June 2021].

Su, H., Bista, M. and Li, M. (2021) Mapping habitat suitability for Asiatic black bear and red panda in Makalu Barun National Park of Nepal from Maxent and GARP models.

Scientific Reports 2021 11:1. [online]. 11 (1), Nature Publishing Group, pp.1–14. Available from: https://www.nature.com/articles/s41598-021-93540-x [Accessed 16 January 2022].

Svensson, G.P. et al. (2011) Chemical ecology and insect conservation: optimising pheromone-based monitoring of the threatened saproxylic click beetle Elater ferrugineus. *Journal of Insect Conservation 2011 16:4*. [online]. 16 (4), Springer, pp.549–555. Available from: https://link.springer.com/article/10.1007/s10841-011-9440-5 [Accessed 18 January 2022].

Swenson, J.E. et al. (2007) Brown bear body mass and growth in northern and southern Europe. *Oecologia 2007 153:1*. [online]. 153 (1), Springer, pp.37–47.

Swenson, J.E., Sandegren, F. and Soderberg, A. (1998) Geographic expansion of an increasing brown bear population: evidence for presaturation dispersal. *Journal of Animal Ecology*. [online]. 67 (5), John Wiley & Sons, Ltd, pp.819–826. Available from: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1046/j.1365-2656.1998.00248.x [Accessed 11 July 2021].

Taberlet, P. et al. (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear populationse. *Molecular Ecology*. [online]. 6 (9), John Wiley & Sons, Ltd, pp.869–876.

Taberlet, P. and Bouvet, J. (1992) Bear conservation genetics. *Nature 1992 358:6383*. [online]. 358 (6383), Nature Publishing Group, pp.197–197. Available from: https://www.nature.com/articles/358197a0 [Accessed 26 September 2021].

Tanentzap, A.J., Walker, S. and Theo Stephens, R.T. (2017) Better practices for reporting on conservation. *Conservation Letters*. [online]. 10 (1), Wiley-Blackwell, pp.146–152.

Tee, T.L. et al. (2020) An evaluation of noninvasive sampling techniques for Malayan sun bears. *https://doi.org/10.2192/URSUS-S-20-00004.1*. [online]. 2020 (31e16), International Association for Bear Research and Management, pp.1–12. Available from: https://bioone.org/journals/ursus/volume-2020/issue-31e16/URSUS-S-20-00004.1/An-evaluation-of-noninvasive-sampling-techniques-for-Malayan-sun-bears/10.2192/URSUS-S-20-00004.1.full [Accessed 4 January 2022].

Thompson, W.L. (2004) *Sampling rare or elusive species: concepts, designs, and techniques for estimating population parameters Population English Edition*. [online]. 34 (3). Available from: http://www.loc.gov/catdir/toc/ecip0417/2004008006.html [Accessed 28 June 2021].

Thomsen, P.F. and Willerslev, E. (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*. [online]. 183, Elsevier, pp.4–18. Available from: https://www.sciencedirect.com/science/article/pii/S0006320714004443 [Accessed 13 March 2019].

Thuo, D. et al. (2019) Food from faeces: Evaluating the efficacy of scat DNA metabarcoding in dietary analyses. *PLOS ONE*. [online]. 14 (12), Public Library of Science, p.e0225805. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0225805 [Accessed 30 October 2021].

Townzen, J.S., Brower, A.V.Z. and Judd, D.D. (2008) Identification of mosquito bloodmeals using mitochondrial cytochrome oxidase subunit I and cytochrome b gene sequences. *Medical and Veterinary Entomology*. [online]. 22 (4), pp.386–393. Available from: http://doi.wiley.com/10.1111/j.1365-2915.2008.00760.x [Accessed 23 January 2018].

Travaini, A. et al. (2013) Evaluation of neophobia and its potential impact upon predator control techniques: A study on two sympatric foxes in southern Patagonia. *Behavioural Processes*. [online]. 92, Elsevier, pp.79–87.

Tsaparis, D. et al. (2014) Non-invasive genetic study and population monitoring of the brown bear (Ursus arctos) (Mammalia: Ursidae) in Kastoria region – Greece. *http://dx.doi.org/10.1080/00222933.2013.877992*. [online]. 49 (5–8), Taylor & Francis, pp.393–410. Available from: https://www.tandfonline.com/doi/abs/10.1080/00222933.2013.877992 [Accessed 9 July 2021].

Tsuda, Y. et al. (2009) *Dispersal of a Blow Fly, Calliphora nigribarbis, in Relation to the Dissemination of Highly Pathogenic Avian Influenza Virus Jpn. J. Infect. Dis*. 62.

Vafidis, J. et al. (2021) Mapping scrub vegetation cover from photogrammetric point clouds: A useful tool in reserve management. *Ecology and Evolution*. [online]. 11 (11), John Wiley & Sons, Ltd, pp.6789–6797. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.7527 [Accessed 18 January 2022].

Veloz, S.D. (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*. [online].

Vynne, C. et al. (2012) Factors influencing degradation of DNA and hormones in maned wolf scat. *Animal Conservation*. [online]. 15 (2), John Wiley & Sons, Ltd, pp.184–194. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1469-1795.2011.00503.x [Accessed 30 October 2021].

Walker, R.E. and Craighead, F.L. (1997) Analyzing Wildlife Movement Corridors in Montana Using GIS. [online]. Available from: https://www.researchgate.net/publication/297734026 [Accessed 17 January 2022].

Wall, R. (1998) Estimates of population density and dispersal in the blowfly Lucilia sericata (Diptera: Calliphoridae) Sustainable livestock development View project Ticks and Tick Borne Diseases View project. *Article in Bulletin of Entomological Research*. [online].

Wang, F. et al. (2018) Does one size fit all? A multispecies approach to regional landscape corridor planning. *Diversity and Distributions*. [online]. 24 (3).

Wang, S. et al. (2021) Methodology of fish eDNA and its applications in ecology and environment. *Science of The Total Environment*. [online]. 755, Elsevier, p.142622.

Warren, D.L. and Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*. [online]. 21 (2), John Wiley & Sons, Ltd, pp.335–342.

Wegge, P., Pokheral, C.Pd. and Jnawali, S.R. (2004) Effects of trapping effort and trap shyness on estimates of tiger abundance from camera trap studies. *Animal Conservation forum*. [online]. 7 (3), Cambridge University Press, pp.251–256. Available from: https://www.cambridge.org/core/journals/animal-conservation-forum/article/abs/effects-of-trapping-effort-and-trap-shyness-on-estimates-of-tiger-abundance-from-camera-trap-studies/3584B5AED9CF8467CF866B7DECE4957B [Accessed 17 October 2021].

Weldon, L. et al. (2020) A comparison of European eel Anguilla anguilla eDNA concentrations to fyke net catches in five Irish lakes . *Environmental DNA*. [online].

Whiteman, A. et al. (2017) Identifying key denning habitat to conserve brown bear (Ursus arctos) in Croatia. *Wildlife Research*. [online]. 44 (4), p.309. Available from: http://www.publish.csiro.au/?paper=WR16164 [Accessed 1 April 2019].

Wich, S.A. and Koh, L.P. (2018) Conservation drones: Mapping and monitoring biodiversity. *Conservation Drones: Mapping and Monitoring Biodiversity*. [online]. Oxford University Press, pp.1–118. Available from:

https://books.google.com/books/about/Conservation_Drones.html?id=5P5cDwAAQBAJ [Accessed 18 January 2022].

Wiegand, T., Revilla, E. and Moloney, K.A. (2005) Effects of Habitat Loss and Fragmentation on Population Dynamics. *Conservation Biology*. [online]. 19 (1), John Wiley & Sons, Ltd, pp.108–121.

Wiesner, R.J., Rüegg, J.C. and Morano, I. (1992) Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues. *Biochemical and biophysical research communications*. [online]. 183 (2), Biochem Biophys Res Commun, pp.553–559. Available from: https://pubmed.ncbi.nlm.nih.gov/1550563/ [Accessed 22 March 2022].

Wilson, K.A. et al. (2011) Prioritizing conservation investments for mammal species globally. *Philosophical Transactions of the Royal Society B: Biological Sciences*. [online]. 366 (1578), The Royal Society , pp.2670–2680.

Wong, M.K.-S., Nakao, M. and Hyodo, S. (2020) Field application of an improved protocol for environmental DNA extraction, purification, and measurement using Sterivex filter. *Scientific Reports 2020 10:1*. [online]. 10 (1), Nature Publishing Group, pp.1–13. Available from: https://www.nature.com/articles/s41598-020-77304-7 [Accessed 31 October 2021].

Wright, M.N. and Ziegler, A. (2015) ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*. [online]. 77 (1), American Statistical Association.

Wu, J. et al. (2015) Phylogeographic and Demographic Analysis of the Asian Black Bear (Ursus thibetanus) Based on Mitochondrial DNA. *PLOS ONE*. [online]. 10 (9), Public Library of Science, p.e0136398. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0136398 [Accessed 26 March 2022].

Yu, L. et al. (2007) Analysis of complete mitochondrial genome sequences increases phylogenetic resolution of bears (Ursidae), a mammalian family that experienced rapid speciation. [online]. Available from: http://www.biomedcentral.com/1471-2148/7/198 [Accessed 10 May 2019].

Zhang, Z. et al. (2004) A Greedy Algorithm for Aligning DNA Sequences. *https://home.liebertpub.com/cmb*. [online]. 7 (1–2), Mary Ann Liebert, Inc. , pp.203–214. Available from:

https://www.liebertpub.com/doi/abs/10.1089/10665270050081478 [Accessed 21 November 2021].

Appendix I: Supplementary material

A. Habitat Suitability Modelling

A.1 Three most important variables for each model:

The shape descriptions for the response curves followed the categories used in (van Gils et al., 2014).

Table 13. Brown bear SDMs in the two raster cell size groups and four categories of environmental variables and percentage contribution of the three largest contributors.

Model	Predictor 1	Permutation importance	Curve shape [c]	Predictor 2	Permutation importance	Curve shape [c]	Predictor 3	Permutation importance	Curve shape [c]
SP_tp	Elevation	93.65	Bell	Topography	6.35	na	-	-	-
SP_lc	Elevation	85.05	Bell	Land cover	12.06	na	Topography	2.90	na
SP_hi	Elevation	66.47	Bell	Dist. from roads	15.18	Curvln -	Distance from urban areas	13.98	r-skew

[a] Bio 14: Precipitation of driest month

[b] Bio 3: Isothermality

[c] Types of curves: Bell, bell-shaped; r-skew, right-skewed; l-skew, left-skewed; curvln-, curvilinear negative; curvl+, curvilinear positive; na, not applicable to categorical predictors, as described by van Gils et al. (2014).

A.2. MaxEnt model outputs for the Species Presence models

Topography (Variables: elevation and topographic position)

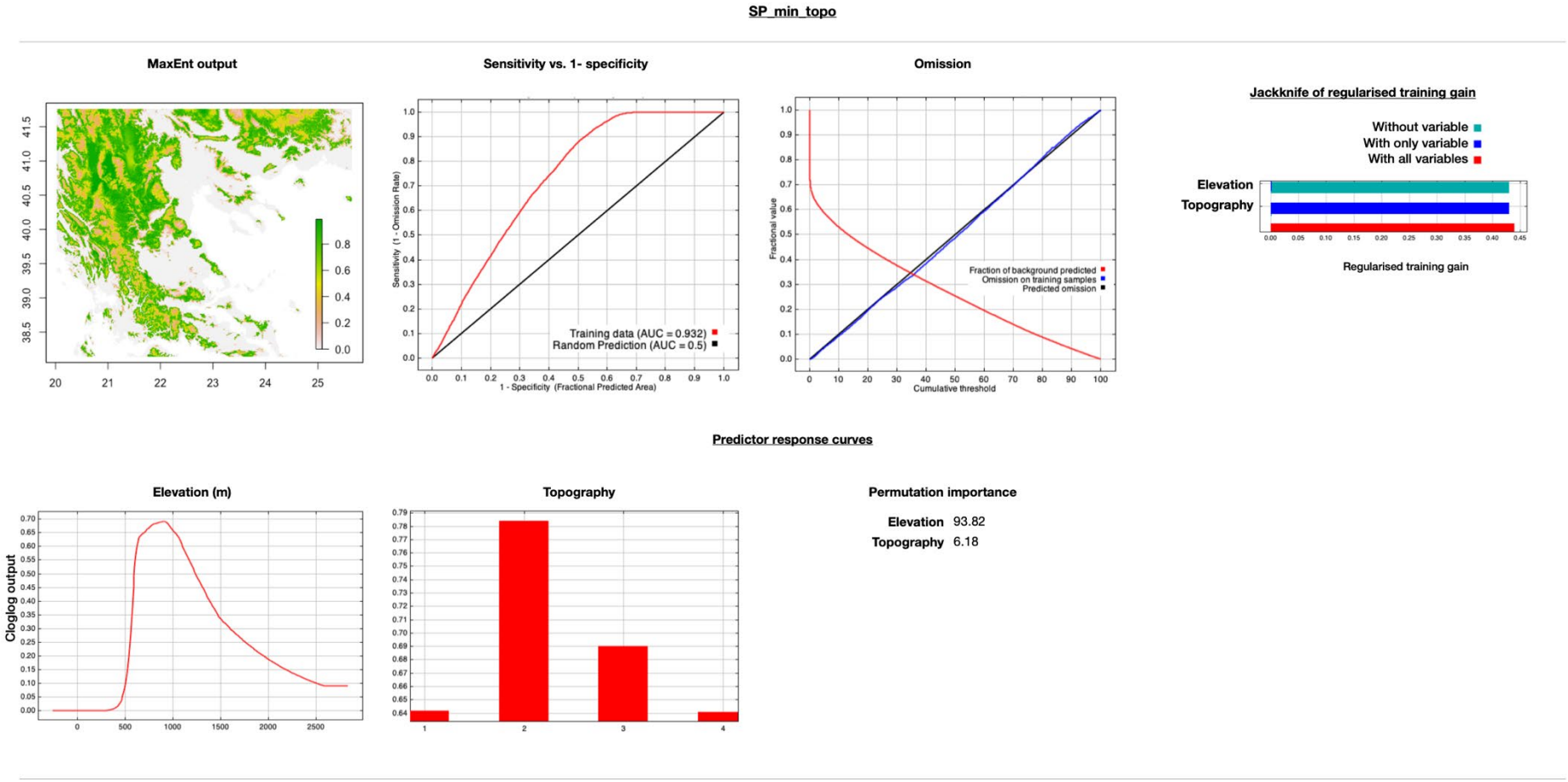


Figure 42. Species Presence MaxEnt model outputs: Topography model (Variables: elevation and topographic position)

Land Cover (Variables: elevation, topographic position and land cover type)

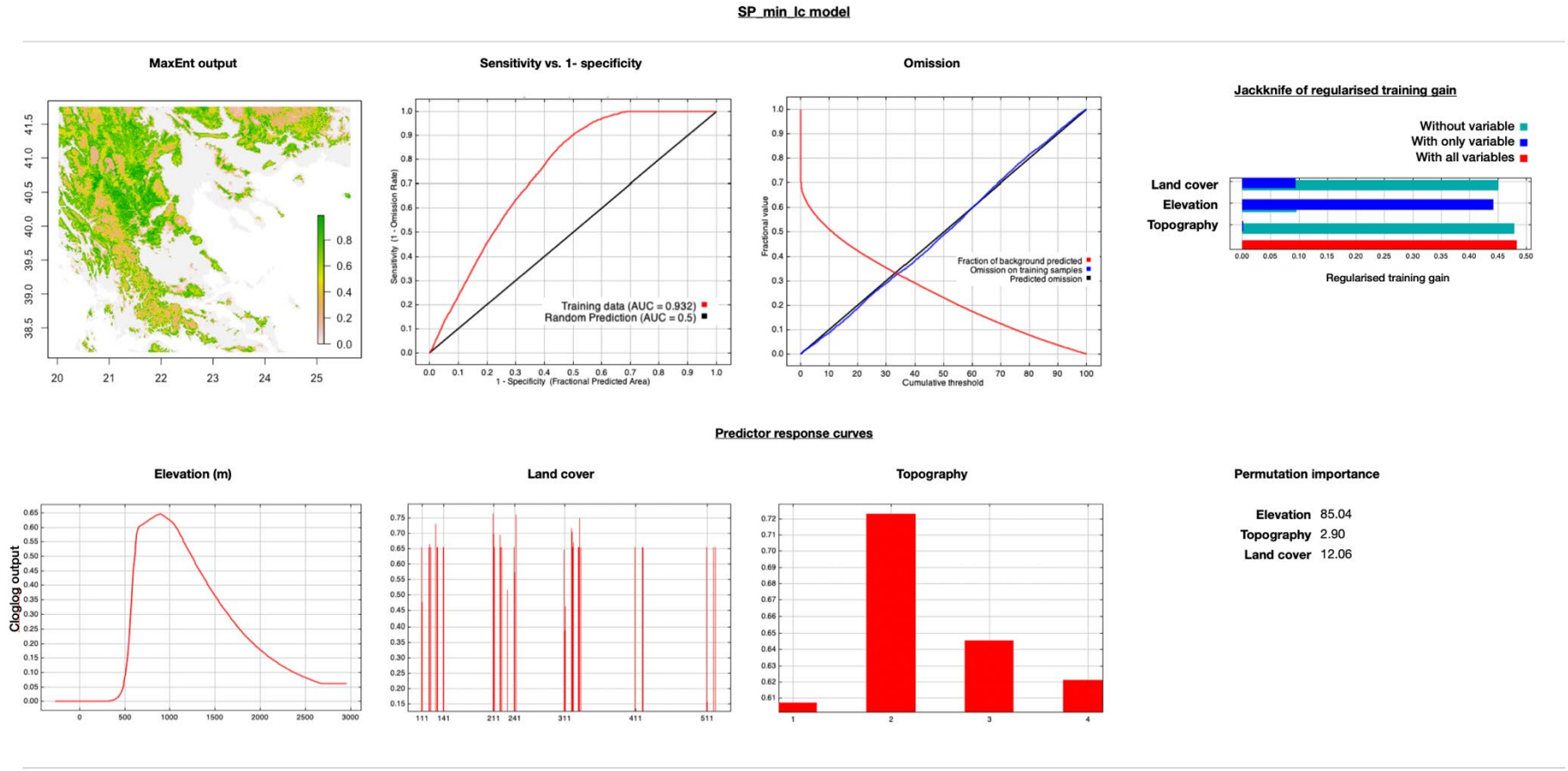


Figure 43. Species Presence MaxEnt model outputs: Land Cover model (Variables: elevation, topographic position and land cover type)

Human Impact (Variables: elevation, topographic position, land cover type, distance from major roads and distance from urban areas)

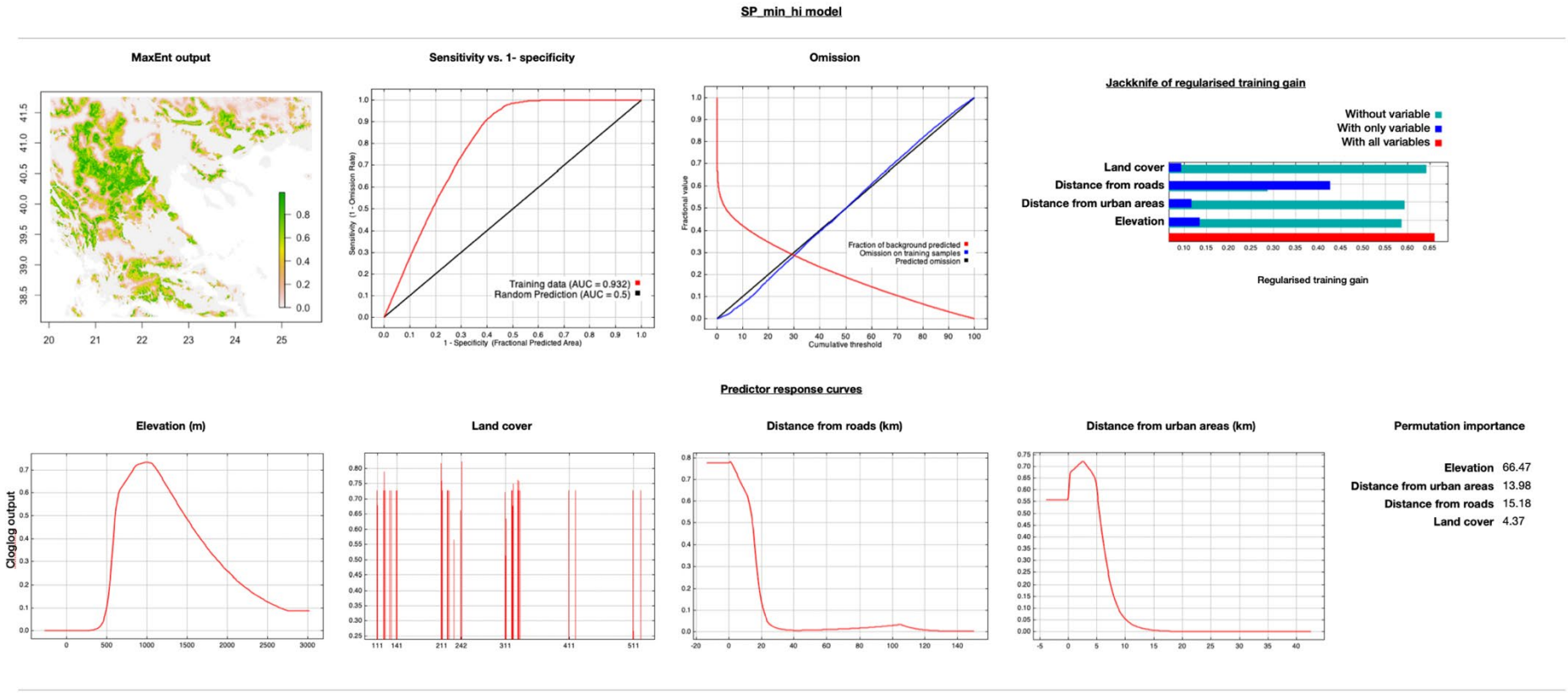


Figure 44. Species Presence MaxEnt model outputs: Human Impact model (Variables: elevation, topographic position, land cover type, distance from major roads and distance from urban areas).

A.3. Binary model comparisons

Results of binary model comparison showing the differences between model pair coverage and model overlap (Table 14).

Table 14. Comparison of suitable patches predicted by each binary model, along with the areas of overlap and areas uniquely predicted by only one of the modes.

A. Comparison of expert knowledge models

Model	Topography **EK_tp versus EK_tp(Pl)**		Land cover **EK_lc versus EK_lc(Pl)**		Human impact **EK_hi versus EK_hi(Pl)**	
Percent cover overlap	99.98		85.93		44.77	
Percent cover of IUCN dataset						
	EK_tp	**EK_tp(Pl)**	**EK_lc**	**EK_lc(Pl)**	**EK_hi**	**EK_hi(Pl)**
Percentage of study area covered by each model	38.11	38.10	38.23	36.19	43.29	20.53
Percentage of study area covered only by this model	0.01	0.00	3.34	5.38	23.91	1.16

B. Comparison of Expert Knowledge models with Species Presence models

Model	Topography **EK_tp VS SPmin_tp**		Land cover **EK_lc VS SPmin_lc**		Human impact **EK_hi VS SPmin_hi**	
Percent cover overlap	58.93		45.11		22.34	
Percent cover of IUCN dataset						
	EK_tp	**SPmin_tp**	**EK_lc**	**SPmin_tp**	**EK_hi**	**SPmin_tp**
Percentage of study area covered by each model	38.11	22.45	38.10	19.08	43.29	10.70
Percentage of study area covered only by this model	15.66	0.00	20.98	1.84	33.62	1.03

Results of binary model comparison with the IUCN dataset showing the differences in each model's coverage and model overlap (Table 15).

Table 15. Comparison of model predictions (suitability ≥50) with IUCN dataset. All percentage (%) values are in relation to the percentage cover of the total IUCN dataset extent in the study area.

Expert Knowledge models (HSI_EK)	**Topograhy**	**Land cover**	**Human impact**
Percentage IUCN dataset covered by the model prediction	86.87	85.03	87.28
Percentage of IUCN dataset not covered by the model prediction	13.13	14.97	12.72
Model prediction outside the extent of the IUCN dataset	58.12	60.41	77.41
Permutation Importance models (HSI_PI)	**Topograhy**	**Land cover**	**Human impact**
Percentage IUCN dataset covered by the model prediction	86.85	84.12	82.11
Percentage of IUCN dataset not covered by the model prediction	13.15	15.88	17.89
Model prediction outside the extent of the IUCN dataset	58.11	53.56	39.79
Species Presence models – fine resolution (SP_min)	**Topograhy**	**Land cover**	**Human impact**
Percentage IUCN dataset covered by the model prediction	54.79	46.77	23.61
Percentage of IUCN dataset not covered by the model prediction	45.21	53.23	76.39
Model prediction outside the extent of the IUCN dataset	30.63	25.83	17.09

B. Single-species iDNA monitoring

B.1. Synthetic DNA sequence

The following sequence was manufactured by the IDT ((gBlocks® Gene Fragments, Integrated DNA Technologies, Inc). Sequencefound in the cytochrome b region, mtDNA:

5'-TTCCTAGCCATACACTATACACCAGACACAACCGCAGCTTTTTCATCGGTCACCCACATTTG
CCGAGACGTTCACTACGGGTGAGTTATCCGATATGTACATGCAAATGGAGCCTCCATCTTCTTT
ATCTGCCTATTTATGCACGTAGGACGGGGCCTGTACTATGGCTCATACCTATTCCCAGAAACAT
GAAACATTGGCATTATTCTCCTATTTACAATTATAGCCACCGCATTTATAGGATACGTCCTACCC
TGGGGCCAAATGTCCTTCTGAGGAGCGACTGTCATCACCAATCTACTATCGGCCATTCCCTACA
TCGGAACGGACCTGGTAGAATGAATCTGAGGGGGCTTTTCCGTAGATAAGGCGACCCTAACA
CGATTCTTTGCTTTCCACTTTATTCTCCCGTTCATCATCCTAGCACTAGCAGCAGTCCATCTATTG
TTCCTACACGAAACAGGATCTAACAACCCCTCTGGAATCCCATCTGACTCAGA-3'

B.2. Sanger results BLAST

The Sanger sequencing primary sequence was checked against the NCBI database using BLAST (Table 16). The results are found below. The amplicon matched with two extinct bear species, the Deninger's bear (*Ursus deningeri*) and the Kudaro bear (*Ursus kudarensis*) with 100% Percent Identity, and two dik-dik species (*Madoqua kirkii* and *Madoqua guentheri*) with 92-94% Percent Identity (disregarded due to the fact that they are not present in the area and not matched fully to the amplicon).

Table 16. NCBI BLAST results of the primary sequence from the Snager sequencing analysis.

Description	*Scientific Name*	*Per. ident*	*Acc. Len*	*Accession*
Ursus deningeri kudarensis mitochondrion, partial genome	Ursus kudarensis	100.00	16815	MH605139.1
Ursus kudarensis praekudarensis mitochondrion, partial genome	Ursus kudarensis praekudarensis	100.00	16816	MW491935.1
Ursus kudarensis kudarensis mitochondrion, complete genome	Ursus kudarensis kudarensis	100.00	16814	MW491934.1
Ursus arctos mitochondrion, partial genome	Ursus arctos	97.50	17022	MH255807.1
Ursus maritimus voucher F-2374 mitochondrion, partial genome	Ursus maritimus	97.50	17020	OK001279.1
Ursus arctos voucher F-2296 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001278.1
Ursus arctos voucher IK-1 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001277.1
Ursus arctos voucher F(R)-217 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001274.1
Ursus arctos voucher F(R)-276 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001269.1
Ursus arctos voucher F(R)-247 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001268.1
Ursus arctos voucher F(R)-219 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001267.1
Ursus arctos voucher F(R)-19 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001266.1
Ursus arctos voucher F(R)-18 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001265.1
Ursus arctos voucher F(R)-7/2 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001264.1
Ursus arctos voucher F(R)-248 mitochondrion, partial genome	Ursus arctos	97.50	17020	OK001262.1
Ursus arctos from China mitochondrion, partial genome	Ursus arctos	97.50	16446	OK512981.1
Ursus arctos voucher VNHM 40626 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512979.1
Ursus arctos voucher VNHM 40601 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512978.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16447	OK512977.1
Ursus arctos voucher VNHM 40648 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512976.1
Ursus arctos voucher VNHM 40604 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512975.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16447	OK512974.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16447	OK512973.1
Ursus arctos haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512972.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16447	OK512971.1

Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16447	OK512966.1
Ursus arctos voucher IPAE 705/514 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512965.1
Ursus arctos voucher TP5-01-1 haplogroup 4 mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512964.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16418	OK512955.1
Ursus arctos from Canada mitochondrion, partial genome	Ursus arctos	97.50	16446	OK512954.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16442	OK512953.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16306	OK512952.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16399	OK512950.1
Ursus arctos voucher AMNH 30422 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512949.1
Ursus arctos voucher FAM 95630 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512948.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16447	OK512947.1
Ursus arctos voucher FAM 95642 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512946.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16441	OK512944.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16434	OK512943.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16427	OK512942.1
Ursus arctos voucher FAM 95612 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512941.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16443	OK512940.1
Ursus arctos voucher FAM 95596 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512939.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16446	OK512938.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16424	OK512936.1
Ursus arctos voucher FAM 95671 haplogroup 3c mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512935.1
Ursus arctos voucher FAM 95603 haplogroup 3c mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512934.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16437	OK512933.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16390	OK512932.1
Ursus arctos voucher FAM 30770-F haplogroup 3c mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512931.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16415	OK512928.1
Ursus arctos voucher FAM 95640 haplogroup 3c mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512927.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16440	OK512926.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16443	OK512925.1
Ursus arctos from Canada mitochondrion, partial genome	Ursus arctos	97.50	16443	OK512924.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16440	OK512923.1
Ursus arctos voucher KU 23034 haplogroup 4 mitochondrion, complete genome	Ursus arctos	97.50	16445	OK512920.1
Ursus arctos voucher CMN 28972 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512919.1

Ursus arctos voucher FAM 95665 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512918.1
Ursus arctos voucher CMN 42381 haplogroup 4 mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512917.1
Ursus arctos from USA mitochondrion, partial genome	Ursus arctos	97.50	16446	OK512915.1
Ursus arctos from Canada mitochondrion, partial genome	Ursus arctos	97.50	16424	OK512914.1
Ursus arctos from Canada mitochondrion, partial genome	Ursus arctos	97.50	16443	OK512912.1
Ursus arctos voucher MMZ S34928 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512910.1
Ursus arctos voucher MMZ S1396 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512909.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16442	OK512908.1
Ursus arctos voucher MMZ S159009 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512907.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16439	OK512906.1
Ursus arctos voucher MMZ S34972 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512905.1
Ursus arctos voucher MMZ S34958 haplogroup 3b mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512904.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16444	OK512903.1
Ursus arctos voucher MMZ S84888 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512902.1
Ursus arctos voucher MMZ S34934 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512901.1
Ursus arctos voucher MMZ S84887 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512900.1
Ursus arctos voucher MMZ S66341 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512899.1
Ursus arctos voucher MMZ S59248 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512898.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16443	OK512897.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16424	OK512896.1
Ursus arctos voucher MMZ S6039 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512895.1
Ursus arctos from Mongolia mitochondrion, partial genome	Ursus arctos	97.50	16429	OK512894.1
Ursus arctos voucher MMZ S34945 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512893.1
Ursus arctos voucher MMZ S1395 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16447	OK512892.1
Ursus arctos voucher MMZ S2073 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512891.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16442	OK512890.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16442	OK512889.1
Ursus arctos voucher MMZ S22359 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512888.1
Ursus arctos voucher MMZ S14882 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512887.1
Ursus arctos from Russia mitochondrion, partial genome	Ursus arctos	97.50	16443	OK512886.1

Ursus arctos voucher MMZ S66362 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512885.1
Ursus arctos voucher MMZ S113733 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512884.1
Ursus arctos voucher MMZ S-66359 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512883.1
Ursus arctos voucher IPAE 621/9 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512882.1
Ursus arctos voucher IPAE 107/06 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512881.1
Ursus arctos voucher IPAE 871/1 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512880.1
Ursus arctos voucher IPAE 719/449 haplogroup 3a mitochondrion, complete genome	Ursus arctos	97.50	16448	OK512879.1
Ursus arctos JBB-32K mitochondrial DNA, nearly complete genome	Ursus arctos	97.50	17012	LC595633.1
Madoqua kirkii cytochrome b (cytb) gene, complete cds; mitochondrial	Madoqua kirkii	94.00	1140	JF489137.1
Madoqua guentheri isolate BM32A cytochrome b (cytb) gene, partial cds; mitochondrial	Madoqua guentheri	94.00	300	HM209242.1
Madoqua guentheri cytochrome b (cyt b) gene, mitochondrial gene encoding mitochondrial protein, partial cds	Madoqua guentheri	94.00	300	AF030598.1
Madoqua kirkii voucher FL0193 cytochrome b (cytb) gene, partial cds; mitochondrial	Madoqua kirkii	92.00	306	MN124227.1

C. Occupancy

C1. Correlation of top predictors

Table of predictor correlation analysis.

Table 17. Correlation matrix for top predictors. Positive correlations shown in green and negative correlations show in red. Survey covariates (transect length and survey year) were excluded from the analysis.

	Non-irrigated arable land	Broadleaved forest	Natural grassland	Transitional woodland	Canyons	Flat/gentle slopes	Steep slopes	Ridgetops	Mean elevation	Mean elevation SD	Mean distance from major roads	Mean distance from major roads SD	Mean distance form urban areas	Mean distance form urban areas SD
Non-irrigated arable land	1	-0.41	-0.22	-0.28	-0.52	0.67	-0.65	-0.54	-0.47	-0.58	-0.38	-0.07	-0.33	0.36
Broadleaved forest		1	-0.17	0.09	0.63	-0.63	0.57	0.59	0.53	0.62	0.16	-0.04	0.28	-0.3
Natural grassland			1	-0.08	0.1	-0.28	0.3	0.21	0.28	0.18	0.15	-0.03	0.11	-0.1
Transitional woodland				1	0.31	-0.38	0.38	0.26	0.18	0.28	0.28	0.03	0.2	-0.23
Canyons					1	-0.79	0.65	0.84	0.64	0.79	0.43	-0.08	0.29	-0.34
Flat/gentle slopes						1	-0.97	-0.83	-0.77	-0.89	-0.49	0.06	-0.36	0.41
Steep slopes							1	0.69	0.73	0.83	0.47	-0.04	0.34	-0.4
Ridgetops								1	0.71	0.8	0.37	-0.08	0.29	-0.33
Mean elevation									1	0.75	0.52	-0.04	0.33	-0.41
Mean elevation SD										1	0.44	-0.04	0.29	-0.35
Mean dist, major roads											1	0.16	0.3	-0.37
Mean dist, major roads SD												1	0.09	-0.06
Mean distance form urban areas													1	-0.87
Mean distance form urban areas SD														1

C2. Optimised occupancy models

Figure of optimised occupancy models.

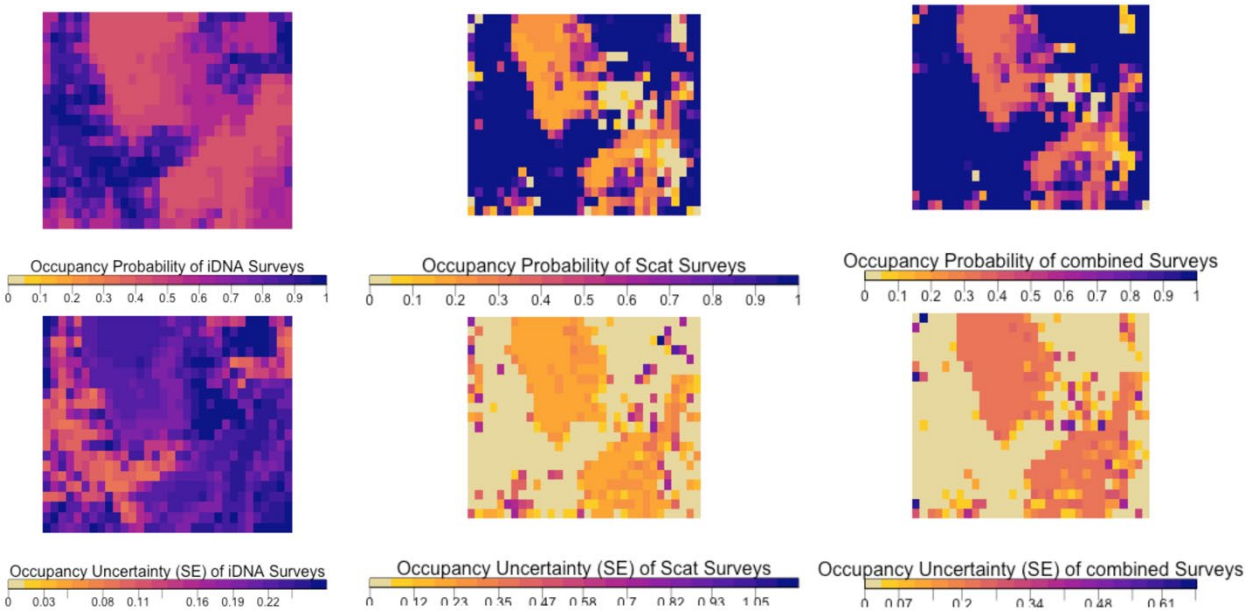


Figure 45. Optimised models: Occupancy probability and uncertainty for iDNA (left), scat survey (middle) and combined dataset (right).

Appendix II: Publications

[1] Savvantoglou, A., Mertzanis, Y., Bird, D., & Steer, M. D. (2017). A GIS approach to identifying connectivity potential between brown bear (*Ursus arctos*) habitat in northern Greece. *18th Hellenic Forestry Congress & International Workshop*, 1422–1432.

A GIS APPROACH TO IDENTIFYING CONNECTIVITY POTENTIAL BETWEEN BROWN BEAR (*Ursus arctos*) HABITAT IN NORTHERN GREECE

Savvantoglou, Angeliki[1*]; Mertzanis, Yorgos[2]; Bird, David[1]; Steer, Mark D.[1]

[1]The University of the West of England, Health and Applied Sciences, Bristol, BS16 1QY, angeliki.savvantoglou@uwe.ac.uk
[2]Callisto, Wildlife & Nature Conservation Society (NGO) – 123, Mitropoleos str. – GR-54621_Thessaloniki, Greece

Abstract

The brown bear (*Ursus arctos*) distribution in Greece is divided between two main large populations, one situated in the Pindos Mountain Range (NW Greece) and the other in the Rhodope Mountain Complex (NE Greece). As a result of recent recolonisation, a number of understudied, separate populations are found in other parts of mainland Greece. This study modelled landscape connectivity of brown bear habitat between recently recolonised areas and larger populations with regard to population dispersal potential. Using Geographical Information System (GIS) software, the species' home range and a Habitat Suitability Index (HSI) of the study area were integrated to generate an estimate of connectivity potential among known populations. The analysis suggests that recently recolonised areas may be functioning as stepping stones, connecting the core population nuclei, and identified areas which are likely to be important in terms of providing landscape connectivity.

***Keywords**: landscape ecology; fragmentation; habitat corridors; species distribution*

Introduction

European populations of brown bear (*Ursus arctos arctos*), are showing a stable or in some cases increasing trend in many parts of their range (Kaczensky *et al.* 2012; Chapron *et al.* 2014). However, the long-term survival of such a space-demanding species depends both on the management of their habitats and the linkages between suitable habitat patches (Swenson *et al.* 2011; Kaczensky *et al.* 2012). Large carnivores are particularly vulnerable to habitat loss and degradation because of their low population densities and large habitat requirements (Kaczensky *et al.* 2012; Chapron *et al.* 2014).

Landscape connectivity has been a significant focus for conservation efforts across Europe in the 21st century. Wildlife corridors connecting habitat patches promote the movement of individuals across populations, allowing gene flow and thus reducing homogeneity and the potential for inbreeding (Bennett, 2003). Furthermore, corridors can provide escape routes in case of a stochastic catastrophic event such as a wildfire (Bennett, 2003). Digital models of potential suitable patches and corridors, can facilitate an initial, rapid examination of extensive study areas and highlight areas of interest. In species conservation, when both functional connectivity and structural connectivity for the focal species between patches are considered together, the resulting model becomes more robust (Velez-Liendo *et al.* 2014; Almpanidou *et al.* 2014). Habitat Suitability Indices (HSI), introduced by the U.S. Fish and Wildlife Services (1981), are used to assess a habitat's fitness in regards to a single species. Based on the ecological behaviour of the target species, an HSI assigns scores to assess the quality of variables such as land cover, topography and distance from risk areas. HSIs can be useful in spatial studies as they can provide an insight into the animal's dispersal and habitat utilisation (Kusak and Huber 1998; Chouvardas *et al.* 2013; Almpanidou *et al.* 2014).

In Greece, there are two core brown bear populations found in the Pindos Mountain Range (NW Greece) and the Rhodope Mountain Complex (NE Greece) (Mertzanis, 2012). In Pindos, the population reaches the lowest latitudinal range of the continent (Mertzanis 2012). The northwest population is connected populations in the Former Yugoslav Republic of Macedonia (FYROM) (n = 160-200 bears) and Albania (n = 180-200 bears), while the northeast population is linked to Bulgaria (n = 530-590 bears) (Kaczensky *et al.* 2012; Chapron *et al.* 2014). In addition to the two core areas, a number of areas in mainland Greece which were within the species' historical range have been re-

colonised by brown bears in the last 40-60 years (Mertzanis 2012). Although conservation efforts have resulted in a recent population increase (Chapron *et al.* 2014), bears in Greece are still listed as Endangered in the Greek Red Data Book on Threatened Wildlife species (Mertzanis *et al.* 2009).

This was an exploratory study based on literature review and expert opinion, using digital spatial data to assess the suitability of habitats across Greece, the southern Former Yugoslav Republic of Macedonia (FYROM) and the southernmost part of Bulgaria. It attempted to use freely available spatial data to identify potential corridors, linking the newly re-colonised habitat patches and core areas. Specifically, the study focuses on the connectivity of the newly re-colonised areas and their role as large-scale stepping-stones or corridors that could facilitate the exchange of individuals between the two core areas. Furthermore, it discusses the advantages and disadvantages of this method and its potential role in aiding conservation plans on a national and international scale.

Materials and Methods

Study Area

This study focused on the smaller, recently recolonised areas in central Northern Greece extending between the core populations in the west (Pindos region) and in the east (Rhodope region), as well as southern regions of FYROM and Bulgaria; an area of 54,250km^2 (Figure 1). The study did not consider the full extent of the core areas because the focus was on the connectivity between these areas rather than the quality of the core areas themselves. For the purposes of this study, the massifs and their surrounding habitats were grouped into nine regions (Figure 1).

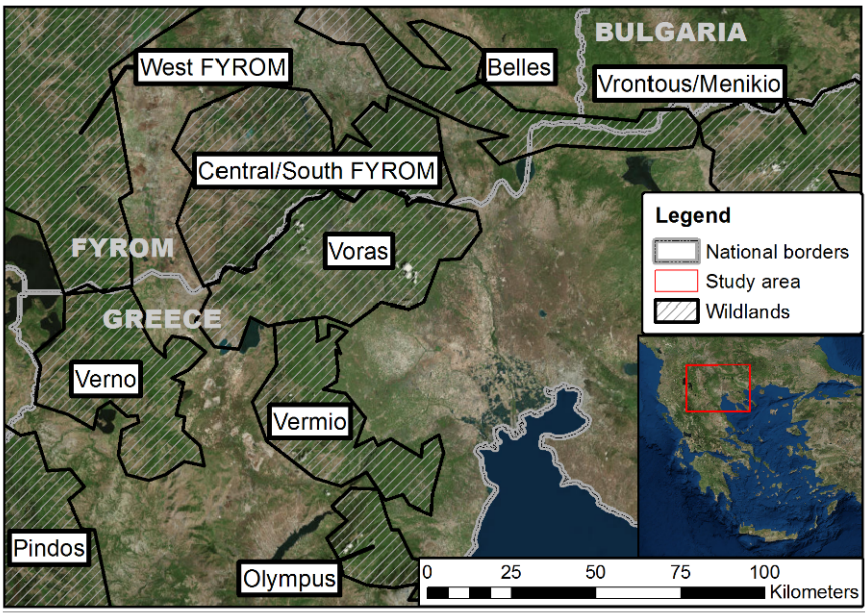


Figure 1. Study area (size: 54,250km2) and nine Wildland areas (Pindos, Verno, West FYROM, Central/South FYROM, Voras, Belles, Vrontous/Menikio, Vermio, Olympus).

Connectivity modelling

This study utilized the GIS tool CorridorDesigner, which was developed specifically for the purposes of modelling habitat suitability assessments and corridor design (Majka *et al.* 2007), operated within ArcGIS v.10.2.2 (ESRI, 2014)

A Habitat Suitability Index (HSI) was created by drawing on published data to evaluate the relative importance of four key variables (land cover, altitude, topography and distance from urban areas) and weighting these according to their importance for the target species (see tables 1-3). Land cover data was obtained from the Corine Land Cover (CLC) dataset (resolution 100 meters) (EEA 2000) to ensure data consistency for the three countries. Altitude was derived from a Digital Elevation Model (DEM; resolution 25 meters) obtained from the European Environment Agency (EEA 2013). A number of studies highlight altitudinal preferences in brown bears, especially regarding denning sites with the average altitude being reported as 863m above sea level in Croatia (Kusak and Huber 1998, see also Kanellopoulos *et al.* 2006). The optimum altitudinal range for this HSI was deemed to be between 800 m and 1700 m, as 800 m is the lowest elevation used significantly by brown bears (Kanellopoulos *et al.* 2006), while 1700 m is the average altitude where the tree line ends (low food availability and canopy cover). The lower HSI altitude groups (0-800 m) are associated mainly with foraging. There is a relative paucity of data concerning how brown bears specifically use the topography of a landscape, apart from denning sites often being located in steep canyons (Mertzanis, *et al.*, 2005). Bears often select paths of least resistance to conserve energy and thus may avoid steep terrain (Nawaz *et al.,* 2014). A Topographic Position Raster, created from the DEM, was used to generate four different categories of slope terrain (Table 2). Proximity to areas of human activity, has been identified as an important factor affecting the suitability of habitats for large carnivores (Chapron *et al.* 2014). All human activity areas found in the CLC dataset were merged to create an urban layer (Table 1). Three distance groups were identified (Table 2).

Whilst the road network is often used in HSI assessments, this study did not include the road network within the corridor analysis as it was not possible to obtain a layer of consistent quality for the entire study area. Alternatively, the road network was taken into account in the final assessment by identifying areas where roads bisect potential corridors. It is important to note that while highways and railways seem to affect bear movement negatively, forestry roads are consistently used by bears as they provide a more energy-efficient passageway within suitable habitat (Huber *et al.* 2005).

The relative importance of each of the four variables were weighted against each other in terms of their impact on brown bears and following the methods of Majka *et al.* (2013; see figure 2 for weightings).

The analysis also required an estimation of home range. Brown bear home ranges vary greatly, even amongst Southern European populations (see Kusak and Huber 1998; Kanellopoulos *et al.* 2006; Mertzanis *et al.* 2011). Brown bear home range telemetry studies in Greece varied from 25-30 km^2 for adult females with cubs of that year to as much as 507 km^2 for adult males (Kanellopoulos *et al.* 2006, Mertzanis *et al.* 2011). In an attempt to select an area suitable for both sexes, this study defined the home range as 150 km^2.

Table 1. Habitat suitability scores assigned to land cover categories (see Kusak & Huber, 1998; Kanellopoulos *et al.*, 2006; Mertzanis *et al.*, 2008; Paralikidis *et al.*, 2009; Karamanlidis *et al.*, 2014; Can *et al.*, 2014; Savvantoglou 2015).

CORINE code	CORINE Layer	Score	Justification
1.1-1.4.2	Urban areas (codes 1.1 to 1.4.2)	**0**	**Areas of high human activity**
2.1.1	Non-irrigated arable land	50	Occasionally used for feeding, but not suitable for breeding
2.1.2	Permanently irrigated land	**30**	**Potentially occasional feeding areas, but not suitable for breeding.**
2.1.3	Rice fields	10	Not suitable
2.2.1	Vineyards	**30**	**Potentially occasional feeding areas, but not suitable for breeding.**
2.2.2	Fruit trees and berry plantations	30	Very good source of food, but plantations associated with intensive agriculture. Occasional visits. Not suitable for breeding
2.2.3	Olive groves	**0**	**Unsuitable food source. Not suitable for breeding**
2.3.1	Pastures	50	Occasionally used for feeding, but not suitable for breeding
2.4.1	Annual crops associated with permanent crops	**30**	**Occasionally used for feeding, but not suitable for breeding**
2.4.2	Complex cultivation	30	Occasionally used for feeding, but not suitable for breeding
2.4.3	Land principally occupied by agriculture, with significant areas of natural vegetation	**60**	**Frequent use for feeding, possible breeding potential**
2.4.4	Agro-forestry areas	**80**	**Consistent use for feeding and breeding**
3.1.1	Broad-leaved forest	100	Consistent use for feeding and breeding
3.1.2	Coniferous forest	**80**	**Consistent use for feeding and breeding**
3.1.3	Mixed forest	90	Consistent use for feeding and breeding
3.2.1	Natural grassland	**60**	**Frequent use for feeding, possible breeding potential**
3.2.2	Moors and heathland	60	Frequent use for feeding, possible breeding potential
3.2.3	Sclerophyllous vegetation	**60**	**Frequent use for feeding, possible breeding potential**
3.2.4	Transitional woodland shrub	50	Occasionally used for feeding, but not suitable for breeding
3.3.1	Beaches, dunes, and sand plains	**0**	**Not suitable**
3.3.2	Bare rock	80	Consistent use for feeding and breeding
3.3.3	Sparsely vegetated areas	**20**	**Bears might cross in search for good habitat types**
3.3.4	Burnt areas	0	Not suitable
4.1.1	Inland marshes	**20**	**Bears might cross in search for good habitat types**
5.1	Water courses and bodies	0	Not suitable
Other habitat types from CLC2000		All missing CLC habitat types were not present in the study area	

Data processing

The *Habitat Suitability Tool* analysed the four different layers according to their weightings and produced a habitat suitability layer that reflected the given scores. The *Patch Tool* was used to reclassify all areas with habitat quality above 60 (60 = lowest value associated with occasional use and breeding) and group them according to their size (Majka *et al.* 2007). A suitable area smaller than 150 km^2 was classified as '*smaller than breeding patch*'. An area between 150 km^2 and 750 km^2 was considered a '*breeding patch*' and would be able to support a single individual or a mother with cubs. Any suitable area larger than 750 km^2 was considered a '*population patch*', an area that could support five individuals (5x the home range of the species; Majka *et al.* 2013). Finally, '*Create Corridor*

Model' developed the corridor analysis where the most suitable pathways between the patches were selected and displayed as potential linkages (Majka *et al.* 2007). See figures 2 and 3 for illustrations of the entire process.

Table 2. Habitat suitability scores of altitude, topography and distance form urban areas (synthesized from Kusak and Huber, 1998, Mertzanis *et al.* 2008, Kanellopoulos *et al.* 2006, Mertzanis *et al.* 2011, Savvantoglou 2015)

		Scores	Justification
Altitude (meters above sea level)	0 - 400	20	Rarely used in search for food
	400 - 800	60	Frequent use for feeding, possible breeding potential
	800 - 1700	100	Ideal altitudinal range
	1700-2889	50	Tree line ends and alpine vegetation begins. Occasionally used for feeding
Slope	Canyon bottom	40	Based on observations due to lack of data on this type of landscape structure in Greece
	Flat-gentle slope	80	Less energetically costly. Consistent use for feeding and breeding
	Steep slope	50	Use for commuting between places. Narrow field of vision and energetically costly. Occasionally used for feeding, not suitable for breeding.
	Ridge top	90	Best field of vision. Path of least resistance. Consistent use for feeding and breeding.
Distance from urban areas (meters)	0 - 500	50	Infrequent feeding. Not suitable for breeding.
	500 - 1500	60	Frequent use for feeding, potential breeding
	1500 and above	100	Ideal minimum distance

Table 3. Habitat suitability variable weights

Variable	Weight (%)	Justification
Land cover	65	**Highest contributor to habitat selection for Brown bears**
Topography	15	Preference for energy efficient terrain, but use of steeper areas to den or find water courses
Altitude	**10**	**Not as important as topography as bears are often observed foraging in low altitude**
Distance from urban areas	10	Small score due to the increasingly high instances of bears trespassing into inhabited areas to forage on human waste or fruit trees

Results

The tool identified fourteen potential corridors linking the wildlands, with some of the patches sharing more than one linkage (figure 3). All patches were connected to at least one other patch, while the largest number of corridors was found around the Vernio wildland (n=6). The Vrontous/Menikio region was the only patch connected to just one other patch (Belles). The average number of corridors between patches (median = 2) suggests a relatively high degree of connectivity between the brown bear populations across the study area. The length of the corridors varied from 3 km (linkage zone 4, figure 3) to 82 km long (linkage zone 5).

The model suggests a that ecological linkages exist between the Rhodope mountain range (east of Vrontous/Menikio) in the East and the Pindos mountain range in the West of Greece, and another

route connecting the FYROM populations with the newly recolonised Olympus patch via Mt. Vermio and Mt. Voras. A secondary corridor linking Mt. Olympus with the Pindos mountain range was identified, but in this case the corridor is much longer and narrower. The Belles patch in itself provides an extensive corridor due to its shape and is further connected via a short, wide corridor (approx.. 6 x 9 km) to the C/S FYROM patch. The transport network often cuts through the species' habitat, adding another barrier to bear movement and this requires further study.

The analysis suggests that bears moving between Voras and the eastern Greek population (Rhodope-Vrountous-Menikio-Belles) may be moving across two or three countries; possibly entering into Bulgaria (Vrontous - Belles), crossing to FYROM (Belles - West FYROM patch) and from that back to Greece (Voras patch).

A few remarks need to be made with regards to the application of this HSI model. Firstly, no presence/absence data was used to create this model. Presence data improves the design and refines it according to verified patterns of actual bear movement. However, this study focuses purely on the creation of a connectivity model using freely available spatial data with an aim to create a system of rapid examination of a given study area without the existence of prior field surveys, but merely the understanding of an animal's ecology and behaviour. Secondly, a few of the identified suitable habitat patches do not correspond with evidence of bear activity. Specifically, the highlighted areas in Figure 2 are not colonised by bears nor has there been evidence of bear activity, but the model has labelled them as areas of suitable habitat. This is because this model maps suitable habitat no matter whether it is colonised by bears or not. As bears are not present in these patches, the corridor modelling of these areas was excluded from the corridor design. However, this approach may provide a projection of future colonisations within the study area.

Information on the linkage zones and the corridors within them in relation to their function, physical attributes and HS values has been summarized in Table 5.

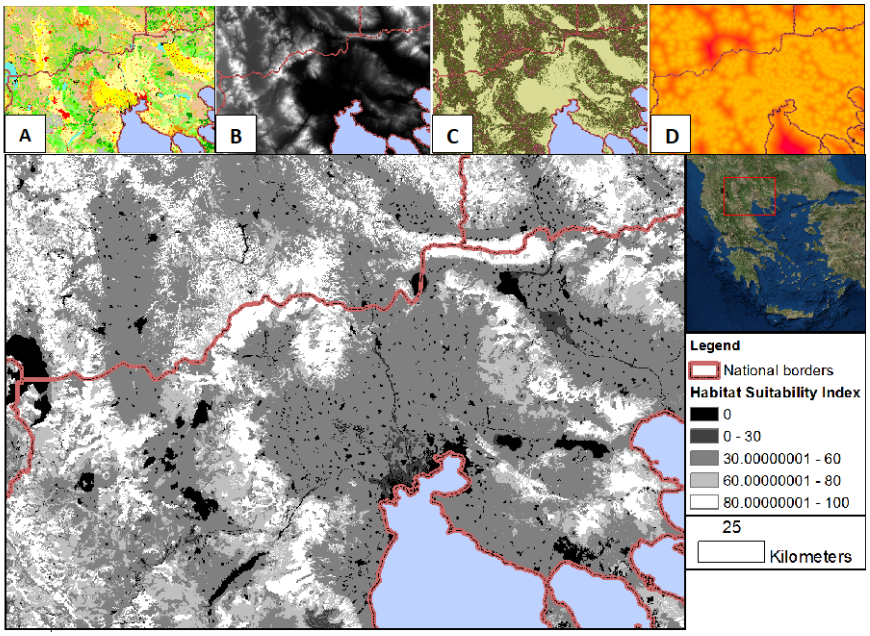


Figure 2. Corridor model: CLC (A), Altitude (B), Topography (C) and Distance from urban areas (D) developed the Habitat Suitability Index (main panel). HSI scores represent: 100 = optimal habitat associated with highest survival and reproductive

success; 80 = Lowest score typically associated with successful breeding; 60 = Lowest score associated with consistent use and breeding; 30 = Lowest score associated with occasional use for non-breeding activities; < 30 = Avoided HSI (Majka *et al*. 2013) (Layer sources: EEA 2000, EEA 2013).

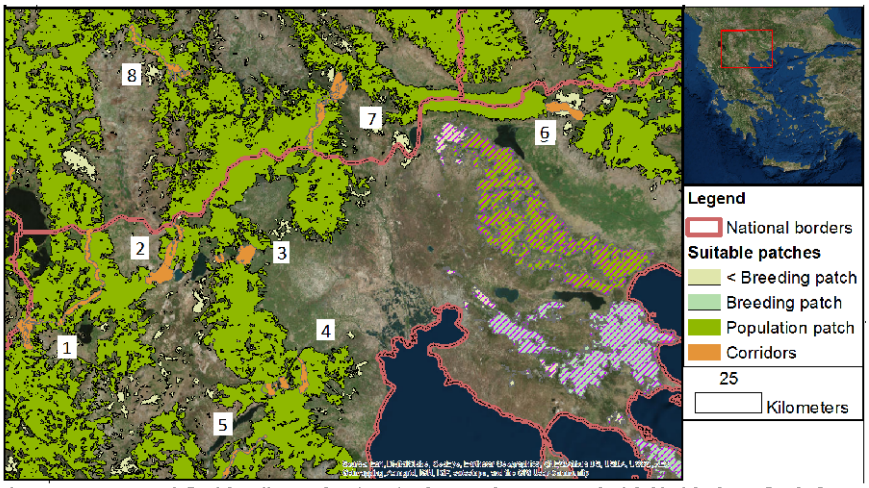


Figure 3. Home range defined breeding patches (green). The areas known not to be inhabited by bears (hatched) were excluded from the corridor model to reduce processing time. The final model revealed the potential areas likely to be serving as corridors for bear dispersal (orange). Numbers represent linkage zones between wildlands.

Table 5. Summary of linkage areas. See figure 3 for illustration of linkage zone numbers.

Linkage zone		Pindos - Vemo	Vemo - Voras	Voras - Vermio	Vermio - Olympus	Pindos - Olympus	Belles-Vrontous/Menikio	Belles - FYROM	FYROM
Linkage zone no		1	2	3	4	5	6	7	8
Signs of bear use	Confirmed	x						x	x
	Unconfirmed indications		x	x					
	Not Confirmed				x	x	x		
Type of connection	C-C								x
	S-R	x							
	C-R		x						
	R-R			x	x		x		
	C-R					x		x	
Corridor attributes	Length	9km	6km	3km	3km	82km	15km	9km	27km
	Min width	2km	4km	1km*	0.5km*	0.25km	1.5km	5.5km	0.5km
	Primary land cover (CLC)	2.4.4	2.1.1	2.4.4/3.2.2**	2.4.4/ 3.2.2**	2.4.4	3.2.2	3.2.2	2.4.4
	Connecting countries	-	-	-	-	-	Yes	Yes	-
	Overall HS value	74	65	65	75	78	68	65	79
	Bisected by major roads	-	A29	E86	E90	E92, E65	E79	A1	E75, E65
	SD of HS value	15.31	16.33	16.32	18.84	15.25	14.11	16.9	16.84

Discussion

This was a pilot study, assessing the connectivity potential for the brown bear dispersal across its Greek range and neighbouring habitat patches in FYROM and Bulgaria. The purpose of this desk-based connectivity study was to identify the potential linkage areas, allowing for a remote assessment of the study area.

It is widely understood that an HSI model is only as good as the input data (Majka *et al.* 2013). The aim of this study was to demonstrate the use of spatial corridor modelling as a tool for landscape conservation and identify its role in conservation planning. Indeed the CorridorDesigner software was able to identify those areas between the habitat patches where the habitat is, according to this HSI, able to support bear dispersal. Some of the limitations of this study have already been highlighted in the methods section, while below there are additional points regarding the reliability of the model. However, it is important to emphasise on the efficiency of such a model as an introductory step to a larger project assessing and improving connectivity between brown bear populations in its Greek range. This study was completed with freely available layers and without the use of species presence data. In the early planning stages of landscape connectivity in Greece, this study provides a rapid habitat assessment, highlighting the areas potentially acting as bear crossings. Further ground-truthing

in collaboration with researchers from Bulgaria and the FYROM, would allow for an assessment of bear activity within the highlighted corridors.

Study area connectivity

The model suggests that, in terms of human disturbance and geographical factors, the two core areas might be potentially joined via a series of smaller habitat patches and corridors. The mean length of all the corridors identified is 19.5 km, while the median is 9km. Interestingly, a number of the smaller patches, pivotal in the exchange of individuals between the East and West of the country, are areas of recent recolonisation. The Voras, Belles, and Vermio patches may be acting as larger scale stepping-stones, facilitating the movement of bears between larger populations. Perhaps, given this information, the importance of these smaller populations should be foregrounded, with a view to raising their profile in future brown bear conservation management plans. To support this recent range expansion, surveys into the intermediate areas (potential linkage areas between patches) would help to pinpoint those areas that best accommodate bear dispersal and highlight barriers in connectivity. This model shows that in order to create a robust network where bears can freely move between patches individually or across generations, the areas that require attention first are the linkage zones 1, 2, 6 and 7 (core to core area). Secondly, conserving and improving the corridors in zones 3, 4 and 5 would support the bear expansion to the south of Greece (Vermio and Olympus).

Comments on variables affecting brown bear dispersal

Possibly the two weakest points of every ecological modelling study are the lack of quality input data and information on the species life history and ecology (Almpanidou *et al.* 2014, Velez-Liendo *et al.* 2014). Regarding GIS data, detailed and homogenous spatial information is hard to obtain, especially when it comes to cross-country studies. Even though the CLC data (derived from a 25x25m Landsat-7 satellite imagery) is categorised into 44 different habitat types, a more detailed HSI model for a focal species with distinct preferences would require more detailed information on plant species and canopy cover. Additionally, very little information was obtained on how brown bears specifically use the topography of a landscape. These limitations might have led to inaccuracies in the suitable habitat patches and corridors.

Human-Bear Conflict

Wide-ranging foraging and breeding activities result in brown bears having large home ranges, often causing a variety of unfavourable human-bear interactions. It is expected that as human population increases, HBC will escalate, endangering bears across their range (Can *et al.* 2014). Therefore, it is very important for an HSI and corridor design to include areas of high risk for HBC. This study only included distance from urban areas in terms of possible HBC issues. Future modelling would benefit from information on the positions of farms, beehives and mountainous areas where livestock is taken to graze. This would help highlight the areas where HBC prevention measures could be taken and further evaluate the degree of HBC risk for the proposed corridors.

Home Range

The selection of the home range size for this study (150km^2) was done using a relatively low value for male bears (Kanellopoulos *et al.* 2006) but large enough for a female adult bear with cub(s) of the year (Mertzanis *et al.* 2005). Using three different home range groups; one for adult males (>150km^2), one for solitary females (<150km^2) and one for females with cubs (=150km^2), resulting in three different sets of patches and corridors may provide more nuanced results.

Finally, this model was designed for *U. a. arctos*, taking into account its spatial behaviour as studied in SW Europe. The classification scores and weighting of variables may differ between countries, as the spatial behaviour observed in the same bear species can vary greatly between populations (e.g. Kusak and Huber 1998, Kanellopoulos *et al.* 2006, Mertzanis *et al.* 2011). The brown bear is an umbrella species, with their large home ranges inhabited by many other animals. Although the conservation of bear habitat often benefits other smaller species, it is not necessarily the case that

the predicted corridors will aid the dispersal of other animals in the same way they could do for bears (Majka *et al.* 2013). A more holistic conservation project could attempt to create a number of HSIs for different species, project the different sets of corridors on a single map and highlight the overlapping corridors to reveal the habitat linkages between suitable areas.

Conclusions

The spatial model revealed a number of suitable patches and connecting corridors between brown bear habitats in Greece, the southern FYROM and south-western Bulgaria. Areas of recent recolonisation were found to be connected to two or three other patches, allowing the movement of bears between larger populations. This was a pilot study testing the function, advantages and disadvantages of using specialised software to map areas of suitable patches and habitat connectivity based on the spatial behaviour of a focal species in Greece. As a next step ground-truthing would be required to assess the function of the suggested areas as active corridors. Nevertheless, the potentially high degree of connectivity between suitable brown bear patches is a very encouraging find. In the face of habitat fragmentation, the future of brown bear conservation in Greece will rely greatly on the protection and expansion of wildlife corridors. If proven largely accurate, the use of this type of habitat modelling could aid the design of conservation management plans for brown bears and other umbrella species.

Acknowledgements

Sarah Almond and Paul Satchell provided significant technical support. We are grateful to two anonymous referees for useful comments.

References

Almpanidou, V., Mazaris, A.D., Mertzanis, Y., Avraam, I., Antoniou, I., Pantis, J.D. and Sgardelis, S.P., 2014. Providing insights on habitat connectivity for male brown bears: A combination of habitat suitability and landscape graph-based models. For Ecol Model. 286 (1): 37-44.

Bennett, A.F., 2003. Linkages in the Landscape: The Role of Corridors and Connectivity in Wildlife Conservation. 2nd Edition. IUCN, Gland, Switzerland and Cambridge, UK, pp. 3-18.

Can, Ö.E., D'Cruze, N., Garshelis, D.L., Beecham, J. and Macdonald, D.W., 2014. Resolving Human-Bear Conflict: A Global Survey of Countries, Experts, and Key Factors. For Conserv Lett. 7 (6): 501–513.

Chouvardas, D., Vrahnakis, M. S., Bousbouras, D., Evangelou, C., Lampou, E. and Georgiadis, L., 2013. Modelling habitat suitability of agro-silvopastoral landscapes for Brown bear (*Ursus arctos*). For J Environ Prot Ecol 14 (1): 162-171.

Chapron. *et al.,* 2014. Recovery of large carnivores in Europe's modern human-dominated landscapes. For Science. 346: 1517-1519 and supplement: 1-18.

Huber, D., Dečak, D., Frković, A., Grubešić, M., Iviček, B., Kulić, B., Sertić, D. and Štahan, Ž., 2005. Brown bear management plan for the Republic of Croatia [online]. Ministry of Agriculture, Forestry and Water Management, Department for Hunting and Ministry of Culture, Department for Nature Protection.

Kaczensky, P., Chapron, G., von Arx, M., Huber, D., Andrén, H., and Linnell, J., 2012. Status of large carnivores in Europe – update 2012 [online]. Status, management and distribution of large carnivores – bear, lynx, wolf & wolverine – in Europe. Part 1 - Summary Report. European Commission.

Kanellopoulos, N., Mertzanis, G., Korakis, G. and Panagiotopoulou, M., 2006. Selective habitat use by brown bear (*Ursus arctos* L.) in northern Pindos, Greece. For J Biol Res-Thessalon. 5: 23-33.

Karamanlidis, A.A., Stojanov, A., Hernando, Miguel de Gabriel, Ivanov, G., Kocijan, I., Melovski, D., Skrbinsek, T. and Zedrosser, A., 2014. Distribution and genetic status of brown bears in FYR Macedonia: implications for conservation. Acta Theriol. 59 (1): 119-128.

Kusak, J. and Huber, D., 1998. Brown bear habitat quality in Gorski kotar, Croatia. For Ursus. 10 (1): 281-291.

Majka, D., Jenness, J. and Beier, P., 2007. CorridorDesigner: ArcGIS tools for designing and evaluating corridors. Available from http://corridordesign.org.

Majka, D., Jenness, J. and Beier, P., 2013. Corridor Designer: Conceptual steps for designing wildlife corridors. Available from: http://corridordesign.org/designing_corridors/

Mertzanis, G., Ioannis, I., Mavridis, A., Nikolaou, O., Riegler, S., Riegler, A. and Tragos, A., 2005. Movements, activity patterns and home range of a female brown bear (*Ursus arctos*, L.) in the Rodopi Mountain Range, Greece. For Belg J Zool. 135 (2): 217-221.

Mertzanis, G., Kallimanis, A.S., Kanellopoulos, N., Sgardelis, S.P., Tragos, A. and Aravidis, I., 2008. Brown bear (*Ursus arctos* L.) habitat use patterns in two regions of northern Pindos, Greece - management implications. For J Nat Hist. 42 (5): 301-315.

Mertzanis G., Mazaris Ant., Sgardelis St., Aravidis El., Giannakopoulos Al., Godes C., Riegler S., Riegler A. and Tragos Ath., 2011. Telemetry as a Tool to Study Spatial Behaviour and Patterns of Brown Bears as Affected by the Newly Constructed Egnatia Highway – N. Pindos - Greece. In: Ondrej Krejcar *ed.* (2011) Modern Telemetry. InTech, Croatia, pp. 307-328.

Mertzanis, G. (2012) Bear - Greece. In: Kaczensky, P., Chapron, G., von Arx, M., Huber, D., Andrén, H., and Linnell, J., 2012. Status of large carnivores in Europe – update 2012 [online]. Status, management and distribution of large carnivores – bear, lynx, wolf & wolverine – in Europe. Part 2 - Country by country data. European Commission.

Nawaz, M.A., Martin, J., Swenson, J.E., 2014. Identifying key habitats to conserve the threatened brown bear in the Himalaya. For Biol Cons. 170: 198–206.

Paralikidis, N.P., Papageorgiou, N.K., Kontsiotis, V.J. and Tsiompanoudis, A.C., 2010. The dietary habits of the Brown bear (*Ursus arctos*) in western Greece. For Mamm Biol. 75 (1): 29-35.

Savvantoglou, A., 2015. A pilot study: The distribution of the brown bear *(Ursus arctos)* meta-population of Mt. Voras, N. Greece. An estimation of the human-bear conflicts and risks. Placement report, BSc. University of the West of England, UK and NGO Callisto, Greece.

Swenson, J.E., Taberlet, P. and Bellemain, E., 2011. Genetics and conservation of European brown bears *Ursus arctos*. For Mammal Rev. 41 (2): 87-98.

U.S. Fish and Wildlife Services, Department of Interior, 1981. Ecological Services Manual - Standards for the Development of Habitat Suitability Index Models. Part 103 ESM, Release no 1-18.

Velez-Liendo, X., Adriaensen, F. and Matthysen, E., 2014. Landscape assessment of habitat suitability and connectivity for Andean bears in the Bolivian Tropical Andes. For Ursus. 25 (2): 172-187.