

**Predicting areas of potential conflicts between bearded vultures
(*Gypaetus barbatus*) and wind turbines in the Swiss Alps**

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Sergio Vignali

von Manciano (IT)

Leiter der Arbeit:
Prof. Dr. R. Arlettaz
PD Dr. V. Braunisch
Institut für Ökologie und Evolution

**Predicting areas of potential conflicts between bearded vultures
(*Gypaetus barbatus*) and wind turbines in the Swiss Alps**

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Sergio Vignali

von Manciano (IT)

Leiter der Arbeit:
Prof. Dr. R. Arlettaz
PD Dr. V. Braunisch
Institut für Ökologie und Evolution

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Bern, 21. September 2021

Der Dekan
Prof. Dr. Z. Balogh



**Predicting areas of potential conflicts between bearded vultures
(*Gypaetus barbatus*) and wind turbines in the Swiss Alps**

Doctoral Thesis by Sergio Vignali

Division of Conservation Biology

Institute of Ecology and Evolution

University of Bern, Switzerland



Summary, general introduction, and general discussion of this thesis are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>. Chapter 1, Chapter 2, and Chapter 3 include their own license in the cover page.

Layout & Figures: Sergio Vignali

Illustrations: Eléonore Tran

Package logo: Sergio Vignali

CONTENTS

Summary	9
General introduction	11
Chapter 1: Modelling the habitat selection of the bearded vulture to predict areas of potential conflict with wind energy development in the Swiss Alps	27
Chapter 2: <i>SDMtune</i> : An R package to tune and evaluate species distribution models	75
Chapter 3: A predictive flight-height model for avoiding future conflicts between an emblematic raptor and wind energy development in the Swiss Alps	113
General discussion	179
Acknowledgments	191
Declaration of consent	193

SUMMARY

The alarming increase in global temperature observed over the last hundred years, driven by the use of fossil fuels, has prompted a shift towards “greener” energy production. An extensive expansion of wind power exploitation is expected in the coming years, which makes its effect on vulnerable species an issue of growing conservation concern. Among the wildlife affected by wind turbines, vultures are probably the most vulnerable avian ecological guild. They have experienced a sharp decline during the last decades and their survival in many areas is the result of targeted recovery and conservation actions. The bearded vulture (*Gypaetus barbatus*) represents an emblematic example. After having been extirpated from the European Alps, the species once again inhabits its former habitat, thanks to the massive long-lasting effort of a dedicated reintroduction programme. There are concerns, however, that the sprawl of wind turbines in the Alpine massif will jeopardise this successful population recovery. The main goal of this PhD thesis was therefore to predict areas in the Swiss Alps where conflicts between bearded vulture conservation and wind energy development are likely to occur, thus allowing for a more biodiversity-friendly spatial planning of wind turbines. Using a spatially explicit modelling framework with combined information of casual observations and GPS data, I predicted species’ potential distribution as well as its flight behaviour in relation to landscape, wind, and foraging conditions. First, I investigated the species ecological requirements in relation to season and age and translated these into distribution maps covering the whole Swiss Alpine arc. Here the focus was on evaluating the ability of the models to predict the possible future expansion of the species, a crucial point for anticipating potential conflicts arising from the spread of wind energy. During this process, I secondly had to delve into methodological challenges, especially with regard to taking objective decisions for model tuning. Based on the example of modelling the distribution of the bearded vulture, I introduced a new genetic algorithm for hyperparameters tuning, which drastically reduces computation time while achieving a model performance comparable or equal to that obtained with standard methods. Moreover, I generalised the developed routines so as to make them applicable to the most common species distribution modelling techniques and compiled the solutions in an R package now available to the scientific community. Thirdly, I explored the flight height patterns of bearded vultures to identify key factors driving low-height flight activity and delineated areas where the species is likely to fly within the critical height range that is typically swept by the blades of modern wind turbines.

Overall, I found that food availability is an important driver of both distribution and low-

height flight activity of bearded vultures. Habitat selection differed between seasons and between age classes during the cold season. While food availability and geological substrates were the main drivers of the distribution during the warm season, I observed a shift in the requirement of adult birds in the cold season, where habitat selection was mainly influenced by climatic conditions. This suggests that adult birds may be constrained by favourable winter conditions for the selection of breeding territories. Combining the ecological requirements of both age classes and seasons I found that 40% of the Swiss Alps offers suitable habitat for the species. The model trained with species data collected between 2004 and 2014 was able to accurately predict new breeding territories established in 2015 – 2019, and thus adequately delineated areas where the spreading population will likely to occur in the future and where conflicts with wind energy development might arise. The flight-height analysis of the GPS-tagged birds revealed that bearded vultures mainly fly within the critical height range swept by the turbine blades (77.5% of GPS locations), which poses the species at high risk of collision. Flying at low heights most frequently occurred along south exposed mountainsides and in areas with a high probability of ibex (*Capra ibex*) presence, a key food source for bearded vulture. Synthesising the information on bearded vulture distribution with the flight height behaviour allowed identifying and mapping areas where the species is likely to fly at risky height within its habitat. This high resolution, spatially explicit information represents a valuable tool for planners involved in wind energy development as well as a first basis for detailed impact assessments, while the methodological framework I developed represents a transferable approach for scientists studying potential conflicts between the development of aerial infrastructure and other target organisms.

GENERAL INTRODUCTION

The rapid demographic and economic growth of recent decades has been coupled with an increasing worldwide need for energy and a consequent sharp rise in CO₂ emissions (Dong et al., 2018). This scenario has raised awareness and concern about the negative effects of greenhouse gases on the environment and renewable energy sources have been introduced as alternatives to fossil fuels. Accordingly, several countries have started to support green energy production by means of energy policies. The European Union has set ambitious and specific targets in this direction: the aim to produce 20% of its energy from renewable sources by 2020 (Directive 2009/28/EC) was recently extended to 32% by 2030 (Directive (EU) 2018/2001), with the challenging goal to become the first “climate-neutral” continent by 2050 (European Commission, 2019). Various technologies have been developed over time to use renewable energy sources, including those to harness wind power. Given the rapid development of the wind energy industry (Lee & Zhao, 2020), wind power is likely to become one of the energy sources that will lead the transition towards green energy production worldwide. Also in Switzerland wind energy is being actively developed. Although not many wind turbines have been erected as to 2020, the government aims to reach an yearly wind energy production of 4300 GWh by 2050 (Bundesamt für Raumentwicklung ARE, 2017). Considering that less than 200 GWh per year are currently produced, the erection of numerous new wind turbines is envisioned in the coming years. However, around two third of the Swiss territory consists of mountain areas, representing a fragile ecosystem characterised by highly specialized species communities. There is thus a risk that the spread of wind turbines within these areas will have a negative impact on this valuable habitat and the wildlife it harbours.

Wind energy – wildlife conflicts

There is a general agreement that wind energy is a clean and environmentally friendly energy source. Although this holds true when compared to conventional fossil fuels, it still has negative effects on humans, climate, and especially on the environment and wildlife (rewied for example in Dai, Bergot, Liang, Xiang, & Huang, 2015; Leung & Yang, 2012). Given the rapid development of the wind industry, ignoring the negative effects, even those that so far seem only minor, could cause serious consequences in the future (Dai et al., 2015; Leung & Yang, 2012). Many studies have shown the negative impact on wildlife, whether marine or terrestrial. Offshore wind farms impact both flying vertebrates and aquatic fauna, with negative effects ranging from behavioural

changes to displacement and direct mortality of fish, marine mammals, sea turtles, birds and bats (Goodale & Milman, 2016). Offshore wind energy plants are still less numerous than their onshore counterparts (Lee & Zhao, 2020) as their location is largely limited to shallow water. However, new technologies have been developed with new floating turbines, which can be moored in deeper water. This technology allows to erect wind turbines far from the coast, paving the way for a further development of this sector, while their impact on wildlife remains still unclear and needs further research (Farr et al., 2021). Also onshore wind turbines cause several negative effects on wildlife. Numerous studies have shown their impact on insects (Voigt, 2021), bats (Arnett et al., 2008, 2015; Frick et al., 2017; Hein & Schirmacher, 2016; Rydell et al., 2010; Wellig et al., 2018), and birds (Coppes et al., 2020; Drewitt & Langston, 2006; Madders & Whitfield, 2006), with the main negative effects being displacement (Dohm et al., 2019; Millon et al., 2018; Taubmann et al., 2021) and mortality (Aschwanden et al., 2018; Johnson et al., 2002; Perold et al., 2020; Zimmerling & Francis, 2016). Among birds, large soaring raptors are the most affected by collision (Carrete et al., 2009; Dahl et al., 2012; de Lucas et al., 2012; Ferrer et al., 2012; Heuck et al., 2019; Katzner et al., 2017; Madders & Whitfield, 2006). Most raptors, such as vultures, are long-lived species with a low reproductive rate, often coupled with a delayed maturity. With this demographic predisposition, even a small increase in mortality could be detrimental (Beston et al., 2016; Carrete et al., 2009; Schaub et al., 2009; Watson et al., 2018). Moreover, vultures have a limited forward visual field which limits their vision ahead (Martin et al., 2012) and consequently increases the risk of collision with turning turbine blades. In addition, vultures need specific wind conditions to sustain their soaring flight, conditions that are often available in areas selected for wind energy development (Poessel, Brandt, et al., 2018; Rushworth & Krüger, 2014). As vultures have experienced a sharp decline over the last decades (Ogada et al., 2012, 2016; Safford et al., 2019), there is a serious concern that the expected development of wind energy will pose a serious threat to them. This, together with the above mentioned traits, makes vultures suitable model species for assessing potential conflicts between wind turbine construction and raptor conservation.

The bearded vulture

The bearded vulture is a long-lived large soaring raptor with a bone-based diet (mainly marrow) (Margalida et al., 2009; Margalida & Villalba, 2017) that inhabits several mountain areas of the Old World. It breeds on cliffs and manifests a late sexual maturity with a remarkably delayed age at first successful reproduction (10.4-11.4 and 8-9 years on average in the Pyrenean and Alpine

populations, respectively) (Antor et al., 2007; López-López et al., 2013; Lörcher & Hegglin, 2020). Although bearded vultures usually lay two eggs, only one chick survives so reproductive rate is typically low (Margalida et al., 2003, 2004; Schaub et al., 2009). The species is sparsely distributed across the southern part of the Palearctic region and in the southern and eastern parts of the Afrotropical region. Based on morphological characteristics, Hiraldo et al. (1984) distinguished two subspecies: *Gypaetus barbatus barbatus* occurring in Nord Africa and Eurasia, and *Gypaetus barbatus meridionalis* present in South and East Africa. There are clear morphological differences between the two subspecies which probably result from adaptation to different environmental conditions, as the subspecies *barbatus* inhabits mountainous regions with a harsher climate (Delibes et al., 1984; Margalida, Negro, et al., 2008). Many bearded vulture populations have suffered strong declines and their distributional ranges have shrunk over time, as observed in Nepal (Acharya et al., 2010) and in southern Africa (S. C. Krüger et al., 2014). Within Europe the species was extirpated from the Alps (Mingozzi & Estève, 1997), the Balkans (Andevski, 2013), Andalucía in Spain (Simón et al., 2007), and Sardinia in Italy (Schenk et al., 2004). Towards the end of the twentieth century isolated breeding populations of bearded vultures remained only in the Pyrenees (Spain and France), Crete, and Corsica (Andevski, 2013; Margalida, Heredia, et al., 2008). Given the dramatic situation, a large-scale reintroduction project started in 1978 to reinstate the species in the Alpine arc (Coton & Estève, 1990; Frey & Walter, 1989). Regular releases of juvenile birds started in Austria in 1986, followed by continental France (1987), Switzerland and Italy (1991). Lately, the project was extended to Spain (2006) and Corsica (2016) with the aim of fostering the connection among the different European populations. Since the first successful breeding event in 1997 (Heuret & Rouillon, 1998), thanks to this ambitious and successful programme, 308 chicks naturally fledged in the Alps (Llopis, 2020) and the population reached an estimated size of 300 individuals in 2020 (Lörcher & Hegglin, 2020).

The decline of the European bearded vulture populations occurring throughout the nineteenth and twentieth centuries was largely driven by direct and indirect human persecution (Mingozzi & Estève, 1997; Schenk et al., 2004; Simón et al., 2007). Shooting and poisoning was widely adopted for pest control and game management while specimens were collected for taxidermy (trophies and museums) (Mingozzi & Estève, 1997; Schenk et al., 2004; Simón et al., 2007). These threats have not completely disappeared nowadays. Poisoning, which obviously also affects other scavengers (Berny et al., 2015; Margalida, 2012; Margalida et al., 2013; Pantović & Andevski, 2018), is listed as a critical threat to the bearded vulture in the European recovery action

plan (Izquierdo, 2017). Indeed, Margalida et al. (2008) found an increase in mortality due to intentional and unintentional poisoning in Europe since the start of the releasing project. Moreover, Krüger et al. (2015) ascribed half of the bearded vulture fatalities found during their study in Southern Africa to poisoning events. Although poisoning is still the major source of mortality worldwide, disturbance is also recognized as a threat to the species, since it negatively affects breeding success (Arroyo & Razin, 2006). Further threats are lead exposure (Hernández & Margalida, 2009; Margalida et al., 2013), the use of veterinary drugs like diclofenac (Acharya et al., 2010), and collisions with aerial infrastructures such as powerlines or other aerial cables (Margalida, Heredia, et al., 2008). Recently, a new threat, wind energy development, potentially impacts the species, adding to the many threats listed above.

The bearded vulture is considered globally as near threatened by the IUCN Red List (BirdLife International, 2017), vulnerable in Europe (BirdLife International, 2015a) and critically endangered in Switzerland (Keller, Gerber, et al., 2010) where it is also included among the priority species requiring a targeted recovery programme (Keller, Aye, et al., 2010).

Bearded vulture and wind energy in Switzerland

Given their delayed sexual maturity and the low reproductive rate, bearded vultures are particularly vulnerable to any anthropical source of mortality. Schaub et al. (2009) prospected that an increase of 50% in the annual mortality would lead to a decline of the Alpine bearded vulture population. If the development of wind energy within the Swiss Alps doesn't account for this emblematic species, there is a high risk that the successful reintroduction project will be jeopardised. The Swiss concept for wind energy development excludes buffer areas of five kilometers around nesting locations known in 2014 (Bundesamt für Raumentwicklung ARE, 2017) as a measure to limit collision risk for bearded vultures. Similarly, Horch et al (2013) proposed to exclude buffer areas of 15 kilometers around breeding and releasing sites. Since 2014 several new breeding sites occurred as consequence of the bearded vulture population expansion, demonstrating that this method falls short to safely protect the species. As Switzerland hosts most of the breeding territories of the Alpine population (Lauper, 2020), more effective planning is required. A planning that can effectively protect both the present and future bearded vulture population.

Assessing conflicts between wind energy and wildlife

Several methods have been used to estimate the impact of wind energy on wildlife. Concerning

conflicts related to avifauna, many studies have focused on the effect of a certain wind energy project. Data are either collected by direct observations from vantage points or indirectly through radio telemetry or GPS devices to estimate densities, space use and behaviour of the studied species in the surrounding of wind turbines and thus assess the negative effects they might face from the realisation of a wind energy project (Strickland et al., 2011). For example, the collected data can be used to estimate the risk of collision, either empirically (Hoover & Morrison, 2005) or by using tailored collision risk models (Masden & Cook, 2016), or to assess displacement (Dahl et al., 2012; Dohm et al., 2019). Even though these methods are crucial for assessing a project's impact prior to its implementation or for evaluating the negative effects after its construction, they are mostly site-specific and cannot be generalised to other areas (but see Coppes et al., 2020). Other approaches look at the problem at a different scale. They aim to provide information which allows identifying broad areas of conflict, thus enabling strategic wind energy development planning (e.g. Braunisch, Coppes, Bächle, & Suchant, 2015). The simplest of these approaches relies on banning wind energy development within buffer areas around sensitive locations (Bright et al., 2008; Janss et al., 2010; Venter et al., 2019). Other approaches equate areas of conflict with areas where sensitive species occur (Allinson, 2017; Bright et al., 2008; Mc Guinness et al., 2015). Although the latter two are valid methods to delineate conflict zones, they fall short with regard to two key aspects, namely to estimate conflicts in areas where data are deficient or to anticipate conflicts in situations of non-equilibrium of the population, e.g. when dealing with reintroduced species that do not yet occupy all suitable habitat. Predictive models in contrast, can be used to project their predictions also onto areas where data are missing or, in the case of species distribution models (SDMs), to predict the potential distribution of a species, and are thus powerful tools for anticipating future conflicts.

However, for constructing valid models, many decisions have to be taken at various stages of the model building process, all of them potentially affecting the model predictions (Guisan & Thuiller, 2005). An important decision is the selection of environmental variables which might be particularly challenging when the ecology of the species is barely known or a choice should be made among ecologically equally relevant predictors. Moreover, many of the available predicting model algorithms require defining the values of their hyperparameters before model training, values that are specific of the modelling problem and dataset. Tuning the model hyperparameters is very time-consuming because it involves training and evaluating as many models as the possible combinations of their hyperparameters' values. Therefore, there is a need of new methods to efficiently automate those decisions and objectively decide among predictors.

Scope of the thesis

The main objective of the thesis was to develop a spatial modelling framework that allows identifying and anticipating conflict zones where the construction of wind energy infrastructure potentially interferes with vulnerable species, especially raptors. Using the example of the bearded vulture in the Swiss Alps, the aim was to 1) enhance methodological approaches to predict the spatio-temporal behavior of the target species and 2) provide a precise spatial planning tool that can support decision makers in avoiding turbine placement in key areas for species conservation. To achieve this goal I first had to refine our understanding of the species' ecological requirements and then translate these into detailed spatially explicit information on species occurrence and behavior. Best practice suggests that wind turbines should be located away from areas of conservation importance, where endangered species occur and concentrate their activities (Arnett et al., 2015; Drewitt & Langston, 2006). However, when dealing with released species, such as the bearded vulture, it must be considered that they may spread into new, previously unconsidered areas, thus generating unforeseen conflicts. At first, I assessed the ecological requirements of the species and predicted its current and potential distribution across the entire Swiss Alpine arc so as to identify, on a broad-scale, the areas where environmental conditions are suitable for the bearded vulture, i.e. areas where attention should be focused from the perspective of wildlife-friendly wind turbine planning. I then delved into the methodological problems that arose during the analysis of the bearded vulture distribution, developed new methodological approaches to perform variable selection and hyperparameters tuning for SDMs, and integrated them into a novel R package released in CRAN. On a second step, to refine this model and predict hotspots of potential conflict within the area of species occurrence, I analysed the flight height patterns of bearded vultures in relation to environmental, wind, and foraging conditions. Finally, I combined predictions of habitat selection and flight behavior to provide evidence-based and spatially explicit information on potential airspace conflicts between bearded vulture and wind energy that can assist policy makers, environmental consultants, wildlife managers, and wind energy companies in the selection of areas for wind turbine installation that are as little detrimental to the species as possible, already at an early stage of planning.

Thesis outline

This thesis consists of three chapters.

Chapter 1 focusses on the ecological requirements and potential distribution of the bearded vulture in the Swiss Alps. Starting from casual observations of bearded vultures I studied the habitat selection of the species accounting for potential differences between immatures and adult individuals, i.e. distinguishing between the prospecting and settling phase, and between the warm and cold season, representing different weather conditions and consequential constraints with regard to foraging and reproduction. Given that the species has been reintroduced in the Alps, released birds are gradually recolonising their former range. Areas that have not been recolonised yet might see the establishment of the species in the coming decades due to a rapidly expanding population. Therefore, one of the main objectives was to properly evaluate the capacity of the models to correctly predict the future spatial expansion of the species in Switzerland. A problem that arises when using casual observations collected without systematic sampling scheme is that the data might potentially be biased towards easy accessible areas (Fourcade et al., 2014). For that endeavor we tested different bias-correction methods in order to avoid flawed predictions. The outputs of this Chapter are summarised in maps representing the probability of bearded vulture occurrence in each season and age class. I also combined the information of each season and age class into a final map that ranks suitable habitat according to the probability that an area would be used year round by all age classes, thereby highlighting sensitive areas and thus providing a useful and readable tool for stakeholders.

Chapter 2 describes in detail the novel methodologies I developed for the species distribution models applied in Chapter 1, so as to make the routines available to other researchers and practitioners and thus enhance reproducible research. The main contributions of this Chapter are the automation of variable selection processes based on the signal produced by several evaluation metrics (data-driven variable selection) and the use of a genetic algorithm to reduce the computation time required to select among numerous hyperparameters' values. Although in my work I used a single modelling approach, i.e. Maxent (Phillips et al., 2006), I generalised these routines to make them applicable to the most common SDM algorithms. Moreover, I complemented each algorithm by interactive charts showing in real-time the effect of each implemented step on model performance to ease the understanding of the underlying process.

In Chapter 3 I analysed a large dataset of GPS locations collected by bearded vultures within Switzerland to understand which environmental conditions drive the flight height patterns of the species. Based on the assumption that flight activity occurring below 200 meters above ground level would be highly risky for the species if a wind turbine were present, I modelled the probability of a bearded vulture flying below this critical height given environmental, foraging, and wind conditions. The model was projected to the entire Swiss Alpine range and combined with the outputs of Chapter 1 to obtain detailed spatially explicit information of flight activity within the species habitat. In that way I could delineate areas suitable for the species where flights would likely occur at risky height from the perspective of the species conservation. This map embraces the whole complexity of habitat selection and habitat use and allows a better siting of wind turbines within the Swiss Alps.

REFERENCES

- Acharya, R., Cuthbert, R., Sagar Baral, H. E. M., & Chaudhary, A. (2010). Rapid decline of the Bearded Vulture *Gypaetus barbatus* in Upper Mustang, Nepal. *Forktail*, **26**(26), 117–120.
- Allinson, T. (2017). Introducing a New Avian Sensitivity Mapping Tool to Support the Siting of Wind Farms and Power Lines in the Middle East and Northeast Africa. In *Wind Energy and Wildlife Interactions* (pp. 207–218). Springer International Publishing.
https://doi.org/10.1007/978-3-319-51272-3_12
- Andevski, J. (2013). *Vulture Conservation in the Balkan Peninsula and Adjacent Regions*. Vulture Conservation Foundation.
- Antor, R. J., Margalida, A., Frey, H., Heredia, R., Lorente, L., & Sesé, J. A. (2007). First Breeding Age in Captive and Wild Bearded Vultures *Gypaetus barbatus*. *Acta Ornithologica*, **42**(1), 114–118. <https://doi.org/10.3161/068.042.0106>
- Arnett, E. B., Baerwald, E. F., Mathews, F., Rodrigues, L., Rodríguez-Durán, A., Rydell, J., Villegas-Patraca, R., & Voigt, C. C. (2015). Impacts of wind energy development on bats: A global perspective. In *Bats in the Anthropocene: Conservation of Bats in a Changing World* (pp. 295–323). Springer International Publishing. https://doi.org/10.1007/978-3-319-25220-9_11
- Arnett, E. B., Brown, W. K., Erickson, W. P., Fiedler, J. K., Hamilton, B. L., Henry, T. H., Jain, A., Johnson, G. D., Kerns, J., Koford, R. R., Nicholson, C. P., O’Connell, T. J., Piorkowski, M. D., & Tankersley, R. D. (2008). Patterns of Bat Fatalities at Wind Energy Facilities in North America. *Journal of Wildlife Management*, **72**(1), 61–78.
<https://doi.org/10.2193/2007-221>
- Arroyo, B., & Razin, M. (2006). Effect of human activities on bearded vulture behaviour and

-
- breeding success in the French Pyrenees. *Biological Conservation*, **128**(2), 276–284.
<https://doi.org/10.1016/j.biocon.2005.09.035>
- Aschwanden, J., Stark, H., Peter, D., Steuri, T., Schmid, B., & Liechti, F. (2018). Bird collisions at wind turbines in a mountainous area related to bird movement intensities measured by radar. *Biological Conservation*, **220**, 228–236. <https://doi.org/10.1016/j.biocon.2018.01.005>
- Berny, P., Vilagines, L., Cugnasse, J. M., Mastain, O., Chollet, J. Y., Joncour, G., & Razin, M. (2015). VIGILANCE POISON: Illegal poisoning and lead intoxication are the main factors affecting avian scavenger survival in the Pyrenees (France). *Ecotoxicology and Environmental Safety*, **118**, 71–82. <https://doi.org/10.1016/j.ecoenv.2015.04.003>
- Beston, J. A., Diffendorfer, J. E., Loss, S. R., & Johnson, D. H. (2016). Prioritizing Avian Species for Their Risk of Population-Level Consequences from Wind Energy Development. *PLOS ONE*, **11**(3), e0150813. <https://doi.org/10.1371/journal.pone.0150813>
- BirdLife International. (2015). *Gypaetus barbatus*. *The IUCN Red List of Threatened Species 2015: e.T22695174A60116752*. Downloaded on 13 April 2021.
- BirdLife International. (2017). *Gypaetus barbatus* (amended version of 2017 assessment). *The IUCN Red List of Threatened Species: e.T22695174A118590506*. <https://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T22695174A118590506.en>. Downloaded on 13 April 2021.
- Braunisch, V., Coppes, J., Bächle, S., & Suchant, R. (2015). Underpinning the precautionary principle with evidence: A spatial concept for guiding wind power development in endangered species' habitats. *Journal for Nature Conservation*, **24**(C), 31–40.
<https://doi.org/10.1016/j.jnc.2015.01.003>
- Bright, J., Langston, R., Bullman, R., Evans, R., Gardner, S., & Pearce-Higgins, J. (2008). Map of bird sensitivities to wind farms in Scotland: A tool to aid planning and conservation. *Biological Conservation*, **141**(9), 2342–2356. <https://doi.org/10.1016/j.biocon.2008.06.029>
- Bundesamt für Raumentwicklung ARE. (2017). *Konzept Windenergie. Basis zur Berücksichtigung der Bundesinteressen bei der Planung von Windenergieanlagen*. Bern.
- Carrete, M., Sánchez-Zapata, J. A., Benítez, J. R., Lobón, M., & Donázar, J. A. (2009). Large scale risk-assessment of wind-farms on population viability of a globally endangered long-lived raptor. *Biological Conservation*, **142**(12), 2954–2961.
<https://doi.org/10.1016/J.BIOCON.2009.07.027>
- Coppes, J., Braunisch, V., Bollmann, K., Storch, I., Mollet, P., Grünschachner-Berger, V., Taubmann, J., Suchant, R., & Nopp-Mayr, U. (2020). The impact of wind energy facilities on grouse: a systematic review. *Journal of Ornithology*, **161**(1), 1–15.
<https://doi.org/10.1007/s10336-019-01696-1>
- Coton, C., & Estève, R. (1990). La réintroduction du gypaète barbu dans les Alpes. *Rev. Ecol. (Terre Vie)*, **5**(9), 227–241.
- Dahl, E. L., Bevanger, K., Nygård, T., Røskaft, E., & Stokke, B. G. (2012). Reduced breeding

- success in white-tailed eagles at Smøla windfarm, western Norway, is caused by mortality and displacement. *Biological Conservation*, **145**(1), 79–85. <https://doi.org/10.1016/j.biocon.2011.10.012>
- Dai, K., Bergot, A., Liang, C., Xiang, W. N., & Huang, Z. (2015). Environmental issues associated with wind energy - A review. *Renewable Energy*, **75**, 911–921. <https://doi.org/10.1016/j.renene.2014.10.074>
- de Lucas, M., Ferrer, M., Bechard, M. J., & Muñoz, A. R. (2012). Griffon vulture mortality at wind farms in southern Spain: Distribution of fatalities and active mitigation measures. *Biological Conservation*, **147**(1), 184–189. <https://doi.org/10.1016/j.biocon.2011.12.029>
- Delibes, M., Hiraldo, F., & Calderón, J. (1984). Age and geographic variation in the Bearded Vulture *Gypaetus barbatus* (Linnaeus, 1758). *Bonner Zoologische Beiträge*, **35**(1–3), 71–90.
- Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources. *Official Journal of the European Union*, **L 328**, 82–209.
- Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC. (2009). *Official Journal of the European Union*, **L 140**, 16–62.
- Dohm, R., Jennelle, C. S., Garvin, J. C., & Drake, D. (2019). A long-term assessment of raptor displacement at a wind farm. *Frontiers in Ecology and the Environment*, **17**(8), 433–438. <https://doi.org/10.1002/fee.2089>
- Dong, K., Hochman, G., Zhang, Y., Sun, R., Li, H., & Liao, H. (2018). CO2 emissions, economic and population growth, and renewable energy: Empirical evidence across regions. *Energy Economics*, **75**, 180–192. <https://doi.org/10.1016/j.eneco.2018.08.017>
- Drewitt, A. L., & Langston, R. H. W. (2006). Assessing the impacts of wind farms on birds. *Ibis*, **148**(s1), 29–42. <https://doi.org/10.1111/j.1474-919X.2006.00516.x>
- European Commission. (2019). *Communication from the commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. The European Green Deal. COM/2019/640 final*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0640>
- Farr, H., Ruttenberg, B., Walter, R. K., Wang, Y.-H., & White, C. (2021). Potential environmental effects of deepwater floating offshore wind energy facilities. *Ocean & Coastal Management*, **207**, 105611. <https://doi.org/10.1016/j.ocecoaman.2021.105611>
- Ferrer, M., de Lucas, M., Janss, G. F. E., Casado, E., Muñoz, A. R., Bechard, M. J., & Calabuig, C. P. (2012). Weak relationship between risk assessment studies and recorded mortality in wind farms. *Journal of Applied Ecology*, **49**(1), 38–46. <https://doi.org/10.1111/j.1365-2664.2011.02054.x>
- Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping Species Distributions

-
- with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. *PLoS ONE*, **9**(5), e97122. <https://doi.org/10.1371/journal.pone.0097122>
- Frey, H., & Walter, W. (1989). The Reintroduction of the Bearded Vulture *Gypaetus barbatus* into the Alps. *Raptors in the Modern World*, 341–344.
- Frick, W. F., Baerwald, E. F., Pollock, J. F., Barclay, R. M. R., Szymanski, J. A., Weller, T. J., Russell, A. L., Loeb, S. C., Medellin, R. A., & McGuire, L. P. (2017). Fatalities at wind turbines may threaten population viability of a migratory bat. *Biological Conservation*, **209**, 172–177. <https://doi.org/10.1016/j.biocon.2017.02.023>
- Goodale, M. W., & Milman, A. (2016). Cumulative adverse effects of offshore wind energy development on wildlife. *Journal of Environmental Planning and Management*, **59**(1), 1–21. <https://doi.org/10.1080/09640568.2014.973483>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Hein, C. D., & Schirmacher, M. R. (2016). Impact of wind energy on bats: a summary of our current knowledge. *Human-Wildlife Interactions*, **10**(1), 19–27. <https://doi.org/https://doi.org/10.26077/x7ew-6349>
- Hernández, M., & Margalida, A. (2009). Assessing the risk of lead exposure for the conservation of the endangered Pyrenean bearded vulture (*Gypaetus barbatus*) population. *Environmental Research*, **109**(7), 837–842. <https://doi.org/10.1016/j.envres.2009.05.001>
- Heuck, C., Herrmann, C., Levers, C., Leitão, P. J., Krone, O., Brandl, R., & Albrecht, J. (2019). Wind turbines in high quality habitat cause disproportionate increases in collision mortality of the white-tailed eagle. *Biological Conservation*, **236**, 44–51. <https://doi.org/10.1016/j.biocon.2019.05.018>
- Heuret, J., & Rouillon, A. (1998). Première reproduction réussie de Gypaètes barbus *Gypaetus barbatus* issus de réintroduction dans les Alpes (Haute-Savoie, France): observations comportementales du couple et du jeune. *Nos Oiseaux*, **45**, 199–207.
- Hiraldo, F., Delibes, M., & Calderón, J. (1984). Comments on the taxonomy of the Bearded Vulture *Gypaetus barbatus* (Linnaeus, 1758). *Bonner Zoologische Beiträge*, **35**, 91–95.
- Hoover, S. L., & Morrison, M. L. (2005). Behavior of red-tailed hawks in a wind turbine development. *Journal of Wildlife Management*, **69**(1), 150–159.
- Horch, P., Schmid, H., Guélat, J., & Liechti, F. (2013). *Konfliktpotenzialkarte Windenergie – Vögel Schweiz: Teilbereich Brutvögel, Gastvögel und Vogelschutzgebiete gemäss WZVV. Erläuterungsbericht. Aktualisierung 2013*. Schweizerische Vogelwarte, Sempach.
- Izquierdo, D. (2017). *European Union single species recovery plan for the Palearctic population of Bearded Vulture (Gypaetus barbatus barbatus)*.
- Janss, G. F. E., de Lucas, M., Whitfield, D. P., Lazo, A., & Ferrer, M. (2010). The precautionary

- principle and wind-farm planning in Andalucía. *Biological Conservation*, **143**(8), 1827–1828. <https://doi.org/10.1016/j.biocon.2010.02.016>
- Johnson, G. D., Erickson, W. P., Strickland, M. D., Shepherd, M. F., Shepherd, D. A., & Sarappo, S. A. (2002). Collision Mortality of Local and Migrant Birds at a Large-Scale Wind-Power Development on Buffalo Ridge, Minnesota. *Wildlife Society Bulletin*, **30**(3), 879–887. <https://www.jstor.org/stable/3784243>
- Katzner, T. E., Nelson, D. M., Braham, M. A., Doyle, J. M., Fernandez, N. B., Duerr, A. E., Bloom, P. H., Fitzpatrick, M. C., Miller, T. A., Culver, R. C. E., Braswell, L., & DeWoody, J. A. (2017). Golden Eagle fatalities and the continental-scale consequences of local wind-energy generation. *Conservation Biology*, **31**(2), 406–415. <https://doi.org/10.1111/cobi.12836>
- Keller, V., Aye, R., Müller, W., Spaar, R., & Zbinden, N. (2010). Species of national conservation concern in Switzerland: revision 2010. Die prioritaeren Vogelarten der Schweiz: Revision 2010. *Ornithologische Beobachter*, **107**(4), 265–285.
- Keller, V., Gerber, A., Schmid, H., Volet, B., & Zbinden, N. (2010). *Rote Liste Brutvögel. Gefährdete Arten der Schweiz, Stand 2010*. Bundesamt für Umwelt, Bern, und Schweizerische Vogelwarte, Sempach UmweltVollzug Nr. 1019. 53 S.
- Krüger, S. C., Allan, D. G., Jenkins, A. R., & Amar, A. (2014). Trends in territory occupancy, distribution and density of the Bearded Vulture *Gypaetus barbatus meridionalis* in southern Africa. *Bird Conservation International*, **24**(2), 162–177. <https://doi.org/10.1017/S0959270913000440>
- Krüger, S. C., Simmons, R. E., & Amar, A. (2015). Anthropogenic activities influence the abandonment of Bearded Vulture (*Gypaetus barbatus*) territories in southern Africa. *The Condor*, **117**(1), 94–107. <https://doi.org/10.1650/CONDOR-14-121.1>
- Lauper, M. (2020). Update on the Alpine bearded vulture population and the International Bearded Vulture Monitoring. *InfoGIPETO*, **37**, 2–5.
- Lee, J., & Zhao, F. (2020). GWEC Global Wind Report. In *Global Wind Energy Council*. www.gwec.net
- Leung, D. Y. C., & Yang, Y. (2012). Wind energy development and its environmental impact: A review. *Renewable and Sustainable Energy Reviews*, **16**(1), 1031–1039. <https://doi.org/10.1016/J.RSER.2011.09.024>
- Llopis, A. (2020). The Bearded Vulture EEP: results 2020. *InfoGIPETO*, **37**, 6–8.
- López-López, P., Zuberogoitia, Í., Alcántara, M., & Gil, J. A. (2013). Philopatry, natal dispersal, first settlement and age of first breeding of bearded vultures *Gypaetus barbatus* in central Pyrenees. *Bird Study*, **60**(4), 555–560. <https://doi.org/10.1080/00063657.2013.842537>
- Lörcher, F., & Hegglin, D. (2020). Life and death in the bearded vulture in the Alps. *InfoGIPETO*, **37**, 9–12.
- Madders, M., & Whitfield, D. P. (2006). Upland raptors and the assessment of wind farm

-
- impacts. *Ibis*, **148**(s1), 43–56. <https://doi.org/10.1111/j.1474-919X.2006.00506.x>
- Margalida, A. (2012). Baits, budget cuts: A deadly mix. *Science*, **338**, 192. <https://doi.org/10.1126/science.338.6104.192-a>
- Margalida, A., Arlettaz, R., & Donázar, J. A. (2013). Lead ammunition and illegal poisoning: Further international agreements are needed to preserve vultures and the crucial sanitary service they provide. *Environmental Science and Technology*, **47**(11), 5522–5523. <https://doi.org/10.1021/es401544j>
- Margalida, A., Bertran, J., Boudet, J., & Heredia, R. (2004). Hatching asynchrony, sibling aggression and cannibalism in the Bearded Vulture *Gypaetus barbatus*. *Ibis*, **146**(3), 386–393. <https://doi.org/10.1111/j.1474-919X.2004.00261.x>
- Margalida, A., Bertran, J., & Heredia, R. (2009). Diet and food preferences of the endangered Bearded Vulture *Gypaetus barbatus*: a basis for their conservation. *Ibis*, **151**(2), 235–243. <https://doi.org/10.1111/j.1474-919X.2008.00904.x>
- Margalida, A., Garcia, D., Bertran, J., & Heredia, R. (2003). Breeding biology and success of the Bearded Vulture *Gypaetus barbatus* in the eastern Pyrenees. *Ibis*, **145**(2), 244–252. <https://doi.org/10.1046/j.1474-919X.2003.00148.x>
- Margalida, A., Heredia, R., Razin, M., & Hernández, M. (2008). Sources of variation in mortality of the Bearded Vulture *Gypaetus barbatus* in Europe. *Bird Conservation International*, **18**(1), 1–10. <https://doi.org/10.1017/S0959270908000026>
- Margalida, A., Negro, J. J., & Galván, I. (2008). Melanin-based color variation in the Bearded Vulture suggests a thermoregulatory function. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **149**(1), 87–91. <https://doi.org/10.1016/j.cbpa.2007.10.017>
- Margalida, A., & Villalba, D. (2017). The importance of the nutritive value of old bones in the diet of Bearded vultures *Gypaetus barbatus*. *Scientific Reports*, **7**(1), 1–5. <https://doi.org/10.1038/s41598-017-08812-2>
- Martin, G. R., Portugal, S. J., & Murn, C. P. (2012). Visual fields, foraging and collision vulnerability in Gyps vultures. *Ibis*, **154**(3), 626–631. <https://doi.org/10.1111/j.1474-919X.2012.01227.x>
- Masden, E. A., & Cook, A. S. C. P. (2016). Avian collision risk models for wind energy impact assessments. *Environmental Impact Assessment Review*, **56**, 43–49. <https://doi.org/10.1016/j.eiar.2015.09.001>
- Mc Guinness, S., Muldoon, C., Tierney, N., Murray, A., Egan, S., & Crowe, O. (2015). Bird Sensitivity Mapping for Wind Energy Developments and Associated Infrastructure in the Republic of Ireland. *BirdWarchIreland*, 120 pp.
- Millon, L., Colin, C., Brescia, F., & Kerbirou, C. (2018). Wind turbines impact bat activity, leading to high losses of habitat use in a biodiversity hotspot. *Ecological Engineering*, **112**, 51–54. <https://doi.org/10.1016/j.ecoleng.2017.12.024>

- Mingozzi, T., & Estève, R. (1997). Analysis of a historical extirpation of the bearded vulture *Gypaetus barbatus* (L.) in the western Alps (France-Italy): Former distribution and causes of extirpation. *Biological Conservation*, **79**(2–3), 155–171. [https://doi.org/10.1016/S0006-3207\(96\)00110-3](https://doi.org/10.1016/S0006-3207(96)00110-3)
- Ogada, D., Keesing, F., & Virani, M. Z. (2012). Dropping dead: causes and consequences of vulture population declines worldwide. *Annals of the New York Academy of Sciences*, **1249**(1), 57–71. <https://doi.org/10.1111/j.1749-6632.2011.06293.x>
- Ogada, D., Shaw, P., Beyers, R. L., Buij, R., Murn, C., Thiollay, J. M., Beale, C. M., Holdo, R. M., Pomeroy, D., Baker, N., Krüger, S. C., Botha, A., Virani, M. Z., Monadjem, A., & Sinclair, A. R. E. (2016). Another Continental Vulture Crisis: Africa's Vultures Collapsing toward Extinction. *Conservation Letters*, **9**(2), 89–97. <https://doi.org/10.1111/conl.12182>
- Pantović, U., & Andevski, J. (2018). *Review of the problem of poison use and vulture poisoning in the Balkan Peninsula*. Vulture Conservation Foundation, Netherlands.
- Perold, V., Ralston-Paton, S., & Ryan, P. (2020). On a collision course? The large diversity of birds killed by wind turbines in South Africa. *Ostrich*, **2020**(3), 228–239. <https://doi.org/10.2989/00306525.2020.1770889>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**(3–4), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Poessel, S. A., Brandt, J., Mendenhall, L., Braham, M. A., Lanzone, M. J., McGann, A. J., & Katzner, T. E. (2018). Flight response to spatial and temporal correlates informs risk from wind turbines to the California Condor. *The Condor*, **120**(2), 330–342. <https://doi.org/10.1650/CONDOR-17-100.1>
- Rushworth, I., & Krüger, S. (2014). Wind farms threaten southern Africa's cliff-nesting vultures. *Ostrich*, **85**(1), 13–23. <https://doi.org/10.2989/00306525.2014.913211>
- Rydell, J., Bach, L., Dubourg-Savage, M.-J., Green, M., Rodrigues, L., & Hedenström, A. (2010). Bat Mortality at Wind Turbines in Northwestern Europe. *Acta Chiropterologica*, **12**(2), 261–274. <https://doi.org/10.3161/150811010X537846>
- Safford, R., Andevski, J., Botha, A., Bowden, C. G. R., Crockford, N., Garbett, R., Margalida, A., Ramírez, I., Shobrak, M., Tavares, J., & Williams, N. P. (2019). Vulture conservation: The case for urgent action. *Bird Conservation International*, **29**(1), 1–9. <https://doi.org/10.1017/S0959270919000042>
- Schaub, M., Zink, R., Beissmann, H., Sarrazin, F., & Arlettaz, R. (2009). When to end releases in reintroduction programmes: demographic rates and population viability analysis of bearded vultures in the Alps. *Journal of Applied Ecology*, **46**(1), 92–100. <https://doi.org/10.1111/j.1365-2664.2008.01585.x>
- Schenk, H., Aresu, M., & Marras, N. (2004). Feasibility Study on the Reintroduction of the Bearded Vulture (*Gypaetus barbatus*) in Sardinia. In H. Frey, G. Schaden, & M. Bijleveld van Lexmond (Eds.), *Bearded Vulture Annual Report* (pp. 80–104). Foundation for the

Conservation of the Bearded Vulture (F.C.B.V) Wassenaar, The Netherlands.

- Simón, M. Á., Couto, S., Carrasco, A. L., García-baquero, M. J., Godino, A., Eugenio, J., Hernández, F. J., Jiménez, E., Liñán, M., López, M., Padial, J. M., César, J., Bautista, F., Barco, M., & Macías, E. (2007). The reintroduction of the Bearded Vulture *Gypaetus barbatus* in Andalusia , southern Spain. *Vulture News*, **56**, 29–40.
- Strickland, D., Arnett, E., Erickson, W., Johnson, G., Morrison, M., Shaffer, J., & Warren-Hicks, W. (2011). *Comprehensive Guide to Studying Wind Energy/Wildlife Interactions. Prepared for the National Wind Coordinating Collaborative, Washington, D.C., USA.*
- Taubmann, J., Kämmerle, J.-L., Andrén, H., Braunisch, V., Storch, I., Fiedler, W., Suchant, R., & Coppes, J. (2021). Wind energy facilities affect resource selection of capercaillie *Tetrao urogallus*. *Wildlife Biology*, **2021**(1). <https://doi.org/10.2981/wlb.00737>
- Venter, J. A., Martens, F. R., & Wolter, K. (2019). Conservation Buffer Sizes Derived from Movement Data of Adult Cape Vultures (*Gyps coprotheres*) in South Africa. *African Zoology*, **54**(2), 115–118. <https://doi.org/10.1080/15627020.2019.1600428>
- Voigt, C. C. (2021). Insect fatalities at wind turbines as biodiversity sinks. *Conservation Science and Practice*, e366. <https://doi.org/10.1111/csp2.366>
- Watson, R. T., Kolar, P. S., Ferrer, M., Nygård, T., Johnston, N., Hunt, W. G., Smit-Robinson, H. A., Farmer, C. J., Huso, M., & Katzner, T. E. (2018). Raptor Interactions With Wind Energy: Case Studies From Around the World. *Journal of Raptor Research*, **52**(1), 1–18. <https://doi.org/10.3356/jrr-16-100.1>
- Wellig, S. D., Nusslé, S., Miltner, D., Kohle, O., Glazot, O., Braunisch, V., Obrist, M. K., & Arlettaz, R. (2018). Mitigating the negative impacts of tall wind turbines on bats: Vertical activity profiles and relationships to wind speed. *PLoS ONE*, **13**(3), e0192493. <https://doi.org/10.1371/journal.pone.0192493>
- Zimmerling, J. R., & Francis, C. M. (2016). Bat mortality due to wind turbines in Canada. *The Journal of Wildlife Management*, **80**(8), 1360–1369. <https://doi.org/10.1002/jwmg.21128>

CHAPTER 1



Modelling the habitat selection of the bearded vulture to predict areas of potential conflict with wind energy development in the Swiss Alps



Chapter 1 of this thesis is published in *Global Ecology and Conservation*:
<https://doi.org/10.1016/j.gecco.2020.e01405> under the terms of the Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Modelling the habitat selection of the bearded vulture to predict areas of potential conflict with wind energy development in the Swiss Alps

Sergio Vignali¹, Franziska Lörcher^{2,3,4}, Daniel Hegglin^{2,3}, Raphaël Arlettaz¹ & Veronika Braunisch^{1,5}

- ¹ Division of Conservation Biology, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012, Bern, Switzerland
- ² Stiftung Pro Bartgeier, Wuhrstrasse 12, CH-8003, Zurich, Switzerland
- ³ SWILD, Wuhrstrasse 12, CH-8003, Zuerich, Switzerland
- ⁴ Vulture Conservation Foundation, Wuhrstrasse 12, CH-8003, Zurich, Switzerland
- ⁵ Forest Research Institute of Baden-Wuerttemberg, Wonnhaldestrasse 4, D-79100, Freiburg, Germany

Published in *Global Ecology and Conservation* (2021) **25**: e01405

<https://doi.org/10.1016/j.gecco.2020.e01405>

ABSTRACT

Global warming impels countries to dramatically reduce their release of greenhouse gas emissions and increase their reliance on green energy, notably wind power. Yet, without cautious planning, the sprawl of wind turbines could negatively impact biodiversity, especially flying vertebrates that are otherwise already threatened. Inherent risks for vulnerable and endangered species are usually mitigated by banning constructions within buffer areas around nesting locations. This approach, however, neglects species' range dynamics and particularly falls short of protecting expanding populations, as in the case of natural returns or reintroduction programmes. We present here an alternative approach to mitigate wildlife-infrastructure conflicts, applying it to the bearded vulture, a species reintroduced in the European Alps. Combining casual observations and GPS locations of tagged individuals, we built several predictive distribution models with respect to bearded vulture age class and season and tested for models' ability to correctly predict its future expansion in the Alps. Although immature and adult birds showed different habitat selection patterns, both in summer and winter, wide areas of the Swiss Alps (40%) offer suitable habitat. The above combined information enabled correctly predicting today's use by breeding bearded vultures of previously unused areas. This study not only provides a detailed analysis of the bearded vulture's ecological requirements in the Alps but also helps delineating areas where conflicts with wind energy production and other aerial infrastructure will likely occur in Switzerland. The resulting maps provide a large-scale planning tool that companies, landscape planners and wildlife managers can use in any environmental risk assessments.

Keywords: environmental impact assessment, growing population, reintroduction program, risk mitigation, spatial planning, wildlife-human conflicts, wind energy

INTRODUCTION

Climate change and the increase in energy demand are leading to a progressive shift towards fossil-free energy production worldwide. With the Paris Climate Agreement ratified in 2015, nations have agreed to diminish global warming by adopting several measures, notably increasing the share of their total energy consumption with at least 20% renewable energy. Among the available renewable energy sources, wind power has developed rapidly during the last decades (Leung & Yang, 2012) and is now playing a key role in the energy transition, having reached a worldwide capacity of 651 GW in 2019 (Lee & Zhao, 2020).

Although wind energy does not release greenhouse gases into the atmosphere, wind turbines can have a negative impact on wildlife, especially on flying vertebrates, i.e. bats (Arnett et al., 2015; Rydell et al., 2010) and birds (Barrios & Rodríguez, 2004; Carrete et al., 2009). The main negative effects are mortality resulting from collisions with rotor blades and related power lines, and displacement caused by disturbance or habitat loss (Drewitt & Langston, 2006; Madders & Whitfield, 2006). Large soaring raptors are highly affected by collisions (Carrete et al., 2009; Dahl et al., 2012; de Lucas et al., 2012; Ferrer et al., 2012; Katzner et al., 2017). Possible explanations include: 1) a low reproductive rate which, usually combined with a late sexual maturity, makes each additional source of mortality detrimental (Beston et al., 2016; Carrete et al., 2009; Watson et al., 2018), 2) a limited visual field in the direction of movement, which reduces the perception of vertical obstacles ahead (Martin et al., 2012), and 3) the fact that the wind industry often deploys on areas with landforms and wind conditions similar to those selected by raptor species (Katzner et al., 2012; Poessel, Brandt, et al., 2018; Rushworth & Krüger, 2014). Collision with wind turbines is thus considered a major or even critical threat for some vulture species (Botha et al., 2017).

Vultures represent a highly vulnerable ecological guild (Buechley & Şekercioğlu, 2016). Their populations have steadily and dramatically declined over the last few decades in many regions (Ogada et al., 2012, 2016; Safford et al., 2019), with few exceptions (e.g. griffon vulture (*Gyps fulvus*) in Western Europe (Safford et al., 2019)). Conservation action has thus been taken to avoid global or local extinction of many vulture species. For example, the withdrawal of diclofenac as veterinary drug prevented a further decline of vultures in India, Nepal and Pakistan (Cuthbert et al., 2011; Galligan et al., 2014; Prakash et al., 2012) while a release programme was necessary to save the California condor (*Gymnogyps californianus*) from extinction in North America (Walters et al., 2010). Other reintroduction programmes were initiated in Europe to

reinstate or reinforce vulture populations, including the cinereous vulture (*Aegypius monachus*) in France (Eliotout et al., 2007) and Bulgaria (Stoynov et al., 2019), and the griffon vulture in Southern France (Sarrazin et al., 1994), in the Balkans (Demerdzhiev et al., 2014; Stoynov et al., 2018), in Sardinia (Italy, Aresu et al., 2020) and in Cyprus. The bearded vulture (*Gypaetus barbatus*) was similarly reintroduced into the European Alps (Robin et al., 2004; Schaub et al., 2009).

The massive expansion of the wind industry in recent times represents a new rapidly growing source of threat for vultures and other raptors, which calls for a meticulous spatial planning of wind plants deployment. In Europe, the precautionary principle is usually applied to protect endangered species potentially affected by wind turbine infrastructure (Braunisch et al., 2015; Kriebel et al., 2001). This typically results in wind turbines being excluded from buffer areas around sensitive locations, principally nesting sites, with buffer radii defined based on expert knowledge or estimated according to species' home range size (Bright et al., 2008; Janss et al., 2010; Venter et al., 2019). However, this approach is limited and cannot safely protect the populations of vulnerable species for several reasons. Firstly, it is static. As it is based on the extant knowledge at a specific point of time, it inherently lacks predictive power and thus cannot address nor anticipate future conflicts, which is especially problematic for highly dynamic, expanding or declining populations (Braunisch et al., 2015). Secondly, it normally accounts for habitat selection at one life stage, namely breeding, but species could use habitats diversely at different life stages or times of the year (Hirzel et al., 2004; S. Krüger et al., 2014). Spatial and temporal variations in occurrence probability are thus usually not considered, with potentially dramatic consequences for both collision-prone species and wind energy planners: expanding populations may lose potentially suitable habitat while planners may face new, unforeseen restrictions because an endangered species suddenly appears in an area that had been pre-selected for a wind farm while the species was not yet inhabiting it. This may particularly be the case for reintroduced species that have not yet re-colonized parts of their historical distribution range. A good planning instrument identifying areas of actual and future potential conflict with wind turbine installations and other aerial infrastructures should be able to predict future circumstances and incorporate information on species' ecological requirements at different life stages and times of the year.

In our study, we focused on season and age-related variations in habitat selection of an endangered vulture species, the bearded vulture, a long-lived cliff-nesting bird that was extirpated from many European countries at the beginning of the 20 century (Mingozzi & Estève, 1997). An

ambitious large-scale reintroduction programme started in 1978 to reinstate bearded vulture populations in the Alpine arc. Releases of juvenile birds began in Austria as early as 1986 (Frey & Walter, 1989), and continued regularly in three additional countries: continental France (1987), Italy and Switzerland (1991), with more recent reintroductions in Spain (2006) and Corsica (2016). Since 1986, 227 individuals have been released from a captive population while 272 chicks naturally fledged in the wild in the Alps, all the latter originally stemming from reintroduced ancestors. If Schaub et al. (2009) evidenced a high annual survival probability for the Alpine population (ca 0.88 during the first year, 0.96 afterwards), they feared that a 50% increase in the annual mortality rate would result in a population decline (Schaub et al., 2009). This means that even a few additional fatalities per year, as feared with the ongoing and future expansion of wind turbines throughout the Alps, would reverse the steady population growth and jeopardize this long-term reintroduction program whose success has been so inspirational.

Regarding species' ecological requirements in the Alpine range, if Hirzel et al. (2004) already distinguished two different age classes: juveniles/immatures (i.e. individuals typically in an exploring phase) and subadults/adults (phase of territory settlements), their model was restricted to an area of 5200 km² in the SW Swiss Alps and did not account for seasonality. The present model is not only based on much more information, including GPS tracked individuals in addition to casual observations, but it encompasses the whole Swiss Alps and considers two time-periods. Its objective is also not only to predict habitat suitability throughout the Swiss Alps but also to properly evaluate the models' potential to predict the future spatial expansion of the spreading Alpine bearded vulture population from a conservation perspective. In effect, identifying broad-scale areas that are – and will be – potentially suitable for the species in the Alps is the only way to prevent conflicts with the development of wind energy and other aerial infrastructures.

MATERIAL AND METHODS

Study area and species

The study area encompassed the entire Swiss Alpine range (Fig. 1), covering 25,808 km², which represents about 13.4% of the total area of the Alpine massif (192,753 km²). Four of the six biogeographical regions present in Switzerland (Gonseth et al., 2001) were included: Northern Alps, Inner Western Alps, Inner Eastern Alps, and Southern Alps. Elevation ranges from 191 to 4611 m a.s.l., with an average of ca 1720 m a.s.l. The area is characterized by highly heterogeneous

climate conditions, with an oceanic climate in the Northern Alps, continental and subcontinental conditions in the Inner Alps, and Insubric climate in the Southern Alps. The variety of climate conditions together with variegated geological substrates lead to a fairly inhomogeneous landscape across the study area. Since the beginning of the reintroduction programme, the bearded vulture population has been steadily increasing within the Swiss Alpine range. After first successful reproductions in 2007 in both the Western and Eastern Swiss Alps (Biollaz et al., 2011), the number of breeding pairs has increased to 21 breeding territories in 2019. The bearded vulture is a long-lived species characterized by a low fecundity rate (clutch size of two eggs with obligate siblicide) (Margalida et al., 2003, 2004; Schaub et al., 2009) and a delayed age at first successful breeding (on average 11.4 years in the Pyrenees) (Antor et al., 2007). It feeds mostly on bones (Margalida & Villalba, 2017). It is listed as critically endangered in Switzerland (Keller, Gerber, et al., 2010) and vulnerable in Europe (BirdLife International, 2015b).

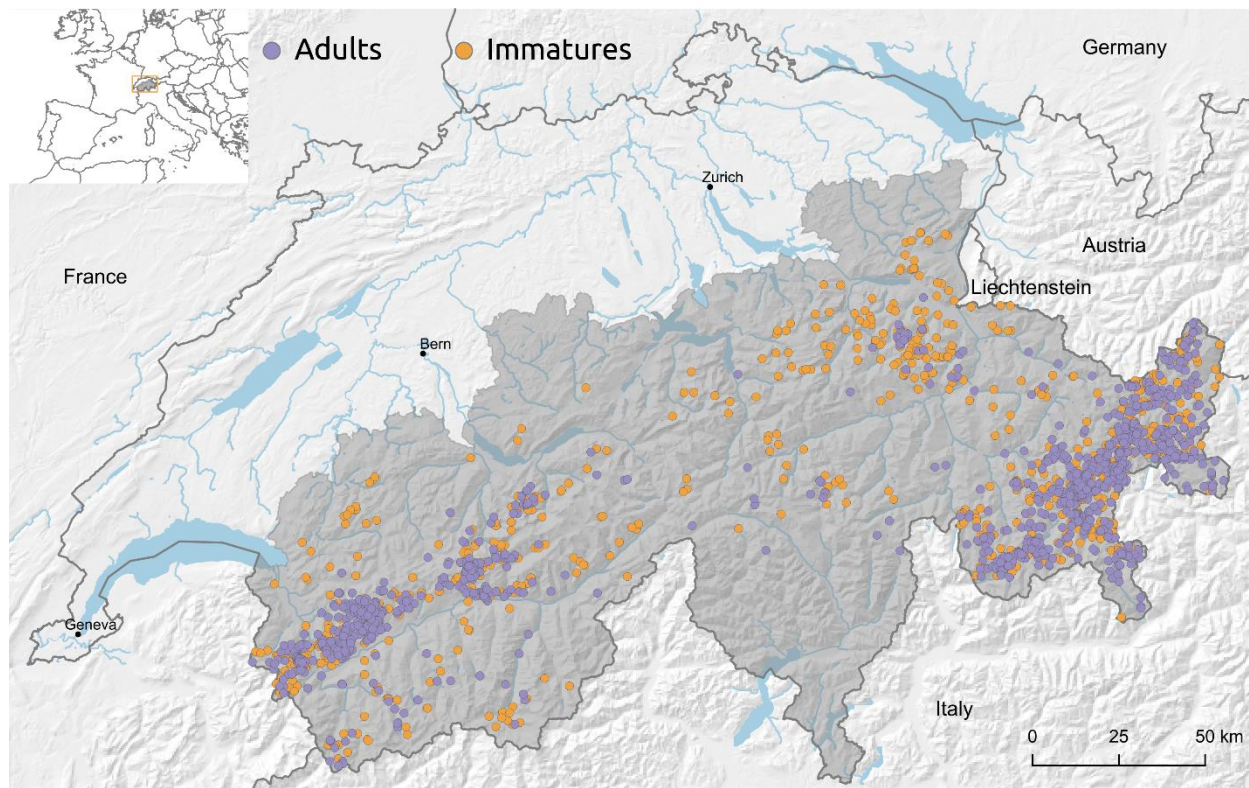


Figure 1: Map of the study area (grey shaded) that encompasses the whole Swiss Alps, with casual observations of adult (violet) and immature (orange) bearded vultures between 2004 and 2014 (unfiltered datasets).

Environmental predictors

Environmental predictors were chosen based on extant literature (e.g. Hirzel et al., 2004) and ecological knowledge of the species. The original data sources consisted of digital maps in a vector or raster format. Vector data were first converted into raster data with 25 m spatial resolution. Given that the bearded vulture habitat selection is probably driven by habitat quality over a vast area, we accounted for the conditions within a circular moving window of 564 m radius (i.e. 1 km²), corresponding to the sight-field scale of the species (Hirzel et al., 2004), by calculating the average (continuous variables) or the proportion (Boolean variables) within this range. The resulting raster maps were then resampled at 100 m spatial resolution. In that way we compiled a total of 31 environmental variables that can be grouped into six categories: 1) climate, 2) topography, 3) food availability, 4) geological substrate, 5) anthropogenic infrastructure, and 6) land cover and land use (Table 1). Solar radiation was calculated with the Area Solar Radiation tool available in the ArcGIS software version 10.2, taking 2009 as reference year. Chamois and ibex occurrence probabilities were modelled with Maxent version 3.4.1 (Phillips et al., 2006, 2017) (see supplementary material) and used directly as proxies for the potential availability of chamois and ibex carcasses. Sheep and goat densities were estimated by dividing the number of sheep and goats occurring in each community (taken as the average stock size between 2004 and 2014) by the area of alpine pastures present on the territory of the same community.

Data on wind turbines

Locations of operational wind turbines were gathered from the inventory provided by the Swiss Federal Office of Energy (Bundesamt für Energie BFE, 2014). We considered only tall wind plants within the study area, excluding small wind turbines (hub height 9–26 m, blade diameter 4–18 m) as risks of collision with low rotors is unlikely for this species. An additional dataset with the planned wind turbines was provided by the Swiss Foundation for Landscape Protection (*Windparkkarte Schweiz - Stiftung Landschaftsschutz Schweiz*, n.d.).

Table 1: Environmental predictors used for the analysis of the bearded vulture potential distribution across the Swiss Alps with unit of measurement, abbreviation and data source (warm season: May to October; cold season: November to April).

Category	Description	Unit	Abbreviation	Source
Climate	Average precipitation in the cold season	mm	prec122	WSL ^a
	Average precipitation in the warm season	mm	prec57	WSL
	Average solar radiation in the cold season	WH/m ²	solar_rad122	DHM ^b
	Average solar radiation in the warm season	WH/m ²	solar_rad57	DHM
	Average temperature in the cold season	°C	temp122	WSL
	Average temperature in the warm season	°C	temp57	WSL
	Average wind speed at 100 m above ground	m/s	windspeed100	BFE ^c
Topography	<i>Sine</i> of the aspect	-1 to 1	eastness	DHM
	<i>Cosine</i> of the aspect	-1 to 1	northness	DHM
	Standard deviation (SD) of the altitude	m	roughness	DHM
	Slope	degree	slope	DHM
	Topographic Position Index ^d	Index	tpi	DHM
Food	Ibex occurrence probability	0 to 1	ibex	CSCF ^e
	Chamois occurrence probability	0 to 1	chamois	CSCF
	Sheep and goat warm season density	N/ha	sheep_goat	BsF ^f
Geology	Limestone frequency	%	limestone	GeoKarten500 ^g
	Gneiss frequency	%	gneiss	GeoKarten500
	Granite frequency	%	granite	GeoKarten500
	Other rock frequency ^h	%	other_rocks	GeoKarten500
Anthropogenic	Distance from skilifts	m	skilift_d	Vector 25 ⁱ
	Distance from cableways, not including ski-lifts	m	cableway_d	Vector 25
	Distance from roads and railways	m	road_rail_d	Vector 25
Land cover	Bush frequency	%	bush	Vector 25
	Forest frequency	%	forest	Vector 25
	Glacier frequency	%	glacier	Vector 25
	Marshland frequency	%	marsh	Vector 25
	Grassland and unproductive vegetation frequency	%	other_veg	Vector_25 + BsF
	Permanent culture frequency	%	permanent	Vector 25
	Rock frequency	%	rock	Vector 25
	Scree frequency	%	scree	Vector 25
	Distance from lakes and rivers	m	water_d	Vector 25

^a Federal Institute for Forest, Snow and Landscape Research WSL; available upon request: www.wsl.ch.

^b Digital Height Model of Switzerland (Swisstopo): https://shop.swisstopo.admin.ch/en/products/height_models/dhm25.

^c Swiss Wind Atlas (Bundesamt für Energie BFE, 2016).

^d Topographic position index according to Wilson (1984).

^e Centre suisse de cartographie de la faune (CSCF): <http://www.cscf.ch/cscf/de/home.html>.

^f Federal Administration for Statistic Switzerland (BsF): <https://www.bfs.admin.ch/bfs/de/home/statistiken.html>.

^g Geo Maps: <https://shop.swisstopo.admin.ch/de/products/maps/geology/GK500>.

^h This environmental predictor includes gravels, sands, marl, conglomerate, and sandstone.

ⁱ Digital Cartographic Model of Switzerland (Swisstopo): <https://shop.swisstopo.admin.ch/en/products/maps/national/vector/smv25>.

Species data

Three different data sources were used for the analyses: casual observations, locations collected by birds tagged with Global Position System (GPS) devices, and nesting site locations. Casual observations, collected across the entire Swiss Alps between 2004 and 2014, were used to model the potential distribution of the bearded vulture. However, such observations, obtained without systematic sampling schemes, might suffer from a sampling bias towards easily accessible areas (Fourcade et al., 2014) and we had to account for it. Several techniques, usually referred to as thinning or filtering, have been successfully adopted to reduce the effect of sampling bias (Aiello-Lammens et al., 2015; Varela et al., 2014), but their application cannot be generalized to all datasets. The best way to assess if the application of a given filter to the data improves the model performance is to train a model with filtered data and evaluate it against an independent, systematically collected dataset (Boria et al., 2014). We therefore used GPS locations from tagged bearded vultures, which can be considered free from any bias, to evaluate different filtering methods applied to the casual observations. Moreover, the dataset containing the precise location of all known nesting sites in 2019 (irrespective of breeding success that year) was used to verify if our models, trained using the casual observations from 2004 to 2014, could correctly identify breeding areas that established after this period.

Casual observations

Observations of bearded vulture were obtained from two databases: IBM (International Bearded Vulture Monitoring) and Ornitho.ch. IBM collects all the observations of the species in an international framework, coordinating the monitoring activities of the different partners of the reintroduction project. In its database are entered only observations which have been verified by local coordinators. Ornitho.ch is the Swiss birding online platform, into which amateur bird watchers and professional ornithologists enter their observation data. Each of the retrieved bearded vulture observations (N = 3890) included date, precise location and photographs of the bird. The portraits were inspected to determine bird's age and accordingly classify the observations into two age classes: juveniles/immatures and subadults/adults. Juveniles/immatures included all individuals with a maximum age of three years (i.e. dark-brownish head and plumage, hereafter referred to as immatures), subadults/adults all older individuals (i.e. light colored head and underparts, hereafter referred to as adults). Moreover, we distinguished between observations collected during the warm season (May–October) and observations collected during the cold season

(November–April), producing a total of four distinct datasets that were analyzed separately. For each dataset, we randomly retained only one observation per 100×100 m raster pixel (hereafter referred to as unfiltered dataset, Fig. 1) to which different filtering methods were then applied.

GPS data

Within the alpine reintroduction programme framework, 81 captive-bred, young bearded vultures equipped with a GPS device were released since 2005. The GPS tag was fixed with a leg loop harness that included a weak breaking point (Hegglin et al., 2004). In addition, 10 wild-hatched fledglings (in 2016–2019) and one adult bird (2017) were tagged with the same method. All GPS data collected between 2004 and 2014 were from immatures. Therefore, assuming that the casual observations of adult bearded vultures are affected by the same sample bias as that of immature birds, we tested the different filtering techniques on the GPS locations of immatures, and then applied the best performing filter to the adult dataset. As we were not only interested in knowing the accuracy of contemporary models but also of their ability to predict future circumstances, we split the GPS positions into two time frames: fixes collected from 2004 to 2014, covering the same time period as casual observations, and fixes collected between 2015 and 2019, during which the species had been expanding into new areas not previously occupied. For this second time frame, GPS data from adults were also available (i.e. the adult tagged in 2017 and nine tags active for more than 36 months) and therefore tested in our final models.

The GPS data were prepared as follows: 1) for each tagged bird we excluded all the GPS fixes recorded during the first eight weeks, to limit the bias in favor of the releasing or nesting site; 2) all GPS positions collected outside of the Swiss study area, i.e. abroad, were excluded; 3) a subsample of two observations per month was randomly extracted from the remaining locations (we opted for two observations in order to obtain an amount of records comparable with the testing dataset held apart from the casual observations, see Table 3); 4) we merged the subsamples generated from each bird and randomly retained only one observation per raster pixel. This process was repeated for each season, age, and time frame, generating six datasets (warm and cold seasons for immatures during 2004–2014, and for immatures and adults during 2015–2019).

Filtering procedure

In order to identify the best filter type and intensity for our observation datasets, we tested two different techniques, namely the geographic filter (Aiello-Lammens et al., 2015), which selects

only one location within a predefined geographic distance, thereby removing spatial clumping, and the environmental filter (Varela et al., 2014), which subdivides the environmental gradient (e.g. of potentially bias-prone variables) into regular bins and selects only one location per bin, thereby reducing clumped samples in particular environmental conditions. Each filtering method was applied with three different intensity levels (i.e. geographic filter: different radii; environmental filter: different bin sizes). The geographic filter was performed using the function `thin` implemented in the `spThin` R package (Aiello-Lammens et al., 2019), and testing distances equal to 250, 500, and 1000 m. We tested two environmental filters, one built following the recommendation of Castellanos et al. (2019) using the first two Principal Component axes derived from all environmental variables used in our study as environmental gradients, and the other built using the distance from roads and the distance from cableways/ski-lifts, two variables defining site-accessibility in the Alps and thus hypothesized to affect sampling intensity. Both environmental filters were applied dividing the range of their variables into 200, 100, and 80 equal-sized bins. This way we tested a total of nine filters in addition to the unfiltered dataset. The filtering methods were tested separately for the warm and the cold seasons, as mountain regions are less accessible during winter due to the harsh conditions (i.e. low ambient temperature, strong wind and deep snow cover), with the sampling bias changing accordingly. The performance of each filter was finally evaluated with the independent datasets generated from GPS locations using the area under the receiver operating characteristic (ROC) curve (AUC) (Fielding & Bell, 1997).

Modelling approach

Since 1986, the first year with releases, the bearded vulture is recolonizing the Alpine range from the few sites of reintroduction scattered across the massif. In this context, data on species absence are unreliable (in terms of reflecting habitat unsuitability) given that in the future the species will probably settle in areas that have not yielded any observation yet. We therefore selected the maximum entropy approach, implemented in `Maxent` version 3.4.1 (Phillips et al., 2006, 2017) that is widely used to assess the distribution of a species when only presence data is available. This method discriminates environmental conditions at presence locations against the overall conditions prevailing in the study area (i.e. at background locations) to find the probability distribution that maximizes the entropy while fitting the best possible distribution of the species data.

Each of the ten datasets prepared from the casual observations of immature bearded vultures per season (unfiltered dataset plus the nine filtered datasets, see above) was randomly split into two

parts. One part, comprising 70% of the observations, was used for model training, and the remaining part (30%) was used as a testing dataset to evaluate the final model. The training dataset was further split into five random folds to perform cross validation and each model was trained using the same set of 10,000 randomly selected background locations. The whole analysis was conducted in R (R Core Team, 2019) through the RStudio software (RStudio Team, 2018) by using the R package `SDMtune` (see Vignali et al., 2020, for the explanation of the functions used in this analysis). For each dataset, the model was selected according to the following procedure. First, we reduced the initial set of environmental variables by removing highly correlated predictors (Spearman's $|rs| \geq 0.7$ based on 30,000 random locations) using the function `varSel`. Model hyperparameters were set to default values (i.e. feature classes: linear, quadratic, product and hinge; regularization multiplier: 1, maximum number of iterations: 500). We used the permutation importance to rank variable contributions and the mean training AUC across the five cross validation folds as performance metric. Second, we fine-tuned the models' hyperparameters with the `optimizeModel` function using only the selected variables, checking for the increase in the mean validation AUC across the five cross validation folds. The best set of hyperparameters was searched among the following values: 1) feature classes combinations: lq, lh, lqp, lqh, lph, lqph, with l representing linear, q quadratic, p product and h hinge; 2) regularization multiplier ranging from 0.1 to 2.9 with increments of 0.2; 3) maximum number of iterations: 300, 500 or 700. Third, starting from the model with the tuned hyperparameters combination, we removed variables with a low permutation importance ($<2\%$) by means of the `reduceVar` function, using the Jackknife approach to control for the decrease in the mean validation AUC across the five cross validation folds. Finally, we trained the final, optimized model with all data (i.e. without cross validation).

We evaluated the model's performance by computing the AUC for two testing datasets: the one generated from the same pull of data (i.e. 30% of data held apart), and the one prepared from the GPS locations collected in the same time period (between 2004 and 2014). These two different evaluation datasets were used to select the best filtering approach, and consequently the best model, for the given season. The best season-specific filter was then applied to the corresponding dataset of casual observations of adult bearded vultures and the same method as described above was applied to perform variable and model selection.

Each of the final models was projected to the full extent of the study area by applying the cloglog transformation (Phillips et al., 2017) to the raw output of the model. The four resulting raster maps, representing the predicted probability of species occurrence, were aggregated by

selecting the maximum value of the four overlapping pixels. This way we summarized the results to a final unique map that accounts for the requirements of both age classes during both seasons. Probability maps were converted into presence/absence maps using the threshold value that maximized the sum of sensitivity (i.e. proportion of presence locations correctly predicted) and specificity (i.e. proportion of absence locations correctly predicted) on the training dataset. This threshold has been suggested as a good criteria to convert continuous probabilities into a binary map when only presence data are available (Canran Liu et al., 2013). As a final step, we created a map of potential bearded vulture – wind turbine conflicts (hereafter referred to as potential conflict map) by summing the values of the four presence/absence maps. This map has pixel values ranging from zero to four, with zero indicating areas not suitable for the species (i.e. predicted as absence in each of the four models), and four suitable areas for both age classes during both seasons, and therefore areas particularly sensitive with regard to potential threats to the species.

RESULTS

Species data

A total of 2364 casual observations were retained in the unfiltered datasets, with 630 and 474 observations of immatures during the warm and cold season, respectively, and 719 and 541 observations of adults during the two seasons, respectively (Fig. 1). The number of locations sampled from the GPS fixes varied depending on the number of tagged birds (44 used in the analysis) available for the respective time frame, season and age class (Table 2). On average, a GPS device was active for 25 months with a range of 4–69 months.

Filtering

Applying a filter to the casual observations generally improved the predictions on the independent datasets generated from the GPS locations, for both seasons (Fig. 2). However, the environmental filter built using the distance from roads and the distance from cableways/ski-lifts outperformed the other tested filters. Accounting for the highest testing AUC values on the GPS dataset, we selected the environmental filter created with 80 equal bins (e80 in Fig. 2) for the cold season. For the warm season, the environmental filters with 100 and 200 equal bins had a similar effect. We decided for the one with 200 equal bins (e200 in Fig. 2) given that it holds a narrowed difference between training and testing AUC. Contrasting the bias-corrected predictions against those generated with the unfiltered datasets indicates an over-prediction of vulture occurrence by the

latter in most of the study area, especially in the better accessible places (Fig. 3a–b). This effect was less substantial during the warm season (occurring in 58% of the total area) compared to the cold season (in 81% of the total area). Regarding binary predictions of presence and absence (Fig. 3c–d), the application of the filter had an opposite effect for the two seasons. During the warm season the major effect was a reduction of the area predicted as presence (6.9% of the area predicted as presence by the model trained with the unfiltered dataset converted to absence, and 3.9% from absence converted to presence). In contrast, during the cold season, the major effect was an increase of the area predicted as presence (2.1% of the area predicted as presence converted to absence, and 8.1% from absence converted to presence).

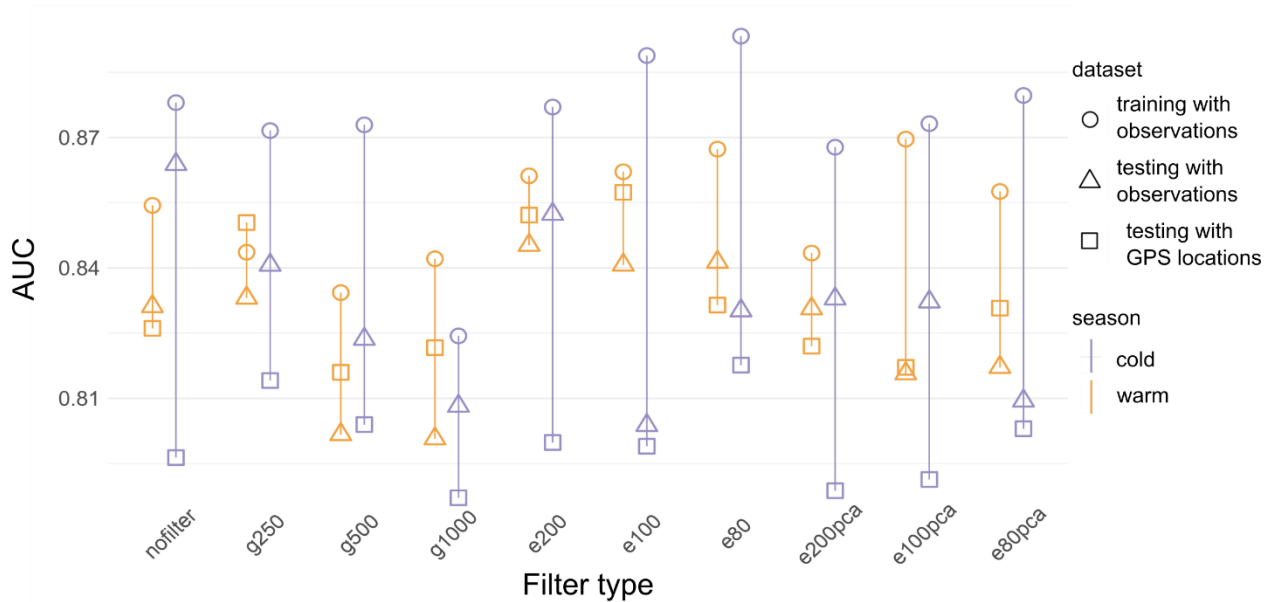


Figure 2: Evaluation of the models trained with the datasets generated by applying the different filtering methods and intensities (the prefixes g and e refer to geographical and environmental filters, respectively, see Material and Methods for details) to the casual observations of immature bearded vultures during the warm (May–October, in orange) and cold (November–February, in violet) season, respectively, provided as the area under the receiver operating characteristics curve AUC. Circles represent the AUC values for the training datasets, triangles for the testing datasets held apart from the casual observations, and squares for the independent GPS locations collected in 2004–2014.

Table 2: Sample size of the bearded vulture GPS locations used in the analysis. Each dataset was randomly sub-sampled from the total amount of GPS fixes collected within the study area, retaining two locations per month after excluding the first eight weeks of recordings due to proximity to release site during the first weeks following fledging. Locations were sampled separately per age class, season, and time periods (in parenthesis the number of tagged birds from which the sample has been extracted).

Age class	Season	2004–2014	2015–2019
Immature	Warm	237 (N=20)	363 (N=25)
Adult		NA	178 (N=10)
Immature	Cold	203 (N=18)	281 (N=23)
Adult		NA	154 (N=10)

Models' predictions

All four models performed well with regard to both training (≥ 0.861) and testing (≥ 0.830) AUC, underlining their high accuracy in predicting the occurrence probability of bearded vultures independently of age class or season. Yet, compared to the adult birds, the occurrence of immature bearded vultures was slightly less well predicted. When considering the AUC for the GPS datasets collected during 2015–2019, models for the warm season performed better than the models for the cold season (Table 3). The ten breeding sites that were known until 2014 all fall inside the projected potential conflict area. After 2014, twelve new breeding sites were settled as a result of range expansion and all but one were correctly predicted. According to the potential conflict map, about 40% of the Swiss Alps (10,244 km² ranging from 224 to 4420 m a.s.l.) offers suitable habitat for the bearded vulture. In general, the area was larger for immature birds (7097 km² and 5362 km² in the warm and cold season, respectively) than for the adults (6060 km² and 5421 km² in the warm and cold season, respectively).

Ecological requirements

The environmental variables that contributed most (permutation importance) to explaining bearded vulture's habitat in the warm season were the probability of ibex occurrence, gneiss frequency and the frequency of geological substrates other than granite, limestone, and gneiss (42.7%, 20.3%, and 9.3% for immature birds, and 34.2%, 15.9%, and 9.5% for adult birds, respectively, Fig. 5a). Vulture occurrence increased with an increasing availability of ibex, whereas gneiss and other rock types were avoided in favor of limestone substrate (Fig. S1-2). During the cold season, the most

important environmental variables driving the habitat selection differed for the two age classes. Whereas immature bearded vultures selected areas with a high probability of Chamois occurrence (22%), an intermediate solar radiation (12%) and a low forest frequency (8.5%) (Fig. 5b), low levels of precipitation (21.5%), intense solar radiation (16.6%), and a high probability of ibex presence (13.4%) were most important for adult birds (Fig. S4).

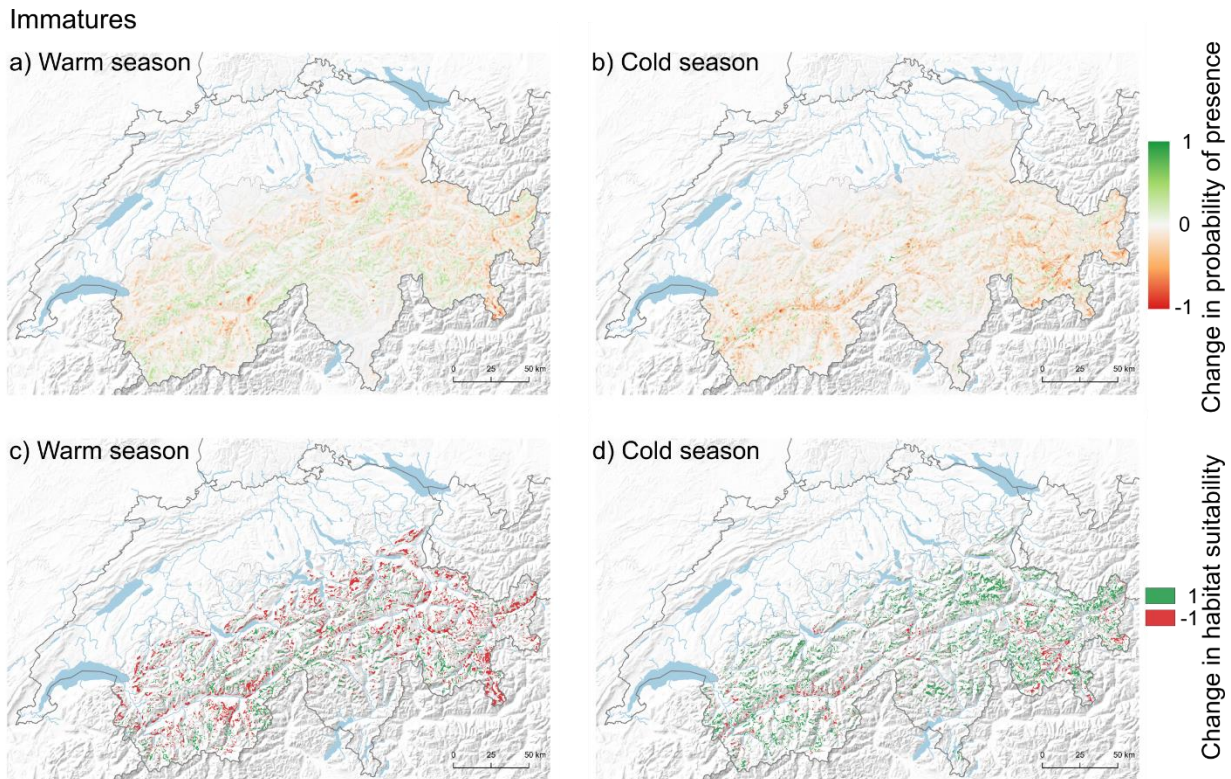


Figure 3: Error maps showing the difference between the model results based on the filtered and the unfiltered datasets, with positive and negative values showing where filtering increased or decreased the predicted occurrence probability of immature bearded vultures, respectively, during the warm (a, May–October) and cold (b, November–April) season. Increases in probability, due to the use of the filter, are displayed as a green gradient, while decreases are displayed as a red gradient. The last two maps show the changes occurring in the derived presence/absence maps (see Material and Methods for more detail) for the warm (c) and cold (d) season. The presence/absence map produced with the unfiltered dataset was subtracted from the one produced with the filtered dataset. Changes from suitable to unsuitable are colored in red, while changes from unsuitable to suitable are colored in green.

Wind turbines

Within the study area 15 big to medium size wind turbines (hub height 46–119 m, blade diameter 40–102 m) are present, 10 thereof arranged in three wind farms consisting of 2–4 wind turbines each, and five standing singly (Bundesamt für Energie BFE, 2014). Of those, two single wind turbines and two wind farms are located within the area predicted suitable for the bearded vulture (i.e. potential conflict map with values greater than zero). Furthermore, from the 106 wind turbines planned but not yet erected in the Swiss Alps, 35 are located inside suitable bearded vulture habitat according to our model predictions.

Table 3: Model performance measured as the area under the receiving operating characteristic (ROC) curve (AUC) (Fielding & Bell, 1997) of the four models predicting habitat suitability for immature and adult bearded vultures in the warm and cold season, respectively. The AUC is provided for the training (Training AUC) and testing (Testing AUC) datasets, and for the testing dataset sampled from the GPS locations (Testing AUC GPS) collected during 2015–2019 (in parenthesis the number of observations included in each dataset).

Age class	Season	Training AUC	Testing AUC	Testing AUC GPS
Immature	Warm	0.861 (N=417)	0.845 (N=179)	0.809 (N=363)
Adult		0.909 (N=474)	0.879 (N=203)	0.818 (N=178)
Immature	Cold	0.893 (N=276)	0.830 (N=118)	0.759 (N=281)
Adult		0.899 (N=265)	0.893 (N=113)	0.714 (N=154)

DISCUSSION

We estimated the potential distribution of the bearded vulture across the Swiss Alpine range, accounting for differences between the prospecting and settling phases, and between the warm and cold seasons. This delivered new insights into the ecological requirements of the species, which allowed delineating areas with a high – contemporary and future – probability of species occurrence, i.e. where conflicts with wind energy development will likely occur. Given that our models were mainly based on simple casual observations, which are today readily available from existing data-collection platforms where private naturalists report their findings, the approach used here would be easily transferable to other contexts. This study also establishes, however, that despite their great value for conservation purposes, casually collected data should not be used without cautiously checking for inherent issues and biases that may otherwise lead to flawed predictions.

Filtering

Reducing or eliminating the spatial biases affecting observational data should be a prerequisite to any species distribution models (SDM). This is because sampling bias has been proven to artificially inflate evaluation metrics (Hijmans, 2012; Veloz, 2009). Although the effect of the filtering techniques varies with their intensity (Castellanos et al., 2019; Varela et al., 2014), data are often filtered using a predefined filter size, without exploring the effect of its gradient intensity on the specific dataset (Galante et al., 2018; Chunlong Liu et al., 2019; Rose et al., 2020; Santangeli et al., 2019). In our study, we tested different procedures and intensities and introduced a new alternative option by restrictively applying the environmental filter to the two variables that define site-accessibility (road and cableways/ski-lift network). This turned out to provide the best performing filter type in terms of model performance on the independent dataset. The dataset sampled from the GPS locations collected during 2015–2019 allowed us to further estimate the model’s ability to predict a future population expansion into areas that had apparently not been occupied previously. This is crucial for the bearded vulture since its Alpine population has not yet reached carrying capacity while the objective of the whole modelling exercise was to delineate sensitive areas from the viewpoint of wind industry development.

Ecological requirements

The ecological requirements of bearded vultures vary according to both age class and season. While the requirements of immature and adult birds were quite similar during the warm season (comparable permutation importance of the three top ranked environmental variables and similar response curves), clear differences emerged between seasons and between age classes within the cold season. Food availability and geological substrate mainly determined the distribution during the warm season. Whereas the species avoids areas characterized by gneiss and rocks of “other substrate types” (i.e. other than limestone, granite and gneiss) (Fig. 5 and S1-2), areas dominated by limestone were positively selected (Fig. S1-2). This is additionally supported by the high percent contribution of this variable (Fig. S5) which ranks second for adults and third for immatures, corroborating former findings (Hirzel et al., 2004). Indeed, limestone regions are characterized by fine-grained screes that offer the best conditions for bone-breaking, a key feeding strategy of bearded vultures (Hirzel et al., 2004).

During the cold season food availability was again among the most important environmental variables, but with a noticeable difference between immature and adult birds. For

immature bearded vultures, chamois presence was most important, while for adult birds ibex presence was ranked as the third most important variable. This shift of food availability from the first to the third rank suggests that the habitat selection of adults in the cold season may be driven more by climatic conditions than food supply. Average precipitation and solar radiation were indeed the environmental variables with the highest permutation importance, with a negative and positive response type, respectively. Conceivably, this may be because adult birds could be constrained by the availability of favorable winter conditions in nest surroundings. In effect, the breeding season in this species starts in the middle of the winter, with egg laying between December and February, i.e. when environmental conditions are particularly harsh, with a probable preference for warm and dry breeding cliffs.

The few discrepancies between this study and that by Hirzel et al. (2004) are, first, that chamois did not play a key role in the latter, contrary to ibex. This is maybe due to the fact that Hirzel's et al. investigations were carried out in the southwestern Swiss Alps where chamois is uniformly widespread, while there is high spatial variability in chamois occurrence when considering the entire Swiss Alpine range. Second, limestone frequency was the most important variable for both immature and adult birds in that previous model. If we also evidenced a key role of limestone, this condition was less important than food availability during the warm season or climatic conditions during the cold season. This discrepancy could again be explained by both a greater extension of our study area and because we distinguished between seasonal requirements. Alternatively, our observational dataset was more recent, which may less well reflect species' key ecological preferences given that bearded vultures have meanwhile started to colonize less optimal areas due to population growth and hence expansion. Sheep and goat densities, as an alternative food source, were discarded during the variable selection process. This was not a real surprise because domestic ungulates provide exploitable carcasses almost exclusively during the warm season: in winter they are in stables.

Distribution maps and wind energy development

The differences in habitat selection between the two age classes and seasons translated into distribution maps that somehow spatially differ in terms of habitat suitability. Specifically, if highly suitable areas fairly widely overlap between seasons and age classes (Fig. 4a–d), the synthetic map obtained by taking the maximum pixel value from each of the four maps (Fig. 4e) embraces the whole complexity of situations and provides basic information for conservation management.

Furthermore, the potential conflict map (Fig. 4f) ranks suitable areas into four classes and provides a readily useable tool for landscape planning and environmental impact assessments. This map indeed can be considered as a spatial-explicit estimation of the “sensitivity” of a given area from the perspective of bearded vulture conservation. In effect, it highlights sensitive areas that may easily be overlooked otherwise, in particular if bearded vultures are not yet regularly spotted there, i.e. it informs whether planning a wind turbine in a given area is a sensible option from the viewpoint of wildlife protection. However, this does not mean that outside these sensitive areas (i.e. in areas tagged with zero probability of risk) collisions with wind turbines can entirely be excluded, and that impact assessments would not be required. It simply means that the likelihood of a wind turbine-project to be opposed because of conflicts with bearded vulture conservation would be much lower there than in the other, more sensitive zones (1–4).

The Swiss Wind energy concept, which delineates areas suitable for wind energy production, had proposed exclusion zones within a 5 km radius around all bearded vulture nesting sites known at the time (N = 10; 2014), aiming to reduce a priori the risk of collisions of bearded vultures with turbine blades in future installations (Bundesamt für Raumentwicklung ARE, 2017). Yet, the rapid expansion of the population has led to 22 breeding pairs in 2019 in Switzerland. Although not yet considered by the wind energy concept, all these novel sites are correctly predicted by our spatial models (only one nest location is situated 50 m beyond an area predicted as suitable for the species). It should be noticed that we obtained a correct prediction of the novel nesting sites even if our models have been trained using only observations spanning from 2004 to 2014 (the last being the reference year in the wind concept). This demonstrates the power of predictive models to project and anticipate conflict zones, paving the way for more accurate estimates of critical sites for wind energy development. Further support for such models is provided by the fact that the breeding site of nine “pairs” have changed over time, with relocation distances ranging from 280 to 3632 m (median 1081 m), with all but one of these new sites (40 eyries for 22 territories) falling within the areas delineated by the potential conflict map. For all these reasons, a simple delineation of buffer areas around known bearded vulture nest sites is not a meaningful way to guarantee the protection of a reestablishing breeding population.

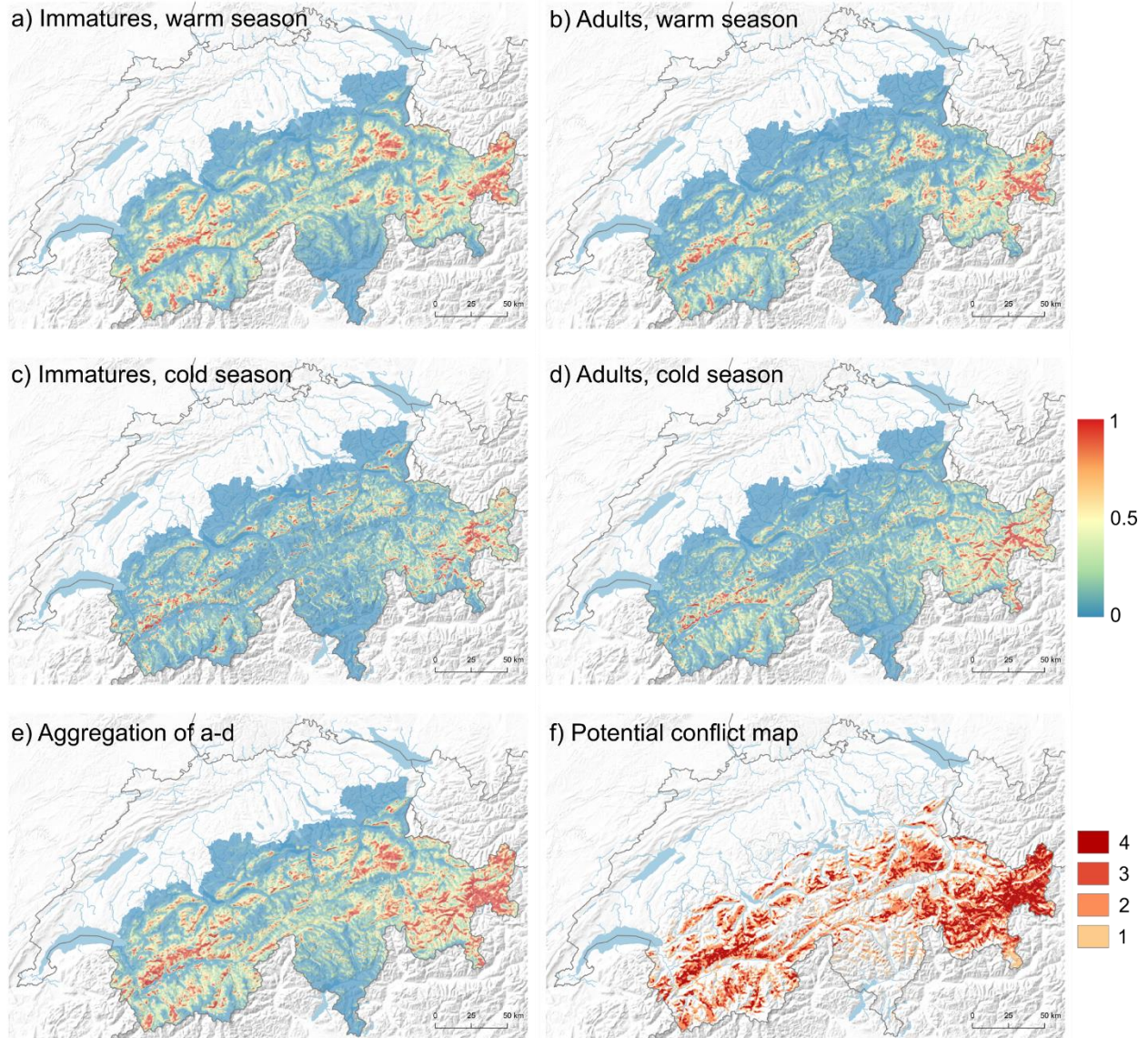


Figure 4: Predicted probability (gradient from blue, zero probability to red, high probability) of bearded vulture occurrence, projected to the entire study area for immatures during the warm (a) and cold (c) season, and for adults during the warm (b) and cold (d) season, respectively. The map e results from the aggregation of the four raster maps a-d, taking the maximum value out of the four overlapping pixels and f is the potential conflict map (see Material and Methods for more detail); the red scale in f indicates the number of models out of four that predict that area as suitable. The map f can thus be considered as a spatial-explicit estimation of the “sensitivity” of a given area from the perspective of bearded vulture conservation.

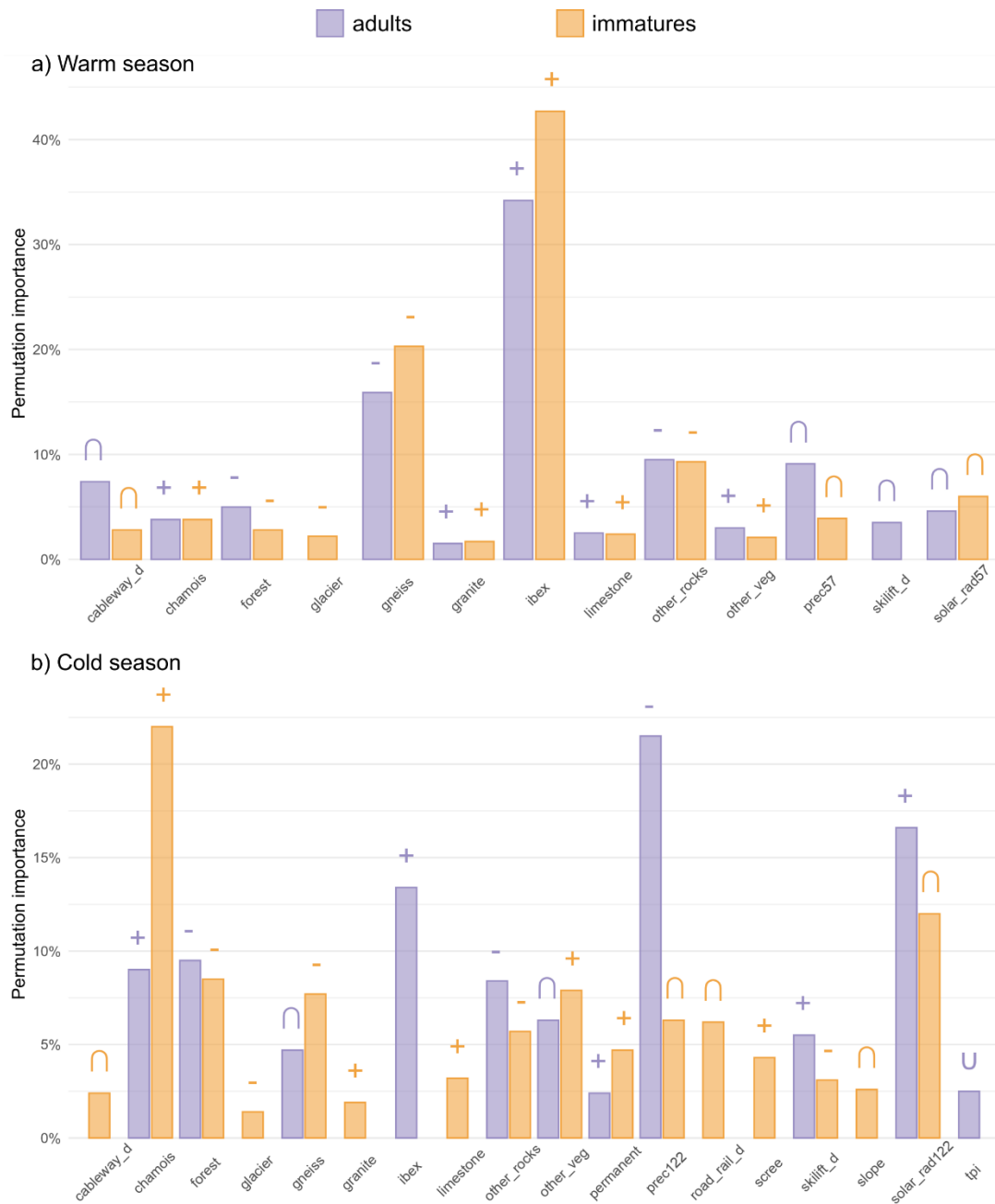


Figure 5: Variable importance, given as permutation importance, of the environmental variables retained in the final models predicting species occurrence in the a) warm and b) cold season of adult (violet) and immature (orange) bearded vultures. The symbol above the bar indicates the response type for the univariate model trained using every variable separately (for the response curves see Figures S1-4), with +, -, ∩ or ∪ indicating a positive, negative, unimodal or bimodal response, respectively. The relative (normalized %) permutation importance is computed by randomly shuffling the values of each environmental predictor at a time for both training and background locations, evaluating the model with the shuffled data, and measuring the drop in training AUC. See Table 1 for the variable abbreviations and descriptions. The variable importance measured as Maxent percent contribution [sensu Phillips, (2017)] is provided in Fig. S5.

A recent Europe-wide study ranked Switzerland among the countries with the lowest per capita wind energy production (Iten & Nipkow, 2019), but the Swiss government has planned to boost, i.e. subsidize this industry with the objective to reach a yearly energy output of 4,3 TWh by 2050 (Bundesamt für Raumentwicklung ARE, 2017). Areas with constant and sufficient wind speed in terms of economical profitability are rare, and mostly located in the Jura mountains and in the ridges and passes of the Alps – the latter being within the potential range of the bearded vulture.

Numerous new wind turbine projects are thus expected in the coming years, with already 106 new planned wind turbines within the Swiss Alps, and several additional projects being currently evaluated. Many of these new plants ($N = 71$) are outside of our potential conflict map but sometimes not far from the areas predicted as suitable for bearded vulture (minimum distance 7 m, maximum distance 17,557 m, median value 649 m). This calls for much better, i.e. evidence-based land-planning strategies. In this context, the present model represents a real asset.

Potentials and limitations

Our potential conflict map delineates areas that are suitable for the species, incorporating requirements of all life stages throughout the year cycle. Accounting for age-related habitat selection patterns is crucial for a species like the bearded vulture that explores vast areas during its teenage and becomes more and more territorial when becoming sexually mature (Margalida et al., 2016). Season-related requirements are not less important given that winters in mountain environments can be harsh, limiting the availability of suitable nesting sites and food resources. The potential conflict map delineates suitable areas in a conservative way, meaning that even if an area is predicted suitable for the species, it may not necessarily represent a zone of conflict with wind energy development. This occurs for instance when bearded vultures overfly some areas only during commuting, i.e. at altitudes far above the air space potentially swept by the rotor blades of wind turbines. On the other hand, a conservative approach avoids overlooking species-relevant areas, thus providing a broad-scale framework that will serve as a base for refined risk assessments of potential collision risks locally.

The single-species approach used here is of course inherently limited. There are other vulnerable species that should also be considered when estimating the collision risks generated by wind turbines, particularly soaring raptors with similar life-history traits and flight behaviour. The golden eagle (*Aquila chrysaetos*), for example, is another large resident raptor present in the Swiss

Alps that is potentially at risk given its large home range size and extensive daily movements. In this respect, combined spatial analyses conducted on the potentially most impacted mountain species would deliver a comprehensive spatial overview of the foreseen conflict areas to better inform management.

CONCLUSION

Here we demonstrate that, when dealing with reintroduced or spreading species, the use of buffer areas around extant nesting locations for precluding any potential threats – such as the collision risks generated by wind turbines – does not guarantee a protection of all key areas relevant for a breeding population. Our approach allows not only identifying key species' ecological requirements but also predicting its future range expansions into previously unused areas. The resulting habitat suitability maps and the potential conflict map represent a set of predictive spatial projections of future conflicts with wind energy production. The potential conflict map accounts not only for the breeding fraction of the population but also for dispersing and roaming immature individuals, thus considering the whole life cycle for enhancing conservation. We stress that our maps are intended for general guidance purposes but can in no case replace the environmental impact assessments required for planning new wind turbines. We nevertheless hope our maps will help wind energy developers as well as planners of aerial infrastructures such as electric power lines, cable cars and ski-lifts, at an early stage of the planning process, to identify areas where bearded vulture would encounter a high risk of collision, which might compromise a project from the onset.

ACKNOWLEDGMENTS

We thank the Stiftung Pro Bartgeier, the International Bearded vulture Monitoring, and the Swiss Ornithological Institute for providing the bearded vulture observations and the Stiftung Landschaftschutz Schweiz for the data regarding the wind turbines' placement. We also thank the Stiftung Pro Bartgeier, the Nationalpark Hohe Tauern, the Vulture Conservation Foundation, Vautours en Baronnies, LIFE GypConnect, LIFE GypHelp, ASTERS, LPO Grands Causses, Parc Naturel Régional du Vercors and the Parco Naturale Alpi Marittime for sharing the GPS data. This research was supported by the Swiss Federal Office for Energy, the Swiss Federal Office for the Environment and the following foundations: Parrotia Stiftung, Margarethe und Rudolf Gsell-

Stiftung, Alfons und Mathilde Suter-Caduff Stiftung, WWF Switzerland, Beat und Dieter Jutzler Stiftung, University of Bern Forschungsstiftung, Stiftung Dreiklang für ökologische Forschung und Bildung, Sophie und Karl Binding Stiftung, Stiftung Temperatio, Ernst Göhner Stiftung, Steffen Gysel Stiftung für Natur und Vogelschutz and Samy Harshallanos.

REFERENCES

- Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, **38**(5), 541–545. <https://doi.org/10.1111/ecog.01132>
- Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2019). *spThin: Function for Spatial Thinning of Species Occurrence Records for Use in Ecological Models*. R package version 0.1.0.1. <https://CRAN.R-project.org/package=spThin>.
- Antor, R. J., Margalida, A., Frey, H., Heredia, R., Lorente, L., & Sesé, J. A. (2007). First Breeding Age in Captive and Wild Bearded Vultures *Gypaetus barbatus*. *Acta Ornithologica*, **42**(1), 114–118. <https://doi.org/10.3161/068.042.0106>
- Aresu, M., Rotta, A., Fozzi, A., Campus, A., Muzzeddu, M., Secci, D., Fozzi, I., De Rosa, D., & Berlinguer, F. (2020). Assessing the effects of different management scenarios on the conservation of small island vulture populations. *Bird Conservation International*, **31**(1), 1–18. <https://doi.org/10.1017/S0959270920000040>
- Arnett, E. B., Baerwald, E. F., Mathews, F., Rodrigues, L., Rodríguez-Durán, A., Rydell, J., Villegas-Patracá, R., & Voigt, C. C. (2015). Impacts of wind energy development on bats: A global perspective. In *Bats in the Anthropocene: Conservation of Bats in a Changing World* (pp. 295–323). Springer International Publishing. https://doi.org/10.1007/978-3-319-25220-9_11
- Barrios, L., & Rodríguez, A. (2004). Behavioural and environmental correlates of soaring-bird mortality at on-shore wind turbines. *Journal of Applied Ecology*, **41**(1), 72–81. <https://doi.org/10.1111/j.1365-2664.2004.00876.x>
- Beston, J. A., Diffendorfer, J. E., Loss, S. R., & Johnson, D. H. (2016). Prioritizing Avian Species for Their Risk of Population-Level Consequences from Wind Energy Development. *PLOS ONE*, **11**(3), e0150813. <https://doi.org/10.1371/journal.pone.0150813>
- Biollaz, F., Posse, B., Denis, S., Gabbud, B., Jordan, N., Mettez, S., Michellod, B., & Arlettaz, R. (2011). Première reproduction réussie de Gypaètes barbus *Gypaetus barbatus* issus de réintroduction dans les Alpes (Haute-Savoie, France): observations comportementales du couple et du jeune. *Nos Oiseaux*, **58**, 3–12.
- BirdLife International. (2015). *Gypaetus barbatus*. *The IUCN Red List of Threatened Species 2015: e.T22695174A60116752*. Downloaded on 07 March 2020.
- Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce

- sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, **275**, 73–77. <https://doi.org/10.1016/j.ecolmodel.2013.12.012>
- Botha, A. J., Andevski, J., Bowden, C. G., Gudka, M., Safford, R. J., Tavares, J., & Williams, N. P. (2017). Multi-species action plan to conserve African-Eurasian vultures. In *CMS Raptors MOU Technical Publication No. 5. CMS Technical Series No. 35. Coordinating Unit of the CMS Raptors MOU, Abu Dhabi, United Arab Emirates*.
- Braunisch, V., Coppes, J., Bächle, S., & Suchant, R. (2015). Underpinning the precautionary principle with evidence: A spatial concept for guiding wind power development in endangered species' habitats. *Journal for Nature Conservation*, **24**, 31–40. <https://doi.org/10.1016/j.jnc.2015.01.003>
- Bright, J., Langston, R., Bullman, R., Evans, R., Gardner, S., & Pearce-Higgins, J. (2008). Map of bird sensitivities to wind farms in Scotland: A tool to aid planning and conservation. *Biological Conservation*, **141**(9), 2342–2356. <https://doi.org/10.1016/j.biocon.2008.06.029>
- Buechley, E. R., & Şekercioğlu, Ç. H. (2016). The avian scavenger crisis: Looming extinctions, trophic cascades, and loss of critical ecosystem functions. *Biological Conservation*, **198**, 220–228. <https://doi.org/10.1016/j.biocon.2016.04.001>
- Bundesamt für Energie BFE. (2014). *Windenergieanlagen*. <http://data.geo.admin.ch/ch.bfe.windenergieanlagen/>
- Bundesamt für Energie BFE. (2016). *Windatlas Schweiz: Jahresmittel der modellierten Windgeschwindigkeit und Windrichtung*. Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK. www.bfe.admin.ch
- Bundesamt für Raumentwicklung ARE. (2017). *Konzept Windenergie. Basis zur Berücksichtigung der Bundesinteressen bei der Planung von Windenergieanlagen*. Bern.
- Carrete, M., Sánchez-Zapata, J. A., Benítez, J. R., Lobón, M., & Donázar, J. A. (2009). Large scale risk-assessment of wind-farms on population viability of a globally endangered long-lived raptor. *Biological Conservation*, **142**(12), 2954–2961. <https://doi.org/10.1016/J.BIOCON.2009.07.027>
- Castellanos, A. A., Huntley, J. W., Voelker, G., & Lawing, A. M. (2019). Environmental filtering improves ecological niche models across multiple scales. *Methods in Ecology and Evolution*, **10**(4), 481–492. <https://doi.org/10.1111/2041-210X.13142>
- Cuthbert, R., Taggart, M. A., Prakash, V., Saini, M., Swarup, D., Upreti, S., Mateo, R., Chakraborty, S. S., Deori, P., & Green, R. E. (2011). Effectiveness of Action in India to Reduce Exposure of Gyps Vultures to the Toxic Veterinary Drug Diclofenac. *PLoS ONE*, **6**(5), e19069. <https://doi.org/10.1371/journal.pone.0019069>
- Dahl, E. L., Bevanger, K., Nygård, T., Røskoft, E., & Stokke, B. G. (2012). Reduced breeding success in white-tailed eagles at Smøla windfarm, western Norway, is caused by mortality and displacement. *Biological Conservation*, **145**(1), 79–85. <https://doi.org/10.1016/j.biocon.2011.10.012>
- de Lucas, M., Ferrer, M., Bechard, M. J., & Muñoz, A. R. (2012). Griffon vulture mortality at

- wind farms in southern Spain: Distribution of fatalities and active mitigation measures. *Biological Conservation*, **147**(1), 184–189. <https://doi.org/10.1016/j.biocon.2011.12.029>
- Demerdzhiev, D., Hristov, H., Dobrev, D., Angelov, I., & Kurtev, M. (2014). Long-term population status, breeding parameters and limiting factors of the Griffon vulture (*Gyps fulvus* Hablizl, 1783) population in the Eastern Rhodopes, Bulgaria. *Acta Zoologica Bulgarica*, **66**(3), 373–384.
- Drewitt, A. L., & Langston, R. H. W. (2006). Assessing the impacts of wind farms on birds. *Ibis*, **148**(s1), 29–42. <https://doi.org/10.1111/j.1474-919X.2006.00516.x>
- Eliotout, B., Lécuyer, P., & Duriez, O. (2007). Premiers résultats sur la biologie de reproduction du Vautour Moine *Aegypius monachus* en France. *Alauda*, **75**(3), 253–264.
- Ferrer, M., de Lucas, M., Janss, G. F. E., Casado, E., Muñoz, A. R., Bechard, M. J., & Calabuig, C. P. (2012). Weak relationship between risk assessment studies and recorded mortality in wind farms. *Journal of Applied Ecology*, **49**(1), 38–46. <https://doi.org/10.1111/j.1365-2664.2011.02054.x>
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**(1), 38–49.
- Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. *PLoS ONE*, **9**(5), e97122. <https://doi.org/10.1371/journal.pone.0097122>
- Frey, H., & Walter, W. (1989). The Reintroduction of the Bearded Vulture *Gypaetus barbatus* into the Alps. *Raptors in the Modern World*, 341–344.
- Galante, P. J., Alade, B., Muscarella, R., Jansa, S. A., Goodman, S. M., & Anderson, R. P. (2018). The challenge of modeling niches and distributions for data-poor species: a comprehensive approach to model complexity. *Ecography*, **41**(5), 726–736. <https://doi.org/10.1111/ecog.02909>
- Galligan, T. H., Amano, T., Prakash, V. M., Kulkarni, M., Shringarpure, R., Prakash, N., Ranade, S., Green, R. E., & Cuthbert, R. J. (2014). Have population declines in Egyptian Vulture and Red-headed Vulture in India slowed since the 2006 ban on veterinary diclofenac? *Bird Conservation International*, **24**(3), 272–281. <https://doi.org/10.1017/S0959270913000580>
- Gonseth, Y., Wohlgemuth, T., Sansonnens, B., & Buttler, A. (2001). *Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard* (p. 48). Umwelt Materialien Nr. 137 Bundesamt für Umwelt, Wald und Landschaft.
- Hegglin, D., Wehrle, M., & Aebischer, A. (2004). Satellite tracking of Bearded Vultures - the project « BEARDED VULTURE ON THE MOVE ». *Annu. Rep. Found. Conserv. Bearded Vultures*, 52–55.
- Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology*, **93**(3), 679–688. <https://doi.org/10.1890/11-0826.1>

- Hirzel, A., Posse, B., Oggier, P.-A., Crettenand, Y., Glenz, C., & Arlettaz, R. (2004). Ecological requirements of reintroduced Bearded vultures colonizing the Alps : implications for release policy. *Journal of Applied Ecology*, **41**(6), 1103–1116. <https://doi.org/10.1111/j.0021-8901.2004.00980.x>
- Iten, T., & Nipkow, F. (2019). *Solar- und Windenergieproduktion der Schweiz im Europäischen Vergleich*. Schweizerische Energie-Stiftung.
- Janss, G. F. E., de Lucas, M., Whitfield, D. P., Lazo, A., & Ferrer, M. (2010). The precautionary principle and wind-farm planning in Andalucía. *Biological Conservation*, **143**(8), 1827–1828. <https://doi.org/10.1016/j.biocon.2010.02.016>
- Katzner, T. E., Brandes, D., Miller, T., Lanzone, M., Maisonneuve, C., Tremblay, J. A., Mulvihill, R., & Merovich, G. T. (2012). Topography drives migratory flight altitude of golden eagles: implications for on-shore wind energy development. *Journal of Applied Ecology*, **49**(5), 1178–1186. <https://doi.org/10.1111/j.1365-2664.2012.02185.x>
- Katzner, T. E., Nelson, D. M., Braham, M. A., Doyle, J. M., Fernandez, N. B., Duerr, A. E., Bloom, P. H., Fitzpatrick, M. C., Miller, T. A., Culver, R. C. E., Braswell, L., & DeWoody, J. A. (2017). Golden Eagle fatalities and the continental-scale consequences of local wind-energy generation. *Conservation Biology*, **31**(2), 406–415. <https://doi.org/10.1111/cobi.12836>
- Keller, V., Gerber, A., Schmid, H., Volet, B., & Zbinden, N. (2010). *Rote Liste Brutvögel. Gefährdete Arten der Schweiz, Stand 2010*. Bundesamt für Umwelt, Bern, und Schweizerische Vogelwarte, Sempach UmweltVollzug Nr. 1019. 53 S.
- Kriebel, D., Tickner, J., Epstein, P., Lemons, J., Levins, R., Loechler, E. L., Quinn, M., Rudel, R., Schettler, T., & Stoto, M. (2001). *Environmental Health Perspectives*, **109**(9), 871–876. <https://doi.org/https://doi.org/10.1289/ehp.01109871>
- Krüger, S., Reid, T., & Amar, A. (2014). Differential Range Use between Age Classes of Southern African Bearded Vultures *Gypaetus barbatus*. *PLoS ONE*, **9**(12), 1–18. <https://doi.org/10.1371/journal.pone.0114920>
- Lee, J., & Zhao, F. (2020). GWEC Global Wind Report. In *Global Wind Energy Council*. www.gwec.net
- Leung, D. Y. C., & Yang, Y. (2012). Wind energy development and its environmental impact: A review. *Renewable and Sustainable Energy Reviews*, **16**(1), 1031–1039. <https://doi.org/10.1016/J.RSER.2011.09.024>
- Liu, Canran, White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, **40**(4), 778–789. <https://doi.org/10.1111/jbi.12058>
- Liu, Chunlong, Comte, L., Xian, W., Chen, Y., & Olden, J. D. (2019). Current and projected future risks of freshwater fish invasions in China. *Ecography*, **42**(12), 2074–2083. <https://doi.org/10.1111/ecog.04665>
- Madders, M., & Whitfield, D. P. (2006). Upland raptors and the assessment of wind farm

- impacts. *Ibis*, **148**(s1), 43–56. <https://doi.org/10.1111/j.1474-919X.2006.00506.x>
- Margalida, A., Bertran, J., Boudet, J., & Heredia, R. (2004). Hatching asynchrony, sibling aggression and cannibalism in the Bearded Vulture *Gypaetus barbatus*. *Ibis*, **146**(3), 386–393. <https://doi.org/10.1111/j.1474-919X.2004.00261.x>
- Margalida, A., Garcia, D., Bertran, J., & Heredia, R. (2003). Breeding biology and success of the Bearded Vulture *Gypaetus barbatus* in the eastern Pyrenees. *Ibis*, **145**(2), 244–252. <https://doi.org/10.1046/j.1474-919X.2003.00148.x>
- Margalida, A., Pérez-García, J. M., Afonso, I., & Moreno-Opo, R. (2016). Spatial and temporal movements in Pyrenean bearded vultures (*Gypaetus barbatus*): Integrating movement ecology into conservation practice. *Scientific Reports*, **6**(1), 1–12. <https://doi.org/10.1038/srep35746>
- Margalida, A., & Villalba, D. (2017). The importance of the nutritive value of old bones in the diet of Bearded vultures *Gypaetus barbatus*. *Scientific Reports*, **7**(1), 1–5. <https://doi.org/10.1038/s41598-017-08812-2>
- Martin, G. R., Portugal, S. J., & Murn, C. P. (2012). Visual fields, foraging and collision vulnerability in Gyps vultures. *Ibis*, **154**(3), 626–631. <https://doi.org/10.1111/j.1474-919X.2012.01227.x>
- Mingozzi, T., & Estève, R. (1997). Analysis of a historical extirpation of the bearded vulture *Gypaetus barbatus* (L.) in the western Alps (France-Italy): Former distribution and causes of extirpation. *Biological Conservation*, **79**(2–3), 155–171. [https://doi.org/10.1016/S0006-3207\(96\)00110-3](https://doi.org/10.1016/S0006-3207(96)00110-3)
- Ogada, D., Keesing, F., & Virani, M. Z. (2012). Dropping dead: causes and consequences of vulture population declines worldwide. *Annals of the New York Academy of Sciences*, **1249**(1), 57–71. <https://doi.org/10.1111/j.1749-6632.2011.06293.x>
- Ogada, D., Shaw, P., Beyers, R. L., Buij, R., Murn, C., Thiollay, J. M., Beale, C. M., Holdo, R. M., Pomeroy, D., Baker, N., Krüger, S. C., Botha, A., Virani, M. Z., Monadjem, A., & Sinclair, A. R. E. (2016). Another Continental Vulture Crisis: Africa's Vultures Collapsing toward Extinction. *Conservation Letters*, **9**(2), 89–97. <https://doi.org/10.1111/conl.12182>
- Phillips, S. J. (2017). *A Brief Tutorial on Maxent*. http://biodiversityinformatics.amnh.org/open_source/maxent/
- Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, **40**(7), 887–893. <https://doi.org/10.1111/ecog.03049>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**(3–4), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Poessel, S. A., Brandt, J., Mendenhall, L., Braham, M. A., Lanzone, M. J., McGann, A. J., & Katzner, T. E. (2018). Flight response to spatial and temporal correlates informs risk from wind turbines to the California Condor. *The Condor*, **120**(2), 330–342.

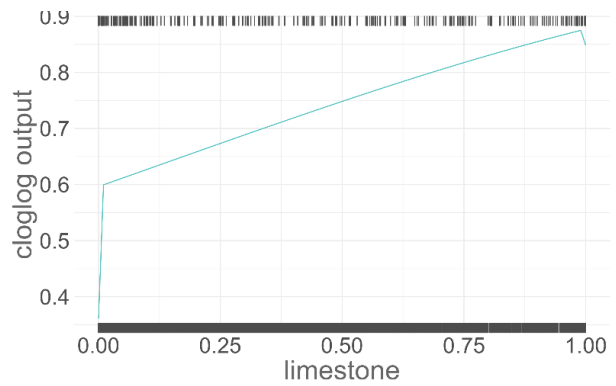
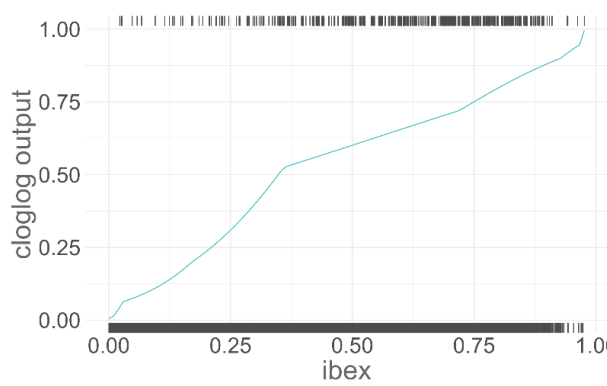
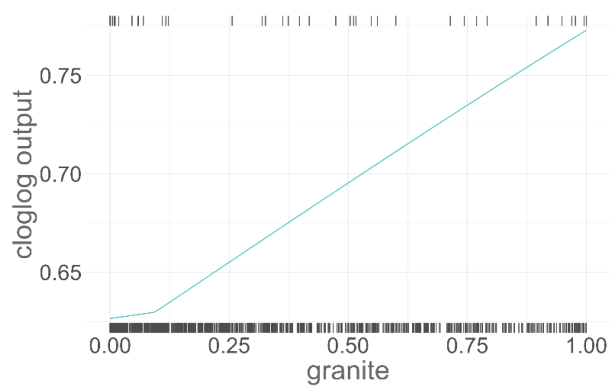
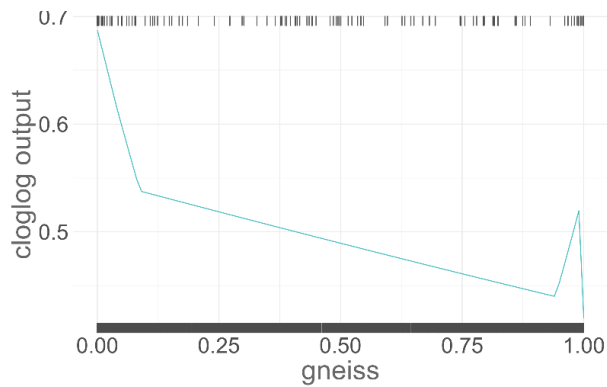
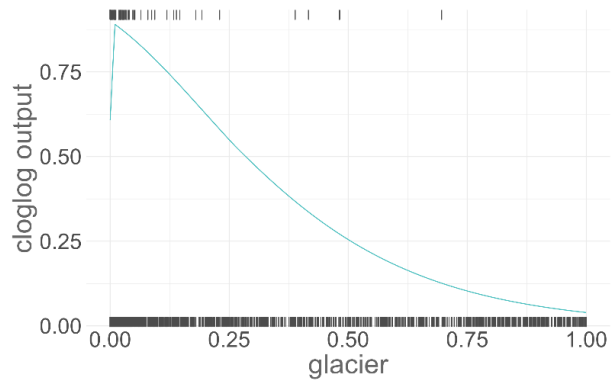
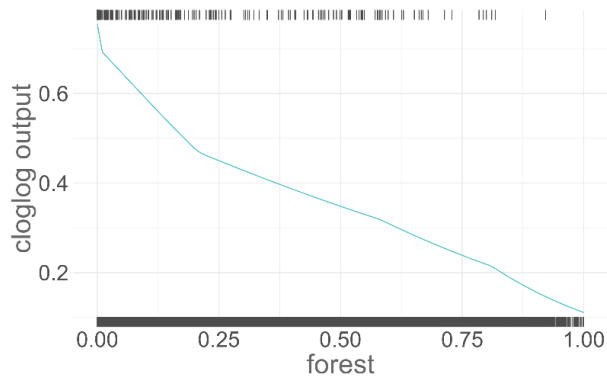
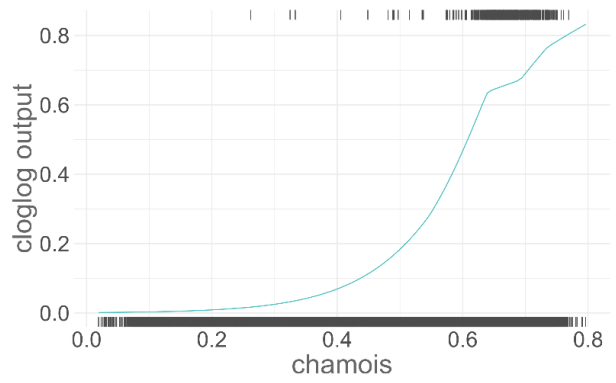
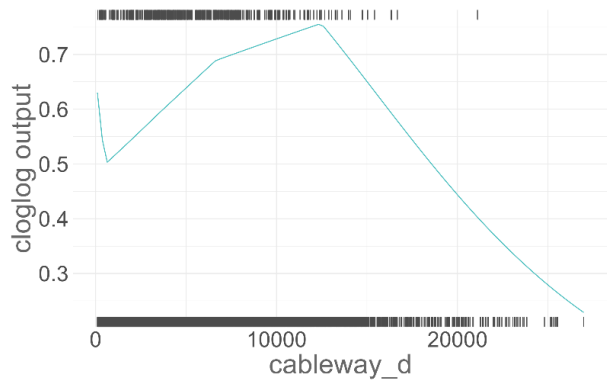
<https://doi.org/10.1650/CONDOR-17-100.1>

- Prakash, V., Bishwakarma, M. C., Chaudhary, A., Cuthbert, R., Dave, R., Kulkarni, M., Kumar, S., Paudel, K., Ranade, S., Shringarpure, R., & Green, R. E. (2012). The Population Decline of Gyps Vultures in India and Nepal Has Slowed since Veterinary Use of Diclofenac was Banned. *PLoS ONE*, **7**(11), e49118. <https://doi.org/10.1371/journal.pone.0049118>
- R Core Team. (2019). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*. <https://www.r-project.org/>
- Robin, K., Müller, J. P., Pachlatko, T., & Buchli, C. (2004). Das Projekt zur Wiederansiedlung des Bartgeiers in den Alpen ist 25-jährig: Ein Überblick. *Ornithologische Beobachter*, **101**(1), 1–18.
- Rose, J. P., Halstead, B. J., & Fisher, R. N. (2020). Integrating multiple data sources and multi-scale land-cover data to model the distribution of a declining amphibian. *Biological Conservation*, **241**, 108374. <https://doi.org/10.1016/j.biocon.2019.108374>
- RStudio Team. (2018). *RStudio: Integrated Development Environment for R*.
- Rushworth, I., & Krüger, S. (2014). Wind farms threaten southern Africa’s cliff-nesting vultures. *Ostrich*, **85**(1), 13–23. <https://doi.org/10.2989/00306525.2014.913211>
- Rydell, J., Bach, L., Dubourg-Savage, M.-J., Green, M., Rodrigues, L., & Hedenström, A. (2010). Bat Mortality at Wind Turbines in Northwestern Europe. *Acta Chiropterologica*, **12**(2), 261–274. <https://doi.org/10.3161/150811010X537846>
- Safford, R., Andevski, J., Botha, A., Bowden, C. G. R., Crockford, N., Garbett, R., Margalida, A., Ramírez, I., Shobrak, M., Tavares, J., & Williams, N. P. (2019). Vulture conservation: The case for urgent action. *Bird Conservation International*, **29**(1), 1–9. <https://doi.org/10.1017/S0959270919000042>
- Santangeli, A., Girardello, M., Buechley, E., Botha, A., Minin, E. Di, & Moilanen, A. (2019). Priority areas for conservation of Old World vultures. *Conservation Biology*, **33**(5), 1056–1065. <https://doi.org/10.1111/cobi.13282>
- Sarrazin, F., Bagnolini, C., Pinna, J. L., Danchin, E., & Clobert, J. (1994). High survival estimates in a reintroduced population of Griffon Vultures Restoration and conservation of scavengers populations View project LIFE Gypconnect View project. *The Auk*, **111**(4), 853–862. <https://doi.org/10.2307/4088817>
- Schaub, M., Zink, R., Beissmann, H., Sarrazin, F., & Arlettaz, R. (2009). When to end releases in reintroduction programmes: demographic rates and population viability analysis of bearded vultures in the Alps. *Journal of Applied Ecology*, **46**(1), 92–100. <https://doi.org/10.1111/j.1365-2664.2008.01585.x>
- Stoyanov, E., Kmetova-Biro, E., Stoyanov, G., Peshev, H., Ivanov, I., Stoev, I., Bonchev, L., Vangelova, N., Nikolova, Z., Iankov, L., Parvanov, D., & Grozdanov, A. (2018). Population boost of the griffon Vulture *Gyps fulvus* (Hablizl, 1783) (Accipitridae) in Bulgaria Based on Reintroductions. *Acta Zoologica Bulgarica*, **Suppl. 12**, 59–65.

- Stoynov, E., Peshev, H., Stoyanov, G., Ivanov, I., Parvanov, D., & Grozdanov, A. (2019). Past and Present State of the Cinereous Vulture (*Aegypius monachus*) and Feasibility Analysis for its Reintroduction in Bulgaria. *Ecologia Balkanica, Special Edition 2*, 30–43.
- Varela, S., Anderson, R. P., García-Valdés, R., & Fernández-González, F. (2014). Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography*, **37**(11), 1084–1091. <https://doi.org/10.1111/j.1600-0587.2013.00441.x>
- Veloz, S. D. (2009). Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, **36**(12), 2290–2299. <https://doi.org/10.1111/j.1365-2699.2009.02174.x>
- Venter, J. A., Martens, F. R., & Wolter, K. (2019). Conservation Buffer Sizes Derived from Movement Data of Adult Cape Vultures (*Gyps coprotheres*) in South Africa. *African Zoology*, **54**(2), 115–118. <https://doi.org/10.1080/15627020.2019.1600428>
- Vignali, S., Barras, A. G., Arlettaz, R., & Braunisch, V. (2020). *SDMtune* : An R package to tune and evaluate species distribution models. *Ecology and Evolution*, **10**(20), 11488–11506. <https://doi.org/10.1002/ece3.6786>
- Walters, J. R., Derrickson, S. R., Michael, D., Haig, S. M., Marzluff, J. M., & Wunderle, J. M. (2010). Status of the California Condor (*Gymnogyps Californianus*) and Efforts to Achieve Its Recovery. *The Auk*, **127**(4), 969–1001. <https://doi.org/10.1525/auk.2010.127.4.969>
- Watson, R. T., Kolar, P. S., Ferrer, M., Nygård, T., Johnston, N., Hunt, W. G., Smit-Robinson, H. A., Farmer, C. J., Huso, M., & Katzner, T. E. (2018). Raptor Interactions With Wind Energy: Case Studies From Around the World. *Journal of Raptor Research*, **52**(1), 1–18. <https://doi.org/10.3356/jrr-16-100.1>
- Wilson, J. D. (1984). Determining a TOPEX score. *Scottish Forestry*, **38**(4), 251–256.
- Windparkkarte Schweiz - Stiftung Landschaftsschutz Schweiz*. (n.d.). Retrieved July 26, 2020, from <https://www.sl-fp.ch/de/stiftung-landschaftsschutz-schweiz/dokumentation/windparkkarte-schweiz-66.html>

SUPPLEMENTARY MATERIAL

Immatures, warm season



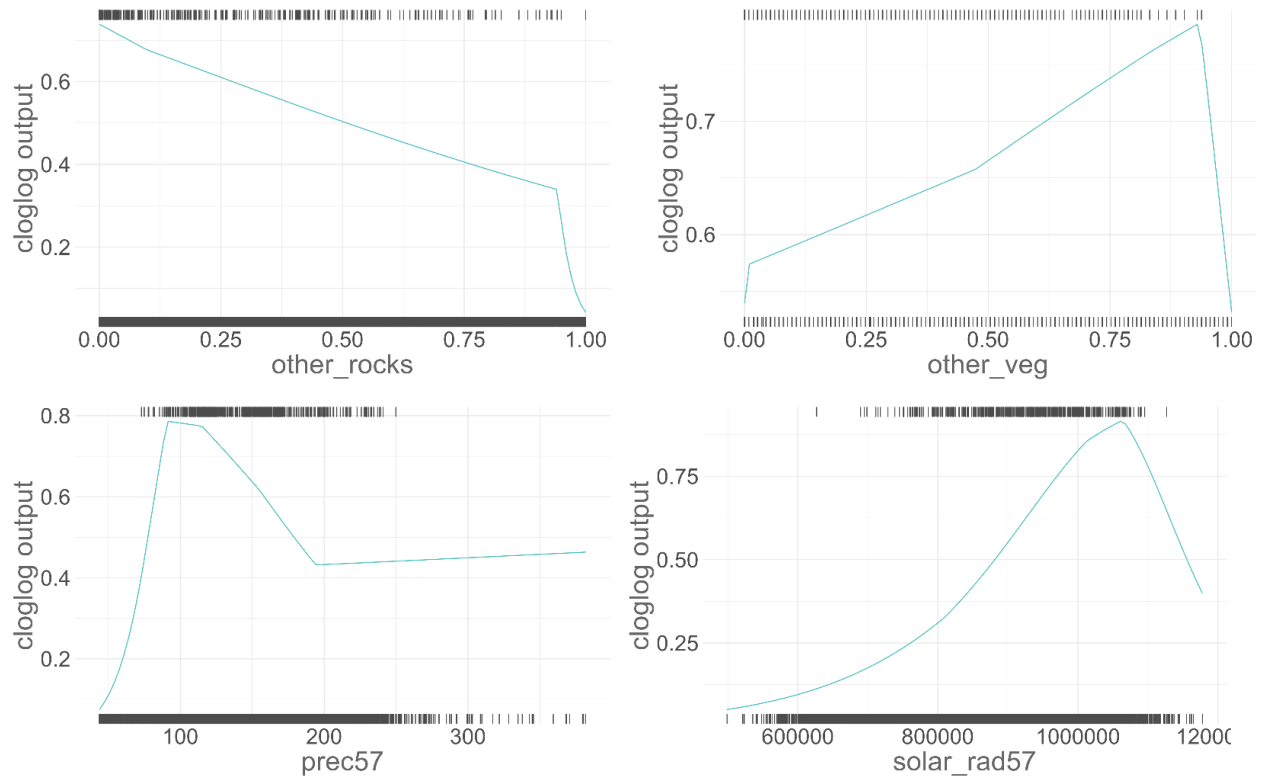
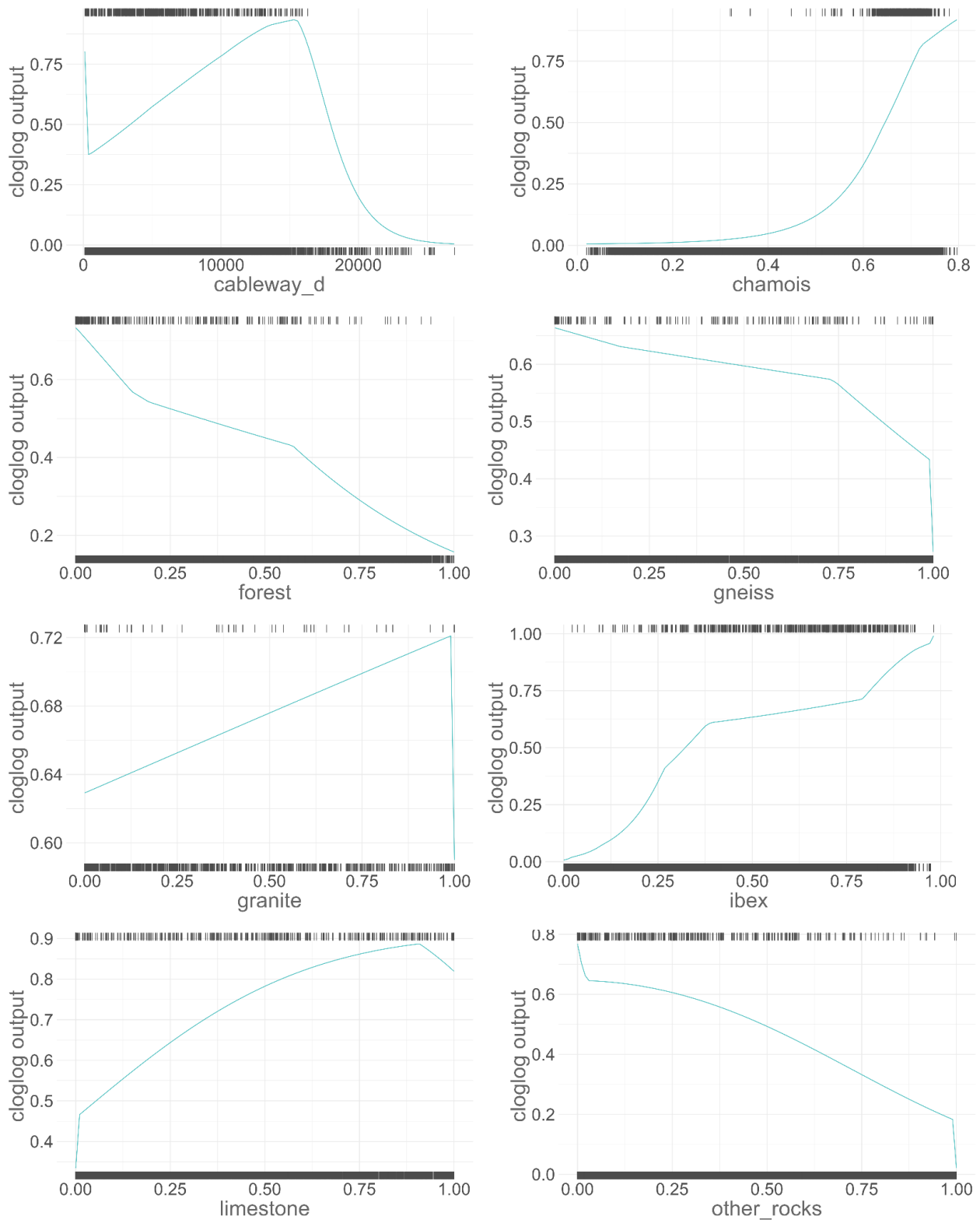


Figure S1: Response of immature bearded vultures to the environmental variables included in the final model for the warm season. The response curves show the change in cloglog prediction for a model trained using only the displayed variable (i.e. univariate model). The values of the variable at presence locations are shown on the top of the plot, the values at background locations on the bottom. See Table 1 for the variable abbreviations and descriptions.

Adults, warm season



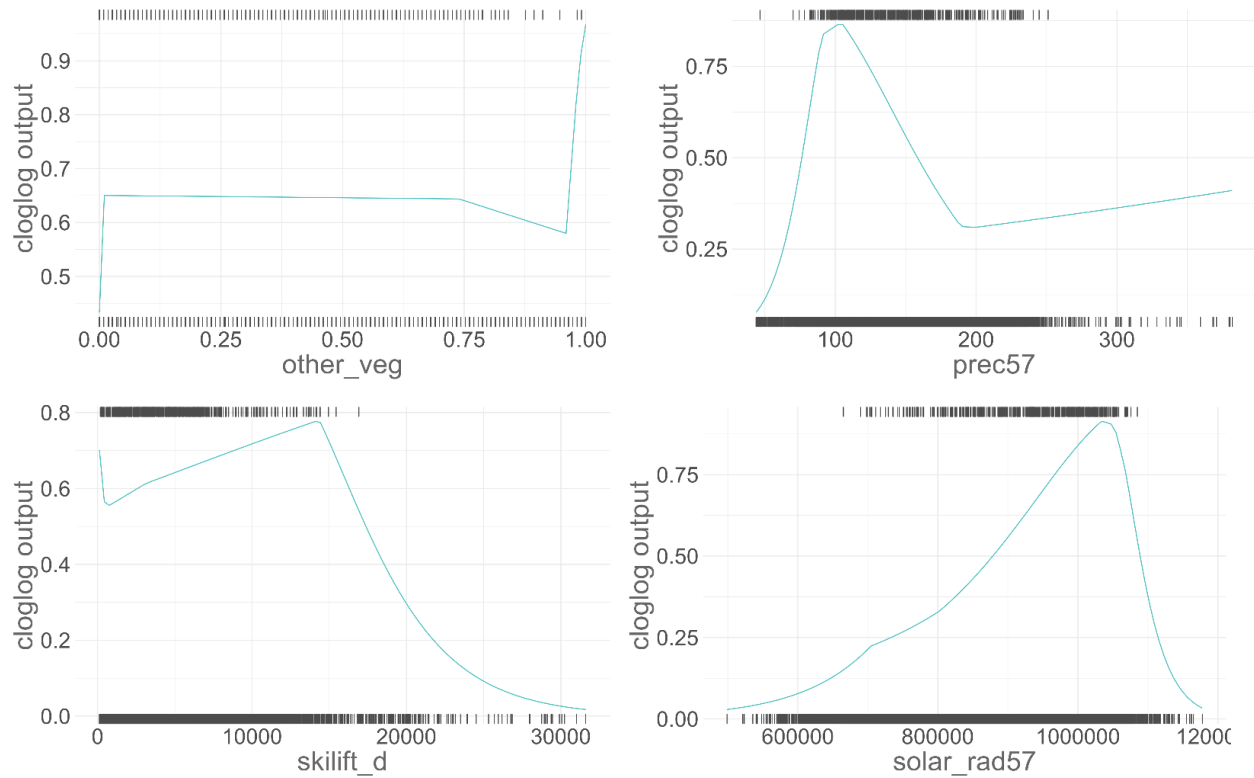
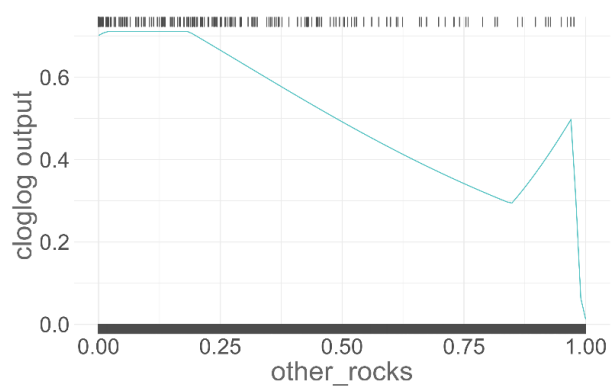
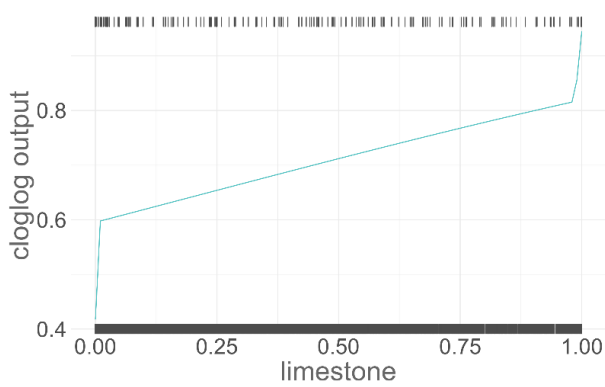
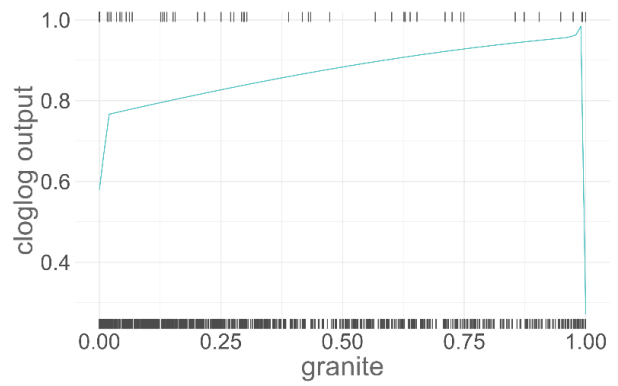
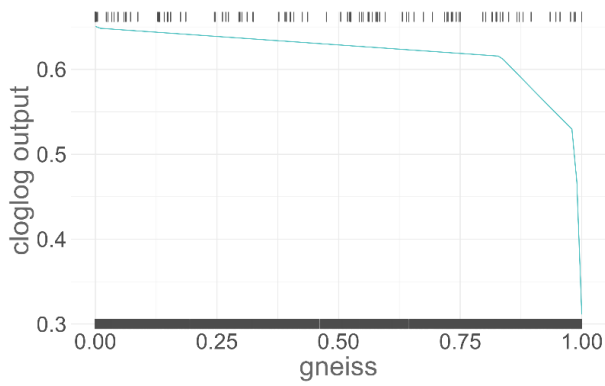
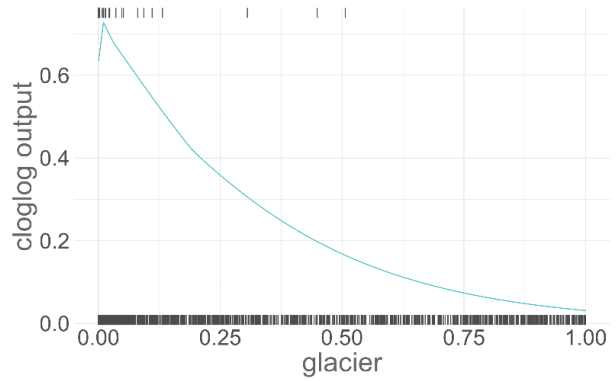
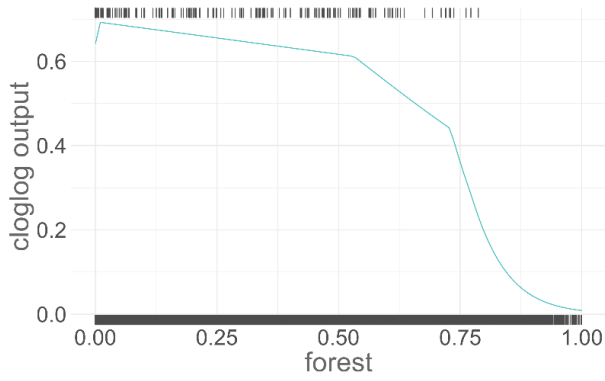
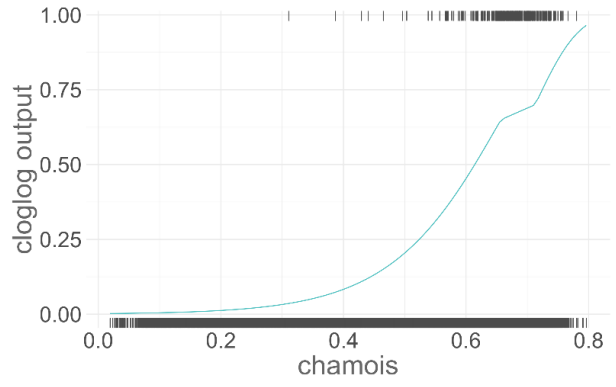
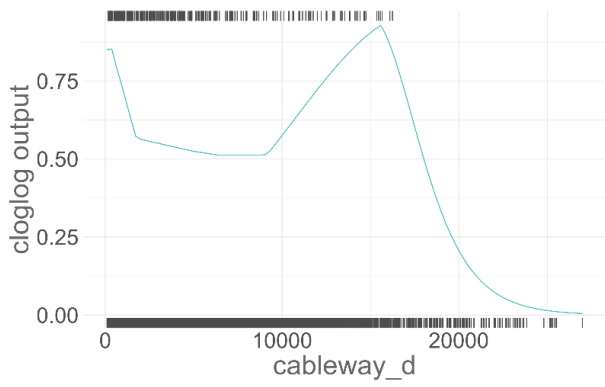


Figure S2: Response of adult bearded vultures to the environmental variables included in the final model for the warm season. The response curves show the change in cloglog prediction for a model trained using only the displayed variable (i.e. univariate model). The values of the variable at presence locations are shown on the top of the plot, the values at background locations on the bottom. See Table 1 for the variable abbreviations and descriptions.

Immatures, cold season



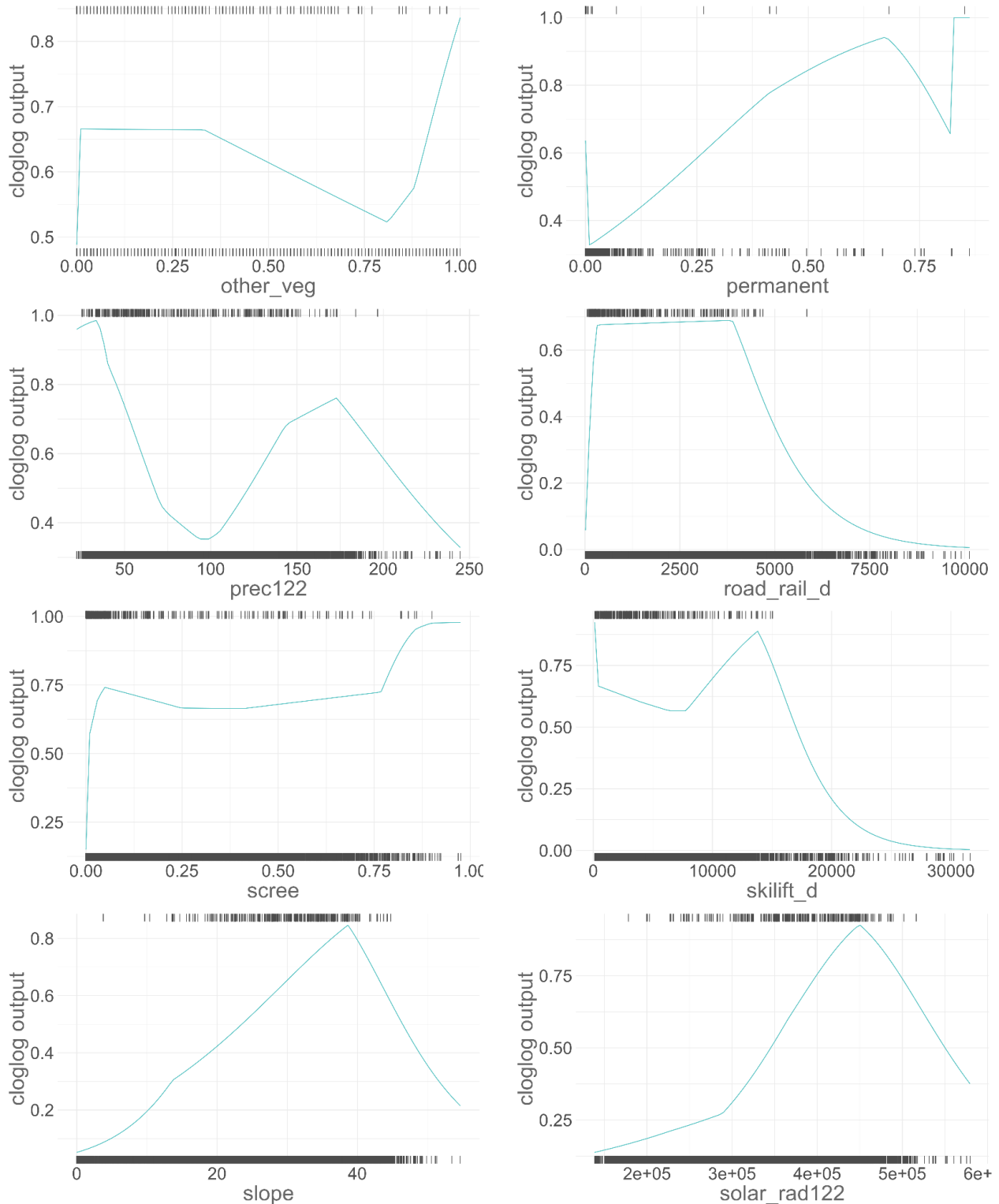
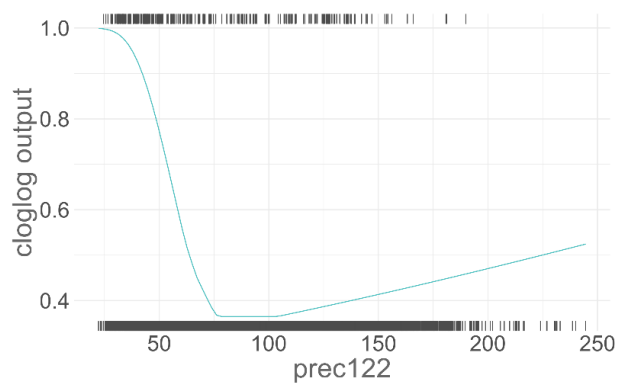
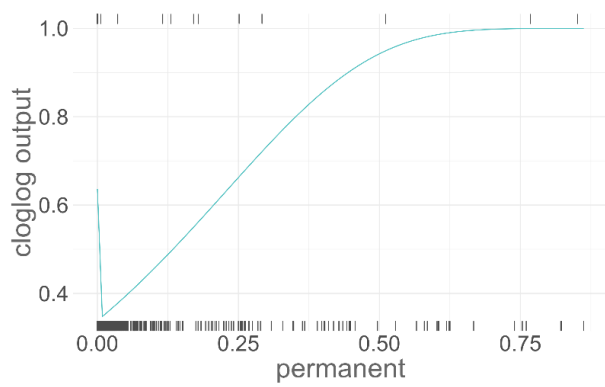
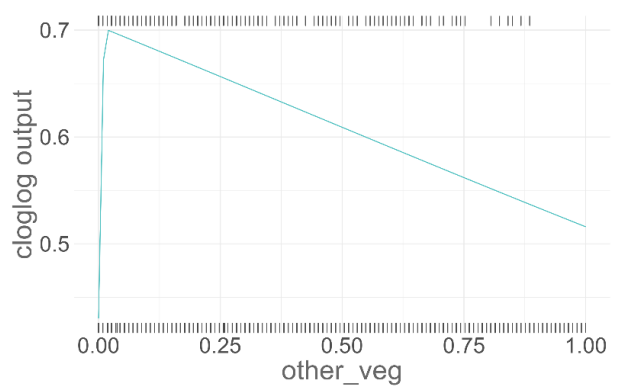
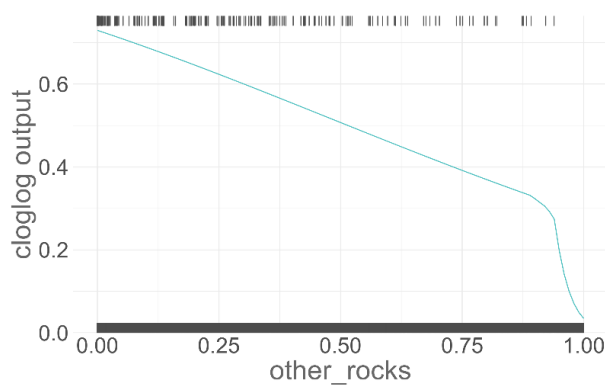
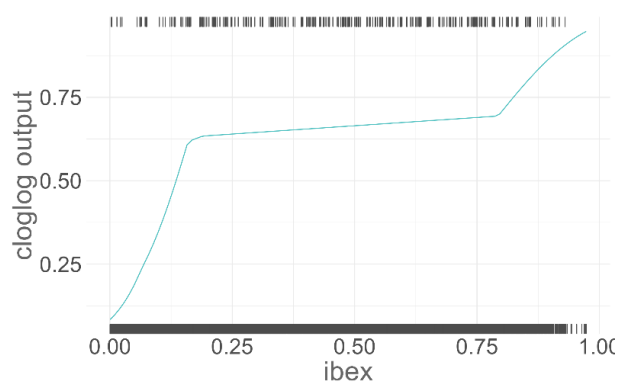
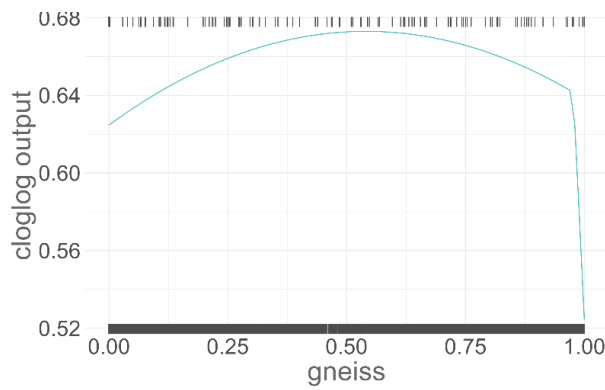
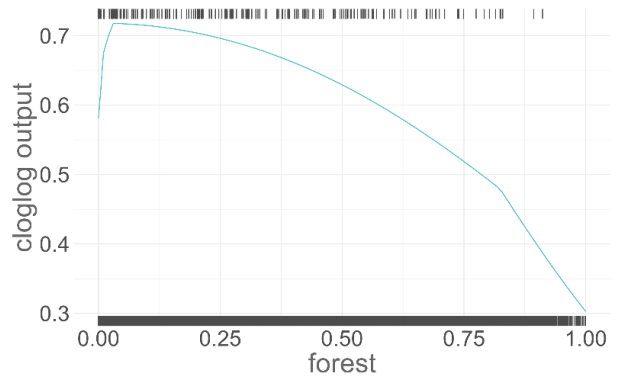
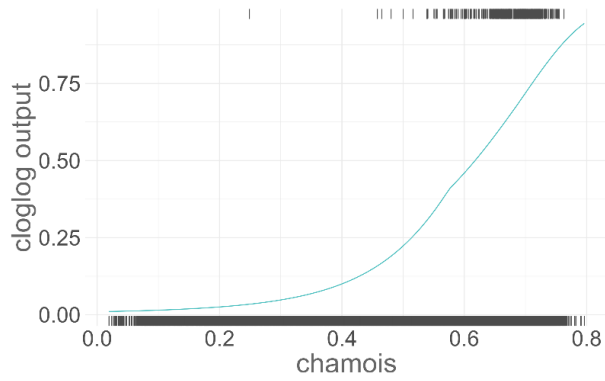


Figure S3: Response of immature bearded vultures to the environmental variables included in the final model for the cold season. The response curves show the change in cloglog prediction for a model trained using only the displayed variable (i.e. univariate model). The values of the variable at presence locations are shown on the top of the plot, the values at background locations on the bottom. See Table 1 for the variable abbreviations and descriptions.

Adults, cold season



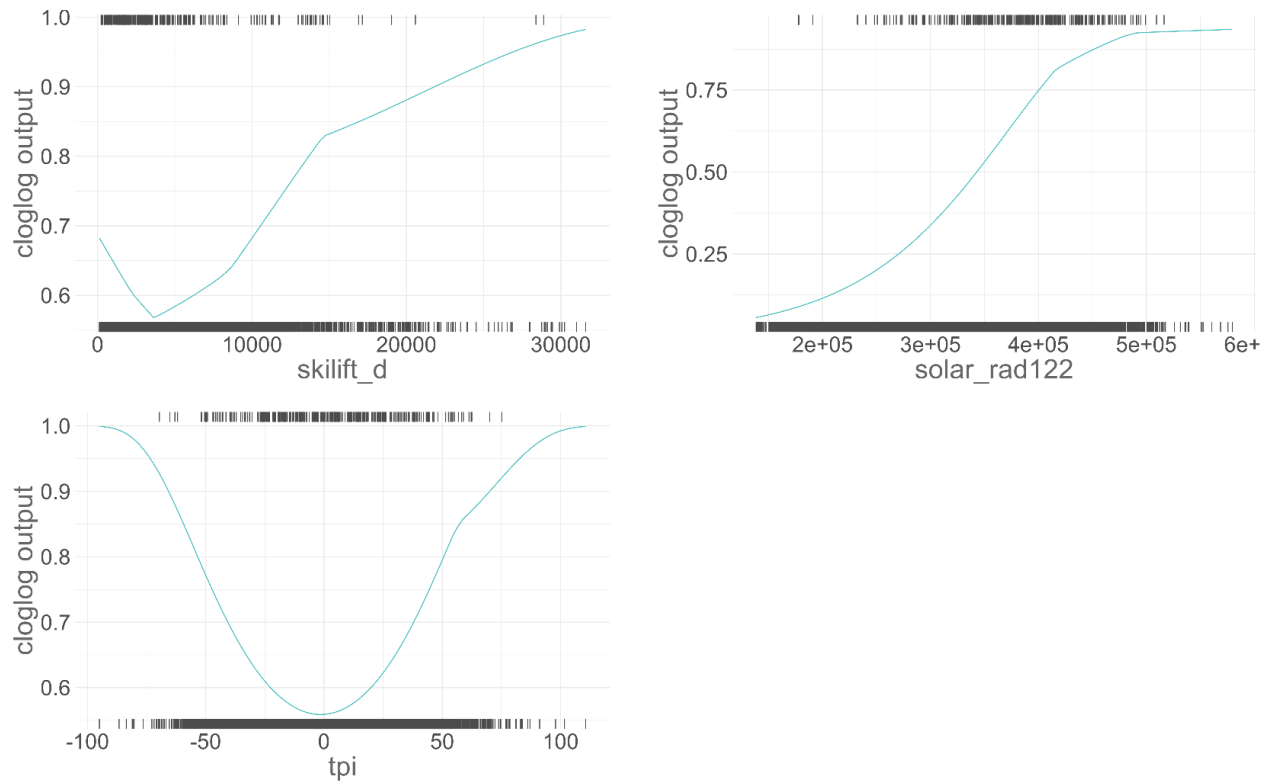


Figure S4: Response of adult bearded vultures to the environmental variables included in the final model for the cold season. The response curves show the change in cloglog prediction for a model trained using only the displayed variable (i.e. univariate model). The values of the variable at presence locations are shown on the top of the plot, the values at background locations on the bottom. See Table 1 for the variable abbreviations and descriptions.

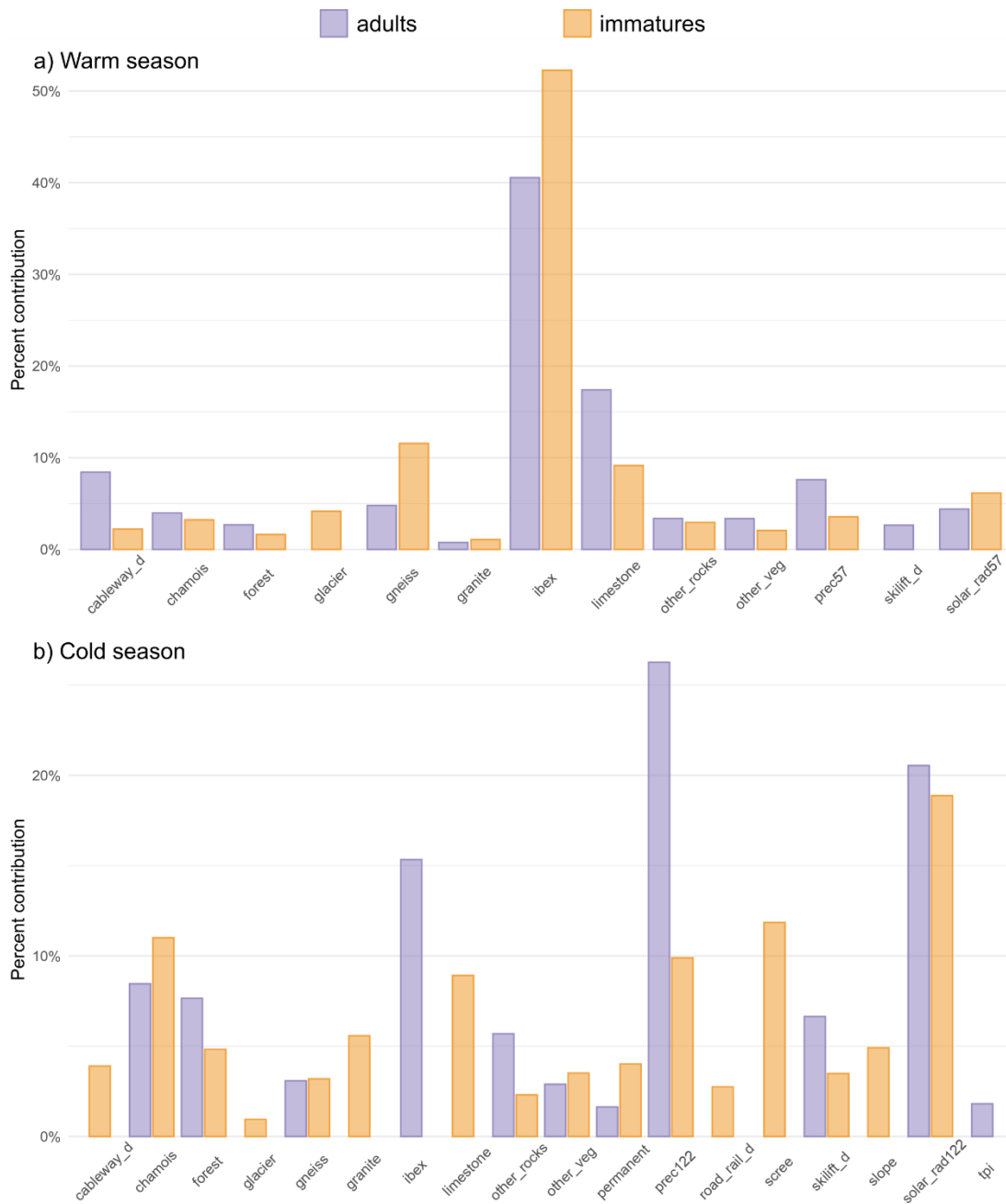


Figure S5: Variable importance, given as Maxent percent contribution, of the environmental variables retained in the final model for adult (violet) and for immature (orange) bearded vultures in the a) warm and b) cold season.

Table S1: Environmental predictors used to model the occurrence probability of ibex and chamois across the Swiss Alps with unit, abbreviation and data source (warm season: May to October; cold season: November to April).

Category	Description	Unit	Abbreviation	Source
Climate	Average precipitation in the cold season	mm	prec122	WSL ¹
	Average precipitation in the warm season	mm	prec57	WSL
	Average solar radiation	WH/m ²	solar_rad	DHM ²
	Average temperature in the cold season	°C	temp122	WSL
	Average temperature in the warm season	°C	temp57	WSL
Topography	<i>Sine</i> of the aspect	-1 to 1	eastness	DHM
	<i>Cosine</i> of the aspect	-1 to 1	northness	DHM
	Standard deviation (SD) of the altitude	m	roughness	DHM
	Slope	degree	slope	DHM
	Topographic Position Index ³	Index	tpi	DHM
Geology	Limestone frequency	%	limestone	GeoKarten500 ⁴
	Gneiss frequency	%	gneiss	GeoKarten500
	Granite frequency	%	granite	GeoKarten500
	Other rock frequency	%	other_rocks	GeoKarten500
Anthropogenic	Distance from skilifts	m	skilift_d	Vector 25 ⁵
	Distance from cableways, not including ski-lifts	m	cableway_d	Vector 25
	Distance from roads and railways	m	road_rail_d	Vector 25
Land cover	Bush frequency	%	bush	Vector 25
	Forest frequency	%	forest	Vector 25
	Glacier frequency	%	glacier	Vector 25
	Marshland frequency	%	marsh	Vector 25
	Grassland and unproductive vegetation frequency	%	other_veg	Vector_25 + BsF ⁶
	Permanent culture frequency	%	permanent	Vector 25
	Rock frequency	%	rock	Vector 25
	Scree frequency	%	scree	Vector 25
	Distance from lakes and rivers	m	water_d	Vector 25
Distance from rock areas steeper than 45°	m	rock45_d	Vector 25 + DHM	

¹ Federal Institute for Forest, Snow and Landscape Research WSL; available upon request: www.wsl.ch.

² Digital Height Model of Switzerland (Swisstopo): https://shop.swisstopo.admin.ch/en/products/height_models/dhm25.

³ Topographic position index according to Wilson (1984).

⁴ Geo Maps: <https://shop.swisstopo.admin.ch/de/products/maps/geology/GK500>.

⁵ Digital Cartographic Model of Switzerland (Swisstopo): <https://shop.swisstopo.admin.ch/en/products/maps/national/vector/smv25>.

⁶ Federal Administration for Statistic Switzerland (BsF): <https://www.bfs.admin.ch/bfs/de/home/statistiken.html>.

Ibex and Chamois occurrence probability

Ibex and chamois occurrence probability were modeled following the same procedure as described in methods. The datasets containing observations of the two species were provided by “info fauna” (Centre suisse de cartographie de la faune (CSCF): <http://www.cscf.ch/cscf/de/home.html>). The ibex dataset contained 7576 observations and was modelled using 10000 background locations. The chamois dataset included 18232 observations and was modelled using 20000 background locations.

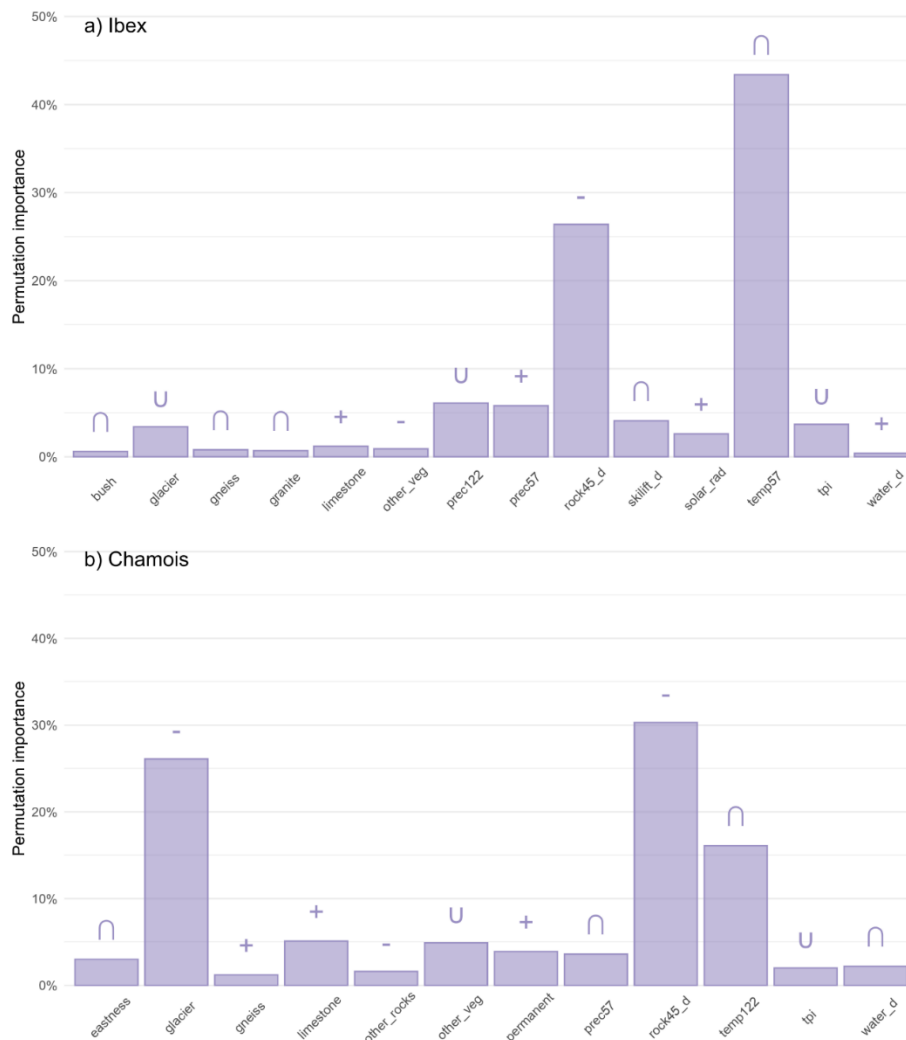
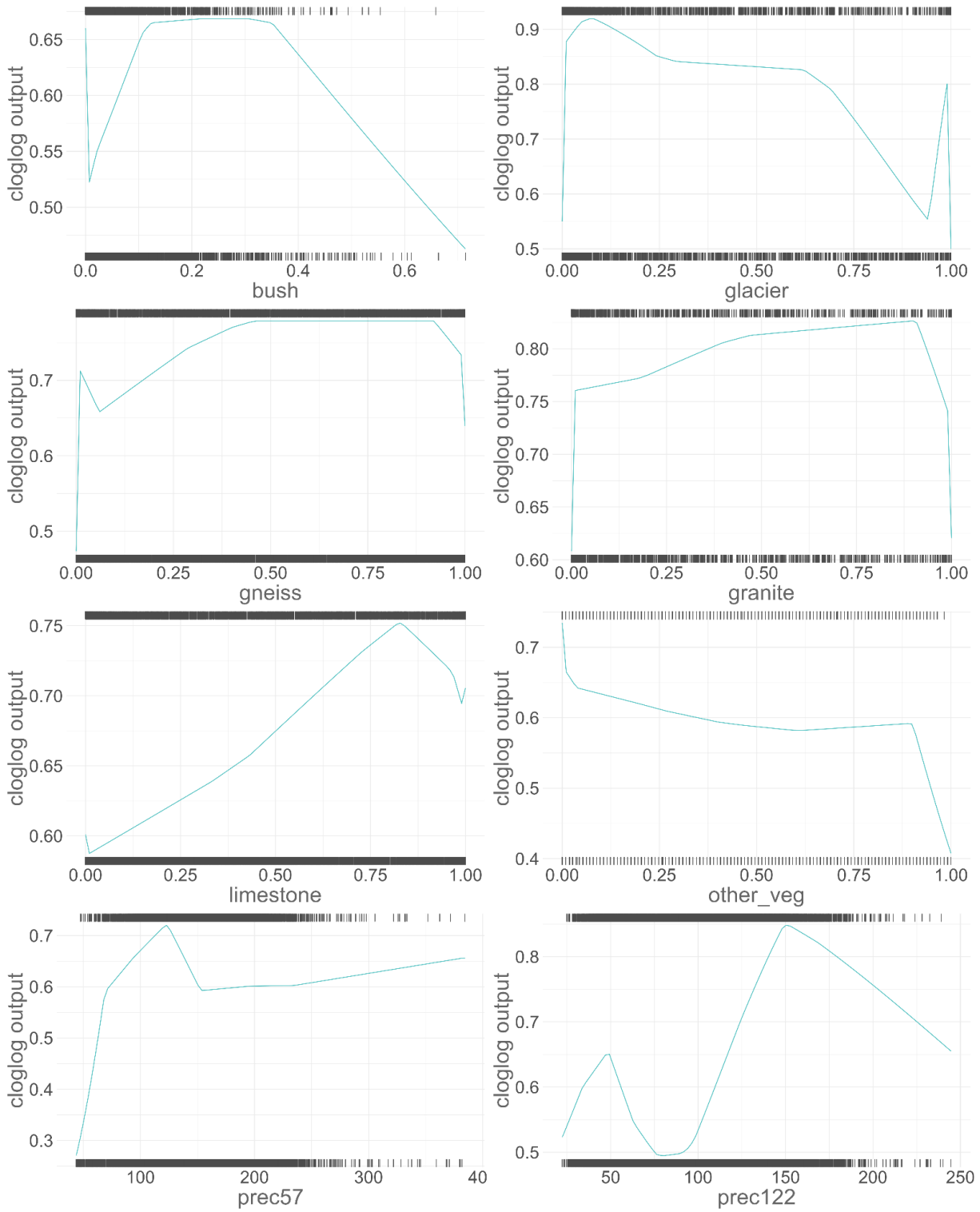


Figure S6: Variable importance, given as permutation importance, of the environmental variables retained in the final models predicting species occurrence for a) ibex and b) chamois. The symbol above the bar indicates the response type for the univariate model trained using every variable separately (see Figures S7-8), with +, -, ∩ or ∪ indicating a positive, negative, unimodal or bimodal response, respectively. The relative (normalized percentage) permutation importance is computed by randomly shuffling the values of each environmental predictor at a time for both training and background locations, evaluating the model with the shuffled data, and measuring the drop in training AUC. See Table S1 for the variable abbreviations and descriptions.

Ibex



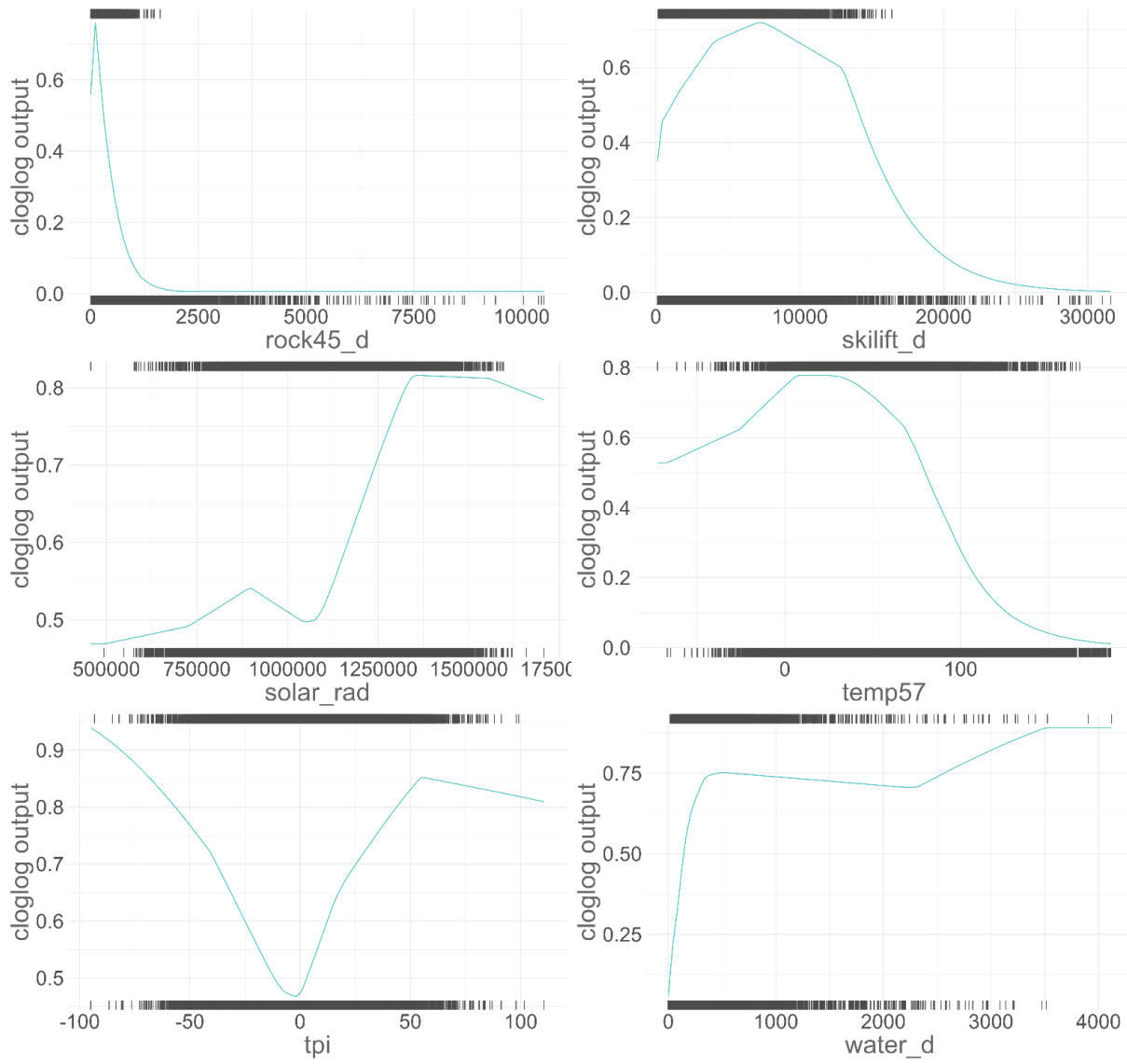
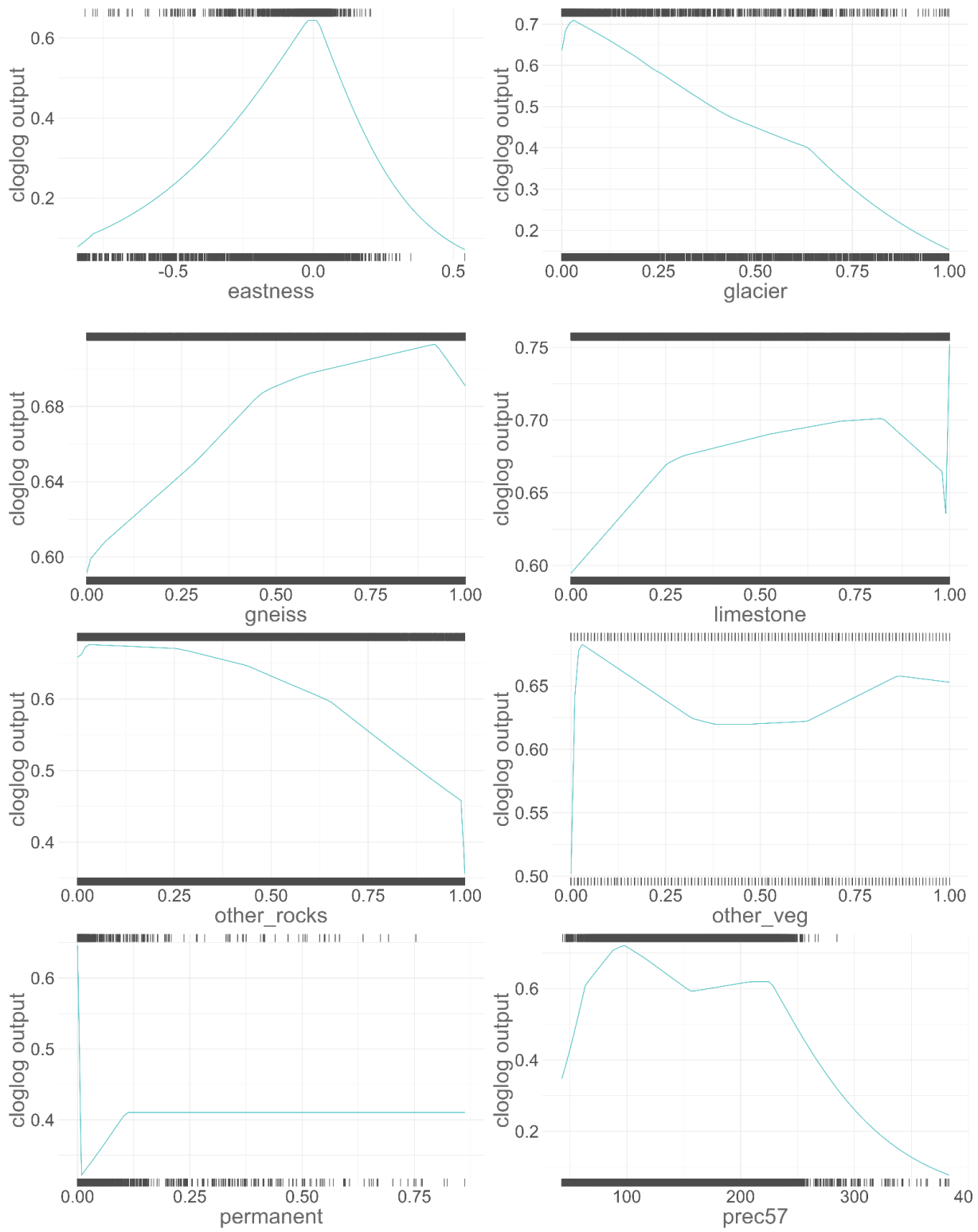


Figure S7: Response of ibex to the environmental variables included in the final model. The response curves show the change in cloglog prediction for a model trained using only the displayed variable (i.e. univariate model). The values of the variable at presence locations are shown on the top of the plot, the values at background locations on the bottom. Table S1 for the variable abbreviations and descriptions.

Chamois



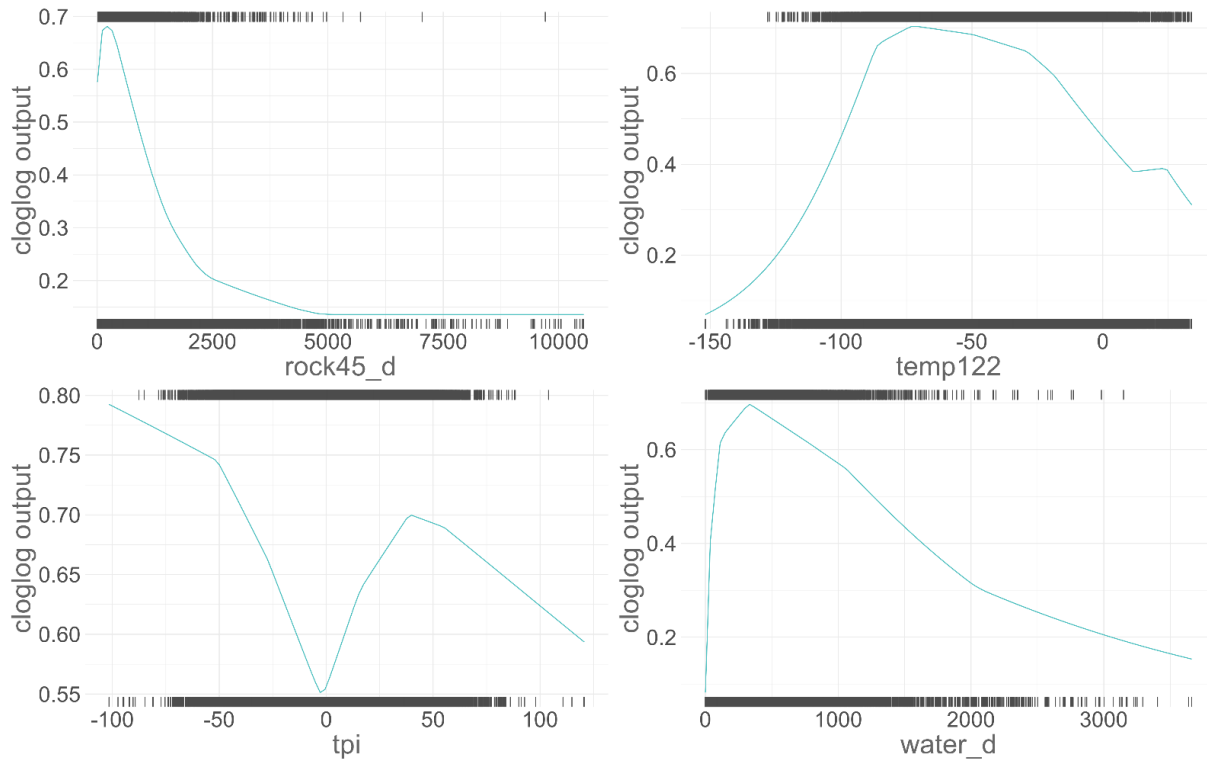


Figure S8: Response of chamois to the environmental variables included in the final model. The response curves show the change in cloglog prediction for a model trained using only the displayed variable (i.e. univariate model). The values of the variable at presence locations are shown on the top of the plot, the values at background locations on the bottom. Table S1 for the variable abbreviations and descriptions.

REFERENCES

Wilson, J. D. (1984). Determining a TOPEX score. *Scottish Forestry*, **38**(4), 251–256.

CHAPTER 2



***SDMtune*: An R package to tune and evaluate species distribution models**



Chapter 2 of this thesis is published in Ecology and Evolution: <https://doi.org/10.1002/ece3.6786> under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>.

***SDMtune*: An R package to tune and evaluate species distribution models**

Sergio Vignali¹, Arnaud G. Barras¹, Raphaël Arlettaz¹ & Veronika Braunisch^{1,2}

¹ Division of Conservation Biology, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012, Bern, Switzerland

² Forest Research Institute of Baden-Wuerttemberg, Wonnhaldestrasse 4, D-79100, Freiburg, Germany

Published in *Ecology and Evolution* (2020) **10**: 11488–11506

<https://doi.org/10.1002/ece3.6786>

ABSTRACT

Balancing model complexity is a key challenge of modern computational ecology, particularly so since the spread of machine learning algorithms. Species distribution models are often implemented using a wide variety of machine learning algorithms that can be fine-tuned to achieve the best model prediction while avoiding overfitting. We have released *SDMtune*, a new R package that aims to facilitate training, tuning, and evaluation of species distribution models in a unified framework. The main innovations of this package are its functions to perform data-driven variable selection, and a novel genetic algorithm to tune model hyperparameters. Real-time and interactive charts are displayed during the execution of several functions to help users understand the effect of removing a variable or varying model hyperparameters on model performance. *SDMtune* supports three different metrics to evaluate model performance: the area under the receiver operating characteristic curve, the true skill statistic, and Akaike's information criterion corrected for small sample sizes. It implements four statistical methods: artificial neural networks, boosted regression trees, maximum entropy modeling, and random forest. Moreover, it includes functions to display the outputs and create a final report. *SDMtune* therefore represents a new, unified and user-friendly framework for the still-growing field of species distribution modeling.

Keywords: ecological niche model, fine-tuning, genetic algorithm, machine learning, model complexity, variable selection

INTRODUCTION

Species distribution models (SDMs) encompass a variety of methods used to predict the occurrence of a species from the environmental conditions at a given site, thus providing a proxy of habitat suitability (Warren & Seifert, 2011). These methods are increasingly used in various fields of ecology (Guisan & Thuiller, 2005), often with the aim of guiding decision-making in species conservation management and planning (Guisan et al., 2013). Indeed, SDMs represent a crucial and cost-effective tool to identify current important areas for threatened species, and to forecast ecosystem impacts of rapid human-induced environmental change (Beaumont et al., 2016; Elith et al., 2010; Franklin, 2013; Guillera-Aroita et al., 2015; Guisan et al., 2013; Schwartz et al., 2006). Among the variety of available algorithms, machine learning approaches are becoming increasingly popular, facilitated by the recent availability of high computational power, and due to their ability to fit complex nonlinear relationships without requiring an a priori definition of a data model (Breiman, 2001). However, there still are many decisions to be made at various steps of the model building process that can influence the final output (Guisan & Thuiller, 2005). For example, the amount of complexity should be cautiously controlled to avoid models that underfit or overfit the underlying data (Merow et al., 2014; Warren & Seifert, 2011).

In general, the amount of complexity of a model depends on the number of chosen predictors and their transformations (Merow et al., 2014). Moreover, each machine learning algorithm has a series of parameters, known as hyperparameters. In contrast to model parameters, which are estimated from the data during model training, hyperparameters have a fixed value that must be defined before model training. Even if most machine learning algorithms have predefined default values, the optimal value of each hyperparameter is unknown, as it is specific to the modeling problem and the dataset. However, its choice affects model complexity and/or performance. For example, in a neural network, the maximum number of iterations controls the amount of iterations executed by its optimization algorithm. This value does not affect model complexity but if it is too low the algorithm might not converge, thus generating a model with lower performance. On the other hand, increasing the size of the hidden layer increases the number of parameters of the model and consequently its complexity, which in turn might affect its performance. In a Maxent model (Phillips et al., 2006), the amount of regularization controls overfitting by shrinking some parameters toward zero which consequently penalizes model complexity. Although several authors have stressed the importance of inspecting the

hyperparameters because default settings did not always yield an optimal performance (Elith et al., 2010; Merow et al., 2013; Warren et al., 2014; Warren & Seifert, 2011), the time-consuming task of comparing models trained with a multitude of possible combinations of hyperparameters' values (e.g., Zeng et al., 2016) may discourage many researchers from doing so in practice.

In order to optimize model complexity and performance, both the predictors used to build the model and the values of hyperparameters should be carefully selected which represents a challenge given the often numerous possible options. The new package *SDMtune* described here offers a framework to build and systematically tune SDMs. The package includes utilities that help R users (R Core Team, 2019) all along the analysis process, from data preparation to graphical representation of the results and reporting. In particular, it contains dedicated functions to perform variable selection and hyperparameter tuning. Hyperparameter tuning, also called hyperparameter optimization, is a process usually based on a trial and error experiment during which several models with different values of the hyperparameters are trained and evaluated in order to identify which combination yields the best performance. The simplest algorithm for hyperparameter tuning, grid search, trains and compares models with all possible combinations of the defined hyperparameters' values and can thus be a very time-consuming process. While other available R packages contain functions for tuning one (e.g., *ENMeval* (Muscarella et al., 2014), *wallace* (Kass, Pinilla-Buitrago, et al., 2018)), *kuenm* (Cobos et al., 2019) or several statistical model types (e.g., *biomod2* (Thuiller et al., 2019), *sdm* (Naimi & Araújo, 2016), *zoon* (Golding et al., 2018) and *caret* (Kuhn et al., 2019)), functions for data-driven variable selection are not always included and the hyperparameter tuning is always based on grid search or random search algorithms. *SDMtune* offers an alternative that relies on a genetic algorithm for exploring the hyperparameter configuration space (Lessmann et al., 2005; Young et al., 2015), applicable to the most commonly used SDM algorithms. This method significantly reduces the time required to find a near-optimal or the optimal model configuration. As an additional advantage, all functions for selecting the variables and tuning the hyperparameters are supported by an interactive real-time displayed chart that shows the change in model performance during the different steps of function execution. The chart is created in the RStudio (RStudio Team, 2018) viewer pane using the open source library *Chart.js* (<https://www.chartjs.org>), thus facilitating the understanding of the underlying algorithm action through a graphical representation of the output and avoiding the user's feeling of handling a black box that usually comes up when dealing with complex methods.

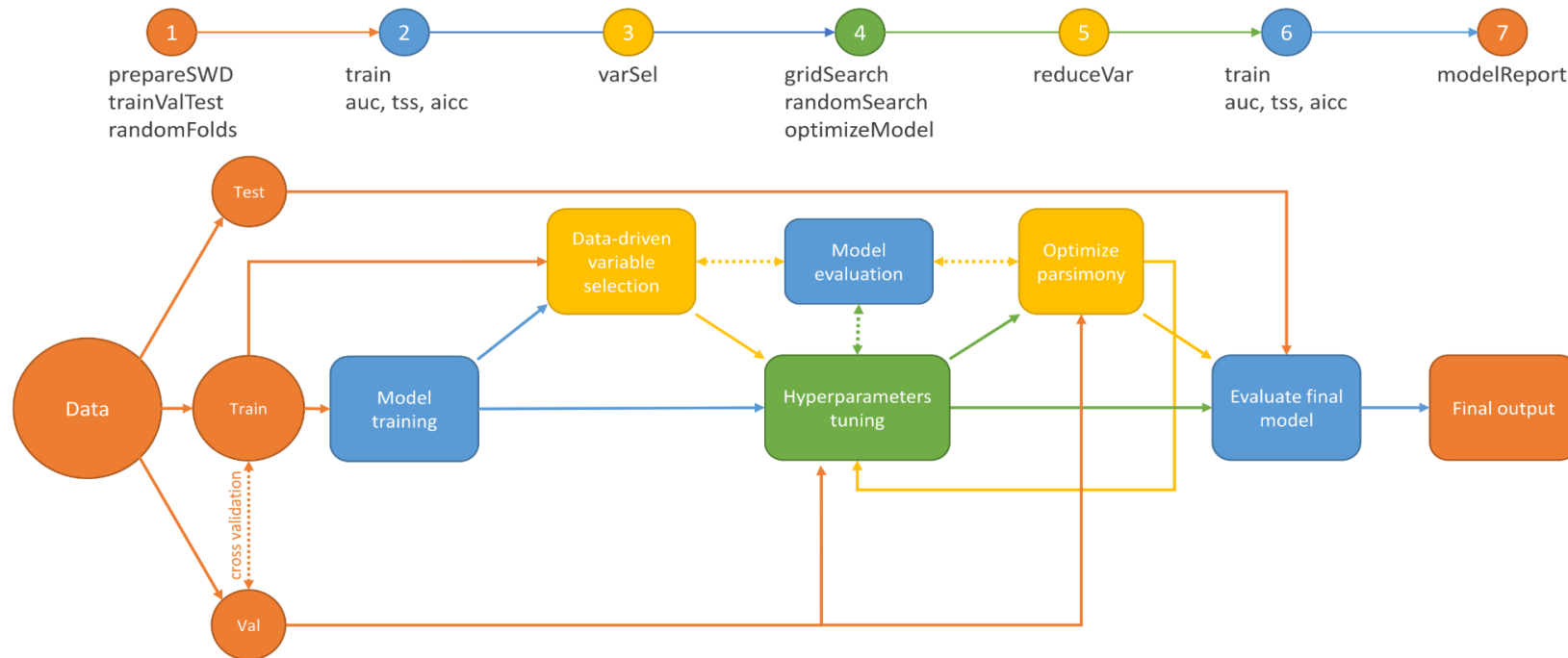


Figure 1: Package workflow illustrating the seven steps of the model tuning process. The functions required to perform the different steps are given in the headline. The different colors indicate different types of steps with: orange: preparation of data and results; blue: model training and evaluation; yellow: variable selection; green: hyperparameter tuning. Dashed connections represent an iterative process.

PACKAGE WORKFLOW AND DESCRIPTION

In this section, we present a possible use of the *SDMtune* package that covers a complete analysis in seven steps (Figure 1): (1) preparing data for the analysis; (2) training and evaluating a model; (3) performing variable selection; (4) tuning model hyperparameters; (5) optimizing model parsimony; (6) evaluating the final model; and (7) generating an output report. Users can combine the available functions in a way that best suits them. For example, step 4 could be repeated after step 5 to further fine-tune model hyperparameters.

Preparing data for the analysis

SDMtune uses a special object to compile the data for the analysis. This object, called SWD (samples with data, a format similar to the one used by the Maxent software), bundles all the information related to each record (name of the species, coordinates of the species' presence and absence/background locations, and the values of the environmental variables at these locations), thereby reducing the risk of mistakes in further analyses.

Before starting the analysis the user should decide which evaluation strategy to use. *SDMtune* provides two methods: (1) simple hold-out validation and (2) k-fold cross-validation. The k folds for the cross-validation can be created either randomly, using the provided **randomFolds** function, or spatially/environmentally clustered, using functions included in the packages *ENMeval* and *blockCV* (Valavi et al., 2019): In this case, *SDMtune* will internally convert the folds into the required format. The selected validation strategy is used to perform the variable selection and/or tune the model hyperparameters in order to optimize the model performance and address overfitting. When tuning the hyperparameters, several models with different configurations are trained and evaluated in an iterative process that aims at improving the predictive performance on the validation dataset, or—if cross-validation is used—on the arithmetic mean of the evaluation metric across all folds. During this process, part of the information contained in the validation dataset is inevitably transferred into the trained model, even if the validation data are not directly used to train the model (Chollet & Allaire, 2018; Müller & Guido, 2016). It is therefore advisable to hold apart an independent partition of the data, that is, the testing dataset, to obtain an unbiased evaluation of the final model (Hastie et al., 2009; Merow et al., 2014).

The selection of a metric and a validation strategy should therefore be done early in the model tuning process, because it has implications on how the data should be split before training the first model. Note that the AICc score is computed using all the observation locations (Warren

& Seifert, 2011) and does not require to partition the observation data into training and validation.

Table 1. Overview of the hyperparameters that can be tuned per statistical method and underlying package.

Method	R package	Hyperparameters	Default value
ANN	<i>nnet</i> (Venables & Ripley, 2002)	Size of hidden layer	-
		Weight decay	0
		Initial random weights	0.7
		Number of iterations	100
BRT	<i>gbm</i> (B. Greenwell et al., 2019)	Number of trees	100
		Interaction depth	1
		Shrinkage	0.1
		Bag fraction	0.5
ME	<i>dismo</i> (Hijmans et al., 2017)	Feature class combinations	lqph ^a
		Regularization multiplier	1
	<i>Maxnet</i> (Phillips, 2017b)	Number of iterations	500
		Feature class combinations	lqph
RF	<i>randomForest</i> (Liaw & Wiener, 2002)	Number of randomly sampled variables	floor(sqrt(#variables))
		Number of trees	500
		Minimum size of terminal nodes	1

Note: The meaning of each hyperparameter can be found in the respective package documentation and default values, when available, are provided in the last column.

^a (l) linear, (q) quadratic, (p) product, and (h) hinge.

Training and evaluating a model

Currently, four machine learning methods are available (Table 1): artificial neural networks (ANN), boosted regression trees (BRT), maximum entropy (ME), and random forest (RF). Two different implementations of the ME method can be selected: “Maxent” to use the Java implementation (version 3.4.1 or higher) and “Maxnet” for the R implementation using the maxnet package (Phillips et al., 2006, 2017). There are specific arguments of the train function that can be used to set the model hyperparameters. By default, these arguments are set to the same values as implemented in the dependent packages.

A trained model can be evaluated using one of the three implemented metrics: (1) the area under the receiver operating characteristic (ROC) curve (AUC) (Fielding & Bell, 1997), (2) the true skill statistic (TSS) (Allouche et al., 2006), and (3) Akaike's information criterion corrected for small sample sizes (AICc, only for ME method) (Burnham & Anderson, 2004; Warren & Seifert, 2011). It should be noted that AICc is a relative measure describing the model fit in relation to complexity (parsimony) but holds no information on predictive performance. It can thus only be used to compare competing models trained using the same data but not for final model evaluation.

Performing the variable selection

When the environmental variables used to train the model are highly correlated, it is difficult to interpret the model output, especially the relative importance of the variables and their response curves. A common practice is thus to select a subset of variables among which collinearity falls below a predefined threshold. A reasonable approach to remove highly correlated variables is to base the selection on expert knowledge, that is, retaining the environmental variable that is most likely to be ecologically meaningful for the target species. When this is unknown, an alternative approach is a “data-driven” variable selection that uses the information contained in the data to select the variable with the highest explanatory value among those that are highly correlated. The function **varSel** iterates through several steps: Starting from a trained model, it checks if the variable ranked as the most important (using the permutation importance or, optionally for Maxent models, the percent contribution (Phillips, 2017a)) is correlated with any of the other variables, using a given correlation criterion (e.g., Spearman's rho) and correlation threshold. If so, a leave-one-out Jackknife test is performed, starting with the full model, and among all correlated variables the one that decreases least model performance on the training dataset is discarded. A new model without this variable is then trained and again checked for highly correlated variables. The process is repeated until the correlations among all retained variables fall below the predefined threshold. During the execution of the function **varSel**, a real-time chart shows which variable is removed and the relative effect on the model performance.

Tuning the model hyperparameters

Tuning the model hyperparameters is a long process, as it requires testing many combinations of the hyperparameters in order to identify the best performing model. The simplest tuning method, known as “grid search,” is implemented in the function **gridSearch**. The user has the possibility to

define a set of possible values for one or several hyperparameters, out of which the function will create all possible combinations. The function also returns the value of the chosen evaluation metric so that the user can see the effect of varying the hyperparameters on the model performance.

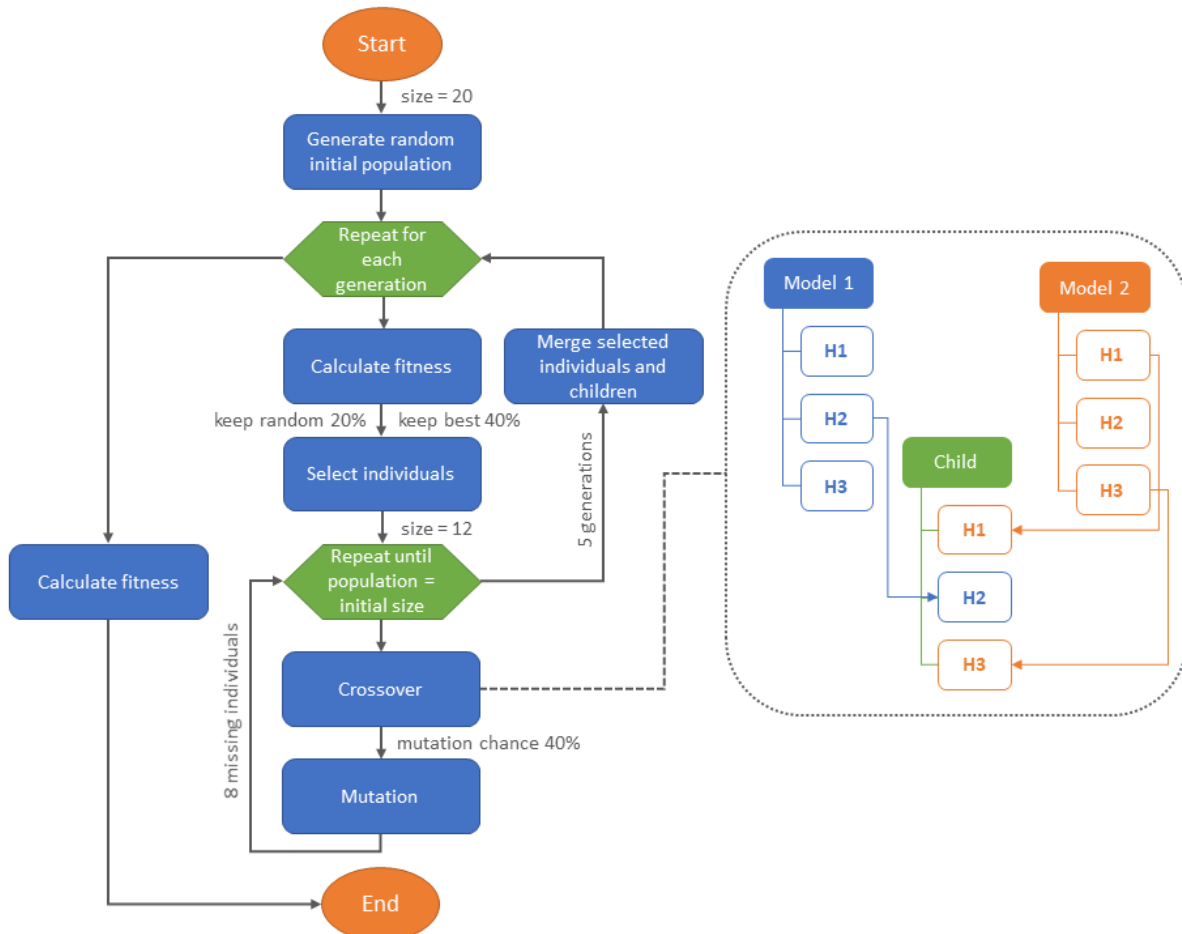


Figure 2: Flowchart illustrating the steps of the genetic algorithm implemented in the function `optimizeModel`, with orange ovals representing the begin and the end of the algorithm, blue boxes the main operations executed by the algorithm, and the green hexagons the iteration loops. In gray are provided the default values used by the function, with “size” indicating the initial population size; “keep best” the proportion of best models retained; “keep random” the proportion of less performing models retained; “mutation chance” the probability that a mutation event occurs. Keep best and keep random are provided as proportion of the initial population size. The dotted box shows an example of crossover during which two models, randomly selected from the selected “individuals”, are combined to generate a child model that inherits the first and third hyperparameters’ values from Model 2 and the second from Model 1. When the number of generations is zero, the flowchart represents the algorithm implemented in the function `randomSearch`.

Grid search is based on a brute force method that results in a very time-consuming process with high computational costs. A possible alternative is to randomly select some hyperparameters' combinations among the user-defined values (Bergstra & Bengio, 2012). This approach is implemented in the **randomSearch** function that usually finds a better performing model compared with the starting one. However, the disadvantage of the grid search and random search methods is that they do not use any information acquired during the iteration through the hyperparameter configuration space in order to improve the model performance. The function **optimizeModel** applies a genetic algorithm (Holland, 1992) instead, to more quickly optimize the combination of the hyperparameters (an example of a genetic algorithm used to define hyperparameters and architecture of a deep neural network is presented by Miikkulainen et al. (2019)). The algorithm (Figure 2) starts by generating a random initial “population” of models (using the **randomSearch** algorithm), with a given “population size”. The “fitness” of the population is measured with the chosen evaluation metric computed on the validation dataset and models are ranked accordingly. During the evaluation of the “fitness,” underfitting is controlled by ensuring that models for which the evaluation metric computed for the validation dataset is higher than the one computed for the training dataset are ranked in the last positions. At this point starts, the selection process during which some models (“individuals”) are selected according to their “fitness” from the initial “population” to create the first “generation.” There are two selection criteria. At first, a predefined proportion of the “fittest” models (i.e., models ranked in the first positions) is retained. Afterward, a small portion of the poor performing models (i.e., those not selected as “fittest”) is randomly retained in order to keep more variation in the population and reduce the possibility that the algorithm falls in a local optimum. The retained models are then submitted to the optimization process: they are “bred” (i.e., combined) to create other “individuals” and to reach again the predefined “population” size. In this process, two models, called “parents,” are randomly selected from the retained models (“selected individuals”) to “breed” and generate a “child.” This new model will randomly inherit a value for each hyperparameter from one of the “parents,” a process called “crossover.” A “mutation” chance with a predefined probability is added to increase the variation in the population. When the “mutation” is triggered one of the model's hyperparameter is randomly selected and its value is randomly sampled from those available but not included in the “parents.” Once the population reaches the defined size, the “fitness” is calculated again, and the process is repeated for the number of generations specified in the function. The user can set all the arguments: population size, number of generations, fractions of both best

and worst performing models to be retained at each generation as well as the probability of mutation during crossover episodes, but default values—that will work in most cases—are also defined. All the functions described in this section come with a real-time chart showing the model performance while the algorithm is running in the background.

Optimizing model parsimony

As soon as an optimal hyperparameter combination has been selected, we may want to reduce model complexity by removing some environmental variables ranked as less important. The function **reduceVar** automates this with two alternative approaches: (a) removing all the variables with an importance lower than a given threshold in a stepwise fashion, starting from the variable with the lowest importance; (b) removing the variables only if the model performance does not decrease compared to the initial model, according to a given evaluation metric. In the second case, a leave-one-out Jackknife test is performed. If removing one of the variables ranked below the given threshold does not decrease the performance of the model on the validation dataset compared to the initial model, that variable is discarded. A new model, trained without this variable, is checked again with the Jackknife test, and the process is repeated until all the variables with an importance lower than the given threshold are either retained or discarded. A real-time chart showing the removed variable together with its relative effect on model performance is generated during the execution of the function.

Evaluating the final model

At this point, after the variable set has been optimized (**varSel** and **reduceVar**) and the hyperparameters of the model have been tuned (**gridSearch**, **randomSearch**, or **optimizeModel**) the model can be evaluated on the held apart testing dataset, which was never used during the tuning procedure, using one of the functions that compute the chosen metric (i.e., AUC or TSS). Another possibility would be to train a new model using the selected variables and hyperparameter combinations with the full dataset (i.e., without applying cross-validation or data partitioning) and evaluate it on the held apart testing dataset (Chollet & Allaire, 2018). This way the model can avail of a greater amount of information and might thus be able to generalize better.

Creating the output

There are several functions for visualizing the model results and predictions. The user can plot the

response curves, the variable importance, the ROC curve, project the predicted relative probability of species occurrence to the extent of the environmental variables, plot and save the results as a map with either continuous (relative occurrence probabilities) or binary (presence-absence) values based on various threshold values. *SDMtune* implements its own algorithm to make predictions for “Maxent” models without calling the Java software. This results in a much faster execution that allows considerably speeding up projections, which is particularly useful when models are evaluated using the AICc, as this metric requires computing the Maxent raw output for the full geographic extent of the environmental variables. The prediction produced by our algorithm may—in some cases—differ marginally from the output of the Java implementation, which could result in only slightly different prediction values. Finally, the **modelReport** function creates a report similar to the one produced by the Maxent software, for all methods.

PERFORMANCE ASSESSMENT OF GENETIC ALGORITHM

We evaluated the performance of the genetic algorithm in terms of time-saving and model accuracy for the four SDM-methods available in *SDMtune* by comparing the output of the **optimizeModel** and **gridSearch** functions. We used the **virtualSp** dataset provided with the package. This dataset contains simulated presence, absence, and background locations generated with the package *virtualspecies* (Leroy et al., 2016). For artificial neural network, boosted regression trees, and random forest we used the presence and absence datasets, while for the maximum entropy method we used the presence and background datasets. The maximum entropy method was performed with the “Maxnet” implementation. In all cases, a 10-fold cross-validation was used as validation strategy and the AUC was used as evaluation metric. As first step, we trained a model with default hyperparameters' values (for artificial neural network we used an inner layer of a size equal to the number of environmental variables), and then executed the two functions testing 1200 possible hyperparameters' combinations (Table A1, for the **optimizeModel** function we used default arguments). The results of the analysis are presented in Table 2. In all cases, the **optimizeModel** functions found a near-optimal solution in a significantly reduced amount of time.

Table 2. Performance assessment of the **gridSearch** compared to the **optimizeModel** function for model tuning regarding execution time (expressed as HH:MM:SS) and evaluation metric (on the training dataset “Train AUC”, the validation dataset “Val AUC”, given as arithmetic mean across the folds of a 10-folds cross validation) for the four methods implemented in *SDMtune*. Models were trained using the **virtualSp** dataset available with the package and 1200 possible hyperparameters’ combinations. Presence and background locations were used for the Maxnet method, presence and absence locations for the other methods.

Method	Default model		Genetic algorithm		Time	Grid search		
	Train AUC	Val AUC	Train AUC	Val AUC		Train AUC	Val AUC	Time
ANN	0.8600	0.8619	0.9839	0.9590	00:11:44	0.9814	0.9615	05:51:33
BRT	0.9873	0.9750	0.9905	0.9779	00:01:33	0.9892	0.9787	00:29:45
RF	1	0.9724	1	0.9740	00:02:16	1	0.9735	00:48:03
Maxnet	0.8681	0.8561	0.8710	0.8565	00:17:49	0.8702	0.8567	05:01:21

EXAMPLE OF APPLICATION: BEARDED VULTURE IN THE SWISS ALPS

To demonstrate possible applications of *SDMtune*, we used 1947 observation locations of the Bearded vulture (*Gypaetus barbatus*) collected in Switzerland between 2004 and 2017. The occurrences were gathered from two databases: the International Bearded Vulture Monitoring (IBM) database and ornitho.ch, the official birding exchange platform in Switzerland. Clumped observations were removed with a spatial thinning of 250 m using the *spThin* package (Aiello-Lammens et al., 2015). We randomly split the observations into two partitions and used 80% (1363 observations) as training dataset and the remaining 20% (584) as testing dataset. A set of 39 environmental predictors that might be relevant for the species was prepared for the analysis, as using numerous predictors together with a large amount of species observations allows for a better illustration of the advantages and time-saving functionalities provided by our package. The variables included information on topography, climate, geology, anthropogenic infrastructure, land cover, and food availability, referring to Hirzel et al. (2004). All predictors were prepared as raster maps with a resolution of 100×100 m, with each cell containing the average value of the respective variable within a 1 km² circular moving window (a list of the variables is provided in Appendix A, Table A2). The whole analysis was conducted using R version 3.6.0 (R Core Team, 2019).

We performed the data-driven variable selection using the function **varSel** on the initial set of 39 predictors. As a first step, we trained a model using the “Maxent” method with default settings (i.e., linear, quadratic, product and hinge as feature class combinations, regularization multiplier equal to 1, 10,000 background locations and 500 iterations) and the 39 environmental variables.

We then used the **varSel** function to execute the variable selection using the percent contribution to rank variable importance and the AUC as performance metric. The function arguments were set to check for Spearman's correlation coefficients $|rs|$ greater than or equal to 0.7, based on 30,000 random background locations (Table A3).

Table 3. Performance of the **gridSearch** compared to the **optimizeModel** function for model tuning regarding execution time (expressed as HH:MM:SS) and evaluation metric (on the training dataset “Train AUC”, the validation dataset “Val AUC” and the difference between both “Diff AUC”, given as arithmetic mean of the 4-folds cross validation) on the case example data of the bearded vulture.

h	Algorithm	Exec. time	Train AUC	Val AUC	Diff AUC	FC^a	reg	iter
75	gridSearch	01:29:24	0.8687	0.8581	0.0106	lqph	3.0	500
	optimizeModel	01:06:50	0.8687	0.8581	0.0106	lqph	3.0	500
150	gridSearch	02:29:16	0.8687	0.8581	0.0106	lqph	3.0	500
	optimizeModel	01:16:25	0.8687	0.8581	0.0106	lqph	3.0	500
300	gridSearch	04:56:44	0.8691	0.8584	0.0107	lqph	2.9	500
	optimizeModel	01:15:09	0.8691	0.8581	0.0110	lqph	2.8	500
600	gridSearch	10:57:43	0.8707	0.8588	0.0119	lqph	2.7	700
	optimizeModel	01:18:46	0.8707	0.8588	0.0119	lqph	2.7	700
1200	gridSearch	21:14:45	0.8706	0.8588	0.0118	lqph	2.8	900
	optimizeModel	01:06:58	0.8700	0.8550	0.0149	lqph	1.9	700

Note: The models were trained using the Maxent method.

The number of tested hyperparameters' combinations is given by “h”. A description of the exact hyperparameters' combinations is provided in Appendix A, Table A5. “FC” represents the feature class combination, “reg” the regularization multiplier and “iter” the number of iterations for the best performing model.

^a FC: (l) linear, (q) quadratic, (p) product, and (h) and hinge.

Starting with the model trained using the 28 selected variables (i.e., the output of the **varSel** function, Table A4), we conducted a simple experiment to investigate the performance of the **optimizeModel** compared to the **gridSearch** function in terms of execution time and best hyperparameter combination. We selected the AUC as the performance metric running a fourfold cross-validation. The folds were created by randomly splitting the training dataset with the function **randomFolds**. For the **optimizeModel** function, we used the default arguments: a population size

of 20 models, five generations, kept 40% of the best performing models, randomly retained 20% of the less performing ones and used a mutation chance of 40%. We tested different sets of hyperparameters (Table A5 and Figure A1), varying the feature class combinations, the regularization multiplier and the number of iterations. The results illustrate how using the **optimizeModel** function tremendously reduces computation time while providing a near-optimal solution when the number of hyperparameter combinations increases (Table 3). In our experiment, with 1200 possible hyperparameter combinations, the execution time dropped from 21 hr 14 min and 45 s using **gridSearch** to 1 hr 6 min and 58 s using **optimizeModel** with a similar predictive performance of the resulting models (mean validation AUC across the fourfold of 0.8588 and 0.8550, respectively).

Table 4. Comparison of model performance between models tuned using the genetic algorithm and grid search implemented in the **optimizeModel** and **gridSearch** function respectively, on the case example data of the Bearded vulture.

Step	Dataset	Model performance	
Starting model	Training	0.863	
	Validation	0.846	
	Testing	0.836	
		Genetic algorithm	Grid search
Hyperparameter tuning	Training	0.870	0.871
	Validation	0.856	0.859
	Testing	0.848	0.853
Optimize parsimony	Training	0.865	0.868
	Validation	0.856	0.859
	Testing	0.846	0.854
Final model evaluation	Training	0.862	0.867
	Testing	0.846	0.855

Note: Performance is evaluated with the AUC metric on the training, validation and testing datasets as mean AUC of the 4-fold cross validation at different steps of the modelling workflow. For the final model evaluation the model is trained merging training and validation datasets. The testing dataset refers to the dataset held apart and not used during the hyperparameter tuning and optimize parsimony steps and the starting model is the model trained after removing highly correlated variables.

In a next step, we investigated whether the final evaluation of the resulting models provided similar results. For this purpose, we selected the models with the optimized hyperparameters' combination (i.e., the output of the **optimizeModel** and **gridSearch** functions run with 1200 different hyperparameters' combinations). We used the **reduceVar** function to test if some variables with low contribution could be removed without affecting the validation AUC. We considered the Maxent percent contribution to rank the environmental variables, a threshold of 2% for variable removal and used the Jackknife approach. We could remove nine and seven environmental variables, respectively, without reducing the mean validation AUC (Table A6 and Figure A2).

Finally, we trained a model using the full training dataset without cross-validation, the selected environmental variables and the best hyperparameter configuration found by the two functions. We estimated the performance of these tuned models on the held apart testing dataset, obtaining very similar results (Table 4).

DISCUSSION

Most of the algorithms supported by the package have predefined default values for the hyperparameters, while ANN requires the size of the hidden layer to be specified (Table 1). Default values are not necessarily the best choice for any given dataset and modeling problem, and a tuning procedure can improve model performance considerably. For example, the default hyperparameters' values of the Maxent algorithm were derived based on an empirical tuning experiment conducted on 226 species (Phillips & Dudík, 2008), however, several authors found that these values were not always optimal for their specific datasets (Anderson & Gonzalez, 2011; Merow et al., 2013; Radosavljevic & Anderson, 2014; Warren et al., 2014; Warren & Seifert, 2011). While dedicated R packages are available for fine-tuning Maxent's hyperparameters, like *ENMeval* (Muscarella et al., 2014), *wallace* (Kass, Vilela, et al., 2018), and *kuenm* (Cobos et al., 2019), this process can be very time consuming (Table 2 and 3) and limiting, especially when performed for multiple species. With *SDMtune*, we introduce a genetic algorithm that drastically reduces the computation time of hyperparameter tuning while achieving an optimal or near-optimal model configuration.

While the **gridSearch** function can be preferred for tuning a single or a few hyperparameters, it quickly comes to its limits when testing numerous hyperparameters'

combinations. In this case, the **randomSearch** function may represent a valid and time-efficient alternative for finding a better model compared to the one trained with default settings. By taking a random subsample from predefined hyperparameters' combinations, it has to train only n models (with n equal to the population size, i.e., 20 by default). During this process, it may select the best combination simply by chance. This chance, however, decreases with an increasing amount of possible combinations. The function **optimizeModel**, in contrast, can achieve a better result in a reasonable amount of time, as it makes use of the information generated during the tuning process, thanks to the genetic algorithm. This function also trains a reduced amount of models compared to **gridSearch** with the amount depending on: (1) the population size; (2) the fractions of both best and worst performing models to be retained at each generation; (3) the number of generations, which results in 60 models when using the default settings.

Although there is no rule of thumb to decide when **optimizeModel** should be preferred to **gridSearch** or **randomSearch**, the choice can be supported by considering four important factors. The first and most important factor is the time necessary to train one model, which in turn depends on the sample size, the number of predictors, the selected method, and the setting of the hyperparameters. For instance, increasing the number of trees in RF or BRT increases the computation time as does decreasing the amount of regularization or using hinge or threshold future class combinations in ME methods. The second is the selected evaluation strategy: when k-fold cross-validation is performed, the required time to train one model is inflated by the factor k . Moreover, using k-fold cross-validation requires to compute the evaluation metric k times and compute their mean value, and this marginally increases the overall computation time. The third factor is the selected evaluation metric. To compute, the AICc is necessary to get the prediction for the whole study area which might take a long time in the case of large extents. The fourth factor is the number of hyperparameters' combinations used for the tuning procedure. Considering these aspects, the total amount of time necessary to tune the hyperparameters with the function **gridSearch** can be roughly estimated based on the time necessary to train and evaluate one model multiplied by the number of possible hyperparameters' combinations. Furthermore, the two functions could also be used in conjunction: the **optimizeModel** function returns n models, with n equal to the size of the predefined “population” of models (20 by default), ordered by decreasing model performance. The user could inspect the values of the hyperparameters of the returned models and further refine them using **gridSearch**. Finally, in case different hyperparameter-configurations result in models with equal or similar values of the evaluation metric, the selection

of one among the competing models can be based on further exploration, for example, by computing multiple evaluation metrics. It should be stressed that hyperparameter tuning is limited to the values of the hyperparameters defined by the user and thus is not exhaustive. What we defined “best model” refers to the best model among those trained with all the possible combinations of the predefined values. Therefore, the definition of these values determines the quality of the final model.

The genetic algorithm implemented in the function **optimizeModel** relies on some parameters that govern the optimization process (i.e., population size, number of generations, fractions of both best and worst performing models to be retained at each generation, and probability of mutation during crossover episodes), which are provided with default values. We defined these values based on a deep understanding of the algorithm and after testing it multiple times on varying datasets. Nevertheless, although these values could have been hard coded into the source code, we decided to provide a more flexible function making them available as arguments. In the performance assessment of the genetic algorithm and in the example of application presented here (Table 2 and 3), default values worked when testing as much as 1200 predefined hyperparameters' combinations. In case of a similar or higher amount of hyperparameters' combinations, these values might require small adjustments to introduce more variability, for instance by increasing the population size and the probability of mutation.

With the implementation of the genetic algorithm, we introduced a new way of hyperparameters optimization in the field of SDMs. This way could be extended further by testing different modifications. For example, in our implementation only one model is created during the “crossover” event, but two “sibling” models could be produced instead. Furthermore, other optimization algorithms, like the Bayesian optimization, could be implemented. With our **optimizeModel** function, we provide a first implementation of a new algorithm that can be extended in future releases of the package.

Not only the tuning of hyperparameters, but also the selection of environmental variables for SDMs has gained attention in recent years (Jueterbock et al., 2016; Warren et al., 2014; Zeng et al., 2016). Despite the fact that highly correlated environmental variables are not a problem when the aim of the study is prediction in the same extent of the observed data, reducing collinearity is recommended in order to reduce model complexity and increase the interpretability of the predictors (Dormann et al., 2013; Merow et al., 2014). In addition, although the degree of accepted model complexity varies according to the modeling scope(s) (Halvorsen, 2012; Halvorsen et al.,

2015), it has been pointed out that models might perform best when trained with a reduced number of predictors (Brun et al., 2020; Halvorsen et al., 2015). Even though the selection should be driven by the knowledge of the modeled species, this might be difficult when the user must decide among several a priori ecologically relevant predictors for the species, or if the ecology of the species is poorly known. Cobos et al. (2019), with their package *kuenm*, provide a framework that enables tuning several models starting with different sets of environmental variables. Yet, this process still requires predefining the predictor sets. Warren et al. (2014) described a method where environmental variables are removed in a stepwise approach that accounts for regularization tuning, variable importance, and improvements in the AICc metric. A similar approach has been implemented in the package *MaxentVariableSelection* (Jueterbock, 2015), used by Jueterbock et al. (2016) to model the effect of climate change on the Arctic seaweed (*Fucus distichus*). In both examples, all predictors with a contribution-score lower than a given threshold and predictors highly correlated with the most important variable were removed simultaneously at each step. Given that removing a variable affects the contribution-score of the remaining predictors and therefore their resulting rank, our functions for data-driven variable selection remove only one variable at a time. For the same reason, removing highly correlated variables and variables with low contribution is performed by two distinct functions and not combined into the same process, as described in the previous examples. Furthermore, instead of relying merely on a variable's rank of importance for deciding which one to retain, our functions base the selection on a leave-one-out Jackknife test, while controlling the desired performance metric. Note that the **varSel** function aims at maintaining the value of the selected metric for the training dataset (i.e., removes the variables that decreases least the evaluation metric) while the **reduceVar** function aims to at least maintain the value of the selected metric for the validation dataset (i.e., removes a variable if the evaluation metric does not drop). The reasons are, first, that highly correlated predictors should be removed before performing any tuning, and second, that optimizing the selected metric for the training dataset allows capturing the information contained in the data, which is especially important if ecological selection criteria are lacking. The over- or underfitting can then be controlled later by fine-tuning the hyperparameters. On the other hand, removing variables with low predictive contribution aims to reduce model complexity and increase model generalization, which is why the validation dataset is used.

There are other R packages which include functions for variable selection. *Caret*, for instance, implements several methods based, among others, on simulated annealing, recursive

elimination, or a genetic algorithm. Whereas these methods aim at identifying the best subset of the available variables, our implementations address different problems: **varSel** removes variables to reduce collinearity, and **reduceVar** removes variables that contribute least to the model to increase parsimony. The functions for data-driven variable selection can be particularly useful when the fitted model is extrapolated in space or time. In such cases, the currently prevailing correlations among the environmental variables may differ from those observed in the new time periods or geographical areas (Braunisch et al., 2013), causing unexpected predictions (Warren et al., 2014). This risk is reduced with a reduced number of predictors. Moreover, reducing the number of predictors may limit overfitting, and thus result in a model that generalizes better and thus yields more accurate predictions for data not used during training. The selection of a threshold to reduce the number of variables with the function **reduceVar** is quite arbitrary. If the aim is to remove as many variables as possible while preserving model performance, the threshold could be set to 100 and the Jackknife method must be selected. On the contrary, if the user, based on his expertise, judges a certain variable as ecologically important for the species and wants to retain it in the model, he could define a threshold that is lower than the importance of this variable. Nevertheless, the functions presented in this article should not be applied blindly. Therefore, *SDMtune* provides interactive real-time charts to visualize every step of the algorithms with the idea that the user further evaluates the validity of the final output.

These charts are particularly useful for two reasons. First, because they are updated in real time, they confirm that the running function is working properly and is not frozen at some unknown step. This is especially important for functions that take long to be executed. Second, because they are interactive, different types of information can be provided without overloading a single graph, since extra information is embedded in a tooltip that appears when the user hovers over a specific element of the chart. Interactive real-time charts are well known and used in other fields that represent the state-of-the-art of machine learning, and available in few R packages such as *keras* (Allaire & Chollet, 2020) which allows the user to build complex deep learning models.

CONCLUSION

The new R package *SDMtune* enables data-driven variable selection and hyperparameters tuning within a unified and user-friendly framework. The core functions provide interactive real-time charts that represent the effect of each step of the implemented algorithms in relation to the model

performance and allow a deeper understanding of the automated processes. The new functions we present in this paper (i.e., genetic algorithm for hyperparameter tuning and automated variable selection) are implemented in a framework that also integrates functions already available in other packages. This unification, combining all required functions in a single package, offers the advantage for the user to learn a unique framework instead of jumping from one package to the other, each time having to adapt data structures. Currently, *SDMtune* supports three evaluation metrics (i.e., AUC, TSS, and AICc) and four modeling methods (i.e., ANN, BRT, RF, and ME) and more can be easily added in future releases.

Despite providing comprehensive descriptions and visual illustration of the functions, we still stress that users should be familiar with their data and the selected algorithm used to train their model. Particular attention should be paid to preparing the data before modeling. *SDMtune* also offers functions to prepare the data, but it is upon the user's knowledge and expertise to decide upon the most appropriate way to partition and filter the dataset, accounting for sample size and possible sampling biases, or which metric is best to evaluate the model in relation to the modeling objectives. In this respect Araújo et al. (2019) defined best-practice standards for SDMs stressing the importance of evaluating models with a temporally or spatially independent dataset (Araújo et al., 2019: Supplement S2.4B). For this reason, *SDMtune* supports functions well developed in other packages (*blockCV* and *ENMeval*) to produce such data partitions. These best-practices have recently gained importance and have been integrated in the ODMAP standard protocol (Zurell et al., 2020) that provides a workflow for reproducible and good quality analyses.

The package documentation provides a more complete description of all the available functions, and the articles hosted on the package webpage (<https://consbiol-unibern.github.io/SDMtune/>) describe meaningful examples of application in various fields of ecological research. These examples are also included in the package and accessible through the vignettes.

INSTALLATION

The package *SDMtune* is available in the CRAN repository at <https://CRAN.R-project.org/package=SDMtune> and can be installed in R with the command `install.packages("SDMtune")`. The package is under development and the source code is hosted in GitHub (<https://github.com/ConsBiol-unibern/SDMtune>). We encourage future users to provide

feedback and report bugs by opening an issue on the GitHub platform.

ACKNOWLEDGMENTS

We thank the Stiftung Pro Bartgeier, the International Bearded Vulture Monitoring, and the Swiss Ornithological Institute for providing the Bearded vulture locations used in the case study. This research was supported by the Swiss Federal Office for Energy, the Swiss Federal Office for the Environment and the following foundations: Parrotia Stiftung, Margarethe und Rudolf Gsell-Stiftung, Alfons und Mathilde Suter-Caduff Stiftung, WWF Switzerland, Beat und Dieter Jutzler Stiftung, UniBern Forschungsstiftung, Stiftung Dreiklang für ökologische Forschung und Bildung, Sophie und Karl Binding Stiftung, Stiftung Temperatio, Ernst Göhner Stiftung, Steffen Gysel Stiftung für Natur und Vogelschutz, and Samy Harshallanos.

REFERENCES

- Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, **38**(5), 541–545. <https://doi.org/10.1111/ecog.01132>
- Allaire, J. J., & Chollet, F. (2020). *keras; R Interface to “Keras”. R package version 2.3.0.0*. <https://CRAN.R-project.org/package=keras>.
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**(6), 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, **222**(15), 2796–2811. <https://doi.org/10.1016/j.ecolmodel.2011.04.011>
- Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O’Hara, R. B., Zimmermann, N. E., & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, **5**(1), eaat4858. <https://doi.org/10.1126/sciadv.aat4858>
- Beaumont, L. J., Graham, E., Duursma, D. E., Wilson, P. D., Cabrelli, A., Baumgartner, J. B., Hallgren, W., Esperón-Rodríguez, M., Nipperess, D. A., Warren, D. L., Laffan, S. W., & VanDerWal, J. (2016). Which species distribution models are more (or less) likely to project broad-scale, climate-induced shifts in species ranges? *Ecological Modelling*, **342**, 135–146. <https://doi.org/10.1016/j.ecolmodel.2016.10.004>
- Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization. *Journal of*

- Machine Learning Research*, **13**(10), 281–305.
- Braunisch, V., Coppes, J., Arlettaz, R., Suchant, R., Schmid, H., & Bollmann, K. (2013). Selecting from correlated climate variables: a major source of uncertainty for predicting species distributions under climate change. *Ecography*, **36**(9), 971–983. <https://doi.org/10.1111/j.1600-0587.2013.00138.x>
- Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, **16**(3), 199–215. <https://doi.org/10.1214/ss/1009213726>
- Brun, P., Thuiller, W., Yohann Chauvier, |, Loïc Pellissier, |, Wüest, R. O., Wang, Z., Niklaus, |, & Zimmermann, E. (2020). Model complexity affects species distribution projections under climate change. *Journal of Biogeography*, **47**(1), 130–142. <https://doi.org/10.1111/jbi.13734>
- Bundesamt für Energie BFE. (2016). *Windatlas Schweiz: Jahresmittel der modellierten Windgeschwindigkeit und Windrichtung*. Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK. www.bfe.admin.ch
- für Umwelt, Verkehr, Energie und Kommunikation UVEK.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, **33**(2), 261–304. <https://doi.org/10.1177/0049124104268644>
- Chollet, F., & Allaire, J. J. (2018). *Deep learning with R* (1st ed.). Manning.
- Cobos, M. E., Townsend Peterson, A., Barve, N., & Osorio-Olvera, L. (2019). Kuenm: An R package for detailed development of ecological niche models using Maxent. *PeerJ*, **7**(2), e6281. <https://doi.org/10.7717/peerj.6281>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**(1), 027–046. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**(4), 330–342. <https://doi.org/10.1111/j.2041-210x.2010.00036.x>
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**(1), 38–49. <https://doi.org/10.1017/S0376892997000088>
- Franklin, J. (2013). Species distribution models in conservation biogeography: developments and challenges. *Diversity and Distributions*, **19**(10), 1217–1223. <https://doi.org/10.1111/ddi.12125>
- Golding, N., August, T. A., Lucas, T. C. D., Gavaghan, D. J., van Loon, E. E., & McInerney, G. (2018). The zoon r package for reproducible and shareable species distribution modelling. *Methods in Ecology and Evolution*, **9**(2), 260–268. <https://doi.org/10.1111/2041-210X.12858>

- Greenwell, B., Boehmke, B., Cunningham, J., & GBM Developers. (2019). *gbm: Generalized Boosted Regression Models*. R package. <https://CRAN.R-project.org/package=gbm>.
- Guillera-Aroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., McCarthy, M. A., Tingley, R., & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, **24**(3), 276–292. <https://doi.org/10.1111/geb.12268>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., ... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, **16**(12), 1424–1435. <https://doi.org/10.1111/ele.12189>
- Halvorsen, R. (2012). A gradient analytic perspective on distribution modelling. *Sommerfeltia*, **35**(1), 1–165. <https://doi.org/10.2478/v10208-011-0015-3>
- Halvorsen, R., Mazzoni, S., Bryn, A., & Bakkestuen, V. (2015). Opportunities for improved distribution modelling practice via a strict maximum likelihood interpretation of MaxEnt. *Ecography*, **38**(2), 172–183. <https://doi.org/10.1111/ecog.00565>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. Springer.
- Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2017). *dismo: Species Distribution Modeling*. <https://cran.r-project.org/package=dismo>.
- Hirzel, A., Posse, B., Oggier, P.-A., Crettenand, Y., Glenz, C., & Arlettaz, R. (2004). Ecological requirements of reintroduced Bearded vultures colonizing the Alps : implications for release policy. *Journal of Applied Ecology*, **41**(6), 1103–1116. <https://doi.org/10.1111/j.0021-8901.2004.00980.x>
- Holland, J. H. (1992). *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press.
- Jueterbock, A. (2015). R package *MaxentVariableSelection: selecting the Best Set of Relevant Environmental Variables along with the Optimal Regularization Multiplier for Maxent Niche Modeling*. <https://CRAN.R-project.org/package=MaxentVariableSelection>.
- Jueterbock, A., Smolina, I., Coyer, J. A., & Hoarau, G. (2016). The fate of the Arctic seaweed *Fucus distichus* under climate change: An ecological niche modeling approach. *Ecology and Evolution*, **6**(6), 1712–1724. <https://doi.org/10.1002/ece3.2001>
- Kass, J. M., Pinilla-Buitrago, G. E., Vilela, B., Aiello-Lammens, M. E., Muscarella, R., Merow, C., & Anderson, R. P. (2018). *wallace: A Modular Platform for Reproducible Modeling of*

- Species Niches and Distributions*. R package. <https://CRAN.R-project.org/package=wallace>.
- Kass, J. M., Vilela, B., Aiello-Lammens, M. E., Muscarella, R., Merow, C., & Anderson, R. P. (2018). Wallace: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. *Methods in Ecology and Evolution*, **9**(4), 1151–1156. <https://doi.org/10.1111/2041-210X.12945>
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Team, R. C., Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., & Hunt, T. (2019). *caret: Classification and Regression Training*. R package. <https://CRAN.R-project.org/package=caret>.
- Leroy, B., Meynard, C. N., Bellard, C., & Courchamp, F. (2016). virtualspecies, an R package to generate virtual species distributions. *Ecography*, **39**(6), 599–607. <https://doi.org/10.1111/ecog.01388>
- Lessmann, S., Stahlbock, R., & Crone, S. F. (2005). Optimizing hyperparameters of support vector machines by genetic algorithms. *Proceedings of the 2005 International Conference on Artificial Intelligence, IC-AI'05*, **1**, 74–80.
- Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, **2**(3), 18–22.
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., Thuiller, W., Wüest, R. O., Zimmermann, N. E., & Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, **37**(12), 1267–1281. <https://doi.org/10.1111/ecog.00845>
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, **36**(10), 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., Raju, B., Shahrzad, H., Navruzyan, A., Duffy, N., & Hodjat, B. (2019). Chapter 15 - Evolving Deep Neural Networks. In R. Kozma, C. Alippi, Y. Choe, & F. C. Morabito (Eds.), *Artificial Intelligence in the Age of Neural Networks and Brain Computing* (pp. 293–312). Academic Press. <https://doi.org/10.1016/B978-0-12-815480-9.00015-3>
- Müller, A. C., & Guido, S. (2016). *Introduction to machine learning with Python : a guide for data scientists*. O'Reilly.
- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (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*, **5**(11), 1198–1205. <https://doi.org/10.1111/2041-210X.12261>
- Naimi, B., & Araújo, M. B. (2016). sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography*, **39**(4), 368–375. <https://doi.org/10.1111/ecog.01881>

- Phillips, S. J. (2017a). *A Brief Tutorial on Maxent*. Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/. (p. 29).
- Phillips, S. J. (2017b). *maxnet: Fitting “Maxent” Species Distribution Models with “glmnet”*. *R Package*. <https://CRAN.R-project.org/package=maxnet>.
- Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, **40**(7), 887–893. <https://doi.org/10.1111/ecog.03049>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**(3–4), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, **31**(2), 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- R Core Team. (2019). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*. <https://www.r-project.org/>
- Radosavljevic, A., & Anderson, R. P. (2014). Making better MAXENT models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography*, **41**(4), 629–643. <https://doi.org/10.1111/jbi.12227>
- RStudio Team. (2018). *RStudio: Integrated Development Environment for R*.
- Schwartz, M. W., Iverson, L. R., Prasad, A. M., Matthews, S. N., & O’Connor, R. J. (2006). PREDICTING EXTINCTIONS AS A RESULT OF CLIMATE CHANGE. *Ecology*, **87**(7), 1611–1615. [https://doi.org/10.1890/0012-9658\(2006\)87\[1611:PEAARO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[1611:PEAARO]2.0.CO;2)
- Thuiller, W., Georges, D., & Breiner, F. (2019). *biomod2: Ensemble Platform for Species Distribution Modeling*. *R package*. <https://CRAN.R-project.org/package=biomod2>.
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). block CV: An r package for generating spatially or environmentally separated folds for k -fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, **10**(2), 225–232. <https://doi.org/10.1111/2041-210X.13107>
- Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. (Fourth Edi). Springer.
- Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**(2), 335–342. <https://doi.org/10.1890/10-1171.1>
- Warren, D. L., Wright, A. N., Seifert, S. N., & Shaffer, H. B. (2014). Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. *Diversity and Distributions*, **20**(3), 334–343. <https://doi.org/10.1111/ddi.12160>
- Wilson, J. D. (1984). Determining a TOPEX score. *Scottish Forestry*, **38**(4), 251–256.

- Young, S. R., Rose, D. C., Karnowski, T. P., Lim, S.-H., & Patton, R. M. (2015). Optimizing deep learning hyper-parameters through an evolutionary algorithm. *Proceedings of the Workshop on Machine Learning in High-Performance Computing Environments*, 1–5. <https://doi.org/10.1145/2834892.2834896>
- Zeng, Y., Low, B. W., & Yeo, D. C. J. (2016). Novel methods to select environmental variables in MaxEnt: A case study using invasive crayfish. *Ecological Modelling*, **341**, 5–13. <https://doi.org/10.1016/j.ecolmodel.2016.09.019>
- Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., ... Merow, C. (2020). A standard protocol for reporting species distribution models. *Ecography*, **43**(9), 1–17. <https://doi.org/10.1111/ecog.04960>

APPENDIX A
R code to reproduce the performance assessment of the genetic algorithm

The output presented in the article has been produced using R version 3.6 in a Linux operating system. Results might be slightly different with other versions of R or operating systems due to possible different random numbers generated after setting the seed.

```

library(SDMtune)
# Set general seed for all experiments
seed = 186546

## Load and prepare data-----
files <- list.files(path = file.path(system.file(package = "dismo"), "ex"),
                    pattern = "grd", full.names = TRUE)
predictors <- raster::stack(files)
p_coords <- virtualSp$presence
a_coords <- virtualSp$absence
data <- prepareSWD(species = "Virtual species", p = p_coords, a = a_coords,
                  env = predictors[[1:8]])
folds <- randomFolds(data, k = 10, seed = seed)

## ANN experiment 1200 hyperparameters-----
# Train starting model ==> size inner layer = number of variables
set.seed(seed)
model_ann <- train("ANN", data = data, size = 8, folds = folds)
auc(model_ann)
auc(model_ann, test = TRUE)

# 1200 hyperparameters' combinations
h_ann <- list(size = 2:81, decay = c(0.01, 0.05, 0.1, 0.3, 0.5),
             maxit = c(100, 500, 1000))
nrow(expand.grid(h_ann)) == 1200 # Make sure there are 1200 combinations

# Genetic Algorithm
om_ann <- optimizeModel(model_ann, hypers = h_ann, metric = "auc",
                       seed = seed)
om_ann@results[1:5, ]
# Grid Search
set.seed(seed)
gs_ann <- gridSearch(model_ann, hypers = h_ann, metric = "auc",
                    save_models = FALSE)
head(gs_ann@results[order(-gs_ann@results$test_AUC), ])

## BRT experiment 1200 hyperparameters-----
# Train starting model
set.seed(seed)

```

```

model_brt <- train("BRT", data = data, folds = folds)
auc(model_brt)
auc(model_brt, test = TRUE)

# 1200 hyperparameters' combinations
h_brt <- list(n.trees = seq(40, 1020, 20), interaction.depth = 1:4,
             shrinkage = seq(0.05, 0.1, 0.01))
nrow(expand.grid(h_brt)) == 1200 # Make sure there are 1200 combinations

# Genetic Algorithm
om_brt <- optimizeModel(model_brt, hypers = h_brt, metric = "auc",
                       seed = seed)
om_brt@results[1:5, ]

# Grid Search
gs_brt <- gridSearch(model_brt, hypers = h_brt, metric = "auc",
                    save_models = FALSE)
head(gs_brt@results[order(-gs_brt@results$test_AUC), ])

## RF experiment 1200 hyperparameters-----
# Train starting model
set.seed(seed)
model_rf <- train("RF", data = data, folds = folds)
auc(model_rf)
auc(model_rf, test = TRUE)

# 1200 hyperparameters' combinations
h_rf <- list(ntree = seq(420, 1000, 20), mtry = 3:6, nodesize = 1:10)
nrow(expand.grid(h_rf)) == 1200 # Make sure there are 1200 combinations

# Genetic Algorithm
om_rf <- optimizeModel(model_rf, hypers = h_rf, metric = "auc",
                      seed = seed)
om_rf@results[1:5, ]

# Grid Search
gs_rf <- gridSearch(model_rf, hypers = h_rf, metric = "auc",
                   save_models = FALSE)
head(gs_rf@results[order(-gs_rf@results$test_AUC), ])

## Maxnet experiment 1200 hyperparameters-----
# Train starting model
bg_coords <- virtualSp$background
data <- prepareSWD(species = "Virtual species", p = p_coords,
                  a = bg_coords, env = predictors[[1:8]])
folds <- randomFolds(data, k = 10, only_presence = TRUE, seed = seed)
model_mx <- train("Maxnet", data = data, folds = folds)

```

```

auc(model_mx)
auc(model_mx, test = TRUE)

# 1200 hyperparameters' combinations
h_mx <- list(reg = seq(0.1, 4.88, 0.02),
            fc = c("l", "lh", "lqp", "lqph", "lqpht"))
nrow(expand.grid(h_mx)) == 1200 # Make sure there are 1200 combinations

# Genetic Algorithm
om_mx <- optimizeModel(model_mx, hypers = h_mx, metric = "auc",
                      seed = seed)
om_mx@results[1:5, ]

# Grid Search
gs_mx <- gridSearch(model_mx, hypers = h_mx, metric = "auc",
                   save_models = FALSE)
head(gs_mx@results[order(-gs_mx@results$test_AUC), ])

```

Table A1: Hyperparameter values used during the performance assessment of the genetic algorithm.

Method	Hyperparameters' values
ANN	size = 2:81, decay = c(0.01, 0.05, 0.1, 0.3, 0.5), maxit = c(100, 500, 1000)
BRT	n.trees = seq(40, 1020, 20), interaction.depth = 1:4, shrinkage = seq(0.05, 0.1, 0.01)
RF	ntree = seq(420, 1000, 20), mtry = 3:6, nodesize = 1:10
ME (Maxnet)	reg = seq(0.1, 4.88, 0.02), fc = c("l", "lh", "lqp", "lqph", "lqpht")

Note: The values are provided using the R code to generate them.

Table A2: Complete list of the environmental variables used in the case example of the Bearded vulture with metric and data source (warm season: May to October; cold season: November to April).

Variable	Abbreviation	Source
Average altitude (m a.s.l.)	altitude_av	DHM ¹
Distance from anthropogenic areas (m)	an_area_d_av	Vector 25 ²
Bush frequency (%)	bush_fr *	Vector 25
Limestone frequency (%)	compact_limestone_fr *	GeoKarten500 ³
Distance from cableways without ski-lifts (m)	c_way_d_av *	Vector 25
<i>Sine</i> of the aspect (-1 to 1)	eastness *	DHM
Forest frequency (%)	forest_fr *	Vector 25
Glacier frequency (%)	glacier_fr *	Vector 25
Gneiss frequency (%)	gneiss_fr *	GeoKarten500
Granite frequency (%)	granite_fr *	GeoKarten500
Ibex density in the warm season (N/ha)	ibex_summer_density_av *	BAFU ⁴ + CSCF ⁵
Ibex density in the cold season (N/ha)	ibex_winter_density_av *	BAFU + CSCF
Distance from lakes (m)	lake_d_av *	Vector 25
Marshland frequency (%)	marsh_fr *	Vector 25
<i>Cosine</i> of the aspect (-1 to 1)	northness_av *	DHM
Open forest frequency (%)	open_forest_fr	Vector 25
Other rock frequency (%)	other_rocks_fr *	GeoKarten500
Grassland and unproductive vegetation frequency (%)	other_soil_fr *	Vector_25 + BfS ⁶
Permanent culture frequency (%)	permanent_fr *	Vector 25
Average precipitation in the cold season (mm)	prec_122_av *	WSL ⁷
Average precipitation in the warm season (mm)	prec_57_av *	WSL
Distance from rivers and creeks (m)	riv_cre_d_av *	Vector 25
Distance from roads and railways (m)	ro_rail_d_av *	Vector 25
Rock frequency (%)	rock_fr	Vector 25
Distance from rock steeper than 45° (m)	rock45_d_sv	Vector 25 + DHM
Standard deviation (SD) of the altitude (m)	roughness *	DHM
Chamois occurrence probability (0 to 1)	rupicapra_hs_av *	CSCF
Scree frequency (%)	scree_fr *	Vector 25
Sheep and goat warm season density (N/ha)	sheep_goat_d_av *	BsF
Distance from ski-lifts (m)	ski_d_av *	Vector 25
Slope (degree)	slope_av	DHM
Frequency of slopes steeper than 30° (%)	slope30_fr	DHM
Average solar radiation in the cold season (WH/m ²)	solar_rad122_av	DHM
Average solar radiation in the warm season (WH/m ²)	solar_rad57_av *	DHM
Average temperature in the cold season (°C)	tave_122_av	WSL ⁷
Average temperature in the warm season (°C)	tave_57_av	WSL
Topographic Position Index ⁸ (index)	tpi_av *	DHM
Distance from water (m)	water_d_av	Vector 25
Average wind speed at 100 m above ground (m/s)	wind_100_av *	BFE ⁹

Note: Variables were calculated as average value within a circular moving window with $r=564\text{m}$ (1km^2) on raster data of 100 x 100m resolution. The 28 variables selected by the *varSel* function are marked using the * symbol.

¹ Digital Height Model of Switzerland (Swisstopo): https://shop.swisstopo.admin.ch/en/products/height_models/dhm25.

² Digital Cartographic Model of Switzerland (Swisstopo): <https://shop.swisstopo.admin.ch/en/products/maps/national/vector/smv25>.

³ Geo Maps: <https://shop.swisstopo.admin.ch/de/products/maps/geology/GK500>.

⁴ Distribution of ibex colonies: <https://www.bafu.admin.ch/bafu/de/home/themen/biodiversitaet/zustand/karten.html>.

⁵ Centre suisse de cartographie de la faune (CSCF): <http://www.cscf.ch/cscf/de/home.html>.

⁶ Federal Administration for Statistic Switzerland (BsF): <https://www.bfs.admin.ch/bfs/de/home/statistiken.html>.

⁷ Federal Institute for Forest, Snow and Landscape Research WSL; available upon request: www.wsl.ch.

⁸ Topographic position index according to Wilson (1984).

⁹ Swiss Wind Atlas (Bundesamt für Energie BFE, 2016).

Table A3: Spearman' correlation coefficient of highly correlated environmental variables ($r_s \leq -0.7$ and $r_s \geq 0.7$). For variable codes see Table A2.

Var1	Var2	r_s
tave_122_av	tave57_av	0.996
altitude_av	tave57_av	-0.996
altitude_av	tave_122_av	-0.995
forest_fr	open_forest_fr	0.993
slope_av	slope30_fr	0.970
riv_cre_d_av	water_d_av	0.925
rock_fr	rock45_d_av	-0.903
roughness	slope_av	0.882
roughness	slope30_fr	0.863
altitude_av	ro_rail_d_av	0.826
ro_rail_d_av	tave57_av	-0.826
northness_av	solar_rad122_av	-0.814
rock45_d_av	slope_av	-0.814
ro_rail_d_av	tave_122_av	-0.814
rock45_d_av	slope30_fr	-0.813
an_area_d_av	ro_rail_d_av	0.803
solar_rad122_av	solar_rad57_av	0.797
an_area_d_av	tave57_av	-0.774
an_area_d_av	tave_122_av	-0.764
altitude_av	an_area_d_av	0.762
altitude_av	scree_fr	0.738
scree_fr	tave57_av	-0.734
scree_fr	tave_122_av	-0.732
open_forest_fr	tave122_av	0.731
open_forest_fr	tave57_av	0.731
altitude_av	open_forest_fr	-0.718
forest_fr	tave_122_av	0.718
forest_fr	tave57_av	0.718
rock45_d_av	roughness	-0.717
rock_fr	scree_fr	0.707
altitude_av	forest_fr	-0.705

Table A4: List of the environmental variables selected by the *varSel* function and relative percent contribution rounded to the first decimal place before executing the *optimizeModel* function.

Variable	P. contribution
bush_fr	0.7
compact_limestone_fr	3.5
c_way_d_av	1.1
eastness_av	0.6
forest_fr	0.8
glacier_fr	0.6
gneiss_fr	1.4
granite_fr	1.3
ibex_summer_density_av	9.8
ibex_winter_density_av	41.3
lake_d_av	0.1
marsh_fr	0.0
northness_av	3.9
other_rocks_fr	0.5
other_soil_fr	1.4
permanent_fr	0.8
prec_122_av	11.2
prec_57_av	1.9
riv_cre_d_av	0.4
ro_rail_d_av	0.1
roughness	0.5
rupicapra_hs_av	2.4
scree_fr	4.2
sheep_goat_d_av	1.4
ski_d_av	1.4
solar_rad57_av	0.9
tpi_av	6.7
wind_100_av	1.0

Table A5: Hyperparameter values used during the hyperparameter tuning experiment with h: number of tested hyperparameter combinations, FC: feature class combinations with linear (l), quadratic (q), product (p) and hinge (h) feature classes, reg: regularization multiplier and iter: number of iterations.

h	FC	reg	iter
75	c("lq", "lh", "lqp", "lqh", "lqph")	seq(0.2, 3, 0.2)	500
150	c("lq", "lp", "lh", "lqp", "lqh", "lqph")	seq(0.2, 5, 0.2)	500
300	c("lq", "lp", "lh", "lqp", "lqh", "lqph")	seq(0.1, 5, 0.1)	500
600	c("lq", "lp", "lh", "lqp", "lqh", "lqph")	seq(0.1, 5, 0.1)	c(500, 700)
1200	c("lq", "lp", "lh", "lqp", "lqh", "lqph")	seq(0.1, 5, 0.1)	seq(300, 900, 200)

Note: the values are provided using the R code to generate them. In the **optimizeModel** function, in order to have consistent results, we set the *seed* argument to 186546 (a randomly generated number).

Table A6: List of the remaining environmental variables after the execution of the *reduceVar* function (Fig. A2) and relative percent contribution rounded to the first decimal place.

Variable	Percent contribution	
	optimizeModel	gridSearch
bush_fr	-	0.9
compact_limestone_fr	7.7	6.8
c_way_d_av	1.7	1.9
eastness_av	1.4	1.4
forest_fr	1.3	1.4
gneiss_fr	1.5	1.7
granite_fr	1.5	1.3
ibex_summer_density_av	19.0	18.5
ibex_winter_density_av	6.8	7.0
northness_av	6.8	6.8
other_rocks_fr	2.2	3.2
other_soi_fr	-	0.8
permanent_fr	0.7	0.7
prec_122_av	24.6	24.2
prec_57_av	2.5	2.4
roughness	1.5	1.0
rupicapra_hs_av	3.5	3.5
scree_fr	9.0	9.9
ski_d_av	1.3	1.3
solar_rad57_av	5.2	3.4
wind_100_av	1.7	1.9

Note: The model parsimony optimization was performed based on the output of the **optimizeModel** and **gridSearch** functions respectively, executed to tune 1200 possible combinations of hyperparameters.

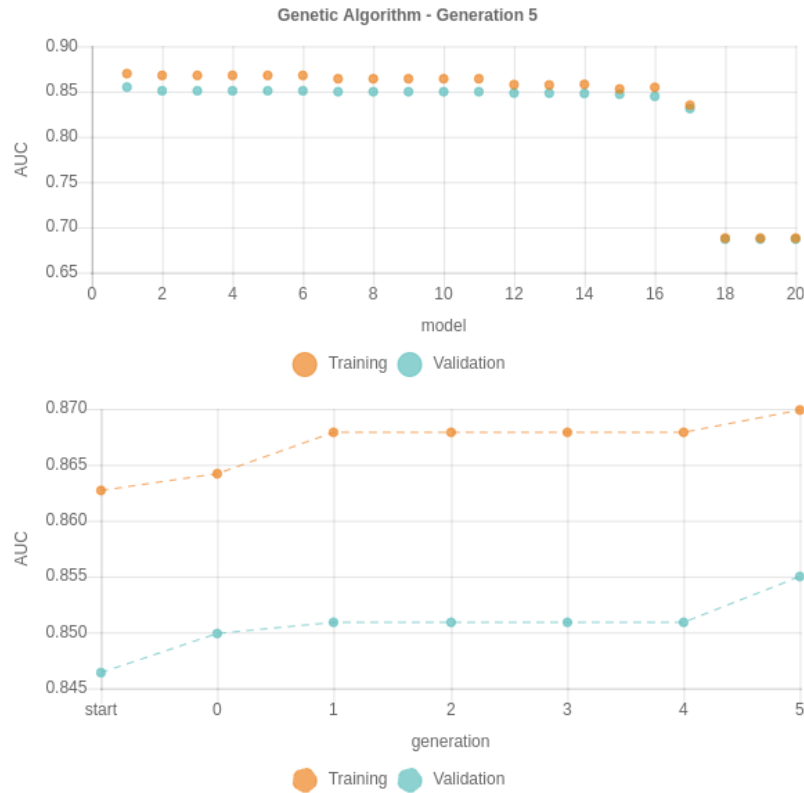


Figure A1. Snapshot of the real-time chart after executing the **optimizeModel** function (with a “population size” of 20 models and five model generations) using 1200 different hyperparameter combinations on the case example data of the Bearded vulture. The scatterplot on top shows the training AUC (in orange) and the validation AUC (in blue) of the 20 ranked models at the end of the fifth generation, given as the arithmetic mean of the 4-folds cross validation. The line plot at the bottom shows the increase in model performance (based on the AUC on both training and validation dataset) at each generation with start: the starting model before running the optimization process, 0: the best performing model after the random population is created, 1-5: the best performing models in each of the five generations of the optimization process.

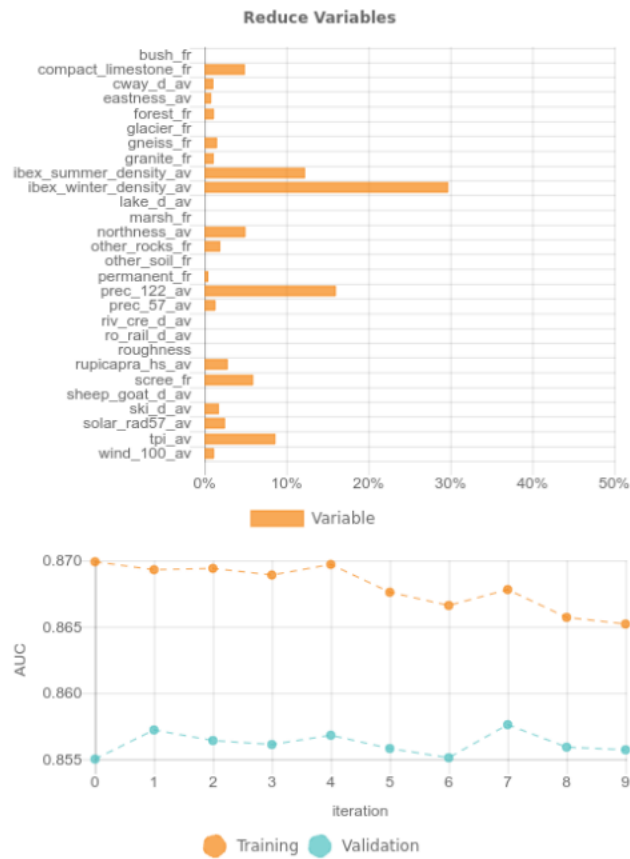


Figure A2. Snapshot of the real-time chart after executing the `reduceVar` function on the case example data of the Bearded vulture. The bar chart on top shows the 28 uncorrelated environmental variables with the percent contribution of the retained environmental variables at the end of the selection process, calculated according to Phillips et al. (2006). The line chart at the bottom shows the change in model performance (based on the AUC on both the training and validation dataset, given as arithmetic mean of the 4-folds cross validation) at each iteration where a single variable is removed. In the RStudio viewer pane the chart is interactive and when the user hovers over the line or bar chart a tooltip reports the variable that has been removed and the relative model performance values. Variable codes are provided in Table A2.

CHAPTER 3



A predictive flight-height model for avoiding future conflicts between an emblematic raptor and wind energy development in the Swiss Alps



Chapter 3 of this thesis is the submitted version of the article to Royal Society Open Science prior to peer review. It has been published by the journal after the PhD defence: <https://doi.org/10.1098/rsos.211041> under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>.

A predictive flight-height model for avoiding future conflicts between an emblematic raptor and wind energy development in the Swiss Alps

Sergio Vignali¹, Franziska Lörcher^{2,3,4}, Daniel Hegglin^{2,3}, Raphaël Arlettaz¹ & Veronika Braunisch^{1,5}

¹ Division of Conservation Biology, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012, Bern, Switzerland

² Stiftung Pro Bartgeier, Wuhrstrasse 12, CH-8003, Zurich, Switzerland

³ SWILD, Wuhrstrasse 12, CH-8003, Zuerich, Switzerland

⁴ Vulture Conservation Foundation, Wuhrstrasse 12, CH-8003, Zurich, Switzerland

⁵ Forest Research Institute of Baden-Wuerttemberg, Wonnhaldestrasse 4, D-79100, Freiburg, Germany

Under review

ABSTRACT

The deployment of wind energy should contribute to the societal shift towards a massive reduction of greenhouse gas emissions. Yet, wind energy and large birds, notably soaring raptors, both depend on suitable wind conditions. Conflicts in airspace use may thus arise between wind energy development and wildlife protection due to the risks of collisions of birds with the blades of wind turbines. Using locations of GPS-tagged bearded vultures, a rare scavenging raptor reintroduced into the Alps, we built a spatially-explicit model to predict potential areas of conflict with future wind turbines deployment in the Swiss Alps. We modelled the probability of bearded vultures flying within the range of rotor-swept heights of wind turbines as a function of wind and environmental conditions, including food supply (wild ungulates presence). Flight activity within the blade-swept heights of wind turbines was generally high, concentrating on south-exposed mountainsides, especially in areas where ibex carcasses have a high occurrence probability, with critical areas covering vast expansions throughout the Swiss Alps. Our model provides a spatially-explicit decision tool that will guide authorities and energy companies for planning the deployment of wind farms without jeopardising the chances of survival of emblematic Alpine wildlife.

Keywords: wind energy, risk mitigation, wildlife-human conflicts, spatial planning, predictive modelling, vulture conservation

INTRODUCTION

The evident negative impacts of global warming on our economy and the biosphere have led many countries to tackle the energy crisis and define objectives for reducing their fossil fuel consumption. Recently, the European Union has launched the “European Green Deal” (European Commission, 2019), a new broad strategy that aims to drastically reduce greenhouse gas emissions and decouple the economic growth from the exploitation of natural resources. The goal of this EU’s initiative is to become “climatically neutral” by 2050 and this will involve, among others, the extensive use of renewable sources of energy. Yet, along with the claim to progressively rely exclusively on clean energy, another main target of this new deal is to preserve biodiversity. If the EU commits to combat biodiversity erosion, it remains to demonstrate that the expansion of renewable energy sources can proceed without jeopardising biodiversity conservation.

Wind energy, together with solar energy, will likely lead the green energy revolution worldwide (Lee & Zhao, 2020). However, its negative effects on biodiversity have been the subject of an intense debate (Arnett et al., 2015; Drewitt & Langston, 2006; Katzner et al., 2019; Leung & Yang, 2012; Northrup & Wittemyer, 2013), known as the “green-green dilemma” (Dulluri & Rat, 2019; Straka et al., 2020; Voigt, 2021; Voigt et al., 2019): even though wind energy can help to reduce greenhouse gas emissions, it may represent a new threat to sensitive wildlife whose protection may in turn hamper the development of this energy sector.

Many of the species most affected by the large-scale expansion of the wind industry are already endangered or threatened, in particular among flying vertebrates such as birds (Drewitt & Langston, 2006; Madders & Whitfield, 2006) and bats (Frick et al., 2017; Wellig et al., 2018). Large soaring diurnal raptors are often the main avian victims of the blades of wind turbines (Barrios & Rodríguez, 2004; Perold et al., 2020; Smallwood & Thelander, 2008; Thaxter et al., 2017). This is particularly concerning as they have a late sexual maturity and a low reproductive rate. Hence, even a slight increase in their mortality rates can exert strong negative impacts on their population dynamics (Beston et al., 2016; Carrete et al., 2009; Schaub et al., 2009; Watson et al., 2018). Wind energy facilities are often erected in regions where landforms and climate generate favourable conditions to support the soaring flight of vultures (Poessel, Brandt, et al., 2018; Rushworth & Krüger, 2014), either via thermal or orographic lifts. The limited frontward visual field of diurnal raptors (Martin et al., 2012), which reduces their ability to perceive obstacles appearing in their direction of movement, further exacerbates collision risks. Moreover, when

soaring in a wind updraught, a raptor follows an ascending spiral (in case of thermals) or an eight-shaped (in case of orographic lifts) (Pennycuick, 1973), which may suddenly expose it to a rotating blade that was still invisible a few seconds ago due to the -frequent change of flight azimuth (Allinson, 2017).

Reconciling the transition towards a genuinely greener energy production thus necessitates a rigorous and strategic planning that satisfies the dual objective of a smooth energetic transition without jeopardising wildlife survival. To prevent any detrimental impacts of the wind industry development on endangered species, wildlife managers and wind energy companies need adequate planning tools to avoid deploying wind facilities in areas where major conflicts with biodiversity preservation will occur.

Different approaches have been used as planning tools to mitigate the risks encountered by flying vertebrates, spanning from mere delineations of buffer areas around sensitive locations (Bright et al., 2008; Janss et al., 2010; Venter et al., 2019), through the compilation of distribution areas of sensitive species (Allinson, 2017; Bright et al., 2008; Mc Guinness et al., 2015), to more complex methods that account for fine-grained habitat use and/or flight behaviour of potentially impacted bat and bird species (Balotari-Chiebao et al., 2018; Murgatroyd et al., 2021; Reid et al., 2015; Tikkanen et al., 2018; Wellig et al., 2018). The first approach is fairly imprecise, with for instance buffer areas created around nesting locations while neglecting habitat selection at other life stages. It is furthermore static, being incapable of accounting for range expansion caused by increases in population sizes of potentially affected species (Braunisch et al., 2015). The second method equates areas of species presence with areas of potential conflict; not accounting for actual fine-grained species-habitat associations it remains coarse but can be valuable for identifying broad areas of potential conflicts. The third method is the most sophisticated and also the most informative one, as spatially-explicit predictive models allows extrapolating to areas for which data about species presence may be deficient. Moreover, when relying on individual-based data such as radio- or GPS-tracking, it enables delineating areas of potential conflict with an unprecedented precision, most notably when providing information about the height above ground at which birds fly. This approach opens the gate towards 3D spatial modelling aimed to mitigate if not avoid conflicts between flying vertebrates and future wind facilities development.

The aim of this study was to predict areas of the Swiss Alps where bearded vultures (*Gypaetus barbatus*) are likely to fly within the critical height range that is typically swept by the blades of modern wind turbines. The bearded vulture is a long-lived scavenger listed as vulnerable

in Europe (BirdLife International, 2015b). It is still critically endangered in Switzerland (Keller, Gerber, et al., 2010). Extirpated from many European countries in the early twentieth century (Mingozzi & Estève, 1997), the species has been reintroduced into the Alps since the 1980s, with a steadily growing population that progressively recolonises its former historical range (Hirzel et al., 2004). Several cases of collisions (including fatalities) with aerial anthropogenic structures have been reported in this re-established population (Izquierdo, 2017; Lörcher & Hegglin, 2020), including with wind turbines which may represent a new major source of hazard into the future (Vignali et al., 2021). In effect, Schaub et al. (2009) have shown that even a slight increase in mortality would push the Alpine population of bearded vultures below demographic self-sustainability. If a spatial model we have recently developed predicts species' potential distribution, including future expansion across the Swiss Alps (Vignali et al., 2021), the present model adds a vertical dimension to these projections. In effect, such a model would refine the prediction of potential conflicts, as the actual use of the airspace, i.e. the flight height with respect to the blade-swept range is accounted for.

Starting from a large dataset of GPS locations collected from tagged individuals, we thus modelled the probability that bearded vultures fly within the critical, blade-swept height range of wind turbines and identified the environmental and topographic variables that drive flight height selection. The model was projected to the entire Swiss Alpine range and combined with the previously modelled potential distribution of the species (Vignali et al., 2021) in order to show the joint probability of bearded vultures flying at risky heights within suitable habitat. The resulting map provides useful spatial information to delineate areas where the species would be at risk of colliding with wind turbine blades and therefore represents a useful decision tool for planning the deployment of wind power plants across the Swiss Alpine range while minimising their potential impacts on emblematic biodiversity.

METHODS

Study area and environmental variables

We modelled the flight height of bearded vulture across the entire Swiss Alpine range, defined as four of the six biogeographical regions of Switzerland (Gonseth et al., 2001): Northern Alps, Inner Western Alps, Inner Eastern Alps, and Southern Alps. We used environmental variables that represent land cover characteristics, geology, topography, food availability and wind conditions

(Table 1). Land cover information was extracted from the digital cartographic model of Switzerland (Vector25, <https://www.swisstopo.admin.ch/en/geodata/maps/smv/smv25.html>). This vector layer was converted into a raster map with 25 m spatial resolution and reclassified to represent the following ten classes: orchards, forest, bush, scree, anthropic areas, marshland, water, rock, glacier, and remaining areas not included in the other classes (Table S1). The geological features were derived from the simplified geotechnical map of Switzerland which was provided as digitised vector map by the University of Bern (<https://biblio.unibe.ch/maps/bis/publications/dl-oef21.html>) and represents the types of the topmost rock strata (<https://data.geo.admin.ch/ch.swisstopo.geologie-geotechnik-gk200/> (Bundesamt für Statistik, 1967)). The shapefile was converted into a raster map with 25 m spatial resolution and reclassified into four classes: areas dominated by limestone, granite, gneiss, and remaining geological substrates (Table S2).

Table 1: Environmental predictors used to model the probability of bearded vultures flying below 200 m a.g.l. (i.e. within the flight height range swept by wind turbine blades) across the Swiss Alps, with indication of unit of measurement, abbreviation, and data source.

Category	Description	Unit	Abbreviation	Source
Land cover			landcover	Vector 25 ^a
Geology			geology	gk200 ^b
Topography	<i>Sine</i> of the aspect	-1 to 1	eastness	DHM25 ^c
	<i>Cosine</i> of the aspect	-1 to 1	northness	DHM25
	Slope	degree	slope	DHM25
	Slope unevenness	Index	slope_unev	DHM25
	Topographic Position Index ^d	Index	tpi	DHM25
Food	Ibex occurrence probability	0 to 1	ibex	Vignali et al. (2021)
	Chamois occurrence probability	0 to 1	chamois	Vignali et al. (2021)
Climate	Average wind speed at 100 m above ground	m/s	windspeed	BFE ^e

^a Digital Cartographic Model of Switzerland: <https://www.swisstopo.admin.ch/en/geodata/maps/smv/smv25.html>.

^b Simplified geotechnical map of Switzerland (Bundesamt für Statistik, 1967).

^c Digital Height Model of Switzerland: <https://www.swisstopo.admin.ch/en/geodata/height/dhm25.html>.

^d Topographic position index according to Wilson (1984).

^e Swiss Wind Atlas (Bundesamt für Energie BFE, 2016).

Topography was characterised with five raster layers extracted from a digital elevation model with a spatial resolution of 25 m (DHM25, <https://www.swisstopo.admin.ch/en/geodata/height/dhm25.html>). The aspect of the study area was represented by the deviation from east and north (*sine* and *cosine* of aspect, respectively). Terrain characteristics were incorporated by using the Topographic Position Index (TPI, Wilson, 1984) and the slope unevenness, which describe the elevation or slope of a cell relative to the surrounding terrain, respectively (both calculated within a moving window of nine pixels). Northness and eastness were calculated with ArcGIS 10.2, TPI and slope unevenness were derived using the *raster* package in R (Hijmans, 2019). Food availability was described using the modelled probability of chamois and ibex occurrence, the two main providers of bones for bearded vultures, which thus served as a proxy for food supply (for methodological details see Vignali, Lörcher, Hegglin, Arlettaz, & Braunisch, 2021, Appendix A). Finally, average wind speed at 100 m a.g.l. was extracted from Swiss Wind Atlas (Bundesamt für Energie BFE, 2016). Pairwise Spearman's correlations between all continuous environmental variables were $|r_s| < 0.6$, calculated based on 10,000 random locations. Categorical variables (i.e. land cover and geology) were one-hot encoded while continuous variables were normalised using the mean and standard deviation derived from the training dataset.

Species data and data processing

Between 2005 and 2020, as part of the Alpine reintroduction programme, 97 bearded vultures have been equipped with GPS loggers (battery or solar-powered) fitted with a leg loop harness (Hegglin et al., 2004). All birds but one were tagged as fledglings, 81 thereof captive-bred and 16 wild-hatched. In addition, one adult bird, released in 1999, was tagged in 2017 after recapture, rehabilitation and re-release. Loggers from different manufacturers and relying on various power sources were deployed whilst GPS locations were collected with a very heterogenous schedule. For example, some devices were programmed to collect bursts with high frequency resolution (1 Hz) as long as the bird was moving and the battery was sufficiently charged. Others collected GPS locations at 1-min resolution under similar conditions, while some devices recorded data with even lower temporal resolution. Since we were interested in modelling the flight height above ground level, we selected only data collected by GPS devices that simultaneously recorded information on both flight altitude and instantaneous ground speed so that non-flight locations could be excluded from the analysis (see below). Following the definition of Péron et al. (2020) we define “flight

height” as the distance between the bird and the ground level, and “flight altitude” as the distance relative to a reference surface like the earth ellipsoid or the mean sea level. Some of the devices provided flight altitude estimates relative to the mean sea level while others measured it relative to the earth ellipsoid. In the latter case, flight altitude measures were converted to altitude relative to the mean sea level using the method described by Poessel et al. (2018, Appendix S1). The altitude relative to the mean sea level was then used to estimate the flight height above the ground level by subtracting the ground elevation extracted from the digital elevation model at each GPS location.

Several authors described the problem of negative flight height values estimated from GPS locations (see for example Katzner et al. (2012); Péron et al. (2017, 2020); Poessel et al. (2018)). Negative flight height values are essentially due to the sum of errors in the measure of the altitude and/or position provided by the GPS tag and in the interpolation of the digital elevation model used to calculate the flight height. Visual inspection of GPS locations collected at 1 Hz resolution showed that most negative flight height values occurred close to steep slopes suggesting a significant influence of the position error in generating negative values. To reduce the position error we culled our data by removing observations with a Horizontal Dilution of Precision (HDOP) ≥ 10 (when the HDOP was provided), which correspond to an error of about 30 m (Katzner et al., 2012) or by discarding all locations with an error ≥ 30 m using the position error provided by the manufacturer. Moreover, we retained only locations with a flight height within the range of -50–4000 m (Poessel, Brandt, et al., 2018), assuming that values outside of this range were probably generated by an erroneous measure of the flight altitude.

In order to ensure that the locations retained for our analysis were all collected from flying vultures we considered a combination of two criteria. First, we selected only GPS locations recorded during the day, from sunrise to sunset, using the R package *suncalc* (Thieurmél & Elmarhraoui, 2019). Second, we removed all locations whose instantaneous ground speed was less than 2 m/s (Poessel, Brandt, et al., 2018). This second criterion might have removed some valid flying positions, but we preferred to be conservative and avoid the risk of including non-flight locations. Finally, for each bird we selected only observations collected within the Swiss Alpine range, removed all GPS locations recorded during the first eight weeks after fledging to reduce a potential bias related to the release event, and randomly sampled one observation per minute in the case of bursts collected at 1 Hz resolution. This last step was necessary to reduce potential autocorrelation among locations and avoid an overrepresentation of the vultures that collected high temporal resolution data.

Table 2: GPS-tagged birds included for modelling the flight height of bearded vultures in the Swiss Alps with mention of the country of first release (or subsequent recapture), origin (C: captive-bred; W: wild-fledged), year of fledging, sex (M: Male; F: Female; U: Unknown), manufacturer of the transmitter, number of GPS locations retained after data cleaning (N), total number of tracking days within the Swiss Alpine range, percent of locations below 200 m a.g.l. (%), and number of GPS locations retained after randomly subsampling one location per minute (S).

Bird ID	Country	Origin	Year	Sex	Manufacturer	N	Days	%	S
BG1071	CH	C	2020	F	e-obs	4,855	102	82.92	1,667
BG1068	CH	C	2020	M	e-obs	17,036	85	92.14	1,948
BG1003	CH	C	2018	F	e-obs	831,178	664	84.00	24,356
BG1001	CH	C	2018	M	e-obs	537,155	591	77.22	18,244
BG964	CH	C	2017	M	e-obs	959,434	689	74.83	27,357
BG899	CH	C	2016	M	Microwave	40,771	600	75.95	39,961
BG900	CH	C	2016	M	Microwave	17,330	292	73.04	17,006
BG841	CH	C	2015	F	Microwave	7,263	372	69.95	7,113
BG838	CH	C	2015	F	e-obs	1,386	339	80.38	1,358
BG802	CH	C	2014	M	Microwave	39,759	1,308	73.20	39,112
BG797	CH	C	2014	M	Microwave	12,302	985	74.69	12,162
BG321*	CH	C	1999	F	Ornitela	3,764	263	58.95	671
BG1031	FR	C	2019	F	Ornitela	3,358	32	58.87	421
BG1022	FR	C	2019	M	Ornitela	2,782	1	70.13	77
BG980	FR	C	2018	M	Ornitela	9,779	26	63.43	472
BG983	FR	C	2018	M	Ornitela	5,789	13	72.62	268
BG905	FR	C	2016	M	e-obs	4,001	5	73.41	255
W361	FR	W	2020	U	Ornitela	4,710	19	67.56	303
W356	FR	W	2020	U	Ornitela	170	44	84.71	170
W346	FR	W	2020	U	Ornitela	20,172	15	81.60	666
W285	FR	W	2019	F	Ornitela	139,526	173	74.28	4,734
W284	FR	W	2019	F	Ornitela	29,788	58	60.37	1,306
W313	FR	W	2019	F	Ornitela	11,563	10	68.36	345
W251	FR	W	2018	M	Ornitela	14,482	109	61.89	840
W209	FR	W	2017	M	Ornitela	99,304	137	71.68	3,594
W196	FR	W	2016	F	Ornitela	46,655	986	83.41	12,056
BG998	AT	C	2018	M	Ornitela	44,431	277	76.93	3,152
BG843	AT	C	2015	M	e-obs	31,206	311	58.60	7,214
BG840	AT	C	2015	M	e-obs	407	76	74.55	407
W349	IT	W	2020	F	Ornitela	78	13	75.64	77

* Tagged as adult bird in 2017

Modelling approach

Tracking animals with GPS devices has expanded over the last years and new generation loggers are able to collect many locations at fine temporal resolution (López-López, 2016). Processing a large amount of data is challenging and computationally expensive. Often data are heavily subsampled not only to reduce autocorrelation problems but also meet computational capacities of classical statistical approaches (Murgatroyd et al., 2021; Poessel, Brandt, et al., 2018). On the other hand, more recent techniques like machine learning algorithms, and especially artificial neural networks, can capture complex nonlinear relationships present in the data but they require large datasets. In this regard many tools have been developed to speed up computation with graphics processing unit GPU acceleration and create data pipelines to efficiently pre-process data before model training. Artificial neural networks gained popularity in many fields of biology during the last decade, including behavioural classification from tri-axial acceleration data (Nathan et al., 2012; Resheff et al., 2014) or from GPS data (Browning et al., 2018; Maekawa et al., 2020). In order to make use of all information included in the data and also develop a method that easily scales to potentially very large datasets we used a deep feedforward neural network to model the probability of a bearded vulture flying within a given height-range at a given location. Considering the still ongoing trend of increasing heights in newly constructed, modern wind turbines, we decided for a threshold of 200 m (hereafter referred to as critical height), below which the flight of a bird is deemed to be at potential risk of collision with the rotor blades (see also Murgatroyd et al., 2020; Reid, Krüger, Whitfield, & Amar, 2015). The flight height was converted to a binary response with 1 being a location within the critical height range. Our model was defined and trained within the tensorflow framework (Martín in Abadi et al., 2016) and using the *keras* R package (Allaire & Chollet, 2020). The Keras application programming interface (API) allows great flexibility in defining the architecture of a neural network. We used two hidden layers connected by a dropout layer and a single unit as output of the network that used a sigmoid activation function (for model implementation, see R code in supplementary material). A dropout layer acts as a regularization layer by randomly deactivating some units during training, thus reducing the risk of overfitting the training data (Srivastava et al., 2014). The model was trained to minimise the binary cross entropy loss function using the Adam optimiser.

During the modelling process, we first conducted a grid search experiment to identify the best model architecture, varying the number of units in the hidden layers independently from 16 to 512, each time doubling the number of units (i.e. 16, 32, 64, 128, 256 and 512), and searching the

rate of the dropout layer in the range 0.2 – 0.7 with increments of 0.1, which resulted in a total of 216 different model configurations. We trained the model on 70% of the data, used the remaining 30 % for validation and stopped model training when the area under the operating characteristic (ROC) curve (AUC) (Fielding & Bell, 1997) computed for the validation dataset did not increase for more than 10 epochs.

In a second step we used a cross-validation approach to evaluate the ability of the model to generalise among different individual birds and for different zones of the study area. Using the best model configuration, identified with the random search experiment, we trained 30 different models for 20 epochs, each time leaving out the locations collected from a different bird on which the model predictions were then evaluated. Similarly, we ran a spatial block cross validation dividing the study area into spatial blocks created with the *blockCV* R package (Valavi et al., 2019). The analysis of the spatial autocorrelation among continuous variables, conducted using the function *spatialAutoRange*, suggested a block size with a minimum side-length of 5719 m. We decided for 10 km (Fig. S1) so as to verify the ability of the model to generalise across wider areas. Using the blocks we randomly partitioned the GPS locations within them, into five cross-validation folds.

As a third step we investigated the contribution and the marginal effect that each environmental variable had on the model predictions. The contribution of different variables was estimated via their permutation importance using the *vip* R package (B. M. Greenwell & Boehmke, 2020) and measuring the drop in AUC, while the marginal effect of the environmental variables was investigated using individual conditional expectation (ICE) (Goldstein et al., 2015) and partial dependence (PD) (Friedman, 2001) plots created with the *pdp* R package (B. M. Greenwell, 2017). ICE curves are generated for a given variable in the dataset, the range of which is subdivided into a grid of n equally spaced values. For one observation in the dataset, predictions are made by varying the focal variable within the grid while the other variables are kept constant, thus creating a single ICE curve. The process is repeated for each observation, generating as many ICE curves as there are observations in the dataset. This is a good method to show complex interactions among variables while the overall effect is shown by the PD curve, which simply represents the average of all ICE curves.

Finally, we evaluated model uncertainty by means of a bagging procedure (Shu & Burn, 2004). We sampled the training dataset (70% of the locations) 30 times, with replacement, and used the validation dataset to stop model training when the validation AUC didn't increase for more than 10 epochs. We then used the 30 trained models to project predictions to the full extent

of the study area and used the mean of the 30 produced maps as final prediction. A 95% credible interval was also created to identify areas where model predictions are more uncertain. The full extent prediction was visually evaluated by field experts (R.A., D.H. and F.L.) to verify that known areas where the species flies close or far from the ground were correctly identified by the model.

Conflict map for landscape planning

The map obtained from the above modelling represents the probability of a bearded vulture flying within the critical height range swept by turbine blades, this regardless of whether the habitat conditions are suitable for the bearded vulture, i.e. independent of the probability of species occurrence. To identify areas within actual species' suitable habitat in which there would exist a risk of collision in case of wind turbine installation, we combined the output of our previously developed habitat suitability model (see Vignali et al., 2021, figure 4e, reported also in Fig. 3c) with the output of the model described in this article. The joint probability of species occurrence and flying within the critical height range (Fig. 3e) was calculated by taking the product of the two raster maps (Reid et al., 2015).

We also converted the predicted probabilities of a bearded vulture flying at risky heights into a binary map by means of the threshold which held a sensitivity of 95%. The resulting binary map was then intersected with the potential conflict map described in Vignali et al. (2021, Fig. 4d, reported also in Fig 3d) in order to delineate the areas within the habitat where the species flies within the critical height range which would be particularly prone to collisions (Fig. 3f). Hereafter this synthetic map is referred to as high-risk conflict map.

The whole analysis was run in R (R Core Team, 2020) version 4.0.2 through the RStudio software (RStudio Team, 2021).

RESULTS

Tracking data

Among all bearded vultures GPS-tagged by the Alpine reintroduction programme, 32 individuals fulfilled the requirements for being included in the analysis. Two out of these 32 individuals also had to be discarded because they yielded only five and nine GPS locations in Switzerland, respectively. The retained 30 bearded vultures had been tagged in all four countries involved in the Alpine reintroduction project (i.e. Switzerland, France, Italy, and Austria), with the data used for

the analysis having been collected from September 2014 to the end of December 2020. The number of collected locations, as well as the amount of tracking days, varied significantly among tagged individuals, with larger sample sizes in birds released within Switzerland, and lower sample sizes in birds that only occasionally visited the study area, stemming from release sites in the neighbouring countries (Table 2). The number of tracking days within the Swiss Alpine range varied from 1 to 1308 per individual while the duration of the tracking period per individual varied according to the lifetime of the solar-battery system, any device loss or deficiency, or in case of a bird's death. Some devices (N=9) recorded data for more than 3 years. This is approximately the age at which juvenile bearded vultures change their behaviour, shifting from an exploring phase to a phase of territory establishment (Hirzel et al., 2004; Vignali et al., 2021). A total of 2,939,411 GPS locations were retained after data cleaning, of which 77.5% were collected below 200 m a.g.l. (average proportions varying between individuals from 58.6% to 92.1%). After applying the above subsampling procedure, flight height was finally modelled based on 227,313 GPS locations.

Model architecture and predictions

The best model configuration identified during the grid search experiment was a deep feedforward neural network with 256 units in the first hidden layer, a dropout rate of 60%, and 32 units in the second hidden layer. This model had an AUC value of 0.726 for the training dataset and 0.719 for the validation dataset. Overall, the model was able to generalise well across birds, which was indicated by a mean training and testing AUC of 0.716 (SD=0.003) and 0.721 (SD=0.045), respectively (Table S3), which was comparable to the performance of the model trained using all birds. Similarly, the model showed a good ability to generalise across different regions of the study area with a mean training and testing AUC of 0.719 (SD=0.005) and 0.710 (SD=0.016), respectively (Table S4).

The combination of the potential conflict map (Fig. 3d) with binary representation of the probability of flying within the critical height below 200 m a.g.l. (Fig. 3b) revealed that about 77% of the area suitable for the species is likely to be overflowed within the critical height range (Fig. 3f). This area, ranging from 278 to 4502 m a.s.l., represents 30.6% (7871 km², Table 3) of the overall extension of the Swiss Alpine massif.

Table 3: Percentage of shares of the different levels of sensitivity (increasing from 1 to 4) predicted in the potential conflict map (Fig. 3e) and the high risk conflict map (i.e. where the bearded vulture is likely to fly within the critical height range, i.e. below 200 m a.g.l., Fig. 3f) in the whole Swiss Alps. The last column reports the shares of areas where the bearded vulture is likely to fly above the critical height within the habitat.

Level of sensitivity	Potential conflict map	High risk conflict map	Remaining
1	12.3	8.2	4.1
2	11.0	8.5	2.5
3	7.2	5.8	1.3
4	9.2	8.0	1.2
Total	39.7	30.6	9.2

Relative contribution of different variables

The environmental conditions that mainly drove the probability of a bearded vulture flying within the critical height range were steepness of the terrain and food availability (permutation importance of 29.6 and 19.7%, respectively) (Fig. 1).

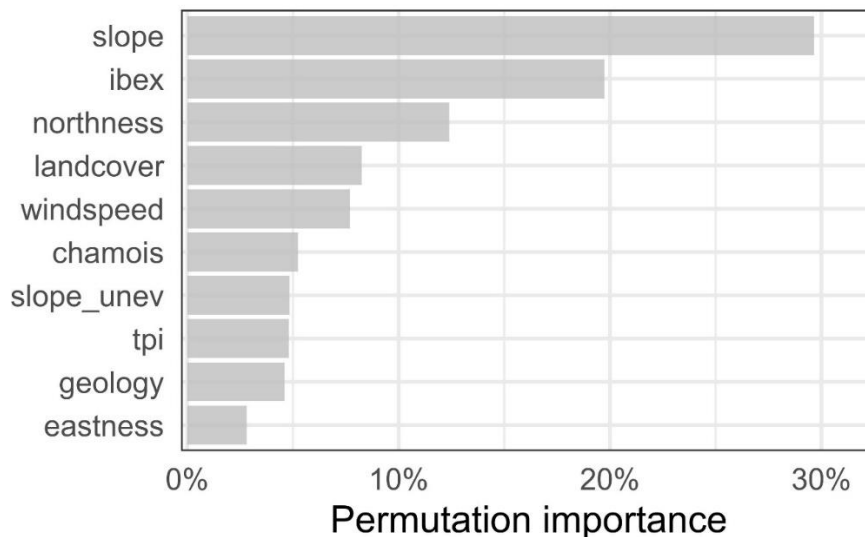


Figure 1 Permutation importance of the environmental variables used to model the probability of bearded vultures flying below 200 m a.g.l. Permutation importance is presented as the drop in training AUC (%) when randomly permuting the values of the respective variable within their empirical range. Variable abbreviations are given in Table 1.

Bearded vultures were more likely to fly at lower height (<200 m a.g.l.) not only when approaching steeper slopes but also in areas with a high probability of ibex presence (Fig. 2), i.e. sectors where it is more likely to find ibex carcasses. This pattern is evidenced not only by the PD curves, but also by the increasing concentration of the ICE curves with increasing values of the two variables. Overall, the probability of flying within the critical height range was lower over north-facing mountainsides than over south-facing mountainsides and higher over areas dominated by scree, rocks, and glaciers compared to the remaining land cover conditions (i.e. forest, anthropic areas, water bodies, etc). Flying within the critical height range was also more likely to occur in areas typically exposed to stronger winds compared to areas with weaker winds.

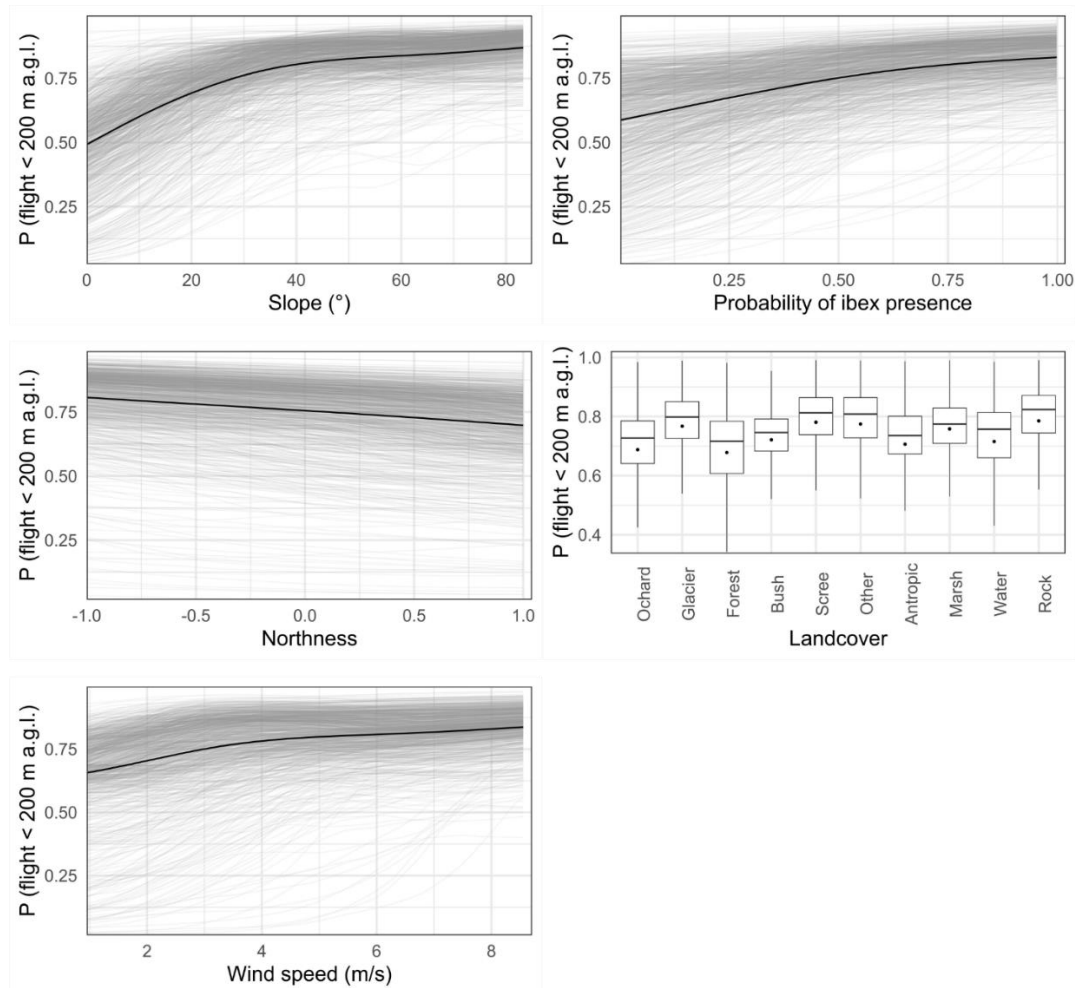


Figure 2 Marginal effect of the five most important environmental variables for predicting the probability of a bearded vulture flying below 200 m a.g.l. In grey are plotted 1000 randomly sampled individual conditional expectation (ICE) curves (Goldstein et al., 2015) and in blue the partial dependence (PD) curve (Friedman, 2001). For land cover, a categorical variable, each boxplot shows the ICE values without outliers and the blue dot the value of the PD. The curves for the remaining environmental variables are shown in Fig. S4.

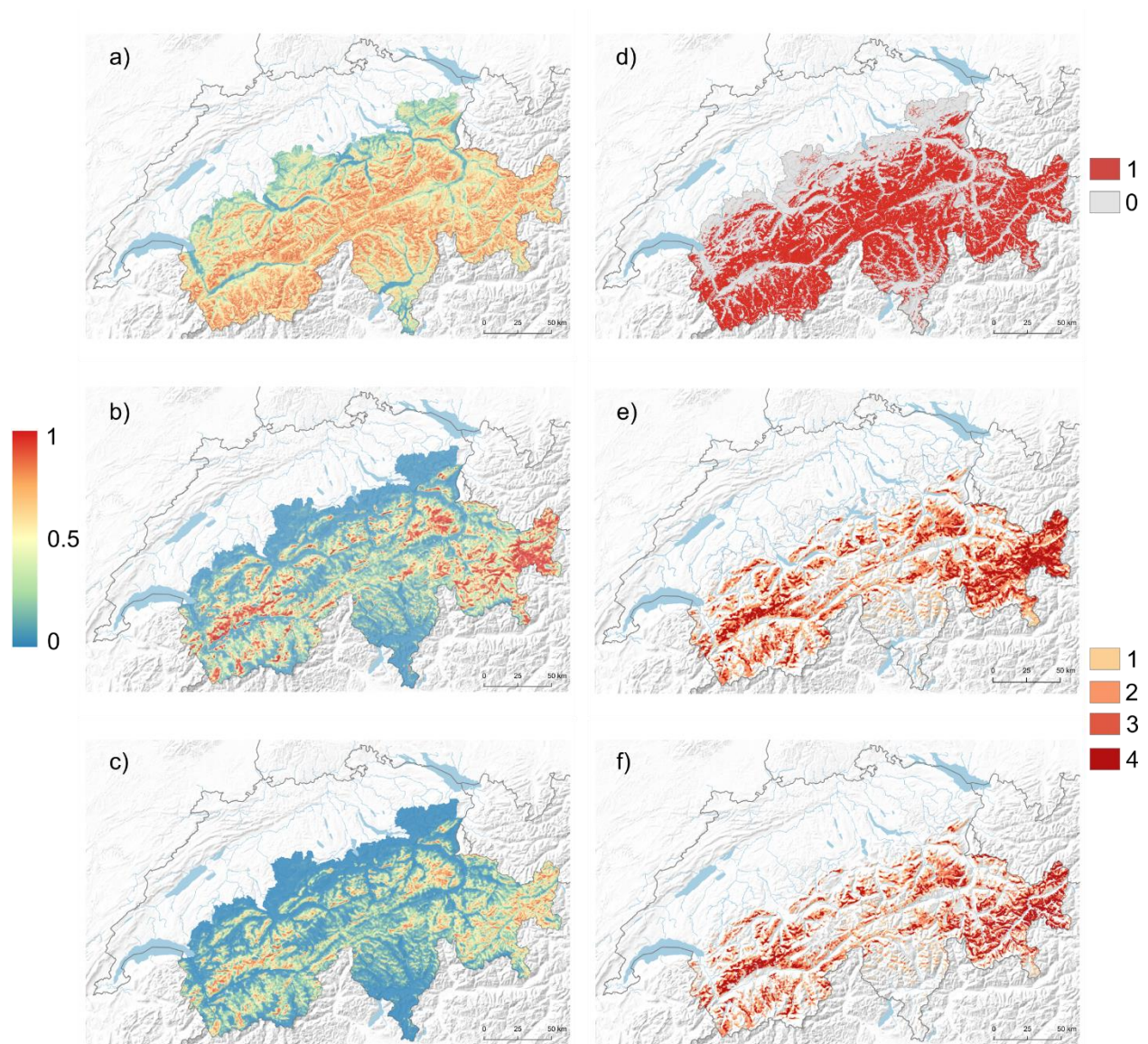


Figure 3 Predictions of bearded vulture occurrence and the probability of the species flying within the risky height range (< 200m a.g.l.) that potentially generate risks of collisions with the blades of wind turbines in the Swiss Alps. The maps show a) the predicted probability of a bearded vulture flying below 200 m a.g.l. calculated as the mean prediction of a 30-bagging procedure and extrapolated to the whole Swiss Alpine range (gradient from blue, zero probability; to red, high probability); b) the probability of bearded vulture occurrence described in Vignali et al. (2021, Fig. 4e); c) joint probability of occurrence and flying below 200m a.g.l., calculated as the product of map a and b. The map showed in d) is the translation of a) into a binary response using the threshold for which 95% of the locations occurring at risky heights are correctly predicted (the areas with a high probability that a bearded vulture flies within the critical height range are shown in red); e) the “potential conflict map” described in Vignali et al. (2021, Fig. 4f); f) the “high-risk conflict map” calculated as the product of d and e.

DISCUSSION

The wildlife vs wind energy conflict model developed here extends commonly applied approaches of predicting areas of potential collision risk with wind turbine blades based on mere species' spatial occurrences by adding a vertical dimension. In effect, its predictions are refined in the sense that it specifically predicts in which areas bearded vultures would effectively fly within the vertical segment of airspace swept by the blades of wind turbines if deployed. As the majority of the GPS locations (77.5%) indicated flight activity of bearded vultures at heights lower than 200 m a.g.l. (Fig. S3), bearded vultures in the Swiss Alps seem to be active most of the time in the dangerous height range. This concurs with former findings by Rushworth & Krüger (2014) who estimated that south African bearded vultures spend 74.7% of their foraging time below 200 m a.g.l. As a result, our final high-risk conflict map shows that 76.9% of the area of suitable habitat across the Swiss Alps (i.e. 30.5% of the entire Swiss Alpine range) may incur some potential collision risk (Fig. 3f, Table 3).

Steep south-facing slopes and areas with high probability of ibex presence, i.e. providing carcasses potentially exploitable by bearded vultures, offered the best conditions for low-height flight (Fig. 1-2). Terrain steepness and exposition are indeed two key factors explaining the formation of updraughts, i.e. both thermals and orographic uplifts. Thermals are generated by unequal heating of the earth surface: solar radiation heats up certain land cover types faster than others (e.g. dark rocky outcrops), thus generating columns of warm air that rise from the ground. Orographic uplifts stem from the upward deviation of wind blowing against mountainsides. Although the former occur mostly along sun-exposed rocky slopes and over flat areas (Shamoun-Baranes et al., 2003), the latter are generated exclusively along slopes and ridges. While thermals in lowlands are furthermore easily disrupted by dominant winds, mountain regions supply a year-round source of uplifts, thus providing optimal conditions to support the flight of raptors exhibiting high wing-loading such as vultures (Shepard & Lambertucci, 2013). In line with this, various studies have showed that raptors fly relatively low over ridges and steep slopes (Hanssen et al., 2020; Katzner et al., 2012; Murgatroyd et al., 2021) where orographic uplift is more likely to occur. As shown here, bearded vultures are no exception to that rule.

While static and/or dynamic environmental conditions are frequently included to model flight behaviour in relation to potential collision risk with wind turbines (Hanssen et al., 2020; Katzner et al., 2012; Murgatroyd et al., 2021; Péron et al., 2017; Reid et al., 2015), the spatial distribution of potential food supply has – to the best of our knowledge - never been considered so far. Our model ranked this environmental variable as the second most important one in explaining

the probability of flying within the critical height range. Vultures, and large soaring birds in general, are constrained in their movements by the availability of updraughts. By soaring into ascending air currents for gaining height and then gliding to another location (Pennycuick, 2008), they achieve an energy-efficient commuting pathway (Duriez et al., 2014; Ruxton & Houston, 2004). However, the optimal height to be gained with soaring is likely to result from a trade-off between mobility for long-range horizontal displacement and the ability to inspect the ground for locating food, the latter diminishing with height (Shepard et al., 2011). This might be particularly crucial for bearded vultures which, given their peculiar bone-based diet (Margalida et al., 2009; Margalida & Villalba, 2017), have to locate carrion and small parts of carcasses that may be easily overseen. Ibex carcasses represent the most important food supply of Alpine bearded vultures, to an extent that their distribution is largely driven by the presence of that ungulate species (Hirzel et al., 2004; Vignali et al., 2021). Our model could even capture that ecological requirement since the probability of flying within the dangerous height range was always high over areas with high probability of ibex presence, regardless of the other environmental conditions (Fig. 2). That a similar pattern was not found for chamois, another important food source for Alpine bearded vultures, is not surprising because chamois' distribution is much more uniform as the species is less dependent on a rocky substrate than ibex.

We used two different validation approaches to assess the ability of our model to generalise across individuals and geographic regions within the Swiss Alps. In this respect, the leave-one-bird-out cross validation is important to ensure that model predictions are not biased by individuals with a prevailing number of GPS bearings. Only three birds out of 30 (i.e. W361, BG1022, and BG980, Table S3) caused some marked drop of the AUC when used to evaluate the model. These birds had all been tagged in France, visiting only occasionally the southwestern most part of the study area. These notwithstanding, a few birds yielded GPS locations exclusively from one part of the study area (e.g. only in western, central, or eastern Swiss Alps), while most of them roamed across the entire Swiss Alpine massif. The similarly good AUC values obtained across the individuals confirm the reliability of our model extrapolations to the entire study area. Finally, the spatial block cross validation further reinforced the validity of the model throughout the study area, given that it performed equally well for predicting the locations in the spatial blocks not used for model training (Telford & Birks, 2009).

The AUC values yielded by validation were not particularly high, although comparable with those obtained by Reid et al. (2015) for the bearded vulture population inhabiting southern

Africa. A model with an AUC greater than 0.7 is considered to be sufficiently accurate for discriminating positive from negative classes (i.e. distinguishing between locations where bearded vultures flew below vs above the critical height range) (Hosmer & LemeshowStanley, 2000). Since bearded vultures may fly both above and below the critical height range within any given place, there will always be situations where the model correctly predicts one class and therefore incorrectly predicts the other, resulting in low AUC values. In addition, there are plenty of situations for a mountain raptor, especially near sheer cliffs and steep slopes, where a small horizontal displacement can significantly change the height a.g.l., and thus the respective flight height class. Similarly, even tiny inaccuracies in GPS bearings collected in the vicinity of cliffs can affect the allocation to one of the two height flight classes. Therefore, if the AUC remains a valid threshold-independent metric useful to evaluate the overall model performance – which is the reason why we relied on it to tune the model’s hyperparameters and to check the ability of the model to generalise across birds and areas – one always benefits from a final, complementary validation by visually inspecting how the model performs in known areas where the birds have been regularly observed. Finally, a bagging procedure was carried out, which showed that model predictions remained stable over repetitions (Fig. S5).

The projected map showing the probability of bearded vultures flying below the critical height range swept by wind turbine blades (Fig. 3a) revealed that wide areas of the Swiss Alps are potentially prone to collisions. The probability of flight at risky heights was particularly high along mountainsides and ridges. This is not surprising since valleys are overflown at high elevation, notably during commuting relocations. While this probability map gives a general overview of the areas with environmental conditions favouring low flight heights throughout the Swiss Alps, only the map resulting from the joint probability of species occurrence (Braunisch et al., 2015) and of flying below the critical height range (Fig. 3e) encompasses the whole complexity of the species-habitat associations, including ecological requirements and flight behavioural routines. By intersecting these two probabilistic maps, we could filter out areas within the species’ habitat extension where it is unlikely that bearded vultures would fly within the critical height range, i.e. mainly the valley bottoms, while we managed to highlight critical hotspots of potential conflict with wind energy development.

Selecting a threshold to convert a probability map into a binary map is always a critical step because it eventually determines the classification skills of the model. The Youden index, maximising sensitivity plus specificity, is probably the most frequently used threshold approach

for classifying species distribution models (Canran Liu et al., 2013). It has already been applied to flight height models (Murgatroyd et al., 2021). However, when the conservation of an endangered and vulnerable species is at stake, the ability of a model to correctly predict the risk-class is the most important aspect, even if it comes at the expense of the accuracy in predicting the other class. Applying the principle of precaution, we therefore chose a threshold that held a true positive rate of 95% since we considered it especially important to correctly predict locations where the species flew within the critical height range. By doing so we were fairly conservative, accepting the risk that some flight locations above 200 m a.g.l. were wrongly classified into the critical flight height range. The resulting probability map (Fig. 3b) might thus overrepresent critical areas, which is a minor issue from a conservation and risk assessment viewpoint. Therefore, the map showing areas with high risks of conflict (Fig. 3f) is a refinement of our previous conflict map (Vignali et al., 2021) since it eventually subtracts the areas over which the species flies high above the ground. Although we adopted a conservative approach, 9,2% of the areas previously classified as potential conflict, which corresponds to around 2,358 km², is most likely not overflown within the critical height range.

Our model was developed in a framework particularly suitable for very large datasets. In effect, modern tracking devices are capable of collecting data at high temporal resolution, thereby introducing new challenges for their analysis (López-López, 2016). The use of artificial neural network approaches is a possible solution to address this challenge. First, contrary to other classical statistical methods, it does not require an *a priori* definition of the functional forms for each relationship between predictors. (Breiman, 2001) as complex non-linear relationships among variables are learned directly from the data. Second, it can take advantage of specific libraries developed to create efficient data pipelines (see for example the *tensorflow dataset* library and its R implementation (Allaire et al., 2020)). Data pipelines serve for transformations like normalization, standardization, and one-hot encoding of categorical variables on batches of data that are then fed into the neural network. This way it is not required to apply each transformation to the entire dataset but rather to the single batches, and the data can be loaded in batches directly from a file or database (see code in Supplementary material).

Although innovative in several respects, our approach focuses only one species potentially affected by wind turbines deployment. We believed, however, that a similar method could be readily applied to any other raptor species, if not to other soaring birds such as storks or herons, for which flight heights might also be decisive. Of course, a combination of species-specific predictive

models obtained from different emblematic species potentially impacted by the wind industry development would be a major step towards a biodiversity-friendly spatial planning of these novel anthropogenic infrastructures that not only obliterate our landscapes but may also affect wildlife survival. Policy makers and land-use planners, wind energy promoters and conservation biologists would all benefit from the rapid development of such comprehensive decision tools. Wind energy companies, in particular, could from the onset evaluate whether their investments would be at risk of not obtaining official approval. It must be explicitly stressed, however, that models such as the one presented here can inform spatial planning but in no way represent substitutes to in-situ environmental impact assessments that are prerequisites for any infrastructure project development.

ACKNOWLEDGMENTS

We thank the Stiftung Pro Bartgeier, the Nationalpark Hohe Tauern, the Vulture Conservation Foundation, Vautours en Baronnies, LIFE GypConnect, LIFE GypHelp, ASTERS, LPO Grands Causses, Parc Naturel Régional du Vercors, the Parco Naturale Alpi Marittime, and the Parco Nazionale dello Stelvio for sharing the GPS data. We are grateful to Olivier Roth for the English translation of table S2.

REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ... Zheng, X. (2016). TensorFlow: A System for Large-Scale Machine Learning. *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, 265–283.
- Allaire, J. J., & Chollet, F. (2020). *keras*; R Interface to “Keras”. R package version 2.3.0.0. <https://CRAN.R-project.org/package=keras>.
- Allaire, J. J., Tang, Y., & Ushey, K. (2020). *tfdatasets*: Interface to “Tensorflow” Datasets. R package version 2.2.0. <https://CRAN.R-project.org/package=tfdatasets>.
- Allinson, T. (2017). Introducing a New Avian Sensitivity Mapping Tool to Support the Siting of Wind Farms and Power Lines in the Middle East and Northeast Africa. In *Wind Energy and Wildlife Interactions* (pp. 207–218). Springer International Publishing. https://doi.org/10.1007/978-3-319-51272-3_12
- Arnett, E. B., Baerwald, E. F., Mathews, F., Rodrigues, L., Rodríguez-Durán, A., Rydell, J.,

- Villegas-Patracá, R., & Voigt, C. C. (2015). Impacts of wind energy development on bats: A global perspective. In *Bats in the Anthropocene: Conservation of Bats in a Changing World* (pp. 295–323). Springer International Publishing. https://doi.org/10.1007/978-3-319-25220-9_11
- Balotari-Chiebao, F., Brommer, J. E., Saurola, P., Ijäs, A., & Laaksonen, T. (2018). Assessing space use by pre-breeding white-tailed eagles in the context of wind-energy development in Finland. *Landscape and Urban Planning*, **177**, 251–258. <https://doi.org/10.1016/j.landurbplan.2018.05.012>
- Barrios, L., & Rodríguez, A. (2004). Behavioural and environmental correlates of soaring-bird mortality at on-shore wind turbines. *Journal of Applied Ecology*, **41**(1), 72–81. <https://doi.org/10.1111/j.1365-2664.2004.00876.x>
- Beston, J. A., Diffendorfer, J. E., Loss, S. R., & Johnson, D. H. (2016). Prioritizing Avian Species for Their Risk of Population-Level Consequences from Wind Energy Development. *PLOS ONE*, **11**(3), e0150813. <https://doi.org/10.1371/journal.pone.0150813>
- BirdLife International. (2015). *Gypaetus barbatus*. *The IUCN Red List of Threatened Species 2015: e.T22695174A60116752*. Downloaded on 07 March 2020.
- Braunisch, V., Coppes, J., Bächle, S., & Suchant, R. (2015). Underpinning the precautionary principle with evidence: A spatial concept for guiding wind power development in endangered species' habitats. *Journal for Nature Conservation*, **24**(C), 31–40. <https://doi.org/10.1016/j.jnc.2015.01.003>
- Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, **16**(3), 199–215. <https://doi.org/10.1214/ss/1009213726>
- Bright, J., Langston, R., Bullman, R., Evans, R., Gardner, S., & Pearce-Higgins, J. (2008). Map of bird sensitivities to wind farms in Scotland: A tool to aid planning and conservation. *Biological Conservation*, **141**(9), 2342–2356. <https://doi.org/10.1016/j.biocon.2008.06.029>
- Browning, E., Bolton, M., Owen, E., Shoji, A., Guilford, T., & Freeman, R. (2018). Predicting animal behaviour using deep learning: GPS data alone accurately predict diving in seabirds. *Methods in Ecology and Evolution*, **9**(3), 681–692. <https://doi.org/10.1111/2041-210X.12926>
- Bundesamt für Energie BFE. (2016). *Windatlas Schweiz: Jahresmittel der modellierten Windgeschwindigkeit und Windrichtung*. Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK. www.bfe.admin.ch
- Bundesamt für Statistik. (1967). *Vereinfachte Geotechnische Karte der Schweiz. Nach der Karte 1: 200'000 der Schweizerischen Geotechnischen Kommission aus den Jahren 1963-67* (p. 17).
- Carrete, M., Sánchez-Zapata, J. A., Benítez, J. R., Lobón, M., & Donázar, J. A. (2009). Large scale risk-assessment of wind-farms on population viability of a globally endangered long-lived raptor. *Biological Conservation*, **142**(12), 2954–2961. <https://doi.org/10.1016/J.BIOCON.2009.07.027>

- Drewitt, A. L., & Langston, R. H. W. (2006). Assessing the impacts of wind farms on birds. *Ibis*, **148**(s1), 29–42. <https://doi.org/10.1111/j.1474-919X.2006.00516.x>
- Dulluri, A., & Rat, D. (2019). The Green-Green Dilemma: Reconciling the Conflict between Renewable Energy and Biodiversity Coastal Adaptation and Ecosystem Resilience in Odisha View project. *The Journal of Health, Environment, & Education*, **11**, 6–15. <https://doi.org/10.18455/19002>
- Duriez, O., Kato, A., Tromp, C., Dell’Omo, G., Vyssotski, A. L., Sarrazin, F., & Ropert-Coudert, Y. (2014). How Cheap Is Soaring Flight in Raptors? A Preliminary Investigation in Freely-Flying Vultures. *PLoS ONE*, **9**(1), e84887. <https://doi.org/10.1371/journal.pone.0084887>
- European Commission. (2019). *Communication from the commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. The European Green Deal. COM/2019/640 final*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0640>
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**(1), 38–49. <https://doi.org/10.1017/S0376892997000088>
- Frick, W. F., Baerwald, E. F., Pollock, J. F., Barclay, R. M. R., Szymanski, J. A., Weller, T. J., Russell, A. L., Loeb, S. C., Medellin, R. A., & McGuire, L. P. (2017). Fatalities at wind turbines may threaten population viability of a migratory bat. *Biological Conservation*, **209**, 172–177. <https://doi.org/10.1016/j.biocon.2017.02.023>
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, **29**(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation. *Journal of Computational and Graphical Statistics*, **24**(1), 44–65. <https://doi.org/10.1080/10618600.2014.907095>
- Gonseth, Y., Wohlgemuth, T., Sansonnens, B., & Buttler, A. (2001). *Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard* (p. 48). Umwelt Materialien Nr. 137 Bundesamt für Umwelt, Wald und Landschaft.
- Greenwell, B. M. (2017). pdp: An r package for constructing partial dependence plots. *The R Journal*, **9**(1), 241–246. <https://doi.org/10.32614/RJ-2017-016>
- Greenwell, B. M., & Boehmke, B. C. (2020). Variable Importance Plots-An Introduction to the vip Package. *The R Journal*, **12**, 343–366. <https://doi.org/10.32614/RJ-2020-013>
- Hanssen, F., May, R., & Nygård, T. (2020). High-Resolution Modeling of Uplift Landscapes can Inform Micrositing of Wind Turbines for Soaring Raptors. *Environmental Management*, **66**(3), 319–332. <https://doi.org/10.1007/s00267-020-01318-0>
- Hegglin, D., Wehrle, M., & Aebischer, A. (2004). Satellite tracking of Bearded Vultures - the project « BEARDED VULTURE ON THE MOVE ». *Annu. Rep. Found. Conserv. Bearded*

Vultures, 52–55.

Hijmans, R. J. (2019). *raster: Geographic Data Analysis and Modeling. R package* (p.).

<https://cran.r-project.org/package=raster>.

Hirzel, A., Posse, B., Oggier, P.-A., Crettenand, Y., Glenz, C., & Arlettaz, R. (2004). Ecological requirements of reintroduced Bearded vultures colonizing the Alps : implications for release policy. *Journal of Applied Ecology*, **41**(6), 1103–1116. <https://doi.org/10.1111/j.0021-8901.2004.00980.x>

Hosmer, D. W., & LemeshowStanley. (2000). *Applied Logistic Regression* (2nd ed.). John Wiley & Sons.

Izquierdo, D. (2017). *European Union single species recovery plan for the Palearctic population of Bearded Vulture (Gypaetus barbatus barbatus)*.

Janss, G. F. E., de Lucas, M., Whitfield, D. P., Lazo, A., & Ferrer, M. (2010). The precautionary principle and wind-farm planning in Andalucía. *Biological Conservation*, **143**(8), 1827–1828. <https://doi.org/10.1016/j.biocon.2010.02.016>

Katzner, T. E., Brandes, D., Miller, T., Lanzone, M., Maisonneuve, C., Tremblay, J. A., Mulvihill, R., & Merovich, G. T. (2012). Topography drives migratory flight altitude of golden eagles: implications for on-shore wind energy development. *Journal of Applied Ecology*, **49**(5), 1178–1186. <https://doi.org/10.1111/j.1365-2664.2012.02185.x>

Katzner, T. E., Nelson, D. M., Diffendorfer, J. E., Duerr, A. E., Campbell, C. J., Leslie, D., Vander Zanden, H. B., Yee, J. L., Sur, M., Huso, M. M. P., Braham, M. A., Morrison, M. L., Loss, S. R., Poessel, S. A., Conkling, T. J., & Miller, T. A. (2019). Wind energy: An ecological challenge. *Science*, **366**(6470), 1206–1207. <https://doi.org/10.1126/science.aaz9989>

Keller, V., Gerber, A., Schmid, H., Volet, B., & Zbinden, N. (2010). *Rote Liste Brutvögel. Gefährdete Arten der Schweiz, Stand 2010*. Bundesamt für Umwelt, Bern, und Schweizerische Vogelwarte, Sempach UmweltVollzug Nr. 1019. 53 S.

Lee, J., & Zhao, F. (2020). GWEC Global Wind Report. In *Global Wind Energy Council*.

Leung, D. Y. C., & Yang, Y. (2012). Wind energy development and its environmental impact: A review. *Renewable and Sustainable Energy Reviews*, **16**(1), 1031–1039.

<https://doi.org/10.1016/J.RSER.2011.09.024>

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, **40**(4), 778–789.

<https://doi.org/10.1111/jbi.12058>

López-López, P. (2016). Individual-Based Tracking Systems in Ornithology: Welcome to the Era of Big Data. *Ardeola*, **63**(1), 103–136. <https://doi.org/10.13157/arla.63.1.2016.rp5>

Lörcher, F., & Hegglin, D. (2020). Life and death in the bearded vulture in the Alps. *InfoGIPETO*, **37**, 9–12.

Madders, M., & Whitfield, D. P. (2006). Upland raptors and the assessment of wind farm

- impacts. *Ibis*, **148**(s1), 43–56. <https://doi.org/10.1111/j.1474-919X.2006.00506.x>
- Maekawa, T., Ohara, K., Zhang, Y., Fukutomi, M., Matsumoto, S., Matsumura, K., Shidara, H., Yamazaki, S. J., Fujisawa, R., Ide, K., Nagaya, N., Yamazaki, K., Koike, S., Miyatake, T., Kimura, K. D., Ogawa, H., Takahashi, S., & Yoda, K. (2020). Deep learning-assisted comparative analysis of animal trajectories with DeepHL. *Nature Communications*, **11**(1). <https://doi.org/10.1038/s41467-020-19105-0>
- Margalida, A., Bertran, J., & Heredia, R. (2009). Diet and food preferences of the endangered Bearded Vulture *Gypaetus barbatus*: a basis for their conservation. *Ibis*, **151**(2), 235–243. <https://doi.org/10.1111/j.1474-919X.2008.00904.x>
- Margalida, A., & Villalba, D. (2017). The importance of the nutritive value of old bones in the diet of Bearded vultures *Gypaetus barbatus*. *Scientific Reports*, **7**(1), 1–5. <https://doi.org/10.1038/s41598-017-08812-2>
- Martin, G. R., Portugal, S. J., & Murn, C. P. (2012). Visual fields, foraging and collision vulnerability in Gyps vultures. *Ibis*, **154**(3), 626–631. <https://doi.org/10.1111/j.1474-919X.2012.01227.x>
- Mc Guinness, S., Muldoon, C., Tierney, N., Murray, A., Egan, S., & Crowe, O. (2015). Bird Sensitivity Mapping for Wind Energy Developments and Associated Infrastructure in the Republic of Ireland. *BirdWarchIreland*, 120 pp.
- Mingozzi, T., & Estève, R. (1997). Analysis of a historical extirpation of the bearded vulture *Gypaetus barbatus* (L.) in the western Alps (France-Italy): Former distribution and causes of extirpation. *Biological Conservation*, **79**(2–3), 155–171. [https://doi.org/10.1016/S0006-3207\(96\)00110-3](https://doi.org/10.1016/S0006-3207(96)00110-3)
- Murgatroyd, M., Bouten, W., & Amar, A. (2021). A predictive model for improving placement of wind turbines to minimise collision risk potential for a large soaring raptor. *Journal of Applied Ecology*, **58**(4), 857–868. <https://doi.org/10.1111/1365-2664.13799>
- Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., & Getz, W. M. (2012). Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *The Journal of Experimental Biology*, **215**, 986–996. <https://doi.org/10.1242/jeb.058602>
- Northrup, J. M., & Wittemyer, G. (2013). Characterising the impacts of emerging energy development on wildlife, with an eye towards mitigation. *Ecology Letters*, **16**(1), 112–125. <https://doi.org/10.1111/ele.12009>
- Pennycuik, C. J. (1973). The Soaring Flight of Vultures. *Scientific American*, **229**(6), 102–109. <https://doi.org/10.1038/scientificamerican1273-102>
- Pennycuik, C. J. (2008). Soaring behaviour and performance of some East African birds, observed from a motor-glider. *Ibis*, **114**(2), 178–218. <https://doi.org/10.1111/j.1474-919X.1972.tb02603.x>
- Perold, V., Ralston-Paton, S., & Ryan, P. (2020). On a collision course? The large diversity of

- birds killed by wind turbines in South Africa. *Ostrich*, **2020**(3), 228–239.
<https://doi.org/10.2989/00306525.2020.1770889>
- Péron, G., Calabrese, J. M., Duriez, O., Fleming, C. H., García-Jiménez, R., Johnston, A., Lambertucci, S. A., Safi, K., & Shepard, E. L. C. (2020). The challenges of estimating the distribution of flight heights from telemetry or altimetry data. *Animal Biotelemetry*, **8**(1), 5.
<https://doi.org/10.1186/s40317-020-00194-z>
- Péron, G., Fleming, C. H., Duriez, O., Fluhr, J., Itty, C., Lambertucci, S., Safi, K., Shepard, E. L. C., & Calabrese, J. M. (2017). The energy landscape predicts flight height and wind turbine collision hazard in three species of large soaring raptor. *Journal of Applied Ecology*, **54**(6), 1895–1906. <https://doi.org/10.1111/1365-2664.12909>
- Poessel, S. A., Brandt, J., Mendenhall, L., Braham, M. A., Lanzone, M. J., McGann, A. J., & Katzner, T. E. (2018). Flight response to spatial and temporal correlates informs risk from wind turbines to the California Condor. *The Condor*, **120**(2), 330–342.
<https://doi.org/10.1650/CONDOR-17-100.1>
- Poessel, S. A., Duerr, A. E., Hall, J. C., Braham, M. A., & Katzner, T. E. (2018). Improving estimation of flight altitude in wildlife telemetry studies. *Journal of Applied Ecology*, **55**(4), 2064–2070. <https://doi.org/10.1111/1365-2664.13135>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reid, T., Krüger, S., Whitfield, D. P., & Amar, A. (2015). Using spatial analyses of bearded vulture movements in southern Africa to inform wind turbine placement. *Journal of Applied Ecology*, **52**(4), 881–892. <https://doi.org/10.1111/1365-2664.12468> @ [10.1111/\(ISSN\)1365-2664.ENDANGEREDSPECIES](https://doi.org/10.1111/(ISSN)1365-2664.ENDANGEREDSPECIES)
- Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., & Nathan, R. (2014). AcceleRater: A web application for supervised learning of behavioral modes from acceleration measurements. *Movement Ecology*, **2**(1), 27. <https://doi.org/10.1186/s40462-014-0027-0>
- RStudio Team. (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.
- Rushworth, I., & Krüger, S. (2014). Wind farms threaten southern Africa's cliff-nesting vultures. *Ostrich*, **85**(1), 13–23. <https://doi.org/10.2989/00306525.2014.913211>
- Ruxton, G. D., & Houston, D. C. (2004). Obligate vertebrate scavengers must be large soaring fliers. *Journal of Theoretical Biology*, **228**(3), 431–436.
<https://doi.org/10.1016/j.jtbi.2004.02.005>
- Schaub, M., Zink, R., Beissmann, H., Sarrazin, F., & Arlettaz, R. (2009). When to end releases in reintroduction programmes: demographic rates and population viability analysis of bearded vultures in the Alps. *Journal of Applied Ecology*, **46**(1), 92–100.
<https://doi.org/10.1111/j.1365-2664.2008.01585.x>
- Shamoun-Baranes, J., Leshem, Y., Yom-Tov, Y., & Liechti, O. (2003). Differential Use of

- Thermal Convection by Soaring Birds Over Central Israel. *The Condor*, **105**(2), 208–218.
<https://doi.org/10.1093/condor/105.2.208>
- Shepard, E. L. C., & Lambertucci, S. A. (2013). From daily movements to population distributions: Weather affects competitive ability in a guild of soaring birds. *Journal of the Royal Society Interface*, **10**(88), 20130612. <https://doi.org/10.1098/rsif.2013.0612>
- Shepard, E. L. C., Lambertucci, S. A., Vallmitjana, D., & Wilson, R. P. (2011). Energy Beyond Food: Foraging Theory Informs Time Spent in Thermals by a Large Soaring Bird. *PLoS ONE*, **6**(11), e27375. <https://doi.org/10.1371/journal.pone.0027375>
- Shu, C., & Burn, D. H. (2004). Artificial neural network ensembles and their application in pooled flood frequency analysis. *Water Resources Research*, **40**(9), W09301.
<https://doi.org/10.1029/2003WR002816>
- Smallwood, K. S., & Thelander, C. (2008). Bird Mortality in the Altamont Pass Wind Resource Area, California. *JOURNAL OF WILDLIFE MANAGEMENT*, **72**(1), 215–223.
<https://doi.org/10.2193/2007-032>
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, **15**, 1929–1958. <http://jmlr.org/papers/v15/srivastava14a.html>
- Straka, T. M., Fritze, M., & Voigt, C. C. (2020). The human dimensions of a green–green-dilemma: Lessons learned from the wind energy — wildlife conflict in Germany. *Energy Reports*, **6**, 1768–1777. <https://doi.org/10.1016/j.egy.2020.06.028>
- Telford, R. J., & Birks, H. J. B. (2009). Evaluation of transfer functions in spatially structured environments. *Quaternary Science Reviews*, **28**(13–14), 1309–1316.
<https://doi.org/10.1016/j.quascirev.2008.12.020>
- Thaxter, C. B., Buchanan, G. M., Carr, J., Butchart, S. H. M., Newbold, T., Green, R. E., Tobias, J. A., Foden, W. B., O'Brien, S., & Pearce-Higgins, J. W. (2017). Bird and bat species' global vulnerability to collision mortality at wind farms revealed through a trait-based assessment. *Proceedings of the Royal Society B: Biological Sciences*, **284**(1862), 20170829.
<https://doi.org/10.1098/rspb.2017.0829>
- Thieurmel, B., & Elmarhraoui, A. (2019). *suncalc: Compute Sun Position, Sunlight Phases, Moon Position and Lunar Phase. R package version 0.5.0.* <https://CRAN.R-project.org/package=suncalc>.
- Tikkanen, H., Rytkönen, S., Karlin, O. P., Ollila, T., Pakanen, V. M., Tuohimaa, H., & Orell, M. (2018). Modelling golden eagle habitat selection and flight activity in their home ranges for safer wind farm planning. *Environmental Impact Assessment Review*, **71**, 120–131.
<https://doi.org/10.1016/j.eiar.2018.04.006>
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). block CV: An r package for generating spatially or environmentally separated folds for k -fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, **10**(2), 225–232.
<https://doi.org/10.1111/2041-210X.13107>

- Venter, J. A., Martens, F. R., & Wolter, K. (2019). Conservation Buffer Sizes Derived from Movement Data of Adult Cape Vultures (*Gyps coprotheres*) in South Africa. *African Zoology*, **54**(2), 115–118. <https://doi.org/10.1080/15627020.2019.1600428>
- Vignali, S., Lörcher, F., Hegglin, D., Arlettaz, R., & Braunisch, V. (2021). Modelling the habitat selection of the bearded vulture to predict areas of potential conflict with wind energy development in the Swiss Alps. *Global Ecology and Conservation*, **25**, e01405. <https://doi.org/10.1016/j.gecco.2020.e01405>
- Voigt, C. C. (2021). Insect fatalities at wind turbines as biodiversity sinks. *Conservation Science and Practice*, e366. <https://doi.org/10.1111/csp2.366>
- Voigt, C. C., Straka, T. M., & Fritze, M. (2019). Producing wind energy at the cost of biodiversity: A stakeholder view on a green-green dilemma. *Journal of Renewable and Sustainable Energy*, **11**(6), 063303. <https://doi.org/10.1063/1.5118784>
- Watson, R. T., Kolar, P. S., Ferrer, M., Nygård, T., Johnston, N., Hunt, W. G., Smit-Robinson, H. A., Farmer, C. J., Huso, M., & Katzner, T. E. (2018). Raptor Interactions With Wind Energy: Case Studies From Around the World. *Journal of Raptor Research*, **52**(1), 1–18. <https://doi.org/10.3356/jrr-16-100.1>
- Wellig, S. D., Nusslé, S., Miltner, D., Kohle, O., Glazot, O., Braunisch, V., Obrist, M. K., & Arlettaz, R. (2018). Mitigating the negative impacts of tall wind turbines on bats: Vertical activity profiles and relationships to wind speed. *PLoS ONE*, **13**(3), e0192493. <https://doi.org/10.1371/journal.pone.0192493>
- Wilson, J. D. (1984). Determining a TOPEX score. *Scottish Forestry*, **38**(4), 251–256.

SUPPLEMENTARY MATERIAL

Table S1: Reclassification table used to define the land cover classes used in the model from the classes provided by the Vector25, layer pri25_a. OBJECTVAL was the field used for the reclassification.

Class	OBJECTVAL	Label
1	Z_BaumS	Orchard, vineyard, and horticulture areas
	Z_ObstAn	
	Z_Reben	
2	Z_Wald	Forest
	Z_WaldOf	
3	Z_Gebue	Brush
4	Z_GerGeb	Scree
	Z_GerGle	
	Z_Geroel	
	Z_GerWa	
5	Z_GsPist	Others
	Z_Uebrig	
6	Z_HaPist	Anthropic areas
	Z_KiGrub	
	Z_LeGrub	
	Z_Siedl	
	Z_StauDa	
7	Z_StauMa	Marsh
	Z_SteBru	
	Z_SumGeb	
	Z_Sumpf	
8	Z_SumWa	Water, lakes, and rivers
	Z_SumWaO	
9	Z_Fluss	Rock
	Z_See	
10	Z_Fels	Rock
10	Z_Glet	Glacier

Table S2: Reclassification of the classes defined in the simplified geotechnical map of Switzerland (Bundesamt für Statistik, 1967) used to define the classes within the geology environmental variable. GT_ID was the field used for the reclassification.

Class	GT_ID	Label	Short description
1	1	Others	Lake
1	2	Others	Glacier
1	3	Others	Sands to silts, usually clay, partly calcareous loess, often with attachments (ground moraines) or debris, fine to coarse, mixed with sand, silt, and clay (upper moraines)
1	4	Others	Clayey silts and clays, sometimes with inclusions of sand to gravel (sea soil clay, alluvial clay, hillside clay)
1	5	Others	Gravels and sands, clean or silty, sometimes slightly cemented (rubble of the ice age)
1	6	Others	Gravels and sands, mostly clean, sometimes with clayey-silty overlaps or retention, as well as extensive gravel deposits (current stream deposits)
1	7	Others	Larger areas with angular rubble, often by block size (landslide material, scree)
1	8	Others	Marl with intercalations of weakly consolidated sandstones, partly predominant, and occasional conglomerates or shell sandstones
1	9	Others	Marl and shale clay to shale, with limestone, dolomite and sandstone benches, sometimes with layers of gypsum and anhydrite
1	10	Others	Red, lime-free sandstones to sandy shale
1	11	Others	Ferruginous, mostly lime-free, lean to rich clay, often including iron ore, silica sand, refractory clay
1	12	Others	Marl with intercalations of medium strong consolidated sandstones, partly predominant, and occasionally of conglomerates
1	13	Others	Conglomerates, weakly to moderately consolidated, with abundant to predominant sandstone and marl layers
1	14	Others	Conglomerates, weakly to moderately consolidated, always accompanied by sandstone and marl layers
1	15	Others	Conglomerates and breccias, highly strengthened, with different proportions of arkoses and sandstones, sometimes with sandy phyllite and volcanic rocks
1	16	Others	Slate to phyllites, often sandy, with deposits of sandstones and breccias to conglomerates
1	17	Others	Marl to lime phyllites, with inclusions of sandstones, partly predominant
2	18	Limestone	Lime phyllites to lime mica slate, frequently with inclusions of sandy limestone, dolomites, quartzites, low lime phyllites and green schists
2	19	Limestone	Limes in general in massive formation, often with marly interlayers, partly with siliceous limestones and green sandstones
2	20	Limestone	Limes, sand limes to pebble limes, slated, with layers of marl shale and lime phyllites, some with cherts and dolomites, rarely with rare breccia to conglomerate structure
1	21	Others	Significant layers of marl shale and marl limes
2	22	Limestone	Dolomite and cell lime, partly. with gypsum layers
3	23	Granite	Granite, quartz diorite, quartz syenite and diorite, predominantly homogeneous
3	24	Granite	Quartz porphyry, porphyrites and porphyry tufa, moderately to slightly slated
3	25	Granite	Quartzite, massive or platy to slaty
4	26	Gneiss	Two-mica to biotite gneisses, often with abundant feldspar, partly with amphibolites and hornblende-bearing gneisses
4	27	Gneiss	Slated, often sericite-rich conglomerates and breccias
4	28	Gneiss	Sericite-chlorite gneiss to shale, homogeneous or heterogeneous
4	29	Gneiss	Green shale, with transitions to basic rocks
4	30	Gneiss	Serpentinites, rarely with transitions to peridotites or olivine rocks

Table S3: Model evaluation based on a leave-one-bird-out cross validation. Each model was trained without the GPS locations of one bird at a time and evaluated with the held apart locations of the respective bird using the area under the operating characteristic (ROC) curve AUC (Fielding & Bell, 1997) as performance metric. Each line in the table represents the results of one model, with BirdID indicating the removed bird (for BirdIDs see Table 2), the AUC evaluated for the training and testing dataset, and the number of locations included in the testing dataset (N).

Bird ID	Training AUC	Testing AUC	N
BG905	0.714	0.826	255
W349	0.713	0.802	78
W346	0.711	0.802	666
W356	0.717	0.795	170
BG1068	0.717	0.768	1,948
BG983	0.717	0.762	268
W313	0.716	0.745	345
W196	0.714	0.745	12,056
BG838	0.718	0.737	1,358
BG1003	0.711	0.728	24,356
BG1001	0.716	0.723	18,244
BG964	0.716	0.723	27,357
W251	0.719	0.722	840
BG900	0.717	0.721	17,006
W285	0.713	0.720	4,734
BG899	0.717	0.716	39,961
BG840	0.712	0.716	407
BG841	0.712	0.713	7,113
W209	0.715	0.706	3,594
BG1031	0.718	0.705	421
BG998	0.718	0.695	3,152
W284	0.716	0.695	1,306
BG843	0.718	0.694	7,214
BG802	0.722	0.691	39,112
BG1071	0.717	0.684	1,667
BG321	0.711	0.676	671
BG797	0.721	0.673	12,162
W361	0.713	0.656	303
BG1022	0.712	0.654	77
BG980	0.719	0.639	472

Table S4: Model evaluation using a spatial block cross validation. The study area was divided in spatial blocks of 10 x 10 km that were randomly partitioned into five folds (Fig. S1). The folds were then used to run the cross validation and the models were evaluated with the area under the operating characteristic (ROC) curve AUC (Fielding & Bell, 1997) (in parenthesis the number of GPS locations included in each dataset).

Fold	Training AUC	Testing AUC
1	0.719 (N = 183,139)	0.703 (N = 44,174)
2	0.726 (N = 184,238)	0.689 (N = 43,075)
3	0.719 (N = 179,605)	0.716 (N = 47,708)
4	0.721 (N = 184,260)	0.711 (N = 43,053)
5	0.712 (N = 178,010)	0.732 (N = 49,303)

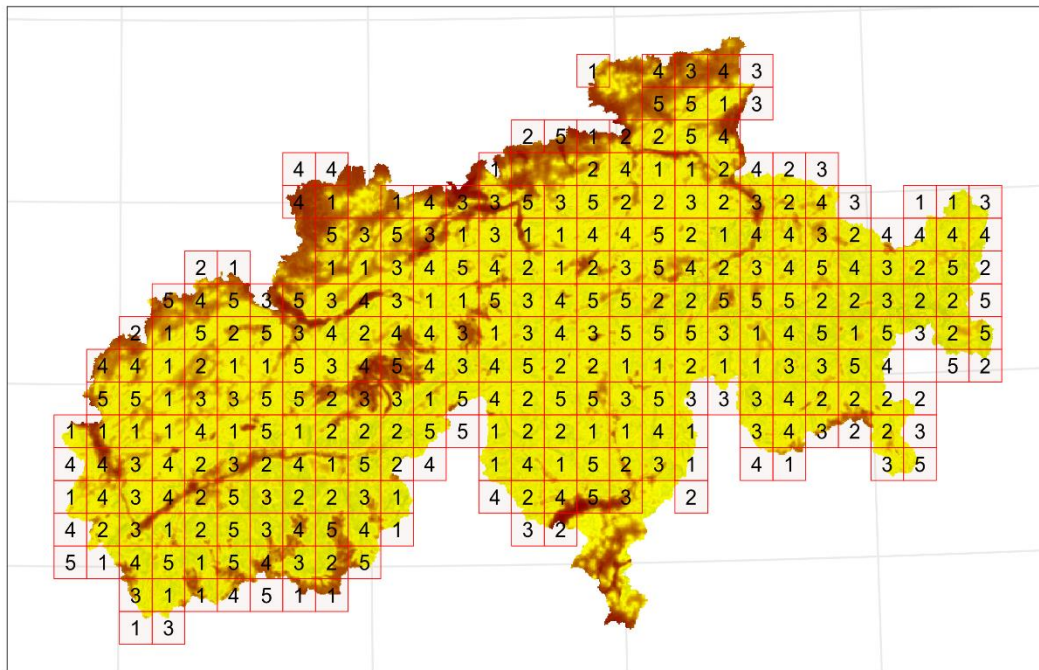


Figure S1 Spatial block partitioning with spatial blocks of 10x10 km. Each block was randomly assigned to one of the five cross-validation folds.

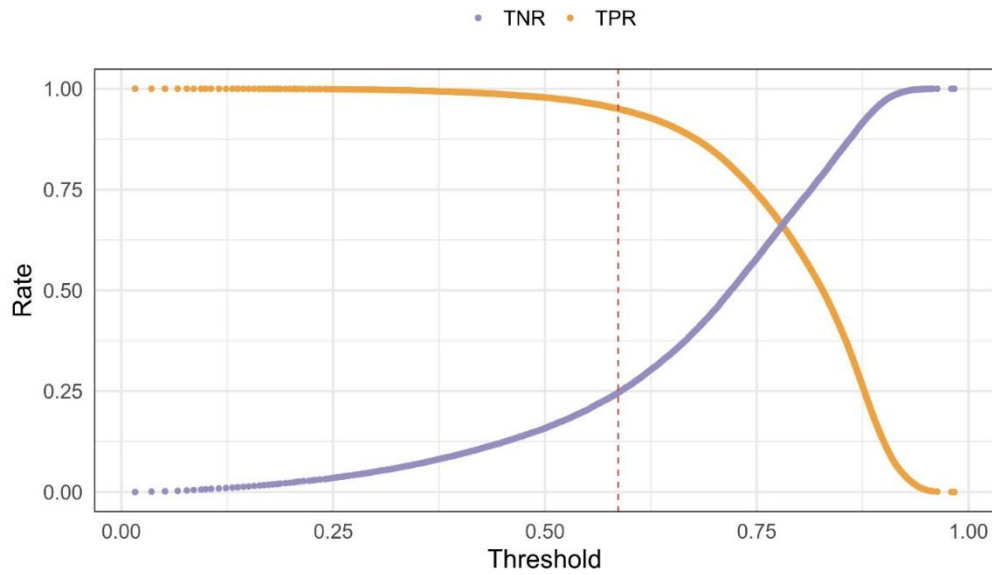


Figure S2 Sensitivity (TNR: true positive rate) and specificity (TPR: true negative rate) under varying thresholds computed for the training dataset from the averaged probability of a 30-bagging procedure. The vertical dashed red line shows the selected threshold at 0.586, for which 95% of the locations occurring below 200 m a.g.l. are correctly predicted.

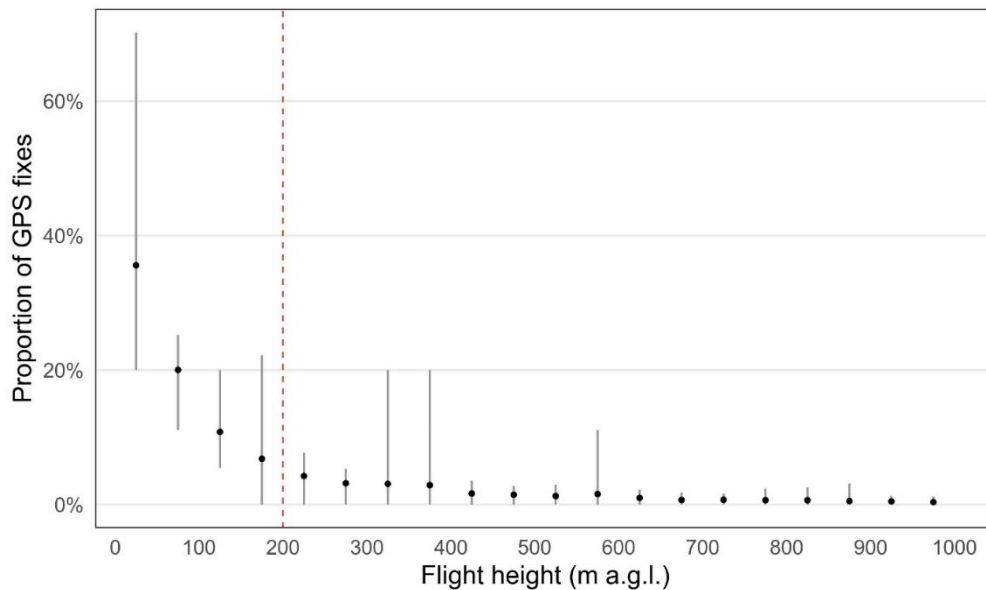


Figure S3 Flight height breakdown. Proportion of bearded vulture GPS locations recorded in flight per bird within each 50 m height bin (1.93% of the locations were collected above 1000 m a.g.l., not shown in the graph). Each line represents the minimum maximum range and the blue dot the mean value. The dashed red line is the critical threshold of 200 m.

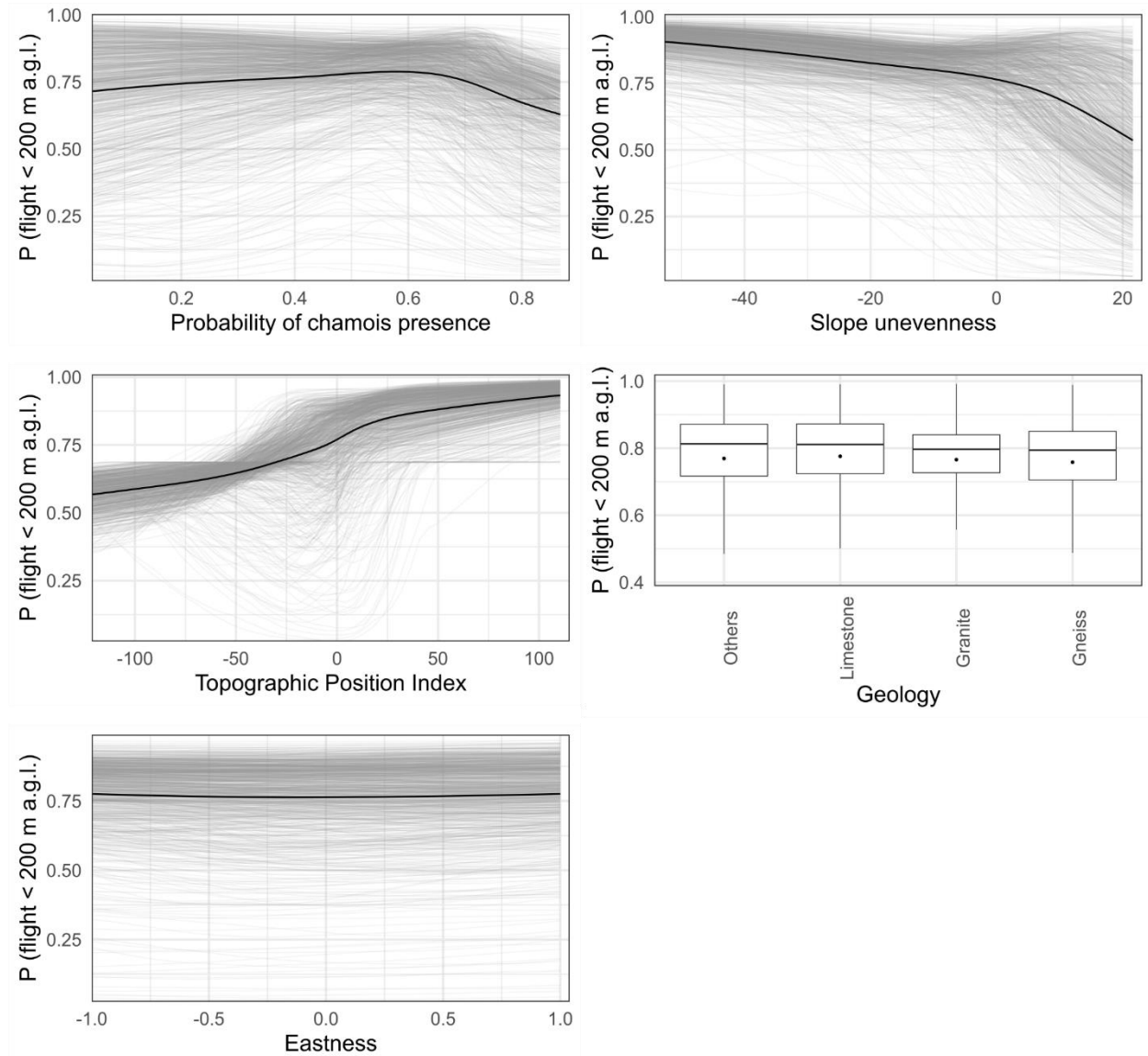


Figure S4 Marginal effects of the environmental variables (not shown in Fig. 2) used for predicting the probability of a bearded vulture flying below 200 m a.g.l. In grey are plotted 1000 randomly sampled individual conditional expectation (ICE) curves (Goldstein et al., 2015) and in blue the partial dependence (PD) curve (Friedman, 2001). For geology, a categorical variable, each boxplot shows the ICE values without outliers and the blue dot the value of the PD.

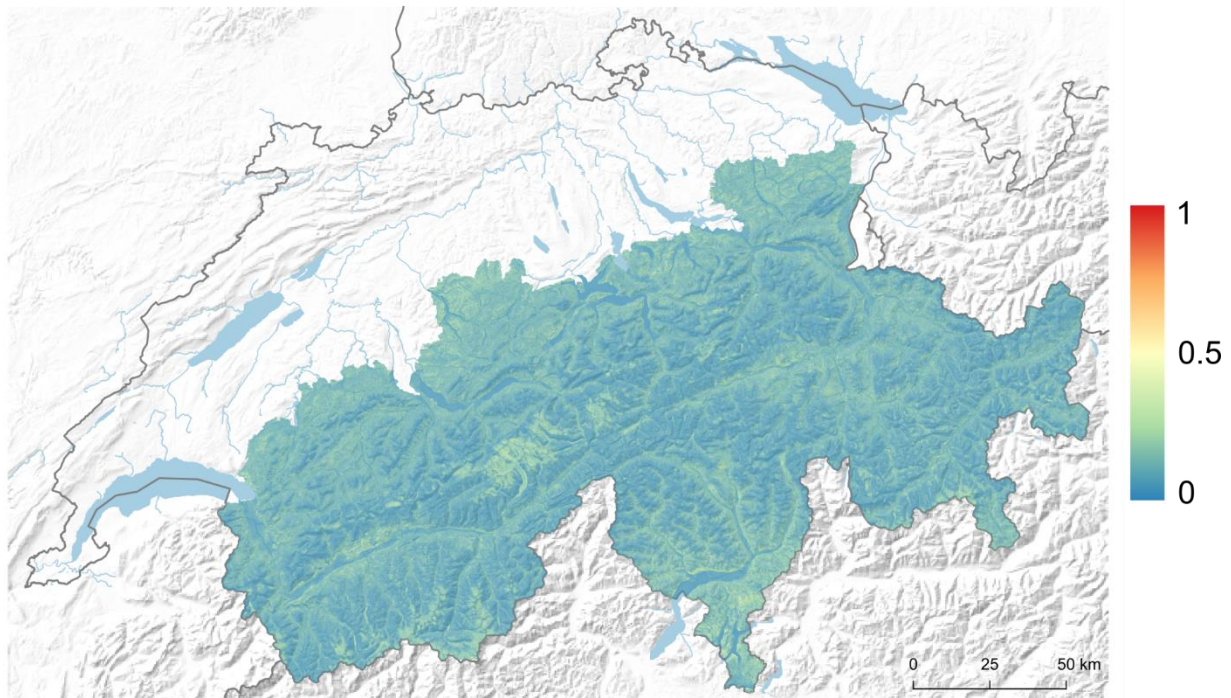


Figure S5 Model (un)certainty illustrated as the magnitude of the 95% credible interval calculated from the 30-bagging procedure for the predicted probability of a bearded vulture flying below 200 m a.g.l. within the Swiss Alpine range (gradient from blue = zero to red = one).

REFERENCES

- Bundesamt für Statistik. (1967). Vereinfachte Geotechnische Karte der Schweiz. Nach der Karte 1: 200'000 der Schweizerischen Geotechnischen Kommission aus den Jahren 1963-67 (p. 17).
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**(1), 38–49. <https://doi.org/10.1017/S0376892997000088>
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, **29**(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation. *Journal of Computational and Graphical Statistics*, **24**(1), 44–65. <https://doi.org/10.1080/10618600.2014.907095>

R code to reproduce the analysis

```
## *****
## Title: CODE FOR ARTICLE
## Purpose: Reproduce analysis presented in the article
## Author: Sergio Vignali
##.*****
# The code has five sections and each section can be executed without
# having to run the previous ones. There is an initial part that loads
# libraries, data, and create a utility function. This part of the code
# needs to be executed at least once all the times a new R session is
# created.

# *****
# SETUP----
# *****
# R VERSION 4.0.2

## Load libraries and data----
# Libraries
library(reticulate) # Version 1.18
library(tensorflow) # Version 2.2.0
library(keras) # Version 2.3.0.0
library(tfdatasets) # Version 2.2.0.9000, at the moment of writing it
# was not possible to save the specifications using
# the released version
library(tidyverse) # Version 1.3.0
hp <- import("tensorboard.plugins.hparams.api")

# The tensorflow python version was 2.4.1, check it with the following
# command
tf$version$VERSION

# Load data, expected to be in the project folder and named
# "bv-data.csv.gz"
dataset <- read_csv(
  file = "bv-data.csv.gz",
  col_types = "ccddddddici"
)
dataset %>% glimpse()

# Create indexes for training (70%) and validation (30%) datasets
set.seed(25)
idx <- sample(nrow(dataset), size = 0.3 * nrow(dataset))

## Utility function----
# This function creates and compile a new model given the hyperparameters'
# values, a dataset, and the feature specifications
```

```

getModel <- function(units_1l, units_2l, rate, ds, spec) {
  input <- layer_input_from_dataset(ds)
  output <- input %>%
    layer_dense_features(feature_columns = dense_features(spec)) %>%
    # First hidden layer
    layer_dense(units = units_1l, activation = "relu",
                kernel_initializer = initializer_he_normal(),
                kernel_constraint = constraint_maxnorm(3)) %>%
    # Dropout
    layer_dropout(rate = rate) %>%
    # Second hidden layer
    layer_dense(units = units_2l, activation = "relu",
                kernel_initializer = initializer_he_normal(),
                kernel_constraint = constraint_maxnorm(3)) %>%
    layer_dense(units = 1, activation = "sigmoid")

  model <- keras_model(input, output)

  # Compile model
  model %>% compile(
    loss = "binary_crossentropy",
    optimizer = optimizer_adam(lr = 0.001),
    metrics = tf$keras$metrics$AUC(name = "auc")
  )

  return(model)
}

## Custom colors and theme----
# Custom colors
grey <- "#969696"

# Custom theme
theme_custom <- function() {
  theme_bw(base_size = 28) +
  theme(
    text = element_text(colour = "black"),
    axis.ticks = element_blank(),
    strip.background = element_rect(fill = "black", color = "black"),
    strip.text = element_text(color = "white", size = rel(1),
                              margin = margin(l = 5, r = 5)),

    # Legend
    legend.key = element_rect(fill = "white", color = NA),
    legend.position = "top",
    plot.title = element_text(margin = margin(b = 6), face = "bold")
  )
}

```

```

# *****
# 1) HYPERPARAMETERS TUNING----
# *****
## Create folder to save results----
folder <- "hyperparameter-tuning"
dir.create(folder)
# Create folder to save logs from tensorflow
logdir <- file.path(folder, "logs")
dir.create(logdir)
# Create folder to save model weights
dir.create(file.path(folder, "weights"))

## Prepare data pipeline----
# Remove unused columns
data <- dataset %>% select(-bird, -fold_id)
# Split training (70%) and validation (30%) datasets
val_df <- data[idx, ]
train_df <- data[-idx, ]

# Feature specifications
spec <- train_df %>%
  # Define target as response variable
  feature_spec(target ~ .) %>%
  # Normalise numeric variables
  step_numeric_column(
    all_numeric(), normalizer_fn = scaler_standard()
  ) %>%
  # One hot encode categorical variables
  step_categorical_column_with_vocabulary_list(
    landcover,
    vocabulary_list = as.character(1:10)
  ) %>%
  step_categorical_column_with_vocabulary_list(
    geology,
    vocabulary_list = as.character(1:4)
  ) %>%
  step_indicator_column(landcover, geology) %>%
  # Fit the object
  fit()

# Save feature specifications
saveRDS(spec, file.path(folder, "spec.Rds"))

# Create datasets
train <- train_df %>%
  tensor_slices_dataset() %>% # Convert dataframe to tensor
  dataset_use_spec(spec) %>% # Apply specifications
  dataset_batch(batch_size = 256,

```

```

        drop_remainder = TRUE) %>% # Batch data
dataset_prefetch(1)                # Pre-fetch one batch

# This is used for evaluation, note that trailing elements in the last
# batch are not removed
train_eval <- train_df %>%
  tensor_slices_dataset() %>%
  dataset_use_spec(spec) %>%
  dataset_batch(batch_size = 1024) %>% # Faster during evaluation
  dataset_prefetch(1)

val <- val_df %>%
  tensor_slices_dataset() %>%
  dataset_use_spec(spec) %>%
  dataset_batch(batch_size = 1024) %>% # Faster during evaluation
  dataset_prefetch(1)

# In case of very large datasets, instead of load the data in memory, is
# possible to use one of the following functions:
# make_csv_dataset()
# text_line_dataset()
# sqlite_dataset()

## Define hyperparameters grid----
# Number of units in the first hidden layer
NUM_UNITS_1L <- hp$HParam(
  "num_units_1l",
  hp$Discrete(list(16, 32, 64, 128, 256, 512))
)
# Number of units in the second hidden layer
NUM_UNITS_2L <- hp$HParam(
  "num_units_2l",
  hp$Discrete(list(16, 32, 64, 128, 256, 512))
)
# Dropout rate
DROPOUT <- hp$HParam(
  "dropout",
  hp$Discrete(list(0.2, 0.3, 0.4, 0.5, 0.6, 0.7))
)

# Write hyperparameters' grid, for each combination save training and
# validation
# AUC
with(tf$summary$create_file_writer(logdir)$as_default(), {
  hp$hparams_config(
    hparams = list(NUM_UNITS_1L, NUM_UNITS_2L, DROPOUT),
    metrics = list(hp$Metric("train_auc", display_name = "Train AUC"),
                  hp$Metric("val_auc", display_name = "Val AUC"))
  )
}

```

```
    )
  )
})

## Run hyperparameters tuning experiment-----
# This takes few hours and results might be slightly different due to the
# stochasticity of the model's weight initialization by Keras and the
# random dropout.

# Start tensorboard
tensorboard(logdir)

# Loop through the hyperparameters' grid
trial <- 1
for(units_1l in NUM_UNITS_1L$domain$values) {
  for(units_2l in NUM_UNITS_2L$domain$values) {
    for (rate in DROPOUT$domain$values) {
      cat("Trial", trial, "- units 1L =", units_1l, "- units 2L =",
          units_2l, "rate = ", rate, "\n")
      # Define hyperparameters
      hparams <- dict(
        NUM_UNITS_1L = units_1l,
        NUM_UNITS_2L = units_2l,
        DROPOUT = rate
      )
      cat("Trial", trial, "- units 1L =", units_1l, "- units 2L =",
          units_2l, "rate = ", rate, "\n")
      # Create the model
      model <- getModel(
        units_1l = units_1l,
        units_2l = units_2l,
        rate = rate,
        ds = train,
        spec = spec
      )

      # Save model weights in the following file, file name is coded as:
      # weights_number-units-first-layer_number-units-second-layer_dropout-
      # rate.hdf5
      file_weights <- file.path(
        folder,
        "weights",
        paste0("weights_", units_1l, "_", units_2l, "_",
              sub("\\.", "", rate), ".hdf5")
      )

      # Logs of each model are saved inside the run_dir folder coded as:
      # number_units_first_layer-number_units_second_layer-dropout_rate
```

```

name <- paste(units_1l, units_2l, sub("\\.", "", rate), sep = "-")
run_dir <- file.path(logdir, name)

# Define Callbacks
callbacks <- list(
  # Tensorboard
  callback_tensorboard(log_dir = run_dir),
  # Save model weights only when validation auc improves
  callback_model_checkpoint(
    filepath = file_weights,
    monitor = "val_auc",
    save_best_only = TRUE,
    save_weights_only = TRUE,
    mode = "max",
    verbose = 0
  ),
  # Stop training when validation auc doesn't increase for more than
  # 10 epochs. Also recover best model weights at the end of training
  callback_early_stopping(
    monitor = "val_auc",
    patience = 10,
    verbose = 0,
    mode = "max",
    restore_best_weights = TRUE
  )
)

# Train model
model %>% fit(
  train %>% dataset_shuffle(buffer_size = 1000L), # Shuffle training
data validation_data = val, # Don't shuffle validation data
  epochs = 100,
  callbacks = callbacks,
  verbose = 0
)

# Evaluate model
results <- list(
  train = model %>% evaluate(train_eval, verbose = 0),
  val = model %>% evaluate(val, verbose = 0)
)

# Log values in tensorboard
with(tf$summary$create_file_writer(run_dir)$as_default(), {
  # Write the values used in this trial
  hp$hparams(hparams)
  # Write evaluation metrics in the logs

```

```

        tf$summary$scalar("train_auc", results$train["auc"], step = 1L)
        tf$summary$scalar("val_auc", results$val["auc"], step = 1L)
    })

    trial <- trial + 1
  }
}

# Download the results from tensorboard at the end of the experiment. The
# best model configuration in this run was:
# - Number of units in the first hidden layer: 256
# - Number of units in the second hidden layer: 32
# - Dropout rate: 0.6

# *****
# 2) LEAVE ONE BIRD OUT CROSS VALIDATION----
# *****
## Prepare data for cross validation----
# Create folders to save results
folder <- "leave-one-bird-out-cv"
dir.create(folder)
# Create folder to save model weights
dir.create(file.path(folder, "weights"))
# Create folder to save logs from tensorflow
logdir <- file.path(folder, "logs")
dir.create(logdir)
# Create folder to save feature specifications
dir.create(file.path(folder, "specs"))

# Prepare data, remove unused column
data <- dataset %>% select(-fold_id)
birds <- unique(data$bird)

## Use tensorflow to log results----
BIRD = hp$HParam("bird", hp$Discrete(as.list(birds)))

with(tf$summary$create_file_writer(logdir)$as_default(), {
  hp$hparams_config(
    hparams = list(BIRD),
    metrics = list(hp$Metric("train_auc", display_name = "Train AUC"),
                  hp$Metric("val_auc", display_name = "Val AUC"))
  )
})

## Run cross validation----
# This takes a while and results might be slightly different due to the
# stochasticity of the model's weight initialization by Keras and the

```

```

# random dropout.

# Start tensorboard
tensorboard(logdir)

# Loop through the birds
for(bird in BIRD$domain$values) {
  cat("Testing using bird", bird, "\n")
  hparams <- dict(BIRD = bird)

  train_df <- data[data$bird != bird, ] %>%
    select(-bird) # Remove column bird
  val_df <- data[data$bird == bird, ] %>%
    select(-bird) # Remove column bird

  # Feature specifications
  spec <- train_df %>%
    # Define target as response variable
    feature_spec(target ~ .) %>%
    # Normalise numeric variables
    step_numeric_column(
      all_numeric(), normalizer_fn = scaler_standard()
    ) %>%
    # One hot encode categorical variables
    step_categorical_column_with_vocabulary_list(
      landcover,
      vocabulary_list = as.character(1:10)
    ) %>%
    step_categorical_column_with_vocabulary_list(
      geology,
      vocabulary_list = as.character(1:4)
    ) %>%
    step_indicator_column(landcover, geology) %>%
    # Fit the object
    fit()

  # Save feature specifications
  saveRDS(spec, file.path(folder, "specs", paste0("spec-", bird, ".Rds")))

  # Create datasets
  train <- train_df %>%
    tensor_slices_dataset() %>% # Convert dataframe to tensor
    dataset_use_spec(spec) %>% # Attach specifications
    dataset_batch(batch_size = 256,
                  drop_remainder = TRUE) %>% # Batch
    dataset_prefetch(1) # Pre-fetch to speed up
  computation
  val <- val_df %>%

```

```
tensor_slices_dataset() %>%
dataset_use_spec(spec) %>%
dataset_batch(batch_size = 1024) %>% # Faster during evaluation
dataset_prefetch(1)

model <- getModel(
  units_1l = 256,
  units_2l = 32,
  rate = 0.6,
  ds = train,
  spec = spec
)

# Save model weights, file name will be: weights_removed-bird-id.hdf5
file_weights <- file.path(
  folder, "weights", paste0("weights_", bird, ".hdf5")
)

# Logs of each model are saved inside the run_dir folder coded with the
# name of the removed bird
run_dir <- file.path(logdir, bird)

# Define callbacks
callbacks <- list(
  # Tensorboard
  callback_tensorboard(log_dir = run_dir),
  # Save model weights only when validation auc improves
  callback_model_checkpoint(
    filepath = file_weights,
    monitor = "val_auc",
    save_best_only = TRUE,
    save_weights_only = TRUE,
    mode = "max"
  )
)

# Train model
model %>% fit(
  train %>% dataset_shuffle(buffer_size = 1000L),
  validation_data = val,
  epochs = 20,
  callbacks = callbacks,
  verbose = 0
)

# Recover best weights
model %>% load_model_weights_hdf5(file_weights)
```

```

# Note that trailing element in the last batch are not removed
train <- train_df %>%
  tensor_slices_dataset() %>%
  dataset_use_spec(spec) %>%
  dataset_batch(batch_size = 1024) %>% # Faster during evaluation
  dataset_prefetch(1)

# Evaluate model
results <- list(
  train = model %>% evaluate(train, verbose = 0),
  val = model %>% evaluate(val, verbose = 0)
)

# Log values in tensorboard
with(tf$summary$create_file_writer(run_dir)$as_default(), {
  # Write the values used in this trial
  hp$hparams(hparams)
  # Write evaluation metrics in the logs
  tf$summary$scalar("train_auc", results$train["auc"], step = 1L)
  tf$summary$scalar("val_auc", results$val["auc"], step = 1L)
})
}

## Update results----
# Add column with number of GPS fixes of the removed bird
# From tensorboard download the csv file, name it "lobo-results.csv", and
# save it inside the working folder (i.e. "leave-one-bird-out-cv")
df <- read.csv(file.path(folder, "lobo-results.csv"))

# Count number of observation per bird
count <- data %>% count(bird)
# Match bird columns
ix <- match(df$bird, count$bird)

# Add column with number of fixes of the removed bird and order by
# descending validation AUC
df <- cbind(
  bird = df$bird,
  n = count[ix, "n"],
  df %>% select(-bird)
) %>% arrange(desc(Val.AUC))

# Save updated data frame
write_csv(df, file = file.path(folder, "lobo-results.csv"))

# *****
# 3) SPATIAL BLOCKS CROSS VALIDATION----
# *****

```

```
## Prepare data for cross validation----
# Create folder to save results
folder <- "spatial-blocks-cv"
dir.create(folder)
# Create folder to save logs from tensorflow
logdir <- file.path(folder, "logs")
dir.create(logdir)
# Create folder to save model weights
dir.create(file.path(folder, "weights"))
# Create folder to save feature specifications
dir.create(file.path(folder, "specs"))

# Prepare data, remove unused column
data <- dataset %>% select(-bird)

## Use tensorflow to log results----
FOLD = hp$HParam("fold", hp$Discrete(list(1, 2, 3, 4, 5)))

with(tf$summary$create_file_writer(logdir)$as_default(), {
  hp$hparams_config(
    hparams = list(FOLD),
    metrics = list(hp$Metric("train_auc", display_name = "Train AUC"),
                  hp$Metric("val_auc", display_name = "Val AUC"))
  )
})

## Run cross validation----
# This takes a while and results might be slightly different because of the
# stochasticity of the model's weight initialization by Keras and the
# random dropout.

# Start tensorboard
tensorboard(logdir)

# Loop through the folders
for(fold in FOLD$domain$values) {

  cat("Testing using fold", fold, "\n")
  hparams <- dict(FOLD = fold)

  run_dir <- file.path(logdir, fold)

  train_df <- data[data$fold_id != fold, ]
  val_df <- data[data$fold_id == fold, ]

  # Feature specifications
  spec <- train_df %>%
```



```

# Define target as response variable
feature_spec(target ~ .) %>%
# Normalise numeric variables
step_numeric_column(
  all_numeric(), normalizer_fn = scaler_standard()
) %>%
# One hot encode categorical variables
step_categorical_column_with_vocabulary_list(
  landcover,
  vocabulary_list = as.character(1:10)
) %>%
step_categorical_column_with_vocabulary_list(
  geology,
  vocabulary_list = as.character(1:4)
) %>%
step_indicator_column(landcover, geology) %>%
# Fit the object
fit()

# Save feature specifications
saveRDS(spec, file.path(folder, "specs", paste0("spec-", fold, ".Rds")))

# Create datasets
train <- train_df %>%
  tensor_slices_dataset() %>% # Convert dataframe to tensor
  dataset_use_spec(spec) %>% # Attach specifications
  dataset_batch(batch_size = 256,
                drop_remainder = TRUE) %>% # Batch
  dataset_prefetch(1) # Pre-fetch to speed up
computation
val <- val_df %>%
  tensor_slices_dataset() %>%
  dataset_use_spec(spec) %>%
  dataset_batch(batch_size = 1024) %>% # Faster during evaluation
  dataset_prefetch(1)

model <- getModel(
  units_1l = 256,
  units_2l = 32,
  rate = 0.6,
  ds = train,
  spec = spec
)

# Save model weights, file name is coded as: weights_fold-used-fold.hdf5
file_weights <- file.path(
  folder, "weights", paste0("weights_fold_", fold, ".hdf5")
)

```

```
# Define callbacks
callbacks <- list(
  # Tensorboard
  callback_tensorboard(log_dir = run_dir),
  # Save model weights only when validation auc improves
  callback_model_checkpoint(
    filepath = file_weights,
    monitor = "val_auc",
    save_best_only = TRUE,
    save_weights_only = TRUE,
    mode = "max"
  )
)

# Train model
model %>% fit(
  train %>% dataset_shuffle(buffer_size = 1000L),
  validation_data = val,
  epochs = 20,
  callbacks = callbacks,
  verbose = 0
)

# Recover best weights
model %>% load_model_weights_hdf5(file_weights)

# Note that trailing element in the last batch are not removed
train_eval <- train_df %>%
  tensor_slices_dataset() %>%
  dataset_use_spec(spec) %>%
  dataset_batch(batch_size = 1024) %>% # Faster during evaluation
  dataset_prefetch(1)

# Evaluate model
results <- list(
  train = model %>% evaluate(train_eval, verbose = 0),
  val = model %>% evaluate(val, verbose = 0)
)

# Log values in tensorboard
with(tf$summary$create_file_writer(run_dir)$as_default(), {
  # Write the values used in this trial
  hp$hparams(hparams)
  # Write evaluation metrics in the logs
  tf$summary$scalar("train_auc", results$train["auc"], step = 1L)
  tf$summary$scalar("val_auc", results$val["auc"], step = 1L)
})
```

```

}

## Update results----
# Add column with number of birds and GPS fixes in each fold
# From tensorboard download the csv file, name it "sb-results.csv", and
# save it inside the working folder (i.e. "spatial-blocks-cv")
# !IMPORTANT! Make sure that the output is sorted by ascending fold id
# before download it
df <- read.csv(file.path(folder, "sb-results.csv"))

ids <- df$fold
l <- length(ids)
train_birds <- vector("integer", length = l)
train_n <- vector("integer", length = l)
val_birds <- vector("integer", length = l)
val_n <- vector("integer", length = l)

for (i in 1:l) {
  # Train dataset
  x <- dataset[dataset$fold_id != ids[i], ] %>% count(bird)
  train_birds[i] <- length(x$bird)
  train_n[i] <- sum(x$n)
  # Test dataset
  x <- dataset[dataset$fold_id == ids[i], ] %>% count(bird)
  val_birds[i] <- length(x$bird)
  val_n[i] <- sum(x$n)
}

# Add columns with number of fixes and birds
df <- cbind(
  fold = df$fold,
  train_birds = train_birds,
  val_birds = val_birds,
  train_n = train_n,
  val_n = val_n,
  df %>% select(-fold)
)

# Save data frame with metrics
write_csv(df, file = file.path(folder, "sb-results.csv"))

# *****
# 4) BAGGING----
# *****

## Prepare data----
# Create folder to save results
folder <- "bagging"
dir.create(folder)

```

```
# Create folder to save model weights
dir.create(file.path(folder, "weights"))
# Create folder to save feature specifications
dir.create(file.path(folder, "specs"))

# Remove unused columns
data <- dataset %>% select(-bird, -fold_id)

# Recreate training and validation datasets used in the hyperparameters'
# tuning experiment
train_df <- data[-idx, ]
val_df <- data[idx, ]

## Run bagging procedure-----
n <- 30
# Create empty matrix to store results
output <- matrix(0, nrow = nrow(train_df), ncol = n)

# This takes quite long
set.seed(25)
for (i in 1:n) {
  cat("Bagging", i, "|", n, "\n")
  ix <- sample(nrow(train_df), size = nrow(train_df), replace = TRUE)
  ds <- train_df[ix, ]

  # Feature specifications
  spec <- ds %>%
    # Define target as response variable
    feature_spec(target ~ .) %>%
    # Normalise numeric variables
    step_numeric_column(
      all_numeric(), normalizer_fn = scaler_standard()
    ) %>%
    # One hot encode categorical variables
    step_categorical_column_with_vocabulary_list(
      landcover,
      vocabulary_list = as.character(1:10)
    ) %>%
    step_categorical_column_with_vocabulary_list(
      geology,
      vocabulary_list = as.character(1:4)
    ) %>%
    step_indicator_column(landcover, geology) %>%
    # Fit the object
    fit()

  # Save feature specifications
  saveRDS(spec, file.path(folder, "specs", paste0("spec_", i, ".Rds")))
}
```

```

# Create datasets
train <- ds %>%
  tensor_slices_dataset() %>% # Convert dataframe to tensor
  dataset_use_spec(spec) %>% # Attach specifications
  dataset_batch(batch_size = 256,
                 drop_remainder = TRUE) %>% # Batch
  dataset_prefetch(1) # Pre-fetch to speed up
computation
val <- val_df %>%
  tensor_slices_dataset() %>% # Convert dataframe to tensor
  dataset_use_spec(spec) %>% # Attach specifications
  dataset_batch(batch_size = 1024) %>% # Batch
  dataset_prefetch(1) # Pre-fetch to speed up
computation

model <- getModel(
  units_1l = 256,
  units_2l = 32,
  rate = 0.6,
  ds = train,
  spec = spec
)

# Save model weights, file name will be: weights_trial.hdf5
file_weights <- file.path(
  folder, "weights", paste0("weights_", i, ".hdf5")
)

# Define callbacks
callbacks <- list(
  callback_early_stopping(
    monitor = "val_auc",
    patience = 10,
    verbose = 0,
    mode = "max",
    restore_best_weights = TRUE
  ),
  # Save model weights when validation auc improves
  callback_model_checkpoint(
    filepath = file_weights,
    monitor = "val_auc",
    save_best_only = TRUE,
    save_weights_only = TRUE,
    mode = "max"
  )
)
)

```

```
# Train model
model %>% fit(
  train %>% dataset_shuffle(buffer_size = 1000L),
  validation_data = val,
  epochs = 100,
  verbose = 0,
  callbacks = callbacks
)

# Make predictions
output[, i] <- model %>% predict(keras_array(train_df %>% select(-
target)))
}

## Select threshold----
# Take the mean of the bagging predictions
preds <- apply(output, 1, mean)

# Use the dismo package to get the confusion matrix for varying thresholds
using
# the training dataset
eval <- dismo::evaluate(
  p = preds[train_df$target == 1],
  a = preds[train_df$target == 0]
)

id <- max(which(eval@TPR >= 0.95))
eval@t[id]
eval@TPR[id]

# Reproduce Figure S4
data.frame(
  x = eval@t,
  tpr = eval@TPR,
  tnr = eval@TNR
) %>%
  ggplot(mapping = aes(x = x, y = tpr)) +
  geom_point(mapping = aes(color = "TPR")) +
  geom_point(mapping = aes(y = tnr, color = "TNR")) +
  scale_color_manual(values = c("TPR" = "#f1a340", "TNR" = "#998ec3")) +
  geom_vline(xintercept = eval@t[id], color = "#d73027", linetype =
"dashed") +
  labs(x = "Threshold", y = "Rate", color = "") +
  theme_bw(base_size = 18) +
  theme_custom()
ggsave(
  filename = "tpr-tnr.jpeg",
  device = "jpeg",
```

```

path = folder,
width = 250,
height = 150,
units = "mm",
dpi = 300
)

# *****
# 5) VARIABLE IMPORTANCE AND CURVES----
# *****
# Use the best model of the hyperparameters' tuning experiment for the
variable
# importance and the curves
library(vip)
library(pdp)

## Prepare data----
# Create folder to save results
folder <- "plots"
dir.create(folder)

# Remove unused columns
data <- dataset %>% select(-bird, -fold_id)

# Recreate training dataset used in the hyperparameters' tuning experiment
train_df <- data[-idx, ]

# Load feature specifications
spec <- readRDS(file.path("hyperparameter-tuning", "spec.Rds"))

train <- train_df %>%
  tensor_slices_dataset() %>% # Convert dataframe to tensor
  dataset_use_spec(spec) %>% # Attach specifications
  dataset_batch(batch_size = 1024) %>% # Batch
  dataset_prefetch(1)

## Recreate model----
model <- getModel(
  units_1l = 256,
  units_2l = 32,
  rate = 0.6,
  ds = train,
  spec = spec
)

# Load weights from tuning experiment
model %>% load_model_weights_hdf5(
  file.path("hyperparameter-tuning", "weights", "weights_256_32_06.hdf5")
)

```

```
)

# Sanity check
model %>% evaluate(train)

# Wrap function used to make predictions
pred_wrapper <- function(object, newdata) {
  newdata <- as.data.frame(newdata)
  predict(object, x = keras_array(newdata)) %>%
    as.numeric()
}

## Permutation importance----
set.seed(25)
vi <- vi(
  object = model,
  method = "permute",
  pred_wrapper = pred_wrapper,
  train = train_df %>% select(-target),
  target = train_df$target,
  metric = "auc",
  reference_class = 1,
  type = "difference",
  progress = "text"
)

# Plot Variable Importance
vi %>%
  # Normalise values to percentage and create factor to maintain same order
  mutate(Importance = Importance / sum(Importance),
         Variable = fct_reorder(Variable, Importance)) %>%
  ggplot(mapping = aes(x = Variable, y = Importance)) +
  geom_col(fill = "grey", alpha = 0.8) +
  scale_y_continuous(labels = scales::percent,
                    expand = expansion(mult = c(0.01, 0.1))) +
  coord_flip() +
  labs(y = "Permutation importance", x = "") +
  theme_custom()
ggsave(
  filename = "variable-importance.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)
```



```

# Save variable importance
saveRDS(vi, file.path(folder, "variable-importance.Rds"))

## ICE and PD curves-----
### Continuous variables-----
#### Slope-----
# ICE curves
ice_slope <- partial(
  model,
  pred.var = "slope",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  type = "classification",
  progress = "text"
)

# Save ICE object for slope
saveRDS(ice_slope, file.path(folder, "ice-slope.Rds"))

# PD curve
pdp_slope <- pdp::average_ice_curves.ice(ice_slope)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_slope %>%
  structure(class = "data.frame") %>% # Remove class ice
  filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
  ggplot(mapping = aes(x = slope, y = yhat)) +
  geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
  geom_line(data = pdp_slope, mapping = aes(x = slope, y = yhat),
            color = "black", size = 1) +
  scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
  labs(x = "Slope (°)", y = "P (flight < 200 m a.g.l.)") +
  theme_custom()
ggsave(
  filename = "slope-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)

#### Ibex-----
# ICE curves

```

```
ice_ibex <- partial(
  model,
  pred.var = "ibex",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  type = "classification",
  progress = "text"
)

# Save ICE object for ibex
saveRDS(ice_ibex, file.path(folder, "ice-ibex.Rds"))

# PD curve
pdp_ibex <- pdp::average_ice_curves.ice(ice_ibex)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_ibex %>%
  structure(class = "data.frame") %>% # Remove class ice
  filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
  ggplot(mapping = aes(x = ibex, y = yhat)) +
  geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
  geom_line(data = pdp_ibex, mapping = aes(x = ibex, y = yhat),
            color = "black", size = 1) +
  scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
  labs(x = "Probability of ibex presence",
       y = "P (flight < 200 m a.g.l.)") +
  theme_custom()
ggsave(
  filename = "ibex-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)

#### Northness----
# ICE curves
ice_northness <- partial(
  model,
  pred.var = "northness",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
```

```

    type = "classification",
    progress = "text"
  )

# Save ICE object for northness
saveRDS(ice_northness, file.path(folder, "ice-northness.Rds"))

# PD curve
pdp_northness <- pdp:::average_ice_curves.ice(ice_northness)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_northness %>%
  structure(class = "data.frame") %>% # Remove class ice
  filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
  ggplot(mapping = aes(x = northness, y = yhat)) +
  geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
  geom_line(data = pdp_northness, mapping = aes(x = northness, y = yhat),
            color = "black", size = 1) +
  scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
  labs(x = "Northness", y = "P (flight < 200 m a.g.l.)") +
  theme_custom()
ggsave(
  filename = "northness-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)

#### Wind speed----
# ICE curves
ice_windspeed <- partial(
  model,
  pred.var = "windspeed",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  type = "classification",
  progress = "text"
)

# Save ICE object for wind speed
saveRDS(ice_windspeed, file.path(folder, "ice-windspeed.Rds"))

```

```
# PD curve
pdp_windspeed <- pdp::average_ice_curves.ice(ice_windspeed)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_windspeed %>%
  structure(class = "data.frame") %>% # Remove class ice
  filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
  ggplot(mapping = aes(x = windspeed, y = yhat)) +
  geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
  geom_line(data = pdp_windspeed, mapping = aes(x = windspeed, y = yhat),
            color = "black", size = 1) +
  scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
  labs(x = "Wind speed (m/s)", y = "P (flight < 200 m a.g.l.)") +
  theme_custom()
ggsave(
  filename = "windspeed-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)

#### Chamois----
# ICE curves
ice_chamois <- partial(
  model,
  pred.var = "chamois",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  type = "classification",
  progress = "text"
)

# Save ICE object for chamois
saveRDS(ice_chamois, file.path(folder, "ice-chamois.Rds"))

# PD curve
pdp_chamois <- pdp::average_ice_curves.ice(ice_chamois)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_chamois %>%
  structure(class = "data.frame") %>% # Remove class ice
```

```

filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
ggplot(mapping = aes(x = chamois, y = yhat)) +
geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
geom_line(data = pdp_chamois, mapping = aes(x = chamois, y = yhat),
          color = "black", size = 1) +
scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +
scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
labs(x = "Probability of chamois presence",
     y = "P (flight < 200 m a.g.l.)") +
theme_custom()
ggsave(
  filename = "chamois-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)

#### Slope unevenness-----
# ICE curves
ice_slope_unev <- partial(
  model,
  pred.var = "slope_unev",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  type = "classification",
  progress = "text"
)

# Save ICE object for slope unevenness
saveRDS(ice_slope_unev, file.path(folder, "ice-slope-unevenness.Rds"))

# PD curve
pdp_slope_unev <- pdp::average_ice_curves.ice(ice_slope_unev)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_slope_unev %>%
  structure(class = "data.frame") %>% # Remove class ice
  filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
  ggplot(mapping = aes(x = slope_unev, y = yhat)) +
  geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
  geom_line(data = pdp_slope_unev, mapping = aes(x = slope_unev, y = yhat),
            color = "black", size = 1) +
  scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +

```

```
scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
labs(x = "Slope unevenness",
     y = "P (flight < 200 m a.g.l.)") +
theme_custom()
ggsave(
  filename = "slope-unevenness-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)

#### TPI----
# ICE curves
ice_tpi <- partial(
  model,
  pred.var = "tpi",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  type = "classification",
  progress = "text"
)

# Save ICE object for TPI
saveRDS(ice_tpi, file.path(folder, "ice-tpi.Rds"))

# PD curve
pdp_tpi <- pdp::average_ice_curves.ice(ice_tpi)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_tpi %>%
  structure(class = "data.frame") %>% # Remove class ice
  filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
  ggplot(mapping = aes(x = tpi, y = yhat)) +
  geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
  geom_line(data = pdp_tpi, mapping = aes(x = tpi, y = yhat),
           color = "black", size = 1) +
  scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
  labs(x = "Topographic Position Index",
       y = "P (flight < 200 m a.g.l.)") +
  theme_custom()
ggsave(
  filename = "tpi-curve.jpeg",
```

```

device = "jpeg",
path = folder,
width = 250,
height = 150,
units = "mm",
dpi = 300
)

#### Eastness----
# ICE curves
ice_eastness <- partial(
  model,
  pred.var = "eastness",
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  type = "classification",
  progress = "text"
)

# Save ICE object for eastness
saveRDS(ice_eastness, file.path(folder, "ice-eastness.Rds"))

# PD curve
pdp_eastness <- pdp::average_ice_curves.ice(ice_eastness)

# There are too many ice curves to be plotted, take 1000 randomly
set.seed(25)
ice_eastness %>%
  structure(class = "data.frame") %>% # Remove class ice
  filter(yhat.id %in% sample(nrow(train_df), size = 1000)) %>%
  ggplot(mapping = aes(x = eastness, y = yhat)) +
  geom_line(mapping = aes(group = yhat.id), alpha = 0.1, color = grey) +
  geom_line(data = pdp_eastness, mapping = aes(x = eastness, y = yhat),
            color = "black", size = 1) +
  scale_x_continuous(expand = expansion(mult = c(0, 0.02))) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.02))) +
  labs(x = "Eastness",
       y = "P (flight < 200 m a.g.l.)") +
  theme_custom()
ggsave(
  filename = "eastness-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 150,
  units = "mm",
  dpi = 300
)

```

```
)

### Categorical variables----
#### Land cover----
# ICE curves
ice_landcover <- partial(
  model,
  pred.var = "landcover",
  pred.grid = data.frame(landcover = as.character(1:10)),
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  cats = c("landcover", "geology"),
  type = "classification",
  progress = "text"
)

# Save ICE object for land cover
saveRDS(ice_landcover, file.path(folder, "ice-landcover.Rds"))

# PD curve
pdp_landcover <- pdp::average_ice_curves.ice(ice_landcover)

# Plot curves
ice_landcover %>%
  structure(class = "data.frame") %>% # Remove class ice
  ggplot(mapping = aes(x = landcover, y = yhat)) +
  geom_boxplot(outlier.shape = NA) + # Remove outliers
  coord_cartesian(ylim = quantile(ice_landcover$yhat, c(0.03, 1))) +
  geom_point(data = pdp_landcover, mapping = aes(x = landcover, y = yhat),
            color = "black", size = 1.5) +
  scale_x_discrete(labels = c("1" = "Ochard",
                              "2" = "Forest",
                              "3" = "Bush",
                              "4" = "Scree",
                              "5" = "Other",
                              "6" = "Antropic",
                              "7" = "Marsh",
                              "8" = "Water",
                              "9" = "Rock",
                              "10" = "Glacier")) +
  labs(x = "Landcover", y = "P (flight < 200 m a.g.l.)") +
  theme_custom() +
  theme(axis.text.x = element_text(angle = 90))
ggsave(
  filename = "landcover-curve.jpeg",
  device = "jpeg",
  path = folder,
```



```

width = 250,
height = 150,
units = "mm",
dpi = 300
)

#### Geology----
# ICE curves
ice_geology <- partial(
  model,
  pred.var = "geology",
  pred.grid = data.frame(geology = as.character(1:4)),
  pred.fun = pred_wrapper,
  train = train_df %>% select(-target),
  prob = TRUE,
  cats = c("landcover", "geology"),
  type = "classification",
  progress = "text"
)

# Save ICE object for geology
saveRDS(ice_geology, file.path(folder, "ice-geology.Rds"))

# PD curve
pdp_geology <- pdp::average_ice_curves.ice(ice_geology)

# Plot curves
ice_geology %>%
  structure(class = "data.frame") %>% # Remove class ice
  ggplot(mapping = aes(x = geology, y = yhat)) +
  geom_boxplot(outlier.shape = NA) + # Remove outliers
  coord_cartesian(ylim = quantile(ice_geology$yhat, c(0.03, 1))) +
  geom_point(data = pdp_geology, mapping = aes(x = geology, y = yhat),
             color = "black", size = 1.5) +
  scale_x_discrete(labels = c("1" = "Others",
                              "2" = "Limestone",
                              "3" = "Granite",
                              "4" = "Gneiss")) +
  labs(x = "Geology", y = "P (flight < 200 m a.g.l.)") +
  theme_custom() +
  theme(axis.text.x = element_text(angle = 90))
ggsave(
  filename = "geology-curve.jpeg",
  device = "jpeg",
  path = folder,
  width = 250,
  height = 155,
  units = "mm",

```

```
    dpi = 300  
)
```

GENERAL DISCUSSION

The negative effects of wind energy on wildlife have been extensively studied (Drewitt & Langston, 2006; Goodale & Milman, 2016; Hein & Schirmacher, 2016; Madders & Whitfield, 2006; Voigt, 2021) and there is a broad consensus that additional mortality due to collision with the rotor blades is highly detrimental to K-strategy bird species (Carrete et al., 2009; de Lucas et al., 2012). However, there is a lack of area-wide high-resolution spatial information on habitat use and flight behaviour of vulnerable species which hinders to effectively integrate species protection into strategic windfarm-planning. The predictive models I developed provide crucial information not only for site screening at the beginning of the planning process but also for micro-siting, i.e. the deployment of wind turbines in sites where they pose the least risk to target species.

Modelling potential conflicts

By combining spatially explicit information on habitat selection (Chapter 1) and flight height behaviour (Chapter 3) of the bearded vulture, I developed an evidence-based methodological framework to anticipate areas of potential conflicts with regard to wind energy deployment which, however, can be applied to any detrimental aerial infrastructure (Fig. 1). Several studies have aimed to spatially predict the distribution of potential conflicts arising from wind energy development for different species, either based on GPS locations of radio-tagged animals (Reid et al., 2015; Thaxter et al., 2019; Vasilakis et al., 2016, 2017) or on nest locations (Heuck et al., 2019; Murgatroyd et al., 2021; Tack & Fedy, 2015), but not so many combined information of habitat selection and flight behaviour into a single framework (Reid et al., 2015). This, however, is crucial for comprehensively capturing potential collision risks of vulnerable bird species. Although the collision risk is likely to increase in areas that are highly suitable for a species (Heuck et al., 2019), habitat suitability per se, which is usually calculated as the mere probability of species occurrence, does not include information on flight height behaviour and therefore also includes areas where birds mainly fly far above ground level. To the best of my knowledge, there is only one study that combines these two sources of information to predict areas of conflicts on a large extent (Reid et al., 2015). Compared to their results, my study goes a step further and introduces synthetic, ready-to-use maps for practical application, delineating areas with different “sensitivity” levels from the perspective of bearded vulture conservation. Whereas the “potential conflict map” defines areas where conflicts are likely to occur, the “high-risk conflict map” refines it and facilitates the identification of areas within the species’ range that are suitable for wind energy development but

less likely to be detrimental to the species.

Furthermore, I accounted for the potential future expansion of the bearded vulture population (Chapter 1). This aspect is paramount in the context of wind energy development in areas harbouring species with highly dynamic or expanding populations, where new areas of conflict, neglected by extant knowledge, are very likely to appear. For the bearded vulture in the Alps we are observing an ongoing recolonisation of its former range, confirmed by the number of new territories and the location of new breeding sites that year after year add to the known ones. 2014 was the last year included in the casual observations used to train my models, corresponding to the reference year of the breeding locations considered in the Swiss wind energy concept (Bundesamt für Raumentwicklung ARE, 2017). All new breeding territories that were established after 2014, or relocated over time, appeared within the areas predicted as suitable for the species, except for one location that was very close to it, thus confirming the validity of my approach to anticipate conflict zones.

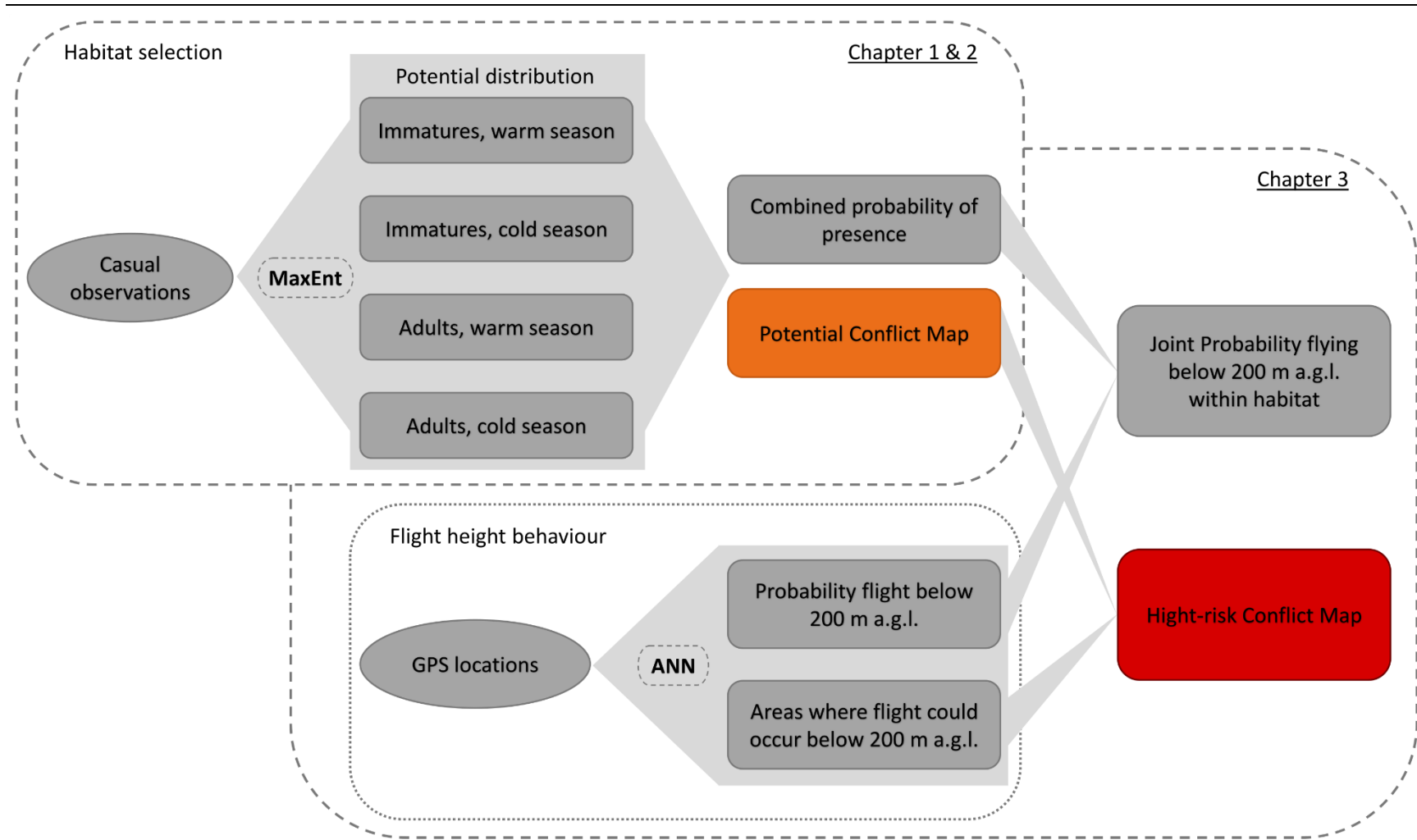


Figure 1: Methodological framework used for predicting conflicts between bearded vultures and wind energy development in the Swiss Alps. In light grey is shown the information flow with the data (ovals) feeding the predictive models (dashed labels, MaxEnt: maximum entropy and ANN: artificial neural network) to produce information on different aspects of bearded vulture ecology and project it as maps (rectangles). Coloured rectangles represent the two main deliveries, i.e. maps delineating areas of potential conflict on a broad scale (orange) and hotspots with high risk conflict (red).

Bearded vulture ecology

My approach also contributes to better understanding bearded vulture ecology. Inhabiting mountain regions in several areas of the Old world, there is a strong evidence that bearded vultures actively select steep slopes, cliffs, and ridges (Hirzel et al., 2004; Reid et al., 2015) which provide sites sheltered from adverse weather conditions and therefore optimal for both roosting and nesting (Hirzel et al., 2004). Moreover, these terrain features are optimal to generate updraughts which can efficiently support the flight of soaring birds and thus characterise areas where they predominantly fly at lower height (Hanssen et al., 2020; Katzner et al., 2012; Murgatroyd et al., 2021). My results corroborate these findings, but further highlight the importance of food supply for low-flying activity (i.e. below 200 m a.g.l., Chapter 3), specifically a high probability of ibex presence (*Capra ibex*), which is considered a proxy for the availability of ibex carcasses. I linked this association to the bearded vulture's ability to detect carrion. Shepard et al. (2011) suggested a similar pattern for the Andean condor (*Vultur gryphus*). Also other vulture species like the griffon vulture (*Gyps fulvus*) have been observed to fly at lower heights over areas with high density of ungulates (Houston, 1974), although they may also be able to locate food sources indirectly by flying high and observing other individuals (Houston, 1974). Bearded vultures instead feed mainly on smaller large bones of medium sized ungulates (Margalida et al., 2009) that might be difficult to detect when flying far above the ground. For this reason, the species may track areas where ibex colonies concentrate at low flight height, regardless of other environmental conditions.

Food supply has also been identified as one of the main drivers of bearded vulture distribution, especially for immature birds (Chapter 1; Hirzel et al., 2004) and adults during the warm season (Chapter 1). Adult bearded vultures however changed their “priorities” during the cold season when climatic conditions became more important, probably because favourable winter conditions are required in the breeding territory. The fact that food supply is of critical importance for both habitat selection and flight height behaviour has a clear direct implication for wind energy development, which should be avoided in areas hosting ibex colonies. Moreover, given its critical importance, food supply should always be considered when investigating a species’ habitat selection. If this information is not available at sufficiently high resolution or in a format useful to run predictive models, a valid proxy may be used instead, for example by modelling the distribution of the main source of food using predictive models and survey data. For this study, for example, I modelled the probability of ibex and chamois occurrence (*Rupicapra rupicapra*) which proved to be a valid proxy for food supply (see supplementary material in Chapter 1).

Methodological challenges

In all Chapters of this dissertation I emphasised key methodological challenges and contributed with novel approaches to address them objectively and efficiently. In Chapter 1 and 3 I built several models to predict either the potential distribution or the flight height behaviour of my target species so as to anticipate potential conflicts arising from detrimental aerial infrastructure. In these contexts, a model that can achieve accurate predictions from existing data is of prime importance, with hyperparameters tuning becoming a prerequisite for increasing model performance. Although several authors have stressed that the use of default values of the hyperparameters may not be optimal (Elith et al., 2010; Merow et al., 2013; Warren et al., 2014; Warren & Seifert, 2011), researchers may be discouraged by the time needed to test numerous hyperparameters combinations, notably when studies are conducted for several species, for multiple datasets, or dataset partitions. The new genetic algorithm for supporting hyperparameters tuning in SDMs that I presented in Chapter 2 drastically reduces computation time while achieving model performances comparable or equal to those obtained by testing all the possible combinations of the default hyperparameters values. Indeed, it leverages the information acquired during the tuning procedure and quickly discovers a valid solution. For example, when testing several methods to reduce the bias introduced by nonsystematic sampling of species data (Chapter 1), each repeated for two seasons, the new algorithm allowed tuning the hyperparameters for each trial in a reasonable time, saving valuable time for further progressing with my research.

A second major challenge was to efficiently analyse large datasets of GPS locations, as used for predicting the flight height of bearded vultures in Chapter 3. I addressed this issue by selecting the *tensorflow* (Martín Abadi et al., 2016) and *Keras* (Allaire & Chollet, 2020) libraries which allow an easy and computationally optimised implementation of state-of-the-art models in the framework of artificial intelligence, tailored for large datasets. I also performed hyperparameters tuning in this context and provided the R code to show how to run such analyses with the *tensorboard* library (Martín Abadi et al., 2015). This code can serve as a template for reproducing similar analyses on large datasets requiring efficient data pipelines, a growing problem given the increasing availability of high temporal resolution GPS data (López-López, 2016).

Finally, I was not only interested on model performance but also on model interpretability to properly elucidate bearded vulture ecological requirements. Machine learning algorithms like MaxEnt, which I used in Chapter 1 of my thesis, can fit complex nonlinear relationships but their interpretation is difficult when many variables are used or when variables are highly correlated

(Halvorsen, 2012; Halvorsen et al., 2015). Therefore, model complexity should be controlled to produce more reliable and better interpretable models (Brun et al., 2020; Halvorsen et al., 2015). I achieved this goal by developing two routines, one for selecting among highly correlated variables and the other for removing variables with low contribution without affecting model performance. Compared to expert-based selection among correlated or low-performing variables, which is particularly difficult when information on the species ecology is lacking, these routines perform their task in an objective way, thus providing reproducible results.

Since reproducibility is also a main prerequisite for credibility of any research, I made all methodological solutions conceived in my dissertation available to the scientific community, first by releasing the package *SDMtune* in CRAN and second by providing the R code to fully reproduce the analysis conducted in Chapter 3.

Further perspectives

Although I believe that my research contributes significantly to the conservation of the bearded vulture in the Swiss Alps in relation to wind energy construction, first, by providing new insights on its ecological requirements, and second, by delivering a ready-to-use spatial planning tool, some questions are still open. There are three points which, in my opinion, deserve to be further explored:

1. The risk of a bird colliding with the blades of a wind turbine depends, among others, on two factors, namely its flight height and the frequency with which it visits the area. While flight height has been studied in detail in the framework of this PhD thesis, information on the areas where bearded vultures concentrate their movement is still missing. Identifying areas that are intensively used by the species would allow discovering where, within the species' habitat, the development of wind energy would most likely be particularly detrimental to the bearded vulture.
2. Understanding if and under which weather conditions bearded vultures perceive and thus can avoid aerial obstacles is of paramount importance to guide mitigation measures. Assessing under which combinations of wind, precipitation and fog collision risk increases would allow providing detailed recommendations to temporarily shut down wind turbines located within the bearded vulture habitat.

-
3. Finally, GPS locations could be used in conjunction with collision risk models (CRMs) to estimate mortality rates at operational and planned wind turbines (Vasilakis et al., 2016, 2017). This information, together with the parameters of the population dynamic model of Schaub et al. (2009) would provide new insights on the overall effects of wind turbine construction on the bearded vulture population.

Conclusions

The long-term large-scale bearded vulture reintroduction programme in the Alps launched in 1978 is one of the successful examples of how humans can restore populations and thus at least partly repair the damage they have caused to nature in the past. Many resources, both financial and human, have been invested during the 43 years of this ongoing project. It is therefore crucial that any decisions affecting the establishing bearded vulture population or its habitat are carefully assessed to avoid any potential risk of jeopardising this project. Since more than one third of the Alpine breeding territories are in the Swiss Alps (Lauper, 2020) Switzerland assumes a critical role in the conservation of the species. Wind energy is rapidly developing worldwide (Lee & Zhao, 2021), and many new wind turbines are to be erected in Switzerland to achieve the country's energy objectives (Bundesamt für Raumentwicklung ARE, 2017). Therefore, as part of the planning process, it is a paramount task to account for potential conflicts with the bearded vulture population, and to guarantee the coexistence of the species with wind energy exploitation. To this end, decision makers should be able to easily access all information needed to evaluate potential risks already during the initial planning stage. To date, the available information is insufficient to safely protect the species (Bundesamt für Raumentwicklung ARE, 2017; Horch et al., 2013). With this thesis I have contributed to filling this gap and developed several maps that provide spatial information tailored to the problem: With Chapter 1 I enhanced the knowledge of the present and future distribution of the species to provide broad information of risk-areas where conflicts with wind energy development can arise. Furthermore, I accounted for differences between seasons and age classes to provide a complete picture encompassing the requirements of all life stages in different seasons. In Chapter 3 I gained new insights on the species flight behaviour and integrated them in the results of Chapter 1. That way I could refine the previously identified risk-areas. All these deliveries are ready to use and easy to combine with any other kind of relevant information in any GIS application. Although my maps have been developed for estimating conflicts with wind energy development, they may also be valid for identifying air space conflicts induced by other aerial

infrastructures such as powerlines, ski lifts and cable cars. However, it is important to stress that these maps are intended to support planning but can in no case replace the required environmental risk assessment. Yet, they can provide general guidance and, for example, help to avoid that resources are invested in areas identified as high risk.

Although I focused my research on the bearded vulture, there are other vulnerable species occurring in the Alps that might also be seriously impacted by wind energy expansion. The golden eagle (*Aquila chrysaetos*) is resident in the Swiss Alps and the griffon vulture is a frequent visitor, especially during the summer. Both perform extensive daily movements and are therefore potentially at high risk. Given the similar movement behavior, the developed maps may also be valid to some extent for those two species, but this requires further investigation. Moreover, there are other mountain bird species that should be considered, like the red billed chough (*Pyrrhocorax pyrrhocorax*), the capercaillie (*Tetrao urugallus*), the black grouse (*Lyrurus tetrix*) and the rock ptarmigan (*Lagopus muta*), with the latter not predominantly suffering from collision risk, but from effective habitat loss through avoiding wind energy infrastructure (Coppes et al., 2020). A more comprehensive study including these species would provide a more comprehensive overview of the problem and a better spatial tool for management. Furthermore, this study is limited to the Swiss Alps, while the problem should be addressed for the entire Alpine Arc. By providing the methodological framework for this purpose, I hope that the neighbouring countries will develop similar models and take measures to protect the species. The bearded vulture is an emblematic species that fascinates many people visiting the Alps. I sincerely wish that the results elaborated in this thesis can contribute effectively to the conservation of this species.

REFERENCES

- Abadi, Martín, Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ... Zheng, X. (2016). TensorFlow: A System for Large-Scale Machine Learning. *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, 265–283.
- Abadi, Martín, Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., ... Research Google. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*.
- Allaire, J. J., & Chollet, F. (2020). *keras; R Interface to "Keras". R package version 2.3.0.0*. <https://CRAN.R-project.org/package=keras>.

-
- Brun, P., Thuiller, W., Chauvier, Y., Pellissier, L., Wüest, R. O., Wang, Z., & Zimmermann, N. E. (2020). Model complexity affects species distribution projections under climate change. *Journal of Biogeography*, **47**(1), 130–142. <https://doi.org/10.1111/jbi.13734>
- Bundesamt für Raumentwicklung ARE. (2017). *Konzept Windenergie. Basis zur Berücksichtigung der Bundesinteressen bei der Planung von Windenergieanlagen*. Bern.
- Carrete, M., Sánchez-Zapata, J. A., Benítez, J. R., Lobón, M., & Donázar, J. A. (2009). Large scale risk-assessment of wind-farms on population viability of a globally endangered long-lived raptor. *Biological Conservation*, **142**(12), 2954–2961. <https://doi.org/10.1016/J.BIOCON.2009.07.027>
- Coppes, J., Braunisch, V., Bollmann, K., Storch, I., Mollet, P., Grünschachner-Berger, V., Taubmann, J., Suchant, R., & Nopp-Mayr, U. (2020). The impact of wind energy facilities on grouse: a systematic review. *Journal of Ornithology*, **161**(1), 1–15. <https://doi.org/10.1007/s10336-019-01696-1>
- de Lucas, M., Ferrer, M., Bechard, M. J., & Muñoz, A. R. (2012). Griffon vulture mortality at wind farms in southern Spain: Distribution of fatalities and active mitigation measures. *Biological Conservation*, **147**(1), 184–189. <https://doi.org/10.1016/j.biocon.2011.12.029>
- Drewitt, A. L., & Langston, R. H. W. (2006). Assessing the impacts of wind farms on birds. *Ibis*, **148**(s1), 29–42. <https://doi.org/10.1111/j.1474-919X.2006.00516.x>
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**(4), 330–342. <https://doi.org/10.1111/j.2041-210x.2010.00036.x>
- Goodale, M. W., & Milman, A. (2016). Cumulative adverse effects of offshore wind energy development on wildlife. *Journal of Environmental Planning and Management*, **59**(1), 1–21. <https://doi.org/10.1080/09640568.2014.973483>
- Halvorsen, R. (2012). A gradient analytic perspective on distribution modelling. *Sommerfeltia*, **35**(1), 1–165. <https://doi.org/10.2478/v10208-011-0015-3>
- Halvorsen, R., Mazzoni, S., Bryn, A., & Bakkestuen, V. (2015). Opportunities for improved distribution modelling practice via a strict maximum likelihood interpretation of MaxEnt. *Ecography*, **38**(2), 172–183. <https://doi.org/10.1111/ecog.00565>
- Hanssen, F., May, R., & Nygård, T. (2020). High-Resolution Modeling of Uplift Landscapes can Inform Micrositing of Wind Turbines for Soaring Raptors. *Environmental Management*, **66**(3), 319–332. <https://doi.org/10.1007/s00267-020-01318-0>
- Hein, C. D., & Schirmacher, M. R. (2016). Impact of wind energy on bats: a summary of our current knowledge. *Human-Wildlife Interactions*, **10**(1), 19–27. <https://doi.org/10.26077/x7ew-6349>
- Heuck, C., Herrmann, C., Levers, C., Leitão, P. J., Krone, O., Brandl, R., & Albrecht, J. (2019). Wind turbines in high quality habitat cause disproportionate increases in collision mortality of the white-tailed eagle. *Biological Conservation*, **236**, 44–51. <https://doi.org/10.1016/j.biocon.2019.05.018>

- Hirzel, A., Posse, B., Oggier, P.-A., Crettenand, Y., Glenz, C., & Arlettaz, R. (2004). Ecological requirements of reintroduced Bearded vultures colonizing the Alps : implications for release policy. *Journal of Applied Ecology*, **41**(6), 1103–1116. <https://doi.org/10.1111/j.0021-8901.2004.00980.x>
- Horch, P., Schmid, H., Guélat, J., & Liechti, F. (2013). *Konfliktpotenzialkarte Windenergie – Vögel Schweiz: Teilbereich Brutvögel, Gastvögel und Vogelschutzgebiete gemäss WZVV. Erläuterungsbericht. Aktualisierung 2013*. Schweizerische Vogelwarte, Sempach.
- Houston, D. C. (1974). Food searching in griffon vultures. *African Journal of Ecology*, **12**(1), 63–77. <https://doi.org/10.1111/j.1365-2028.1974.tb00107.x>
- Katzner, T. E., Brandes, D., Miller, T., Lanzone, M., Maisonneuve, C., Tremblay, J. A., Mulvihill, R., & Merovich, G. T. (2012). Topography drives migratory flight altitude of golden eagles: implications for on-shore wind energy development. *Journal of Applied Ecology*, **49**(5), 1178–1186. <https://doi.org/10.1111/j.1365-2664.2012.02185.x>
- Lauper, M. (2020). Update on the Alpine bearded vulture population and the International Bearded Vulture Monitoring. *InfoGIPETO*, **37**, 2–5.
- Lee, J., & Zhao, F. (2021). GWEC Global Wind Report. In *Global Wind Energy Council*.
- López-López, P. (2016). Individual-Based Tracking Systems in Ornithology: Welcome to the Era of Big Data. *Ardeola*, **63**(1), 103–136. <https://doi.org/10.13157/arla.63.1.2016.rp5>
- Madders, M., & Whitfield, D. P. (2006). Upland raptors and the assessment of wind farm impacts. *Ibis*, **148**(s1), 43–56. <https://doi.org/10.1111/j.1474-919X.2006.00506.x>
- Margalida, A., Bertran, J., & Heredia, R. (2009). Diet and food preferences of the endangered Bearded Vulture *Gypaetus barbatus*: a basis for their conservation. *Ibis*, **151**(2), 235–243. <https://doi.org/10.1111/j.1474-919X.2008.00904.x>
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, **36**(10), 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Murgatroyd, M., Bouten, W., & Amar, A. (2021). A predictive model for improving placement of wind turbines to minimise collision risk potential for a large soaring raptor. *Journal of Applied Ecology*, **58**(4), 857– 868. <https://doi.org/10.1111/1365-2664.13799>
- Reid, T., Krüger, S., Whitfield, D. P., & Amar, A. (2015). Using spatial analyses of bearded vulture movements in southern Africa to inform wind turbine placement. *Journal of Applied Ecology*, **52**(4), 881–892. [https://doi.org/10.1111/1365-2664.12468@10.1111/\(ISSN\)1365-2664.ENDANGEREDSPECIES](https://doi.org/10.1111/1365-2664.12468@10.1111/(ISSN)1365-2664.ENDANGEREDSPECIES)
- Schaub, M., Zink, R., Beissmann, H., Sarrazin, F., & Arlettaz, R. (2009). When to end releases in reintroduction programmes: demographic rates and population viability analysis of bearded vultures in the Alps. *Journal of Applied Ecology*, **46**(1), 92–100. <https://doi.org/10.1111/j.1365-2664.2008.01585.x>
- Shepard, E. L. C., Lambertucci, S. A., Vallmitjana, D., & Wilson, R. P. (2011). Energy Beyond Food: Foraging Theory Informs Time Spent in Thermals by a Large Soaring Bird. *PLoS*

ONE, **6**(11), e27375. <https://doi.org/10.1371/journal.pone.0027375>

- Tack, J. D., & Fedy, B. C. (2015). Landscapes for energy and wildlife: Conservation prioritization for golden eagles across large spatial scales. *PLoS ONE*, **10**(8), e0134781. <https://doi.org/10.1371/journal.pone.0134781>
- Thaxter, C. B., Ross-Smith, V. H., Bouten, W., Clark, N. A., Conway, G. J., Masden, E. A., Clewley, G. D., Barber, L. J., & Burton, N. H. K. (2019). Avian vulnerability to wind farm collision through the year: Insights from lesser black-backed gulls (*Larus fuscus*) tracked from multiple breeding colonies. *Journal of Applied Ecology*, **56**(11), 2410–2422. <https://doi.org/10.1111/1365-2664.13488>
- Vasilakis, D. P., Whitfield, D. P., & Kati, V. (2017). A balanced solution to the cumulative threat of industrialized wind farm development on cinereous vultures (*Aegypius monachus*) in south-eastern Europe. *PLoS ONE*, **12**(2), e0172685. <https://doi.org/10.1371/journal.pone.0172685>
- Vasilakis, D. P., Whitfield, D. P., Schindler, S., Poirazidis, K. S., & Kati, V. (2016). Reconciling endangered species conservation with wind farm development: Cinereous vultures (*Aegypius monachus*) in south-eastern Europe. *Biological Conservation*, **196**, 10–17. <https://doi.org/10.1016/j.biocon.2016.01.014>
- Voigt, C. C. (2021). Insect fatalities at wind turbines as biodiversity sinks. *Conservation Science and Practice*, e366. <https://doi.org/10.1111/csp2.366>
- Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**(2), 335–342. <https://doi.org/10.1890/10-1171.1>
- Warren, D. L., Wright, A. N., Seifert, S. N., & Shaffer, H. B. (2014). Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. *Diversity and Distributions*, **20**(3), 334–343. <https://doi.org/10.1111/ddi.12160>

ACKNOWLEDGMENTS

At first, I would like to thank my supervisors, Veronika Braunisch and Raphaël Arlettaz. It has been a long time since 2014 when I first read about the bearded vulture project in the Conservation Biology website from my computer in Italy. Since then, you gave me the opportunity to join your group and involved me in several projects before I could finally start this PhD in 2017. Thank you for your trust, support, and advice which guided me throughout this journey.

I sincerely thank Olivier Duriez and Loïc Pellissier who accepted to be the external examiners of this thesis, and Chaterine Peichel who agree to chair the defence.

I am grateful to all project partners that provided the financial support: the Swiss Federal Office for Energy, the Swiss Federal Office for the Environment and the following foundations: Parrotia Stiftung, Margarethe und Rudolf Gsell-Stiftung, Alfons und Mathilde Suter-Caduff Stiftung, WWF Switzerland, Beat und Dieter Jutzler Stiftung, University of Bern Forschungsstiftung, Stiftung Dreiklang für ökologische Forschung und Bildung, Sophie und Karl Binding Stiftung, Stiftung Temperatio, Ernst Göhner Stiftung, Steffen Gysel Stiftung für Natur und Vogelschutz and Samy Harshallanos. I would also like to thank the Stiftung Pro Bartgeier, the Swiss Ornithological Institute, the Nationalpark Hohe Tauern, the Vulture Conservation Foundation, Vautours en Baronnies, LIFE GypConnect, LIFE GypHelp, ASTERS, LPO Grands Causses, Parc Naturel Régional du Vercors and the Parco Naturale Alpi Marittime for providing data. Thanks in particular to Franziska Lörcher and Daniel Hegglin for their support and the helpful brainstormings.

The Conservation Biology team deserves special thanks. Many people have been part of this amazing group over the years and it is impossible to mention you all but thank you for the enjoyable coffee breaks and aperos. I am particularly grateful to Jaime Resano-Mayor, Arnaud Barras, and Daniel Slodowicz for their friendship and helpful inputs, and to Olivier Roth and Jean-Yves Humbert for the stimulating discussions. Thanks also to Eléonore Tran for painting the illustrations included in this thesis.

I would like to thank my parents, Marzia Rossi and Andrea Vignali, for their support and for giving me the determination and self-confidence to achieve my ambitions. Big thanks to my grandfather Rolando Rossi who passed on to me his interest in nature and especially in birds.

Acknowledgments

Finally, I want to thank my partner Tina Richard who was always there for me, especially in difficult times, to motivate me, listen to me, help me, advise me, and support me. Thank you for your love and great adventures we shared over the years.

Declaration of consent

on the basis of Article 18 of the PromR Phil.-nat. 19

Name/First Name: Vignali Sergio

Registration Number: 17-127-648

Study program: Ecology & Evolution CUSO

Bachelor Master Dissertation

Title of the thesis: Predicting areas of potential conflicts between bearded vultures (*Gypaetus barbatus*) and wind turbines in the Swiss Alps

Supervisors: Prof. Dr. Raphaël Arlettaz
PD Dr. Veronika Braunisch

I declare herewith that this thesis is my own work and that I have not used any sources other than those stated. I have indicated the adoption of quotations as well as thoughts taken from other authors as such in the thesis. I am aware that the Senate pursuant to Article 36 paragraph 1 litera r of the University Act of September 5th, 1996 and Article 69 of the University Statute of June 7th, 2011 is authorized to revoke the doctoral degree awarded on the basis of this thesis.

For the purposes of evaluation and verification of compliance with the declaration of originality and the regulations governing plagiarism, I hereby grant the University of Bern the right to process my personal data and to perform the acts of use this requires, in particular, to reproduce the written thesis and to store it permanently in a database, and to use said database, or to make said database available, to enable comparison with theses submitted by others.

Place/Date:

Bern, den 15. Juni 2021

Signature

