

Mémoire

Auteur : Vanhussel, Margaux

Promoteur(s) : Magain, Nicolas; 12754

Faculté : Faculté des Sciences

Diplôme : Master en biologie des organismes et écologie, à finalité approfondie

Année académique : 2020-2021

URI/URL : <http://hdl.handle.net/2268.2/12605>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

VANHUSSEL MARGAUX

ASSESSMENT OF ECOLOGICAL FACTORS
INFLUENCING THE WINTER HABITAT SUITABILITY
OF *ERITHACUS RUBECULA* ACROSS EUROPE
THROUGH ECOLOGICAL NICHE MODELING

THESIS FOR THE COMPLETION OF A MASTER
DEGREE IN BIOLOGY OF ORGANISMS AND ECOLOGY,
WITH AN IN-DEPTH FOCUS ON ETHOLOGY



DRAWING : LANGEBRAKE C.

UNIVERSITÉ DE LIÈGE
DEPARTMENT OF BIOLOGY, ECOLOGY,
EVOLUTION
FACULTY OF SCIENCE
SEPTEMBER 2021



SUPERVISOR : LIEDVOGEL M.
CO-SUPERVISOR : MAGAIN N.

Acknowledgments

I would like to take those few lines to express all my gratitude and respect to all the people who helped me realize this master thesis. First of all, I would like to thank Miriam Liedvogel, my supervisor, for giving me this amazing opportunity, including me in her team, but also providing me all the support, freedom, and knowledge to improve the quality of this project. This work would also not have been possible without the great support of Corinna Langebrake who made me discover all the secrets of the robins but also for all her knowledge shared during her field and all the other moments spent by her side. Furthermore, the realization of my model would never have been possible without the support of Guillermo F., Flavien C., and Georg M. whom professionalism and expertise were more than necessary. Thank you also to Nicolas Magain for the co-supervision and support through all the different steps of this thesis. On a personal aspect, I would also like to express my thankfulness towards Philippine for joining and supporting me through the joy and excitement of this adventure and sometimes the stress, discouragement, and doubt a master thesis can make you live. And of course, thank you to my mother who was always supportive of the decisions I took.

Having the opportunity to realize my master thesis in another country and with a team of such passionate people was an amazing experience from which I learned a lot. And in a more general aspect, this project is also the end of wonderful studies, so I would like to thank all the teachers and assistants who shared their passion and all the classmates who contributed to enjoy the best moments and survive the more difficult ones especially through this last year.

Finally, I would like to express my gratitude to the person reading this master thesis, whether for providing feedback, learning more about my subject, or evaluating my work.

Summary

Assessment of ecological factors influencing the winter habitat suitability of *Erithacus rubecula* across Europe through ecological niche modeling – Vanhussel Margaux

2021 - Department of biology, ecology, and evolution with the supervision of M. Liedvogel and N. Magain

Erithacus rubecula (European robin) is a common songbird from Europe and found in North Africa and Western Asia. Easily recognizable with its typical orange throat and melodious song, its winter ecological niche is not fully understood yet. Moreover, the European robin is what we call a "facultative non-breeding partial migrant", meaning that in one population, some individuals will show a migratory behavior whereas others will not. In addition, that status can vary from one year to another for a certain proportion of the population. Across this master thesis, the goal was to determine which abiotic and, less importantly, biotic factors are driving the winter habitat suitability of the European robin and what impact those factors could potentially have on its migratory behavior. I worked with the ecological niche modeling tool in R studio and Maxent/ GLM as algorithms to achieve this goal. The occurrences data were downloaded from the eBird Observation Dataset loaded on GBIF; the abiotic variables were downloaded from the WorldClim website and the biotic variables from the land-cover MOD44B. Once the variables and the algorithm settings were selected, models were fitted, assessed, statistically evaluated, and the habitat suitability maps were plotted. The results showed that even though the abiotic factors were the primary concern of this study, they explain 57% of the winter habitat suitability of the robin, which is less than was expected. Land-cover was thus also greatly contributing to characterize the ecological niche of the robin.

Moreover, the temperature seasonality, the mean temperature of the coldest quarter, and the absence of land-cover had the most significant impact on the habitat suitability. Studies led in Belgium, Spain, and Portugal found that females and juveniles had greater chances to express a migratory behavior. Nevertheless, we can suggest that a high seasonality over the year, temperatures below 0.5°C and above 16.53°C, and a lack of vegetation superior to 40% could highly contribute to greater chances of witnessing migratory behavior of an individual. However, whether those variables impact the robin directly or indirectly requires further investigations which were not realized in the context of this master thesis.

Résumé

Évaluation des facteurs écologiques influençant l'habitat hivernal d'*Erithacus rubecula* en Europe grâce à la modélisation de niches écologiques – Vanhussel Margaux

2021 - Département de biologie, écologie et évolution sous la supervision de M. Liedvogel et N. Magain

Erithacus rubecula (Rouge-gorge) est un passereau fréquemment trouvé en Europe ainsi qu'en Afrique du Nord et en Asie de l'Ouest. Il est reconnaissable à sa gorge orange et son chant mélodieux. Sa niche écologique hivernale, elle, n'est pas complètement connue. De plus, Le rouge-gorge présente un comportement appelé 'migrant partiel non reproducteur facultatif'. Cela signifie qu'au sein d'une même population, certains individus montreront un comportement migratoire alors que d'autres seront résidents. Néanmoins, une proportion variable de cette population va présenter des changements de stratégie d'une année à l'autre. À travers ce mémoire, l'objectif était de déterminer quels facteurs abiotiques, et dans un second temps biotiques, influençaient l'habitat hivernal du rouge-gorge ; et quels impacts ces facteurs pouvaient potentiellement avoir sur le comportement migratoire. Pour répondre à ces questions, j'ai utilisé la modélisation de niche écologique sur R studio à l'aide des algorithmes Maxent et GLM comme outils. Les points de présence des rouges-gorges ont été téléchargé depuis la base de données de eBird Observation, elle-même disponible via GBIF. Les variables abiotiques ont quant à elles été téléchargées depuis le site Worldclim et les variables biotiques depuis le module land-cover MOD44B. Une fois ces variables et les algorithmes sélectionnés, les modèles ont été ajustés. Les résultats montrent que malgré une hypothèse majoritairement axée sur l'importance des variables abiotiques, ces dernières semblent expliquer 57% de l'habitat hivernal du rouge-gorge ce qui est moins qu'attendu. Le couvert végétal (et donc les variables biotiques) participeraient donc plus dans la caractérisation de la niche écologiques hivernale du rouge-gorge que prévu. De plus, il semblerait que la saisonnalité thermique, les températures moyennes du quadrimestre le plus froid ainsi que l'absence de végétation sont les variables présentant le plus d'impact sur la qualité de l'habitat. Bien-sûr, nous savions déjà de par les multiples études réalisées en Belgique, Espagne et Portugal que les femelles et les juvéniles présentaient de plus grandes chances de montrer un comportement migratoire à la sortie de la période de reproduction. Néanmoins, il est envisagé qu'une forte saisonnalité thermique au cours de l'année, des températures hivernales inférieures à 0.5°C et supérieures à 16.53°C ainsi qu'un manque de végétation supérieur à 40% contribuent fortement à observer un comportement migratoire chez un individus. Toutefois, que ces variables impactent directement ou indirectement les rouges-gorges nécessitent de plus amples recherches qui n'ont pas été réalisées dans le cadre de ce mémoire.

Figure and table titles

Figure 1: Habitat occupation of the European robin (*Erithacus rubecula*). Orange: breeding, purple: year-round, blue: non-breeding. (Collar 2020) 2

Figure 2: Distribution of the Northern robin population and direction of migratory routes during autumn migration (adapted from Remisiewicz, 2002)5

Figure 3: Threshold model of the variation between obligate residents and obligate migrants from a partially migrant population, which compose the dichotomous trait. The threshold model describes how a specific but unknown continuously distributed variable (=liability or propensity to migrate) will influence the proportion of facultative partial migrants of a partial migrant population depending on environmental effect and/or individual condition (Pulido et al.,1996; Pulido et al.,2011). 7

Figure 4: Graphical representation of the fundamental and realized ecological niche issued of the interactions between biotic/ abiotic factors and movement capacity; Box 5: Explanation on why ENM technics relies mainly on the fundamental niche.8

Figure 5: Presence points of the robins in December- January - February based on the GBIF database once the points had been cleaned (see section 2.3)11

Table 6: Description of the 19 bioclimatic variables from <http://wordclim.org>.....13

Figure 7: Main steps of the ENM approach. The presence data are used to generate artificial absences and extract correlated variables values to fit the model and assess the winter habitat suitability through the predictions16

Table 8: Selected variable description based on O’Donnell & Ignizio’ s paper (2012). Results range are the values from the birdlife version of each variable corresponding to the best habitat suitability for the robin (based on the threshold, see Table 11a & 12b). Tavg.i is the average temperature (°C) for the given month; Tmax.i and Tmin.i are the maximum and minimum temperatures for a given month; PPTi are the monthly total precipitations for a given month; SD is the standard deviation; Tkavg.i is the average temperature (°K) for the given month.....18

Figure 9: 3D map of Europe to see the terrain variation (image source: Sketchfab, Jerry 2021).....19

Figure 10: a) Mean of the predictions made by the model for the whole Europe version (GLM + Maxent). b) Predictions of the occurrence probabilities by GLM for the whole Europe version. c) Predictions of the occurrence probabilities made by Maxent for the whole Europe version. 1 is the highest probability of occurrence and 0 is the absence of probability. d) Mean of the predictions made by the model for the birdlife version (GLM + Maxent). e) Predictions of the occurrence probabilities by GLM for the birdlife version. e) Predictions of the occurrence probabilities made by Maxent for the birdlife version.....20

Table 11.a: Statistical performances of the algorithms and the model for the whole Europe version.....21

Table 11.b: Statistical performances of the algorithms and the model for the birdlife version.....21

Contents

1. INTRODUCTION	2
1.1 ECOLOGY OF THE EUROPEAN ROBIN	2
1.1.1 DISTRIBUTION RANGE:	2
1.1.2 FOOD DIET AND TERRITORIAL BEHAVIOR	3
1.2 MIGRATORY BEHAVIOR	4
1.2.1 DEFINITION	4
1.2.2 AND FOR THE ROBIN?	5
1.2.3 THRESHOLD MODEL, AN ILLUSTRATION TO FACULTATIVE PARTIAL MIGRATION	7
1.3 NICHE MODELING	8
1.4 OBJECTIVES OF THE MASTER THESIS	10
2. MATERIAL AND METHODS	11
2.1 SPECIES DATABASE	11
2.2 CLIMATIC AND BIOTIC DATABASE	11
2.3 CLEANING PROCESS AND VARIABLES SELECTION	12
2.4 MODELING	13
3. RESULTS	17
3.1 VARIABLES SELECTION	17
3.2 MODEL FIT AND PREDICTIONS	19
4. DISCUSSION	22
4.1 WINTER ECOLOGICAL NICHE CHARACTERIZATION OF THE EUROPEAN ROBIN	22
4.2 BIOLOGICAL CHARACTERIZATION OF THE NICHE	23
4.3 COMPARING MODEL ACCURACY WITH A PREVIOUS STUDY IN SPAIN	23
4.4 NICHE OCCUPATION	24
4.5 WHAT IMPACT DOES THIS HAVE ON MIGRATORY STATUS?	25
4.6 JUSTIFICATION OF THE METHOD AND LIMITS TO THE MODEL	25
5. CONCLUSION AND PERSPECTIVES	28
REFERENCES	
APPENDICES	

**Assessment of ecological factors influencing
the winter habitat suitability of *Erithacus
rubecula* across Europe through ecological
niche modeling**

Margaux Vanhussel

University of Liège

1. INTRODUCTION

1.1 Ecology of the European robin

1.1.1 DISTRIBUTION RANGE:

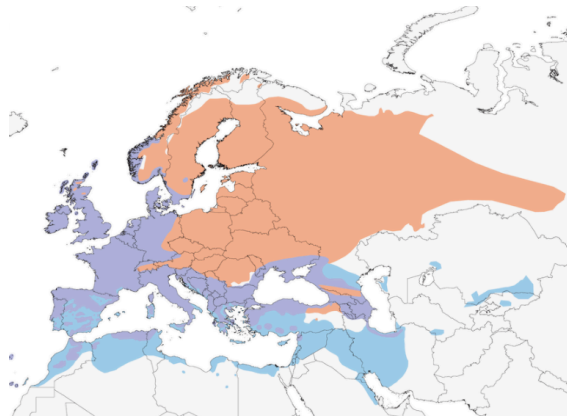


Figure 1: Habitat occupation of the European robin (*Erithacus rubecula*). Orange: breeding, purple: year-round, blue: non-breeding. (Collar 2020)

From the *Muscicapidae* family, the European robin (hereafter “robin”) is one of the most common breeding birds in Europe, often found in our gardens and easily recognizable with its orange breast. Despite its principal occupancy in Europe, its breeding range also includes Western Asia, such as Russia, Turkey, Iraq, and North Africa. This species occupies up to 23 900 000 km² for around 130 to 200 billion individuals (¹). However, this distribution is not constant throughout the year. Depending on the region, the robin population is composed of entirely residents (southern distribution range), migrants (northern distribution range, primarily Norway, Sweden, and Denmark), or both phenotypes occurring in varying frequencies, which is then called a partially migratory population (Adriaensen &

Dhondt, 1990). Most eastern and northern habitats are only suitable during the breeding season, and robins have to migrate to escape unfavorable winter conditions (see Figure 1, orange). Therefore, partial migratory populations, varying in the frequency of their phenotype, possibly according to the latitude, are found in Central and Southern Europe (see Figure 1, purple and blue), and entirely resident populations occur in the South of Europe throughout the year (Collar, 2020).

The phenotypic appearance of the robin is described as "olive-brown above, with orange face and breast fringed by a band of pale blue-grey on neck side to breast side; buff lower flanks and white belly to vent; bill blackish, legs pinkish-brown" (Collar, 2020).

The robin is a species that has successfully adapted to various environments. It is common to find robins in places providing low and bushy vegetation with more shaded regions in central Europe. Robins can also be found in areas without mid/high vegetation cover and mostly dead leaves, or mixed forests with resinous and deciduous species, such as mid-dense forests dominated by oak and cultivated olive trees (e.g., in Spain) and bushes, hedges, riverbanks, or gorges (Catry *et al.*, 2004; Géroudet, 1951). Finally, in places with higher human population density, such as in England, robins are also found in gardens allowing for enough space to keep a reasonable distance from the houses. During the

¹ BirdLife International. (2021). *Species factsheet: Erithacus rubecula*. <http://datazone.birdlife.org/species/factsheet/european-robin-erithacus-rubecula>

winter, the robin prefers gardens to the wooded areas (Géroudet, 1951). The winter habitat is thus fundamental for survival and reproductive rate, especially for the residents since they could protect that territory until the breeding period. According to Adriaensen & Dhondt's hypotheses (1990), scarceness in food availability and lack of protection against predators are the main reasons for not having many resident robins in the woodland during winter. However, these hypotheses are only applicable to the residential robins outside the breeding period, especially in central Europe.

First regarding the food aspect, the robin's diet is composed of invertebrates but can also forage fruits as preparation for migration. These food sources tend to decrease during winter, forcing residents to leave the forests and, in urban areas, move closer to human infrastructures such as dunghills or bird tables with specific food for birds or leftovers and bread. Indeed, by artificially feeding robins in the woods in mid-August and through the whole winter, robins within the artificially fed areas tended to leave the woods for their non-breeding territory later (Adriaensen & Dhondt, 1990; Géroudet, 1951).

Second, vegetation cover contributes to protection from predation, and with the onset of autumn, leaves-fall starts, reducing that protective effect. Populations of migratory robins wintering from the Middle East and Northern Africa tend to occupy farmlands, dense areas in thick plantations, other low vegetation places, and woodlands (Collar, 2020). One study suggests that food availability is the main reason for departure in autumn, and vegetation cover as a shelter for predation being only a second driver (Adriaensen & Dhondt 1990). Another reason for winter habitat choice is the competition between residents/migrants and between sexes; this will be covered in more detail later (Adriaensen & Dhondt, June 1990; Campos *et al.*, 2011).

1.1.2 FOOD DIET AND TERRITORIAL BEHAVIOR

The robin can be defined as a generalist in terms of food choices. Therefore, it feeds on berries such as blackberries, gooseberries, yew, juniper, blueberries but also invertebrates such as larvae of insects such as ants, bees, sawflies, butterflies, but also beetles, earwigs, Diptera, earthworms, etc. They may also feed on acorns or small vertebrates such as fish or lizards (Collar 2020; Géroudet, 1951). Their diet is diverse and can vastly fluctuate depending on the habitat and the season. In Spain, during winter, individuals from a holm oak forest mainly feed on invertebrates such as ants, beetles, larvae, but also acorns in varying quantities. Near farmlands, the invertebrate ratio is almost similar to woodlands, but fruits replace acorns. Finally, the diet in Mediterranean shrubs seems to be mainly composed of juicy fruits and acorn endosperm. Invertebrates are consumed mainly when fruits are scarce (Collar, 2020). According to another study led by Debussche & Isenmann (1985), in the South of France during the non-breeding period, robins will mainly consume invertebrates in early autumn, late winter, and at the beginning of spring; a higher quantity of fruits in mid-autumn; mostly acorn endosperm through winter. In Germany, the shrub population, on the other hand, is mainly eating beetles (60%) but also flies, millipedes, or spiders (Collar, 2020). However, as explained above, that diet is highly variable depending on the habitat. Moreover, if invertebrates are present consistently in their habitats, such as in farmlands or holm oak forests, the availability of the fruits is much more variable depending on one year to another. This inconsistency forces robins to adapt their behavior and number quickly based on the food availability (Telleria *et al.*, 2012).

In correlation of the two Adriaensen & Dhondt's hypotheses (1990), is the territorial behavior. Both males and females express winter territorial behavior through vocalizations, intimidating postures with their orange chest put forward, and, if necessary, attack. This behavior aims to defend and occupy a territory that will allow protection against predators and provide food (Lack, 1946). However, territorial behavior does not come without costs, and vocally defending the territory can increase the risk of detection by predators (Hinde, 1956). Therefore, the winter territory has a different purpose than during the breeding season, and its size is smaller: between 0.27 and 3.12 hectares during the breeding season and only 0.08 to 1.18 hectares in autumn and winter (Adriensen & Dhondt, June 1990). A trade-off between the energy spent on the defense of a significant territory and the ability to occupy a new territory in spring after returning from migration could then explain this size reduction (Lack, 1946).

For migrants that occupy different areas during the breeding and non-breeding period, the situation is different. Residents of the lowest latitudes already occupy many high value territories when migrants arrive. At their arrival in autumn, migrants from breeding ranges at higher latitudes are thus relayed to lower quality areas (Catry *et al.*, 2004; Campos *et al.*, 2011). Moreover, in one study in the South of Portugal, Catry *et al.* (2004) witnessed latitudinal segregation and habitat distribution based on sex, age, and size. Indeed, subordinate individuals as females, juveniles, and smaller males often occupied shrub habitats. Two hypotheses can explain this segregation. The first one is the social dominance hypothesis, where the most dominant males occupy the better-quality territories. The second is the specialized habitat theory, with segregation in habitats adapted depending on sex and morphology. However, this first hypothesis is not confirmed yet since females also possess reproductive hormones generating a territorial behavior outside the breeding period (Campos *et al.*, 2011; Lack, 1946; Schwabl, 1990). In addition to that territorial behavior are the "floaters". Those individuals are not territorial and thus present a weaker strategy. They do not own direct access to food resources and must then often enter occupied territories. They are frequently found in places with more shrubs to protect against predators and the territorial robins whose territory is used for foraging food. Consequentially, those places are usually higher in the density of individuals (Campos *et al.*, 2011; Catry *et al.*, 2004).

1.2 Migratory behavior

1.2.1 DEFINITION

Birds are found across all continents, featuring various adaptations that enabled them to occupy even the most hostile environments, from the freezing poles to the dry deserts. One excellent adaptation to changing seasonal environments is the migratory behavior. Birds can then breed in very productive habitats during the summer while leaving areas with a low carrying capacity during the winter to increase their survival chance. In reaction to that fluctuation, many bird species evolved various migration strategies, differing, for example, in the migration distance from short to very long distances between their breeding and non-breeding sites (Hegemann *et al.*, 2019) or in terms of what sub-populations engage in migratory behaviors.

A transitional stage between obligate residential and obligate migrating populations exists, defined as partial migratory behavior. This behavior is typical among birds, and it occurs when the same population consists of migratory and residential phenotypes (from 1 to 99% of the population). A partial migratory population evolves when food availability fluctuates during the year and is lower but not zero during the non-breeding season (Chapman *et al.*, 2011; Pulido *et al.*, 2011). Endogenous and exogenous factors influence the partial migratory state. A few elements influencing their behavior are specific to the individual, and others are external biotic and abiotic factors. The first one consists of the individual condition/dominance (sex influence, age, and more), competition for resources, predation, etc. The second one depends on environmental factors such as weather conditions (Hegemann *et al.*, 2019). Those two phenotypes of residents and migrants both possess their success rate even though the residential behavior tends to give a better survival rate and reproductive success. This does not mean that both migratory and residential behavior possess an equal pay-off as it mostly depends on the individual itself. Many authors agree that those strategies are probably under a threshold mechanism (Adriensen & Dhondt, Oct. 1990; Pulido *et al.*, 1996).

There are three forms of partial migration. The first one is called non-breeding partial migration. The breeding site is common to the entire population, but some individuals migrate during the non-breeding period. The second one is the partial breeding migration, where the population gathered during the non-breeding period, and some migrate to another site for reproduction. Those two forms can be divided into two intrinsic or extrinsic sub-categories. Either the individuals will keep the same strategy their whole life (migration/residency), or their status will vary each year depending on the environmental and individual conditions (Hegemann *et al.*, 2019). Those two sub-categories are respectively called "obligate partial migration" and "facultative partial migration". Finally, the "skipped breeding partial migration" is a rare form of partial migration. In this case, most of the population migrate to breed, but some individuals will stay on the non-breeding site depending on the years (Chapman *et al.*, 2011).

1.2.2 AND FOR THE ROBIN?

Robins are expressing the facultative non-breeding partial migratory behavior. Indeed, several biotic and abiotic factors can influence its migratory behavior from one year to another, and the robins which migrate will leave the breeding site after the breeding period. Among those factors, sex, age, and size can act as segregating factors (Campos *et al.*, 2011). One extrinsic factor (i.e., factor depending mainly on external influences) would be how females, subordinate to males, tend to migrate in higher frequency and farther than males (Catry *et al.*, 2004). On the other hand, an intrinsic factor (i.e., mainly internal influences acting on the individual) would be that males with higher fat and muscle score tend to remain on the breeding ground year-round (Campos *et al.*, 2011). Moreover, the robin is considered a "niche-tracker". This means it will track the same environmental conditions between its breeding and non-breeding period (Fandos & Tellería, 2020).

The facultative partial migratory behavior explains well how the migratory strategy is more a "best out of bad" strategy. For instance, being resident in a partially migratory population in Belgium is a significant advantage for reproduction

and survival compared to migrants. It increases the chances of survival two to three times since it enables the individuals to keep the most favorable territory and reduces the probability of dying during migration (Adriensen & Dhondt, Oct. 1990). However, harsher winters can reduce survival and reproduction rates, but they remain higher than those of migratory individuals, at least in Belgium. Nevertheless, this illustrates how facultative partial migratory behavior is under selection and is thus an excellent strategy for adapting to changing climatic conditions.

Furthermore, this reproductive success unbalanced by the individual condition maintaining the two phenotypes enables the population to remain successful if conditions for residents are detrimental and *vice versa* (Adriensen & Dhondt, Oct. 1990). Thus, for robins, three categories of migratory behavior occur: the residents who occupy and defend their territory all year round; the "locally wintering birds" staying in the same areas but leaving their territory; and the migrants (Adriensen & Dhondt, Oct. 1990). For this chapter, we will mainly focus on the third category.

Robin migratory behavior across the distribution range generally divides the north-eastern populations dominated by migrants (Belgium, Germany, Norway, etc.) from the South-West, where the frequency of the resident phenotype dominates. Migratory routes can be summarized as the westerly route towards the Iberian Peninsula, the Mediterranean route; the Apennine route; the cross-Mediterranean route from Morocco to Turkey, and potentially a route to the Balkans (see Figure 2; Ambrosini *et al.*, 2016). Nonetheless, those routes are not entirely defined, as some overlaps can occur between robin populations (Remisiewick, 2002). The migration distance is also quite fluctuating because the primary driver of migration seems to be temperature, but some environmental factors such as the Mediterranean Sea can influence the temperature as well, resulting in maritime and continental conditions (Ambrosini, 2016).

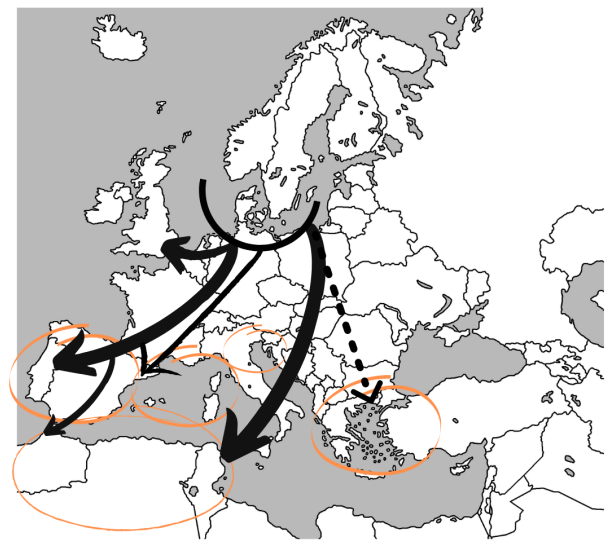


Figure 2: Distribution of the Northern robin population and direction of migratory routes during autumn migration (adapted from Remisiewick, 2002)

With the facultative partial migration and the three strategies, we can therefore understand how flexible and fast robins can adapt to a changing environment. Facultative partial migrants can keep their territory or not, change their migratory distance and strategy and modify their departure time. For instance, blackcap (*Sylvia atricapilla*) has increased its wintering habitat further into Nordic countries, especially in Britain and Ireland. A recent study has shown how artificial feeding by humans and climate change altered their winter ecology and morphology (Berthold *et al.*, 1992; Plummer *et al.*, 2015; Van Doren *et al.*, 2021)).

1.2.3 THRESHOLD MODEL, AN ILLUSTRATION TO FACULTATIVE PARTIAL MIGRATION

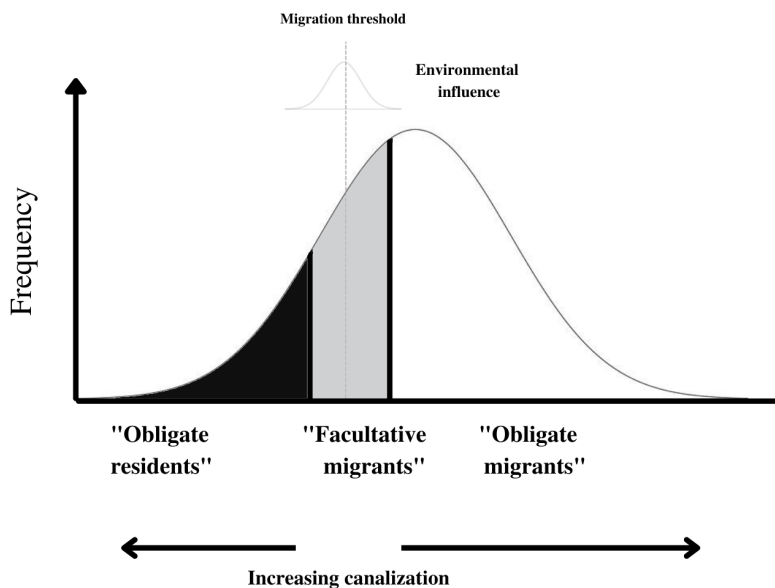


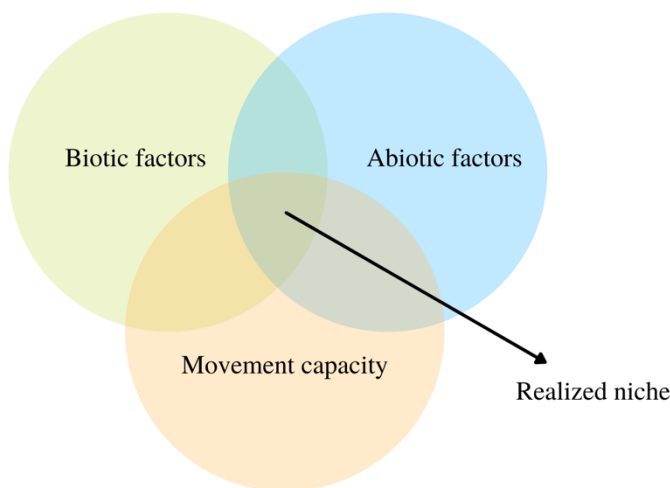
Figure 3: Threshold model of the variation between obligate residents and obligate migrants from a partially migrant population, which compose the dichotomous trait. The threshold model describes how a specific but unknown continuously distributed variable (=liability or propensity to migrate) will influence the proportion of facultative partial migrants of a partial migrant population depending on environmental effect and/or individual condition (Pulido *et al.*,1996; Pulido *et al.*,2011).

We now know that endogenous and exogenous factors influence partial migration. This thesis will mainly discuss exogenous factors, but it is also essential to understand endogenous factors to analyze partial migration within the context of the threshold model.

The threshold model for partial migration is a "(quantitative) genetic model to describe the inheritance and evolution of the incidence of migratory behavior, i.e., whether a bird is migratory or resident" (Pulido *et al.*, 2011). The model assumes that the liability (propensity to migrate) follows a Gaussian distribution that determines the expression of the dichotomous trait (the migratory phenotype) through the threshold. The threshold is genetically determined and can be influenced by environmental factors and individual conditions (see Figure 3). This progression is led by an increased canalization of the phenotype plasticity and decreased environmental sensitivity and *vice versa* (Pulido *et al.*, 2011). According to the model, the migration threshold is where canalization is at its minimum, and this provides an area of a facultative partial migratory phenotype where individuals can change their strategy depending on individual and environmental conditions. In a single migrant population, the migratory behavior of one individual is not visible until its liability (= a specific but unknown variable) crosses the threshold or the threshold shifts due to environmental factors, such as temperature, resource availability, etc. In figure 3, the "obligate residents" are the farthest individuals from the threshold and thus, keep their strategy under unfavorable conditions for their phenotype, e.g., this would be colder winters for residents. The same is true for the "obligate migrants" who will also not be as sensitive to the environmental conditions and keep their migratory behavior also under the least favorable environmental conditions. If individuals in that "facultative migrant" window possess a liability constrained, these birds are more plastic, and their migratory behavior can more easily shift between resident and migratory depending on environmental conditions (Pulido *et al.*,1996; Pulido *et al.*, 2011).

The threshold model is thus a great concept to explain the evolution of migratory traits in a facultative partial migratory population. In an attempt to study the inheritance of migratory behavior in robins, Biebach (1983) characterized nocturnal migratory restlessness behavior (= Zugunruhe) of selected and crossbred lines, specifically lines selected for migratory (mig x mig) or resident behavior (res x res), as well as crossbred birds (mig x res). These results showed that offspring in migratory selection lines resembled their parental phenotype in 89%; whereas offspring in selection lines for resident behavior, only 53% of the offspring resembled the parental phenotype. Those numbers indicate how a single phenotype in a sub-population can induce both migratory and residential behavior. Having both phenotypes in the progeny led to the conclusion that a genetic polymorphism was the source with an inheritance at 0.52 of the migratory behavior for that population of European robin. The rest of that ratio is then probably driven by environmental factors (Biebach,1983).

1.3 Niche modeling



The use of those abiotic variables in the ENM has the purpose to get close to the species' actual distribution while keeping a reasonable degree of freedom. This is expressed under so called "species response curves". Together, those variables selected as the most important for characterizing the species ecological niche constitute an ecological model. This allows to predict the species ecological niche as close to reality as possible. In addition to abiotic factors, biotic factors and the movement capacity of the species will still influence its distribution and sometimes not allow a species to establish itself in an otherwise physiologically suitable area. So, those results will impact the species' distribution by limiting it, such as the competition for a specific site, regulating it by altering its physiology or acting as a resource such as consumable compounds (Guisan *et al.*, 2017; Jiménez *et al.*, 2019).

Figure 4: Graphical representation of the fundamental and realized ecological niche issued of the interactions between biotic/ abiotic factors and movement capacity; Box 5 : Explanation on why ENM technics relies mainly on the fundamental niche

On our planet, different habitats are occupied by various species populating particular niches. The ecological niche concept can explain species complexity and coexistence, which Hutchinson describes as a "quantitative description of the range of environmental conditions that allow a population to persist in some location" (Hutchinson, 1957). The characterization of an ecological niche is twofold: first as fundamental niche through abiotic conditions necessary for a species' survival; and as realized niche, which is the correlation between a species' abiotic, biotic values and their

movement capacity (see Figure 4). Those two parts constitute the "duality principle of Hutchinson," which is the correlation between the geographical areas adapted to a species and the places where the species will occur. Indeed, even though a species has specific areas where the environmental conditions allow it to live (fundamental niche), other factors such as interspecies competition, other species presence, etc. prevent it from occupying that place (realized niche; Colwell & Rangel, 2009; Zurell *et al.*, 2020).

This concept is an important tool in studying a species repartition at past/present/future times. But also, the comprehension of its ecology, invasive or endangered character, or even how it could react to climate change.

This technique is called "ecological niche modeling" (ENM) or "species distribution model" (SDM). From now on, we will only use the term "ENM" in the present work.

For determining if a species can be found in a particular area, three conditions need to be met, as illustrated in Figure 4:

1. The species must have reached the area in question (see "movement capacity from Figure 4) (Guisan *et al.*, 2017).
2. The abiotic environment must be suitable for the ecophysiological conditions of the species, also named *habitat suitability* which means that environmental conditions must allow for the species' survival and its maintenance (Guisan *et al.*, 2017).
3. The biotic environment must be appropriate to the species. i.e., the species must maintain itself and survive despite the interactions with other species sharing the same physical environment. The maximal carrying capacity, for instance, is one of the conditions. It is the maximal capacity of individuals per species that an area can tolerate before the species starts to decline (Guisan *et al.*, 2017).

From the perspective of the fundamental niche, individuals are spreading around an environment according to a specific environmental gradient and following their physiological adaptations acquired through time. Indeed, each species (and in fact each individual) has its tolerance interval to different abiotic variables and disperses itself from those variables, but this is beyond the scope of our study (Smith *et al.*, 2019). Most of the species' abundance is found within the most suitable values and gradually decreases when getting away from those values, following a Gaussian curve though many other forms can arise, such as sigmoids or unimodals. Then, each variable characterizing the niche has a particular effect on the species' physiology, and a synergic result can appear from the interactions between several variables.

1.4 Objectives of the master thesis

This master's thesis is complementary to a field study characterizing partial migratory behavior in individually phenotyped (radiotags and color bands) and genotyped (whole-genome resequencing) European robins (*Erithacus rubecula*) in Northern Germany, where this behavior's genetic basis and dynamics are investigated. The essential purpose of this thesis is to use ecological niche modeling (ENM) as a tool to describe winter habitat suitability of the robin and explain which habitat favors which migratory behavior, where it is beneficial to stay, and, if necessary, where to leave. Through my master thesis, I will thus try to clarify several questions about the robin's ecology. First, which abiotic factors form its winter ecological niche participate and constrain the robin to certain places of its total habitat. Second, which biological explanations can be connected to those environments (such as physiology, food, and habitat type). And finally, study if my results can participate in the explanation of the facultative partial migratory behavior.

Citizen science observation data, used in this study since they are easily accessible and numerous, will serve to evaluate the overall environmental conditions that allow the robin to spend the winter across whole Europe. Once I know which predictors to use, I will model and predict the winter habitat suitability. This will enable the characterization of the winter habitat and evaluate the importance of the variables. Finally, Telleria *et al.* (2012) have already studied the winter habitat suitability of the robin in Spain with ring recoveries and validated the niche modeling method based on monitoring efforts throughout Spain. Since we already have specific knowledge of the winter habitat in that part of Europe, I will use their results to evaluate the accuracy of my predictions.

2. MATERIAL AND METHODS

2.1 SPECIES DATABASE

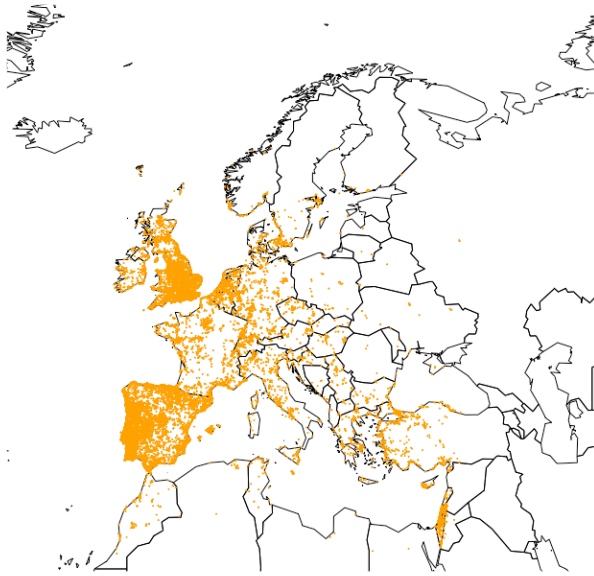


Figure 5: Presence points of the robins in December- January - February based on the GBIF database once the points had been cleaned (see section 2.3)

This study focuses on the populations located in Europe and sporadically in North Africa outside the breeding period (see Figure 5; longitude: [-10,38] and latitude: [28,64] following the EPSG 4326 projection) (Collar, 2020). Occurrence data were downloaded in GBIF (Global Biodiversity Information Facility), from December to February between 1970 and 2019 ⁽²⁾. These data are recorded from the EOD - eBird Observation Dataset and then loaded on GBIF. The raw database contained 80 070 observations.

GBIF is an international network and data infrastructure providing open access data for everyone, requiring a dataset about a species occurrence and abundance for any projects or research, whether terrestrial or marine ⁽³⁾. However, GBIF only provides presence points, whereas many models also

need absence points to correct predictions. One workaround solution is to generate "artificial" absence points called background data or pseudo-absence (Guisan *et al.*, 2017).

2.2 CLIMATIC AND BIOTIC DATABASE

Bioclimatic variables (from 1970 to 2000) used for the niche modeling analyses were extracted from Worldclim v.2.1 (<https://worldclim.org/>; Fick & Hijmans, 2017). Worldclim provides a set of climate maps available at resolutions ranging from 1 km² to 340 km². It includes three types of maps. Firstly, there is the "historical climate data" that comprises minimal/maximal and mean temperatures (°C), precipitations (mm), solar radiation (kJ m⁻² day⁻¹), wind speed (m s⁻¹), and water vapor pressure (kPa). Secondly, 19 bioclimatic variables (see Table 6) are derived from the "historical climate data". For instance, isothermality (BIO 3) is the ratio between the mean diurnal range and the temperature annual range * 100. Another one is the mean temperature of the coldest quarter (BIO 11) which is the "approximate mean temperatures that prevail during the coldest quarter" (O'Donnell & Ignizio, 2012). Thirdly, there is the "future climate data" where min/max temperature and precipitations are averaged depending on the one hand on the nine global climate

² GBIF.org (14 October 2020) GBIF Occurrence Download <https://doi.org/10.15468/dl.amqum6>

³ GBIF. (n.d.). What is GBIF?. <https://www.gbif.org/what-is-gbif>

models using mathematical equations to characterize the transfer of energy occurring in the Earth system ⁽⁴⁾. Finally, for the future, four shared socio-economic pathways, « Socio-economic scenarios used to derive emissions scenarios without and with climate policies » for over 20 years (2021-2040, 2041-2060, 2061-2080, 2081-2100) ⁽⁵⁾. In the context of my research project, I focused on those mentioned above 19 bioclimatic variables at a 10 km resolution.

Land-cover variables are further relevant to be included in the model. Here I used variables from the Annual global automated MODIS vegetation continuous field between 2000 and 2010 (DiMiceli *et al.*, 2011), kindly provided by our collaborator Guillermo Fandos. Those variables represent the percentage of tree cover, non-tree vegetation, landcover, landcover diversity, and bare graduations of the Earth's surface at a 1 km resolution.

2.3 CLEANING PROCESS AND VARIABLES SELECTION

The following steps were made with the program R studio (R version 4.0.5 (2021-03-31)).

a) Variables selection

In a first step, all values needed to be adapted to the extent of Europe. The “crop” function from the *raster* package (Robert, 2021) was applied to obtain variable maps to the same extent (corresponding to the European land-cover variable extent).

Since some of the 19 bioclimatic variables are correlated, if the correlation between multiple variables is too high, this will exaggerate the variance of regression parameters and induce bias in selecting the best predictors (Dormann *et al.*, 2013). In order to avoid these multicollinearity issues, first I made a principal component analysis (PCA) and calculated a correlation matrix. Second, I selected variables based on robin ecology knowledge (as recommended in Fourcade *et al.*, 2018) and removed all the variables having a Pearson correlation greater than 0.7 (as recommended in Dormann *et al.*, 2013). Third, after variable reduction/condensation, I chose which variables from each correlated group were most suitable for explaining robin ecology.

b) Presence points cleaning

The occurrence database was cleaned from erroneous data such as observations near an institution, centroids, missing values, or duplicates. For this, the function “cleanCoordinate” from the *CoordinateCleaner* package was used (Zizka *et al.*, 2019). In addition, only one occurrence per pixel of the 10km resolution was kept with the *dismo* package (Hijmans *et al.*, 2020).

⁴ NOAA (National Oceanic and Atmospheric Administration) Climate. (n.d.). *Climate Models*. <https://www.climate.gov/maps-data/primer/climate-models>

⁵ WordClim. *Future climate data*. (n. d.). <https://worldclim.org/data/cmip6/cmip6climate.html>

Bio 1: <u>Annual mean temperature</u> (°C)
Bio 2: <u>Mean diurnal range (mean of monthly (maximal temperature – minimal temperature))</u> (°C)
Bio 3: <u>Isothermality (BIO2/BIO7) (*100)</u>
Bio 4: <u>Temperature seasonality (standard deviation *100)</u>
Bio 5: <u>Maximum temperature of warmest month</u> (°C)
Bio 6: <u>Minimum temperature of coldest month</u> (°C)
Bio 7: <u>Temperature annual range (BIO5–BIO6)</u> (°C)
Bio 8: <u>Mean temperature of wettest quarter</u> (°C)
Bio 9: <u>Mean temperature of driest quarter</u> (°C)
Bio 10: <u>Mean temperature of warmest quarter</u> (°C)
Bio 11: <u>Mean temperature of coldest quarter</u> (°C)
Bio 12: <u>Annual precipitation</u> (millimeters)
Bio 13: <u>Precipitation of wettest month</u> (millimeters)
Bio 14: <u>Precipitation of driest month</u> (millimeters)
Bio 15: <u>Precipitation seasonality (coefficient of variation)</u>
Bio 16: <u>Precipitation of wettest quarter</u> (millimeters)
Bio 17: <u>Precipitation of driest quarter</u> (millimeters)
Bio 18: <u>Precipitation of warmest quarter</u> (millimeters)
Bio 19: <u>Precipitation of coldest quarter</u> (millimeters)

Table 6: Description of the 19 bioclimatic variables from <http://wordclim.org>

2.4 MODELING

a) Model fitting

To model the species ecological niche, 10 cross-validations were realized using 70% of the data to calibrate the model and 30% to evaluate it. The evaluation metrics were the area under the curve (AUC)/ true skill statistic (TSS) /Kappa of Cohen (Allouch *et al.*, 2006; Freeman & Moisen, 2008; Jiménez-Valverde, (2011).

To calibrate the models, two algorithms were chosen: GLM from the *stats* package (R core team, 2021) and Maxent from the *maxnet* package (Phillips, 2017). The explanatory variables were the bioclim and landcover variables, and the response variables were the presence/"artificial" absences. For GLM, the data were run under the quadratic predictor mode and the binomial family. Furthermore, the primary setting to choose for Maxent was the model complexity, also called "feature classes". As the best setting is not uncommonly the one with the most complicated features (i.e., functions and transformations applied to the variables), Maxent was run with the "ENMevaluate" function (*ENMeval* package; Muscarella *et al.*, 2014) to assess the most suitable one (i.e., the one with the lowest Akaike information criterion

(AIC) score). For the parameters, the method used was the checkerboard2 and the feature classes 'L', 'LQ', 'LQP', 'LQPTH' with the algorithm present in the java software maxent.

b) Background and pseudo-absence points

As these techniques also require absences or background point, the idea was to try two ranges of "artificial" absences; one applied to the whole Europe and North Africa and the second to the birdlife range. Then see if it was influencing the predictions. Two types of "artificial absences" were compiled: one by creating 10 000 background points with the "randomPoints" function (*dismo* package; Hijmans *et al.*, 2020) on the whole of Europe (Robert *et al.*, 2020), and the second version using pseudo-absence data created with the non-breeding distribution area of the robin, called "birdlife range ". The necessary information for this birdlife version was obtained through a request to the Birdlife website ⁽⁶⁾ under an ESRI File Geodatabase format (data are under the Geographic Coordinate System and World Geodetic System 1984 (WGS84)) from which seasonal numbers (coded 1 for resident; 2 for breeding; 3 for non-breeding; 4 for passage) were extracted. As our study focused on winter habitat, we only retained data for seasonal numbers 1 and 3. To generate 10 000 absence points within the birdlife range with the "randomPoints" function, a mask from the polygons corresponding to the seasonal numbers 1 and 3 was created. The 10 000 absence points were generated inside that "birdlife range" (Robert *et al.*, 2020). For the GLM, a 10 km buffer was applied around the presence points, and the pseudo-absences were randomly generated outside those buffer points. This buffer was not applied for Maxent, and all 10 000 background points were randomly generated inside the "birdlife range ". The bioclim and land-cover variables from the pseudo-absence and background points were then extracted for both versions of background and pseudo-absence points (see Figure 7 for visual explanation of the method).

c) Predictions

For each version, GLM and Maxent models were individually used to predict the habitat suitability. In order to account for model uncertainty (Araújo & New, 2007), a consensus model was generated by applying a mean, weighted mean, and median (see Figure 7 for visual explanation). The two algorithms and the model thus generated predictions at the present time on Europe/ North Africa and the birdlife range.

Moreover, once the models were generated with both algorithms, their response curves were plotted. Based on the threshold estimated from the sensitivity and specificity and those response curves, the minimal and maximal values characterizing the winter habitat suitability were extracted with the "partial_response" function (*mecofun* package; Zurell, 2020) and the mean of each min./max. value for GLM and Maxent was calculated.

d) Model assessment and statistical evaluation

The thirty other percent of the "data Split Sample" was used here for that assessment. Firstly, an Akaike Information Criterion (AIC) was performed for the GLM, and both algorithms and model predictions were evaluated with the "evalSDM" function (*mecofun* package; Zurell, 2020) with the AUC/TSS/Kappa/Sensitivity/Specificity tests. Depending on the model used, the estimation of the habitat suitability threshold (i.e., where it is suitable for a robin to be or not)

⁶ BirdLife International. (n.d.). Species distribution data request. <http://datazone.birdlife.org/species/requestdis>

can vary. It is then necessary to identify that threshold for each model via the MaxSens+Spec method (as recommended in Liu et al., 2013 for presence-only data) with the “threshold.method : evalSDM” function (*dismo* package; Robert *et al.*, 2020). That function maximizes the sensibility and specificity sum by reducing the error rate for the true positives and true negatives errors as much as possible. Moreover, with the “ENMevaluate” function (*ENMeval* package; Muscarella *et al.*, 2014), I also evaluated the contribution of the variables to the model with Maxent.

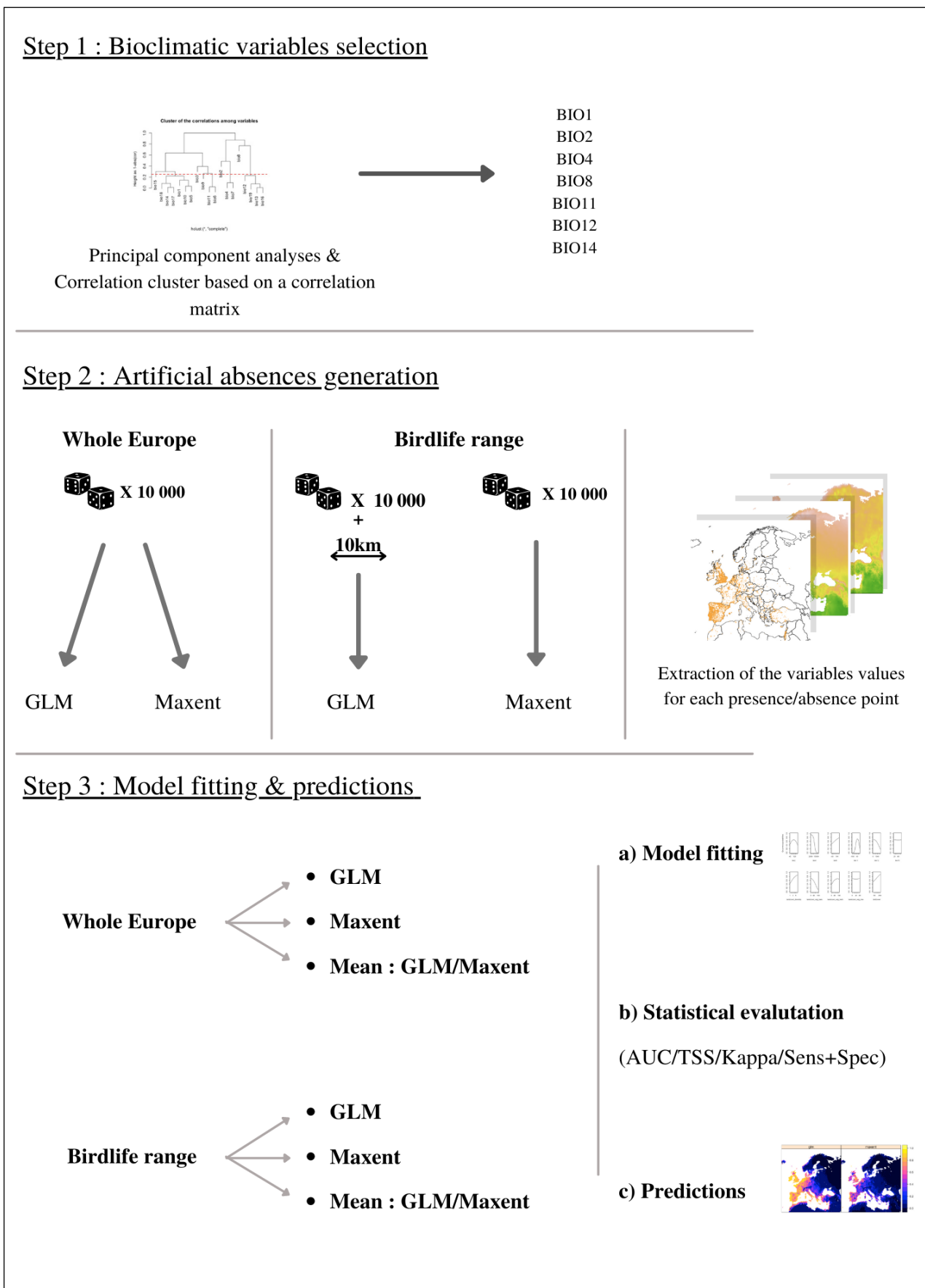


Figure 7: Main steps of the ENM approach. The presence data are used to generate artificial absences and extract correlated variables values to fit the model and assess the winter habitat suitability through the predictions

3. RESULTS

3.1 VARIABLES SELECTION

PCA was used to graphically illustrate the relationship and interaction of the 19 bioclimatic variables. The two first components explained 74,7% of the data (PCA1 = 0,487; PCA2 = 0.260). Based on the correlation circle and the correlation cluster (see Appendix 1), there were five main clusters (BIO 15- BIO/14/17/18/1/10/5 - BIO 9/6/11 - BIO 4/7 - BIO 12/19/13/16) and three independent variables (BIO 2/3/8) below the 0.7 threshold from the Pearson correlation. Following the ecology of the robin (Fourcade *et al.*, 2018) and the objectives of this thesis, one variable from each cluster and two out of the three independent variables were chosen: **BIO 1/2/4/8/11/12/15** (see Table 8.).

The mean values from GLM and Maxent response curves from the birdlife version were used to provide a general idea of the habitat suitability range of the robin for each variable (see Table 8 and Appendices 3 & 4 for more details). **Bio 1** represented the average amount of energy the ecosystem will receive throughout the year (O'Donnell & Ignizio, 2012). This variable was a generalist that allowed to highlight how yearly temperature influences the repartition of the robin. The habitat suitability values were [-3,38°; 19,74°], and Maxent did not react to that variable. **Bio 2**, is the average value between the monthly maximal and minimal temperatures over a year. This indicated if the species could exist in an extensive temperature range or not (O'Donnell & Ignizio, 2012). The value range of the habitat was: [4,67°; 15,54°]. **Bio 4** evaluated the fluctuation of temperature over each year. Based on the ratio between the studied temperature variation over a certain period compared to the standard deviation and the monthly average temperature. The higher the value, the more variability of temperature in the environment (O'Donnell & Ignizio, 2012). For both GLM versions, the probabilities decreased gradually following a sigmoid. This variable, based on Maxent, was contributing at 54,23% and 43,20% to the model depending on the version (*see* Appendices 2a & 2b). The value range of the habitat was: [216,54; 710,25]. **Bio 8** was the mean temperature values based on the three consecutive months with the highest cumulative precipitation total of the year (O'Donnell & Ignizio, 2012). The value range of the habitat was: [-3,23°; 23,20°]. **Bio 11**, just as Bio 8, was calculated based on the three coldest months of the year (O'Donnell & Ignizio, 2012). Its value range of the habitat was: [0,48°; 16,53°].

Variable code	Variable description and range	Units	Calculation	Results range
BIO 1	Annual mean temperature [-5°; 25°]	°C	$Bio\ 1 = \frac{\sum_{i=1}^{i=12} T_{avg_i}}{12}$	[-3,38°; 19,74°]
BIO 2	Annual mean diurnal range [2°; 18°]	°C	$Bio\ 2 = \frac{\sum_{i=1}^{i=12} (T_{max_i} - T_{min_i})}{12}$	[4,67°; 15,54°]
BIO 4	Temperature seasonality (<i>coefficient of deviation</i>) [200; 1400]	°K * 100	$Bio\ 4a = \frac{SD\{T_{kavg_1}, \dots, T_{kavg_{12}}\}}{(Bio\ 1 + 273.15)} \times 100$	[216,54; 710,25]
BIO 8	Mean temperature of the wettest quarter [-10°; 25°]	°C	$Bio\ 8 = \frac{\sum_{i=1}^{i=3} T_{avg_i}}{3}$	[-3,23°; 23,20°]
BIO 11	Mean temperature of the coldest quarter [-20°; 20°]	°C	$Bio\ 11 = \frac{\sum_{i=1}^{i=3} T_{avg_i}}{3}$	[0,48°; 16,53°]
BIO 12	Annual precipitation values [0; 3500]	mm	$Bio\ 12 = \sum_{i=1}^{i=12} PPT_i$	[14,95; 1454,74]
BIO 15	Precipitation seasonality [0; 120]	Mm*100	$Bio\ 15 = \frac{SD\{PPT_1, \dots, PPT_{12}\}}{1 + (Bio\ 12 / 12)} \times 100$	[7,94; 123,21]

Table 8. Selected variable description based on O'Donnell & Ignizio's paper (2012). Results range are the values from the birdlife version of each variable corresponding to the best habitat suitability for the robin (based on the threshold, see Table 11a & 12b). T_{avg_i} is the average temperature (°C) for the given month; T_{max_i} and T_{min_i} are the maximum and minimum temperatures for a given month; PPT_i are the monthly total precipitations for a given month; SD is the standard deviation; T_{kavg_i} is the average temperature (°K) for the given month

Bio 12 was the total amount of precipitation that occurred. The value range of the habitat was: [14,95; 1454,74]. This could have given a clue on the importance of water availability for the robin (O'Donnell & Ignizio, 2012). For the results, the higher the amount of precipitation is, the lower the suitability values are. **Bio 15** calculated the amount of precipitation per month evaluated through a whole year. The value range of the habitat was: [7,94; 123,21]. In the first version of the modeling, the values for the probabilities were constant for GLM, whereas, for Maxent, they were lower, around 60 to 100 mm. Regarding the land-cover variables, they contributed to the model at ~22% for the whole Europe version and ~40% for the birdlife model. Both versions had the bare landcover vegetation as the major landcover variable contribution (see Appendix 2a; 2b).

3.2 MODEL FIT AND PREDICTIONS

As a reminder, the difference between the whole Europe and the birdlife versions is the studied area and the “artificial” absence points. For the whole Europe version, the mean model (i.e., the model from the mean of GLM and Maxent) returned, in general, slightly lower scores than the two algorithm models, except for the specificity (see Table 11a). The birdlife version is the opposite since the mean model always had higher values except for the sensitivity (see Table 11b). For the birdlife version, the highest AUC was with the mean model (AUC= 0,924), and for the whole Europe version, it was the Maxent model (AUC= 0,917). Based on Elith (2002), all versions and models provided a



Figure 9. 3D map of Europe to see the terrain variation (image source: Sketchfab, Jerry 2021)

good performance since their AUC value was always above 0.75, the minimum value acceptable in reserve planning. To better understand the predictions, the robin's winter habitat was divided into five areas: Western - Southern - Northern - Eastern Europe, and the Mediterranean basin (see Appendix 8 for the detailed map of the areas).

Regarding the predictions, the code returned an occurrence probability map, also called habitat suitability map, based on the modeling (see Figure 10). From yellow to purple ([0.5; 1]) were the highest winter habitat

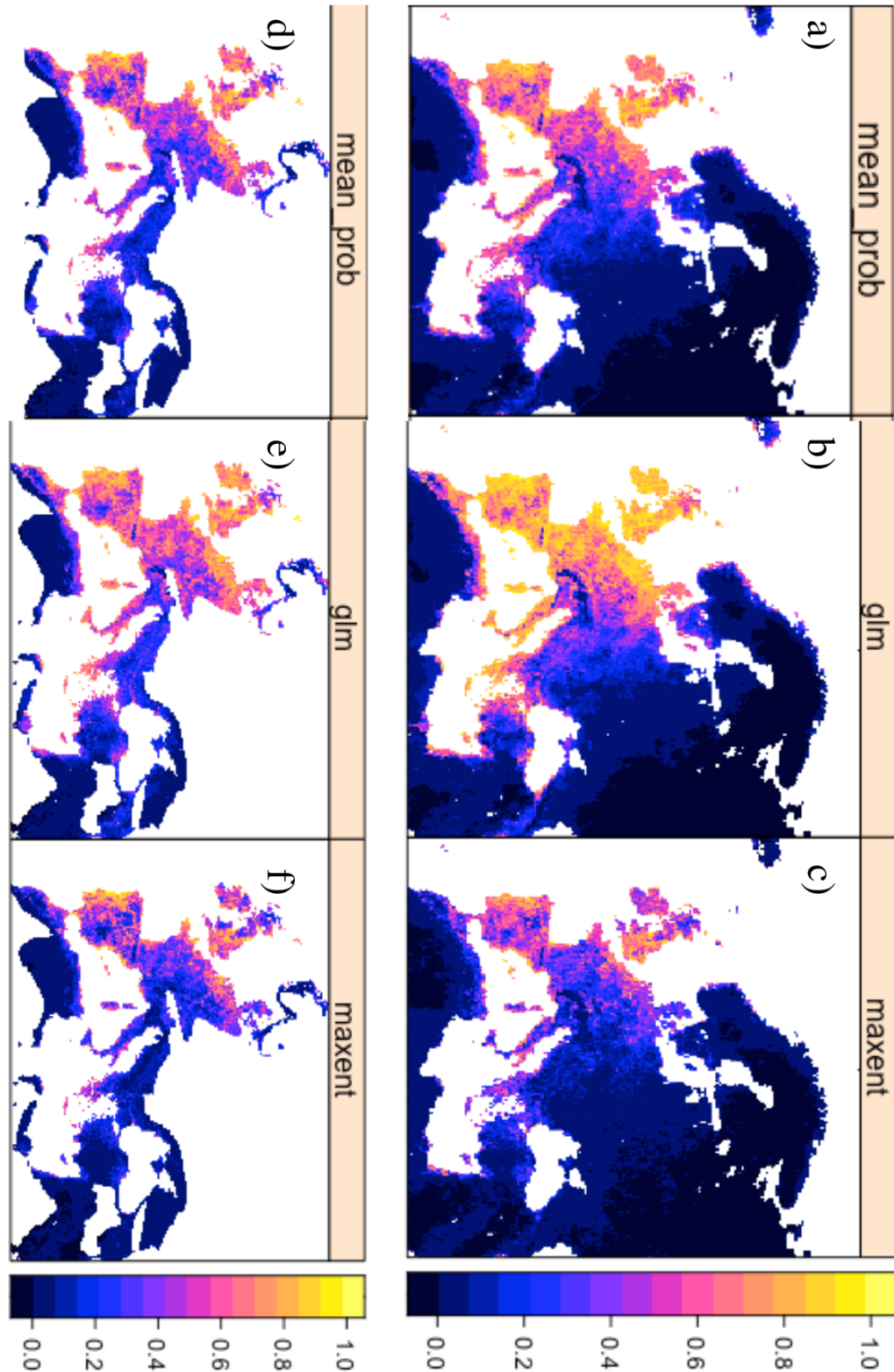


Figure 10. a) Mean of the predictions made by the model for the whole Europe version (GLM + Maxent). b) Predictions of the occurrence probabilities by GLM for the whole Europe version. c) Predictions of the occurrence probabilities made by Maxent for the whole Europe version. 1 is the highest probability of occurrence and 0 is the absence of probability. d) Mean of the predictions made by the model for the birdlife version (GLM + Maxent). e) Predictions of the occurrence probabilities by GLM for the birdlife version. e) Predictions of the occurrence probabilities made by Maxent for the birdlife version.

Results

suitability areas and from purple to black ([0; 0.49]) the less suitable areas.

Since it was recommended that I base my discussion on the birdlife version and that the mean model had the best statistical scores (AUC= 0.924; TSS= 0,698; Kappa = 0,698), those results were used to characterize the winter habitat suitability of the robin. All the results from the suitability map are based on visual evaluation, and an ad hoc threshold value of 0.7 was determined to assess the highest suitability areas and the maxSens+Spec threshold of 0,44 for the lowest suitability areas (see Table 11b). Therefore, habitats with the highest scores for suitability were globally located in the United Kingdom, Portugal, Nederland, Spain, and the coasts of Western/ Southern Europe and the Mediterranean basin (see Figure 10.d). Low-temperature seasonality (BIO 4: [216,5; 710,2]) and high mean temperature of the coldest quarter (BIO 11: [0,5°; 16,5°]) had the highest percentage contribution value. The areas with the lowest probabilities were mainly correlated with a higher relief in those regions, such as the Italian and Dinaric alps and vast plains, as shown in Figure 9. For Northern Europe, the habitat was mainly suitable around the coasts and islands, but the suitability was globally lower than ~0,4. The Mediterranean basin suitability values were low except around some parts of the coasts, and Libya was almost entirely unsuitable. For Eastern Europe, the habitat suitability was relatively minor and below ~0,4.

Algorithm	AUC	TSS	Kappa	Sensitivity	Specificity	Threshold
GLM	0,909	0,695	0,698	0,958	0,738	0,40
Maxent	0,919	0,698	0,701	0,921	0,777	0,20
	AUC	TSS	Kappa	Sensitivity	Specificity	Threshold
Mean	0,917	0,695	0,697	0,914	0,781	0,38
Median	0,917	0,695	0,697	0,914	0,781	0,38
Weighted mean	0,917	0,695	0,697	0,914	0,781	0,38

Table 11.a: Statistical performances of the algorithms and the model for the whole Europe version

Algorithm	AUC	TSS	Kappa	Sensitivity	Specificity	Threshold
GLM	0,837	0,526	0,526	0,855	0,670	0,48
Maxent	0,862	0,548	0,548	0,819	0,729	0,27
	AUC	TSS	Kappa	Sensitivity	Specificity	Threshold
Mean	0,924	0,698	0,698	0,842	0,857	0,440
Median	0,924	0,698	0,698	0,842	0,857	0,440
Weighted mean	0,924	0,698	0,698	0,842	0,857	0,440

Table 11.b: Statistical performances of the algorithms and the model for the birdlife version

4. Discussion

4.1 WINTER ECOLOGICAL NICHE CHARACTERIZATION OF THE EUROPEAN ROBIN

With my research project, I headed out to characterize abiotic factors focally, and less importantly, land-cover factors (biotics) to identify the main drivers of the winter ecological niche of the robin, a partial migrant. I based my discussion on the results from the birdlife version. Since the mean model of that version had the highest AUC value, the visual evaluation of the suitability map is based on that one (see Figure 10d). When evaluating the lowest suitable places (> 0.44) inside the winter robin habitat, the BIO 12 and BIO 11 variable maps (see Appendix 7) seemed to be correlated with those places. Indeed, annual precipitation values above 1500 mm/year and mean temperature of the coldest quarter below $\sim 0^\circ$ found in the highest altitudes might be associated with an unfavorable environment, disabling food foraging for the robin (Biebach, 1983).

Another conclusion based on the suitability map was that since some places are only occupied during the breeding period or the non-breeding period (see Figure 1; blue and orange areas), it could be interesting to investigate the variables that could contribute based on the value range (see Table 8). First, I noticed a place only occupied during the breeding season in the north of Turkey/Armenia, whereas all around the area is occupied outside the breeding period (see Figure 1, orange). According to the variable maps (see Appendix 7), the environmental condition that seems to most impact that place is visually estimated to be a high-temperature seasonality (BIO 4) around 1000, which is outside the winter range tolerated by the robin according to my predictions. Regarding the biotic variables, however, they don't seem to have differences that could explain the absence of robins. So, in this part of the robin distribution range, it seems that abiotic factors could thus have an important role in the migration decision making (Ambrosini, 2016). However, further niche modeling analyses comparing the breeding and wintering ecological niche and variables importance would be needed, such as Ruiz-Sánchez *et al.* method (2015).

In contrast, North Africa, an area intensely occupied outside the breeding period this time, seems to influence the robin presence mainly based on land-cover factors. Indeed, when visually comparing the suitability map to the variables map (see Appendix 7), the bare and herb vegetation land-cover seem to be the major drivers for explaining this high suitability. This might be influenced by the anti-predator strategy, one of the main factors influencing the winter habitat selection of the migrant robins (Cuadrado, 1997). To validate this hypothesis, the mean model should be run at a local scale.

In conclusion, depending on the studied area, different factors might impact the migratory status of the bird. In this case, abiotic variables could play a substantial role in the decision to leave the area, but land-cover could be a significant criterion for deciding on a migratory destination. However, these are just assumptions based on visual cues and the relief might have an impact there too.

4.2 BIOLOGICAL CHARACTERIZATION OF THE NICHE

Based on a study conducted in North America, the temperature likely acts as a proximal factor influencing avian populations through food availability (Emlen *et al.*, 1986). The winter diet of the robin is dominated by invertebrates and fruits, both heavily dependent on seasonality and thus correlated to temperature seasonality (BIO4), putatively impacting the partial migratory behavior of the robin. In fact, fruit production is less essential for plant survival than leaf production; thus, fruits are only produced if weather conditions are favorable during the year. Consequentially, fruit availability relies on more extended periods and is less predictable (Pérez Tris, 2001). Therefore, shifting to invertebrates when fruits are scarce is possible for this species. However, locally low temperature, such as BIO 11, could be a limiting factor when invertebrates are the only type of food available or the only source rich enough in nutrients to allow robins to resist harsh winters. Moreover, as previously described, the robin has a " niche tracking behavior " (Fandos & Tellería, 2020).

Consequently, food availability could be a segregating factor for migration departure decisions correlated with a low-density vegetation area related to anti-predatory behavior and despotic distribution (Pérez Tris, 2001). This means that the most extrinsic favorable conditions for an individual to stay would be a low-temperature seasonal environment with temperatures during the coldest months between 0,5° to 16,53°C and bare vegetation land-cover below 40%. Moreover, in areas where resources are limited during winter, only fitter individuals can persist. Indicators for a higher fitness could be intrinsic factors such as age, sex, and fat score (Catry *et al.*, 2004).

4.3 COMPARING MODEL ACCURACY WITH A PREVIOUS STUDY IN SPAIN

Tellería *et al.* (2012) explored the effectiveness of niche modeling approaches by comparing actual abundance distribution with ENM predictions for the robin in Spain. Since their predictions are positively correlated with the winter habitat suitability of the robin in Spain, I visually examined my own predictions in comparison to the prediction map from Tellería *et al.* Though they did not mention the resolution of their Worldclim variables, they generated a layer of short-wave winter radiation derived from a digital elevation model at a 1 km pixel size. Their resolution was thus higher than mine (although they had a smaller number of occurrences than my study), thus not directly comparable, but still allowed for an overall assessment of the prediction's accuracy. There seem to be a higher habitat suitability area around the coasts and the Balearic Islands for both Tellería *et al.* and my predictions. However, the lowest habitat suitability between my predictions and Tellería did vary. Tellería *et al.* predicted larger areas with low chances in the North of Spain, whereas my model predicted them in the South-West part of Spain. This can be explained by their study involving a small sample size, and thus places where robins were witnessed in my database but missed in their study.

Moreover, the use of different predictors significantly impacts the predictions, and the significant difference was the absence of the human footprint impact in my data (Regos *et al.*, 2019). And finally, they only used Maxent, where I

used the mean of GLM and Maxent. In conclusion, my predictions seemed entirely accurate based on the habitat suitability map from Telleria *et al.* with primarily differences in the specificity (see Appendix 9). However, such a comparison requires additional analysis to be proven statistically meaningful.

4.4 NICHE OCCUPATION

In the context of my master's research, a brief overview of the niche occupation could be interesting in the context of habitat loss and migratory behavior. Indeed, even though it is essential to study the impact of habitat loss in endangered species, studying how a changing environment impacts generalists and 'close to humans' species could also be relevant. If generalist species are impacted, they might considerably impact the community (Davey *et al.*, 2013). Moreover, still unoccupied habitat could reveal crucial information on the facultative partial migratory behavior, but no studies have focally analyzed this yet. Although, it seems like seasonal migration of temperate North American passerine does not influence the range size (Pegan & Winger, 2020). Another reason for unoccupied habitat could also be a regional change of the competitive interactions due to local adaptations of other species to changing habitat/environment (Lenoir *et al.*, 2010)

When juxtaposing the presence map (see Appendix 10) with the winter habitat suitability map (see Figure 10d), on the one hand, we can find robins in places where the suitability habitats, according to the predictions, are relatively low or even outside the range provided by Birdlife (⁷). For instance, some robins were witnessed in Sweden, Poland, or even Lithuania, outside their normal non-breeding area. In this case, the model based on Europe (see Figure 10.a) is more relevant to explain the robin's fundamental ecological niche. One option to explain this phenomenon could be that some individuals sometimes take a new migratory route and end up in typically unsuitable or newly suitable places. For example, this happens with blackcaps showing newly evolving migratory strategies by a novel migratory direction migrating NW from central Europe to southern Britain in autumn, likely due to new suitable places created by artificial bird feeding and climate change (Delmore *et al.*, 2020; Van Doren *et al.*, 2021). A further explanation could be that those individuals decided not to leave and survived due to rare but acceptable environmental conditions during the departure period.

On the other hand, in countries such as France or Ireland presenting decent, suitable habitats (values above the 0.44 threshold), robins are present in fewer numbers than expected by the suitability map. Since my study focused on the fundamental niche, other locally biotic factors not considered in my model could explain this lower number, such as lower access to food due to high competition (Ricklefs, 2010). However, the absence of human observations could also impact the predictions, and consequently, it was impossible to disentangle these factors and more precisely characterize the niche occupation based on my model, and more complex models are needed.

• ⁷ BirdLife International. (2021). Species factsheet: *Erithacus rubecula*. <http://datazone.birdlife.org/species/factsheet/european-robin-erithacus-rubecula>

4.5 WHAT IMPACT DOES THIS HAVE ON MIGRATORY STATUS?

In a study led by Ambrosini *et al.* (2016), the impact of winter temperatures on the migratory behavior of the robin was investigated at a continental scale with bird ringing data. Each individual was classified into a resident or migrant (threshold value of 0.571 km below which the individual was characterized as a resident). Their results showed two main clusters of facultative partial migrants: Belgium/United-Kingdom with 56,3% of migrants; Germany/Czech Republic/Poland with 99,3% of migrants.

Visual comparison based on the suitability and variable maps (see Figure 10d and Appendix 7) is made between the different propensity of migrant/resident areas and the environmental factors selected by my model. The purpose is to have an idea of the variable's values at those places of medium and high migration propensity (i.e., BIO 4; BIO 11; landcover_veg_bare) and evaluate their potential impact on the facultative partial migratory behavior. When around half of the population express a migratory behavior, the environmental conditions associated are a temperature seasonality around 400 and temperature of the coldest quarter between $\sim [0^{\circ}\text{C}; 7^{\circ}\text{C}]$. When compared with the range evaluated in Table 8, those values are relatively close to the medium values the robin can tolerate. However, when almost the whole population is migratory, the temperature seasonality is around $[400; 800]$, and the temperatures of the coldest quarter are around $[-5^{\circ}\text{C}; 5^{\circ}\text{C}]$. In those areas, the seasonality is thus higher, and the temperatures are also colder. Both variables reach values outside the range tolerated by the robin (see Table 8). Higher temperature seasonality and cold periods could have an impact on the proportion of resident individuals surviving on the breeding ground. Since facultative partial migration is also under a genetic influence, the migratory behavior could provide a better fitness (Dhondt, 1983). Regarding the landcover variables, the bare landcover vegetation is similar for both areas. Moreover, the percentage of the herb is lower ($\sim 20\%$) in the second area, whereas the first area has values around 80-100%. Since a higher residential status is associated with low vegetation habitats (i.e., parks and gardens), the propensity of migrants could also be attributed to that factor (Adriaensen & Dhont, Oct 1990).

Based on those results, a preliminary conclusion could be that higher seasonality associated with relatively low temperature during the coldest months, and few herb coverages could be unsuitable for the robins to stay outside the breeding period. Moreover, and as described in section 4.1, climatic factors could play a more substantial role in the departure decision, whereas land-cover could impact assessing the non-breeding area. However, those data are only based on the model of the winter habitat and should request further studies in the field.

4.6 JUSTIFICATION OF THE METHOD AND LIMITS TO THE MODEL

ENM is an approach used to establish the highest suitable places/ conditions for a species to be found and not its actual distribution. The robin is ideally suited to study the influence of environmental factors on partial migratory behavior since it has a wide distribution range and exhibits significant ecological plasticity. Moreover, it is a widespread species, which occurs across a wide distribution range in high numbers and is thus well suited for ENM approaches. This is an essential prerequisite for model accuracy, which I want to assess in this section explicitly.

Before deciding the final settings to study the winter habitat suitability, many trials and thoughts were invested. Two regression-based approaches (GLM and generalized additive models (GAM)) were first chosen for their generalist aspect. GAM was known to be more precise in its responses since it is using smoother algorithms. However, GLM can be sufficient and returned the same results as GAM (Guisan *et al.*, 2017). This was the case with this study, and this is why only GLM was selected. Two other more complex and suitable approaches for this subject were RandomForest, a machine learning method based on decision trees, and Maxent (Guisan *et al.*, 2017). Since Maxent was more suitable for a presence-only database, I decided to use it.

The other major decision to make was regarding the presence-only database. The presence data provided for my master subject were downloaded from the eBird observations loaded on GBIF. Another source could have been the ringing data, as Telleria (2012) did. However, since this study had an exploratory purpose rather than a targeted area analysis (e.g., assessing the winter habitat suitability of the robins in Northern Germany) or analysis on the migratory routes, GBIF was preferred. Though GBIF is a great tool to access many observations easily, its major drawback is the lack of genuine absences. For instance, absences are needed when evaluating the commission rate (false presence) of predictions (Li & Guo, 2013). One way to avoid this problem was to generate those "artificial" absences. However, "artificial" absences are never as precise as genuine absences, and it can induce some bias, such as underestimating the favorable places (Fernandes *et al.*, 2019). Therefore, I selected two types for my master's subject: background data and pseudo-absences. Indeed, since Maxent is a machine-learning algorithm, it requires background points from the whole environment available, whereas GLM needs more precise pseudo-absences.

The background absences are less precise than the pseudo-absences since those points are generated randomly in Europe (i.e., the first version/whole Europe version) or inside the birdlife range (i.e., the second version/ birdlife version) except at the presence points. On the other hand, the pseudo-absences are already more precise since the points were generated to the same extent as the background points but not in a 10 kilometers perimeter around each presence point. This last value was selected based on the winter behavior of the robin (e.g., the territory it occupies and the surroundings where it could search for extra food).

Regarding the number of "artificial" absences, it could also influence the accuracy of predictions. In their paper, Barbet-Massin *et al.* (2012) studied the influence of this number and concluded that both GLM and Maxent were giving better scores with around 10 000 "artificial" absences and presence/absence weighted equally.

For the statistical tests used to assess the accuracy of the ecological niche model, they are based on a confusion matrix composed of four categories: true positives, true negatives, false positives, and false negatives. Derived from that matrix is the commission (i.e., the ability to predict the presence of robins correctly) and omission (i.e., the ability to predict the absence of robins correctly) error rates, respectively known as specificity and sensitivity (Guisan *et al.*, 2017). AUC is commonly used for ENM since it is independent of the threshold, which means that its value does not rely on the threshold selected for the confusion matrix and the sensitivity/specificity (Jiménez-Valverde, 2011). Moreover, it is

suitable for non-binary results. However, it appears that summarizing the ROC into a single AUC measure can provide less information on the predictions than sensitivity/specificity and the tests derived from them (TSS, Kappa). However, disrespect to the AUC theory seems to be done by generating “artificial” absences since its objectives are to produce and evaluate realized distributions that are not feasible without true absences (Jiménez-Valverde (2011)). The use of AUC as evaluation of the model accuracy is thus increasingly debated, and other methods such as Cohen's kappa and the TSS could be more interesting (Fernandes *et al.*, 2019). However, the use of Cohen's kappa is also debated as it relies on the prevalence of the models and has been created to compare models and not assess the quality of a model's predictions (Tooth & Ottenbacher, 2004). Since neither AUC nor Kappa and TSS is completely optimal, I based my model selection on the three.

As previously mentioned, GBIF is a beneficial tool to counter-act shortfalls in terms of data availability (such as the museum collections or scientific surveys). However, it possesses some spatial and human-induced bias. By gathering species information across the globe, it can distort the biodiversity pattern based on the scale, the species, and the data mobilized for the study. Indeed, depending on the country, financial supports and policy can affect data collection and availability. The research in the United Kingdom or Sweden, for instance, is well-financed and can therefore invest more into data collection. The result at a species range can thus be impacted since more individuals will be registered in those countries. However, it can thus induce bias in the predictions, which could partially explain why habitat is more suitable in the Netherland and England. A way to compensate for this bias is using "bias files" that contain information such as population density, sampling intensity of the studied species, or traffic infrastructure (Beck *et al.*, 2014).

Nonetheless, those biases do not explain all the occupations related to humans. For example, as described in section 1.2.2, the winter habitat of the robin is also driven by human activities such as artificial feeding. It is thus not a surprise to find better habitat suitability around big cities, especially in the UK, where I have already described this shift in the blackcap winter habitat (Berthold *et al.*, 1992; Plummer *et al.*, 2015; Van Doren *et al.*, 2021).

Therefore, my predictions appear to be statistically good, but caution must be taken regarding accuracy since the human footprint is yet not included and "artificial" absence data were used. Moreover, the literature does not advice to generate “artificial” absences only on the winter birdlife range as the “artificial” absences must represent the complete environmental range of the robin. But since the robin has a niche tracking behavior, this should not have a major impact on the predictions and generating “artificial” absences at a to large extent is not recommended either (Barbet-Massin *et al.*, 2012; Fandos & Tellería, 2020). However, the significant amount of presence data gathered could balance those defaults since I had more than 80 000 presence points and still had ~10 000 after the cleaning process. In the future, the primary setting to change would be the addition of the human footprint and more adequate statistical tests such as the Boyce index, which is more suitable for presence-only data.

5. Conclusion and perspectives

In conclusion, the purpose of this master thesis was to assess the winter habitat suitability of the European robin and its potential impact on its migratory behavior through an ecological niche modeling approach and the use of climatic (i.e., abiotic) and land-cover (i.e., biotic) variables. Despite many ecological studies led on the European robin at a local scale, with this continent-wide niche modeling analysis, we now have a clearer vision of the contribution of climatic and land-cover factors to the winter habitat of the European robin. Temperature seasonality and temperature of the coldest quarter were the two main climatic variables, whereas, for the land-cover variables, the bare land-cover had the most significant contribution. The suitability map also provided preliminary results on biotic and abiotic variables' influence on migratory decision making.

From broader scales perspectives, my results could help understand and predict the migratory behavior of the robin and select populations for genetic sampling and sequencing to understand the genetic basis of the highly variable migratory behavior in European robins. My model could also provide a general basis for future analyses on species distribution based on their niche. For instance, in modeling past and future climate projections, e.g., to accomplish genetic or demographic studies based on single nucleotide polymorphism (SNP) data and understand how different robin populations can react to changing climatic conditions in the face of climate change.

REFERENCES

Papers and books

- Adriaensen, F. & Dhondt, A.A (1990). Territoriality in the continental European robin, *Erithacus rubecula rubecula*. *Ardea*, 459-465.
- Adriaensen, F., & Dhondt, A. A. (1990). Population dynamics and partial migration of the european robin (*Erithacus rubecula*) in different habitats. *The Journal of Animal Ecology*, 59(3), 1077.
<https://doi.org/10.2307/5033>
- 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>
- Ambrosini, R., Cuervo, J. J., Feu, C. du, Fiedler, W., Musitelli, F., Rubolini, D., Sicurella, B., Spina, F., Saino, N., & Møller, A. P. (2016). Migratory connectivity and effects of winter temperatures on migratory behaviour of the European robin *Erithacus rubecula*: A continent-wide analysis. *Journal of Animal Ecology*, 85(3), 749-760. <https://doi.org/10.1111/1365-2656.12497>
- Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22(1), 42-47. <https://doi.org/10.1016/j.tree.2006.09.010>
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many?: How to use pseudo-absences in niche modeling?. *Methods in Ecology and Evolution*, 3(2), 327-338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, 19, 10-15.
<https://doi.org/10.1016/j.ecoinf.2013.11.002>
- Berthold, P., & Querner, U. (1981). Genetic basis of migratory behavior in european warblers. *Science*, 212(4490), 77-79. <https://doi.org/10.1126/science.212.4490.77>

References

- Berthold, P., Helbig, A. J., Mohr, G., & Querner, U. (1992). Rapid microevolution of migratory behaviour in a wild bird species. *Nature*, 360(6405), 668-670. <https://doi.org/10.1038/360668a0>
- Biebach, H. (1983). Genetic determination of partial migration in the european robin (*Erithacus rubecula*). *The Auk*, 100(3), 601-606. <https://doi.org/10.1093/auk/100.3.601>
- Campos, A. R., Catry, P., Ramos, J., & Robalo, J. I. (2011). Competition among European Robins *Erithacus rubecula* in the winter quarters: Sex is the best predictor of priority of access to experimental food resources. *Ornis Fennica*, 88, 226-233.
- Catry, P., Campos, A. R., Granadeiro, J. P., Neto, J. M., Ramos, J., Newton, J., & Bearhop, S. (2016). Provenance does matter: Links between winter trophic segregation and the migratory origins of European robins. *Oecologia*, 182(4), 985-994. <https://doi.org/10.1007/s00442-016-3725-z>
- Catry, P., Campos, A., Almada, V., & Cresswell, W. (2004). Winter segregation of migrant European robins *Erithacus rubecula* in relation to sex, age and size. *Journal of Avian Biology*, 35(3), 204-209. <https://doi.org/https://doi.org/10.1111/j.0908-8857.2004.03266.x>
- Chapman, B. B., Brönmark, C., Nilsson, J.-Å., & Hansson, L.-A. (2011). The ecology and evolution of partial migration. *Oikos*, 120(12), 1764-1775. <https://doi.org/https://doi.org/10.1111/j.1600-0706.2011.20131.x>
- Collar, N. (2020). *European Robin (Erithacus rubecula)*, version 1.0. In *Birds of the World* (J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.eurrob1.01>
- Colwell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences*, 106(Supplement_2), 19651–19658. <https://doi.org/10.1073/pnas.0901650106>
- Cuadrado, M. (1997). Why are migrant Robins (*Erithacus rubecula*) territorial in winter? : The importance of the anti-predatory behaviour. *Ethology Ecology & Evolution*, 9(1), 77-88. <https://doi.org/10.1080/08927014.1997.9522904>
- Davey, C. M., Devictor, V., Jonzén, N., Lindström, Å., & Smith, H. G. (2013). Impact of climate change on communities : Revealing species' contribution. *Journal of Animal Ecology*, 82(3), 551-561. <https://doi.org/10.1111/1365-2656.12035>

References

- Debussche, M., & Isenmann, P. (1985). Frugivory of transient and wintering European robins *Erithacus rubecula* in a Mediterranean region and its relationship with ornithochory. *Ecography*, 8(2), 157-163. <https://doi.org/10.1111/j.1600-0587.1985.tb01166.x>
- Delmore, K. E., Van Doren, B. M., Conway, G. J., Curk, T., Garrido-Garduño, T., Germain, R. R., Hasselmann, T., Hiemer, D., van der Jeugd, H. P., Justen, H., Lugo Ramos, J. S., Maggini, I., Meyer, B. S., Phillips, R. J., Remisiewicz, M., Roberts, G. C. M., Sheldon, B. C., Vogl, W., & Liedvogel, M. (2020). Individual variability and versatility in an eco-evolutionary model of avian migration. *Proceedings of the Royal Society B: Biological Sciences*, 287(1938), 20201339. <https://doi.org/10.1098/rspb.2020.1339>
- DiMiceli, C.M., Carroll, M.L., Sohlberg, R.A., Huang, C., Hansen, M.C. and Townshend, J.R.G. (2011). *Annual global automated MODIS vegetation continuous fields (MOD44B) at 1x1 km spatial resolution for data years beginning day 65, 2000–2010*. University of Maryland, College Park, MD, USA.
- Dhondt, A. A. (1983). Variations in the number of overwintering stonechats possibly caused by natural selection. *Ringing & Migration*, 4(3), 155-158. <https://doi.org/10.1080/03078698.1983.9673800>
- 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), 27-46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.
- Elith, J. (2002). Quantitative methods for modeling species habitat: Comparative performance and an application to australian plants. In S. Ferson & M. Burgman (Éds.), *Quantitative Methods for Conservation Biology* (p. 39-58). Springer. https://doi.org/10.1007/0-387-22648-6_4
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. *Diversity and Distributions*, 17(1), 43-57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>.
- Emlen, J. T., DeJong, M. J., Jaeger, M. J., Moermond, T. C., Rusterholz, K. A., & White, R. P. (1986). *Density trends and range boundary constraints of forest birds along a latitudinal gradient*. *The Auk*, 103(4), 791-803. <https://doi.org/10.1093/auk/103.4.791>

References

- Fandos, G., & Tellería, J. L. (2018). Range compression of migratory passerines in wintering grounds of the Western Mediterranean: Conservation prospects. *Bird Conservation International*, 28(3), 462-474. <https://doi.org/10.1017/S0959270917000120>.
- Fandos, G., & Tellería, J. L. (2020). Seasonal niche-tracking behaviour of two partially migratory passerines. *Ibis*, 162(2), 307-317. <https://doi.org/10.1111/ibi.12721>
- Fernandes, R. F., Scherrer, D., & Guisan, A. (2019). Effects of simulated observation errors on the performance of species distribution models. *Diversity and Distributions*, 25(3), 400-413. <https://doi.org/10.1111/ddi.12868>
- Fourcade, Y., Besnard, A. G., & Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Biogeography*, 27(2), 245-256. <https://doi.org/10.1111/geb.12684>
- Freeman, E. A., & Moisen, G. G. (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, 217(1-2), 48-58. <https://doi.org/10.1016/j.ecolmodel.2008.05.015>.
- Géroudet, P. (1951). Le rouge gorge familier *Erithacus rubecula* (L.) Les passereaux d'Europe, des coucous aux merles (p 308-315) ». Delachaux et Niestlé, tome 1.
- Glutz von Blotzheim, U. N., Bauer, K. M., Bezzel, E. (1985). *Handbuch der Vögel Mitteleuropas*, vols 1-14. Aula, Wiesbaden.
- Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat suitability and distribution models: with Applications in R*. Cambridge University Press.
- Günther, A., Einwich, A., Sjulstok, E., Feederle, R., Bolte, P., Koch, K.-W., Solov'yov, I. A., & Mouritsen, H. (2018). Double-Cone Localization and Seasonal Expression Pattern Suggest a Role in Magnetoreception for European Robin Cryptochrome 4. *Current Biology*, 28(2), 211-223., <https://doi.org/10.1016/j.cub.2017.12.003>
- Hegemann, A., Fudickar, A. M., & Nilsson, J.-Å. (2019). A physiological perspective on the ecology and evolution of partial migration. *Journal of Ornithology*, 160(3), 893-905. <https://doi.org/10.1007/s10336-019-01648-9>

References

- Hinde, A. (1956). The biological significance of the territories of birds. *Ibis*, 98(3), 340-369.
<https://doi.org/https://doi.org/10.1111/j.1474-919X.1956.tb01419.x>
- Hutchinson, G.E. (1957). Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415-427. <http://dx.doi.org/10.1101/SQB.1957.022.01.039>
- Jiménez-Valverde, A., (2011). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, 21, 498–507.
- Jiménez, L., Soberón, J., Christen, J. A., & Soto, D. (2019). On the problem of modeling a fundamental niche from occurrence data. *Ecological Modelling*, 397, 74-83.
<https://doi.org/10.1016/j.ecolmodel.2019.01.020>
- Lack, D. (1946). *The life of the robin*, London :H. F. & G. Witherby LTD.
- Lack, D. (1951). Geographical variation in *Erithacus rubecula*. *Ibis*, 93(4), 629-630.
<https://doi.org/https://doi.org/10.1111/j.1474-919X.1951.tb05469.x>
- Lenoir, J., Gégout, J.-C., Guisan, A., Vittoz, P., Wohlgemuth, T., Zimmermann, N. E., Dullinger, S., Pauli, H., Willner, W., & Svenning, J.-C. (2010). Going against the flow : Potential mechanisms for unexpected downslope range shifts in a warming climate. *Ecography*, 33(2), 295-303.
<https://doi.org/10.1111/j.1600-0587.2010.06279.x>
- Li, W., & Guo, Q. (2013). How to assess the prediction accuracy of species presence-absence models without absence data?. *Ecography*, 36(7), 788-799. <https://doi.org/10.1111/j.1600-0587.2013.07585.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>
- Morales, N. S., Fernández, I. C., & Baca-González, V. (2017). MaxEnt's parameter configuration and small samples: Are we paying attention to recommendations? A systematic review. *PeerJ*, 5, e3093.
<https://doi.org/10.7717/peerj.3093>
- O'Donnell, M.S., and Ignizio, D.A., (2012), *Bioclimatic predictors for supporting ecological applications in the conterminous United States: U.S. Geological Survey Data Series 691*, 10 p.

References

- Pegan, T. M., & Winger, B. M. (2020). The influence of seasonal migration on range size in temperate North American passerines. *Ecography*, 43(8), 1191-1202. <https://doi.org/10.1111/ecog.05070>
- Pérez Tris, J. (2001). Migración y Sedentarismo en los Paseriformes Forestales Ibéricos: Perspectivas ecológicas y evolutivas, [PhD thesis, Universidad Complutense de Madrid], https://www.ucm.es/data/cont/media/www/pag-33472/Tesis_2001_pereztris.pdf#page=103
- Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications*, 19(1), 181-197. <https://doi.org/10.1890/07-2153.1>
- Plummer, K. E., Siriwardena, G. M., Conway, G. J., Risely, K., & Toms, M. P. (2015). Is supplementary feeding in gardens a driver of evolutionary change in a migratory bird species?. *Global Change Biology*, 21(12), 4353-4363. <https://doi.org/10.1111/gcb.13070>
- Pulido, F. (2011). Evolutionary genetics of partial migration - the threshold model of migration revis(it)ed. *Oikos*, 120(12), 1776–1783. <https://doi.org/10.1111/j.1600-0706.2011.19844.x>
- Pulido, F., Berthold, P., & van Noordwijk, A. J. (1996). *Frequency of migrants and migratory activity are genetically correlated in a bird population: Evolutionary implications*. Proceedings of the National Academy of Sciences, 93(25), 14642-14647. <https://doi.org/10.1073/pnas.93.25.14642>
- Regos, A., Gagne, L., Alcaraz-Segura, D., Honrado, J. P., & Domínguez, J. (2019). Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Scientific Reports*, 9(1), 4221. <https://doi.org/10.1038/s41598-019-40766-5>
- Remisiewicz, M. (2002). The spatio-temporal pattern to Robin (*Erithacus rubecula*) migration—evidence from ringing recoveries. *Ardea*, 90. 489-502.
- Remsen, J. V. (2010). Chapter 6: Subspecies as a meaningful taxonomic rank in avian classification - las subespecies como un rango taxonómico significativo en la clasificación de las aves. *Ornithological Monographs*, 67(1), 62-78. <https://doi.org/10.1525/om.2010.67.1.62>
- Ricklefs, R. E. (2010). Evolutionary diversification, coevolution between populations and their antagonists, and the filling of niche space. *Proceedings of the National Academy of Sciences*, 107(4), 1265-1272.

References

- Ruiz-Sánchez, A., Renton, K., Landgrave-Ramírez, R., Mora-Aguilar, E. F., & Rojas-Soto, O. (2015). Ecological niche variation in the Wilson's warbler *Cardellina pusilla* complex. *Journal of Avian Biology*, 46(5), 516-527. <https://doi.org/10.1111/jav.00531>
- Schwabl H., Silverin B. (1990) Control of Partial Migration and Autumnal Behavior. In: Gwinner E. (eds) Bird Migration. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-74542-3_10
- Smith, A. B., Godsoe, W., Rodríguez-Sánchez, F., Wang, H.-H., & Warren, D. (2019). Niche estimation above and below the species level. *Trends in Ecology & Evolution*, 34(3), 260-273. <https://doi.org/10.1016/j.tree.2018.10.012>
- Tellería, J. L., Santos, T., Refoyo, P., & Muñoz, J. (2012). Use of ring recoveries to predict habitat suitability in small passerines: Ring recoveries used to predict bird distribution. *Diversity and Distributions*, 18(11), 1130-1138. <https://doi.org/10.1111/j.1472-4642.2012.00900.x>
- Tooth, L.R. & Ottenbacher, K.J. (2004) The k statistic in rehabilitation research: an examination. *Archives of Physical Medicine and Rehabilitation*, 85, 1371–1376.
- Van Doren, B. M., Conway, G. J., Phillips, R. J., Evans, G. C., Roberts, G. C. M., Liedvogel, M., & Sheldon, B. C. (2021). Human activity shapes the wintering ecology of a migratory bird. *Global Change Biology*, 27(12), 2715-2727. <https://doi.org/10.1111/gcb.15597>
- Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svantesson, S., Wengström, N., Zizka, V., & Antonelli, A. (2019). Coordinate-Cleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, 7. doi: 10.1111/2041-210X.13152, Rpackage version 2.0-18
- Zurell, D., Zimmermann, N. E., Gross, H., Baltensweiler, A., Sattler, T., & Wüest, R. O. (2020). Testing species assemblage predictions from stacked and joint species distribution models. *Journal of Biogeography*, 47(1), 101-113. <https://doi.org/10.1111/jbi.13608>

References

Websites

- BirdLife International. (2021). Species factsheet: *Erithacus rubecula*. <http://datazone.birdlife.org/species/factsheet/european-robin-erithacus-rubecula>
- BirdLife International. (n.d.). Species distribution data request. <http://datazone.birdlife.org/species/requestdis>
- BirdLife species factsheet. (n. d.). European Robin (*Erithacus rubecula*). <http://datazone.birdlife.org/species/factsheet/european-robin-erithacus-rubecula>
- Fick, S.E. and R.J. Hijmans, (2017). WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37 (12): 4302-4315.
- GBIF.org (2020, October 14) GBIF Occurrence Download <https://doi.org/10.15468/dl.amqum6>
- GBIF. (n.d.). *What is Darwin Core, and why does it matter?*. <https://www.gbif.org/darwin-core>
- GBIF. (n.d.). *What is GBIF?*. <https://www.gbif.org/what-is-gbif>
- GISGeography. (2021, June 8). *World Geodetic System (WGS84)*. <https://gisgeography.com/wgs84-world-geodetic-system/>
- Jerry, F. (2015, January 14). *Europe Relief Map*. Sketchfab. <https://sketchfab.com/3d-models/europe-relief-map-9f8e318d93b246fdb2cdab2ce4f07c03>
- National Center for Ecological Analysis and Synthesis. (n.d.). *Overview of Coordinate Reference Systems (CRS) in R*. <https://www.nceas.ucsb.edu/sites/default/files/2020-04/OverviewCoordinateReferenceSystems.pdf>
- NOAA (National Oceanic and Atmospheric Administration) Climate. (n.d.). *Climate Models*. <https://www.climate.gov/maps-data/primer/climate-models>
- Riahi, K., van Vuuren, DP., Kriegler, E., O'Neill, B. (2016). *The Shared Socio-Economic Pathways (SSPs): An Overview*, [Poster], https://unfccc.int/sites/default/files/part1_iiasa_rogelj_ssp_poster.pdf
- WordClim. *Future climate data*. (n. d.). <https://worldclim.org/data/cmip6/cmip6climate.html>
- WordClim. *Global climate and weather data*. (n. d.). <https://worldclim.org/data/cmip6/cmip6climate.html>

References

Package

- Bioclim2: Bioclim - bioclimatic variables - simplified in jjvanderwal/climates: methods for working with weather & climate. (s. d.). <https://rdr.io/github/jjvanderwal/climates/man/bioclim2.html>
- Hijmans, R. J., Phillips, S., Elith, J., Leathwick, J. (2020) dismo: Species Distribution Modeling. R package version 1.3-3. <https://CRAN.R-project.org/package=dismo>
- Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J., Uriarte, M. and R.P. Anderson (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods in Ecology and Evolution*, 5(11), 1198-1205.
- Robert J. Hijmans (2021). raster: Geographic Data Analysis and Modeling. R package version 3.4-13. <https://CRAN.R-project.org/package=raster>
- RStudio Team (2019). RStudio: Integrate Development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.
- Zurell, D. (2020). mecofun: useful functions for macroecology and species distribution modelling version 0.0.0.9. University of Potsdam, Potsdam. <https://gitup.uni-potsdam.de/macroecology/mecofun>
- Zizka A, Silvestro D, Andermann T, Azevedo J, Duarte Ritter C, Edler D, Farooq H, Herdean A, Ariza M, Scharn R, Svanteson S, Wengstrom N, Zizka V, Antonelli A (2019). CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, -7. doi: 10.1111/2041-210X.13152 (URL: <https://doi.org/10.1111/2041-210X.13152>), R