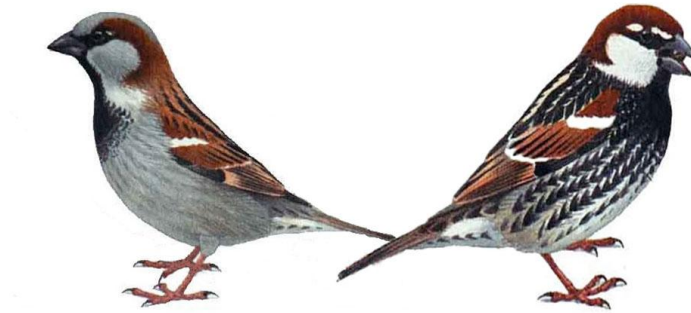# Fine- and broad-scale premating isolation between two closely related species, the house sparrow (*Passer domesticus*) and the Spanish sparrow (*P. hispaniolensis*)

My Hanh Tu

Master of Science Thesis

2012

Centre for Ecological and Evolutionary Synthesis

Department of Biology

Faculty of Mathematics and Natural Sciences

University of Oslo, Norway

Table of Contents

Acknowledgements

This thesis was written at the Center for Ecological and Evolutionary Synthesis (CEES) at the Department of Biology, University of Oslo, under the supervision of Glenn-Peter Sætre, Fredrik Haas and Jo Skeie Hermansen.

First of all, I want to thank Glenn for accepting me into his group and giving me a project I absolutely love. You're an amazing supervisor and you've always been there when I've needed you. In these two years you've taught me so much and your support and guidance have kept me going. You're such a great teacher and students are lucky to have you as a supervisor and professor.

Fredrik, you're probably the most positive and enthusiastic person I've ever known. Your love for birds has really rubbed off on me and you've shaped me to become a "birdlover" (I didn't even know it was possible!). Fieldwork with you has been a blast and I've learned so much. Thank you for opening your home to me and giving me all the help and guidance even though we're not even in the same country! You're a great teacher and inspiration to me.

To the Sparrow group: I couldn't be prouder for being a part of this group. Everyone have been more than great and supported me in every way. I can't find words to express how grateful I am to all of you. A special thanks goes out to Fabrice for helping me with the statistics and giving me comments on my thesis. Richard, it's like you can see when I'm struggling and you just swoop in and save the day. Jo, for making my thesis better. Jorunn, for keeping up with my constant whining (and crying) in the study room. Finally, Cassie! You're such a great friend, always encouraging and helping me whenever I've needed it. You calm me down when I'm stressed (You give the best pep talks!) and you're just always there for me. I really don't know what I would do without you! (And Maria González, I miss you and I'm looking forward to see you again soon!)

I also want to say thanks to my friends (Helena, Kjersti and Anette) and studyroom-mates for supporting me all this time.

My family (dad, mom, My Anh, Chi, and Jørgen), thank you so much for the support and encouragement (and really good food) throughout the years. I love you, and I hope I've made you proud. Lastly, Sverre for giving me something to look forward to everyday and comforting me whenever I've needed it. I couldn't have asked for anyone better - You're the best!

Abstract

Speciation involves the evolution of reproductive barriers that maintain the genetic integrity of species. Premating barriers, in particular, have been referred to as one of the most important mechanisms in keeping species segregated. In this study, I investigate such premating isolation between two closely related species, the house sparrow (*Passer domesticus*) and Spanish sparrow (*P. hispaniolensis*). These two species live in sympatry over most of the Spanish sparrow's breeding range, but apparently, interbreeding between them is either infrequent or absent. During seven weeks of fieldwork over two breeding seasons, I studied the two species in an area where they are sympatric, but remain distinct, namely Badajoz, Spain. To examine the potential premating barriers that keep them isolated, a series of methods were implemented. Transect counts and radio-telemetry was conducted to investigate habitat use, and these methods mutually demonstrated significant habitat segregation between the species. House sparrows stay in closer proximity to human constructions, while maintaining a larger distance to water-sources, whereas Spanish sparrows do the opposite staying closer to water and further away from buildings and human activity. Further, observations of breeding colonies at a larger spatial scale revealed a between-species difference in choice of nest site with pure house sparrow colonies residing closer to cities than colonies including Spanish sparrows. By investigating phenology, I found a minimum difference of 13 days in timing of breeding, consistent with studies in aviaries, which would reduce gene flow significantly. Hence, the opportunities for hybridization are strongly reduced due to spatial and temporal isolation. Finally, ecological niche models indicate dry climate as the main driving force for the distribution of Spanish sparrow which can point to fundamental ecological differences between the species. Overall, my results suggest multiple ecological differences between the two species that most likely contribute towards isolating them, explaining why they do not interbreed extensively in areas of sympatry.

Introduction

Ernst Mayr (1942) defined species as "groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups". When two such groups meet in sympatry, and apparently do not interbreed (at least not at high rates), barriers have to be present between them in order for them to retain their genetic integrity. Identifying reproductive barriers that prevents gene flow between species is therefore vital for the understanding of speciation.

Reproductive barriers are mechanisms that block genes from flowing between taxa. Such isolation can be classified into three major categories depending on when in the reproductive cycle they act: premating, postmating prezygotic and postzygotic barriers. Of the three categories, premating barriers have often been referred to as the more important, critical and effective barriers against gene flow (e.g. Coyne and Orr 2004). They are mechanisms that can prevent or reduce the gene flow before postzygotic barriers even come into play. This gives the premating barriers an important role as the first defense to prevent different species from initiating courtship or copulation. Such barriers have been reported between many taxa, such as plants (Kay 2006; Martin and Willis 2007), amphibians (Littlejohn 1965), insects (Bailey et al. 2004; Morrow et al. 2000; Koeniger and Koeniger 2000) and birds (Balakrishnan and Sorenson 2006; Price 2007). Premating barriers can be further divided into three sub-categories, namely habitat isolation, temporal isolation and behavioral isolation (Coyne and Orr 2004).

Habitat isolation, is the set of mechanisms that arise when species exhibit preferences or adaptations for separate habitats, and consequently reduces the likelihood of reproductive encounters between the species (Funk et al. 2002). It may however not always rest solely on the organism's own preference, but also on competition that forces them to use different niches in areas of sympatry and on their ability to find the habitats to which they are best adapted (Coyne and Orr 2004). Basically, there are two forms of habitat isolation: Microspatial isolation occurs when members of the two species occupy the same general area, but their reproductive encounters are reduced by their different preference or adaptations to different parts of the area, such as different host plants. Macrospatial isolation, however, is when members of the two species cannot interbreed because their habitats are allopatric (Coyne and Orr 2004). The main difference between the two types is that in microspatial isolation the species regularly encounter each other, whilst for species with macrospatial

isolation; the contact between the species is highly reduced. Finally, a mix of both forms of habitat isolation can restrict gene flow between species, as for the crickets *Gryllus pennsylvanicus* and *G. firmus* (Rand and Harrison 1989). These two closely related species are adapted to different habitats, the former preferring loamy soils and the latter sandy soils. Their different habitats are largely allopatric, but their ranges overlap in areas that hold patches of both habitats. Interestingly, in such areas, they are largely restricted to their respective preferred soil type resulting in a mosaic of populations. Their preference or adaptation for different soil type therefore acts as a barrier preventing gene flow between them (Rand and Harrison 1989). It is important to keep in mind that habitat isolation, like all forms of prezygotic isolation, can only work as a strong barrier during the period when pair formation occurs. For instance, many bird species overlap in their ranges in the winter and often appear in mixed flocks, but as soon as spring arrives and they are ready to breed, they split up and change to having different breeding ranges (Coyne and Orr 2004; Funk et al. 2002).

Another isolating mechanism, temporal isolation, reduces gene flow because the species differ in mating period or season. Such isolation can either depend on different responses to environmental cues, for instance tidal or lunar cycles, which is often the case for marine species, or alternatively, it can depend on responses to different environmental cues, such as availability of food resources (Coyne and Orr 2004). The latter represents the case for the closely related cactus finch (*Geospiza scandens)* and the medium ground finch (*G. fortis)* on the Galápagos Islands of Daphne and Daphne Major. The breeding of the cactus finch is closely linked to the flowering of *Opuntia echios* before the annual rains, while the medium ground finch breeds only after the annual rains begin (Grant and Grant 1996; Millington and Grant 1984). Temporal isolation, however, has been suggested to be relatively unimportant (Mayr 1963 in Coyne and Orr 2004). Mayr argued that in animals, the contribution of such barriers to the preservation of reproductive isolation between species is largely unknown. Furthermore, he pointed out that in cases where temporal isolation was broken down due to unusual weather conditions, reproductive isolation was still preserved by other factors. The role and potential significance of temporal isolation in reducing gene flow therefore clearly needs further investigation.

Finally, behavioral isolation represents all species differences that lead to lack of attraction between opposite sexes of different species, and therefore prevents initiation of courtship or copulation between them. A typical scenario would be one sex (usually the male)

having a signal or trait (song, acoustic signals, plumage) that stimulates a preference in the other conspecific sex (usually the female) but does not stimulate individuals of other species (Coyne and Orr 2004). One of the best-known studies of behavioral isolation is on the wing melanin pattern in the western white butterfly (*Pieris occidentalis)* and the checkered white butterfly (*P. protodice)* (Wiernasz and Kingsolver 1992). They are closely related, live sympatrically in the western United States, and are morphologically similar. However, they do not hybridize in nature, even though their hybrids have proven to be more or less viable. The males of the two species differ in wing patterns; with those of *P. occidentalis* having much darker forewings. Field observations reveal that *P. occidentalis* females discriminate against courting heterospecific males relative to conspecific males. However, when experimentally darkening the forewings of *P. protodice* males, they were significantly more acceptable to the *P. occidentalis* females (Wiernasz and Kingsolver 1992).

Much of our knowledge regarding the evolution of reproductive isolation is derived from the rigorously studied *Drosophila* fruit flies. This has led to the revelation of two important patterns. Firstly, reproductive isolation increases with phylogenetic distance between the hybridizing species (Tubaro and Lijtmaer 2002). Studies have shown that in many taxa, the degree of hybridization may increase with increasing phenotypic similarities between the species in traits such as coloration, size, shape, song or courtship display (Grant and Grant 1992; Nagel and Schluter 1998; Seehausen and Alphen 1998). These studies suggest that species may have a hard time discriminating their own species from other species due to similarities in such traits. However, hybridization might not always result from failure in species recognition. It may also result from a scarcity of members of the other sex forcing heterospecific mating to happen (the so called Hubb's principle; cf. Hubbs 1955). Hybridization can then be seen as a "best-of-a-bad-job" strategy (Burke and Arnold 2001; Hettyey and Pearlman 2003; Nuechterlein and Buitron 1998). Another hypothesis, suggested by Pierotti and Annett (1993), proposes that the rate of hybridization is influenced by parental care. Females may choose mates of another species that provides more or better paternal care than mates of their own species. An interesting example can be taken from Veen et al.'s (2001) study of the collared (*Ficedula albicollis)* and pied flycatcher (*F. hypoleuca*). In this study, female collared flycatchers pair with male pied flycatchers despite the fact there is a plentiful supply of male collared flycatchers. This happens especially in late breeding season and the heterospecific pairing has shown to yield 40 % more young than homospecific pairs. When hybrids formed between two species are of low fertility and fecundity, costs are to be

paid which leads to selection favoring the evolution of premating mechanisms that reduces hybridization (Saetre et al. 1997). Interestingly, female collared flycatchers reduce this cost of hybridization by engaging in extra-pair copulations with nearby conspecifics. In fact, an astonishing 60 % of all chicks of female collared flycatchers paired with male pied flycatchers are collared flycatchers and not hybrids (Veen et al. 2001).

Secondly, reproductive isolation is greater between sympatric than allopatric species with the same phylogenetic distance (Tubaro and Lijtmaer 2002). Two populations can be defined as being sympatric "…if individuals of each are physically capable of encountering one another (…) with moderately high frequency" (Bush and Howard 1986 in Rice and Salt 1990). As many closely related species live sympatrically, sharing the same geographic areas, gene flow between the species is highly possible and likely.

In this project, I investigate potential premating barriers on both a fine and broad geographic scale between two closely related species, the house sparrow (*Passer domesticus)* and the Spanish sparrow *(P. hispaniolensis)*. The house sparrow is a well-known, widely distributed bird with a close relationship to human settlements. Originating in the Middle East, it has followed the spread of agriculture some 3,000-7,000 years ago (Saetre et al. 2012). The Spanish sparrow, on the other hand, seems to be less associated with human habitation. The two species have a large sympatric range, with the distribution of the Spanish sparrow contained almost entirely within the distribution of the house sparrow (Figure 1). Interbreeding is usually very limited between the two taxa, but episodes of extensive hybridization have been documented in some areas, especially where the abundance of one of the species is low (Summer-Smith 1988). The latter is likely to have been the case in the past in Italy where hybridization between them has led to the formation of a homoploid hybrid species (that is, a hybrid species with the same number of chromosomes as the parent species). This viable hybrid taxon, the Italian Sparrow (*P. italiae*) (Elgvin et al. 2011; Hermansen et al. 2011), resides on the entire Italian Peninsula and some Mediterranean Islands. It occurs in sympatry with the Spanish sparrow on the Gargano Peninsula in southeast Italy and meets the house sparrow in a narrow contact zone in the Alps. Interbreeding with its parent species is infrequent (with the house sparrow in the Alps) or apparently absent (with the Spanish sparrow on Gargano) (Hermansen et al. 2011). Still, although hybridization evidently has led to the formation of a viable hybrid taxon, the parent species are able to live in sympatry without much hybridization and introgression occurring. Thus, the study of potential premating barriers between the house sparrow and the Spanish sparrow in an area of

sympatry is very much of interest and essential for the understanding of speciation in this species complex.
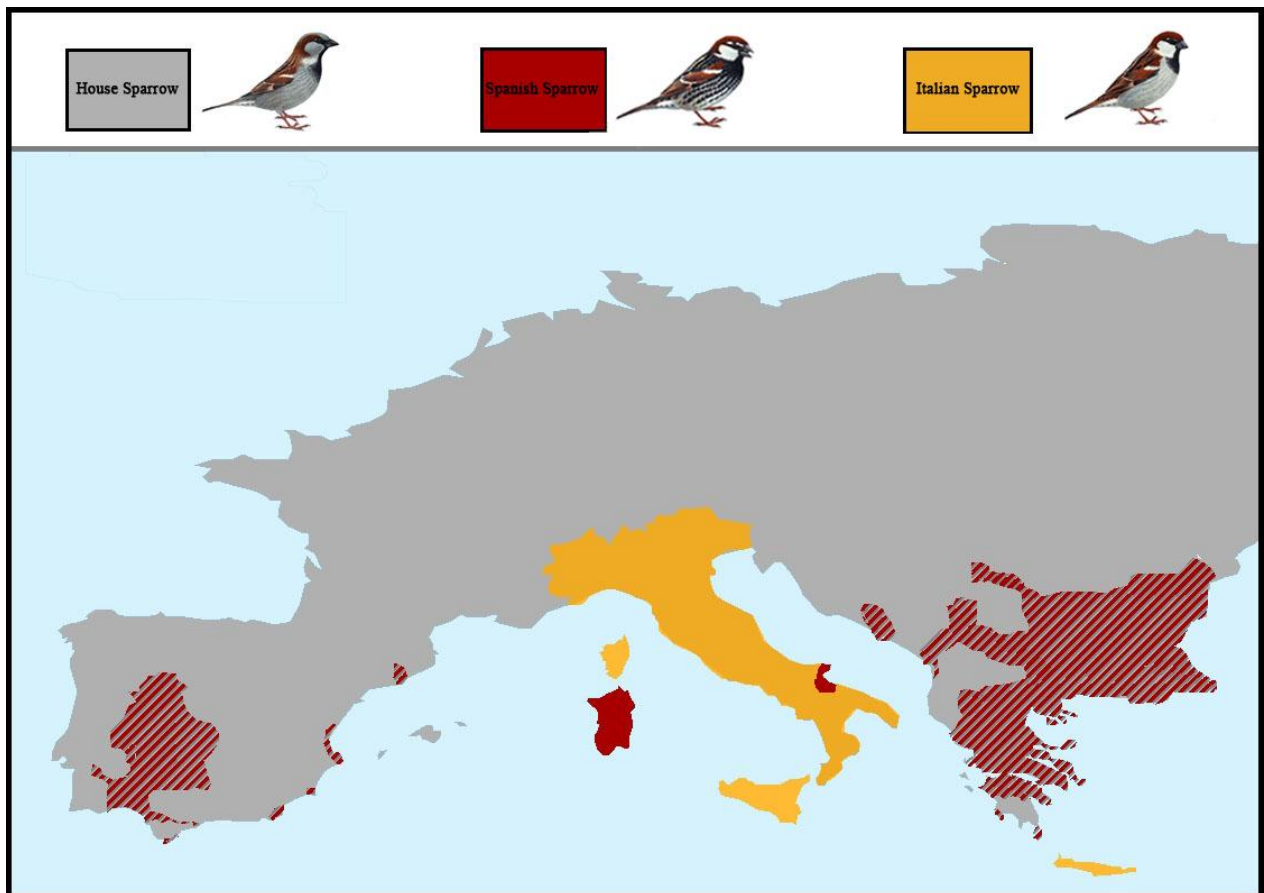
Figure 1. Breeding distribution of three *Passer* sparrow species North of the Mediterranean Sea: The house sparrow (grey), Italian sparrow (yellow) and Spanish sparrow (red). The hatched, red/grey areas indicate regions where the Spanish sparrow and house sparrow live in sympatry (Based on The EBCC Atlas of European Breeding Birds: Their Distribution and Abundance, Hagemeijer and Blair 1997).

No F1-hybrids have been found in the sympatric population in the study area in Spain, although some back-crossed individuals were detected (Hermansen, Haas and Saetre, unpublished data). Two general hypotheses may explain the lack of hybrids: (1) the house sparrow and the Spanish sparrow hybridize, but their offspring do not survive. I consider this hypothesis relatively unlikely due to the aforementioned successful homoploid hybrid speciation event in Italy (Hermansen et al. 2011) and the reported episodes of extensive hybridization elsewhere (Summers-Smith 1988). Also, in general, successful formation of F1-hybrids is rather common in birds although problems may arise in the later generations (e.g. Price 2007). This is also strongly supported by studies of the two sparrow species conducted

in aviaries in which viable F1-crossess have been obtained (Alonso 1984; Haas and Eroukhmanoff, unpublished data). (2) Premating barriers are strong. Since, I *a priori* consider the first hypothesis relatively unlikely, my working hypothesis is that premating barriers are important in separating the two sparrow species. Thus, this study aims to investigate factors that may prevent them from interbreeding.

There are in fact some differences between the house sparrow and the Spanish sparrow that may potentially contribute to the reduction of gene flow. A difference in plumage (explained later) and small differences in song may act as potential sexual barriers (Summers-Smith 1988). However, here I will focus on potential ecological barriers, including spatial and temporal segregation related to differences in habitat use, choice of breeding site, and timing of breeding.

A range of methods were used to study potential premating barriers between the two species. These include transects and radio-telemetry to investigate potential microspatial isolation; observations at breeding colonies to look for potential species differences in choice of nest site; examination of breeding phenology by catching birds to investigate beak color (males) and presence/absence of a breeding patch (females), traits that indicate when the bird is in breeding mode and finally; ecological niche modeling to investigate potential macrospatial isolation.

Materials and Methods

Study Taxa

The house sparrow

Due to its commensal relationship with humans, the house sparrow is one of the most widely distributed species of land birds in the world (Summers-Smith 1988). Its natural range includes large parts of the world such as Europe, North Africa, parts of Asia and the Indian subcontinent. From these parts, it has been introduced by humans, both intentionally and accidently, to virtually the rest of the world. Over most of its breeding range, the house sparrow occurs only in manmade habitats, such as farmland and cities (Saetre et al. 2012).

The male house sparrow has a small black bib with no black flanking on the body side. Its nape and crown are grey, and it has greyish cheeks. The females and juveniles are in contrast quite dull in plumage, having a light brown color above and pale greyish brown underparts (Summers-Smith 1963).

The house sparrow's domed nests are usually found under the roof of houses, or in any other holes and crevices in buildings, as well as behind the creepers on the walls of buildings. Another preferred nesting site is inside of the nests of large birds, such as storks and birds of prey, commonly seen in the study area in Spain. Even though it can build a nest practically anywhere, the construction of the nest is essentially the same, consisting of an exterior of hard dead grass or straw and an inside layer of softer materials (mostly feathers, but also occasionally hair, string, fine grasses, paper and bast) (Anderson 2006; Summers-Smith 1963). The house sparrow breeds in colonies, and feeds in homospecific or heterospecific flocks. It commonly spends a great deal of time on the ground foraging on seeds, especially in open areas. Seeds of cultivated crops are favoured, such as oats, wheat, barley, corn, millet and rice. But as an opportunistic bird, it takes advantage of a variety of food sources that come to hand (Anderson 2006; Summers-Smith 1988).

The Spanish sparrow

Breeding in tightly packed colonies and foraging in flocks that can number in thousands of individuals outside the breeding season, makes the Spanish sparrow one of the most social Palaearctic passerine bird. Also known as the willow sparrow, this species breeds in a rather thin band of latitude from southern Europe and Northern Africa eastwards to Turkestan. It

thus has a rather wide distribution, and over most of its range it is in sympatry with the house sparrow (Summers-Smith 1988) (Figure 1).

Striking in appearance, the Spanish sparrow has a large black bib spreading sideways on the breast and going over to heavy black flanking along the body sides. Its crown, nape and sides of the neck are chestnut red, with brightly white-colored cheeks. Females and juveniles are very similar to house sparrows, and it takes a trained eye to distinguish them (Summers-Smith 1988; Saetre and Tu, personal observation).

The Spanish sparrow makes use of open, deciduous woodland, plantations and roadside trees during the breeding season, and is above all found near water-sources, such as rivers. Where the Spanish sparrow occurs sympatrically with the house sparrow, their nests are in for the most part in trees and bushes, both openly and similar to the house sparrow, in the under-storey of nests of birds of prey, storks, herons and corvids. However, wherever the Spanish sparrows are free of house sparrows, they consume the role of their close relative, nesting in holes and creepers on man-made structures (Summers-Smith 1988). As a colonial nester, the breeding colonies in some places can be incredibly large; Gavrilov (1963) recorded colonies of 100,000 – 800,000 nests in Kazakhstan. The nests are firmly attached to trees, and consist of green grass, small plants, straws or leafy twigs taking an irregular, coarsely spherical form. The Spanish sparrow feeds predominantly on seeds of grasses and cultivated cereals (wheat, millet, barley and oats), and where it is widely distributed it is considered a serious crop pest (Summers-Smith 1988).

Study Site

The fieldwork was conducted during a six-week period in March-April 2011, and one week in April-May 2012. I made use of two study areas: 1) a major part of the largest province in Spain, Badajoz (38°52′44″N, 6°58′1″W, hereafter Badajoz province), where I by car crisscrossed the roads in search for Spanish and house sparrow breeding colonies, 2) the Marzala farm (4 km^2, 38°38′55.42″N, 7°12′55.53″W), mainly consisting of open oak tree forests and farmland, giving the area both farmland as well as semi-natural habitats (Figure 2).
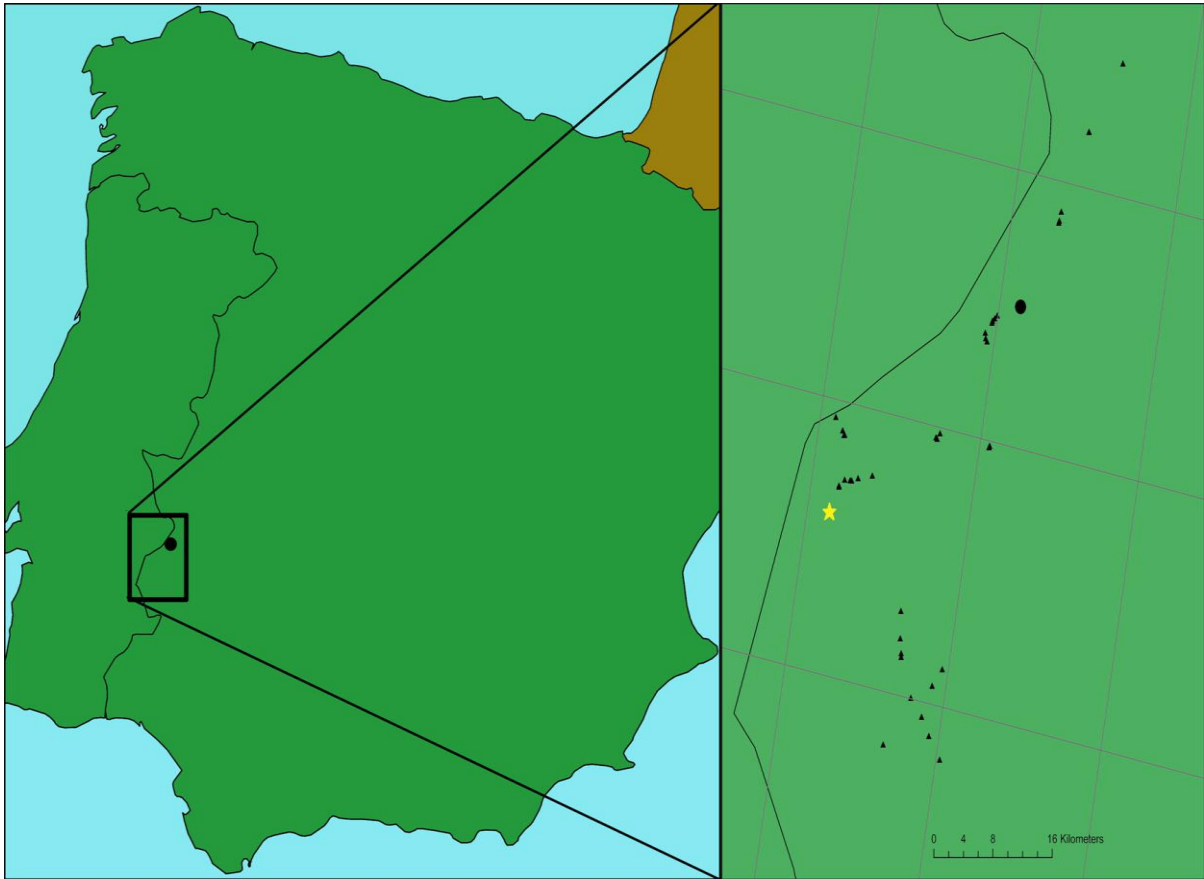
Figure 2. A map of the Iberian Peninsula including a zoomed in panel of the study area. The black dots in both panels show the location of Badajoz city, the yellow star represent Marzala farm and finally, the small black triangles represent the locations of identified sparrow colonies in Badajoz province.

Habitat Use

Radio-telemetry

Radio-tracking was conducted during March-April 2011 at Marzala Farm. For practical reasons we were not able to perform this in 2012. A total of 5 birds: 3 house sparrows and 2 Spanish sparrows were caught using mist nets at two spots within the farm for the purpose of radio-tracking. The equipment used was from Biotrack (Biotrack Ltd, Wareham, Dorset, United Kingdom), and consisted of radio-transmitters (PicoPip Ag392), a receiver and an antenna. The radio-transmitters weigh approximately 1 gram, i.e. less than four percent of the weight of an average house or Spanish sparrow, and have a range of 200-600 meters. The transmitters were attached to the birds with leg-loop harnesses (Rappole and Tipton 1991). The making and the dimensioning of the leg-loop harnesses were done according to Naef Daenzer's (2007) allometric function based on body mass data, but slightly modified to fit the transmitter birds in the best possible way. The procedure includes attaching two elastic strings

(loops) on both sides of the sender with superglue with a harness loop span of 50 mm. The result is a backpack-like radio-transmitter that fits well on the bird. The transmitter is then attached to the bird by pulling the loops over their thighs and placing the transmitter itself on the rump. It is of great importance to properly attach the radio-tags to avoid negative effects such as entanglement of the wing, leg or bill, which can lead to immobilization and fatal harm. The birds were tracked 1-5 times a day for a period of 1-2 weeks. The respective locations were recorded using a GPS (Garmin Norge AS, Dilling, Norway).

By using Google Earth version 6.2.2.6613 (Google Inc., Mountain View, CA 94043, USA), we measured the distance between the individuals' tracked locations and the nearest farm building/construction, river and pond.

Transects

I monitored the distribution of house and Spanish sparrows by systematically walking through the network of roads and paths at the Marzala Farm study area (approx. 4 km^2) and recording the position of the sparrows we observed (road transects). For any given observation we noted the approximate number of birds (most observations concerned birds in flocks), the species (pure house sparrow, pure Spanish or mixed flocks) and the habitat type (Oak tree forest, near buildings, near water-sources). The location of each observation site was recorded using a GPS. The road transects were repeated each week for the study period of 2011. A total of 5 transect walks were carried out.

The coordinates for all observations were exported to ArcGIS 9.3.1. (Esri, Redlands, CA), where they were plotted on a map of the study site. In addition, Google Earth was used to measure distance to the nearest farm building/construction, as well as river and pond.

Choice of Breeding Site

Observations of breeding colonies were made for an area consisting of large parts of Badajoz province during April-May 2012. A car was utilized for a three-day road-trip to find stork nests (a common nesting site for sparrows) along roads containing colonies of either house sparrows, Spanish sparrows, or a mix of both species. Quantification of the species was based on observation and a rough count of nests. A total of 38 colonies were recorded of which 18 were of mixed species and 20 of pure house sparrows.

Google Earth was utilized to measure the distance between the different colonies and the nearest city and crop field.

Breeding Phenology

A total of 193 sparrows (House sparrow: 53 females and 74 males; Spanish sparrow: 31 females and 35 males) were caught using mist nets at five different spots within the Marzala Farm study area in 2011. In addition to measures of beak height, beak length, tarsus length, wing length and bib size, I recorded phenological traits that indicate the breeding status of the bird. In males we recorded the beak color, ranging from pale (pre-breeding) to black (breeding mode) and in females the presence or absence of an incubation patch. When incubation is initiated the female plucks off feathers of her belly to produce a bare skin patch that keeps the eggs warm.

Hybrids

Catching

Male individuals were classified as pure house sparrow, Spanish sparrow or (if any) hybrid based on the plumage coloration (described above). Females were only classified to species (no hybrid class) as they are more similar in phenotype. In addition, in a parallel genetic study, all the individuals included in my study were genotyped for species diagnostic SNP-markers (Hermansen, Haas and Saetre, in prep). The information from that study is used here to help classify birds as pure Spanish, pure house or F1-hybrids.

Nest boxes

To investigate the reproductive success of the potential heterospecific pairs relative to pure pairs of both species, we wanted to measure and compare the date of first egg, clutch size, hatching success and fledging success. 125 nest boxes for the sparrows were set up around the Marzala Farm at the start of 2011, scattered over five stations with 28 boxes each. These five stations were situated both close to the farm-houses and in more semi-natural sites. However, the sparrows did not make use of the nest boxes the first breeding season of 2011. Although some of the nest boxes were utilized the second season, we have no usable data.

Permissions for radio-tracking, catching and measuring birds were obtained from the appropriate authorities in Spain.

Ecological Niche Modeling (ENM)

To predict the potential distribution of the Spanish sparrows and the ecological factors that may explain their distribution, I made use of ecological niche modeling. Ecological niche models were constructed for the Spanish sparrow using the maximum entropy method implemented by the program MaxEnt version 3.3.3k (Phillips et al. 2004; 2006). This method finds the model that best predicts the probability distribution of a species as a function of environmental variables. The "best" probability distribution is the one that is consistent with the empirical presence data and the environmental conditions connected to it, and also makes the least constrained assumptions about the rest of the distribution (a distribution as close to uniform as possible) (Phillips et al. 2004). Ecological niche models use geo-referenced point locality data (presence) in combination with environmental data (i.e. climate and land cover) in the form of GIS data layers to identify areas that have the most suitable habitat for the focal species (Glor and Warren 2011; Soberón and Peterson 2005). I only modeled the distribution of Spanish sparrows due to the fact that the distribution of house sparrow is so strongly associated with human settlements, and it is basically found everywhere humans reside. Thus, making a model for such a species would not give any fruitful results. A more useful approach would be to utilize absence-points (i.e. where the house sparrow does not exist), but such data are difficult to obtain.

MaxEnt requires all layers to have the same projection, resolution and geographic boundaries.

Data

Presence data, with a resolution of 10 x 10 km, for the Spanish sparrow on the Iberian Peninsula and the Canary Islands were extracted from breeding bird atlases (Spain - Martín and Del Moral 2003; Portugal – Rufino 1989) All presence points deemed by the authors to indicate probable or confirmed breeding were digitalized and transferred to ArcGIS 9.3.1. A total of 897 presence locations were acquired.

The land cover information was taken from the European land cover and land-use classification CORINE (Coordination of Information on the Environment) (www.eea.europa.eu/data-and-maps/data). These layers have a resolution of 100m x 100m.

Climate and altitude data was taken from the WorldClim (19 climate and 1 altitude layer, www.worldclim.org) database. The climate data are biologically relevant temperature and precipitation layers set to a resolution of 2.5 arc-minutes.

Data preparation

As the land cover information comes in one single layer, we had to split and modify it into multiple layers, each containing a single land cover class. The 12 most common land cover layers were extracted, which together cover 85 % of the Iberian Peninsula and the Canary Islands. These layers have a resolution of 100m x 100m, so to match the bioclimatic layers of 2.5 arc-minutes, modifications in ArcGIS 9.3.1 (Spatial Analyst Tools,) were made. After the transformation, each cell in a given layer contained information on the coverage of a specific land cover class. Also, Worldclim data was re-projected to get the same projection as the CORINE Land cover data (ETRS_1989_LAEA_L52_M10).

Collinearity is a well-known problem with these kinds of spatial data (reviewed in Dormann et al. 2012) because it may result in the misidentification of relevant predictors in an analysis. To reduce the problem with highly correlated variables we constructed a correlation matrix with all 32 environmental layers (Spatial Analyst Tools, ArcGIS 9.3.1). We set the threshold for acceptance of a variable at $r > 0.75$. When a pair of variables exhibited $r > 0.75$ we eliminated the variable we considered to be less biologically relevant and/or harder to interpret (cf. Rissler and Apodaca 2007). Using threshold values to eliminate highly correlated variables in spatial data sets to reduce the problems of collinearity have been shown to be as efficient as several of the more elaborate methods developed to deal with the same problem (Dormann et al. 2012) After the elimination process, a total of seven bioclimatic- and twelve land cover variables (Table 1; Table 2) were retained for further analyses.

Table 1. An overview of the bioclimatic variables included in the modeling (www.worldclim.org/bioclim).

Bioclimatic variables
BIO3 = Isothermality (Monthly/annual temperature range)
BIO4 = Temperature seasonality (standard deviation*100)
BIO8 = Mean temperature of wettest quarter
BIO9 = Mean temperature of driest quarter
BIO15 = Precipitation seasonality (Coefficient of variation)
BIO16 = Precipitation of wettest quarter
BIO17 = Precipitation of driest quarter.

Table 2. An overview of the land cover variables included in the modeling (European Commission 1994).

Land cover variables

Lc12: Non-irrigated arable land = Cereals, legumes, fodder crops, root crops and fallow land

Lc13: Permanently irrigated land = crops irrigated permanently or periodically, using a permanent infrastructure (irrigation channels, drainage network)

Lc17: Olive groves = areas planted with olive trees

Lc20: Complex cultivation patterns = Juxtaposition of small parcels of diverse annual crops, pasture and/or permanent crops

Lc21: Land principally occupied by agriculture, with significant areas of natural vegetation = Areas principally occupied by agriculture, with significant areas of natural vegetation

Lc22: Agro-forestry areas = Annual crops or grazing land under the wooded cover of forestry species

Lc23: Broad-leaved forest = Vegetation formation composed principally of trees, including shrub and bush understories, where broad-leaved species predominate

Lc24: Coniferous forest = Vegetation formation composed principally of trees, including shrub and bush understories, where coniferous species predominate

Lc25: Mixed forest = Vegetation formation composed principally of trees, including shrub and bush understories, where neither broad-leaved nor coniferous species predominate

Lc26: Natural grasslands = Low productivity grassland, often situated in areas of rough uneven ground

Lc28: Sclerophylous vegetation = Bushy sclerophylous vegetation, includes maquis and garrigue.

Lc29: Transitional woodland-shrub = Bushy or herbaceous vegetation with scattered trees.

Modeling and evaluation

 I ran two models, one to predict the distribution of Spanish sparrows on the Iberian Peninsula and Canary Islands, and one to predict the distribution in south, and central Europe and Turkey. In the first model, I used both climate- and land cover data. In the second model I only used climate data due to the lack of land cover data for Greece.

In both runs I let the models train on a subsample of the present points extracted from the Spanish and Portuguese breeding bird atlases. Based on the training result MaxEnt extrapolated the distribution of Spanish sparrows in Iberia and Canary Islands, and across Europe and Turkey, respectively. The settings for the runs: 15 replicates were created and then averaged in combination with withholding 30 % of the data for testing (random test percentage) (cf. Thuiller et al. 2004; Virkkala et al. 2005), allowing us to test the performance of the model while taking advantage of all available data without having an independent dataset; I used subsample as the replicated run type as we have a large dataset; 5000 iterations to allow the model to have enough time for convergence to avoid over- and under-prediction; and finally, random seed for random sets of test/training for each run. Finally, the resulting ASCII file from MaxEnt was exported to ArcGIS and viewed.

The models were tested for accuracy and predictive power by using the area under the curve (AUC) of a receiver operating characteristic (ROC) plot (cf. Luoto et al. 2007; Thuiller et al. 2004; Pearson et al. 2004; Virkkala et al. 2005) and true skill statistic (TSS) (Allouche et al. 2006; Coetzee et al. 2009). The AUC is a threshold-independent method, and is the probability that a randomly chosen presence point will be scored above a randomly chosen absence point. A model that has the same performance as a random one will be given an average AUC of 0.5, and a perfect model will achieve a value closer to 1 (Fielding and Bell 1997). The AUC-values were extracted from the MaxEnt output files for the 15 replicates (for each model) and averaged. Further, TSS, a threshold-dependent method, was also used to evaluate the model. The data for estimating TSS was extracted from the html file, sample predictions and background predictions from the MaxEnt output files. To estimate the sensitivity and specificity (below) of the models we used the logistic prediction values for the sample predictions, logistic value for background predictions and the 10 percentile threshold value located in the html file. Sensitivity and specificity was then later used to calculate TSS = Sensitivity + Specificity-1. Finally, the TSS for all the replicates (for each model) was averaged. The scores are as following: $0 - 0.40$ = Poor; $0.40 - 0.55$ = Fair; $0.55 - 0.70$ =

Good; $0.70 - 0.85 =$ Very good; $0.85 - 0.99 =$ Excellent, and finally; $0.99 - 1 =$ Perfect (Monserud and Leemans 1992)

MaxEnt tested the importance of the variables included in the model by conducting a jackknife test of test gain.

Sensitivity (the ability to correctly predict absence) = ((Number of values above the threshold value in the logistic prediction values for test data)/ (Number of values above the threshold value in the logistic prediction values for test data + Number of values below the threshold value in the logistic prediction values for test data))

Specificity (the ability to correctly predict absence) = ((Number of values below the threshold value in logistic value)/ (Number values above the threshold value in logistic value + Number of values below the threshold value in logistic value))

Statistical Analyses

For the statistical analyses, SPSS version 15 (SPSS, Chicago, Ill., USA) was used.

Habitat use

Nested ANOVA was performed for all the radiotracked individuals' nearest distance to a building, river and pond to test if the respective variables have any effect on the individual's use of habitat. To avoid the risk of pseudo-replication, the "individual" variable was nested within species as a random factor. We then included distance to building, river and pond as covariates.

To assess if distance to building, river and pond have any effect on the distribution of the flocks (Pure house sparrow, pure Spanish sparrow and mixed) from transect data; three ANOVA analyses were performed with the distances to building, pond and river as response variables.

Choice of breeding site

I performed a type III generalized linear model (GLZ) with a probit link function to test if the distances to nearest building, city and field have any effect on the probability of a colony being of either pure house sparrows or containing any Spanish sparrow. Additionally, a Levene's test of equality of error variances was conducted to compare variances between pure

house sparrow colonies and mixed colonies in any of the three covariates. I did not include the two non-stork's nest colonies in this model.

Breeding phenology

I used the phenology data for beaks and incubation patch and plotted the 95 % confidence intervals for the mean.

Results

Habitat Use

The radio-tagged individuals showed clear differences in habitat use with respect to species. The house sparrows (n = 3) were at all times (when tracked) within a short range of farm buildings or other human constructions, such as pigsties and feeders. Moreover, none of the individuals were ever found far away from their previous tracked locations, indicating a small home range (Figure 3; Figure 4; Figure 5). In contrast, the Spanish sparrows (n = 2) were often further away from buildings, and were usually found closer to rivers or ponds (Figure 6; Figure 7). These results were further supported by the nested ANOVA analysis, where a significant effect of distance to pond was found ($F_{1,4}$ = 28.826, P = 0.001), the Spanish sparrow being closer to ponds than the house sparrow (Figure 8). The two species also differ in their distance to farm buildings, the house sparrow keeping a close proximity ($F_{1,4}$ = 14.297, P = 0.001) (Figure 8). Similarly, a species-difference in distance to river was detected, where Spanish sparrows occurred closer to rivers than house sparrows ($F_{1,4}$ = 9.415, P = 0.001) (Figure 8). "Species" explains on average 53.1 % of the total variation found in the model (Table 3).

The spatial distribution of birds in flocks (Pure house sparrow, pure Spanish sparrow and mixed) (Figure 9) follow the same general pattern as the individual tracked birds. The pure house sparrow flocks are significantly closer to farm buildings than pure Spanish sparrow flocks ($F_{2,80}$ = 23,429, P = 0.001) (Figure 10). Interestingly, the mixed flocks showed a tendency to be found at an intermediate distance to farm buildings (but with large variation). Distance to pond also showed a difference between the pure flocks with the Spanish sparrows being closer to ponds than house sparrows ($F_{2,80}$ = 6,979, P = 0.001) (Figure 10). The mixed flocks showed some tendency of being furthest away from ponds, but again with large variation (Figure 10). We found no significant effect of distance to river for the road transect analysis.

Choice of Breeding Site

The type III generalized linear model (GLZ) showed that the colonies containing any Spanish sparrows (i.e. mixed colonies) were significantly further away from cities compared to the colonies with no Spanish sparrows (i.e. pure house sparrow colonies) ($\chi^2_{1,\ 37}$= 4.915, P = 0.027) (Figure 11). However, we found no significant effect of distance to field (Unequal

variances, Levene's test = 0.034). Furthermore, field observations revealed a clear difference between the two species within the mixed colonies. In all mixed colonies, the house sparrow seemed to occupy the nest spots closest to the center of the stork's nest, while the Spanish sparrow occupied the spots on the outer parts of the nest. Finally, I also observed two large pure Spanish sparrow colonies with at least 1500-2000 pairs that were not in stork's nests.

Breeding Phenology

The means and 95 % confidence intervals of the date when females showed incubation patches and the males black beaks are presented in Table 4. We identified a minimum of 13 days difference in breeding time between the two species (Figure 12). The house sparrow males already had black beaks when we started our fieldwork 3rd of March 2011, which means they were in breeding mode even before we arrived. The first female house sparrows to have a present incubation patch was documented on the 25th of March. In contrast, the first documented Spanish sparrow male to gain black beak was on 16th of March (i.e. min. 13 days after the house sparrows). We never documented the presence of incubation patch in the Spanish sparrow females.

Hybrids

No F1-hybrids were caught during the period of catching. Only a few back-crossed individuals were identified through genotyping (House sparrows with introgressed Spanish alleles; Hermansen, Haas and Saetre, in prep).

Ecological Niche Modeling (ENM)

I made two models that predict the distribution of the Spanish sparrow, displaying areas of habitat suitability on a map. The first model for the Spanish sparrow on the Iberian Peninsula is characterized by a high average AUC statistic of 0.822 with an average standard deviation of 0.007, and a fair average TSS-value of 0.548 (Figure 13; Table 5). The model predicts higher habitat suitability in the central and southern parts of the Iberian Peninsula and the Canary Islands (Figure 14), and is primarily driven by the environmental variable BIO17 = Precipitation of driest quarter (63.6 %) (Table 6). This is supported by the jackknife test on the importance of the variables included in the model (Figure 15). The response curve for BIO17 shows that the presence probability of Spanish sparrow decreases with higher precipitation (Figure 16).

The second model predicting the distribution across Europe has a high AUC of 0.822 with a standard deviation of 0.010 and a TSS-value of 0.550 (Figure 17; Table 5). The model predicts high habitat suitability in Spain, Sardinia (Italy), Sicily (Italy), Greece, Cyprus and Turkey (Figure 18). Similar to the previous model, the variable BIO17 is identified as the most important variable for explaining the distribution of Spanish sparrow (Table 7), this again supported by the jackknife test (Figure 19). Figure 20 shows a similar response curve for the BIO17.
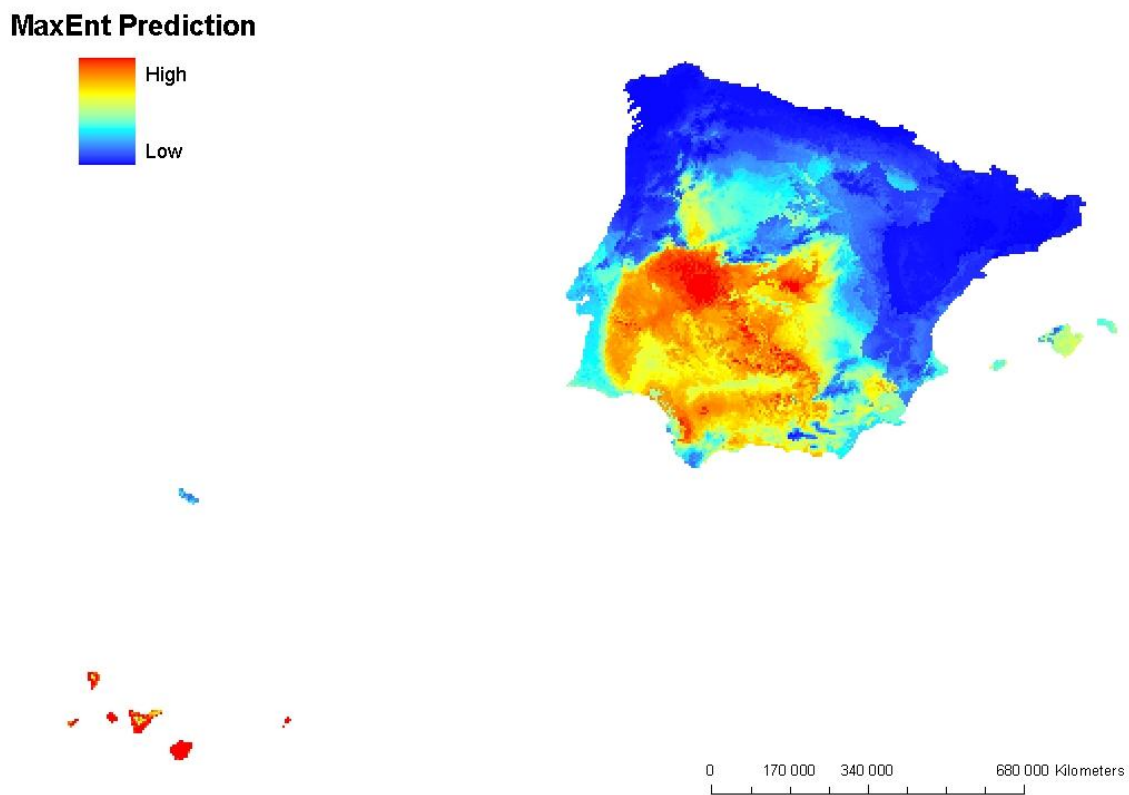

MaxEnt Prediction

High

Low

0   170 000   340 000   680 000 Kilometers

Figure 14. The ENM for the Spanish sparrow on the Iberian Peninsula, the Balearic Islands and the Canary Islands. The suitabilities are listed in a graded coloration from red (= highest suitability) to blue (= lowest suitability).
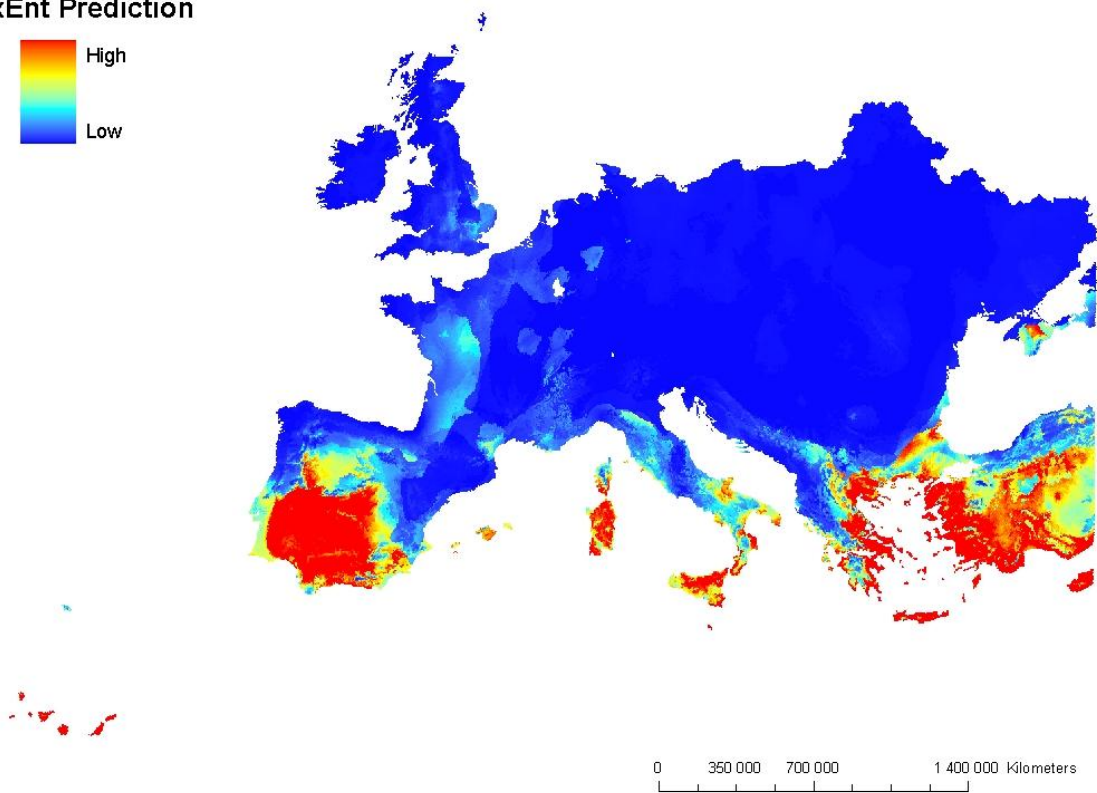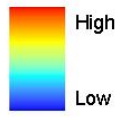
Figure 18. The ENM for the Spanish Sparrow in Europe. The suitabilities are listed in a graded coloration from red (=highest suitability) to blue (=lowest suitability)

Discussion

Although the house sparrow and Spanish sparrow are similar in many ways, I did find clear differences between them in traits that likely function as premating barriers, and thus may help to explain the preservation of their species integrity. I have investigated premating reproductive barriers including micro-habitat differences, choice of breeding site, breeding phenology and broad-scale distribution patterns. The house sparrow is a human-commensal bird and its habitat use and choice of breeding site is typically close to buildings and cities. In contrast, the Spanish sparrow keeps a larger distance from humans and takes advantage of more semi-natural habitats, often near rivers and ponds. Investigation of their breeding phenology (color of beak and incubation patch) gave evidence of a large difference in breeding time.

Differences in Habitat Use

Studies of habitat use by wild animals frequently use radio tracking. However, estimation of habitat use based on radio-tracking of animals often poses large difficulties. First of all, a large sample size is hard to obtain (Aebischer and Robertson 1993). This is mainly caused by the economical costs of the equipment and the time consumption of the tracking. Although I was only able to obtain data for a total of 5 birds, the apparent difference between the species was statistically supported and the results were consistent with the road transect data and field observations. Another potential problem with radio tracking is a bias in sampling sites. All the tracked house sparrow individuals were from one catching spot, namely near the main farm building of Marzala farm, while Spanish sparrows were caught near bushes. This might have influenced the results. On the other hand, the road transects displayed similar patterns, as did the tracking of nest sites in Badajoz province. The house sparrow's close proximity to farm buildings as well as to cities reflects their commensal relationship with humans. The Spanish sparrow's closeness to rivers (i.e. semi-natural habitat) and ponds may reflect a preference for open bodies of water, but perhaps more importantly, rivers are typically surrounded by dense bushes that the sparrows thrive in. However, a direct comparison between the species cannot be made due to the fact that both species were tracked for the same time period albeit being at different stages in the breeding cycle. Hence, whereas the house sparrows were tracked in the midst of their breeding period, the Spanish sparrows were tracked partly before and partly at the start of their breeding period. This makes it difficult to compare the results directly. For

instance, the Spanish sparrow's larger home range might partly be an artifact of the stage they were tracked. Nevertheless, the results display clear spatial segregation that would reduce encounters between the species during breeding time. Thus, habitat segregation seems to represent a barrier between the species that may reduce gene flow.

Differences in Choice of Nest Site

In Badajoz, the white stork (*Ciconia ciconia)* is a very common bird. Their large nests are usually placed on top of electricity masts along roads, and can contain large colonies of sparrows. The two Spanish sparrow colonies I observed were not in stork's nests, but in oak-tree forests with nests placed in trees. Interestingly, these colonies were the largest recorded ones, containing no less than a few thousand pairs. These sparrows feed off crops in large flocks and gather nesting materials nearby. The Spanish sparrows seem to be nomadic, choosing new breeding sites frequently (following resources), seeing as the locals said they had never seen the large colonies there before. This also corresponds to the fact that the Spanish sparrow colonies found on Marzala farm in 2011 had all scattered and were not to be found in 2012.

In accordance with results from habitat use, pure house sparrow colonies were significantly closer to cities, while colonies also containing Spanish sparrows were further away. I also detected a within-stork nest difference in the mixed colonies. In such colonies, the house sparrows place their nests in the center of the stork's nest, while Spanish sparrows seem to occupy the ones in the outer parts of the large nest. This is likely the result of the fact that the house sparrow begins breeding first and occupies the best sites with presumably lower predation risk. Additionally, mixed colonies such as these may provide a platform in which hybridization can occur. However, as an immense number of Spanish sparrows breed in the above mentioned colonies, a large proportion can more or less be excluded from the possibility of contributing to gene flow between the species, and gene flow from Spanish sparrows into the smaller house sparrow populations may be more likely.

A bias in sampling should also be considered here, as we only traveled on roads giving an over-representation of colonies that have an easy access. In particular, I consider it likely that I would have missed a large number of pure Spanish sparrow colonies by this sampling strategy.

Difference in Breeding Time

Before starting the breeding period, the two species usually amass in large heterospecific flocks (i.e. winter flocks). This naturally allows a lot of contact between the species. However, when their breeding seasons are closing in, they split from the large winter flocks and form smaller flocks (Anderson 2003; Summers-Smith 1988). This is when they start looking for nest sites and possibly start finding their mates. As I detected a minimum of 2 weeks distance in breeding time between the two species it is reasonable to think that that the timing of leaving the winter flocks and then forming conspecific flocks would serve as a reproductive barrier. I do not know exactly how large this gap is since the house sparrows had already started breeding when we started our fieldwork, so the gap in breeding time may be larger than 2 weeks, which would further strengthen the barrier. The large distance between the species' breeding time has also been supported by aviary studies (Haas and Eroukhmanoff, unpublished data). A difference in breeding time in birds usually reflects their response to different environmental cues, such as food availability (i.e. they make use of different food resources for their young), but I did not investigate this. I cannot ignore the fact that their breeding seasons do overlap to some extent, leaving a possibility of hybridization. However, these sparrows are usually monogamous and breed with the same mate throughout the breeding season (Anderson 2006).

Ecological Niche Modeling and Evaluation

Studies of reproductive isolation usually focus on areas of sympatry and ignore geographic isolation as part of total reproductive isolation. Ecological differences on a broad-scale, however, are important in speciation and ought to be considered (Kay 2006). The two models for the Spanish sparrow based on the distribution data clearly suggest that the distribution of the Spanish sparrow is principally driven by dry habitats. The fact that one variable can explain 63.6 % of the driving force behind its distribution indicates niche-conservatism. In contrast, the house sparrow's driving force, being tightly associated with human settlements and urban habitats, can most probably be explained by many diverse climate and land-cover variables. As the two species' distributions are likely to be explained by different variables, it is reasonable to assume there also are fundamental differences in ecology. The Spanish sparrows' more limited distribution also contributes to the reduction of contact between the species on a broad scale.

The models I made had high AUC statistics and usable TSS values, and they also fit well with the actual distribution of the Spanish sparrow (Figure 1). The reason for evaluating the model with both methods is due to the relatively new criticism of the well-used AUC. This evaluation method has been used by a wide range of distribution modeling studies, and has long been considered to be the standard method to assess the predictive power of the models (Luoto et al. 2007; Thuiller et al. 2004; Virkkala et al. 2005). However, Lobo et al. 2008 harshly criticize the method saying that the AUC value is partly determined by the grain size (scale) and the range of distribution compared with the extent of the study area. Thus, the AUC may increase due to the disproportionate relationship between a large study area and a species' narrow geographical range. Secondly, it weighs omission (= misclassification of absences) and commission errors (= misclassification of presences) equally, which gives both errors the same importance (Elith et al. 2006, Phillips et al. 2006). One should consider a misclassification of absences (commission errors) a more severe drawback than a misclassification of presence according to Lobo et al. (2008). TSS has been suggested as an alternative, which is supposed to correct for both drawbacks: the problem of prevalence and equal weighting of omission and commission errors (Allouche et al. 2006).

In distribution modeling, there are also other pitfalls to consider. There are certainly several kinds of biases present, such as the bias in sampling, which usually occurs near roads or places where the accessibility is easy (Edwards et al. 2006; Wollan et al. 2008). Obviously, if the recoded presences do not represent the actual distribution of a species adequately, the quality of the modeling will be affected (Edwards et al. 2006). Overall, my strategy has attempted to make a model that has included the most important variables and is evaluated with methods considered good for such modeling.

Relative Importance of Premating Barriers

Other studies have combined different mechanisms of isolation (behavioral, temporal, spatial, hybrid viability) to estimate the relative contribution of different barriers to reproductive isolation. These include studies of *Drosophila* species (Coyne and Orr 1989; 1997), *Asellus aquaticus* isopods (Eroukhmanoff et al. 2011), *Chorthippus* grasshoppers (Bailey et al. 2004), *Mimulus* monkeyflowers (Ramsey et al. 2003), *Costus* gingers (Kay 2006) and *Ficedula* flycatchers (Saetre and Saether 2010). Coyne and Orr's studies on *Drosophila* yielded important conclusions in relation to sympatric species and the relative strength of barriers. They show that at low genetic distances prezygotic barriers are usually far stronger than

postzygotic barriers in sympatric species (Coyne and Orr 1989). This pattern may be explained by reinforcement of prezygotic isolation by natural selection against sterile or inviable hybrids. The house sparrow and the Spanish sparrow, however, are able to produce fertile and viable hybrids. What is then maintaining the seemingly strong premating barriers between them? One possibility may be that hybrids formed may be disadvantaged, not because they have reduced fertility or viability, but because they fall between two different ecologies which make them unfit for any habitat (Coyne and Orr 1989). This is consistent with the lack of hybrids and the large difference between habitat use of the house sparrow and Spanish sparrow found in this study. Moreover, assortative mating is further strengthened if hybrids have an intermediate or a different breeding time relative to its parental species and/or if the courtship behavior of the hybrids makes it hard for them to secure mates (Coyne and Orr 1989). However, data on active mate choice in this species complex is lacking.

In nature, isolating mechanisms act sequentially, so that a given barrier can only decrease the potential gene flow that is not already excluded by earlier acting barriers (Coyne and Orr 1989). In addition, isolating barriers may work synergistically and so, two joint barriers may contribute more to isolation than one relatively strong barrier (Eroukhmanhoff et al. 2011). Multiple barriers between the species were indeed found in this study, together with the lack of F1-hybrids, I therefore expect premating barriers to be strong. However, this study does not yield data on which the relative contribution of different premating barriers can be quantified. The extent of sexual isolation via mate choice is still unknown, and it would have been better to estimate the habitat separation on a fine scale during the same time of the breeding seasons of the two species, at the time when they are choosing mates. The temporal difference would also be more accurately estimated by using the actual timing of pair formation. However, the evidence presented here nevertheless highlights the importance of premating barriers for the maintenance of this species complex.

The hybrid Italian sparrow is a living proof that isolating mechanisms present today have somehow been broken down in the past. If this event happened before the house sparrow became strongly commensal, the house sparrow and the Spanish sparrow were likely to be more similar ecologically. Specifically, they may have had more similar habitats creating a suitable habitat for the hybrid to be born into. Timing of breeding was also likely to be more synchronized since commensalism promotes food availability, thus leading to longer and earlier reproductive periods (see Saetre et al. 2012 and references therein). It is then likely that as the house sparrow became more commensal, habitat- and temporal isolation that was

reduced or absent at that time may have over time led to a stronger differentiation in ecology leading to the barriers that are present today. It may be possible to test this hypothesis using ancient DNA analysis of fossil sparrows, provided one could find adaptive genetic changes associated with adaptation to commensalism.

Conclusion

In this project, I have investigated multiple aspects of reproductive isolation between the house sparrow and the Spanish sparrow, all the way from the local scale to the broad geographic scale. All the methods I applied to accomplish this yielded results that showed significant segregation between the species. This thesis therefore presents evidence suggesting multiple potential isolation mechanisms that reduce gene flow between the two sparrow species, keeping the interbreeding at a minimum. I find that spatial segregation on both a fine scale and on a large geographic scale constricts the contact between the species. Moreover, the ecological niche models suggest fundamental ecological differences. Non-synchronized timing of breeding also represents a significant barrier. Although I covered many aspects of the isolating mechanisms between the house sparrow and the Spanish sparrow, further questions have been generated. Is there any sexual isolation between the species? Does the sharing of stork's nests affect the probability of hybridizing? More generally, how does the structure of the landscape affect contact between the species? Finding answers to these questions may complete our understanding of reproductive isolation in this species complex.

References

Aebischer, N. J., P. A. Robertson, and R. E. Kenward. 1993. Compositional analysis of habitat use from animal radio-tracking data. Ecology **74**:1313-1325.

Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology **43**:1223-1232.

Alonso, J. C. 1984. On the molt of Spanish sparrows (*Passer hispaniolensis)* and house sparrows (*Passer domesticus)* in Iberia. Journal Fur Ornithologie **125**:209-223.

Anderson, T. R. 2006. Biology of the ubiquitous house sparrow: from genes to populations. Oxford Univeristy Press, Inc.

Bailey, R. I., C. D. Thomas, and R. K. Butlin. 2004. Premating barriers to gene exchange and their implications for the structure of a mosaic hybrid zone between *Chorthippus brunneus* and *C. jacobsi* (Orthoptera : *Acrididae*). Journal of Evolutionary Biology **17**:108-119.

Balakrishnan, C. N. and M. D. Sorenson. 2006. Song discrimination suggests premating isolation among sympatric indigobird species and host races. Behavioral Ecology **17**:473-478.

Burke, J. M. and M. L. Arnold. 2001. Genetics and the fitness of hybrids. Annual Review of Genetics **35**:31-52.

Coetzee, B. W. T., M. P. Robertson, B. F. N. Erasmus, B. J. van Rensburg, and W. Thuiller. 2009. Ensemble models predict important bird areas in southern Africa will become less effective for conserving endemic birds under climate change. Global Ecology and Biogeography **18**:701-710.

Coyne, J. A. and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. Evolution **43**:362-381.

Coyne, J. A. and H. A. Orr. 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution **51**:295-303.

Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach. 2012. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography **35** : 001- 020.

Edwards, T. C., Jr., D. R. Cutler, N. E. Zimmermann, L. Geiser, and G. G. Moisen. 2006. Effects of sample survey design on the accuracy of classification tree models in species distribution models. Ecological Modelling **199**:132-141.

Elgvin, T. O., J. S. Hermansen, A. Fijarczyk, T. Bonnet, T. Borge, S. A. Saether, K. L. Voje, and G.-P. Saetre. 2011. Hybrid speciation in sparrows II: a role for sex chromosomes? Molecular Ecology **20**:3823-3837.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography **29**:129-151.

Eroukhmanoff, F., A. Hargeby, and E. I. Svensson. 2011. The role of different reproductive barriers during phenotypic divergence of isopod ecotypes. Evolution **65**:2631-2640.

Fielding, A. H. and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation **24**:38-49.

Funk, D. J., K. E. Filchak, and J. L. Feder. 2002. Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica **116**:251-267.

Gavrilov, E. I. 1963. The biology of the Eastern Spanish sparrow, *Passer hispaniolensis transcaspicus Tschusi*, in Kazakhstan. Journal of the Bombay Natural History Society **60**:301-317.

Glor, R. E. and D. Warren. 2011. Testing ecological explanations for biogeographic boundaries. Evolution **65**:673-683.

Grant, P. R. and B. R. Grant. 1992. Hybridization of bird species. Science **256**:193-197.

Grant, P. R. and B. R. Grant. 1996. Speciation and hybridization in island birds. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **351**:765-772.

Hagemeijer, E. J. M. and M. J. E. Blair. 1997. The EBCC atlas of european breeding birds: their distribution and abundance. T & A. D Poyser, London.

Hermansen, J. S., S. A. Saether, T. O. Elgvin, T. Borge, E. Hjelle, and G. P. Saetre. 2011. Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow. Molecular Ecology **20**:3812-3822.

Hettyey, A. and P. B. Pearman. 2003. Social environment and reproductive interference affect reproductive success in the frog *Rana latastei*. Behavioral Ecology **14**:294-300.

Hubbs, C. L. 1955. Hybridization between fish species in nature. Systematic Zoology **4**:1-20.

Kay, K. M. 2006. Reproductive isolation between two closely related hummingbird-pollinated neotropical gingers. Evolution **60**:538-552.

Koeniger, N. and G. Koeniger. 2000. Reproductive isolation among species of the genus *Apis*. Apidologie **31**:313-339.

Littlejohn, M. J. 1965. Premating isolation in the Hyla Ewingi complex (*Anura, Hylidae*). Evolution **19**:234-243.

Lobo, J. M., A. Jimenez-Valverde, and R. Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography **17**:145-151.

Luoto, M., R. Virkkala, and R. K. Heikkinen. 2007. The role of land cover in bioclimatic models depends on spatial resolution. Global Ecology and Biogeography **16**:34-42.

Martín, A. and J. C. Moral. 2004. Atlas de las aves reproductoras de España. Organismo Autónomo Parques Nacionales.

Martin, N. H. and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. Evolution **61**:68-82.

Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press.

Millington, S. J. and P. R. Grant. 1984. The breeding ecology of the cactus finch *Geospiza scandens* on isla Daphne Major, Galapagos. Ardea **72**:177-188.

Monserud, R. A. and R. Leemans. 1992. Comparing global vegetation maps with the kappa-statistic. Ecological Modelling **62**:275-293.

Morrow, J., L. Scott, B. Congdon, D. Yeates, M. Frommer, and J. Sved. 2000. Close genetic similarity between two sympatric species of tephritid fruit fly reproductively isolated by mating time. Evolution **54**:899-910.

Naef-Daenzer, B. 2007. An allometric function to fit leg-loop harnesses to terrestrial birds. Journal of Avian Biology **38**:404-407.

Nagel, L. and D. Schluter. 1998. Body size, natural selection, and speciation in sticklebacks. Evolution **52**:209-218.

Nuechterlein, G. L. and D. Buitron. 1998. Interspecific mate choice by late-courting male western grebes. Behavioral Ecology **9**:313-321.

Pearson, R. G., T. P. Dawson, and C. Liu. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. Ecography **27**:285-298.

Phillips, S. J., M. Dudík, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Page 83 Proceedings of the twenty-first international conference on Machine learning. ACM, Banff, Alberta, Canada.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling **190**:231-259.

Pierotti, R. and C. A. Annett. 1993. Hybridization and male parental investment in birds. Condor **95**:670-679.

Price, T. 2007. Speciation in birds. Roberts and Co.

Ramsey, J., H. D. Bradshaw, and D. W. Schemske. 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (*Phrymaceae*). Evolution **57**:1520-1534.

Rand, D. M. and R. G. Harrison. 1989. Ecological genetics of a mosaic hybrid zone - mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. Evolution **43**:432-449.

Rappole, J. H. and A. R. Tipton. 1991. New harness design for attachment of radio transmitters to small passerines. Journal of Field Ornithology **62**:335-337.

Rice, W. R. and G. W. Salt. 1990. The evolution of reproductive isolation as a correlated character under sympatric conditions - experimental-evidence. Evolution **44**:1140-1152.

Rissler, L. J. and J. J. Apodaca. 2007. Adding more ecology into species delimitation: Ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). Systematic Biology **56**:924-942.

Rufino, R. 1989. Atlas das aves que nidificam em Portugal Continental. Ministério do Plano e da Administração do Territorio, Serviço Nacional de Parques, Reservas e Conservação da Natureza.

Saetre, G. P., T. Moum, S. Bures, M. Kral, M. Adamjan, and J. Moreno. 1997. A sexually selected character displacement in flycatchers reinforces premating isolation. Nature **387**:589-592.

Saetre, G. P., S. Riyahi, M. Aliabadian, J. S. Hermansen, S. Hogner, U. Olsson, M. F. G. Rojas, S. A. Saether, C. N. Trier, and T. O. Elgvin. 2012. Single origin of human commensalism in the house sparrow. Journal of Evolutionary Biology **25**:788-796.

Saetre, G. P. and S. A. Saether. 2010. Ecology and genetics of speciation in *Ficedula* flycatchers. Molecular Ecology **19**:1091-1106.

Seehausen, O. and J. J. M. van Alphen. 1998. The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). Behavioral Ecology and Sociobiology **42**:1-8.

Soberon, J. and A. Townsend Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics **2**:1-10.

Summers-Smith, D. 1963. The House-Sparrow. (Collins).

Summers-Smith, J. D. 1988. The sparrows. T. & A.D. Poyser, Calton, England.

Thuiller, W., M. B. Araujo, and S. Lavorel. 2004. Do we need land-cover data to model species distributions in Europe? Journal of Biogeography **31**:353-361.

Tubaro, P. L. and D. A. Lijtmaer. 2002. Hybridization patterns and the evolution of reproductive isolation in ducks. Biological Journal of the Linnean Society **77**:193-200.

Veen, T., T. Borge, S. C. Griffith, G. P. Saetre, S. Bures, L. Gustafsson, and B. C. Sheldon. 2001. Hybridization and adaptive mate choice in flycatchers. Nature **411**:45-50.

Virkkala, R., M. Luoto, R. K. Heikkinen, and N. Leikola. 2005. Distribution patterns of boreal marshland birds: modelling the relationships to land cover and climate. Journal of Biogeography **32**:1957-1970.

Wiernasz, D. C. and J. G. Kingsolver. 1992. Wing melanin pattern mediates species recognition in *Pieris occidentalis*. Animal Behaviour **43**:89-94.

Wollan, A. K., V. Bakkestuen, H. Kauserud, G. Gulden, and R. Halvorsen. 2008. Modelling and predicting fungal distribution patterns using herbarium data. Journal of Biogeography **35**:2298-2310.

Tables

Table 3. Sum of squares, F-ratio, P-value and variance components for distances to pond, building and river of the radiotracking.

	Sum of Squares (SS)	$F_{1,4}$	P	Variance components (VC)
Pond	802157.380	28.826	0.000	67.8 %
Building	171446.374	14.297	0.000	51.0 %
River	101225.374	9.415	0.000	40.6 %

Table 4. Mean date and standard deviation (in days, hours: minutes: seconds) for each phenology trait.

Species	Sex	Beak Color/ Incubation Patch	Mean	SD	N
House Sparrow	M	Black	17.03.11	7 13:00:51.820	74
	M	Pale	-	-	-
	F	Present	27.03.11	1 03:42:46.127	12
	F	Absent	14.03.11	6 12:54:00.781	41
Spanish Sparrow	M	Black	25.03.11	4 04:41:53.598	14
	M	Pale	12.03.11	6 09:16:54.528	21
	F	Present	-	-	-
	F	Absent	18.03.11	8 06:15:10.855	31

Table 5. TSS, AUC, and AUC standard deviation for the Iberian Peninsula- and Europe model replicates.

Replicate	Iberian Peninsula		Europe	
	TSS	*AUC*	*TSS*	*AUC*
1	0.549	0.817	0.537	0.819
2	0.566	0.835	0.566	0.838
3	0.547	0.819	0.528	0.817
4	0.563	0.833	0.659	0.819
5	0.567	0.824	0.581	0.836
6	0.516	0.816	0.537	0.822
7	0.533	0.825	0.516	0.804
8	0.530	0.812	0.560	0.833
9	0.579	0.831	0.521	0.814
10	0.558	0.832	0.523	0.810
11	0.522	0.811	0.540	0.816
12	0.564	0.821	0.539	0.818
13	0.545	0.817	0.539	0.824
14	0.545	0.820	0.539	0.835
15	0.540	0.820	0.573	0.828
Average	0.548	0.822	0.550	0.822

Table 6. Overview of the relative contributions of the variables to the MaxEnt model for the Iberian Peninsula.

Variable	Percent Contribution
BIO17	63.6
Lc22	6.2
BIO15	5.3
BIO9	3.8
BIO3	3.4
Lc12	2.4
BIO8	2
Lc17	1.9
Lc20	1.8
Lc25	1.6
Lc24	1.6
BIO4	1.3
Lc26	1.3
BIO16	1.3
Lc23	0.6
Lc29	0.6
Lc13	0.5
Lc28	0.4
Lc21	0.3

Table 7. Overview of the relative contributions of the variables to the MaxEnt model for Europe.

| Variable	Percent Contribution
BIO17	70.5
BIO9	10.8
BIO15	6.1
BIO8	4.3
BIO3	3.9
BIO4	2.3
BIO16	2.1

Figures
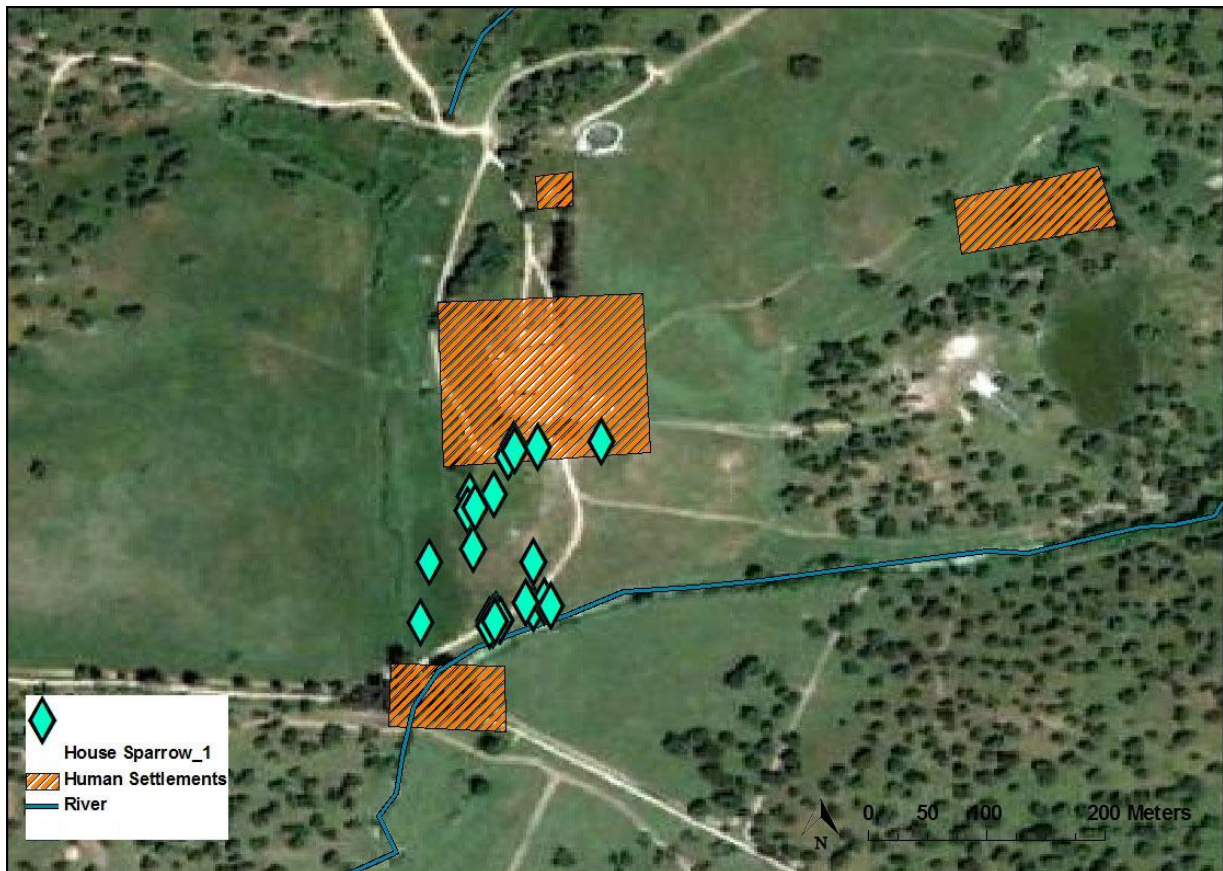
Figure 3. Map of radiotracking of house sparrow individual 1. The blue diamonds represent the tracked locations of the house sparrow; hatched orange areas show where the human constructions (such as pig feeders, farm buildings and stables) were located, and finally; the blue line marks the river.
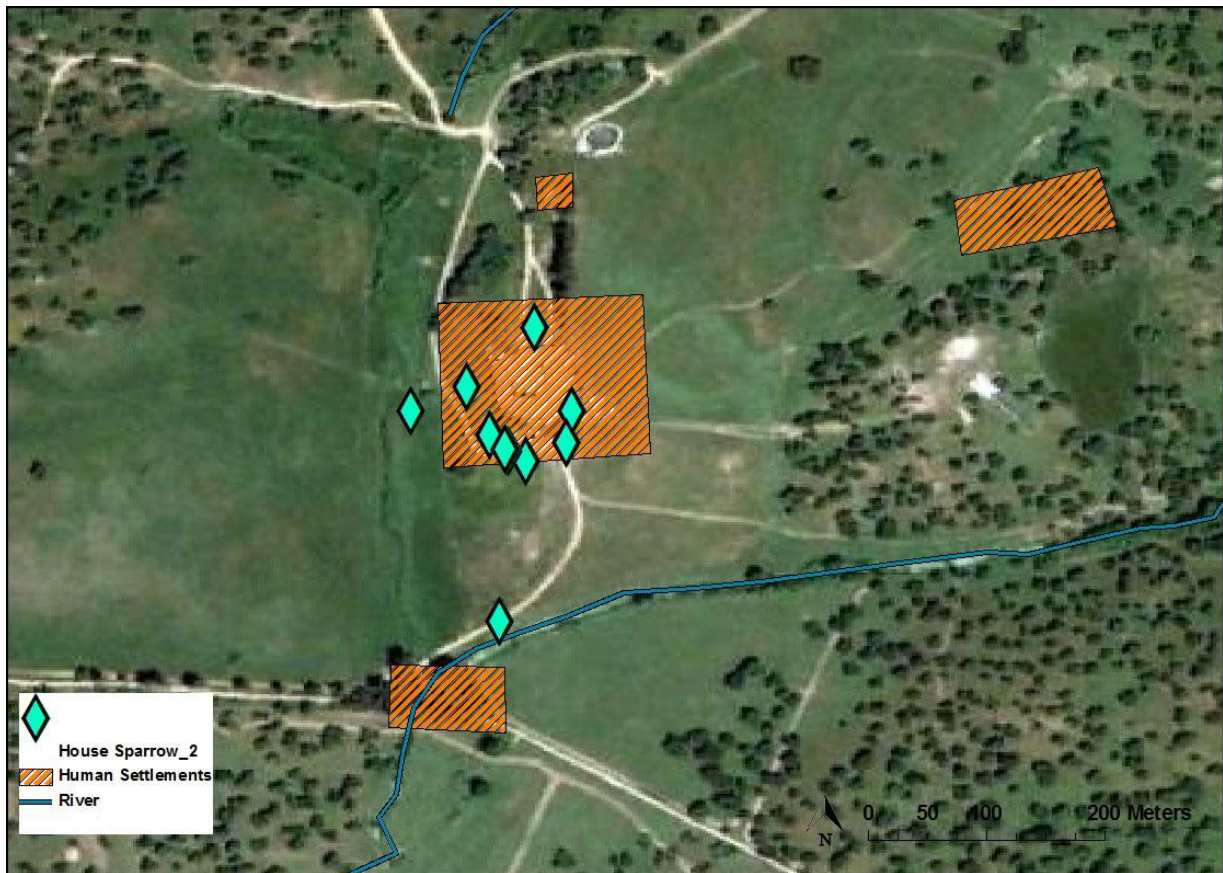
Figure 4. Map of radiotracking of house sparrow individual 2. The blue diamonds represent the tracked locations of the house sparrow; hatched orange areas show where the human constructions (such as pig feeders, farm buildings and stables) were located, and finally; the blue line marks the river.
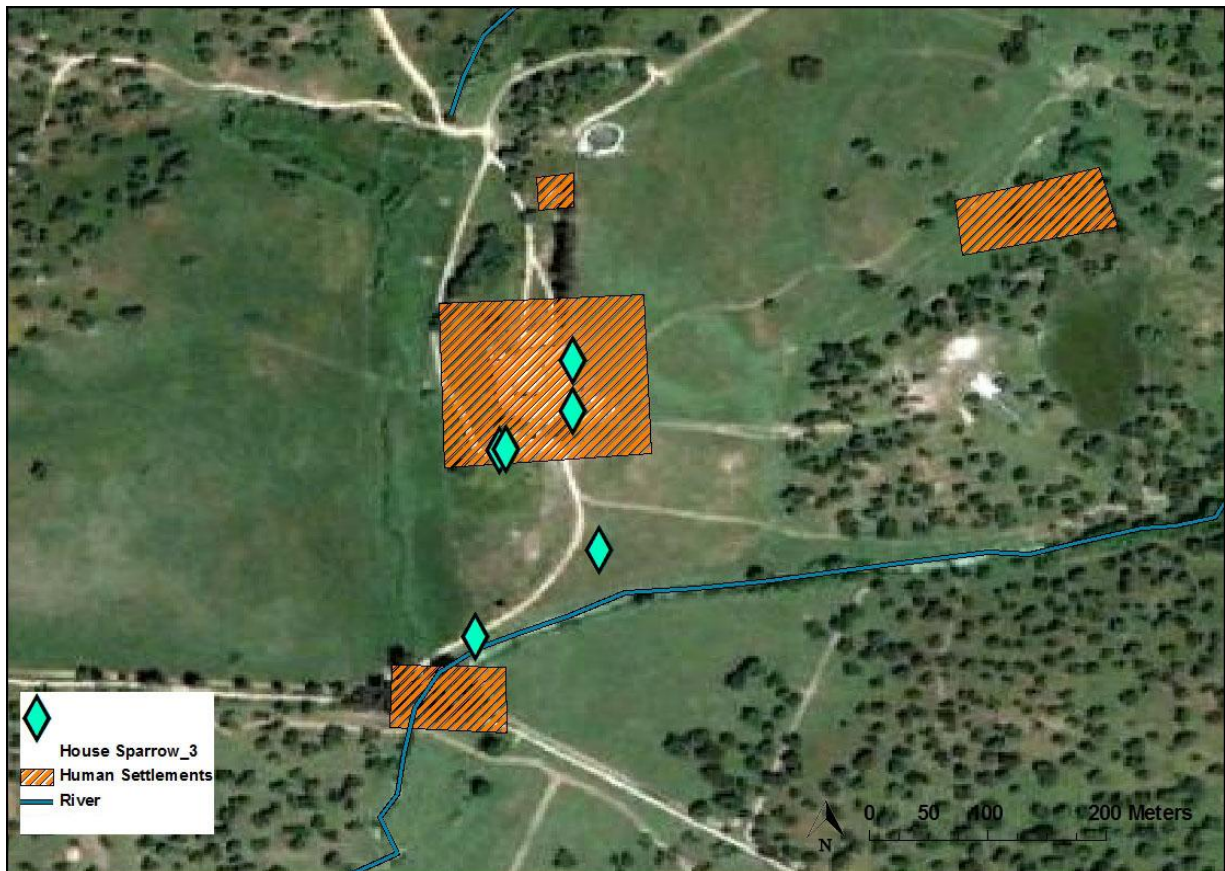
Figure 5. Map of radiotracking of house sparrow individual 3. The blue diamonds represent the tracked locations of the house sparrow; hatched orange areas show where the human constructions (such as pig feeders, farm buildings and stables) were located, and finally; the blue line marks the river.
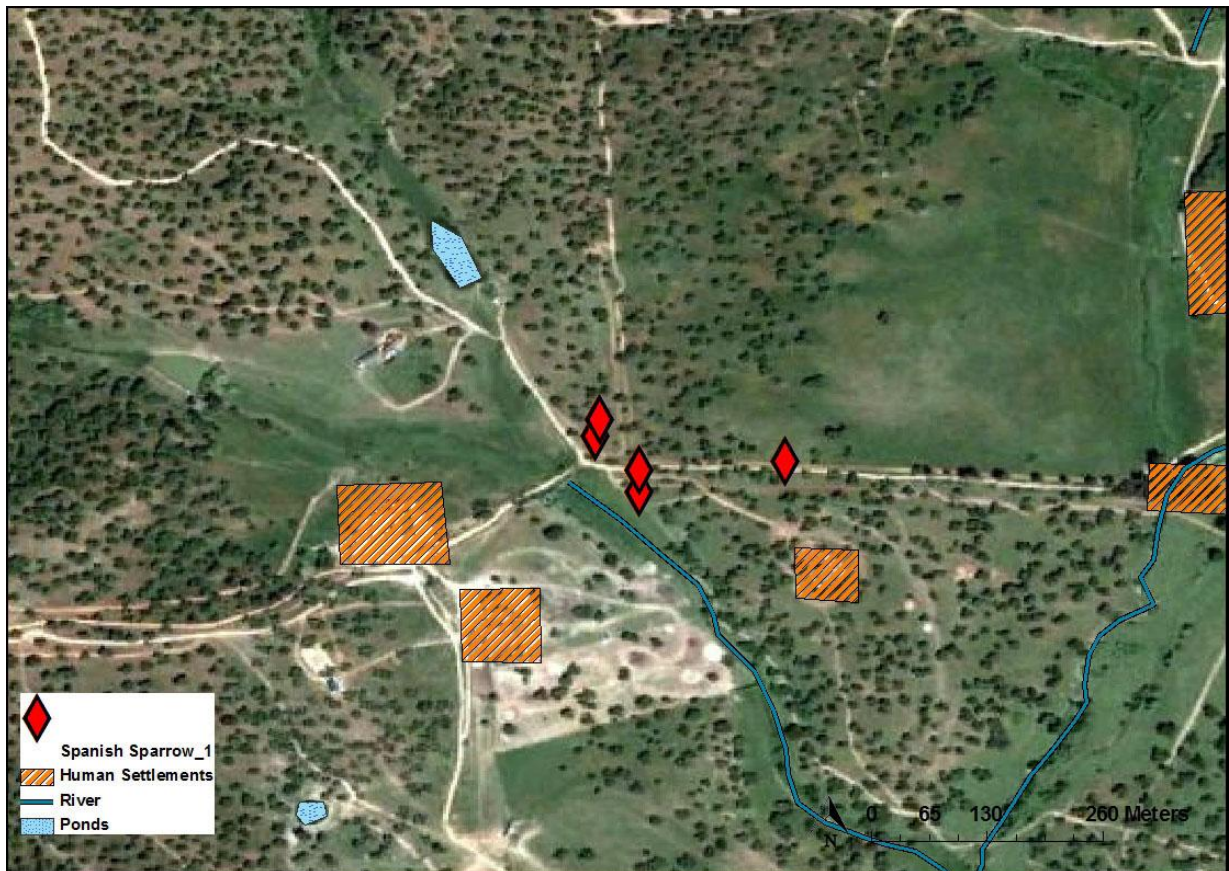
Figure 6. Map of radiotracking of Spanish sparrow individual 1. The red diamonds represent the tracked locations of the Spanish sparrow; hatched orange areas show where the human constructions (such as pig feeders, farm buildings and stables) were located; the blue line marks the river, and finally; the light blue dotted polygons represent the ponds.
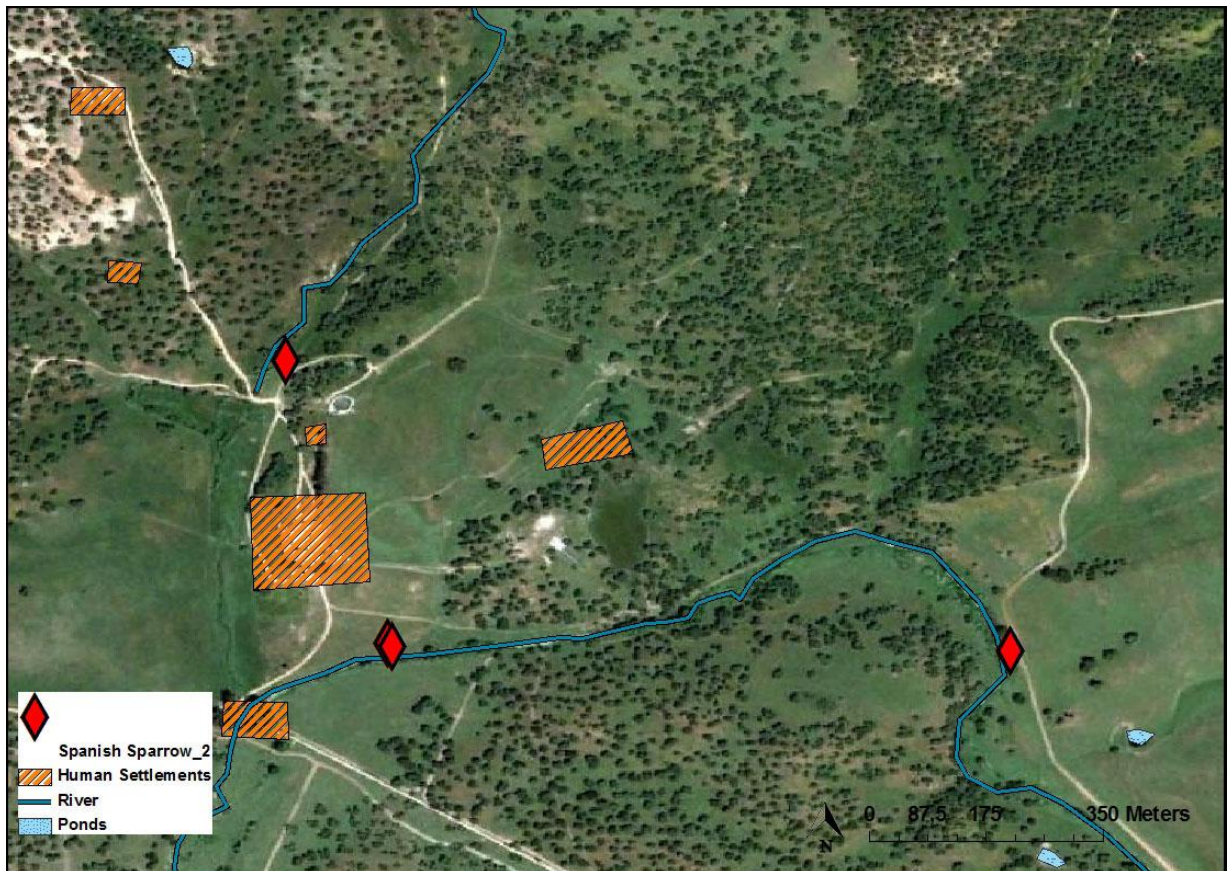
Figure 7. Map of radiotracking of Spanish sparrow individual 2. The red diamonds represent the tracked locations of the Spanish sparrow; hatched orange areas show where the human constructions (such as pig feeders, farm buildings and stables) were located; the blue line marks the river, and finally; the light blue dotted polygons represent the ponds.
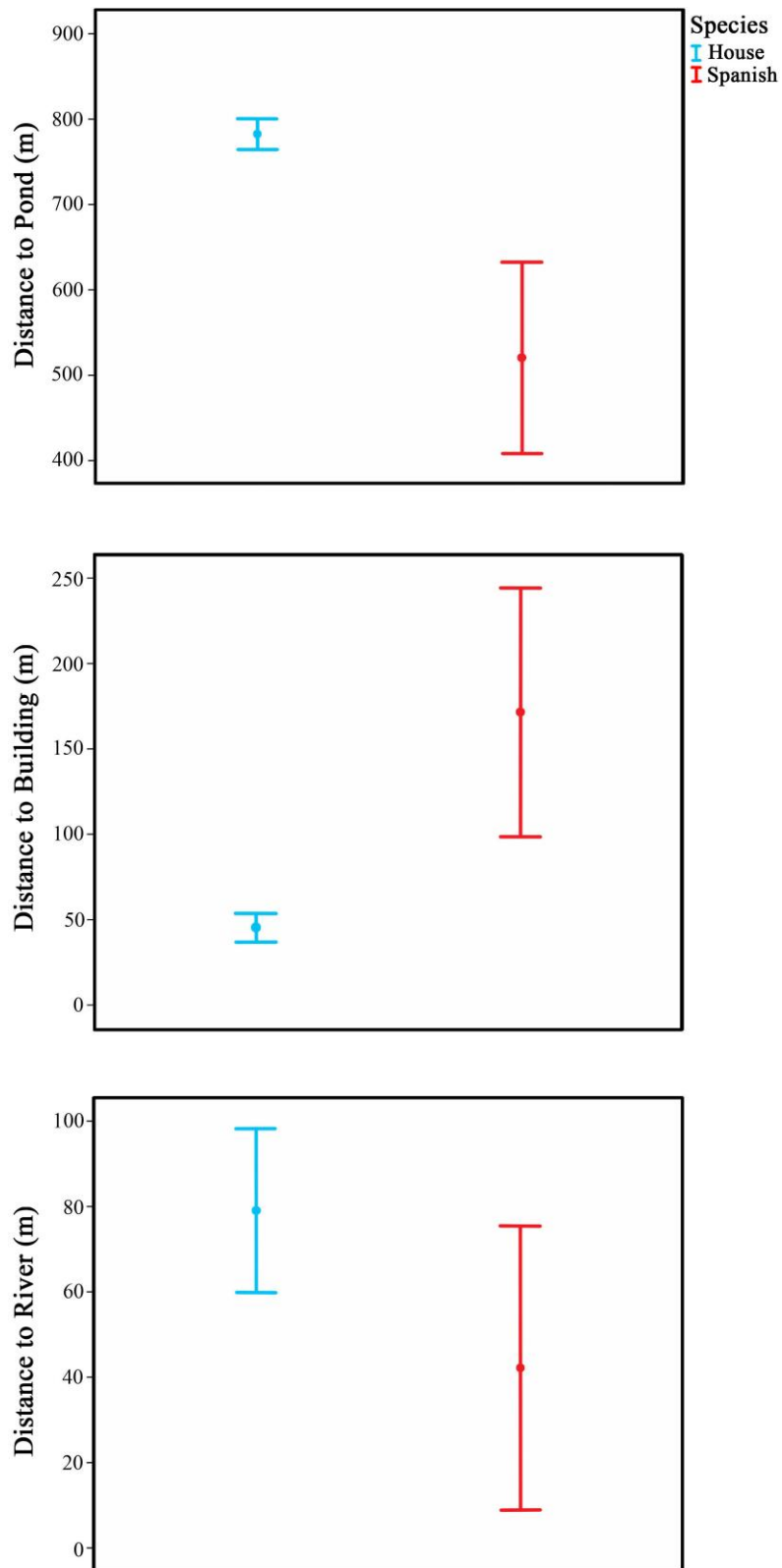
Figure 8. Mean and 95 % confidence intervals of distances (m) to nearest pond, building/human construction and river for radiotracked house sparrow (blue) and Spanish sparrow (red).
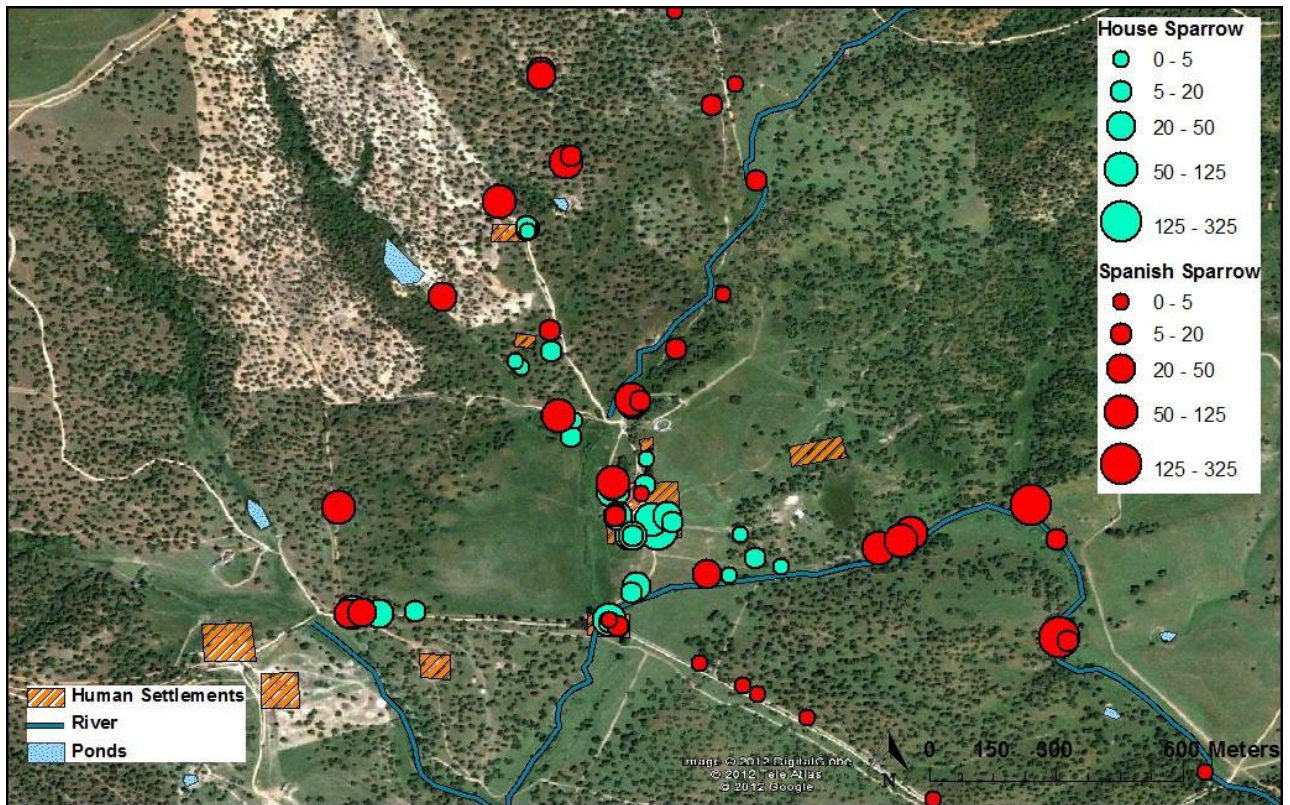
Figure 9. Map of the road transects. The blue circles represent the house sparrow flocks; the red circles represent the Spanish sparrows; hatched orange areas show where the human constructions (such as pig feeders, farm buildings and stables) were located; the blue line marks the river, and finally; the light blue dotted polygons represent the ponds.
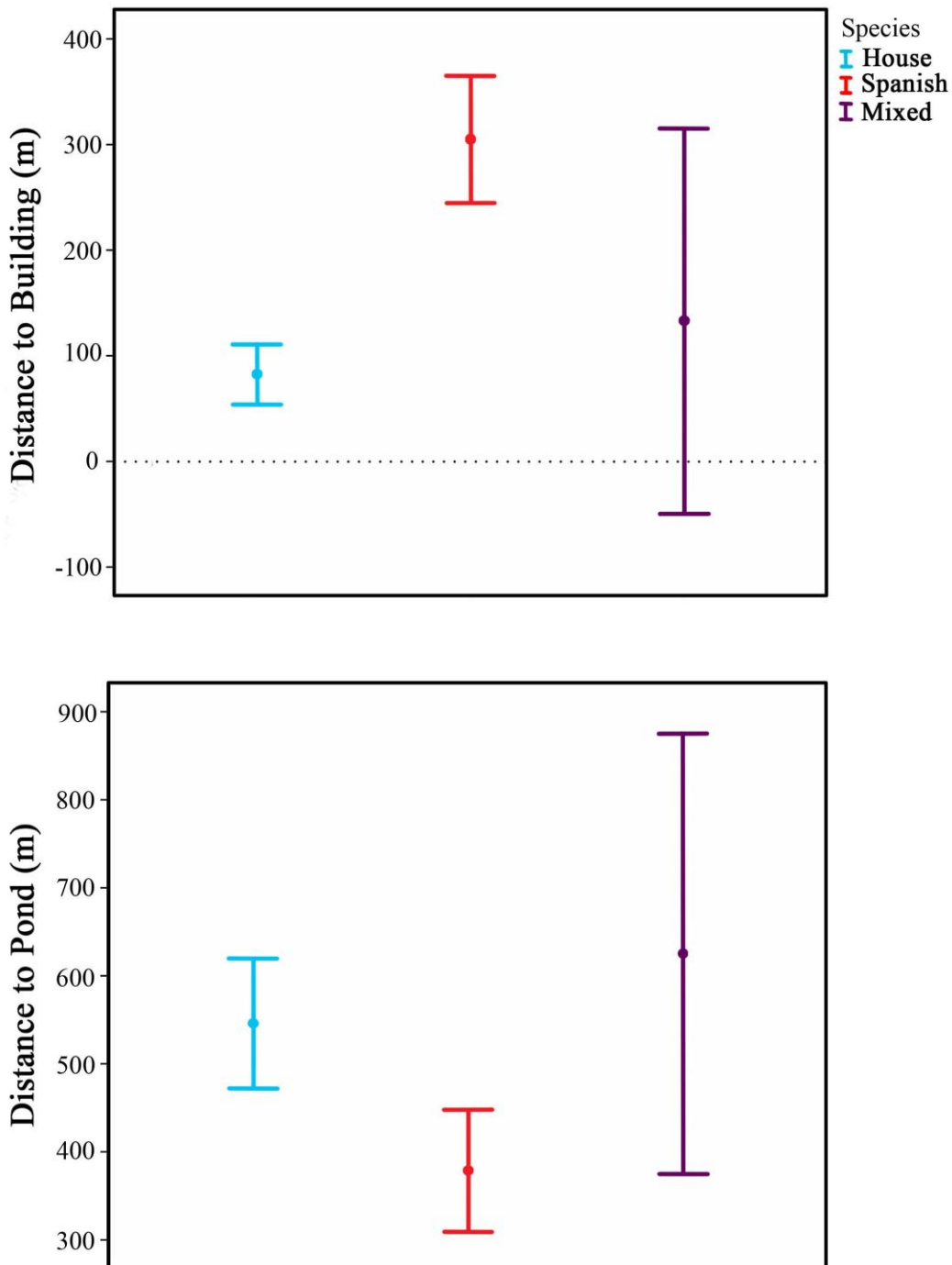
Figure 10. Mean and 95 % confidence intervals of distances (m) to nearest building/human construction and pond for the pure house sparrow (blue), pure Spanish sparrow (red) and mixed flocks (purple). The negative values of the confidence interval of the distance to building in the mixed flocks is an artifact resulting from the Gaussian assumption, and the dotted line indicates the minimum possible distance to any building.
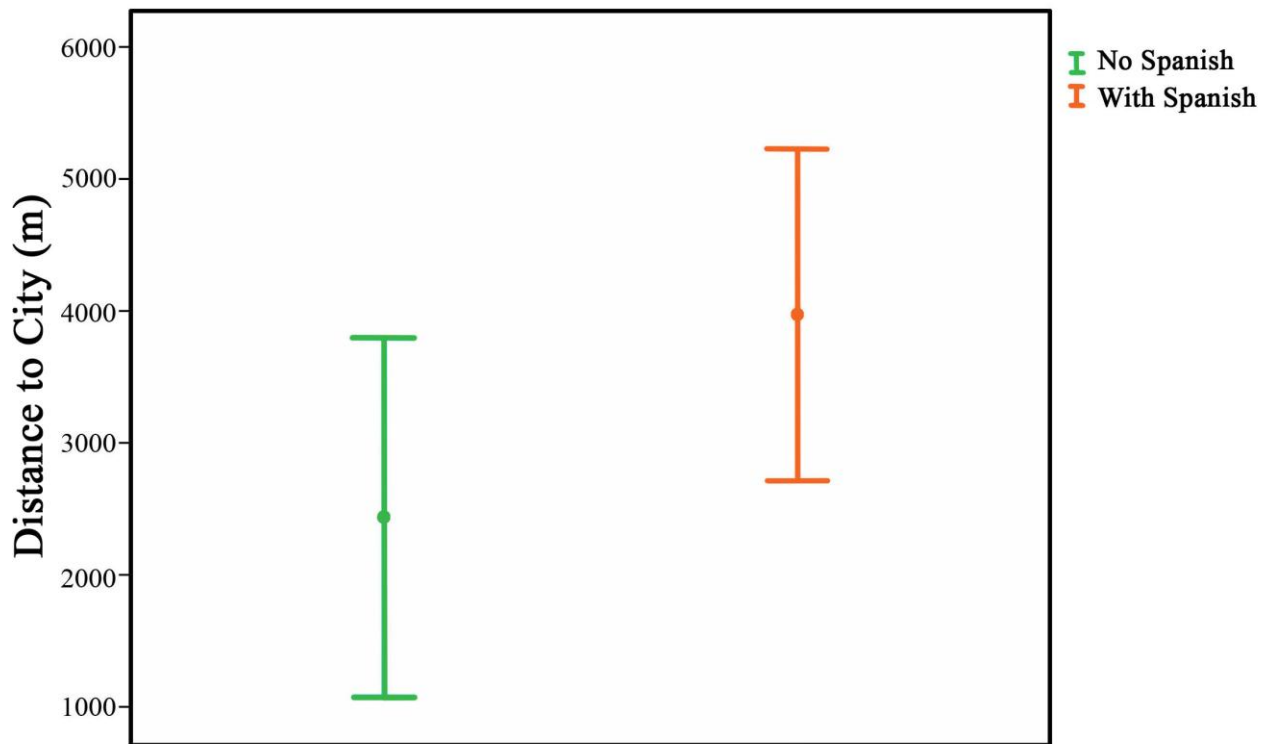
Figure 11. Mean and 95 % confidence intervals of distances (m) to nearest city for colonies with no Spanish sparrows (green) and with Spanish sparrow (Orange).
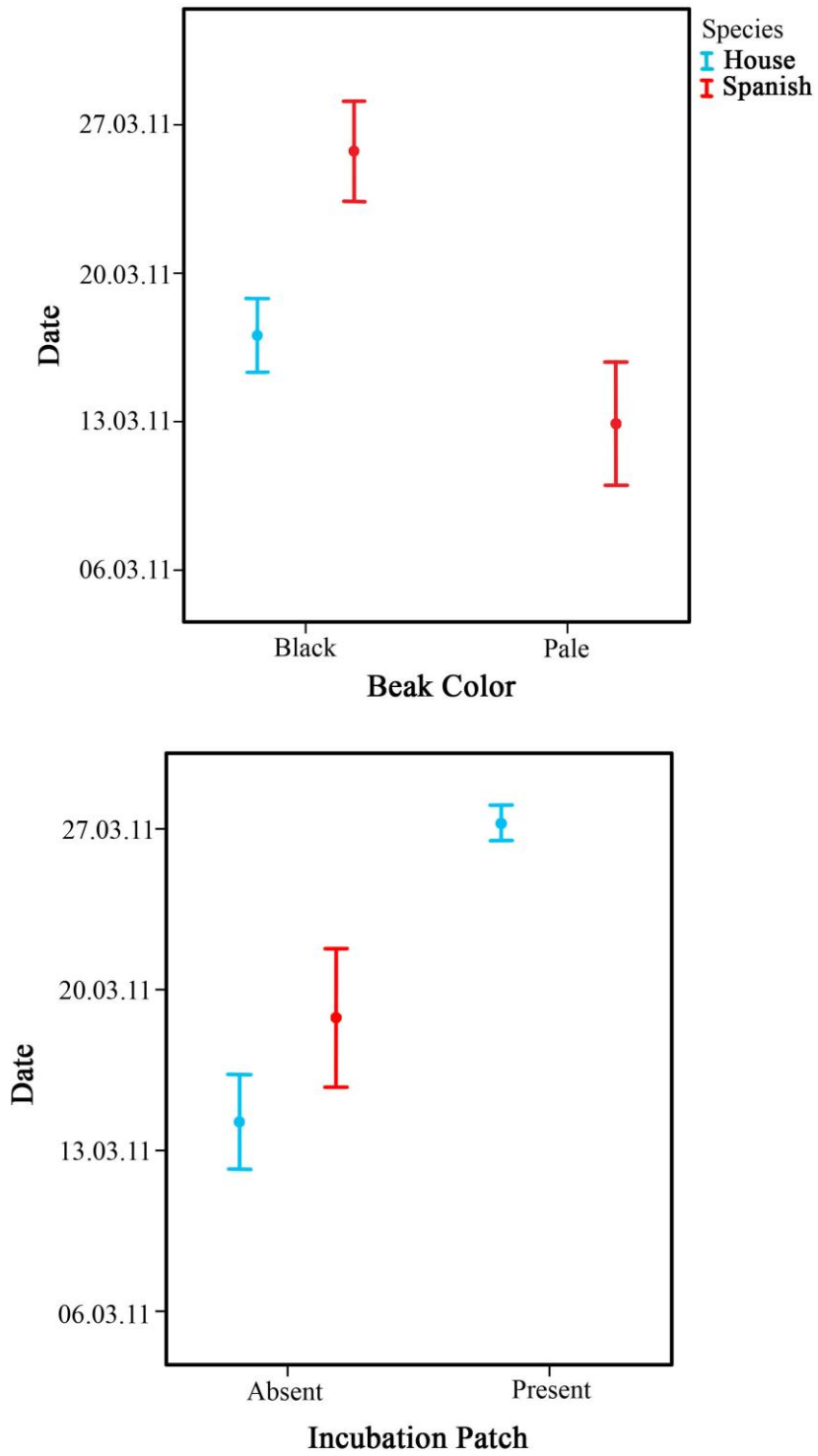
Figure 12. Mean (± SD) dates for beak color in males (top) and incubation patch in females (bottom) for house sparrow (blue) and Spanish sparrow (red).
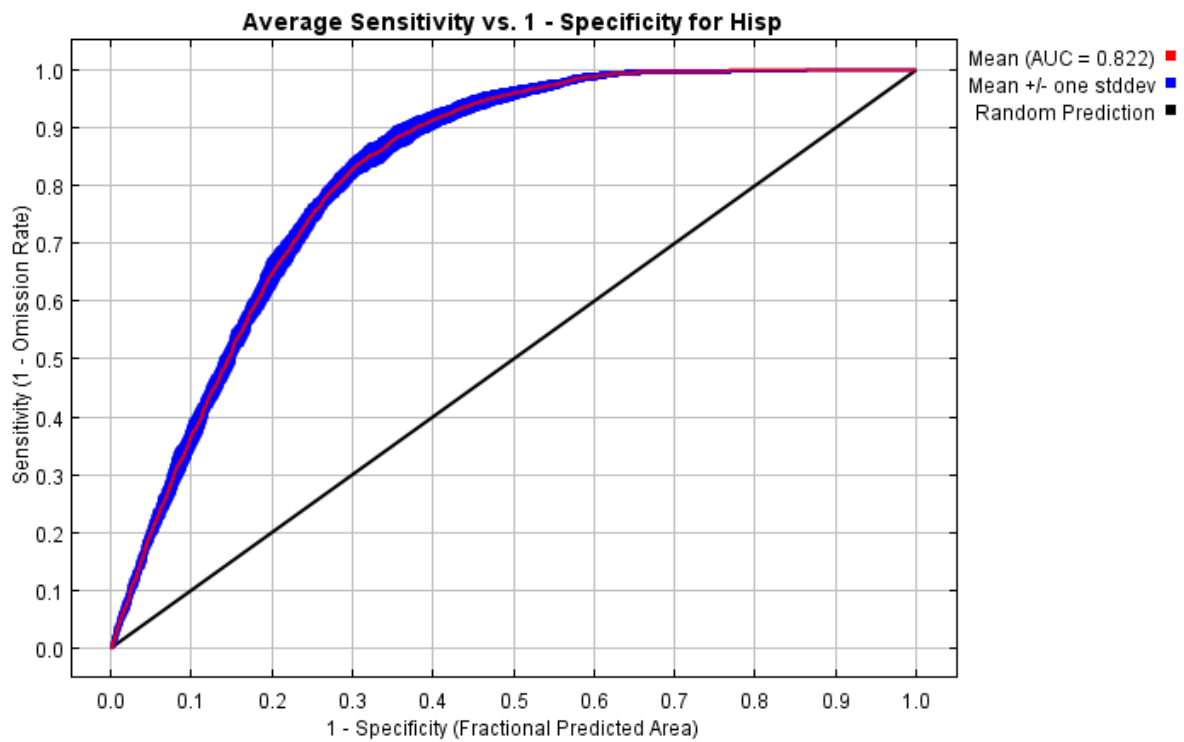
Figure 13. The receiver operating characteristic (ROC) curve for the average test data (Iberian Peninsula model). On the X-axis is the 1 − Specificity (the ability to correctly predict absence) and on the Y-axis is the Sensitivity (the ability to correctly predict presence). The red curve represents the mean AUC (= 0.922), and the blue shows the mean +/- one standard deviation (= 0.007). The black line represents the value of a randomly predicted model.
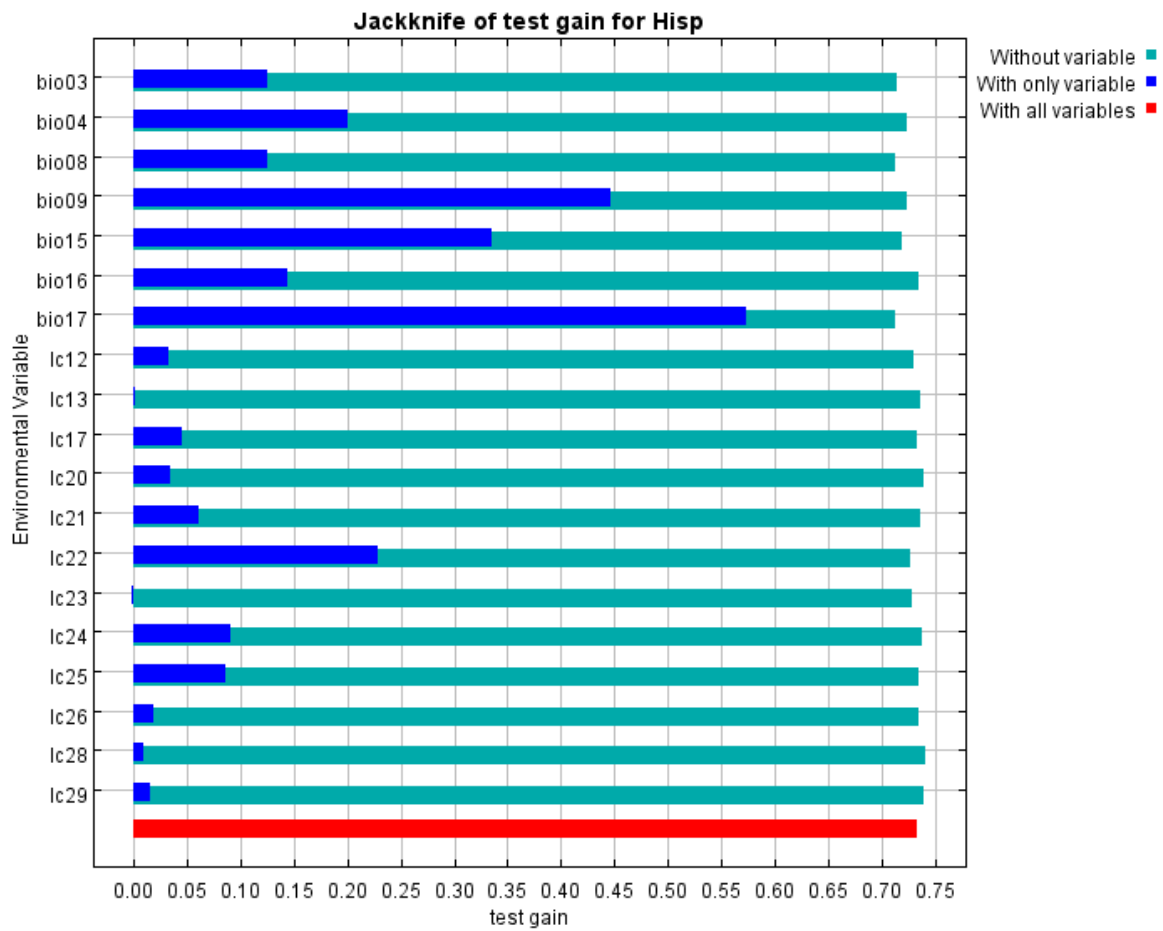
Figure 15. A jackknife test of variable importance by using the test gain (Iberian Peninsula model). The figure shows the relative contribution of each variable to the model. The dark blue bars shows the gain (goodness of fit) when the variable is used in isolation, while the light blue bars how the gain is decreased when the variable is not used in the model. The red bar at the bottom represents the gain of the model.
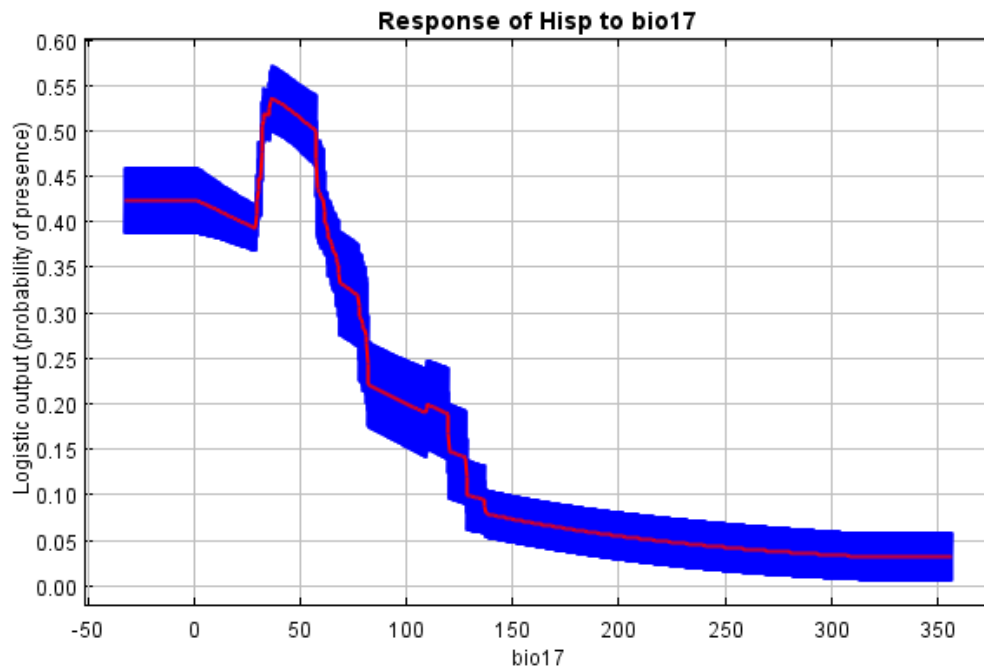
Figure 16. The single variable response of Spanish sparrow to BIO17 = Precipitation of driest quarter (Iberian Peninsula-model). The figure shows how probability of presence (Y-axis) changes as the precipitation is varied (X-axis). The red curve shows the mean response of the 15 replicate runs and blue represents the mean +/- one standard deviation.
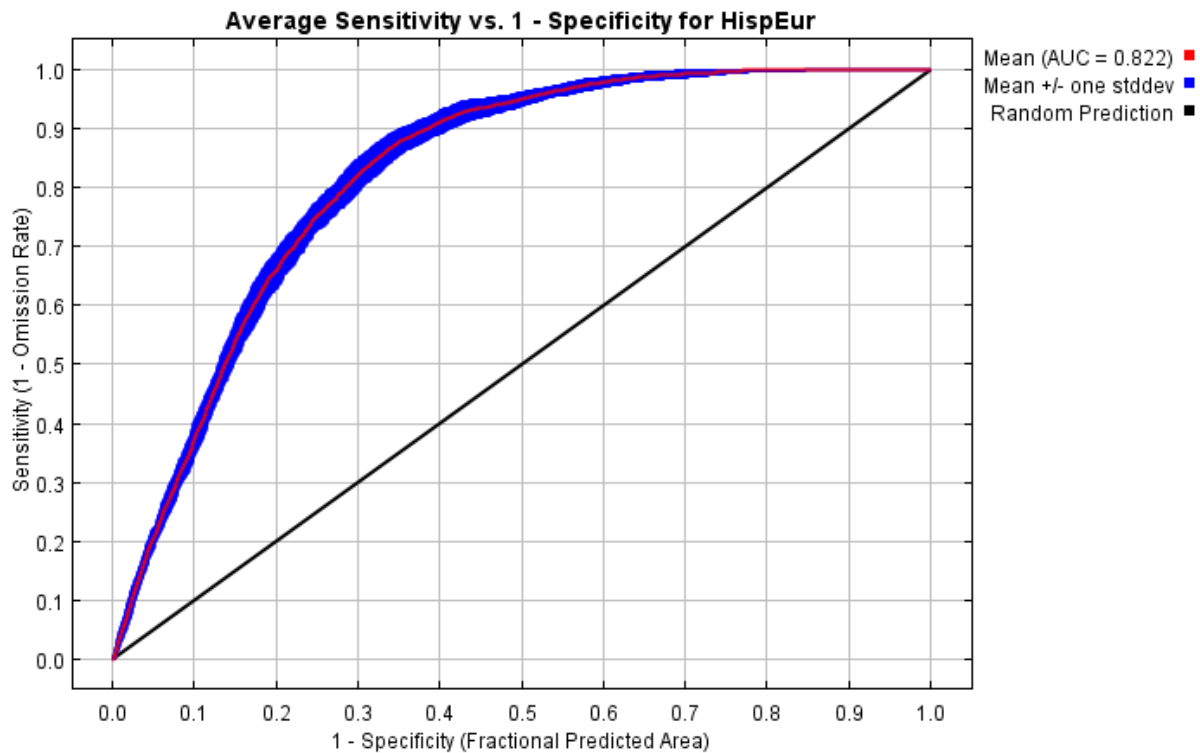
Figure 17. The receiver operating characteristic (ROC) curve for the average test data (Europe-model). On the X-axis is the 1 − Specificity (the ability to correctly predict absence) and on the Y-axis is the Sensitivity (the ability to correctly predict presence). The red curve represents the mean AUC (= 0.822), and the blue shows the mean +/- one standard deviation (= 0.010). The black line represents the value of a randomly predicted model.
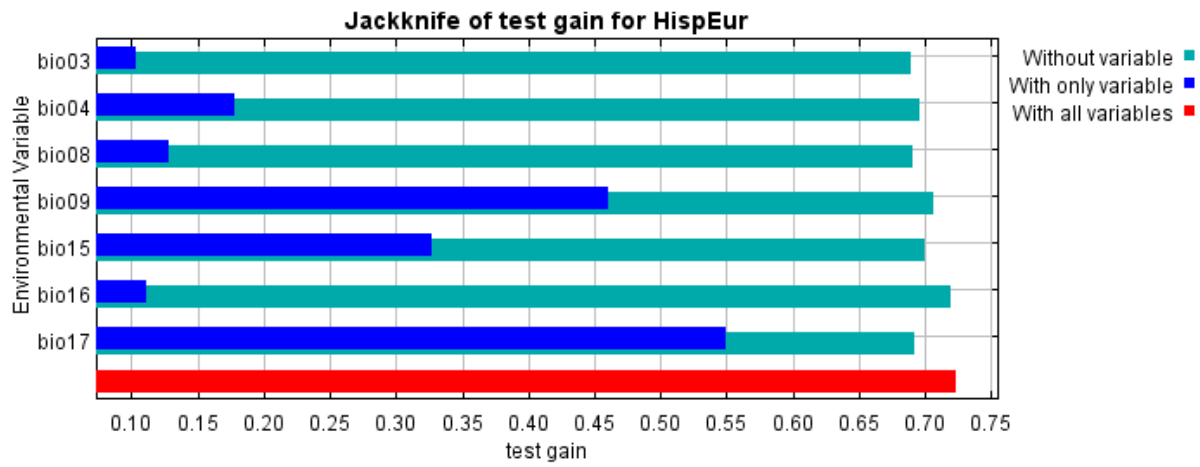
Figure 19. A jackknife test of variable importance by using the test gain (Europe model). The figure shows the relative contribution of each variable to the model. The dark blue bars show the gain (goodness of fit) when the variable is used in isolation, while the light blue bars show how the gain is decreased when the variable is not used in the model. The red bar at the bottom represents the gain of the model.
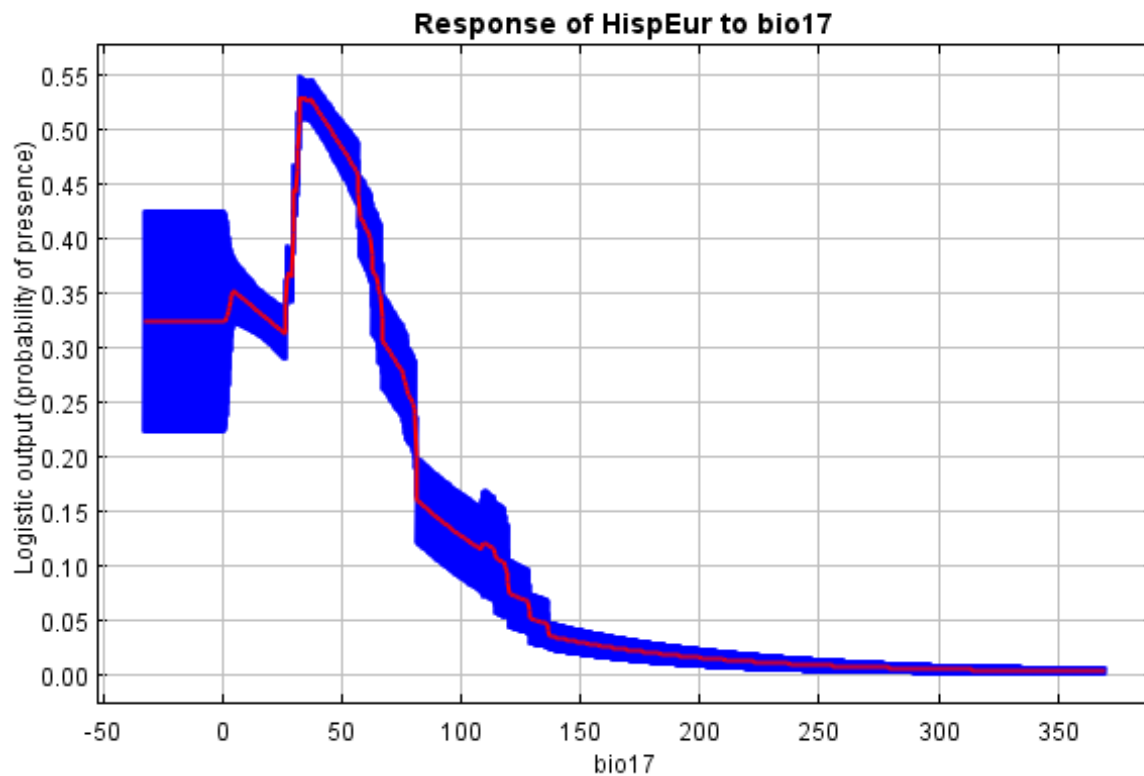
Figure 20. The single variable response of Spanish sparrow to BIO17 = Precipitation of driest quarter (Europe-model). The figure show how probability of presence (Y-axis) changes as the precipitation is varied (X-axis). The red curve show the mean response of the 15 replicate runs and blue represent the mean +/- one standard deviation.