

Actual and future predicted occupancy of the Black Kite in Spain

- Case study of the course Monitoring and occupancy modelling -

Felix Nößler

2nd semester in master program: Ecology, Evolution and Conservation (M.Sc.),

Registration number: 810578

August 15, 2021

Module Quantitative conservation biogeography

Prof. Dr. Damaris Zurell & Dr. Guillermo Fandos-Guzman

ABSTRACT

The Black Kite is an abundant bird of prey in Spain. The main goal is to assess which site covariates are important for the occurrence of the Black Kite in Spain. Furthermore, an actual distribution map of the Black Kite in Spain is presented and compared to the distribution under the assumption of an increase of annual mean temperature and a decrease of total precipitation in 80 years. The selection of site covariates is based on ecological knowledge of the focal species and an exploratory analysis with a Random Forests model. The eBird data set is filtered for repeated surveys and used as the data source for fitting the static occupancy model.

1 Introduction

Climate change affects all levels of biodiversity (Bellard et al. 2012; Garcia et al. 2014). A main task in ecological research is to build accurate models to predict the biological response to climate change (Urban et al. 2016; Araújo and Rahbek 2006).

One of the affected groups by the climate change are birds of prey. Birds of prey play a crucial role in ecosystems. Very often they are top predators. They can be regarded as flagship species, because they are very vulnerable

to human activity, because of their role as top predator in food chains and because of the attractiveness of their behavior to humans. Furthermore, they provide regulating, supporting and cultural ecosystem services (Donázar et al. 2016).

A raptor species with a very wide distribution is the Black Kite, *Milvus migrans* (Boddaert, 1783). The distribution of this species ranges from Western Europe to East Asia and Australia (BirdLife International 2021). The species is mainly migratory and the European

population stays in the winter in sub-Saharan Africa. The individuals leave the breeding area between July and October and come back between February and May (Panuccio et al. 2014; BirdLife International 2021). In this case study the focus lies on the Black Kite population in Spain.

Citizen science data are often used in species distribution modelling, because of the high data availability. However there are some challenges in using these data sets, for example the spatio-temporal bias of the observations. For instance people tend to watch for birds more in the breeding season and around their own home (Robinson et al. 2018; Reich et al. 2018). Two techniques are available to overcome this difficulties. First, it is possible to filter out low quality observations. Second, it is possible to apply statistical techniques to address sampling bias and observational heterogeneity (Steen et al. 2019). In this study the second option is used. Observational heterogeneity and imperfect detection is tackled with an extension of species distribution model, the so-called occupancy model (MacKenzie, Nichols, et al. 2002).

The model is used to learn more about the ecology of the focal species. This knowledge can be used in conservation planning.

The main questions are (1) which site covariates are most important for the occurrence of the Black Kite and (2) what happens to the Black Kite population in Spain under future climate conditions.

2 Methods

Data of the species occurrences are from eBird data (Sullivan et al. 2009; Cornell Lab of Ornithology 2021). The dataset is filtered for repeated surveys, that were done three to ten times in the same area between April and June of 2019. Only standing or travelling surveys with a total distance up to 5 km with one to five observers were used. To get an overview a map with all observations was produced using QGIS Development Team (2021).

Data preparation, selection of covariates, model selection, model evaluation and prediction were done with R Core Team (2021). Exploratory analysis of the covariates were done in Van Rossum and Drake (2009). An ODMAP-Protocol (Zurell et al. 2020) describing typical steps in species distribution modelling can be found in the electronic appendix.

Spatial resolution of the model

The home range size differs between younger not breeding individuals, that are one to seven years old, so-called floaters, (Blas et al. 2009) and the breeding individuals. Floaters had in south Spain an home range size over 300 km², breeding females 43 km² and breeding males 80 km² (Tanferna et al. 2013). However, these home range values were observed with radio-tracking of individuals and later on the calculation of the minimum convex polygon. The main activity is probably more restricted to centre of the area.

A resolution of 2.5 minutes (roughly 21.5 km²) was chosen. It is possible to argue that

the size is too small for the large home range sizes of the Black Kite. However through spatial subsampling the data loss is higher and the connection between the observations and the site variables becomes less strong.

Selection of detection covariates

Detection covariates that are present in the eBird data are used. These are the day of the year, time when observation started, duration of observation, walking distance and protocol type (standing or travelling). The two most important covariates from the Random Forests model (see section on exploratory analysis of covariates) were used in the final model.

Selection of site covariates

First a preliminary selection of site covariates was compiled. These set include all bioclimatic variables (Fick and Hijmans 2017), land cover fractions, as the tree cover, the grass cover and the bare soil cover (Buchhorn, Smets, et al. 2020), and the distance between the centre of the raster cells to the closest land-fill and the closest river or lake.

In the next step ecological hypotheses were formulated based on existing ecological knowledge from the literature. Some of the covariates were selected based on these hypotheses. As an additionally step, all preliminary covariates were used to fit a Random Forests model and to get the importance of all the covariates for predicting the detection probability of the Black Kite. The most important covariates from these analysis were added to the list

of covariates.

The covariates from the ecological hypotheses and from the Random Forests model were tested for collinearity with the Variance Inflation Factor (as implemented in Heiberger 2020). Highly collinear covariates were removed. Maps of the final site covariates can be found in the appendix (see Figure 7). The complete selection process is shown in Figure 1 and further explained in the next sections.

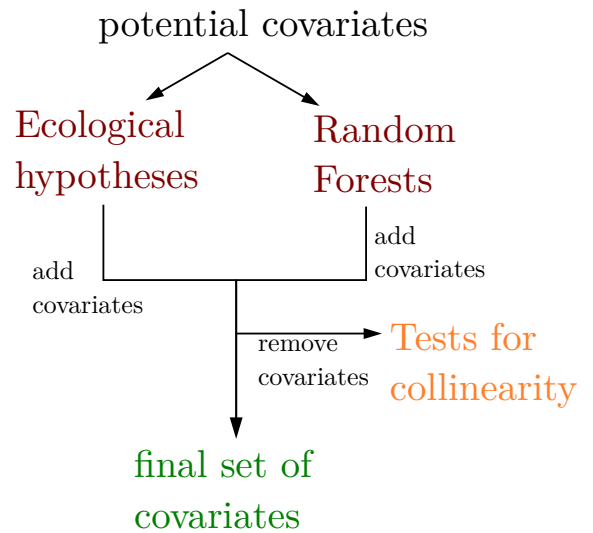


Figure 1: Procedure of selecting site covariates for the occupancy model

Data preparation of site covariates

For all raster operations Hijmans (2021) was used. Bioclimatic variables were download from Fick and Hijmans (2017) in the target resolution. Tree cover, herbaceous vegetation cover and bare soil cover were downloaded (Buchhorn, Lesiv, et al. 2020; Buchhorn, Smets, et al. 2020). The raster images were merged to one file and then cropped to mainland of Spain. The resolution was resam-

pled to the target resolution using a bilinear interpolation.

Land cover data was retrieved in vector format (Copernicus Land Monitoring Service 2018) and cropped to the mainland of Spain. The distance between each centre of the raster cell to all polygons of landfills and rivers or lakes was calculated with Bivand and Rundel (2020). Afterwards, the minimum distance was saved.

Ecological justification of site covariates

The selection of covariates is based on existing ecological knowledge about the focal species. A list of all site covariates and a corresponding simplified hypothesis can be found in Table 1.

The Black Kite uses a variety of feeding sources, for example birds, fish, crayfish, insects, carrion, vegetable matter and smaller mammals (Sergio and Boto 1999; Vinuela and Veiga 1992; BirdLife International 2021). Part of the prey is found in wetlands and marshes and it was proposed that the Black Kite has a habitat binding to wetlands and marshes (Veiga and Hiraldo 1990; Tanferna et al. 2013). Accordingly, the minimal distance to lakes and river was calculated and used as a site covariate. The distance to the closed lake or river was used, because it is more important that lakes or rivers are nearby than that the Black Kite was actually observed in cell with a lake or a river.

It has been shown that Black Kites visit landfills for feeding (De Giacomo and Guerrieri 2008; Blanco 1994). Therefore the distance

to the closest landfill were analysed. The landfill cover was not used, because in most of the raster cells landfills are not present and therefore the model fitting will not be optimal.

Black Kites breed in branches of trees. However, closed woodlands are avoided (Tanferna et al. 2013). Therefore the hypothesis is that the occupancy of the Black Kite follows a unimodal distribution with regard to the tree cover, where the occupancy is high at low to intermediate tree cover. In contrast, there should be a positive relationship with the grass cover, because the Black Kite is searching for prey in open landscapes.

Table 1: Site covariates and simplified hypothesized response of occupancy of the Black Kite, arrows symbolize the expected occupancy probability with higher values of the respective covariate

Site covariates
Annual Mean Temperature ↘
Annual precipitation ↗
Tree cover ↘
Grass cover ↗
Bare soil cover ↘
Distance to closest river or lake ↘
Distance to closest landfill ↘

Exploratory analysis of site and detection covariates

An exploratory analysis of the importance of the site and detection covariates was done using the machine learning method Random Forests (Ho 1995). All covariates as described

before were used. These include all bioclim variables, land cover fractions, cover of water bodies, cover of landfills, distance to closest water body, distance to closest landfill and all detection covariates. From each grid cell one observation was selected and related to the site and detection covariates. Because of the higher number of non-detection an balanced Random Forest Classifier was used (Lemaître et al. 2017, with default parameters). Preparing the data, standardization of data and splitting in train and test data (80 % and 20 %) was done with Pedregosa et al. (2011), Harris et al. (2020), and McKinney (2010). 500 simulation were done to account variability in the model fit to the different data sets (each time one of the up to ten observations per grill cell was chosen). The mean and the standard error of the importance of all covariates was calculated. The effect of the most important covariates on the detection probability was tested. Therefore all other covariates were set to the mean and only the focal covariate was changed. The python script for conducting the analysis can be found in the appendix.

Selection of best model

A nullmodel, a model with only detection covariates, a model with only site covariates and a full model (site and detection covariates) were compared with the Akaike information criterion (AIC) and the Akaike information criterion corrected for small sample size (AICc) with Fiske and Chandler (2011). The model with the lowest AIC and AICc was cho-

sen.

To compare the stability of the model an average model was build from the best model. Subsets of the best model were chosen with Bartoń (2020), all detection covariates were used as fixed terms (present in all subsetted models). Models were sorted according to the lowest AIC and selected with cumulative Akaike weight (Wagenmakers and Farrell 2004) larger than 95 %. The selected models are weighted with the AICs and averaged with Bartoń (2020).

Model evaluation

The predictive power of the best model is evaluated with the R^2 -value (Nagelkerke 1991). A parametric bootstrapping approach is used to access the goodness of fit ("parboot" function as implemented in Fiske and Chandler 2011). From the test results the \hat{c} -value was calculated. If this values is noticeably larger than one, it indicates overdispersion of the model. With this \hat{c} -value a quasi-Akaike information criterion corrected for small sample size (qAICc) was calculated to see if there are big differences to the calculated AICc. Another goodness of fit as proposed by MacKenzie and Bailey (2004) was carried out (implemented in Mazerolle 2020).

Prediction

Predictions were done with the best model for the mainland of Spain. A map was produced with the predicted occupancy and the standard error of the prediction.

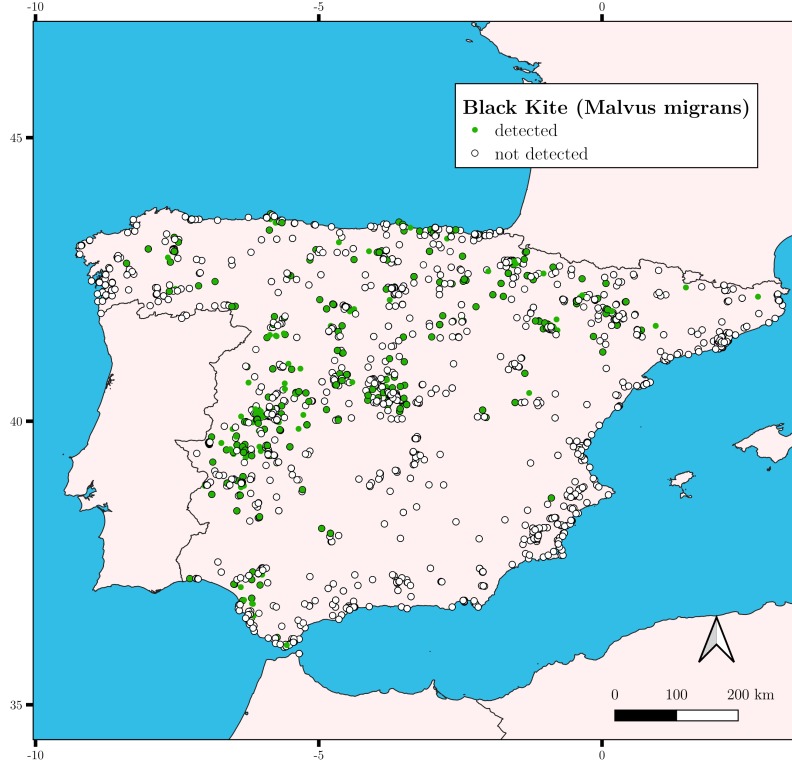


Figure 2: Detection of the Black Kite in the filtered eBird data set in Spain, each dot represents one observation within the repeated surveys, created with QGIS Development Team (2021)

Prediction of future climate conditions were done accordingly to existing knowledge about the climate change. The annual temperature will likely rise around 3 °C and the annual rainfall will decrease around 10 % from 2020 to 2100 (State Meteorological Agency (AEMET) 2021). Iturbide et al. (2021) show under the SSP2 4.5 scenario (Riahi et al. 2017) an increase between 2 and 3 °C and a decrease of total precipitation between 5 % and 20 % in different areas of Spain (mean values of 32 models, 2081-2100, relative to 1995-2014). The model agreement is high for the temperature increase, however the signal is not robust for the decrease in total precipitation. These data set is used to model the future occupancy of

the Black Kite. The interactive atlas provides regional predictions on a spatial scale of 95 km edge length of a raster cell. GeoTIFFs were downloaded for the change in annual mean temperature and the annual total precipitation (change from 1995-2014 to 2081-2100). These raster files were cropped to the mainland of Spain and resampled to the target resolution (2.5 minutes of a degree). Maps of the changes in the target resolution can be seen in the appendix (see Figure 8). The change in temperature and precipitation were applied to the respective bioclimatic variables (Fick and Hijmans 2017).

We will only incorporate the habitat suitability in relation to climate factors and omit

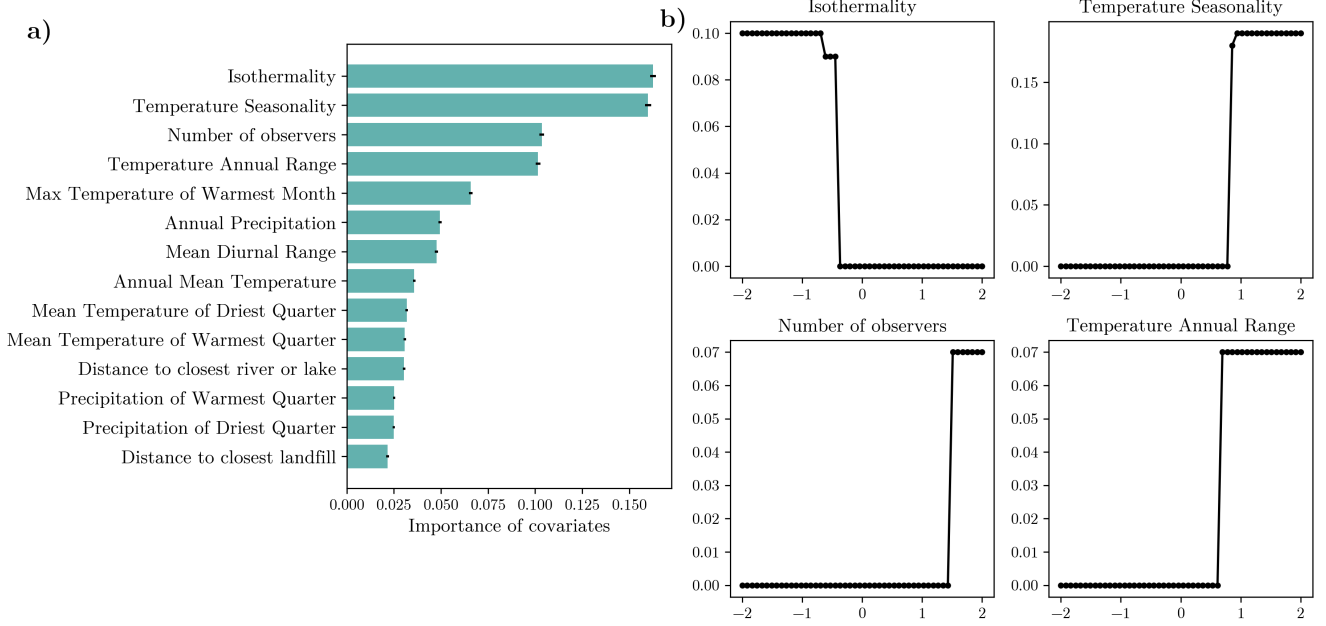


Figure 3: Relationship between covariates and detection probability, (a) importance of the covariates in the Random Forests model (only covariates with importance > 0.02 are included), mean importance values of 500 simulation are shown, black lines represent the standard error, and (b) influence of the four most important covariates on the detection probability of the Black Kite in Spain, the y-axis shows the probability, the x-axis the standardized value of the covariate, created with Hunter (2007)

all biological mechanisms that are proposed to play a role for prediction by Urban et al. (2016) like demography or evolution of a population. It is therefore a simplistic model, but the goal is to catch the main trend.

The effect of the site and detection covariates were tested. Accordingly, all other covariate than the target covariate were set to the mean value and the predictions were done in the range of the target covariate that the model was fitted to. Occupancy probabilities were produced for this range the uncertainty was measured with the standard error.

3 Results

A map with all filtered observation from the eBird data set is shown in Figure 2. It can be observed that the main distribution is in central Spain. The Black Kite was detected from the coast of the Gulf of Cadiz northwards up to the western Pyrenees. However, in the south-eastern part of Spain detections of the Black Kites are rare.

Exploratory analysis of covariates

The Random Forests model could explain close to 100 % of the detection of the Black Kite in the train and the test data set. An overview of the most important covariates can be found

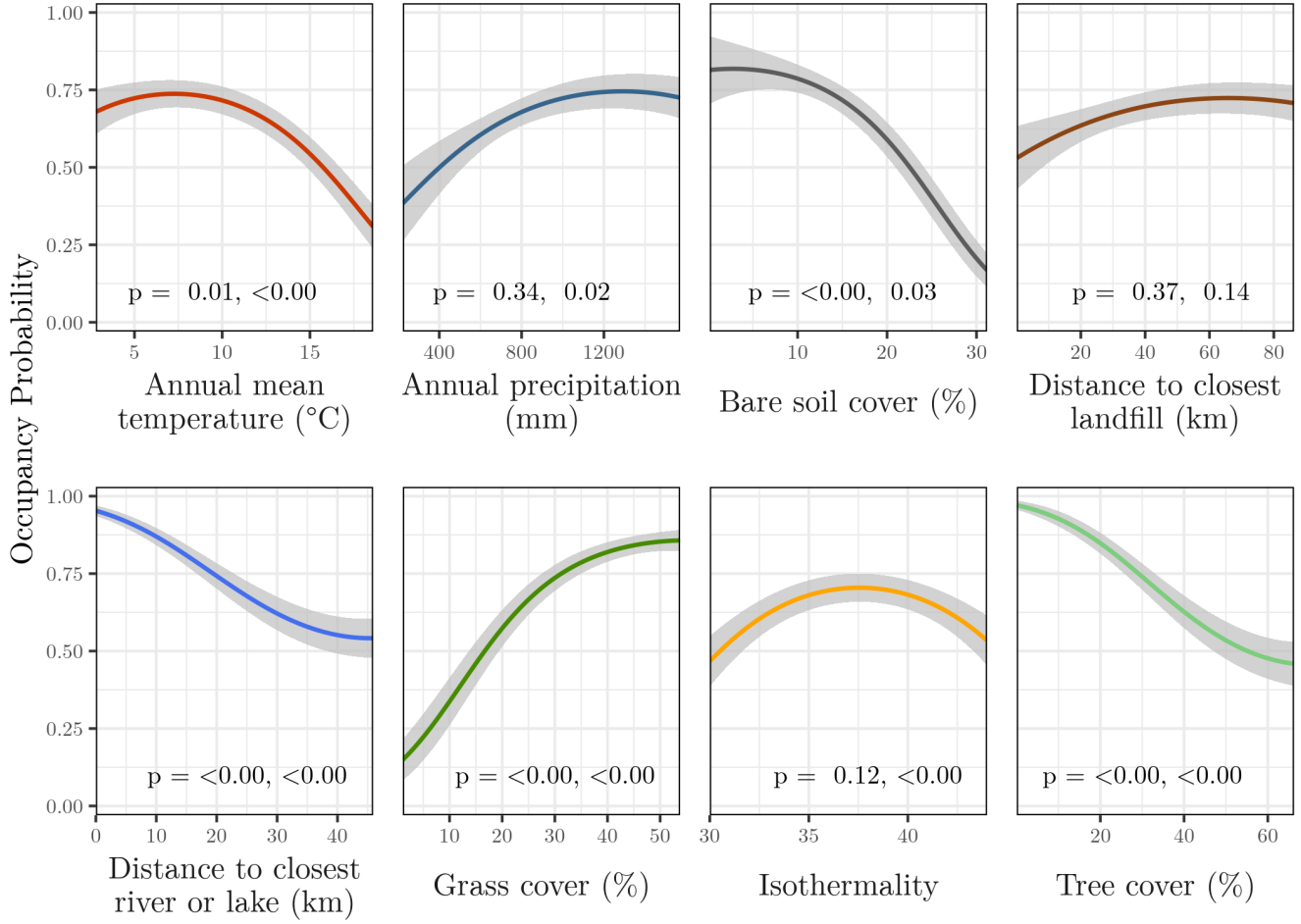


Figure 4: Predicted occupancy of the best model in response to site covariates, for each plot only the focal site covariate was varied, grey areas represent the standard error, the p values show the significance of the first and the second coefficient, created with Wickham (2016)

in Figure 3a.

Especially important are covariates that deal with temperature evenness over the year like isothermality, temperature seasonality and temperature annual range. The only detection covariate that really plays a role is the number of observers. More observers improve the chance to detect a Black Kite (see Figure 3b). Land cover fractions are not represented under the most important covariates. The annual precipitation is sixth influential covariate.

The distance to the closest river or lake and the distance to the closest landfill play minor role.

The four most important factor, namely the isothermality, temperature seasonality, temperature annual range and maximal temperature of warmest month were added as preliminary covariates.

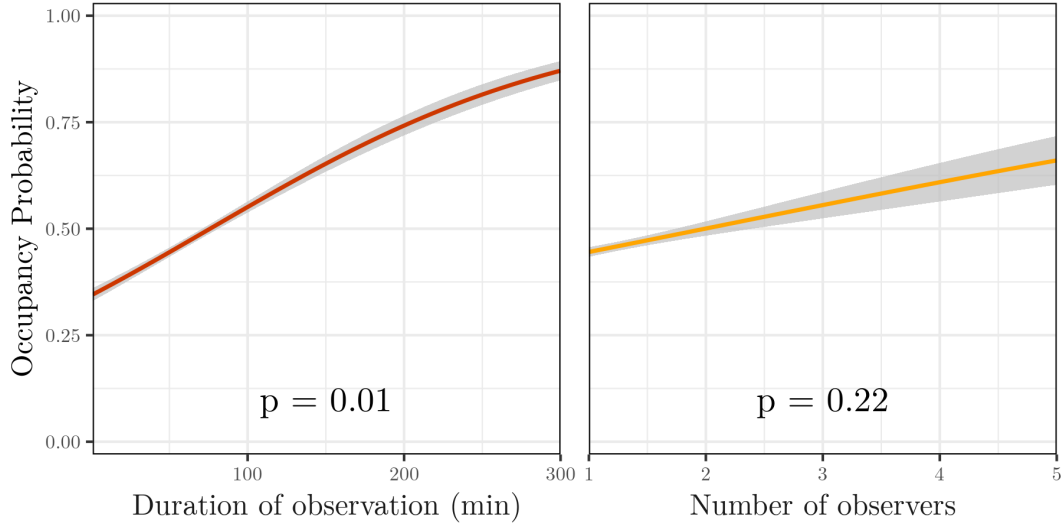


Figure 5: Predicted occupancy in response to detection covariates, for each plot only the focal detection covariate was varied, grey areas represent the standard error, the p values show the significance of the coefficient, created with Wickham (2016)

Occupancy model

Seven site covariates were analysed from the ecological hypotheses and four were added from the Random Forests model. Because of the high collinearity temperature seasonality, temperature annual range and maximal temperature of warmest month were excluded from the final list of covariates. After that the variance inflation factor was for all covariates below 1.9. The highest correlation exists between the tree cover and the annual precipitation (Pearson $\rho = 0.62$).

The final best model consists of 20 parameters and has an AIC of 4578. The second best model with only the site covariates has an AIC that is 143 higher (delta AIC). The AICc gives roughly the same result. In the best model there are two detection covariates, namely the number of observers and the duration of observation. All site covariates are

fitted as polynomials of two degrees. These are the annual mean temperature, the isothermality, the annual precipitation, the bare soil cover, the grass cover, the tree cover, the distance to closest river or lake and the distance to closest landfill. All estimated values of the site covariates are significant different from zero, except the first coefficients of the isothermality and the annual precipitation and both coefficients of the distance to the closest landfill (see Figure 4). The coefficient of the duration of observation is significant different from zero, however the number of observers not (see Figure 5).

The occupancy probability shows a positive relation with the grass cover and the annual precipitation. A unimodal relationship can be seen between the isothermality and the annual mean temperature and the occupancy probability of the Black Kite. Bare soil cover, distance to closest lake or river and a high tree

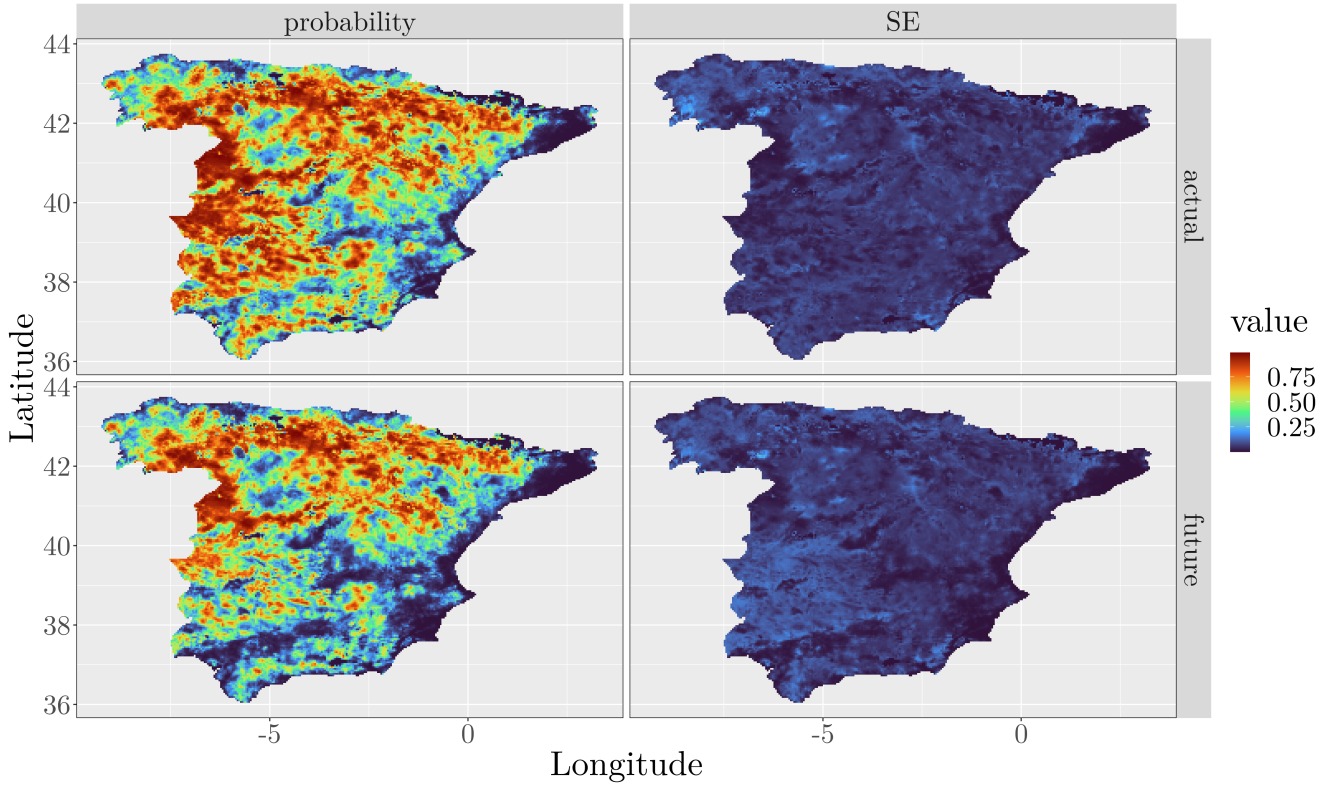


Figure 6: Occupancy map under actual and future climate conditions predicted with the best model, the occupancy probability is shown in the left maps, the corresponding standard error in the plots on the right side, created with Wickham (2016)

cover have a negative influence on the occupancy probability. Both coefficients of the distance to the closest landfill are not significant that is why we cannot see any trend here (see Figure 4).

An average model was build from three single models. These are one model with all site covariates except the distance to closest landfill, a model with all site covariates and a third model with all site covariates except the annual precipitation and the closest distance to landfills. These results go inline with the non-significance of the coefficients of the covariates closest distance to landfills and annual precipitation in the full model.

The explained variance in the detection of the Black Kite is 35 % (R^2 -value).

The model predicts that 522 out of 1032 grid cells with observations are occupied (51 %, median best unbiased predictor from occurrence state). In 503 grid cells was the Black Kite detected (49 %). The MacKenzie and Bailey (2004) goodness-of-fit test shows that the model adequately fits the data ($\hat{c} = 0.83$, $p = 1$). Also the parametric bootstrap test does not indicate a lack of fit ($\hat{c} = 0.97$, $p = 0.86$). The \hat{c} -values are below zero, for this reason no QAIC and QAICc are calculated (they have the same values as the AIC and AICc when setting the \hat{c} -value to one).

The mean occupancy probability under actual climate conditions is 52.6 % (16748 grid cells with occupancy probability ≥ 0.5). The mean occupancy probability decreases and is 37.8 % under future climate conditions (10290 grid cells with occupancy probability ≥ 0.5). This decrease in occupancy is also visible in the map especially between Sevilla and Madrid (see Figure 6). In 92 % of all grid cells there is a lower occupancy probability under future climate conditions, in contrast in only 8 % of all grid cells the occupancy probability is higher under future climate conditions.

The duration of the observation has a positive influence on the occupancy probability. In contrast, the number of observers does not show a significant result (see Figure 5).

4 Discussion

Our results imply that the climate change may negatively affect the occupancy of the Black Kite in Spain.

annual mean temperature, Annual precipitation, bare soil cover, tree cover, grass cover, distance to closest river or lake in line with hypothesis

distance to closest landfill coefficients not significant, even opposite trend large landfills have to be covered due to eu law

duration of observation the only significant detection covariate \rightarrow maybe other detection covariate more appropriate like the weather at the day of the observation

general reduction of occupancy of the black kite under future climate conditions, very coarse scale of climate prediction, climate is not the only factor that will change, ecology of species may also change.

References

- Araújo, M. B. and C. Rahbek (2006). “How does climate change affect biodiversity?” In: *Science* 313.5792, pp. 1396–1397. DOI: 10.1126/science.1131758.
- Bartoń, K. (2020). *MuMin: Multi-Model Inference*. URL: <https://cran.r-project.org/package=MuMin>.
- Bellard, C., C. Bertelsmeier, P. Leadley, W. Thuiller, and F. Courchamp (2012). “Impacts of climate change on the future of biodiversity”. In: *Ecology Letters* 15.4, pp. 365–377. DOI: 10.1111/j.1461-0248.2011.01736.x.
- BirdLife International (2021). *Species factsheet: Milvus migrans*. URL: <http://datazone.birdlife.org/species/factsheet/black-kite-milvus-migrans/> (visited on 07/19/2021).
- Bivand, R. and C. Rundel (2020). *rgeos: Interface to Geometry Engine - Open Source ('GEOS')*. URL: <https://cran.r-project.org/package=rgeos>.
- Blanco, G. (1994). “Seasonal abundance of Black kites associated with the rubbish dump of Madrid, Spain”. In: *Journal of Raptor Research* 28.4, pp. 242–245.
- Blas, J., F. Sergio, and F. Hiraldo (2009). “Age-related improvement in reproductive performance in a long-lived raptor: A cross-sectional and longitudinal study”. In: *Ecography* 32.4, pp. 647–657. DOI: 10.1111/j.1600-0587.2008.05700.x.
- Buchhorn, M., M. Lesiv, et al. (2020). “Copernicus Global Land Cover Layers: Collection 2”. In: *Remote Sensing* 12.6, p. 1044. DOI: 10.3390/rs12061044.
- Buchhorn, M., B. Smets, et al. (2020). *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe*. DOI: 10.5281/zenodo.3939050.
- Copernicus Land Monitoring Service (2018). *Corine Land Cover (CLC) 2018*. European Union, European Environment Agency (EEA). URL: <https://land.copernicus.eu/pan-european/corine-land-cover/clc2018>.
- Cornell Lab of Ornithology (2021). *eBird Basic Dataset*. Ithaca, New York. URL: <https://ebird.org/data/download> (visited on 03/01/2021).
- De Giacomo, U. and G. Guerrieri (2008). “The feeding behavior of the Black Kite (*Milvus migrans*) in the rubbish dump of Rome”. In: *Journal of Raptor Research* 42.2, pp. 110–118. DOI: 10.3356/JRR-07-09.1.
- Donázar, J. A. et al. (2016). “Roles of Raptors in a Changing World: From Flagships to Providers of Key Ecosystem Services”. In: *Ardeola* 63.1, pp. 181–234. DOI: 10.13157/arla.63.1.2016.rp8.
- Fick, S. E. and R. J. Hijmans (2017). “WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas”. In: *International Journal of Climatology* 37.12, pp. 4302–4315. DOI: 10.1002/joc.5086.

- Fiske, I. and R. Chandler (2011). “unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance”. In: *Journal of Statistical Software* 43.10, pp. 1–23. DOI: 10.18637/jss.v043.i10.
- Garcia, R. A., M. Cabeza, C. Rahbek, and M. B. Araújo (2014). “Multiple dimensions of climate change and their implications for biodiversity”. In: *Science* 344.6183. DOI: 10.1126/science.1247579.
- Harris, C. R. et al. (2020). “Array programming with NumPy”. In: *Nature* 585.7825, pp. 357–362. DOI: 10.1038/s41586-020-2649-2.
- Heiberger, R. M. (2020). *HH: Statistical Analysis and Data Display: Heiberger and Holland*. URL: <https://cran.r-project.org/package=HH>.
- Hijmans, R. J. (2021). *raster: Geographic Data Analysis and Modeling*. URL: <https://cran.r-project.org/package=raster>.
- Ho, T. K. (1995). “Random decision forests”. In: *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR* 1, pp. 278–282. DOI: 10.1109/ICDAR.1995.598994.
- Hunter, J. D. (2007). “Matplotlib: A 2D graphics environment”. In: *Computing in Science & Engineering* 9.3, pp. 90–95. DOI: 10.1109/MCSE.2007.55.
- Iturbide, M. et al. (2021). *Repository supporting the implementation of FAIR principles in the IPCC-WG1 Atlas*. DOI: 10.5281/zenodo.3691645. URL: <https://github.com/IPCC-WG1/Atlas> (visited on 08/13/2021).
- Lemaître, G., F. Nogueira, and C. K. Aridas (2017). “Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning”. In: *Journal of Machine Learning Research* 18.17, pp. 1–5. URL: <http://jmlr.org/papers/v18/16-365>.
- MacKenzie, D. I. and L. L. Bailey (2004). “Assessing the fit of site-occupancy models”. In: *Journal of Agricultural, Biological, and Environmental Statistics* 9.3, pp. 300–318. DOI: 10.1198/108571104X3361.
- MacKenzie, D. I., J. D. Nichols, et al. (2002). “Estimating site occupancy rates when detection probabilities are less than one”. In: *Ecology* 83.8, pp. 2248–2255. DOI: 10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2.
- Mazerolle, M. J. (2020). *AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c)*. URL: <https://cran.r-project.org/package=AICcmodavg>.
- McKinney, W. (2010). “Data Structures for Statistical Computing in Python”. In: *Proceedings of the 9th Python in Science Conference*. Ed. by S. van der Walt and J. Millman, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.
- Nagelkerke, N. J. D. (1991). “A note on a general definition of the coefficient of determination”. In: *Biometrika* 78.3, pp. 691–692. DOI: 10.1093/biomet/78.3.691.

- Panuccio, M., N. Agostini, U. Mellone, and G. Bogliani (2014). “Circannual variation in movement patterns of the Black Kite (*Milvus migrans migrans*): A review”. In: *Ethology Ecology and Evolution* 26.1, pp. 1–18. DOI: 10.1080/03949370.2013.812147.
- Pedregosa, F. et al. (2011). “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12, pp. 2825–2830. URL: <http://jmlr.org/papers/v12/pedregosa11a.html>.
- QGIS Development Team (2021). *QGIS Geographic Information System*. QGIS Association. URL: <https://www.qgis.org>.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.r-project.org/>.
- Reich, B. J., K. Pacifici, and J. W. Stallings (2018). “Integrating auxiliary data in optimal spatial design for species distribution modelling”. In: *Methods in Ecology and Evolution* 9.6, pp. 1626–1637. DOI: 10.1111/2041-210X.13002.
- Riahi, K. et al. (2017). “The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview”. In: *Global Environmental Change* 42, pp. 153–168. DOI: 10.1016/j.gloenvcha.2016.05.009.
- Robinson, O. J., V. Ruiz-Gutierrez, and D. Fink (2018). “Correcting for bias in distribution modelling for rare species using citizen science data”. In: *Diversity and Distributions* 24.4, pp. 460–472. DOI: 10.1111/ddi.12698.
- Sergio, F. and A. Boto (1999). “Nest dispersion, diet, and breeding success of Black Kites (*Milvus migrans*) in the Italian pre-Alps”. In: *Journal of Raptor Research* 33.3, pp. 207–217.
- State Meteorological Agency (AEMET) (2021). *Climate projections for the XXI Century*. URL: http://www.aemet.es/en/serviciosclimaticos/cambio%7B%5C_%7Dclimat/ (visited on 08/02/2021).
- Steen, V. A., C. S. Elphick, and M. W. Tingley (2019). “An evaluation of stringent filtering to improve species distribution models from citizen science data”. In: *Diversity and Distributions* 25.12, pp. 1857–1869. DOI: 10.1111/ddi.12985.
- Sullivan, B. L. et al. (2009). “eBird: A citizen-based bird observation network in the biological sciences”. In: *Biological Conservation* 142.10, pp. 2282–2292. DOI: 10.1016/j.biocon.2009.05.006.
- Tanferna, A., L. López-Jiménez, J. Blas, F. Hiraldo, and F. Sergio (2013). “Habitat selection by Black kite breeders and floaters: Implications for conservation management of raptor floaters”. In: *Biological Conservation* 160, pp. 1–9. DOI: 10.1016/j.biocon.2012.12.031.
- Urban, M. C. et al. (2016). “Improving the forecast for biodiversity under climate change”. In: *Science* 353.6304. DOI: 10.1126/science.aad8466.

- Van Rossum, G. and F. L. Drake (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace. ISBN: 1441412697.
- Veiga, J. P. and F. Hiraldo (1990). “Food habits and the survival and growth of nestlings in two sympatric kites (*Milvus milvus* and *Milvus migrans*)”. In: *Ecography* 13.1, pp. 62–71. DOI: 10.1111/j.1600-0587.1990.tb00590.x.
- Vinuela, J. and J. P. Veiga (1992). “Importance of rabbits in the diet and reproductive success of black kites in southwestern Spain”. In: *Ornis Scandinavica* 23.2, pp. 132–138. DOI: 10.2307/3676440.
- Wagenmakers, E. J. and S. Farrell (2004). “AIC model selection using Akaike weights”. In: *Psychonomic Bulletin and Review* 11.1, pp. 192–196. DOI: 10.3758/BF03206482.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN: 978-3-319-24277-4. URL: <https://ggplot2.tidyverse.org>.
- Zurell, D. et al. (2020). “A standard protocol for reporting species distribution models”. In: *Ecography* 43.9, pp. 1261–1277. DOI: 10.1111/ecog.04960.

Supplementary Information

The appendix consists of extra graphics and code for repeating the analysis.

Graphics

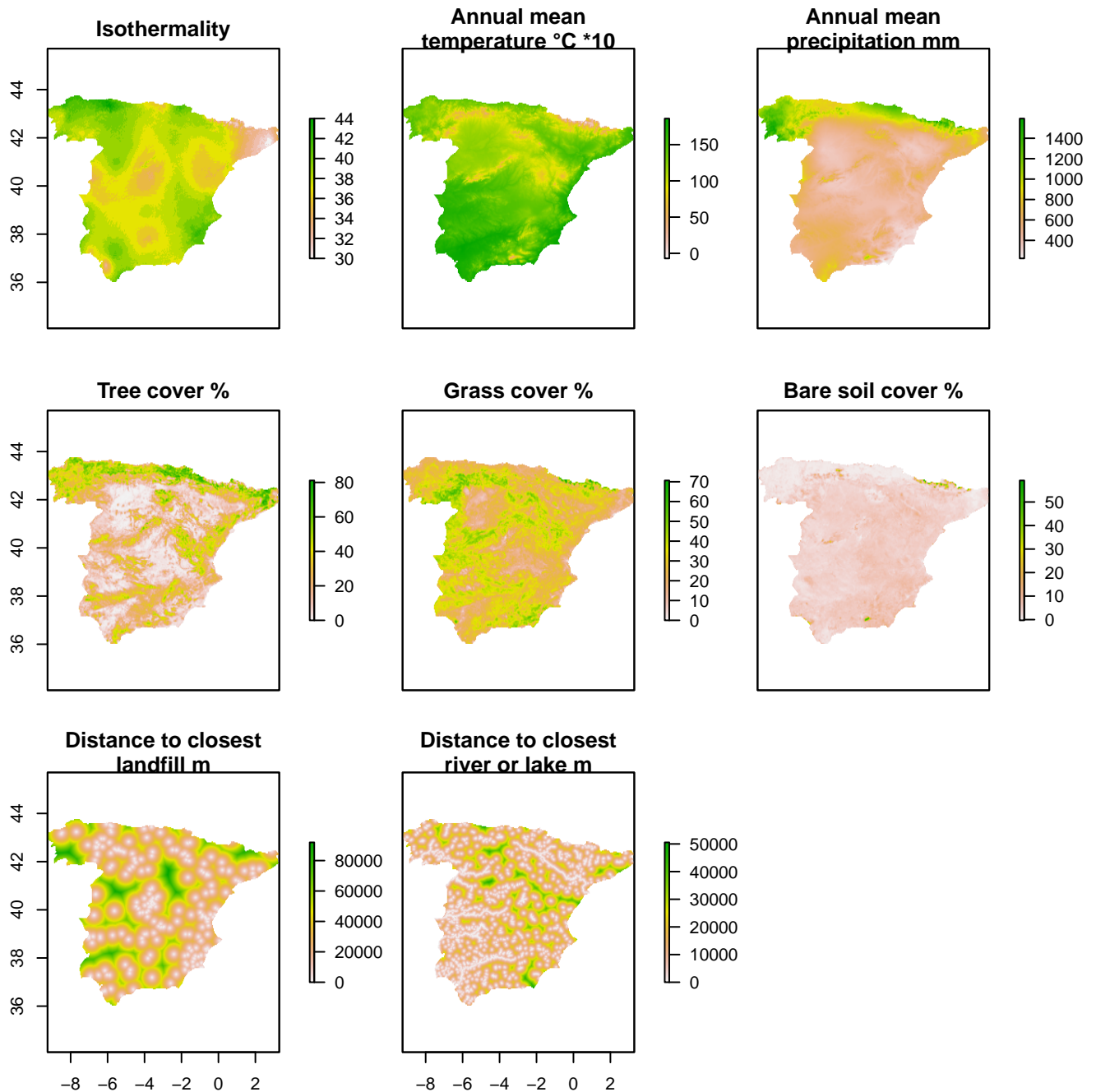


Figure 7: Maps of all site covariates, created with Hijmans (2021)

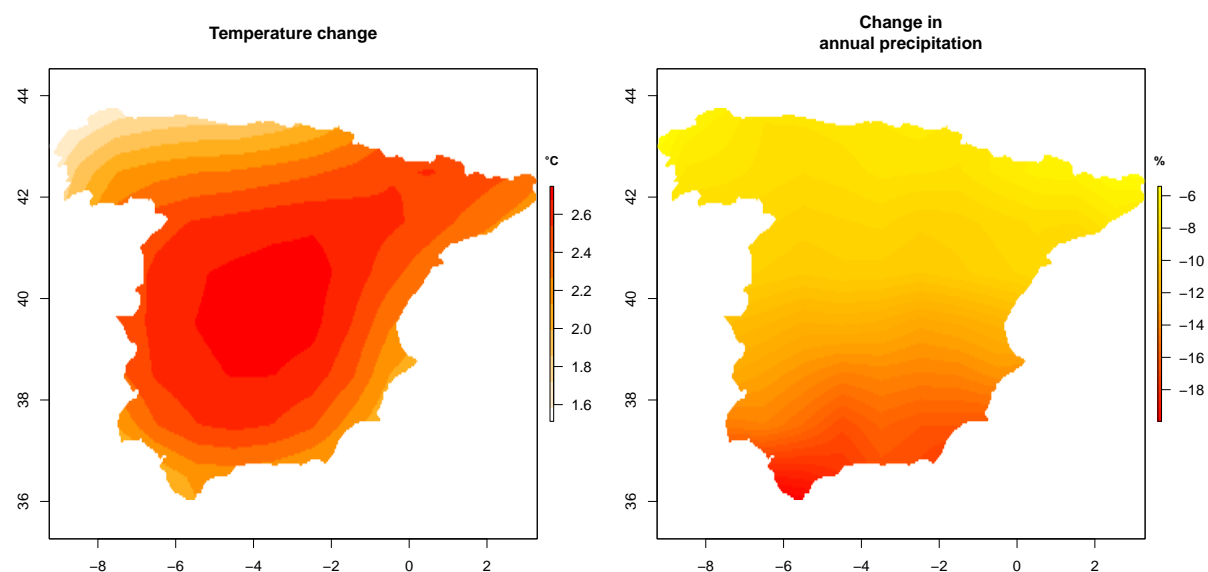


Figure 8: Change in the annual mean temperature and the annual precipitation between 1995-2014 and 2081-2100 with the SSP2 4.5 pathway, mean of 34 models (Iturbide et al. 2021), created with Hijmans (2021)

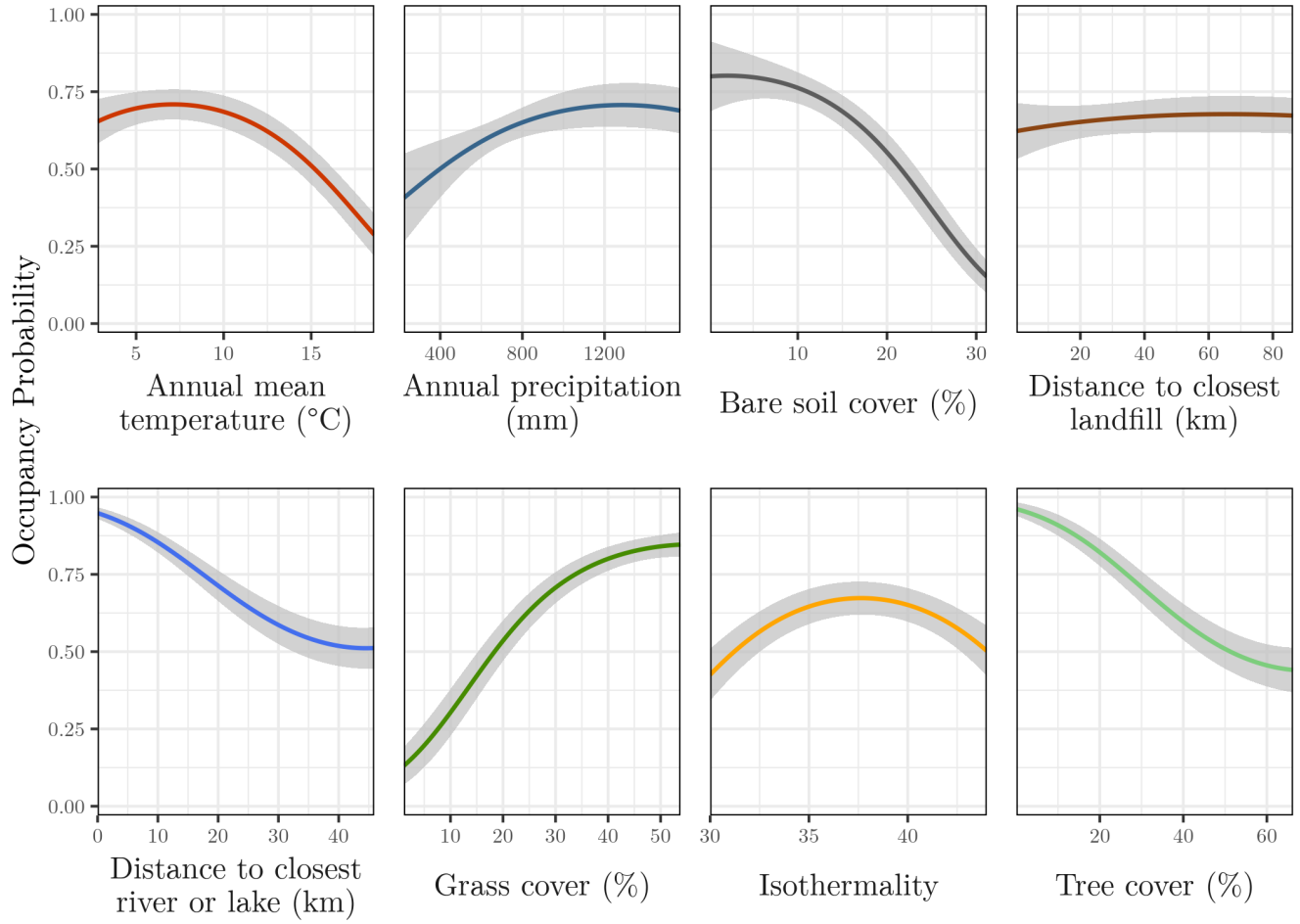


Figure 9: Predicted occupancy of the average model in response to site covariates, for each plot only the focal site covariate was varied, grey areas represent the standard error, created with Wickham (2016)

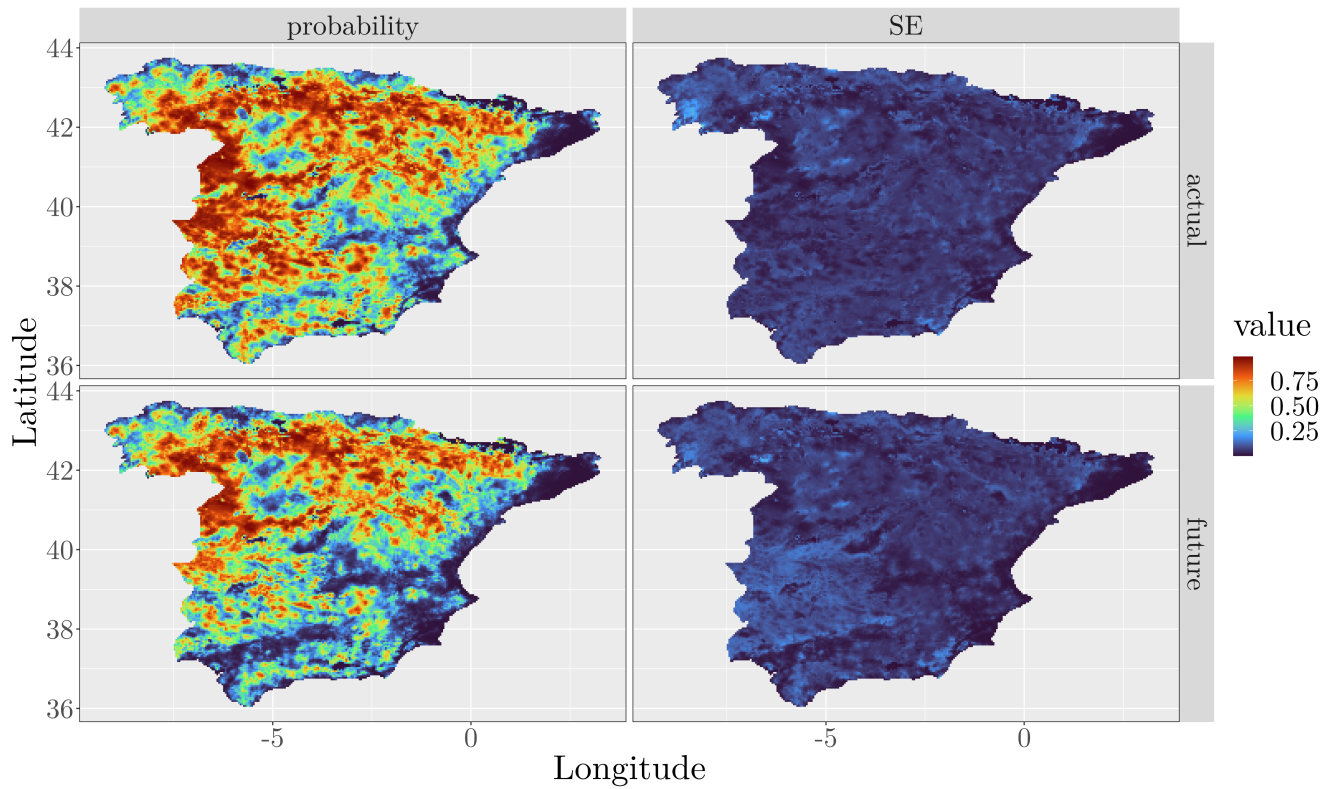


Figure 10: Occupancy map under actual and future climate conditions predicted with the average model, the occupancy probability is shown in the left maps, the corresponding standard error in the plots on the right side, created with Wickham (2016)

Code

The data and the R Code can be found at
https://github.com/FelixNoessler/QCB_Black_Kite

The R-script for running the analysis is also embedded here: