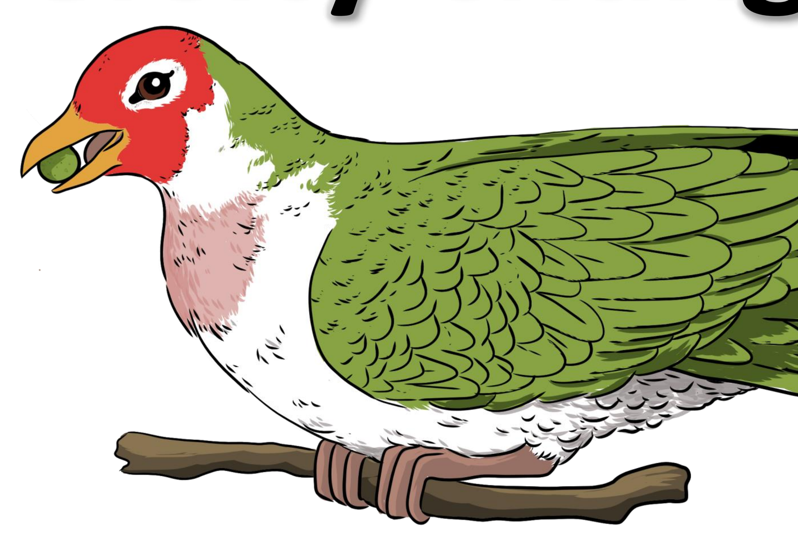
Towards predicting temporal biodiversity change from static data



Friederike J.R. Wölke[1], Carmen D. Soria[1], Gabriel R. Ortega[1], Karel Šťastný[2], Vladimír Bejček[2], Ivan Mikuláš[2], Mutsuyuki Ueta[3] & Petr Keil[1]

[1]Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Praha-Suchdol, Czech Republic
[2]Department of Ecology, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Praha-Suchdol, Czech Republic
[3]Japan Bird Research Association, Fuchu-shi, Tokyo, Japan

Background

The world is experiencing significant environmental transformations, impacting biodiversity and ecosystem functions. Obtaining temporal biodiversity data is challenging due to cost and monitoring limitations. We aim to predict temporal trends in species occupancy using static data, which are considered fixed once recorded, unlike dynamic data requiring continuous updates.

Objective

We analyzed static snapshots of species spatial distributions and their covariates from four breeding bird atlases sampled in two time periods (pre-2000 and post-2000). Our goal was to determine the predictive strength of static covariates and the overall predictive ability of the model.

What are static data and how are they linked to temporal change?

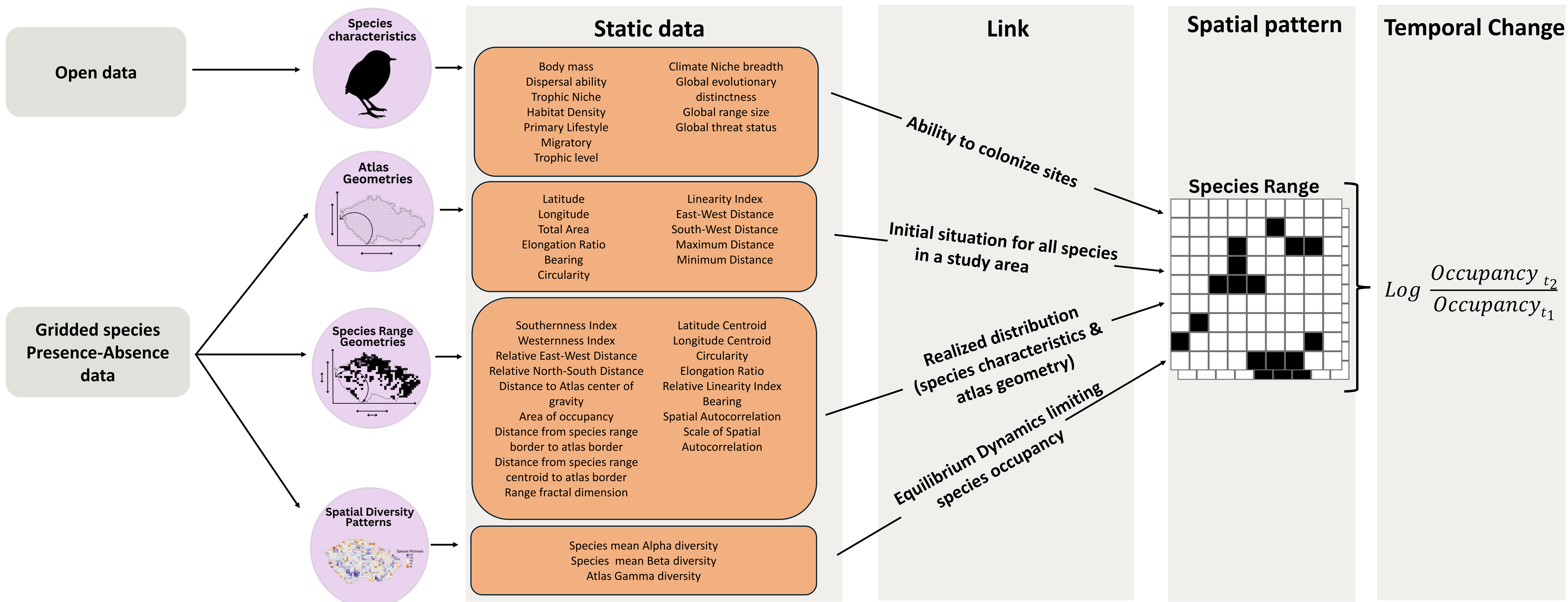


Figure 1: **What are static data and how are they linked to temporal change?**

Static data include all characteristics and metrics assumed to be independent of time (i.e., they stay constant). These data can be linked to dynamical temporal changes, via a) Species characteristics (e.g., body mass, dispersal ability, habitat niche breadth), b) Atlas limits (e.g., geographical location, spatial extent, number of atlas squares), c) Species range structure (e.g., range area, range geometry), and d) diversity metrics (e.g., average spatial diversity for each species).

These factors characterize the initial situation for all species in a study area and individual species' situations and may influence the ability and probability of species to adapt or disperse to new sites. Thus, they may explain a species' colonization potential and hence infer temporal change in occupancy (log ratio of AOO between time 1 and time 2).

We assumed that the spatial pattern of a species range in the grid is the product of the underlying population-level processes (colonization, extirpation, birth, death) and thus can be utilized as a proxy to infer temporal change dynamics.

Methods

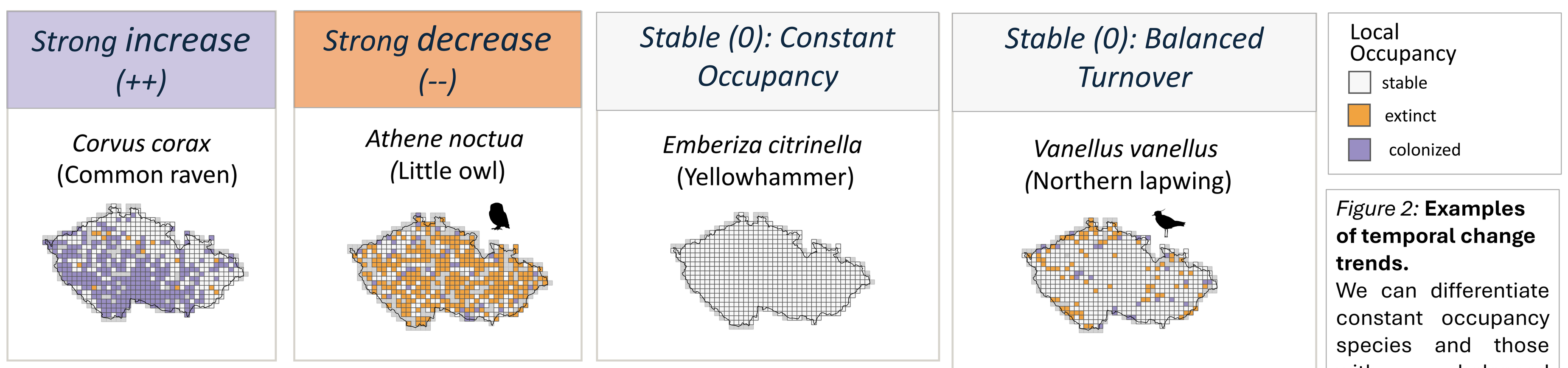


Figure 2: **Examples of temporal change trends**

Figure 2: **Examples of temporal change trends.** We can differentiate constant occupancy species and those with balanced turnover. Those very different underlying processes are both masked by a stable occupancy trend.

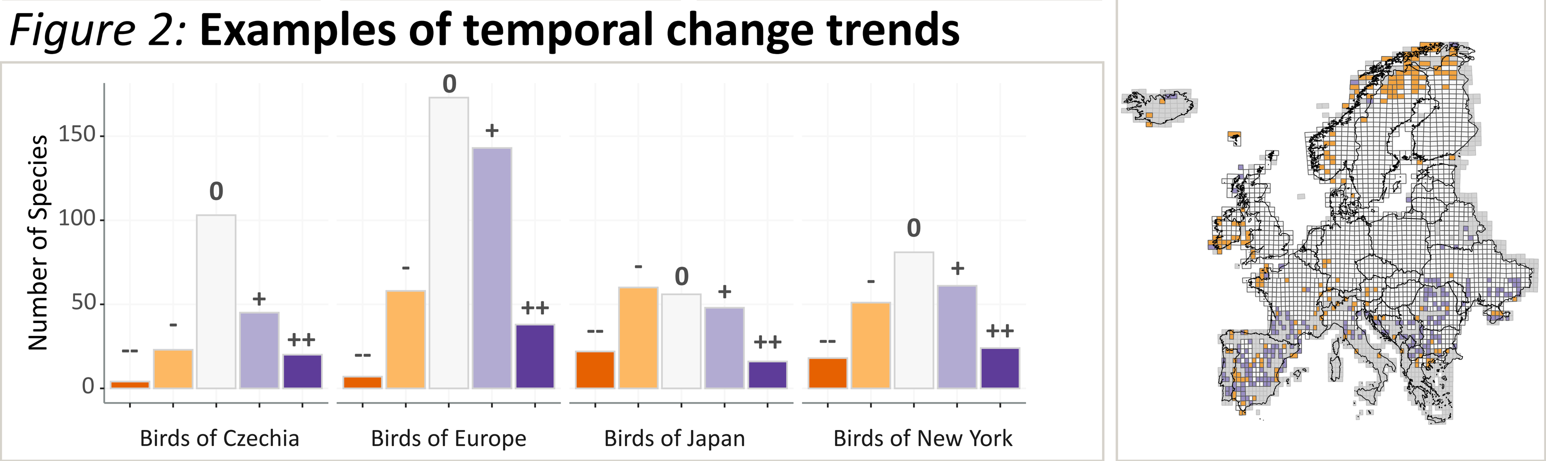


Figure 3: **Regional occupancy trends**

Figure 3: **Regional occupancy trends** by category for birds in Czechia, Europe, Japan and New York state for the first sampling periods.

Data:
We used high-quality breeding bird atlas data from Czech Republic, Japan, New York State, and Europe (N = 841). Data were aggregated over several years for each sampling period. We calculated predictors characterizing species range geometry and structure, diversity metrics, and atlas coverage. External data from various sources were integrated to extract species traits. Temporal change was defined as the log ratio of the area of occupancy (AOO) between the two sampling periods.

Model:
We developed two complex random forest models to predict the log ratio of AOO. Each model utilized 54 covariates (83 predictors) derived from the two sampling periods for both past and future changes. We used ten resampling splits and k-fold repeated cross-validation with 10 folds and 5 repeats to ensure robust model results. We identified the top predictors based on their importance across resampling runs, using variable permutation and mean squared error (MSE) decrease (Figure 4). The models' predictive performance was then evaluated against new data (Figure 5).

Results

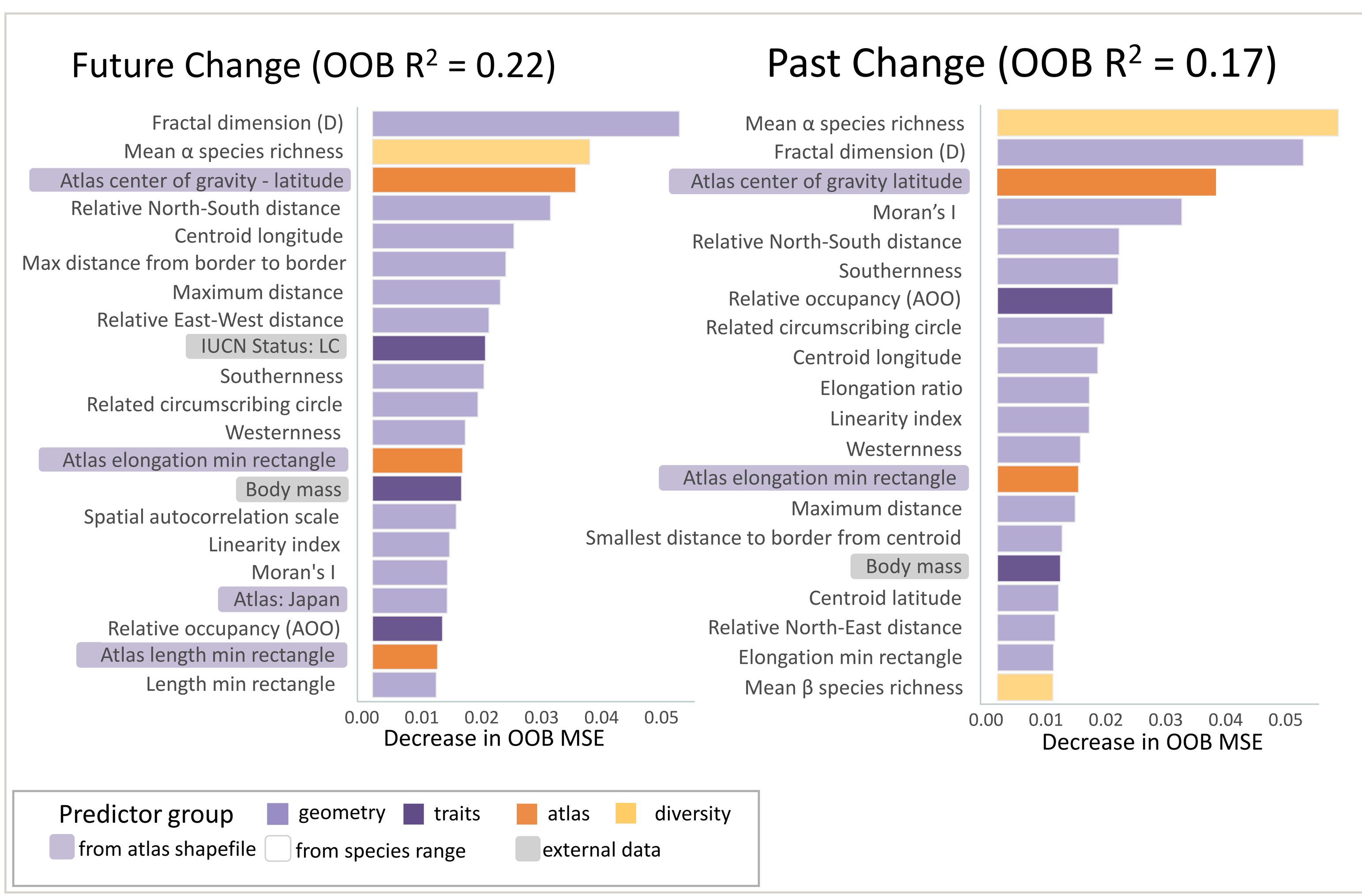


Figure 4: **Average variable importance inferred from the random forest model for a) future change and b) past change for all birds in the study.**

Importance was calculated based on permutation across resamples with k-fold repeated cross-validation. Predictor names are colored by their origin (purple = calculated from the gridded atlas shapefiles, grey = external data, white = calculated from individual species ranges) and bars are colored by predictor category (geometry, traits, atlas characteristics, diversity).

Figure 4: **Predicting change from static data: Top Predictors**

Discussion & Conclusions

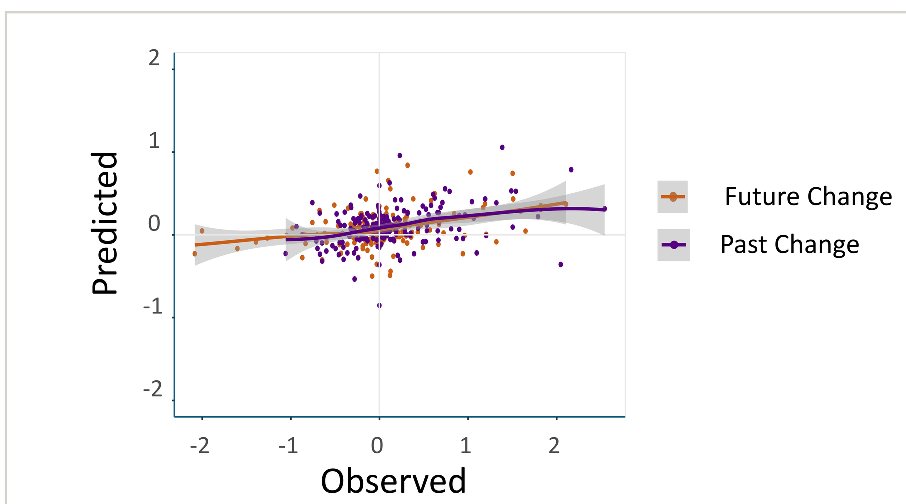


Figure 5: **Predicted versus observed change.** Demonstrating that static patterns are only partially able to predict temporal change, with better predictive strength for future changes.

Figure 5: **Predicted vs. Observed**

Although static patterns are only partially able to predict temporal change, we found that the predictive strength is lower when predicting past change as compared to future change.

Interestingly, geometric constraints of the study area and the species distribution enhance model performance significantly, suggesting that processes of temporal change leave signals in static data. Reasons for the weak model performance are potentially the diverse underlying and patterns of stable species (Figure 2), which may be diluting the spatial imprints of temporal change.
Further investigations will involve exploring stable species and adapting the model for the integration of unstructured species range data.

GitHub >>

Presenter contact: Friederike Wölke, wolke@fzp.czu.cz, @FriederikeWolke (X), https://friedarosa.github.io/

Funding: This study is funded by CZU, the Internal Grant Agency of CZU (Call 2024-2025) & the European Union (ERC, BEAST, 101044740) || Bird illustration provided by Alexandra Koch