Individual-based modelling application for intraspecific variation in habitat selection of white storks (*Ciconia ciconia*) in central Europe.

Master's Thesis

for Achievement of the Academic Degree

Master of Science (Tropical Forestry)

as a part of Erasmus Mundus Joint Masters in Global Forestry (GLOFOR)

at the Institute of International Forestry and Forest Products

Faculty of Environmental Sciences

Technische Universität Dresden

Jannatul Ferdous

Matriculation number: 5259573

Supervisors:

Dr. Ronny Peters (Forest Biometrics and Systems Analysis, TU Dresden)

Dr. Elina Takola (Department of Landscape Ecology, UFZ)

Tharandt, December 2025

Technische Universität Dresden

UFZ HELMHOLTZ Zentrum für Umweltforschung

Acknowledgement

I would like to express my heartfelt gratitude to my supervisors, Dr. Elina Takola and Dr. Ronny Peters, for their continuous guidance, support, and invaluable feedback throughout this study. Their mentorship played a crucial role in shaping the direction and depth of my research.

I am also deeply grateful to the Chair of Tropical and International Forestry at Technische Universität Dresden and to the Erasmus Mundus Global Forestry (GLOFOR) programme committee for their steady support and guidance throughout my academic journey.

My sincere appreciation goes to the Department of Landscape Ecology at UFZ and the BIOECOS research group for providing access to essential software licenses, computational resources, and research infrastructure that made this work possible.

Special thanks to the University of Copenhagen and the University of Padova, where I completed the first and second semesters of my Erasmus Mundus Global Forestry (GLOFOR) programme. The academic training and support I received at both institutions formed a substantial foundation for this master's degree and has contributed greatly to my overall learning experience.

I would like to thank my colleague Johanna Esguerra for her support, particularly for helping me identify and access the necessary resources during the initial stages of this thesis.

Last but not least, my warmest thanks go to Sreejani and Biplob for their unwavering mental and emotional support throughout the entire thesis journey.

Declaration by the candidate

I, Jannatul Ferdous, hereby declare that this thesis entitled **"Individual-based modelling application for intraspecific variation in habitat selection of white storks (*Ciconia ciconia*) in central Europe."** is an authentic record of my research conducted under the supervision of Dr. Elina Takola at the Helmholtz Centre for Environmental Research (UFZ) and Dr. Ronny Peters at the Technical University of Dresden. The results embodied in this thesis have not been submitted for any other degree or diploma award.

Tharandt, 10 December 2025

Place and Date

Signature

Permitted/not permitted for lending

Dresden

Head of the Examination Board

Declaration of Library Access and Usage Rights

I, Jannatul Ferdous, hereby entitle the library "Sächsische Landesbibliothek- Staats-und Universitätsbibliothek Dresden and Zweigbibliothek Forstwesen", to make available on their premises my Master thesis for all kinds of library usage, including reference library.

Tharandt, 10 December 2025

Place and Date

Signature

Abbreviations

NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
ed	Edge Density
pd	Patch Density
shei	Shannon's Evenness Index
shdi	Shannon's Diversity Index
ent	Shannon Entropy
RSF	Resource Selection Function
IBM	Individual-based Model
LULC	Land-use and Land-cover

Table of Contents

List of Figures

List of Tables

Abstract (English)

Understanding how white storks (*Ciconia ciconia*) interact with dynamic agricultural landscapes requires integrating empirical habitat selection data with mechanistic models that explain individual behavioural variability. The aim of this study was to This study examines how landscape transformation shapes white stork habitat selection, movement decisions, and breeding status using an integrated modelling approach. Using GPS tracking data from 33 storks in Germany, this study quantified resource selection with mixed-effects Resource Selection Functions (RSFs) using the empirical data and embedded these model coefficients into an Individual-Based Model (IBM) to evaluate how landscape structure influences movement, energy balance, and breeding status. RSF showed that vegetation indices (NDVI, NDWI) provided weak explanatory power, largely due to the narrow environmental gradients and rapid vegetation turnover in the study area. Instead, land-use and land-cover (LULC) classes were the primary determinants of habitat use, with strong selection for herbaceous croplands and opportunistic use of impervious surfaces. Storks also preferred landscapes with high edge density and low evenness, indicating a reliance on fragmented agricultural systems that maximise prey accessibility.

The IBM reproduced these empirical selection patterns and revealed that LULC classes directly influences energetic performance. Simulated breeding outcomes increased in years with greater availability of preferred cropland habitats, demonstrating clear linkages between landscape composition and reproductive fitness. However, strong individual variability reduced predictive accuracy in both the empirical RSF (mean AUC = 0.58) and IBM-derived RSF models (overall mean AUC = 0.51). Breeder-specific models performed substantially better (AUC up to 0.61), reinforcing that combining behavioural states masks true selection signals.

Overall, the RSF-IBM integration demonstrated that white stork habitat use is shaped primarily by LULC classes and landscape structure rather than vegetation indices, and that breeding status is sensitive to temporal changes in landscape composition. These results highlight the ecological risks posed by homogenisation of agricultural systems and underscore the need for conservation strategies that prioritise fragmented open agricultural landscapes, and individual variability across populations.

Keywords: Individual based model, Resource selection function, White stork, Agricultural landscape, Habitat selection

Zusammenfassung (Deutsch)

Das Verständnis dafür, wie Weißstörche (Ciconia ciconia) mit dynamischen Agrarlandschaften interagieren, erfordert die Verknüpfung empirischer Habitatwahl-Daten mit mechanistischen Modellen, die individuelle Verhaltensvariabilität erklären können. Ziel dieser Studie war es zu untersuchen, wie Landschaftstransformation die Habitatwahl, Bewegungsentscheidungen und den Brutstatus des Weißstorchs beeinflusst, basierend auf einem integrierten Modellierungsansatz. Unter Verwendung von GPS-Tracking-Daten von 33 Störchen in Deutschland quantifizierte die Studie die Ressourcennutzung mit gemischten Resource-Selection-Function-Modellen (RSFs) auf Grundlage empirischer Daten und implementierte diese Modellkoeffizienten anschließend in ein Individualbasiertes Modell (IBM), um zu bewerten, wie Landschaftsstruktur Bewegung, Energiebilanz und Brutstatus beeinflusst. Die RSF-Analysen zeigten, dass Vegetationsindizes (NDVI, NDWI) aufgrund geringer Umweltgradienten und schneller Vegetationsdynamik im Untersuchungsgebiet nur eine geringe Erklärungskraft besaßen. Stattdessen erwiesen sich Landnutzungs- und Landbedeckungsklassen (LULC) als wichtigste Determinanten der Habitatnutzung, wobei eine deutliche Präferenz für krautige Ackerflächen und eine opportunistische Nutzung versiegelter Bereiche erkennbar war. Zudem bevorzugten Störche Landschaften mit hoher Kantenlänge und geringer Gleichverteilung, was auf eine Abhängigkeit von fragmentierten Agrarsystemen hindeutet, die eine hohe Beutezugänglichkeit ermöglichen.

Das IBM reproduzierte diese empirischen Selektionsmuster und zeigte, dass LULC-Klassen den energetischen Zustand der Individuen direkt beeinflussen. Simulierte Bruterfolge nahmen in Jahren mit einem größeren Anteil bevorzugter Ackerflächen zu, was klare Zusammenhänge zwischen Landschaftszusammensetzung und reproduktiver Fitness belegt. Allerdings verringerte eine starke individuelle Variabilität die Vorhersagegenauigkeit sowohl in den empirischen RSF-Modellen (mittlere AUC = 0,58) als auch in den aus dem IBM abgeleiteten RSF-Modellen (mittlere AUC = 0,51). Modelle, die ausschließlich auf Brutvögel beschränkt waren, erzielten deutlich bessere Leistungen (AUC bis 0,61), was unterstreicht, dass die Vermischung unterschiedlicher Verhaltenszustände wahre Selektionssignale überdeckt.

Insgesamt zeigt die RSF-IBM-Integration, dass die Habitatnutzung des Weißstorchs in erster Linie durch LULC-Klassen und Landschaftsstruktur geprägt wird und dass der Brutstatus sensibel auf zeitliche Veränderungen der Landschaftszusammensetzung reagiert. Die Ergebnisse verdeutlichen die ökologischen Risiken einer zunehmenden Homogenisierung landwirtschaftlicher Systeme und betonen die Notwendigkeit von Schutzstrategien, die

fragmentierte offene Agrarlandschaften und die individuelle Variabilität innerhalb von Populationen stärker berücksichtigen.

Schlagwörter:
Individuelles agentenbasiertes Modell, Ressourcenwahlfunktion, Weißstorch, Agrarlandschaft, Habitatselektion

Chapter 1. Introduction

Understanding why animals select certain habitats and avoid others has become increasingly important in ecological and evolutionary research, particularly for informing conservation strategies. Habitat selection is defined as the behavioural process by which animals either actively choose particular habitats or passively remain in environments where they persist (Boyce & McDonald, 1999; Matthiopoulos et al., 2020; Northrup et al., 2022). It is influenced by factors such as resource availability and population density, meaning that individual space use is shaped by both environmental conditions and the presence of other organisms. This makes the process inherently density-dependent and sensitive to how it is measured (Northrup et al., 2022). Furthermore, human-altered landscapes can interfere with these natural patterns by limiting movement and dispersal. Habitat fragmentation may reduce dispersal tendencies and trap individuals in unsuitable environments, ultimately impacting their fitness and survival (Cayuela et al., 2020). These insights highlight the importance of developing habitat models that account for both behavioural processes and the ecological effects of anthropogenic change.

Animal movement is not only a response to environmental structure but also a driver of ecological processes, such as foraging efficiency, predator-prey interactions, and species distribution (Schlägel et al., 2020). Several studies have reported the impact of anthropogenic habitat alteration on the behaviour of different species (Cayuela et al., 2020; Gilbert et al., 2016) and many of these traditional studies assumed that all individuals of the same species are ecologically equivalent (Gomez et al., 2025), which is only valid if differences in resource use among individuals are uncommon, minimal, or have little impact on broader ecological processes (Bolnick et al., 2003; Des Roches et al., 2017). But this assumption overlooks important ecological differences among individuals, which may include distinct feeding preferences, risk-taking behaviours, or functional roles, such as the task division seen in social insects (Jeltsch et al., 2025). According to Rohwäder and Jeltsch (2022), populations with a mix of bold and shy individuals were more resilient to fragmentation, while behaviourally homogeneous groups, especially those dominated by cautious individuals, experienced greater negative impacts. Moreover, the habitat alteration can also shape the behaviour of individuals, which can also be passed through genetics to future offspring (Cayuela et al., 2020). This variation is particularly relevant in human-dominated landscapes, where habitat availability and quality can fluctuate at fine spatial and temporal scales.

Among all other animals, birds function as both important conservation priorities and reliable indicators of agricultural change, largely because of their responsiveness to land-use dynamics and their relative ease of observation and study (Fraixedas et al., 2020; Ormerod & Watkinson, 2000). However, over the last decades, the farmland birds have been declining globally, resulting from agricultural intensification and land-use changes. Particularly in Europe, the Common Agricultural Policy (CAP) has played a key role in transforming agricultural practices, including changes in how crops and grasslands are managed, the arrangement and specialisation of crops, and the reduction of non-crop features like hedgerows and field margins. These alterations have had substantial effects on the habitats and populations of farmland birds (Jerrentrup et al., 2017). Although the impact of intensive agriculture differs among bird species, it is generally associated with reduced bird occurrence across a wide range of species (Rigal et al., 2023; Siriwardena et al., 2000). Evidence from Central and Eastern Europe, where less intensive farming persisted until EU accession, further confirms this trend: despite increased conservation funding, farmland bird populations declined following intensification associated with CAP-driven reforms, suggesting that the ecological costs of intensification seem to overshadow the benefits of current conservation schemes (Butler et al., 2010; Reif et al., 2024). Specifically, recent findings indicate that not only rare species, but also common birds, are significantly affected by land-use change, with common species often showing the strongest declines in response to agricultural intensification (K. P. Davis et al., 2023) . This underscores the need for conservation strategies that address a better understanding of different species responses to changing landscapes, including the common farmland birds like the white stork.

The white stork serves as a representative species for agricultural habitat biodiversity and acts as an umbrella species, offering conservation benefits to a range of organisms that inhabit farmland landscapes (Pestka et al., 2023). As a species that commonly feeds in farmland and nests near human-made structures, it exemplifies how human-driven land-use patterns can influence both movement behaviour and habitat (Standfuß et al., 2022), which also makes it vulnerable to food scarcity resulting from habitat alteration and anthropogenic barriers in movement corridors (Rotics et al., 2016). They select foraging habitats based on vegetation structure (Standfuß et al., 2022) and food availability, typically favouring open areas with low herbaceous cover (Latus & Kujawa, 2005; Pestka et al., 2023). In Europe, they are opportunistic predators that prey on small mammals, amphibians, reptiles, earthworms, and insects, and they show a strong preference for traditionally managed pastures, meadows, and

shallow field wetlands (Pestka et al., 2023). Later in the breeding season, they are also frequently observed feeding on mowed meadows and freshly ploughed fields, often following tractors to forage more efficiently. Additionally, white storks display behavioural difference depending on age, life stage, sex, colony size, and prior reproductive success (Vergara et al., 2006; Zurell et al., 2018). Their tendency to maintain site fidelity is strongly influenced by age and breeding status; for example, adult storks typically exhibit high nest-site fidelity during the breeding season (Vergara et al., 2006; Zurell et al., 2018) but may show reduced fidelity outside of it (Berthold et al., 2002; Flack et al., 2016). Despite being a common breeding species in Europe, they started declining during the late 1940s, mostly due to landscape alteration, agricultural intensification, increased human settlement, as well as climate change in the wintering grounds (Itonaga et al., 2011; Pestka et al., 2023). But recently, their numbers have been seen to increase in some regions (Itonaga et al., 2011) as a result of the various conservation practices adopted in many regions of Europe. Initiatives like establishment of artificial nesting platforms (Santopaolo et al., 2013) and 'European Stork Villages' founded by EuroNatur (Ferger & Schwaderer, 2016) have contributed in the ascending population trend of white storks in Europe. Despite the recent recovery, understanding their habitat preferences through both statistical and individual-based modelling is critical for designing future conservation actions that can maintain and reinforce this positive population trajectory. Moreover, white storks are among the most intensively monitored bird species in Europe, with long-term GPS tracking and strong citizen-science engagement, enabling robust spatial and temporal analyses (Berthold et al., 2002). These characteristics make the white stork an ideal candidate for investigating habitat preference using both statistical and individual-based modelling approaches.

Studying how animals use space and select resources provides valuable insights into their essential habitat needs. To uncover the factors influencing resource selection, it is important to examine how animals interact with their environment across different ecological settings (Cornelsen et al., 2024). Resource Selection Functions (RSFs) are essential tools in ecological research used to quantify how animals select resources. They are defined as functions proportionate to the probability of a resource being used. RSFs were developed alongside quantitative models of natural selection, as both rely on similar statistical methods (Boyce & McDonald, 1999; A. J. Davis et al., 1994). RSF models are commonly built using generalised linear models (GLMs), particularly logistic regression, and are used to estimate habitat preference based on observed use of resources. It is also a form of habitat suitability index

(HSI) but differs from HSI models by being statistically derived from empirical data rather than expert opinion. When integrated with GIS, RSFs become powerful tools for conservation, land-use planning, and ecological risk assessment. Despite their utility, RSF models may not always perform well across different regions or time periods due to behavioural or ecological variability, making predictive accuracy the most critical measure of model performance. (Boyce et al., 2002b).

Individual-based modelling (IBM) is a mechanistic approach that represents ecological systems as collections of individuals interacting with each other and their environment. Unlike density-based models that rely on population averages, IBMs assign traits, behaviours, and decision rules to individuals, allowing system-level dynamics to emerge from their interactions (Grimm & Railsback, 2013; Stadtländer, 2012). This approach enables researchers to capture ecological complexity by linking individual variability and behavioural processes to broader population and ecosystem outcomes.

Since the 1990s, IBMs have been increasingly applied across ecology, with notable contributions in habitat selection, animal movement, and species management. Railsback et al. (2003) showed how IBMs outperform conventional density-based models in explaining stream trout habitat selection. Carter et al. (2015) developed a spatially explicit IBM for tigers (*Panthera tigris*) in Chitwan National Park, successfully replicating patterns of reproduction, mortality, dispersal, and territoriality. In disease ecology, Gritter et al. (2024) applied a GPS-based IBM to mule deer (*Odocoileus hemionus*), linking habitat use, home ranges, and social behaviour to contact rates relevant for disease spread. Recent advances have also highlighted IBMs as tools for scaling up from individual energetics to community coexistence (Szangolies et al., 2024), predicting animal distributions under global change (Gomez et al., 2025), and incorporating behavioural variation, such as risk-taking or foraging personalities in fragmented landscapes (Dammhahn et al., 2022; Rohwäder & Jeltsch, 2022). These applications draw on diverse data sources, including telemetry and GPS tracking, behavioural experiments, habitat surveys, and long-term ecological monitoring, underscoring the versatility of IBM in integrating empirical evidence into mechanistic models.

The key strength of IBM lies in its ability to mechanistically link individual behaviour to higher-level ecological patterns (DeAngelis & Mooij, 2005; Grimm & Railsback, 2007; Jeltsch et al., 2025), making it particularly relevant for studying complex processes such as species interactions, population dynamics, and ecosystem resilience. More recently, this perspective

has been advanced under the concept of individual-based global change ecology (IBGCE), which stresses the role of individual variability, adaptive behaviour, and eco-evolutionary feedback in shaping biodiversity outcomes under anthropogenic pressures (Jeltsch et al., 2025). With improvements in computational power, remote sensing, and big-data analytics, IBMs are now capable of incorporating individual-level processes at larger spatial and temporal scales. This positions them not only as valuable tools for addressing specific ecological or epidemiological questions but also as key components of a predictive framework to inform conservation, disease management, and sustainability planning in an era of global change.

A recent literature has highlighted growing concerns that statistical approaches in habitat selection research are often developed ignoring the ecological theory and practical application, which can lead to confusion, misinterpretation, and reduced usefulness for both ecologists and conservation practitioners (Northrup et al., 2022). Although Resource Selection Functions (RSFs) can quantify population-level habitat preferences, they assume that all individuals behave similarly and that landscapes remain static, making them poorly suited for evaluating how behaviour and resource availability interact in dynamic agricultural systems. Conversely, mechanistic Individual-Based Models (IBMs) can simulate decision-making, energy balance, and movement under changing conditions, but most existing stork IBMs rely on theoretical rules or assumed energetic parameters and are not informed by empirical habitat selection data (Johst et al., 2001; Zurell et al., 2015). This limitation is evident even in the most recent conservation-focused applications. A recent study in the UK used an IBM to evaluate management interventions for the reintroduced white stork population, but the authors acknowledged that their model relied on assumed demographic parameters and explicitly called for more empirical data to improve future projections (Mayall et al., 2023). This reveals a clear knowledge gap that no existing model integrates an RSF directly as the movement decision rule within a stork IBM, meaning that empirically observed habitat preferences have yet to be embedded into the mechanistic simulation of individual behaviour of white storks. This gap Is important to investigate as white storks experience strong spatial and temporal variation in resource distribution, and their reproductive output is tightly linked to foraging efficiency. Without linking empirical habitat preferences to mechanistic movement and fitness simulations, conservation planning risks relying on statistical predictions that overlook behavioural drivers of resource use. Therefore, integration of RSF and IBM is necessary to creates a unique approach that reproduces how individuals respond to landscape structure, improving biological realism and reducing the gap between statistical habitat selection patterns

and simulated movement outcomes. Overall, it provides a unified modelling approach that links behavioural ecology, landscape structure, and population-level fitness outcomes, which traditional RSFs and purely mechanistic models cannot achieve independently.

1.1 Aim and Goals

The aim of this study is to understand how white storks interact with an intensively managed agricultural landscape shaped by long-term land-use change and ongoing shifts in habitat structure by identifying the environmental features that shape their habitat preferences. Specifically, this research seeks to clarify whether stork habitat use is influenced by vegetation indices, landscape configuration and land-use patterns, or whether it emerges from individual behavioural variability across the population. A further focus is to evaluate how changes in the availability of preferred habitats may alter their resource use and breeding status. This idea is based on the concept that breeding directly reflects an individual's ability to convert energy and resources into viable offspring, which is the fundamental measure of biological fitness (Matthiopoulos et al., 2020). By combining empirical evidence of habitat selection with simulations of individual behaviour, the study aims to provide an integrated ecological perspective on how landscape transformation influences both movement decisions and reproductive performance in white storks. This aim leads to two goals of this study:

1. To identify the key environmental factors that shape resource selection in white storks.
2. To evaluate how landscape change influences individual foraging and resource-selection patterns.

To address these two goals, the study first used RSFs to quantify how key environmental factors influence habitat selection. These empirical selection patterns were then incorporated into an IBM, where fixed and random effects shaped movement behaviour alongside ecological rules governing energy gain and survival. This hybrid approach allowed individuals' foraging behaviour to translate into fitness outcomes: breeding status (breeder or non-breeder) and mortality.

Chapter 2. Materials and Methods

2.1 Overview of the Methodology

The study used empirical GPS tracks from 33 white storks in Germany to quantify how landscape characteristics shape their habitat use. The GPS locations were processed to derive vegetation indices (NDVI, NDWI) and a set of landscape configuration and fragmentation metrics, which were then linked to the birds' used-available steps to model resource selection using logistic regression (RSF) (section 2.4). The fixed and random effects estimated from the RSF were subsequently incorporated into the Individual-Based Model (IBM) (section 2.5) as the movement decision rules, complemented by habitat-specific energy allocation and ecological processes such as site fidelity. This integration allowed simulated individuals to make movement and foraging decisions that reflect empirically observed selection patterns while accumulating energy that determines their breeding outcomes.

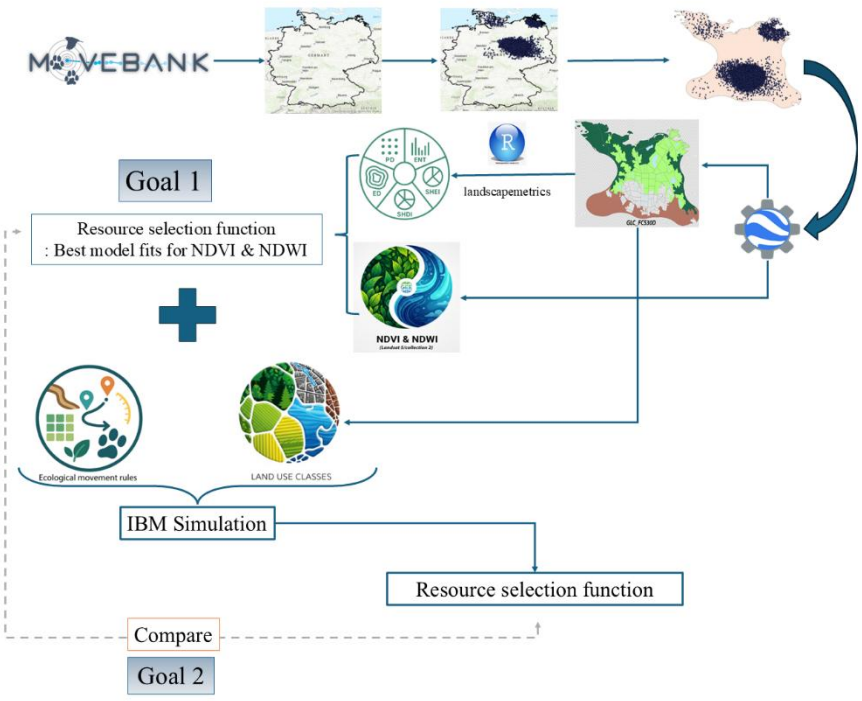

Figure 1. Working framework starting from the data import from Movebank, data processing and then model processing according to the study goals.

2.2 Data Sources and Description

2.2.1 Movebank Database

The study has used one stork GPS dataset downloaded from the Movebank database (Kranstauber et al., 2011). Movebank is a global online platform designed to support the collection, management, sharing, and long-term archiving of animal movement and bio-logging data. Established in 2007 and maintained by the Max Planck Institute of Animal Behaviour in collaboration with the University of Konstanz, the Ohio State University, the North Carolina Museum of Natural Sciences, and the Senckenberg Society for Nature Research, Movebank provides a centralised, standardised, and secure data platform for working with animal-tracking information from around the world (Wikelski et al., 2025).

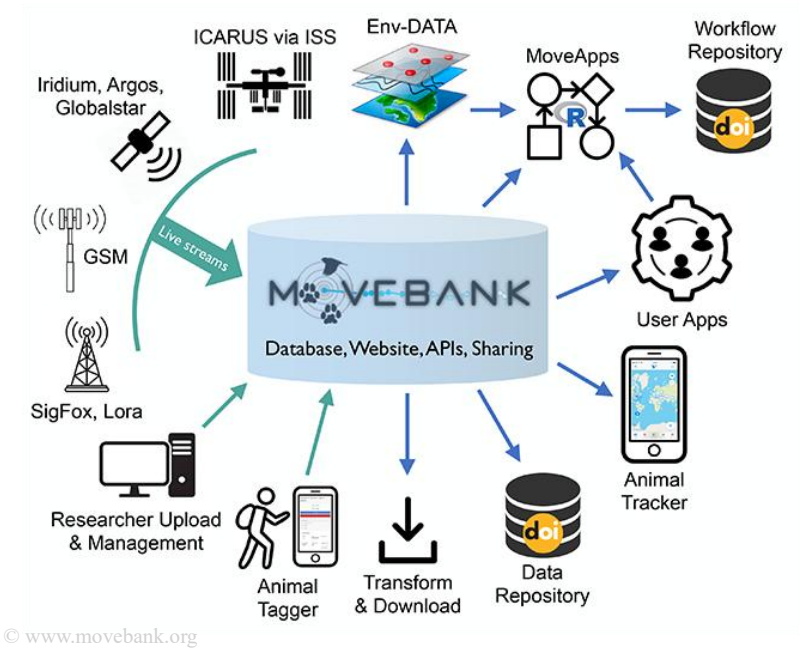

Figure 2. Conceptual diagram of Movebank data architecture showing how animal movement data are collected from multiple tracking technologies (e.g., GPS, GSM, Argos, ICARUS), uploaded and managed by researchers, linked with environmental layers (Env-DATA), and processed through Movebank's database, APIs, analytical tools, and user applications.

The Movebank database accommodates a wide range of data types generated by animal-borne sensors, including GPS, Argos, GSM, VHF, acoustic, and light-level telemetry, as well as measurements from accelerometers, magnetometers, temperature loggers, and other physiological or environmental sensors. Each record is timestamped and linked to individual animals through detailed metadata describing tag specifications, deployment methods, and

study context. This harmonized structure ensures compatibility across species, tracking technologies, and research programs (Kranstauber et al., 2011). Movebank primarily serves researchers, conservationists, and wildlife managers, but it is also used by government agencies, NGOs, and educational institutions to support ecological monitoring, environmental policy, and species-protection initiatives (Kays et al., 2022). Data owners retain full control over access rights and can choose to share datasets privately within collaborations or publicly for broader scientific use. The platform supports thousands of active studies worldwide, facilitating large-scale comparative analyses of migration, habitat selection, and behavioural ecology.

2.2.2 MPIAB Argos white stork tracking (1991-2017)

The MPIAB Argos white stork tracking dataset, collected from 1991 to 2017, was initiated by the Max Planck Institute of Ornithology and its collaborators (Berthold et al., 2022). The dataset primarily relies on Argos satellite telemetry and was later expanded to include GSM-based tracking, with some individuals involved in experimental studies. It contains spatiotemporal location data for 214 individual white storks tagged across Belgium, Germany, Greece, Israel, Poland, Russia, South Africa, Spain, and Switzerland (Berthold et al., 2022).

For this thesis, a subset of the MPIAB data was used, consisting of breeding-season locations of individuals that remained within the German national boundary for at least one complete breeding season between 1992 and 2017.

2.2.3 Movement Data Subset

The subset was generated in ArcGIS Pro 3.3.2 using the Spatial Join tool by filtering the tracking data with the German national boundary shapefile. The dataset spans from April 1992 to August 2012, containing a total of 1,302 GPS records. Tracking intervals are irregular, and two sensor types were used, primarily Argos Doppler-shift transmitters. Among the tagged individuals, one was identified as male, while 32 had unidentified sex. Age classes include 3 adults, 12 juveniles in their first winter, 1 juvenile, and 17 individuals of unspecified age. Regarding living status, 5 individuals were wild, 10 were raised or kept in animal care centres, and 17 had no recorded status (Data summary: Appendix table S1).

2.3 Environmental Variables and Landscape Metrics

Environmental variables were derived to represent habitat composition, vegetation condition, and landscape structure across the study area. All spatial analyses were conducted using Google Earth Engine (GEE), ArcGIS Pro 3.3.2, and R (version 4.3.1).

2.3.1 Used-Available Points

To establish the sampling framework for the Resource Selection Function (RSF), a used-available design was implemented. Used points represent the actual GPS locations where individual storks were recorded during the breeding season, while available points represent potential locations within the accessible area that were not used but could have been visited. For each used point, an annular sampling region between 1 km and 50 km from the point was generated to represent the potential movement space of the individual. The used and available points were created in R using the 'sf' package (Pebesma, 2016). For each used location, ten available points were generated, maintaining a fixed 10:1 ratio between available and used records for each individual. Each available point was assigned to the same year and breeding season as its corresponding used point to preserve temporal consistency. Metadata such as bird ID, tag ID, year, and month were preserved for each sampled point. All used points (true GPS points) were assigned a value of 1, while all available points were assigned a value of 0. The combined dataset of used and available points was exported as a GeoPackage and a shape file (.shp), which were subsequently used for extracting NDVI, NDWI, and landscape variables in Google Earth Engine and for computing landscape metrics in R.

Before environmental variable extraction, the study area was spatially defined by creating a 50 km buffer around all used and available points using the Buffer tool in ArcGIS Pro. This buffer was used to represent the potential area available for exploration during the breeding season and to ensure that both used and available locations were captured within a realistic ecological range (Gilbert et al., 2016). All subsequent environmental layers were exported to this buffered extent to maintain consistent spatial boundaries across datasets.

2.3.2 LULC, NDVI and NDWI data extraction

Land-use and land-cover (LULC) data were obtained from the GLC_FCS30D global dataset at 30 m spatial resolution, available for July in the years 1990, 1995, 2000, 2001, and 2012 (Zhang et al., 2024). To ensure temporal consistency with the tracking dataset (1992–2012), each RSF year was matched to the closest available LULC year. Each raster was projected to WGS 84 (EPSG:4326) and clipped to the buffered study area before further analysis.

For vegetation and surface-water indices, monthly Normalised Difference Vegetation Index (NDVI) and Normalised Difference Water Index (NDWI) values were derived from Landsat 5 TM and Landsat 7 ETM+ Collection 2 Level-2 surface-reflectance data for the period 1992-2012. The imagery was pre-processed in GEE by masking clouds, shadows, and snow using the QA_PIXEL band and rescaling digital numbers using the official Landsat surface-reflectance calibration factors. Monthly median composites were generated for the months July and August, from which NDVI and NDWI were computed using the standard band ratios (Eq. i and ii).

Equation (i):

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

Equation (ii):

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

where:

NIR = surface reflectance in the near-infrared band

Red = surface reflectance in the red visible band

Green = surface reflectance in the green visible band

The resulting NDVI and NDWI values were extracted at each used and available point within a 250 m buffer, representing local vegetation productivity and surface-water availability during the breeding season. These point-based values were exported as CSV tables and merged with the RSF dataset for model fitting.

Subsequently, continuous NDVI and NDWI layers were generated for July-August of 2012 2018, and 2024 at 30 m spatial resolution, along with LULC rasters for 2018 and 2024. The NDVI and NDWI composites for 2018 and 2024 were derived from Landsat 8 OLI and Landsat 9 OLI-2 surface reflectance imagery, while the LULC layers were obtained from the GLC_FCS30D dataset. All raster layers were exported from Google Earth Engine as GeoTIFFs and then used as environmental inputs in the IBM to simulate white stork movement and habitat selection dynamics under changing landscape conditions.

2.3.3 Calculation of Landscape Metrics

Landscape configuration and diversity were quantified to assess habitat heterogeneity within the breeding range of white storks. The analysis was based on the LULC rasters exported from Google Earth Engine at 30 m spatial resolution. These LULC layers served as the categorical input for calculating five landscape-level metrics using the landscapemetrics package (Hesselbarth et al., 2018, 2019) in R. For each point, a 500 m circular buffer was created to represent the local habitat neighbourhood. The LULC raster corresponding to the appropriate year was cropped and masked to each buffer extent, and five landscape-level metrics were computed according to the theoretical foundation used in FRAGSTAT software (McGarigal & Marks, 2012), which is the theoretical base for the R package landscapemetrics.

Patch Density (pd) represents the number of discrete patches per unit area and quantifies habitat fragmentation within the landscape. Borderline or edge patches are excluded from the computation.

Edge Density (ed) measures the total length of class boundaries per unit area and reflects the amount of edge present within the landscape. Only "true" edges between different land-cover classes are included, while boundary contributions are handled according to the user-defined settings.

Shannon's Diversity Index (shdi) measures the compositional diversity of the landscape. It represents the negative sum of each class's proportional area times its own proportion and is typically used as a relative measure to compare diversity across landscapes.

Shannon's Evenness Index (shei) expresses how evenly land-cover classes are distributed. It is obtained by dividing the observed Shannon's Diversity Index by its maximum possible value for the given number of classes, using proportional areas based on the total landscape area excluding internal background.

Entropy (ent) quantifies the configurational complexity of the landscape. It measures the uncertainty in class adjacencies by evaluating the joint probabilities of neighbouring land-cover classes. Higher values indicate more heterogeneous and spatially complex patterns (Nowosad & Stepinski, 2019).

The pd, ed, shdi and shei were calculated following equations (iii), (iv), (v), and (vi) (McGarigal et al., 2015; McGarigal & Marks, 2012) and Ent was calculated following equation (vii) (Nowosad & Stepinski, 2019).

Equation (iii): $pd(ha^{-1}) = \frac{N}{A}(10^4)$

Equation (iv): $ed(mha^{-1}) = \frac{E}{A}(10^4)$

Equation (v): $shdi = -\sum_{i=1}^{m} P_i \, ln(P_i)$

Equation (vi): $shei = \frac{-\sum_{i=1}^{m} P_i ln(P_i)}{ln(m)}$

Equation (vii): $ent = \sum_{i=1}^{m} \sum_{j=1}^{m} P_{ij} \ln(P_{ij})$

Where:

N = total number of patches in the landscape;

E = total length (m) of edge in landscape;

A = total landscape area (m^2);

P_i = proportion of the landscape occupied by patch type (class) i;

P_{ij} = proportion of adjacencies between class i and class j ;

m = number of land-cover classes

The selection of the landscape matrices is grounded in the white stork's behaviour as a central-place forager that depends on open, heterogeneous, and human-modified landscapes rich in accessible prey (Alonso et al., 1991; Zurell et al., 2018). These metrics collectively quantify the fragmentation, diversity, and spatial arrangement of habitats that influence stork foraging efficiency, movement costs, and breeding status.

2.4 Resource Selection Function

The breeding-season habitat preferences of white storks were quantified using a Resource Selection Function (RSF). RSFs estimate the relative probability of a location being selected as a function of environmental covariates by comparing "used" GPS locations with "available" points sampled from the accessible area (A. J. Davis et al., 1994). But prior to the statistical model fit, all continuous variables were tested for multicollinearity.

2.4.1 Multicollinearity Assessment

Multicollinearity was evaluated prior to RSF model selection using a combination of Pearson correlation matrices and Variance Inflation Factors (VIF). VIF is a simple diagnostic used to

detect multicollinearity when two or more predictor variables in a regression model are strongly correlated with each other. A high VIF means the variable is not giving unique information to the model because it overlaps strongly with other predictors. A VIF of <3 is generally acceptable for ecological regression models, indicating that multicollinearity is low enough that each predictor contributes sufficiently distinct information to the analysis. All continuous predictors (NDVI, NDWI, pd, ed, shdi, shei, and ent) were examined for pairwise correlations, which revealed moderate to strong associations among the landscape structure metrics, which was due to the dependency among composition and configuration of landscape indices (Appendix S2).

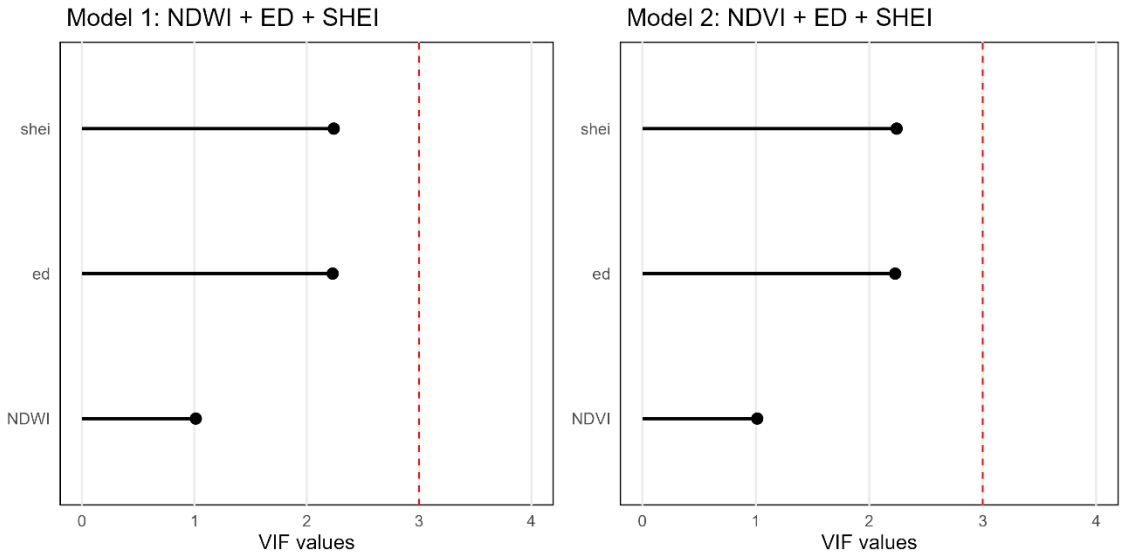

Figure 3. Variance inflation factors (VIF) for two combinations of compatible variables for RSF model. These examples demonstrate the multicollinearity screening applied consistently across the complete set of candidate model structures to verify that predictor combinations satisfied acceptable VIF criteria.

VIF values were computed using simple binomial GLMs fitted with the 'car' package in R (Fox et al., 2001), following recommended practice for mixed-effects RSF modelling (Zuur et al., 2010). The full global model containing all continuous predictors exhibited a very high VIF values (VIF>10), indicating unacceptable collinearity. Therefore, predictors were not combined arbitrarily. Only predictor sets with VIF < 3 were retained for subsequent RSF model evaluation (Figure 2). This ensured that coefficient estimates were numerically stable and biologically interpretable. A total of 40 RSF model combinations were tested before selecting the best model with the lowest AIC value.

2.4.2 Fitting the Statistical Model

Resource selection was modelled using a mixed-effects logistic regression formulation, in which the probability of a location being used was expressed as a function of environmental predictors.

The conditional probability of use was linked to a linear predictor through the logit link function of equation (viii), where the linear predictor was defined as in equation (ix). This formulation allows each individual bird to have its own baseline selection tendency and its own functional response to NDVI or NDWI, where they were included. The RSF itself is expressed as the exponential of the fixed-effect component, as shown in equation (x).

Let $Y_i$denote the binary response for location i:

$$Y_i = \begin{cases} 1, & \text{if location } i \text{ is a used GPS fix} \\ 0, & \text{if location } i \text{ is a randomly sampled available point} \end{cases}$$

Equation (viii): $\text{logit}\{Pr(Y_i = 1)\} = \eta_i$

Equation (ix): $\eta_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + b_{0ji} + b_{1ji} Z_i$

Equation (x): $w(x) = exp\ (\beta_0 + \sum_{k=1}^{p} \beta_k x_{ik})$

where:

x_{ik} = the k-th fixed-effect predictor for observation i

β_k = fixed-effect coefficients

b_{0ij} = random intercept for individual j

b_{1ij} = random slope for NDVI or NDWI (when included)

Z_i = vegetation index (NDVI or NDWI) for observation i

RSF models were fitted using the 'glmmTMB' package with a binomial distribution and logit link in R (Brooks et al., 2017). All continuous variables (NDVI, NDWI, pd, ed, shdi, shei, ent) were standardised to 0 mean and 1 variance prior to modelling to improve numerical stability.

Model selection was performed across 40 candidate structures, representing different ecologically plausible combinations of vegetation indices, LULC classes, and landscape metrics. Competing formulations were ranked using Akaike's Information Criterion (AIC).

Convergence diagnostics were examined for each model, and candidate models exhibiting boundary estimates, singular random-effects structures, and degenerate Hessians were excluded from further consideration.

2.4.3 K-Fold Mean Validation

Model performance was evaluated using a k-fold cross-validation method, as this method has been highly recommended for the validation of RSFs (Boyce et al., 2002a; Wiens et al., 2008). As the dataset comprised only 33 individuals, five folds were used as it provided the best compromise between training sample size, stability of parameter estimates, and the need to retain sufficient individuals in each test fold.

Cross-validation was implemented using group-wise folding, such that all observations from a given bird were assigned to the same fold. This prevented information leakage across training and testing sets and ensured that predictive performance reflected the model's ability to generalise across individuals rather than across repeated locations within the same individual. For each fold i, 80% of the individuals in the dataset were partitioned into the training set, and the remaining 20% of individuals were used as the testing set (Boyce et al., 2002a). Moreover, the predictive performance for each fold was quantified using AUC (Area Under the ROC Curve) values, which were computed using the R package 'pROC' (Robin et al., 2010).

A Receiver Operating Characteristic (ROC) curve is a threshold-independent diagnostic plot that evaluates how well a model distinguishes between used and available locations. It is constructed by plotting the true positive rate (sensitivity) against the false positive rate (1-specificity) across all possible probability thresholds. A model with good discrimination will show a ROC curve that rises steeply toward the upper-left corner, indicating high sensitivity with low false-positive rates. The Area Under the ROC Curve (AUC) summarises this performance into a single metric ranging from 0.5 (no better than random) to 1.0 (perfect discrimination).

2.5 Individual-Based Model

The individual-based model was created from the integration of RSF estimates and ecological rules to address the second goal of the study. The model is described following the standard Overview, Design concepts and Details (ODD) protocol (Grimm et al., 2020)

2.5.1 Purpose

The purpose of this Individual-Based Model (IBM) is to simulate the resource selection, movement, and energy balance of individual white storks during the breeding season (March 1st to August 31st) of specific years. It integrates mixed-effects Resource Selection Functions (RSFs), high-resolution environmental rasters (LULC, NDVI, NDWI, fragmentation indices), and biologically motivated behaviour like site fidelity, nest pull, foraging range, and daily activity periods. In RSF integration, fixed effects as well as individual bird random effects are considered to reflect the individual variability within the bird population

The IBM aims to predict how changes in landscape structure influence individual foraging behaviour and energy accumulation, and whether these energetic differences allow individuals to reach the threshold required for producing an offspring, which is termed successful breeding. In this model, breeding outcomes represent the simulated end-of-season status of each individual, which are: successful breeder (energy accumulation > fixed breeding threshold), failed breeder (non-breeder) (energy accumulation < fixed breeding threshold) or mortality (energy deficit), used as a fitness-relevant measure of landscape effects.

2.5.2 Entities, State Variables, and Scales

i. Entities

The model includes three entities

- BirdAgent: individuals reflecting the 33 white storks
- Environment: raster-based landscape providing habitat attributes and food availability.

ii. State Variables

- BirdAgent: represents the individual white storks. Its behaviour (movement, resource selection, energy balance, breeding decision, and survival) is driven by the variables listed in Table 1.

Table 1. State Variables of the “BirdAgent” Entity, including variable type, unit, description, and associated model functions.

Variable	Type	Unit	Description	Determined by
lon, lat	float	decimal degrees	Current geographical position of the bird	Movement submodel (great-circle step); boundary handling
bearing_deg	float	degrees (0 to 360)	Current movement direction	von Mises distribution turning angle + previous bearing
step_mu_m	float	meters	Mean step length for movement	Behaviour-parameter table (breeding vs nonbreeding) or per-bird override
step_sigma_m	float	meters	SD of step length	Behaviour-parameter table
turn_kappa	float	unitless concentration	Movement persistence; higher values = straighter paths	Behaviour-parameter table
energy_cum	float	energy units	Cumulative net energy over season	Energy budget submodel
day_energy_gain	float	energy units/day	Total food energy collected during the day	Food intake submodel

day_energy_cost	float	energy units/day	Total movement + maintenance cost	Energy cost submodel
alive	bool	unitless	Whether the individual is alive	Mortality rule (DEATH_THRESHOLD)
bred	bool	unitless	Whether the individual qualifies as a breeder	15-day rolling energy window rule
daily_net_history	list of float	energy units	120-day rolling record of daily net energy	Daily energy update submodel
start_lon, start_lat	float	decimal degrees	Nest location used for site fidelity	Movebank nest extraction
step_count	int	count	Number of movement steps taken	Scheduler
behavior mode	string	unitless	Determines movement, foraging radius, nest pull	Breeding rule
foraging_radius_m	float	meters	Soft upper limit for movement away from nest	Behaviour-parameter table

- Environment: represents the spatial landscape and provides covariates for movement decisions and energy availability.

Table 2. State Variables of the "Environment" Entity, including variable type, unit, description, and associated model functions or model input.

Variable	**Type**	**Unit**	**Description**	**Determined by**
lulc	2D array (int)	class code	LULC class at each cell	LULC raster input
ndvi	2D array (float)	normalized index	Vegetation greenness -1 to +1	NDVI raster input
ndwi	2D array (float)	normalized index	Surface water index -1 to +1	NDWI raster input
frag	2D array (float)	index-specific	Fragmentation metrics (ed, pd, shdi, shei, ent)	Optional raster inputs
food_base	2D array (float)	energy units	Baseline cell-level food value derived from LULC + NDVI + NDWI + fragmentation	Food initialisation submodel
food_stock	2D array (float)	energy units	Current food availability that depletes with consumption	Energy gain + 2-day renewal rule
transform, crs	raster metadata	unitless	Spatial transform for lat, lon to row, col	Rasterio package (python version 2.7.16)
last_renewal	datetime	unitless	Timestamp of last food stock reset	Scheduler

iii. Scales

- Temporal resolution (input data): irregular time interval. Some birds have one data point per day, some have less, and some have two GPS points a day.
- Temporal resolution (output data): 30 minutes per step.
- Activity period: 05:00 to 20:00 daily (Sunrise to Sunset)
- Season length: 1 March to 31 August
- Spatial resolution: 30 m
- Simulation duration: distinct single years (multiple years possible, but in this study, single year simulations for 2012, 2018 and 2024 were implemented, considering the computational efficiency and to avoid memory errors)

2.5.3 Process Overview and Scheduling

Each day proceeds in the following sequence:

i. Daily reset (05:00): Breeding birds return to their nest location; nonbreeders continue from their last position.

ii. Daytime steps (05:00–20:00): For each bird, in each 30-minute step:

- Sample step length: At each time step, the movement distance is drawn from a normal distribution centred on the behaviour-specific mean step length, allowing realistic variability in step size.
- Sample turning angle: At each time step, a turning deviation is drawn from a von Mises distribution centred on zero, with concentration parameter κ (kappa) controlling the degree of directional persistence. This deviation is added to the previous bearing to determine the new movement direction.
- Compute proposed new position using great-circle movement.
- Apply foraging radius penalty if distance from nest exceeds behavioural threshold.
- Apply boundary correction if outside raster.
- Sample environmental covariates at the proposed location.
- Compute RSF linear predictor and modify with:

- site fidelity (breeders only): start from the same location every day and every year.
- scheduled nest pull (breeders only): every two hours or 120 minutes the breeding birds visit their nest (Zurell et al., 2018).

- Compute energy intake (proportional to food stock).
- Update the food stock of the visited cell.
- Compute energy costs (flight + baseline physiology).
- Update state variables and position.

iii. End-of-day update (20:00):

- Compute daily net energy and append to daily_net_history.
- Check mortality threshold.
- After ≥15 days, evaluate breeding status using 15-day energy sum.

iv. Food renewal: Every two days, food_stock is reset to food_base.

v. End of season: Export movement tracks and bird summaries (alive, bred, energy balance).

2.5.4 Design Concepts

i. Basic Principles: Movement is governed by RSF-derived habitat suitability, behavioural constraints (site fidelity, nest pull, foraging radius), and energetic feedback. Energy acquisition depends on food availability, which depletes locally. Breeding emerges from sustained energetic surplus.

ii. Emergence: The model generates emergent patterns in spatial use, energy trajectories, survival, and breeding outcomes. Breeders and nonbreeders differ in movement extent due to imposed behavioural rules and energy constraints. For breeders, the highest foraging radius is set to be 28.1km, and for non-breeders it's 48.2km (Gilbert et al., 2016).

iii. Adaptation: Agents do not adapt or learn. Although breeding status modifies behaviour, the decision rule is fixed and energy-based, which also depends on the landscape variables, not adaptive learning.

iv. Objectives: Agents implicitly maximise daily net energy through their movement choice. Fitness outcomes (breeding and survival) follow from this habitat selection process.

v. Sensing: Agents explicitly sense habitat attributes and food availability at their proposed next cell.

vi. Stochasticity: Stochasticity enters via step-length sampling, turning-angle sampling, boundary reflections, food encounters, and resulting energy balances. These processes create variation among individuals.

vii. Interaction: There are no direct behavioural interactions between individuals. However, birds interact indirectly via the shared food environment, as foraging at a cell reduces its food level until the next renewal, influencing the conditions encountered by other individuals. This constitutes indirect exploitation competition.

viii. Learning: Not included.

ix. Prediction: Not applicable.

x. Collective: None.

xi. Observation: Outputs include per-bird trajectories, step length, timeline, daily and cumulative energy, survival, breeding status and also yearly summaries. Additional figures display combined tracks for breeders and nonbreeders for each year.

2.5.5 Initialisation

i. Environmental layers (LULC, NDVI, NDWI) are loaded for the specific year. This study did not use fragmentation layers as they were not available.

ii. All layers are reprojected to LULC resolution.

iii. Baseline food availability food_base is computed from LULC, NDVI, NDWI, and fragmentation indices. Constant energy scores are allocated for each LULC class, but for food availability calculation, NDVI gain, NDWI gain and fragmentation indices gain are also included in the calculation.

iv. food_stock is initialised as a copy of food_base.

v. Nest coordinates are extracted from the first timestamps in the Movebank dataset.

vi. Fixed and random RSF effect tables are merged to create per-bird RSF coefficient vectors.

vii. Each BirdAgent is placed at its nest with zero accumulated energy, random initial bearing, and non-breeding state.

2.5.6 Input Data

i. Annual LULC rasters

ii. Annual NDVI and NDWI rasters.

iii. Fragmentation rasters (optional).

iv. Fixed-effect RSF coefficient CSV.

v. Random-effect RSF coefficient CSV.

vi. Movebank data file containing bird IDs and nest coordinates.

2.5.7 Sub-models

i. Movement Submodel

Movement is executed for each bird at every 30-minute daytime interval (05:00–20:00). Movement consists of (i) sampling a step length, (ii) sampling a turning deviation, (iii) computing a proposed new position using a great-circle update, (iv) applying behavioural movement constraints, and (v) handling boundary collisions.

- Step length sampling: Step distance is drawn from a normal distribution following the equation (xi), whose parameters depend on the behavioural status of the agent (breeding or non-breeding)

Equation (xi): $s_t \sim N(\mu_s, \sigma_s^2)$

where:

s_t= step length at time t(m)

μ_s= behaviour-specific mean step length (m)

σ_s= behaviour-specific SD (m)

Negative values are truncated to zero.

- Turning angle sampling: Directional change follows a von Mises distribution centred on 0 with a concentration parameter κ:

Equation (xii): $\Delta\theta_t \sim \text{vonMises}(0, \kappa)$

where:

κ: directional persistence (higher means straighter movement), from turn_kappa.

- Bearing update: The bearing update applies a von Mises-distributed angular deviation ($\Delta\theta_t$ from equation *xii*) to the previous movement direction, creating a correlated random walk with behaviour-dependent directional persistence following the equation *(xiii)*.

Equation (xiii): $\theta_t = (\theta_{t-1} + \Delta\theta_t) mod 2\pi$

where:

θ_t= bearing at time *t* (radians)

θ_{t-1}= previous bearing

- Great-circle displacement: The model uses a spherical Earth approximation to compute the proposed coordinates. New coordinates (ϕ_t, λ_t) are computed using equations *(xiv)* and *(xv)*.

Equation (xiv):

$$\phi_t = arcsin(sin\ \phi_{t-1} cos\ \frac{s_t}{R} + cos\ \phi_{t-1} sin\ \frac{s_t}{R} cos\ \theta_t)$$

Equation (xv):

$$\lambda_t = \lambda_{t-1} + arctan2(sin\theta_t sin\frac{s_t}{R} cos\phi_{t-1}, cos\frac{s_t}{R} - sin\phi_{t-1} sin\phi_t$$

where:

ϕ_t, λ_t=latitude and longitude (radians)

$R = 6{,}371{,}000$: Earth radius (m)

s_t= step length

θ_t= updated bearing

- Soft foraging-radius constraint: The model applies a continuous exponential penalty only when the distance exceeds the behaviour-specific home-range radius. The position is recomputed using equation *(xvi)*.

Equation (xvi): $s_t^* = s_t exp\ [-\gamma(d_{\text{nest}} - r_{max})]$if $d_{\text{nest}} > r_{max}$

where:

s_t^*= adjusted step length

d_{nest}= distance from proposed point to nest (m)

r_{max}= behavioural max-foraging radius (m)

$\gamma = 0.001$: decay coefficient

- Boundary correction: If the proposed coordinates fall outside raster bounds, bearing is reversed:

Equation (xvii): $\theta_t = \theta_{t-1} + \pi \pm U(-\delta, \delta)$

where:

θ_t= Updated bearing time at time t

π = 180° direction reversal

$U(-\delta, \delta)$ = uniform random jitter applied to avoid repeated collisions

$\delta = 30°$(angular jitter)

The model attempts up to three such corrections. If all three redirected bearings still result in coordinates outside the raster bounds, the agent remains in its current cell for that time step and resumes normal movement evaluation on the next step.

ii. RSFModel

The RSF linear predictor LP_t represents the relative habitat suitability of the proposed step. It is a log-relative selection weight: locations with higher values are predicted by the fitted RSF model to be more likely to be selected by that individual in real data.

Each proposed location is evaluated through an RSF-derived linear predictor using the per-bird coefficient vector (fixed effects + individual random effects) obtained from the mixed-effects RSF. The habitat quality of a candidate cell at time t is calculated following equation (xviii).

Equation (xviii): $LP_t = \beta_0 + \beta_{\text{NDWI}} \cdot NDWI_t + \beta_{\text{LULC}[c_t]} + \sum_k \beta_k \, x_{k,t}$

where:

LP_t= linear predictor at time t

$NDWI_t$= NDWI value at the cell

LULC$[c_t]$= LULC class at the cell

$x_{k,t}$= additional covariates (fragmentation metrics)

β= RSF coefficients (fixed + random effects)

Missing covariates contribute zero (fragmentation covariates for this study contribute zero).

iii. Behavioural Modifiers

The RSF score is adjusted by site-fidelity and nest-pull behaviours.

- Site fidelity (breeders only): This penalises long distances from the nest. In this model, the $\delta_h$is fixed for all the birds, but site fidelity weight (w_{home}) is calculated for the breeders following the equation *(xix),* where for non-breeders $w_{\text{home}} = 1$, ensuring that $\delta_h$has no effect on their movement behaviour.

Equation (xix): $w_{\text{home}} = exp\,(-\delta_h d_{\text{nest}})$

where:

w_{home}= site fidelity weight

δ_h= site-fidelity decay coefficient

d_{nest}= distance from proposed location to nest

- Nest pull (breeders only, every 120 minutes): Activated only when simulation time satisfies:

Equation (xx): $(t_{\text{minutes}} mod 120) = 0$

$t_{\text{minutes}} = 60 \times \text{hour}(t) + \text{minute}(t)$. This is a simulation time converter. It standardises time into a single numerical scale so that periodic events (eg. nest pull every 2 hours) can be evaluated cleanly and consistently during the simulation.

The additional directional reward is:

Equation (xxi): $w_{\text{pull}} = exp\ (-\alpha \cdot d_{\text{nest}})$

where:

w_{pull}= nest-return weight

α= nest-pull strength in BEHAVIOR_PARAMS

This reward is a mathematical way to represent site fidelity. Breeders tend to return to the nest, so the model biases movement slightly in that direction.

- Final RSF score

Equation (xxii):

$$Score_t = LP_t + log\ (w_{\text{home}} + 10^{-12}) + log\ (w_{\text{pull}} + 10^{-12})$$

where:

$Score_t$=Final RSF score

LP_t = habitat-quality score from the RSF linear predictor

w_{home}= weight for staying within the home range (site fidelity at larger scale)

w_{pull}= weight for nest-directed movement (fine-scale nest fidelity). The 10^{-12}term prevents taking the log of zero

This final score is the composite decision value that the agent uses to evaluate whether the proposed step is acceptable. It merges: habitat preference (RSFmestimates), home-range fidelity (tendency not to roam infinitely) and nest attraction (for breeders during the season).

iv. Food Intake Submodel

Food base calculation: the food base is computed following the equation *(xxiii)*.

For each cell i:

Equation (xxiii): $F_{cell(i)} = E_{\mathrm{LULC}(i)} + 0.3 \cdot NDVI_i + 0.6 \cdot NDWI_i$

where:

$F_{cell(i)}$= baseline food availability

$E_{\mathrm{LULC}(i)}$= LULC class energy value

$NDVI_i$= NDVI at cell *i*

$NDWI_i$= NDWI st cell *i*

This formulation combines a LULC based energy component with two vegetation-derived indicators that approximate fine-scale foraging conditions. The weighted NDVI and NDWI terms (0.3 and 0.6) adjust food availability based on vegetation greenness and surface moisture, allowing wetter or greener cells to contribute proportionally more potential food than dry or bare areas. The LULC energy value ensures that different habitat types (e.g., cropland, grassland) start with distinct base resource levels consistent with known stork foraging preferences. The $E_{\mathrm{LULC}(i)}$ is assigned for each LULC class separately.

After energy extraction, the food stock in that cell is reduced accordingly, as described in Equation (*xxiv*) and at each movement step, the bird obtains energy from the current cell based on the available food stock in that location. The energy gained at time *t* is calculated as equation (*xxv*).

Equation (xxiv): $F_{\mathrm{cell}} = max\ (0, F_{\mathrm{cell}} - G_t)$

Equation (xxv): $G_t = \eta F_{\mathrm{cell}}$

where:

G_t= energy gained at time *t*

η= food extraction proportion (fixed at 0.25 for model simplification)

F_{cell}= food stock or food score of a certain cell

To avoid long-term depletion and to reflect natural replenishment processes in agricultural landscapes, all food stocks are reset every two days. This periodic reset prevents irreversible local starvation in the model and maintains ecological plausibility while keeping the model tractable. This ensures that food availability cannot drop below zero.

In this sub-model, first, each cell $F_{\text{cell}(i)}$ is assigned a baseline food score. During the simulation, the current food stock $F_{\text{cell}}$is updated as energy $G_t$is extracted by the bird

v. Energy Cost Submodel

At each movement step, the total energetic cost consists of three components: a distance-dependent flight cost, a physiological maintenance cost, and a foraging cost. The combined cost at time t is

Equation (xxvi): $C_t = (c_d \cdot d_t) + c_p + c_f$

where:

c_d = distance cost scaling factor (0.0002)

d_t= step length (distance travelled in the current step)

c_p= physiological cost per step

c_f= foraging cost per step

- Physiological cost (c_p): This represents the baseline metabolic energy required to maintain bodily functions and is allocated evenly across all daily time steps:

Equation (xxvii): $c_p = \frac{E_{\text{day}}}{n_{\text{steps}}} \cdot \frac{\rho_{\text{phys}}}{\rho_{\text{tot}}}$

where:

c_p= physiological cost per movement step

E_{day}= baseline daily energy requirement (fixed at 10)

n_{steps}= number of time steps per day (48 steps per day for 30-min step recording)

ρ_{phys}= physiological energy fraction = 0.75 (as physiological activity requires 75% of total energy gain)

ρ_{tot}= total energy fraction = 1.0 (sum of all energy requirements)

This ensures that 75% of daily required energy is reserved for physiological maintenance.

- Foraging cost:

Equation(xxviii): $c_f = \frac{E_{\text{day}}}{n_{\text{steps}}} \cdot \frac{\rho_{\text{for}}}{\rho_{\text{tot}}}$

$\rho_{\text{for}} = 0.10$(foraging fraction)

Together, physiological (75%) and foraging (10%) costs represent fixed components of the daily energetic budget, with the remaining fraction implicitly covering other behavioural expenditures.

- Net energy: The net energy gained during a movement step is calculated with equation (*xxix*).

Equation (xxix): $N_t = G_t - C_t$

where:

G_t= energy gained from the cell at step *t*

C_t= total energetic cost at step *t*

- Cumulative energy: The cumulative energy balance evolves as equation (*xxx*).

Equation (xxx): $E_t = E_{t-1} + N_t$

where:

E_t = Cumulative energy

E_{t-1}= Cumulative energy from the previous step

N_t = Step level net gain

Cumulative energy determines whether an individual reaches the threshold required for breeding initiation, fails to meet it, or drops below 60% and dies

vi. Daily Update and Mortality Submodel

Cumulative energy reserves E_t are updated at each time step by adding net energy N_t. An individual dies when its reserves fall below a fixed threshold. To evaluate daily energetic performance, the model aggregates step-level net energy into a daily total N_{day}.

Daily net energy:

Equation (xxxi): $N_{\text{day}} = \sum_{\text{day}} N_t$

where:

N_{day}= Total net energy per gained or lost per day

A rolling history of the most recent 120 days is stored to avoid excessive memory usage.

Mortality: An individual dies when its cumulative energy reserve falls below a critical threshold.

Equation (xxxii): $E_t < \varepsilon_d \times (-E_{day})$

where:

ε_d=Death threshold proportion = 0.6

If *(xxxii)* is true, then the agent doesn't record any further movement. Which means if the total energy falls under 60% then the bird dies and does not take part in the rest of the simulation.

vii. Breeding Submodel

Breeding in the model is triggered when an individual maintains a sustained positive energy balance over a rolling 15-day window. An agent qualifies as a breeder when the cumulative net energy gained across the window exceeds a minimum surplus threshold like in equation (*xxxiii*).

Equation (xxxiii): $\sum_{i=W-14}^{W} N_t \geq \varepsilon_b E_{day}$W

where:

N_t= net energy at each time step

W= length of the breeding assessment window (15 days)

$\varepsilon_b = 0.30$= minimum surplus proportion required for breeding

E_{day}= daily energy requirement (10 units)

This condition ensures that only individuals that consistently achieve a net positive energy surplus of at least 30% above the baseline requirement over 15 days can breed. The threshold reflects the biological expectation that breeding requires sustained energetic stability rather

than short-term spikes in food acquisition. Individuals failing to meet this surplus are classified as non-breeders.

2.5.8 Further Analysis

The IBM was simulated for the years 2012, 2018 and 2024. The changes in LULC classes were checked using the R ‘sf’ package (Pebesma, 2016).

The IBM outputs were then downsampled to reduce the temporal resolution in order to avoid the overfitting of a certain LULC type. The downsampled data were then fitted into another RSF to check the variability of resource use in different landscape scenarios.

Chapter 3.Results

3.1 General Description of Environmental Covariates

All environmental covariates were examined using descriptive statistics and exploratory visualisations to understand their distributions, ranges, and potential redundancies. This step ensures that predictors represent meaningful ecological gradients and do not introduce structural issues such as extreme skewness or multicollinearity. All continuous variables were standardised prior to RSF fitting.

3.1.1 Vegetation Index and Wetness Index Covariates

The NDVI values ranged from -0.45 to 0.55, with a mean of 0.12 across the study area, indicating substantial spatial heterogeneity in vegetation. The distribution was slightly left-skewed (Figure 4a), with most observations falling within the low to moderate NDVI range. This pattern suggests that the landscape was primarily composed of open or semi-open habitats rather than densely vegetated areas.

The NDWI ranged from -0.49 to 0.44, with a mean value of -0.13, indicating an overall dry soil. The data distribution was right-skewed, with the upper tail extending to positive NDWI reflects the presence of moist fields and shallow-water environments. The distribution was clearly bimodal (Figure 4b).

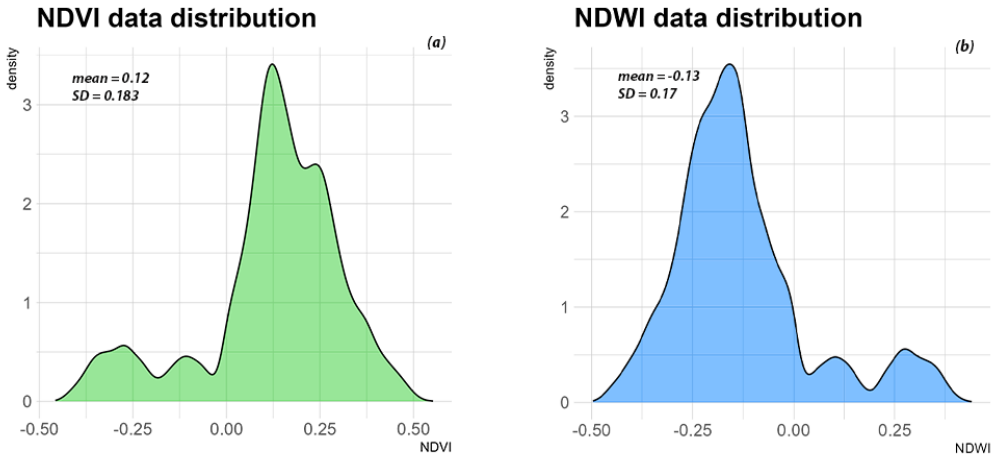

Figure 4. Data distribution of NDVI (a) and NDWI (b) from the used-available points.

3.1.2 Landscape Structure Covariates

shdi

The Shannon Diversity Index (shdi) values ranged from 0 to 2.27, with a mean of 0.87, indicating a landscape with moderate compositional heterogeneity. The distribution was unimodal and right-skewed (Figure 5a), showing that most locations contained a balanced mixture of a few dominant LULC types, whereas highly simplified or highly heterogeneous patch mosaics were less frequent.

shei

The Shannon Evenness Index (shei) values ranged from 0 to 1, with a mean of 0.44. The distribution was unimodal and slightly left-skewed (Figure 5b), and most observations concentrated between 0.40 and 0.60, indicating that the landscape generally exhibited moderate evenness in the relative proportions of land-cover classes. It suggests that the landscapes had a mix of dominant classes, but highly uneven landscapes dominated by a single land-cover type were uncommon.

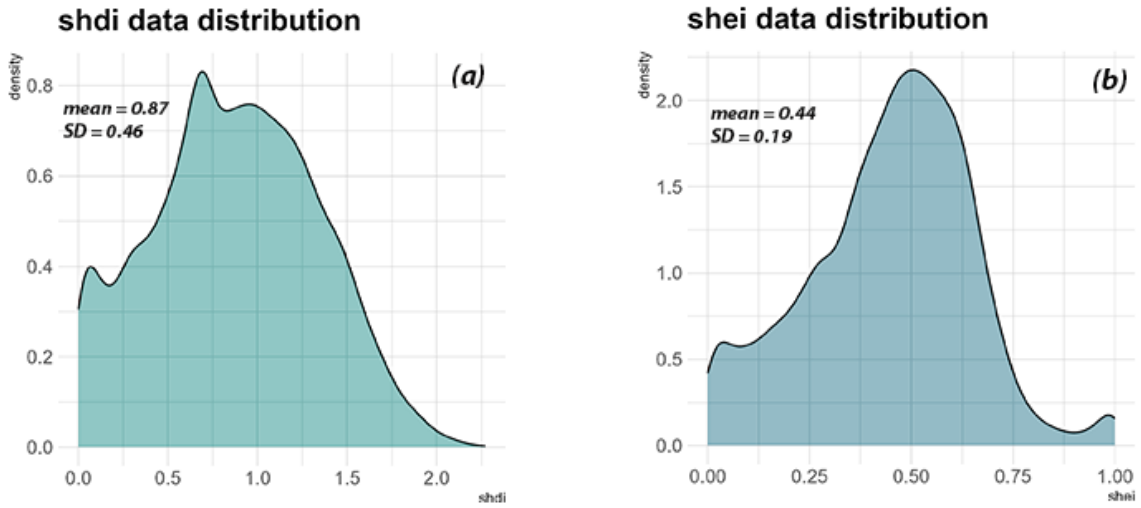

Figure 5. Data distribution of shdi (a) and shei (b) from the used-available points.

3.1.3 Landscape Fragmentation Metrics Covariates

ed

The Edge Density (ed) exhibited a heavily right-skewed distribution (Figure 6a), with the highest value of 3.75×10^7 m/ha with a mean of 1.18×10^7 m/ha. Higher values correspond to strongly fragmented areas with dense land-cover transitions.

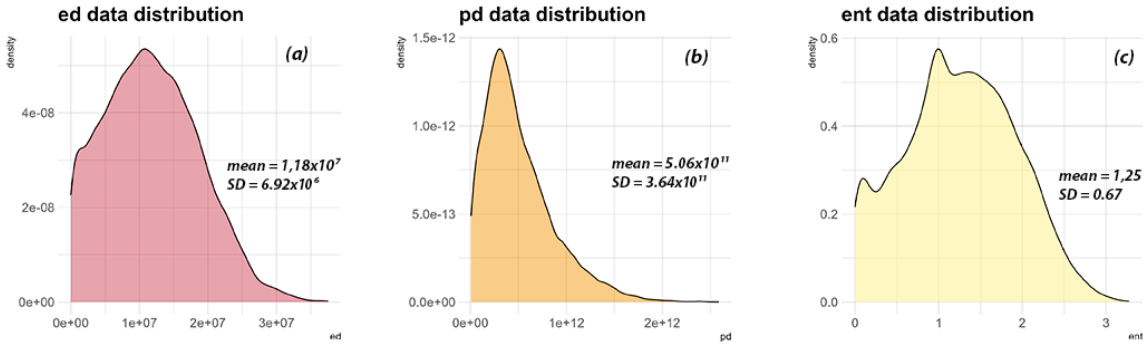

Figure 6. Distribution of landscape fragmentation metrices ed (a), pd (b) and ent (c).

pd

Patch Density (pd) also showed a pronounced right skew (Figure 6b), reflecting a landscape containing many small and irregularly distributed patches. High pd values indicate fine-grained mosaics typical of intensively managed agricultural systems. The locations exhibited a mean pd of 5.06×10^{11} patches ha^{-1}, ranging maximum to 2.59×10^{12} patches ha^{-1}. Together with ed, pd quantifies the spatial structure underlying fragmentation intensity.

ent

The Entropy (ent) showed a mean value of 1.25, ranging from 0 to 3.27. This wide range indicates that the sampled locations captured both simple, low-complexity patches and distinctly heterogeneous landscape configurations. The distribution was unimodal with moderate spread (Figure 6c), showing no extreme skew but a clear density concentration around the mid-entropy range.

The overall descriptive statistics of all environmental covariates are summarised in Table 1.

Table 3. Descriptive statistics of all covariates including their mean, median, standard deviation (SD), interquartile range (IQR) and min-max values.

variable	mean	median	SD	IQR	min	max
NDVI	0.123363	0.144668	0.183304	0.178307	-0.45748	0.551794
NDWI	-0.13077	-0.16045	0.169263	0.16534	-0.49684	0.440805
shei	0.445907	0.467684	0.19992	0.266379	0	0.999995
shdi	0.871588	0.869225	0.461036	0.668262	0	2.272154
ed (m/ha)	$1.18x10^{7}$	$1.14x10^{7}$	$6.92x10^{6}$	$1.0x10^{7}$	0	$3.75x10^{7}$
pd (ha^{-1})	$5.06x10^{11}$	$4.22x10^{11}$	$3.64x10^{11}$	$4.45x10^{11}$	$8.54x10^{9}$	$2.59x10^{12}$
ent	1.25288	1.249728	0.666163	0.965846	0	3.274728

3.1.4 LULC Classes

The overall frequency distribution of LULC classes was strongly biased, with a clear dominance of agricultural habitats (Figure 7). Herbaceous cover cropland was the most abundant class in the landscape, followed by rainfed cropland. These classes formed the main background matrix through which storks moved and in which the available points were generated. Other cropland categories, such as irrigated cropland, occurred less frequently, whereas the forest classes were represented only by small proportions of the overall dataset, except for the open evergreen broadleaved forest (Open EG BrL Forest). Non-vegetated classes like impervious surfaces, wetlands, marshes, grasslands, and water bodies were similarly sparse.

This distribution confirms that the study landscape is dominated by the open agricultural systems, with all other land-cover types occurring only as minor or patchy components of the broader environment.

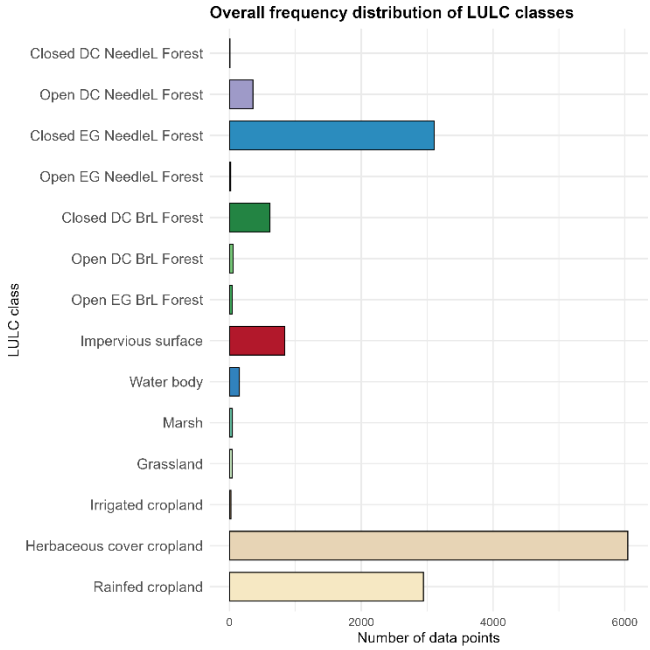

Figure 7. Overall frequency distribution of LULC classes across the landscape. LULC class names were abbreviated (DC=Deciduous, EG=Evergreen, NeedleL=Needle-leaved, BrL=Broad-leaved).

3.1.5 Used Steps Distribution in Environmental Covariates

The used steps points per individual spanned a mean NDVI range of -0.2 to 0.3, with most birds clustering between 0.1 and 0.3 (Figure 8a). This concentration indicates that individuals predominantly used areas with low to moderate vegetation greenness. In contrast, mean NDWI values at used steps ranged from -0.3 to -0.1 for the majority of individuals (Figure 8b), showing a clear tendency toward relatively dry surfaces. Only a small subset of birds was associated with locations exhibiting moderate NDWI values.

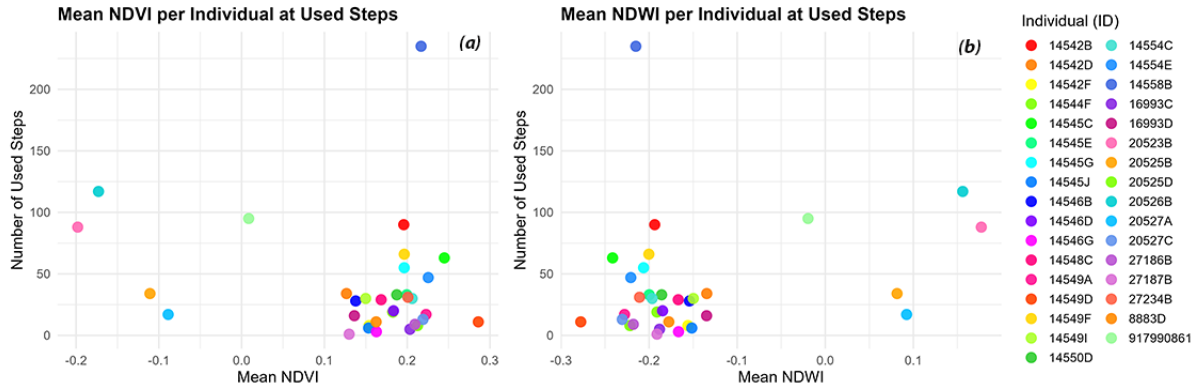

Figure 8. Distribution of mean NDVI (a) and NDWI (b) across used steps per individual. Each individual is identified by their unique ID number in the legend on the right.

Furthermore, mean shei values at used steps were widely dispersed, spanning roughly 0.35 to 0.58, with no strong clustering around a single landscape-evenness level. This spread indicates that storks used areas with varying degrees of patch evenness, from moderately uneven to relatively balanced landscapes, without a consistent preference across individuals (Figure 9a).

Similarly, mean shdi values ranged broadly from 0.7 to 1.2, revealing substantial individual-level variation in the compositional diversity of used locations. Some birds foraged in relatively simple patch mosaics, while others used areas with more complex and diverse land-cover mixtures (Figure 9b).

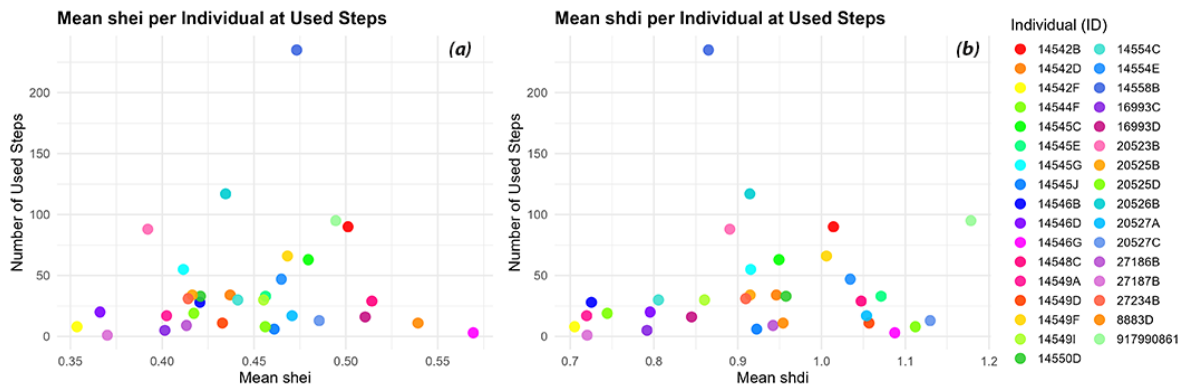

Figure 9. Distribution of mean shei (a) and shdi (b) across used steps per individual. Each individual is identified by their unique ID number in the legend on the right.

Fragmentation metrics demonstrated heterogeneous patterns among individuals. ed and pd exhibited weak clustering around intermediate values, with most used steps occurring in landscapes of moderate edge and patch density (Figure 10a & 10b). However, the wide overall spread and presence of outliers indicate substantial individual variation rather than a strong, uniform selection signal. The ent showed no appreciable clustering, with individuals distributed broadly across the full heterogeneity gradient (Figure 10c).

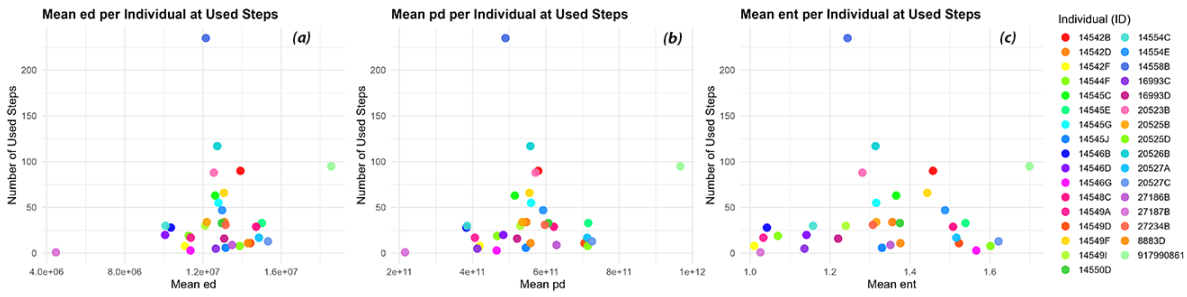

Figure 10. Distribution of mean ed (a), pd (b) and ent (c) across used steps per individual. Each individual is identified by their unique ID number in the legend on the right.

Used steps were disproportionately concentrated in herbaceous cover cropland, indicating that most individuals foraged primarily within these open agricultural habitats. Rainfed cropland also accounted for a substantial share of the used locations for many birds. A small number of individuals showed occasional use of open evergreen broadleaf forest, but these occurrences were infrequent compared to the dominant cropland classes (Figure 11).

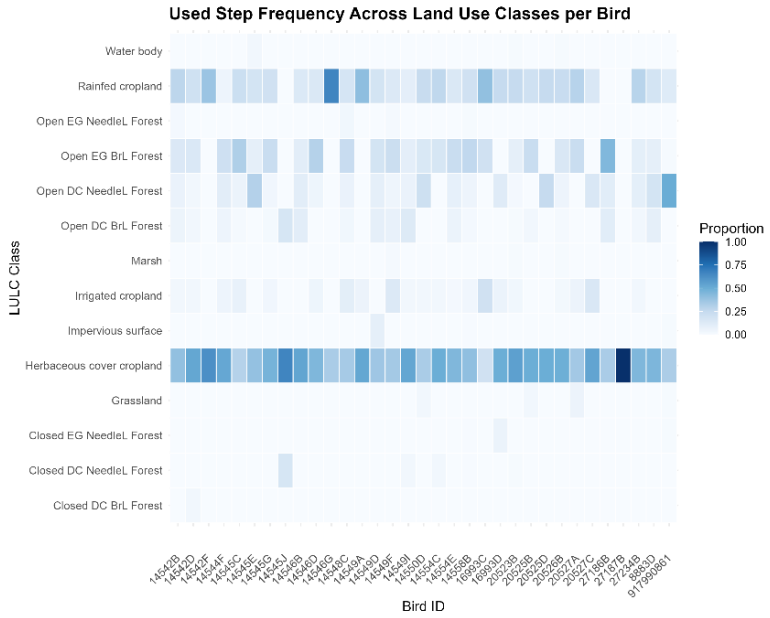

Figure 11. Used steps frequency (%) per individual across LULC classes. LULC class names were abbreviated (DC=Deciduous, EG=Evergreen, NeedleL=Needle-leaved, BrL=Broad-leaved). The individuals are identified with their unique IDs on the x-axis.

3.2 Resource Selection by White Storks

Resource selection patterns showed a clear association between the individual movement decisions and the environmental conditions available along their trajectories. Across individuals, used steps consistently aligned with specific habitat types and landscape characteristics, indicating that storks do not use the landscape randomly but respond to measurable gradients in land-cover, vegetation structure, and spatial configuration. These patterns form the basis for the subsequent assessment of model performance and covariate-specific effects.

3.2.1 Best Model to Describe the Resource Selection

Model selection was initially guided by AIC, but the model with the second-lowest AIC was retained because it provided the most ecologically coherent set of effects. The structure of the model with lowest AIC was: Used ~ NDWI +LULC_class + ed_scaled +shei_scaled+ (1 +NDVI | bird_id) (AIC = 8490.3). However, the NDWI distribution across the study area

(Figure 4b) showed consistently low moisture conditions, meaning the birds had limited access to higher-NDWI patches. Under such limited availability, the probability of use is strongly shaped by small differences in moisture, making the ecologically structured model more appropriate despite its slightly higher AIC. So, the model: Used ~ NDVI +LULC_class + ed_scaled +shei_scaled+ (1 +NDVI | bird_id) was selected as the best model (AIC = 8494.1).

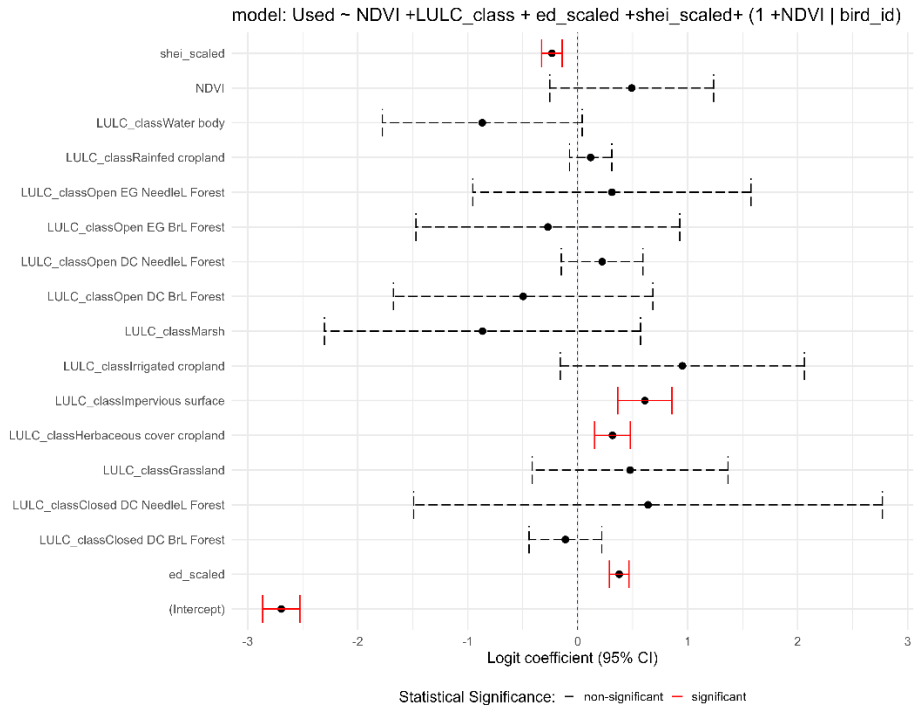

Figure 12. The best selected RSF model containing the combination of NDVI, LULC classes, ED and PD. The significant coefficients are represented with red solid lines and non-significant coefficients are represented with black dashed lines. LULC class names were abbreviated (DC=Deciduous, EG=Evergreen, NeedleL=Needle-leaved, BrL=Broad-leaved).

The model coefficients show a preference for a distinct habitat type (Figure 12). Although the vegetation index NDVI shows a positive influence on habitat selection, it is not significant. The choice of habitat selection seems to be influenced mostly by LULC classes, landscape configuration and fragmentation metrics. The model also reveals a significant positive coefficient for the two LULC classes, herbaceous cropland and impervious surface, which indicates the stork's preference for these habitats.

Moreover, the ed had a significant positive coefficient, implying that birds favoured landscapes with a higher density of boundaries, which likely provide diverse foraging opportunities. In contrast, shei was significantly negative, indicating avoidance of highly heterogeneous, evenly

mixed mosaics. This combination suggests that storks favour fragmented edges within dominant habitat types but not fully interspersed multi-habitat mosaics.

3.2.1.1 K-fold Mean Model Validation

ROC curves generated from five-fold cross-validation showing the predictive performance of the mixed-effects RSF model using NDVI, LULC class, ed, and shei. Three folds (fold 1, 4 and 5) exhibit moderate distinction (AUC > 0.60), indicating detectable habitat selection patterns in those subsets, while the remaining two folds lie close to the 1:1 diagonal, reflecting weak or random-level predictive ability. The variability among folds highlights strong behavioural and environmental heterogeneity within the dataset, resulting in a low overall mean AUC (= 0.58) despite moderate predictive performance in several folds.

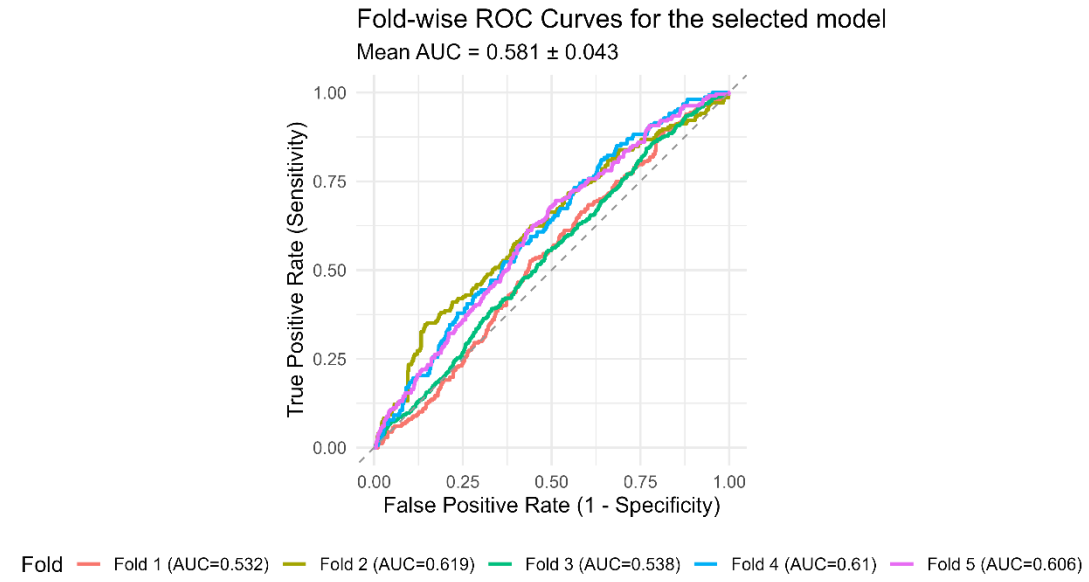

Figure 13. Fivefold-wise ROC curves for the selected habitat selection model. The fold 2, 4 and 5 are indicating slightly better predictive performance than the random model.

3.3 Resource Use of Individually Simulated Birds

3.3.1 Yearly Simulation Summary

The IBM outputs showed clear change in the numbers of successful breeders and non-breeders across the simulated years (Figure 14), indicating that reproductive performance and, therefore, overall fitness were sensitive to landscape changes over time. The number of breeders significantly increased in the year 2018 in comparison to the year 2012. The year 2024 had a slight reduction in the number of breeders than 2018 but it was still substantially higher than the year 2012.

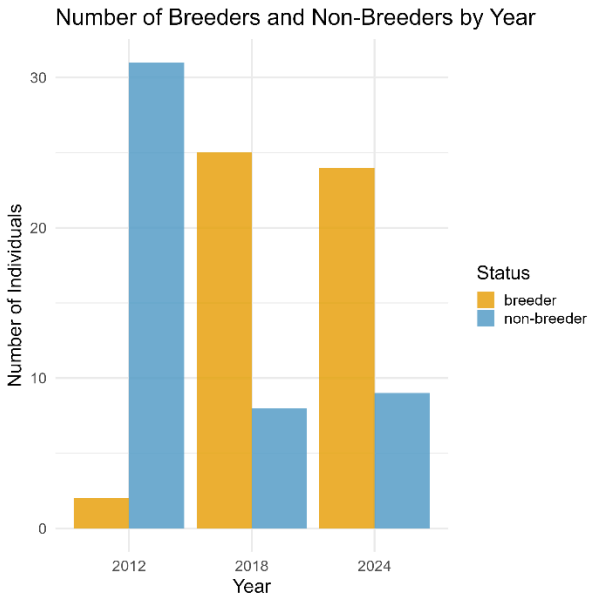

Figure 14. The number of breeders and non-breeders over the simulated years: 2012, 2018 and 2024.

The comparison of LULC classes across the simulation years revealed a consistent directional pattern of landscape change. Between 2012 and 2018 (Figure 15a), several classes showed noticeable increases, including rainfed cropland (10), herbaceous cropland (11), closed deciduous broad-leaved forest (62), closed evergreen needle-leaved forest (72), impervious surfaces (190), and water bodies (210). Over the same period, irrigated cropland (20), closed evergreen broad-leaved forest (52), open evergreen needle-leaved forest (71), lichens and mosses (140), flooded flats (183), and bare areas (200) showed declines. A similar pattern was observed between 2018 and 2024 (Figure 15b). The same set of LULC classes showed increases (rainfed cropland, herbaceous cropland, closed deciduous and evergreen needle-leaved forests, impervious surfaces, and water bodies), while the same classes continued to decline, alongside additional decreases in grassland, sparse vegetation, and tidal flats. Although the magnitude of change differed between periods, the direction of change remained consistent, indicating a stable trajectory of land-cover transitions in the study region.

Although the directional patterns of land-cover change were similar in both periods, the magnitude and ecological consequences differed. The 2012-2018 interval showed strong gains in the two positively selected habitats, herbaceous cropland and impervious surfaces, resulting in higher energetic intake and the highest number of breeders in 2018. In contrast, the 2018-2024 period exhibited weaker increases in these preferred habitats and larger declines in potential foraging habitats like grassland, sparse vegetation and shallow wet areas which

together reduced foraging opportunities and movement efficiency. As a result, fewer individuals reached the energetic threshold required for breeding, leading to a slight decline in breeder numbers in 2024 (Figure 14).

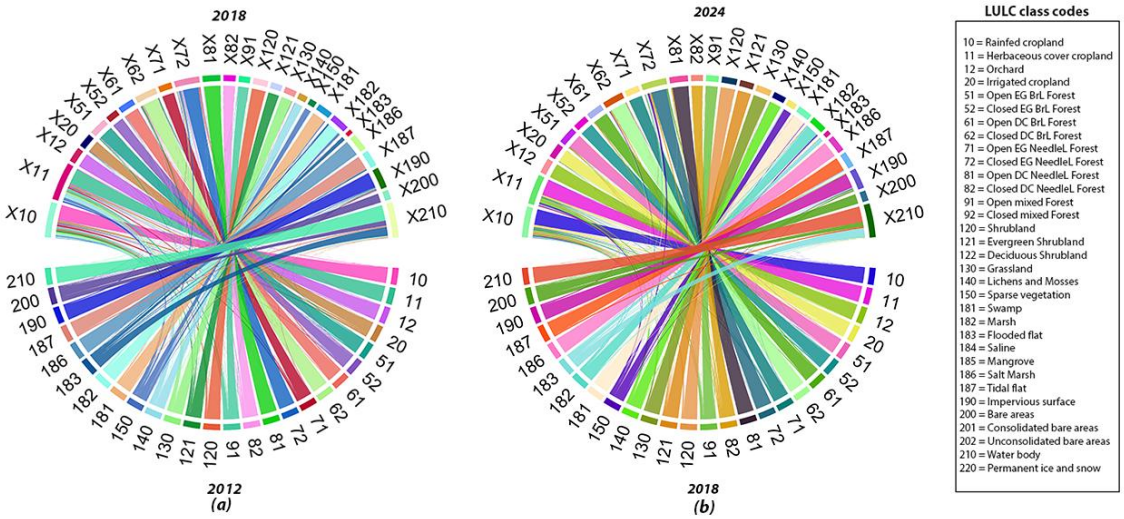

Figure 15. The LULC class conversions from 2012 to 2018 (a) and 2018 to 2024 (b). The LULC classes are expressed as the numeric value of their cover class. LULC class names were abbreviated (DC=Deciduous, EG=Evergreen, NeedleL=Needle-leaved, BrL=Broad-leaved).

The movement trajectories in 2012 revealed marked differences between breeders and non-breeders (Figure 16). Breeding individuals showed spatially constrained movement patterns centred around a relatively small core area, reflecting nesting-site fidelity and the limited foraging radius typical of birds provisioning nests (Figure 16a). In contrast, non-breeders exhibited far more dispersed and wide-ranging movements across the landscape, with no single dominant activity centre (Figure 16b). This broad spatial spread indicates the absence of reproductive constraints and suggests higher exploratory behaviour among non-breeding individuals. This behavioural divergence remained consistent across all simulated years among breeders and non-breeders (Appendix figures S2, S3).

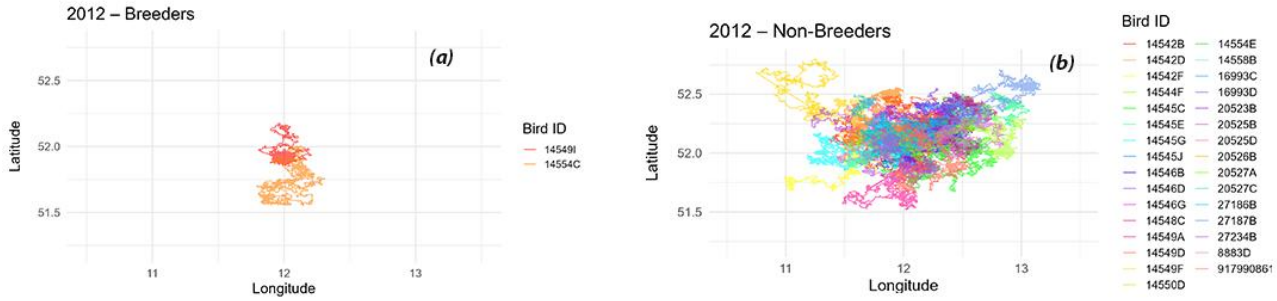

Figure 16. The movement of breeders (a) and non-breeders (b) across the landscape, with Longitude on the x-axis and Latitude on the y-axis. Each individual is identified by their unique ID number in the legend on the right.

3.3.2 Best Model to Describe the Resource Selection for Simulated Birds

The model selection was based on the lowest AIC. The lowest AIC model supports the same combination of variables from the selected RSF model from the original bird locations.

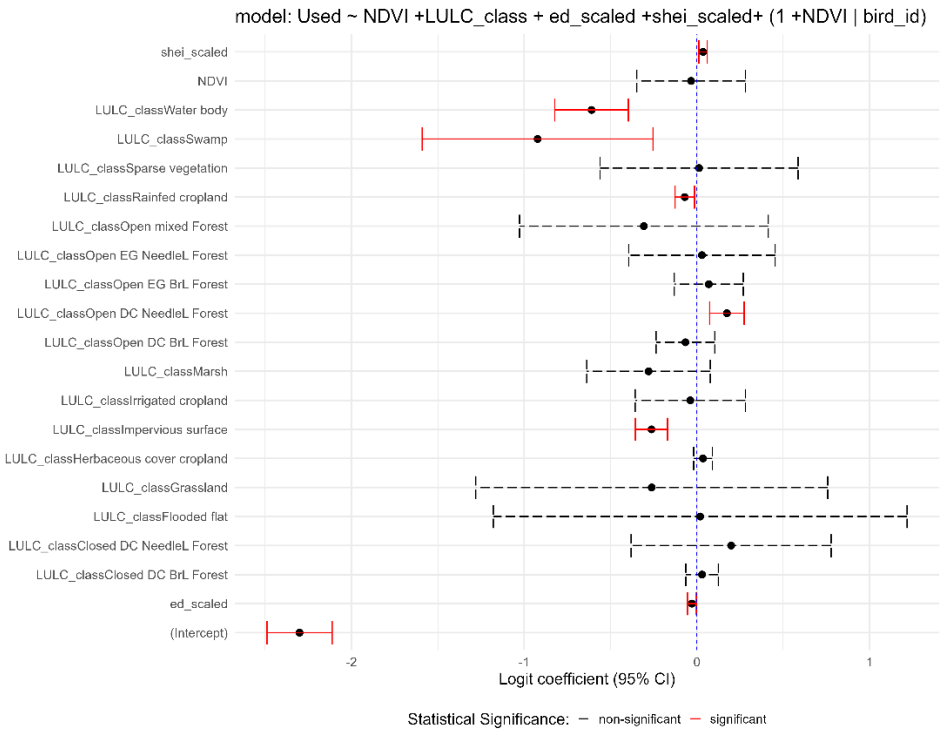

Figure 17. The best model to describe the resource selection of simulated birds containing the combination of NDVI, LULC classes, ed and shei. The significant coefficients are represented with red solid lines and non-significant coefficients are represented with black dashed lines. LULC class names were abbreviated (DC=Deciduous, EG=Evergreen, NeedleL=Needle-leaved, BrL=Broad-leaved).

The model (Figure 17) indicated no measurable effect of NDVI on habitat selection, with estimates centred around zero. In contrast, LULC classes showed clearer signals. Water bodies and swamps exhibited strong negative coefficients, indicating consistent avoidance. Rainfed cropland and impervious surfaces also showed slightly negative but statistically significant effects. Only the open deciduous needle-leaved forest class displayed a positive and significant coefficient, suggesting limited but detectable selection for this habitat type.

Landscape configuration metrics contributed weakly to selection patterns. The ed showed a small negative but statistically significant effect that nearly overlapped the neutral line, while the shei produced a similarly marginal but statistically significant estimate close to zero. Overall, LULC categories explained more variation in simulated habitat use than vegetation indices or landscape structure.

Although the model showed significant influences, it performed no better than a random model in k-fold mean validation and projected a mean AUC of 0.509 with a highest AUC of 0.519 (Figure 18).

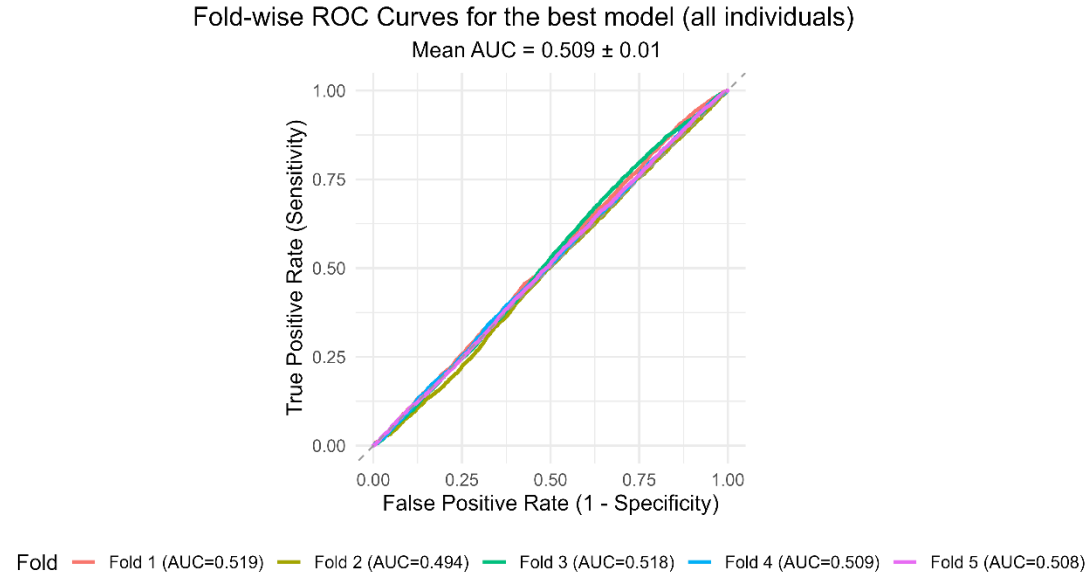

Figure 18. k-fold mean validation for the best RSF model of IBM simulated birds. This model included all the individuals irrespective of their breeding status.

The RSF models for breeders and non-breeders also supported the same model and represented lowest AIC for the combination of NDVI, LULC classes, shei and ed.

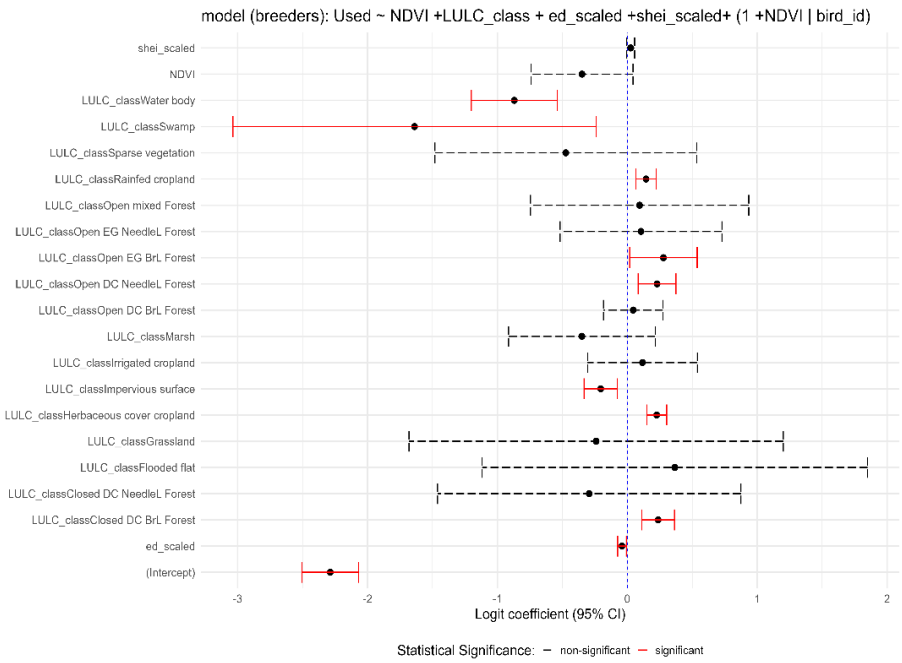

Figure 19. The best model to describe the resource selection of simulated birds (breeders only) containing the combination of NDVI, LULC classes, ed and shei. The significant coefficients are represented with red solid lines and non-significant coefficients are represented with black dashed lines. LULC class names were abbreviated (DC=Deciduous, EG=Evergreen, NeedleL=Needle-leaved, BrL=Broad-leaved).

The best model for breeders (Figure19) showed significant preference for rainfed croplands, open deciduous needle leaved forests, open evergreen broadleaved forest, herbaceous croplands and closed deciduous broadleaved forests. They also showed significant negative coefficients for swamps, impervious surfaces, deep water, indicating avoidance of these LULC classes.

This breeder-only model performed better in k-fold mean validation than the overall model. It projected a mean AUC of 0.56 but the highest AUC was 0.614 for fold 5 only (Figure 20), which shows that the model has moderate predictability for a single group.

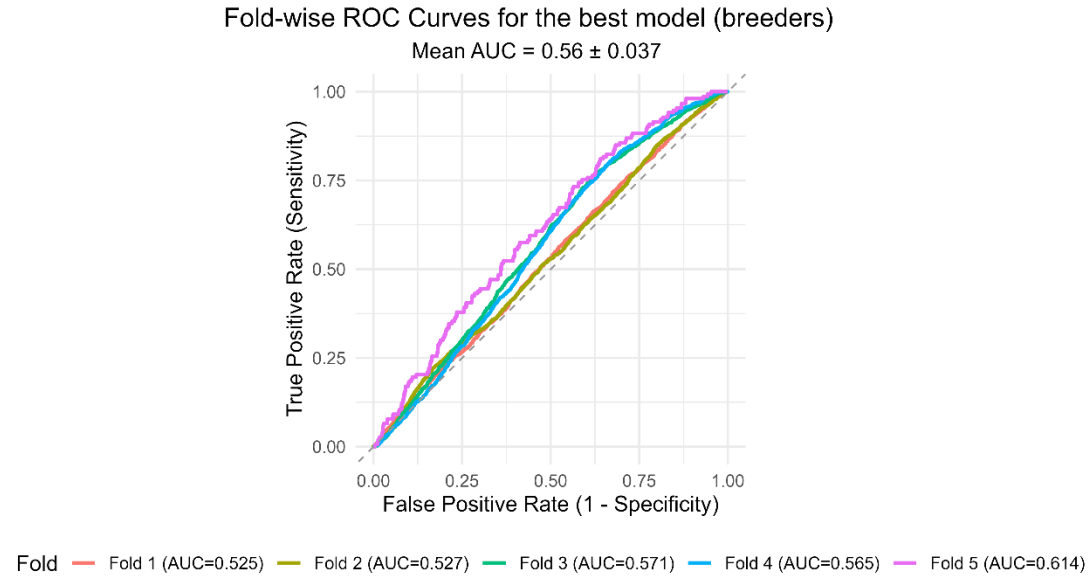

Figure 20. k-fold mean validation for the best RSF model of for IBM simulated breeding birds. This model included all the individuals irrespective of their breeding status.

For non-breeding birds they showed significant positive coefficients for herbaceous cover croplands, closed evergreen needle leaved forest, open deciduous needle leaved forest closed deciduous needle leaved forest and bare areas (Figure 21).

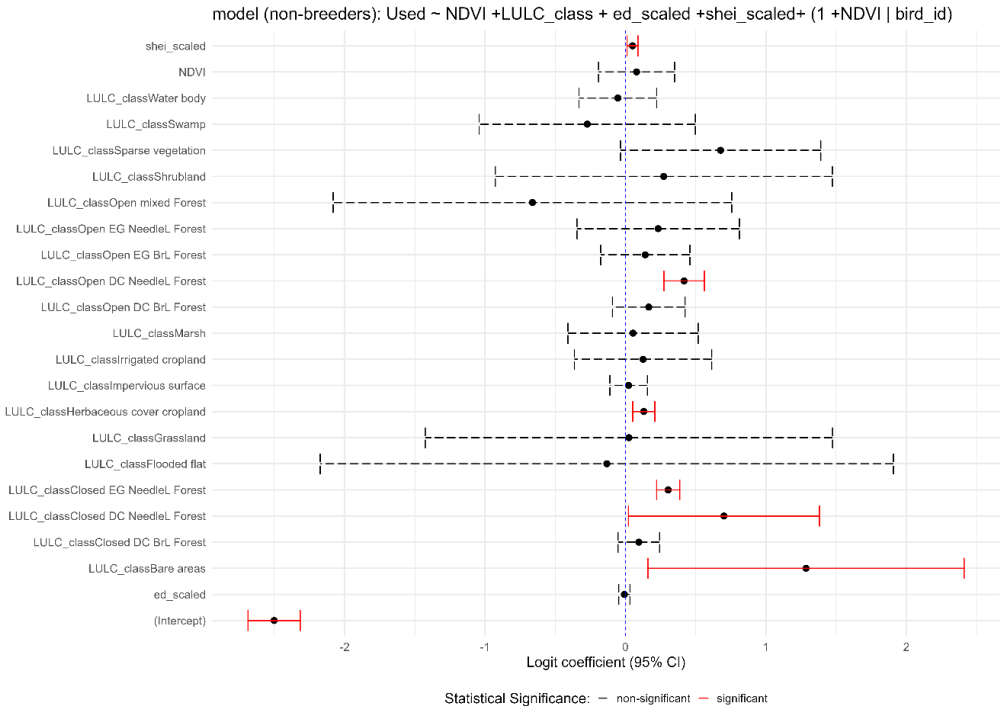

Figure 21. The best model to describe the resource selection of simulated birds (non-breeders only) containing the combination of NDVI, LULC classes, ed and shei. The significant coefficients are represented with red solid lines and non-significant coefficients are represented with black dashed lines. LULC class names were abbreviated (DC=Deciduous, EG=Evergreen, NeedleL=Needle-leaved, BrL=Broad-leaved)

The model with non-breedeers also did not perform any better than the random model in k-fold cross validation. Which shows individual variability can weaken the model predictive power. (Figure 22).

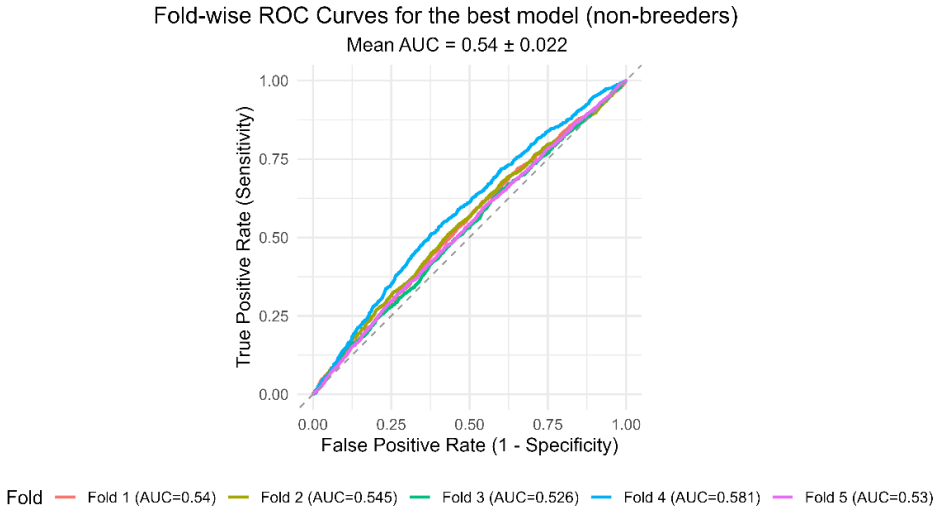

Figure 22. k-fold mean validation for the best RSF model of for IBM simulated breeding birds. This model included all the individuals irrespective of their breeding status.

Chapter 4. Discussion

4.1 Role of NDVI and NDWI in Stork Habitat Selection

The weak influence of NDVI and NDWI in the RSF models reflects both the structure of the study landscape and the scale at which movement was analysed. As shown in the Results, the NDWI and NDVI based models had nearly equivalent AIC support, and NDWI was initially considered because white storks regularly forage in wetlands, shallow flooded meadows, and moist agricultural fields (Pestka et al., 2023; Tsachalidis & Goutner, 2002). However, NDWI values in the study area showed very little spatial variation (Figure 4b), indicating that broad moisture gradients were weakly expressed in this landscape. Under such conditions, selection of moist habitats is difficult to detect statistically as there is little gradient to respond to (Zeller et al., 2012), which explains why NDWI contributed little to the model despite its slightly lower AIC. Moreover, NDVI values in the area fell largely within a moderate range, corresponding to the short vegetation structure characteristic of pastures, croplands and grasslands habitats, which are the primary foraging ground of storks due to high prey detectability of the habitats (Alonso et al., 1991; Standfuß et al., 2022).

Although storks forage mainly in open agricultural habitats, NDVI does not strongly differentiate between short-vegetation classes with comparable greenness values, such as pastures, croplands and grasslands. Previous work has shown that NDVI can perform poorly in grassland-dominated agricultural systems because mowing, harvesting and rapid regrowth introduce strong temporal noise into NDVI time series (Standfuß et al., 2022, 2024). In contrast, NDVI tends to be more informative at broader spatial scales, for example along migratory flyways, where storks track large-scale primary productivity gradients (Martin et al., 2021).

4.2 Habitat Selection in White Storks is Driven by LULC Class Preferences

White storks in this study exhibited clear and consistent habitat preferences in their breeding season. They showed a clear preference for herbaceous croplands and impervious surfaces. There are several studies supporting our result, which indicate croplands as the general foraging ground of storks during the breeding season (Alonso et al., 1991; Gilbert et al., 2016). One of the foundational studies also pointed out that white storks consistently select herbaceous and open-ground habitats such as tall-grass pastures, ploughed cereal fields, and open ash groves,

as these environments concentrate prey and provide greater visibility and accessibility (Alonso et al., 1991). Moreover, a study from the Warminska Refuge Natura 2000 site in Poland revealed that traditional agricultural mosaics composed of pastures and meadows promote white stork nesting (Pestka et al., 2023), underscoring the importance of herbaceous croplands in their habitat selection.

Furthermore, the storks were also found to prefer impervious surfaces in this study, which is supported by some recent research claiming that the storks are opportunistic foragers (Standfuß et al., 2022) and they can swiftly forage in new areas if food resources are not enough, like in landfills (Gilbert et al., 2016; Tryjanowski et al., 2018). In a study from Iberia, the researchers found that white storks were actively foraging in impervious-surface environments such as landfills during their breeding and non-breeding season due to the availability of rodents as their prey (Gilbert et al., 2016). This preference for impervious surfaces is also supported by a study conducted in Madrid, Spain, which documented a long-term shift of white storks movement toward landfill-dominated landscapes, where artificial food subsidies drive both nesting density and foraging activity (López-García & Aguirre, 2023).

4.2.1 Simulated Birds Showed Preference for Different Forest Types

The IBM-simulated birds revealed a preference for open deciduous needle-leaved forest in the overall model. Although, the model had a weak statistical fit, the breeder and non-breeder specific models also indicated a preference for different forest types. Where forests are generally avoided by the storks, there have been studies where the storks have been recorded foraging in the forests. A study in the forests in Poland revealed that a single individual stork got used to foraging in the forest for preying on lizards (Tryjanowski et al., 2018). Furthermore, (Zurell et al., 2018), noticed that while breeding storks generally avoided the forest, breeders from one particular region actually preferred broadleaved and mixed forests. Likewise, despite having a weak predictability of the model, the non-breeding birds also showed a preference for forested areas. Although evidence for white stork use of forested habitats is limited, their opportunistic foraging behaviour and substantial individual-level variation suggest that some birds may occasionally exploit forest patches. Observations from Tryjanowski et al., (2018) and Zurell et al., (2018) indicate that behavioural differences among individuals can extend their foraging into less typical environments, including certain forested areas.

4.3 Storks Prefer Structured Openness in Landscape Configuration

The landscape-configuration coefficients in our study indicate that white storks exhibit a distinct structural preference characterised by the combined selection for high edge density (ed) and avoidance of evenly mixed landscapes (shei). This pattern aligns with the species foraging ecology and supports the idea that storks rely on a specific blend of openness and structure that maximises prey accessibility while preserving efficient movement across the landscape. A study from East Germany on the nest locations of white storks revealed that storks prefer moderately fragmented with high density of habitat boundaries and heterogeneous landscapes (Latus et al., 2000), which aligns with the positive coefficient of ed in our study. Although they only mentioned the diversity of habitat and did not explicitly mention the evenness of habitat distribution, they mentioned that the stork habitat preference is highly influenced by the presence of grasslands, which indicated their bias for a single habitat type. Furthermore, a study in Brandenburg, Germany, revealed that stork presence was highest in landscapes where grasslands and croplands were highly fragmented into numerous small patches, generating high boundary density between habitat types (Latus & Kujawa, 2005), which exactly supports our outcomes.

4.4 Landscape Structure Can Influence Breeding Status as a Fitness Proxy

The IBM outputs showed a clear difference in the number of successful breeders in the simulation years. The number of breeders increased when the preferred habitats increased (Figure 15) in the following years with highest number of breeders in 2018 which decreased slightly in 2024. It indicated that landscape structure can indirectly shape reproductive outcomes in white storks. This is supported by evidence from (Onmuş et al., 2012), that the number of breeding pairs and fledgling output in white storks strongly correlates with habitat features that determine foraging opportunities. Moreover, some other studies also supports that the breeding success of storks depend on their surrounding habitat quality and presence of preferred habitats or habitats with higher prey availability (Bialas et al., 2020; Nowakowski, 2003).

Additionally, simulated differences between breeders and non-breeders align with GPS tracking studies showing that white storks adopt highly variable movement strategies (Zurell et al., 2018). As successful breeders are anchored to a specific nest site due to their higher site fidelity (Vergara et al., 2006), they are forced to rely entirely on the surrounding local

landscape. This lack of mobility makes them highly sensitive to habitat quality. In resource-poor environments, the energy spent during repeated commutes to and from the nest compounds quickly, leading to rapid deficits. This central-place constraint makes breeders particularly vulnerable to landscape changes that alter field size, edge availability, or crop management.

Overall, the IBM results indicates that the spatial arrangement of preferred foraging habitats can determine whether individuals can gain more energy than they spend. When high-quality habitats are close to the nest and clustered, birds maintain positive net energy and when these patches are distant, travel costs tend to exceed gains and breeding becomes unlikely. This reflects movement-ecology theory linking landscape configuration to fitness through its effects on movement costs and access to resources (McGarigal et al., 2016; Nathan et al., 2008).

4.5 Factors Influencing the RSF Predictive Performance

The weak predictive performance of the RSF model (mean AUC = 0.581) likely reflects several contributing factors. The simplified model structure and limited set of environmental predictors may not fully capture the primary drivers of stork habitat selection. Additionally, the narrow geographic scope of the study area reduces environmental variability, which can limit the model's ability to detect selection signals. It is also plausible that storks in this region respond to habitat features or behavioural cues not represented in the current set of covariates. Several studies have reported that meteorological variables like precipitation and temperature influences breeding phenology and migratory behaviour of white stork (Gyalus et al., 2022; Tobolka et al., 2018; Vaitkuviene et al., 2015). Moreover, RSF in this study used static NDVI, which is reportedly a weak predictor as storks response better to the dynamic nature of the agricultural landscape indicating the need for dynamic NDVI layers (Standfuß et al., 2024). Furthermore, the simplification of the of landscape-configuration variables in the IBM is another reason of weak predictive power of the models. To reduce computational load, the IBM was run without landscape-configuration rasters (ed, pd, shei, shdi, ent), and this omission likely further limited the predictive power of the resulting models.

Beyond these methodological constraints, another plausible reason behind the model weak predictability can be the individual variability of the storks. Pooling all individuals into a single model forces the estimation procedure to reconcile fundamentally different movement rules and habitat-use strategies, which dilutes true selection signals and results in low overall model

fitness. However, among the five model validation folds, three folds created an AUC > 0.60, which indicates that the model can detect habitat selection patterns in those three groups

A similar pattern was also shown by the IBM output-derived RSF model. IBM exported a mixed dataset of breeders and non-breeders, and accommodating both of the groups in a single overall model substantially reduced the model fitness (mean AUC = 0.509). This result aligns with findings from other species, although not from white storks, showing that habitat selection varies across behavioural states. When behaviourally distinct individuals are modelled together, the contrasting decision rules dilute the selection signal and suppress overall model accuracy (Chatterjee et al., 2024; Leclerc et al., 2016). However, Carlson et al., (2021), demonstrated that storks exhibit high individual niche specialisation, where distinct individuals consistently prefer specific subsets of the environment rather than following a general species-level pattern. Consequently, treating all individuals as ecologically equivalent violates the assumption of a shared environmental niche, resulting in the poor fit observed in our mixed model.

Furthermore, the behavioural difference was evident when the IBM-derived RSF with only breeding individuals improved the model fitness (mean AUC = 0.56) with one-fold resulting in an AUC > 0.60, which shows that the model shows moderate predictability for a set of individuals. This result aligns with the findings of the home range study of white storks in Central Germany, where the model with breeding birds had better explanatory power than the model with non-breeders (Zurell et al., 2018). Another national-scale analysis by Bonnet-Lebrun et al., (2025), showed that model performance varied markedly across regions, not because of environmental noise but because of habitat preferences differed among and within populations.

Despite these predictive limitations, the hybrid modelling approach retains significant potential for future application. The finding that breeder-only models achieved notably higher AUC values suggests that an underlying selection signal is present yet currently obscured by the model simplification, confounding effects of individual variability and incomplete spatial covariates. Future iterations of this framework would likely benefit from integrating landscape-configuration metrics to better capture the complexity of the landscape and improve the explanatory power of IBM-derived RSFs.

4.6 Implication for Conservation

White storks have benefited from a long history of successful conservation initiatives across Europe. Measures such as securing the electricity lines with insulating materials (Czujkowska et al., 2017), the installation of artificial nesting platforms (Czujkowska et al., 2017; Santopaolo et al., 2013) and the establishment of "European Stork Villages" by EuroNatur (Ferger & Schwaderer, 2016) have played a substantial role in supporting the population recovery of white storks in Europe. However, to continue this positive trend, the breeding grounds of storks need to be protected. As central-place foragers, breeding storks are spatially constrained to a limited radius around their nests (Alonso et al., 1991; Gilbert et al., 2016; Zurell et al., 2018). Because storks generally exhibit strong site fidelity, returning to the same territories annually, they are unable to easily relocate if their local foraging grounds are degraded. Moreover, breeding success is closely linked to site fidelity; individuals that exhibit stronger fidelity to high-quality breeding areas generally achieve higher reproductive success (Vergara et al., 2006).

The RSF–IBM approach used in this study reinforces this pattern by showing that only breeders displayed consistent and ecologically interpretable habitat associations, whereas non-breeders exhibited far more variable and opportunistic space use. This distinction is critical as treating all storks as ecologically interchangeable risks directing conservation efforts toward areas that are frequently used but contribute little to long-term population viability (Carlson et al., 2021). For effective spatial planning, habitats specifically and repeatedly selected by breeders should therefore be prioritised, as these signals are more robust, ecologically meaningful, and potentially more geographically transferable than habitat associations derived from mixed populations (Zurell et al., 2018).

In practice, preserving these breeding habitats requires a shift towards managing landscape configuration and quality rather than just protecting nests. To support high prey accessibility, traditional agricultural practices such as periodic mowing of meadows should be continued, as short vegetation heights are critical for stork foraging efficiency (Johst et al., 2001). On the other hand, the intensification of agriculture, particularly the conversion of grasslands into stork unfavourable habitats like dense monocultures must be discontinued within core breeding area, as these habitats become physically less accessible to storks during the breeding season (Bjedov et al., 2023; Kosicki et al., 2006). Furthermore, conservation strategies should introduce or restore landscape heterogeneity by reviving the dried or drained wetlands.

Finally, the management of anthropogenic food sources requires a precautionary approach. While open landfills currently provide preys like rodents and attracts stork population (Gilbert et al., 2016; López-García & Aguirre, 2023), they can possess substantial risk, specially to the nestlings who show clear biochemical signs of pollution exposure. Exposure to pollutants from landfills areas can affect their nervous system and put the birds are under multi-source environmental stress (Bjedov et al., 2023). With EU directives mandating the closure of open landfill sites, conservation efforts must prioritize the restoration of natural wetlands, grasslands to compensate for this loss of resources, ensuring that storks are not left without food alternatives (López-García & Aguirre, 2023).

Chapter 5. Conclusion

This study demonstrates that white stork habitat selection during the breeding season is driven far more by land-cover composition and landscape structure than by vegetation indices. NDVI and NDWI provided little explanatory power because the study landscape lacked strong gradients in moisture or vegetation productivity, and because rapid agricultural turnover introduces noise that reduces their usefulness. In contrast, LULC classes and landscape fragmentation matrices, emerged as the dominant drivers of habitat use. Storks consistently selected open croplands and areas with high boundary density, which maximise prey accessibility and minimise movement costs, while avoiding uniformly mixed mosaics and moisture-rich habitats that were largely absent in the region.

By embedding RSF coefficients into an IBM, this work shows that these habitat preferences scale mechanistically into differences in energy accumulation and, ultimately, breeding outcomes. Years with higher proportions of preferred herbaceous cropland habitat supported more simulated breeders, linking landscape composition directly to fitness. This result aligns with the theory that central-place foragers are constrained by the quality and configuration of habitats surrounding nest sites.

A major constraint identified in this study is the strong individual variability. Both empirical and simulated RSFs displayed weak predictive performance when all birds were pooled, not because habitat selection was absent but because individuals expressed distinct strategies. Breeder-only models significantly improved predictive accuracy, underscoring that separating behavioural states is essential when modelling habitat selection in species with diverse movement strategies. This has direct implications for conservation. Management based on population-level averages risks overlooking the habitats that are actually relevant for reproductive success.

Methodologically, the RSF-IBM integration used here provides a robust pathway toward mechanistic, behaviourally grounded ecological forecasting. Although simplified landscape-configuration layers in the IBM limited predictive strength, the integrated modelling approach is well-suited for future improvements. Incorporating dynamic fragmentation metrics, realistic food-availability surfaces, and year-to-year landscape transitions would enable scenario testing under different agricultural policies or climate-driven vegetation changes.

Overall, the findings indicate that conservation strategies for white storks should prioritise maintaining herbaceous croplands, preserving fragmented field edges, and protecting heterogeneous agricultural mosaics that support efficient foraging. Because breeders are particularly sensitive to the local landscape surrounding nest sites, conservation planning should treat them as the primary ecological indicator rather than modelling all individuals as behaviourally equivalent. As agricultural systems continue to intensify and homogenise, failing to account for these behavioural and structural dependencies will undermine habitat suitability and reproductive performance. This study establishes a mechanistic foundation for future conservation planning and highlights the value of integrating empirical habitat selection with individual-based simulations to understand species responses in rapidly changing landscapes.

5.1 References

Alonso, J. C., Alonso, J. A., & Carrascal, L. M. (1991). Habitat selection by foraging white storks, Ciconia ciconia, during the breeding season. *Canadian Journal of Zoology*, *69*(7), 1957–1962. https://doi.org/10.1139/z91-270

Berthold, P., Bossche, W. V.D., Jakubiec, Z., Kaatz, C., Kaatz, M., & Querner, U. (2002). Long-term satellite tracking sheds light upon variable migration strategies of White Storks (Ciconia ciconia). *Journal Für Ornithologie*, *143*(4), 489–495. https://doi.org/10.1046/j.1439-0361.2002.02044.x

Berthold, P., Kaatz, C., Kaatz, M., Querner, U., Bossche, W. van den, Chernetsov, N., Fiedler, W., & Wikelski, M. (2022). Data from: Study "MPIAB Argos white stork tracking (1991-2017)." *Movebank Data Repository*. https://doi.org/https://www.doi.org/10.5441/001/1.k29d81dh

Bialas, J. T., Dylewski, Ł., & Tobolka, M. (2020). Determination of nest occupation and breeding effect of the white stork by human-mediated landscape in Western Poland. *Environmental Science and Pollution Research*, *27*(4), 4148–4158. https://doi.org/10.1007/s11356-019-06639-0

Bjedov, D., Velki, M., Kovačić, L. S., Begović, L., Lešić, I., Jurinović, L., Mikuska, T., Sudarić Bogojević, M., Ečimović, S., & Mikuška, A. (2023). White Stork (Ciconia ciconia) Nestlings Affected by Agricultural Practices? Assessment of Integrated Biomarker Responses. *Agriculture*, *13*(5), 1045. https://doi.org/10.3390/agriculture13051045

Bolnick, D. I., Svanbäck, R., Fordyce, J. A., Yang, L. H., Davis, J. M., Hulsey, C. D., & Forister, M. L. (2003). The Ecology of Individuals: Incidence and Implications of Individual Specialization. *The American Naturalist*, *161*(1), 1–28. https://doi.org/10.1086/343878

Bonnet-Lebrun, A., Lemaire-Patin, R., Gendre, N., Dugué, H., Dupuy, J., Pedespan, A., Brand, A., Garnier, L., & Barbraud, C. (2025). Habitat availability is insufficient to explain regional variations in white stork breeding habitat preference. *Oikos*, *2025*(9). https://doi.org/10.1002/oik.11214

Boyce, M. S., & McDonald, L. L. (1999). Relating populations to habitats using resource

selection functions. *Trends in Ecology & Evolution*, *14*(7), 268–272. https://doi.org/10.1016/S0169-5347(99)01593-1

Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. . (2002a). Evaluating resource selection functions. *Ecological Modelling*, *157*(2–3), 281–300. https://doi.org/10.1016/S0304-3800(02)00200-4

Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. A. (2002b). Evaluating resource selection functions. *Ecological Modelling*, *157*(2–3), 281–300. https://doi.org/10.1016/S0304-3800(02)00200-4

Brooks, M., Bolker, B., Kristensen, K., Maechler, M., Magnusson, A., Skaug, H., Nielsen, A., Berg, C., & van Bentham, K. (2017). glmmTMB: Generalized Linear Mixed Models using Template Model Builder. In *CRAN: Contributed Packages*. https://doi.org/10.32614/CRAN.package.glmmTMB

Butler, S. J., Boccaccio, L., Gregory, R. D., Vorisek, P., & Norris, K. (2010). Quantifying the impact of land-use change to European farmland bird populations. *Agriculture, Ecosystems & Environment*, *137*(3–4), 348–357. https://doi.org/10.1016/j.agee.2010.03.005

Carlson, B. S., Rotics, S., Nathan, R., Wikelski, M., & Jetz, W. (2021). Individual environmental niches in mobile organisms. *Nature Communications*, *12*(1), 4572. https://doi.org/10.1038/s41467-021-24826-x

Cayuela, H., Besnard, A., Cote, J., Laporte, M., Bonnaire, E., Pichenot, J., Schtickzelle, N., Bellec, A., Joly, P., & Léna, J. P. (2020). Anthropogenic disturbance drives dispersal syndromes, demography, and gene flow in amphibian populations. *Ecological Monographs*, *90*(2), 1–25. https://doi.org/10.1002/ecm.1406

Chatterjee, N., Wolfson, D., Kim, D., Velez, J., Freeman, S., Bacheler, N. M., Shertzer, K., Taylor, J. C., & Fieberg, J. (2024). Modelling individual variability in habitat selection and movement using integrated step-selection analysis. *Methods in Ecology and Evolution*, *15*(6), 1034–1047. https://doi.org/10.1111/2041-210X.14321

Cornelsen, K. A., Elphinstone, A., & Jordan, N. R. (2024). Digging deeper: habitat selection within the home ranges of a threatened marsupial. *Animal Behaviour*, *212*, 149–174. https://doi.org/10.1016/j.anbehav.2024.03.010

Czujkowska, A., Królak, T., Łukasik, D., Olszewski, A., Rzępała, M., Tarłowski, A., Wierzbicka, A., & Zbyryt, A. (2017). *A guide to the protection of the white stork* (A. Wierzbicka (ed.)).

Davis, A. J., Manly, B. F. J., McDonald, L. L., & Thomas, D. L. (1994). Resource Selection by Animals: Statistical Design and Analysis for Field Studies. *The Journal of Animal Ecology*, *63*(3), 745. https://doi.org/10.2307/5247

Davis, K. P., Sofaer, H. R., & Pejchar, L. (2023). Land cover differentially affects abundance of common and rare birds. *Global Change Biology*, *29*(11), 2999–3009. https://doi.org/10.1111/gcb.16700

DeAngelis, D. L., & Mooij, W. M. (2005). Individual-Based Modeling of Ecological and Evolutionary Processes. *Annual Review of Ecology, Evolution, and Systematics*, *36*(1), 147–168. https://doi.org/10.1146/annurev.ecolsys.36.102003.152644

Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., Schweitzer, J. A., & Palkovacs, E. P. (2017). The ecological importance of intraspecific variation. *Nature Ecology & Evolution*, *2*(1), 57–64. https://doi.org/10.1038/s41559-017-0402-5

Ferger, S. W., & Schwaderer, G. (2016). The european stork villages - A successful combination of nature conservation and sustainable development of rural areas. In *The White Stork. Studies in Biology, Ecology and Conservation* (pp. 219 – 226). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85031284067&partnerID=40&md5=d74c212552da189b66f6843d7c0c141c

Flack, A., Fiedler, W., Blas, J., Pokrovsky, I., Kaatz, M., Mitropolsky, M., Aghababyan, K., Fakriadis, I., Makrigianni, E., Jerzak, L., Azafzaf, H., Feltrup-Azafzaf, C., Rotics, S., Mokotjomela, T. M., Nathan, R., & Wikelski, M. (2016). Costs of migratory decisions: A comparison across eight white stork populations. *Science Advances*, *2*(1), 1–7. https://doi.org/10.1126/sciadv.1500931

Fox, J., Weisberg, S., & Price, B. (2001). car: Companion to Applied Regression. In *CRAN: Contributed Packages*. https://doi.org/10.32614/CRAN.package.car

Fraixedas, S., Lindén, A., Piha, M., Cabeza, M., Gregory, R., & Lehikoinen, A. (2020). A state-of-the-art review on birds as indicators of biodiversity: Advances, challenges, and

future directions. *Ecological Indicators*, *118*(July), 106728. https://doi.org/10.1016/j.ecolind.2020.106728

Gilbert, N. I., Correia, R. A., Silva, J. P., Pacheco, C., Catry, I., Atkinson, P. W., Gill, J. A., & Franco, A. M. A. (2016). Are white storks addicted to junk food? Impacts of landfill use on the movement and behaviour of resident white storks (Ciconia ciconia) from a partially migratory population. *Movement Ecology*, *4*(1), 7. https://doi.org/10.1186/s40462-016-0070-0

Gomez, S., English, H. M., Alegre, V. B., Blackwell, P. G., Bracken, A. M., Bray, E., Evans, L. C., Gan, J. L., Grecian, W. J., Roberts, C. G., Harju, S. M., Hejcmanová, P., Lelotte, L., Marshall, B. M., Matthiopoulos, J., Mnenge, A. M. J., Niebuhr, B. B., Ortega, Z., Pollock, C. J., … Börger, L. (2025). Understanding and predicting animal movements and distributions in the Anthropocene. In *Journal of Animal Ecology*. John Wiley and Sons Inc. https://doi.org/10.1111/1365-2656.70040

Grimm, V., & Railsback, S. F. (2007). Individual-based Modeling and Ecology. *Landscape Ecology*, *22*(2), 317–318. https://doi.org/10.1007/s10980-006-9008-z

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S. A., Milles, A., Nabe-Nielsen, J., Polhill, J. G., Radchuk, V., Rohwäder, M.-S., Stillman, R. A., Thiele, J. C., & Ayllón, D. (2020). The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. *Journal of Artificial Societies and Social Simulation*, *23*(2). https://doi.org/10.18564/jasss.4259

Gyalus, A., Lovászi, P., Végvári, Z., & Csörgő, T. (2022). Effects of climate variables on the White Stork (Ciconia ciconia L.) productivity in a long term study. *Ornis Hungarica*, *30*(2), 61–74. https://doi.org/10.2478/orhu-2022-0020

Hesselbarth, M. H. K., Sciaini, M., Nowosad, J., & Hanss, S. (2018). landscapemetrics: Landscape Metrics for Categorical Map Patterns. In *CRAN: Contributed Packages*. https://doi.org/10.32614/CRAN.package.landscapemetrics

Hesselbarth, M. H. K., Sciaini, M., With, K. A., Wiegand, K., & Nowosad, J. (2019). landscapemetrics : an open-source R tool to calculate landscape metrics. *Ecography*, *42*(10), 1648–1657. https://doi.org/10.1111/ecog.04617

Itonaga, N., Köppen, U., Plath, M., & WallschläGer, D. (2011). Declines in breeding site fidelity in an increasing population of White Storks Ciconia ciconia. *Ibis*, *153*(3), 636–639. https://doi.org/10.1111/j.1474-919X.2011.01131.x

Jeltsch, F., Roeleke, M., Abdelfattah, A., Arlinghaus, R., Berg, G., Blaum, N., De Meester, L., Dittmann, E., Eccard, J. A., Fournier, B., Gaedke, U., Gallagher, C., Govaert, L., Hauber, M., Jeschke, J. M., Kramer-Schadt, S., Linstädter, A., Lucke, U., Mazza, V., … Zurell, D. (2025). The need for an individual-based global change ecology. *Individual-Based Ecology*, *1*, 1–18. https://doi.org/10.3897/ibe.1.148200

Jerrentrup, J. S., Dauber, J., Strohbach, M. W., Mecke, S., Mitschke, A., Ludwig, J., & Klimek, S. (2017). Impact of recent changes in agricultural land use on farmland bird trends. *Agriculture, Ecosystems & Environment*, *239*, 334–341. https://doi.org/10.1016/j.agee.2017.01.041

Johst, K., Brandl, R., & Pfeifer, R. (2001). Foraging in a patchy and dynamic landscape: Human land use and the white stork. *Ecological Applications*, *11*(1), 60–69. https://doi.org/https://doi.org/10.1890/1051-0761(2001)011[0060:FIAPAD]2.0.CO;2

Kays, R., Davidson, S. C., Berger, M., Bohrer, G., Fiedler, W., Flack, A., Hirt, J., Hahn, C., Gauggel, D., Russell, B., Kölzsch, A., Lohr, A., Partecke, J., Quetting, M., Safi, K., Scharf, A., Schneider, G., Lang, I., Schaeuffelhut, F., … Wikelski, M. (2022). The Movebank system for studying global animal movement and demography. *Methods in Ecology and Evolution*, *13*(2), 419–431. https://doi.org/10.1111/2041-210X.13767

Kosicki, J. Z., Profus, P., Dolata, P. T., & Tobółka, M. (2006). Food composition and energy demand of the White Stork Ciconia ciconia breeding population . Literature survey and preliminary results from Poland. In *The White Stork in Poland: studies in biology, ecology and conservation* (pp. 169–183).

Kranstauber, B., Cameron, A., Weinzerl, R., Fountain, T., Tilak, S., Wikelski, M., & Kays, R. (2011). The Movebank data model for animal tracking. *Environmental Modelling and Software*, *26*(6), 834–835. https://doi.org/10.1016/j.envsoft.2010.12.005

Latus, C., & Kujawa, K. (2005). The effect of land cover and fragmentation of agricultural landscape on the density of white stork (Ciconia ciconia L.) in Brandenburg, Germany. *Polish Journal of Ecology*, *53*(4), 535–543.

Latus, C., Kujawa, K., & Glemnitz, M. (2000). *The influence of landscape structure on white stork's Ciconia ciconia nest distribution* (P. A. N. M. i I. Zoologii & M. of the E. O. U. (2 ; 1999 ; Gdańsk) (eds.); Vol. 35, Issue 1, pp. 97–102). Muzeum i Instytut Zoologii PAN. http://rcin.org.pl/miiz/dlibra/publication/edition/45252

Leclerc, M., Vander Wal, E., Zedrosser, A., Swenson, J. E., Kindberg, J., & Pelletier, F. (2016). Quantifying consistent individual differences in habitat selection. *Oecologia*, *180*(3), 697–705. https://doi.org/10.1007/s00442-015-3500-6

López-García, A., & Aguirre, J. I. (2023). White Storks nest at high densities near landfills changing stork nesting distributions in the last four decades in Central Spain. *Ornithological Applications*, *125*(2), 1–13. https://doi.org/10.1093/ornithapp/duad009

Martin, B., Onrubia, A., & Ferrer, M. (2021). Climate Change and the Spatiotemporal Variation in Survival of a Long-Distance Migrant (White Stork , Ciconia ciconia) across Western Europe. *MDPI Birds*, *2*, 362–380. https://doi.org/https://doi.org/10.3390/birds2040027

Matthiopoulos, J., Fieberg, J., & Aarts, G. (2020). Species-Habitat Associations: Spatial data, predictive models, and ecological insights. In *Species-Habitat Associations: Spatial data, predictive models, and ecological insights*. University of Minnesota Libraries. https://doi.org/10.24926/2020.081320

Mayall, E., Groves, L., Kennerley, R., Hudson, M., & Franco, A. (2023). Demographic consequences of management actions for the successful reintroduction of the White Stork Ciconia ciconia to the UK. *Bird Conservation International*, *33*, e47. https://doi.org/10.1017/S0959270922000466

McGarigal, K., Cushman, S., Neel, M. C., & Ene, E. (2015). FRAGSTATS: spatial pattern analysis program for categorical maps. In *University of Massachusetts* (Vol. 97331, Issue 503). http://www.umass.edu/landeco/research/fragstats/fragstats.html

McGarigal, K., & Marks, B. (2012). FRAGSTATS Manual: spatial pattern analysis program for quantifying landscape structure PNW-GTR-351. *URL: Ftp://Ftp. Fsl. Orst. Edu/Pub/Fragstats*, *August*, 122. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:FRAGSTATS+:+Spatial+Pattern+Analysis+Program+for+Quantifying+Landscape+Structure#2

McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C., & Cushman, S. A. (2016). Multi-scale habitat selection modeling: a review and outlook. *Landscape Ecology*, *31*(6), 1161–1175. https://doi.org/10.1007/s10980-016-0374-x

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, *105*(49), 19052–19059. https://doi.org/10.1073/pnas.0800375105

Northrup, J. M., Vander Wal, E., Bonar, M., Fieberg, J., Laforge, M. P., Leclerc, M., Prokopenko, C. M., & Gerber, B. D. (2022). Conceptual and methodological advances in habitat-selection modeling: guidelines for ecology and evolution. *Ecological Applications*, *32*(1). https://doi.org/10.1002/eap.2470

Nowakowski, J. J. (2003). Habitat Structure and Breeding Parameters of the White Stork Ciconia ciconia in the Kolno Upland (NE Poland). *Acta Ornithologica*, *38*(1), 39–46. https://doi.org/10.3161/068.038.0109

Nowosad, J., & Stepinski, T. F. (2019). Information theory as a consistent framework for quantification and classification of landscape patterns. *Landscape Ecology*, *34*(9), 2091–2101. https://doi.org/10.1007/s10980-019-00830-x

Onmuş, O., AĞAOĞLU, Y., & Gül, O. (2012). Environmental Factors Affecting Nest-Site Selection and Breeding Success of the White Stork (Ciconia ciconia) in Western Turkey. *The Wilson Journal of Ornithology*, *124*, 354–361. https://doi.org/10.2307/41480753

Ormerod, S. J., & Watkinson, A. R. (2000). Editors' Introduction: Birds and Agriculture. *Journal of Applied Ecology*, *37*(5), 699–705. https://doi.org/10.1046/j.1365-2664.2000.00576.x

Pebesma, E. (2016). sf: Simple Features for R. In *CRAN: Contributed Packages*. https://doi.org/10.32614/CRAN.package.sf

Pestka, Z., Zbyryt, A., Menderski, S., & Jakubas, D. (2023). Habitat suitability mapping of white stork Ciconia ciconia in one of its key European breeding areas. *Ecological Indicators*, *151*(October 2021), 110278. https://doi.org/10.1016/j.ecolind.2023.110278

Reif, J., Gamero, A., Hologšková, A., Aunins, A., Chodkiewicz, T., Hristov, I., Kurlavičius, P., Leivits, M., Szép, T., & Voříšek, P. (2024). Accelerated farmland bird population

declines in European countries after their recent EU accession. *Science of The Total Environment*, *946*(February), 174281. https://doi.org/10.1016/j.scitotenv.2024.174281

Rigal, S., Dakos, V., Alonso, H., Auniņš, A., Benkő, Z., Brotons, L., Chodkiewicz, T., Chylarecki, P., de Carli, E., del Moral, J. C., Domşa, C., Escandell, V., Fontaine, B., Foppen, R., Gregory, R., Harris, S., Herrando, S., Husby, M., Ieronymidou, C., … Devictor, V. (2023). Farmland practices are driving bird population decline across Europe. *Proceedings of the National Academy of Sciences*, *120*(21), 2017. https://doi.org/10.1073/pnas.2216573120

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2010). pROC: Display and Analyze ROC Curves. In *CRAN: Contributed Packages*. https://doi.org/10.32614/CRAN.package.pROC

Rohwäder, M. S., & Jeltsch, F. (2022). Foraging personalities modify effects of habitat fragmentation on biodiversity. *Oikos*, *2022*(12). https://doi.org/10.1111/oik.09056

Rotics, S., Kaatz, M., Resheff, Y. S., Turjeman, S. F., Zurell, D., Sapir, N., Eggers, U., Flack, A., Fiedler, W., Jeltsch, F., Wikelski, M., & Nathan, R. (2016). The challenges of the first migration: movement and behaviour of juvenile vs. adult white storks with insights regarding juvenile mortality. *Journal of Animal Ecology*, *85*(4), 938–947. https://doi.org/10.1111/1365-2656.12525

Santopaolo, R., Godino, G., Golia, S., Mancuso, A., Monterosso, G., Pucci, M., Santopaolo, F., & Gustin, M. (2013). Increase of white stork Ciconia ciconia population attracted by artificial nesting platforms in Calabria, Italy. *Conservation Evidence*, *10*, 67–69.

Schlägel, U. E., Grimm, V., Blaum, N., Colangeli, P., Dammhahn, M., Eccard, J. A., Hausmann, S. L., Herde, A., Hofer, H., Joshi, J., Kramer-Schadt, S., Litwin, M., Lozada-Gobilard, S. D., Müller, M. E. H., Müller, T., Nathan, R., Petermann, J. S., Pirhofer-Walzl, K., Radchuk, V., … Jeltsch, F. (2020). Movement-mediated community assembly and coexistence. *Biological Reviews*, *95*(4), 1073–1096. https://doi.org/10.1111/brv.12600

Siriwardena, G. M., Crick, H. Q. P., Baillic, S. R., & Wilson, J. D. (2000). Agricultural land-use and the spatial distribution of granivorous lowland farmland birds. *Ecography*, *23*(6), 702–719. https://doi.org/10.1111/j.1600-0587.2000.tb00314.x

Standfuß, I., Geiß, C., Nathan, R., Rotics, S., Scacco, M., Kerr, G., & Taubenböck, H. (2022). Time series enable the characterization of small-scale vegetation dynamics that influence fine-scale animal behavior – an example from white storks' foraging behavior. *Remote Sensing in Ecology and Conservation*, *8*(3), 391–408. https://doi.org/10.1002/rse2.251

Standfuß, I., Geiß, C., Senaratne, H., Kerr, G., Nathan, R., Rotics, S., & Taubenböck, H. (2024). Assessing cumulative uncertainties of remote sensing time series and telemetry data in animal-environment studies. *Landscape Ecology*, *39*(2), 7. https://doi.org/10.1007/s10980-024-01804-4

Tobolka, M., Dylewski, L., Wozna, J. T., & Zolnierowicz, K. M. (2018). How weather conditions in non-breeding and breeding grounds affect the phenology and breeding abilities of white storks. *Science of The Total Environment*, *636*, 512–518. https://doi.org/10.1016/j.scitotenv.2018.04.253

Tryjanowski, P., Grzywaczewski, G., & Zbyryt, A. (2018). Foraging of white stork ciconia ciconia in forests-The heritage of an ancient behaviour? *Polish Journal of Ecology*, *66*(3), 250–256. https://doi.org/10.3161/15052249PJE2018.66.3.005

Tsachalidis, E. P., & Goutner, V. (2002). Diet of the White Stork in Greece in Relation to Habitat E. *Waterbirds*, *30*(1), 417–423. https://doi.org/https://doi.org/10.1675/1524-4695(2002)025[0417:DOTWSI]2.0.CO;2

Vaitkuviene, D., Dagys, M., Bartkeviciene, G., & Romanovskaja, D. (2015). The effect of weather variables on the White Stork (Ciconia ciconia) spring migration phenology. *Ornis Fennica*, *92*(1), 43–52. https://doi.org/10.51812/of.133867

Vergara, P., Aguirre, J. I., Fargallo, J. A., & Davila, J. A. (2006). Nest-site fidelity and breeding success in White Stork Ciconia ciconia. *Ibis*, *148*(4), 672–677. https://doi.org/10.1111/j.1474-919X.2006.00565.x

Wiens, T. S., Dale, B. C., Boyce, M. S., & Kershaw, G. P. (2008). Three way k-fold cross-validation of resource selection functions. *Ecological Modelling*, *212*(3–4), 244–255. https://doi.org/10.1016/j.ecolmodel.2007.10.005

Wikelski, M., Davidson, S., & Kays, R. (2025). *Movebank: archive, analysis and sharing of animal movement data. Hosted by the Max Planck Institute of Animal Behavior.*

www.movebank.org

Zeller, K. A., McGarigal, K., & Whiteley, A. R. (2012). Estimating landscape resistance to movement: A review. *Landscape Ecology*, *27*(6), 777–797. https://doi.org/10.1007/s10980-012-9737-0

Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X., & Liu, L. (2024). GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. *Earth System Science Data*, *16*(3), 1353–1381. https://doi.org/10.5194/essd-16-1353-2024

Zurell, D., Eggers, U., Kaatz, M., Rotics, S., Sapir, N., Wikelski, M., Nathan, R., & Jeltsch, F. (2015). Individual-based modelling of resource competition to predict density-dependent population dynamics: a case study with white storks. *Oikos*, *124*(3), 319–330. https://doi.org/10.1111/oik.01294

Zurell, D., von Wehrden, H., Rotics, S., Kaatz, M., Groß, H., Schlag, L., Schäfer, M., Sapir, N., Turjeman, S., Wikelski, M., Nathan, R., & Jeltsch, F. (2018). Home Range Size and Resource Use of Breeding and Non-breeding White Storks Along a Land Use Gradient. *Frontiers in Ecology and Evolution*, *6*(JUN). https://doi.org/10.3389/fevo.2018.00079

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, *1*(1), 3–14. https://doi.org/10.1111/j.2041-210x.2009.00001.x

Chapter 6. Appendix

6.1 Raw data summary and column names

Source Data: MPIAB Argos white stork tracking (1991-2017) (Movebank Study ID 7431347)

Subset: This subset contains the yearly breeding season (Mar-Aug) data of 33 individuals who stayed within the German boundary in any year for their entire breeding season.

Total GPS records: 1,302

Time span: April 1992 - August 2012

Months covered: March to August (breeding season)

Sensor types: 2 (mostly argos-doppler-shift)

Time interval: Irregular

Sex: 1 male (ID:14549A), 32 unknown

Age: 3 Adults, 12 juveniles (1st winter), 1 juvenile, 17 unknown

Living status: 5 Wild, 10 raised or kept in animal care centre, 17 not specified

Table S1. Column descriptions of the raw data file (only that contains data inputs).

Column Name	**Description**
event.id	Unique identifier for each tracking record
visible	Indicates whether the point is marked as visible
timestamp	Timestamp of the GPS/ARGOS record (UTC)
location.long	Longitude in decimal degrees
location.lat	Latitude in decimal degrees
algorithm.marked.outlier	Flag if automatically marked as outlier
argos.altitude	Estimated altitude above sea level (meters)
argos.best.level	ARGOS best level quality estimate
argos.calcul.freq	Frequency used in ARGOS Doppler calculation
argos.iq	Index of Quality from ARGOS system
argos.lat1	First possible latitude from ARGOS
argos.lat2	Second possible latitude from ARGOS
argos.lc	ARGOS Location Class (accuracy level) (higher location class (LC) = better estimation, lower LC or class = A/B/Z
argos.lon1	First possible longitude from ARGOS
argos.lon2	Second possible longitude from ARGOS
argos.nb.mes	Number of messages received by satellite
argos.nb.mes.120	Number of 120-bit messages received
argos.nopc	Number of position calculations attempted
argos.pass.duration	Duration of satellite pass (seconds)

argos.valid.location.algorithm	Flag if location is valid per ARGOS algorithm
manually.marked.valid	TRUE if manually marked as valid location
sensor.type	Type of sensor used (e.g., argos-doppler-shift or GPS)
individual.taxon.canonical.name	Scientific name of the tracked species
tag.local.identifier	Unique identifier of the tag on the bird
individual.local.identifier	Unique identifier of the individual bird

6.2 Correlation matrix for model covariates

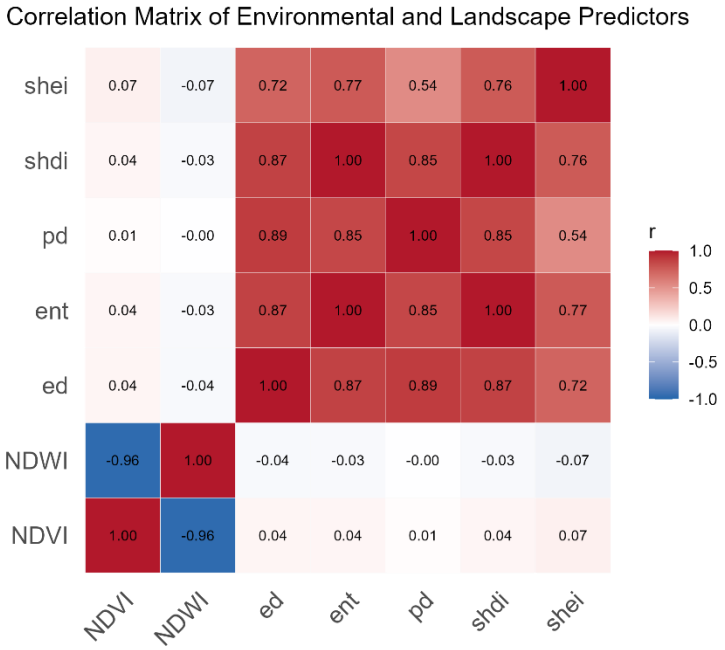

Figure S1. Correlation matrix for model covariates where ed= edge density, pd= patch density, shei = Shannon's evenness index, shdi = Shannon's diversity index, ent = Shannon entropy.

6.3 Breeders and non-breeders track over the years

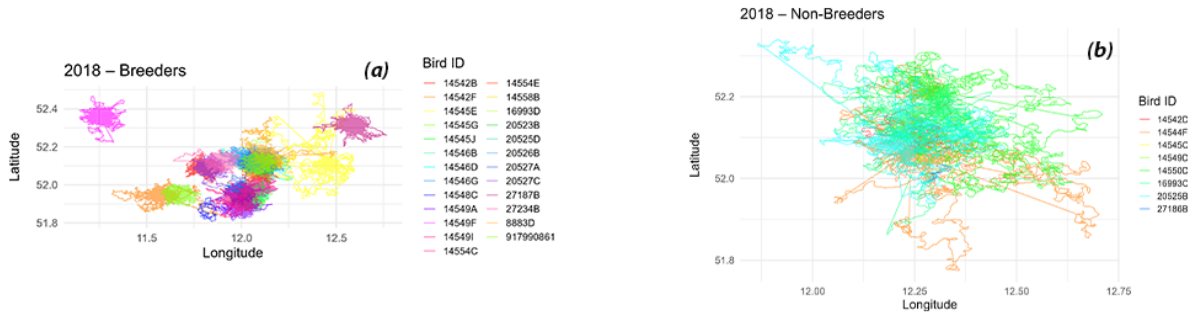

Figure S2. The movement of breeders (a) and non-breeders (b) across the landscape for the year 2018, with Longitude on the x-axis and Latitude on the y-axis. Each individual is identified by their unique ID number in the legend on the right.

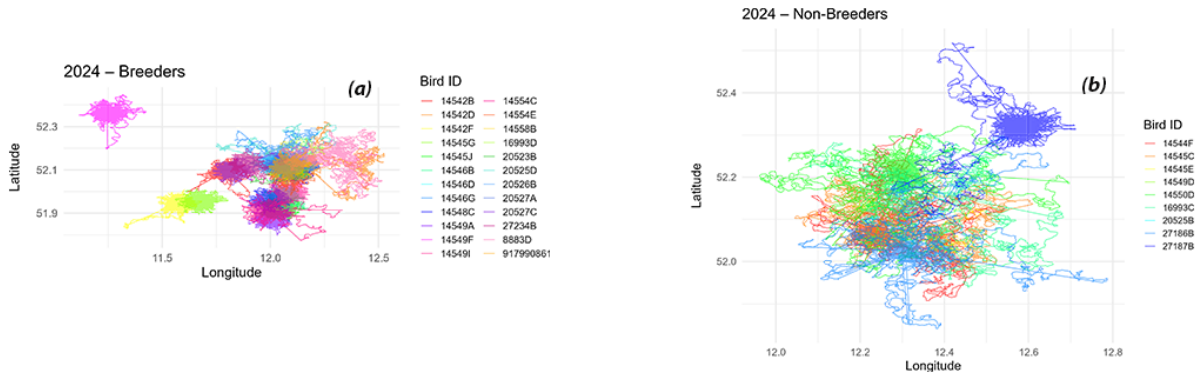

Figure S3. The movement of breeders (a) and non-breeders (b) across the landscape for the year 2024, with Longitude on the x-axis and Latitude on the y-axis. Each individual is identified by their unique ID number in the legend on the right.

6.4 AIC values for all the model combinations

Table S2. AIC values for all model combinations. Column RSF1 corresponds to RSF model with empirical data, RSF2 corresponds the IBM output induced RSF model with all birds, RSF3 corresponds to breeder only RSF models and RSF4 corresponds to non-breeders only RSF model AICs. The green highlighted cells indicates the lowest AIC. The yellow highlighted cell indicates the selected model for the RSF1 set as it was ecologically supported for the study region.

Model Combinations	**AIC**			
	RSF1	RSF2	RSF3	RSF4
Used ~ NDWI + LULC_class + ed_scaled + shei_scaled + (1 + NDWI | bird_id)	8490.3	107194.1	62275.5	44640.5

Used ~ NDVI + LULC_class + ed_scaled + shei_scaled + (1 + NDVI | bird_id)	8494.1	106840.6	62035.5	44606.4
Used ~ NDWI + LULC_class + ed_scaled + (1 + NDWI | bird_id)	8513.5	107204.7	62279.2	44645.7
Used ~ NDWI + ed_scaled + shei_scaled + (1 + NDWI | bird_id)	8514.6	107247.5	62443.7	44665.1
Used ~ NDWI + LULC_class + pd_scaled + (1 + NDWI | bird_id)	8517.5	107200.5	62273.6	44645.7
Used ~ NDVI + LULC_class + ed_scaled + (1 + NDVI | bird_id)	8517.6	106847	62039.2	44610.9
Used ~ NDVI + ed_scaled + shei_scaled + (1 + NDVI | bird_id)	8520	106934.4	62181.3	44652
Used ~ NDVI + LULC_class + pd_scaled + (1 + NDVI | bird_id)	8521.9	106846.5	62037.5	44611
Used ~ NDWI + LULC_class + shdi_scaled + (1 + NDWI | bird_id)	8530.3	107211.9	62295	44646
Used ~ NDWI + LULC_class + ent_scaled + (1 + NDWI | bird_id)	8530.5	107211.9	62295	44645.9
Used ~ NDVI + LULC_class + shdi_scaled + (1 + NDVI | bird_id)	8534.7	106841.3	62038.3	44611.2
Used ~ NDVI + LULC_class + ent_scaled + (1 + NDVI | bird_id)	8534.8	106841.3	62038.2	44611.4
Used ~ NDWI + ed_scaled + (1 + NDWI | bird_id)	8538.9	107258.4	62451.9	44736.6
Used ~ NDVI + ed_scaled + (1 + NDVI | bird_id)	8544.3	106943.3	62184.4	44655.6
Used ~ NDWI + pd_scaled + (1 + NDWI | bird_id)	8545.3	107250.7	62423.4	44754
Used ~ NDVI + pd_scaled + (1 + NDVI | bird_id)	8551.1	106936.8	62170	44656.1
Used ~ NDWI + LULC_class + (1 + NDWI | bird_id)	8556.5	107210	62293	44644.4
Used ~ NDWI + LULC_class + shei_scaled + (1 + NDWI | bird_id)	8557.7	107211.6	62294.1	44640.4
Used ~ NDWI + shdi_scaled + (1 + NDWI | bird_id)	8560.4	107283	62491.5	44756.9
Used ~ NDWI + ent_scaled + (1 + NDWI | bird_id)	8560.5	107280.3	62491.2	44744.1
Used ~ NDVI + LULC_class + (1 + NDVI | bird_id)	8560.7	106845	62040.2	44613.4
Used ~ NDVI + LULC_class + shei_scaled + (1 + NDVI | bird_id)	8562	106843.3	62042.2	44604.6
Used ~ NDVI + shdi_scaled + (1 + NDVI | bird_id)	8566.3	106950.3	62199.4	44655
Used ~ NDVI + ent_scaled + (1 + NDVI | bird_id)	8566.4	106951	62199.9	44655
Used ~ NDWI + shei_scaled + (1 + NDWI | bird_id)	8586.6	107291.5	62506.8	44734.6
Used ~ NDWI + (1 + NDWI | bird_id)	8587.3	107291.8	62519.4	44746.5
Used ~ NDVI + shei_scaled + (1 + NDVI | bird_id)	8592.2	106950.9	62199.4	44652.1
Used ~ NDVI + (1 + NDVI | bird_id)	8592.8	106949.6	62199.9	44653.5
Used ~ LULC_class + ed_scaled + pd_scaled + (1 | bird_id)	8643.4	113723.1	62668.8	50858.3
Used ~ LULC_class + ed_scaled + (1 | bird_id)	8643.8	113725.4	62672.6	50856.4
Used ~ LULC_class + pd_scaled + (1 | bird_id)	8645.3	113722.7	62667.4	50856.8
Used ~ LULC_class + shdi_scaled + (1 | bird_id)	8659.1	113726.4	62680	50858.1
Used ~ LULC_class + ent_scaled + (1 | bird_id)	8659.2	113726.4	62680	50858.2
Used ~ ed_scaled + (1 | bird_id)	8671.2	113820.4	62844.8	50938.2
Used ~ pd_scaled + (1 | bird_id)	8675.4	113808.5	62817.4	50937.3
Used ~ LULC_class + (1 | bird_id)	8686.9	113725	62678.3	50858.8
Used ~ LULC_class + shei_scaled + (1 | bird_id)	8688.4	113724.8	62680.2	50850.8
Used ~ shdi_scaled + (1 | bird_id)	8692.7	113828.7	62869.3	50938.2
Used ~ ent_scaled + (1 | bird_id)	8692.8	113828.7	62869.3	50938.3
Used ~ shei_scaled + (1 | bird_id)	8722.8	113833.8	62878.2	50935.4

6.5 Model outputs for best RSF models

Table S3. Standard outputs for RSF models from empirical data

RSF1 (RSF from empirical data) Selected model				
Covariates	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.69494	0.08632	-31.219	< 2e-16
NDVI	0.49086	0.37985	1.292	0.196277
Rainfed cropland	0.11754	0.09809	1.198	0.230827
Herbaceous cover cropland	0.31595	0.08294	3.809	0.000139
Grassland	0.47656	0.45390	1.050	0.293751
Marsh	-0.86521	0.73309	-1.180	0.237910
Impervious surface	0.61057	0.12574	4.856	1.20e-06
Irrigated cropland	0.95165	0.56635	1.680	0.092894
Water body	-0.86736	0.46345	-1.872	0.061270
Open evergreen broadleaf forest	-0.27087	0.61196	-0.443	0.658037
Open deciduous broadleaf forest	-0.49564	0.60177	-0.824	0.410141
Closed deciduous broadleaf forest	-0.11201	0.16899	-0.663	0.507419
Open evergreen needleleaf forest	0.31085	0.64526	0.482	0.629983
Open deciduous needleleaf forest	0.22185	0.18938	1.171	0.241413
Closed deciduous needleleaf forest	0.63967	1.08708	0.588	0.556245
ed_scaled	0.37747	0.04544	8.308	< 2e-16
shei_scaled	-0.23468	0.04770	-4.920	8.67e-07

Table S4. Standard outputs for IBM derived RSF models with all bird (breeders and non-breeders).

RSF2 (RSF from IBM output data-all birds) best model with lowest AIC				
Covariates	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.29975	0.09615	-23.917	< 2e-16
NDVI	-0.03494	0.16042	-0.218	0.827566
Rainfed cropland	-0.07049	0.02859	-2.466	0.013680
Herbaceous cover cropland	0.03512	0.02756	1.274	0.202489
Irrigated cropland	-0.03802	0.16293	-0.233	0.815510
Open evergreen broadleaf forest	0.06898	0.10222	0.675	0.499788
Open deciduous broadleaf forest	-0.06637	0.08689	-0.764	0.444927
Closed deciduous broadleaf forest	0.03007	0.04806	0.626	0.531522
Open evergreen needleleaf forest	0.02893	0.21629	0.134	0.893584
Open deciduous needleleaf forest	0.17406	0.05091	3.419	0.000629
Closed deciduous needleleaf forest	0.19866	0.29540	0.673	0.501257
Open mixed forest	-0.30741	0.36732	-0.837	0.402649
Shrubland	-0.65393	0.59288	-1.103	0.270046
Grassland	-0.26132	0.51987	-0.503	0.615197

Sparse vegetation	0.01271	0.29245	0.043	0.965330
Swamp	-0.92241	0.34097	-2.705	0.006825
Marsh	-0.27987	0.18258	-1.533	0.125310
Flooded flat	0.02980	0.60869	0.049	0.960948
Impervious surface	-0.26302	0.04755	-5.532	3.17e-08
Bare areas	0.54496	0.54029	1.009	0.313142
Water body	-0.60996	0.10878	-5.607	2.05e-08
ed_scaled	-0.02863	0.01315	-2.177	0.029454
shei_scaled	0.03563	0.01223	2.914	0.003566

Table S5. Standard outputs for IBM derived RSF models with only breeders.

RSF3 (RSF from IBM output data-breeders only) best model with lowest AIC				
Covariates	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.28706	0.11131	-20.547	< 2e-16
NDVI	-0.34839	0.20012	-1.741	0.081702
Rainfed cropland	0.14408	0.04005	3.597	0.000322
Herbaceous cover cropland	0.22661	0.03881	5.839	5.25e-09
Irrigated cropland	0.11678	0.21575	0.541	0.588313
Open evergreen broadleaf forest	0.27785	0.13241	2.099	0.035861
Open deciduous broadleaf forest	0.04525	0.11653	0.388	0.697768
Closed deciduous broadleaf forest	0.23700	0.06446	3.677	0.000236
Open evergreen needleleaf forest	0.10525	0.31781	0.331	0.740517
Open deciduous needleleaf forest	0.22904	0.07418	3.088	0.002018
Closed deciduous needleleaf forest	-0.29328	0.59589	-0.492	0.622595
Open mixed forest	0.09470	0.42868	0.221	0.825168
Grassland	-0.24005	0.73502	-0.327	0.743981
Sparse vegetation	-0.47339	0.51413	-0.921	0.357177
Swamp	-1.63787	0.71333	-2.296	0.021671
Marsh	-0.34988	0.28805	-1.215	0.224505
Flooded flat	0.36562	0.75695	0.483	0.629089
Impervious surface	-0.20502	0.06522	-3.143	0.001670
Water body	-0.86998	0.16892	-5.150	2.60e-07
ed_scaled	-0.04084	0.01725	-2.367	0.017934
shei_scaled	0.02525	0.01554	1.625	0.104180

Table S6. Standard outputs for IBM derived RSF models with only non-breeders.

RSF4 (RSF from IBM output data-nonbreeders only) best model with lowest AIC				
Covariates	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.502763	0.094876	-26.379	< 2e-16
NDVI	0.080582	0.138017	0.584	0.55932
Herbaceous cover cropland	0.130969	0.040063	3.269	0.00108
Irrigated cropland	0.125854	0.249359	0.505	0.61376

Open evergreen broadleaf forest	0.140877	0.162284	0.868	0.38535
Open deciduous broadleaf forest	0.165913	0.131796	1.259	0.20808
Closed deciduous broadleaf forest	0.095186	0.075718	1.257	0.20871
Open evergreen needleleaf forest	0.232361	0.295210	0.787	0.43122
Closed evergreen needleleaf forest	0.304700	0.041888	7.274	3.49e-13
Open deciduous needleleaf forest	0.417217	0.073376	5.686	1.30e-08
Closed deciduous needleleaf forest	0.701807	0.346750	2.024	0.04297
Open mixed forest	-0.663597	0.724414	-0.916	0.35964
Shrubland	0.272456	0.612288	0.445	0.65633
Grassland	0.025197	0.739756	0.034	0.97283
Sparse vegetation	0.678331	0.363779	1.865	0.06223
Swamp	-0.269719	0.392477	-0.687	0.49194
Marsh	0.052907	0.236781	0.223	0.82319
Impervious surface	0.022879	0.068229	0.335	0.73738
Bare areas	1.284934	0.574484	2.237	0.02531
Water body	-0.053864	0.141189	-0.381	0.70283
ed_scaled	-0.007663	0.020406	-0.376	0.70728
shei_scaled	0.050407	0.019781	2.548	0.01083

6.6 Full IBM model code

```
#Import Python libraries

from __future__ import annotations
import os, math, random, warnings
from dataclasses import dataclass, field
from typing import Dict, Tuple, List
from datetime import datetime, timedelta
import numpy as np
import pandas as pd
import rasterio
from rasterio.transform import rowcol, xy
import rioxarray as rxr
import xarray as xr
import matplotlib.pyplot as plt
from scipy.special import expit
from tqdm import tqdm
warnings.filterwarnings("ignore", category=FutureWarning)

# ===== Define File Paths==========

LULC_PATH = "D:/4th Semester/thesis/models/GeoTiffs/New/LC_2012_07.tif"
NDVI_PATH = "D:/4th
Semester/thesis/models/GeoTiffs/NDVI_NDWI/NDVI_2012_JulAug_30m.tif"
NDWI_PATH = "D:/4th
Semester/thesis/models/GeoTiffs/NDVI_NDWI/MNDWI_2012_JulAug_30m.tif"


# Fragmentation layers (optional leave {} if not available)
FRAG_PATHS = {}
POINTS_PATH = "D:/4th Semester/thesis/models/RSF/MPIAB_inside_data_xl.xlsx"
RANDOM_RSF_PATH = "D:/4th Semester/thesis/models/RSF/rsf_random_effects_by_bird.csv"
FIXED_RSF_PATH  = "D:/4th Semester/thesis/models/RSF/rsf_fixed_effects.csv"
OUTPUT_DIR = "D:/4th Semester/thesis/models/IBM/Output"


# ================ Model Constants ======================

STEP_MINUTES_DEFAULT = 30       #30 minutes interval step record
FOOD_RENEW_DAYS = 2              # food renewal time in days
BREED_WINDOW_DAYS = 15          #number of energy accumulation days for breeding

#Energy spending percentages
ENERGY_REQUIREMENT = {"physio": 0.75, "foraging": 0.10, "flying": 0.10, "other": 0.05}
REQUIREMENT_TOTAL = float(sum(list(ENERGY_REQUIREMENT.values())))
BREED_MIN_EXCESS = 0.30 #Breeding status = breeder if energy surplus = 30% atleast
DEATH_THRESHOLD = 0.60 #Death if net energy <60%

DEFAULT_LULC_ENERGY = {
   10: 3.0, 11: 3.0, 20: 2.5, 51: 1.0, 61: 1.5, 62: 0.5, 71: 0.5,
   72: 0.1, 81: 0.1, 82: 0.1, 130: 3.0, 182: 4.0, 190: 1.5, 210: 0.1
}   #Energy scores per LULC class


ENERGY_SCALARS = {
   "ndvi_gain_scale": 0.3, # Scaling factor converting NDVI (0–1) into additional food availability
```

```
    "ndwi_gain_scale": 0.6,
    "fragmentation_gain_scale": 0.7,
    "distance_cost_scale": 0.0002, # Energy cost per metre of movement (flying penalty)
"baseline_daily_units": 10.0,
}
# Tiled processing size to avoid MemoryError
TILE = 2048


# ===== Activity & nest behavior =====
DAY_START_HOUR = 5          # start of active day
DAY_END_HOUR   = 20         # end of active day

HOME_DECAY = 0.0005         # site fidelity strength (higher = tighter home range)
NEST_PULL_INTERVAL_MIN = 120  # every 2 hours
NEST_PULL_ALPHA = 0.003       # strength of nest-pull reward
NEST_PULL_ACTIVE_ONLY_IF_BRED = True  # True = only after bird is marked 'bred'

BEHAVIOR_PARAMS = {
    "breeding": {
        "step_mu_m": 350.0,     # typical step length during breeding
        "step_sigma_m": 180.0,  # step variability
        "turn_kappa": 3.0,      # higher = straighter persistence
        "nest_pull_alpha": 0.0008,  # smoother home attraction strength
        "max_foraging_radius_m": 28100.0}, #max foraging radius for breeders in meter
    "nonbreeding": {
        "step_mu_m": 2000.0,
        "step_sigma_m": 800.0,
        "turn_kappa": 1.0,
        "nest_pull_alpha": 0.0,       # no nest pull if not breeding
        "max_foraging_radius_m": 48200.0
    }
}


# ================ Helpers ============================

def ensure_dir(path: str):                  #output directory creator
    os.makedirs(path, exist_ok=True)

def softmax(x, temperature=1.0):
    z = (x - np.nanmax(x)) / max(temperature, 1e-6)
    e = np.exp(z)
    return e / np.sum(e)

def within_bounds(shape, r, c):          #Preventing out-of-bound pixel lookups
    h, w = shape
    return 0 <= r < h and 0 <= c < w

# --- Movement helpers ---
def haversine_m(lon1, lat1, lon2, lat2):
    R = 6371000.0
    phi1, phi2 = math.radians(lat1), math.radians(lat2)
    dphi = math.radians(lat2 - lat1)
    dlambda = math.radians(lon2 - lon1)
    a = math.sin(dphi/2)**2 + math.cos(phi1)*math.cos(phi2)*math.sin(dlambda/2)**2
```

```
    return 2 * R * math.asin(min(1.0, math.sqrt(a)))

def bearing_deg(lon1, lat1, lon2, lat2):
    phi1, phi2 = math.radians(lat1), math.radians(lat2)
    dlambda = math.radians(lon2 - lon1)
    y = math.sin(dlambda) * math.cos(phi2)
    x = math.cos(phi1)*math.sin(phi2) - math.sin(phi1)*math.cos(phi2)*math.cos(dlambda)
    brng = (math.degrees(math.atan2(y, x)) + 360.0) % 360.0
    return brng

def von_mises_weight(prev_bearing_deg, new_bearing_deg, kappa):
    # proportional weight in (0,1]; add tiny epsilon to avoid log(0)
    th1 = math.radians(prev_bearing_deg)
    th2 = math.radians(new_bearing_deg)
    # np.i0 is the modified Bessel function of order 0
    return max(1e-12, math.exp(kappa * math.cos(th2 - th1)) / (2 * math.pi * np.i0(kappa)))

def move_along_bearing(lon, lat, dist_m, bearing_deg):
    """Great-circle move from (lon,lat) by dist_m at bearing_deg. Returns (lon2, lat2)."""
    R = 6371000.0
    d = dist_m / R
    brg = math.radians(bearing_deg)
    lat1 = math.radians(lat)
    lon1 = math.radians(lon)

    lat2 = math.asin(math.sin(lat1) * math.cos(d) + math.cos(lat1) * math.sin(d) * math.cos(brg))
    lon2 = lon1 + math.atan2(math.sin(brg) * math.sin(d) * math.cos(lat1),
                             math.cos(d) - math.sin(lat1) * math.sin(lat2))
    return (math.degrees(lon2), math.degrees(lat2))


# ================ Environment ==========================
@dataclass
class Environment:
    lulc_path: str
    ndvi_path: str
    ndwi_path: str
    frag_paths: Dict[str, str]
    lulc: xr.DataArray = field(init=False)
    ndvi: xr.DataArray = field(init=False)
    ndwi: xr.DataArray = field(init=False)
    frag: Dict[str, xr.DataArray] = field(init=False, default_factory=dict)
    transform: rasterio.Affine = field(init=False)
    crs: str = field(init=False)
    food_base: np.ndarray = field(init=False)
    food_stock: np.ndarray = field(init=False)
    last_renewal: datetime = field(init=False)


    def __post_init__(self):
        if os.path.exists("./data/nests_auto.csv"):
            try:
                nests = pd.read_csv("./data/nests_auto.csv")
                minx, miny, maxx, maxy = (
                    nests["lon"].astype(float).min(),
```

```
                nests["lat"].astype(float).min(),
                nests["lon"].astype(float).max(),
                nests["lat"].astype(float).max(),
            )
            buffer = 0.5
            minx -= buffer; miny -= buffer; maxx += buffer; maxy += buffer
            with rasterio.open(self.lulc_path) as src:
                window = src.window(minx, miny, maxx, maxy)
                self.lulc = rxr.open_rasterio(self.lulc_path, masked=True, window=window).squeeze()
        except Exception as e:
            print(f"Cropping failed, loading full raster: {e}")
            self.lulc = rxr.open_rasterio(self.lulc_path).squeeze()
    else:
        self.lulc = rxr.open_rasterio(self.lulc_path).squeeze()

    self.ndvi = rxr.open_rasterio(self.ndvi_path).squeeze().rio.reproject_match(self.lulc)
    self.ndwi = rxr.open_rasterio(self.ndwi_path).squeeze().rio.reproject_match(self.lulc)

    for k, p in self.frag_paths.items():
        self.frag[k.upper()] = rxr.open_rasterio(p).squeeze().rio.reproject_match(self.lulc)

    with rasterio.open(self.lulc_path) as src:
        self.transform = src.transform
        self.crs = src.crs.to_string() if src.crs else ""

    self._init_food_tiled()


# ------- Tiled min/max helper for frag normalization -------
def _min_max_tiled(self, arr: xr.DataArray) -> Tuple[float, float]:
    a = arr.values
    h, w = a.shape
    mn = np.inf
    mx = -np.inf
    for r0 in range(0, h, TILE):
        r1 = min(h, r0 + TILE)
        for c0 in range(0, w, TILE):
            c1 = min(w, c0 + TILE)
            block = a[r0:r1, c0:c1]
            bmn = np.nanmin(block)
            bmx = np.nanmax(block)
            if bmn < mn: mn = bmn
            if bmx > mx: mx = bmx
    if not np.isfinite(mn): mn = 0.0
    if not np.isfinite(mx): mx = 1.0
    if mx - mn < 1e-9:
        mx = mn + 1e-9
    return float(mn), float(mx)


# ------- Tiled food init to avoid giant temporaries -------
def _init_food_tiled(self):
    lulc = self.lulc.values.astype(np.int32, copy=False)
    h, w = lulc.shape
    base = np.zeros((h, w), dtype=np.float32)
```

```
        # set base by LULC
        for cls, e in DEFAULT_LULC_ENERGY.items():
            mask = (lulc == cls)
            base[mask] = float(e)

        # NDVI + NDWI in tiles
        ndvi = self.ndvi.values  # keep original dtype; cast per block
        ndwi = self.ndwi.values
        for r0 in range(0, h, TILE):
            r1 = min(h, r0 + TILE)
            for c0 in range(0, w, TILE):
                c1 = min(w, c0 + TILE)

                ndvi_block = ndvi[r0:r1, c0:c1].astype(np.float32, copy=False)
                np.clip(ndvi_block, 0.0, 1.0, out=ndvi_block)
                base[r0:r1, c0:c1] += ENERGY_SCALARS["ndvi_gain_scale"] * ndvi_block

                ndwi_block = ndwi[r0:r1, c0:c1].astype(np.float32, copy=False)
                np.clip(ndwi_block, 0.0, 1.0, out=ndwi_block)
                base[r0:r1, c0:c1] += ENERGY_SCALARS["ndwi_gain_scale"] * ndwi_block


# Fragmentation composite (normalized per-layer, then mean)
# precompute min/max per layer

        if self.frag:
                        stats = {}
            for name, arr in self.frag.items():
                stats[name] = self._min_max_tiled(arr)

            for r0 in range(0, h, TILE):
                r1 = min(h, r0 + TILE)
                for c0 in range(0, w, TILE):
                    c1 = min(w, c0 + TILE)
                    comps = []
                    for name, arr in self.frag.items():
                        mn, mx = stats[name]
                        block = arr.values[r0:r1, c0:c1].astype(np.float32, copy=False)
                        block = (block - mn) / (mx - mn)
                        comps.append(block)
                    if comps:
                        frag_idx = np.nanmean(np.stack(comps, axis=0), axis=0).astype(np.float32)
                        base[r0:r1, c0:c1] += ENERGY_SCALARS["fragmentation_gain_scale"] * frag_idx

        self.food_base = base
        self.food_stock = base.copy()
        self.last_renewal = None

    def renew_food_if_due(self, now):
        if self.last_renewal is None:
            self.last_renewal = now
            return
        if (now - self.last_renewal).days >= FOOD_RENEW_DAYS:
```

```
            self.food_stock = self.food_base.copy()
            self.last_renewal = now

    def sample_cell(self, lon, lat):
        r, c = rowcol(self.transform, lon, lat)
        r, c = int(r), int(c)
        if not within_bounds(self.lulc.shape, r, c):
            return {}
        out = {
            "LULC": float(self.lulc.values[r, c]),
            "NDVI": float(np.clip(self.ndvi.values[r, c], 0, 1)),
            "NDWI": float(np.clip(self.ndwi.values[r, c], 0, 1)),
            "FOOD": float(self.food_stock[r, c]),
        }
        for k, arr in self.frag.items():
            out[k] = float(arr.values[r, c])
        out["row"], out["col"] = r, c
        return out

    def neighbors_within(self, lon, lat, radius_cells=2):
        r0, c0 = rowcol(self.transform, lon, lat)
        r0, c0 = int(r0), int(c0)
        out = []
        for dr in range(-radius_cells, radius_cells + 1):
            for dc in range(-radius_cells, radius_cells + 1):
                rr, cc = r0 + dr, c0 + dc
                if within_bounds(self.lulc.shape, rr, cc):
                    x, y = xy(self.transform, rr, cc)
                    out.append((rr, cc, x, y))
        return out

    def consume_food(self, r, c, amount):
        if within_bounds(self.food_stock.shape, r, c):
            self.food_stock[r, c] = max(0.0, self.food_stock[r, c] - amount)

# ================ RSF Model (fixed + random) =========

@dataclass
class RSFModel:
    """Per-bird RSF coefficients (already combined fixed + random)."""
    coeffs: Dict[str, float]

    def _get(self, name: str) -> float:
        aliases = [name, name.upper(), name.lower(), name.capitalize()]
        if name == "(Intercept)":
            aliases += ["Intercept", "(intercept)", "intercept"]
        for a in aliases:
            if a in self.coeffs:
                return self.coeffs[a]
        return 0.0

    def linear_predictor(self, cov: Dict[str, float]) -> float:
        lp = 0.0
        lp += self._get("(Intercept)")
```

```
        ndwi = cov.get("NDWI")
        if ndwi is not None and np.isfinite(ndwi):
            lp += self._get("NDWI") * ndwi

        lulc = cov.get("LULC")
        if lulc is not None and np.isfinite(lulc):
            try:
                code = int(lulc)
                for key in (f"LULC_class{code}", f"LULC_{code}"):
                    if key in self.coeffs:
                        lp += self.coeffs[key]
                        break
            except Exception:
                pass

        cov_upper = {k.upper(): v for k, v in cov.items()}
        for k, beta in self.coeffs.items():
            ku = k.upper()
            if ku in {"(INTERCEPT)", "INTERCEPT", "NDWI"} or ku.startswith("LULC_") or
ku.startswith("LULC_CLASS"):
                continue
            if ku in cov_upper and np.isfinite(cov_upper[ku]):
                lp += beta * cov_upper[ku]

        return float(lp)
# ================ RSF Helpers ========================

def _read_fixed_effects(path: str) -> Dict[str, float]:
    fx = pd.read_csv(path)
    fx.columns = [c.strip().lower() for c in fx.columns]
    if not {"term", "coef"}.issubset(fx.columns):
        raise ValueError("Fixed-effects file must have columns: term, coef")
    fx = fx.dropna(subset=["term", "coef"])
    out = {}
    for _, r in fx.iterrows():
        term = str(r["term"]).strip()
        try:
            val = float(r["coef"])
        except Exception:
            continue
        out[term] = val
    if "intercept" in out and "(Intercept)" not in out:
        out["(Intercept)"] = out.pop("intercept")
    return out

def _wide_or_pivot_randoms(df: pd.DataFrame) -> pd.DataFrame:
    cols_lower = [c.lower() for c in df.columns]
    if {"bird_id", "term", "coef"}.issubset(cols_lower):
        bid = next(c for c in df.columns if c.lower() == "bird_id")
        term = next(c for c in df.columns if c.lower() == "term")
        coef = next(c for c in df.columns if c.lower() == "coef")
        wide = df.pivot_table(index=bid, columns=term, values=coef, aggfunc="first").reset_index()
        wide.columns.name = None
        return wide.set_index(bid)
    if "bird_id" in cols_lower:
```

```
        bid = next(c for c in df.columns if c.lower() == "bird_id")
        return df.set_index(bid)
    raise ValueError("Random-effects file must include a 'bird_id' column or long format (bird_id,
term, coef).")

def _combine_fixed_random_for_bird(fixed: Dict[str, float], rrow: pd.Series) -> Dict[str, float]:
    """
    Generic combiner:
    - If rrow already contains full coefficients (e.g., '(Intercept)', 'NDWI', 'LULC_class20', ...),
      use those values directly. Fixed terms fill any missing keys.
    - Else if rrow has 're_*' columns, add them as deltas to matching fixed terms.
    """
    cols_lower = {str(c).lower() for c in rrow.index}
    has_full_intercept = any(c in cols_lower for c in ["(intercept)", "intercept"])
    has_full_ndwi = "ndwi" in cols_lower
    has_any_lulc = any(str(k).lower().startswith("lulc") for k in rrow.index)

    if has_full_intercept or has_full_ndwi or has_any_lulc:
        combined = {}
        for k, v in rrow.items():
            if pd.isna(v): continue
            if isinstance(v, (int, float, np.floating)):
                combined[str(k)] = float(v)
        if "Intercept" in combined and "(Intercept)" not in combined:
            combined["(Intercept)"] = combined.pop("Intercept")
        # carry over fixed for missing terms
        for fk, fv in fixed.items():
            if fk not in combined:
                combined[fk] = fv
        return combined

    combined = dict(fixed)
    for k, v in rrow.items():
        if pd.isna(v): continue
        k_low = str(k).lower()
        if not k_low.startswith("re_"):
            continue
        base = k_low[3:]
        if base in {"intercept", "(intercept)"}:
            key = "(Intercept)"
        else:
            key = None
            for fk in fixed.keys():
                if fk.lower() == base:
                    key = fk
                    break
            if key is None and base.startswith("lulc"):
                key = base
        try:
            delta = float(v)
        except Exception:
            continue
        if key is None:
            continue
        combined[key] = combined.get(key, 0.0) + delta
```

```
    if "Intercept" in combined and "(Intercept)" not in combined:
        combined["(Intercept)"] = combined.pop("Intercept")
    return combined

# ================ Load Birds (fixed + random) =========

def load_birds_from_points(points_file: str,
                           fixed_effects_csv: str,
                           random_effects_csv: str,
                           step_minutes: int):
    # --- nests ---
    ext = os.path.splitext(points_file)[1].lower()
    df = pd.read_excel(points_file, dtype=str) if ext in ['.xls', '.xlsx'] else pd.read_csv(points_file,
dtype=str)
    df.columns = [c.strip().lower().replace('.', '_').replace(' ', '_') for c in df.columns]

    if {'individual_local_identifier', 'location_long', 'location_lat', 'timestamp'}.issubset(df.columns):
        df['timestamp'] = pd.to_datetime(df['timestamp'], errors='coerce')
        df = df.sort_values(['individual_local_identifier', 'timestamp'])
        nests = df.groupby('individual_local_identifier').first().reset_index()
        nests = nests.rename(columns={'individual_local_identifier': 'bird_id',
                                      'location_long': 'lon',
                                      'location_lat': 'lat'})
        ensure_dir('./data')
        nests[['bird_id', 'lon', 'lat']].to_csv('./data/nests_auto.csv', index=False)
        print(f" Auto-created ./data/nests_auto.csv with {len(nests)} birds:")
    elif {'bird_id', 'lon', 'lat'}.issubset(df.columns):
        nests = df.copy()
        print(f" Using provided nests CSV with {len(nests)} birds.")
    else:
        raise ValueError(" Input must contain either Movebank or preprocessed columns.")

    nests['bird_id'] = (nests['bird_id'].astype(str).str.strip()
                        .str.replace(r'\.0$', '', regex=True)
                        .str.upper().str.replace(r'\s+', '', regex=True))

    # --- fixed (global) ---
    fixed = _read_fixed_effects(fixed_effects_csv)

    # --- randoms / per-bird table ---
    rnd = pd.read_csv(random_effects_csv)
    rnd.columns = [c.strip() for c in rnd.columns]
    bid_col = next((c for c in rnd.columns if c.lower() == "bird_id"), None)
    if bid_col is None and {"bird_id", "term", "coef"}.issubset({c.lower() for c in rnd.columns}):
        bid_col = next(c for c in rnd.columns if c.lower() == "bird_id")
    if bid_col is None:
        raise ValueError("Random-effects file must include a 'bird_id' column (wide or long).")
    rnd[bid_col] = (rnd[bid_col].astype(str).str.strip()
                    .str.replace(r'\.0$', '', regex=True).str.upper()
                    .str.replace(r'\s+', '', regex=True))
    rnd_wide = _wide_or_pivot_randoms(rnd)  # index=bird_id

    print("\n Matching IDs:")
    print("  Nests IDs:", nests['bird_id'].tolist()[:5])
```

```
    print("  RSF IDs:", list(rnd_wide.index[:5]))
    print(f"  Total overlap: {len(set(nests['bird_id']) & set(rnd_wide.index))} birds")

    birds = []
    for _, row in nests.iterrows():
        bid = str(row['bird_id']).strip().upper()
        if bid not in rnd_wide.index:
            print(f"Skipping bird {bid} — no RSF row found.")
            continue

        rrow = rnd_wide.loc[bid]
        combined = _combine_fixed_random_for_bird(fixed, rrow)
        rsf_model = RSFModel(coeffs=combined)

        b = BirdAgent(
            bird_id=bid,
            rsf=rsf_model,
            start_lon=float(row['lon']),
            start_lat=float(row['lat']),
            step_minutes=step_minutes,
            step_mu_m=float(row.get('step_mu_m', 300.0)),
            step_sigma_m=float(row.get('step_sigma_m', 150.0)),
            turn_kappa=float(row.get('turn_kappa', 2.0)),
        )
        birds.append(b)

    print(f" Loaded {len(birds)} birds for simulation.")
    return birds


# =============== Bird Agent ==========================

@dataclass
class BirdAgent:
    bird_id: str
    rsf: RSFModel
    start_lon: float
    start_lat: float
    step_minutes: int = STEP_MINUTES_DEFAULT
    alive: bool = True

    lon: float = field(init=False)
    lat: float = field(init=False)
    bearing_deg: float = field(default_factory=lambda: random.uniform(0, 360))

    energy_cum: float = 0.0
    day_energy_gain: float = 0.0
    day_energy_cost: float = 0.0
    bred: bool = False
    step_count: int = 0

    step_mu_m: float = 300.0
    step_sigma_m: float = 150.0
    turn_kappa: float = 2.0
    daily_net_history: List[float] = field(default_factory=list)
```

```
def __post_init__(self):
    self.lon = self.start_lon
    self.lat = self.start_lat

def reset_to_nest(self):
    """Snap bird back to its nest (for daily/yearly site fidelity)."""
    self.lon = self.start_lon
    self.lat = self.start_lat
    # Optionally reset bearing:
    # self.bearing_deg = random.uniform(0, 360)

def step(self, env: Environment, now: datetime) -> Dict[str, float]:
    """
    New step logic:
      - Sample a real step length (m) and turning deviation (von Mises).
      - Compute new (lon,lat) by great-circle move.
      - Apply RSF + site fidelity + (optional) smooth nest pull every 2h in daytime.
      - Keep daily reset elsewhere (unchanged).
    """
    # ---- choose behavioural mode ----
    mode = "breeding" if self.bred else "nonbreeding"
    p = BEHAVIOR_PARAMS[mode]

    # If user provided per-bird overrides, keep them as fallbacks
    mu = float(getattr(self, "step_mu_m", p["step_mu_m"])) or p["step_mu_m"]
    sg = float(getattr(self, "step_sigma_m", p["step_sigma_m"])) or p["step_sigma_m"]
    kappa = float(getattr(self, "turn_kappa", p["turn_kappa"])) or p["turn_kappa"]

    # ---- sample distance and bearing ----
    dist_m = max(0.0, np.random.normal(mu, sg))
    turn_rad = np.random.vonmises(0.0, kappa)  # deviation from current bearing
    new_bearing = (self.bearing_deg + math.degrees(turn_rad)) % 360.0

    # ---- propose new coordinate ----
    new_lon, new_lat = move_along_bearing(self.lon, self.lat, dist_m, new_bearing)

    # ---- pull back softly if foraging too far from nest (soft arena, not a hard cap) ----
    nest_dist_after = haversine_m(self.start_lon, self.start_lat, new_lon, new_lat)
    arena = p["max_foraging_radius_m"]
    if arena > 0:
        # exponential penalty beyond the arena radius
        if nest_dist_after > arena:
            over = nest_dist_after - arena
            penalty = math.exp(-0.001 * over)  # gentle decay; tweak if needed
            # pull slightly toward nest by shortening the step
            adj_dist = dist_m * penalty
            new_lon, new_lat = move_along_bearing(self.lon, self.lat, adj_dist, new_bearing)
            nest_dist_after = haversine_m(self.start_lon, self.start_lat, new_lon, new_lat)

    # ---- read environment at proposed location ----
    cov = env.sample_cell(new_lon, new_lat)

    # --- Handle out-of-bounds movement by bouncing back ---
```

```
if not cov:
    # Try turning ~180° (reverse) plus a small random deviation
    new_bearing = (self.bearing_deg + 180 + np.random.uniform(-30, 30)) % 360

    # Try up to 3 reflection attempts before giving up
    for _ in range(3):
        new_lon, new_lat = move_along_bearing(self.lon, self.lat, dist_m, new_bearing)
        cov = env.sample_cell(new_lon, new_lat)
        if cov:
            break
        # if still out of bounds, pick a new random bearing
        new_bearing = np.random.uniform(0, 360)

    # if still invalid after retries, stay in place
    if not cov:
        new_lon, new_lat = self.lon, self.lat
        cov = env.sample_cell(new_lon, new_lat)
        if not cov:
            return {}


# ---- RSF linear predictor ----
lp = self.rsf.linear_predictor(cov)

# ---- site fidelity ----
if self.bred:
    home_weight = math.exp(-HOME_DECAY * nest_dist_after)
else:
    home_weight = 1.0  # no site fidelity for non-breeders
score = lp + math.log(home_weight + 1e-12)


# ---- smoother nest pull: breeders vs non-breeders ----
if now is not None and (DAY_START_HOUR <= now.hour <= DAY_END_HOUR):
    minutes_since_midnight = now.hour * 60 + now.minute
    pull_tick = (minutes_since_midnight % NEST_PULL_INTERVAL_MIN == 0)

    # --- Breeders: apply nest pull every 2 hours ---
    if self.bred and NEST_PULL_ACTIVE_ONLY_IF_BRED and pull_tick:
        alpha = BEHAVIOR_PARAMS["breeding"]["nest_pull_alpha"]
        if alpha > 0:
            pull_weight = math.exp(-alpha * nest_dist_after)
            score += math.log(pull_weight + 1e-12)

    # Non-breeders: explicitly no nest pull
    elif (not self.bred) and pull_tick:
        pass


# ---- energy & bookkeeping (unchanged logic) ----
food_gain = cov.get("FOOD", 0.0)
take = 0.25 * food_gain  # same as before
env.consume_food(int(cov["row"]), int(cov["col"]), take)
```

```
        flying_cost = ENERGY_SCALARS["distance_cost_scale"] * dist_m
        slices_per_day = (24 * 60) / self.step_minutes
        physio_cost = (ENERGY_SCALARS["baseline_daily_units"] / slices_per_day) *
ENERGY_REQUIREMENT["physio"] / REQUIREMENT_TOTAL
        foraging_cost = (ENERGY_SCALARS["baseline_daily_units"] / slices_per_day) *
ENERGY_REQUIREMENT["foraging"] / REQUIREMENT_TOTAL

        gain = take
        cost = flying_cost + physio_cost + foraging_cost
        net = gain - cost

        # commit new position
        self.lon, self.lat = new_lon, new_lat
        self.bearing_deg = new_bearing
        self.day_energy_gain += gain
        self.day_energy_cost += cost
        self.energy_cum += net
        self.step_count += 1


        return {
            "lon": self.lon, "lat": self.lat, "dist_m": dist_m, "bearing": new_bearing,
            "gain": gain, "cost": cost, "net": net, "lulc": cov.get("LULC", np.nan)
        }
    def end_of_day_update(self):
        day_req = ENERGY_SCALARS["baseline_daily_units"]
        net_today = self.day_energy_gain - self.day_energy_cost


        # Track daily net for window-based breeding rule
        self.daily_net_history.append(float(net_today))
        if len(self.daily_net_history) > 120:  # keep a bounded history
            self.daily_net_history = self.daily_net_history[-120:]

        met_fraction = net_today / max(day_req, 1e-6)

        died = self.energy_cum < (DEATH_THRESHOLD * -day_req)  # unchanged death test

        # Reset daily tallies
        g, c = self.day_energy_gain, self.day_energy_cost
        self.day_energy_gain = 0.0
        self.day_energy_cost = 0.0

        if died:
            self.bred = False
            self.alive = False

        bred_today = False
        return g, c, met_fraction, died, bred_today



# ================ Simulation ============================

@dataclass
```

```
class SimulationConfig:
    start_year: int = 2012
    end_year: int = 2022
    season_start_mmdd: Tuple[int, int] = (3, 1)   # March 1
    season_end_mmdd: Tuple[int, int] = (8, 31)    # August 31
    step_minutes: int = STEP_MINUTES_DEFAULT
    neighbor_radius_cells: int = 2
    rng_seed: int = 42

@dataclass
class Scenario:
    lulc_percent_delta: Dict[int, float] = field(default_factory=dict)
    fragmentation_multiplier: float = 1.0

class IBM:
    def __init__(self, env: Environment, birds: List[BirdAgent], cfg: SimulationConfig, scenario:
Scenario):
        self.env = env
        self.birds = birds
        self.cfg = cfg
        self.scenario = scenario
        random.seed(cfg.rng_seed)
        np.random.seed(cfg.rng_seed)

    def apply_scenario(self):
        if not self.scenario.lulc_percent_delta:
            return
        lulc = self.env.lulc.values.copy().astype(int)
        unique, counts = np.unique(lulc, return_counts=True)
        uniq_list = unique.tolist()
        count_map = dict(zip(uniq_list, counts.tolist()))

        target_counts = count_map.copy()
        for cls, pct in self.scenario.lulc_percent_delta.items():
            base = count_map.get(cls, 0)
            delta = int(round(base * (pct / 100.0)))
            target_counts[cls] = max(0, base + delta)

        donors = [c for c in uniq_list if c not in self.scenario.lulc_percent_delta]
        for cls, tgt in target_counts.items():
            cur = int((lulc == cls).sum())
            if tgt > cur:
                need = tgt - cur
                donor_counts = {d: int((lulc == d).sum()) for d in donors}
                donor_total = sum(donor_counts.values())
                if donor_total <= 0:
                    continue
                for d in donors:
                    if need <= 0: break
                    give = int(round(need * donor_counts[d] / donor_total))
                    if give <= 0: continue
                    rr, cc = np.where(lulc == d)
                    if len(rr) == 0: continue
                    take_n = min(give, len(rr), need)
                    idx = np.random.choice(len(rr), size=take_n, replace=False)
```

```
                lulc[rr[idx], cc[idx]] = cls
                need -= take_n

        self.env.lulc.values[:] = lulc
        if self.scenario.fragmentation_multiplier != 1.0:
            for k, arr in self.env.frag.items():
                arr.values[:] = arr.values * self.scenario.fragmentation_multiplier
        # re-init food with tiles
        self.env._init_food_tiled()

    def run(self, years: List[int], output_dir: str):
        ensure_dir(output_dir)
        for year in years:
            print(f"\n=== Simulating {year} ===")
            self._run_year(year, output_dir)


    def _run_year(self, year: int, output_dir: str):
        self.env._init_food_tiled()
        for b in self.birds:
            b.alive = True
            b.energy_cum = 0.0
            b.day_energy_cost = 0.0
            b.day_energy_gain = 0.0
            b.bred = False
            b.step_count = 0
            b.reset_to_nest()   # ensure same starting point every year

        yout_dir = os.path.join(output_dir, "steps", str(year))
        fig_dir = os.path.join(output_dir, "figures", str(year))
        sum_dir = os.path.join(output_dir, "summaries")
        ensure_dir(yout_dir); ensure_dir(fig_dir); ensure_dir(sum_dir)

        season_start = datetime(year, *self.cfg.season_start_mmdd)
        season_end = datetime(year, *self.cfg.season_end_mmdd, 23, 59)
        step = timedelta(minutes=self.cfg.step_minutes)
        t = season_start

        writers = {}
        for b in self.birds:
            fpath = os.path.join(yout_dir, f"{b.bird_id}.csv")
            writers[b.bird_id] = open(fpath, "w", encoding="utf-8")
            writers[b.bird_id].write("timestamp,step,lon,lat,dist_m,bearing,gain,cost,net,lulc\n")

        day_counter = 0
        breed_checked = set()
        deaths = set()

        steps_total = int(((season_end - season_start).total_seconds() // (self.cfg.step_minutes * 60)) + 1)
        pbar = tqdm(total=steps_total)

        last_day = season_start.date()  # to detect day changes

        while t <= season_end:
            # daytime-only movement gate
```

```
    if (t.hour < DAY_START_HOUR) or (t.hour > DAY_END_HOUR):
        t += step
        pbar.update(1)
        continue

    self.env.renew_food_if_due(t)

    if t.date() != last_day:
        for b in self.birds:
            if not b.alive:
                continue
            # Only confirmed breeders return to nest each new day
            if b.bred:
                b.reset_to_nest()
            # Non-breeders continue from their current position (no reset)
        last_day = t.date()

    for b in self.birds:
        if not b.alive:
            continue
        rec = b.step(self.env, t)
        if rec:
            writers[b.bird_id].write(
                f"{t.isoformat()},{b.step_count},"
                f"{rec['lon']:.6f},{rec['lat']:.6f},{rec['dist_m']:.2f},{rec['bearing']:.2f},"
                f"{rec['gain']:.4f},{rec['cost']:.4f},{rec['net']:.4f},"
                f"{int(rec['lulc']) if not math.isnan(rec['lulc']) else ''}\n"
            )


    # end-of-day energy updates: check at last active slot; since nights are skipped

    if t.hour == DAY_END_HOUR:
        day_counter += 1

        # === End-of-day energy update ===
        for b in self.birds:
            if not b.alive:
                continue
            g, c, frac, died, _ = b.end_of_day_update()
            if died:
                b.alive = False
                deaths.add(b.bird_id)

        # === Breeding decision after window ===
        # Evaluate only after enough days have passed
        if day_counter >= BREED_WINDOW_DAYS:
            for b in self.birds:
                if not b.alive:
                    continue
                # sustained energy surplus over last BREED_WINDOW_DAYS
                b.bred = self._should_breed(b)
    t += step
    pbar.update(1)
pbar.close()
```

```
    for f in writers.values():
        f.close()

    rows = []
    for b in self.birds:
        rows.append({"bird_id": b.bird_id,
                     "alive": b.alive,
                     "bred": b.bred,
                     "energy_cum": round(b.energy_cum, 3),
                     "step": b.step_count})
    df = pd.DataFrame(rows)
    df.to_csv(os.path.join(sum_dir, f"summary_{year}.csv"), index=False)

    self._plot_year_tracks(year, yout_dir, fig_dir)
    self._plot_year_energy(year, sum_dir, fig_dir)
    self._plot_year_tracks_by_status(year, yout_dir, fig_dir, self.birds)


def _plot_year_tracks(self, year: int, yout_dir: str, fig_dir: str):
    plt.figure(figsize=(8, 8))
    for birdfile in os.listdir(yout_dir):
        if not birdfile.endswith('.csv'):
            continue
        df = pd.read_csv(os.path.join(yout_dir, birdfile))
        if df.empty:
            continue
        plt.plot(df['lon'], df['lat'], linewidth=0.6, alpha=0.7)
    plt.title(f"Tracks — {year}")
    plt.xlabel("Longitude")
    plt.ylabel("Latitude")
    plt.tight_layout()
    out = os.path.join(fig_dir, f"tracks_{year}.png")
    plt.savefig(out, dpi=200)
    plt.close()

def _plot_year_energy(self, year: int, sum_dir: str, fig_dir: str):
    df = pd.read_csv(os.path.join(sum_dir, f"summary_{year}.csv"))
    alive = int(df['alive'].sum())
    bred = int(df['bred'].sum())
    plt.figure(figsize=(6, 4))
    plt.bar(["Alive", "Bred"], [alive, bred])
    plt.title(f"Year {year}: Alive & Bred counts")
    plt.tight_layout()
    out = os.path.join(fig_dir, f"energy_summary_{year}.png")
    plt.savefig(out, dpi=200)
    plt.close()

def _should_breed(self, b: BirdAgent) -> bool:
    if not b.alive:
        return False
    window = BREED_WINDOW_DAYS
    day_req = ENERGY_SCALARS["baseline_daily_units"]
    if len(b.daily_net_history) < window:
        return False
```

```
        recent = b.daily_net_history[-window:]
        window_net = float(sum(recent))

        required = BREED_MIN_EXCESS * day_req * window
        return window_net >= required

    def _plot_year_tracks_by_status(self, year: int, yout_dir: str, fig_dir: str, birds: List[BirdAgent]):
        breeders, nonbreeders = [], []
        for b in birds:
            fpath = os.path.join(yout_dir, f"{b.bird_id}.csv")
            if not os.path.exists(fpath) or os.path.getsize(fpath) == 0:
                continue
            df = pd.read_csv(fpath)
            if b.bred:
                breeders.append(df)
            else:
                nonbreeders.append(df)

        # --- Breeders ---
        if breeders:
            plt.figure(figsize=(8, 8))
            for df in breeders:
                plt.plot(df["lon"], df["lat"], linewidth=0.6, alpha=0.7)
            plt.title(f"Breeder Tracks — {year}")
            plt.xlabel("Longitude")
            plt.ylabel("Latitude")
            plt.tight_layout()
            plt.savefig(os.path.join(fig_dir, f"tracks_breeders_{year}.png"), dpi=200)
            plt.close()

        # --- Non-breeders ---
        if nonbreeders:
            plt.figure(figsize=(8, 8))
            for df in nonbreeders:
                plt.plot(df["lon"], df["lat"], linewidth=0.6, alpha=0.7)
            plt.title(f"Non-breeder Tracks — {year}")
            plt.xlabel("Longitude")
            plt.ylabel("Latitude")
            plt.tight_layout()
            plt.savefig(os.path.join(fig_dir, f"tracks_nonbreeders_{year}.png"), dpi=200)
            plt.close()


# ================ Main Runner ============================

# Single year runner : for test
def run_simulation():
    env = Environment(LULC_PATH, NDVI_PATH, NDWI_PATH, FRAG_PATHS)
    birds = load_birds_from_points(POINTS_PATH, FIXED_RSF_PATH, RANDOM_RSF_PATH,
STEP_MINUTES_DEFAULT)
    cfg = SimulationConfig()
    scenario = Scenario()
    ibm = IBM(env, birds, cfg, scenario)
    ibm.apply_scenario()
    ibm.run([2012], OUTPUT_DIR)
```

```
# Multiple year runner loop
def run_simulation_multi(years: list[int]):
    """
    Run the IBM for multiple years, automatically swapping raster layers.
    Requires rasters named in the pattern:
        LC_<year>_07.tif
        NDVI_<year>_JulAug_30m.tif
        MNDWI_<year>_JulAug_30m.tif
    """
    for year in years:
        print(f"\n Starting simulation for {year}...")

        # --- Construct year-specific raster paths ---
        LULC = f"D:/4th Semester/thesis/models/GeoTiffs/New/LC_{year}_07.tif"
        NDVI = f"D:/4th
Semester/thesis/models/GeoTiffs/NDVI_NDWI/NDVI_{year}_JulAug_30m.tif"
        NDWI = f"D:/4th
Semester/thesis/models/GeoTiffs/NDVI_NDWI/MNDWI_{year}_JulAug_30m.tif"

        # --- Check raster existence before proceeding ---
        missing = False
        for f in [LULC, NDVI, NDWI]:
            if not os.path.exists(f):
                print(f"Missing raster file: {f}")
                missing = True
        if missing:
            print("   Skipping this year.\n")
            continue

        # --- Initialize environment + birds ---
        env = Environment(LULC, NDVI, NDWI, FRAG_PATHS)
        birds = load_birds_from_points(POINTS_PATH, FIXED_RSF_PATH, RANDOM_RSF_PATH,
STEP_MINUTES_DEFAULT)
        cfg = SimulationConfig(start_year=year, end_year=year)
        scenario = Scenario()
        ibm = IBM(env, birds, cfg, scenario)
        ibm.apply_scenario()

        # --- Run simulation for this year ---
        out_dir_year = os.path.join(OUTPUT_DIR, str(year))
        ibm.run([year], out_dir_year)
        print(f" Finished simulation for {year}. Results saved to {out_dir_year}\n")

if __name__ == "__main__":
        run_simulation_multi([2018])
```