Modelling Species Distribution with R-INLA (and INLABRU?)

*Barber, Conesa, Gomez-Rubio, Lopez-Quilez,
Martinez-Minaya, Paradinas, Pennino*

RUBRICA Y LIMITACIÓN DE RESPONSABILIDAD

Este libro está escrito por mis amigos y cuando lo acabemos nos vamos a correr una buena juerga.... te apuntas???

Firmado:

J. Xavier Barber Vallés

28996841M

Contents

Spatial (areal & geostatistical) SDM (Antonio y Maria)   13

Preferential Sampling in the context of SDM (David)   17

Excess of zeroes in the context of SDM (Joaquin)   19

Barriers and non-stationarity models (David y Joaquin)   21

Misalignment Data (Xavi)   23

Presence-only (INLABRU) (Xavi y Iosu)   25

List of Figures

List of Tables

Preface

Este es nuestro hijo, después de tiempo intentando concebirlo, finalmente estamos en la primera ecografía (guimo a Maria i Iosu, ¿os acordais? jejeje)

Así que ahora toca currar, unos más que otros, y a mi me toca ir pidióos cada dia un poco más de esfuerzo.

Yo sé que lo vamos a conseguir, así que sin prisa pero sin pausa... que nos queda un largo camino por andar juntos para ver como nuestra criatura (o engendro) sale adelante en tiempo y forma

1

Species Distribution Models: a recent history (María)

Capítulo en el que María nos hará un repaso de los SDM y la necesidad de seguir avanzando por este camino

1.1 The Unavoidable Globalization of Knowledge

Biodiversity, that magical mosaic of life that unfolds before us at all scales and in every corner of the planet, has always been a subject of fascination and study. From the earliest naturalists who roamed forests and deserts, to the modern biologists using satellites and algorithms for their research, there has been an insatiable desire to understand how and why species distribute themselves in space and time.

This is where the importance of Species Distribution Models (SDM) comes into play. These models aim not only to understand the current geographical distribution of species but also to predict how they might react and move in response to changes in the environment, such as climate change or habitat transformation due to human activities.

The use of SDMs has increased over the past few decades due to the growing need to make informed decisions in areas like conservation, resource manage-

ment, and territorial planning. The implications extend beyond the academic world; public policies, industries, and society at large benefit directly from the knowledge these models provide.

So, if SDMs are such a powerful and necessary tool, why the insistence on communicating about them in English?

The answer lies in the unavoidable globalization of knowledge. We live in a world where information flows faster than ever, and often, the lingua franca of science is English. By publishing and communicating findings in this language, a global reach is achieved, allowing scientists from all over the world to access, understand, critique, and build upon that foundation.

However, this doesn't diminish the value of disseminating in other languages. In fact, it's essential to bring science closer to non-English speaking communities. But, to establish international and cooperative dialogue, the adoption of English becomes a crucial tool.

In this book, we will not only delve into the theory and practice of SDMs but also address their global significance and how these models intertwine in an interconnected world. Thus, through the lens of English, we expand the boundaries of knowledge and strengthen international collaboration towards a common goal: understanding and preserving planetary biodiversity.

2

Bayesian Inference: Integrated Nested Laplace Approximation & Variational Bayes. (Virgilio)

Aquí Virgilio nos marcará el camino de la notación, utilizando como base su libro Bayesian Inference nos hará una intro a la Bayesiana para que los lectores entinedan de qué va esto y que son eso de los parámetros/hiperparámetros y porque "en Bayesiano" y porque con "INLA".

2.1   Introduction: Navigating the Bayesian Landscape

In the vast universe of statistical methodologies, Bayesian Inference stands as a beacon, illuminating the path toward a deeper understanding of data and the underlying processes that generate them. This probabilistic framework, named after the Reverend Thomas Bayes, offers a fundamentally different approach to understanding uncertainty compared to traditional frequentist methods.

At its core, Bayesian Inference elegantly melds prior knowledge or beliefs with observed data, resulting in updated or "posterior" beliefs about the phenomena in question. This continuous cycle of learning, updating, and refining

stands in stark contrast to the rigid hypothesis testing of classical statistics. Instead of isolated point estimates, Bayesian methods provide entire distributions of probabilities, enabling richer interpretations and more nuanced decision-making.

Despite its age-old foundations, Bayesian Inference has witnessed a renaissance in the late 20th and early 21st centuries. Advances in computational power and algorithms, particularly Markov chain Monte Carlo (MCMC) methods, have made intricate Bayesian models computationally tractable. As a result, fields ranging from genetics to economics, and from ecology to astrophysics, are increasingly turning to Bayesian methodologies to solve complex problems.

In this book, we embark on a journey through the Bayesian landscape. From foundational principles to advanced modeling techniques, we will dissect the Bayesian paradigm, aiming to provide readers with both the theoretical underpinnings and practical tools to navigate the Bayesian world confidently.

Join us as we delve into the intuitive and powerful world of Bayesian Inference, a realm where data, prior beliefs, and probability converge to shape our understanding of the world around us.

3

SDM with GLM and GAM (covariates) (María y David)

Aquí David utilizará su amplio material de sus cursos para crear un capítulo donde haga una introducción a los GLM y GAM en el ámbito de los SDM e INLA.

3.1 Charting the Ecological Landscape with GLM and GAM

Species Distribution Models (SDMs) have long been a cornerstone in ecological research, conservation planning, and environmental decision-making. These models, which aim to correlate the occurrence or abundance of species with environmental predictors, are essential tools for forecasting how species might respond to environmental changes, be they natural shifts or anthropogenic impacts.

Within the broad spectrum of available modeling techniques, Generalized Linear Models (GLM) and Generalized Additive Models (GAM) stand out as both foundational and transformative. While GLM extends the familiar linear regression to accommodate non-normally distributed response data, GAM further generalizes the relationship between predictors and responses by al-

lowing for non-linearities via smooth functions. These characteristics make them particularly suited for capturing the complexities inherent in ecological datasets.

However, with their strengths come intricacies in interpretation, model selection, and diagnostics. Striking a balance between model complexity and ecological interpretability is an art as much as it is a science.

This book is conceived as a comprehensive guide to understanding and applying GLM and GAM in the context of SDMs. We will navigate through the theoretical underpinnings of these models, their implementation, and the challenges and solutions associated with their use. Real-world examples and case studies will illuminate the path, ensuring that the reader not only grasps the mathematical nuances but also appreciates their practical implications.

Embark with us on this exploration of the confluence of statistics and ecology, where GLM and GAM become powerful lenses through which we can better understand the distribution of life on our planet.

3.2 GLM

```r
# Load required libraries
#install.packages("stats")
library(stats)

# Generate a hypothetical dataset
set.seed(123)
n <- 100
data <- data.frame(
  presence = rbinom(n, 1, 0.5),
```

```
  env1 = rnorm(n),
  env2 = rnorm(n)
)

# Fit a GLM
glm_model <- glm(presence ~ env1 + env2, data=data, family="binomial")

# Summary of the GLM
summary(glm_model)
```

```
##
## Call:
## glm(formula = presence ~ env1 + env2, family = "binomial", data = data)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.1426     0.2040  -0.699   0.4845
## env1         -0.3646     0.2186  -1.668   0.0954 .
## env2         -0.0719     0.2189  -0.328   0.7426
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 138.27  on 99  degrees of freedom
## Residual deviance: 135.34  on 97  degrees of freedom
## AIC: 141.34
##
## Number of Fisher Scoring iterations: 4
```

3.3 GAM

```
# Load required libraries
#install.packages("mgcv")
library(mgcv)
```

```
## Loading required package: nlme
```

```
## This is mgcv 1.8-42. For overview type 'help("mgcv-package")'.
```

```
# Generate a hypothetical dataset (same as before)
set.seed(123)
n <- 100
data <- data.frame(
  presence = rbinom(n, 1, 0.5),
  env1 = rnorm(n),
  env2 = rnorm(n)
)

# Fit a GAM
gam_model <- gam(presence ~ s(env1) + s(env2), data=data, family="binomial")

# Summary of the GAM
summary(gam_model)
```

```
##
## Family: binomial
## Link function: logit
```

```
##
## Formula:
## presence ~ s(env1) + s(env2)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.1241     0.2034   -0.61    0.542
##
## Approximate significance of smooth terms:
##          edf Ref.df Chi.sq p-value
## s(env1)    1      1  2.781  0.0954 .
## s(env2)    1      1  0.108  0.7426
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.00925   Deviance explained = 2.12%
## UBRE = 0.41339  Scale est. = 1         n = 100
```

4

Spatial Data (data type and Maps) (Virgilio)

De nuevo VIrgilio nos deleitará con su saber sobre estadística espacial y enseñará el camino a los lectores, pero sin entrar en analizar datos, un capítulo con los conocimientos necesarios en R para poder plotear mapas chulos, que paquetes usar, como cargar un mapa que viene de GIS, etc… Todo lo que un *practicioner* debe saber antes de enfrentarse a datos espaciales como tal. ¿Cómo lo ves Virgilio?

4.1   Mapping the Confluence of Spatial Data and Species Distributions

The canvas on which species paint their intricate patterns of distribution is inherently spatial. As we attempt to model these patterns, understanding the character and nuances of spatial data becomes imperative. It is the matrix of coordinates, polygons, and raster cells that gives life to the ecological narratives of species existence and abundance.

In this chapter, "Spatial Data (data type and Maps)", we embark on a journey to explore the multifaceted world of spatial data within the context of Species Distribution Models (SDMs) using R-INLA. From vector data delineating

precise locations to raster data capturing continuous environmental gradients, each data type offers unique insights into species-habitat relationships. Beyond mere data types, the realm of spatial data also encompasses the visual wonders of maps, transforming abstract numbers into tangible representations that reveal underlying ecological stories.

R-INLA, with its cutting-edge capabilities for Bayesian inference, holds particular promise for spatial modeling. Its prowess in handling complex spatial structures ensures that our SDMs are not only statistically robust but also ecologically meaningful. As we delve into R-INLA's tools for spatial data analysis, we'll uncover techniques to seamlessly integrate data types, craft intricate maps, and ultimately harness the full potential of spatial information in our models.

Guided by practical examples, code snippets, and vivid visualizations, this chapter seeks to be both a primer for the uninitiated and a deep dive for the experienced, bridging the gap between spatial data intricacies and the powerful modeling framework of R-INLA.

Join us as we traverse this spatial dimension, illuminating the path to more informed, accurate, and insightful species distribution models.

5

Spatial (areal & geostatistical) SDM (Antonio y Maria)

Vale, y aquí Antonio nos dara una lección magistral de Geoestadística y Areal Data... no sé si algo de patrones puntuales y nos pondrá ejemplos de código *R* para que los lectores puedan hacer sus primeras predicciones o su primer modelo de BYM...

¿cómo lo veis Antonio y María?

5.1 Navigating the Spatial Landscape in Species Distribution Models

Spatial structure – a fundamental component in understanding species distributions. In ecology, it's long been acknowledged that observations in close proximity tend to be more similar than those farther apart, a concept encapsulated by Tobler's First Law of Geography. Embracing this spatial dependence is crucial, especially when modeling species distributions, as overlooking it can lead to erroneous conclusions.

In this chapter titled "Spatial (areal & geostatistical)", we dive deep into the intricacies of areal and geostatistical data and their role in Species Distribution Models (SDMs). While areal data deals with aggregate information over

13

spatial units (like polygons or regions), geostatistical data handles continuous variation across space, often relying on point-referenced data. Each has its unique challenges and opportunities when integrated into SDMs.

R-INLA, or Integrated Nested Laplace Approximations in R, emerges as a powerful tool in this spatial realm. Not only does it provide a robust framework for hierarchical modeling, but its capabilities in handling spatial structures – both areal and geostatistical – are unparalleled.

Throughout this chapter, we'll explore the theoretical underpinnings of spatial data, understand their implications in SDMs, and master the techniques to effectively harness them using R-INLA. Step-by-step examples, real-world applications, and insightful visualizations will guide our journey, ensuring a comprehensive grasp of these spatial nuances.

As we chart this spatial course, our goal is to equip readers with the knowledge and skills to construct more accurate, refined, and ecologically meaningful species distribution models. So, let's embark on this spatial exploration, where the synergy of space, species, and sophisticated modeling techniques unveils the patterns of life.

5.2 Geostatistical Species Distribution Model with R-INLA

5.3 Install & Load Necessary Packages

5.3.1 Generate Hypothetical Data

```
# Create a spatial grid
grid <- expand.grid(x=seq(0, 10, length=20), y=seq(0, 10, length=20) )
coordinates(grid) <- ~x+y
x=seq(0, 10, length=20)
y=seq(0, 10, length=20)
# Simulate some environmental predictor
grid$env1 <- with(grid, sin(x)*cos(y))


# Simulate presence-absence data with some spatial structure
set.seed(123)
grid$presence <- rbinom(nrow(grid), 1, plogis(-2 + 0.5*grid$env1 +
                  with(grid, rnorm(n=nrow(grid), mean=0, sd=0.5))))
```

5.3.2 Fit a Geostatistical Model Using R-INLA

```
# Define the formula
formula <- presence ~ 1 + env1 + f(x, y, model="matern", scale.model="range")
```

```
# Fit the model
result <- inla(formula, data=grid, family="binomial", control.predictor=list(compute=TRUE)
```

5.3.3   Plot the Results

```
# Plot the observed presences
spplot(grid, "presence", main="Observed Presence", col.regions=c("white", "red"), at=c(-0.

# Plot the predicted probabilities
pred.grid <- grid
pred.grid$predicted <- result$summary.fitted.values$mean[1:nrow(grid)]
spplot(pred.grid, "predicted", main="Predicted Presence Probability", col.regions=colorRam

# Plot the spatial random effect (this captures the spatial structure not explained by the
spatial.effect <- result$summary.random$f.x[, "mean"]
matrix.effect <- matrix(spatial.effect, nrow=sqrt(length(spatial.effect)), ncol=sqrt(lengt
image(matrix.effect, main="Spatial Effect", xaxt="n", yaxt="n", col=colorRampPalette(c("bl
```

6

Preferential Sampling in the context of SDM (David)

Y aquí ya empezamos con nuestra investigación. Lo primero que vamos a mostrar es el preferencial, ya que en la gran mayoría de an´lsiis espacial que se haga en ecología este tipo de enfoque seguro que tiene cabida.

6.1 Unraveling the Intricacies of Preferential Sampling in Species Distribution Models

When it comes to modeling species distributions, one might assume that the data at hand is random and unbiased. However, this is often not the case. Many a time, the locations of sampled data are not independent of the underlying process we wish to study. This phenomenon, where sampling locations are influenced by the very process of interest, is termed 'preferential sampling'. And in the world of Species Distribution Models (SDMs), it poses both challenges and opportunities.

In this chapter, "Preferential Sampling in the context of SDM", we will delve deep into the complexities arising from non-random sampling patterns and their implications for modeling species distributions using R-INLA. Acknowledging and addressing preferential sampling is crucial, as neglecting it can

lead to biased inferences, potentially misguiding conservation efforts, policy decisions, or scientific understanding.

But why R-INLA? The framework's advanced capabilities for Bayesian inference, combined with its adeptness in handling intricate spatial structures, make it an invaluable tool for tackling the challenges posed by preferential sampling. Through R-INLA, we can incorporate the potential dependence between the sampling locations and the underlying spatial process, leading to more accurate and robust SDMs.

Armed with real-world examples, hands-on coding exercises, and vivid visual illustrations, we aim to provide readers with a comprehensive understanding of preferential sampling and its nuances. By the end of this chapter, you will have gained both the theoretical knowledge and practical skills to recognize, assess, and account for preferential sampling in your species distribution modeling endeavors.

So, let us embark on this journey to understand the subtle dance between where we sample and what we discover, ensuring that our SDMs truly reflect the intricate tapestry of life on Earth.

7

Excess of zeroes in the context of SDM (Joaquin)

Vale, aquí Joaquin nos va a presenta run chapter en el que se muestren los ZIP y los Hurdel … ¿así te va bien Joaquín?

7.1 Navigating the Zeroes - Challenges and Opportunities in Species Distribution Models

In the intricate realm of Species Distribution Models (SDMs), data often speaks in whispers and shouts. Among the most vocal, yet perplexing aspects of this data, is the occurrence of zeroes. Far from being inconsequential, these zeroes – often representing absences or undetected presences – can dominate datasets and play a pivotal role in shaping our inferences about species distributions.

"Excess of zeroes in the context of SDM" is not just a mathematical quirk but a biological and observational reality. It can arise due to various reasons: genuine absences, detection failures, or the patchy nature of many ecological phenomena. While zeroes can provide invaluable insights, an overabundance without proper handling can lead to misinterpreted models, undermining the reliability of predictions and conservation strategies derived from them.

R-INLA, with its prowess in Bayesian modeling and its flexibility to incorporate complex data structures, emerges as a beacon in addressing this challenge. This chapter aims to elucidate the intricacies of datasets with an excess of zeroes and offers robust methodologies within the R-INLA framework to tackle them effectively.

Through a blend of theory, real-world case studies, and hands-on R-INLA tutorials, we will explore various modeling approaches such as zero-inflated models and hurdle models. These techniques, tailored for datasets with a surfeit of zeroes, ensure that our SDMs are both statistically sound and ecologically meaningful.

Join us in this deep dive into the world of zeroes, as we unravel their mysteries and harness their full potential in the quest to accurately model and understand the patterns of life on our planet.

8

Barriers and non-stationarity models (David y Joaquin)

Y ya que David y Joaquín habrán cogido ritmo por el capítulo 6 y 7, pues aunarán esfuerzos para crear el capítulo de las Barreras.

8.1 Beyond Uniformity – Embracing Barriers and Non-Stationarity in SDMs

In the vast landscapes that species inhabit, uniformity is often more the exception than the rule. The natural world is replete with barriers – be they mountain ranges, rivers, or anthropogenic constructions – that can impede or redirect species movements and distributions. Coupled with this is the concept of non-stationarity, where the relationships between species and their environment aren't consistent across space, reflecting the ever-changing tapestry of ecological interactions.

The chapter "Barriers and non-stationarity models" is dedicated to untangling these complexities in the context of Species Distribution Models (SDMs). While traditional models often operate under the assumption of spatial stationarity, real-world scenarios frequently challenge these assumptions. It becomes essential, then, to adapt and evolve our modeling techniques to more faithfully

represent the intricate dance of species across heterogeneous landscapes.

R-INLA, with its refined capabilities for Bayesian spatial modeling, offers a promising avenue for addressing these challenges. Through its flexible framework, we can incorporate barriers directly into our models, ensuring they influence species distributions in a biologically meaningful manner. Likewise, R-INLA's adeptness at handling non-stationary relationships ensures that our models remain sensitive to the varying dynamics between species and their environment across different spatial scales.

In this chapter, theory meets practice. Supported by empirical examples, hands-on exercises, and vivid illustrations, we will guide readers through the nuances of modeling barriers and non-stationarity using R-INLA. By journey's end, you'll be equipped with a richer understanding and a robust toolkit to capture the multifaceted realities of species distributions in our ever-diverse world.

Let us embark on this exploration, where barriers become more than just obstacles, and where changing spatial dynamics are not challenges but opportunities to deepen our ecological insights.

9

Misalignment Data (Xavi)

Pues aquí yo (Xavi, por si alguien tiene duda) desempolvaré el código y lo actulizaré así como mostreros de forma clara cómo se puede esquivar esto del misaligment.

9.1 The Challenge of Misalignment - Navigating Data Inconsistencies in SDMs

In the intricate task of modeling species distributions, the harmonization of data is pivotal. Yet, ecologists and modelers often grapple with datasets that are like pieces of different jigsaw puzzles, not quite fitting seamlessly together. This misalignment, be it in terms of spatial resolution, grid orientation, or data source discrepancies, poses unique challenges in constructing reliable and accurate Species Distribution Models (SDMs).

"Misalignment Data" is more than a mere technical inconvenience; it's a reflection of the varied scales and methods of data collection inherent in ecological research. Overlooking or inadequately addressing data misalignment can introduce biases, leading to erroneous inferences and predictions, with potential ramifications for conservation planning and species management.

Enter R-INLA. With its sophisticated capabilities for Bayesian inference and a framework adaptable to diverse data structures, R-INLA emerges as an

indispensable tool for handling misalignment issues. This chapter sheds light on the complexities arising from data misalignment in SDMs and presents advanced methodologies within the R-INLA environment to adeptly address them.

Guided by detailed case studies, hands-on exercises, and clear visual demonstrations, we will journey through strategies for reconciling mismatched data, ensuring that the resulting SDMs are both statistically robust and ecologically meaningful. From interpolation techniques to multi-resolution modeling, the chapter aims to empower readers with a comprehensive toolkit for tackling misalignment challenges head-on.

Join us in this deep dive into the world of misalignment, as we navigate the nuances, pitfalls, and solutions, ensuring that our species distribution models remain a faithful mirror to the complex realities of the natural world.

10

Presence-only (INLABRU) (Xavi y Iosu)

Vale, aquí yo si que voyu a tener que estudiar, espero que Iosu me ayude bastante... o ese es mi deseo.

Este capítulo fue una de las cosas que nos piderieron los Revisores, y por eso lo hemos incluido

10.1 Unraveling the Enigma of Presence-Only Data with INLABRU

In the vast repositories of ecological data, presence-only datasets stand out as both a treasure trove and a challenge. They represent locations where a species is known to occur, but without the balancing act of absence data, offering a one-sided view of the distribution landscape. While abundant and often easier to collect, the intrinsic biases and nuances of presence-only data have historically posed analytical challenges in the realm of Species Distribution Models (SDMs).

The chapter "Presence-Only with INLABRU" delves deep into this specific data type, guiding readers through its intricacies and the advanced methodologies available within the INLABRU framework for its effective utilization.

INLABRU, as an extension of R-INLA, emerges as a beacon of hope in this context, providing tools specifically tailored to tackle the challenges and harness the potential of presence-only data.

Through a combination of theoretical insights, hands-on exercises, and empirical examples, we will explore how INLABRU facilitates the creation of robust and insightful SDMs from presence-only datasets. From understanding potential sampling biases to creating pseudo-absence points and integrating environmental covariates, this chapter aims to offer a comprehensive guide to modeling with presence-only data using INLABRU.

Beyond just techniques and methodologies, we will also address the ecological and conservation implications of the insights derived from such models. The goal is not only to produce accurate models but also to derive meaningful interpretations that can guide conservation and management strategies.

Embark with us on this enlightening journey through the world of presence-only data. Together, we will navigate its challenges, embrace its potential, and unlock new horizons in species distribution modeling using the combined might of R-INLA and INLABRU.

11

Spatio-temporal species distribution (Iosu)

Y como Iosu ya habrá calentado con el chapter 10, entonces ahora cogerá sus papers de Spatio-Temporal y nos hará un coqueto capítulo.

11.1 Navigating the Dynamic Tapestry of Spatio-temporal Species Distribution

The ever-shifting landscapes of our planet, influenced by a myriad of ecological and anthropogenic factors, ensure that species distributions are not static snapshots but dynamic canvases evolving over time. To capture this essence of change and spatial intricacy, spatio-temporal modeling emerges as a critical tool, allowing us to understand not just where a species might be found, but also when and how its distribution might shift.

In the chapter "Spatio-temporal Species Distribution," we embark on a journey to explore the confluence of space and time in the realm of Species Distribution Models (SDMs). Recognizing that species distributions are a product of both present conditions and historical legacies, and that they may hint at future trajectories, we dive deep into methodologies that capture these complex dynamics.

With the combined prowess of R-INLA and INLABRU, we are equipped with a powerful suite of tools to tackle the challenges of spatio-temporal modeling. These frameworks, designed with flexibility and precision in mind, allow for the integration of diverse data sources and the crafting of models that reflect the true dynamism of ecological systems.

Through theoretical discussions, real-world case studies, and hands-on tutorials, this chapter aims to guide readers through the intricacies of spatio-temporal SDMs. From understanding temporal trends and seasonality to predicting future distributions under changing scenarios, we'll explore the myriad ways in which time adds depth and dimension to our spatial models.

As we navigate the interplay of space and time, we also confront the broader implications of our findings. How can spatio-temporal models inform conservation strategies, habitat management, and ecological forecasting? These questions, and more, will be addressed as we journey through the dynamic world of spatio-temporal species distribution.

So, let us set forth on this exploration, where time becomes more than a mere backdrop, and space is understood in its ever-changing context, all through the lens of R-INLA and INLABRU.

12

Joint Species Distribution models (Xavi)

Pues me ha tocado a mi, como era lógico, escribir el capítulo de modelizar Varias Espcies a la vez. A ver si conseguimos que el famoso $\alpha_{2|1}$ nos sale bien y utilizamos lo que nunca hemos llegado a publicar de este tema.

12.1   Synergy in Nature - Delving into Joint Species Distribution Models

Ecological landscapes are a symphony of interactions, where the presence, abundance, or absence of one species often hinges on the intricate web of relations with others. Species do not exist in isolation; they are part of a larger community, governed by competition, mutualism, predation, and a myriad of other interspecies interactions. Capturing this interconnectedness is not just a matter of ecological authenticity but of model accuracy and predictive power.

Enter "Joint Species Distribution Models" (JSDMs). By moving beyond the traditional single-species modeling approach, JSDMs provide a holistic view, allowing for the simultaneous modeling of multiple species, thereby integrating the direct and indirect interactions that shape community structures and spatial patterns.

In this chapter, we unravel the complexities and opportunities associated with JSDMs. Utilizing the robust frameworks of R-INLA and INLABRU, we will dive into the methodologies that allow for the integration of multiple species data sets, shedding light on the shared environmental drivers and potential biotic interactions.

Through a blend of theoretical discussions, illustrative examples, and hands-on exercises, we aim to equip readers with the knowledge and tools to construct, interpret, and apply JSDMs. We will delve into the nuances of model specification, explore methods to disentangle species interactions from environmental effects, and address the challenges of high dimensionality and potential collinearity.

The overarching goal is to harness the power of JSDMs to provide deeper insights into community ecology, inform conservation strategies, and predict changes in biodiversity under various scenarios, all while tapping into the synergies and interdependencies that define nature.

Join us on this enlightening journey, as we navigate the multifaceted world of joint species distributions, guided by the combined capabilities of R-INLA and INLABRU.

13

Multiple sources of information: dependent-independent data and more. (Xavi y Antonio)

Vale, este es el tema que empezamos pre-pandemia y que ahora lo lleva Mario; pero bueno esperemos que publiquen ponto ese paper y así poder aporvechar ese material para hacer ese capítulo.

13.1 Confluence of Data Streams - Harnessing Multiple Sources of Information in SDMs

In the mosaic of species distribution modeling, data is the bedrock upon which insights are sculpted. But not all data sources are created equal, nor do they operate in isolation. From field surveys to remote sensing and citizen science contributions, each stream of data offers its unique vantage point, complete with intrinsic strengths and inherent biases. The true challenge, and opportunity, lies in weaving these disparate threads into a coherent, robust, and enlightening tapestry of understanding.

This chapter, "Multiple Sources of Information (dependent-independent data and more)," embarks on the journey of dissecting, interpreting, and integrating

varied data types within the realm of Species Distribution Models (SDMs). Recognizing the value of both dependent and independent datasets, as well as the intricacies that come with blending various sources, we will explore methodologies that celebrate this diversity of information.

With the combined analytical might of R-INLA and INLABRU, we are uniquely positioned to handle such multi-faceted datasets. These platforms, with their flexibility and precision, serve as the nexus where different data types converge, are analyzed in tandem, and produce insights greater than the sum of their parts.

Guided by rich examples, hands-on exercises, and deep theoretical insights, we'll traverse the nuances of data integration, from understanding potential correlation structures to addressing biases and scaling discrepancies. Emphasis will be placed on understanding the nature of each data type, ensuring that their integration is not just statistically sound but ecologically meaningful.

As we wade through the confluence of data streams, the broader implications for conservation, species management, and ecological understanding will come to the fore. This chapter is not just an exploration of data but a testament to the enriched insights that arise when multiple voices of information sing in harmony.

Dive in with us, as we navigate the intricate dance of data integration, harnessing the full spectrum of information to paint a richer, more nuanced picture of species distributions with the aid of R-INLA and INLABRU.

14

Other issues: downscaling, spatial confounding, multiple likelihoods, big data and mapping the world. (Xavi, Antonio y artistas invitados)

Pues la verdad es que no sé aquí que pondremos... supongo que las "open questions"... a ver que tal sale ... o si llegamos desfondados a este capítulo.

14.1 Navigating the Multifaceted Landscape of Advanced SDMs

The journey of understanding species distributions is rife with complexities, not merely in the ecological phenomena they depict, but in the analytical intricacies we encounter. Beyond the foundational principles and standard methodologies, a plethora of nuanced issues await the discerning modeler – challenges that, when addressed, can significantly elevate the depth, accuracy, and utility of Species Distribution Models (SDMs).

In this chapter, **Other Issues (downscaling, spatial confounding, multiple likelihoods, big data and mapping the world)**, we step off the beaten path to explore these often-underappreciated yet crucial facets of SDM. Each of these issues, whether it's the art of downscaling, the challenge of spatial confounding, the intricacies of handling multiple likelihoods, the enormity of big data, or the ambitious endeavor of mapping the world, represents a unique puzzle demanding tailored solutions.

Harnessing the unparalleled capabilities of R-INLA and INLABRU, we embark on this intricate journey. These powerful tools, with their adaptability and depth, are primed to tackle these challenges head-on, turning potential roadblocks into opportunities for enriched insights.

Guided by a blend of theory, practical exercises, and illustrative examples, we will:

- **Delve into the subtleties of downscaling**, understanding its necessity and methodologies in the context of SDMs.
- **Unravel the complexities of spatial confounding**, ensuring our models remain robust and ecologically meaningful.
- **Explore the world of multiple likelihoods**, optimizing our models for diverse data sources and ecological realities.
- **Grapple with big data**, leveraging its vastness without drowning in its depth.
- And, with ambition in our hearts, **explore the potential and challenges of mapping the world**, providing holistic, global insights into species distributions.

Beyond the techniques and tools, we'll keep an ever-watchful eye on the ecological and conservation implications of our endeavors, ensuring our models remain grounded in their ultimate purpose – understanding, preserving, and coexisting with the myriad species that share our planet.

Embark with us on this advanced expedition through the multi-dimensional world of SDMs, harnessing the combined prowess of R-INLA and INLABRU

to navigate and illuminate these intricate terrains.