Master Final Thesis

Using multivariate autoregressive state-space models to examine stock structure of horse mackerel in the North Atlantic

Ibsa Bernárdez Pousa

Master's Degree in Statistical Techniques

Curso 2025-2026

Final Thesis Proposal

Galician title: Uso de modelos autorregresivos multivariantes en espazo de estados para examinar a estrutura poboacional do xurelo no Atlántico Norte
Espanish title: Uso de modelos autorregresivos multivariantes en espacio de estados para examinar la estructura poblacional del jurel en el Atlántico Norte
Title: Using multivariate autoregressive state-space models to examine stock structure of horse mackerel in the North Atlantic
Modality: Modality B
Author: Ibsa Bernárdez Pousa, Universidad de A Coruña
Director: Javier Roca Pardiñas, Universidade de A Coruña
Mentors: Adriana Nogueira Gassent, Instituto Español de Oceanografía (CSIC); Gersom Costas, Instituto Español de Oceanografía (CSIC)
Brief summary of the thesis Horse mackerel is managed as three different stocks in the North Atlantic. The student will examine the population structure of the different stocks through time and space to determine whether they are part of the same larger population or not and how fishing or environmental changes affect their dynamics. Specifically he will ask (1) if the population abundance data support the existing management boundaries or if there are alternative grouping that receive more support, (2) if the subpopulation (if they exist) experience independent environmental variability or correlated variability. For that, he will mathematically formulate different hypothesis about the population structure via different multivariate autoregressive state-space models (MARSS).
Recommendations:
Further observations:

Don Javier Roca Pardiñas, Director académico de la Universidade de A Coruña, doña Adriana Nogueira Gassent, Tutora en la empresa de Instituto Español de Oceanografía (CSIC), y don Gersom Costas, Tutor en la empresa de Instituto Español de Oceanografía (CSIC), informan que el Trabajo Fin de Máster titulado

Using multivariate autoregressive state-space models to examine stock structure of horse mackerel in the North Atlantic

fue realizado bajo su dirección por don Ibsa Bernárdez Pousa para el Máster en Técnicas Estadísticas. Estimando que el trabajo está terminado, dan su conformidad para su presentación y defensa ante un tribunal. Además, Don Javier Roca Pardiñas y don Ibsa Bernárdez Pousa

√ sí ☐ no

autorizan a la publicación de la memoria en el repositorio de acceso público asociado al Máster en Técnicas Estadísticas.

En Vigo, a 3 de junio de 2026.

El director:
Don Javier Roca Pardiñas

ROCA PARDIÑAS JAVIER - 77592181W

Firmado digitalmente por ROCA PARDIÑAS JAVIER - 77592181W
Fecha: 2026.06.02 23:08:31 +02'00'

La tutora:
Doña Adriana Nogueira Gassent

NOGUEIRA GASSENT ADRIANA - DNI 14307098V

Firmado digitalmente por NOGUEIRA GASSENT ADRIANA - DNI 14307098V
Nombre de reconocimiento (DN): c=ES, o=CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS, ou=CERTIFICADO ELECTRONICO DE EMPLEADO PUBLICO, ou=CENTRO OCEANOGRAFICO DE VIGO, ou=14307098V, serialNumber=IDCES-14307098V, sn=NOGUEIRA GASSENT, givenName=ADRIANA, cn=NOGUEIRA GASSENT ADRIANA - DNI 14307098V
Fecha: 2026.06.03 14:59:32 +02'00'

El tutor:
Don Gersom Costas

Firmado por COSTAS BASTIDA GERSOM - DNI ***5070** el día

El autor:
Don Ibsa Bernárdez Pousa

BERNARDEZ POUSA IBSA - 48117627V

Firmado digitalmente por BERNARDEZ POUSA IBSA - 48117627V
Fecha: 2026.06.03 20:46:59 +02'00'

Declaración responsable. Para dar cumplimiento a la Ley 3/2022, de 24 de febrero, de convivencia universitaria, referente al plagio en el Trabajo Fin de Máster (Artículo 11, Disposición 2978 del BOE núm. 48 de 2022), **el/la autor/a declara** que el Trabajo Fin de Máster presentado es un documento original en el que se han tenido en cuenta las siguientes consideraciones relativas al uso de material de apoyo desarrollado por otros/as autores/as:

- Todas las fuentes usadas para la elaboración de este trabajo han sido citadas convenientemente (libros, artículos, apuntes de profesorado, páginas web, programas,...)
- Cualquier contenido copiado o traducido textualmente se ha puesto entre comillas, citando su procedencia.
- Se ha hecho constar explícitamente cuando un capítulo, sección, demostración,... sea una adaptación casi literal de alguna fuente existente.

Y, acepta que, si se demostrara lo contrario, se le apliquen las medidas disciplinarias que correspondan.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my mentors, Adriana Nogueira Gassent and Gersom Costas, for their constant support, patience and dedication throughout the entire process. Their reviews, comments and suggestions have been essential in improving both the analysis and the interpretation of the results. Thank you for helping me to better understand the problem, for guiding me in the application of the models and for taking the time to review every part of the work in such detail.

Secondly, I would like to thank my thesis director, Javier Roca Pardiñas, for his guidance and support with the methodological aspects of this Master's Thesis. His advice was essential in ensuring the correct application of the MARSS and sdmTMB models, as well as in structuring the statistical analysis and resolving the difficulties that arose during the course of the project.

I would also like to thank Martin O. Lindegren and Federico Maioli for their collaboration and assistance in standardising the data and estimating biomass indices using sdmTMB. Their contributions were invaluable in developing the methodological aspects of this work and in deriving the indices subsequently used in the MARSS models.

Finally, I would like to thank everyone who, in one way or another, contributed to the development of this project, whether through technical assistance, discussion of ideas or support during the various phases of the work. This Master's Thesis would not have been possible without that environment of collaboration and learning.

Contents

Abstract

Spanish abstract

La información sobre la estructura poblacional es esencial para una gestión pesquera eficaz. El jurel del Atlántico Norte (*Trachurus trachurus*) se gestiona tradicionalmente en tres grandes unidades: el stock del mar del Norte, que abarca el canal de la Mancha oriental y el mar del Norte meridional y central; el stock occidental, desde el mar Cantábrico y el golfo de Vizcaya hasta el norte del mar del Norte y el mar de Noruega; y el stock meridional, asociado a las aguas atlánticas de la península ibérica.

Este trabajo utiliza modelos autorregresivos multivariantes de espacio de estados (MARSS) para examinar las trayectorias de biomasa del jurel en distintas regiones del Atlántico Norte y evaluar si apuntan a una dinámica común o a unidades diferenciadas. Para ello, se recopilaron datos de campañas científicas de arrastre de fondo y se calcularon índices de biomasa mediante modelos espaciotemporales que consideran la heterogeneidad espacial. Posteriormente, se formularon varias hipótesis en MARSS para evaluar estructuras poblacionales alternativas, seleccionando los modelos más parsimoniosos según el criterio de información de Akaike corregido (AICc). También se incorporaron covariables ambientales y pesqueras, como la NAO y las capturas comerciales. Los resultados apoyaron tres dinámicas poblacionales diferenciadas en el Atlántico Nordeste, aunque no coinciden completamente con las delimitaciones de ICES. Los mejores modelos incluyeron las capturas comerciales y la NAO como covariables principales. En conjunto, MARSS es útil para contrastar hipótesis de estructura de stock, pero los resultados obtenidos deberían validarse en futuros estudios mediante la incorporación de técnicas complementarias de delimitación de stocks.

English abstract

Accurate information on population structure is essential for effective fisheries management. The North Atlantic horse mackerel (*Trachurus trachurus*) is traditionally managed across three units: the North Sea stock, which mainly covers the eastern English Channel, southern and central North Sea; the Western Stock, extending from the Cantabrian Sea and the Bay of Biscay to the northern North Sea and the Norwegian Sea; and the Southern Stock, associated with the Atlantic waters of the Iberian Peninsula, mainly west of Spain and Portugal.

This research uses multivariate autoregressive spatial state models (MARSS) to examine horse mackerel biomass trajectories in different regions of the North Atlantic and to assess whether these trajectories point to a common population dynamic or to distinct population units. To this end, data from scientific bottom trawl surveys were compiled to calculate biomass indices. These indices were estimated using spatiotemporal models to account for spatial heterogeneity and sampling design. Subsequently, multiple hypotheses were formulated in the MARSS model to evaluate different population structures, selecting the most parsimonious models based on the Akaike Information Criterion (AICc).

Furthermore, environmental and fisheries covariates, such as the Atlantic Oscillation Index (NAO) or commercial catches. Results supported the existence of three different population dynamics in the North-East Atlantic, although do not fully coincide with ICES delineations. The best models included commercial catches and the NAO as key covariates, although their effect should be interpreted with caution. Overall, MARSS is useful for testing hypotheses regarding stock structure, but future work should incorporate other delineation techniques.

Galician abstract

A información sobre a estrutura poboacional é esencial para unha xestión pesqueira eficaz. O xurelo do Atlántico Norte (*Trachurus trachurus*) xestiónase en tres unidades: o stock do mar do Norte, que abrangue a canle da Mancha oriental e o mar do Norte meridional e central; o stock occidental, desde o mar Cantábrico e o golfo de Biscaia ata o norte do mar do Norte e o mar de Noruega; e o stock meridional, asociado ás augas atlánticas da península ibérica, principalmente ao oeste de España e Portugal.

Este traballo utiliza modelos autorregresivos multivariantes espazo-estado (MARSS) para examinar as traxectorias de biomasa do xurelo en distintas rexións do Atlántico Norte e avaliar se apuntan a unha dinámica común ou a unidades diferenciadas. Para iso, recompiláronse datos de campañas científicas de arrastre de fondo e calculáronse índices de biomasa mediante modelos espazo-temporais que consideran a heteroxeneidade espacial. Posteriormente, formuláronse varias hipóteses no MARSS para avaliar estruturas poboacionais alternativas, seleccionando os modelos máis parsimoniosos seguindo o criterio de información de Akaike corrixido (AICc). Tamén se incorporaron covariables ambientais e pesqueiras, como a NAO e as capturas comerciais. Os resultados apoiaron a existencia de tres dinámicas poboacionais distintas no Atlántico nordeste, aínda que non coinciden completamente coas delimitacións do ICES. Os mellores modelos incluíron as capturas comerciais e a NAO como covariables principais, aínda que o seu efecto debe interpretarse con cautela. En xeral, MARSS é útil para probar hipóteses sobre a estrutura poboacional, pero os traballos futuros deberían incorporar outras técnicas de delimitación.

Chapter 1

Introduction

1.1 Background on horse mackerel in the North Atlantic

Figure 1.1: Horse mackerel (*Trachurus trachurus*)

The Atlantic horse mackerel (Figure 1.1) called *Trachurus trachurus* belongs to the Carangidae family and occurs throughout the eastern Atlantic, from Norwegian waters to western Africa, as well as in the Mediterranean Sea (Froese and Pauly, 2015). It is a pelagic species that forms schools and is commonly found over the continental shelf, making it one of the most broadly distributed shelf species in the Northeast Atlantic, where its distribution partially overlaps with other species of the genus Trachurus, including *T. picturatus* (Bowdich, 1825) and *T. mediterraneus* (Steindachner, 1868) in the north and west of Iberian peninsula.

Atlantic horse mackerel emerged as one of the three most important pelagic species in the European fisheries sector during the 1980s and 1990s. This development was partly driven by the decline in the availability of other fishery resources due to overexploitation, which consequently increased the commercial value of alternative species (Abaunza et al. 2003).

Historically, horse mackerel was frequently discarded in fisheries intended for human consumption because of its relatively low commercial value. With the expansion of small-mesh industrial fisheries, the species was also increasingly targeted for reduction purposes. However, during the 1970s, Eastern European fleets began exploiting horse mackerel for direct human consumption fisheries (Postuma, 1978).

Horse mackerel is a long-lived species, with individuals capable of reaching up to 40 years of age, although ageing methods remain somewhat uncertain (Abaunza et al. 2003). It may grow to approximately 60 cm in length, although individuals are more commonly found within the 15–40 cm range (Smith-Vaniz, 1986). Growth is particularly rapid during the first years of life but slows considerably after the age of three.

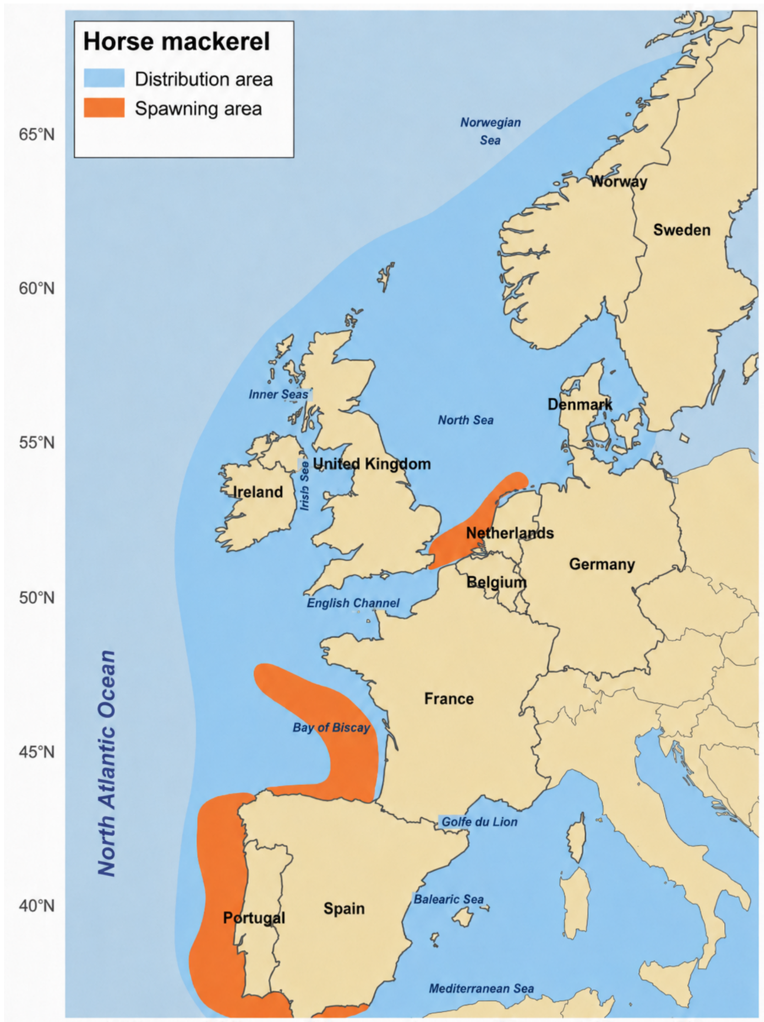

Figure 1.2: Distribution and spawning of Horse mackerel (*Trachurus trachurus*)

As we can see in (Figure 1.2), horse mackerel appears to distribute among the Northeast Atlantic waters. Despite its commercial importance, substantial knowledge gaps remain regarding the development and reproductive biology of *T. trachurus* (Abaunza et al. 2003).

Sexual maturity is generally reached at lengths between 16 and 25 cm, corresponding to approximately 2–4 years of age, although in some region's maturity may occur after only one year (Abaunza et al. 2003).

The species presents well-defined spawning, feeding, and overwintering grounds. Migration patterns are believed to be primarily influenced by prey availability and water temperature. In autumn, when water temperatures decrease below approximately 10 °C, *T. trachurus* migrates from feeding areas in the Norwegian Sea and the North Sea towards overwintering areas further south. These include the western English Channel for the North Sea stock (Lockwood and Johnson 1977; Macer 1974, 1977), areas along the continental slope surrounding the British Isles (Macer, 1977), and the eastern Bay of Biscay shelf edge for the western stock (Eaton, 1983).

Horse mackerel appears to avoid waters colder than 8 °C during overwintering (Polonsky, 1965), and feeding activity ceases when temperatures fall below 9 °C (Herrmann, pers. comm.; Abaunza et al. 2003). Quarterly surveys indicate that North Sea horse mackerel are most abundant in the southern North Sea during spring and summer, although juveniles remain in the area somewhat longer than adults.

Juvenile horse mackerel are *pelagic feeders* that prey mainly on planktonic organisms such as euphausiids and copepods (Macer, 1977). As individuals grow larger, they feed increasingly on demersal prey, and small fish become progressively more important in their diet (Eaton, 1983).

1.2 Stock delineation and population structure

ICES is the International Council for the Exploration of the Sea, an intergovernmental marine science organization that provides impartial evidence on the state and sustainable use of seas and oceans. Its goal is to advance and share scientific understanding of marine ecosystems and to use this knowledge to provide advice for conservation, management and sustainability. Moreover, its work extends not only to the Atlantic Ocean but also to the Arctic, the Mediterranean Sea and the Black Sea **(ICES, 2026)**.

The concept of a stock plays a central role in fisheries science, as it represents the basic unit used to analyse the dynamics of exploited populations and to design management measures. Generally, the term "population" has been associated with the concept of "fish population" (Waldman, 1999). However, the term "fish population" is often used to refer to the members of a species that are commercially exploited through fishing activities (Shaklee and Currens, 2003).

According to Gulland (1971) and Hilborn & Walters (1992), a thorough understanding of the stock structure is necessary to evaluate the fishing pattern and the condition of the resource. However, there is no single universally accepted definition, as its meaning can vary depending on the study's objectives and the discipline from which it is approached (Carvalho and Hauser, 1994). In general terms, various interpretations of the concept have been proposed, such as genetic, phenotypic, contingent, or harvest stock, which differ in the degree of biological or functional integrity they consider (Gauldie, 1988; Carvalho and Hauser, 1994; Secor, 1999; Hare, 2005).

From a genetic perspective, a stock consists of individuals of the same species that reproduce with one another at random and maintain a certain degree of genetic independence when isolated, at least during the reproductive period. A phenotypic stock, on the other hand, refers to groups within a species that exhibit differences in certain observable traits, whether due to environmental or genetic factors, or a combination of both. A contingent stock combines these concepts to some extent, describing groups of individuals that maintain spatial and temporal coherence through their own migratory patterns, distinct from those of other groups. Finally, the harvest stock is defined from a more applied perspective as a locally accessible fishery resource whose exploitation does not directly affect the abundance of neighboring resources.

The division of the horse mackerel into different stocks was established based on several complementary sources of information, such as the spatial and seasonal distribution of fishing activity, patterns of egg and larval distribution, data from acoustic and trawl surveys, and differences in parasite infestation levels (ICES 2015).

Historically, tagging studies were implemented in 1994 and further studies were carried out in

1997 using genetic information based on allozymes and morphometry (ICES, 1998). Later refinements of horse mackerel population units were mainly informed by the EU-funded HOMSIR project (2000–2003). This project adopted a multidisciplinary approach, combining several genetic techniques, such as allozymes, mitochondrial DNA and microsatellites, with parasites as biological tags, body morphometry, otolith shape analysis and comparisons of life-history traits, including growth, reproduction and distribution (Abaunza et al. 2008).

The post-HOMSIR period introduced more targeted approaches to improve horse mackerel stock identification, especially for the uncertain boundary between the Western and North Sea stocks. In 2015, the PFA/WMR pilot study tested two methods: chemical fingerprinting using gas chromatography and genetic analysis based on microsatellite discovery through next-generation sequencing. Although both approaches were considered promising, they could not separate the Western and North Sea stocks conclusively (Brunel et al. 2016).

Further progress was made by Farrell and Carlsson (2018), within the EAPO Northern Pelagic Working Group project. This study expanded the sampling design across several years and spawning seasons and developed additional microsatellite markers. It detected significant and temporally stable genetic differentiation between samples from the southern North Sea and other areas, suggesting that the southern North Sea component represents a distinct biological unit.

Finally, the most recent and powerful approach has been the application of whole-genome sequencing (WGS), developed within the EAPO Northern Pelagic Working Group framework and applied by Fuentes-Pardo et al. (2020, 2023). Using genome-wide SNPs, these studies identified subtle but significant genetic differentiation, including a distinct southern North Sea population, a northern Atlantic group including western Ireland, the northern Spanish shelf and northern Portugal, and a more southern component around southern Portugal and North Africa.

Consequently, without stronger evidence, the Western and North Sea populations cannot be assumed to be fully independent. Finally, in (Figure 1.3), the evolution before and after 2004 of the stock delineation can be seen.

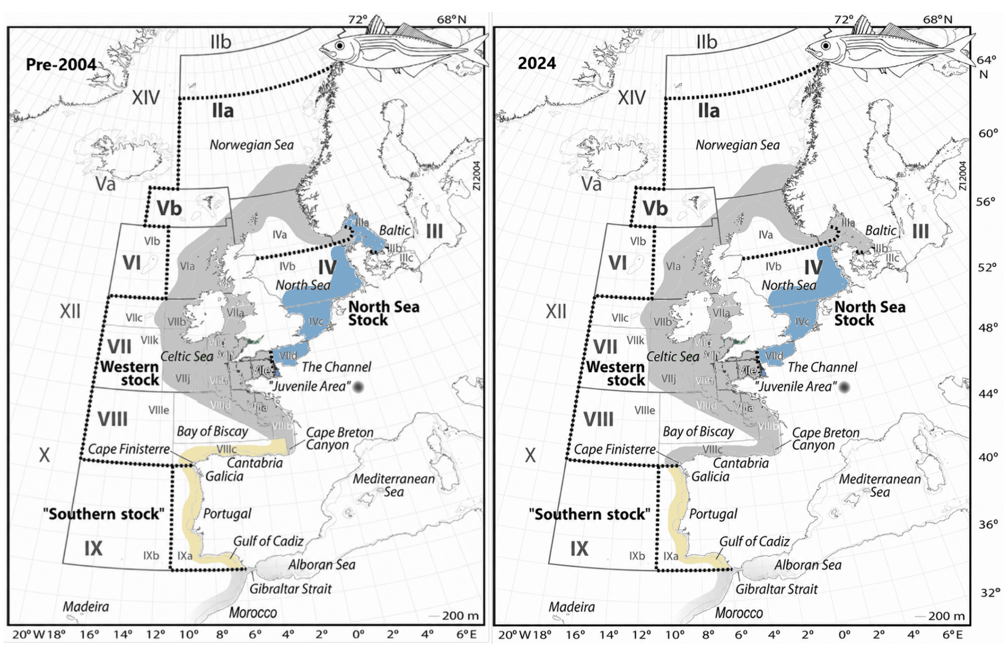

Figure 1.3: Horse mackerel stock delineation evolution before 2004 and in 2024.

1.3 Horse mackerel management units in ICES

As mentioned in stock delineation, stock boundaries must be clearly defined when conducting a stock assessment, and each stock should be assessed separately. In the ICES area, stock definitions are mainly based on egg distribution. The stocks are named after their spawning areas. The southern stock is spawning west of Portugal and Spain, the western is spawning in northern part of the Bay of Biscay, west of Ireland and Scotland and the southern stock is spawning in the southern part of the North Sea.

In the Northeast Atlantic horse mackerel are assessed and managed as three separate stocks (Figure 1.4). The North Sea stock in the eastern English Channel and southern North Sea area (ICES divisions 4.b, 4.c and 7.d), the southern stock was defined as that found in the Atlantic waters of the Iberian Peninsula (ICES Division 9.a), and the western stock on the northeast continental shelf of Europe, stretching from the Bay of Biscay in the south to Norway in the north (ICES Subarea 8 and Divisions 2a, 3a, 4a, 5b, 6a, and 7a–c, e–k).

The first formal delineation of horse mackerel stocks was established in 1987, based on the assumed existence of three distinct spawning areas corresponding to the Southern, Western and North Sea stocks. However, ICES working groups noted that the available information on eggs and larvae was not sufficient to demonstrate that these were truly independent biological stocks (Abaunza et al. 2008). In 1989, Divisions 2.a and 4.a were reallocated to the Western stock because their migration pattern was considered more similar to the Western stock than to the North Sea stock. Later, the HOMSIR project (2000–2003) proposed a further realignment between the Western and Southern stocks, with Division 8.c reallocated from the Southern to the Western stock. Despite these revisions, uncertainty around the stock boundaries has persisted, especially in potential mixing areas such as divisions 4.a or 7.d regarding the Western and Southern stocks.

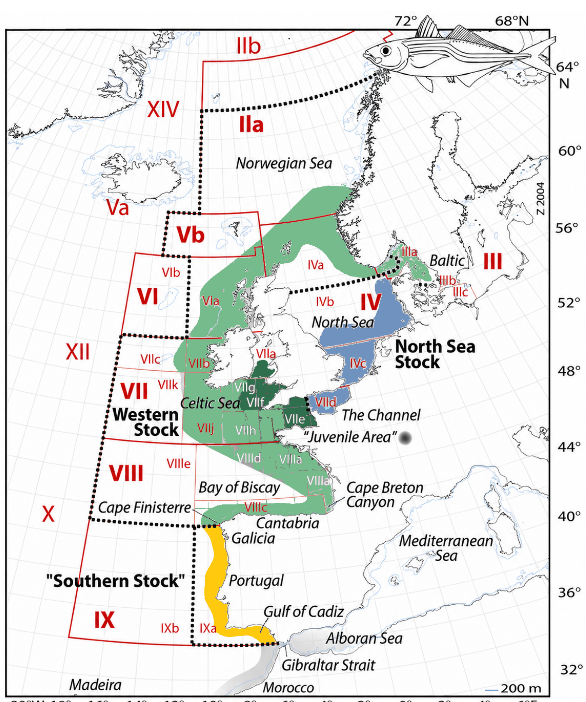

Figure 1.4: Spatial distribution of the three main horse mackerel (*Trachurus trachurus*) population units in the Northeast Atlantic. The Western Stock is shown in green, the North Sea Stock in blue and the Southern Stock in orange. Red labels represent ICES divisions and subdivisions.

1.4 Scientific Bottom Trawl Surveys (SBTS) and survey indices

Scientific bottom-trawl surveys (SBTS) are sampling processes used to estimate fish abundance by providing independent information on the number and weight of fish in a specific area and time (Jardim and Ribeiro, 2007).

Unlike commercial trawling, SBTS operates under standardised and brief sampling durations (Rufener et al. 2021) and uses sampling design techniques to uniformly sample whole regions and communities. Their extended time-series, frequent sampling (usually at least once a year), regional coverage, and diversity of the taxa sampled make them exceptional.

These surveys have been useful in understanding species co-occurrence and trophic relationships (Carrol et al. 2019; Selden et al. 2018), long-term community change (Jennings et al. 2002), biodiversity change and species on the move (Dencker et al. 2017; Magurran et al. 2015; Pecuchet et al. 2016), and community biomass changes (Friedland et al. 2020) within their survey regions, in addition to gathering vital biological data to support fish stock assessments.

Within the ICES area, the International Bottom Trawl Survey Working Group (IBTSWG) coordinates multispecies bottom-trawl surveys in the North Sea and northeastern Atlantic, ensuring consistent and comparable data for analysing spatial and temporal changes in fish populations and assemblages for stock assessment purpose (ICES 2019).

Traditionally, survey indices were often calculated using design-based methods, which are based on the survey sampling design and commonly rely on stratified estimators (Cochran, 1977; Pennington and Grosslein, 1978; Smith, 1996). In this approach, biomass is estimated directly from the survey design, for instance through stratified means by area, depth stratum or ICES division (ICES, 2019). The resulting index is based on the observed catches within predefined strata and assumes that the sampling design adequately represents the spatial distribution of the stock (Cochran, 1977; Pennington and Grosslein, 1978).

More recently, stock applications have moved towards model-based standardization. Instead of relying only on predefined strata, model-based approaches use statistical models to account for spatial heterogeneity, differences in sampling coverage, gear effects, depth, environmental conditions and temporal variation (Shelton et al. 2014; Thorson et al. 2015; Yalcin et al. 2023). This is particularly relevant for mobile and spatially heterogeneous species such as *Trachurus trachurus*, whose biomass may be unevenly distributed and variable across years, seasons and regions (Abaunza et al. 2003; Fuentes-Pardo et al. 2020; Farrell et al. 2024).

1.5 Spatio-temporal models

Geostatistical generalized linear mixed models, or geostatistical GLMMs, provide a suitable framework for this type of data. These models include spatially or spatiotemporally correlated random effects, which account for unmeasured processes that make nearby observations more similar in space or in both space and time (Rue and Held, 2005; Diggle and Ribeiro, 2007; Cressie and Wikle, 2011; Thorson and Kristensen, 2024). In the same way that random intercepts can account for correlation among observations within the same group, spatial and spatiotemporal random effects account for correlation among observations located close to each other.

A common way to represent spatial random effects is through Gaussian random fields. Under this approach, the random effects describing the spatial pattern are assumed to follow a multivariate normal distribution, and their dependence structure is defined by a covariance function, such as the exponential or Matérn function (Cressie, 1993; Chilés and Delfiner, 1999; Diggle and Ribeiro, 2007).

However, these models can become computationally demanding when the number of observations is large, because they require operations with large covariance matrices.

To reduce this computational burden, several approaches have been developed, including predictive processes (Banerjee et al. 2008; Latimer et al. 2009), nearest-neighbour Gaussian processes (Datta et al. 2016; Finley et al. 2022), and the stochastic partial differential equation approach, usually referred to as SPDE (Lindgren et al. 2011). The SPDE approach is particularly relevant because it approximates a Matérn correlation structure and produces sparse precision matrices, allowing much more efficient computation (Rue and Held, 2005; Lindgren et al. 2011). This framework has been widely used through R-INLA (Rue et al. 2009; Lindgren et al. 2011; Lindgren and Rue, 2015) and through implementations in TMB, Template Model Builder (Kristensen et al. 2016; Thorson et al. 2015c; Thorson, 2019a; Osgood-Zimmerman and Wakefield, 2023; Thorson and Kristensen, 2024).

1.6 MARSS framework

The MARSS package is an R package designed to fit linear multivariate autoregressive state-space models (MARSS) with Gaussian errors to time-series data (Holmes et al. 2012). These models are widely used to study linear stochastic dynamic systems that change over time and are applied in fields such as ecology (Olivença et al. 2022), economics (Bhadury et al. 2019), engineering (Pavlyuk, 2017) or biology (Ubeda et al. 2023 & Nogueira et al. 2018).

One of its main advantages is that it uses the Expectation-Maximization algorithm, or EM algorithm, as an alternative method for maximum-likelihood estimation. This approach is especially useful in ecological and environmental researchs, where datasets often contain many irregularly spaced missing observations or where observation error variance is difficult to know or estimate in advance (Schnute, 1994). Taking these two limitations into account, the three main contributions are listed below.

First, for some types of multivariate state-space models, the MARSS package uses the Expectation-Maximization algorithm, usually called the EM algorithm (Metaxoglou and Smith, 2007; Shumway and Stoffer, 1982). This algorithm can be slower than other optimization methods, but it is often more stable and robust, especially when the dataset contains many missing observations. The MARSS package also allows constrained parameter estimation meaning that some parameters can be fixed or forced to have the same value within the model matrices (Holmes, 2010).

Second, MARSS makes model specification easier because the model written in R corresponds directly to the model written in matrix form without writing additional code. This is important because, in other packages, the user often has to define specific matrices or write custom functions for each particular model. Thus, it is useful for users who are not experts in programming or matrix algebra. MARSS also allows degenerate multivariate models, where some observation or process variances can be set equal to zero. This makes it possible to include deterministic components in the model or to reformulate models with longer time lags or moving-average errors within the MARSS framework.

Third, the MARSS package includes tools for comparing different state-space models. These tools are based on Akaike's Information Criterion, or AIC, which is used to evaluate the support that the data provide for each model. In general, models with lower AIC values are considered to have stronger support. However, standard AIC and AICc can be biased when applied to hierarchical or autoregressive state-space models. For this reason, MARSS provides alternative AIC calculations based on bootstrapping methods, such as innovations bootstrapping (Cavanaugh and Shumway, 1997; Stoffer and Wall, 1991) and parametric bootstrapping (Holmes, 2010). The package also includes functions to calculate confidence intervals and to correct bias in estimated parameters.

Finally, the package includes a detailed user guide (Holmes et al. 2021) that explains how to use MARSS and presents several case studies based on ecological data. Although these examples are mainly focused on ecology, multivariate time-series analysis is useful in many disciplines, so researchers from other fields can also benefit from them.

1.7 Motivation and objectives

The aim of this research is to evaluate whether the biomass patterns observed in horse mackerel are consistent with the current ICES management stocks and with the expected population structure of the species. In particular, this study examines whether the available survey data support the current stock delineation used for stock assessment or suggest alternative biological groupings.

To achieve this objective, we used scientific bottom-trawl survey data from the ICES IBTS surveys (ICES 2019) covering the horse mackerel distributions of the three stocks. 1) We estimate spatio-temporal indices for each of the surveys 2) These standardized indices are subsequently used to formulate and compare alternative hypotheses regarding horse mackerel population structure through multivariate autoregressive state-space (MARSS) models. 3) Then, environmental and commercial catch covariates are explored to assess their potential influence on horse mackerel population dynamics.

Chapter 2

Materials and Methods

2.1 Study area

The study area (Figure 2.1) covers the northeastern Atlantic waters, where, according to ICES, the three stocks of *Trachurus trachurus* are found. These are the southern stock, located mainly around the Iberian Peninsula; the western stock, distributed throughout the waters of the Atlantic west of France, the Bay of Biscay, the British Isles, and the northern part of the Norwegian Sea; and the North Sea stock, associated mainly with the North Sea and the eastern part of the English Channel (Farell et al. 2020).

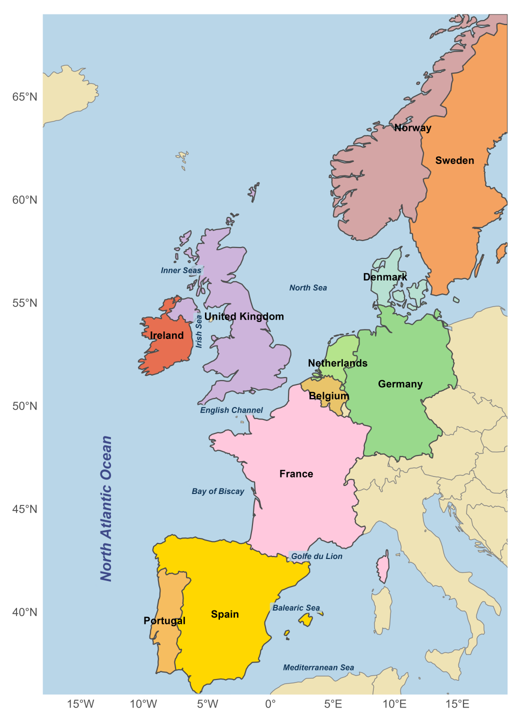

Figure 2.1: Study area for *Trachurus trachurus* in the northeast Atlantic.

2.1.1 Data Sources

Data used in this study were obtained from fishery-independent scientific surveys conducted in the northeastern Atlantic within the distribution area of the three ICES management stocks of *Trachurus trachurus*: the Western stock, the Northern stock and the Southern stock (Table 2.1).

The raw survey data were downloaded from the ICES DATRAS database. This database primarily compiles data from studies of bottom trawl surveys (IBTS) coordinated by ICES expert groups. The data used include biological, spatial and sampling information for each haul, such as catch, species, haul position, year, survey, ICES area and swept area. Catch and swept area variables are required to calculate the species biomass, the so-called Catch Per Unit of Area (CPUA). Finally, the downloaded data were processed in R using the DATRASextra package, which facilitates the downloading, cleaning, standardisation, analysis and visualisation of bottom-trawl survey data.

Table 2.1: Scientific bottom-trawl surveys by stock, country, ICES area and time series.

Stock	**Survey**	**Country**	**ICES division**	**Time-series**
Western Stock	NS-IBTS*	Germany, Denmark, Great Britain, Netherlands, Norway, Sweden	4.a, 6.a	1967–2024
	SP-PORC	Spain	7.b, 7.c, 7.k	2001–2024
	SP-NORTH*	Spain	8.b, 8.c	1990–2024
	FR-EVHOE	France	8.a–d, 7.f–j	1997–2024
	IE-IGFS	Ireland	6.a, 7.b, 7.c, 7.g, 7.j, 7.k	2003–2024
	UK-SWC	United Kingdom	6.a, 7.b	1985–2010
	UK-SCOWCGFS	United Kingdom	6.a, 7.b	2011–2024
	FR-WCGFS	France	7.e, 7.d	2018–2024
Northern Stock	NS-IBTS	Germany, Denmark, Great Britain, Netherlands, Norway, Sweden	4.b, 4.c, 7.d	1967–2024
	FR-CGFS	France	7.d, 4.c	1988–2024
Southern Stock	SP-NORTH	Spain	9.aN	1990–2024
	PT-IBTS	Portugal	9.a	2002–2024
	SP-ARSA	Spain	9.a	1996–2024

Note: NS-IBTS* is divided into two different surveys, one part corresponding to the Western stock (WS) and the other to the Northern stock (NS). SP-NORTH* also includes one WS survey and one SS survey.

The scientific bottom-trawl surveys used in this thesis covered the distribution areas of the WS, NS and SS. The details of each survey, including the corresponding stock, country, ICES division and time series, are shown in (Table 2.1).

The thirteen scientific survey series considered (Figure 2.2) follow stratified sampling designs, although their spatial coverage, temporal extent and sampling protocols differ among areas. For this reason, DATRAS raw data was processed with DATRASextra.

Although some survey series started before 2000 and others continued beyond 2024, the dataset used in the analysis was truncated to the common period 2000–2024 (Table 2.1). This temporal restriction was applied to improve comparability among surveys and stocks, ensuring a more consistent spatial and temporal coverage across the time series used for standardization and subsequent modelling. Prior to analysis, catches were standardized to weight per unit area based on the area swept by each haul (kg/km^2).

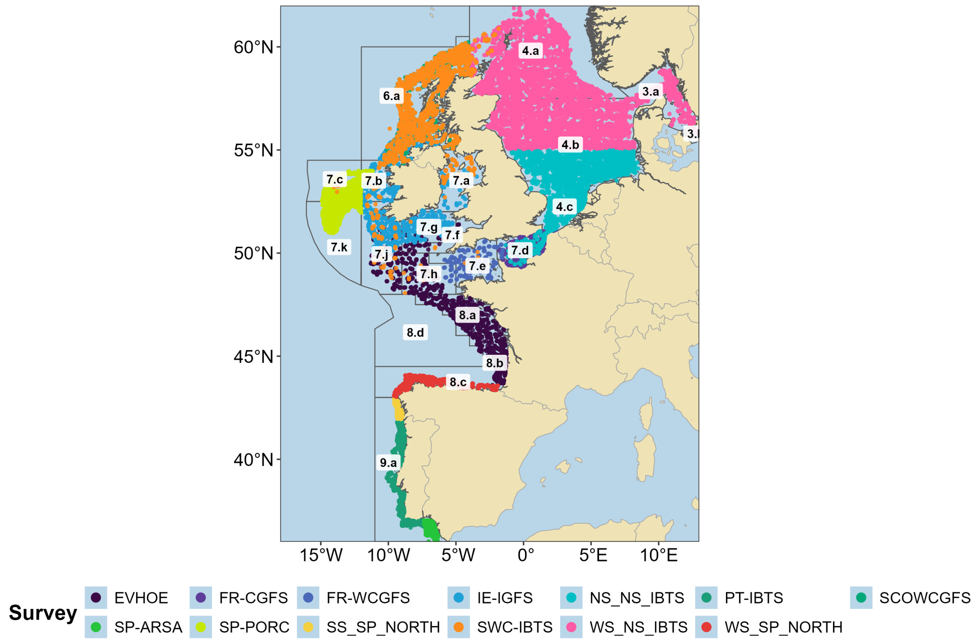

Figure 2.2: Survey distribution of Horse mackerel with ICES divisions.

2.1.2 Covariates

To investigate the potential effects of environmental and fishing factors on horse mackerel dynamics, a set of covariates were included in the modelling framework. These covariates were selected to represent climate variability and fishing pressure, factors that may influence patterns of biomass, recruitment, distribution and population dynamics.

The main environmental covariate considered was the North Atlantic Oscillation (NAO) index (Figure 2.3). The NAO is a large-scale climatic oscillation defined by the difference in atmospheric pressure between the Icelandic low and the Azores high. Because this index is associated with variability in atmospheric circulation, oceanographic conditions, and temperature patterns across the North Atlantic, it has frequently been used as an indicator of climate forcing in marine ecological research. Positive NAO values indicate a stronger pressure gradient between the Azores High and the Icelandic Low, which is usually associated with stronger westerly winds, changes in storm tracks and altered oceanographic conditions. In contrast, negative NAO values indicate a weaker pressure gradient and different atmospheric and oceanographic conditions across the North Atlantic. In this thesis, the NAO was considered as an indicator of environmental variability that could affect the distribution and population dynamics of horse mackerel (National Centers for Environmental Information, 2026).

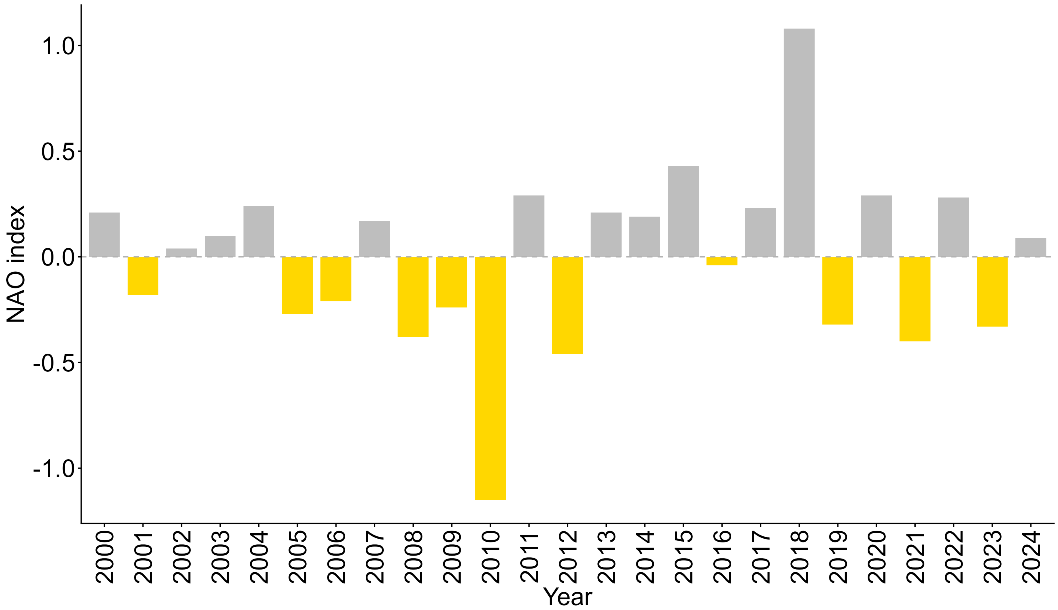

Figure 2.3: North Atlantic Oscillation (NAO) over years.

To account for fishing pressure, commercial catches (Figure 2.4) were included as fishery-related covariates. These variables were intended to represent the direct effect of exploitation on stock dynamics. In general, higher catches may reflect greater fishing pressure, which can influence stock dynamics by altering the temporal trajectory of the population. However, catches should be interpreted cautiously as they are not only determined by fishing pressure, but also depend on stock availability, catchability, fleet behaviour and management decisions.

(Figure 2.4) shows clear differences among the three stock components. The Western stock had the highest catches during most of the study period, although catches declined markedly after the early 2000s and reached very low values in the final years. The Southern and North Sea stocks showed lower catches overall, with more moderate interannual variation. These differences indicate that fishing activity has not been homogeneous across the three management units.

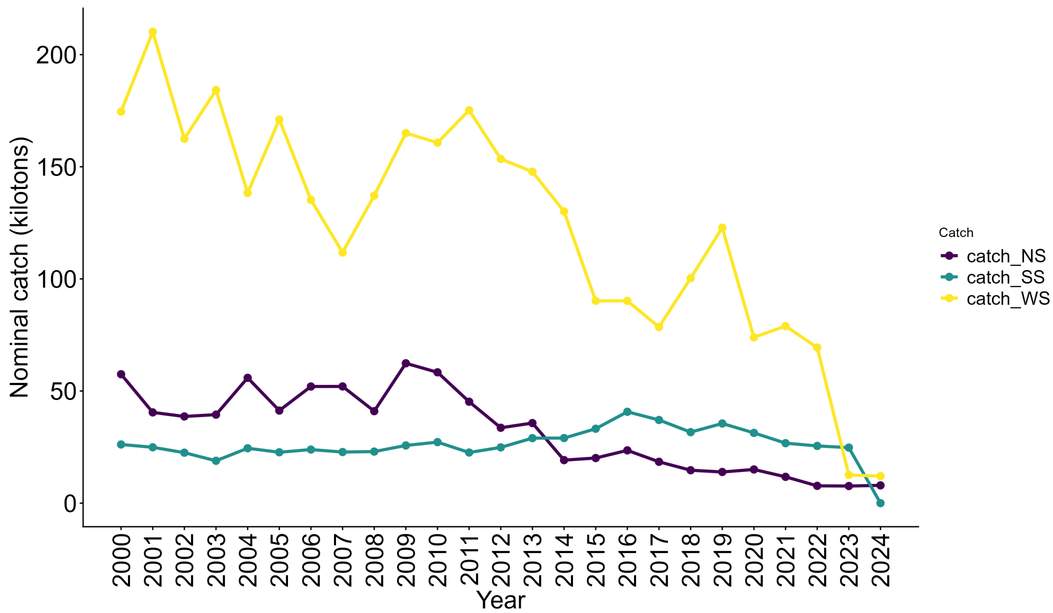

Figure 2.4: Fishery commercial catches over the years.

It should be noted that, in this thesis, both covariates were also analysed simultaneously.

2.2 Statistical methodology

2.2.1 Standardisation using sdmTMB

One of the most recent tools used for model-based standardization is sdmTMB, an R package for fitting spatial and spatiotemporal linear models. Anderson et al. (2024) describe sdmTMB as a framework based on Template Model Builder (Kristensen et al. 2016), Gaussian Markov random fields (Rue and Held 2005, Lindgren et al. 2011) and the Stochastic Partial Differential Equation approach (Lindgren et al. 2022).

Since the objective is to obtain standardized biomass indices from spatially distributed survey hauls, the model must account not only for annual differences in biomass, but also for spatial dependence. For this reason, the sdmTMB formulation can be introduced from the general spatial Gaussian random field GLMM. This model provides the basis for including fixed effects, such as year, and spatial random effects. A spatial Gaussian random field GLMM can be written as:

$$E[y_s] = \mu_s \mu_s = g^{-1}(\eta_s)\eta_s = X_s\beta + \omega_s$$

where y_s is the observation at spatial location s, μ_s is its expected value, g^{-1} is the inverse link function, $X_s\beta$ represents the fixed-effect component, and ω_s is the spatial random field (Anderson et al. 2024).

In this thesis, the response variable was biomass per unit area, expressed as *wgt_cpua*. This variable represents the biomass (catch per swept area) and was used as the measure of biomass density at each haul location. The models were fitted including spatial variables such as longitude, latitude and ICES area; temporal variables such as year, month, quarter and season and survey variables such as haul identity, gear, swept area and depth.

For each haul i, the observed response can be denoted as:

$$y_i = wgt\_cpua_i$$

where y_i is the observed biomass per unit area at haul i. Each observation is associated with a spatial location and a sampling year.

The spatial location is represented as:

$$s_i = (X_i, Y_i)$$

where X_i and Y_i are projected spatial coordinates obtained from the haul longitude and latitude. It is necessary to transform spatial coordinates into an equidistant projection, such as UTM, so that distances are measured consistently across the entire study area (Pebesma, 2018). This transformation is useful because spatial models require coordinates expressed in distance units rather than angular degrees. These coordinates were then used to construct the spatial mesh (Lindgren 2023).

The spatial mesh (Lindgren 2023) provides a representation of the study area and is used to approximate the spatial random field. This random field represents how biomass varies continuously across the study area. However, it is not possible to estimate this field at every possible location in the sea, because space is continuous and would contain infinitely many points. Instead, sdmTMB approximates the spatial field using a finite spatial mesh. The model estimates the spatial effect at the vertices of this mesh and then uses those estimated values to obtain the spatial effect at the actual haul locations and at the prediction grid locations.

In standard geostatistical spatial models, spatial correlation is usually defined as a function of the Euclidean distance between sampling locations. Anderson et al. (2024) describe the Matérn covariance function using the Euclidean distance between spatial points, with correlation decreasing as distance increases. However, in marine applications, Euclidean distance does not always reflect real ecological connectivity, since two sites may be geographically close in a straight line but physically separated by land. For this reason, a barrier mesh was used to reduce unrealistic spatial correlation across land by reducing spatial correlation across land barriers.

The biomass observations were modelled using a Tweedie distribution with a log link (Tweedie 1984). The Tweedie distribution is appropriate for survey biomass data because these data are often continuous and may contain zero values. Thus, for each haul i:

$$y_i \sim \text{Tweedie}(\mu_i, \phi, p)$$

with a log-link:

$$\log(\mu_i) = \eta_i$$

where μ_i is the expected biomass per unit area, ϕ is the dispersion parameter, p is the Tweedie power parameter, and η_i is the linear predictor.

The linear predictor of the model can be expressed as:

$$\eta_i = \beta_{year(i)} + \omega(s_i) + \varepsilon(s_i, t_i)$$

where $\beta_{year(i)}$ is the fixed effect corresponding to the year in which haul i was sampled, $\omega(s_i)$ is the spatial random field at location s_i, and $\varepsilon(s_i, t_i)$ is the spatiotemporal random field associated with location s_i and year t_i.

The term $\omega(s_i)$ represents the spatial random field. This component captures spatial structure that is relatively stable through time. It represents areas that tend to have consistently higher or lower biomass across the survey region after accounting for annual differences. This field is modelled as a Gaussian Markov random field (Rue and Held 2005; Lindgren et al. 2011), meaning that the spatial effects are assumed to be normally distributed and that each location is mainly related to neighbouring locations. The spatial dependence is described using a Matérn covariance structure (Whittle 1954; Matérn 1986), which defines how correlation decreases with distance: nearby locations are expected to be more similar than distant locations. Then, this Matérn field is approximated through the SPDE approach (Lindgren et al. 2022) over the spatial mesh, making it possible to model continuous spatial variation in biomass in a computationally efficient way.

The term $\varepsilon(s_i, t_i)$ represents the spatiotemporal random field. This component captures year deviations from the average spatial pattern. It allows the spatial distribution of biomass to vary from year to year.

The fixed-effect component was specified using year as a categorical variable and without a global intercept. This means that the model estimates a separate annual effect for each year in the time series. The vector of annual effects can be written as:

$$\beta = \begin{bmatrix} \beta_{2000} \\ \beta_{2001} \\ \vdots \\ \beta_{2024} \end{bmatrix}$$

Each element β_t represents the mean biomass level for year t on the log scale, after accounting for the spatial and spatiotemporal components of the model. Therefore, the fitted model standardizes the observed haul-level biomass data by combining three main components: an annual fixed effect, a spatial random field and a year spatiotemporal random field. In simplified form, the model can be written as:

$$\log(\mu_i) = \beta_{year(i)} + \omega(s_i) + \varepsilon(s_i, t_i)$$

After fitting the model, predictions were made over a regular spatial grid covering the survey area. The prediction grid represents a set of spatial locations:

$$s_j^* = (X_j^*, Y_j^*)$$

where s_j^* denotes prediction cell j. This grid was restricted to the survey domain. The grid was then repeated for each year of the study period, so that predictions could be obtained for every combination of grid cell and year.

For each prediction location j and year t, the model estimates the expected biomass density:

$$\widehat{\mu}_{j,t}$$

This prediction is obtained by applying the inverse of the log link:

$$\widehat{\mu}_{j,t} = \exp(\widehat{\eta}_{j,t})$$

where:

$$\widehat{\eta}_{j,t} = \widehat{\beta}_t + \widehat{\omega}(s_j^*) + \widehat{\varepsilon}(s_j^*, t)$$

The annual standardized biomass index was then obtained by integrating predicted biomass density over the prediction grid. It as an area-weighted population index, obtained by predicting over a grid covering the area of interest and summing the predicted biomass, with optional bias correction when transforming from the link scale to the response scale (Thorson and Kristensen, 2016). The index for year t can be written as:

$$I_t = \sum_{j=1}^{J} \widehat{\mu}_{j,t} A_j$$

where I_t is the standardized biomass index for year t, $\widehat{\mu}_{j,t}$ is the predicted biomass density at grid cell j in year t, A_j is the area represented by grid cell j, and J is the total number of prediction grid cells.

In addition to estimating indices for the full survey domain, indices were also obtained by ICES area or subdivisions. For a given area a, the area-specific biomass index can be expressed as:

$$I_{a,t} = \sum_{j \in a} \widehat{\mu}_{j,t} A_j$$

where the summation is restricted to grid cells belonging to area a. This allowed standardized annual biomass indices to be produced for the spatial units needed in the subsequent population-structure analysis.

Uncertainty was also obtained for the annual indices. Therefore, each standardized index was represented by a point estimate and an uncertainty interval:

$$I_t, L_t, U_t$$

where I_t is the estimated annual biomass index, L_t is the lower uncertainty bound and U_t is the upper uncertainty bound.

Overall, the practical application of sdmTMB in this thesis followed four main steps. First, haul-level survey observations were prepared and transformed into projected spatial coordinates. Second, a spatial mesh and a barrier mesh were constructed to represent marine spatial correlation. Third, a Tweedie spatial or spatiotemporal model was fitted to biomass per unit area, including year as a fixed effect and spatial and spatiotemporal random fields. Finally, predictions were made over a regular spatial grid and integrated to obtain standardized annual biomass indices.

Thus, the indices used later in the MARSS models were not simple averages of observed catches. They were model-based standardized estimates that accounted for annual variation, spatial structure, biased sampling and spatial heterogeneity.

2.2.2 MARSS

The population structure of *Trachurus trachurus* was analysed using multivariate autoregressive state-space models, implemented with the MARSS package in R (Mildenberger et al. 2025). As mentioned before, this modelling framework is especially suitable for ecological time series, and it separates the unobserved population process from the observation process (Harvey 1989; Durbin and Koopman 2001; Holmes et al. 2012). In other words, the model distinguishes between the true underlying biomass dynamics and the variability associated with survey observations. This application follows the framework described by Holmes et al. (2025), where MARSS models are used to combine different time-series data and to identify spatial population structure by comparing alternative hypotheses about the number and grouping of underlying subpopulations.

The candidate MARSS models were defined according to different latent-state structures, each representing an alternative assumption about the population structure of *Trachurus trachurus* in the Northeast Atlantic. These model structures are based on the current delineation of the Western stock (WS), North Sea stock (NS) and Southern stock (SS), as well as on alternative assignments of boundary areas. The full set of candidate structures is summarized in (Table 2.2) and described below:

- **H1:** one single state (all survey series together)
- **H2:** three latent states (WS; NS and SS)
- **H3:** three latent states (WS + Division 9.aN; NS; SS)
- **H4:** three latent states (WS; NS; SS + Division 8.c)
- **H5:** three latent states (WS + Division 7.d; NS; SS)
- **H6:** two latent states (WS + NS grouped; SS)
- **H7:** two latent states (WS + SS grouped; NS)
- **H8:** three latent states (WS; NS + Divisions 3.a and 4.a; SS)
- **H9:** three latent states (WS + Divisions 9.aN and 7.d; NS; SS)

In this research, the observations correspond to the log-transformed standardized biomass indices obtained from the scientific bottom-trawl survey series. The data matrix used in the MARSS models can be written as:

$$Y = \begin{bmatrix} y_{1,2000} & y_{1,2001} & \cdots & y_{1,2024} \\ y_{2,2000} & y_{2,2001} & \cdots & y_{2,2024} \\ \vdots & \vdots & \ddots & \vdots \\ y_{13,2000} & y_{13,2001} & \cdots & y_{13,2024} \end{bmatrix}$$

where each row represents one survey time series n and each column represents one year T. So, it is an $(n \times T)$ matrix. Therefore, the observation matrix has 13 rows, corresponding to each of the survey series, and 25 columns, corresponding to the period 2000–2024. The 13 observed time series were ordered as follows:

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \\ y_{6,t} \\ y_{7,t} \\ y_{8,t} \\ y_{9,t} \\ y_{10,t} \\ y_{11,t} \\ y_{12,t} \\ y_{13,t} \end{bmatrix} = \begin{bmatrix} WS\_NS\text{-}IBTS \\ WS\_SP\text{-}PORC \\ WS\_SP\text{-}NORTH \\ WS\_FR\text{-}EVHOE \\ WS\_IE\text{-}IGFS \\ WS\_SWC\text{-}IBTS \\ WS\_SCOWGFS \\ NS\_NS\text{-}IBTS \\ NS\_FR\text{-}CGFS \\ NS\_FR\text{-}WCGFS \\ SS\_PT\text{-}IBTS \\ SS\_SP\text{-}NORTH \\ SS\_SP\text{-}ARSA \end{bmatrix}_t$$

where WS refers to the Western stock, NS to the North Sea stock and SS to the Southern stock.

The latent state vector is denoted by:

$$X = \begin{bmatrix} x_{1,2000} & x_{1,2001} & \cdots & x_{1,2024} \\ x_{2,2000} & x_{2,2001} & \cdots & x_{2,2024} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,2000} & x_{m,2001} & \cdots & x_{m,2024} \end{bmatrix}$$

where m is the number of underlying population trajectories assumed in each hypothesis (states). Therefore, m changes depending on the stock-structure hypothesis being tested. For example, a panmictic hypothesis (Hypothesis 1) has only one latent state, whereas the current ICES stock delineation (Hypothesis 2) has three latent states: Western, North Sea and Southern.

1. **Model without covariates**

   The first group of models was fitted without environmental or fishing covariates. These models were used to evaluate whether the observed biomass time series were better explained by one common latent population trajectory or by several trajectories.

   The process equation is:

   $$x_t = Bx_{t-1} + u + w_t, \qquad \text{where} \qquad w_t \sim MVN(0, Q)$$

   In this equation, x_t represents the latent biomass state at year t, that is, the underlying population trajectory that is not directly observed. The matrix B describes the temporal dependence between the state at time t and the state at time $t-1$. Therefore, each latent trajectory follows a random process:

   $$x_t = Bx_{t-1} + u + w_t$$

   The model uses the default MARSS structure, which corresponds to an identity matrix. For a three-state model, this is:

   $$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, the process equation becomes a multivariate model:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \\ w_{3,t} \end{bmatrix}$$

or, equivalently:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{bmatrix} = \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \\ w_{3,t} \end{bmatrix}$$

This means that the latent biomass in a given year depends on the latent biomass in the previous year, plus an average growth rate and a process error.

The vector u represents the growth rate of each latent population trajectory. It can be interpreted as the mean tendency of the log-biomass trajectory to increase, decrease or remain stable through time (Holmes, 2001; Holmes et al. 2007; Humbert et al. 2009). Different structures were tested for u. In the first structure, called "unequal", each latent state has its own trend. For a three-state model it means that the Western, North Sea and Southern latent trajectories are allowed to have different average growth rates:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

In the second structure, called "equal", all latent states share the same mean trend. This structure assumes that the different latent population stocks may fluctuate differently through time but have the same average growth rate:

$$u = \begin{bmatrix} u \\ u \\ u \end{bmatrix}$$

For some three-state hypotheses, specifically the current stock delineation (Hypothesis 2) model and the model in which the English Channel is assigned to the Western stock (Hypothesis 5), an intermediate structure was also tested for u. This structure assumes that the Western and North Sea latent states share the same average trend, while the Southern state has a different one:

$$u = \begin{bmatrix} u_{WN} \\ u_{WN} \\ u_{SS} \end{bmatrix}$$

The matrix Q represents the variance-covariance matrix of the process errors. These errors describe real temporal variability in the latent population dynamics that is not explained by the deterministic part of the model. The Q matrix captures unexplained fluctuations caused by unobserved processes.

The alternative structures tested for Q were described by Holmes et al. (2012), including "diagonal and equal", "diagonal and unequal", "equalvarcov" and "unconstrained". These structures allow different assumptions about whether process errors are independent or correlated, and whether variances are shared or state specific.

The most flexible one was the "unconstrained" structure. It allows each latent state to have its own process variance and also allows covariances between states. The covariance terms allow the model to detect whether different latent population trajectories tend to fluctuate together:

$$Q = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}$$

A second structure was "diagonal and unequal". It assumes that the latent states have different process variances but that their process errors are independent:

$$Q = \begin{bmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & q_3 \end{bmatrix}$$

A third structure was "diagonal and equal". It assumes that all latent states have the same amount of process variability and that there is no covariance between them:

$$Q = \begin{bmatrix} q & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & q \end{bmatrix}$$

Finally, the "equalvarcov" structure was also tested, allowing a more constrained covariance structure where variances and covariances are shared according to the MARSS specification. This provides an intermediate option between a fully unconstrained covariance matrix and a completely diagonal matrix.

$$Q = \begin{bmatrix} q & c & c \\ c & q & c \\ c & c & q \end{bmatrix}$$

The observation equation was:

$$y_t = Zx_t + a + v_t, \qquad \text{where} \qquad v_t \sim MVN(0, R)$$

Here, y_t is the vector of observed log-biomass indices in year t, Z is the design matrix linking each survey time series to one latent state, a is a scaling vector, and v_t is the observation error.

Thus, the observation equation:

$$\begin{bmatrix} WS\_NS\text{-}IBTS \\ WS\_SP\text{-}PORC \\ WS\_SP\text{-}NORTH \\ WS\_FR\text{-}EVHOE \\ WS\_IE\text{-}IGFS \\ WS\_SWC\text{-}IBTS \\ WS\_SCOWGFS \\ NS\_NS\text{-}IBTS \\ NS\_FR\text{-}CGFS \\ NS\_FR\text{-}WCGFS \\ SS\_PT\text{-}IBTS \\ SS\_SP\text{-}NORTH \\ SS\_SP\text{-}ARSA \end{bmatrix}_t = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{WS,t} \\ x_{NS,t} \\ x_{SS,t} \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \\ v_{4,t} \\ v_{5,t} \\ v_{6,t} \\ v_{7,t} \\ v_{8,t} \\ v_{9,t} \\ v_{10,t} \\ v_{11,t} \\ v_{12,t} \\ v_{13,t} \end{bmatrix}$$

The matrix Z is the key element used to test biological hypotheses about stock structure into statistical models (Hinrichsen, 2009; Hinrichsen and Holmes, 2009; Ward et al. 2010; Holmes et al. 2012). Each row of Z corresponds to one survey series, and each column corresponds to one latent population trajectory. A value of 1 indicates that a given survey is assigned to a given latent state (population), whereas a value of 0 indicates that it is not. For example, under the panmictic model (Hypothesis 1), all surveys are assigned to a single latent trajectory. This means that all 13 survey series are assumed to reflect one common horse mackerel population dynamic:

$$Z_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Under the current ICES stock delineation (Hypothesis 2), three latent states are assumed: Western, North Sea and Southern. The corresponding Z matrix would be:

$$Z_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Where the first seven surveys are assigned to the Western state, the next three to the North Sea state, and the last three to the Southern state.

The remaining Z matrices represent alternative hypotheses about stock structure. In Z_3, the survey SS_SP-NORTH is assigned to the Western state instead of the Southern state, testing the hypothesis that area 9.a.N is more closely related to the Western stock. In Z_4, WS_SP-NORTH is assigned to the Southern state, testing whether area 8.c behaves more similarly to the Southern stock. In Z_5, NS_FR-WCGFS is assigned to the Western state, testing whether the English Channel is more consistent with Western dynamics. In Z_6, Western and North Sea surveys are grouped into a single latent state, while Southern surveys form a second state. In Z_7, Western and Southern surveys are grouped together, while North Sea surveys form a separate state. In Z_8, WS_NS-IBTS is assigned to the North Sea state, testing whether the northern part of the North Sea behaves more like the North Sea stock. Finally, Z_9 combines two changes: SS_SP-NORTH and NS_FR-WCGFS are both assigned to the Western state.

The vector a was specified using the "scaling" option. This is important because the observed survey series were obtained from different surveys that encompass different countries, vessels, gears and sampling protocols. MARSS does not estimate all the elements of vector a, but instead sets the first coefficient associated with each latent state to zero and estimates the remaining scale coefficients corresponding to the surveys that observe that same state (Holmes et al. 2012; Holmes et al. 2021).

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \end{bmatrix} = \begin{bmatrix} 0 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ 0 \\ a_9 \\ a_{10} \\ 0 \\ a_{12} \\ a_{13} \end{bmatrix}$$

The observation error matrix R was specified as a diagonal matrix with one observation variance for each survey:

$$R = \begin{bmatrix} r_1 & 0 & 0 & \cdots & 0 \\ 0 & r_2 & 0 & \cdots & 0 \\ 0 & 0 & r_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{13} \end{bmatrix}$$

This assumes that observation errors are independent among surveys, but that each survey can have its own observation variance. Therefore, surveys with more variable or uncertain observations are allowed to have larger observation error variance than more stable surveys.

2. **Model with NAO as an environmental covariate**

The second group of models extended the previous formulation by including the North Atlantic Oscillation (NAO) index as an environmental covariate (NCEI, 2026). The NAO was used as an indicator of large-scale climatic variability in the North Atlantic, which may influence horse mackerel dynamics through changes in temperature, circulation, recruitment success, prey availability or spatial distribution. With the inclusion of the NAO, the process equation becomes:

$$x_t = Bx_{t-1} + u + Cc_t + w_t, \qquad \text{where} \qquad w_t \sim MVN(0, Q)$$

The observation equation remains unchanged:

$$y_t = Zx_t + a + v_t, \qquad \text{where} \qquad v_t \sim MVN(0, R)$$

The new term is:

$$Cc_t$$

where c_t is the covariate value at time t, and C is the matrix of regression coefficients that measures the effect of the covariate on the latent states. In the NAO models, the covariate matrix c has one row and 25 columns:

$$c = \begin{bmatrix} NAO_{2000} & NAO_{2001} & \cdots & NAO_{2024} \end{bmatrix}$$

The NAO series was filtered to the period 2000–2024 and then scaled. Scaling is important because it places the covariate on a standardized scale, making the estimated coefficient easier to interpret and improving numerical stability.

For the panmictic model (Hypothesis 1), where there is one latent state, the C matrix is:

$$C_1 = \begin{bmatrix} \beta_{NAO} \end{bmatrix}$$

Therefore, the process equation is:

$$x_t = x_{t-1} + u + \beta_{NAO} NAO_t + w_t$$

In this case, β_{NAO} measures whether years with higher or lower NAO values are associated with increases or decreases in the latent horse mackerel biomass trajectory.

For a three-state model, the C matrix used in the script is:

$$C = \begin{bmatrix} \beta_{NAO} \\ \beta_{NAO} \\ \beta_{NAO} \end{bmatrix}$$

MARSS constrains the same parameter “nao” so NAO coefficient should be equal across latent states. Therefore, the model assumes that the NAO has a common effect on all latent population trajectories, although each state can still have its own dynamics depending on the selected structures of Z, u and Q.

For a three-state model, the process equation can be written as:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{bmatrix} = \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} \beta_{NAO} \\ \beta_{NAO} \\ \beta_{NAO} \end{bmatrix} NAO_t + \begin{bmatrix} w_{1,t} \\ w_{2,t} \\ w_{3,t} \end{bmatrix}$$

Thus, the NAO model evaluates whether a common environmental factor improves the explanation of the latent biomass trajectories. If the inclusion of the NAO reduces AICc relative to the equivalent model without covariates, this would suggest that climatic variability provides additional information for explaining the temporal dynamics of horse mackerel biomass.

However, the interpretation of the NAO coefficient must be cautious. A positive coefficient would indicate that higher NAO values are associated with increases in the latent biomass trajectory, whereas a negative coefficient would indicate the opposite. Nevertheless, this relationship does not necessarily imply a direct causal effect, because environmental effects on fish populations may operate through delayed or indirect processes.

3. **Model with catches as a fishing-pressure covariate**

The third group of models included commercial catches as a covariate representing fishing pressure. As in the NAO models, the observation equation remained the same, while the process equation was expanded to include a covariate effect:

$$x_t = Bx_{t-1} + u + Cc_t + w_t$$

In this case, c_t represents commercial catches. Unlike the NAO model, where there is a single annual covariate common to all states, the catch covariate was aggregated according to each stock-structure hypothesis. This means that the catch series used in each model depends on the way surveys and areas are grouped in the corresponding Z matrix.

For example, under the current ICES stock delineation (Hypothesis 2), catches were aggregated into three covariate series:

$$c_t = \begin{bmatrix} catch_{WS,t} \\ catch_{NS,t} \\ catch_{SS,t} \end{bmatrix}$$

where $catch_{WS,t}$, $catch_{NS,t}$ and $catch_{SS,t}$ represent the catches assigned to the Western, North Sea and Southern stock areas, respectively. Therefore, for a three-state model, the covariate matrix c has three rows and 25 columns:

$$c = \begin{bmatrix} catch_{WS,2000} & catch_{WS,2001} & \cdots & catch_{WS,2024} \\ catch_{NS,2000} & catch_{NS,2001} & \cdots & catch_{NS,2024} \\ catch_{SS,2000} & catch_{SS,2001} & \cdots & catch_{SS,2024} \end{bmatrix}$$

The C matrix for a three-state catch model was specified as diagonal:

$$C = \begin{bmatrix} \beta_{catch,WS} & 0 & 0 \\ 0 & \beta_{catch,NS} & 0 \\ 0 & 0 & \beta_{catch,SS} \end{bmatrix}$$

This structure means that the catch covariate for each stock component affects only its corresponding latent state. For example, catches assigned to the Western stock affect the Western latent trajectory, but not directly the North Sea or Southern latent trajectories. The zero values outside the diagonal impose this restriction.

The process equation for a three-state model with catches can therefore be written as:

$$\begin{bmatrix} x_{WS,t} \\ x_{NS,t} \\ x_{SS,t} \end{bmatrix} = \begin{bmatrix} x_{WS,t-1} \\ x_{NS,t-1} \\ x_{SS,t-1} \end{bmatrix} + \begin{bmatrix} u_{WS} \\ u_{NS} \\ u_{SS} \end{bmatrix} + \begin{bmatrix} \beta_{catch,WS} & 0 & 0 \\ 0 & \beta_{catch,NS} & 0 \\ 0 & 0 & \beta_{catch,SS} \end{bmatrix} \begin{bmatrix} catch_{WS,t} \\ catch_{NS,t} \\ catch_{SS,t} \end{bmatrix} + \begin{bmatrix} w_{WS,t} \\ w_{NS,t} \\ w_{SS,t} \end{bmatrix}$$

After multiplying C by c_t, this becomes:

$$\begin{bmatrix} x_{WS,t} \\ x_{NS,t} \\ x_{SS,t} \end{bmatrix} = \begin{bmatrix} x_{WS,t-1} \\ x_{NS,t-1} \\ x_{SS,t-1} \end{bmatrix} + \begin{bmatrix} u_{WS} \\ u_{NS} \\ u_{SS} \end{bmatrix} + \begin{bmatrix} \beta_{catch,WS} catch_{WS,t} \\ \beta_{catch,NS} catch_{NS,t} \\ \beta_{catch,SS} catch_{SS,t} \end{bmatrix} + \begin{bmatrix} w_{WS,t} \\ w_{NS,t} \\ w_{SS,t} \end{bmatrix}$$

For two-state Hypothesis, such as the model grouping Western and North Sea stocks together (Hypothesis 6), the covariate structure was adapted accordingly. For example, in the Z_6 hypothesis, the latent states are:

$$x_t = \begin{bmatrix} x_{WS+NS,t} \\ x_{SS,t} \end{bmatrix}$$

and the covariate vector is:

$$c_t = \begin{bmatrix} catch_{WS+NS,t} \\ catch_{SS,t} \end{bmatrix}$$

with:

$$C = \begin{bmatrix} \beta_{catch,WS+NS} & 0 \\ 0 & \beta_{catch,SS} \end{bmatrix}$$

For the panmictic model (Hypothesis 1), all catches are aggregated into a single catch series:

$$c_t = \begin{bmatrix} catch_{ALL,t} \end{bmatrix} \qquad \text{and} \qquad C = \begin{bmatrix} \beta_{catch,ALL} \end{bmatrix}$$

Thus, the catch models test whether fishing pressure, selected according to each hypothesised stock structure, improves the explanation of the latent biomass dynamics. As with the NAO models, the interpretation of catch coefficients must be cautious. A negative catch coefficient would be consistent with a decrease in latent biomass associated with higher catches, whereas a positive coefficient could reflect situations where catches are higher in years of greater biomass availability. Therefore, catch should not be interpreted only as a direct pressure variable, but also as a variable that may be partly driven by abundance (Arreguín-Sánchez 1996; Harley et al. 2001), fishing effort (Liu and Heino 2013) or management decisions.

4. **Model with catches and NAO**

The fourth group of models included both commercial catches and the North Atlantic Oscillation index as covariates. These models were used to evaluate whether the combination of a fishery-related covariate and a large-scale environmental covariate improved the explanation of the latent biomass dynamics. As in the previous covariate models, the observation equation remained unchanged:

$$y_t = Zx_t + a + v_t, \qquad \text{where} \qquad v_t \sim MVN(0, R)$$

The process equation was extended to include both covariate effects:

$$x_t = Bx_{t-1} + u + Cc_t + w_t, \qquad \text{where} \qquad w_t \sim MVN(0, Q)$$

In this case, c_t contains both NAO and commercial catch information. The NAO was included as a common annual environmental covariate, while catches were aggregated according to each stock-structure hypothesis. Therefore, for a three-state model, the covariate vector can be written as:

$$c_t = \begin{bmatrix} NAO_t \\ catch_{WS,t} \\ catch_{NS,t} \\ catch_{SS,t} \end{bmatrix}$$

where NAO_t represents the scaled North Atlantic Oscillation index in year t, and $catch_{WS,t}$, $catch_{NS,t}$ and $catch_{SS,t}$ represent the scaled catch series assigned to the Western, North Sea and Southern components, respectively.

In the script, all covariates were scaled before being included in the MARSS model. This places NAO and catch on comparable standardized scales and improves numerical stability during parameter estimation.

For a three-state model, the coefficient matrix C was specified as:

$$C = \begin{bmatrix} \beta_{NAO} & \beta_{catch,WS} & 0 & 0 \\ \beta_{NAO} & 0 & \beta_{catch,NS} & 0 \\ \beta_{NAO} & 0 & 0 & \beta_{catch,SS} \end{bmatrix}$$

This structure means that the NAO effect is constrained to be equal across the three latent trajectories, whereas catch effects are estimated separately for each component. In other words, NAO is assumed to have a common effect on the Western, North Sea and Southern trajectories, while each catch series affects only its corresponding latent state. The zero values outside the catch diagonal impose the restriction that catches assigned to one component do not directly affect the latent biomass trajectory of the other components.

For a three-state model, the process equation can therefore be written as:

$$\begin{bmatrix} x_{WS,t} \\ x_{NS,t} \\ x_{SS,t} \end{bmatrix} = \begin{bmatrix} x_{WS,t-1} \\ x_{NS,t-1} \\ x_{SS,t-1} \end{bmatrix} + \begin{bmatrix} u_{WS} \\ u_{NS} \\ u_{SS} \end{bmatrix} + \begin{bmatrix} \beta_{NAO} & \beta_{catch,WS} & 0 & 0 \\ \beta_{NAO} & 0 & \beta_{catch,NS} & 0 \\ \beta_{NAO} & 0 & 0 & \beta_{catch,SS} \end{bmatrix} \begin{bmatrix} NAO_t \\ catch_{WS,t} \\ catch_{NS,t} \\ catch_{SS,t} \end{bmatrix} + \begin{bmatrix} w_{WS,t} \\ w_{NS,t} \\ w_{SS,t} \end{bmatrix}$$

After multiplying C by c_t, this becomes:

$$\begin{bmatrix} x_{WS,t} \\ x_{NS,t} \\ x_{SS,t} \end{bmatrix} = \begin{bmatrix} x_{WS,t-1} \\ x_{NS,t-1} \\ x_{SS,t-1} \end{bmatrix} + \begin{bmatrix} u_{WS} \\ u_{NS} \\ u_{SS} \end{bmatrix} + \begin{bmatrix} \beta_{NAO}NAO_t + \beta_{catch,WS}catch_{WS,t} \\ \beta_{NAO}NAO_t + \beta_{catch,NS}catch_{NS,t} \\ \beta_{NAO}NAO_t + \beta_{catch,SS}catch_{SS,t} \end{bmatrix} + \begin{bmatrix} w_{WS,t} \\ w_{NS,t} \\ w_{SS,t} \end{bmatrix}$$

Thus, the combined model evaluates whether latent biomass dynamics are better explained when both large-scale climatic variability and fishery-related variation are considered simultaneously. The NAO coefficient represents a shared environmental effect, while the catch coefficients represent stock-specific associations between commercial catches and the corresponding latent biomass trajectories.

For two-state hypotheses, the covariate structure was adapted to the corresponding grouping of latent states. For example, in Hypothesis 6, where the Western and North Sea stocks are grouped into a single latent trajectory and the Southern stock is kept separate, the covariate vector is:

$$c_t = \begin{bmatrix} NAO_t \\ catch_{WS+NS,t} \\ catch_{SS,t} \end{bmatrix}$$

with:

$$C = \begin{bmatrix} \beta_{NAO} & \beta_{catch,WS+NS} & 0 \\ \beta_{NAO} & 0 & \beta_{catch,SS} \end{bmatrix}$$

For the panmictic model, all catches were aggregated into a single catch series and the model contained only one latent state. In this case:

$$c_t = \begin{bmatrix} NAO_t \\ catch_{ALL,t} \end{bmatrix}$$

and:

$$C = \begin{bmatrix} \beta_{NAO} & \beta_{catch,ALL} \end{bmatrix}$$

Therefore, the models including both catch and NAO test whether the inclusion of both covariates improves model fit relative to models without covariates or models including only one covariate. As in the previous models, the interpretation of the estimated coefficients must be cautious. Catch may represent fishing pressure, but it may also reflect biomass availability, fishing effort or management decisions. Similarly, NAO may capture broad environmental variability, but its effect on horse mackerel biomass may be indirect or delayed. Consequently, these covariates were interpreted as statistical predictors of latent biomass dynamics rather than as direct causal mechanisms.

2.2.3 Tested models

Nine hypotheses of European horse mackerel (*Trachurus trachurus*) population structure were evaluated using the standardized biomass index values estimated for the study regions in the Northeast Atlantic (Table 2.2). The different hypotheses were implemented by modifying the matrix in the observation equation of the MARSS model (Hinrichsen, 2009; Hinrichsen and Holmes, 2009; Ward et al. 2010). This matrix defines how each observed biomass time series is linked to the underlying latent population trajectory. Therefore, regions assigned to the same latent trajectory were assumed to share a common population dynamic, whereas regions assigned to different trajectories were allowed to follow independent biomass trends.

Table 2.2: Nine hypotheses of horse mackerel population structure using survey indices from 10 regions tested by manipulation of the Z matrix in the MARSS models.

Hypothesis	**Number of populations (Z)**	**Scenarios (Z = SS, WS, NS)**
(1) Panmixia	1	(1) All surveys
(2) Current delineation	3	(1) 9a; (2) 3a4a567acek8abdc; (3) 4bc7d
(3) 9aN is part of WS	3	(1) 9a; (2) 3a4a567acek8abdc9aN; (3) 4bc7d
(4) 8c is part of SS	3	(1) 9a8c; (2) 3a4a567acek8abd; (3) 4bc7d
(5) English Channel is part of WS	3	(1) 9a; (2) 3a4a567acedk8abdc; (3) 4bc
(6) Correlation in WS and NS	2	(1) 9a; (2) 3a4a567acek8abdc4bc7d
(7) Correlation in SS and WS	2	(1) 9a3a4a567acek8abdc; (2) 4bc7d
(8) North part of North Sea is part of the NS	3	(1) 9a; (2) 567acek8abdc; (3) 4abc7d
(9) 9aN and English Channel are part of WS	3	(1) 9a; (2) 3a567acdek8abdc9aN; (3) 4bc

Note: NS is Northern Stock, WS is Western Stock, SS is Southern Stock.

The hypotheses tested (Figure 2.5) were formulated by exploring different combinations of stock assignments. Although the current ICES management units for horse mackerel were used as the initial reference framework, alternative configurations were also considered in order to assess whether the observed biomass dynamics supported the existing stock boundaries or suggested different groupings among regions. The nine population structure hypotheses tested are:

1. Panmictic population — one population. This hypothesis assumes a single common trajectory for all regions. It assumes that all biomass time series reflect the same underlying population dynamic, which would be consistent with a high degree of connectivity across the study area.

2. Current stock delineation — three populations. This hypothesis represents the current perception of three main management units: the North Sea stock, the WS, and the Southern stock. Under this scenario, the three stocks are treated as separate population units.

3. Northern Division 9a as part of the WS — three populations. This hypothesis keeps three population stocks but assigns the northern part of Division 9a, represented by the Spanish northern survey area, to the WS instead of the Southern stock. It tests whether this area shows biomass dynamics closer to the WS.

4. Division 8c as part of the Southern stock — three populations. This hypothesis assigns the northern Spanish shelf area in Division 8c to the Southern stock rather than to the WS. It evaluates whether the Cantabrian area is more closely associated with the Southern stock dynamics.

5. English Channel as part of the WS — three populations. This hypothesis assigns the English Channel survey area to the WS instead of the North Sea stock. It tests whether this area behaves more like part of the WS.

6. Western and North Sea stocks combined — two populations. This hypothesis groups the Western and North Sea regions into a single trajectory, while keeping the Southern stock separate. It evaluates whether the Western and North Sea areas share a common biomass dynamic.

7. Western and Southern stocks combined — two populations. This hypothesis groups the Western and Southern regions into a single trajectory, while keeping the North Sea stock separate. It tests whether the boundary between the Western and Southern stocks is weak or whether these two areas may behave as a single population component.

8. Northern North Sea area assigned to the North Sea stock — three populations. This hypothesis assigns the northern North Sea survey area, initially included within the WS, to the North Sea stock. It evaluates whether this northern area shows dynamics more consistent with the North Sea stock than with the WS.

9. Northern Division 9a and English Channel assigned to the WS — three populations. This hypothesis combines two possible boundary adjustments: the northern part of Division 9a is assigned to the WS instead of the Southern stock, and the English Channel is assigned to the WS instead of the North Sea stock. This model tests whether both areas are better represented as part of the WS.

For each population structure hypothesis, alternative assumptions about the population growth rates, represented by the vector , were tested. For all configurations, two general structures were considered: an equal structure, in which all latent population trajectories shared the same growth rate, and an unequal structure, in which each latent trajectory had its own estimated growth rate.

For hypotheses Z2 and Z5, an additional intermediate structure was tested. In this case, the Western and North Sea stocks were constrained to share the same growth rate, while the Southern stock was allowed to have a different growth rate. This structure was included to evaluate whether the Western and North Sea stocks could share similar temporal dynamics while remaining distinct from the Southern stock.

Different structures were also tested for the process variance-covariance matrix , which describes the unexplained variability in the latent population trajectories. Four structures were considered: unconstrained, allowing all variances and covariances among latent trajectories to be estimated freely; diagonal and unequal, assuming no covariance among trajectories but allowing each trajectory to have its own process variance; diagonal and equal, assuming no covariance among trajectories and a common process variance; and equalvarcov, assuming a common variance and a common covariance among trajectories.

The observation variance-covariance matrix , which represents the observation error associated with the biomass indices, was kept independent among observed series. This assumption was adopted because the biomass indices were derived from different surveys, regions and sampling conditions.

Overall, this modelling strategy makes it possible to test explicit hypotheses about horse mackerel population structure. It provides a basis for assessing whether current ICES management boundaries are consistent with the temporal dynamics observed.

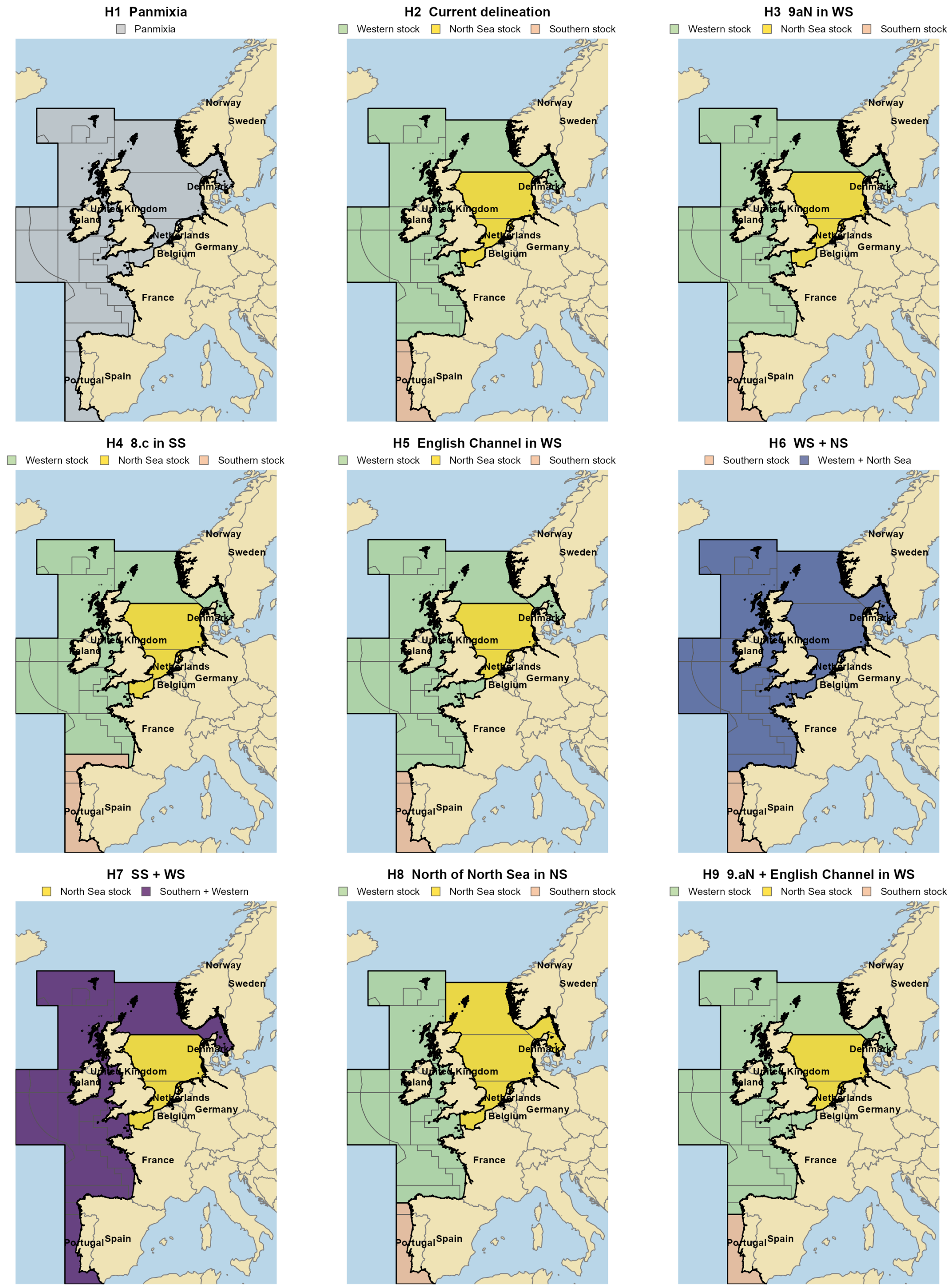

Figure 2.5: Nine hypotheses tested for horse mackerel (*Trachurus trachurus*) in the Northeast Atlantic. Colours indicate the areas covered by the different population units under each hypothesis: green represents the Western stock (WS), yellow the North Sea stock (NS), peach the Southern stock (SS), grey Panmixia, and purple or blue combined stock groups. Areas with the same colour are assumed to share the same stock population.

2.2.4 Model selection

The MARSS package includes an unbiased AIC measure, the AICc, calculated using innovations bootstrapping (Cavanaugh and Shumway, 1997; Stoffer and Wall, 1991) and parametric bootstrapping methods (Holmes, 2010). All combinations of stock-structure hypotheses, u structures and Q structures were fitted using MARSS with the EM algorithm (Metaxoglou and Smith, 2007; Shumway and Stoffer, 1982).

Model selection was carried out to identify the most parsimonious MARSS structure among the hypotheses evaluated. Following the logic described by Holmes et al. (2012), this approach allows different model structures to be compared by balancing model fit and complexity. Therefore, the selected model is the one that provides the best compromise between explanatory capacity and parsimony, avoiding overly flexible models that may fit noise in the data rather than represent robust population patterns.

The candidate models differ in several ways. First, they differ in the Z matrix, which defines the assumed population structure. Second, they differ in the u vector, which controls whether latent states share the same average trend or have separate trends. Third, they differ in the Q matrix, which controls whether latent states have independent or correlated process variability. Finally, in the NAO and Catch models, they also differ in the inclusion and structure of the covariate effect through the C matrix and the covariate matrix c.

For each fitted model, the Akaike Information Criterion (Cavanaugh and Shumway, 1997; Stoffer and Wall, 1991) corrected for small sample sizes, AICc, was calculated. The model with the lowest AICc value was considered to have the strongest relative support from the data. In addition, differences in AICc relative to the best-supported model were calculated as:

$$\Delta AICc_i = AICc_i - AICc_{\min}$$

where $AICc_i$ represents the criterion value for model i, and $AICc_{\min}$ corresponds to the lowest AICc value obtained among all compared models. Models with $\Delta AICc < 2$ were interpreted as having similar support to the top-ranked model. When several models showed comparable support, the simplest model was selected as the preferred model, that is, the model with the lowest number of estimated parameters.

Chapter 3

Results

3.1 Standardized survey biomass indices using sdmTMB

This section presents the standardized biomass indices obtained for the different survey series and for the three main stocks. The standardized biomass indices show substantial spatial and temporal variability among surveys and management stocks. The WS (Figure 3.3) generally presents the highest biomass levels, although important fluctuations are observed between surveys and years. The NS (Figure 3.1) shows more moderate biomass estimates with relatively variable temporal trends, while the SS (Figure 3.2) is characterized by lower biomass levels and higher interannual variability in some surveys. Differences among surveys likely reflect variations in spatial coverage, sampling design, and the heterogeneous distribution of horse mackerel across the study area. Overall, the results suggest that biomass dynamics are not fully homogeneous among regions, which may have implications for the current stock delineation and management structure.

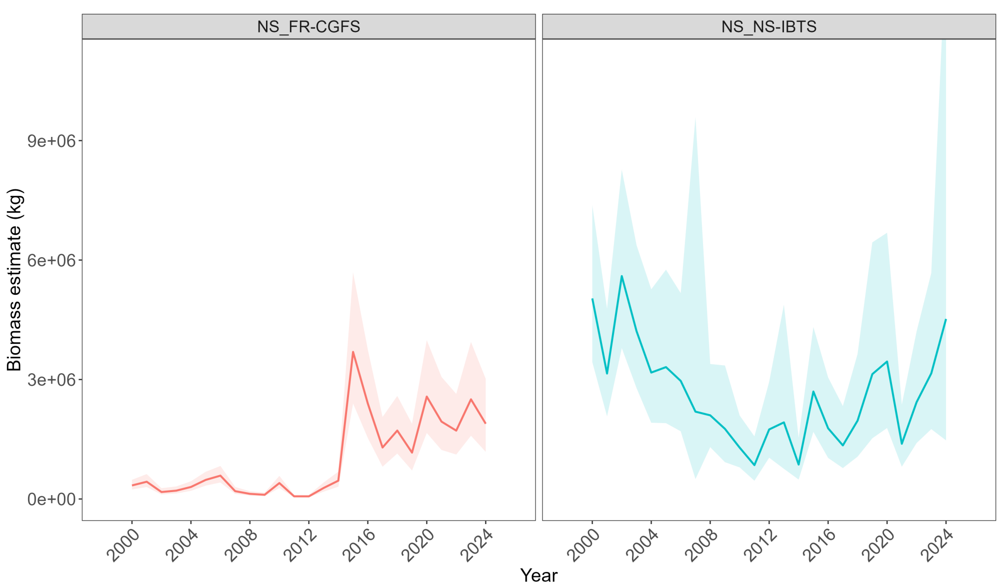

Figure 3.1: Biomass indices estimated for the surveys covering the NS distribution (FR-CGFS and NS-IBTS). Each panel represents one survey series, with the solid line showing the estimated biomass trend and the shaded area representing the associated uncertainty. Both surveys have the same time series (2000-2024).

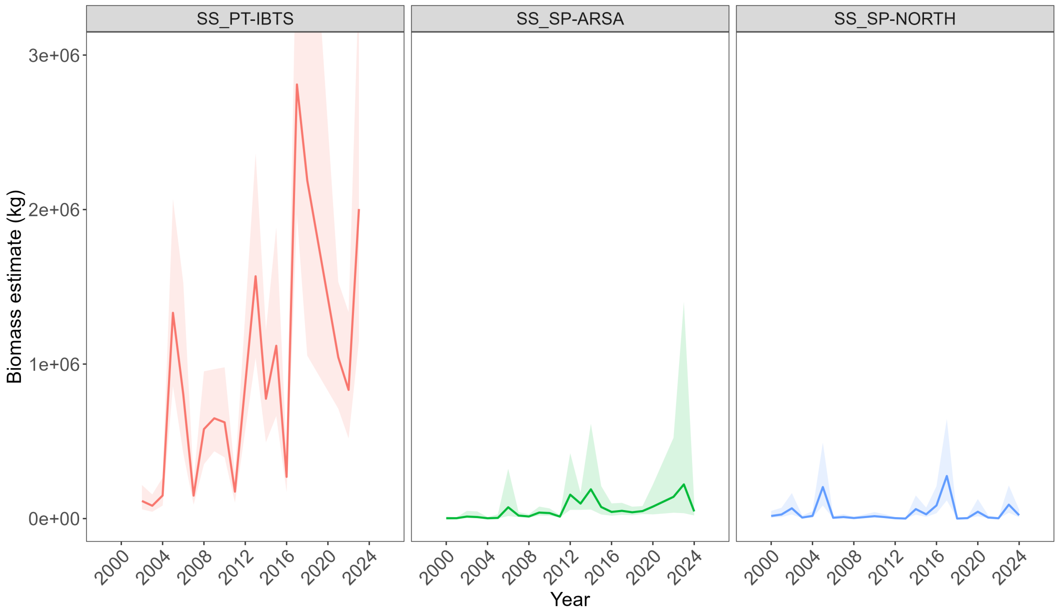

Figure 3.2: Biomass indices covering the distribution for the SS (PT-IBTS, SP-ARSA and SP-NORTH). Each panel represents one survey series, with the solid line showing the estimated biomass trend and the shaded area representing the associated uncertainty. The first survey (2002-2024) has a different time series than the second and third one (2000-204).

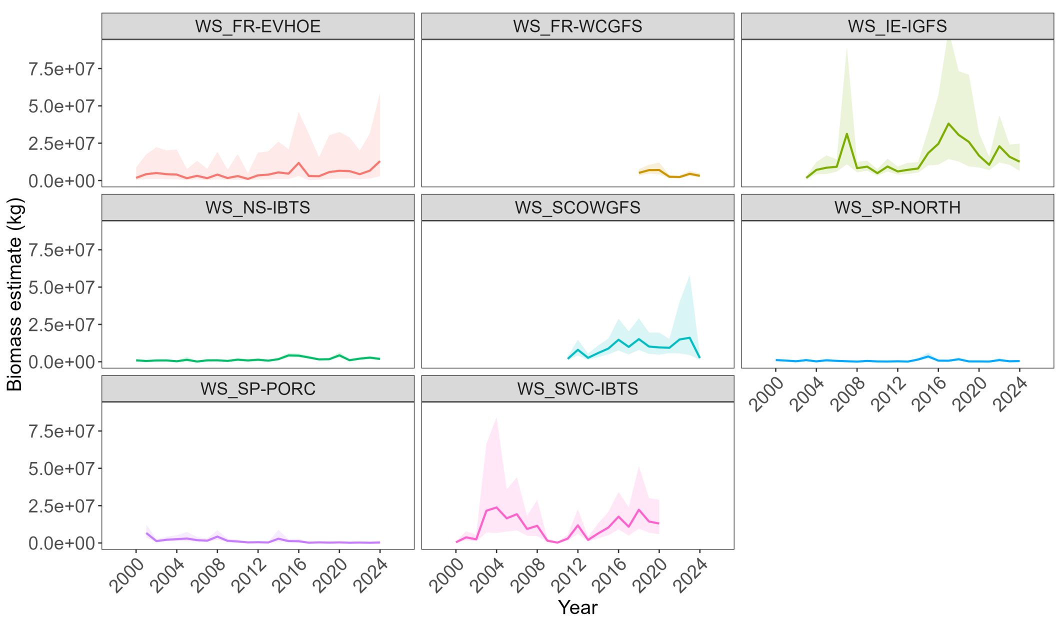

Figure 3.3: Biomass indices covering the distribution of the WS. Each panel represents one survey series, with the solid line showing the estimated biomass trend and the shaded area representing the associated uncertainty. All of the surveys have the same time series (2000-2024) except for the SWC-IBTS (2000-2020), SP-PORC (2001-2024), SCOWGFS (2010-2024), IE-IGFS (2002-2024) and FR-WCGFS (2018-2024).

Because of the previous heterogeneity issue, we use logarithmically transformed data (Figure 3.4).

The log-transformed standardized biomass indices show more comparable temporal patterns among surveys and management stocks. In general, most surveys exhibit relatively stable trends over time, although some fluctuations are still evident, particularly in the SS surveys. The WS generally maintains higher biomass levels than the other stocks, while the NS shows intermediate values and comparatively smoother dynamics. The log transformation reduces the influence of extreme biomass values and makes temporal trends more comparable among surveys, facilitating the evaluation of similarities and differences in population dynamics.

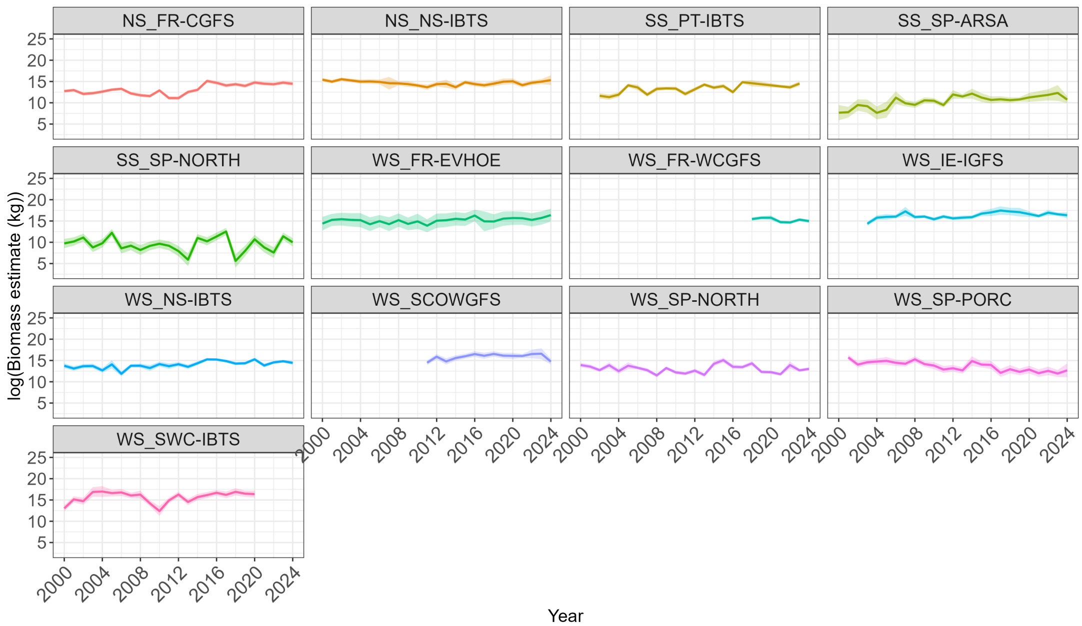

Figure 3.4: Log-Biomass index of the bottom-trawl surveys (2000-2024)

3.2 Results of MARSS models

We evaluated a total of 320 MARSS models with different stock structure hypotheses, process-error structures and covariate combinations, including models with catch, NAO, both covariates, and no covariates (Table 3.1). Model comparison based on AICc indicated strongest support for models under Hypothesis 9, which assumes three latent population trajectories. The best-supported model was Model 1, which included both catch and NAO as covariates, with a diagonal and equal process-error structure, unequal growth rates among trajectories, three latent states and 36 estimated parameters. This model had the lowest AICc value (AICc = 731.96; ΔAICc = 0). The second-ranked model was also based on Hypothesis 9 but included only catch as a covariate (AICc = 733.80; ΔAICc = 1.83). Therefore, although the Model 1 including both catch and NAO provided the best overall fit, the Model 2 with catch covariate also received substantial support.

The highest-ranked models consistently supported Hypothesis 9. The first four models all retained the same three-trajectory stock structure, differing mainly in the covariates included and in the assumed process-error structure. Models based on Hypothesis 5 appeared lower in the ranking, with ΔAICc values above 5.99, indicating weaker support compared with Hypothesis 9. Models without covariates were also less supported than models including both covariates but still received more support than models with only NAO as a covariate, suggesting that part of the temporal variation in the biomass indices was better explained when both covariates were included.

In the best-supported model, with both covariates, three latent trajectories were estimated. The initial states were $x_{0,1} = 13.49$, $x_{0,2} = 15.50$ and $x_{0,3} = 11.02$. The growth-rate parameters differed among the three trajectories, indicating that the WS, NS and SS did not follow the same average temporal trend during the study period. The Western component showed a slightly positive trend ($u_{WS} = 0.0354$), the WS component showed a slightly negative trend ($u_{NS} = -0.0170$), and the Southern component showed the strongest positive trend ($u_{SS} = 0.1380$). The process-error structure of the best model was diagonal and equal, with $Q_{\text{diag}} = 0.0107$. This means that the three latent trajectories were modelled as having independent temporal process variation, but with the same process variance across components. In other words, the unexplained interannual variability was not allowed to be correlated among trajectories, although the magnitude of this variability was assumed to be common to all of them. The effects of catch differed among trajectories. The estimated catch effect was positive for the WS, although small in magnitude ($\text{Catch}(c)_{WS} = 0.0247$), whereas it was negative for the SS and NS ($\text{Catch}(c)_{SS} = -0.0951$; $\text{Catch}(c)_{NS} = -0.1201$), so the effect of commercial catches on latent biomass dynamics was not homogeneous among stocks. The NAO effect in the best model was estimated as positive and common across trajectories (NAO = 0.0708). This suggests that years with higher NAO values were associated with higher latent biomass dynamics in the model. Since the best-supported model included both catch and NAO, the results suggest that combining a fisheries-related covariate with a large-scale environmental covariate improved the explanation of the observed biomass indices compared with models including only one covariate or no covariates.

The best catch model, Model 2, also supported Hypothesis 9 and estimated three latent trajectories. Its estimated initial states were $x_{0,1} = 13.61$, $x_{0,2} = 15.81$ and $x_{0,3} = 11.04$. The growth-rate pattern was similar to that of the best model, with a positive trend for the Western component ($u_{WS} = 0.03174$), a negative trend for the WS component ($u_{NS} = -0.03097$), and a clearly positive trend for the Southern component ($u_{SS} = 0.13767$). The process variance was higher than in the best model, with $Q_{\text{diag}} = 0.02072$, suggesting that, when NAO was not included, a larger proportion of the temporal variability had to be absorbed by the process-error term. Catch effects again differed among trajectories, with a very small positive effect for the WS ($\text{Catch}(c)_{WS} = 0.00184$) and negative effects for the SS ($\text{Catch}(c)_{SS} = -0.13618$) and NS ($\text{Catch}(c)_{NS} = -0.08511$).

The best model without covariates, Model 12, retained the same Hypothesis 9 structure but had lower support. This model estimated three initial states, x0.1 = 13.58, x0.2 = 15.21 and x0.3 = 11.39, and a common growth parameter of u = 0.0482 for all of the three populations. Its process variance was substantially higher (Q.diag = 0.0556) than in the covariate models. This indicates that, in the absence of catch and NAO, more of the temporal variation in the observed biomass indices remained unexplained and had to be captured by the process-error component.

Overall, the updated MARSS model comparison provides strong support for a three-trajectory structure corresponding to Hypothesis 9. The best model included both catch and NAO, although the catch-only model was also closely supported. The results therefore suggest that the biomass indices are best described by three distinct latent population stocks with different growth rates, and including catch and NAO together improves model fit by reducing unexplained temporal variability.

Table 3.1: Top 12 MARSS models.

Model	Hypothesis	Covariates	Q	u	Z	K	AICc	ΔAICc
1	9	Catch and NAO	Diagonal and equal	Unequal	3	36	731.96	0
2	9	Catch	Diagonal and equal	Unequal	3	35	733.80	1.83
3	9	Catch and NAO	Equal variance covariance	Unequal	3	35	734.42	2.46
4	9	Catch	Equal variance covariance	Unequal	3	34	735.74	3.78
5	5	Catch and NAO	Diagonal and equal	Equal West and North	3	35	737.96	5.99
6	9	Catch and NAO	Unconstrained	Unequal	3	39	738.55	6.58
7	9	Catch and NAO	Diagonal and equal	Unequal	3	34	739.09	7.13
8	5	Catch and NAO	Diagonal and equal	Unequal	3	36	739.22	7.26
9	5	Catch and NAO	Equal variance covariance	Unequal	3	35	740.15	8.18
10	9	Catch and NAO	Diagonal and unequal	Equal	3	34	740.37	8.40
11	9	Catch and NAO	Diagonal and equal	Equal	3	34	740.37	8.40
12	9	None	Diagonal and equal	Equal	3	28	740.40	8.43

Table 3.2: Best MARSS models (all covariates, catch and no covariates).

Model 1 – Z9 (Catch and NAO)		**Model 2 – Z9 (Catch)**		**Model 12 – Z9 (No covariates)**	
Parameter	**Estimate**	**Parameter**	**Estimate**	**Parameter**	**Estimate**
$x_{0,1}$	13.49	$x_{0,1}$	13.61	$x_{0,1}$	13.58
$x_{0,2}$	15.50	$x_{0,2}$	15.81	$x_{0,2}$	15.21
$x_{0,3}$	11.02	$x_{0,3}$	11.04	$x_{0,3}$	11.39
u_{WS}	0.0354	u_{WS}	0.03174	u	0.0482
u_{NS}	-0.0170	u_{NS}	-0.03097	Q_{diag}	0.0556
u_{SS}	0.1380	u_{SS}	0.13767		
Q_{diag}	0.0107	Q_{diag}	0.02072		
Catch(c)–WS	0.0247	Catch(c)–WS	0.00184		
Catch(c)–NS	-0.0951	Catch(c)–NS	-0.13618		
Catch(c)–SS	-0.1201	Catch(c)–SS	-0.08511		
NAO	0.0708				

3.2.1 Models without covariates

The MARSS model without covariates estimated three latent biomass trajectories under the alternative stock structure proposed in Hypothesis 9 (Figure 3.5). This model separated the time series into three main stocks: the NS excluding the English Channel survey (7.d), the SS excluding the northern Spanish area (9.aN), and a third component grouping the WS with the northern Spanish area (division 9.aN) and the English Channel (division 7.d).

The first trajectory, corresponding to the NS without FR-CGFS survey, showed a moderate decline from 2000 to around 2011–2012. After this period, the trajectory stabilized and then displayed a gradual recovery, reaching values close to those observed at the beginning of the time series by the final years. This pattern suggests that the NS experienced a temporary reduction in latent biomass during the first half of the study period, followed by partial recovery in recent years.

The second trajectory, associated with the SS excluding SP-NORTH survey, showed the clearest long-term increase. Latent log-biomass increased steadily from the beginning of the series until approximately 2015, after which the trend became more stable but continued to increase slightly. This indicates a progressive improvement in the biomass signal for the southern component, although the exclusion of SP-NORTH suggests that this northern area may not follow the same dynamics as the rest of the WS.

The third trajectory grouped the WS together with SP-NORTH and the English Channel survey. This component remained relatively stable during the first part of the period, with a slight decline until around 2011, followed by a marked increase between 2013 and 2016. After this increase, the trajectory stabilized at a higher level, with only minor fluctuations in the most recent years. This grouping suggests a certain degree of synchrony among these areas, supporting the hypothesis that they may share similar population dynamics.

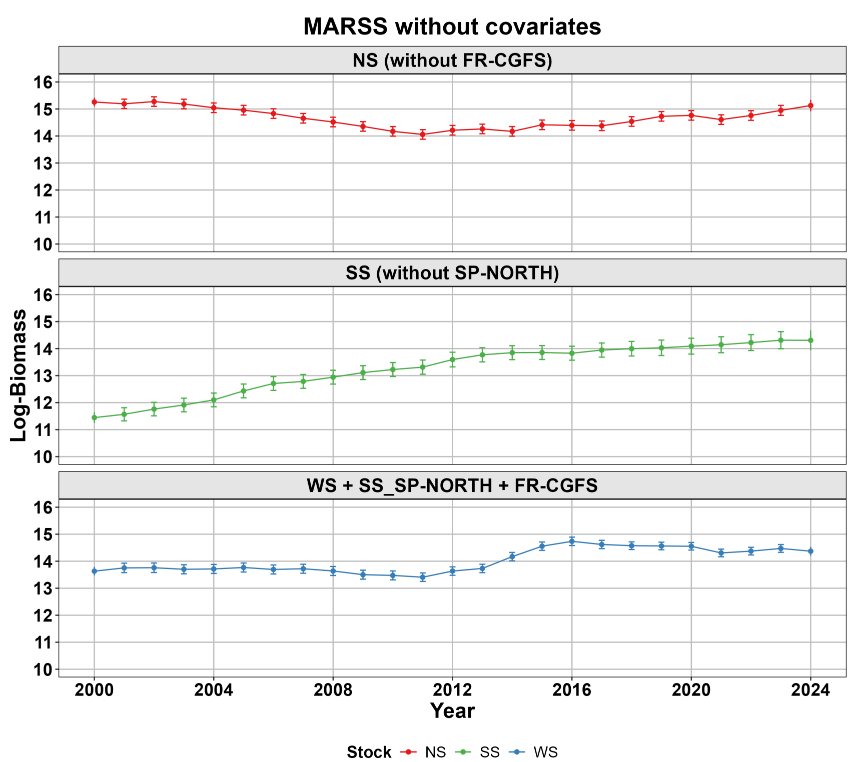

Figure 3.5: Best MARSS model without covariates (Hypothesis 9).

3.2.2 Models with covariates

The MARSS model including both catch and NAO as covariates estimated three latent biomass trajectories under the alternative stock structure proposed in Hypothesis 9 (Figure 3.6). Overall, the temporal patterns were very similar to those obtained in the model without covariates, indicating that the inclusion of catch and NAO did not substantially change the main structure of the latent biomass dynamics. However, this model was better supported according to AICc, suggesting that these covariates helped explain part of the temporal variability in the biomass indices.

The first trajectory, corresponding to the NS excluding the English Channel survey, showed a gradual decline from the beginning of the time series until approximately 2011. After this minimum, the trajectory increased steadily during the second half of the study period. This pattern suggests that the NS experienced a period of decreasing latent biomass in the 2000s, followed by a progressive recovery in more recent years.

The second trajectory, associated with the SS excluding SP-NORTH, showed a clear and sustained increase throughout the study period. The increase was particularly marked during the first half of the time series and then became more gradual after approximately 2014–2015. This indicates that the SS followed a consistently positive latent biomass trend, even after accounting for the effects of catch and NAO.

The third trajectory grouped the WS together with SP-NORTH and the English Channel survey. This component remained relatively stable during the first part of the period, with a slight decline until around 2011, followed by a clear increase between 2013 and 2016. After this increase, the trajectory stabilized at a higher level, with only small fluctuations in the most recent years. The results support the idea that SP-NORTH and FR-CGFS showed biomass dynamics more similar to the WS than to their original management stocks.

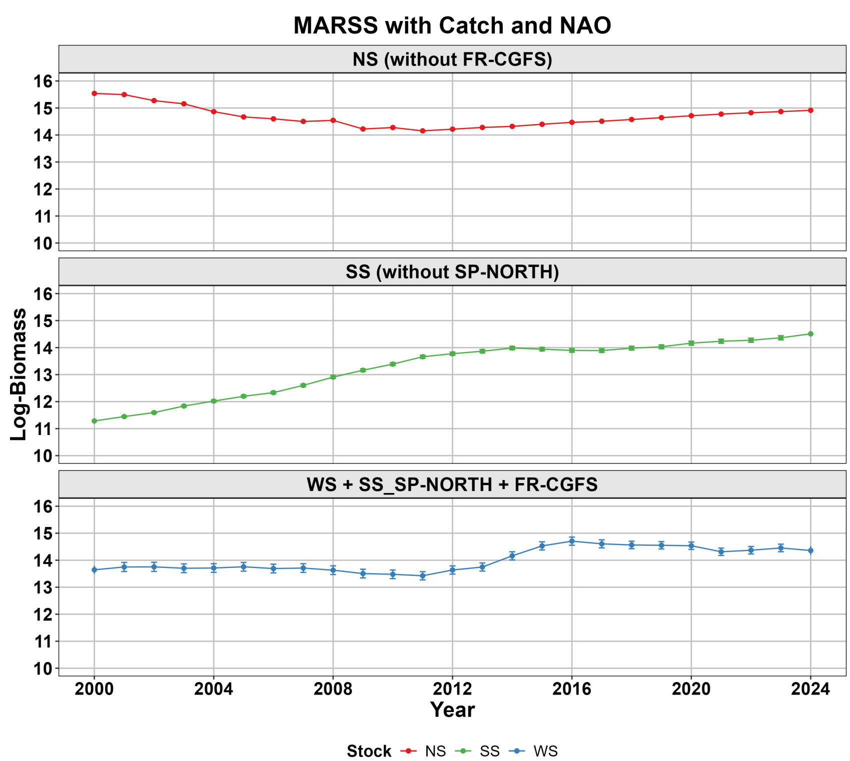

Figure 3.6: Best MARSS model with Catch and NAO (Hypothesis 9).

The MARSS model including catch as a covariate also estimated three latent biomass trajectories under the alternative stock structure proposed in Hypothesis 9 (Figure 3.7)(Figure 20). This was the second-best ranked model in the overall comparison, with a ΔAICc of 1.83 relative to the best model including both catch and NAO. Therefore, the catch-only model also received strong support, suggesting that commercial catch explained an important part of the temporal variation in the observed biomass indices.

The first trajectory, corresponding to the NS excluding the English Channel survey (NS without FR-CGFS), showed a gradual decline from the beginning of the time series until around 2010–2011. After this period, the trajectory increased progressively, reaching higher values again in the most recent years. This pattern indicates that the NS experienced a decline during the first half of the study period, followed by a gradual recovery.

The second trajectory, associated with the SS excluding SP-NORTH, showed a clear and sustained increase throughout the study period. The increase was especially marked from 2000 to approximately 2014, after which the trajectory continued to rise more gradually. This suggests a strong positive biomass trend in the ss, which remained evident after accounting for the effect of catch.

The third trajectory grouped the WS together with SP-NORTH and the English Channel survey (WS + SP-NORTH + FR-CGFS). This component remained relatively stable during the first part of the period, with a slight decline until around 2011–2012. It then increased markedly between 2013 and 2016 and subsequently stabilized at a higher level, with only minor fluctuations. Compared with the model without covariates and the model including both catch and NAO, this trajectory appears smoother.

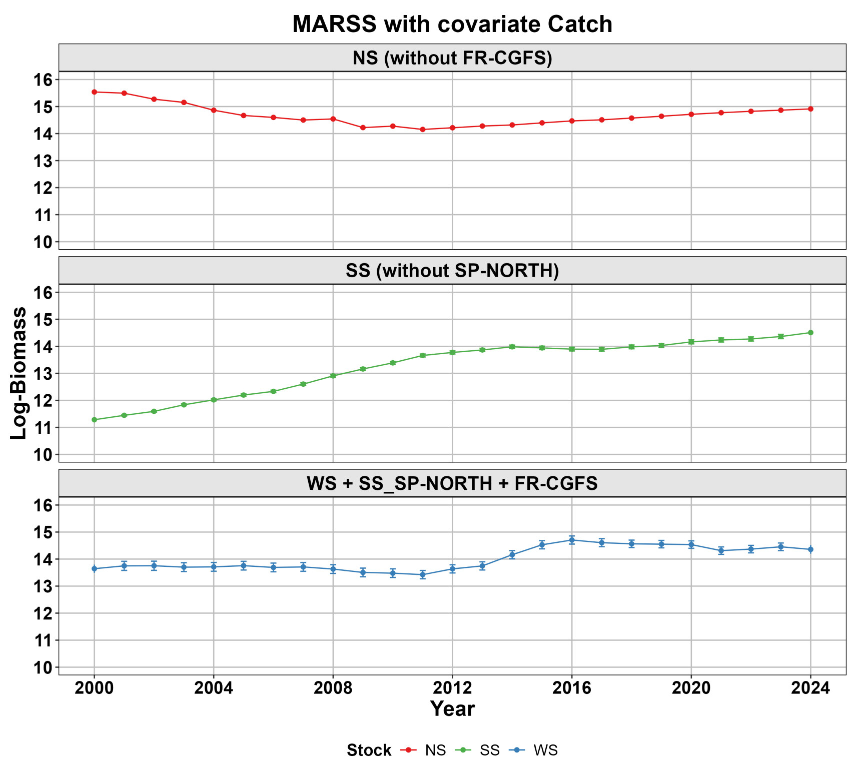

Figure 3.7: Best MARSS model with Catch effect (Hypothesis 9).

Chapter 4

Discussion

Currently, horse mackerel (*Trachurus trachurus*) is managed as three separate stocks in the Northeast Atlantic: the WS, covering Subarea 8 and divisions 2.a, 3.a, 4.a, 5.b, 6.a, 7.a–c and 7.e–k; the WS, covering divisions 4.b, 4.c and 7.d; and the WS, corresponding to Division 9.a (ICES 2025a; ICES 2025b; ICES 2025c).

Understanding the population structure of fish stocks is essential to maintain adequate harvest strategies (Moritz, 2002). The population dynamics of *Trachurus trachurus* in the Northeast Atlantic were examined using bottom-trawl survey indices standardized with sdmTMB. This method was used because it follows a model-based approach, in which biomass indices are estimated through a statistical model that accounts for the spatial and spatio-temporal structure of the data. This approach was considered more appropriate than a purely design-based approach, which relies on detailed information about the original survey design and sampling scheme. Given that multiple surveys were combined in the analysis and that complete design information was not consistently available for all surveys, a model-based framework provided a more robust and coherent alternative. The model-based approach was more suitable because surveys differed in spatial coverage, sampling and temporal length, so it offered a more consistent basis for comparing the WS, NS and SS over the period 2000–2024.

The standardized biomass indices were used, to evaluate whether the three management units of *Trachurus trachurus* in the Northeast Atlantic show sufficiently similar temporal dynamics to be considered a single population for management purposes, or whether they support differentiated population dynamics with MARSS models. MARSS models were then used to compare alternative population structure hypotheses for horse mackerel. The best supported model was based on Hypothesis 9 and included both catch and NAO as covariates.

The inclusion of commercial catches should be interpreted with caution. Catch effects were estimated separately for each latent trajectory, and they were not equal among the three stocks. In the best-supported model, the catch effect was slightly positive for the WS trajectory, whereas it was negative for the SS and NS trajectories. At first, one might expect a negative relationship between catches and biomass, because commercial fishing effort can contribute to biomass decline. However, catches may also be higher in years when biomass dynamics increase or when management advice allows greater fishing opportunities (Punt y Hilborn 1996). In that case, catch does not only act as a direct pressure variable but may also partly follow the underlying biomass dynamics. Therefore, the support for a positive catch in the WS should be understood as evidence that fishing-related information improves the explanation of the latent biomass trajectories, but not as a unique causal relationship for biomass dynamics.

Many studies have shown that fish populations may respond to environmental variability (Maureaud et al. 2020), particularly through changes in recruitment, survival, prey availability or spatial distribution (Ubeda et al. 2023; Nogueira et al. 2018). Environmental drivers may be relevant for horse mackerel because the species has a wide geographic distribution, a long spawning season and seasonal movements between spawning, feeding and overwintering areas (ICES 2025). In this thesis, NAO was included as a large-scale climatic covariate. The best-supported model included both catch and NAO, and the estimated NAO effect was positive and common across the three latent trajectories. This suggests that years with higher NAO values were associated with higher latent biomass dynamics in the model. However, this effect should be interpreted cautiously, because NAO is a broad climatic index and may reflect several indirect environmental processes rather than a single factor acting on horse mackerel.

Our results support treating *Trachurus trachurus* in the Northeast Atlantic as three differentiated population units for management. These three trajectories showed distinct temporal dynamics, long-term growth rates and catch stock covariates, indicating that the main population stocks did not follow the same biomass trend during the study period (2000-2024). The results obtained suggest that horse mackerel biomass dynamics are not fully homogeneous with the current ICES stock delineation, as the northern part of Division 9.a, represented by SP-NORTH survey, and Division 7.d of the English Channel, represented by FR-CGFS survey, were better associated with the WS than with their current stock assignment.

These results are consistent with previous findings regarding boundaries between the WS and SS. Before the HOMSIR project, ICES stock delineation was mainly based on spawning areas and egg distribution, but no clear boundary in egg distribution was recognizable between the WS and SS. The HOMSIR project addressed this uncertainty using a multidisciplinary stock identification approach, including genetic markers, morphometrics, otolith shape, parasites as biological tags, and life-history traits (HOMSIR project 2003). Its main conclusion for the Northeast Atlantic was that the previous boundary between the WS and SS needed to be revised, with the SS restricted to the western Atlantic coast of the Iberian Peninsula south of Cape Finisterre, and the WS extending from Cape Finisterre northwards along the western European coast to Norway. This revision is relevant because Cape Finisterre approximately corresponds to the limit between ICES Divisions 8.c and 9.a. Therefore, areas north of this boundary, such as the Cantabrian Sea and northern Galicia in Division 8.c, were interpreted as more closely related to the WS, whereas Division 9.a remained within the SS. Abaunza et al. (2008) reported that horse mackerel from the Portuguese coast, corresponding mainly to Division 9.a, differed from those in more northern Atlantic areas, whereas individuals from the Cantabrian Sea and northern Galicia, in division 8.c, showed characteristics closer to those of the WS. Taken together, my findings provide additional support for the inclusion of Division 8.c within the WS.

More recent genetic evidence indicated that the WS population may extend southwards into Division 9.a, while the SS appears to be more clearly represented by samples from the southern part of 9.a (Farell et al. 2024). Therefore, the association of SP-NORTH with the WS suggests that northern 9.a may contain a stronger Western connection than assumed by the current ICES delineation. The second relevant boundary area is the English Channel. Farrell et al. (2024) carried out genetic approach research using different marker panels to evaluate whether individuals sampled in boundary areas were assigned to the WS or NS populations. Their results showed that samples from the eastern English Channel, corresponding to Division 7.d, were not exclusively of NS origin. Depending on the marker panel used, 35–38% of individuals from Division 7.d were assigned to the WS population. In the western English Channel, corresponding to Division 7.e, the Western connection was even stronger, with 86–89% of individuals assigned to the WS. These findings indicate that the English Channel should not be interpreted as a purely NS area, but rather as a transition or mixed zone with both WS and NS.

An additional extension would be to explore models with a more flexible autoregressive matrix B. In the MARSS models fitted in this thesis, each latent trajectory depended only on its own previous state, meaning that the population stocks were modelled as independent temporal processes. A more complex formulation could allow off-diagonal elements in the B matrix, so that the previous state of one trajectory could influence another trajectory. For instance, this could be useful for testing whether changes in the Western trajectory are followed by changes in the English Channel or northern 9.a stocks in subsequent years, which would be consistent with movement, delayed connectivity or shared dynamics between neighbouring areas. In this way, the model could evaluate not only whether areas share similar biomass trends, but also whether temporal changes in one component are associated with later changes in another.

The MARSS approach (Homes et al. 2012) used in this thesis provides a formal statistical framework for testing alternative population structure hypotheses from multiple biomass time series to assess population dynamics. Therefore, different biological hypotheses about stock structure can be translated into alternative statistical models and compared using AIC (Cavanaugh and Shumway, 1997; Stoffer and Wall, 1991). However, MARSS does not directly identify the biological mechanisms or environmental factors causing the observed population structure (Nogueira et al. 2018). However, the patterns of synchrony, divergence and correlation among latent trajectories can help infer which mechanisms or processes are likely to be relevant (Ward et al. 2010). Results should be interpreted as statistical evidence of coherent biomass dynamics, rather than as direct proof of biological population boundaries. It is also important to recognize that any modelling approach represents a best-fit solution rather than a perfect description of the biological system (Stringham et al. 2003). The selected MARSS model identifies the hypothesis with the greatest relative support among the candidate models considered, but it does not rule out the influence of other ecological, or fisheries processes not included in the analysis.

MARSS can combine several observed time series through the scaling parameter a (Holmes et al. 2012; Holmes et al. 2020). This parameter was necessary because several standardized survey biomass indices were assigned to the same latent trajectory in some hypotheses. The scaling parameter allows each observed survey series to have its own average level while still being linked to the same underlying latent biomass trajectory. Therefore, a is not used to directly correct for differences in survey design, country, gear or catchability. Instead, it allows the model to align survey series so that the comparison among hypotheses is based mainly on whether their temporal dynamics are coherent with the same latent trajectory. For this reason, the interpretation of the Z matrix and the scaling parameter a must be considered together. The Z matrix defines the biological hypothesis being tested by assigning each observed survey series to a latent population stock. The scaling parameter a then allows each survey series to be assigned to the state or stock and accounts for differences in the average level.

Future work should implement further stock identification methods for horse mackerel. As mentioned before, MARSS provided statistical evidence among standardized biomass indices, but it cannot identify the biological origin of horse mackerel. This limitation is particularly relevant in this study, specifically regarding the northern Division 9a and the division 7.d of the English Channel, where the SP-NORTH and FR-CGFS surveys were linked to the WS. The approach developed by Farrell et al. (2024) provides a useful and up to date framework for this purpose, as its research was able to distinguish individuals from the WS and NS populations with high accuracy and was used to quantify mixed stock composition in areas such as Division 7.d and 7.e. Then, future studies may continue to apply genetic assignment methods to these boundary areas to estimate the proportion of WS, SS and NS individuals present in each study area. A similar strategy could be applied as complementary approach for this thesis. For northern 9.a, genetic assignment would help determine whether the WS biomass trajectory reflects a real southward extension of the WS or a mixture between the WS and SS. In Division 7.d, it would be useful to assess whether the FR-CGFS survey follows the same trend as the WS, as the eastern part of the English Channel contains a significant western component. Nevertheless, future applications should not only rely on genetics. The HOMSIR project showed that horse mackerel stock identification benefits from a holistic approach, combining genetic markers with otolith shape, body morphometrics, parasites as biological tags and life-history traits (HOMSIR 2003). These methods could provide complementary information for this research: genetics can assign individuals to population of origin, otoliths and morphometrics can detect phenotypic differences linked to growth environment, parasites can indicate species habitat or movements, and life-history traits can reveal differences in growth, reproduction and distribution. Combining these approaches would make it possible to evaluate whether the latent trajectories identified by MARSS correspond to true biological population units or to mixed survey connections in transition areas. Therefore, the MARSS results should be considered as complementary evidence rather than as a complete stock identification on their own.

Overall, this thesis shows that combining model-based survey standardization and MARSS modelling provides a useful framework for evaluating stock structure in *Trachurus trachurus* from a population-dynamics perspective. The standardized bottom-trawl survey indices revealed heterogeneous biomass patterns among areas, and the MARSS comparison showed that these patterns were better represented by three differentiated latent trajectories, whose boundaries differed from those of the current management units. In addition, further genetic and other stock identification techniques, would help to further validate the population structure identified here and support the development of biologically meaningful management units.

Chapter 5

Conclusions

The main conclusions of this thesis are:

- The shift from a design base approach towards a model base approach, with sdmTMB, allowed biomass indices to be estimated from bottom-trawl survey data while accounting for spatial and spatio-temporal heterogeneity. It was particularly relevant as surveys differed in spatial coverage, time series and sampling.

- The best-fitting model showed three distinct state processes (trends in biomass), suggesting that the current ICES stock delineation for horse mackerel in the Northeast Atlantic is not entirely accurate. The results support the recognition of three population units: (1) the southern stock, (2) the western stock (encompassing the northern part of Atlantic Iberian waters (subdivision 9.a.n) and the English/French Channel), and (3) the North Sea stock.

- The best-model included NAO and catch as covariates suggesting that commercial catches and NAO contributed to explaining part of the temporal variation in the biomass indices. Catch may be influenced not only by fishing effort, but also by biomass availability and management decisions, while NAO may represent indirect environmental processes affecting recruitment, distribution or survival.

- Future research should adopt a holistic framework that integrates time-series modelling with a multidisciplinary suite of stock identification tools, including genetics, otolith shape analysis, morphometrics, parasite tagging and life-history trait analysis, to achieve a comprehensive understanding of population structure and dynamics.

Bibliography

[1] Abaunza Martínez P (2007) Teoría y práctica en la identificación de stocks de peces de interés comercial. El jurel (*Trachurus trachurus*) como ejemplo de una aproximación holística a la identificación de stocks. Tesis doctoral, Universidad del País Vasco.

[2] Abaunza P, Gordo L, Karlou-Riga C, Murta A, Zimmermann C, Hammer C, Lucio P, Iversen SA, Molloy J, Gallo E (2003) Growth and reproduction of horse mackerel, *Trachurus trachurus* (Carangidae).

[3] Abaunza P, Murta AG, Campbell N, Cimmaruta R, Comesaña AS, Dahle G, García Santamaría MT, Gordo LS, Iversen SA, MacKenzie K, Magoulas A, Mattiucci S, Molloy J, Nascetti G, Pinto AL, Quinta R, Ramos P, Sanjuan A, Santos AT, Zimmermann C (2008) Stock identity of horse mackerel (*Trachurus trachurus*) in the Northeast Atlantic and Mediterranean Sea: Integrating the results from different stock identification approaches. Fisheries Research 89(2):196–209. https://doi.org/10.1016/j.fishres.2007.09.022

[4] Anderson SC, Ward EJ, English PA, Barnett LAK, Thorson JT (2022) sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. https://doi.org/10.1101/2022.03.24.485545

[5] Banerjee S, Gelfand AE, Finley AO, Sang H (2008) Gaussian predictive process models for large spatial data sets. Journal of the Royal Statistical Society: Series B 70(4):825–848.

[6] Bhadury S, Ghosh S, Kumar P (2019) Nowcasting GDP growth using a coincident economic indicator for India.

[7] Bowdich SL (1825) Fishes of Madeira. In: Excursions in Madeira and Porto Santo during the autumn of 1823, while on his third voyage to Africa, pp 121–125, 233–238.

[8] Brunel T, Farrell ED, Kotterman M, Kwadijk C, Verkempynck R, Chen C, Miller D (2016) Improving the knowledge basis for advice on North Sea horse mackerel: Developing new methods to get insight on stock boundaries and abundance. Wageningen Marine Research Report C092/16. IMARES Wageningen UR.

[9] Carroll G, et al. (2019) A review of methods for quantifying spatial predator–prey overlap. Global Ecology and Biogeography 28:1561–1577.

[10] Carvalho GR, Hauser TJ (1994) Molecular genetics and the stock concept in fisheries. Reviews in Fish Biology and Fisheries 4:326–350.

[11] Cavanaugh JE, Shumway RH (1997) A bootstrap variant of AIC for state-space model selection. Statistica Sinica 7:473–496.

[12] Chilès JP, Delfiner P (1999) Geostatistics: Modeling spatial uncertainty. John Wiley & Sons.

[13] Cochran WG (1977) Sampling techniques. 3rd ed. John Wiley & Sons.

[14] Cressie N (1993) Statistics for spatial data. John Wiley & Sons.

[15] Cressie NAC, Wikle CK (2011) Statistics for spatio-temporal data. John Wiley & Sons.

[16] Cronne L (2017) SISP 15: Manual of the IBTS North Eastern Atlantic surveys. https://doi.org/10.17895/ICES.PUB.3519

[17] Datta A, Banerjee S, Finley AO, Gelfand AE (2016) Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. Journal of the American Statistical Association 111(514):800–812. https://doi.org/10.1080/01621459.2015.1044091

[18] Dencker TS, et al. (2017) Temporal and spatial differences between taxonomic and trait biodiversity in a large marine ecosystem: Causes and consequences. PLOS ONE 12:e0189731.

[19] Diggle PJ, Ribeiro PJ (2007) Model-based geostatistics. Springer.

[20] Dransfeld L, Dwane O, Molloy J, Gallagher S, Reid DG (2005) Estimation of mackerel (*Scomber scombrus* L., 1758) and horse mackerel (*Trachurus trachurus* L., 1758) daily egg production outside the standard ICES survey area. ICES Journal of Marine Science 62:1705–1710.

[21] Durbin J, Koopman SJ (2001) Time series analysis by state space methods. Oxford University Press.

[22] Eaton DR (1983) Scad in the North-East Atlantic. MAFF Laboratory Leaflet No. 56. MAFF.

[23] Farrell ED, Carlsson J (2018) Genetic stock identification of Northeast Atlantic horse mackerel, *Trachurus trachurus*. Report prepared for the members of the Northern Pelagic Working Group.

[24] Farrell ED, Fuentes-Pardo AP, Murray DS, Andersson L (2024) A working document on the genetic stock identification of horse mackerel, *Trachurus trachurus*, for the ICES Benchmark Workshop on Horse Mackerel and Boarfish Stocks (WKBHMB). Version 2.0. Working document.

[25] Finley AO, Datta A, Banerjee S (2022) spNNGP R package for nearest neighbor Gaussian process models. Journal of Statistical Software 103(5). https://doi.org/10.18637/jss.v103.i05

[26] Friedland KD, et al. (2020) Changes in higher trophic level productivity, diversity and niche space in a rapidly warming continental shelf ecosystem. Science of the Total Environment 704:135270.

[27] Froese R, Pauly D (2015) Welcome to FishBase. In: Froese R, Pauly D (eds) FishBase, version 10/2015. http://www.fishbase.org/manual/English/PDF/Welcome_to_FishBase_RFroese_etal2015.pdf

[28] Fuentes-Pardo AP, Pettersson M, Sprehn CG, Andersson L, Farrell E (2020) Population structure of the Atlantic horse mackerel (*Trachurus trachurus*) revealed by whole-genome sequencing. Report prepared for the members of the Northern Pelagic Working Group and the Pelagic Advisory Council.

[29] Fuentes-Pardo AP, Farrell ED, Pettersson M, Sprehn CG, Andersson L (2023) The genomic basis and environmental correlates of local adaptation in the Atlantic horse mackerel (*Trachurus trachurus*). Evolutionary Applications 16:1201–1219.

[30] Gauldie RW (1988) Tagging and genetically isolated stocks of fish: A test of one stock hypothesis and the development of another. Journal of Applied Ichthyology 4:168–173.

[31] Gulland JA (1971) Ecological aspects of fishery research. Advances in Ecological Research 7:115–176.

[32] Hare JA (2005) The use of early life stages in stock identification studies. In: Cadrin SX, Friedland KD, Waldman JR (eds) Stock identification methods: Applications in fishery science. Elsevier, pp 89–117.

[33] Harley SJ, Myers RA, Dunn A (2001) Is catch-per-unit-effort proportional to abundance? Canadian Journal of Fisheries and Aquatic Sciences 58(9):1760–1772. https://doi.org/10.1139/f01-112

[34] Harvey AC (1989) Forecasting, structural time series models and the Kalman filter. Cambridge University Press.

[35] Hilborn R, Walters CJ (1992) Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman & Hall.

[36] Hinrichsen RA (2009) Population viability analysis for several populations using multivariate state-space models. Ecological Modelling 220(9–10):1197–1202.

[37] Hinrichsen RA, Holmes EE (2009) Using multivariate state-space models to study spatial structure and dynamics. In: Cantrell RS, Cosner C, Ruan S (eds) Spatial ecology. CRC/Chapman Hall.

[38] Holmes EE (2001) Estimating risks in declining populations with poor data. Proceedings of the National Academy of Sciences of the United States of America 98(9):5072–5077.

[39] Holmes EE, Sabo JL, Viscido SV, Fagan WF (2007) A statistical approach to quasi-extinction forecasting. Ecology Letters 10(12):1182–1198.

[40] Holmes E, Ward EJ, Scheuerell MD (2021) Analysis of multivariate time series using the MARSS package. Version 3.11.4. https://doi.org/10.5281/ZENODO.5781847

[41] Holmes E, Ward EJ, Wills K (2012) MARSS: Multivariate autoregressive state-space models for analyzing time-series data. The R Journal 4(1):11. https://doi.org/10.32614/RJ-2012-002

[42] Holmes E, Ward EJ, Wills K (2012) MARSS: Multivariate autoregressive state-space models for analyzing time-series data. The R Journal 4(1):11. https://doi.org/10.32614/RJ-2012-002

[43] HOMSIR (2003) A multidisciplinary approach using genetic markers and biological tags in horse mackerel (*Trachurus trachurus*) stock structure analysis. Final report, QLK5-CT1999-01438.

[44] Humbert JY, Mills LS, Horne JS, Dennis B (2009) A better way to estimate population trends. Oikos 118(12):1940–1946.

[45] ICES (1998) Report of the Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy, 9–18 September 1997, ICES Headquarters, Copenhagen. ICES CM 1998/Assess:6, Part 1.

[46] ICES. (2005). *ICES FishMap species factsheet: Horse mackerel*. International Council for the Exploration of the Sea. https://www.ices.dk/about-ICES/projects/EU-RFP/EU%20Repository/ICES%20FishMap/ICES%20FishMap%20species%20factsheet-horsemackerel.pdf

[47] ICES (2015) Report of the Workshop on Maturity Staging of Mackerel and Horse Mackerel (WKMSMAC2), 28 September–2 October 2015, Lisbon, Portugal. ICES CM 2015/SSGIEOM:17.

[48] ICES (2024) Benchmark workshop on horse mackerel and boarfish (WKBHMB). ICES Scientific Reports 6(8):296 pp. https://doi.org/10.17895/ices.pub.2500

[49] ICES (2025a) Stock annex: Horse mackerel (*Trachurus trachurus*) in divisions 4.b-c and 7.d (Southern and central North Sea, eastern English Channel). ICES Stock Annexes. https://doi.org/10.17895/ICES.PUB.29655221

[50] ICES (2025b) Stock annex: Horse mackerel (*Trachurus trachurus*) in Subarea 8 and divisions 2.a, 3.a, 4.a, 5.b, 6.a, 7.a–c, e–k (the Northeast Atlantic). ICES Stock Annexes. https://doi.org/10.17895/ICES.PUB.29655605

[51] ICES (2025c) Horse mackerel (*Trachurus trachurus*) in Division 9a (Atlantic Iberian waters). ICES Stock Annexes. https://doi.org/10.17895/ICES.PUB.26004166

[52] ICES (2026a) DATRAS: Download data products. ICES. https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx

[53] ICES (2026b) DATRAS: Who we are. ICES. https://www.ices.dk/about-ICES/who-we-are/Pages/Who-we-are.aspx

[54] Jardim E, Ribeiro PJ Jr (2007) Geostatistical assessment of sampling designs for Portuguese bottom trawl surveys. Fisheries Research 85(3):239–247.

[55] Jennings S, et al. (2002) Long-term trends in the trophic structure of the North Sea fish community: Evidence from stable-isotope analysis, size-spectra and community metrics. Marine Biology 141:1085–1097.

[56] Kristensen K, Nielsen A, Berg CW, Skaug H, Bell BM (2016) TMB: Automatic differentiation and Laplace approximation. Journal of Statistical Software 70(5):1–21. https://doi.org/10.18637/jss.v070.i05

[57] Latimer AM, Banerjee S, Sang H Jr, Mosher ES, Silander JA Jr (2009) Hierarchical models facilitate spatial analysis of large data sets: A case study on invasive plant species in the northeastern United States. Ecology Letters 12(2):144–154. https://doi.org/10.1111/j.1461-0248.2008.01270.x

[58] Lindgren F (2023) fmesher: Triangle meshes and related geometry tools. R package version 0.1.2. https://CRAN.R-project.org/package=fmesher

[59] Lindgren F, Rue H (2015) Bayesian spatial modelling with R-INLA. Journal of Statistical Software 63(1):1–25. https://doi.org/10.18637/jss.v063.i19

[60] Lindgren F, Bolin D, Rue H (2022) The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running. Spatial Statistics 50:100599. https://doi.org/10.1016/j.spasta.2022.100599

[61] Lindgren F, Rue H, Lindström J (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B 73(4):423–498. https://doi.org/10.1111/j.1467-9868.2011.00777.x

[62] Lockwood SJ, Johnson PO (1977) Horse mackerel. Laboratory Leaflet No. 38. MAFF Directorate of Fisheries Research, Lowestoft.

[63] Macer CT (1977) Some aspects of the biology of the horse mackerel [*Trachurus trachurus* (L.)] in waters around Britain. Journal of Fish Biology 10:51–62.

[64] Magurran AE, Dornelas M, Moyes F, Gotelli NJ, McGill B (2015) Rapid biotic homogenization of marine fish assemblages. Nature Communications 6:8405.

[65] Matérn B (1986) Spatial variation. 2nd ed. Springer.

[66] Maureaud AA, Palacios-Abrantes J, Kitchel Z, Mannocci L, Pinsky ML, Fredston A, Beukhof E, Forrest DL, Frelat R, Palomares MLD, Pecuchet L, Thorson JT, Van Denderen PD, Mérigot B (2024) FISHGLOB_data: An integrated dataset of fish biodiversity sampled with scientific bottom-trawl surveys. Scientific Data 11(1):24. https://doi.org/10.1038/s41597-023-02866-w

[67] Metaxoglou K, Smith A (2007) Maximum likelihood estimation of VARMA models using a state-space EM algorithm. Journal of Time Series Analysis 28:666–685.

[68] Mildenberger TK, Maioli F, Berg CW (2025) DATRASextra: Extra functions for DTU Aqua's DATRAS R package. R package. https://github.com/tokami/DATRASextra

[69] Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51:238–254.

[70] National Centers for Environmental Information (2026) North Atlantic Oscillation (NAO). National Oceanic and Atmospheric Administration. https://www.ncei.noaa.gov/access/monitoring/nao/

[71] Nogueira A, Tolimieri N, González-Troncoso D (2019) Using multivariate state-space models to examine commercial stocks of redfish (*Sebastes* spp.) on the Flemish Cap. Canadian Journal of Fisheries and Aquatic Sciences 76(2):208–216. https://doi.org/10.1139/cjfas-2017-0435

[72] Olivença DV, Davis JD, Voit EO (2022) Inference of dynamic interaction networks: A comparison between Lotka–Volterra and multivariate autoregressive models. Frontiers in Bioinformatics 2:1021838. https://doi.org/10.3389/fbinf.2022.1021838

[73] Osgood-Zimmerman A, Wakefield J (2023) A statistical review of Template Model Builder: A flexible tool for spatial modelling. International Statistical Review 91(2):318–342. https://doi.org/10.1111/insr.12534

[74] Pavlyuk D (2017) Short-term traffic forecasting using multivariate autoregressive models. Procedia Engineering 178:57–66. https://doi.org/10.1016/j.proeng.2017.01.062

[75] Pebesma E (2018) Simple features for R: Standardized support for spatial vector data. The R Journal 10(1):439–446. https://doi.org/10.32614/RJ-2018-009

[76] Pecuchet L, Törnroos A, Lindegren M (2016) Patterns and drivers of fish community assembly in a large marine ecosystem. Marine Ecology Progress Series 546:239–248.

[77] Pennington M, Grosslein MD (1978) Accuracy of abundance indices based on stratified-random trawl surveys. ICNAF Research Document 78/VI/77.

[78] Polonsky AS (1965) The horse mackerel of the Eastern Atlantic and its fishery. Rybnoe Khozyaistvo 41(6):8–10; 41(7):13–15.

[79] Postuma KH (1978) Immigration of southern fish into the North Sea. Rapports et Procès-Verbaux du Conseil International pour l'Exploration de la Mer 172:225–229.

[80] Punt A, Hilborn R (1996) Biomass dynamic models: User's manual. Computerized Information Series Fisheries No. 10. FAO.

[81] Rue H, Held L (2005) Gaussian Markov random fields: Theory and applications. CRC Press.

[82] Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society: Series B 71(2):319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

[83] Rufener M-C, Kristensen K, Nielsen JR, Bastardie F (2021) Bridging the gap between commercial fisheries and survey data to model the spatiotemporal dynamics of marine species. Ecological Applications 31:e02453.

[84] Schnute T (1994) A general framework for developing sequential fisheries models. Canadian Journal of Fisheries and Aquatic Sciences 51:1676–1688.

[85] Secor DH (1999) Specifying divergent migrations in the concept of stock: The contingent hypothesis. Fisheries Research 43:13–34.

[86] Selden RL, Batt RD, Saba VS, Pinsky ML (2018) Diversity in thermal affinity among key piscivores buffers impacts of ocean warming on predator–prey interactions. Global Change Biology 24:117–131.

[87] Shaklee JB, Currens K (2003) Genetic stock identification and risk assessment. In: Hallerman E (ed) Population genetics: Principles and applications for fisheries scientists. American Fisheries Society, pp 291–328.

[88] Shelton AO, Thorson JT, Ward EJ, Feist BE (2014) Spatial semiparametric models improve estimates of species abundance and distribution. Canadian Journal of Fisheries and Aquatic Sciences 71(11):1655–1666.

[89] Shumway RH, Stoffer DS (1982) An approach to time series smoothing and forecasting using the EM algorithm. Journal of Time Series Analysis 3(4):253–264.

[90] Smith SJ (1996) Analysis of data from bottom trawl surveys. In: Assessment of groundfish stocks based on bottom trawl survey results. NAFO Scientific Council Studies No. 28.

[91] Smith-Vaniz WF (1986) Carangidae. In: Whitehead PJP, Bauchot M-L, Hureau M, Neilsen J, Tortonese E (eds) Fishes of the North-eastern Atlantic and the Mediterranean, vol 2. UNESCO, pp 815–844.

[92] Steindachner F (1868) Ichthyologischer Bericht über eine nach Spanien und Portugal unternommene Reise. (V. Fortsetzung.) Übersicht der Meeresfische an den Küsten Spanien's und Portugal's. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch-Naturwissenschaftliche Classe 57(1):351–424.

[93] Stoffer DS, Wall KD (1991) Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter. Journal of the American Statistical Association 86(416):1024–1033.

[94] Stringham TK, Krueger WC, Shaver PL (2003) State and transition modeling: An ecological process approach. Journal of Range Management 56:106–113. https://doi.org/10.2307/4003893

[95] Thorson JT (2019) Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210:143–161. https://doi.org/10.1016/j.fishres.2018.10.013

[96] Thorson JT, Kristensen K (2016) Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. Fisheries Research 175:66–74. https://doi.org/10.1016/j.fishres.2015.11.016

[97] Thorson JT, Kristensen K (2024) Spatio-temporal models for ecologists. CRC Press.

[98] Thorson JT, Skaug HJ, Kristensen K, Shelton AO, Ward EJ, Harms JH, Benante JA (2015a) The importance of spatial models for estimating the strength of density dependence. Ecology 96(5):1202–1212. https://doi.org/10.1890/14-0739.1

[99] Thorson JT, Shelton AO, Ward EJ, Skaug HJ (2015b) Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science 72(5):1297–1310.

[100] Tweedie MCK (1984) An index which distinguishes between some important exponential families. In: Ghosh JK, Roy J (eds) Statistics: Applications and new directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference. Indian Statistical Institute, pp 579–604.

[101] Úbeda J, Nogueira A, Tolimieri N, Vihtakari M, Elvarsson B, Treble M, Boje J (2023) Using multivariate autoregressive state-space models to examine stock structure of Greenland halibut in the North Atlantic. Fisheries Management and Ecology 30(5):521–535. https://doi.org/10.1111/fme.12639

[102] Waldman JR (1999) The importance of comparative studies in stock analysis. Fisheries Research 43:237–246.

[103] Ward EJ, Chirakkal H, González-Suárez M, Aurioles-Gamboa D, Holmes EE, Gerber L (2010) Inferring spatial structure from time-series data: Using multivariate state-space models to detect metapopulation structure of California sea lions in the Gulf of California, Mexico. Journal of Applied Ecology 47(1):47–56.

[104] Whittle P (1954) On stationary processes in the plane. Biometrika 41(3/4):434. https://doi.org/10.2307/2332724

[105] Yalcin S, Thorson JT, Punt AE (2023) Exploring the limits of spatiotemporal and design-based index standardization under reduced survey sampling. ICES Journal of Marine Science 80(9):2368–2382.