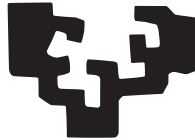


eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

**From habitat to management: a simulation
framework for improving statistical methods
in fisheries science**

Leire Citores Martinez

2021

Supervised by:

Leire Ibaibarriaga Contreras

Dae-Jin Lee

A thesis submitted for the degree of Philosophy Doctor in
Mathematics & Statistics



**From habitat to management: a simulation
framework for improving statistical methods
in fisheries science**

Leire Citores Martinez

2021

Supervised by:
Leire Ibaibarriaga Contreras
Dae-Jin Lee

A thesis submitted for the degree of Philosophy Doctor in
Mathematics & Statistics

This research was supported by a PhD grant from AZTI, by the Basque Government through the BERC 2014-2017 and BERC 2018-2021 program, by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013-0323 and by the Spanish Ministry of Science, Innovation and Universities: BCAM Severo Ochoa accreditation SEV-2017-0718 and through the MTM2017-82379-R project funded by (AEI/FEDER, UE) and acronym “AF-TERAM”.

Itsasoari lotuta ikustezin jardun duten saregile, sardinera, itsasoko emazte eta orokorrean gure herrien zaintzaile diren emakume guztiei.

Esker onak - Acknowledgements

Lehenik eta behin eskerrak eman nahi dizkiet zuzendari izan ditudan Leire Ibaibarriagari eta Dae-Jin Leeri. Muchas gracias por los consejos, paciencia y en general todo lo que me habéis enseñado, y por el ánimo y apoyo constante que me habéis transmitido. Eskerrik asko lan honetan jarritako energia guztiagatik eta urte haue-tan guztietan eskainitako gertutasun eta babesagatik. Mila esker nire tutore izan den Inmaculada Arostegiri ere, matematika ikasle nintzen garaitik, bioestatistikako praktiketan hasteko aukera eskaini eta bide honetan gertu eta laguntzeko prest egon izanagatik. I would also like to thank Altea Lorenzo and Richard Hillary for the revision and favourable report of this thesis.

Ezin ahaztu lan hau burutzeko bidean jasotako beste laguntza asko. Eskerrik asko Andres Uriarte zure jakintza partekatzeagatik, errebisioengatik, ideiak emateko prest azaltzeagatik eta arrantza eta itsasoarekiko interesa gogoz transmititzeagatik. Gracias a Guillem Chust por la ayuda para sumergirme en el mundo de los modelos de distribución de especies, gracias por las discusiones sobre este tema. Eskerrik asko AZTIko ModFish matematikari taldeari ere, erreferente eta babesa izatearren, eta bereziki Sonia Sanchezi, sardinaren inguruko elkarlana erreza egiteagatik. I would also like to thank Mark Brewer from Scotland for the opportunity of doing a three month stay in BioSS and the collaboration for some of the work in this thesis. Muchas gracias a Altea por ser tan buena anfitriona en Aberdeen, por hacer que todo fuese más fácil y agradable y gracias por compartir tu tiempo con Carmen y conmigo, gracias a las dos.

Eskerrak eman nahi nizkieke, hemen AZTIIn, Sukarrietan eta Pasaian, eta Bilbon BCAMen izan ditudan ikaskide eta lankideei. Mila esker Maite, Arkaitz, Elsa,

Nagore eta Sukarrietako bekadunen gelatik pasa diren beste askori, giro aparta sortu eta lana aurrera ateratzeko leku atsegina egitearren, eta bereziki eskerrik asko Natalia eta Itsasori, lankide, pixukide eta lagun ere izatearren eta oraindik ere ematen didazuen babesagatik ere. Bi zentruetan gertu euki dudan Mireni ere eskerrik asko partekatutako tartetxo eta elkarrizketengatik, eta baita Amaiari ere, Eskoziako es-tantzia alaitzearren eta tesi bukaerarako aholku guztiengatik. Muchas gracias a Joaquín por los consejos, apoyo y buena compañía en el BCAM. Muchas gracias también a las R-Ladies, Diana, Lore, Amaia I., Amaia A. y Coté por las ideas, charlas interesantes y apoyo, fue corto, pero me alegro de haber compartido esos momentos y teneros como referente también en este camino.

Bukatzeko, mila esker bide honetako momenturen batean babeseta eta indarra eman didan inguruko jende guztiari. Nire eskerrik beroenak eman nahi dizkiet, kuadrilakideei eta gertuko bidelagunei, emandako energia eta erakutsitako interesagatik, zertan nabilen ondo ulertu gabe itsu-itsuan nigan sinistearren, eta izan duten pazientziagatik ere. Eta nola ez etxekoei, anaia, ama eta aitari, heldu izateko bidean emandako balioengatik, eredu izateagatik, por toda la ayuda y apoyo incondicional, eskerrik asko egindako guztiagatik!

Mila esker, gracias, thank you!

Contents

| | |
|--|-----------|
| Laburpena | ix |
| Summary | xv |
| 1 Introduction | 1 |
| 1.1 Monte Carlo simulation in fisheries science | 1 |
| 1.2 Case study: Bay of Biscay sardine & ICES context | 3 |
| 1.3 State of the art | 8 |
| 1.3.1 Regression models for species distribution models | 8 |
| 1.3.2 Statistical inference on stock assessment models | 12 |
| 1.3.2.1 Automatic Differentiation | 15 |
| 1.3.2.2 Markov chain Monte Carlo | 16 |
| 1.3.3 Management strategy evaluation | 18 |
| 1.4 Objectives | 22 |
| 1.5 Structure | 22 |
| 2 Modeling species presence-absence in the ecological niche theory framework using shape-constrained GAMs | 25 |
| 2.1 Introduction | 25 |
| 2.2 Material and methods | 29 |
| 2.2.1 Regression models for presence-absence data | 29 |
| 2.2.2 Simulation | 31 |
| 2.2.2.1 Environmental gradient | 32 |
| 2.2.2.2 Species responses | 32 |
| 2.2.2.3 Sampling | 33 |
| 2.2.2.4 Presence-absence data | 33 |

| | | |
|----------|---|-----------|
| 2.2.2.5 | Model fit | 34 |
| 2.2.2.6 | Performance statistics | 35 |
| 2.3 | Results | 36 |
| 2.4 | Case studies | 39 |
| 2.4.1 | Thermal niche for sardine eggs | 40 |
| 2.4.2 | Spawning habitat of three pelagic species | 41 |
| 2.5 | Discussion | 45 |
| 3 | Stock assessment models for sardine in the Bay of Biscay | 51 |
| 3.1 | Introduction | 51 |
| 3.2 | Material and methods | 53 |
| 3.2.1 | Data for sardine in the Bay of Biscay | 53 |
| 3.2.1.1 | Catch data | 55 |
| 3.2.1.2 | Abundance indices | 56 |
| 3.2.1.3 | Biological data | 58 |
| 3.2.2 | Modelisation for sardine in the Bay of Biscay | 61 |
| 3.2.2.1 | Common elements | 61 |
| 3.2.2.2 | SS approach | 65 |
| 3.2.2.3 | a4a approach | 68 |
| 3.2.2.4 | Bayesian approach | 70 |
| 3.3 | Results | 74 |
| 3.3.1 | SS approach | 74 |
| 3.3.2 | a4a approach | 78 |
| 3.3.3 | Bayesian approach | 81 |
| 3.3.4 | Comparison | 85 |
| 3.4 | Discussion | 88 |
| 4 | Uncertainty estimation and model selection in stock assessment models with non-parametric effects on fishing mortality | 93 |
| 4.1 | Introduction | 93 |
| 4.2 | Material and methods | 95 |
| 4.2.1 | Operating model and scenarios | 96 |
| 4.2.2 | Data simulation | 99 |
| 4.2.3 | Implemented approaches | 99 |
| 4.2.3.1 | Multivariate normal approach | 99 |
| 4.2.3.2 | MCMC in ADMB | 100 |

| | | |
|----------|---|------------|
| 4.2.3.3 | Parametric Bootstrapping | 100 |
| 4.2.4 | Performance evaluation | 101 |
| 4.2.5 | AIC and BIC accuracy | 101 |
| 4.3 | Results | 102 |
| 4.4 | Discussion | 109 |
| 5 | Development of a full-feedback MSE for the Bay of Biscay sardine | 113 |
| 5.1 | Introduction | 113 |
| 5.2 | Material and methods | 115 |
| 5.2.1 | Operating model | 115 |
| 5.2.1.1 | Population dynamics | 116 |
| 5.2.1.2 | Fleet dynamics | 118 |
| 5.2.2 | Management procedure | 118 |
| 5.2.2.1 | Harvest control rules | 119 |
| 5.2.3 | Uncertainty sources | 120 |
| 5.2.3.1 | Process uncertainty | 120 |
| 5.2.3.2 | Initial population uncertainty | 121 |
| 5.2.3.3 | Observation error | 121 |
| 5.2.3.4 | Stock assessment error | 123 |
| 5.2.4 | Simulations and software | 123 |
| 5.2.4.1 | Developed code | 124 |
| 5.2.5 | Performance indicators | 125 |
| 5.3 | Results | 126 |
| 5.3.1 | Base case | 126 |
| 5.3.2 | Initial population uncertainty | 128 |
| 5.3.3 | Assessment and observation error | 130 |
| 5.3.4 | All uncertainty sources together | 133 |
| 5.3.5 | Rule selection | 135 |
| 5.4 | Discussion | 136 |
| 6 | Conclusions and further work | 141 |
| 6.1 | Contributions | 148 |
| | References | 151 |
| | Annex A | 181 |

| | |
|----------------|------------|
| Annex B | 183 |
| Annex C | 189 |
| Annex D | 195 |

Laburpena

Monte Carlo simulazioa ordenagailu bidezko esperimentera datza. Pseudo-ausazko laginketan oinarritzen da eta azpian dagoen "egia"ren ezagutza da bere indargune nagusia. Metodo estatistikoen propietateak aztertzea ahalbidetzen du, analitikoki oso konplexua edo ezinezkoa den kasuetan ere. Azken hamarkadetan ahalmen konputazionalan eman diren aurrerapenei esker, Monte Carlo simulazioak garrantzi handia hartu du jakintza arlo askotan. Horietako bat arrantza zientzia da, disziplina anitzeko jakintza arloa, arrantza sistemaren ezagutzan eta kudeaketan ikertzea helburu duena. Tesi honetan metodo estatistikoen hobekuntzak garatu dira Monte Carlo simulazioa baliatuz, arrantza zientziarekin lotutako hiru arlo ezberdinetan: espezieen banaketa ereduetan, stocken ebaluaketa ereduetan eta kudeaketa aholkularitzan. Landutako metodo guztiak Bizkaiko Golkoko sardinaren kasu ikerketara aplikatu ditugu (**1. Kapitulu**a).

Landutako lehenengo gaia nitxo ekologikoaren teoriarekin bat datozen espezieen banaketa eredu (Species Distribution Modelling, SDM) garapena izan da (**2. Kapitulu**a). Eredu hauek espezieen presentzia edo ugaritasuna ingurumen baldintzekin lotzen dute, bai eta ezaugarri espazialekin ere. Espezieen habitatak hobeto ezagutu eta etorkizuneko baldintza klimatikoetan gerta litezken banaketak auresatea dute helburu. Nitxo ekologikoaren teoriaren arabera, espezie baten banaketa ingurumen aldagai batean zehar moda bakarrekoa da. Hala ere, eskuragarri dauden metodoek ez dute hau kontuan hartzen modu esplizituan eta ondorioz, zentzu ekologikorik gabeko emaitzak lortzen dira batzuetan. Erregresio ereduetan oinarrituz, forma mugatuko eredu gehigarri orokortuak (Shape Constrained Generalized Additive Models, SC-GAMs) proposatu ditugu nitxo ekologikoaren teoriarekin bat egingo duten SDMAk garatu ahal izateko. Eredu hauek eredu gehigarri orokortuen (GAM) es-

parru berekoak dira, forma mugatzeko aukera eskaintzen dutelarik. SC-GAMen bitartez, estimatutako kurba eskala aurrealean ahurra izateko muga ezartzeak, GAMen malgutasuna mantendu eta moda bakarra izateko baldintza betetzea ahalbidetzen du, aipatutako teoriarekin bat eginez.

Monte Carlo simulazioa erabili dugu proposatutako SC-GAM ereduak ohikoak diren ondorengo ereduarekin konparatzeko: GLM ereduak, soilik kurba simetrikoak estimatu ditzaketenak, GAM ereduak, malgutasun maila altua eta baxua finkatuz, desiratutako forma lortzeko ohiko estrategian oinarrituta eta "Plateau" deituriko nitxo klimatiko ereduak, zatikako funtzio bat doitzen duena, malda gorakorra, lautada eta malda behekorra dituenak. Konparaketa egiteko, lehenik nitxo ekologikoarekin bat datozen presentzia probabilitate kurbak sortu genituen Beta funtzioa erabiliz. Parametro ezberdinak konbinatuz lau kurba ezberdin definitu genituen: simetrikoa, platikurtikoa, ezkerrean alboratua eta eskuinera alboratua. Kurba bakoitzerako 1000 behaketa sortu genituen 4 modu ezberdinetan: ingurumen aldagai osoan zehar ausaz laginduz, modaren inguruan probabilitate handiagoarekin laginduz, eta aldagairen balio baten azpitik edo gainetik lagintzeko zero probabilitatea ezarriz. Presentzia/ausentzia datuak simulatzeko `gamlss.dist` Rko paketeko puztutako beta banaketa erabili genuen, teorikoki sortutako kurben balioetan zentratuz eta 0 eta 1 balio gehiagoren agerpena ahalbidetuz. Ereduak doitzeko, egiantza handieneko prozedura erabili genuen `scam` Rko paketearen bitartez eta "component-wise boosting" prozedura `mboost`en bidez. Ereduen funtzionamendua neurtzeko, errore karratuaren batez besteko erroa eta estimen ziurgabetasuna erabili genituen bai eta aurreikusitako kurbek nitxo ekologikoaren teoriarekin zuten adostasun maila ere, horretarako bigarren deribatuak kalkulatuak.

Egindako simulazio ikerketak, SC-GAM ereduak doikuntza-egokitasuna eta nitxo ekologikoaren teoriarekiko adostasun mailaren arteko oreka ona erakutsi zuten. SC-GAMek GLM eta GAM ereduak baino hobeto funtzionatu zuten, teorikoki sortutako kurbengandik gertuago zeuden emaitzak lortuz. "Plateau" ereduak forma egokiak estimatzeko gai izan zen baina lortutako emaitzak simulazioen ziurgabetasunarekiko ezegonkorragoak izan ziren. Inplementazioei dagokienez, `scam` kodifikatzeko erraza, eraginkorra eta azkarra izan zen, `mboost` bestalde, motelagoa izan zen bariantza estimatzen, baina emaitza egonkorragoak lortu zituen.

SC-GAM ereduak Bizkaiko golko sardinaren arrautzen datu errealei aplikatu zitzaizkien.

Lortutako emaitzak, ereduaren malgutasuna eskuz finkatzea behar duten ereduekin lortutakoak bezain onak izan ziren. Honela, proposatutako SC-GAM ereduak espezieen banaketa eredugintzan oso erabilgarriak direla esan dezakegu, forma mugatzea behar duten eta behar ez dute aldagaiak eredu berean erabili daitezkeelarik.

Stocken ebaluaketa izan da landutako bigarren arloa. Stocka eremu jakin batean bizi diren espezie bereko eta ezaugarri biologiko antzekoak dituzten arrainen bilduma da, auto-ugaltzeko gaitasuna duena. Arrain espezieak aztertu eta kudeatzeko unitate biologiko garrantzitsuena da. Stocken ebaluaketaren helburua iraganeko eta gaur egungo stockaren egoera estimatzea da, bai eta lortutako ziurgabetasuna ere. Horretarako, stockaren dinamika deskribatuko duen eredu sortzen da, arrantza prozesua esplizituki kontuan hartzen duena. Stocken ebaluaketa ereduak Prozesu Ezkurtuko Ereduak (Hidden Process Models) motakoak dira eta bi denbora serieren bilakaera deskribatzen dute paraleloki: egoera prozesua, kasu honetan populazioaren dinamika dagokiona, eta behaketa prozesua, stockari buruzko datu bilketari dagokiona. Eskuragarri dauden datu guztiak integratu daitezke, egiantza bateratua eratuz. Ereduaren egoera ezezagunen gaineko inferentzia egiteko, metodo frekuentistak edo Bayesiarrak erabili daitezke.

Bizkaiko golkoko sardinaren stocka analitikoki ebaluatzen du urtero ICES (International Council for the Exploration of the Sea) erakundeak 2017 ezker, Stock Synthesis (SS) softwarean inplementatutako adinkako egitura duen ebaluaketa ereduaren bitartez. Parametroen estimazio frekuentista diferentziazio automatikoan (Automatic Differentiation, AD) oinarrituta dago eta ADMB (Automatic Differentiation Model Builder) erabiltzen du. **3. Kapitulu**an eredu honen xehetasunak azaldu ditugu eta beste bi ereduarekin konparatu dugu: a4a (assessment for all) ereduarekin, ADn oinarritutakoa baina egitura simpleagoa duena, eta ad-hoc eraikitako eredu Bayesiarra, JAGSen (Just Another Gibbs Sampler) idatzia inferentziarako Markov katearen Monte Carlo metodoak (Markov Chain Monte Carlo, MCMC) erabiliz. Hiru ereduak oinarri berbera dute populazioaren dinamika dagokionez, ordea, datuen erabilera (harrapaketa totala sartzeko edo a priori informazio erabiltzeko), banaketa estatistiko (lognormala edo multinomiala adinkako datuetarako) edo estimazio metodo (AD, MCMC edo prozedura iteratiboak) aldetik ezaugarri ezberdinak dituzte. Lortutako emaitzen arabera, eredu guztiekin lortutako estimak bat egin zuten eta beraz aztertutako hiru aukerak dira baliagarriak Bizkaiko golkoko sardinaren egoera estimatzeko. Eredu Bayesiarraren garapenak, ebaluaketa ofiziala

(SS) "kutxa beltza" ez den eredu batekin konparatzea ahalbidetu zuen, ereduaren ezaugarri eta ekuazio guztiak guk garatu genituelarik. Proposatutako beste aukera, a4a, sinplea eta azkarra izan zen aldetik, milaka iterazio egin behar diren simulazio ikerketetarako aukera egokien bezala identifikatu genuen, ebaluaketa eredu konplexuagoak ordezkatzuz.

Ziurgabetasuna ondo neurtzea garrantzi handikoa da stocken ebaluaketa prozesuetan. Geroz eta interes handiagoa hartzen ari den gaia da, izan ere, hainbat espeziaren gainbehera ekarri du kudeatzeko orduan ziurgabetasuna behar bezala kontuan ez hartzeak. Tesi honetan Monte Carlo simulazioa erabili dugu ziurgabetasun hori estimatzeko metodo ezberdinak konparatzeko (**4. Kapitula**). Gainera, ebaluaketa ereduaren arrantza hilkortasun tasa estimatzeko, funtzio ez parametrikoko gehitzearen efektua ere aztertu dugu. Simulazio ikerketetarako a4a ereduaren erabili dugu, aurreko atalean ondorioztatutakoa jarraituz. Ziurgabetasuna estimatzeko hiru hurbilketa konparatu ditugu: aldagai anitzeko banaketa normala, *bootstrap* parametrikoa eta ADMBen implementaturiko MCMC algoritmo bat. Gehienetan ebaluaketa ereduak funtzio parametrikoko erabiltzen dituzte, kasu askotan datuen konplexutasuna azaltzeko gai ez direnak. Honela, funtzio ez parametrikoen erabilera bultzatu izan da ereduaren zehaztasuna hobetze aldera, ereduaren aukeraketa prozesua zaildu dezaken arren. Ziurgabetasuna estimatzeko metodoak konparatzeko, ereduaren konplexutasunaren arabera bost kasu ezberdin aztertu genituen, arrantza hilkortasun (F) parametroaren leuntze funtzioaren formaren arabera: egitura banakor klasikoetatik, *tensor product spline*etan oinarritutako egitura ez banakorretara. Azken hauek, *mgcv* Rko paketearen eskuragarri dauden leuntze funtzioen bitartez implementatu dira, askatasun gradu ezberdinak erabiliz.

Monte Carlo simulaziorako egiazko balioak sortzeko, definitutako kasu bakoitzetarako Bizkaiko golkoko sardinaren datu errealekin doitu genituen ereduak eta harrapaketen eta indizeen datuak sortu genituen, egiazko balioei errore lognormala biderkatuz. 100 datu multzo sortu genituen kasu bakoitzetarako, eta 1000 iterazio gorde genituen ziurgabetasun estimazio metodo bakoitzetik. Metodo bakoitzaren funtzionamendua konparatzeko, besteak beste, estimatutako biomasa ugaltzailea edo harrapagarritasuna bezalako parametroen estima puntuala, bariantza, aldakuntza koefizientea, estaldura probabilitatea edo alborapen erlatiboa kalkulatu genituen. Literaturan aurkitutakoarekin bat eginez, gure emaitzen arabera MCMC izan zen metodo fidagarriena. Arrantza hilkortasun parametrarako konplexutasun maila ezberdineko

ereduen ziurgabetasun estimak alderatzerakoan, funtzio parametrikoak ez bezala, funtzio ez parametrikoak forma malguak estimatzeko gai izan zirela ikusi genuen. Hala ere, kasu konplexuenetan parametro kopuruak nabarmen egin zuen gora eta ziurgabetasuna handitu egin zen, aldaketa koefiziente eta alborapen handiagoak erakutsiz azken urte eta adinen estimetan, eta estaldura probabilitate okerragoak emanez, MCMC metodoa salbuespen. Eredu aukeraketarako AIC eta BIC irizpideak ere testatu genituen, BICak ere egokia aukeratzeko zehaztasun handiagoa duela ondorioztatuz.

Tesi honetan landutako azken arloa kudeaketa aholkua izan da. Stocken kudeaketarako aholku zientifikoa emateko modu tradizionala ebaluaketa eredu on bat garatu eta emaitza horietatik epe motzerako kudeaketa ekintzak definitzean datza. Prozesu hau urtero egiten da eta ondorioz, urtetik urtera harrapaketa aholku oso aldakorrek suertatu daitezke. Azken hamarkadetan, modu tradizionalaren ordeztuz, kudeaketa estrategien ebaluaketa (Management Strategy Evaluation, MSE) deituriko prozedura indarra hartzen ari da, Monte Carlo simulazio bidez kudeaketa ekintza ezberdinak ebaluatzen dituen, ekintza hauek sistema errealean inplementatu aurretik. Prozedura honetan etorkizunera proiektatzen da stockaren sistema osoa: populazio eta arantzaren dinamikak, behaketa prozesua, stockaren ebaluaketa, kudeaketa erabakiak eta euren inplementazioa. MSEaren emaitzak kudeaketa estrategiak konparatzeko erabiltzen dira, aurrez ezarritako helburuak bete, jasagarritasuna adibidez, eta ziurgabetasun iturri ezberdinen aurrean egonkortasuna erakusten dutenak aukeratuz.

Lan honetan feedback osoko MSEa garatu dugu Bizkaiko golkoko sardinaren stockerako (**5. Kapitulua**). MSE prozesuetan kontuan hartu beharreko ziurgabetasun iturriak zeintzuk diren zehaztea oraindik ere erronka bat den arren, literaturaren arabera beharrezkoak diren hiru iturri sartu genituen gure lanean: prozesu, parametro eta behaketa ziurgabetasunak hain zuzen ere. Gainera, SS ebaluaketa eredu ofiziala Monte Carlo simulazioaren barruan sartu genuen. Hasierako populazioa definitzen duten parametroen ziurgabetasuna, ebaluaketa ereduko MCMC iterazioen bitartez sartu genuen, aurreko atalean ziurgabetasuna estimatzeko metodo eraginkorra dela ikusi baitgenuen. Horretarako, SS ereduak duen MCMC egiteko aukera baliatu genuen (ADMBen oinarritutakoa) eta 1000 iterazioa gorde genituen, MSEa hasierako 1000 populazio ezberdinetan oinarrituz. Behaketa errorea banaketa multinomial eta lognormalen bitartez sartu genuen, SS ereduarekin konfiantza mantenduz, eta prozesu errorea urteroko erreklutamenduei errore lognor-

mala biderkatuz. Ziurgabetasun iturri hauek txandaka sartu genituen euren efektua banaka aztertu ahal izateko, bai eta iturri ezberdinen konbinaketena ere. 1000 simulazio egin genituen aukera bakoitzerako. Kudeaketa estrategien konparaketarako adierazle ezberdinak kalkulatu genituen: biomasa ugaltzailea muga batetik behera egoteko arrisku maila, arrantza debekatzera iristeko probabilitatea, urtetik urterako harrapaketen aldaketa edo biomasa ugaltzailearen mediana besteak beste. FLBEIA paketea erabili genuen simulazioak egiteko, modulu bidez egituratua dago eta horri esker gure kasu ikerketarako funtzio bereziak sortu ahal izan genituen: errore multinomial ez biderkakorra adinkako datuekin erabili ahal izateko funtzioa eta SS ebaluaketa eredu simulazio prozesuaren barruan doitu ahal izateko funtzioa.

Ziurgabetasun iturriak MSEan sartzeak biomasa ugaltzailea muga batetik behera egoteko arriskua handitzea eta batez besteko harrapaketen jaitziera eragiten duela ikusi genuen, geroz eta iturri gehiago sartu, orduan eta nabarmenagoak ziren eragin hauek. Ebaluaketa eredu Monte Carlo simulazioaren barne doitzearen ondorioz, biomasaren goi eta azpi estimazioak ematen zirela ikusi ahal izan genuen, biomasa erreala baino estima baxuagoak ematen zituen ebaluaketa ereduak lehen urteetan, eta errealak baina altuagoak gainontzeko urteetan. Honek biomasa ugaltzailea muga batetik behera egoteko arriskua handitzen zuen bai eta proiektzio urteetan biomasa horren bariabilitatea ere. Harrapaketak debekatzera iristeko probabilitatearen jaitziera ere ikusi genuen, SS ereduak estimatutako biomasen bariabilitatea egiazkoa baina txikiagoa izatearen ondorio. Behaketa errore multinomial eta lognormalak gehitzean, biomasa ugaltzailearen bariabilitateak eta arrantza debekatzeko probabilitateak gora egin zuten, bai eta deskribatutako alborapen efektuak ere. Bizkaiko golkoko sardinarentzat lan honetan testatutako bi ustiapen arauak, F an oinarritutako ICES arauak eta harrapaketen oinarritutako arauak, ICESek definitutako arrisku maila ($Risk3 < 0.05$) betetzen zuten prozesu errorea soilik sartutakoan. Gainontzeko ziurgabetasun iturriak sartzean ordea, ez zen baldintza hori betetzen. Honela, ziurgabetasun horiek kontuan hartzean, arrisku maila hori bete ahal izateko harrapaketa totalak baxuagoak ziren. Ustiapen arau onenaren aukeraketa lan honen helburuetatik kanpo zegoen arren, arau ezberdinak konparatu genituen aipatutako ziurgabetasun iturri guztiak kontuan hartuz, aurrez ezarritako helburuen araberrako arau onenaren aukeraketa nola egin daiteken erakusteko.

Oro har, Monte Carlo simulazioa baliagarria izan da tesi honen helburuak bete ahal izateko, datu erreal eta emaitza analitikoak soilik baliatuz lortu ezingo genituenak.

Summary

Monte Carlo simulation consists in computer experiments that involve creating data by pseudo-random sampling with the objective of obtaining empirical results for complex problems. It has shown to be a powerful tool to study the performance of statistical methods. The strength of Monte Carlo simulation relies on the underlying known "truth" which allows measuring and assessing the properties of the statistical method under study, that are often difficult to be derived analytically. Due to the advances in computational power during the last decades, Monte Carlo simulation has become relevant in many scientific disciplines. One of them is fisheries science, a multidisciplinary area that aims at understanding and managing fisheries. In this thesis Monte Carlo simulation was used to improve statistical methodology related to three different fields of fisheries science: species distribution models, stock assessment models and management advice. All the methods were applied to the case study of sardine in the Bay of Biscay (**Chapter 1**).

The first field of fisheries sciences addressed in this thesis was the development of species distribution models (SDM) that meet the ecological niche theory framework (**Chapter 2**). SDMs relate species occurrences or abundances with environmental conditions and spatial characteristics and can be used to have a better understanding of species habitats or to predict their distribution under future climate scenarios. According to the ecological niche theory, species distributions are expected to show unimodal relationships with respect to environmental gradients. However, most of the available methods do not account for this expected shape explicitly, and result in not meaningful or ecologically interpretable results. Focusing on regression-based models, we propose using shape-constrained generalised additive models (SC-GAMs) to build SDMs under the ecological niche theory. These models are based on the same

statistical framework as generalised linear models (GLMs) or generalised additive models (GAMs) but include the possibility to add shape-constraints. In particular, imposing concavity constraints in the predictor scale in a logistic GAM leads to flexible response curves while retaining the unimodality required by the niche theory.

We used Monte Carlo simulation to compare the proposed SC-GAM approach with respect to commonly used GLMs, which can only estimate symmetric response curves, GAMs, fixing a high and a low number of knots to limit the flexibility of the response curve as a commonly used strategy to obtain the desired shape, and a climate envelope model ("Plateau") that fits a piecewise function consisting of an increasing slope, followed by a plateau and a final decreasing slope. For that, first, we generated theoretical response curves that meet the ecological niche theory using the Beta function. Combining different values for the shape parameters, we defined four different curves: symmetric, platycurtic, left skewed and right skewed. For each of the curves, we generated samples of 1000 observations according to four sampling options: sampling randomly along the whole range, sampling with a higher probability around the mode, and having a sampling probability of zero above or below a specific value of the gradient. Presence/absence data were simulated following a beta inflated distribution centred at the generated theoretical probabilities, allowing for zero and one inflation, as implemented in the `gamlss.dist` R package. SC-GAMs fitting was performed using a maximum likelihood approach with the `scam` R package and a component-wise boosting approach with `mboost`. The performance of the methods was evaluated in terms of the Root Mean Square Error, the uncertainty around estimates and the agreement with the niche theory through the computation of the second derivatives of the predicted curves.

The simulation study showed that the proposed SC-GAM approach resulted in a good balance between the goodness-of-fit and the agreement with the niche theory. SC-GAMs performed better than GLM and GAMs obtaining fitted curves that were closer to the generated theoretical curves. The "Plateau" method was able to provide the correct shapes but results were less robust to the uncertainty of the replicates. The `scam` implementation was easy to code, efficient and fast while the `mboost` approach was more robust but slower for variance estimation.

The SC-GAM approach was also applied to real egg abundance data of sardine in the Bay in Biscay. The results obtained were as good as the outputs from models

that needed manual tuning of parameters for controlling the desired shapes. Therefore, it was concluded that proposed SC-GAMs can be readily applied for modelling both constrained and unconstrained responses in a unified SDM framework.

The second field of fisheries science tackled in this thesis was stock assessment. A stock is defined as a collection of fish of the same species distributed in a given area that is large enough for self-reproducing and shows similar biological characteristics. It forms the most important biological unit for regulation and monitoring of a fish species. The objective of stock assessment is to estimate the past and current status of a stock along with the corresponding uncertainty. This involves constructing a stock assessment model that describe the dynamics of the stock along time incorporating explicitly the fishing process. Stock assessment models belong to the general class of statistical models called Hidden Process Models. These models describe the evolution of two time series in parallel: the state process, that in this case corresponds to the stock population dynamics, and the observation process, that is related to the data collected on the stock. All available data can be integrated constructing a joint likelihood. Estimation methods to conduct inference on the unknown states and model parameters can be based either on frequentist or Bayesian approaches.

Since 2017, sardine in the Bay of Biscay is analytically assessed yearly by ICES (International Council for the Exploration of the Sea) using an age-structured stock assessment model implemented in Stock Synthesis (SS) software. Parameter estimation is based on automatic differentiation (AD) using Automatic Differentiation Model Builder (ADMB) within a frequentist framework. In **Chapter 3** we detail the basis of this official assessment model and compare it with two alternative approaches: the a4a (assessment for all) model, that is also based on AD but intends to be simpler in terms of implementation, and an ad-hoc developed Bayesian model coded in JAGS (Just Another Gibbs Sampler) that uses Markov Chain Monte Carlo (MCMC) methods for inference. The three models were based on the same population dynamics, but differ on some underlying assumptions (the use of the total catch data or the incorporation of prior information), on statistical distributions for the observations process (multinomial or lognormal distributions for age-structured data), and on estimation methods (the use of AD or MCMC methods, or the need for iterative tuning processes for some variance parameters' estimation in SS). We concluded that obtained estimates were all consistent and that the three approaches presented could be used to estimate Bay of Biscay sardine's stock status similarly.

The development of the ad-hoc Bayesian model tailored for this specific case study provided the opportunity to compare the official assessment to a non "black-box" approach where all the assumptions and equations were self-developed. Alternatively, a4a was identified as a fast and simple approach that could be appropriate to replace the official assessment in simulation studies where thousands of iterations must be run.

The correct quantification of uncertainty in stock assessment is a crucial issue, with an increasing interest in incorporating it into management processes due to many management failures related to the lack of consideration of different uncertainty sources. In this thesis we used Monte Carlo simulation to compare different methodologies for uncertainty estimation in stock assessment models with non-parametric effects on fishing mortality (**Chapter 4**). Using the a4a approach, that was identified as the most suitable option for simulation, three different approaches for uncertainty estimation were tested: the multivariate normal distribution, parametric bootstrapping, and an MCMC algorithm implemented in ADMB. Stock assessment models have typically relied on parametric functions that may lack enough flexibility to capture all the data features and the inclusion of non-parametric smooth functions have been advocated to avoid model miss-specification, although model selection could be more difficult. In order to study the performance of the uncertainty estimation methods depending on the model complexity, five different scenarios were defined depending on the shape of the smooth function of the fishing mortality (F). They range from a classical age-year separable structure to a non-separable tensor product based structure, and they were implemented using smoothing functions (splines and tensor products) available in `mgcv` R package with different degrees of freedom.

For the Monte Carlo simulation, true values were generated by fitting the real data from the Bay of Biscay sardine case study to the defined model in each scenario and simulated catch and indices-at-age data were generated by multiplying lognormal errors to the true values. For each scenario 100 datasets were simulated, and 1000 iterations were saved for each of the uncertainty estimation methods. The methods' performance was evaluated for several quantities of interest, such as spawning stock biomass (SSB) or catchabilities, in terms of point estimates, variance, coefficients of variation, coverage probabilities, relative bias or skewness. In agreement with previous studies in the literature, we concluded that MCMC was the most reliable method. When uncertainty estimation was compared across models of different lev-

els of complexity for fishing mortality, we found that non-parametric functions were able to estimate flexible shapes that parametric models could not capture. However, the most flexible configurations increased the number of parameters and resulted in greater uncertainty, with greater CVs and relative bias at terminal years and ages and also in a deterioration of coverage probabilities, except for the MCMC method. AIC and BIC criteria for model selection were also tested in terms of accuracy in selecting the correct model, obtaining better results for BIC.

The third and final field of fisheries science addressed in this thesis was fisheries management advice. The traditional approach for providing scientific advice for management consists in developing the best possible assessment and subsequently translating these into short-term management actions. This is repeated on an annual basis and can result in highly variable recommendations from year to year. In the last decades an alternative approach known as management strategy evaluation (MSE) has emerged. It entails evaluating by means of Monte Carlo simulation different management actions before they are put in place. This approach involves the projection into future years of the entire management system: the real dynamics of the population and fishery, the observation process, the stock assessment, the management decision and its implementation. The results are then used to compare and select management strategies that outperform with respect to pre-specified objectives and that are robust to different sources of uncertainty.

In this work we developed a full-feedback MSE for sardine in the Bay of Biscay (**Chapter 5**). Although the decision on what sources of uncertainty need to be included in an MSE remains a major challenge, we considered the least three sources of uncertainty that should be included according to the literature, namely, process, parameter and observation uncertainties. The official SS assessment model was also incorporated into the Monte Carlo simulation. Uncertainty on parameters that define the initial population was introduced via MCMC iterations, identified as the most robust method for uncertainty estimation in stock assessment models, using the MCMC option in SS (inherited from ADMB) and saving 1000 output iterations as 1000 different initial populations for the MSE conditioning. Observation error was included through multinomial and lognormal distributions in order to be consistent with the SS assessment model assumptions and process error was represented by lognormal errors in the annual recruitments. Different sources of uncertainty were added gradually to provide an analysis of the impact of each of them. 1000

simulation were run for each of the MSE scenarios including different uncertainty sources. The performance indicators computed for comparison were the risk of SSB being below a limit level, the probability of closure of the fishery, the interannual catch variability and the median SSB. FLBEIA package was used to perform these simulations, which has a modular implementation facilitating the development of new functions by the users. We developed several functions for our case study, such as a function to include the SS assessment model within the MSE loop or a function to generate the non-multiplicative multinomial error for age-structured data observation.

The gradual inclusion of uncertainties resulted in a successive increase of the probability of SSB being below the limit biomass reference point B_{lim} and a decrease of the average catches. The fitting of the stock assessment within the MSE simulation led to a bias in the estimated SSB by the SS model, negative in the first years and positive afterwards, resulting in an increase of the risk to be below B_{lim} and of the variability of the SSB in the projection years. A decrease on the probability of closure was observed, related to the fact that the SSB variability estimated by SS was lower than the real one. When multinomial and lognormally distributed observation errors were included, variability in SSB, risks and probability of closure increased, as well as the described bias effect. The two harvest control rules implemented for the Bay of Biscay sardine in this work, the F-based ICES rule and an alternative catch-based rule, were considered precautionary in the base case MSE (only process uncertainty included), but not when incorporating all the proposed sources of uncertainty. So, we concluded that when all the relevant uncertainty sources were included the ICES precautionary criterion ($Risk3 < 0.05$) was met with a lower level of total catches. Although the selection of the best harvest control rule for Bay of Biscay sardine was out of the scope of this work, we compared several additional harvest control rules, to illustrate how the best rule could be selected by managers according to predefined objectives by accounting for the most important uncertainty sources for this stock.

In general terms, Monte Carlo simulation has been an adequate tool to accomplish the objectives of this thesis that definitely could not have been achieved using only available real data or analytical solutions.

Chapter 1

Introduction

1.1 Monte Carlo simulation in fisheries science

Exact analytical properties of statistical methods are often difficult to derive, specially when wrong assumptions are made or several sources of uncertainty are incorporated. In these cases, an alternative approach is Monte Carlo simulation. It consists in computer experiments that involve creating data by pseudo-random sampling with the objective of obtaining empirical results of methods' performance. The strength of these simulation studies relies on the underlying known "truth" which allows measuring and assessing the properties of the evaluated methods.

First modern Monte Carlo experiments were found at the end of the nineteenth century, being Laplace's approximation of the irrational number π the first documented application of the Monte Carlo simulation (Johansen 2010). However, due to the cost of carrying these physical experiments, it was not until the 1940s, when the random number generators were introduced by the physics community, that Monte Carlo simulations became more practical. During the 1980s, the increased popularity of Bayesian methods and the need to evaluate complex high-dimensional integrals led to a revolution in Monte Carlo simulation methodology by the statistics community (Gelfand and Smith 1990).

Monte Carlo simulation is an intuitive and easy to understand tool that can determinate unknown solutions of complex problems very efficiently in contrast to deterministic approaches (Atanassov and Dimov 2008). Moreover, Monte Carlo simulation algorithms can be easily parallelized resulting in a linear speed-up (Rosenthal 2000).

As a drawback, obtained results are not exact and are subject to the number of iterations performed. Computational limits are the main disadvantages for the use of Monte Carlo simulation. However, due to the vast advances in computational power during the last decades (Nordhaus 2007), Monte Carlo simulation have become more attractive and more frequently used in many areas, such as medicine (Concato and Feinstein 1997), finance (Brandimarte 2014), chemistry (Jorgensen and Duffy 2000) or social science (Carsey and Harden 2013).

Monte Carlo simulation has been also extensively used in fisheries science (Garcia et al. 2013, Ianelli 2002, Oksanen and Minchin 2002, Punt et al. 2002, Yin and Sampson 2004). This discipline aims at understanding and managing fisheries (Hart et al. 2002). Subsequently, it is a multidisciplinary area, which encompasses disciplines such as oceanography, marine biology, meteorology, conservation, ecology, population dynamics, statistics, economics, management, and many others. It is a relatively modern science moving nowadays towards holistic approaches to fisheries assessment and management, which take into account ecosystem aspects in order to enhance sustainability.

This thesis uses Monte Carlo simulation to improve the statistical methodology of three different fields of fisheries science:

1. Habitat models, through the development of species distribution models that allow understanding how species interact with their environment.
2. Stock-assessment models, which making use of available data of an exploited species population are able to estimate its current and past status along with the corresponding uncertainties.
3. Long-term management advice, where based on the status of a population and related sources of uncertainties different management alternatives are studied in order to find the best strategy in the long-term.

Methodological improvements in these three research topics can provide managers better and valuable information to oversee marine species in space and time. The state of the art of each of these three topics is reviewed this chapter.

As pointed out by Harrison (2010), there is no single Monte Carlo simulation method, but most of them, as well as the work in this thesis, follow a similar structure. Let

θ denote the vector of the quantity (or quantities) of interest. The general steps of Monte Carlo simulation are::

1. Choose or create the deterministic model for the input, which will represent the reality.
2. Add stochasticity to input values, generating pseudorandom samples (iterations) from selected statistical distributions.
3. Apply the method of interest to the generated pseudo-random inputs and obtain the desired numerical output, $\hat{\theta}_i$.
4. Repeat steps 2 and 3 n times.
5. Perform statistical analysis on collected n output sets $(\hat{\theta}_1, \dots, \hat{\theta}_n)$.

The last step, statistical analysis, involves the computation of performance statistics that compare the obtained outputs (hat $\theta_1, \dots, \hat{\theta}_n$) with respect to the underlying "truth" (θ) to evaluate the statistical method being tested. Common performance statistics measure the bias, the precision and the accuracy either in absolute or in relative terms (Walther and Moore 2005). The most common performance statistics include among others:

- Relative bias, defined as $\frac{\sum(\hat{\theta}_i - \theta)/\theta}{n}$.
- Coverage probability, is defined as $\Pr(\theta \in PI_{\hat{\theta}})$ the proportion (out of the n simulations) of 90% confidence or probability intervals (PI) that contain the true value.
- Root Mean Square Error, $RMSE = \sqrt{\frac{\sum(\theta - \hat{\theta}_i)^2}{n}}$.

1.2 Case study: Bay of Biscay sardine & ICES context

European sardine (*Sardina pilchardus*) is a small pelagic fish that has a lifespan of up to 10 years. Maturity is achieved in their first year of life and spawning occurs throughout the year, with two peaks in spring and in autumn/winter. This species diet consists of zooplankton and phytoplankton and it is prey of a range of fish and marine mammal species, such as, common dolphins or harbour porpoises. European sardine measures from 13 to 25cm but a decreasing trend in body length has been

detected since the 2000s (Véron et al. 2020).

The species has a wide distribution in the Northeast Atlantic extending from the Celtic Sea and North Sea in the north to Mauritania in the south. Populations of Madeira, the Azores and the Canary Islands are at the western limit of the distribution (Parrish et al. 1989). Sardine is also found in the Mediterranean and the Black Seas.

Assessment and management advice for sardine in the Northeast Atlantic have been developed under the ICES (International Council for the Exploration of the Sea, <http://www.ices.dk>) auspices. ICES is an intergovernmental marine science organization whose main goal is to advance and share scientific understanding of marine ecosystems and the services they provide and to use this knowledge to generate state-of-the-art advice for meeting conservation, management, and sustainability goals. Based on scientific work and data, ICES provides scientific advice for a wide range of recipients, including ICES member countries and various international organizations and commissions, such as the European Commission (EC).

The ICES advisory process starts when a request for advice is received from a client. Then, expert groups collect the data, carry out the corresponding analysis and draft a first scientific/technical response to the request. Expert group reports are peer-reviewed by independent experts and then reviewed by an advice drafting group. The draft advice is discussed and finally approved by the Advisory Committee (ACOM) and it is delivered to the client. An important part of ICES advice regards the management of the exploitation of living marine resources. ICES classifies six main categories; from "stocks with quantitative assessments" (category 1), to "negligible landings stocks and stocks caught in minor amounts as bycatch" (category 6). In between there are stocks with qualitative assessments or stock assessments that only indicate trends. ICES advice is aligned with the reformed Common Fisheries Policy (CFP) that includes the fundamental objective to progressively restore and maintain fish stocks above biomass levels capable of producing the maximum sustainable yield (Salomon et al. 2014). In addition, all ICES advice is consistent with the precautionary approach, according to which all the populations need to be maintained within safe biological limits (UN 1995).

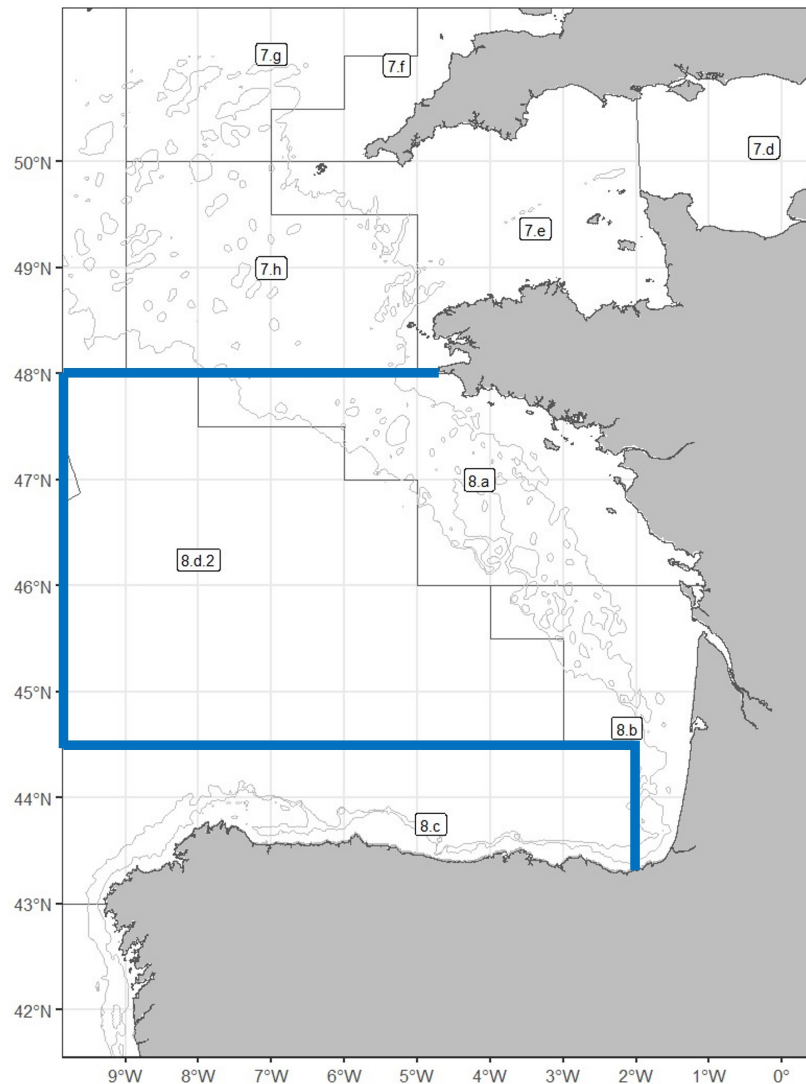


Figure 1.1: ICES areas and divisions in the Bay of Biscay. The stock of sardine in the Bay of Biscay is distributed in divisions 8 a,b,d.

The stock is the most important biological entity for regulating and monitoring a fish species. A unit stock is an arbitrary collection of populations of fish of the same species that is large enough to be essentially self-reproducing with members of the collection showing similar patterns of growth, migration, and dispersal located in a given area (Hilborn and Walters 1992). Sardine in the Northeast Atlantic is separated in three stocks or management units: Iberian sardine that is distributed along

ICES divisions 8c and 9a, sardine in the Bay of Biscay (BoB) that is distributed along ICES divisions 8a,b,d and sardine in southern Celtic Sea and English Channel that is distributed in ICES subarea 7 (Figure 1.1). This thesis is focused on sardine in the Bay of Biscay. Until 2017, this stock was considered together with the stock from subarea 7. Although there was no strong evidence in any direction both regions had individuals of all the life stages, biological traits seemed to be different and it was decided to separate them into two stocks and work on the Bay of Biscay sardine stock assessment model due to poor data availability in subarea 7. The connectivity of this species from the northern Bay of Biscay to the eastern Gulf of Cadiz has been also analysed (Silva et al. 2019), concluding that the flow rates between three large areas were low. These three main areas, Bay of Biscay, the northern Spanish and Portuguese waters and the Gulf of Cadiz, with their own recruitment hotspots, are considered to be host self-sustained weakly connected populations forming a metapopulation.

Concerning human activities, sardine in the Bay of Biscay is exploited mainly by France and Spain. For French fleet, both purse seiners and pelagic trawlers target sardine in French waters. Purse seiners operate mainly in coastal areas while trawlers are not allowed to fish within 3 nautical miles from the coast. Both pair trawlers and purse seiners operate close to their base harbour when targeting sardine. French catches increased from 1983 with around 4 thousand tonnes to 2008 with 21 thousand tonnes, reaching the series maximum in 2018 with more than 25 thousand tonnes. The highest catches are taken in the summer months and almost all are taken in southwest Brittany. Spanish catches are taken by purse seines from the Basque Country operating mainly in division 8b. Spanish landings peaked in 1998 and 1999 with almost 8 thousand tonnes but decreased to below 1 thousand tonnes in 2010. Spanish catches increased again, reaching 16 thousand tonnes in 2014. The Spanish fishery takes place mainly between October and March. The strong increase of Spanish catches in 8b is mainly due to an increase of the fishing effort, taking into account the low level of the Iberian sardine stock in 8c and some monthly closures of the southern fishery.

The population present in the Bay of Biscay is monitored by two surveys carried out in spring. On the one hand, the PELGAS acoustic surveys are conducted yearly and provide sardine estimates in terms of total biomass and number of individuals by age since 2000 (Doray et al. 2018). On the other hand, since 1999 BIOMAN

surveys provide yearly a sardine egg abundance index (Santos et al. 2018) with the Daily Egg Production Method (DEPM, Lasker 1985)), and since 2011, every three years, sardine spawning stock biomass is estimated since 2011 in the SAREVA or BIOMAN surveys (ICES 2020b).

The stock assessment and management of sardine in the Bay of Biscay has evolved along the last decades (Figure 1.2). ICES provided advice for sardine in the combined 8.a,b,d and 7 regions for the first time in 2013. In previous years, exploratory assessments were carried out during the working group on horse mackerel, anchovy and sardine (WGHANSA) and cohort tracking analyses were conducted from 2013 to 2016 to derive total mortality (Z). The advice was based on available survey indices following the guidelines for ICES category 3.

In 2017, an ICES benchmark workshop for evaluating and improving the data and assessment methodology was carried out (ICES 2017a in contrast with the previous advice one in 2013 (ICES 2013)), it was considered that various time series of fishery data in the Bay of Biscay monitored from commercial vessels and surveys were long enough to provide an analytical assessment using a catch-at-age statistical stock assessment model. The model selected for the assessment by WGHANSA was Stock Synthesis (SS, Methot and Wetzel 2013). However, obtained absolute estimates were approved only as relative indicators of stock status. Thus, the sardine stock was classified as category 2 stock.

In 2019 an Inter-benchmark process was carried out for this stock with the objective of evaluating the stock assessment focusing on retrospective bias, data revisions and reference points' update (ICES 2019b). The SS model was revised and changes concerning selectivity estimation and the stock-recruitment relationship were applied. Although some issues on the catchability coefficients still remained, the stock was upgraded from ICES category 2 to category 1 due to the improvement in terms of retrospective bias and other diagnostics.

Given that there is not agreed management plan for this stock, ICES advice is based on the maximum sustainable yield (MSY) approach. No total allowable catch is set by the European Commission for this stock and the fisheries appear to be regulated by market price (ICES 2019e); some fisheries have their own local management, limiting total catch or fish sizes to sustain prices. There is ongoing discussion

about a possible management plan through the South West Waters Advisory Council (<https://www.cc-sud.eu/index.php/en/>) but there is not an agreement yet.

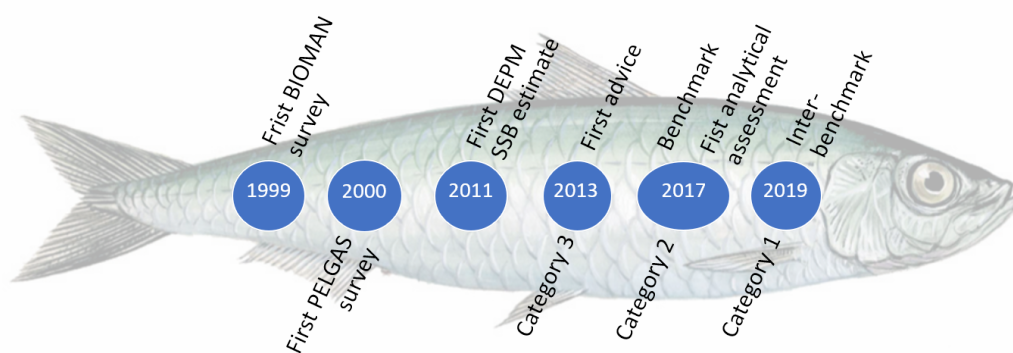


Figure 1.2: Main milestones in the assessment and management advice of sardine in the Bay of Biscay.

1.3 State of the art

In this section, we review the most relevant statistical methodology of the three fisheries science fields that will be dealt with in this thesis.

1.3.1 Regression models for species distribution models

In recent years there has been an increased interest in understanding the impact of environmental changes on species distribution. During the last decades many studies have been published on this topic (Beaugrand et al. 2008, Bruge et al. 2016, Chust et al. 2014, Erauskin-Extramiana et al. 2019, Reygondeau and Beaugrand 2011, Villarino et al. 2015). Since 1950 the atmosphere and ocean have warmed, and climate change is considered an unequivocal reality (IPCC 2013). This reality has an impact across all marine ecosystems, latitudes and trophic levels and may lead to some species local extinctions and invasions (Cheung et al. 2013). Climate change is associated to multiple ecosystem changes such as shifts in temperature, circulation, stratification, nutrient input, oxygen content or ocean acidification among others (Doney et al. 2012). As a result, fish stocks can be impacted, affecting their productivity, habitat suitability or population-dynamics. Many efforts are being made to see how fish species abundance have been affected by climate change during the

last decades or to predict how distributions may change (Koenigstein et al. 2016, Orio et al. 2019, Rijnsdorp et al. 2009). The fishery activity could be also affected. For instance, according to Cheung et al. (2009), an increase in the catch potential in high latitudes and a decrease in tropical and subtropical regions is expected.

Species distribution models (SDM) are important statistical tools in ecology that relate species occurrence or abundance with information on environmental conditions and spatial characteristics of locations where the species are found (Elith and Leathwick 2009). These models can be used to predict or to have a better understanding of the species distribution (Halvorsen 2012, Petitpierre et al. 2017). They are widely used in several fields, such as ecology, evolutionary biology and conservation (Guisan et al. 2013, Peterson et al. 2011, Zimmermann et al. 2009).

Data for SDM is ideally composed by records of species presence and absence locations together with abundance records for the locations where the species are found. Sometimes only opportunistic or biased spatial data are available, containing presence data without information on absences or abundances. In such cases it is a common practice to generate pseudo-absence data (Iturbide et al. 2015). When both, absence data and some measure of abundance, are available for modelling, a two stage approach is recommended (Austin and Cunningham 1981, Borchers et al. 1997).

The focus of the most highly cited papers on SDM is the development of novel methods to better predict environmental suitability for species and the improvement of model performance (Barbosa and Schneck 2015). Still, most of the available statistical methods used to model species distributions result in responses along environmental gradients that do not fulfil the conditions of ecological niche theory, and consequently are not always ecologically meaningful or interpretable.

In this thesis we will focus on the commonly used regression-based approaches for proposing improvements on SDM. Generalised linear models (GLM) are well-established statistical models in SDM (Guisan et al. 2002). GLMs are parametric models which assume that the relationships or pattern of interest can be expressed in a specific formula involving a fixed number of unknown parameters. However, choosing a specific functional form to model the relationship between some variables of interest is often difficult, if not impossible. Non-parametric smoothing methods

are able to let the data suggest the appropriate form based on some infinite dimensional collection of functions. The degree of smoothness of the resulting form can usually be either controlled or estimated. Non-parametric models, where the model structure is not fixed beforehand are well-established in regression models like generalised additive models (GAMs) (Härdle 1994).

GAMs (Hastie and Tibshirani 1990, Wood 2006a) are a generalisation of GLMs, where the linear predictor, $g(x)$, is a smooth function of the explanatory variable. For a single explanatory variable, we have the form:

$$g(x) = \beta_0 + f(x), \quad (1.1)$$

where $f(x)$ is a smooth function. There are several ways to represent $f(x)$, from kernel smoothing or local linear methods to splines-based regression methods. We describe the latter approaches, where $f(x)$ is given by a sum of some basis functions. Hence for a single covariate x , we have:

$$f(x) = \sum_{k=1}^K \theta_k B_k(x), \quad (1.2)$$

where θ_k are the regression coefficients and $B_k(x)$ a basis function of x . There are several choices for the basis functions (e.g. polynomials of a certain order, natural splines, cubic splines or B-splines). Splines are flexible tools for smoothing in general. A spline of degree d is a function formed by connecting polynomial segments of degree d so that the function is continuous, the function has $d - 1$ continuous derivatives, and the d th derivative is constant between knots. B-splines (de Boor 1972) are a popular choice given that they are easy to compute and they have good numerical properties. In regression splines, estimated regression coefficients, $\hat{\theta}_j$, are obtained by least squares (i.e. by minimizing the residual sum of squares) and hence the shape of a spline can be controlled by carefully choosing the number of knots and their exact locations in order to allow flexibility (e.g. fix the locations of k knots at quantiles of x), and avoid overfitting where the trend changes little. However, in many situations, choosing the number of knots and their locations is a very difficult problem to solve.

Alternatively, smoothing splines find the solution of f which minimizes:

$$\sum_{i=1}^n (y_i - f(x_i))^2 + \lambda \int f''(x)^2 dx,$$

where the minimizer $f(x)$ is a natural cubic spline, with knots at each sample point x_1, \dots, x_n , and $\lambda \int f''(x)^2 dx$ is the roughness penalty. The parameter λ , controls the amount of smoothness and takes values $0 < \lambda < \infty$, for $\lambda \rightarrow \infty$; large values of λ result in strong penalisation (a straight line in the limit) and for values λ close to 0 the resulting fit is a wiggly function. The selection of λ can be performed by (generalised) cross-validation or information criteria such as Akaike or Bayesian information criteria (Akaike 1974, Schwarz 1978). However, the main drawback of smoothing splines is the dimensionality for large n (Green and Silverman 1993).

In contrast to smoothing splines, low-rank approximations have been proposed in the literature (see Ruppert et al. 2003 for a complete overview), which are called penalized regression splines. For instance, thin plate regression splines (Wood 2003) are constructed by a simple transformation and truncation of the basis that arises from the solution of the thin plate spline smoothing problem. P-splines (Eilers and Marx 1996) are also a low-rank approximation and a simpler alternative to smoothing splines. They consider moderately large B-spline basis functions of a size smaller than the observations and modify the penalty term by a discrete order difference penalty on adjacent coefficients, i.e. the difference operator acts on the regression coefficients, $\Delta\theta_j = \theta_j - \theta_{j-1}$, $\Delta^2\theta_j = \Delta(\Delta\theta_j) = \theta_j - 2\theta_{j-1} + \theta_{j-2}$ and in general $\Delta^d\theta_j = \Delta(\Delta^{d-1}\theta_j)$ (see Eilers et al. 2015 for further details).

As in GLMs more than one explanatory variable can be incorporated as well as interactions between covariates with different degrees of smoothness. This can be done using *tensor products* which are based on a set of basis functions that allow having different marginal smoothness penalties (Wood 2006a). Other constructions for interactions, such as the mentioned thin plate regression splines are also available but are not invariant to the rescaling of any of the covariates. Thus, *tensor product* smooths are especially useful for a general framework where variables may be measured in different units.

Compared to GLMs, GAMs are capable of fitting more complex ecological response

shapes and are used extensively in SDM (Elith et al. 2006). The most popular R package to fit GAMs is `mgcv` (Wood 2019) which includes several methods for the construction of smoothing functions. Thus, the package can be either used to fit generalised additive models with the implemented estimation methods (penalized likelihood or quasilielihood maximization with smoothness selection by GCV, GACV, gAIC/UBRE or (RE)ML (Wood 2019)) or can be used to construct the desired structures of the smoothing function to incorporate in a more complex models (as done in Chapter 3).

Apart from `mgcv`, another package that implements non-parametric smoothing functions have been used; the package `mboost` (Hothorn et al. 2018). It is based on the the so-called model-based boosting. This is a gradient descent algorithm for optimizing general risk functions using component-wise penalized least squares for fitting GAMs, based on P-spline base-learners (Bühlmann and Hothorn 2007).

1.3.2 Statistical inference on stock assessment models

The objective of stock assessment is to describe the past and current status of a stock. This involves constructing a stock assessment model that describes the dynamics of the stock along time and then, conducting statistical inference on the unknown parameters of that model based on collected data. Besides natural processes such as birth, natural mortality or growth, stock assessment models need to incorporate explicitly the fishing process that affects the fish stock over time.

The models available for assessing fish stocks range from simple to complex based on the available data for a given stock (Edwards et al. 2012a). Among the simplest models, the most widely used ones are biomass dynamic models or production models (Beverton and Holt 1957, Pella and Tomlinson 1969, Schaefer 1954). They describe the fluctuations of the stock biomass as a function of previous year biomass, catch and surplus production, ignoring any age or size structure. Alternatively, statistical catch-at-age (SCA) models (Deriso et al. 1985, Fournier and Archibald 1982, Schnute 1994) use age disaggregated data to describe the population dynamics in terms of the number of individuals in each age group. SCA models are considered more realistic, but they are also more complex and imply a greater number of parameters to be estimated, being more data-demanding. Between these two levels of

complexity, there are other types of stock assessment models, such as virtual population analysis (Pope and Shepherd 1982) or delay-difference models (Deriso 1980). However, stock assessment models are continuously evolving (Maunder and Piner 2015). In the last years, efforts have been directed towards developing models for data-limited stocks (Dowling et al. 2019), while more complex models are incorporating spatial structure (Goethel et al. 2011, Punt et al. 2000) or assessing multiple stocks simultaneously (Punt and Kennedy 1997, Punt et al. 2011).

Stock assessment models are calibrated by using observations collected from the fishery, scientific surveys and species biology. The stock assessment modelling approach of using all available data in a single analysis is known as integrated analysis (Maunder and Punt 2013 and references therein). It was first formulated by Fournier and Archibald (1982) and mathematically, it consists of constructing a joint likelihood for all the observed data that can later be used in any likelihood-based framework. Integrated analysis is composed of the following three layers:

- Population dynamics model: It describes the population dynamics in terms of essential population factors such as stock abundance and mortality, and it can include process errors.
- Data and observation model: It relates the data, such as, survey abundance index, catch, weight, age composition or other available information, to the model predictions. This includes the likelihood function, which represents the sampling process.
- Estimation method: It is the algorithm to provide point and uncertainty estimates of the model quantities of interest.

Integrated analysis belongs to a large group of statistical models called Hidden Process Models (HPM). They describe the evolution of two (usually multivariate) time series evolving in parallel: the state process that represents the unknown (hidden) underlying state of nature, and the observation process that refers to the observations from the state time series (Buckland et al. 2007, Newman et al. 2006, Thomas et al. 2005). In the context of integrated analysis for stock assessment, the state process corresponds to the population dynamics model, whereas the observation process is related to the data and observation model (Aeberhard et al. 2018).

In a formal definition of a HPM, let \mathbf{n}_t denote the state vector at time t and \mathbf{y}_t the associated measurements at time t of this state vector for t in $1, \dots, T$. For instance \mathbf{n}_t could be the vector of abundances by age at time t and \mathbf{y}_t the numbers-at-age estimated in a research survey at that time instant t . The evolution of the state process and the links between the state and observation vectors are described by probability density functions (pdfs). More precisely, a HPM is defined by three different pdfs: the pdf that generates the initial state, $g_0(\mathbf{n}_0|\boldsymbol{\theta})$, the pdf that describes the evolution of the state vector from one time period to the next, $g_t(\mathbf{n}_t|\mathbf{n}_{t-1}, \mathbf{n}_{t-2}, \dots, \mathbf{n}_0, \boldsymbol{\theta})$, and the pdf that relates the observation vector to the state vector, $f_t(\mathbf{y}_t|\mathbf{n}_t, \boldsymbol{\theta})$, being $\boldsymbol{\theta}$ the vector of the model parameters. State-space models (SSM) are a particular case of HPMs where the state process depends only on the previous state, i.e. the pdf that describes the evolution of the state vector along time is simply $g_t(\mathbf{n}_t|\mathbf{n}_{t-1}, \boldsymbol{\theta})$.

The estimation method used to conduct inference on the unknown model parameters $\boldsymbol{\theta}$ and the state vectors \mathbf{n}_t of HPMs in general, and integrated analysis in particular, can be based either on frequentist or on Bayesian methods (Maunder and Punt 2013). While one of the most appealing features of HPMs is that they can account explicitly for both process and measurement errors, this can lead to quite complex models for which inference is a challenge that involves high-dimensional integration. Auger-Méthé et al. (2020) provide a thorough review of different inference methods, and include an in-depth tutorial about how to fit these models in R. Some of the most common methods include Markov chain Monte Carlo (MCMC) methods (Brooks et al. 2011, Gamerman and Lopes 2006, Gilks et al. 1996), sequential Monte Carlo (SMC) methods (Doucet et al. 2001) or Laplace approximation (Fournier et al. 2012). In the particular case of stock assessment models, Aeberhard et al. (2018) review the work done on stock assessment to fully utilize the SSM framework. Other papers that have framed stock assessment models as state-space models can be found in Mäntyniemi et al. (2013; 2015), Meyer and Millar (1999), Millar and Meyer (2000) or Nielsen and Berg (2014).

In what follows, we expand on two of the estimation methods used for stock-assessment that will be used in this thesis: Automatic Differentiation (AD) and Markov chain Monte Carlo (MCMC).

1.3.2.1 Automatic Differentiation

Common methods in frequentist statistical inference, such as MLS (Minimum least squares) or MLE (maximum likelihood estimation), are based on the minimization/maximisation of an objective function and thus, require the calculation of the derivatives. The computation of derivatives can be classified in four type of methods: manually working and coding derivatives, numerical differentiation, symbolic differentiation, and automatic differentiation (Baydin et al. 2018). In the case of stock assessment models, manual and symbolic options cannot be applied as usually they are not closed-form models, whereas numerical methods can be inaccurate due to the high number of parameters. As an alternative, automatic differentiation (AD) can be applied to highly non-linear models with a large number of parameters. AD is based in two main ideas (Griewank et al. 1989):

- The function is split into a sequence of elementary operations with known derivatives.
- The application of the chain rule in the forward or reverse mode, being the latter superior in terms of computational effort but may need more storage.

Efforts on AD go back to the 1950s although it gained popularity in the 1980s due to the improvements in programming languages and the possibility to implement an efficient reverse mode (Baydin et al. 2018). AD has been employed in many different areas such as computational fluid dynamics, weather modelling, engineering design optimization or biomedicine as listed in the community web www.autodiff.org, however, it has not been extensively used in statistical inference.

In the past decade, motivated by fish stock assessment models with hundreds of parameters and highly non-linear objective functions, Automatic Differentiation Model Builder (ADMB Fournier et al. 2012), was developed as a tool for statistical parameter estimation for non-linear statistical models. It is based on Automatic Differentiation for computation of derivatives combined with a quasiNewton optimizer and a simplified C++ template language for model development. Using AD, high order derivatives can be easily calculated which facilitates optimization of the Laplace approximation (Fournier et al. 2012) for the marginal likelihood in complex models with random effects. Uncertainty of estimated parameters can be evaluated through various methods such as profile likelihood, delta method or MCMC. It is available

as a free open-source software (<http://www.admb-project.org>).

Although it was originally developed to fit stock assessment models, it has spread within other scientific fields, ecology, medicine or economy. Moreover, based on the established ADMB, Template Model Builder (TMB Kristensen et al. 2016) was developed as an easier and faster R package for the implementation of non-linear random effects models, combining existing high-performance libraries and easy access to parallel computation (<http://tmb-project.org>).

1.3.2.2 Markov chain Monte Carlo

Opposed to frequentist approaches, where the parameter vector to be estimated (θ) is unknown and fixed, in Bayesian statistics θ is treated as a random variable. Considering the Bayes' theorem, the posterior distribution of the parameter vector given the data can be expressed as:

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)},$$

where $P(\theta)$ represents the prior distribution of θ before any data is observed, $P(y|\theta)$ is the likelihood of the observations $y = y_1, \dots, y_n$ and $P(y)$ is the marginal distribution of y that can be expressed as :

$$P(y) = \int P(y|\theta)P(\theta) d\theta .$$

Solving this integral analytically is sometimes infeasible. Markov chain Monte Carlo (MCMC) methods (Gilks et al. 1996) is one of the methods that can be used to draw samples from a target density, such as the posterior distribution in Bayesian statistics, i.e. $\pi(\theta) = P(\theta|y)$. Metropolis-Hastings (MH, Chib and Greenberg 1995) is one of the first and most popular algorithms for MCMC, together with the Gibbs sampler (Casella and George 1992), popular for multivariate target distributions. The idea behind these algorithms is the updating of the parameter vector, $\theta \in \mathbb{R}^D$, iteratively until the equilibrium target distribution is reached. According to the MP algorithm a candidate parameter value x^* is generated from an arbitrary jumping proposal distribution, $J(x|x_t)$, which is accepted as the next value of the chain (x_{t+1}) with a probability $\alpha = \min(1, \frac{\pi(x^*)J(x_t|x^*)}{\pi(x_t)J(x^*|x_t)})$. Although any proposal distribution should work to reach converge, the speed will depend on the chosen proposal

distribution. The key of the second algorithm, the Gibbs sampler, is that it relies on univariate conditional distributions, where all except one variable have fixed and updated values as proposal distribution, resulting in an acceptance probability of $\alpha = 1$. D random variables are simulated sequentially from these univariate distributions, which is usually more simple than the simulation of a D -dimensional vector from the full joint distribution. Look at Algorithm 1 for the general structure of the method.

Algorithm 1 Gibbs sampling

```

Initialize  $x^0 \in \mathbb{R}^D$ 
for iteration  $n = 1, \dots, N$  do
  sample  $x_1^n \sim \pi(x_1 | x_2^{n-1}, x_3^{n-1}, \dots, x_D^{n-1})$ 
  sample  $x_2^n \sim \pi(x_2 | x_1^n, x_3^{n-1}, \dots, x_D^{n-1})$ 
  ...
  sample  $x_D^n \sim \pi(x_D | x_1^n, x_2^n, \dots, x_{D-1}^n)$ 
end for

```

An important issue of algorithms for MCMC is when to start and when to stop them in order to obtain stationarity and convergence. First draws may not be close to the target distribution. A common practice regarding this issue is to use a burn-in period, i.e., discarding early iterations and start collecting samples which are sufficiently close to the stationary distribution. Apart from that, in order to avoid autocorrelation, thinning is also commonly used, which implies discarding all but every k th sample. Concerning the stopping point, the longer the chain the better estimates can be obtained, however, there are several convergence diagnostics that help deciding when to stop. One of the most popular diagnostic is the Gelman-Rubin test (Gelman et al. 1992); based on parallel chains, the within-chain variance and the between-chain variances are estimated and the ratio between these to estimates is compared to one. The recommendation is to stop the algorithm when this ratio is close to one, with a general cut-off value of 1.1 (Gelman et al. 2013). Diagnostics based in spectral density are also popular, such as, Geweke's test (Geweke et al. 1991), which compares the means of the sampled parameter on two different parts of the chain. If the means are close enough it is assumed that the two samples, usually the last part and an earlier part, come from the same distribution. Another spectral density-based diagnostic is the Heidelberg and Welch test which consists in applying an stationarity test successively. First it is applied to the whole chain, if the null hypothesis of stationarity is rejected, the test is repeated without the first

10%, 20%, ... of the chain. If stationarity is not accepted before reaching the 50% it is considered a failure of stationarity. Graphical convergence diagnostics are also easy and popular tools, such as trace plots; showing the realization of the Markov chain at each iteration so that it can be easily visualized if the chain is well mixed or is stuck in some part, or autocorrelation plots; showing values of the correlation of the sampled parameter with itself at differing time lags, having that as lag increases, values should be close to zero for a fast mixing. A thorough review of convergence an diagnostic for MCMC can be found in Roy (2020).

For state-space models, Bayesian approaches can be implemented using several available flexible software packages such as WinBUGS, OpenBUGS or JAGS (Just Another Gibbs Samples) from the BUGS family (Bayesian inference Using Gibbs Sampling, Lunn et al. 2012) or more recent tools using Hamiltonian Montecarlo such as STAN (Carpenter et al. 2017). These are most widely used generic software for Bayesian inference, all using MCMC algorithms.

Applications in stock assessment field can be found for different types of models, such as, simple production models (Meyer and Millar 1999), two-stage biomass dynamics (Ibaibarriaga et al. 2008), age-structured-length-based (Mäntyniemi et al. 2013), mark-recapture models (Michielsens et al. 2006) or the JABBA tool (Mourato et al. 2018).

1.3.3 Management strategy evaluation

In general, the poor management of fish stocks or natural resources, has led to depletion or population extinction in many cases. More than the half of all assessed fish stocks may be in need of rebuilding (Worm et al. 2009). In 2018 from all stocks in the Northeast Atlantic more than the 30% were overexploited (the exploitation level was above the MSY fishing mortality) according to the STECF report (Scientific, Technical and Economic Committee for Fisheries, STECF 2020). Thus, the general objective is to reduce this percentage by providing advice for more species and providing better advice. The Common Fisheries Policy (CFP) establishes general objectives and rules for fisheries management in Europe (Salomon et al. 2014), defining the total allowable catch (TAC) as one of the core tools for management.



Figure 1.3: General scheme of the process to produce management advice.

The general steps followed for management advice are shown in Figure 1.3. Firstly the population to be managed is identified and data on biology, abundance and exploitation of the stock are collected. These data can be obtained from scientific surveys and commercial catches, however, not all stocks have scientific monitoring and sometimes only commercial data are available. Based on the availability and quality of the data, the status of the stock is assessed, usually through international scientific collaboration, given that stocks can be distributed along several countries. This assessment is ideally performed integrating all available data in mathematical models but not always is possible and trend-based or qualitative assessments can be required. Based on the obtained stock assessment, management advice is produced using estimates from mathematical models and experts opinion in order to meet pre-specified objectives. Finally, the scientific management advice can be applied to the corresponding population, however, the implementation of the advice is subjected to several factors, such as the agreement between countries or other political considerations, that are out of the scientific process.

The traditional approach for providing scientific advice for the management of fish stocks has been based on the development of a good assessment of the resource, that estimates the past and present stock status, and the subsequent short-term forecasts in order to derive a management advice (Butterworth 2007). This is usually repeated on an annual basis, that might result in highly variable recommendations from year to year. Recent advances in fisheries science have been linked to the development of evaluation tools for management actions of harvested stocks in a virtual world,

without their direct application to the real system (Bunnefeld et al. 2011). Management strategy evaluation (MSE, Butterworth and Punt 1999, Kell et al. 2006, Smith 1994) is a simulation framework that evaluates the performance of alternative fisheries management measures in terms of trade-offs between pre-specified objectives. It involves the projection of the entire management system into future years by means of Monte Carlo simulation (McAllister et al. 1999). The aim is to identify a management strategy that meets these management objectives while being robust to the system's uncertainties (Punt et al. 2016, Sainsbury et al. 2000).

The MSE framework is composed of two systems: the true system and the observed one (Kell et al. 2007). The true system is represented by a mathematical model called the operating model (OM), which represents the real dynamics of the fishery system. It is composed of the stock and the fleet dynamics (biological and fishing processes) and other potential covariates that could impact the system. The management procedure (MP) represents the observed system and includes: 1) an observation model that mimics the data monitoring from the true population and fishery in the operating model, 2) stock assessment which gives the stock status based on observed data and 3) decision of management actions based on the stock assessment output. These management actions are usually represented by Harvest Control Rules (HCR) that are mathematical formulas for pre-agreed management actions as a function of variables related to the stock status (ICES 2006) that determine target levels of fishing mortality, total catch or effort. HCRs are defined depending on the available knowledge of fish stocks. The MP is linked to the OM by the implementation of the decided management actions, which will have an impact on the dynamics of the fishery system (Figure 1.4).

One of the strengths of MSE is that it includes different sources of uncertainty in the components of the described system. Francis and Shotton (1997) identified the main sources of uncertainty in fisheries in general. Later, Punt et al. (2016) distinguished that the major sources of uncertainty in MSE in particular are: 1) process uncertainty, related to stochasticity on demographic characteristics 2) parameter uncertainty, coming from variability on estimated parameters used to condition the case study 3) observation error in the data, due to sampling inaccuracy, 4) uncertainty when conducting assessment, related to the estimation process of the assessment model and 5) implementation error, coming from the variability in the fulfilment of the defined management policy.

Depending on the way the observation and assessment errors are included in the MSE, the ICES Study Group on Management Strategies differentiated three types of MSEs (Figure 1.4). A "full-feedback" MSE incorporates the stock assessment in the Monte Carlo simulation, a "shortcut approach" MSE uses an emulator of the stock assessment and an "empirical" MSE is based directly on stock status indices collected from the population and/or the fishery. Being able to incorporate the assessment model in the MSE of a specific stock can represent a substantial improvement for the correct characterization of the entire process and the related uncertainty, which is crucial for a robust management plan. However, the implementation of a full-feedback approach can be tedious and very time consuming.

Among the available software for MSE, we will use **FLBEIA** (Bio-Economic Impact Assessment using FLR, Garcia et al. 2017) that has been applied for a large number of case studies across Europe (Nielsen et al. 2018). **FLBEIA** is an R package (R Core Team 2018) built on the basis of FLR libraries (Kell et al. 2007). It is implemented modularly, which allows for increased complexity. Inside the OM it has modules concerning biological, fleets and covariates dynamics, and in the MP there are modules for observation, assessment and advice processes. Each of the module contains functions that the user can choose or new ones can be coded in R and used within **FLBEIA** to fulfil case specific requirements, as done in this thesis.

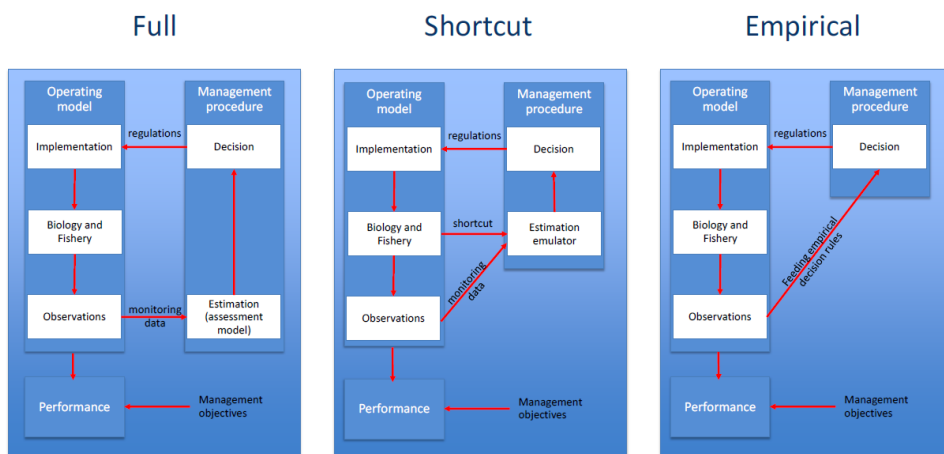


Figure 1.4: Schematic representation of different forms of MSE, modified from Punt et al. (2016) and taken from ICES (2020a).

1.4 Objectives

The general objective of this thesis is to use Monte Carlo simulation to test the improved statistical models for species distribution models, stock assessment models and management advice of fish stocks, taking as a case study the stock of sardine in the Bay of Biscay.

In particular, the specific objectives of this thesis are:

- To propose and test via Monte Carlo simulation a habitat modelling approach that meets the ecological niche theory.
- To compare the established stock assessment model for our case study to other alternative models and to evaluate, using Monte Carlo simulation, uncertainty estimation methods when introducing flexibility through non-parametric smoothing in fishing mortality.
- To develop a full-feedback MSE for our case study introducing different uncertainty sources through Monte Carlo simulation.

1.5 Structure

This thesis is structured in six chapters, with four annex documents that include supplementary material.

After the general introduction provided in this first chapter (Chapter 1), a proposal for species distribution modelling using shape-constrained GAMs is presented, together with real data applications for sardine and other species (Chapter 2). Then, we summarise the current assessment methodology for sardine and we present two alternatives using a simpler approach based on AD and a Bayesian ad-hoc approach (Chapter 3). In the next chapter, we evaluate the uncertainty estimation in stock assessment models with non-parametric effects on fishing mortality using different estimation approaches (Chapter 4). In the last main chapter, we use the stock assessment results for conditioning population dynamics for MSE, and we include the assessment model explicitly in the process together with other uncertainty sources

to evaluate the effect of each of uncertainty source in an MSE (Chapter 5). Each of these main four chapters include their own introduction, material and methods, results, and discussion. Finally, we summarize the main conclusions of the work and propose further work (Chapter 6). A list of the scientific contributions derived from the work presented in this dissertation is also provided.

Modeling species presence-absence in the ecological niche theory framework using shape-constrained GAMs

In this chapter we use Monte Carlo simulation to evaluate the use of shape-constrained generalised additive models (SC-GAMs) to build species distribution models under the ecological niche theory framework. Based on the simulation study and a real data application, we compare their performance with respect to other regression models without shape-constraints (such as standard GLMs and GAMs with varying degrees of freedom) and also to models based on so-called “Plateau” climate-envelopes. The proposed approach is applied to model the thermal niche of sardine eggs in the Bay of Biscay and to characterize the egg distribution at the spawning of other two pelagic species as a function of several environmental variables.

2.1 Introduction

Species distribution models (SDM) associate locations of species occurrence or abundance with the characteristics of their environment in order to predict geographical distributions (Elith and Leathwick 2009). A variety of statistical methods have been applied to species distribution modelling (e.g. Guisan and Zimmermann 2000, Merow et al. 2014) such as regression-based models (Guisan et al. 2002, Hastie and Tibshirani 1990), environmental envelopes (BIOCLIM, Busby (1991), Cerdeira et al. (2018)), mechanistic approaches (CLIMEX, Kriticos et al. (2015)), neural networks (Pearson et al. 2002) and maximum entropy models (MAXENT, Phillips et al. (2006)). However, most of these methods can result in species responses along en-

environmental gradients that are convex or multimodal, and consequently not ecologically meaningful or otherwise difficult to interpret (see below for further discussion).

It has been often claimed that species distribution models need a stronger theoretical background (see Austin 2002, Elith and Leathwick 2009, Jiménez-Valverde et al. 2008 for a detailed review). Recently, several authors have attempted to clarify the relationship between species distribution models and the concept of ecological niche (Kearney 2006, Peterson et al. 2011, Pulliam 2000, Soberon and Nakamura 2009). Although the debate is still open (Halvorsen 2012), it is agreed that the resulting statistical model should be ecologically plausible (Elith and Leathwick 2009). According to the ecological niche theory, species distributions should provide unimodal relationships with respect to environmental gradients (Hutchinson 1957). When environmental conditions become less favourable, various stages of the life cycle (feeding, growth and reproduction) are affected, resulting in a lower presence of the species (Austin 1987, Helaouet and Beaugrand 2009). Hutchinson (1957) defined the niche as an “n-dimensional hypervolume”, where the dimensions are environmental states within which a species can survive. Hutchinson (1957) also distinguished “fundamental” from “realized” niches, to define the conditions under which species could survive and those where they live, respectively: (1) the fundamental niche is determined by the physiological range of tolerance of the species to environmental factors in the absence of biotic interactions (e.g. competition, predation or parasitism), and (2) the realized niche is the part of the fundamental niche occupied by the species, given factors such as the presence of competitors/predators and dispersal limitations of the species (Soberón and Arroyo-Peña 2017). As a result, the realized niche tends to be smaller than the fundamental niche (Soberón and Arroyo-Peña 2017). Although the fundamental niche should be unimodal, the realized niche can be bimodal when the centre of the niche gradient is affected by interspecific competition or when the species is not occupying the most suitable habitat due to dispersal limitation (Austin 2002). However, scarce species data and the heterogeneous distribution of species occurrence along gradients are the most problematic situation leading to multimodal and ecologically non-meaningful relationships with environmental variables. For instance, data on the occurrence of a fish species which spawns in two river mouths, separated by a latitudinal distance, can easily lead to a bimodal distribution along a temperature gradient.

The concept of niche has evolved after the 80’s and incorporates the impacts of

the organism on environmental factors (Chase and Leibold 2003) to better explain competition and species coexistence (Pocheville 2015). For the pragmatic purpose of modelling species distribution, this can include several types of variables, as well as those defining the niche namely, direct variables, resource variables, and indirect variables (Austin and Smith 1990, Guisan and Zimmermann 2000, Huston 1994). Direct variables are those environmental factors having a direct physiological impact on the species but are not consumed, typical examples being pH affecting plant growth or temperature affecting fish growth. Indirect variables do not have a direct physiological impact, but might be highly correlated with the species through the combination of related factors effects. For example, elevation can affect species presence through the combined effect of atmospheric pressure, temperature and UV radiation, and have ecophysiological implications. Resource variables refer to limiting factors (i.e. essential resources consumed by the species, such as food and oxygen) and biotic interactions (competition, predation or mutualism). The first two types (direct and indirect) of variables are within the group of variables that do not interact dynamically with the species and hence are not affected by species abundance. These were termed “scenopoetic” variables by Hutchinson (1978). In contrast, resource variables interact with the species and are affected by species presence and abundance. In the context of species distribution models, several authors (Austin 1980; 2007, Austin and Smith 1990) have discussed the shape of response curves and how this depends on the variable type. While there is no theoretical expectation regarding the shape with respect to indirect variables, they advocated that the fundamental niche as a function of direct variables should be unimodal (symmetric or not), and for limiting factors should be logistic or Michaelis-Menten saturation curves. SDMs based on non-scenopoetic variables might require more elaborate mathematical methods to include species interaction (Peterson et al. 2011). Thus, species distribution models need to combine environmental variables that are expected to meet the ecological niche theory with other explanatory variables having no shape restrictions.

Commonly used methods to build species distribution models in the ecological niche theory framework include regression-based methods, such as generalised linear models (GLMs) and generalised additive models (GAMs). They have been well-documented, both theoretically and empirically (Coudun and Gegout 2006, Guisan et al. 2002, Lehmann et al. 2002, Scott et al. 2002). Generalised linear models (Guisan et al. 2002, McCullagh and Nelder 1989) are widely used in statistical

ecology as a simple parametric technique that may allow symmetric bell-shaped ecological response curves (Coudun and Gegout 2006, Jamil and Ter Braak 2013). However, this can be too restrictive as often non-symmetric responses have been observed (Austin 2007, Huisman et al. 1993). Generalised additive models are also very popular as semi-parametric and more flexible regression-like approaches (Austin 2002, Heikkinen and Makipaa 2010). Pedersen et al. (2019) proposed an extension of GAMS called hierarchical GAMS (HGAMS) to model intergroup variability in ecology; these models allow smooth functions to vary between groups and can be used to test if the smooth functions are common across groups. In general, GAMS and related extensions allow flexible non-symmetric shapes, but they can result in implausible response curves, contrary to the ecological niche theory framework. Current practice tends to use low degree smoothing functions, such as splines with a low number of knots, in order to obtain response curves in agreement with the niche theory (Chust et al. 2014). However, restrictions on the number of knots and/or the degrees of freedom (by altering the smoothing parameter within GAMS, say) do not guarantee this aim, and a visual evaluation of resulting fitted curves is still required.

Other attempts to build species distribution models under ecological niche theory include Beta functions (Minchin 1987) and Huisman-Olff-Fresco (HOF) curves (Huisman et al. 1993), fitting unimodal and monotonic response curves with or without symmetry. A simulation study by Oksanen and Minchin (2002) concluded that HOF curves obtained better results than Beta functions and Gaussian response models which provided biased or inappropriate models. However, they are only allowed for single-variable analysis. Alternatively, the “Plateau” method proposed by Brewer et al. (2016) is an environmental envelope model based on a concave piece-wise polynomial function. While providing an ecologically meaningful method (unimodal even if not symmetric), this approach can be easily extended to multiple environmental variables accounting for potential interactions between the climatic variables.

Shape-constrained generalised additive models (or simply SC-GAMS, Pya and Wood 2014) are based on the same statistical framework as GLMs and GAMS regression methods, but they allow us to incorporate monotonicity and concavity shape-constraints in the component functions of the linear predictor of the GAMS. Imposing concavity constraints should be an effective alternative to fitting non-symmetric parametric response curves, while retaining the unimodality constraint, required by the ecological niche theory, for direct variables and limiting factors. Recently, several

successful applications of shape-constrained models to incorporate prior knowledge about the shape of the response curve along variables of interest have been found related to animal activity, pollution mortality, tree height-diameter relationships or petroleum engineering (Guevara et al. 2018, Hofner et al. 2016, Schmidt et al. 2018).

The objective of this work is to assess the performance of SC-GAMs in fitting species distribution models under the ecological niche theory in comparison with other approaches. We considered two different implementations of SC-GAMs: the maximum likelihood implementation from the `scam` R (R Core Team 2018) package (Pya 2018); and the component-wise boosting approach from the `mboost` R package (Hothorn et al. 2018). First, we conducted a simulation study to assess performance in terms of goodness-of-fit and agreement with the ecological niche theory—comparisons with respect to GLMs, GAMs with different degrees of smoothness, and the “Plateau” method. All methods were evaluated within a real case study, modelling the probability of the presence of sardine eggs in the Bay of Biscay as a function of sea surface temperature. Secondly, SC-GAMs were used to model egg distribution at the spawning of three pelagic species as a function of several environmental gradients, combining direct and indirect variables, and accounting for model selection and validation.

2.2 Material and methods

2.2.1 Regression models for presence-absence data

We considered six different approaches for fitting species distribution models. In all of them, species presence-absence data were modelled as a function of an environmental variable x . Let Y be the response variable coming from a Binomial distribution with probability of presence $p(x)$. The *logit* transformation of $p(x)$ is a function of the environmental variable x (presented here using a single explanatory variable for simplicity):

$$\log\left(\frac{p(x)}{1-p(x)}\right) = g(x). \quad (2.1)$$

The simplest model is a binary logistic generalised linear model (GLM, McCullagh and Nelder 1989, Oksanen et al. 2001, Ter Braak and Looman 1986) where the linear predictor is a second order polynomial of the environmental variable:

$$g(x) = \beta_0 + \beta_1 x + \beta_2 x^2. \quad (2.2)$$

For $\beta_2 < 0$, this results in a unimodal and symmetric relationship between the species response and the environmental variable.

Generalised additive models (GAMs, Hastie and Tibshirani 1990, Wood 2006a) are a generalisation of GLMs, where the linear predictor is a smooth function of the explanatory variable. In a binary regression model with a logit link, we have the form:

$$g(x) = \beta_0 + f(x), \quad (2.3)$$

where $f(x)$ is a smooth function that can be represented from kernel smoothing or local linear methods to splines-based regression methods (see chapter 1).

The R package `mgcv` (Wood 2019) is the most popular R package to fit GAMs. The use of GAMs has already been proposed in the literature on habitat modelling and ecological niche theory (Chust et al. 2014). Generally, the species response curve is not constrained to a particular shape but instead is controlled by limiting the flexibility of the model by selecting the number of knots. GCV (Generalised Cross-Validation) criterion is used for smoothing parameters estimation as a default method in the used `mgcv` package.

The methods proposed in this chapter, SC-GAMs, are based on generalised additive models, allowing us to impose shape-constraints on the linear predictor function. In Bollaerts et al. (2006) or Eilers (2017) an algorithm based on asymmetric penalties in an iterative procedure is proposed. A similar approach is considered in Pya and Wood (2014) using shape constraints (monotonicity, concavity/convexity or mixed-typed constraints) with B-splines on the first or second derivatives of the smooth terms. The latter methods are implemented in the R package `scam` in a more general framework, e.g. including bivariate tensor product smooths (Pya 2018).

For fitting species distribution models in agreement with the ecological niche theory, we imposed concavity constraints in the linear predictor scale ($f''(x) \leq 0$) for which the condition $\theta_j \leq \theta_{j-1}$ suffices (see Pya and Wood 2014 for further details). As proved in Annex A, this implies unimodal probability response curves. The implementation of the method allows for an automatic selection of the smoothing

parameters by calling the *gam* function in the R package *mgcv*. However, we found the algorithm fails to converge in some situations. This issue is discussed in the next section.

Another method we considered is the so-called model-based boosting. Boosting is a gradient descent algorithm for optimizing general risk functions using component-wise penalized least squares for fitting GAMs (see Bühlmann and Hothorn 2007, Hothorn et al. 2010 for further details). Boosting is a popular ensemble method in machine learning, where multiple learners (usually known as base learners) are trained to solve the same problem. In the particular case of modelling species distributions, shape constraints are implemented in the package *mboost* through the base-learner *bmono*, based on P-spline base-learners with an additional symmetric penalty in second-order differences on the linear predictor scale, as in Bollaerts et al. (2006). The optimal number of boosting iterations can be achieved via cross-validated estimation of the empirical risk for hyper-parameter selection. For more technical details about theoretical aspects and software implementation, see Hothorn et al. (2018) or Hofner et al. (2014).

Finally, the “Plateau” method, proposed by Brewer et al. (2016), performs climate envelope fitting via an explicitly defined concave shape on the linear predictor scale. This shape consists of an increasing slope, a possible plateau, and a decreasing slope. In the univariate case, the envelope function is defined as a piece-wise function:

$$g(x) = \begin{cases} \alpha_1 + \beta_1 x & x \leq -\alpha_1/\beta_1, \\ \beta_0 & -\alpha_1/\beta_1 < x < \alpha_2/\beta_2, \\ \alpha_2 + \beta_2 x & x \geq \alpha_2/\beta_2, \end{cases} \quad (2.4)$$

where $\beta_1 > 0$, $\beta_2 < 0$ are increasing and decreasing slopes, α_1 , α_2 are intercepts and β_0 is the *plateau* value.

2.2.2 Simulation

In order to evaluate and compare the performance of the proposed approaches for fitting species distribution models, we carried out a Monte Carlo simulation study.

First, four different theoretical response curves depending on a single environmen-

tal variable were generated within the simulation model, which are considered as the true curves for performance statistics computation. Afterwards, sampling and observation errors were introduced and presence-absence data sets were generated based on the underlying theoretical probability curves. The simulated data sets were then fitted according to the proposed models. Finally, the goodness-of-fit and the concordance of the fitted model with the ecological niche theory were measured through several performance statistics, described in section 2.2.6 below.

2.2.2.1 Environmental gradient

The real environmental variable used for data simulation was the sea surface temperature (SST) in the Atlantic Ocean in 1999 (Edwards et al. 2012b). These data are arranged on a grid with a spatial resolution of 1x1 degrees (1489 data points) covering the region between 40° and 63° in latitude and -70° and 2° in longitude. The average SST in the selected data is 9.92°C with a standard deviation of 5.35°C and minimum and maximum values of -2°C and 20.6°C respectively.

2.2.2.2 Species responses

Theoretical species response curves along the environmental gradient of SST (x) followed the generalised Beta function proposed by Minchin (1987):

$$b(x) = \begin{cases} \frac{P_0}{d} \left(\frac{x-m}{r} + b\right)^\alpha \left(1 - \left(\frac{x-m}{r} + b\right)\right)^\gamma & m - rb < x < m + r(1 - b) \\ 0 & \text{otherwise} \end{cases} \quad (2.5)$$

where m is the location of the optimum, P_0 is the maximum probability of occurrence at the mode, r is the range of occurrence along the gradient and α and γ are shape parameters. The additional parameters b and d depend only on α and γ and are introduced to reduce the complexity of the formula ($b = \alpha/(\alpha + \gamma)$ and $d = b^\alpha(1-b)^\gamma$). Combining different values for the shape parameters, we generated 4 distinct curves representing different plausible scenarios: a symmetric curve (denoted as *curve1*, with $\alpha = 4$, $\gamma = 4$), a platykurtic curve (denoted as *curve2*, with $\alpha = 0.1$, $\gamma = 0.4$), a left skewed curve (denoted as *curve3*, with $\alpha = 1.5$, $\gamma = 0.5$) and a right-skewed curve (denoted as *curve4*, with $\alpha = 1$, $\gamma = 4$) (Figure 2.1). All scenarios were generated with the same maximum probability of occurrence ($P_0 = 1$), location of optima ($m = 6.95$ °C) and range of occurrence ($r = 10$ °C).

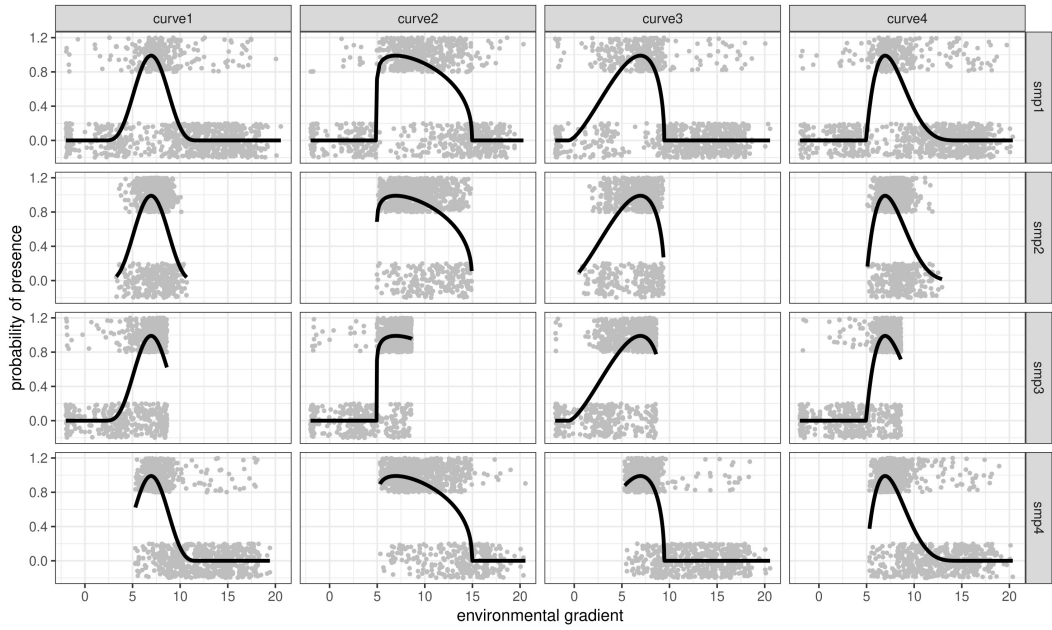


Figure 2.1: Columns are the true curves (*curve1*–*curve4*) and rows are the generated presence-absence data (grey dots) by sampling scenarios (*smp1*–*smp4*) for a single replicate.

2.2.2.3 Sampling

For each type of curve, we generated a sample of 1000 observations according to four different sampling schemes. As a first sampling option (*smp1*) samples were generated randomly along with the whole range of the environmental gradient with the same probability at all locations. In the second sampling scheme (*smp2*), the sampling probability is proportional to $b(x)$, so that the probability of sampling locations is higher around the theoretical response curve mode than in the tails. The last two options (*smp3* and *smp4*) account for the cases where the whole range of the environmental gradient is not observed, having a sampling probability of zero above (or below) a specific value of the gradient (see the rows of Figure 2.1).

2.2.2.4 Presence-absence data

The presence-absence data y was generated via a Bernoulli distribution with probability of occurrence $p(x)$, which is a noisy version of $b(x)$ in equation (2.5). In order to mimic the effect of (unobserved or unmodelled) environmental variables other than x , $p(x)$ was drawn from a beta inflated distribution (*BEINF*, allowing

for zero and one inflation) implemented in the R package `gamlss.dist` (Rigby and Stasinopoulos 2005, Stasinopoulos et al. 2019):

$$p(x) = \text{BEINF}(b(x), \sigma, \nu, \tau), \quad (2.6)$$

where the theoretical occurrence probability $b(x)$ is the mean of the distribution, $\sigma = 0.1$ is the scale parameter and $\nu = \tau = 0.1$ are parameters modelling the probabilities of zero and one respectively ($\nu = p_0/p_2, \tau = p_1/p_2$, where $p_2 = 1 - p_0 - p_1$ and p_0 and p_1 are probabilities of zero and one respectively). For further details on *BEINF* parametrization see Stasinopoulos et al. (2019).

2.2.2.5 Model fit

For each type of curve and each sampling scheme, 100 replicated data sets were generated (a total of 1600 data sets). Each generated data set, with 1000 observations each, was fitted using the proposed methods. Table 2.1 summarizes the six approaches considered (namely “*GLM*”, “*GAMhk*”, “*GAMlk*”, “*SCAMfixSP*”, “*boost*” and “*Plateau*”) and includes: the functional form of the model; constraints (if any); type of basis function (or base learner in the case of *boost*); penalty (yes or no); and finally the corresponding R packages and specific functions.

It is important to state some options we fixed in performing the simulations: i) for GAM methods we consider a low number of knots ($K = 3$, in *GAMlk*) and a higher number of knots ($K = 10$, in *GAMhk*), following Chust et al. (2014) for illustrative purposes; ii) for the SC-GAM’s implementation in the R package `scam`, we found several convergence problems in the current implementation (*scam* version 1.2-4), and hence we decided to remove the penalty from the model by fixing the smoothing parameter (with the argument `sp`) to 10^{-4} and controlling the smoothness with a fixed number of knots in the construction of the model bases; iii) boosting is a computationally more expensive method but overcomes the convergence problems in *scam* (see Annex B for implementation details and code).

For each sampling scenario, fitted values were obtained along with the corresponding sampled environmental gradient interval while predictions were computed for the whole gradient interval. Analyses were performed using the computing environment R (R Core Team 2018).

Table 2.1: Summary of the six approaches considered.

| Method | Functional form | Constraints | Basis | Penalty | Package | Function |
|----------------|-------------------------------------|----------------------------|--------------------------|---------|----------------------|--------------------------|
| <i>GLM</i> | $\beta_0 + \beta_1 x + \beta_2 x^2$ | $\beta_2 < 0$ | 2nd order polynomial | No | <code>stats</code> | <code>glm</code> |
| <i>GAMlk</i> | $\beta_0 + \sum_{j=1}^K f_j(x)$ | $K = 3$ | tprs | Yes | <code>mgcv</code> | <code>gam</code> |
| <i>GAMhk</i> | $\beta_0 + \sum_{j=1}^K f_j(x)$ | $K = 10$ | tprs | Yes | <code>mgcv</code> | <code>gam</code> |
| SCAMfixSP | $\beta_0 + \sum_{j=1}^K f_j(x)$ | $f''(x) \leq 0$ | B-splines with concavity | No | <code>scam</code> | <code>scam</code> |
| <i>boost</i> | $\beta_0 + \sum_{j=1}^K f_j(x)$ | $f''(x) \leq 0$ | concavity constraint | Yes | <code>mboost</code> | <code>gamboost</code> |
| <i>Plateau</i> | See Eq. (2.4) | $\beta_1 > 0, \beta_2 < 0$ | piece-wise parametric | No | <code>plateau</code> | <code>fit.glm.env</code> |

2.2.2.6 Performance statistics

The goodness-of-fit of each method was evaluated in terms of the Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{\sum (p - \hat{p})^2}{n}}, \quad (2.7)$$

where p is the real, theoretical probability, \hat{p} is the estimated probability and n is the sample size ($n = 1000$ in this case).

The level of agreement with the ecological niche theory was evaluated in terms of the concavity constraint. Second derivatives along the environmental gradient were approximated via finite differentiation. Negative second derivatives for the predicted curves along the whole environmental gradient indicate that the concavity restriction is respected on the linear predictor scale, while positive values at some point would indicate that a non-concave shape has been estimated. When concavity is held, we looked at the first derivatives, computing the number of changes of sign of the fitted curve, to evaluate whether the method was capable of estimating a global maximum, as defined in the theoretical curve, or not.

Uncertainty around estimated curves was compared by means of estimated variances of predicted values along with the whole range of each curve. Coverage probabilities were computed as the percentage of theoretical values along the whole gradient that

fell inside the estimated 90% confidence intervals in each replicate (Morris et al. 2019).

2.3 Results

The six modelling approaches were applied to each replicated data set for each type of curve and sampling scheme. The proposed shape-constrained GAM methods (“SCAMfixSP”, “boost”) as well as the “Plateau” method does satisfy the concavity restriction, resulting in unimodal response curves, and show closer estimated probabilities to the true theoretical response curve compared to the rest of the methods, as illustrated in Figure 2.2 for a single replicate and single scenario. In contrast, the “GAMhk” method, the most flexible option, does not fulfil the concavity restriction and neither of the “GLM” and “GAMlk” methods are capable of detecting the maximum. Estimated probabilities with these last two methods are far from the theoretical curve, mainly for unsampled environmental gradient values (Figure 2.2).

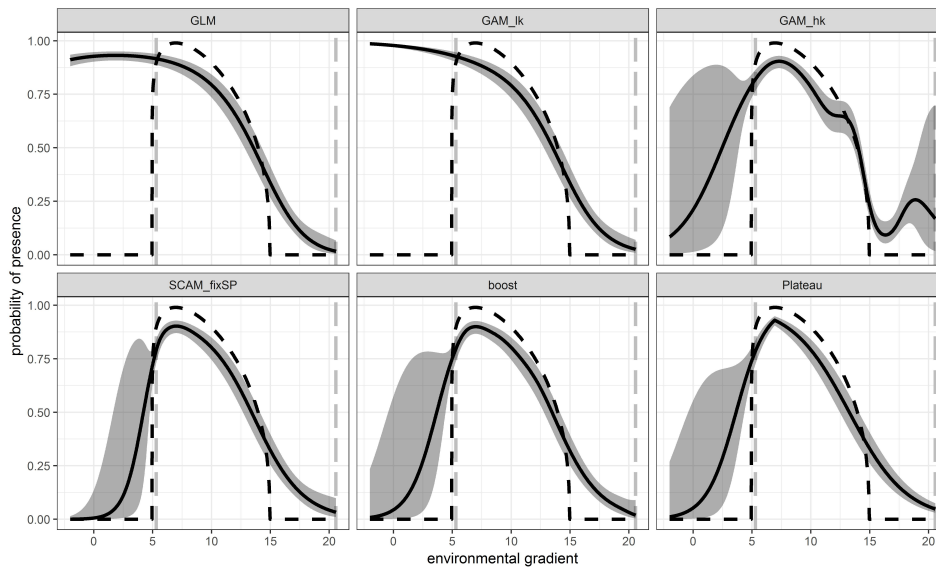


Figure 2.2: Predicted response curves by method for *curve2* and *smp4* for a single replicate. Dashed black lines represent the true theoretical response curves and solid lines represent obtained fitted curves with their corresponding 90% confidence intervals in gray. Vertical gray dashed lines represent the sampling range. Each panel corresponds to a particular method.

In order to summarize the performance statistics for all scenarios and methods, median and 0.1, 0.25, 0.75 and 0.9 percentile values across the 100 replicates were computed. The SC-GAM methods (“SCAMfixSP”, “boost”) and the “Plateau” method all satisfy concavity restrictions in all cases, assuring unimodal response curves are estimated in every scenario. They are able to detect a single global maximum in more than 80% of the replicates in most of the scenarios, with the “boost” method having the highest success percentages on detecting global maxima for all scenarios (Table 2.2). Furthermore, SC-GAMs result in better performance in terms of RMSE, giving lower values than the rest of the methods, except for the most flexible “GAMhk” method, which gives the lowest RMSE values (Figure 2.3). However, when using “GAMhk”, estimated curves almost never satisfy the concavity restriction (only 40 fitted curves out of 1600 simulations are concave). The “GLM” and “GAMlk” methods are able to fit concave curves only for sampling options *smp1* and *smp2* and result in worse RMSE values than the shape-constrained methods. For the rest of the sampling options (*smp3* and *smp4*), these methods are not able to always fit concave curves, and when concavity does hold, global maxima are not detected in most cases (Table 2.2).

Table 2.2: Percentage of replicates for each scenario and method for which estimated response curves are concave in the linear predictor scale and percentage of fitted curves that detect a single global maximum.

| | curve1 smp1 | curve1 smp2 | curve1 smp3 | curve1 smp4 | curve2 smp1 | curve2 smp2 | curve2 smp3 | curve2 smp4 | curve3 smp1 | curve3 smp2 | curve3 smp3 | curve3 smp4 | curve4 smp1 | curve4 smp2 | curve4 smp3 | curve4 smp4 | |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----|
| GLM | | | | | | | | | | | | | | | | | |
| concave % | 100 | 100 | 85 | 0 | 100 | 100 | 15 | 100 | 100 | 100 | 100 | 100 | 0 | 100 | 100 | 19 | 0 |
| max detected % | 100 | 100 | 47 | 0 | 100 | 100 | 8 | 100 | 100 | 100 | 100 | 100 | 0 | 100 | 100 | 6 | 0 |
| GAM lk | | | | | | | | | | | | | | | | | |
| concave % | 100 | 100 | 88 | 0 | 100 | 100 | 15 | 100 | 100 | 100 | 100 | 100 | 0 | 100 | 100 | 18 | 1 |
| max detected % | 100 | 100 | 2 | 0 | 100 | 57 | 0 | 12 | 100 | 100 | 2 | 0 | 100 | 74 | 0 | 0 | 0 |
| GAM hk | | | | | | | | | | | | | | | | | |
| concave % | 0 | 2 | 0 | 0 | 0 | 31 | 0 | 2 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| max detected % | 0 | 2 | 0 | 0 | 0 | 20 | 0 | 1 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| SCAM fix SP | | | | | | | | | | | | | | | | | |
| concave % | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| max detected % | 100 | 100 | 100 | 70 | 100 | 100 | 94 | 69 | 100 | 100 | 100 | 100 | 35 | 100 | 100 | 100 | 100 |
| Boost | | | | | | | | | | | | | | | | | |
| concave % | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| max detected % | 100 | 100 | 100 | 99 | 100 | 98 | 93 | 82 | 100 | 100 | 100 | 100 | 87 | 100 | 100 | 100 | 100 |
| Plateau | | | | | | | | | | | | | | | | | |
| concave % | 100 | 100 | 100 | 100 | 100 | 100 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| max detected % | 100 | 100 | 98 | 92 | 97 | 44 | 18 | 68 | 100 | 93 | 74 | 78 | 100 | 89 | 63 | 87 | 87 |

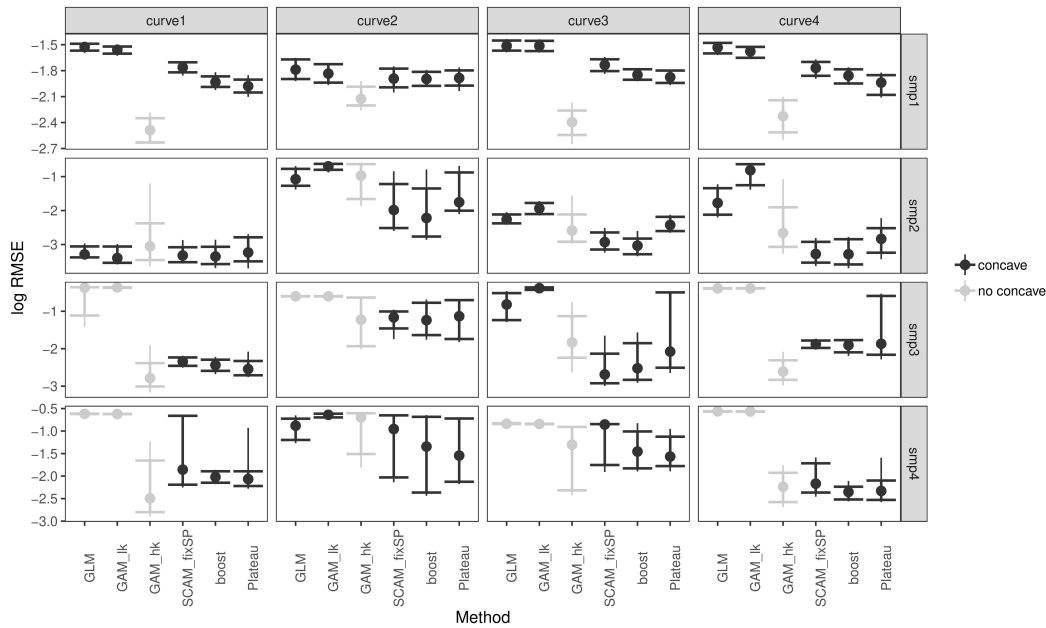


Figure 2.3: Logarithm of RMSE for all curves and sampling scenarios. Points are median RMSE values across replicates and vertical thick lines represent the 75% interquartile range, while thin lines represent the 90% interquartile range. Black elements: all obtained fitted curves are concave; grey elements: not all fitted curves are concave. Each column corresponds to a curve type and each row to a sampling scenario.

Concerning uncertainty indicators, standard deviations were computed for each data-point and were used to compute 90% confidence intervals around the estimated curves. Among the methods that are able to estimate concave shapes, SC-GAM methods have higher coverage percentages (percentage of true theoretical values that fall inside these confidence intervals, see Figure 2.4) in comparison to “GLM” and “GAMlk” methods. The most flexible method, “GAMhk”, shows the highest coverage percentages. However, we have noted that the underlying fitted curves are often not concave. Although in most scenarios, the “Plateau” method and proposed shape-constrained GAM methods show similar results (overlapping intervals), the “Plateau” method presents higher variability in results, while the “boost” method shows more stable interquartile ranges across replicates (see Figures 2.3 and 2.4). Note that obtained coverage percentages are low in all cases due to the introduced zero and one inflated error, making the estimated maximum probability lower than the theoretically fixed value ($P_0 = 1$), and estimated curve tails greater than 0 (see

Figure 2.2).

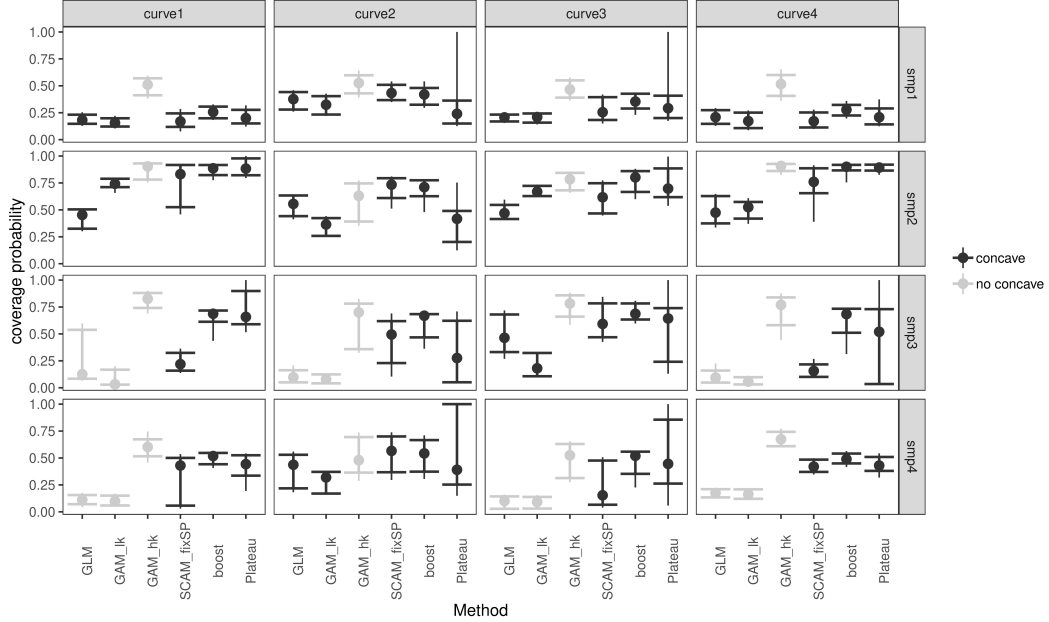


Figure 2.4: Coverage percentages for each scenario and method. Thick lines represent the 75% interquartile range, while thin lines represent the 90% interquartile range. Points represent median values. Black color represents that all obtained fitted curves are concave and grey color means that not all estimated curves are concave. Each column corresponds to a curve and each row to a sampling scenario.

2.4 Case studies

The proposed methods were also used to model the spawning habitat of some fish species in two different case studies. In the first case study, the six modelling approaches were tested and compared in univariable analysis, modelling the occurrence of sardine (*Sardine pilchardus*) eggs in the Bay of Biscay as a function of Sea Surface Temperature (SST). In the second case study, the use of the proposed shape constrained methods was extended to more than one variable. An illustration of the use of concavity restrictions for some variables in a more complex and realistic case study is provided.

2.4.1 Thermal niche for sardine eggs

Several studies have attempted to identify the main environmental variables and timing that determine sardine spawning and found that temperature was an important factor (Bernal et al. 2007, Planque et al. 2007).

We analysed the presence of sardine eggs as a function of sea surface temperature (SST) using data collected in the BIOMAN survey (Santos et al. 2018). This survey is conducted yearly in May in order to estimate the spawning stock biomass of anchovy in the Bay of Biscay by the Daily Egg Production Method (DEPM, Lasker 1985, Parker 1980). In addition, in some years the DEPM is also used to estimate the spawning stock biomass of sardine (see ICES 2017b technical report). We compiled data from the years 1999, 2002, 2008, 2014 and 2017, for which the full DEPM was applied for sardine. At each sampling location, presence-absence data of sardine eggs, geographical position (longitude and latitude), and environmental variables such as SST, were recorded. In total, 3472 data points were used for the model fitting. The presence-absence data distribution along the environmental gradient for this case study is similar to *smp2* scenarios in the simulation study, with overlapping distributions of presences and absences (see Figures C1 and C2 in Annex C for presence-absence data densities).

From the six proposed methods, “GLM” and “GAMhk” result in a convex and a multimodal response curve respectively, that are incompatible with the niche theory. Shape-constrained methods give concave unimodal curves which do agree with the niche theory (see Figure 2.5). The GAM method with fewer degrees of freedom (“GAMlk”) results in a monotone decreasing function. When predicting for temperatures lower than observed the predicted probabilities of presence continue to increase, being far from the expected bell-shaped response curve.

For each method, we computed the optimum temperature and the range of tolerance temperatures. The optimum was the value of the gradient with the highest estimated probability of presence and the tolerance was determined by the range of gradient where the predicted probability of species occurrence was higher than half of the maximum value for predicted probability (Schröder et al. 2005). For “GLM” and “GAMlk” the lower limit of the tolerance range could not be computed or can be considered $-\infty$ given that the estimated curve is monotonically increasing for

decreasing values of the gradient. The optimum SST is estimated around 12.5°C for these two methods while for the rest of the methods it located around 13.5°C. The obtained tolerances with shape-constrained methods and “GAMhk” methods are very similar giving a range from around 12 to 18°C.

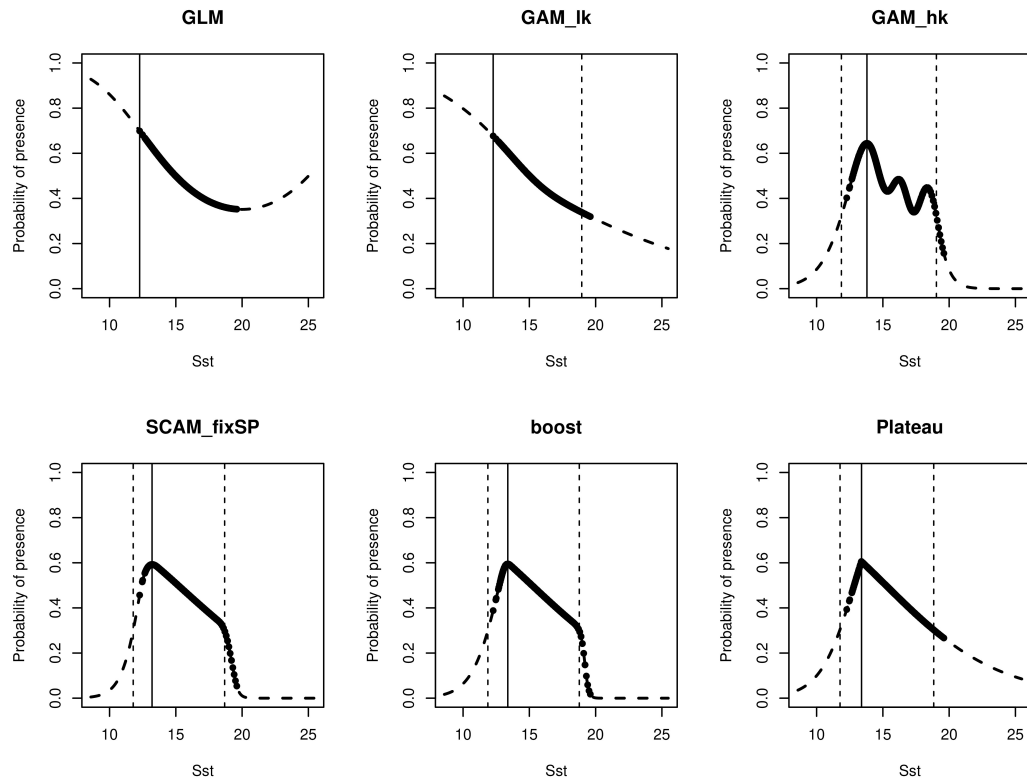


Figure 2.5: Fitted response curves (in black) and predicted curves (dashed curves) for sardine egg real presence-absence data along the SST environmental gradient. Vertical lines represent the optimum and dashed vertical lines tolerance limits. Each panel corresponds to a specific method.

2.4.2 Spawning habitat of three pelagic species

Often when fitting species distribution models, the spatio-temporal coverage of the data is limited and does not cover the range of the environmental gradient that determines the biogeographic species area (Austin 2007). In those cases, the species response is truncated and cannot be modelled adequately. The ample coverage of the ICES triennial mackerel egg survey makes it an exception. Since 1977, the survey

has been conducted every three years between January and July and covers a large area from southern Spain to the north of Scotland, to estimate the total annual egg production of the western Atlantic mackerel stock (ICES 2018a, Lockwood et al. 1981). The egg presence-absence and abundance data collected during the survey have been used to characterize the spawning habitat of mackerel: see Borchers et al. (1997), Bruge et al. (2016), Brunel et al. (2018). Within the framework of an EU programme (INDICES, EU Study 97/017), the samples collected during the 1998 triennial survey were reanalyzed and eggs and larvae of other fish species were quantified (Ibaibarriaga et al. 2007). We applied SC-GAMs to model the egg distribution of three of these species: European anchovy (*Engraulis encrasicolus*); sardine (*Sardine pilchardus*); and Atlantic mackerel (*Scomber scombrus*). Their performance was compared with respect to the other methods considered. For each sampling location of presence-absence of eggs, we compiled environmental and depth data. Environmental data were extracted from the NCEP Global Ocean Data Assimilation System, GODAS (Derber and Rosati 1989), which provides gridded 4D data with a monthly temporal resolution and a vertical resolution of 10m on $0.333^\circ \times 1^\circ$ latitude-longitude grid points of sea surface temperature (SST), salinity (SSS), the temperature at 205 m (temp205), the difference between surface temperature and temperature at 205 m (temp dif) and oceanic mixed layer (dbss obml). Depth data were obtained from the bathymetric database ETOPO1 from NOAA using the package `marmap` (Pante and Simon-Bouhet 2013) in R (R Core Team 2018) and introduced in log scale (`logbathy`).

We applied the “SCAMfixSP” method, which allows constructing models as a combination of shape-constrained variables and non-restricted variables. Among the variables available for these case study, all of them were treated as direct variables (Austin 2007), and therefore introduced with shape constraints, except for bathymetry, which was considered to be an indirect variable, and so introduced without shape restriction. Variable selection was based on AICc, as defined in Barton (2009), selecting for each species the model with the lowest AICc, after removing the variables that were not significant in univariable analysis. Depth was selected for all species models. Additionally, salinity, surface temperature and temp_dif were also selected for anchovy, obtaining a model fit with 61.1% of explained deviance. In the sardine model salinity and temperature at 205m were included obtaining 33.7% of explained deviance, while for mackerel salinity and temp_dif were selected for the final model with 29.97% of explained deviance. All selected variables and AICc

values for each species are shown in Table 2.3.

All used variables except for depth were introduced in the models with the concavity restriction on the linear prediction scale, assuring this way that the ecological niche theory was met. These variables' response curves of the selected direct variables (sea surface temperature (SST), salinity (SSS), the temperature at 205 m (temp205), the difference between surface temperature and temperature at 205 m (temp_dif) and the oceanic mixed layer (dbss obml)) are monotonic or unimodal, presenting a single optimum at most. The optimum salinity value was estimated at 35.3 psu for anchovy and 35.5 psu for mackerel, while for sardine the whole range could not be captured, resulting in a monotonic decreasing response curve (Figure 2.6). The optimum along the temp205 variable was estimated at 12.2°C for sardine. All marginal response curves for these variables and each species can be found in the supplementary material (Annex C, Figures C3, C4, C5).

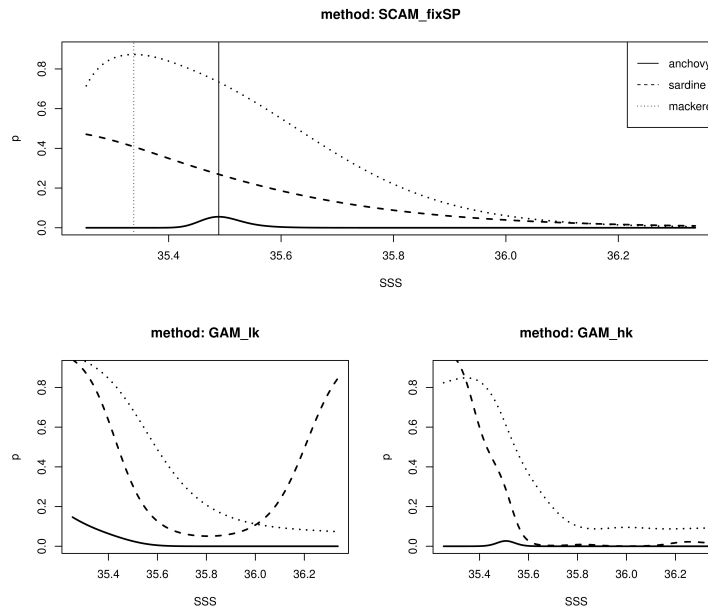


Figure 2.6: Predicted marginal response curves along salinity variable (SSS) for anchovy, sardine and mackerel fitted with three different methods; a proposed SC-GAM method (“SCAMfixSP”) and no restricted GAM approaches with $k=3$ (“GAMlk”) and $k=10$ (“GAMhk”).

This proposed SC-GAM approach was also compared with other unrestricted meth-

ods for this multivariate case study. Presence-absence data for these three species with the same selected explanatory variables were also fitted using more common GAM approaches (“GAMlk” and “GAMhk”), showing that marginal response curves are not in agreement with the ecological niche theory—some estimated response curves do not satisfy the unimodality condition (Figure 2.6).

Validation for these models was conducted via k -fold cross-validation (with $k = 5$). The data set was divided into k equally sized groups (Hijmans 2012), using 80% of randomly selected observations to run the model and the remaining 20% for validation, iteratively for each fold. Accuracy indicators, such as AUC (Area Under the Receiver Operating Characteristic (ROC) curve) (Fielding and Bell 1997, Raes and ter Steege 2007), sensitivity (true predicted presences) and specificity (true predicted absences) were computed for each k random subsets and then averaged. The threshold for presence-absence classification for each species was obtained as the values maximizing sensitivity plus specificity. Obtained AUC, sensitivity, and specificity indicators are above 70% for the three species (Table 2.3) and are similar to the values obtained when using all data without a cross-validation process, showing good out-of-sample performance of the models.

Table 2.3: For each species, selected variables in the final model (using method “SCAMfixSP”), AICc, explained deviance (%), AUC, specificity and sensitivity (%) derived from the whole data set (All data), and specificity and sensitivity (%) derived from the cross-validation process (CV).

| Species | Selected variables | AICc | expl.dev (%) | AUC | All data (%) | | CV (%) | |
|----------|---------------------------|---------|--------------|------|--------------|------|--------|------|
| | | | | | spec | sens | spec | sens |
| Anchovy | SSS,SST,temp_dif,logbathy | 375.75 | 61.10 | 0.92 | 90 | 93 | 90 | 93 |
| Sardine | SSS,temp205,logbathy | 899.35 | 33.70 | 0.80 | 79 | 82 | 79 | 82 |
| Mackerel | SSS,temp_dif,logbathy | 1322.62 | 29.90 | 0.77 | 73 | 81 | 73 | 81 |

Predicted occurrence probabilities for each species have been mapped, using for prediction environmental variables from GODAS for June 1998. Apart from optimum detection for each explanatory environmental variable, extrapolated maps allow us to identify the spawning distribution of each species (Figure 2.7). For mackerel, the north-west part of the map shows a high probability area, although presences were not collected in this area during 1998. However, it has been reported that these species do lay in this area in recent years (Bruge et al. 2016), which confirms the reliability of the model in this area. For anchovy, areas close to the coast in the

Bay of Biscay are detected as locations with a high probability of presence, while for sardine, this area is wider, extending it along the Portuguese coast and up to the Celtic Sea (Figure 2.7).

This case study data set was also analyzed using the “boost” method, which is also capable of dealing with restricted and unrestricted variables. Results were similar to those described and can be found in the supplementary material (Annex C, Figure C6).

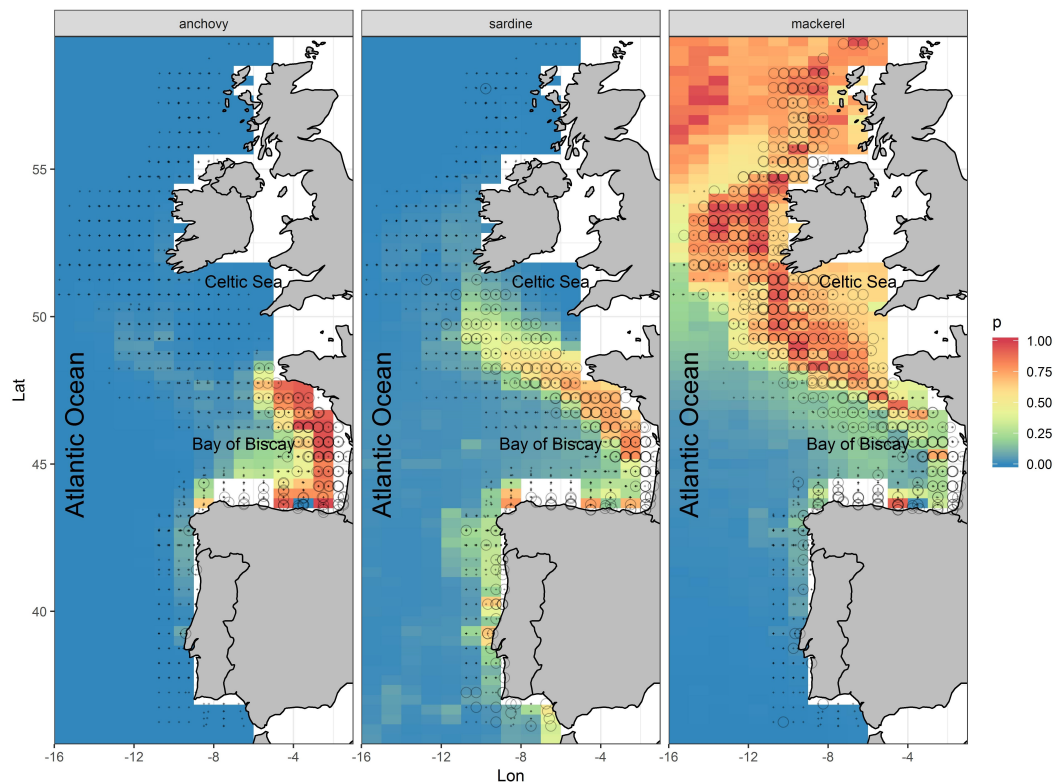


Figure 2.7: Predicted occurrence probabilities (p) in each map cell along with presences (circles) and absences (small dots) for each species in the north-east Atlantic.

2.5 Discussion

This study proposes SC-GAMs for species distribution models under the ecological niche theory framework. This emerges as a new approach in the centre-ground between pure statistical fitting and process-based (or mechanistic) models that apply

physiological thresholds (Martínez et al. 2015) or takes into account factors affecting spatial population dynamics such as species interactions, reproduction, mortality and migration rate; see the comparison in Melle et al. (2014), Robinson et al. (2011). Our proposed model has been tested by simulation for various types of theoretical curves and sampling schemes and has been applied successfully to real case studies. The performance has been compared to other regression models without shape-constraints (GLMs and GAMs with different degrees of freedom (Guisan et al. 2002, Hastie and Tibshirani 1990, McCullagh and Nelder 1989)) and to models based on climate envelopes such as “Plateau” (Brewer et al. 2016).

SC-GAMs are based on the same statistical framework as GLMs and GAMs that are commonly used to fit species distribution models (Guisan et al. 2002). According to the simulation results, in several scenarios, mainly when the range of the environmental gradient was not fully covered, “GLM” and “GAMlk” methods were not able to approximate correctly the underlying theoretical niche model. Increasing the degrees of freedom of the GAM (“GAMhk”) helped to estimate curves that were closer to the true theoretical curve; however, due to random noise, fitted curves were mostly multimodal and not concave which renders them implausible under the ecological niche theory framework. An essential challenge when modelling the relationship between species occurrence and environmental drivers is to capture the signal and to differentiate it from sampling and environmental noise (Burnham and Anderson 2003). Therefore, for all models in general, and for GAMs in particular, determining the appropriate model complexity is critical both for robust inference and for accurate prediction. Excessive flexibility can lead to overfitted models where resulting patterns can be spurious and affected by noise, and predictions based on such models can be biased and unreliable (Burnham and Anderson 2003). Many authors have favoured simpler versus more complex models (Merow et al. 2014 and references therein), suggesting that researchers should constrain the complexity of their models based on the study objective, attributes of the data, and an understanding of how these interact with the underlying biological processes. Austin (2002) suggested that complex functions produced by GAMs could be replaced by an equivalent parametric function, simpler and ecologically easier to interpret. In practice, other authors have manually changed the degrees of freedom of the smoothing functions to achieve simpler curves (e.g., Bruge et al. (2016), Brunel et al. (2018)) or even unimodal or monotonic shapes following the ecological niche theory (Chust et al. 2014). In that context, SC-GAMs automatically provide response curves in

agreement with the niche theory. In the simulations, obtained fits were closer to the underlying theoretical curves in comparison to “GLM” and “GAMk” approaches, and in scenarios where the sampling did not cover the whole environmental range, results were similar or even better than those obtained with the most flexible GAM.

SC-GAMs were also compared to “Plateau” (Brewer et al. 2016) which is a regression model based on climatic envelopes. “Plateau” can provide the correct shape with variance estimates from the hessian in a fast way. The extension to the multivariable functions is straightforward and more variables and their potential interactions can be readily incorporated. The simulations indicated that there were no differences regarding the performance in terms of agreement with the ecological niche theory. Both the “Plateau” and the SC-GAMs satisfied the concavity restrictions and estimated the maximum correctly. However, the simulation results showed that shape-constrained models were more robust across replicates, with less uncertainty in point estimation and coverage probabilities.

The two SC-GAMs implementations tested in this study present statistically sound methods that allow for robust estimation, model comparison, and prediction. However, they exhibited some differences in terms of uncertainty estimation, computing time and ease of use. The “boost” approach seemed to be more robust to the generated uncertainties and showed more stable and narrower intervals for RMSE values and for coverage probabilities. Variance estimation in this approach is performed through bootstrapping which implies a high computational cost. Alternatively, the “SCAM” approach builds on the framework of unconstrained generalised additive models (Wood 2006b), being computationally efficient (Pya and Wood 2014). In addition, it uses almost the same syntax as in `mgcv` R package which facilitates its use.

SC-GAMs provide a unified framework to deal with different types of variables in species distribution models. Direct variables and limiting factors are expected to have a unimodal shape (symmetric or not), whereas there is no theoretical expectation regarding direct variables. However, sometimes, there might be exceptions in which the realized niche is not unimodal with respect to environmental gradients (Austin 2002). In those cases, the comparison between shape-constrained and unconstrained methods could help to better disentangle the factors defining the ecological niche of the species. When modelling species distribution based only on

niche theory, results are limited by strong assumptions such as unlimited dispersal of species, and no consideration of competition processes between species, population dynamics and adaptation of the species (*sensu* population fitness).

The extent and resolution of the data are crucial to obtain an adequate characterization of the niche of a species (Peterson et al. 2011). If the range of the environmental gradient does not cover the limits of the species, the species response is truncated and determining the actual shape of the response will be difficult (Austin 2007). Thuiller et al. (2004) found that this could be especially problematic on the tails of the species response curves, yielding spurious projections. In our simulations, the performance of the shape-unconstrained methods was worse when the range of the environmental gradient was not fully observed. In most of the cases they were not able to fit concave curves, the single maximum was not found and presented high RMSE values. However, shape-constrained methods performed similarly regardless of the type of sampling. Therefore, adding the shape constraints warranted that the species distribution model was ecologically meaningful within the observed range of the environmental variable, and facilitating its subsequent use for extrapolation and prediction.

Methods have been also tested in two different real case studies. The first case study shows that shape-constrained methods can solve issues arising with the other methods, as concluded with the simulation study. Optimum SST values and tolerance ranges obtained by SC-GAMs in the presented real case study are very similar to those reported in Bernal et al. (2007). They compiled data from all the available ichthyoplankton surveys in the Northeast Atlantic and found that spawning is restricted mainly to the shelf area and in a range of temperatures between 12°C and 17°C. Stratoudakis et al. (2007) detected that spawning seasonality varies with latitude following temperature gradients. The preferred temperatures for spawning were identified between 14 and 15°C, while temperatures below 12°C and above 16°C were avoided. In the Bay of Biscay, thermal preference at the surface was found between 12°C and 15°C (Planque et al. 2007). The second case study involves the incorporation of several variables in order to find species probabilities of occurrence combining different types of variables. Results are similar to those reported in other studies that needed a manual selection of smoother parameters such as the probability of the presence of anchovy eggs along with salinity or sea surface temperature reported in Erauskin-Extramiana et al. (2019) or the estimated optima for mackerel

spawning along the salinity gradient in Brunel et al. (2018). It is shown that the framework of SC-GAMs enables us to fit both unconstrained and shape-constrained shapes for each of the included variables depending on the type and prior knowledge. It also allows us to test the shape of each predictor consistently with the expected ecological theory as suggested in Austin (2007).

We consider that proposed SC-GAMs can be readily applied for fitting distribution models and are useful tools for modelling communities of a large number of species, as they result in a good balance between the goodness of fit and agreement with the ecological niche theory. They can incorporate multiple explanatory variables with or without interaction, both shape-constrained and unconstrained, depending on the nature of the variables involved. Thus, SC-GAMs offer the possibility of investigating, for example, the effect of climate change on multiple species without requiring sophisticated and time-consuming mechanistic models that depend on detailed knowledge of vital rates and life traits for each species. Future applications of SC-GAMs in the context of ecological models could go beyond the examples shown in this work. Bivariable smooths with concavity restrictions would allow a better understanding of the interactions between environmental variables, as in Brewer et al. (2016). SC-GAMs could also be extended to include response shapes varying per grouping level as in HGAMs (Pedersen et al. 2019). In this case, several species could be modeled together including interactions between the explanatory variables and the species as a factor obtaining a common effect and different response curves for each species. Multivariate adaptive regression splines (MARS, Friedman et al. 1991) are also claimed to have a strong performance for multiresponse species distribution models (Leathwick et al. 2006). Shape constraints could be also introduced, for unimodality condition in the response curve, to obtain comparable results with SC-GAMs. SDMs can be also fitted in a Bayesian framework, allowing to incorporate prior knowledge of species ecology (Golding and Purse 2016, Paradinas et al. 2015) or prior information on response curve shapes (Fraaije et al. 2015 - Appendix3) using INLA as a tool to perform full Bayesian analysis of latent Gaussian models using Integrated Nested Laplace Approximation (Rue et al. 2014).

Stock assessment models for sardine in the Bay of Biscay

In this chapter, we describe the data and the stock assessment model that serves as a basis to provide management advice for sardine in the Bay of Biscay. In addition, we present an alternative model that will be used throughout the thesis and an ad-hoc model developed specifically for this stock. The results of the three modelling approaches are compared, while the main methodological differences and the benefits of each of them are discussed.

3.1 Introduction

Since 2017 sardine in the Bay of Biscay is assessed yearly by the ICES Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA) using an age-structured model implemented in Stock Synthesis (SS, Methot and Wetzel 2013). The methodology was first established in the Benchmark Workshop on Pelagic Stocks (WKPELA) carried out in 2017 (ICES 2017a) and it was revised in 2019 in the Inter-Benchmark Process on sardine in the Bay of Biscay (IBPSARDINE, ICES 2019b).

Stock Synthesis (SS) is a statistical age-structured population modelling framework (Methot and Wetzel 2013, Punt and Maunder 2013) that can be cast within the integrated analysis framework (Maunder and Punt 2013). It was first developed in the 1980s by Richard Methot for the assessment of northern anchovy (Methot 1986; 1989) and since then, it has been continuously evolving, becoming one of the most commonly used stock assessment models nowadays. It is extensively used

in the United States of America (Dichmont et al. 2016), but it is common also in Tuna-Regional Fisheries Management Organizations, Australia and Europe (Methot and Wetzel 2013). Species such as Western Atlantic Bluefin tuna, Pacific sardine, northern hake or Iberian sardine are currently assessed using this software. SS is designed to flexibly incorporate multiple data sources and can deal with a wide range of models, encompassing from data-limited stocks (Wetzel and Punt 2011) to complex age-structured or length-structured models distributed across multiple areas and seasonal patterns (Methot 2009). Most parameters can change over time in response to environmental and ecosystem factors.

This flexibility of SS and its ability to handle a wide range of model configurations, make it a complex program that requires advanced technical skills. Assessment for all (a4a) (Jardim et al. 2014) is a newer alternative statistical model for age-structured fish stock assessment designed to be flexible in terms of the model structure while being simple for implementation and an easy to use option. The a4a model makes use of the R's syntax for model building (Chambers and Hastie 1992) and is framed into the FLR platform (Kell et al. 2007). It was developed as part of the a4a initiative of the European Commission Joint Research Centre that aimed at providing an accessible and versatile tool that could lead to an increase in the number of stocks assessed in Europe (Jardim et al. 2014). Moreover, taking advantage of the simplicity and flexibility for models implementation, it has been used to test spatial heterogeneity of meta-populations, such as sardine in the Bay of Biscay, comparing the assessment results of a meta-population with the combined results of its components (Jardim et al. 2018).

Another alternative that allows exploring new options out of the pre-specified configurations of the above-mentioned software is building an ad-hoc model (Dichmont et al. 2016). It enables controlling and modifying model equations at any level and even integrating new data sources. The ad-hoc stock assessment model developed in this chapter mimics the official SS assessment and is explored as a more flexible and transparent option within the Bayesian inference framework.

The main objective of this work is to detail the basis of the official SS assessment model used by the Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA) in 2017 and 2018 (ICES 2017c; 2018b) and compare it to two alternative stock assessment models: one based on the a4a approach and an ad-hoc

Bayesian model. The three approaches are age-structured models and hence, they share the same underlying population dynamics model and rely on the same data. The main differences between the three approaches lie in the observation model and the estimation method.

3.2 Material and methods

3.2.1 Data for sardine in the Bay of Biscay

Data to fit stock assessment models originate from different sources. Fishery dependent data provide information on the activity of the fishing vessels, such as landings, discards, their length or age composition and fishing effort exerted by the fleets. Alternatively, additional data might be available from specific research surveys, like acoustics, trawl, egg production or tagging surveys (Gunderson 1993). The advantage of research surveys is that they can provide abundance indices independent from the commercial or recreational activity of the fleets. In addition, they provide additional information on the biology and ecology of the species. Thus, research surveys are critical ingredients for stock assessment (Hilborn 1992) and have proved to be particularly useful for the management of small pelagic fish such as sardine, due to their particular life-history traits (Barange et al. 2009).

Data available to develop an age-structured stock assessment model for sardine in the Bay of Biscay have been collected by different research institutes and have been compiled by ICES (2017a; 2018b). Data are organised according to their type (Table 3.1). In what follows, first we summarise data from the fishery consisting of total catch and catch-at-age. Then, we present data from the research surveys, namely, annual total biomass and age composition from the PELGAS survey (Doray et al. 2018), annual total abundance index based on the egg counts from the BIOMAN survey (Santos et al. 2018) and triennial spawning stock biomass (SSB) index from the application of the daily egg production method (DEPM, Lasker 1985, Parker 1980) in the BIOMAN survey. Finally, we introduce biological data such as weight-at-age in the catch and stock weight-at-age and maturity-at-age recorded in the PELGAS survey. The period ranges from 2000 to 2018. However, not all data sources include measurements for every year as summarized in Figure 3.1. The time series of catch-at-age and weight-at-age in the catch started in 2002, and they are not available in the assessment year (in this case, 2018). Additionally, the DEPM SSB estimates are

only provided every three years since 2011.

Table 3.1: Data used for the assessment of sardine in the Bay of Biscay.

| Symbol | Description | Year range | Age range |
|----------------------|---|------------------|-----------|
| \hat{C}_y^{tot} | Catch in tonnes | 2000-2018 | - |
| $\hat{C}_{a,y}$ | Catch-at-age in numbers | 2002-2017 | 0-6+ |
| $\hat{I}_{a,y}^{ac}$ | PELGAS acoustic index at-age in numbers | 2000-2018 | 1-6+ |
| \hat{I}_y^{tot} | PELGAS acoustic biomass index in tonnes | 2000-2018 | - |
| \hat{I}_y^{egg} | BIOMAN egg count index in numbers | 2000-2018 | - |
| \hat{I}_y^{depm} | BIOMAN DEPM SSB index in tonnes | 2011, 2014, 2017 | - |
| $W_{a,y}$ | Stock weight-at-age in kg | 2000-2018 | 1-6+ |
| $W_{a,y}^C$ | Catch weight-at-age in kg | 2002-2017 | 0-6+ |
| $M_{a,y}$ | Natural mortality year ⁻¹ | 2000-2018 | 0-6+ |
| $m_{a,y}$ | Maturity-at-age proportion | 2000-2018 | 0-6+ |
| PF | Proportion of fishing mortality before spawning | - | - |
| PM | Proportion of natural mortality before spawning | - | - |

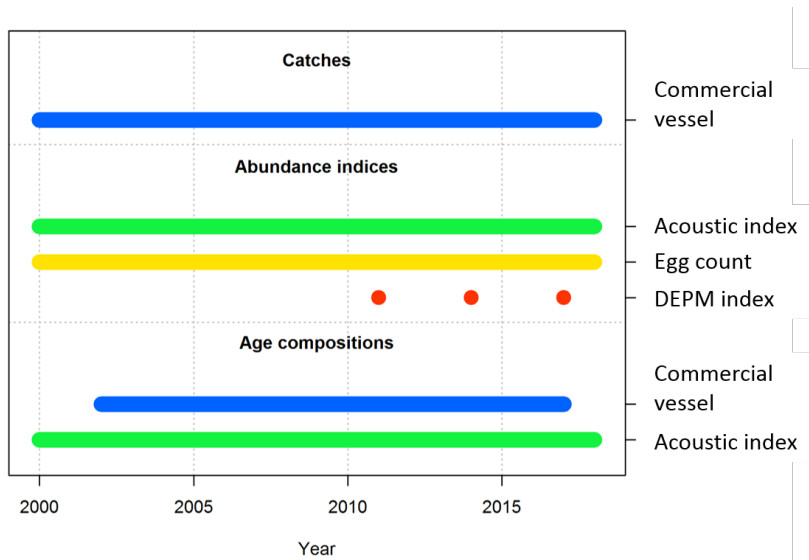


Figure 3.1: Summary of the time range coverage of the data used for the assessment of sardine in the Bay of Biscay.

3.2.1.1 Catch data

Commercial catch data for sardine in the Bay of Biscay include total catch from 2000 to 2018 and catch-at-age data from 2002 to 2018. The total catch in weight has increased from around 15 thousand tonnes at the start of the time series, to around 30 thousand tonnes during the last decade (Figure 3.2). This increase in total catch is mainly due to the increase of Spanish landings as reported in ICES (2018b). Discards are considered negligible and therefore, catches are equal to landings. Usually, catch data are not available for the assessment year, but it is necessary to include some value to fit the model until the assessment year. This allows us to estimate the status of the stock in the assessment year and provide the most up-to-date management advice. One possible option is to assume that catch in the assessment year corresponds to status quo fishing mortality (i.e. fishing mortality in the assessment year is the same as the year before $F_{2018} = F_{2017}$). Other options can be based on the total catch advised by ICES for that year or on preliminary catch data for the assessment year in comparison with the historical series. In this case, the total catch in 2018 is taken as 32,776 tonnes, which is based on the F status quo (ICES 2018b).

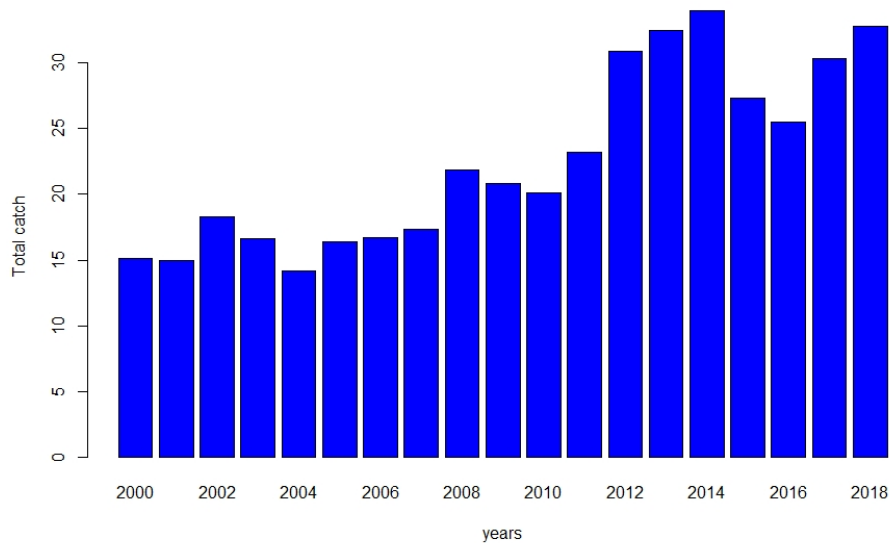


Figure 3.2: Total catch (in thousand tonnes) of sardine in the Bay of Biscay from 2000 to 2018.

Catch-at-age refers to number of caught individuals at each age group (Figure 3.3). The values range from a maximum amount of 321 million fish to less than 1 million fish depending on the year and show a general increase during the last years (see ICES 2018b for details). The most exploited age classes are 1 and 2, with a slight increase up to age 3 on the proportion of older fishes at the middle of the time series. No catch-at-age data are given to the model for the assessment year (2018) as they are not available. Therefore, the catch age composition in the assessment year will be based on some assumption for the age selection pattern of the fishery in that year.

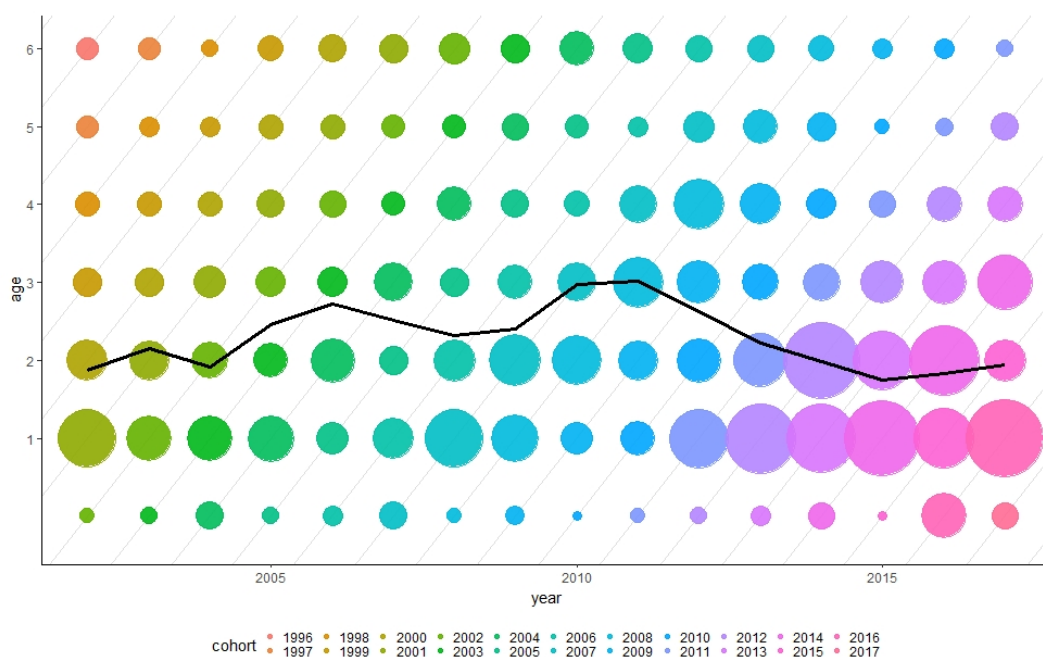


Figure 3.3: Catch-at-age data for sardine in the Bay of Biscay from 2002 to 2017. Bubble sizes are proportional to the number of individuals caught in each year (x axis) and age class (y axis). The black line represents the weighted average age of captured fish in each year.

3.2.1.2 Abundance indices

Two research surveys (PELGAS and BIOMAN) carried out yearly in spring provide the four abundance indices used for the assessment of sardine in the Bay of Biscay (see Chapter 1 for details). The PELGAS acoustic survey provides total biomass and age composition estimates, whereas BIOMAN gives an annual abundance index

and a triennial SSB estimate from the application of the DEPM.

The age-structured abundance index obtained from the acoustic PELGAS survey is given in numbers of individuals by age class (Figure 3.4). The numbers-at-age range from a maximum measurement of 8661 million fish to 8 million fish, depending on the year and the age class (see ICES 2018b for details). This information allows to track the cohorts and identify strong year classes such as 2008, 2012, 2014 or 2016, and weaker ones like the 2005 cohort. The most abundant age classes are 1 and 2, while age 0 is not recorded in these spring surveys. The average age decreases along time from age 3 at the beginning of the time series to age 1.5 in 2018.

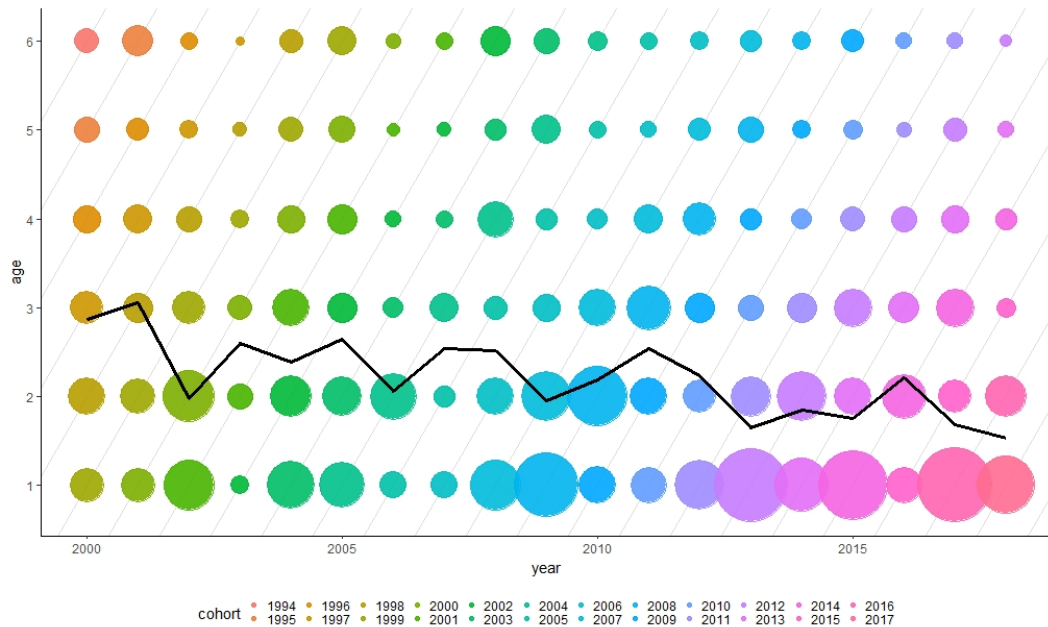


Figure 3.4: PELGAS age-structured abundance index data for sardine in the Bay of Biscay. Bubble sizes are proportional to the number of individuals in each year (x-axis) and age class (y-axis) and colours represent different cohorts. The black line represents the weighted average age for each year.

Aggregated abundance indices consist of the total biomass index from PELGAS and the egg count and DEPM SSB indices from BIOMAN. PELGAS total biomass index and the BIOMAN SSB index from the DEPM are given in thousand tonnes while the egg count is given in billions of eggs (Figure 3.5). None of the time series

shows any clear contrasting trend. PELGAS highest biomass values correspond to 2002 and 2004 with a new peak in 2017, while BIOMAN egg count shows a peak in 2005, followed by 2010 and 2008 estimates. In both cases 2003 and 2007 present among the lowest values of each times series. Regarding the SSB estimates from the DEPM, there are only three indices available, from which 2014 presents the highest SSB value.

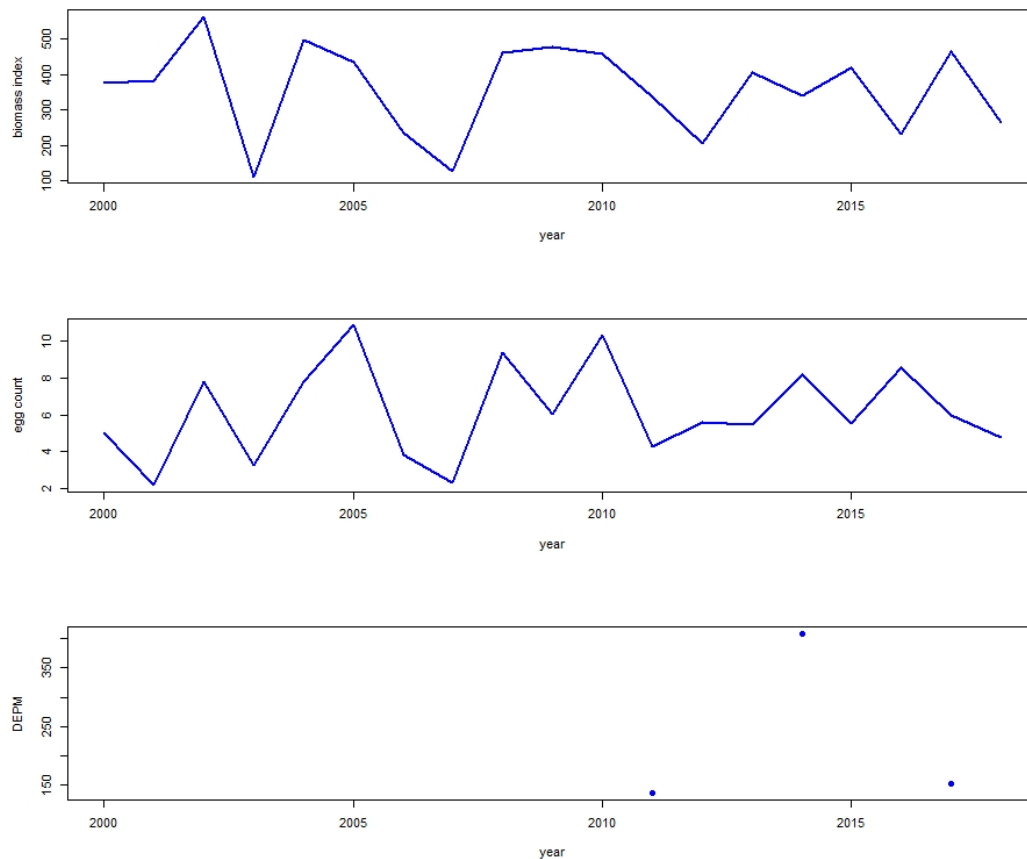


Figure 3.5: Aggregated abundance indices for sardine in the Bay of Biscay. From top to bottom: PELGAS biomass index (thousand tonnes), BIOMAN egg count index (billion of eggs) and BIOMAN DEPM SSB index (thousand tonnes).

3.2.1.3 Biological data

Weights-at-age are obtained both from the monitoring of the fishery and from the PELGAS acoustic survey. The former is used in the stock assessment model as

the catch weight-at-age, while the latter represents the stock weight-at-age at the survey time. Weights are given in kg with the highest values around 0.12kg for the oldest fishes (Figure 3.6). Both catch and stock weights show a decreasing trend along time for all the age classes, with higher values at the beginning of the time series and lower weights in the most recent years as studied in Véron et al. (2020). Weights-at-age for age 0 in the survey are set to 0 given that this age class is not observed.

Maturity-at-age data are also obtained from the PELGAS survey and represent the proportion of mature individuals found in each year and age class. All age 0 individuals are non-mature while the proportion of mature individuals at age 2 and older is almost 1. Maturity of age 1 individuals shows more variability, ranging between 0.4 and 0.8 with the largest value in the year 2006 and lowest in 2012 (Figure 3.7).

The annual rate of natural mortality constitutes one of the crucial and most influential processes in stock assessment models (Maunder and Piner 2015). However, usually, there are no direct estimates available and alternative indirect methods, such as those described by Kenchington (2014) and Then et al. (2015), must be used. Natural mortality for sardine in the Bay of Biscay was established by ICES (2017a). First, several natural mortality-at-age vectors were computed based on the models by Gislason et al. (2010) and Lorenzen (1996) and they were re-scaled by different values to obtain a range of potential options. Then, the best estimates of natural mortality-at-age were selected based on the lowest AIC value of several stock assessment fits under different natural mortality scenarios. The final time-invariant vector of annual rates of natural mortality for ages 0 to 6+ corresponded to the Gislason’s expected natural mortality-at-age vector (Gislason et al. 2010) multiplied by 0.9 (Table 3.2).

Table 3.2: Annual rate of natural mortality-at-age of sardine in the Bay of Biscay.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6+ |
|-----------------------|--------|-------|-------|-------|-------|-------|-------|
| $M_a(\text{yr}^{-1})$ | 1.0171 | 0.692 | 0.546 | 0.475 | 0.435 | 0.412 | 0.400 |

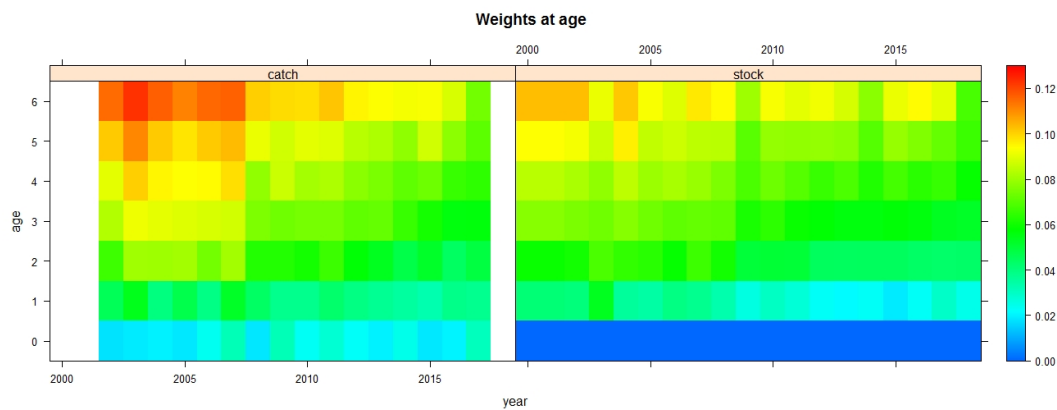


Figure 3.6: Weight-at-age in the catch (in the left) and weight-at-age in the stock obtained from PELGAS survey at spawning time (in the right) for sardine in the Bay of Biscay. Cell colours indicate the weight value (in kg) for each year (x-axis) and age class (y-axis) according to the colour scale in the right.

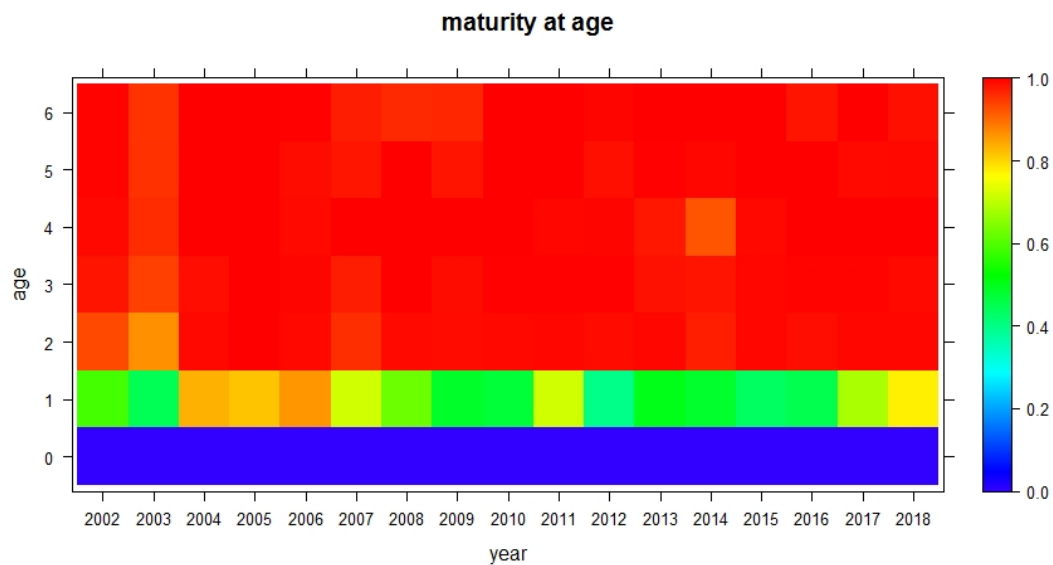


Figure 3.7: Maturity-at-age for sardine in the Bay of Biscay, obtained from the PELGAS survey. Cell colours represent the proportion of mature individuals at each year (x axis) and age class (y axis) according to the colour scale in the right.

3.2.2 Modelisation for sardine in the Bay of Biscay

In this section, we first describe the common elements of the three modelling approaches, with special emphasis on the foundation of age-structured models, and then, we cover the specificities of each approach.

3.2.2.1 Common elements

Let $N_{a,y}$ be the total number of individuals at-age a at the beginning of year y . The index $y = 1, \dots, Y$ represents the assessment years and the index $a = a_r, \dots, A-1, A$ the age groups, being a_r the age at which recruitment occurs (i.e. age at which new individuals enter into the population) and A the age of the plus-group (i.e. the age group including individuals of age A and older). This plus-group age A can be based either on the ability to predict the age of an individual or on the age above which very few individuals are observed. For sardine in the Bay of Biscay, a_r is established at 0 and the plus-group is set at age 6+.

The transition along time of the number of individuals at each age group is given by the survival equation. For $y \geq 2$:

$$N_{a,y} = N_{a-1,y-1} e^{-Z_{a-1,y-1}}, \text{ for } a = 1, \dots, 5 \quad (3.1)$$

and

$$N_{6+,y} = N_{5,y-1} e^{-Z_{5,y-1}} + N_{6+,y-1} e^{-Z_{6+,y-1}}, \quad (3.2)$$

where $Z_{a,y}$ is the total annual mortality rate of individuals of age a in year y that can be decomposed as the sum of the fishing mortality $F_{a,y}$ and the natural mortality $M_{a,y}$ as follows:

$$Z_{a,y} = F_{a,y} + M_{a,y}. \quad (3.3)$$

For sardine in the Bay of Biscay annual natural mortality-at-age is assumed to be time-invariant (i.e. $M_{a,y} = M_a$) and equal to the values given in Table 3.2. In turn, fishing mortality by age and year ($F_{a,y}$) is modelled assuming a separable structure:

$$F_{a,y} = f_y s_a, \quad (3.4)$$

where f_y is the year component that refers to the fishing intensity in year y and s_a is the age component that represents the fishery selection pattern at age a . For sardine, age classes 3 and older are assumed to be fully selected by the fishery. So, the shape of the age selection pattern is assumed to increase from age 0 to 2 and is kept constant from age 3 to 6+ (ICES 2017a).

From the number of individuals in the population $N_{a,y}$, the Baranov's catch equation (Baranov 1918) defines the total number of captured individuals at age a in year y , $C_{a,y}$, as the proportion of dead individuals attributed to fishing:

$$C_{a,y} = N_{a,y}(1 - e^{-Z_{a,y}}) F_{a,y}/Z_{a,y}. \quad (3.5)$$

Then, the catch proportion-at-age a in year y , $CP_{a,y}$, can be calculated as:

$$CP_{a,y} = C_{a,y} / \sum_{a=0}^{6+} C_{a,y}, \quad (3.6)$$

and the total catch in weight in year y , C_y^{tot} , is defined as:

$$C_y^{tot} = \sum_{a=0}^{6+} C_{a,y} W_{a,y}^c,$$

where $W_{a,y}^c$ is the weight-at-age a in year y of the caught fish.

From these basic population dynamic equations, other derived quantities of interest are Total Biomass ($B_{y,t}$) and Biomass of age 1 and older individuals ($B1_{y,t}$) at time instant t of year y :

$$B_{y,t} = \sum_{a=0}^{6+} N_{a,y} e^{(-Z_{a,y}t)} W_{a,y,t}, \quad (3.7)$$

$$B1_{y,t} = \sum_{a=1}^{6+} N_{a,y} e^{(-Z_{a,y}t)} W_{a,y,t}, \quad (3.8)$$

where $W_{a,y,t}$ is the weight-at-age a of the population at time instant t of year y . Similarly, Spawning Stock Biomass (SSB) is the total weight of all individuals in a

fish stock that is capable of reproducing and is computed as:

$$SSB_{y,t_{spw}} = \sum_{a=0}^{6+} N_{a,y} e^{-(F_{a,y} PF + M_{a,y} PM)} W_{a,y,t_{spw}} m_{a,y}, \quad (3.9)$$

where t_{spw} is the breeding time instant, PF is the proportion of fishing mortality before spawning, PM is the proportion of natural mortality before spawning (both set to 0 in this case study) and $m_{a,y}$ is the proportion of mature individuals at-age a in year y .

The number of individuals entering the population at the start of each year y is named recruitment and is denoted by R_y . In this case study, new individuals enter the population at age zero ($a_r = 0$), so recruits are age 0 individuals at the beginning of the year ($R_y = N_{0,y}$). Recruitment can be estimated as individual parameters for each year or can be modelled using a stock-recruitment relationship. Stock-Recruitment relationships (S-R) associate the parental fish stock, usually the spawning stock biomass, and the subsequent number of recruits produced by that spawning (Hilborn and Walters 1992, Quinn and Deriso 1999). This idea is common to all self renewable populations. Two of the most common S-R models are the Beverton-Holt (Beverton and Holt 1957) and the Ricker model (Ricker 1954), but more general or alternative models have been proposed by Deriso (1980), Schnute (1985) or Barrowman and Myers (2000) among others. See Needle (2001) for a summary on S-R models and Subbey et al. (2014) or Sharma et al. (2019) for recent discussions and developments on S-R models.

Given the population dynamics described above, the abundance indices are assumed to be linearly related to the available population at survey time. The numbers-at-age from the acoustic survey PELGAS at age a in year y , $I_{a,y}^{ac}$, is modelled through the following equation:

$$I_{a,y}^{ac} = Q_{ac} q_a N_{a,y} e^{-Z_{a,y} t_{ac}}, \quad (3.10)$$

where t_{ac} represents the timing of the acoustic survey, Q_{ac} is the overall catchability coefficient of the acoustic survey and $q_a \in [0, 1]$ is the survey selectivity at age a . For sardine in the Bay of Biscay, q_a is kept fixed and equal to 1 from ages 2 to 5 (ICES 2017a). Given that age 0 is not observed in this survey, only the survey selectivities

for ages 1 and 6+ must be estimated.

From the above, the proportion-at-age a in year y from the acoustic survey is given by:

$$IP_{a,y} = I_{a,y}^{ac} / \sum_{a=0}^{6+} I_{a,y}^{ac}. \quad (3.11)$$

Similarly, the aggregated acoustic biomass index from the PELGAS acoustic survey I_y^{tot} is assumed to be linearly related to the population biomass at survey time:

$$I_y^{ac} = Q_{tot} B_{y,t_{ac}}, \quad (3.12)$$

and the annual egg abundance index I_y^{egg} and the triennial DEPM index I_y^{depm} from the BIOMAN survey are linearly related to the population SSB at their respective time instants:

$$I_y^{egg} = Q_{egg} SSB_{y,t_{egg}}, \quad (3.13)$$

$$I_y^{depm} = Q_{depm} SSB_{y,t_{depm}}, \quad (3.14)$$

where Q_{tot} , Q_{egg} and Q_{depm} are the catchability parameters for each of the aggregated indices. Given that all the surveys are carried out during spring approximately at the same time, it is further assumed that all surveys time instants are equal to 0.4 ($t_{ac} = t_{egg} = t_{depm} = 0.4$).

Model structure and statistical distribution particularities for recruitment, first-year population numbers, fishing mortality parameters and catch and indices observation equations are detailed in the following sections for each modelling approach. The hat symbol ($\hat{}$) will be used to represent observed data and distinguish them from the model quantities.

3.2.2.2 SS approach

Stock synthesis model specifications as setting up by WKPELA (ICES 2017a) for the Bay of Biscay sardine case study are described below.

Population dynamics model:

Annual **recruitments** (R_y) are estimated yearly and are modelled as lognormal (LN) deviations from a constant mean value in log-scale (R_m) with a standard deviation of $\sigma_R = 0.55$.

$$R_y \sim LN(\ln(R_m), \sigma_R), \text{ for } y = 1, \dots, Y. \quad (3.15)$$

Since there are no data providing information on the recruitment in the assessment year, recruitment in that year is assumed to be the historical geometric mean ($R_{2018} = R_{GM}$).

The **initial population numbers** (i.e. population numbers-at-age in the first year of the assessment, 2000) are calculated by estimating an initial equilibrium population modified by the age composition data in the first year of the assessment (Methot and Wetzel, 2013). The initial equilibrium population is derived assuming an initial catch of 13 000 tons, which is the average catch between the years 1990 and 1999.

Fishing mortality is estimated using the hybrid method described by Methot and Wetzel (2013). This method does a Pope's approximation (Pope 1972) to provide initial values and then iteratively adjusts the continuous F values to closely approximate the observed total catch.

In SS fishery age selectivity (s_a) is such that the parameter for each age is modelled as a random walk from the previous age (Methot 2009). Selectivity-at-age 0 is not estimated and is used as the reference age against which subsequent changes occur. The rest selectivity estimates are computed from a set of internally defined parameters (positive values are associated with increasing selectivity between ages, and negative values are associated with decreasing selectivity between ages) with the property that the maximum selectivity is equal to 1. In order to mimic the sardine dome-shaped selectivity pattern with an increasing or flat shape from ages 0 to 3,

selectivity parameters are estimated for ages 1, 2, and 3, while selectivity at ages 3 to 5 in the fishery are bound, meaning that parameters for ages 4 and 5 are not estimated but assumed to be equal to the parameter estimated for age 3.

Observation model:

SS assumes that the absolute level of **total catch** is known precisely so that the annual fishing mortality values are tuned to approximately match the total catch biomass by year (standard deviation in log-scale $\sigma_C = 0.05$):

$$\widehat{C}_y^{tot} \sim LN(C_y^{tot}, \sigma_C = 0.05), y = 1, \dots, Y. \quad (3.16)$$

The observed fishery age composition, $\widehat{CP}_{1...6+,y}$, derived from **catch-at-age** data is assumed to follow a multinomial distribution depending on the population age composition as follows:

$$\widehat{CP}_{1...6+,y} \sim Multinomial(n_C, CP_{1,y}, \dots, CP_{6+,y}), y = 1, \dots, Y. \quad (3.17)$$

The sample size (n_C) sets the precision of the age composition data. If the multinomial error model is strictly correct and the observations in a sample are independent, n_C should correspond to the actual number of fish in the age samples. However, age samples are usually autocorrelated and it is difficult to obtain reliable values for n_C . Although the levels of age sampling for the sardine stock are high in both the fishery and the acoustic survey, approximate sample sizes are calculated taking into account the harmonic mean of expected sample sizes provided by the model. The sample size for fishery age compositions is set equal to 54.

Similarly, the **acoustic survey age composition** is assumed to follow a multinomial distribution:

$$\widehat{IP}_{1...6+,y} \sim Multinomial(n_I, IP_{1,y}, \dots, IP_{6+,y}), y = 1, \dots, Y, \quad (3.18)$$

where $\widehat{IP}_{1...6+,y}$ is the observed age composition derived from the acoustic survey numbers-at-age. The sample size for acoustic survey age composition (n_I) is fixed to 40, following the same process as for the fishery age composition described above. Similarly to the fishery selectivity, selectivity for this abundance at-age index (q_a) is

such that the parameter for each age is estimated as a random walk from the previous age. However, this applies only to ages 1, 2, and 6+ in the survey. Selectivity at ages 2 to 5 years in the survey are bound, meaning that parameters for ages 3, 4 and 5 are not estimated but assumed to be equal to the parameter estimated for age 2. Selectivity-at-age 0 is not estimated and is used as the reference age against which subsequent changes occur. This selectivity parameterization mimics a dome-shaped pattern, with an increasing or flat shape from ages 0 to 2 and a declining or flat shape at the 6+ group.

The three **aggregated abundance indices** by year are assumed to follow a log-normal distribution centred at the modelled abundance index in log scale with a standard deviation coming from an iterative process where these values are updated iteratively until approximate convergence of the root mean squared error of each aggregated indices (McAllister and Ianelli 1997).

In particular, the **PELGAS biomass index** \hat{I}_y^{ac} is lognormally distributed:

$$\hat{I}_y^{ac} \sim LN(\ln(I_y^{ac}), \sigma_{I_y^{ac}}), y = 1, \dots, Y, \quad (3.19)$$

where the standard deviations ($\sigma_{I_y^{tot}}$) are first set to the values estimated by the survey and are then increased after the tuning iterative process. The resulting values range between 0.3 and 0.7 depending on the year.

The **egg count index** \hat{I}_y^{egg} is modelled as:

$$\hat{I}_y^{egg} \sim LN(\ln(I_y^{egg}), \sigma_{I_y^{egg}}), y = 1, \dots, Y, \quad (3.20)$$

where the standard deviation $\sigma_{I_y^{egg}} = 0.448$ is constant along years and is derived from the iterative process described above.

Similarly, the **DEPM index** \hat{I}_y^{dep} follows a lognormal distribution:

$$\hat{I}_y^{dep} \sim LN(\ln(I_y^{dep}), \sigma_{I_y^{dep}}), y = 1, \dots, Y, \quad (3.21)$$

where the standard deviation is constant along years and is set as $\sigma_{I_y^{dep}} = 0.567$.

Estimation method:

The version of SS used in this work is 3.24 (Methot 2009). The model and a graphical user interface are available at the NOAA Fisheries Stock Assessment Toolbox website: <http://nft.nefsc.noaa.gov/>. Although the first SS version was coded in Fortran, subsequent versions have been programmed and compiled using Automatic Differentiation Model Builder (ADMB, Fournier et al. (2012)). This allows SS to benefit from rapid model convergence and uncertainty estimation through inverse Hessian MCMC approaches are also available (Methot and Wetzel 2013). Model inspection and figures showed in this section are generated using the `r4ss` package (Taylor et al. 2019).

3.2.2.3 a4a approach

The a4a model for sardine in the Bay of Biscay in this work is set to mimic the official SS assessment.

In general, the a4a model structure is defined via five submodels: a model for recruitment, a model for the initial age-structured population, a model for the fishing mortality-at-age, a model for the abundance indices' catchabilities and a model for the variances of the catch-at-age and abundance indices observation equations that are also estimated within the model. All these submodels have linear forms that have to be defined. Specifications of each of these submodels for the sardine case study are described below.

Population dynamics model:

Recruitment R_y is estimated as single and independent parameters for each year y . Similarly, **initial population** numbers-at-age $N_{a,1}$ are estimated as single and independent parameters for each age a .

In the separable structure of the **fishing mortality**, the last value for the year component is fixed to be equal to the previous year ($f_{2018} = f_{2017}$), otherwise this parameter could not be correctly estimated given that there is no catch-at-age information in the last year. Fishery age selectivity (s_a) is restricted so that $s_3 = s_4 = s_5 = s_{6+}$, in order to get the same selectivity shape as in SS.

Observation model:

The a4a approach does not include any observation equation for the **total catch**. **Catches-at-age** are assumed to follow lognormal distributions independent by age and year:

$$\widehat{C}_{a,y} \sim LN(\ln(C_{a,y}), \sigma_C), y = 1, \dots, Y, a = 0, \dots, 6+. \quad (3.22)$$

Although the standard deviation of the observation equation for catch-at-age is one of the submodels of a4a that can be specified for each case study, for sardine the default option of constant variance across ages and years are used (i.e. σ_C not dependent on age or year). This standard deviation σ_C is also estimated in the model fit.

The **acoustic survey age structured index** is also assumed to follow a lognormal distribution:

$$\widehat{I}_{a,y}^{ac} \sim LN(\ln(I_{a,y}^{ac}), \sigma_{I^{ac}}), y = 1, \dots, Y, a = 1, \dots, 6+, \quad (3.23)$$

where the standard deviation $\sigma_{I^{ac}}$ is assumed to be constant across all years and ages and will be estimated by the model. The survey age selectivity q_a is restricted so that $q_2 = q_3 = q_4 = q_5$, to get the same selectivity shape as in SS.

Besides age-disaggregated indices, the a4a package only allows to include total biomass indices or biomass indices over a given age range. Thus, the egg abundance and DEPM **aggregated indices** cannot be related to SSB. As an approximation, and given that age 0 is not present in these surveys, the three aggregated indices are modelled as B1+ indices:

$$\widehat{I}_y^{tot} \sim LN(\ln(B1_{y,tac} Q_{tot}), \sigma_{I^{tot}}), y = 1, \dots, Y, \quad (3.24)$$

$$\widehat{I}_y^{egg} \sim LN(\ln(B1_{y,t_{egg}} Q_{egg}), \sigma_{I^{egg}}), y = 1, \dots, Y, \quad (3.25)$$

$$\widehat{I}_y^{dep} \sim LN(\ln(B1_{y,t_{dep}} Q_{dep}), \sigma_{I^{dep}}), y = 1, \dots, Y. \quad (3.26)$$

where $\sigma_{I_{tot}}$, $\sigma_{I_{egg}}$ and $\sigma_{I_{depm}}$ are standard deviations and Q_{tot} , Q_{egg} and Q_{depm} are survey catchabilities for each index. Both standard deviations and catchabilities are assumed constant across all years and will be estimated by the model.

The R commands used to define the fishing mortality, catchabilities and variance submodels within the a4a approach can be found in Annex D.

Estimation method:

Parameter estimation in a4a is performed via optimization of the likelihood function and conducted using automatic differentiation through ADMB (Fournier et al. 2012). Thus, Markov Chain Monte Carlo (MCMC, Gilks et al. 1996) methods can be also used inherited from ADMB capabilities. Uncertainty estimates are based on the variance-covariance matrix computed from the inverse Hessian returned by ADMB, and could be also obtained from the MCMC methods available in ADMB. The **FLa4a** (Jardim 2017) R package has been used to fit the model, perform diagnostics and generate output figures.

3.2.2.4 Bayesian approach

In contrast with the two previous approaches where specific stock assessment software have been used, in this section, we present an ad-hoc Bayesian counterpart that mimics the official SS assessment. The model is developed using JAGS (Plummer 2003).

In this approach standard deviations of lognormal distributions are represented as the square root of one divided by the precision ($\sigma = \sqrt{1/\tau}$), given that the selected tool, JAGS, works with precisions.

Population dynamics model:

Annual **recruitments** (R_y) are lognormally distributed with mean μ_{rec} and precision τ_{rec} :

$$R_y \sim LN(\mu_{rec}, \sqrt{1/\tau_{rec}}), y = 1, \dots, Y. \quad (3.27)$$

The year component of the **fishing mortality** (f_y) is assumed to be lognormally distributed with mean μ_f and precision τ_f :

$$f_y \sim LN(\mu_f, \sqrt{1/\tau_f}), y = 1, \dots, Y. \quad (3.28)$$

The age selectivity of the fishery for age 3 and older is fixed at 1 ($s_a = 1$ for $a \in \{3, 4, 5, 6+\}$) and only selectivity for ages 0, 1 and 2 is estimated.

Observation model:

Catch-at-age in numbers is assumed to follow a lognormal distribution, centred at $C_{a,y}$ coming from the Baranov equation and precision τ_C :

$$\widehat{C}_{a,y} \sim LN(\ln(C_{a,y}), \sqrt{1/\tau_C}), y = 1, \dots, Y, a = 0, \dots, 6+. \quad (3.29)$$

Total catch in biomass is assumed to follow a lognormal distribution, centred at $\ln(C_y^{tot})$ and precision $\tau_{C^{tot}}$:

$$\widehat{C}_y^{tot} \sim LN(\ln(C_y^{tot}), \sqrt{1/\tau_{C^{tot}}}), y = 1, \dots, Y. \quad (3.30)$$

Acoustic survey indices, both age structured and biomass aggregated indices, are assumed to follow lognormal indices with mean values modelled as indicated in equation (3.10) and (3.12) and setting $q_a = 1$ for $a \in \{2, 3, 4, 5\}$:

$$\widehat{I}_{a,y}^{ac} \sim LN(\ln(I_{a,y}^{ac}), \sqrt{1/\tau_{I^{ac}}}), y = 1, \dots, Y, a = 1, \dots, 6+, \quad (3.31)$$

$$\widehat{I}_y^{tot} \sim LN(\ln(I_y^{tot}), \sqrt{1/\tau_{I^{tot}}}), y = 1, \dots, Y. \quad (3.32)$$

Egg count index is modelled through a lognormal distribution, with the mean defined as in equation (3.13):

$$\widehat{I}_y^{egg} \sim LN(\ln(I_y^{egg}), \sqrt{1/\tau_{I^{egg}}}), y = 1, \dots, Y. \quad (3.33)$$

DEPM index is modelled through a lognormal distribution, with the mean defined as in equation (3.14) :

$$\widehat{I}_y^{depm} \sim LN(\ln(I_y^{depm}), \sqrt{1/\tau_{I^{depm}}}), y \in \{2011, 2014, 2017\}. \quad (3.34)$$

Prior distributions:

The prior distributions for all the unknowns are established as follows:

$$\begin{aligned}
\mu_{rec} &\sim LN(\nu_{rec}, \sqrt{1/\tau_{\nu_{rec}}}), \\
\tau_{rec} &\sim Gamma(\alpha_{rec}, \beta_{rec}), \\
N_{a,1} &\sim LN(\mu_{N1}, \sqrt{1/\tau_{N1}}), a = 0, \dots, 6+, \\
\mu_f &\sim LN(\nu_f, \sqrt{1/\tau_{\nu_f}}), \\
\tau_f &\sim Gamma(\alpha_f, \beta_f), \\
s_a &\sim LN(\mu_s, \sqrt{1/\tau_s}), a \in \{0, 1, 2\}, \\
Q_{ac} &\sim LN(\mu_{Q_{ac}}, \sqrt{1/\tau_{Q_{ac}}}), \\
Q_{tot} &\sim LN(\mu_{Q_{tot}}, \sqrt{1/\tau_{Q_{tot}}}), \\
q_a &\sim LN(\mu_q, \tau_q), a \in \{1, 6+\}, \\
Q_{egg} &\sim LN(\mu_{Q_{egg}}, \tau_{Q_{egg}}), \\
Q_{depn} &\sim LN(\mu_{Q_{depn}}, \tau_{Q_{depn}}), \\
\tau_C &\sim Gamma(\alpha_C, \beta_C), \\
\tau_{C^{tot}} &\sim Gamma(\alpha_{C^{tot}}, \beta_{C^{tot}}), \\
\tau_{Iac} &\sim Gamma(\alpha_{Iac}, \beta_{Iac}), \\
\tau_{I^{tot}} &\sim Gamma(\alpha_{I^{tot}}, \beta_{I^{tot}}), \\
\tau_{Iegg} &\sim Gamma(\alpha_{Iegg}, \beta_{Iegg}), \\
\tau_{Idepn} &\sim Gamma(\alpha_{Idepn}, \beta_{Idepn}),
\end{aligned}$$

The hyperparameters of the prior distributions are selected to be centred at values considered reasonable and with substantial dispersion, reflecting little prior information. The precision hyperparameters τ of the lognormal priors are set equal to 1, resulting in a coefficient of variation (CV) of 1 for all the unknown parameters. The shape α and rate β hyperparameters are set at 16 and 1 respectively, which lead to prior distributions with a CV of 0.25 for the precisions of the observation equations. The exception is the gamma prior distribution set for the total catch that is established with parameters that gave a CV of 0.05 for the total catch, emulating the high precision assumption on the total catch in the SS approach. All the prior distributions, their corresponding hyperparameters and the 95% central probability

intervals are summarized in Table 3.3.

Table 3.3: Prior distributions, hyperparameters and 95% central probability intervals for the Bayesian approach for sardine in the Bay of Biscay.

| Distribution | Hyperparameters | | 95% Prob. Int. |
|--|-------------------------|-------------------------|-----------------|
| $\mu_{rec} \sim LN(\nu_{rec}, \sqrt{1/\tau_{\nu_{rec}}})$ | $\nu_{rec} = 13$ | $\tau_{\nu_{rec}} = 1$ | (62320,3140724) |
| $\tau_{rec} \sim Gamma(\alpha_{rec}, \beta_{rec})$ | $\alpha_{rec} = 16$ | $\beta_{rec} = 1$ | (9.15 ,24.74) |
| $N_{a,1} \sim LN(\mu_{N1}, \sqrt{1/\tau_{N1}})$ | $\mu_{N1} = 13$ | $\tau_{N1} = 1$ | (62320,3140724) |
| $\mu_f \sim LN(\nu_f, \sqrt{1/\tau_{\nu_f}})$ | $\nu_{fy} = \ln(0.1)$ | τ_{ν_f} | (0.01,0.71) |
| $\tau_f \sim Gamma(\alpha_f, \beta_f)$ | $\alpha_{fy} = 16$ | $\beta_{fy} = 1$ | (9.15 ,24.74) |
| $s_a \sim LN(\mu_s, \sqrt{1/\tau_s}), a \in \{0, 1, 2\}$ | $\mu_s = 0$ | $\tau_s = 1$ | (0.14 ,7.10) |
| $\tau_C \sim Gamma(\alpha_C, \beta_C)$ | $\alpha_C = 16$ | $\beta_C = 1$ | (285.8,533.1) |
| $\tau_{C^{tot}} \sim Gamma(\alpha_{C^{tot}}, \beta_{C^{tot}})$ | $\alpha_{C^{tot}} = 40$ | $\beta_{C^{tot}} = 0.1$ | (9.15,24.74) |
| $\tau_{Iac} \sim Gamma(\alpha_{Iac}, \beta_{Iac})$ | $\alpha_{Iac} = 16$ | $\beta_{Iac} = 1$ | (9.15,24.74) |
| $\tau_{I^{tot}} \sim Gamma(\alpha_{I^{tot}}, \beta_{I^{tot}})$ | $\alpha_{I^{tot}} = 16$ | $\beta_{I^{tot}} = 1$ | (9.15,24.74) |
| $\tau_{Iegg} \sim Gamma(\alpha_{Iegg}, \beta_{Iegg})$ | $\alpha_{Iegg} = 16$ | $\beta_{Iegg} = 1$ | (9.15,24.74) |
| $\tau_{Idep} \sim Gamma(\alpha_{Idep}, \beta_{Idep})$ | $\alpha_{Idep} = 16$ | $\beta_{Idep} = 1$ | (9.15,24.74) |
| $Q_{ac} \sim LN(\mu_{Q_{ac}}, \sqrt{1/\tau_{Q_{ac}}})$ | $\mu_{Q_{ac}} = 0$ | $\tau_{Q_{ac}} = 1$ | (0.14,7.10) |
| $Q_{tot} \sim LN(\mu_{Q_{tot}}, \sqrt{1/\tau_{Q_{tot}}})$ | $\mu_{Q_{tot}} = 0$ | $\tau_{Q_{tot}} = 1$ | (0.14,7.10) |
| $q_a \sim LN(\mu_q, \tau_q), a \in \{1, 6+\}$ | $\mu_q = 0$ | $\tau_q = 1$ | (0.14 ,7.10) |
| $Q_{egg} \sim LN(\mu_{Q_{egg}}, \tau_{Q_{egg}})$ | $\mu_{Q_{egg}} = 0$ | $\tau_{Q_{egg}} = 1$ | (0.14,7.10) |
| $Q_{dep} \sim LN(\mu_{Q_{dep}}, \tau_{Q_{dep}})$ | $\mu_{Q_{dep}} = 0$ | $\tau_{Q_{dep}} = 1$ | (0.14,7.10) |

Estimation method:

The model is implemented in JAGS (Plummer 2003) and is run from R (R Core Team 2018) using the package R2jags (Su et al. 2015). 150000 iterations are run for each of the 2 chains with a thinning of 100 and a burn-in period of 50000. Convergence has been tested by visual inspection and the Geweke test (Z scores to test the equality of means between first and last part of a chain), the Gelman and Rubin test (tests within and between chain variance) and the Heidelberg and Welch stationarity test (Cramer-von Mises statistic to accept or reject that the Markov chain is from a stationary distribution) has been also conducted. These convergence diagnostics are checked using the coda package (Plummer et al. 2019). The developed JAGS code for this approach can be found in Annex D.

3.3 Results

3.3.1 SS approach

SS approach results presented in this section correspond to the official stock assessment estimates provided in the ICES WGHANSA 2018 annual report (ICES 2018b) for Bay of Biscay sardine.

The SS model converged adequately and the model estimates were considered reliable. The final convergence indicator that measures the change in the log-likelihood was equal to $3.08579e-6$, below the recommended convergence criterion of < 0.0001 . The contributions of each model component to the total likelihood are shown in Table 3.4.

Table 3.4: Contributions of each model component to the total negative log likelihood for the SS approach.

| | -log likelihood |
|----------------------|-----------------|
| TOTAL | 67.63 |
| Catch | 48e-10 |
| Equil_catch | 51e-6 |
| Survey | -13.23 |
| Age_comp | 74.75 |
| Recruitment | 6.09 |
| Forecast_Recruitment | 0.00 |
| Parm_priors | 0.00 |
| Parm_softbounds | 0.01 |
| Parm_devs | 0.00 |

The most meaningful assessment outputs, namely, time series of recruitment, SSB, total catch and fishing mortality (averaged from age 2 to 5) are shown in Table 3.5. Spawning stock biomass was estimated around 122 thousand tonnes in the first year 2000 and it increased to a maximum of more than 170 thousand tonnes in the year 2005. After various years of successive decreases, SSB increased slightly in the last two years (Figure 3.8). Estimated recruitment is variable over the years and has been oscillating between 2200 million individuals in 2005 and 8700 in 2008. The probability intervals of recruitment in the last years are larger representing the larger

uncertainty in the last years of the assessment because the cohorts have not been observed along their whole trajectories. The fishing mortality reflects two periods of different exploitation levels: until 2011 the time series was below 0.25, while in the last 7 years it has increased being around 0.4. (Table 3.5).

Table 3.5: Annual recruitment (million of individuals), SSB (in tonnes), total catch (in tonnes) and fishing mortality (averaged for ages 2 to 5) estimated using the SS approach.

| | Recruitment | SSB | Catch | F (2-5) |
|------|-------------|--------|-------|---------|
| 2000 | 4570 | 122462 | 15097 | 0.144 |
| 2001 | 5475 | 116007 | 15005 | 0.145 |
| 2002 | 3591 | 143372 | 18277 | 0.165 |
| 2003 | 3914 | 135227 | 16607 | 0.132 |
| 2004 | 7338 | 148814 | 14197 | 0.124 |
| 2005 | 2216 | 173287 | 16360 | 0.123 |
| 2006 | 3556 | 154630 | 16741 | 0.135 |
| 2007 | 7139 | 134657 | 17323 | 0.143 |
| 2008 | 8673 | 143645 | 21821 | 0.201 |
| 2009 | 3382 | 136855 | 20855 | 0.166 |
| 2010 | 2482 | 150824 | 20127 | 0.162 |
| 2011 | 4220 | 121343 | 23208 | 0.216 |
| 2012 | 7496 | 88495 | 30900 | 0.382 |
| 2013 | 5226 | 90798 | 32489 | 0.407 |
| 2014 | 7307 | 93923 | 33943 | 0.466 |
| 2015 | 2670 | 88921 | 27284 | 0.353 |
| 2016 | 7834 | 89522 | 25498 | 0.402 |
| 2017 | 8282 | 109062 | 30318 | 0.410 |
| 2018 | | 125246 | | 0.388 |

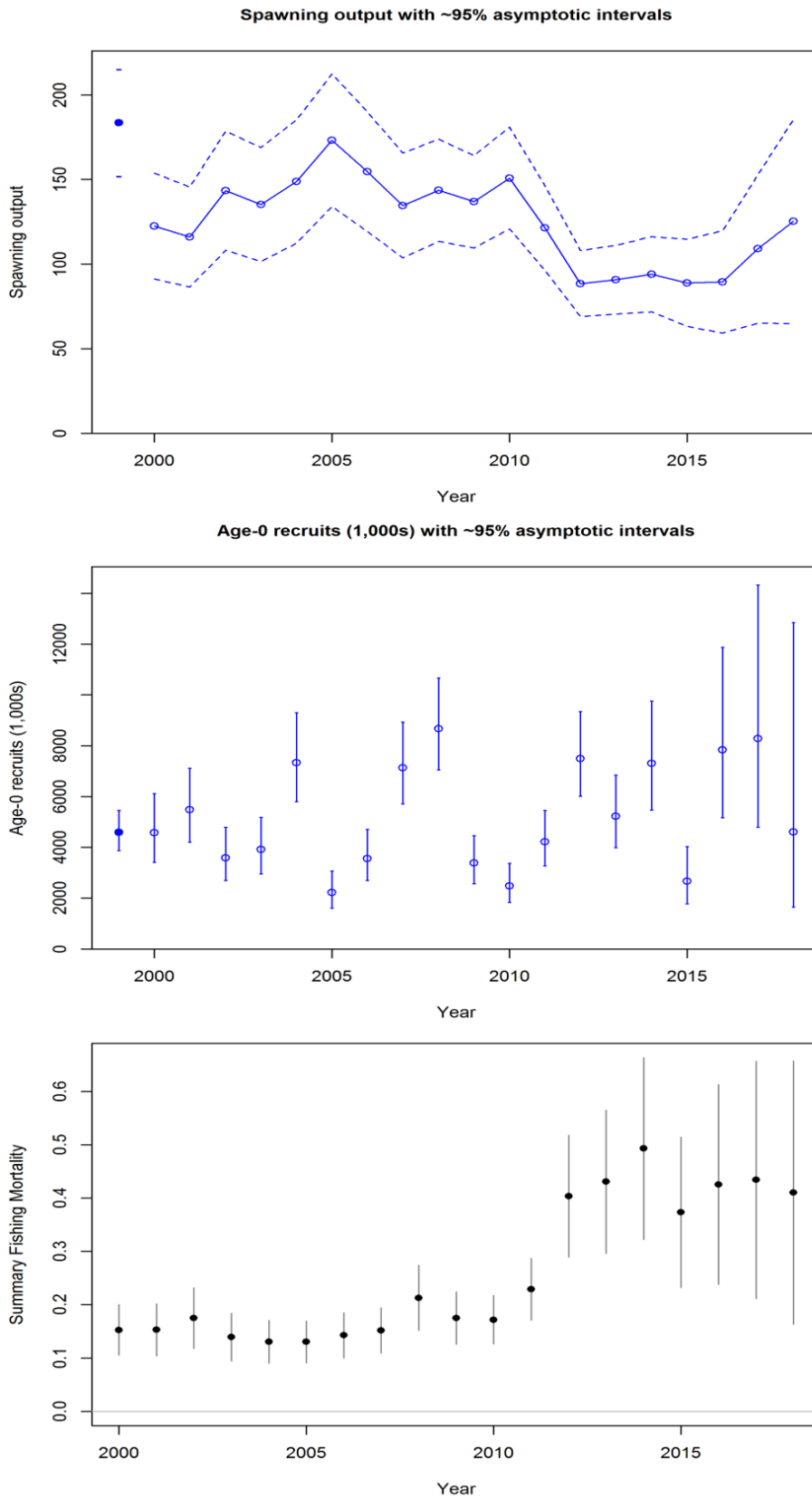


Figure 3.8: Main outputs plots, from top to bottom, SSB, recruitment and fishing mortality (averaged for ages 2 to 5), along with 90% confidence intervals, estimated using the SS approach and generated with the `r4ss` package.

All parameter estimates are given in Table 3.9. The fishery age selectivity s_a was almost equal to zero for age 0 as this age class is caught in very small quantities and then increased at ages 1 and 2, being flat for age 3 and older (Figure 3.9). The survey age selectivity at age 1 was around 0.5 and flat at 1 for ages 2-5, but then decreased again for the 6+ group until 0.8, resulting in a dome-shaped pattern (Figure 3.9). The catchability parameters of the PELGAS and DEPM indices were both above 2. These values are quite high as they indicate that the actual biomass of the population is less than half of the biomass indicated by the indices. This issue was discussed intensively by (ICES 2018b) and lead ICES to consider the assessment results only in relative terms.

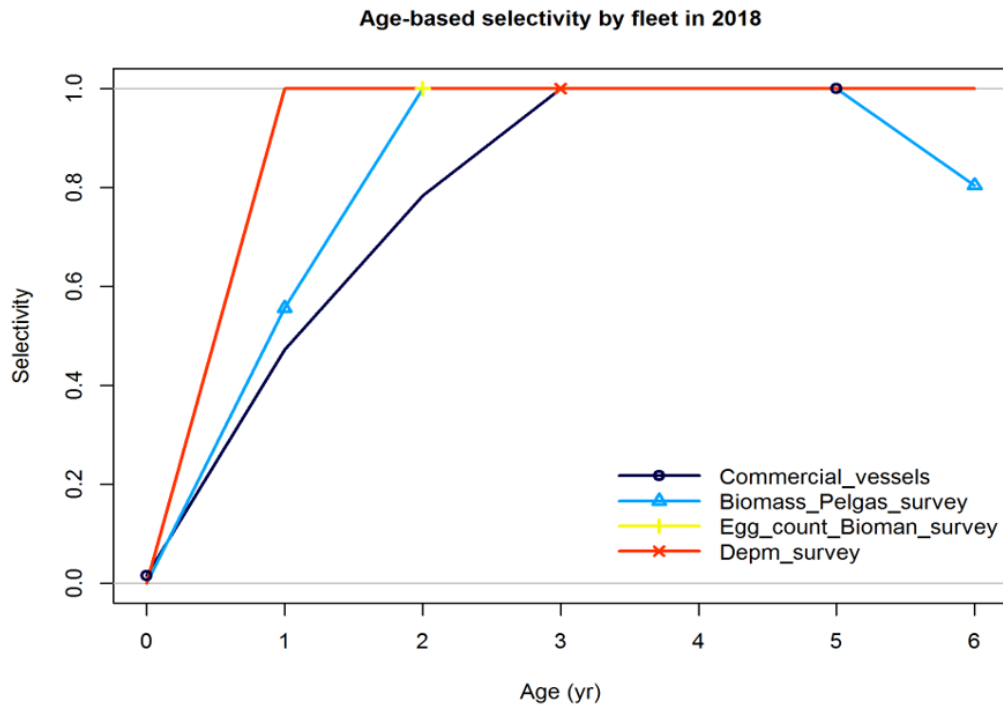


Figure 3.9: Selectivities across ages estimated using the SS approach and generated with the `r4ss` package.

For model diagnostics, the residuals plots of age structured data, such as catch-at-age and acoustic index at-age, were inspected and no significant patterns were detected (Figure 3.10).

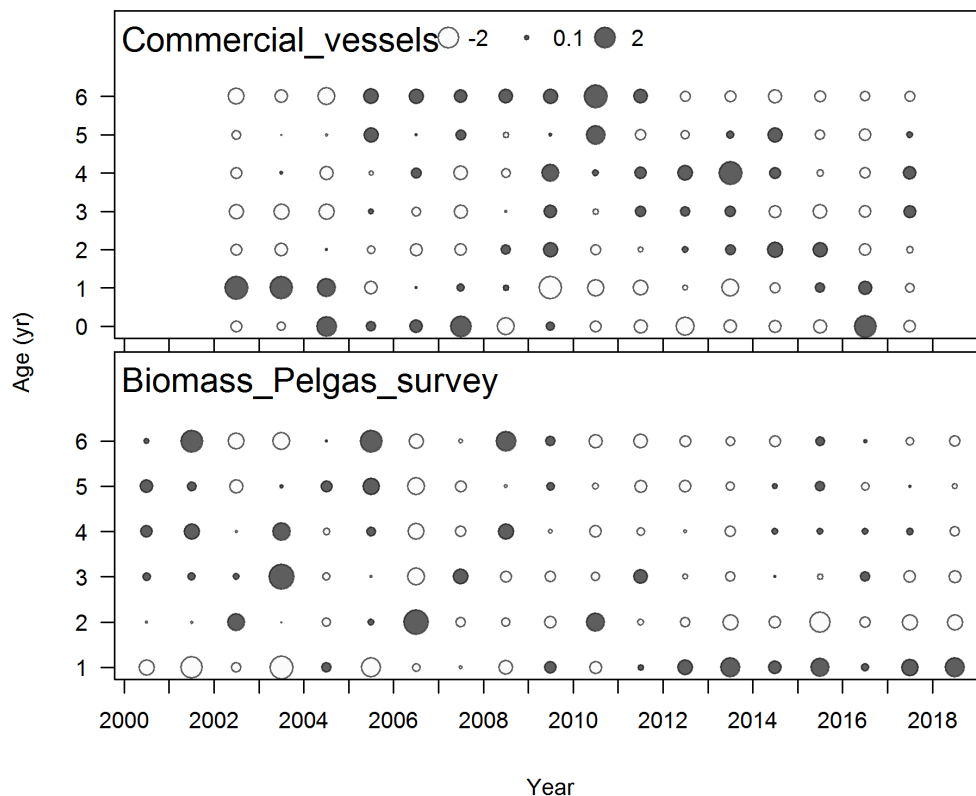


Figure 3.10: Standardized residuals plots for age structured data (catch-at-age data at the top, acoustic survey data at the bottom).

3.3.2 a4a approach

The FL4a package includes a flag on convergence, 0 for true and 1 for false, coming from ADMB convergence indicators. The present model did converge adequately according to this indicator, with a final negative log-likelihood equal to 203.56.

Estimates obtained using the a4a approach are similar in trends to the results obtained by the SS approach. All estimated parameters, both point estimates and variances, are shown in Table 3.9, while most relevant outputs of stock status are summarised in Table 3.6. Spawning biomass in 2002 (first time series year used in this approach) was estimated at 111 thousand tonnes. After attaining the minimum value in 2012, SSB has increased in recent years (Figure 3.11). Recruitment is highly variable along time and shows upwards and downwards peaks like 2016 and 2015 or 2005 respectively. The average fishing mortality for ages 2-5 is estimated below

0.25 until 2011, increasing for the last years up to almost 0.5 in 2014 (Figure 3.11). Although total catch in biomass is not included in the observation model, it can be derived from the catch-at-age estimated by the model and the observed catch weight-at-age (Table 3.6 and Figure 3.11). Uncertainty in the total catch is derived from estimated catch-at-age variances.

Table 3.6: Recruitment (million of individuals), SSB (in tonnes), total catch (in tonnes) and fishing mortality (averaged for ages 2 to 5) estimated using the a4a approach.

| | Recruitment | SSB | Catch | F (2-5) |
|------|-------------|--------|-------|---------|
| 2002 | 2705 | 111001 | 16695 | 0.210 |
| 2003 | 3374 | 104552 | 16185 | 0.177 |
| 2004 | 6888 | 115749 | 14111 | 0.168 |
| 2005 | 2164 | 145285 | 19349 | 0.192 |
| 2006 | 3016 | 131509 | 19949 | 0.207 |
| 2007 | 7802 | 110943 | 18336 | 0.195 |
| 2008 | 7274 | 129546 | 16960 | 0.184 |
| 2009 | 4070 | 128736 | 23581 | 0.218 |
| 2010 | 2275 | 138138 | 18884 | 0.171 |
| 2011 | 3573 | 115854 | 20422 | 0.205 |
| 2012 | 7237 | 83818 | 23258 | 0.306 |
| 2013 | 6206 | 88751 | 29935 | 0.402 |
| 2014 | 8731 | 96937 | 35478 | 0.498 |
| 2015 | 1990 | 98455 | 19939 | 0.235 |
| 2016 | 11172 | 107784 | 33536 | 0.490 |
| 2017 | 7076 | 134739 | 35498 | 0.423 |
| 2018 | | 139308 | | 0.423 |

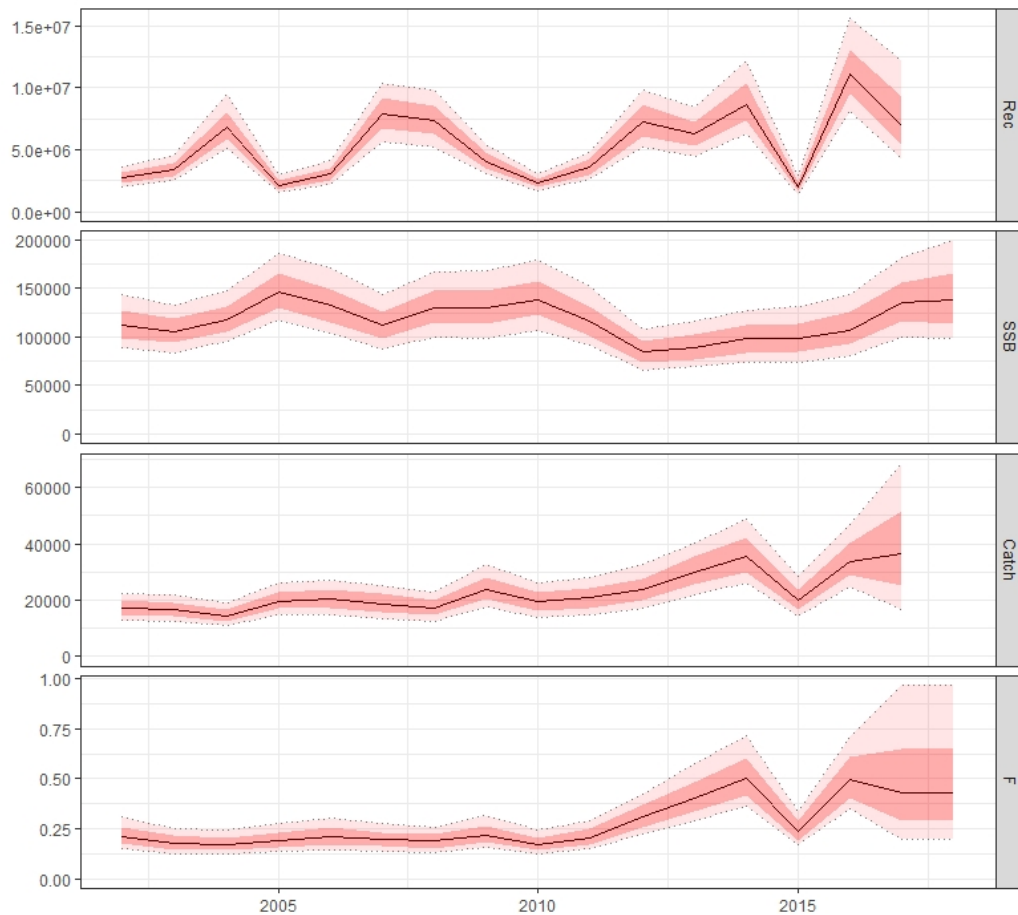


Figure 3.11: Main outputs plots, from top to bottom, recruitment, SSB, catch and fishing mortality (averaged for ages 2 to 5), along with 75% (red) and 90% (light red) confidence intervals estimated using the a4a approach and generated with the FLa4a package.

For model diagnostics, the residuals plots obtained using the FLa4a package did not show any significant pattern for catch estimates while a year related pattern was detected at the start of the time series for the acoustic index (Figure 3.12). Note that both age-structured and aggregated data, in this case, were assumed to follow lognormal distributions independent by age and year.

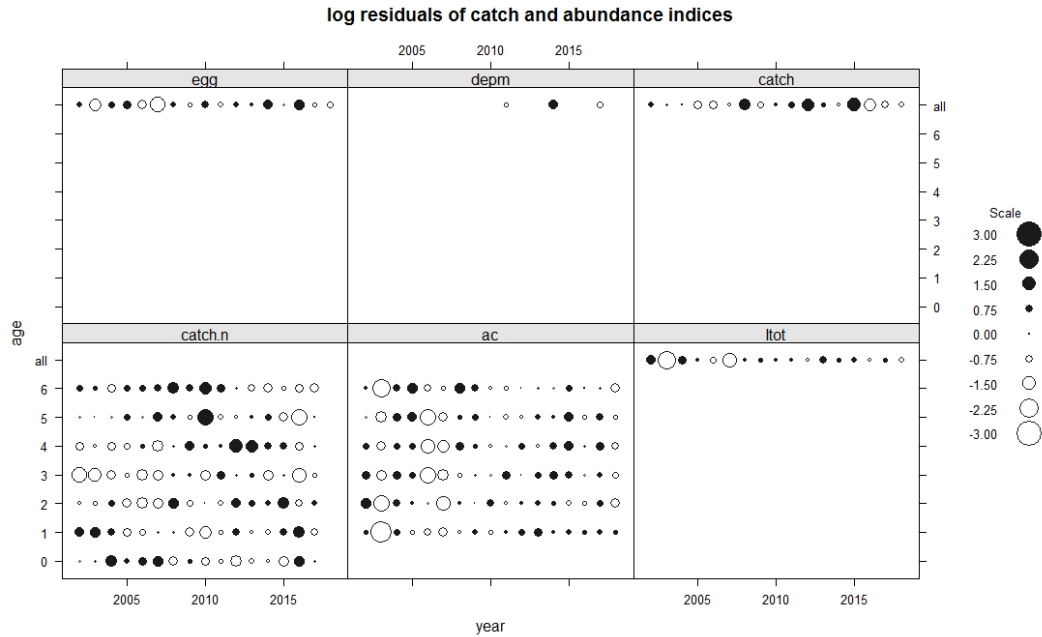


Figure 3.12: Standardized residuals plots for the a4a approach, generated with the FLa4a package.

3.3.3 Bayesian approach

Visual inspections of the autocorrelation and trace plots (Figure 3.13) did not indicate bad mixing or any other convergence issue of the MCMC draws. All the parameters passed the Geweke test that tests the equality of means between the first and last part of a chain (Figure 3.14), the Gelman and Rubin test for within and between chain variance and the Heidelberg and Welch stationarity test. Therefore, the MCMC samples were considered to represent adequately the posterior distributions of the parameters.

Results obtained with the Bayesian approach consist of posterior distributions for all the model parameters as well as for other derived quantities. The point estimates computed as the median of the posterior distribution, and the corresponding variances for all the parameters are listed in Table 3.9. Stock status was summarised in terms of the posterior distributions of recruitment, SSB, total catch and fishing mortality (Figure 3.15). Besides the interannual changes of the estimates, this plot represents also the associated uncertainty. In general, the most recent years have

wider distributions, reflecting a larger uncertainty. Table 3.7 summarizes the medians of these posterior distributions. When comparing the posterior distributions of recruitment and the annual component of the fishing mortality with respect to the prior distributions, posterior distributions are less dispersed and are centred at different values, indicating that data are providing information on model parameters (Figure 3.15).

Standardised residuals which were computed using posterior median values as point estimates, did not show any significant pattern for catch estimates while a year related pattern was detected for the acoustic index (Figure 3.16).

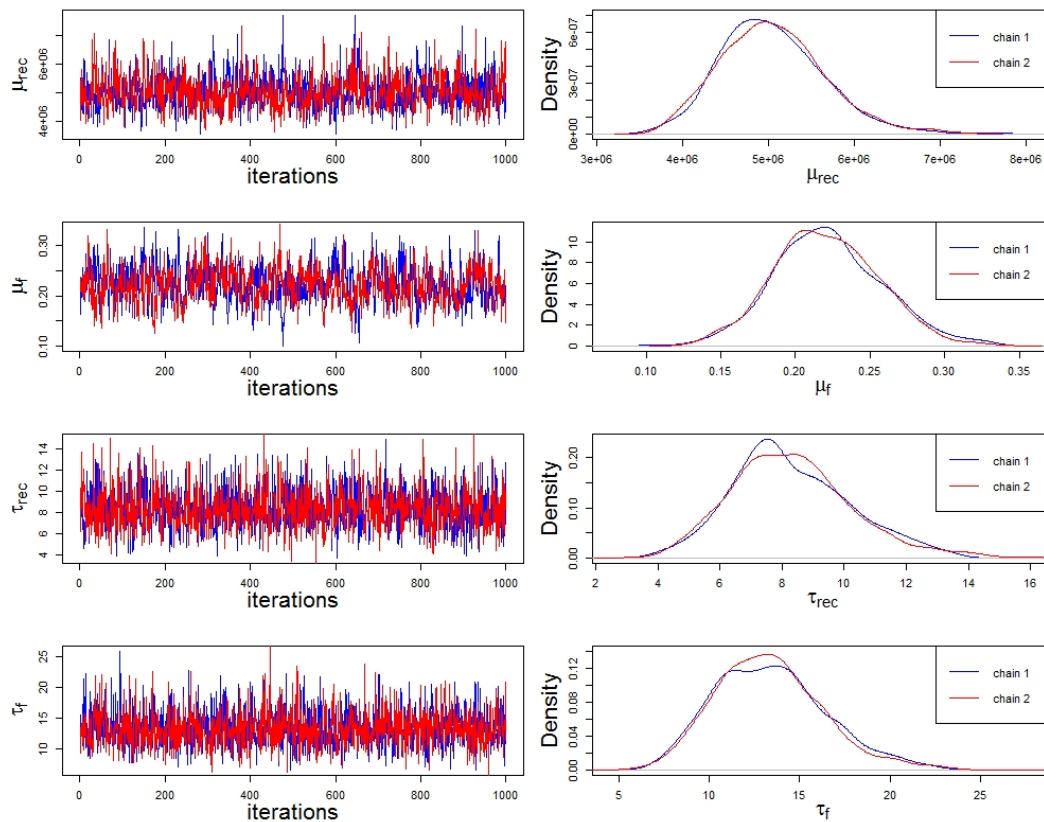


Figure 3.13: Trace plot of MCMC draws (left) and estimated posterior density functions (right) for four parameters (μ_{rec} , μ_f , τ_{rec} , τ_f) for two chains (chain 1 in blue and chain 2 in red).

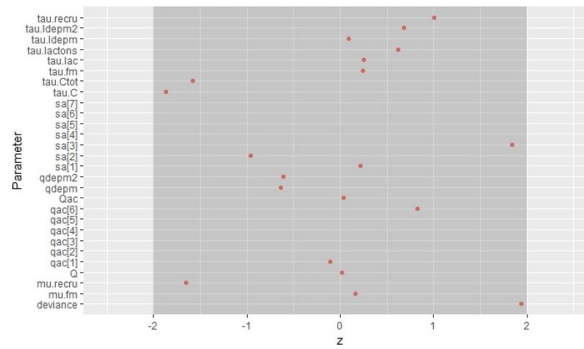


Figure 3.14: Geweke test's Z values for estimated parameters with the Bayesian approach.

Table 3.7: Posterior medians of recruitment (millions of individuals), SSB (in tonnes), total catch (in tonnes) and fishing mortality (averaged for ages 2 to 5) using the Bayesian approach.

| | Recruitment | SSB | Catch | F(2-5) |
|------|-------------|--------|-------|--------|
| 2000 | 4309 | 106254 | 15276 | 0.169 |
| 2001 | 4444 | 102499 | 15185 | 0.165 |
| 2002 | 2895 | 123665 | 18251 | 0.191 |
| 2003 | 3488 | 111621 | 16721 | 0.160 |
| 2004 | 6094 | 123842 | 14364 | 0.153 |
| 2005 | 2569 | 143625 | 16662 | 0.151 |
| 2006 | 3508 | 134171 | 16967 | 0.165 |
| 2007 | 8304 | 120518 | 17537 | 0.165 |
| 2008 | 8008 | 142180 | 21436 | 0.198 |
| 2009 | 5036 | 138633 | 21148 | 0.168 |
| 2010 | 3001 | 155570 | 20111 | 0.154 |
| 2011 | 4506 | 136710 | 22873 | 0.189 |
| 2012 | 8206 | 101377 | 29800 | 0.313 |
| 2013 | 6913 | 105444 | 31976 | 0.335 |
| 2014 | 9284 | 114102 | 33517 | 0.359 |
| 2015 | 2814 | 118358 | 26422 | 0.247 |
| 2016 | 10677 | 123249 | 25597 | 0.288 |
| 2017 | 6903 | 153771 | 29988 | 0.274 |
| 2018 | 5149 | 154266 | 32419 | 0.292 |

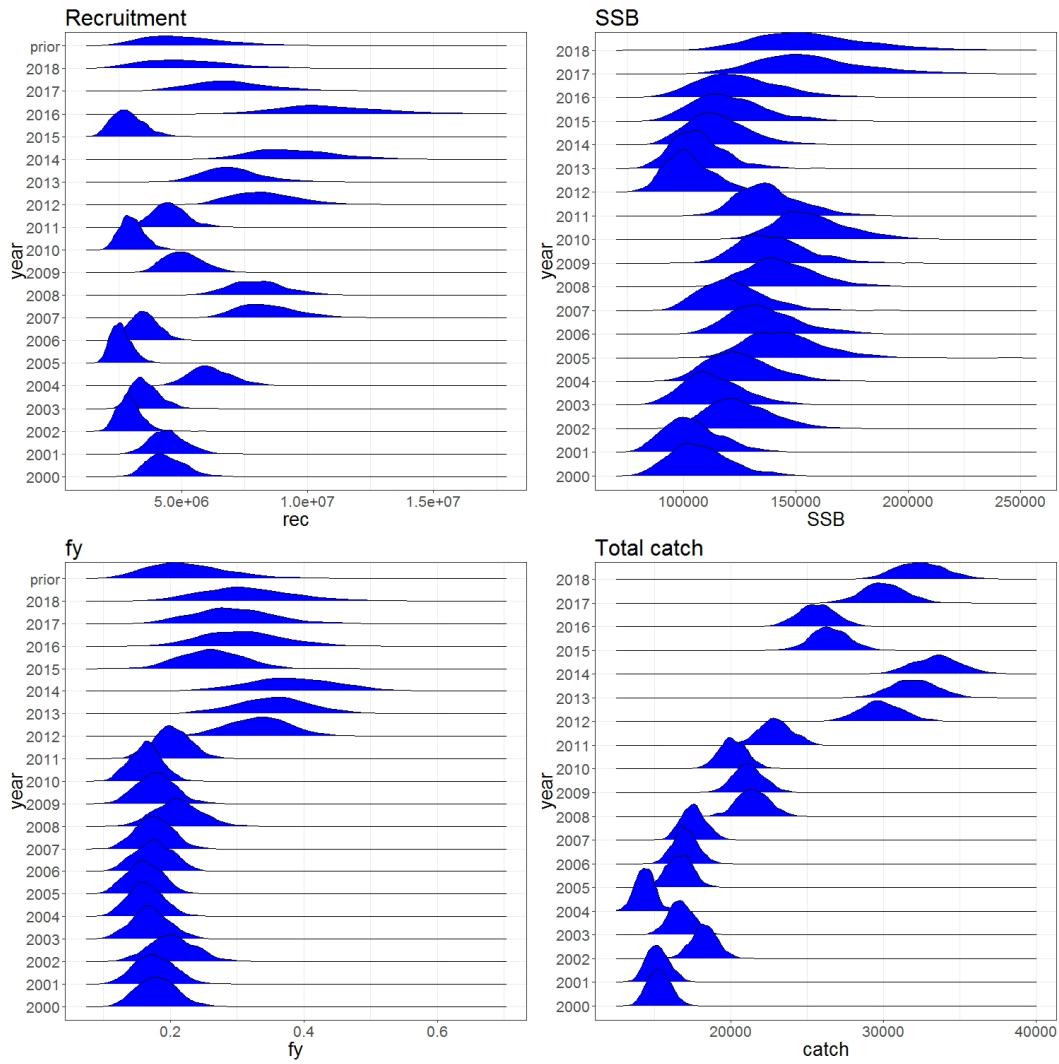


Figure 3.15: Posterior distributions for recruitment, SSB, fishing mortality year component (f_y) and total catch, along years (y axis). For recruitment and f_y parameters the prior distributions are plotted at the top of each panel.

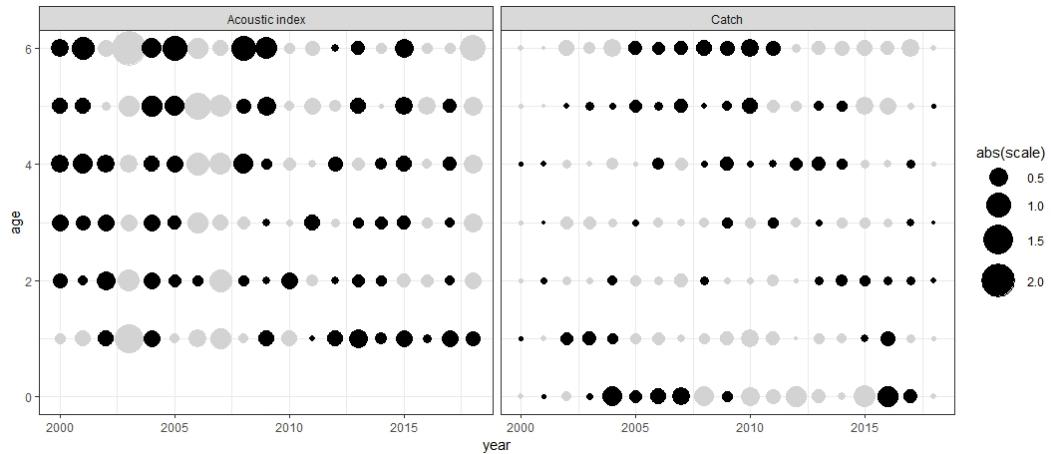


Figure 3.16: Standardized residuals plots for the Bayesian approach. Sizes of the dot represent the absolute value of the residuals, while positive values are coloured in black and negatives in grey.

3.3.4 Comparison

The a4a approach and the Bayesian ad-hoc model were implemented to mimic the features of the official ICES stock assessment model based on the SS approach. Therefore, the comparison between the results of the three approaches took also as reference the SS approach. The three modelling approaches relied on the same population dynamics model but differed on the statistical distributions used for modelling observations. The a4a and Bayesian approaches assumed lognormal distributions for all the observation equations including age-structured catch and abundance indices, while SS modelled the age-structured observations through the multinomial distribution (Table 3.8). However, the observation equations for the aggregated indices were not exactly the same, since the a4a model did not allow to include SSB indices and were replaced by the biomass of age 1 and older. In addition, observation equations for the total catch were only included in the SS and Bayesian ad-hoc models.

Estimated key quantities for describing stock status (SSB, recruitment and fishing mortality) showed similar trends with the three approaches. The estimates did not take exactly the same values, at the beginning of the period the SS approach estimated the highest abundance while the SSB estimate in the final year was higher for the Bayesian approach than the SS approach. In terms of fishing mortality, similar estimates were obtained for the first year, while the final year estimates were

around 0.4 for a4a and SS approach and lower for the Bayesian estimate, with the median around 0.3. However, probability intervals around these values overlapped along with the whole time series, meaning that obtained results were consistent between three approaches (Figure 3.17). Concerning estimated uncertainties, a4a showed the widest confidence intervals being this effect higher in the most recent part of the time series. Note that a4a approach did not include any prior information as in the Bayesian approach nor initial assumptions for some parameters as in SS the approach, such as selectivities' shapes, initial catch in equilibrium or variance parameters for aggregated catch and indices.

Table 3.8: Assumed statistical distributions for each observation dataset for each of the three approaches (SS, a4a and Bayesian ad-hoc).

| | SS | a4a | Bayesian ad-hoc |
|--------------------------|-------------|--------------|-----------------|
| $\widehat{C}_{a,y}$ | multinomial | lognormal | lognormal |
| \widehat{C}_y^{tot} | lognormal | Not included | lognormal |
| $\widehat{I}_{a,y}^{ac}$ | multinomial | lognormal | lognormal |
| \widehat{I}_y^{tot} | lognormal | lognormal | lognormal |
| \widehat{I}_y^{egg} | lognormal | lognormal | lognormal |
| \widehat{I}_y^{depm} | lognormal | lognormal | lognormal |

The rest of the estimated parameters were summarized in Table 3.9 including point estimates and the corresponding standard deviations. Estimated catchability values for all aggregated indices ($Q_{tot}, Q_{egg}, Q_{depm}$) were very similar, slightly higher for the a4a approach, as well as the Q_{ac} for the age-structured index, estimated just with the a4a and Bayesian approach. Selectivity for this age-structured index, fixed to 1 for ages 2 to 5, was estimated for all approaches obtaining values between 0.55 and 0.6 for q_1 and between 0.68 and 0.8 for q_6 . Selectivities for the fishery (commercial vessel) obtained with a4a and the Bayesian approach were also very similar to the ones estimated by SS (Figure 3.9). Finally, estimated variances, for indices and catches, did not show a big difference among approaches. In SS these values were not estimated but provided to the model as inputs, however, the resulting estimated σ s by the other two approaches were not far from these values (Table 3.9).

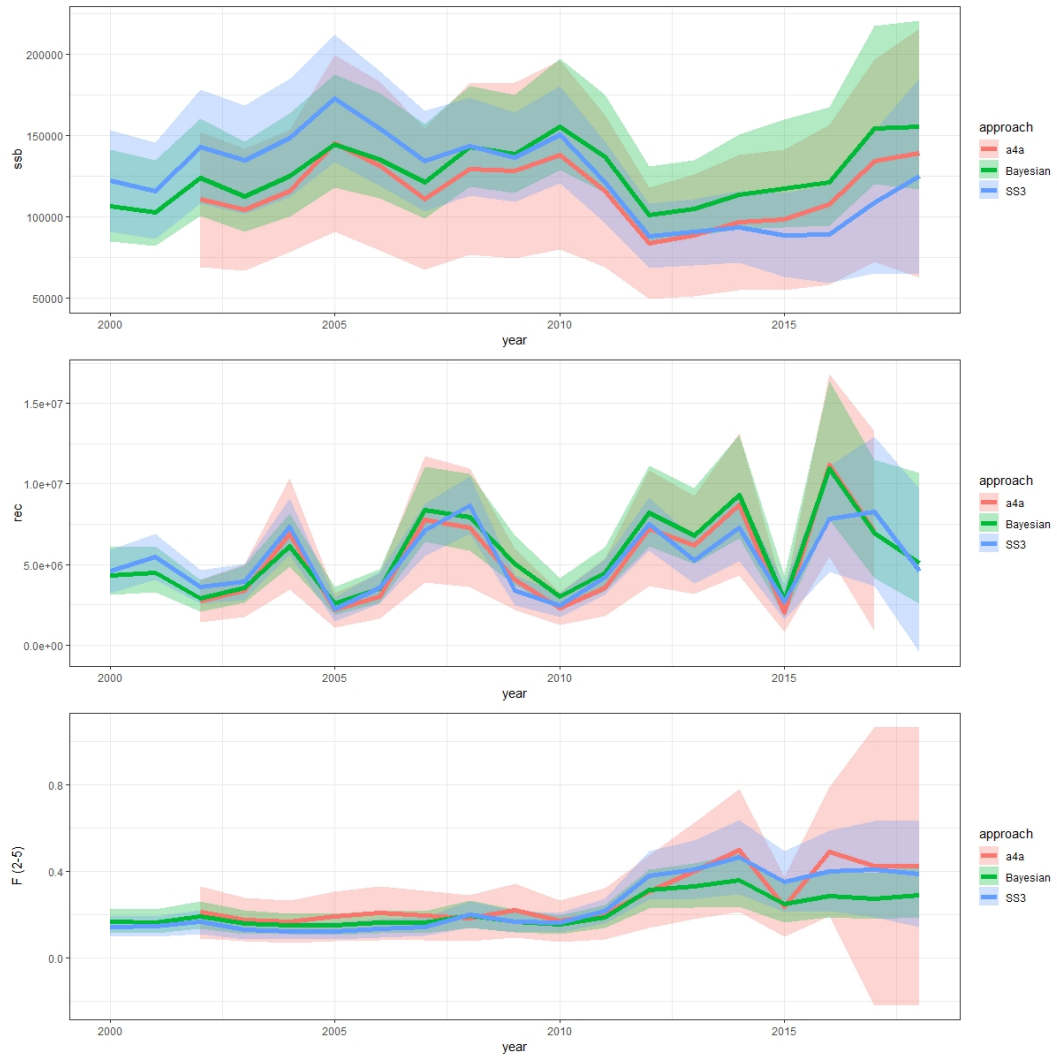


Figure 3.17: Estimated SSB, recruitment and fishing mortality (averaged for ages 2-5) with each of the assessment approaches, along with the corresponding 95% confidence interval for SS and a4a approaches and 95% credible interval for the Bayesian approach.

Table 3.9: Point estimates of model parameters, along with the corresponding standard deviation in brackets, obtained with each of the approaches (SS, a4a and Bayesian). Numbers in grey are given values and values not estimated by the model are marked with ”-”.

| | SS | a4a | Bayesian |
|----------------------------|-------|---------------|---------------|
| s_0 | 0.016 | 0.009 (0.004) | 0.010 (0.002) |
| s_1 | 0.472 | 0.412 (0.173) | 0.496 (0.076) |
| s_2 | 0.783 | 0.727 (0.309) | 0.773 (0.115) |
| q_1 | 0.556 | 0.587 (0.124) | 0.590 (0.114) |
| q_6 | 0.804 | 0.794 (0.188) | 0.683 (0.146) |
| Q_{ac} | - | 3.621 (0.537) | 3.177 (0.486) |
| Q_{tot} | 2.987 | 3.229 (0.468) | 2.865 (0.381) |
| Q_{egg} | 0.462 | 0.598 (0.083) | 0.598 (0.097) |
| $Q_{dep\text{m}}$ | 1.895 | 2.076 (0.828) | 1.920 (1.145) |
| σ_C | - | 0.455 (0.034) | 0.482 (0.033) |
| $\sigma_{C^{tot}}$ | 0.05 | - | 0.050 (0.004) |
| σ_{Iac} | - | 0.721 (0.054) | 0.679 (0.042) |
| $\sigma_{I^{tot}}$ | 0.480 | 0.438 (0.079) | 0.332 (0.036) |
| $\sigma_{I^{egg}}$ | 0.477 | 0.405 (0.075) | 0.480 (0.081) |
| $\sigma_{I^{dep\text{m}}}$ | 0.595 | 0.598 (0.283) | 0.805 (0.451) |

3.4 Discussion

Having an analytical stock assessment is a step forward towards the adequate management of any stock. On the one hand, it allows to define the stock status with respect to biological reference points and evaluate the effect of management measures in the short term. On the other hand, it serves to set the basis for evaluating tactical decisions in the medium or long term. The Bay of Biscay sardine stock was assessed analytically for the first time in 2017 when a statistical-catch-at-age stock assessment model implemented in SS was proposed and accepted as part of an ICES benchmark process (ICES 2017a). These settings were used to provide management advice in 2017 and 2018 (ICES 2017c; 2018b). In this chapter, we have reviewed this official assessment implementation. As shown by the model diagnostics, the official SS assessment summarised well all the data sources and provided meaningful estimates of the quantities of interest. However, some difficulties were faced to

accept the obtained absolute level of the assessment, due mainly to the fact that the estimated catchability parameter for the acoustic index was too high (ICES 2017a). This led ICES to consider the results just in relative terms and to classify the stock in Category 2 (stocks with analytical assessments and forecasts that are only treated qualitatively). However, since then, the assessment has continued evolving. During the Inter-benchmark process held in 2019 (ICES 2019b), several improvements were proposed and implemented in the SS model, such as changes concerning selectivity estimation and the stock-recruitment relationship. Consequently, the stock was upgraded to Category 1 (stocks with quantitative assessments) and served to restart the dialogue with the stakeholders on the development of a management plan.

As an alternative to the official SS assessment, we have explored two approaches: the a4a approach that intends to be a simpler approach in terms of implementation, and an ad-hoc developed model that aims at testing the Bayesian inference paradigm. Both approaches were designed to mimic the official SS assessment. Although some of the assumptions for each of the approaches differ, resulting trends in time for quantities of interest were found to be similar. Statistical distributions used for modelling age-structured observations were also different, having lognormal distributions in the a4a and Bayesian approach and multinomial distributions in SS. However, all approaches resulted in similar estimates concerning age structured catchability and selectivity parameters. Aggregated observations were all modelled through lognormal distributions, in the SS approach the variances were given to the model while in the other two approaches they were estimated obtaining similar values to the given ones. Concerning confidence intervals for the main quantities of interest, i.e. SSB, recruitment and fishing mortality, they were wider for the a4a approach which does not include any prior information as in the Bayesian approach nor initial assumptions for some parameters as in the SS approach. The estimated status of the stock in the terminal year in terms of SSB was higher for the Bayesian approach and lowest for SS, however in general terms, the three different approaches presented here were capable of estimating similar trends in time of stock status, as well as absolute values of estimated parameters.

The election of a software or method for stock assessment has been discussed extensively in the literature (Deroba et al. 2014, Dichmont et al. 2016, Schnute et al. 1997, Thorson 2019). Besides reasons such as data availability and model complexity or flexibility (Edwards et al. 2012a), there might be additional considerations

like the technical skills of the analysts (Berkson et al. 2009), time or computational constraints and the history of the assessments in a region or the locations where the developers of the tools are based (Dichmont et al. 2016). The range of potential assessment tools goes beyond existing packages and software since ad-hoc stock assessment models can also be self-developed for a specific stock, as presented in this chapter or in de Moor and Butterworth (2007), Fernandez et al. (2010), Ibaibarriaga et al. (2008). Each of the different options for stock assessment presents strengths and weaknesses.

The main advantages of using general stock assessment programs can be the confidence in the correct coding of the model due to the testing of several users, the ease to explore different available and implemented model configurations, a faster obtaining of results, easy to peer-reviewing if reviewers are familiar with the selected software or the help of a large user community for improvement or error detection. This can be the case of SS, which is a well-known software with has a big users community and a large number of model configurations implemented. However, as noted in Dichmont et al. (2016), the most flexible and complex software are more likely to be treated as a "black box" and the possibility of mistakes by non-experts is a concern. Most of this type of packages provide results that may look correct in any case and detecting an erroneous configuration can be difficult. This was also our experience with the SS assessment of sardine in the Bay of Biscay, where small configuration issues had to be corrected as we gained experience with the tool.

Many authors have advocated simpler models (Cotter et al. 2004, Hilborn 2003). The a4a initiative can fit into this context. The a4a approach used in this chapter presents a balance between simplicity of use and model flexibility allowing to explore different configurations and combination of sub-models in an easy and fast way (Jardim et al. 2014). The age effect, year effect and cohort effect are easily modelled for any of the submodels allowing also to incorporate non-linear effects. Setting up each submodel is done using R's syntax for equations, which is intuitive and makes it accessible for any R user. ADMB, used by a4a to perform statistical analysis, is also a flexible tool that allows to incorporate different estimation methods such as a maximum likelihood or MCMC methods. All this makes a4a a very versatile option for stock assessment modelling, being able to incorporate external covariates and multiple sources of uncertainty (Scott et al. 2016). It is also an intuitive framework, where fishery scientists can easily and efficiently translate ideas into mathematical

models and thus is presented as useful interface for addressing a range of plausible hypotheses through model averaging (Millar et al. 2014). As shown in this work, it is appropriate for single species and single fleet stock assessments, however, multi-species or multi fleet case studies can not be implemented in a4a.

User-developed ad-hoc code, such as the Bayesian approach presented in this work provides the flexibility to explore configurations or new ideas that are not implemented in other packages. Moreover, the "black box" issue is avoided and aspects of the assessment that are specific for the case study can be incorporated without being restricted by the package's options, which could be restricting innovation in stock assessment science (Dichmont et al. 2016). On the contrary, implementing the model is more time-consuming and requires further testing. Simulation testing like in Deroba et al. (2014) can be used to check that the results are coherent and that the estimation method is working appropriately under the model assumptions.

Another issue that differentiates presented stock assessment approaches is the procedure for the estimation of some parameters, such as sample size for the age-structured data and standard deviations for aggregated indices in our models. a4a and the Bayesian approach are able to estimate these parameters within the model, while SS relies on "tuning algorithms". These tuning algorithms consist of specifying an initial value for the tuning parameter, running the model, adjusting the value of the parameter and repeating the process until a small change is detected. These processes can be time-consuming, difficult to replicate, convergence is rarely defined and uncertainty on these parameters cannot be obtained. In our case study several parameters were obtained following this process and thus no uncertainty could be obtained for these parameters. These tuning process is conducted independently of standard parameter estimation and could be replaced and improved by incorporating mixed-effects (Thorson 2019). On the other hand, Bayesian approaches can easily incorporate the estimation of these parameters that are obtained by "tuning algorithms" in SS. In this work, standard deviations for indices lognormal distributions, are considered parameters with their corresponding prior distributions and are estimated using standard techniques, being able to obtain the uncertainty as well. Thorson (2019) recommends converting existing stock assessment packages to use modern statistical tools such as TMB (Kristensen et al. 2016) or Stan (Carpenter et al. 2017), however, this can be difficult for a model like SS with thousand of code lines.

In fact, estimating uncertainty is crucial in stock assessment (Hilborn 1992). All of the packages or approaches are able to quantify estimation uncertainty, although methods may differ among packages. Uncertainty estimation is conditioned on the data and the model structure and can be obtained using different statistical tools such as the delta method, bootstrap, likelihood profiling or Markov chain Monte Carlo (MCMC). Despite the potential for estimating the uncertainty of MCMC methods (Magnusson et al. 2012, Stewart et al. 2012), they are not very extended given that most of these models have to be written ad-hoc. Their use is often restricted to the capabilities inherited from ADMB or TMB, i.e. SS and a4a, which are based on ADMB, can be run with an MCMC option (detailed at Monnahan et al. 2014), obtaining the uncertainty estimations through this method. The Bayesian approach in this work was coded in JAGS from where the MCMC estimates were obtained with a run time of a few minutes. However long run times are usually needed for integrated stock assessment models with higher complexity or longer time-series. The improvement of MCMC methods is proposed as a solution in Monnahan and Kristensen (2018) and Monnahan et al. (2019), where the use of the no-U-turn sampler (NUTS, Hoffman and Gelman 2014) within ADMB or TMB is presented. The NUTS algorithm is executed by Stan (Carpenter et al. 2017) which is compared to JAGS through a simulation study in Monnahan et al. (2017), concluding that the preferred software depends on model characteristics, being Stan more efficient in most of the studied cases, but remarking that "JAGS clearly remain a valuable tool when run-time is not prohibitive".

The resulting uncertainties from stock assessment models are then used for management advice and risk assessment, and could affect when the evaluation of different management actions. In the next chapter, three different methods for uncertainty estimation are compared and their performance is studied under different model complexity scenarios using the a4a approach as the most simple option for dealing with several distinct models. The management advice is studied in the last chapter of the thesis, using the results obtained in these previous two chapters for conditioning the MSE and for introducing model uncertainty in the process explicitly.

Chapter 4

Uncertainty estimation and model selection in stock assessment models with non-parametric effects on fishing mortality

In this chapter we describe the Monte Carlo simulation study carried out to compare three different methods for uncertainty estimation in stock assessment models: multivariate normal distribution, bootstrap (without and with relative bias correction) and Markov chain Monte Carlo (MCMC). In all approaches model fitting is carried out using the a4a framework and the generated population is an emulation of the Bay of Biscay sardine population.

4.1 Introduction

The current management system for marine fish and shellfish resources relies in stock assessments to set appropriate exploitation levels for commercial stocks. Stock assessments provide estimates of population and fleet dynamics, as well as past and current status of fish populations. This information constitutes the main product provided by science to policy makers and managers, in order for these to take evidence supported management decisions. Therefore, in this framework, stock assessment models should provide accurate estimates of the relevant parameters including their uncertainty. Such information is important to communicate the levels of confidence on scientific results and to allow the development of risk analysis (Francis and Shotton 1997, Hilborn et al. 2001).

Brooks and Deroba (2015) classified the uncertainty of stock assessment estimates into three broad categories: observation, structural and estimation uncertainty. Observation uncertainty is inherent to the input data whereas structural uncertainty is related to the model configuration describing the dynamics of the population. Estimation uncertainty is conditioned on the data and the model structure and can be quantified using different statistical tools such as the delta method, bootstrap, likelihood profiling or Markov chain Monte Carlo (MCMC). According to Patterson et al. (2001) these tools can be classified within the frequentist, likelihood or Bayesian paradigms. Previous works have compared the performance of different uncertainty methods using real or simulated data sets. Magnusson et al. (2012) compared the delta method, bootstrap and MCMC using simulated data and suggested that MCMC was the most reliable method given the dataset and the assessment method. Stewart et al. (2012) compared maximum likelihood (MLE) and Bayesian methods, concluding that MLE approximation under or overestimates the upper and lower portion of long tailed distributions. MacCall (2013) presented the delta method as a quick and practical solution for estimating precision of assessment quantities and in Elvarsson et al. (2014) bootstrap method was compared to Hessian-based approximations and proposed a comparison with MCMC methods as future work.

Stock assessment models have typically relied on parametric functions that can be easy to implement, but may lack enough flexibility to capture all the data features (Hillary 2012, Maunder and Harley 2011). Non-parametric models where the model structure is not fixed beforehand are well-established in regression models like generalized additive models (Hardle 1994). Similar smooth functions are being progressively introduced into assessment models (Aarts and Poos 2009, Thorson and Taylor 2014) because they capture the underlying nonlinear structure of parameters like selectivity or catchability (Crone et al. 2013, Maunder and Piner 2015). However, little is known about how this might affect the uncertainty estimates. In general, the model fit improves when using smooth functions. However, the confidence intervals of the smooth functions are wider at the data extremes where less observations are available (Marra and Wood 2012). This is expected to affect especially highly complex functions and the final assessment years, where some cohorts have not been observed in full. Aarts and Poos (2009) found wider confidence intervals at the beginning of the time series and they suggest that uncertainty in the last years of the assessment could be underestimated. In other studies, such as Thorson and Taylor

(2014), uncertainty increased both at the beginning and at the end of the time series depending on the shape of the selectivity. In the first case confidence intervals are derived from a multivariate distribution generated from maximum likelihood estimations and the corresponding hessian matrix using the percentile method and in the second case uncertainty is reported as 80% simulation intervals. This diverges from the usual approach in GAM's where the confidence intervals provided for the smooths are computed using a Bayesian approach (Marra and Wood 2012, Wahba 1983, Wood 2006b). So, it seems important not only to understand the uncertainty level when smooth functions are introduced into stock assessment models, but also to analyse if they differ depending on the estimation method.

The inclusion of smooth functions has been advocated as a way forward to avoid model miss-specification (Maunder and Harley 2011), although it makes the model selection more difficult. Different criteria like AIC (Akaike 1974) and BIC (Schwarz 1978) are used to determine which parameters and which shapes are more appropriate in a specific assessment model. AIC tends to select models with a larger number of parameters while BIC tends to choose simpler models (Dziak et al. 2012). Some examples in fisheries include Wang and Liu (2006) or Butterworth and Rademeyer (2008), but their use in models with smooth functions have not been compared.

The main objective of this work is to compare various uncertainty estimation methods and model selection statistics when nonlinear functions are included into the assessment model, specially, when fishing mortality is modelled by smooth functions depending on age and year.

4.2 Material and methods

A comparison of three uncertainty estimation methods, approximate multivariate normal distribution based on the Hessian matrix, posterior probability intervals from Markov chain Monte Carlo and parametric bootstrap (without and with bias-correction), is done within a Monte Carlo simulation framework (Deroba et al. 2014) using the R package Assessment for all (**a4a**) (Jardim et al. 2014). This is a new statistical framework for age-based fish stock assessment designed to be flexible in terms of model structure, using R's syntax for model building. For this study, smooth functions have been introduced for fishing mortality, which range from the classical separable age and year effect as factors to bi-dimensional smooths allowing

the interaction between age and year.

4.2.1 Operating model and scenarios

In this study the operating model that represents the true population dynamics, is based on the a4a age-structured model fitted to sardine (*Sardina pilchardus*) in the Bay of Biscay (WGHANSA, 2015). Real data on this stock consists of 13 years (2002-2014) of landings and catch-at-age (ICES 2015), numbers-at-age and weights-at-age from the acoustic survey PELGAS (Massé et al. 2016) and an abundance index from the Daily Egg Production Method survey BIOMAN (Santos et al. 2018; 2011). There is observation data from surveys for every year in the model, being each survey an independent data source.

Let $N_{a,y}$ denote the number of individuals of age a at the beginning of year y . Then, according to a4a framework (Jardim et al. 2014) the population dynamics are described by the usual equations:

$$N_{a,y} = N_{a-1,y-1}e^{-Z_{a-1,y-1}}, \quad (4.1)$$

$$N_{A,y} = N_{A-1,y-1}e^{-Z_{A-1,y-1}} + N_{A,y-1}e^{-Z_{A,y-1}}, \quad (4.2)$$

where A represents the plus group with individuals aged A and older, which in this case is $6+$. The total mortality of age a individuals during year y ($Z_{a,y}$) is decomposed as the sum of fishing mortality and natural mortality ($Z_{a,y} = F_{a,y} + M_{a,y}$).

According to the Baranov equation, catch-at-age ($C_{a,y}$) are computed as the fraction of fishes dying each year due to fishing:

$$C_{a,y} = N_{a,y}(1 - e^{-Z_{a,y}})F_{a,y}/Z_{a,y}. \quad (4.3)$$

The a4a model observations equations include observed catch-at-age, an age-structured abundance index and it might also include an aggregated abundance index. Catch-at-age, $\hat{C}_{a,y}$, are assumed to be lognormally (LN) distributed with variance σ_C :

$$\hat{C}_{a,y} \sim LN(C_{a,y}, \sigma_C). \quad (4.4)$$

The age-structured and aggregated abundance indices are also assumed to be log-

normally distributed with variances $\sigma_{I^{ac}}$ and $\sigma_{I^{egg}}$:

$$\hat{I}_{a,y}^{ac} \sim LN(Q_{ac}N_{a,y}, \sigma_{I^{ac}}), \quad (4.5)$$

$$\hat{I}_y^{egg} \sim LN(Q_{egg}SSB_y, \sigma_{I^{egg}}), \quad (4.6)$$

where Q_{ac} and Q_{egg} are survey catchabilities and SSB_y denotes spawning stock biomass in year y .

Natural mortality is assumed known, although it can have a different value for each year and age class. Therefore, the parameters to be estimated are related to recruitment, number of individuals at initial year, fishing mortality, survey catchabilities and observation variances components. The a4a framework allows a variety of structures for each of these components: from the simplest constant value case to more complex and flexible structures represented by smooth functions depending on age, year or any additional covariate. Parameters are log transformed for estimation.

In the present study, five different scenarios (Table 4.1) have been considered based on the fishing mortality (F) shape depending on age and year, where smoothing is introduced. In the first scenario ages and years are taken as categorical covariates while the rest of the scenarios include smoothing functions with different levels of complexity over ages and years. The first three scenarios represent a separable structure for F while the last ones are non-separable and allow interactions between age and year. Different levels of smoothing were tested by increasing the number of knots in the smoothers. Implementation of smooth functions in a4a is done through the `mgcv` package (Wood 2006a) in R. `mgcv` package is used to construct the structure of the selected smooth function with a given number of knots, i.e. the only thing imported from `mgcv` are the design matrices, which are created using ‘`smooth.construct`’ type functions from this package for each particular smoothing option. a4a uses `s()` and `te()` methods which are used to set up the smoothers. `s()`, given a fixed number of knots, sets up fixed degrees of freedom unpenalised thin plate spline for one dimension, while `te()`, given also the number of knots, sets up an unpenalised tensor product of cubic splines for multi-dimensional cases. Thus, smoothness is fixed by the user for the generation of the design matrix through ‘`mgcv`’ given that the parameter estimation method is done in ADMB and it has not been implemented as a method to estimate smoothing parameter numbers inside the assessment model.

Table 4.1: F structure definition in each scenario.

| Scenario | Name | Notation in R | Description |
|----------|---------------------------|---|--|
| S1 | <i>age x year</i> | $\sim \text{factor}(\text{age}) + \text{factor}(\text{year})$ | Classical age-year separable structure. |
| S2 | <i>high order splines</i> | $\sim \text{s}(\text{age}, k=6) + \text{s}(\text{year}, k=8)$ | Separable structure based on splines with 5 degrees of freedom for age and 7 for year. |
| S3 | <i>low order splines</i> | $\sim \text{s}(\text{age}, k=3) + \text{s}(\text{year}, k=4)$ | Separable structure based on splines with 3 degrees of freedom for age and 4 for year. |
| S4 | <i>high order tensor</i> | $\sim \text{te}(\text{age}, \text{year}, k=c(6,5))$ | Non-separable structure based on a tensor product with 29 degrees of freedom. |
| S5 | <i>low order tensor</i> | $\sim \text{te}(\text{age}, \text{year}, k=c(3,3))$ | Non-separable structure based on a tensor product with 8 degrees of freedom. |

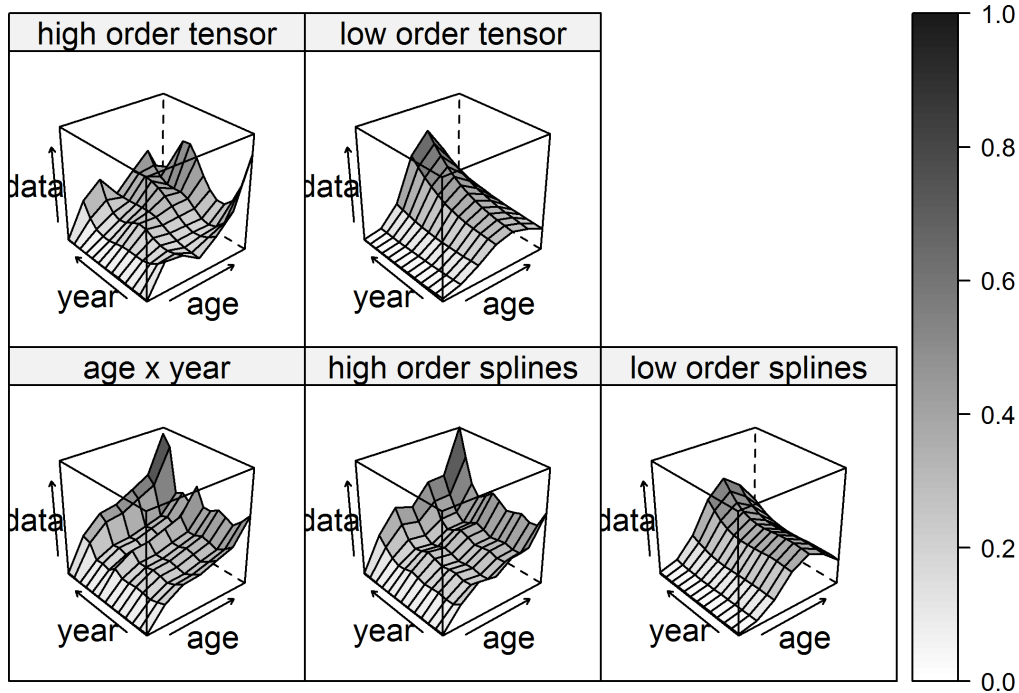


Figure 4.1: Selected fishing mortality shapes for each scenario as function of age and year.

For the sake of simplicity, constant catchabilities, and constant observations variances, across ages and years were assumed. Recruitment and initial year numbers-at-age are estimated for each year and each age respectively. Under these assumptions the models for each scenario were fitted to sardine data. F shapes derived from these structures for each scenario are shown in Figure 4.1. The estimated parameters were taken as the true values for the operating models in the simulation study.

4.2.2 Data simulation

True values were generated fitting real data to the defined operating model for each of the scenarios. Taking estimates of catch and indices (an acoustic index and a biomass index) from this initial run for each scenario as true values, new simulated datasets, consisting of catch-at-age and indices data, were generated multiplying a lognormal error. The coefficient of variation of these errors was set as 25% for catch data and 20% for indices data so that these values remain constant across all scenarios (Magnusson and Hilborn 2007). For each scenario 100 simulated datasets were generated.

4.2.3 Implemented approaches

For each simulated dataset, the a4a assessment model was fitted using three different approaches for uncertainty estimation. The assessment settings were equal to the operating model used for data simulation, i.e. same parametrizations were used in the operating model and the assessment model of each scenario (Magnusson et al. 2012).

The a4a stock assessment model is fitted with maximum likelihood and implemented in R/FLR/ADMB, (Kell et al. 2007, Team 2015). Automatic differentiation is used for likelihood maximization through ADMB (Fournier et al. 2012). Convergence from ADMB was checked obtaining that 100% of the fits had converged.

4.2.3.1 Multivariate normal approach

The first estimation method is based on the general maximum likelihood theory, according to which the maximum likelihood estimates are asymptotically unbiased, normally distributed with variance given by the inverse of the hessian matrix (Vaart and W. 1998). Thus, the empirical distribution of the parameters was approximated

by obtaining a random sample of size 1000 from a multivariate normal distribution centred on the maximum likelihood point estimates with variance-covariance matrix given by the inverse of the hessian matrix. This variance-covariance matrix accounts for the correlation between the estimated parameters. Derived quantities, such as SSB, are computed for each sample, so that their empirical distribution can be derived.

4.2.3.2 MCMC in ADMB

Markov Chain Monte Carlo (MCMC) methods, used to sample from complex multi-dimensional distributions, are widely used in Bayesian inference. ADMB has an option for Bayesian analysis using MCMC, (Monnahan et al. 2014), which has been implemented in the a4a framework. It consists in sampling the maximum likelihood surface using MCMC, with a multivariate normal distribution as the proposal function. Although priors are not specified for this study, a flat prior based on parameters bounds is always assumed in ADMB (Millar 2011). Note that ADMB developers suggest using MCMC results with caution given that there is no documentation available about this implementation. They recommend comparing results using another MCMC software and looking at standard diagnostics. Both recommendations have been followed in this work, looking at convergence diagnostics and obtaining equivalent results in JAGS (Plummer 2003), verifying this way that obtained MCMC results are valid.

In order to determine the number of draws and the thinning for MCMC convergence was checked using original datasets for the fittings. Posterior traces and autocorrelations were checked and Geweke and Raftery and Lewis's diagnostic tests were performed and passed. 100,000 draws were obtained with a thinning of 100. Thus, a sample of size 1000 was saved for subsequent analysis.

4.2.3.3 Parametric Bootstrapping

A parametric bootstrap (Efron and Tibshirani 1994) was performed for each simulated dataset. From the a4a model estimates 1000 catch and indices bootstrap samples were generated according to the observation equations in (4.4) to (4.6). For each bootstrap sample the model is fitted obtaining 1000 estimates. A relative bias

correction (BC) algorithm has been applied to these results. The applied algorithm is a relative bias correction and acceleration algorithm, setting the acceleration coefficient to zero. It adjusts for differences between the median of the bootstrap percentile density function and the estimate obtained with the original data sample (Efron and Tibshirani 1994).

4.2.4 Performance evaluation

Performance of these methods was evaluated for the following estimated parameters and derived quantities: SSB_y (spawning stock biomass for each year y), R_y (recruitment for each year y), $F_{0,y}$, $F_{2,y}$, $F_{6+,y}$ (fishing mortality for ages 0, 2 and 6+ for each year y), Q_{ac} , Q_{egg} (survey catchabilities) and σ_C , σ_{Iac} , σ_{Iegg} (observations variances). For each scenario and each simulated dataset (i), the uncertainty estimation of each parameters (in general denoted as θ) has been measured in terms of:

- Point estimate, $\hat{\theta}_i$, defined as the median of the 1000 iterations saved in simulation i .
- Coefficient of variation, $CV_{\hat{\theta}_i}$, defined as the cv (ratio of the standard deviation to the mean) of the 1000 iterations saved in simulation i .
- Relative bias, $B_{\hat{\theta}_i}$, defined as $(\hat{\theta}_i - \theta)/\theta$, where θ are parameters' true values.
- Skewness, $SK_{\hat{\theta}_i}$, defined as $\mu_{3\theta_i}/\sigma_{\theta_i}^3$, where μ_3 is the third central moment and σ_{θ_i} is the standard deviation of the 1000 iterations saved in simulation i .
- Coverage probability, $CP_{\hat{\theta}}$, is defined as $\Pr(\theta \in PI_{\theta_i})$ the proportion (out of the 100 simulations) of 90% confidence intervals (computed using 5th and 95th percentiles of the 1000 iterations saved in simulation i) that contain the true value.

Point estimates, coefficient of variations, relative bias and skewness measures defined above have been analysed through their corresponding medians and 5th and 95th quantiles across simulation.

4.2.5 AIC and BIC accuracy

In order to evaluate the accuracy of AIC and BIC in choosing the correct model, all simulated datasets have been fitted using the 5 different models defined for each scenario and the AICs and BICs have been computed. In each case the model with

the lowest value of either AIC or BIC was defined as the selected model, computing afterwards the proportion of cases in which a correct selection was done for each scenario.

4.3 Results

All methods (bootstrap, bias corrected bootstrap, maximum likelihood and MCMC) give nearly the same point estimates for SSB, recruitment, fishing mortality and catchability parameters, but different for the estimated variability for catch and indices. The median across the 100 simulations of the relative bias for SSB are slightly positive (around 0.01) but the 90% confidence intervals do not show a trend toward positive or negative relative bias (Figure 4.2). The 90% confidence intervals of the relative bias increase slightly for the last two years, related with higher coefficient of variations, being (-0.12, 0.12) in the last year. The empirical distributions of the CVs of SSB increase from less than 5% in the first years to nearly 10% in the last year (Figure 4.2). When comparing the different approaches the 'multivariate normal' and the two types of bootstraps show nearly the same coefficients of variation while the MCMC option in ADMB results in nearly 15% higher CVs in most of the cases. When comparing scenarios, High order tensor scenario (S4) presents slightly higher values in the last two years and especially for the MCMC approach.

Similar results were obtained in terms of relative bias and CV's for recruitments (Figure 4.3). The 90% confidence intervals of relative bias for recruitment, around (-0.1, 0.1), do not indicate systematic under or overestimation. The increase of the confidence intervals of relative bias in the last year ranges from a ratio of 3.5 with respect to middle years in scenario S1 and S2 to nearly a ratio of 5 in scenario S4. For the CV's in recruitment estimates larger values than for biomass were found, especially in scenario S4 where median CVs in the last year are 5 times greater than those for SSB, reaching values from 45% to 60% depending on the method. For the rest of the years this difference is lower (around 2.5 times) taking CV values near 6%.

Medians across simulations for catchability parameters (Q_{ac} and Q_{egg}) do not show significant relative bias in any of the scenarios and methods with 90% confidence intervals around (-0.08, 0.07) for Q_{ac} and (-0.1, 0.1) for Q_{egg} (Figure 4.4). Observation variances are systematically underestimated in the maximum likelihood (ML) fit. As a result, generated samples for bootstrapping, where estimated variances are used,

present lower variances, and when replicating the fitting it is again underestimated, obtaining even lower observation variance estimates than in the ‘multivariate normal’ approach. The bias corrected version corrects this underestimation for these three parameters, with no significant effects in the rest of estimates. Figure 4.4 shows relative bias median values taken over the 100 simulations where negatives values are detected for bootstrap and multivariate normal approaches, indicating this underestimation. Most of the 90% confidence intervals for these parameters include the zero value, except for σ_C and σ_{Iac} using Bootstrap method. For multivariate normal approach median values are higher but still negative ranging from medians near -0.12 for σ_C to -0.05 for σ_{Iegg} . BC and MCMC approach seem to perform better in term of bias, showing nearly null median values. Except for scenario S4, where the two bootstrap approaches and the multivariate normal approach show lower median relative bias values for σ_C , similar results were obtained for the rest of scenarios. In all scenarios and with all methods σ_{Iegg} show wider relative bias confidence intervals, around 2.5 times σ_{Iac} intervals.

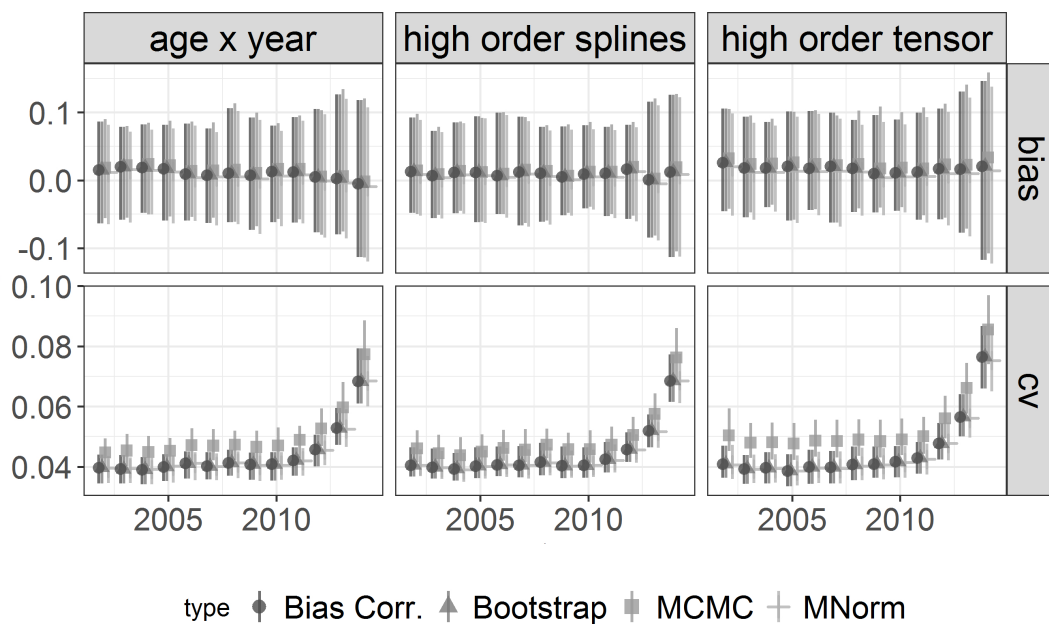


Figure 4.2: Medians (points) and 90% CIs (vertical bars) of the relative bias (top row) and the coefficient of variation (bottom row) of SSB for each year (x-axis). From left to right scenarios age x year, high order splines, and high order tensor (S1, S2, and S4) are represented in each column. Each point symbol and colour corresponds to an estimation approach.

Regarding the CVs of the catchability and observation variances, Q_{egg} presents slightly higher median CVs (around 6.5%) than Q_{ac} (4.5%) with wider 90% confidence intervals for all scenarios and methods (Figure 4.4). σ_C and σ_{Iac} present similar CV values with a common pattern where BC method shows lower values with wider 90% confidence intervals around (6.5-8.5)% in contrast with non-overlapping intervals for bootstrap approach. For σ_{Iegg} a similar pattern is observed with higher CVs around 23%.

Fishing mortality estimates do not show any trend in bias with 90% confidence intervals narrower than (-0.3, 0.3) in most of the cases (Figure 4.5). A small increase of the intervals width can be seen for first and last ages as well as an increase for the last year (30% - 40% increase with respect to other ages or years), especially for High order tensor scenario (S4), where fishing mortality is modelled as a tensor product with high flexibility. In this last scenario estimates for the first age in the

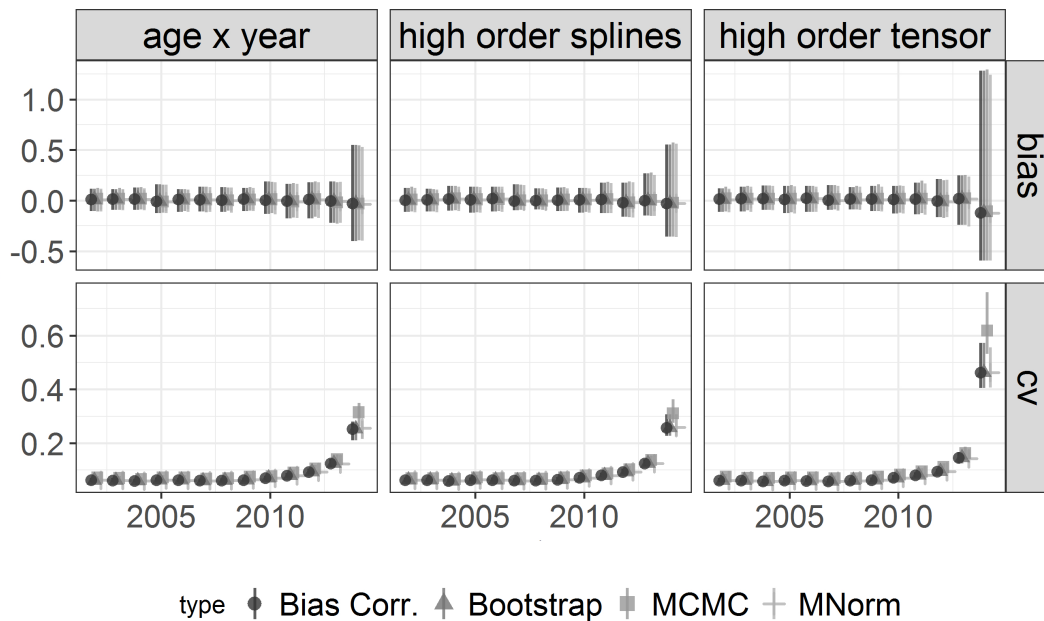


Figure 4.3: Medians (points) and 90% CIs (vertical bars) of the relative bias (top row) and the coefficient of variation (bottom row) of recruitment for each year (x-axis). From left to right scenarios age x year, high order splines, and high order tensor (S1, S2, and S4) are represented in each column. Each point symbol and colour corresponds to an estimation approach.

last year shows a more than 3 times wider confidence interval in contrast to middle years, ranging from -0.5 to 1.5. Concerning CV medians for F estimates (Figure 4.6), different shapes are observed depending on the scenario, having similar structures across methods. Scenarios where F was modelled with a separable submodel present a flatter shape, with median CV values around 11% and a small increase for the last year, taking values up to 16%. Scenarios where F was modelled as non-separable the increase in median CVs was not only for last year but also for first years and first and last ages, ranging from 12% for middle years and ages to 50% for the most extreme value.

All selected parameter distributions seem to be symmetric as the skewness statistic range between -1 and 1 (Figure 4.7).

Coverage probabilities for parameters in the last year (Table 4.2) for the 'multivariate normal' approach and bootstrapping are lower, taking values between 0.75 and 0.85 while MCMC method gives coverage probabilities above 0.9, as shown in Figure

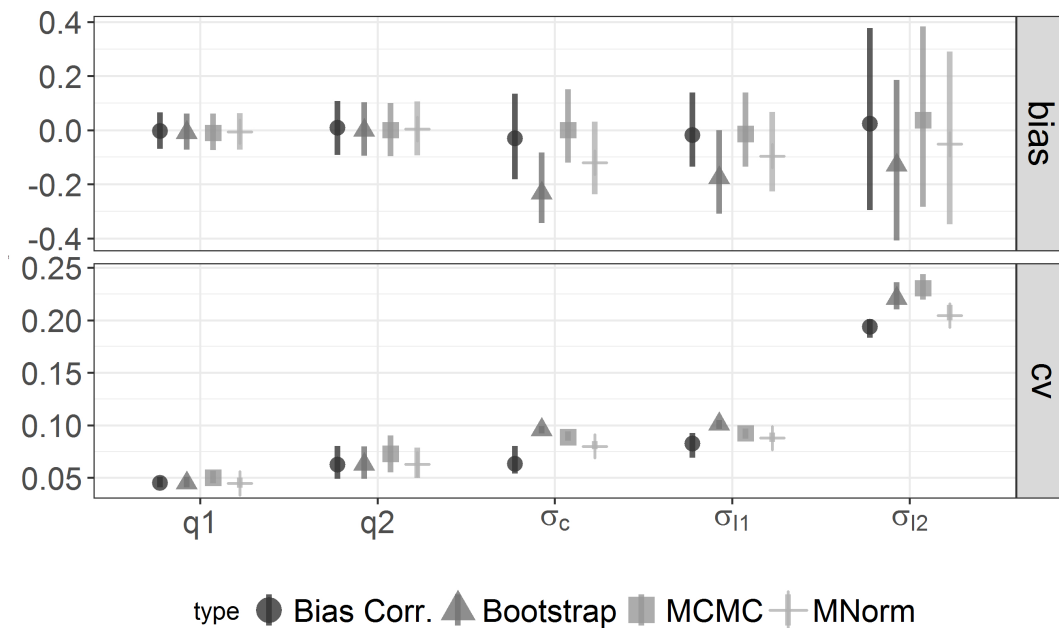


Figure 4.4: Medians (points) and 90% CIs (vertical bars) of the relative bias (top row) and the coefficient of variation (bottom row) of catchabilities and observation variances for scenario S2. Each point symbol and colour corresponds to an estimation approach.

4.8 for the whole time series of SSB, recruitment and fishing mortality parameters. Comparisons of coverage probabilities that are not shown in the table (for the rest of the years) are similar to the presented values. For biased estimations, such as observation variances, coverage probabilities are very low, mostly below 0.5, except

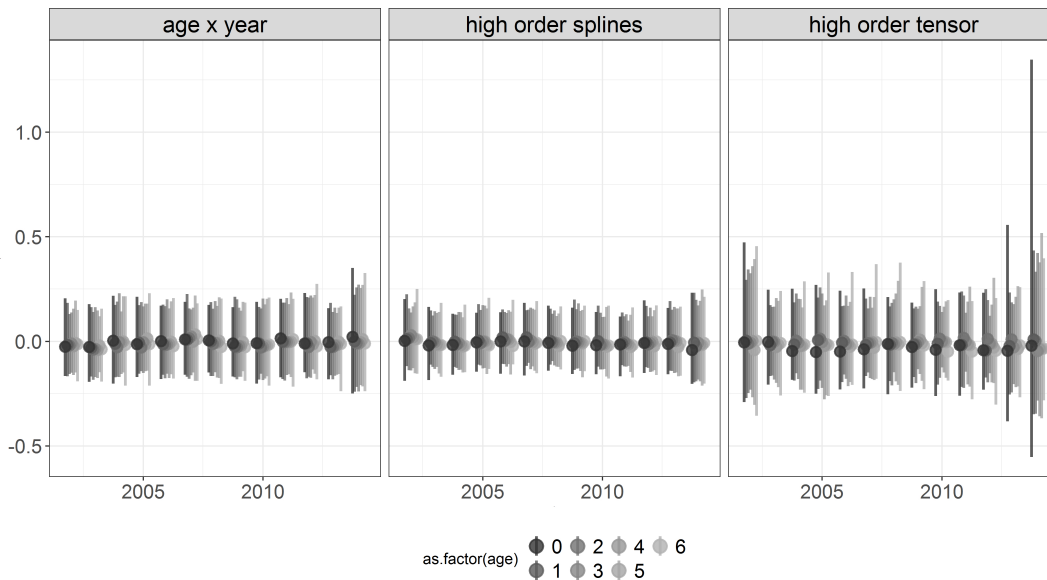


Figure 4.5: Medians (points) and 90% CIs (vertical bars) of the relative bias fishing mortality for each year (x-axis) for the MCMC approach. From left to right scenarios age x year, high order splines, and high order tensor (S1, S2, and S4) are represented in each column. Each colour corresponds to an age.

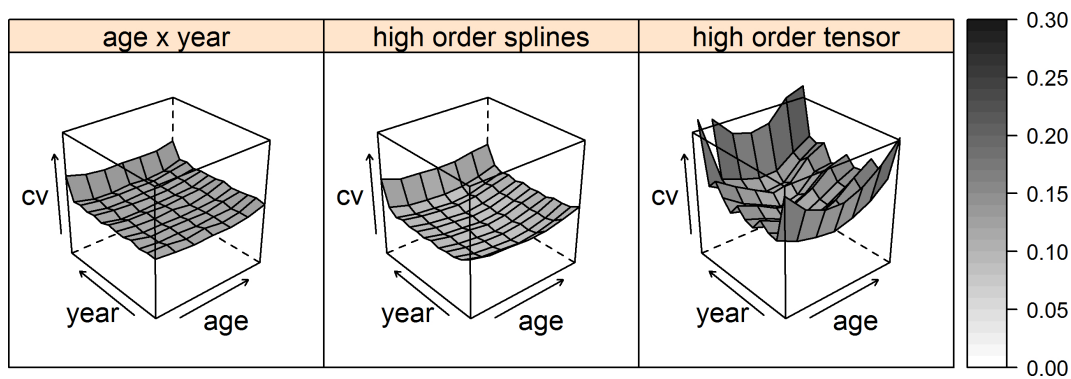


Figure 4.6: Median coefficients of variation of fishing mortality as a function of age and year for the MCMC approach. From left to right the panels correspond to scenarios age x year, high order splines, and high order tensor (S1, S2, and S4)

for the MCMC approach where values near 0.9 are reached.

When comparing across methods, the MCMC shows a more stable performance in face of more complex F models. The results obtained were always close to 0.9, including when the F model complexity increased to a tensor product (S4). The other methods show a deterioration of the coverage in S4 when comparing with S1 and S2.

Concerning AIC and BIC accuracy evaluation for model selection, in the case of AIC selection criteria, the scenario with the smallest number of parameter show the lowest model selection accuracy, while the scenario with the highest number of parameters, low order splines scenario, has been selected correctly with an accuracy of 100% (Figure 4.9). The rest of scenarios present an accuracy of around 75-85%. For BIC selection criteria, all scenarios except the classical separable one, were selected correctly 100% of the times. For the classical separable scenario, the scenario with high order splines was selected 75% of the times. This scenario is the most similar one in terms of number of parameters and separability assumption. In misselection cases AIC and BIC values appear to be very close.

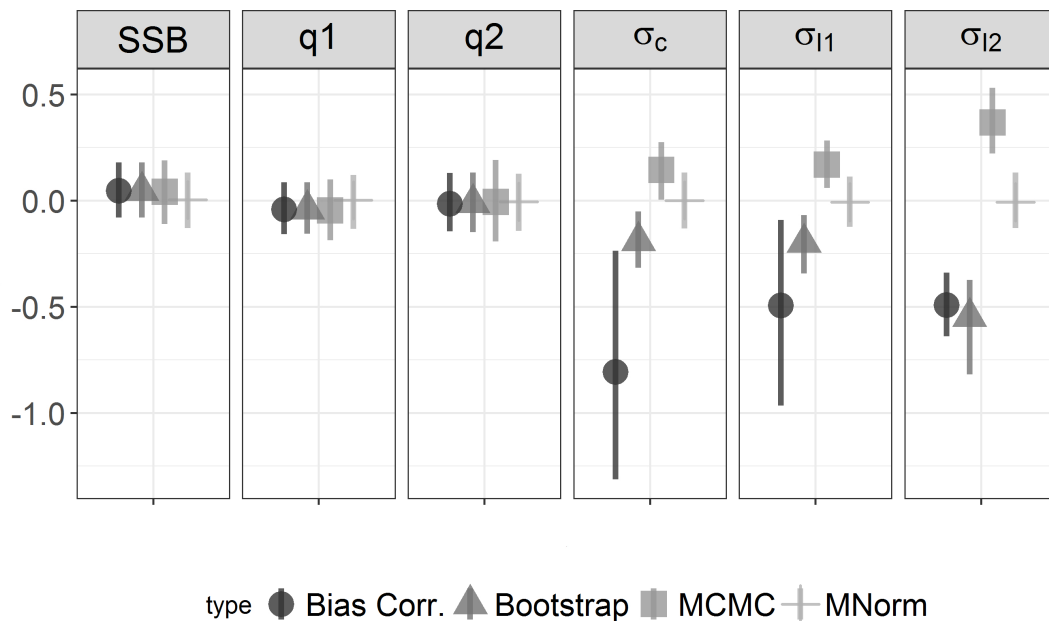


Figure 4.7: Medians (points) and 90% CIs (vertical bars) of the skewness SSB in the last year, catchabilities and observation variances for scenario S2. Each point symbol and colour corresponds to an estimation approach.

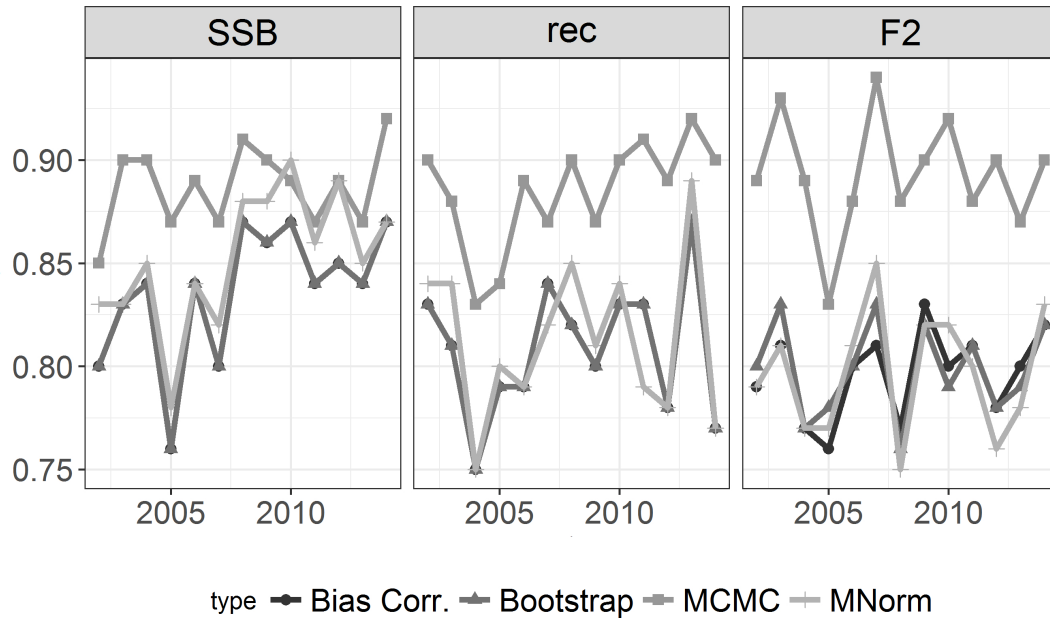


Figure 4.8: Coverage probabilities for each year and each method for SSB, recruitment (rec) and fishing mortality-at-age 2 (F2) parameters in scenario S4.

Table 4.2: Coverage probabilities for each scenario and each estimation method for catchabilities, observation standard deviations and for SSB, R, fishing mortality-at-ages 0, 2, and 6+ in the last year.

| Scenario | Method | SSB | R | f_0 | f_2 | f_6 | Q_{ac} | Q_{egg} | σ_C | σ_{Iac} | σ_{Iegg} |
|----------|---------------------|------|------|-------|-------|-------|----------|-----------|------------|----------------|-----------------|
| S1 | Relative bias corr. | 0.89 | 0.80 | 0.77 | 0.83 | 0.85 | 0.86 | 0.86 | 0.24 | 0.86 | 0.79 |
| S1 | Bootstrap | 0.89 | 0.80 | 0.78 | 0.82 | 0.83 | 0.87 | 0.86 | 0.00 | 0.54 | 0.75 |
| S1 | MCMC | 0.91 | 0.89 | 0.88 | 0.89 | 0.86 | 0.92 | 0.91 | 0.92 | 0.89 | 0.83 |
| S1 | MNorm | 0.91 | 0.80 | 0.78 | 0.82 | 0.83 | 0.87 | 0.88 | 0.35 | 0.75 | 0.78 |
| S2 | Relative bias corr. | 0.86 | 0.85 | 0.90 | 0.88 | 0.91 | 0.92 | 0.90 | 0.54 | 0.79 | 0.83 |
| S2 | Bootstrap | 0.86 | 0.85 | 0.90 | 0.89 | 0.92 | 0.92 | 0.93 | 0.14 | 0.40 | 0.75 |
| S2 | MCMC | 0.92 | 0.90 | 0.97 | 0.92 | 0.94 | 0.93 | 0.94 | 0.91 | 0.93 | 0.90 |
| S2 | MNorm | 0.87 | 0.85 | 0.92 | 0.88 | 0.92 | 0.92 | 0.91 | 0.50 | 0.65 | 0.82 |
| S3 | Relative bias corr. | 0.85 | 0.84 | 0.88 | 0.89 | 0.88 | 0.91 | 0.92 | 0.89 | 0.73 | 0.85 |
| S3 | Bootstrap | 0.85 | 0.84 | 0.88 | 0.89 | 0.86 | 0.91 | 0.92 | 0.37 | 0.38 | 0.74 |
| S3 | MCMC | 0.89 | 0.88 | 0.91 | 0.90 | 0.90 | 0.93 | 0.95 | 0.91 | 0.93 | 0.89 |
| S3 | MNorm | 0.85 | 0.85 | 0.88 | 0.89 | 0.89 | 0.91 | 0.93 | 0.74 | 0.60 | 0.82 |
| S4 | Relative bias corr. | 0.87 | 0.77 | 0.74 | 0.82 | 0.90 | 0.88 | 0.89 | 0.05 | 0.76 | 0.84 |
| S4 | Bootstrap | 0.87 | 0.77 | 0.73 | 0.82 | 0.89 | 0.89 | 0.90 | 0.00 | 0.38 | 0.74 |
| S4 | MCMC | 0.92 | 0.90 | 0.88 | 0.90 | 0.94 | 0.90 | 0.90 | 0.84 | 0.93 | 0.88 |
| S4 | MNorm | 0.87 | 0.77 | 0.75 | 0.83 | 0.90 | 0.87 | 0.90 | 0.18 | 0.64 | 0.80 |
| S5 | Relative bias corr. | 0.90 | 0.87 | 0.86 | 0.86 | 0.81 | 0.87 | 0.93 | 0.78 | 0.76 | 0.85 |
| S5 | Bootstrap | 0.90 | 0.87 | 0.86 | 0.86 | 0.83 | 0.87 | 0.93 | 0.29 | 0.35 | 0.70 |
| S5 | MCMC | 0.92 | 0.90 | 0.89 | 0.89 | 0.88 | 0.92 | 0.94 | 0.90 | 0.94 | 0.89 |
| S5 | MNorm | 0.90 | 0.85 | 0.86 | 0.87 | 0.80 | 0.88 | 0.93 | 0.62 | 0.61 | 0.83 |

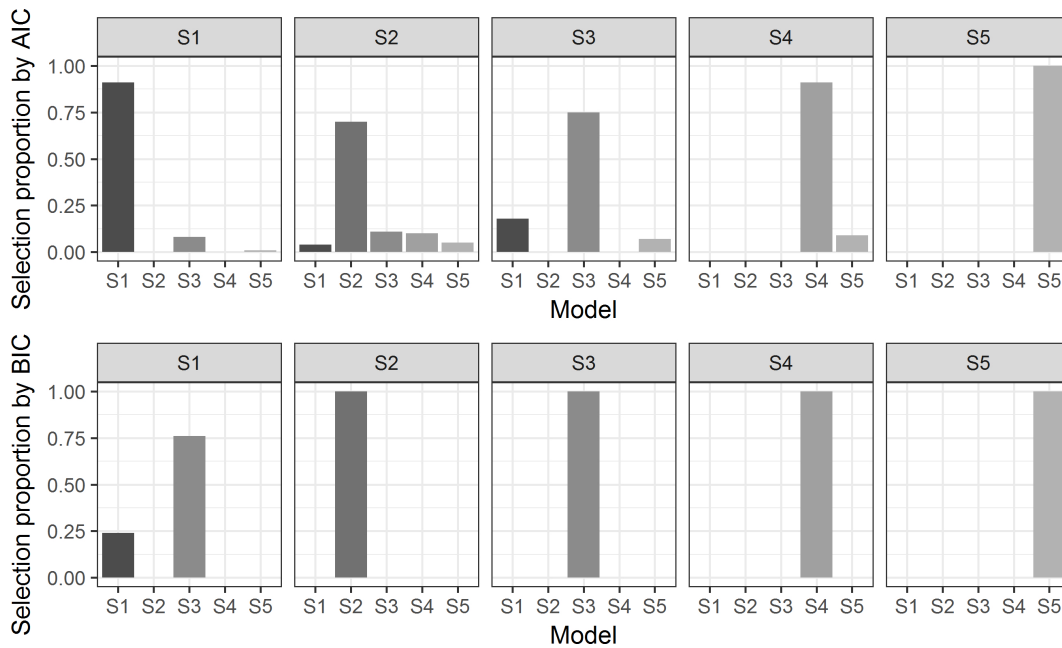


Figure 4.9: Selections proportion when using AIC (top) and when using BIC (bottom). Each panel corresponds to a scenario from which data was generated and x-axis represents the range of eligible models.

4.4 Discussion

The interest in incorporating uncertainty in management processes is increasing given that the lack of consideration of different sources of uncertainty has caused many failures in fisheries management (Hilborn and Peterman 1996). Thus, it becomes crucial to quantify correctly the uncertainty in fisheries assessments. In this work different methodologies were compared, focusing on estimation uncertainty. The work is focused on the introduction of non-parametric functions to model fishing mortality in stock assessment models. In addition, the accuracy of AIC and BIC in model selection has been analysed.

General results from the comparison of methods agreed with previous studies: Magnusson et al. (2012) identified MCMC method as the most reliable one and Stewart et al. (2012) suggest that MLEs generally misestimate skewed distributions in the

tails which does not happen when using MCMC methods. As in Magnusson et al. (2012) the bias correction algorithm for bootstrapping resulted in an improvement of results. However, in comparison to Magnusson et al. (2012) results, in this study it was not detected a systematic underestimation of coverage probabilities for MCMC method.

Parametric models are sometimes considered too restrictive to capture the complex dynamics of fish populations and the fleets exploiting them (Hillary 2012). In Fronczyk et al. (2011) a Bayesian non-parametric approach based on a mixture model for the joint distribution of log-reproductive success and stock biomass is proposed and compared with simpler parametric and semi-parametric models for North Atlantic cod data. They conclude that the non-parametric model outperforms the simpler ones. In Hillary (2012) a Ricker model was compared with a non-parametric alternative and found the latter had better performance.

Most recent works have focused on introducing more flexibility in fishing mortality estimates using different approaches. In Fernandez et al. (2010) autocorrelation processes are used to model a more flexible selectivity and in Nielsen and Berg (2014) the incorporation of a correlation parameter to capture the temporal smooth development in selectivity is proposed. In these approximations a single parameter is included and is estimated from the data instead of pre-specifying the smoothness degree. Other works, such as Thorson and Taylor (2014) or Martell and Stewart (2013) have studied the usage of non-parametric functions to model selectivity in assessment models, finding in the first case, that non-parametric models have less relative bias and greater precision when the parametric function is misspecified. In the second case the authors suggested that when there is no precise knowledge about the fishery or catch data, adopting a flexible selectivity, such as an age-based selectivity interpolated over age and year using a bicubic spline, may be more appropriate than assuming constant selectivity.

Some available software for stock assessment include options for non-parametric modelling. Stock Synthesis (Methot and Wetzel 2013) includes a non-parametric submodel for size selectivity, using waypoints and a set of linear segments while MULTIFAN-CL (Fournier et al. 1998) uses cubic splines to model selectivity.

As mentioned above, non-parametric functions, splines and their tensor products in this study, present advantages in terms of the ability to capture flexible shapes in comparison with parametric options. It has been seen that more flexible shapes result in greater uncertainties, finding greater CVs and relative bias at terminal years and at terminal ages for most complex scenarios, probably due to a larger number of estimated parameters. Coverage probabilities were also found to deteriorate for the most complex model, although not for MCMC method. As in our case study, the extra freedom in non-parametric models used in Hillary (2012) resulted in higher uncertainty level. The differences in uncertainty quantification could have an effect in management advice and risk assessment when evaluating the consequences of different management actions under uncertainty.

The changing level of uncertainty estimated for last periods related to the degree of flexibility in submodels could be handled using model selection. In this work the most commonly used criteria (AIC and BIC) have been evaluated in terms of accuracy in selecting the correct model resulting in a better accuracy for BIC. Thorson et al. (2013) propose a stepwise model selection using AIC to select the degree of smoothness for time-varying parameters and Maunder and Harley (2011) state the need for an alternative to AIC and BIC criteria and perform cross validation model selection to determine non-parametric selectivity curves. An alternative to selecting a single model is performing model averaging, which eliminates the need for selecting a ‘best’ model and rejecting all alternative assumptions as proposed in Millar et al. (2014) and references therein. Model selection or model averaging are not very common practices in stock assessment yet, although the incorporation of these kind of approaches is an issue with an increasing interest (Anderson et al. 2017).

Chapter 5

Development of a full-feedback MSE for the Bay of Biscay sardine

In this chapter we develop a full-feedback Management Strategy Evaluation(MSE) for the Bay of Biscay sardine. We start from a simple Monte Carlo simulation approach, where only process uncertainty is considered, and we introduce gradually assessment, observation and initial population uncertainties. We study the effects of each of these uncertainty sources in the MSE process in terms of the main performance statistics, such as biological risk, expected biomass or catches, and we discuss their impact in the development of a management plan for this stock.

5.1 Introduction

The aim of performing MSE for fish stocks is to evaluate by Monte Carlo simulation the effect of different management actions before they are put in place. This allows to assess their performance with respect to the management objectives and to select strategies that are robust to different sources of uncertainty (Punt et al. 2016).

The success of a MSE depends, among others, on how well uncertainty is represented (Butterworth and Punt 1999, Punt et al. 2016). However, identifying and deciding on what are the specific sources of uncertainty that will be included in the MSE remains a major challenge. Kraak et al. (2010) claimed that the choice of uncertainty sources in MSE often seems to be quite arbitrary and does not necessarily reflect a lack of scientific knowledge of the relevant process. From the experience gained in the development of a management procedure for Southern Bluefin Tuna, Kolody

et al. (2008) recognized that there is no simple prescription for quantifying uncertainty and commended efforts to ensure that a broad range of uncertainty sources are included.

Punt et al. (2016) distinguished five sources of uncertainty in a MSE (as listed in the introduction in section 1.3.3) and considered that a MSE should include minimally process uncertainty, parameter uncertainty and observation error. Examples incorporating process uncertainty can be found in Irwin et al. (2008) which used a multinomial distribution for recruitment proportions, in Punt et al. (2008) where temporally autocorrelated error structures were used or in Garcia et al. (2013) through an autoregressive recruitment model with normally distributed uncertainty. Concerning parameter uncertainty examples, it was included in the stock–recruitment relationship through a Bernoulli distribution for "high" recruitment probability and a multivariate normal distribution for Ricker parameters in Irwin et al. (2008) while the posterior distributions of the Bayesian statistical catch-at-age assessment model were taken in Garcia et al. (2013) to include parameter uncertainty in the initial population. Observation error was included in Needle (2008) through lognormal errors for catches and survey data, as well as in Garcia et al. (2013) who added observation error in the ageing by using a multinomial distribution. Implementation uncertainty was also found in Punt et al. (2008) or Irwin et al. (2008) where normally distributed error were included.

Concerning the assessment process for a full-feedback approach, Garcia et al. (2013) did incorporate the full assessment in some scenarios for the *Pagellus bogaraveo* case study, where XSA model (Shepherd 1999) was applied to observed data within the MSE Monte Carlo simulation. The same assessment model was also incorporated in Needle (2008). Other stock assessment models have been also used in MSE works, such as SAM model for North Sea cod (ICES 2020a), Stock Synthesis for West Coast groundfish (Punt 2003) or ad-hoc models such as the age-structured assessment model developed with ADMB in A'mar et al. (2009). This last work, reports a bias between the true SSB and the estimated by the assessment model, which is another issue to be taken into account when a full-feedback approach is implemented. Thus, when evaluating the performance of a "full feedback" MSE, statistics such as the bias or mean square error of the assessment model need to be included, in order to measure how well the stock assessment method estimates quantities of interest for management (Patterson and Kirkwood 1995, Punt et al.

2001). For data-limited case studies several works have reported biased results in stock assessment models (Chen et al. 2003, Wetzel and Punt 2011). Not reliable catch data, due to discards or misreporting issues, is considered as a source of bias in assessment (Pomarede et al. 2010) and fishery-independent, survey-based management procedures have been proposed as an alternative (De Oliveira and Butterworth 2004, Pomarede et al. 2010). However, none of them explores explicitly the effect of introducing such sources of uncertainty on the management actions' performance.

Given the intensive computational requirements for a full-feedback approach, alternative shortcut options are usually implemented which try to mimic the stock assessment model error. However there is a lack of guidance on how to select the appropriate values for the common approach of using a stochastic process with an assumed level of autocorrelated estimation error. Wiedenmann et al. (2015) analysed the implications of this autocorrelation for MSE. Shortcut approaches can be also achieved by adding estimation error based on the variance covariance matrix of a reference assessment model or using analytical retrospective analysis (ICES 2020a).

The main objective of this work is to develop a full-feedback MSE for sardine in the Bay of Biscay, while quantifying the impact of each of the different sources of uncertainty introduced in the MSE process. The operating model has been conditioned according to the population and fleet dynamics of sardine in the Bay of Biscay. Two different types of harvest control rules were considered: the ICES advice rule and a catch-based rule. Process, parameter, observation and assessment uncertainties were included one by one, in order to analyse their effect separately. This allowed us to identify the major sources of risk and discuss future steps towards the final adoption of a management plan for this stock.

5.2 Material and methods

5.2.1 Operating model

The settings for representing the real dynamics for the Bay of Biscay sardine case study were based on the 2019 stock assessment model (ICES, 2019a) using SS (Stock Synthesis; Methot and Wetzel, 2013). This model assumes a single area, a single fishery, a unique season by year and both genders combined. Spawning stock biomass (SSB) is computed on 1st January. The population is age-structured with

age classes from age 0 to 6+ (the plus group is set at age 6) and the assessment period goes from 2000 to 2019, being the later an interim year in the assessment (so only provisionally assessed). See section "SS approach" in Chapter 3 for a detailed description of the SS stock assessment model settings. In the MSE framework, the historical part was equal to the assessment period (2000-2019), from which estimates of numbers-at-age, fishing mortality-at-age, catch-at-age in numbers and total catch in tonnes were computed as the medians of resulting MCMC iterations from the SS stock assessment model. Then, the population was projected forward for 25 years (2020-2044).

5.2.1.1 Population dynamics

Annual recruitment R_y was modelled as a function of spawning stock biomass SSB_y in year y , according to a segmented regression stock-recruitment model with the breakpoint fixed at B_{lim} (56,300 tonnes from IBPSardine: ICES, 2019b). The segmented regression model was considered the most appropriate model for the simulations as it makes the least assumptions about the shape of the curve out of the observed ranges. It is formulated as a piecewise linear regression (see Figure 5.1) which must pass through the origin and take a constant value after the breakpoint as follows:

$$R_y = \begin{cases} aSSB_y & SSB_y < B_{lim}, \\ aB_{lim} & SSB_y \geq B_{lim}, \end{cases} \quad (5.1)$$

where the slope parameter a is estimated from the historical period. In our case study $a = 87$, which corresponds to the median value across all a parameters estimated for each of the MCMC iterations from the SS assessment model.

Regarding the rest of biological parameters, natural mortality-at-age was set constant along time at the same values as in the historical series, as agreed in the benchmark (ICES, 2017a). Maturity-at-age was set to the average of historical values rounded off to two decimals (Table 5.1). Weights-at-age for the stock show a decreasing trend in the historical period (ICES 2019b, Véron et al. 2020). Thus, for the projection period, weight-at-age for the stock were assumed to be constant and equal to the average of the last 5 years (2015-2019) (Table 5.1). Similarly, given the

decreasing trend of the weight-at-age in the catch for all ages in the historical series, weight-at-age in the catch were taken as the average of the period 2014-2018 (Table 5.1). In this case, weight-at-age in the catch in 2019 were not used given that in the assessment year, weight-at-age in the catch are not known and are assumed to be the average over the last three years.

Table 5.1: Maturity-at-age proportion, stock weight and catch weight-at-age values in kg used for projection years.

| age | maturity | stock weight | catch weight |
|-----|----------|--------------|--------------|
| 0 | 0.00 | 0.00 | 0.02 |
| 1 | 0.69 | 0.02 | 0.04 |
| 2 | 1.00 | 0.04 | 0.05 |
| 3 | 1.00 | 0.05 | 0.06 |
| 4 | 1.00 | 0.06 | 0.06 |
| 5 | 1.00 | 0.07 | 0.07 |
| 6 | 1.00 | 0.08 | 0.08 |

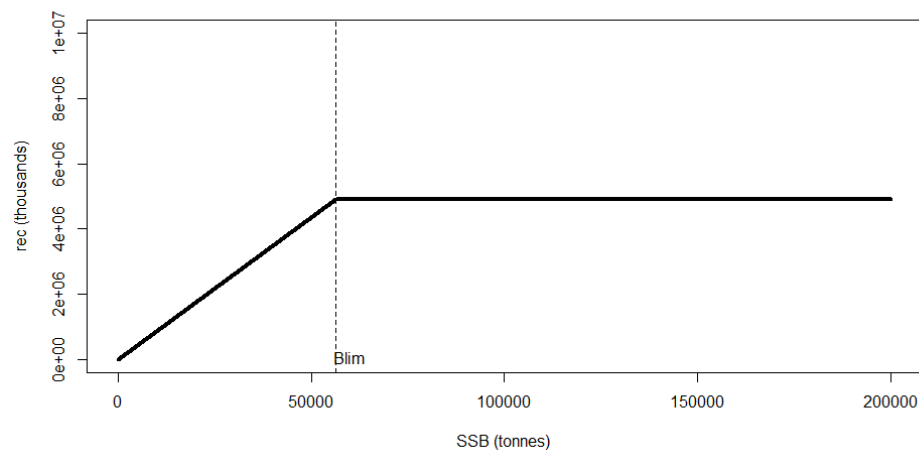


Figure 5.1: Graphical representation of the implemented segmented regression stock-recruitment model.

5.2.1.2 Fleet dynamics

A single fleet with one métier was considered for the Bay of Biscay sardine case study. This represented the international catches by France and Spain, the two main countries participating in this fishery. Given that no effort data were available, effort was fixed to 1 along the past historical period (2000-2019). Discards in this fishery are negligible, so they were assumed to be equal to zero and catches were equal to the landings.

The Cobb-Douglas production function (Cobb and Douglas, 1928, Clark, 1990) establishes the annual catches as a function of effort and biomass as follows:

$$C_{a,y} = q_{a,y}^{CD} E_y^\alpha B_{a,y}^\beta, \quad (5.2)$$

where $C_{a,y}$ stands for catches in year y and age a , q_a^{CD} is the catchability-at-age parameter, E represents the effort and B the biomass. The catchability-at-age parameters for this function, q_a^{CD} , were calculated yearly (assuming $\alpha = \beta = 1$) in the historical period as the ratio between catch and biomass by age at the middle of the year, when catches were assumed to be taken instantaneously. For the simulation period, yearly catchability-at-age parameters for the Cobb-Douglas model were assumed to be constant along time and equal to the average over the last five years (2015-2019), $q_a^{CD} = (0.01, 0.20, 0.31, 0.39, 0.39, 0.38, 0.38)$. In the simulation period, based on the true biomass and the Cobb-Douglas parameters, the effort needed to get the TAC established by the harvest control rule is computed for each year.

5.2.2 Management procedure

Data needed for the stock assessment were generated from the operating model defined above and were used as input for the management procedure. The management procedure consisted of conducting the stock assessment, doing the short-term forecast and applying a harvest control rule for management advice. Subsequently, the Monte Carlo simulation was closed by subtracting the advised catches from the real population in the operating model.

5.2.2.1 Harvest control rules

Two types of harvest control rules (HCRs) were implemented to set the TAC for management year $y+1$ based on the expected SSB \widehat{SSB}_{y+1} that is obtained through the short-term forecast.

The first one is the ICES MSY advice rule (IcesHCR; ICES 2019a). This is an F-based rule that sets the fishing mortality rate of no more than F_{msy} , while maintaining the stock above B_{lim} with at least 95% probability. In particular, the TAC for year $y+1$ corresponds to a fishing mortality that is set as a function of \widehat{SSB}_{y+1} as follows (see Figure 5.2 for a graphical representation):

$$F_{y+1} = \begin{cases} 0, & \text{if } \widehat{SSB}_{y+1} < B_{lim}, \\ F_{msy} \frac{\widehat{SSB}_{y+1}}{MSY B_{trigger}}, & \text{if } B_{lim} < \widehat{SSB}_{y+1} < MSY B_{trigger}, \\ F_{msy}, & \text{if } \widehat{SSB}_{y+1} > MSY B_{trigger}, \end{cases} \quad (5.3)$$

where $F_{msy} = 0.453 \text{ year}^{-1}$, $MSY B_{trigger} = 78,700$ tonnes and $B_{lim} = 56,300$ tonnes are the reference points for this stock (ICES 2019c).

The second type of harvest control rule is a catch-based rule (CbasedHCRtacmax), where the TAC for year $y+1$ is set as a function of \widehat{SSB}_{y+1} as follows (see Figure 5.2 for graphical representation):

$$TAC_{y+1} = \begin{cases} 0, & \text{if } \widehat{SSB}_{y+1} < B_0, \\ TAC_{min} + \gamma(\widehat{SSB}_{y+1} - B_0), & \text{if } B_0 < \widehat{SSB}_{y+1} < B_1, \\ TAC_{max}, & \text{if } \widehat{SSB}_{y+1} > B_1. \end{cases} \quad (5.4)$$

This rule depends on five parameters (the biomass trigger points B_0 and B_1 , the minimum and maximum TAC allowed by the rule TAC_{min} and TAC_{max} , and the exploitation rate γ) that result in different rules. To ensure that the rule is continuous at the biomass trigger point B_1 , B_1 was defined as:

$$B_1 = B_0 + \frac{TAC_{max} - TAC_{min}}{\gamma}. \quad (5.5)$$

The biomass trigger points B_0 and B_1 were taken as B_{lim} and $MSY B_{trigger}$ respectively. The rest of the parameters (γ , TAC_{max} and TAC_{min}) can be fixed to

different values, which will define the specific shape of the HCR. Combining different values for all described parameters on the HCRs definition, lots of different rules could be set and tested. The analysis of comparing different rules is out of the scope of this work. In order to see the effect of introducing different sources of uncertainty on the MSE process, we focused on the particular catch-based rule with $TAC_{min} = 7,500$, $TAC_{max} = 35,000$ and $\gamma = 0.5$. This rule was selected from a previous work, where it showed a good performance in terms of risks without a full-feedback MSE approach (Uriarte et al. 2020).

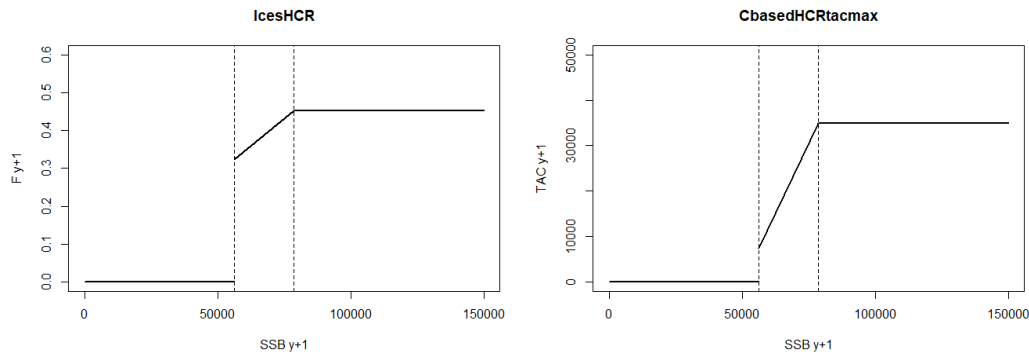


Figure 5.2: Graphical representation of the implemented harvest control rules (IcesHCR and CbasedHCRtacmax).

5.2.3 Uncertainty sources

The MSE framework allows to take into account the full range of uncertainty sources (Punt et al. 2016). In this case study we included the following sources of uncertainty: process uncertainty in the recruitment dynamics, initial population uncertainty through the uncertainty in the parameters that define the historical population, observation uncertainty in the catch and in the abundance indices and uncertainty due to the stock assessment model.

5.2.3.1 Process uncertainty

Uncertainty in the recruitment process was implemented by adding a lognormally distributed error (with zero mean and variance equal to the variance of the residuals from the model fit) around expected recruitment for each iteration.

5.2.3.2 Initial population uncertainty

Initial population uncertainty, referred as "parameter uncertainty" in Punt et al. (2016), was introduced by incorporating into the MSE conditioning the uncertainty of the parameters estimated by the stock assessment model that define the initial population and propagating these uncertainties to other derived quantities in the conditioning.

The official SS assessment was run activating the MCMC option, which in addition to point estimates, returns a value for each parameter for each MCMC iteration. This capability is inherited from ADMB (Monnahan et al. 2014) which samples the maximum likelihood surface using MCMC, with a multivariate normal distribution as the proposal function. As recommended by ADMB developers, MCMC results from ADMB were compared to another MCMC software (compared to JAGS in Chapter 3) obtaining consistent results for our case study model. A thin-in interval of 10 and a burn-in of 1000 iterations was set, obtaining a final set of 1000 iterations.

Output from each MCMC iteration was taken as a different population to condition the operating model. All the estimated parameters (population numbers-at-age, fishing mortalities, selectivities and indices' catchabilities) take the corresponding different values from each MCMC iteration. Stock recruitment parameter's estimation was done for each iteration as well as the computation of Cobb-Douglas catchability parameters. Thus, initial population uncertainty was also propagated to the conditioning process via other parameters.

5.2.3.3 Observation error

Observation error was introduced for the catch-at-age data and for the abundance indices needed to run the stock assessment model, namely, numbers-at-age and total biomass from PELGAS acoustic survey and egg count and DEPM (Daily Egg Production Method) SSB index from BIOMAN survey. For the historical period (2000-2019) catch-at-age and abundance indices were equal to the values estimated by the SS official 2019 assessment (WGHANSA2019). For the projection years (2020-2044) these indices were generated according to the same assumptions as in the assessment model as described below.

Catch-at-age were generated according to a multinomial observation error:

$$\widehat{C}_{0,\dots,6+,y} \sim \text{Multinomial}(n_C, CP_{0,y}, \dots, CP_{6+,y}) * \sum_{a=0}^{6+} C_{a,y}, \quad (5.6)$$

where $CP_{a,y} = \frac{C_{a,y}}{\sum_{a=0}^{6+} C_{a,y}}$ denotes the true catch proportion for age group a and year y and n_C is the sample size that determines the variability of the observation error and was set to the same value as used in the SS assessment model.

The aggregated abundance indices were generated according to a lognormal observation error centred at the true expected index as follows:

$$\widehat{I}_y^{tot} \sim LN(I_y^{tot}, \sigma^{tot}), \quad (5.7)$$

$$\widehat{I}_y^{egg} \sim LN(I_y^{egg}, \sigma^{egg}), \quad (5.8)$$

$$\widehat{I}_y^{depm} \sim LN(I_y^{depm}, \sigma^{depm}), \quad (5.9)$$

where \widehat{I} stands for the observed indices with error, I denote the true expected index and σ 's are the standard deviations of the indices which were taken from the input given to the SS assessment model: $\sigma^{tot} = 0.418$, $\sigma^{egg} = 0.448$, $\sigma^{depm} = 0.567$. Each of the expected indices I was the product between the catchability parameter ($Q^{tot} = 4.20$, $Q^{egg} = 43.40$, $Q^{depm} = 1.78$, assumed to be constant across years) estimated in the WGHANSA2019 assessment and the corresponding total abundance quantity (SSB, biomass or population numbers) as detailed in Chapter 3 section "SS approach".

The numbers-at-age indices from the acoustic survey were generated according to a multinomial observation error:

$$\widehat{I}_{1,\dots,6+,y}^{ac} \sim \text{Multinomial}(n_I, IP_{1,y}^{ac}, \dots, IP_{6+,y}^{ac}) * \sum_{a=1}^{6+} I_{a,y}^{ac}, \quad (5.10)$$

where \widehat{I} are index observed with error, $IP_{a,y}^{ac} = \frac{I_{a,y}^{ac}}{\sum_{a=1}^{6+} I_{a,y}^{ac}}$ are indices proportions for age a and year y and n_I is the sample sizes given to the SS assessment model, which determine here the variability of the observation error we are introducing to

age structured data.

Weight-at-age and maturity data were assumed to be observed without uncertainty. Observed total catch was also generated with no error, following the SS stock assessment model assumptions, which fits the data with nearly no error in total catch.

5.2.3.4 Stock assessment error

A full assessment process was implemented by the inclusion within the management procedure of the SS assessment model that is yearly used in the corresponding ICES working group WGHANSA (ICES 2019c). For each projection year in the MSE Monte Carlo simulation, input data for the assessment model was taken from the operating model, then the observation process (with or without error) was applied and fitting the SS assessment model an estimation of the parameters of interest was obtained and provided for the following steps in the loop.

Through this assessment process an estimated time series of SSB was obtained until year y (assessment year). In order to obtain the SSB for the management year ($y+1$) a short-term forecast of one year was carried out. The short-term forecast followed the same approach as established in the ICES working group (ICES 2019c). The biological parameters for the forecast were taken as the average of the last three observed years, i.e., from year $y - 3$ to year $y - 1$. Concerning the fishing mortality, catches in year y were assumed equal to the advised TAC for this year. The resulting \widehat{SSB}_{y+1} was then used for the HCR in the management procedure.

5.2.4 Simulations and software

Simulations were carried out using the FLBEIA (García et al., 2017), projecting the population forward for 25 years (2020-2044).

Uncertainty sources were introduced gradually starting from a process with perfect observation including only uncertainty on recruitment (R). Initial population uncertainty (I), observation (O) and full assessment error (F) were introduced separately in different steps to evaluate their individual effects. Finally, in the last step, all uncertainty sources were included. Therefore, the following scenarios were defined:

- R: Recruitment (base case)
- RI: Recruitment + initial population
- RF: Recruitment + full assessment
- RFO: Recruitment + full assessment + observation error
- RIFO: Recruitment + initial population + full assessment + observation error

1000 iterations were run for each of the listed scenarios (R, RI, RF, RFO, RIFO).

5.2.4.1 Developed code

As the `FLBEIA` package does not incorporate the option to implement a full-feedback MSE including the SS assessment model, a specific R function was developed to use SS inside the Monte Carlo simulations. The function calls SS and writes the input and output files in the needed formats, transforming FLR objects into SS files and vice versa. All files used for a SS assessment need to be provided, from where data, control and additional files are read using the `r4ss` (Taylor et al. 2019) package. For each projection year observed new catch and indices data are included into the files and the SS executable is called. Apart from saving the estimated SSB and F in the last year for the management procedure part, estimated catchabilities, selectivities recruitment, fishing mortalities and SSB complete time series from SS runs are saved in the "covars" component of the OM, in order to check the performance of the assessment process (i.e. detection of biased assessment).

Part of the developed code was added in a tutorial within the `FLBEIA` documentation (https://flr-project.org/doc/FLBEIA_Incorporating_SSassessment_MP.html). This function is specific for our case study and should be adapted to be used for another stock.

A new function to incorporate multinomial observation error was also developed. Nowadays `FLBEIA` only incorporates functions for multiplicative errors, which need to be fixed at the beginning of the process for every projection year. In order to implement multinomial error for age-structured data, it is necessary to generate such error in each projection year and cannot be generated at the beginning. The developed code allows introducing this type of error and reading the age-structure of

the true population or catch and sampling from a multinomial distribution in each projection year. The sample size to be used for each dataset is read from SS files and fixed in the conditioning part as a covariate.

5.2.5 Performance indicators

In order to assess the effect of the inclusion of different sources of uncertainty in the MSE process, and to evaluate the performance of each HCR, several performance indicators were defined related to both biological status and exploitation status. In particular, we computed the median catch and median SSB across projection years ($y > 2019$) and iterations ($i = 1, \dots, N_{iter}$):

$$Median_catch = med(C_{y,i}), \quad (5.11)$$

$$Median_SSB = med(SSB_{y,i}). \quad (5.12)$$

Average standard deviations across projection years and iterations for catch and SSB:

$$Mean_sd_catch = \frac{\sum_{i=1}^{N_{iter}} \sqrt{\frac{\sum_{y=2020}^{2044} (C_{y,i} - \bar{C}_i)^2}{N_{yproj} - 1}}}{N_{iter}}, \quad (5.13)$$

$$Mean_sd_SSB = \frac{\sum_{i=1}^{N_{iter}} \sqrt{\frac{\sum_{y=2020}^{2044} (SSB_{y,i} - \bar{SSB}_i)^2}{N_{yproj} - 1}}}{N_{iter}}. \quad (5.14)$$

Risk of type 3 or maximum probability of the SSB falling below B_{lim} in the projection period:

$$Risk3 = max \left(\frac{\sum_{i=1}^{N_{iter}} I(SSB_{y,i} < B_{lim})}{N_{iter}} \right). \quad (5.15)$$

Average probability of fishery closure in any year of the projection period:

$$Closure = \frac{\sum_{i=1}^{N_{iter}} \sum_{y=2020}^{2044} I(TAC_{y,i} = 0)}{N_{iter} N_{yproj}}, \quad (5.16)$$

where I is an indicator function that takes the value 1 if the condition within the brackets is fulfilled and 0 otherwise and N_{yproj} denotes the number of years in the projection period.

Apart from that, the coefficient of variation (CV) of SSB, total catch and recruitment was computed across all iterations for each year and bias of the assessment output from the real population was also measured.

5.3 Results

5.3.1 Base case

In the base case (R), when only recruitment uncertainty is included in the MSE process, median SSB stabilises around 89,500 tonnes for IcesHCR and around 94,200 for CbasedHCRtacmax, whereas median catch reaches around 26,800 tonnes for IcesHCR and around 25,200 tonnes for CbasedHCRtacmax (Figure 5.3). In most of the projection period, catch is smaller in CbasedHCRtacmax than in IcesHCR. As a result, SSB is larger in CbasedHCRtacmax than in IcesHCR. Uncertainty around SSB and catch is represented by the interval conformed by the 5th and 95th quantiles out of the performed 1000 iterations. Uncertainty around SSB is similar in both rules, as both intervals have similar width. On the contrary, the upper limit of the total catch interval in CbasedHCRtacmax is trimmed due to the maximum TAC imposed by this rule.

The performance indicators allow to summarise the behaviour of each rule and scenario along the projection period (Figure 5.4). Median catches along the whole projection period for IcesHCR and CbasedHCRtacmax are 27,200 and 25,600 tonnes respectively, while median SSB along the projection period are 90,100 and 94,700 tonnes respectively. Average standard deviation values are similar in both rules for SSB and catch, being the uncertainty in catch slightly larger in CbasedHCRtacmax than in IcesHCR. The risk of being below B_{lim} is 0.55 for IcesHCR and 0.025 for CbasedHCRtacmax. This means that in the base case scenario (R), only CbasedHCRtacmax is precautionary according to the ICES principles (Risk3 <0.05). Concerning the fishery closures, the probability of closure is larger for IcesHCR (0.009) than for CbasedHCRtacmax (0.001).

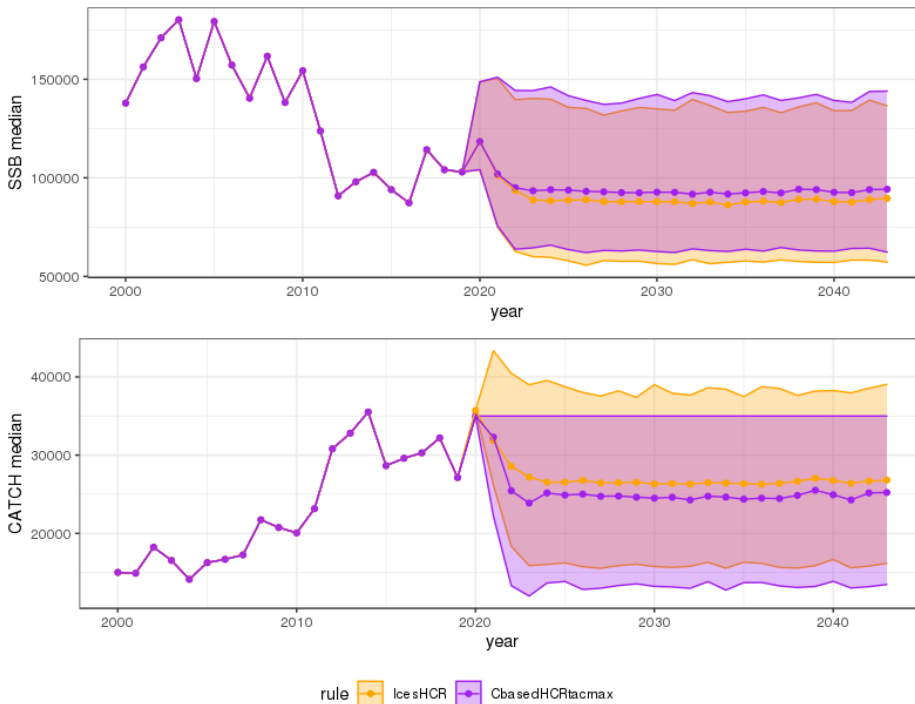


Figure 5.3: Median SSB (in the top) and total catch (in the bottom) along with 5th and 95th quantiles (shaded ribbon) for historical and projection years for the two harvest control rules (IcesHCR and CbasedHCRtacmax) in the base case scenario (R).

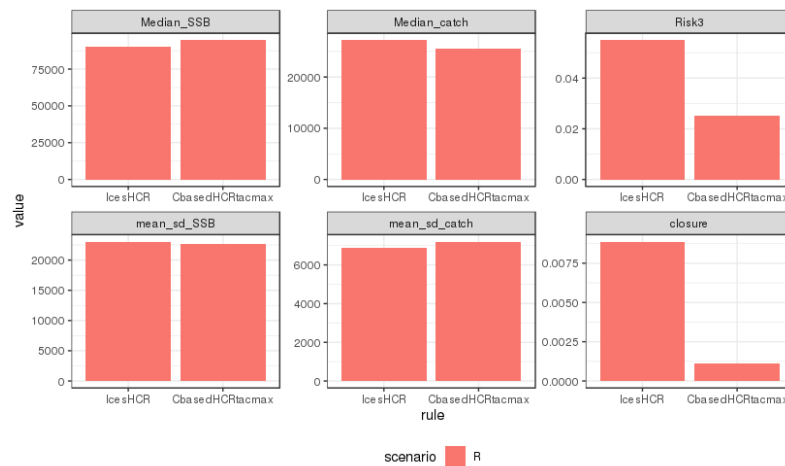


Figure 5.4: Performance indicators for the base case (R) for the two harvest control rules (IcesHCR and CbasedHCRtacmax).

5.3.2 Initial population uncertainty

Initial population uncertainty through the MCMC output of the SS assessment model was propagated along the conditioning of the operating model. This resulted in additional uncertainty in some of the parameters used in the projection period, such as the parameters of the stock-recruitment model or the catchability parameters of the Cobb-Douglas model. In the base case (R), the median value of the slope parameter of the hockey-stick stock-recruitment model was 87. When incorporating the initial uncertainty (scenario RI) this value ranged from 72 to 108, with 5th and 95th quantiles at 79 and 98 (Figure 5.5). Uncertainty of the catchability parameters of the Cobb-Douglas model in the historical period are shown in Figure 5.6. Given the catch-at-age structure of the assessment model, this parameter has an increasing shape from age 0 to age 3 and a constant shape for ages above 3. These catchability values show an increasing trend along time, with values around 0.17 during the first historical years to values around 0.4 in the most recent years.

When incorporating uncertainty on initial population (scenario RI), SSB and catches along the projection period become more uncertain, showing wider intervals around median values for both rules (Figure 5.7). In terms of performance indicators, median SSB and catch values are similar to the base case (scenario R) for both rules. However, the additional uncertainty in the RI scenario results in higher risk of being below B_{lim} (from 0.055 to 0.062 for IcesHCR and from 0.025 to 0.026 for Cbased-HCRtacmax) and higher probability of closure (from 0.009 to 0.013 for IcesHCR and from 0.001 to 0.003 for CbasedHCRtacmax) (Figure 5.11). In terms of variability, the standard deviations of SSB and catches along years increase slightly (Figure 5.11), while coefficients of variation for SSB, catches and recruitment, computed across iterations for each year, increase compared to the base case (Figure 5.12). These effects are similar for both rules.

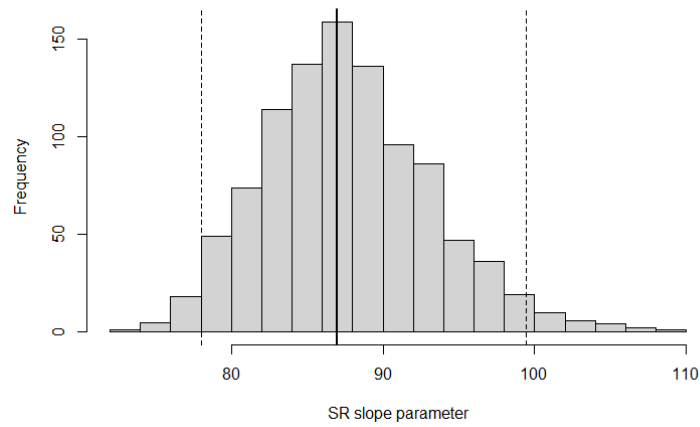


Figure 5.5: Histogram for the segmented regression S-R model's slope parameter when initial population uncertainty is introduced. The solid line represents the median and the dashed lines 5th and 95th quantiles.

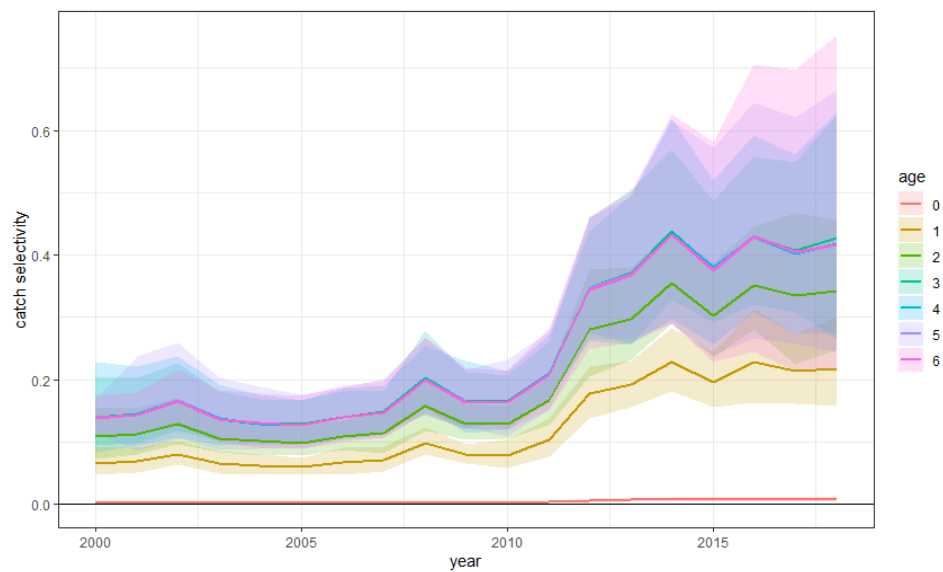


Figure 5.6: Catchability-at-age parameters of the Cobb-Douglas model during the historical period when initial population uncertainty is introduced (RI scenario). The solid lines represent the median values and the shaded ribbons the 5th and 95th quantiles.



Figure 5.7: SSB (top row) and total catch (bottom row) median values along with 5th and 95th quantiles (shaded ribbon) for historical and projection years for the two harvest control rules (in columns) in the base case (R) and in the RI scenarios.

5.3.3 Assessment and observation error

When the SS assessment model is incorporated in the MSE process (scenario RF), a bias effect has been found. Compared to the real SSB in the operating model, the SSB estimated by the SS assessment model within the MSE process for each projection year is negatively biased during the first years of the projection period and positively biased afterwards (Figure 5.8). Thus, at the beginning of the projection period SSB is underestimated by the stock assessment, resulting in a lower TAC in the RF scenario than in the base case scenario (R), that makes the true SSB to be higher (Figure 5.10). On the contrary, during the rest of the projection years, SSB is overestimated by the assesment model, resulting in a higher TAC than that corresponding to the real SSB (R scenario), and subsequently giving a lower SSB than in the base case (Figure 5.10).

Performance indicators show that catch and SSB median values in the RF scenario are similar compared to the base case (R scenario), while the risk of being below B_{lim} gets two times larger (Figure 5.11). Uncertainty around median SSB gets higher,

both in terms of standard deviation across across years (Figure 5.11) and in terms of computed CV values across iterations (Figure 5.12). However, when looking at total catch, uncertainty decreases with respect to the base case, both in terms of computed standard deviations and coefficient of variations (Figures 5.11 and 5.12). This unexpected decrease of the uncertainty on total catches when introducing the assessment process was due to a lower uncertainty in the SSB estimated by the SS stock assessment model in the assessment year when compared to the real SSB from the operating model (Figure 5.13). As a result, the advised TAC based on the SSB estimated in the stock assessment present lower uncertainty than the advised TAC when no assessment model is used.

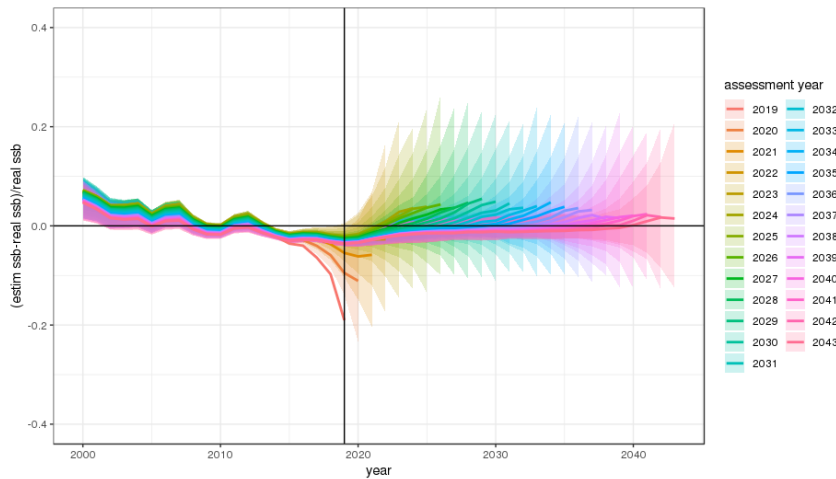


Figure 5.8: Assessment median bias (lines) along with 5th and 95th quantiles (shaded ribbon) for real SSB for each assessment year (colours) for the CbasedHCRtacmax and no observation and initial population uncertainty (RF scenario).

Besides the assessment error, when introducing also the observation error (scenario RFO), very similar effects are obtained in terms of bias in SSB, with higher absolute median bias values and larger uncertainty (Figure 5.9). The trajectories of SSB and catch have wider intervals around medians, meaning that the resulting SSB and catch values are more uncertain (Figure 5.10). This is also reflected in larger computed standard deviations and CVs (Figures 5.11 and 5.12 respectively). The risk of being below B_{lim} increases around 0.03 for both rules and the probability of closure increases around 0.003.

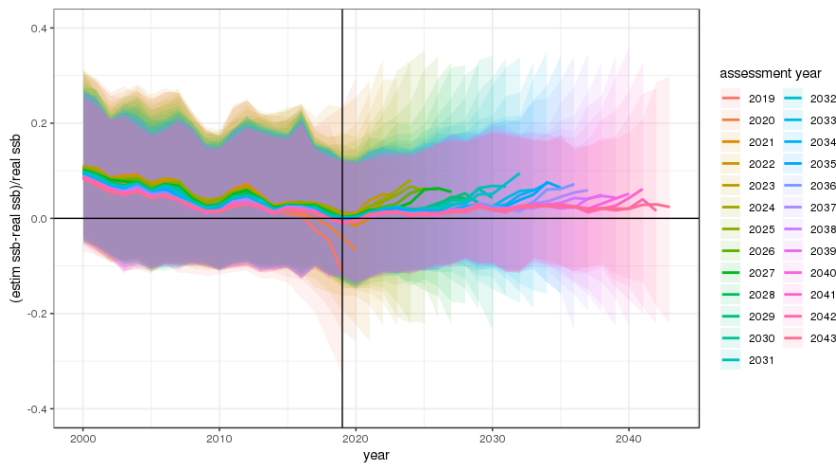


Figure 5.9: Assessment median bias (lines) along with 5th and 95th quantiles (shaded ribbon) for real SSB for each assessment year (colours) for the CbasedHCRtacmax and observation error but not initial population uncertainty (RFO scenario).

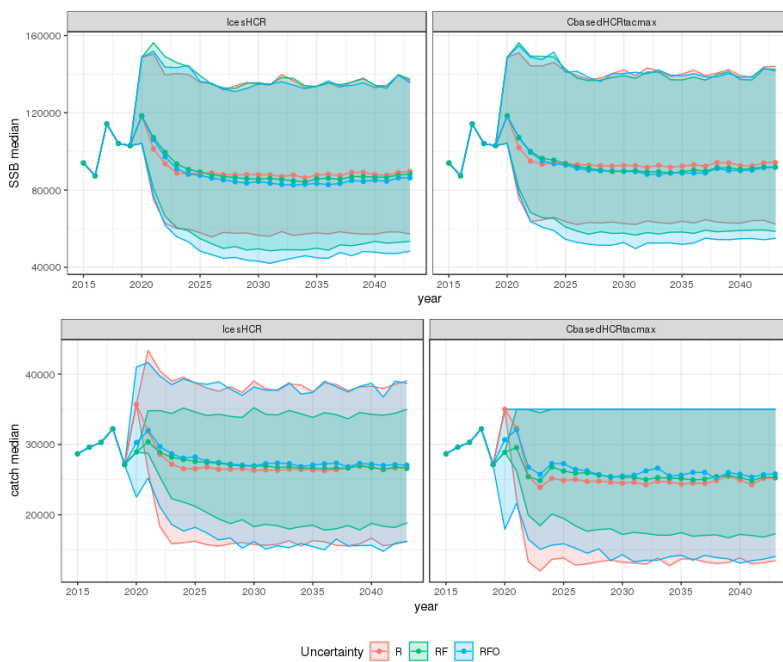


Figure 5.10: Total catch and SSB median values along with 5th and 95th quantiles (shaded ribbon) for historical and projection years for the base case scenario (R) and for RF and RFO scenarios for the two harvest control rules (IcesHCR and CbasedHCRtacmax).

5.3.4 All uncertainty sources together

As a final step, all described sources of uncertainty were included all together (scenario RIFO). As in the previous steps, median catch and SSB values for the whole projection period were similar across scenarios. Median catch tends to increase as uncertainty is included, being the largest when all sources of uncertainty are included (Figure 5.11).

In terms of uncertainty, the computed standard deviation in SSB is higher than in the other scenarios. This is also reflected in a higher risk of SSB being below B_{lim} . None of the HCRs is precautionary when all the uncertainty sources are considered. The probability of SSB being below B_{lim} raises to 0.16 for the IcesHCR and to 0.085 for the CbasedHCRtacmax. The probability of closure gets increased as well being the scenario with the highest values, with a probability of 0.02 for IcesHCR and 0.006 for the CbasedHCRtacmax. Concerning the uncertainty of catches, the standard deviation of median catches for the whole projection period is higher in the RIFO scenario than in the RFO scenario. However, in comparison to the base case, the standard deviation of catch in the RIFO scenario is not higher than the R scenario, due to the uncertainty decreasing effect of the assessment process (Figure 5.11).

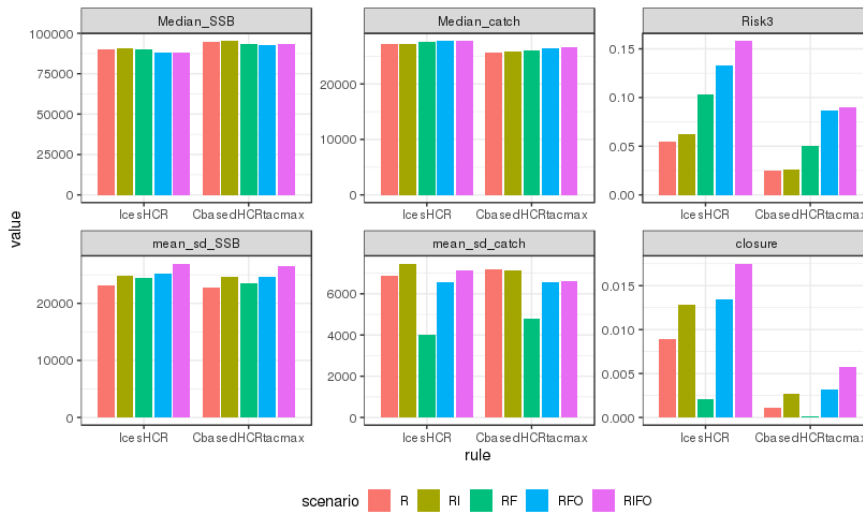


Figure 5.11: Performance indicators for the base case scenario (R) and for RI, RF, RFO and RIFO scenarios for the two harvest control rules (IcesHCR and Cbased-HCRtacmax).

Concerning the CV values across iterations for each projection year (Figure 5.12), this last scenario shows the highest values for SSB and recruitment for both rules, while total catch CVs do not show this trend due to the uncertainty decreasing effect of the assessment process described above.

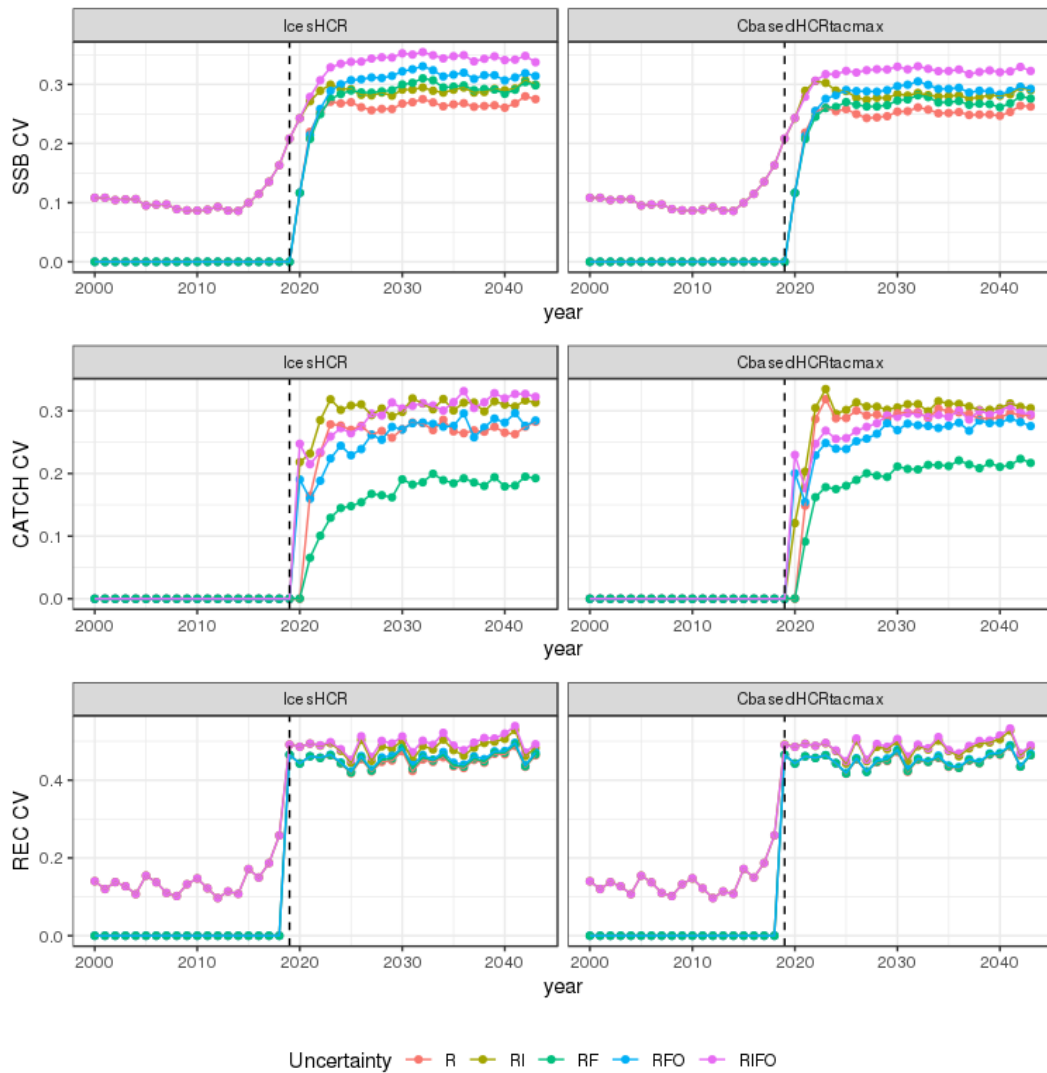


Figure 5.12: Coefficient of variation of SSB, total catch and recruitment for historical and projection years for the base case scenario (R) and for RI, RF, RFO and RIFO scenarios for the two harvest control rules (IcesHCR and CbasedHCRtacmax).

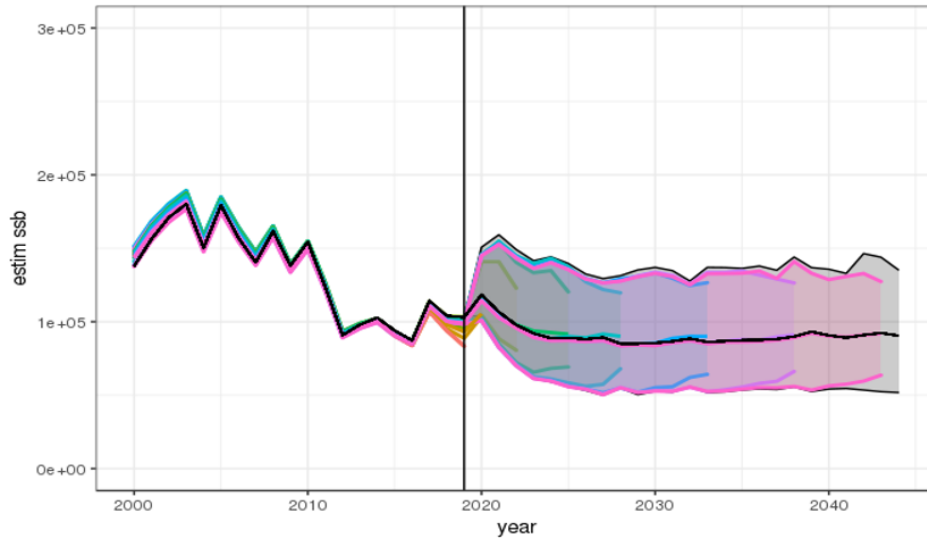


Figure 5.13: Real SSB median values (black line) along with 5th and 95th quantiles (shaded grey ribbon) for historical and projection years for the base case (R). Colored lines represent estimated SSB medians by the SS stock assessment model along with 5th and 95th quantiles (color shaded ribbons), for some assessment years along the projection years.

5.3.5 Rule selection

Although the election of the best rule was out of the scope of this work, we show how performance indicators can be used to compare the relative performance of several rules in order to achieve the management objectives established for that stock.

When incorporating all the sources of uncertainty presented above (RIFO scenario), the two rules used in this work were not precautionary according to ICES principles ($\text{Risk3} < 0.05$). Thus, for illustrative purposes regarding the rule selection, we evaluated additional alternative rules fixing a maximum TAC of 25,000 or 30,000 tonnes for the above described IcesHCR, or changing the gamma value from 0.2 to 0.6 for the CbasedHCRTacmax rule. We evaluated these rules in terms of the six performance statistics presented above. The selected indicators were scaled to the $[0,1]$ interval, using the transformation $x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)}$, in order to represent them jointly in a radar plot (Figure 5.11). This allows selecting the best rule, according to the management objectives that should be agreed by the stakeholders before the analysis. We can see in this example that the Risk3 decreases when the TACmax or

gamma values decrease, obtaining precautionary rule with the CbasedHCRtacmax with $\gamma < 0.4$. The increment of γ values results in higher median catches with higher risks and variability along with lower median SSB values. Moreover, we observe that imposing a maximum TAC to IcesHCR decreases the median catches as well as the risk and the probability of closure and increases the median SSB level respect to the IcesHCR with no TAC limit.

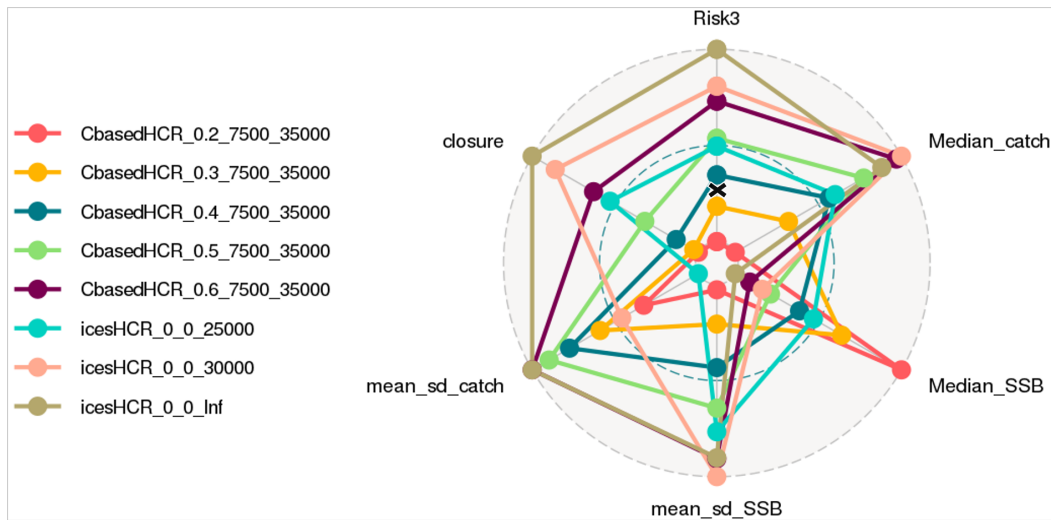


Figure 5.14: Scaled performance indicators for alternative harvest control rules for scenario RIFO. Rule names are composed by the type of rule (IcesHCR or Cbased-HCRtacmax), gamma value, minimum TAC value and maximum TAC value. The black cross in the Risk3 axis represents the scaled value of 0.05.

5.4 Discussion

In this chapter a full-feedback MSE has been developed for sardine in the Bay of Biscay. As recommended in the existing literature, different sources of uncertainty have been sequentially incorporated into the Monte Carlo simulations (Punt et al. 2016), being the assessment uncertainty the one that had a larger impact in the risk estimates.

In comparison to the most traditional management approach based on providing the ‘best assessment’ of the resource, one of the main advantages of MSE is that all the sources of uncertainty can be taken into account. However, characterizing uncer-

tainty remains a key challenge for the effective use of MSE (Butterworth and Punt 1999, Punt et al. 2016). Punt et al. (2016) distinguished five different sources of uncertainties in a MSE and considered that minimally process uncertainty, parameter uncertainty and observation error should always be included. In this work these three uncertainty sources were included. Process uncertainty was introduced through variation in annual recruitment and formed the base case for comparison with the other scenarios. This type of uncertainty is commonly included adding independent errors to expected values (Garcia et al. 2013, Jardim et al. 2010) and environmental autocorrelation effects can be also be accounted for in more complex approaches (Dichmont et al. 2006, Punt et al. 2014, Tommasi et al. 2017, Wayte 2013). Parameter uncertainty was included by incorporating the uncertainty in the parameters that define the initial population, coming from the last assessment's MCMC output as recommended by Punt et al. (2016), which was identified as the most reliable uncertainty estimation method in Chapter 4 in agreement with Magnusson et al. (2012). Furthermore, parameter uncertainty was propagated to other parameters of the MSE like the stock-recruitment model parameters and catchability-at-age parameters of the Cobb-Douglas model, conforming the initial population uncertainty. Observation error was included when generating annual data according to the assessment model assumptions. In general, as more uncertainty sources were included, the risk to be below B_{lim} and the level catches increased, while spawning stock biomass decreased. This means that the ICES precautionary criterion ($Risk3 < 0.05$) will be met for a lower level of total catches in comparison with the base case, where only process error were considered.

Ideally, the management strategy should be simulated as it would be applied in reality. Thus, MSE for model-based HCRs should explicitly incorporate the annual stock assessment model and the short-term forecast in the MSE, thus having a full-feedback MSE according to ICES terminology (ICES 2020a). In our case study, the inclusion of the annual assessment model in this MSE process had several implications. On the one hand, a bias effect was detected in the estimated final year SSB by the SS model, negative in the first three years and positive afterwards in a range of 0-5% median bias without observations error and 5-10% when introducing observation error. On the other hand, the variability of the estimated SSB by the assessment model was reduced in contrast to the true SSB. The former increased the risk of being below B_{lim} , whereas the later reduced the variability in total catches and, as a result, the probability of closures. As noted by ICES (2019d), the full

feedback approach can be useful to identify bias or lags in the assessment output. Garcia et al. (2013) found that XSA model incorporated in their MSE estimated an overexploited stock when it had healthy status and A'mar et al. (2009) found a bias on the estimated SSB by the assessment model of around 10%. Potential explanations of the assessments' output bias can be related to differences between the stock assessment model assumptions and the operating model or to noisy data that prevent the model parameters to be estimated reliably. The lower variation in catches of model-based management strategies have been already found in other case studies (Butterworth and Punt 1999), which is desirable in order to avoid marketing problems. For model-free management strategies, this effect is minimized when maximum interannual catch variations are added to the HCR to seek for catch stability along time. These latter option can be easier to test but may lack a basis to determine strategies for long-term equilibrium of the resource.

The use of the full-feedback approach is still a hot topic and divergent views about the pros and cons of full versus shortcut approaches were gathered in ICES (2020a). The full-feedback approach is very time consuming and could present additional convergence issues, while the uncertainty in the shortcut approach could be inappropriately characterized. As a compromise solution both approaches could be combined: the shortcut approach could help for initial and fast explorations of a wide set of options, but it is recommended to confirm the outcomes with a full-feedback MSE, particularly when comparing alternative operating models (ICES 2020a). Following this line, the next-generation assessment models should be designed to facilitate their use in MSE. On the one hand, they should be used easily as an OM for an MSE and, on the other hand, they should incorporate a "quick-mode" in order to act as the estimation model in the assessment process, without so much computational demand (Punt et al. 2020). Developers of the SS model used in this work, are already working on a R-based system to conduct MSE within the SS framework (Hoyle et al. 2020).

Other sources of uncertainty not considered in this work but that could be very important in MSEs are structural uncertainty (Brodziak and Piner 2010, Stewart and Martell 2015) and implementation uncertainty (Dichmont et al. 2008). Regarding structural uncertainty, different scientists tend to construct different models for the same stock (Stewart and Martell 2015), being the structure of the assessment model a potential uncertainty source to incorporate in MSE processes. Implementation un-

certainty represents situations such as decision-makers not implementing the TACs suggested, or other situations related to discard regulations (Punt et al. 2016). This can be implicitly acknowledged by setting lower catch limits than the advised TAC or introducing uncertainty on the TAC overshoot, assumed to be non-reported which can alter the HCRs performance reaching different SSB levels (Garcia et al. 2013).

The simulation framework was used to compare two alternative HCRs: the F-based ICES MSY advice rule that aims at fishing at F_{msy} and an alternative catch-based HCR. None of them were precautionary when incorporating all proposed uncertainty sources. In the case of the ICES advice rule, this might indicate differences in the uncertainty sources and the methodology used to define the reference points and the ones used in MSE and should be further explored. The major detected effect was related to assessment bias. This effect was reduced in the catch-based rule due to the maximum catch which restricts the exploitation even if the SSB is overestimated by the assessment. Among the tested options, precautionary results are only obtained with the catch-based HCR with a maximum TAC, when no assessment uncertainty is included, or when the γ values are reduced to 0.3 or below for this rule.

FLBEIA resulted to be an effective tool to carry out MSE for sardine in the Bay of Biscay. The composability of the software allowed us to use the functions already built in, but also to code new functions to satisfy specific needs, such as the stock assessment model or the multinomial observation error for the age composition. However, we did not take advantage of the potential of FLBEIA to incorporate uncertainty in the biological parameters such as weights, maturity or mortality-at-age or to analyse the economic impact of the HCRs (Garcia et al. 2013, Prellezo et al. 2016). The configuration of a MSE including biological, ecosystem, fleet and economic dynamics along with all possible sources of uncertainty can result in a very complex model. It has been shown that the relationship between model complexity and the ability to provide understandable and useful results has a bell shape (Grimm and Railsback 2005), so complexity should be chosen with regards to question of interest and objectives of the model to reach the optimum (Leutha et al. 2016). In this case, these additional options provided by FLBEIA were postponed for future work.

Chapter 6

Conclusions and further work

Monte Carlo simulation has been shown to be a powerful tool to evaluate statistical methods (Morris et al. 2019) and to test assumptions that cannot be evaluated experimentally. Simulation can also facilitate objective and transparent processes for decision-making (Goethel et al. 2019) as the performance of potential measures or strategies can be tested before their implementation. This thesis has focused on testing by means of Monte Carlo simulation statistical methods in three different fields of fishery science, namely habitat modelling, assessment and management of fish stocks. In what follows we provide the main conclusions in relation to each specific objective:

- *To propose and test via Monte Carlo simulation a habitat modelling approach that meets the ecological niche theory.*

We addressed the study of habitats of marine species through regression-based species distribution models (SDM). According to the ecological niche theory, species distribution models should provide unimodal relationships with respect to environmental gradients (Hutchinson 1957). However, most of the methods currently used for SDM do not account explicitly for this unimodal relationship and thus, can result in shapes that are not ecologically meaningful or interpretable. In order to address this issue, we proposed the use of shape-constrained Generalised additive models (SC-GAMs) that by imposing concavity shape-constraints in the response curves lead to unimodality along the chosen environmental gradients. The proposed methodology presented a good balance between the goodness of fit and agreement with ecological niche theory when applied to various simulated species distributions.

We demonstrated that SC-GAM performed better than commonly used regression-based approaches, like Generalised linear models (GLMs) and generalised additive models (GAMs) of different degrees of flexibilities, obtaining fitted curves that were closer to the real responses. When the tested datasets did not cover the whole range of the environmental gradient, similar or even better results were obtained in comparison with the most flexible GAMs, which tended to estimate multimodal or convex shapes that did not match with the underlying simulated shape. SC-GAMs were also compared to the “Plateau” environmental envelope model (Brewer et al. 2016). This method was able to provide the correct shapes in agreement with ecological niche theory, but resulted to be less robust to replicates with higher uncertainty.

The two evaluated implementations of SC-GAMs, the maximum likelihood approach in `scam` and the component-wise boosting approach in `mboost`, presented differences in terms of uncertainty estimation, computing time and ease of use. The `mboost` approach was more robust to introduced uncertainties but slower due to the high computational cost for variance estimation, while the `scam` implementation was computationally efficient and faster. In addition, `scam` is based in the popular `mgcv` package syntax, which can facilitate its use.

Based on the simulation study, we concluded that proposed SC-GAMs are appropriate and useful tools for SDM that enable fitting both unconstrained and shape-constrained responses in a unified framework and that can be readily applied for modelling a large number of species. As a first application, egg data from our case study, sardine in the Bay of Biscay, was used to model the spawning thermal habitat of this species. Application to other two pelagic species was carried out, concluding that results obtained through SC-GAMs for these species, were as good as reported results from corresponding models that need manual tuning of parameters to obtain the desired shape.

- To compare the established stock assessment model for our case study to other alternative models and to evaluate, using Monte Carlo simulation, uncertainty estimation methods when introducing flexibility through non-parametric smoothing in fishing mortality.

Having reliable estimates of the status of exploited fish stocks and the associated uncertainty is crucial for an adequate management of the resources. Among the

large number of available stock assessment models the selection of a model for a particular stock can be based on the available data or complexity level, but also on the technical skills of the analyst, computational constraints or the history of the assessment in the region. The well established Stock Synthesis (SS) model was selected by the ICES working group WGHANSA (ICES 2017c) for the first analytical stock assessment of sardine in the Bay of Biscay. In this thesis we compared it to two alternative approaches: the simpler a4a approach and an ad-hoc Bayesian approach. By making the configuration of the two alternative models as similar as possible to the official SS approach, we showed that the three approaches provided similar trends along time for the quantities of interest, such as biomass or fishing mortality, in spite of differences in the underlying assumptions or structures inherent to the type of model. The selected SS model incorporated the total catch data and used a multinomial distribution for age-structured data, however it had the limitation of relying on iterative tuning processes for some variance parameters' estimation. The newer a4a approach did not incorporate observation equations for the total catch and the age-structured data followed independent lognormal distributions, but it was able to estimate all variance parameters. In addition, it was fast and simple to implement, while very flexible in terms of parameter modelization. The proposed Bayesian approach, incorporated prior distributions and was also able to estimate variances directly. In contrast to SS and a4a that are well-known software that allow implementing a large number of configurations tested by a big community of users, the ad-hoc Bayesian model was developed and tailored for this specific case study. This provided the opportunity to compare official assessment results to a non "black-box" approach where all assumptions and equations were self-developed. We concluded that obtained estimates were all consistent and that the presented three approaches could be used to estimate Bay of Biscay sardine's stock status similarly. Furthermore, we identified a4a as a fast and simple approach that could be appropriate to replace the official assessment in simulation studies where thousands of iterations must be run.

Nowadays most stock assessment models are able to quantify estimation uncertainty, however, they use different methods. We carried out a Monte Carlo simulation study to compare different methods for uncertainty estimation in stock assessment models with non-parametric effects on fishing mortality. Using a4a, which was identified above as the most suitable for simulation, the assessment models were fitted with three different approaches for uncertainty estimation: the multivariate normal distri-

bution, the Markov chain Monte Carlo (MCMC) algorithm implemented in ADMB and the parametric bootstrap. In agreement with previous studies in the literature, we concluded that MCMC was the most reliable method. When uncertainty estimation was compared across models of different levels of complexity for fishing mortality, we found that non-parametric functions, such as splines and tensor products, were able to estimate flexible shapes that parametric models cannot capture. However, the most flexible configurations increased the number of parameters and resulted in greater uncertainty, with greater CVs and relative bias at terminal years and ages. Deterioration of coverage probabilities was also observed for the most complex scenarios, except for the MCMC method. In order to handle the degree of flexibility and the goodness of fit, model selection is usually based on AIC or BIC criteria. Both indicators were tested in terms of accuracy in selecting the correct model, obtaining better results for BIC.

The results from this study enabled having a better understanding of uncertainty estimation in stock assessments, identifying MCMC as the best option for quantifying uncertainty in real case studies that will have an impact on the management strategy evaluation process.

- *To develop a full-feedback management strategy evaluation for our case study introducing different uncertainty sources through Monte Carlo simulation.*

Managing a fish stock requires making decisions about possible actions. Management Strategy Evaluation (MSE) aims at evaluating by means of simulation the effect of different management actions before they are put in place in order to find strategies that meet the management objectives and are robust to different sources of uncertainty. This requires that the most relevant sources of uncertainty are correctly identified and the ones to be included in the MSE are properly selected. We used Monte Carlo simulation to provide a detailed analysis of the impact of the most relevant uncertainty sources for sardine in the Bay of Biscay. Starting from the base case that only incorporated uncertainty on recruitment, we added gradually additional sources of uncertainty like parameter uncertainty, observation uncertainty and estimation uncertainty, that resulted in a successive increase of the probability of SSB being below the limit reference point B_{lim} and a decrease of the average catches.

One of the major achievements in this chapter was the incorporation of the annual SS stock assessment model within the MSE loop, providing a full-feedback MSE for these stock. This allowed us to detect a bias in the estimation of SSB by the SS model, negative in the first three years and positive afterwards, that resulted in an increase of the risk of being below B_{lim} and of the variability in SSB. In addition, the SSB variability estimated by SS was lower than the real one, which led to a decrease in the probability of closure in comparison with the non full-feedback approach. The bias induced by the SS model has not been fully understood yet, and deserves further work.

For the rest of uncertainty sources, when multinomial and lognormally distributed observation errors were included, variability in SSB, risks and probability of closure increased, as well as the bias effect in the SSB estimated from the SS stock assessment. Similarly, the incorporation of initial population uncertainty, given by the MCMC output of the most recent SS assessment, resulted in an increment of the risk to be below B_{lim} , and of other indicators such as, the probability of closure and the mean standard deviation of catches and SSB along the projection year.

In general, adding uncertainty led to larger risks. So, we concluded that the ICES precautionary criterion ($Risks3 < 0.05$) would be obtained with a lower level of total catches in comparison with an MSE that does not include the studied sources of uncertainty. The two rules implemented for the Bay of Biscay sardine in this work, were considered precautionary by the base case MSE, but not when incorporating all the proposed sources of uncertainty. Although the selection of the best harvest control rule for BoB sardine was out of the scope of this work, we provided an example where several additional HCRs were compared, so that the best rule could be selected according to predefined management objectives.

The MSE was developed in FLBEIA. The modular structure allowed us to create complementary functions to incorporate the most relevant uncertainty sources for our case study. In particular, we created a function to include the SS assessment model within the MSE simulation or a function to incorporate non-multiplicative multinomial errors for age-structured data observation. The code developed for our case study is publicly available, although it is a case-specific code, it could be adapted to other stocks assessed by the SS model.

In general terms, we can say that Monte Carlo simulation has been a very useful framework for the all the studied fishery science fields. It allowed testing the modelization of ecological premises that could not have been possible using only real data. Through the generation of several types of responses along an environmental gradient based on a particular ecological theory and using Monte Carlo simulation we could test the proposed method, being able to compare the resulting responses with the true underlying theoretical shape and evaluating the performance of the our proposal. Moreover, the simulation framework allowed comparing different uncertainty estimation methods under controlled variability and population structures of different levels of complexity, which are not known for the real case study populations. Accuracy of model selection criteria could be also tested thanks to the knowledge of the real underlying "truth". Monte Carlo simulation has shown to be also a key tool for management advice, where complexity is being gradually incorporated and tested, as done here, in order to obtain reliable results, before their application to the real world. High computational power was sometimes required by the presented simulation studies, however, this was alleviated thanks to the easy parallelization of Monte Carlo simulation. This framework has provided the needed type of results for each of the objectives of this thesis and we definitely will incorporate in future studies that could be reinforced by it.

Further work

This thesis have resulted in considerable advances concerning the addressed issues, from which further research topics have emerged.

Firstly, species distribution modelling in agreement with the ecological niche theory has been addressed from a frequentist approach. Bayesian approaches are also popular for SDM (Golding and Purse 2016, Paradinas et al. 2015) and the proposal of SC-GAMs in this thesis, could be extended to a Bayesian framework. INLA (Rue et al. 2014) is increasingly used for Bayesian SDM, for which a SC-GAMs implementation could be proposed and tested. This could be done either imposing the needed constraints in the linear predictor scale as done in Chapter 1, or taking advantage of the Bayesian framework and using informative priors for certain parameters in order to obtain the desired response shapes in agreement with the niche theory. For

this purpose the interface for `mgcv` and `JAGS` (Wood et al. 2017) could be also tried and extended.

Concerning the assessment and management of fish stocks, the bias found when introducing the SS stock assessment in the MSE remained unresolved. A self-testing study for this model could help to understand if the bias is due to model misspecification or to some other aspect related to the simulation assumptions. Moreover, the factors impacting the magnitude of bias could be also studied. Some ideas have been already proposed, such as issues related to the stock-recruitment relationship, differences between the stock-recruitment used for the conditioning the MSE and the stock-recruitment in the assessment model, or changes in the production of the stock that no longer are in concordance with the fitted stock-recruitment at the beginning of the time series.

This assessment bias has an effect on management strategy evaluation that should be studied for a better decision making process, through a new simulation study where the generated bias would be known and controlled. In this line, short-cut approaches could be compared to the full-feedback MSE. This would require to develop a short-cut approach that can capture all the characteristics of the stock assessment model as realistically and thoroughly as possible, while being faster and efficient to run.

Finally, in line with fisheries sciences trend towards holistic approaches, the linkage between habitat modelling and assessment and management process could be addressed. The use of very complex ecosystem based models, which take in to account, economical, social and ecological aspects, is expanding rapidly. However, increasing model complexity does not always lead to a better understanding of the system and may increase uncertainty and model errors. Thus, caution is suggested when using them for management due to a possible lack of robustness, especially for unknown future climate scenarios (Storch et al. 2017). As an intermediate more simple approach, species interaction with the ecosystem could be studied through the proposed SDM using SC-GAMs, obtaining simple response curves for future climate change scenarios, that could be incorporated in the a4a stock assessment framework as covariates, and also in FLBEIA for MSE, taking advantage of the unified syntax of FLR and `mgcv` libraries.

6.1 Contributions

The scientific and technical material produced during the period of this doctoral thesis are detailed in this section:

First author scientific articles

- L. Citores, S. Sanchez-Maróño, L. Ibaibarriaga, A. Uriarte. (In preparation). "Development of a full-feed back MSE for the Bay of Biscay sardine".
- L. Citores, L. Ibaibarriaga, D.J. Lee, M.J. Brewer, M. Santos, G. Chust. (2020). "Modelling species presence-absence in the ecological niche theory framework using shape-constrained generalized additive models". *Ecological Modelling*, 418, 108926.
- L. Citores, L. Ibaibarriaga, E. Jardim. (2017). "Uncertainty estimation and model selection in stock assessment models with non-parametric effects on fishing mortality". *ICES Journal of Marine Science*, 75(2), 585-595.

Other scientific articles

- S. Sanchez, A. Uriarte, L. Ibaibarriaga, L. Citores. (2021). "Adapting simple index-based catch rules for data-limited stocks to short-lived fish stocks' characteristics". *Frontiers in Marine Science*, 8, 550.
- A. Silva, S. Garrido, L. Ibaibarriaga, L. Pawlowski, I. Riveiro, V. Marques, F. Ramos, E. Duhamel, M. Iglesias, P. Bryère, A. Mangin, L. Citores, P. Carrera, A. Uriarte. (2019). "Adult-mediated connectivity and spatial population structure of sardine in the Bay of Biscay and Iberian coast". *Deep Sea Research Part II: Topical Studies in Oceanography*, 159, 62-74.
- E. Villarino, J. R. Watson, B. Jönsson, J. M. Gasol, G. Salazar, S. G. Acinas, M. Estrada, R. Massana, R. Logares, C. R. Giner, M. C. Pernice, M. P. Olivar, L. Citores, J. Corell, N. Rodríguez-Ezpeleta, J. L. Acuña, A. Molina-Ramírez, J. I. González-Gordillo, A. Cózar, E. Martí, J. A. Cuesta, S. Agustí, E. Fraile-Nuez, C. M. Duarte, X. Irigoien, G. Chust. (2018). "Large-scale ocean connectivity and planktonic body size". *Nature communications*, 9(1), 1-13.
- E. Jardim, M. Eero, A. Silva, C. Ulrich, L. Pawlowski, S. J. Holmes, L. Ibaibarriaga, J. A. De Oliveira, I. Riveiro, N. Alzorriz, L. Citores, F. Scott, A. Uriarte, P. Carrera, E. Duhamel, I. Mosqueira. (2018). "Testing spatial heterogeneity with stock assessment models". *PloS one*, 13(1), e0190791.

Scientific reports

- Workshop on Data-limited Stocks of Short-Lived Species (WKDLSSLS2) (2020). ICES Scientific Reports. 2:99. 119 pp. <http://doi.org/10.17895/ices.pub.5984>
- Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA). (2020). ICES Scientific Reports. 2:41. 655 pp. <http://doi.org/10.17895/ices.pub.5977>
- Inter-benchmark process on sardine (*Sardina pilchardus*) in the Bay of Biscay (IBP-Sardine). (2019). ICES Scientific Reports. 1:80. 50 pp. <http://doi.org/10.17895/ices.pub.5552>
- Workshop on the Iberian Sardine Management and Recovery Plan (WKSARMP). (2019). Lisbon, Portugal. ICES Scientific Reports. 1:18. 168 pp. <http://doi.org/10.17895/ices.pub.5251>
- Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA). (2017). Bilbao, Spain. ICES CM 2017/ACOM:17.640 pp.
- Assessment for All initiative (a4a) – Workshop on development of MSE algorithms with R/FLR/a4a. (2017). Ispra, Italy. Publications Office of the European Union, Luxembourg, 2017, ISBN 978-92-79-71290-6, <https://doi.org/10.2760/18924>.
- Benchmark Workshop on Pelagic Stocks (WKPELA). (2017). Lisbon, Portugal. ICES CM 2017/ACOM:35. 278 pp.
- Workshop on Atlantic Sardine (WKSAR). (2016). Lisbon, Portugal. ICES CM 2016/ACOM:41. 351 pp.

Conferences

- L. Citores, S. Sánchez, A. Uriarte, L. Pawlowski, E. Duhamel, L. Ibaibarriaga. From data-limited to data-rich: the evolution of sardine in the bay of Biscay (Oral communication). ICES Annual Science Conference, Gothenburg, Sweden, 2019. (Presented by A. Uriarte).
- L. Citores, S. Sánchez, A. Uriarte, L. Pawlowski, E. Duhamel, L. Ibaibarriaga. From data-limited to data-rich: Management Strategy Evaluation for the sardine in the bay of Biscay (Oral communication). IV Jornadas de Estudiantes de la Sociedad Española de Biometría. UCM, Albacete, Spain, 2019.
- L. Citores, L. Ibaibarriaga, M. Brewer, D.J. Lee, G. Chust. Modelling probability of presence curves in the ecological niche theory framework: The case study of sardine egg data in the Bay of Biscay (Oral communication). Biostatnet 4th General Meeting. Santiago de Compostela, Spain, 2019.

- L. Citores, L. Ibaibarriaga, M. Brewer, D.J. Lee, G. Chust. Modelling presence-absence response curves in the ecological niche theory framework (Oral communication). International Statistical Ecology Conference, ISEC18. University of St Andrews, St Andrews, Scotland, 2018.
- L. Citores, L. Ibaibarriaga, M. Brewer, D.J. Lee, G. Chust. Modelling presence-absence response curves in the ecological niche theory framework (Oral communication). Young Researchers Using Statistics. Royal Statistical Society, Aberdeen, Scotland, 2018.
- L. Citores, L. Ibaibarriaga, D.J. Lee, G. Chust. Preliminary results on modelling presence-absence response curves in the ecological niche theory framework (Oral communication). III Jornadas de Estudiantes de la Sociedad Española de Biometría. BCAM, Bilbao, Spain, 2018.
- L. Citores, L. Ibaibarriaga, L. Pawlowski, A. Uriarte, D.J. Lee. A Bayesian stock assessment model for the Sardine in the Bay of Biscay (Oral communication). Second Bilbao Data Science Workshop. BCAM, Bilbao, Spain, 2017.
- L. Citores, L. Ibaibarriaga, L. Pawlowski, A. Uriarte, D.J. Lee. A statistical catch at age stock assessment model for Sardine in the Bay of Biscay (Oral communication). XVI Spanish Biometric Conference. IMUS, Sevilla, Spain, 2017.
- L. Citores, L. Ibaibarriaga, L. Pawlowski, A. Uriarte, D.J. Lee. A Bayesian stock assessment model for the Sardine in the Bay of Biscay (Poster). Valencia International Bayesian Analysis Workshop. UV, Valencia, Spain, 2017.
- L. Citores, L. Ibaibarriaga, D.J. Lee. Statistical modelling for fish meta-population management (Oral communication). First Bilbao Data Science Workshop. BCAM, Bilbao, Spain, 2016.
- L. Citores, L. Ibaibarriaga. Comparison of different methods to estimate stock assessment models uncertainty (Oral communication). II Jornadas de Estudiantes de la Sociedad Española de Biometría. SEB, Barcelona, Spain, 2016.

References

- Aarts, G. and Poos, J. J. (2009). Comprehensive discard reconstruction and abundance estimation using flexible selectivity functions. *ICES Journal of Marine Science*, 66:763–771.
- Aeberhard, W. H., Mills Flemming, J., and Nielsen, A. (2018). Review of state-space models for fisheries science. *Annual Review of Statistics and Its Application*, 5:215–235.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19:716–723.
- A’mar, Z. T., Punt, A. E., and Dorn, M. W. (2009). The impact of regime shifts on the performance of management strategies for the gulf of alaska walleye pollock (*Theragra chalcogramma*) fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, 66(12):2222–2242.
- Anderson, S. C., Cooper, A. B., Jensen, O. P., Minto, C., Thorson, J. T., Walsh, J. C., Afflerbach, J., Dickey-Collas, M., Kleisner, K. M., Longo, C., Osio, G. C., Ovanado, D., Mosquera, I., Rosenberg, A. A., and Selig, E. R. (2017). Improving estimates of population status and trend with superensemble models. *Fish and Fisheries*, 00:1–10.
- Atanassov, E. and Dimov, I. T. (2008). What Monte Carlo models can do and cannot do efficiently? *Applied Mathematical Modelling*, 32(8):1477–1500.
- Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Flemming, J. M., Nielsen, A., Petris, G., et al. (2020). An introduction to state-space modeling of ecological time series. *arXiv preprint arXiv:2002.02001*.

- Austin, M. (1980). Searching for a model for use in vegetation analysis. *Vegetatio*, 42(1-3):11–21.
- Austin, M. (1987). Models for the analysis of species' response to environmental gradients. *Vegetatio*, 69:35–45.
- Austin, M. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, 157:101–118.
- Austin, M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, 200:1–19.
- Austin, M. and Smith, T. (1990). A new model for the continuum concept. In *Progress in theoretical vegetation science*, volume 83(1-2), pages 35–47. Springer.
- Austin, M. P. and Cunningham, R. (1981). Observational analysis of environmental gradients. *Proceedings of the Ecological Society of Australia*. 11: 109-19 24.
- Barange, M., Bernal, M., Cercole, M. C., Cubillos, L. A., Daskalov, G. M., Cunningham, C. L., de Oliveira, J. A., Dickey-Collas, M., Gaughan, D. J., Hill, K., et al. (2009). Current trends in the assessment and management of stocks. In *Climate change and small pelagic fish*, pages 191–255. Cambridge University Press.
- Baranov, F. (1918). On the question of the biological basis of fisheries. *Nauchnyi Issledovatel'skii Ikhtiologicheskii Institut Isvestia*, 1:81–128.
- Barbosa, F. G. and Schneck, F. (2015). Characteristics of the top-cited papers in species distribution predictive models. *Ecological Modelling*, 313:77–83.
- Barrowman, N. J. and Myers, R. A. (2000). Still more spawner-recruitment curves: the hockey stick and its generalizations. *Canadian Journal of Fisheries and Aquatic Sciences*, 57(4):665–676.
- Barton, K. (2009). Mumin: multi-model inference. R package. <http://r-forge.r-project.org/projects/mumin/>.
- Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M. (2018). Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18(153):1–43.

- Beaugrand, G., Edwards, M., Brander, K., Luczak, C., and Ibanez, F. (2008). Causes and projections of abrupt climate-driven ecosystem shifts in the north atlantic. *Ecology letters*, 11(11):1157–1168.
- Berkson, J., Hunt, K., Whitehead, J., Murie, D., Kwak, T., and Boreman, J. (2009). Is there a shortage of fisheries stock assessment scientists? *Fisheries*, 34(5):217–219.
- Bernal, M., Stratoudakis, Y., Coombs, S., Angelico, M., De Lanzós, A. L., Porteiro, C., Sagarminaga, Y., Santos, M., Uriarte, A., Cunha, E., Valdés, L., and Borchers, D. (2007). Sardine spawning off the european atlantic coast: characterization of and spatio-temporal variability in spawning habitat. *Progress in Oceanography*, 74(2-3):210–227.
- Beverton, R. J. and Holt, S. J. (1957). On the dynamics of exploited fish populations. *Fisheries Investigation, Ministry of Agriculture, London, UK*, 19(2).
- Bollaerts, K., Eilers, P., and van Mechelen, I. (2006). Simple and multiple p-splines regression with shape constraints. *British Journal of Mathematical and Statistical Psychology*, 59:451–469.
- Borchers, D., Buckland, S., Priede, I., and Ahmadi, S. (1997). Improving the precision of the daily egg production method using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(12):2727–2742.
- Brandimarte, P. (2014). *Handbook in Monte Carlo simulation: applications in financial engineering, risk management, and economics*. John Wiley & Sons.
- Brewer, M. J., O’Hara, R. B., Anderson, B. J., and Ohlemüller, R. (2016). Plateau: a new method for ecologically plausible climate envelopes for species distribution modelling. *Methods in Ecology and Evolution*, 7(12):1489–1502.
- Brodziak, J. and Piner, K. (2010). Model averaging and probable status of north pacific striped marlin, *tetrapturus audax*. *Canadian Journal of Fisheries and Aquatic Sciences*, 67(5):793–805.
- Brooks, E. N. and Deroba, J. J. (2015). When data are not data: the pitfalls of post hoc analyses that use stock assessment model output. *Canadian Journal of Fisheries and Aquatic Sciences*, 72:634–641.
- Brooks, S., Gelman, A., Jones, G., and Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*. CRC press.

- Bruge, A., Alvarez, P., Fontán, A., Cotano, U., and Chust, G. (2016). Thermal niche tracking and future distribution of atlantic mackerel spawning in response to ocean warming. *Frontiers in Marine Science*, 3:86.
- Brunel, T., Van Damme, C. J., Samson, M., and Dickey-Collas, M. (2018). Quantifying the influence of geography and environment on the northeast atlantic mackerel spawning distribution. *Fisheries oceanography*, 27(2):159–173.
- Buckland, S. T., Newman, K. B., Fernández, C., Thomas, L., and Harwood, J. (2007). Embedding population dynamics models in inference. *Statistical Science*, 22(1):44–58.
- Bühlmann, P. and Hothorn, T. (2007). Boosting algorithms: Regularization, prediction and model fitting. *Statistical Science*, 22(4):477–505.
- Bunnefeld, N., Hoshino, E., and Milner-Gulland, E. J. (2011). Management strategy evaluation: a powerful tool for conservation? *Trends in ecology & evolution*, 26(9):441–447.
- Burnham, K. P. and Anderson, D. R. (2003). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media.
- Busby, J. (1991). BIOCLIM—a bioclimate analysis and prediction system. *Nature Conservation: Cost Effective Biological Surveys and Data Analysis (eds Margules, C.R. and Austin, M.P.)*, CSIRO:64–68.
- Butterworth, D. and Punt, A. (1999). Experiences in the evaluation and implementation of management procedures. *ICES Journal of Marine Science*, 56(6):985–998.
- Butterworth, D. S. (2007). Why a management procedure approach? some positives and negatives. *ICES Journal of Marine Science*, 64(4):613–617.
- Butterworth, D. S. and Rademeyer, R. A. (2008). Statistical catch-at-age analysis vs adapt-vpa: the case of gulf of maine cod. *ICES Journal of Marine Science*, 65:1717–1732.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1):1–32.
- Carsey, T. M. and Harden, J. J. (2013). *Monte Carlo simulation and resampling methods for social science*. Sage Publications.

- Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, 46(3):167–174.
- Cerdeira, J. O., Monteiro-Henriques, T., Martins, M. J., Silva, P. C., Alagador, D., Franco, A. M., Campagnolo, M. L., Arsénio, P., Aguiar, F. C., and Cabeza, M. (2018). Revisiting niche fundamentals with Tukey depth. *Methods in Ecology and Evolution*, 9(12):2349–2361.
- Chambers, J. and Hastie, T. (1992). *Statistical models, Chapter 2, Statistical models in S*. Wadsworth and Brooks.
- Chase, J. M. and Leibold, M. A. (2003). *Ecological niches: linking classical and contemporary approaches*. University of Chicago Press.
- Chen, Y., Jiao, Y., and Chen, L. (2003). Developing robust frequentist and Bayesian fish stock assessment methods. *Fish and Fisheries*, 4(2):105–120.
- Cheung, W. W., Lam, V. W., Sarmiento, J. L., Kearney, K., Watson, R., and Pauly, D. (2009). Projecting global marine biodiversity impacts under climate change scenarios. *Fish and fisheries*, 10(3):235–251.
- Cheung, W. W., Watson, R., and Pauly, D. (2013). Signature of ocean warming in global fisheries catch. *Nature*, 497(7449):365–368.
- Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American statistician*, 49(4):327–335.
- Chust, G., Castellani, C., Licandro, P., Ibaibarriaga, L., Sagarminaga, Y., and Irigoien, X. (2014). Are Calanus spp. shifting poleward in the North Atlantic? A habitat modelling approach. *ICES Journal of Marine Science*, 71(2):241–253.
- Concato, J. and Feinstein, A. R. (1997). Monte Carlo methods in clinical research: applications in multivariable analysis. *Journal of investigative medicine: the official publication of the American Federation for Clinical Research*, 45(6):394–400.
- Cotter, A., Burt, L., Paxton, C., Fernandez, C., Buckland, S., and Pan, J.-X. (2004). Are stock assessment methods too complicated? *Fish and Fisheries*, 5(3):235–254.
- Coudun, C. and Gegout, J.-C. (2006). The derivation of species response curves with gaussian logistic regression is sensitive to sampling intensity and curve characteristics. *Ecological Modelling*, 199(2):164–175.

- Crone, P., Maunder, M., Valero, J., McDaniel, J., and Semmens, B. (2013). *Selectivity: theory, estimation and application in fishery stock assessment models*.
- de Boor, C. (1972). *A Practical Guide to Splines*. Springer.
- de Moor, C. and Butterworth, D. (2007). Base case assessment of the south african sardine resource. *Marine and Coastal Management Document MCM/2007/SWP/SWG-PEL/06*. 30pp.
- De Oliveira, J. and Butterworth, D. S. (2004). Developing and refining a joint management procedure for the multispecies south african pelagic fishery. *ICES Journal of Marine Science*, 61(8):1432–1442.
- Derber, J. and Rosati, A. (1989). A global oceanic data assimilation system. *Journal of Physical Oceanography*, 19(9):1333–1347.
- Deriso, R., Quinn Ii, T., and Neal, P. (1985). Catch-age analysis with auxiliary information. *Canadian Journal of Fisheries and Aquatic Sciences*, 42(4):815–824.
- Deriso, R. B. (1980). Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(2):268–282.
- Deroba, J. J., Butterworth, D. S., Methot, R. D., De Oliveira, J. A. A., Fernandez, C., Nielsen, A., Cadrin, S. X., et al. (2014). Simulation testing the robustness of stock assessment models to error: some results from the ices strategic initiative on stock assessment methods. *ICES Journal of ine Science*, 72:19–30.
- Dichmont, C., Deng, A., Punt, A., Ellis, N., Venables, W., Kompas, T., Ye, Y., Zhou, S., and Bishop, J. (2008). Beyond biological performance measures in management strategy evaluation: bringing in economics and the effects of trawling on the benthos. *Fisheries Research*, 94(3):238–250.
- Dichmont, C. M., Deng, A. R., Punt, A. E., Venables, W., and Haddon, M. (2006). Management strategies for short-lived species: The case of australia’s northern prawn fishery: 1. accounting for multiple species, spatial structure and implementation uncertainty when evaluating risk. *Fisheries research*, 82(1-3):204–220.
- Dichmont, C. M., Deng, R. A., Punt, A. E., Brodziak, J., Chang, Y.-J., Cope, J. M., Ianelli, J. N., Legault, C. M., Methot Jr, R. D., Porch, C. E., et al. (2016). A review of stock assessment packages in the United States. *Fisheries Research*, 183:447–460.

- Doney, S. C., Ruckelshaus, M., Emmett Duffy, J., Barry, J. P., Chan, F., English, C. A., Galindo, H. M., Grebmeier, J. M., Hollowed, A. B., Knowlton, N., et al. (2012). Climate change impacts on marine ecosystems. *Annual review of marine science*, 4:11–37.
- Doray, M., Petitgas, P., Romagnan, J. B., Huret, M., Duhamel, E., Dupuy, C., Spitz, J., Authier, M., Sanchez, F., Berger, L., et al. (2018). The PELGAS survey: ship-based integrated monitoring of the bay of biscay pelagic ecosystem. *Progress in Oceanography*, 166:15–29.
- Doucet, A., De Freitas, N., and Gordon, N. (2001). An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo methods in practice*, pages 3–14. Springer.
- Dowling, N. A., Smith, A. D., Smith, D. C., Parma, A. M., Dichmont, C. M., Sainsbury, K., Wilson, J. R., Dougherty, D. T., and Cope, J. M. (2019). Generic solutions for data-limited fishery assessments are not so simple. *Fish and Fisheries*, 20(1):174–188.
- Dziak, J. J., Coffman, D. L., Lanza, S. T., and Li, R. (2012). Sensitivity and specificity of information criteria: Technical Report Series 12–119. *The Methodology Center and Department of Statistics, Penn State, The Pennsylvania State University*, pages 1–30.
- Edwards, C., Hillary, R., Levontin, P., Blanchard, J., and Lorenzen, K. (2012a). Fisheries assessment and management: a synthesis of common approaches with special reference to deepwater and data-poor stocks. *Reviews in Fisheries Science*, 20(3):136–153.
- Edwards, K., Barciela, R., and Butenschön, M. (2012b). Validation of the nemo-ersem operational ecosystem model for the north west european continental shelf. *Ocean Sciences Discussions*, 8(6):983–1000.
- Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- Eilers, P. and Marx, B. (1996). Flexible smoothing with B -splines and penalties. *Statistical Science*, 11:89–121.
- Eilers, P., Marx, B., and Durbán, M. (2015). Twenty years of P-splines. *SORT-Statistics and Operations Research Transactions*, 39(2):149–186.

- Eilers, P. H. (2017). Uncommon penalties for common problems. *Journal of Chemometrics*, 31(4):e2878.
- Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., et al. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2):129–151.
- Elith, J. and Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics*, 40.
- Elvarsson, B., Taylor, L., Trenkel, V. M., Kupca, V., and Stefansson, G. (2014). A bootstrap method for estimating bias and variance in statistical fisheries modelling frameworks using highly disparate datasets. *African Journal of Marine Science*, 36:99–110.
- Erauskin-Extramiana, M., Alvarez, P., Arrizabalaga, H., Ibaibarriaga, L., Uriarte, A., Cotano, U., Santos, M., Ferrer, L., Cabré, A., Irigoien, X., and Chust, G. (2019). Historical trends and future distribution of anchovy spawning in the bay of biscay. *Deep Sea Research Part II: Topical Studies in Oceanography*, 159:169–182.
- Fernandez, C., Cervino, S., Perez, N., and Jardim, E. (2010). Stock assessment and projections incorporating discard estimates in some years: an application to the hake stock in ices divisions viiic and ixa. *ICES Journal of Marine Science*, 67:1185–1197.
- Fielding, A. H. and Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, 24(1):38–49.
- Fournier, D. and Archibald, C. P. (1982). A general theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(8):1195–1207.
- Fournier, D. A., Hampton, J., and Sibert, J. R. (1998). Multifan-cl: a length-based, age-structured model for fisheries stock assessment, with application to south pacific albacore, thunnus alalunga. *Canadian Journal of Fisheries and Aquatic Sciences*, 55:2105–2116.
- Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., et al. (2012). Ad model builder: using automatic differenti-

- ation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 27:233–249.
- Fraaije, R. G., ter Braak, C. J., Verduyn, B., Breeman, L. B., Verhoeven, J. T., and Soons, M. B. (2015). Early plant recruitment stages set the template for the development of vegetation patterns along a hydrological gradient. *Functional Ecology*, 29(7):971–980.
- Francis, R. and Shotton, R. (1997). Risk in fisheries management: a review. *Can J. Fish. Aquat. Sci.*, 54:1699–1715.
- Friedman, J. H. et al. (1991). Multivariate adaptive regression splines. *The annals of statistics*, 19(1):1–67.
- Fronczyk, K., Kottas, A., and Munch, S. (2011). Flexible modeling for stock-recruitment relationships using Bayesian nonparametric mixtures. *Environmental and Ecological Statistics*.
- Gamerman, D. and Lopes, H. F. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. CRC Press.
- Garcia, D., Sánchez, S., Prellezo, R., Urtizberea, A., and Andrés, M. (2017). Flbeia: A simulation model to conduct bio-economic evaluation of fisheries management strategies. *SoftwareX*, 6:141–147.
- Garcia, D., Urtizberea, A., Diez, G., Gil, J., and Marchal, P. (2013). Bio-economic management strategy evaluation of deepwater stocks using the flbeia model. *Aquatic Living Resources*, 26(4):365–379.
- Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.
- Gelman, A., Rubin, D. B., et al. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.
- Geweke, J. et al. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, volume 196. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN.

- Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). Introducing Markov Chain Monte Carlo. *Markov chain Monte Carlo in practice*, 1:19.
- Gislason, H., Daan, N., Rice, J. C., and Pope, J. G. (2010). Size, growth, temperature and the natural mortality of marine fish. *Fish and Fisheries*, 11(2):149–158.
- Goethel, D. R., Lucey, S. M., Berger, A. M., Gaichas, S. K., Karp, M. A., Lynch, P. D., Walter III, J. F., Deroba, J. J., Miller, S., and Wilberg, M. J. (2019). Closing the feedback loop: on stakeholder participation in management strategy evaluation. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(10):1895–1913.
- Goethel, D. R., Quinn, T. J., and Cadrin, S. X. (2011). Incorporating spatial structure in stock assessment: movement modeling in marine fish population dynamics. *Reviews in Fisheries Science*, 19(2):119–136.
- Golding, N. and Purse, B. V. (2016). Fast and flexible Bayesian species distribution modelling using gaussian processes. *Methods in Ecology and Evolution*, 7(5):598–608.
- Green, P. J. and Silverman, B. W. (1993). *Nonparametric regression and generalized linear models: a roughness penalty approach*. Chapman and Hall/CRC.
- Griewank, A. et al. (1989). On automatic differentiation. *Mathematical Programming: recent developments and applications*, 6(6):83–107.
- Grimm, V. and Railsback, S. F. (2005). *Individual-based modeling and ecology*, volume 8. Princeton university press.
- Guevara, J., Zadrozny, B., Buoro, A., Lu, L., Tolle, J., Limbeck, J., Wu, M., and Hohl, D. (2018). A hybrid data-driven and knowledge-driven methodology for estimating the effect of completion parameters on the cumulative production of horizontal wells. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Guisan, A., Edwards Jr, T. C., and Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling*, 157(2-3):89–100.
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., Ferrier, S.,

- Kearney, M. R., Possingham, H. P., and Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology letters*, 16(12):1424–1435.
- Guisan, A. and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, 135(2-3):147–186.
- Gunderson, D. R. (1993). *Surveys of fisheries resources*. John Wiley & Sons.
- Halvorsen, R. (2012). A gradient analytic perspective on distribution modelling. *Sommerfeltia*, 35(1):1–165.
- Hardle, W. (1994). Applied nonparametric regression. *Cambridge university press*.
- Harrison, R. L. (2010). Introduction to Monte Carlo simulation. In *AIP conference proceedings*, volume 1204(1), pages 17–21. American Institute of Physics.
- Hart, P. J., Reynolds, J. D., Hart, P. J., and Reynolds, J. D. (2002). *Handbook of fish biology and fisheries*, volume 1. Wiley Online Library.
- Hastie, T. J. and Tibshirani, R. J. (1990). Generalized additive models. *Monographs on Statistics and Applied Probability*, 43.
- Heikkinen, J. and Makipaa, R. (2010). Testing hypotheses on shape and distribution of ecological response curves. *Ecological Modelling*, 221(3):388–399.
- Helaouet, P. and Beaugrand, G. (2009). Physiology, ecological niches and species distribution. *Ecosystems*, 12(8):1235–1245.
- Hijmans, R. J. (2012). Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, 93(3):679–688.
- Hilborn, R. (1992). Current and future trends in fisheries stock assessment and management. *South African Journal of Marine Science*, 12(1):975–988.
- Hilborn, R. (2003). The state of the art in stock assessment: where we are and where we are going. *Scientia Marina*, 67(S1):15–20.
- Hilborn, R., Maguire, J.-J., Parma, A. M., and Rosenberg, A. A. (2001). The precautionary approach and risk management: can they increase the probability of successes in fishery management? *Canadian Journal of Fisheries and Aquatic Sciences*, 58:99–107.

- Hilborn, R. and Peterman, R. M. (1996). The development of scientific advice with incomplete information in the context of the precautionary approach. *FAO Fisheries Technical Paper*, 63(606):77–102.
- Hilborn, R. and Walters, C. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media.
- Hillary, R. (2012). Practical uses of non-parametric methods in fisheries assessment modelling. *Marine and freshwater research*, 63(7):606–615.
- Hoffman, M. D. and Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623.
- Hofner, B., Kneib, T., and Hothorn, T. (2016). A unified framework of constrained regression. *Statistics and Computing*, 26(1-2):1–14.
- Hofner, B., Mayr, A., Robinzonov, N., and Schmid, M. (2014). Model-based boosting in R: A hands-on tutorial using the R package mboost. *Computational Statistics*, 29(3–35).
- Hothorn, T., Bühlmann, P., Kneib, T., Schmid, M., and Hofner, B. (2010). Model-based boosting 2.0. *Journal of Machine Learning Research*, 11:2109–2113.
- Hothorn, T., Bühlmann, P., Kneib, T., Schmid, M., and Hofner, B. (2018). *mboost: Model-Based Boosting*. <https://CRAN.R-project.org/package=mboost>.
- Hoyle, S., Maunder, M., and A’mar, Z. (2020). Frameworks for the next generation of general stock assessment models: 2019 capam workshop report. *New Zealand Fisheries Assessment Report*, page 39.
- Huisman, J., Olf, H., and Fresco, L. (1993). A hierarchical set of models for species response analysis. *Journal of Vegetation Science*, 4(1):37–46.
- Huston, M. A. (1994). *Biological diversity: the coexistence of species*. Cambridge University Press.
- Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22:415–427.
- Hutchinson, G. E. (1978). *An introduction to population ecology*. Hew Haven Connecticut Yale University Press.

- Ianelli, J. N. (2002). Simulation analyses testing the robustness of productivity determinations from west coast pacific ocean perch stock assessment data. *North American Journal of Fisheries Management*, 22(1):301–310.
- Ibaibarriaga, L., Fernández, C., Uriarte, A., and Roel, B. A. (2008). A two-stage biomass dynamic model for bay of biscay anchovy: a Bayesian approach. *ICES Journal of Marine Science*, 65(2):191–205.
- Ibaibarriaga, L., Irigoien, X., Santos, M., Motos, L., Fives, J., Franco, C., Lago de Lanzós, A., Acevedo, S., Bernal, M., Bez, N., Eltink, G., Faniha, A., Hammer, C., Iversen, S., Milligan, S., and Reid, G. (2007). Egg and larval distributions of seven fish species in north-east atlantic waters. *Fisheries Oceanography*, 16(3):284–293.
- ICES (2006). Report of the workshop on guidelines for management strategy evaluations (WKG MSE). 23 - 27 January 2006, ICES Headquarters. ICES CM 2006/ACFM:15. 157 pp.
- ICES (2013). Report of the benchmark workshop on pelagic stocks (wkpela). 4-8 February 2013, Copenhagen, Denmark. ICES CM 2013/ACOM:46. 483 pp.
- ICES (2015). Report of the working group on southern horse mackerel, anchovy and sardine (WGHANSA). 24-29 June 2015, Lisbon, Portugal. ICES CM 2015/ACOM:16.612 pp.
- ICES (2017a). Report of the benchmark workshop on pelagic stocks (WKPELA). 6-10 February 2017, Lisbon, Portugal. ICES CM 2017/ACOM:35.
- ICES (2017b). Report of the working group on acoustic and egg surveys for sardine and anchovy in ices areas 7, 8, and 9. WGACEGG Report 2016 Capo, Granitola, Sicily, Italy. 14-18 November 2016. ICES CM 2016/SSGIEOM:31. 326 pp.
- ICES (2017c). Report of the working group on southern horse mackerel, anchovy and sardine (WGHANSA). 24-29 June 2017, Lisbon, Portugal. ICES CM 2017/ACOM:17.
- ICES (2018a). Report of the working group on mackerel and horse mackerel egg surveys. WGMEGS Report 2018 Dublin, Ireland. 9-13 April 2018. ICES CM 2018/EOSG: 17.

- ICES (2018b). Report of the working group on southern horse mackerel, anchovy and sardine (WGHANSA). 26–30 June 2018, Lisbon, Portugal. ICES CM 2018/A-COM:17.659 pp.
- ICES (2019a). Advice basis. In Report of the ICES Advisory Committee. ICES Advice 2019 section 1.2.
- ICES (2019b). Inter-benchmark process on the sardine (*sardina pilchardus*) in the bay of biscay (IBPSARDINE). ICES Scientific Reports. 1:80. 50 pp.
- ICES (2019c). Report of the working group on southern horse mackerel, anchovy and sardine (WGHANSA). ICES Scientific Reports. 1:34. 653 pp. <http://doi.org/10.17895/ices.pub.4983>.
- ICES (2019d). Report of the workshop on guidelines for management strategy evaluations (WKG MSE2). ICES Scientific Reports. 1:33. 162 pp. <http://doi.org/10.17895/ices.pub.5331>.
- ICES (2019e). Stock annex: Sardine in divisions 8abd.
- ICES (2020a). Report of the workshop on guidelines for management strategy evaluations (WKG MSE3). ICES Scientific Reports. 2:116. 112 pp. <http://doi.org/10.17895/ices.pub.7627>.
- ICES (2020b). Working group on acoustic and egg surveys for sardine and anchovy in ices areas 7, 8 and 9 (WGACEGG; outputs from 2019 meeting). ICES Scientific Reports. 2:44. 490 pp.
- IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Irwin, B. J., Wilberg, M. J., Bence, J. R., and Jones, M. L. (2008). Evaluating alternative harvest policies for yellow perch in southern lake michigan. *Fisheries Research*, 94(3):267–281.
- Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., and Gutiérrez, J. M. (2015). A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*, 312:166–174.
- Jamil, T. and Ter Braak, C. J. (2013). Generalized linear mixed models can detect unimodal species-environment relationships. *PeerJ*, 1:e95.

- Jardim, E. (2017). *a4a: Assessment for all*. http://www.flr-project.org/doc/Statistical_catch_at_age_models_in_FLa4a.html.
- Jardim, E., Cervino, S., and Azevedo, M. (2010). Evaluating management strategies to implement the recovery plan for iberian hake (*merluccius merluccius*); the impact of censored catch information. *ICES Journal of Marine Science*, 67(2):258–269.
- Jardim, E., Eero, M., Silva, A., Ulrich, C., Pawlowski, L., Holmes, S. J., Ibaibarriaga, L., De Oliveira, J. A., Riveiro, I., Alzorritz, N., et al. (2018). Testing spatial heterogeneity with stock assessment models. *PloS one*, 13(1):e0190791.
- Jardim, E., Millar, C. P., Mosqueira, I., Scott, F., Osio, G. C., Ferretti, M., Alzorritz, N., and Orío, A. (2014). What if stock assessment is as simple as a linear model? the a4a initiative. *ICES Journal of Marine Science*, 72(1):232–236.
- Jiménez-Valverde, A., Lobo, J. M., and Hortal, J. (2008). Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, 14(6):885–890.
- Johansen, A. (2010). Monte Carlo Methods. In Peterson, P., Baker, E., and McGaw, B., editors, *International Encyclopedia of Education (Third Edition)*, pages 296–303. Elsevier, third edition edition.
- Jorgensen, W. L. and Duffy, E. M. (2000). Prediction of drug solubility from Monte Carlo simulations. *Bioorganic & Medicinal Chemistry Letters*, 10(11):1155–1158.
- Kearney, M. (2006). Habitat, environment and niche: what are we modelling? *OIKOS*, 115(1):186–191.
- Kell, L. T., De Oliveira, J. A., Punt, A. E., McAllister, M. K., and Kuikka, S. (2006). Operational management procedures: an introduction to the use of evaluation frameworks. In *Developments in Aquaculture and Fisheries Science*, volume 36, pages 379–407. Elsevier.
- Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R. M., Jardim, E., et al. (2007). Flr: an open-source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science*, 64:640–646.
- Kenchington, T. J. (2014). Natural mortality estimators for information-limited fisheries. *Fish and Fisheries*, 15(4):533–562.

- Koenigstein, S., Mark, F. C., Gößling-Reisemann, S., Reuter, H., and Poertner, H.-O. (2016). Modelling climate change impacts on marine fish populations: process-based integration of ocean warming, acidification and other environmental drivers. *Fish and Fisheries*, 17(4):972–1004.
- Kolody, D., Polacheck, T., Basson, M., and Davies, C. (2008). Salvaged pearls: lessons learned from a floundering attempt to develop a management procedure for southern bluefin tuna. *Fisheries Research*, 94(3):339–350.
- Kraak, S. B., Kelly, C. J., Codling, E. A., and Rogan, E. (2010). On scientists' discomfort in fisheries advisory science: the example of simulation-based fisheries management-strategy evaluations. *Fish and Fisheries*, 11(2):119–132.
- Kristensen, K., Nielsen, A., Berg, C., Skaug, H., and Bell, B. (2016). Template model builder tmb. *Journal of Statistical Software*, 70(5):1–21.
- Kriticos, D. J., Maywald, G. F., Yonow, T., Zurcher, E. J., Herrmann, N. I., and Sutherst, R. (2015). Exploring the effects of climate on plants, animals and diseases. *CLIMEX Version*, 4:184.
- Lasker, R. (1985). An egg production method for estimating spawning biomass of pelagic fish: application to the northern anchovy, *engraulis mordax*. *NOAA Technical Report*, page NMFS 36pg. 99 pp.
- Leathwick, J., Elith, J., and Hastie, T. (2006). Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological modelling*, 199(2):188–196.
- Lehmann, A., Overton, J. M., and Austin, M. (2002). Regression models for spatial prediction: their role for biodiversity and conservation. *Biodiversity & Conservation*, 11(12):2085–2092.
- Lockwood, S., Nichols, J., and Dawson, W. A. (1981). The estimation of a mackerel (*scomber scombrus* l.) spawning stock size by plankton survey. *Journal of Plankton Research*, 3(2):217–233.
- Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of fish biology*, 49(4):627–642.

- Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. CRC press.
- MacCall, A. D. (2013). Use of the delta method to evaluate the precision of assessments that fix parameter values. *Fisheries Research*, 142:56–60.
- Magnusson, A. and Hilborn, R. (2007). What makes fisheries data informative? *Fish and Fisheries*, 8:337–358.
- Magnusson, A., Punt, A. E., and Hilborn, R. (2012). Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and mcmc. *Fish and Fisheries*, 14:325–342.
- Mäntyniemi, S., Uusitalo, L., Peltonen, H., Haapasaari, P., and Kuikka, S. (2013). Integrated, age-structured, length-based stock assessment model with uncertain process variances, structural uncertainty, and environmental covariates: case of central baltic herring. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(9):1317–1326.
- Mäntyniemi, S. H., Whitlock, R. E., Perälä, T. A., Blomstedt, P. A., Vanhatalo, J. P., Rincón, M. M., Kuparinen, A. K., Pulkkinen, H. P., and Kuikka, O. S. (2015). General state-space population dynamics model for Bayesian stock assessment. *ICES Journal of Marine Science*, 72(8):2209–2222.
- Marra, G. and Wood, S. N. (2012). Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics*, 39:53–74.
- Martell, S. and Stewart, I. (2013). Towards defining good practices for modeling time-varying selectivity. *Fisheries Research*, 158:84–95.
- Martínez, B., Arenas, F., Trilla, A., Viejo, R. M., and Carreño, F. (2015). Combining physiological threshold knowledge to species distribution models is key to improving forecasts of the future niche for macroalgae. *Global change biology*, 21(4):1422–1433.
- Massé, J., Duhamel, E., Petitgas, P., Doray, M., and Huret, M. (2016). Spring acoustic surveys: PELGAS survey. pelagic surveys series for sardine and anchovy in ices areas viii and ix (WGACEGG). *Towards an Ecosystem Approach. ICES Cooperative Research Report*, 332.

- Maunder, M. N. and Harley, S. J. (2011). Using cross validation model selection to determine the shape of nonparametric selectivity curves in fisheries stock assessment models. *Fisheries Research*, 110:283–288.
- Maunder, M. N. and Piner, K. R. (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science*, 72:7–18.
- Maunder, M. N. and Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, 142:61–74.
- McAllister, M., Starr, P., Restrepo, V., and Kirkwood, G. (1999). Formulating quantitative methods to evaluate fishery-management systems: what fishery processes should be modelled and what trade-offs should be made? *ICES Journal of Marine Science*, 56(6):900–916.
- McAllister, M. K. and Ianelli, J. N. (1997). Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(2):284–300.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*, volume 37. CRC press.
- Melle, W., Runge, J., Head, E., Plourde, S., Castellani, C., Licandro, P., Pierson, J., Jonasdottir, S., Johnson, C., Broms, C., Debes, H., Falkenhaus, T., Gaard, E., Gislason, A., Heath, M., Niehoff, B., Nielsen, T. G., Pepin, P., Stenevik, E. K., and Chust, G. (2014). The north atlantic ocean as habitat for *Calanus finmarchicus*: Environmental factors and life history traits. *Progress in Oceanography*, 129:244–284.
- Merow, C., Smith, M. J., Edwards Jr, T. C., Guisan, A., McMahon, S. M., Normand, S., Thuiller, W., Wüest, R. O., Zimmermann, N. E., and Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37(12):1267–1281.
- Methot, R. (1986). Synthetic estimates of historical abundance and mortality for northern anchovy, *engraulis mordax*. NMFS, Southwest Fish. Cent., Admin. Rep. LJ.
- Methot, R. (2009). User manual for stock synthesis. *NOAA Fisheries, Seattle, USA*.

- Methot, R. D. (1989). Synthetic estimates of historical abundance and mortality for northern anchovy. In *American Fisheries Society Symposium*, volume 6, pages 66–82.
- Methot, R. D. and Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142:86–99.
- Meyer, R. and Millar, R. B. (1999). Bayesian stock assessment using a state–space implementation of the delay difference model. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(1):37–52.
- Michielsens, C. G., McAllister, M. K., Kuikka, S., Pakarinen, T., Karlsson, L., Romakkaniemi, A., Perä, I., and Mäntyniemi, S. (2006). A Bayesian state space mark recapture model to estimate exploitation rates in mixed-stock fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(2):321–334.
- Millar, C. P., Jardim, E., Scott, F., Osio, G. C., Mosqueira, I., and Alzorriz, N. (2014). Model averaging to streamline the stock assessment process. *ICES Journal of Marine Science*, 72:93–98.
- Millar, R. B. (2011). Maximum likelihood estimation and inference: with examples in r, sas and admb. *John Wiley & Sons*.
- Millar, R. B. and Meyer, R. (2000). Bayesian state-space modeling of age-structured data: fitting a model is just the beginning. *Canadian Journal of Fisheries and Aquatic Sciences*, 57(1):43–50.
- Minchin, P. R. (1987). Simulation of multidimensional community patterns: Towards a comprehensive model. *Vegetatio*, 71:145–156.
- Monnahan, C. C., Branch, T. A., Thorson, J. T., Stewart, I. J., and Szuwalski, C. S. (2019). Overcoming long Bayesian run times in integrated fisheries stock assessments. *ICES Journal of Marine Science*, 76(6):1477–1488.
- Monnahan, C. C. and Kristensen, K. (2018). No-u-turn sampling for fast Bayesian inference in ADMB and TMB: Introducing the adnuts and tmbstan R packages. *PloS one*, 13(5):e0197954.

- Monnahan, C. C., Muradian, M. L., and Kuriyama, P. T. (2014). A guide for Bayesian analysis in AD Model Builder. <http://www.admb-project.org/developers/mcmc/>.
- Monnahan, C. C., Thorson, J. T., and Branch, T. A. (2017). Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods in Ecology and Evolution*, 8(3):339–348.
- Morris, T., R White, I., and Crowther, M. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11):1–29.
- Mourato, B., Winker, H., Carvalho, F., and Ortiz, M. (2018). Stock assessment of Atlantic blue marlin (*Makaira nigricans*) using a Bayesian state-space surplus production model JABBA. *Collect. Vol. Sci. Pap. ICCAT*, 75(5):1003–1025.
- Needle, C. L. (2001). Recruitment models: diagnosis and prognosis. *Reviews in Fish Biology and Fisheries*, 11(2):95–111.
- Needle, C. L. (2008). Management strategy evaluation for north sea haddock. *Fisheries Research*, 94(2):141–150.
- Newman, K. B., Buckland, S., Lindley, S., Thomas, L., and Fernandez, C. (2006). Hidden process models for animal population dynamics. *Ecological Applications*, 16(1):74–86.
- Nielsen, A. and Berg, C. W. (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research*, 158:96–101.
- Nielsen, J. R., Thunberg, E., Holland, D. S., Schmidt, J. O., Fulton, E. A., Bastardie, F., Punt, A. E., Allen, I., Bartelings, H., Bertignac, M., et al. (2018). Integrated ecological–economic fisheries models—evaluation, review and challenges for implementation. *Fish and Fisheries*, 19(1):1–29.
- Nordhaus, W. D. (2007). Two centuries of productivity growth in computing. *The Journal of Economic History*, 67(1):128–159.
- Oksanen, J., LääRä, E., Tolonen, K., and Warner, B. G. (2001). Confidence intervals for the optimum in the gaussian response function. *Ecology*, 82:1191–1197.
- Oksanen, J. and Minchin, P. R. (2002). Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling*, 157:119–129.

- Orio, A., Bergström, U., Florin, A.-B., Lehmann, A., Šics, I., and Casini, M. (2019). Spatial contraction of demersal fish populations in a large marine ecosystem. *Journal of Biogeography*, 46(3):633–645.
- Pante, E. and Simon-Bouhet, B. (2013). marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS One*, 8(9):e73051.
- Paradinas, I., Conesa, D., Pennino, M. G., Muñoz, F., Fernández, A. M., López-Quílez, A., and Bellido, J. M. (2015). Bayesian spatio-temporal approach to identifying fish nurseries by validating persistence areas. *Marine Ecology Progress Series*, 528:245–255.
- Parker, K. (1980). A direct method for estimating northern anchovy engraulis mordax spawning biomass.
- Parrish, R., Serra, R., and Grant, W. (1989). The monotypic sardines, sardina and sardinops: their taxonomy, distribution, stock structure, and zoogeography. *Canadian Journal of Fisheries and Aquatic Sciences*, 46(11):2019–2036.
- Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., et al. (2001). Estimating uncertainty in fish stock assessment and forecasting. *Fish and Fisheries*, 2:125–157.
- Patterson, K. and Kirkwood, G. (1995). Comparative performance of adapt and laurec–shepherd methods for estimating fish population parameters and in stock management. *ICES Journal of Marine Science*, 52(2):183–196.
- Pearson, R. G., Dawson, T. P., Berry, P. M., and Harrison, P. (2002). Species: a spatial evaluation of climate impact on the envelope of species. *Ecological modelling*, 154(3):289–300.
- Pedersen, E., Miller, D., Simpson, G. L., and Ross, N. (2019). Hierarchical generalized additive models: an introduction with mgcv. *PeerJ*.
- Pella, J. J. and Tomlinson, P. K. (1969). A generalized stock production model. *Inter-American Tropical Tuna Commission Bulletin*, 13(3):416–497.
- Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., and Araújo, M. B. (2011). *Ecological niches and geographic distributions (MPB-49)*, volume 56. Princeton University Press.

- Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C., and Guisan, A. (2017). Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Global Ecology and Biogeography*, 26(3):275–287.
- Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190(3-4):231–259.
- Planque, B., Bellier, E., and Lazure, P. (2007). Modelling potential spawning habitat of sardine (*sardina pilchardus*) and anchovy (*engraulis encrasicolus*) in the bay of biscay. *Fisheries Oceanography*, 16(1):16–30.
- Plummer, M. (2003). Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*, p. 125. Vienna.
- Plummer, M., Best, N., Cowles, K., and Vines, K. (2019). *Package ‘coda’*. <http://CRAN.R-project.org/package=coda>.
- Pocheville, A. (2015). The ecological niche: history and recent controversies. pages 547–586.
- Pomarede, M., Hillary, R., Ibaibarriaga, L., Bogaards, J., and Apostolaki, P. (2010). Evaluating the performance of survey-based operational management procedures. *Aquatic Living Resources*, 23(1):77–94.
- Pope, J. (1972). An investigation of the accuracy of virtual population analysis using cohort analysis. *ICNAF Research Bulletin*, 9(10):65–74.
- Pope, J. and Shepherd, J. (1982). A simple method for the consistent interpretation of catch-at-age data. *ICES Journal of Marine Science*, 40(2):176–184.
- Prellezo, R., Carmona, I., and García, D. (2016). The bad, the good and the very good of the landing obligation implementation in the bay of biscay: A case study of basque trawlers. *Fisheries Research*, 181:172–185.
- Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology letters*, 3(4):349–361.
- Punt, A. E. (2003). Evaluating the efficacy of managing west coast groundfish resources through simulations. *Fishery Bulletin*, 101(4):860–873.

- Punt, A. E., A'mar, T., Bond, N. A., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A., Haltuch, M. A., Hollowed, A. B., and Szuwalski, C. (2014). Fisheries management under climate and environmental uncertainty: control rules and performance simulation. *ICES Journal of Marine Science*, 71(8):2208–2220.
- Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A., and Haddon, M. (2016). Management strategy evaluation: best practices. *Fish and Fisheries*, 17(2):303–334.
- Punt, A. E., Dorn, M. W., and Haltuch, M. A. (2008). Evaluation of threshold management strategies for groundfish off the us west coast. *Fisheries Research*, 94(3):251–266.
- Punt, A. E., Dunn, A., Elvarsson, B. T., Hampton, J., Hoyle, S. D., Maunder, M. N., Methot, R. D., and Nielsen, A. (2020). Essential features of the next-generation integrated fisheries stock assessment package: A perspective. *Fisheries Research*, 229:105617.
- Punt, A. E. and Kennedy, R. B. (1997). Population modelling of tasmanian rock lobster, *jasus edwardsii*, resources. *Marine and Freshwater Research*, 48(8):967–980.
- Punt, A. E. and Maunder, M. N. (2013). Stock synthesis: Advancing stock assessment application and research through the use of a general stock assessment computer program. *Fisheries Research*, (142):1–2.
- Punt, A. E., Pribac, F., Walker, T. I., Taylor, B. L., and Prince, J. D. (2000). Stock assessment of school shark, *galeorhinus galeus*, based on a spatially explicit population dynamics model. *Marine and Freshwater Research*, 51(3):205–220.
- Punt, A. E., Smith, A. D., and Cui, G. (2001). Review of progress in the introduction of management strategy evaluation (mse) approaches in australia's south east fishery. *Marine and Freshwater Research*, 52(4):719–726.
- Punt, A. E., Smith, A. D., and Cui, G. (2002). Evaluation of management tools for australia's south east fishery. 2. how well can management quantities be estimated? *Marine and Freshwater Research*, 53(3):631–644.
- Punt, A. E., Smith, D. C., and Smith, A. D. (2011). Among-stock comparisons for improving stock assessments of data-poor stocks: the “robin hood” approach. *ICES Journal of Marine Science*, 68(5):972–981.

- Pya, N. (2018). *scam: Shape Constrained Additive Models*. <https://CRAN.R-project.org/package=scam>.
- Pya, N. and Wood, S. N. (2014). Shape constrained additive models. *Statistics and Computing*, 25(3):543–559.
- Quinn, T. J. and Deriso, R. B. (1999). *Quantitative fish dynamics*. oxford university Press.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Raes, N. and ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, 30(5):727–736.
- Reygondeau, G. and Beaugrand, G. (2011). Future climate-driven shifts in distribution of *Calanus finmarchicus*. *Global Change Biology*, 17(2):756–766.
- Ricker, W. E. (1954). Stock and recruitment. *Journal of the Fisheries Board of Canada*, 11(5):559–623.
- Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape,(with discussion). *Applied Statistics*, 54:507–554.
- Rijnsdorp, A. D., Peck, M. A., Engelhard, G. H., Möllmann, C., and Pinnegar, J. K. (2009). Resolving the effect of climate change on fish populations. *ICES journal of marine science*, 66(7):1570–1583.
- Robinson, L., Elith, J., Hobday, A., Pearson, R., Kendall, B., Possingham, H., and Richardson, A. (2011). Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. *Global Ecology and Biogeography*, 20(6):789–802.
- Rosenthal, J. S. (2000). Parallel computing and Monte Carlo algorithms. *Far east journal of theoretical statistics*, 4(2):207–236.
- Roy, V. (2020). Convergence diagnostics for Markov Chain Monte Carlo. *Annual Review of Statistics and Its Application*, 7:387–412.
- Rue, H., Martino, S., Lindgren, F., Simpson, D., Riebler, A., and Krainski, E. T. (2014). *INLA: Functions which allow to perform full Bayesian analysis of latent*

Gaussian models using Integrated Nested Laplace Approximation. <http://www.R-INLA.org>.

Ruppert, D., Wand, M., and Carroll, R. (2003). *Semiparametric Regression*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Sainsbury, K. J., Punt, A. E., and Smith, A. D. (2000). Design of operational management strategies for achieving fishery ecosystem objectives. *ICES Journal of Marine Science*, 57(3):731–741.

Salomon, M., Markus, T., and Dross, M. (2014). Masterstroke or paper tiger - the reform of the eus common fisheries policy. *Marine Policy*, 47:76–84.

Santos, M., Uriarte, A., Boyra, G., and Ibaibarriaga, L. (2018). Anchovy depm surveys 2003–2012 in the bay of biscay (subarea8): Bioman survey series. In Massé, J., Uriarte, A., Angélico, M. M., and Carrera, P. (Eds.). *Pelagic survey series for sardine and anchovy in ICES subareas 8 and 9 –Towards an ecosystem approach*. ICES Cooperative Research Report No. 332. 268.

Santos, M., Uriarte, A., and Ibaibarriaga, L. (2011). Spawning stock biomass estimates of the bay of biscay anchovy (*engraulis encrasicolus*, l.) in 2010 applying the daily egg production method. *Revista de Investigación Marina*, 18(5):76–91.

Schaefer, M. B. (1954). Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Inter-American Tropical Tuna Commission Bulletin*, 1(2):23–56.

Schmidt, M., Breidenbach, J., and Astrup, R. (2018). Longitudinal height-diameter curves for norway spruce, scots pine and silver birch in norway based on shape constraint additive regression models. *Forest Ecosystems*, 5(1):9.

Schnute, J. (1985). A general theory for analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences*, 42(3):414–429.

Schnute, J., Richards, L., and Olsen, N. (1997). Statistics, software, and fish stock assessment. In *Int. Symp. on Fishery Stock Assessment Models for the 21st Century, Anchorage, Alaska, EEUU. 8-11 Oct 1997*.

Schnute, J. T. (1994). A general framework for developing sequential fisheries models. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(8):1676–1688.

- Schröder, H. K., Andersen, H. E., and Kiehl, K. (2005). Rejecting the mean: Estimating the response of fen plant species to environmental factors by non-linear quantile regression. *Journal of Vegetation Science*, 16(4):373–382.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6:461–464.
- Scott, F., Jardim, E., Millar, C. P., and Cerviño, S. (2016). An applied framework for incorporating multiple sources of uncertainty in fisheries stock assessments. *PloS one*, 11(5):e0154922.
- Scott, J., Heglund, P., Morrison, M., Haufler, J., Raphael, M., Wall, W., and Samson, F. (2002). Predicting species occurrences: issues of scale and accuracy.
- Sharma, R., Porch, C. E., Babcock, E. A., Maunder, M. N., and Punt, A. E. (2019). Recruitment: Theory, estimation, and application in fishery stock assessment models. *Fisheries Research*.
- Shepherd, J. (1999). Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. *ICES Journal of Marine Science*, 56(5):584–591.
- Silva, A., Garrido, S., Ibaibarriaga, L., Pawlowski, L., Riveiro, I., Marques, V., Ramos, F., Duhamel, E., Iglesias, M., Bryere, P., et al. (2019). Adult-mediated connectivity and spatial population structure of sardine in the bay of biscay and iberian coast. *Deep Sea Research Part II: Topical Studies in Oceanography*, 159:62–74.
- Smith, A. (1994). Management strategy evaluation: the light on the hill. *Population dynamics for fisheries management*, pages 249–253.
- Soberón, J. and Arroyo-Peña, B. (2017). Are fundamental niches larger than the realized? testing a 50-year-old prediction by hutchinson. *PloS one*, 12(4):e0175138.
- Soberon, J. and Nakamura, M. (2009). Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, 106 Suppl 2:19644–50.
- Stasinopoulos, M., Rigby, B., Akantziliotou, C., Heller, G., Ospina, R., and Stasinopoulos, M. M. (2019). *Package ‘gamlss.dist’*.

- STECF (2020). Monitoring the performance of the common fisheries policy (STECF-adhoc-20-01). Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-18115-6.
- Stewart, I. J., Hicks, A. C., Taylor, I. G., Thorson, J. T., Wetzel, C., and Kupschus, S. (2012). A comparison of stock assessment uncertainty estimates using maximum likelihood and Bayesian methods implemented with the same model framework. *Fisheries Research*, 142:37–46.
- Stewart, I. J. and Martell, S. J. (2015). Reconciling stock assessment paradigms to better inform fisheries management. *ICES Journal of Marine Science*, 72(8):2187–2196.
- Storch, L. S., Glaser, S. M., Ye, H., and Rosenberg, A. A. (2017). Stock assessment and end-to-end ecosystem models alter dynamics of fisheries data. *PLoS one*, 12(2):e0171644.
- Stratoudakis, Y., Coombs, S., de Lanzós, A. L., Halliday, N., Costas, G., Caneco, B., Franco, C., Conway, D., Santos, M. B., Silva, A., et al. (2007). Sardine (*sardina pilchardus*) spawning seasonality in european waters of the northeast atlantic. *Marine Biology*, 152(1):201–212.
- Su, Y.-S., Yajima, M., Su, M. Y.-S., and SystemRequirements, J. (2015). *Package ‘R2jags’*. <http://CRAN.R-project.org/package=R2jags>.
- Subbey, S., Devine, J. A., Schaarschmidt, U., and Nash, R. D. (2014). Modelling and forecasting stock–recruitment: current and future perspectives. *ICES Journal of Marine Science*, 71(8):2307–2322.
- Taylor, I., Stewart, I., Hicks, A., Garrison, T., Punt, A., Wallace, J., Wetzel, C., Thorson, J. T., Takeuchi, Y., Ono, K., et al. (2019). Package ‘r4ss’. <https://cran.r-project.org/web/packages/r4ss/r4ss.pdf>.
- Team, R. C. (2015). A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.
- Ter Braak, C. J. and Looman, C. W. (1986). Weighted averaging, logistic regression and the gaussian response model. *Vegetatio*, 65(1):3–11.
- Then, A. Y., Hoenig, J. M., Hall, N. G., Hewitt, D. A., and editor: Ernesto Jardim, H. (2015). Evaluating the predictive performance of empirical estimators of natural

- mortality rate using information on over 200 fish species. *ICES Journal of Marine Science*, 72(1):82–92.
- Thomas, L., Buckland, S. T., Newman, K. B., and Harwood, J. (2005). A unified framework for modelling wildlife population dynamics. *Australian & New Zealand Journal of Statistics*, 47(1):19–34.
- Thorson, J. T. (2019). Perspective: Let’s simplify stock assessment by replacing tuning algorithms with statistics. *Fisheries Research*, 217:133–139.
- Thorson, J. T. and Taylor, I. G. (2014). A comparison of parametric, semi-parametric, and non-parametric approaches to selectivity in age-structured assessment models. *Fisheries Research*, 158:74–83.
- Thorson, J. T., Zhou, S., E. Punt, A., and Smith, A. D. M. (2013). A stepwise-selected spline approximation to time-varying parameters, with application to occupancy modelling. *Methods in Ecology and Evolution*, 4:123–132.
- Thuiller, W., Brotons, L., Araújo, M. B., and Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, 27(2):165–172.
- Tommasi, D., Stock, C. A., Pegion, K., Vecchi, G. A., Methot, R. D., Alexander, M. A., and Checkley Jr, D. M. (2017). Improved management of small pelagic fisheries through seasonal climate prediction. *Ecological Applications*, 27(2):378–388.
- Uriarte, A., Sánchez, S., Citores, L., , and Ibaibarriaga, L. (2020). Planteamiento y evaluación mediante mse de reglas de explotación de la sardina del golfo de vizcaya. Elaborado por AZTI para Eusko Jaurlaritza - Gobierno Vasco, Dpto. Agricultura, Pesca y Alimentación, Viceconsejería de Política e Industria Alimentaria, Dirección de Innovación y Desarrollo Tecnológico (Cliente). 48 pp (+ anexo B).
- Vaart, V. D. and W., A. (1998). *Asymptotic Statistics*. Cambridge University Press.
- Véron, M., Duhamel, E., Bertignac, M., Pawlowski, L., and Huret, M. (2020). Major changes in sardine growth and body condition in the bay of biscay between 2003 and 2016: temporal trends and drivers. *Progress in Oceanography*, 182:102274.
- Villarino, E., Chust, G., Licandro, P., Butenschön, M., Ibaibarriaga, L., Larrañaga, A., and Irigoien, X. (2015). Modelling the future biogeography of north atlantic

zooplankton communities in response to climate change. *Marine Ecology Progress Series*, 531:121–142.

Wahba, G. (1983). Bayesian "confidence intervals" for the cross-validated smoothing spline. *Journal of the Royal Statistical Society, Series B (Methodological)*:133–150.

Walther, B. A. and Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, 28(6):815–829.

Wang, Y. and Liu, Q. (2006). Comparison of akaike information criterion (aic) and Bayesian information criterion (bic) in selection of stock recruitment relationships. *Fisheries Research*, 77:220–225.

Wayte, S. E. (2013). Management implications of including a climate-induced recruitment shift in the stock assessment for jackass morwong (*nemadactylus macropterus*) in south-eastern australia. *Fisheries Research*, 142:47–55.

Wetzel, C. R. and Punt, A. E. (2011). Performance of a fisheries catch-at-age model (stock synthesis) in data-limited situations. *Marine and Freshwater Research*, 62(8):927–936.

Wiedenmann, J., Wilberg, M. J., Sylvia, A., and Miller, T. J. (2015). Autocorrelated error in stock assessment estimates: implications for management strategy evaluation. *Fisheries Research*, 172:325–334.

Wood, S. N. (2003). Thin plate splines regression. *Journal of the Royal Statistical Society*, 65(1):95–114.

Wood, S. N. (2006a). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Wood, S. N. (2006b). On confidence intervals for generalized additive models based on penalized regression splines. *Australian & New Zealand Journal of Statistics*, 48:445–464.

Wood, S. N. (2019). "mgcv" Mixed GAM Computation Vehicle with Automatic Smoothness Estimation, R package. <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>.

- Wood, S. N., Pya, N., and Säfken, B. (2017). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, 111(516):1548–1563.
- Worm, B., Hilborn, R., Baum, J. K., Branch, T. A., Collie, J. S., Costello, C., Fogarty, M. J., Fulton, E. A., Hutchings, J. A., Jennings, S., et al. (2009). Rebuilding global fisheries. *science*, 325(5940):578–585.
- Yin, Y. and Sampson, D. B. (2004). Bias and precision of estimates from an age-structured stock assessment program in relation to stock and data characteristics. *North American Journal of Fisheries Management*, 24(3):865–879.
- Zimmermann, N. E., Yoccoz, N. G., Edwards, T. C., Meier, E. S., Thuiller, W., Guisan, A., Schmatz, D. R., and Pearman, P. B. (2009). Climatic extremes improve predictions of spatial patterns of tree species. *Proceedings of the National Academy of Sciences*, 106(2):19723–19728.

Annex A

In this annex we proof that imposing concavity constraints in the linear predictor scale, with a logit link function, result in unimodal probability curves, as mentioned in Chapter 2:

Let $p(x), x \in \mathbb{R}$ be the estimated response probability curve obtained when imposing concavity restrictions in the linear predictor scale, with $h = \text{logit}$ as link function.

Then,

$h(p(x))$ is concave $\xrightarrow{1)}$ $h(p(x))$ is quasiconcave $\xrightarrow{2)}$ $h^{-1}(h(p(x))) = p(x)$ is quasiconcave $\xrightarrow{3)}$ $p(x)$ is unimodal

1)

Definition 2.2 (Avriel et al. 1988): A function f defined on the convex set $C \in \mathbb{R}^n$ is called concave if for every $x_1, x_2 \in C$ and $0 \leq \lambda \leq 1$ we have $f(\lambda(x_1 + (1 - \lambda)x_2)) \geq \lambda f(x_1) + (1 - \lambda)f(x_2)$.

Theorem 3.1 (Avriel et al. 1988): Let f be defined on the convex set $C \in \mathbb{R}^n$. It is a quasiconcave function if and only if $f(\lambda(x_1 + (1 - \lambda)x_2)) \geq \min(f(x_1), f(x_2))$ for every $x_1, x_2 \in C$ and $0 \leq \lambda \leq 1$.

It is clear that a concave function is also quasiconcave (not vice versa), since $f(\lambda(x_1 + (1 - \lambda)x_2)) \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \geq \min(f(x_1), f(x_2))$

2)

Proposition 3.2 (Avriel et al. 1988): Let ϕ be a quasiconcave function defined on $C \in \mathbb{R}^n$ and let f be a nondecreasing function on $D \in \mathbb{R}$, containing the range of ϕ . Then the composite function $f\phi(x)$ is also quasiconcave.

Thus, this proof can be generalized to any link function h whose inverse h^{-1} is

nondecreasing (in our case $h(p(x))$ is quasiconcave, and h^{-1} (antilogit function) is a non decreasing function obtaining that the composite $h^{-1}(h(p(x))) = p(x)$ is quasiconcave).

3)

Proposition 3.8 (Avriel et al. 1988): Let f be defined on the interval $C \in \mathbb{R}$ and suppose that it attains its maximum at a point $x^* \in C$. Then f is quasiconcave if and only if it is unimodal on C .

Proofs for Theorem 3.1, Proposition 3.2 and Proposition 3.8 can be found in Avriel et al. (1988).

References

Avriel, M., Diewert, W.E., Schaible, S., Zang, I. Generalized Concavity. *Plenum Press*, 1988.

Annex B

In this annex we present the code for data generation and fitting with each of the proposed methods in Chapter 2. Data for this example is generated using the provided code, and no external dataset needs to be uploaded.

First, we load the required R libraries and we fix the random number generation seed.

```
library(coenocliner)
library(scam)
library(gamlss)
library(mboost)
library(plateau)
set.seed(18)
```

Data generation

For this example, both the environmental variable and the response probability curve for a certain species are simulated. The environmental variable x , is simulated from a normal distribution with $\mu = 12$ and $\sigma = 3$ (note that the simulation study presented in the work is based on real Sea Surface Temperature data). The response probability curve along this variable is generated using the Minchin's function (generalised Beta function), implemented in the `Coenocliner` package with a maximum probability of presence of 1 around $x = 14$, a range of 10 and parameters $\alpha = \gamma = 4$ that control the shape of the curve (Figure B1).

In order to mimic real observations, errors are introduced in the generated curve

using the zero and one inflated BEINF distribution from the `gamlss.dist` package (<https://cran.r-project.org/web/packages/gamlss.dist/gamlss.dist.pdf>), and finally presence absence data are simulated from a binomial distribution with probabilities coming from the generated response curve. Figure B1 shows the theoretical curve and the resulting presence-absence datapoints after the data generating process.

```
#generate einvironmental gradient x with nn points:
nn<-200
x<-sort(rnorm(nn,12,3))

#define paramters for Minchins generalised beta function
Ao <- 1;m <- median(x)+2;r <- 10;a <- 4;g <- 4
params <- list(m = m, A0 = Ao, r = r, alpha = a, gamma = g)

#generate probability curve with the Beta function and
#defined parameters
p_real_all <- as.matrix(coenocline(x, responseModel = "beta",
params = params,countModel = "bernoulli", expectation = TRUE)+1e-6)

#add error from a zero and one inflated dsitribution (BEINF)
sig<-0.1
expit<-rBEINF(length(p_real_all), mu=p_real_all, sigma=sig, nu=sig,
tau=sig)

#generate presence absecence data from the previous probabilities
pa = rbinom(length(p_real_all),1,expit)

#create data frame with generated data for fitting
df = data.frame(y=pa,x=x,real=p_real_all,real_error=expit)
```

Data fitting

Generated data are fitted using 6 different approaches: GLM, GAM.lk, GAM.hk, SCAM.fixsp, boost and Plateau, using the code provided below. Resulting fitted

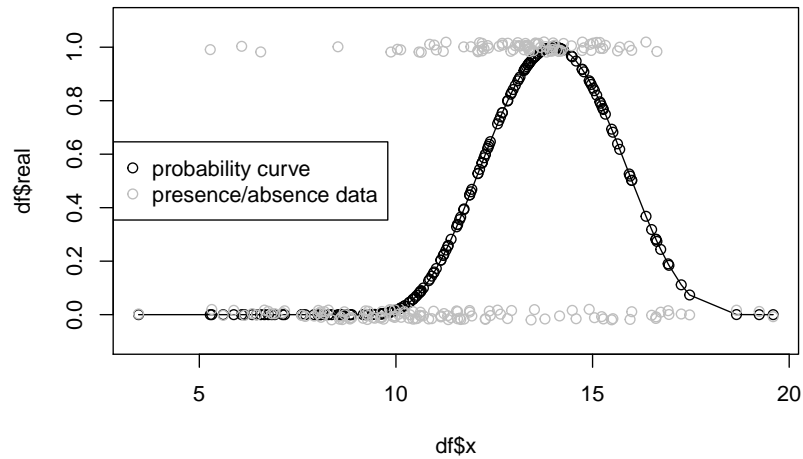


Figure B1: Generated response curve using Minchin's generalised Beta function (black) and simulated presence/absence datapoints (grey).

response curves with each of the methods for this particular example are shown in Figure B2. For this particular simulated dataset all methods are able to detect the maximum around the same value ($m \sim 14$), however they result in different shapes, with GAM.hk not fulfilling the unimodality condition.

#GLM code

```
reg1<-glm( y~x+I(x^2),data=df,family="binomial")
```

For GAM fitting (with `mgcv` package) default smoothing basis (thin plate regression splines) have been used, as found in literature (Chust et al., 2014). No potential change in the estimated curve shape is expected when changing basis in terms of penalized basis type (from default, to penalised splines, `bs="tp"`) or in terms of the order of the penalty (`m=3`, third order penalty instead of second order, `m=2`) (Figure B3).

#GAM_lk k=3 code

```
reg2<-gam(y ~ s(x,k=3), family=binomial(link="logit"),data=df)
```

#GAM_hk k=10 code

```
reg3<-gam(y ~ s(x,k=10), family=binomial(link="logit"),data=df)
```

```
#different basis
reg32<-gam(y ~ s(x,k=10,bs="ps",m=2) ,family=binomial(link="logit"),
data=df)
reg33<-gam(y ~ s(x,k=10,bs="ps",m=3) ,family=binomial(link="logit"),
data=df)
```

For SCAM fitting the smoothing parameter was fixed due to the instability of results when trying to estimate the smoothing parameter. For some of the datasets, convergence was not held. As illustration, a new dataset was simulated where estimating the smoothing parameter with SCAM or fixing it to 1 lead to no convergent results while fixing it to a low number lead to stable values (Figure B4). Thus, regression splines were used, controlling the smoothness with the number of knots, fixed to 8 for this study and fixing the smoothing parameter to a small value (0.0001), so that convergence was reached for all datasets.

```
#SCAM_fixsp code
reg4<-scam(y ~ s(x,k=8,bs="cv",m=2), family=binomial(link="logit"),
data=df,sp=0.0001)
```

When using GAMBOOST the base learner *bmono* was selected, being the only one that is able to include concavity restrictions. The number of knots and degrees of freedom are set to 15 so that the fitting is flexible enough and not dependent on these parameters. It was seen that increasing the number of knots or degrees of freedom did not improve the fit but did increase computational time.

```
#boosting code
reg5<-gamboost(factor(y) ~ bmono(x, knots = 15, df = 15,constraint =
c("concave")),data=df, family = Binomial(link = c("logit")))
```

Plateau method needs some initial values for the optimisation process. They were set to 1 arbitrarily. If convergence is not held, different initial values have to be selected.

```
#Plateau code
initial.pars <- rep(1,5)
reg6<-fit.glm.env(y="y",x.clim=c("x"),initial.pars.input=initial.pars,
data=df)
```

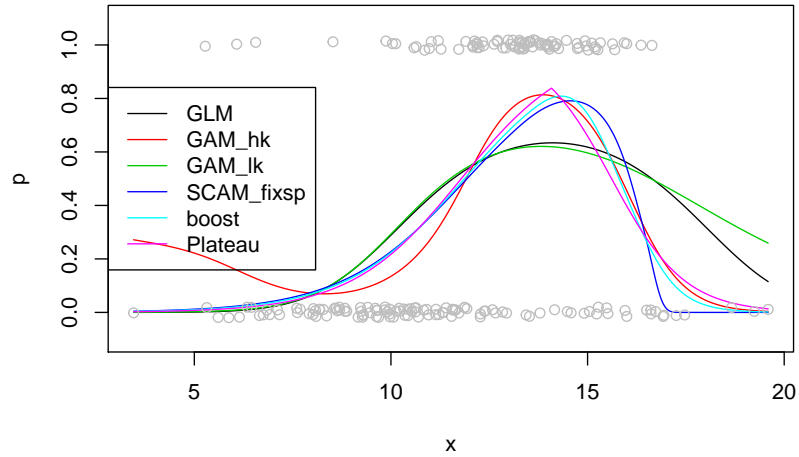


Figure B2: Estimated response curves for each method and simulated presence absence data (grey).

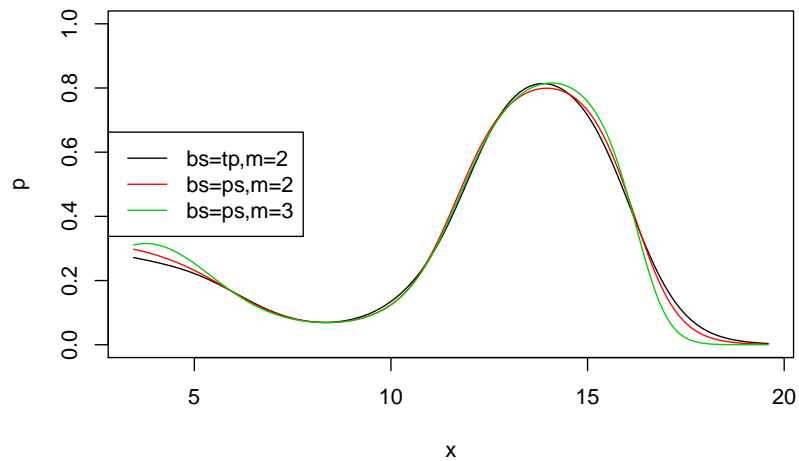


Figure B3: Estimated response curves with different smoothing basis for GAM fittings. The black line corresponds to the default option, thin plate regression splines, whereas the red and green lines represent penalised splines with the order of the penalty equal to 2 and 3 respectively.

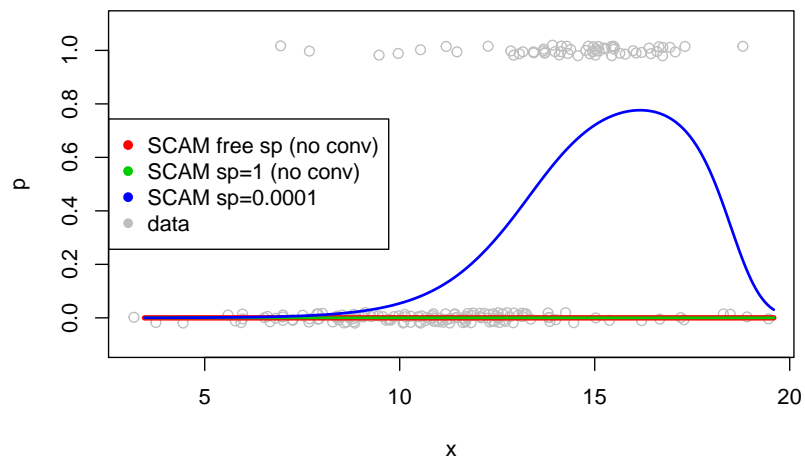


Figure B4: Estimated response curves with SCAM when estimating the smoothing parameter (red), when fixing it to 1 (green) or to 0.0001 (blue).

Annex C

In this annex we provide additional plots mentioned in Chapter 2.

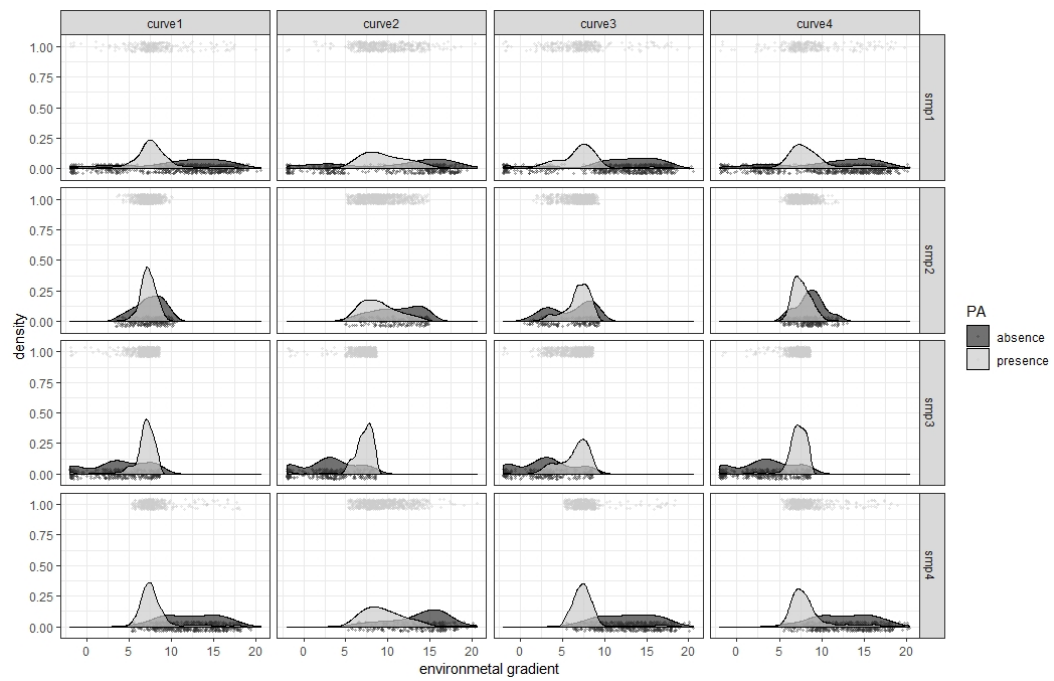


Figure C1: Presence data points and density (in grey) and absence data points and density (in black) along the environmental gradient for each simulation scenario for a single replicate.

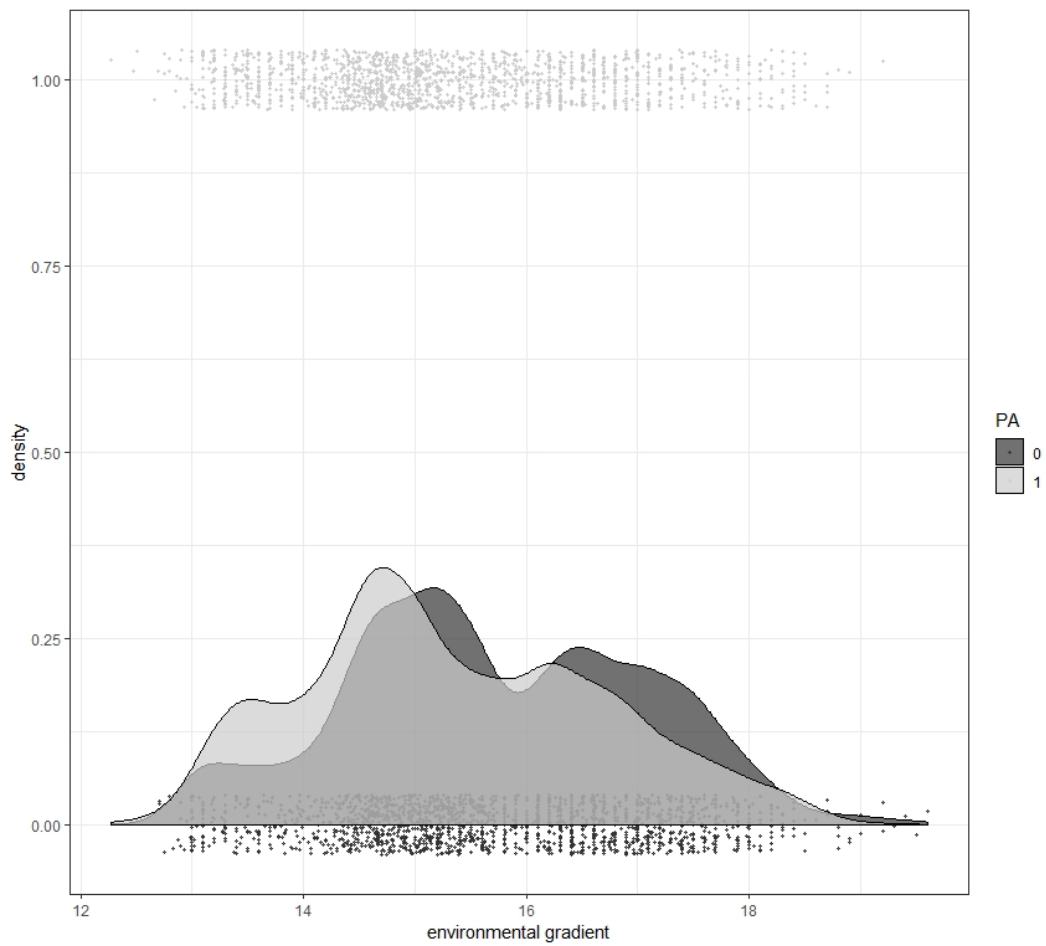


Figure C2: Presence data points and density (in grey) and absence data points and density (in black) along the environmental gradient (SST in this case) for real data used in "Thermal niche for sardine eggs" case study (section 4.1).

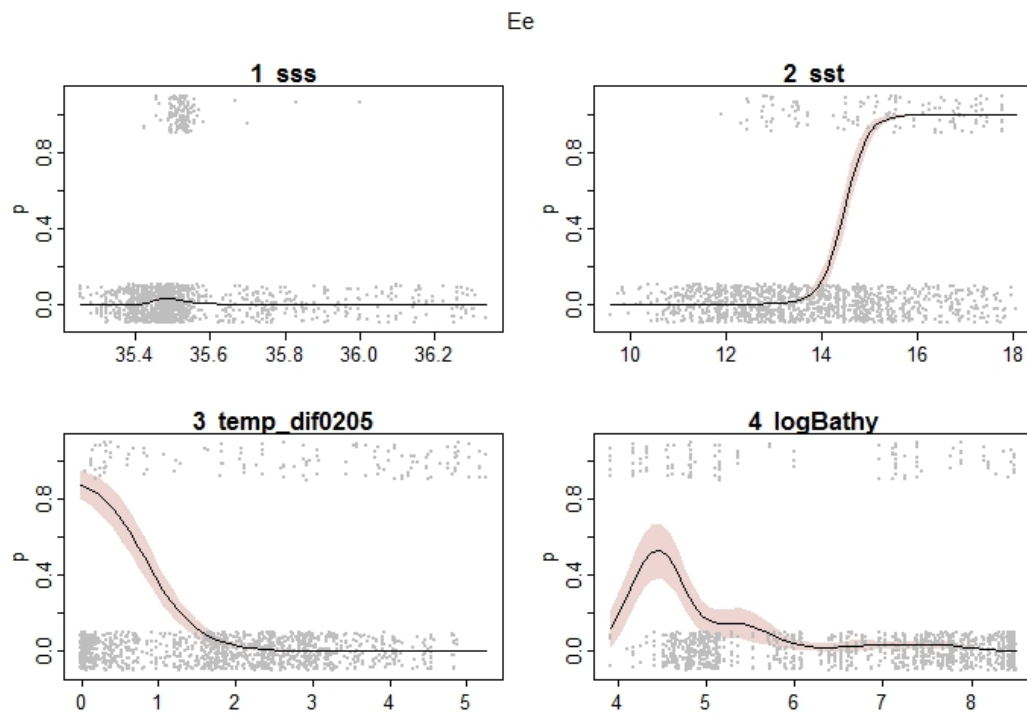


Figure C3: Anchovy (*Engraulis encrasicolus*) egg real presence-absence data (grey points), fitted response curves (in black), with the 95% CI (shaded area) along the selected 4 environmental variables.

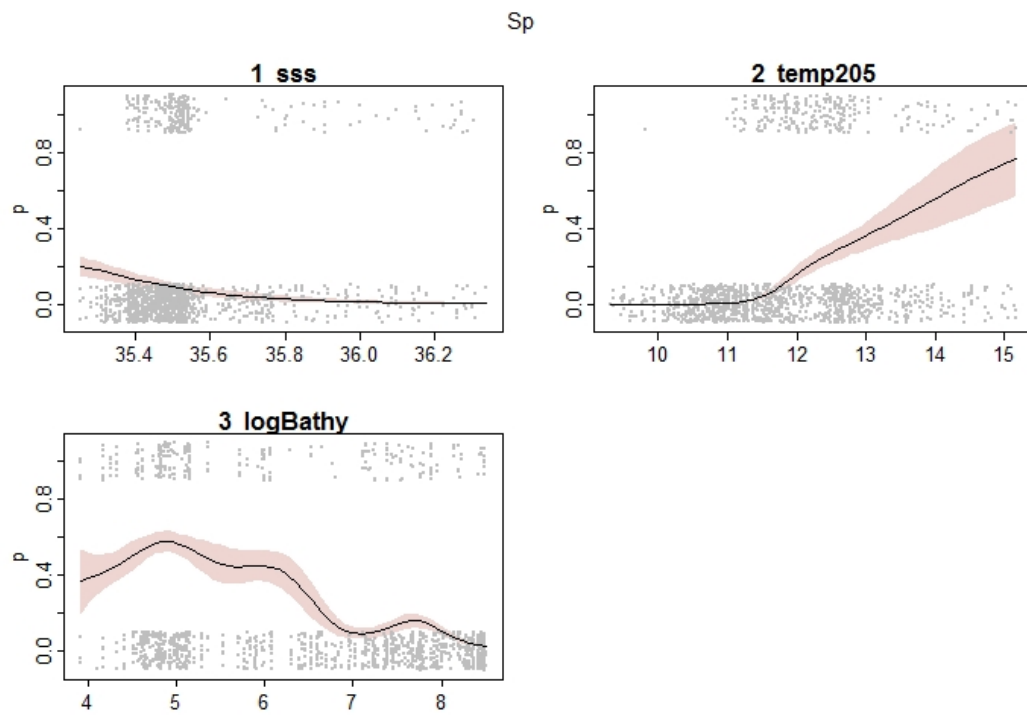


Figure C4: Sardine (*Sardine pilchardus*) egg real presence-absence data (grey points), fitted response curves (in black), with the 95% CI (shaded area) along the selected 3 environmental variables.

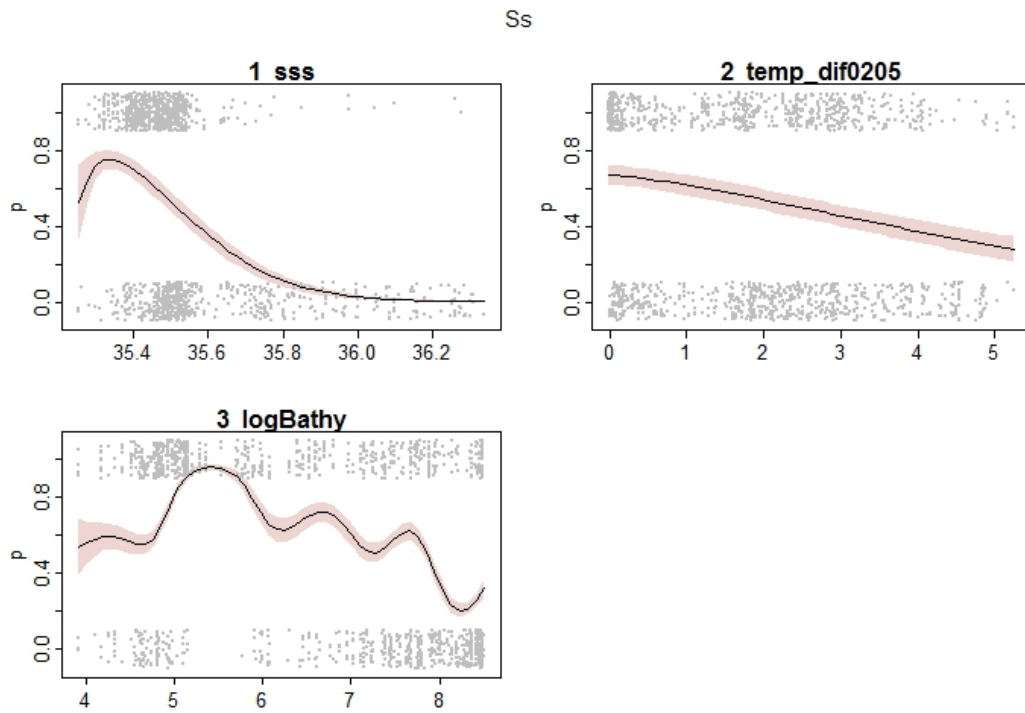


Figure C5: Mackerel (*Scomber scombrus*) egg real presence-absence data (grey points), fitted response curves (in black), with the 95% CI (shaded area) along the selected 3 environmental variables.

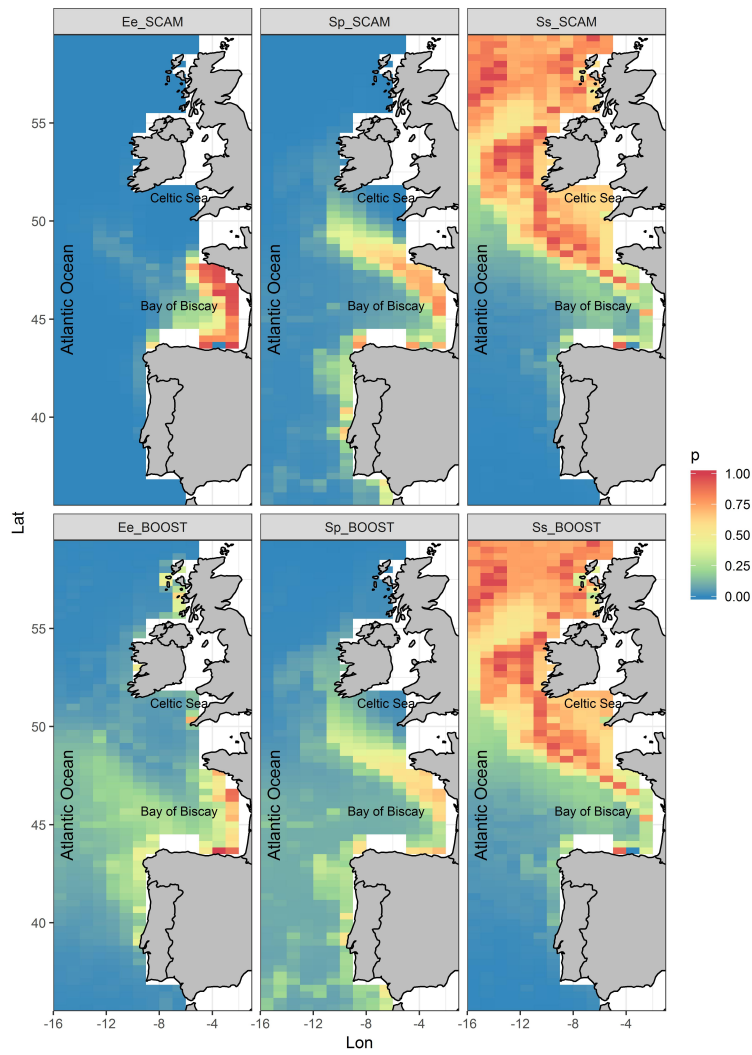


Figure C6: Predicted occurrence probabilities (p) in each map cell for each species (from left to right, anchovy, sardine and mackerel) with SCAM method (up) and Boosting method (bottom).

Annex D

JAGS code for the Bayesian approach of the Bay of Biscay sardine assessment described in Chapter 3:

```
model{

##Recruitment

mu.recru~dlnorm(logmu.sr,tau.sr)
tau.recru~dgamma(a.recru,b.recru)

#Priors on first row and column
for(y in 1:(Y)) {
N[y,1]~dlnorm(log(mu.recru),tau.recru) }
N[Y+1,1] <- N[Y,1]

for(a in 2:(A)) {
N[1,a]~dlnorm(logmu.n1,tau.n1)}

##Population dynamics

for(y in 2:(Y+1)){
```

```

#survival eq.
for(a in 2:(A-1)){N[y,a] <- N[y-1,a-1]*exp(-Z[y-1,a-1])}

N[y,A]<-N[y-1,A-1]*exp(-Z[y-1,A-1]) +N[y-1,A]*exp(-Z[y-1,A])
}

#catch equation:
for(y in 1:Y){

for(a in 1:A) {
Ca[y,a]<-N[y,a]*(1-exp(-Z[y,a]))*F[y,a]/Z[y,a]
catch[y,a]<-Ca[y,a]*WC[y,a]
}

logmu.C[y]<-log(sum(catch[y,1:A]))

#total catch by year

Ctot[y]<-sum(Ca[y,1:A])

#total catch in tons lognormal

Ctons[y]~dlnorm(logmu.C[y],tau.Ctot)
reslogC[y]<-log(Ctons[y])-logmu.C[y]/(sqrt(tau.Ctot))

}

for(y in 1:(Y)){

#CP[y,1:A] ~ dmulti(Cprop[y,],ss)

for(a in 1:A){
C[y,a]~dlnorm(log(Ca[y,a]),tau.C)
reslogCa[y,a]<-(log(C[y,a])-log(Ca[y,a]))/(sqrt(tau.C))
}
}

```

```

}

##Observation equations indices:

for(y in 1:Y){

for(a in 1:A){
tons[y,a]<-N[y,a]*WI[y,a]*mat[y,a]
tonsI[y,a]<-N[y,a]*exp(-Z[y,a]*t_depm)*WI[y,a]*mat[y,a]
tonsIac[y,a]<-N[y,a]*exp(-Z[y,a]*t_depm)*WI[y,a]}

SSB[y]<-sum(tons[y,1:A])
SSBI[y]<-sum(tonsI[y,1:A])
BI[y]<-sum(tonsIac[y,1:A])

#indicecs numbers
for(a in 2:A) {
Ia[y,a-1]<-N[y,a]*exp(-Z[y,a]*t_ac)*qac[a-1]*Q
index[y,a-1]<-Ia[y,a-1]*WI[y,a]
}

logmu.Iac[y]<-log(Qac*BI[y])

#pelgas acoustic

Iactons[y]~dlnorm(logmu.Iac[y],tau.Iactons)
reslogIactons[y]<-(log(Iactons[y])-logmu.Iac[y])/sqrt(tau.Iac)

for(a in 1:(A-1)){
Iac[y,a]~dlnorm(log(Ia[y,a]),tau.Iac)
reslogIac[y,a]<-((log(Iac[y,a])-log(Ia[y,a])))/sqrt(tau.Iac)
}

```

```
}

```

```
#depm egg
logmu.Idepm[y]<-log(SSBI[y]*qdepm)
Idepm[y]~dlnorm(logmu.Idepm[y],tau.Idepm)
#reslogIdepm[y]<-(log(Idepm[y])-logmu.Idepm[y])/sqrt(tau.Idepm)
}
```

```
#DEPM
for(y in 1:length(idx.depm2)){
logmu.Idepm2[y]<-log(SSBI[idx.depm2[y]]*qdepm2)
Idepm2[y]~dlnorm(logmu.Idepm2[y],tau.Idepm2)
#reslogIdepm2[y]<-(log(Idepm2[y])-logmu.Idepm2[y])/sqrt(tau.Idepm2)
}
```

```
##F

```

```
for(y in 1:Y){
for(a in 1:A){

F[y,a]<-fy[y]*sa[a]
Z[y,a]<-F[y,a]+M[a]
}
}
```

```
##PRIORS

```

```
mu.fm~dlnorm(logmu.f,tau.f)
tau.fm~dgamma(a.recru,b.recru)

for(y in 1:(Y)){ fy[y]~dlnorm(log(mu.fm),tau.fm)}
```

```

#fix s[refages]=1 as fixed in ss3
for(a in refages) {sa[a]<-1}
for(a in 1:(refages[1]-1)){sa[a]~dlnorm(logmu.s,tau.s)}
for(a in (refages[length(refages)]+1):A){sa[a]<-1}

qdep1~dlnorm(logmu.qdep1,tau.qdep1)
qdep2~dlnorm(logmu.qdep2,tau.qdep2)

#fix qac[refageq]=1 as fixed in ss3
for(a in refageq) {qac[a]<-1}
for(a in 1:(refageq[1]-1)){qac[a]~dlnorm(logmu.s,tau.s)}
for(a in (refageq[length(refageq)]+1):(A-1)){
qac[a]~dlnorm(logmu.s,tau.s)}

Q~dlnorm(logmu.qac,tau.qac)
Qac~dlnorm(logmu.qac,tau.qac)

#precisions

tau.C~dgamma(a.C,b.C)
tau.Ctot~dgamma(a.Ctot,b.Ctot)

tau.Iac~dgamma(a.Iac,b.Iac)
tau.Iactons~dgamma(a.Iactons,b.Iactons)

tau.Idep1~dgamma(a.Idep1,b.Idep1)
tau.Idep2~dgamma(a.Idep2,b.Idep2)

for(y in 1:length(idx.dep2)){
muIdep2[y]<-exp(logmu.Idep2[y])}

#####quantities to save#####

```

```

sdC<-1/sqrt(tau.C)
sdI<-1/sqrt(tau.Iac)
sdCtot<-1/sqrt(tau.Ctot)

for(a in 1:A){
NO[a]<-N[1,a]}

for(y in 1:Y){

for(a in 1:(A-1)){
muIac[y,a]<-Ia[y,a]}

for(a in 1:A){
muC[y,a]<-Ca[y,a]}

Catch[y]<-exp(logmu.C[y])

muIdep[m][y]<-exp(logmu.Idep[m][y])

rec[y]<-N[y,1]
}
}

```

R commands used to define the fishing mortality, catchabilities and variance submodels the a4a approach of the Bay of Biscay sardine assessment described in Chapter 3:

```

fmodel<-~factor(replace(year,year>2017,2017))+
factor(replace(age,age>3&age<7,3))

qmodel<-list(~factor(replace(age,age>2&age<6,2)),~1,~1,~1)

vmodel<-list(~1,~1,~1,~1,~1)

```