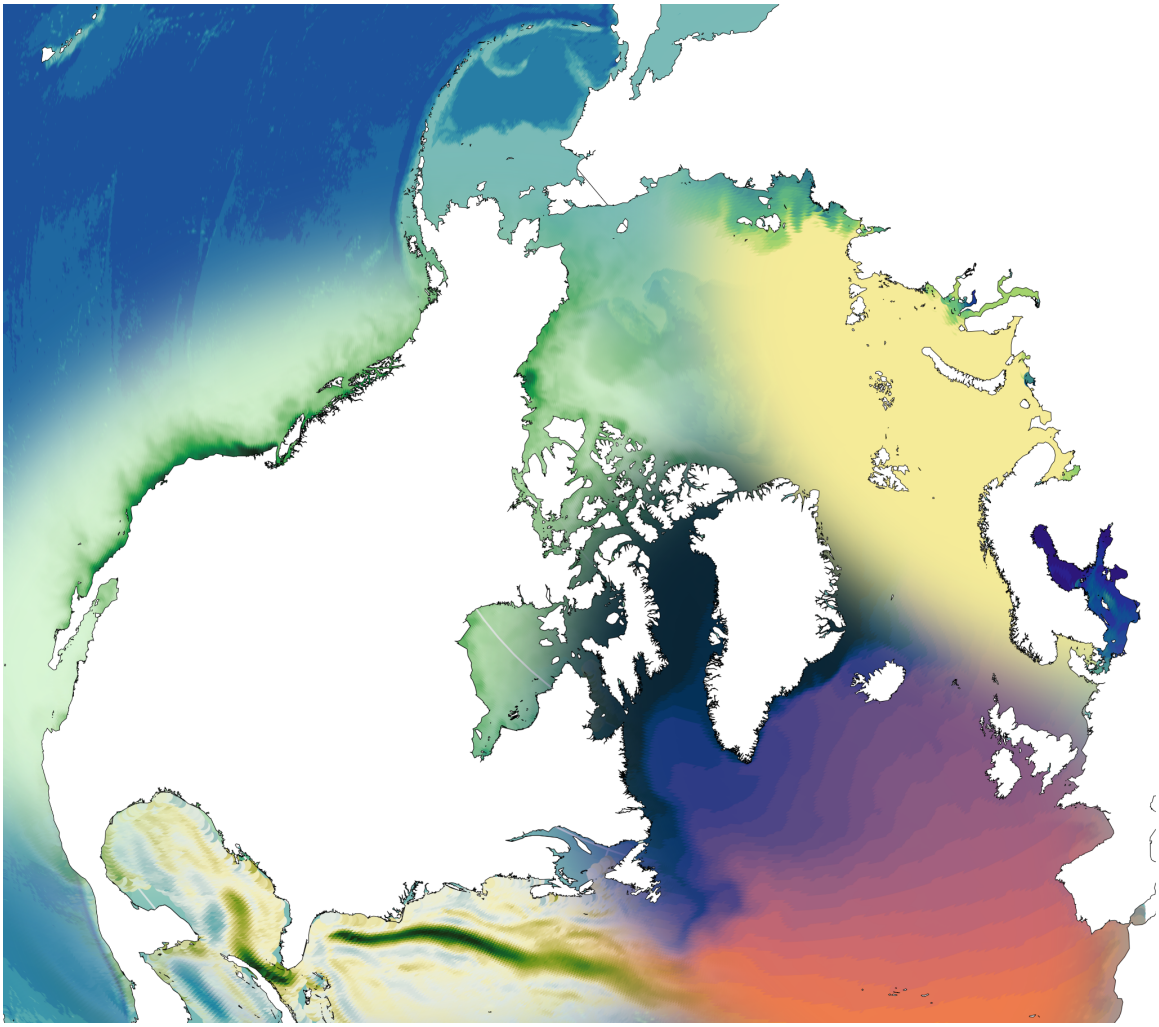
Biogeography and community assembly of demersal fishes across continental shelf seas



Dissertation

zur Erlangung des Doktorgrades (Dr. rer. Nat)
der Mathematisch-Naturwissenschaftlichen-Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Liam MacNeil, M.Sc.
March 2025, Kiel

Submission of the dissertation: 14.03.2025
Oral examination: 30.04.2025

Prof. Dr. Eric Achterberg, Chair

Prof. Dr. Thorsten Reusch, Supervisor

Prof. Dr. Marco Scotti, First reviewer

Prof. Dr. Katja Heubel, Second reviewer

Contents

Summary

Life is not distributed randomly on our Ocean planet. Randomness does appear at varying scales of time and space, but beneath this, life is organized by the environment, by evolutionary history, and by human-induced disturbances. These effects are multifaceted and often viewed as scale dependent. But only in recent years have scientific observations touched large expanses of the oceans and made it possible to test regional- and global-scale drivers which shape biodiversity in the marine realm. Fish are among the most valuable model taxon for testing ecological patterns in the oceans due to their high diversity, dispersal capacity, and the burgeoning availability of data on their distributions, biomass, life-histories, and evolutionary histories. Understanding how different processes shaped fish communities in the modern ocean can also generate insight into how anthropogenic impacts may yet reshape the ecological functioning of the biosphere.

The primary objective of this dissertation is to disentangle patterns of marine fish biomass (Chapter 2), geographic breakpoints in community composition (Chapter 3), and drivers of trait compositions which are proxies for species ecological functions (Chapter 4). To do this, I collated and analyzed scientific bottom trawls which are available throughout continental shelf seas in the Northern Hemisphere. The first manuscript (Chapter 2) focuses on biomass distributions of four demersal species which dominate biomass in the Baltic Sea to describe spatiotemporal biomass shifts and the primary environmental covariates to each species. The second and third manuscripts (Chapter 3–4) scale outwards to include scientific trawls taken in the North Atlantic and northeast Pacific oceans. These works focus on community-wide partitions in species composition (Chapter 3) and large-scale drivers of species traits acting as proxies for ecological functions (Chapter 4), respectively. Together, I show (i) evidence for substantial changes in species biomass distributions in the Baltic Sea when modelling biomass at yearly and seasonal scales, (ii) large-scale biogeographic divides in fish community networks weighted by differing co-occurrence frequencies, and (iii) that fish traits have been primarily shaped by environmental filtering, with a unique degree of evolutionary conservatism in the northeast Pacific. Scaling from a marginal sea to continental-scale ocean shelves, and from select species to community-wide data, this dissertation offers novel insight into seasonal and macroecological patterns which explain the distribution of highly diverse fishes in the modern oceans.

Zusammenfassung

Das Leben ist nicht zufällig auf unserem Ozeanplaneten verteilt. Der Zufall tritt zwar in verschiedenen Maßstäben von Zeit und Raum auf, aber darunter ist das Leben durch die Umwelt, die Evolutionsgeschichte und durch vom Menschen verursachte Störungen organisiert. Diese Auswirkungen sind vielfältig und werden oft als maßstabsabhängig angesehen. Aber erst in den letzten Jahren haben wissenschaftliche Beobachtungen große Teile der Ozeane erfasst und es ermöglicht, die regionalen und globalen Einflussfaktoren auf die biologische Vielfalt im Meer zu untersuchen. Fische gehören zu den wertvollsten Modelltaxa, wenn es darum geht, ökologische Muster in den Ozeanen zu untersuchen, da sie sehr vielfältig sind, sich gut ausbreiten können und immer mehr Daten über ihre Verteilung, Biomasse, lebensgeschichtliche Merkmale und Evolutionsgeschichte zur Verfügung stehen. Wenn man versteht, wie verschiedene Prozesse die Fischgemeinschaften im modernen Ozean geformt haben, kann man auch Erkenntnisse darüber gewinnen, wie anthropogene Einflüsse die ökologische Funktionsweise der Biosphäre umgestalten können.

Das Hauptziel dieser Dissertation besteht darin, Muster der Biomasse von Meeresfischen (Kapitel 2), geografische Bruchstellen in der Zusammensetzung der Lebensgemeinschaften (Kapitel 3) und Einflussfaktoren auf die Merkmalszusammensetzung, die stellvertretend für die ökologischen Funktionen der Arten stehen, zu entschlüsseln (Kapitel 4). Zu diesem Zweck sammelte und analysierte ich wissenschaftliche Grundschleppnetze, die in allen kontinentalen Schelfmeeren der nördlichen Hemisphäre verfügbar sind. Das erste Manuskript (Kapitel 2) konzentriert sich auf die Biomasseverteilung von vier Grundfischarten, die die Biomasse in der Ostsee dominieren, um räumlich-zeitliche Biomasseverschiebungen und die wichtigsten Umweltkovariablen für jede Art zu beschreiben. Im zweiten und dritten Manuskript (Kapitel 3-4) werden wissenschaftliche Schleppnetzfänge im Nordatlantik und Nordostpazifik untersucht. Diese Arbeiten konzentrieren sich auf die gemeinschaftsweite Aufteilung der Artenzusammensetzung (Kapitel 3) bzw. auf die großmaßstäblichen Triebkräfte von Artenmerkmalen, die als Stellvertreter für ökologische Funktionen dienen (Kapitel 4). Gemeinsam zeige ich (i) Beweise für erhebliche Veränderungen in der Verteilung der Biomasse von Arten in der Ostsee, wenn die Biomasse auf jährlicher und saisonaler Ebene modelliert wird, (ii) großräumige biogeografische Trennungen in Fischgemeinschaftsnetzwerken, die durch unterschiedliche Häufigkeiten des gemeinsamen

Auftretens gewichtet werden, und (iii) dass Fischmerkmale in erster Linie durch Umweltfilterung geformt wurden, mit einem einzigartigen Grad an evolutionärem Konservatismus im Nordostpazifik. Durch die Skalierung von einem Randmeer zu kontinentalen Schelfgebieten und von ausgewählten Arten zu gemeinschaftsweiten Daten bietet diese Dissertation neue Einblicke in saisonale und makroökologische Muster, die die Verbreitung hoch diverser Fische in den modernen Ozeanen erklären.

1. Introduction

1.1. Spatiotemporal scaling in ecological patterns

Earth is riddled with patterns of natural phenomena. The realization that life is patchy, forming and aggregating along gradients and across environments, has disproven many assumptions of 19th-century naturalists based on uniformity or randomness (Legendre & Fortin, 1989). Agents of ecological and evolutionary change occur at varying spatial and temporal scales, each with some amount of distinctive patterning. No universal scale for observation exists, rather the mechanisms controlling local and regional patterns stand in hierarchies that are linked (Wiens, 1989); so, we must choose our scale(s) of interest and recognize the observational filter we are imposing towards understanding patterns and the underlying processes (Levin, 1992). The evidence for immense spatial heterogeneity in nature challenged conventional approaches in ecology which had developed in the 20th century, relying on deductive experimental tools in relatively small study systems (Ricklefs & Jenkins, 2011). Though valuable, these methods rarely scaled from small local systems and explained macroecological patterns of ecological communities in the wild (Lawton, 1999). As bigger, broader datasets have become available, and especially since plate tectonics became widely accepted in the 1960s as a structuring force in Earth's natural history, the importance of recovering geographic patterns across populations and communities has become increasingly clear for explaining past and future ecosystems by incorporating life-history patterns or evolutionary origins (Briggs, 2007; Ricklefs & Jenkins, 2011).

1.2. The power of biogeography

Biogeography focuses on the large-scale distribution of living and dead organisms. As a field of study, biogeography has centered on the multiple factors which explain a species preferred area (abundances), environmental tolerances and dispersal limits (range extents), or the relative roles of abiotic and biotic drivers, historical contingencies (e.g., mass extinctions, plate tectonics) or evolutionary histories on their distribution (Brown & Lomolino, 1998). The expansive spatial domain of biogeographic research means it typically extends beyond the purview of studies in community ecology and evolution (Ricklefs & Jenkins, 2011). Biogeography is rooted back to some of the earliest scientific efforts (e.g., Humboldt, 1806), and although it has been claimed to be distinct from macroecology (Blackburn & Gaston,

2002), biogeography has long focused on analyzing patterns of abundance, diversity, distribution across continental or global scales including their environmental, phylogenetic, and functional drivers (Wallace, 1876). Perhaps having developed in relative isolation to research in ecology or evolution— like marine and terrestrial ecology themselves (Webb, 2012)— biogeography became confined in scope (Wiens & Donoghue, 2004). It is my opinion that "macroecology" is a neologism that has been given status as a new branch of science but is fundamentally rooted in biogeography. Although naming disciplines can clarify research objectives (Lawton, 1999), this should not obscure the overarching goals to reach shared understandings of the natural world.

1.3. Links between biogeography and ecological theory

Studies of island biogeography have been seminal in developing ecological theory. The classic equilibrium theory— defining (Melanesian) island richness as a function of species colonization dynamics and extinction rates affected by isolation and geographic area (MacArthur & Wilson, 1967)— set biogeography and ecology more broadly on more rigorous, quantitative footing[1]. Many more ecological hypotheses grew from these advances, recognizing that dynamic equilibrium cannot, for example, generally explain community composition in nature and that many oceanic archipelago fossil records contained more diverse communities until recent human migrations only centuries ago (Brown & Lomolino, 2000). Beyond geography, the concept of community assembly rules emerged to explain patchworks of New Guinea island bird communities, primarily through dispersal and biological interactions creating competitive exclusion of similar taxa (Diamond, 1975). Though incomplete, this work contributed to expanding questions about the strength of different processes shaping the identity and abundance structure of ecological communities. If community assembly is the sum of these processes, the niche concept has attempted to combine the role of environment and ecology to explain and predict species distributions, with a "Hutchinsonian" niche formalizing this combination geometrically in an n-dimensional hypervolume (Hutchinson, 1957). Like equilibrium theory in island biogeography, niche processes have been transformative in ecologists thinking of resource availability, competition, and environmental tolerances to explain biogeographic patterns (Holt, 2009).

[1] It is widely unappreciated, and at the time ignored, that a "theory of island biogeography" based on population equilibrium was independently derived by a PhD student, Eugene G. Monroe, nearly 20 years earlier at Cornell University (Munroe, 1948; Brown & Lomolino, 1989).

By scaling up observations beyond local experiments, biogeography has generated ecological hypotheses which even question the role of species interactions. The unified neutral theory of biodiversity and biogeography stipulates that for trophically similar organisms (e.g., herbivores), competitive equivalence reigns such that they're ecologically identical, and variations in abundance or distribution are due to random variations in demography through births, deaths, migration, and speciation— that is, by ecological drift (Hubbell, 2001). This is a profound assertion as it minimizes the role of interspecific competition and of ecological fitness, which have been cornerstones of the niche concept and of evolution by natural selection; it also elevates the explanatory power of chance. As far as it is unifying, most physicists would be dissatisfied (Abrams, 2001) in that its scope was originally restricted to explaining species richness and the distribution of relative abundances. But neutrality has frequently fit biogeographic data of species abundances quite well, and where failing (e.g., McGill, 2003), provided useful null models to highlight the potential role of niche processes (Rosindell *et al.* 2012). Combinations of niche, neutral, and stochastic processes have also been proposed (Tilman, 2004). For community assembly, ecological neutrality has become a powerful framework for testing the strength of ecological, environmental, or evolutionary influences on biogeography.

1.4. Biogeography in the oceans

These developments in ecological theory have been enormously productive but biased in their application towards land: the vast ocean realm is less studied than the terrestrial world (Hughes *et al.* 2021) and many aspects of marine biogeography remain fundamentally mysterious. A standard biogeographic example in the oceans (and on land) is the latitudinal diversity gradient, peaking in the tropics and diminishing towards the poles (Hillebrand, 2004; Tittensor *et al.* 2010). The proposed causes of this gradients emergence and stability have been multifaceted, including the older age of tropical seas which describe temperate taxa as merely nested and dispersal out of the tropics as recent, greater diversification rates, or by increased niche diversification supported by the larger area of tropical seas and through more biotic interactions (Mittelbach *et al.* 2007). By comparison, A relatively simple community model based on neutral processes of demography, dispersal constraints, and latitudinal gradients in habitat area with temperature-driven species turnover can produce stable richness gradients (Tittensor & Worm, 2016).

The more specific regional differences in species richness have also helped resolve evolutionary histories. For example, although many conceptual models have been proposed to explain the geography of speciation and dispersal, much support has been given to the centre of origin conceptual model (Briggs, 2003; Bowen *et al.* 2016) where most successful speciation events occur in high diversity areas from which species flow outwards. Major origin centres have been identified based on biogeographic patterns of species endemism and geological history, with centres of faunal origins in the modern ocean dominated by the Indo-Pacific and its "Coral Triangle" linking the Philippines, Indonesia, and New Guinea (Briggs & Bowen, 2013). For cold-adapted biota, the North Pacific has been attributed as the greatest engine of extant taxa followed by the Antarctic for extant Austral taxa, and lesser importance given to the western (Caribbean) Tropical Atlantic (Briggs, 2003). Both the continuous biogeographic patterns in latitudinal richness and binned partitions of the oceans based on faunal endemicity are biological windows in the potential role of evolutionary history and contemporary environmental conditions in shaping ecological communities in space and time. Theory aside, biogeography can also inform our understanding of biological invasions, as species from highly diverse centres could naturally flow outwards under favorable environmental conditions (Briggs, 2007) and be accelerated by human-mediated pathways (Seebens *et al.* 2013). To understand where and how biodiversity has changed and may yet be altered, focusing our scientific efforts on the most accessible and well sampled regions in the ocean can help disentangle the signatures of evolutionary history and the modern drivers of marine biodiversity.

1.5. Continental shelf seas

The continental shelf seas (Figure 1.1) represent some of the major biologically productive (Behrenfeld *et al.* 2006) and diverse (Tittensor *et al.* 2010) zones in the global ocean, harboring a substantial magnitude of taxa in relatively small areas and shallow depths. The shelf seas border continental land masses and are characterized by average depths ~200 m, distinct oceanographic currents, sea surface temperatures, and primary production that recurs annually (Sherman, 1991). The ecological communities contained within shelf seas and coastal regions contain many repeatedly sampled groups, including many sessile groups (e.g., invertebrates, Álvarez-Noriega *et al.* 2020; seagrasses, Strydom *et al.* 2023), but also highly mobile fish species (Edgar & Stuart-Smith, 2014). The growing data records of shelf sea ecosystems plus their ecological and economic importance make these valuable regions in the oceans to study

the explanatory power of different assembly mechanisms (environment, phylogeny, traits) across populations or communities. Broader environmental shifts are already in motion (e.g., surface temperature rise, acidification, noise pollution, eutrophication) and portend ecological changes which could substantially alter these ecosystems (He & Silliman, 2019).

Figure 1.1 Continental shelves in the Northern Hemisphere (shaded blue) outlined at 1000 m isobaths using GEBCO gridded bathymetry (Weatherall et al. 2015), covering a majority of the continental shelves and margins north of the equator (Vion & Menot, 2009).

1.6. Fish as test cases

Continental shelf seas have experienced the greatest proportions of human fishing efforts in living memory (Rousseau *et al.* 2024) and through time immemorial owing to the density of human settlement (Fagan, 2017). This feature is a two-way mirror: through anecdotes, paleoenvironmental records, written histories, and contemporary scientific sampling, big datasets have been amassed to characterize broadscale changes in marine communities through time (Jackson *et al.* 2011). Probably owing to their iconic natural histories intertwined with human societies, and their continued economic and nutritional importance as food, fish are a

frequent choice for species distribution models (Melo-Merino *et al.* 2020). As mobile adults, marine fishes can expand into diverse environments, colonizing far-flung ecosystems from their site-of-birth, and appear to have radiated in taxonomically diverse groups (Mora *et al.* 2008). As such, fish communities are replete with patterns for testing biogeographic hypotheses and specific aspects of community assembly, for example the striking latitudinal species richness gradient inverted to speciation rates (Rabosky *et al.* 2018) or the inverse gradient of demersal biomass (van Denderen *et al.* 2023).

Demersal (i.e., bottom-associated) fishes have supported major fisheries globally (van Denderen *et al.* 2024), however widespread bottom trawling fisheries have depleted many targeted demersal fish stocks in shelf seas (Worm *et al.* 2009; Amoroso *et al.* 2018). Although these modern fishery records contain decades— even beyond a century (e.g., Thurstan *et al.* 2010)— of documented catches, they are biased towards target species and opportunistic sample sites typically influenced by the proximity to port and market incentives. These aspects of commercial fishery records limit their applicability in testing for statistical patterns in space or time, which are fundamentally based on assumptions of observational independence. To overcome these limitations and provide unbiased abundance estimates, scientific (i.e., fisheries-independent) trawl data have been developing for decades (Maureaud *et al.* 2021), employing similar gear types and vessels across years, thereby maintaining catchability, and randomized sampling designs (NRC, 2000).

1.7. Structure of the dissertation

This work builds from regional scientific trawl data outwards to incorporate trawls taken across the continental shelf seas of the Northern Hemisphere, towards addressing the following three research questions in chapters:

1. What biomass changes have occurred in the major Baltic Sea demersal fishes within the past two decades?
2. Where do biogeographic boundaries occur in demersal fish communities across the Northern Hemisphere's shelf seas?
3. What are the primary drivers of trait biogeography in demersal fish communities?

Each chapter increases in data complexity and broadly aims to explain how fish demersal fish communities change in time (Chapter 2) and space (Chapter 3–4). The explanatory variables

for each chapter build from spatiotemporal representations of oceanographic conditions and habitat (Chapter 2) to networks of community composition (Chapter 3), and finally a combination of environmental, evolutionary, and human-driven factors (Chapter 4). The position of each chapter can be viewed against the spatiotemporal scaling of other marine phenomena (Figure 1.2), showing how the role of oceanographic variables unifies each chapter in theory, while they differ in considering seasonality and yearly variability versus broad macroecological explanations.

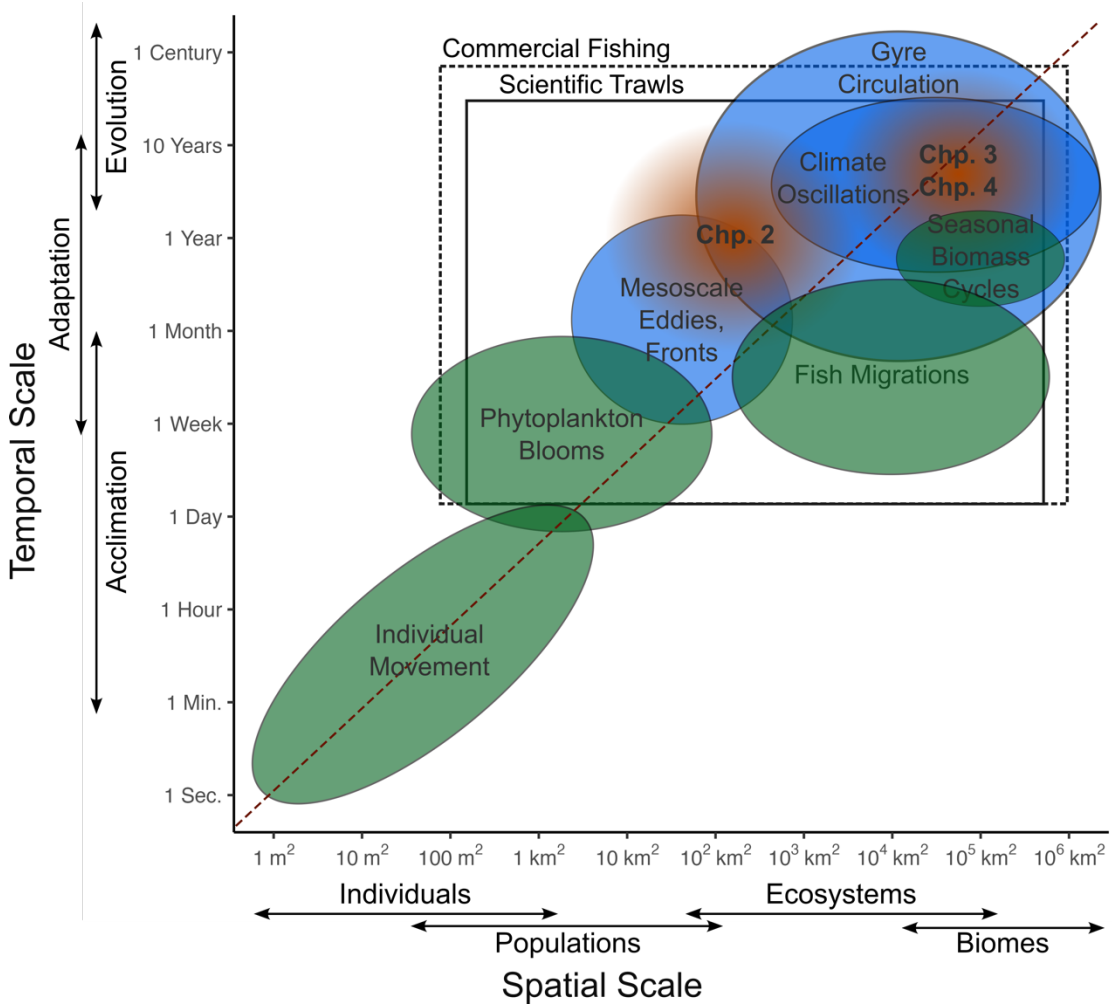


Figure 1.2 Spatial and temporal scaling of select ecological (green), oceanographic and climatic (blue) phenomena in the oceans. Spatial and temporal scaling of select ecological (green), oceanographic and climatic (blue) phenomena in the oceans. As a 2–D simplification, each polygon displays the assumed primary spatiotemporal variation of each process; however, it could also be imagined in 3–D, with peaks and troughs of some chosen response variable (e.g., probabilities of frequency, duration; Stommel, 1963). The x-axis represents spatial extent logarithmically spanning meters to millions of kilometers and individual-level variation to biomes encompassing ocean basins. The y-axis shows time from seconds to one century, including relatively rapid behavioral acclimation by individuals, longer-term adaptation of groups through natural selection and ultimately species evolution. The approximate spatiotemporal coverage for scientific trawl data across continental shelf seas in the Northern Hemisphere is shown in a black box, with commercial fishing shown in dashed black lines. The approximate position of each chapter (Chp. 1 → 3) are highlighted in orange.

1.8. References

Abrams, P. A. (2001). A world without competition. *Nature*, *412*(6850), 858–859. https://doi.org/10.1038/35091120

Álvarez-Noriega, M., Burgess, S. C., Byers, J. E., Pringle, J. M., Wares, J. P., & Marshall, D. J. (2020). Global biogeography of marine dispersal potential. *Nature Ecology & Evolution*, *4*(9), 1196–1203. https://doi.org/10.1038/s41559-020-1238-y

Amoroso, R. O., Pitcher, C. R., Rijnsdorp, A. D., McConnaughey, R. A., Parma, A. M., Suuronen, P., Eigaard, O. R., Bastardie, F., Hintzen, N. T., Althaus, F., Baird, S. J., Black, J., Buhl-Mortensen, L., Campbell, A. B., Catarino, R., Collie, J., Cowan, J. H., Durholtz, D., Engstrom, N., … Jennings, S. (2018). Bottom trawl fishing footprints on the world's continental shelves. *Proceedings of the National Academy of Sciences*, *115*(43). https://doi.org/10.1073/pnas.1802379115

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., Milligan, A. J., Falkowski, P. G., Letelier, R. M., & Boss, E. S. (2006). Climate-driven trends in contemporary ocean productivity. *Nature*, *444*(7120), 752–755. https://doi.org/10.1038/nature05317

Blackburn, T.M. & Gaston, K.J. (2002) Macroecology is distinct from biogeography. *Nature*, *418*, 723.

Bowen, B. W., Gaither, M. R., DiBattista, J. D., Iacchei, M., Andrews, K. R., Grant, W. S., Toonen, R. J., & Briggs, J. C. (2016). Comparative phylogeography of the ocean planet. *Proceedings of the National Academy of Sciences*, *113*(29), 7962–7969. https://doi.org/10.1073/pnas.1602404113

Briggs, J. C. (2003). Marine centres of origin as evolutionary engines. *Journal of Biogeography*, *30*(1), 1–18. https://doi.org/10.1046/j.1365-2699.2003.00810.x

Briggs, J. C. (2007). Marine biogeography and ecology: Invasions and introductions. *Journal of Biogeography*, *34*(2), 193–198. https://doi.org/10.1111/j.1365-2699.2006.01632.x

Briggs, J. C., & Bowen, B. W. (2013). Marine shelf habitat: Biogeography and evolution. *Journal of Biogeography*, *40*(6), 1023–1035. https://doi.org/10.1111/jbi.12082

Brown, J. H., & Lomolino, M. V. (1989). Independent Discovery of the Equilibrium Theory of Island Biogeography. *Ecology*, *70*(6), 1954–1957. https://doi.org/10.2307/1938125

Brown, J. H., & Lomolino, M. V. (1998) Biogeography, 2nd edn. Sunderland, MA. Sinauer Associates.

Brown, J. H., & Lomolino, M. V. (2000). Concluding remarks: Historical perspective and the future of island biogeography theory. *Global Ecology & Biogeography*, *9*, 87–92.

Diamond, J.M. (1975). Assembly of species communities. In: M.L. Cody & J.M. Diamond (Eds.) Ecology and evolution of communities. Cambridge: *Harvard University Press*, Cambridge, Massachusetts, USA.

Edgar, G. J., & Stuart-Smith, R. D. (2014). Systematic global assessment of reef fish communities by the Reef Life Survey program. *Scientific Data*, *1*(1), 140007. https://doi.org/10.1038/sdata.2014.7

Fagan, B. (2017). Fishing: How the Sea Fed Civilization. *Yale University Press*, New Haven, Connecticut, USA.

He, Q., & Silliman, B. R. (2019). Climate Change, Human Impacts, and Coastal Ecosystems in the Anthropocene. *Current Biology*, *29*(19), R1021–R1035. https://doi.org/10.1016/j.cub.2019.08.042

Hillebrand, H. (2004). On the Generality of the Latitudinal Diversity Gradient. *The American Naturalist*, *163*(2), 192–211. https://doi.org/10.1086/381004

Holt, R. D. (2009). Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*, *106*(supplement_2), 19659–19665. https://doi.org/10.1073/pnas.0905137106

Hubbell, S. P. (2001). The unified neutral theory of biodiversity and biogeography. *Princeton University Press*, Princeton, New Jersey, USA.

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, *44*(9), 1259–1269. https://doi.org/10.1111/ecog.05926

Humboldt, A. 1806. Essai sur la géographie des plantes; accompagné d'un tableau physique des régions équinoxales, accompagné d'un tableau physique des régions e´quinoctiales. – Schoel and Co., Paris.

Hutchinson, G. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427.

Jackson, J. B. C., Alexander, K., & Sala, E. (2011). Shifting baselines: The past and the future of ocean fisheries. *Island Press*, Washington, DC, USA

Lawton, J. H. (1999). Are There General Laws in Ecology? *Oikos*, *84*(2), 177. https://doi.org/10.2307/3546712

Legendre, P., & Fortin, M. J. (1989). Spatial pattern and ecological analysis. *Vegetatio*, *80*(2), 107–138. https://doi.org/10.1007/BF00048036

Levin, S. A. (1992). The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology*, *73*(6), 1943–1967. https://doi.org/10.2307/1941447

MacArthur, R. H., & Wilson, E. O. (1967) The theory of island biogeography, Monographs in Population Biology, no. 1. *Princeton University Press*, Princeton, New Jersey, USA

Maureaud, A., Frelat, R., Pécuchet, L., Shackell, N., Mérigot, B., Pinsky, M. L., Amador, K., Anderson, S. C., Arkhipkin, A., Auber, A., Barri, I., Bell, R. J., Belmaker, J., Beukhof, E., Camara, M. L., Guevara-Carrasco, R., Choi, J., Christensen, H. T., Conner, J., … T. Thorson, J. (2021). Are we ready to track climate-driven shifts in marine species across international boundaries? - A global survey of scientific bottom trawl data. *Global Change Biology*, *27*(2), 220–236. https://doi.org/10.1111/gcb.15404

McGill, B. J. (2003). A test of the unified neutral theory of biodiversity. *Nature*, *422*(6934), 881–885. https://doi.org/10.1038/nature01583

Melo-Merino, S. M., Reyes-Bonilla, H., & Lira-Noriega, A. (2020). Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. *Ecological Modelling*, *415*, 108837. https://doi.org/10.1016/j.ecolmodel.2019.108837

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., Harrison, S. P., Hurlbert, A. H., Knowlton, N., Lessios, H. A., McCain, C. M., McCune, A. R., McDade, L. A., McPeek, M. A., Near, T. J., Price, T. D., Ricklefs, R. E., Roy, K., Sax, D. F., … Turelli, M. (2007). Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecology Letters*, *10*(4), 315–331. https://doi.org/10.1111/j.1461-0248.2007.01020.x

Munroe, E.G. (1948) The geographical distribution of butterflies in the West Indies. PhD Dissertation. Cornell University, Ithaca, New York.

Mora, C., Tittensor, D. P., & Myers, R. A. (2008). The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1631), 149–155. https://doi.org/10.1098/rspb.2007.1315

National Research Council (NRC). (2000). Improving the collection, management, and use of marine fisheries data. *National Academies Press*, Washington, DC, USA.

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T. J., Coll, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, *559*(7714), 392–395. https://doi.org/10.1038/s41586-018-0273-1

Ricklefs, R. E., & Jenkins, D. G. (2011). Biogeography and ecology: Towards the integration of two disciplines. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1576), 2438–2448. https://doi.org/10.1098/rstb.2011.0066

Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., & Etienne, R. S. (2012). The case for ecological neutral theory. *Trends in Ecology & Evolution*, *27*(4), 203–208. https://doi.org/10.1016/j.tree.2012.01.004

Rousseau, Y., Blanchard, J. L., Novaglio, C., Pinnell, K. A., Tittensor, D. P., Watson, R. A., & Ye, Y. (2024). A database of mapped global fishing activity 1950–2017. *Scientific Data*, *11*(1), 48. https://doi.org/10.1038/s41597-023-02824-6

Seebens, H., Gastner, M. T., & Blasius, B. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters*, *16*(6), 782–790. https://doi.org/10.1111/ele.12111

Sherman, K. (1991). The Large Marine Ecosystem Concept: Research and Management Strategy for Living Marine Resources. *Ecological Applications*, *1*(4), 349–360. https://doi.org/10.2307/1941896

Stommel, H. (1963). Varieties of Oceanographic Experience: The ocean can be investigated as a hydrodynamical phenomenon as well as explored geographically. *Science*, *139*(3555), 572–576. https://doi.org/10.1126/science.139.3555.572

Strydom, S., McCallum, R., Lafratta, A., Webster, C. L., O'Dea, C. M., Said, N. E., Dunham, N., Inostroza, K., Salinas, C., Billinghurst, S., Phelps, C. M., Campbell, C., Gorham, C., Bernasconi, R., Frouws, A. M., Werner, A., Vitelli, F., Puigcorbé, V., D'Cruz, A., … Serrano, O. (2023). Global dataset on seagrass meadow structure, biomass and production. *Earth System Science Data*, *15*(1), 511–519. https://doi.org/10.5194/essd-15-511-2023

Thurstan, R. H., Brockington, S., & Roberts, C. M. (2010). The effects of 118 years of industrial fishing on UK bottom trawl fisheries. *Nature Communications*, *1*(1), 15. https://doi.org/10.1038/ncomms1013

Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences*, *101*(30), 10854–10861. https://doi.org/10.1073/pnas.0403458101

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., & Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, *466*(7310), 1098–1101. https://doi.org/10.1038/nature09329

Tittensor, D. P., & Worm, B. (2016). A neutral-metabolic theory of latitudinal biodiversity. *Global Ecology and Biogeography*, *25*(6), 630–641. https://doi.org/10.1111/geb.12451

van Denderen, D., Maureaud, A. A., Andersen, K. H., Gaichas, S., Lindegren, M., Petrik, C. M., Stock, C. A., & Collie, J. (2023). Demersal fish biomass declines with temperature across productive shelf seas. *Global Ecology and Biogeography*, *32*(10), 1846–1857. https://doi.org/10.1111/geb.13732

van Denderen, P. D., Jacobsen, N., Andersen, K. H., Blanchard, J. L., Novaglio, C., Stock, C. A., & Petrik, C. M. (2024). Estimating Fishing Exploitation Rates to Simulate Global Catches and Biomass Changes of Pelagic and Demersal Fish. *Earth's Future*, *12*(10), e2024EF004604. https://doi.org/10.1029/2024EF004604

Vion, A., & Menot, L. (2009), Continental margins between 140m and 3500m depth. http://geonetwork.vliz.be/geonetwork/srv/eng/catalog.search#/metadata/dc5fd0a6-7e19-41a8-bd46-498e5f8926d1.

Wallace, A.R. (1876). The Geographical Distribution of Animals, *Cambridge University Press*. Cambridge, UK.

Weatherall, P., Marks, K. M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J. E., Rovere, M., Chayes, D., Ferrini, V., & Wigley, R. (2015). A new digital bathymetric model of the world's oceans. *Earth and Space Science*, *2*(8), 331–345. https://doi.org/10.1002/2015EA000107

Webb, T. J. (2012). Marine and terrestrial ecology: Unifying concepts, revealing differences. *Trends in Ecology & Evolution*, *27*(10), 535–541. https://doi.org/10.1016/j.tree.2012.06.002

Wiens, J. A. (1989). Spatial Scaling in Ecology. *Functional Ecology*, *3*(4), 385. https://doi.org/10.2307/2389612

Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, *19*(12), 639–644. https://doi.org/10.1016/j.tree.2004.09.011

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., Jensen, O. P., Lotze, H. K., Mace, P. M., McClanahan, T. R., Minto, C., Palumbi, S. R., Parma, A. M., Ricard, D., Rosenberg, A. A., … Zeller, D. (2009). Rebuilding Global Fisheries. *Science*, *325*(5940), 578–585. https://doi.org/10.1126/science.1173146

2. Spatial change of dominant Baltic Sea demersal fish across two decades

This chapter is based on a paper which is in review.

MacNeil, L., Madiraca, F., Otto, S., Scotti, M. Spatial change of dominant Baltic Sea demersal fish across two decades. *Authorea. 3rd* review in *Ecology and Evolution*.

2.1. Abstract

The range and biomass distribution of marine fish species offer insights into their underlying niches. Quantitative data are rare compared to occurrences and remain underused in species distribution models (SDMs) to explore realized niches— the actual space occupied by a species shaped by abiotic and biotic factors. Local densities drive differences in species contributions to ecological processes and ecosystem function rather than through presence alone. If a species growth rate is strongly controlled by macro-environmental conditions, then predicting geographical abundance or densities should be possible. We collated twenty years (2001-2020) of standardized scientific bottom trawl data to fit several versions of hierarchical generalized additive models using biomass (kg km^{-2}) of four dominant demersal species (Common dab, European flounder, European plaice, Atlantic cod) within yearly and seasonal (winter and autumn) time windows. Covariates were represented with trawl-level geographic information (position, depth) and high-resolution oceanographic features. This work illustrates species-specific spatiotemporal biomass patterns across two decades and demonstrates superior predictive performance with seasonally variable smoothing terms, revealing seasonally different responses to oceanographic predictors. Firstly, we find relative stasis in Common dab biomass which is linked to the macro-environmental salinity gradient in the western Baltic Sea but with different temperature responses across seasons. Secondly, we show both European flounder and plaice have increased in biomass in the western Baltic Sea with different seasonal relationships to bottom temperature, and that flounder switches between salinity conditions based on season during spawning/feeding periods. Lastly, both juvenile and adult Atlantic cod life stages are shown to have declined most significantly in the Bornholm Deeps and the Gdańsk Deeps. For cod, we conclude that biomass was less reliably predicted in comparison to the other major Baltic demersals studied here, warranting dynamic fishing covariates as a formerly major commercially fished group. These models approach more dynamic species distribution models and are increasingly valuable to constrain uncertainties in biogeographic forecasting which often rely on annually averaged response curves, occurrence data and suitability maps which rarely discriminate between areas of high and low biomass areas in space and time.

2.2. Introduction

The Baltic Sea is among the world's largest inland brackish seas (Leppäranta and Myrberg, 2009). It exhibits relatively low genetic diversity of species (Johannesson and André, 2006) and contains a species-poor fish community that has been heavily relied upon and exploited by humans for centuries, analogous to other intensely exploited marginal seas (Jackson *et al.* 2001; Lotze *et al.* 2006). This shallow, human-modified ecosystem underwent regime shifts driven by overfishing and eutrophication in the 1980–1990s (Österblom *et al.* 2007; Tomczak *et al.* 2013) and is pressingly vulnerable to anticipated climate and environmental changes (Reusch *et al.* 2018). The collapse of Eastern Baltic cod populations illustrates the ecological consequences of these complex dynamics, driven by a combination of low oxygen and salinity, high temperatures, nutrient runoff, and fishing pressure (Möllman *et al.* 2009). The diminished role of cod as a Baltic Sea predator has been related to shifts in distribution and abundance in some flatfish considered as demersal (i.e., bottom associated) competitors (Orio *et al.* 2020; Haase *et al*. 2020). Flatfish represent a diverse assemblage within the Baltic fish community (Smoliński and Radtke, 2017), are of high economic value (e.g., Hyytiäinen *et al.* 2021) and play a relevant ecological role linking upper to lower trophic levels (Scotti *et al.* 2022).

The contemporary geographic distributions of demersal fish species offer key insights into their environmental preferences, range limits, and determine the area within which biotic interactions occur (Guisan *et al.* 2013). As human-driven environmental and climate changes are expected to alter these conditions further, understanding the drivers of species distributions and abundances or biomass is an urgent task (Urban *et al.* 2016). The controls over species distributions are related to a multi-dimensional niche concept (Grinnell, 1914; Hutchinson, 1957). The fundamental niche is often conceived as the determining environmental factors in a species ability to persist under a range of environmental conditions; the realized niche combines abiotic and biotic factors, which influence species ability to establish and achieve positive growth rates (Godsoe *et al.* 2017). The relative importance of abiotic factors and biotic interactions in structuring species niches and resulting biogeography has been increasingly debated with larger empirical datasets and more powerful modelling tools (Wiens, 2011; Pigot and Tobias, 2013; Cadotte and Tucker, 2017; Ovaskainen *et al.* 2017). Species distribution models (SDMs; also referred to as ecological niche models, see also Peterson and Soberón, 2012) have been widely adopted in terrestrial and marine ecology to correlate occurrence data types (presence only, presence-absence) to long-term averaged climate conditions (Guisan and

Thuiller, 2005). As such, SDMs generally relate abiotic effects to species distributions (i.e., fundamental niche; Elith and Leathwick, 2009), which have shown large effects beyond local scales (Soberón, 2007; Thuiller *et al.* 2015; King *et al.* 2021) including across the west-east gradients of temperature and salinity in the Baltic Sea (Pecuchet *et al.* 2016). In the marine realm, SDMs have been used to assess species distributions across broad environmental gradients, invasion risk from alien species, and searching for effective ways to better inform conservation strategies (Robinson *et al.* 2017; Melo-Merino *et al.* 2020). Accurately representing the niches and distributions of important fish species will help prioritize conservation efforts within the Baltic Sea, which displays unique eco-evolutionary constraints between the North Sea-Baltic Sea transition zone (Johannesson *et al.* 2020) and is replete with complex international governance challenges (VanDeveer, 2011).

Quantitative data based on species abundances or biomass are typically less widely available than occurrences (Hortal *et al.* 2015) and have remained underused to disentangle niche relationships (e.g., Howard *et al.* 2014; Rousseau and Betts, 2022), especially in marine environments (Young and Carr, 2015; Erauskin-Extramiana *et al.* 2019; Waldock *et al.* 2022). Species make unequal contributions to ecological processes and ecosystem function through different local densities, rather than through presence alone (Ehrlén and Morris, 2015), and quantitative-based measures can represent an important signal of population growth or collapse (e.g., Ceballos *et al.* 2020). However, abundance patterns do not always appear strongly correlated to species-environment suitability estimates from occurrence-based SDMs (VanDerWal *et al.* 2009; Weber *et al.* 2017; Dallas and Hastings, 2018). If a species growth rate is strongly controlled by macro-environmental conditions and affected by steep environmental gradients, for which there is strong evidence to be acting on the western and central Baltic Sea demersal fish community composition (Pecuchet *et al.* 2016; Frelat *et al.* 2018), then predicting geographical abundance or biomass density patterns should be possible (Waldock *et al.* 2022).

In this work we applied three versions of hierarchical generalized additive models (HGAMs) to predict biomass densities (kg km^{-2}) of four demersal fish species (Common dab, European flounder, European plaice, and Atlantic cod) in the Baltic Sea. Fish biomass densities were derived from internationally coordinated, bi-annual (first and fourth quarter) fisheries-independent trawl data covering 2001-2020 and differentiated by ontogenetic size classes for Atlantic cod. The HGAMs were constructed to fit and predict biomass densities at yearly and

seasonal (biannual) time windows, with each successive model version incorporating greater complexity in predictor-response relationships, building from solely geographic predictors to include seasonal abiotic (sea bottom temperature, salinity, and oxygen) predictors. The HGAM versions which combined geography and seasonal abiotic predictors offered the highest performing models, clarifying seasonal differences in species-specific relationships to environment and predicting spatiotemporal changes in species biomass densities. Accurately summarizing the changing geographic distribution of species densities across highly mobile fish groups could help refine predictions of future shifts under environmental and climate changes.

2.3. Methods

2.3.1. Study Region

The Baltic Sea extends from the northeast Atlantic Ocean shelf into northern Europe and is characterized by strong gradients of salinity and species diversity (Leppäranta and Myrberg, 2009). The study region includes ICES subdivisions 21-29 covering the majority of the Baltic Sea, ranging from the marine-brackish transitions zones (Kattegat, Øresund, and western Baltic Sea) to the Baltic Proper (Arkona Basin, eastern Bornholm Deeps, and eastern Gotland Basin). Seafloor depth is generally shallow (<50 m); three prominently deep areas include Bornholm Deeps (>70 m), Gdańsk Basin (>100 m), and eastern Gotland Basin (>175 m; Figure 2.1). Each of these regions represents important spawning and nursery grounds for many demersal fishes (HELCOM, 2021).

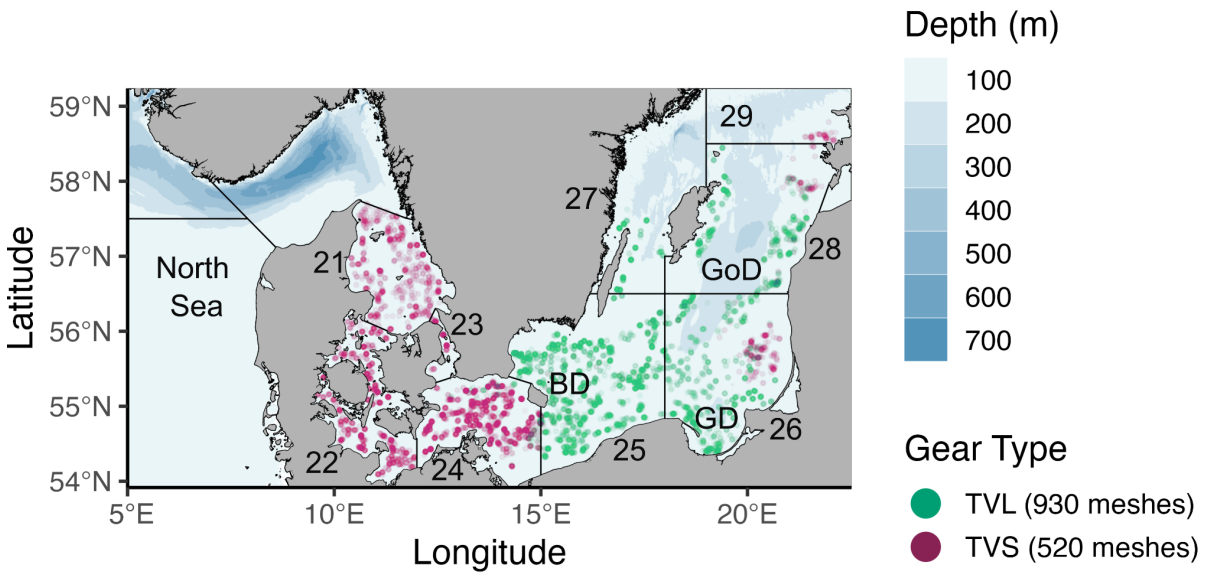


Figure 2.1 The Baltic Sea study region on the European shelf and linked to the North Sea sampling points are colored by gear size/type and underlying bathymetry (Weatherall et al. 2015). The labelled ICES subdivisions (21-29) delimit boundary area and deep basins are annotated including Bornholm Deeps (BD), Gdańsk deeps (GD) and Gotland deeps (GoD).

2.3.2. Observational Data

The Baltic International Trawl Survey (BITS) is a multi-decadal program collecting standardized, fisheries-independent demersal trawls (ICES, 2017). Surveys are regularly bi-annual, occurring in Q1 (Feb-Mar; winter) and Q4 (Oct-Nov; autumn) using standardized gears (TVL 930 meshes, TVS 520 meshes) at randomly stratified stations based on ICES subdivisions (Figure 2.1). Sampling design is described in more detail within ICES protocols (ICES, 2017). We extracted BITS exchange data (haul and catch information) covering bi-annual trawls between 2001-2020, chosen as the beginning of international gear standardization (ICES, 2017) and to match the time window available for our environmental covariates. Observations corresponding to negative bottom oxygen values— produced to represent hydrogen sulfide concentrations in the SCOBI model (Eilola *et al.* 2009)— were removed as they represented rare (<1.3%) extremes in the oxygen data. All subsequent analyses were conducted in the R programming language (R Core Team, 2023). From the full exchange data, hauls were removed which: 1) Did not record haul duration, longitude, or latitude, 2) were not designated as "valid" (<15 or >90 minutes) or "additional", 3) did not record species length class, or 4) did not record catch weight (kg). The four dominant demersal fish species in the Baltic Sea focused on here include Atlantic cod (*Gadus morhua*), European plaice (*Pleuronectes platessa*), European flounder (*Platichthys flesus*), and Common dab (*Limanda*

limanda). We note however that a distinct flounder species (Platichthys solemdali) predominates towards the northern Baltic proper (Momigliano *et al.* 2018), but our limited trawl coverage in these areas suggests European flounder constitute the largest fraction of our observations. For all hauls, zeros were imputed where these species were not recorded, creating a zero-inflated dataset (22.1%) familiar to many ecological studies (Martin *et al.* 2005). The final dataset contains 8991 individual trawls, averaging 450 trawls per year (Figures 6.1-6.5). We standardized catch weight by total swept area (km^2) as a measure of effort (Berg *et al.* 2019) and the unit for modelling represents a measure of species-specific biomass density (kg km^{-2}; hereafter "biomass").

2.3.3. Covariates

Environmental layers were collected to represent oceanographic conditions in the Baltic Sea at bottom depths (Table 2.1). Oceanographic variables derived from high-resolution (~2 nautical miles) physics and biogeochemistry reanalysis products of the NEMO-Nordic ice-ocean model coupled to the Swedish Coastal and Ocean Biogeochemical (SCOBI) model (Eilola *et al.* 2009; Liu *et al.* 2017) available in Copernicus Marine Environmental Service (CMEMS)[2]. All oceanographic variables were downloaded as monthly means. To capture seasonal patterns and better match environmental conditions during trawls, each variable was averaged grid-cellwise by values within the first (Q1; winter) and fourth (Q4; autumn) quarter of each year, weighted by the proportion of trawl observations of each month. We selected abiotic variables at the sea bottom (temperature, oxygen, salinity) from each model as they shape the ecology and distribution of demersal fishes in the Baltic Sea (e.g., Smoliński and Radtke, 2017; Rau *et al.* 2019; Orio *et al.* 2019; Brander, 2022; see also Table 2.1). In preparation of modelling, trawl data were spatially thinned to covariate resolution (~2 nautical miles) to reduce overfitting due to sampling bias of repeated observations (Boria *et al.* 2014). Trawl data were then geographically matched to covariates.

[2] https://marine.copernicus.eu/. Accessed May, 2022.

Table 2.1 Environmental covariates overview for HGAMs, including their expected impacts on demersal fish biomass supported with example references.

Variable	Unit	Source	Data Product	Expected Effect
Sea bottom temperature	Potential temperature (°C)	CMEMS	physics reanalysis	Somatic growth and recruitment rates (Olsson *et al.* 2012)
Sea bottom salinity	PSU (‰)	CMEMS	physics reanalysis	Osmoregulation stress (Kültz, 2015); egg buoyancy and larval mortality (Nissling and Westin, 1997)
Sea bottom oxygen	Dissolved O_2 concentration (mg L^{-1})	CMEMS	biogeochemistry reanalysis	Feeding rates and somatic growth (Brander, 2022)

2.3.4. Hierarchical Generalized Additive Models (HGAMs)

We fit hierarchical generalized additive models (HGAMs; Pedersen *et al.* 2019) onto the spatiotemporal dataset of demersal fish biomass (kg km^{-2}) to test the importance of geography and dynamic, bottom-associated abiotic variables (temperature, oxygen, and salinity) in predicting different species biomass distributions. We split predictions in hierarchical species groups, with an ontogenetic divide between adult (≥35 cm) and juvenile (<35 cm) cod (Valentinsson *et al.* 2019). HGAMs were fit in *mgcv* (Wood, 2017) with the *bam* function using fast restricted maximum likelihood (fREML) with discretized covariates to speed model fitting. To produce parsimonious model versions, variable selection included null space penalization for each smoothing term, penalizing non-informative variables towards zero. Model diagnostics were inspected with mgcv functions *gam.check* for model residuals (Figure 6.6) and *concurvity* as a generalized indication of collinearity in model terms (Dormann *et al.* 2013), which remained below 0.8 (e.g., Leonardi *et al.* 2022; Figure 6.7) and GAMs show robust performance despite concurvity (Wood, 2008). Due to the zero-inflated nature of the dataset, biomass values for the i^{th} observation for species group j (B_{ij}) were modelled using a Tweedie distribution with a log-link function. The μ_{ij} represents the mean biomass value for species j

at a particular time (year and season) and location for the i^{th} observation. The ϕ_j and ρ_j represent a dispersion and power parameter, respectively, that modify the mean-variance relationship for each species *j* in the Tweedie distribution. An error term was assumed for each model version (ϵ_i).

$$B_{ij} \sim Tweedie(\mu_{ij}, \phi_j, \rho_j) \text{ (1)}$$

HGAM version-I (eq. 2) includes a random intercept (β_{season}) by season with geographic predictors for expected mean species biomass, log transformed, combining into a spatiotemporal term a tensor product factor smooth (bs = "fs") between spatial position (longitude, latitude) and year encoded as a factor (e.g., Zabihi-Seissan *et al.* 2024), with unique smoothers per species (by = $species_j$). To incorporate variations by depth, we add a tensor product smooth for the interaction between depth and year (continuous) with species-specific smoothers.

$$log(\mu_{ij}) = \beta_{season} + ti(latitude_i, longitude_i, year_i, by = species_j) + te(depth_i, year_i, by = species_j) + \epsilon_i \text{ (2)}$$

HGAM version-II (eq. 3) builds on version-I, including the primary spatiotemporal term of position and year varying by species, adding species-specific smoothers between depth and each abiotic covariate (temperature, oxygen, salinity) — we found repeating the $te(depth_i, year_i, by = species_j)$ term with abiotic covariates produced almost complete concurvity (>0.96). The separate smoothers allowed group-level variations to each covariate using thin plate regression splines (bs = "tp").

$$log(\mu_{ij}) = \beta_{season} + ti(latitude_i, longitude_i, year_i, by = species_j) + s(depth_i, by = species_j) + s(temperature_i, by = species_j) + s(oxygen_i, by = species_j) + s(salinity_i, by = species_j) + \epsilon_i \text{ (3)}$$

HGAM version-III (eq. 4) extends the relationship to the environment further to permit seasonally independent responses to covariates and different smoothing penalties by species. For model versions-II and -III, we penalize the square second derivative (m = 2) of group-level smoothers to limit excessive curvature (Pedersen *et al.* 2019). Note that we exclude seasonal random smoothers in version-III for the intercept, as these are included in the factor smooths for depth, temperature, oxygen, and salinity.

$$log(\mu_{ij}) = ti(latitude_i, longitude_i, year_i, by = species_j) + s(depth_i, season, by = species_j) + s(temperature_i, season, by = species_j) + s(oxygen_i, season, by = species_j) + s(salinity_i, season, by = species_j) \text{ (4)}$$

The default values for the number of basis functions in smoothers (k) was chosen, maximizing the allowed degrees of freedom (k-1) and for greater variation in individual smoothers. This was decided as even substantially larger effective degrees of freedom (EDF) compared to k (e.g., $EDF_{ti(lon, lat, year)} = 124$ versus $k' = 50$) remained indicated as problematically significant, implying an inadequate number of basis functions in *gam.check*. Despite this, changing k values for individual variables left the diagnostic plots and partial dependence curves unchanged, thus we maintained k values to a default (Table 6.1). To inspect the apparent issue of significant basis dimension, we plotted model residuals against each covariate for each model version (Figure 6.8-6.9) and did not find any clear residual dependence to geographic, temporal, or environmental structure. We note here that the *gam.check* output is a heuristic, and this apparent significance could be due to missing or unmeasured variables to explain species biomass (e.g., spatially-resolved biotic interactions). Trawl observations are clustered in time (year and season), leading us to incorporate an AR1 autocorrelation matrix through the rho parameter in the bam function. The value of rho was determined for each model form (version I = 0.17, II = 0.21, III = 0.15) using the R package *itsadug* (Van Rij *et al.* 2015).

2.3.5. Model Evaluation and Prediction

Models were assessed through k-fold cross validation, with data split into train-test k-folds (k = 10) for fitting and evaluation, respectively (Figure 6.10). Model performance metrics compared observed versus predicted biomass using three regression metrics presented in Waldock *et al.* (2022): Discrimination (Pearson's correlation coefficient), accuracy (Mean Absolute Error; MAE), and dispersion ($\sigma_{est}/\sigma_{obs}$)[3]. After model evaluation under cross validation, final predictions for mapping incorporated all observations (Step 5, option 1 in Roberts *et al.* 2017), favoring prediction quality over error estimation although we expect reasonable error estimates from cross validation. We visualize predicted species biomass patterns as the difference in average biomass within 0.5° grid cells between the last five years (2016-2020) and the first five years (2001-2005) of the trawl survey. Together with the regression metrics, which illustrate an individual model's predictive ability, we also compared model parsimony using AIC (Akaike's information criterion; Akaike, 1974) of final models used for prediction mapping and generated partial dependence curves to assess the relationship of each species to geographic and abiotic covariates using the R package *gratia* (Simpson,

[3] Variance of estimated (σ_{est}) and observed (σ_{obs}) biomass density.

2018). All modelling specifications are detailed in supplementary materials including R script for replication with an ODMAP protocol (Appendix S1-S2; Zurell *et al.* 2020).

2.4. Results

2.4.1. Models Performance

Model predictions measured on k-fold cross validation indicated relatively similar discrimination ($\bar{x}_r$ = 0.71 ± 0.14), precision ($\bar{x}_\sigma$ = 0.59 ± 0.24), and accuracy ($\bar{x}_{MAE}$ = 0.51 ± 0.14), between model versions but we did observe stepwise improvement by including abiotic covariates and at seasonal resolution. The highest predictive performance for every metric was achieved by HGAM version-III including seasonal abiotic covariates. Individual performance metrics for each model and time period are shown in Figure 2.2 and Table 6.2. Across species, Common dab biomass was uniformly predicted with the greatest discrimination, accuracy, and precision. Secondarily, European plaice and flounder were predicted with similar discrimination and precision, but plaice displayed greater MAE in biomass predictions. Both juvenile and adult cod were similar to flounder and plaice in terms of biomass discrimination and precision but rarely predicted at observed values (high MAE). A similar stepwise gain in model parsimony was found by comparing AIC, with the lowest values estimated for model version-III including seasonal abiotic covariates. (Table 6.2).

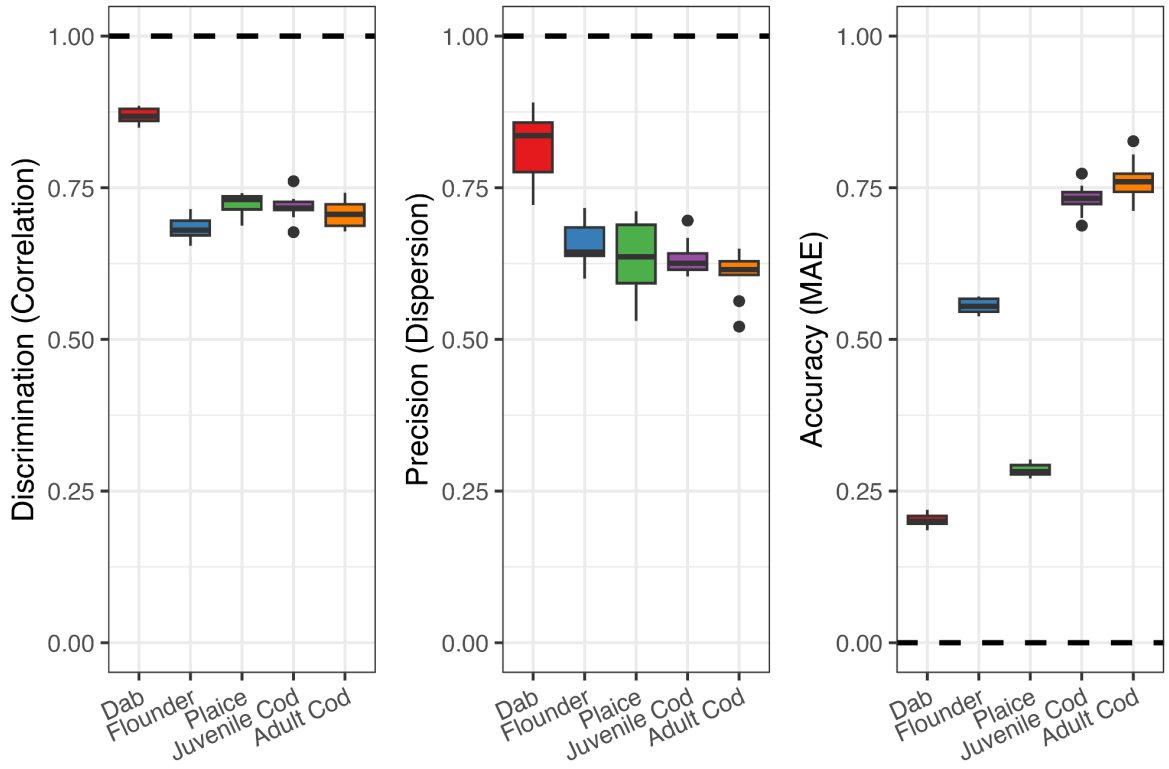


Figure 2.2 Regression metrics of HGAM version-III performance, including geographic and seasonal abiotic predictors, based on discrimination (Pearson's correlation coefficient), accuracy (Mean Absolute Error; MAE), and precision (ratio between variance of estimated and observed biomass density). Boxplot central lines indicate median values and boxes' interquartile range symbolizes 25% and 75% percentiles. Dotted lines for each panel mark the respective optimal metric values.

2.4.2. Seasonal Partial Effects

Assessing the partial effects on the mean response of demersal fish biomass in the Baltic Sea indicates values can differ considerably across depths and oceanographic conditions, and by season (Figure 2.3). The spatiotemporal partial effects are shown by species in Appendix S1 (Figures 6.11-6.16). By depth, mean demersal biomass generally responded positively to deeper waters up to 50-75 m, including in HGAM version-II, capturing strong effects of the shallower western Baltic Sea region (ICES subdivisions 21-24; Figure 6.17). Dab and flounder were the notable exceptions, declining in biomass towards 75 m most strongly in autumn. The abiotic response curves generally differed by season and species, with bottom temperatures producing the greatest differences. Dab displays a positive relationship to temperature in winter and a negative trend in autumn. Flounder showed higher partial effects at bottom temperatures ~8–11°C in winter and only positive partial effects <5°C in autumn whereas plaice shows

relatively higher average biomass at temperatures between 5–10°C in winter and responds variably across a wide temperature range in autumn. Both juvenile and adult cod vary similarly to bottom temperatures, with declining average biomass at higher temperatures in winter and with generally higher biomass at temperatures up to ~9°C in autumn. Species tend to respond positively to increasing bottom oxygen concentrations in both seasons, however each decline in biomass at the highest concentrations >10 mg L-1. The partial effects for bottom salinity are generally positive in winter for all species between 10–30 PSU, but this relationship flips in autumn with higher partial effects for flounder, and cod in fresher condition (<10 PSU) In autumn, these relationships change most strongly in flounder, which decline monotonically in average biomass at higher salinities.

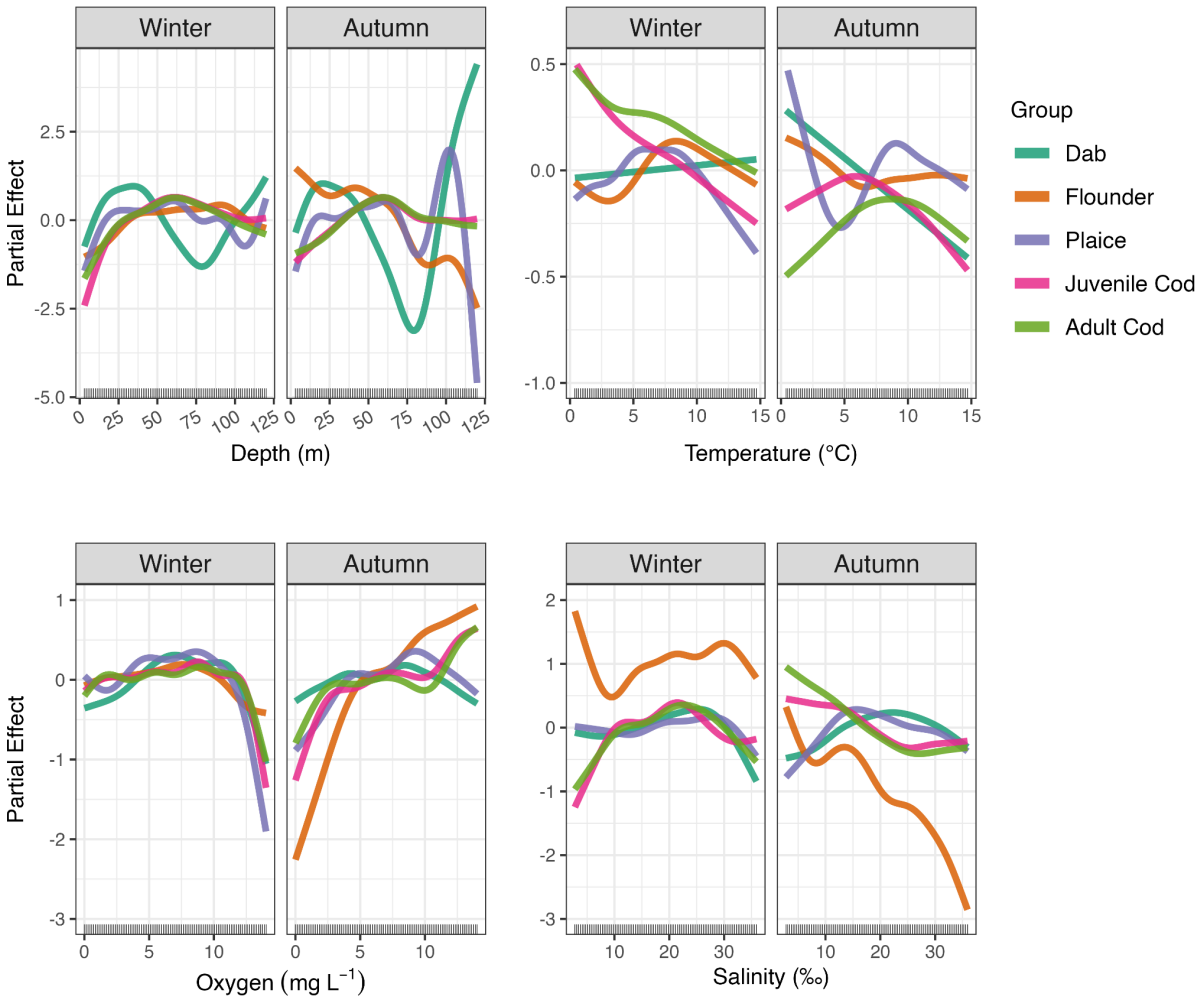


Figure 2.3 Partial dependence curves generated from HGAM version-III, including geographic and seasonal abiotic (sea bottom) predictors. For each variable, species-level random smoothers allow individually unique penalties (wiggliness) and independent shape by season. Rugs mark the observed data coverage on the x-axis for each covariate.

2.4.3. Spatiotemporal Patterns of Species Biomass

Spatiotemporal patterns in species biomass predictions were broadly species specific in our HGAMs and are shown as the total change in biomass per grid cell in Figure 2.4. Yearly species biomass predictions are shown in Figures 6.18–6.22. First, Common dab shows little discernible spatial changes in biomass, instead characterized by relative stasis and with minor biomass fluctuations confined to the western Baltic Sea. Second, European flounder increased widely across the western Baltic Sea and relative declines in the northern Baltic proper—

however we expect these apparent declines to have been influenced by unequal sampling coverage in the earliest years of the trawl survey (Figure 6.19). Third, European plaice have increased in biomass almost uniformly throughout the western Baltic Sea towards present. Fourth, both juvenile and adult cod declined in biomass in several deeper regions in the Baltic Sea: Both cod life stages declined in the Bornholm Deeps, most substantial among juveniles, and near the Gdańsk Deeps where adults showed the greatest declines in biomass. Outside these deeper areas and the northern Baltic proper, both cod life stages show some relative increases in biomass.

2.5. Discussion

This work depicts changing spatial patterns of demersal fish biomass and offers descriptions of their seasonal relationship to abiotic conditions throughout the past two decades in the Baltic Sea. Atlantic cod is arguably the most widely studied species in the Baltic Sea (e.g., Hüssy, 2011; Neuenfeldt *et al.* 2020; Svedäng *et al.* 2022). Our model results confirm and expand previously described spatiotemporal patterns, based on similar trawl data, of cod in the eastern Baltic Sea: A theme of spatial contraction and quantitative decline in juveniles and adults towards present (Casini *et al.* 2019; Orio *et al.* 2019). This work clearly maps a striking decline of juvenile and adult cod life stages from nursery areas near the Gdańsk deeps and the erosion of biomass in spawning areas of the Bornholm Deeps compared to the early 2000s— a period already following precipitous declines in cod populations in the Baltic during the 1980s-1990s (Möllmann *et al.* 2021; Eero *et al.* 2023). These spatiotemporal predictions align with related models of Eastern Baltic cod biomass declines in the 1990s, which were accompanied by widespread depreciations in cod body index and highly concentrated near the Bornholm Deeps (Lindmark *et al.* 2023).

Across flatfishes, the relative stasis of dab in space reflects average patterns where biomass is strongly filtered by the salinity gradient and restricted to Kattegat, Øresund, and parts of the Danish Straits in the western Baltic Sea (Heessen *et al.* 2015). Declines of flounder biomass in the northern Baltic proper are a more recent trend of the 2010s but exaggerated by limited coverage by BITS in the early 2000s (Figure 2.4; Figure 6.19) but compared to 1980–1990s biomass appears to have increased substantially (Orio *et al.* 2019)— caveated by potential catchability biases in different gear standardizations prior to 2001 (Orio *et al.* 2017). The similar shifts of plaice and flounder, increasing in the western Baltic Sea towards present were

clearly not limited by a similar environmental filter as dab, reinforcing important questions about the role of increased competition by demersal fishes contributing to the continued decline of cod populations (Orio *et al.* 2020). By comparison, springtime trawl data from 2004-2016 in the western Baltic Sea indicate that missing biotic variables relevant to flatfish biomass include quantitative distributions of food sources such as arthropods, annelids, and molluscs (Rau *et al.* 2019). This limit to our study is in part compensated through bottom oxygen concentrations, the availability of which is a prerequisite for the growth of demersal food sources (Villnäs *et al.* 2012). Based on the performance of our best model version incorporating seasonal abiotic covariates (HGAM version-III), cod is less reliably predicted than the other major demersals considered here; we expect that as a formerly major commercial target in the Baltic, dynamic fishing effort or nominal catch data are needed at relatively fine resolution (e.g., downscaled from global estimates Rousseau *et al.* 2024) to incorporate the negative impacts particularly on cod biomass as covariates.

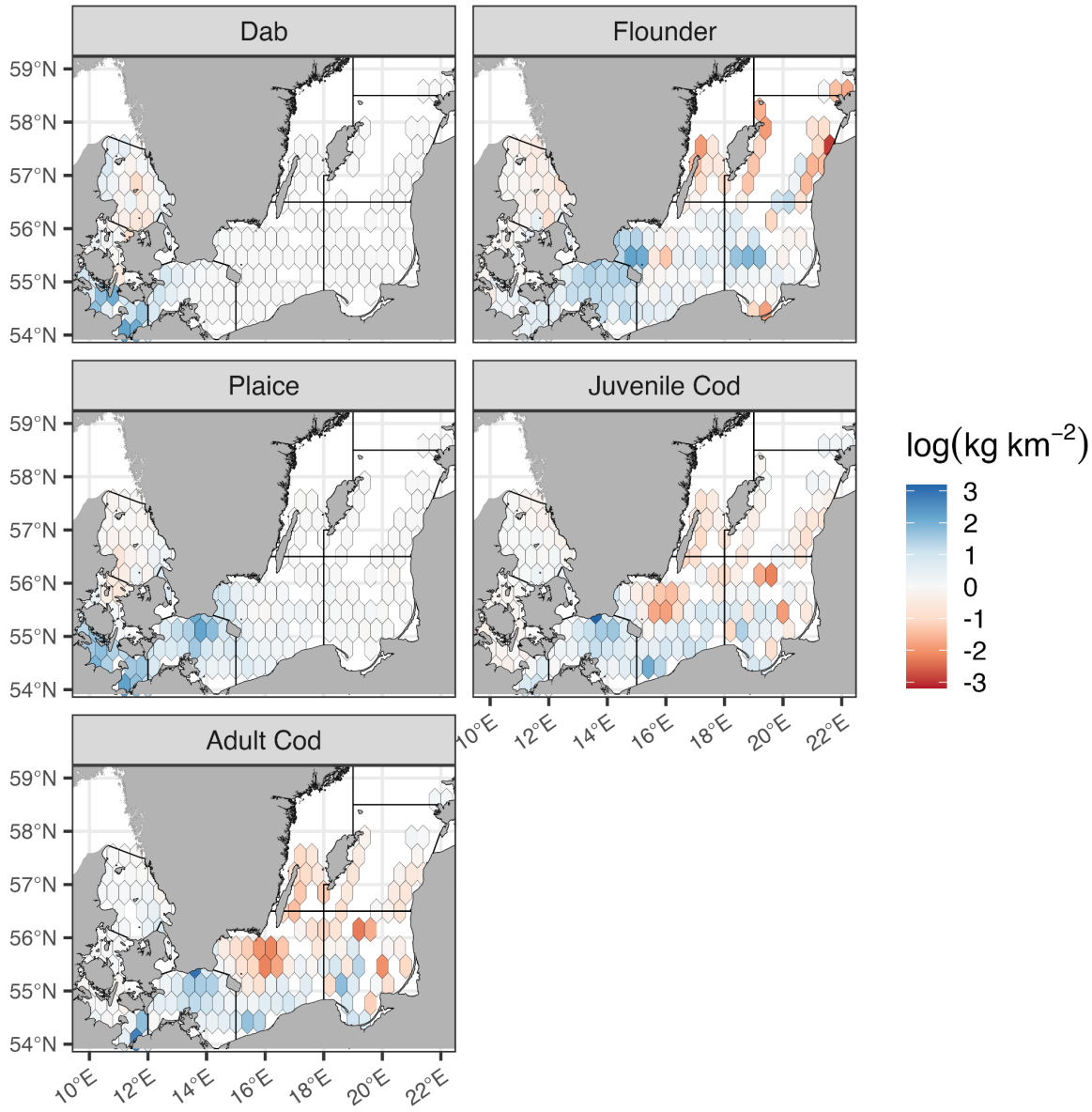


Figure 2.4 Species biomass predictions, expressed as log(kg km^{-2}), calculated as the difference in mean grid cell biomass between the last five years (2016-2020) and the first five years (2001-2005) of the trawl survey.

Driven by the strong environmental gradients in the Baltic Sea, particularly salinity, population differentiation in multiple fish species has emerged with strong geographic clustering (e.g., Johannesson *et al.* 2020). This creates relevant physiological and thus niche differences that generate intraspecific variation (Smith *et al.* 2019) and yet clear geographic divides rarely exist to split data by subpopulations (e.g., Zhang *et al.* 2021). So, despite the richness of the multi-decadal trawl survey presented here, important subpopulation factors are missing in our fish groups: At least two cod stocks (East and West) in the Baltic Sea bear different niche breadth

in their tolerances to salinity (e.g., 14-36 PSU) and temperature (Righton *et al.* 2010), and at least one example of mutation in hemoglobin proteins appears adaptive to different temperature and oxygen conditions (Andersen *et al.* 2009). In space, these niche differences can filter larvae through differences in neutral buoyancy as western Baltic cod eggs sink at lower salinities (<20 PSU, Nissling and Westin, 1997), becoming exposed to hypoxic waters <2 ml L^{-1} O_2. But a definitive geographic boundary between subpopulations is complicated by population mixing in the Arkona Basin (ICES subdivision 24) recorded independently from otoliths (Hüssy *et al.* 2016) and single nucleotide polymorphisms (SNPs; Weist *et al.* 2019). A combination of trawl data, otoliths, and genomic microsatellites deepens the degree of coexistence in the Arkona Basin, confirming population overlap in space and partitioned by depth, with western stocks occupying shallower zones and the eastern stock predominating in deeper waters (Schade *et al.* 2022). Similarly, European (*P. flesus*) and Baltic flounder (*P. solemdali*) display different adaptations to salinity, oxygen, and temperature— although our trawls rarely cover areas dominated by Baltic flounder with relatively few observations surrounding Öland and Gotland in the Baltic proper (Momigliano *et al.* 2018). We expect the plaice distributions described here to include two major stocks roughly with a western stock predominating throughout Kattegat, Øresund and the western Baltic Sea, while the second stock prevails in the Baltic proper (Ulrich *et al.* 2013). More broadly, limitations in observation data to resolve subpopulations with relevant niche differences persist across the SDM literature (Meynard *et al.* 2019). These aspects limit the generality of our HGAM descriptions and suggest improvement of species range forecasting under climate change with evolutionary markers (Aguirre-Liguori *et al.* 2021); yet these data remain sparse by comparison to species occurrences, abundances or biomass densities.

Despite these limits, we resolve seasonal differences that arguably capture important life histories. Flounder biomass responded positively at higher salinities (> 10 PSU; Feb–March), just prior to or during spawning (Orio *et al.* 2017), but this relationship flips during autumn (October–November), likely reflecting seasonal feeding periods when both European and Baltic flounder appear to use shallow, fresher coastal waters prior to late autumn/early winter spawning migrations (Nissling & Dahlman, 2010; Heessen *et al.* 2015). For dab and plaice, we build on work previously focused on the German sections of the BITS survey (e.g., Rau *et al.* 2019), generally showing that the estimated partial effects, unsurprisingly, do not hold across seasons and that species occupy different ranges of their realized niches. This is shown also in the apparent flip of the relationship between dab biomass and bottom temperatures between

winter and autumn. For cod, it appears that the biomass of juveniles and adults respond fundamentally similar to each other against temperature, oxygen, and salinity conditions but at different magnitudes. But these apparent differences can result in substantial spatiotemporal variation in larvae survival (Huwer *et al.* 2014), and local populations could exhibit distinct environmental preferences within the general patterns shown here. A key takeaway from this work is the species-specific, seasonally variable response curves. These data can be repurposed as inputs to help parametrize and constrain uncertainties in distribution forecasting under scenarios of climate change or fishing pressures (Nascimento *et al.* 2023), instead of fixed annual or long-term average response curves derived from occurrence data and suitability maps.

This work's mapping of species biomass consists of an interpolation within a well sampled range. However, model transferability (i.e., extrapolation beyond training region or timeframe), which indicates the generality of any model's representation of ecological mechanisms, remains a challenge to demonstrate for large-scale models (Heikkinen *et al.* 2012; Yates *et al.* 2018). Limited transferability in SDMs is associated with poor representation of species underlying biology and ecology (Lee-Yaw *et al.* 2022). The characteristics of transferable models or species traits that affect transferability in 129 North American bird species were explored by Rousseau and Betts (2022), revealing widespread failures in spatial transferability to predict species abundances and cautioning extrapolation beyond trained areas. We attempted to constrain model overfitting through spatial thinning, k-fold cross validation, and by focusing on core abiotic variables (e.g., temperature, salinity, oxygen) which have causal links to species distributions and achieved biomass (Wenger and Olden, 2012). Future work on transferability in SDMs will also benefit from testing scale (i.e., covariate grid size) sensitivity of predictors and response variables (Lu and Jetz, 2023). Altogether, applying quantitative data as a response variable such as our trawl biomass data, where sufficiently available, could provide greater model discrimination ability than typical occurrence data through a closer link to underlying processes (e.g., growth, mortality, competition) and thus increase transferability and biological realism (Waldock *et al.* 2022; Rousseau and Betts, 2022). Towards that aim, we observe that our HGAMs can reliably predict biomass patterns which have been individually documented in past studies, but generally underpredict biomass, a common feature to many ecological models applied across large spatial or temporal extents (Waldock *et al.* 2022; Palacio and Clark, 2023).

As a Large Marine Ecosystem (Sherman, 1991), the Baltic Sea is among the best studied marine regions, which afforded our work unique access to fisheries-independent trawling data that are consistent, standardized, and comparable across sites and years. However, better coverage of shallow (<20 m) coastal regions could improve detection and understanding of population dynamics in critical spawning and nursery areas, especially in the western Baltic Sea (Kraufvelin *et al.* 2018). Sample coverage can be expanded to new sites or understudied depths by combining scientific surveys with commercial trawl data but presents further challenges of data integration: In the western Baltic Sea, scientific trawls are comparatively brief (less than one hour) whereas commercial trawls can run up to nine hours (Rufener *et al.* 2021). Expanding and improving trawl data collection in the Baltic Sea can support more finer scale, temporally dense observations to underpin seasonal or monthly response curves for a wider variety of fish species, aiming to disentangle the local and regional outcomes which might be likely under the pressing environmental changes bearing on the Baltic Sea.

2.6. Author contributions

Liam MacNeil: Conceptualization (Lead); Data curation (Lead); Formal analysis (Lead); Writing – original draft (Lead). **Frane Madiraca:** Conceptualization (Supporting); Formal analysis (Supporting); Writing – review & editing (Supporting). **Saskia A. Otto:** Conceptualization (Supporting); Writing – review & editing (Supporting). **Marco Scotti:** Conceptualization (Supporting); Data curation (Supporting); Funding acquisition (Lead); Methodology (Supporting); Project administration (Lead); Writing – review & editing (Supporting).

2.7. References

Aguirre-Liguori, J. A., Ramírez-Barahona, S., & Gaut, B. S. (2021). The evolutionary genomics of species' responses to climate change. *Nature Ecology & Evolution*, *5*(10), 1350–1360. https://doi.org/10.1038/s41559-021-01526-9.

Akaike, H. (1974). A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, *19*(6), 716-723.

Andersen, Ø., Wetten, O. F., De Rosa, M. C., Andre, C., Carelli Alinovi, C., Colafranceschi, M., Brix, O., & Colosimo, A. (2009). Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1658), 833–841. https://doi.org/10.1098/rspb.2008.1529.

Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, *275*, 73–77. https://doi.org/10.1016/j.ecolmodel.2013.12.012.

Brander, K. (2022). Support for the hypothesis that growth of eastern Baltic cod is affected by mild hypoxia. A comment on Svedäng et al. (2022). *ICES Journal of Marine Science*, *79*(7), 2155–2156. https://doi.org/10.1093/icesjms/fsac070.

Cadotte, M. W., & Tucker, C. M. (2017). Should Environmental Filtering be Abandoned? *Trends in Ecology & Evolution*, *32*(6), 429–437. https://doi.org/10.1016/j.tree.2017.03.004.

Casini, M., Tian, H., Hansson, M., Grygiel, W., Strods, G., Statkus, R., Sepp, E., Gröhsler, T., Orio, A., & Larson, N. (2019). Spatio-temporal dynamics and behavioural ecology of a "demersal" fish population as detected using research survey pelagic trawl catches: The Eastern Baltic Sea cod (*Gadus morhua*). *ICES Journal of Marine Science*, *76*(6), 1591–1600. https://doi.org/10.1093/icesjms/fsz016.

Dallas, T. A., & Hastings, A. (2018). Habitat suitability estimated by niche models is largely unrelated to species abundance. *Global Ecology and Biogeography*, *27*(12), 1448–1456. https://doi.org/10.1111/geb.12820.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x.

Ehrlén, J., & Morris, W. F. (2015). Predicting changes in the distribution and abundance of species under environmental change. *Ecology Letters*, *18*(3), 303–314. https://doi.org/10.1111/ele.12410.

Eilola, K., Meier, H. E. M., & Almroth, E. (2009). On the dynamics of oxygen, phosphorus and cyanobacteria in the Baltic Sea; A model study. *Journal of Marine Systems*, *75*(1–2), 163–184. https://doi.org/10.1016/j.jmarsys.2008.08.009.

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159.

Erauskin-Extramiana, M., Arrizabalaga, H., Hobday, A. J., Cabré, A., Ibaibarriaga, L., Arregui, I., Murua, H., & Chust, G. (2019). Large-scale distribution of tuna species in a warming ocean. *Global Change Biology*, *25*(6), 2043–2060. https://doi.org/10.1111/gcb.14630.

Eero, M., Brander, K., Baranova, T., Krumme, U., Radtke, K., & Behrens, J. W. (2023). New insights into the recent collapse of Eastern Baltic cod from historical data on stock health. *PLOS ONE*, *18*(5), e0286247. https://doi.org/10.1371/journal.pone.0286247.

Frelat, R., Orio, A., Casini, M., Lehmann, A., Mérigot, B., Otto, S. A., Sguotti, C., & Möllmann, C. (2018). A three-dimensional view on biodiversity changes: Spatial, temporal, and functional perspectives on fish communities in the Baltic Sea. *ICES Journal of Marine Science*, *75*(7), 2463–2475. https://doi.org/10.1093/icesjms/fsy027.

Godsoe, W., Jankowski, J., Holt, R. D., & Gravel, D. (2017). Integrating Biogeography with Contemporary Niche Theory. *Trends in Ecology & Evolution*, *32*(7), 488–499. https://doi.org/10.1016/j.tree.2017.03.008.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x.

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., … Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424–1435. https://doi.org/10.1111/ele.12189.

Grinnell, J. (1917). The Niche-Relationships of the California Thrasher. *The Auk*, *34*(4), 427–433. https://doi.org/10.2307/4072271.

Haase, K., Orio, A., Pawlak, J., Pachur, M., & Casini, M. (2020). Diet of dominant demersal fish species in the Baltic Sea: Is flounder stealing benthic food from cod? *Marine Ecology Progress Series*, *645*, 159–170. https://doi.org/10.3354/meps13360.

Heikkinen, R. K., Marmion, M., & Luoto, M. (2012). Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography*, *35*(3), 276–288. https://doi.org/10.1111/j.1600-0587.2011.06999.x.

HELCOM (2021). Essential fish habitats in the Baltic Sea – Identification of potential spawning, recruitment and nursery areas. Helsinki Commission, 38 pp.

Heessen, H. J. L., Daan, N., & Ellis, J. R. (Eds.). (2015). *Fish atlas of the Celtic Sea, North Sea, and Baltic Sea: Based on international research-vessel surveys*. Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-878-0.

Hortal, J., Bello, F. de, Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M. and Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*. *46*, 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400.

Howard, C., Stephens, P. A., Pearce-Higgins, J. W., Gregory, R. D., & Willis, S. G. (2014). Improving species distribution models: The value of data on abundance. *Methods in Ecology and Evolution*, *5*(6), 506–513. https://doi.org/10.1111/2041-210X.12184.

Hüssy, K. (2011). Review of western Baltic cod (Gadus morhua) recruitment dynamics. *ICES Journal of Marine Science*, *68*(7), 1459–1471. https://doi.org/10.1093/icesjms/fsr088

Hüssy, K., Hinrichsen, H.-H., Eero, M., Mosegaard, H., Hemmer-Hansen, J., Lehmann, A., & Lundgaard, L. S. (2016). Spatio-temporal trends in stock mixing of eastern and western Baltic cod in the Arkona Basin and the implications for recruitment. *ICES Journal of Marine Science: Journal Du Conseil*, *73*(2), 293–303. https://doi.org/10.1093/icesjms/fsv227.

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposium*, *22*, 415–427.

Huwer, B., Hinrichsen, H., Böttcher, U., Voss, R., & Köster, F. (2014). Characteristics of juvenile survivors reveal spatio-temporal differences in early life stage survival of Baltic cod. Marine Ecology Progress Series, 511, 165–180. https://doi.org/10.3354/meps10875.

Holt, R. D. (2009). Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*, *106*(Supplement_2), 19659–19665. https://doi.org/10.1073/pnas.0905137106.

Hyytiäinen, K., Bauer, B., Bly Joyce, K., Ehrnsten, E., Eilola, K., Gustafsson, B. G., Meier, H. E. M., Norkko, A., Saraiva, S., Tomczak, M., and Zandersen, M. (2021). Provision of aquatic ecosystem services as a consequence of societal changes: The case of the Baltic Sea, *Population Ecology*, *63*, 61–74. https://doi.org/10.1002/1438-390X.12033.

ICES Baltic International Fish Survey Working Group (2017). Manual for the Baltic International Trawl Surveys (BITS), *Series of ICES Survey Protocols SISP 7- BITS*. 99 pp.

Jackson, J. B. C., Kirby, M. X., Berger, W. H., Bjorndal, K. A., Botsford, L. W., Bourque, B. J., Bradbury, R. H., Cooke, R., Erlandson, J., Estes, J. A., Hughes, T. P., Kidwell, S., Lange, C. B., Lenihan, H. S., Pandolfi, J. M., Peterson, C. H., Steneck, R. S., Tegner, M. J., & Warner, R. R. (2001). Historical Overfishing and the Recent Collapse of Coastal Ecosystems. *Science*, *293*(5530), 629–637. https://doi.org/10.1126/science.1059199.

Johannesson, K., & André, C. (2006). Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology*, *15*(8), 2013–2029. https://doi.org/10.1111/j.1365-294X.2006.02919.x.

Johannesson, K., Le Moan, A., Perini, S., & André, C. (2020). A Darwinian Laboratory of Multiple Contact Zones. *Trends in Ecology & Evolution*, *35*(11), 1021–1036. https://doi.org/10.1016/j.tree.2020.07.015.

King, T. W., Vynne, C., Miller, D., Fisher, S., Fitkin, S., Rohrer, J., Ransom, J. I., & Thornton, D. H. (2021). The influence of spatial and temporal scale on the relative importance of biotic vs. Abiotic factors for species distributions. *Diversity and Distributions*, *27*(2), 327–343. https://doi.org/10.1111/ddi.13182.

Kraufvelin, P., Pekcan-Hekim, Z., Bergström, U., Florin, A.-B., Lehikoinen, A., Mattila, J., Arula, T., Briekmane, L., Brown, E. J., Celmer, Z., Dainys, J., Jokinen, H., Kääriä, P., Kallasvuo, M., Lappalainen, A., Lozys, L., Möller, P., Orio, A., Rohtla, M., …

Olsson, J. (2018). Essential coastal habitats for fish in the Baltic Sea. *Estuarine, Coastal and Shelf Science*, *204*, 14–30. https://doi.org/10.1016/j.ecss.2018.02.014.

Kültz, D. (2015). Physiological mechanisms used by fish to cope with salinity stress. *Journal of Experimental Biology*, *218*(12), 1907–1914. https://doi.org/10.1242/jeb.118695.

Leppäranta, M., & Myrberg, K. (2009). Physical oceanography of the Baltic Sea. Praxis Publishing Ltd, Chichester.

Lee-Yaw, A. J., McCune, J., Pironon, S., & N. Sheth, S. (2022). Species distribution models rarely predict the biology of real populations. *Ecography*, *2022*(6). https://doi.org/10.1111/ecog.05877.

Leonardi, M., Boschin, F., Boscato, P., & Manica, A. (2022). Following the niche: The differential impact of the last glacial maximum on four European ungulates. *Communications Biology*, *5*(1), 1038. https://doi.org/10.1038/s42003-022-03993-7.

Lindmark, M., Anderson, S. C., Gogina, M., & Casini, M. (2023). Evaluating drivers of spatiotemporal variability in individual condition of a bottom-associated marine fish, Atlantic cod (*Gadus morhua*). *ICES Journal of Marine Science*, *80*(5), 1539–1550. https://doi.org/10.1093/icesjms/fsad084.

Liu, Y., Meier, H. E. M., & Eilola, K. (2017). Nutrient transports in the Baltic Sea – results from a 30-year physical–biogeochemical reanalysis. *Biogeosciences*, *14*(8), 2113–2131. https://doi.org/10.5194/bg-14-2113-2017.

Lotze, H. K., Lenihan, H. S., Bourque, B. J., Bradbury, R. H., Cooke, R. G., Kay, M. C., Kidwell, S. M., Kirby, M. X., Peterson, C. H., & Jackson, J. B. C. (2006). Depletion, Degradation, and Recovery Potential of Estuaries and Coastal Seas. *Science*, *312*(5781), 1806–1809. https://doi.org/10.1126/science.1128035.

Lu, M., & Jetz, W. (2023). Scale-sensitivity in the measurement and interpretation of environmental niches. *Trends in Ecology & Evolution*, S0169534723000071. https://doi.org/10.1016/j.tree.2023.01.003.

Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., & Possingham, H. P. (2005). Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. Ecology Letters, 8(11), 1235–1246. https://doi.org/10.1111/j.1461-0248.2005.00826.x.

Melo-Merino, S. M., Reyes-Bonilla, H., & Lira-Noriega, A. (2020). Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. *Ecological Modelling*, *415*, 108837. https://doi.org/10.1016/j.ecolmodel.2019.108837.

Meynard, C. N., Leroy, B., & Kaplan, D. M. (2019). Testing methods in species distribution modelling using virtual species: What have we learnt and what are we missing? *Ecography*, *42*(12), 2021–2036. https://doi.org/10.1111/ecog.04385.

Möllmann, C., Diekmann, R., Müller-Karulis, B., Kornilovs, G., Plikshs, M., & Axe, P. (2009). Reorganization of a large marine ecosystem due to atmospheric and anthropogenic pressure: A discontinuous regime shift in the Central Baltic Sea. *Global Change Biology*, *15*(6), 1377–1393. https://doi.org/10.1111/j.1365-2486.2008.01814.x.

Möllmann, C., Cormon, X., Funk, S., Otto, S. A., Schmidt, J. O., Schwermer, H., Sguotti, C., Voss, R., & Quaas, M. (2021). Tipping point realized in cod fishery. *Scientific Reports*, *11*(1), 14259. https://doi.org/10.1038/s41598-021-93843-z.

Momigliano, P., Denys, G. P. J., Jokinen, H., & Merilä, J. (2018). *Platichthys solemdali* sp. nov. (Actinopterygii, Pleuronectiformes): A New Flounder Species From the Baltic Sea. *Frontiers in Marine Science*, *5*, 225. https://doi.org/10.3389/fmars.2018.00225.

Nascimento, M. C., Husson, B., Guillet, L., & Pedersen, T. (2023). Modelling the spatial shifts of functional groups in the Barents Sea using a climate-driven spatial food web model. *Ecological Modelling*, *481*, 110358. https://doi.org/10.1016/j.ecolmodel.2023.110358.

Neuenfeldt, S., Bartolino, V., Orio, A., Andersen, K. H., Andersen, N. G., Niiranen, S., Bergström, U., Ustups, D., Kulatska, N., & Casini, M. (2020). Feeding and growth of Atlantic cod (*Gadus morhua* L.) in the eastern Baltic Sea under environmental change. *ICES Journal of Marine Science*, *77*(2), 624–632. https://doi.org/10.1093/icesjms/fsz224.

Nissling, A., & Westin, L. (1997). Salinity requirements for successful spawning of Baltic and Belt Sea cod and the potential for cod stock interactions in the Baltic Sea. *Marine Ecology Progress Series*, *152*, 261–271. https://doi.org/10.3354/meps152261.

Nissling, A., & Dahlman, G. (2010). Fecundity of flounder, Pleuronectes flesus, in the Baltic Sea—Reproductive strategies in two sympatric populations. Journal of Sea Research, 64(3), 190–198. https://doi.org/10.1016/j.seares.2010.02.001.

Olsson, J., Bergström, L., & Gårdmark, A. (2012). Abiotic drivers of coastal fish community change during four decades in the Baltic Sea. *ICES Journal of Marine Science*, 69(6), 961–970. https://doi.org/10.1093/icesjms/fss072.

Orio, A., Bergström, U., Casini, M., Erlandsson, M., Eschbaum, R., Hüssy, K., Lehmann, A., Ložys, L., Ustups, D., & Florin, A.-B. (2017). Characterizing and predicting the distribution of Baltic Sea flounder (Platichthys flesus) during the spawning season. Journal of Sea Research, 126, 46–55. https://doi.org/10.1016/j.seares.2017.07.002.

Orio, A., Bergström, U., Florin, A., Lehmann, A., Šics, I., & Casini, M. (2019). Spatial contraction of demersal fish populations in a large marine ecosystem. *Journal of Biogeography*, *46*(3), 633–645. https://doi.org/10.1111/jbi.13510.

Orio, A., Bergström, U., Florin, A.-B., Šics, I., & Casini, M. (2020). Long-term changes in spatial overlap between interacting cod and flounder in the Baltic Sea. *Hydrobiologia*, *847*(11), 2541–2553. https://doi.org/10.1007/s10750-020-04272-4.

Österblom, H., Hansson, S., Larsson, U., Hjerne, O., Wulff, F., Elmgren, R., & Folke, C. (2007). Human-induced Trophic Cascades and Ecological Regime Shifts in the Baltic Sea. *Ecosystems*, *10*(6), 877–889. https://doi.org/10.1007/s10021-007-9069-0.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, *20*(5), 561–576. https://doi.org/10.1111/ele.12757.

Palacio, R. D., & Clark, J. S. (2023). Incorporating intraspecific variation into species responses reveals both their resilience and vulnerability to future climate change. *Ecography*, *11*, e06769. https://doi.org/10.1111/ecog.06769.

Pecuchet, L., Törnroos, A., & Lindegren, M. (2016). Patterns and drivers of fish community assembly in a large marine ecosystem. *Marine Ecology Progress Series*, *546*, 239–248. https://doi.org/10.3354/meps11613.

Peterson, A. T., & Soberón, J. (2012). Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right. *Natureza & Conservação*, *10*(2), 102–107. https://doi.org/10.4322/natcon.2012.019.

Pigot, A. L., & Tobias, J. A. (2013). Species interactions constrain geographic range expansion over evolutionary time. *Ecology Letters*, *16*(3), 330–338. https://doi.org/10.1111/ele.12043.

Radchuk, V., Kramer-Schadt, S., & Grimm, V. (2019). Transferability of Mechanistic Ecological Models Is About Emergence. *Trends in Ecology & Evolution*, *34*(6), 487–488. https://doi.org/10.1016/j.tree.2019.01.010.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rau, A., Lewin, W.-C., Zettler, M. L., Gogina, M., & Von Dorrien, C. (2019). Abiotic and biotic drivers of flatfish abundance within distinct demersal fish assemblages in a brackish ecosystem (western Baltic Sea). *Estuarine, Coastal and Shelf Science*, *220*, 38–47. https://doi.org/10.1016/j.ecss.2019.02.035.

Reusch, T. B. H., Dierking, J., Andersson, H. C., Bonsdorff, E., Carstensen, J., Casini, M., Czajkowski, M., Hasler, B., Hinsby, K., Hyytiäinen, K., Johannesson, K., Jomaa, S., Jormalainen, V., Kuosa, H., Kurland, S., Laikre, L., MacKenzie, B. R., Margonski, P., Melzner, F., … Zandersen, M. (2018). The Baltic Sea as a time machine for the future coastal ocean. *Science Advances*, *4*(5), eaar8195. https://doi.org/10.1126/sciadv.aar8195.

Righton, D., Andersen, K., Neat, F., Thorsteinsson, V., Steingrund, P., Svedäng, H., Michalsen, K., Hinrichsen, H., Bendall, V., Neuenfeldt, S., Wright, P., Jonsson, P., Huse, G., van der Kooij, J., Mosegaard, H., Hüssy, K., & Metcalfe, J. (2010). Thermal niche of Atlantic cod *Gadus morhua*: Limits, tolerance and optima. *Marine Ecology Progress Series*, *420*, 1–13. https://doi.org/10.3354/meps08889.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. https://doi.org/10.1111/ecog.02881.

Robinson, N. M., Nelson, W. A., Costello, M. J., Sutherland, J. E., & Lundquist, C. J. (2017). A Systematic Review of Marine-Based Species Distribution Models (SDMs) with Recommendations for Best Practice. *Frontiers in Marine Science*, *4*, 421. https://doi.org/10.3389/fmars.2017.00421.

Rousseau, J. S., & Betts, M. G. (2022). Factors influencing transferability in species distribution models. *Ecography*, *2022*(7). https://doi.org/10.1111/ecog.06060.

Rousseau, Y., Blanchard, J. L., Novaglio, C., Pinnell, K. A., Tittensor, D. P., Watson, R. A., & Ye, Y. (2024). A database of mapped global fishing activity 1950–2017. Scientific Data, 11(1), 48. https://doi.org/10.1038/s41597-023-02824-6.

Rufener, M., Kristensen, K., Nielsen, J. R., & Bastardie, F. (2021). Bridging the gap between commercial fisheries and survey data to model the spatiotemporal dynamics of marine species. *Ecological Applications*, *31*(8). https://doi.org/10.1002/eap.2453.

Schade, F. M., Weist, P., Dierking, J., & Krumme, U. (2022). Living apart together: Long-term coexistence of Baltic cod stocks associated with depth-specific habitat use. *PLOS ONE*, *17*(9), e0274476. https://doi.org/10.1371/journal.pone.0274476.

Sherman, K. (1991). The Large Marine Ecosystem Concept: Research and Management Strategy for Living Marine Resources. *Ecological Applications*, *1*(4), 349–360. https://doi.org/10.2307/1941896.

Simpson, G.L. (2018). R Package: gratia. Ggplot-based graphics and other useful functions for GAMs fitted using Mgcv, 0.1-0 (Ggplot-based graphics and utility functions for working with GAMs fitted using the mgcv package).

Smith, A. B., Godsoe, W., Rodríguez-Sánchez, F., Wang, H.-H., & Warren, D. (2019). Niche Estimation Above and Below the Species Level. *Trends in Ecology & Evolution*, *34*(3), 260–273. https://doi.org/10.1016/j.tree.2018.10.012.

Smoliński, S., & Radtke, K. (2017). Spatial prediction of demersal fish diversity in the Baltic Sea: Comparison of machine learning and regression-based techniques. *ICES Journal of Marine Science*, *74*(1), 102–111. https://doi.org/10.1093/icesjms/fsw136.

Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, *10*(12), 1115–1123. https://doi.org/10.1111/j.1461-0248.2007.01107.x.

Svedäng, H., Savchuk, O., Villnäs, A., Norkko, A., Gustafsson, B. G., Wikström, S. A., & Humborg, C. (2022). Re-thinking the "ecological envelope" of Eastern Baltic cod (*Gadus morhua*): Conditions for productivity, reproduction, and feeding over time.

ICES Journal of Marine Science, *79*(3), 689–708. https://doi.org/10.1093/icesjms/fsac017.

Thuiller, W., Pollock, L. J., Gueguen, M., & Münkemüller, T. (2015). From species distributions to meta-communities. *Ecology Letters*, *18*(12), 1321–1328. https://doi.org/10.1111/ele.12526.

Ulrich, C., Boje, J., Cardinale, M., Gatti, P., LeBras, Q., Andersen, M., Hemmer-Hansen, J., Hintzen, N. T., Jacobsen, J. B., Jonsson, P., Miller, D. C. M., Nielsen, E. E., Rijnsdorp, A. D., Sköld, M., Svedäng, H., & Wennhage, H. (2013). Variability and connectivity of plaice populations from the Eastern North Sea to the Western Baltic Sea, and implications for assessment and management. *Journal of Sea Research*, *84*, 40–48. https://doi.org/10.1016/j.seares.2013.04.007.

Urban, M. C., Bocedi, G., Hendry, A. P., Crozier, L. G., Meester, L. D., Godsoe, W., Gonzalez, A., Hellmann, J. J., Holt, R. D., Huth, A., Johst, K., Krug, C. B., Leadley, P. W., Palmer, S. C. F., Pantel, J. H., Schmitz, A., Zollner, P. A., & Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. *Science*, *353*(6304), aad8466. https://doi.org/10.1126/science.aad8466.

Valentinsson, D., Ringdahl, K., Storr-Paulsen, M., Madsen, N. (2019). The Baltic Cod Trawl Fishery: The Perfect Fishery for a Successful Implementation of the Landing Obligation? In: Uhlmann, S. S., Ulrich, C., & Kennelly, S. J. (Eds.). (2019). *The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries*. Springer International Publishing. https://doi.org/10.1007/978-3-030-03308-8.

VanDerWal, J., Shoo, L. P., Johnson, C. N., & Williams, S. E. (2009). Abundance and the Environmental Niche: Environmental Suitability Estimated from Niche Models Predicts the Upper Limit of Local Abundance. *The American Naturalist*, *174*(2), 282–291. https://doi.org/10.1086/600087.

VanDeveer, S. D. (2011). Networked Baltic environmental cooperation. *Journal of Baltic Studies*. *42*, 37–55. https://doi.org/10.1080/01629778.2011.538516.

Van Rij, J., Wieling, M., Baayen, R.H., van Rijn, D. (2015). itsadug: Interpreting time series and autocorrelated data using GAMMs. R Package Version. 2.4.1.

Villnäs, A., Norkko, J., Lukkari, K., Hewitt, J., & Norkko, A. (2012). Consequences of Increasing Hypoxic Disturbance on Benthic Communities and Ecosystem Functioning. *PLOS ONE*, *7*(10), e44920. https://doi.org/10.1371/journal.pone.0044920.

Weatherall, P., Marks, K. M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J. E., Rovere, M., Chayes, D., Ferrini, V., & Wigley, R. (2015). A new digital bathymetric model of the world's oceans. *Earth and Space Science*, 2(8), 331–345. https://doi.org/10.1002/2015EA000107.

Weber, M. M., Stevens, R. D., Diniz-filho, J. A. F., & Grelle, C. E. V. (2017). Is there a correlation between abundance and environmental suitability derived from ecological

niche modelling? A meta-analysis. *Ecography*, 40(7), 817–828. https://doi.org/10.1111/ecog.02125.

Weist, P., Schade, F. M., Damerau, M., Barth, J. M. I., Dierking, J., André, C., Petereit, C., Reusch, T., Jentoft, S., Hanel, R., & Krumme, U. (2019). Assessing SNP-markers to study population mixing and ecological adaptation in Baltic cod. *PLOS ONE*, *14*(6), e0218127. https://doi.org/10.1371/journal.pone.0218127.

Wenger, S. J., & Olden, J. D. (2012). Assessing transferability of ecological models: An underappreciated aspect of statistical validation: Model transferability. *Methods in Ecology and Evolution*, *3*(2), 260–267. https://doi.org/10.1111/j.2041-210X.2011.00170.x.

Wiens, J. J. (2011). The niche, biogeography and species interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1576), 2336–2350. https://doi.org/10.1098/rstb.2011.0.

Wood, S. N. (2008). Fast Stable Direct Fitting and Smoothness Selection for Generalized Additive Models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *70*(3), 495–518. https://doi.org/10.1111/j.1467-9868.2007.00646.x.

Wood, S.N. (2017). Generalized additive models: An introduction with R. CRC Press.

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., … Sequeira, A. M. M. (2018). Outstanding Challenges in the Transferability of Ecological Models. *Trends in Ecology & Evolution*, *33*(10), 790–802. https://doi.org/10.1016/j.tree.2018.08.001.

Young, M., & Carr, M. H. (2015). Application of species distribution models to explain and predict the distribution, abundance and assemblage structure of nearshore temperate reef fishes. *Diversity and Distributions*, *21*(12), 1428–1440. https://doi.org/10.1111/ddi.12378.

Zabihi-Seissan, S., Baker, K. D., Stanley, R. R. E., Tunney, T. D., Beauchamp, B., Benoit, H. P., Brickman, D., Chabot, D., Cook, A., Deslauriers, D., Koen-Alonso, M., Lawlor, J., Le Bris, A., Mullowney, D. R. J., Roux, M., Skanes, K. R., Wang, Z., & Pedersen, E. J. (2024). Interactive effects of predation and climate on the distributions of marine shellfish in the Northwest Atlantic. *Oikos*, e10524. https://doi.org/10.1111/oik.10524.

Zhang, Z., Kass, J. M., Mammola, S., Koizumi, I., Li, X., Tanaka, K., Ikeda, K., Suzuki, T., Yokota, M., & Usio, N. (2021). Lineage-level distribution models lead to more realistic climate change predictions for a threatened crayfish. *Diversity and Distributions*, *27*(4), 684–695. https://doi.org/10.1111/ddi.13225.

Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., … Merow, C. (2020). A standard protocol for reporting species

distribution models. *Ecography*, *43*(9), 1261–1277. https://doi.org/10.1111/ecog.04960.

3. Network-based bioregionalization of demersal fish in continental shelf seas

This chapter is based on a paper which is in review.

MacNeil, L., Scotti, M. Network-based bioregionalization of demersal fish in continental shelf seas. *Authorea. 1st* review in *Ecography*.

3.1. Abstract

Biogeographical partitioning of ecological communities has been renewed in recent decades to illustrate broad distributional patterns. In the oceans, observational datasets have grown substantially and open new access to test bioregional patterns beyond the classical fixed thresholds of endemism to differentiate regions. This work combines a recently collated dataset of 29 different scientific bottom trawl surveys spanning 21 years with network-based clustering to illustrate biogeographical partitions of vast tracts of the northern hemisphere's continental shelf seas. Our work contributes to testing bioregionalization patterns in demersal fishes using observational data, totaling >2.5 million species records and >2000 species, with bipartite network clustering weighted by species occurrence frequencies. We propose eight major bioregions across shelf seas which fall along the longest geographical axis in each shelf region and against continua of species richness gradients, endemicity, and phylogenetic turnover rates. These patterns capture known biogeographical boundaries (e.g., North Sea–Baltic Sea, Cape Hatteras) alongside potential transition areas deduced from uncertainty estimates based on shared network nodes between bioregions. The most species-rich areas include the Southeast US Shelf, Temperate Pacific, Northeast Atlantic Shelf, and the Outer European Shelf—corresponding to relatively high endemicity. However, the relatively species-poor partitions including the Baltic Sea and the North & Celtic Seas display comparatively low endemicity (<8%), illustrating apparent statistical differences in partitions captured by bipartite networks and occurrence frequencies that would otherwise be missed using a fixed endemic criterion. Our proposed bioregionalization can be compared against the growing availability of species occurrence data, dispersal limitations, or other quantitative observations of ecological communities.

3.2. Introduction

Seeking large-scale biogeographic patterns is centuries old (Zimmerman 1777; von Humboldt 1806). Modern ecological datasets combining taxonomic and phylogenetic information have renewed efforts in recent years to classify distinct biological communities spread across ecosystems (Ebach & Parenti 2015; Morrone 2018). These methods have redrawn Alfred Russel Wallace's 19th–century zoogeographic regions (Holt *et al.* 2013), Australia's floral geography (Burbidge 1960; Ebach *et al.* 2015), and detected biogeographic rearrangements driven by human impacts (e.g., Rueda *et al.* 2024). In dynamic marine environments, identifying natural geographical entities of taxonomic groups or communities is of a different order in difficulty: Marine ecosystems are less intensely studied than on land (Hughes *et al.* 2021) and constitute vast semi-permeable spaces connected through migration and dispersal (Cowen *et al.* 2007; Álvarez-Noriega *et al.* 2020).

Global biogeographical partitions of the oceans have been developed based on oceanographic boundary currents and surface primary production regimes (Longhurst 2010). For coastal and shelf zones, broad-scale delimitations are made using oceanographic and bathymetric features of areas >200,000 km^2 (Large Marine Ecosystems; Sherman 1991) and combinations of community composition assessments with expert opinion (Ecoregions; Spalding *et al.* 2007). Continental shelf marine ecosystems are often distinct in environmental profiles from the greater open ocean with the prominent influence of prevailing boundary currents and primary productivity (Reygondeau *et al.* 2013). By comparison, progress differentiating major biogeographical partitions based on observational ecological datasets remains underdeveloped.

The growth of species occurrence data has broadened our power to describe biogeography of numerous groups and test the explanatory strength of ecological or evolutionary hypotheses in marine communities. For example, it has helped to identify relationships between temperature and diversity (Tittensor *et al.* 2010), biomass (Van Denderen *et al.* 2023), or the direction of trophic control (Frank *et al.* 2007). On regional scales, benthic invertebrate communities divide strongly in the northwest Atlantic shelf based on oceanographic connectivity, dispersal limitations, and species thermal tolerances (Pappalardo *et al.* 2015). Globally, brittle stars show roughly a dozen major biogeographic regions with strong transition zones (Victorero *et al.* 2023). The degree of endemicity has been a primary criterion to define bioregionalizations into a hierarchical system of realms, regions, and provinces based on a fixed threshold (10%) of

endemic communities, genera, and species, respectively (Costello *et al.* 2017; Morrone 2018). Endemicity links contemporary patterns to the potential evolutionary innovativeness of a biogeographical partition, dispersal limitations, and ecological niche constraints (Bowen *et al.* 2013; Cowman *et al.* 2017; Victorero *et al.* 2023). Yet highly mobile fish species range widely across continental shelves, occupying the richest fishing grounds on Earth (Rousseau *et al.* 2024), and creating substantial variability within shelf seas where taxa frequently overlap.

If distinct bioregions exist in fish communities across continental shelf seas, identification would require extensive, standardized, and repeated data collection with which to test the robustness of biogeographical partitions. To this challenge, the modern growth of scientific bottom trawls has now routinely sampled substantial tracts of continental shelf seas in the northern hemisphere (Maureaud *et al.* 2021). Despite these advances, biogeographical partitions based on abundant fish communities in shelf seas have relied upon a fixed species endemism to define provinces shallower than 200 m depth (Briggs 1974; Toonen *et al.* 2016), and data-driven partitions based on the growing scientific trawl data between regions have not been proposed.

Among the methods used to delimit biogeographical zones, hierarchical clustering of species turnover rates and network-based community detection are widely popular due to their computational efficiency and accuracy in reconstructing known regional divides (Kreft & Jetz 2010; Bloomfield *et al.* 2018). Network-based methods have shown strength in detecting more granular biogeographical features than hierarchical clustering, especially in transitional biomes (e.g., Bloomfield *et al.* 2018), including reproducing the Wallace's zoogeographic realms and continental-wide plants biomes in the United States (Vilhena & Antonelli 2015). A bipartite network (site × species) connecting species and site nodes is a suitable tool to represent biogeographical distributions, as it can be divided into distinct communities based on the arrangement (i.e., topology) and weight of the links (Dormann *et al.* 2009).

In this work, we focus on biogeographical partitioning— bioregionalization— across demersal fish communities throughout continental shelf seas in the northern hemisphere using a weighted bipartite network. This approach uses a fast network-based clustering of scientific bottom trawl data spanning 21 years, producing stable bioregions at large spatial scales. We also calculate the strength of each site's connection to its respective bioregion to identify possible transition zones in the network structure and compare bioregionalizations to rates of

turnover in phylogenetic lineages. Substantial differences are observed in demersal fish communities across their longest geographical axis: Bioregions generally divide by latitude in the North American coastal shelves and by longitude in Europe. These patterns appear in part explained by differing species endemicity, phylogenetic turnover, and species occurrence frequency. These proposed bioregions match some pre-existing marine Ecoregions and biogeographic provinces, but with exceptions that illustrate the wide-ranging pools of demersal fish communities. Together these findings depict a broad biogeographical partitioning across continental shelf ecosystems amidst the torrent of ecological and environmental variability in the coastal oceans.

3.3. Methods

3.3.1. Observational data

We compiled scientific bottom trawl data using 29 different surveys from FISHGLOB (Maureaud *et al.* 2024; Table 8.1), spanning large tracts of the continental shelf seas in the northern hemisphere (Figure 3.1). We consider trawls between 1999–2020, with the majority of surveys (17/29) having already begun in 1999 (Figure 8.1). All data preparation and subsequent analyses were done in the R programming language (R Core Team 2023). The FISHGLOB data are pretreated to harmonize and validate taxonomic names against the World Register of Marine Species (WoRMs), plus all observations are matched with trawl-level information and units are standardized (Maureaud *et al.* 2024). We add a spatiotemporal filter which removes trawls below 2% coverage in a 23,320 km^2 grid cell (size 7; Maureaud *et al.* 2024), this trimmed 4% of observations. The final dataset contains >2 million individual observations, 138,240 trawls and 2090 species across 21 years.

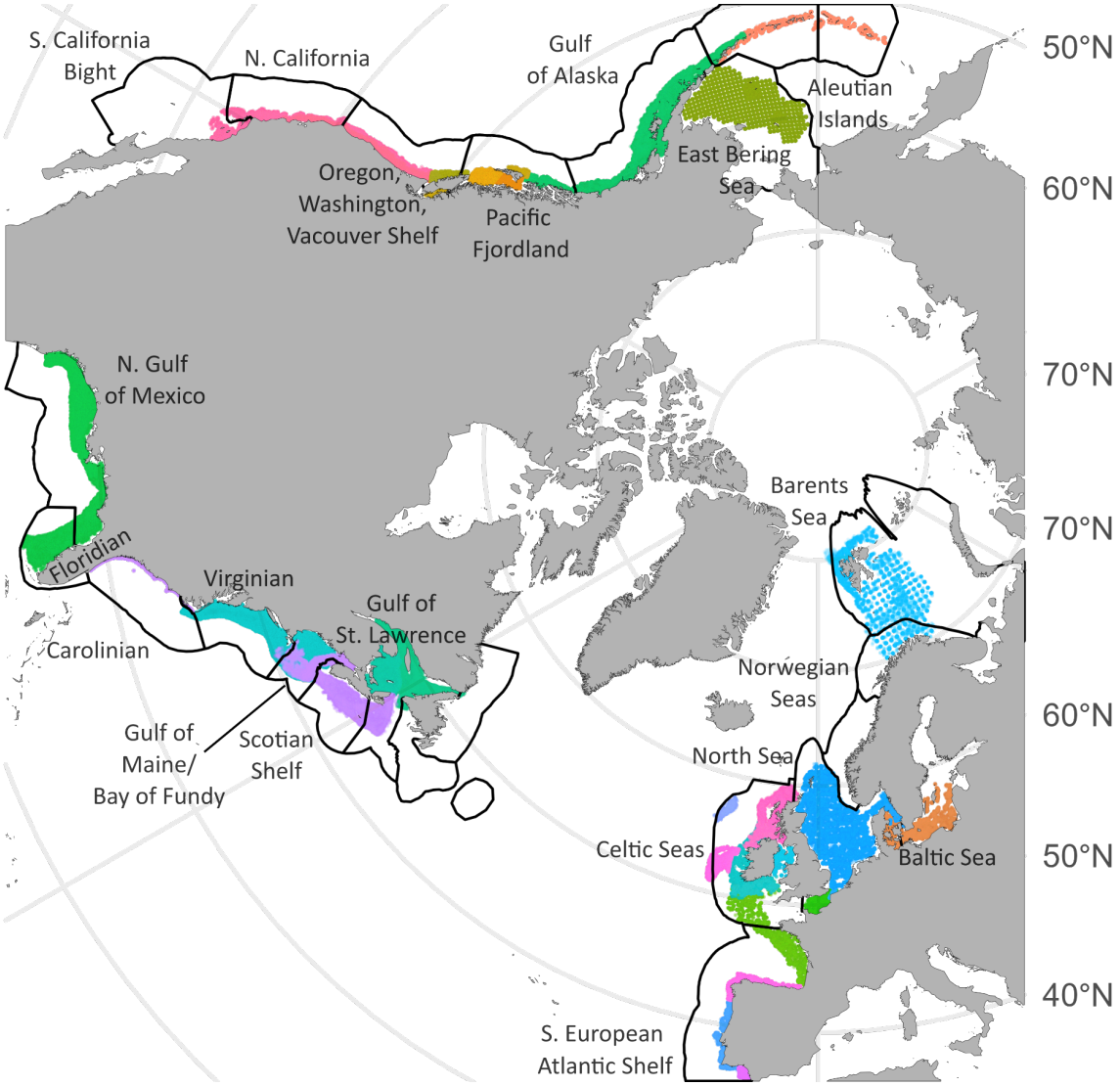


Figure 3.1 Scientific bottom trawl data spanning 21 years (1999–2020) colored by survey identity underneath marine Ecoregions (Spalding et al. 2007). See Table 8.1 and Figure 8.1 for survey names, acronyms, geographic designations and yearly trawl frequencies.

All observations were prepared for bioregionalization by gridding trawls into 100 km (10,000 km^2) and 200 km (40,000 km^2) hexagonal grids in the *sf* package (Pebesma 2018) to test for spatial sensitivity in our results (e.g., [a]Daru *et al.* 2020). Each grid cell was given a grid identification (ID) for further analysis. At each spatial scale, we rarefied trawl data down to the first quartile ($n_{1stQ,\ 100\ km} = 92$; $n_{1stQ,\ 200\ km} = 124$) and median ($n_{median,\ 100\ km} = 190$; $n_{median,\ 200\ km} = 337$) number of trawls per grid cell to equalize sampling efforts. We additionally test bioregionalization without rarefaction using unweighted approaches described below.

3.3.2. Network-based bioregionalization

Bioregionalization was tested using weighted bipartite (site × species) networks (Beckett, 2016). We constructed networks in the *igraph* package (Csardi & Nepusz 2006), creating edge lists from the species list and associated grid IDs. Edge lists were converted to undirected graphs and weights were generated based on species occurrence frequency in each grid cell and added to graph edges. We also tested bioregionalization using unweighted, undirected networks without grid cell rarefaction as described above. To identify modules (i.e., bioregions) in bipartite networks, we partitioned underlying community structure to define highly connected areas with shared species (i.e., endemics) using the Leiden clustering algorithm (Traag *et al.* 2019)— an extension of the popular Louvain algorithm (Blondel *et al.* 2008), using iterative node partitioning, refinement and aggregation while optimizing for network modularity (Newman 2006). The resolution parameter was set equal to two, defining a minimum density threshold for bioregions, however we search across a range of resolution values to show the resulting variability in bioregion number (Figure 8.2). We chose the resolution parameter because it was the minimum threshold to reproduce the North Sea-Baltic Sea divide where the salinity gradient powerfully filters the fish community (Pecuchet *et al.* 2016)— a larger resolution parameter threshold produced a patchwork of spatially fragmented bioregions. The algorithm was run across 100 iterations.

To test the sensitivity of our proposed bioregions to algorithm choice, we also applied the infomap algorithm with 100 iterations based on the map equation using random walks of connections between nodes (Rosvall & Bergstrom 2008). We directly compare algorithms using the *robin* package (Policastro *et al.* 2021) to assess partition stability under a degree preserving randomization of network structure, increasingly randomizing network edges by 5% until 50% perturbation is reached; we summarize these perturbations using a normalized variation of information (VI) indicator of stability (Meilă 2007). An interval testing procedure (Pini & Vantini 2016) was used to compare and calculate p-values of VI between algorithms at sequential levels of network perturbation (Figure 8.3). More stable network partitions (i.e., bioregions) should be more robust to network perturbations. We find that although the infomap-derived bioregions produced similar network modularity scores as the Leiden algorithm (Table 8.2), it displayed greater spatial fragmentation and instability (Figure 8.3), thus these maps are shown in the supplement (Figures 8.4–8.6). To test if the observed bioregion patterns cluster together with greater density than could be observed randomly, the

achieved modularity scores of the Leiden algorithm were compared against those of three null models in the *bipartite* package (Dormann *et al.* 2008) using shuffle, swap, and vaznull algorithms. Each null model differs in the imposed constraints during randomization. The shuffle algorithm randomizes the bipartite networks but maintains the number of species associated to each site (connectance) and network topology, the swapping algorithm randomizes species frequencies but maintains the overall values at each site and preserves connectance, while the vaznull algorithm is similarly constrained as the swapping algorithm but less strictly preserves connectance by assigning species-specific probabilities to frequencies (Vázquez *et al.* 2007); each null model was run for 999 iterations (Figure 8.7).

To identify areas within bioregions which share a high degree of species composition with another bioregion and are thus uncertain in partition membership or represent potential ecological transition zones (Morrone 2024), we calculate a participation coefficient (PC; Guimerà & Amaral 2005) for every i^{th} site node (i.e., grid cell):

$$PC_i = 1 - \sum_{s=1}^{N_m} \left(\frac{k_{is}}{k_i}\right) \quad (1)$$

Where N_m is the number of bioregions detected, s is the bioregion identity, k_{is} indicates the number of links for node i in bioregion s, and k_i denotes the total number of links for node i (i.e., total degree). The PC represents the diversity of shared species between bioregions and broadly describes the strength of a nodes' links to its bioregion. High PC values indicate many shared species with another bioregion while low PC values imply highly unique species sets. The density of shared species links between different bioregions are visualizable in the bipartite network (Figure 3.2). Across all our defined bioregions we characterized total area size (km^2; Figure 8.8), demersal fish species richness (Figure 8.9), endemicity, the average PC value, depth ranges, and the number of intersecting Ecoregions (Spalding *et al.* 2007) to reappraise their usefulness as ecological boundaries for highly mobile fish communities. We also broadly compare bioregionalization to phylogenetic turnover rates (Figure 8.10) by placing species names into a phylogeny of ray-finned fishes (Rabosky *et al.* 2018) and calculating turnover rates in phylogenetic lineages using the *epm* package (Title *et al.* 2022) across 200 km grid cells tested at 250 and 500 km focal radius searches for multi-site dissimilarities (Baselga, 2013).

3.4. Results

3.4.1. Demersal fish bioregions

In total, our bioregionalization indicates eight distinct bioregions. Across the continental shelves, latitudinal gradients prevail on the North American continent. On the Pacific coast, bioregions divide into a Temperate Pacific region— incorporating the California Current region from the Southern California Bight to the Pacific Northwest (Oregon, Washington, and Vancouver Shelf)— and a Northern Pacific region ranging the Gulf of Alaska coast, the Aleutian Islands, and the East Bering Sea (Figure 3.2). On the Atlantic coast, the Gulf of Mexico and southeastern US shelves (Floridian, Carolinian) cluster together in a large range south of Cape Hatteras (Southeast US Shelf). The northeast US and Canadian shelves cluster together north of Cape Hatteras across the Virginian coast, the Gulf of Maine, the Bay of Fundy, and the majority of the Scotian Shelf (Northeast Atlantic Shelf). The more northern bioregions divide at the eastern Scotian Shelf near the Laurentian Channel and the Gulf of St. Lawrence, which clusters with the Norwegian and Barents Seas (Sub-Arctic Atlantic). On the European shelves, bioregion partitions appear most strongly related to longitude, dividing between outer regions exposed to the Atlantic (Outer European Shelf), the North and Celtic Seas, and the Baltic Sea.

The bioregions delineated from weighted bipartite networks were highly stable across grid cell resolutions and rarefaction depths. Network modularity was similar using Leiden and infomap algorithms but the fomer was faster and produced bioregions that were less spatially fragmented and more stable (Figures 8.3–8.6)— the infomap converged closer to the Leiden proposed bioregions with weighted links (Figure 8.5); The unweighted bioregionalizations across both algorithms were unable to reproduce the Baltic Sea–North Sea divide and had higher uncertainty (i.e., higher PC values) in bioregion identity (Figures 8.4, 8.6). We show our proposed bioregionalization using 200 km grid cells and the less restrictive rarefaction depth ($n_{1stQ,\ 200\ km} = 124$) to clearly visualize bioregion patterns as the more restrictive approach ($n_{median,\ 200\ km} = 337$) presented large gaps within shelves and excluded the Norwegian and Barents Seas data (Figure 3.2).

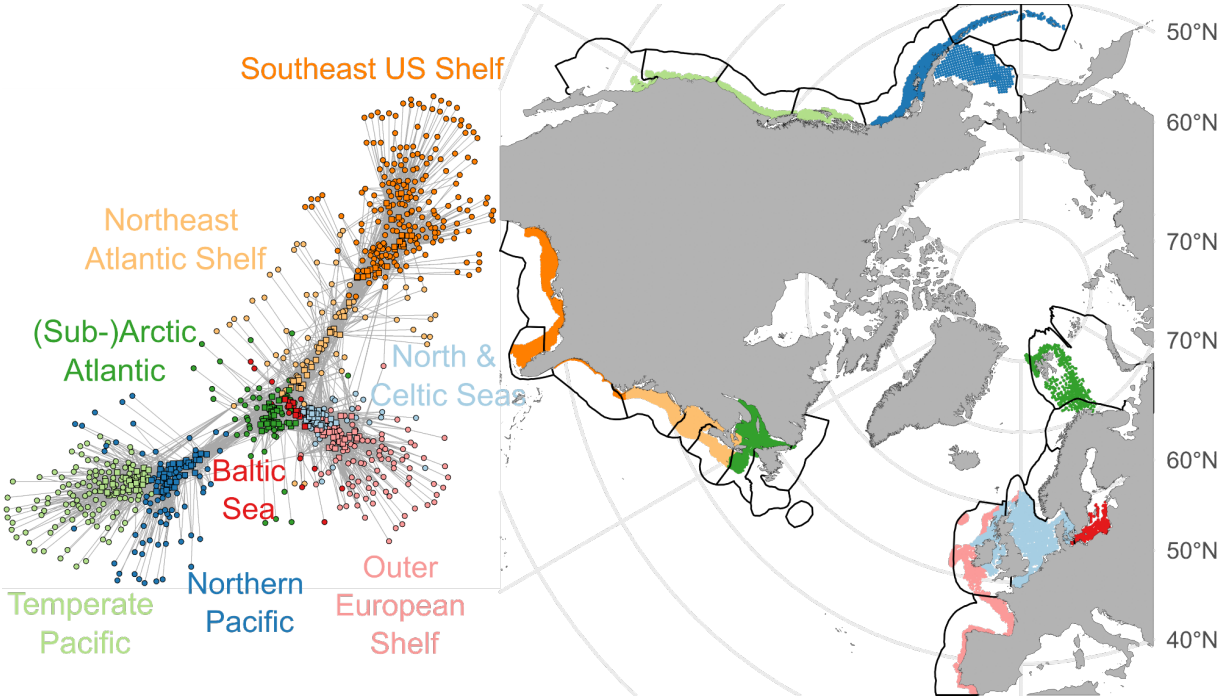


Figure 3.2 The networked bioregions of demersal fish across the continental shelf seas in the Northern Hemisphere. The bipartite network structure illustrates the distribution of sites × species links (squares × circles) colored by defined bioregions, with the map displaying their geographical distribution alongside Ecoregions (Spalding et al. 2007). These results display the outcomes at a 200 km grid cell resolution rarefied less restrictively (1st quartile), however the bioregions appear across different resolutions and rarefaction depths.

The largest bioregions were found in the Arctic and Sub-Arctic (Northern Pacific and Sub-Arctic Atlantic) and on the longitudinally broad North and Celtic Seas (Table 3.1). As a marginal sea on the European Shelf, the Baltic Sea bioregion was less than half the size of the relatively small Northeast Atlantic Shelf. The dominant latitudinal and longitudinal axes of differentiation in North American and European shelves, respectively, tracked differences in species richness and declining species endemicity (Table 3.1). In North America, species richness declined at higher latitudes and endemics became rarer, whereas this pattern in European seas was observed moving eastward away from Atlantic shelf margins to the Celtic Seas, North Sea, and the Baltic Sea. Bioregions cover wide depth ranges including shallow coastal environments and beyond the continental shelf breaks >1000 m depth (Table 3.1). Absolute differences in depth ranges across bioregions are partly an artifact of survey extent, but important relative differences are observable between bioregions that span continental shelf breaks (Temperate Pacific, Northern Pacific, Northeast Atlantic Shelf, (Sub-)Arctic Atlantic,

Outer European Shelf) and those located within continental shelves or marginal seas (Southeast US Shelf, North & Celtic Seas, Baltic Sea).

Table 3.1 Overview of proposed bioregions, including area, species richness, endemics, total depth range, average participation coefficient (PC), and the number of intersected Ecoregions (Spalding et al. 2007).

Bioregion	Area (km^2)	Species Richness	Endemic Species	Endemicity	Depth Range (m)	Mean PC	Eco-regions
Temperate Pacific	969,948	293	142	48%	18–1428	0.30	4
Northern Pacific	1,558,845	205	50	24%	11–984	0.40	3
Southeast US Shelf	1,004,589	478	296	62%	2–500	0.14	4
Northeast Atlantic Shelf	796,743	284	36	13%	7–1475	0.22	3
(Sub-)Arctic Atlantic	1,212,453	133	11	8%	24–1357	0.37	4
Outer European Shelf	953,307	229	78	34%	10–787	0.44	3
North & Celtic Seas	1,350,999	129	9	7%	3–450	0.39	4
Baltic Sea	346,410	54	2	3%	3–146	0.31	2

3.4.2. Spatial gradients and potential transition areas

Faunal links between bioregions show a high degree of connectedness as neighboring bioregions share the highest number of species (Figure 3.3). The strongest connections are within ocean basins; however, a small number of species are distributed widely between regional shelves (Figure 3.3). The Northeast Atlantic Shelf contains the highest number of

species overlap, primarily with the species-rich Southeast Atlantic Shelf, but also with the European shelves. The European seas are highly interconnected, but the relatively species-poor Baltic Sea is reflected by its low frequency of shared species.

All global bioregions were replicable but the fidelity of individual sites (grid cells) to its assigned bioregion varied (Figure 3.4). The broad patterns in PC values— the degree of species overlap with another bioregion— and thereby variability in bioregion identity increased with latitude on both North American shelves and was mixed across the European shelves with highest values in the North & Celtic Seas and Outer European Shelf bioregions (Table 3.1). Grid-cell level patterns in PC values differed between regions and illustrates finer-scale evidence for demersal fish transition zones. On the Pacific Shelf, the highest PC values (>0.6) were found in the Aleutian Islands and the East Bering Sea, implying a potential divide between the Gulf of Alaska and the Aleutian Islands fish communities. The areas in the Pacific Northwest region (Oregon, Washington, Vancouver Coast and Shelf; Pacific Fjordland) contain individual grid cells with relatively high PC values (>0.6) and are otherwise surrounded by relatively low PC value grid cells (<0.2) which cluster strongly with this proposed Temperate Pacific bioregion. On the Atlantic side of the North American continent, the highest PC values (>0.6) indicating a potential transition area appear at more northern latitudes on the Scotian Shelf near the divide of the Northeast Atlantic Shelf and the Gulf of St. Lawrence which contains the (Sub-)Arctic Atlantic proposed bioregion. Intermediate PC values (0.4–0.5) are found around Cape Hatteras at the divide between our Southeast US Shelf and the Northeast Atlantic Shelf bioregions. Overall, the greatest magnitude and extent of high PC values (>0.6) are found on the European shelves where the outer shelf areas and the southern/eastern North Sea display the highest overlap in species composition.

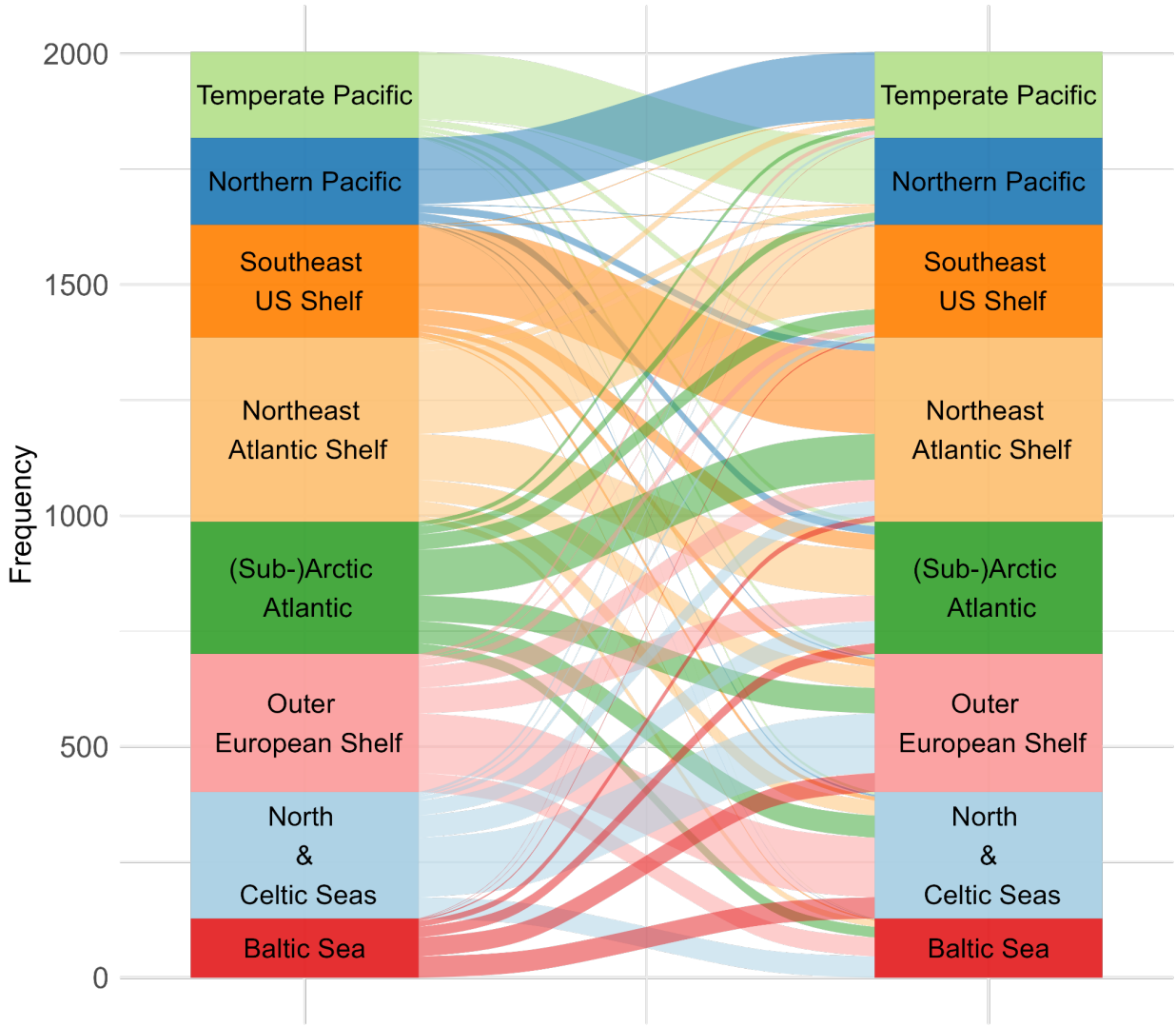


Figure 3.3 Alluvial plot depicting the frequency of shared species across bioregions. Line thickness relates to cumulative shared species, showing widespread connections between bioregions but with nearby groups sharing the largest number of species.

From this variation, these biogeographical partitions of demersal fish communities appear to fall along shelf-specific continua of species richness and endemism (Table 3.1, Figure 8.9). Comparing our proposed bioregionalization to spatial turnover rates of phylogenetic lineages shows a patchwork of differences between continental shelves (Figure 8.10). The primary zones of lineage turnover are typically found within bioregions and not exactly aligned with zones of high PC values: namely in the Pacific Northwest (Oregon, Washington, and Vancouver Shelf), the eastern Gulf of Mexico, between the Scotian Shelf and Gulf of St. Lawrence, the Bay of Biscay, and the Baltic Sea (Figure 8.9).

3.5. Discussion

Demersal fish communities separate into bioregions across the continental shelf seas in the northern hemisphere primarily along their longest continuous geographical axis: Bioregions divide by latitude in the North American coastal shelves and by longitude in Europe (Figure 3.2). The exception to this finding is in the Norwegian/Barents Seas, where missing data in the Norwegian Seas due to quality assurance and quality control flags (Maureaud *et al.* 2024) are potentially contributing to its broad clustering with the Subarctic Gulf of St. Lawrence. A first-order description of the general partitioning is geographic: that dissimilarity increases with distance (Tobler 1970). This geographic simplification is powerful, but the underlying drivers to these patterns are likely multifaceted. The evolutionary origin location of species, geographic barriers, dispersal limitations, and ecological niche dynamics all shape biogeographic patterns (Wiens & Donoghue 2004; Briggs & Bowen 2013). Similar compilations of trawl data as are analyzed here illustrate that demersal biomass increases towards colder temperatures along the latitudinal gradient of the northern hemispheres shelf seas (Van Denderen *et al.* 2023); combining trawl data with assumed trait values revealed spatial gradients in life-history categories— primarily on the Atlantic shelves— where a fast-slow continuum related to temperature dominated the structure of fish communities (Beukhof *et al.* 2019). Regionalizations based on species range maps and larval dispersal models are strikingly similar to our biogeographical partitions (Legrand *et al.* 2024), but its most notable discrepancy is in grouping the US shelf around Cape Hatteras, where our observed trawl data reflects known biogeographic junctures with greater differences between communities than what larval dispersal limitations alone suggest.

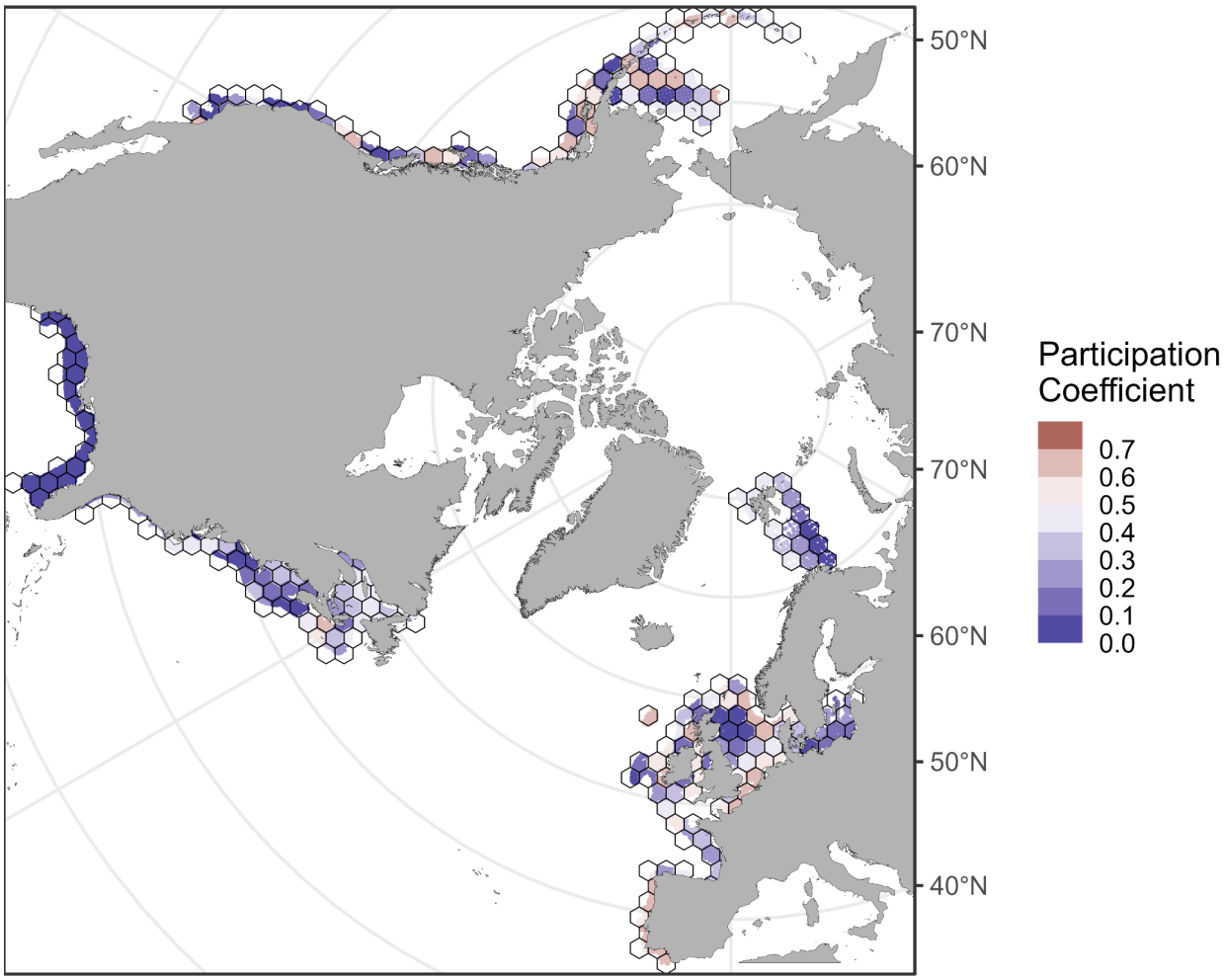


Figure 3.4 The participation coefficient (PC; equation 1) values coloring the points within grid cells (200 km); high PC values indicate areas containing high species overlap with another bioregion, low PC values represent areas with unique species sets rarely contained in another bioregion.

From our analyses, an apparent mixture of species richness gradients, endemism, differing occurrence frequencies, and phylogenetic turnover characterize spatial partitions. The gradient of declining species richness and endemism across the respective geographical axes of each continent are the primary coincidental factors with biogeographic partitions. However, by testing weighted and unweighted bipartite networks across two clustering algorithms we show how including demersal fish occurrence frequencies contribute to identifying the Baltic Sea-North Sea and Temperate-Northern Pacific divides at different spatial resolutions and reducing bioregion uncertainty shown in PC values. Turnover of phylogenetic lineages (Figure 8.10–8.11) appears secondarily related to bioregionalization and we speculate is strongly affected by local environmental gradients, with high lineage turnover in the Baltic Sea where there is strong salinity filtering (Peuchet *et al.* 2016); in the eastern Gulf of Mexico along the Mississippi delta

region between the Mississippi and the Texas estuarine coastal environments with strongly plume-driven seasonal salinity fronts (Mendelssohn *et al.* 2017; Justić *et al.* 2022); and in the Pacific Northwest (Oregon, Washington, and Vancouver Shelf) south of Vancouver Island and the Juan de Fuca Strait at the northern limit of the California Current system (Juan-Jordá *et al.* 2009). But altogether it is only in the Baltic Sea that relatively high lineage turnover rates correspond to a bioregion partition. We note however that our calculation of phylogenetic turnover is limited as only 75% (859/1144) of species in the trawl record matched phylogenetic branch tips from Rabosky *et al.* (2018) despite taxonomic harmonization in Maureaud *et al.* (2024).

Taking the example of the North American shelves in the Atlantic Ocean, our proposed bioregions based on bipartite networks display a broad similarity in demersal fish communities north of Cape Hatteras and across a significant portion of the Scotian Shelf despite important local differences in community composition (Gabriel 1992). These results demonstrate the wide-ranging distributions of demersal fishes which, in terms of spatial scale, are broadly nested between the province and Ecoregion classification by Spalding *et al.* (2007). Each bioregion intersects multiple Ecoregions, but it is only the Southeast US Shelf bioregion which corresponds to an ecological realm: Warm Temperate Northwest Atlantic (Spalding *et al.* 2007). The combination of the northern Gulf of Mexico with the Floridian and Carolinian Ecoregions is not entirely surprising given their similarity in ocean temperature and links to the Gulf Stream through the Loop Current (Richardson 2001). The near-shore limit of trawl samples also likely influenced the grouping of the Southeast US coast south of Cape Hatteras—itself a benchmark biogeographic boundary located where the Gulf Stream is largely driven offshore and related to shifts in ecological communities (Briggs 1974). Viewed against vast ecological realms (Spalding *et al.* 2007), realms of species endemicity across kingdoms (Costello *et al.* 2017), and fish-specific provinces of endemicity (Toonen *et al.* 2016), we illustrate underlying biogeographical structure to demersal fish communities across shelf seas.

Identifying bioregions or distinct ecological zonations where they exist in the oceans depends on the relevant level of ecological hierarchy: Either by integrating broad spatial patterns from groups across food webs (Costello *et al.* 2017) or by targeting specific communities (e.g., reef fish; Kulbicki *et al.* 2013). We have taken the latter approach using the modern growth of fisheries-independent survey data to describe the biogeography of bottom-associated fishes. Statistical methods available for defining bioregions are increasingly diverse, including range

limits and hierarchical clustering of taxonomies (Kreft & Jetz 2010) and phylogenies ([b]Daru *et al.* 2020), or clustering of network structures (Bloomfield *et al.* 2018) as we demonstrate in this work. It is also possible to use multi-stage modeling to hierarchically link species composition and representations of environment for probabilistic outputs of bioregions (e.g., Hill *et al.* 2020); however, this approach demands large computational budgets for big datasets unlike our network-based approach which executes in seconds, has numerous available algorithms (Csardi & Nepusz 2006) and has relatively straightforward uncertainty estimates like participation coefficients (Guimerà & Amaral 2005).

Access to a vast number of bottom trawl records with user-friendly data cleaning and quality assurance steps has quickly opened new avenues for analyzing biogeographical patterns. From biogeography, questions can be drawn about ecology and evolution, including the role of environmental niches to explain species sorting across variable environments (Ricklefs & Jenkins 2011) and whether marine provinces of high endemicity are origin hotspots for speciation or the persistence of ancient lineages (Briggs 2003; Toonen *et al.* 2016). The conventional 10% species endemism definition of biogeographical provinces of marine fishes omits, for example, the Baltic Sea (Briggs & Bowen 2012), where the salinity gradient is known to drive genetic divergence across taxa (Johannesson et al. 2020) and speciation through genetic isolation in the case of Baltic Flounder (*Platichthys solemdali* sp.; Momigliano *et al.* 2017, 2018). Together these questions are trying to explain the processes of community assembly across space (Leibold *et al.* 2022). However, we caution that sufficient care be taken in treating spatial or temporal autocorrelation before attributing causal processes to the menagerie of biogeographical patterns in the natural world (Warren *et al.* 2014).

Bioregions more specifically can be useful boundaries to test against temporally dense observations from tracking tags, for example, by comparing spatiotemporal distributions including depth distributions that can be related to regional mesoscale phenomena influencing behaviour (Arostegui *et al.* 2022). It has been suggested by invasion scenario simulations that Wallace's classic zoogeographic delineation of mammals, amphibians, and birds (Wallace 1876) are strongly weakened by contemporary species invasions and homogenizing community compositions (Bernardo-Madrid *et al.* 2019). The human impacts of overfishing, habitat quality through bottom trawling, and pollution on coastal shelf ecosystems prompts similar questions to test for temporal boundary shifts, mixing, or displacement of historically distinct communities (McLean *et al.* 2021; Gordó-Vilaseca *et al.* 2023; Leroy *et al.* 2023), and

the future makeup of distinct ecological zones across continental shelf seas. The proposed bioregions shown here are naturally permeable as individuals can travel or be transported great distances, but we show strong modular structure in bipartite networks to support a broad bioregionalization among well-sampled, diverse, and abundant demersal fishes across shelf seas.

3.6. Data archiving statement

The observational trawl data are available through FISHGLOB (FISHGLOB – Fish biodiversity under global change (ucsc.edu)). Phylogenetic data are available through fishtree.org (FishLife (fishtree.org)). The relevant R scripts and resultant bioregion shapefiles are freely available (https://zenodo.org/uploads/13628921).

3.7. Author contributions

Liam MacNeil: Conceptualization (Lead); Data curation (Lead); Formal analysis (Lead); Writing – original draft (Lead). **Marco Scotti:** Conceptualization (Supporting); Data curation (Supporting); Funding acquisition (Lead); Methodology (Supporting); Project administration (Lead); Writing – review & editing (Supporting).

3.8. References

Álvarez-Noriega, M., Burgess, S. C., Byers, J. E., Pringle, J. M., Wares, J. P., & Marshall, D. J. (2020). Global biogeography of marine dispersal potential. *Nature Ecology & Evolution*, *4*(9), 1196–1203. https://doi.org/10.1038/s41559-020-1238-y.

Arostegui, M. C., Gaube, P., Woodworth-Jefcoats, P. A., Kobayashi, D. R., & Braun, C. D. (2022). Anticyclonic eddies aggregate pelagic predators in a subtropical gyre. *Nature*, *609*(7927), 535–540. https://doi.org/10.1038/s41586-022-05162-6.

Baselga, A. (2013). Multiple site dissimilarity quantifies compositional heterogeneity among several sites, while average pairwise dissimilarity may be misleading. *Ecography*, *36*(2), 124–128. https://doi.org/10.1111/j.1600-0587.2012.00124.x.

Beukhof, E., Frelat, R., Pecuchet, L., Maureaud, A., Dencker, T. S., Sólmundsson, J., Punzón, A., Primicerio, R., Hidalgo, M., Möllmann, C., & Lindegren, M. (2019). Marine fish traits follow fast-slow continuum across oceans. *Scientific Reports*, *9*(1), 17878. https://doi.org/10.1038/s41598-019-53998-2.

Bernardo-Madrid, R., Calatayud, J., González-Suárez, M., Rosvall, M., Lucas, P. M., Rueda, M., Antonelli, A., & Revilla, E. (2019). Human activity is altering the world's zoogeographical regions. *Ecology Letters*, *22*(8), 1297–1305. https://doi.org/10.1111/ele.13321.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

Bloomfield, N. J., Knerr, N., & Encinas-Viso, F. (2018). A comparison of network and clustering methods to detect biogeographical regions. *Ecography*, *41*(1), 1–10. https://doi.org/10.1111/ecog.02596.

Bowen, B. W., Rocha, L. A., Toonen, R. J., & Karl, S. A. (2013). The origins of tropical marine biodiversity. *Trends in Ecology & Evolution*, *28*(6), 359–366. https://doi.org/10.1016/j.tree.2013.01.018.

Briggs, J. C., (1974). Marine zoogeography. *McGraw-Hill*, New York, 475 pp.

Briggs, J. C., (2003). Guest Editorial: Marine Centres of Origin as Evolutionary Engines, *Journal of Biogeography*, *3*(1), 1–18.

Briggs, J. C., & Bowen, B. W. (2012). A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography*, *39*(1), 12–30. https://doi.org/10.1111/j.1365-2699.2011.02613.x.

Briggs, J. C., & Bowen, B. W. (2013). Marine shelf habitat: Biogeography and evolution. *Journal of Biogeography*, *40*(6), 1023–1035. https://doi.org/10.1111/jbi.12082.

Costello, M. J., Tsai, P., Wong, P. S., Cheung, A. K. L., Basher, Z., & Chaudhary, C. (2017). Marine biogeographic realms and species endemicity. *Nature Communications*, *8*(1), 1057. https://doi.org/10.1038/s41467-017-01121-2.

Cowen, R., Gawarkiewicz, G., Pineda, J., Thorrold, S., & Werner, F. (2007). Population Connectivity in Marine Systems: An Overview. *Oceanography*, *20*(3), 14–21. https://doi.org/10.5670/oceanog.2007.26.

Cowman, P. F., Parravicini, V., Kulbicki, M., & Floeter, S. R. (2017). The biogeography of tropical reef fishes: Endemism and provinciality through time. *Biological Reviews*, 92(4), 2112–2130. https://doi.org/10.1111/brv.12323.

Csardi, G., Nepusz, T. (2006). The igraph software package for complex network research. *Complex Systems*, 1695. https://igraph.org. Version 2.0.3.

[a]Daru, B. H., Farooq, H., Antonelli, A., & Faurby, S. (2020). Endemism patterns are scale dependent. *Nature Communications*, *11*(1), 2115. https://doi.org/10.1038/s41467-020-15921-6.

[b]Daru, B. H., Karunarathne, P., & Schliep, K. (2020). phyloregion: R package for biogeographical regionalization and macroecology. *Methods in Ecology and Evolution*, *11*(11), 1483–1491. https://doi.org/10.1111/2041-210X.13478.

Dormann, C. F., Gruber, B., & Fründ, J. (2008). Introducing the bipartite Package: Analysing Ecological Networks. *R News* 8, 8–11. Version 2.20.

Dormann, C. F., Frund, J., Bluthgen, N., & Gruber, B. (2009). Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. *The Open Ecology Journal*, *2*(1), 7–24. https://doi.org/10.2174/1874213000902010007.

Ebach, M. C., & Parenti, L. R. (2015). The dichotomy of the modern bioregionalization revival. *Journal of Biogeography*, *42*(10), 1801–1808. https://doi.org/10.1111/jbi.12558.

Ebach, M. C., González-Orozco, C. E., Miller, J. T., & Murphy, D. J. (2015). A revised area taxonomy of phytogeographical regions within the Australian Bioregionalisation Atlas. *Phytotaxa*, *208*(4), 261. https://doi.org/10.11646/phytotaxa.208.4.2.

Frank, K. T., Petrie, B., & Shackell, N. L. (2007). The ups and downs of trophic control in continental shelf ecosystems. *Trends in Ecology & Evolution*, *22*(5), 236–242. https://doi.org/10.1016/j.tree.2007.03.002.

Gabriel, W. L. (1992). Persistence of Demersal Fish Assemblages Between Cape Hatteras and Nova Scotia, Northwest Atlantic. *Journal of Northwest Atlantic Fishery Science*, *14*, 29–46. https://doi.org/10.2960/J.v14.a2.

Gordó-Vilaseca, C., Stephenson, F., Coll, M., Lavin, C., & Costello, M. J. (2023). Three decades of increasing fish biodiversity across the northeast Atlantic and the Arctic Ocean. *Proceedings of the National Academy of Sciences*, *120*(4), e2120869120. https://doi.org/10.1073/pnas.2120869120.

Guimerà, R., & Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature*, *433*(7028), 895–900. https://doi.org/10.1038/nature03288.

Hill, N., Woolley, S. N. C., Foster, S., Dunstan, P. K., McKinlay, J., Ovaskainen, O., & Johnson, C. (2020). Determining marine bioregions: A comparison of quantitative approaches. *Methods in Ecology and Evolution*, *11*(10), 1258–1272. https://doi.org/10.1111/2041-210X.13447.

Holt, B. G., Lessard, J.-P., Borregaard, M. K., Fritz, S. A., Araújo, M. B., Dimitrov, D., Fabre, P.-H., Graham, C. H., Graves, G. R., Jønsson, K. A., Nogués-Bravo, D., Wang, Z., Whittaker, R. J., Fjeldså, J., & Rahbek, C. (2013). An Update of Wallace's Zoogeographic Regions of the World. *Science*, *339*(6115), 74–78. https://doi.org/10.1126/science.1228282.

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, *44*(9), 1259–1269. https://doi.org/10.1111/ecog.05926.

Humboldt, A. 1806. Essai sur la géographie des plantes; accompagné d'un tableau physique des régions équinoxales, accompagné d'un tableau physique des régions e´quinoctiales. – Schoel and Co., Paris.

Johannesson, K., Le Moan, A., Perini, S., & André, C. (2020). A Darwinian Laboratory of Multiple Contact Zones. *Trends in Ecology & Evolution*, *35*(11), 1021–1036. https://doi.org/10.1016/j.tree.2020.07.015.

Juan-Jordá, M. J., Barth, J. A., Clarke, M. E., & Wakefield, W. W. (2009). Groundfish species associations with distinct oceanographic habitats in the Northern California Current. *Fisheries Oceanography*, *18*(1), 1–19. https://doi.org/10.1111/j.1365-2419.2008.00489.x.

Justić, D., Kourafalou, V., Mariotti, G., He, S., Weisberg, R., Androulidakis, Y., Barker, C., Bracco, A., Dzwonkowski, B., Hu, C., Huang, H., Jacobs, G., Le Hénaff, M., Liu, Y., Morey, S., Nittrouer, J., Overton, E., Paris, C. B., Roberts, B. J., … Wiggert, J. (2022). Transport Processes in the Gulf of Mexico Along the River-Estuary-Shelf-Ocean Continuum: A Review of Research from the Gulf of Mexico Research Initiative. *Estuaries and Coasts*, *45*(3), 621–657. https://doi.org/10.1007/s12237-021-01005-1.

Kulbicki, M., Parravicini, V., Bellwood, D. R., Arias-Gonzàlez, E., Chabanet, P., Floeter, S. R., Friedlander, A., McPherson, J., Myers, R. E., Vigliola, L., & Mouillot, D. (2013). Global Biogeography of Reef Fishes: A Hierarchical Quantitative Delineation of Regions. *PLoS ONE*, *8*(12), e81847. https://doi.org/10.1371/journal.pone.0081847.

Kreft, H., & Jetz, W. (2010). A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, *37*(11), 2029–2053. https://doi.org/10.1111/j.1365-2699.2010.02375.x.

Legrand, T., Fragkopoulou, E., González-Trujillo, J. D., & Assis, J. (2024). The global biogeography of coastal oceanographic connectivity. *PeerJ,* https://doi.org/10.21203/rs.3.rs-4319669/v1.

Leibold, M. A., Govaert, L., Loeuille, N., De Meester, L., & Urban, M. C. (2022). Evolution and Community Assembly Across Spatial Scales. *Annual Review of Ecology, Evolution, and Systematics*, *53*(1), 299–326. https://doi.org/10.1146/annurev-ecolsys-102220-024934.

Leroy, B., Bellard, C., Dias, M. S., Hugueny, B., Jézéquel, C., Leprieur, F., Oberdorff, T., Robuchon, M., & Tedesco, P. A. (2023). Major shifts in biogeographic regions of freshwater fishes as evidence of the Anthropocene epoch. *Science Advances*, *9*(46), eadi5502. https://doi.org/10.1126/sciadv.adi5502.

Longhurst, A. (2010), *Ecological Geography of the Sea*, Academic Press, London.

Maureaud, A. A., Frelat, R., Pécuchet, L., Shackell, N., Mérigot, B., Pinsky, M. L., Amador, K., Anderson, S. C., Arkhipkin, A., Auber, A., Barri, I., Bell, R. J., Belmaker, J., Beukhof, E., Camara, M. L., Guevara-Carrasco, R., Choi, J., Christensen, H. T., Conner, J., … T. Thorson, J. (2021). Are we ready to track climate-driven shifts in marine species across international boundaries? - A global survey of scientific bottom trawl data. *Global Change Biology*, *27*(2), 220–236. https://doi.org/10.1111/gcb.15404.

Maureaud, A. A., Palacios-Abrantes, J., Kitchel, Z., Mannocci, L., Pinsky, M. L., Fredston, A., Beukhof, E., Forrest, D. L., Frelat, R., Palomares, M. L. D., Pecuchet, L., Thorson, J. T., Van Denderen, P. D., & Mérigot, B. (2024). FISHGLOB_data: An integrated dataset of fish biodiversity sampled with scientific bottom-trawl surveys. *Scientific Data*, *11*(1), 24. https://doi.org/10.1038/s41597-023-02866-w.

McLean, M., Mouillot, D., Maureaud, A. A., Hattab, T., MacNeil, M. A., Goberville, E., Lindegren, M., Engelhard, G., Pinsky, M., & Auber, A. (2021). Disentangling tropicalization and deborealization in marine ecosystems under climate change. *Current Biology*, *31*(21), 4817-4823.e5. https://doi.org/10.1016/j.cub.2021.08.034.

Meilă, M. (2007). Comparing clusterings—An information based distance. *Journal of Multivariate Analysis*, *98*(5), 873–895. https://doi.org/10.1016/j.jmva.2006.11.013.

Momigliano, P., Jokinen, H., Fraimout, A., Florin, A.-B., Norkko, A., & Merilä, J. (2017). Extraordinarily rapid speciation in a marine fish. *Proceedings of the National Academy of Sciences*, *114*(23), 6074–6079. https://doi.org/10.1073/pnas.1615109114.

Momigliano, P., Denys, G. P. J., Jokinen, H., & Merilä, J. (2018). Platichthys solemdali sp. nov. (Actinopterygii, Pleuronectiformes): A New Flounder Species From the Baltic Sea. *Frontiers in Marine Science*, *5*, 225. https://doi.org/10.3389/fmars.2018.00225.

Morrone, J. J. (2018). The spectre of biogeographical regionalization. *Journal of Biogeography*, *45*(2), 282–288. https://doi.org/10.1111/jbi.13135.

Morrone, J. J. (2024). Why biogeographical transition zones matter. *Journal of Biogeography*, *51*(4), 544–549. https://doi.org/10.1111/jbi.14632.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103.

Pappalardo, P., Pringle, J. M., Wares, J. P., & Byers, J. E. (2015). The location, strength, and mechanisms behind marine biogeographic boundaries of the east coast of North America. *Ecography*, *38*(7), 722–731. https://doi.org/10.1111/ecog.01135.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal, 10(1), 439. https://doi.org/10.32614/RJ-2018-009. Version 1.0.16.

Pecuchet, L., Törnroos, A., & Lindegren, M. (2016). Patterns and drivers of fish community assembly in a large marine ecosystem. *Marine Ecology Progress Series*, *546*, 239–248. https://doi.org/10.3354/meps11613.

Pini, A., & Vantini, S. (2016). The Interval Testing Procedure: A General Framework for Inference in Functional Data Analysis. *Biometrics*, *72*(3), 835–845. https://doi.org/10.1111/biom.12476.

Policastro, V., Righelli, D., Carissimo, A., Cutillo, L., & Feis, I., De. (2021). ROBustness In Network (robin): An R Package for Comparison and Validation of Communities. *The R Journal*, *13*(1), 292. https://doi.org/10.32614/RJ-2021-040. Version 1.2.0.

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T. J., Coll, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, *559*(7714), 392–395. https://doi.org/10.1038/s41586-018-0273-1.

Reichardt, J., & Bornholdt, S. (2006). Statistical Mechanics of Community Detection. *Physical Review E*, *74*(1), 016110. https://doi.org/10.1103/PhysRevE.74.016110.

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., & Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, *27*(4), 1046–1058. https://doi.org/10.1002/gbc.20089.

Richardson, P. L. (2001). Florida current, gulf stream, and labrador current. in Encyclopedia of ocean sciences, Ed. Steele J. H. (Amsterdam: Elsevier Ltd), 1054–1064.

Ricklefs, R. E., & Jenkins, D. G. (2011). Biogeography and ecology: Towards the integration of two disciplines. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1576), 2438–2448. https://doi.org/10.1098/rstb.2011.0066.

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105.

Rousseau, Y., Blanchard, J. L., Novaglio, C., Pinnell, K. A., Tittensor, D. P., Watson, R. A., & Ye, Y. (2024). A database of mapped global fishing activity 1950–2017. *Scientific Data*, *11*(1), 48. https://doi.org/10.1038/s41597-023-02824-6.

Rueda, M., González-Suárez, M., & Revilla, E. (2024). Global biogeographical regions reveal a signal of past human impacts. *Ecography*, *2024*(3), e06762. https://doi.org/10.1111/ecog.06762.

Sherman, K. (1991). The Large Marine Ecosystem Concept: Research and Management Strategy for Living Marine Resources. *Ecological Applications*, *1*(4), 349–360. https://doi.org/10.2307/1941896.

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. *BioScience*, *57*(7), 573–583. https://doi.org/10.1641/B570707.

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., & Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, *466*(7310), 1098–1101. https://doi.org/10.1038/nature09329.

Title, P. O., Swiderski, D. L., & Zelditch, M. L. (2022). ECOPHYLOMAPPER: An R package for integrating geographical ranges, phylogeny and morphology. *Methods in Ecology and Evolution*, *13*(9), 1912–1922. https://doi.org/10.1111/2041-210X.13914. Version 1.1.2.

Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. – *Economic Geography*, *46*, 234–240.

Toonen, R. J., Bowen, B. W., Iacchei, M., & Briggs, J. C. (2016). Biogeography, Marine. In *Encyclopedia of Evolutionary Biology* (pp. 166–178). Elsevier. https://doi.org/10.1016/B978-0-12-800049-6.00120-7.

Traag, V. A., Waltman, L., & Van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*(1), 5233. https://doi.org/10.1038/s41598-019-41695-z.

Van Denderen, D., Maureaud, A. A., Andersen, K. H., Gaichas, S., Lindegren, M., Petrik, C. M., Stock, C. A., & Collie, J. (2023). Demersal fish biomass declines with temperature across productive shelf seas. *Global Ecology and Biogeography*, *32*(10), 1846–1857. https://doi.org/10.1111/geb.13732.

Vázquez, D. P., Melián, C. J., Williams, N. M., Blüthgen, N., Krasnov, B. R., & Poulin, R. (2007). Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, *116*(7), 1120–1127. https://doi.org/10.1111/j.0030-1299.2007.15828.x.

Victorero, L., Samadi, S., O'Hara, T. D., Mouchet, M., Delavenne, J., Leprieur, F., & Leroy, B. (2023). Global benthic biogeographical regions and macroecological drivers for ophiuroids. *Ecography*, *2023*(9), e06627. https://doi.org/10.1111/ecog.06627.

Vilhena, D. A., & Antonelli, A. (2015). A network approach for identifying and delimiting biogeographical regions. *Nature Communications*, *6*(1), 6848. https://doi.org/10.1038/ncomms7848.

Wallace, A.R. (1876). The Geographical Distribution of Animals, *Cambridge University Press*. Cambridge.

Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, *19*(12), 639–644. https://doi.org/10.1016/j.tree.2004.09.011.

Mendelssohn, I. A., Byrnes, M. R., Kneib, R. T., Vittor, B. A. (2017). Coastal Habitats of the Gulf of Mexico, In: Ward, C. H. (Ed.). *Habitats and Biota of the Gulf of Mexico: Before the Deepwater Horizon Oil Spill: Volume 1: Water Quality, Sediments, Sediment Contaminants, Oil and Gas Seeps, Coastal Habitats, Offshore Plankton and Benthos, and Shellfish*. Springer New York. https://doi.org/10.1007/978-1-4939-3447-8.

Warren, D. L., Cardillo, M., Rosauer, D. F., & Bolnick, D. I. (2014). Mistaking geography for biology: Inferring processes from species distributions. *Trends in Ecology & Evolution*, *29*(10), 572–580. https://doi.org/10.1016/j.tree.2014.08.003.

Zimmermann, E.A.G. 1777: Specimen Zoologiæ Geographicæ. Quadru-pedum. 685 pp. Theodorum Haak et Socios, Lugduni Batavorum.

4. Environmental filtering drives widespread trait convergence in marine demersal fishes

This chapter is based on a draft which has been submitted to *Nature Ecology & Evolution*.

MacNeil, L., McLean, M., Tittensor, D.P., Bayer, T., Reusch, T.B.H., Scotti, M. Environmental filtering drives widespread trait convergence in marine demersal fishes, Submitted to *Nature Ecology & Evolution*.

4.1. Abstract

Understanding the processes that shape the distribution of biodiversity in the oceans is central for predicting and conserving ecosystems under global change. Although a vast literature exists on drivers of species diversity, the geographical patterns and drivers of ecosystem functioning, and in particular the traits that shape this functioning, remain relatively unexplored. Here we aggregate trawl and trait information (body size, habitat, reproduction, trophic ecology, and growth) for 1150 demersal fish species on continental shelf seas throughout the Northern Hemisphere to test the relative importance of environmental, evolutionary, and anthropogenic drivers in shaping trait distributions at ocean-basin scales. We find that environmental filtering has strongly shaped the functional makeup of fish communities while, in contrast to expectations, trait conservatism across evolutionary lineages appears uniquely strong in the Pacific Ocean but less important in the Atlantic. Although fishing can alter traits and deplete populations, it shows no explanatory power in describing trait compositions. The widespread role of environmental conditions in shaping fish traits highlights the potential sensitivity of community functioning to environmental and climate change, and sheds new light on the potential for trait-based conservation strategies.

4.2. Introduction

Disentangling the processes that govern the distribution, composition, and maintenance of biodiversity are key questions in ecology and conservation. While the distribution of species diversity has received considerable attention, the distributions and potential drivers of individual traits and functional roles are far less understood (Violle *et al.* 2014). Ecological traits are proxies for an individual's functioning in an ecosystem (McGill *et al.* 2006; Violle *et al.* 2007). Trait diversity (and hence functional diversity) reflects differing interactions with the surrounding environment (Luiz *et al.* 2012; Bruelheide *et al.* 2018) and is a potential window into the large-scale drivers generating different trait compositions (McLean *et al.* 2021; Bosch *et al.* 2021) and community-wide responses to disturbance (Moretti & Legg, 2009; Mouillot *et al.* 2013; [a]Beukhof *et al.* 2019). Growing datasets of trait variations across domains of life (Carmona *et al.* 2021) offer new possibilities for exploring the distribution of ecological functions and functional diversity at different spatial scales (Messier *et al.* 2010). The biogeography of traits can also generate insight into how anthropogenic impacts may reshape the ecological functioning of the biosphere and explore novel conservation strategies.

Given that traits likely impact species abilities to colonize and persist in a given ecosystem, trait biogeography can test hypotheses about the mechanisms governing community assembly. The processes shaping the contemporary composition and distribution of life occur through multiple dispersal, environmental, and biotic filters (Leibold *et al.* 2022). Two contrasting explanatory frameworks include (i) convergent similarity, where large-scale environmental filters generate adaptive similarity, and (ii) divergent partitioning, where biotic filters (e.g. competing taxa) limit niche or trait similarity and drive niche partitioning (e.g., MacArthur & Levins, 1967; Diamond, 1975; Samuels & Drake, 1997; Chase, 2003). In theory, evolutionary history should provide insight into community assembly processes generating trait evolution, as recently diverged taxa are expected to be more similar (Webb *et al.* 2002; Kraft *et al.* 2007). Convergent similarity and divergent partitioning may not be mutually exclusive, as different assembly processes can produce patterns of either scenario across environmental gradients (Spasojevic & Suding, 2012), and whether trait compositions deviate from a neutral model (Hubbell, 2001) has become a standard test for the apparent strength of any biodiversity pattern.

Marine fishes are a useful model taxon for testing patterns of community assembly as the most diverse vertebrate group with high dispersal capacity, and widespread availability of data on

traits and evolutionary histories (Villéger *et al.* 2017; Rabosky *et al.* 2018; McLean *et al.* 2021; Bosch *et al.* 2021). For marine fishes, environment and evolutionary history are not isolated factors: human-induced pressures including fishing have significantly altered marine ecosystems (IPBES, 2019) and can also drive abrupt shifts in body size and other life-history traits through harvest selection (Darimont *et al.* 2009). In the marine realm, the broad spatial patterns of trait evolution and the relative role of environmental conditions, fishing pressure, and evolutionary history are largely unknown. Here, we address this gap by testing the relative effect of these different drivers on trait compositions in demersal fish communities throughout continental shelves in the Northern Hemisphere. We test the generality of the relationship between environment, fishing pressure, and evolutionary history for fish trait compositions across the Pacific and Atlantic Oceans using standardized scientific trawl surveys.

4.3. Methods

To test how environment, fishing, and evolutionary history shape the trait composition of demersal fish, we assembled scientific bottom trawls and species traits, constructed trait spaces using statistical differences in species-trait compositions, characterized broad spatial differences in environmental conditions and fishing pressure, and estimated differences in evolutionary histories through phylogenetic turnover. We also assessed linear and non-linear relationships between mean trait space and each covariate, the spatial scaling of these relationships, the impact of sampling intensity, and their deviation from null and neutral models of trait evolution.

4.3.1. Scientific trawl data

We used scientific bottom trawl data from FISHGLOB (Maureaud *et al.* 2024; Supplementary Table 9.1), representing 29 different surveys across Northern Hemisphere continental shelves (Fig. 4.1). We retained trawls between 1999–2021, given that a majority of surveys (17/29) had already begun in 1999 (Supplementary Fig. 9.1). All data preparation and subsequent analyses used the R programming language (R Core Team, 2024). Details on pretreatments to exclude rare trawls and spatiotemporal filters for limiting outliers are in Supplement S1 (Chapter 9). For further analyses, we constructed equal-area grids to equalize sampling effort (Supplementary Fig. 9.2) by rarefying to a defined number of trawls across three different spatial resolutions of 100 km^2, 200 km^2 hexagonal grid cells, and Ecoregions (Spalding *et al.* 2007); we randomly sampled trawls at two levels of sampling intensity based on the 1st quartile and median number trawls for each resolution (Supplementary Table 9.2).

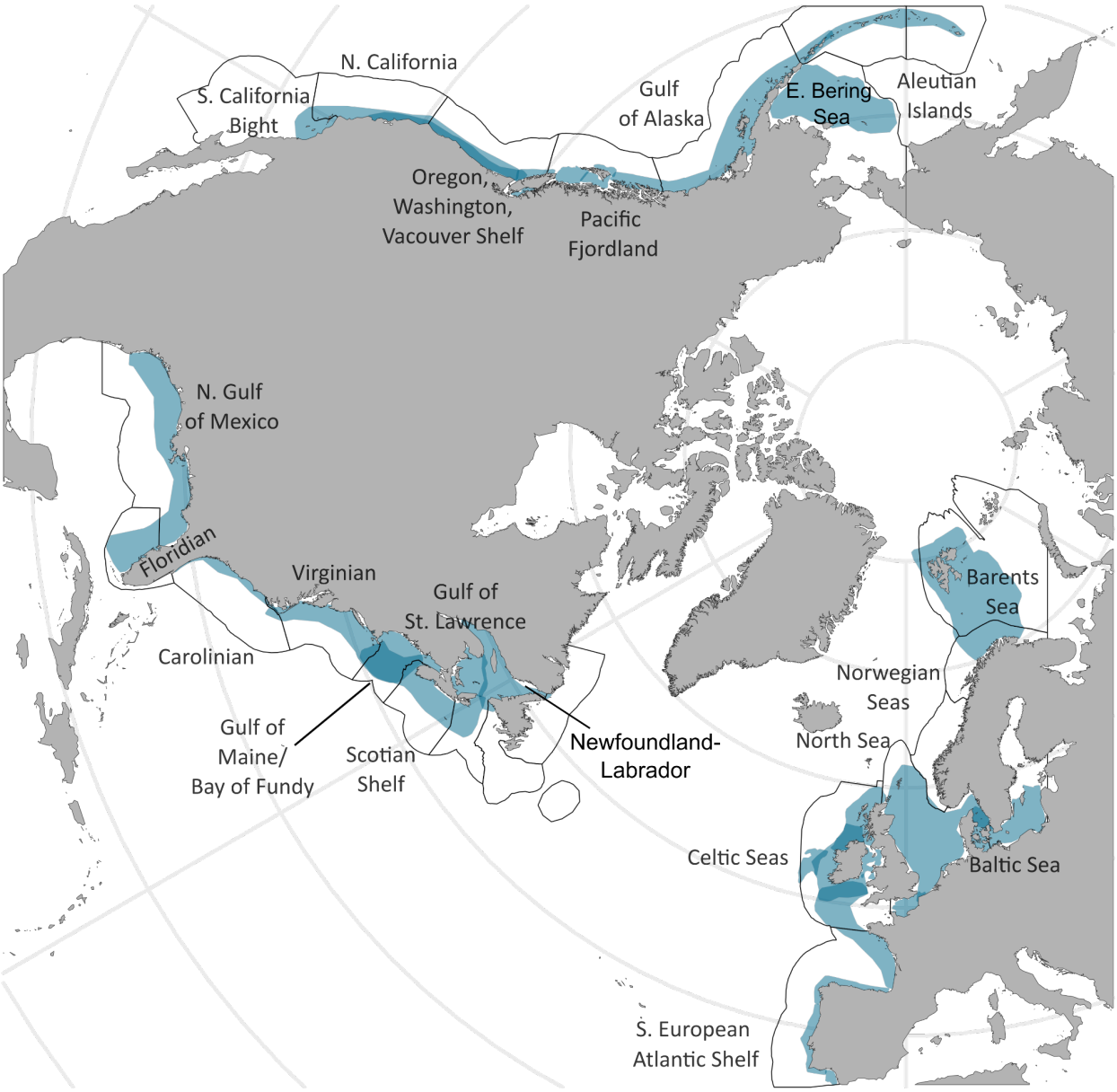


Figure 4.1 The spatial footprint of scientific trawl data ranging from 1999 to 2021. Blue polygons represent 10 km buffers surrounding each unique survey, with regions containing overlapping survey coverage in darker blues. Marine Ecoregions are plotted as black polygons and annotated.

4.3.2. Trait data

Eight traits were extracted from [b]Beukhof *et al.* (2019) for species found in the North Atlantic and Pacific continental shelves based on FishBase (Froese and Pauly, 2019). These represent multiple dimensions of fish ecology and life history (Villéger *et al.* 2017; [c]Beukhof *et al.* 2019), including habitat (depth associations), reproduction (spawning type, age at maturity), trophic ecology (trophic level, feeding mode), growth (Von Bertalanffy growth coefficient K), lifespan

(maximum age), and size (maximum length). Body size, growth, lifespan, age at maturity and trophic level are continuous traits, while the remaining traits are categorical (factors). Age-related traits showed high pairwise correlations (R = 0.65) and similar axis loadings in a principal component analysis, so we excluded lifespan, leaving all remaining continuous traits correlated at R < 0.48. Body size, growth, and age at maturity were highly skewed and were transformed with a natural logarithm. All traits are listed as proportions in Supplementary Fig. 9.3. A total of 78% of species (1320 of 1672) from the trawls were matched with the trait dataset ([b]Beukhof *et al.* 2019). To impute missing species traits, we adopted the structural equation model of Thorson *et al.* (2023) by estimating trait covariance of phylogenetic relatives for missing taxa.

4.3.3. Environment, fishing pressure, and phylogeny

Oceanographic variables were extracted at 0.05° resolution using 2000-2020 means from BIO-ORACLE v3.0 (Assis *et al.* 2024) and included mean ocean depth, mean and seasonal range of bottom temperature and salinity, and mean and seasonal ranges of surface chlorophyll-a concentration. To represent fishing pressure, we extracted effective fishing effort (kW x days at sea) from Rousseau *et al.* (2024) at 0.5° resolution for the timespan 1999–2017, excluding all pelagic-targeting fleets. These spatial files were aggregated to mean values at different spatial scales (100 km^2, 200 km^2, Ecoregions) defined by the trawl data. To represent evolutionary history for each species, we pruned a time-calibrated phylogeny of ray-finned fishes (Rabosky *et al.* 2018) from *fishtree* (Chang *et al.* 2019) to retain only species found within trawls. In total, 1533 species from the trawl dataset were found in the phylogenetic tip names which were dominated by Teleosts (> 99%), excluding predominantly chondrichthyans (n = 169) with one hagfish and two lamprey species. This pruned dataset defined by available species phylogenies (n_{trawls} = 132,275), is the basis for the following analysis. All covariates are visualized in the supplemental material (Supplementary Fig. 9.4).

4.3.4. Trait space

We calculated trait space for all species using a Principal Coordinates Analysis (PCoA) on a species × trait Gower's dissimilarity matrix, which can handle categorical and continuous trait values. We applied an adaptation of Gower's dissimilarity (*gawdis*; De Bello *et al.* 2021) which limits the unequal contribution of traits by weighting individual traits that highly correlate with multi-trait dissimilarities. We also calculated the richness of unique trait combinations (functional entities) using the *mFD* package (Magneville *et al.* 2022). To reduce

hyperdimensionality in trait spaces and ensure dimensionality reduction accurately captures the original species differences, we used the first six PCoA axes to represent trait space. This dimensionality was selected by testing the mean absolute difference (MAD) between species differences in the global Gower's dissimilarity matrix (trait space) and Euclidean distances of trait space across the first ten PCoA axes (Maire *et al.* 2015). We found that differences were minimized at six dimensions, explaining 77% of total variance while being significantly related to the global dissimilarities (MAD = 0.032, Mantel's R = 0.90, $p < 0.001$). Six dimensions is typical within the wide range of optimal dimensions across different ecological datasets (Moulliot *et al.* 2021). The first and second axes are the primary focus of visualization here and explain 46% of the variance and are also significantly related to the global dissimilarities (Mantel's R = 0.74, $p < 0.001$).

To spatially characterize species-trait combinations in trait space, we calculated the mean position (centroids) of each spatial unit (100 km^2, 200 km^2, Ecoregions) from random sampling without replacement of trawls to summary statistic values (described above) and bootstrap this process 99 times to calculate PCoA centroids with standard errors (Supplementary Fig. 9.5–9.7). The first six axes of trait space were highly similar across spatial resolutions and sampling intensities (Supplementary Fig. 9.8), so we hereafter only visualize trait space patterns at 200 km^2 using 133 trawls per grid cell (1st quartile of trawls) because it is the most spatially complete. To check for yearly variability in trait space centroids, we randomly sampled species-traits from each grid cell within each year (200 km^2, 1st quartile number of trawls), limiting outliers by considering only grid cells that contain at least 10 species in a year (Supplementary Fig. 9.9).

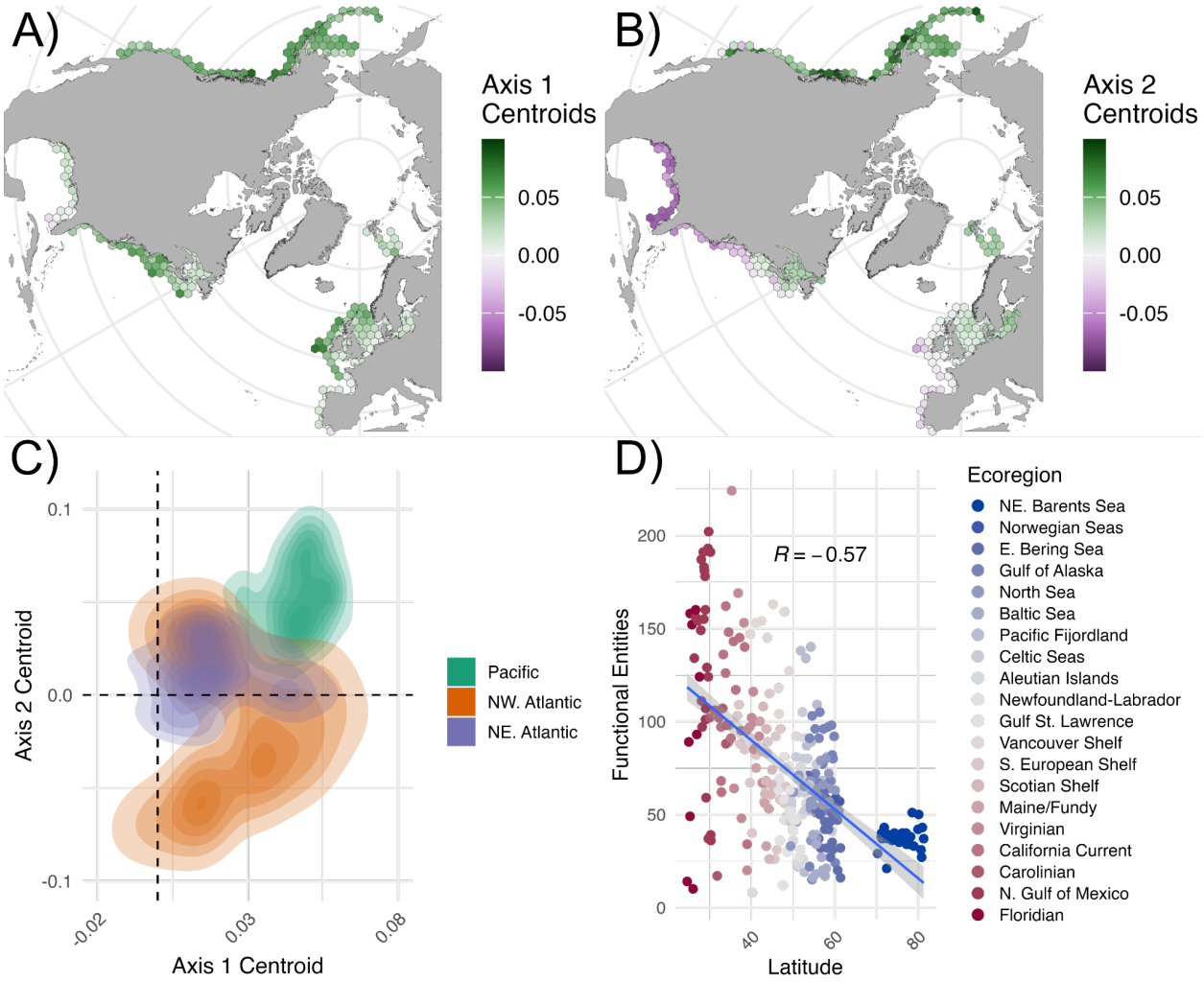


Figure 4.2 Geographic trait space and functional entities of demersal fish. The PCoA centroids at 200 km^2 for A–C) PCoA axis 1–2. D) The latitudinal gradient in the number of functional entities per grid cell for the associated Ecoregion (Supplementary Table 9.3).

4.3.5. Relationship between environment, fishing, and phylogeny to trait space

We compared relationships between Euclidean distances of trait centroids in the first six axes of trait space to Euclidean distances in environment, fishing pressure, and phylogenetic turnover. Due to their differing scales, environmental variables were scaled by their mean and standard deviations before calculating pairwise Euclidean distances. Phylogenetic turnover was calculated as the mean pairwise Jaccard dissimilarity in lineage turnover (Leprieur *et al.* 2012) using the *betapart* package (Baselga & Orme, 2012). Phylogenetic turnover is bounded between 0–1, ranging from identical to complete lineage replacement. We show the linear relationships (Pearson's R, covariance R^2, two-sided t-test p-value) between trait dissimilarities and environment, fishing, and evolutionary history, but also explored non-linear relationships

using several hierarchical generalized additive models (HGAMs; Pedersen *et al.* 2019) in Supplement S2 (MacNeil, 2025).

4.3.6. Null and neutral models

We compared observed trait dissimilarity to a null and a neutral model. We constructed the null model by spatially randomizing species-traits occurrences while maintaining species richness using the *picante* package (*randomizeMatrix* "*richness*" null model; Kembel *et al.* 2010). We then calculated the correlation between PCoA 1 centroids and latitude across 99 bootstraps (200 km^2, 133 trawls per grid cell) against this null model to test if trait space patterns could be produced by species richness. To test whether the relationship between observed trait centroids and environment, fishing pressure, or evolutionary history was stronger than expected by neutral drift, we simulated trait evolution using the PCoA 1 axes as a continuous trait proxy under a neutral model based on Brownian motion (Mazel *et al.* 2017; McLean *et al.* 2021) across the phylogenetic tree of ray-finned fishes (Rabosky *et al.* 2018) in the *geiger* package (*fitContinuous*, Pennell *et al.* 2014). This estimates a rate of trait variance that we used to simulate (PCoA 1) values across the phylogenetic tree under neutral drift for 999 repetitions with a root value of zero (*sim.char*, Pennell *et al.* 2014). Simulated neutral trait values for each species were then randomly sampled across 200 km^2 grid cells with 133 trawls per grid cell and bootstrapped across 99 repetitions identically to observed trait values. We then calculated standard effect sizes (SES; Mazel *et al.* 2017) between the Euclidean distance ($dist$) of observed and neutral traits:

$$SES = \frac{dist(observed) - dist(neutral)}{\sigma_{neutral}} \quad (1)$$

Where $\sigma_{neutral}$ represents the standard deviation of neutral trait centroids across 99 bootstraps. Values of SES above zero indicate that observed trait dissimilarities are more different than those expected under neutral drift while values below zero indicate that observed trait dissimilarities are more similar than those expected under neutral drift (Mazel *et al.* 2017; McLean *et al.* 2021). We also recalculated the relationship between dissimilarities in environment, fishing pressure, phylogenetic turnover, and neutral traits to test if observed relationships differ significantly from a neutral expectation.

4.4. Results

4.4.1. A fast-slow trait space

The first two PCoA axes of the global pool in species-traits varied across a fast-slow life history continuum (Supplement Fig. 9.10), where a latitudinal gradient emerged such that (sub)tropical regions were associated with fast growing and maturing species at lower trophic levels, which were smaller in size and often herbivorous and planktivorous; temperate/polar latitudes more often contained slow growing and maturing species within higher trophic levels which were larger in size with generalist feeding strategies. Trait space values are more strongly related to latitude than expected by a richness-constrained null model, suggesting the underlying patterns are not only influenced by the latitudinal diversity gradient (Supplementary Fig. 9.11). Geographically, PCoA 1 centroids were strongly positive and strikingly similar across ocean basins (Fig. 4.2A, C) although a latitudinal gradient in trait spaces appears strongest in the northwest Atlantic Ocean (Fig. 4.2A–B, D). Mean trait space also broadly represents yearly trait space variability (Supplementary Fig. 9.9).

4.4.2. Environment is a greater driver of trait composition than evolutionary history

For all pairwise gridcell comparisons, trait dissimilarities are always positively correlated to environmental dissimilarity ($R = 0.41$, $R^2 = 0.16$, $p < 2e^{-16}$, Fig. 4.3 A). Although we focused on the linear relationships to describe trait space patterns, we also find this positive relationship across different HGAMs across spatial scales (100 km^2, 200 km^2, Ecoregions) (Supplement S2). In contrast, trait dissimilarities were unrelated ($R = 0.064$, $R^2 < 0.0045$, $p = 0.064$) to dissimilarities in effective fishing pressure at each spatial scale. This positive relationship between traits and environmental differences is present across 20 Ecoregions in the Northern Hemisphere, albeit with varying strength by region and latitude (Fig. 4.3B, D). The positive relationship was strongest in the Atlantic Ocean ($\underline{x}_R = 0.52$), where the northwest Atlantic Ocean shelves, and the (sub)tropical latitudes (25–35°N; the Gulf of Mexico and southeastern US shelf) in particular, contained stronger associations between trait and environmental dissimilarities. This pattern is weaker on the northeast Atlantic shelves ($\underline{x}_R = 0.39$), but still greater than phylogenetic turnover. These relationships are weakest in the Pacific Ocean (Fig. 4.3B), where, similarly to the northeast Atlantic, the strongest relationship to environment occurs at the highest latitudes (East Bering Sea; 55–65°N) and are otherwise comparatively

weak across the vast latitudinal range from the Gulf of Alaska through the California Current (35–60°N).

Phylogenetic turnover shows generally weaker associations to trait dissimilarity, but was found to be strongly positive in the Pacific Ocean ($\underline{x}_R = 0.46$, $\underline{x}_{R^2} = 0.21$) and negative on the southeastern US shelf (Carolinian, Floridian) and in the Gulf of Mexico ($\underline{x}_R = -0.33$, $\underline{x}_{R^2} = 0.11$). The HGAMs support the strongest effects for phylogenetic turnover on trait composition in the Pacific Ocean and at the coarsest, Ecoregion scale (Supplement S2). Across all regions, phylogenetic turnover is weakly related to environmental dissimilarity (Supplementary Fig. 9.12).

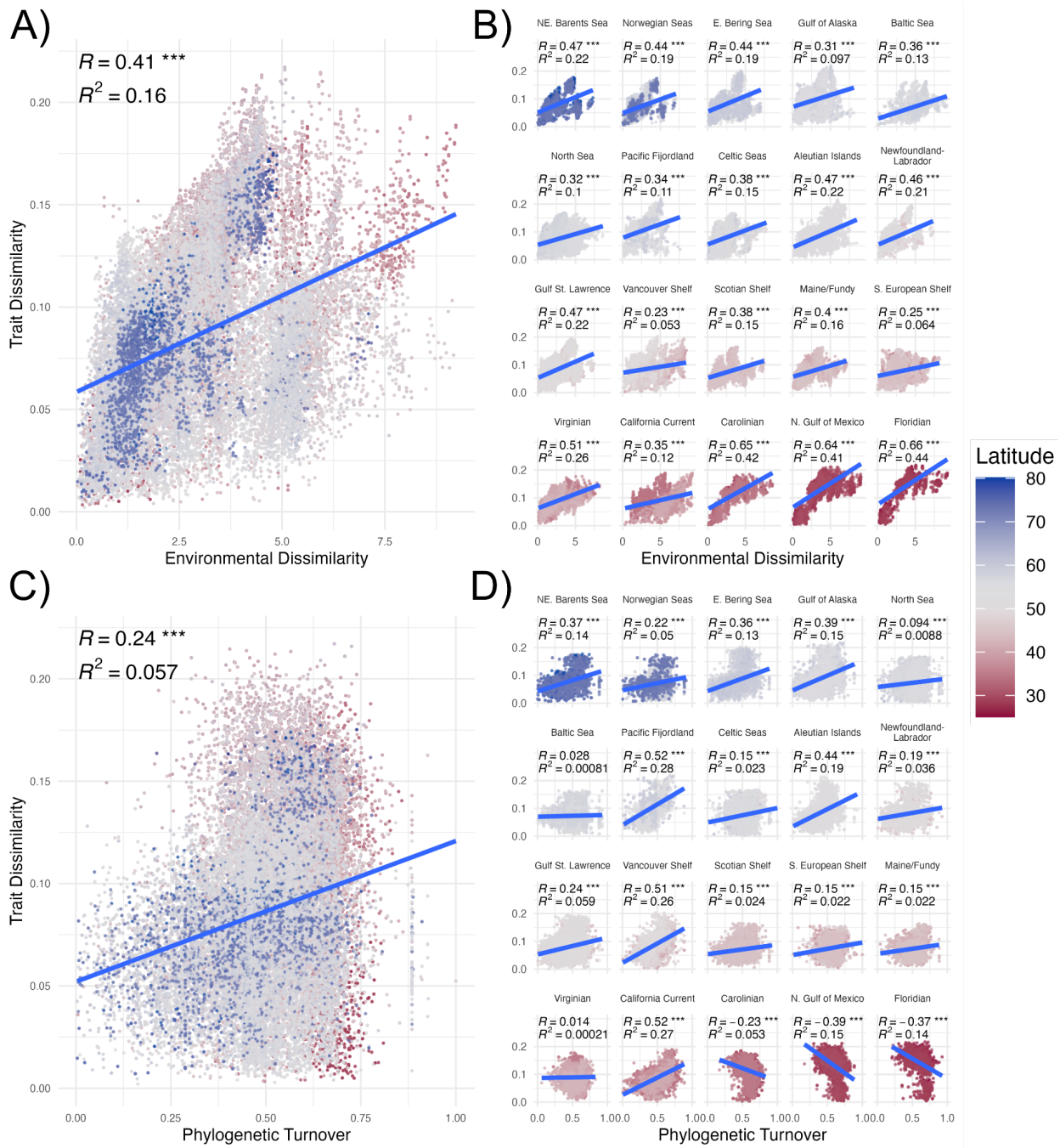


Figure 4.3 Trait dissimilarities are more strongly related to environment (A–B) than to evolutionary history (C–D). Each point represents grid cell pairwise comparisons for all grid cells (A, C) and within Ecoregions (B, D), colored by latitude. Pearson's correlation coefficient (R), covariance (R^2), and p-values (*** $p < 0.001$, no asterisks $p > 0.1$) for Pearson's R are reported in panels.

4.4.3. Does neutral drift explain marine fish traits?

Examining standard effect sizes between dissimilarities of observed and simulated (neutral) traits, a substantial fraction of observed trait compositions surround zero and could possibly be

generated under neutral drift (Fig. 4.4). However, the majority (~81%) of grid cell comparisons produced negative standard effect sizes, suggesting that most trait compositions are more similar than expected under trait evolution governed by ecological drift. The evidence for trait divergence from similarity beyond neutral drift is comparatively rare, however it does occur most frequently near the margins of each continental shelf (Gulf of Alaska, Florida coast, S. European Shelf) and some coastal seas (Pacific Fjordland and Gulf of Maine/Fundy; Fig. 4.4A–B). Despite the potential for neutral drift to explain trait compositions based on standard effect sizes, observed trait differences are more strongly related to environment or evolutionary history than expected under ecological drift (Supplementary Fig. 9.13).

4.5. Discussion

This study documents widespread evidence that contemporary environmental conditions are a main driver of trait composition in demersal marine fishes. We show that trait and environmental differences are positively related from grid cells to Ecoregions— scaling spatially across three orders of magnitude (100 km^2 → Ecoregions)— throughout vast sections of the Northern Hemisphere's shelf seas. The negative relationship between phylogenetic turnover and traits on the southeastern US shelves gives additional evidence for trait convergence across distant evolutionary lineages (Cadotte *et al.* 2013). This widespread relationship implies that although trait compositions vary substantially across environments, they generally converge under similar environmental conditions. Trait compositions are more similar than expected by neutral drift. Convergence could be generated either by environmental filtering in modern oceans from traits which were originally widespread, or through evolutionary convergence where traits evolved independently under similar conditions (Losos, 2011; Mazel *et al.* 2017).

Strongly positive associations of phylogenetic turnover and trait dissimilarities in the Pacific are evidence for phylogenetic conservatism in traits and the importance of evolutionary history. The Pacific is the oldest ocean basin and tropical reef fish biodiversity has been shaped by habitat transformations during tectonic shifts some 140 million years ago (Ma). before the North Atlantic basin had opened some 60 Ma. later (Leprieur *et al.* 2016). However, the Pacific shelf seas along the North American continent contain families that are, on average, 18 Ma. younger than those in the Atlantic (Supplementary Fig. 9.14), potentially reflecting that trait conservatism is due to limited evolutionary time for accumulating adaptive change. These

shelves are also comparatively young, having become ice-free only after sea-level rise and the opening of the Bering Strait at the end of the Younger-Dryas ~13 ka ago (Peltier, 1994). Except the East Bering Sea, these relatively young shelf seas are also narrower than those in the Atlantic, and oceanographic gradients are potentially less frequent barriers (Jönsson & Watson, 2016) to dispersal and environmental filtering. Evolutionary history is clearly important in determining the pool of species adaptations (Mittelbach & Schemske, 2015) and defining the regions from which groups disperse (Mora *et al.* 2003), however many traits have evolved independently in marine fishes (Floeter *et al.* 2018) and it appears that convergence across distant lineages towards a similar trait makeup could be common.

While the role of biotic interactions also remains challenging to resolve at local scales, evolutionary history is a potential signal of biotic filtering as close relatives competing for similar resources may eventually limit coexistence, but this mechanism for community assembly appears weakly supported by observations (Gerhold *et al.* 2015; [b]Cadotte *et al.* 2017; McLean *et al.* 2021). However, we have provided novel insights that the northeastern shelves in the Pacific Ocean appear to harbour trait compositions that have been uniquely constrained by evolutionary history. In explaining why this is not the case in the North Atlantic Ocean, phylogenetic and trait dispersions could be disconnected in marine fishes because more evolutionary time has passed for many lineages to generate adaptive similarity (Supplementary Fig. 9.14), or high rates of trait evolution, the varying spatial scale of species pools, or that competition between relatives is less important in structuring ecological communities ([b]Cadotte *et al.* 2017). In fact, phylogenetic diversity has been shown to poorly track functional diversity on land (mammals, birds) and in the oceans (Mazel *et al.* 2018).

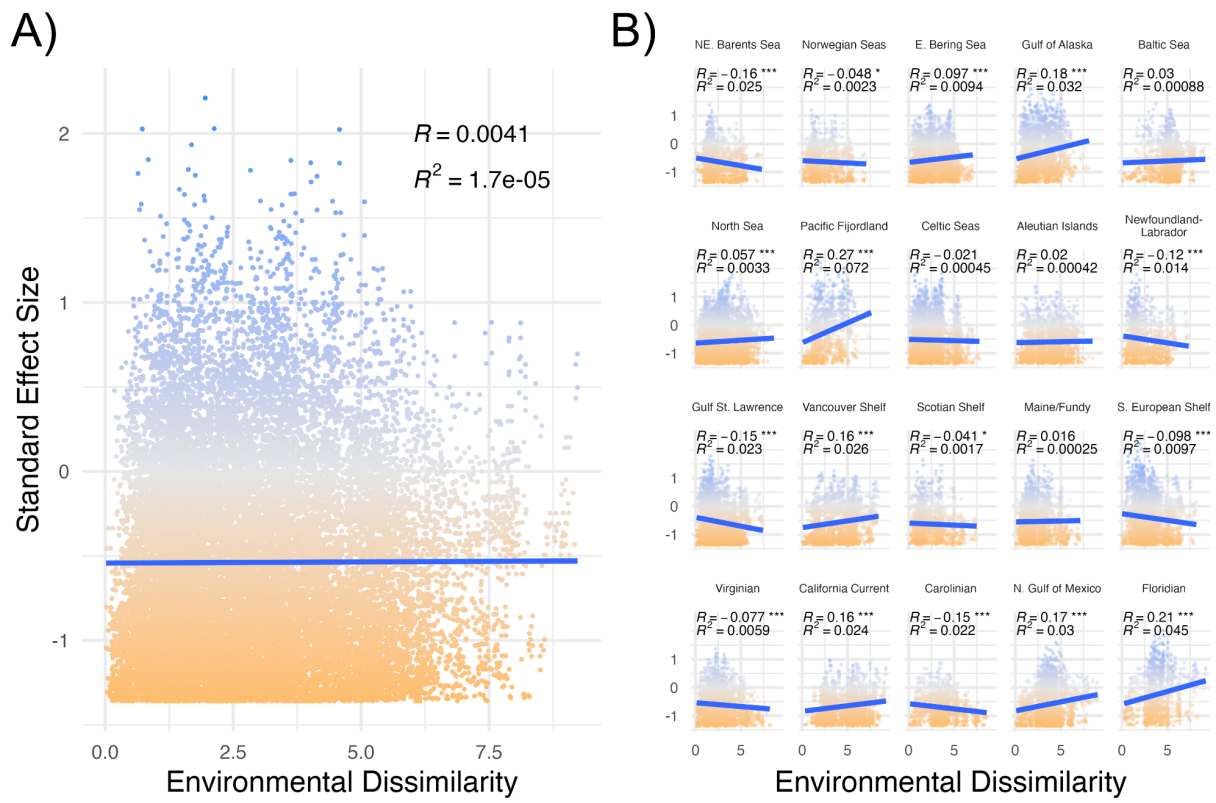


Figure 4.4 Trait compositions are generally more similar than expected by neutral drift. The standard effect sizes (equation 1) between observed and simulated (neutral) traits are visualized against environmental dissimilarities. Positive values (blue) indicate trait compositions more different from neutral drift and negative values (orange) indicate evidence for trait compositions more similar than expected under neutral drift. Each point represents 200 km^2 grid cell pairwise comparisons (A) and within Ecoregions (B). Pearson's correlation coefficient (R), covariance (R^2), and p-values p-values (*** $p < 0.001$, ** $p < 0.05$, * $p < 0.1$) for Pearson's R are reported in panels.

Environmental filtering as a process driving community-wide trait convergence has been criticized for a lack of clarity in its underlying mechanisms and for potentially overlooking other confounding processes that can produce similar patterns (Kraft *et al.* 2015; Cadotte & Tucker, 2017). A precondition for a role of environmental filtering is the establishment of widespread relationships between species traits and environment (Cadotte & Tucker, 2017), which we demonstrate here. However, the precise mechanisms by which environmental conditions have filtered demersal fish traits by limiting some functional groups' ability to colonize and persist across environmental gradients is largely unknown. This process would select for fewer trait combinations and support convergence in trait evolution— however there

is clearly substantial variation in our observations (Fig. 4.3) which indicates potential deviations of individual communities from this overarching pattern.

The disconnect between traits and fishing pressure might have multiple explanations. First, overall trait composition could be largely unaltered by harvest selection removing individual species over years or decades if communities have been shaped by environmental and evolutionary dynamics over centennial or geological scales (Briggs, 2003); second, it is also likely that the lack of trawl-level trait data obscures changing species traits, widespread declines in body size and earlier maturation being among the most significant (Heino *et al.* 2015). However, the trait spaces constructed here and based on fixed species trait values are not completely static as we do observe prominent shifts in yearly PCoA centroids in the northwest Atlantic Ocean (Supplementary Fig. 9.9; southeastern US shelf, Gulf of St. Lawrence, Newfoundland-Labrador). Here, on average, trait compositions after 2005 consistently contained fewer large species which grow and mature slowly and are within higher trophic levels. Without abundance-weighted traits across the dozens of trawl surveys used here— accounting for differences in effective sampling efforts and species-specific catchability (Walker *et al.* 2017)— we cannot yet disentangle the annual or cumulative effects of fishing pressure on these otherwise powerful community-wide datasets.

Trait-based macroecology is always constrained by trait availability and here, in a similar manner, some fish traits are missing in our work due to a lack of data availability, including diel activity patterns and social behaviors or predation defense mechanisms (Villéger *et al.* 2017). Each of these traits may capture different strategies adopted across environments or clades, however an analysis of the functional vulnerability of demersal fish in the northeast Atlantic found discrete traits added little explanatory power in trait spaces (Coulon *et al.* 2023). Crucially, in marine trawl data, we are largely missing intraspecific traits which differentially interact with the surrounding environment (Des Roches *et al.* 2018) and are a potential link to evolutionary history by incorporating adaptive constraints. Finally, to move towards a collection of traits that can disentangle trait-mediated variations in fitness across marine ecosystems, we need more traits describing dynamic rates (e.g., biological growth, reproduction, survival) throughout an organism's life history to infer specific ecosystem properties like ecological fitness (Streit & Bellwood, 2023) and demographic trade-offs (Laughlin *et al.* 2020).

4.6. Conclusions

Traits in demersal marine fishes are strongly related to environmental differences and appear to have converged substantially in overall composition across the Northern Hemisphere's continental shelf seas. The notable exception to this pattern is found in the northeastern shelves of the Pacific Ocean, where lineages are comparatively young and phylogenetic conservatism of traits appears more common. Amidst this important regional variation, there is high similarity in marine fish trait compositions, which suggests that ecosystem function-based conservation strategies could potentially complement species-based approaches. The processes shaping trait diversity in the oceans have broad implications for the ecosystem functioning of marine communities and conservation decisions focused on protecting trait-mediated functions or rare functional groups. Functional diversity has already been significantly altered across terrestrial and aquatic life from anthropogenic change (Carmona *et al.* 2021)— that environment is a primary driver of trait compositions in marine fishes implies environmental changes could significantly rearrange trait diversity in the future.

4.7. Data availability

All data are publicly available for scientific trawls in FISHGLOB (https://fishglob.sites.ucsc.edu/), species traits in PANGAEA ([b]Beukhof *et al.* 2019), environmental covariates through BIO-ORACLE v.3.0 (https://www.bio-oracle.org/), and phylogeny through the Fish Tree of Life (https://fishtreeoflife.org/).

4.8. Code availability

The code to conduct the main text analyses is available on request from the authors, and the details and code for hierarchical generalized additive models are available in Supplement S2 (MacNeil, 2025).

4.9. Author contributions

Liam MacNeil: Conceptualization (Lead); Data curation (Lead); Methodology (Lead); Formal analysis (Lead); Writing – original draft (Lead); **Matthew McLean:** Conceptualization (Supporting); Methodology (Supporting); Writing – review & editing (Supporting). **Derek Tittensor**: Writing – review & editing (Supporting). **Till Bayer:** Data curation (Supporting); Methodology (Supporting). **Thorsten Reusch:** Writing – review & editing (Supporting).

Marco Scotti: Conceptualization (Supporting); Data curation (Supporting); Funding acquisition (Lead); Methodology (Supporting); Project administration (Lead); Writing – review & editing (Supporting).

4.10. References

Assis, J., Fernández Bejarano, S. J., Salazar, V. W., Schepers, L., Gouvêa, L., Fragkopoulou, E., Leclercq, F., Vanhoorne, B., Tyberghein, L., Serrão, E. A., Verbruggen, H., & De Clerck, O. (2024). BIO-ORACLE v3.0. Pushing marine data layers to the CMIP6 Earth System Models of climate change research. *Global Ecology and Biogeography*, *33*(4), e13813. https://doi.org/10.1111/geb.13813

Baselga, A., & Orme, C. D. L. (2012). betapart: An R package for the study of beta diversity. *Methods in Ecology and Evolution*, *3*(5), 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

[a]Beukhof, E., Dencker, T., Pecuchet, L., & Lindegren, M. (2019). Spatio-temporal variation in marine fish traits reveals community-wide responses to environmental change. *Marine Ecology Progress Series*, *610*, 205–222. https://doi.org/10.3354/meps12826

[b]Beukhof, E., Dencker, T. S., Palomares, M. L. D., Maureaud, A. (2019). A trait collection of marine fish species from North Atlantic and Northeast Pacific continental shelf seas. *Pangaea, https://doi.org/10.1594/PANGAEA.900866.*

[c]Beukhof, E., Frelat, R., Pecuchet, L., Maureaud, A., Dencker, T. S., Sólmundsson, J., Punzón, A., Primicerio, R., Hidalgo, M., Möllmann, C., & Lindegren, M. (2019). Marine fish traits follow fast-slow continuum across oceans. *Scientific Reports*, *9*(1), 17878. https://doi.org/10.1038/s41598-019-53998-2

Bosch, N. E., Wernberg, T., Langlois, T. J., Smale, D. A., Moore, P. J., Franco, J. N., Thiriet, P., Feunteun, E., Ribeiro, C., Neves, P., Freitas, R., Filbee-Dexter, K., Norderhaug, K. M., Garcıa, A., Otero-Ferrer, F., Espino, F., Haroun, R., Lazzari, N., & Tuya, F. (2021). Niche and neutral assembly mechanisms contribute to latitudinal diversity gradients in reef fishes. *Journal of Biogeography*, *48*(11), 2683–2698. https://doi.org/10.1111/jbi.14237

Briggs, J. C. (2003). Marine centres of origin as evolutionary engines. *Journal of Biogeography*, *30*(1), 1–18. https://doi.org/10.1046/j.1365-2699.2003.00810.x

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S. M., Botta-Dukát, Z., Chytrý, M., Field, R., Jansen, F., Kattge, J., Pillar, V. D., Schrodt, F., Mahecha, M. D., Peet, R. K., Sandel, B., Van Bodegom, P., Altman, J., Alvarez-Dávila, E., … Jandt, U. (2018). Global trait–environment relationships of plant communities. *Nature Ecology & Evolution*, *2*(12), 1906–1917. https://doi.org/10.1038/s41559-018-0699-8

Cadotte, M., Albert, C. H., & Walker, S. C. (2013). The ecology of differences: Assessing community assembly with trait and evolutionary distances. *Ecology Letters*, *16*(10), 1234–1244. https://doi.org/10.1111/ele.12161

[a]Cadotte, M. W., & Tucker, C. M. (2017). Should Environmental Filtering be Abandoned? *Trends in Ecology & Evolution*, *32*(6), 429–437. https://doi.org/10.1016/j.tree.2017.03.004

[b]Cadotte, M. W., Davies, T. J., & Peres-Neto, P. R. (2017). Why phylogenies do not always predict ecological differences. *Ecological Monographs*, *87*(4), 535–551. https://doi.org/10.1002/ecm.1267

Carmona, C. P., Tamme, R., Pärtel, M., De Bello, F., Brosse, S., Capdevila, P., González-M., R., González-Suárez, M., Salguero-Gómez, R., Vásquez-Valderrama, M., & Toussaint, A. (2021). Erosion of global functional diversity across the tree of life. *Science Advances*, *7*(13), eabf2675. https://doi.org/10.1126/sciadv.abf2675

Chase, J. M. (2003). Community assembly: When should history matter? *Oecologia*, *136*(4), 489–498. https://doi.org/10.1007/s00442-003-1311-7

Chang, J., Rabosky, D. L., Smith, S. A., & Alfaro, M. E. (2019). An R package and online resource for macroevolutionary studies using the ray-finned fish tree of life. *Methods in Ecology and Evolution*, *10*(7), 1118–1124. https://doi.org/10.1111/2041-210X.13182

Coulon, N., Lindegren, M., Goberville, E., Toussaint, A., Receveur, A., & Auber, A. (2023). Threatened fish species in the Northeast Atlantic are functionally rare. *Global Ecology and Biogeography*, *32*(10), 1827–1845. https://doi.org/10.1111/geb.13731

Darimont, C. T., Carlson, S. M., Kinnison, M. T., Paquet, P. C., Reimchen, T. E., & Wilmers, C. C. (2009). Human predators outpace other agents of trait change in the wild. *Proceedings of the National Academy of Sciences*, *106*(3), 952–954. https://doi.org/10.1073/pnas.0809235106

De Bello, F., Botta-Dukát, Z., Lepš, J., & Fibich, P. (2021). Towards a more balanced combination of multiple traits when computing functional differences between species. *Methods in Ecology and Evolution*, *12*(3), 443–448. https://doi.org/10.1111/2041-210X.13537

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

Diamond, J.M. (1975). Assembly of species communities. In: M.L. Cody & J.M. Diamond (Eds.) Ecology and evolution of communities. Cambridge: *Harvard University Press*, Cambridge, Massachusetts, USA.

Floeter, S. R., Bender, M. G., Siqueira, A. C., & Cowman, P. F. (2018). Phylogenetic perspectives on reef fish functional traits. *Biological Reviews*, *93*(1), 131–151. https://doi.org/10.1111/brv.12336

Froese, R., Pauly, D. (2019). FishBase; www.fishbase.org.

Gerhold, P., Cahill, J. F., Winter, M., Bartish, I. V., & Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, *29*(5), 600–614. https://doi.org/10.1111/1365-2435.12425

Heino, M., Díaz Pauli, B., & Dieckmann, U. (2015). Fisheries-Induced Evolution. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 461–480. https://doi.org/10.1146/annurev-ecolsys-112414-054339

Hubbell, S. P. (2001). The unified neutral theory of biodiversity and biogeography. *Princeton University Press*, Princeton, New Jersey, USA.

IPBES. (2019). Global assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Brondízio, E. S., Settele, J., Díaz, S., Ngo, H. T. (eds). *IPBES secretariat*, Bonn, Germany, ISBN: 978-3-947851-20-1

Jönsson, B. F., & Watson, J. R. (2016). The timescales of global surface-ocean connectivity. *Nature Communications*, *7*(1), 11239. https://doi.org/10.1038/ncomms11239

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, *26*(11), 1463–1464. https://doi.org/10.1093/bioinformatics/btq166

Kraft, N. J. B., Cornwell, W. K., Webb, C. O., & Ackerly, D. D. (2007). Trait Evolution, Community Assembly, and the Phylogenetic Structure of Ecological Communities. *The American Naturalist*. *170*, 271–283.

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, *29*(5), 592–599. https://doi.org/10.1111/1365-2435.12345

Laughlin, D. C., Gremer, J. R., Adler, P. B., Mitchell, R. M., & Moore, M. M. (2020). The Net Effect of Functional Traits on Fitness. *Trends in Ecology & Evolution*, *35*(11), 1037–1047. https://doi.org/10.1016/j.tree.2020.07.010

Leibold, M. A., Govaert, L., Loeuille, N., De Meester, L., & Urban, M. C. (2022). Evolution and Community Assembly Across Spatial Scales. *Annual Review of Ecology, Evolution, and Systematics*, *53*(1), 299–326. https://doi.org/10.1146/annurev-ecolsys-102220-024934

Leprieur, F., Albouy, C., De Bortoli, J., Cowman, P. F., Bellwood, D. R., & Mouillot, D. (2012). Quantifying Phylogenetic Beta Diversity: Distinguishing between 'True' Turnover of Lineages and Phylogenetic Diversity Gradients. *PLoS ONE*, *7*(8), e42760. https://doi.org/10.1371/journal.pone.0042760

Leprieur, F., Descombes, P., Gaboriau, T., Cowman, P. F., Parravicini, V., Kulbicki, M., Melián, C. J., De Santana, C. N., Heine, C., Mouillot, D., Bellwood, D. R., & Pellissier, L. (2016). Plate tectonics drive tropical reef biodiversity dynamics. *Nature Communications*, *7*(1), 11461. https://doi.org/10.1038/ncomms11461

Losos, J. B. (2011). Convergence, adaptation, and constraint. *Evolution*, *65*(7), 1827–1840. https://doi.org/10.1111/j.1558-5646.2011.01289.x

Luiz, O. J., Madin, J. S., Robertson, D. R., Rocha, L. A., Wirtz, P., & Floeter, S. R. (2012). Ecological traits influencing range expansion across large oceanic dispersal barriers: Insights from tropical Atlantic reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1730), 1033–1040. https://doi.org/10.1098/rspb.2011.1525

MacArthur, R., & Levins, R. (1967). The Limiting Similarity, Convergence, and Divergence of Coexisting Species. *The American Naturalist*, *101*(921), 377–385. https://doi.org/10.1086/282505

MacNeil, L. (2025). Supplemental material for hierarchical GAMs. Zenodo. https://doi.org/10.5281/zenodo.15001873

Magneville, C., Loiseau, N., Albouy, C., Casajus, N., Claverie, T., Escalas, A., Leprieur, F., Maire, E., Mouillot, D., & Villéger, S. (2022). mFD: An R package to compute and illustrate the multiple facets of functional diversity. *Ecography*, *2022*(1), ecog.05904. https://doi.org/10.1111/ecog.05904

Maire, E., Grenouillet, G., Brosse, S., & Villéger, S. (2015). How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Global Ecology and Biogeography*, *24*(6), 728–740. https://doi.org/10.1111/geb.12299

Mazel, F., Wüest, R. O., Gueguen, M., Renaud, J., Ficetola, G. F., Lavergne, S., & Thuiller, W. (2017). The Geography of Ecological Niche Evolution in Mammals. *Current Biology*, *27*(9), 1369–1374. https://doi.org/10.1016/j.cub.2017.03.046

Mazel, F., Pennell, M. W., Cadotte, M. W., Diaz, S., Dalla Riva, G. V., Grenyer, R., Leprieur, F., Mooers, A. O., Mouillot, D., Tucker, C. M., & Pearse, W. D. (2018). Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature Communications*, *9*(1), 2888. https://doi.org/10.1038/s41467-018-05126-3

McGill, B., Enquist, B., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, *21*(4), 178–185. https://doi.org/10.1016/j.tree.2006.02.002

McLean, M., Stuart-Smith, R. D., Villéger, S., Auber, A., Edgar, G. J., MacNeil, M. A., Loiseau, N., Leprieur, F., & Mouillot, D. (2021). Trait similarity in reef fish faunas across the world's oceans. *Proceedings of the National Academy of Sciences*, *118*(12), e2012318118. https://doi.org/10.1073/pnas.2012318118

Messier, J., McGill, B. J., & Lechowicz, M. J. (2010). How do traits vary across ecological scales? A case for trait-based ecology. *Ecology Letters*, *13*(7), 838–848. https://doi.org/10.1111/j.1461-0248.2010.01476.x

Mittelbach, G. G., & Schemske, D. W. (2015). Ecological and evolutionary perspectives on community assembly. *Trends in Ecology & Evolution*, *30*(5), 241–247. https://doi.org/10.1016/j.tree.2015.02.008

Mora, C., Chittaro, P. M., Sale, P. F., Kritzer, J. P., & Ludsin, S. A. (2003). Patterns and processes in reef fish diversity. *Nature*, *421*(6926), 933–936. https://doi.org/10.1038/nature01393

Moretti, M., & Legg, C. (2009). Combining plant and animal traits to assess community functional responses to disturbance. *Ecography*, *32*(2), 299–309. https://doi.org/10.1111/j.1600-0587.2008.05524.x

Mouillot, D., Graham, N. A. J., Villéger, S., Mason, N. W. H., & Bellwood, D. R. (2013). A functional approach reveals community responses to disturbances. *Trends in Ecology & Evolution*, *28*(3), 167–177. https://doi.org/10.1016/j.tree.2012.10.004

Mouillot, D., Loiseau, N., Grenié, M., Algar, A. C., Allegra, M., Cadotte, M. W., Casajus, N., Denelle, P., Guéguen, M., Maire, A., Maitner, B., McGill, B. J., McLean, M., Mouquet, N., Munoz, F., Thuiller, W., Villéger, S., Violle, C., & Auber, A. (2021). The dimensionality and structure of species trait spaces. *Ecology Letters*, *24*(9), 1988–2009. https://doi.org/10.1111/ele.13778

Peltier, W. R. (1994). Ice Age Paleotopography. *Science*, *265*(5169), 195–201. https://doi.org/10.1126/science.265.5169.195

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876. https://doi.org/10.7717/peerj.6876

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., Alfaro, M. E., & Harmon, L. J. (2014). geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, *30*(15), 2216–2218. https://doi.org/10.1093/bioinformatics/btu181

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T. J., Coll, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, *559*(7714), 392–395. https://doi.org/10.1038/s41586-018-0273-1

R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rousseau, Y., Blanchard, J. L., Novaglio, C., Pinnell, K. A., Tittensor, D. P., Watson, R. A., & Ye, Y. (2024). A database of mapped global fishing activity 1950–2017. *Scientific Data*, *11*(1), 48. https://doi.org/10.1038/s41597-023-02824-6

Samuels, C.L., & Drake, J.A. (1997). Divergent perspectives on community convergence. *Trends in Ecology and Evolution*, *12*, 427–432.

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. *BioScience*, *57*(7), 573–583. https://doi.org/10.1641/B570707

Spasojevic, M. J., & Suding, K. N. (2012). Inferring community assembly mechanisms from functional diversity patterns: The importance of multiple assembly processes. *Journal of Ecology*, *100*(3), 652–661. https://doi.org/10.1111/j.1365-2745.2011.01945.x

Streit, R. P., & Bellwood, D. R. (2023). To harness traits for ecology, let's abandon 'functionality.' *Trends in Ecology & Evolution*, *38*(5), 402–411. https://doi.org/10.1016/j.tree.2022.11.009

Thorson, J. T., Maureaud, A. A., Frelat, R., Mérigot, B., Bigman, J. S., Friedman, S. T., Palomares, M. L. D., Pinsky, M. L., Price, S. A., & Wainwright, P. (2023). Identifying direct and indirect associations among traits by merging phylogenetic comparative methods and structural equation models. *Methods in Ecology and Evolution*, *14*(5), 1259–1275. https://doi.org/10.1111/2041-210X.14076

Villéger, S., Brosse, S., Mouchet, M., Mouillot, D., & Vanni, M. J. (2017). Functional ecology of fish: Current approaches and future challenges. *Aquatic Sciences*, *79*(4), 783–801. https://doi.org/10.1007/s00027-017-0546-z

Violle, C., Navas, M., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, *116*(5), 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x

Walker, N. D., Maxwell, D. L., Le Quesne, W. J. F., & Jennings, S. (2017). Estimating efficiency of survey and commercial trawl gears from comparisons of catch-ratios. *ICES Journal of Marine Science*, *74*(5), 1448–1457. https://doi.org/10.1093/icesjms/fsw250

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics*, *33*(1), 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

5. Synthesis & Outlook

Throughout this dissertation, I have tested the explanatory drivers of ecological data from a marginal shelf sea in the northeast Atlantic Ocean, to shelf seas across the Pacific and Atlantic Oceans. The logical flow of these chapters increased in the dimensionality of covariates from seasonal environmental predictors to network graphs of community composition, to representations of environment, evolutionary history, and fishing. While the first two chapters, respectively, focused on quantitative differences between species and inherent differences in the composition of species identities for demersal fish, the third chapter incorporates attributes of species phenotypes and life-histories using traits as proxies. This is also a stepwise expansion of ecological realism in the dissertation, first viewing only quantitative data on select demersal species, then partitioning community-wide data based on differing co-occurrence frequencies, and finally testing community-wide effects of environment, evolution, and human-driven pressures on the trait composition in marine fishes. Altogether, environmental differences emerged as the primary influence shaping species fundamental niches across seasons, environmental gradients appear to have left a signature in community-wide composition patterns, and broad environmental differences were key to explaining the compositions of fish traits. By comparing between oceans, Chapter 3 also revealed a novel difference in Atlantic and Pacific trait makeup: While trait convergence driven by environment appears widespread throughout Atlantic shelf seas, a unique signature of evolutionary conservatism in traits emerged in the Pacific. To reemphasize a discussion point from Chapter 3, while geographic differences in the relationships to environment are relevant and evolutionary history can be an important constraint on trait composition especially, the central role of environment in structuring marine fish communities in modern oceans implies future environmental and climate changes could significantly rearrange species and functional groups (Lotze *et al.* 2019).

Each chapter herein revealed ecological features and practical lessons for understanding marine demersal fish communities. The ecology viewpoint for Chapter 2 is that demersal fish biomass has changed substantially in the Baltic Sea, the distribution of these changes confirms declines and contractions in Atlantic cod, coincident with increased and expanded biomass of both flounder and plaice. The modelling viewpoint is that the most accurate predictions of demersal fish biomass are captured by deconstructing species-covariate relationships at yearly and seasonal resolutions. Although static species distribution models have been important for

model developments in ecology, they are increasingly questionable for reflecting ecological properties such as genetic diversity or population fitness (Lee-Yaw *et al.* 2022) and forecasting into this century (Briscoe *et al.* 2019). For Chapter 3, we show how commonly fixed endemicity thresholds miss important boundary zones and bioregions across continental shelf seas, and the degree of differentiation in species occurrence frequencies reveal strong latitudinal gradients in demersal fish communities. This finding is also practical, given that unweighted bipartite networks, which ignore changing species occurrence frequency, miss important boundary areas where demersal communities generally shift in composition (e.g., Baltic Sea, Cape Hatteras). Chapter 4 provided novel and direct evidence that environment has driven convergence in trait compositions of marine demersal fish and that the northeast Pacific shelves harbour community-wide trait patterns which have been constrained by evolutionary history. The comparison between the Atlantic and Pacific which was focused in more detail in Chapter 4 also indicates the importance of geological and evolutionary timescales in shaping the ecological (trait) components of marine fishes, as Pacific fish families appear to have diverged more recently than those in the Atlantic, and may not have had sufficient evolutionary time to accumulate adaptive changes which seem to have driven convergent similarity in Atlantic fish communities. An important practical lesson learned in Chapter 3 included the value in comparing observed and simulated (null/neutral) trait compositions to estimate the strength of convergence. This step clarified that statistically, the assembly of trait compositions in demersal fish lineages are not generated by purely random or neutral variations in demography during trait evolution.

There are potential links to identify between the chapters in this dissertation— primarily between Chapters 3 and 4 given they evaluate the same dataset and regions. First, the latitudinal gradient of bioregions along the North American coasts is also underpinned by a fast-slow continuum of species life-history traits. It is also likely that the large-scale divides in community composition are, at least in part, shaped by an environmental filter which affects species dispersal, growth, reproduction and ecological fitness (Cadotte & Tucker, 2017). As shown in Chapter 4, this process appears most pronounced on the Atlantic shelves and might explain why only two bioregions were detected across the vast range of the Pacific shelf region, as species communities appear to more widely coexist. This could also be linked to the younger ages of demersal fish communities in the Pacific which we show in Chapter 4, as processes shaping species niches are ongoing and species with contrasting niches may coexist for long periods of time (e.g., a century, Pearman *et al.* 2008). However, advances in molecular

techniques push us to be open to rapid evolutionary change which may alter species niches on timeframes (i.e., decades) which are classically thought to be caused by ecological drivers (Carroll *et al.* 2007; MacColl, 2011; Momigliano *et al.* 2017). In the European seas, a fast-slow life history continuum is also present, however the shelf bioregions largely divide by longitude, which represents a substantial gradient in average temperature and depth. Although we do not focus on individual life-history strategies across shelves, it has been found that this gradient is associated with shifts in species communities dominated by low generation times and those with high fecundity— i.e., an opportunistic–periodic gradient (Pecuchet *et al.* 2017). It seems combining large-scale biogeographic divides and finer spatial patterns in species and life-history compositions could be a productive path to understand the causes of major zonal differences in the semi-permeable oceans. To begin testing for differentiation in ecological communities at varying time scales, movement tracking data is the next logical link to scale down to the level of an individual, and to evaluate the effects of dispersal, life-history, and neutral or environmental filters on community distributions and makeup (Schlägel *et al.* 2020).

The patterns evaluated in this dissertation are manifold, and the analysis of ecological patterns is fundamentally the description of variation: The outstanding challenge facing scientists is to bridge gaps between pattern and process, which may be mismatched in spatial scale (Levin, 1992). This is an especially acute issue in ecology and in the life sciences more generally, where observational data has grown enormously in recent decades and classic descriptive methods— unrooted from a mechanistic framing, i.e., a data-generating process (Li *et al.* 2023)— are likely to fail in achieving generalized predictive power (MacNeil, 2008). However, a select number of processes, such as those shared between community ecology and the Modern Synthesis of Evolution (Selection, Drift, Speciation, Dispersal), can produce a panoply of ecological patterns without direct signatures of causation (Vellend*,* 2010). The concept of community assembly featured in Chapter 4 offers an explanatory framework for widespread functional patterns and which narrows the probable causes of trait convergence (i.e., modern environmental filtering or independent evolution of similarity). But attributing the precise cause of these patterns might be yet impractical as the role of historical contingencies which filter different marine lineages (e.g., founder events) are poorly resolved across oceans.

A useful example for practically applying causal modelling in the natural sciences has been borrowed from computer science and the work of Judea Pearl: The structural causal model framework imports causal thinking into variable selection (applicable to any common

statistical model) to ultimately constrain spurious parameter estimates (Pearl, 2009). Briefly, it is based on variable selection criteria which block non-causal pathways, guided by direct acyclic graphs, that indirectly obscure the relationship between a given causal factor and the underlying response variable (X → Y; backdoor selection). This is a form of bias correction which has been explored in a mock ecological dataset to show that the effects of forestry on a simulated terrestrial species are significantly underestimated without causal (backdoor) selection (Arif & MacNeil, 2022). Causal inference has been adopted more rigorously in earth sciences than in life sciences, for example to disentangle effects of global teleconnections in climate phenomena on the path of the Jet Stream in the Northern Hemisphere (Runge *et al.* 2019). This is surely due to a deeper mathematical, and mechanistic understanding of physical phenomena (e.g., Navier-Stokes equation in describing fluid dynamics) compared to biological systems. But for ecology, these methods are still limited, and with the expanding set of potential environmental or biological covariates now available in the natural sciences it seems at first impractical to scale covariate selection in this way— although some automated selection procedures are improving the accessibility of these methods (Textor *et al.* 2017). So, the problem remains, building scalable causal modeling structures in ecology is an ongoing effort.

I will dedicate a final argument about causation in ecology. A substantial challenge has arisen to the genetic (and genomic) blueprint vision which peaked after the highly publicized draft of the human genome project in 2001. It is best shown in Phillip Ball's recent book (Ball, 2024), which argues that the limited predictability of almost *everything* in nature (from individuals to communities) derives from genes as the key explanatory variable. Genes are remarkable molecular artifacts (palimsests, even; Teasdale *et al.* 2017), and essential material for constructing life. But the gap between genes (coding or non-coding) and proteins to traits and behaviors is fraught, and it is these latter factors which are likely to produce agency (Ball, 2024) and to be selected on evolutionary timescales. The physiological and behavioral responses of an organism to its environment might imprint itself on genes and inject doubt into classical views of genes as a one-way causal blueprint (Noble, 2024). This has been something of a quiet revolution, and it is no Lamarckian-redux, but a refocusing of focal variables for ecological models. It is a forceful argument for the use of higher-order variables such as physiological and behavioural processes to approach causation and robust predictability where it is possible. A realignment away from baseline units represented by genes, genomes, or individual morphological traits could help bring about a genuine paradigm shift (Kuhn, 1962) in ecology.

Unravelling patterns in nature depends on how things occur and change in space and time. As a discipline, biogeography contains invaluable tools and a rich literature on the spatiotemporal dynamics of ecological communities, dead or alive. It has been a wellspring for ecological theory and has been pushed into the modern era of statistical ecology. This dissertation has aimed to recover biogeographic patterns and predictors on regional and ocean-basin scales, using fishes as a model group which have been studied scientifically for centuries. Throughout, I have shown that individual species, and larger communities, have been intricately shaped by their environments despite tremendous regional variation. As global datasets continue to amass, biogeography can support us in patching together the mosaic of life on this Ocean planet, towards understanding its abundant causes and preserving its diversity.

5.1 References

Arif, S., & MacNeil, M. A. (2022). Predictive models aren't for causal inference. *Ecology Letters*, *25*(8), 1741–1745. https://doi.org/10.1111/ele.14033

Ball, P. (2024). How Life Works: A User's Guide to the New Biology*, Macmillan*. 560 pages.

Briscoe, N. J., Elith, J., Salguero-Gómez, R., Lahoz-Monfort, J. J., Camac, J. S., Giljohann, K. M., Holden, M. H., Hradsky, B. A., Kearney, M. R., McMahon, S. M., Phillips, B. L., Regan, T. J., Rhodes, J. R., Vesk, P. A., Wintle, B. A., Yen, J. D. L., & Guillera-Arroita, G. (2019). Forecasting species range dynamics with process-explicit models: Matching methods to applications. *Ecology Letters*, *22*(11), 1940–1956. https://doi.org/10.1111/ele.13348

Cadotte, M. W., & Tucker, C. M. (2017). Should Environmental Filtering be Abandoned? *Trends in Ecology & Evolution*, *32*(6), 429–437. https://doi.org/10.1016/j.tree.2017.03.004

Carroll, S. P., Hendry, A. P., Reznick, D. N., & Fox, C. W. (2007). Evolution on ecological time-scales. *Functional Ecology*, *21*(3), 387–393. https://doi.org/10.1111/j.1365-2435.2007.01289.x

Kuhn, T. S. (1962). The structure of scientific revolutions. *University of Chicago Press*. 222 pages.

Lee-Yaw, A. J., L. McCune, J., Pironon, S., & N. Sheth, S. (2022). Species distribution models rarely predict the biology of real populations. *Ecography*, *2022*(6), e05877. https://doi.org/10.1111/ecog.05877

Levin, S. A. (1992). The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology*, *73*(6), 1943–1967. https://doi.org/10.2307/1941447

Li, F., Ding, P., & Mealli, F. (2023). Bayesian causal inference: A critical review. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *381*(2247), 20220153. https://doi.org/10.1098/rsta.2022.0153

Lotze, H. K., Tittensor, D. P., Bryndum-Buchholz, A., Eddy, T. D., Cheung, W. W. L., Galbraith, E. D., Barange, M., Barrier, N., Bianchi, D., Blanchard, J. L., Bopp, L., Büchner, M., Bulman, C. M., Carozza, D. A., Christensen, V., Coll, M., Dunne, J. P., Fulton, E. A., Jennings, S., … Worm, B. (2019). Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. *Proceedings of the National Academy of Sciences*, *116*(26), 12907–12912. https://doi.org/10.1073/pnas.1900194116

MacColl, A. D. C. (2011). The ecological causes of evolution. *Trends in Ecology & Evolution*, *26*(10), 514–522. https://doi.org/10.1016/j.tree.2011.06.009

MacNeil, M. A. (2008). Making empirical progress in observational ecology. *Environmental Conservation*, *35*(3), 193–196. https://doi.org/10.1017/S0376892908004888

Momigliano, P., Jokinen, H., Fraimout, A., Florin, A.-B., Norkko, A., & Merilä, J. (2017). Extraordinarily rapid speciation in a marine fish. *Proceedings of the National Academy of Sciences*, *114*(23), 6074–6079. https://doi.org/10.1073/pnas.1615109114

Noble, D. (2024). Genes are not the blueprint for life, *Nature,* 626, 254-255.

Pearl, J. (2009) Causality: models, reasoning and inference, 2nd edition. UK: Cambridge University Press; Cambridge.

Pearman, P. B., Guisan, A., Broennimann, O., & Randin, C. F. (2008). Niche dynamics in space and time. *Trends in Ecology & Evolution*, *23*(3), 149–158. https://doi.org/10.1016/j.tree.2007.11.005

Pecuchet, L., Lindegren, M., Hidalgo, M., Delgado, M., Esteban, A., Fock, H. O., Gil De Sola, L., Punzón, A., Sólmundsson, J., & Payne, M. R. (2017). From traits to life-history strategies: Deconstructing fish community composition across European seas. *Global Ecology and Biogeography*, *26*(7), 812–822. https://doi.org/10.1111/geb.12587

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., Van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., … Zscheischler, J. (2019). Inferring causation from time series in Earth system sciences. *Nature Communications*, *10*(1), 2553. https://doi.org/10.1038/s41467-019-10105-3

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

Teasdale, M. D., Fiddyment, S., Vnouček, J., Mattiangeli, V., Speller, C., Binois, A., Carver, M., Dand, C., Newfield, T. P., Webb, C. C., Bradley, D. G., & Collins, M. J. (2017). The York Gospels: A 1000-year biological palimpsest. *Royal Society Open Science*, *4*(10), 170988. https://doi.org/10.1098/rsos.170988

Textor, J., Van Der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. H. (2017). Robust causal inference using directed acyclic graphs: The R package 'dagitty.' *International Journal of Epidemiology*, dyw341. https://doi.org/10.1093/ije/dyw341

Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, *85*(2), 183–206. https://doi.org/10.1086/652373

6. Appendix S1: Spatial change of dominant Baltic Sea demersal fish across two decades

Supplement for Chapter 2.

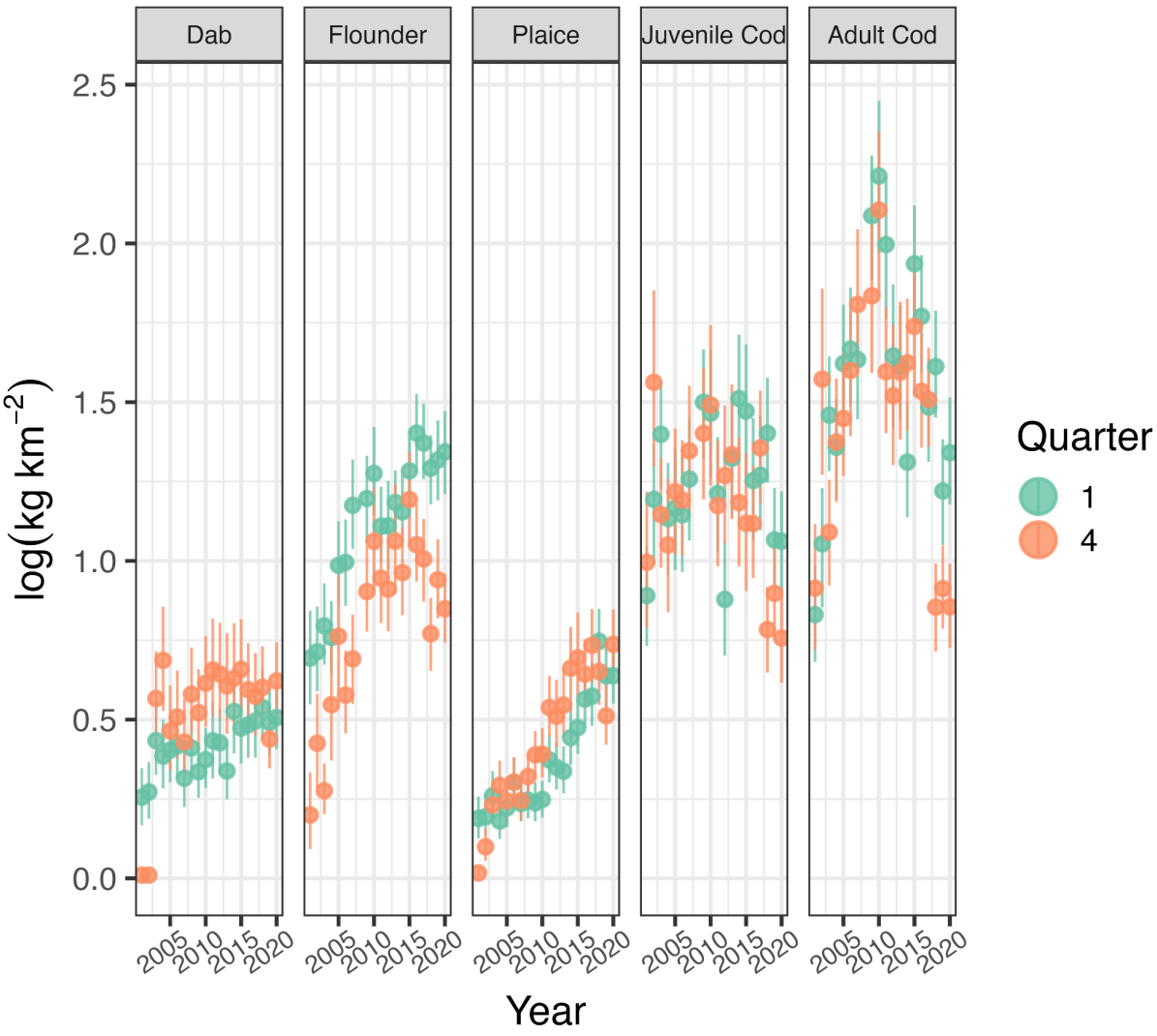


Figure 6.1 Trawl biomass densities throughout the 20-year time series as bootstrap means with 95% confidence intervals, colored by yearly quarter and faceted by species.

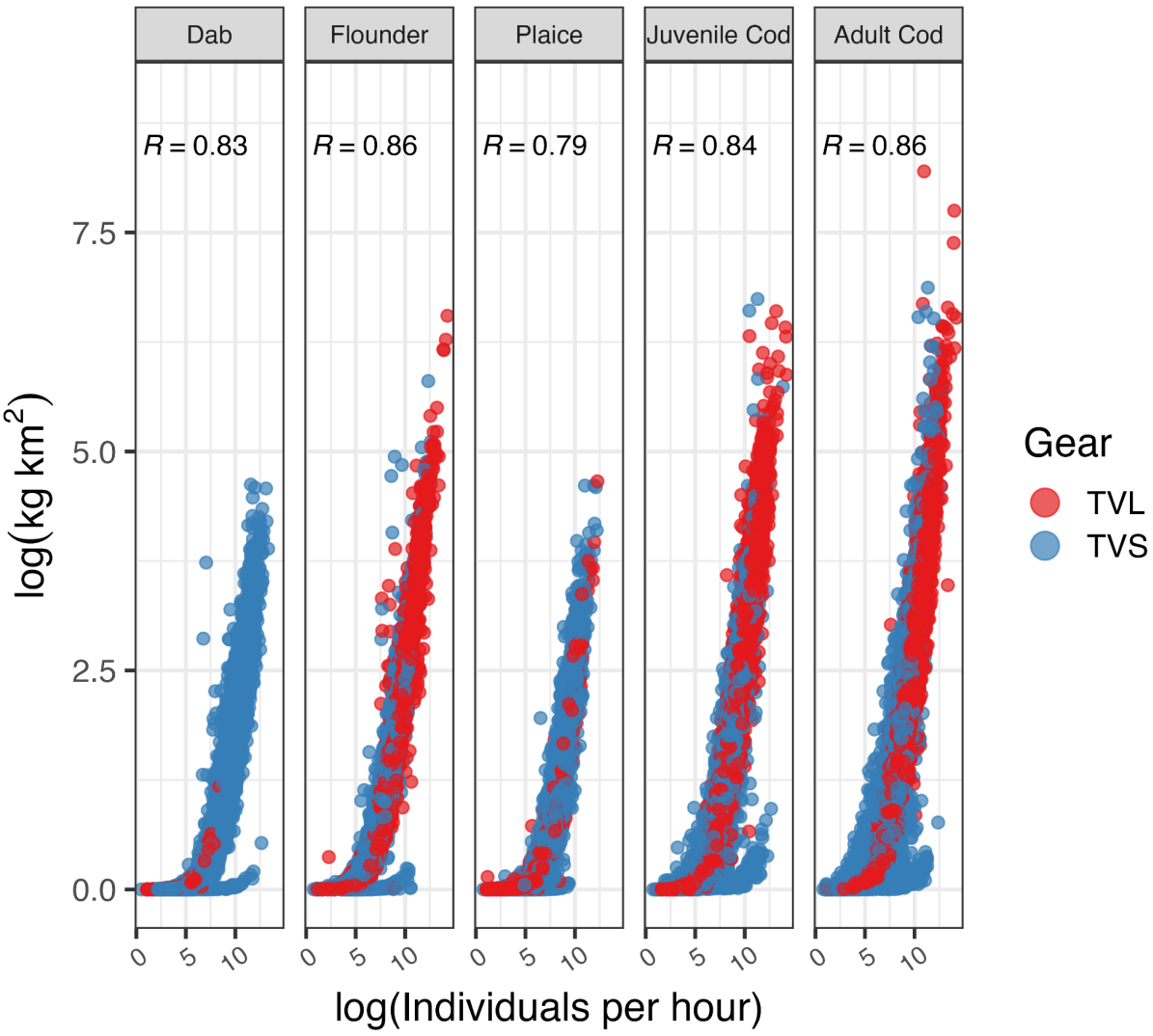


Figure 6.2 Biomass densities (kg km-2) are strongly correlated (R = Pearson's correlation coefficient) to a conventional catch-per-unit-effort metric based on abundance counts. Trawls are colored by gear type and faceted by species.

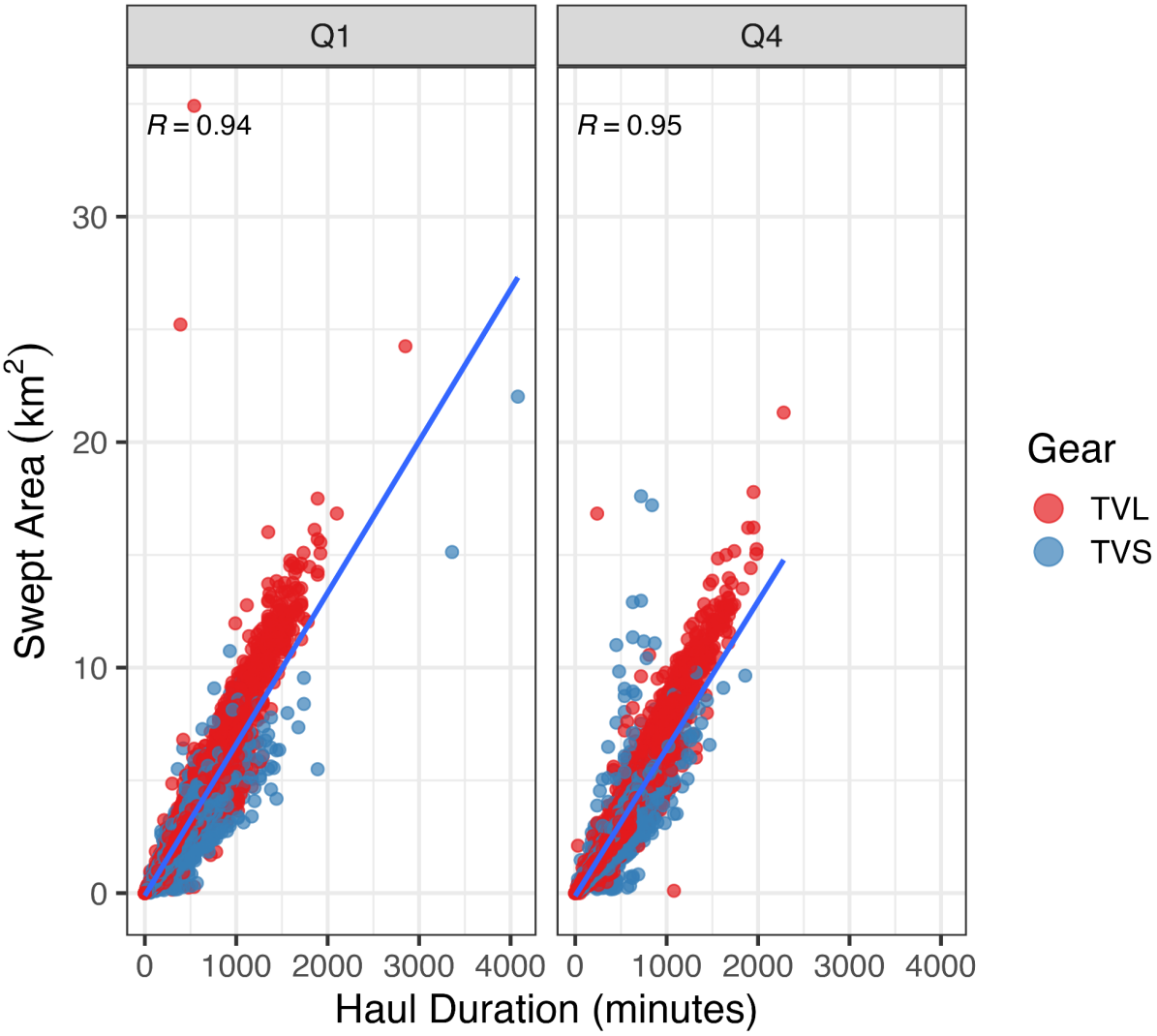


Figure 6.3 Swept area (km2) representing total covered area during trawl is strongly correlated (R = Pearson's correlation coefficient) to haul duration (summed by station by year). Trawls are colored by gear type and faceted by quarter.

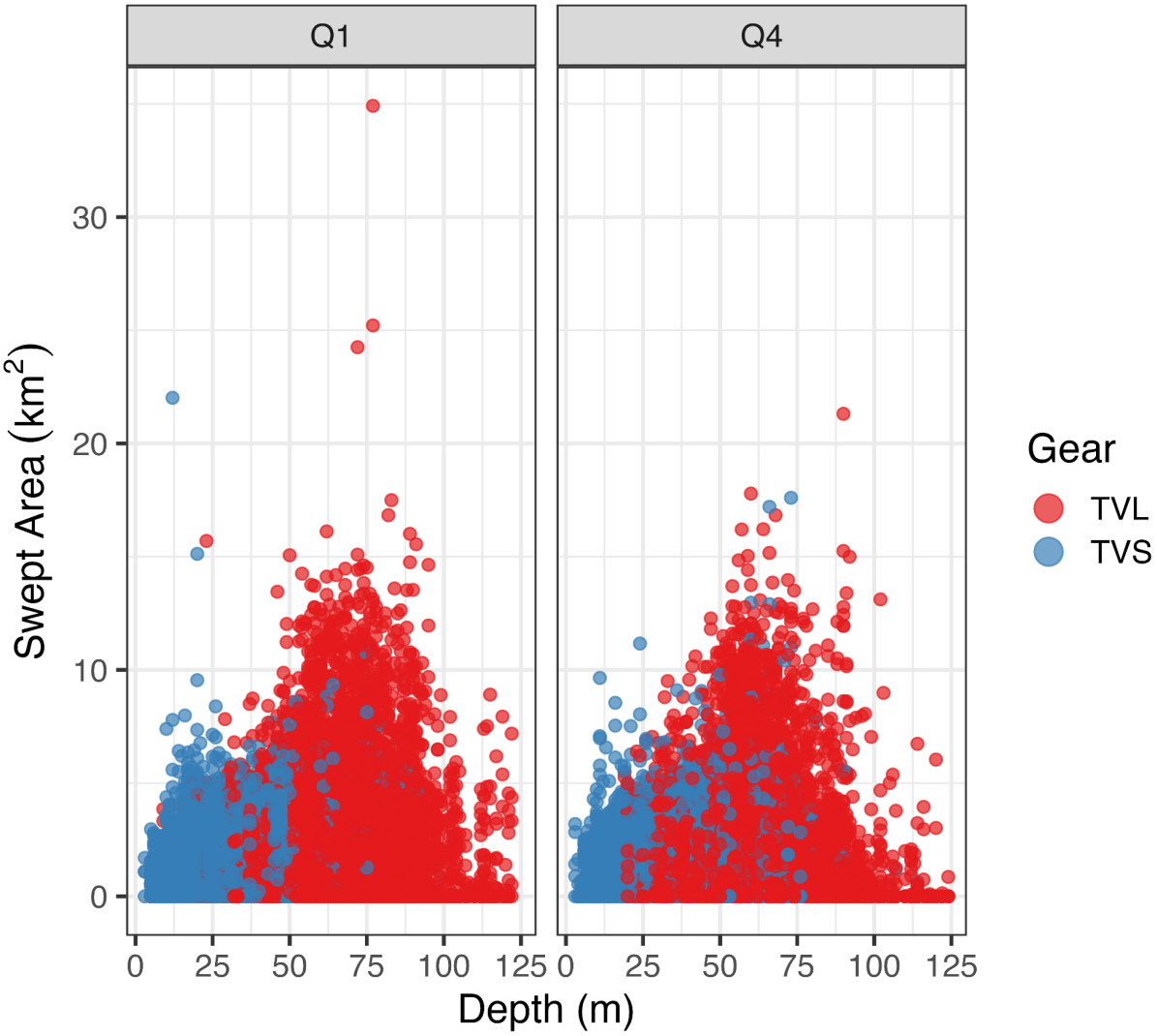


Figure 6.4 Swept area (km2) representing total covered area is distributed across depths, illustrating the partial separation of smaller TVS gear types in more shallow environments in the western Baltic Sea, and the larger TVL gear types used elsewhere in deeper regions. Trawls are colored by gear type and faceted by quarters.

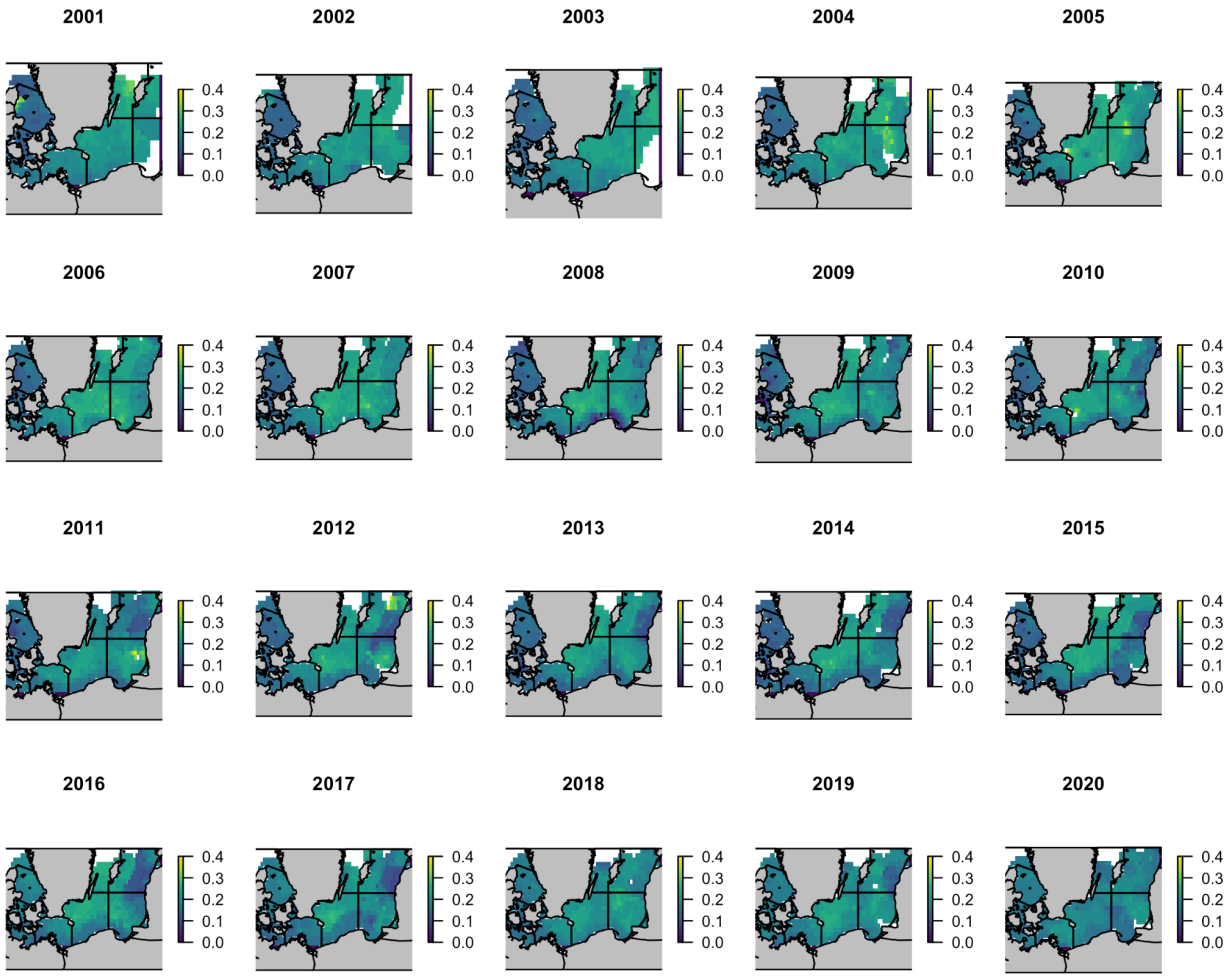


Figure 6.5 Door-based swept area (km2) corresponding to an estimate of trawl effort described in Berg et al. (2019). Interpolated (0.2° × 0.2°) by inverse distance weighting constrained by maximum distance of 50 km and an inverse distance power (idp = 2) prioritizing nearby points. Swept area was relatively stable throughout the 20-year time series.

Table 6.1 Diagnostic checks for each model version (I, II, III) including the allowed degrees of freedom set by basis function number (k), the effective degrees of freedom (EDF) used for each term, and a p-value estimate based on the simulation in the gam.check function in mgcv (Wood, 2017). Significant p-values (<0.05) suggest potentially inadequate basis function number for an individual smoother and motivated further checks of model residuals in Figures 6.8-6.9.

Smooth terms	# of basis functions (k)	EDF	F-statistic	p-value
Model-I				
ti(lon, lat, year)				

Dab	124	66.8	66.8	<0.001
Flounder	124	77.2	16.5	<0.001
Plaice	124	89	62.8	<0.001
(Juvenile) Cod	124	83.3	31.9	<0.001
(Adult) Cod	124	79.7	34.6	<0.001
te(year, depth)				
Dab	24	14.5	11.2	<0.005
Flounder	24	17.5	7.9	<0.005
Plaice	24	20.7	18.8	<0.015
(Juvenile) Cod	24	17.7	31.4	<0.001
(Adult) Cod	24	16.5	22.2	<0.001
Model II				
ti(lon, lat, year)				
Dab	124	64.3	35.5	<0.001
Flounder	124	80.9	21.5	<0.001
Plaice	124	90.4	48.5	<0.001
(Juvenile) Cod	124	88.6	26.1	<0.001
(Adult) Cod	124	84.9	29.4	<0.001
s(depth, species)	225	63.9	5.9	0.3
s(temperature, species)	225	28.1	1.9	0.04
s(oxygen, species)	225	59.4	4.6	0.005
s(salinity, species)	50	36.7	11.3	<0.001
Model III				
ti(lon, lat, year)				
Dab	124	64.7	36.6	<0.001
Flounder	124	83.5	22.9	<0.001
Plaice	124	90.9	50.6	<0.001
(Juvenile) Cod	124	89.6	27.3	<0.001
(Adult) Cod	124	87.1	30.9	<0.001
s(depth, species)				
Winter	50	29.7	20.8	0.8
Autumn	50	43.6	29.3	0.8
s(temperature, species)				
Winter	50	12.2	0.9	0.02
Autumn	50	11.4	1.2	0.03
s(oxygen,				

species)				
Winter	50	34.1	9.2	<0.001
Autumn	50	26.1	9.9	<0.001
s(salinity, species)				
Winter	50	35.4	7.1	0.1
Autumn	50	30.7	7.5	0.09

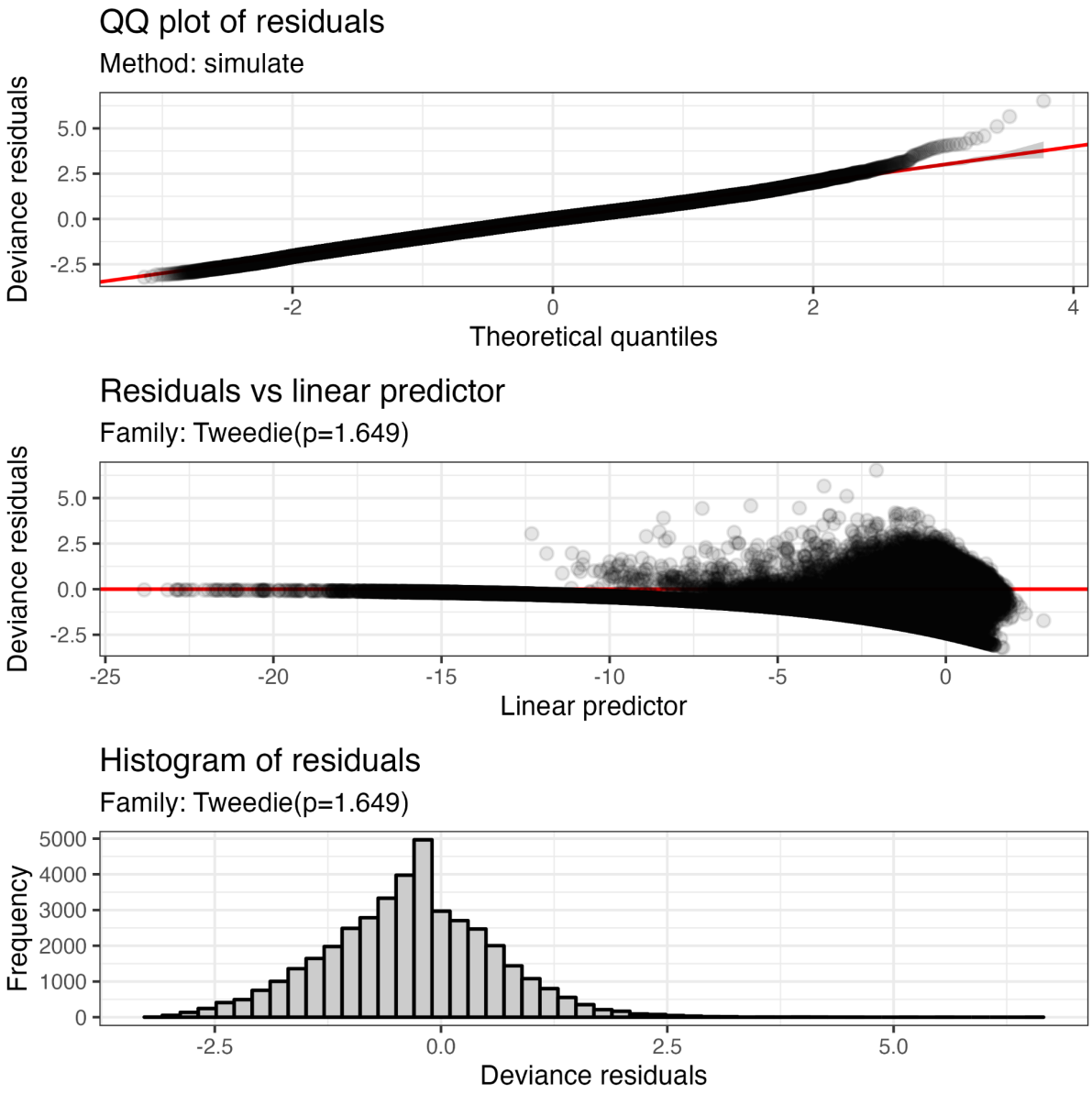


Figure 6.6 Diagnostic appraisal of HGAM version-III simulating QQ plots including normal standard errors, comparing the residuals to linear predictor where the zeros are the lower bound of deviance in the residuals, and finally a histogram of the HGAM residuals. An additional diagnostic plot comparing the response variable versus fitted values is shown below in Figure 6.7 including species-specific differences across 10-fold cross validation.

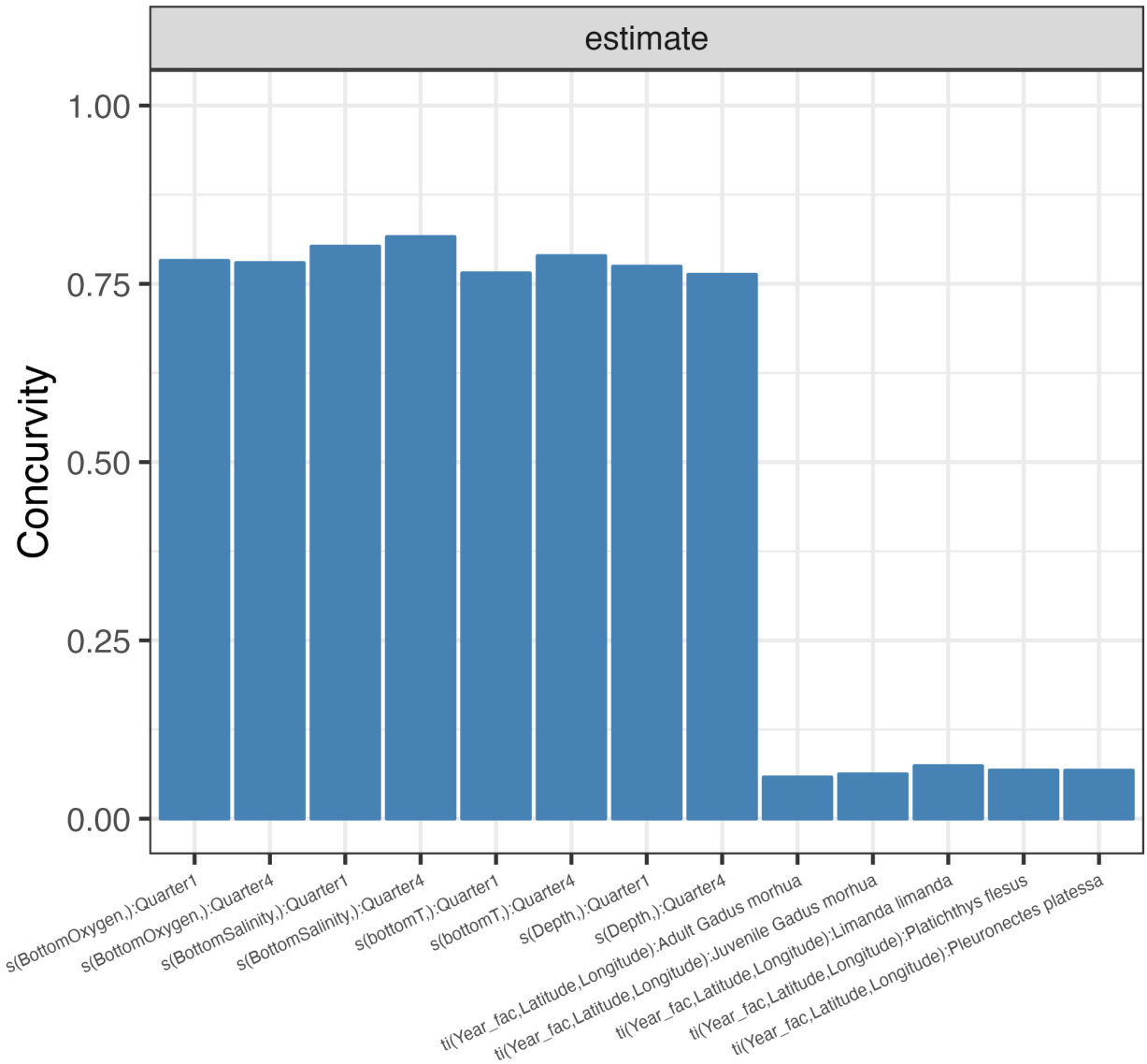


Figure 6.7 HGAM version-III model concurvity for smooth terms (Simpson, 2018), where a value of 1 indicates complete non identifiability (analogous to maximal collinearity). No consensus threshold exists for excluding terms based on concurvity, but we note that all values for all models are beneath 0.8 (e.g., Leonardi et al. 2022).

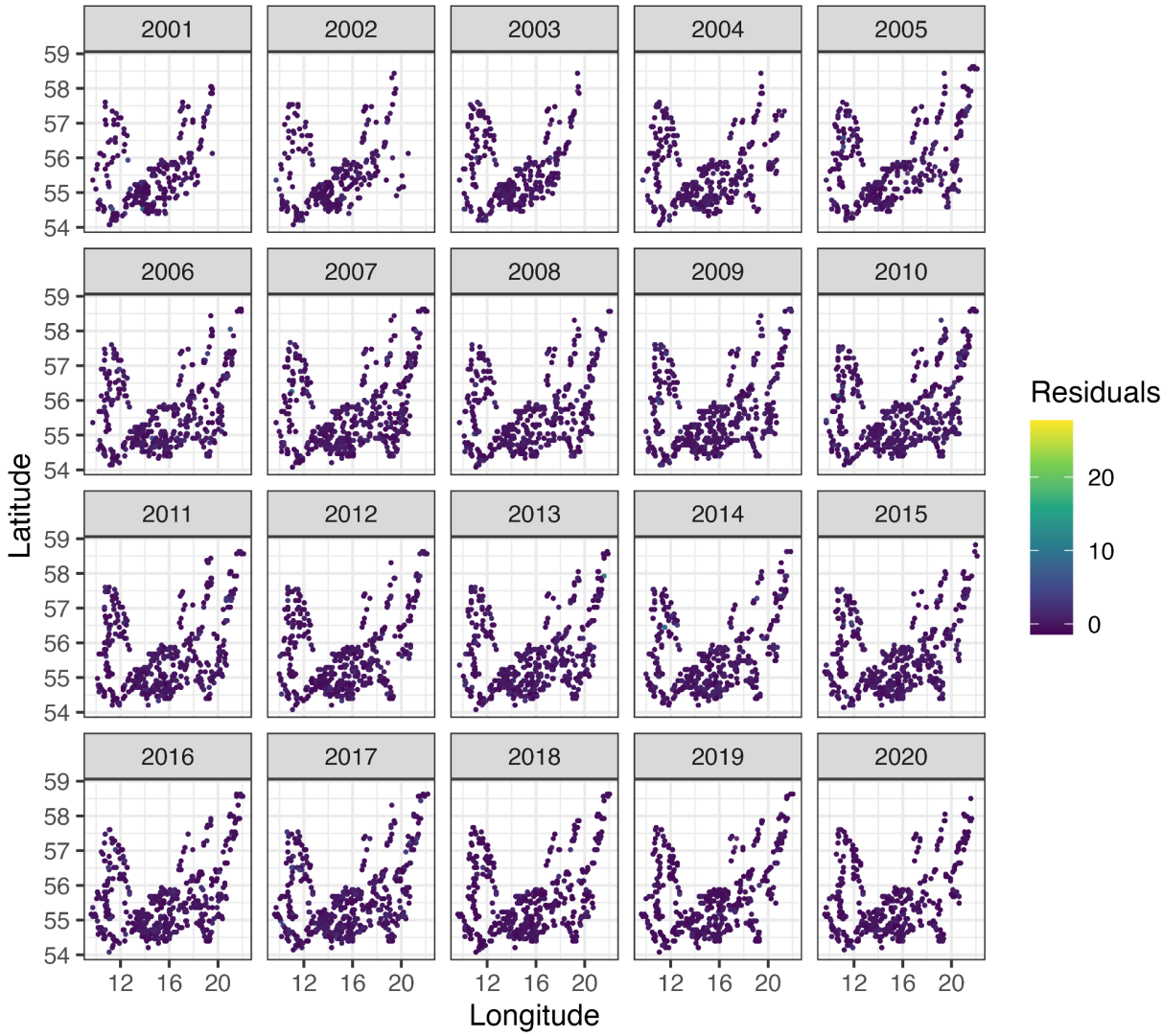


Figure 6.8 Model residuals (HGAM version-III) plotted geographically (Longitude versus Latitude) and by year, with residual magnitude filling point color. No clear model residual patterns against space or time are evident.

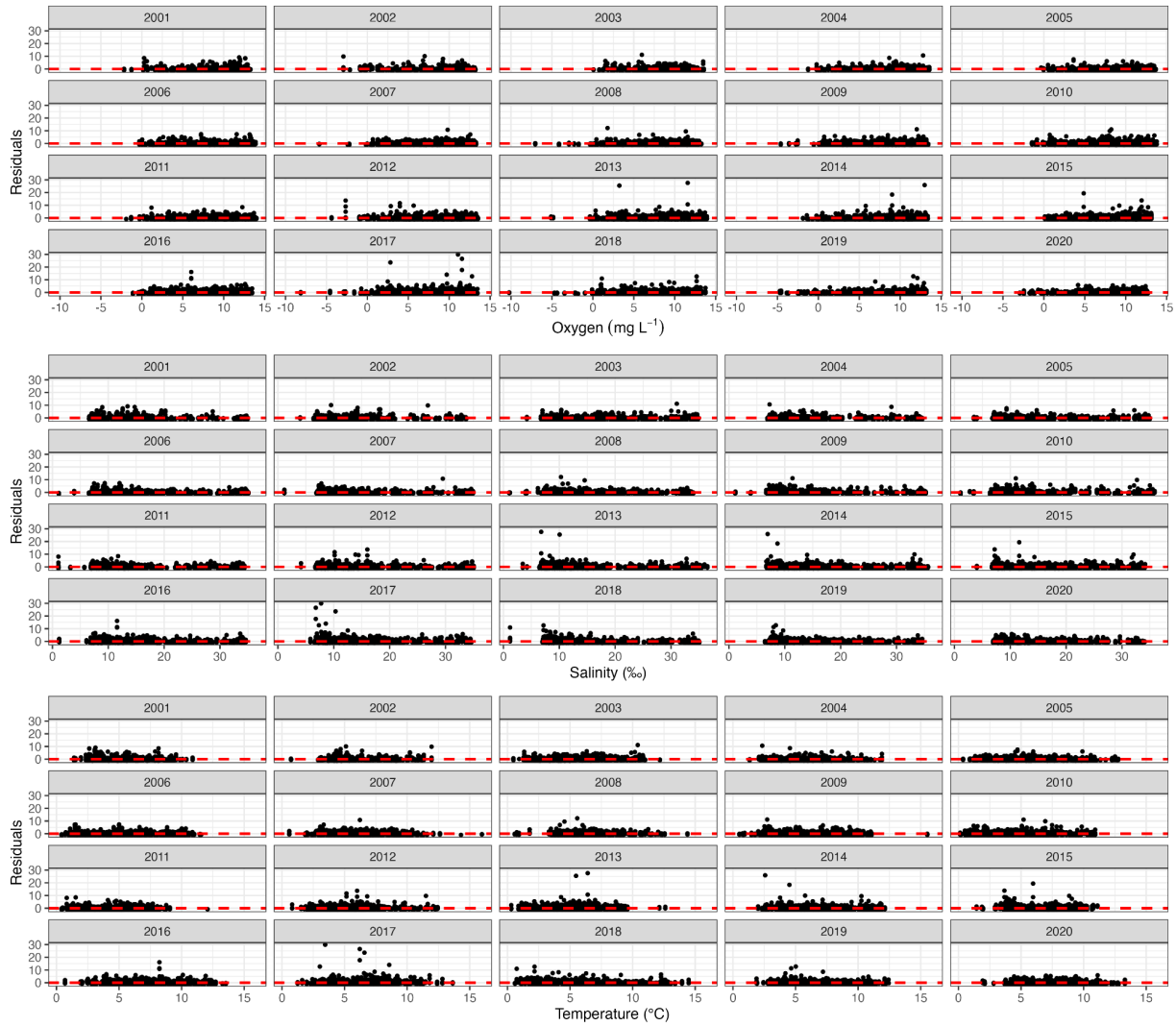


Figure 6.9 Model residuals (HGAM version-III) plotted against environmental covariates and by year. The red dotted line is set at zero (y = 0) and altogether no dependencies between model residuals and abiotic covariates are apparent.

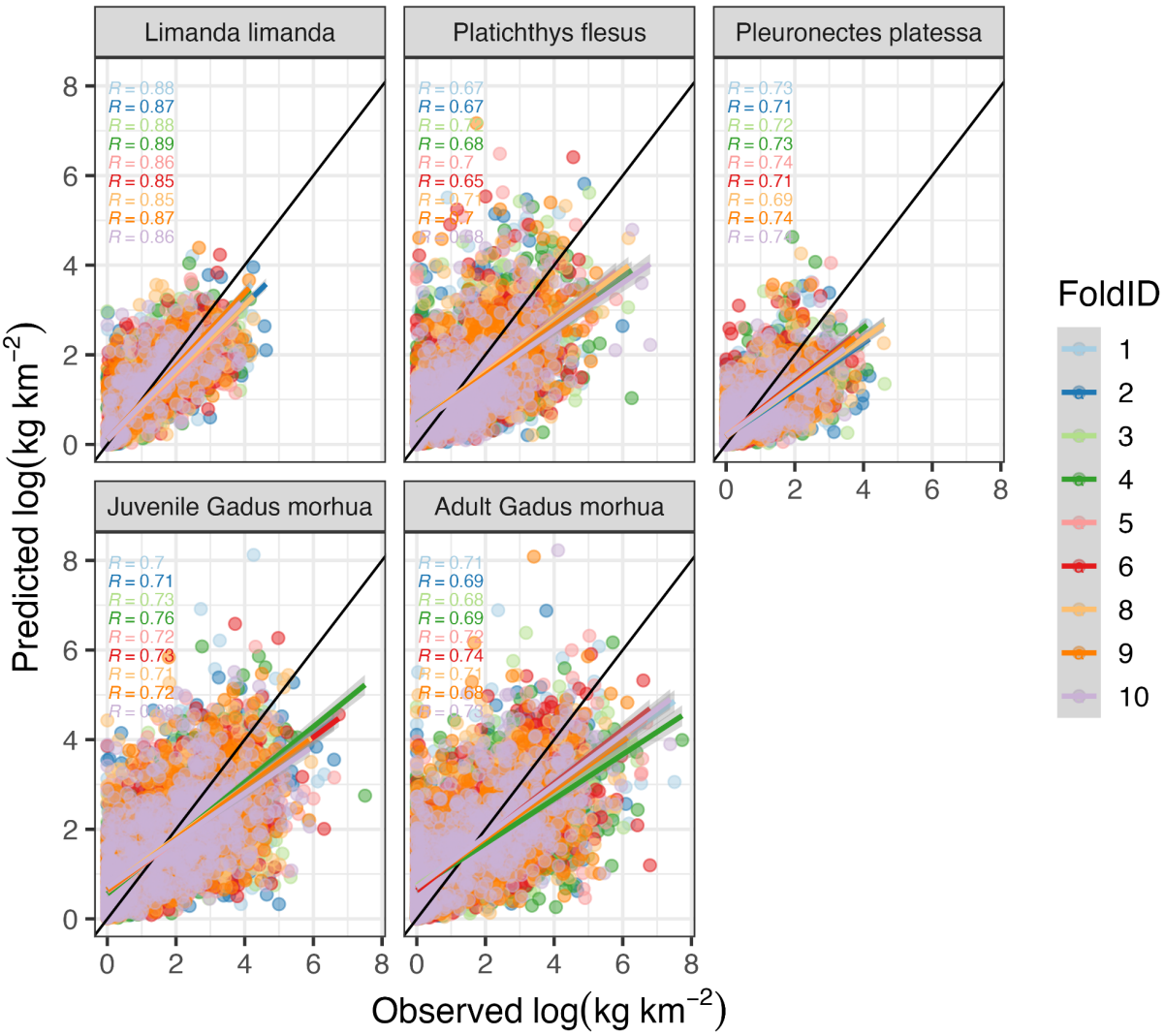


Figure 6.10 Diagnostic appraisal of HGAM version-III for predictive performance in 10-fold cross validation, comparing observed versus fitted biomass on the natural log scale. Points and linear corrections are colored by k-fold, with Pearson's correlation coefficient (R) recorded in each panel for each k-fold. The 9th fold model run failed to converge and was excluded. The solid black line represents a perfect linear relationship for comparison. The average linear correlations (with standard deviations) represent our measure of discrimination recorded in Table 6.2. The two additional measures, precision and accuracy, also derive from these diagnostics with point cloud dispersion capturing precision and the absolute error in prediction values capturing accuracy (Table 6.2).

Table 6.2 HGAM performance by species group and model version (I, II, III) against regression metrics for predictive performance (discrimination, precision, accuracy) and parsimony (AIC). The optimal regression metric values (O_v) are as follows: Discrimination (O_v=1), Precision (O_v=1), Accuracy (O_v=0).

Model	Group	Discrimination (Correlation)	Precision (Dispersion)	Accuracy (MAE)	Deviance explained	AIC
Version-I (Geographic)	Dab	0.83 ± 0.016	0.75 ± 0.042	0.23 ± 0.015	61.0%	54218
	Flounder	0.52 ± 0.03	0.31 ± 0.019	0.70 ± 0.02		
	Plaice	0.68 ± 0.028	0.53 ± 0.064	0.30 ± 0.012		
	(Juvenile) Cod	0.56 ± 0.022	0.41 ± 0.029	0.88 ± 0.022		
	(Adult) Cod	0.68 ± 0.016	0.49 ± 0.027	0.85 ± 0.024		
Version-II (+ Abiotic)	Dab	0.86 ± 0.011	0.81 ± 0.049	0.20 ± 0.014	62.1%	53107
	Flounder	0.59 ± 0.021	0.48 ± 0.031	0.65 ± 0.013		
	Plaice	0.71 ± 0.028	0.60 ± 0.041	0.28 ± 0.010		
	(Juvenile) Cod	0.70 ± 0.017	0.59 ± 0.030	0.75 ± 0.025		
	(Adult) Cod	0.70 ± 0.013	0.56 ± 0.018	0.78 ± 0.019		
Version-III (+ Seasonal Abiotic)	Dab	0.87 ± 0.013	0.83 ± 0.056	0.20 ± 0.010	65.2%	50211
	Flounder	0.68 ± 0.026	0.67 ± 0.015	0.56 ± 0.013		
	Plaice	0.74 ± 0.013	0.63 ± 0.055	0.28 ± 0.013		
	(Juvenile) Cod	0.73 ± 0.0224	0.62 ± 0.035	0.74 ± 0.024		
	(Adult) Cod	0.72 ± 0.019	0.61 ± 0.032	0.78 ± 0.021		

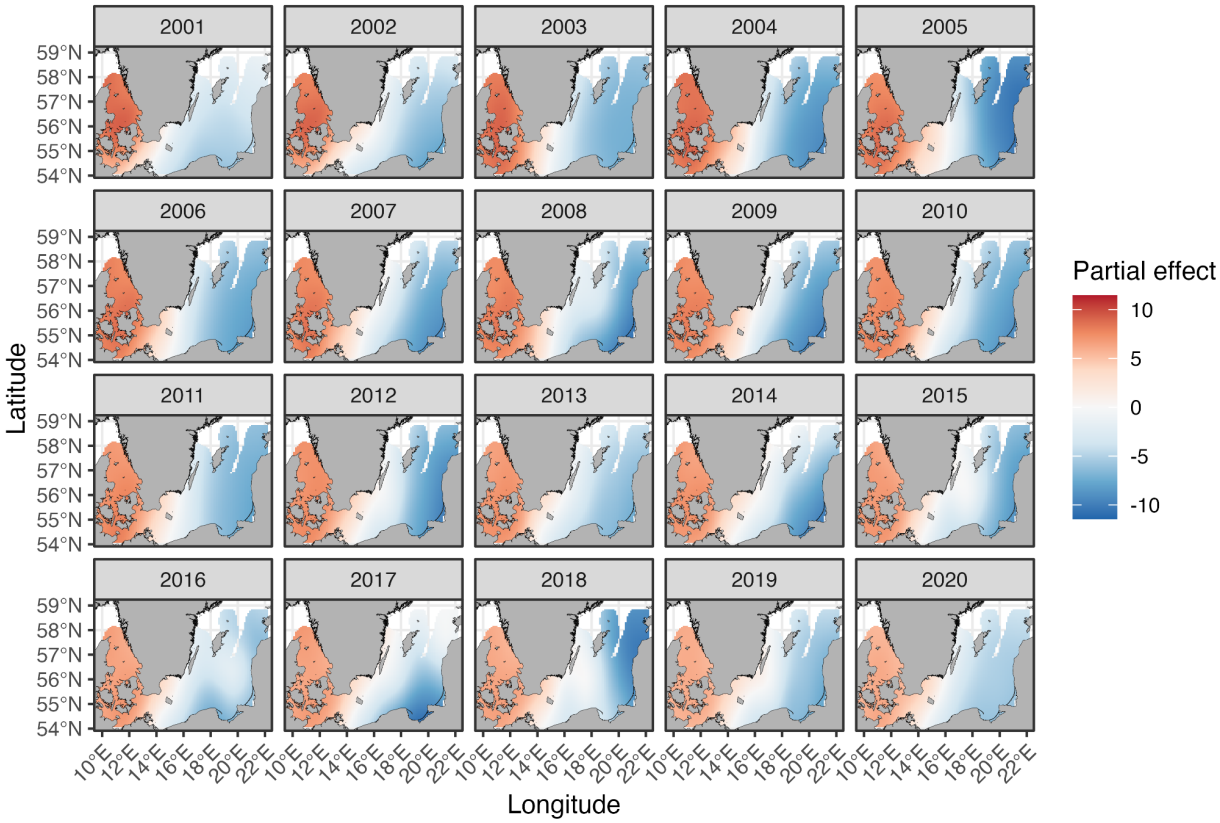


Figure 6.11 Common dab (*Limanda limanda*) spatiotemporal partial dependence (2-d smooth; Simpson, 2018) for tensor smooth product (te(lon, lat, year, by = species)) in HGAM version-III. The 2-D smooth maps are cropped by the 0.5° buffer polygon to avoid extrapolation into unsampled areas.

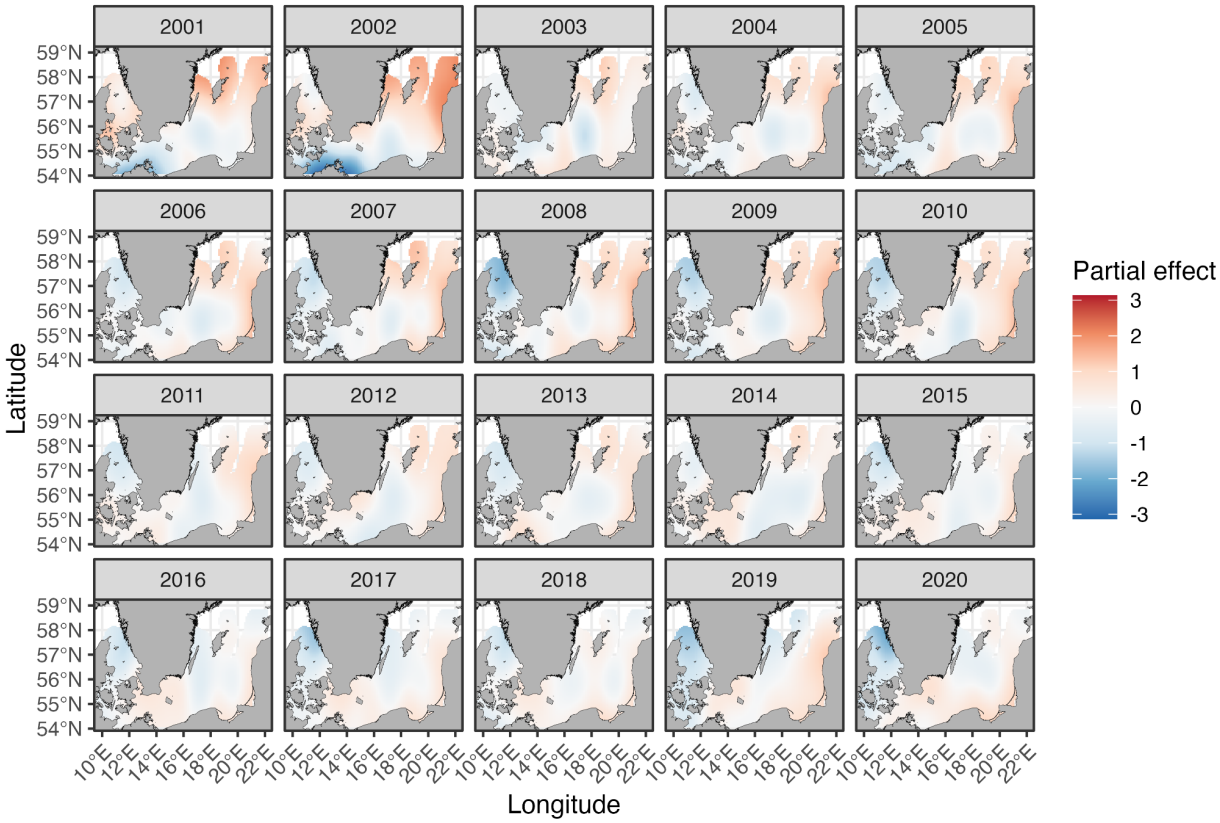


Figure 6.12 European flounder (*Platichthys flesus*) spatiotemporal partial dependence (2-d smooth; Simpson, 2018) for tensor smooth product (te(lon, lat, year, by = species)) in HGAM version-III. The 2-D smooth maps are cropped by the 0.5° buffer polygon to avoid extrapolation into unsampled areas.

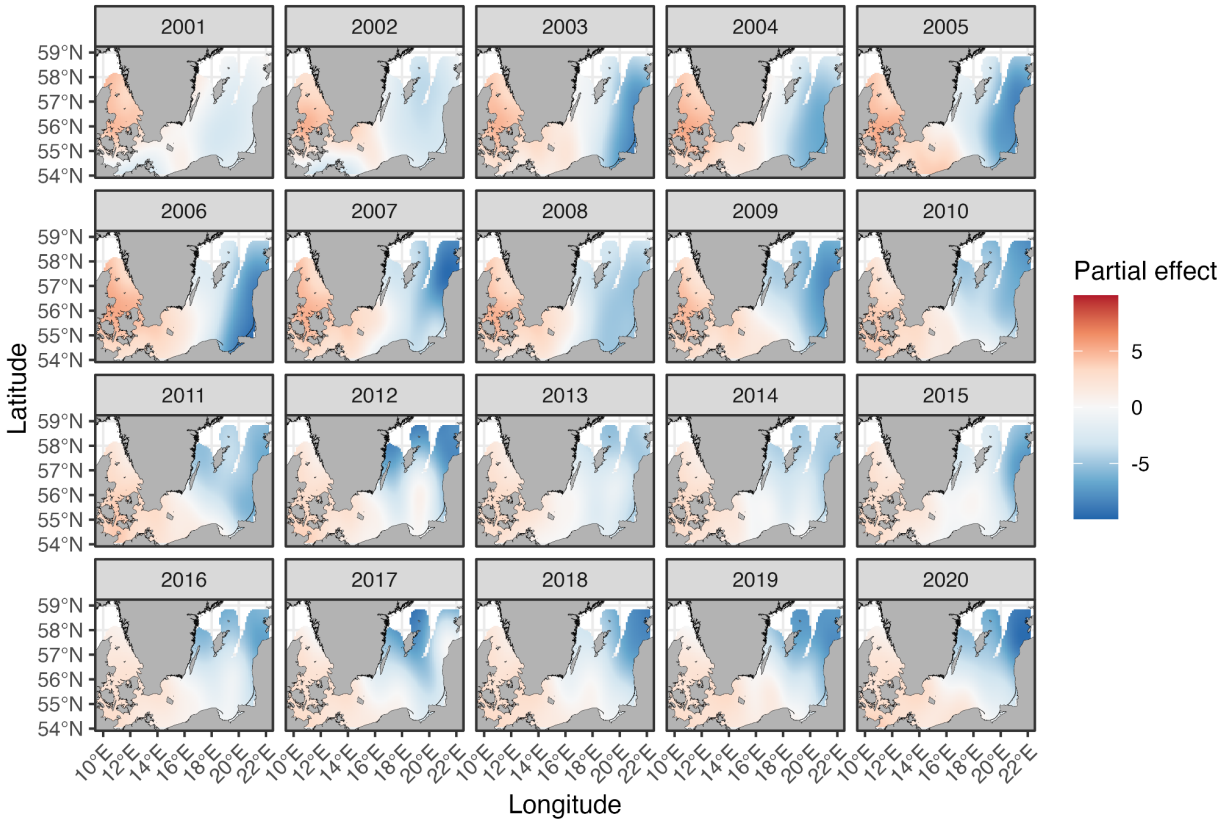


Figure 6.13 European plaice (*Pleuronectes platessa*) spatiotemporal partial dependence (2-d smooth; Simpson, 2018) for tensor smooth product (te(lon, lat, year, by = species)) in HGAM version-III. The 2-D smooth maps are cropped by the 0.5° buffer polygon to avoid extrapolation into unsampled areas.

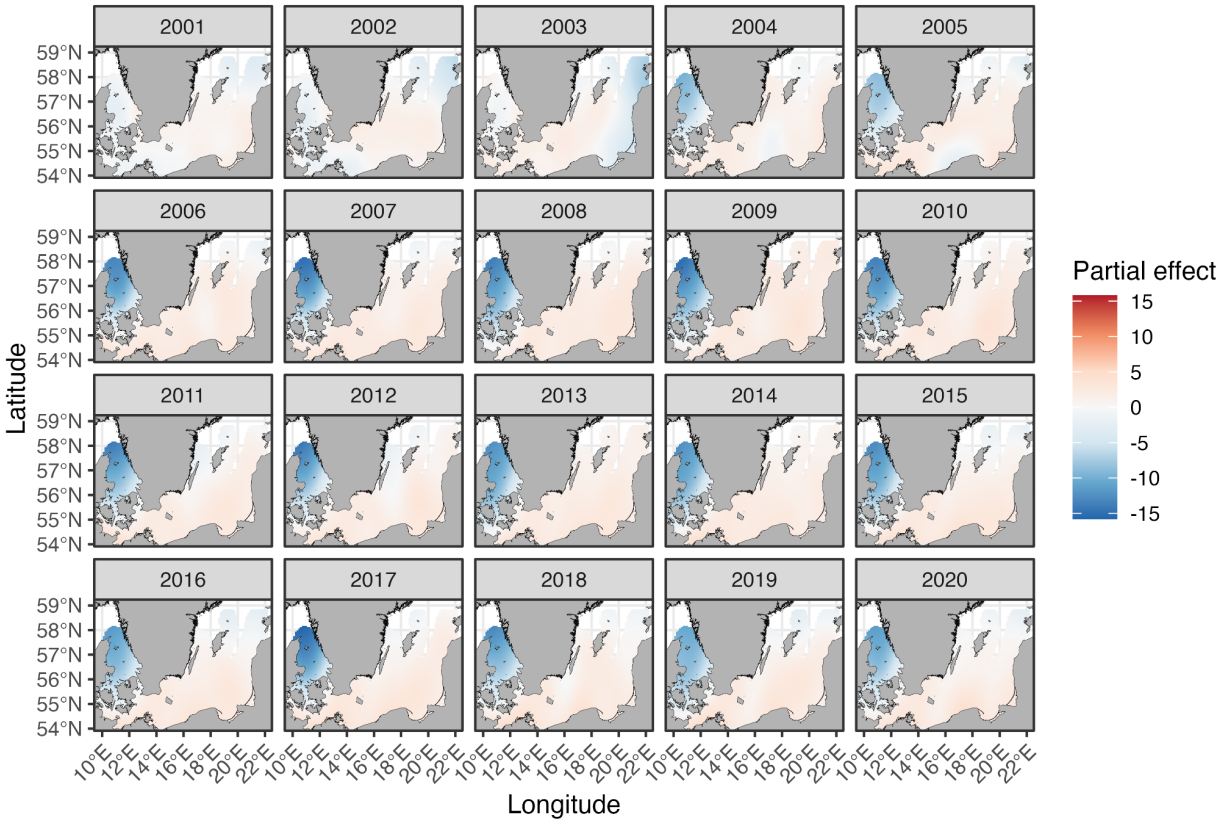


Figure 6.14 Juvenile Atlantic cod (*Gadus morhua,* <35 cm) spatiotemporal partial dependence (2-d smooth; Simpson, 2018) for tensor smooth product (te(lon, lat, year, by = species)) in HGAM version-III. The 2-D smooth maps are cropped by the 0.5° buffer polygon to avoid extrapolation into unsampled areas.

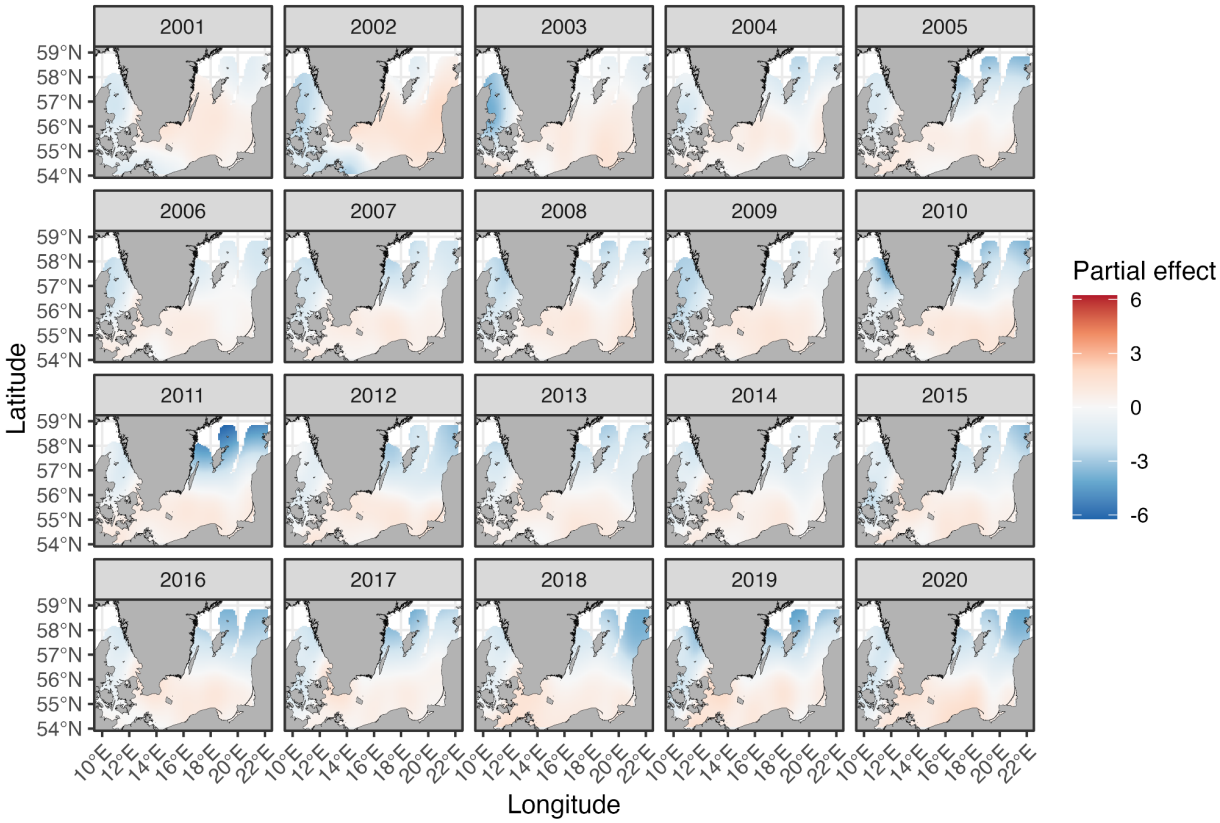


Figure 6.15 Adult Atlantic cod (*Gadus morhua,* ≥35 cm) spatiotemporal partial dependence (2-d smooth; Simpson, 2018) for tensor smooth product (te(lon, lat, year, by = species)) in HGAM version-III. The 2-D smooth maps are cropped by the 0.5° buffer polygon to avoid extrapolation into unsampled areas.

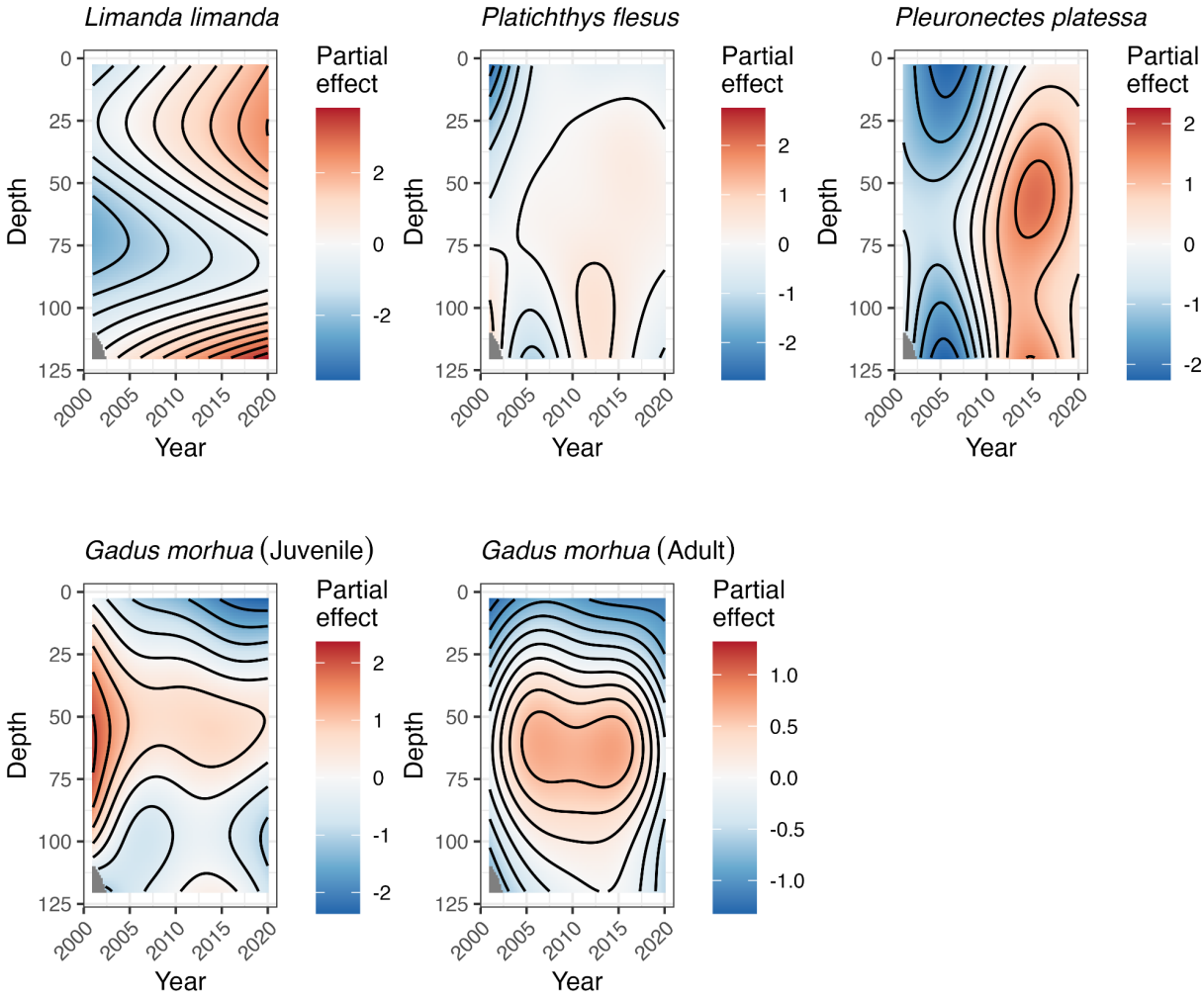


Figure 6.16 Spatiotemporal partial dependence (2-d smooth; Simpson, 2018) for tensor smooth product (te(depth, year, by = species)) in HGAM version-I. Depth is measured in metres. This term was simplified to s(depth, species) in subsequent model versions (II and III) to avoid high concurvity and did not impact model predictive performance or abiotic partial dependence curves. The greyed-out corner at deeper depths (110-125 m) from 2001-2002 inclusively represents a distance (default dist = 0.01) constraint in the smooth estimation to avoid spurious extrapolation.

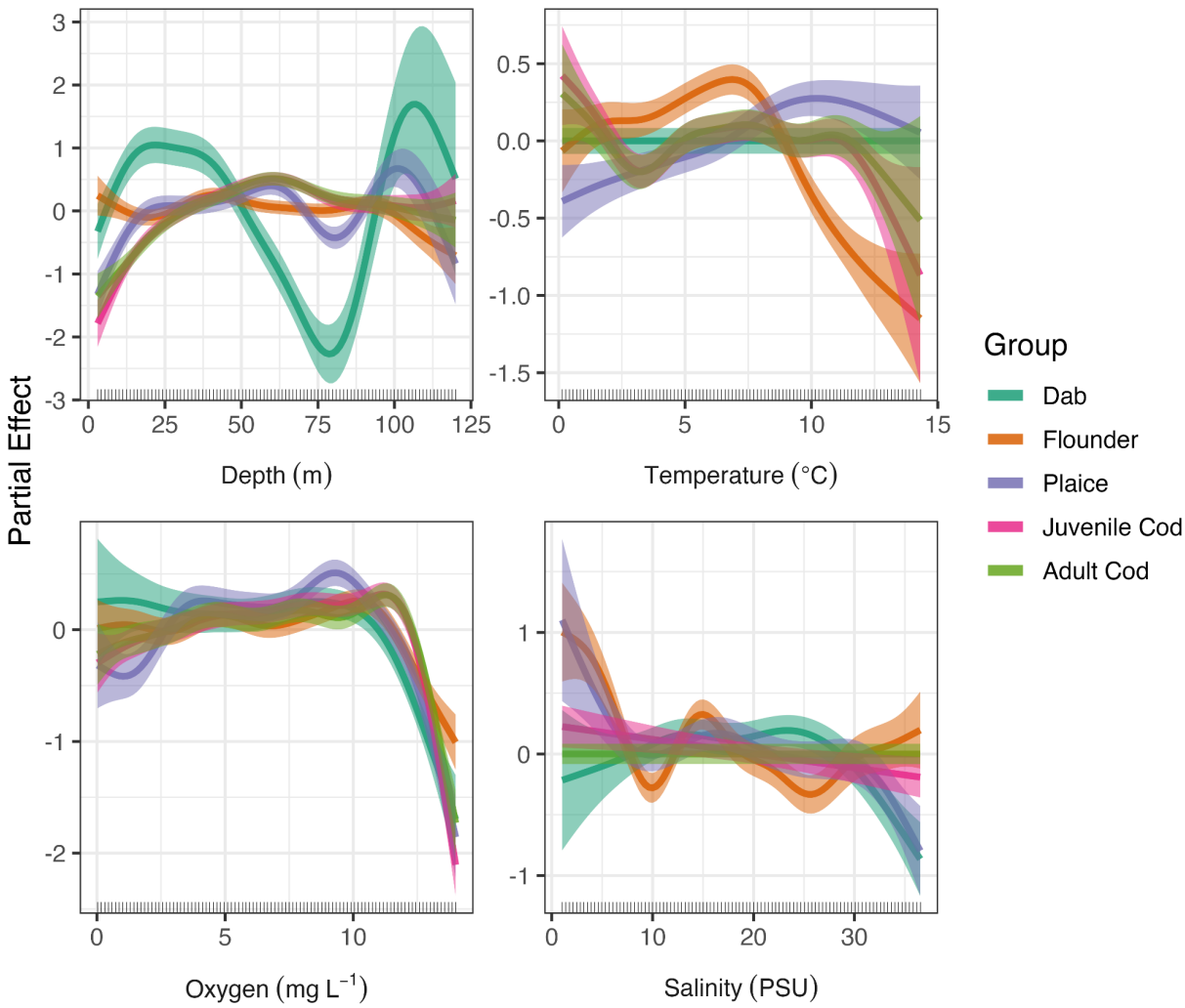


Figure 6.17 Spatiotemporal partial dependence (Simpson, 2018) for factor smooths of depth and abiotic covariates in HGAM version-II, ignoring seasonality. This model version outperformed the geographic-only model (HGAM version-I) but was inferior to HGAM version-III (seasonal covariates) based on predictive performance and AIC. 95% confidence intervals are shown for each species partial effect, and x-axis rugs mark the observed data coverage on the x-axis for each covariate.

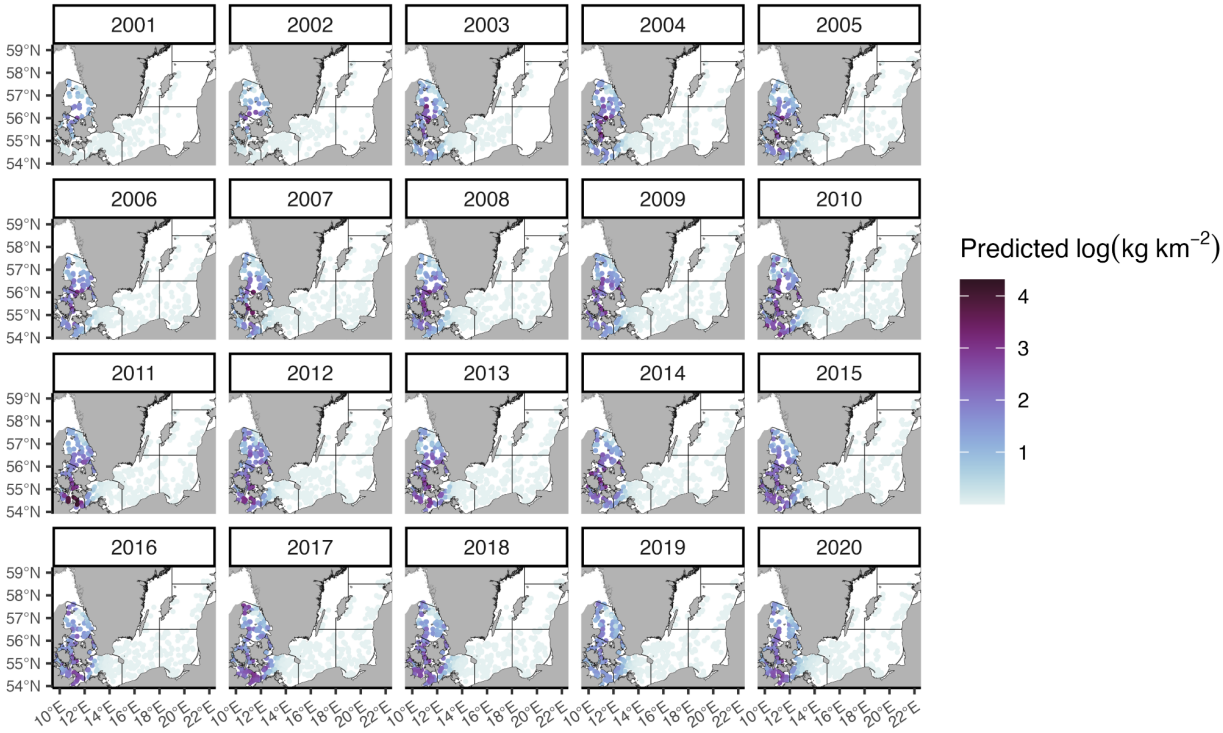


Figure 6.18 Common dab (*Limanda limanda*) biomass predictions for HGAM version-III, expressed as log(kg km^{-2}), for the full time series (2001-2020).

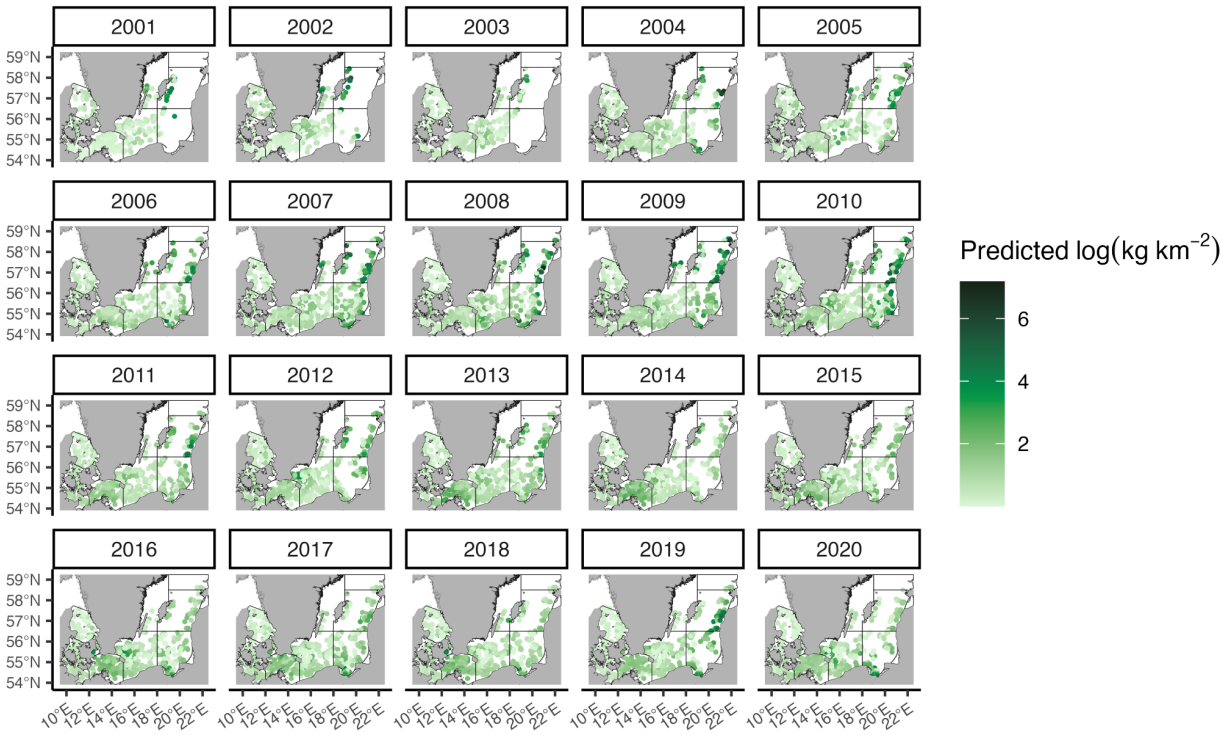


Figure 6.19 European flounder (*Platichthys flesus*) biomass predictions HGAM version-III, expressed as log(kg km^{-2}), for the full time series (2001-2020).

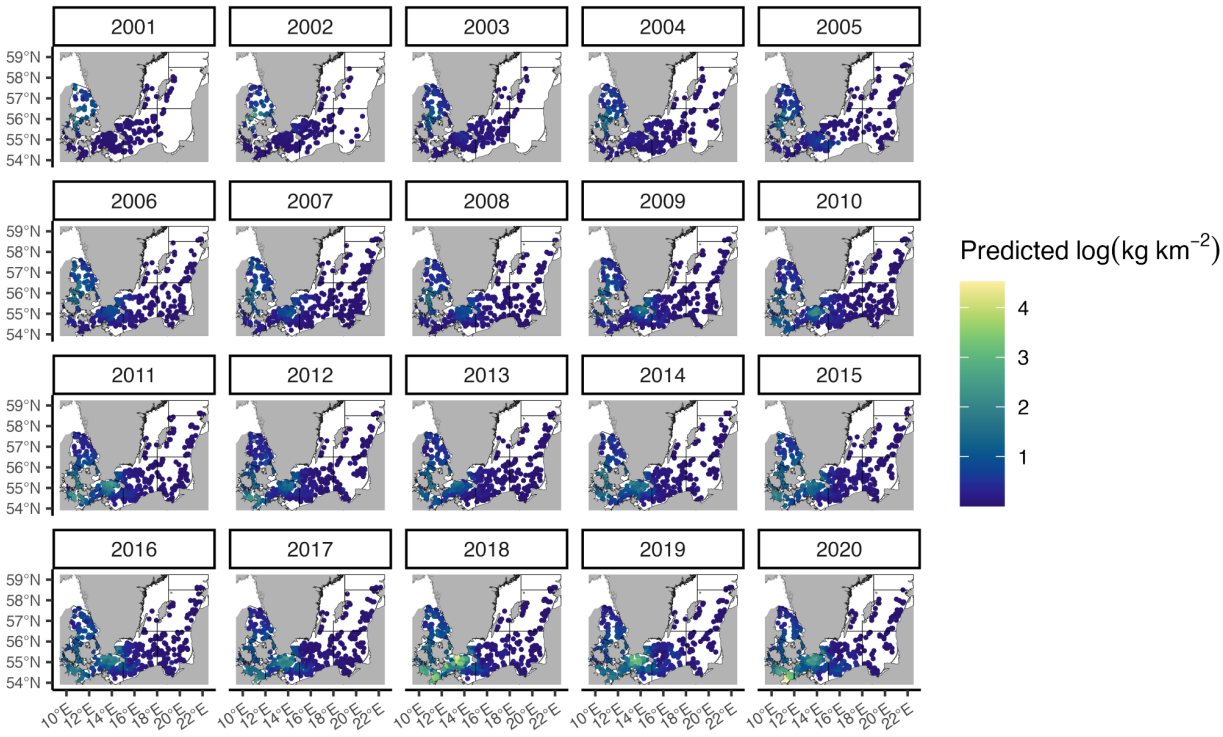


Figure 6.20 European plaice (*Pleuronectes platessa*) biomass predictions HGAM version-III, expressed as log(kg km^{-2}), for the full time series (2001-2020).

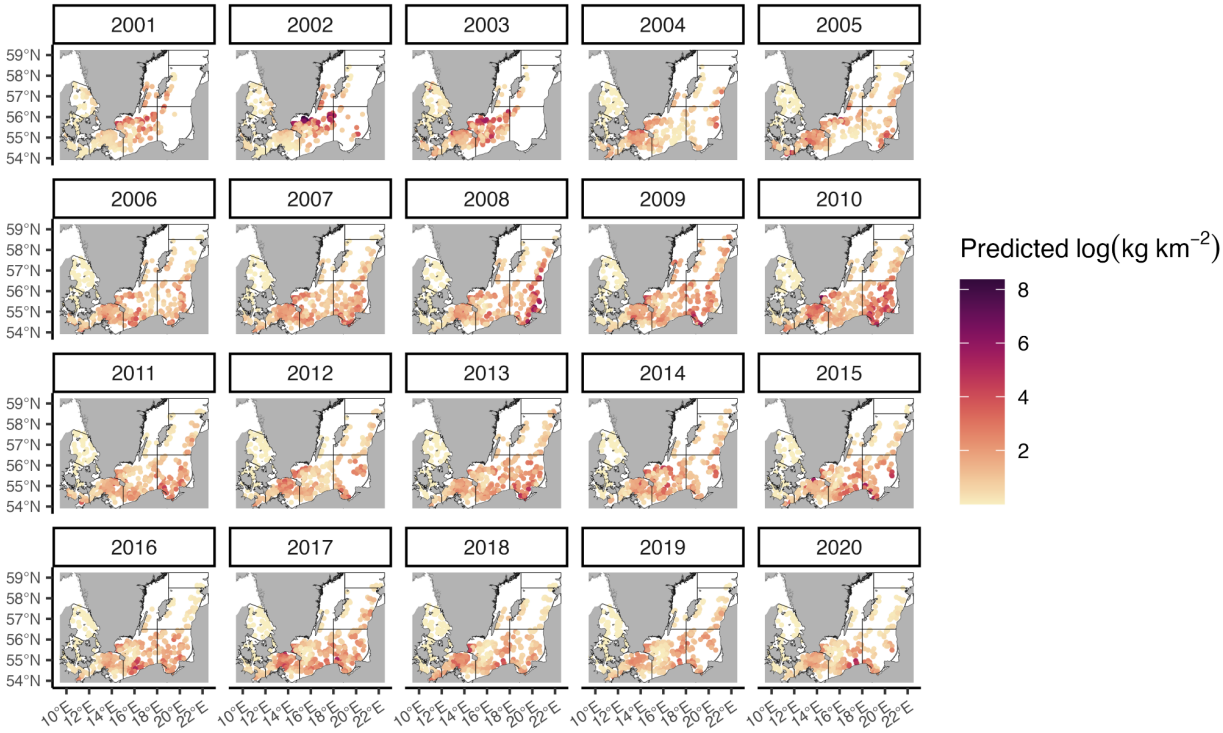


Figure 6.21 Juvenile Atlantic cod (*Gadus morhua*) biomass predictions HGAM version-III, expressed as log(kg km^{-2}), for the full time series (2001-2020).

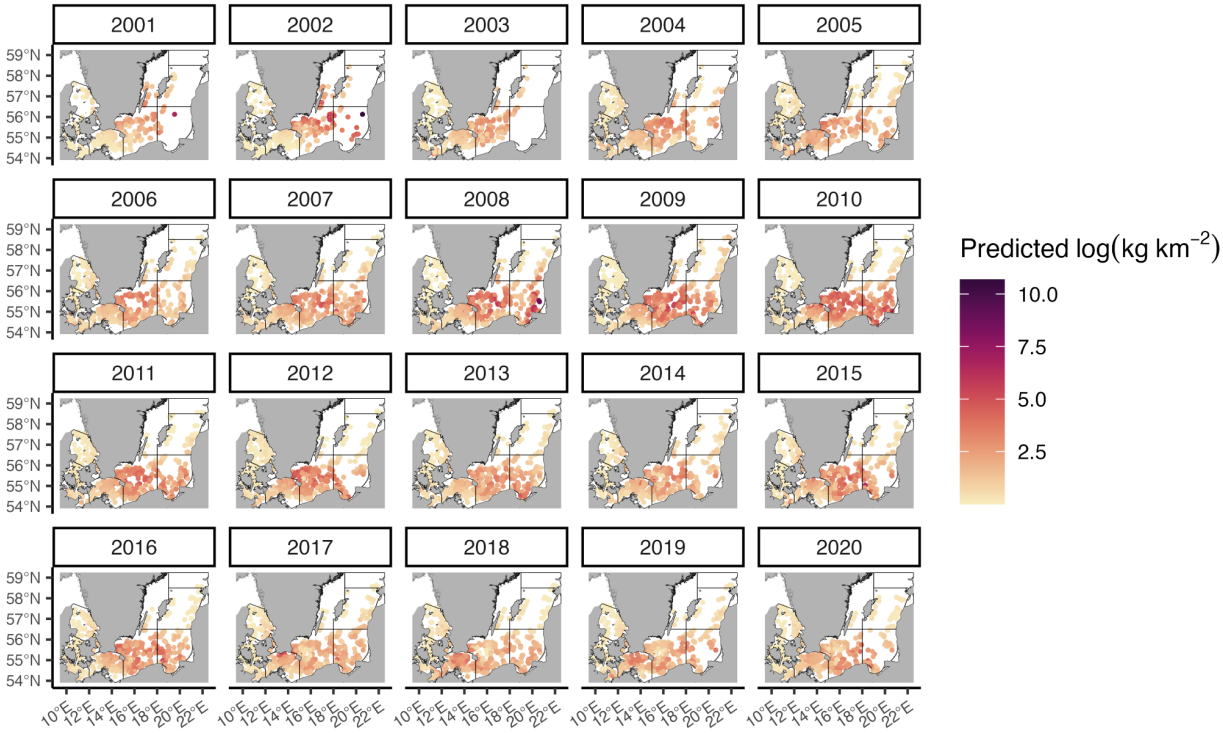


Figure 6.22 Adult Atlantic cod (*Gadus morhua*) biomass predictions HGAM version-III, expressed as log(kg km^{-2}), for the full time series (2001-2020).

6.1 References

Berg, C. W., Brun, M., Börjesson, P., Chaves, C., Degel, H., Lynam, C. P., Martinez, I., Schuchert, P., Soni, V., Velaso, F., Villamor, A., & Wieland, K. (2019). Workshop on Methods to develop a swept-area based effort index (WKSABI). International Council for the Exploration of the Sea (ICES). ICES Scientific Report Vol. 1 No. 3 https://doi.org/10.17895/ices.pub.4902

Leonardi, M., Boschin, F., Boscato, P., & Manica, A. (2022). Following the niche: The differential impact of the last glacial maximum on four European ungulates. *Communications Biology*, *5*(1), 1038. https://doi.org/10.1038/s42003-022-03993-7

Simpson, G.L. (2018). R Package: gratia. Ggplot-based graphics and other useful functions for GAMs fitted using Mgcv, 0.1-0 (Ggplot-based graphics and utility functions for working with GAMs fitted using the mgcv package).

7. Appendix S2: ODMAP Protocol following Zurrel *et al.* (2020): Spatial change of dominant Baltic Sea demersal fish across two decades

Supplement for Chapter 2.

Overview

Authorship

Contact: lmacneil@geomar.de

Model Objective

Mapping and interpolation

Target output: Spatially continuous biomass density (kg km^{-2}).

Focal Taxa

Atlantic Cod (*Gadus morhua*), European Plaice (*Pleuronectes platessa*), European Flounder (*Platichthys flesus*), Common Dab (*Limanda limanda*).

Location

Baltic Sea (marginal sea in the northeast Atlantic Ocean) including Kattegat, Øresund, western Baltic Sea, and Baltic proper (ICES subdivisions 21-29).

Scale of Analysis

Spatial extent (lon/lat): 9.420922, 22.51811, 53.91644, 59.23877 (xmin, xmax, ymin, ymax)

Spatial resolution:

1. Approximately 2 nautical miles
    a. Delta Longitude: 3’ 20” or 0.05556 degrees.
    b. Delta Latitude: 2’ or 0.03333 degrees.

Temporal extent: 2001-2020

Temporal resolution: Yearly, biannual

1. Q1 (Feb-Mar), 2001-2020
2. Q4 (Oct-Nov), 2001-2020

Boundary: Polygon, ICES subdivisions 21-29

Biodiversity Data
Observation Type: Fisheries-independent demersal trawl survey (DATRAS-BITS).

Response Data Type: Catch weight (biomass) density (kg km^{-2}).

Predictors
Types: Geographic, temporal, oceanographic.

Hypothesis
Raw occurrence data are unable to distinguish correlates to a species' realized niche, where a species achieves positive growth rates and contributes more to ecological functioning. To predict quantitative patterns in demersal fish in the Baltic Sea, seasonally-resolved covariates provide the best predictive quality to approximate a species' realized niche.

Model assumptions:
1. No systematic observation bias.
2. Independence of species observations.
3. Availability of all important predictors; key explanatory variables are available and incorporated in the model.
4. Predictors are free of error.
5. Niche stability/constancy, niche conservatism.

Algorithms
Modelling techniques: Three versions of Hierarchical Generalized Additive Models (HGAMs) allowing species-level variation to geographic and seasonal abiotic predictors.

Model complexity: Our multi-model approach included an incremental gain in model complexity from simpler geographic-only HGAMs (lon, lat, depth, year) to geographic plus abiotic covariates, and finally including seasonal abiotic covariates. Each model used a zero-inflated distribution assumption (Tweedie) and models were compared using AIC values to identify model parsimony ranks.

Model averaging: Model evaluation metrics (discrimination, precision, and accuracy; Waldock *et al.* 2022) are based on average achieved scores in k-fold cross validation. Final predictions are visualized as the difference between average species-specific biomass in the last five years (2016-2020) and the first five years (2001-2005) of the BITS survey. During the revision process, we also compared biomass differences between the first (2001–2010) and last (2011–2020) decades alongside a median rate of change per grid cell (kg km^{-2} yr^{-1}). All patterns are calculated within a 0.5° gridcell.

Workflow
Modelling: Trained HGAMs based on biomass density, analogous to Species Abundance Models (SAMs; e.g, Waldock *et al.* 2022; Chardon *et al.* 2022), for each demersal fish species hierarchically across the Baltic Sea region (Kattegat, Øresund, western Baltic Sea, Baltic proper). Each model contained a spatiotemporal geographic term to limit autocorrelation and included varying degrees of species-level variation through shared or independent smoothing penalties following Pedersen *et al.* (2019). Each model was evaluated across random k-folds (k = 10). Final predictions for mapping incorporated all observations (Step 5, option 1 in Roberts *et al.* 2017), favoring prediction quality over error estimation although we expect reasonable error estimates from cross validation.

Software
R versions 4.2.1 (R Core Team, 2022); packages 'stats' (R Core Team, 2022), 'raster' ([a]Hijmans *et al*. 2022), 'terra' ([b]Hijmans *et al*. 2022), 'tidyverse' (Wickham *et al.* 2019), 'mgcv' (Wood, 2023), 'oce' (Kelley, 2018), 'sf' (Pebesma, 2018), 'gratia' (Simpson, 2018)
Code availability: https://github.com/LiamMacNeil/Balt_Demersal_SDMs

Data availability:

1. DATRAS-BITS (Baltic International Trawl Survey) exchange data and swept-area assessment (https://datras.ices.dk/data_products/download/download_data_public.aspx).
2. Copernicus marine service environmental layers
    a. Baltic Sea Physics Reanalysis | Copernicus Marine MyOcean Viewer.
    b. Baltic Sea Biogeochemistry Reanalysis | Copernicus Marine MyOcean Viewer.

Data

Biodiversity data
Taxon names: Atlantic Cod (*Gadus morhua*)*,* Common Dab (*Limanda limanda*)*,* European Flounder (*Platichthys flesus*)*,* European Plaice (*Pleuronectes platessa*).

Taxonomic reference system: Linnean

Ecological level: Species

Data sources: Fisheries-independent demersal trawl survey (Baltic International Trawl Survey; BITS), routinely collected with standardized gears during Q1 (Feb-Mar) and Q4 (Oct-Nov) spanning 2001-2020.

S**ampling design:** Coordinated international demersal trawls (BITS) following ICES (2017) using standardized gears (TVL 930 meshes, TVS 520 meshes) at randomly stratified stations based on ICES subdivisions. Sampling design is described in more detail within ICES (2017).

Sample Size: A total of 8991 individual trawls are included.

Regional mask: All data are clipped to a boundary region of ICES subdivisions 21-29 and covariates further masked by a buffer polygon 0.5° within the radius of trawl observations.

Scaling: Trawl observations were spatially thinned to the resolution of environmental predictors (~2 nautical miles) for each species, during each modelling time window, to reduce the spatial autocorrelation from repeated observations.

Data cleaning/filtering:

1. Observations from DATRAS-BITS exchange data and swept area assessments were removed that did not report (1) lon/lat, (2) station number, (3) catch weight (CatCatchWgt), (4) species length class (LngtClass), or (5) door-based swept area (SweptAreaDSKM2; Berg *et al.* 2019).

2. Only trawls designated by DATRAS-BITS as “Valid”, “Additional”, or “Calibration” are considered.
3. Trawls durations below 15 minutes and longer than 90 minutes were excluded.
4. Only observations produced from two standard gear types (TVL 930 meshes, TVS 520 meshes) were included.
5. DATRAS-BITS were trimmed to within 2001-2020, inclusively. The year 2001 represents the time of implementation of standard gear types and the year 2020 represents the final year available for swept area assessments (Berg *et al.* 2019).

Background data: Zeros were imputed for all trawls where our focal species were not recorded, totalling ~22% in frequency.

Errors and biases:

1. Coastal regions (e.g., <20 m depth) are poorly covered by the DATRAS-BITS survey, thus we expect poorer prediction accuracy in these areas, otherwise major basins throughout the Baltic Sea (western Baltic Sea, Arkona Basin, eastern Bornholm Basin, and eastern Gotland Basin) are richly covered.
2. Misidentification is expected to be low, given all four species are common focal taxa for demersal surveys. Flounder is a possible exception, because a new Baltic species was only recently demarcated (Momigliano *et al.* 2018), which potentially affected taxonomic identification afterwards in the most northern extent of our study area (Gulf of Riga, and northern Baltic proper) where this species occurs.
3. Data gaps (missing and outliers) for calculating swept area (km^2; estimate of effort) do exist and are caused predominantly by the lack of gear geometry sensors during deployment (Berg *et al.* 2019). The richest data availability, supporting the most reliable swept area estimates, are provided by Germany, Denmark, and Sweden covering ICES subdivisions 22-26 (western Baltic Sea, Øresund, Arkona Basin, eastern Bornholm Basin). Errors outside these regions (ICES subdivisions 27-29; Baltic proper) are possibly affected by unreliable estimates (AS7).

Data partitioning
Training data: Randomized 10-fold cross validation where k-1 folds are iteratively used for model training.

Validation/Test data: Performance is assessed for each model on withheld k-fold.

Predictor variables
Predictor variables:

1. Physics Reanalysis
    a. Sea water potential temperature at sea floor (bottomT).
    b. Sea water salinity at sea floor (Sb).

2. Biogeochemistry Reanalysis
    a. Mole concentration of dissolved molecular oxygen in sea water at sea floor (O_2b).

Data sources:
Copernicus marine service oceanographic layers
    c. Baltic Sea Physics Reanalysis | Copernicus Marine MyOcean Viewer.

d. Baltic Sea Biogeochemistry Reanalysis | Copernicus Marine MyOcean Viewer.

Spatial extent (lon/lat): 9.420922, 22.51811, 53.91644, 59.23877 (xmin, xmax, ymin, ymax)

Spatial resolution:
1. Approximately 2 nautical miles
    a. Delta Longitude: 3’ 20” or 0.05556 degrees
    b. Delta Latitude: 2’ or 0.03333 degrees

Temporal extent: 2001-2020

Temporal resolution: Yearly, biannual
1. Q1 (Feb-Mar), 2001-2020
2. Q4 (Oct-Dec), 2001-2020

Coordinate reference system: longlat; EPSG:4326

Data processing: We extracted monthly values for every oceanographic variable and created weighted averages (weighted by observation frequency) for each quarter of each year.

Errors and biases: No outstanding errors or biases are expected from data processing.

1. For each reanalysis file (numerical model), a quality information document indicates the following patterns in environmental layers:
    a. Bottom oxygen displayed relatively high root-mean-square deviation from observed values in the southern Baltic proper.

Dimension reduction: NA

Model

Variable pre-selection
Variable pre-selection: We focused on bottom-associated variables to pre-select variables which shape environments most often occupied by these demersal fish. Further, we pre-selected core variables (temperature, salinity, oxygen) which have been previously described and attributed to impact Baltic demersal niches (e.g., Rau *et al.* 2012; Smoliński and Radtke, 2017; Orio *et al.* 2019; Brander, 2022).

Multicollinearity
Multicollinearity: We assessed variable concurvity– a generalization of collinearity– for each model and found no variable exceeded 0.8. No absolute threshold exists to exclude variables based on concurvity, but GAMs are strongly robust to biases due to concurvity (Wood, 2008).

Model settings

GAM: All HGAMs were built in the *bam* function in *mgcv*; all versions shared a Tweedie distribution assumption with a log-link function, plus discretized covariates and fast restricted maximum likelihood (fREML) for efficient computation. Parsimonious model configurations were determined by null space penalization. Variable smoothers were constructed for geographic and abiotic covariates using tensor smooth products for interactions (lon x lat x year; depth x year) species-level variation was incorporated using species-specific smoothing penalties (by = species) in HGAM-II for depth and abiotic covariates. For HGAM version-III (seasonal abiotic covariates), species-level variation was permitted to vary independently across seasons using factor smooths (bs ="fs"). Random intercepts were also included across model versions without factor smooths (HGAM-I, HGAM-II) through random effect smoothers (bs = "re"). For model versions-II (geography + abiotic covariates) and -III (geography + seasonal abiotic covariates), we penalized the square second derivative ($m = 2$) of group-level smoothers to limit excessive curvature (Pedersen *et al.* 2019). To address temporal autocorrelation in model residuals, we added an AR1 autoregressive process based on the observed lag using the *rho* parameter in the *bam* function.

Model estimates
Coefficients: Coefficient estimates for GAMs

Parameter uncertainty: Coefficient variation within cross-validation.

Variable importance: NA

Model selection
GAM: Automatic regularization with 'null space penalization' (https://www.rdocumentation.org/packages/mgcv/versions/1.8-42/topics/step.gam), which penalizes each smoothing term and effectively removes the predictor from model predictions if the smoothing term tends toward zero.

Model ensembles: NA

Analysis and Correction of non-independence
Spatial autocorrelation: A spatiotemporal model term, varying independently by species, was included in each model to incorporate spatiotemporal dependencies in the data.

Nested data: All trawl data were thinned to covariate grid level. Hierarchical groupings including species- and seasonal-level (quarter) variation.

Assessment

Performance statistics
Metrics have been chosen from review in Waldock *et al.* (2022)

Performance on training data: Pearson's correlation coefficient (Discrimination), Mean Absolute Error (MAE; Accuracy), dispersion (σ_{est} / σ_{obs})[4].

[4] Variance of estimated (σ_{est}) and observed (σ_{obs}) biomass density

Performance on validation data: Pearson's correlation coefficient (Discrimination), Mean Absolute Error (MAE; Accuracy), dispersion (σ_{est} / σ_{obs}).

Performance on test data: Pearson's correlation coefficient (Discrimination), Mean Absolute Error (MAE; Accuracy), dispersion (σ_{est} / σ_{obs}).

Plausibility check
Response shapes: Partial response curves (Simpson, 2018).

Prediction

Prediction output
Prediction unit: Species-level catch weight (biomass) density (kg km^{-2}).

Post-processing: No post-prediction calibration was applied (*sensu* Dormann, 2020) as we do not use conventional classification-based species distribution models with a probabilistic interpretation.

Uncertainty quantification
Algorithmic uncertainty: Random cross-validation (k = 10) within training-test set splits where model performance metrics are the averages (± 95 and 5% quantiles).

7.1 References

Berg, C. W., Brun, M., Börjesson, P., Chaves, C., Degel, H., Lynam, C. P., Martinez, I., Schuchert, P., Soni, V., Velaso, F., Villamor, A., & Wieland, K. (2019). Workshop on Methods to develop a swept-area based effort index (WKSABI). International Council for the Exploration of the Sea (ICES). ICES Scientific Report Vol. 1 No. 3 https://doi.org/10.17895/ices.pub.4902.

Brander, K. (2022). Support for the hypothesis that growth of eastern Baltic cod is affected by mild hypoxia. A comment on Svedäng et al. (2022). *ICES Journal of Marine Science*, *79*(7), 2155–2156. https://doi.org/10.1093/icesjms/fsac070

Chardon, N. I., Nabe-Nielsen, J., Assmann, J. J., Dyrholm Jacobsen, I. B., Guéguen, M., Normand, S., & Wipf, S. (2022). High resolution species distribution and abundance models cannot predict separate shrub datasets in adjacent Arctic fjords. *Diversity and Distributions*, *28*(5), 956–975. https://doi.org/10.1111/ddi.13498.

Dormann, C. F. (2020). Calibration of probability predictions from machine-learning and statistical models. *Global Ecology and Biogeography*, *29*(4), 760–765. https://doi.org/10.1111/geb.13070

[a]Hijmans, R.J. (2022). Geographic Data Analysis and Modeling. R package version v. 3.6.14.

[b]Hijmans, R. J., Bivand, R., Forner, K., Ooms, J., & Pebesma, E. (2022). R package version v. 1.7.3. https://rspatial.org/terra/

ICES (2017). SISP 7 BITS Manual for the Baltic International Trawls Surveys. https://doi.org/10.17895/ICES.PUB.2883

Kelley, D. *"The Oce Package" In Oceanographic Analysis with R*. R package version v. 1.7.10, 91–101 (Springer, 2018).

Momigliano, P., Denys, G. P. J., Jokinen, H., & Merilä, J. (2018). *Platichthys solemdali* sp. nov. (Actinopterygii, Pleuronectiformes): A New Flounder Species From the Baltic Sea. *Frontiers in Marine Science*, *5*, 225. https://doi.org/10.3389/fmars.2018.00225

Orio, A., Bergström, U., Florin, A., Lehmann, A., Šics, I., & Casini, M. (2019). Spatial contraction of demersal fish populations in a large marine ecosystem. *Journal of Biogeography*, *46*(3), 633–645. https://doi.org/10.1111/jbi.13510

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data, *The R Journal*, 10(*1*) 439-446. R package version v. 1.0.9.

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876. https://doi.org/10.7717/peerj.6876

Rau, A., Lewin, W.-C., Zettler, M. L., Gogina, M., & von Dorrien, C. (2019). Abiotic and biotic drivers of flatfish abundance within distinct demersal fish assemblages in a brackish ecosystem (western Baltic Sea). *Estuarine, Coastal and Shelf Science*, *220*, 38–47. https://doi.org/10.1016/j.ecss.2019.02.035

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. https://doi.org/10.1111/ecog.02881

Simpson, G.L. (2018). R Package: gratia. Ggplot-based graphics and other useful functions for GAMs fitted using Mgcv, 0.1-0 (Ggplot-based graphics and utility functions for working with GAMs fitted using the mgcv package).

Smoliński, S., & Radtke, K. (2017). Spatial prediction of demersal fish diversity in the Baltic Sea: Comparison of machine learning and regression-based techniques. *ICES Journal of Marine Science*, *74*(1), 102–111. https://doi.org/10.1093/icesjms/fsw136

Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W. L., Edgar, G. J., Mouillot, D., Tjiputra, J., & Pellissier, L. (2022). A quantitative review of abundance-based species distribution models. *Ecography*, *2022*(1), ecog.05694. https://doi.org/10.1111/ecog.05694.

Wickham *et al.*, (2019). Welcome to the Tidyverse. Journal of Open Source Software, 4(43), R package version v. 1.3.2 (2022), 1686, https://doi.org/10.21105/joss.01686.

Wood, S. N. (2008). Fast Stable Direct Fitting and Smoothness Selection for Generalized Additive Models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *70*(3), 495–518. https://doi.org/10.1111/j.1467-9868.2007.00646.x
Wood, S. (2023). Package 'mgcv'. R package version 1.8.41.

Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., … Merow, C. (2020). A standard protocol for reporting species distribution models. *Ecography*, *43*(9), 1261–1277. https://doi.org/10.1111/ecog.04960.

8. Supplement: Network-based bioregionalization of demersal fish in continental shelf seas

Supplement for Chapter 3.

Table 8.1 Overview of the scientific bottom trawl surveys from Maureaud et al. (2024) including their associated acronyms shown in Figure 8.1, the quarters sampled, and the geographic region targeted.

Survey	Acronym	Quarter	Region
West coast triannual trawl survey	WCTRI	2,4	Pacific northwest (USA/Canada)
West coast annual trawl survey	WCANN	2,3,4	Pacific northwest (USA/Canada)
Scotland west coast international bottom trawl survey	SWC-IBTS	1,4	Scottish west coast
Spanish Porcupine bank trawl survey	SP-PORC	3,4	Porcupine Bank (Irish Shelf)
Spanish north trawl survey	SP-NORTH	3,4	Northern Spanish coast
SP-ARSA	SP-ARSA	1,4	Gulf of Cadiz (Spain)
Southeast United States trawl survey	SEUS	2,3,4	Southeast United States
Scotian shelf trawl survey	SCS	1,2,3	Scotian Shelf (Canada)
Rockall plateau survey	ROCKALL	3	Rockall Plateau (European shelf)
Portuguese international bottom trawl survey	PT-IBTS	4	West Portuguese coast
North Sea international bottom trawl survey	NS-IBTS	1,3	North Sea, Skagerrak
Norwegian international bottom trawl survey	NOR-IBTS	3,4	Norwegian Seas
Northern Ireland groundfish survey	NIGFS	1	Irish Sea

Northeast United States trawls survey	NEUS	4	Northeast US coast
Ireland international groundfish survey	IE-IGFS	4	Irish Sea
Gulf of St. Lawrence south trawl survey	GSL-S	3,4	Southern Gulf of St. Lawrence (Canada)
Gulf of St. Lawrence north trawl survey	GSL-N	2,3,4	Northern Gulf of St. Lawrence (Canada)
Gulf of Alaska trawl survey	GOA	2,3	Gulf of Alaska (USA)
Gulf of Mexico trawl survey	GMEX	2,3,4	Northern Gulf of Mexico (USA)
France coordinated groundfish survey	FR-CGFS	4	Eastern English Channel (France)
EVHOE trawl survey	EVHOE	4	Bay of Biscay and Celtic Seas (France)
East Bering Sea trawl survey	EBS	2,3	East Bering Sea (USA)
Department of Fisheries and Oceans west coast Vancouver Island survey	DFO-WCVI	1	Pacific northwest (Canada)
Department of Fisheries and Oceans west coast Haida Gwaii survey	DFO-WCHG	1	Pacific northwest (Canada)
Department of Fisheries and Oceans west coast Strait of Georgia survey	DFO-SoG	1	Pacific northwest (Canada)
Department of Fisheries and Oceans west coast Queen Charlotte Sound survey	DFO-QCS	1	Pacific northwest (Canada)
Department of Fisheries and Oceans west coast Hecate Strait survey	DFO-HS	1	Pacific northwest (Canada)
Baltic international trawl survey	BITS	1,4	Baltic Sea

Aleutian Islands trawl survey	AI	2,3	Aleutian Islands (USA)

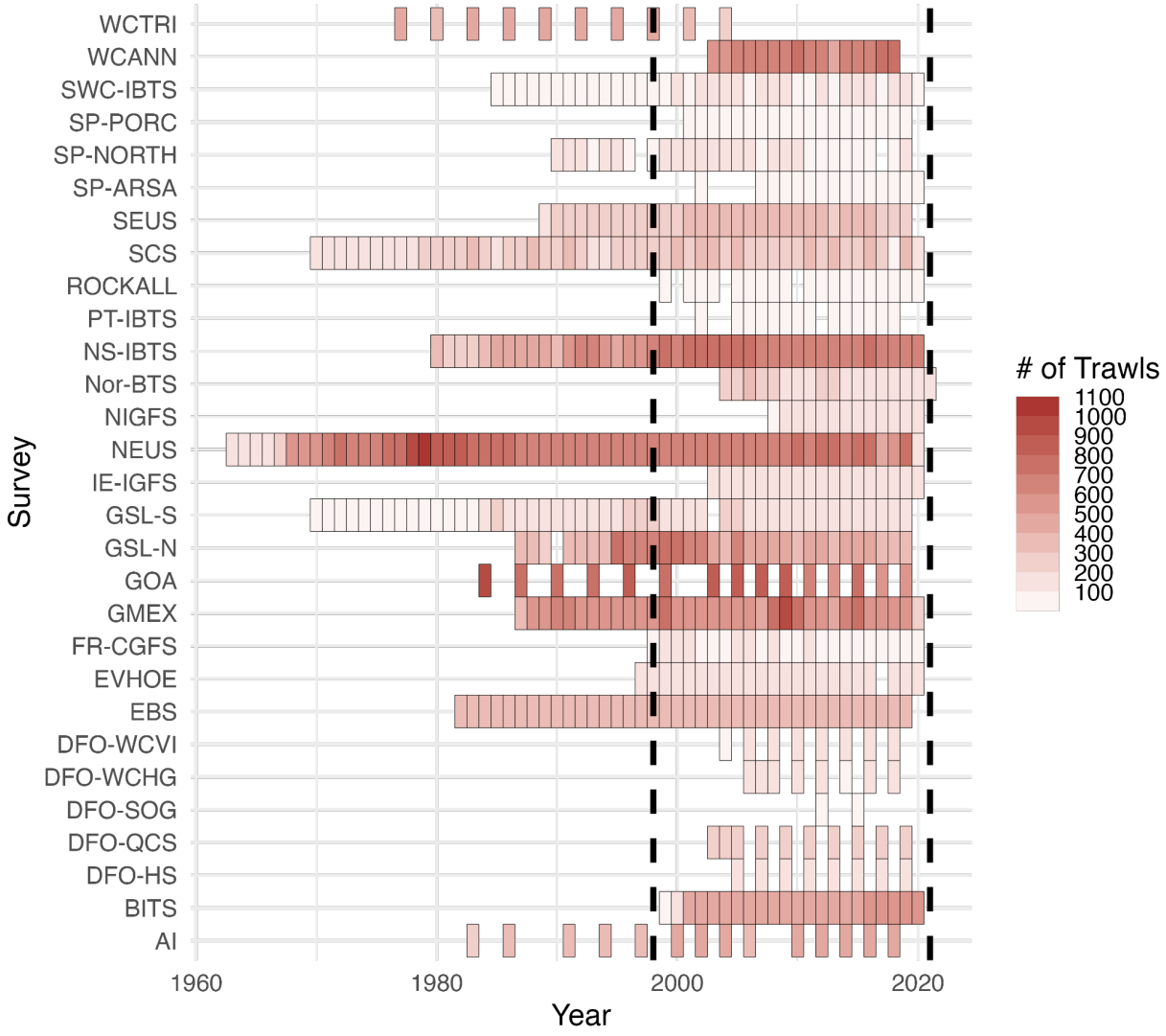


Figure 8.1 Gantt chart for number of trawls per survey across total years available. Here the vertical dashed lines bound the years included in our study (1999-2020). The majority of surveys had already begun (17/29) in 1999.

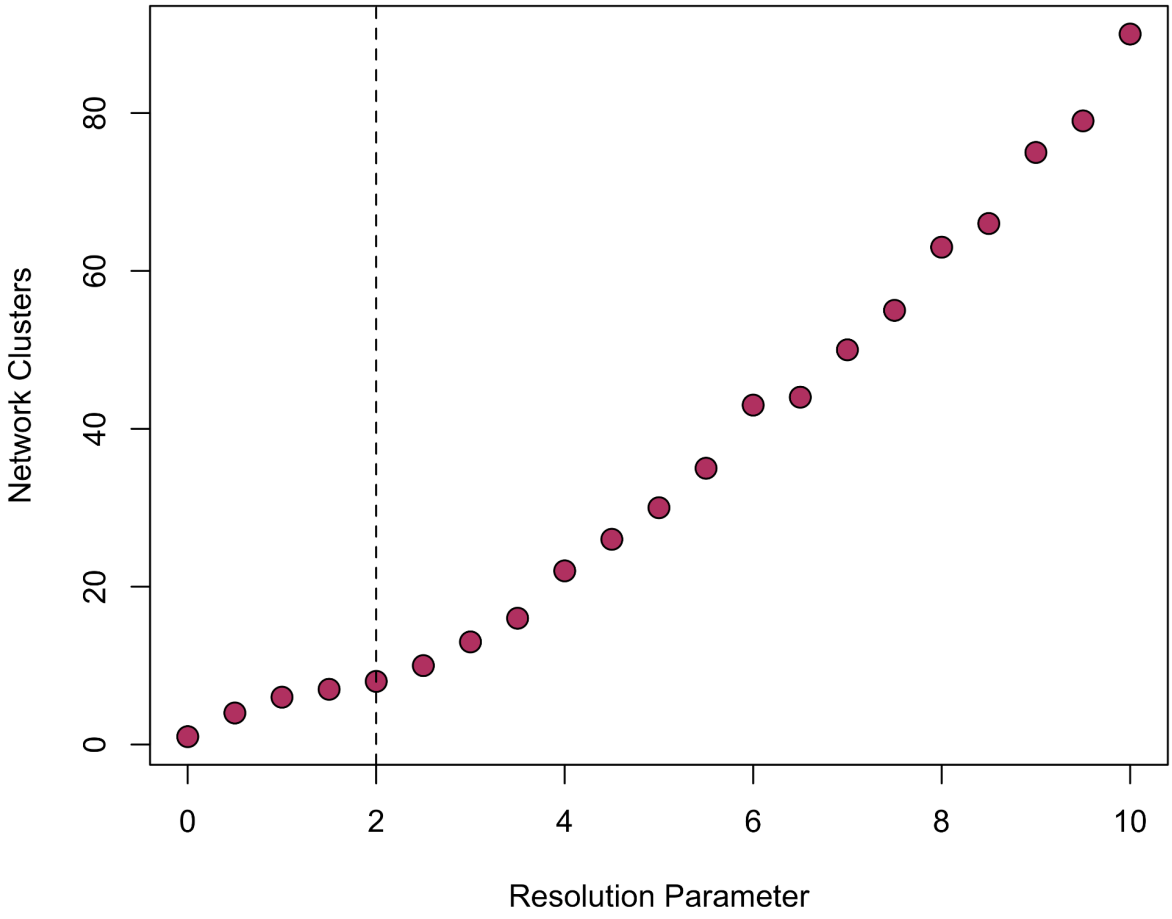


Figure 8.2 Searching across the resolution parameter— defining a minimum density threshold for clusters to be defined— in the Leiden clustering algorithm (Traag et al. 2019). The vertical dashed line indicates the value used as it was the minimum value to reproduce the North Sea-Baltic Sea divide where the salinity gradient powerfully filters the fish community (Pecuchet et al. 2016)— a larger resolution parameter threshold produced a patchwork of clusters without spatial continuity. This plot displays the number of network clusters identified at 200 km resolution and rarefied less restrictively ($n_{1stQ,\ 200\ km} = 124$), however the results were effectively identical at 100 km resolution ($n_{1stQ,\ 100\ km} = 92$) and at more restrictive rarefaction depths ($n_{median,\ 100\ km} = 194$; $n_{median,\ 200\ km} = 337$).

Table 8.2 The overview of bipartite network clustering scales at spatial resolution and rarefaction depth to which each grid cell is randomly sampled. Network modularity scores were measured in igraph (Csardi & Nepusz 2006); network modularity indicates how strongly modular and connected the proposed community partitions appear, with values closer to one representing total modularity and distinctiveness, and values closer to zero showing more diffuse and less strongly connected clusters.

Algorithm	Nodes	Resolution	Rarefaction	Modularity
Leiden	Weighted	100 km	1st Quartile	0.60
Leiden	Weighted	100 km	Median	0.59
Leiden	Weighted	200 km	1st Quartile	0.61
Leiden	Weighted	200 km	Median	0.59
Leiden	Unweighted	200 km	NA	0.58
Infomap	Weighted	100 km	1st Quartile	0.60
Infomap	Weighted	100 km	Median	0.60
Infomap	Weighted	200 km	1st Quartile	0.62
Infomap	Weighted	200 km	Median	0.58
Infomap	Unweighted	200 km	NA	0.63

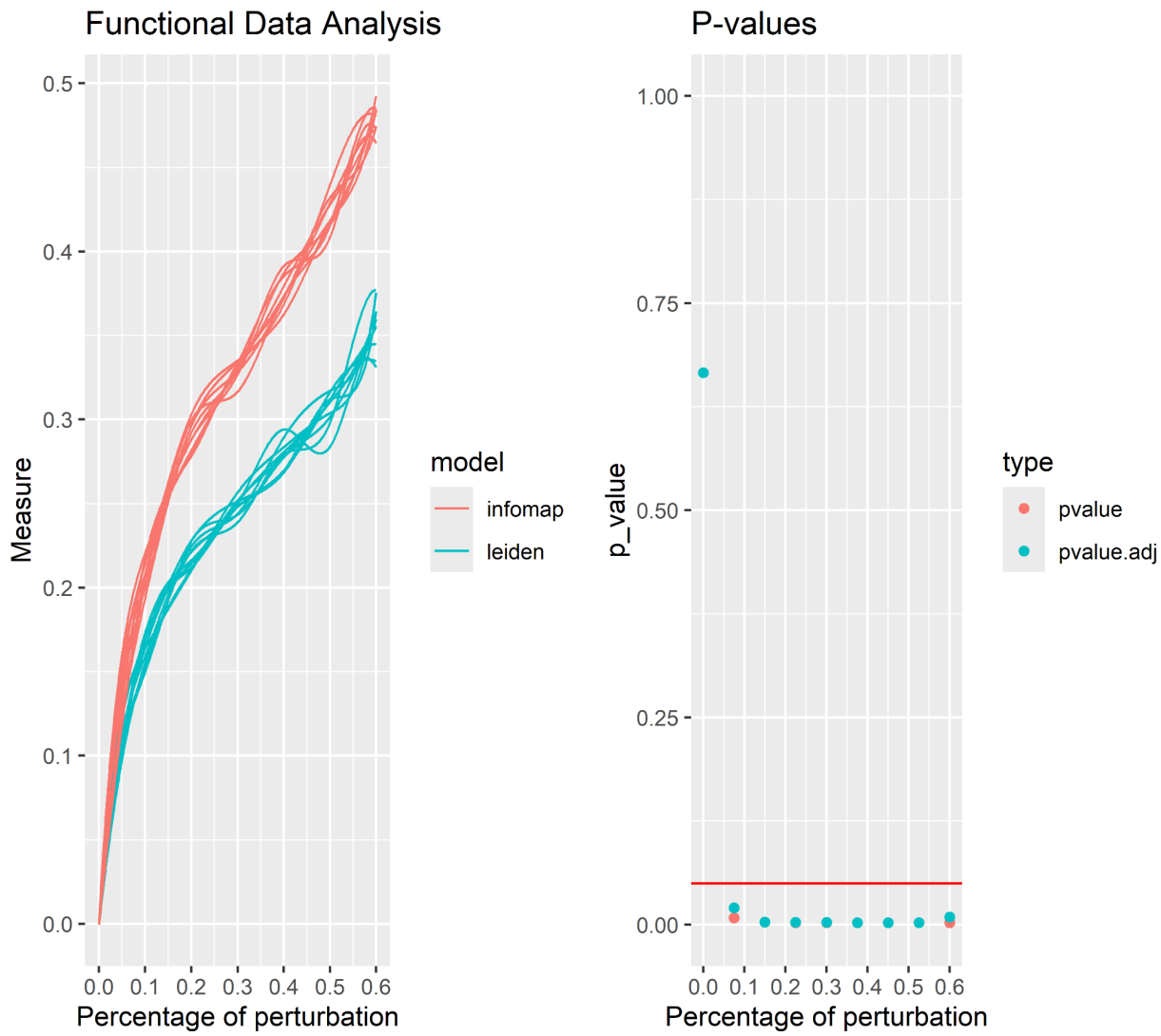


Figure 8.3 A statistical comparison of the Leiden (Traag et al. 2019) and infomap (Rosvall & Bergstrom 2008) algorithms using the robin package (Policastro et al. 2021). This comparison was made at 200 km grid cell resolution and with the less restrictive rarefaction ($n_{1stQ,\ 200\ km}$ = 124), using a weighted network for clustering. The left panel displays a variation of information (VI) indicator (Meilă 2007) across 5% sequential increases in network perturbations using a degree preserving randomization of network edges, showing less information loss and greater stability in partitions from the Leiden algorithm. The left panel displays the statistical significance of VI values between algorithms at perturbation values including adjusted p-values for the multiple comparisons generated during basis expansions (Pini & Vantini 2016). By perturbing the original network, less information is lost with the Leiden algorithm partitions, which is significantly outperforming the infomap algorithm after 5% perturbation.

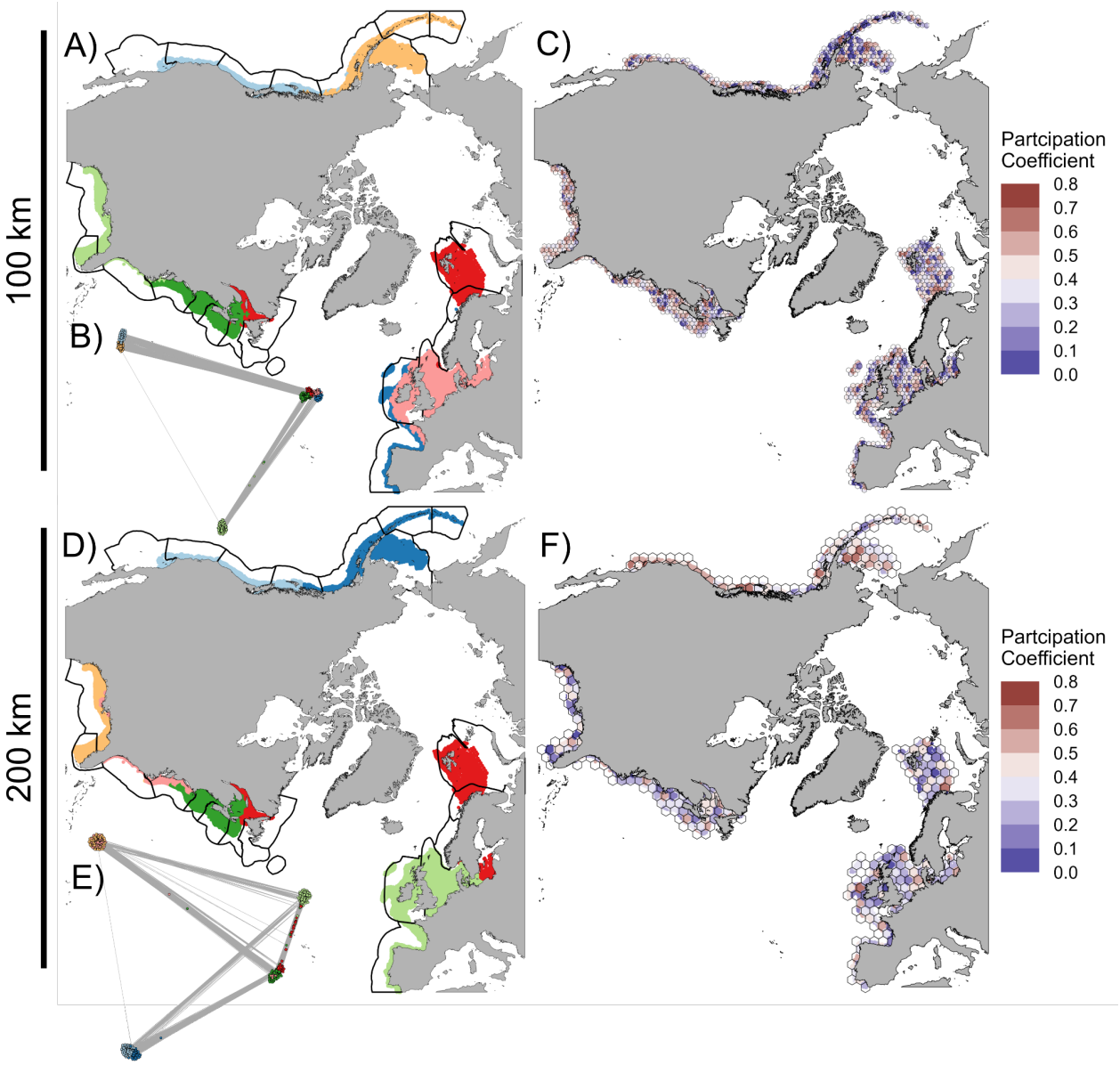


Figure 8.4 Bioregionalization based on non-rarefied data points using an unweighted Leiden algorithm at 100 km (A-C) and 200 km (D-F) shown underneath ecoregions (Spalding et al. 2007). B, E) The bipartite networks (site × species) are colored by delineated bioregions, showing strong geographic clustering with the denser network structure. The unweighted bioregionalization approach largely reproduces the patterns in the weighted approach but misses the Baltic Sea–North Sea divide at finer spatial resolution (A). Such a divide is illustrated by the participation coefficient (PC; equation 1) maps (C, F) that visualize the degree of shared species between bioregions, illustrating potential transition zones at high PC values.

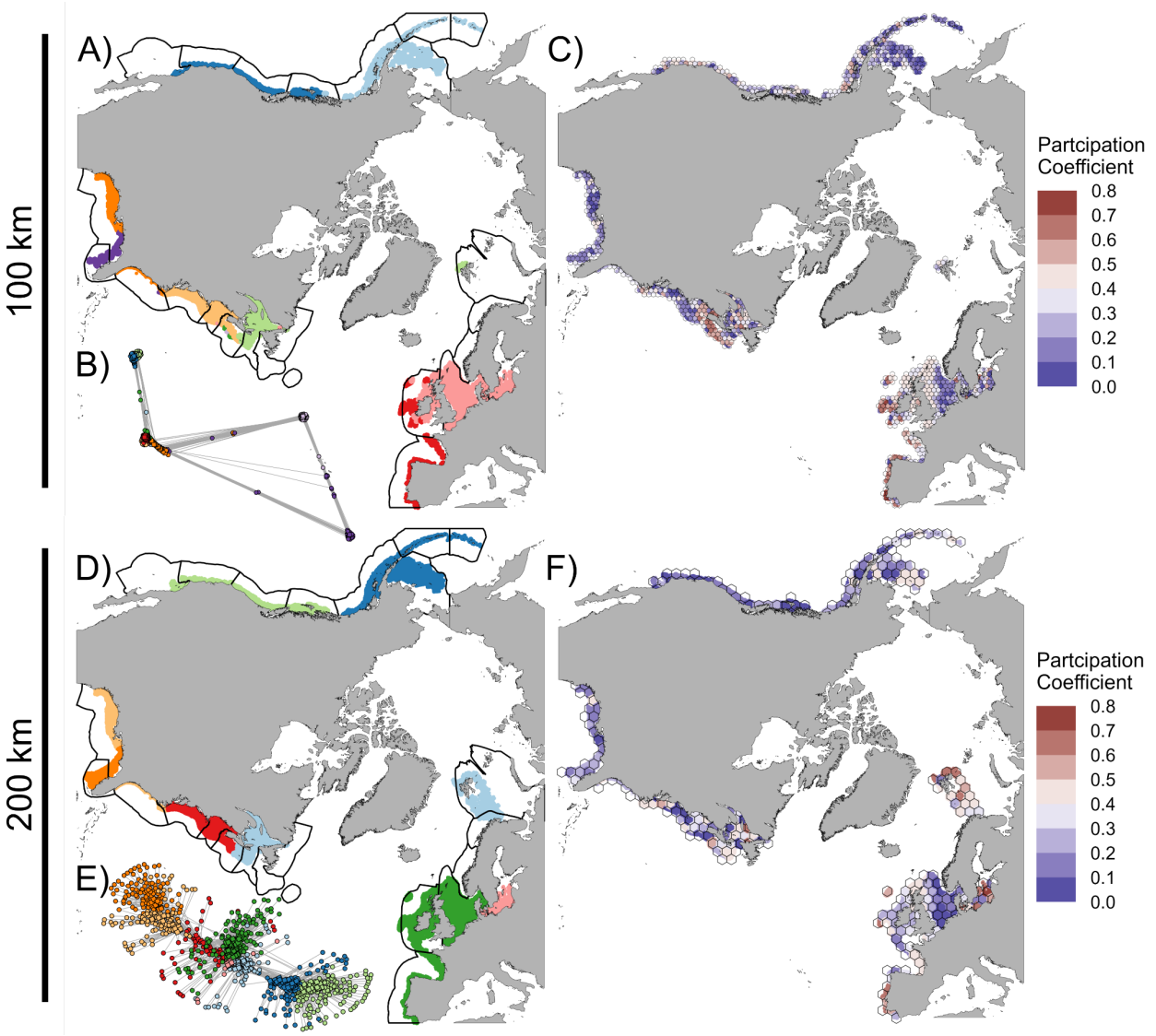


Figure 8.5 Bioregionalization based on less restrictively rarefied data points ($n_{1stQ,\ 100\ km}$ = 92; $n_{1stQ,\ 200\ km}$ = 124) at 100 km (A-C) and 200 km (D-F) grid cell resolutions using a weighted infomap algorithm based on the map equation (Rosvall & Bergstrom 2008) shown underneath ecoregions (A and D; Spalding et al. 2007). B, E) The bipartite networks (site × species) are colored by delineated bioregions, showing strong geographic clustering with the dense network structure. The weighted bioregionalization approach largely reproduces the patterns in the Leiden algorithm at 200 km spatial resolution (Figure 2) but it aggregates the European shelf and splits the Gulf of Mexico near the Mississippi River delta despite relatively low participation coefficient (PC) values; however we note that the Southeast US Shelf on the Atlantic coast (Florida to North Carolina at Cape Hatteras) is grouped together with the western Gulf of Mexico. At finer, 100 km spatial resolution, the patterns are similar to coarser 200 km resolution while missing the North Sea–Baltic Sea divide; the Norwegian/Barents Seas data were excluded during rarefaction. C, F) The participation coefficient (PC; equation 1)

visualizing the degree of shared species nodes between bioregions; potential transition zones occur at high PC values.

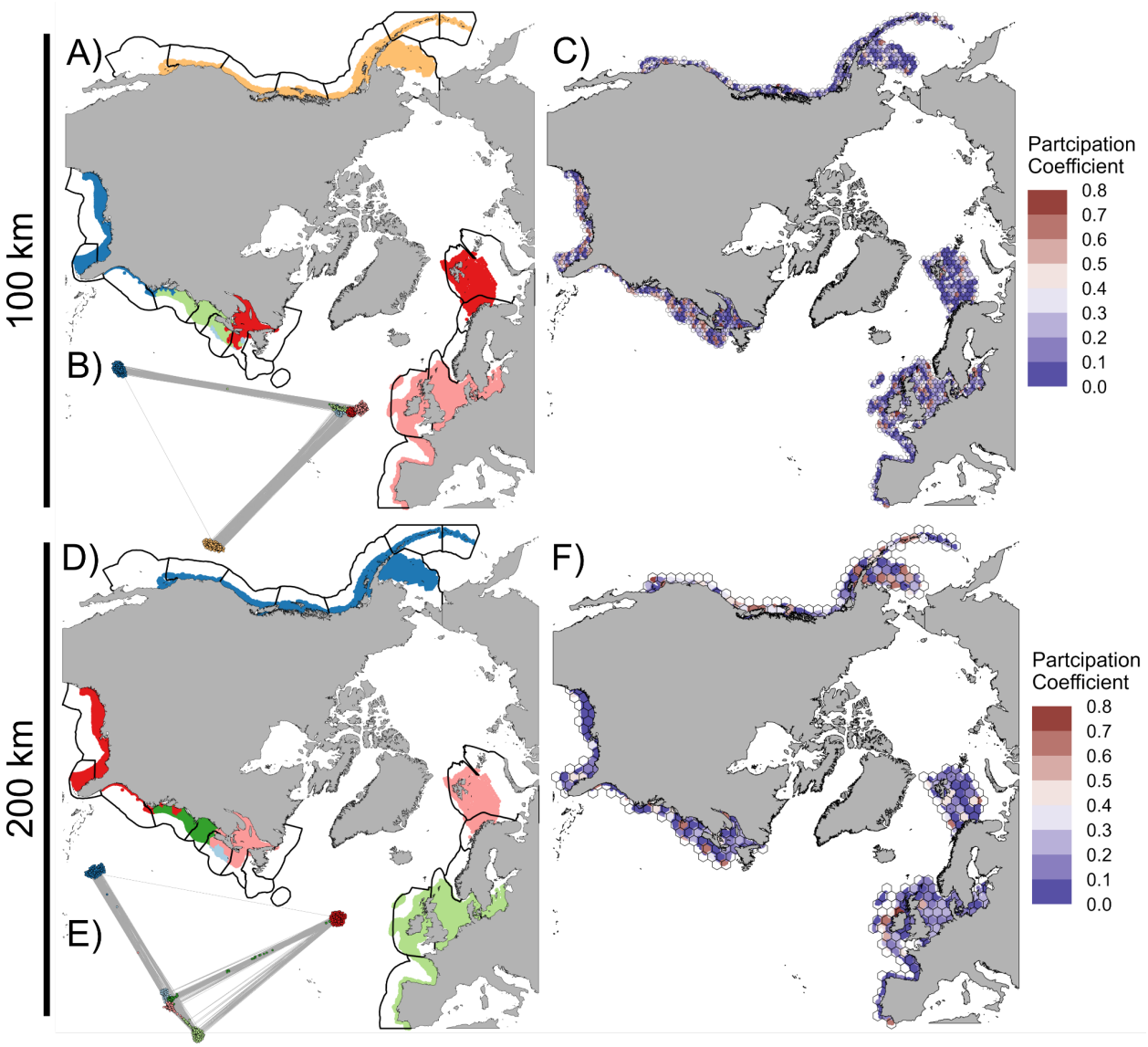


Figure 8.6 Bioregionalization based on non-rarefied data points at 100 km (A-C) and 200 km (D-F) grid cell resolutions using an unweighted infomap algorithm based on the map equation (Rosvall & Bergstrom 2008) shown underneath ecoregions (A and D; Spalding et al. 2007). B, E) The bipartite networks (site × species) colored by delineated bioregions show geographic clustering with the dense network structure, but with smaller apparent bioregions and the absorption of the entire Pacific coast as a single bioregion. The unweighted bioregionalization approach reproduces the broad differences between regions found in the Leiden algorithm (Figure 2) but aggregates the European shelf, splits the northeast Atlantic shelves into smaller local bioregions at coarser spatial resolutions, and misses the Baltic Sea–North Sea divide. C, F) The participation coefficient (PC; equation 1) visualizing the degree of shared species nodes between bioregions; potential transition zones occur at high PC values.

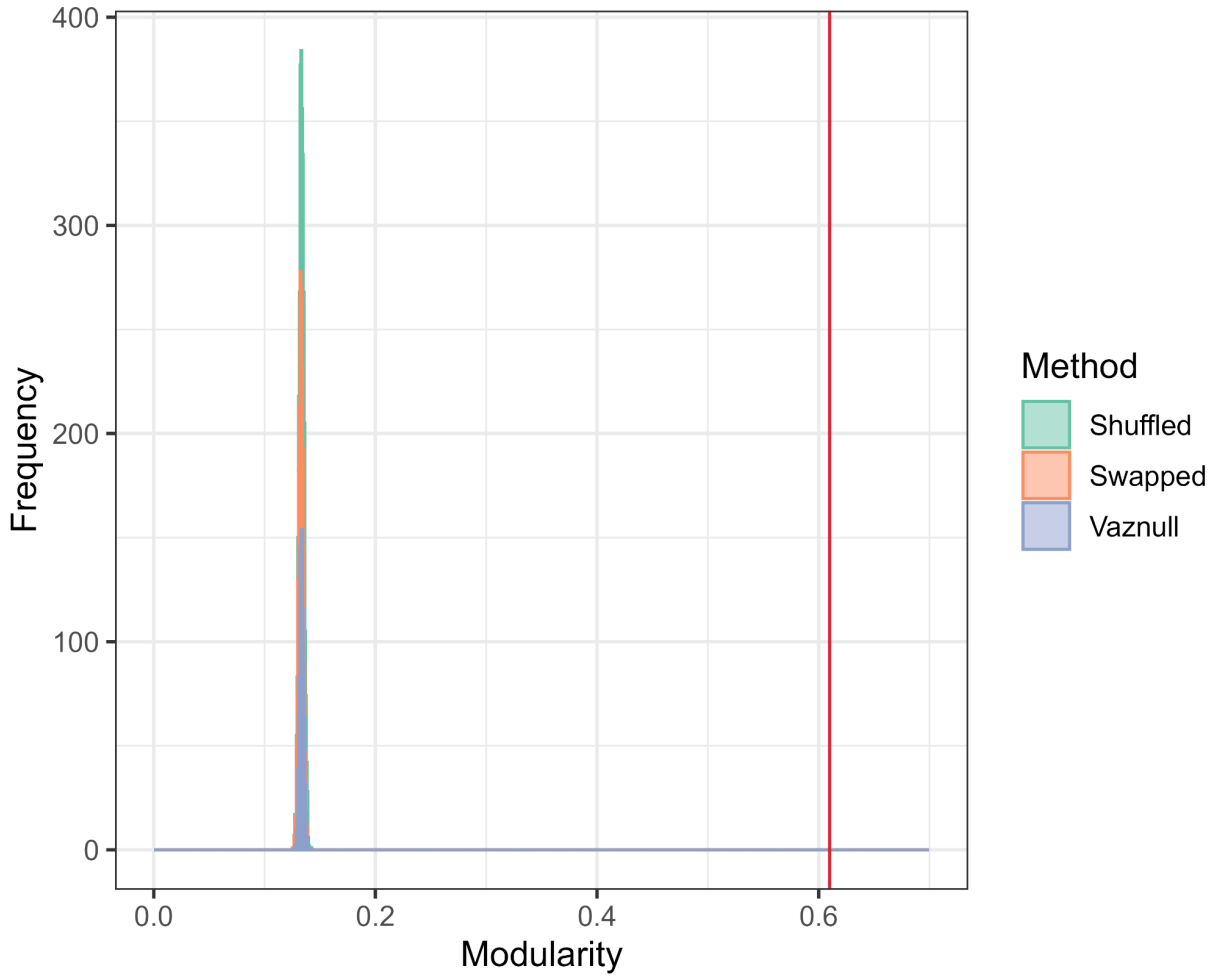


Figure 8.7 Observed modularity values from Leiden clustering ($n_{1stQ,\ 200\ km} = 124$) of bioregions (vertical red line) lies far beyond the distribution of modularity values from Leiden clustering (Traag et al. 2019) of null models across shuffled, swapped, and vaznull algorithms (Vázquez et al. 2007; Dormann et al. 2008).

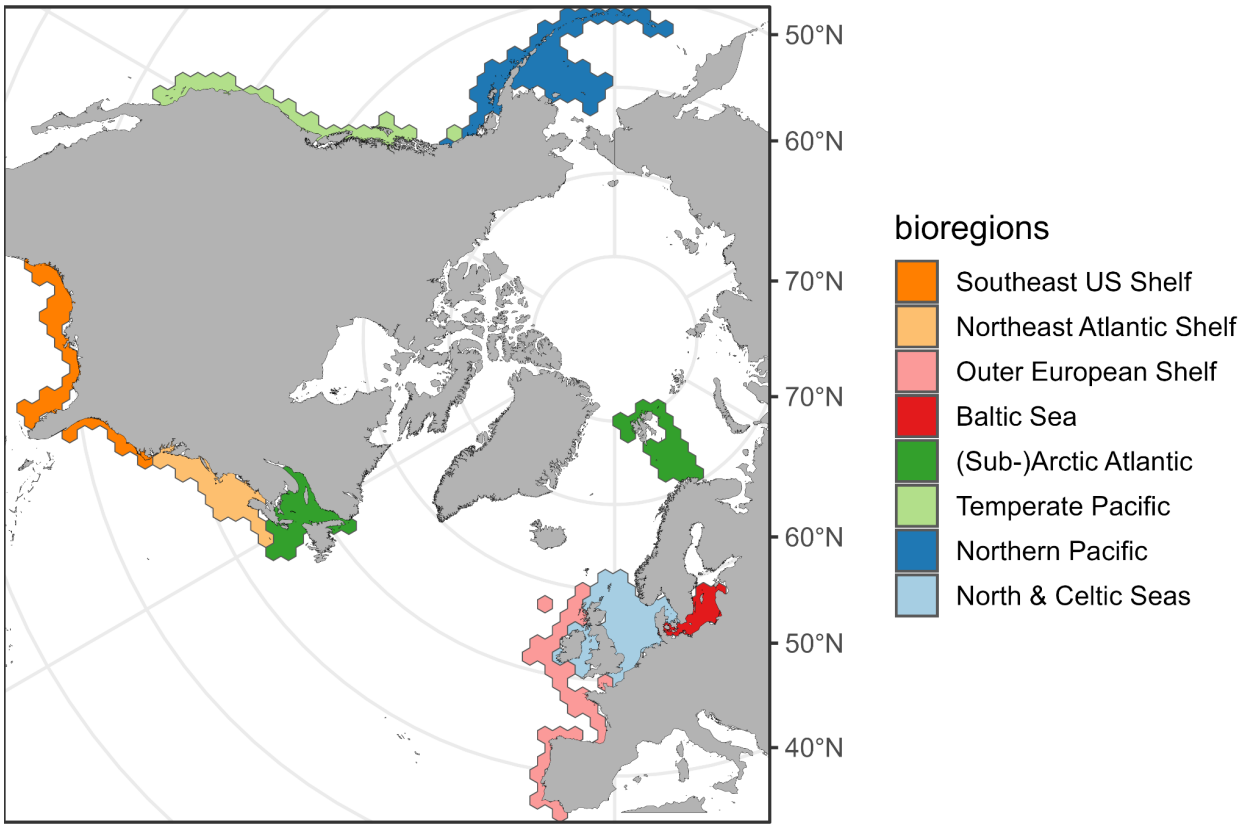


Figure 8.8 Polygonized bioregions using the Leiden algorithm ($n_{1stQ,\ 200\ km}$ = 124) to calculate total area (reported in Table 3.1). The absolute values are expected to change at finer resolutions (e.g., 100 km grid cells) and with different rarefaction depths.

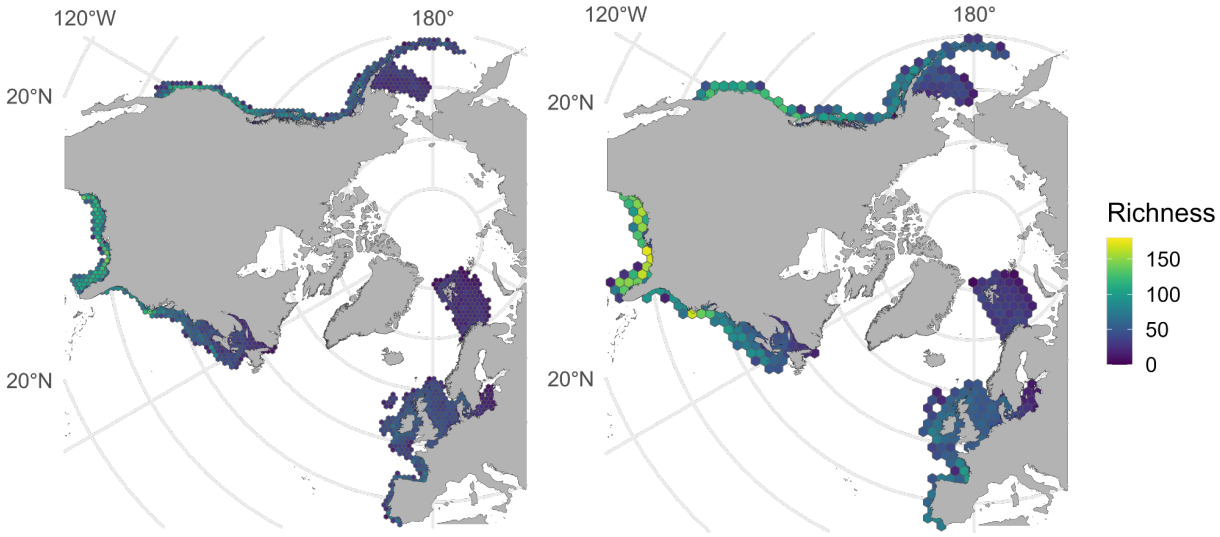


Figure 8.9 Taxonomic richness at (left) 100 km grid cell resolution and (right) 200 km grid cell resolution. Calculations were made in the EcoPhyloMapper (epm) R package (Title et al. 2022).

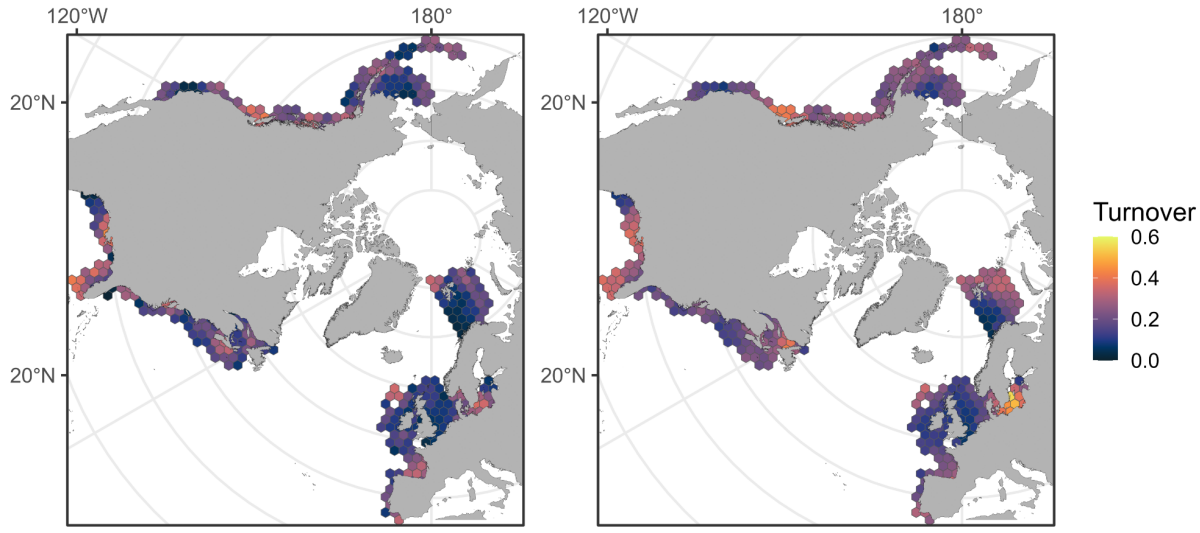


Figure 8.10 Turnover rates of phylogenetic lineages based on Rabosky et al. (2018) phylogeny of ray-finned fishes. Turnover was calculated in the EcoPhyloMapper (epm) R package (Title et al. 2022) using a moving focal window (left) 250 and (right) 500 km radius search to calculate multi-site dissimilarity at 200 km resolution (Baselga 2013). The 250 km search radius spans a median of 4 grid cells and 1-6 grid grid cells overall. The 500 km search radius spans a median of 10 grid cells and 3-18 grid grid cells overall. The smaller, more local search window illustrates finer grained patterns in lineage turnover rates; however, the broader search radius illustrates greater turnover rates in the northern Pacific (Gulf of Alaska), Pacific Northwest (Oregon, Washington, and Vancouver Shelf), eastern Gulf of Mexico, the (sub-)Arctic Atlantic region in the Gulf of St. Lawrence, and in the Baltic Sea. We note however that only 75% (859/1144) of species in the trawl record matched phylogenetic branch tips despite taxonomic harmonization in Maureaud et al. (2024).

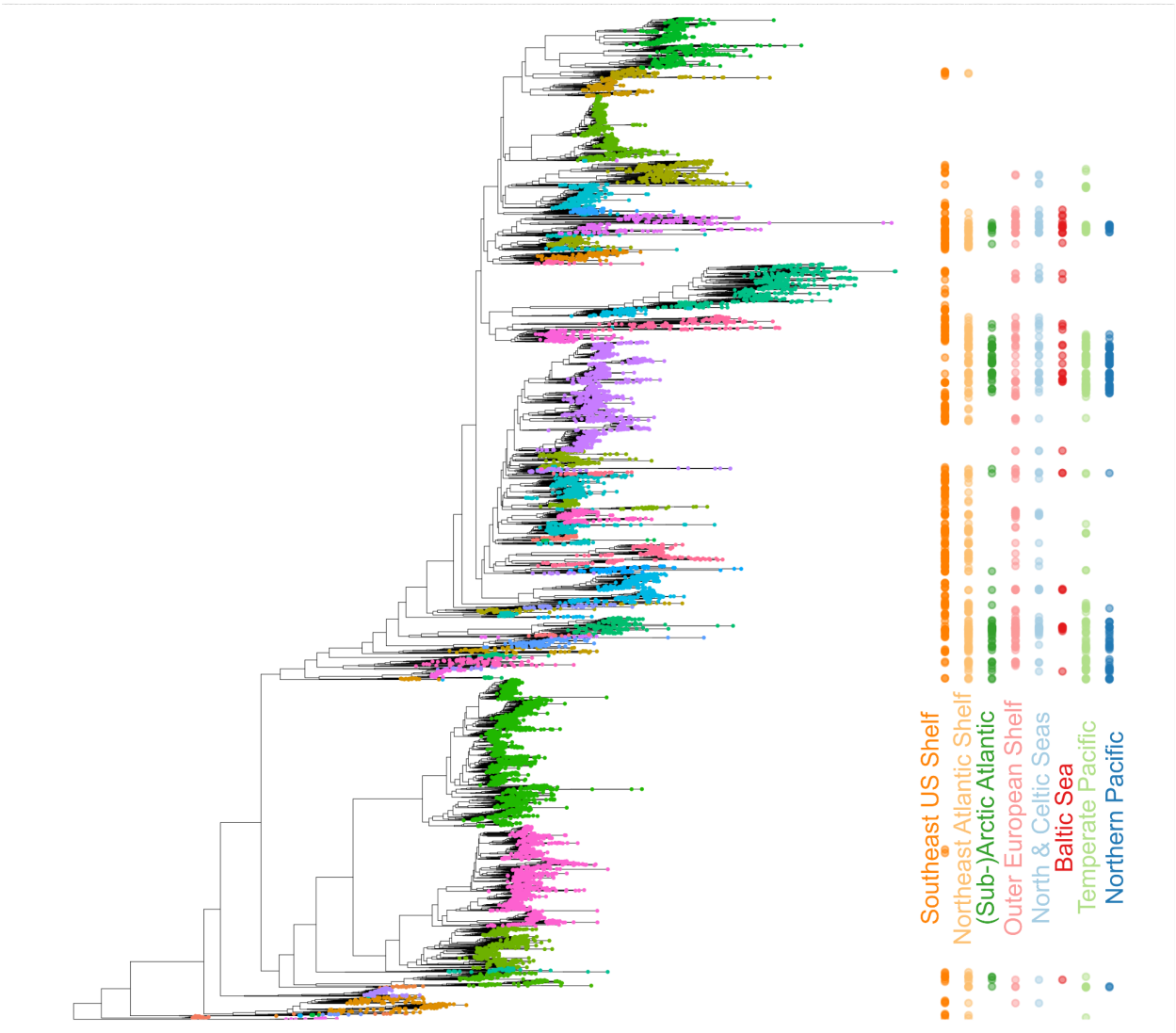


Figure 8.11 The phylogenetic coverage across the ray-finned fishes (Rabosky et al. 2018) shown as points for our proposed demersal fish bioregions. The phylogenetic tree branch tips are colored by taxonomic order. We also report the most common taxonomic orders per bioregion below in Table 8.3.

Table 8.3 The top five taxonomic orders in each bioregion, including their individual proportion and the summed proportion of each order.

Bioregion	Order	Species Proportion	Summed Proportion
Temperate Pacific	*Pleuronectiformes*	36%	89%
	Perciformes	34%	
	Gadiformes	11%	
	Chimaeriformes	6%	
	Squaliformes	3%	
Northern Pacific	*Pleuronectiformes*	48%	99%
	Perciformes	26%	
	Gadiformes	22%	
	Osmeriformes	4%	
	Salmoniformes	<1%	
Southeast US Shelf	*Eupercaria incertae sedis*	29%	72%
	Perciformes	14%	
	Pleuronectiformes	13%	
	Scombriformes	8%	
	Carangiformes	7%	
Northeast Atlantic Shelf	*Gadiformes*	30%	85%
	Pleuronectiformes	22%	
	Perciformes	18%	
	Clupeiformes	11%	
	Squaliformes	5%	

(Sub-)Arctic Atlantic	*Perciformes*	29%	96%
	Pleuronectiformes	29%	
	Gadiformes	27%	
	Osmeriformes	6%	
	Clupeiformes	5%	
Outer European Shelf	*Gadiformes*	36%	81%
	Pleuronectiformes	21%	
	Perciformes	13%	
	Carcharhiniformes	6%	
	Acanthuriformes	4%	
North & Celtic Seas	*Pleuronectiformes*	33%	90%
	Gadiformes	32%	
	Perciformes	16%	
	Clupeiformes	6%	
	Callionymiformes	4%	
Baltic Sea	*Pleuronectiformes*	41%	98%
	Gadiformes	29%	
	Clupeiformes	17%	
	Perciformes	12%	
	Mulliformes	<1%	

8.1 References

Baselga, A. (2013). Multiple site dissimilarity quantifies compositional heterogeneity among several sites, while average pairwise dissimilarity may be misleading. *Ecography*, *36*(2), 124–128. https://doi.org/10.1111/j.1600-0587.2012.00124.x

Csardi, G., Nepusz, T. (2006). The igraph software package for complex network research. *Complex Systems*, 1695. https://igraph.org.

Denderen, P. D., & Mérigot, B. (2024). FISHGLOB_data: An integrated dataset of fish biodiversity sampled with scientific bottom-trawl surveys. *Scientific Data*, *11*(1), 24. https://doi.org/10.1038/s41597-023-02866-w

Dormann, C. F., Gruber, B., & Fründ, J. (2008). *Introducing the bipartite Package: Analysing Ecological Networks*. *8*.

Maureaud, A. A., Palacios-Abrantes, J., Kitchel, Z., Mannocci, L., Pinsky, M. L., Fredston, A., Beukhof, E., Forrest, D. L., Frelat, R., Palomares, M. L. D., Pecuchet, L., Thorson, J. T., Van Meilă, M. (2007). Comparing clusterings—An information based distance. *Journal of Multivariate Analysis*, *98*(5), 873–895. https://doi.org/10.1016/j.jmva.2006.11.013

Pecuchet, L., Törnroos, A., & Lindegren, M. (2016). Patterns and drivers of fish community assembly in a large marine ecosystem. *Marine Ecology Progress Series*, *546*, 239–248. https://doi.org/10.3354/meps11613

Pini, A., & Vantini, S. (2016). The Interval Testing Procedure: A General Framework for Inference in Functional Data Analysis. *Biometrics*, *72*(3), 835–845. https://doi.org/10.1111/biom.12476

Policastro, V., Righelli, D., Carissimo, A., Cutillo, L., & Feis, I., De. (2021). ROBustness In Network (robin): An R Package for Comparison and Validation of Communities. *The R Journal*, *13*(1), 292. https://doi.org/10.32614/RJ-2021-040

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T. J., Coll, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, *559*(7714), 392–395. https://doi.org/10.1038/s41586-018-0273-1

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. *BioScience*, *57*(7), 573–583. https://doi.org/10.1641/B570707

Title, P. O., Swiderski, D. L., & Zelditch, M. L. (2022). EcoPhyloMapper: An R package for integrating geographical ranges, phylogeny and morphology. *Methods in Ecology and Evolution*, *13*(9), 1912–1922. https://doi.org/10.1111/2041-210X.13914

Traag, V. A., Waltman, L., & Van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*(1), 5233. https://doi.org/10.1038/s41598-019-41695-z

Vázquez, D. P., Melián, C. J., Williams, N. M., Blüthgen, N., Krasnov, B. R., & Poulin, R. (2007). Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, *116*(7), 1120–1127. https://doi.org/10.1111/j.0030-1299.2007.15828.x

9. Supplemental S1: Trait convergence appears widespread in marine demersal fishes

Supplement for Chapter 4.

Supplement S2 can be found as a Markdown .pdf in Zenodo:
https://doi.org/10.5281/zenodo.14989962

FISHGLOB data are pretreated to harmonize and validate taxonomic names against the World Register of Marine Species (WoRMs, 2025), plus all observations are matched with trawl-level information and units are standardized (Maureaud *et al.* 2024). We include the spatiotemporal filter from Maureaud *et al.* (2024) to remove rare trawls from 23,320 km^2 grid cells intersecting trawl data (2% threshold), this trimmed 4% of observations and 1% of species.

Table 9.1 Scientific bottom trawl surveys from Maureaud et al. (2024) including their associated acronyms shown in Figure 9.1, the quarters sampled, and the geographic region targeted.

Survey	Acronym	Quarter	Region
West coast triannual trawl survey	WCTRI	2, 4	Pacific northwest (USA/Canada)
West coast annual trawl survey	WCANN	2, 3, 4	Pacific northwest (USA/Canada)
Scotland west coast international bottom trawl survey	SWC-IBTS	1, 4	Scottish west coast
Spanish Porcupine bank trawl survey	SP-PORC	3, 4	Porcupine Bank (Irish Shelf)
Spanish north trawl survey	SP-NORTH	3, 4	Northern Spanish coast
SP-ARSA	SP-ARSA	1, 4	Gulf of Cadiz (Spain)
Southeast United States trawl survey	SEUS	2, 3, 4	Southeast United States
Scotian shelf trawl survey	SCS	1, 2, 3	Scotian Shelf (Canada)
Rockall plateau survey	ROCKALL	3	Rockall Plateau (European shelf)
Portuguese international bottom trawl survey	PT-IBTS	4	West Portuguese coast

North Sea international bottom trawl survey	NS-IBTS	1, 3	North Sea, Skagerrak
Norwegian international bottom trawl survey	NOR-IBTS	3, 4	Norwegian Seas
Northern Ireland groundfish survey	NIGFS	1	Irish Sea
Northeast United States trawls survey	NEUS	4	Northeast US coast
Ireland international groundfish survey	IE-IGFS	4	Irish Sea
Gulf of St. Lawrence south trawl survey	GSL-S	3, 4	Southern Gulf of St. Lawrence (Canada)
Gulf of St. Lawrence north trawl survey	GSL-N	2, 3, 4	Northern Gulf of St. Lawrence (Canada)
Gulf of Alaska trawl survey	GOA	2, 3	Gulf of Alaska (USA)
Gulf of Mexico trawl survey	GMEX	2, 3, 4	Northern Gulf of Mexico (USA)
France coordinated groundfish survey	FR-CGFS	4	Eastern English Channel (France)
EVHOE trawl survey	EVHOE	4	Bay of Biscay and Celtic Seas (France)
East Bering Sea trawl survey	EBS	2, 3	East Bering Sea (USA)
Department of Fisheries and Oceans west coast Vancouver Island survey	DFO-WCVI	1	Pacific northwest (Canada)
Department of Fisheries and Oceans west coast Haida Gwaii survey	DFO-WCHG	1	Pacific northwest (Canada)
Department of Fisheries and Oceans west coast Strait of Georgia survey	DFO-SoG	1	Pacific northwest (Canada)

Department of Fisheries and Oceans west coast Queen Charlotte Sound survey	DFO-QCS	1	Pacific northwest (Canada)
Department of Fisheries and Oceans west coast Hecate Strait survey	DFO-HS	1	Pacific northwest (Canada)
Baltic international trawl survey	BITS	1, 4	Baltic Sea
Aleutian Islands trawl survey	AI	2, 3	Aleutian Islands (USA)

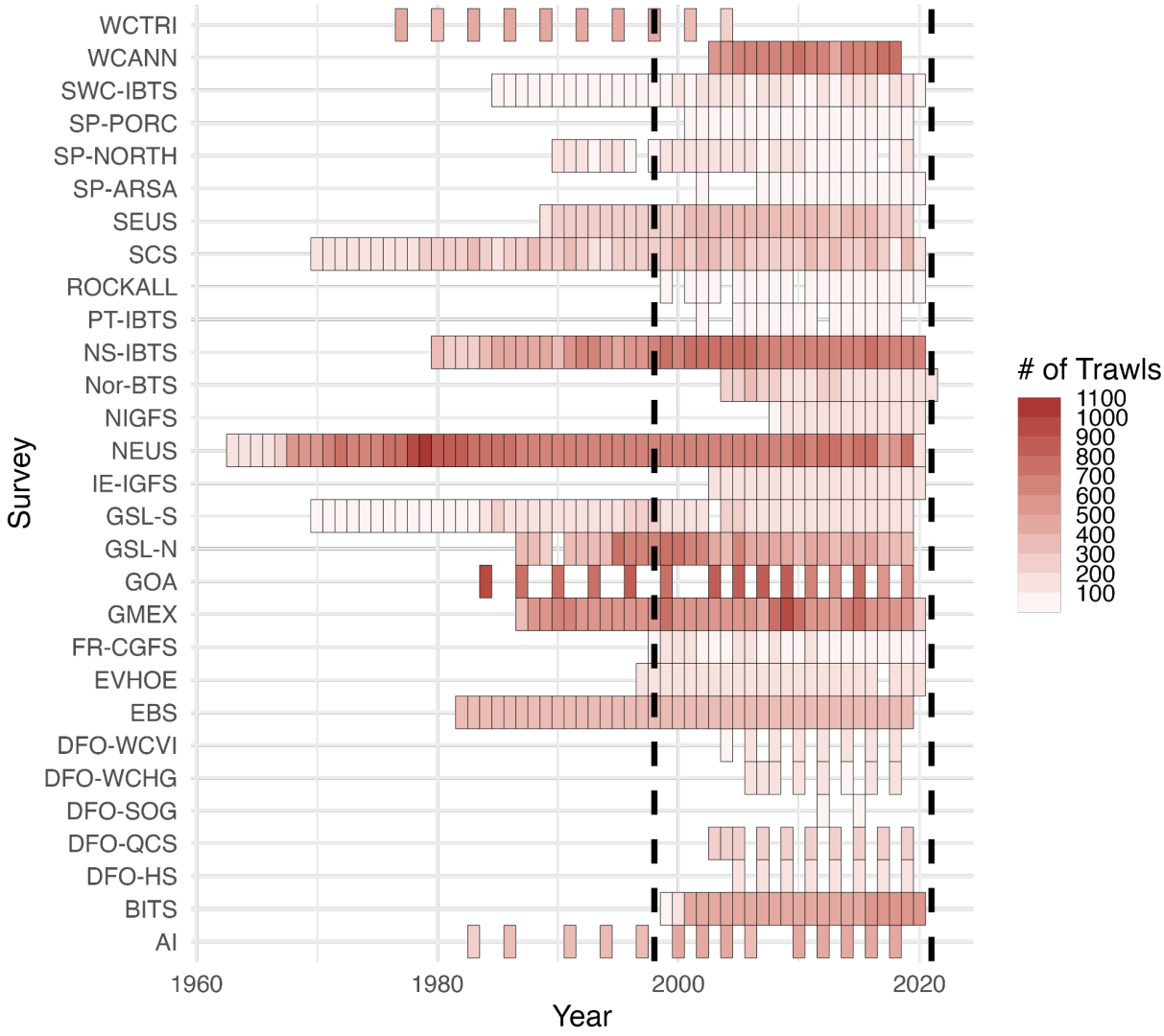


Figure 9.1 Gantt chart for number of trawls per survey across total years available. Here the vertical dashed lines bound the years included in our study (1999-2020). The majority of surveys had already begun (17/29) in 1999.

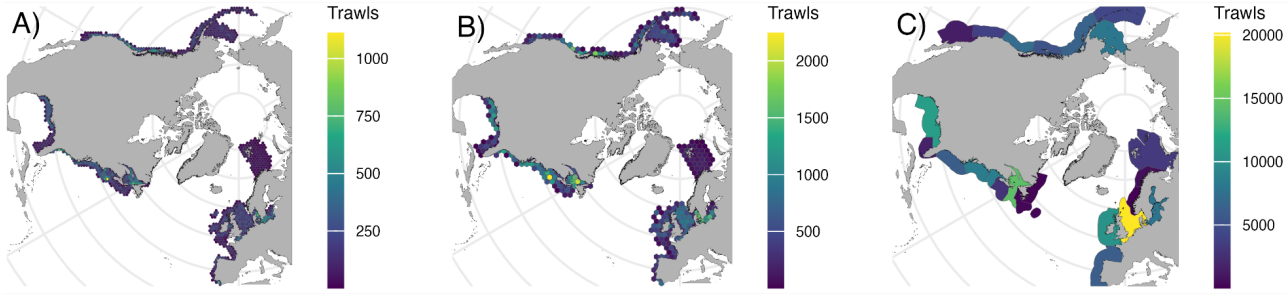


Figure 9.2 The number of scientific trawls at A) 100 km^2, B) 200 km^2 and at Ecoregion (Spalding et al. 2007) spatial resolutions. The relatively high sample intensity (1999–2021) occurs in several areas, including the Pacific northwest at the Canada/USA border, the northwest Atlantic especially north of Cape Hatteras, and in Europe across the Celtic, North, and Baltic Seas.

Table 9.2 The summary statistics of trawls per spatial resolution used for calculating trait centroids to compare against the global pool.

Resolution	1st Quartile	Median
100 km^2	90	187
200 km^2	133	357
Ecoregions	1558	5410

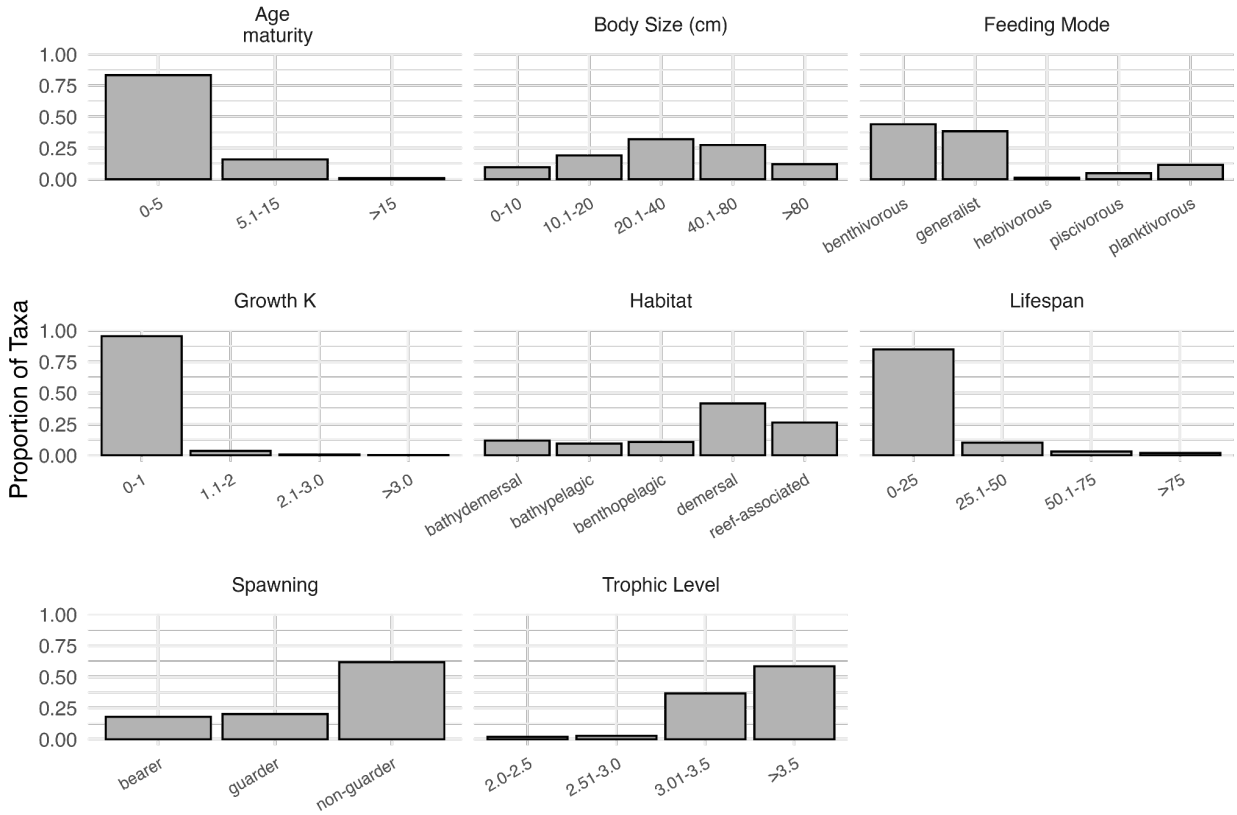


Figure 9.3 Eight fish traits used from Beukhof et al. (2019), and imputed missing data (22%) from Thorsen et al. (2023), as proportions of all taxa from all trawls. We note that lifespan is excluded for trait space analysis due to high collinearity with age at maturity (R = 0.65) and similar loadings on the first axis of a principal component analysis.

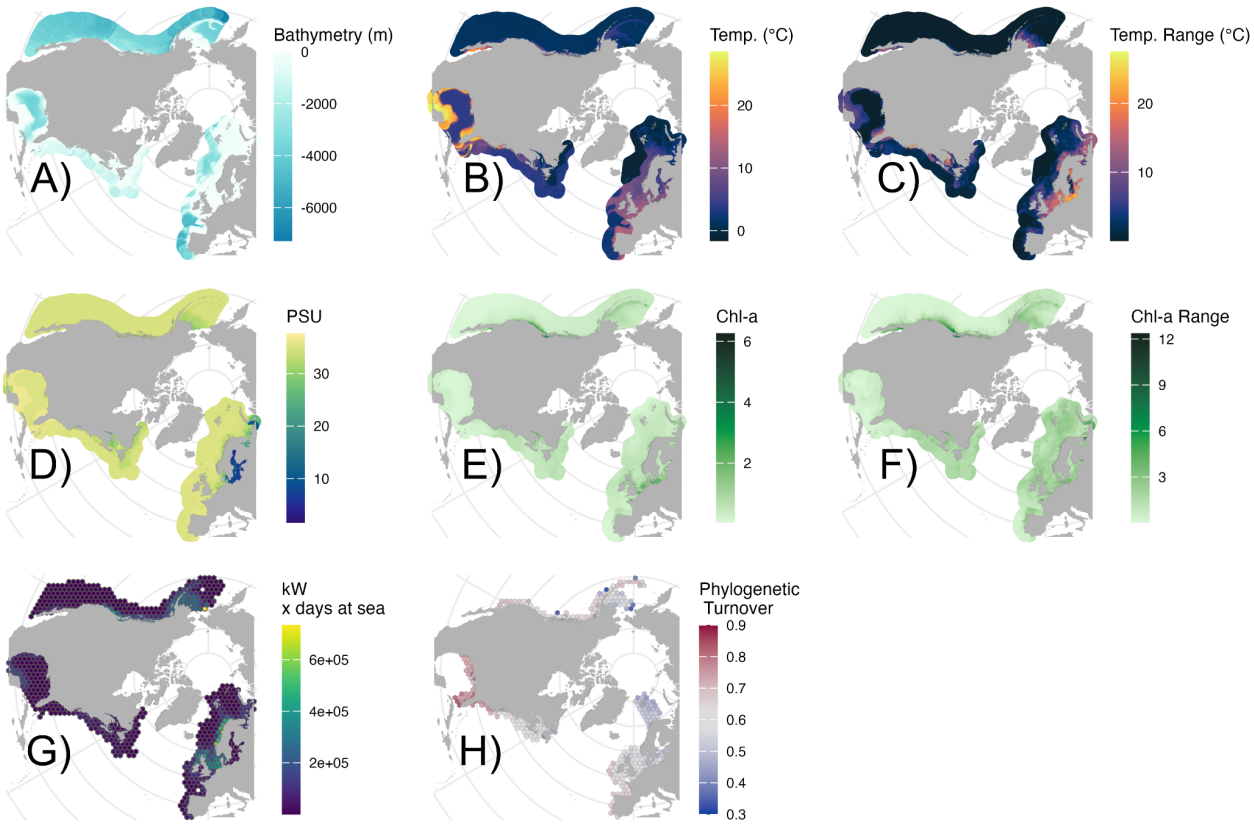


Figure 9.4 Covariates for comparing and modelling relationships to trait dissimilarities visualized across Large Marine Ecosystem where possible (all except H). Panels A–F represent environmental covariates (Assis et al. 2024) combined into broad environmental dissimilarity values including A) depth, B) mean bottom temperature, C) annual bottom range, D) mean bottom salinity, E) mean surface chlorophyll-a concentration, and F) annual chlorophyll-a range. Shown at 200 km^2, G) Average yearly demersal fishing pressure as effective effort (Rousseau et al. 2024), and H) phylogenetic turnover component of Jaccard's beta diversity as representing differences in evolutionary history in trawl data.

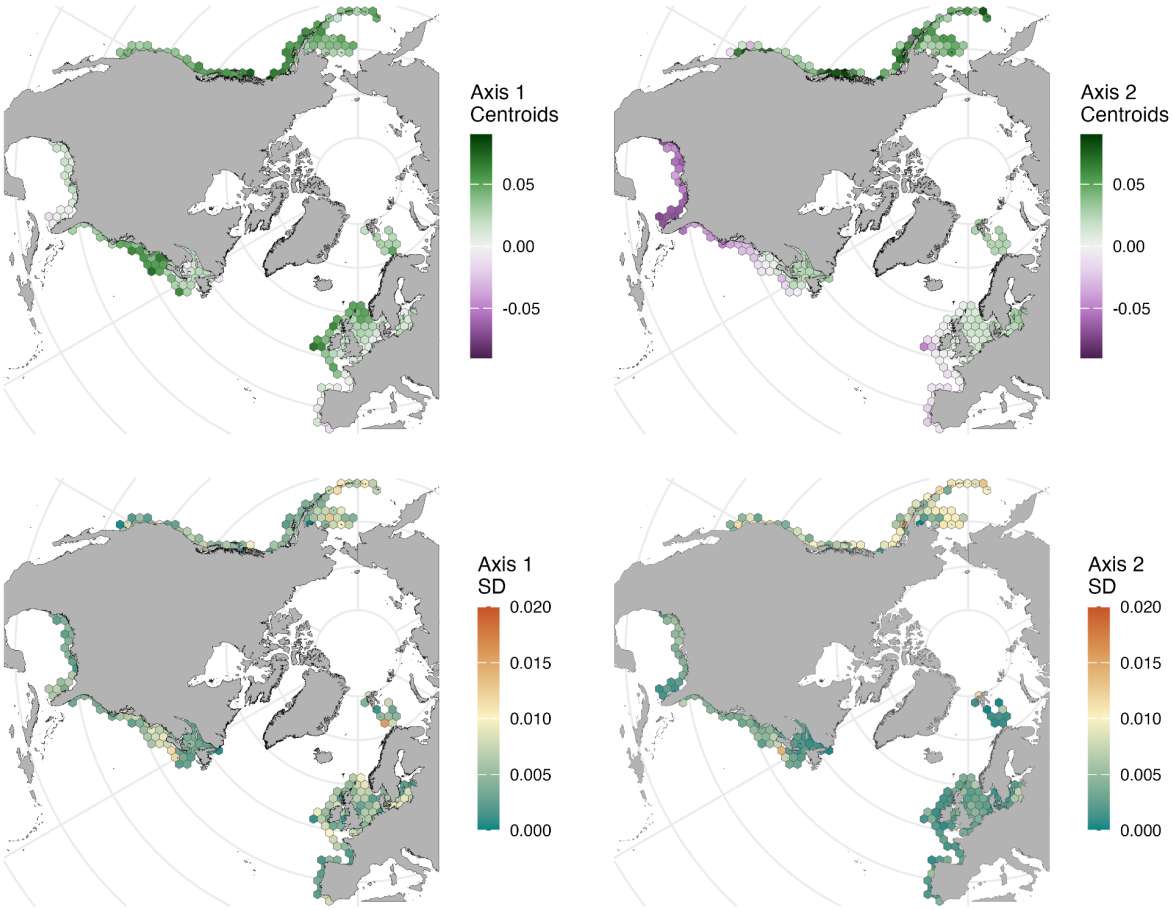


Figure 9.5 PCoA Axes 1–2 centroids including standard deviations (SD) calculated from 99 bootstraps (standard errors) randomly sampling 133 trawls per grid cell at 200 km^2.

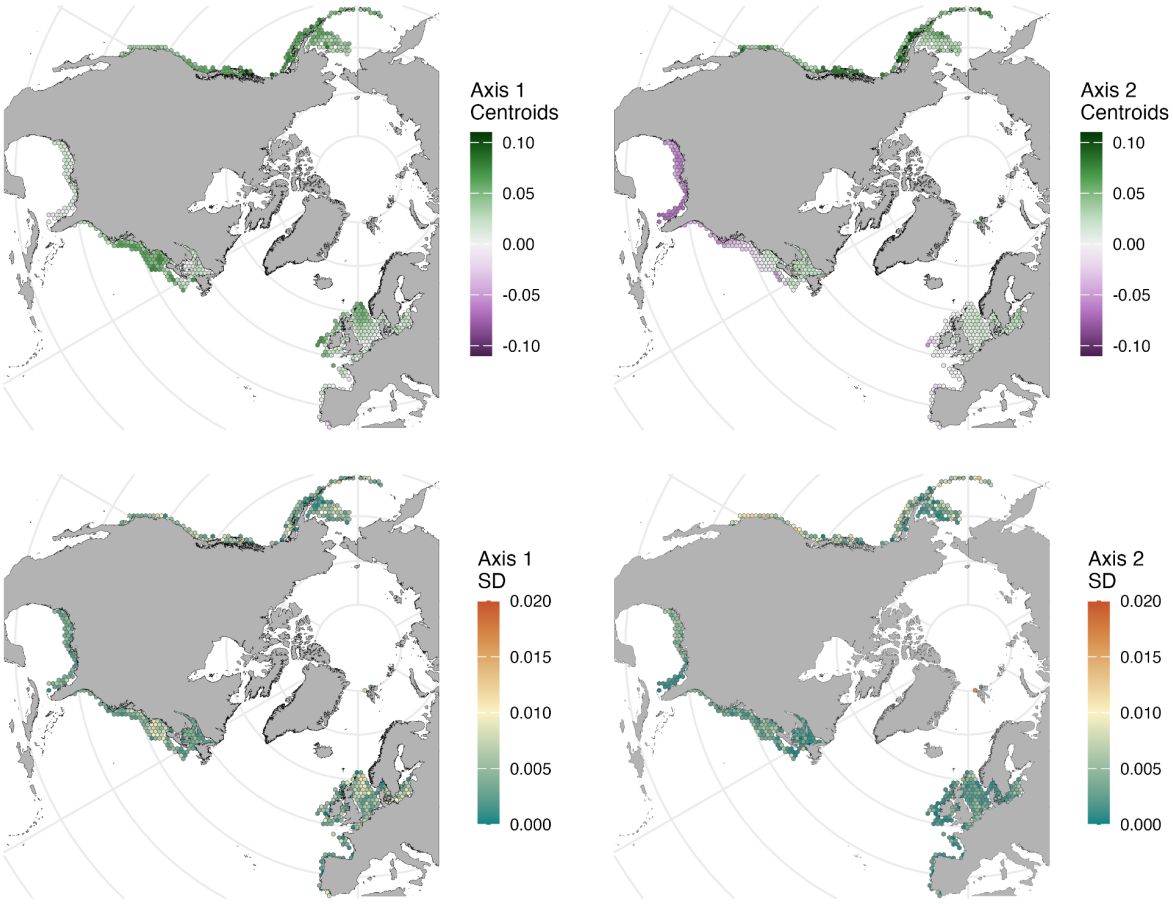


Figure 9.6 PCoA Axes 1–2 centroids including standard deviations (SD) calculated from 99 bootstraps (standard errors) calculated from 99 bootstraps randomly sampling 90 trawls per grid cell at 100 km^2.

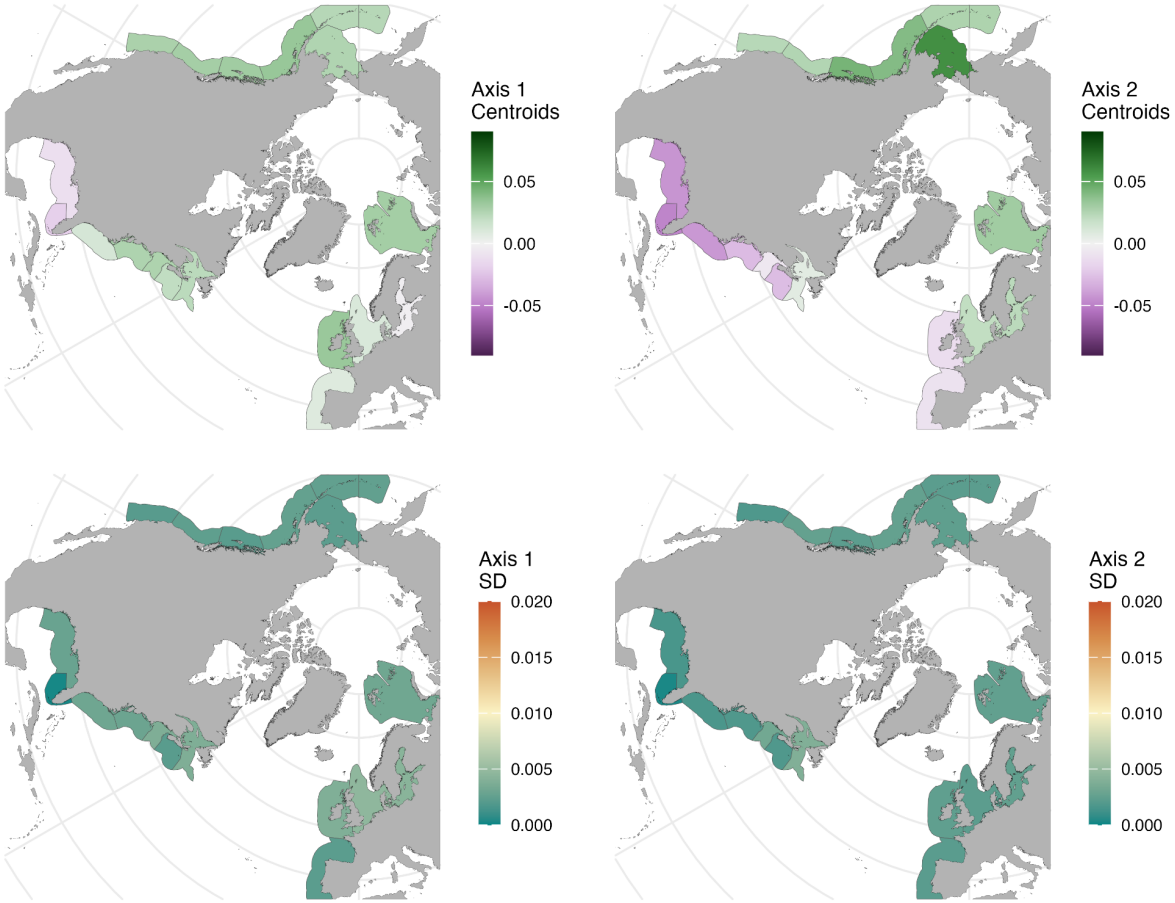


Figure 9.7 PCoA Axes 1–2 centroids including standard deviations (SD) calculated from 99 bootstraps (standard errors) calculated from 99 bootstraps randomly sampling 1558 trawls per Ecoregion (Spalding et al. 2007).

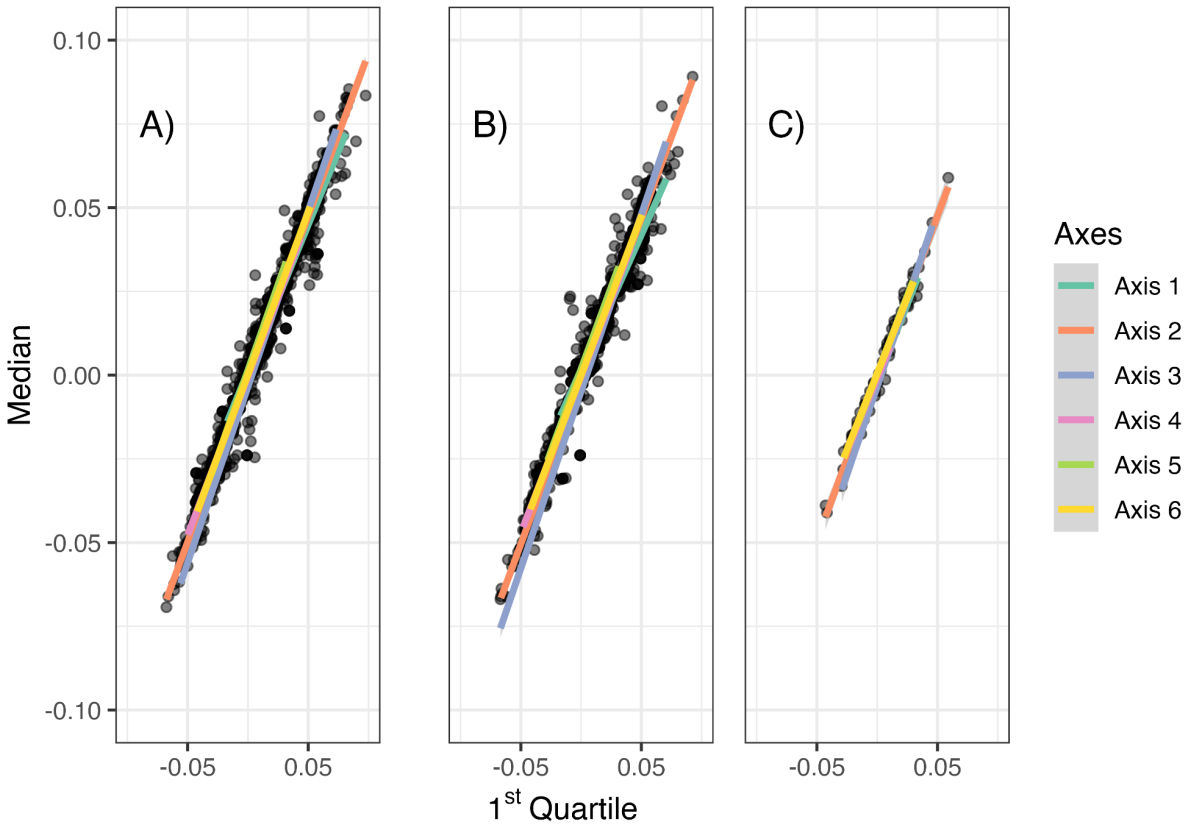


Figure 9.8 Marine fish trait space position (Axes 1–6) is robust to spatial size and sampling intensity. Trait space centroid position derived from principal coordinate analysis of pairwise Gower dissimilarities (gawdis function) is highly similar at A) 100 km^2, B) 200 km^2, and C) Ecoregion (Spalding et al. 2007) spatial scales when analyzed with equal number of trawls per grid cell filtered for less (1st Quartile) and more (Median) restrictive sampling intensities.

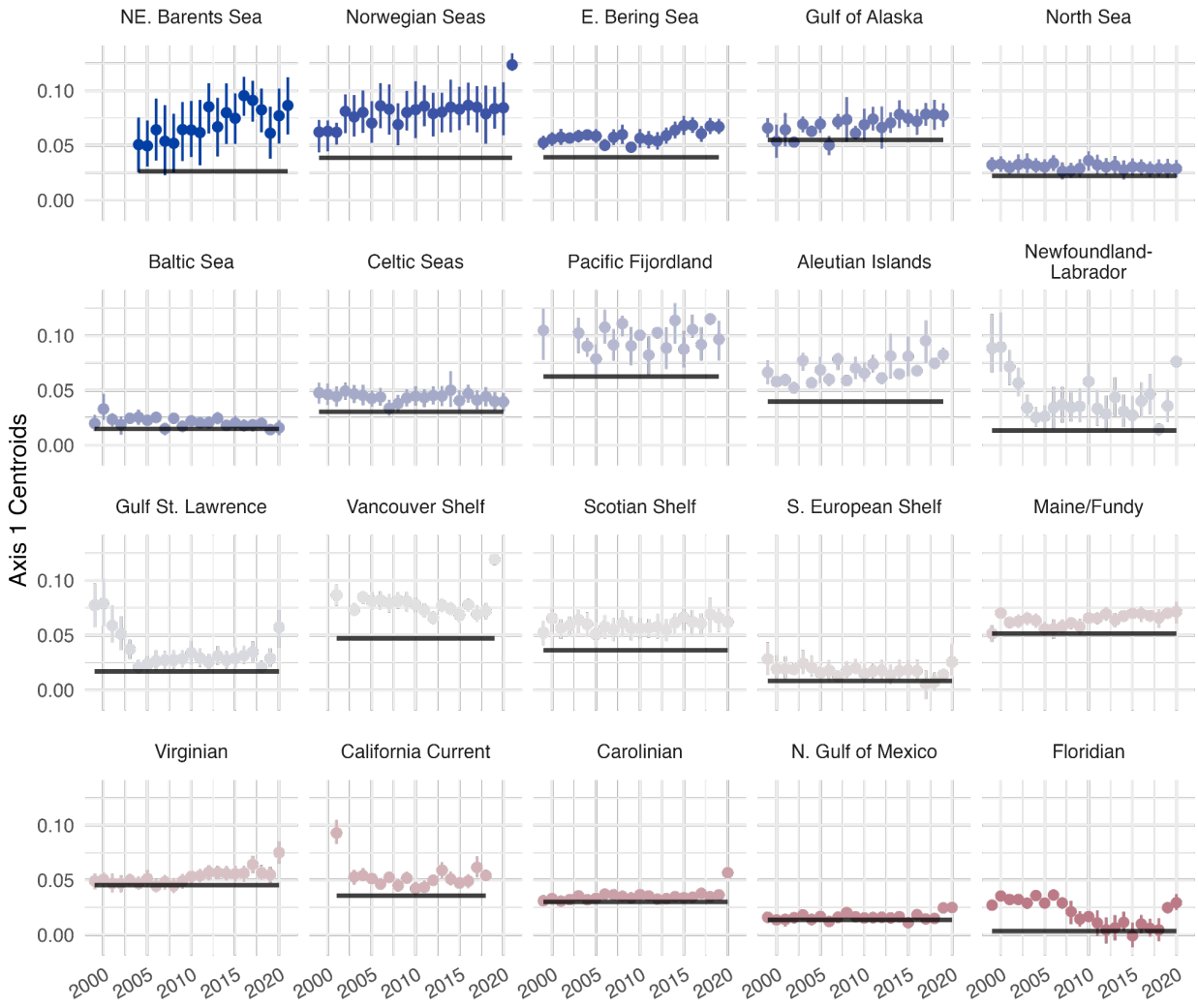


Figure 9.9 Yearly average trait space centroids calculated for 200 km^2 grid cells sampled to the 1st quartile number of trawls per grid cell per year (36 trawls) and bootstrapped for 99 repetitions. Points are colored by the mean latitude of their grid cells, which are averaged for each year by the Ecoregion they are within (facets). The trait space centroids for each Ecoregion based on randomly sampled trawls from all years (Figure 2, 200 km^2 and 133 trawls per grid cell) are shown as black horizontal lines. Yearly trait space is generally stable and mirrors values from all years combined (solid line) with notable exceptions in the northwest Atlantic Ocean shelves, where the Gulf of St. Lawrence and Newfoundland-Labrador diverged in trait composition starting around 2005, on the Floridian coast (eastern Gulf of Mexico) in 2007, and in the Barents and Norwegian Seas in 2005.

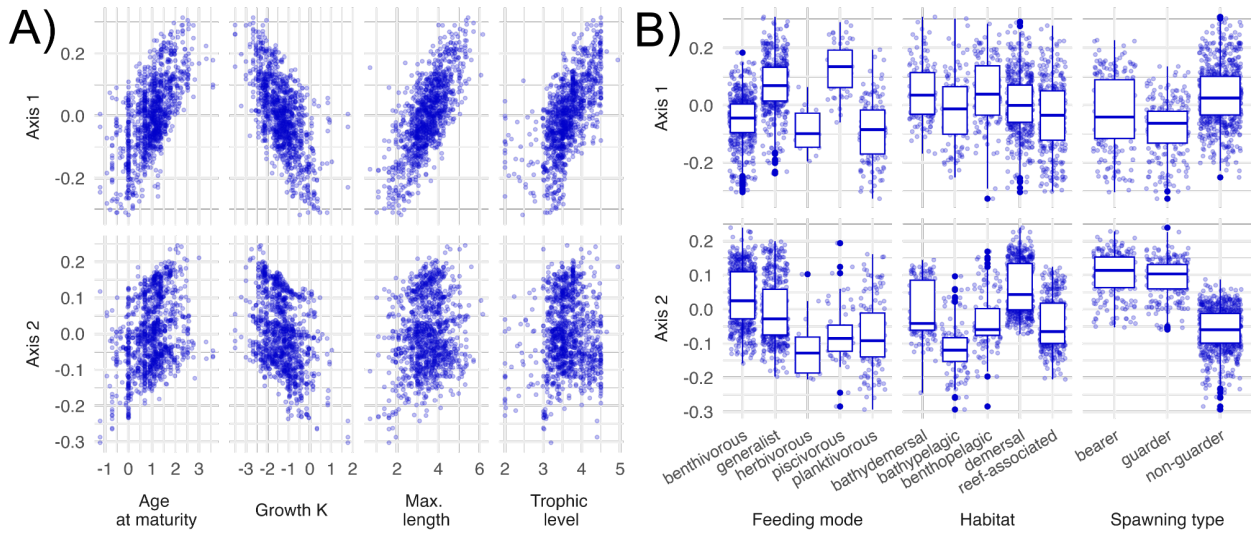


Figure 9.10 A "fast-slow" continuum of species traits. Negative axis values correspond to species traits frequently found at (sub)tropical latitudes which are fast growing and maturing at lower trophic levels, smaller in size and often herbivorous and planktivorous; positive axis values correspond to species traits frequently found at temperate/polar latitudes which are slow growing and maturing within higher trophic levels, are larger in size and often have generalist feeding strategies. This displays the relationship between species trait space centroids (Axes 1–2) to A) each continuous trait (age at maturity, maximum age, von Bertalanffy growth coefficient— "Growth K"— and maximum length are log-transformed) and B) discrete traits by factor level.

Table 9.3 The 24 Ecoregions (Spalding et al. 2007) which intersect trawl point data (n = 24). The group column denotes ecoregion names used in the analysis and asterisks highlight ecoregions which were gathered together used for visualization because they are adjacent and displayed identical trait space patterns.

Ecoregion	Group
Aleutian Islands	Aleutian Islands
Baltic Sea	Baltic Sea
Carolinian	Carolinian
Celtic Seas	Celtic Seas
Eastern Bering Sea	E. Bering Sea
Floridian	Floridian
Gulf of Alaska	Gulf of Alaska
Gulf of Maine/Bay of Fundy	Maine/Fundy
Gulf of St. Lawrence - Eastern Scotian Shelf	Gulf of St. Lawrence
North American Pacific Fijordland	Pacific Fjordland
Northern California	California Current*
Northern Grand Banks - Southern Labrador	Newfoundland-Labrador*
Northern Gulf of Mexico	Gulf of Mexico
Oregon, Washington, Vancouver Coast and Shelf	Vancouver Shelf*
Puget Trough/Georgia Basin	Vancouver Shelf*
Scotian Shelf	Scotian Shelf
South European Atlantic Shelf	S. European Shelf
Southern California Bight	California Current*
Southern Grand Banks - South Newfoundland	Newfoundland-Labrador*
Virginian	Virginian
North Sea	North Sea
Southern Norway	Norwegian Seas*
Northern Norway and Finnmark	Norwegian Seas*
North and East Barents Sea	NE. Barents Sea

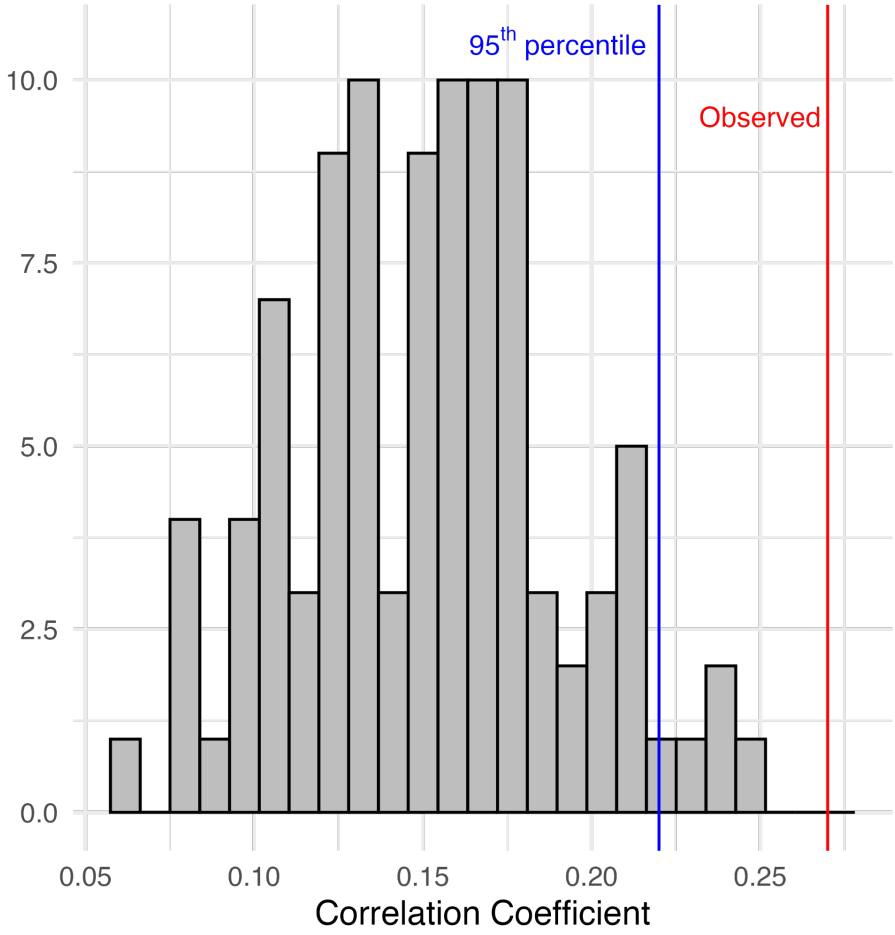


Figure 9.11 The relationship (Pearson's R) between trait space and latitude is stronger than by a null model that maintains species richness. The observed correlation coefficient is greater than the 95th-percentile derived from 99 bootstrap replications of random trawl resampling (200 km^2, 133 trawls per grid cell) from the null model. This confirms that species richness alone does not explain the underlying trait space values.

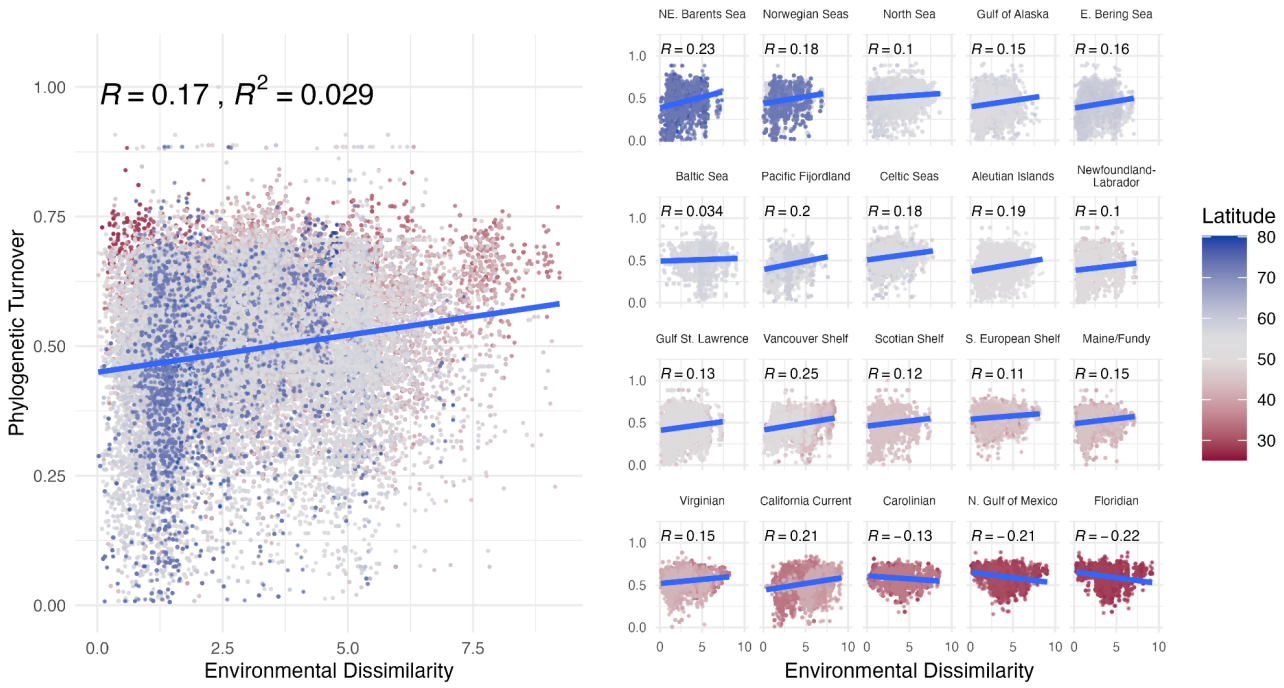


Figure 9.12 Environmental dissimilarity has a weak and mixed relationship to (Jaccard's) dissimilarity in lineage turnover. Each point represents grid cell pairwise comparisons for all grid cells (left) and within Ecoregions (right; Spalding et al. 2007), colored by latitude. Pearson's correlation coefficient (R) is reported in panels.

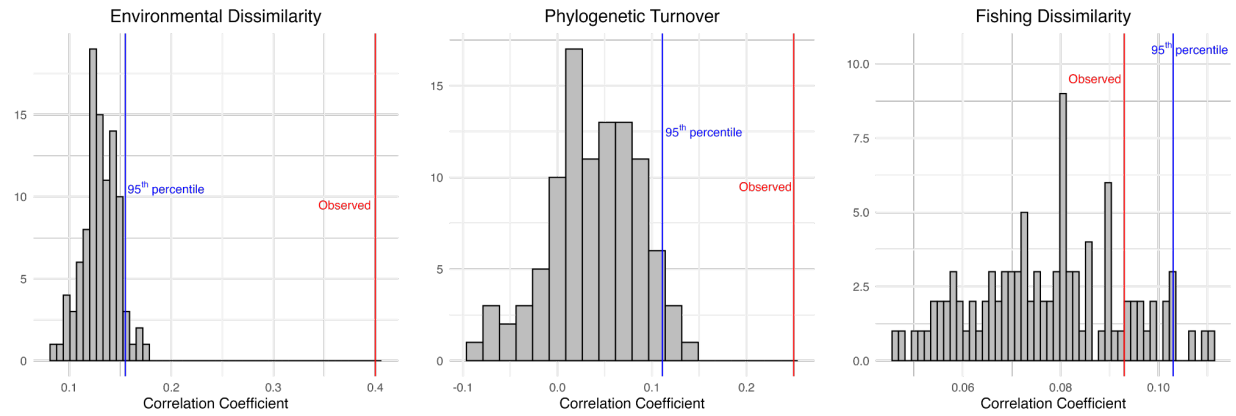


Figure 9.13 The relationship (Pearson's R) between observed trait dissimilarities and environmental dissimilarities and phylogenetic turnover are stronger than expected under neutral evolution (Brownian motion simulated traits), but weak in fishing dissimilarity. Histograms display 99 bootstrap replications randomly sampling simulated neutral traits at 200 km^2 and with 133 random trawls per grid cell. The 95th percentile of correlation coefficients between neutral trait dissimilarities and each facet are shown in blue with observed value in red.

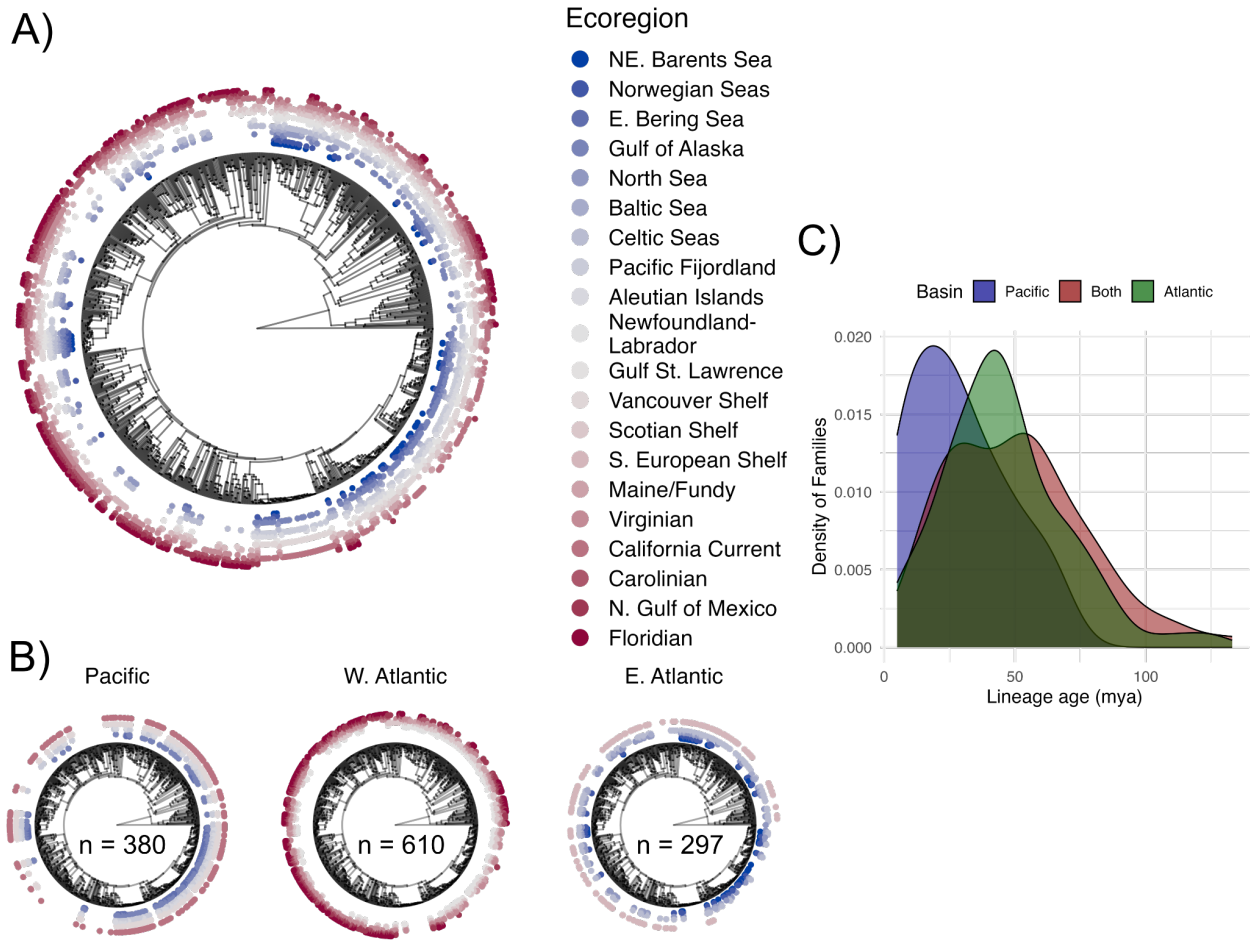


Figure 9. 14 A) The pruned phylogenetic tree (Chang et al. 2019) matching species found in the trawl data (n = 1150) visualizing the distribution of all lineages and B) spread across each ocean region. Lineage occurrence is colored by the respective Ecoregion and arranged by mean latitude. Broadly, lineages found in temperate and polar regions are nested from (sub)tropical zones— this pattern is strongest in the Pacific, whereas many more lineages in the Atlantic are differentiated by latitude. C) The time to diverge for fish families is calculated as the distance of the most recent common ancestor to the tree root, subtracted from the maximum branch distance to obtain ages.

9.1 References

Assis, J., Fernández Bejarano, S. J., Salazar, V. W., Schepers, L., Gouvêa, L., Fragkopoulou, E., Leclercq, F., Vanhoorne, B., Tyberghein, L., Serrão, E. A., Verbruggen, H., & De Clerck, O. (2024). BIO-ORACLE v3.0. Pushing marine data layers to the CMIP6 Earth System Models of climate change research. *Global Ecology and Biogeography*, *33*(4), e13813. https://doi.org/10.1111/geb.13813

Beukhof, E., Dencker, T. S., Palomares, M. L. D., Maureaud, A. (2019). A trait collection of marine fish species from North Atlantic and Northeast Pacific continental shelf seas. *Pangaea, https://doi.org/10.1594/PANGAEA.900866.*

Chang, J., Rabosky, D. L., Smith, S. A., & Alfaro, M. E. (2019). An R package and online resource for macroevolutionary studies using the ray-finned fish tree of life. *Methods in Ecology and Evolution*, *10*(7), 1118–1124. https://doi.org/10.1111/2041-210X.13182

Maureaud, A. A., Palacios-Abrantes, J., Kitchel, Z., Mannocci, L., Pinsky, M. L., Fredston, A., Beukhof, E., Forrest, D. L., Frelat, R., Palomares, M. L. D., Pecuchet, L., Thorson, J. T., Van Denderen, P. D., & Mérigot, B. (2024). FISHGLOB_data: An integrated dataset of fish biodiversity sampled with scientific bottom-trawl surveys. *Scientific Data*, *11*(1), 24. https://doi.org/10.1038/s41597-023-02866-w

Rousseau, Y., Blanchard, J. L., Novaglio, C., Pinnell, K. A., Tittensor, D. P., Watson, R. A., & Ye, Y. (2024). A database of mapped global fishing activity 1950–2017. *Scientific Data*, *11*(1), 48. https://doi.org/10.1038/s41597-023-02824-6

Sherman, K. (1991). The Large Marine Ecosystem Concept: Research and Management Strategy for Living Marine Resources. *Ecological Applications*, *1*(4), 349–360. https://doi.org/10.2307/1941896

Simpson, G.L. (2018). R Package: gratia. Ggplot-based graphics and other useful functions for GAMs fitted using Mgcv, 0.9-2 (Ggplot-based graphics and utility functions for working with GAMs fitted using the mgcv package).

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. *BioScience*, *57*(7), 573–583. https://doi.org/10.1641/B570707

Thorson, J. T., Maureaud, A. A., Frelat, R., Mérigot, B., Bigman, J. S., Friedman, S. T., Palomares, M. L. D., Pinsky, M. L., Price, S. A., & Wainwright, P. (2023). Identifying direct and indirect associations among traits by merging phylogenetic comparative methods and structural equation models. *Methods in Ecology and Evolution*, *14*(5), 1259–1275. https://doi.org/10.1111/2041-210X.14076

Wood, S. N. (2013). On p-values for smooth components of an extended generalized additive model. *Biometrika*, *100*(1), 221–228. https://doi.org/10.1093/biomet/ass048

Wood, S.N. (2017). Generalized additive models: An introduction with R. *CRC Press*.

List of Figures

List of Tables

Acknowledgements

I owe my deepest thanks to Dr. Marco Scotti, for giving me a chance in the Ecosystem Modelling Group. Without him, and without his boundless generosity, this work would not have been possible.

To Dr. Thorsten Reusch, I am also thankful for acting as my supervisor at GEOMAR and for modelling a consummate professional scientist. To Dr. Rainer Froese, thank you for your knowledge sharing and support. The members of the Ecosystem Modelling Group made all this worthwhile: Dr. Maysa Ito, Dr. Léa Joly, Dr. Eva Papaioannou, Dr. Marcela Nascimento, Anna Waffender, and many more students and visiting fellows— I thank you. To the scientists and students in RD3, thank you for supporting a top-tier research environment.

To the GEOMAR maintenance staff, who never called security on me for playing violin in the office after hours, I also thank you.

To my family, and to Isabelle, no thanks would be too much.

Declaration

I hereby declare that:

1. apart from the supervisor's guidance, the content and design of the dissertation is my own, original work;
2. this dissertation has not already been submitted either partially or wholly as part of a doctoral degree to another examining body;
3. each manuscript chapter has been submitted for scientific publication in peer-reviewed journals;
4. this dissertation been prepared subject to the Rules of Good Scientific Practice of the German Research Foundation;
5. I have never had an academic degree has ever been withdrawn.

[x]

March, 12, 2025