Journal: ICES

Article Type: Original Article

Title: "Combining fisheries surveys to inform marine species distribution modelling"

Authors and Affiliations:

Meadhbh Moriarty[1,2,3,4]

Suresh A. Sethi[3,5]

Debbi Pedreschi[6]

T. Scott Smeltz[2,3]

Chris McGonigle[1]

Bradley P. Harris[3]

Nathan Wolf[3]

Simon P.R. Greenstreet[4]

1. Geography and Environmental Sciences Research Institute, Ulster University, United Kingdom
2. New York Cooperative Fish and Wildlife Research Unit, Cornell University, USA
3. Fisheries, Aquatic Science, and Technology (FAST) Laboratory, Alaska Pacific University, Anchorage, AK, USA
4. Marine Scotland Science, 375 Victoria Rd, Aberdeen AB11 9DB
5. U.S. Geological Survey, New York Cooperative Fish and Wildlife Research Unit, Cornell University, USA
6. Marine Institute, Rinville, Oranmore, Co. Galway, Ireland.

Abstract (228 words)

Ecosystem-scale examination of fish communities typically involves creating spatiotemporally-explicit relative abundance distribution maps using data derived from multiple fishery-independent surveys. However, survey sampling performance varies by vessel and sampling gear, which may influence estimated species distribution patterns. Using generalised additive mixed models, the effect of different gear-vessel combinations on relative abundance estimates at length are investigated using European fisheries-independent groundfish survey data. We constructed a modelling framework for evaluating relative efficiency of multiple survey gear-vessel combinations and examined 19 disparate surveys for 254 species-length combinations across the northeast Atlantic. Space-time variables explained the majority of the variation in catches when combining data across different gears or vessels for 181 of 254 species-length cases, indicating that for many species, models could successfully characterize distribution patterns by combining data from disparate surveys. Variables controlling for catch efficiency differences across gear-vessel combinations explained substantial variation in catches for 127 of 254 species-length data sets. In such cases, models that fail to control for gear efficiencies across surveys can mask changes in the spatial distribution of species. Estimated relative differences in catch efficiencies grouped strongly by gear type, but did not exhibit a clear pattern across species' functional forms, suggesting difficulty in predicting the potential impact of gear efficiency differences when combining data across surveys to assess species' distributions and highlighting the importance of modelling approaches that can control for gear differences.

Keywords

Catchability; gear efficiency; fisheries independent assessment; Generalised Additive Mixed Model (GAMM); survey standardisation; species distribution modelling

1. Introduction

As ecosystem-based management in the marine environment advances, fisheries policies increasingly require consideration of both target and non-target species in assessing the state of fisheries and fishing impacts on marine ecosystems (e.g. the European Union (EU) Marine Strategy Framework Directive (MSFD; EC 2008; 2010; 2017), Common Fisheries Policy (CFP; EC, 2013), United States Magnuson–Stevens Fishery Conservation and Management Act (US, 1996; 2006), etc.). This transition to ecosystem-based management has led to a need for greater understanding and detailed information on the distribution of a broad spectrum of fish species across large spatial scales, such as large marine ecosystems or ecoregions (Kelley & Sherman 2018).

Fisheries-independent groundfish surveys sample both commercial and non-target fish species, often providing the only data source available to estimate relative abundances for non-commercial species (Poos et al., 2013). These surveys tend to be discrete monitoring programmes, operating at local scales usually associated with the exclusive economic zones of countries managing the surveys. To obtain information on fish distributions at large marine ecosystems scales, therefore, requires integration across national jurisdictional boundaries and multiple disparate surveys that may differ in terms of spatial coverage, survey vessel, season, types of fishing gear, and survey protocols. Amalgamating such data into a single cohesive analysis is difficult because of potential differences in gear efficiency among different length-classes and species of fish (Fraser et al., 2007; Walker et al., 2017), types of survey gear, and vessels that vary in their fishing power (Dann et al., 2005).

Estimates of species' latent abundance, and hence species-at-length catchability coefficients, are rarely available in fisheries survey data. In isolation, each individual survey provides estimates of species' relative abundance at sampled locations and can provide assessment of the spatial distribution of fish within the survey domain. Problems may arise, however, when two or more surveys need to be combined to assess species' distributions. If gear efficiencies vary between different surveys, then estimates of species relative abundance provided by each survey may not be compatible. Failure to understand, or ignoring, how gear efficiency differs between surveys may lead to incoherent abundance estimates when merging surveys together to conduct assessments at large spatial scales. To perform such assessment, therefore, requires quantification of gear efficiency for different species, different size classes of fish, and different gears.

The traditional approach to estimating gear efficiency is through paired field experiments, where two vessels fish side by side and compare catches (Somerton et al., 1999; Zhou et al., 2014). Such experiments are costly to conduct and are generally implemented over limited spatial and temporal scales. However, where different survey domains overlap spatially, there may be opportunity to utilize species distribution modelling to complement, or even replace, field-based estimation of gear efficiencies (e.g. Ono et al., 2018); thereby providing a convenient framework for handling data from disparate surveys that can be regularly updated as new survey data become available. Statistical modelling of species distributions from large data sets is no longer limited by insufficient computing capacity. The use of such models offers an opportunity of overcoming challenges in combining data across surveys with varying gear efficiencies to enable extensive study of marine species distributions across large spatial scales.

Here we build from previous gear efficiency modelling efforts (Walker et al., 2017, Zhou et al., 2014) with an aim to advance the tools available for combining information across disparate fisheries surveys towards informing the spatial ecology of marine species. The spatial scale, the number of species assessed, the interaction between the gear-vessel combinations, and the spatial and temporal variation inherent within European fisheries surveys presents unique challenges requiring a new approach. Utilizing Generalised Additive Mixed Models (GAMMs); we analyse the proportion of variance explained by the differences in gear efficiency and the spatial–temporal variation in abundance of 135 species, in three length categories, collected in the 19 northeast Atlantic groundfish surveys with 24 different gear-vessel combinations. Here we focus on bottom trawl gears, namely otter trawls and beam trawls, as others have previously focused on combining acoustic measurements with habitat data to gain inference about the abundance of fish and infer on bottom trawl gear efficiencies (Kotwicki et al., 2018). Three length categories were chosen to (1) capture the main intra-specific length-related catchability differences described in previous studies (Fraser et al., 2007; Walker et al., 2017), (2) broadly reflect trophic guilds in marine fish communities (ICES, 2017), and (3) reflect the main size classes of fish either retained in commercial trawls or that escape through the mesh (Piet et al., 2009). The 24 gear-vessel combinations were chosen to best reflect the perceived differences in rigging and standard operating procedures carried out by different countries in their national surveys (Table 1). By understanding which species in our length categories are affected by variations among gears and vessels, our primary goal is to develop a consistent approach for combining groundfish surveys to facilitate marine ecosystem monitoring at large spatial scales. Using the GAMMs to control for differences in gear efficiency among surveys, we also generate estimates of spatial and temporal trends of relative abundance for species among different length categories throughout the northeast Atlantic to inform marine fish community ecological analyses (covering three ICES marine ecoregions/large marine ecosystems: Greater North Sea, Celtic Seas, and Bay of Biscay and the Iberian Coastal; Spalding et al., 2007). Finally, we conclude with a discussion of high priority information needs to further improve understanding of gear efficiency within marine fisheries survey data.

2. Methods

2.1 Fisheries Surveys

Data for most European groundfish surveys are uploaded and maintained on the ICES "Database of Trawl Surveys" (DATRAS). Data for surveys carried out in the Northeast Atlantic were recently subjected to a quality assurance and quality audit (QAQA) process (Moriarty et al., 2017; Greenstreet and Moriarty 2017a; 2017b; Moriarty et al., 2019), to ensure their adequacy to meet monitoring and assessment requirements under the EU MSFD (EC, 2008; 2010; 2017). These standard monitoring programme data products, along with data for four Spanish surveys, which underwent the same QAQA process but were not fully uploaded to DATRAS, were used in this study to obtain maximum spatial and temporal coverage and include the widest possible range of survey types for modelling (Table 1). Each survey data product includes the number of fish caught ($C_{i,s,l}$) of a species (s) at length (l), for each trawl sample (i), along with the vessel and fishing gear (g), tow location, date, depth and swept area (E). The fishing gear (g), included information from vessels that were expected to fish differently based on their gear configuration information. For example, both French and Irish vessels surveying in the Celtic Seas region use a GOV gear. However, the French surveys use double sweeps, and the Irish surveys rotate between a standard GOV survey gear (ICES 2015) and a double sweep with 16-inch bobbins, depending on the substrate (Table 1). The fish abundance data were organized into three broad length categories (lc), small unfished (<23cm), intermediate transition (23 - 35cm), and large fished (>35cm). Groundfish surveys only record those species and lengths caught (i.e. presence only data). Data rows for zero catches were added to the full data set where species at length were not reported in any given sample. To ensure constant and equivalent distance units, survey sample latitude - longitude coordinates were converted to eastings and northings (X, Y) using R package "Rgdal" (Bivand et al., 2018). Date (t) was incremented in quarterly time bins starting from quarter 4 (Oct – Dec) 2003, which was assigned time step $t = 1$, while the quarter 1 (Jan – Mar) 2004 was assigned time step $t = 2$, and so on.

2.2 Exploring Sources of Variation in Survey Abundance at Length Data

Generalised Additive Mixed Models (GAMMs) were used to account for non-linear spatial and temporal trends in fish density while simultaneously estimating gear efficiency using a modelling framework adapted from Walker et al. (2017). Survey catches were modelled as counts, with separate regressions for each species-length bin combination. Many species had a preponderance of zero catches. Initial exploration casting GAMMs for all species within Poisson, negative binomial, and zero-inflated Poisson frameworks showed that Poisson models provided a poor fit and failed to accommodate over-dispersion in catch data. Negative binomial and zero-inflated Poisson models showed similar fits for non-schooling species, but schooling species violated the assumption of independence required by Poisson processes. Consequently, we analyzed catches as Negative Binomially (NB) distributed GAMMs fit using the "mgcv" package (Wood 2004; 2011) in the R statistical programming environment (R Core Team 2017). The full model for a given species and length category catch data set had the form:

$C_i \sim NB(\mu_i, k)$

with $\mathrm{E}[C_i] = \mu_i = e^{\log(E_i) + s(X_i, Y_i, t_i) + zg_{(i)}}$ 1,

where C_i is the number of fish of a given species in a given length category caught in the i^{th} sample (fishing event), k is the negative binomial shape parameter representing the degree of overdispersion, $\log(E_i)$ is the log of swept area for fishing event i which was included as an offset to account for varying fishing effort among trips, $s(X_i, Y_i, t_i)$ denotes a multivariate smoothing function to represent spatio-temporal trends in catch data, and $zg_{(i)}$ are i.i.d. normally distributed random effects for gear-vessel combinations associated with fishing events. The space-time smoothing model component, $s(X_i, Y_i, t_i)$, was specified as a tensor product smoother for which the associated basis functions were cast as cubic splines with shrinkage (i.e., $te(X_i, Y_i, t_i, bs = "cs")$ in mgcv formulaic notation), a formulation which can accommodate data on different scales (Wood 2004; 2011). Gear-vessel combination was treated as a random effect, as opposed to a fixed effect, because variation among catch efficiencies is the primary feature of interest, and because this approach also aids in model convergence by reducing the number of fitted parameters. The spatiotemporal smoother describes the underlying estimated distribution of species across space and time; whereas the random effect controls for variation among gear efficiency when combining disparate survey data sets. To facilitate model convergence, we excluded data on species-at-length for which any given length category was sampled by fewer than two gear-vessel combinations or was sampled fewer than 100 times. The full model was compared to a reduced model that included space-time covariates, but which did not account for the effect of gear-vessel combinations (i.e. the gear-vessel combination random effect was dropped) in order to assess the impact on species distribution modelling inference when gear is ignored. Comparisons of full and reduced model fits were assessed using Akaike's information criterion (AIC). The full model was further assessed for reliability using visual tests and a chi squared goodness of fit test. To substantiate that our GAMM models can effectively differentiate between the random gear-vessel effects and the spatial and temporal variation in the abundance of demersal fish in the north east Atlantic region, we performed a simulation-estimation experiment (Supplemental Material S2).

2.3 Interpretation of models

To interpret the importance of gear efficiency versus spatiotemporal distribution patterns in explaining variation in survey data, we utilized variance components analysis. This analysis partitions total variation in the fitted data among the three modelled components: gear efficiency, spatiotemporal distribution, or unexplained residual variation. Accordingly, when the gear component constitutes the preponderance of model variation for a given species and/or length category, we conclude that gear efficiency varies widely across gears and surveys. In contrast, when location and time make up the majority of model variability for a given species, we conclude that catches are more strongly influenced by the ecology of the fish, rather than the differences in gear efficiency.

A non-metric multidimensional scaling (nMDS) unconstrained ordination technique using Euclidean distances was employed to explore how each species within the assemblages varied with estimates of gear efficiencies among gear-vessel coefficients and length classes from our models. Species were grouped by taxonomic order as a proxy for functional forms to examine if there was a pattern in estimates of gear efficiencies in species groups with similar morphological or ecological attributes. The gear-vessel coefficients were conditioned into a matrix, where the Scottish vessel with a GOV gear type was used as a reference gear, and the difference was calculated for each other gear-vessel combination. Permutational multivariate analysis of variance (PERMANOVA) was used to test the differences between the gear-vessel coefficients derived for each species in each length class from our full models for similar gear types. A clustering criterion that minimizes the amount of variance within in the gear-vessel groups was implemented (Ward, 1963). Euclidean distance was used and the *p*-value was set to 0.05. The nMDS and PERMANOVA routines were implemented in R (R Core Team 2017) using the “vegan” package (Oksanen et al., 2017).

3. Results

Data for 135 fish species were available from otter trawl surveys across the northeast Atlantic, whereas beam trawl surveys operate in a much more limited area within the North and Irish Seas (Figure 1). The surveys carried out in the Irish Sea have the highest degree of spatial and temporal overlap, whereas survey overlap is more limited in the Bay of Biscay and Iberian Coast region (Figure 1).

Two hundred and fifty-four full GAMMs were fit to 132 species in up to three length categories (Figure 2). For fishes in the smallest size class (<23cm), the full model was fit to 109 species, and 23 species had insufficient data based on the criteria described in Methods (Section 2.2). For fishes in the intermediate transition category (23-35cm), the full model was fit to 85 species, and 47 species had insufficient data. For the largest size class (>35cm), the full model was fit to 60 species, and 72 species had insufficient data.

In 39/254 models, the unexplained variance was greater than the explained variance (Figure 2). In 237/254 of the species-length combinations, the full model, which controlled for differences in gear-vessel combinations, improved the deviance explained over the reduced model (Table S1.1). 250/254 full models had a lower AIC score than the reduced model. In the cases where the full estimates did not improve inference, the differences in the amounts of deviance explained and the AIC scores between the full and reduced models were small (Table S1.1).

In 215/254 full models, over 50% of the variation in the data can be explained, suggesting that this framework is an effective way of calculating variance in latent species abundance over a large spatial scale. In 181/254 full models, location (X, Y) and time (t) components of the model explained over 50% of the variation in the data, suggesting that catch rates are strongly driven by the ecology of the fish, while the random effect of fishing gear on a given vessel(g) at a given length category (l) generally plays a smaller role in explaining variance. Indeed, in 51 of these 181 models, the overall variance explained is >50%, but the variance explained by gear is <1%. As an example, for common dab *(Limanda limanda)* in the <23cm length class, the random effect of fishing gear on a given vessel (g) explains 0.007% of the variance, while location (X, Y) and time (t) components explained 62.2% of the variance (Figure 3a/b). In this case, the reduced model, where location (X, Y) and time (t) components explained 61.1% of the variance, performed similarly to the full model (Supplemental Material 1 Table S1.1).

In 37/254 full models, the overall variance explained is >50%, and the gear component explains between 1% and 5% of the variation, suggesting that gear efficiency varies across gears and vessel combinations but has relatively little influence on catch performance. For example, for the thorny skate (*Amblyraja radiata*) in the 23-35cm length class, the random effect of fishing gear (g) explained 3.7% of the variance, while location (X, Y) and time (t) components of the full model explained 68.7% of the variance. While the estimated variance component for gear effects was smaller than the space-time components, the effect of fishing gear can be seen in the difference in spatial pattern between the full and reduced models (Figure 3d).

In 127/254 full models the overall variance explained is >50%, and the gear component explains more that 5% of the variation, suggesting that gear efficiency for these species-at-length varies substantially across gear and vessel combinations. For example, for sole (*Solea solea)* in the 23-35cm length class, the random effect of fishing gear (g) explained 8.6% of the variance, while location (X, Y) and time (t) components of the full model explained 46.5% of the variance in the data (Figure 3e/f). In this case, the output of the full model highlights the importance of understanding the effect of fishing gear in assessing the distribution of this species.

To assess the difference in inference gleaned from the full and reduced models, we further explored the spatial-temporal pattern of sole (*Solea solea)* in the 23-35cm category. While the general pattern is similar in the full and reduced models (Figure 4), the reduced model suggests the presence of intermediate-sized sole off of the coasts of Spain and Portugal; whereas the full model suggests that there are no intermediate-sized sole in these areas. When examined more specifically, we see that for the entire area, the sole data is 88% zero values, but for the southern part of the study area, where Spain and Portugal survey, the sole data is 96.5% zero values. Consequently, we can conclude that the reduced model is likely to overestimate the abundance in this area, and that this overestimation is likely an artefact of not accounting for gear.

Aggregating over the entire distribution of sole, there is a steadier rate of movement in the centre of mass in the population estimated from the full model, while the movement in the centre of mass in the population estimated from the reduced model is more variable (Figure 5a). The centre of mass metric highlights the eastward movement in the population in the full model, which is not the case in the reduced model (Figure 5b). The inference from the simulations suggests that the full model should be more capable of capturing the direction of movement than the reduced model ((Supplemental 2, Figure S2.4).

Unsurprisingly, nMDS highlights that the estimated gear coefficients vary considerably by gear types (Figure 6a; PERMANOVA test for differences in gears: $F = 2.36, R^2 = 0.18, p - value = 0.001$). However, gear coefficients are largely consistent within gear type, indicating stable catch efficiencies within gear types regardless of the survey country of origin or vessel. The GOV, beam trawls, and baca trawls gear-vessels tended to group most closely in their estimated gear coefficients, whereas other gears tended to differ more widely. The GOV has the highest level of variance and is the most widely used gear within the region. The beam trawl surveys have a high level of spatial overlap with the surveys that use the GOV gear in the North Sea and the rockhopper trawl in the Irish Seas. The baca trawl has very limited spatial overlap with other gears as it is used exclusively by the Spanish in the Bay of Biscay and Iberian Coast region. There is no clear pattern emerging in the estimated relative difference in catch efficiencies across species functional form (Figure 6b).

4. Discussion

Understanding how gear efficiency impacts fishery independent survey sampling is required for robust multi-survey species distribution modelling of both commercial and non-commercial species and is a key factor in determining absolute abundance estimates for commercial stocks (Kasatkina & Ivanova, 2009; Maunder & Piner, 2014). The aim of the analyses presented here is to provide an overall understanding how species are affected by the rigging of individual vessels to guide future ecosystem-scale species distribution modelling and examinations of fish communities. Our models support the derivation of relative species abundance estimates, and they provide information on gear efficiency of 24 gear-vessel combinations seasonally for three length groups chosen to reflect the main intra-specific length-related differences described in previous catchability studies (Fraser et al., 2007) in this region. This provides a modelling workflow to combine data across surveys that controls for potential gear-vessel-specific differences in catchability. The flexible framework provided here may be adapted to the end users' needs; for example, different length categories may be applied to answer specific ecological questions. We caution; however, that the gear efficiency coefficients used in this analysis were estimated using a 10-year historical time span and are only valid under the conditions for which they are calculated. As such, any efforts to employ them for correcting individual survey-species catches need take this into account (Arreguín-Sánchez, 1996).

In 15% (39/254) of models, the unexplained variance is higher than the explained variance (Figure 2). Given that it is unlikely for a species to be randomly distributed in space and time, this high unexplained variance is likely due to the rareness of the species within a given length category (i.e. there are not enough samples to describe the latent species distribution). Species that are rarely caught may not be rare in the environment, but instead may be particularly poorly sampled (i.e. low gear efficiency) in the survey trawl gear. Sampling of fish in the marine environment by fishing gear is known to be imperfect (Fraser et al., 2007, Zhou et al., 2014, Walker et al., 2017). This means additional considerations may need to be addressed during sampling and data analysis, such as joint dynamic species distribution modelling (Thorson et al., 2016). Reliable inference depends on sampling methods that produce reasonable odds of detection given presence, where no estimator will be particularly helpful when applied to data on populations or species that are ‘‘invisible’’ to collection gear (MacKenzie et al., 2006).

The estimated variance components from our models show that in 35% of cases (88/254), the location and time components explained most of the variation in the data, while the gear component explained relatively little variation (≤ 5%; Figure 2). This suggests that in such circumstances, the spatial-temporal distribution of these species can be estimated using combined survey data. Where the modelled gear component is especially small, particularly in relation to the location and time component, use of raw survey catch data from multiple surveys provides a reasonably accurate representation of temporal and spatial variation in species' abundances (by length category) at large spatial scales. The common dab (Figure 3a/b.), highlights a circumstance in which little variance can be attributed to gear effects, and we see a consequent small difference in inference in the temporal and spatial trends between the full and reduced models. The variance explained by the gear is <1% while the spatial and temporal components explain 62.2% of the variance. Thus, this species (by length category) abundance appears to be less impacted by the effects of gear as the catch rates are likely driven by the ecology of the fish. The variation that is attributable to gear effects is smaller than that attributed to space and time in most of our GAMM models, but the nature of the gear effects are not randomly distributed throughout the study area or throughout the year. They are instead systematically distributed by seasonal surveys. This regularity in the differences may impact species distribution inference at large scales. Simulations (S2a) for species demonstrating substantial movements in distribution attributed 5.7 % of model variance to gear, even when no gear effect was included. This suggests that some of the variance associated with location and time may be attributed to gear, but inferences from full and reduced models were similar. Conversely, when there is a strong gear effect (S2b) then the full model improves inference of abundance estimates and direction of population centre of mass movements over the reduced model (Supplemental Material 2).

Not accounting for gear may lead to incorrect estimates of relative abundance or species' distributions. Data analysed here suggest that gear effects on catches across disparate surveys are not uncommon, whereby in half of our full models (127/254), the gear component explained more that 5% of the total variation in survey catches, while overall variance explained is >50%. Our examination of the distribution of sole provides demonstration of the potential importance of controlling for gear effects when attempting to combine data across surveys for some species. The variance explained by gear in this case was 8.6%, while the spatial and temporal components of the model accounted for 46.5% of the variance. Consequently, we found substantial differences in relative abundance trends between models which control for gear effects compared to reduced models which ignore gear effects in combining data across surveys (Figures 3d/e, 4, 5). Importantly, failure to control for gear differences across surveys for this species would mask differences in the spatial distribution of the stock across commercial fishing areas, as well as mask ecosystem-scale population shifts to the east (Figure 5). It may be valid to pool across surveys in assessing species distributions for many species-size combinations; however, there are differences evident across gear types and it is not clear a priori for which species gear differences matter (Figure 6b). Thus, a sensible workflow when combining data across surveys may be to implement models that control for gear type as demonstrated here and then subsequently evaluate whether gear differences account for a substantial portion of the variation in catches.

Northeast Atlantic waters are currently surveyed by 12 countries carrying out 19 different surveys designed with individual goals and objectives and using different vessels and a variety of gears (Table 1). ICES facilitates survey coordination and collaboration through working groups to make the surveys as comparable as possible. The North Sea bottom trawl surveys have led the way in terms of minimising gear efficiency issues caused by differences in vessels and by ensuring survey overlap and similarity among gears (ICES, 2015). There is a large body of work ongoing in ICES survey groups (e.g. WGBEAM, International Bottom Trawl Survey Working Group; IBTSWG) to minimise survey variability; however, assessing relative gear efficiency at the scale examined here highlights the need for comparative experiments to help achieve a more coherent understanding of gear efficiency within fisheries independent survey data. This is particularly relevant in the Bay of Biscay, where overlapping or paired tows between the Spanish Baca Trawl and Portuguese Norwegian Campelen Trawl and the Spanish Baca Trawl and French Grande Overture Vertical Trawl would help to improve inferences of species relative abundance obtained from these different gears (Figure 6a). Analyses herein provide further understanding of the differences in gear efficiency between trawl gears used by different surveys for species sampled across the northeast Atlantic.

Information on the abundance and distribution of organisms is a fundamental knowledge need for fisheries management. Data on predator and prey abundances by different age and size classes can inform species status assessments as well as provide information on the interactions among species and size classes, providing understanding about the impact of fishing on fish communities (Fraser et al., 2007; e.g. Large Fish Indicator). This study provides an approach to facilitate comparability between catches from different surveys and gears, providing a framework to assist in integrating data across countries, regions, and sampling programs towards maximizing the use of available information to inform species' abundance and spatial distribution assessments.

Acknowledgements

The collaboration between MM, TSS, NW, BPH and SAS was made possible by the Fulbright-Marine Institute funding from the Irish Fulbright Commission, for which the authors are very grateful. SPRG was supported by the Scottish Government's scheduals of service SLA/20452. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Arreguín-Sánchez, F. 1996. Catchability: a key parameter for fish stock assessment. Reviews in fish biology and fisheries, *6*(2), 221-242.

Bivand, R., Keitt, T., and Rowlingson, B. 2019 rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.4-3. https://CRAN.R-project.org/package=rgdal

Dann, N., Heessen, H., terHofstede, R. 2005 North Sea Elasmobranchs: distribution, abundance and biodiversity. Theme Session on Elasmobranch Fisheries Science ICES CM 2005/N:06

EC. 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). Off. J. Eur. Union 164, 19–40.

EC. 2010. Commission decision of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters. Official Journal of the European Union, L232: 14–24.

EC. 2013. Regulation (EU) No 1380/2013 of the European Parliament and of the Council of 11 December 2013 on the Common Fisheries Policy, amending Council Regulations (EC) No 1954/2003 and (EC) No 1224/2009 and repealing Council Regulations (EC)

No 2371/2002 and (EC) No 639/2004 and Council Decision 2004/585/EC. Official Journal of the European Union, L354: 22–61.

EC. 2017. Commission Decision (EU) 2017/848 of 17 May 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, and repealing Decision 2010/477/EU. Official Journal of the European Union, 18.5.2017 L 125: 43–74.

Fraser, H. M., Greenstreet, S. P. R., and Piet, G. J. 2007. Taking account of catchability in groundfish survey trawls: implications for estimating efficiency of survey and commercial trawl gears estimating demersal fish biomass. ICES Journal of Marine Science, 64: 1800–1819.

Greenstreet, S.P.R., and Moriarty, M., 2017a. Ospar Interim Assessment 2017 Fish Indicator Data Manual (Relating to for Version 2 of the Groundfish Survey Monitoring and Assessment Data Product). Scottish Marine and Freshwater Science Report Vol 8 No. 17 DOI: 10.7489/1985-1

Greenstreet, S.P.R., and Moriarty, M., 2017b. Manual for Version 3 of the Groundfish Survey Monitoring and Assessment Data Product. Scottish Marine and Freshwater Science Report Vol.8 No.18 DOI:10.7489/1986-1

ICES. 2015 Manual for the International Bottom Trawl Surveys. Series of ICES Survey Protocols SISP 10 - IBTS IX. 86 pp.

ICES. 2017. Interim Report of the Working Group on Biodiversity Science (WGBIODIV), 6–10 March 2017, Venice, Italy. ICES CM 2017/SSGEPD:01. 14 pp.

Kasatkina, S., & Ivanova, V. (2009). Modelling study of catchability properties of research and commercial trawls to identify sources of uncertainty in resource surveys indices. *ICES CM*, *1*, 13.

Kelley, E., & Sherman, K. (2018). Trends of the Large Marine Ecosystem assessment and management approach as reflected in the literature. *Ocean & Coastal Management*, *155*, 104-112.

Kotwicki, S., Ressler, P.H., Ianelli, J.N., Punt, A.E., and Horne, J.K., 2018 ‘Combining data from bottom-trawl and acoustic-trawl surveys to estimate an index of abundance for semipelagic species‘ Can. J. Fish. Aquat. Sci. 75: 60–71 (2018) dx.doi.org/10.1139/cjfas-2016-0362

MacKenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey LL, et al. (2006) Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. New York: Academic Press. 324 p.

Maunder, M. N., & Piner, K. R. (2014). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science*, *72*(1), 7-18.

Moriarty, M., Greenstreet, S.P.R., Rasmussen, J., de Boois, I., (2019) Assessing the State of Demersal Fish to Address Formal Ecosystem Based Management Needs: Making Fisheries Independent Trawl Survey Data ‘Fit for Purpose’ , *Frontiers in Marine Science* 6,162 10.3389/fmars.2019.

Moriarty, M., Greenstreet, S.P.R. and Rasmussen, J. (2017). Derivation of Groundfish Survey Monitoring and Assessment Data Products for the Northeast Atlantic Area. Scottish Marine and Freshwater Science Report Vol 8. No 16. DOI:10.7489/1984-1

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs

E., and Wagner, H. (2017). vegan: Community Ecology Package. R package version 2.4-5.

Ono, K., Ianelli, J. N., McGilliard, C. R., and Punt, A. E. 2017. Integrating data from multiple surveys and accounting for spatio-temporal correlation to index the abundance of juvenile Pacific halibut in Alaska. – ICES Journal of Marine Science, 75: 572–584.

Piet, G. J., van Hal, R., and Greenstreet, S. P. R. 2009. Modelling the direct impact of bottom trawling on the North Sea fish community to derive estimates of fishing mortality for non-target fish species. ICES Journal of Marine Science, 66: 1985–1998.

Poos, J. J., Aarts, G., Vandemaele, S., Willems, W., Bolle, L. J., and van Helmond, A. T. M. 2013. Estimating spatial and temporal variability of juvenile North Sea plaice from opportunistic data. Journal of Sea Research, 75: 118–128.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Somerton, D., Ianelli, J., Walsh, S., Smith, S., Godø, O. R., and Ramm, D. 1999. Incorporating experimentally derived estimates of survey trawl efficiency into the stock assessment process: a discussion. ICES Journal of Marine Science, 56: 299–302.

Spalding, Mark D., Helen E. Fox, Gerald R. Allen, Nick Davidson, Zach A. Ferdana, M. A. X. Finlayson, Benjamin S. Halpern et al., "Marine ecoregions of the world: a bioregionalization of coastal and shelf areas." *BioScience* 57, no. 7 (2007): 573-583.

Thorson, J.T., Ianelli, J.N., Larsen, E.A., Ries, L., Scheuerell, M.D., Szuwalski, C., and Zipkin, E.F., 2016. Joint dynamic species distribution models: a tool for community ordination and spatiotemporal monitoring. Global Ecology and Biogeography, (2016) 25, 1144–1158

US. 1996. Magnuson-Stevens Fishery Conservation and Management Sustainable Fisheries Act of 1996 Public Law 104-297

US. 2006. Magnuson-Stevens Fishery Conservation and Management Reauthorization Act of 2006 Public Law 109-479, 109th Congress, Stat. pp. 3575–3665.

Walker, N. D., Maxwell, D. L., Le Quesne, W. J. F., and Jennings, S. 2017. Estimating efficiency of survey and commercial trawl gears from comparisons of catch-ratios. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsw250.

Ward, J. H., Jr. (1963), "Hierarchical Grouping to Optimize an Objective Function", *Journal of the American Statistical Association*, 58, 236–244.

Wood, S.N. (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association. 99:673-686.

Wood, S.N. (2008) Fast stable direct fitting and smoothness selection for generalized additive models. Journal of the Royal Statistical Society (B) 70(3):495-518

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36

Zhou, S., Klaer, N. L., Daley, R. M., Zhu, Z., Fuller, M., and Smith, A. D. M. 2014. Modelling multiple fishing gear efficiencies and abundance for aggregated populations using fishery or survey data. ICES Journal of Marine Science, 71: 2436–2447.

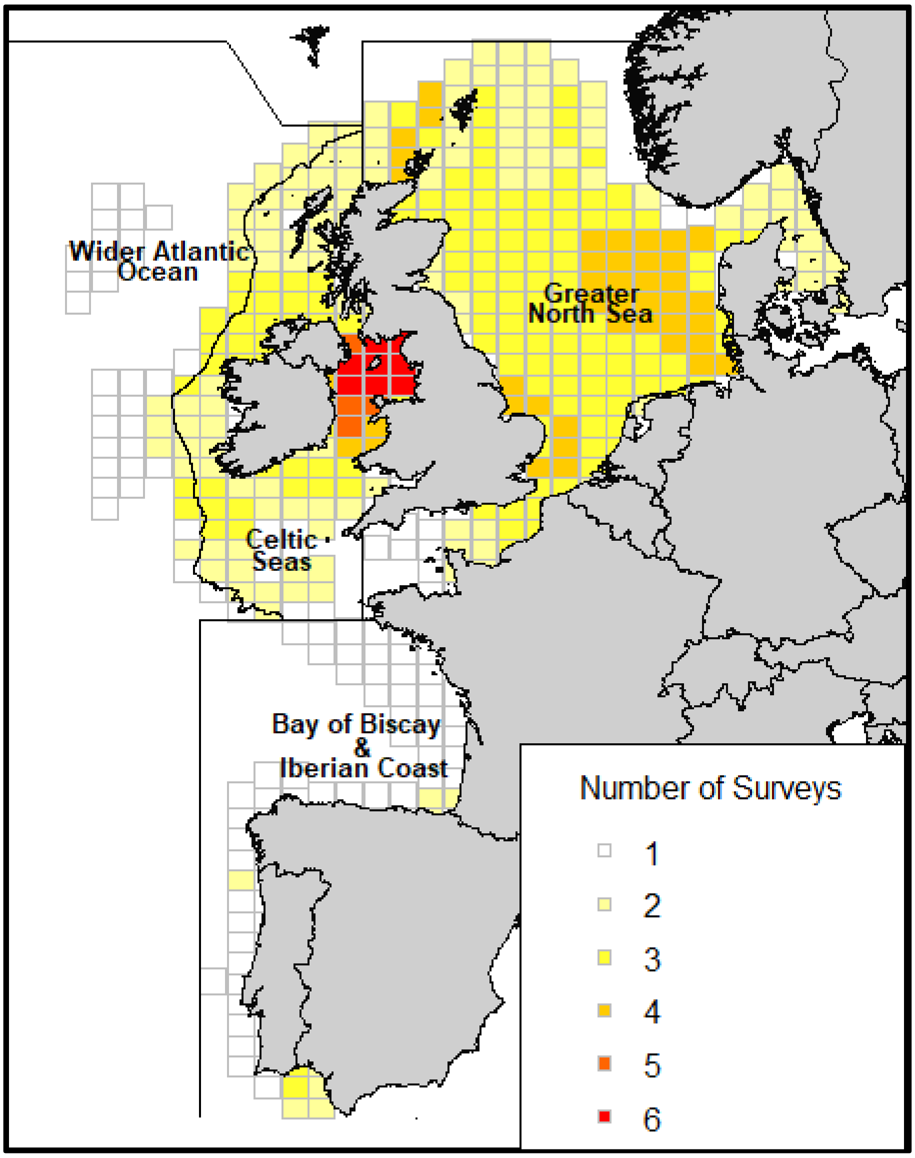

Figure 1: *Fisheries independent survey coverage across the northeast Atlantic. Thick black line shows Oslo/Paris convention (OSPAR) boundaries. Number of surveys operating in each ICES statistical rectangle is depicted by a different colour. See Table 1 for list of surveys.*

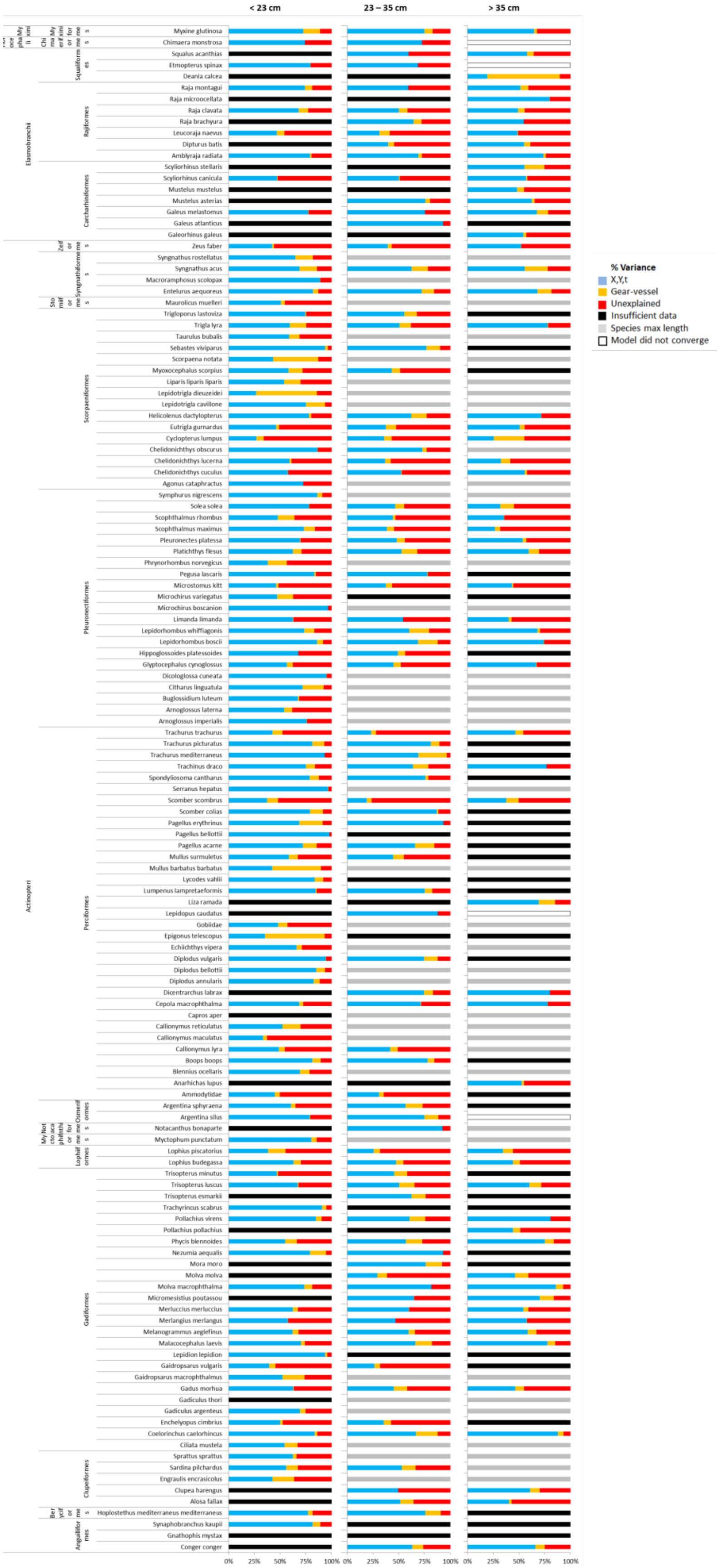

Figure 2: Summary of the proportion of variance explained from full model components for each length category (<23cm; 23 - 35cm and >35cm) and *species, grouped in taxonomic order. X, Y and time (t) variance components are represented by blue bars, gear-vessel*

components by orange bars, and unexplained variance by red bars. Black bars indicate insufficient data to fit a model for a given species-size combination, and white bars indicate model convergence failed. Finally, grey bars indicate a given length size bin is larger than the maximum observed length of a species.

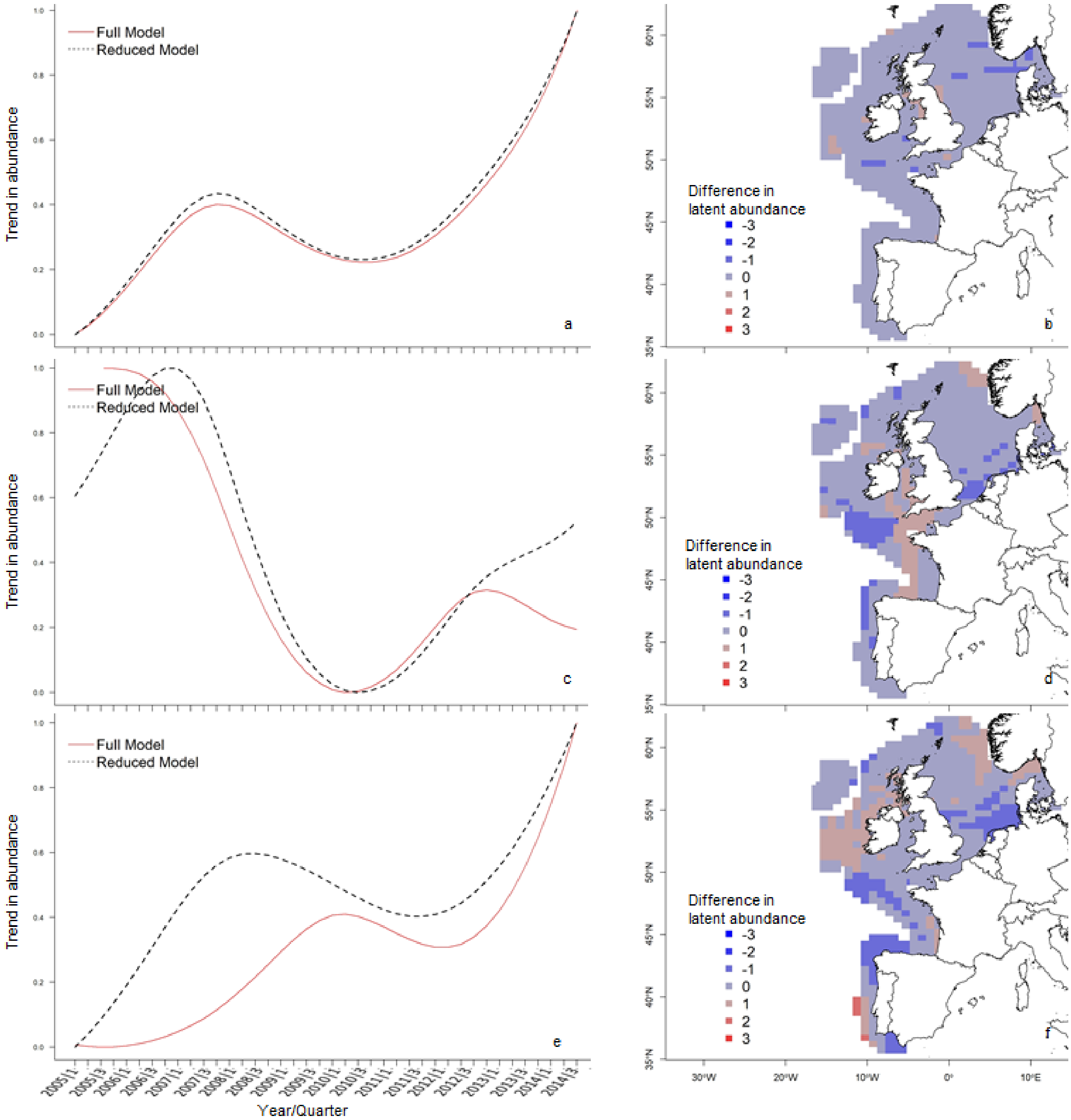

Figure 3: Top row (a,b): Common dab (Limanda limanda) < 23cm, highlighting an example of a species where the reduced and full model perform similarly as the variance explained by gear is very small (0.007%). Middle row (c,d) Thorny skate (Amblyraja radiata) 23 – 35cm, highlighting an example of a species with between 1-5% variance explained by gear. Bottom row (e,f) Sole (Solea solea) 23-35cm, highlighting an example of a species with >5% variance explained by gear. Left column (a,c,e): Estimated domain-wide species' abundance trends for the full model which controls for gear differences across surveys, versus the reduced model which does not control for gears. A large discrepancy between the curves indicates gear differences across surveys may impact inference about species' abundance and distributions. Right column (b,d,f): Differences in predicted species' relative mean

abundance between the full and reduced models. Dark colours represent large discrepancies between the models, indicating differences in gears across surveys may influence estimated species' distributions if not accounted for.

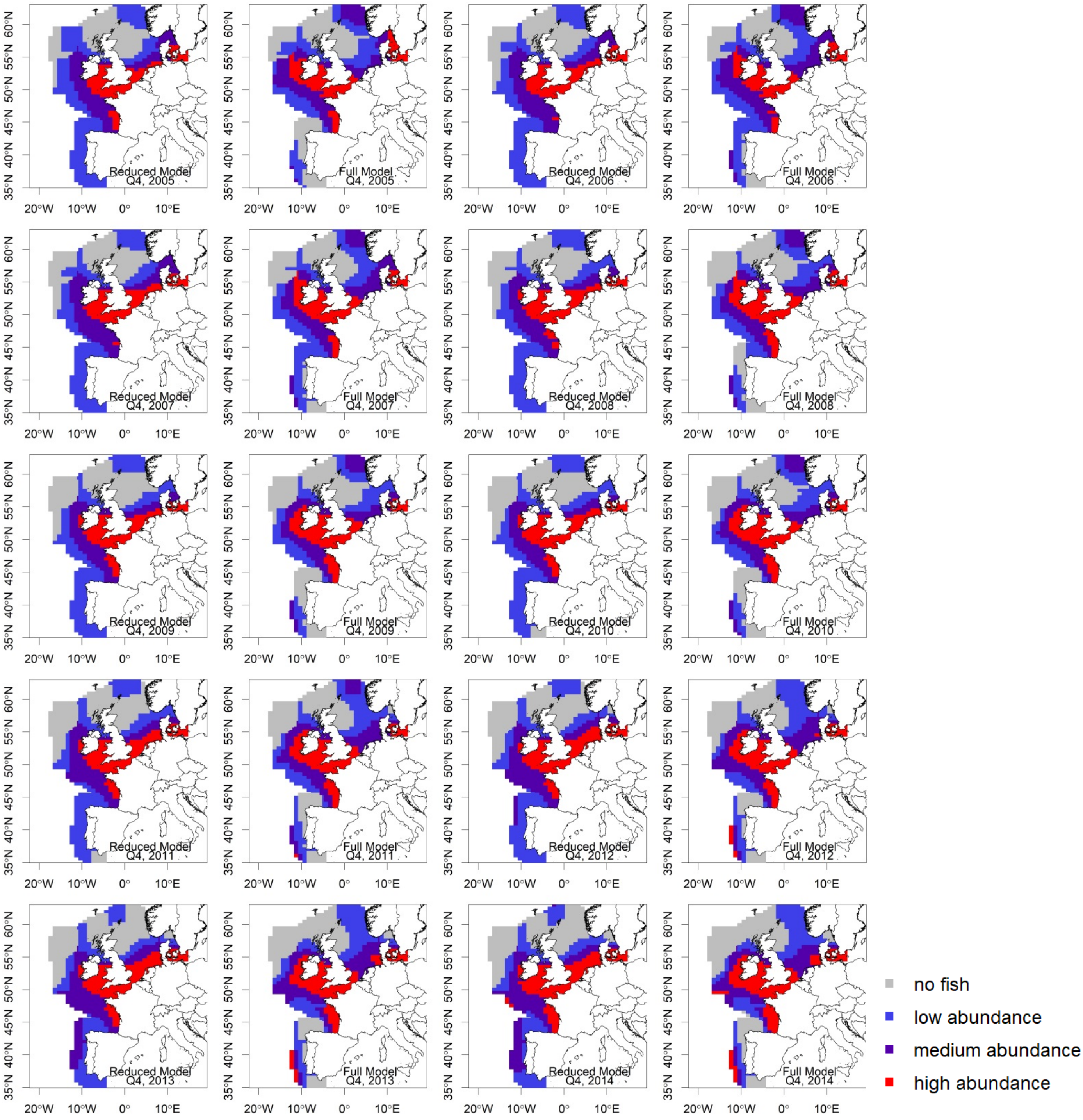

Figure 4. Spatial-temporal pattern in quarter 4 (Oct.-Dec.) for each year of sole (Solea solea) 23-35cm from the reduced model on the left and the full model on the right. Abundance is depicted as "low" in the 1st-2nd quantile, "medium" in the 2nd – 3rd quantile, and high in the 3rd-4th quantile.

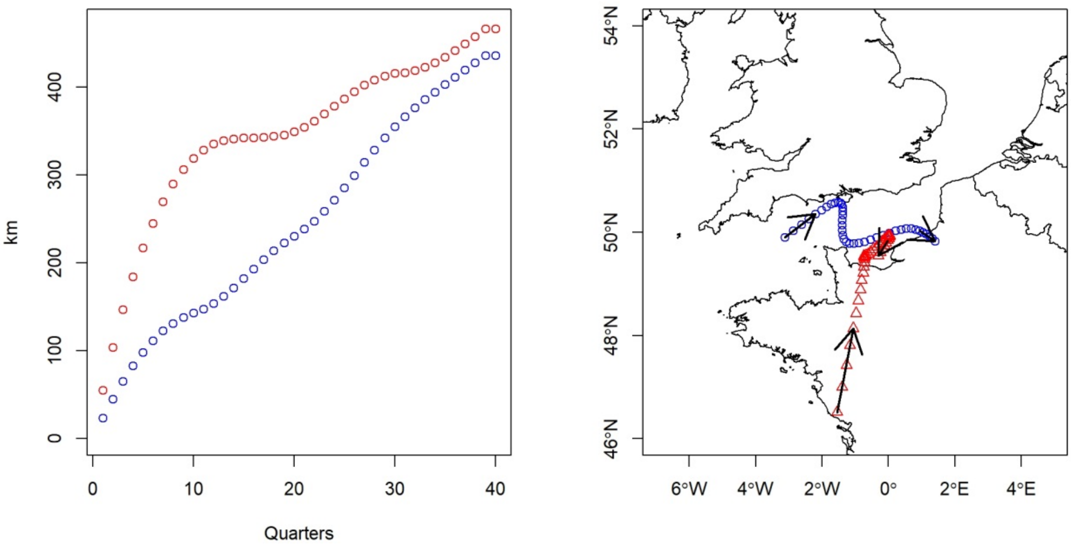

Figure 5. Summary of difference in inference from the spatial-temporal pattern of sole (Solea solea) 23-35cm from the full and reduced models. (a.) Cumulative movement from the centre of mass from the start of the time series for the full model (blue circles) and the reduced model (red triangles). (b.) Centre of mass of the abundance of the fish from the full model (blue circles) and the reduced model (red triangles).

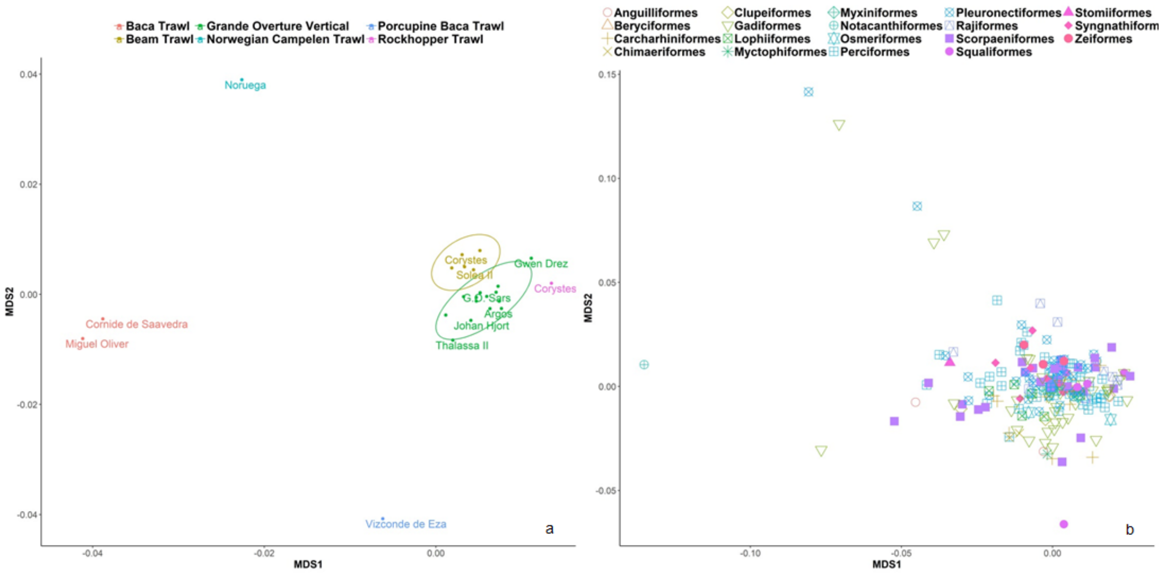

Figure 6. Non-metric multidimensional scaling (nMDS) plots describing how the gear-vessel coefficients varied by survey or by taxonomic grouping (Stress = 0.102). (a) Gear coefficients grouped by trawl type (colours) and survey research vessel name (labels). Points more closely situated are more similar in terms of their gear-vessel coefficients. Ellipses indicate the 95% confidence intervals for clusters of each gear type. (b) Gear coefficients grouped by taxonomic order as a proxy for species' functional form.

Survey Acronym	DATRAS Acronym	Subregion	Country	Start Year	End Year	Quarter	Vessels	Gear Type	Mesh size (mm)	Haul Duration (min) $\bar{x} \pm s$	Distance Towed (km) $\bar{x} \pm s$	Wing Swept Area (km^2) $\bar{x} \pm s$	Data Source	DOI
GNSIntOT1	IBTS	Greater North Sea	International	1983	2017	1	Multiple Ships	Otter (GOV)					DATRAS	10.7489/1922-1
							G.O.Sars		20	29±4	3.4±0.6	0.06±0.01		
							Argos		20	30±4	3.6±0.5	0.07±0.01		
							Dana		20	30±2	3.6±0.3	0.07±0.01		
							Dana (Sweden)		20	30±2	3.4±03	0.07±0.01		
							CEFAS Endeavour (Netherlands)		20	29±3	3.6±0.4	0.0±0.017		
							Haakon Mosby		20	29±3	2.9±0.3	0.06±0.1		
							Mimer		20	29±3	3.3±0.3	0.06±0.01		
							Scotia III		20	31±6	3.6±0.7	0.07±0.01		
							Thalassa II		20	30±1	3.6±0.4	0.06±0.01		
							Tridens II		20	30±4	3.8±0.5	0.07±0.01		
							Walther Herwig III		20	30±2	3.8±0.4	0.07±0.01		
GNSIntOT3	IBTS	Greater North Sea	International	1998	2016	3	Multiple Ships	Otter (GOV)					DATRAS	10.7489/1923-1
							Argos		20	30±1	3.5±0.2	0.07±0.01		
							Dana		20	29±4	3.6±0.5	0.07±0.01		
							Dana (Sweden)		20	30±1	3.4±0.2	0.07±0.01		
							CEFAS Endeavour		20	29±3	3.5±0.4	0.07±0.01		
							Haakon Mosby		20	27±5	3.2±0.6	0.07±0.01		
							Johan Hjort		20	27±6	3±0.8	0.06±0.02		
							Scotia III		20	29±4	3.3±0.5	0.06±0.01		
							Walther Herwig III		20	29±4	3.6±0.7	0.07±0.01		
GNSFraOT4	FR CGFS	Greater North Sea	France	1988	2016	4	Thalassa II, Gwen Drez	Otter (GOV)					DATRAS	10.7489/1959-1
							Gwen Drez		20	29±3	2.9±0.5	0.03±0.01		
							Thalassa II		20	29±2	3.4±0.3	0.05±0.01		
CSScoOT1	SWC-IBTS	Celtic Sea	Scotland	1985	2016	1	Scotia II	Otter (GOV)	20	56±10	7.4±1.7	0.15±0.03	DATRAS	10.7489/1957-1
							Scotia III		20	30±6	3.4±0.7	0.07±0.01		
CSScoOT4	SWC-IBTS	Celtic Sea	Scotland	1997	2016	4	Scotia II	Otter (GOV)	20	56±10	6.6±1.7	0.13±0.03	DATRAS	10.7489/1924-1
							Scotia III		20	29±3	3.4±0.4	0.06±0.01		
CSIreOT4	IE-IGFS	Celtic Sea	Ireland	2003	2016	4	Celtic Explorer	Otter (GOV)	20	30 ± 2	3.6±0.3	0.07±0.01	DATRAS	10.7489/1925-1
CSNIrOT1	NIGFS	Celtic Sea	Northern	1992	2016	1	Corystes, Lough	Otter (ROT)					DATRAS	10.7489/1961-

			Ireland				Foyle						1
							Corystes		20	58±9	5.3±0.9	0.08±0.01	
							Lough Foyle		20	59±6	5.±0.6	0.08±0.01	
CSNIrOT4	NIGFS	Celtic Sea	Northern Ireland	1992	2016	4	Corystes, Lough Foyle	Otter (ROT)					NDB (92-07) DATRAS (08-15)	10.7489/1962-1
							Corystes		20	19±1	1.9±0.01	0.03±0.02		
							Lough Foyle		20	50±18	4.7±1.6	0.07±0.02		
CS/BBFraOT4	EVHOE	Celtic Sea/Bay of Biscay	France	1997	2016	4	Thalassa II	Otter (GOV)	20	30 ± 1	3.6±0.2	0.07±0.01	NDB (92-07) DATRAS (08-15)	10.7489/1958-1
BBIC(n)SpaOT4	SP-North	Bay of Biscay and Iberian Coast	Spain	1993	2014	4	F deP Navarro	Otter (BACA)	20	30	2.7±0.1	0.05	NDB	Not released, no DOI
							Cornide de Saavedra		20	30	2.8±0.2	0.05		
BBIC(s)SpaOT1	SP-ARSA	Bay of Biscay and Iberian Coast	Spain	1990	2015	1	F deP Navarro	Otter (BACA)	20	60	5.6±0.2	0.1 ± 0.02	NDB	Not released, no DOI
							Cornide de Saavedra		20	60	5.6±0.4	0.1±0.01		
BBIC(s)SpaOT4	SP-ARSA	Bay of Biscay and Iberian Coast	Spain	1997	2014	4	F deP Navarro	Otter (BACA)	20	60	5.5±0.3	0.09±0.02	NDB	Not released, no DOI
							Cornide de Saavedra		20	60	5.5±0.3	0.1±0.01		
BBICPorOT4	PT-IBTS	Bay of Biscay and Iberian Coast	Portugal	2001	2014	4	Capricornio, Noruega	Otter (NCT)	20	29±3	3.1±0.4	0.05±0.01	DATRAS	10.7489/1963-1
WAScoOT3	Rockall	Wider Atlantic	Scotland	1999	2016	3	Scotia III	Otter (GOV)	20	30±3	3.4±0.4	0.07±0.01	DATRAS	10.7489/1960-1
WASpaOT3	SP-PORC	Wider Atlantic	Spain	2001	2015	3	Vizconda de Eza	Otter (PBACA)	20	24 ± 4	2.7±0.5	0.07±0.02	NDB	Not released, no DOI
GNSNetBT3	BTS	Greater North Sea	The Netherlands	1999	2016	3	Isis, Tridens II	Beam (8m)					DATRAS	10.7489/1967-1
							Isis		40	30±2	3.8±0.3	0.03		
							Tridens II		40	34±11	4.5±1.4	0.04±0.01		

GNSEngBT3	BTS	Greater North Sea	England	1990	2016	3	Corystes	Beam (4m)	40	29±3	3.7±0.6	0.01	DATRAS	10.7489/1966-1
							Endevour		40	28±4	3.5±0.6	0.01		
GNSGerBT3	BTS	Greater North Sea	Germany	1998	2016	3	Solea I	Beam (7m)	40	30±3	3.5±0.5	0.03	DATRAS	10.7489/1965-1
							Solea II		40	30 ± 2	3.3±0.3	0.02		
CSEngBT3	BTS VIIa	Celtic Sea	England	1993	2015	3	Corystes, Endevour	Beam (4m)					DATRAS	10.7489/1964-1
							Corystes		40	28 ± 5	3.8±0.5	0.02		
							Endevour		40	28 ± 4	3.5±0.6	0.01		

Table 1. List of individual surveys considered in the derivation of the Oslo/Paris convention (OSPAR) Groundfish Survey Monitoring and Assessment data products. Survey acronyms reflect sub-region/country/gear/quarter, except CS/BB in the French EVHOE survey acronym to denote a survey that extends across two sub-regions, the Celtic Seas and Bay of Biscay. Data product start and end years reflect the period when surveys were deemed sufficiently established with consistent standardised methodology (Moriarty et al., 2017). NDB refers to national database. For this study we subset the data from 2004 – 2015 for continuous spatial coverage across the northeast Atlantic, the information on mesh size, haul duration, distance towed, and wing swept area reflect the data included in this study from 2004-2015 where $\overline{x}$ is the sample mean and s is the sample standard deviation.

Notes on fishing gear exceptions

S = Standard Gear B = Bobbins used D = Double Sweeps I2 = Ground gear D with 16-inch bobbins R = Rockhopper

<u>*Grande Overture Vertical Trawl*</u>

1. *Scotland uses R.V. Scotia III on five surveys WAScoOT3; CSScoOT4; CSScoOT1; GNSIntOT3; GNSIntOT1. For the west coast surveys (CSScoOT4 / CSScoOT1/ WAScoOT3) they use a "S" and "I2" gear for to deal with rocky habitat. In North Sea surveys (GNSIntOT3; GNSIntOT1), Scotland uses an "S" and a "B" exception.*
2. *Sweden uses a standard GOV ("S") on R.V Argos and R.V. Mimer in both North Sea surveys (GNSIntOT3 and GNSIntOT1).*
3. *Denmark uses an "S" gear and an "R" exception in both surveys on R.V. Dana II in both North Sea surveys (GNSIntOT3 and GNSIntOT1).*
4. England uses a standard GOV ("S") gear in the North Sea (GNSIntOT3) on R.V. CEFAS Endeavour.
5. *The Netherlands uses a standard GOV ("S") gear in the North Sea (GNSIntOT1) on R.V. Tridens II. R.V. CEFAS Endeavour was used in quarter 1 by Netherlands when Tridens broke down.*
6. *Norway uses an "S" gear and "D" exception on R.V. G.O. Sars and R.V. Johan Hjort in the North Sea (GNSIntOT3/ GNSIntOT1). When the R.V Haakon Mosby has been used only a standard gear is noted.*
7. *France uses a GOV gear in the North Sea (GNSFraOT4) on R.V. Gwen Drez, no exception is noted, however, the gear is smaller than the standard gear in the North Sea. France uses Thalassa II on two surveys CS/BBFraOT4 and GNSIntOT1. For the west coast surveys (CS/BBFraOT4) they use ground gear "D" while in the North Sea surveys (GNSIntOT1), a standard gear is used.*
8. *Germany uses a standard gear on R.V Walther Herwig III in the North Sea (GNSIntOT3/ GNSIntOT1).*
9. *Ireland uses an "S" and "I2" gear for west coast survey (CSIreOT4) to deal with rocky habitat in line with Scotland on R.V. Celtic Explorer.*

<u>*Beam Trawl*</u>

10. *The Netherlands uses R.V. Tridens II and R.V Isis in the GNSNetBT3 survey. Both ships use an 8m beam with a tickler but Tridens II has a different set up to Isis.*
11. *Germany uses a 7m Beam trawl with a 5m tickler chain on R.V. Solea II during GNSGerBT3.*
12. *England uses a 4m Beam trawl during both her CSEngBT3 and GNSEngBT3 surveys on R.V. Corystes and CEFAS Endeavour in 2014 and 2015 with the same rigging on both ships.*

Rockhopper Trawl

13. *The Rockhopper Otter Trawl in used by Northern Ireland in the CSNIrOT4 / CSNIrOT1 on R.V. Corystes.*

Baka Trawl

14. *Spain uses a Baka trawl on 3 surveys (BBIC(s)SpaOT4 / BBIC(s)SpaOT1 / BBIC(n)SpaOT4) on R.V. Cornide de Saavedra.*
15. *Spain uses a Porcupine Baka trawl on 1 survey (WASpaOT3) on R.V. Vizconde de Eza.*

Norwegian Campelen Trawl

16. *Portugal reports B and R gear exceptions on R.V Noruega.*

Supplemental Material 1:

Table S1.1 Variance explained (%) and AIC scores for all models in all length classes. This includes a folder containing the same information in Figure3 for every species/length combination.

Supplemental Material 2: Can GAMMs differentiate effects of gear efficiency from spatial and temporal variation in abundance in demersal fish?

Supplemental Material 3:

File S3.1: Example of R Scripts for fitting model

File S3.2: Example of R Scripts for generating data and fitting simulated models (S1 conditions only)