Fish functional groups of the North Atlantic and Arctic Oceans

Authors: Murray S.A. Thompson1, Izaskun Preciado2, Federico Maioli3,4,5, Valerio Bartolino6, Andrea Belgrano3,5,6, Michele Casini3,4, Pierre Cresson7, Elena Eriksen8, Gema Hernandez-Milian9, Ingibjörg G. Jónsdóttir10, Stefan Neuenfeldt11, John K. Pinnegar1, Stefán Ragnarsson10, Sabine Schückel12,13, Ulrike Schückel13, Brian E. Smith14, Maria Ángeles Torres15, Thomas J. Webb16, and Christopher P. Lynam1

Correspondence: murray.thompson@cefas.gov.uk

1 Centre for Environment, Fisheries and Aquaculture Science (Cefas), Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK
2 Instituto Español de Oceanografía (IEO-CSIC), Centro Oceanográfico de Santander, Severiano Ballesteros 16 39004 Santander, Spain
3 Department of Aquatic Resources, Institute of Marine Research, Swedish University of Agricultural Sciences, Turistgatan 5, SE-45330 Lysekil, Sweden.
4 University of Bologna, Department of Biological, Geological and Environmental Sciences, Laboratory of Marine Biology and Fisheries, Viale Adriatico 1/N, 61032 Fano, Italy
5 Swedish Institute for the Marine Environment (SIME), University of Gothenburg, Seminariegatan 1F, SE-41313 Gothenburg, Sweden
6 Gothenburg Global Biodiversity Centre, University of Gothenburg, P.O. Box 463, SE-405 30 Gothenburg, Sweden
7 IFREMER, Channel and North Sea Fisheries Research Unit, 150 Quai Gambetta F-62200 Boulogne sur Mer, France
8 Institute of Marine Research, Bergen, Norway
9 Centre Oceanographic Vigo-Spanish National Research Council (COV-CSIC), Subida Radio Faro 50, 36390 Vigo, Spain
10 Marine and Freshwater Research Institute, Fornubúðum 5, 220 Hafnarfjörður, Iceland
11 National Institute of Aquatic Resources, Technical University of Denmark, Lyngby, Denmark

https://doi.org/10.5194/essd-2024-102
Preprint. Discussion started: 17 June 2024
© Author(s) 2024. CC BY 4.0 License.
Abstract

International efforts to assess the status of marine ecosystems have been hampered by insufficient observations of food web interactions across many species, their various life stages, and geographic ranges. Hence, we collated data from multiple databases of fish stomach contents from samples taken across the North Atlantic and Arctic Oceans containing 944,129 stomach samples from larvae to adults, with 14,196 unique interactions between 227 predator species and 2158 prey taxa. We use these data to develop a data-driven, reproducible approach to classifying broad functional feeding guilds and then apply these to fish survey data from the Northeast Atlantic shelf seas to reveal spatial and temporal changes in ecosystem structure and functioning. In doing so, we construct predator-prey body size scaling models to predict the biomass of prey functional groups, e.g., zooplankton, benthos, and fish, for different predator species. These predictions provide empirical estimates of species- and size-specific feeding traits of fish, such as predator-prey mass ratios, individual prey mass, and the biomass contribution of different prey to predator diets. The functional groupings and feeding traits provided here help to further resolve our understanding of interactions within marine food webs and support the use of trait-based indicators in biodiversity assessments. The data used and predictions generated in this study are published on the Cefas Data Hub at: https://doi.org/10.14466/CefasDataHub.149 (Thompson et al., 2024).
Keywords: feeding guilds; species traits; marine ecosystem structure and functioning; food web indicator; biodiversity; trophic interactions; stomach contents data.

1 Introduction

Human pressures are affecting global patterns in marine ecosystem structure and functioning, from species distributions and their body sizes, to how and where energy fluxes through food webs (Daufresne et al., 2009; du Pontavice et al., 2020; Kortsch et al., 2015). Grouping organisms into broad functional feeding guilds and assessing change in their populations has been widely advocated to assess marine ecosystem status, gauge sustainable levels of human pressure, and inform management interventions (ICES, 2018; Rombouts et al., 2013; Tam et al., 2017). Yet, international efforts to develop indicators of marine ecosystem status have been hampered by the large number of observations needed to understand food web processes involving many species, their various life stages, and geographic ranges.

Despite the complexity of natural food webs, their structure and dynamics are largely determined by the size of the interacting organisms because predators are systematically larger than their prey, especially in aquatic ecosystems (Brose et al., 2006; Petchey et al., 2008; Woodward et al., 2005). The behaviour of organisms relating to their evolutionary history, including adaptations for specific foraging strategies and habitats, also affects predator-prey interactions (Brose et al., 2019; Link, 2004; Pecuchet et al., 2020; Pomeranz et al., 2019). Fish fulfil many different roles in an ecosystem (Katara et al., 2021), often starting life as planktivores which feed lower in the food web, with some species developing into intermediate (e.g., benthivores) and higher predator feeding guilds (e.g., piscivores) which consume larger prey as they grow. Fish feeding guild biomasses are also responsive to human pressures including fishing and climate change (Garrison & Link, 2000a, 2000b; Thompson et al., 2020).

Analysing change in fish feeding guild biomass can, therefore, provide simultaneous information on ecosystem structure and functioning, by revealing temporal change in the spatial distribution of, e.g.,
planktivore biomass (structure) which is also indicative of change in energy flux between plankton and fish (functioning). Change in feeding guild species richness provides a measure of functional redundancy where, for instance, relatively low and decreasing values highlight areas where a function is supported by only a few species and is at risk. Internationally coordinated surveys with observations on fish species distributions, body sizes and biomass provide extensive data to assess change in fish feeding guilds (Lynam & Ribeiro, 2022). Complementing these, many fish stomach content data that contain information on food web interactions have been collected across the North Atlantic and Arctic Oceans (Arroyo et al., 2017; Cachera et al., 2017; ICES, 1997; Pinnegar, 2019; Smith & Link, 2010; Torres et al., 2013). However, idiosyncrasies in how feeding information has been quantified and reported across different stomach content databases have hampered efforts to depict general feeding traits. Specifically, four feeding guilds relevant to fish are used in OSPAR (The Convention for the Protection of the Marine Environment of the North-East Atlantic) and EU’s MSFD (Marine Strategy Framework Directive) reporting processes, but without an agreed, data-driven method to categorise predators into them (i.e., planktivores, sub-apex demersal, sub-apex pelagic and apex predators; Boschetti et al., 2021; Walmsley et al., 2016; see also https://oap.ospar.org/en/resource-catalogue/enumeration-tables/cemp-enumeration-tables/). Any macroecological assessment of feeding guilds may therefore be confounded because changes could have a methodological basis. This has constrained the contribution of food web indicators to marine ecosystem status assessment. The body size scaling relationship between predators and their prey offers a means to depict general feeding traits across many species, their various life stages, and geographic ranges that have so far been lacking. Predator-prey body size scaling is also important to measure empirically because it can affect the pathway and quantity of energy flux through a food web (Barnes et al., 2010; Brose et al., 2019; Nakazawa et al., 2011; Schneider et al., 2012). Predators feeding on relatively small prey with high predator-prey mass ratios (PPMR) can dampen strong oscillatory dynamics and thus help to maintain stability in food webs (Otto et al., 2007; Rooney et al., 2006) and ecosystem functioning...
Moreover, both predator and prey taxonomy and their traits can be used to predict interactions with typically high PPMRs that may be particularly important (Brose et al., 2019; Reum et al., 2019). For instance, we anticipate that planktivorous and benthivorous fish will have some of the highest PPMRs, especially those that remain in the same feeding guild through ontogeny, while fish species which develop into piscivores could see the largest change (a decrease) in their PPMR across their body size range.

Here we make use of predator-prey body size scaling relationships to draw on multiple stomach content databases and derive comparable fish feeding trait information for the North Atlantic and Arctic Oceans. This new data collation contains observations from 944,129 fish stomachs collected between 1836 - 2020. We use these feeding traits to categorise fish into feeding guilds in a way that is conducive to their application internationally, across ecosystems. Feeding guilds are then applied to survey data collected from across the northeast Atlantic shelf seas to demonstrate macroecological patterns in ecosystem structure and functioning relevant to status assessment advocated by OSPAR.

We test the following hypotheses: i) intra and interspecific body size scaling for predator species is dependent on prey group (e.g., plankton, benthos, and fish prey); ii) multiple distinct feeding guilds are evident based on feeding trait data; iii) feeding guilds capture significant spatio-temporal trends in survey data. Our aim was twofold: to generate empirical estimates of fish feeding traits that could help improve understanding of changes in marine ecosystem structure and functioning; and to achieve international consensus on the best approach to assessing feeding guilds across ecosystems within the OSPAR Maritime Area and in a way that can be readily extended to other areas (e.g., Northeast US continental shelf) and organisms (e.g., invertebrates and mammals).

### 2 Methods

#### 2.1 Stomach contents data

We draw together stomach contents data primarily collected from the North Atlantic shelf seas, with important contributions from the Baltic, Barents and Norwegian Seas (Fig. 1, S1). These data were...
sourced from a combination of previously published and unpublished data including DAPSTOM (Pinnegar, 2019), ICES Year of the Stomach (Daan, 1981; ICES, 1997), the Northeast US continental shelf (Smith & Link, 2010), Northern Spanish shelf (Arroyo et al., 2017), Gulf of Cadiz (Torres et al., 2013), Swedish-, Icelandic-, Norwegian-, French- (Cachera et al., 2017; Timmerman et al., 2020; Travers-Trolet, 2017; Verin, 2018) and German-led surveys (e.g., FishNet, https://www.nationalpark-wattenmeer.de/wissensbeitrag/fishnet/). We have included stomach contents data from outside the OSPAR Area (i.e., Northeast US continental shelf and Baltic Sea) to demonstrate the wider applicability of our approach to defining feeding guilds and because those data have been used to classify feeding guilds previously (Garrison & Link, 2000a). The full data collation contains observations from larvae (<1 g) to adults (up to 351 kg), representing 14,196 unique interactions between 227 predator species and 2158 prey taxa (https://doi.org/10.14466/CefasDataHub.149; Thompson et al., 2024). We provide a summary of data sources, spatial and temporal ranges, and sample distributions in Table 1. All data processing and subsequent analyses were conducted in R version 4.02 (R Core Team, 2020). Predator and prey taxonomy were processed using the “taxize” package (Chamberlain et al., 2020) and assigned to ‘zooplankton’, ‘benthos’, ‘fish’, ‘nekton’ and ‘other’ functional groups after Webb & Vanhoorne (2020) using the “worrms” package (Chamberlain, 2019).
Figure 1: Number of stomach samples on a 1° longitudinal by 1° latitudinal grid.

Table 1. A summary of data sources, spatial and temporal ranges, and sample distributions (see also Fig. S1).

<table>
<thead>
<tr>
<th>Source</th>
<th>Temporal range</th>
<th>Latitudinal range</th>
<th>Longitudinal range</th>
<th>n stomachs</th>
<th>n predator taxa</th>
<th>n prey taxa</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAPSTOM</td>
<td>1836 to 2016</td>
<td>44 to 80</td>
<td>-57 to 50</td>
<td>89500</td>
<td>113</td>
<td>838</td>
</tr>
<tr>
<td>France</td>
<td>2009 to 2018</td>
<td>49 to 51</td>
<td>-2 to 2</td>
<td>895</td>
<td>16</td>
<td>254</td>
</tr>
<tr>
<td>Germany</td>
<td>2019 to 2020</td>
<td>54 to 55</td>
<td>8 to 9</td>
<td>312</td>
<td>21</td>
<td>67</td>
</tr>
<tr>
<td>Iceland</td>
<td>1992 to 1992</td>
<td>63 to 67</td>
<td>-27 to -10</td>
<td>32744</td>
<td>22</td>
<td>506</td>
</tr>
<tr>
<td>ICES - Baltic</td>
<td>1963 to 2014</td>
<td>54 to 60</td>
<td>10 to 24</td>
<td>66829</td>
<td>1</td>
<td>124</td>
</tr>
<tr>
<td>ICES - North Sea</td>
<td>1980 to 2013</td>
<td>51 to 62</td>
<td>-6 to 12</td>
<td>251006</td>
<td>29</td>
<td>781</td>
</tr>
<tr>
<td>Norway</td>
<td>2004 to 2020</td>
<td>69 to 82</td>
<td>-9 to 51</td>
<td>56406</td>
<td>3</td>
<td>348</td>
</tr>
<tr>
<td>Spain</td>
<td>1988 to 2019</td>
<td>36 to 44</td>
<td>-10 to -1</td>
<td>181494</td>
<td>97</td>
<td>354</td>
</tr>
<tr>
<td>Sweden</td>
<td>2013 to 2013</td>
<td>56 to 59</td>
<td>8 to 13</td>
<td>268</td>
<td>1</td>
<td>52</td>
</tr>
<tr>
<td>USA</td>
<td>1973 to 2019</td>
<td>35 to 45</td>
<td>-76 to -65</td>
<td>264675</td>
<td>58</td>
<td>258</td>
</tr>
</tbody>
</table>

2.2 Quantifying intra- and interspecific body size scaling relationships between predators and different prey functional groups

Prey count and biomass observations (wet weight in grams) are needed to estimate predator-prey mass ratios (PPMR), but these were available for only 56% of the stomach contents data. Therefore, to make use of all the data (Fig. 1; Table 1) when assigning fish to feeding guilds, a linear mixed effect model of predator-prey body size scaling was constructed to estimate prey counts or biomass where
either were unavailable; using only data where taxonomy for both predator and prey was resolved (i.e., to species and functional group, respectively), and predator body size, prey body size and prey counts were all available. Major axis regression following Brose et al. (2019) was not suitable because we needed to make predictions of prey body size and minimise the squared residuals in the response (Legendre, 1998). Log_{10} transformed individual prey mass (wet weight in g) was fit as the response, with an interaction term between log_{10} transformed predator body mass (wet weight in g). Random intercepts and slopes were fit for both predator taxa or prey functional group to test if either interacted with predator body mass. Random intercepts were fit for datasets to test for systematic differences in how data were generated: grouped into those from Spain, USA, and ICES, which each follow their own protocols for measuring prey biomass, and all others which represent a mixture of methods from across studies. Random intercepts were also fit for years, sites (a 3 by 3 grid across the study area based on splitting the gradients of longitude and latitude into 3 equal lengths; Fig S1) and the number of stomachs sampled with two levels, 1 = from multiple stomachs and 2 = were stomach samples were pooled (some stomach samples contained in DAPSTOM and ICES Year of the Stomach were pooled at the point of collection for size classes of predator species). We use a Student’s t-distribution to account for heavy tails in the distribution of the response. The Akaike Information Criterion (AIC) was used on nested models to assess the importance of all predictors. The full model had the lowest AIC by >2 units meaning all predictors were retained (Table S1). Model diagnostic plots were performed using the R package DHARMa (Hartig, 2022).

The ‘full’ model was used to predict the mean individual body mass of prey functional groups for predator species of a given size. This enabled us to make use of many observations in DAPSTOM, for instance, which have recorded prey counts but no prey biomass. In such cases, we estimated the biomass of each prey taxa by multiplying the predicted mean individual body mass for their functional group by the observed prey count. Where prey counts were missing, e.g., much of the data from Smith & Link (2010; USA data in Fig. S1), we estimate these by dividing the observed biomass of each prey taxa by the predicted mean individual prey mass for their functional group. We provide R script and...
the data underlying our model, with an example showing how to predict individual prey size based on a list of predator taxa, predator sizes and prey functional groups ([https://github.com/MurraySAThompson/fish-feeding-traits-glmm](https://github.com/MurraySAThompson/fish-feeding-traits-glmm)). Our full model has temporal and spatial information as random effects because we were interested in developing general feeding traits for the study area, irrespective of spatial and temporal gradients. However, the significance of spatial and temporal random effects (Table S1) suggests future work exploring environmental change drivers of predator-prey scaling could be fruitful. All linear mixed effects models were fit using the `glmmTMB` R package (Brooks et al., 2017).

2.3 Classifying predator feeding guilds based on feeding trait data

Predators were categorised by species and size. We use 20 equal size bins to categorise predator mass along a log_{10} transformed gradient from 0.1 micrograms to 190 tonnes, capable of capturing organisms from plankton to blue whales (Table S2). Data for each species size class was then estimated across all available stomach samples (Fig. 1; Table 1), with means calculated for % prey functional group biomass, biomass weighted PPMR (after Reum et al., 2019) and mean individual prey mass (Table S3).

We used directly observed data where available and predictions from our predator-prey body size scaling models where data were missing. Feeding guilds were assigned based on cluster analysis using the ‘ward D2’ agglomeration method on Bray-Curtis dissimilarities between predator diets available in the R `stats` package (R Core Team, 2020).

We compared different methods to classifying feeding guilds where the dissimilarity matrix used in the cluster analysis was generated using either: 1) the biomass of prey taxa (Garrison & Link, 2000a); 2) prey taxa occurrence (Thompson et al., 2020) or; 3) via a novel method where dissimilarities are based on broad feeding traits (henceforth, the biomass, occurrence and trait methods). Feeding traits were log_{10} transformed mean individual prey mass (g), log_{10} transformed mean biomass weighted PPMR, and the mean % biomass contribution to the stomach contents of zooplankton (including fish <0.5g), benthos, nekton (other than fish) and fish (all fish prey ≥ 0.5g); with all variables rescaled to...
values of or between 0 and 1. We tested for differences between these methods to classifying feeding
guilds by comparing them after re-sampling ($n = 1000$) subsets of the data ($n = 30$ unique stomach
samples per predator). Predators with fewer than 30 samples were not classified into feeding guilds.

Compositional change in predators between successively reclassified feeding guilds was used to
determine the ability of each method to consistently classify similar predators in the same guild.

Compositional change was measured using the distance to centroid following analysis of multivariate
homogeneity of groups dispersions (Anderson, 2006). The method with the lowest mean distance to
centroid was determined to have the most robust feeding guild classifications as determined using
analysis of variance tests. First, we tested whether compositional change across feeding guilds was
non-random for each method: distance to centroid (i.e., compositional change) was the response,
with ‘Guild’, and ‘Data’ (i.e., a factor identifying each unique re-sampling event) as predictors. We then
tested for significant differences between methods: distance to centroid (i.e., compositional change)
was the response, with ‘Method’, ‘Guild’, and ‘Data’ as predictors. Significant predictors were
determined using the $F$-test on nested models. Targeted tests for differences between the mean
distance to centroid across methods were carried out using Tukey’s all-pairwise comparisons that
corrects for multiple comparisons in the “multcomp” package (Hothorn et al., 2016).

The ability to classify common feeding guilds across ecosystems (e.g., sub-apex and apex predators)
rather than area-specific guilds (e.g., a feeding guild unique to the North Sea) is another important
quality for a feeding guild indicator to exhibit. We assessed how important spatial gradients were in
our three different approaches to classifying feeding guilds. First, we generated latitudinal and
longitudinal coordinate centroids for each predator by taking a mean across their stomach samples.

We then took a mean across these predator centroids to generate a centroid for all the data, and also
means across these predator centroids but grouped by feeding guild and method to generate method-
specific guild centroids. Next, we measured the distance between the overall data centroid to the
method-specific guild centroids using the geosphere package (Hijmans et al., 2021) and summed
distances for each method. A large sum of distances for a method to the overall data centroid would
indicate that feeding guilds were area-specific, largely made up of predators found close together, and thus spatial gradients would be important determinants of feeding guild structure. The method with the lowest sum of distances to the overall data centroid was deemed to be least affected by spatial gradients and thus preferred.

Four feeding guilds have been called for in OSPAR and MSFD guidance, i.e., planktivores, sub-apex demersal, sub-apex pelagic and apex predators (Boschetti et al., 2021; Walmsley et al., 2016; see also https://oap.ospar.org/en/resource-catalogue/enumeration-tables/cemp-enumeration-tables/), without consensus on how to categorise predators into these guilds. We use four feeding guilds here to help bridge this gap and so that we can elegantly capture a broad set of ecosystem components while exploring guild responses in biomass and species richness in the survey data. Changing the number of feeding guilds could be justified, depending on the question, and is straightforward to implement by taking a higher or lower split in the classification tree. We see this as a strength of our approach because feeding guilds are hierarchically structured much like how taxonomic or other trait information has been organised. We provide a table which details the branches for up to five feeding guilds so future assessments can choose which level of complexity suits their need. We also present axis scores from a non-metric multidimensional scaling analysis of the dissimilarities used in our cluster analysis which provide a more nuanced understanding (i.e., bounded data as opposed to categorical) of different predator feeding traits in relation to others. Moreover, because it is a data-driven, reproducible approach, new information can be systematically integrated to 1) further resolve differences in feeding traits, 2) feeding guild composition and 3) test if changes in predator feeding traits provides evidence for spatially or temporally flexible classifications.

2.4 Using feeding guilds to capture spatiotemporal trends in survey data

The new feeding guild classifications have been applied to processed otter trawl survey data for the Northeast Atlantic shelf seas collected between 1997–2020 (Lynam & Ribeiro, 2022) to reveal spatial and temporal patterns in feeding guild responses. These survey data have been processed specifically
to support state indicators, with observations for the biomass of species size classes standardised to
the area swept for each haul. Survey data corresponding with all our stomach content data, from north
of Norway, Icelandic waters, the Baltic Sea and eastern shelf seas of the USA, have not yet been
standardised and processed in the same way, hence why we haven’t included them here. Extending
this work to assess change in ecosystem structure and function across the study region covered by the
stomach contents data (Fig. 1) represents a key area for future development.

Compared with quarter 2 and 3 (April – September), data from quarters 1 and 4 (January-March and
October – December, respectively) typically have longer time-series available over much of the study
region and so were preferentially selected. Where data from quarters 1 or 4 were not available, otter
trawl data from other quarters were used. Table S4 provides information on the surveys used and
their spatial and temporal ranges. The temporal assessment covers 1997-2020 because the majority
of the surveys considered have at least a near complete time-series covering that period. Longer time-
series do exist for some surveys but including these data would mean we are looking at long-term
change for some areas, but shorter-term change for others which could confound interpretation.

Spatial and temporal change in feeding guild responses were determined for the Greater North Sea,
Celtic Seas, Bay of Biscay and Iberian Coast, and the wider Atlantic. The assessment strata used here
replicate those used for the OSPAR food web indicators: mean-maximum length and size-composition
in fish communities (Lynam et al., 2022; Lynam & Piet, 2022).

Kendall’s τ trend analysis was used to identify areas of significant temporal change in feeding guild
responses based on the relationship between mean haul-level values of feeding biomass and species
richness for each assessment strata and year. Kendall’s τ scores of −1 to +1 represent a 100%
probability of a decreasing or increasing trend, respectively. By using Kendall’s τ, which is rank-based
and non-parametric, we can detect correlations which may be non-linear. Stomach contents data,
prey size predictions, haul-level estimates of feeding guild biomass and their species richness along
with Kendall’s τ correlation coefficients and p have all been made available
(https://doi.org/10.14466/CefasDataHub.149; Thompson et al., 2024).
3 Results

3.1 Unique intra- and interspecific body size scaling relationships between predators and prey functional groups

There were significant differences in the log_{10} transformed scaling relationship between predator and prey body mass captured by the interactions between predator body mass and predator species and prey functional group combinations (Fig. 3). These results support our first hypothesis that predator species can have unique intra- and interspecific body size scaling relationships with different prey functional groups. Fish prey tended to be the biggest, meaning fish-fish interactions tended to have higher intercepts and slopes (Fig. 3a, c, d, h) and thus the lowest mean PPMR, with predator species of the same size consuming relatively small benthic and zooplankton prey (Fig. 3b, e, f, g, i, j, k, l). These models enabled us to estimate prey biomass, counts and predator-prey mass ratios across the different stomach contents datasets, species and size classes useful for feeding guild classifications (Table S3).
Figure 3: A subset of the 498 unique combinations of predator species and prey functional group body mass scaling relationships on log_{10} transformed axes. Blue points = observed values for specific relationships, red points = model predictions for specific relationships, grey points = all observed values. The final plot ‘Predator species and prey groups’ shows the different scaling relationships across the preceding plots, ordered by prey size at maximum predator size with a dashed 1:1 line to show scaling relationships were generally sublinear (i.e., prey increased less than their predators per unit increase in body size).

3.2 Feeding trait data reveal multiple distinct feeding guilds

Feeding guilds captured significant variation in the composition of predators for each cluster-based method (Table 2), confirming our second hypothesis that multiple feeding guilds can be delineated
from the analysis of feeding traits. The occurrence method had the most robust feeding guilds with the lowest compositional change in predators following re-sampling (mean distance to centroid = 0.13), followed by the trait (mean distance to centroid = 0.22) and then the biomass methods (mean distance to centroid = 0.34; randomly generated feeding guild mean distance to centroid ranged between 0.6 – 0.61; Fig. S2; Table S5). The trait method had the lowest sum of distances to the data centroid (2,655 km) followed by the biomass (7,034 km) and occurrence methods (8,757 km; Fig. S3). The trait method was therefore preferred because it could identify multiple distinct feeding guilds even where we consider small subsets of predator stomach contents (n = 30 stomach samples) while being the least affected by spatial gradients in prey taxa.

The four feeding guilds identified using the trait method have been named based on the % biomass of prey functional groups as follows: planktivores, benthivores, bentho-piscivores and piscivores (Fig. 4). Differences between feeding guilds were related to predator size, which correlated positively with piscivory and negatively with planktivory (Fig. S4). Small size classes of species often occur in the planktivore guild, moving to another guild as they increase in size, with multiple medium- to larger size-classes of a species often in the same guild (Table S3). Typically, the biggest fish within and across feeding guilds had the highest PPMR (hence the sublinear relationship in Fig. 3, where prey increased less than their predators per unit increase in body size), yet piscivores were typically the biggest and had the lowest PPMR on average. This apparent contradiction is largely because small piscivores had some of the lowest PPMR values, whereas big planktivores and benthivores had some of the highest values (Fig. 4; Table S3).
Table 2. Analysis of variance results across nested models for each cluster-based method and across methods. Data for biomass, occurrence and trait methods were used to test whether guild (i.e., cluster-based groups) and data (i.e., a factor identifying each unique re-sampling event) captured significant variation in the composition of predators. Data for all was used to test for differences between the different methods (see also Table S5; Fig. S2). The term dropped column indicates which term was dropped from each model (blanks indicate no predictors were dropped), AIC reveals change in model fit, and p-values from F-tests highlight significant change in model fit.

<table>
<thead>
<tr>
<th>Data</th>
<th>Term dropped</th>
<th>Df</th>
<th>AIC</th>
<th>F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass</td>
<td>Guild</td>
<td>3</td>
<td>-18942</td>
<td>67.42</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>999</td>
<td>-19400</td>
<td>1.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Occurrence</td>
<td>Guild</td>
<td>3</td>
<td>-14526</td>
<td>592.83</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>999</td>
<td>-15784</td>
<td>2.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Trait</td>
<td>Guild</td>
<td>3</td>
<td>-14526</td>
<td>592.83</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>999</td>
<td>-15784</td>
<td>2.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>All</td>
<td>&lt;none&gt;</td>
<td></td>
<td>-48990</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Guild</td>
<td>3</td>
<td>-47509</td>
<td>483.29</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Method</td>
<td>2</td>
<td>-43609</td>
<td>3113.44</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>999</td>
<td>-49756</td>
<td>1.19</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Figure 4: Differences between feeding guilds in predator size, individual prey size, predator-prey mass ratio (PPMR) and the % biomass contribution of different prey functional groups to stomach contents. Points represent means for species size classes (Table S3), error bars represent standard error. Bottom panel: PPMR increases with predator body size within (solid, coloured) and across (black, dashed line) feeding guilds.
3.3 Feeding guilds capture significant spatiotemporal trends in survey data

When assigning feeding guilds in the survey data, we were able to classify 92% of the biomass which included 122 species size classes. Many rare predators observed in the survey data (n = 366, representing 8% of the surveyed biomass) remain unclassified due to insufficient stomach contents data (Table S6). The perspective of change in the survey data is therefore weighted towards predators contributing most to community biomass and ecosystem functioning. We found clear spatial structure and regions of contrasting temporal change in feeding guild biomasses and their species richness (Figs. 5-6), confirming our third hypothesis. For instance, significant and spatially extensive temporal decreases in planktivore feeding guild biomass lower in the food web were evident in the Celtic Seas and Bay of Biscay where the biomass of the bentho-piscivore and piscivore feeding guilds higher in the food web has increased (Fig. 5). Benthivore biomass has increased in the southern North Sea, where there has been little change in other feeding guilds. Planktivore, bentho-piscivore and piscivore biomass have all decreased in at least one assessment strata in the northern North Sea. Regions of temporal change in species richness were also different across feeding guilds (Fig. 6). For instance, over large areas in the Celtic Seas, Bay of Biscay and northern North Sea where there was relatively limited change in planktivore species richness, the species richness of benthivores, bentho-piscivores and piscivores all increased (see Fig. S5 for changes in unclassified biomass).
Figure 5: Mean spatial distribution (top row) and temporal change (bottom row) in feeding guild biomass by assessment strata based on otter trawl data. Significant temporal trends are highlighted with a black border and coloured blue or red to depict a decreasing or increasing trend, respectively (Kendall’s τ scores of –1 to +1 represent a 100% probability of a decreasing or increasing trend).

Figure 6: Mean spatial distribution (top row) and temporal change (bottom row) in feeding guild species richness by assessment strata based on otter trawl data. Significant temporal trends are highlighted with a black border and coloured blue or red to depict a decreasing or increasing trend, respectively (Kendall’s τ scores of –1 to +1 represent a 100% probability of a decreasing or increasing trend).
4 Discussion

Our predator-prey body size scaling models enabled predictions of mean individual prey mass for different predator species and prey functional groups across the Atlantic and Arctic Oceans. With these predictions we could estimate key fish feeding traits across species and sizes, such as predator-prey mass ratios and the relative biomass contribution of zooplankton, benthos and fish prey to predator diets. Our feeding trait-based approach to categorising feeding guilds was also less susceptible to large spatial gradients in the composition of prey than previously used methods (Garrison & Link, 2000a; Thompson et al., 2020). By using feeding guilds to assess routinely collected survey data, we revealed contrasting patterns of change in the biomass and species richness of fish feeding at different levels within the food web across the Northeast Atlantic shelf seas. For example, extensive decreases in the biomass of consumers lower in the food web (planktivores) were evident in the Celtic Seas and Bay of Biscay and this contrasted with biomass increases higher up (benthopiscivores and piscivores; Fig 5). These contrasting patterns of change in how biomass and species are distributed across the food web indicate regions of temporal change in marine ecosystem structure and functioning as the relative importance of different energy pathways changes across the fish assemblage. This is because changes in species richness and the distribution of biomass across the food web can influence nutrient uptake and the efficiency of communities in converting nutritional resources into biomass (Cardinale et al., 2012; Wang & Brose, 2018; Worm et al., 2006).

A wide range of prey sizes could be consumed by a given predator species and size and this variability was generally captured well by our models (Fig. 3). Omnivory was also ubiquitous with all prey groups occurring in the diet of all feeding guilds, albeit to markedly different levels (Fig. 4; Table S3). This, along with the variability in our re-sampled feeding guild classifications (Fig. S2), highlights the plasticity of fish feeding behaviour with some individuals of, e.g., a typically planktivorous species having consumed mostly fish prey. Feeding guilds provide a necessary simplification of this complexity as a means to indicate change in marine food webs across ecosystems by taking the typical behaviour of a species size class. We have also provided more nuanced information, including empirical
estimates of predator-prey mass ratios, relative contributions of different prey groups, and from multivariate analysis on the dissimilarity of species size class feeding traits (Table S3; Fig. S4) which could help further unpick what such change means. This information is also widely applicable for quantifying and predicting the effects of different fishing and climate regimes where changes in the distribution of species and body sizes are anticipated (Kleisner et al., 2016; Lotze et al., 2019; Spence et al., 2021; Thompson et al., 2023).

Previous studies which identified feeding guilds used differing approaches, relied on a subset of the stomach contents data we use here and assessed different survey data to one-another (Garrison & Link, 2000a, 2000b; Thompson et al., 2020). It was therefore not clear which method would be optimal as an indicator across ecosystems. Moreover, differences in how feeding trait information has been quantified and reported has hampered the synthesis of stomach contents data. Here, we brought those different stomach contents data together by developing models to predict prey biomass and counts where information was missing. We were then able to compare different approaches to classifying feeding guilds across the datasets and apply the optimal approach to the same survey data. This approach has helped achieve international consensus on how to assess feeding guilds across ecosystems within the OSPAR Area (Thompson, Lynam, et al., 2023). The work we present here represents a development of that pilot indicator, with improvements in how we estimated prey weights (i.e., improved modelling framework that made use of more stomach contents data) which affected feeding guild classifications; and the use of Kendall’s τ trend analysis on the classified survey data (as opposed to Pearson’s correlation coefficients) in order to detect correlations which may be non-linear. Many of the temporal patterns we present here are similar to those in the pilot indicator. The most notable differences appear for feeding guild species richness where values were relatively low (i.e., where relatively little change could have a large influence) in the central and northern North Sea.

We use a trend-based assessment rather than one based on reference limits. Setting thresholds for feeding guilds was outside the remit of this study and will require reference limits for Good
Environmental Status to be established. This could be done through synthesising empirical evidence on feeding guild response to environmental change and human pressure (e.g., following Garrison & Link, 2000b; Thompson et al., 2020) with dynamical modelling capable of predicting ecosystem state under varying levels of human pressure (Link et al., 2010; Rossberg et al., 2017; Samhouri et al., 2010), for instance. Such work would benefit from data products capable of integrating information from different surveys, e.g., acoustic and various trawl gears, that catch a broad range of species and sizes including smaller pelagic species which are likely under-sampled by demersal trawls (Kotwicki et al., 2018; Nnanatu et al., 2020). Developing understanding of the drivers in spatial and temporal change in fish feeding behaviour (Table S1) could also be fruitful, potentially leading to the quantification of energy fluxes from different habitats (e.g., benthos and pelagic), regions, species, and seasons, for instance. Such work could help develop understanding of the connections between the wider suite of indicators which draw on different assemblages and often rely on data collected at different times and at different spatial scales (e.g., Preciado et al., 2023). The inventory of feeding interactions could also continue to develop, e.g., via surveys targeting areas and predators with limited information (e.g., Fig. S5; Table S6), via DNA metabarcoding of stomach contents (Jakubavičiute et al., 2017), inference from similar predators (Gray et al., 2015; Hicks et al., 2019), biotracers (Pethybridge et al., 2018) and predictive modelling (Hernvann et al., 2022; Link, 2004; Petchey et al., 2008). Variability in digestion rates driven by environmental gradients such as temperature and differences in prey sizes and tissue composition could also be modelled to help improve estimates of biomass flux across the food web (Temming & Herrmann, 2003).

Using change in functionally distinct feeding guilds to assess environmental status has been widely advocated to fulfil OSPAR and the Marine Strategy Framework Directive requirements (Boschetti et al., 2021; ICES, 2018; Rombouts et al., 2013; Tam et al., 2017; Walmsley et al., 2016), but international consensus on how to do this has been lacking. We drew together data and expertise from across the North Atlantic and Arctic Oceans to help achieve this. In doing so, we have made empirical estimates of a range of key fish feeding traits that are widely applicable in marine ecosystem science, identified
robust feeding guilds, and revealed how and where ecosystem structure and function was changing across the OSPAR Area. Such information will be critical to help better quantify and predict the effect of human pressures, such as fishing and climate change, on global patterns in marine ecosystem structure and functioning.

5 Conclusions

We make use of multiple stomach contents databases to predict species- and size-specific feeding traits for fish across the North Atlantic and Arctic Oceans. We then developed a repeatable, data-driven workflow that categorises fish based on these feeding traits and show how they can be applied to robustly define the feeding guilds required in OSPAR and the Marine Strategy Framework Directive guidance. This has provided an indicator capable of revealing change in ecosystem structure and function across the OSPAR Area based on routinely collected survey data. For instance, we reveal significant and spatially extensive temporal changes across the food web, with decreases in the biomass of smaller planktivorous fish which contrasted with increases in the biomass of larger more piscivorous fish. The information we have generated can be tailored to fulfil other specific evidence needs, such as improving the parametrisation of ecosystem models and quantifying sustainable levels of human pressure. Our study provides evidence supporting a candidate food web indicator for the OSPAR Area that can be readily extended to other areas and organisms.

6 Competing interests

The contact author has declared that none of the authors has any competing interests.

7 Acknowledgements

Euromarine funded the foresight workshop EFIMBA; MT and CPL were supported via the European Union’s Horizon 2020 research and innovation programme under grant agreement No 869300.
"FutureMARES", the Natural Environment Research Council grant NE/V017039/1 and Cefas Seedcorn ‘Next generation Cefas biodiversity science: from individuals to ecosystems’ (DP433). IP and MAT were supported by the Spanish Ministry for Ecological Transition and Demographic Challenge through the ESMARES project; SS and US were supported via the FishNet project financed by the European Maritime and Fisheries Fund (EMFF), the Federal Government and the State of Schleswig-Holstein; AB was supported by the Swedish Agency for Marine and Water Management (SwAM). Ifremer’s samples were collected during two projects, (1) the CHARM III project within the scope of the INTERREG IVA France-(Channel)-England cross-border European Cooperation Program, co-financed by the European Regional Development Fund (ERDF) and managed by the region Haute Normandie, and (2) the VARITROPH project, funded by European Marine Fisheries Fund and France Filière Pêche.

8 Authors’ contribution statement

MSAT, CL, IP, AB conceived the project; MSAT, IP, VB, AB, MC, PC, EE, GH-M, JJ, FM, SR, SS, US, BES, MAT and CL were involved in designing the methodology and collating the data; MSAT analysed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

9 Data availability

The data we make use of are largely publicly available, cited, and displayed alongside URLs, where applicable. We also provide all observed fish stomach contents data used in our modelling, predictions of fish feeding traits from our linear mixed effects models, haul-level estimates of feeding guild species richness, numbers of fish and their biomass based on scientific trawl surveys from the Northeast Atlantic and the temporal trend values useful for indicator reporting processes (https://doi.org/10.14466/CefasDataHub.149; Thompson et al., 2024).
10 References


Chamberlain, S. (2019). *worrms: World Register of Marine Species (WoRMS) Client* (R package version 0.4.0.)


