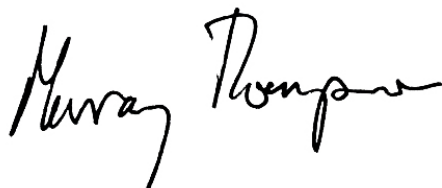
6 February 2025

Dear Earth System Science Data Editorial Office,

Many thanks for your email dated 27/12/2024, detailing two reviewers' comments on our manuscript "Fish functional groups of the North Atlantic and Arctic Oceans". We appreciate the time taken to provide these reviews and some thoughtful suggestions to help us improve the manuscript further.

Below, we have responded, point by point, to the reviewers' comments. We have also made extensive changes to improve and clarify our methodology and more rigorously interrogate our results. The latter includes additional analyses which we describe below.

This work has not been published elsewhere, nor has it been accepted or placed under consideration for publication in another journal or book. The submission of this work has been approved by all the authors listed.

We would like to thank you, the two reviewers and your colleagues in the Editorial Office for dealing with our manuscript in such a thorough and efficient manner. We hope you find this significantly revised and improved version acceptable for publication in Earth System Science Data.

Yours sincerely,

Dr Murray Thompson

Principal Marine Conservation Scientist

Cefas, Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK

Tel: +44(0) 1502 524438 | murray.thompson@cefas.gov.uk | http://www.cefas.co.uk

Comments from reviewers are highlighted in bold and italicised, our responses are in regular font.

RC1:

Thompson et al. present a work where they:

-collate different pre-existing databases of fish stomach contents into a single one, including almost one million "predator' entries from the North Atlantic shelf and arctic seas over the 1936-2020 time period;

-uniformize trait variables in the database (prey individual size, prey biomass, and prey density) by modeling missing entries using a predator-prey mass scaling predictive model

-develop a new categorization of predators (species X mass) into fish trophic guilds based on a broader diversity of trophic traits computed from the uniformized database

-apply the method to a survey dataset from the N-E Atlantic shelf.

The datasets are impressive, the methods are rich, and most are very well-detailed. The application example sufficiently illustrates the potential of the databases while not spoiling future works and investigations. Overall, the datasets could be of use to many academics as well as marine managers, and the overall approach could be inspiring for other biodiversity research based on trophic traits.

We thank RC1 for their kind words and recognising the potential contribution our manuscript offers to biodiversity research.

However, the manuscript needs to be repurposed to align the claimed objectives with the provided datasets and codes.

Suppose the purpose is to produce a uniform dataset of fish stomach contents, which, despite different protocols, enables the display of more information than solely prey occurrence, i.e., species and size-specific feeding traits. In that case, this purpose has been reached as both the codes (glmm) and the databases (before and after uniformization) are provided along with the manuscript, and the methods to simulate the missing data are provided with sufficient details. However, the end products do not fulfill the expectations of the title and claimed objectives of the paper's introduction (i.e., fish guilds).

As claimed by the title and introduction, the purpose is to provide a new categorization of fish guilds based on a broader set of trophic traits (rather than just occurrence and biomass). In that case, the objectives and the provided datasets/codes are inconsistent.

The database for fish trophic guilds, which should be central, is provided for now only as a table in the supplementary material (while a csv is needed for this info to be reused). The codes to generate the clustering that could be used to update or check the robustness of guilds over other regions are not provided. The categorization efficiency based on feeding traits has been tested

against previously developed approaches with similar objectives. Still, we are provided only with very few results, although this should be a major motive for the study. Why this new one is more robust is not sufficiently demonstrated (only through the distance to the centroid, for which we have very few results provided).

Making the manuscript a better fit for the latter point is an easy fix. The introduction should be more informative, not only to address beyond the very specialized marine fisheries scientists but also to better account for previous works on this topic. This is not the first attempt to provide an estimator for fish guilds. Still, it seems this one is expected to leverage databases to be more informative (by simulating the missing traits) rather than by working with the minimally common data in the heterogeneous databases (occurrences and biomass). It is only in the late discussion that previous attempts are acknowledged.

We thank the reviewers for the extensive, detailed and extremely useful input. We have made the following extensive revisions in response:

We now provide the feeding guild tables as part of the published data (see 'feeding guilds.csv', https://data.cefas.co.uk/view/21771) as well as R code and some example data to reproduce our cluster analyses and the tests for differences between feeding guild classification approaches (https://github.com/MurraySAThompson/fish-feeding-guild-classifcation). We have added a sensitivity analysis as part of our R code, where we use information from only directly observed feeding traits and from the DAPSTOM, ICES Year of the Stomach and northeast US continental shelf data (i.e., those that have published prey taxa information) to provide additional support for our decision to select the trait over the biomass and occurrence methods. This demonstrates that across ecosystem comparisons of feeding guilds, which is a fundamental requirement for the indicator, is less reliable using the biomass and occurrence approaches because they are more susceptible to change in prey taxa composition at large spatial scales. Put another way, we demonstrate that the trait approach classified more predators into feeding guilds that were common across ecosystems (e.g., secondary consumers, sub-apex and apex predators), relative to the biomass and occurrence approaches which grouped predators into feeding guilds that were more area-specific (e.g., dominated by species found in specific ecosystems such as the North Sea or northeast US continental shelf).

We agree that the effectiveness of different approaches to reliably and robustly classify predators into feeding guilds is important. Sections 2.3 (Methods) and 3.2 (Results) are dedicated to it, as is much of the 'fish-feeding-guild-classifcation' R code we provide above. We also now state explicitly in our hypotheses that:

'iii) the effectiveness to reliably and robustly classify predators into feeding guilds applicable across ecosystems varies due to whether classifications are based on the biomass of prey taxa, prey taxa occurrence, or broad feeding traits (i.e. PPMR, mean prey body-mass, and the % biomass contribution of different prey functional groups)'.

The primary motive of our study is to provide comparable feeding traits and a single means to classify fish into functional feeding groups across ecosystems rather than compare approaches to classifying groups, which is secondary. Evidence in support of our decision to focus on the trait approach is provided in the text throughout section 3.2, in Table 2, while the associated plots and supporting tables (including those from our new sensitivity analysis, Table S5 and Figure S3) remain in the Supplementary Material because they play a supporting rather than central role.

We have revised the Introduction in line with the suggestions of RC1. The key references which report previous attempts to classify feeding guilds for the region were already cited in the Introduction as well as the Discussion, but it is now clear that the approaches they developed are tested against one-another in this study. In the second paragraph of our Introduction, text has been revised to:

'Despite differing methods being used to classify fish into feeding guilds, their biomasses have been shown to respond to human pressures including fishing and climate change (Garrison & Link, 2000a, 2000b; Thompson et al., 2020)'

And in the hypotheses in our Introduction, we provided an additional hypothesis (iii, see above, there are now four in total) so that it is explicit that we test different approaches to classifying feeding guilds against one-another. We have updated our results to reflect this change with the following text:

'… These results also confirm our third hypothesis that the effectiveness to reliably and robustly classify predators into feeding guilds applicable across ecosystems varies due to whether classifications are based on the biomass of prey taxa, prey taxa occurrence, or broad feeding traits. Results from our sensitivity analysis using only directly observed prey count and weight information reveals that the trait approach had both the most robust feeding guilds and lowest sum of distances to the data centroid, providing further support for our decision to use it to assess change in survey data (Table S5; Fig. S3).'

Specific comments

Unfunctional web links l 95 and l138

We have removed the first link because the citations were already sufficient, and the link displayed lots of other unrelated information which could be confusing. Link on l138 corrected.

paragraph 2.2 and throughout the ms: I found referring to prey and predator individual mass as "body size" confusing. This is probably a common term, but it is improper.

Corrected to individual body mass and for brevity clearly defined as: 'The body-mass scaling relationship between individual predators and individual prey (henceforth, predator-prey body-mass scaling)'

paragraph 2.2 the glmm explanation needs to be clarified: what are the predictors and the response formula? random and fixed terms? all is here but it is only when I checked the formula on the GitHub that I got a clear idea. the glmm formula should be written in plain, and the text should be polished to be better understandable.

We now provide the GLMM formula in full and revised the text for clarity.

Fig 2: many misspelling in it; also, what is the pink category?

Spellings corrected. Figure 2 is now referenced near the start of section 2.4 and appears earlier in the text. We have also simplified the figure, moving information from the pink circles to the initial survey circles on the left. Arrows are now defined in the legend as 'flow of information'.

thoughout the mansucript, a zooplanktivorus fish is a predator, even if found in another fish's stomach. this is an example of the marine jargon's problem of the paper. the "Prey "and "predator" terms should be defined upfront: it would be much more transparent for the reader if, from the beginning, preys were defined as items found in stomach contents and predators as the sampled fish.

We thank the reviewer for this point which will help with clarity. We now define predators and prey in section 2.1 where the data are described, as suggested.

'The full data collation contains observations from larval to adult predators (i.e. fish whose stomach contents have been sampled, ranging from <1g to 351 kg), representing 14,196 unique interactions between 227 predator species and 2158 prey taxa (i.e., prey are defined as organisms found in stomach contents; https://doi.org/10.14466/CefasDataHub.149; Thompson et al., 2024).'

How are uncertainties accounted for in the predictions?
This is a key point. We use the model's fitted values (i.e. best guess) when predicting prey mass for our feeding guild classifications because we are interested in mean individual prey mass while the simulated data help validate the model and are useful in food web research more broadly.

We have now clarified this in our Methods:

In section 2.2: 'We provide R script and the data underlying our model, with an example showing how to predict mean individual prey size (i.e. generate fitted values) based on a list of predator taxa, predator body mass and prey functional groups (https://github.com/MurraySAThompson/fish-feeding-traits-glmm). Here we also demonstrate how to simulate data using the uncertainty measured by our model to help gauge its performance and because variability in individual prey masses is useful in food web research more broadly (Brose et al., 2019; Pomeranz et al., 2019b; Scott et al., 2014).'

And in section 2.3: 'We used directly observed data where available and predictions (i.e. the fitted values) from our predator-prey body-mass scaling models where data were missing'.

Fig 3: higher intercepts and slopes for the biggest fish-fish interactions (l313) cannot be observed in Fig 3.

Corrected. The intercepts and slopes are visible in Fig. 3 panel m, represented by the fitted lines labelled a, c, d, and h.

RC2, Andrew Edwards:

Thompson et al. compile an impressive data set of stomach samples from several countries from the Northeast Atlantic shelf seas. They use this to provide empirical estimates of species- and size-specific feeding traits of fish. This is an interesting study that appears to be well thought out, and do not have any major technical concerns (number 3 might be the only one that requires some work).

We thank Andrew Edwards for his support.

Main comments and suggestions

1. Line 81: suggest defining 'feeding guild species richness', as it is not intuitive whether this refers to the species richness of the predators or of the prey. Line 83 suggests it is referring to prey species, but given a feeding guild is a group of predator species that feed on similar prey, the species richness of the guild would seem to be the number of predator species.

Corrected to: 'Change in the species richness of predators in feeding guilds…'

2. Lines 42 and 262 mention the work being reproducible. It is commendable that the raw data appear to be fully available on the Cefas Data Hub website (in a 47 Mb .csv file!). I checked the GitHub site for the code (L193), and it appears to be well documented and usable by interested parties, which is commendable. Given this, I think it should be mentioned why the data (and therefore the analyses) stop in 2020. If the full workflow was reproducible then it should have been fairly easy to update the analyses. I appreciate that the collating of data into the big .csv file is likely hard to somewhat automate, given the different databases involved, and this might be the barrier to easily updating the analyses (and you have to cut the data off at some point in time). Or maybe all the data were simply not yet available. I'm thinking ahead about whether the analyses could be updated to operationalise the work, given the suggestion of a candidate food web indicator (L501), which presumably would need to be continually updated if it were to be used in fisheries or ecosystem management.

This is a great point. We make use of the Lynam and Ribeiro (2022) dataproduct which was specifically processed for indicator assessments in the recent OSPAR Quality Status Report (2023), with observations for the biomass of species size classes standardised to the area swept for each haul. It was beyond the scope of this study to update the dataproduct with more recent survey data. We now provide the R code necessary to append the feeding guild information to the survey data so that when the dataproduct is updated for the next assessment (Intermediate Assessment planned in 2028), the analysis can be readily updated. We also now clarify this in, section 2.4: 'We also provide the necessary R code (see 'appending_fish_feeding_trait_and_survey_data.R', https://github.com/MurraySAThompson/fish-feeding-guild-classifcation) so that our feeding guilds can be readily appended to new survey data when available and processed as required.'

3. I think the modelling to fill in missing data (L163-167) is sound, but was a sensitivity analysis conducted to compare results when using only data for which all variables are available? I did not see one but it would make the work more convincing. It would at least highlight the impact of the assumptions, and such results may depend on the relative differences between partial and complete data sets (and how those difference differ between areas or guilds).

We now provide a sensitivity analysis (Figure S3; Table S5; and https://github.com/MurraySAThompson/fish-feeding-guild-classifcation) by testing whether we come to the same conclusion that the trait approach outperforms the occurrence and biomass approaches when using only directly observed data. We do this using information from only directly observed feeding traits (i.e., with no information from our predator-prey body-mass scaling models) and from the DAPSTOM, ICES Year of the Stomach and northeast US continental shelf data to provide additional support for our decision to select the trait over the biomass and occurrence methods. Based on these data, the trait approach outperformed the occurrence and biomass approaches in both the ability to generate common feeding guilds across ecosystems and in providing consistent

groupings on small (n = 30) subsets of stomach samples per predator taxa body-mass class (Table S5 and Figure S3).

4. Figure 2 is not referenced in the text, and should appear earlier on than it currently does. The red circles are not defined, nor are the arrows. I think the arrows represent flows of information, but the thin red arrows are a little unclear. Maybe the red circles should just be another step along the black arrows? The rest of the diagram is clearer.

Figure 2 is now referenced near the start of section 2.4 and appears earlier in the text. We have also simplified the figure, moving information from the pink circles to the initial survey circles on the left. Arrows are now defined in the legend as 'flow of information'.

5. L310: The sentence "There were significant differences in the log10 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)" is very unwieldly. I think it could be reduced to something like "There were significant differences in the log10 transformed scaling relationship *of* predator and prey body mass between the different predator-prey functional group combinations (Fig. 3)" The "captured by the interactions between predator body mass and predator species" is somewhat implicit and is not needed, unless I am missing the intent of the sentence. The Figure 3 caption is clearer.

We agree this revision is clearer and have further simplified it to: 'There were significant differences in the predator-prey body-mass scaling relationships between the different combinations of predators and prey functional groups (Fig. 3).'

6. Figure 3: panels are really small given the amount of detail; similarly for Figures 5 and 6. Also, it is a little unintuitive that prey is on the y-axis and predator on the x-axis, given what is being shown is the predator-prey mass ratio (and we normally think of a y-value being divided by an x-value). I guess I'm also used to thinking of predator-prey models where prey would be on the x-axis. I checked some predator-prey mass ratio papers and this does seem to be the usual way to plot these, but maybe just add a comment somewhere. L347 confirms that a higher PPRM is a lower point in Figure 3. It's a little confusing.

We have now simplified the axes and plot labels in Figs 3-4 and reorganised Figs 5-6 (they are now portrait with labels appearing within each panel) so each panel has more space. We have altered the colours of Fig. 3 for contrast to make the points clearer and added the following note within the legend to note why we have organised the axes the way we have: 'Predator individual body mass was fit as the predictor of prey individual body mass in our model, hence appearing on the x- and y-axes, respectively.'

7. L320: How is the subset of 498 combinations chosen to be shown in the figure? This should be stated, as it is a very small subset of the full results.

We now state why we selected these to plot in the legend of Figure 3 and point to our published R script where we provide the means to plot any unique combination of predator species and prey functional group: 'We selected this subset because they represent important predators of fish, benthos and zooplankton across ecosystems (see https://github.com/MurraySAThompson/fish-feeding-traits-glmm for R script to reproduce any unique predator species and prey functional group combination).'

8. L320: Suggest referring explicitly to the panels (and call the bottom one (m)), and explain earlier in the caption that (a)-(l) are ordered by prey size at maximum predator size (it's confusing to

mention that when talking about panel (m)). And the grey points are only in (m), yes? Could be more explicit.

We now refer explicitly to the panels, including the newly labelled panel m, explain earlier in the caption how they are ordered and point out that the now white points in panel m represent all observed values (observed values across all panels are now white points with predictions coloured dark blue, appearing over a grey background for contrast).

9. Figure 4. Again, labelling of panels might help, especially as the first three are a bit different to the second three (but on first reading, only four types of plot are described).

Panels now labelled and explicitly referred to in the legend.

Minor comments:

L103: it might be clearer to change "Predators feeding on relatively small prey with high predator-prey mass ratios (PPMR) can dampen strong oscillatory dynamics" to "Predators feeding on relatively small prey yield high predator-prey mass ratios (PPMR) that can dampen strong oscillatory dynamics". The prey do not have high PPMRs.

Agreed, corrected.

L126 mentions extending to other areas such as the Northeast US continental shelf, but such data are already included.

We now clarify this by adding 'where applicable and routinely collected survey data exist but were outside the scope of our assessment' because routine fish survey data are collected from the Northeast US continental shelf but not in a way that is directly comparable to the survey data we currently make use of in the NE Atlantic, and comparing across the different survey data was beyond the scope of this project. In section 2.4, we clarify this further and highlight this as a key area for future development: '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.'

L134: DAPSTOM needs defining, and I think a few other acronyms might also.

Now defined when first cited.

L135 and Table 1: Figure 1 shows data off the east coast of Labrador, Canada, but no Canadian data sources are mentioned, just Northeast US continental shelf. This is not accurate if the term is meant to include the Canadian data.

Those data are contained in DAPSTOM and this is depicted in Fig. S1.

L169: the model descriptions would help by also having a simple equation of the models (in word form is fine as is often done for such models), as it would summarise the approach.

Now included.

L177 'were' should be 'where'

Corrected.

L381 and L382: 'lower in the food web' is a bit superfluous given the 'planktivore' reference, though maybe '(i.e., lower in the food web)' would work. Similarly for 'higher in the food web'.

Corrected, as suggested.

L381: The Celtic Seas and Bay of Biscay need defining or highlighting on a map.

These are now depicted in a map (Figure S2) which accompanies Table S4 and pointed to in our Methods, section 2.4: 'Table S4 provides information on the surveys used and their spatial and temporal ranges and Figure S2 depicts survey locations within OSPAR regions (e.g., Celtic Sea, North Sea).'

L407 and likely elsewhere: Suggest hyphenating 'body size' when used as an adjective.

Now hyphenated throughout where we refer to predator-prey body-mass scaling

L408: Really it's just 'parts' of the Atlantic and Arctic Oceans, not the full extent (which is how it reads). L415 mention of 'Northeast Atlantic shelf seas' is more correct.

We agree the feeding traits are applicable to areas where we have stomach contents data and have revised this to: 'North Atlantic and Arctic Ocean's shelf seas'.

L491: "We make use" and "We then developed". Be consistent with tense.

Checked and corrected.

While making our revisions, we found some incorrect values for our trait model AIC and *F*-values in Table 2 which had been mistakenly copied from the rows above. These have been corrected and did not affect our conclusions.