1 Answers to reviewer 2

Author Response

Dear reviewer,
Thank you for your comments and taking the time to review the paper. Before I get to answering your specific comments, I want to respond to what I agree is a very general criticism of the paper that should be properly addressed, which is related to using data from different geographic areas. The issue is that there is very little data on MeHg bioaccumulation at the base of the food web, and almost nothing in the North Sea. This complicates model validation and is the reason why the decision was made to simply try to use all data available. However, 2 recent papers came out with a lot of data, of which one paper, by McClelland et al., 2024 samples 476 benthic animals from two locations in the Canadian Arctic. While this is a different location, it does allow us to evaluate if certain patterns between feeding strategies are consistent in samples from the same geographic location. Because of this, I suggest to expand the paper by a third component, where I evaluate if we see an effect of feeding strategy on MeHg bioaccumulation. I suggest to add the analyses below:

Suggested edit

A third test of our hypothesis using field data from a single study to see if the same dynamics as occured in our model are present in these observations. We used MeHg bioaccumulation and trophic level data from 476 individuals across 53 taxa of benthic invertebrates as published by McClelland et al. (2024). These data were selected as it is the laregest study we could find with both trophic level and MeHg concentrations. When several animals of the same group where sampled, the dataset presents mean values per group per location, which we use as one datapoint in our analyses. Although feeding strategies in the dataset were broadly aligned with our classifications, we reassigned them to match the functional groups in our model. For example, shrimps were categorized as generalist feeders, which group is not present in McClelland et al. (2024), and isopods, which can be small benthic predators, were labeled as deposit feeders because their prey type is not represented in our model.
The data is sampled from two locations in the Canadian Arctic, Cape Bathurst (CB), which has a depth of 22 m and is located at 70°41′42.79″ N, 128°50′21.34″ W, and the eastern coast of Herschel Island in the Mackenzie Trough (MT), which has a depth of 116 m and is located at 69°36′44.96″ N, 138°33′45.25″ W. It must be noted this dataset is selected as it is extensive, but the region does have notable differences to the North Sea, where our model is run. It has extensive ice cover in winter, it is colder and is geographically distant from the model location. While the Arctic ecosystem differs from the North Sea, the use of a single, internally consistent dataset allows us to isolate feeding strategy effects better than using the global dataset and verify our conclusions.

Author Response

I will add this to the methods section at line 218

Suggested edit

Evaluation of the model using a single dataset

To isolate the effect of the feeding strategy on MeHg bioaccumulation, we first transformed MeHg concentrations to their natural logartitm and fit a linear model with trophic level as predictor using the base R `lm()` function. The significance of the deviation from the predicted MeHg concentration at the trophic level was assessed using a one-sample t tests. To improve interpretability, we calculated the percentage differences using $\text{Percentage difference} = 100 \times \left(\frac{\text{MeHg}_{\text{obs}}}{\text{MeHg}_{\text{pred}}} - 1 \right)$ based on the residuals of the linear fit. This is visualized on a bar graph showing the percentage difference in MeHg concentration caused by the feeding strategy. The errors bars represent the mean ± 1 Standard Error (SE). The same analysis was then performed to estimate differences in MeHg bioaccumulation related to phylum.

As a final test, linear models were fitted on the natural logarithm of bioaccumulated MeHg concentrations using trophic level, phylum, and feeding strategy as predictor variables (using the `lm()` function in R). Estimated marginal means (EMM) for each feeding strategy were calculated with the `emmeans()` function of the emmeans package and compared against the overall mean to assess deviations. This analysis was also performed separately for the MT and CB locations to verify the consistency of the effects of the feeding strategy. The EMMs where transformed to a percentage difference with the earlier used equation and the estimated percentage difference due to feeding strategy and its signifance is shown.

Author Response

I will add this to the result and discussion about the literature study result section (3.1). At line 351.

Suggested edit

The fit of the linear model against the natural logarithm of the bioaccumulated MeHg is shown in Fig. 1. The R^2 is similar with 0.43 and 0.45 in the CB and MT respectively, while the bioaccumulation is a bit lower in the CB ($\log(\text{MeHg}_{\text{BA}})=0.137+1.14*\text{TL}$) compared to that in the MT ($\log(\text{MeHg}_{\text{BA}})=0.256+1.39*\text{TL}$), where MeHg_{BA} is the bioaccumulated MeHg in ng Hg mg^{-1} d.w. and TL is the trophic level. The influence of the feeding strategy on MeHg bioaccumulation based on the results of McClelland et al. (2024) is shown in Table 1. While the only significant effect is deposit feeders in the MT having less MeHg than would be expected on their trophic positions, some other effects are consistent, albeit not significant in both locations. The strongest effect is that filter feeders have consistently higher MeHg (residuals are 0.7 in the CB and 0.8 in the MT), while deposit feeders have lower MeHg (residuals are -0.2 in the CB and -0.5 in the MT). The results of the same analyses for phyla are shown in Table 2. Here we see two consistent significant effects. Molluscs have elevated MeHg levels (residuals are 0.61 in the CB and 0.51 in the MT) while arthropods have reduced MeHg values (residuals are -0.35 in the CB and -0.30 in the MT). The percentage difference in MeHg bioaccumulation per feeding strategy is visualised in Fig. 2 and per phyla in Fig. 3. The average percentage difference between observed and the expectation based on trophic level is 102% and 128% in the CB and MT respectively for filter feeders while deposit feeders have 19 and 37% less MeHg than would be predicted based on trophic level alone in the CB and MT respectively. If we repeat the analysis per Phylum we see that molluscs have highly elevated MeHg levels with an increase of

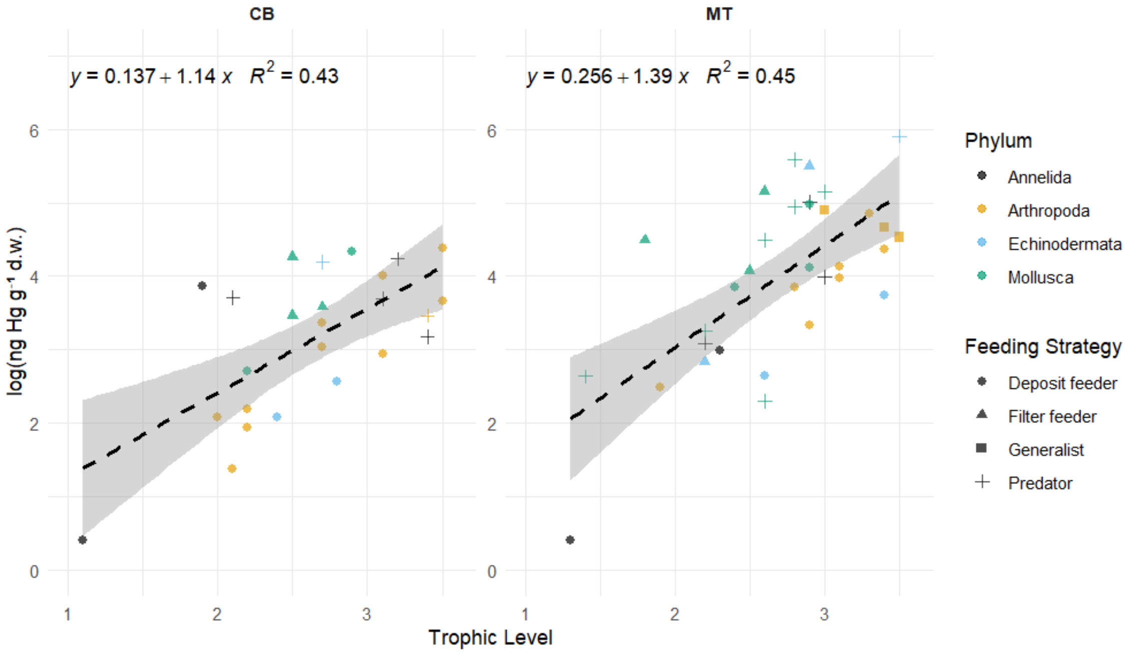


Figure 1: The linear fitted model between the natural logarithm of the bioaccumulated MeHg in ng Hg mg^{-1} d.w. and the Trophic Level in the data presented by McClelland et al. (2024). For extra clarity the different Phyla shown with different colors while the different feeding strategies are marked with different symbols. In both the CB and MT setups there positive relationship between trophic level and the bioaccumulation of MeHg, but R^2 is only 0.43 and 0.45 in the CB and MT respectively, so it does not explain the full variation in bioaccumulation.

66% and 85% respectively in the CB and MT. The largest reduction in observed MeHg compared to the predicted values based on trophic level is in Arthropoda, here there is a decrease compared to the predicted values of 29% and 26% in the CB and MT respectively.

The results of the final analyses are shown in Table 3. Despite the lower sample size, which reduces statistical power, the results indicate that filter feeders consistently have higher MeHg levels than predicted based on their trophic position and phyla, while deposit feeders tend to have lower MeHg concentrations. These results are stronger in the MT with a change of 118% and -40% in filter and deposit feeders respectively than in the CB with a change of 7.2% and -14.8% in filter and deposit feeders respectively. It must be stated that this final analysis is included to address potential concern between the co-correlation of phyla and feeding strategy, but the problem of the reduced sample size has to be addressed. In the CB, where the increase in MeHg in filter feeders is rather low after correcting for both trophic level and feeding strategy, there are only 3 filter feeders, which are molluscs, and they make up 3/5 mollusc samples in this location, meaning that results should be seen with skepticism as filter feeders and molluscs have too much overlap. On the other hand, in the MT, there are 5 filter feeders from multiple phyla (Mollusca and Echinodermata) and the effect is considerably stronger with filter feeders having 118% more MeHg than would be expected based on their trophic level and phyla.

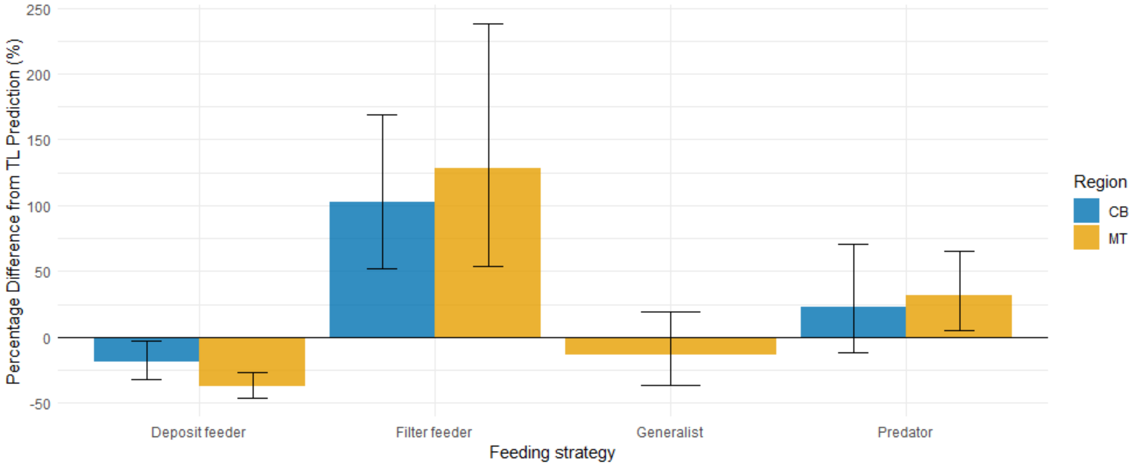


Figure 2: Percentage difference from trophic level predicted MeHg concentrations by feeding strategy, with error bars showing 1 SE. In both CB and MT regions, filter feeders have elevated MeHg levels relative to trophic level based expectations, while deposit feeders are reduced. Predators display higher MeHg than predicted, though the effect is smaller than in filter feeders; in CB, this increase does not exceed one SE. Generalist feeders have a slight reduction compared to expectations, but this is well within one SE, and were not present in CB for cross-region comparison.

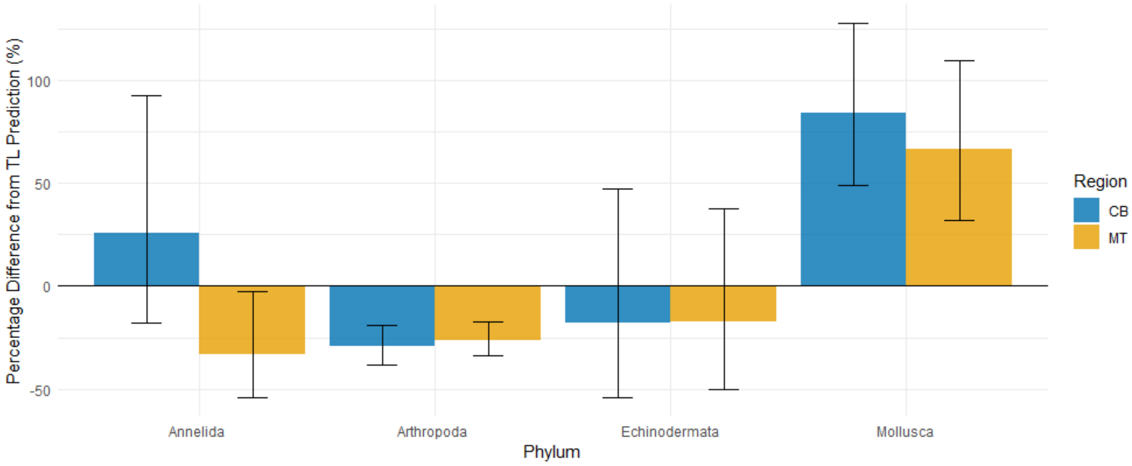


Figure 3: Percentage difference from the predicted MeHg bioaccumulated based on trophic level per phyla, the error bars represent the mean 1 SE. The notable phyla are Mollusca and Arthropoda, while Mollusca have a notable increase in MeHg bioaccumulation compared to the prediction of 85% and 66% respectively in the CB and MT, there is a reduction of 26% and 29% in Arthropoda in the CB and MT respectively. Annelida are inconsistent with an increase in the CB and decrease in the MT compared to the predictions. Echinodermata have a mean reduction compared to the prediction in both the CB and the MT, but the SE is much larger than the mean effect.

Table 1: Mean residuals (±SE) of log(MeHg) by feeding strategy and region, after trophic level correction. Significant deviations ($p < 0.05$) are marked with *.

Region	Feeding Strategy	n	Mean Residual ± SE	p-value
CB	Deposit feeder	16	-0.208 ± 0.181	0.268
CB	Filter feeder	3	0.704 ± 0.286	0.133
CB	Predator	6	0.203 ± 0.331	0.568
MT	Deposit feeder	15	-0.467 ± 0.159	0.011*
MT	Filter feeder	5	0.824 ± 0.395	0.105
MT	Generalist	3	-0.143 ± 0.319	0.698
MT	Predator	12	0.277 ± 0.226	0.247

Table 2: Mean residuals (±SE) of log(MeHg) by phylum and region, after trophic level correction. Significant deviations ($p < 0.05$) are marked with *.

Region	Phylum	n	Mean Residual ± SE	p-value
CB	Annelida	6	0.229 ± 0.424	0.612
CB	Arthropoda	11	-0.349 ± 0.137	0.0294*
CB	Echinodermata	3	-0.198 ± 0.584	0.767
CB	Mollusca	5	0.611 ± 0.211	0.0446*
MT	Annelida	5	-0.405 ± 0.377	0.343
MT	Arthropoda	12	-0.304 ± 0.111	0.0196*
MT	Echinodermata	5	-0.188 ± 0.509	0.730
MT	Mollusca	13	0.509 ± 0.231	0.0482*

Table 3: The effect of feeding strategy on MeHg bioaccumulation per Region compared to the prediction accounting for both trophic level and feeding strategy. Significant ($p < 0.05$) is marked with *. There is still a consistent increase in filter feeders and a consistent decrease in deposit feeders. This is effect is larger and significant in the MT with a relative percentage increase of 118% in filter feeders and a decrease of 40% in deposit feeders.

Feeding Strategy	**% Diff (MT)**	**p-value (MT)**	**% Diff (CB)**	**p-value (CB)**
Deposit feeder	-40.0	0.034*	-14.8	0.888
Filter feeder	118.0	0.034*	7.2	0.888
Generalist	-25.9	0.563	–	–
Predator	3.0	0.895	9.4	0.888

Author Response

Updated statistics due to extra data in the analyses

In addition to the extra papers added, a final small change was made: the assimilation efficiency of the deposit feeder was set to 0.12, as this is more in line with the literature presented by Dutton and Fisher (2012). To clarify, this and the addition of two more papers, of course, change the statistics and also highlight your concern about different base levels of Hg. The new statistical analyses are presented in Table 4. Because the new data has a higher baseline of MeHg, it indeed changes (in this case reduces) the agreement between the model and the observations in absolute terms. It does, however, not mean that the correlation between feeding strategies and elevated MeHg levels is different, as this remains consistent. To increase transparency of model fit, I will both show the R^2 based on the Pearson correlation coefficient and the Residual-Based R^2.

Suggested edit

Finally, the goodness of fit of the model is shown both based on the Pearson correlation coefficient and based on the residuals. It is important to note the difference between the R^2 based on the Pearson correlation coefficient ($\mathrm{R^2}_{\mathrm{Pearson}}$) and the Residual-Based R^2 ($\mathrm{R^2}_{\mathrm{Residual}}$). They are defined as $R^2_{Pearson} = (\mathrm{corr}(y_{\mathrm{mod}}, y_{\mathrm{obs}}))^2$ and the $R^2_{Residual} = 1 - \frac{\sum(y_{\mathrm{obs}} - y_{\mathrm{mod}})^2}{\sum(y_{\mathrm{obs}} - \bar{y}_{\mathrm{obs}})^2}$, where y_{obs} are the mean observed values and y_{mod} are the mean modeled value per feeding strategy. The $\mathrm{R^2}_{\mathrm{Pearson}}$ describes how well a linear relationship between our model and the data holds, thus, if this is high, the models explain the observed differences between feeding strategies well, but it is not effected by the absolute concentrations. This means that the $\mathrm{R^2}_{\mathrm{Pearson}}$ evaluates if our model accurately represents differences between feeding strategies, but it is not affected by a bias in the data due to mismatch between the baseline of MeHg bioaccumulation. The $\mathrm{R^2}_{\mathrm{Residual}}$ measures how well the modeled values match the observations, including absolute values, which is both affected by the ability of the model to predict the effect of feeding strategy on bioaccumulation, and by overal concentrations of iHg and MeHg in biota. Both $\mathrm{R^2}_{\mathrm{Pearson}}$ and $\mathrm{R^2}_{\mathrm{Residual}}$ give a value below 1. Closer to 1 means a better fit while below 0 means that the the model does not outperform estimating the data based on the mean.

Suggested edit

The $\mathrm{R^2}_{\mathrm{Pearson}}$ is high ($>$0.86) for MeHg in all setups and exceeds 0.99 in the AS model, indicating that the model captures the relative differences between feeding strategies well. For iHg, performance is lower, particularly in the Northern North Sea ($\mathrm{R^2}_{\mathrm{Pearson}} = 0.24$). The ability to reproduce absolute bioaccumulation is more limited. Only the AS model in the Southern North Sea shows good agreement ($\mathrm{R^2}_{\mathrm{Residual}} = 0.96$), while all other setups yield $\mathrm{R^2}_{\mathrm{Residual}} < 0$, suggesting that using the mean of the observations would outperform the model.

This can be explained, as baseline MeHg levels vary between sampling regions. Notably, the AS model in the Southern North Sea performs well both in reproducing overall MeHg levels and in explaining variability across feeding strategies. Even when excluding predators and top predators, $\mathrm{R^2}_{\mathrm{Pearson}}$ remains high (0.80), suggesting that feeding strategy effects are captured across trophic levels and are not just driven by high MeHg levels in predatorial feeding strategies. In contrast, the Northern North Sea has a high $\mathrm{R^2}_{\mathrm{Pearson}}$ (=0.94) but low $\mathrm{R^2}_{\mathrm{Residual}}$ ($<$0) so it captures the effect of feeding strategies while not being able to resplicate absolute MeHg concentrations.

Table 4: Statistical analysis of model performance for iHg and MeHg levels by feeding strategy for Southern North Sea (SNS) and Northern North Sea (NNS).

	SNS				**NNS**			
	iHg		**MeHg**		**iHg**		**MeHg**	
	N. Bias	**BF10**	**N. Bias**	**BF10**	**N. Bias**	**BF10**	**N. Bias**	**BF10**
Suspension	0.18	0.11	0.85	0.011	-0.18	0.11	0.17	0.0072
Filter	1.48	0.47	-0.33	0.046	1.45	0.069	-0.71	0.054
Deposit	1.01	0.078	-0.44	0.046	0.34	0.034	-0.77	0.052
Generalist	1.31	0.18	-0.39	0.047	1.23	0.14	-0.75	0.055
Predator	0.41	0.081	-0.42	0.10	0.07	0.079	-0.71	0.074
Top predator	-0.22	0.089	-0.80	0.51	-0.46	0.10	-0.93	0.59
Predator (AS)	0.41	0.081	-0.31	0.093	0.07	0.079	-0.75	0.11
Top predator (AS)	-0.22	0.089	-0.12	0.35	-0.46	0.10	-0.67	0.45
Overall Model Performance								
RMSE	40		132		40		147	
NRMSE	0.36		0.35		0.35		0.39	
R^2_{Pearson}	0.61		0.86		0.24		0.94	
R^2_{Residual}	<0		<0		<0		<0	
RMSE (AS)	40		22.8		27		108	
NRMSE (AS)	0.36		0.061		0.26		0.29	
R^2_{Pearson} (AS)	0.61		>0.99		0.24		>0.99	
R^2_{Residual} (AS)	<0		0.96		<0		<0	

2 Direct answers on reviewer comments

Reviewer Comment

Since the primary objective of the study is to model Hg bioaccumulation, I recommend that the model evaluation be presented as part of the Results and Discussion rather than the Methods. This change would strengthen the narrative and reduce redundancy—many of the points currently discussed in Section 3 could be streamlined. I suggest restructuring Section 3.1 to serve as the model evaluation, followed by subsequent sections explaining key discrepancies between model output and observations (currently in Section 4).

Author Response

I agree and I will give the manuscript a restructuring of the result sections. As you mentioned, the model evaluation should be seen as results. We will expand the evaluation to expand on the evaluation of the Hg cycling and pelagic bioaccumulation, a suggested edit is presented below, as the manuscript was very low on the evaluation of pelagic Hg cycling and bioaccumulation. Then I would merge the model and result section together as section 3. After this we will then discuss the results of how the model and literature study agree or disagree.

The benefit of doing this is that we can than state the AS model is better, and therefore focus the rest of the paper on using this model.

1. **Introduction**
2. **Methods**
   (a) The model (Current Sections 2.2–2.10)
   (b) Explain the model evaluation using all data (Current Sections 2.1 and 2.11)
   (c) Explain the Evaluation of the model conclusion using a single dataset (based on the use of the McClelland et al. (2024) dataset)
3. **Results**
   (a) Model evaluation: Hg cycling and pelagic bioaccumulation (described below)
   (b) Discuss patterns that occur in the model (Current Sections 3.1–3.4)
   (c) Evaluate the observed patterns using all data (Current Sections 4 and 5)
   (d) Evaluation of the observed patterns using a single dataset (based on McClelland et al. (2024), as described above)
4. Discuss the observed rol of the feeding strategy based on the results (Discussed below).
5. **Model limitations** (Current Section 6)
6. **Future work**: Further improvement of this model (Discussed below)
7. **Summary and conclusion (Current section 7)**

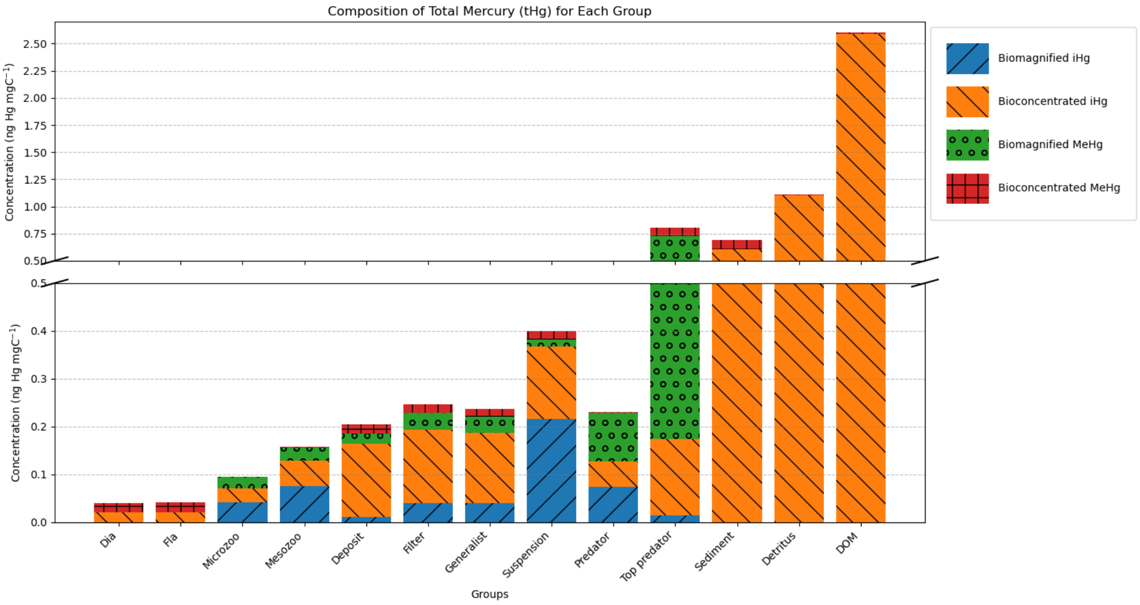


Figure 4: Modeled bioconcentration and biomagnification of iHg and MeHg. Partitioning to detritus and DOM is colored as bioconcentration. The y-axis is cut to show the high and low values. Notably is the high iHg to mgC ratio of detritus and DOM, leading to elevated iHg in suspension feeders. Additionally, higher trophic level animals have higher biomagnified MeHg.

Author Response

We will also update Fig. 6 to show the bioaccumulation in the AS model, as shown in Fig. 4. This because this is the better performing version of the model and we can therefore better anlyse this version.

Author Response

I would also add this section ad the end of the model evaluation segment

Suggested edit

2.1 Model Limitations and Potential Improvements

Evaluation of the GOTM-MERCY-ECOSMO setup reveals performance strengths and weaknesses of the model. There is a strong correlation between the modeled and observed effect of the feeding strategy on the bioaccumulation of MeHg in both the Northern and Southern North Sea setups. However, the Southern North Sea setup outperforms the Northern setup in both pelagic Hg cycling and megabenthos bioaccumulation assessments. In the Northern North Sea setup, the mesozooplankton tHg levels are higher than expected, although this could not be properly validated due to a lack of measurements. What is more notable is our model results show a reduction in MeHg bioaccumulation in the Northern North Sea compared to the Southern North Sea, while this is not present in the observations presented by McClelland et al. (2024). Since the Northern North Sea does have a good agreement with observations in the correlation between the feeding strategy and the bioaccumulation of MeHg, but the overall model performance is lower due to lower than expected MeHg concentrations in all benthic groups, it is likely that the model underestimates the flux of MeHg from the pelagic to the benthic system. In the

Southern North Sea setup, macrobenthos can feed directly on the phytoplankton bloom which results in a major exchange of organic material and MeHg between the benthic and pelagic food webs. In the Northern North Sea this is not possible, and the consumption of detritus, either by filtering it from the water or feeding on it from the sediment, is the main food source for megabenthos. The higher performance of the model in the Southern North Sea compared to the Northern North Sea indicates that this model could be further improved by enhancing the MeHg dynamics in deeper water. Further model evaluation is needed to determine whether the observed underperformance is due to limitations in the representation of sediment Hg chemistry, pelagic Hg speciation in deep waters, interactions between megabenthos, detritus, and MeHg, or a combination of these factors. A potential improvement here would be to implement a variable partioning coefficient of iHg and MeHg to detritus to account for the particle-concentration effect as is discussed by Coquery and Cossa (1995) for the Hg in the North Sea.

Author Response

I will add this to the beginning of the model evaluation section to show the Hg cycling and bioaccumulation in the pelagic is in line with observations.

Suggested edit

2.2 Evaluation of the Hg cycling and pelagic bioaccumulation

The marine cycling and speciation of Hg, in addition to the bioaccumulation in phytoplankton and zooplankton, is an essential driver of the bioaccumulation of iHg and MeHg in the benthic food web. Observed and modelled dissolved tHg concentration, the percentage of tHg that is MeHg, and the Hg content of phytoplankton and zooplankton is shown in Table 5. The concentration of dissolved tHg and the percentage of MeHg of dissolved tHg are compared to observations by Coquery and Cossa (1995), while the bioaccumulation of tHg in phytoplankton and zooplankton is compared to observations by Nfon et al. (2009). It must be noted that the observations by Nfon et al. (2009) are not from the North Sea itself, but from the better-studied nearby Baltic Sea. The average dissolved tHg concentration is 1.7 and 2.1,pM in the Northern and Southern North Sea, respectively. This is well within 1 standard deviation of the 1.7±0.7 pM observed by Coquery and Cossa (1995). The MeHg concentration was observed to be between 0.5 and 4.3% of tHg, with an average of 3% in the North Sea. The percentage MeHg in our model is 2.3% and 2.0% on average, which falls well within that range.

For bioaccumulation, we could not find separate reliable measurements of MeHg and iHg in phytoplankton and zooplankton in the North Sea, and we therefore evaluated the tHg content. This was measured in diatoms to be 10±5 ng Hg mg^{-1}. This means that the mean bioaccumulation in our model in diatoms is lower, with 5.8 ng Hg mg^{-1} and 9.0 ng Hg mg^{-1} in the Northern and Southern North Sea, respectively, but still within 1 standard deviation of the measurements. Observations labeled as zooplankton and mysis were compared to our modeled microzooplankton and mesozooplankton, respectively. All modeled values fall within 1 standard deviation of the observed tHg concentration, with one exception: mesozooplankton in the Northern North Sea, which is 13.5% more than 1 standard deviation above the observations. This is mostly driven by a high iHg content, as the MeHg content is similar in microzooplankton and mesozooplankton.

This similarity in MeHg is caused because, even though mesozooplankton have a higher trophic level, they prefer to feed on larger diatoms which have less MeHg than smaller flagellates, which are preferred by microzooplankton. The high iHg content, especially in the Northern North Sea, is caused by the consumption of detritus by zooplankton in the model. While there is a shortage of data on bioaccumulation at the base of the food web, especially in the North Sea, which complicates model evaluation, the dissolved tHg concentration, the percentage of MeHg, and the tHg content of phytoplankton and zooplankton agree well with observations. With the exception of the 13.5% elevated tHg content in Northern North Sea mesozooplankton, all modeled values fall within 1 standard deviation of the observations. Because of this, we conclude that the model replicates marine Hg cycling and bioaccumulation at the base of the food web in line with observations, with the caveat that we do not have measurements of zooplankton in the Northern North Sea to verify or reject the elevated levels in that setup.

Table 5: Dissolved total Hg (pM), MeHg (% of tHg), and Hg concentrations in biota (ng Hg mg^{-1} d.w.) across North Sea regions.

	Observed	NNS	SNS
tHg$_{\text{dissolved}}$ (pM)	1.7 ± 0.7	1.7 ± 0.26	2.0 ± 0.28
MeHg (% of tHg)	3 (0.5–4.3)	2.3 ± 0.23	2.0 ± 0.31
Diatoms (ng Hg mg^{-1})	10 ± 5	7.0 ± 1.1	8.3 ± 1.6
Flagellates (ng Hg mg^{-1})		13.9 ± 3.0	14.3 ± 3.0
Microzooplankton tHg (ng Hg mg^{-1})	37.5 ± 31.3	67.4 ± 29.3	40.3 ± 11.4
Microzooplankton MeHg (ng Hg mg^{-1})		7.1 ± 2.1	10.5 ± 2.7
Mesozooplankton tHg (ng Hg mg^{-1})	62.5 ± 12.5	86.7 ± 15.1	72.3 ± 19.6
Mesozooplankton MeHg (ng Hg mg^{-1})		6.9 ± 2.6	10.5 ± 1.7

Author Response

I would suggest to add the below expansion of the discussion about what drives the role in feeding strategy.

Suggested edit

Combining the results of the model and the literature studies is difficult due to the high uncertainty in most drivers, including the organic material content of dry weight and result should be viewed with skepticism. The data used in this study was prepared by McClelland et al. (2024) to mimic consumption by predators: for small arthropods, their skin was not removed, but for gastropods and bivalves, the shell was not taken into account for the weight as predators would typically not eat this. The concentration of MeHg per unit energy is arguably the key measure in bioaccumulation. Predators need to ingest a specific energy amount, so if a prey is composed of half organic material and half non-organic components, such as shell, its MeHg content per dry weight is halved. However, predators would consume double the dry weight to obtain the energy, and thus the same MeHg. In general, the energy appears to be consistent with Ash Free Dry Weight (AFDW), as such ideally we would normalize all measurement of MeHg bioaccumulation per AFDW (Weil et al., 2019).
Unfortunately, doing this conversion reliably on published data is not possible as AFDW varies with the age and body size of animals, which information is not generally available (Eklöf et al., 2017). That being said, we find it convincing that both our model results, the literature study in which we aggregate all measurements, and the literature study where we take samples from a single study all suggest that filter feeders have elevated MeHg.

The difference in the uptake of MeHg and organic carbon is also the main driver in our model of the difference caused between the bioaccumulation in filter and deposit feeders. In the model this is mostly driven by the lower feeding efficiency of MeHg presented by (Dutton & Fisher, 2012). While the lower assimilation efficiency of MeHg from sediment is understudied, it is supported by the limited available literature. In addition to the direct work by Dutton and Fisher (2012) analysing transfer efficiency of MeHg in the sediment to biota, it is also demonstrated that the concentration of MeHg in the sediment correlates poorly with the MeHg concentration in all biota, except infauna (Buckman et al., 2019; Chen et al., 2009). If MeHg would easily be transferred from the sediment to deposit feeders, there would be a much stronger correlation between sediment and deposit feeder MeHg concentrations.
The elevated iHg levels in suspension feeders are not related to this. There is no literature on the assimilation efficiency of MeHg when DOM is consumed, but the parameter is not altered in our model compared to when other food is consumed. The increased level of iHg in our model must, therefore, be caused by another driver. In our model, this is because the partitioning of iHg to DOM is stronger than that to detritus. As mentioned earlier, the partitioning of iHg to detritus and DOM is based on the K_{ow} which has the values of log10(6.4) and log10(6.6) for detritus and DOM respectively. So while biota bioaccumulates MeHg stronger than iHg, DOM has a stronger binding to iHg than to MeHg. Since iHg is generally much more common (about 98% of tHg in our model) this results in a much higher bioaccumulation of iHg than MeHg. This aligns with the observations by Orani et al. (2020), which found MeHg to be only between 1% and 29% in DOM consuming sponges.

Author Response

Additionally I would add the following component to the concluding remarks

Suggested edit

Filter feeders and molluscs typically accumulate more MeHg than other organisms at similar trophic levels. This pattern is consistent not only in our models but also in available data. This raises a hypothesis that expanding bivalve populations, as seen in mussel or oyster farming, might affect MeHg bioaccumulation in higher trophic levels. This is supported by the observations that fish in lakes invaded by zebra mussels have higher Hg levels than fish in lakes without zebra mussels Blinick et al. (2024). However, such ecological alterations also impact other bioaccumulation factors like biomass distribution and trophic interactions. While our findings support the role of filter feeders and molluscs in MeHg dynamics and higher bioaccumulation in top predators, the complexity of ecological situations requires further case specific studies to understand if and when bivalve communities lead to increased MeHg transfer.

Modeling studies can help our understanding of the factors influencing MeHg bioaccumulation, but ability to accurately predicht MeHg bioaccumulations needs to be carefully validated. Our findings reveal that filter-feeding molluscs and DOM-utilizing suspension feeders have different Hg bioaccumulation patterns compared to other megabenthos. Modeling bivalve aquaculture or DOM-consuming suspension feeders can help explore their potential role in altering MeHg bioaccumulation. Understanding how functional traits like feeding strategy influence MeHg transfer remains key to improving both predictive models and environmental risk assessments.

Reviewer Comment

The purpose of Figure 3 is unclear. It is not evident why the authors chose to use Hg data from different ecosystems and plot them against trophic level (referred to as feeding strategy in the figure). Since ecosystems differ in baseline inorganic Hg and MeHg concentrations, the MeHg–trophic level relationship should be examined within each ecosystem independently.

Author Response

The reason we showed this was to be able to compare the relationship between Trophic Level and bioaccumulation between the base model and the AS model with the relationship in observations. While the original intention was the show that while MeHg bioaccumulates, iHg does not, I agree that this is thus standard accepted that showing this is not necessary. Because the figures does allow to quickly compare the slope of the bioaccumulation with the model I would put the figure in the Supplementary information. Here an interested reader can see how the slope of the bioaccumulation of MeHg and iHg compare to the global dataset while not making it a core part of the paper. Instead we performed the ecosystem specific analyses based on the McClelland et al. (2024) dataset of which Fig. 1 will be shown in the paper.

Reviewer Comment

For model evaluation, I strongly suggest plotting modeled versus observed concentrations of speciated Hg (inorganic and MeHg) for each modeled feeding strategy. This would provide a clearer and more direct assessment of model performance.

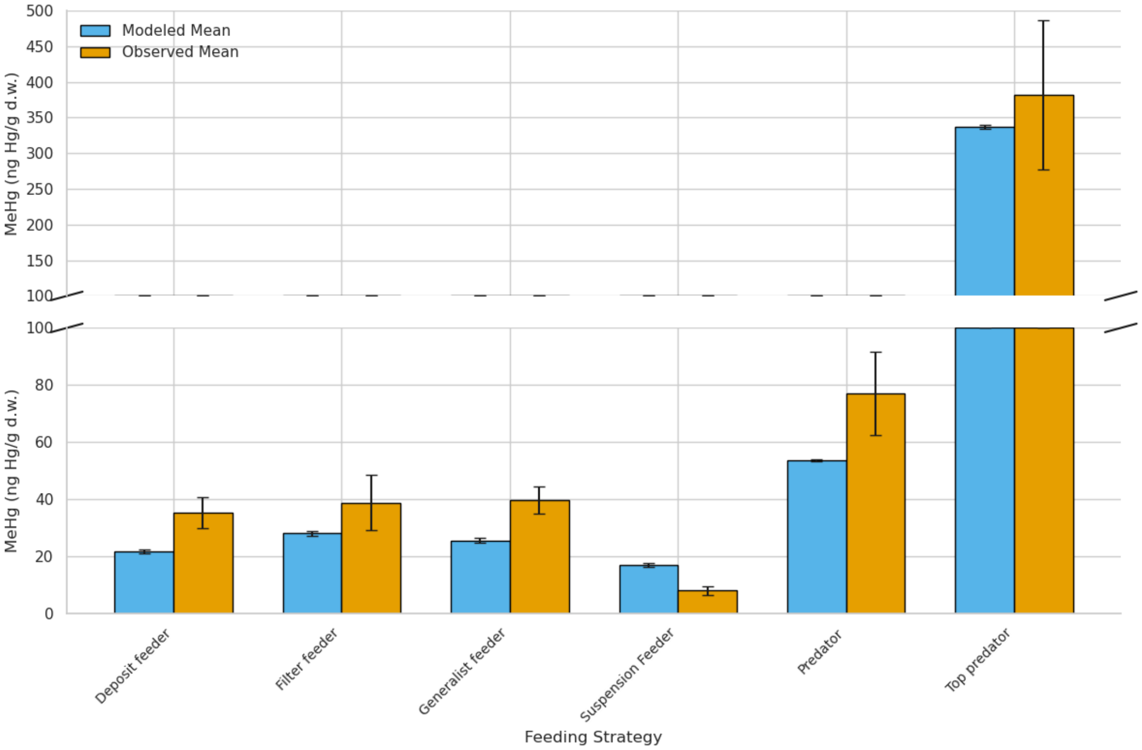


Figure 5: Mean bioaccumulation of MeHg in both model and observations is shown, with error bars representing 1 SE. The model's predictions do not consistently match observations within 1 SE, yet they display a comparable trend: Top Predators have the highest MeHg levels, followed by predators, with generalists and filter feeders showing similar MeHg, which are higher than those found in deposit feeders. Both model and observations show that suspension feeders have the lowest MeHg levels among the feeding strategies.

Author Response

We will do this by making a barplot for both MeHg and iHg with a broken axis for MeHg so we can show both higher and lower trophic levels. I would also suggest to only do this for the MeHg so it can show both lower and higher trophic levels. We will show Fig. 5 for MeHg and Fig. 6 for iHg. Based on your comment another suggestion would be to show it as is shown in Fig. 7, but I think the bar plots give the same information in this case in a way that is very easy to undestand for readers, even readers that are not used to seeing model data.

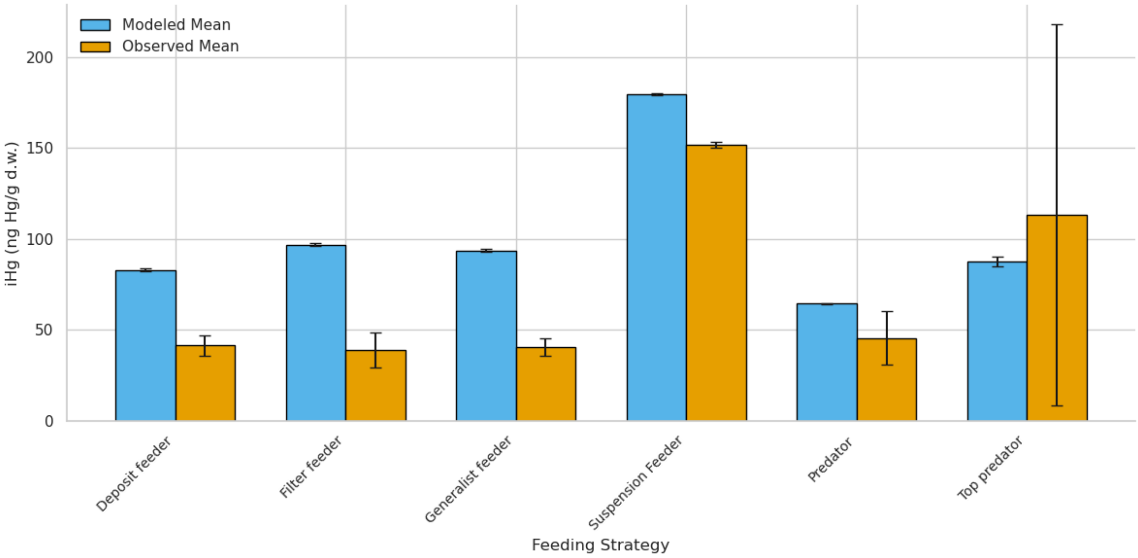
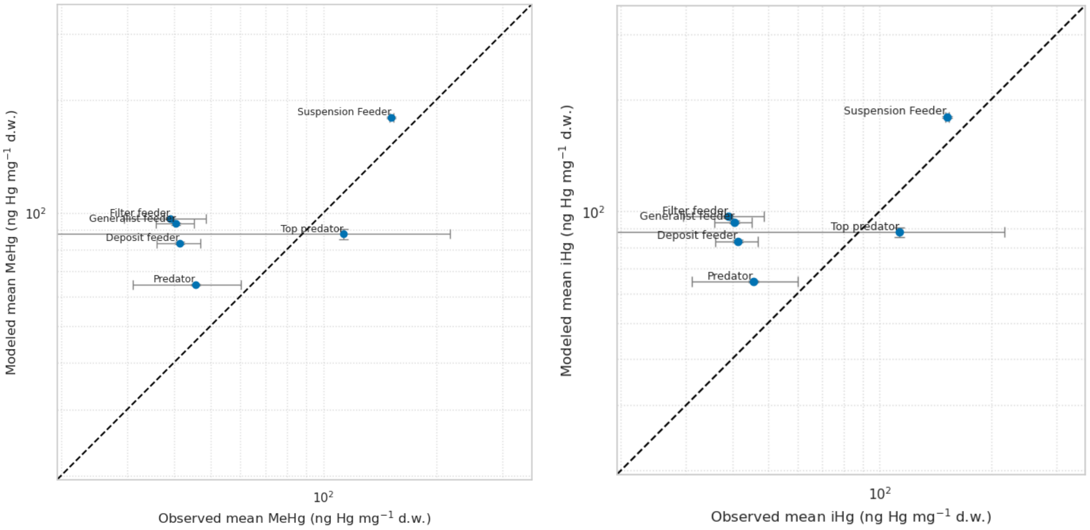


Figure 6: Mean bioaccumulation of iHg in both the model and observations. Error bars represent 1 SE. The model accurately shows increased iHg levels in suspension feeders. However, for all other feeding categories, except top predators, the iHg content is overestimated. In top predators, observed iHg levels are higher and not reflected in the model, though these elevated iHg levels have a very high SE.


(a)

(b)

Figure 7: Direct plot of modeled and obseved MeHg (a) and iHg (b).

Reviewer Comment

Section 3.2.3, which addresses the effect of feeding strategy on bioaccumulation, is central to the manuscript's aims, yet it is not discussed in sufficient depth. In contrast, the manuscript devotes substantial space to explaining Hg vs. trophic level patterns (Section 3.2.4), which are already well-established in the literature. I recommend condensing the discussion in 3.2.4 and focusing more on how feeding strategies influence MeHg and inorganic Hg transfer, particularly in benthic food webs.

Author Response

I addition to the suggestion to restructure made above, I would further adress this comment by merging section 3.2.4 about the effect of trophic level and 3.2.5 about the allometric scaling. The whole reason why this model performs better is because it better captures the trophic dynamics. These dynamcis are, as you mention already well understood and thus not a novel result of this model. This combined with the removal of Fig. 3 would shorten the part of this paper dedicated to this. Combined with the above discussed expansion of the drivers behind the importance of feeding strategy

Reviewer Comment

Figure 4 is difficult to interpret. It is unclear whether the data are empirical or simulated. A more straightforward approach might be to present Hg concentrations across feeding strategies as a bar chart with error bars. If the intent is to show correlations between feeding strategies, a correlation coefficient would be more appropriate.

Author Response

I will remove this plot. The bar plots that you suggested convay the same message in a clearer way, that there are differences in iHg and MeHg bioaccumulation. Combining this with a plot that compares it to the model saves space and a reducent figure.

Reviewer Comment

Section 3.3, on allometric scaling, should appear earlier in the manuscript. When reading Sections 3.2.3 and 3.2.4, I repeatedly found myself wondering about the effects of allometric scaling on the results. Figures 7 and 8 could be consolidated to allow readers to compare model performance with and without allometric scaling more clearly.

Author Response

I agree and I hope the suggested restructering would improve the manuscript in your opinion. Here I will show Fig. 7 and Fig. 8 in the first part of the Result section. Then the difference between the 2 models can be seen. This can be compared in to observations in the third part of the result section. This would allow for a good comparison between the previous version of the model and the AS model, and an easy comparison to field observations that the AS model is an improvement in higher trophic leel compared to the previous version of the model. I will update the images as shown belowin Fig. 8

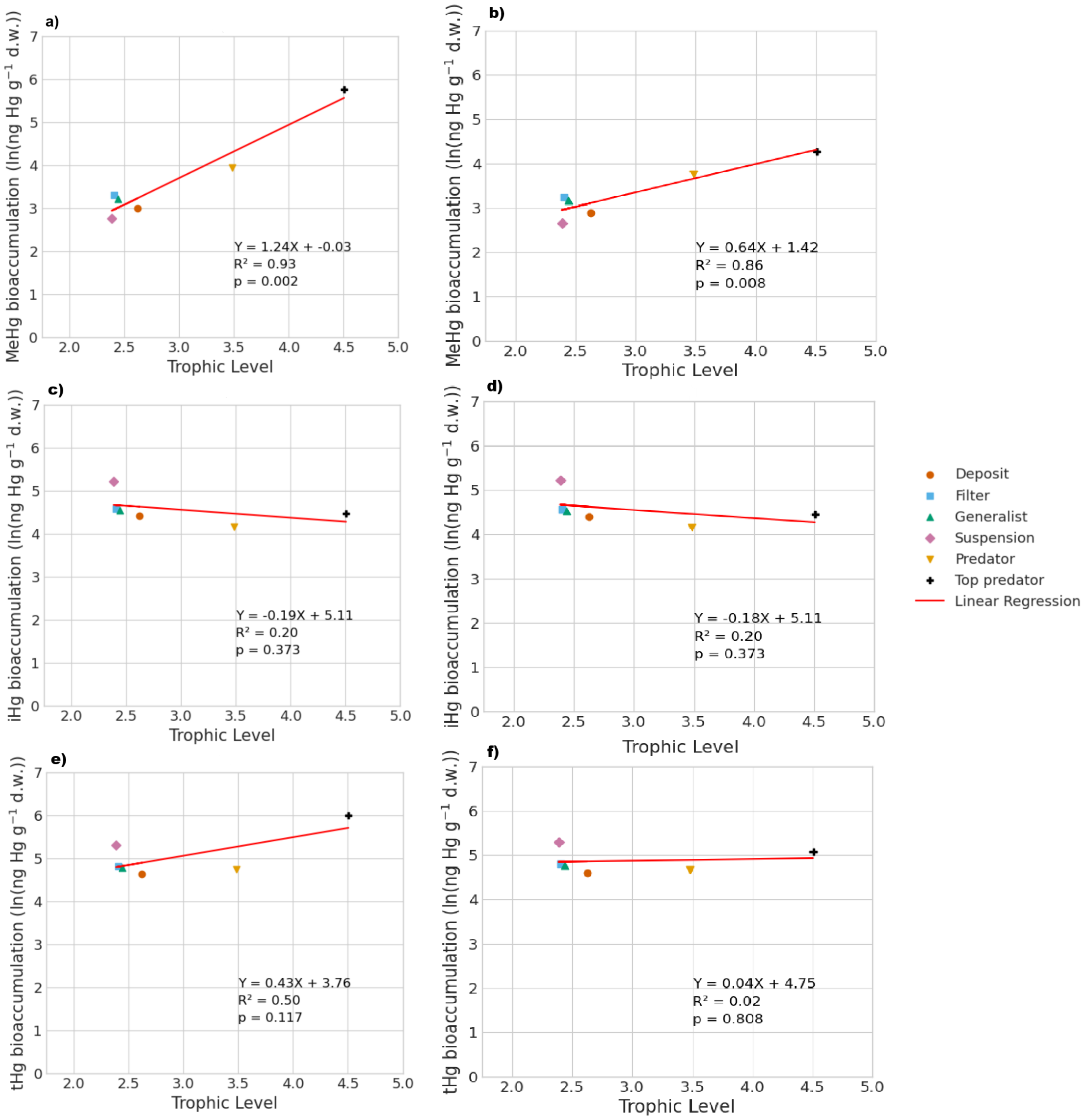


Figure 8: The inluence of trophic level on the bioaccumulation of MeHg, iHg, and tHg in both the AS (panels a, c, e) and the base model (panels b, d, f). In the AS model, the relationship with trophic level is stronger, where ln(MeHg)=1.24TL-0.03, compared to the base model, which is ln(MeHg)=0.64TL+1.42. TL represents trophic level, and MeHg is expressed in ng Hg g^{-1} d.w. For iHg, the bioaccumulation patterns are nearly identical, with ln(MeHg)=-0.19TL+5.11 for the AS model and ln(MeHg)=-0.18TL+5.11 for the base model, both showing a weak inverse correlation with trophic level, largely due to higher iHg levels in low trophic level feeders. In terms of tHg, there is a higher increase in bioaccumulation in the AS model (ln(MeHg)=0.43TL+3.76) compared to the base model (ln(MeHg)=0.04TL+4.175), driven by the stronger association between MeHg and trophic level in the AS model.

Reviewer Comment

Lines 340–350: This content would be better integrated into the allometric scaling section.

Author Response

Agreed, I will move this to the evaluation (section 3(a) as suggested above.)

Reviewer Comment

As the authors note, the model is implemented for the North Sea, yet many of the empirical datasets used for comparison originate from other regions. This mismatch raises concerns about the validity of the model evaluation. Comparing model output to observations from ecologically distinct systems—each with different baseline Hg and MeHg levels, food web structures, and biogeochemical conditions—complicates interpretation and undermines the credibility of the evaluation. I strongly recommend either (1) limiting the model evaluation to observed data from the North Sea, or (2) running separate models parameterized for the specific ecosystems from which the empirical data are drawn.

Author Response

I agree with this concern of data being drawn from different locations. Unfortunately, there is not enough data from the North Sea to evaluate the model purely on data from the North Sea. I think the biggest concern is that there might be a cocorrelation between certain areas with high MeHg and the feeding strategies that are commen there. I hope that involving the evaluation component purely based on the McClelland et al. (2024) data adresses this concern to some degree. The model is not purely designed to be a predictive model for different megabenthos species in the North Sea, rather it is aimed to analyse if feeding strategy is a significant driver of MeHg bioaccumulation. The aim is to show that running a model in an idealized 1D setup resembling coastal conditions results in differences in the bioaccumulation caused by feeding strategies. Notably, our model shows higher MeHg in filter feeders compared to deposit feeders and extremely elevated levels of iHg in suspension feeders consuming DOM. We find it convincing the same pattern occurs in our model, the global dataset and the McClelland et al. (2024) dataset, but I fully agree that better measurements are necessary to validate models before they can be used in a predictive capacity. Our hope is that this manuscripts helps the message that studies to measrurements of biaoccumulation in the benthic food web can be relevant to increase our understandig of bioaccumulation in higher trophic levels.

References

Blinick, N. S., Link, D., Ahrenstorff, T. D., Bethke, B. J., Fleishman, A. B., Janssen, S. E., Krabbenhoft, D. P., Nelson, J. K. R., Rantala, H. M., Rude, C. L., & Hansen, G. J. A. (2024). Increased mercury concentrations in walleye and yellow perch in lakes invaded by zebra mussels.

Buckman, K. L., Seelen, E. A., Mason, R. P., Balcom, P., Taylor, V. F., Ward, J. E., & Chen, C. Y. (2019). Sediment organic carbon and temperature effects on methylmercury concentration: a mesocosm experiment. *The Science of the total environment*, *666*, 1316.

Chen, C. Y., Dionne, M., Mayes, B. M., Ward, M., Darron, Sturup Stefan, & Brian, J. P. (2009). Mercury Bioavailability and Bioaccumulation in Estuarine Food Webs in the Gulf of Maine. *Environ. Sci. Technol.*, *43*, 1804–1810.

Coquery, M., & Cossa, D. (1995). Mercury speciation in surface waters of the north sea. *Netherlands Journal of Sea Research*, *34*(4), 245–257.

Dutton, J., & Fisher, N. S. (2012). Bioavailability of sediment-bound and algal metalsto killifish Fundulus heteroclitus. *Aquatic biology*, *16*, 85–96.

Eklöf, J., Austin, Å., Bergström, U., Donadi, S., Eriksson, B. D., Hansen, J., & Sundblad, G. (2017). Size matters: Relationships between body size and body mass of common coastal, aquatic invertebrates in the Baltic Sea. *PeerJ*, *2017*(1), e2906.

McClelland, C., Chételat, J., Conlan, K., Aitken, A., Forbes, M. R., & Majewski, A. (2024). Methylmercury dietary pathways and bioaccumulation in Arctic benthic invertebrates of the Beaufort Sea. *Arctic Science*, *10*(2), 305–320.

Nfon, E., Cousins, I. T., Järvinen, O., Mukherjee, A. B., Verta, M., & Broman, D. (2009). Trophodynamics of mercury and other trace elements in a pelagic food chain from the Baltic Sea.

Orani, A. M., Vassileva, E., Azemard, S., & Thomas, O. P. (2020). Comparative study on Hg bioaccumulation and biotransformation in Mediterranean and Atlantic sponge species. *Chemosphere*, *260*, 127515.

Weil, J., Trudel, M., Tucker, S., Brodeur, R. D., & Juanes, F. (2019). Percent ash-free dry weight as a robust method to estimate energy density across taxa. *Ecology and Evolution*, *9*(23), 13244–13254.