

## Response to Referees' Comments

### Referee #1

This manuscript presents an impressive multiproxy dataset. The paper addresses a timely topic: how climate variability might affect diel vertical migration of mesopelagic fishes and potential implications for carbon export. The manuscript contains lots of information. The Discussion contains many ideas. The paper has clear potential. However, in its current form the manuscript is structurally diffuse and not sufficiently focused on the central question it aims to address. The headline inference chain: warming → reduced DVM → reduced biological pump efficiency—is repeatedly asserted but not yet fully supported with a sufficiently tight mechanistic bridge and uncertainty treatment for a top-ranked journal. The Methods are also very lengthy, disproportionately long in the main text, and key assumptions are not stress-tested.

Thank you for your time and effort in reviewing our paper and for your kind remarks.

In this work, we present a new methodological approach that can allow testing the hypothesis that warming may reduce fish-DVM. DVM is viewed as a process that mechanistically contributed to the biological carbon pump (BCP); testing this is not within the scope of our study, but it has been supported by numerous studies in the past (e.g. Aumont et al. 2018; Pinti et al. 2023; Aksnes et al. 2023; all these are already cited in the Introduction). Briefly, phytoplankton takes up carbon dioxide from the atmosphere to photosynthesize, and the carbon then passes through the food web to fish, including mesopelagic DVM-performing fish/ These mesopelagic fish, through DVM, export the carbon directly into the twilight zone. In the absence of DVM, carbon would be exported from the euphotic zone fully exposed and would thus most likely be remineralized on the way. Instead of this, the carbon that is consumed at the surface by fishes performing DVM is protected and cannot be remineralized while it remain within the bodies of these fishes. When the fishes dive again, and respire, excrete or die in the deep, they inject the carbon directly into the mesopelagic zone, by-passing hundreds of meters of remineralization. In this way, a reduced DVM would be expected to lead to reduced BCP. We can clarify this point further in the Discussion.

Regarding the Methods section, we understand that it is quite long, but these are really all the methodological information needed to replicate our study. We can propose however to move methodological information and results that are secondary to our message into the supplement.

We address your specific comments one-by-one below.

The manuscript needs a substantial rewrite and restructuring to sharpen the key inference. At present, the narrative is diluted by extensive, mixed threads (preservation/taphonomy, modern analogues, age models and multiple paleoceanographic events) that are not always clearly separated into what is measured vs what is inferred. The impression is that the central inference becomes difficult to evaluate. Therefore, I strongly recommend a structural rewrite that follows an explicit hypothesis-testing logic. For example, in the Introduction, it would be to define the core question and hypotheses/predictions concisely (what would you expect to observe in otolith isotopes if DVM weakens/strengthens? what would you expect in

productivity/export proxies?). in the Methods, keep only what is essential for evaluating the central claims in the main text; move methodological background and secondary proxy details to Supplement where possible.

Thank you for this suggestion. We will indeed move the secondary methodological information and secondary results to the supplement.

The results hold separate proxy reconstructions from derived interpretations; avoid mixing inference with description.

We will separate these clearly. Thank you.

In the Discussion, focusing on one main thread would be more efficient. inferred vertical habitat/DVM → mechanisms → implications for export (cautiously) → alternative explanations → limitations.

Thank you! Yes, this will be better. We will restructure the Discussion following the suggested structure.

I feel that the connection between DVM reduction and the biological pump narrative is one step too far without a tighter bridge. This is the most important contribution the paper wants to make, but currently it reads stronger than the data can directly support. The manuscript does not directly measure biological pump efficiency (which was what I was most interested in); it infers ecosystem function from inferred depth habitat changes, metabolic reasoning, and productivity indicators. This is great, but that may be plausible and the logic needs to be made explicit and bounded.

In this paper we do not attempt to prove that DVM contributes to BCP. This has been the topic of other studies in the past, and it has been quite well supported (please see also our response to this above). Our aim is to propose a new method for detecting changes in fish DVM patterns in the geological past, which is something that has never been done before. Specifically, in the abstract we state “We present a multi-proxy, ecosystem-level assessment of paleoenvironmental changes in the Eastern Mediterranean during the Middle Pleistocene (marine isotope stages MIS 23–18; 923–756 ka B.P.) and use the carbon and oxygen isotopic composition of fossil fish otoliths to assess the impacts of these changes on DVM and their possible implications on the biological pump.” Also in the last paragraph of the Introduction, we say “We investigate here changes in the DVM patterns of two of the most common mesopelagic fishes ...”. We will check again the wording throughout the manuscript and make sure to be as clear as possible on this point.

I recommend a clear distinction between direct observations (otolith  $\delta^{18}\text{O}/\delta^{13}\text{C}$ , foraminiferal isotopes, TOC/productivity proxies) and derived quantities (lifetime-average depth; inferred DVM intensity; inferred pump efficiency) in the manuscript. And replace causal language, such as replacing “consequently reduced pump efficiency” with appropriately conditional phrasing unless you provide a quantitative bridge (“consistent with / may imply / compatible with”).

We already have separate subsections for each of these methods and their results, and a quite detailed Methods section that clearly states that lifetime-average depth is modeled. Regarding DVM intensity, this is an interpretation and therefore only addressed in the Discussion section

already. As explained above, the link between DVM and BCP is not the topic of this study and has investigated and supported by other studies.

Also, please define precisely what you mean by “biological pump efficiency” in this context and which proxy observations are intended to constrain this important key character.

See above. We do not provide proxy observations for biological pump efficiency.

In terms of definition, according to De la Rocha & Passow (2014): “The biological pump is the set of processes by which inorganic carbon (e.g., carbon dioxide) is fixed into organic matter via photosynthesis and then sequestered away from the atmosphere generally by **transport into the deep ocean**. This may be accomplished by the passive sinking of particulate organic matter, through the **vertical migration of zooplankton**, or the downwelling of surface waters rich in dissolved organic matter. In addition to concentrating carbon in the deep sea, the biological pump also significantly affects the distribution of a number of different chemical constituents of ocean water. There is keen interest in being able to predict both the overall capacity and the efficiency of the biological pump in different places and at different times (including in the future). The physical environment, the type of phytoplankton present, the activities of zooplankton, the presence of biominerals and clay minerals, and the structure of the food web all play important roles in determining both the capacity and efficiency of the biological pump on local and regional scales, complicating efforts to portray the biological pump in models.”

As such, DVM is part of BCP. We can add this definition in the Introduction.

Reference:

De la Rocha, C.L., Passow, U. (2014) 8.4 – The biological pump. In: Heinrich D. Holland, Karl K. Tuerkian. Treatise on Geochemistry (2<sup>nd</sup> edition). Elsevier. pp. 93–122. <https://doi.org/10.1016/B978-0-08-095975-7.00604-5>.

Maybe one explicit “prediction vs observation” table would be helpful. what patterns would support DVM weakening vs community reorganization vs changes in water-mass structure? Just a recommendation, but this is not necessary if the manuscript structure is improved significantly. I do think alternative hypotheses are needed. For example, warming-driven community reorganization is another strong candidate for your results. Even if warming coincides with a change in inferred vertical habitat, warming can also drive community-level reorganization. E.g., composition, dominance, size structure, turnover that alters ecosystem function without requiring the specific mechanism proposed (i.e., your reduced DVM per se). This alternative (or complementary) pathway needs explicit discussion.

I recommend adding a dedicated subsection that contrasts your interpretation with other plausible mechanisms and relevant recent work, for example:

Salvatteci et al. (2022, Science) emphasize warming impacts on ecosystem and community organization and trophic structure. Britten & Sibert (2020, Nat. Commun.) on climate-related changes in pelagic communities. Lin et al. (2023, Science Advances) demonstrated long records showing warming effects on mesopelagic fish production/diversity and community reorganization.

The studies you are suggesting are very important, but the impact of climate warming on fish community composition has already been explored by us on different works and are beyond the scope of this study. Salvatteci et al. (2022) investigated climatic impacts on fish

community composition in the SE Pacific by reconstructing the past fish fauna across late Pleistocene–Holocene using scales and vertebrates, and detected a shift in the composition of the assemblages during the last interglacial. Britten & Sibert (2020) investigated fish production and trophic transfer efficiency in the Early Eocene Climate Optimum based on a fossil fish ichthyolith (teeth and scales) accumulation rates. Finally, Lin et al. (2023) reconstructed the mesopelagic fish assemblages over the last 460,000 years in the western Pacific through fossil otoliths and related fish diversity and accumulation rates to climatic variability.

In Agiadi et al. (2018), we had already reconstructed the fish community composition in the eastern Mediterranean (to species level and from the same site as the present study) based on otoliths across Pleistocene glacial–interglacial periods and associated shifts in community composition with the expansion and contraction of the biogeographic ranges of North Atlantic species (including mesopelagic species), particularly after 1.5 Ma. Then, in Agiadi et al. (2023) we went back to the same site and reconstructed the abundance and body size of mesopelagic fishes across middle Pleistocene glacial–interglacial interval, demonstrating warming induced a decline of ~35% in mesopelagic fish body size at the community level through distribution range shifts leading to increased relative abundance of smaller mesopelagic fishes. We refer to our previous studies already in the text, specifically:

- In the end of the Introduction: “Although some aspects of the response of marine organisms to the paleoclimatic changes that occurred in the Mediterranean during the Pleistocene have been studied so far ... fish communities (Agiadi et al., 2011, 2018; Girone and Varola, 2001), as well as the body size of ... fishes (Agiadi et al., 2023)...”
- In the Discussion, we address these alternative possibilities (also citing our previously published results): “Moreover, there is no decline in absolute or relative abundance of myctophid otoliths in the sediments corresponding to interglacial periods of the Lindos Bay Formation (Agiadi et al., 2018, 2023).”

We must stress at this point that our otolith isotopic analyses are performed on monospecific samples. Therefore, the results reflect the DVM of those two species and should not be directly affected by changes in fish community composition. Indeed, food-web restructuring could impact the DVM of *Ceratoscopelus maderensis* and *Hygophum benoiti* (since DVM is motivated by trophic interactions) and therefore could be detected as shift in the single-species oxygen isotopic ratios.

We can add further explanation to clarify this point in the Discussion.

Reference:

Salvatteci, R., Schneider, R.R., Galbraith, E., Field, D., Blanz, T., Bauersachs, T., Crosta, X., Martinez, P., Echevin, V., Scholz, F., Bertrand, A. (2022) Smaller fish species in a warm and oxygen-poor Humboldt Current system. *Science* 375(6576): 101–104, <https://doi.org/10.1126/science.abj0270>.

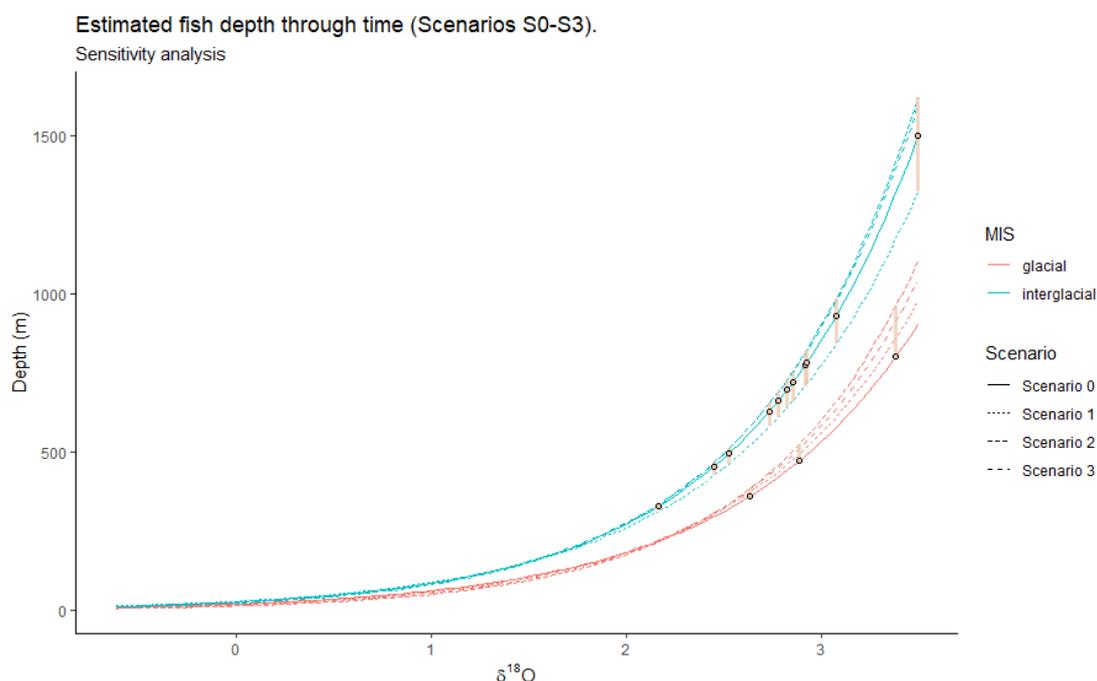
Lin, C.-H., Wei, C.-L., Ho, S.L., Lo, L. (2023) Ocean temperature drove changes in the mesopelagic fish community at the edge of the Pacific Warm Pool over the past 460,000 years. *Science Advances* 9(27), <https://doi.org/10.1126/sciadv.adf0656>.

Britten, G.L., Sibert, E.C. (2020) Enhanced fish production during a period of extreme global warmth. *Nature Communications* 11: 5636, <https://doi.org/10.1038/s41467-020-19462-w>.

Your lifetime-average depth model is plausible, but assumptions are rather strong and should be stress-tested. The lifetime-average depth reconstruction from otolith  $\delta^{18}\text{O}$  is an interesting approach, but it relies on strong assumptions that should be made transparent and, ideally,

tested. My concerns are that the model depends on the assumed  $\delta^{18}\text{O}$ –depth relationship and the assigned ecological depth ranges of the foraminiferal calibrators used to fit the exponential curve. These assumptions are especially consequential in the Eastern Mediterranean, where stratification and water-mass structure can vary substantially over time. Maybe adding a sensitivity analysis showing how inferred fish depths change under reasonable alternative parameterizations would help.

Thank you for this suggestion. We have now included a sensitivity analysis using narrower and wider depth ranges as well as alternative sampling distributions (uniform and Gaussian). While absolute depth estimates vary slightly, they remain within the confidence intervals and the overall pattern and interpretation remain the same. The new version of the code including the sensitivity analysis is available at <https://doi.org/10.5281/zenodo.18835015>



The otolith  $\delta^{13}\text{C}$  interpretation is actually confounded by size/weight effects already observed. The result reports a significant correlation between otolith weight/size and otolith  $\delta^{13}\text{C}$  for both taxa. This is critical; it complicates the interpretation of  $\delta^{13}\text{C}$  shifts as primarily metabolic or habitat-driven signals across MIS intervals. We need a control for size/weight statistically (e.g., model  $\delta^{13}\text{C}$  with size as a covariate; use residuals; ANCOVA-style comparisons), or demonstrate that key temporal patterns persist in size-matched subsets across intervals. Otherwise,  $\delta^{13}\text{C}$ -based conclusions should be framed more cautiously. i.e., “consistent with” rather than diagnostic.

Indeed, we report in the Results section already that “there is a significant correlation between otolith weight and otolith  $\delta^{13}\text{C}$  for both *Hygophum benoiti* 395 ( $\rho = 0.45$ ,  $p = 0.04$ ) and *Ceratoscopelus maderensis* ( $\rho = 0.80$ ,  $p = 0.0003$ ).” In the Discussion, we do not report on any temporal patterns apart from a small decline in  $\delta^{13}\text{C}_{\text{benoiti}}$  in MIS 18, which we do not interpret at all, precisely because the signal is confounded by the size-isotopic ratio relationship and because we have too few points per species to analyze this statistically. Instead, we discuss this relationship as a whole for the entire dataset: “overall, the larger otolith specimens exhibit higher  $\delta^{13}\text{C}$  values, suggesting that, as the animals grew, they

consumed fewer and/or switched to lower-trophic-level items. A possible explanation could be that younger individuals consumed much, primarily detritus, whereas adults became more selective, consuming mesozooplankton.”

That said, we understand that the sentence “no important changes in the otolith  $\delta^{13}\text{C}$  values of the two fish species are observed over time” may have caused this misunderstanding, so we propose to revise that to “no important changes in the otolith  $\delta^{13}\text{C}$  values of the two fish species are observed since the otolith  $\delta^{13}\text{C}$  signal is confounded by its correlation with fish body size”.

Line 42. The statement that *Hygophum benoiti* and *Ceratoscopelus maderensis* are “two of the most common mesopelagic fishes worldwide” is definitely not correct. These taxa may be very common in the Mediterranean sediment otolith record (e.g., Lin et al. 2017, 2018; and related work on Mediterranean distributions), but global mesopelagic dominance patterns differ substantially by region (e.g., Schwarzhans 2013: Gulf of Guinea/Azores case study).

They are species with broad distribution that are among the most common in the North Atlantic and the Mediterranean. Of course other species are dominant in other regions. Of course we will rephrase.