

# Final Author comment for RC2

This paper investigates an interesting approach to monitoring the AMOC that applies machine learning to derive information from Argo float profiles, and other data. I am not an expert in machine learning and my review is from the perspective of an oceanographer.

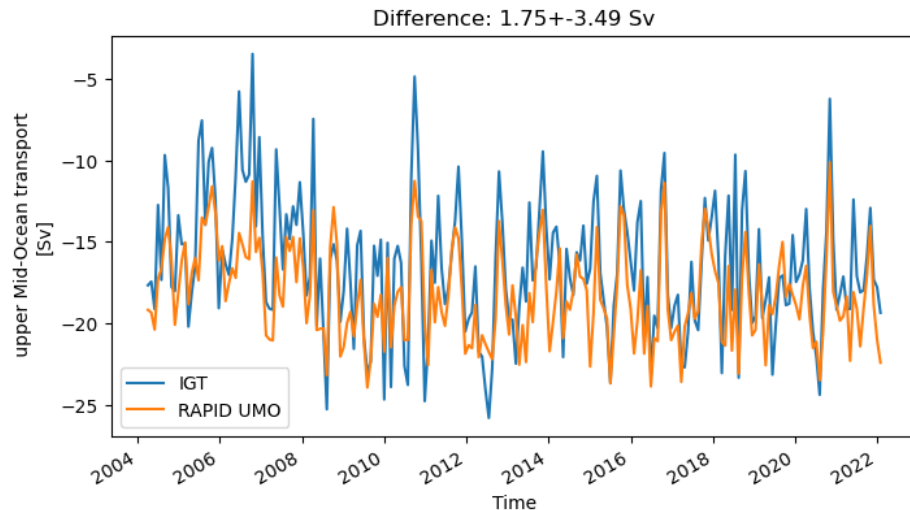
The authors find that in the model the machine learning technique can make accurate estimates of the AMOC, however, the amount of training data is much greater than is currently available from real observations. Thus, the only prospect for applying the method to estimate the real AMOC would be to train the method on model data. In the discussion, the authors suggest that models are not sufficiently realistic for this to be done now, though this is not analysed. The work is novel and interesting but the presentation is not always easily understandable and some of the analysis seems to confuse different questions. I recommend a major revision

There are two parts to the paper that I think need to be improved.

1). Section 4.3. This should be the most important part of the paper as it focuses on the contribution to the AMOC that depends upon the mooring measurements. However, I found this section confusing and the main variable under consideration was not clearly defined

a) The "RAPID like AMOC", sometimes also referred to in the manuscript as the "geostrophic AMOC" which is the focus of the analysis in section 4.3 is not clearly defined. On line 132 use the term "interior geostrophic transport", I think this is the most accurate description and it would be better throughout. Labelling it as "AMOC" is misleading.

- Response: We agree with the reviewer's comment that the interior geostrophic transport is not well defined in our manuscript. We replaced all occurrences of geostrophic or RAPID-like AMOC with "interior geostrophic transport". We kept the 'RAPID-like' to highlight that the transport is not a sum of gridboxes in an ocean model. Furthermore, we have added an appendix (Appendix A) which describes in detail how we used the information in McCarthy et al. (2015) to calculate the interior geostrophic transport. We ask the reviewer to share their opinion on whether this should stay in the appendix or be moved into the published code repository. We are convinced that this description will make the manuscript more comprehensible.
- We validated our approach by applying the interior geostrophic transport calculation from Appendix A to the published RAPID moorings dataset "ts\_gridded.nc" and compared the time series to the published upper mid-ocean transport from the "moc\_transports.nc" dataset. We expect the main difference between the time series to the western boundary wedge. We find that the difference has an average of 1.75Sv and a standard deviation of 3.49Sv. McCarthy et al. (2015) states "The array measures components of the Antilles Current and the Deep Western Boundary Current in combination from Abaco Island to WB2 (WB3) with a mean strength of 1 (4) Sv with a standard deviation of 3 (10) Sv.". Overall, these statistics validate our calculations and we suspect the small error is due to either non-published or from us not implemented assumptions in the RAPID calculation.



b). The AMOC is usually defined as the maximum of the overturning stream function so the text on line 251 should be "strength of the stream function at the grid box closest to 1000m". Then on line 255 "we also use an interior geostrophic transport time series".

- Response: We thank the reviewer for their feedback and have made changes to the text to improve its precision. We incorporated the suggestions of the reviewer and decided to replace total AMOC with only AMOC throughout the manuscript, based on the reviewer's comment in 1a), where we dropped the usage of the geostrophic AMOC. To overview all changes we will add the track changes version to our next manuscript upload.

c) Note too that the RAPID "upper mid-ocean time series" usually includes the western boundary wedge. Smeed et al 2018 presented only the geostrophic part east of the mooring WB2 and referred to that as "gyre recirculation".

- Response: We agree with the reviewer that the former version of our manuscript lacked distinction between the RAPID upper mid-ocean time series and our interior geostrophic transport. We changed the text accordingly. In the beginning of section 4.3 we made clear that it is only part of the umo "...can help to effectively reconstruct the interior geostrophic transport of the RAPID upper mid-ocean component..." (l. 450). When describing the interior geostrophic transport, we explicitly stated that we exclude the WBW and removed the reference to the upper mid-ocean. We changed in line 2123 "For the interior geostrophic transport, we calculate the RAPID upper mid-ocean transport and exclude the western boundary wedge, therefore this is similar to the gyre recirculation in Smeed et al. (2018). To be as similar as possible to the RAPID interior transport, we took the RAPID calculation for the RAPID moorings (McCarthy et al., 2015) and mapped this to the VIKING20X output to share as many assumptions as possible. "

d) On line '536' it is stated that "the RAPID-like geostrophic AMOC, mainly represents the southward deeper return brach of the AMOC" . I think this is incorrect, but the variable is not defined so I am not sure. Normally the southward deep transport should be equal to the AMOC

- Response: We thank the reviewer for this finding and agree that "southward deeper return branch" is the wrong phrase for what was meant. We changed this sentence in

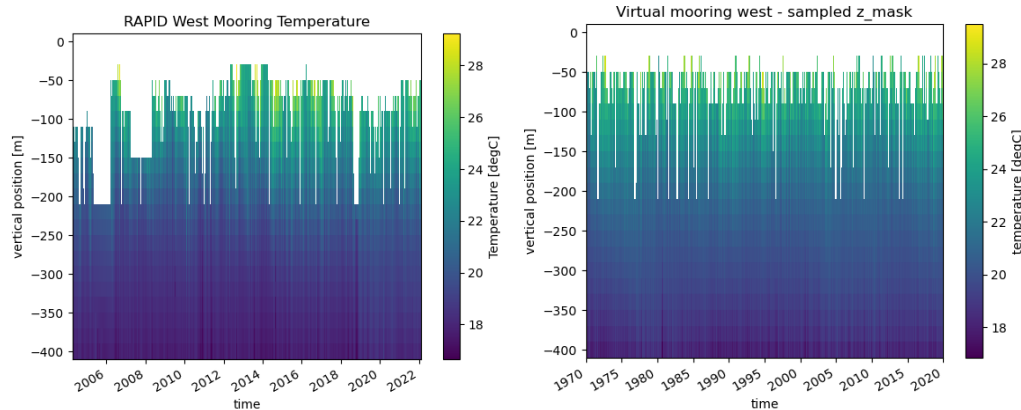
line 450 into “The contribution of RAPID like interior geostrophic transport we find is to be predominantly negative transport.”

e) When calculating geostrophic transport it is necessary to choose a reference velocity at one level. How was this done in this case? In the RAPID calculation this is done so that the total net transport is zero, so the reference velocity is also influenced by the Ekman and Florida Straits transports.

- Response: We thank the reviewer for this comment and want to connect this to our answer for point 1a). We add the information about the reference level in the appendix describing the calculation of the interior geostrophic transport. As we are not calculating the total RAPID timeseries we do not enforce a total net zero transport but rather choose reference levels. The velocity calculated from the vertical shear of northward velocity is the surface. Furthermore, we decided to take -4820 meters as level of no transport (see Appendix A Eq. A4) because as mentioned in section “The external transport:” in McCarthy et al. 2015 this is the depth of the deepest RAPID instruments. A deeper level of no transport would include the Antarctic Bottom Water transport which is not part of the RAPID array.

f) I did not understand why a reduced sampling near the surface to mimic the RAPID observations was done. Surely we want to know how well the ML reconstruction can estimate the actual geostrophic transport? How the missing data from the moorings affects the RAPID estimate is interesting but separate question. The analysis is confounding two different things. For this paper it would be better to focus only on the ML technique.

- Response: We thank the reviewer for questioning the sampling of missing data for virtual moorings. But we would like to keep this decision because our work aims at the academic question, whether machine learning could be a helpful method to incorporate real-world Argo floats in real-world basin-wide array measurement campaigns aimed at the quantification of the AMOC? The goal is to mimic the challenge an ML approach using real-world data would face as closely as possible. Hence, we use the approx. same data quality (therefore missing mooring data) and only change the amount of data (using ocean models) for the ML training process. Otherwise, we would need to estimate the error that is made by the missing mooring values, which would give this work a flavor of a model vs RAPID comparison. We would like to avoid such a direction and as the reviewer said, focus on the ML technique.
- We want to show how the missing data in the near-surface of the RAPID moorings limits the availability of measurements. The left plot show the RAPID mooring west from the downloadable data product ‘ts\_gridded.nc’ from the RAPID website. The white areas show non-available measurements. On the right hand side we show how we mimic this distribution of missing data in the virtual moorings that are extracted from an ocean simulation.



## 2) Section 4.2. "Importance of individual components for the AMOC reconstruction"

a) This section seems to be confusing two questions. The first question is what components of the circulation contribute most to AMOC variability and the second is which data is most useful for the ML reconstruction.

- Response: We thank the reviewer for this comment. We fully agree that the first question regarding the contribution of the different components to AMOC variability is not within the scope of this paper. See our response to 2.b) for details.

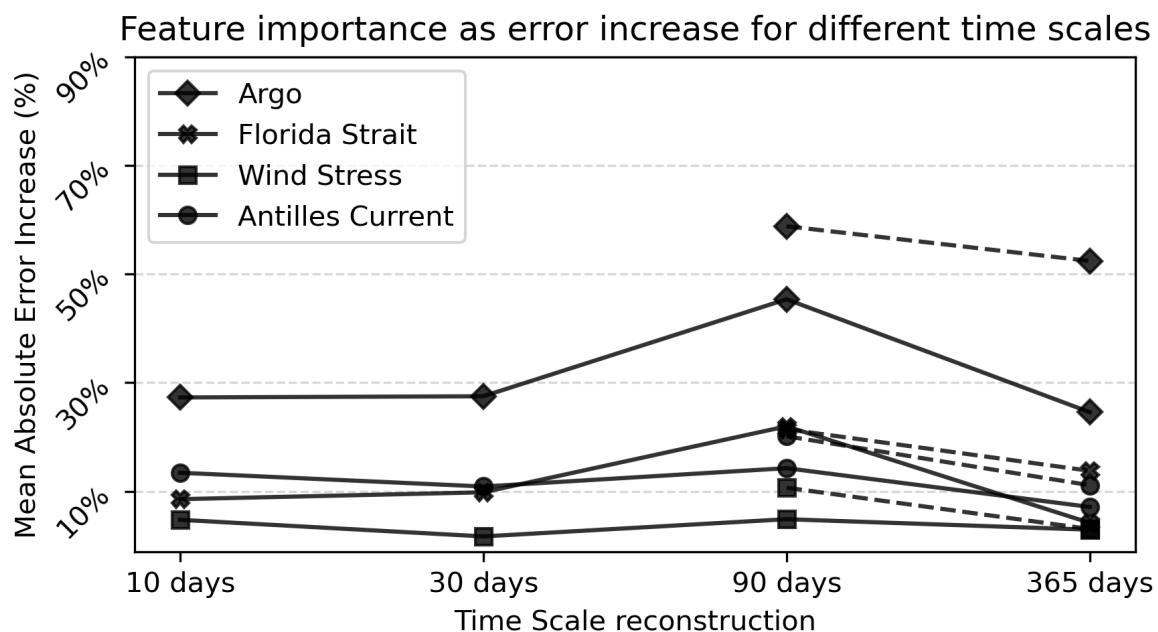
b) There are already quite a few papers that have discussed the first question. In particular Moat et al 2020 discuss how Ekman transport is important at short timescales and that at long time scales most variability is from the mid-ocean transport (see their Figure 2). So the results in Figure 7 do not seem surprising

- Response: We thank the reviewer for this comment and agree with their perspective. Our goal was to show that we are in line with the current research. The fact itself is, for sure, not surprising. We want to highlight that our data-driven reconstruction utilizes the expected correlations, shown by the previous papers, to give some validation of our method. We rewrote the text to reduce the confusion to this point, explicitly stating that the finding itself is not surprising but necessary for validation. We added the following two sentences after the second paragraph: (l. 426)"On short time scales the high importance of zonal wind stress is not surprising and agrees with current literature (Moat et al. 2020). However, a high importance of wind stress on shorter time scales partially validates the learned AMOC reconstruction, as it utilizes expected and known correlations."

c) It would be much more interesting if the authors instead examined how much different data contributed to the interior geostrophic transport. Is the surface stress or the Florida Straits transport contributing to the skill in the reconstruction of this component?

- Response: We thank the reviewer for this valuable feedback. We executed the same experiment as in section 4.2 with the interior geostrophic transport. The result figure is attached to this response. The importance of wind stress is constantly the lowest across all time scales, as expected. Most of the importance is assigned to the virtual Argo profiles for all time scales. A smaller artifact can be found in the seasonal scales when the importance of the Florida Current is larger than the Antilles Current's

importance. Our interpretation is that the Florida Current has a strong seasonal signal, which is evident in a correlation coefficient of  $-0.45$  between the seasonal highpass filtered interior geostrophic transport and the Florida Current transport (consistent with Frajka-Williams et al., 2016). Hence, from Florida Current Transport, the algorithm may infer the season which is not explicitly provided via any inputs. In our opinion, however, the manuscript would not benefit from adding these results, also considering the manuscript's length, as a figure or as an appendix. We decided to add the following sentence to the end of section 4.2: (l. 444)"When calculating the feature importance with regard to the interior geostrophic AMOC, not shown in this manuscript, we observed a similar pattern, the Florida Current gains importance for seasonal scales, due to the strong correlation to the AMOC (Frajka-Williams et al. 2016). Despite this, the importance for all other inputs is constant, for the interior geostrophic AMOC across all time scales with virtual Argo profiles being the most important."



Other comments:

I found the paper quite long (740 lines excluding figures, tables, references and the abstract) and there are many places where the text could be shortened. E.g. in the introduction "Zilberman et al. (2020) grouped Argo profiles into  $6^\circ \times 6^\circ$  cells in the Pacific to create a uniform coverage of Argo profiles which could be used for further computation" seems tangential and could be removed. Is it necessary to say (about Argo floats) "Data are transmitted through a satellite connection while the float drifts at the surface for a few hours."? Shortening the text will make the manuscript easier to read.

- Reponse: We thank the reviewer for their comment and agree that the manuscript was at some parts overextended. We removed the mentioned parts in the introduction (Zilberman, Transmission of Argo signals), also we removed the Deep Argos from the introduction and shortened the RAPID section. We sharpened the

result section especially in section 4.1 by removing what we considered non-essential information. Overall, we were able to reduce the manuscript to ~620 and believe that this improved its quality. We add a track changes version of the manuscript in next upload to explicitly show the removed parts.

line 115 I do not understand "we also use positions of the RAPID moorings for information about the deeper layers."

- Response: We changed the text to highlight our experiment (see section 4.4) about the influence of measurements deeper than the Argo profiles. The sentence and the following were meant to describe the three possible options, including no deeper measurements, including the deeper part of the RAPID moorings, or assume that all Argos would be Deep Argos ranging up to 6000m depth.

line 155 "Figure ??"

- Response: We thank the reviewer for finding this, and we fixed the reference.

line 193 please provide a citation for "graph data structure". Many readers, like me, will not be expert in the techniques of machine learning and so citations are particularly important. Similarly for "explainable AI (X-AI) techniques" on line 485

- Response: We thank the reviewer for this helpful comment. We have added appropriate citations for "graph data structure" (doi: 10.1146/annurev-ecolsys-102209-144718) and "explainable AI (X-AI) techniques" (doi: 10.1145/3561048) in the revised manuscript (l. 159).

line 223. The statement "The reconstruction uses the concatenation of the density values from the Argo profiles for the upper 2000 meters and the derivation of the meridional velocity w.r.t. the depth computed with the RAPID mooring locations as information deeper than 2000 meters". is confusing. A concatenation of density and velocity seems odd.

- Response: We thank the reviewer for this feedback and agree that we missed to describe that the input of a neural network does not necessarily share the measurement units. The learning process is based on the statistics of the input variables and therefore can process differently measured variables as shown in Figure 2 C). We made changes to text explaining briefly how measurements with different units can be processed by the same neural network and reference experiment 4.4 that investigates different data sources below 2000 meters depth.
- Line 187: "Please note that the inputs to our neural network must not have the same physical units, since the learning process of the neural networks is based on the statistics of the input variables and not on their semantics. In section 4.4, we investigate the effectiveness of deep information by first testing the removal of the additional mooring data. Next, we assume that all Argo floats would reach up to 6000 meters."

Line 294 I do not understand what is meant by "For the virtual Argo profiles, the goal is to train an embedding (black box in Figure 2 B)) that maps a set of Argo profiles into a hidden space in where similar ocean states are near each other even though their spatial distribution of observations may be different." What is "an embedding"? I think "in where" should be "in which"

- Response: We understand that the term "embedding" is not a familiar oceanography term and is also not well explained in the manuscript. We provide a thorough explanation of the term embedding with a better reference to a manifold learning survey article (Meilă and Zhang,2024). Therefore we restructured section 3.1 such that it follow the schematic from Figure 2. First we describe why Argo profiles with their variable amount and unstructuredness are a challenge to Neural Networks which expect fixed-sized inputs. Next, we show the potential of an embedding and connect the term to mathematical vector spaces that are connected to the ocean manifold. We think this improved the section and gives a better introduction to the embedding term as it is used in machine learning.
- In this answer letter, we want to give a more technical explanation of what we understand as embedding and which task we want it to fulfill. Given a set of Argo profiles with different locations and times the goal of the processing is to identify a structured numerical representation of the underlying ocean state which is independent from the number of profiles. We choose to represent the ocean state by an **embedding** which maps said set of Argo profiles onto a space of n-dimensional floating-point vectors. Often, the term embedding is used for the n-dimensional vector, but also for the vector space, and for the mapping. Constructing a suitable embedding has two main aspects. First, the same ocean state observed by different spatial and temporal distributions of Argo profiles should result in the same (i.e. negligible Euclidean distance) embeddings. Second, two embeddings should be similar (i.e. small Euclidean distance) if the corresponding ocean states are similar, too. A successful embedding creates a vector representation of the manifold of ocean states that can be used to generalize for unseen but similar ocean states in the evaluation. However, the semantics of these intermediate vectors are not trivial as the exact mapping from the manifold of ocean states to the vector representation is a data-drivenly learned and non-linear function. Without any outside constraints (see black box in Figure 2B) the content of the vectors is not interpretable. The discrepancy between the degree of freedom of the ocean state manifold and the usually smaller dimensionality of the embedding, which limits the information content, can be understood as a low-pass-filter in the training process, enhancing the generalization of the mapping function to focus the mapping function on the largest variabilities in the manifold.
- We used parts of the upper bullet point in the manuscript in line 249 as continuation of the original second paragraph elaborating on the need for such a fixed-sized representation: "A classical binning comes with many assumptions about shape, distance, and connectivity of the bins. This requires a thorough testing of hyperparameters for classical binning approaches, which may be in the end only valid in the study area. However, we aim in this work for a data-driven mapping function which identifies a structured numerical representation (n-dimensional vector) of the underlying ocean state which is independent from the coverage of Argo profiles. Often, the term embedding is used for such a n-dimensional vector, but also for the vector space, and for the mapping (Meilă and Zhang,2024). Constructing a suitable embedding (see black box in Figure 2B) has two main aspects. First, the same ocean state observed by different spatial and temporal distributions of Argo profiles should result in the same (i.e. negligible Euclidean distance) embeddings. Second, two embeddings should be similar (i.e. small Euclidean distance) if the corresponding ocean states are similar, too. A successful embedding creates a

vector representation of the manifold of ocean states that can be used to generalize for unseen but similar ocean states in the evaluation.”

Table 1 In the last line I suppose "WS" should be "ZW"?

- Response: Thanks for finding this mistake. It was a relict from an older version, we changed from "WS" to "ZW".

Line 354 is the naming of "test, validation, and training periods" standard? Often "test" and "validation" have similar meaning.

- Response: The naming is the standard in the field of machine learning. We understand that we missed an explanation and added a clarifying sentence. In short, the validation set is excluded from the training and is used to verify the choice of hyperparameters, while the test data is only used for the final evaluation.