Response to Reviewers for Paper: "Dynamical reconstruction of the upper-ocean state in the Central Arctic during the winter period of the MOSAiC Expedition."

Reviewer 2: *A review of "Dynamical reconstruction of the upper-ocean state in the Central Arctic during the winter period of the MOSAiC Expedition" by Kuznetsov and co-authors. My background in is modeling and data assimilation of the Arctic Ocean, but not so much sub-mesoscale oceanography.*

The manuscript exploits a very dense measurement campaign from the ambitious MOSAiC ice camp somewhere in the central Arctic and assimilates it into a numerical model of very high resolution. As there is no other realistic ocean model of similar resolution set up in the central Arctic to my knowledge, the study stands out by its very high originality.

The closest relatives of such studies were the pioneering ocean forecasts by the Harvard Group in the early 90's where cruise data were assimilated by Optimal Interpolation and fed into a forecast model running onboard during the cruise (Robinson et al. 1996), which did demonstrate forecasting skills. The MOSAiC was however on a slower path at the speed of the sea ice drift and was not in control of the trajectory, so the ambitions stay realistically on a lower level, that of performing an oceanographic process study, which is interesting in its own right.

The authors did interpolate the temperature and salinity profiles obtained along the whole 4-months experimental period into 3D fields and assimilated these data into a bespoke model set up specifically for the area of the experiment, and then discuss the dynamical features of the simulated ocean fields.

Response: Thank you very much to the reviewer for such detailed and constructive comments. We want to avoid misunderstandings from the outset and clarify the method we used. The observed data were indeed extended onto the model grid to further fit the model to this data (lines 217-219 of the original manuscript). However, it occurred only at grid nodes no more than 1 km from the observation point. The number of grid nodes constrained the influence; with a horizontal resolution of 1 km, the observational data impacted no more than 3 model nodes. Conversely, with a resolution of 0.25 km, the data's effect was confined to the 30 nearest nodes and spanned less than 1 km. In other words, with a distance of 20 km between the buoys, the zone not covered by the data spanned 19 km. Moreover, there was no vertical interpolation as such since observational data with a resolution of 1 m were used for data from autonomous profiles. For other data, from the CTD chain buoys, the data from the observation horizons were expanded vertically by 3 m. In both cases, the expansion of the zone of influence of the data on the model (1 km horizontally and 3 m vertically) was implemented to avoid the formation of sharp horizontal fronts when crossing buoy trajectories and to prevent strong vertical instabilities when crossing trajectories of different types of instruments, which could lead to an unstable model solution. This approach is not about interpolating data and then using these 3D interpolated fields to nudge the model, as the nudging was performed in relatively limited spatial areas.

Reviewer 2: *A major assumption of the study is therefore the synoptic* *nature of the measurements assimilated ("quasi-steady-state" is not mentioned until line 203 in the methods description, which is very late for an important assumption). Further on, the model results reveal that this assumption was actually wrong, as acknowledged in the conclusion. This is a major weakness of the study that should be recognised from the start to avoid some unnecessary suspense.*

Response: Application of the quasi-steady-state assumption is now included in the abstract in the revised manuscript. Additionally, more justification for the application is added to the methods section.

The absence of full 3D interpolation (interpolation over the entire space between the data, as mentioned above) provided us with the opportunity to apply the quasi-steady-state assumption. This is precisely because the distance between the buoys and observation points is greater than the characteristic size of submesoscale eddies (7 km) in this area. Specifically, because there is no horizontal interpolation between the data, we use this approximation, assuming that sub-mesoscale processes do not influence each other on mesoscale scales (50 km). The example we provide of the interaction of eddies clearly demonstrates that using any horizontal interpolation of temporally separated data can lead to challenges in identifying submesoscale processes. It should also be noted that the dynamics shown in the latter part of the article, which the reviewer refers to, were simulated in a free run when the data did not constrain the model, and the quasi-steady-state assumption was not applied.

Reviewer 2 *Another major weakness is the odd choice of the interpolation method. The inverse distance method is not able to de-cluster observations in an irregular network such as the ones obtained here during a nearly random drift, and generates very erratic extrapolation features, so it should not be used other than with regularly spaced measurements. It is possible that the inverse distance worked well when combined with a clustering method briefly referred to, but it is a priori an ill-informed choice (See Zimmermann et al. 1999 for a thorough comparison, as well as modern examples of interpolation with Bourgain and Gascard 2011 and Troupin et al. 2012).*

Response: As we have already noted, there was no 3D interpolation for all nodes of the model. We used the method described in the technique solely to smooth the effect of high-resolution data and to smooth data from different sources; this is more like extrapolation.

Reviewer 2: *I will argue later that the study may have been improved by extracting only the large scales variations from the data and perturbing randomly the mesoscales.*

Response: The experiment with random perturbations is interesting as a sensitivity analysis of the model. Similar experiments were conducted in the earlier stages of developing mathematical modeling methods, including the FESOM model family. It should also be noted that, due to the absence of continuous interpolation across the entire area, we are implementing in this paper considering only that the perturbations are not random but are based on data.

Reviewer 2: *The resulting maps of interpolated values, the only energy* *source of the simulation are not shown, which casts the shadow of a doubt on the realism of all the results obtained throughout the paper. Are we looking at observed mesoscale features or quasi-random (sampling-dependent) perturbations of a homogeneous density field?*

Response: Figures 5, 6, and 8 show the resulting maps and transects, reconstructed using the model. We do not have interpolation maps as noted above. We have a 'spaghetti plot' from the data with an increased nudging radius of up to 1 km. An example can be seen in the article https://doi.org/10.5194/essd-14-4901-2022, Figure 7a.

Reviewer 2: *The validation against independent observations is unconvincing because these measurements are taken in the vicinity of the ice camp and are not representative of the remote unobserved areas. The authors correctly recognise that the validation is poor at the cross-over points but still boast uncritically the success of the validation in several places. This is not a major problem since the model is used for process studies which do not require any accuracy but the text gives a misleading impression of accuracy.*

Response: Independent observational data (near daily temperature and salinity profiles from a microstructure profiler, MSS) were obtained 20 km from the buoys, based on which we adjusted the model. Indeed, they are located at the center of the pattern, but, as previously noted, we do not employ any data interpolation at such scales. The model is not simply superimposed onto the fields. The data from the ship's CTD (located in the vicinity of MSS) used for nudging were primarily collected on different days from when the MSS measurements occurred. Moreover, CTD profiles measured from the ship were conducted once a week, while microstructure profiles were taken almost every day. The data is independent: it is distinct from the rest and was not collected concurrently with the ship's measurements.

Reviewer 2: *The authors do not use any numerical model reanalysis nor climatology as background values, which is probably for the best to avoid additional artefacts.*

There are other aspects of the experimental setup that are should be clearly explained upfront in the paper rather than admitted too late in the discussion section. One is that the ocean is completely shielded from the atmosphere by an idealised ice cover, so that the only source of momentum in the model is the nudging to temperature and salinity.

Response: We discuss ice in the model description lines 125-129. The influence of ice in the model occurs through the parameterization of friction in the surface layer and the parameterization of turbulent exchange coefficients in the closure equation. The suggestion that the ice drift does not serve as a source of momentum due to ice dynamics is added to the model description in the revised manuscript. The idealized ice cover is now mentioned in the abstract.

Reviewer 2: *Another one is the breach of the continuity equation by the nudging, which contradicts the assertion that the assimilation is physically consistent.*

Response: In the methods section, we've included a note stating that nudging violates the principle of continuity. It's essential to emphasize, however, that our application of data nudging is confined to specific observational points, rather than uniformly across the entire region. This targeted approach avoids major complications while establishing initial conditions for a free simulation phase.

Reviewer 2: *The data assimilation method itself is admittedly very rudimentary (nudging), but contains unexpected complications that are not justified at all: using different relaxation times for temperature and salinity and the odd-looking vertical relaxation coefficient in Eq (4). If these complications were necessary then the authors should explain what led to them.*

Response: 1. We use the same coefficients for temperature and salinity. Perhaps the reviewer referring to the Ck coefficient ? However, the index 'k' pertains exclusively to the type of instrument, as detailed in the text, and not to a differentiation between temperature and salinity. 2. The 'odd-looking' vertical relaxation coefficient is applied to the first group of data obtained from buoys that record temperature/salinity at five levels. This is to ensure that the model (in the absence of 3D interpolated fields, and with only point data available) does not experience a shock from nudging at a single point. We employed the method described in the article to smooth the effect of nudging. The text added to the article states: Thus, the model's nudging occurs in the vicinity of the observation point +-3 meters, but the strength of the nudging decreases with distance from the observation point.

Reviewer 2: *The paper writing is overall quite poor, even though the English is good, the explanations and justifications are often vague and the logic is not obvious. This is particularly true of the introduction, which reads as a long enumeration of unconnected facts. So the paper needs a thorough revision of the text to remove all the loose ends and strengthen the logic.*

Response: We have revised the introduction to provide a better flow, leading up to the objectives of the manuscript. Selected citations have been added or removed.

Before the paper is acceptable for publication, the authors should provide visual evidence that the interpolated fields obtained by the inverse distance method are making sense as a quasi-steady-state estimate of the water masses or if any random perturbation of the homogeneous initial fields would have led to the same conclusions.

Response: As noted in the manuscript on lines 217-218, the nudging term was incorporated solely for grid nodes located in close proximity to observation points. Thus, we do not have interpolated fields obtained by the inverse distance method; we only have scattered data + interpolated/extrapolated values within a radius of 1 km from a specific measurement. The effect of these for nudging was expanded by forming a matrix indicating the presence or absence of data within a 1 km radius of observation points. However, data from a neighbouring buoy located 20 km away does not affect nudging in the vicinity of 1 km from another buoy. The exception is moments when the trajectories of the buoys intersect in space but not in time. In such cases, the method used in the vicinity of 1 km from the observations is practically irrelevant; the main task here is exclusively to smooth out potential pseudo-fronts caused, for example, by different phases of the internal wave. Random perturbations of the observed data will not lead us to the same results.

Reviewer 2: *The introduction should be completely re-written to prepare the reader for the experiments at hand and formulate more precise goals than to "extend current knowledge of submesoscale dynamics". The conclusions are just as vague: they are mostly reflecting a posteriori on the limitations of the experiments rather than highlight the newly gained insights related to the vertical EKE profiles.*

The paper has important scientific merits in spite of the abundant flow of criticism coming below, so I believe that it should appear after major revisions: new experiments would be an improvement but are not compulsory. However there should be a restructuring of the text, better explanations and a new figure showing horizontal interpolated Temperature and Salinity fields.

Response: The manuscript is rewritten in accordance with the reviewer's suggestions.

Reviewer 2: *Detailed comments: The abstract does not work as an abstract because it lacks most of the basic elements of context (What? Where? When? How?). On the contrary the five first lines do not belong in an abstract, but more in the introduction, and can be safely removed.*

L8: The model is a major element of the study. The reader needs to know what kind of model is used: its nature (ocean general circulation without active sea ice), its name, the mesoscale-resolving resolution.

The time period of the study is missing, at least the season would be useful to know.

L12: Indications like East and West make no sense unless you mention the name of the area: the Nansen Basin, Amundsen Basin?

L12: 'high variability' is also blue sky to the reader. High with respect to what?

L16 'the fields can be used for further analysis'. That statement is very vague and should be made more specific once we have an impression of the degree of realism of the interpolated data.

Response: The abstract is now rewritten, taking into account the reviewer's comments.

Reviewer 2: *The introduction is an accumulation of facts taken from the literature. Although all of them are interesting in their own right, they cover a too broad scope to frame sufficiently the scientific context of the present study. They also read like an itemised notes from a literature review with no indication whether the findings will be revised by this study or not, and most of them are not. The logical succession of these facts is also left to the imagination of the reader.*

- L.50: Typically "An analysis of the dynamics of baroclinic vortices [...] is given in Sokolovskiy and Verron (2013)" does not tell whether this analysis is in any way related to the present paper. If the discussion does not loop back to it, then please remove it from the introduction.

- L. 68: "Very high horizontal resolution" is too vague. Are they eddy-resolving, permitting, or event in the non-hydrostatic assumption?

- L. 85: there are more than one interpolation technique. Since this part is criticising interpolation techniques, it is the adequate place to mention the one that will be used in this paper.

- L89 to 97 the whole paragraph is a very cumbersome justification for using rudimentary rather than advanced data assimilation. If we trust your argument as it stands, there is no advantage to advanced data assimilation methods at all (nudging is more practical and yields better results) and nobody should ever be using anything else than nudging. Obviously you do not need to upset the whole data assimilation community to justify your choice of method. It is sufficient to state that 1) the costs and the complexity are not affordable in your case, plus 2) that the data coverage by a single quasi-random track is very unusual, so you lack evidence that advanced data assimilation is cost-effective in your case. Please rewrite the paragraph to better justify the choice of nudging.

- L98: The goal of the study "extend current knowledge of submesoscale dynamics" is too vague. It is impossible to verify whether this goal has been attained or not. Please make it more precise.

Response: The introduction is now rewritten, taking into account the reviewer's comments.

Reviewer 2: *- L115: Indicate already here the vertical coordinate of the model is sigma rather than in Section 2.3.*

- L125: There is no thermodynamical effect of sea ice on the ocean, the next section will indicate that the ice drift is a constant value. A missing piece of information here is the sea ice area coverage, which seems to be 100% thus sheltering completely the ocean from the atmosphere. It should be made clear that there is no direct effect of the atmosphere on the ocean and that, after mentioning the constant lateral boundary conditions, there is no input of momentum to the model apart from the nudging term. As recognised somewhere far down in the manuscript.

Response: In the model description, specifically lines 125-129, we address the topic of ice. It is clarified that in our model, ice drift does not act as a source of momentum; instead, its impact is limited to friction and serving as an upper boundary in turbulence closure. This clarification is now included in the model description. The influence of ice in the model occurs through the parameterization of friction in the surface layer and the parameterization of turbulent exchange coefficients in the closure equation.

Reviewer 2: *- L139: Is the value of 0.7 m/s set for the whole period and the whole model domain? Please explain why you have not made it more realistic.*

Response: This parameter is used across the entire domain and throughout the entire period of the model's nudging. This value describes the upper limit in our parameterization of vertical turbulence and sufficiently well represents turbulent mixing for our task. Since we use a 'quasi-steady-state' approximation, this parameter remains unchanged. More sophisticated parameterizations of ice dynamics and thermodynamics might lead to a more realistic description of the turbulent layer beneath the ice. However, we lack information about the ice roughness at the water-ice interface, realistic small-scale wind conditions, snow distribution, and many other factors that determine the upper boundary.

Reviewer 2: - *L141: Why do you choose this definition of the mixed layer depth and why a minimum of 20 meters?*

Response: Because according to our observations, we did not see a mixed layer thinner than 20 meters, the specific number is not important for the task definition, but it simplifies and speeds up the model. Delving into the numerical methods of the specific task goes beyond the scope of this article. This is one of the commonly used definitions of MLD, which does not play a crucial role for the specific task.

Reviewer 2: - *L159: Please indicate here the nature of the model boundary conditions. Not later.*

Response: The nature of the boundary conditions is described above; they are solid boundaries. If what is meant are the boundaries between zones where there are mostly data and where there are none (line 380), then this is the result of our experiments and is not related to the model setup.

Reviewer 2: - *L165: A boarder situation map with some topographic features would help understanding where we are. And where are the North and the East.*

Response: A map of the experiment region is now added to Figure 1.

Reviewer 2: - *Figure 1a) is too small to discern all the details. I cannot see the cyan rectangle, maybe because I am colour-blind, but I suspect there is too much information on this sub-plot.*

Response: The cyan color is now changed.

Reviewer 2: - *Figure 1b) shows a wide spread of T/S profiles, but only one density profile, which leaves us to imagine what the spread entails in terms of density changes. Can you include the spread of density still keeping the clarity of the plot.*

Response: Changes have been made to Figure 1, including the map of the region, spreads of density and n2.

Reviewer 2: - *L190: The duration of the experiment, 4 months, should have been mentioned earlier in the abstract and the introduction.*

Response: Time frames have been added to the abstract.

Reviewer 2: - *L195: The "ambivalence" is only a redundancy from the point of view of interpolation, but you could have exploited these crossing points as temporal information to calculate the errors related to the "quasi-stationary" assumption.*

Response: A good suggestion that can be implemented in subsequent publications on these data and model. It is included in the relevant section of the article.

Reviewer 2: - *L198: the "quasi-steady-state" assumption is only mentioned in the "Nudging" section, when you cannot avoid it any longer, although it has been implicitly a major assumption since the beginning of the paper. Please formulate it upfront in the introduction and reflect on its implications for the study.*
Response: "quasi-steady-state" assumption is introduced in introduction and abstract now.

Reviewer 2: - *L200: high drift speed compared to the water velocity. The drift speed has been set to 0.7 m/s above, the velocities of 1cm/s are only mentioned in the discussion section.*

Response: Thank you for pointing this out, the average drift speed during the described period was about 12 cm/s. However, a value of 0.7 was used in the parameterization of the upper boundary in the turbulent closure. Here 0.7 refers to the upper boundary condition for the dynamic wind speed in the turbulent energy budget equation. It is not the ice drift speed as it might have seemed from the article's text. This has been corrected in the new version.

Reviewer 2: - *L206: I can understand that submesoscale features of size 10 km located hundreds of kilometres apart are independent, but the mixed layer depths may change a lot within 4 months, please kill the suspense and indicate that this will be discussed later.*

Response: Line L206 is be modified.

Reviewer 2: - *L210: The nudging term acts on temperature and salinity but the model currents are only corrected progressively through geostrophic adjustment, which makes the model inconsistent during the adjustment time (this is - by the way - an aspect better handled by advanced data assimilation than nudging), what is the typical timescale of this adjustment in your case?*

Response: It takes four model months to reach a stable solution with the specified nudging coefficients. This was explained in the experiment description section and detailed in the figure with schematic of conducted simulations.

Reviewer 2: - *L215: Why use two different relaxation time scales for temperature and salinity? What does that mean for the dynamical adjustment of the model to density changes? Can you at least indicate the values of the two time scales (Trelax comes later, but I cannot locate Srelax in the text)*

Response: The relaxation time scales for temperature and salinity are the same, as mentioned in the above comment. We alter the temperature and salinity, thereby changing the density, which leads to changes in baroclinic pressure and, consequently, alterations in the dynamic characteristics of the system. Due to the gradual change in temperature and salinity, the system gradually adapts its dynamics. Trelax - Temporal relaxation coefficient (line 239). it is the same for temperature and salinity. We do not have any Srelax.

Reviewer 2: - *L221: Moving at 0.7 m/s, 2 minutes correspond to 80 meters (is this what you meant with "horizontal resolution"?) and are often within the same model mesh cell.*

Response: The actual ice drift changes over time; 80 meters is the distance between two measurements. Indeed, multiple measurements can be located within a single model cell, which is one of the reasons why we smooth high-resolution data onto a model node using the inverse distance weighting method.

Reviewer 2: - *L228: The mathematics of spatial interpolation have progressed significantly since 1968. Maybe the inverse distance method combined with the kd-tree does perform well, but the choice is not justified here.*

Response: Since our goals and objectives do not include the creation of 3D interpolated fields, and since our primary concern is the smoothness of data at the points where buoy trajectories intersect, as well as to slightly extend the nudging area around a single node again for the smoothness of the nudging impact on the model, the specific method of distributing observations across model grid nodes is not important to us, and considering that this combination of methods is quite common in data processing for unstructured grid modeling, the use of this method seems sufficient to us, and moreover, a detailed discussion of interpolation methods is clearly out of the scope of this work.

Reviewer 2: - *L230: The sharp transition between the cells that do and do not participate in the nudging should be mentioned here rather than in the end of the article.*

Response: Grid cells in the model do not participate in nudging, only the nodes where passive tracers such as temperature and salinity are calculated. The fact that only a part of the nodes are used for nudging is mentioned on lines 217-219: "The term responsible for nudging was included only for grid nodes in the immediate vicinity of observations. To do this, a mask of nodes has been precalculated for each type of observation and is explained below." Areas with and without data are indicated in Figure 1a.

Reviewer 2: - *Eq (4) looks like an inverse square distance interpolation in the vertical dimension but goes to zero in the separations between observed levels at depths. This seems excessively complex in the circumstances. Not relaxing between two observed levels seems prone to inconsistencies (unstable density profiles between two observed levels), a more intuitive solution would have been to perform vertical interpolation of the SIT profiles to the model levels (linear or cubic splines), ensuring the density increases with depths, and then relax with a single coefficient. The adequacy of the vertical interpolation should be better justified.*

Response: We could have managed without extending the nudging area vertically and used nudging in the model only on 5 layers, but this would have required significantly reducing the computational time step to overcome sharp changes. And yes, we do not nudge to artificially interpolated data where it doesn't exist, and that is precisely why the values are zero between layers where data is present. Even if this leads to minor instabilities in the initial moments of time, the model, due to its own dynamics, mixes out this instability. Considering that such changes are local and limited to the area where nudging is performed (in a very limited number of nodes and vertical levels), it is not critical for the model's instability.

What the reviewer suggests could be tried in future work, but this would first lead to much more dangerous horizontal instabilities, when somehow interpolated data from 5 levels to a full profile would create significant baroclinic horizontal gradients leading to artificial giant velocities. Moreover, such vertical interpolation should assume the presence of background information about possible vertical characteristics of the water column, and in the case of a eddy, such smoothing simply negates any sense in further using the model, as all heterogeneities in the form of vortices will disappear because the model will be forcibly nudged to a fictional profile.

Reviewer 2: - *Eq(4) Is the surface temperature relaxed to the freezing point temperature or is that already handled by the FESOM model?*

Response: Nudging in the FESOM-C model occurs only in places where there are data, as described in the article. In our model, there is no nudging to the freezing temperature or restoration of surface salinity, as in global models.

Reviewer 2: - *L238: Note here that a relaxation time of one day is considered very strong relaxation in practice.*

Response: The relaxation time depends on the model's tasks and application area. If we are considering a global task with coarse resolution, then one day might be a large value, but for tasks with high resolution in the coastal zone, modeling tides or internal waves, the relaxation time will be significantly less. What practice is the reviewer referring to, and why is one day considered very strong? For this particular task, one day was a successful choice.

Reviewer 2: - *L240: I imagine that the maximum distance changes together with the maximum number of values but please specify explicitly. Also mention the size of the largest and smallest neighbourhood tested.*

Response: The maximum distance is defined as 1 kilometer, see line 228 of the article.

Reviewer 2: - *L244: The deeper profiles are nudged over shorter distances than the shallow SIT profiles, making their effect probably negligible. This is counter-intuitive since the length scales are longer at depths. Please explain.*

Response: SIT buoys do not provide any profiles! SIT buoys have only 5 horizons, with the maximum depth being 100 meters. See the detailed description of the instruments on lines 174 - 179 of the original article, where the maximum variability is observed. Considering that SIT buoys provide data every 2 minutes, and instruments that provide incomplete vertical profiles once a day, and complete profiles at best once a week, the effect of profiles on the nudging of surface layers is significantly less, simply due to their insignificant quantity. At the same time, zones deeper than 100 meters are determined exclusively by profiles.

Reviewer 2: - *L247: No reason is given why the OC/PS profiles vertical relaxation is also different from the SIT profiles. Is it because these profiles have higher vertical resolution than the model?*

Response: SIT buoys do not provide any profiles! For two different types of data (profiles (PS/OC CTD) and individual horizons (SIT buoys)), we have two slightly different methods of nudging, one for profiles and one for data from individual horizons.

Reviewer 2: - *L247: The model is nudged towards invariant temperature and salinity fields interpolated from the SIT profiles. These interpolated maps being the only external forcing of the model, they should be shown at a representative depth, for example the salinity above the halocline (20m or 50m) and the temperature at 100 m.*

Response: Refer to the above response about interpolating data across the entire space, and that nudging occurs only within a distance of up to a kilometer from the data. In spaces between buoys, there is a lack of data and interpolation, so there are no two-dimensional maps as such. The only thing that can be shown is scattered plots of data, examples of which are shown in the data links.

Reviewer 2: *- L249-250: These technical details can be removed. - L256: The "small time resolution" of the instruments should probably be called "low frequency sampling".*

Response: changed in revised manuscript.

Reviewer 2: *- L260: The PS-CTD profile in Figure 1 is not a typical profile as it is both warmer and more saline than all the other profiles, so it is not obvious if this profile is overall more stratified or more unstable than the profiles that will be assimilated later. If you plot all density profiles as thin lines, you can give an indication if the assimilation will have a stabilising or destabilising effect overall.*

Response: Thin profiles refer to profiles from the MSS instrument. Since most of MSS measurements are concentrated in the northern part of the drift, they stand out to one side of the profile that was used for the initial conditions. Figure 1 was changed.

Reviewer 2: *- L265: It is very well that the authors admit the breach of the continuity principle, but this should have been admitted earlier in the methods description. More advanced multivariate data assimilation methods may mitigate that problem, which could be worth noting in the discussions. Alternatively, the authors could have calculated geostrophic current velocity increments from the density gradients caused by the nudging term, by analogy with the Cooper and Haines (1996) method.*

Response: We are unable to calculate geostrophic current velocity using the nudging term, as continuous density fields from observations are not available to us. Additionally, we lack interpolated fields derived from the data (as mentioned earlier). However, it's important to note that our use of Data Nudging is restricted to specific observation locations, rather than being applied across the entire area. This localized application prevents any significant issues during the phase of determining initial conditions for a free simulation. The corresponding changes have been made to the manuscript.

Reviewer 2: *- Figure 4 sections show discontinuities in the vertical salinity profiles, are these real or are there only a few bands in the (very small) colour scale?*

Response: It's unclear which discontinuities are being referred to. If the discussion pertains to Figure 4a, then the data are linearly interpolated between profiles from observations, as noted in the figure description.

Reviewer 2: *- L286: This sentence is a very contorted way to acknowledge the temporal evolution of the ocean variables. Again, it is regrettable that the cross-over differences have not been exploited as mentioned earlier.*

Response: The sentence is rewritten in the revised manuscript. In such cases, the model points are aligned with at least two separate observations of the same variable at the same location, highlighting the limitations of the quasi-stationary approximation assumption. Typically, the model strives to replicate the smoothed values derived from these overlapping observations.

Reviewer 2: *- L297: "Same order variability at the model grid scale" this variability is not visible in Figure 6, has it been smoothed?*

Response: Thank you, that was a poor formulation. Removed.

Reviewer 2: *- Section 3.2 T/S reconstruction contains lengthy descriptions. Please reconsider if they can be shortened. - L300-301: Unclear sentence about the slope being "deeper than 40 m". Please rephrase.*

Response: Rephrased: Both sections reveal an increase in the Mixed Layer (ML) depth from northwest to southeast. However, the slope of the isopycnals from west to east is less consistent below approximately 40 meters compared to the north-south section.

Reviewer 2: *- L304: Why is the bulge associated with mesoscale features? - L306: "characterise the simulated system as isotropic". Unclear as well, please rephrase.* **Response:** Removed, this analysis is beyond the scope of the article.

Reviewer 2: *- L315: Only here do the authors first admit that the EKE collapses by construction of the model. Knowing that the nudged EKE is resulting from an interpolated dataset containing both temporal and spatial variations and not strictly "steady state", I would expect that the resulting EKE would initially be too high. Remember that the interpolated fields are not shown so the readers are free to imagine what has come in there. So please show the interpolated fields and indicate - even roughly - the a priori expected range of values of EKE that should be reasonable.*

Response: The absence of interpolation fields has already been noted earlier. The model is not nudged to EKE, only to temperature and salinity. From the observed profiles of temperature and salinity, we cannot estimate EKE. For comparison with the expected values of EKE, references to relevant works are provided.

Reviewer 2: *- Figure 7a) shows tiny mesoscale features but a lot of the areas are white. It should be rotated (the X and Y axes have not meaning anyway) and cropped to maximise the useful area.*

Response: The X and Y axis labels are switched (an error in the figure). The figure was cropped to fit all elements, including the area over which averaging was done in stereographic coordinates. In geographic coordinates, it's difficult to assess scales.

Reviewer 2: *- Figure 7b) I cannot see the yellow solid lines. Try making them thicker.*

Response: The figure has been redone.

Reviewer 2: *- Figure 8 also has too much white area. Rotate and crop for clarity.*

Response: The figure has been redone, but not rotated. These are stereographic coordinates; rotating it would only confuse its orientation relative to Figure 1.

Reviewer 2: *- Figure 9 is very nice and even shows internal waves that are not discussed in the text. This could be added if space permits.*

Response: The figure pertains to the dynamics of individual eddies; a discussion of internal waves would require a separate section.

Reviewer 2: *- L345-349: Is the description of the eddies movements necessary for the rest of the paper?*

Response: In our opinion, yes.

Reviewer 2: *- L359-361: This argument does reach any conclusion, so I will give you mine. The ice drift change direction in the Northern part but the model forcing was constant, however the data coverage is more complete where the ice drift is sinuous. This means that the data sampling affects the simulated vortices, which should be more abundant to the top part of the graphs.*

Response: Thank you for your contribution, it is included in the new version of the manuscript: "In the northern part, the ice drift changes direction, but the model forcing remained constant. However, data coverage is more comprehensive where the ice drift is sinuous. This implies that data sampling influences the simulation of vortices, which are expected to be more prevalent in the upper portion of the graphs."

Reviewer 2: *- Section 4.3 "Method limitation" does not discuss much the limitations arising from the results but mostly limitations by construction that could have been flagged upfront in the method description instead of leaving the readers wonder about them throughout the paper. Paragraph 369-378 is probably the only proper discussion of the results and should stay there.*

Response: The section is divided into two parts; the first part has been moved to the methods section.

Reviewer 2: *- L374: What do you mean by "displaced"?*

Response: mixed layer. changed.

Reviewer 2: *- L379-385: According to the previous paragraph, the large-scale gradients can be trusted as "instantaneous" (or synoptic), but not so much the small scales. So a solution following Robinson et al. 1996 could be beneficial here: the large-scale component of the interpolated data can be used for nudging, while discard the small-scales, which can be excited by random mesoscale perturbations all over the model domain, thereby removing the "internal boundary".*

Response: We executed the process in two steps. First, we reconstructed with a coarse resolution (1 km mesh), then with a high resolution (250 m), as described in the experiment section, avoiding interpolation. In contrast to random (unknown) mesoscale perturbations, we utilized actual data.

Reviewer 2: *- Same paragraph: The same remark applies to the vertical interpolation since you have used a similar square distance function with only 3 meter characteristic depths. An "internal boundary" in the vertical may a priori have more adverse effects on this study.*

Response: As noted in the manuscript and previous answers, we do not use interpolation between layers. Therefore, data from 75 m depth have no influence on the nudging at 50 m depth.

Reviewer 2: *- L384-395: I thought that you already did a sensitivity analysis to the influence radius, did you not test a larger radius?*

Response: You are correct. It was a leftover from the draft version. The sentence has been removed to avoid confusion.

Reviewer 2: *- L388: It is clear that 1cm/s is much smaller than the 0.7 m/s ice drift but that could be noted upfront. The general direction of the currents could be indicated as well for information.*

Response: An example of the velocity distribution is shown in Figure 8. It is extremely difficult to indicate the main direction of the currents in this case.

Reviewer 2: - *L410 The number 630.000 may seem quite impressive but still does not make a proper synoptic measurement campaign. Please acknowledge that.*

Response: To avoid a discussion about oceanographic synoptic measurements which is unnecessary in this manuscript, we prefer to omit the comment on non-synoptic measurements in this context, especially since it would contradict the overview of this expedition: Overview of the MOSAiC expedition: Physical oceanography. Elem Sci Anth, 10: 1. DOI:https://doi.org/10.1525/elementa.2021.00062.

Reviewer 2: - *L412 As noted earlier, the measurements are not independent if they are located within one Rossby radius of the assimilated profiles. There are spatial autocorrelations that reduce the significance of the validation.*
Response: The response regarding the independence of data has already been given earlier.

Reviewer 2: - *L417: I would not claim that these are "dynamically consistent" as long as the measurements are collected along a 4-months trajectory across moving eddies. You have criticised the "quasi-stationary" assumption earlier so you should moderate this claim accordingly.*

Response: Corrected.

Reviewer 2: - *L421-422: Here would be the adequate place to recap these insights. I can note the vertical maxima of EKE in the Atlantic layer and the halocline. These findings may have been noted by earlier studies but it is still good to confirm or contradict earlier papers.* **Response:** Summary is extended.

Reviewer 2: - *Code and data availability: Are the in situ profiles publicly available?*

Response: All MOSAiC expedition data are publicly available and open. The data can be reached by the links in section 2.4 Observational Data, or visit the PANGAEA database and also view data on the ITP profiles on the WHOI website.

Response: All noted typos are addressed in the revised version of the manuscript.

Typos:

- l28: add a comma between Basin and Zhao.
- L. 80: DN is undefined at this point.
- l82: close the parenthesis after Fang et al. 2023 (submitted).
- L.121: Li et al. misses a year.
- L138: Define ML as Mixed Layer.
- Eq (1) indicate that z is the depth, positive downwards.
- Figure 4: The subplots labels a, b, and c are wrong in the caption. They should be b, c, and d.
- L303 "low-salinity (high-density) intrusions": should this rather be "low-density"?
- L340 "and about 5 km" is missing an "is".
- L425: "The rest".
- The reference to Sokolovskiy and Vernon has duplicate title but no journal name nor volume number.

References :

Bourgain, P., & Gascard, J. (2011). The Arctic Ocean halocline and its interannual variability from 1997 to 2008. Deep Sea Research Part I: Oceanographic Research Papers, 58(7), 745–756. https://doi.org/10.1016/j.dsr.2011.05.001

Cooper, M., & Haines, K. (1996). Altimetric assimilation with water property conservation. J. Geophys. Res, 101, 1059–1077.

Robinson, A. R., H. G. Arango, A. J. Miller, A. Warn-Varnas, P.-M. Poulain, and W. G. Leslie (1996), Real-time operational forecasting on ship-board of the Iceland-Faeroe Frontal variability, Bull. Am. Meterol. Soc.,77, 243–259.

Troupin, C, Barth, A, Sirjacobs, D, Ouberdous, M, Brankart, J.-M, Brasseur, P, Rixen, M, Alvera Azcarate, A, Belounis, M, Capet, A, Lenartz, F, Toussaint, M.-E, & Beckers, J.-M. (2012). Generation of analysis and consistent error fields using the Data Interpolating Variational Analysis (Diva). Ocean Modelling, 52-53, pp. 90-101.

Zimmerman, D., Pavlik, C., Ruggles, A. et al. An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting. Mathematical Geology 31, 375–390 (1999). https://doi.org/10.1023/A:1007586507433