Referee #2

We thank Referee #2 for the detailed and exhaustive review of our study and for the valuable feedback.

All line numbers in this response refer to the updated manuscript.

Review of "Surface Factors Controlling the Volume of Accumulated Labrador Sea Water" by Kostov, Messias, Mercier, Marshall, and Johnson

The manuscript reconstructs accumulation and variability of the Labrador Sea Water (LSW) using the ECCO state estimate. Specifically, it examines contributions of local and remote surface momentum and buoyancy fluxes to the LSW variability. In general, the present findings are very similar to those reported in several previous studies.

We have expanded the literature review in the introduction (lines 32-35, 41-59, 69-77, 124-130, 135-140) citing previous studies that have used observations and/or models of various degrees of complexity to explore the climatology and variability of LSW.

As we acknowledge in lines 131-136: "A number of previous studies have applied the adjoint of the MITgcm to exploring sources of ocean variability (e.g., Czeschel et al., 2010; Heimbach et al, 2011; Pillar et al., 2016; Jones et al., 2018; Smith and Heimbach, 2019; Loose et al., 2020; Kostov et al. 2021; Boland et al. 2021; Kostov et al., 2022), but we are the first to use this framework for reconstructing variability in the volume of LSW accumulated in the Labrador Sea. In addition, our method allows us to attribute historical watermass anomalies to different surface boundary conditions and to identify some of the physical mechanisms that govern the LSW volume budget in the ECCOv4 state estimate (Forget et al., 2015; Fukumori et al., 2017; Heimbach et al., 2019)."

We have now added a citation to Czeschel et al. 2010 (line 132). In lines 403-412, we have expanded our comparison to the results of Jones et al. (2018) and Loose et al. (2020):

"Another important feature in the sensitivity to wind stress is the pattern around the coast of Iceland. We see that a westward wind stress anomaly along the south coast and an eastward wind stress anomaly along the northern coast of Iceland promote a larger LSW volume at a lead time of three years (Figs. 4a, 5a). Jones et al. (2018) identify a similar pattern in the sensitivity of Labrador Sea heat content to wind stress. In addition, Loose et al. (2020) see this pattern in the sensitivity of heat transport across the Iceland-Scotland Ridge to wind stress. Loose et al. (2020) highlight Ekman transport along the Icelandic Coast as the mechanism behind this pattern and argue that it can generate onshore convergence or divergence and hence a pressure anomaly along the coast. This pressure anomaly is quickly communicated around Iceland as a coastal wave and affects the geostrophic transport between Iceland and Scotland, heat transport convergence in the Nordic Seas (Loose et al., 2020), and subsequent water mass transformation. Through Denmark Strait (Fig. 1), temperature and salinity anomalies from the Nordic Seas, are then exported back to the eastern subpolar gyre, where they can precondition LSW formation. "

That is in addition to the already existing comparisons and parallels drawn in lines 397 and 568.

We already stated in lines 569-570: “Our findings are thus similar to the results of Jones et al. (2018), who highlight similar enhanced sensitivity of total heat content in the Labrador Sea to wind stress along the African and West European shelves.”

Presumably, the novelty is in the use of the linear adjoint technique for this topic, noting that the adjoint had been used similarly for other applications. Overall, I feel somewhat on the fence about this manuscript.

We have tried to be more explicit about the new contributions that our study presents and put them in context. For example, in lines 136-140 we now say “Our results suggest that the upper limb of the AMOC exerts a strong lagged influence on the rate of LSW formation, a feature seen in some but not all general circulation models (Ortega et al., 2017; Li et al., 2019). It is noteworthy and novel that we identify this causal relationship between the upper AMOC limb and LSW in the ECCO state estimate constrained with historical surface boundary conditions and observations of the real ocean (Forget et al., 2015).”

We have also edited lines 563-564 following your suggestion, and there we once again highlight some of the novelty: “These novel results challenge the traditional view that wintertime cooling in the Labrador Sea is the dominant driver of interannual variability in LSW volume in the Labrador Sea.” (Previously that sentence referred to NADW).

We also elaborate on the importance of our contribution in lines 579-585 : “Some general circulation models show a significant lagged correlation between the upper AMOC limb and LSW, where the former leads the latter in time (Ortega et al., 2017; Li et al., 2019). We are the first to identify this as a causal relationship and an oceanic teleconnection in a state estimate constrained with observations (Forget et al., 2015). Our “Traffic Controller” sensitivity pattern is a geographical fingerprint associated with this oceanic teleconnection that relates flow along the upper AMOC limb to LSW formation and storage in the Labrador Sea. Surface wind stress and density anomalies can act to divert and redirect the transport of warm and saline subtropical water which is necessary for the formation of LSW in the Labrador and Irminger Seas.”

We furthermore want to clarify that our approach is fundamentally different from previous studies of watermass budgets, and that the difference extends beyond the use of the adjoint. In lines 124-130, we now say, “Our approach is different from the watermass formation and transformation framework of Walin (1982), Speer and Tziperman (1992), and Desbruyeres et al. (2019) who also use surface fluxes in their analysis. Our main constraint for defining LSW is based on vertical stratification (a component of the potential vorticity, PV), while we also define generous potential density bounds to help identify the watermass. This is similar to the definition of LSW used in Li et al. (2019). Another important difference between our framework and Speer and Tziperman (1992) is that we consider the immediate and delayed impacts of both local and remote surface heat and freshwater fluxes, as well as surface wind stress across the entire Atlantic – Arctic region.”

It keeps jumping from one point to another with quite a few figure panels, not all of which are discussed in detail. There are many small spatial scale features in the figures and their interpretations are not usually prvided.

Figure 2 is now Figure 3, and we have added further discussion in lines 371-378 following your recommendation and that of Referee 1: "For example, salt fluxes, heat fluxes, and wind stress, all, contribute to the positive LSW volume anomalies in the early 2000s. We suggest that the 2008 relative increase in LSW explored by Yashayaev and Loder (2009) can be attributed primarily to surface heat fluxes (Fig. 3a). These attribution results hold true across our three reconstructions. However, all releases of ECCOv4 seem to underestimate the magnitude of the 2008 relative increase in LSW, where the underestimation is most pronounced in release 4. We find that the subsequent 2010-2011 decline in LSW volume is dominated by heat fluxes and wind stress, while the 2012 recovery is attributable to both heat and salt fluxes (Fig. 3a). The well-documented increase in LSW volume after 2015 (Yashayaev and Loder, 2017) seems to be primarily related to wind stress anomalies and, to a smaller extent, surface heat fluxes."

In addition, following your recommendation, we have described in greater detail the small spatial scale features in Figures 4 and 5 in lines 398-412: "We also see sensitivity to zonal wind stress along the boundaries of the Labrador Sea (Figs. 4a, 5a), which can affect the strength of the boundary current and its exchange with the interior (Czeschel, 2004).We further notice that at a lead time of three years, LSW volume is sensitive to zonal wind stress just south of Greenland (Figs. 4a, 5a) in the region of the Greeland Tip Jet, where zonal winds can affect the rate of deep convection in the Irminger Sea (Pickart et al., 2003b).

Another important feature in the sensitivity to wind stress is the pattern around the coast of Iceland. We see that a westward wind stress anomaly along the south coast and an eastward wind stress anomaly along the northern coast of Iceland promote a larger LSW volume at a lead time of three years (Figs. 4a, 5a). Jones et al. (2018) identify a similar pattern in the sensitivity of Labrador Sea heat content to wind stress. In addition, Loose et al. (2020) see this pattern in the sensitivity of heat transport across the Iceland-Scotland Ridge to wind stress. Loose et al. (2020) highlight Ekman transport along the Icelandic Coast as the mechanism behind this pattern and argue that it can generate onshore convergence or divergence and hence a pressure anomaly along the coast. This pressure anomaly is quickly communicated around Iceland as a coastal wave and affects the geostrophic transport between Iceland and Scotland, heat transport convergence in the Nordic Seas (Loose et al., 2020), and subsequent water mass transformation. Through Denmark Strait (Fig. 1), temperature and salinity anomalies from the Nordic Seas, are then exported back to the eastern subpolar gyre, where they can precondition LSW formation."

I worry about the caveats listed on l.463-473, especially the robustness of the findings – though they are similar to the ones found in other studies.

We wanted to state these caveats openly, and that is why we tried to describe them in detail. At the same time, as you say, putting the results into the context of previous literature and tying them to physical mechanisms gives us further confidence in the robustness and relevance of our findings.

There seem to be larger NAO events (Fig. B1) in 1995, 1996, and 2000. Why weren't those chosen? Can additional runs, including for the summertime, be included, increasing the ensemble size? I think the robustness / fidelity of the "traffic controller" pattern should be quantified.

In lines 233-237, we have now justified the choice of years towards the end of the ECCOv4 release 2 timeseries for defining objective functions: "The selected three years are close to the end of the ECCOv4 release 2 state estimate because that allows us to compute lagged sensitivity over longer periods leading up to these three years. We use the sensitivity patterns for the ECCOv4 release 2 objective functions to reconstruct not only LSW variability in release 2, but also to attempt reconstructing LSW in the more recent ECCO release 3 (not shown) and in release 4, that extends further in time until the end of 2017."

We had limited computational resources early on and currently no access to a computational budget for this project. We have added two more runs with summertime objective functions (e.g., the new Fig.2, Fig. 3, Figure 11), but we cannot extend the ensemble further by considering objective functions in additional years.

Also, what is the reason for using freshwater fluxes for the forward experiment? Would use of momentum and heat fluxes from Fig. 4 produce similar results?

We have now clarified in lines 432-435: "The spatial pattern of the sensitivity to surface freshwater fluxes is very similar to the sensitivity to heat fluxes (Fig. 4). However, compared to heat content anomalies, freshwater anomalies do not directly trigger air-sea feedback mechanisms that are excluded from our sensitivity patterns. We thus choose to apply the sensitivity pattern from Fig. 4d as a perturbation to the background rainfall…"

The sensitivity to momentum (meridional and zonal windstress) in Fig. 4 has a different geographical fingerprint compared to surface buoyancy fluxes. At the same time, we also see that the NAC pathway stands out in Fig. 5a.

Other comments and suggestions:

l.19-20: This sentence is stronger than what is said in the text on l.409-411.

We have rephrased line 20 to read, "predict a limited yet substantial and significant fraction of LSW variability"

l.23: Rephrase the last part of this sentence.

We have rephrased lines 22-24 to read, "We point out the important role of key processes that promote the formation of LSW both in the Irminger and Labrador Seas: buoyancy loss and preconditioning along the NAC pathway, in the Iceland Basin, the Irminger Sea, and the Nordic Seas."

Fig. 1 caption: Indicate what "R" is both here and elsewhere.

We have indicated that R stands for "Pearson correlation coefficient" when it appears in lines 331, 343, 344, 361, 388, 546, and 553, some of which are figure captions.

l.89-91: Include spatial resolution information.

We added: "The global state estimate has a nominal 1° horizontal resolution and 50 vertical levels." on line 150.

l.106: Delete the comma after "Lozier".

Fixed (Line 173)

l.116: “in Appendix A”, that is delete “the”.

Fixed (Line 197)

l.121: Use larger parentheses for the most outside ones.

Fixed (Line 203).

l.127: “Figs.”

Fixed (Line 208)

l.132-135: I am not following what is said here. How does a simple scaling reduce noise? Also, despite what is said here, the units of 10^14 m^3 seem to be in use throughout the manuscript.

We have added several sentences to clarify this in lines 213-219: “Instead, we multiply the LSW volume by a large nondimensional factor of 1500. The rescaling increases the magnitude of the sensitivity patterns output by the adjoint. This eliminates some (but not all) of the numerical noise that arises when the adjoint of the MITgcm outputs the sensitivity of our objective function to surface boundary conditions. The rescaling helps as some sources of numerical noise in the adjoint have a magnitude independent of the magnitude of the objective function and its sensitivity. We then divide the lagged sensitivity patterns that the adjoint outputs by 1500 when we post-process them offline, so that the post-processed sensitivity has units of m^3 per forcing, assuming that the forcing is sustained over a single model time step set to 1 hour.”

l.137: “measure” -> “compute”.

Fixed (line 221).

l.140-141: How do these seasonal objective functions relate to monthly LSW calculations, say, as plotted in Fig. 1b?

We have added text to clarify this: “That is, we assume that a monthly objective function from March is representative of January and February monthly objective functions, too. Similarly, an August objective function is assumed to be a good substitute for July and September monthly objective functions. In addition to the 2006-2007 set of seasonally representative objective functions, for comparison, we do adjoint calculations with wintertime objective functions computed in 2006 and 2011, and summertime objective functions computed in 2005 and 2010.” (Lines 225-230)

l.144: Please refer to much earlier works on this topic.

We have also cited Dickson et al. 1996 and Rhein et al., 2017 (Line 231)

l.154: Subscripts “N” and “E” do not appear to refer to “zonal” and “meridional”. Please clarify.

We have now clarified this: “zonal (positive eastward)” and “meridional (positive northward)” (Lines 247, 365, 735) and similarly, in line 381.

l.179-181: How does this impact the adjoint optimization? Does it mean that the surface fluxes are not being adjusted?

The surface forcing remains the adjusted one, just as in the ECCO run. We have now clarified this in line 279: "However, following our approach, the model's forward trajectory remains the same as in the optimized ECCO state estimate."

Figs. 2 & 3: They can be combined to a single 6-panel figure.

Done. See the new combined Fig. 3

l.271: Both here and elsewhere, specify what is meant by "short lead times".

We have changed this to read "at lead times shorter than a year" (lines 391, 393, 529, 729, 730).

l.280: "3c" -> "2c".

The figures and their captions were merged in the new combined Fig. 3, where this typo is no longer present.

l.299: Use "NAO" as it was introduced earlier.

Fixed (Line 420).

l.307: What does "themselves" refer to?

To avoid confusion, we have dropped the word "themselves" in line 423.

l.317: Why is this particular scaling used? Also, why such a small perturbation? Is it because the adjoint has to be linear and it cannot accommodate a larger perturbation?

We have tried to clarify this in lines 437-439: "However, we multiply the pattern by a factor of (-10^{-22}), so that the rescaled perturbation is of order $10^{-8}\ m\ s^{-1}$ (or $\sim 10^{-5}\ kg\ m^{-2}\ s^{-1}$), which is comparable to the standard deviation in January surface freshwater fluxes between different years (see Fig. 8b)."

l.317-318: What is the reason for imposing only one-month perturbation and only in January 2000?

We have tried to explain our motivation more clearly in lines 439-445: "Perturbations applied during the winter are distributed over a deeper mixed layer, which further enhances their persistence and triggers a large response. This motivates our choice to branch the experiment from the ECCO state estimate (our control run) in January. On the other hand, the perturbation we prescribe is based on a sensitivity pattern from a winter-time objective function, so it would not be appropriate to extend the prescribed forcing anomaly beyond the winter season. We thus apply the perturbation only throughout January 2000 and explore its impact over the subsequent years. We have chosen the period 2000-2008 because it is marked by a resumption in the formation and storage of relatively larger volumes of LSW (Fig. 2). Hence, launching our experiment in 2000 allows us to explore this regime of enhanced LSW production."

l.320-321: This adjustment is unclear. Why are the poles adjusted, rather than considering an area-average adjustment?

We have explained the motivation behind our choice in greater detail in lines 446-449: “We adjust the amplitude of our positive and negative poles in the applied perturbation pattern such that the net input of freshwater is zero. Adjusting the poles of the applied forcing pattern is more physically consistent than redistributing the freshwater imbalance as a uniform area-averaged offset over the North Atlantic. Unlike a uniform redistribution of the imbalance, the poles of the sensitivity pattern are aligned with dynamical barriers such as inter-gyre boundaries.”

l.322-323: I cannot see this “deceleration”.

We have clarified this in lines 450-474: “When comparing against the unperturbed ECCO state estimate, we see that the “Traffic Controller” pattern affects the SSH gradients in the subpolar gyre (Fig. 6a). In the climatology (contour lines in Fig. 6a), the SSH decreases in the direction towards the Labrador Sea and Southeast Greenland. In contrast, the SSH anomaly in our experiment is more positive in the Labrador Sea and more negative along the NAC pathway (Fig. 6a). Contours of the SSH anomalies can be used to infer changes in the geostrophic component of surface currents (Jones et al., 2023). It is also important to note that in this freshwater flux experiment, we keep wind stress unperturbed, and hence do not change ageostrophic wind-driven transport relative to the control state. Therefore, the SSH anomalies in Fig. 6a tell us that our “Traffic Controller” perturbation *decelerates* NAC transport towards the western subpolar gyre (blue arrow in Fig. 6a) relative to the control run. However, the “Traffic Controller” increases northeastward transport towards the Iceland Basin and the GIN Seas (red arrow in Fig. 6a) compared to the control simulation. In addition, there is an increase in the southward transport along the East Greenland Shelf (Fig. 6a).”

l.328: delete “water”.

Fixed (Line 480).

Fig. 6 caption: “cm” -> “m”. Also, “SSS” has not been defined yet.

We corrected the unit typo and defined SSS in the Fig. 6 caption.

l.355 & 357: Both here and elsewhere, no need to repeat “at each model grid point”.

We have fixed this in line 497 and the Fig. 8 caption.

l.358 & 359: Fig. 8 is for January, so winter is shown, correct?

We added “in the winter season compared to the summer” (line 506)

l.361: Please update the date for Petit et al. citation, both here and elsewhere.

Updated on lines 508-509; 591; and updated in the references.

l.362: “ …. Surface buoyancy flux …. ”.

We added the word “buoyancy”. (Line 509)

Fig. 9 caption: To be precise, Fig. 4g is for 61 months, not 5 years.

We have changed the Fig. 9 caption, as suggested.

l.386-387: What does "short characteristic lead time" mean?

We have changed that to read "characteristic lead times that peak within two years" (Line 528)

l.391: "drive even more delayed response" why?

We have added a possible explanation: "This is consistent with the findings of Kostov et al. (2022) who suggest that LSW volume responds to surface perturbations along the Labrador Sea Western Boundary on timescales set by the propagation of signals from the western to the eastern subpolar gyre and back." (Lines 533-535)

l.392: Figure 10 panels have large magnitude differences. Please discuss implications. Are all regions equally important?

We have added a discussion in Lines 536-540: "Surface boundary conditions over different regions make contributions of very different magnitudes to intermonthly (Fig. 10) and interannual (Fig. F2) LSW variability. Among the four regions that we focus on, surface boundary conditions along the NAC pathway make the largest contribution (Figs. 10c and F2c). This further highlights the importance of the "Traffic Controller" pattern that we identify and its role in driving LSW volume anomalies via alterations to the strength and the pathway of NAC transport."

Fig. 10: Panel titles do not seem to match what these plots are. They are supposed to be contributions of the surface flux components into these regions. Not the other way around. Also, l.401: "panel b" -> "panel c".

We corrected the Fig. 10 panel titles and caption.

l.405-406: What is the reason for the "1 year" choice?

Our skill decreases drastically if we extend our annual-mean prediction two years into the future (not shown). This sharp decline in prediction skill is not a surprise, as the responses to surface heat and momentum fluxes along the NAC exhibit peaks at lead times of roughly 1.5 to 2.5 years (Fig. 10c and F2c). (Lines 555-558)

Fig. 11: Panel title does not match what is shown.... 1 year into the future, not longer than 1 year.

The Fig. 11 panel title was fixed.

l.425-426: I am not sure if this sentence is correct. The traditional view concerns (multi-)decadal time scales. The present work does not really cover that time horizon.

We have changed the text to read “These novel results challenge the traditional view that wintertime cooling in the Labrador Sea is the dominant driver of interannual variability in LSW volume in the Labrador Sea.” (Lines 563-564). Previously, the sentence vaguely referred to “North Atlantic Deepwater (NADW) masses” which inappropriately implied that we were referring to the entire low branch of the upper AMOC cell. Thank you for pointing this out.

l.427: “significant” statistically?

Yes, we added “statistically” (Line 565-566)

l.430: Delete parentheses for Pillar et al.’s year.

Changed (Line 568)

l.439: Delete “be”.

Changed (Line 577)

l.456: “plays a key role” -> “contributes”.

Changed. (Line 602)

l.469” “on” -> “in”.

Fixed (Line 614)

Fig. A1 caption: “in a” -> “is” ?

We rephrased that text (Fig. A1 caption)

l.503: What does “This” refer to?

We have changed the text to read “The steep maximum slope of the activation function raises the question …” (Line 649)

l.526: No need to repeat the figure information again.

Fixed (Line 672)

l.549: Again, why this citation? There are many seminal ones on this topic.

We have also cited Dickson et al. 1996 and Rhein et al., 2017 (Line 695)

l.580: Why does the skill peak around 2.5 years? Why is it rather low early on? Also, the decline of skill is rather small. So, does not really justify the cutoff at 6.5 years.

It is beyond the scope of our study to establish the full set of processes that explain the change in reconstruction skill with time. The reconstruction skill is low for lead times shorter than 2 years because various processes (such as the Traffic Controller mechanism) affect LSW volume at lead times beyond 2 years.

The decline of skill is indeed small after 6.5 years. What is more important is that using a history of forcing longer than 6.5 years does not improve the reconstruction. Instead it adds more noise, even if the skill does not deteriorate.

l.581: “for all points on” -> “in”.

Fixed (line 727)

Appendix D: Is this Appendix really needed?

Thank you for this question! When we presented this work at meetings prior to submission there were always questions about the sensitivity at short lead times. That is why we decided that a subset of the readers may want to see these results, and we kept them in the Appendix.

l.604: “Irminger”.

Fixed (Line 749)

Fig. F1: Why not just use one panel only with different colors for regions? In the caption: “Labrador”.

Following your recommendation, we have merged the different panels in Fig. F1, and individual regions are distinguished using colors. We have also fixed the typo in the Fig. F1 caption.

Fig. F2 caption: Same comments as for Fig. 10 caption above.

We have updated the Fig. F2 panel titles and caption following your recommendation.

l.665: “the helpful” -> “helpful”.

In this case “for the helpful discussions” may be correct. (Line 810)

l.678: Delete the first sentence.

We have deleted the first sentence. See the updated paragraph starting on line 823.

l.682: Is this language still acceptable for the journal?
The relevant code developed in this study will be made publicly available