Reply: comments 2nd reviewer

1. L210: Systematic errors such as satellite roll, and phase errors, could be much bigger sources of error on SSH than Karin random error. Can the authors elaborate on the reason why they choose to focus on Karin random error on SSH data and observable wavelengths in SSH, eddy kinetic energy, and strain rate?

You are correct, systematic errors can add >20 cm to the SWOT SSH signal. However, the SWOT Project's simulations demonstrate that 1) we are efficient in the removal of these contributions (Dibarboure et al., 2022), and 2) These simulated systematic errors affect wavelengths that are larger than the mesoscales observed in this manuscript, so the impact on our statistics is minimal. Karin random error is the key component for the small-scale gradients observed here. We have added a sentence on this in our SWOT simulator section.

<u>Old text</u>:
Line 209: "Dibarboure et al. (2022) estimate that the systematic errors alone contribute tens of centimetres in SSH. "

New text:
Dibarboure et al. (2022) estimate that systematic errors alone contribute tens of centimetres in SSH. However the SWOT Project's cross-calibration techniques are efficient in reducing these simulated systematic errors, and since their wavelengths are larger than the mesoscales observed in this manuscript, the impact on our eddy statistics is minimal.

2. L391-L395 The paper's findings don't seem to align with Sasaki et al. (2014)'s research regarding the energy feeding dynamics between different mesoscale EKE during various seasons. If the paper doesn't adequately address this discrepancy, it could be a point of critique.

The explanation for this difference may be due to the specificity of this zone, and that our simulation is only 1-year long, so the "seasonal" calculations are dominated by large eddy events. Our analysis shows that seasonal dynamics in the Agulhas Retroflection are weak compared to mesoscale events and do not strictly follow the paradigm of mixed-layer instabilities, contrary to Sasaki et al., 2014. For further detail please refer to our response to the Community comment 3. This point is also clarified in the revised version of the manuscript.

3. L233-244 U-Net's General Application: The authors show that U-net performs well even with different models and in different zones, leading to similar Root Mean Square Error (RMSE) and variance of SSH residuals. This indicates the potential universal applicability of the U-net method across various regions and circumstances. While the authors show U-Net's promise, the discussion lacks a thorough analysis of why U-Net performs similarly across various regions and models. It's not clear how U-Net achieves this consistency, which might lead to skepticism about the broad application of U-Net in other environments. Moreover, the potential improvements that could be achieved by training U-Net specifically for different regions could be discussed in more detail.

The U-net works on data statistics and aims to minimise the gap between the noisy ssh and the simulated noise. Its performance is best in its training domain, which is the case for the noise added to the NATL60 with the SWOT simulator which is well-characterised. One key point is that the SWOT simulated noise that is added to the MITgcm is based on the same spectral estimates as those used for the U-net training: it is not regional, it is a global spectral estimate. The noise mitigation from another training region is thus quite efficient. On the other hand, if the input SSH is too different from the one used for the training, the performance will decrease. For this, the NATL60 SSH used for training and the MITGCM SSH before noise mitigation are normalised with respect to their variance and mean. Then once the noisy SSH is mitigated with the U-net methods, the normalisation is reapplied to find a SSH similar to the input one.

Now, the noise on the real data may be very different from the simulated one, so we need to recharacterize the noise on the real data. Then, we will need to simulate new SSH with the actual noise and retrain the U-net with this new dataset for efficient random noise mitigation.

We have added a more detailed discussion on this in the final paragraph of the discussion: see response to comment 5 below.

4. L383 The study found that Natal Pulses, despite being the strongest events in terms of Eddy Kinetic Energy (EKE), had only a small impact on the strain, particularly on the pseudo-DUACS strain. This could be criticized if it contradicts the existing understanding of the effects of Natal Pulses on ocean dynamics or if the paper does not provide a sufficient explanation for this finding.

This is an interesting point. To our knowledge, the effect of Natal Pulses on an ocean strain field has not previously been investigated. However, Natal Pulses are large, slow-moving structures compared to the small scales. So although they have a strong SSH signature in their core, and induce gradients and currents only at their edges, they evolve more slowly in time. Figures 5 and 6 highlight that in Box 1, the large-scale CMEMS strain rate and the cross-term strain both increase slightly as the Natal Pulse passes, but this is compensated by the smothering of the small-scale strain. This probably warrants a more detailed study in some future work, maybe with real SWOT data.

5. SWOT Observability: The authors employ the U-net noise mitigation technique and assess its impact on different models. Despite the efficacy of the U-net method, it does slightly underestimate the signal amplitude. Also, there is no clear discussion on the limitations of using the U-net technique and the reasons behind its efficacy.

Thank you. We have added some of these points to the final discussion:

<u>Old text</u>
Lines 565-570: The U-Net methodology has benefits, but it needs a training data set that will not be readily available for real data. However, our study shows that the parameters tuned for the eNAtl60 model (North Atlantic zone) are still providing good results in another region, period, and with data from another simulation that accounts for different ocean dynamics. Early SWOT data in 2023 will be in a different period, and we have to correct for residual errors without a training dataset. Based on the promising results of this work, the U-net noise mitigation technique can be applied to reduce small-scale noise, and it will be interesting to see the outcome and correct potential differences.

New text:
Finally, the U-Net noise mitigation is very promising. The technique slightly reduces the signal (by a few percent) but does not over-smooth the gradients, and retains the main structures and anomalies up to the coast. This is a strong benefit for eddy diagnostics and provides better noise mitigation with respect to the other filters analysed by Treboutte et al. (2022). Our study shows that the U-net trained for the eNAtl60 model (North Atlantic zone) is still providing good results in the Agulhas region in a different period, and with data from

another simulation that accounts for different ocean SSH dynamics. One of the reasons for the U-Net efficiency is that the U-net was trained in the North Atlantic for different wave heights and seasons trying to be as representative as possible for dynamics in other regions. Besides, the simulated SWOT random errors are based on global spectral estimates and are not regionally varying. This may not be the case for the early SWOT data in 2023, with potentially a geographically-varying random error. In applying the U-net technique, we also take care that the input SSH is not too different from the one used for the training, to maintain a good performance. For this, the training SSH and the SSH are both normalised with respect to their variance and mean before the U-net technique is applied, and then the inverse normalisation is applied to recover the correct input SSH. One of the disadvantages of the U-Net method is that it needs a solid training data set that will not be readily available for the early SWOT data. However, the promising results of this Agulhas study, based on a different model and region, are very encouraging. The parameters derived from the U-net technique trained on simulated data and noise may be used for the first random noise correction for SWOT, and it will be interesting to see the outcome and correct for potential differences when applying it to the real data.