This manuscript presents a new methodology for downscaling Sea Surface Salinity (SSS) fields based on deep learning models. The ultimate goal is the application of the algorithm to satellite observations; however, the present study is restricted to model-generated data in order to demonstrate the potential of the proposed methodology.

The manuscript is generally well written. The methodology is adequately described, and the results are presented in a reasonably clear and rigorous manner. The proposed algorithm deserves publication after several revisions and clarifications.

General Comments:

Although the current manuscript is primarily intended as a proof-of-concept study, it would benefit from a more comprehensive discussion of the challenges that may arise when applying the methodology to real satellite observations.

Some of the issues that are likely to be encountered, but are not sufficiently discussed in the manuscript, include the following:

1. Use of high-resolution SSH: The manuscript relies on high-resolution SSH information. However, prior to the launch of SWOT, SSH observations were primarily derived from conventional altimetry, which has a spatial resolution comparable to that of SSS (approximately 0.25°). This limitation should be explicitly acknowledged. The authors should clarify whether the intended application relies on SWOT observations, which would inevitably constrain the temporal coverage of the downscaled SSS products.

2. Use of high-resolution SST: The use of high-resolution SST is strongly affected by cloud coverage. SST can be retrieved from microwave sensors, which provide observations at resolutions comparable to SSS, or from infrared sensors, which offer much higher spatial resolution but only under cloud-free conditions. This limitation should be discussed. In particular, the authors could highlight that, under cloudy conditions, the SSH-only version of the algorithm may become the preferred alternative, thereby emphasizing the added value of that approach.

3. Impact of SSS uncertainties: Errors in SSS observations, particularly in coastal regions, are a critical issue. These errors are not limited to random noise but often include systematic biases and other retrieval artifacts. It is therefore important to discuss how sensitive the proposed methodology is to such errors. The manuscript would benefit from a sensitivity analysis evaluating the impact of different noise levels and error structures on the reconstruction.

4. Training strategy using observations: An important aspect that should be discussed is how the model would be trained using real observations. The most widely used reference salinity dataset within the satellite community is provided by the Argo network. However, Argo observations do not provide high-resolution spatial coverage and are particularly sparse in coastal regions, which are among the most relevant application areas of the proposed

method. The authors should discuss how this limitation could affect the training and validation of the algorithm.

5. Uncertainty estimation and ensemble information: More generally, I believe the manuscript does not fully exploit the capabilities of the proposed methodology. Since the method generates an ensemble of possible reconstructions, the spread among ensemble members (e.g., their standard deviation) could potentially be used as an estimate of reconstruction uncertainty. Such information would be extremely valuable to the community. In addition, it is not entirely clear why the selected output corresponds to the realization closest to the ensemble mean rather than the ensemble mean itself. The rationale behind this choice should be explained.

Line-Specific Comments

L92–96: Please explain the impact of the normalization procedure (zero mean and unit variance) on the results.

The manuscript also states that this approach reduces computational cost. Could the authors clarify this point? Computational cost typically depends on matrix dimensions, yet the proposed approach appears to increase the number of matrices being processed. Is the method parallelized? Additionally, does the dimension reductton has any impact on the accuracy of the final reconstruction?

Figure 1: Consider including the standard deviation of SSS at the low resolution used in the right plot of the figure. This would help illustrate the loss of variability associated with the loss of resolution.

L70 and Figure 2: Have the authors considered using Sea Level Anomaly (SLA) instead of SSH? SLA generally captures clearer information on mesoscale eddies and related structures.

L108: Please specify the masking procedure that is applied.

L111: Replace “both” with “both together”.

L197: What is meant by an “MSE-type algorithm”? Please clarify.

L310: The reported differences appear relatively small and are of similar magnitude to the improvements claimed for other methods earlier in the manuscript.

L313 and Figure 7: Figure 7 does not clearly demonstrate that “DIFF-SST-SSH-GE sharpens fronts and filaments relative to DIFF-SST-SSH and RESAC.” In the pdf file the four plots look almost identical.

Figure 8 (L320–L325): While the figure shows improvements in some parts of the gradient distribution, there is also a significant negative difference, indicating that DIFF-SST-SSH-GE captures fewer pixels with weak gradients than observed in the reference distribution. Could this be interpreted as evidence of excessive sharpening? This negative difference is not discussed in the manuscript and deserves further explanation.

Figure 9 (L325–L332): Figure 9 provides additional evidence that DIFF-SST-SSH-GE may overestimate small-scale variability. At nearly all temporal scales, the amplitude of the reconstructed signal exceeds that of the reference field. This point should be discussed more thoroughly.

General comment:
It would be useful to provide a clearer discussion of the added value of RESAC relative to DIFF-SST-SSH.

L336: Please justify the statement: “their gradient magnitudes remain, on average, weaker than those of the reference fields by roughly 10%.” It is not clear how this estimate was obtained from the presented results.