I would like to thank the authors for the very interesting work and the thorough validation of the results. However I believe there is overstatements in the presentation of the paper compared to what the authors actually did, that would require some clarification before publication. The paper is presented as a reconstruction of gridded salinity fields from observation, but it is in fact an upscaling of the IAP1 product from 1 degree to 0.25 degree resolution. I have three main comments.

1 - This paper uses data already gridded by Global Circulation Models (GCM) as inputs (IAP1 and SSW). The IAP1 data is also "interpolated into unified monthly and 0.25°×0.25°spatial resolution fields" (L131-L132). Therefore the following sentences are overstatements because they are not mentioning the presence of GCM in the inputs already gridded at 0.25x0.25 resolution :
—> L94-L96 : “This paper explores the possibility of using a machine learning approach in which in situ observations and remote sensing data are merged to reconstruct the salinity changes at a 0.25°×0.25°horizontal resolution and a monthly temporal resolution from the surface down to a depth of 2000 m.”
—> L500-L503 : “This study used an FFNN approach to reconstruct a high-resolution (0.25°×0.25°) ocean subsurface salinity dataset (1–2000 m) for the period 1993–2020, in which the spatial and temporal information (time, longitude, latitude, depth), and satellite remote sensing data (ADT, SST, SSW) were input into the reconstruction. By training the functional relationship between input variables and truth values (observed salinity), the reconstruction model was established.”
To correct this would require a change of title, it is not a “reconstruction” but an “upscaling”. And also correct the abstract, introduction, and conclusion, by stating first that the network is taking IAP1 as an input. I believe IAP1 is the main source of information the network is learning from. An analysis of the relative importance of each input for each output could answer this question (see figure 10 and 11 of Pauthenet et al 2022).

2 - The “truth” data is made of in situ profiles gridded by simple arithmetic averaging (L143). It is referred to as the “in situ observations” several time in the paper, which I find confusing. Here are three examples :
—> The statement L86 to L89 leads to believe that the present study does use in situ profiles directly as an input as opposed to previous studies : “First, some studies (Lu et al. 2019; Su et al. 2020; Wang et al. 2021) used Argo gridded data rather than in situ
salinity observations data as the “truth” to train the machine learning model, and thus the reconstruction error in the Argo gridded data is embedded in the final reconstruction.”

—> L94 : “in situ observations” refer to the gridded 0.25x0.25 field here.

—> L503 : “truth values (observed salinity)”

Could you please differentiate clearly between the in situ profiles and the “truth” gridded fields of salinity? A gridded field is not “observation”.

3 - Figure 9 shows the RMSE between IAP0.25, IAP1, Armor3D and EN4 against “observation data” (L355). But is this observation data the “0.25°×0.25°, 1-month, and 41-level grid averages” or the in situ profiles? if it is the grid average, this presentation is biased towards IAP0.25 as it is the only product that was trained with the grid average, so the other products are expected to have higher RMSE. If it is the in situ profiles compared to colocated profiles from EN4, Armor3D, IAP0.25 and IAP1, then it would be interesting to see the RMSE between the in situ profiles and the grid average used a “truth” data. Could you add this RMSE profile to the figure?

The IAP1 data is also “interpolated into unified monthly and 0.25°× 0.25° spatial resolution fields” (L131-L132). Could you plot the RMSE between this regridding of IAP1 and in situ profiles on figure 9 too? That would be a good justification for using the neural network for this upscaling.

Best regards,

Etienne Pauthenet

Reference :