

We thank the Editor for the clear guidance on the revision of the manuscript. We will provide detailed answers to the referees' comments separately.

### **Additional Evidence for Code Correctness**

As requested, to provide evidence of the code's correct implementation, we have performed a series of idealized experiments analysing misfits and residuals. Instead of a full twin experiment, we designed specific tests for the assimilation of SLA tracks and Temperature/Salinity profiles. We specifically analysed the sensitivity of the residuals to the prescribed observational error. For these tests, synthetic observations ( $y$ ) were generated at the analysis time ( $t_a$ ), allowing for a direct assessment.

With  $x_b$  as the model background and  $\delta x$  as the increments, the analysis ( $x_a$ ) is defined as:

$$x_a = x_b + \delta x$$

The misfits ( $d$ ) and residuals ( $r$ ) are then:

$$d = y - H[x_b]$$

$$r = y - H[x_a]$$

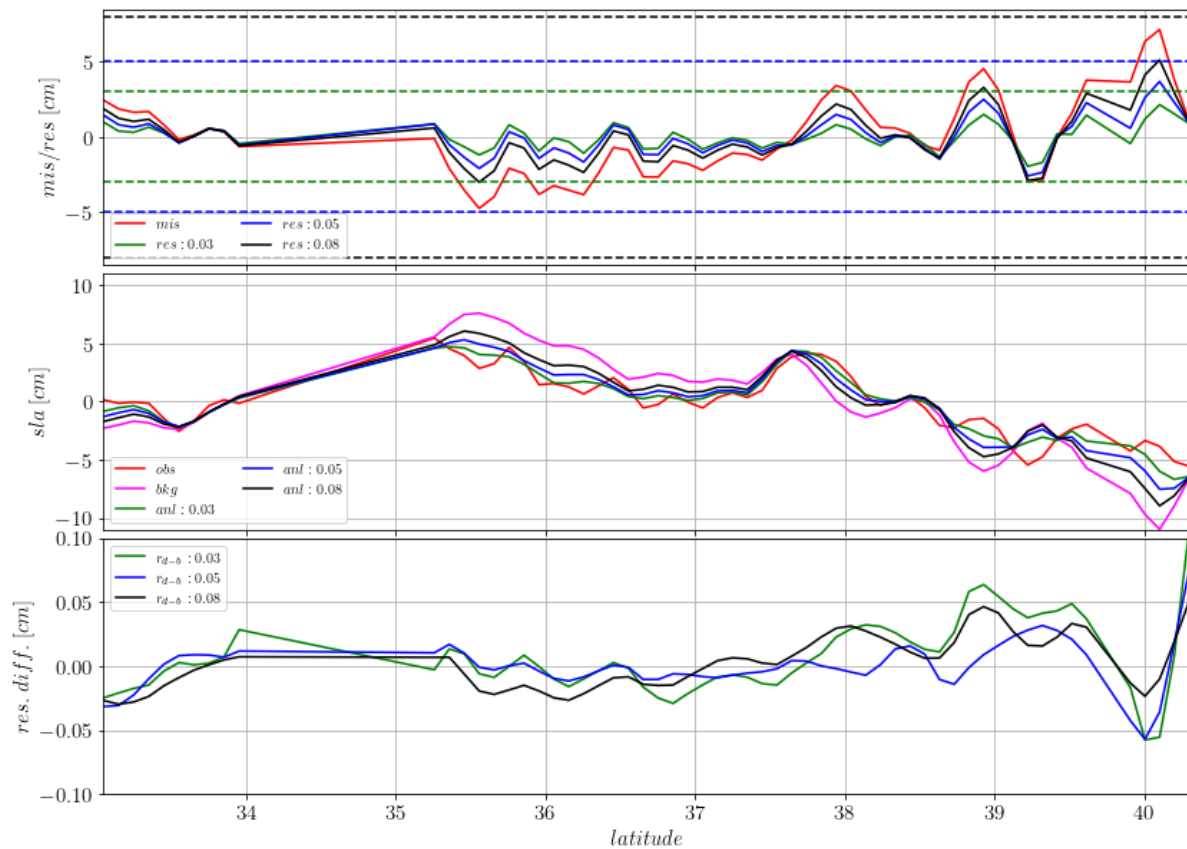
Consistent with our realistic experiments, the observation operator ( $H$ ) is a bilinear interpolator. We have chosen not to include these idealized experiments in the main manuscript as we believe they are not central to its narrative. However, we agree they provide a crucial check of the code's functionality and are happy to share them here as supplementary proof of correctness. We believe these additional analyses address the Editor's concerns and provide sufficient evidence of the correctness and robustness of the code used in our manuscript.

### **Test 1: SLA track Assimilation**

To specifically verify the SLA operators, we assimilated a synthetic SLA track representing the data after bias and tidal signal removal. We performed the minimization using both the barotropic model and the dynamic height operators and using 3 different observational error values: 3 cm (as in the submitted manuscript); 5 cm and 8 cm.

Figure 1 shows the results. The top panel shows that the analysis residuals (coloured lines) are consistently smaller than the initial misfit (red line) and fall below the corresponding prescribed observational error (dashed lines). The central panel visualizes the background, the synthetic truth, and the resulting analyses. In the bottom panel the differences between the residuals ( $r|_{DH} - r|_{BM}$ ) confirm that the difference in SLA residuals produced by the two operators is minimal. This is expected, as the primary impact of the choice of operator is on

the projected temperature and salinity profiles, as shown in Figure 3 of the submitted manuscript.

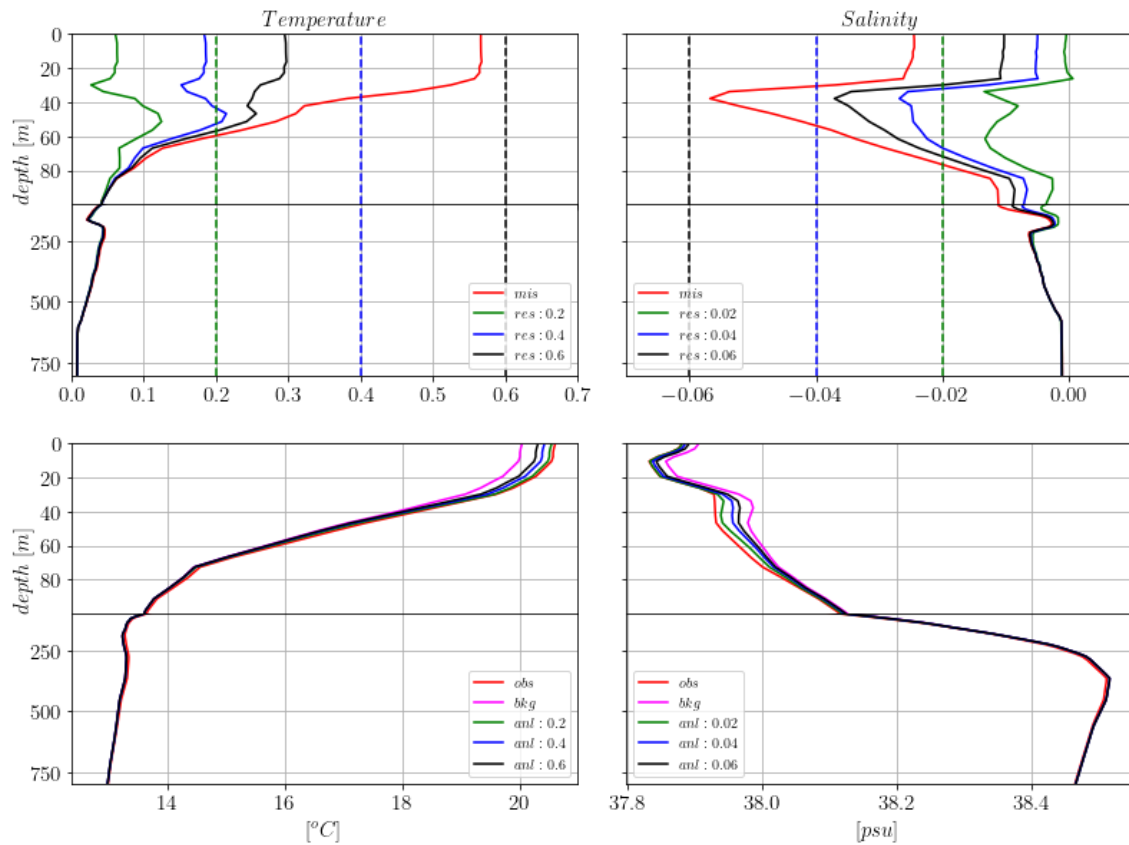


**Figure 1:** SLA Assimilation Test. **Top panel:** Initial misfit (red) and analysis residuals (solid-coloured lines) for experiments with varying observational errors (dashed lines). **Middle panel:** Synthetic “true” observation (red), the model background (magenta), and the resulting analyses (coloured lines). **Bottom panel:** Difference between residuals obtained using the barotropic and dynamic height operators. The numbers in the legends indicate the observational error used in the specific experiment ( $m$ ).

## Test 2: Temperature and Salinity Profile Assimilation

While the manuscript's focus is not on in-situ data, we conducted similar tests for Temperature ( $T$ ) and Salinity ( $S$ ) profiles. Assimilating a single  $T$  or  $S$  observation is trivial, but assimilating a vertical profile is more complex because our system (OceanVar) enforces multivariate vertical correlations via pre-computed Empirical Orthogonal Functions (EOFs) in the Background Error Covariance (BEC). To create a clear test, we generated the synthetic observations to yield misfits that are consistent with the vertical structure of the BEC. In sensitivity experiments, we assimilated these profiles with prescribed observational errors of 0.2, 0.4, and 0.6 °C for temperature, and corresponding errors of 0.02, 0.04, and 0.06 psu for salinity.

The results in Figure 2 confirm the code functions correctly. For both temperature (left) and salinity (right), the analysis residuals decrease as the prescribed observational error is reduced, consistently remaining below the error threshold.



**Figure 2:** Temperature and Salinity Profile Assimilation Test. **Top Panels:** Initial misfit (red) and analysis residuals (coloured lines) for Temperature (left) and Salinity (right) for different prescribed observational errors. **Bottom Panels:** Synthetic "true" profiles (red), model background (magenta), and the resulting analysis profiles (coloured lines).