Updates to the global FOAM system including model and data assimilation changes

February 16, 2024

Technical Report No: 659

Davi M. Carneiro[1], Jennifer Waters[1], Daniel Lea[1], Matthew Martin[1], James While[1], Anthony Weaver[2], Arthur Vidard[3], Catherine Guiavarc'h[1], Dave Storkey[1], David Ford[1], Ed Blockley[1], Jonathan Baker[1], Keith Haines[4], Martin Price[1], Mike Bell[1], and Richard Renshaw[1]

[1]Met Office, Exeter, UK

[2]CERFACS, Toulouse, France

[3]Inria, Grenoble, France

[4]NCEO and University of Reading, Reading, UK

Executive summary

Global ocean/sea-ice monitoring, forecasting and reanalysis systems at the Met Office are used to provide information to various users including defence, marine navigation and science users. Improvements to the ocean/sea-ice data assimilation (DA) system are implemented and tested in a research and development framework and then handed over to operational teams for implementation. This report describes and evaluates the impact of model and data assimilation changes implemented in the latest global ocean/sea-ice model GOSI9 on the performance of both 1/4º (ORCA025) and 1/12º (ORCA12) configurations in the Forecast Ocean Assimilation Model (FOAM). For the model configuration, GOSI9 uses NEMOv4.0.4 instead of NEMOv3.6, applying the TEOS10 formulation for the seawater equation of state, with improved parameter settings. The sea-ice model has also been changed in GOSI9 from CICE to SI^3, although their configurations remain quite similar. Major DA updates have also been made, including an improved specification of observation errors for sea surface temperature (SST) and sea level anomaly (SLA), the use of only a short length-scale for temperature as part of the background error specifications, an increase in the number of minimisation iterations, the calculation of horizontal and vertical normalisation factors separately, changes to the Mean Dynamic Topography used to assimilate SLA data, and finally the implementation of physically-based checks, based on a Brunt-Väisälä stability threshold, to reject temperature (T) and salinity (S) increments where the water column becomes unstable. Additionally, a multigrid assimilation has been implemented to increase the DA computational efficiency.

GOSI9 FOAM runs have been conducted in 2019 for both ORCA025 and ORCA12, and compared with the results from GO6 runs, which consist of sea ice model, ocean model, and DA components from the previous version of the system. For both ORCA025 and ORCA12, major improvements are seen in the GOSI9 observation-minus-background statistics relative to those from GO6, with GOSI9 Root Mean Squared Differences decreasing globally by ~20% and ~25% for SLA and sub-surface temperatures, respectively. The SLA and sub-surface temperatures in GOSI9 are much better represented in western boundary current regions, and especially along the Antarctic Circumpolar Current. The main drawback of the GOSI9 configuration is related to localised water column instabilities, exacerbated by DA in the absence of a long length-scale component of the background errors for temperature. These localised water column instabilities are present in the Mediterranean Outflow region, but they are largely mitigated by rejecting T/S increments based on a Brunt-Väisälä stability threshold. For the sea ice, although substantial improvements are seen in the Arctic over summer, the sea-ice concentration in GOSI9 is slightly worse overall than in GO6, particularly in the Antarctic.

Additional runs have been conducted, covering 2019, withholding T/S profile data to assess the improvements of GOSI9 in comparison to GO6 from a reanalysis perspective. The purpose of these runs was to represent the pre-Argo period. Unlike GO6, the absence of T/S profile assimilation in GOSI9 does not significantly affect the SST and SLA results, and the T/S statistics are much less degraded in GOSI9 relative to GO6. Hence, much more consistent

results for ocean transports and heat content are achieved in the GOSI9 run withholding T/S profile observations when compared to the equivalent GO6 run. These improvements in GOSI9, even when there is no T/S profile assimilation, are due to removing the propagation of the SLA signal onto large-scale subsurface temperatures, leading to a promising configuration for improved reanalysis runs. From an ocean forecast perspective, these GOSI9 updates will also be implemented in the next scientific upgrade of the operational suite, which is expected to improve not only FOAM but also the fully coupled numerical weather prediction and the seasonal forecast systems, benefiting their wide range of users.

Contents

Executive summary 2
Contents 4
1. Introduction 5
2. Overview of GO6 - the current version of the FOAM system 7
3. Model updates in GOSI9 9
3.1 NEMO 4.0.4 Package 9
3.2 TEOS10 Equation of State 10
3.3 Sea Ice Model 11
4. Data assimilation changes in GOSI9 12
4.1 Assimilation interface with SI3 12
4.2 Updates of NEMOVAR configuration 12
4.3 Brunt-Vaisala verifications for T/S increments 17
4.4 Decomposed normalisation factors 20
4.5 Multigrid assimilation 21
4.6 Summary of changes between GOSI9 and GO6 23
5. Results 24
5.1 GOSI9 and GO6 comparisons 24
5.1.1 ORCA025 24
5.1.2 ORCA12 33
5.2 Potential impact of GOSI9 changes on pre-Argo reanalysis 41
6. Conclusions 51
References 53
Appendix A – Impacts of the model and assimilation changes on GOSI9 56
Appendix B – MDT updates 60
Appendix C – ORCA025 T/S statistics in marginal seas 64
Appendix D – ORCA12 T/S statistics in marginal seas 65

1. Introduction

Global ocean/sea-ice monitoring, forecasting and reanalysis systems at the Met Office are used to provide information to various users including defence, marine navigation and science. They are also included as part of the models used to produce short-range weather and seasonal forecasts so that important coupled ocean/sea-ice/atmosphere/land exchanges can be represented. An important aspect of the ocean/sea-ice system, FOAM (Forecast Ocean Assimilation Model), and the ocean/sea-ice/atmosphere/land systems used for coupled Numerical Weather Prediction (NWP) and seasonal forecasting (GloSea) is their initialisation. The ocean/sea-ice components of the models are initialised using a common data assimilation (DA) framework for the three types of forecasts, as well as for ocean/sea-ice reanalysis. The FOAM and coupled NWP systems both apply this ocean/sea-ice DA procedure separately within their own frameworks: FOAM runs its own DA; the coupled NWP system runs the ocean/sea-ice DA as part of a weakly coupled DA framework (which also includes atmosphere/land DA); and the GloSea system uses initial conditions produced by the coupled NWP system. The only major difference between the FOAM and coupled NWP ocean/sea-ice DA systems is the time-window used to collect observations and apply the DA, which is set to be 24 hours for FOAM and 6 hours for coupled NWP.

Improvements to the ocean/sea-ice DA system are implemented and tested in a research and development framework and then handed over to operational teams for implementation. Upgrades to model versions of the global FOAM system, the coupled NWP system and the GloSea seasonal forecasting system are coordinated so that the initial conditions are appropriate for the forecast model used operationally. The GloSea system also requires a long reanalysis of the ocean to initialise the reforecasts needed for real-time forecast calibration, and the system used to produce this reanalysis needs to be consistent with the real-time forecasting system. The model development process in the Met Office is organised through GC (Global Coupled) model versions which consist of specific versions and configurations of atmosphere (GA), land (GL), ocean (GO) and sea-ice (GSI) models. A new GC version has recently been developed called GC5, which includes the GOSI9 ocean and sea-ice components, and the initialisation of forecasting systems using this new model version needs to be implemented and tested before it can be run operationally.

Alongside the model developments described above, development of the ocean/sea-ice DA capability is a continuous process. This includes development of underpinning DA algorithms, inclusion and testing of new observation types, as well as improvements to the error covariances and other ancillary information needed by the DA. The DA software we use is called NEMOVAR (Waters et al., 2015) and is developed collaboratively with ECMWF, INRIA and CERFACS.

This report describes the implementation and testing of DA in the GOSI9 ocean/sea-ice components of the GC5 model. Section 2 provides a brief overview of the current operational

FOAM system which is based on the GO6 model version. Section 3 then gives an overview of the model changes since then. These changes include moving to a new sea-ice model called SI^3, a more recent version of the NEMO model (NEMO4.0.4) with a new equation of state, and various improvements to parameter settings, in particular to improve the realism of ocean simulations in the Southern Ocean. In section 4, we provide a description of the DA changes we have made. These include changes related to getting the DA to work correctly with the model changes, e.g. implementing sea-ice concentration DA capability in SI^3 and making sure the new ocean model state variables are compared with the observations correctly. We also go into some detail regarding updates to the DA configuration including a more recent version of NEMOVAR, which includes: changes to the observation and background error covariances, an increase in the number of minimisation iterations, the implementation of Brunt-Väisälä stability checks to reject temperature (T) and salinity (S) increments in regions where DA would make the water column unstable, changes to the Mean Dynamic Topography (MDT) used to assimilate altimeter sea level anomaly (SLA) data, and other changes related to improving the computational efficiency of the DA. For each of these changes we demonstrate the impact the change has on the quality of the system. We then show the overall impact of the changes in one-year experiments with two global model configurations, one at 1/4° horizontal resolution and one at 1/12° resolution. We also assess the potential impact of the changes on reanalyses before the Argo period by running another set of experiments with the 1/4° system withholding T/S profile data. Finally, we summarise the impact of the model and DA changes and draw conclusions.

2. Overview of GO6 - the current version of the FOAM system

The model and data assimilation components of the current version of FOAM are described in some detail by Barbosa Aguiar et al. (2024), together with an assessment of the impact of changing the horizontal resolution from 1/4° to 1/12°. Here we briefly summarise the main characteristics of the model and DA to provide context for the changes being tested.

2.1 Ocean and sea-ice models

As mentioned above, the FOAM system currently uses the GO6 model which is described by Storkey et al. (2018) and Ridley et al. (2018) for the ocean and sea-ice components respectively. Two configurations are available, one at 1/4° horizontal resolution called ORCA025 and one at 1/12° horizontal resolution called ORCA12. These are both run using a tripolar horizontal grid where the two poles in the northern hemisphere are placed over land to avoid singularities in the part of the grid where computations are carried out. Both configurations have 75 vertical levels, with about 1 m vertical resolution in the top 10 m of the ocean. The bathymetries used in the two configurations were derived differently, with more smoothing of the bathymetry applied in the ORCA025 configuration. A non-linear free surface is used in which the cell thicknesses throughout the water column are allowed to vary with time. Vertical mixing is modelled using the Turbulent Kinetic Energy (TKE) scheme. Convection in the model is parameterized by significantly increasing the vertical diffusivity where the water column is unstable. River inputs to the ocean are specified using a monthly climatological dataset. Freshwater inputs from melting icebergs are represented through an online Lagrangian iceberg model.

The sea-ice model in the current FOAM version is the Los Alamos National Laboratory sea-ice model CICE. It has multi-layer, energy-conserving thermodynamics with four layers of ice and one of snow. The ice rheology is modelled with an elastic-viscous-plastic approach, and the model has multi-category ice thicknesses with 5 categories. The impact of surface melt ponds is included.

2.2 Data assimilation

In FOAM, the NEMOVAR system is used to perform multi-variate, incremental 3DVar DA with first-guess-at-appropriate-time (FGAT). Model values are interpolated to the observation locations at the correct time during a short model forecast. The observation minus model values (called the innovations) are fed into the NEMOVAR code which then generates a set of changes to be made to bring the model closer to the observations (called the analysis increments). These increments are then added into the model using Incremental Analysis Updates (IAU). It is worth mentioning that ORCA12 DA is currently performed at ORCA025

resolution since it is very computationally expensive to run the DA at ORCA12 full resolution.

The observation types assimilated operationally in FOAM are in situ and L2 satellite sea surface temperature (SST) data, altimeter SLA data, in situ profiles of temperature and salinity from various platforms (including Argo profiles, moored buoys, XBTs, CTDs, marine mammals, gliders), and sea-ice concentration (SIC) data from SSMI/S satellite sensors. Variational observation bias correction is carried out for satellite SST data using the scheme described in While and Martin (2019) and for SLA data using the scheme described by Lea et al. (2008). A recent description of the way NEMOVAR is implemented in FOAM is included in Barbosa Aguiar et al. (2024).

Important inputs to NEMOVAR include the observation and background error covariances. Observation errors are assumed to be uncorrelated and we specify the observation error variances through seasonally and spatially varying estimates produced using Hollingsworth and Lonnberg (1986) based on data from a previous reanalysis.

Background error covariances are assumed to be correlated horizontally, vertically and between variables. The inter-variable correlations are represented by physically-based balance relationships, and variables used in the 3DVar scheme are transformed to a set of assumed mutually uncorrelated control variables, namely temperature, unbalanced salinity and unbalanced sea surface height (SSH). After generating the analysis increments, these variables are transformed back to the state variables of temperature, salinity, SSH and the horizontal velocities (increments to the velocities are geostrophically-balanced with the other increments). Sea-ice concentration is analysed separately from the other ocean variables.

Horizontal correlations in the background errors for the control variables are modelled using an implicit diffusion operator (Weaver et al., 2016). The horizontal correlations in temperature and unbalanced salinity are both represented by a combination of two Gaussian functions, each with a different length-scale (Mirouze et al., 2016). The small-scale errors have a length-scale which depends on the first baroclinic Rossby radius, while the large-scale errors have a 400 km length-scale. Unbalanced SSH has a 400 km length-scale, while sea-ice concentration has a 25 km length-scale.

Vertical correlations in temperature and unbalanced salinity are parameterised near the surface using the model background mixed-layer depth. The surface vertical length-scale is specified to be the depth of the current local mixed layer, with the length-scale decreasing to be twice the vertical grid spacing at the base of the mixed layer and below it. This allows surface data to have a strong influence within the mixed layer, with little influence below it.

3. Model updates in GOSI9

3.1 NEMO 4.0.4 Package

NEMO 4.0.4 package, used in GOSI9, includes various updates relative to NEMO 3.6. These are described in detail in Guiavarc'h et al. (2023). Two of the most significant changes to the model are the change to the equation of state and the new sea ice model. These are discussed in sections 3.2 and 3.3 - here we will give an overview of other model changes. There is a new implicit-adaptive vertical advection scheme (Shchepetkin, 2015) which allows implicit vertical advection to be used in regions where the vertical Courant-Friedrichs-Lewy (CFL) condition is likely to be breached. In other regions an explicit vertical advection scheme is used. This development maintains the accuracy of the explicit scheme but allows for a longer time-step. The time-step for FOAM ORCA12 has increased from 180 seconds to 400 seconds and this produces a significant improvement in run-time for the ocean-ice model from approximately 50 minutes to 26 minutes for the 24-hour observation operator task.

A 4^{th} order tracer advection scheme is implemented in the NEMO 4.0.4 package. Using a higher order advection scheme was found to reduce spurious vertical mixing. There was also tuning of the Turbulent Kinetic Energy (TKE) mixing depth between 10°S and 40°S to reduce temperature biases. An increase in the Chlorophyll concentration parameter was also set which reduces the solar radiation penetration depth in the ocean model. In addition, there was some tuning to increase the Antarctic Circumpolar Current (ACC) transport in GOSI9, including an increased topographic drag in the Southern Ocean due to changing the lateral boundary condition from a free-slip to a partial-slip condition (Storkey et al., 2024).

A space and time-dependent version of the Gent-McWilliams scheme (Tréguier et al., 1997) is applied to both ORCA025 and ORCA12. It follows the approach of Hallberg (2013) that eddies should be explicitly represented in parts of the domain where the model resolution is sufficiently fine and parameterised where it is not. Therefore, the Gent-McWilliams scheme is mostly applied to ORCA12 at very high latitudes (poleward of ~60°S and 60°N). Additionally, in GO6 an interactive iceberg model was used, but this was found to be unstable in NEMO 4.0.4, and therefore an iceberg climatology is used in GOSI9.

The GOSI9 model changes described above improved the observation-minus-background SLA statistics in the Southern Ocean relative to GO6 (see Appendix A), particularly the Root Mean Squared Difference (RMSD). However, this Southern Ocean SLA RMSD improvement coming from the model changes contributed only to 37.5% of the total RMSD improvement seen in GOSI9 for this region, while the DA changes (see section 4.2) contributed to the remaining 62.5%. Although there are some improvements to the North Atlantic temperatures at depth and to the Southern Ocean salinities near the surface, the impacts of the GOSI9 model changes on the observation-minus-background statistics for these variables are mostly negligible across the different ocean basins (see Appendix A).

3.2 TEOS10 Equation of State

A significant development for GOSI9 is the change of the equation of state from EOS80 to TEOS10. TEOS10 was adopted as the official description of seawater by the Intergovernmental Oceanographic Commission in 2009 (IOC, SCOR and IAPSO, 2010). The TEOS10 equation of state changes the temperature and salinity variables used in the model and analysis from potential temperature (PT) and practical salinity (PS) to conservative temperature (CT) and absolute salinity (AS). The change of variables allows for a consistent representation of thermodynamic properties in the model.

Observations of temperature (both SST and profiles) are generally provided as in-situ temperature measurements while observation of salinity are provided on the practical salinity scale. We already convert temperature profiles observations prior to assimilation from in-situ to potential temperature. A similar conversion is not required for SST as in-situ and potential temperature are the same at the surface. A new pre-processing step has been included in the NEMO observation operator to allow the conversion of temperature observations to conservative temperature and salinity observations to absolute salinity. Conservative temperature is not the same as in-situ temperature at the surface, so it is also necessary to convert SST observations. For the temperature conversion, a corresponding salinity value is required. Where there is a corresponding salinity observation (e.g., for Argo floats), the observed salinity is used in the conversion. When this is not the case, the model salinity is used to convert the temperature observations, which is a similar approach used in GO6 for the conversion from in-situ to potential temperature. In addition to the changes in NEMO, the equation of state used in the NEMOVAR balance operator was also updated to use TEOS10.

Figure 1 illustrates the difference in EOS80 and TEOS10 variables at the surface and at 300 m. The largest differences in potential and conservative temperature are near the surface and in the Mediterranean, Baltic, Black Sea and Caspian Sea. For salinity, the difference between practical and absolute salinity is negative everywhere. In most regions, outside the Baltic and closed seas, the difference is around 0.16. These results demonstrate the importance of dealing with the temperature and salinity variables carefully in the observation operator, data assimilation and model validation.

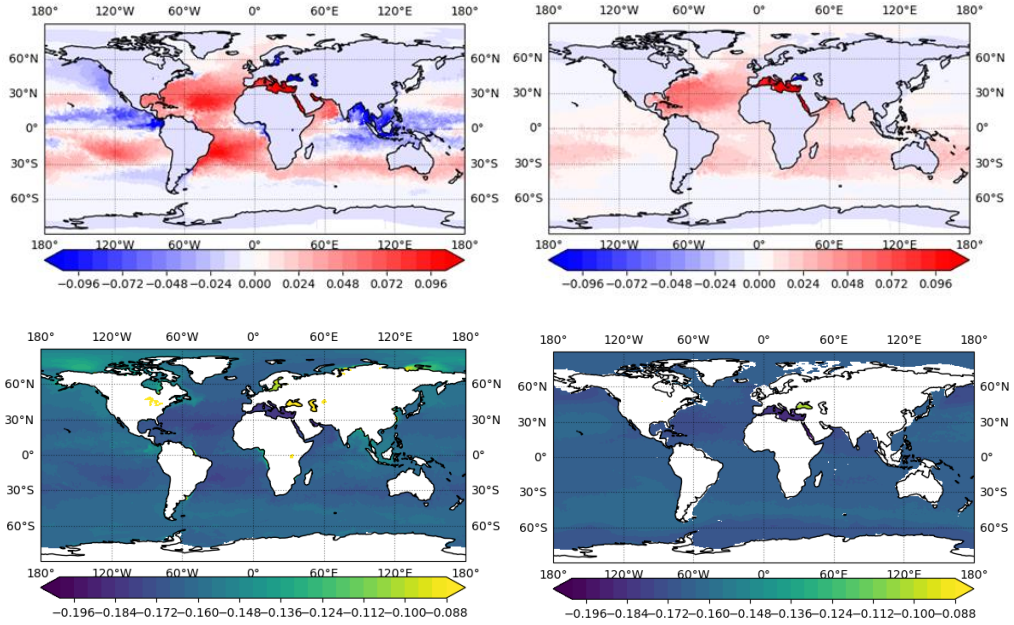

Figure 1: The top plots show the difference in Potential Temperature and Conservative Temperature (PT minus CT) at the surface (left) and 300 m (right). The bottom plots show the difference in Practical Salinity and Absolute Salinity (PS minus AS) at the surface (left) and 300 m (right). All plots are valid on the 1st of February 2015.

3.3 Sea Ice Model

The sea-ice model component of GOSI9 is based upon the native NEMO sea-ice model, SI^3 (Vancoppenolle et al., 2023). SI^3 was developed from the LIM3 model with some functionality merged from CICE, which is the sea-ice model used in GO6. SI^3 was first made available at NEMO version 4.0; it is fully embedded in the code and invoked from within the Surface Boundary Code module. The version of SI^3 used at GOSI9 is based on the NEMO version 4.0.4 release. Aside from the change in the sea-ice model, the physics itself remains largely similar to CICE configurations in GO6. Like CICE, SI^3 is a dynamic-thermodynamic continuum sea ice model that includes an ice thickness distribution, conservation of horizontal momentum, an elastic-viscous-plastic rheology, and energy-conserving halo-thermodynamics (Vancoppenolle et al., 2023). For the GOSI9 configuration, five thickness categories are used to model the sub-grid-scale ice thickness distribution, and an additional ice-free category represents open water. The bounds of the thickness categories are determined using a function of domain-mean ice thickness, which is specified as 2.0 m. This sets the maximum thickness category bounds to 0.00 m, 0.45 m, 1.13 m, 2.14 m, 3.67 m, and 99.0 m.

SI^3 is run on the same grid as the NEMO ocean model component and on every ocean time-step. An advantage of using the ice model native to NEMO is that the interpolation of velocity points required between NEMO (Arakawa C-grid) and CICE (Arakawa B-grid) in previous configurations is no longer necessary. For a more detailed description of sea-ice modelling in GOSI9, see Blockley et al. (2023).

4. Data assimilation changes in GOSI9

4.1 Assimilation interface with SI3

In order to apply sea-ice concentration (SIC) increments within the SI^3 model, the IAU code for SIC in GOSI9 was developed following the same core methodology as in GO6 with CICE (Peterson et al., 2015). Positive ice concentration increments are always added to the SI^3 thinnest category of ice (consisting of ice up to 0.45 m), while negative increments are first removed from the thinnest available category until it reaches zero concentration, and then progressively removed from thicker categories. When ice is removed, a volume of ice associated with the grid point average thickness of ice for that category and change in concentration is removed. New ice is added with a thickness of 0.45 m, which is thicker than frazil ice (i.e., ice added due to freezing of sea water) to prevent immediate melting of new ice.

Finally, all other model prognostic variables in SI^3 are adjusted proportionally to changes in ice concentration and volume, so equivalent values can be maintained after adding SIC increments. This is the case for the snow volume, snow and ice enthalpy, ice pond concentration and volume. The comparison between GO6 and GOSI9 SIC assimilation results will be described for both ORCA025 and ORCA12 in section 5.1.

It is worth highlighting that other ways of applying SIC increments to SI^3 have been tested, for example, assuming that SIC increments should be added respecting the background distribution of concentration across the multiple ice categories. This methodology is referred to as RED by Smith et al. (2015). However, although the total SIC results are as good as the methodology described in Peterson et al. (2015), the sea-ice thickness (SIT) results using RED are much worse when compared to SIT observations available (not shown). According to Smith et al. (2015), additional ice formation, melt and deformation can be represented incorrectly by using RED to spread SIC increments, as these processes act non-linearly on the different ice categories, which might result in unphysical ice volumes. Therefore, the methodology described by Peterson et al. (2015) was chosen to add SIC increments into SI^3 as part of GOSI9 DA.

4.2 Updates of NEMOVAR configuration

A few updates to NEMOVAR parameters and input files were also made in GOSI9. SLA and SST Representation Errors (REs), due to unresolved model scales, were calculated for ORCA025, using a python-based code (available from GitHub[1]) to compute observation errors (Mignac, 2021). The REs were estimated by accounting for the model sub-grid scale variability

[1] https://github.com/MetOffice/ocean_error_covs

present in the observations (Oke and Sakov, 2008), following the steps below:

- Provided that there are enough observations and that the observations adequately span the grid cell, all the observations that fall within each grid cell are identified and their average is calculated for a given time window. At least four observations are required to proceed with the RE calculation in a particular grid cell, otherwise the grid cell is masked for that specific RE calculation.
- The difference between each of the observations and the cell average calculated above is computed.
- The standard deviation of the differences from above is calculated for each model grid cell. This provides an estimate for the cell averaged RE for a particular point in time for each model grid cell.

The SLA REs were calculated every 5 days for 2017-2018, using along-track data from all the available altimeters within a 5-day window, to mitigate sampling errors. For the SST, only night-time VIIRS data was used. This was because it is the reference dataset for the bias correction of SST observations in FOAM, and it has consistent daily global coverage. The SST REs were calculated every 3 days for 2017-2018 using a 3-day window. Both SST and SLA REs were averaged over the seasons to be used as input files for NEMOVAR. After averaging the REs, additional steps were done, such as smoothing the fields and infilling any ocean grid cells where the minimum number of observations, required for the RE calculation, was not met. Since ORCA12 DA is currently performed at ORCA025 resolution, the same SST and SLA REs for ORCA025 were used in ORCA12 assimilation. The ORCA12 REs will be improved in the future when ORCA12 DA is performed at its full spatial resolution.

On top of the seasonally averaged REs, a measurement error of 4 cm was added to SLA observations, resulting in the SLA observation errors used by NEMOVAR in GOSI9. As an example, Figure 2 shows that the new SLA observation errors in GOSI9 are much larger than the old ones in GO6. Apart from within the western boundary currents and the ACC, the SLA observation errors in GO6 are even smaller than the measurement error of 4 cm used in GOSI9. This indicates that the SLA assimilation in GO6 may have been overfitting the observations. However, it is worth noting that using observation time windows greater than 1 day, to reduce sampling errors as done by Oke and Sakov (2008), might also overestimate the REs calculated for GOSI9.

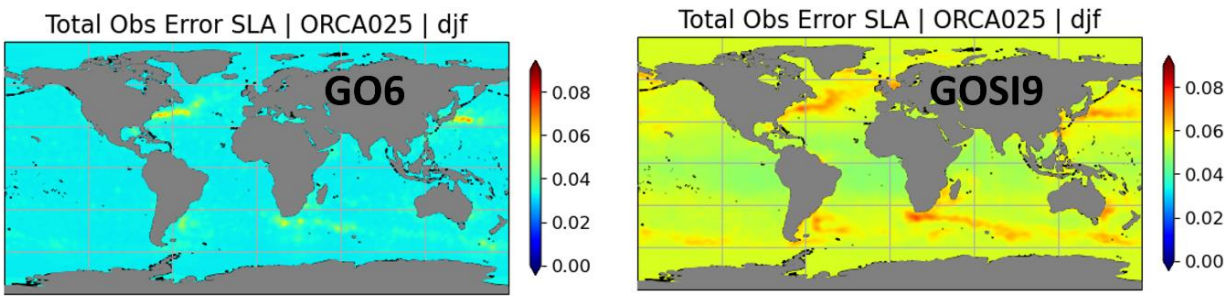

Figure 2: ORCA025 SLA observation errors (m) in winter for GO6 (left) and GOSI9 (right).

The measurement errors for SST are defined for each single observation, using the measurement uncertainties given by the data providers. It is worth mentioning that the use of observation-specific SST measurement error in a global FOAM configuration is a new development in GOSI9. The SST observation errors in GO6 were only represented by seasonal estimates, making them much more generic and less optimal for the assimilation. Therefore, when the new SST REs and measurement errors are used in ORCA025 DA, SST RMSD improvements are seen relative to an equivalent run with old SST observation errors, decreasing from 0.41 to 0.38°C globally when compared to in situ SST measurements (Figure 3).

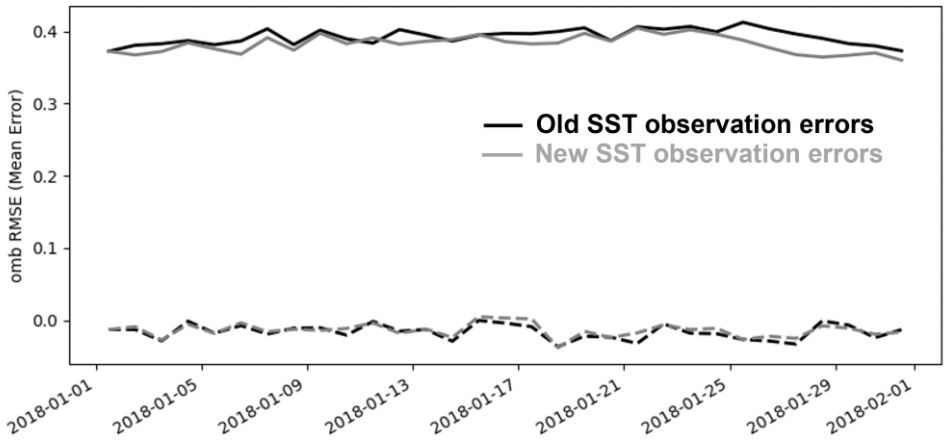

Figure 3: Global SST RMSDs (solid) and mean observation minus background (dashed) in °C when the SST assimilation uses new (grey) and old (black) observation error estimates for January 2018. These statistics were calculated with respect to in situ SST observations.

In addition to changing the SLA and SST observation errors, only the short length-scales of the background errors are used in NEMOVAR to horizontally spread the temperature information in GOSI9. The large-scale DA corrections for temperature are removed to mitigate known issues with the projection of the SLA signal onto large-scale temperature at depth, which can lead to drifts and spurious variability in sub-surface temperatures and heat content, particularly in the pre-Argo period. However, the long length-scales of the background errors are still used to perform large-scale DA corrections in the unbalanced salinity fields, due to the sparsity of salinity observations.

Figure 4 shows that removing the large-scale DA corrections for temperature can improve the temperature statistics globally in ORCA025, both in terms of observation-minus-background RMSDs and mean differences. The major improvements are located at depth, between 250 and 1500 m, and are coming mainly from the Southern Hemisphere, with large improvements in the temperature statistics of the South Pacific also being highlighted in Figure 4. This is also the case for the South Atlantic and the Southern Ocean (see section 5.1). However, applying large-scale DA corrections for the unbalanced salinity is still needed to prevent near-surface drifts, which is also highlighted by the salinity statistics in the South Pacific. Since there are many fewer salinity than temperature observations available for assimilation, the large-scale DA corrections should play a very important role to constrain salinity biases in the model. It is also worth emphasising that the unbalanced salinity is not used in the density computation

that is fed to the dynamic height relationship for computing the SSH balance. Therefore, the SLA signal does not project onto unbalanced salinity, which explains why there are no issues when using a long length-scale for salinity background errors.

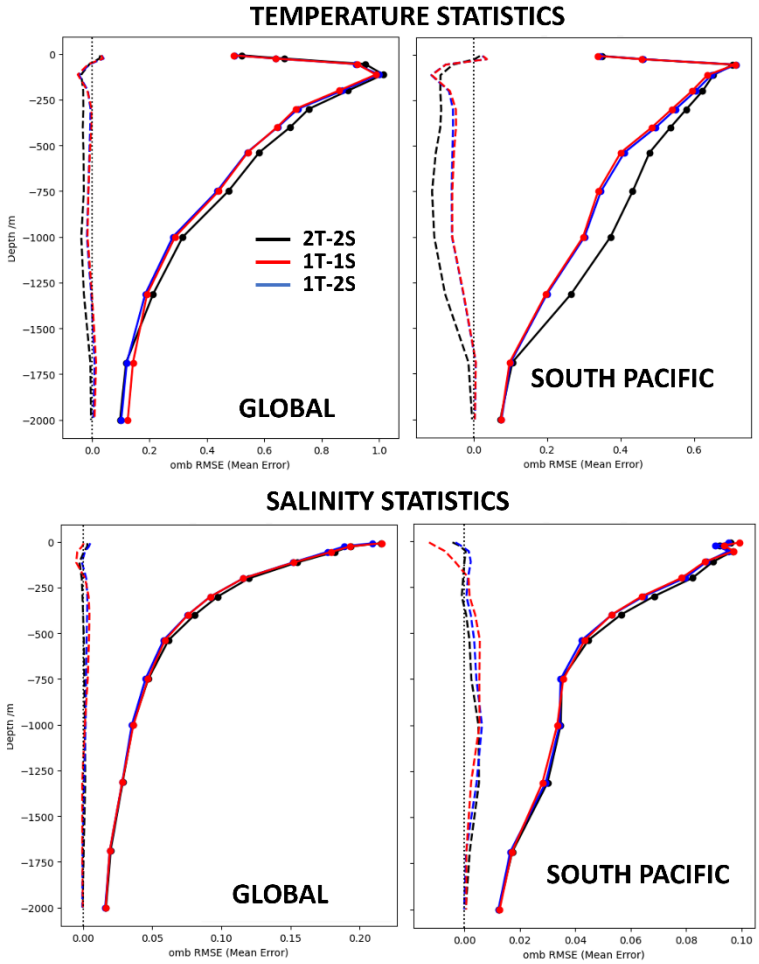

Figure 4: Observation-minus-background RMSD (solid) and mean difference (dashed) for temperature (top) and salinity (bottom) of GOSI9 ORCA025 runs with respect to T/S profile observations. The black line is the run using both short and long length-scales for T and S (2T-2S), the red line is the run using only short length-scales for T and S (1T-1S), and the blue line is the run using only short length-scales for T but both short and long length-scales for S (1T-2S). The statistics were calculated for the period between January and June 2018.

During the GOSI9 implementation, it was also found that 40 inner loop iterations in NEMOVAR were not always enough to guarantee convergence of the 3DVAR cost function, which is minimised iteratively using a B-preconditioned conjugate gradient algorithm (Gürol et al., 2014). This was particularly true in high-dynamic regions, such as the western boundary currents and the ACC. Therefore, the number of inner loop iterations has increased from 40 in GO6 to 120 in GOSI9 for the ocean DA component, which resulted in better SLA RMSDs

(Figure 5), with the largest improvements in ORCA025 coming from the North Atlantic, South Pacific and the Southern Ocean. Although there were clear benefits in increasing the number of inner loop iterations for the SLA statistics, NEMOVAR running time almost doubled, going from an average of ~13 min in GO6 to ~21 min in GOSI9. However, the implementation of multigrid assimilation (see section 4.5) enabled the diffusion operator in NEMOVAR to work with a coarser grid resolution for the long length-scale component of the background error covariances, compensating for the increase in run time caused by a higher number of inner loop iterations in GOSI9.

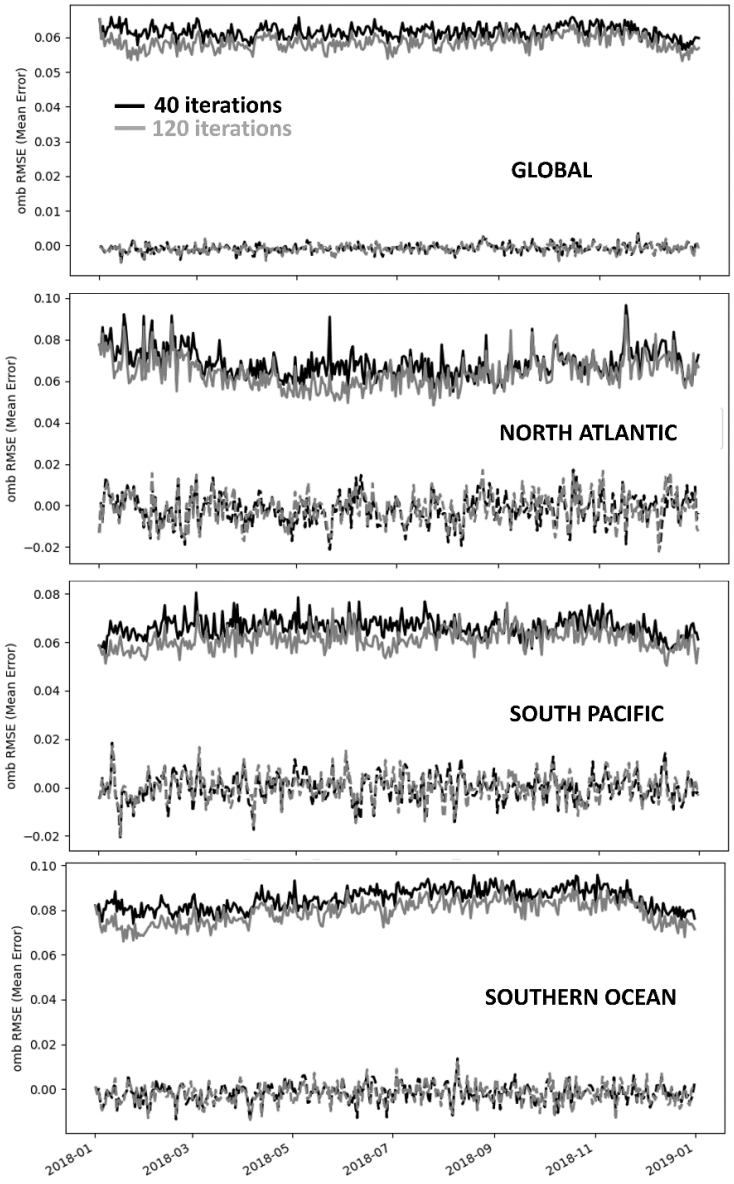

Figure 5: Observation-minus-background RMSD (solid) and mean differences (dashed) for SLA (m) of GOSI9 ORCA025 runs with respect to along-track data from all altimeters available in 2018. The black and grey lines are the runs using 40 and 120 inner loop iterations, respectively.

Another change impacting the SLA assimilation is that the SSH balance is being applied in

GOSI9 throughout the whole water column, whereas in GO6 the SSH balance is not applied within the mixed layer. As shown in Figure 6, this change in the SSH balance, along with the other DA improvements mentioned above, resulted in much better results in the Mediterranean Outflow region for ORCA12 SLA. Although most of the figures in this section correspond to ORCA025 runs, it is important to highlight that the same GOSI9 DA changes have also been applied to ORCA12, and more detailed results for both grid configurations will be presented in section 5.1. Additionally, most of the GOSI9 improvements with respect to GO6 are related to the DA changes presented here (see Appendix A).

It is also worth mentioning that GOSI9 DA changes also include updating the MDT used for SLA assimilation to a more recent version. Since this MDT update is currently being tested, the experiments in the result section do not include this specific change and the evaluation of a new MDT product is discussed separately in Appendix B.

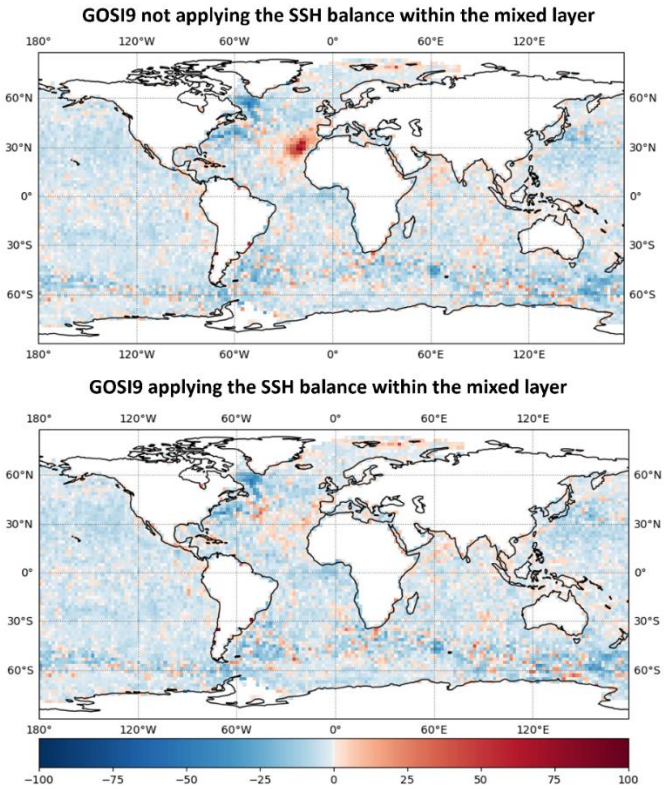

Figure 6: Observation-minus-background RMSD percentage improvement (blue) and degradation (red) in ORCA12 GOSI9 relative to ORCA12 GO6 from January to June 2019. The top plot already has all the DA changes mentioned in this section, except for the SSH balance being extended within the mixed layer, which is included in the bottom plot.

4.3 Brunt-Vaisala verifications for T/S increments

Although the DA changes in section 4.2 resulted in major FOAM improvements (see section 5.1), the fact that a long length-scale is used in GOSI9 to horizontally spread the salinity, but

not the temperature information, may exacerbate water column instabilities in very sensitive regions to T/S increments, such as in deep convection locations, which were already present in GO6.

Therefore, changes were made in the IAU, so that Brunt-Väisälä buoyancy frequencies (N^2) are computed using salinity and temperature fields after their respective increments are added on each time-step of the IAU. T/S increments for the whole water column are then rejected at each grid point if any N^2 value indicates water column instability (i.e., $N^2 < 0$) within a specified depth range. In order to target deep convection areas, the depth range for the Brunt-Väisälä checks was set to be between 400 and 1500 m globally, except for the Mediterranean Outflow region, where the initial depth was chosen to be 150 m. In the Mediterranean Outflow region, a shallower depth range was chosen to avoid water column instabilities starting at the base of the mixed layer.

Figure 7 shows that the temperature field at 1200 m in GOSI9 becomes much less noisy in the Labrador Sea when T/S increments are physically constrained by N^2 values, relative to GO6 and GOSI9 runs where no Brunt-Väisälä checks are applied. In fact, patches of colder water appearing at 1200 m in GO6 and GOSI9 runs with no Brunt-Väisälä checks seem to result from deep convection being triggered by T/S increments which make the water column unstable. The impacts provided by the Brunt-Väisälä verifications of T/S increments in the Labrador Sea are consistent with the improvements seen in the RMSD statistics for this same region in section 5.1.

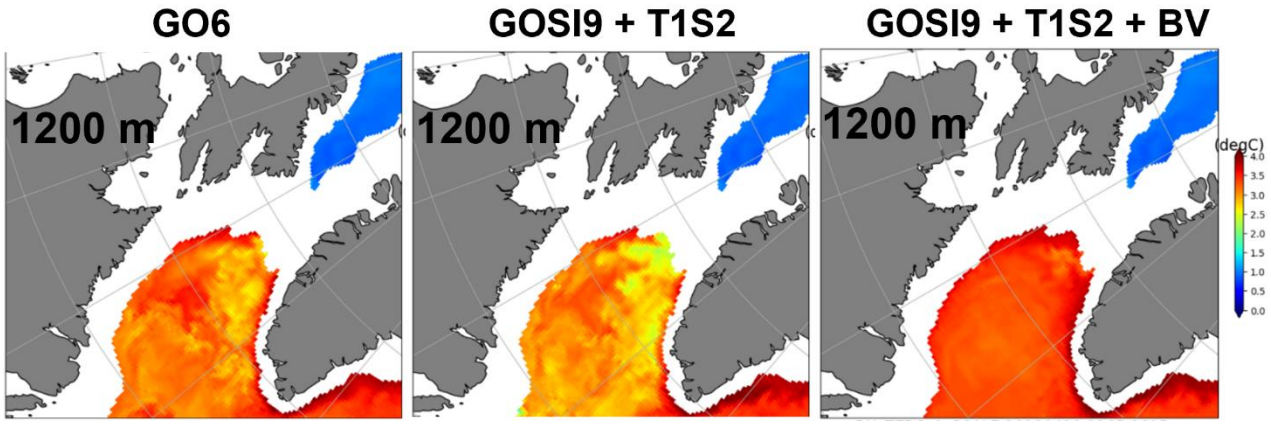

Figure 7: Temperature fields (°C) at 1200 m in the Labrador Sea for GO6 (left), GOSI9 removing the large-scale DA corrections in temperature (middle) and additionally applying Brunt-Väisälä verifications on T/S increments (right). This is a snapshot of 31st of March 2019 after all runs have been initialised on the 1st of January 2019 from the same initial condition.

As mentioned before, applying Brunt-Väisälä checks at shallower depths in the Mediterranean Outflow region also mitigates water column instabilities in GOSI9, particularly in the ORCA12 configuration. Figure 8 highlights the mean T/S profiles, averaged over the Mediterranean Outflow region, for GO6 and GOSI9 ORCA12 runs. In the GOSI9 ORCA12 run without Brunt-Väisälä checks, the T/S profiles in this region are much less stratified near the surface because of recurrent shallow water column instabilities. This leads to worse T/S statistics, between 150

and 500 m, which are then reverted and even improved with respect to GO6 when the Brunt-Väisälä checks are applied for these depths in GOSI9 (Figure 9).

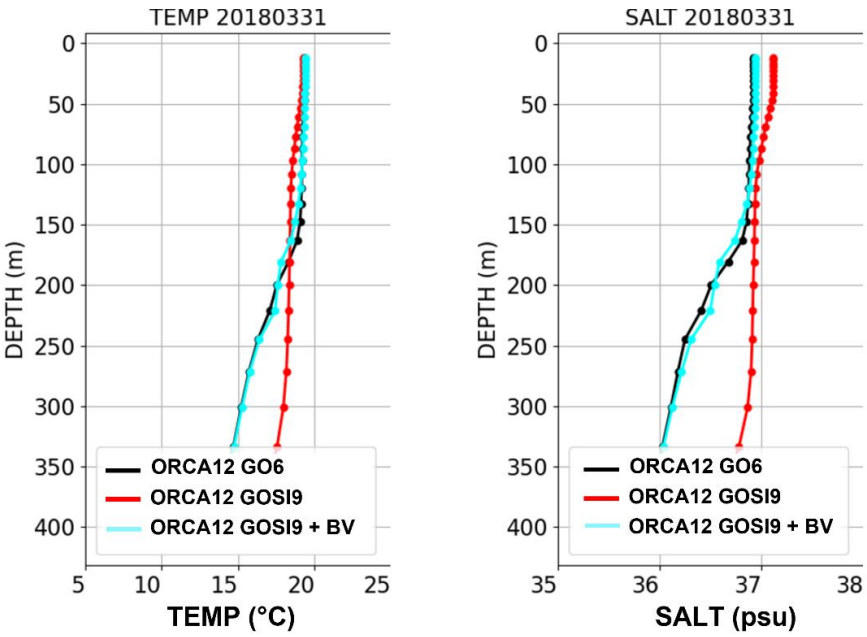

Figure 8: T/S profiles averaged over the Mediterranean Outflow region for ORCA12 runs on the 31st of March 2019. The black, red and cyan lines correspond to ORCA12 GO6, ORCA12 GOSI9 and ORCA12 GOSI9 applying the Brunt-Väisälä verification on T/S increments. All runs have been initialised on the 1st of January 2019 from the same initial condition.

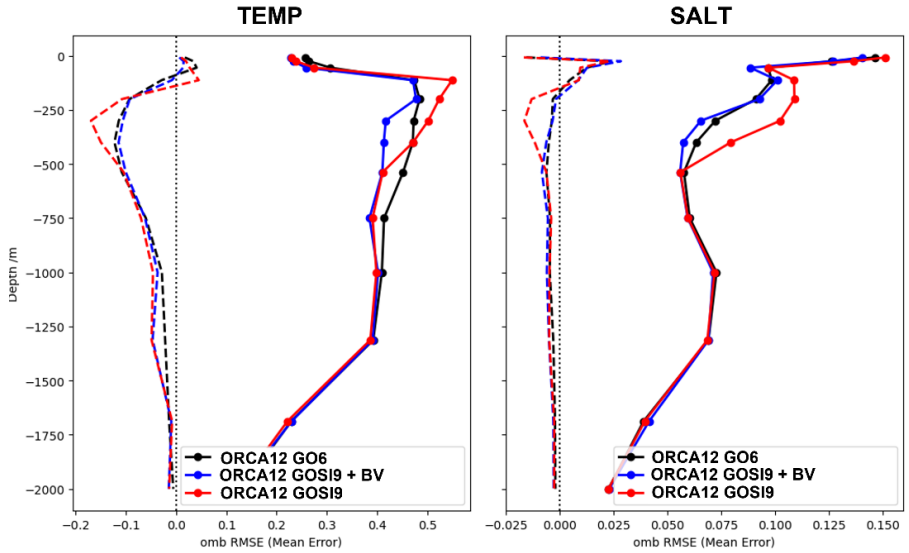

Figure 9: ORCA12 T/S innovation statistics for the Iberia-Biscay-Ireland region, which also includes the Mediterranean Outflow, for the period between January and March 2019. The black, red and blue lines correspond to ORCA12 GO6, ORCA12 GOSI9 and ORCA12 GOSI9 applying the Brunt-Väisälä verification on T/S increments.

4.4 Decomposed normalisation factors

The background error correlations in NEMOVAR are modelled using an implicit diffusion operator (Weaver et al., 2016). This method requires a normalisation of the resulting matrix to ensure that the diagonal elements are equal to one. It is prohibitively expensive to calculate the exact normalisation factors for the full 3D fields for a realistic ocean model domain, such as ORCA025 and ORCA12, and therefore a randomisation method (Weaver et al., 2020) is generally used. This requires the diffusion operator to be applied to 3D fields of random vectors multiple times (5000 random field members are used in our system). Since the normalisation factors depend on the background error correlation length-scales, any changes to these parameters requires a costly re-calculation of the normalisation factors. In practice, this makes it very expensive to use flow dependent correlation length-scales in NEMOVAR. The normalisation factor look-up table (Waters et al., 2015) was a pragmatic solution to allow flow-dependent vertical correlation length-scales in FOAM. The vertical length-scales depend on the local mixed layer depth while the horizontal length-scales are fixed. The look-up table contains the 3D normalisation factors for 50 discrete mixed layer depths and only needs to be calculated once (or whenever the horizontal length-scales change). For a particular cycle of NEMOVAR, the mixed layer depths are calculated from the model fields for that day and then at each horizontal location a profile of normalisation factors is extracted from the field in the look-up table with the corresponding mixed layer depth. While this allows us to have flow-dependent vertical scales, it is still very expensive to perform the initial calculation of the look-up table (particularly for ORCA12) and it uses a significant amount of input/output during the running of NEMOVAR.

Weaver et al. (2020) proposed an alternative approach where the calculation of the horizontal and vertical normalisation factors can be separated, which we refer to as the decomposed normalisation factors. This allows us to calculate a single 3D field of normalisation factors for the horizontal correlations offline. This is calculated using the randomisation method with 5000 members and only needs to be calculated once (or whenever the horizontal length-scales are updated). We then calculate a new 3D field of vertical normalisation factors for the flow-dependent vertical scales on each assimilation cycle. In Weaver et al. (2020), the authors used the randomisation method to calculate these, but in FOAM we use the exact calculation for the vertical normalisation factors. This method is computationally affordable and reduces the input/output requirements used in NEMOVAR compared to the look-up table. Figure 10 illustrates the impact of using the decomposed normalisation factors on the global profile statistics. There is no significant change to the observation-minus-background RMSD or mean difference for temperature and salinity with the decomposed normalisation factors. Similar results are seen for SST and SSH (not shown).

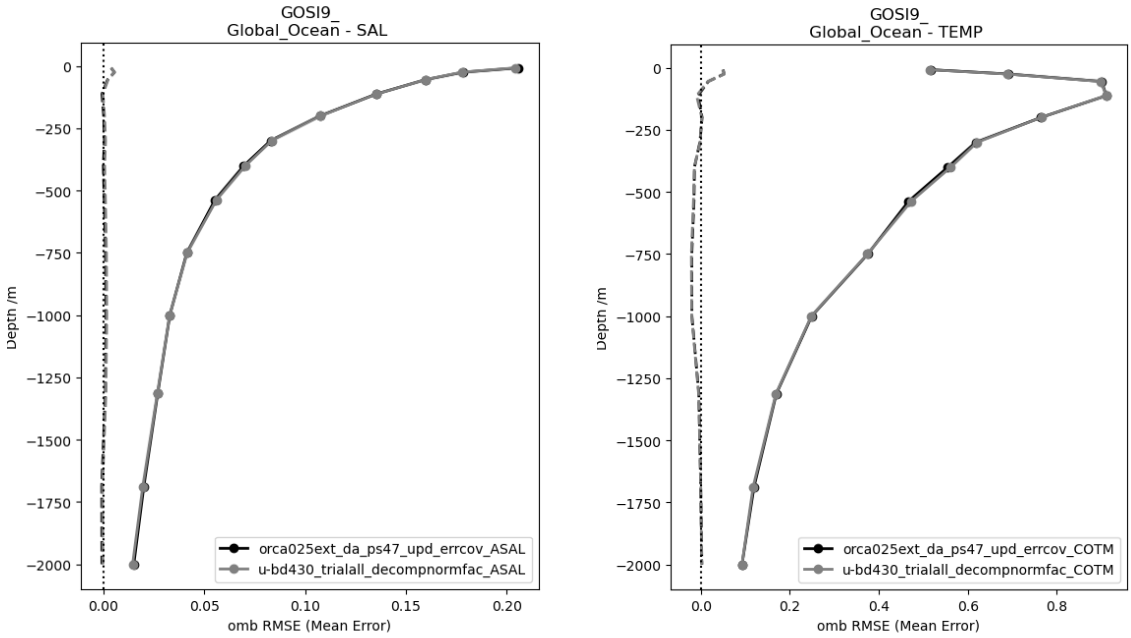

Figure 10: ORCA12 T/S global innovation statistics for 2019. The black line is an experiment using the look-up table normalisation factors and the grey line is the corresponding experiment using the decomposed normalisation factors.

4.5 Multigrid assimilation

Multigrid assimilation using the transfer grid (TRF) functionality of NEMOVAR was implemented for the first time in a Met Office system. One of the most expensive operations in NEMOVAR is the implicit diffusion calculation used for generating the spatial correlation functions, which are employed in the background error modelling. The TRF option in NEMOVAR allows this to be run on a coarser grid, with either two or three times coarsening. The disadvantage is that the correlation functions are smoother, but this is acceptable for the longer background error length-scales which are many multiples of the grid size even in ORCA025. The advantage is a 50% reduction in NEMOVAR runtime for the two times coarsening option in GOSI9, i.e. when the long length-scale of the background errors is applied to the unbalanced salinity but not to the temperature. The three times coarsening option was also tried but it does not reduce the run time further. There are fewer diffusion iterations needed on the coarser grid, but it requires many additional MPI calls to coarsen the grid near to the processor boundaries, due to bigger halo sizes, compared to two times coarsening.

The TRF code operates in three stages: (i) coarsening of the model resolution to transfer fine grid information to the coarse grid; (ii) the implicit diffusion code is run for selected covariance length-scales and variables on the coarse grid; and (iii) the result is then transferred to the fine/standard assimilation grid.

For GOSI9, the TRF was tested with decomposed normalisation factors. To avoid artifacts when using the coarse grid, especially near coasts, we filter the normalisation factors using a Shapiro filter. Comparing increments for one day (with the same background and innovations), there were some differences in increments near the coasts but very little difference in the open ocean (not shown). Furthermore, no significant differences in the innovation statistics were found between 1-year runs of the GOSI9 FOAM performed with and without TRF (not shown). Currently, we have implemented this in ORCA025 for the salinity long length-scale since the temperature long length-scale was removed in GOSI9 (see section 4.2). Using TRF has also the potential to make the assimilation on the ORCA12 native grid feasible, removing the need of performing ORCA12 DA at ORCA025 resolution, and this will be addressed in the future.

4.6 Summary of changes between GOSI9 and GO6

	GO6	GOSI9
Ocean model	NEMO3.6	NEMO4.04
Sea-ice model	CICE	SI^3
Equation of state	EOS80	TEOS10
SST and SLA observation errors	Seasonally and spatially varying estimates produced using the Hollingsworth and Lonnberg method. This is treated as the total error.	Seasonally and spatially varying representation error due to unresolved scales in the model (Oke and Sakov, 2008) + observation-specific measurement error for SST and 4 cm measurement error for SLA.
Background errors	Short and long length-scales are used for both T and S	Long length-scale is not applied for T but is used for S
Inner loop iterations for ocean DA	40	120
SSH balance	Applied below, but not in the mixed layer	Applied through the whole water column
Quality control of T/S increments	None	Rejection of T/S increments based on water column instabilities diagnosed from Brunt-Väisälä buoyancy frequencies
Normalisation factors	Look-up table	Decomposed normalisation factors
TRF	Not implemented	Implemented for ORCA025 DA
MDT	CNES-CLS13 (CNES-CLS18) in the research (operational) suite	Expected to change to CNES-CLS22
ORCA025 NEMOVAR run time	13 minutes using 15 computational nodes	13 minutes using 15 computational nodes

Table 1: List summarising all the model and DA changes between GOSI9 and GO6. Although listed in the table, the MDT update is still under evaluation and is not included in the result section. MDT updates are discussed separately in Appendix B.

5. Results

5.1 GOSI9 and GO6 comparisons

In this section, observation-minus-background statistics for T/S, SLA, SST and SIC, such as the RMSD and mean differences, are shown comparing GO6 and GOSI9 configurations for both ORCA025 and ORCA12. To better quantify the GOSI9 impacts for each variable with respect to GO6, the observation-minus-background RMSD for both GOSI9 and GO6 are compared using the equation below:

$$\alpha = \left(\frac{RMSD_{GOSI9} - RMSD_{GO6}}{RMSD_{GO6}}\right) * 100 \qquad (1)$$

where α represents an RMSD improvement (degradation) percentage ratio of GOSI9 relative to GO6 in case α is negative (positive). The statistics correspond to 2019, which is the year that trial runs have been conducted for both FOAM GO6 and GOSI9. The initial condition for both FOAM GO6 and GOSI9 is the same, which comes from a previous GO6 run, as well as the atmospheric forcing, which comes from the Met Office operational NWP system.

It is worth mentioning that the temperature and salinity RMSD results for GO6 and GOSI9 are calculated from the EOS80 and TEOS10 variables, respectively. The magnitude of the errors is expected to be consistent whether using TEOS10 or EOS80. We investigated the impact of converting between absolute and practical salinity on the observation-minus-background values and found that it has a very small impact of the order of 0.001, which is much smaller than the salinity differences and RMSDs between GO6 and GOSI9 presented here.

5.1.1 ORCA025

The SLA statistics are clearly improved in GOSI9 compared to GO6 for ORCA025. Figure 11 shows that there is a global reduction in the RMSD from 0.063 m in GO6 to 0.056 m in GOSI9, with the mean differences remaining largely unaffected relative to the SLA observations. RMSD improvements are more significant from April 2019 onwards, indicating that there is a spin-up period of ~3 months for the SLA to adjust to the new DA configurations in GOSI9.

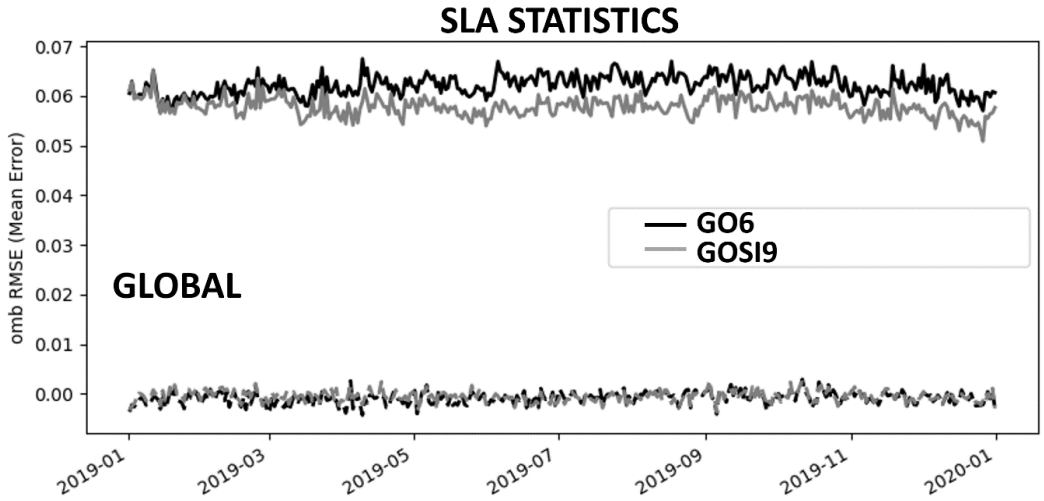

Figure 11: 2019 ORCA025 observation-minus-background SLA statistics (m) for GO6 (black) and GOSI9 (grey), calculated against along-track altimeter observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

Except for the Mediterranean outflow and a few coastal regions, SLA RMSD improvements of ~20% are found in GOSI9 relative to GO6 in all ocean basins (Figure 12). It is worth highlighting that the most significant SLA RMSD improvements in GOSI9 are found within western boundary current regions, such as the Gulf Stream and Kuroshio Current, and along the ACC path, which can reach up to a 50% RMSD improvement.

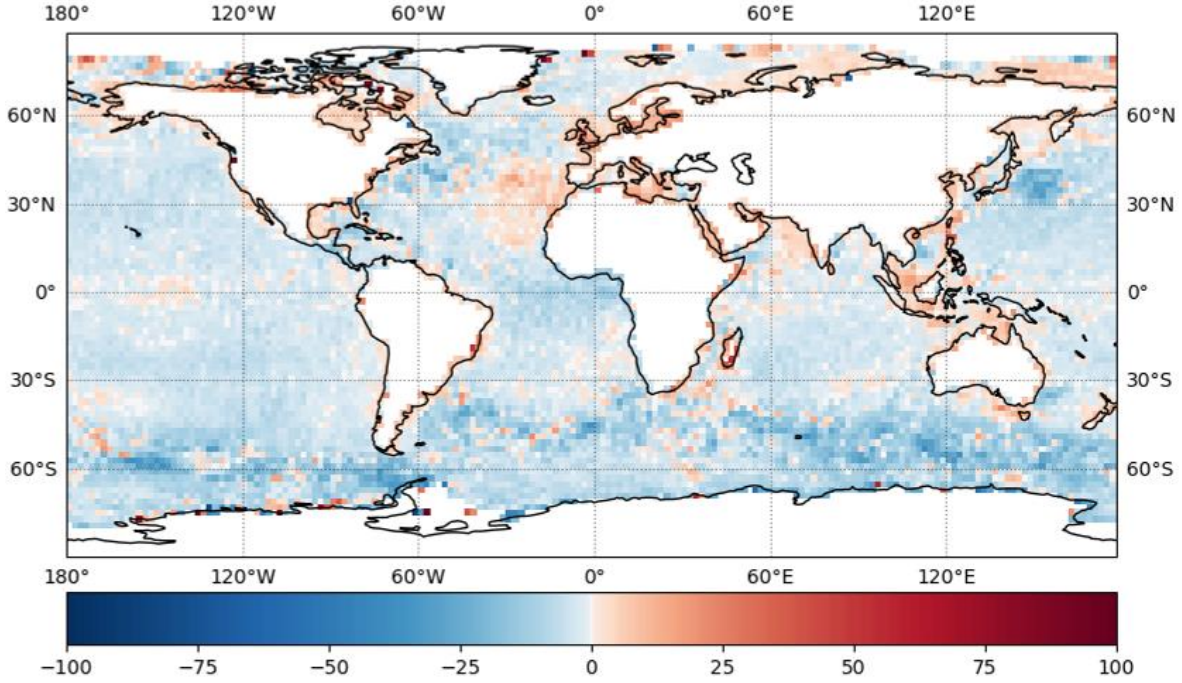

Figure 12: SLA RMSD improvement (blue) and degradation (red) of ORCA025 GOSI9 relative to GO6 for 2019, shown as percentages. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

On a global scale, the SST improvements in ORCA025 GOSI9 relative to GO6 are fractionally smaller than the SLA improvements. As shown in Figure 13, daily RMSDs between GOSI9

and GO6 are very similar globally, although slightly lower in GOSI9 throughout the year. However, when looking at the spatial variation of RMSD improvements (Figure 14), GOSI9 SST enhancements are consistently found in all ocean basins, particularly in the tropics, reaching RMSD improvements of about 50% with respect to GO6. Slight improvements in GOSI9 SSTs are also consistently seen when compared to surface temperature measurements from profile observations (Figure 15).

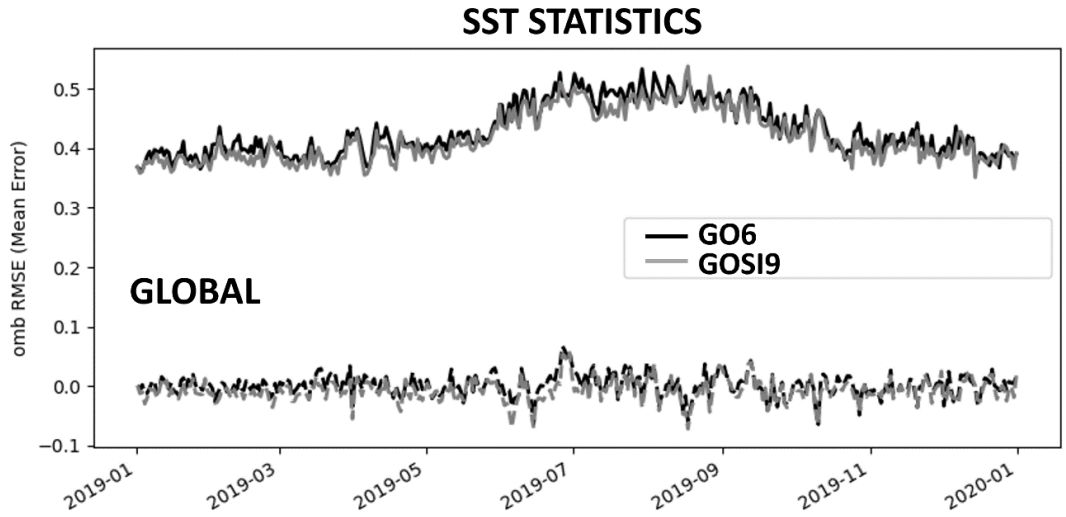

Figure 13: 2019 ORCA025 observation-minus-background SST statistics (°C) for GO6 (black) and GOSI9 (grey), calculated against in situ SST drifters. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

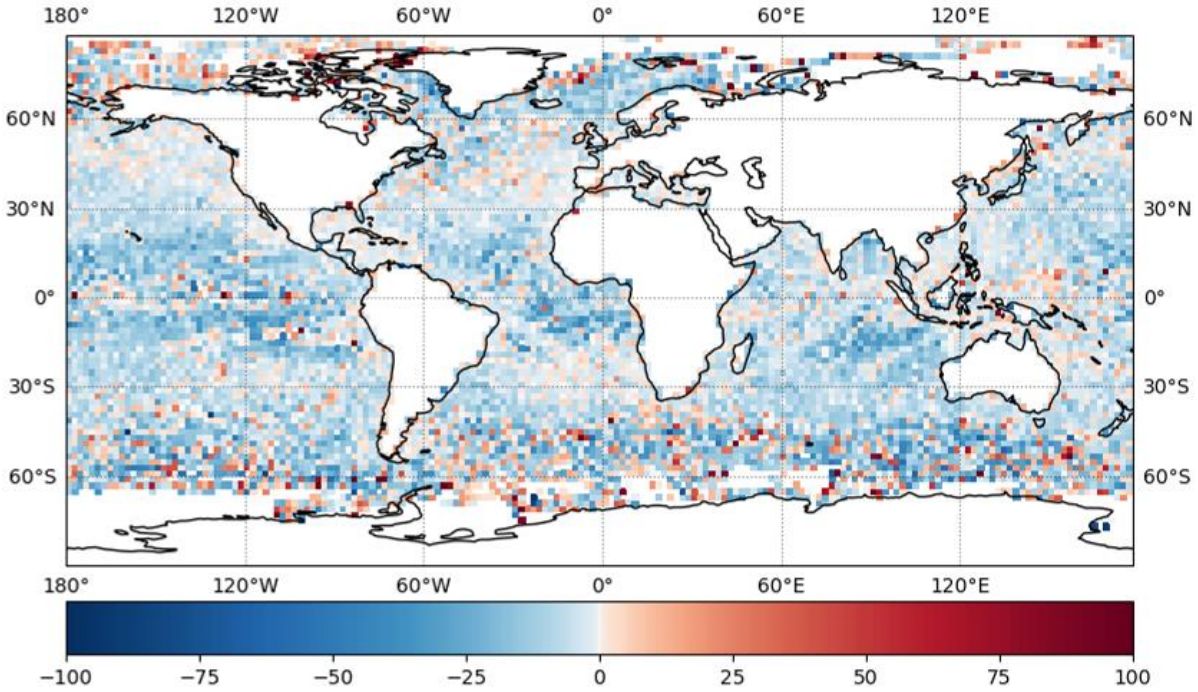

Figure 14: SST RMSD improvements (blue) and degradations (red) of ORCA025 GOSI9 relative to GO6 for 2019, shown as percentages. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

The GOSI9 DA changes in ORCA025 have led to a significant decrease in the global RMSDs and biases of the model temperatures at depth, particularly between 250 and 1500 m, relative to profile observations when compared to GO6 (Figure 15). These improvements are largely driven by the substantial impacts of the GOSI9 changes in the Southern Hemisphere, particularly in the South Pacific, South Atlantic and Southern Ocean basins. In the Southern Ocean and South Pacific, the GOSI9 temperature RMSDs are halved when compared to those from GO6 at depth. This reinforces the idea that removing large-scale DA corrections for temperature leads to improvements in how the SLA observation information projects onto model temperatures at depth in GOSI9 (see also Figure 4). Despite the largest impacts being in the Southern Hemisphere, improvements in the T statistics are also noted in the Arctic, North Atlantic and North Pacific (Figure 15), as well as in marginal seas (see Appendix C), especially for the RMSD statistics. Looking at the globally averaged RMSD improvements in temperature as a function of depth relative to GO6 (Figure 16), it is also clear that these GOSI9 improvements gradually increase throughout 2019 for depths between 500 and 1500 m.

For salinity, GOSI9 improvements relative to GO6 are not so significant as for temperature (Figure 17). It is worth emphasising that a long length-scale is still used to propagate the unbalanced salinity increments in GOSI9, as done in GO6. Although the global salinity statistics are only slightly better in GOSI9, a few regions stand out showing a consistent RMSD decrease throughout the water column, such as the South Pacific and the Southern Ocean, where most significant GOSI9 impacts are also seen for temperature. This indicates that removing the large-scale DA corrections for temperature has an indirect effect on improving the sub-surface salinity structure in the Southern Hemisphere. However, even though RMSDs are improved in the Southern Hemisphere, the salinity biases slightly increase in GOSI9 when compared to GO6 in the same region, especially near the surface. The globally averaged RMSD improvements in salinity, shown as a function of depth (Figure 18), reinforces that GOSI9 improves the sub-surface salinity in 2019 when compared to GO6. However, unlike temperature, the sub-surface salinity RMSD improvements are steady, not showing any clear trends throughout the year.

In addition to salinity improvements in the Southern Hemisphere, smaller but positive impacts in GOSI9 salinities are also seen in the North Atlantic and North Pacific (Figure 17), as well as in coastal regions (see Appendix C), particularly in the sub-surface RMSDs.

As already mentioned in section 3.1, it is worth emphasising that the model changes only contribute marginally to the improvements seen in GOSI9, except for the SLA improvements in the Southern Ocean. The DA changes (see section 4.2) are the ones responsible for most of the GOSI9 RMSD decrease for SLA, SST and T/S with respect to GO6. See Appendix A for a more detailed breakdown of the impacts between model and assimilation changes on GOSI9.

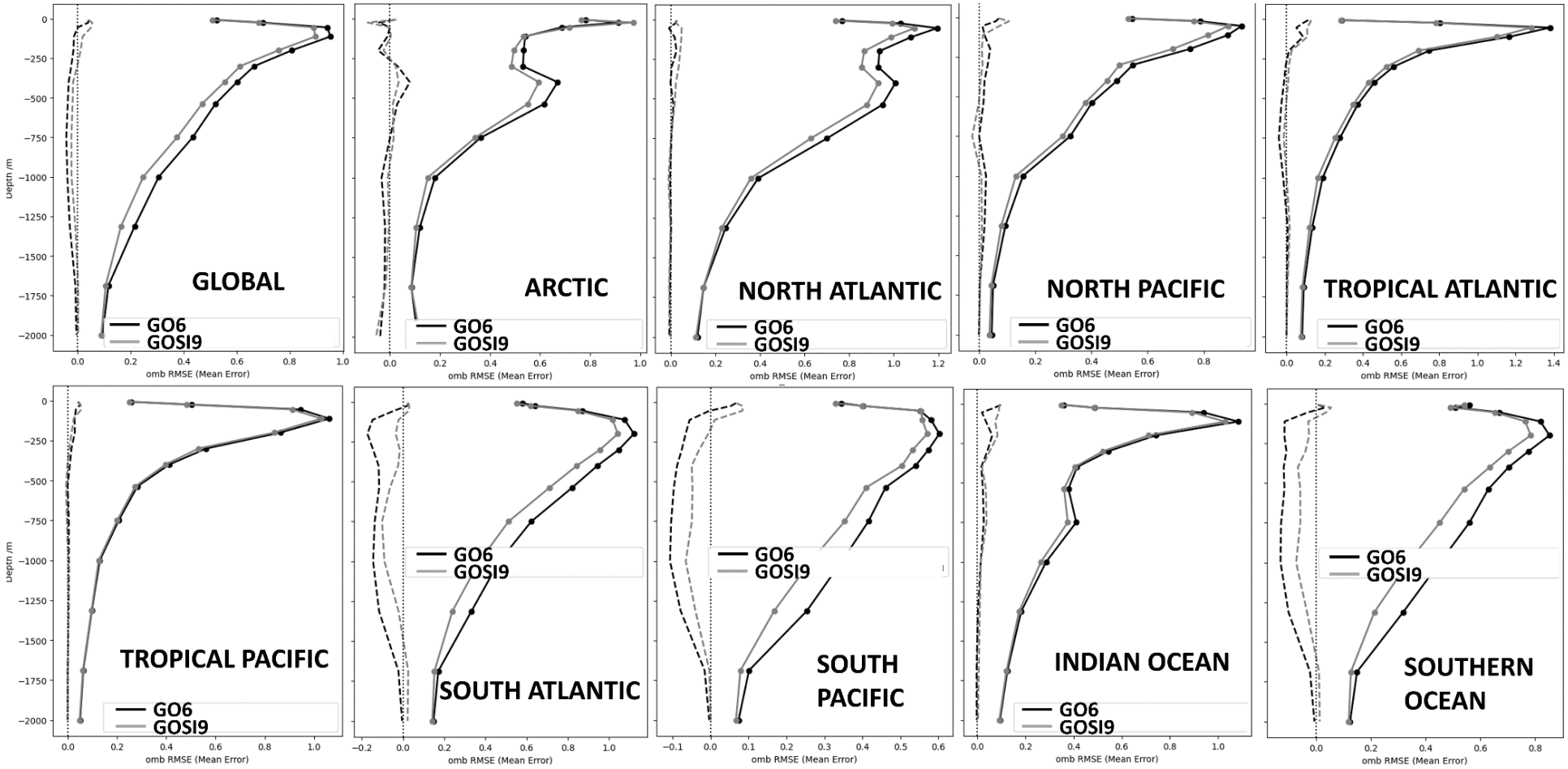

Figure 15: 2019 ORCA025 observation-minus-background temperature statistics (°C) for GO6 (black) and GOSI9 (grey) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

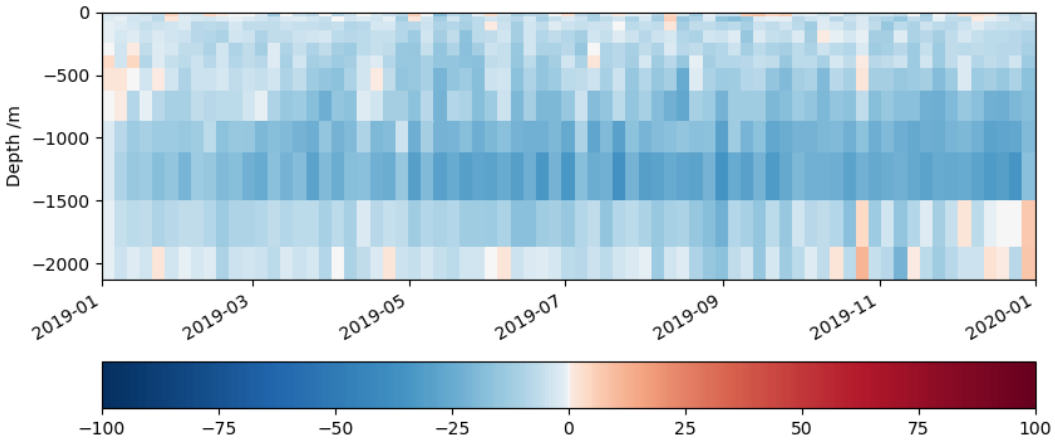

Figure 16: Hovmöller of RMSD temperature improvements (blue) and degradations (red) of ORCA025 GOSI9 relative to GO6, globally averaged as a function of depth. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

In terms of sea ice, there is a significant reduction of SIC RMSDs in GOSI9 relative to GO6 between July and September, when both systems are compared to satellite observations (Figure 19). A known issue in GO6 is the excessive melting of Arctic ice introduced by DA over summer, so there is less sea ice than is observed during this season. This causes a sudden increase in the RMSD and a positive observation-minus-background bias in the Arctic summer. This issue no longer happens in GOSI9, with Arctic SIC RMSDs and mean differences dropping from 0.10 and 0.05 in GO6 to 0.04 and 0.01 in GOSI9 over summer, respectively. The GOSI9 SIC RMSD improvements over Arctic summer are very consistent spatially relative to GO6, occurring not only in the ice pack, where a ~30% RMSD decrease is found, but also in regions near the ice edge (Figure 20). When SIC increments are added to the SI^3 model in GOSI9, the SIC DA changes are proportionally propagated to other model prognostic variables, so equivalent values are maintained between all prognostic variables. This might explain why the new ice introduced by DA is more resilient to melting in GOSI9, leading to clear Arctic summer SIC improvements.

Despite the significant SIC improvements in Arctic summer, GOSI9 SIC results are slightly worse than in GO6 in other Arctic seasons, as well as throughout the year in the Antarctic (Figure 19). As shown in Figure 20, SIC RMSD degradations in GOSI9 relative to GO6 are more dominant in the austral (boreal) winter for the Antarctic (Arctic) when compared to summer, spreading to most of the sea-ice covered regions. A current limitation of GOSI9 SIC assimilation is that, although a new sea-ice model (SI^3) is employed, it still uses the background error covariances derived from GO6 (i.e., from a CICE run). Therefore, this inconsistency might help explain the slightly worse performance of GOSI9 SIC assimilation with respect to GO6.

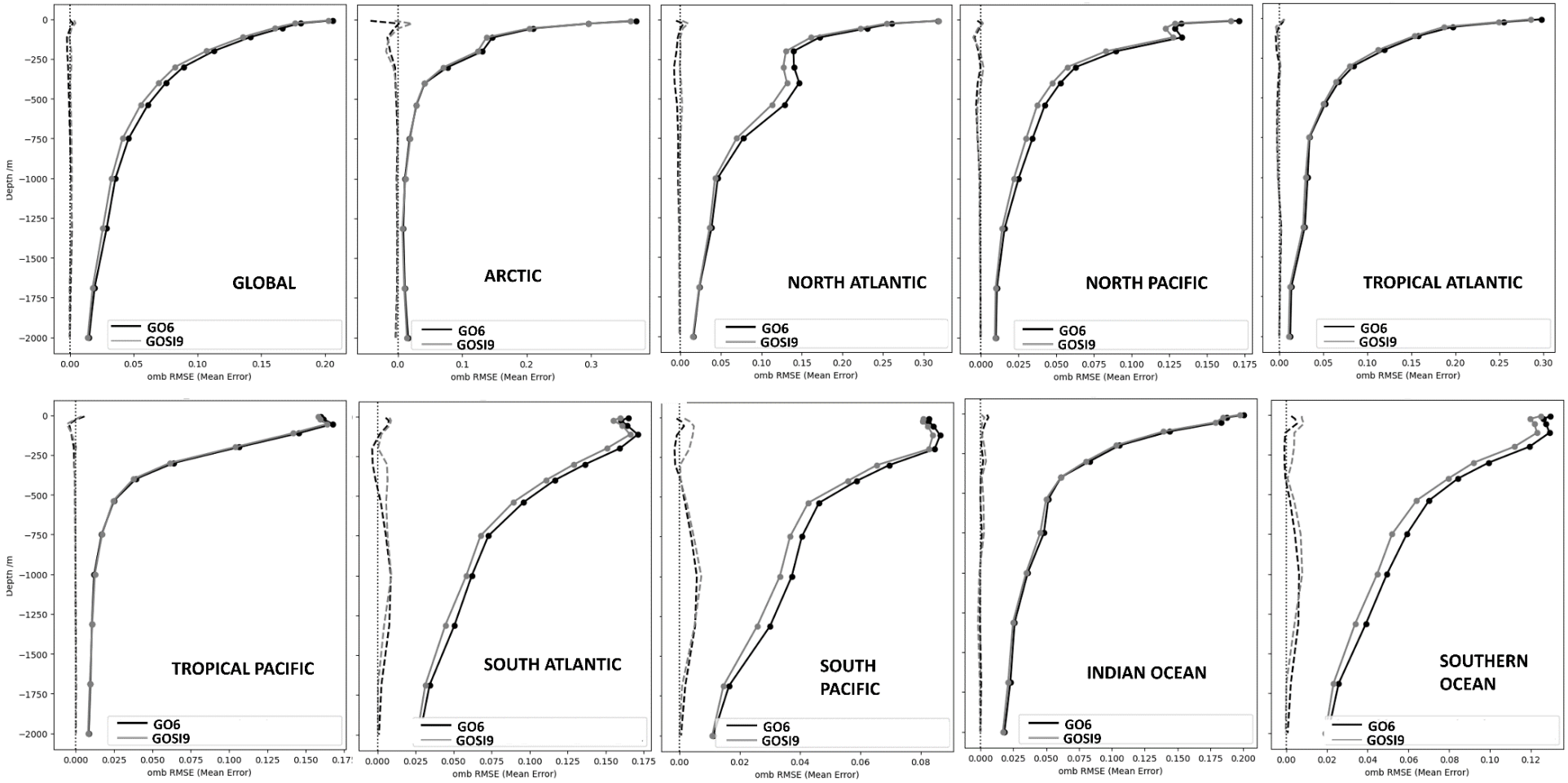

Figure 17: 2019 ORCA025 observation-minus-background salinity statistics for GO6 (black) and GOSI9 (grey) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

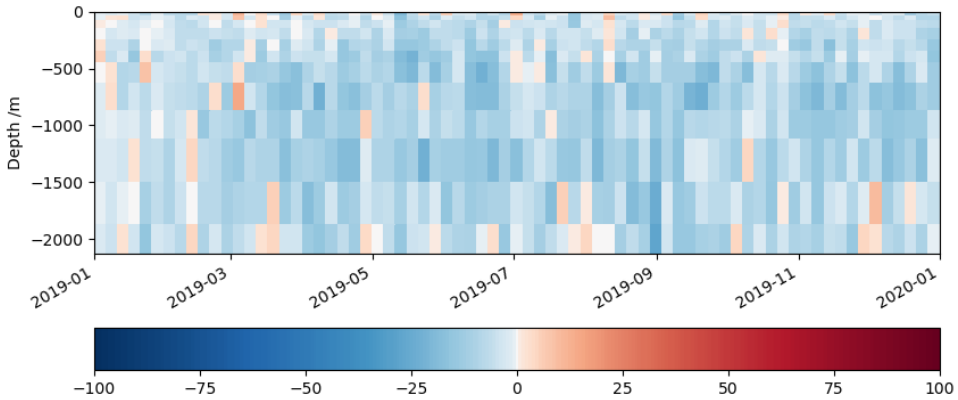

Figure 18: Hovmöller of RMSD salinity improvements (blue) and degradations (red) of ORCA025 GOSI9 relative to GO6, globally averaged as a function of depth. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

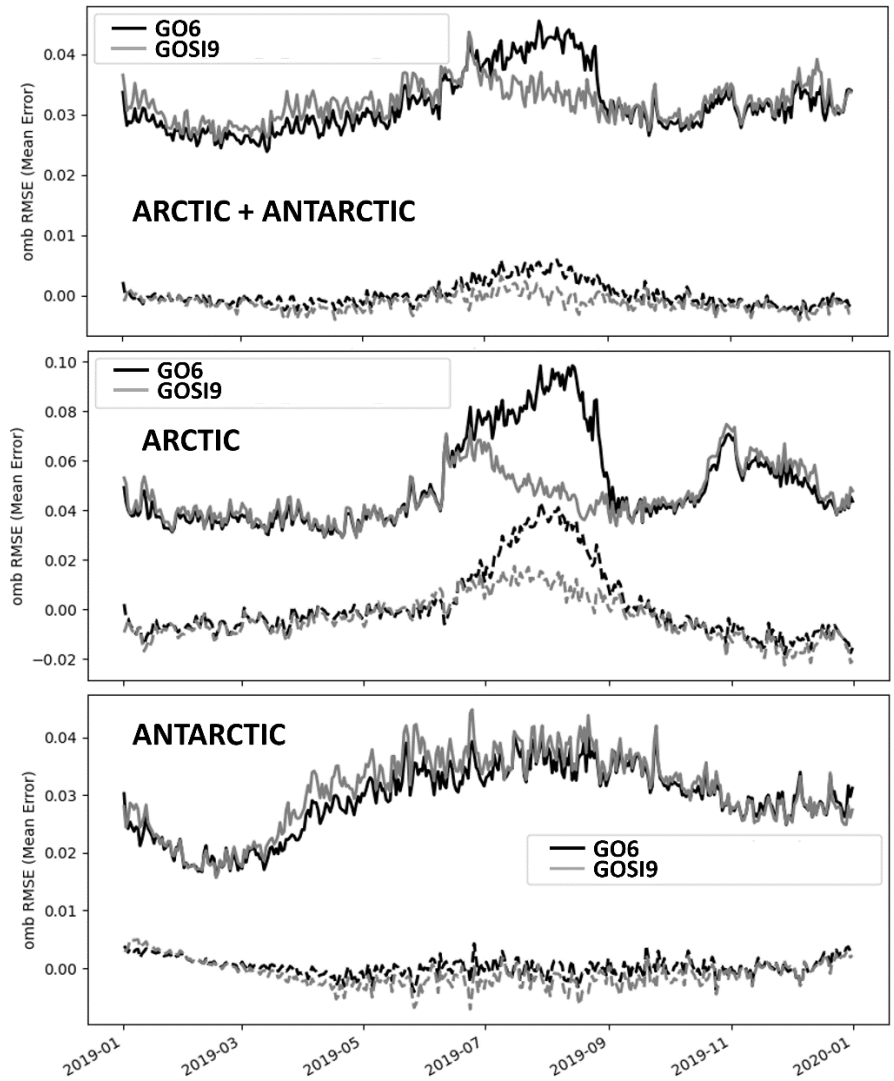

Figure 19: 2019 ORCA025 observation-minus-background SIC statistics for GO6 (black) and GOSI9 (grey), calculated against satellite observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

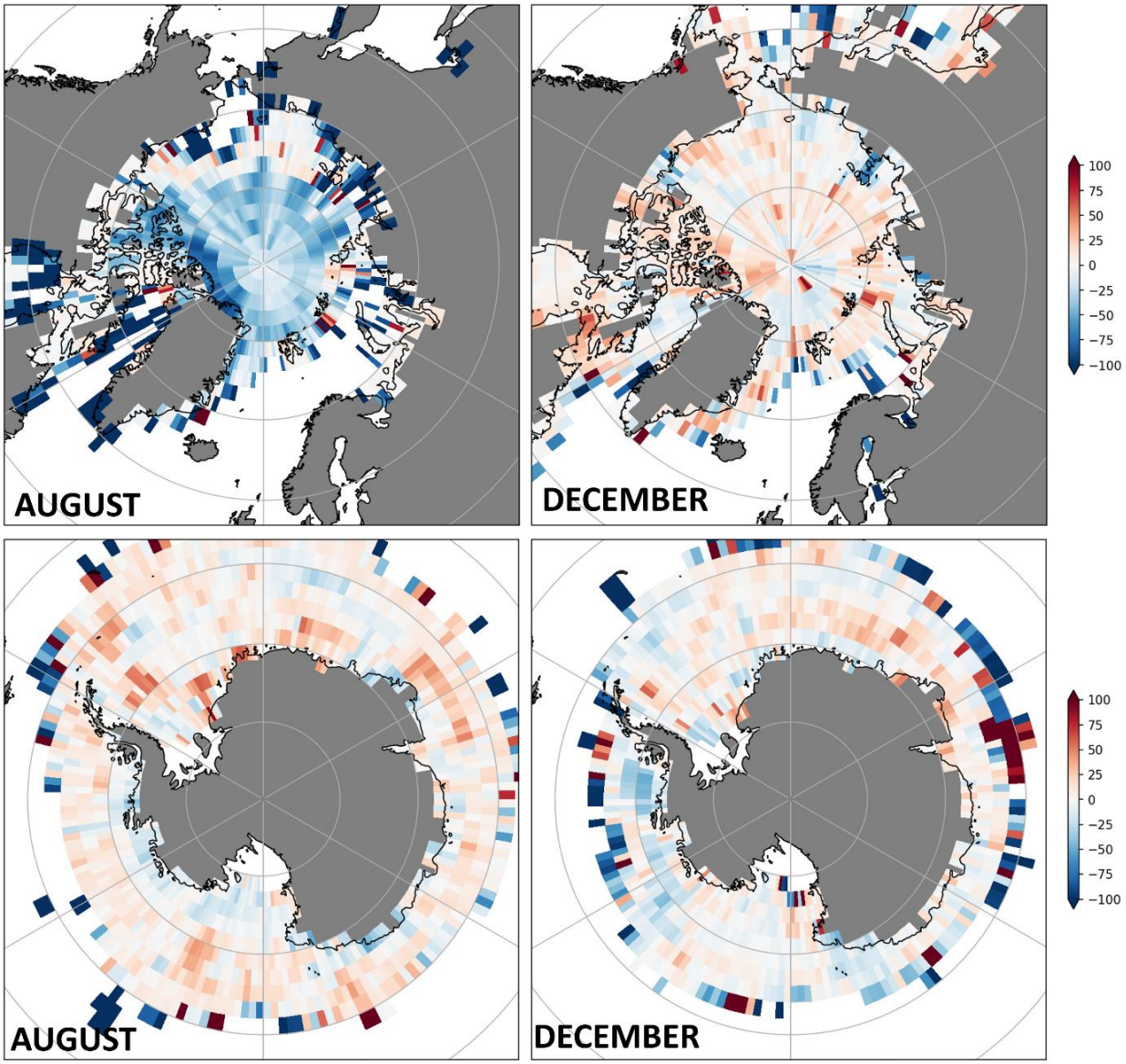

Figure 20: SIC RMSD improvements (blue) and degradations (red) of ORCA025 GOSI9 relative to GO6 for the Arctic (top) and Antarctic (bottom), shown as percentages, for August (left) and December (right). The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

5.1.2 ORCA12

As seen in ORCA025, SLAs are also clearly improved for ORCA12 in GOSI9 when compared to GO6, showing a global RMSD decrease from 0.067 to 0.058 m (Figure 21). ORCA12 SLA improvements in GOSI9 are consistent throughout the year, particularly from April 2019 onwards. This highlights that, similar to ORCA025, a spin-up of ~3 months is required for the ORCA12 SLAs to fully reflect the positive impacts caused by GOSI9 DA changes. Spatially, ORCA12 SLA RMSD improvements can easily reach up to 40% in GOSI9 relative to GO6, especially in the western boundary regions and along the ACC path (Figure 22). Although smaller than in the western boundary regions, ORCA12 SLA RMSD improvements are noted almost everywhere in GOSI9, with the main exceptions being the Mediterranean outflow, a few coastal regions, and enclosed seas.

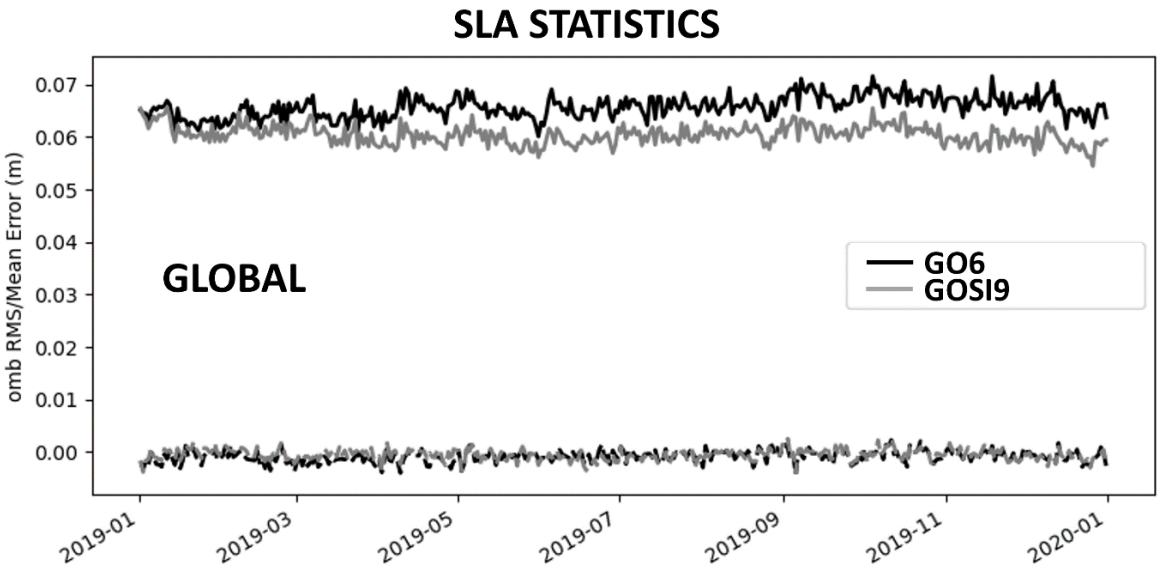

Figure 21: 2019 ORCA12 observation-minus-background SLA statistics (m) for GO6 (black) and GOSI9 (grey), calculated against along-track altimeter observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

In terms of SST statistics, GOSI9 ORCA12 performs slightly better than its GO6 version globally (Figure 23), with RMSDs decreasing for the global ocean from 0.46 to 0.42°C. These SST RMSD improvements are consistent spatially, particularly in the tropical regions, where a ~30% RMSD improvement in ORCA12 GOSI9 can be found across all ocean basins relative to GO6. However, it is worth highlighting that small RMSD degradations are seen for the SSTs in the Mediterranean Outflow and the Arctic. Although removing large-scale DA temperature corrections can lead to significant FOAM improvements, particularly in the Southern Hemisphere, it has also been shown in section 4.3 that a drawback of this change is that it can exacerbate localised water column instabilities in very sensitive regions to T/S increments, such as the Mediterranean Outflow.

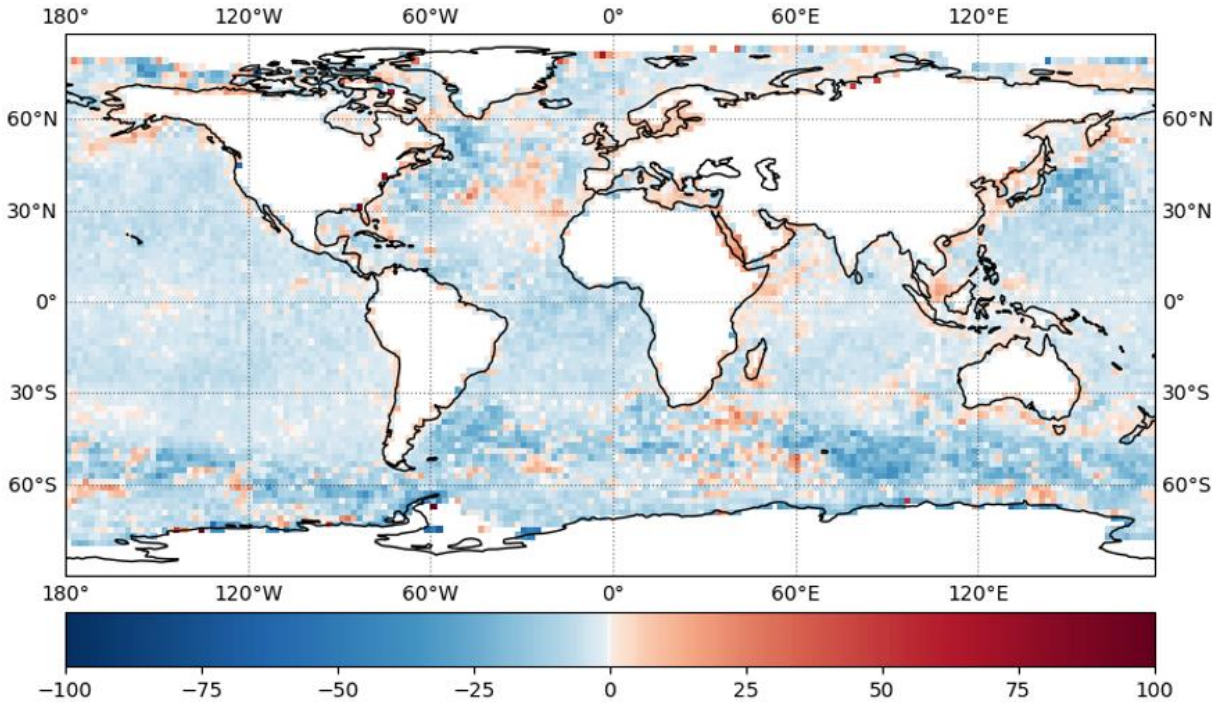

Figure 22: SLA RMSD improvement (blue) and degradation (red) of ORCA12 GOSI9 relative to GO6 for 2019, shown as percentages. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

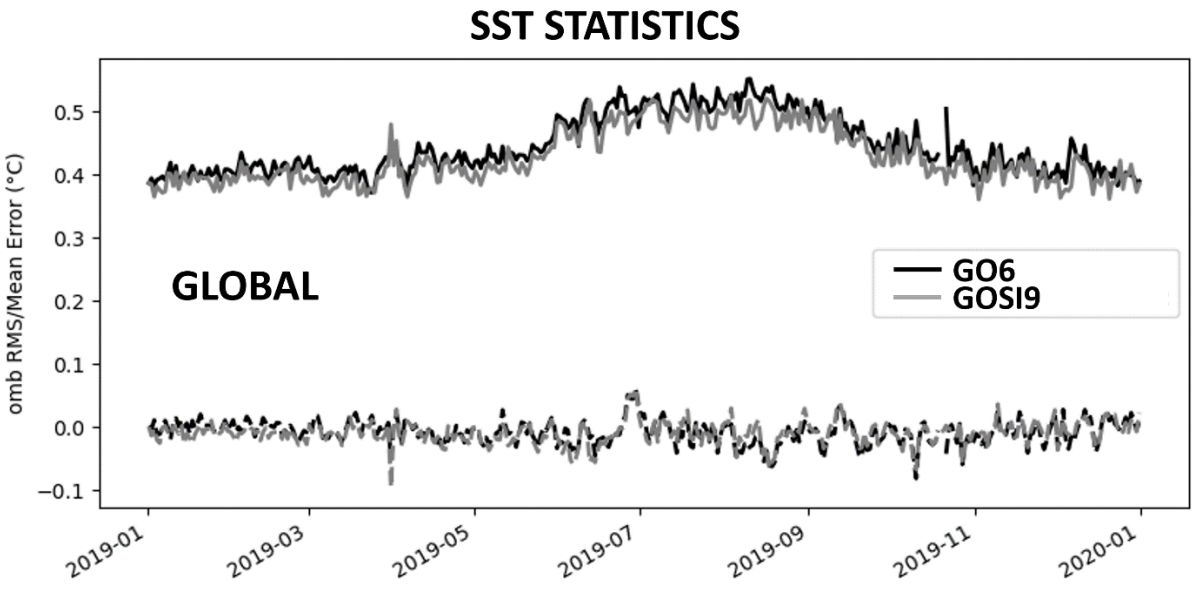

Figure 23: 2019 ORCA12 observation-minus-background SST statistics (°C) for GO6 (black) and GOSI9 (grey), calculated against in situ SST drifters. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

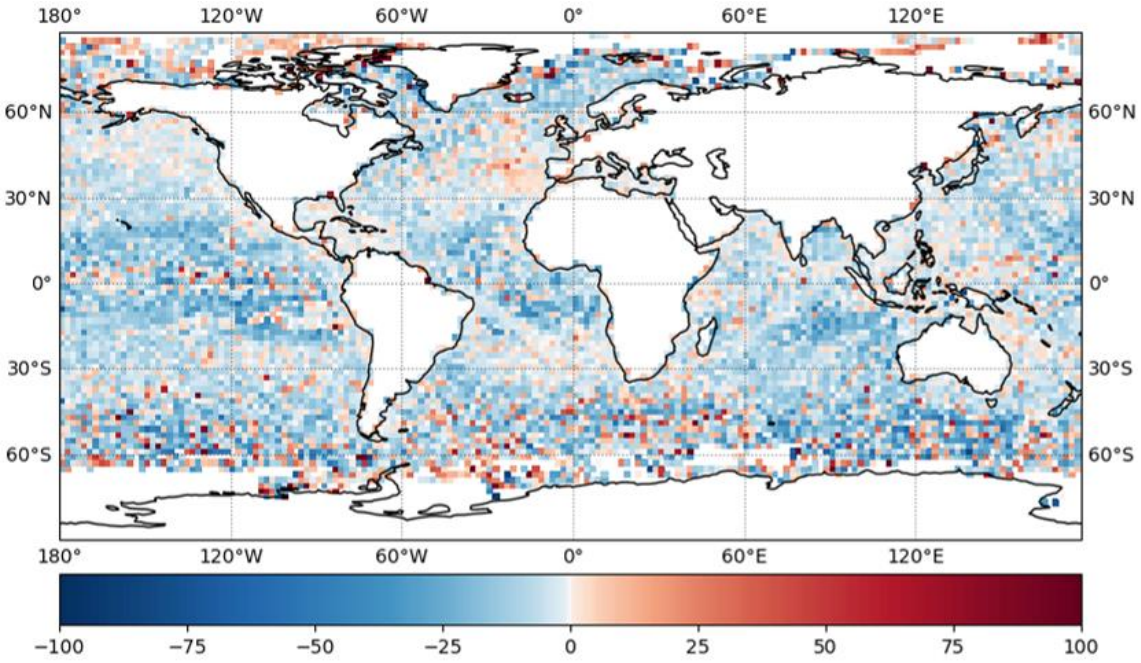

Figure 24: SST RMSD improvement (blue) and degradation (red) of ORCA12 GOSI9 relative to GO6 for 2019, shown as percentages. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

As expected, based on GOSI9 ORCA025 temperature improvements, there is also a clear positive impact of GOSI9 DA changes decreasing the RMSDs and temperature biases at depth for ORCA12 when compared to its GO6 version (Figure 25). These temperature improvements are very significant in the Southern Hemisphere, particularly in the South Pacific and Southern Ocean, but also extending to other regions, such as the North Atlantic and North Pacific, which is consistent with ORCA025 results (see section 5.1.1). Consequently, when ORCA12 temperature statistics are globally averaged, there is a consistent GOSI9 RMSD reduction for depths between 250 and 1500 m relative to GO6. For marginal seas, however, GOSI9 ORCA12 impacts on temperature are mixed, generally showing RMSD improvements at depth, but mean differences are slightly degraded with respect to profile observations (see Appendix D).

Although temperature improvements clearly increase throughout 2019 in GOSI9 ORCA025 (Figure 16), the same trend is not seen in GOSI9 ORCA12, despite consistent RMSD improvements of ~25% relative to GO6 occurring after the first quarter of 2019 for depths between 500 and 1500 m (Figure 26). It is also worth mentioning that there is a discontinuity in GOSI9 RMSD improvements between temperatures above and below 1500 m, which is noted in ORCA025 (Figure 16) but is particularly clear in ORCA12 (Figure 26). The depth of 1500 m coincides with the depth used in the assimilation for the dynamic height relationship, which relates density changes to SSH changes and vice-versa. Further investigation is needed to better understand the impacts of using different target depths for the dynamic height relationship on the T/S assimilation results, but this is beyond the scope of this report.

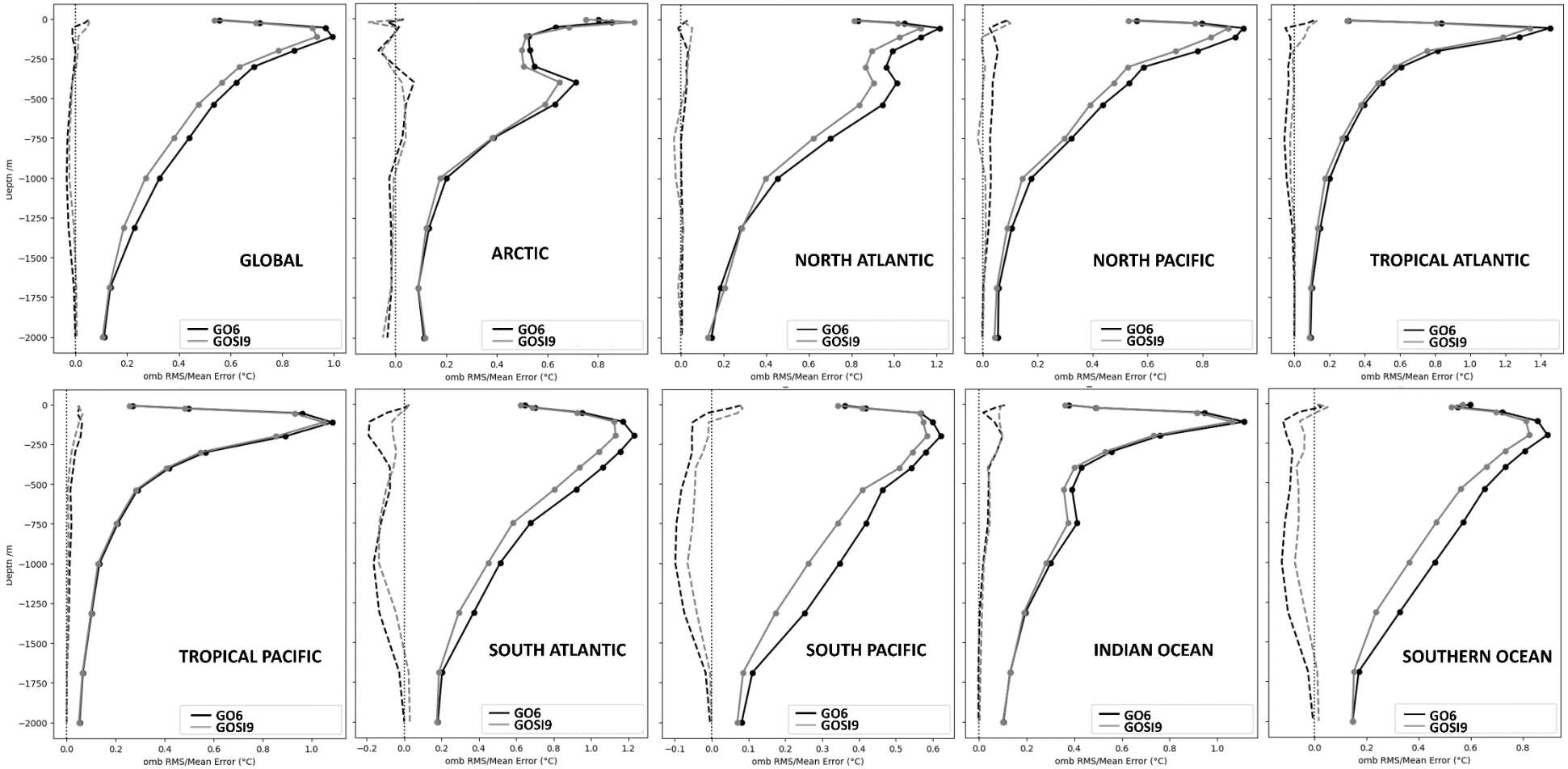

Figure 25: 2019 ORCA12 observation-minus-background temperature statistics (°C) for GO6 (black) and GOSI9 (grey) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent observation-minus-background mean differences.

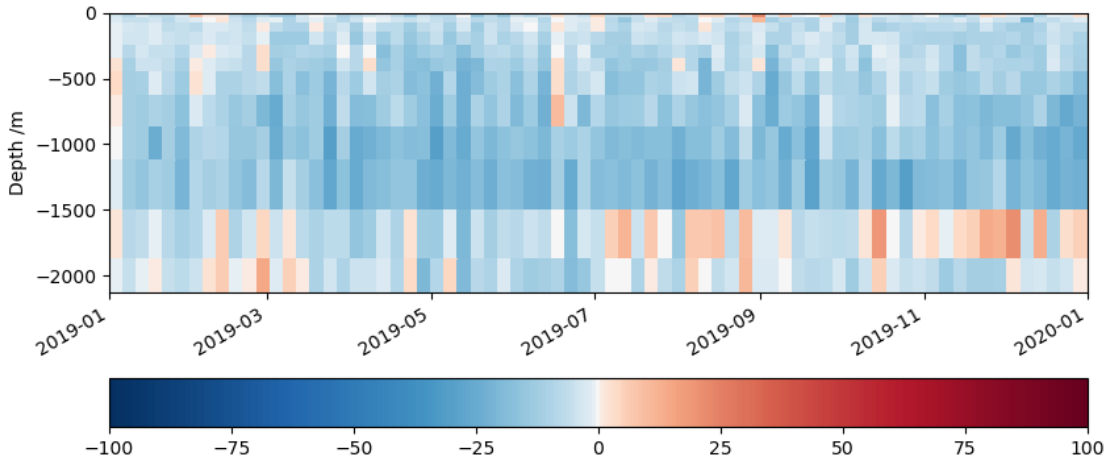

Figure 26: Hovmöller of RMSD temperature improvements (blue) and degradations (red) of ORCA12 GOSI9 relative to GO6, globally averaged as a function of depth. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

When compared to GO6, GOSI9 ORCA12 salinity improvements also occur at depth in the same regions where GOSI9 ORCA025 temperature improvements are observed, such as in the South Pacific, Southern Ocean, North Atlantic and North Pacific (Figure 27), as well as in the Mediterranean Sea (see Appendix D). However, like GOSI9 ORCA025, small near-surface salinity drifts are also seen in GOSI9 ORCA12, particularly in the South Pacific and Southern Ocean. These near-surface salinity drifts in GOSI9 may be avoided if a model bias correction scheme is used within the assimilation (e.g., Balmaseda et al., 2007), which looks like a promising next step to further improve results. Globally, the salinity improvements in GOSI9 are generally consistent over time (Figure 28), particularly between 500 and 1500 m. However, as seen for temperature (Figure 26), salinity RMSD degradations may also occur immediately below 1500 m.

The excessive melting of Arctic summer ice is alleviated in GO6 ORCA12 (Figure 29) when compared to its ORCA025 version (Figure 19). However, SIC improvements are still seen in GOSI9 ORCA12 with respect to GO6 over the Arctic summer, with RMSDs decreasing from ~0.06 to ~0.04 (Figure 29). ORCA12 RMSD percentage improvements of GOSI9 SICs relative to GO6 are identified in the Arctic ice pack over summer, reaching up to 50% in the North of Greenland and in the Canadian Archipelago, and are also present in regions near the ice edge (Figure 30).

These GOSI9 SIC improvements in ORCA12 over the Arctic summer are accompanied by slightly worse results in other Arctic seasons when compared to GO6, particularly in the winter (Figure 29), where SIC RMSD degradations are found in the Central Arctic as well as near the ice edge (Figure 30). The ORCA12 SIC statistics in GOSI9 are also generally worse than in GO6 for the Antarctic, especially in the winter, although RMSD percentage improvements and degradations are quite mixed between GOSI9 and GO6 over distinct Antarctic regions (Figure 30).

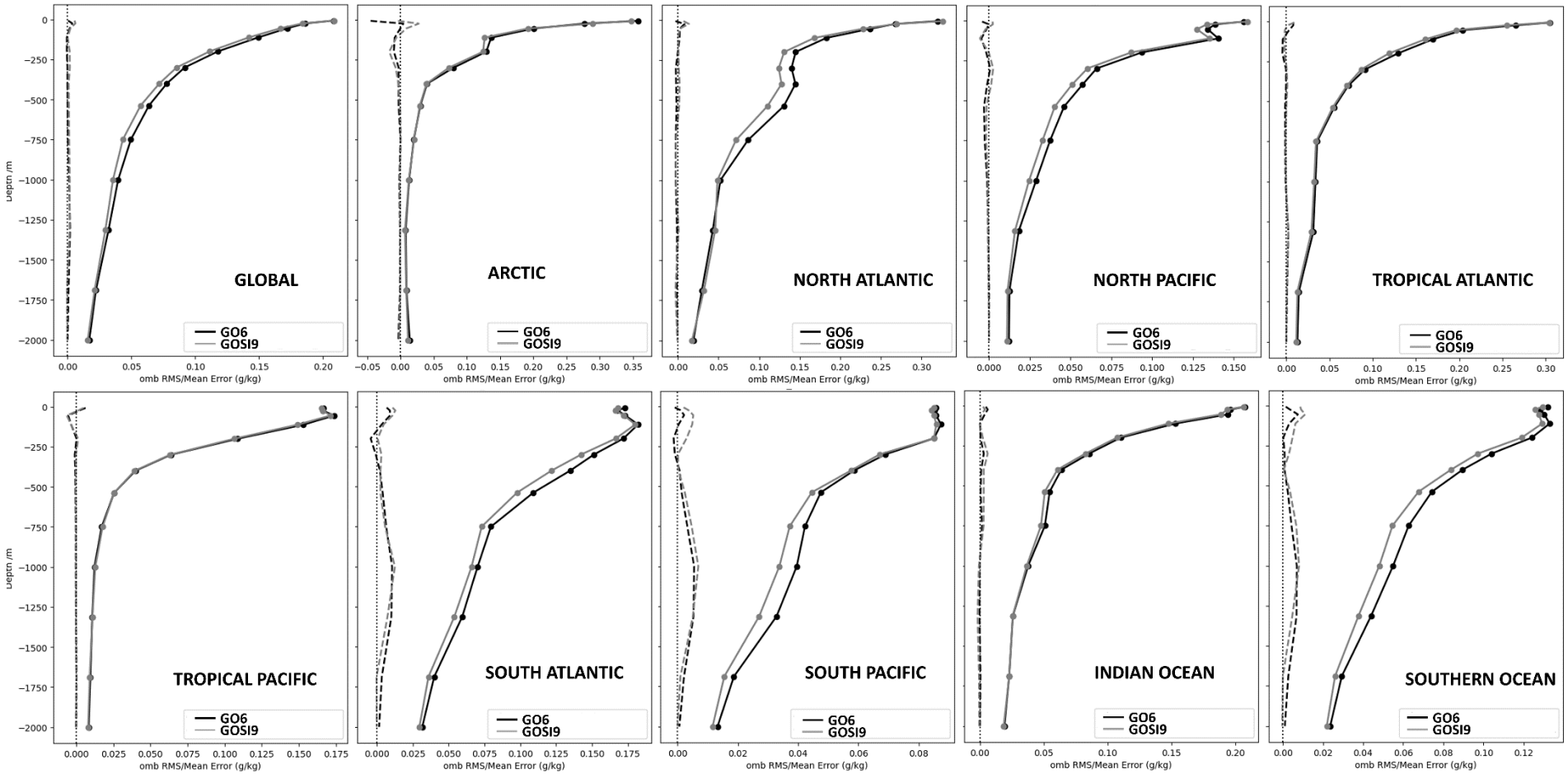

Figure 27: 2019 ORCA12 observation-minus-background salinity statistics for GO6 (black) and GOSI9 (grey) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

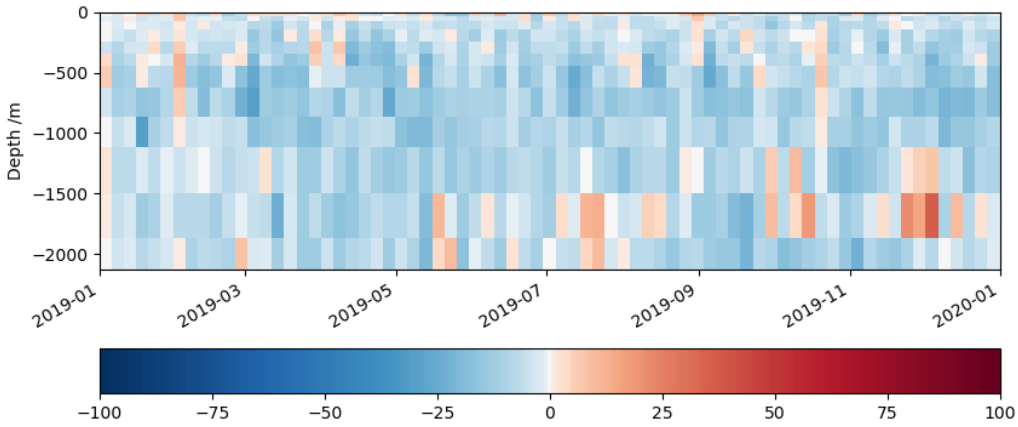

Figure 28: Hovmöller of RMSD salinity improvements (blue) and degradations (red) of ORCA12 GOSI9 relative to GO6, globally averaged as a function of depth. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

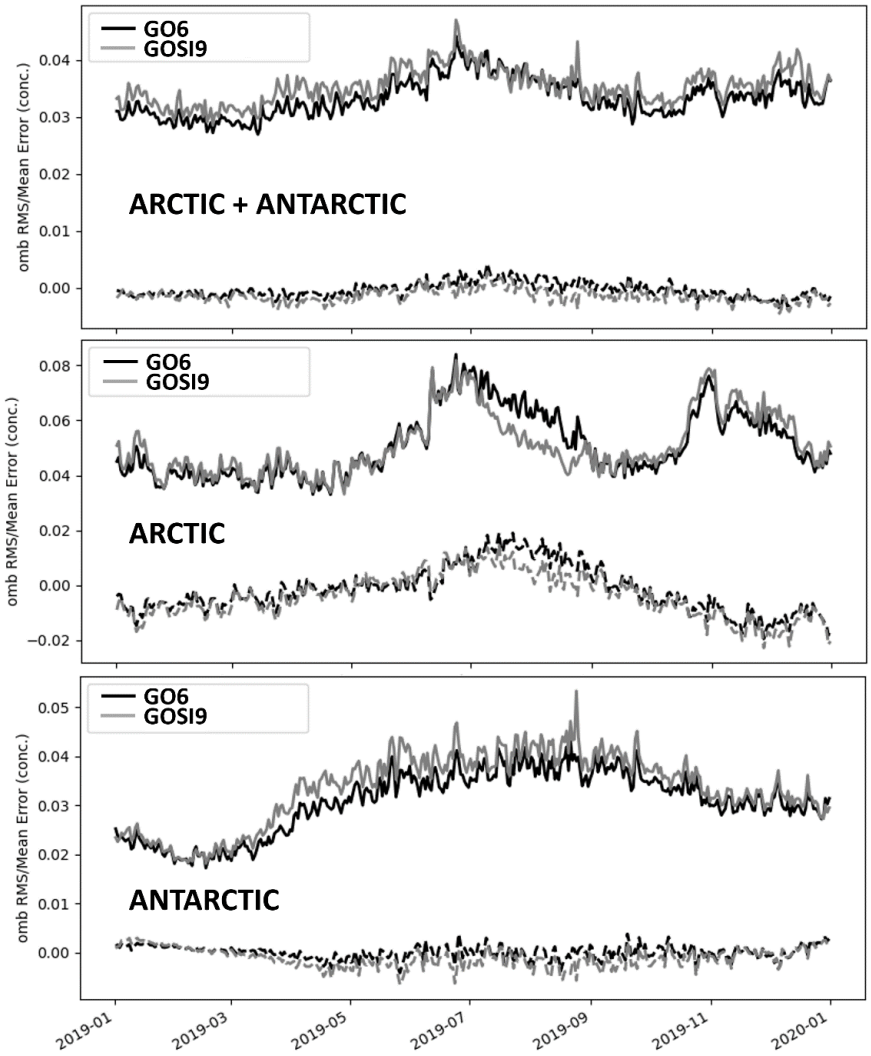

Figure 29: 2019 ORCA12 observation-minus-background SIC statistics for GO6 (black) and GOSI9 (grey), calculated against satellite SIC observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

As in ORCA025, it is worth emphasising that the SIC background error covariances are still derived from a GO6 (CICE) run. Additionally, no sea-ice model tuning has been performed in GOSI9 ORCA12, which reinforces that there is clearly room for SIC improvements coming from both the assimilation and modelling sides in GOSI9 ORCA12.

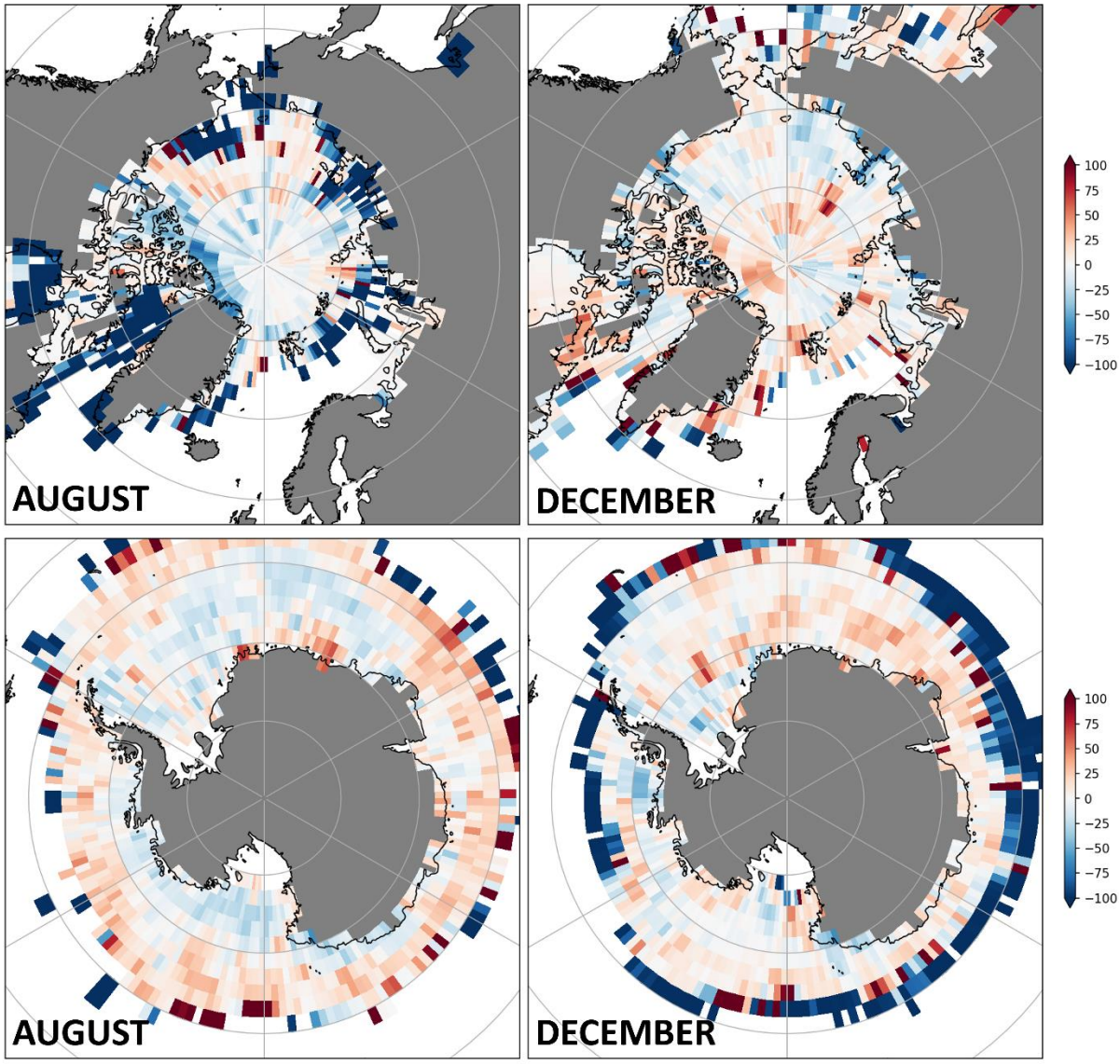

Figure 30: SIC RMSD improvements (blue) and degradations (red) of ORCA12 GOSI9 relative to GO6 for the Arctic (top) and Antarctic (bottom), shown as percentages, for August (left) and December (right). The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

5.2 Potential impact of GOSI9 changes on pre-Argo reanalysis

In this section, we assess the potential impact of the GOSI9 changes on reanalyses before the Argo period by running another set of experiments in 2019 with the ORCA025 system withholding T/S profile observations. Therefore, the following runs are evaluated:

- GO6 trial run shown in section 5.1.1 (GO6)
- GO6 trial run withholding T/S profile data (GO6-NoTS-Prof)
- GOSI9 trial run shown in section 5.1.1 (GOSI9)
- GOSI9 trial run withholding T/S profile data (GOSI9-NoTS-Prof)

As shown in Figure 31, the SLA statistics in GOSI9-NoTS-Prof are better than in GO6-NoTS-Prof. While SLA RMSDs increase in GO6-NoTS-Prof throughout 2019, showing a clear degradation in the SLA statistics without the T/S profile assimilation, the GOSI9-NoTS-Prof experiment holds very similar SLA RMSDs in comparison to the original GOSI9 run, which is even better than the GO6 experiment assimilating T/S profiles. Most of the RMSD degradations in GO6-NoTS-Prof quickly develop at high latitudes, particularly in the Southern Ocean, where the SLAs can reach up to a 100% RMSD degradation in some areas of the ACC relative to GO6 (Figure 32). This is not the case for the GOSI9-NoTS-Prof experiment, which mostly maintains the positive SLA impacts relative to GO6 in all ocean basins, including the Southern Ocean. It is worth highlighting that the SLA statistics in the Mediterranean Outflow are better in GOSI9-NoTS-Prof than in GOSI9 (Figure 32), suggesting that there could still be minor issues in assimilating SLA and T/S profile data together, even after the substantial SLA improvements caused by GOSI9 DA changes.

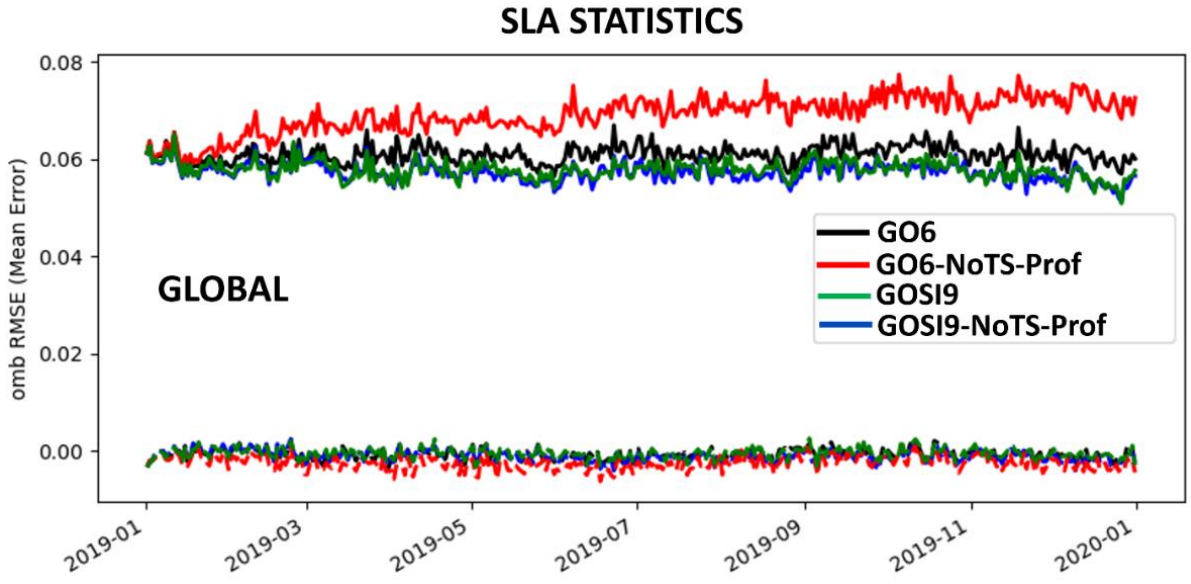

Figure 31: 2019 observation-minus-background SLA statistics (m) for GO6 (black), GO6-NoTS-Prof (red), GOSI9 (green) and GOSI9-NoTS-Prof (blue), calculated against along-track altimeter observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

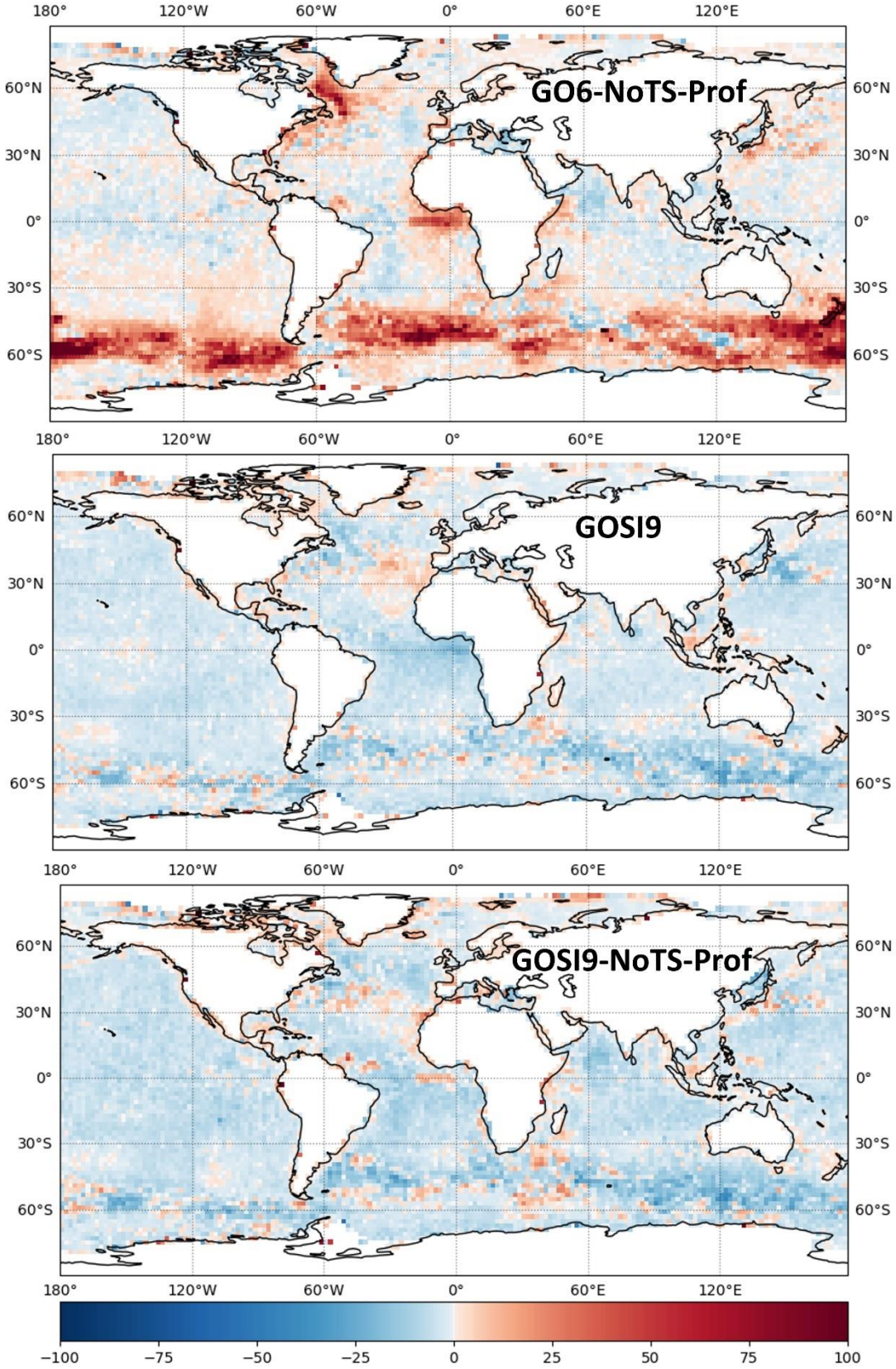

Figure 32: SLA RMSD improvements (blue) and degradations (red) of GO6-NoTS-Prof (top), GOSI9 (middle) and GOSI9-NoTS-Prof (bottom) relative to GO6 for 2019, shown as percentages. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

A very similar behaviour is also observed for SSTs between GOSI9 and GO6 when T/S profile data are not assimilated. While GOSI9-NoTS-Prof maintains very similar SST statistics to GOSI9, even though T/S profiles are not assimilated, the SST statistics in GO6-NoTS-Prof are considerably worse than the original GO6 run (Figure 33). Again, substantial SST RMSD degradations develop in GO6-NoTS-Prof with respect to GO6, particularly near the equator and in the Southern Ocean, which are mostly mitigated in GOSI9-NoTS-Prof (Figure 34). These results are very promising for future pre-Argo ocean reanalysis runs in GOSI9, since both the SST and SLA statistics in GOSI9 are not significantly degraded in the absence of T/S profile assimilation.

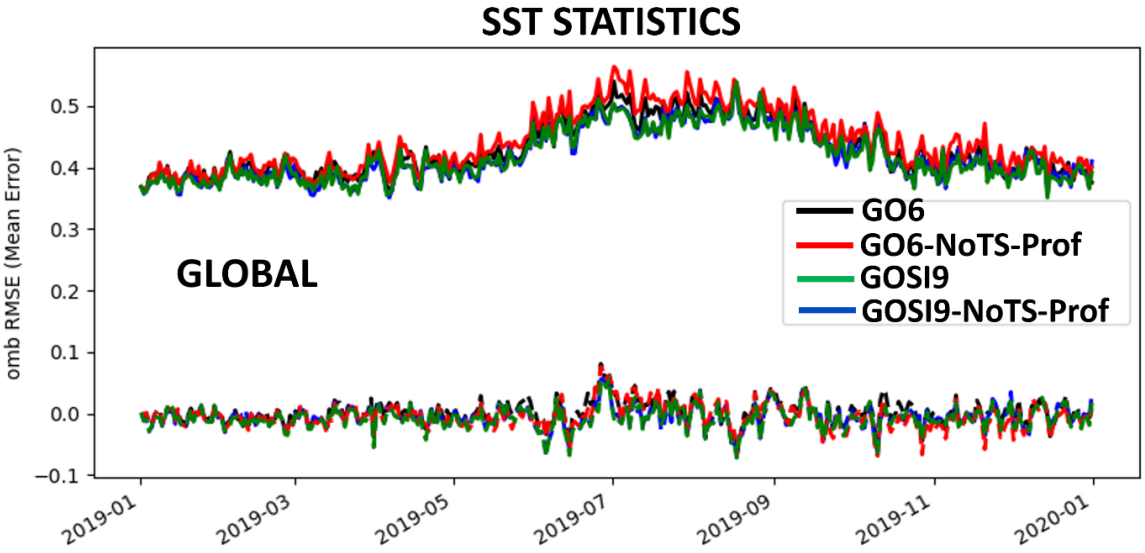

Figure 33: 2019 observation-minus-background SST statistics (°C) for GO6 (black), GO6-NoTS-Prof (red), GOSI9 (green) and GOSI9-NoTS-Prof (blue), calculated against in situ SST drifters. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

Removing the large-scale DA corrections of temperature in GOSI9 significantly mitigates the degradation of sub-surface temperatures, which is a known GO6 issue caused by the large-scale DA propagation of the SLA signal onto model temperatures at depth, especially when there is no T/S profile assimilation. This is highlighted by Figure 35 and Figure 36, which show much smaller degradations in RMSDs and biases for sub-surface temperature and salinity in GOSI9-NoTS-Prof when compared to GO6-NoTS-Prof. As well as in SST and SLA statistics, most of the GOSI9-NoTS-Prof temperature and salinity improvements relative to GO6-NoTS-Prof are found in the Southern Hemisphere, particularly in the South Atlantic, South Pacific and Southern Ocean. Temperature and salinity improvements in GOSI9-NoTS-Prof also extend to the Northern Hemisphere, particularly in the North Atlantic and North Pacific, where sub-surface temperature and salinity biases are much closer to zero when compared to GO6-NoTS-Prof.

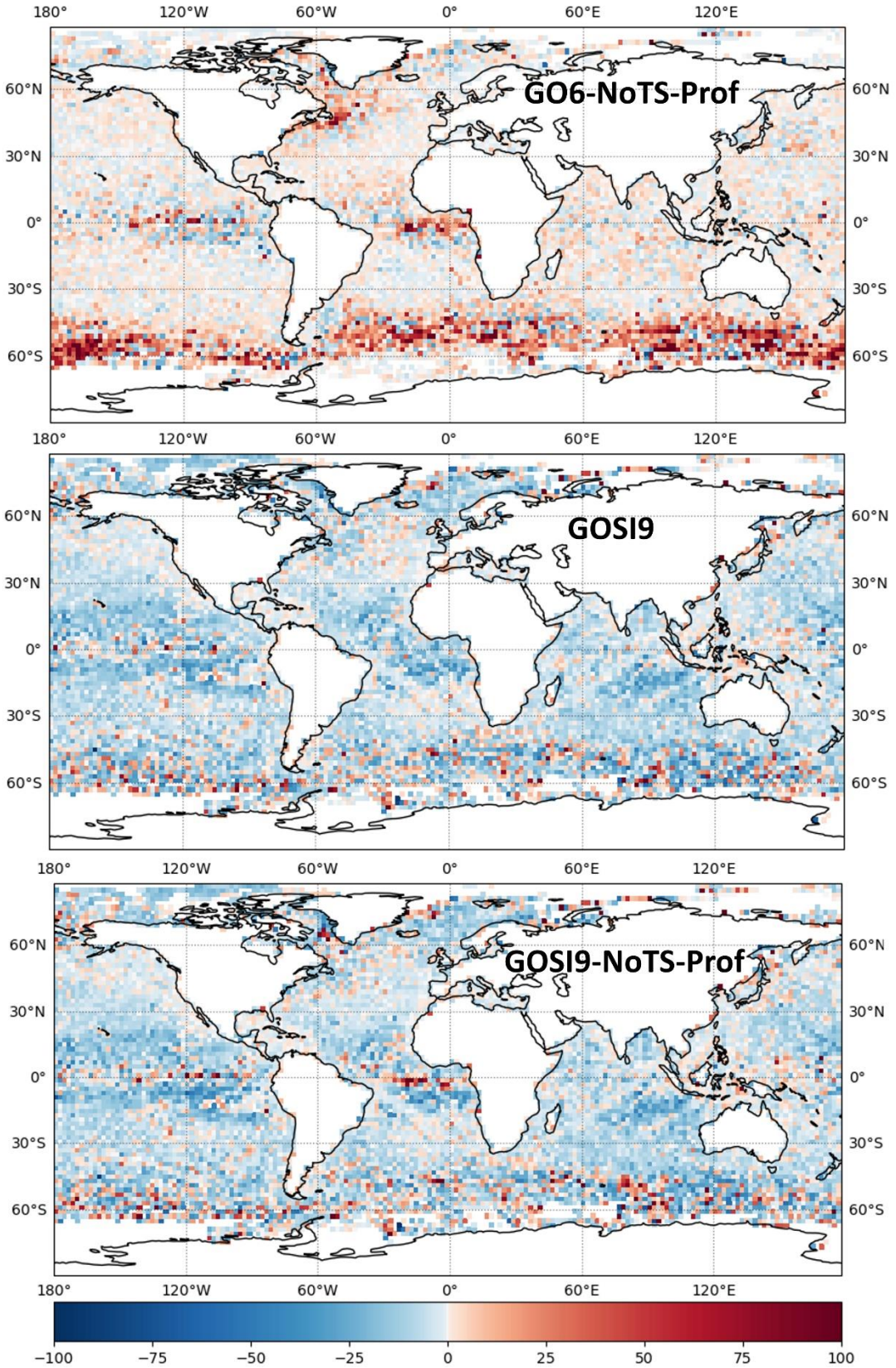

Figure 34: SST RMSD improvements (blue) and degradations (red) of GO6-NoTS-Prof (top), GOSI9 (middle) and GOSI9-NoTS-Prof (bottom) relative to GO6 for 2019, shown as percentages. The RMSD percentage changes between GOSI9 and GO6 are calculated using Equation (1).

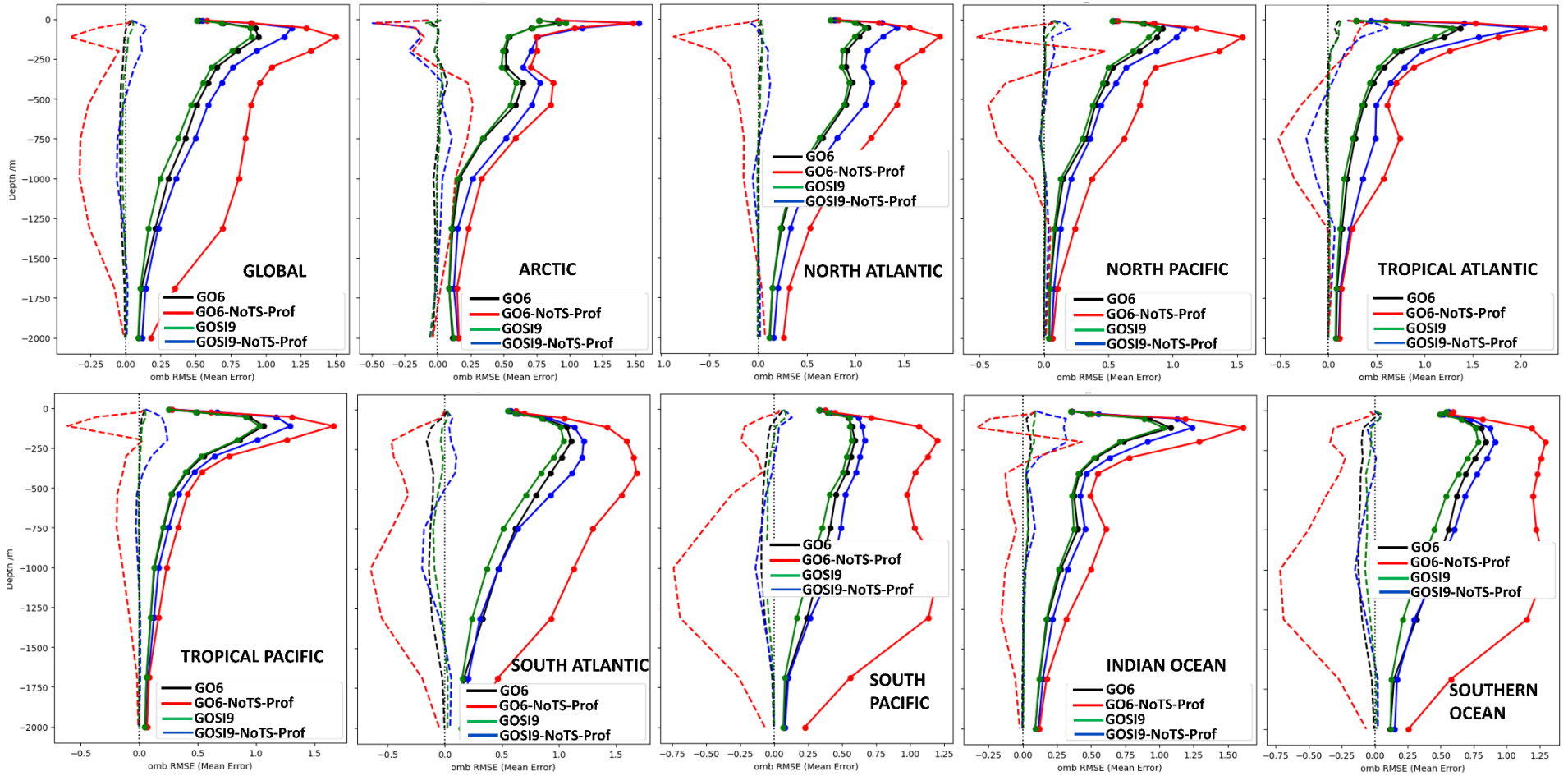

Figure 35: 2019 observation-minus-background temperature statistics (°C) for GO6 (black), GO6-NoTS-Prof (red), GOSI9 (green) and GOSI9-NoTS-Prof (blue) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

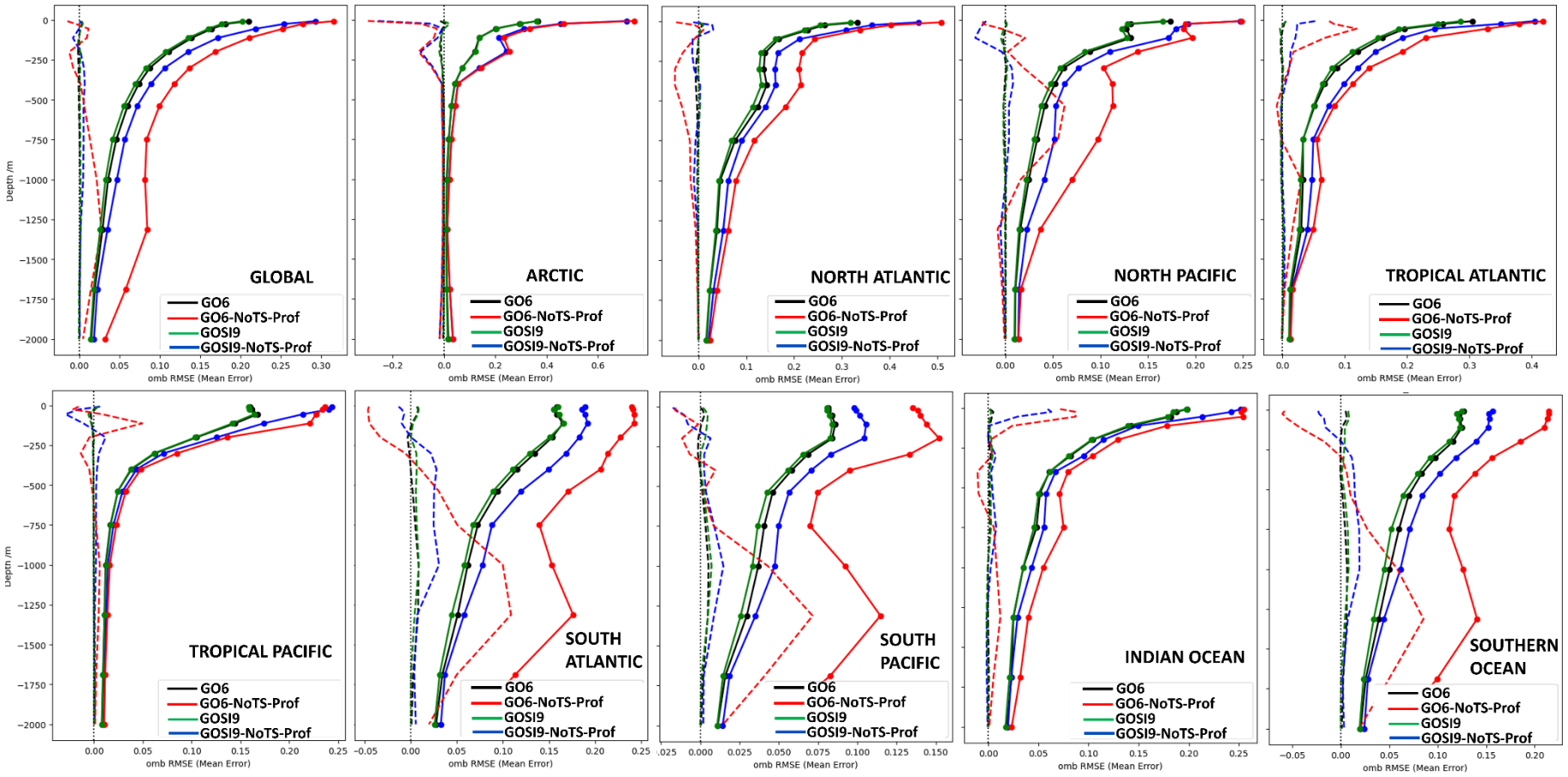

Figure 36: 2019 observation-minus-background salinity statistics for GO6 (black), GO6-NoTS-Prof (red), GOSI9 (green) and GOSI9-NoTS-Prof (blue) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

As a result of these consistent improvements in the sub-surface T/S structure, the globally averaged RMSD in GOSI9-NoTS-Prof is decreased by a factor of 3, between 500 and 1500 m, for both temperature and salinity when compared to GO6-NoTS-Prof (Figure 35 and Figure 36). Similarly, the global mean differences relative to T/S profile observations are much closer to zero in GOSI9-NoTS-Prof relative to GO6-NoTS-Prof, especially between 250 to 1500 m.

In fact, observation-minus-background RMSD comparisons between each experiment and GO6, globally averaged as a function of depth, show that, although GOSI9-NoTS-Prof has an expected T/S degradation with respect to GOSI9, it is much smaller over time than in GO6-NoTS-Prof (Figure 37). In the first half of 2019, the RMSD degradations in GO6-NoTS-Prof quickly reaches 100% relative to GO6 at depths between 1000 m and 2000 m for temperature and salinity, extending to almost the whole water column by the end of the year. In GOSI9-NoTS-Prof, the RMSD degradations with respect to GO6 are much smaller, around 30%, and they remain largely steady throughout the year.

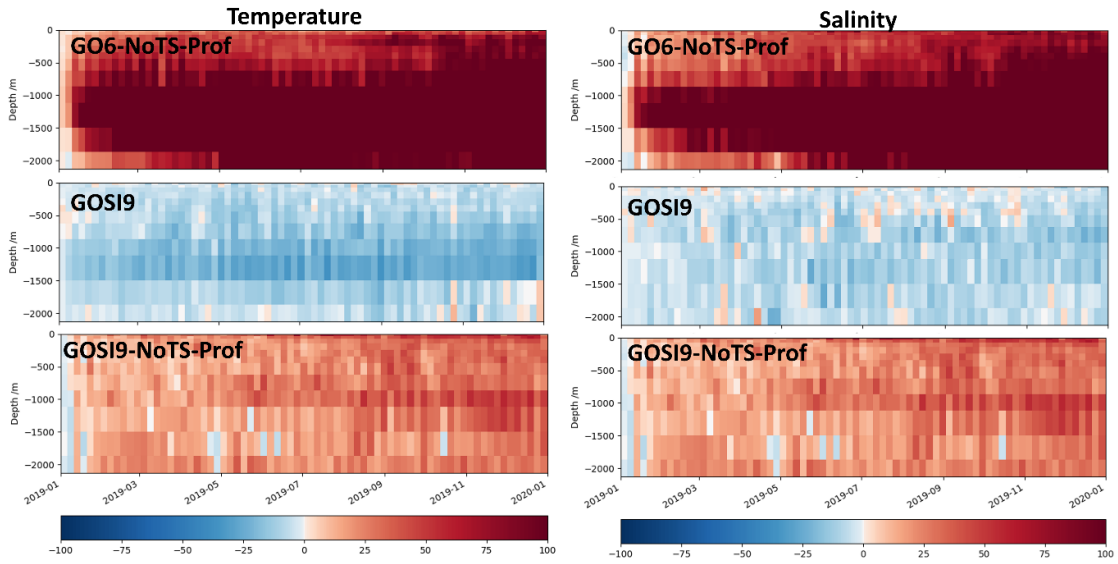

Figure 37: Temperature (left) and salinity (right) RMSD percentage changes of GO6-NoTS-Prof (top), GOSI9 (middle) and GOSI9-NoTS-Prof (bottom) relative to GO6 for 2019. The RMSD changes are calculated as percentages using Equation (1), indicating improvements (degradations) with respect to GO6 when values are negative (positive).

In addition to the plots and statistics presented so far, we also include here some additional diagnostics looking at the impacts of the GOSI9 changes on the Atlantic Meridional Overturning Circulation (AMOC) and heat content, which are important key indicators from a reanalysis perspective. The heat content time series (Figure 38) shows that GOSI9-NoTS-Prof drifts much less than GO6-NoTS-prof when both are compared to their original runs assimilating T/S profile observations. The heat content drifts in GO6-NoTS-prof are significant in both Atlantic and Pacific oceans, particularly at depths between 700 and 2000 m, which is consistent with the depth range where large RMSD degradations in temperature are seen with respect to GO6 (Figure 37).

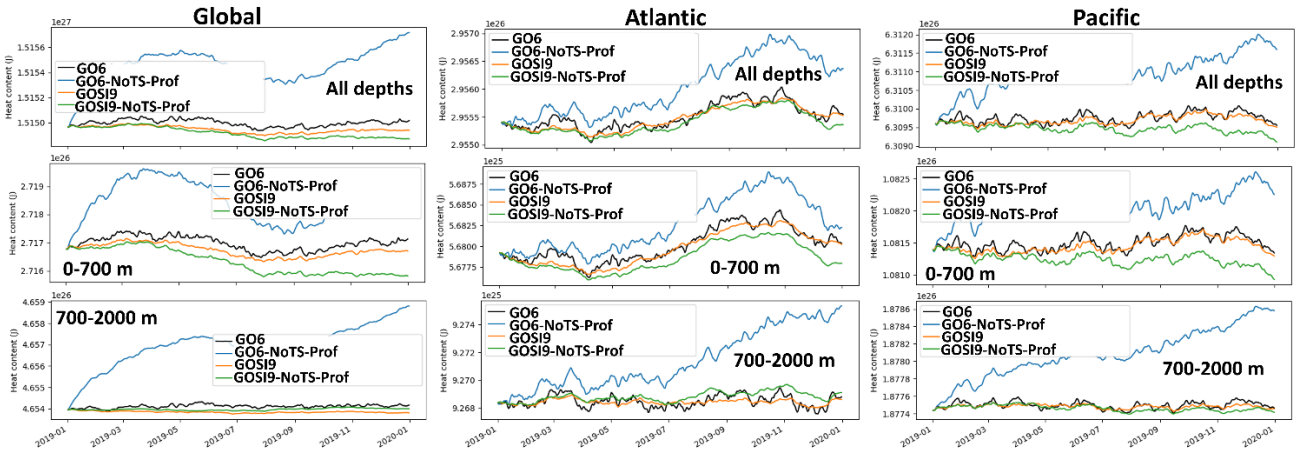

Figure 38: Daily time series of heat content (J) for the global (left), Atlantic (middle) and Pacific (right) oceans, considering all depths (top), 0-700 m (middle) and 700-2000 m (bottom). GO6 (black), GO6-NoTS-Prof (blue), GOSI9 (orange) and GOSI9-NoTS-Prof (green) heat content time series are shown for 2019.

The heat content drift seen in GO6-NoTS-Prof between 700 and 2000 m is largely mitigated in GOSI9-NoTS-Prof, which follows very closely the heat content of the GOSI9 run assimilating T/S profile observations (Figure 38). However, heat content drifts are still present in GOSI9-NoTS-Prof to some extent for depths between 0 and 700 m, in both Atlantic and Pacific oceans. One might expect some heat content drift in the pre-Argo period due to the lack of T/S profile assimilation, and longer GOSI9 reanalysis runs should be conducted to look at these drifts in more detail. In any case, future work is planned to implement a model bias correction scheme in FOAM, so these heat content drifts can be further improved, particularly in the pre-Argo period.

It is also worth highlighting in Figure 38 that the GO6 run assimilating T/S profiles shows large and unrealistic daily heat content variability, which is introduced by the large-scale signal from the SLA assimilation being inconsistently propagated onto the model temperatures at depth. This was noted as a problem by Dong et al. (2021) when seeking to apply an assimilation smoother to FOAM GO6 results. Since these large-scale SLA corrections onto sub-surface temperatures are removed from DA in GOSI9, the daily heat content variability in the new FOAM system is already much smoother than in GO6. This should enable a consistent application of the assimilation smoother, as in Dong et al. (2021), to further improve GOSI9 reanalysis results.

As well as for the heat content, the GOSI9 DA impacts on the AMOC transports are encouraging (Figure 39 and Figure 40). The monthly AMOC transports in GOSI9-NoTS-Prof follow very closely those from the GOSI9 run assimilating T/S profiles, particularly in the Northern Hemisphere, and still show relatively good agreement with RAPID transports at 26.5°N (Figure 39). On the other hand, the AMOC transports in GO6-NoTS-Prof significantly drift, producing increased and unrealistic AMOC transports at 50°N and 26.5°N, which do not agree at all with RAPID transports. Although the drifts in the AMOC transports are largely reduced in GOSI9-NoTS-Prof in the Northern Hemisphere, it is worth noting that differences of ~4 Sv can occur between this run and the GOSI9 run assimilating T/S profiles at 30°S,

particularly for July-November, before converging to similar AMOC transports in December 2019. Longer time series of AMOC transports in the Southern Hemisphere will be evaluated in future GOSI9 reanalysis runs, although previous results have shown that T/S, SST and SLA statistics are clearly improved in the Southern Hemisphere with GOSI9 DA changes. Therefore, it is expected that the ocean circulation will also respond positively to these DA changes in the Southern Hemisphere.

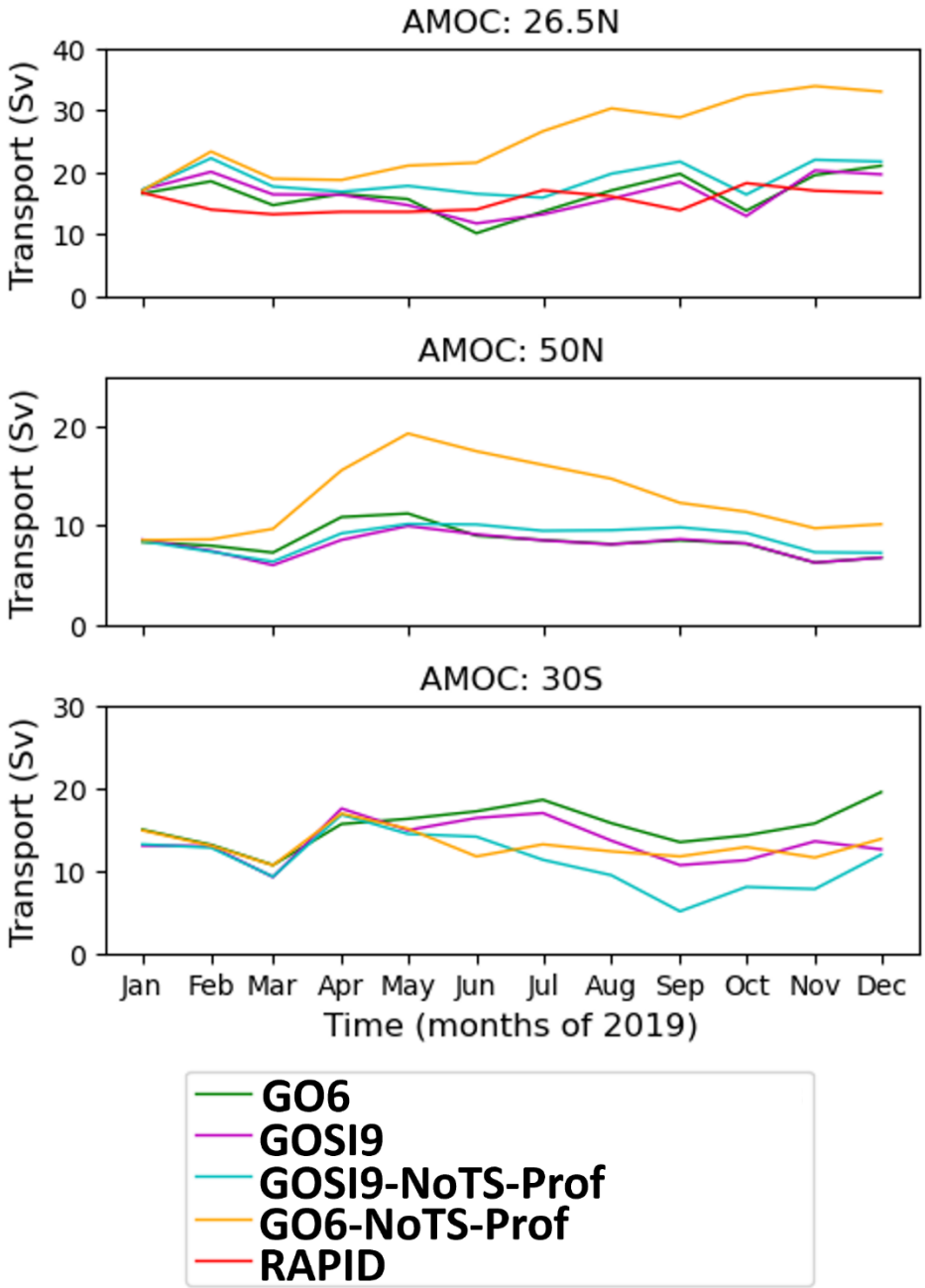

Figure 39: Monthly AMOC transports (Sv) in 2019 at 26.5°N (top), 50°N (middle) and 30°S (bottom) for GO6 (green), GOSI9 (magenta), GOSI9-NoTS-Prof (cyan) and GO6-NoTS-Prof (orange). Note that observed AMOC transports from the RAPID array (red) are also included at 26.5°N.

GOSI9 results are also promising in terms of the AMOC stream function. In the absence of T/S profile assimilation, GOSI9-NoTs-Prof transports are not as intense as its equivalent GO6-

NoTs-Prof run in the Northern Hemisphere, dropping from ~40 Sv to ~25 Sv just north of the equator. It is also worth highlighting that the upper AMOC transports are significantly weakened south of the equator in GO6-NoTS-Prof, particularly between 10°S and the equator. This abrupt disruption in the upper AMOC transports near the equator is significantly improved in GOSI9-NoTS-Prof, where the AMOC seems to be more meridionally coherent across the Atlantic, particularly in the equatorial region.

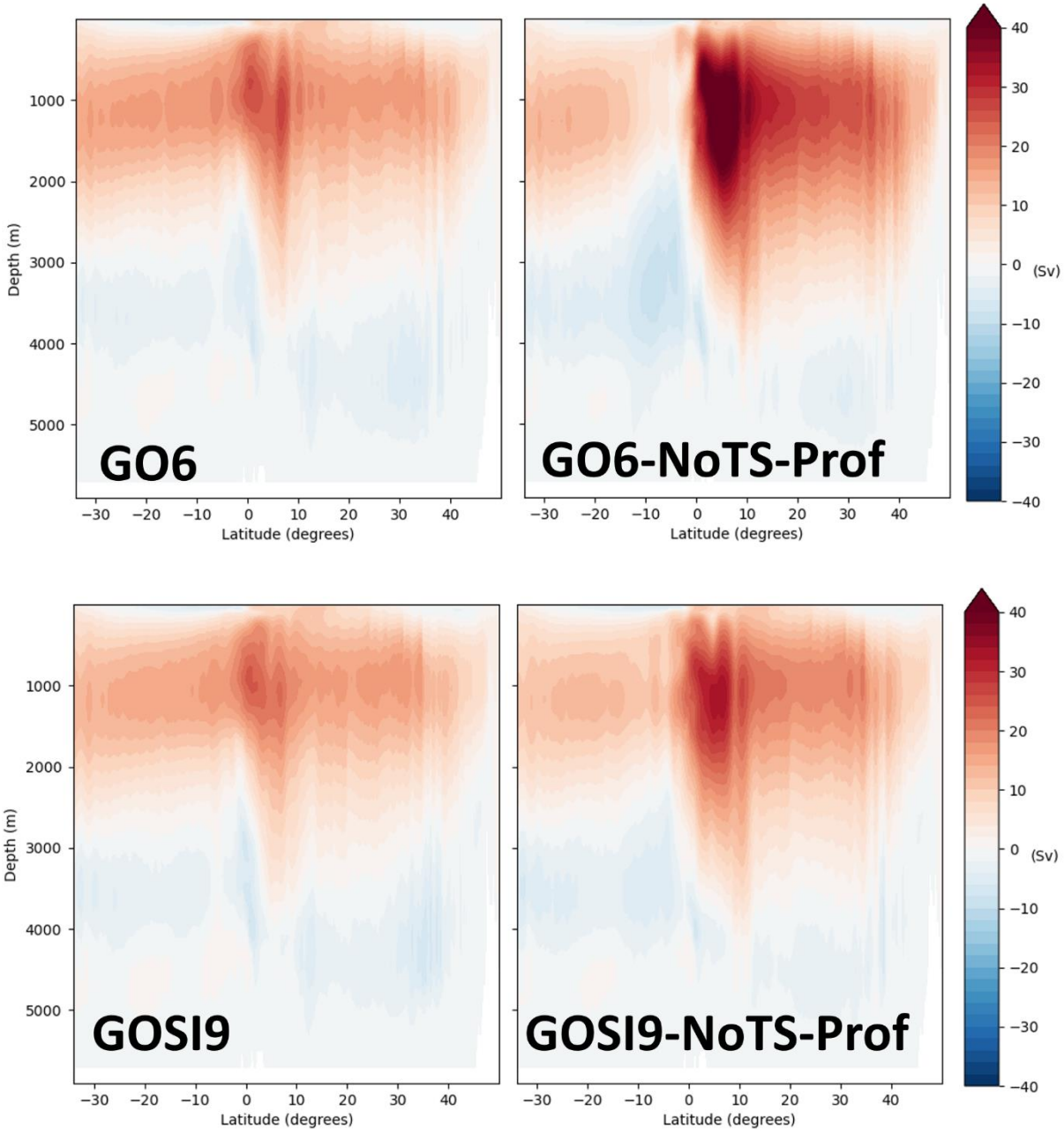

Figure 40: 2019 AMOC stream function (Sv) for GO6 (top left), GO6-NoTS-Prof (top right), GOSI9 (bottom left) and GOSI9-NoTS-Prof (bottom right).

6. Conclusions

In this technical report, the impacts of GOSI9 updates to the global FOAM system, including model and data assimilation changes, were evaluated by comparing the GOSI9 performance of both ORCA025 and ORCA12 against their respective GO6 versions in 2019. Further experiments were also conducted to test the impact of GOSI9 changes on the pre-Argo period by running further GO6 and GOSI9 experiments in 2019 withholding T/S profile data.

For both ORCA025 and ORCA12, the GOSI9 system performs considerably better than GO6 for ocean variables, particularly for SLAs and sub-surface temperatures, with a substantial decrease in global observation-minus-background RMSD and mean difference. This is largely driven by improvements in the Southern Hemisphere. Although they are smaller, SLA and temperature improvements also extend to the Northern Hemisphere, particularly in the North Atlantic and North Pacific. For example, GOSI9 SLAs are clearly improved relative to GO6 in the western boundary currents, such as in the Brazil-Malvinas confluence, Gulf Stream and the Kuroshio current, as well as along the ACC path. The increased number of NEMOVAR inner loop iterations, improvements in the SST and SLA observation error specifications, and removing large-scale DA corrections for model temperatures at depth all played an essential role in producing positive impacts on GOSI9 SLAs and temperatures. It is worth highlighting that these DA changes also produced better RMSD results for sub-surface salinities in the Southern Hemisphere, although near-surface biases are slightly increased in the same regions when GOSI9 salinity statistics are compared to those from GO6.

Even though GOSI9 changes have resulted in major FOAM improvements, the fact that a long length-scale is used to spread the salinity, but not the temperature information, may exacerbate localised water column instabilities in regions sensitive to T/S increments, such as the Mediterranean Outflow. Therefore, Brunt-Väisälä buoyancy frequencies are calculated in GOSI9 and used to reject T/S increments in places where they would make the water column unstable. This largely mitigates the localised T/S DA issues. However, the GOSI9 SLA statistics are still slightly worse in the Mediterranean Outflow relative to GO6.

For sea ice, GOSI9 improvements for both ORCA025 and ORCA12 are mixed when compared to GO6. Substantial GOSI9 SIC improvements are seen in the Arctic over summer, reducing the excessive melting of Arctic summer ice in GO6, particularly in ORCA025. However, for other Arctic seasons, as well as throughout the year in the Antarctic, the GOSI9 SIC results are slightly worse when compared to GO6. Although the sea-ice configurations are intended to be very similar between GOSI9 and GO6, CICE has been replaced by SI^3 in GOSI9. Therefore, there is room for further enhancements in GOSI9 SICs, by improving the model tuning and assimilation configuration, such as the sea-ice background error covariances, which are still derived from a CICE run.

In addition to clear improvements in the FOAM performance, GOSI9 DA updates also have a

large potential to enhance reanalysis runs, particularly in the pre-Argo period. The SLA, SST and T/S statistics from a 1-year GOSI9 run withholding profile observations are significantly better than its equivalent GO6 run, particularly in the Southern Hemisphere. These improvements also extend to heat content and AMOC diagnostics. Both the heat content and AMOC transports are much more temporally consistent in GOIS9 compared to GO6, drifting much less when T/S profile data are not assimilated. These are promising results from a reanalysis perspective, which could lead to more potential for use of the Met Office ocean reanalysis in climate studies.

Since the GOSI9 updates are planned to be implemented in the next scientific upgrade of the operational suite, it is expected that both FOAM and coupled NWP systems will reflect the improvements shown in this technical report. This is expected to benefit a wide range of users, including those from the navigation sector and the Navy. Moreover, the GOSI9 updates will also enable the Met Office teams in Climate Science to run improved ocean reanalyses and seasonal forecasts.

Although large-scale DA corrections for model temperatures have been removed in GOSI9, work is currently being done to improve the length-scales of the background errors. The implementation of a hybrid ensemble/variational assimilation scheme in NEMOVAR, as described in Lea et al. (2022), adds flow-dependent ensemble information to the background errors and improves the assimilation results when compared to the 3DVar scheme. Moreover, a convolutional neural network algorithm is being tested to estimate flow-dependent length-scales of the background errors.

As future work, there are already planned implementations aimed at further improving GOSI9 results. A T/S model bias correction scheme will be implemented to further reduce model drifts in temperature and salinity, particularly in the pre-Argo period. Further work will also be done to keep improving background and observation error covariances, which include better error covariance specifications for sea-ice concentration in both the Arctic and Antarctic. Since the TRF functionality is now implemented in GOSI9, we are also implementing ORCA12 DA at full resolution.

References

Balmaseda, M. A., D. Dee, A. Vidard, and D. L. T. Anderson, 2007: A Multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Quart. J. Roy. Meteor. Soc.*, 133**,** 167–179.

Barbosa Aguiar, A., Bell, M.J., Blockley, W., Calvert, D., Crocker, R., King, R., Lea, D.J., Maksymczuk, J., Martin, M.J., Price, M., Waters, J., While, J., Smout-Day, K., Siddorn, J. The Met Office Forecast Ocean Assimilation Model (FOAM) using a 1/12 degree grid for global forecasts. Submitted to Q.J. Royal Met. Soc., 2024.

Blockley, E., Fiedler, E., Ridley, J., Roberts, L., West, A., Copsey, D., Feltham, D., Graham, T., Livings, D., Rousset, C., Schroeder, D., and Vancoppenolle, M.: The sea ice component of GC5: coupling SI3 to HadGEM3 using conductive fluxes, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-1731, 2023.

Dong, B., K. Haines and M. Martin. (2021) Improved high resolution ocean reanalyses using a simple smoother algorithm. JAMES, https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2021MS002626.

Gent, P.R., McWilliams, J.C. Eliassen–Palm fluxes and the momentum equation in non-eddy-resolving ocean circulation models. J. Phys. Oceanography. 26, 2539–2546. 1996.

Guiavarc'h, C., Storkey, D., Blaker, A., Blockley, E., Megann, A., Hewitt, H., Bell, M., Copsey, D., Sinha, B., Moreton, S. , Mathiot, P., Ann, B. GOSI9: UK Global Ocean and Sea Ice configurations (to be submitted to Geoscientific Model Development). 2023.

Gürol S, Weaver AT, Moore AM, Piacentini A, Arango HG, Gratton S. 2014. b-preconditioned minimization algorithms for variational data assimilation.Q. J. R. Meteorol. Soc.140: 539–556.

Hallberg, R: Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects, Ocean Modelling, 72, 92-103, 2013.

Hollingsworth, A. and Lönnberg, P., 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus, 38A, 111-136.

IOC, SCOR and IAPSO. The international thermodynamic equation of seawater: Calculation and use of thermodynamic properties. Intergovernmental Oceanographic Commission, Manuals and Guides No.56, UNESCO, 196 pp, 2010.

Jousset S., Mulet S., Wilkin J., Greiner E., Dibarboure G. and Picot N. (2022b): "New global Mean Dynamic Topography CNES-CLS-22 combining drifters, hydrological profiles and High Frequency radar data", OSTST 2022, https://doi.org/10.24400/527896/a03-2022.3292.

Jousset S., Aydogdu A., Ciliberti S., Clementi E., Escudier R., Jansen E., Lima L., Menna M., Mulet S., Nigam T., Sanchez A., Tarry D. R., Pascual A., Peneva E., Poulain P.-M. and Taupier-Letage I. (2022b). New Mean Dynamic Topography of the Mediterranean and Black Seas from altimetry, gravity and in-situ data. In preparation.

Lea, D. J., J.-P. Drecourt, K. Haines, M. J. Martin, 2008. Ocean altimeter assimilation with

observational- and model-bias correction. Q. J. Roy. Met. Soc. 134:1761-1774.

Lea, D.J., While, J., Martin, M.J., Weaver, A., Storto, A. and Chrust, M. (2022) A new global ocean ensemble system at the Met Office: Assessing the impact of hybrid data assimilation and inflation settings. Q. J. Roy. Met. Soc, 134: 1996-2030.

Mignac, D. Development of the HL Method and Function Fitting. ErrorCov1.0.0, Zenodo, 4 Mar. 2021, doi:10.5281/zenodo.4580548.

Mirouze, I., E.W. Blockley, D.J. Lea, M.J. Martin and M.J. Bell, 2016. A multiple length scale correlation operator with application to ocean data assimilation Tellus A, 68, 29744, http://dx.doi.org/10.3402/tellusa.v68.29744.

Mirouze, I., & Weaver, A.T. Representation of correlation functions in variational assimilation using an implicit diffusion operator. Quarterly Journal of the Royal Meteorological Society, 136., 2010.

Mulet, S., Rio, M.-H., Etienne, H., Artana, C., Cancet, M., Dibarboure, G., Feng, H., Husson, R., Picot, N., Provost, C., and Strub, P. T.: The new CNES-CLS18 Global Mean Dynamic Topography, Ocean Sci., 17, 789-808, https://doi.org/10.5194/os-17-789-2021, 2021.

Oke P. R., Sakov P. 2008. Representation error of oceanic observations for data assimilation. J Atmos Ocean Technol. 25:1004–1017.

Peterson, K.A., Arribas, A., Hewitt, H.T., Keen, A.B., Lea, D.J. and McLaren, A.J. (2015) Assessing the forecast skill of Arctic sea ice extent in the GloSea4 seasonal prediction system. Climate Dynamics, 44, 147–162. https://doi.org/10.1007/s00382- 014-2190-9.

Ridley, J. K., Blockley, E. W., Keen, A. B., Rae, J. G. L., West, A. E., and Schroeder, D.: The sea ice model component of HadGEM3-GC3.1, Geosci. Model Dev., 11, 713–723, https://doi.org/10.5194/gmd-11-713-2018, 2018.

Rio, M.-H., Mulet, S., Picot, N. New global mean dynamic topography from a GOCE geoid model, altimeter measurements and oceanographic in-situ data. ESA Living Planet Symposium, Proceedings of the conference held on 9-13 September 2013 at Edinburgh in United Kingdom. ESA SP-722. 2-13, p.27, 2013. Bibcode: 2013ESASP.722E..27R

Rio, M.-H., S. Mulet, and N. Picot (2014), Beyond GOCE for the ocean circulation estimate: Synergetic use of altimetry, gravimetry, and in situ data provides new insight into geostrophic and Ekman currents, Geophys. Res. Lett.,41,8918–8925, doi:10.1002/2014GL061773.

Shchepetkin, A. F. An adaptive, Courant-number-dependent implicit scheme for vertical advection in oceanic modelling, Ocean Modelling, 91, 38-69, https://doi.org/https://doi.org/10.1016/j.ocemod.2015.03.006. 2015.

Smith, G. C., Roy, F., Reszka, M., Surcel Colan, D., He, Z., Deacu, D., Belanger, J.-M., Skachko, S., Liu, Y., Dupont, F., Lemieux, J.-F., 745 Beaudoin, C., Tranchant, B., Drevillon, M., Garric, G., Testut, C.-E., Lellouche, J.-M., Pellerin, P., Ritchie, H., Lu, Y., Davidson, F., Caya, A., and Lajoieb, M.: Sea ice forecast verification in the Canadian Global Ice Ocean Prediction System, Q. J. R. M. S., 142, 659–671, https://doi.org/10.1002/qj.2555, 2015.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., Kuhlbrodt, T., Rae, J. G. L., and Sinha, B.: UK Global Ocean GO6 and GO7: a traceable hierarchy of model resolutions, Geosci. Model Dev., 11, 3187–3213, https://doi.org/10.5194/gmd-11-3187-2018, 2018.

Storkey D., Mathiot P., Bell M., Copsey D., Guiavarc'h C., Hewitt H., Ridley J. (2024). Resolution dependence of interlinked biases in the Southern Ocean in the initial spin up of global coupled HadGEM3 models. In preparation.

Tréguier, A. M., Held, I. M., and Larichev, V. D.: Parameterization of Quasigeostrophic Eddies in Primitive Equation Ocean Models, Journal of Physical Oceanography, 27, 567 – 580, 1997.

Waters, J., Lea, D.J., Martin, M.J., Mirouze, I., Weaver, A.T. and While, J. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. Quarterly Journal of the Royal Meteorological Society, 141, 333–349. https://doi.org/10.1002/qj.2388. 2015.

Weaver, A. T., Chrust, M., Ménétrier, B. and Piacentini, A. "An evaluation of methods for normalizing diffusion-based covariance operators in variational data assimilation." Quarterly Journal of the Royal Meteorological Society 147, 289 - 320. 2020.

Weaver A. T., Courtier P. Correlation modelling on the sphere using a generalized diffusion equation. Q. J. R. Meteorol. Soc. 127: 1815–1846. 2001.

Weaver, A. T., Tshimanga, J. and A. Piacentini, 2016. Correlation operators used on an implicitly formulated diffusion equation solved with the Chebyshev iteration. Q. J. Roy. Meteorol. Soc., 142: 455-471.

While, J, Martin, MJ. Variational bias correction of satellite sea-surface temperature data incorporating observations of the bias. Q J R Meteorol Soc. 2019; 145: 2733– 2754. https://doi.org/10.1002/qj.3590.

Vancoppenolle, M., Rousset, C., Blockley, E., Aksenov, Y., Feltham, D., Fichefet, T., Garric, G., Guémas, V., Iovino, D., Keeley, S., Madec, G., Massonnet, F., Ridley, J., Schroeder, D., and Tietsche, S.: SI3, the NEMO Sea Ice Engine (4.2release_doc1.0), Zenodo, https://doi.org/10.5281/zenodo.7534900, 2023.

Appendix A – Impacts of the model and assimilation changes on GOSI9

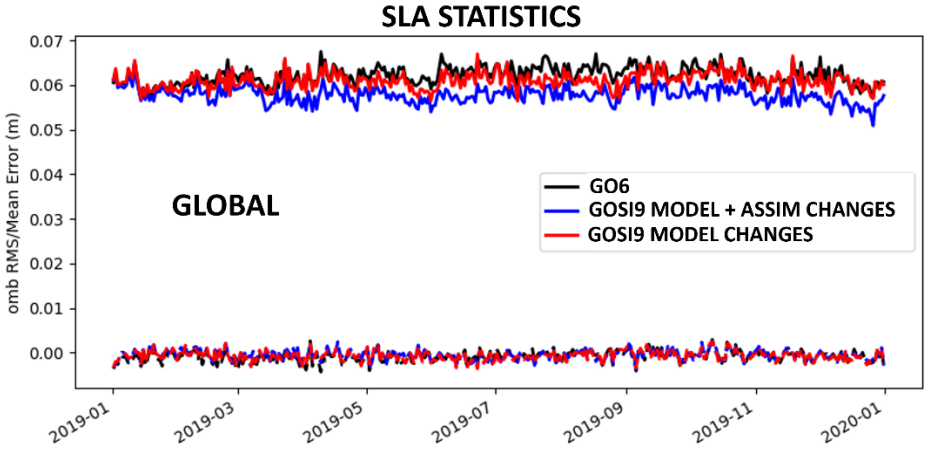

Figure 1A: 2019 observation-minus-background SLA statistics (m) for GO6 (black), GOSI9 with only model changes (red), and GOSI9 with model plus assimilation changes (blue), calculated against along-track altimeter observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

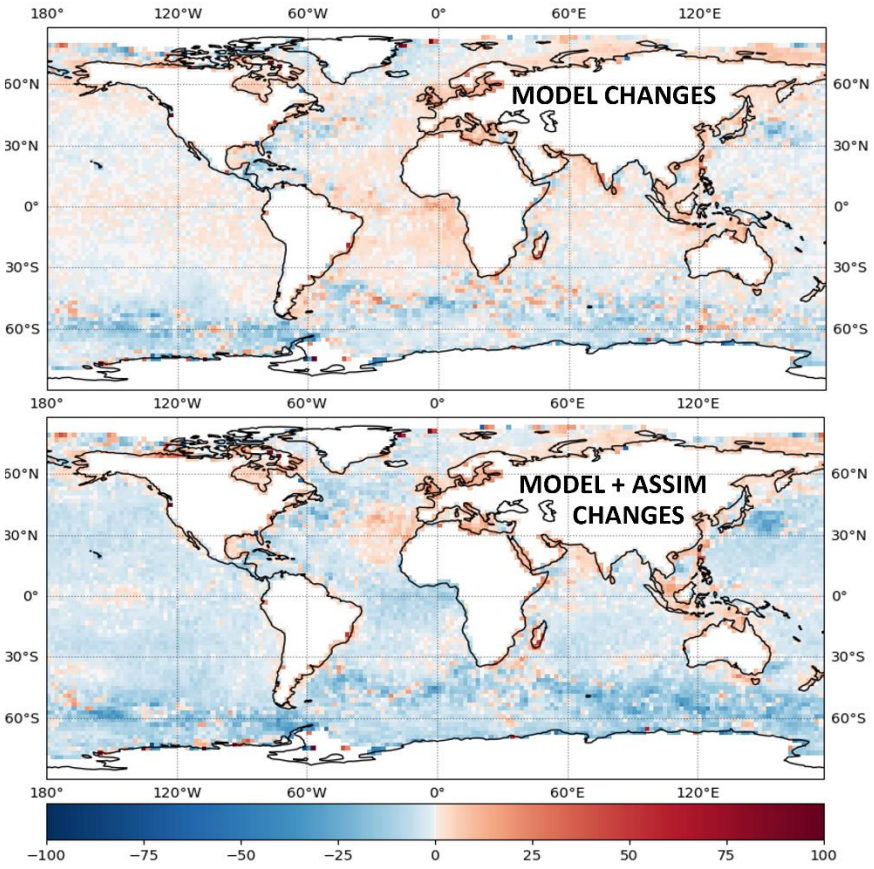

Figure 2A: SLA RMSD improvements (blue) and degradations (red) of GOSI9 with only model changes (top) and GOSI9 with model plus assimilation changes relative to GO6 for 2019, shown as percentages. The RMSD percentage changes are calculated using Equation (1).

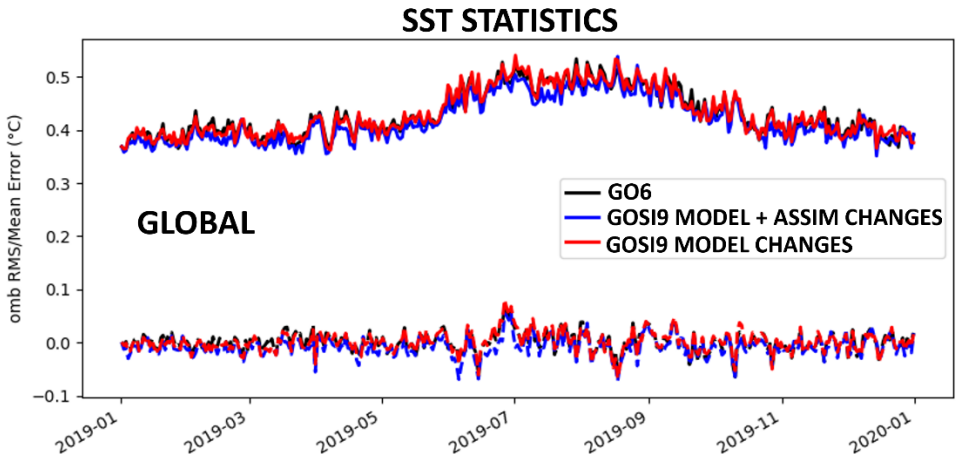

Figure 3A: 2019 observation-minus-background SST statistics (°C) for GO6 (black), GOSI9 with only model changes (red), and GOSI9 with model plus assimilation changes (blue), calculated against in situ SST drifters. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

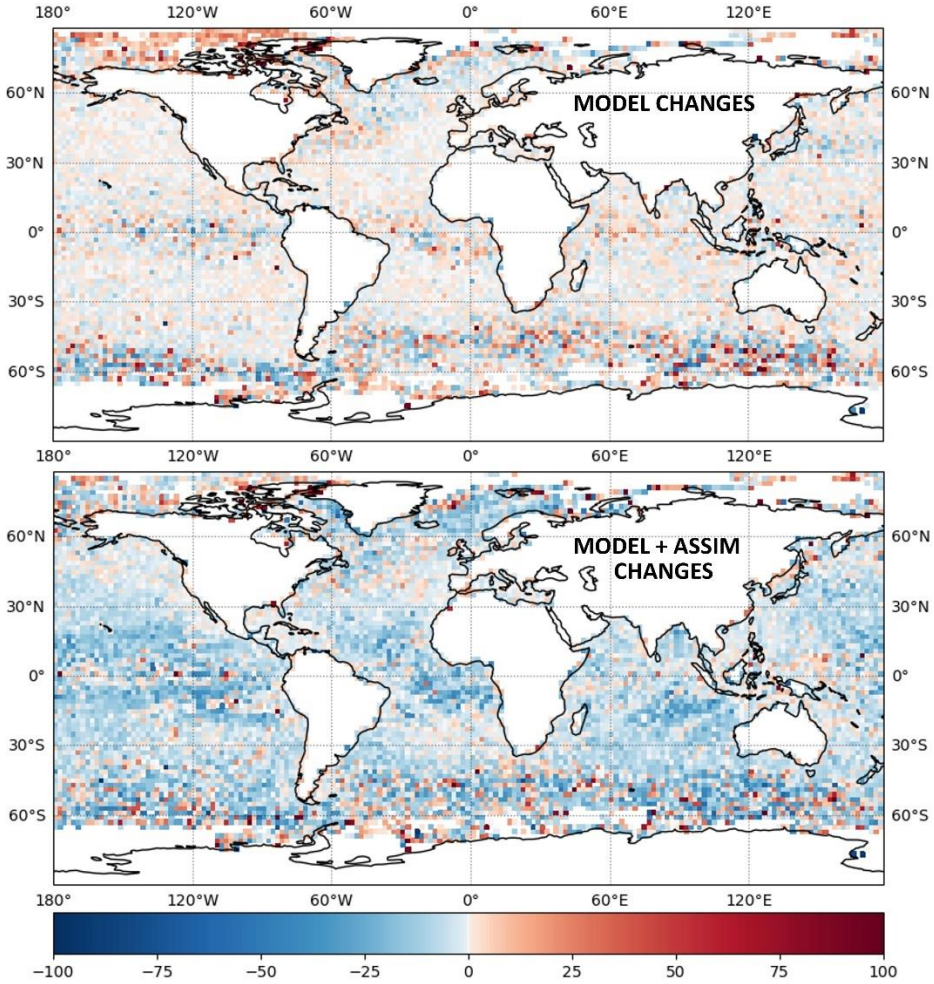

Figure 4A: SLA RMSD improvements (blue) and degradations (red) of GOSI9 with only model changes (top) and GOSI9 with model plus assimilation changes (bottom) relative to GO6 for 2019, shown as percentages. The RMSD percentage changes are calculated using Equation (1).

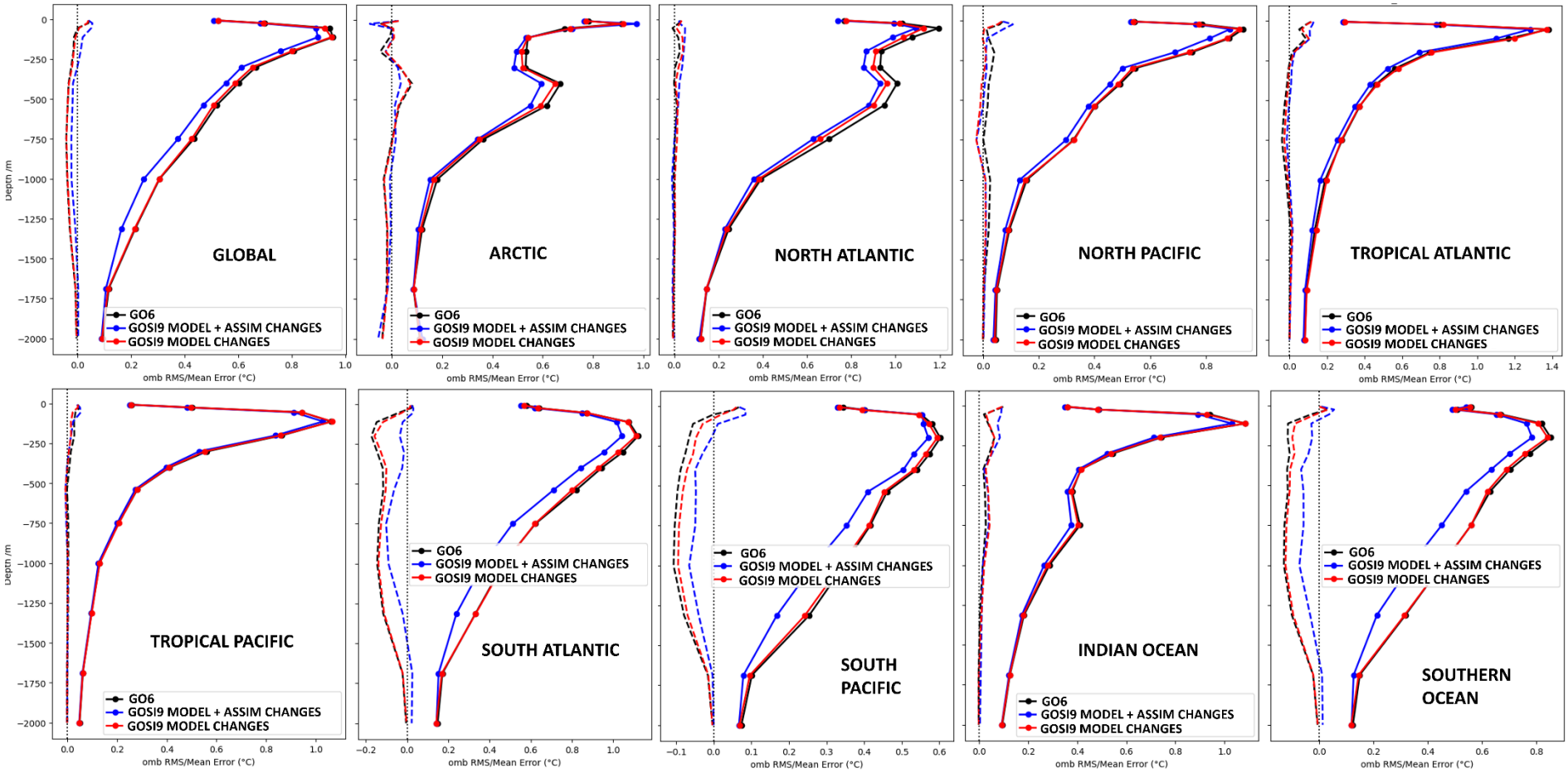

Figure 5A: 2019 observation-minus-background temperature statistics (°C) for GO6 (black), GOSI9 with only model changes (red), and GOSI9 with model plus assimilation changes (blue) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

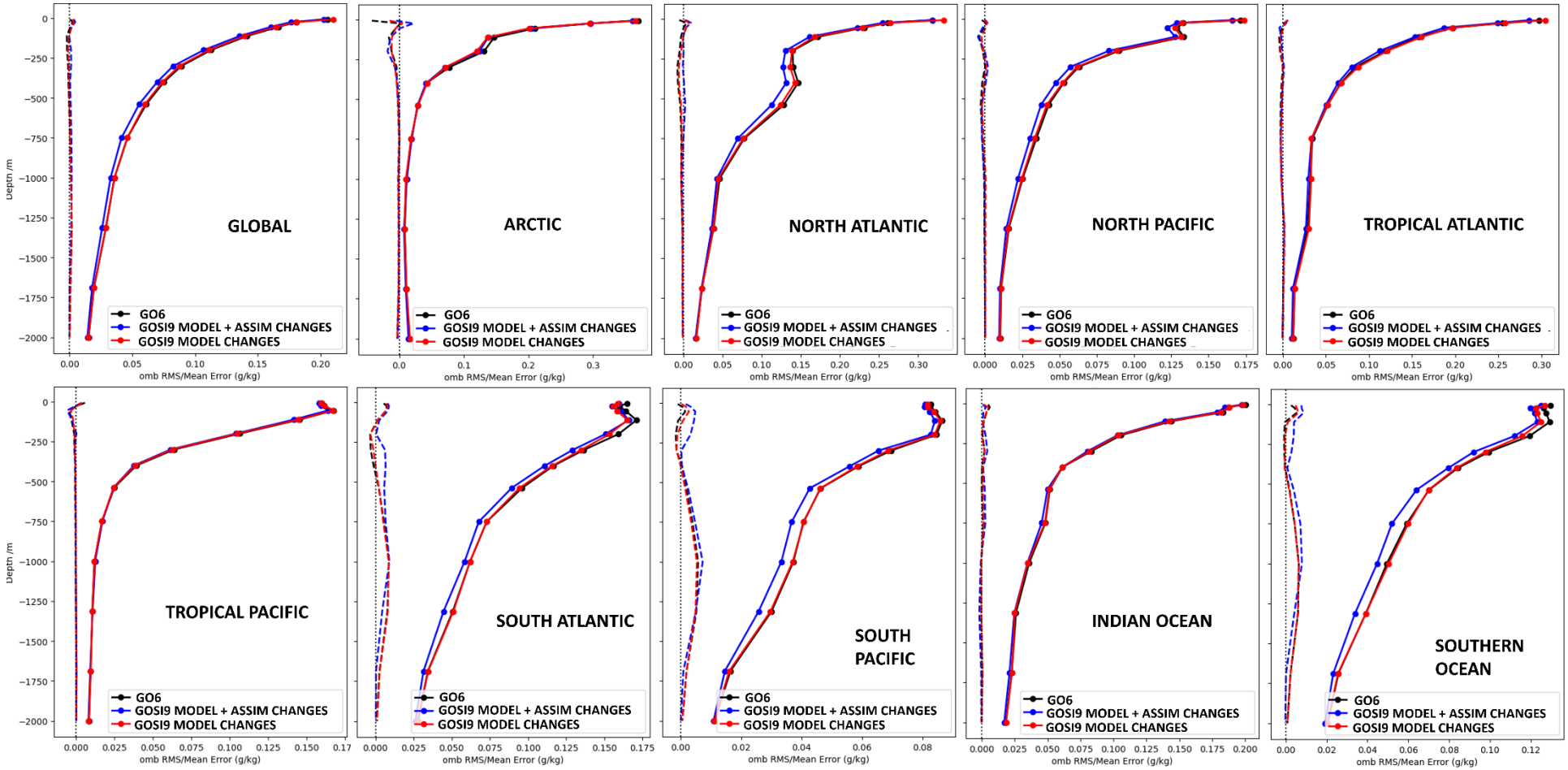

Figure 6A: 2019 observation-minus-background salinity statistics for GO6 (black), GOSI9 with only model changes (red), and GOSI9 with model plus assimilation changes (blue) for different ocean regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

Appendix B – MDT updates

In order to separate the ocean dynamics signal (time varying) from the gravity signal (time constant), altimeter data is provided as SLA with respect to a time mean. To assimilate this data into ocean models requires an MDT, which added to the SLA, gives the observation equivalent to the model SSH.

In the operational suite we have been using the CNES-CLS18 MDT (Mulet et al., 2021) and in the FOAM R&D suite the CNES-CLS13 MDT (Rio et al., 2013; Rio et al., 2014). These MDTs are a combination of the gravity data from the Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) and the Gravity Recovery and Climate Experiment (GRACE) to give the large scales, with in-situ drifter data used to calculate mean geostrophic velocities on small scales. A new global MDT, the CNES-CLS22 (Jousset et al., 2022b), is available which improves on the previous MDTs as it uses more GOCE data, more in-situ data and improved filtering methods. This is combined with a Mediterranean MDT, MDT_CMEMS_2020_MED (Jousset et al., 2022b), given that the correct SSH offset between the Mediterranean Sea and the rest of the ocean is accounted for. The Black Sea is also included. This whole combined MDT is called HYBRID-CNES-CLS22-CMEMS2020. In general, any MDT product will have errors, which will show up as time mean biases. In order to account for this, we include an altimeter observation bias correction scheme (Lea et al., 2008).

We ran a series of trials of the FOAM R&D suite with ORCA025, testing the different MDTs for the period 1 Jan 2019 – 31 Dec 2019. Note for CLS13 trial we used a spun-up altimeter bias but for other MDT tests we started with a zero-altimeter bias. To remove some of the spin-up we excluded the first month of the runs from our assessments (see Table 1B).

Label	MDT
CLS13	CNES-CLS13
CLS18	CNES-CLS18
CLS22	HYBRID-CNES-CLS22-CMEMS2020
CLS22_Medadj*	HYBRID-CNES-CLS22-CMEMS2020*

Table 1B: List of experiments. *As CLS22 but with Mediterranean MDT reduced uniformly by 4 cm.

Preliminary tests with the CLS22 gave similar SLA statistics for all experiments in most regions (Figure 1B).

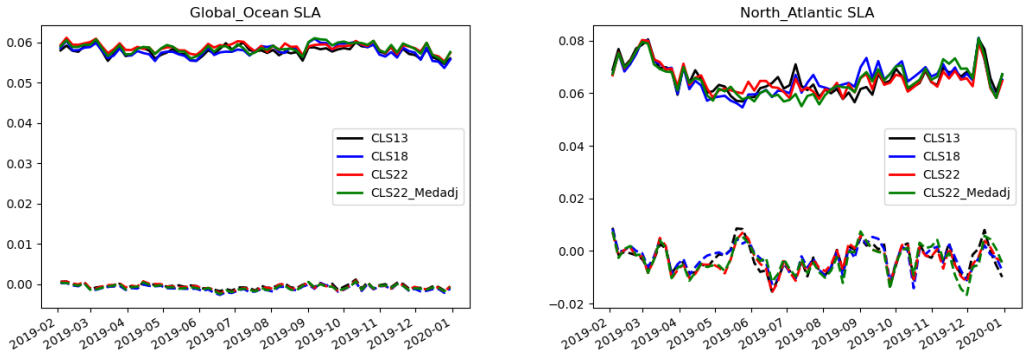

Figure 1B: SLA (m) RMSD (solid lines) and mean (dashed lines) observation minus background time series plots. The left panel shows the globally accumulated statistics and the right panel those in the North Atlantic region.

The results in the Mediterranean Sea were not as good as either the CLS18 or CLS13 MDTs. In particular, there was a large mean offset of 3-4 cm over the whole basin (see Figure 2B).

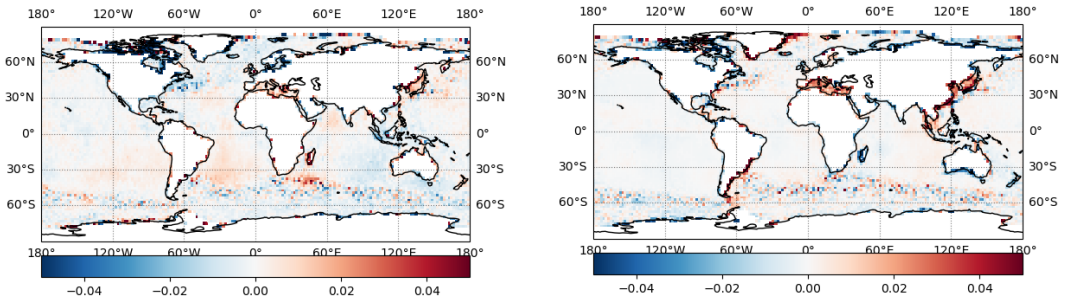

Figure 2B: (Left) Map of SLA (m) mean observation minus background for experiment CLS13. (Right) Difference in SLA (m) mean observation background of CLS22 to CLS13.

This led to a degradation in the profile statistics in the Mediterranean Sea (Figure 3B) with particularly large errors in surface salinity. Note the improvement in the global statistics seen in both CLS18 and CLS22 compared to CLS13. The North Atlantic in particular (Figure 4B) shows the benefits of using the newer MDTs.

All the MDTs are combined with separate Mediterranean MDTs. This needs to be applied with an appropriate offset to the global MDT. For the older MDTs, model mean SSHs were used to calculate the offset, but this was changed in CLS22 and instead the gravity observations themselves were used. This gravity model-based offset may well be more realistic, but our results suggest it is incompatible with our model perhaps due to model bias. A reason for this may be an inaccurate representation of the ocean dynamics in the straits of Gibraltar (13 km wide at the narrowest) with a 1/4° (~25 km) or even 1/12° (~8 km) configuration unable to correctly resolve this. Another potential issue is forcing biases, in particular the freshwater balance in Mediterranean is unlikely to be correct. In view of this, we therefore decided to adjust the CLS22 MDT in the Mediterranean Sea. The approach taken was to use the mean observation minus background for the CLS22 run of 4 cm. Subtracting this from MDT

everywhere in the Mediterranean Sea gives a new MDT CLS22_MedAdj.

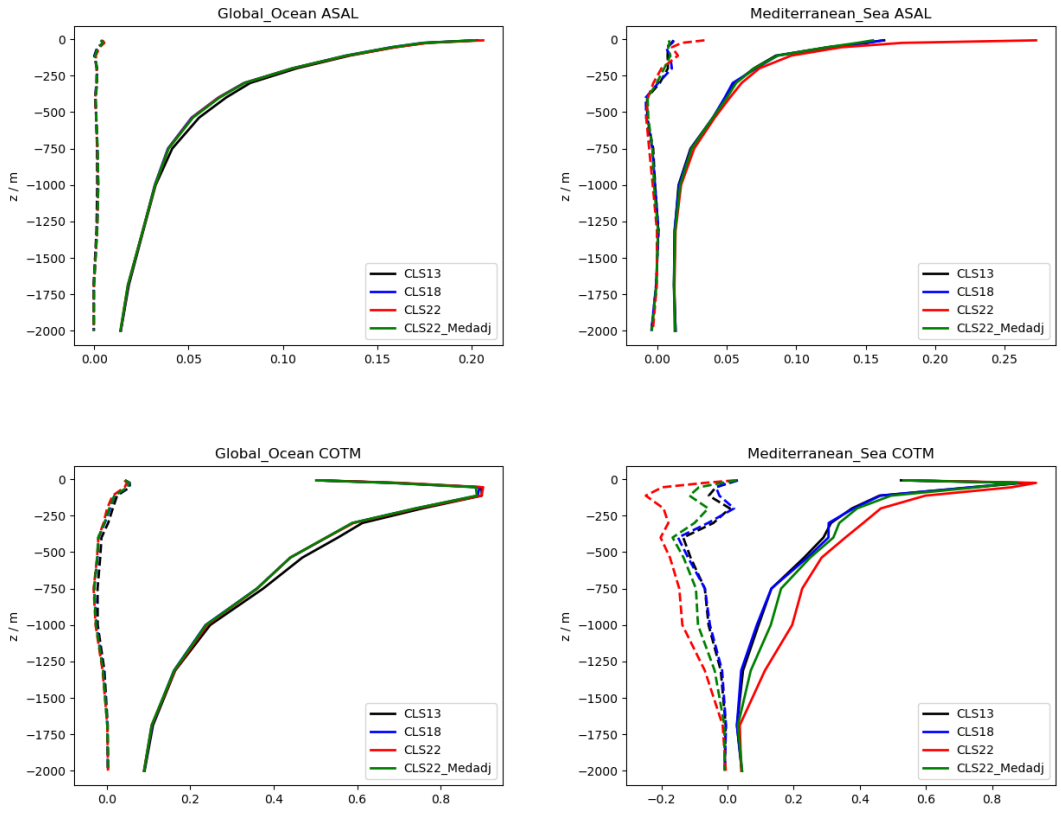

Figure 3B: Absolute Salinity (ASAL) (g kg^{-1}) and Conservative Temperature (COTM) (°C) RMSD (solid lines) and mean (dashed lines) observation minus background as a function of depth. The left panels show the globally accumulated statistics and the right panels those in the Mediterranean region.

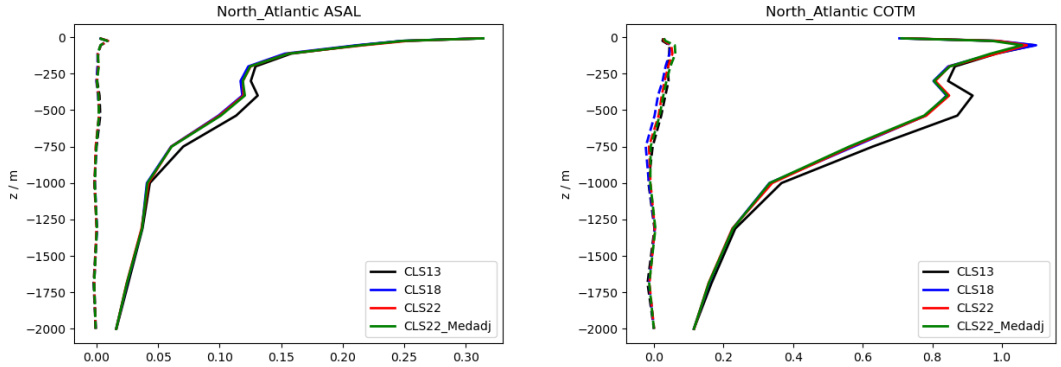

Figure 4B: Absolute Salinity (ASAL) (g kg^{-1}) and Conservative temperature (COTM) (°C) RMSD (solid lines) and mean (dashed lines) observation minus background as a function of depth for the North Atlantic region.

We also saw problems in the Black Sea. Previous MDTs were undefined in the Black Sea. This meant that we were not assimilating altimeter data there. Switching to CLS22 in the SLA assimilation resulted in a large degradation in the profile statistics (Figure 5B). There are likely to be high model biases in the Black Sea similar to those in the Mediterranean due to the difficulties in modelling the flow through the very narrow Bosphorus (1 km to 4 km) with low resolution models. Additionally, since this is the first time we have attempted altimeter assimilation in this region and the Black Sea has a unique vertical structure (with different balance relationships to the global ocean), we decided it was safer to mask this out in our adjusted MDT CLS22_MedAdj.

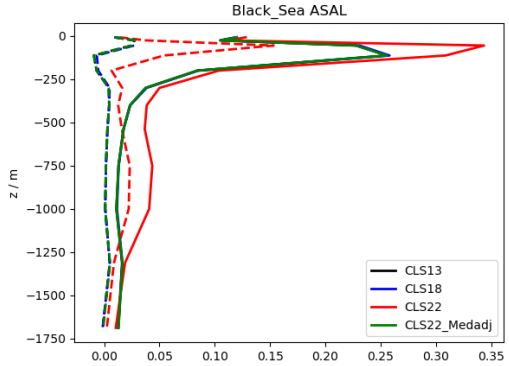

Figure 5B: Absolute Salinity (ASAL) (g kg^{-1}) RMSD (solid lines) and mean (dashed lines) observation minus background as a function of depth for the Black Sea.

Therefore, with CLS22_MedAdj we retain the improved global ocean statistics along with substantial improvements in the Mediterranean Sea in relation to CLS22. We expect the newer MDT to give more benefit at high resolution and near the coast in ORCA12, as well as in even higher resolution shelf models. Trial runs using CLS22_MedAdj in ORCA12 are being conducted to evaluate the impact of this MDT on higher resolution models.

Appendix C – ORCA025 T/S statistics in marginal seas

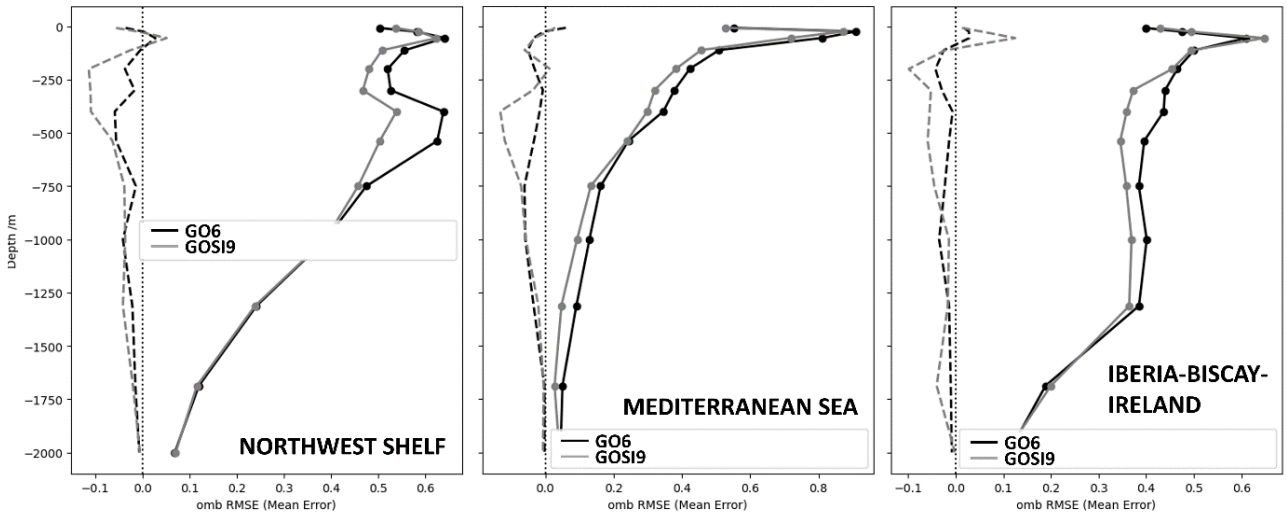

Figure 1C: 2019 ORCA025 observation-minus-background temperature statistics (°C) for GO6 (black) and GOSI9 (grey) for different coastal regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

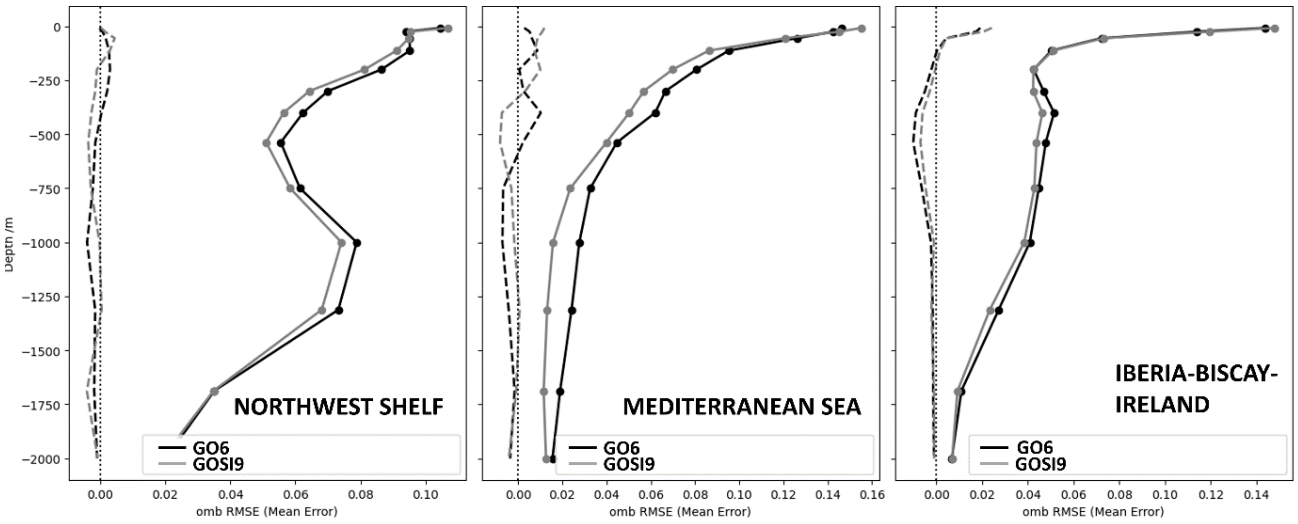

Figure 2C: 2019 ORCA025 observation-minus-background salinity statistics for GO6 (black) and GOSI9 (grey) for different coastal regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

Appendix D – ORCA12 T/S statistics in marginal seas

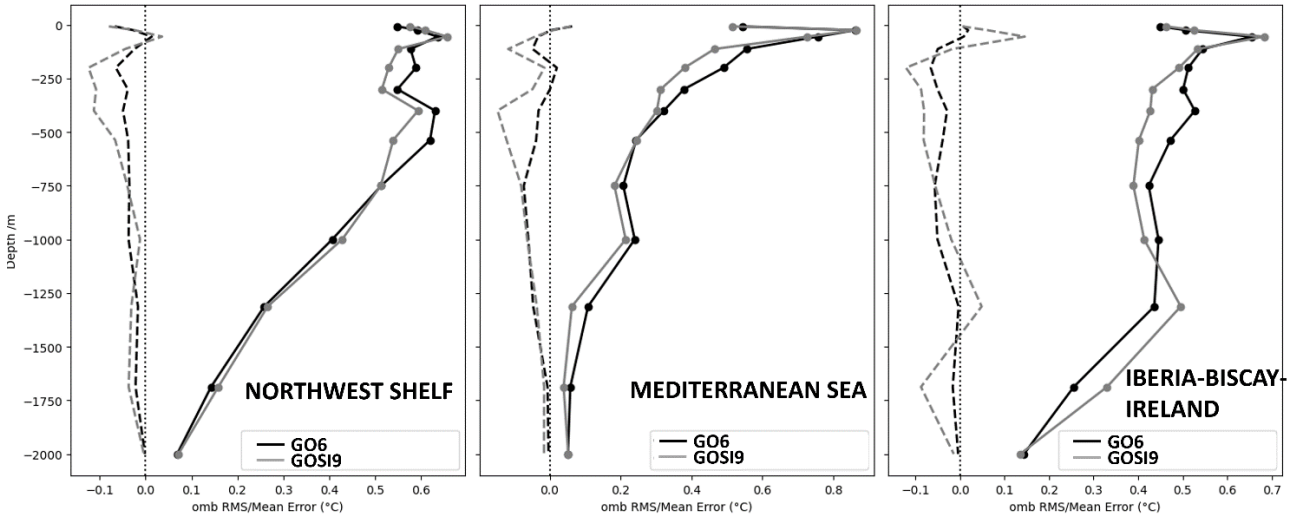

Figure 1D: 2019 ORCA12 observation-minus-background temperature statistics (°C) for GO6 (black) and GOSI9 (grey) for different coastal regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.

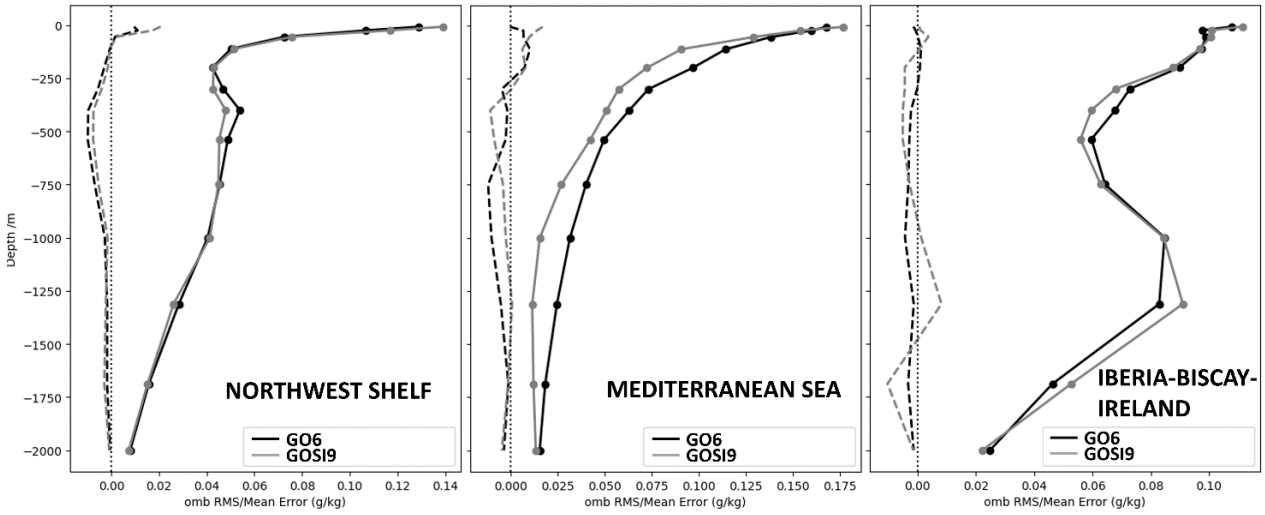

Figure 2D: 2019 ORCA12 observation-minus-background salinity statistics for GO6 (black) and GOSI9 (grey) for different coastal regions, calculated against profile observations. The solid lines correspond to RMSDs, whereas the dashed lines represent mean differences.