Data assimilation sensitivity experiments in the East Auckland Current system using 4D-Var

Rafael Santana[1,2,3], Helen Macdonald[2], Joanne O'Callaghan[1], Brian Powell[4], Sarah Wakes[3], and Sutara H. Suanda[5]

[1]The University of Auckland, Department of Physics, Auckland, 1010, New Zealand
[2]National Institute of Water and Atmospheric Research, Wellington, 6021, New Zealand
[3]University of Otago, Department of Mathematics & Statistics, Dunedin, 9016, New Zealand
[4]Department of Oceanography, University of Hawai'i, Honolulu, HI 96822, United States
[5]University of North Carolina Wilmington, Wilmington, NC, 28403, United States

Correspondence: Rafael Santana (rafacsantana@gmail.com)

Abstract. This study analyses data assimilative numerical simulations in an eddy dominated western boundary current: the East Auckland Current (EAuC). The goal is to assess the impact of assimilating surface and subsurface data into a model of the EAuC via running observing system experiments (OSEs). We used the Regional Ocean Modelling System (ROMS) in conjunction with the 4-dimensional variational (4D-Var) data assimilation scheme to incorporate sea surface height (SSH) and temperature (SST), and subsurface temperature, salinity, and ~~velocities~~ velocity from three moorings located at the upper, mid and lower continental slope using a 7-day assimilation window. Assimilation of surface fields (SSH and SST) reduced SSH root mean square deviation (rmsd) by 25% in relation to the non-assimilative (NoDA) run. The inclusion of velocity subsurface data further reduced SSH rmsd up- and downstream the moorings by 18-25%. By improving the representation of the mesoscale eddy field, data assimilation increased complex correlation between modelled and observed velocity in all experiments by at least three times. However, the inclusion of temperature and salinity slightly decreased the velocity complex correlation. The assimilative experiments ~~had smaller SST rmsd~~ reduced the SST rmsd by 36% in comparison to the NoDA run. The lack of subsurface temperature for assimilation led to larger ~~errors~~ rmsd (>1°C) around 100 m in relation to the NoDA run. Comparisons to independent Argo data ~~showed similar results~~also showed larger errors at 100 m in experiments that did not assimilate subsurface temperature data. Withholding subsurface temperature forces near-surface average negative temperature increments to the initial conditions that are corrected by increased net heat flux at the surface ~~which does not affect waters~~ but this had limited or no effect on water temperature at 100 m depth. Assimilation of mooring temperature generates mean positive increments to the initial conditions that reduces 100 m water temperature rmsd. ~~Larger~~ In addition, negative heat flux and positive wind stress curl ~~was generated~~ were generated near the moorings in experiments that assimilated subsurface temperature data. Positive wind stress curl generates convergence and downwelling ~~which is another way of correcting the upper thermocline cold bias. The larger positive wind stress curl~~ that can correct interior temperature but might also be responsible for decreased velocity ~~correlation in the experiments that assimilated temperature and salinity~~correlations. The few moored CTDs (8) had little impact in correcting salinity ~~, however,~~ in comparison to independent Argo data. However, using doubled decorrelation length scales of tracers and a 2-day assimilation window improved model salinity and temperature in comparison to

~~independent Argo data. In addition, the results were similar to the global reanalysis~~ Argo profiles throughout the domain. This assimilation configuration, however, led to large errors when subsurface temperature data was not assimilated due to incorrect increments to the subsurface. As all reanalyses show improved model-observation skill relative to HYCOM-NCODA ~~which assimilates Argo profiles and was used as boundary condition. HYCOM-NCODA had near zero velocity complex correlation on the mid-slope, whereas all reanalyses showed improved results which highlights~~ (the model boundary conditions), these results highlight the benefit of numerical downscaling to a regional model of the EAuC.

1 Introduction

The East Auckland Current (EAuC) is a western boundary current (WBC) that originates as the reattachment of the subtropical water flow to a continental margin on the New Zealand Northeastern Continental Slope (NZNES) ~~(Stanton et al. (1997) ,~~ (Stanton et al., 1997) (Fig. 1 – see Fig. 5 in Chiswell et al. (2015) for a detailed schematic). The EAuC mean transport was estimated to be 9 Sverdrups (Sv) (Roemmich and Sutton, 1998) with variability at periods longer than 100 days (Stanton and Sutton, 2003). ~~Studies of the EAuC impact on~~ Previous studies suggested that the EAuC intrudes onto the continental shelf ~~found shallow~~ bringing subtropical waters to shallow regions (60 m)~~intrusions of subtropical water~~, possibly driven by ~~EAuC~~ bottom Ekman transport (Zeldis et al., 2004; Santana et al., 2021). Long-term variability ($>$ 100 days) in the EAuC was suggested to be driven by ~~the arrival of~~ baroclinic Rossby waves (Laing et al., 1998; Chiswell, 2001). Santana et al. (2021) observed locally formed mesoscale eddies and their arrival from the east using a one-year time series of in situ and remotely-sensed sea surface height (SSH) and temperature (SST) data. The EAuC is ~~less coherent~~ more eddy-dominated compared to other WBC, where ~~its~~ the ratio between the eddy and mean kinetic energies is larger (0.3) than the East Australian Current's (EAC~~) ratio (Oke et al., 2019).~~ ; 0.1) (Fig. 8 in Oke et al. (2019)). Mesoscale eddies are generated by barotropic and baroclinic instabilities, which ~~are unpredictable~~ limit the predictability of the system (Marchesiello et al., 2003; Feng et al., 2005), and accurate simulation of the EAuC variability needs to incorporate observations, potentially through data assimilation, to realistically represent eddies and other features that vary on short ~~timeframes~~ time-scales (Oke et al., 2005, 2013, 2015; Santana et al., 2020).

Data assimilation (DA) combines observations and numerical ~~models to obtain ocean fields with reduced uncertainty~~model outputs to obtain an improved estimation of the ocean system, called the analysis~~, that better represents the ocean state (e.g. location of mesoscale eddies)~~ . The ocean analysis can be achieved via the Ensemble Kalman Filter method and its variants which have flow-dependent information about the model error statistics but require a large (20+) number of simulations (Bannister, 2017). The ocean analysis can ~~be achieved via~~ also be obtained using calculus of variations through the minimisation, in a least-squares sense, of the difference between model results and observations (Weaver et al., 2003; Di Lorenzo et al., 2007; Moore et al., 2011a). The 4-dimensional variational data assimilation (4D-Var) method reduces ~~model errors~~ model-data misfits over a finite time interval ~~using all observations available~~ (assimilation window) using available observations while preserving dynamical consistency ~~. 4D-Var has notable applications in oceanography , such as: Weaver et al. (2003); Powell et al. (2008); Mazl . Some of those~~ and it has been used in notable studies in oceanography (Powell et al., 2008; Zavala-Garay et al., 2012; Kerry et al., 2016; P

. These studies are realistic 4D-Var applications using the Regional Ocean Modelling system (ROMS), from which different values of ~~time-window~~ assimilation-window length, observational errors, and decorrelation length scales were used for the tests ~~presented in the~~ performed in this current study.

Observing system experiments (OSEs) aim to assess the importance of distinct sets of observations on the quality of different analyses (Oke et al., 2015). By withholding observational subsets, one can estimate the importance of those data in the ocean reanalysis. For instance, Zavala-Garay et al. (2012) used ROMS 4D-Var to show that assimilation of ~~sea surface temperature (SST) and height (SSH) data only~~ SST and SSH data increased temperature error between 350 and 750 m in comparison to a non-assimilative run. The authors found that ~~assimilation of XBT or synthetic CTD data were needed to correct this~~further assimilating temperature from XBTs or synthetic temperature generated from its relation with SSH was needed to constrain the degradation in temperatures between 350 and 750 m. Pasmans et al. (2019) stated that assimilation of temperature and salinity from ocean gliders should be accompanied by surface measurements to prevent the generation of unrealistic instabilities. Siripatana et al. (2020) compared two ~~reanalysis~~reanalyses: (i) assimilated traditional observations only (satellite SSH and SST, and temperature and salinity from Argo floats) and (ii) assimilated traditional observations and data from moorings, gliders and high-frequency radar; and found improved subsurface results of temperature and velocity in the latter experiment.

The EAuC system lacks traditional subsurface data, where Argo profiles are rare and the region can experience months without any water column measurement. Between May 2015 and May 2016, the EAuC region was sampled along Topex/-Poseidon 147 line, ~~when~~ and five mooring lines (named M1 to M5) were deployed ~~from coastal waters to~~ in coastal waters and the continental rise ~~which sampled~~ to measure velocity, temperature and salinity (Fig. 1) and were analysed in Santana et al. (2021). ~~Sampling efforts like those are costly, and deciding~~ During this one year of observations, Santana et al. (2021) identified distinct periods of mesoscale activity in the region driven alternately by anti- and cyclonic eddies, as well as a period of encroachment of the EAuC jet towards the continental slope. Intensive offshore sampling efforts such as these are potentially cost-prohibitive. Therefore, determining the most important variables and locations of interest for ocean sampling is key in the region, ~~specially in this eddy dominated WBC. Moreover, the~~ especially in this eddy-dominated WBC. The EAuC also showed strong velocity shear on the continental mid-slope (Santana et al., 2021), which might indicate the need for subsurface data for accurate velocity simulation.

The goal of this study is to evaluate the impact of ~~subsurface data assimilation~~ assimilating subsurface observations into a model of EAuC system. We conducted a set of OSEs where the most complete simulation assimilated surface fields (SSH and SST), and mooring velocity, temperature and salinity (ASFUVTS) (Fig. 1). The other experiments suppressed observation types in the assimilation algorithm. They withheld velocities (NoUV); subsurface temperature and salinity (NoTS); and all mooring data (NoUVTS). As control, we examined a non-assimilative ~~freely evolving~~ simulation (NoDA). Argo data was left out of all experiments for independent model-data comparison. Sensitivity tests on decorrelation length scales and assimilation window were also conducted which helps deciding the best assimilation configuration for an operational forecast of the EAuC.

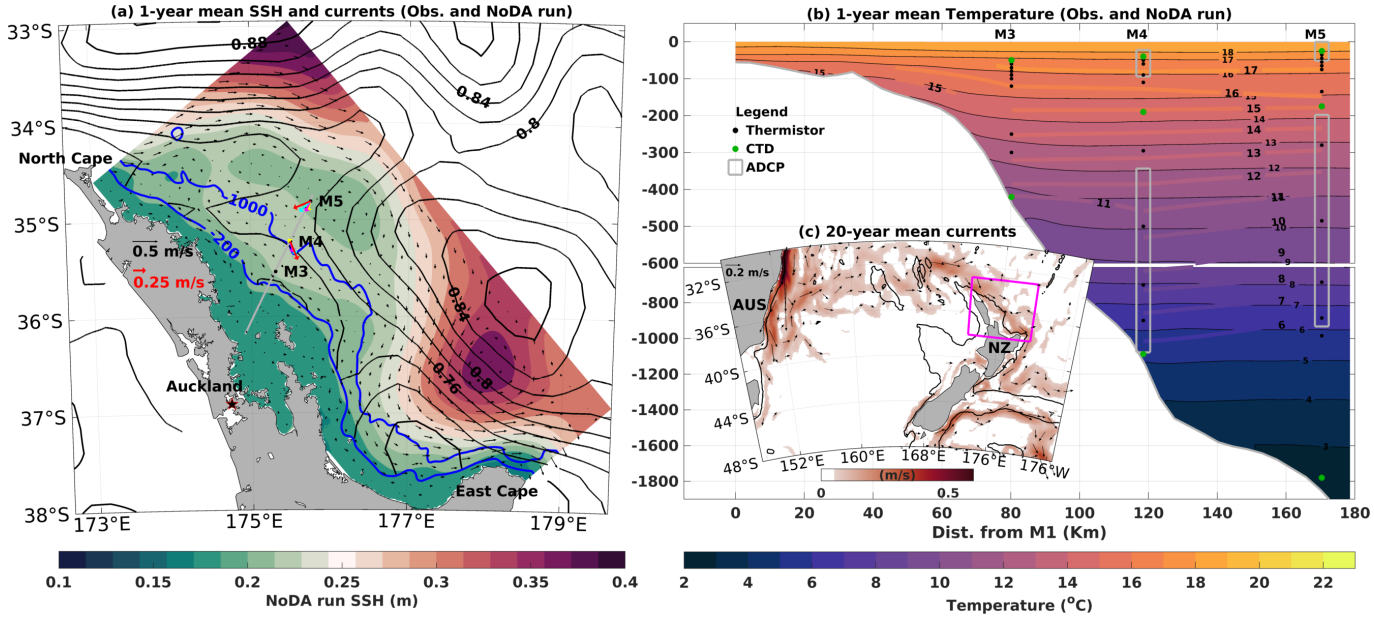

Figure 1. (a) Study area showing the one-year average SSH from AVISO (black contours) and NoDA run (green-red shade), ~~geostrophic velocities from AVISO (blue arrows)~~ and near-surface velocities NoDA run (black arrows). The stations M3 to M5 are indicated along the grey line. The coloured arrows represent the vertically averaged (bin = 280 m) ~~*in situ*~~ in situ velocities ~~centered~~ centred at 140 m (red), 420 m (cyan), 700 m (blue), and 980 m (yellow). ~~The white arrows represent the orientation axes used. The stations M3 to M5 are indicated along the grey line. Integer numbers represent Argo locations for independent data comparison.~~ (b) Temporal averages of temperature from in situ measurements (coloured contours) and the NoDA run (coloured shade). The grey rectangles show regions sampled by ADCPs at M4 and M5 and ~~used in data assimilation. The~~ the vertical positions of thermistors and CTDs are indicated as black and green dots, respectively. (c) Location of the study area (magenta contour) ~~relative to other currents~~ in the ~~Southwestern~~ Southwest Pacific Ocean. The ~~colors~~ colours and arrows represent the ~~one-year~~ 20-year average of geostrophic ~~current speed~~currents. Current speeds lower than 0.05 m/s were masked. The 200- and 1000-meter isobaths are shown as blue and ~~white~~ black contours in (a) and (c).

2 Methods

2.1 Numerical model

We use the Regional Ocean Modeling System (ROMS), a primitive-equation, hydrostatic, and free-surface ocean model that solves the Reynolds-averaged form of the Navier–Stokes equations. ROMS is a fully nonlinear, finite-difference model that uses terrain-following (sigma) vertical coordinates and horizontal orthogonal ~~or curvilinear Arakawa C grid~~ curvilinear coordinates on a staggered Arakawa C-grid (Shchepetkin and McWilliams, 2003, 2005; Haidvogel et al., 2008). The model domain (290 x 150) is rotated 52.14° clockwise to better resolve the NZNES and spans 332 km offshore at the widest point (near North Cape) (Fig. 1). The domain has a horizontal resolution of approximately 2 km, which roughly captures coastline variability and still resolves the continental shelf and slope without large computational cost. The model has 30 vertical sigma layers and

model bathymetry was interpolated from the 250 m resolution bathymetric data set built by the National Institute of Water and Atmospheric Research (NIWA - https://niwa.co.nz/our-science/oceans/bathymetry). We use a vertical discretisation scheme that increases the resolution near the surface and bottom by applying stretching function type 4 and transformation equation option 2 (Shchepetkin and McWilliams, 2005, 2009). The vertical resolution is higher at the upper 200 m (~~from 4 to 30 layers~~4th - 30th layer). On the slope (depth<1000 m), the vertical resolution is higher than 66 m and in the open ocean the thickest level is 233 m (3838.6 m depth). The effect of bottom friction is parameterised using a constant drag coefficient of 3.0×10^{-3} (Lentz, 2008), and the vertical mixing of tracers and momentum is done with K-ϵ (Rodi, 1987), using the generic length scale (GLS) scheme (Umlauf and Burchard, 2003). Baroclinic modes are resolved using a time step of 180 s, while the barotropic time step is 6 s.

Model surface forcing is from the Japanese atmospheric 55-year reanalysis for driving ocean models (JRA55-do, Tsujino et al. (2018)). A previous study demonstrated that JRA55-do had the highest correlation with observed winds in comparison to other atmospheric forcing datasets in the Southwest Pacific (Taboada et al., 2019). Atmospheric forcing fields of wind speed, net shortwave radiation, downward longwave radiation, relative humidity, temperature, rain, and pressure are ~~specified~~ provided every 3 h and are used to compute the surface fluxes of stress, heat and freshwater using the bulk flux parameterisation of Fairall et al. (2003). The model uses initial and boundary conditions of SSH, temperature, salinity, and velocities from HYCOM-NCODA (Chassignet et al., 2009) versions 91.1 and 91.2 which cover ~~period of simulations generated here. de Souza et al. (2021) analysed the performance of four global reanalysis that assimilate SSH, SST and Argo data on the New Zealand waters. The authors found that HYCOM-NCODA (8 km resolution) had higher SSH and SST variability than low-resolution (25 km) satellite surface observations and other reanalyses with similar grid-spacing. HYCOM-NCODA had velocity standard deviation similar to that observed at M4 and M5, even though GLORYS (8 km global ocean reanalysis, (Lellouche et al., 2018)) better represented temperature and salinity profiles in the region.~~ the period focus of this study. Annual average discharge from several rivers are included as lateral forcing in the model. The boundary forcing is applied daily using Chapman (1985) condition for free surface, Shchepetkin condition (Mason et al., 2010) for barotropic velocities and mixed radiation-nudging (Marchesiello et al., 2001) for baroclinic velocities, temperature and salinity. A 5-day nudging coefficient is applied towards the lateral boundaries. The model aims to simulate continental shelf, slope and rise regions, including the offshore extent of the EAuC and its eddy variability. ~~This and the upcoming studies using the simulations analysed here focus on intra-annual variability and tides are not included as forcing~~Tidal variability is not included in this study.

2.2 Observations

Remotely-sensed observations of SSH and SST were used in all assimilative runs ~~of the current~~ in this study. Optimally interpolated gridded maps of daily SSH with 1/4° of horizontal resolution created by AVISO (Ducet et al., 2000) were assimilated into the model. Daily SSH offsets (annual mean of ∼51 cm) were calculated prior to every assimilation cycle as the difference between the spatial average of observations and model daily averages for each assimilative experiment (one-year averages are shown in Fig. 1a as examples of this offset). ~~This offset was~~ These offsets were removed from AVISO observations before assimilation to ensure the comparisons were made on the same reference surface. Daily maps of SST (1/4°) from AVHRR

Pathfinder (Casey et al., 2010) were used for assimilation into the model at a depth of 2 m. SSH and SST data were assimilated daily at 12:00 pm UTC in regions deeper than 200 m. Data points that are within 20 grid points of the boundaries were removed and a total of 123 data points of SSH (or SST) were assimilated per day, with roughly an SSH/SST measurement every ~~12^{th}~~ 12th grid point. We used an assimilation window of seven days (see details in section 2.3) which gives a total of 861 SSH/SST data points per assimilation cycle (~~Fig. ??~~Table 1).

A cross shelf-slope mooring transect collected data from the 6th of May 2015 ~~and~~ to the 21st of May 2016 (Fig 1a,b). In situ observations of temperature, salinity, u and v components of velocity from ~~moorings~~ 3 moorings (M3, M4, and M5) were assimilated into the model (Fig 1b). MicroCAT CTDs were located near the surface and bottom at the three stations and two extras CTDs were located around 200 m at M4 and M5 (green dots in Fig 1b). 26 temperature sensors were evenly distributed between the three stations, with higher density of instruments in the upper 200 m water column (black dots in Fig 1b). Two ~~long range~~ long-range (LR) and two ~~short range~~ short-range (SR) ADCPs were located at M4 and M5 (grey rectangles in Fig 1b). The water column was binned every 15 m (LR) or 4 m (SR) by the ~~upward looking~~ upward-looking ADCPs. Velocity (temperature and salinity) measurements were taken during a period of 2 min. (1 min.) every 10 min or less. The dataset is freely available and can be found in O'Callaghan et al. (2015). All in situ observations were low-pass filtered at a period of 30h to remove tidal signals and their interaction with other processes, as in Kerry et al. (2016). The observations were averaged at time bins of 6 h, centred at 3, 9, 15 and 21 UTC. ~~Mooring velocity observations(magenta line in Fig. ??) were available in the first cycle and were followed by temperature and salinity (cyan and green lines in Fig. ??). Velocity had the largest amount of mooring observations but stratification information (SST, mooring temperature and salinity) was larger when combined (blue line in Fig. ??).~~ Types of observations, their sources and median number of observations per assimilation cycle are shown in Table 1.

~~Non-assimilated~~

Table 1. Types of observations assimilated, their source and median number of observations per assimilation cycle.

Observation type	Obs. source	Obs. error (std)	Median number of obs. per 7-day assim. cycles
SSH	AVISO	0.04 m	861
SST	AVHRR	0.3 °C	861
Velocity	Moorings M3-M5	0.15 m/s	1493
Temperature	Moorings M3-M5	0.8 °C	808
Salinity	Moorings M3-M5	0.16 g/kg	224

Independent data from 30 Argo profiles (Roemmich et al., 2019) ~~sampled the region encompassed by the model domain (integer numbers in Fig. 1a) during the simulation period and~~ were used for model-data independent comparison. More details on model evaluation are described in section 2.4.

~~Number of observations (after quality control and assurance, filtering and averaging) used in each 7-day assimilation cycle. Observations types are shown by colour: AVISO SSH (red), AVHRR SST (black), mooring u and v components of velocity~~

(magenta), mooring temperature (cyan), mooring salinity (green), combined SST, mooring temperature and salinity (blue). Mooring data was used from stations M3, M4 and M5.

2.3 Data assimilation

Data assimilation is applied in a series of time windows using strong constraint 4D-Var, i.e. neglecting model errors (Di Lorenzo et al., 2007; Moore et al., 2011a). Tests using 2, 3, 4, and 7 days as assimilation window were conducted and the OSEs with a 7-day window had the most realistic results in comparison to observations. Results from a 2-day window reanalysis were also analysed in the current study. In this work, 4D-Var is used to adjust the control variables for including initial, boundary, and atmospheric conditions. The ocean analysis is obtained via minimisation of model-to-data discrepancy or cost function (J) given by:

$$J(\delta\mathbf{z}) = \frac{1}{2}\delta\mathbf{z}^{\mathrm{T}}\mathbf{D}^{-1}\delta\mathbf{z} + \frac{1}{2}(\mathbf{G}\delta\mathbf{z}-\mathbf{d})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{z}-\mathbf{d}) \tag{1}$$

where $\delta\mathbf{z}$ is the state vector constituted of increments (corrections) to the initial $\delta\mathbf{x}(t_0)$, lateral $\delta\mathbf{b}(t)$, and surface $\delta\mathbf{f}(t)$ conditions. $\mathbf{D}$ and $\mathbf{R}$ are the background and observation error covariance matrixes, respectively. Superscripts T and -1 represent the transpose and inverse operations, respectively.

$\mathbf{G} \equiv \mathbf{H}\mathbf{M}_f$, where $\mathbf{H}$, is the linearised version of the observation function $\mathcal{H}$ that maps the model-state to the observation time and locations. The operator $\mathbf{M}_f$ denotes the tangent linear operator of the model integration about the forecast (denoted by the subscript f) over the assimilation window. For the transpose, $\mathbf{G}^{\mathrm{T}} \equiv \mathbf{M}_f^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}$, $\mathbf{H}^{\mathrm{T}}$ maps from observation to model-space and the $\mathbf{M}^{\mathrm{T}}$ operation integrates backward in time over the assimilation window.

$\mathbf{G}\delta\mathbf{z}-\mathbf{d}$ represents the mismatch between the tangent linear model fields mapped to the observations ($\mathbf{G}\delta\mathbf{z}$) and the observations $\mathbf{d}$ which is called the innovation vector and is defined as the difference between observations and model background interpolated to observation location. SSH observations had a daily bias (or offset) between data and model reduced from the observations before the innovation values were computed.

The combination of control variables $\delta\mathbf{z}_a$ that minimizes $\boldsymbol{J}$ is yielded iteratively in the subspace spanned by the linear combinations of the observed model variables. The method chosen in the present work is the physical-space statistical analysis system (4D-PSAS) (Moore et al., 2011a), which and the algorithm that minimises the cost function is shown in Fig. 2 on Moore et al. (2011a). 4D-PSAS defines $\delta\mathbf{z}_a$ as:

$$\delta\mathbf{z}_a = \mathbf{D}\mathbf{G}^{\mathrm{T}}(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}}+\mathbf{R})^{-1}\mathbf{d} \tag{2}$$

$\delta\mathbf{z}_a$ can be equivalently written as $\delta\mathbf{z}_a = \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{w}^{\mathrm{a}}$ where $\mathbf{w}^{\mathrm{a}}$ is the sub-space of the model state vector spanned by the observations (i.e. the dual space) and satisfies:

$$(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}}+\mathbf{R})\mathbf{w}^{\mathrm{a}} = \mathbf{d} \tag{3}$$

The dual form has the advantage that the dimension of $\mathbf{w}^a$ is equal to the number of observations which is, in our case, several orders of magnitude smaller than the dimension of the full state vector. Thus, solving (3) may be less demanding than solving equation (2) (Moore et al., 2011a).

In practice, we find an acceptable reduction of the model-to-data discrepancy after 20 iterations (inner loops), when atmospheric and lateral forcings are adjusted every 12 h along with initial conditions at the beginning of the assimilation cycle. For further details of the method and its application, readers are referred to Moore et al. (2011a, b).

The ~~background~~ diagonal background/prior error covariance matrix **D** cannot be completely calculated or stored, it is rather estimated via factorisation (Weaver and Courtier, 2001). **D** takes into consideration the background error standard deviations ~~,~~ of initial, boundary and forcing conditions, spatial decorrelation scales ~~,~~ and normalisation factors ~~(Moore et al., 2011b) . Background~~ (Moore et al., 2011a). The background error standard deviations were calculated from the average of 4-day variances computed from ~~2~~ two years of the NoDA run ocean (initial conditions), boundary, and forcing fields. Horizontal decorrelation length scales for SSH, velocities, active tracers (temperature and salinity) were 100, 50 and 100 kilometres. Vertical decorrelation length scales for velocities and active tracers were 50 and 100 metres. The normalisation factors are the costliest part of the covariance modelling but they are computed only once for each combination of background error standard deviations and spatial decorrelation length scales. The normalisation factors were estimated via randomisation (Fisher and Courtier, 1995) using 7500 iterations. ~~The described background error configuration is valid for the 7-day reanalyses which differs in three aspects from the 2-day reanalysis – see section 2.4 for details .~~ Fig. 3 in de Paula et al. (2021) shows an example of convolution of a unit impulse function with the horizontal (vertical) decorrelation length scale set to 100 km (50 m). For more details on the computation of the background error covariance matrix **D**, please see Moore et al. (2011a).

The observational error covariance matrix **R** is assumed to be diagonal. The standard deviation used was the largest value between an assigned standard deviation and the observation error obtained from the AVISO and AVHRR ~~proucts~~products, with the assigned value being the highest most of the time. This strategy was applied to satellite observations, whereas in situ data used a given standard deviation only. According to values used in the literature (e.g. Moore et al. (2011b); Kerry et al. (2016)), different standard deviations were tested and the values that gave the best results were: 0.04 m for SSH; 0.3 °C for satellite SST; 0.8 °C for subsurface temperature; 0.16 g/kg for subsurface salinity; and 0.15 m/s for water column u and v components of velocities. ~~Large~~ A large standard deviation was attributed to in situ temperature due to strong internal tides in the region and ~~blow down~~ blow-down events that often happened at the three moorings. Depth of the thermistors was obtained assuming an inverted pendulum relationship between CTDs located near the surface and bottom (Stanton and Morris, 2004) but a level of uncertainty remains in the depth of measurements. A summary of the observation standard deviation errors and the median number of observations per 7-day assimilation cycles is shown in Table 1.

2.4 Experiment design and evaluation metrics

~~Five data assimilative~~ Four data assimilation experiments and one free-running simulation (no data assimilation; NoDA) ~~were conducted~~are part of the OSEs. The first data ~~assimilative~~ assimilation experiment incorporated SSH, SST, u and v components of velocity and ~~temperature and~~ mooring temperature and salinity ~~data~~ (ASFUVTS). The other data assimilative experiments

were similar but withheld: subsurface temperature and salinity (NoTS); ~~mooring~~ subsurface velocities (NoUV); and all mooring data (NoUVTS). Another experiment assimilating surface and all mooring data was also conducted using a 2-day assimilation window (ASFUVTS-2days) which was designed to better match the observations. This experiment (ASFUVTS-2days) differed from the other simulations in three aspects. It used 7-day variances to compute background error standard deviations and 200 kilometres (metres) as horizontal (vertical) decorrelation length scales of active tracers (temperature and salinity). The last difference is the assimilation window length which is two days. These modifications made ASFUVTS-2days less comparable to the rest of the ~~OSE~~ OSEs but they were needed in order to achieve the best match between assimilated and independent observations – for results, see sections 3.2 and 3.3. Two other 7-day assimilation window experiments (ASFUVTS-2x and NoUVTS-2x) were run assimilating the same observations as in ASFUVTS and NoUVTS but with the with the background error covariance matrix used in ASFUVTS-2days (doubled decorrelation length scales and 7-day variances). Those experiments showed that using doubled decorrelation length scale of tracers can lead to larger errors in temperature if subsurface data is not available. Table 2 summarises the experiments' configurations.

Table 2. Assimilation experiments and their observations, decorrelation length scales of temperature and salinity and assimilation window.

Experiments	Assimilated observations	Horizontal (vertical) decorr. length scales of tracers	Assim. window
ASFUVTS	SSH, SST, velocity, temperature, and salinity	100 km (metres)	7 days
NoTS	SSH, SST, and velocity	100 km (metres)	7 days
NoUV	SSH, SST, temperature and salinity	100 km (metres)	7 days
NoUVTS	SSH and SST	100 km (metres)	7 days
ASFUVTS-2days	SSH, SST, velocity, temperature, and salinity	200 km (metres)	2 days
ASFUVTS-2x	SSH, SST, velocity, temperature, and salinity	200 km (metres)	7 days
NoUVTS-2x	SSH and SST	200 km (metres)	7 days

The numerical simulations started on the 1st of May 2015 using an ~~interpolated initial condition~~ initial condition interpolated from HYCOM-NCODA and they ran until the 31st of May 2016. The simulations started assimilating SSH and SST observations on the 1st of May, and mooring data between the 8th of May 2015 and the 21st of May 2016. The simulations assimilated SSH and SST until the 31st of May 2016 and 57 (200) assimilation cycles were performed in the 7-day (2-day) reanalyses. The NoDA run was integrated until 31st December 2016. Model results were interpolated to the observations' spatial resolution for evaluation.

The NoDA run and reanalyses were objectively validated using root mean square deviation (rmsd) given by:

$$rmsd\text{rmsd} = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(x_i - y_i)^2}; \tag{4}$$

and linear correlation (r):

$$rr = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}; \quad (5)$$

between observed (x) and modelled (y) results, where $i=1,2,...,n$ are the observation times or locations and the averages $^-$ were applied in time or space. The daily offset between observed and model SSH was removed before performing the SSH rmsd calculation. SSH and SST rmsd were computed using daily averaged model results (analysis) interpolated to observations locations inside the domain. Complex ~~correlation (Kundu, 1976) was calculated~~ vector correlation was computed between simulated and observed 2D velocity vectors. Complex correlation converts a pair of two-dimensional vector series into complex numbers and computes the correlation between their real parts and their relative angular displacement using their imaginary parts (Kundu, 1976). Model bias is defined as the difference between simulated and observed quantities. Statistics computed between model and in situ observations used daily averaged data processed and studied in Santana et al. (2021). ~~30~~ Thirty Argo profiles were available inside the model domain (~~integer numbers in Fig. 1a~~see section 3.3 for details) during the reanalyses period and these were used to provide independent temperature and salinity ~~data~~ observations for model-data comparison. Argo data and model results were linearly interpolated from 0 to 2000 m depth using bins of 10 m. HYCOM-NCODA analyses which assimilated SSH, SST and Argo profiles (Chassignet et al., 2009) were also ~~used in the comparison~~ compared against subsurface mooring and Argo data.

3 Results

3.1 DA impact on surface fields

~~A skillful state estimate should have rmsdto observations smaller than the typically observed~~ Documented periods with anti- (A1 and A2) and cyclonic (C2) mesoscale eddies (Santana et al., 2021) were used to compare the NoDA (no data assimilation – control) and ASFUVTS (assimilating SSH, SST, velocity, temperature and salinity data) runs. ASFUVTS' SSH height fields showed good agreement with observed SSH structures, where the centre of model high and low SSH tended to match with the core of anti- (A1 and A2) and cyclonic (C2) eddies respectively (Fig. 2a,c,e). The NoDA run generated anti- and cyclonic eddies close to the observed ones, however, the eddies' centres were displaced by tens of kilometres (Fig. 2b,d,f). This result can be attributed to SSH and velocity boundary conditions from HYCOM-NCODA which provided accurate SSH at the border and forced the generation of eddies with the correct cyclonicity albeit displaced from the observed eddies. Smaller eddies were modelled in the two simulations due to their higher resolution (~2 km) in comparison to the observations (~25 km) (Fig. 2c,d). Mesoscale eddies that had their centre near the domain boundary were not well simulated by either experiment, such as the cyclonic eddy centred at 34°S and 174°E on 2015/10/09 and the anticyclonic eddy centred at 36°S and 178°E on 2016/04/09 (Fig. 2a,b,e,f). A statistical analysis of the representation of mesoscale eddies (SSH and SST fields) by the numerical experiments is shown in the rest of this section.

The root mean square deviation (rmsd) was calculated in time and space against observations to objectively evaluate SSH and SST generated by the experiments. The rmsd is compared to the observations' standard deviation (std) ~~.~~ which shows

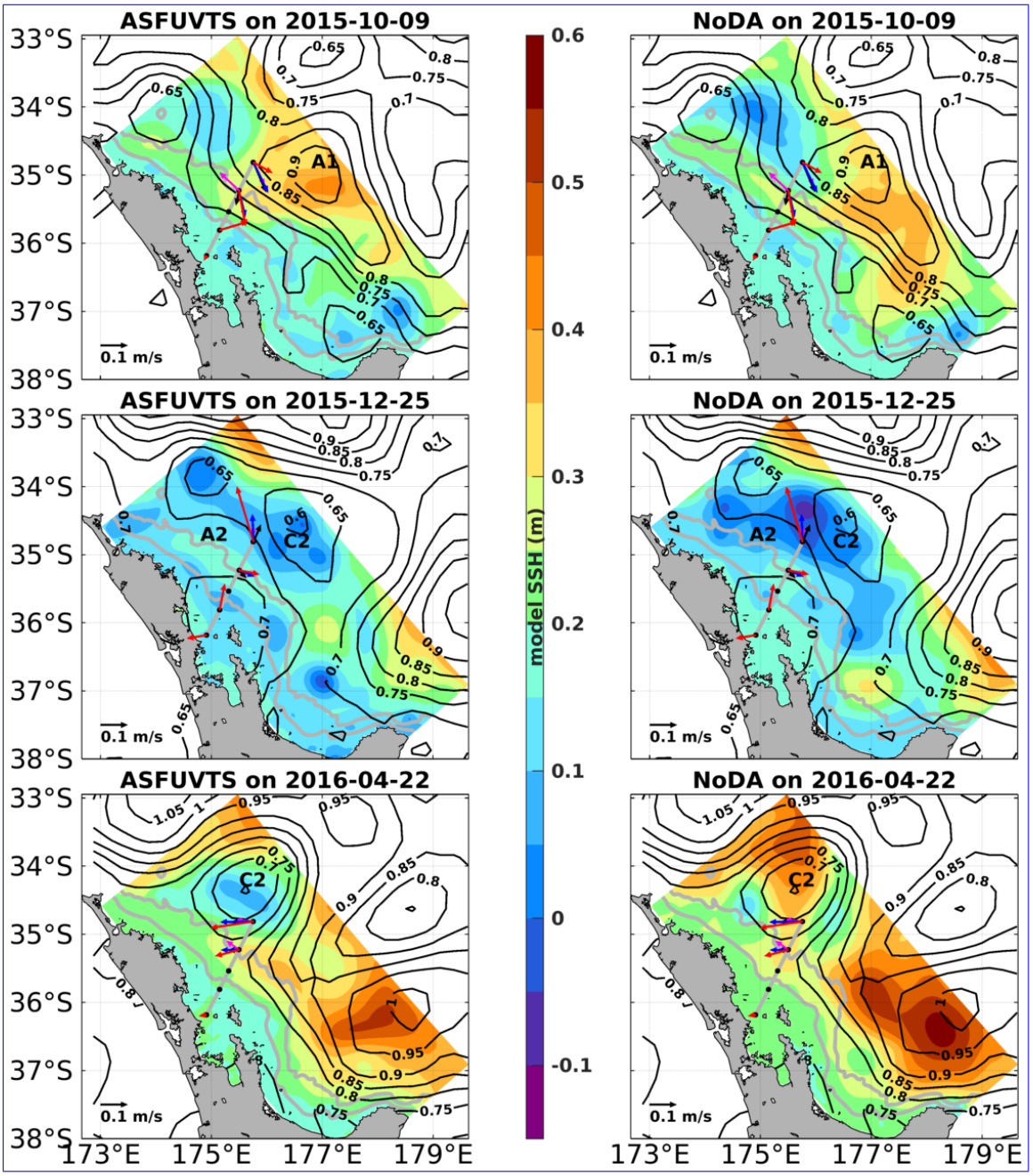

Figure 2. Daily maps of observed SSH from AVISO (black contours), and modelled SSH (coloured shade) from experiments ASFUVTS (left column) and NoDA (right column) on days 2015/10/09 (a,b), 2015/12/25 (c,d), and 2016/04/22 (e,f). The acronyms A1, A2, and C2 represent mesoscale eddies studied in Santana et al. (2021). The coloured arrows show in situ velocities as in Fig. 1a).

regions or periods of larger mesoscale variability that might be harder to accurately simulate observed SSH and SST. All experiments had reduced SSH rmsd on the slope where observed SSH std was smaller and SSH rmsd tended to be higher in the open ocean (larger SSH std) where mesoscale eddies are present (Fig. 3a,b,c,d). The most complete ocean analysis (ASFUVTS) had ~~smaller SSH rmsd~~ reduced SSH rmsd (mean = 0.06 m) in comparison to the ~~observed SSH std for most~~ NoDA SSH rmsd (mean = 0.08 m) for the majority of the domain (Fig. 3a,b). The NoUVTS run had similar average SSH rmsd (0.06 m) to the ASFUVTS run. However~~ASFUVTS had~~ , ASFUVTS further reduced SSH rmsd upstream and downstream of the moorings (Fig. 3b,c). Experiments that assimilated in situ velocities (NoTS, ASFUVTS-2days) also had positive impact on SSH representation up- and downstream of the moorings (not shown). NoDA had ~~larger SSH rmsd~~large SSH rmsd, especially in areas of observed high SSH std – north and east of the moorings and north of the East Cape (178°E, 37°S) (~~0.08-0.14 m) in comparison to observed std (0.07-0.11 m) in regions deeper than 1000 m (~~Fig. 3a,d)~~, where more than 4 mesoscale eddies with life span larger~~ . The dynamics in those regions were largely dominated by more than four mesoscale eddies that persisted for more than a month ~~were observed~~ each (Santana et al., 2021).

The NoDA run had the highest average SST rmsd (0.72 °C) ~~smaller than the average observed SST std (2.39 °C) (Fig. 3e,h). This skill was achieved by the NoDA run's ability to reproduce the seasonal cycle (amplitude of 6 °C), however this simulation had a daily variability different to the observed. The experiment NoDA had the~~ and showed the largest SST rmsd (>1.4 °C) near the model NW boundary, which was also seen in the assimilative runs (Fig. 3f,g,h). Assimilation of surface and subsurface fields (ASFUVTS run) reduced the maximum (∼ 1.0 °C) and average (0.45 °C) SST rmsd values in comparison to the NoDA run (Fig. 3f,h). Withholding subsurface temperature data (NoUVTS runs) had small impacts in the maximum (decrease of ∼ 0.2 °C) and mean (increase of 0.01 °C) SST rmsd (Fig. 3f,g), as well as for NoUV and NoTS runs (not shown).

The NoDA run had larger ~~spatially averaged SSH rmsd (mean = 8 cm)~~ spatial SSH rmsd in comparison to the ~~observed spatially averaged SSH std (mean = 7 cm)~~ ASFUVTS and NoUVTS runs for most of the year-long period of simulation (Fig. 4a). Assimilating subsurface velocity, temperature and salinity further reduced the average SSH rmsd by 14%. The NoDA run average spatial SSH correlation was 0.51 and below 0.2 in some events (Fig. 4b). This shows the lack of skill in simulating the timing and location of the mesoscale eddies in the NoDA run. ~~ASFUVTS and NoUVTS runs had lower temporally averaged spatial SSH rmsd (Fig. 4a) and higher temporally averaged spatial SSH correlation (Fig. 4b) in comparison to the NoDA run. ASFUVTS had even higher spatial SSH correlation which means that including subsurface data for assimilation improved the representation of mesoscale eddies and the EAuC studied in Santana et al. (2021)~~Assimilation of SSH and SST (NoUVTS) improved SSH correlation by 27% and the inclusion of subsurface data further increased the correlation (by 31%) which was above 0.4 virtually during the entire year of simulation. The SSH represents the integral of subsurface density fields which is a good proxy for representation of the whole ocean state and the SSH was better represented when velocity, temperature and salinity observations were assimilated.

The NoDA run had spatial SST rmsd similar ~~or smaller to the observed spatial SST std~~ to the assimilative runs in the first ~~two thirds~~ two-thirds of the simulation period, however, it showed larger errors from Feb. 2016 onwards and reached a maximum of 1.82°C (Fig. 4c). The ASFUVTS run had smaller SST rmsd time mean (0.47°C) in relation to the NoDA run (mean = 0.68°C). Withholding subsurface data (NoUVTS) had little impact ~~in~~ on the SST rmsd (mean = 0.46°C). The NoDA run had small

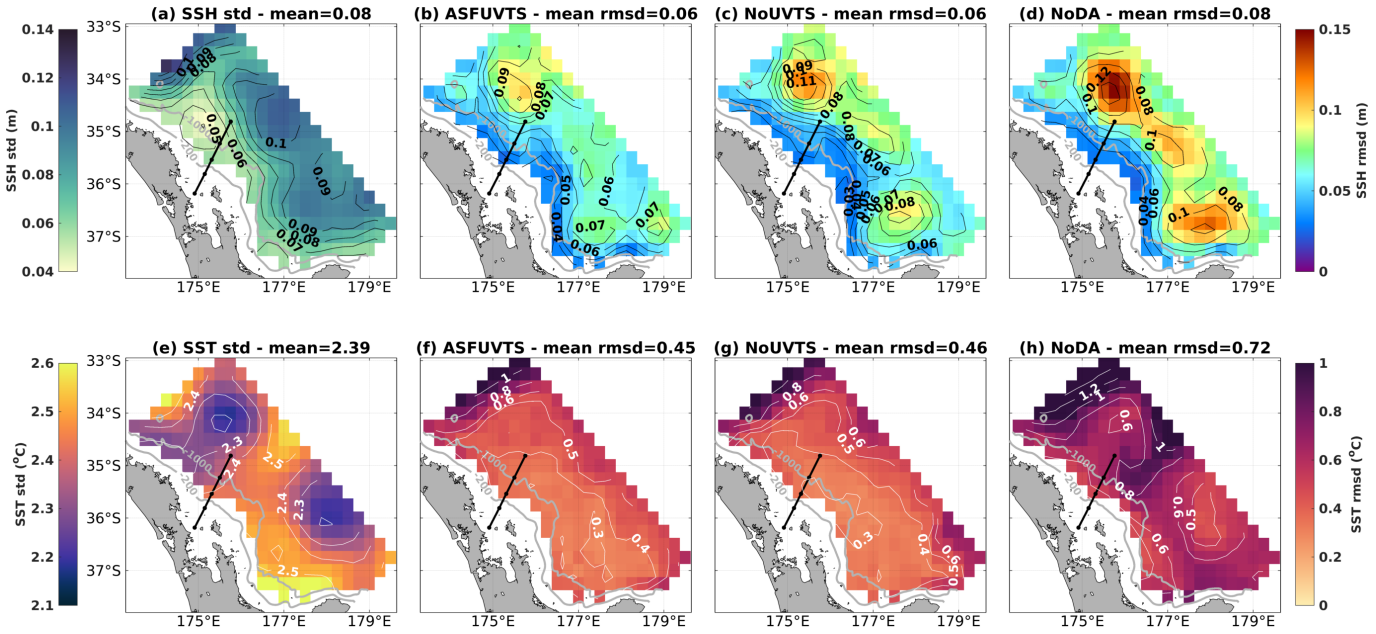

Figure 3. Maps of SSH (top row) and SST (bottom row) statistics. Observed AVISO SSH std (a) and AVHRR SST std (e), modelled ASFUVTS run SSH (b) and SST rmsd (f), NoUVTS SSH (c) and SST rmsd (g), NoDA SSH (d) and SST rmsd (h).

absolute SST difference to observations (<1.0°C) during the first ~~two thirds~~ two-thirds of the timeseries but a larger cold bias (<-1°C) was developed in Apr. 2016 (Fig. 4d). Assimilation of surface fields only (NoUVTS) had absolute SST bias below 0.5°C throughout the year-long period of simulation. ~~Inclusion~~ The inclusion of subsurface data (ASFUVTS) did not have a well-marked impact on model SST bias, most of the surface field correction was done by assimilation of SSH and SST more specifically.

Assimilation of subsurface ~~and surface~~ data (ASFUVTS) ~~had little impact on~~ positively impacted the representation of SST in comparison to the experiment that withheld subsurface observations (NoUVTS). ~~However, a~~ A larger positive impact was ~~seen on the spatial SSH rmsd with respect to AVISO~~shown on SSH rmsd when subsurface data was assimilated. Assimilation of in situ velocities, especially, further reduced the SSH rmsd up- and downstream of the moorings. Main differences between experiments appeared in the subsurface fields comparisons which are shown in the next section.

3.2 DA impact on subsurface fields

The NoDA run had very little velocity complex correlation (< 0.1) with observations at M4 and M5 (Fig. 5a,b). DA improved complex correlation between modelled and observed velocity vectors in all experiments at the two stations. These results can be related to assimilation of SSH which corrects the geostrophic circulation that is responsible for the long-term (> 30 days) upper half water column current variability in the region (Santana et al., 2021). The most complete assimilative run (ASFUVTS) had better results in comparison to ~~NoDA~~ the NoDA run but not against other assimilative runs. Withholding

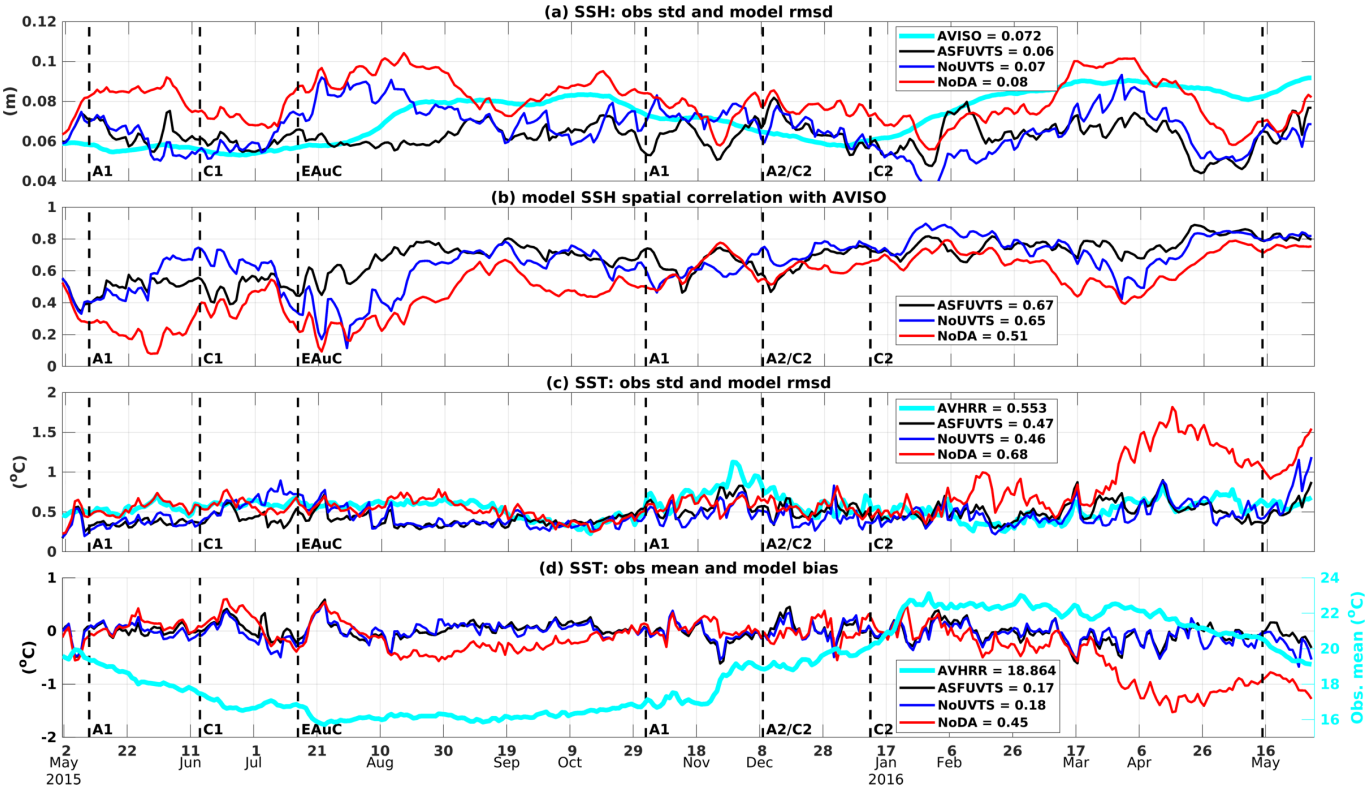

Figure 4. Timeseries of SSH spatial observed std ~~/~~and model rmsd (a), SSH spatial correlation (b), SST spatial observed std ~~/~~and model rmsd (c), and SST spatial mean ~~or~~ (right hand axis) and model bias (left hand axis) (d). The colours represent observations from AVISO/AVHRR (light blue), and model output from NoDA run (red), NoUVTS run (dark blue), and ASFUVTS run (black). Mean std, rmsd and correlation coefficients, and std SST bias are shown in the legend. Mean SST in (d) is shown on the right axis for better visualisation. The acronyms A1, C1, EAuC, A1, A2/C2, and C2 represent ~~the presence of~~ mesoscale structures ~~at~~ near M5 which were studied in Santana et al. (2021).

temperature and salinity (NoTS and NoUVTS) improved upon ASFUVTS, and NoTS had the best overall results. ASFUVTS-2days, on the other hand, had the best overall results. HYCOM-NCODA poorly represented velocity at M4 probably due to less accurate representation of the slope bathymetry ~~(de Souza et al., 2021).~~ (Fig. 11 in de Souza et al. (2021)). At M5, however, HYCOM-NCODA showed velocity complex correlation coefficient similar to the 7-day assimilative runs. Velocity at M5 was more dominated by the mesoscale field and less controlled by topographic constraints when compared to M4 (Santana et al., 2021), which makes velocity easier to simulate even in lower horizontal resolution models.

The assimilative experiments had smaller temperature rmsd at the surface in comparison to the NoDA run at M4 and M5 (coloured markers at the surface in Fig. 5c,d). ~~However, the lack of subsurface temperature for assimilation (NoTS and NoUVTS) led to larger errors (>1°C)~~ Assimilating subsurface temperature (ASFUVTS and NoUV) generated temperature rmsd smaller than the NoDA run rmsd in the upper ~~200~~ 500 m at M4 and M5 ~~, and below 500 m at M5 (pink and dark~~ (black and light blue lines in Fig. 5c,d). ~~Assimilation~~ The lack of subsurface temperature for assimilation (NoTS and NoUVTS),

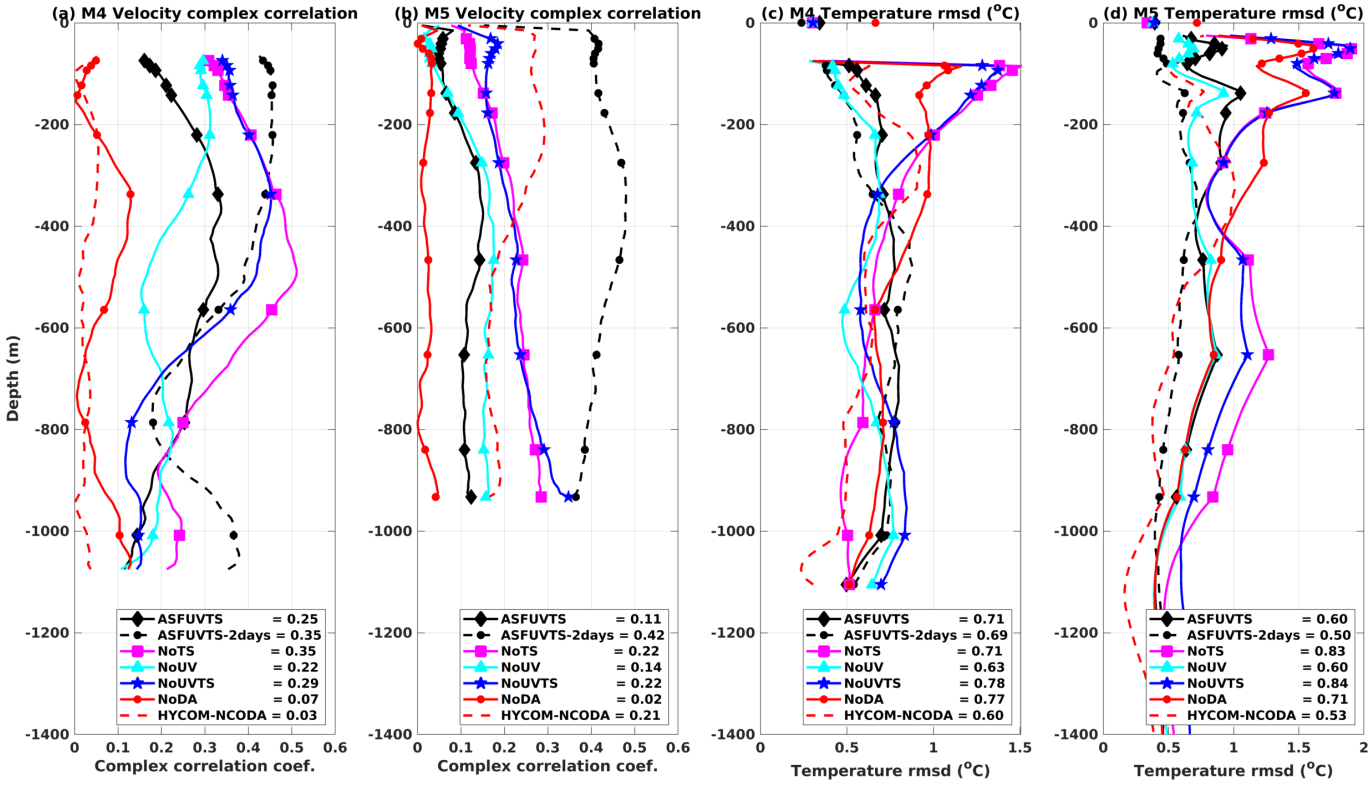

Figure 5. Profiles of complex correlation coefficient between observed and modelled velocity vector at (a) M4 and (b) M5. Profiles of temperature root mean square deviation between observed and modelled temperature at (c) M4 and (d) M5. Black diamonds (ASFUVTS), black dots (ASFUVTS-2days), magenta squares (NoTS), cyan triangles (NoUV), blue stars (NoUVTS) and red dots (NoDA) represent the median depth of thermistors and CTDs at each station, except at the surface where comparisons are against AVHRR SST. HYCOM-NCODA results are shown as dashed red lines. The depth averaged values of complex correlation and temperature rmsd are shown in the ~~legend~~legends.

however, led to larger rmsd (~~ASFUVTS and NoUV)had reduced temperature rmsd in comparison to the NoDA run from surface down to 500~~ >1°C) in the upper 200 m at M4 and M5~~(black and light~~, and below 500 m at M5 (pink and dark blue lines in Fig. 5c,d). The ASFUVTS-2days run had improved temperature rmsd down to 1000 m at M5. HYCOM-NCODA showed improved temperature results compared to the NoDA run (dashed and solid red lines in Fig. 5c,d). ~~This is due to the assimilation of~~ HYCOM-NCODA assimilates temperature and salinity data from Argo floats ~~since the reanalysis did not assimilate data from M4 or M5~~can explain the improved results.

Temperature differences between model and observations at M4 showed small cold bias and relative warming at the beginning of the timeseries for all experiments (Fig. 6). The small ~~bias~~ biases can be attributed to the initial ~~condition on 1st of May 2015~~ and boundary conditions obtained from HYCOM-NCODA analysis ~~, however, differences~~ which shows a relative warming between Jul and Aug 2015 when Argo data was not available in the region (red shade and numbers in Fig. 6e). Larger differences between the OSEs started to appear in Sep. 2015. The ASFUVTS and NoUV runs are colder than observed between

Sep. and Nov. 2015 and Mar. and May 2016 (<~~1~~ -1°C ~~- –~~ second lightest blue shade in Fig. 6~~b,c). NoTS and NoUVTS~~ a,b). The NoUVTS and NoTS (not shown) runs are cooler in the upper 200 m water column from Sep. 2015 onwards (Fig. 6~~d,e~~c). Withholding subsurface temperature for assimilation (~~NoTS and NoUVTS~~ NoUVTS and NoTS) led to the continuous cooling of waters in the top 200 m until the end of the simulation. The NoDA run had a small temperature difference when compared to observations in Sep. 2015 but these increased in Oct. 2015 (Fig. 6~~f~~d). The longest persistent cold bias in the NoDA run started in mid-Mar 2016, ~~which might be~~ associated with the SST cold bias during the same period (Fig. 4d). Cooling of the upper water column seen in the NoTS and NoUVTS runs also occurred in the NoDA run. This effect might be intrinsic to the model configuration, and data assimilation of velocities and/or surface fields only enhanced this bias. ASFUVTS-2days had a marked negative temperature difference between Aug and Sep 2015 that was reduced in the rest of the year (Fig. 6e).

At M5, a mid water column warm bias (>1°C) was simulated in the first half of the year-long period in all ~~experiments~~ OSEs (second lightest red shade Fig. 7a,b,c,d). This also ~~happened on~~ occurred in HYCOM-NCODA between Jun and Sep 2015 (~~not shown~~Fig. 7e) probably due to the lack of Argo floats to the north of 35°S. ASFUVTS had the ~~closest match~~ smallest differences to the observations out of all 7-day assimilation window experiments (Fig. 7~~b~~a). In mid-Aug 2015, a near surface cold bias (<-2°C) was pronounced in all experiments, but further developed in the NoDA, NoTS and NoUVTS runs towards ~~at~~ the end of the simulation period. This is associated with the lack of subsurface temperature assimilation and the ~~intrinsic variability in the model run.~~ near surface cold bias present in the NoDA run. The ASFUVTS (NoUV) run reduced the cold ~~and warm~~(warm) differences to values ~~below 1~~above -1°C (below 2°C) for most of the period of the simulation. It showed that the assimilation of subsurface temperature and salinity had positive impact by reducing temperature biases at M4 and M5. ~~We did not analyse salinity results at~~ ASFUVTS-2days better matched the observations and had the smallest temperature differences at M5 (Fig. 7f). However, this assimilation configuration led to larger temperature differences when mooring temperature was not assimilated. As such, this configuration is not advisable in this region for an operational forecast system as Argo floats can spend months without sampling the region.

Observed salinity timeseries from CTDs located near 200 m depth at M4 and M5 were compared to simulated salinity from three experiments and HYCOM-NCODA (Fig. 8). All models' salinity tended to oscillate near the observed values, and average absolute biases were smaller than 0.1 g/kg. The ASFUVTS run slightly degraded salinity results in comparison to the NoDA run, however, salinity rmsd were smaller than the observation standard deviation (0.16 g/kg). ASFUVTS-2days and HYCOM-NCODA had larger peaks and troughs in salinity timeseries and the largest variances (0.02 g/kg) at M5. This result can be attributed to the frequency of increments applied to the initial conditions – everyday in HYCOM-NCODA and every other day in ASFUVTS-2days. Despite that, ASFUVTS-2days had the best statistical performance amongst the simulations compared which can also be attributed to the more frequent increments to the initial conditions.

Salinity rmsd comparisons at three different depths at M4 and M5 are shown for all ROMS' experiments and HYCOM-NCODA in Table 3. These experiments had salinity rmsd oscillating around 0.16 ± 0.04 g/kg in shallower depths ($\leq$ 220 m) at M4 and M5 stations. The exception to this is ASFUVTS-2days which had the smallest salinity rmsd (<0.10 g/kg) in the upper M4 and M5 depths. HYCOM-NCODA also had salinity rmsd around 0.16 ± 0.04 g/kg in the upper and 200 m depth at both M4 and M5. Salinity rmsd was reduced near the bottom at M4 and M5 (< 0.08 g/kg) in all ROMS' experiments as well as for

HYCOM-NCODA. ASFUVTS-2days, however, had the largest salinity rmsd near the bottom at M5. The bathymetry used in HYCOM-NCODA was shallower than near bottom measurements at M5 and salinity rmsd was not calculated. These OSEs' results suggest that future experiments should reduce salinity observation standard deviation (or error) below a certain depth. For instance, observed salinity error can be reduced to 0.06 g/kg below 1000 m as in Kerry et al. (2016) or an exponential decay can be applied from surface (0.12 g/kg) to 2000 m (0.02 g/kg) (Xie and Zhu, 2010; Mignac et al., 2015; Santana et al., 2020). In the ~~moorings due to the few number (≤3)of CTDs at each station. Instead, model salinity is compared to Argo independent observations in the~~ next section, model temperature and salinity are validated against Argo-independent observations.

Table 3. Salinity root mean square deviation (rmsd) between experiments and CTD observations at M4 and M5.

Experiments	M4 84 m	M4 220 m	M4 1008 m	M5 31 m	M5 177 m	M5 1799 m
ASFUVTS	0.198	0.110	0.068	0.169	0.152	0.047
ASFUVTS-2days	0.079	0.098	0.045	0.064	0.115	0.074
NoTS	0.181	0.103	0.074	0.159	0.133	0.041
NoUV	0.159	0.093	0.061	0.147	0.114	0.045
NoUVTS	0.164	0.085	0.072	0.159	0.111	0.038
NoDA	0.155	0.092	0.066	0.146	0.134	0.045
HYCOM-NCODA	0.172	0.160	0.041	0.132	0.163	shallower bathymetry

3.3 Comparison to independent observations

Thirty non-assimilated temperature and salinity profiles spread in time (Fig. 9) and throughout the domain (Fig. 10e) ~~and in time~~ were used for model-data independent comparison. All experiments had small salinity differences when compared to the first five salinity profiles sampled by Argo floats which can be attributed to HYCOM-NCODA's initial and boundary conditions (Fig. 9). Larger salinity differences, however, appeared in the OSEs' results in late August when simulated salinity was fresher in the upper 200 m depth and saltier underneath. Assimilation of surface and mooring data using a 7-day window did not generate any marked positive/negative impact in salinity in comparison to the NoDA run. ASFUVTS-2days had the smallest differences between simulated and observed salinity (Fig. ~~7) were used for model-data independent comparison. Model~~ 9f) and its results were similar to HYCOM-NCODA (Fig. 9e) which assimilates Argo data. However, fresher salinity differences were simulated in mid May 2016 when mooring salinity data collection had ended.

In this analysis, we included experiments using a 7-day assimilation window and doubled decorrelation length scales for tracers (NoUVTS-2x and ASFUVTS-2x) to show the impact of this approach on simulating temperature and salinity at Argo locations. All experiments' temperature rmsd showed similar ~~trends~~ patterns to the results simulated at M4 and M5. When temperature was not assimilated (~~NoTS and NoUVTS runs~~NoUVTS and NoTS) there was a higher temperature rmsd between 0 and 200 m in comparison to the NoDA run due to a colder bias (<-1°C) that developed in ~~those runs (blue and pink lines in~~ these runs (Fig. 10a,b). Assimilation of subsurface temperature data (ASFUVTS, ASFUVTS-2days and NoUV runs) ~~generated~~

resulted in smaller temperature rmsd in comparison to the NoDA run from the near surface down to 2000 m (~~ASFUVTS) (1000 m in NoUV) (solid black and cyan lines in~~ Fig. 10a,b). HYCOM-NCODA reanalysis which assimilates Argo, SSH and SST data had the overall best temperature representation at the Argo locations. Assimilation of surface and mooring data using doubled decorrelation length scales (ASFUVTS-2x) slightly improved temperature representation near the surface and around 1000 m depth in comparison to ASFUVTS (Fig. 10a,b). However, when subsurface data was not assimilated (NoUVTS-2x) larger temperature rmsd were generated between 200 and 1200 m depth in relation to the NoUVTS run.

Comparisons to Argo salinity measurements showed that all 7-day window experiments and the NoDA run had similar vertical ~~structure~~ structures in the salinity bias. Fresher salinity was modelled in the upper 200 m, and saltier waters were modelled below that, peaking at 600 m, in ~~those~~ these experiments (Fig. 10c). ~~The NoDA run had the smallest salinity rmsd from 0 to 200 m, and ASFUVTS had the largest at that depth (Fig. 10d). It seemed that the 8 CTD observations spread between M3-M5 were not enough to correct salinity in the model domain. On the other hand,~~ ASFUVTS-2days had the smallest salinity bias and rmsd in comparison to the 7-day window experiments ~~. It~~ (Fig. 10d). This suggests that, more frequent increments to the initial conditions and doubled decorrelation length scales of tracers can overcome the small number of ~~CTD~~ salinity observations. ASFUVTS-2days salinity results were close to HYCOM-NCODA reanalysis which assimilated salinity data from Argo floats.

Applying doubled decorrelation length scales of tracers to the 7-day assimilation window experiments slightly reduced rsmd for temperature (3%) and salinity (4%) at the Argo locations' when assimilating mooring temperature and salinity (ASFUVTS-2x) in comparison to using 100 km (100 m) as horizontal (vertical) length scale (Fig. 10). However, when subsurface data was withheld (NoUVTS-2x) it increased the mean temperature (salinity) rmsd by 18% (47%). Larger rmsd occurred when withholding mooring temperature and salinity data using a 2-day assimilation window and double decorrelation length scales of tracers (NoUVTS-2days – not shown).

3.4 Increments to initial and surface conditions

Variability in temperature increments to the initial conditions at M4 revealed oscillation between positive and negative increments through time in all experiments (Fig. 11). The 7-day window experiments that assimilated subsurface temperature and salinity (ASFUVTS and NoUV) had positive increments extending from surface down to 500 m (darker red shade in Fig. 11a,b). In contrast, experiments that did not assimilate temperature and salinity (NoTS and NoUVTS) had ~~large~~ strong positive increments bounded to near the surface (Fig. 11c,d). ASFUVTS-2days ~~had smaller increments throughout the water column~~ did not have large positive increments near the surface but increments tended to vary from positive to negative values between assimilation cycles (Fig. 11e).

Near surface average temperature increments had a distinct difference between experiments that did and did not assimilate subsurface temperature at M4 and M5 (Fig. 12a,b). Assimilation of in situ temperature (ASFUVTS and NoUV) generated larger average positive increments (>0.04°C) around 100 m depth in comparison to the simulations that withheld subsurface temperature (NoTS and NoUVTS). The latter experiments had near surface mean negative temperature increments that

decreased towards zero below 200 m. ASFUVTS-2days had smaller average ~~increments throughout the water column~~ and standard deviation increments near the surface in comparison to the 7-day experiments (Fig. 12a,b).

Surface mean heat flux ~~increments~~ analysis fields varied according to the presence/absence of subsurface temperature for assimilation (Fig. 13). Simulations that withheld subsurface temperature (~~NoTS and NoUVTS) had average positive increments~~ NoUVTS and NoTS) had large positive heat flux near the moorings and in the majority of the domain (Fig. 13a,b). The positive heat flux ~~increments~~ in the NoTS and NoUVTS runs ~~were~~ was related to the ~~near surface~~ near-surface negative temperature increments at M4 and M5 (dark blue and pink lines Fig. 12). The combined positive heat flux ~~increments~~ and negative temperature increments corrected the surface cold bias present in the NoDA run, but not the cooling trend around 200 m. Experiments that assimilated subsurface temperature (ASFUVTS, NoUV and ASFUVTS-2days) had negative average heat flux ~~increment near~~ (or near zero heat flux – ASFUVTS-2days) near the moorings that extended north and south with varying size between experiments (Fig. 13c,d,e). This negative heat flux increment ~~might~~ balance the positive temperature increment at M4 and M5 which aim at correcting both surface and upper thermocline (∼200 m) ~~clod biases.~~ cold biases. The large variability in heat flux analysis fields between experiments (Fig. 13) might be related to the 4-day variances used to compute the background error covariance. The 4-day variances used here include daily cycle and cloud coverage variabilities which generated large heat flux uncertainty and increments.

The ASFUVTS-2days run had smaller negative heat flux ~~increments~~analysis near the moorings. This happened due to the smaller average negative increments (M4) and positive increments (M5) added to the initial condition of temperature at ~~those~~ these stations (black dashed line Fig. 12). The NoDA run average net heat flux ~~showed roughly positive values north~~ was mostly positive north of 35°S and ~~it~~ was negative southwards (Fig. 13f). On average, the atmosphere has a warming effect on the ocean north of 35°S and a cooling effect south of 35°S. This is observed in JRA55-do (Fig. 48 in Tsujino et al. (2018)), in which the NZNES is located near the average zero net heat flux contour line in the atmospheric forcing dataset.

Annual average wind stress is mainly from ~~southwest in all experiments but wind~~ northwest in the 7-day window runs and from west in the ASFUVTS-2days and NoDA runs (vectors in Fig. 14). Wind stress curl showed some spatial variability between the assimilative runs (red and green shade in Fig. 14). Experiments that did not assimilate subsurface temperature and salinity (NoUVTS and NoUV) had average negative wind stress curl at the moorings' location (Fig. 14a,b). Conversely, assimilation of subsurface temperature and salinity (NoUV and ASFUVTS) generated average positive wind stress curl on top of M5 (M3-M5 in ASFUVTS) (Fig. 14c,d). Positive wind stress curl generates convergence and downwelling of warmer water masses that might be acting to prevent the cold bias in the numerical model. These changes in wind stress curl ~~might~~ can also explain why assimilation of subsurface temperature and salinity (ASFUVTS and NoUV) slightly degraded the representation of velocity at M4 and M5 compared to withholding subsurface temperature and salinity (NoTS an NoUVTS). ASFUVTS-2days run had smaller wind stress curl magnitude compared to the other simulations. Its wind stress curl field resembles the NoDA wind field forced with JRA55-do, especially in the regions of negative wind stress curl near the coast around 37.5°S and observed by Taboada et al. (2019) using CCMP data.

4 Discussion and comparison to other studies

In this study, SSH, SST, mooring velocity, temperature and salinity observations were assimilated into an ocean model of the EAuC using 4D-Var with a 7-day ~~window length~~assimilation window. Observing system experiments (OSEs) were conducted in order to elucidate the importance of in situ data assimilation in the EAuC region. ~~4~~ Four OSEs were performed based on the most complete simulation called ASFUVTS which assimilated surface fields (SSH and SST), and mooring velocity, temperature and salinity. The other experiments withheld observation types from assimilation. They removed mooring velocities (NoUV); subsurface temperature and salinity (NoTS); and all mooring data (NoUVTS). A non-assimilative ~~freely evolving~~ simulation (NoDA) was used as the control run. Argo data was left out ~~of all experiments~~ for independent model-data comparison. Another run that assimilated surface and mooring data using a ~~4D-Var~~ 2-day window (ASFUVTS-2days) and aimed at better matching the observations rather than serving as basis for an operational forecast was also evaluated. HYCOM-NCODA outputs, which provided boundary conditions, were also used in the model comparison.

All assimilative experiments showed a reduction in SSH rmsd in comparison to the NoDA run of about 25%. The improvements in SSH were small in comparison to achievements seen in other WBC 4D-Var studies, such as: the East Australian Current (EAC) (∼63% Kerry et al. (2016)) and the Brazil Current (BC) (48% - de Paula et al. (2021)). Smaller correction in SST (rmsd reduction of 37%) when compared to 4D-Var studies in the East Australian Current (EAC) (∼60% Kerry et al. (2016)) and the Brazil Current (BC) (27% - de Paula et al. (2021)). These studies used lower horizontal resolution grids in the open ocean (5 km EAC and 9 km BC) compared to our model spatial grid spacing (2 km) which might explain the differences in performance. According to Sandery and Sakov (2017), increasing model resolution towards the submesocale (from 10 km to 2.5 km) reduces the skill of the analysis and forecast~~generated~~. They suggested that resolving the less predictable submesoscale lowers the predictability of the mesoscale as there is an inverse cascade in the kinetic energy spectrum. Kerry et al. (2020) found larger errors when downscaling from a regional to a coastal domain (750-1000 m resolution) while simulating the cyclonic inshore side and frontal instabilities of the EAC.

All OSEs improved velocity complex correlation by at least ~~3-fold~~ three times at M4 and ~~5-fold~~ five times at M5 in comparison to the NoDA run (coef. < 0.07). ~~Including velocities in the assimilation~~ Assimilation of subsurface velocity (NoTS) simulated flow reversals at depth on the slope (M4) more accurately and it improved on the simulation ~~which only assimilated~~ that assimilated only surface fields (NoUVTS). ~~This was most notable at M4, where the velocity data assists the model in better capturing frequent reverse flows at depth.~~ Counterintuitively, inclusion of temperature and salinity in the 7-day window runs (NoUV and ASFUVTS) degraded the velocity results in comparison to NoUVTS run. This might be associated with ~~the~~ increased positive wind stress curl near the moorings in experiments that assimilated subsurface temperature (ASFUVTS and NoUV runs). Positive wind stress curl causes downwelling of warmer waters and counter-balances the cold bias in the model but degrades velocity results in ASFUVTS and NoUV runs. The wind impact can be reduced by decreasing the number of days to estimate wind variance (uncertainty). Another factor that may have played a role was the low vertical resolution of salinity sensors, which were not enough to correct density fields and generate accurate geostrophic currents. For future data collection strategies, we suggest that higher vertical resolution for salinity data. ~~Changes in the wind stress curl analysis field had a~~

negative impact in the representation of the velocity at M4 and M5 and may be the main cause for that. Larger positive wind stress curl was generated near the moorings in ASFUVTS and NoUV runs. Positive wind stress curl causes downwelling of warmer waters and counter-balances the cold bias in the model but degraded velocity results in ASFUVTS and NoUV runs. The application of doubled decorrelation length scales of tracers improved temperature, salinity and velocity representation when subsurface data was assimilated (ASFUVTS-2x and ASFUVTS-2days). However, this strategy led to larger temperature and salinity errors when subsurface tracer data was not assimilated (NoUVTS-2x and NoUVTS-2days).

The 2-day assimilation window run (ASFUVTS-2days) showed had the highest complex correlation coefficients (>0.35) which combined from a combination of doubled decorrelation length scales, larger model error estimate estimates and more frequent increments. ASFUVTS-2days overcame the low vertical resolution of salinity and temperature observations was more efficient at extracting and spreading mooring observational information and generated the best results when compared to Argo distant locations. temperature and salinity from Argo observations. This ocean reanalysis can be used to estimate EAuC heat and volume transports since it better represents observations. However, this approach must be used with care should not be used if subsurface data availability is low because larger colder biases a large subsurface cold bias can be generated. This is the case for an operational forecast in which Argo floats might not be available for months. In Kerry et al. (2016), higher velocity complex correlation coefficients ($\sim$1) were obtained in a 2-year reanalysis of the EAC. However, the EAC has a more coherent jet (Mata et al., 2000; Bowen et al., 2005; Sloyan et al., 2016) which is well represented by the non-assimilative run (complex correlation $\sim$0.8 in some locations) (Kerry et al., 2016). In contrast, the EAuC has a more eddy-dominated field (2/3 of 1 year) (Santana et al., 2021) which makes it harder for ocean free-running models and reanalyses to capture such variabilityvariabilities.

Marked differences between the experiments arose when model results were compared to subsurface observations of among the experiments appeared in the subsurface temperature. Experiments that withheld in situ temperature and salinity (NoTS and NoUVTS) generated a larger cold bias around 100 m at M4 and M5 in comparison to the NoDA run. At 100 m, errors rmsd were about 1.4°C (1.9°C) in the NoTS (NoUVTS) run at M4 and M5, whereas the NoDA run had temperature rmsd of 1.1°C (1.6°C) at M4 (M5). Assimilation of in situ temperature and salinity (ASFUVTS and NoUV) prevented that and had improved results against the NoDA runreduced the cold bias. The lack of subsurface temperature assimilation also generated similar errors at the top of the thermocline in other regional studies (e.g., Zavala-Garay et al. (2012); Santana et al. (2020)). Zavala-Garay et al. (2012) needed to assimilate XBT or synthetic CTD data to correct that large temperature rmsd ($\sim$2°C) between 200 and 500 m depth. The authors aimed at simulating the EAC variability between years 2001 and 2002, and the Argo project was still beginning, with few sondes in the ocean. In our study, comparisons to HYCOM-NCODA suggest that it is still an open question if assimilating the few Argo temperature and salinity profiles (30) and surface data could would prevent the growth of the 100 m cold bias at M4 and M5. However, it still an open question if this would improve the ROMS 4D-Var results used in this study. These experiments represent a good benchmark to define proper assimilation window assess proper assimilation windows and decorrelation length scales. Doubled length scales of tracers (200 km and 200 m) and a 2-day assimilation window led to colder biases in the experiments that withheld in situ temperature (not shown) which could be a problem when Argo data is not available for a long period. The 7-day assimilation window and smaller length scales (100 km

and 100 m) ~~seem~~ seemed to be a good configuration to well-represent the surface (SSH and SST) and subsurface velocity fields in ~~case of application in~~ an operational forecast system.

In ROMS 4D-Var, spatial decorrelation length scales are set using one value per model variable for the whole domain (Moore et al., 2011a). In reality, decorrelation length scales can vary in space depending on the main drivers of variability in each specific region. For instance, atmospheric heat flux dominates temperature variability at the surface, and SST at M4 has high correlation (>0.7) with points spread hundreds of kilometres across in the NoDA run (Fig. 15a). Nonetheless, this correlation is limited to a few metres below the surface (<20 m) (Fig. 15b). At 200 m depth, the high correlation area covers a region spanning ~50 km across and ~350 m at depth (Fig. 15c,d). In this study, we showed that using 100 km (m) as horizontal (vertical) decorrelation length scales of tracers is a good approach to improve upon the non-assimilative (NoDA) run while preventing the degradation of subsurface temperature fields when water column data is not assimilated. However, more tests can be done while assimilating Argo data and a non-trivial implementation of spatially variable decorrelation length scales is encouraged.

Model temperature comparisons to Argo data showed similar results to ~~those simulated~~ comparisons at M4 and M5. NoTS and NoUVTS runs had a larger cold bias near 100 m depth which generated a higher temperature rmsd (~ 2°C) in comparison to the NoDA run (~ 1.6°C). A warm bias was also evident around 600 m, in all experiments with varying degrees. Assimilation of mooring temperature (ASFUVTS and NoUV) reduced the cold bias and the temperature rmsd (~ 1°C) in comparison to the NoDA run. The shallow-cold and deeper-warm biases are ~~intrisic~~ intrinsic to the NoDA run, which might be associated with biases in the boundary condition from HYCOM-NCODA when Argo data was not available (between Jun and Sep 2015) for assimilation north of 35°S. Even though larger temperature errors were observed in the simulations that withheld mooring temperature and salinity (1.6 - 2.1°C), they were still comparable to freely evolving simulations temperature rmsd (~ 1.9°C) in other studies (e.g., Kerry et al. (2016); Siripatana et al. (2020)) when assessed using Argo temperature.

Data assimilation of subsurface observations using a 7-day window and smaller decorrelation length scales had little impact in correcting salinity at the Argo locations. A small salinity degradation was simulated between 0 and 200 m in the assimilative experiments in comparison to the NoDA run. Negative impact on near surface salinity was also observed when assimilating SST or SSH data only in the Brazil Current region (Santana et al., 2020). Some authors suggest that assimilation of salinity data is needed to constrain the model water column salinity (e.g., Oke and Schiller (2007); Tanajura et al. (2014); Oke et al. (2015)). Assimilation of surface and subsurface data every 2 days with larger decorrelation length scale of tracers (ASFUVTS-2days) was able to considerably correct model salinity below 400 m~~in relation to the NoDA run.~~ . The ASFUVTS-2days run reached error values similar to HYCOM-NCODA which assimilated salinity data from Argo floats.

Increments to temperature initial conditions were positive and deep (200 m) in the experiments that assimilated subsurface temperature (ASUVTS and NoUV). In contrast, withholding in situ temperature (NoTS and NoUVTS) generated negative and shallower (50 m) increments. This generated differences in the atmospheric heat fluxes, where NoTS and NoUVTS had average positive heat flux ~~increment~~ in most of the domain to compensate the negative increments to the temperature initial condition. The combined positive heat flux increments and negative temperature increments corrected the surface cold bias present in the NoDA run, but not the cooling trend around 200 m. Conversely, experiments that assimilated subsurface temperature (ASFU-

VTS and NoUV) had negative average heat flux ~~increment~~ near the moorings. This negative heat flux ~~increment~~ balanced the positive temperature increment at M4 and M5 which aimed at correcting both surface and upper thermocline (∼200 m) cold biases.

ASFUVTS, NoUV and ASFUVTS-2days runs (assimilated subsurface temperature) had net ~~heat flux increment~~ negative or near zero surface heat flux varying from negative ~~(near the moorings and north of the domain)~~ to positive flux ~~in the southeast region of the model domain~~from near to away from the moorings. This variability might be associated with the ~~heat flux~~ 4-day variances which include the daily cycle of radiation and cloud coverage and were used to compute the background error covariance of the forcing heat flux. Alternatively, there is a large spatial variability in heat flux that forced the NoDA run. It was positive north of 35°S and negative southwards. The NZNES is located near the zero heat flux line in the JRA55-do atmospheric forcing used in this study (Fig. 48 in Tsujino et al. (2018)). For instance, the Hawaiian archipelago is also located near a zero line heat flux in the atmospheric forcing product. 4D-Var experiments in the region, also showed large average heat flux increments ($\pm$100 W/m^2), and marked spatial variability with negative increments in the lee of the islands and positive increments to the west of ~~that~~ them (Matthews et al., 2012).

Positive wind stress curl was generated on top of M5 when subsurface temperature data was assimilated (ASFUVTS and NoUV). Positive wind stress curl generates convergence and downwelling of warmer waters which might be associated with corrections of the 200 m cold bias. Positive wind stress curl correction was also observed in the simulations on SE Brazil which reduced the magnitude of upwelling in the region (de Paula et al., 2021). Changes in the wind stress curl in the ASFUVTS and NoUV runs (assimilated subsurface temperature) were also responsible ~~to~~ for the lowered velocity complex correlation at M4 and M5 in comparison to the simulations that withheld subsurface temperature. If more subsurface salinity data were available, the solutions ASFUVTS and NoUV runs could have converged to improved density structures that resulted in ~~better simulated~~ better-simulated currents as we observed in ASFUVTS-2days.

5 Conclusions and future work

By running the OSEs we elucidate the importance of different datasets on the quality of ocean reanalyses. The representation of surface fields and consequent mesoscale eddies was improved by data assimilation of surface data only. The model high spatial resolution (2 km) which starts to ~~solve submesoscale variability~~ resolve submesoscale process might be responsible for the lower skill compared to other 4D-Var regional studies that had lower horizontal resolution ($>$ 5 km). In situ subsurface temperature is of utmost importance to correctly simulate the top of the thermocline - one of the most difficult regions to simulate in ocean models. The lack of subsurface temperature for assimilation (NoTS and NoUVTS) increased the near surface cold bias present in the freely evolving model run (NoDA). Assimilation of mooring temperature (ASFUVTS and NoUV) corrects this cold bias, even at distant Argo locations. Data assimilation using a 2-day window and doubled decorrelation length scales better matched the assimilated and independent observations. This approach must be used with care if subsurface data availability is low (e.g., using Argo floats only as subsurface data) because larger cold bias in the upper-thermocline can be generated (not shown). Nevertheless, all reanalyses showed improved velocity results on the mid-slope when compared to

HYCOM-NCODA, which shows the importance of downscaling to better represent the slope bathymetry and possibly shelf-slope exchange.

The current work is part of a set of experiments that prepare for a data assimilative ocean forecast for the NZNES. The computational cost to produce one day of reanalysis is about 52 min, and 7 min to generate a 7-day forecast using 80 cores on the NeSI supercomputer (https://www.nesi.org.nz/), a Cray XC50. Alongside the operational forecast development, improvements in the NoDA run configuration need to be conducted in order to understand and mitigate temperature and salinity biases. Tests should be conducted using the Mercator-Ocean forecast system (GLORYS) (Lellouche et al., 2018) as boundary conditions since they showed better performance in comparison to HYCOM-NCODA (de Souza et al., 2021). New experiments based on the 7-day window configuration can be ~~conducted~~ done using Argo and glider data. Questions regarding the absence or presence of glider data can be asked and the impact on the velocity field evaluated with observations. Continuous temperature and salinity sampling from ocean gliders would provide enough high vertical resolution and good spatial coverage that would positively impact the simulation of the thermohaline field and ocean currents.

In the future, a posterior check of the consistency of the observation and background error hypotheses (Mattern et al., 2018) can be applied to improve the quality of the ocean reanalyses. Mattern et al. (2018) described how the covariance of residuals and innovations, and the covariance of increments and innovations, should be roughly equal to the assumed observation and background error variances, respectively. Posterior tests of these statistics probe whether the prior assumptions for error variances need adjusting to be consistent with the model intrinsic skill and representativeness error. This method has been applied to the ~~Mercator-Ocean forecast system (GLORYS)~~ GLORYS to obtain better performance (Lellouche et al., 2018). Moreover, an Ensemble 4D-Var approach can be applied to improve the quality of the analysis on the NZNES. This methodology uses several perturbed simulations (Ensemble) to estimate the model error covariance matrix (**D**) (e.g., Pasmans and Kurapov (2019)). This includes spatial and temporal variability to **D** compared to a fixed model covariance matrix used here. This method showed improved representation of glider temperature and salinity observations when compared to 4D-Var using a static **D** (Pasmans et al., 2020).

Code availability. The Regional Ocean Model System (ROMS) is widely known in the modelling oceanographic community and its source code, documentation, and discussion forum can be found at https://www.myroms.org/. Configuration files and scripts for analysing model output can be found on https://doi.org/10.5281/zenodo.7306271

Author contributions. J.O. conducted the sampling experiment. H.M. set up the numerical model and lateral conditions. R.S. analysed the data, included new atmospheric forcing in the model, generated the data assimilative runs, analysed the results and wrote the manuscript. B.P. supervised the improvement of the reanalysis and provided seapy (python library for ocean state estimate). R.S., H.M., J.O., B.P., S.W., and S.S. discussed the results and reviewed the manuscript.

Competing interests. No competing interests are present

Acknowledgements. We thank everyone involved in data collection and curation at NIWA, as well as support from the New Zealand eScience Infrastructure (NeSI) that made this work possible. Funding was provided by NIWA Strategic Science Investment Fund (SSIF) Ocean-Climate Interactions Programme, and Flows and Productivity Programme and MBIE endeavour Smart idea grant CO1X1814. We also thank John Wilkin who provided extra comments that improved the manuscript.

References

Bannister, R.: A review of operational methods of variational and ensemble-variational data assimilation, Quarterly Journal of the Royal Meteorological Society, 143, 607–633, 2017.

Bowen, M. M., Wilkin, J. L., and Emery, W. J.: Variability and forcing of the East Australian Current, Journal of Geophysical Research: Oceans, 110, https://doi.org/10.1029/2004JC002533, 2005.

Casey, K. S., Brandon, T. B., Cornillon, P., and Evans, R.: The Past, Present, and Future of the AVHRR Pathfinder SST Program, pp. 273–287, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-90-481-8681-5_16, 2010.

Chapman, D. C.: Numerical treatment of cross-shelf open boundaries in a barotropic coastal ocean model, Journal of Physical oceanography, 15, 1060–1075, 1985.

Chassignet, E. P., Hurlburt, H. E., Metzger, E. J., Smedstad, O. M., Cummings, J. A., Halliwell, G. R., Bleck, R., Baraille, R., Wallcraft, A. J., Lozano, C., et al.: US GODAE: global ocean prediction with the HYbrid Coordinate Ocean Model (HYCOM), Oceanography, 22, 64–75, 2009.

Chiswell, S. M.: Determining the internal structure of the ocean off northeast New Zealand from surface measurements, New Zealand Journal of Marine and Freshwater Research, 35, 289–306, https://doi.org/10.1080/00288330.2001.9516999, publisher: Taylor & Francis, 2001.

Chiswell, S. M., Bostock, H. C., Sutton, P. J., and Williams, M. J.: Physical oceanography of the deep seas around New Zealand: a review, New Zealand Journal of Marine and Freshwater Research, 49, 286–317, https://doi.org/10.1080/00288330.2014.992918, 2015.

de Paula, T. P., Lima, J. A. M., Tanajura, C. A. S., Andrioni, M., Martins, R. P., and Arruda, W. Z.: The impact of ocean data assimilation on the simulation of mesoscale eddies at Sao Paulo plateau (Brazil) using the regional ocean modeling system, Ocean Modelling, 167, 101 889, https://doi.org/https://doi.org/10.1016/j.ocemod.2021.101889, 2021.

de Souza, J. M. A. C., Powell, B., Castillo-Trujillo, A. C., and Flament, P.: The vorticity balance of the ocean surface in Hawaii from a regional reanalysis, Journal of Physical Oceanography, 45, 424–440, 2015.

de Souza, J. M. A. C., Couto, P., Soutelino, R., and Roughan, M.: Evaluation of four global ocean reanalysis products for New Zealand waters–A guide for regional ocean modelling, New Zealand Journal of Marine and Freshwater Research, 55, 132–155, 2021.

Di Lorenzo, E., Moore, A. M., Arango, H. G., Cornuelle, B. D., Miller, A. J., Powell, B., Chua, B. S., and Bennett, A. F.: Weak and strong constraint data assimilation in the inverse Regional Ocean Modeling System (ROMS): Development and application for a baroclinic coastal upwelling system, Ocean Modelling, 16, 160–187, 2007.

Ducet, N., Le Traon, P. Y., and Reverdin, G.: Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2, Journal of Geophysical Research: Oceans, 105, 19 477–19 498, https://doi.org/https://doi.org/10.1029/2000JC900063, 2000.

Fairall, C. W., Bradley, E. F., Hare, J., Grachev, A. A., and Edson, J. B.: Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm, Journal of climate, 16, 571–591, 2003.

Feng, M., Wijffels, S., Godfrey, S., and Meyers, G.: Do eddies play a role in the momentum balance of the Leeuwin Current?, Journal of Physical Oceanography, 35, 964–975, 2005.

Fisher, M. and Courtier, P.: Estimating the covariance matrices of analysis and forecast error in variational data assimilation, ECMWF Tech. Mem. 220, 1995.

Haidvogel, D. B., Arango, H., Budgell, W. P., Cornuelle, B. D., Curchitser, E., Di Lorenzo, E., Fennel, K., Geyer, W. R., Hermann, A. J., Lanerolle, L., et al.: Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System, Journal of computational physics, 227, 3595–3624, 2008.

Janeković, I., Mihanović, H., Vilibić, I., Grčić, B., Ivatek-Šahdan, S., Tudor, M., and Djakovac, T.: Using multi-platform 4D-Var data assimilation to improve modeling of Adriatic Sea dynamics, Ocean Modelling, 146, 101 538, 2020.

Kerry, C., Powell, B., Roughan, M., and Oke, P.: Development and evaluation of a high-resolution reanalysis of the East Australian Current region using the Regional Ocean Modelling System (ROMS 3.4) and Incremental Strong-Constraint 4-Dimensional Variational (IS4D-Var) data assimilation, Geoscientific Model Development, 9, 3779–3801, 2016.

Kerry, C., Roughan, M., and Powell, B.: Predicting the submesoscale circulation inshore of the East Australian Current, Journal of Marine Systems, 204, 103 286, 2020.

Kundu, P. K.: Ekman Veering Observed near the Ocean Bottom, Journal of Physical Oceanography, 6, 238 – 242, https://doi.org/10.1175/1520-0485(1976)006<0238:EVONTO>2.0.CO;2, 1976.

Laing, A. K., Oien, N. A., Murphy, R., and Uddstrom, M. J.: Coherent signals in the temperature and height of the sea surface off North Cape, New Zealand, New Zealand Journal of Marine and Freshwater Research, 32, 187–202, https://doi.org/10.1080/00288330.1998.9516819, 1998.

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Regnier, C., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Drevillon, M., Garric, G., and Drillet, Y.: Mercator Ocean Global High-Resolution Monitoring and Forecasting System, New Frontiers in Operational Oceanography, pp. 563–592, 2018.

Lentz, S. J.: Observations and a Model of the Mean Circulation over the Middle Atlantic Bight Continental Shelf, Journal of Physical Oceanography, 38, 1203 – 1221, https://doi.org/10.1175/2007JPO3768.1, 2008.

Levin, J., Arango, H. G., Laughlin, B., Hunter, E., Wilkin, J., and Moore, A. M.: Observation impacts on the Mid-Atlantic Bight front and cross-shelf transport in 4D-Var ocean state estimates: Part I—Multiplatform analysis, Ocean Modelling, 156, 101 721, 2020.

López, A. G., Wilkin, J. L., and Levin, J. C.: Doppio–a ROMS (v3. 6)-based circulation model for the Mid-Atlantic Bight and Gulf of Maine: configuration and comparison to integrated coastal observing network observations, Geoscientific Model Development, 13, 3709–3729, 2020.

Marchesiello, P., McWilliams, J. C., and Shchepetkin, A.: Open boundary conditions for long-term integration of regional oceanic models, Ocean modelling, 3, 1–20, 2001.

Marchesiello, P., McWilliams, J. C., and Shchepetkin, A.: Equilibrium structure and dynamics of the California Current System, Journal of physical Oceanography, 33, 753–783, 2003.

Mason, E., Molemaker, J., Shchepetkin, A. F., Colas, F., McWilliams, J. C., and Sangrà, P.: Procedures for offline grid nesting in regional ocean models, Ocean Modelling, 35, 1–15, https://doi.org/https://doi.org/10.1016/j.ocemod.2010.05.007, 2010.

Mata, M. M., Tomczak, M., Wijffels, S., and Church, J. A.: East Australian Current volume transports at 30°S: Estimates from the World Ocean Circulation Experiment hydrographic sections PR11/P6 and the PCM3 current meter array, Journal of Geophysical Research: Oceans, 105, 28 509–28 526, https://doi.org/10.1029/1999JC000121, 2000.

Mattern, J. P., Edwards, C. A., and Moore, A. M.: Improving variational data assimilation through background and observation error adjustments, Monthly Weather Review, 146, 485–501, 2018.

Matthews, D., Powell, B., and Janeković, I.: Analysis of four-dimensional variational state estimation of the Hawaiian waters, Journal of Geophysical Research: Oceans, 117, 2012.

Mazloff, M. R., Heimbach, P., and Wunsch, C.: An eddy-permitting Southern Ocean state estimate, Journal of Physical Oceanography, 40, 880–899, 2010.

Mignac, D., Tanajura, C., Santana, A., Lima, L., and Xie, J.: Argo data assimilation into HYCOM with an EnOI method in the Atlantic Ocean, Ocean Science, 11, 195–213, 2015.

Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Weaver, A. T., and Zavala-Garay, J.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part I–System overview and formulation, Progress in Oceanography, 91, 34–49, 2011a.

Moore, A. M., Arango, H. G., Powell, B., Broquet, G., Edwards, C., Veneziani, M., Foley, D., Doyle, J. D., Costa, D., and Robinson, P.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part II–Performance and application to the California Current System, Progress in Oceanography, 91, 50–73, 2011b.

Moore, A. M., Levin, J., Arango, H. G., and Wilkin, J.: Assessing the performance of an ocean observing, analysis and forecast System for the Mid-Atlantic Bight using array modes, Ocean Modelling, 164, 101 821, 2021.

O'Callaghan, J., Mike, B., and Fiona, E.: Across-shelf mooring array from the northeast shelf of New Zealand, SEANOE, https://doi.org/https://doi.org/10.17882/78971, 2015.

Oke, P., Schiller, A., Griffin, D., and Brassington, G.: Ensemble data assimilation for an eddy-resolving ocean model of the Australian region, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 131, 3301–3311, 2005.

Oke, P. R. and Schiller, A.: Impact of Argo, SST, and altimeter data on an eddy-resolving ocean reanalysis, Geophysical Research Letters, 34, 2007.

Oke, P. R., Sakov, P., Cahill, M. L., Dunn, J. R., Fiedler, R., Griffin, D. A., Mansbridge, J. V., Ridgway, K. R., and Schiller, A.: Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3, Ocean Modelling, 67, 52–70, 2013.

Oke, P. R., Larnicol, G., Fujii, Y., Smith, G. C., Lea, D. J., Guinehut, S., Remy, E., Balmaseda, M. A., Rykova, T., Surcel-Colan, D., et al.: Assessing the impact of observations on ocean forecasts and reanalyses: Part 1, Global studies, Journal of Operational Oceanography, 8, s49–s62, 2015.

Oke, P. R., Roughan, M., Cetina-Heredia, P., Pilo, G. S., Ridgway, K. R., Rykova, T., Archer, M. R., Coleman, R. C., Kerry, C. G., Rocha, C., et al.: Revisiting the circulation of the East Australian Current: Its path, separation, and eddy field, Progress in Oceanography, 176, 102 139, 2019.

Osafune, S., Masuda, S., Sugiura, N., and Doi, T.: Evaluation of the applicability of the Estimated State of the Global Ocean for Climate Research (ESTOC) data set, Geophysical Research Letters, 42, 4903–4911, 2015.

Pasmans, I. and Kurapov, A. L.: Ensemble of 4DVARs (En4DVar) data assimilation in a coastal ocean circulation model, part I: methodology and ensemble statistics, Ocean Modelling, 144, 101 493, 2019.

Pasmans, I., Kurapov, A., Barth, J., Ignatov, A., Kosro, P., and Shearman, R.: Why Gliders Appreciate Good Company: Glider Assimilation in the Oregon-Washington Coastal Ocean 4DVAR System With and Without Surface Observations, Journal of Geophysical Research: Oceans, 124, 750–772, 2019.

Pasmans, I., Kurapov, A. L., Barth, J. A., Kosro, P. M., and Shearman, R. K.: Ensemble 4DVAR (En4DVar) data assimilation in a coastal ocean circulation model. Part II: Implementation offshore Oregon-Washington, USA, Ocean Modelling, 154, 101 681, 2020.

Phillipson, L. and Toumi, R.: Impact of data assimilation on ocean current forecasts in the Angola Basin, Ocean Modelling, 114, 45–58, 2017.

Powell, B.: Quantifying how observations inform a numerical reanalysis of Hawaii, Journal of Geophysical Research: Oceans, 122, 8427–8444, 2017.

Powell, B., Arango, H., Moore, A., Di Lorenzo, E., Milliff, R., and Foley, D.: 4DVAR data assimilation in the intra-Americas sea with the Regional Ocean Modeling System (ROMS), Ocean Modelling, 23, 130–145, 2008.

Rodi, W.: Examples of calculation methods for flow and mixing in stratified fluids, Journal of Geophysical Research: Oceans, 92, 5305–5328, https://doi.org/https://doi-org.ezproxy.otago.ac.nz/10.1029/JC092iC05p05305, 1987.

Roemmich, D. and Sutton, P.: The mean and variability of ocean circulation past northern New Zealand: Determining the representativeness of hydrographic climatologies, Journal of Geophysical Research: Oceans, 103, 13 041–13 054, https://doi.org/10.1029/98JC00583, 1998.

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., Oke, P., Owens, W. B., Pouliquen, S., Purkey, S., et al.: On the future of Argo: A global, full-depth, multi-disciplinary array, Frontiers in Marine Science, 6, 439, 2019.

Sandery, P. A. and Sakov, P.: Ocean forecasting of mesoscale features can deteriorate by increasing model resolution towards the submesoscale, Nature communications, 8, 1–8, 2017.

Santana, R., Costa, F. B., Mignac, D., Santana, A. N., Vidal, V. F. d. S., Zhu, J., and Tanajura, C. A. S.: Model sensitivity experiments on data assimilation, downscaling and tides for the representation of the Cape São Tomé Eddies, Ocean Dynamics, 70, 77–94, https://doi.org/10.1007/s10236-019-01307-w, 2020.

Santana, R., Suanda, S., Macdonald, H., and O'Callaghan, J.: Mesoscale and wind-driven intra-annual variability in the East Auckland Current, Scientific Reports, 9764, https://doi.org/https://doi.org/10.1038/s41598-021-89222-3, 2021.

Shchepetkin, A. F. and McWilliams, J. C.: A method for computing horizontal pressure-gradient force in an oceanic model with a nonaligned vertical coordinate, Journal of Geophysical Research: Oceans, 108, 2003.

Shchepetkin, A. F. and McWilliams, J. C.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean modelling, 9, 347–404, 2005.

Shchepetkin, A. F. and McWilliams, J. C.: Computational kernel algorithms for fine-scale, multiprocess, longtime oceanic simulations, in: Handbook of Numerical Analysis, vol. 14, pp. 121–183, Elsevier, 2009.

Siripatana, A., Kerry, C., Roughan, M., Souza, J. M. A. C., and Keating, S.: Assessing the Impact of Nontraditional Ocean Observations for Prediction of the East Australian Current, Journal of Geophysical Research: Oceans, 125, e2020JC016 580, https://doi.org/https://doi.org/10.1029/2020JC016580, e2020JC016580 2020JC016580, 2020.

Sloyan, B. M., Ridgway, K. R., and Cowley, R.: The East Australian Current and Property Transport at 27°S from 2012 to 2013, Journal of Physical Oceanography, 46, 993–1008, https://doi.org/10.1175/JPO-D-15-0052.1, 2016.

Stanton, B. and Sutton, P.: Velocity measurements in the East Auckland Current northeast of North Cape, New Zealand, New Zealand Journal of Marine and Freshwater Research, 37, 195–204, https://doi.org/10.1080/00288330.2003.9517157, 2003.

Stanton, B. R. and Morris, M. Y.: Direct velocity measurements in the Subantarctic Front and over Campbell Plateau, southeast of New Zealand, Journal of Geophysical Research: Oceans, 109, https://doi.org/https://doi.org/10.1029/2002JC001339, 2004.

Stanton, B. R., Sutton, P. J. H., and Chiswell, S. M.: The East Auckland Current, 1994-95, New Zealand Journal of Marine and Freshwater Research, 31, 537–549, https://doi.org/10.1080/00288330.1997.9516787, 1997.

Taboada, F. G., Stock, C. A., Griffies, S. M., Dunne, J., John, J. G., Small, R. J., and Tsujino, H.: Surface winds from atmospheric reanalysis lead to contrasting oceanic forcing and coastal upwelling patterns, Ocean Modelling, 133, 79–111, https://doi.org/https://doi.org/10.1016/j.ocemod.2018.11.003, 2019.

Tanajura, C., Novaes Santana, A., Mignac, D., Nascimento Lima, L., Belyaev, K., and Ji-Ping, X.: The REMO ocean data assimilation system into HYCOM (RODAS_H): General description and preliminary results, Atmospheric and Oceanic Science Letters, 7, 464–470, 2014.

Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., Danabasoglu, G., Suzuki, T., Bamber, J. L., Bentsen, M., et al.: JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do), Ocean Modelling, 130, 79–139, 2018.

Umlauf, L. and Burchard, H.: A generic length-scale equation for geophysical turbulence models, Journal of Marine Research, 61, 2003.

Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, Quarterly Journal of the Royal Meteorological Society, 127, 1815–1846, 2001.

Weaver, A., Vialard, J., and Anderson, D.: Three-and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks, Monthly Weather Review, 131, 1360–1378, 2003.

Xie, J. and Zhu, J.: Ensemble optimal interpolation schemes for assimilating Argo profiles into a hybrid coordinate ocean model, Ocean Modelling, 33, 283–298, 2010.

Zavala-Garay, J., Wilkin, J., and Arango, H.: Predictability of mesoscale variability in the East Australian Current given strong-constraint data assimilation, Journal of physical oceanography, 42, 1402–1420, 2012.

Zeldis, J. R., Walters, R. A., Greig, M. J., and Image, K.: Circulation over the northeastern New Zealand continental slope, shelf and adjacent Hauraki Gulf, during spring and summer, Continental Shelf Research, 24, 543–561, https://doi.org/10.1016/j.csr.2003.11.007, 2004.

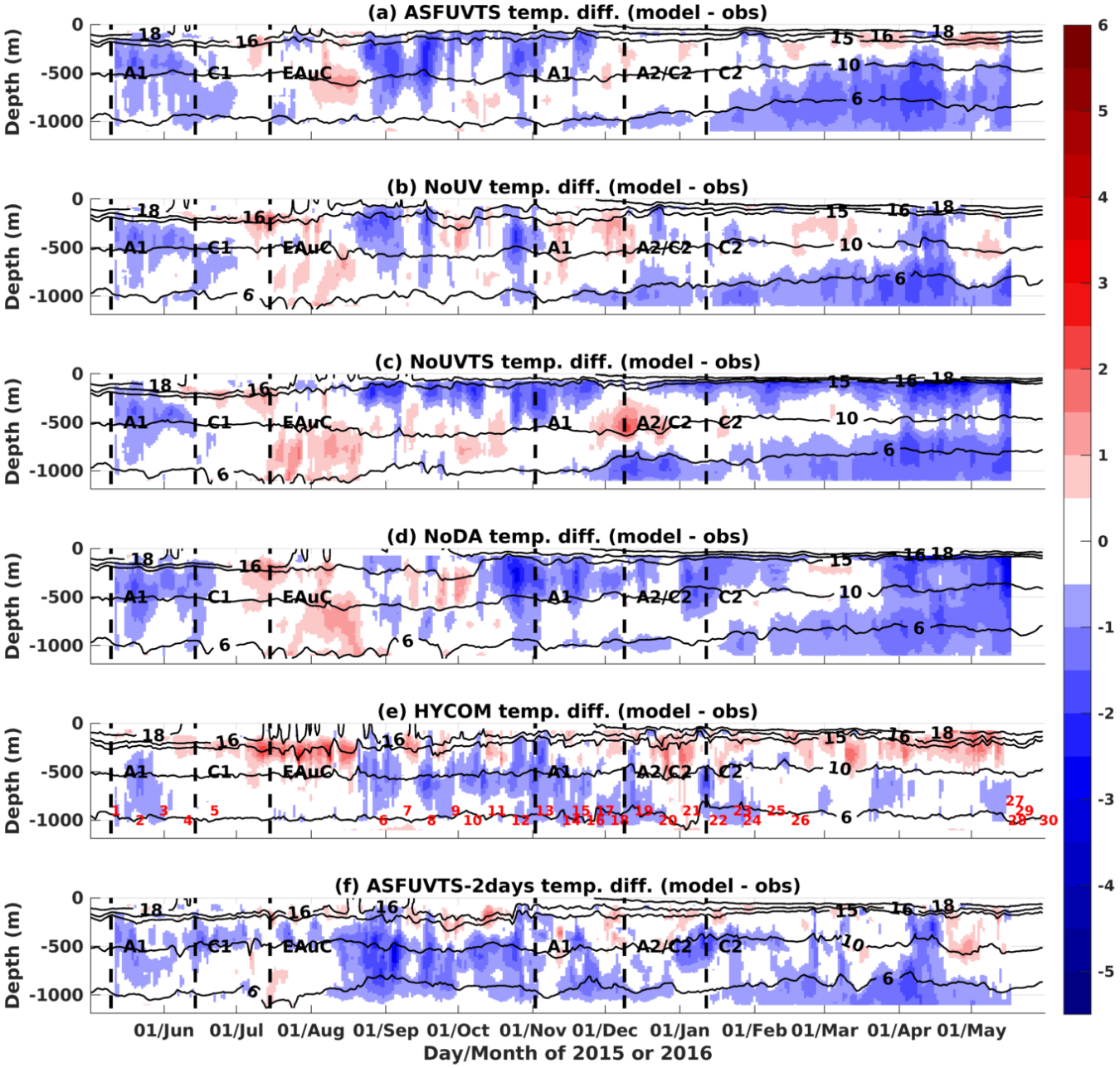

Figure 6. (a) ~~M4~~ Difference between model and observed daily average temperature ~~and difference between modelled and observed temperature~~ at M4 for experiments (~~b~~a) ASFUVTS, (~~e~~b) NoUV, (c) NoUVTS, (d) ~~NoTS~~NoDA, (e) ~~NoUVTS~~ HYCOM-NCODA, and (f) ~~NoDA~~ASFUVTS-2days. The red ~~, blue and cyan~~ numbers in (~~a~~e) show the dates ~~where~~ when Argo profilers sampled the NZNES. The ~~location~~ locations of the Argo floats are shown in Fig. ~~1a and~~ 10e. The acronyms A1, C1, EAuC, A1, A2/C2, and C2 represent mesoscale structures near M5 which were studied in Santana et al. (2021).

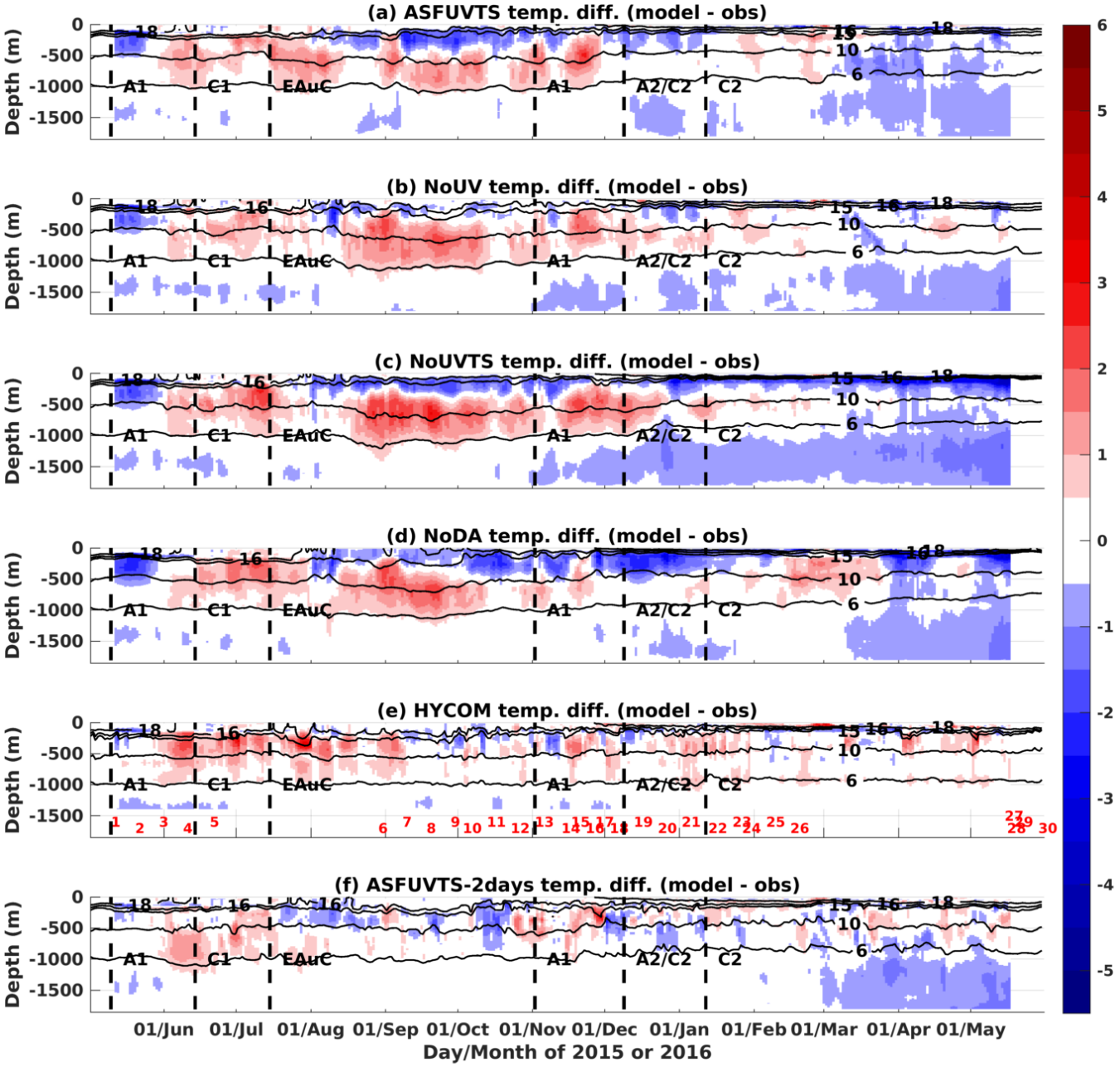

Figure 7. (a) ~~M5 Observed daily average temperature and difference~~ Difference between ~~modelled~~ model and observed daily average temperature ~~from~~ at M5 for experiments (~~b~~a) ASFUVTS, (~~c~~b) NoUV, (c) NoUVTS, (d) ~~NoTS~~NoDA, (e) ~~NoUVTS~~ HYCOM-NCODA, and (f) ~~NoDA~~ASFUVTS-2days. The red ~~, blue and cyan~~ numbers in (~~a~~e) show the dates ~~where~~ when Argo profilers sampled the NZNES. The ~~location~~ locations of the Argo floats are shown in Fig. ~~1a and~~ 10e. The acronyms A1, C1, EAuC, A1, A2/C2, and C2 represent mesoscale structures near M5 which were studied in Santana et al. (2021).

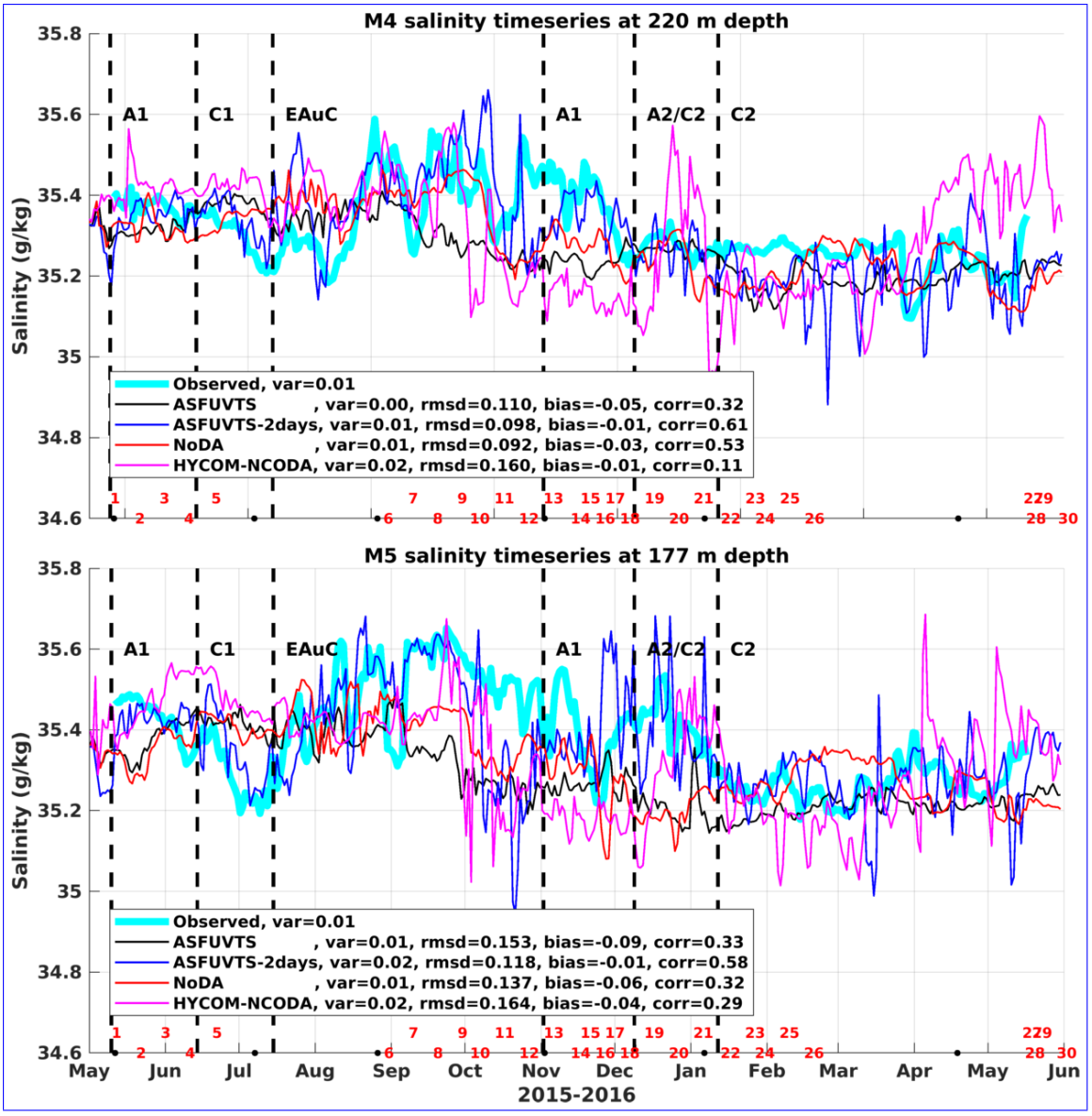

Figure 8. Observed (light blue) and modelled daily average salinity near 200 m depth at M4 (a) and M5 (b) from experiments ASFUVTS (black), ASFUVTS-2days (dark blue), NoDA (red), and HYCOM-NCODA (magenta/pink). Some experiments had their colour changed from previous plots for better visualisation. Var, rmsd, bias and corr represent variance, root mean square deviation, mean bias and correlation coefficient calculated between modelled and observed salinity. Red numbers show the dates when Argo profilers sampled the NZNES, and their locations are shown in Fig. 10e. The acronyms A1, C1, EAuC, A1, A2/C2, and C2 represent mesoscale structures near M5 which were studied in Santana et al. (2021).

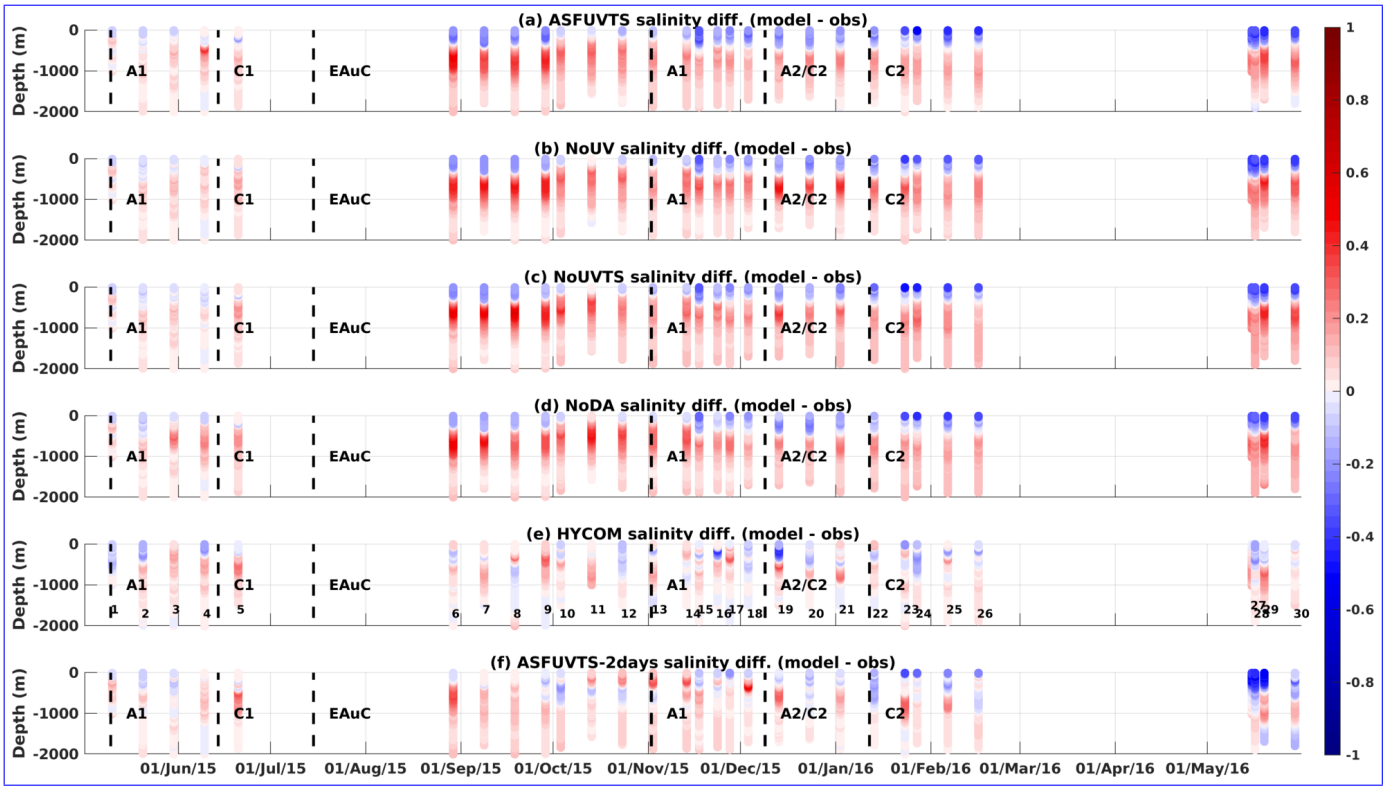

Figure 9. Salinity difference between modelled and Argo observations from experiments (a) ASFUVTS, (b) NoUV, (c) NoUVTS, (d) NoDA, (e) HYCOM-NCODA, (f) ASFUVTS-2days. The numbers in (e) show the dates when Argo profilers sampled the NZNES. The locations of the Argo floats are shown in Fig. 10e. The acronyms A1, C1, EAuC, A1, A2/C2, and C2 represent mesoscale structures near M5 which were studied in Santana et al. (2021).

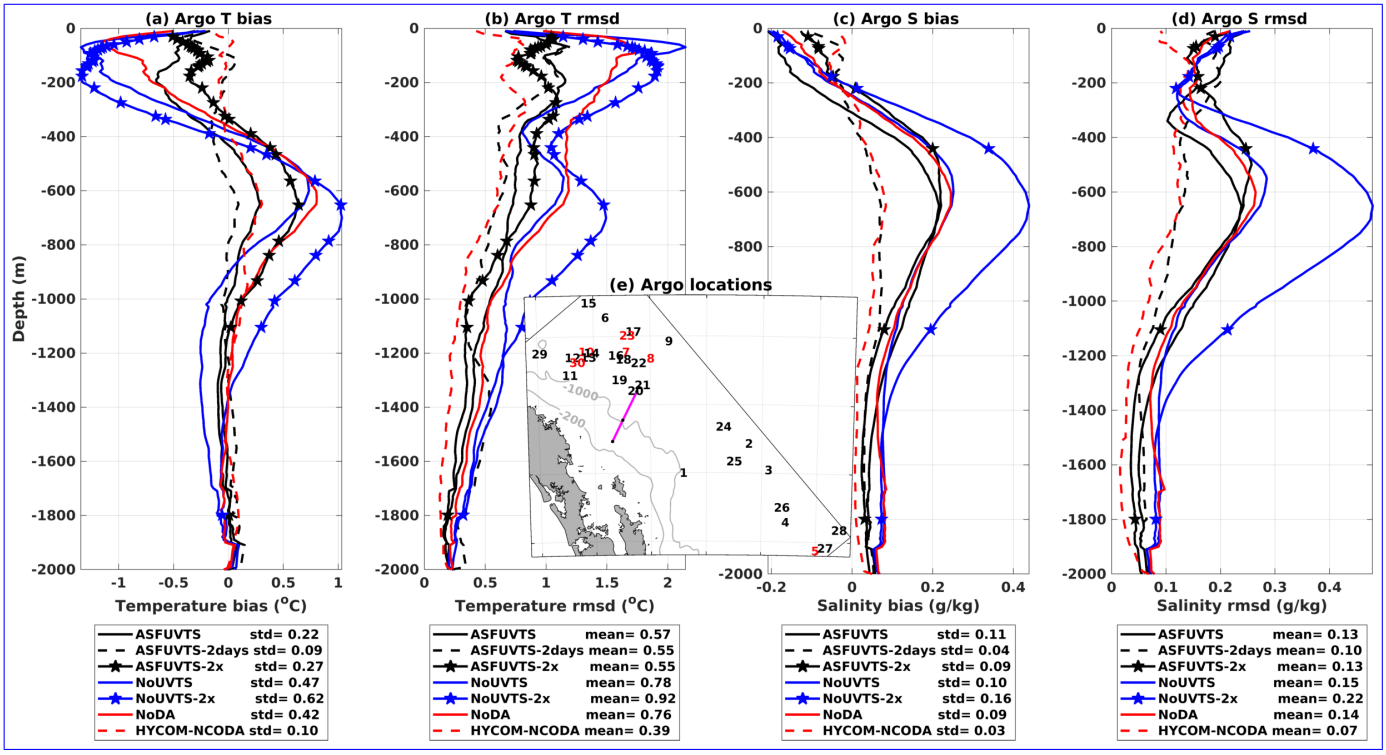

Figure 10. Profiles of mean bias and root mean square deviation (rmsd) between model and independent Argo observations of temperature (a,b) and salinity (c,d). ~~Black diamonds~~ Solid black line (ASFUVTS), ~~black dots and~~ dashed black line (ASFUVTS-2days), ~~magenta squares~~ starred black line (~~NoTS~~ASFUVTS-2x), ~~cyan triangles~~ solid blue line (~~NoUV~~NoUVTS), starred blue ~~stars~~ line (~~NoUVTS~~NoUVTS-2x)~~and~~, red ~~dots~~ line (NoDA), and dashed red line (HYCOM-NCODA). The stars represent the median depth of thermistors and CTDs (CTDs) at M3, M4 and M5 in temperature (salinity) bias and rmsd profiles. ~~HYCOM-NCODA results are~~ Results from experiments NoUV and NoTS were not shown ~~as dashed red lines~~for better visualisation. (e) Map of Argo locations which are represented as numbers in black or red for better visualisation only. The moorings' locations are shown as black dots on the magenta line.

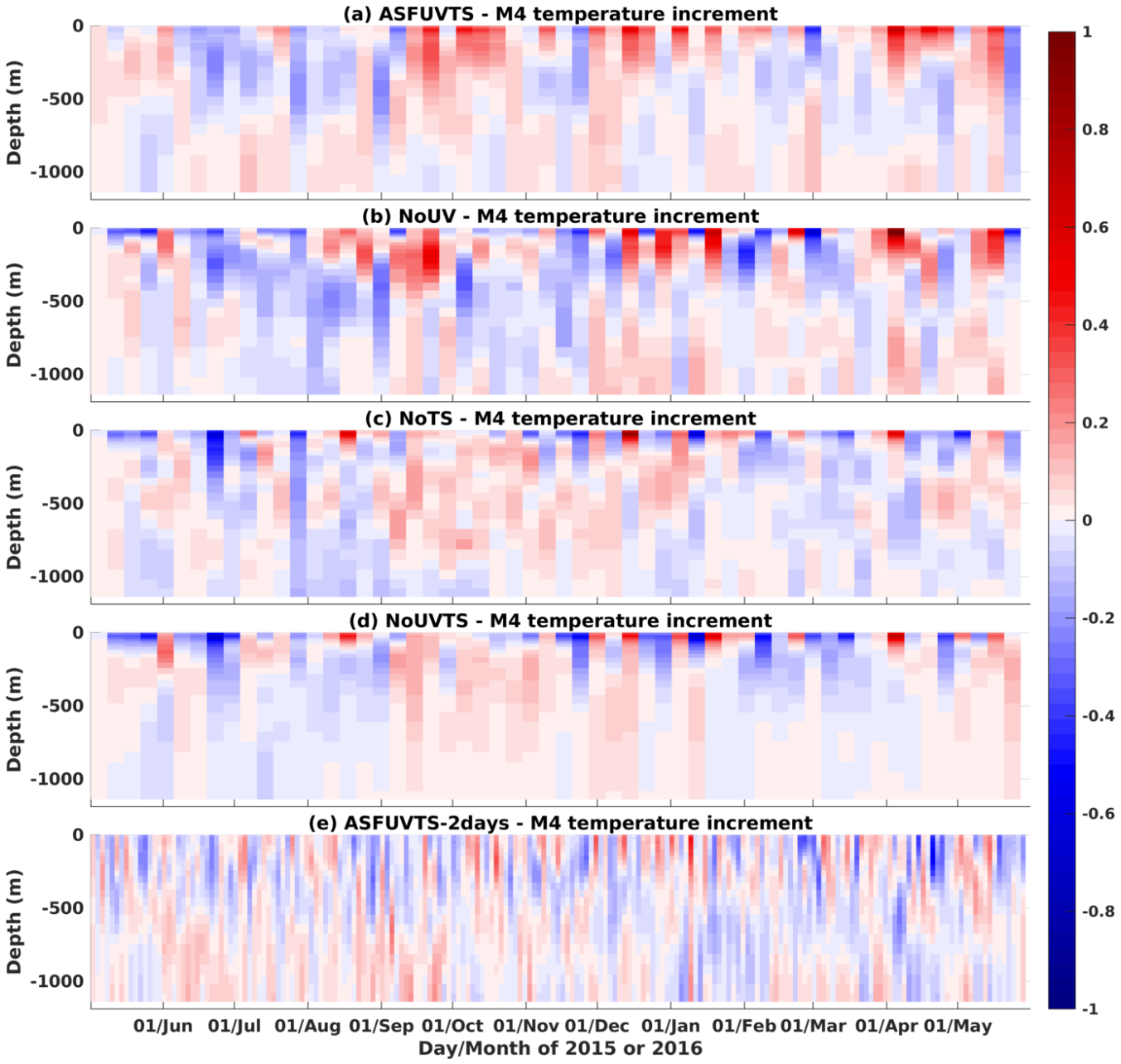

Figure 11. Time variability of temperature increments (°C) to the initial conditions at M4 from experiments (a) ASFUVTS, (b) NoUV run, (c) NoTS, (d) NoUVTS, and (e) ASFUVTS-2days.

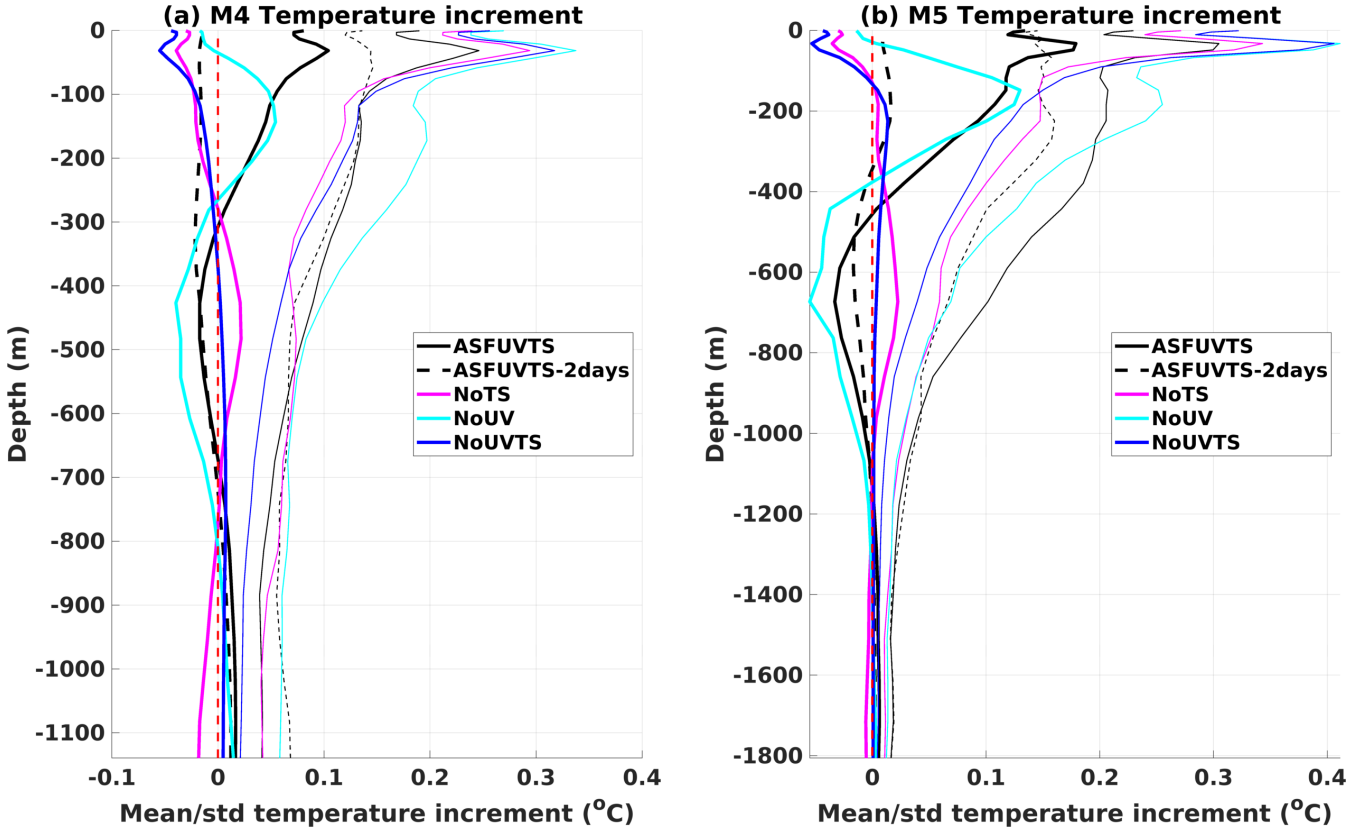

Figure 12. Profile of average (thick lines) and standard deviation (thin lines) increment to initial conditions of temperature at M4 (a) and M5 (b). Solid black line (ASFUVTS), dashed black line (ASFUVTS-2days), pink (magenta) line (NoTS), light blue (cyan) line (NoUV) and dark blue line (NoUVTS). The red dashed line marks the 0 °C increment.

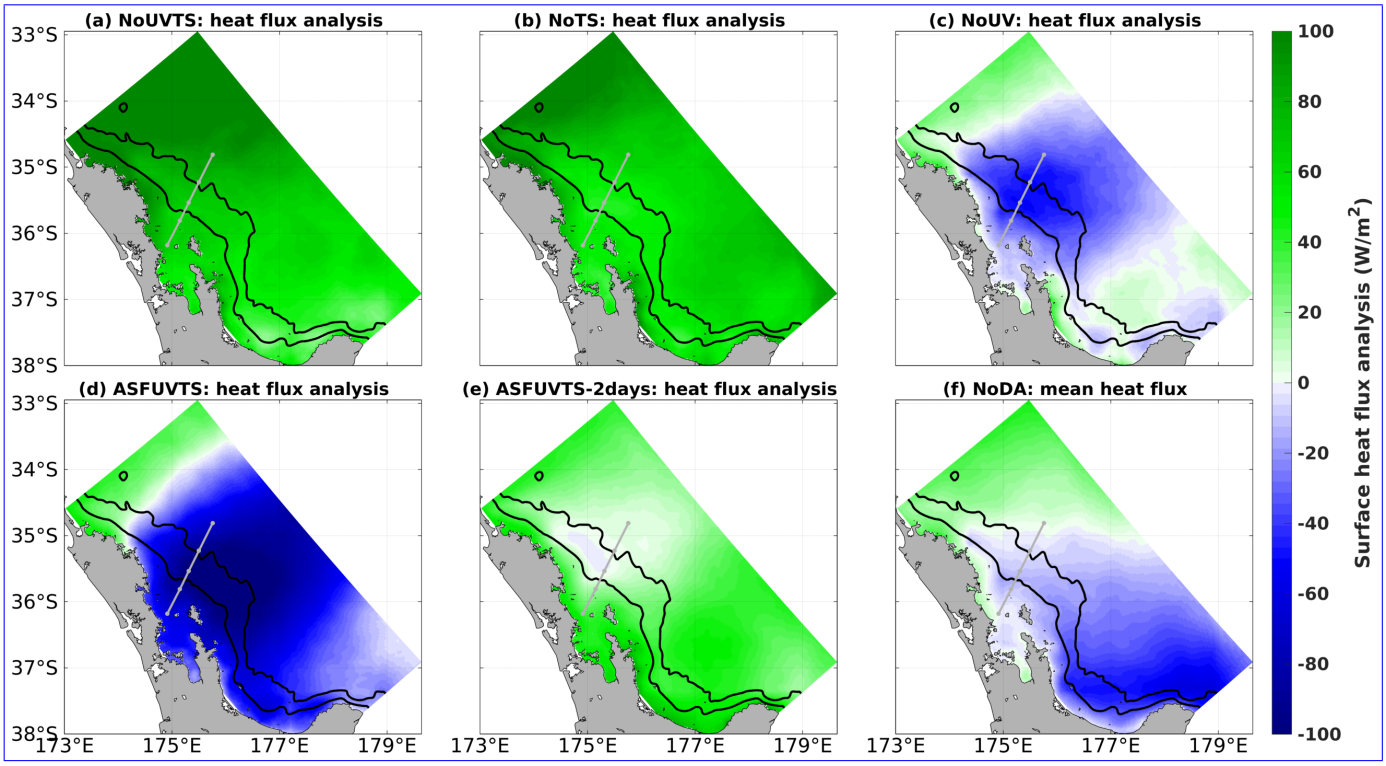

Figure 13. Average surface net heat flux ~~increments~~ analysis (W/m^2) (positive = downward) in experiments (a) NoUVTS, (b) NoTS, (c) NoUV, (d) ASFUVTS, and (e) ASFUVTS-2days. (f) shows mean net heat flux in the freely evolving simulation (NoDA). Green (blue) colours show positive (negative) mean surface heat flux.

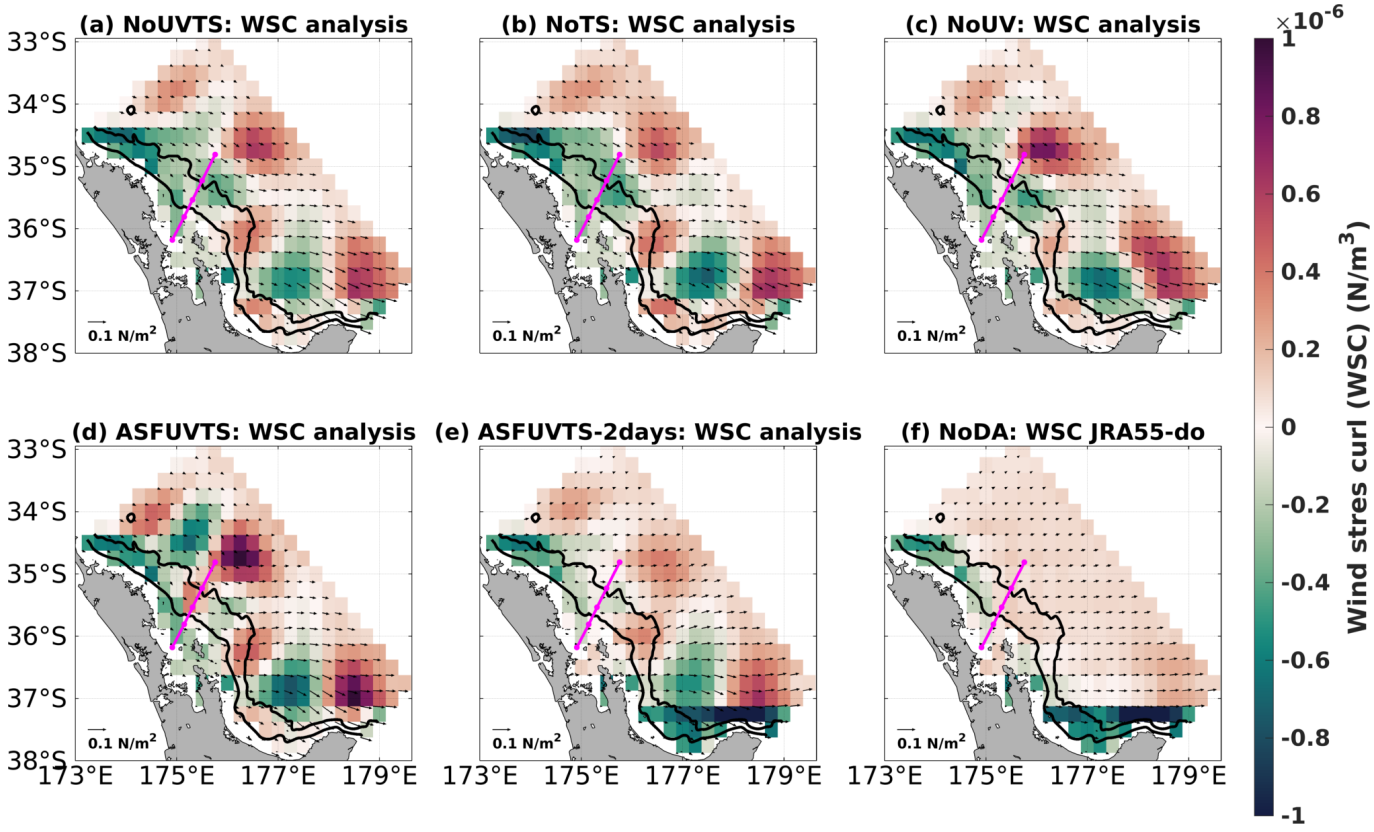

Figure 14. Analysis field of wind stress (N/m^2) and its curl (N/m^3) (red-green shade) in experiments (a) NoUVTS, (b) NoTS, (c) NoUV, (d) ASFUVTS, and (e) ASFUVTS-2days. (f) shows mean wind stress and its curl from JRA55-do. All fields were degraded from 2 km to 1/4° to be similar to JRA55-do horizontal resolution.

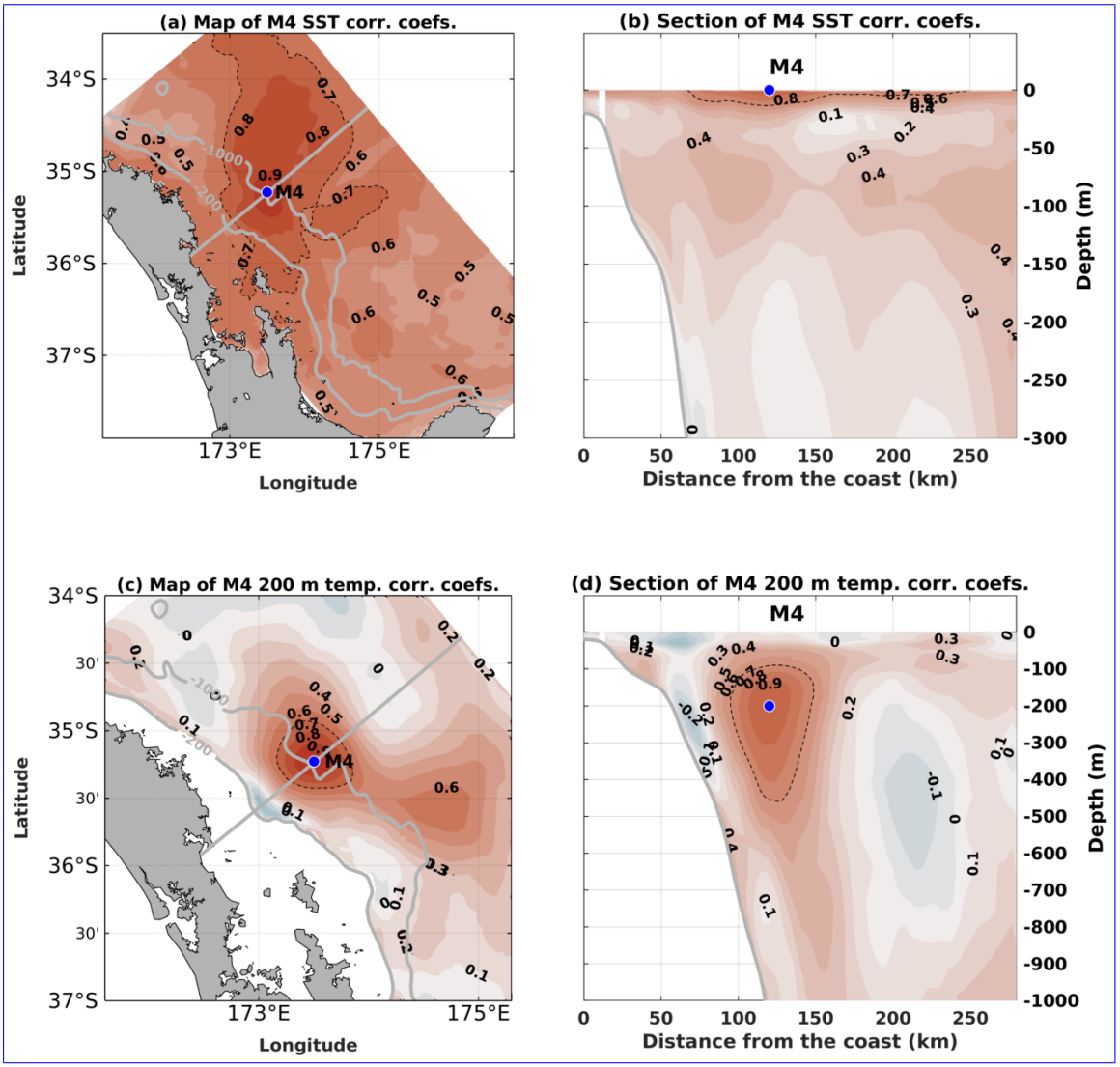

Figure 15. Maps (left panels) and cross-sections (right panels) of linear correlation coefficients between surface (top row) or 200 m (bottom row) temperature at M4 and points in the rest of the domain from the NoDA run. The blue dots show observation locations used to correlate with surrounding model grid points. The dashed black line shows regions with correlation coefficients equal to 0.7. The cross-section used here follows the grid alignment which is different from the moorings' orientation. The y-axis in (b) and (d) is on the right-hand side for better visualisation.