

Next generation of Bluelink ocean reanalysis with multiscale data assimilation: BRAN2020

Matthew A. Chamberlain¹, Peter R. Oke¹, Russell A. S. Fiedler¹, Helen Beggs², Gary B. Brassington³, Prasanth Divakaran²

¹ *CSIRO Oceans and Atmosphere, Hobart, TAS, Australia*

² *Bureau of Meteorology, Melbourne, VIC, Australia*

³ *Bureau of Meteorology, Sydney, NSW, Australia*

Abstract

BRAN2020 is an ocean reanalysis that combines ocean observations with an eddy-resolving, near-global ocean general circulation model, to produce four-dimensional estimates of the ocean state. The data assimilation system employed is ensemble optimal interpolation, implemented with a new multiscale approach that constrains the broad-scale ocean properties and the mesoscale circulation in two steps. The reanalysis spans January 1993 to December 2019, and assimilates observations of *in situ* temperature and salinity; and satellite sea-level anomaly and sea surface temperature. Reanalysed fields from BRAN2020 generally show much closer agreement to observations than all previous versions of this reanalyses, with mis-fits between reanalysed and observed fields reduced by over 30% for some variables. The BRAN2020 dataset is comprised of daily-averaged fields of temperature, salinity, velocity, mixed-layer depth, and sea-level. Reanalysed fields realistically represent all of the major current systems within 75°S and 75°N, excluding processes relating to sea ice, but including boundary currents, equatorial circulation,

southern ocean variability, and mesoscale eddies. BRAN2020 is publicly-available and is intended for use by the research community.

1. Introduction

An ocean reanalysis combines ocean observations with a model to produce gridded estimates of the ocean state. The dataset presented here is the 2020 version of the Bluelink ReANalysis (BRAN2020; 2020 refers to the year in which the configuration was finalised). Previous versions of BRAN include BRAN1 (Oke et al., 2005), BRAN1p5 (Oke et al., 2008), BRAN2 (Schiller et al., 2008), BRAN3 (Oke et al., 2013b), BRAN2015 (Oke et al., 2018, when the naming convention changed), and BRAN2016 (same configuration as BRAN2015 but initialised in the 1990s). Overall, the accuracy of BRAN2020 is much better than all previously presented versions.

There are two main differences between BRAN2020, and all previous versions of reanalyses in this series. The first difference is the adoption of a multiscale data assimilation technique (Chamberlain et al., 2020) that constrains the broad-scale features of the ocean and the mesoscale features separately, in two steps. Chamberlain et al. (2020) showed that this delivers an improvement in accuracy – measured by the mis-fit to assimilated and withheld observations – of up to 35%. The same improvement is realised here, in BRAN2020. The second difference is the assimilation of a larger observational dataset that includes observations from marine mammals, moorings, and ship-borne surveys that were not assimilated into previous versions of BRAN. Previous versions only assimilated observations from the conventional observation platforms (Argo; eXpendable BathyThermographs, XBTs;

23 ship-borne surveys; satellite sea-level anomaly, SLA; and satellite sea-surface
24 temperature, SST). The new sources of data offer greater spatial coverage,
25 with fewer gaps.

26 This paper is intended to provide a description of how BRAN2020 data
27 are generated, with sufficient detail that a user assess whether the data are
28 suitable for their intended application. The structure of the paper is as fol-
29 lows. Section 2 describes the ocean model, data assimilation system, and
30 observations assimilated. The discussion of the data assimilation system
31 includes a summary of the multiscale data assimilation approach, with tech-
32 nical details of the assimilation reported in Appendix A. Section 3 describes
33 the results, including an evaluation of the quality of reanalysed fields and
34 inter-comparisons with previous BRAN experiments. Section 4 describes the
35 output available, access and structure of files, and indicates potential uses of
36 the BRAN2020 dataset.

37 **2. Ocean Reanalysis Methods**

38 In this section we describe the procedures followed to generate BRAN2020.
39 We provide a description of the analysis cycle, how the model and observa-
40 tions are combined; and descriptions of the model, the data assimilation
41 system, and the observations assimilated.

42 *2.1. Analysis cycle*

43 The initial condition for the ocean state at the beginning of BRAN2020
44 is interpolated from climatology, using the 2013 version of the World Ocean
45 Atlas (WOA13; Zweng et al., 2013; Locarnini et al., 2013). The first three
46 months of the integration (October-December 1992; comprising 30 analysis

cycles) are not included in the publicly available dataset. Reanalysed fields during this period are dominated by unphysical adjustments to initial conditions.

The analysis cycle involves alternate running of the ocean model and data assimilation system. For BRAN2020, observations are assimilated every three days. The instantaneous fields for salinity, sea level and horizontal velocity, and a daily-averaged fields of temperature (for the day immediately preceding the analysis time) are used as the background field for the data assimilation step. Daily-averaged temperature is used as the background field instead of instantaneous fields to avoid systematic, regional biases due to the diurnal cycle of temperature near the sea-surface that depends on local time of day. Temperature fields are also converted from potential temperature (used by the model) to *in situ* temperature to match the assimilated observations. This conversion has no impact at the surface, but changes the temperature by about 0.2°C at 2000 m depth.

The differences between the background field and observations are calculated by linearly-interpolating the model fields to the observation locations, yielding a vector of background innovations. Background innovations represent the differences between the model and observations before assimilation. The data assimilation system is used to calculate the analysis field (fully described in Section 2.3). The differences between the analysis field and observations are also calculated, yielding a vector of analysis innovations. Analysis innovations represent the differences between the model and observations after assimilation. The differences between the analysis field and the background fields are here referred to as the increments. Increments include

72 values for all variables of the model state: temperature, salinity, horizontal
73 velocities, and sea-level; for all model grid points. The increments are added
74 to the instantaneous fields at the end of the previous analysis cycle, for all
75 variables, to reinitialise the model for the next analysis cycle. For salinity,
76 horizontal velocities, and sea-level, this means that the model is initialised
77 with the analysis fields. But because temperature is treated slightly dif-
78 ferently, as described above, the initialised temperature does not precisely
79 match the analysis fields.

80 This analysis cycle is repeated every three days for the period of January
81 1993 to December 2019. Background and analysis innovations, for each cycle,
82 are used to quantify the accuracy of the reanalysis in Section 3. Daily-
83 averaged fields of temperature, salinity, velocity, sea-level and mixed layer
84 depth are stored for each day of integration. The mixed layer depth is here,
85 is the depth over which the buoyancy exceeds a threshold of 0.0003 m/s^2 ,
86 as described by Griffies (2012). These daily-averaged fields comprise the
87 BRAN2020 product that is presented here.

88 *2.2. Ocean model*

89 The model configuration used here is the Ocean Forecasting Australian
90 Model, version 3 (OFAM3). A comprehensive description of OFAM3 is pre-
91 sented by Oke et al. (2013a). Briefly, OFAM3 is a near-global, eddy-resolving,
92 z^* configuration of the Modular Ocean Model (version 5, Griffies, 2012), de-
93 veloped principally for the purpose of reanalyses and forecasting upper ocean
94 conditions across the globe, excluding the polar regions. The model grid has
95 0.1° grid resolution between 75°S and 75°N , with 5-m vertical resolution
96 down to 40-m depth and 10-m vertical resolution to 200-m depth. OFAM3

is forced with atmospheric conditions from JRA-55 (Kobayashi et al., 2015; Tsujino et al., 2020) using bulk formulas to calculate surface fluxes of heat, freshwater and momentum. The model is ocean-only, and does not include sea ice. Forcing fields are masked for sea ice extent from the atmospheric reanalysis, to avoid unrealistic surface fluxes that would result from bulk formulas with open water exposed to extreme winter atmospheric conditions at high-latitudes.

Surface salinity is restored to monthly-averaged climatology, using WAO13 (Zweng et al., 2013), with a restoring time-scale of 14 days. Temperature and salinity below 2000 m are also relaxed towards WAO13 climatology (Zweng et al., 2013; Locarnini et al., 2013), with a restoring time-scale of 1 year. The configuration of OFAM3 used here employs the K-epsilon scheme (Rodi, 1987) to calculate unresolved turbulent mixing in the upper ocean.

2.3. Ocean data assimilation system

The assimilation system used to constrain BRAN2020 is EnKF-C (Sakov, 2014), implemented in the Ensemble Optimal Interpolation (EnOI) mode. Technical details, specific to the EnKF-C software including all parameters and settings, are presented in Appendix A. For EnOI, the state of the ocean model, \mathbf{w} (comprised of temperature, salinity, horizontal velocity, and sea-level), is updated by projecting background innovations onto an ensemble of model anomalies, \mathbf{A} , and calculating weightings, \mathbf{c} , for each member. The weights vary horizontally and are recalculated for each analysis cycle. The increment, used to update the ocean state, is a weighted sum of ensemble members:

$$\mathbf{w}^a = \mathbf{w}^b + \mathbf{w}^{inc} \quad (1)$$

$$= \mathbf{w}^b + \sum_{i=1}^n \mathbf{c}_i(x, y) \cdot \mathbf{A}_i(x, y, z), \quad (2)$$

where superscripts a , b , and inc denote analysis, background, and increment fields, respectively; subscripts i denote the i th ensemble member; n is the ensemble size; and x , y , and z are dimensions in space.

The weights, in equation (2), depend on many factors, including the assumed relative size of the observation error variance and background field error variance. These weights also depend on the consistency between the background innovations and the anomalies in the ensemble members. If, for example, there is no combination of ensemble members that can “fit” the background innovations, then the ensemble can be considered rank-deficient - lacking sufficient degrees of freedom to fit the data. In that case, the analyses won’t match the observations well. The details of the assumed observation errors, and the details of the ensembles can have significant impact of the performance of the data assimilation and the overall reanalysis.

For BRAN2020, the ensembles are constructed as anomalies derived with output from previous experiments using free-running models with no data assimilation. A strength of this method is that it is multivariate. This means that observations of one variable can be used to calculate increments for other variables. For example, the assimilation of SLA results in an update to velocity, even when velocity observations are not assimilated. To achieve this, the ensemble fields are used to characterise the statistical relationships - the covariance - between each model variable at each point in space. In practice, the covariability of different variables at different locations are represented by the ensemble members. By constructing the ensembles in this way, we’re

144 making an assumption that the background field errors can be represented
 145 by the sample of anomalies in these ensembles. Moreover, we’re assuming
 146 that there are background field errors on both broad scales and mesoscales.

147 For BRAN2020, the data assimilation is implemented in two steps – the
 148 coarse- and high-resolution assimilation – using two different ensembles at
 149 the different resolutions. The coarse-resolution ensemble is generated by
 150 a one-degree resolution global ocean model (the ocean component of the
 151 Australian Community Climate and Earth System Simulator, ACCESS; Bi
 152 et al., 2013), using a configuration like Kiss et al. (2019) with the sea ice
 153 model distributed with the ocean model (Sea Ice Simulator, Delworth et al.,
 154 2006). The coarse ensemble is constructed by differencing monthly-averaged
 155 fields from a 40-year climatology, with one member constructed for each
 156 month of a free-running 40-year model simulation. This ensemble is intended
 157 to represent the potential broad-scale background field errors. The high-
 158 resolution ensemble is generated by a OFAM3 spinup (Oke et al., 2013a).
 159 The high-resolution ensemble is constructed by centred differences of 3-day
 160 from 3-month averages, with one member constructed for each month of a 12-
 161 year model simulation, without data assimilation. This ensemble is intended
 162 to represent the potential mesoscale background field errors.

163 The first step, the coarse-resolution assimilation, is intended to correct the
 164 broad scales of the ocean state. To calculate increments for the coarse assimi-
 165 lation step, the background field on the high-resolution grid $\mathbf{w}_{high-res}^b$, is first
 166 spatially-averaged onto a one-degree grid, yielding \mathbf{w}_{coarse}^b . The spatially-
 167 averaged background field is compared to observations, yielding the back-
 168 ground innovations for the coarse-resolution assimilation step. EnOI is used

169 to project these innovations onto the 480-member coarse-resolution ensemble
 170 of anomalies to produce a coarse-increment $\mathbf{w}_{coarse}^{inc}$. The increments are inter-
 171 polated back to the 0.1° grid, yielding $\mathbf{w}_{inter-coarse}^{inc}$, and added to $\mathbf{w}_{high-res}^b$.
 172 This gives an updated background field on the high-resolution grid $\mathbf{w}_{high-res}^{updated-b}$:

$$\mathbf{w}_{high-res}^{updated-b} = \mathbf{w}_{high-res}^b + \mathbf{w}_{inter-coarse}^{inc}, \quad (3)$$

173 ready for the second step of the assimilation. We find that this first step
 174 largely constrains the broad scales of the ocean state, without impacting the
 175 mesoscales.

176 The second step, the high-resolution assimilation, is intended to correct
 177 the mesoscales of the ocean state and uses the same configuration as recent
 178 versions of BRAN (Oke et al., 2018). The updated background field on the
 179 high-resolution grid, $\mathbf{w}_{high-res}^{updated-b}$, is compared to observations (the same obser-
 180 vations assimilated in the first step) yielding the background innovations for
 181 the high-resolution assimilation step. EnOI is used to project these innova-
 182 tions onto a 144-member high-resolution ensemble of anomalies to produce
 183 a high-resolution increment $\mathbf{w}_{high-res}^{inc}$. These increments are added to the
 184 respective variables in the model’s restart files, and the model is integrated
 185 forwards in time. We find that this step largely constrains the mesoscales of
 186 the ocean, mainly adjusting the locations and properties of mesoscale eddies,
 187 and adjusting the strength, locations, and properties of currents.

188 2.4. Ocean observations

189 BRAN2020 assimilates observations from satellite SST, satellite SLA, and
 190 *in situ* temperature and salinity. BRAN2020 uses updated sources of obser-
 191 vational data than previous versions of BRAN. Figure 1 is a “Gannt” chart

192 summarising the temporal extent of various observation products that are
193 assimilated into BRAN2020.

194 For most years, the European Space Agency SST Climate Change Initia-
195 tive (CCI-SST, Merchant et al., 2019) is the source of Advanced Very High
196 Resolution Radiometer (AVHRR, Embury et al., 2019a) and Along Track
197 Scanning Radiometer (ATSR, Embury et al., 2019b) SST observations. Data
198 hosted by CCI-SST include updated calibrations and quality controls, span-
199 ning from 1981 to 2016; “best quality” (level 5), day- and night-time tem-
200 peratures at depth (0.2 m) data are assimilated into BRAN (noting the top
201 layer of BRAN is 5 m, so temperature at 0.2 m is equivalent to the simu-
202 lated SST). From year 2017 onwards, AVHRR data (again, day and night
203 time observations) are sourced from NAVO (Naval Oceanographic Office,
204 2014a,b,c,d). Night-time microwave SST data from AMSRE (Gentemann
205 et al., 2010) and AMSR2 (retrieved from the Japan Aerospace Exploration
206 Agency, <ftp://ftp.eorc.jaxa.jp/AMSR2/GDS2.0/>) are assimilated from 2009,
207 where available. Microwave SST data are not as accurate as infrared ob-
208 servations, but microwave SST data are not as affected by clouds (in the
209 absence of rain), thus providing a good complement to infrared-based ob-
210 servations. Night-time SST data from Visible Infrared Imaging Radiometer
211 Suite (VIIRS, Petrenko et al., 2014; NOAA Office of Satellite and Prod-
212 uct Operations, 2019) aboard Suomi National Polar Partnership (NPP) and
213 NOAA-20 satellites, are included from 2012 and 2018 respectively. The VI-
214 IRS SST data have somewhat superior spatial coverage to AVHRR SST, due
215 to the higher spatial resolution (0.75 km to 1.5 km for VIIRS compared with
216 9 km to 30 km)

217 Observations of SLA are assimilated into BRAN2020, using along-track
 218 satellite altimeter data from various satellite platforms that have been avail-
 219 able since the 1990s (Figure 1). SLA data are sourced from the Radar Al-
 220 timeter Database System (RADS Ver. 4, Scharroo et al., 2013), and include
 221 corrections for astronomical tides and inverse barometer effects. SLAs in the
 222 model are referenced to a mean sea-level, calculated from an 18-year mean
 223 of an OFAM3 experiment without data assimilation (the same experiment
 224 that was used to construct the high-resolution ensemble). SLA observations
 225 are not assimilated in water less than 200 m.

226 *In situ* observations of temperature and salinity are assimilated into BRAN2020,
 227 using measurements from Argo floats, surface drifting buoys, XBT (temper-
 228 ature only), sea mammals, moorings, as well as ship-borne surveys, and un-
 229 derway systems. These data are sourced from the Coriolis Ocean Dataset for
 230 ReAnalysis CORA (CORA, versions 5.0 and 5.1, Cabanes et al., 2013) and
 231 from a near-real time (NRT) database maintained at the Australia Bureau of
 232 Meteorology. CORA databases include data that have been processed with
 233 delayed-mode quality control. The NRT database, used for the last couple
 234 of years of BRAN2020 (Figure 1), includes data sourced from the Global
 235 Telecommunications System and the Argo Global Data Assembly Centres.
 236 These data are subject to some NRT quality control and are duplicate-
 237 checked to exclude multiple versions of the same observations (Brassington
 238 et al., 2012).

239 For every assimilated observation, an explicit observation error estimate
 240 is required. Specifically, the standard deviation of observation errors need
 241 to be estimated. Observations are assumed to be unbiased. Observation

242 errors include measurement error and representation error (e.g., Oke and
 243 Sakov, 2008). Because observation errors (especially representation errors)
 244 are relatively poorly known, the estimates used for BRAN2020 were tuned
 245 somewhat, to improve the performance based on ocean reanalyses simulat-
 246 ing recent years. Table 1 summarises all observation classes and the assumed
 247 standard deviations of each observation type. As is common in data assimi-
 248 lation problems, it is possible to force the model ocean state to closely match
 249 observation data, but this can lead to overfitting and may cause unrealistic
 250 artefacts in the model that can degrade the quality of the reanalysis. As a
 251 result, assumed observation errors tend to be larger than some might expect.
 252 The one exception to this here, is the assumed estimate of AVHRR SST
 253 error (Table 1). For BRAN2020, we assume that the standard deviation of
 254 the AVHRR SST error is 0.1°C . In retrospect, based on the recent results
 255 presented by Merchant et al. (2019, their Figure 6) and the assessment in
 256 Section 3, we think that the BRAN2020 performance may have been better
 257 for upper-ocean temperature if we’d assumed a larger error for AVHRR SST
 258 - perhaps with a value between 0.2 and 0.4°C . We plan to assess this for
 259 future versions of BRAN in the future.

260 The observation errors actually used by EnKF-C in the data assimilation
 261 calculations are influenced by a process, called “super-obing.” This is a pre-
 262 processing step applied at the start of each data assimilation cycle to combine
 263 observations of the same type and at approximately the same locations and
 264 times (in most cases, within the same model grid point). The error variance
 265 of a super-observation is calculated from the inverse of the sum of inverse vari-
 266 ances of the individual observations used to construct the super-observation

267 (Sakov, 2014). As a result, the error variance of each super-observation is
 268 often smaller than the assumed observation error variances of the individ-
 269 ual observations. The practice of “super-obing” is common, reducing the
 270 computational cost of the assimilation step, and as a conservative measure,
 271 eliminating some anomalous observations that may represent fluctuations not
 272 represented by the model (e.g., sub-gridscale features).

273 **3. Evaluation and Assessment Ocean Reanalysis**

274 To demonstrate the quality of the BRAN2020 dataset, we present two
 275 classes of metrics. First, we present the model-observation differences com-
 276 puted during the assimilation steps - specifically looking at the background
 277 and analysis innovations. Second, we compare daily-averaged reanalysed
 278 fields (the fields released as part of the BRAN2020 dataset) with observa-
 279 tions, observation-based indices, and observation-based properties.

280 *3.1. Analysis of innovations*

281 Here, we report statistics of the background and analysis innovations, be-
 282 fore and after assimilation. This provides insight into how closely BRAN2020
 283 analyses “fit” the assimilated observations (using the analysis innovations),
 284 and how faithfully the model integrates analyses forward in time (using the
 285 background innovations). Comparison of the analysis and background inno-
 286 vations can help identify potential cases of over-fitting. Ideally, we hope for
 287 small analysis innovations and small background innovations. If analysis in-
 288 novations are small, but background innovations are large, we might suspect
 289 over-fitting.

290 The mean absolute deviations of the analysis and background innovations
 291 are presented in Table 2 and 3, respectively. For BRAN2020, statistics are
 292 presented for decade-averaged and globally-averaged innovations for the SST
 293 and SLA; and for sub-surface temperature and salinity in different depth
 294 ranges. The decade averages are presented for the 1990s, 2000s, and 2010s,
 295 to show how the quality of the reanalysed fields changes with time, as the
 296 observing system changes (Figure 1) and as the ocean variability changes.
 297 For comparison, the same statistics are included for BRAN2016. For al-
 298 most all metrics reported in Table 2 and 3, BRAN2020 fields are in better
 299 agreement with observations than BRAN2016. The only exception to this is
 300 temperature in the top 50 m in the 2000s.

301 The greatest improvement in BRAN2020, compared to BRAN2016, is in
 302 first decades of sub-surface temperature and salinity fields (Table 2 and 3).
 303 For depths greater than 500 m, the analysis and background innovations for
 304 both temperature and salinity are about 30% smaller in BRAN2020 than
 305 they were in BRAN2016. There is greater improvement in the reported in-
 306 novations of the 1990s, though innovations in this decade are higher due
 307 to the sparsity of *in situ* data in this decade. Between 50 and 500 m
 308 depth, the analysis and background innovations are between 20 and 40%
 309 smaller in BRAN2020. For depths shallower than 50 m, the analysis and
 310 background innovations for temperature are 11-14% greater in BRAN2020
 311 than they were in BRAN2016; and for salinity, they are 5-14% smaller in
 312 BRAN2020. For SST, the analysis and background innovations are 7-22%
 313 smaller in BRAN2020; and for SLA, they are 3-6% smaller in BRAN2020
 314 than they were in BRAN2016.

315 The sub-surface fields were known to have relatively large biases in earlier
 316 versions of BRAN (e.g., Oke et al., 2013b). This seems to me mostly elim-
 317 inated with the use of the multiscale data assimilation (Chamberlain et al.,
 318 2020). We think that the slight degradation in temperature, shallower than
 319 50 m, is an indication that the BRAN2020 analyses are over-fitting SST. As
 320 noted in Section 2.4, we think that a better result may have been achieved if
 321 we assumed a larger observation error for AVHRR SST for BRAN2020 (Ta-
 322 ble 1). We note that VIIRS SST is reported as higher in spatial resolution
 323 to AVHRR SST and generally agrees closer to drifting buoy SST (Minnett
 324 et al., 2020; O’Carroll et al., 2019, their Fig. 10), although this is not re-
 325 flected in our error estimates in Table 1 – and so we surmise that post-2012
 326 (when VIIRS data are assimilated) the impact of this slight over-fitting is
 327 much less. Note also, ATSR has less error relative to AVHRR, as assessed
 328 with drifting buoys (Merchant et al., 2019, their Fig. 6), so the assumed
 329 ATSR error of 0.1 °C (Table 1) is a reasonable.

330 To better understand whether BRAN2020 analyses over-fit any of the
 331 assimilated observations, we present differences between the mean absolute
 332 values of background and analysis innovations for BRAN2020 and BRAN2016
 333 in Table 4. Comparison between BRAN2016 and BRAN2020 (in the 2000s,
 334 column 3 of Table 4), suggests that the error growth between analysis cy-
 335 cles during the 2000s is faster for BRAN2020. This may be because either
 336 BRAN2020 over-fits some observations, or because BRAN2016 under-fits
 337 some observations - or perhaps both. We can’t be sure of which explana-
 338 tion is true. But as noted, we think that BRAN2020 over-fits AVHRR SST;
 339 and we are quite sure that BRAN2016 under-fit sub-surface temperature and

340 salinity below 50 m, and particularly below 500 m. These results demonstrate
341 one of the main challenges of performing these big reanalysis experiments.
342 There are many factors to “tune” in these systems, and there are compet-
343 ing goals of “fitting” data as close as possible, but not over-fitting. It looks
344 like for BRAN2020, the configuration is achieving a good balance - fitting
345 the data much more closely than previous versions across most of the ocean
346 state, but perhaps overfitting some variables in some places.

347 Times series of the mean absolute values of the background and analy-
348 sis innovations are presented for SLA, SST, temperature (all depths), and
349 salinity (all depths) in Figure 2. These are the same metrics used to prepare
350 Tables 2, 3, and 4. The low-frequency of the SLA innovations, in Figure 2a,
351 fluctuates modestly in time, with a few step-changes when the composition of
352 satellite altimeters changed (Figure 1). For example, the innovations increase
353 when data from the Topex/Poseidon satellites become unavailable at the end
354 of the 1990s, and they reduce again when data from the Jason/Envisat satel-
355 lites are after 2002.

356 Over the course of BRAN2020, there is a gradual decrease in SST innova-
357 tions (Figure 2b), as the number of satellites increased and sensors improved
358 (O’Carroll et al., 2019). As noted in Figure 1, the main source of SST ob-
359 servations for BRAN2020 is CCI-AVHRR. But BRAN2020 also assimilated
360 microwave SST from 2009, and VIIRS SST from 2012. There is no clear step
361 change when microwave SST is used, but there is a clear reduction in both
362 analysis and background innovations when VIIRS SST data are assimilated
363 from 2012. VIIRS SST data have been observed to generally have smaller
364 standard deviations when compared with drifting buoy SSTs than either

365 AVHRR or AMSR2 SST products (O’Carroll et al., 2019; Helen Beggs, pers.
366 comm.).

367 The constraint of sub-surface temperature and salinity improves consis-
368 tently over the course of BRAN2020 (Figure 2c,d). This is particularly true
369 over the decade of the 2000s directly associated with the expansion of the
370 Argo float array that provides systematic observations of sub-surface ocean
371 properties with near-global coverage (Roemmich et al., 2019). Prior to the
372 Argo array, *in situ* data were scattered in space and sporadic in time. Not
373 only do more *in situ* observations improve analyses - with smaller anal-
374 ysis innovations - but they also result in better quality “forecasts”, with
375 smaller background innovations. Another prominent feature of Figure 2c,d,
376 is a seasonal cycle in the global averaged innovations from the beginning of
377 BRAN2020, until about 2008. The seasonal cycle in innovations largely dis-
378 appears once observations from the global Argo array nominally reached its
379 target density in about 2008 (Wong et al., 2020).

380 The analysis of innovations computed during the assimilation steps of
381 BRAN2020, presented above, is intended to give an indication of the quality
382 of the BRAN2020 product. A few caveats regarding these comparisons are
383 warranted. A long, centered observation window is used for sub-surface ob-
384 servations in the coarse-resolution data assimilation step of BRAN2020 (see
385 Appendix A). This means that some of the observations included to calculate
386 the background innovations are assimilated in previous assimilation cycles.
387 For SST and SLA, the observations used for the background innovations are
388 all independent - having not been used to constrain the reanalysis. But for
389 sub-surface temperature and salinity, this is not always the case.

390 Regarding the comparisons between BRAN2020 and BRAN2016, we again
 391 note that the data sources and volume of data assimilated in each experi-
 392 ment is different, and the reanalyses are initialised and forced differently.
 393 The main difference between BRAN2016 and BRAN2020 is the approach
 394 to data assimilation - with multiscale assimilation used in BRAN2020. To
 395 more clearly attribute the improvements in BRAN2020, Chamberlain et al.
 396 (2020) used the same configuration as BRAN2020 to compare reanalyses
 397 using precisely the same observations and using data assimilation configu-
 398 rations of BRAN2016 and BRAN2020. Their results show small changes in
 399 surface metrics of the ocean state (SST and SLA) in BRAN2020 relative to
 400 the BRAN2016 configuration, similar to those reported here; and significant
 401 improvements in sub-surface temperature and salinity - again, with similar
 402 improvements to those reported above. We therefore conclude that the main
 403 improvements evident in BRAN2020, compared to BRAN2016, relate to the
 404 new multiscale data assimilation approach (Chamberlain et al., 2020).

405 *3.2. Assessment of daily-averaged reanalysed fields*

406 *Point-wise comparisons*

407 The analysis of innovations presented above are based on metrics cal-
 408 culated during the data assimilation steps of BRAN2020. However, the
 409 BRAN2020 dataset that is publicly available is comprised of daily-averaged,
 410 reanalysed fields that are produced by the model in between analysis up-
 411 dates. To assess these fields explicitly, we now compare the daily-averaged
 412 fields with observations for each day of BRAN2020; 9861 daily comparisons

413 from the start of 1993, to the end of 2019.

414 For SLA and SST, and for sub-surface temperature and salinity in depth
415 ranges of 0-50m and 50-500 m, we group the model-data differences into
416 $10 \times 10^\circ$ bins for the period 2008-2019 across the entire model domain. The
417 results for each of these fields are presented in Figures 3, 4, 5, and 6. For
418 each figure, the top panel shows the map of Mean Absolute Deviation (MAD)
419 between observed and daily-mean reanalysed fields; and the bottom panels
420 show the difference between the MAD on day 3 (the day before each assimilation
421 step) and day 1 (the day after each assimilation step) of each analysis
422 cycle. These fields are intended to demonstrate the level of agreement between
423 BRAN2020 fields and observations, and how differences grow in each
424 analysis cycle. We expect that this should provide users with a clear idea of
425 how accurate BRAN2020 fields are for all regions.

426 The MADs between reanalysed and observed SLA on day 1 of each assimilation
427 cycle average 4.5 cm globally, with values of 6-12 cm in western
428 boundary currents (WBCs), 4-5 cm along the Antarctic Circumpolar Current
429 (ACC), and 2-3 cm in equatorial regions and in the more quiescent parts of
430 the ocean (Figure 3). These MADs tend to grow by 1-2 cm in WBCs, by less
431 than 2 cm along the ACC, and by less than 1 cm in the equatorial and less
432 variable regions of the ocean.

433 For SST, The MADs between reanalysed and observed fields on day 1
434 of each assimilation cycle average 0.26°C globally, $0.2\text{-}0.6^\circ\text{C}$ in WBCs, 0.3-
435 0.4°C along the ACC, and with MADs of less than 0.2°C in equatorial and
436 quiescent parts of the ocean (Figure 4). These MADs tend to grow by 0.1-
437 0.3°C in WBCs and along the ACC, and by less than 0.05°C for much of the

438 equatorial ocean. Unlike the growth in SLA, MADs for SST grow significantly
439 in eastern boundary current regions, with increases of 0.1-0.2°C.

440 Figures 5 and 6 show similar distributions of MAD and MAD growth
441 within the assimilation cycle for temperature and salinity, in two depth
442 bands, 0-50 m and 50-500 m. For temperature in the top 50 m, the MADs
443 between reanalysed and observed fields on day 1 of each assimilation cycle
444 average 0.31°C globally, with values of 0.4-0.8°C in WBCs, values of less
445 than 0.3°C along most of the ACC, and less than 0.2°C in the equatorial and
446 quiescent ocean (Figure 5). These MADs for temperature in the top 50 m
447 tend to grow by 0.2-0.3°C in WBCs, by 0.1-0.2°C along the ACC, and by
448 less than 0.1°C for much of the equatorial ocean. Like SST, the upper ocean
449 temperature MAD growth by day 3 is larger in eastern boundary currents,
450 with increases of up to 0.3°C.

451 For temperature between 50 and 500 m depth, the MADs between reanal-
452 ysed and observed fields on day 1 of each assimilation cycle average 0.4°C
453 globally, with values of 0.4-0.6°C in WBCs, less than 0.3°C along most of
454 the ACC, about 0.5°C in equatorial regions, and 0.3°C the quiescent ocean
455 (Figure 5). These MADs tend to grow by 0.3-0.6°C in WBCs, by 0.3-0.4°C
456 along the ACC, and by 0.2-0.3°C for much of the equatorial ocean.

457 For salinity in the top 50 m, the MADs between reanalysed and observed
458 fields on day 1 of each assimilation cycle average 0.08 psu globally, with values
459 of 0.04-0.2 psu in WBCs, less than 0.05 psu along most of the ACC, and less
460 than 0.06 psu in equatorial and quiescent ocean (Figure 6). These MADs for
461 salinity in the top 50 m tend to grow by about 0.1 psu equatorward of 30°S
462 and 30°N, and by less than 0.03 psu in the Southern Ocean.

463 For salinity between 50 and 500 m depth, the MADs between reanal-
464 ysed and observed fields on day 1 of each assimilation cycle average 0.06
465 psu globally, with values of 0.04-0.2°C WBCs, values of less than 0.05 psu
466 along most of the ACC, and values of about 0.06 psu in equatorial regions
467 (Figure 6). Those MADs for salinity between 50 and 500 m tend to grow
468 by about 0.02-0.03 psu in the Pacific Ocean basin, by 0.02-0.1 psu in the
469 Atlantic, by 0.02-0.07 psu in the Indian Ocean, and by less than 0.04 psu in
470 the Southern Ocean.

471 Figure 7 summarises comparisons between reanalysed and observed tem-
472 perature as a function of depth and time. Figure 7a shows profiles of sub-
473 surface temperature MAD with depth for different time periods. Clearly
474 shown is the decrease in MAD from the 1990s to the 2010s (Figure 7a; com-
475 pare the solid and dashed lines) as the number of sub-surface observations
476 increase with data from the Argo float array becoming available (Figure 7c).
477 Separate profiles of averaged MAD (Figure 7a) are presented for regions of
478 high- and low-variability, as indicated in the mask map (Figure 7b; blue de-
479 notes high-variability regions), and averaged over times with and without
480 Argo data. For this analysis, the regions of high variability are where the
481 standard deviation of detrended SLA exceeds 8 cm. The choice of this cut-
482 off value is arbitrary, but this value means that the high-variability region
483 includes all of the WBC regions, the path of the ACC, and some of the
484 high-variability tropical current systems.

485 Figure 7 shows that MAD is lowest in regions of low variability when
486 Argo is available, and highest in high-variability regions pre-Argo. Figure 7d-
487 f shows Hovmöller plots of model-data MAD fields for temperature for all

regions (panel d), high-variability regions (panel e), and low-variability regions (panel f). For temperature, MAD is highest in the upper thermocline and below the surface (~ 50 -200 m) and a seasonal cycle is evident in the pre-Argo period, peaking in the middle of each year. MAD at the surface is relatively low, where reanalysed fields are well-constrained by SST observations. At depth (> 500 m) temperature MAD are low everywhere, even at the beginning of BRAN2020, but reduce further with Argo data.

The profiles of salinity MAD with depth (Figure 8) are similar to those of temperature, with variability decreasing everywhere with Argo and increasing number of global observations. Without the equivalent of high-quality SST observations for salinity, the MAD is high at the surface. There is less difference in salinity MAD between high- and low-variability regions, compared to temperature; most of this difference is in the top 200 m.

501 *Sea-level variability*

To further assess the variability of the reanalysed circulation, we compare estimates of sea-level variability, here quantified using the standard deviation of detrended sea-level (Figure 9), for BRAN2020, a free-running model (Spinup-EI), and an observation-based analysis product (using Aviso-Ssalto/Duacs). Spinup-EI is a simulation with the same global ocean model used for BRAN2020, but is run without data assimilation (Oke et al., 2013a). The run used here is equivalent to the run that is used to construct the high-resolution ensemble described in Section 2.3. The Aviso-Ssalto/Duacs product is a multi-mission product, generated by analysing along-track SLA observations to produce weekly maps of SLA on a $1/4^\circ$ grid, and spans 1993

512 to 2014 (accessed from www.aviso.altimetry.fr); the sea level variability fields
 513 from models are calculated over the same years. The comparisons in Figure 9
 514 show that the simulated SLA variability is very similar for all products. The
 515 model simulation (Spinup-EI) realistically reproduces the key features of the
 516 observed field. However, we consider the comparisons between BRAN2020
 517 and the Aviso-Ssalto/Duacs product to be significantly better. The locations,
 518 spatial-extent, and amplitude of the local maxima of SLA variability in Fig-
 519 ure 9, agree more closely between BRAN2020 and Aviso-Ssalto/Duacs, than
 520 Spinup-EI. Figure 9 confirms that the reanalysed circulation in BRAN2020
 521 is realistic and closely aligned with other observational estimates.

522 *SST variability*

523 Figure 10 shows a comparison of some key SST-based climate indices, cal-
 524 culated using observations (black) and BRAN2020 (red) fields. This includes
 525 comparisons for Niño3, Niño4, the Interdecadal Pacific Oscillation (IPO; as
 526 defined by Henley et al., 2015), and the Indian Ocean Dipole (IOD). There is
 527 a very close match between the observed and reanalysed Niño indices and the
 528 IPO (Figure 10a-c). The agreement for the IOD (Figure 10d) is not as good.
 529 For this comparison, the observed IOD is calculated from a 1°-resolution
 530 HadISST product (Rayner et al., 2003). The discrepancies shown in Fig-
 531 ure 10 for the IOD are mostly associated with the eastern node, a significant
 532 part of which crosses islands of Indonesia (see the boxes in Figure 11a). SST
 533 estimates in different products depend on resolution, particularly in coastal
 534 regions with high variability. As a result, the details of the “observed” and
 535 reanalysed IOD differ in detail. But despite this, the observed and reanalysed

536 IOD are still very consistent.

537 Figure 11 shows maps of SST variability, here quantified as the standard
538 deviation of SST anomalies, for three products: BRAN2020, Spinup-EI, and
539 gridded CCI-SST observation product (version 2.1, Merchant et al., 2019;
540 Good et al., 2019). All estimates are calculated for the period spanning
541 2000-2009, with respect to climatologies from the same period. SST vari-
542 ability is greatest in regions of significant mesoscale eddy variability, such
543 as WBCs and along the path of the ACC. SST variability is also high in
544 places associated with interannual modes, such as the equatorial regions and
545 the midlatitudes, which include the nodes of the SST indices indicated. Fig-
546 ure 11 shows that the SST variability simulated by a model without data
547 assimilation is realistic. But the comparison between the observation-based
548 estimate (CCI-SST) and BRAN2020 shows exceptionally good agreement.
549 In practice, it's difficult to identify any region of disagreement for the colour
550 scales shown here. The only subtle differences, evident in this comparison,
551 is in regions of seasonal sea ice cover, which BRAN does not include.

552 4. Data availability and uses

553 Output from BRAN2020 is available from NCI (Australia National Com-
554 puting Infrastructure) OpenDAP servers (**doi!?**) - **to be added**. Daily
555 averages of temperature, salinity, velocity, and mixed layer depth are avail-
556 able in monthly files in netCDF format. Three components of velocity are
557 provided, however the component of vertical was not saved in the first 5 years
558 of BRAN2020. Monthly and annual averages of the same variables are also
559 available.

560 Creation of the BRAN2020 dataset required significant resources. Apart
561 from the human resources, BRAN2020 was performed on up to 1200 cpus,
562 using 1 TB of memory, over a period of about six months. The BRAN2020
563 dataset itself requires ~ 15 TB of disk-space.

564 In the past, data from BRAN experiments have been used for studies
565 of ocean general circulation (e.g., Schiller et al., 2008; Chiswell et al., 2015;
566 Feng et al., 2016; Oke et al., 2019); regional oceanography (e.g., Schiller
567 et al., 2009; Oke and Griffin, 2011; Pilo et al., 2015; Schiller and Oke, 2015;
568 Oke et al., 2018); the impacts of different components of the ocean observing
569 system on data-assimilating models (e.g., Oke and Schiller, 2007; Oke et al.,
570 2009); and applications to other fields, such as fisheries (e.g., Hartog et al.,
571 2011) and marine biodiversity (e.g., Coleman et al., 2011). We offer the
572 BRAN2020 dataset for use in similar applications.

573 For most metrics, BRAN2020 outperforms all previously presented re-
574 analyses in this series of BRAN products. We attribute the improved perfor-
575 mance of BRAN2020 mainly to the adoption of multiscale data assimilation
576 (Chamberlain et al., 2020). This delivered the greatest benefit for sub-surface
577 temperature and salinity, where the previously-acknowledged bias (Oke et al.,
578 2013b) has been largely eliminated. As a result, agreement between observed
579 and reanalysed temperature and salinity fields below 50 m are of the order
580 of 30% less in BRAN2020, compared to previous versions of BRAN.

581

582 **Acknowledgements**

583 Production of BRAN2020 is supported by the Bluelink Project, a part-

nership between CSIRO, the Australian Bureau of Meteorology and the Australian Department of Defence. Data was sourced from the Integrated Marine Observing System (IMOS, www.aodn.org.au). IMOS is a national collaborative research infrastructure, supported by the Australian Government. CCI L2P SST data from AVHRR and ATSR sensors were produced at the Met Office as part of the European Space Agency (ESA) SST Climate Change Initiative (CCI) Project, funded by ESA. AMSR-E L2P SST data were provided by GHRSSST and REMSS. AMSR2 L2P SST data were provided by GHRSSST and JAXA/EORC. The AVHRR L2P SST data from the Naval Oceanographic Office are made available under Multi-sensor Improved Sea Surface Temperature (MISST) project sponsorship by the Office of Naval Research (ONR). The VIIRS L3U SST data were provided by Group for High Resolution Sea Surface Temperature (GHRSSST) and the National Oceanic and Atmospheric Administration (NOAA). Argo data were collected and made freely available by the International Argo Program and the national programs that contribute to it. (www.argo.ucsd.edu, argo.jcommops.org). The Argo Program is part of the Global Ocean Observing System. This research used computation resources and archives available at the National Computational Infrastructure (NCI), which is located at the Australian National University and supported by the Australian Government.

604

605 **Appendix A: Ensemble Optimal Interpolation**

606 For BRAN2020, EnOI is applied in a two-step multiscale data assimilation
607 approach, as described in Section 2.1 and by Chamberlain et al. (2020).

608 Here, we present the technical details of the data assimilation, with specific
 609 reference to the parameters and settings (in Table 5) used with the open-
 610 source EnKF-C software (Sakov, 2014). We also present the data assimilation
 611 equations to provide context for these parameters and settings.

612 The details in Table 5 include the data assimilation cycle length, the en-
 613 semble size for each assimilation step, the method for ensemble construction,
 614 the localisation radii, the stride, the R-factor, K-factor, observation windows,
 615 and the domain over-which observations are assimilated. Each of these points
 616 are explained below.

617 The cycle length for BRAN2020 is 3 days; meaning that data are assim-
 618 ilated every 3 days. The ensemble size for the coarse- and high-resolution
 619 data assimilation steps are 480 members and 144 members respectively. The
 620 ensemble for the coarse-resolution assimilation is intended to represent broad-
 621 scale anomalies. The ensemble for the high-resolution assimilation is intended
 622 to represent anomalies associated with mesoscale anomalies. In practice, we
 623 would like to make the ensemble sizes as large as possible. These are limited
 624 only by computational resources.

625 For the coarse-resolution assimilation, data is only assimilated within
 626 65°S and 65°N, to avoid impacts of sea ice that are represented in the coarse-
 627 resolution ensemble. For the coarse-resolution assimilation step, Argo data
 628 are assimilated for a 10-day window, centred on the analysis time. For all
 629 other observations, data are assimilated for a 3-day centred window.

630 The EnKF-C data assimilation system, in EnOI mode, solves the analysis
 631 equation:

$$\mathbf{w}^a = \mathbf{w}^b + \mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{w}^b), \quad (4)$$

632 where \mathbf{w} is the ocean state vector, comprised of temperature, salinity, sea-
633 level, and horizontal velocities; \mathbf{P} is the background error covariance matrix
634 - the diagonals of which are the assumed background error variance; \mathbf{R} is
635 the observation error covariance matrix - here assumed to be diagonal, with
636 values that are the assumed observation error variance (from values in Ta-
637 ble 1); \mathbf{H} is an operator that interpolates from model-space to observation
638 locations; and \mathbf{y} is the vector of observations.

639 The matrices \mathbf{P} and \mathbf{R} are never explicitly constructed. Instead, equa-
640 tion (4) is solved by EnKF-C by operating with the ensemble anomalies \mathbf{A} ,
641 and the diagonal elements of \mathbf{R} . In practice, the increments, used to update
642 the model state, can be expressed as:

$$\mathbf{w}^{inc} = \sum_{i=1}^n \mathbf{c}_i(x, y) \cdot \mathbf{A}_i(x, y, z), \quad (5)$$

643 where subscripts i denote the i th ensemble member; n is the ensemble size;
644 x , y , and z are dimensions in space, and \mathbf{c} contains the weights for each
645 ensemble member.

646 In theory, without localisation, \mathbf{c} (in equation 5) does not vary in space.
647 But for most practical applications, with a large state dimension, this works
648 poorly because the ensemble is rank-deficient - lacking a sufficient number of
649 degrees of freedom to appropriately “fit” the background innovations (e.g.,
650 Oke et al., 2007). For BRAN2020, and most other applications of EnOI,
651 localisation is used to increase the effective rank of the ensemble, and to
652 eliminate unrealistic, long-distance ensemble-based covariances. This results
653 in spatially inhomogeneous weightings, with $\mathbf{c}(x, y)$ typically a function of
654 horizontal location.

655 Use of localisation also means that observations beyond the localising
 656 length-scale from a grid point have no impact on analyses at that grid point.
 657 This is exploited by EnKF-C for computational efficiency, by projecting back-
 658 ground innovations onto the ensemble separately for individual horizontal
 659 grid points. This means that instead of a single - practically impossible -
 660 calculation, the assimilation problem is reduced to a large number of smaller
 661 calculations, with one calculation for each horizontal grid point. The lo-
 662 calising function used by EnKF-C is a quasi-Gaussian function, where the
 663 function goes to exactly zero over the localisation radius. The effective e -
 664 folding length-scale of the localising function is ~ 3.5 times smaller than the
 665 localisation length-scale reported here (and elsewhere in the literature).

666 It's expected that the weighting coefficients in \mathbf{c} , vary smoothly in space,
 667 with values for adjacent grid points being very similar. This permits a further
 668 simplification, to reduce the computational cost, with the use of a “stride”
 669 (Sakov, 2014). In practice, weights are only calculated on a subset of grid
 670 points, determined by the stride (where a stride of 3, for example, means
 671 that coefficients are only calculated for every third grid point). When a
 672 stride greater than 1 is used, weights (\mathbf{c}) are linearly interpolated onto the
 673 model grid before constructing the increments, using equation (5).

674 In EnKF-C, weights are calculated by:

$$\mathbf{c} = \mathbf{S}^T(\mathbf{I} + \mathbf{S}\mathbf{S}^T)^{-1}\mathbf{s}, \quad \text{where} \quad (6)$$

$$\mathbf{S} = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{A}/\sqrt{n-1}, \quad \text{and} \quad (7)$$

$$\mathbf{s} = \mathbf{R}^{-1/2}(\mathbf{y} - \mathbf{H}\mathbf{w}^b)/\sqrt{n-1}. \quad (8)$$

675 EnKF-C includes an option to use a “K-factor”. The K-factor is intended

676 to prevent overfitting, by modifying observation variances, following the for-
 677 mula:

$$\sigma_{obs}^2 = [(\sigma_b^2 + \sigma_{obs}^2)^2 + \sigma_b^2 d^2 / K^2]^{1/2} - \sigma_b^2, \quad (9)$$

678 where σ_{obs}^2 and σ_b^2 are the observation and background error variances, and
 679 $d = \mathbf{y} - \mathbf{H}\mathbf{w}^b$, is the background innovation (Sakov and Sandery, 2017).
 680 Using this feature, the assumed variance of observation errors are smoothly
 681 increased so that the corresponding increments do not exceed the specified
 682 number of standard deviations of the ensemble.

683 Another parameter used to tune the data assimilation performance of
 684 EnKF-C is the “R-factor”. The R-factor scales the assumed observation
 685 error variances, affecting the ratio of the background error variance (the en-
 686 semble variance) to observation error variances, and adjusting the impact of
 687 observations on the calculated corrections to the ocean state (Sakov, 2014).
 688 The R-factor is typically increased for applications using larger localisation
 689 radii that include more observations for each calculation. This increases the
 690 relative assumed observation error variance, again, to avoid over-fitting. In
 691 practice, using an R-factor of 4, effectively reduces the amplitude of the en-
 692 semble variance by 4, which is the same as reducing the effective background
 693 error (ensemble spread) by $\sqrt{4}$.

694 The data assimilation parameters are different for the two data assimila-
 695 tion steps (Table 5). The values used for BRAN2020 are those found to give
 696 good data assimilation performance in Chamberlain et al. (2020). Larger
 697 ensembles are feasible for the coarse-resolution data assimilation step, as the
 698 ocean state vectors and ensemble members are much smaller and relatively
 699 inexpensive computationally compared to the high-resolution data assimila-

700 tion step. We found that increasing the number of the *in situ* observations in
701 the coarse data assimilation step and increasing the localisation radius and
702 observation window, improved the efficacy of the course-resolution assimila-
703 tion step to eliminate the sub-surface biases. For this step, the R-factor is
704 increased to moderate the impact of these observations.

705 **References**

- 706 Bi, D., Marsland, S.J., Uotila, P., O’Farrell, S., Fiedler, R., Sullivan, A.,
707 Griffies, S.M., Zhou, X., Hirst, A.C., 2013. ACCESS OM, the ocean-sea
708 ice core of the ACCESS coupled model. Australian Meteorological and
709 Oceanographic Journal 63, 213–232.
- 710 Brassington, G., Freeman, J., Huang, X., Pugh, T., Oke, P., Sandery, P.,
711 Taylor, A., Andreu-Burillo, I., Schiller, A., Griffin, D., et al., 2012. Ocean
712 model, analysis and prediction system: version 2. Journal of Research and
713 Practice in Information Technology 39, 151–164.
- 714 Cabanes, C., Grouazel, A., von Schuckmann, K., Hamon, M., Turpin,
715 V., Coatanoan, C., Paris, F., Guinehut, S., Boone, C., Ferry, N.,
716 de Boyer Montégut, C., Carval, T., Reverdin, G., Pouliquen, S., Le Traon,
717 P.Y., 2013. The cora dataset: validation and diagnostics of in-situ
718 ocean temperature and salinity measurements. Ocean Science 9, 1–18.
719 doi:10.5194/os-9-1-2013.
- 720 Chamberlain, M.A., Oke, P.R., Brassington, G.B., Sandery, P., Divakaran,
721 P., Fiedler, R.A.S., 2020. Multiscale data assimilation in the Bluelink
722 ocean reanalysis (BRAN). Ocean Modelling .
- 723 Chiswell, S.M., Bostock, H.C., Sutton, P.J., Williams, M.J., 2015. Phys-
724 ical oceanography of the deep seas around New Zealand: a review.
725 New Zealand Journal of Marine and Freshwater Research 49, 286–317.
726 doi:10.1080/00288330.2014.992918.

- 727 Coleman, M.A., Roughan, M., Macdonald, H.S., Connell, S.D., Gillan-
728 ders, B.M., Kelaher, B.P., Steinberg, P.D., 2011. Variation in
729 the strength of continental boundary currents determines continent-
730 wide connectivity in kelp. *Journal of Ecology* 99, 1026–1032.
731 doi:<https://doi.org/10.1111/j.1365-2745.2011.01822.x>.
- 732 Delworth, T.L., Broccoli, A.J., Rosati, A., Stouffer, R.J., Balaji, V., Beesley,
733 J.A., Cooke, W.F., Dixon, K.W., Dunne, J., Dunne, K.A., Durachta,
734 J.W., Findell, K.L., Ginoux, P., Gnanadesikan, A., Gordon, C.T., Griffies,
735 S.M., Gudgel, R., Harrison, M.J., Held, I.M., Hemler, R.S., Horowitz,
736 L.W., Klein, S.A., Knutson, T.R., Kushner, P.J., Langenhorst, A.R., Lee,
737 H.C., Lin, S.J., Lu, J., Malyshev, S.L., Milly, P.C.D., Ramaswamy, V.,
738 Russell, J., Schwarzkopf, M.D., Shevliakova, E., Sirutis, J.J., Spelman,
739 M.J., Stern, W.F., Winton, M., Wittenberg, A.T., Wyman, B., Zeng, F.,
740 Zhang, R., 2006. GFDL’s CM2 Global Coupled Climate Models. Part I:
741 Formulation and Simulation Characteristics. *Journal of Climate* 19, 643.
742 doi:[10.1175/JCLI3629.1](https://doi.org/10.1175/JCLI3629.1).
- 743 Embury, O., Bulgin, C., Mittaz, J., 2019a. ESA Sea Surface Tem-
744 perature Climate Change Initiative (SST_cci): Advanced Very High
745 Resolution Radiometer (AVHRR) Level 2 Preprocessed (L2P) Climate
746 Data Record, version 2.1. Centre for Environmental Data Analysis
747 doi:[10.5285/373638ed9c434e78b521cbe01ace5ef7](https://doi.org/10.5285/373638ed9c434e78b521cbe01ace5ef7).
- 748 Embury, O., Bulgin, C., Mittaz, J., 2019b. ESA Sea Surface Tem-
749 perature Climate Change Initiative (SST_cci): Along-Track Scan-
750 ning Radiometer (ATSR) Level 2 Preprocessed (L2P) Climate Data

Record, version 2.1. Centre for Environmental Data Analysis
doi:10.5285/916b93aaf1474ce793171a33ca4c5026.

Feng, M., Zhang, X., Oke, P., Monselesan, D., Chamberlain, M., Matear, R.,
Schiller, A., 2016. Invigorating ocean boundary current systems around
Australia during 1979–2014: As simulated in a near-global eddy-resolving
ocean model. *Journal of Geophysical Research: Oceans* 121, 3395–3408.

Gentemann, C.L., Meissner, T., Wentz, F.J., 2010. Accuracy of satellite sea
surface temperatures at 7 and 11 ghz. *IEEE Transactions on Geoscience
and Remote Sensing* 48, 1009–1018. doi:10.1109/TGRS.2009.2030322.

Good, S., Embury, O., Bulgin, C., Mittaz, J., 2019. ESA Sea Surface
Temperature Climate Change Initiative (SST_cci): Level 4 Analysis Cli-
mate Data Record, version 2.1 . Centre for Environmental Data Analysis
doi:10.5285/62c0f97b1eac4e0197a674870afe1ee6.

Griffies, S., 2012. Elements of the modular ocean model (mom): 2012 release
with updates. GFDL Ocean Group Tech. Rep. 7, NOAA 632.

Griffies, S.M., 2012. Elements of MOM5, GFDL Ocean Group Technical
Report No. 7. NOAA/Geophysical Fluid Dynamics Laboratory , Code
and documentation available online at www.gfdl.noaa.gov.

Hartog, J.R., Hobday, A.J., Matear, R., Feng, M., 2011. Habitat overlap
between southern bluefin tuna and yellowfin tuna in the east coast longline
fishery - implications for present and future spatial management. *Deep
Sea Research Part II: Topical Studies in Oceanography* 58, 746 – 752.
doi:<https://doi.org/10.1016/j.dsr2.2010.06.005>.

- 774 Henley, B.J., Gergis, J., Karoly, D.J., Power, S., Kennedy, J., Folland, C.K.,
775 2015. A Tripole Index for the Interdecadal Pacific Oscillation. *Climate*
776 *Dynamics* 45, 3077–3090. doi:10.1007/s00382-015-2525-1.
- 777 Kiss, A.E., Hogg, A.M., Hannah, N., Boeira Dias, F., Brassington, G.B.,
778 Chamberlain, M.A., Chapman, C., Dobrohotoff, P., Domingues, C.M., Du-
779 ran, E.R., England, M.H., Fiedler, R., Griffies, S.M., Heerdegen, A., Heil,
780 P., Holmes, R.M., Klocker, A., Marsland, S.J., Morrison, A.K., Munroe,
781 J., Oke, P.R., Nikurashin, M., Pilo, G.S., Richet, O., Savita, A., Spence,
782 P., Stewart, K.D., Ward, M.L., Wu, F., Zhang, X., 2019. Access-om2:
783 A global ocean-sea ice model at three resolutions. *Geoscientific Model*
784 *Development Discussions* 2019, 1–58. doi:10.5194/gmd-2019-106.
- 785 Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H.,
786 Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., et al., 2015. The JRA-
787 55 reanalysis: General specifications and basic characteristics. *Journal of*
788 *the Meteorological Society of Japan* 93, 5–48.
- 789 Locarnini, R., Mishonov, A., Antonov, J., Boyer, T., Garcia, H., Baranova,
790 O., Zweng, M., Paver, C., Reagan, J., Johnson, D., Hamilton, M., Seidov,
791 D., 2013. *World Ocean Atlas 2013, volume 1: Temperature*, in: Levitus, S.,
792 Mishonov, A. (Eds.), NOAA Atlas NESDIS 74. U.S. Government Printing
793 Office, Washington, D.C., p. 40 pp.
- 794 Merchant, C.J., Embury, O., Bulgin, C.E., Block, T., Corlett, G.K., Fiedler,
795 E., Good, S.A., Mittaz, J., Rayner, N.A., Berry, D., et al., 2019. Satellite-
796 based time-series of sea-surface temperature since 1981 for climate appli-

797 cations. Scientific data 6, 1–18. doi:[http://doi.org/10.1038/s41597-019-](http://doi.org/10.1038/s41597-019-0236-x)
798 0236-x.

799 Minnett, P.J., Kilpatrick, K.A., Podestá, G.P., Evans, R.H., Szczodrak,
800 M.D., Izaguirre, M.A., Williams, E.J., Walsh, S., Reynolds, R.M., Bai-
801 ley, S.W., Armstrong, E.M., Vazquez-Cuervo, J., 2020. Skin sea-surface
802 temperature from viirs on suomi-npp—nasa continuity retrievals. Remote
803 Sensing 12. doi:10.3390/rs12203369.

804 Naval Oceanographic Office, 2014a. METOP-A AVHRR GAC L2P swath
805 SST data set. Ver. 1.0. PO.DAAC, CA, USA , Dataset accessed 2017–01–
806 01 to 2019–12–31 at <https://doi.org/10.5067/GHMTG-2PN02>.

807 Naval Oceanographic Office, 2014b. METOP-B AVHRR GAC L2P swath
808 SST data set. Ver. 1.0. PO.DAAC, CA, USA , Dataset accessed 2017–01–
809 01 to 2019–12–31 at <https://doi.org/10.5067/GHMTB-2PN02>.

810 Naval Oceanographic Office, 2014c. N-18 AVHRR GAC L2P swath SST
811 data set. Ver. 1.0. PO.DAAC, CA, USA , Dataset accessed 2017–01–01 to
812 2018–05–14 at <https://doi.org/10.5067/GH18G-2PN02>.

813 Naval Oceanographic Office, 2014d. N-19 AVHRR GAC L2P swath SST
814 data set. Ver. 1.0. PO.DAAC, CA, USA , Dataset accessed 2017–01–01 to
815 2018–05–17 at <https://doi.org/10.5067/GH19G-2PN02>.

816 NOAA Office of Satellite and Product Operations, 2019. Sea Surface tem-
817 perature retrievals produced by NOAA/NESDIS/OSPO office from VIIRS
818 sensor. Ver. 2.61 . PO.DAAC, CA, USA , Dataset accessed 2012–01–01 to
819 2019–12–31 at <https://doi.org/10.5067/GHVRS-2PO61>.

820 Oke, P.R., Brassington, G.B., Griffin, D.A., Schiller, A., 2008. The Bluelink
821 Ocean Data Assimilation System (BODAS). *Ocean Modelling* 21, 46–70.
822 doi:10.1016/j.ocemod.2007.11.002.

823 Oke, P.R., Griffin, D.A., 2011. The cold-core eddy and strong upwelling
824 off the coast of New South Wales in early 2007. *Deep Sea Research* 58,
825 574–591. doi:10.1016/j.dsr2.2010.06.006.

826 Oke, P.R., Griffin, D.A., Rykova, T., de Oliveira, H.B., 2018. Ocean
827 circulation in the great australian bight in an eddy-resolving ocean re-
828 analysis: The eddy field, seasonal and interannual variability. *Deep*
829 *Sea Research Part II: Topical Studies in Oceanography* 157-158, 11–26.
830 doi:doi.org/10.1016/j.dsr2.2018.09.012.

831 Oke, P.R., Griffin, D.A., Schiller, A., Matear, R.J., Fiedler, R., Mansbridge,
832 J.V., Lenton, A., Cahill, M., Chamberlain, M.A., Ridgway, K., 2013a.
833 Evaluation of a near-global eddy-resolving ocean model. *Geoscientific*
834 *model development* 6, 591–615. doi:doi:10.5194/gmd-6-591-2013.

835 Oke, P.R., Roughan, M., Cetina-Heredia, P., Pilo, G.S., Ridgway, K.R.,
836 Rykova, T., Archer, M.R., Coleman, R.C., Kerry, C.G., Rocha, C., Schaef-
837 fer, A., Vitarelli, E., 2019. Revisiting the circulation of the East Australian
838 Current: Its path, separation, and eddy field. *Progress in Oceanography*
839 176, 102139. doi:https://doi.org/10.1016/j.pocean.2019.102139.

840 Oke, P.R., Sakov, P., 2008. Representation error of oceanic observations for
841 data assimilation. *Journal of Atmospheric and Oceanic Technology* 25,
842 1004–1017.

843 Oke, P.R., Sakov, P., Cahill, M.L., Dunn, J.R., Fiedler, R., Griffin, D.A.,
844 Mansbridge, J.V., Ridgway, K.R., Schiller, A., 2013b. Towards a dynam-
845 ically balanced eddy-resolving ocean reanalysis: BRAN3. *Ocean Modelling*
846 67, 52–70. doi:10.1016/j.ocemod.2013.03.008.

847 Oke, P.R., Sakov, P., Corney, S.P., 2007. Impacts of localisation in the EnKF
848 and enoi: experiments with a small model. *Ocean Dynamics* 57, 32–45.
849 doi:10.1007/s10236-006-0088-8.

850 Oke, P.R., Sakov, P., Schulz, E., 2009. A comparison of shelf
851 observation platforms for assimilation in an eddy-resolving ocean
852 model. *Dynamics of Atmospheres and Oceans* 48, 121 – 142.
853 doi:10.1016/j.dynatmoce.2009.04.002.

854 Oke, P.R., Schiller, A., 2007. Impact of Argo, SST, and altimeter data on an
855 eddy-resolving ocean reanalysis. *Geophysical Research Letters* 34, L19601.
856 doi:10.1029/2007GL031549.

857 Oke, P.R., Schiller, A., Griffin, D.A., Brassington, G.B., 2005. Ensemble data
858 assimilation for an eddy-resolving ocean model of the Australian region.
859 *Quarterly Journal of the Royal Meteorological Society* 131, 3301–3311.
860 doi:10.1256/qj.05.95.

861 O’Carroll, A.G., Armstrong, E.M., Beggs, H.M., Bouali, M., Casey, K.S.,
862 Corlett, G.K., Dash, P., Donlon, C.J., Gentemann, C.L., Høyer, J.L.,
863 Ignatov, A., Kabobah, K., Kachi, M., Kurihara, Y., Karagali, I., Ma-
864 turi, E., Merchant, C.J., Marullo, S., Minnett, P.J., Pennybacker,
865 M., Ramakrishnan, B., Ramsankaran, R., Santoleri, R., Sunder, S.,

866 Saux Picart, S., Vázquez-Cuervo, J., Wimmer, W., 2019. Observational
867 needs of sea surface temperature. *Frontiers in Marine Science* 6, 420.
868 doi:10.3389/fmars.2019.00420.

869 Petrenko, B., Ignatov, A., Kihai, Y., Stroup, J., Dash, P., 2014.
870 Evaluation and selection of sst regression algorithms for jpss vi-
871 irs. *Journal of Geophysical Research: Atmospheres* 119, 4580–4599.
872 doi:https://doi.org/10.1002/2013JD020637.

873 Pilo, G.S., Oke, P.R., Rykova, T., Coleman, R., Ridgway, K., 2015. Do East
874 Australian Current anticyclonic eddies leave the Tasman Sea? *Journal of*
875 *Geophysical Research: Oceans* 120, 8099–8114.

876 Rayner, N., Parker, D., Folland, C., Horton, E., Alexander, L., Rowell, D.,
877 2003. The global sea-ice and sea surface temperature (HadISST) data sets.
878 *J. Geophys. Res* 108, 2–22.

879 Rodi, W., 1987. Examples of calculation methods for flow and mixing in
880 stratified fluids. *Journal of Geophysical Research: Oceans* 92, 5305–5328.

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

885 Sakov, P., 2014. EnKF-C user guide. arXiv preprint arXiv:1410.1233 , 0–46.

886 Sakov, P., Sandery, P., 2017. An adaptive quality control procedure for data
887 assimilation. *Tellus A* 69. doi:10.1080/16000870.2017.1318031.

- 888 Scharroo, R., Leuliette, E., Lillibridge, J., Byrne, D., Naeije, M., Mitchum,
889 G., 2013. RADS: Consistent Multi-Mission Products, in: Ouwehand, L.
890 (Ed.), 20 Years of Progress in Radar Altimetry, p. 69.
- 891 Schiller, A., Oke, P.R., 2015. Dynamics of ocean surface mixed layer vari-
892 ability in the Indian Ocean. *Journal of Geophysical Research: Oceans* 120,
893 4162–4186. doi:10.1002/2014JC010538.
- 894 Schiller, A., Oke, P.R., Brassington, G.B., Entel, M., Fiedler, R., Griffin,
895 D.A., Mansbridge, J.V., 2008. Eddy-resolving ocean circulation in the
896 Asian-Australian region inferred from an ocean reanalysis effort. *Progress*
897 *in Oceanography* 76, 334–365. doi:10.1016/j.pocean.2008.01.003.
- 898 Schiller, A., Ridgway, K.R., Steinberg, C.R., Oke, P.R., 2009. Dynamics
899 of three anomalous SST events in the Coral Sea. *Geophysical Research*
900 *Letters* 36, L06606. doi:10.1029/2008GL036997.
- 901 Tsujino, H., Urakawa, S., Nakano, H., Small, R.J., Kim, W.M., Yeager,
902 S.G., Danabasoglu, G., Suzuki, T., Bamber, J.L., Bentsen, M., Boen-
903 ing, C., Bozec, A., Chassignet, E., Curchitser, E., Dias, F.B., Du-
904 rack, P.J., Griffies, S.M., Harada, Y., Ilicak, M., Josey, S., Kobayashi,
905 C., Kobayashi, S., Komuro, Y., Large, W.G., Le Sommer, J., Mars-
906 land, S., Masina, S., Scheinert, M., Tomita, H., Valdivieso, M., Ya-
907 mazaki, D., 2020. input4mips.cmip6.omip.mri.mri-jra55-do-1-5-0. version
908 20200916. URL: <https://doi.org/10.22033/ESGF/input4MIPs.15017>,
909 doi:10.22033/ESGF/input4MIPs.15017.
- 910 Wong, A.P., Wijffels, S.E., Riser, S.C., Pouliquen, S., Hosoda, S., Roemmich,

911 D., Gilson, J., Johnson, G.C., Martini, K., Murphy, D.J., et al., 2020. Argo
912 data 1999–2019: Two million temperature-salinity profiles and subsurface
913 velocity observations from a global array of profiling floats. *Frontiers in*
914 *Marine Science* .

915 Zweng, M.M., Reagan, J.R., Antonov, J.I., Locarnini, R.A., Mishonov, A.V.,
916 Boyer, T.P., Garcia, H.E., Baranova, O.K., Paver, C.R., Johnson, D.R.,
917 Seidov, D., Biddle, M., 2013. *World Ocean Atlas 2013, volume 2: Salin-*
918 *ity*, in: Levitus, S., Mishonov, A. (Eds.), *NOAA Atlas NESDIS 74*. U.S.
919 Government Printing Office, Washington, D.C., p. 39 pp.

Table 1: Summary of observations and errors used when assimilated into BRAN. Brackets contain typical values for the assumed observation error standard deviations for each observation classes that read error provided with data. TEM/SAL includes Argo, TEM2/SAL2 includes XBT, marine mammals etc; (tab:obs)

| Observation Type and Platform | Assumed Standard Deviation of Observation Errors |
|----------------------------------|---|
| SST (AVHRR) | 0.1°C |
| SST (AMSRE) | 0.5°C |
| SST (AMSR2) | 0.45°C |
| SST (ATSR) | 0.1°C |
| SST (VIIRS) | 0.15°C |
| SLA | 0.05 m |
| SLA (Geosat) | 0.07 m |
| TEM | 0.5°C |
| TEM2 | 1°C |
| SAL | 0.075 psu |
| SAL2 | 0.15 psu |

Table 2: Decade averages of mean absolute values of analysis innovations for BRAN2020 and BRAN2016. These statistics provide an indication of how closely BRAN analyses “fit” the assimilated observations. (tab:analysis)

| | BRAN2020 | | | BRAN2016 | | |
|-------------------|----------|-------|-------|----------|-------|-------|
| | 1990s | 2000s | 2010s | 1990s | 2000s | 2010s |
| SST | 0.19 | 0.17 | 0.13 | 0.23 | 0.22 | 0.16 |
| SLA | 3.0 | 3.2 | 2.9 | 3.3 | 3.4 | 3.1 |
| Temperature (all) | 0.44 | 0.35 | 0.28 | 0.50 | 0.46 | 0.39 |
| (0-50 m) | 0.50 | 0.39 | 0.29 | 0.47 | 0.35 | 0.29 |
| (50-500 m) | 0.43 | 0.37 | 0.32 | 0.52 | 0.53 | 0.45 |
| (500+ m) | 0.26 | 0.25 | 0.19 | 0.40 | 0.37 | 0.33 |
| Salinity (all) | 0.085 | 0.060 | 0.046 | 0.161 | 0.090 | 0.071 |
| (0-50 m) | 0.122 | 0.086 | 0.059 | 0.198 | 0.103 | 0.072 |
| (50-500 m) | 0.069 | 0.053 | 0.047 | 0.132 | 0.092 | 0.075 |
| (500+ m) | 0.044 | 0.041 | 0.035 | 0.074 | 0.063 | 0.061 |

Table 3: Decade averages of mean absolute values of background innovations for BRAN2020 and BRAN2016. These statistics provide an indication of how well the model integrates BRAN analyses forward in time.(tab:forecast)

| | BRAN2020 | | | BRAN2016 | | |
|-------------------|----------|-------|-------|----------|-------|-------|
| | 1990s | 2000s | 2010s | 1990s | 2000s | 2010s |
| SST | 0.38 | 0.36 | 0.31 | 0.46 | 0.39 | 0.32 |
| SLA | 5.3 | 5.4 | 5.1 | 5.8 | 5.6 | 5.2 |
| Temperature (all) | 0.68 | 0.56 | 0.45 | 0.81 | 0.66 | 0.56 |
| (0-50 m) | 0.73 | 0.57 | 0.42 | 0.75 | 0.50 | 0.43 |
| (50-500 m) | 0.72 | 0.65 | 0.57 | 0.87 | 0.79 | 0.69 |
| (500+ m) | 0.33 | 0.30 | 0.23 | 0.48 | 0.42 | 0.37 |
| Salinity (all) | 0.148 | 0.105 | 0.082 | 0.273 | 0.131 | 0.102 |
| (0-50 m) | 0.229 | 0.175 | 0.129 | 0.349 | 0.181 | 0.134 |
| (50-500 m) | 0.118 | 0.094 | 0.085 | 0.207 | 0.132 | 0.110 |
| (500+ m) | 0.051 | 0.046 | 0.038 | 0.084 | 0.068 | 0.064 |

Table 4: Decade averages of differences between the mean absolute values of background and analysis innovations for BRAN2020 and BRAN2016. These statistics provide an indication of how the errors grow between each analysis cycle. Fast error growth may indicate that analyses are over-fit, or that the observing system is too sparse to adequately constrain the reanalysis. (tab:growth)

| | BRAN2020 | | | BRAN2016 | | |
|-------------------|----------|-------|-------|----------|-------|-------|
| | 1990s | 2000s | 2010s | 1990s | 2000s | 2010s |
| SST | 0.19 | 0.19 | 0.18 | 0.23 | 0.17 | 0.16 |
| SLA | 2.3 | 2.2 | 2.2 | 2.5 | 2.2 | 2.1 |
| Temperature (all) | 0.24 | 0.21 | 0.17 | 0.31 | 0.20 | 0.17 |
| (0-50 m) | 0.22 | 0.18 | 0.13 | 0.28 | 0.15 | 0.14 |
| (50-500 m) | 0.29 | 0.28 | 0.25 | 0.35 | 0.26 | 0.24 |
| (500+ m) | 0.07 | 0.05 | 0.04 | 0.08 | 0.05 | 0.04 |
| Salinity (all) | 0.063 | 0.046 | 0.036 | 0.113 | 0.041 | 0.031 |
| (0-50 m) | 0.107 | 0.089 | 0.071 | 0.152 | 0.078 | 0.062 |
| (50-500 m) | 0.048 | 0.041 | 0.038 | 0.076 | 0.040 | 0.035 |
| (500+ m) | 0.007 | 0.005 | 0.003 | 0.011 | 0.005 | 0.004 |

Table 5: data assimilation parameter values of experiments presented here at each resolution of multiscale data assimilation system for the two experiments presented. (* - note, in practice, the observations windows were 1 day longer from the start of BRAN2020 to 2007.) (tab:da))

| | Coarse data assimilation | High-res. data assimilation |
|------------------------------|-----------------------------------|---|
| Cycle length (days) | 3 | 3 |
| Ensemble size | 480 | 144 |
| Ensemble generation | monthly climatological anomaly | 3-day minus 3-month centered differences |
| Loc. radius (km) | 1600 | 250 |
| Stride | 1 | 3 |
| R-factor | 4 | 1 |
| K-factor | 1 | 2 |
| Updated domain (latitude) | 65°S -65°N | 75°S -75°N |
| Argo obs. window (days*) | 10 | 3 |
| Non-Argo obs. window (days*) | 3 | 3 |

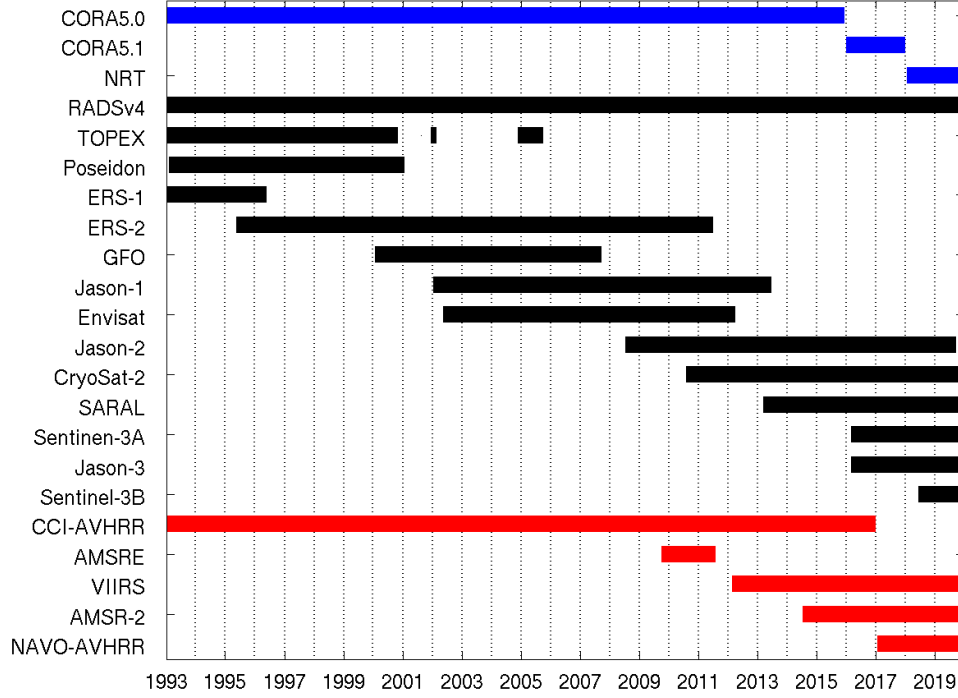


Figure 1: Gantt chart showing when data from different types and sources are assimilated into BRAN2020. Altimeter data are accessed from RADSV4, accessing data from 13 different satellite altimeters (black bars). SST data are accessed from five different sources (red bars), including data from individual satellites (VIIRS, AMSRE, AMSR-2) and from datasets that have combined AVHRR and ATSR data from multiple satellites (CCI-AVHRR, CCI-ASTR and NAVO-AVHRR). In situ data are accessed from three sources - two from Copernicus Marine Environment Monitoring Service (CMEMS; CORIOLIS dataset for Re-Analysis v5 and v5.1; CORA5.0 and CORA5.1) and a near-real-time dataset maintained by the Bureau of Meteorology. (fig:obs)

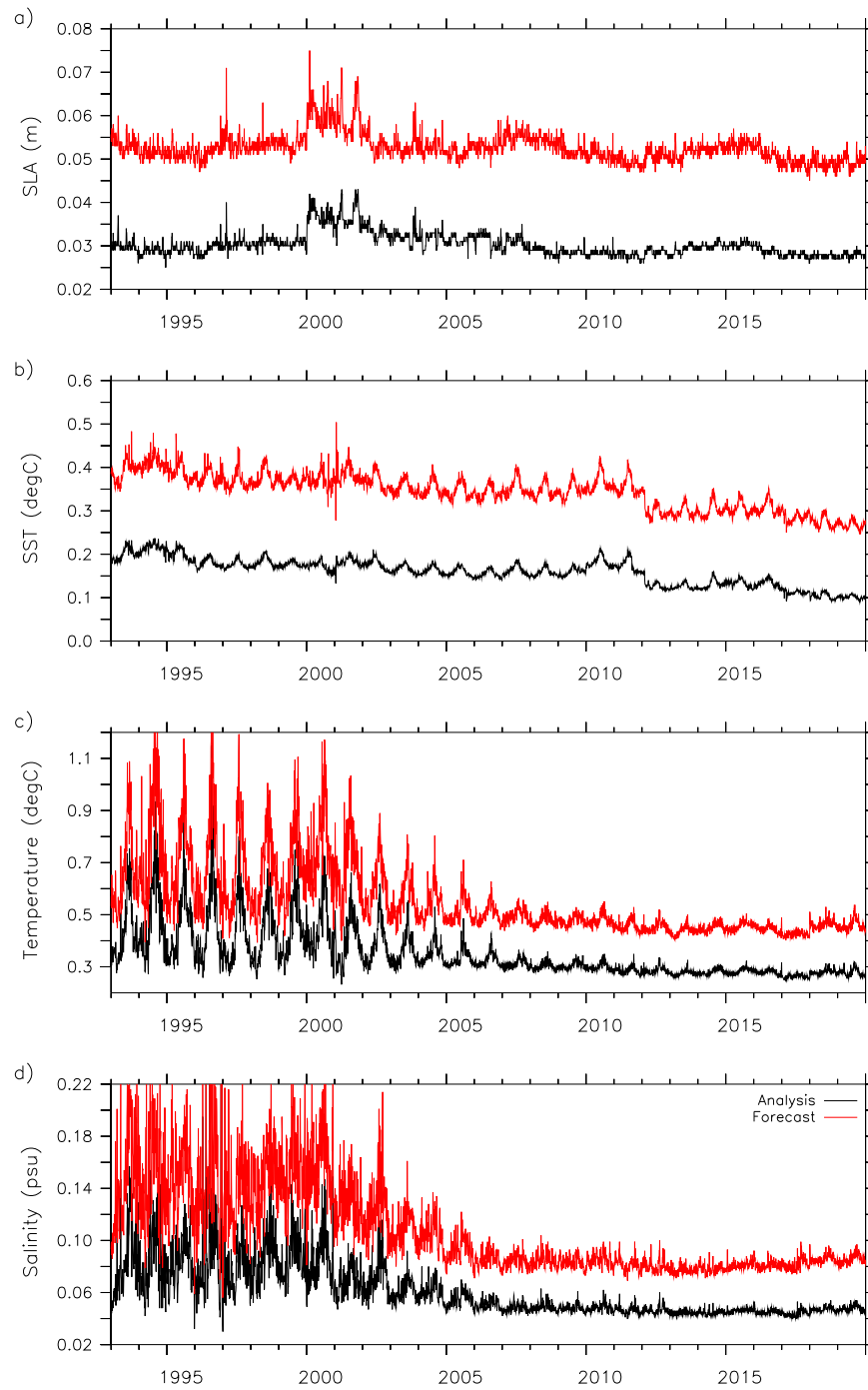


Figure 2: Time-series of global MAD between observations and background fields (red) and analyses (black) for (a) SLA, (b) SST, (c) temperature, and (d) salinity. (fig:da-stats)

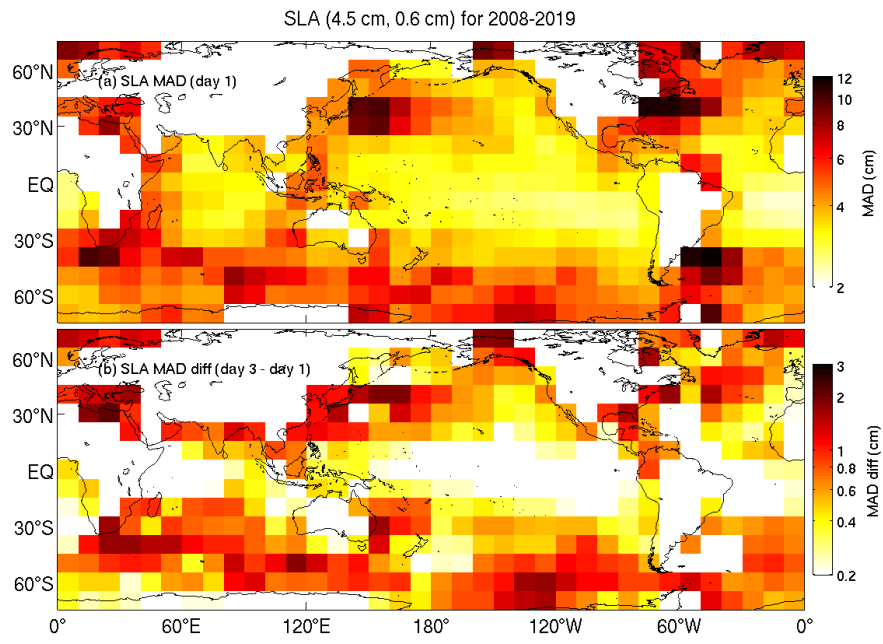


Figure 3: Binned sea level mean absolute deviations, averaged over 2008-2019 from the first day (top) and Day 3 minus Day 1 (bottom). (fig:sla-mad)

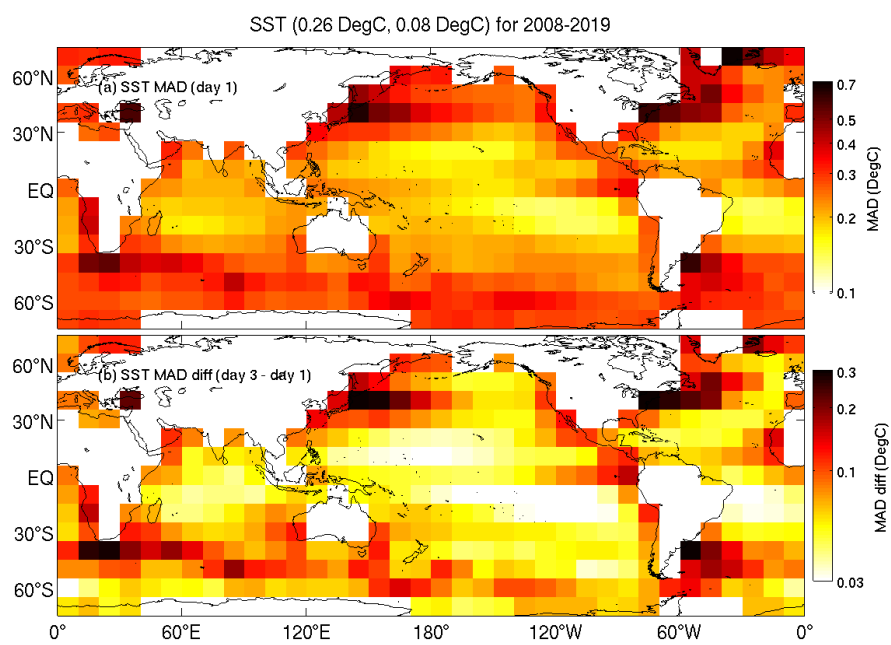


Figure 4: Binned SST mean absolute deviations, averaged over 2008-2019 from the first day (top) and Day 3 minus Day 1 (bottom). (fig:sst-mad)

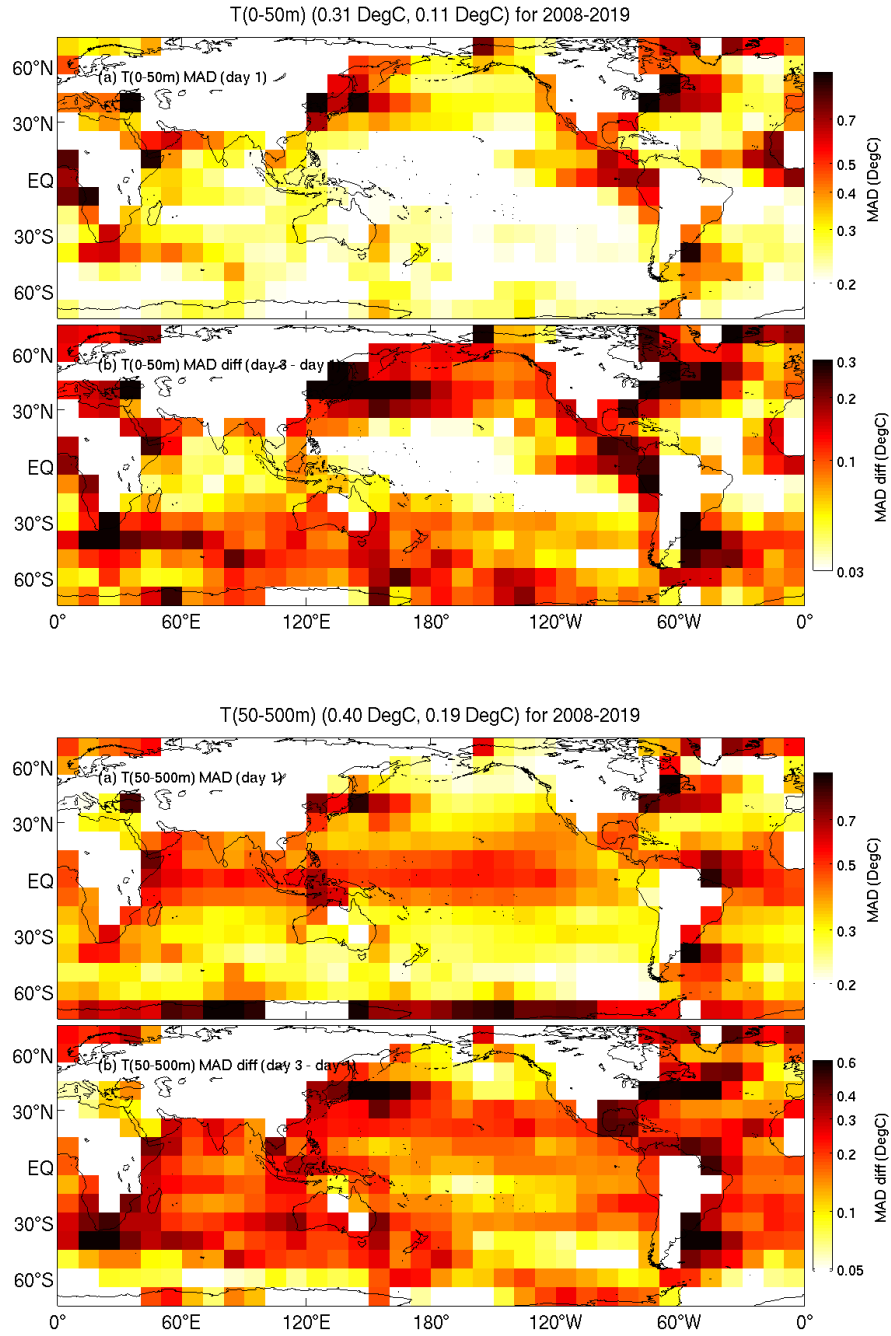


Figure 5: Binned mean absolute deviations of temperatures between 0 and 50 m (top pair) and 50 and 500 m (bottom), averaged over 2008-2019 from the first day (top of pair) and Day 3 minus Day 1 (bottom of pair). (fig:tem-mad)

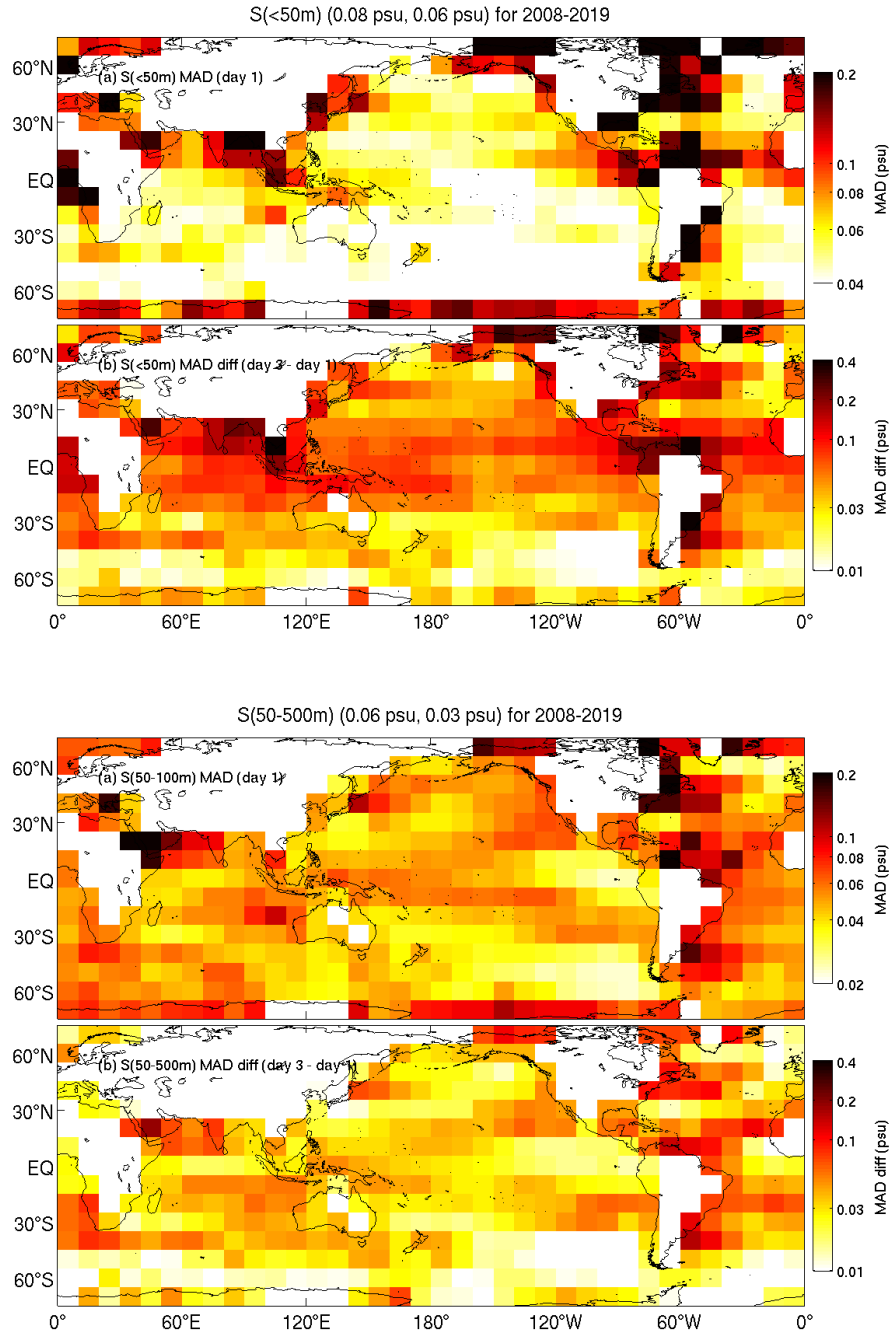


Figure 6: Binned mean absolute deviations of salinity between 0 and 50 m (top pair) and 50 and 500 m (bottom), averaged over 2008-2019 from the first day (top) and Day 3 minus Day 1 (bottom). (fig:sal-mad)

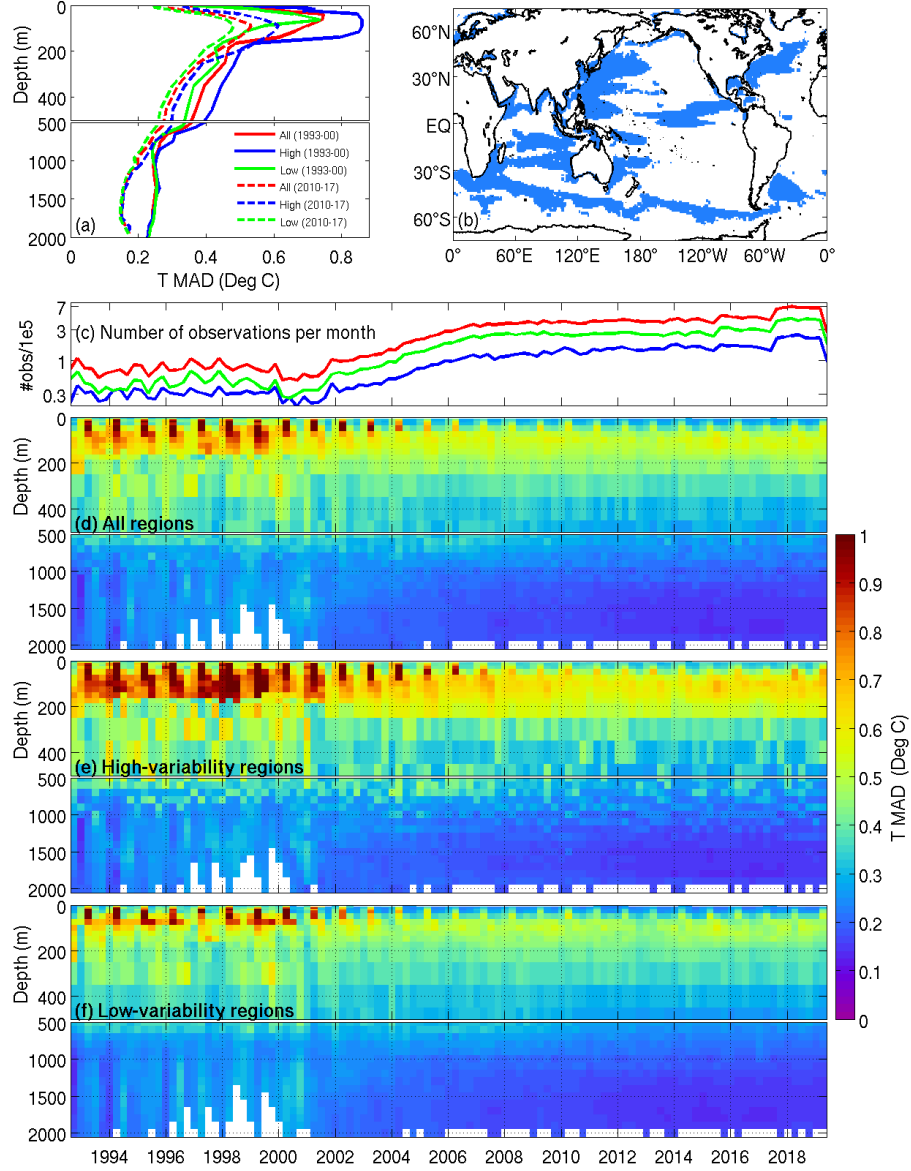


Figure 7: Mean absolute deviations (MAD) of temperature with depth over the course of BRAN2020. Results are shown for regions with high and low variability, as seen in sea level, and combined. MAD profiles in the top left, averaged over times before the full Argo array and after. Time series of the number of observations over BRAN2020 (log scale) is above Hovmöllers of MAD for each region. (fig:Tz-mad)

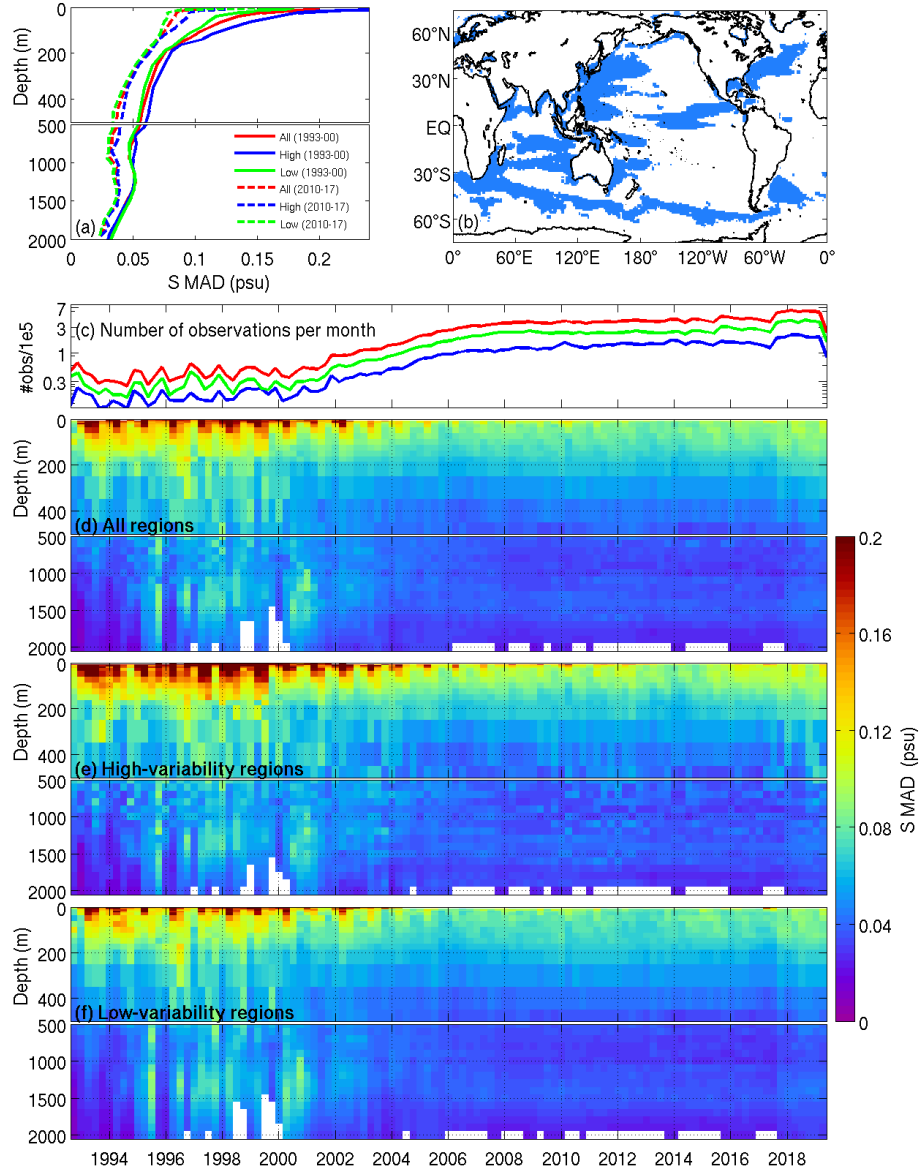


Figure 8: Mean absolute deviations (MAD) of salinity with depth over the course of BRAN2020. Results are shown for regions with high and low variability, as seen in sea level, and combined. MAD profiles in the top left, averaged over times before the full Argo array and after. Time series of the number of observations over BRAN2020 (log scale) is above Hovmöllers of MAD for each region. (fig:Sz-mad)

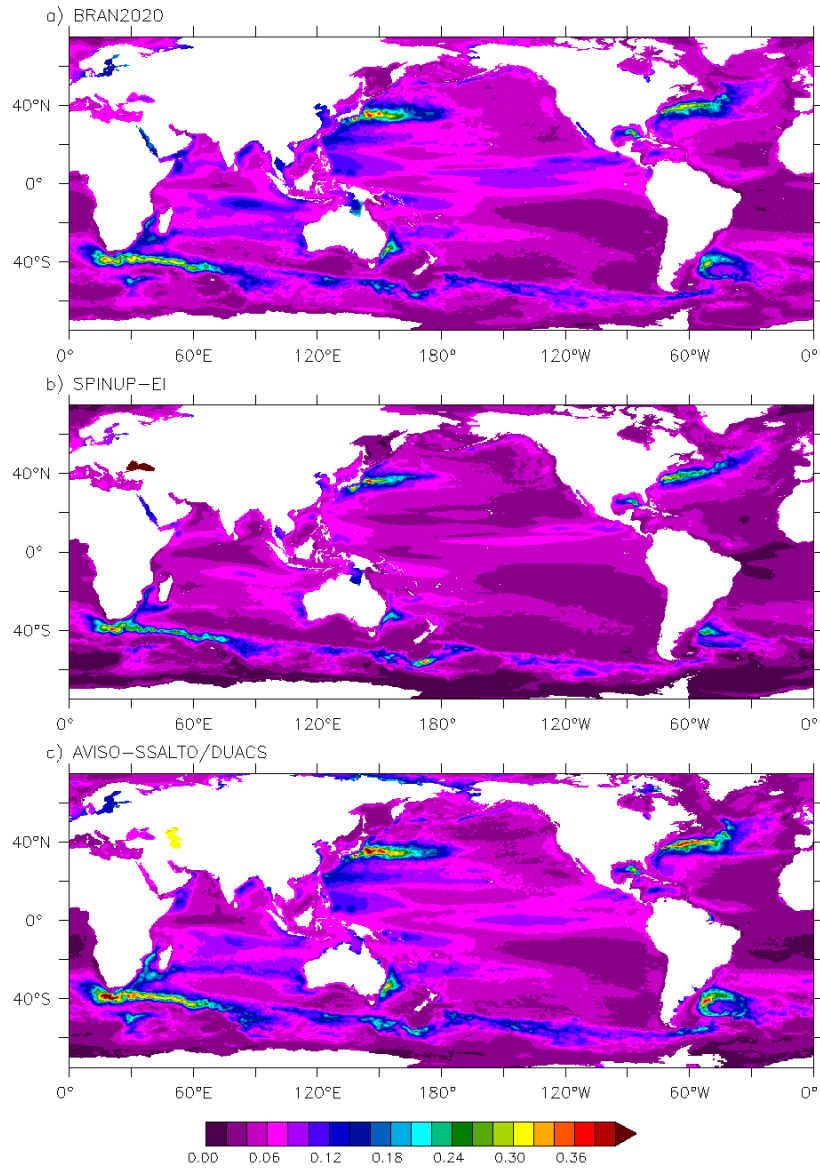


Figure 9: Sea level variability over 1993-2014 (detrended) from BRAN2020 (top), and free-running model (middle) and observations (bottom). (fig:sla-var-global)

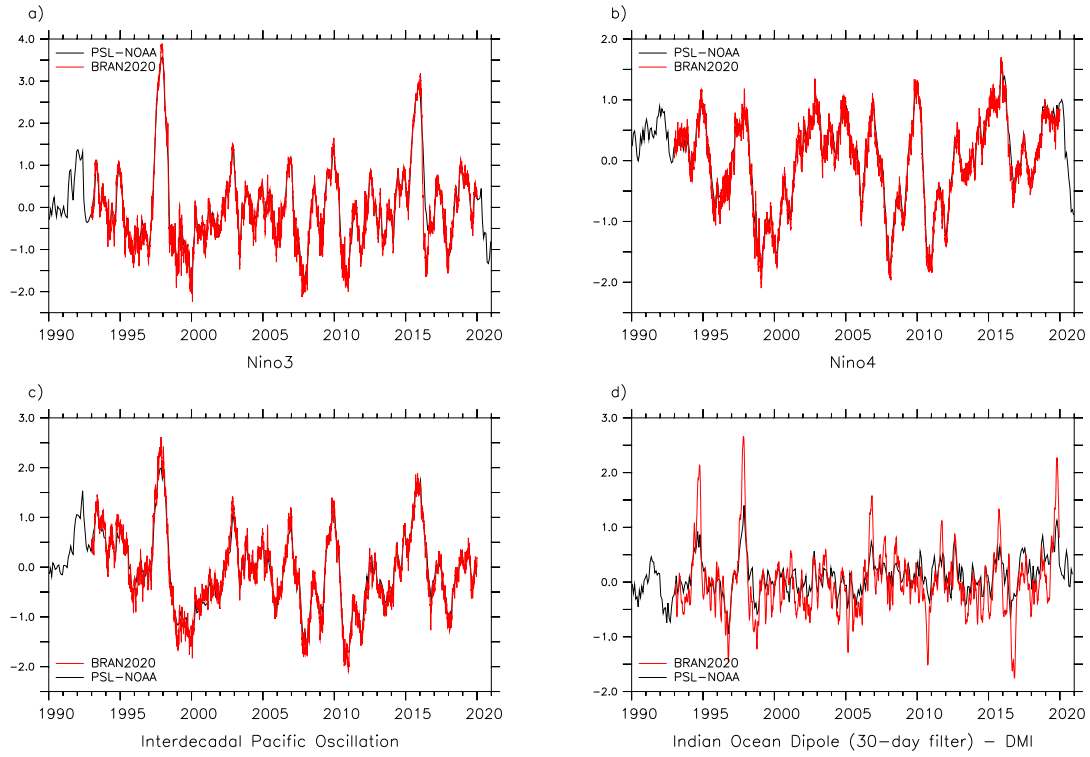


Figure 10: Major climate SST indices a) Niño3 (5°S - 5°N , 210°E - 270°E), b) Niño4 (5°S - 5°N , 160°E - 210°E), c) Interdecadal Pacific Oscillation, and d) the Indian Ocean Dipole (10°S - 10°N , 50°E - 70°E ; 10°S - 0°N , 90°E - 110°E).

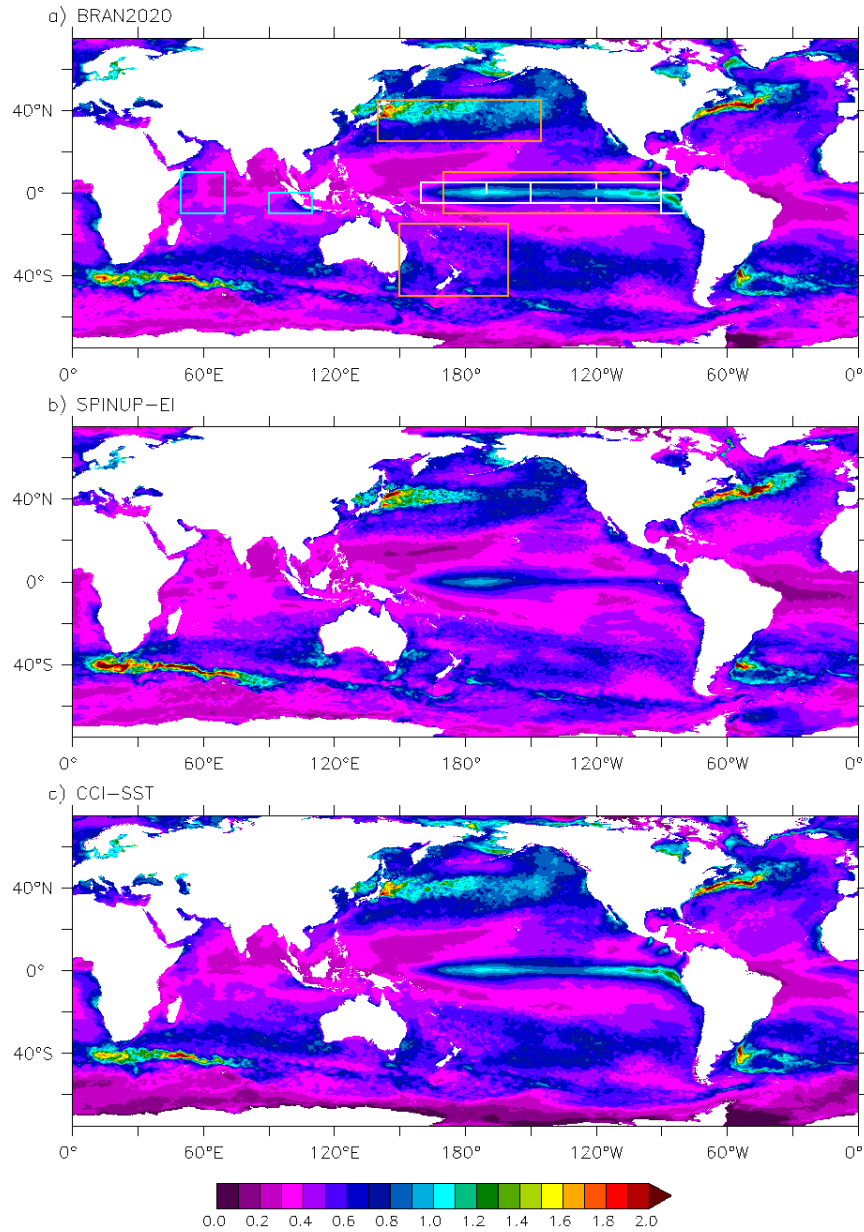


Figure 11: SST variability as standard deviation of climatological anomalies in BRAN2020 (top), the free-running Spinup-EI (middle) and CCI-SST (bottom) over the years 2000-2009. Top panel shows location for climate indices shown in Fig. 10; ENSO SST indices in white, IPO in orange and IOD in blue. (fig:sst-var-map)