

# 1 Comparison of ocean heat content estimated using two eddy- 2 resolving hindcast simulations based on OFES1 and OFES2

3 Fanglou Liao<sup>1,2</sup>, Xiao Hua Wang<sup>2\*</sup>, and Zhiqiang Liu<sup>1,3\*</sup>

4 <sup>1</sup>Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055,  
5 China

6 <sup>2</sup>The Sino-Australian Research Consortium for Coastal Management, School of Science, The University of New South  
7 Wales, Canberra, 2610, Australia

8 <sup>3</sup>Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, 511458, China

9 *Correspondence to:* Zhiqiang Liu (liuzq@sustech.edu.cn) or Xiao Hua Wang (x.h.wang@unsw.edu.au)

10 **Abstract.** In this study, we have compared the ocean heat content (OHC), estimated using two eddy-resolving hindcast  
11 simulations based on Ocean General Circulation Model for the Earth Simulator Version 1 (OFES1) and Version 2  
12 (OFES2). Results from a global objective analysis of subsurface temperature (EN4) were taken as a reference. Both  
13 EN4 and OFES1 suggest that OHC has increased in most regions of the top 2000 m during 1960–2016, which is  
14 mainly associated with the deepening of neutral density surfaces and variations along the neutral density surfaces of  
15 regional importance. Upon comparing the results obtained from the two OFES hindcasts, we found substantial  
16 differences in the temporal and spatial distributions of the OHC, especially in the Atlantic Ocean. A basin-wide heat  
17 budget analysis showed that there was less surface heating for the major basins in the OFES2. The horizontal heat  
18 advection was mostly similar, however, the OFES2 had a significantly stronger meridional heat advection associated  
19 with the Indonesian Throughflow (ITF) above 300 m. Additionally, large discrepancies in the vertical heat advection  
20 were also evinced when the two OFES results were compared, especially at a depth of 300 m in the Indian Ocean. We  
21 inferred that there are large discrepancies in the vertical heat diffusion (cannot be directly evaluated in this study due  
22 to data unavailability), which, along with the different magnitudes of sea surface heat flux and vertical heat advection,  
23 were the major factors responsible for the examined differences in OHC. This work suggests that OFES1 provides a  
24 reasonable multi-decadal estimate of global and basin-integrated warming trends above 700 m, except for the top 300  
25 m for the Pacific Ocean and between 300–700 m for the Indian Ocean. Although the estimates of the global OHC  
26 during 1960–2016 are consistent with observations between 700–2000 m, caution is warranted while examining the  
27 basin-wide multi-decadal OHC variations using OFES1. The seemingly suboptimal OHC estimate based on OFES2,  
28 suggests that any conclusions on long-term climate variations derived from OFES2 might suffer from large drifts,  
29 necessitating audits.

## 30 1 Introduction

31 The global oceans store more than 90% of extra heat that has been added to the Earth since the 1950s, generating a  
32 significant OHC increase (Levitus et al., 2012; IPCC 2013). Therefore, OHC forms an important indicator of climate  
33 change, and it helps estimate the Earth's energy imbalance (Palmer et al., 2011; Von Schuckmann et al., 2016).  
34 Although natural factors such as El Niño–Southern Oscillation (ENSO) and volcanic eruptions can modulate the OHC

35 (Balmaseda et al., 2013; Church et al., 2005), the recent warming trend has been largely induced by the accumulation  
36 of greenhouse gas in the atmosphere (Abraham et al., 2013; Gleckler et al., 2012; Pierce et al., 2006).

37 The OHC increase, being a major concern for both oceanography and climate communities, has attracted a great deal  
38 of attention. Although direct observational records represent the most reliable data for determining the oceanic thermal  
39 state, the available observations are not dense enough in both the temporal and spatial domains, especially for the deep  
40 and abyssal oceans. The number of observations has greatly improved since the launch of a global array of profiling  
41 floats, the Argo, in the 2000s. However, the spatial resolution of the Argo program (i.e., approximately 300 km) is not  
42 high enough to capture mesoscale structures (Sasaki et al., 2020, hereafter **S2020**). There are several approaches for  
43 filling the temporal and spatial gaps in global temperature measurements, which can be used to produce gridded  
44 temperature products for estimating the OHC. Typical approaches include an objective analysis (Good et al., 2013) of  
45 observational data and an ensemble optimal interpolation with a dynamic ensemble (EnOI-DE (Cheng and Zhu, 2016)).  
46 In addition, ocean general circulation models (OGCMs) provide the temperature fields by solving primitive equations  
47 of fluid motion and state. When constrained by observations, a numerical ocean modelling becomes the ocean  
48 reanalysis, which generally lacks dynamical-consistence (the resulting fields satisfy the underlying fluid dynamics and  
49 thermodynamics equations), unless the adjoint method was adopted to use information contained in observations.  
50 Although ocean reanalysis has been widely constructed, unconstrained OGCMs are still an important tool for climate  
51 prediction, for instance, the Coupled Model Intercomparison Project (CMIP). How multi-scale dynamical processes  
52 are represented in these unconstrained models and their implementation of external forcing significantly impact their  
53 OHC estimates.

54 The Ocean General Circulation Model for the Earth Simulator (OFES (Masumoto et al., 2004; Sasaki et al., 2004)),  
55 developed by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) and other institutes, is a well-  
56 known eddy-resolving OGCM, and the hindcast simulation of the OFES Version 1 (OFES1) has been widely used  
57 (Chen et al., 2013; Dong et al., 2011; Du et al., 2005; Wang et al., 2013). The hindcast simulation based on the OFES  
58 Version 2 (OFES2) has now been released with certain improvements over the OFES1 (**S2020**). For example, in a  
59 comparison to the OFES1, the authors found a smaller bias in the global sea surface temperature (SST), sea surface  
60 salinity (SSS), and the water-mass properties of the Indonesian and Arabian seas in the OFES2. To our knowledge,  
61 however, a comparison of the multi-decadal OHC at a basin or global scale between OFES1 and OFES2 is lacking.  
62 As this high-resolution quasi-global hindcast simulation is expected to be widely used in oceanography and climate  
63 communities for examining the state of the ocean in the near future, it is necessary to compare the OHC estimated  
64 using the two OFES as an indicator of the potential improvements in OFES2 over OFES1. Such a study is also  
65 expected to provide insights on the adaptability of the two simulations for OHC-related studies. The finding that  
66 subsurface oceanic fields could be notably different when estimated based on the results of two OFES runs with  
67 different atmospheric forcing, despite their similar results in the near-surface region (Kutsuwada et al., 2019), forms  
68 an added motivation to conduct the envisioned study.

69 The aim of this study is twofold: (1) estimate the OHC in the global ocean and in each major basin using OFES1 and  
70 OFES2, with a primary focus on differences between the two hindcasts; and (2) understand the causes of the  
71 differences between the two hindcasts. To this end, we used the potential temperature  $\theta$  to calculate and compared the

72 OHC from 1960 to 2016 for both the global ocean and the major basins, i.e., the Pacific Ocean, the Atlantic Ocean,  
73 and the Indian Ocean between 64° S and 64° N.

74 In Section 2, we provide a brief description of the data and methods used in this study. In Section 3, we describe and  
75 discuss the differences in OHC in both the temporal and spatial domains. A tentative analysis of the possible causes  
76 of these differences was also conducted. Section 4 summarizes the principal points and the possible extensions  
77 involving factors that were not examined here due to data unavailability, although such factors could be important.  
78 Accordingly, we have added the future scope of this study to improve the associated work.

## 79 **2 Data and Methods**

### 80 **2.1 Data**

81 The potential temperature  $\theta$  from both OFES1 and OFES2 were used to calculate the global and basin OHCs. This  
82 allowed us to compare the OHC estimated from OFES1 and OFES2, along with the estimates from the observation-  
83 based EN4. Although results from EN4 cannot be considered to represent the actual oceanic state, it has been widely  
84 used in OHC-related studies (Allison et al., 2019; Carton et al., 2019; Häkkinen et al., 2016; Trenberth et al., 2016;  
85 Wang et al., 2018). A brief description of the three datasets is given below; readers are referred to Sasaki et al. (2004),  
86 Sasaki et al. (2020), and Good et al. (2013) for a more detailed description.

87 The OFES1 has a horizontal spatial resolution of 0.1° with 54 vertical levels and a maximum depth of 6065 m (Sasaki  
88 et al., 2004). Such a high lateral resolution enables it to resolve mesoscale processes. Following a 50-year  
89 climatological simulation, the hindcast simulation of the OFES1 was integrated forward, with the publicly available  
90 data from 1950 to 2017. The multi-decadal integration made it possible to analyze oceanic fields at temporal scales  
91 from intra-seasonal to multi-decadal. Unlike most other datasets used for the estimation of the OHC, OFES1 is  
92 unconstrained by any observations. Therefore, it can be used to demonstrate the adaptability of high-resolution  
93 numerical modeling without data assimilation in climate studies.

94 OFES2 has the same horizontal spatial resolution of 0.1°. Vertically, there are 105 levels with a maximum depth of  
95 7500 m. OFES1 uses National Centers for Environmental Prediction (NCEP) reanalysis (2.5° × 2.5° (Kalnay et al.,  
96 1996)) for atmospheric forcing on an everyday basis, whereas OFES2 obtains atmospheric forcing from the JRA55-  
97 do Version 08 (55 km × 55 km (Tsujino et al., 2018)) with a temporal resolution of 3 hours. Both the temporal and  
98 spatial resolutions of atmospheric forcing have increased significantly in the OFES2. The OFES2 also incorporates  
99 river runoff and sea-ice models, but polar areas are not included.

100 In the horizontal direction, both OFES1 and OFES2 use a biharmonic mixing scheme to suppress the computational  
101 noise (S2020). The horizontal diffusivity coefficient is equal to  $-9 \times 10^9$  m<sup>4</sup>/s at the equator (S2020) and varies  
102 proportionally with the cube of the cosine of the latitude (personal communication with Hide Sasaki). The OFES2  
103 uses a mixed-layer vertical mixing scheme (Noh and Kim 1999) with parametrization of tidal energy dissipation (Jayne  
104 and St. Laurent 2001; St. Laurent et al., 2002), whereas OFES1 uses the K-profile parameterization (KPP) scheme  
105 (Large et al., 1994). Taking the temperature and salinity of January 1, 1958 from OFES1 as the initial conditions,  
106 OFES2 was integrated forward, with the publically available data from 1958 to 2016. To reduce the computation time,  
107 we subsampled the OFES1 and OFES2 data at every five grid points in the horizontal direction.

108 To evaluate the OHC from the two OFES data, we used EN4 from the UK Meteorological Office Hadley Centre as  
 109 a reference. Note that we used the EN4.2.1 as the EN4 version, with bias-corrected following Levitus et al. (2009).  
 110 The EN4 data can be considered as objective analysis data that is based on observations (Good et al., 2013), with a  
 111 horizontal resolution of  $1^\circ$  and 42 vertical levels down to 5350 m. The EN4 assimilates data mostly from the World  
 112 Ocean Database (WOD) and the Coriolis dataset for ReAnalysis (CORA). Preprocessing and quality checks were  
 113 conducted before the observational data were used to construct this objective analysis product.

114 Although we used the results from EN4 as a reference for evaluating the performance of OFES in simulating the 57-  
 115 year thermal state of the ocean, EN4 cannot be considered to represent the actual ocean state. The main reason is that  
 116 the measurements used to construct the EN4 datasets are sparse and inhomogeneous in both temporal and spatial  
 117 domains, and are insufficient to resolve mesoscale or even sub-mesoscale motions. There are more observations in  
 118 the northern hemisphere compared to the southern hemisphere, and there is also a seasonal bias in the observational  
 119 data density (Abraham et al. 2013; Smith et al. 2015). A larger density of data was generated only after the World  
 120 Ocean Circulation Experiment (WOCE) was conducted in the 1990s and following the launch of the Argo profiling  
 121 floats in the 2000s. Table 1 summarizes the three datasets.

122 **Table 1.** A summary of the OFES1, OFES2 and EN4. The symbol / means “not applicable”.

	OFES1	OFES2	EN4
Model	MOM3	MOM3	/
Horizontal coverage	$75^\circ \text{ S} - 75^\circ \text{ N}$	$76^\circ \text{ S} - 76^\circ \text{ N}$	$83^\circ \text{ S} - 89^\circ \text{ N}$
Horizontal grids	$3600 \times 1500$	$3600 \times 1520$	$360 \times 173$
Vertical levels	54	105	42
Maximum depth	6065 m	7500 m	5350 m
Atmospheric forcing	NCEP	JRA55-do Ver.08	/
Data assimilated	/	/	WOD, CORA
Time span	1950 – 2017	1958 – 2016	1900 – 2021

123  
 124 We considered water from the sea surface to approximately 2000 m, and divided it into three layers: upper (0–300  
 125 m), middle (300–700 m), and lower (700–2000 m). The ocean above 2000 m is often divided into two layers, 0–700  
 126 m and 700–2000 m (or even one: 0–2000 m) (Allison et al., 2019; Häkkinen et al., 2015; Häkkinen et al., 2016; Levitus  
 127 et al., 2012; Zanna et al., 2019). However, our analysis shows that it is necessary to divide it into three layers to reach  
 128 the objective of this study. Similar vertical division can also be seen in Liang et al. (2021).

129 The reasons for ignoring water below 2000 m were mainly fourfold. First, the simulated behavior of the deep and  
 130 abyssal oceans depends on the spin-up of the numerical simulation, which is mostly incomplete (Wunsch 2011), at  
 131 least in the first decade. Second, the observational data used in EN4 are largely confined to the top 2000 m, and some  
 132 available measurements do not even go down to this depth (personal communication with the EN4 UK Meteorological  
 133 Office Hadley Centre). The number of data is significantly lesser in the deep and abyssal oceans. Third, the EN4 data  
 134 that we used here was bias-corrected following Levitus et al. (2009), in which only the ocean above 700 m was  
 135 considered. For instance, the Expendable Bathythermograph (XBT) profiles below 700 m were corrected using the  
 136 correction values provided for 700 m (personal communication from the UK Meteorology Office Hadley Centre).  
 137 Finally, the maximum depth of OFES2 and EN4 differs by more than 2000 m. It was felt that the full-depth OHC,  
 138 estimated using the three datasets, is not highly comparable. However, this does not imply that we can ignore the

139 contribution of the deep ocean; it can play an essential role in regulating the global-ocean thermal state (Desbruyères  
 140 et al. 2016; Desbruyères et al. 2017; Palmer et al. 2011). It is expected that a significantly better understanding of the  
 141 deep and abyssal ocean states will be gained with the implementation of the Deep Argo program, which is partially  
 142 validated by Johnson et al. (2019).

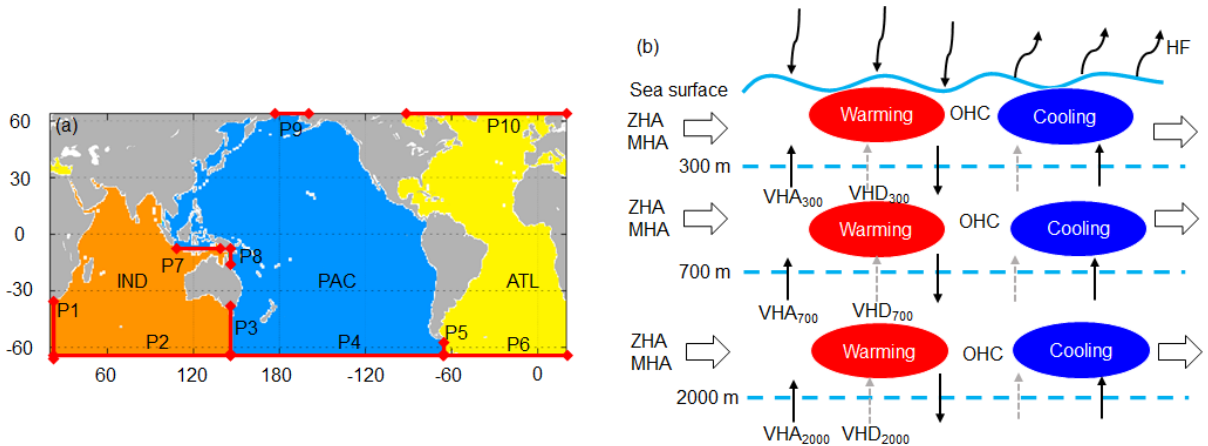
## 143 2.2 Methods

144 We compared the three datasets for the period 1960–2016. In this paper, the OHC represent the OHC anomalies  
 145 relative to the OHC estimates of 1960. At each grid point, the OHC is expressed as follows:

$$146 \text{OHC} = \rho \delta v C_p (\theta - \theta_{1960}) = \rho \delta v C_p \Delta \theta, \quad (1)$$

147 where  $\rho$  is the seawater density ( $\text{kg/m}^3$ ),  $\delta v$  is the grid volume ( $\text{m}^3$ ),  $C_p$  is the specific heat of seawater at constant  
 148 pressure ( $\text{J/kg/}^\circ\text{C}$ ),  $\theta$  is the yearly potential temperature ( $^\circ\text{C}$ ), and  $\theta_{1960}$  is the average potential temperature during  
 149 1960. The total OHC in the upper ocean layer (above 300 m) is the integral of Eq. (1) from 0 to 300 m. Similar  
 150 procedures were applied to the other two layers (300–700 m and 700–2000 m). A value of  $4.1 \times 10^6 \text{ J/m}^3/^\circ\text{C}$  was used  
 151 for the product of  $\rho$  and  $C_p$  (Palmer et al., 2011).

152 OHCs of both global and individual basins were calculated for comparison. Fig. 1 shows the domains of the Pacific,  
 153 Atlantic, and Indian Oceans between  $64^\circ \text{ S}$  and  $64^\circ \text{ N}$ , including their respective marginal seas. Our definition of the  
 154 marginal seas of each major basin may be inconsistent with those of other studies. The major water passages  
 155 connecting the different basins are denoted by red lines in Fig. 1a. Fig. 1b is the schematic of primary processes that  
 156 determine the OHC of an ocean basin.



157  
 158 **Figure 1. (Left)** Domains of the major basins between  $64^\circ \text{ S}$  and  $64^\circ \text{ N}$  and **(right)** a schematic of the primary  
 159 processes controlling the thermal state of an ocean. **(a)** The PAC stands for the Pacific Ocean, the ATL for the Atlantic  
 160 Ocean and the IND for the Indian Ocean. The basin domain is extracted using the gcmfaces package (Forget et al.,  
 161 2015) and then interpolated to the corresponding grid of each product. Grey indicates the land. The red solid lines  
 162 with diamond arrow stand for the water passages connecting different basins. We label it with the capital letter P  
 163 (abbreviation for passage) and a serial number. The horizontal and vertical axis are longitude and latitude, respectively.  
 164 **(b)** We use a light blue solid curve to represent the free sea surface and three dashed lines to indicate the 300 m, 700  
 165 m and 2000 m depth. The curve arrow represents the net heat flux (HF) through the ocean surface. The black hollow  
 166 arrow shows the zonal (ZHA) or meridional (MHA) heat advection. The black thin arrow represents the vertical heat  
 167 advection (VHA) and the grey dash arrow stands for the vertical heat diffusion (VHD). The red ellipse illustrates  
 168 warming water and the blue ellipse cooling water. P1: ( $20^\circ \text{ E}$ ,  $64^\circ \text{ S} - 34.5^\circ \text{ S}$ ); P2: ( $20^\circ \text{ E} - 146.5^\circ \text{ E}$ ,  $64^\circ \text{ S}$ ); P3:

169 (147° E, 64° S – 36.5° S); P4: (147° E – 65.5° W, 64° S); P5: (67° W, 64° S – 55° S); P6: (65° W – 19.5° E, 64° S);  
170 P7: (118.5° E – 138.5° E, 8.5° S); P8: (142° E, 12.5° S – 8° S); P9: (172.5° W – 166.5° W, 64° N); P10: (88° W –  
171 19.5° E, 64° N).  
172

173 In addition,  $\Delta\theta$  at a fixed depth is decomposed into a heave (HV) component (the second term of Eq. (2)) and a  
174 spiciness (SP) component (the third term of Eq. (2)) (Bindoff and McDougall, 1994). HV-related warming or cooling  
175 is manifested as a vertical displacement of the neutral density surfaces (a continuous analog of discretely referenced  
176 potential density surfaces (Jackett and McDougall, 1997)). In general, both the dynamic changes and the change in  
177 the renewal rates of water-masses can induce vertical displacement, generating HV-related warming or cooling  
178 (Bindoff and McDougall, 1994). SP represents warming or cooling as a result of density compensation in  $\theta$  and salinity  
179 ( $S$ ) along the neutral density surfaces. Decomposition of  $\Delta\theta$  helps to better understand the contributions of different  
180 water-masses to generating OHC. The formula for decomposing the potential temperature is given as follows:

$$181 \quad d\theta/dt|_z = - \overbrace{dz/dt|_n}^{\text{HV}} d\theta/dz + \overbrace{d\theta/dt|_n}^{\text{SP}}, \quad (2)$$

182 where  $t$  is the time (year),  $z$  is the depth (m), and  $|_n$  means along the neutral density surface.

183 The program developed by Jackett and McDougall (1997) was used to calculate the neutral densities, HV, and SP.  
184 This code is based on the United Nations Educational, Scientific and Cultural Organization (UNESCO), 1983 for the  
185 computation of fundamental properties of seawater ([http://www.teos-](http://www.teos-10.org/preteos10_software/neutral_density.html)  
186 [10.org/preteos10\\_software/neutral\\_density.html](http://www.teos-10.org/preteos10_software/neutral_density.html)). We used its MATLAB version for our calculations. The main inputs  
187 for this program were  $\theta$  and  $S$ . The code limits the latitude to be between 80° S and 64° N, but we further confined  
188 our investigation domain to 64° from the equator, which avoids comparisons in sea-ice-impacted areas, given that  
189 only OFES2 includes a sea-ice model.

190 To analyze the origin of the differences in OHC from thermodynamic and dynamic perspectives, we calculated the  
191 surface heat flux (HF), zonal heat advection (ZHA), meridional heat advection (MHA), and vertical heat advection  
192 (VHA). Owing to a temporary suspension of the OFES2 data by the JAMSTEC, we could not access the vertical  
193 diffusivity data of OFES2 while preparing this manuscript. Note that OFES1 does not provide such data. This  
194 prevented us from directly comparing the estimates of vertical heat diffusion (VHD) based on OFES1 and OFES2.  
195 Alternatively, we calculated the residual of the total OHC and all the other heat inputs (HF, ZHA, MHA, and VHA),  
196 and used the results as a proxy for VHD. As the horizontal heat diffusion was found to be significantly weaker than  
197 that of ZHA and MHA (not shown), we did not include it in the analysis. A schematic of the primary process is shown  
198 in Fig. 1b. Note that the linear trend in the following sections was calculated using multiple linear regression using  
199 least squares at 95% confidence level.

### 200 **3 Results**

201 The principal objective of this study is to compare the results from OFES1 and OFES2, considering EN4 as an  
202 observation-based reference. We attempted to evaluate if there is any significant difference between the results  
203 obtained from OFES2 and those from one or both of the other two datasets, and if any such difference represents a  
204 real phenomenon that is not present in the other two widely used datasets or it is an unwanted property of the newly

205 released OFES2 simulation. In this section, we compare the three sets of results for the global ocean, along with  
 206 individual cases of the Pacific, Atlantic, and Indian Oceans.

207 **3.1 Temporal evolution of the OHC, HV, and SP from 1960 to 2016**

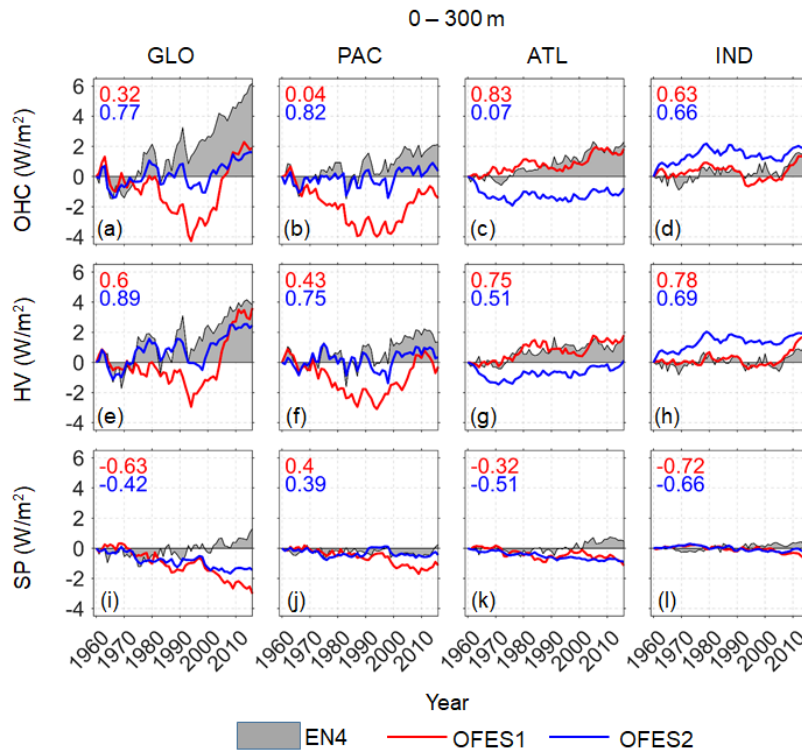
208 **3.1.1 Time series of OHC, HV, and SP**

209 Figures 2–4 illustrate the time series of the total OHC, and its HV and SP components for the upper (0–300 m), middle  
 210 (300–700 m), and lower (700–2000 m) ocean layer, respectively. Note that OHC, HV, and SP were calculated as an  
 211 anomaly relative to the estimates in 1960, and was converted to an equivalent HF applying over the entire surface area  
 212 of the Earth.

213

214 *Upper layer*

215 For the global ocean between 0 and 300 m, all three data indicate cooling from approximately 1963 to 1966 (Fig. 2a),  
 216 which was caused by the volcanic eruption of Mount Agung (Balmaseda et al., 2013). A similar trend of cooling  
 217 during this period is also reported for the upper 700 m (Domingues et al., 2008; Allison et al., 2019) and for both 0–  
 218 700 m and 0–3000 m depth (Achutarao et al., 2007). This short, however, sharp cooling period significantly impacted  
 219 the Pacific Ocean (Fig. 2b). Marked reductions in the OHC associated with strong volcanic eruptions of El Chichón  
 220 in 1982 (a strong El Niño also emerged in 1982–83) and Pinatubo in 1991 were also consistently captured by all three  
 221 data.



222

223 **Figure 2.** Time series of the global and basin-wide OHC (**top**), HV (**middle**) and SP (**bottom**) between 0–300 m  
 224 based on the three datasets. The OHC, HV and SP here are converted to the accumulative heating in  $W/m^2$  applied

225 over the entire surface of Earth. Grey shadow: EN4; red solid line: OFES1; blue solid line: OFES2. Numbers on the  
226 left top corners are the correlation coefficients between the OFES1 (red) or OFES2 (blue) and EN4. The OHC hereafter  
227 is directly calculated from the potential temperature, rather than the sum of the HV and SP.  
228

229 Both EN4 and OFES2, but not OFES1, showed a slowdown in warming in the Pacific Ocean during the 2000s (Fig.  
230 2b). This slowdown of warming in the Pacific corresponds to a sharp warming trend in the upper layer of the Indian  
231 Ocean (Fig. 2d), seen in all the three datasets. This relationship between the Pacific and Indian Oceans could be a  
232 consequence of intensified Indonesian Throughflow (ITF), which increased heat transport from the Pacific to the  
233 Indian Oceans (Lee et al. 2015; Zhang et al. 2018). Note that these two studies considered the top 700 m. However,  
234 the sudden warming of the Indian Ocean was largely confined to the top 300 m, which is indicated by OFES1 and  
235 OFES2 (Fig. 3d). The EN4 showed a clear acceleration of warming trend above 300 m in the global ocean around  
236 2003, which was probably an artifact caused by the transition of the ocean observation network from a ship-based  
237 system to Argo floats (Cheng and Zhu, 2014), although these authors mainly used subsurface temperature data from  
238 the World Ocean Database 2009 (WOD09). Interestingly, a dramatic shift can also be seen in OFES1 (Fig. 2a),  
239 although that OFES1 is not directly constrained by observations. A major difference in this jump between EN4 and  
240 OFES1 is its close association with SP in EN4 (Fig. 2i) compared to HV in OFES1 (Fig. 2e). This spiciness warming  
241 around 2003, derived from EN4, complements the work of Cheng and Zhu (2014).

242 However, several significant differences were observed between the three datasets. Results from EN4 indicated that  
243 the temporal evolution of the warming was approximately linear since ~1970 (Fig. 2a), which was modulated by the  
244 abovementioned climate signals. The OFES1, however, showed that the cooling period persisted almost until the early  
245 1990s, while a stronger linear warming trend appeared afterward (Fig. 2a). This was more than 20 years later than that  
246 indicated by the EN4. In the OFES2, the approximately linear warming trend appeared even later (~2000), the  
247 magnitude of which was approximately the weakest among the three datasets.

248 Compared to OFES1, the temporal profile of the global upper ocean obtained using OFES2 generally agreed better  
249 with that indicated by EN4 (Fig. 2a), which, to some extent, is consistent with the smaller SST bias estimated from  
250 the OFES2 than that from the OFES1 when compared to the World Ocean Atlas 2013 (WOA13) (S2020). However,  
251 the difference between OFES2 and EN4 in magnitude became larger after 1980. This was mainly due to the SP  
252 component (Fig. 2i), with both OFES1 and OFES2 indicating a clear SP cooling episode. This might imply some  
253 discrepancies in the salinity information of these three datasets. In contrast, there was a good agreement between the  
254 HV values of EN4 and OFES2 (Fig. 2e).

255 Clear differences can also be seen for each basin. The OFES1 differed significantly from the other two in the Pacific  
256 Ocean during 1970–1990, with the other two being similar to each other with respect to both HV and SP. In the  
257 Atlantic Ocean, however, the OFES1 agreed quite well with the EN4 in the HV. The two OFES datasets had similar  
258 spiciness in the Atlantic Ocean, but both disagreed with the spiciness of EN4. The HV, estimated using OFES2,  
259 showed poor agreement with both EN4 and OFES1 in the 1960s (Fig. 2g). In the Indian Ocean, OFES1 was much  
260 closer to EN4 than OFES2. The notable deviations of the OFES2 relative to others were mainly generated from the  
261 uniquely strong warming trend in the OFES2 Indian Ocean before ~1980 (Fig. 2d).

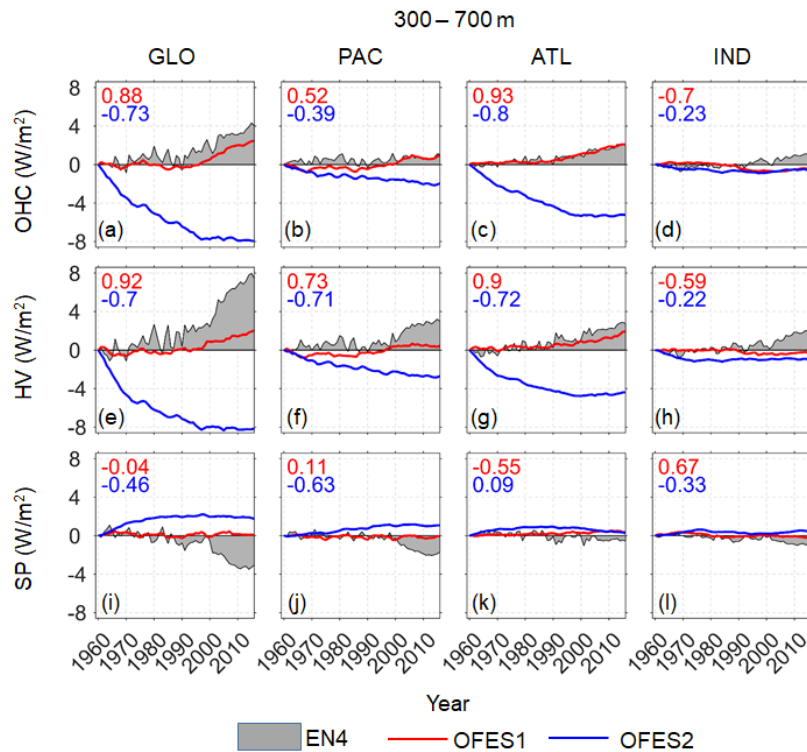


262 A potential issue of the OFES2 is the spin-up, although it was initiated from the calculated temperature and salinity  
 263 fields from OFES1. Without any prior knowledge of the timing of complete spin-up, here we have shown and  
 264 compared the simulated results from 1960, excluding the first two years (1958–1959). It seems that the results obtained  
 265 using OFES2 have a better agreement with EN4 since the 1980s for both Atlantic and Indian Oceans (Fig. 2c, d),  
 266 which is likely to be related to the improvement in spin-up with time. However, in the Pacific Ocean, the OFES2 was  
 267 quite similar to EN4 before 1990, especially its HV component. This, to some extent, might weaken the spin-up  
 268 argument.

269

270 *Middle layer*

271 In the middle ocean layer (300–700 m) (Fig. 3), there were remarkable differences in the OHC and its HV and SP  
 272 components between the OFES2 and the other two datasets, which is most noticeable in the Atlantic Ocean, and lesser  
 273 for the Pacific Ocean; the difference was minor for the Indian Ocean. The OFES2 showed a moderate Pacific cooling  
 274 for almost the entire 57-year period and a strong Atlantic cooling trend until ~2000, with a subsequent hiatus in the  
 275 Atlantic Ocean. The OFES2 indicated that there was a minor cooling in the Indian Ocean during the 1960–70s. In  
 276 OFES2, these uniquely cooling trends were mainly associated with HV because its spiciness was generally more  
 277 positive than as indicated by the other two datasets.



278

279 **Figure 3.** As for Fig.2 but for the middle layer (300–700 m).

280

281 In contrast, both EN4 and OFES1 indicated that the middle layer was relatively stable before the early 1990s (Fig.  
 282 3a). Afterwards, EN4 and OFES1 both showed global ocean and Atlantic Ocean warming (Fig. 3a, c), mostly due to  
 283 an increase in the HV (Fig. 3e, g). Despite such good agreement between EN4 and OFES1, there were notable

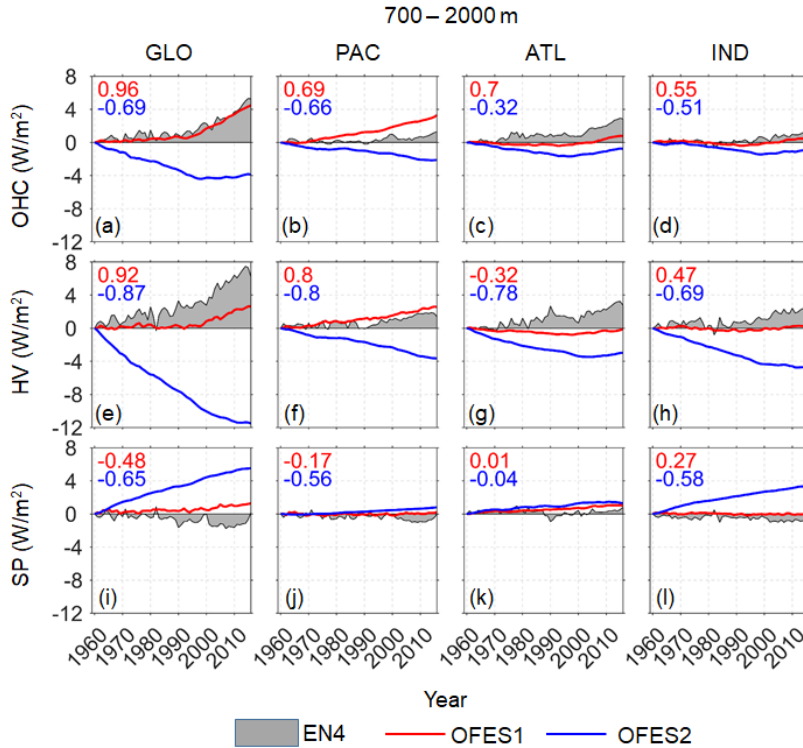
284 differences in their HV and SP components. Compared to the OFES1, there was a stronger positive HV in the EN4  
285 (Fig. 3e–h) and a stronger negative SP in the EN4, particularly after 2000 (Fig. 3i, j). A possible reason for this finding  
286 may be that many more observations have become available since the WOCE was conducted in the late 1990s and  
287 from the Argo since the beginning of the 2000s. This might have led to a systematic trend in the observation-based  
288 dataset EN4. Unlike EN4 and OFES2, the SP variations in OFES1 were almost invisible for almost all the basins. In  
289 addition, the aforementioned significant warming acceleration from the early 2000s to the 2010s in the Indian Ocean  
290 (Fig. 2d) can still be seen in the EN4 (Fig. 3d), however, this was almost invisible in the two OFES datasets.

291 One major cause of the profound differences between OFES2 and the other two datasets may be the spin-up issue.  
292 Indeed, even after 2000, clear differences can be observed in the global ocean. This is expected because the middle  
293 layer takes more time to be completely spun compared to the upper layer. Hence, special caution is required while  
294 investigating the multi-decadal variations or even decadal variations in the recent two decades based on OFES2.

295  
296 *Lower layer*

297 In the lower oceanic layer (700–2000 m) (Fig. 4), the OFES2 was again an outlier among the three datasets. It showed  
298 that the Atlantic and the Indian Oceans experienced cooling from 1960 to the end of the 1990s (Fig. 4c, d), followed  
299 by a slight warming episode. In the Pacific Ocean, however, OFES2 showed cooling over the entire 57-year period  
300 (Fig. 4b). The better agreement between the results from OFES2 and EN4 since the end of the 1990s might be related  
301 to the spin-up issue of the OFES2, at least to some extent. However, the agreement between EN4 and OFES2 was  
302 even better than that in the middle layer (300–700 m), particularly in the Atlantic Ocean. This might weaken the spin-  
303 up argument because it is expected that the middle layer can be more easily spun-up than the lower layer.

304 The variations in OHC determined using OFES1 and EN4 were similar for the global ocean, however, this could be  
305 associated with the cancelation of the substantial differences in the Pacific and Atlantic Oceans (Fig. 4b, c), and in the  
306 HV and SP (Fig. 4e–l). More specifically, there was a larger increase of OHC in the Pacific Ocean, when estimated  
307 using OFES1 than from EN4, however, the latter showed a larger increase of OHC in the Atlantic Ocean. From the  
308 perspective of potential temperature decomposition, EN4 generally showed a stronger increase in HV than OFES1 in  
309 the Atlantic and Indian Oceans (Fig. 4g, h), however, a stronger negative or a weaker positive increase of SP is also  
310 evinced (Fig. 4i–l).



311  
312 **Figure 4.** As for Fig.2 but for the lower layer (700–2000 m).  
313

314 **3.1.2 Temporal evolution of the OHC, HV, and SP trend**

315 Figures 2–4 show the similarities and differences between the three datasets with respect to the time series of OHC,  
316 HV, and SP for the period 1960–2016. In this section, we calculate the linear trend in OHC, HV, and SP over a rolling  
317 window of 10 years for the three datasets following Smith et al. (2015), and the results for the three layers are shown  
318 in Figures 5–7, respectively. Such evaluation has helped us to quantitatively compare the three datasets over each  
319 temporal window.

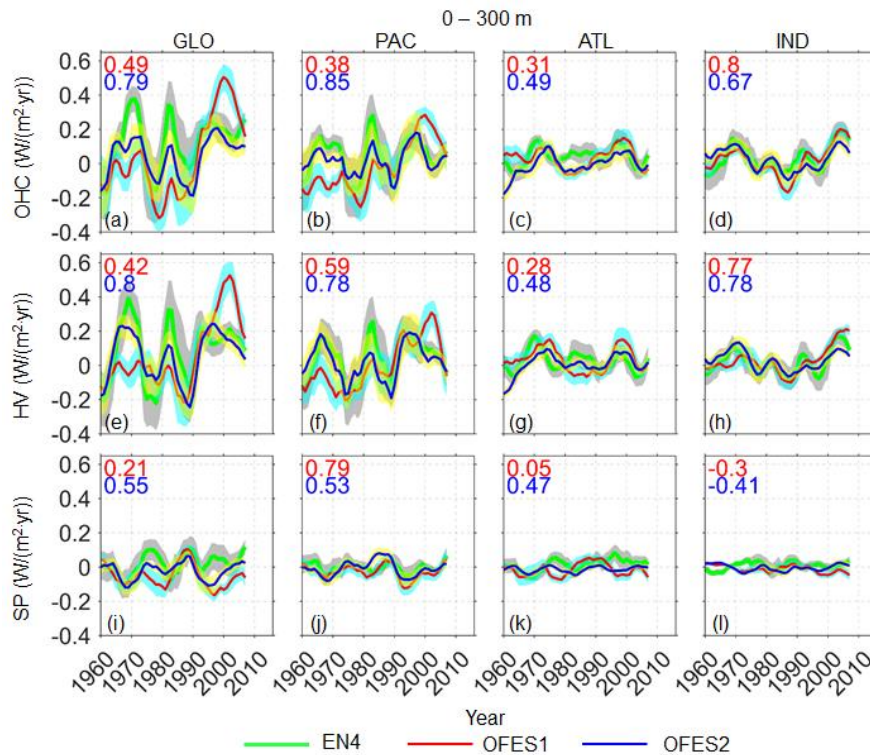
320 *Upper layer*

321 The profile of the 10-year rolling trend of the OHC evaluated based on the three datasets was similar in shape; they  
322 captured most of the peaks and troughs pretty consistently. There was a better agreement among the data for the Indian  
323 Ocean (Fig. 5d) compared to that in the other two basins (Fig. 5b, c), however, notable differences were still observed  
324 in this shallow layer of the Indian Ocean. The rolling trend for the global ocean, estimated from EN4, was mostly  
325 positive, except at the beginning of the 1960s and the end of the 1970s and the 1980s (Fig. 5a). The OFES1 showed a  
326 cooling trend in the global ocean before ~1990; it then indicated a larger warming trend compared to that estimated  
327 from the other two datasets. The OFES2 generally had a better agreement with EN4 for the global ocean, however,  
328 the warming trend was significantly smaller than that estimated using EN4 from the late 1960s to ~1990. Since the  
329 beginning of the 1990s, the disparity in the trend between OFES2 and EN4 was significantly reduced, although the  
330 OFES2 still showed a consistently weaker warming trend. This improved agreement may be attributed to two factors.  
331 First, after running the simulation for approximately 30 years, the OFES2 is expected to have developed better spun-

332 up and, therefore, the associated results were expected to be closer to the actual state. Second, it is also possible that  
 333 the accuracy of the EN4 data increased as more observational data were included, given that oceanographic  
 334 observations have increased significantly since the 1990s (e.g., satellite-based SST measurements and in-situ  
 335 temperature measurements).

336 Among the differences observed between the three datasets, the three extreme trend peaks at approximately 1970,  
 337 1980, and 2000 (Fig. 5a) were particularly prominent, with remarkable differences between OFES and EN4, indicating  
 338 some limitations of unconstrained numerical models in the reproduction of strong climate events. The OFES1 was  
 339 closer to EN4, showing significant warming in the Indian Ocean in the 2000s, whereas OFES2 showed a relatively  
 340 weaker warming trend. The second better agreement between the three datasets was reached for the Atlantic Ocean.

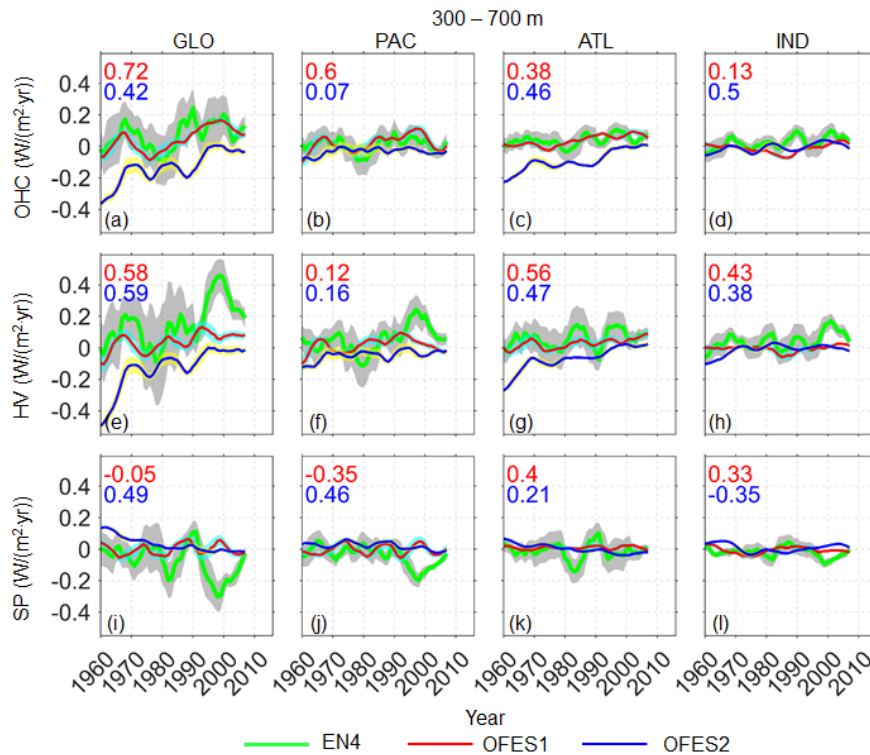
341 It was evinced that HV has dominated the 10-year rolling trend in all basins (Fig. 5e–h), and the major differences  
 342 between the three datasets resulted from the differences in the HV component. In addition, there was an generally out-  
 343 of-phase relationship between the HV and SP trends in the global ocean and the Pacific Ocean. This correspondence  
 344 between the HV and SP is expected for typical stratification in subtropical regions (Häkkinen et al. 2016), with warm  
 345 and salty water overlying cold and fresh water. The OFES1 and OFES2 provided quite similar results for the  
 346 simulation of spiciness, particularly in the individual basins (Fig. 5i–l).



347  
 348 **Figure 5.** Temporal evolution of the 10-year rolling trends in the global and basin OHCs (**top row**), HV (**middle row**)  
 349 and SP (**bottom row**) in the upper layer (0–300 m), based on the three datasets. Numbers in the top left corners are  
 350 the correlation coefficients between the EN4 and the OFES1 (red) or OFES2 (blue). The OHC, HV and SP were  
 351 converted to accumulative heating ( $W/m^2$ ) over the entire surface of the Earth. Thick green line: EN4 (grey shadow:  
 352 95% confidence interval); thin red solid line: OFES1 (cyan shadow: 95% confidence interval); thin blue solid line:  
 353 OFES2 (yellow shadow: 95% confidence interval).  
 354

355 *Middle layer*

356 The variation in the 10-year rolling trend, evaluated based on OFES1 and EN4 datasets, was found to be similar for  
 357 the global (Fig. 6a), Pacific (Fig. 6b), and Atlantic (Fig. 6c) Oceans, however, the latter dataset had a significantly  
 358 larger uncertainty. The OFES2 showed a significantly different and generally cooling trend, especially concentrated  
 359 in the Atlantic Ocean, consistent with Fig. 3. The origin of the notable cooling trend and its weakening with time  
 360 estimated from the OFES2 for the Atlantic Ocean need to be further studied. The cooling trend of the OHC, estimated  
 361 from OFES2, was mostly generated from the HV. In the Pacific Ocean (Fig. 6b), the OFES2 consistently showed a  
 362 weak cooling trend, however, in the middle and late 1960s and after ~1980, both EN4 and OFES1 showed a warming  
 363 trend of similar magnitudes. The results from OFES1 also agreed well with that from the EN4 for the Atlantic Ocean,  
 364 i.e., both indicated a weak warming trend for most of the studied period along with a sporadic cooling trend. However,  
 365 such agreements could represent the compensation results of the significantly different HV and SP components of  
 366 OFES1 and EN4. For example, the EN4 showed a significantly stronger HV warming trend than the OFES1 in the  
 367 Pacific Ocean since the early 1990s, however, in the meantime, the EN4 also indicated a stronger SP cooling trend.  
 368 In the Indian Ocean, EN4 presented a warming trend over much of the 57 years, whereas the two OFES showed weak  
 369 variations and reversals between warming and cooling episodes.



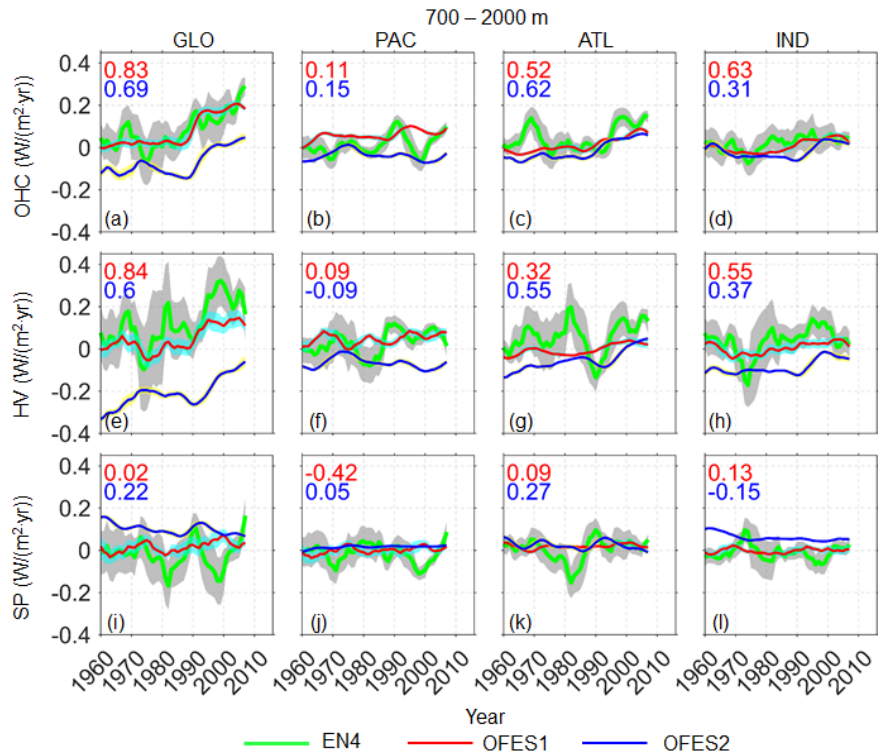
370  
 371 **Figure 6.** As for Fig. 5 but for middle layer (300–700 m).  
 372

373 *Lower layer*

374 As in the middle layer, the OFES2 differed significantly from the other two datasets by displaying a cooling trend in  
 375 the global ocean until approximately 2000 (Fig. 7a). Although OFES2 indicated the appearance of a warming trend in  
 376 the global ocean after ~2000, the intensity was significantly lower than that of EN4 and OFES1. The major differences

377 between the two OFES datasets occurred in the Pacific Ocean (Fig. 7b), and were mostly associated with the HV  
 378 component. Despite the good agreement in the OHC trend between the OFES1 and OFES2 for the Atlantic and Indian  
 379 Oceans (Fig. 7c, d), their HV and SP components were markedly different, especially in the Indian Ocean (Fig. 7h, i).  
 380 The OFES1 and EN4 showed a mostly similar global OHC trend (Fig. 7a); this was because the significant HV and  
 381 SP components canceled each other.

382 To summarize, the OFES2 demonstrated some improvement (better agreement with EN4) over the OFES1 in the  
 383 upper layer (above 300 m), but was more of an outlier below 300 m. It is essential to examine the HV and SP  
 384 components while investigating the OHC trends because different data products might show mostly similar evolution  
 385 of the OHC but substantially different HV and SP.



386  
 387 **Figure 7.** As for Fig. 6 but for the lower layer (700–2000 m).  
 388

389 **3.2 Temporal evolution of the zonal-averaged potential temperature trend**

390 Section 3.1 focused on comparisons of the temporal characteristics of the global and basin-wide OHC, HV, and SP  
 391 estimated from the three datasets. Although both similarities and differences were demonstrated, the comparison in  
 392 the temporal domain lacked spatial information. In this section, we aimed to understand how these similarities and  
 393 differences were distributed in the meridional direction. As a first step, we calculated the 10-year rolling trends in the  
 394 zonal-averaged potential temperature for all three datasets (Figs. 8–10). We also calculated the HV and SP components  
 395 (Supplementary information, Figs. 1–6).

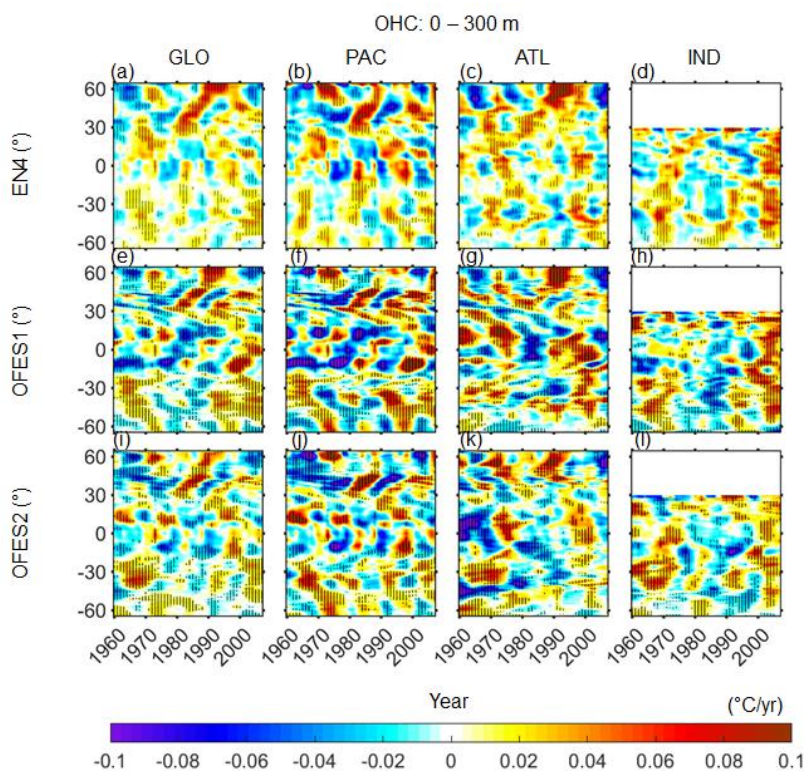
396 The complex patterns shown in Figures 8–10 defy easy interpretation; therefore, we have focused on the large-  
 397 scale patterns of the observed similarities and differences.

398

399 *Upper layer*

400 In general, a reasonable agreement was observed between the three datasets at latitudes of 30–60° N for both Pacific  
401 and Atlantic Oceans (there is no northern high latitude in the Indian Ocean). More specifically, a wave-like cooling  
402 patch propagating from approximately 60° N to 30° N was observed from 1960 to the end of the 1970s in the global  
403 ocean; this propagation was especially evinced in the EN4 and OFES2 data. In addition, there was a northward  
404 propagation of a cooling trend in the 1990s between 30 and 45° N, mainly occurring in the Pacific Ocean. It is  
405 reasonable to attribute these cooling episodes to the volcanic eruptions of Indonesia’s Mount Agung in 1963,  
406 Mexico’s El Chichón in 1982, and the Philippines’ Mount Pinatubo in 1991. The two hindcast simulations were able  
407 to reproduce these climate events.

408 Following these cooling events, there were three subsequent warming trends as the ocean surface temperature  
409 returned to normal after the aerosols released over several years of volcanic eruptions were completely dispersed. Of  
410 these warming trends, the one associated with the El Chichón eruption was the most significant, and there was a clear  
411 northward propagation of this significant warming trend from approximately 30° N to the subpolar areas. Interestingly,  
412 the contributions of SP to this large-scale warming and cooling episodes were comparable to those of the HV  
413 (Supplementary Information, Figs. S1–2), contradicting the general impression that HV is the most dominant  
414 contributor to the potential temperature changes. In fact, the abovementioned propagation of the cooling patch from  
415 approximately 60° N to 30° N during 1960–1970 was, to a larger extent, associated with the SP.



416

417 **Figure 8.** Temporal evolution of 10-year rolling trend of the zonal averaged potential temperature change in the upper  
418 layer of the ocean (0–300 m). **Left to right:** global, Pacific, Atlantic and Indian Ocean. **Top to bottom:** EN4, OFES1

419 and OFES2. Horizontal axis: year; vertical axis: latitude. Stippling indicates the 95% confidence level. The HV and  
420 SP counterparts are in the Supplementary Information, Figs. S1–2.

421  
422 Equatorward of 30°, large differences were observed among the three datasets. Strong cooling was particularly  
423 visible in the OFES1 in the Pacific tropics before around 1990 (Fig. 8f), corresponding to the persistent cooling of the  
424 global ocean and the Pacific Ocean as estimated based on OFEES1 in Fig. 2. The results of OFES2 for the Pacific  
425 Ocean indicated clear differences from the EN4 in the low latitudes before 1980, and then a pattern similar to that of  
426 EN4 was simulated by the OFES2. In the Atlantic tropics, considerable cooling over 1960s was evinced in the OFES2,  
427 which may be the result of poor spun-up in the OFES2. All three datasets captured the Atlantic tropical warming in  
428 the 1970s and from the 1990s to the 2000s, however, the two OFES datasets estimated a stronger intensity than EN4,  
429 especially the OFES1. In addition, OFES1 showed the appearance of significant cooling in the Atlantic tropics during  
430 the 1980s (Fig. 8g). Although a similar contemporary cooling was demonstrated by the OFES2, its cooling center was  
431 shifted several degrees southward. The Atlantic tropical cooling during the 1980s was not notable in EN4. The OFES2  
432 indicated an approximate 20-year (1960–1980) cooling episode in the vicinity of 45 °S in the Atlantic Ocean (Fig. 8k).  
433 A similar cooling trend existed in the 1960s, but with a relatively weaker intensity in EN4. In the Indian Ocean, the  
434 most significant agreement among the three datasets was observed, particularly the intense warming in the 2000s. In  
435 addition, there were some common cooling patterns observed from the 1980s to the 1990s in all three datasets. It was  
436 shown that the HV accounted for more substantial potential temperature changes than the SP, with the latter generally  
437 counteracting the HV (Supplementary Information, Figs. S1–2).

438 A general property of the similarities and differences between these three datasets is the fact that a better agreement  
439 was reached in the poleward of 30° than the latitudes equatorward of 30°. A possible explanation for this latitudinal  
440 dependence is that a deeper thermocline at higher latitudes responded less sensitively to the applied wind stress  
441 (Kutsuwada et al., 2019). Kutsuwada et al. (2019) found certain issues with the NCEP reanalysis wind stress that was  
442 used as atmospheric forcing in OFES1 as it generated a significantly shallower thermocline in the tropical North  
443 Pacific Ocean. Therefore, large negative temperature differences were observed when compared to the real  
444 observations along with the data obtained from OFES version forced by the wind stress from satellite measurements  
445 (QSCAT). The authors also claimed that the JRA-55 wind stress had problems similar to that of the NCEP wind.  
446 Indeed, the intense Pacific cooling patches in Fig. 8f were likely generated from the abnormally shallower thermocline  
447 in the tropical Pacific Ocean, consistent with Kutsuwada et al. (2019), although different temporal periods were  
448 considered.

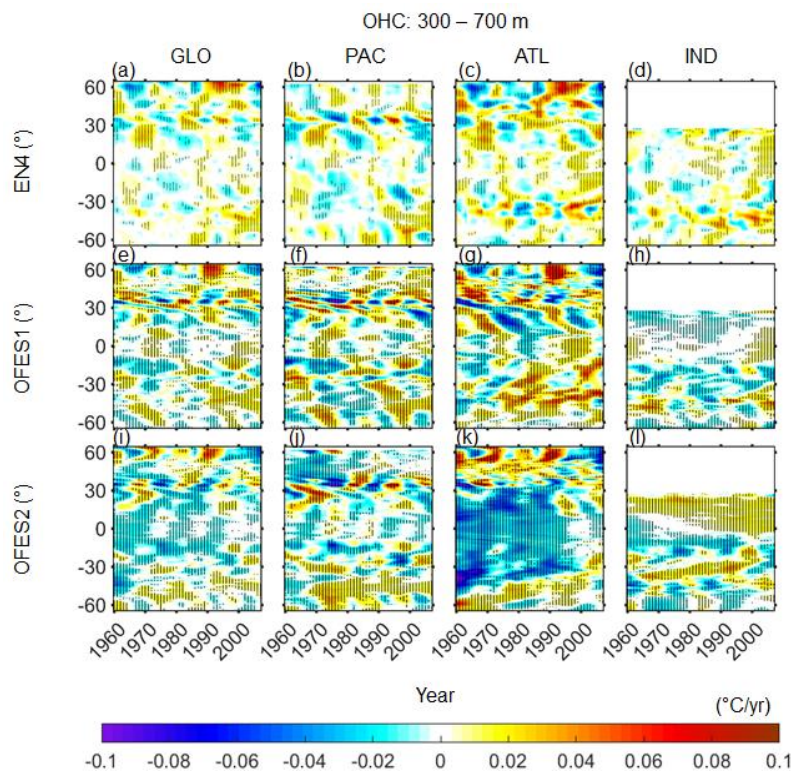
#### 449 450 *Middle layer*

451 In the intermediate layer between 300 and 700 m, the three datasets showed relatively poor agreement compared to  
452 the upper layer. The OFES2 differed from the others by displaying intense cooling before 2000 in the Atlantic Ocean  
453 (Fig. 9k) and a moderate but consistent warming trend in the northern Indian Ocean over almost the entire period (Fig.  
454 9l). In addition, there were large-scale cooling patches in the northern Pacific Ocean (Fig. 9j) and along the Indian  
455 equator (Fig. 9i) from the OFES2, while these cooling patches were not prominent in the other two datasets. These  
456 cooling distributions, obtained from OFES2, further demonstrated the place and timing of the cooling trend shown



457 Fig. 3, which may be partially attributed to the spin-up issue of the OFES2. Some similarities between the OFES2 and  
 458 the other two datasets have emerged in recent decades. For example, similar to EN4 and OFES1, the OFES2  
 459 reproduced the marked warming episodes observed in the high latitudes of the northern Atlantic Ocean during the  
 460 1980s and 1990s along with the subsequent cooling trend (Fig. 9c, g, k).

461 Upon comparing OFES1 with EN4, both similarities and differences can be discerned. The OFES1 generally agreed  
 462 with EN4 in regions located at the north of 30 °N, with some minor differences. However, in the tropics, large  
 463 differences were observed between OFES1 and EN4. For instance, the OFES1 indicated that the northern Indian Ocean  
 464 was mostly cooling (Fig. 9h), however, EN4 reflected alternate warming and cooling episodes (Fig. 9d). Furthermore,  
 465 the intense warming patches of the southern Atlantic Ocean demonstrated by the OFES1 (Fig. 9g) were not apparent  
 466 in EN4 (Fig. 9c). These potential temperature changes mainly resulted from the vertical displacement of the neutral  
 467 density surfaces, i.e., of the HV component (Supplementary Information, Fig. S3). However, the role of SP cannot be  
 468 ignored. This was especially clear in the southern hemisphere of EN4.



469  
 470 **Figure 9.** As for Fig. 8 but for the middle layer (300–700 m).  
 471

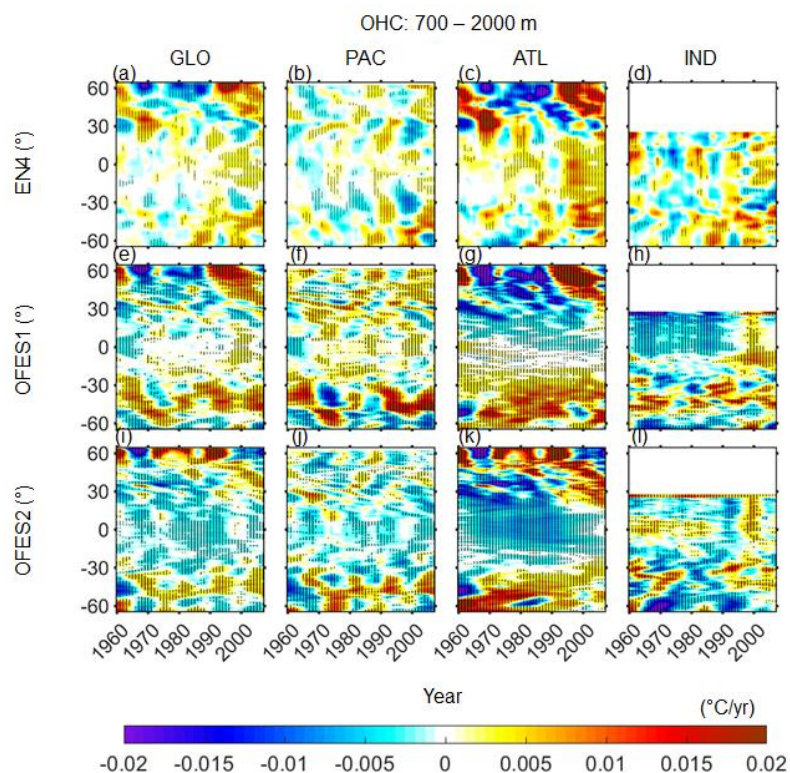
472  
 473 *Lower layer*

474 The northern Atlantic Ocean, especially to the north of 30 °N, dominated the global potential temperature change in  
 475 this lower layer (Fig. 10). This was generally more related to SP, especially in the intense cooling patch  
 476 (Supplementary Information, Fig. S6). Although the OFES1 data agreed well with EN4 in the northern Atlantic Ocean  
 477 (> 30° N), there were considerable differences between OFES1 and EN4. More specifically, OFES1 revealed that  
 478 there were intense HV-associated (Supplementary Information, Fig. S5) warming and cooling in the southern Pacific

479 Ocean during the 1960s and 1970s, however, such trend was not evinced in EN4. In addition, the warming of the  
 480 southern Pacific Ocean was much stronger in OFES1 than in EN4 since approximately 1990, which was associated  
 481 with the strong SP cooling in the southern Pacific Ocean, as revealed in EN4 (Supplementary Information, Fig. S6).  
 482 Moreover, OFES1 demonstrated consistent cooling of the Atlantic tropics, significant warming of the southern  
 483 Atlantic Ocean, and intense cooling of the northern Indian Ocean before the middle of the 1990s, which were not  
 484 evident in the EN4.

485 The OFES2 data captured some warming patterns in the southern hemisphere, similar to the OFES1; it also agreed  
 486 with the other two datasets in terms of the intense warming patches in the northern Atlantic Ocean in 1960s and after  
 487 ~1990. However, the agreement between OFES2 and the others was generally poor. This was most noticeable in the  
 488 cooling episode indicated by the OFES2 at the low and middle latitudes for both the Pacific and Atlantic Oceans,  
 489 especially the latter. OFES2 showed marked SP variations in the northern Atlantic Ocean (>30 °N), but generally  
 490 opposite to that as indicated by EN4. OFES1 indicated moderate SP in a similar warming/cooling pattern to EN4.

491 To summarize, the two OFES datasets had some good agreements with EN4 for the upper ocean layer, however,  
 492 such general agreement was largely confined to the middle-high latitudes. In general, the agreement for the ocean at  
 493 lower levels was poor. Specifically, in the middle ocean layer, the OFES1 displayed a generally reasonable agreement  
 494 with the EN4 for locations north to 30° N, however, large differences were observed elsewhere. In the OFES2,  
 495 intensive cooling patches were simulated, especially in the Atlantic Ocean. Although the spin-up issue may partially  
 496 explain the notable differences between the OFES and EN4 data for ocean water below 300 m, other causes might  
 497 have also contributed toward the examined differences.



498  
 499 **Figure 10.** As for Fig. 8 but for the lower layer (700–2000 m). Note the different colour scales.  
 500

### 3.3 Depth-time distribution of potential temperature, HV, and SP trends

Although we divided the top 2000 m into three layers, some details were lost while considering the averages of individual layers (i.e., the three vertical layers). In this section, we compare the depth-time patterns of the trends with respect to changes in potential temperature ( $\Delta\theta_{\text{OHC}}$ ) and its HV ( $\Delta\theta_{\text{HV}}$ ) and SP ( $\Delta\theta_{\text{SP}}$ ) components (Figs. 11–13).

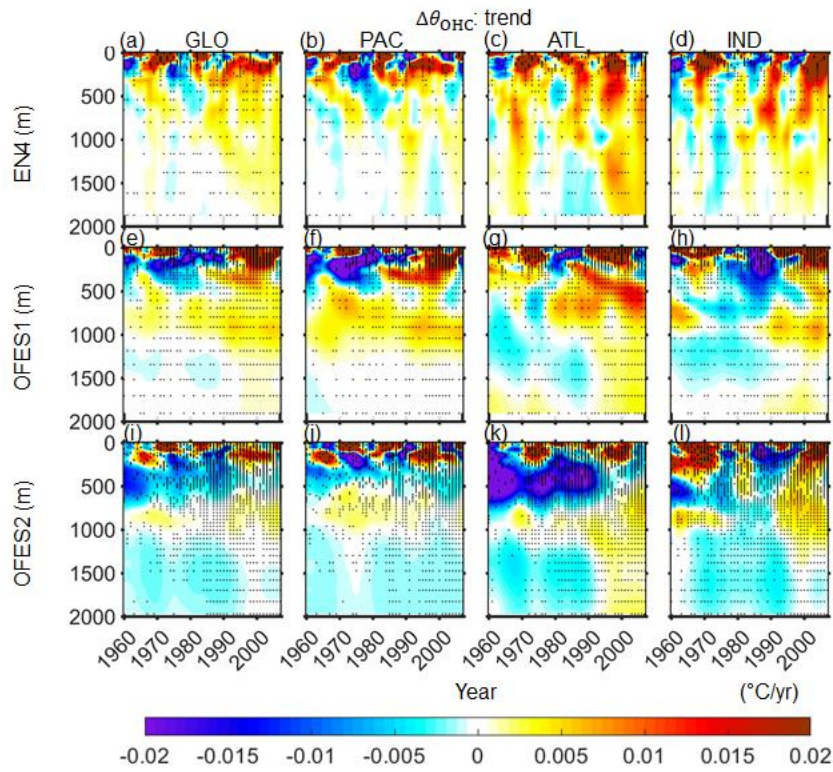
For the global ocean, the upper ocean layer above 300 m accounted for most of the warming or cooling trends (Fig. 11, left column). EN4 showed warming episodes over most of the investigated period, with only a few cooling episodes as a response to certain distinctive climate events. It can be seen that the volcanic eruptions of Mount Agung and El Chichón had a greater impact compared to the eruption of Pinatubo. The aforementioned strong cooling episode in the upper Pacific layer before 1990, which has been estimated from the OFES1, was initiated at a greater depth in the beginning, and subsequently, it terminated at a shallower depth (Fig. 11e). In the middle and lower layers, moderate warming or cooling trend was observed. Specifically, in EN4, moderate warming has extended to approximately 2000 m, since the early 1990s. The OFES1 showed moderate warming between 500 and 1000 m over almost the entire investigated period (Fig. 11e). Additionally, it indicated that since the middle of the 1990s, a weak warming trend has extended to 2000 m. The differences in the results of OFES2 relative to the other two datasets are apparent in the global ocean below approximately 200 m, where cooling is the dominant pattern (Fig. 11i); some weak warming patches between 500 and 1000 m are exceptions (Fig. 11i).

In the Pacific Ocean, the OFES2 had a generally reasonable agreement with EN4 above approximately 200 m, whereas the agreement between OFES1 and EN4 was poorer, despite some similar warming or cooling patches. Further below, EN4 showed alternate warming and cooling trends. The OFES1 reflected consistent warming between 500 and 1200 m, whereas the OFES2 estimated a consistent cooling trend below around 200 m, with some exceptions between 500 and 1000 m. Although beyond the scope of this work, the question of why both OFES1 and OFES2 showed relatively consistent warming trends between 500 and 1000 m near the permanent thermocline necessitates further work.

In the Atlantic Ocean, intense warming or cooling extended to deeper regions than in the Pacific Ocean. More specifically, the strong warming trend in the 1980–90s, estimated from EN4, extended to as deep as approximately 750 m. On the other hand, moderate warming trend extended to 2000 m since the middle of 1990s in EN4. The OFES1 well captured the warming trend of the 1970s and 1990s, along with the subsequent cooling period in the 2000s in the upper layer of the Atlantic Ocean, similar to EN4. However, the OFES1 estimated a strong cooling in the 1980s in the upper layer of the Atlantic Ocean, which was not evinced in the EN4. Interestingly, the OFES1 showed downward propagation of a strong Atlantic warming trend from approximately 200 m to 800 m since the early 1980s. Downward propagation of the cooling trend from approximately 600 m to 1800 m before ~1990 was also evinced in the OFES1 data of the Atlantic Ocean (Fig. 11g). Similar to EN4, a moderate warming trend extended to 2000 m since the middle of the 1990s in OFES1. In the case of OFES2, the most prominent pattern that distinguished it from the others was the extensive cooling patches before ~1990 in the upper and middle layers. In addition, it showed a moderate cooling below 1000 m before 1990. These two extensive cooling patterns in the upper-middle and the lower layers of the Atlantic Ocean, estimated using OFES2, raised the following questions: i) What are the main causes of the two cooling patches exhibited in the OFES2, and ii) Why the cooling patches suddenly terminated at approximately 1990? One possible

538 reason is the improvement in the reanalysis product of the atmospheric forcing since 1990, especially in the surface  
 539 HF and wind stress components, the latter being proved to be essential for subsurface temperature simulations  
 540 (Kutsuwada et al. 2019).

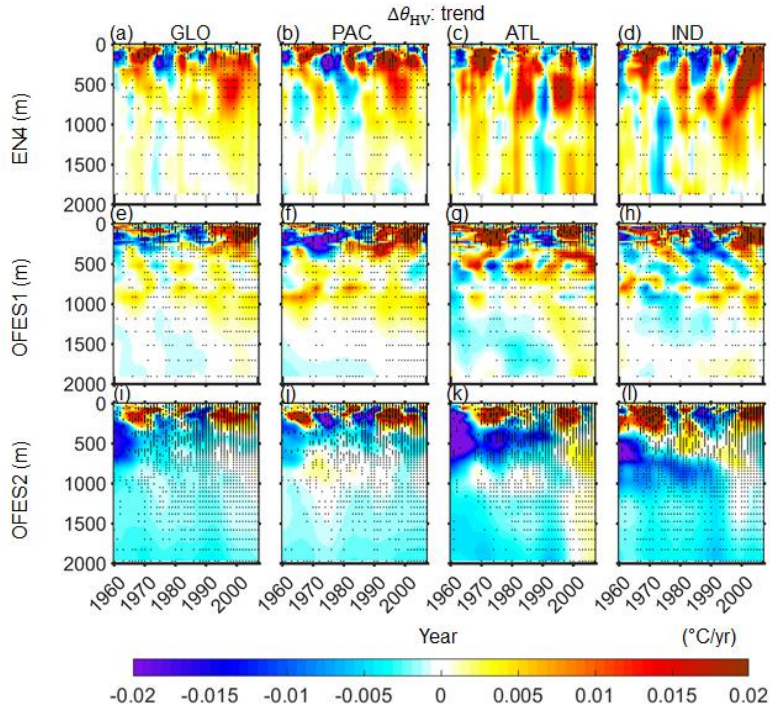
541 In the Indian Ocean, both OFES1 and OFES2 captured the warming trend in the 1960–70s and the 2000s, similar to  
 542 EN4. The OFES1 presented an intense cooling in the upper-middle layer during the 1980s; a similar but less extensive  
 543 and shallower cooling was also evinced in OFES2. But this cooling patch was significantly less prominent in EN4.  
 544 Beneath the upper layer, EN4 presented mostly warming in the Indian Ocean, with a major exception of a cooling  
 545 trend in the 1970s. In the two OFES, cooling pattern was more prominent than warming below 500 m, especially in  
 546 OFES2. However, between 500–1000 m, warming patches were seen in the 1960s and after ~1990, in both OFES1  
 547 and OFES2.



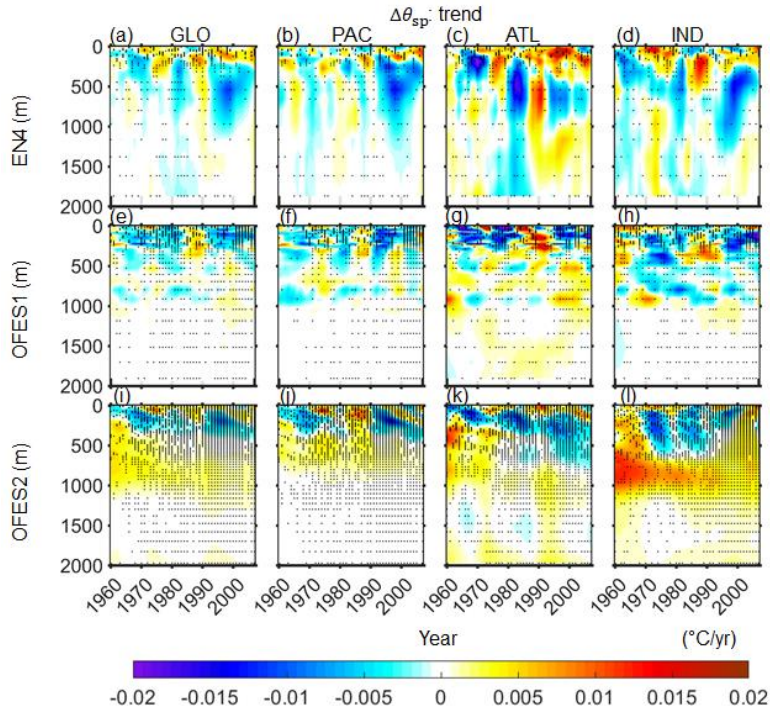
548  
 549 **Figure 11.** Depth-time patterns of the horizontally averaged potential temperature change  $\Delta\theta_{OHC}$  for (left to right) the  
 550 global, Pacific, Atlantic and Indian Oceans. **Top to bottom:** EN4, OFES1 and OFES2. Horizontal axis: year; vertical  
 551 axis: depth in m.  
 552

553 Upon comparing Fig. 11 with Figs. 12 and 13, it is evinced that to a great extent, the HV components dominated the  
 554 OHC variations. For instance, the profound warming and cooling patterns observed in Fig. 11 are mostly associated  
 555 with the HV component. The moderate cooling trend observed below 1000 m in OFES2 was also dominantly related  
 556 to HV. Although the SP was generally weaker and less important than the HV in accounting for the OHC variations,  
 557 its role cannot be ignored. Indeed, intense warming or cooling episodes associated with the SP component were  
 558 observed in EN4 in all major basins. The intensified subsurface SP cooling since the 1990s in the Pacific and Indian  
 559 Oceans, as indicated by EN4, has been particularly interesting, which could be associated with a significant increase

560 in subsurface salinity observations since the 1990s. A possible explanation for the appearance of the intensification of  
 561 SP cooling in the Pacific and Indian Oceans, but not in the Atlantic Ocean, is that the Atlantic Ocean has been better  
 562 observed than the Pacific and Indian Oceans before the 1990s. Another interesting point with regard to the SP is the  
 563 consistent SP warming trend that is observed in OFES2, especially in the Indian Ocean, and not in the other two  
 564 datasets.



565  
 566 **Figure 12.** Depth-time patterns of the horizontally averaged potential temperature change from the HV component,  
 567  $\Delta\theta_{HV}$ , for (left to right) the global, Pacific, Atlantic and Indian Oceans. **Top to bottom:** EN4, OFES1 and OFES2.  
 568 Horizontal axis: year; vertical axis: depth in m.



569

570 **Figure 13.** Depth-time pattern of the horizontally averaged potential temperature change from the SP component,  
 571  $\Delta\theta_{SP}$ , for (left to right) the global, Pacific, Atlantic and Indian Oceans. **Top to bottom:** EN4, OFES1 and OFES2.  
 572 Horizontal axis: year; vertical axis: depth in m.

573

### 574 3.4 Spatial patterns of the potential temperature, HV, and SP trends

575 To gain a more detailed understanding of the similarities and differences between the trends of potential temperature  
 576 estimated from the three datasets, here we have presented the spatial distributions of the potential temperature change  
 577 ( $\Delta\theta_{OHC}$ ), and its HV ( $\Delta\theta_{HV}$ ) and SP ( $\Delta\theta_{SP}$ ) components in the three ocean layers (Figs. 14–16).

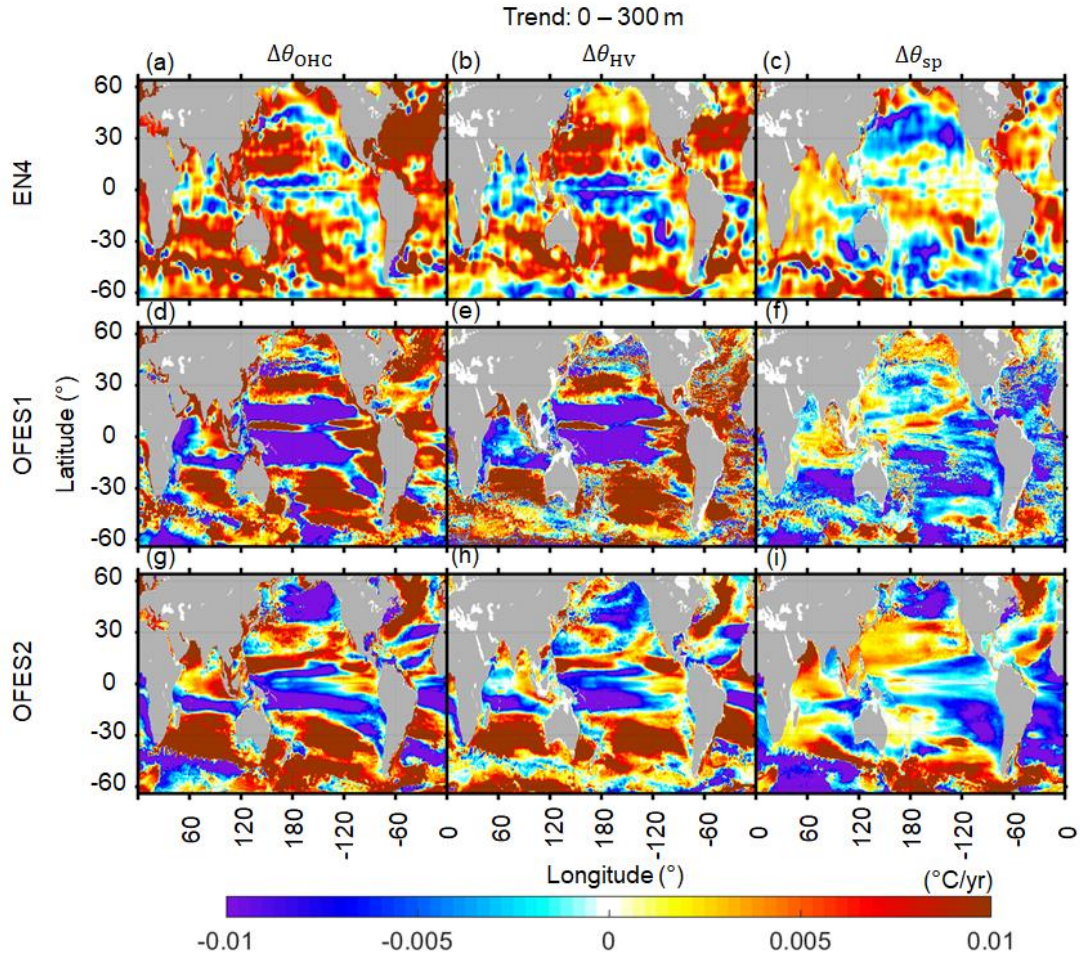
578

#### 579 *Upper layer*

580 Warming was almost ubiquitous in EN4 (Fig. 14a) and was particularly strong in the northern Atlantic Ocean and the  
 581 Southern Ocean. These two warming hotspots are expected from both theories and models. More specifically, the  
 582 shallow ocean ventilation in these two regions could generate faster warming than the global average (Banks and  
 583 Gregory 2006; Durack et al. 2014; Fyfe 2006; Talley 2003). Major cooling appeared in the western Pacific equator,  
 584 along the North Pacific Current, in the southeastern Pacific Ocean, parts of the Argentine Basin, and the southern  
 585 Indian tropics. All of these cooling regions accounts for a small fraction of the global ocean. Similar to EN4, both  
 586 OFES datasets showed significant warming in the subtropics, the high latitudes of the northern Atlantic Ocean, and  
 587 the Arabian Sea of the Indian Ocean. In addition, the OFES1 was similar to EN4 in terms of cooling along the North  
 588 Pacific Current. Despite these similarities, large differences exist between the three datasets. The most significant  
 589 difference was observed in the Pacific tropics. Although EN4 indicated the presence of a zonal band of cooling in the  
 590 Pacific tropics, this zonal band, when estimated using the OFES1 and OFES2 data, was much stronger in intensity

591 and more extensively stretched. It was mainly related to the HV component, especially in the case of OFES1. This  
592 strong cooling pattern in the vicinity of the equator was likely generated because of the poor qualities of the  
593 atmospheric wind stress over certain periods. As mentioned earlier, Kutsuwada et al. (2019) demonstrated that the  
594 NCEP wind stress used for forcing the OFES1 data generated a significantly shallower thermocline in the north Pacific  
595 tropical area, and therefore, negative differences were observed relative to the observations. In the northeast of the  
596 Pacific Ocean, the OFES2, but not the OFES1 and EN4, showed a patch of intense cooling, corresponding to the  
597 cooling pattern in the 1960–70s (Fig. 8j). The OFES2 also showed a couple of large cooling areas in the Atlantic  
598 Ocean (Fig. 14g). In the Indian Ocean, the OFES1 and OFES2 datasets indicated the presence of a patch of intense  
599 cooling in the southern Indian tropic and in the Indian sector of the Southern Ocean. Significant cooling also appeared  
600 in the western part of the north Indian Ocean in OFES1.

601 The decomposition of the changes in potential temperature into HV and SP components showed that the warming  
602 trend, estimated using EN4, was largely the result of isopycnal deepening (HV) in the subtropics. This is consistent  
603 with the finding that the subtropical mode water (STMW) is the primary water-mass accounting for global warming  
604 (Häkkinen et al., 2016), as discussed later. The SP was generally weaker than the HV and tended to counteract the HV  
605 warming, especially in the subtropics. This dampening effect can be easily understood from Fig. 1 of Häkkinen et al.  
606 (2016). For example, in a stratified ocean with warm and salty water overlying cold and fresh water, which is typically  
607 found in the subtropics, complete warming of one water parcel can be considered as the vector sum of warming and  
608 salination component, manifested as a transition from its original isopycnal to a new isopycnal (HV part) and a cooling  
609 and freshening component along the original isopycnal (SP). Two major exceptions of this cancellation between HV  
610 and SP were the northern Atlantic subtropics and the southern Indian Ocean in EN4, where HV warming was mostly  
611 accompanied by the SP warming. The SP warming in the northern Atlantic subtropics was generated owing to a  
612 substantial increase in salinity through evaporation (Curry et al., 2003; Häkkinen et al., 2016). Similarly, we found  
613 widespread positive SP warming in most of the Indian Ocean in EN4, except west to southwest Australia. This SP-  
614 related warming in the northern Indian Ocean dominantly controlled the potential temperature change in EN4,  
615 especially in the Arabian Sea. The most significant SP warming, however, was found in the Indian sector of the  
616 Southern Ocean (may be related to the salination of the Southern Ocean), southern subtropics of the Atlantic Ocean,  
617 and Labrador Sea (Fig. 14c).



618  
 619 **Figure 14.** Spatial distributions of  $\Delta\theta_{\text{OHC}}$  (left column),  $\Delta\theta_{\text{HV}}$  (middle column) and  $\Delta\theta_{\text{SP}}$  (right column), 1960–2016,  
 620 in the top ocean layer (0–300 m). **Top to bottom:** EN4, OFES1 and OFES2.  
 621

622 Comparing the HV components in the three datasets showed that the two OFES simulations were able to reproduce  
 623 some subtropical HV warming patterns, although less accurately in the northern hemisphere. The strong and extensive  
 624 equatorial cooling in the Pacific and Indian Oceans was largely associated with variations in HV in the two OFES  
 625 datasets.

626 The SP in the OFES1 was similar to EN4 in the northern subpolar region of the Pacific Ocean, parts of the northern  
 627 Pacific subtropics, the Labrador Sea, and parts of the northern Indian Ocean. The SP, estimated using OFES2, was  
 628 similar to the estimates from the EN4 in the Labrador Sea and the western Indian Ocean. In general, however, no  
 629 common patterns were observed in most of the global oceans. Neither of the OFES datasets captured the SP warming  
 630 in the western part of the northern Atlantic subtropics. The OFES2, but not EN4 and OFES1, indicated moderate SP  
 631 warming in the North Pacific subtropics and intense SP warming in the Pacific sector of the Southern Ocean,  
 632 respectively. The improvements in SP determined based on the OFES2 dataset over that from the OFES1 in the  
 633 Arabian and Indonesian seas, and not in the Bengal Bay, is partly consistent with **S2020**. The authors demonstrated a  
 634 smaller bias in the water-mass properties of the Arabian and Indonesian seas, however, a large salty bias remained in  
 635 the Bengal Bay in the OFES2.



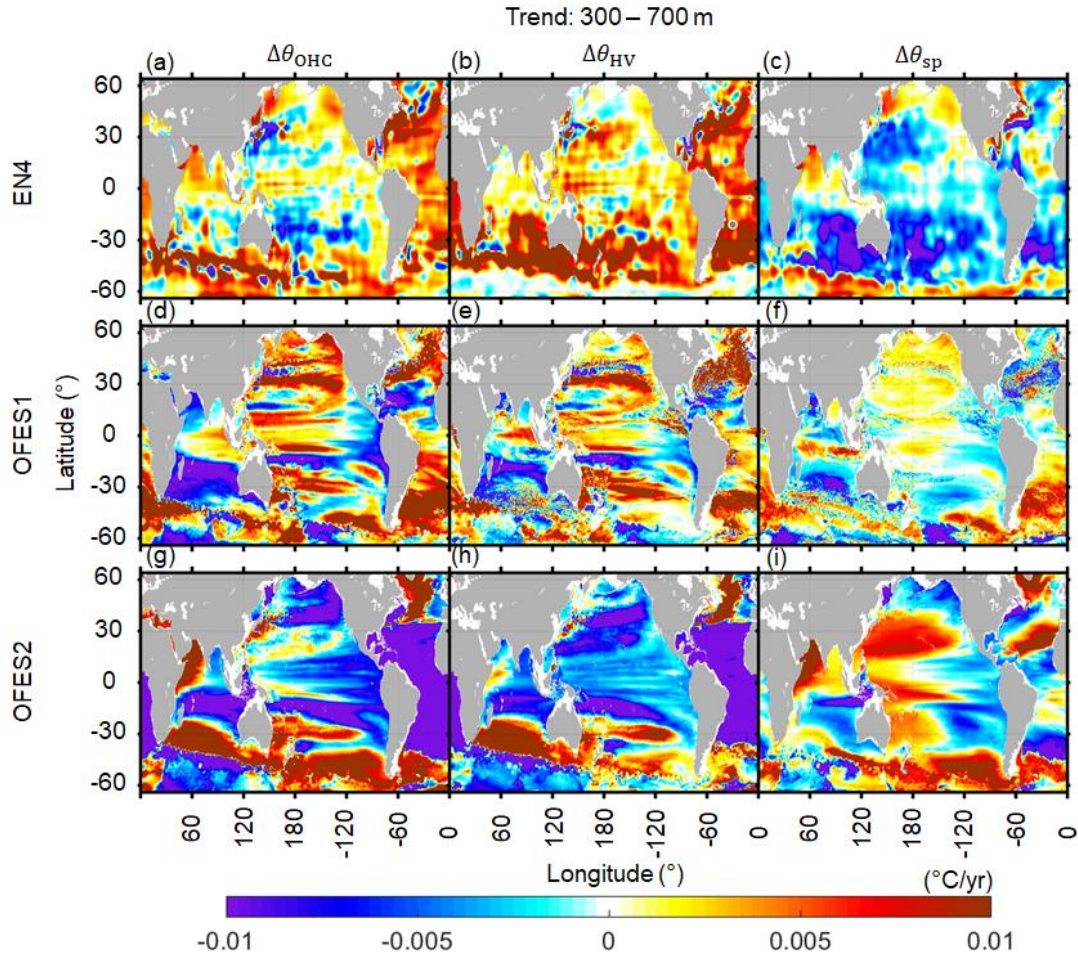
636 In Fig. 2, we show that the SP, estimated using EN4 and OFES2, was largely similar in the upper layer of the Pacific  
637 Ocean. However, the spatial distributions of the SP component in the Pacific Ocean were seldom similar between EN4  
638 and OFES2. In other words, the time series of a basin-wide quantity hides many details.

639  
640 *Middle layer*

641 EN4 showed that the cooling of the ocean was mostly concentrated in the southern Pacific subtropics and the region  
642 associated with the Kuroshio (Fig. 15a). Clear warming trend was observed, accompanied by sporadic cooling patches  
643 in the rest of the global ocean, especially over most of the Atlantic Ocean, in the northern Indian Ocean, and along the  
644 Antarctic Circumpolar Current (ACC) path of the Southern Ocean. The OFES1 dataset could reproduce some warming  
645 patterns in the northern Pacific Ocean, the bulk of the Atlantic Ocean, the eastern part of the northern Indian Ocean,  
646 and parts of the ACC path. However, notable differences were found between OFES1 and EN4. Among these  
647 differences, the most prominent is the intense cooling in the southern Indian Ocean as estimated from OFES1. In  
648 addition, strong cooling patches were also found in the southern Pacific tropics, west to central-south America, in the  
649 northern Atlantic subtropics, in the Arabian Sea, and along parts of the southern edge of the ACC in OFES1. The  
650 pattern in the OFES1 Pacific Ocean clearly appears as zonal bands. Consistent with Fig. 3, intense cooling was  
651 simulated by OFES2 for all major basins, with the most prominent being in the Atlantic Ocean. Large-scale warming  
652 patterns were found in the Kuroshio region, in the southern Pacific and Indian subtropics, in the northern Atlantic  
653 Ocean (north of 35° N), in the western part of the northern Indian Ocean, and in the Pacific and Atlantic sectors of the  
654 Southern Ocean. In general, there were apparent differences between the three datasets when the bulk of the global  
655 ocean was considered. The above 700 m is relatively well observed, especially in the Atlantic Ocean (even back to  
656 1950–60s, Häkkinen et al., 2016). Therefore, it is likely that the OFES2 dataset was an outlier at the analyzed multi-  
657 decadal scale, and there could be some potential problems in the OFES1, for example, in the southern Indian Ocean.

658 Interestingly, EN4 suggested that HV warming was almost ubiquitous in the middle layer (Fig. 15b), especially in  
659 the southern hemisphere, which is consistent with the warming shift toward the southern hemisphere (Häkkinen et al.,  
660 2016). Correspondingly, SP cooling also occupies most of the global ocean (Fig. 15c), with a similar southern shift,  
661 the most prominent being around the east and western regions of Australia. Major SP warming patches were found in  
662 the Sea of Okhotsk, north of the Gulf Stream, in the Arabian Sea, and along the southern edge of the ACC. These  
663 regions are generally associated with strong variations in salinity. Comparing HV and SP estimated based on EN4 and  
664 OFES1 dataset showed that the OFES1 captured some warming patterns in the Pacific and Atlantic, but not in the  
665 Indian subtropics. The agreement of HV for the southern Pacific, Indian tropics, and the Southern Ocean was mostly  
666 poor. In the case of SP, the OFES1 reproduced intense SP cooling in western Australia and the southern Pacific  
667 subtropics, similar to EN4, despite its smaller coverage. However, OFES1 showed almost opposite trends of SP over  
668 most of the global ocean. In OFES2, both HV and SP were strong, however, the basin-wide cooling was mainly  
669 generated as a result of HV. Overall, the OFES2 dataset had a reasonable agreement with EN4 in the southern  
670 subtropics (Pacific and Indian Oceans) in terms of HV. It also had a common HV warming patch in the northern  
671 Atlantic Ocean (north to 35° N) as EN4. With regard to SP, the OFES2 was similar to EN4 in displaying SP warming  
672 in the Arabian Sea and parts of the southern edge of the ACC. In addition, it also captured SP cooling in the eastern

673 Pacific Ocean, along the Gulf Stream path, west of Australia. Except for these similarities, however, OFES2 dataset  
 674 was generally not consistent with that of EN4 in terms of SP.

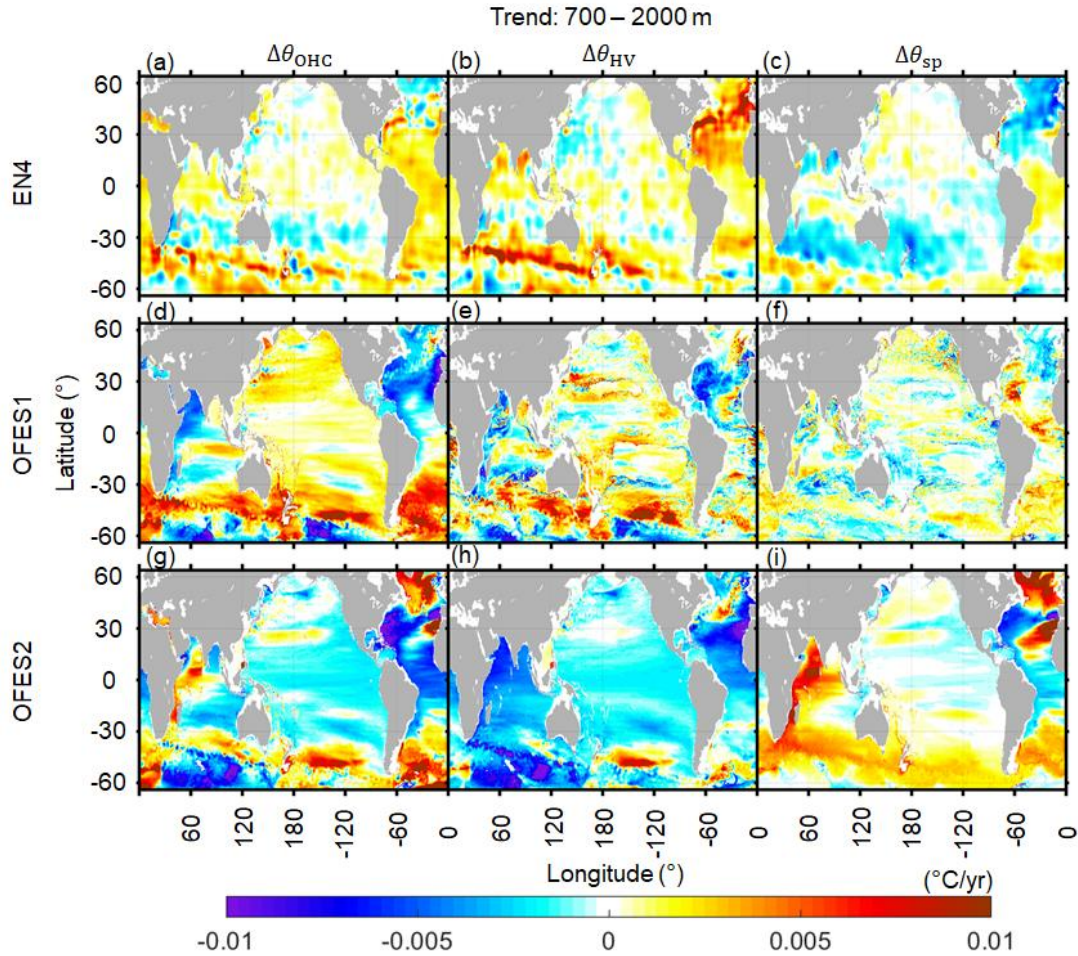


675  
 676 **Figure 15.** As for Fig. 14 but for the middle layer (300–700 m).  
 677

678 *Lower layer*

679 In general, the warming and cooling intensities were significantly weaker in the lower layer compared to that in the  
 680 top two layers, which is consistent with several previous findings that more heat was stored in the upper 700 m than  
 681 at greater depths (Häkkinen et al., 2016; Levitus et al., 2012; Wang et al., 2018; Zanna et al., 2019). EN4 showed  
 682 widespread warming patches in the Southern and Atlantic Oceans, and three large zonal bands of cooling in the  
 683 southern subtropics of the Pacific and Indian Oceans, and in the northern subpolar region of the Atlantic Ocean (Fig.  
 684 16a). Similar to EN4, the OFES1 dataset reflected warming along the northern edge of the ACC and in the southern  
 685 Atlantic Ocean, but the intensity of warming was much stronger for OFES1 than in EN4 (Figs. 16a, d). OFES1  
 686 reflected moderate warming over almost the entire Pacific Ocean, which was not the case in EN4. Significant  
 687 differences between OFES1 and EN4 were also found in the northern Atlantic Ocean, where the OFES1 showed  
 688 extensive cooling compared to the moderate warming in EN4. OFES1 demonstrated strong cooling in the Arabian  
 689 Sea, which is in contrast to negligible variations the Arabian Sea obtained from the EN4. To some extent, the OFES2

690 was similar to the other two datasets in showing warming along the northern edge of the ACC and in the southern  
 691 Atlantic Ocean, south to 30 °S (Fig. 15g), despite the differences in the intensity of warming. It also showed cooling  
 692 in the low and middle latitudes of the Atlantic Ocean, similar to OFES1 but opposite to EN4. The bulk of the Pacific  
 693 Ocean was shown to be cooling in the OFES2 (Fig. 15g), which was almost opposite to the OFES1 results (Fig. 15d)  
 694 and similar to EN4 only in parts of the southern Pacific subtropics (Fig. 15a). Moreover, OFES2 reflected intense and  
 695 widespread cooling in the Indian sector of the Southern Ocean.



696  
 697 **Figure 16.** As for Fig. 14 but for the lower layer (700–2000 m).  
 698

699 In NE4, there was intense HV warming along the northern edge of the ACC in the Indian and Pacific Oceans, and in  
 700 the northern Atlantic Ocean (Fig. 16b), which largely accounted for the total potential temperature variations. HV  
 701 warming was generally accompanied by SP cooling (Fig. 16c). Moderate HV and SP warming coexist in the northern  
 702 Atlantic tropics and the southern Atlantic Ocean in EN4. We found that the OFES1 captured the HV warming pattern  
 703 along the northern edges of the ACC, which to some extent, is consistent with the results from EN4. However, there  
 704 were remarkable differences in OFES1 results from those of EN4, particularly in the northern Atlantic and Indian  
 705 Oceans. In terms of SP, there were some similarities between the OFES1 and EN4; for example, they both had SP  
 706 cooling and warming in the northern and southern Atlantic Ocean, respectively. Among the three datasets, OFES2

707 showed the most extensive and strong HV-associated cooling, except for a patch of HV warming in the Pacific sector  
708 of the Southern Ocean, which was also observed in the other two datasets. The OFES2 estimated intense SP warming  
709 in the Southern Ocean, the western Indian Ocean, and the northern Atlantic subpolar regions. A large-scale patch of  
710 abnormally strong SP warming, associated with the Mediterranean Overflow Water (MOW), was also observed. This  
711 extremely strong SP warming, associated with MOW, is likely the result of the unrealistic spreading of the salty  
712 Mediterranean overflow reported in **S2020**.

713 Besides the above-discussed multi-decadal linear trends, we have demonstrated that (not shown here) the significant  
714 differences between the two OFES datasets and the EN4 were significantly reduced if the period between 2005 and  
715 2016 is considered, during which the two OFES were argued to be well spun-up (**S2020**). In addition, over this 12-  
716 years period, the spatial pattern of the OFES2 showed some improvements over the OFES1 for the upper and middle  
717 layers, however, it was not necessarily true for the lower layer when EN4 was used as a reference. Is this better  
718 agreement a result of better spun-up or it was generated owing to improvements in the reanalysis product of the  
719 atmospheric forcing for the two OFES data? This interesting question requires further exploration in the future.

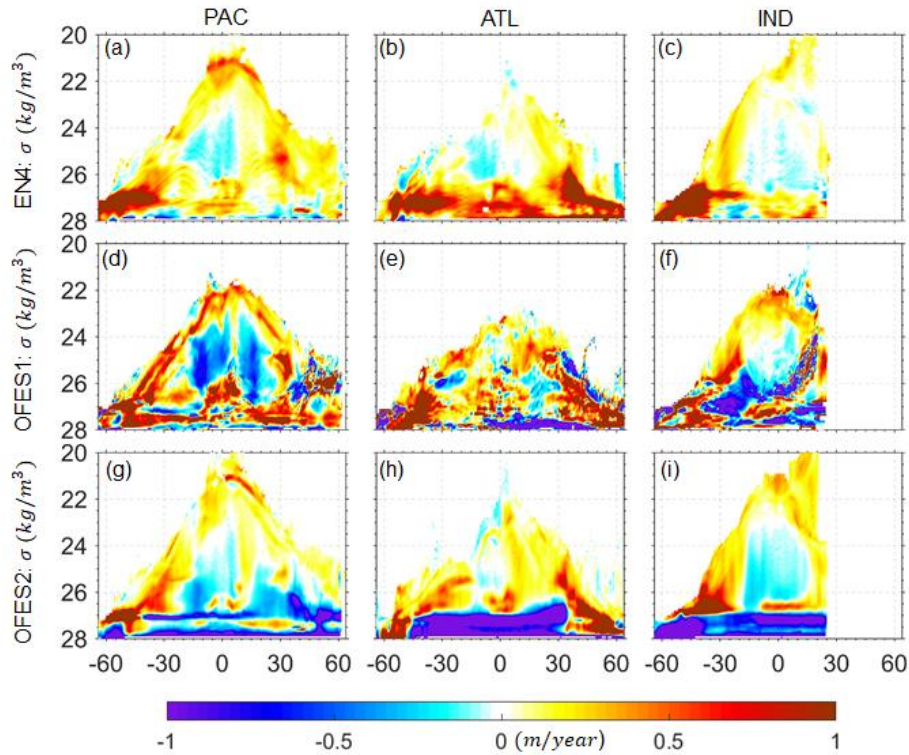
### 720 **3.5 Trends of HV and SP in the neutral density domain**

721 Plotting the HV and SP components in neutral density coordinates provides useful information to analyze the warming  
722 and cooling from the perspective of water-mass. Following Häkkinen et al. (2016), we calculated the linear trend of  
723 the zonal-averaged sinking of the neutral density surfaces in each major basin over 1960–2016 (Fig. 17). We also  
724 calculated the zonal-averaged SP-related warming or cooling along the neutral density surfaces (Fig. 18).

725 Our results, based on the EN4 dataset, were similar to those of Häkkinen et al. (2016), who used an earlier version  
726 of EN4 dataset (i.e., EN4.0.2) and considered the period from 1957 to 2011. More specifically, our EN4 results showed  
727 that bulk HV warming (deepening of neutral density surfaces) was associated with water-mass of over  $26 \text{ kg/m}^3$ , and  
728 was mainly concentrated south of  $30^\circ \text{ S}$ , from the ventilation region at high latitudes to the subtropics. There was one  
729 exception in the Atlantic Ocean, where deepening of isopycnal heavier than  $26 \text{ kg/m}^3$  occurred at all the considered  
730 latitudes. The concentrated warming in the northern Atlantic Ocean was attributed to phase change of the North  
731 Atlantic Oscillation (NAO) from negative in the 1950–60s to positive in the 1990s (Häkkinen et al. 2016; Williams et  
732 al. 2014). As explained by Häkkinen et al. (2016), the significant deepening of neutral density surfaces was associated  
733 with the subtropical mode water (STMW,  $26.0 < \sigma_0 \text{ (kg/m}^3) < 27.0$ ) and the Subantarctic Mode Water (SAMW,  $26.0$   
734  $< \sigma_0 \text{ (kg/m}^3) < 27.1$ ). These vertical displacements of neutral density surfaces probably resulted from heat uptake via  
735 subduction, which subsequently might have spread from these high-latitude ventilation regions. The large vertical  
736 deepening of the STMW and SAMW had subsequently pushed the Subpolar Mode Water (SPMW,  $27.0 < \sigma_0 \text{ (kg/m}^3)$   
737  $< 27.6$ ) and the Antarctic Intermediate Water (AAIW,  $27.1 < \sigma_0 \text{ (kg/m}^3) < 27.6$ ) further down. However, as the vertical  
738 displacement of the STMW/SAMW was larger, its volume would have increased, and the volume of the underlying  
739 SPMW/AAIW decreased (Häkkinen et al., 2016). Besides these significant sinking of neutral density surfaces, there  
740 was generally a shoaling pattern of lower density ( $\sigma_0 \text{ (kg/m}^3)$ ) ranging from 24 to 26, which was mostly concentrated  
741 between the equator and  $30^\circ \text{ S}$ . To a large extent, this shoaling occurred in the central water, for example, in the South  
742 Pacific Central Water (SPCW).

743 In this study, we have not focused on the detailed mechanisms of warming from the perspective of water-mass, as  
 744 has been done in previous studies (Häkkinen et al., 2016). Instead, we have focused on the differences between the  
 745 three datasets with respect to the trends of HV and SP.

746 It can be seen that along the surfaces of the Pacific and Indian Oceans, there was a general appearance of HV warming  
 747 in almost all three datasets. In the Atlantic Ocean, however, the EN4 estimated a sea surface cooling south of 30° S  
 748 and in the northern tropic; the OFES2 also estimated a cooling trend near the surface of the Atlantic tropics. In contrast  
 749 to both EN4 and OFES2, OFES1 showed an intense HV cooling pattern along the Atlantic surface between 30 and  
 750 50° N (Fig. 17e).



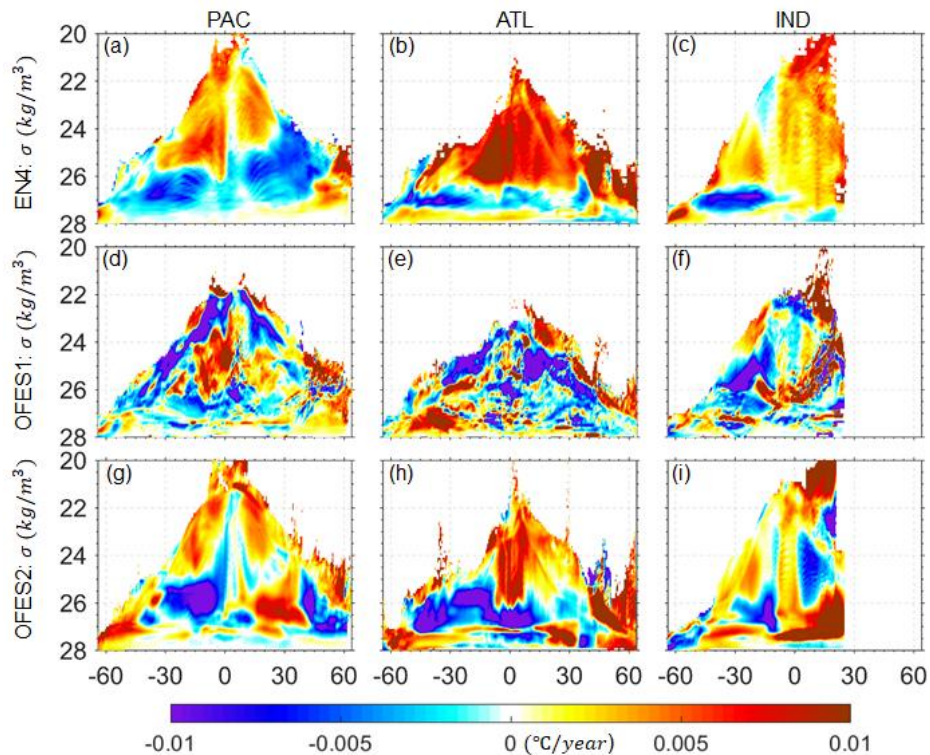
751  
 752 **Figure 17.** Linear trends in the zonal-averaged sinking of the neutral density surfaces in the Pacific (**left column**),  
 753 Atlantic (**middle column**) and Indian (**right column**) Oceans. **Top to bottom:** EN4, OFES1, OFES2. Positive values  
 754 mean deepening of the neutral density surfaces. The calculation was for the water above 2000 m.  
 755

756 South of 30° S, EN4 detected large downward movements, associated with the STMW, SAMW, and AAIW in all  
 757 three basins. In the case of OFES1, the dominant pattern in the three basins was sinking, however, it was surrounded  
 758 by shoaling patches; larger differences from the EN4 were found in the OFES2, which showed significant and  
 759 extensive shoaling patterns, especially in the Atlantic and Indian Oceans. The almost opposite trend in the vertical  
 760 displacements of the neutral density surfaces between the OFES2 and the observation-based EN4 may indicate that  
 761 the changes of properties of water-mass simulated in the OFES2 were unrealistic, at least at this multi-decadal scale.

762 In the ocean interior between 30°S and 30° N, the OFES1 presented shoaling patterns in the Pacific and Indian  
 763 Oceans, however, such shoaling pattern was not prominent in the Atlantic Ocean. Although the shoaling patterns in  
 764 the Pacific and Indian Oceans were also evinced in EN4, their magnitude was generally weaker. The OFES2 had better

765 agreement with EN4 for the shoaling pattern in the southern Pacific subtropics. OFES2 also captured shoaling in the  
 766 Indian Ocean, with similar coverage, however, the intensity was generally stronger. Shoaling in the southern Atlantic  
 767 subtropics was not prominent in OFES2, similar to the OFES1, but different from the EN4.

768 In the north of 30° N, EN4 detected widespread sinking, particularly in the northern Atlantic Ocean. This strong  
 769 sinking in the northern Atlantic Ocean originated mainly from SPMW and STMW. In the EN4 Pacific Ocean, there  
 770 were certain shoaling patches, which were related to the North Pacific Intermediate Water (NPIW). In OFES1, the  
 771 pattern was filled with both sinking and shoaling patches, which defies easy interpretation. However, an apparent  
 772 outlier of OFES1 was the intense shoaling in the northern Atlantic Ocean (mostly below 700 m (Figs. 14–16)), which  
 773 is the opposite of EN4. The shoaling of neutral density surfaces in the OFES2 Pacific Ocean, north to 30° N, was even  
 774 more prominent than that in the OFES1. The OFES2 had a better agreement with EN4 in terms of the sinking patterns  
 775 in the Atlantic Ocean north of 30° N.



776  
 777 **Figure 18.** Linear trends in the zonal-averaged warming or cooling along the neutral density surfaces in the Pacific  
 778 (left column), Atlantic (middle column) and Indian (right column) Oceans. Top to bottom: EN4, OFES1, OFES2.  
 779

780 The major SP warming episodes determined by EN4 in the Pacific Ocean were associated with STUW and Pacific  
 781 Central Water (PCW) in the low and middle latitudes, with a shift toward the southern hemisphere. The northern high-  
 782 latitude SP warming was mainly related to the Pacific Subarctic Intermediate Water (PSIW). The two SP coolings  
 783 were generated from the STMW, accompanying the isopycnal deepenings in Fig. 17(a). HV warming/ SP cooling was  
 784 particularly typical in the subtropical regions, and HV/ SP warming was typical in the subpolar regions, more details  
 785 of which are presented in Häkkinen et al. (2016). An extremely strong SP warming trend occurred in the Atlantic

786 Ocean, resulting from salination via evaporation. In the southern Atlantic Ocean, the pattern of SP cooling was mostly  
787 accompanied by the sinking of the STMW.

788 The SP pattern determined from the OFES1 dataset was quite noisy, and generally had a poor agreement between  
789 OFES1 and EN4 in terms of SP warming, which is likely to result from some issues of simulation of salinity in the  
790 OFES1. As shown in **S2020**, OFES1 was not capable of simulating salty outflows, for example, the outflow through  
791 the Persian Gulf into the Indian Ocean. There were notable improvements in the salinity fields of OFES2 over OFES1,  
792 which has been mainly attributed to the inclusion of river runoff and sea-ice, however, some issues associated with  
793 poor performance in the simulation of the MOW remained. Overall, the SP warming pattern in the density coordinate  
794 was significantly improved in OFES2 compared to OFES1. However, upon combining Figs. 14–16, it is evinced that  
795 the similarities in SP estimation between the OFES2 and EN4 dataset were confined to a small fraction of the global  
796 ocean, mainly in the upper and middle layers of the Labrador Sea, northern Indian Ocean, and Southern Ocean. In  
797 addition, the simulations by OFES2 shared similarities with those of EN4 in showing a patch of SP cooling in the  
798 western part of the northern Atlantic subtropics.

799

### 800 **3.6 A basin-wide heat budget analysis**

801 The primary processes controlling the oceanic thermal state include the net surface HF, ZHA and MHA in the  
802 horizontal direction, and VHA and VHD in the vertical direction (Fig. 1b). Lateral heat diffusion was not considered  
803 here because it was found to play a minor role in our analysis (not shown). Because our focus is on the global and  
804 basin-wide OHC in the three vertical layers, we calculated and compared the inter-basin heat exchange, and the VHA,  
805 integrated over each basin from 1960 to 2016. No vertical heat diffusivity data were available from OFES1. In addition,  
806 the vertical heat diffusivity data from OFES2 were temporarily unavailable because of a security incident when this  
807 manuscript was prepared. This prevented us from directly calculating and comparing the VHD between OFES1 and  
808 OFES2. As an alternative solution, we calculated the residual of the OHC change, along with all associated heat  
809 transport components that contribute to each basin, and used the results as a proxy for VHD. This indirect method  
810 might suffer from some errors; for instance, it includes the impacts of river runoff in the OFES2, however, it can still  
811 provide us with some important information. The calculations are listed in Table 2–4. The related time series of these  
812 surface heat fluxes and heat advection are shown in Supplementary Figs. S7–9.

813

#### 814 *Upper layer*

815 In the Pacific Ocean, the rate of change of the OHC was rather low for both OFES1 and OFES2. The average surface  
816 HF, estimated using the OFES1 dataset, was twice that of OFES2, indicating that heavier heating applies to the OFES1  
817 Pacific Ocean, signifying their differences in atmospheric forcing. Vertically, both datasets indicated a net downward  
818 advection of heat in the Pacific Ocean at 300 m, however, the intensity was much stronger in the OFES1 (different by  
819 approximately  $0.7 \text{ W/m}^2$ ), which may be related to their different wind-forcing sources, as the downward heat  
820 advection in the upper ocean was mainly from the wind-driven Ekman pumping in the subtropical gyres. Indeed,  
821 Kutsuwada et al. (2019) claimed that the NCEP wind stress curl was too strong, and had generated the overly strong  
822 Ekman pumping. There was an increase in the eastward heat advection through the water passage between the

823 Australian mainland and 64° S by 0.150 W/m<sup>2</sup> (P3 in Fig. 1a) in the OFES2, in a comparison to the OFES1. Although  
824 the two OFES datasets indicated that the MHA from the Southern Ocean to the Pacific Ocean (P4) had opposite signs,  
825 the relatively small absolute value indicated that this difference was not essential. The Drake Passage (P5) is the major  
826 water passage, through which heat is exchanged between the Pacific and Atlantic Oceans. There was 0.108 W/m<sup>2</sup>  
827 more heat loss through P5 into the Atlantic Ocean in the OFES1, inferring a stronger ACC from the OFES1 in the  
828 upper ocean. P7 and P8 connect the Pacific and Indian Oceans, and the Indonesian Throughflow (ITF) flows through  
829 P7. The MHA passing through P7 was almost one time stronger in OFES2 than in OFES1, with a difference of 0.637  
830 W/m<sup>2</sup>. This indicated an enhancement of the ITF simulated by the OFES2. This to some extent, agreed with the results  
831 of Sasaki et al. (2018), who showed that the inclusion of a tidal-mixing scheme resulted in an intensification of the  
832 ITF, noting that the tidal-mixing scheme was implemented in OFES2 but not in OFES1. In addition, the OFES1  
833 demonstrated that more heat was transported westward into the Indian Ocean between Papua New Guinea and  
834 Australia (P8), however, the small absolute heat advection indicated that it was not the major cause of the OHC  
835 discrepancy between OFES1 and OFES2. The net heat advection through the Bering Strait (P9) was rather weak in  
836 both datasets. The indirect calculation of the VHD showed that there was net downward heat diffusion at the depth of  
837 300 m in the Pacific Ocean in both the OFES datasets, although the intensity was much stronger (0.747 W/m<sup>2</sup>) in the  
838 OFES1.

839 In the Atlantic Ocean, the OHC increased at an average rate of 0.032 W/m<sup>2</sup> in OFES1, however, it decreased by  
840 0.014 W/m<sup>2</sup> in OFES2. There was net surface heating in the OFES1 Atlantic Ocean, but minor cooling was evinced  
841 in the OFES2. The two OFES datasets were also profoundly different in terms of VHA at 300 m. Specifically, OFES1  
842 showed a net downward heat advection, and OFES2 showed upward and significantly weaker heat advection. Again,  
843 this difference in the VHA was likely the result of different wind stress datasets in the two OFES, as discussed above.  
844 In a comparison to the OFES2, the OFES1 showed an increase in the heat transported from the Atlantic Ocean to the  
845 Indian Ocean through P1 between South Africa and 64° S by 0.158 W/m<sup>2</sup>. As mentioned above, more heat was  
846 advected into the Atlantic Ocean through the Drake Passage (P5) in OFES1. Additionally, there was more heat  
847 advected southward from the Atlantic Ocean to the Southern Ocean in OFES1 (P6). The wide passage connecting the  
848 North Atlantic Ocean to the Arctic Ocean (P10) served as the major channel, through which the Atlantic Ocean  
849 exchanged heat with the Arctic Ocean; the two OFES datasets exhibited similar heat loss. All these differences led us  
850 to conclude that the resulting VHD at 300 m differed by 0.411 W/m<sup>2</sup> (with stronger upward heat diffusion estimated  
851 by the OFES1).

852 For the Indian Ocean, the average rate of increase in OHC, calculated by OFES2, was higher than in OFES1 by 0.009  
853 W/m<sup>2</sup>. The time-averaged surface HF in OFES2 was 0.729 W/m<sup>2</sup> lesser than that in OFES1. Both datasets showed a  
854 net downward heat advection, however, the results obtained from OFES2 were approximately two times stronger. The  
855 small difference in the southward heat advection across 64° S ( P2) only affected the OHC in the upper Indian Ocean  
856 to a small extent. In contrast, the differences in the HF, VHA, and MHA associated with the ITF contributed to the  
857 difference and led us to calculate a remarkable difference of 1.898 W/m<sup>2</sup> in the VHD at a depth of 300 m in the Indian  
858 Ocean. The enhanced ITF is one of the main contributors to the larger increase of the OHC in the upper layer of the  
859 OFES2 Indian Ocean (Fig. 2).



860 To summarize, OFES1 estimated a higher surface HF into the major basins. The VHA was generally downward,  
 861 indicating the essential role of subtropical Ekman pumping in the heat uptake of the upper ocean layer. The differences  
 862 between these two (HF and VHA) contributors were mainly due to the different atmospheric forcing used in the two  
 863 OFES datasets, emphasizing the importance of reliable atmospheric forcing products in numerical ocean modeling.  
 864 Although the different wind stresses could also produce different lateral advection through P1–P10, the horizontal  
 865 heat advection through these passages are largely similar in the two OFES. The most prominent difference in the  
 866 lateral heat advection was associated with the ITF, which was to some extent a result of the adoption of a tidal-mixing  
 867 scheme. This ITF-related difference and the indirectly inferred VHD suggested the significance of the vertical mixing  
 868 scheme in producing the examined differences in OHC.

869 **Table 2.** Time-averaged OHC increasing rate, surface heat flux (HF) and advection of heat through the major water  
 870 passages for the upper layer (0–300 m) of each basin. VHA in this table is at a depth of 300 m. Residual: difference  
 871 between the OHC increase and all the heat flux into a basin, approximately the VHD. All quantities converted to  $W/m^2$   
 872 applied over the entire surface of the Earth. Values smaller than 0.001 are set to 0. Positive means heat gain and  
 873 negative means heat loss.

PACIFIC OCEAN (0–300 m)										
	OHC	HF	VHA	P3	P4	P5	P7	P8	P9	Residual
OFES1	–0.025	2.135	–0.814	1.233	0.011	–0.891	–0.728	–0.162	–0.003	–0.808
OFES2	0.007	1.066	–0.113	1.383	–0.020	–0.783	–1.365	–0.100	0	–0.061
ATLANTIC OCEAN (0–300 m)										
	OHC	HF	VHA	P1	P5	P6	P10	Residual		
OFES1	0.032	0.184	–0.445	–0.823	0.891	–0.085	–0.440	0.749		
OFES2	–0.014	–0.036	0.005	–0.665	0.783	–0.051	–0.388	0.338		
INDIAN OCEAN (0–300 m)										
	OHC	HF	VHA	P1	P2	P3	P7	P8	Residual	
OFES1	0.026	0.195	–0.639	0.823	–0.038	–1.233	0.728	0.162	0.028	
OFES2	0.035	–0.534	–2.091	0.665	–0.012	–1.383	1.365	0.100	1.926	

874

875 *Middle layer*

876 The horizontal and vertical heat transport in the middle layer (300–700 m) of the Pacific Ocean (Tab. 3), estimated by  
 877 OFES1 and OFES2, displayed no significant difference. It can be seen that the ITF was weak for this deeper layer,  
 878 and the differences in the results from OFES1 and OFES2 were small ( $0.084 W/m^2$ ). However, there was heat advected  
 879 or diffused from the upper layer (at 300 m, the top face of the middle ocean layer). There was a difference of  
 880 approximately  $0.747 W/m^2$  in the VHD at a depth of 300 m in the Pacific Ocean and a difference of  $0.701 W/m^2$  in  
 881 the VHA. All these results led us to infer a VHD difference of  $1.295 W/m^2$  at a depth of 700 m in the Pacific Ocean,  
 882 with more heat diffusing downward in OFES1.

883 In the Atlantic Ocean, the average OHC trend, estimated by OFES1, was positive. It was, however, negative in  
 884 OFES2, with a difference of  $0.129 W/m^2$ . A VHA of  $–1.585 W/m^2$  was calculated for OFES2, which was 32% stronger  
 885 than that for OFES1. Additionally, more heat was lost through P1 into the Indian Ocean, and more heat was advected  
 886 into the Atlantic Ocean through the Drake Passage in the OFES1. Differences also existed in the heat advection

887 between the Atlantic Ocean and the Southern Ocean (P6) and the Arctic (P10) Oceans. The vertical heat transport  
 888 (VHA + VHD) at 300 m in the Atlantic Ocean (Tab. 2) was close between the two OFES data. The inferred VHD at  
 889 a depth of 700 m in the Atlantic Ocean was upward in both datasets, although it was stronger by 0.393 W/m<sup>2</sup> in the  
 890 OFES2.

891 The average OHC trend in the Indian Ocean was weakly negative for both OFES1 and OFES2. There was more heat  
 892 (by 0.142 W/m<sup>2</sup>) advected downward at a depth of 700 m in the OFES2. Horizontally, 0.121 W/m<sup>2</sup> more heat was  
 893 acquired from the Atlantic Ocean (through P1) in the OFES1, however, there were only slight differences in the lateral  
 894 heat transport through the other passages connecting the Indian Ocean with other basins. The time-averaged VHD at  
 895 700 m in the Indian Ocean was 0.423 W/m<sup>2</sup> in OFES1 and 1.083 W/m<sup>2</sup> in OFES2.

896 To summarize, the notable cooling trend in the Pacific and Atlantic Oceans (Fig.3), determined using OFES2 was  
 897 mainly generated from vertical heat transport (VHA + VHD) processes. For example, there was a net upward heat  
 898 advection at 300 m in the OFES2 Atlantic Ocean and a stronger downward heat advection at 700 m. As a result, more  
 899 heat was lost vertically in the middle layer of the OFES2 Atlantic Ocean compared to the OFES1 Atlantic Ocean.

900 **Table 3.** As for Tab. 2 but for the middle layer (300–700 m). VHA is at a depth of 700 m in this table.

<b>PACIFIC OCEAN (300–700 m)</b>									
	<b>OHC</b>	<b>VHA</b>	<b>P3</b>	<b>P4</b>	<b>P5</b>	<b>P7</b>	<b>P8</b>	<b>P9</b>	<b>Residual</b>
<b>OFES1</b>	0.017	-0.096	1.208	-0.026	-1.056	0.044	0	0	-1.679
<b>OFES2</b>	-0.034	-0.084	1.247	-0.030	-0.917	-0.040	0	0	-0.384
<b>ATLANTIC OCEAN (300–700 m)</b>									
	<b>OHC</b>	<b>VHA</b>	<b>P1</b>	<b>P5</b>	<b>P6</b>	<b>P10</b>	<b>Residual</b>		
<b>OFES1</b>	0.037	-1.203	-0.770	1.056	0.056	-0.057	1.260		
<b>OFES2</b>	-0.092	-1.585	-0.649	0.917	0.017	-0.102	1.653		
<b>INDIAN OCEAN (300–700 m)</b>									
	<b>OHC</b>	<b>VHA</b>	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P7</b>	<b>P8</b>	<b>Residual</b>	
<b>OFES1</b>	-0.010	-0.519	0.770	-0.043	-1.208	-0.044	0	0.423	
<b>OFES2</b>	-0.013	-0.661	0.649	-0.043	-1.247	0.040	0	1.083	

901  
 902 *Lower layer*  
 903 OFES2 showed cooling in the bottom (700–2000 m) layer of each basin, but OFES1 showed overall warming (Tab.  
 904 4). In the Pacific Ocean, the VHA at 2000 m was downward and had a similar magnitude in the two OFES datasets.  
 905 Owing to the vertical coherence of the ACC, there was intense eastward heat advection through P3 and P5, even below  
 906 700 m, with the OFES2 showing higher advection. The horizontal heat advection through P4 and P7 was relatively  
 907 weak, and it was again larger in OFES2. For example, the MHA passing through P7 was more than one time larger in  
 908 the OFES2. In fact, more heat advected southward into the Indian Ocean through the ITF, which was found in all the  
 909 ocean layers (OFES1 showed a weak northward heat advection in the middle layer). As a result of these differences  
 910 and the estimated VHA and VHD at a depth of 700 m, we calculated a significant difference of approximately 1.252  
 911 W/m<sup>2</sup> in the VHD (in the downward direction) between the two OFES datasets at a depth of 2000 m in the Pacific  
 912 Ocean.

913 Unlike at 2000 m in the Pacific Ocean, OFES2 reflected that there was a significantly stronger downward heat  
 914 advection at 2000 m in the Atlantic Ocean. The dominant horizontal heat advectons were through P1 and P5, with  
 915 the OFES2 showing stronger heat advection at both passages. We estimated a downward heat diffusion at a depth of  
 916 2000 m of 0.216 W/m<sup>2</sup> in the OFES1 Atlantic Ocean and an upward VHD of 0.383 W/m<sup>2</sup> in the OFES2 Atlantic  
 917 Ocean.

918 In the Indian Ocean, the calculated downward heat advection was twice as strong in the OFES1; there were also  
 919 some moderate differences in horizontal heat advection. The resulting VHD at 2000 m was upward in both OFES1  
 920 and OFES2, although it was much greater (by 0.455 W/m<sup>2</sup>) in the latter.

921 In summary, the differences in the lateral heat advection through the major passages P1–P10 in the lower layer were  
 922 small, and the major drivers of the examined OHC differences between OFES1 and OFES2 were generated largely  
 923 from vertical heat transport (VHA + VHD), similar to the situation in the middle layer.

924 **Table 4.** As for Tab. 2 but for the lower layer (700–2000 m). VHA is at a depth of 2000 m.

<b>PACIFIC OCEAN (700–2000 m)</b>									
	<b>OHC</b>	<b>VHA</b>	<b>P3</b>	<b>P4</b>	<b>P5</b>	<b>P7</b>	<b>P8</b>	<b>P9</b>	<b>Residual</b>
<b>OFES1</b>	0.058	-0.126	0.951	-0.04 7	-1.12 0	-0.035	0	0	-1.341
<b>OFES2</b>	-0.037	-0.105	1.146	-0.08 0	-1.29 4	-0.082	0	0	-0.089
<b>ATLANTIC OCEAN (700–2000 m)</b>									
	<b>OHC</b>	<b>VHA</b>	<b>P1</b>	<b>P5</b>	<b>P6</b>	<b>P10</b>	<b>Residual</b>		
<b>OFES1</b>	0.014	-0.029	-0.97 4	1.120	0.066	0.105	-0.216		
<b>OFES2</b>	-0.013	-0.536	-1.05 9	1.294	0.003	-0.031	0.383		
<b>INDIAN OCEAN (700–2000 m)</b>									
	<b>OHC</b>	<b>VHA</b>	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P7</b>	<b>P8</b>	<b>Residual</b>	
<b>OFES1</b>	0.007	-0.241	0.974	-0.03 3	-0.95 1	0.035	0	0.126	
<b>OFES2</b>	-0.018	-0.120	1.059	-0.05 2	-1.14 6	0.082	0	0.581	

925

## 926 **4 Conclusions and Discussion**

927 In this study, we estimated the OHC based on two eddy-resolution hindcast simulations, OFES1 and OFES2, with a  
 928 major focus on estimating their differences. The global observation-based dataset EN4 acted as a reference. The main  
 929 findings of this study are as follows:

930 1. Multi-decadal warming was clearly evinced in most of the global ocean (0–2000 m), especially in the EN4 and  
 931 OFES1 datasets. The warming was dominantly manifested as deepening of the neutral density surfaces (HV  
 932 component), with changes along the neutral surfaces (SP component) of regional importance.

933 2. Significant differences in the OHC (or potential temperature) were found between OFES1 and OFES2; the major  
 934 causes for these were fourfold. First, less surface heating or even cooling applied in OFES2. Second, the ITF was

935 almost one time stronger in the OFES2, especially in the top 300 m. Third, the differences in the intensity of the VHA  
936 were large, particularly at the depth of 300 m in the Indian Ocean. Finally, remarkable differences in the vertical heat  
937 diffusion were inferred.

938 3. Overall, the global and basin-integrated OHC estimates for the period 1960–2016 were reasonable for the top 700  
939 m upon considering the OFES1 results. Below 700 m, multi-decadal climate changes derived from the OFES1 need  
940 careful evaluations even though the estimates of global OHC between 700–2000 m are highly correlated with  
941 observations. The notable differences between OFES2 and EN4 suggest that attention is clearly warranted while  
942 concluding on multi-decadal climate changes based on OFES2.

943 Although we have detailed the OHC differences between the OFES1 and OFES2, and also analyzed the horizontal  
944 and vertical heat transport in an attempt to understand the causes of these differences, further work is required for  
945 improving this field. First, a direct calculation of the VHD is desirable to obtain a more reliable and accurate  
946 comparison between the two OFES. In addition, decomposing the VHD into tidal mixing and mixed-layer vertical  
947 mixing is also an interesting topic, and can help to isolate the effects of tidal mixing in the ocean state. We also expect  
948 to see a detailed comparison of the wind stress from these two datasets over the 57 years. This is inspired by the work  
949 of Kutsuwada et al. (2019) and our detection of the large differences in VHA. Considering the apparent differences in  
950 the SP component among the three datasets, a comprehensive comparison of salinity between both OFES1 and OFES2,  
951 along with observations, was required. This helped the community determine their choice of datasets for their research  
952 purposes.

953 One may argue that the inability to spin-up completely could be the likely cause for the identified differences  
954 between the OFES2 and other datasets since the OFES1 followed a 50-year climatological simulation but OFES did  
955 not. However, large differences between the two OFES datasets can be seen in the temporal evolution of global and  
956 basin OHCs, even during the last two decades. In addition, for example, **S2020** found that the Azores Current was  
957 simulated in the OFES2 in the initial two decades, however, it disappeared after 1970. These, to some extent, weaken  
958 the spin-up argument, although it does not rule out the possibility completely. OFES2 was not expected to be highly  
959 sensitive to the spin-up issue because the starting conditions are from OFES1. There were indeed some improvements  
960 in the OFES2 during the recent decades, for example, from 2005–2016 (not shown here). Two potential explanations  
961 are as follows: First, the model was well spun-up after a couple of decades of integration; second, improvements in  
962 the reanalysis of atmospheric forcing data contributed to improvements in simulation.

963 As mentioned above, results based on EN4 should not be considered as the *truth*. Several factors such as mapping  
964 methods and data assimilated impact the resulting quality of the observational-based product, and might consequently  
965 alter our conclusions. As a preliminary test of robustness, we compared the temporal evolution of the OHC (Fig. S10)  
966 and the spatial patterns of the long-term potential temperature trend (Fig. S11) determined using EN4 and two datasets,  
967 G10 and IAP. G10 is the most up-to-date version of EN4 datasets (EN4.2.2) with bias-corrected following Gouretski  
968 and Reseghetti (2010) and IAP is the dataset from the Institute of Atmospheric Physics (Cheng and Zhu, 2016). The  
969 primary difference between EN4 (bias-corrected following Levitus et al. (2009)) and G10 lies in the bias correction  
970 methods, whereas IAP differs from EN4 in assimilated datasets, mapping methods, and among others. The large  
971 similarities between EN4 and G10 suggest that the different correction methods do not lead to notable differences in

972 the resulting state estimates. On the other hand, there were some differences between the IAP and both EN4 and G10.  
973 This may indicate that the applied mapping method causes some discrepancies among different oceanic products,  
974 which is consistent with Cheng and Zhu (2016). Nonetheless, this preliminary test shows that our primary conclusions  
975 are unlikely to be altered when choosing different observational-based datasets for comparisons.

976 Finally, in absence of any observation-based constraints, the OFES products, especially the OFES1, have captured  
977 some of the warming and cooling trends shown by EN4 and in the literature. However, clear differences between the  
978 two OFES datasets and EN4 suggest the importance of observational data in improving the performance of a hindcast  
979 simulation. The significant differences in the vertical heat diffusion between the two OFES datasets also suggest that  
980 special attention should be given to the validation of the vertical mixing scheme in future ocean modeling.

981  
982 **Author contributions:** F.L conceived the study. All authors contributed to the details of study design. F.L conducted  
983 the calculations and analysis. F.L drafted the manuscript; Z.L and X.H.W improved the writing.

984  
985 **Competing interests:**The authors declare that they have no conflict of interest.

986  
987 **Acknowledgements:** This is publication No. 87 of the Sino-Australian Research Consortium for Coastal Management  
988 (previously the Sino-Australian Research Centre for Coastal Management). This work was supported by the Key  
989 Special Project for Introduced Talents Team of the Southern Marine Science and Engineering Guangdong Laboratory  
990 (Guangzhou; GML2019ZD0210). The authors thank Dr. Peter McIntyre for improving the manuscript. The authors  
991 acknowledge public access to the data used in this paper from the UK Meteorological Office and the JAMSTEC.  
992 Constructive comments from the editor and two anonymous reviewers greatly improved the manuscript.

993  
994 **Code and data availability:** OFES1 and OFES2 are based on the MOM3, available at [https://github.com/mom-](https://github.com/mom-ocean/MOM3)  
995 [ocean/MOM3](https://github.com/mom-ocean/MOM3). Code for decomposing the potential temperature: <http://www.teos-10.org/software.htm>. Original EN4  
996 data: <https://www.metoffice.gov.uk/hadobs/en4/download-en4-2-1.html>. Original OFES1 temperature and salinity  
997 data: [http://apdrc.soest.hawaii.edu/dods/public\\_ofes/OfES/ncep\\_0.1\\_global\\_mmean](http://apdrc.soest.hawaii.edu/dods/public_ofes/OfES/ncep_0.1_global_mmean). Due to a data security incident,  
998 access to the OFES2 data has been temporarily suspended. The data and codes (including the publically available  
999 scripts for completion) needed to reproduce the results of this paper are archived on Zenodo  
1000 (<https://doi.org/10.5281/zenodo.5205444>). The archived data are annual mean values calculated from the original data.

## 1001 **References**

1002 Abraham, J. P., Reseghetti, F., Baringer, M., Boyer, T., Cheng, L., Church, J., Domingues, C., Fasullo, J. T., Gilson,  
1003 J., Goni, G., Good, S., Gorman, J. M., Gouretski, V., Ishii, M., Johnson, G. C., Kizu, S., Lyman, J., MacDonald, A.,  
1004 Minkowycz, W. J., Moffitt, S. E., Palmer, M., Piola, A., Trenberth, K. E., Velicogna, I., Wijffels, S., and Willis, J.: A  
1005 review of global ocean temperature observations: implications for ocean heat content estimates and climate change,  
1006 *Rev. Geophys.*, 51, 450-483, doi.org/10.1002/rog.20022, 2013.

1007

1008 AchutaRao, K. M., Ishii, M., Santer, B. D., Gleckler, P. J., Taylor, K. E., Barnett, T. P., Pierce, D. W., Stouffer, R. J.,  
1009 and Wigley, T. M. L.: Simulated and observed variability in ocean temperature and heat content, *Proc. Natl. Acad.*  
1010 *Sci.*, 104,10768-10773, doi.org/10.1073/pnas.0611375104, 2007.

1011  
1012 Allison, L. C., Roberts, C. D., Palmer, M. D., Hermanson, L., Killick, R. E., Rayner, N. A., Smith, D. M., and Andrews,  
1013 M. B.: Towards quantifying uncertainty in ocean heat content changes using synthetic profiles, *Environ. Res. Lett.*,  
1014 14, 084037, doi.org/10.1088/1748-9326/ab2b0b, 2019.

1015  
1016 Balmaseda, M. A., Trenberth, K. E., and Källén, E.: Distinctive climate signals in reanalysis of global ocean heat  
1017 content, *Geophys. Res. Lett.*, 40, 1754-1759, doi.org/10.1002/grl.50382, 2013.

1018  
1019 Bindoff, N. L., and McDougall, T. J.: Diagnosing climate change and ocean ventilation using hydrographic data, *J.*  
1020 *Phy. Oceanogr.*, 24, 1137-1152, doi.org/10.1175/1520-0485(1994)024<1137:DCCA0V>2.0.CO;2, 1994.

1021  
1022 Carton, J. A., Chepurin, G. A. and Chen, L.: SODA3: A New Ocean Climate Reanalysis, *J. Climate.*, 31, 6967-6983,  
1023 <https://doi.org/10.1175/JCLI-D-18-0149.1>, 2018.

1024  
1025 Banks, H. T., and Gregory, J. M.: Mechanisms of ocean heat uptake in a coupled climate model and the implications  
1026 for tracer based predictions of ocean heat uptake, *Geophys. Res. Lett.*, 33, L07608,  
1027 <https://doi.org/10.1029/2005GL025352>, 2006.

1028  
1029 Carton, J. A., Chepurin, G., A. and Chen, L.: SODA3: A New Ocean Climate Reanalysis, *J. Climate.*, 31, 6967-6983,  
1030 <https://doi.org/10.1175/JCLI-D-18-0149.1>, 2018.

1031  
1032 Carton, J. A., Penny, S. G., and Kalnay, E.: Temperature and salinity variability in the SODA3, ECCO4r3, and ORAS5  
1033 ocean reanalyses, 1993–2015, *J. Climate.*, 32, 2277-2293, doi.org/10.1175/JCLI-D-18-0605.1, 2019.

1034  
1035 Chen, X., Yan, Y., Cheng, X., and Qi, Y.: Performances of seven datasets in presenting the upper ocean heat content  
1036 in the South China Sea, *Adv. Atmos. Sci.*, 30, 1331-1342, doi.org/10.1007/s00376-013-2132-1, 2013.

1037  
1038 Cheng, L., Trenberth, K. E., Palmer, M. D., Zhu, J., and Abraham, J.: Observed and simulated full-depth ocean heat  
1039 content changes for 1970–2005, *Ocean Sci.*, 12, 925-935, doi.org/10.5194/os-12-925-2016, 2016.

1040  
1041 Cheng, L., and Zhu, J.: Artifacts in variations of ocean heat content induced by the observation system changes.  
1042 *Geophys. Res. Lett.*, 41, 7276-7283, <https://doi.org/10.1002/2014GL061881>, 2014.

1043  
1044 Cheng, L., and Zhu, J.: Benefits of CMIP5 Multimodel Ensemble in Reconstructing Historical Ocean Subsurface  
1045 Temperature Variations. *J. Climate.*, 29(15), 5393–5416, <https://doi.org/10.1175/JCLI-D-15-0730.1>, 2016.

1046

1047 Cheng, L., Abraham, J., Goni, G., Boyer, T., Wijffels, S., Cowley, R., Gouretski, V., Reseghetti, F., Kizu, S., Dong,  
1048 S., Bringas, F., Goes, M., Houpert, L., Sprintall, J., and Zhu, J.: XBT Science: Assessment of Instrumental Biases and  
1049 Errors. *Bull. Am. Meteorol. Soc.*, 97(6), 924–933, <https://doi.org/10.1175/BAMS-D-15-00031.1>, 2016.

1050  
1051 Church, J. A., White, N. J., and Arblaster, J. M.: Significant decadal-scale impact of volcanic eruptions on sea level  
1052 and ocean heat content, *Nature*, 438, 74-77, [doi.org/10.1038/nature04237](https://doi.org/10.1038/nature04237), 2005.

1053  
1054 Curry, R., Dickson, B. and Yashayaev, I.: A change in the freshwater balance of the Atlantic Ocean over the past four  
1055 decades. *Nature*, 426, 826-829, <https://doi.org/10.1038/nature02206>, 2003.

1056  
1057 Desbruyères, D., McDonagh, E. L., King, B. A., and Thierry, V.: Global and Full-Depth Ocean Temperature Trends  
1058 during the Early Twenty-First Century from Argo and Repeat Hydrography, *J. Climate.*, 30, 1985-1997,  
1059 [doi.org/10.1175/JCLI-D-16-0396.1](https://doi.org/10.1175/JCLI-D-16-0396.1), 2017.

1060  
1061 Desbruyeres, D., Purkey, S. G., Mcdonagh, E. L., Johnson, G. C. and King, B. A.: Deep and abyssal ocean warming  
1062 from 35 years of repeat hydrography. *Geophys. Res. Lett.*, 43, 10356-10365, [doi.org/10.1002/2016GL070413](https://doi.org/10.1002/2016GL070413), 2016.

1063  
1064 Dong, S., Garzoli, S., and Baringer, M.: The role of interocean exchanges on decadal variations of the meridional heat  
1065 transport in the South Atlantic, *J. Phys. Oceanogr.*, 41, 1498-1511, [doi.org/10.1175/2011JPO4549.1](https://doi.org/10.1175/2011JPO4549.1), 2011.

1066  
1067 Durack, P. J., Gleckler, P. J., Landerer, F. W., and Taylor, K. E.: Quantifying underestimates of long-term upper-  
1068 ocean warming, *Nat. Climate Change.*, 4, 999-1005, <https://doi.org/10.1038/nclimate2389>, 2014.

1069  
1070 Du, Y., Qu, T., Meyers, G., Masumoto, Y., and Sasaki, H.: Seasonal heat budget in the mixed layer of the southeastern  
1071 tropical Indian Ocean in a high-resolution ocean general circulation model, *J. Geophys. Res. Oceans.*, 110, C04012,  
1072 [doi.org/10.1029/2004JC002845](https://doi.org/10.1029/2004JC002845), 2005.

1073  
1074 Emery, W.: Water Types and Water Masses, *Encyclopedia of Ocean Sciences*, 4, 3179-3187,  
1075 [doi.org/10.1006/rwos.2001.0108](https://doi.org/10.1006/rwos.2001.0108), 2001.

1076  
1077 Ernst, W. G.: *Earth systems: processes and issues*. Cambridge University Press, 2000.

1078  
1079 Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., and Wunsch, C.: ECCO version 4: an integrated  
1080 framework for non-linear inverse modeling and global ocean state estimation, *Geosci. Model Dev.*, 8 (10), 3071–3104,  
1081 [doi:10.5194/gmd-8-3071-2015](https://doi.org/10.5194/gmd-8-3071-2015), 2015.

1082  
1083 Fyfe, J.: Southern Ocean warming due to human influence, *Geophys. Res. Lett.*, 33, L19701, [10.1029/2006GL027247](https://doi.org/10.1029/2006GL027247),  
1084 2006.

1085

1086 Gleckler, P. J., Santer, B. D., Domingues, C. M., Pierce, D. W., Barnett, T. P., Church, J. A., Taylor, K. E., Achutarao,  
1087 K., Boyer, T. P., and Ishii, M.: Human-induced global ocean warming on multidecadal timescales, *Nat. Climate*  
1088 *Change.*, 2, 524-529, doi.org/10.1038/nclimate1553, 2012.

1089  
1090 Good, S. A., Martin, M., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and  
1091 monthly objective analyses with uncertainty estimates, *J. Geophys. Res. Oceans.*, 118, 6704-6716,  
1092 doi.org/10.1002/2013JC009067, 2013.

1093  
1094 Gouretski, V., and Reseghetti, F.: On depth and temperature biases in bathythermograph data: Development of a new  
1095 correction scheme based on analysis of a global ocean database. *Deep Sea Res. Part I Oceanogr.*, 57(6), 812-833,  
1096 <https://doi.org/10.1016/j.dsr.2010.03.011>, 2010.

1097  
1098 Häkkinen, S., Rhines, P. B. and Worthen, D. L.: Heat content variability in the North Atlantic Ocean in ocean  
1099 reanalyses, *Geophys. Res. Lett.*, 42, doi.org/10.1002/2015GL063299, 2901-2909, 2015.

1100  
1101 Häkkinen, S., Rhines, P. B., and Worthen, D.: Warming of the global ocean: Spatial structure and water-mass trends,  
1102 *J. Climate.*, 29, 4949-4963, doi.org/10.1175/JCLI-D-15-0607.1, 2016.

1103  
1104 IPCC.: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535pp.,  
1105 doi:10.1017/CBO9781107415324, 2013.

1106  
1107 Jackett, D. R., and McDougall, T. J.: A neutral density variable for the world's oceans, *J. Phys. Oceanogr.*, 27, 237-  
1108 263, doi.org/10.1175/1520-0485(1997)027<0237:ANDVFT>2.0.CO;2, 1997.

1109  
1110 Jayne, S. R., and Laurent, L. C. St.: Parameterizing tidal dissipation over rough topography, *Geophys. Res. Lett.*, 28,  
1111 811-814, doi.org/10.1029/2000GL012044, 2001.

1112  
1113 Johnson, G. C., Purkey, S. G., Zilberman, N. V., and Roemmich, D. Deep Argo Quantifies Bottom Water Warming  
1114 Rates in the Southwest Pacific Basin. *Geophysical Research Letters*, 46(5), 2662-2669,  
1115 <doi.org/10.1029/2018GL081685>, 2019.

1116  
1117 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen,  
1118 J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa,  
1119 A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, *B. Am. Meteorol. Soc.*, 77,  
1120 437-472, doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

1121  
1122 Kutsuwada, K., Kakiuchi, A., Sasai, Y., Sasaki, H., Uehara, K., and Tajima, R.: Wind-driven North Pacific Tropical  
1123 Gyre using high-resolution simulation outputs, *J. Oceanogr.*, 75, 81-93, 10.1007/s10872-018-0487-8, 2019.

1124



1125 Large, W. G., McWilliams, J. C., and Doney, S. C.: Oceanic vertical mixing: A review and a model with a nonlocal  
1126 boundary layer parameterization, *Rev. Geophys.*, 32, 363-403, doi.org/10.1029/94RG01872, 1994.

1127  
1128 Lee, S., Park, W., Baringer, M. O. A., Gordon, L., Huber, B. A., and Liu, Y.: Pacific origin of the abrupt increase in  
1129 Indian Ocean heat content during the warming hiatus, *Nature Geosci.*, 8, 445-449, doi.org/10.1038/ngeo2438, 2015.

1130  
1131 Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O., Garcia, H. E., Locarnini, R. A., Mishonov, A. V., Reagan, J.  
1132 R., Seidov, D., and Yarosh, E. S.: World ocean heat content and thermosteric sea level change (0–2000 m), *Geophys.*  
1133 *Res. Lett.*, 39, 1955-2010, doi.org/10.1029/2012GL051106, 2012.

1134  
1135 Liang, X., Piecuch, C. G., Ponte, R. M., Forget, G., Wunsch, C., and Heimbach, P.: Change of the global ocean vertical  
1136 heat transport over 1993–2010, *J. Climate.*, 30, 5319-5327, doi.org/10.1175/JCLI-D-16-0569.1, 2017.

1137  
1138 Liang, X., Liu, C. R., Ponte, M. and Chambers, D. P.: A Comparison of the Variability and Changes in Global Ocean  
1139 Heat Content from Multiple Objective Analysis Products During the Argo Period, *J. Climate.*, 1-47,  
1140 doi.org/10.1175/JCLI-D-20-0794.1, 2021.

1141  
1142 Liu, C., Liang, X., Chambers, D. P. and Ponte, R. M.: Global Patterns of Spatial and Temporal Variability in Salinity  
1143 from Multiple Gridded Argo Products, *J. Climate.*, 33, 8751-8766, <https://doi.org/10.1175/JCLI-D-20-0053.1>, 2020.

1144  
1145 Liu, M., and T. Tanhua.: Water masses in the Atlantic Ocean: characteristics and distributions, *Ocean Sci.*, 17, 463-  
1146 486, doi.org/10.5194/os-17-463-2021, 2021.

1147  
1148 Noh, Y., and Kim, H. J.: Simulations of temperature and turbulence structure of the oceanic boundary layer with the  
1149 improved near-surface process, *J. Geophys. Res. Oceans.*, 104, 15621-15634, doi.org/10.1029/1999JC900068, 1999.

1150  
1151 O'Connor, B. M., Fine, R. A. and Olson, D. B.: A global comparison of subtropical underwater formation rates, *Deep*  
1152 *Sea Research Part I: Oceanographic Research Papers*, 52, 1569-1590, doi.org/10.1016/J.DSR.2005.01.011, 2005.

1153  
1154 Palmer, M. D., Mcneall, D. J., and Dunstone, N. J.: Importance of the deep ocean for estimating decadal changes in  
1155 Earth's radiation balance, *Geophys. Res. Lett.*, 38, L13707, doi.org/10.1029/2011GL047835, 2011.

1156  
1157 Pierce, D. W., Barnett, T. P., Achutarao, K., Gleckler, P. J., Gregory, J. M., and Washington, W. M.: Anthropogenic  
1158 warming of the oceans: Observations and model results, *J. Climate.*, 19, 1873-1900, doi.org/10.1175/JCLI3723.1,  
1159 2006.

1160  
1161 Sasaki, H., Sasai, Y., Kawahara, S., Furuichi, M., Araki, F., Ishida, A., Yamanaka, Y., Masumoto, Y., and Sakuma,  
1162 H.: A series of eddy-resolving ocean simulations in the world ocean-OFES (OGCM for the Earth Simulator) project,  
1163 *Oceans '04 MTS/IEEE Techno-Ocean '04 (IEEE Cat. No. 04CH37600) 3*, 1535-1541, 2004.

1164

1165 Sasaki, H., Kida, S., Furue, R., Nonaka, M., and Masumoto, Y.: An Increase of the Indonesian Throughflow by  
1166 Internal Tidal Mixing in a High-Resolution Quasi-Global Ocean Simulation, *45*, 8416–8424, 2018.  
1167 doi.org/10.1029/2018GL078040.

1168  
1169 Sasaki, H., Kida, S., Furue, R., Aiki, H., Komori, N., Masumoto, Y., Miyama, T., Nonaka, M., Sasai, Y., and Taguchi,  
1170 B.: A global eddying hindcast ocean simulation with OFES2, *Geosci. Model Dev.*, *13*, 3319–3336,  
1171 doi.org/10.5194/gmd-13-3319-2020, 2020.

1172  
1173 Smith, D. M., Allan, R.P., Coward, A.C., Eade, R., Hyder, P., Liu, C., Loeb, N.G., Palmer, M.D., Roberts, C.D. and  
1174 Scaife, A.A.: Earth's energy imbalance since 1960 in observations and CMIP5 models, *Geophys. Res. Lett.*, *42*, 1205-  
1175 1213, doi.org/10.1002/2014GL062669, 2015.

1176  
1177 Spence, P., Saenko, O. A., Sijp, W., and England, M.: The role of bottom pressure torques on the interior pathways of  
1178 North Atlantic deep water, *J. Phys. Oceanogr.*, *42*, 110-125, doi.org/10.1175/2011JPO4584.1, 2012.

1179  
1180 St. Laurent, L. C., Simmons, H. L., and Jayne, S. R.: Estimating tidally driven mixing in the deep ocean, *Geophys.*  
1181 *Res. Lett.*, *29*, 21–21–21–24, doi.org/10.1029/2002GL015633, 2002.

1182  
1183 Talley, L. D.: Shallow, Intermediate, and Deep Overturning Components of the Global Heat Budget, *J. Phys.*  
1184 *Oceanogr.*, *33*, 530-560, [https://doi.org/10.1175/1520-0485\(2003\)033<0530:SIADOC>2.0.CO;2](https://doi.org/10.1175/1520-0485(2003)033<0530:SIADOC>2.0.CO;2), 2003.

1185  
1186 Trenberth, K. E., Fasullo, J. T., Von Schuckmann, K., and Cheng, L.: Insights into Earth's energy imbalance from  
1187 multiple sources, *J. Climate.*, *29*, 7495-7505, doi.org/10.1175/JCLI-D-16-0339.1, 2016.

1188  
1189 Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., Danabasoglu, G., Suzuki, T., Bamber,  
1190 J. L., Bentsen, M., Böning, C. W., Bozec, A., Chassignet, E. P., Curchitser, E., Boeira Dias, F., Durack, P. J., Griffies,  
1191 S. M., Harada, Y., Ilicak, M., Josey, S. A., Kobayashi, C., Kobayashi, S., Komuro, Y., Large, W. G., Le Sommer, J.,  
1192 Marsland, S. J., Masina, S., Scheinert, M., Tomita, H., Valdivieso, M., and Yamazaki, D.: JRA-55 based surface  
1193 dataset for driving ocean-sea-ice models (JRA55-do), *Ocean Model.*, *130*, 79-139,  
1194 doi.org/10.1016/j.ocemod.2018.07.002, 2018.

1195  
1196 Von Schuckmann, K., Palmer, M. D., Trenberth, K. E., Cazenave, A., Chambers, D. P., Champollion, N., Hansen, J.,  
1197 Josey, S. A., Loeb, N. G., and Mathieu, P. P.: An imperative to monitor Earth's energy imbalance, *Nat. Climate*  
1198 *Change.*, *6*, 138-144, doi.org/10.1038/nclimate2876, 2016.

1199  
1200 Wang, G., Cheng, L., Abraham, J., and Li, C.: Consensuses and discrepancies of basin-scale ocean heat content  
1201 changes in different ocean analyses, *Clim. Dyn.*, *50*, 2471-2487, doi.org/10.1007/s00382-017-3751-5, 2018.

1202  
1203 Wang, X. H., Bhatt, V., and Sun, Y.-J.: Study of seasonal variability and heat budget of the East Australian Current  
1204 using two eddy-resolving ocean circulation models, *Ocean. Dyn.*, *63*, 549-563, doi.org/10.1007/s10236-013-0605-5,  
1205 2013.

1206  
1207 Wunsch, C.: The decadal mean ocean circulation and Sverdrup balance, *J. Mar. Res.*, 69, 417-434, doi.org/  
1208 10.1357/002224011798765303, 2011.  
1209  
1210 Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., and Heimbach, P.: Global reconstruction of historical ocean heat  
1211 storage and transport, *Proc. Natl. Acad. Sci.*, 116, 1126-1131, doi.org/10.1073/pnas.1808838115, 2019.  
1212  
1213 Zhang, Y., Feng, M., Du, Y. H., Phillips, E., Bindoff, N. L., and McPhaden, M. J.: Strengthened Indonesian  
1214 Throughflow Drives Decadal Warming in the Southern Indian Ocean, *Geophys. Res. Lett.*, 45, 6167-6175,  
1215 doi.org/10.1029/2018GL078265, 2018.  
1216  
1217