scientific data

OPEN

DATA DESCRIPTOR

Global climatological data of ocean thermohaline parameters derived from WOA18

Peter C. Chu ✉ & Chenwu Fan

This is a global ocean climatological dataset of 17 thermohaline parameters such as isothermal layer (ITL) depth (h_T), mixed layer (ML) depth (h_D), thermocline gradient (G_T), pycnocline gradient (G_D) determined from temperature (T) and salinity (S) profiles of the National Centers for Environmental Information (NCEI) world ocean atlas 2018 (WOA18) using the double gradient method along with the identity-index (i-index) showing the quality of the determination. With the identified (h_T, h_D) and (G_T, G_D), other parameters such as ITL heat content (H_{ITL}), mixed layer fresh-water content (F_{ML}), maximum thermocline gradient (G_{Tmax}), thermocline depth (h_{th}), temperature at thermocline depth (T_{th}), maximum density gradient G_{Dmax}), pycnocline depth (h_{pyc}), density at pycnocline depth (ρ_{pyc}), and ($h_T - h_D$) (barrier layer if positive or compensated layer if negative). The dataset is located at the NOAA/NCEI website (https://doi.org/10.25921/j3v2-jy50). It provides useful background information for ocean mixed layer dynamics, air-sea interaction, climatological studies. This paper is the only document for the dataset.

Background & Summary

Ocean has upmost layer with near-zero vertical gradient such as isothermal layer (ITL) for temperature (T), or mixed layer (ML) for density (ρ) and underneath layer with strong vertical gradient such as thermocline (for T), or pycnocline (for ρ)[1]. Temperature and salinity (S) are observed in oceanography. Usually, the Thermodynamic Equation of Seawater-2010 (https://www.teos-10.org/) is used to compute ρ from (T, S) data. Thermocline or pycnocline with strong vertical gradient is the transition layer between the vertically quasi-uniform layer from the surface (ITL or ML), and the deep-water layer. The vertically quasi-uniform ITL and ML are caused by intense turbulent mixing near the ocean surface. Such mixing is driven by shear due to surface wind stress and by convection due to heat loss from ocean to atmosphere. The thermocline (pycnocline) resists the turbulent mixing from the ITL (or ML) and limits the ITL (or ML) deepening and heat (or water mass) exchange between the ITL (or ML) and deeper layer due to its strong vertical gradient. The thermocline gradient G_T (or pycnocline gradient G_D) directly affects such exchange[2–4].

The ITL and ML provide dynamic-thermodynamic links and mediates the exchange of momentum, heat, and moisture between the atmosphere and the oceans; and hence plays a key role to affect weather and climate. Such exchange depends on an important parameter, i.e., the ocean isothermal layer depth (ILD) h_T [or mixed layer depth (MLD) h_D)], which determines the heat content (or freshwater content) and mechanical inertia of the layer. Temporal variability of h_T (or h_D) is caused by many processes occurring in the isothermal (mixed) layer such as surface forcing, lateral advection, internal waves, etc., ranging from diurnal, seasonal, to interannual variability[2–4]. Spatial variability of h_T (or h_D) is evident from less than 20 m in summer to more than 500 m in winter in subpolar latitudes[5]. Therefore, determination of the four parameters (h_T, h_D, G_T, G_D) from (T, ρ) profiles becomes important. Note that the two depths (h_T, h_D) are not necessary the same with the occurrence of barrier layer if ($h_T > h_D$)[6–15] (Fig. 1a), and compensated layer if ($h_T < h_D$)[15].

Let a T-profile (or ρ-profile) starting from the ocean surface down to depth z_k be represented by $T(z_k)$ [or $\rho(z_k)$]. $k = 1, 2, \ldots, K$, with z_1 the surface and z_K the bottom of the profile. The corresponding vertical gradient is represented by $G_T(z_k)$ [or $G_D(z_k)$]. The vertical gradient for upper ocean ITL-thermocline (or ML-pycnocline) is depicted by

Department of Oceanography, Naval Postgraduate School, Monterey, CA, 93943, USA. ✉e-mail: pcchu@nps.edu

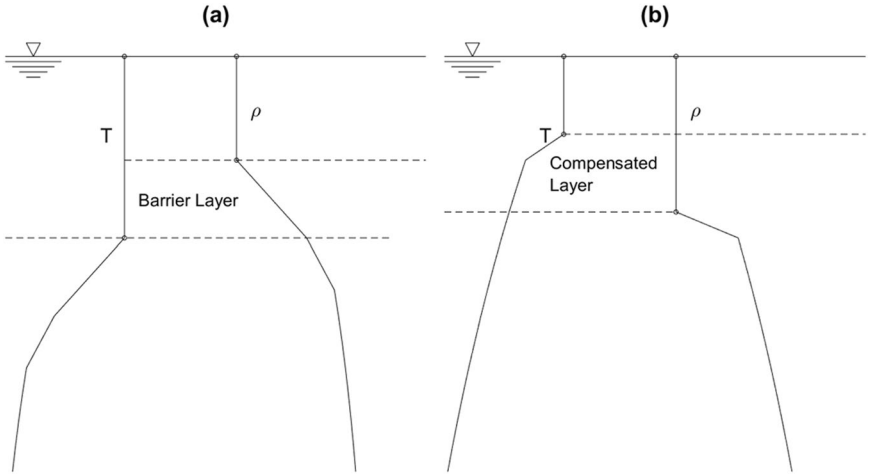

Fig. 1 Schematic illustration of (T, ρ) profiles for occurrence of (**a**) barrier layer when ILD > MLD, and (**b**) compensated layer when ILD < MLD.

$$\begin{cases} G_T(z_k) \approx 0 \text{ for } z_k > -h_T \text{ (in ITL)}, & G_D(z_k) \approx 0 \text{ for } z_k > -h_D \text{ (in ML)} \\ G_T(z_k) \text{ or } G_D(z_k) \text{ evident for } z_k < -h_T \text{ (in thermocline) or } z_k < -h_D \text{ (in pycnocline)} \end{cases} \quad (1)$$

Two types of methodology, single near-zero gradient and double gradients, are available to determine h_T and h_D from vertical T and ρ profiles. The single near-zero gradient method requires either the deviation of T (or ρ) from its value near the surface (i.e., reference level) to be smaller than a certain fixed value, such as 0.8 °C[9], 0.2 °C[15,16] to 0.1 °C[17] for T, (0.03 kg/m^3)[15], (0.125 kg/m^3)[18], to (0.05 kg/m^3)[19] for ρ, or the near-zero vertical gradient to be smaller than a certain fixed value, such as (0.025 °C/m)[6,10], (0.02 °C/m)[20], to (0.015 °C/m)[21] for $G_T(z_k)$. To eliminate or reduce such uncertainty, the split-and-merge (SM)[22], maximum curvature[23–25], optimal linear fitting[26], and maximum angle[27] methods have been developed. The double gradient method is based on the transition of a *near-zero* gradient in the ITL (or ML) to an evident gradient in the thermocline (or pycnocline). The double gradient method with the exponential leap-forward gradient (ELG)[28,29] was used in this study.

After (h_T, h_D) are determined from individual (T, ρ) profiles, the barrier (or compensated) layer depth, ocean heat content (OHC) for ITL (H_{ITL})[30], freshwater content (FWC) for ML (F_{ML}) were calculated. Note that H_{ITL} (or F_{ML}) is the vertical integration of temperature (salinity) profile from the surface (z = 0) down to the base of the ITL ($z = -h_T$) [or ML ($z = -h_D$)] rather than to fixed depths such as (H_{700}, F_{700}) for the upper 700 m. Obviously, H_{ITL} and F_{ML} are new and different from traditionally defined OHC (or FWC) in the oceanographic community with fixed depth intervals from the ocean surface such as 0–150 m (in the Indian Ocean)[31], 0–300 m[32], 0–400 m[33], 0–700 m[34], 0–750 m[35], 0–2000 m[36], deep layer OHC such as below 2000 m[37], and the full layer OHC[38]. Interested readers are referred to two excellent review papers[39,40].

From vertical gradient data below ITL (or ML) $G_T(z_k)$, $z_k < -h_T$ [$G_D(z_k)$, $z_k < -h_D$], other thermohaline parameters can be identified from each T-profile (ρ-profile), such as maximum temperature (density) gradient, thermocline (pycnocline) depth, temperature (density) at thermocline (pycnocline) depth. The quality of the (h_T, h_D) data are estimated by the identification index[29] (i-index, see Error Estimation Section). Altogether, this dataset, containing 17 thermohaline parameters with i-index, is located at the NCEI website (Data Citation 1) for public use.

Global climatological (annual mean and monthly mean) data of ocean thermohaline parameters can be established through two approaches: (1) analysing climatological (T, S) profiles such as earlier version of WOA18 to obtain climatological (h_T, h_D) data[5,9,13,41], and (2) analysing observational profiles to get synoptic thermohaline parameters[29,30] and then using optimal interpolation[42], Kalmen filter[43], or optimal spectral decomposition[44] to produce gridded climatological thermohaline parameters. At present, we cannot estimate how big the difference is if using these two approaches. We can estimate only after the two approaches have been used. In this study, we take the first approach to derive climatological thermohaline parameter data from the NOAA/NCEI World Ocean Atlas 2018 (WOA18)[45] annual and monthly mean temperature and salinity (S) profiles with regular 102 vertical levels (Table 1).

Furthermore, on the base of the single gradient method (i.e., near-zero gradient in the mixed layer), serval global climatological datasets of ocean (h_T, h_D) were produced from the earlier version of WOA18 such as the NOAA/NCEI Mixed Layer Depth Data[5], the Naval Research Laboratory Mixed Layer Depth (NMLD) Climatologies[41], and comprehensive Mixed Layer Data for the Indian Ocean[13]. All these datasets don't include any parameters below the ITL/ML. On the base of the double-gradient method, this dataset[45] contains more thermohaline parameters from WOA18 such as thermocline gradient (G_T), pycnocline gradient (G_D), ITL heat content (H_{ITL}), mixed layer fresh-water content (F_{ML}), maximum thermocline gradient (G_{Tmax}), thermocline depth (h_{th}), temperature at thermocline depth (T_{th}), maximum density gradient G_{Dmax}), pycnocline depth (h_{pyc}), density at pycnocline depth (ρ_{pyc}), and ($h_T - h_D$) (barrier layer if positive or compensated layer if negative).

Standard Level	Standard Depth (m)	Standard Level	Standard Depth (m)	Standard Level	Standard Depth (m)	Standard Level	Standard Depth (m)
1	0	27	250	53	1300	79	3200
2	5	28	275	54	1350	80	3300
3	10	29	300	55	1400	81	3400
4	15	30	325	56	1450	82	3500
5	20	31	350	57	1500	83	3600
6	25	32	375	58	1550	84	3700
7	30	33	400	59	1600	85	3800
8	35	34	425	60	1650	86	3900
9	40	35	450	61	1700	87	4000
10	45	36	475	62	1750	88	4100
11	50	37	500	63	1800	89	4200
12	55	38	550	64	1850	90	4300
13	60	39	600	65	1900	91	4400
14	65	40	650	66	1950	92	4500
15	70	41	700	67	2000	93	4600
16	75	42	750	68	2100	94	4700
17	80	43	800	69	2200	95	4800
18	85	44	850	70	2300	96	4900
19	90	45	900	71	2400	97	5000
20	95	46	950	72	2500	98	5100
21	100	47	1000	73	2600	99	5200
22	125	48	1050	74	2700	100	5300
23	150	49	1100	75	2800	101	5400
24	175	50	1150	76	2900	102	5500
25	200	51	1200	77	3000		
26	225	52	1250	78	3100		

Table 1. Standard vertical depths of WOA18 data.

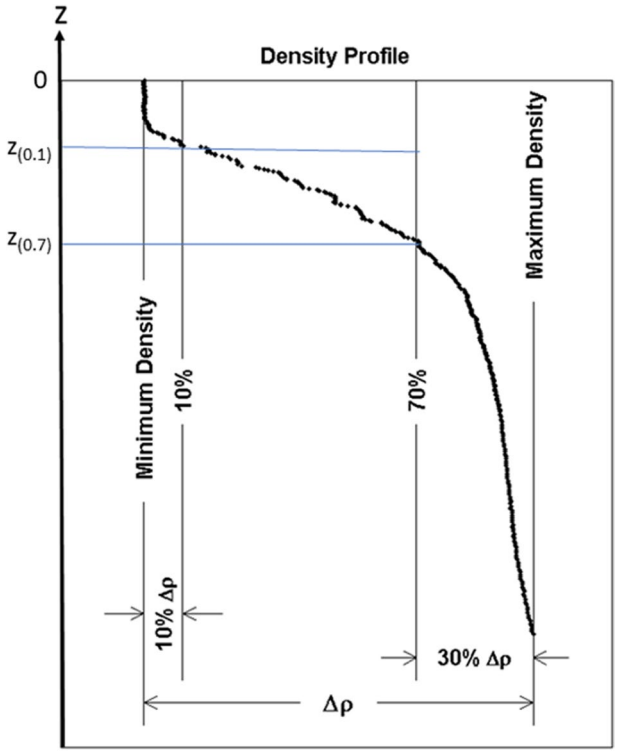

Fig. 2 Illustration for determination of $z_{(0.1)}$ and $z_{(0.7)}$ for ρ-profile.

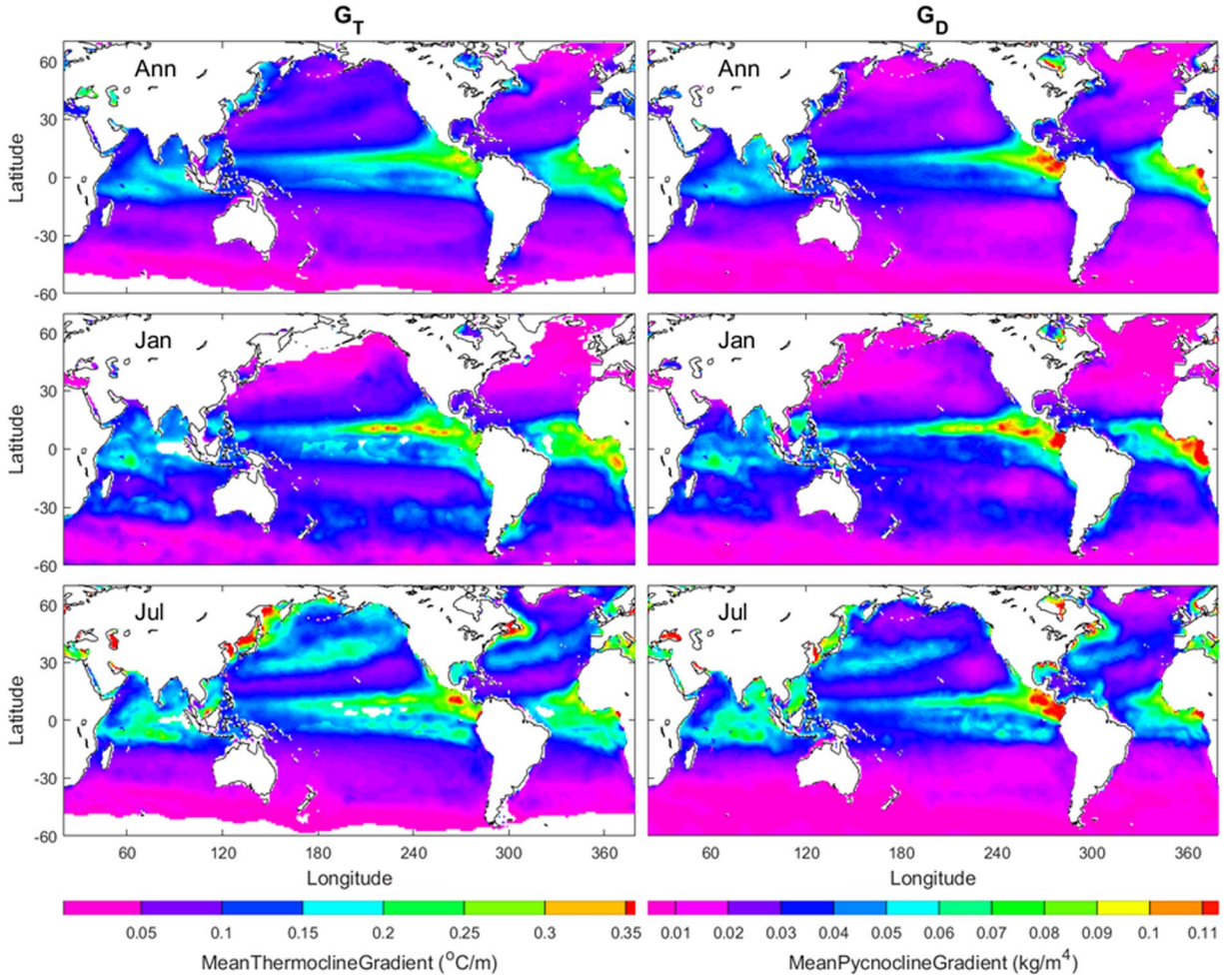

Fig. 3 Annual, January, and July mean (G_T, G_D) maps with the left panels for G_T and the right panels for G_D. The areas in open oceans with the white colour indicate low quality of the identification. The land and the areas of low quality of the identification are represented by white with the land enclosed by black curves (i.e., coasts).

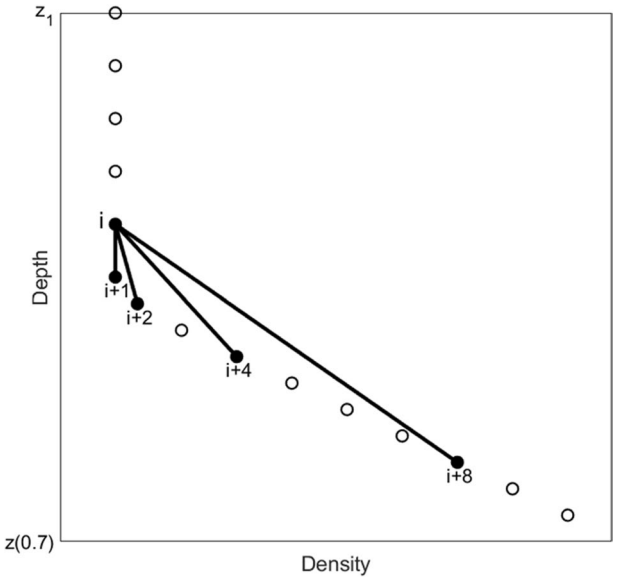

Fig. 4 Illustration of the exponentially leap-forward gradient (ELG) method.

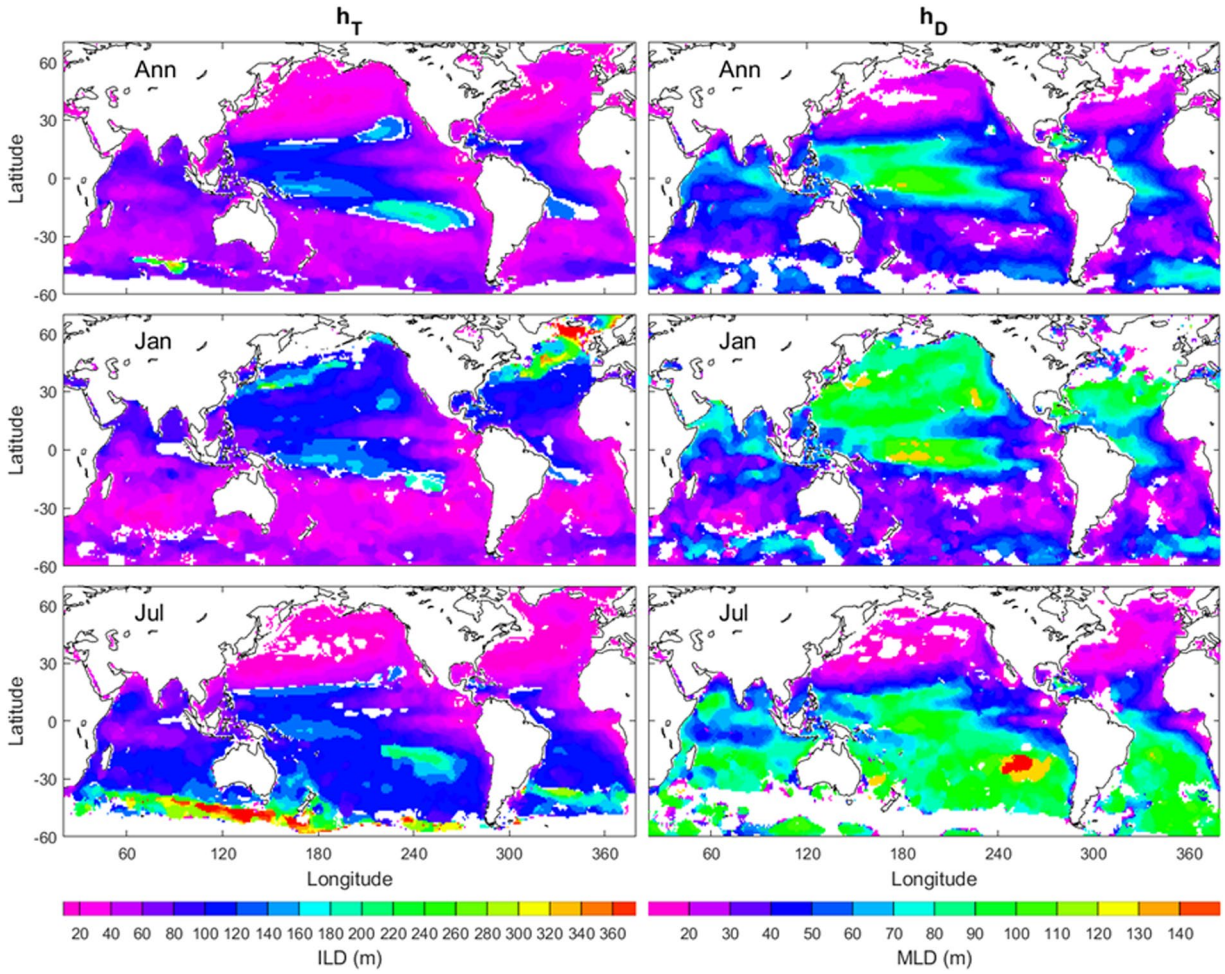

Fig. 5 Annual, January, and July mean (h_T, h_D) maps with the left panels for h_T and the right panels for h_D. The areas in open oceans with the white colour indicate low quality of the identification. The land and the areas of low quality of the identification are represented by white with the land enclosed by black curves (i.e., coasts). The red bull-eye in the southeast Pacific in the right-lower panel indicates the area of $h_D = 150$ m.

Among them, ITL heat content (H_{ITL}), and mixed layer fresh-water content (F_{ML}) are different from the commonly used fixed-depth heat and freshwater contents.

Methods

These methods can be found in our related work[28,29].

Main part of pycnocline (or Thermocline). The ρ-profile [$\rho(z_k)$] is taken for illustration. Let the depths corresponding to $\rho_{\min}$ and ρ_{max} be z_1 and z_K. The vertical density difference, $\Delta\rho = \rho_{max} - \rho_{\min}$, represents the total variability. Theoretically, the variability is 0 in ML and large in pycnocline beneath the ML. It is reasonable to identify the main part of the pycnocline between the two depths: $z_{(0.1)}$ and $z_{(0.7)}$, with the difference to $\rho_{\min}$ as $0.1\Delta\rho$ and $0.7\Delta\rho$ (Fig. 2), respectively. Here, $\rho(z_{(0.1)}) = \rho_{\min} + 0.1\Delta\rho$, and $\rho(z_{(0.7)}) = \rho_{\min} + 0.7\Delta\rho$.

Pycnocline and thermocline gradients. The data between $z_{(0.1)}$ and $z_{(0.7)}$ is rearranged into [$\rho_i, i = 0, 1, 2, \ldots, I$] with [$\rho_0 = \rho(z_{(0.1)})$, $\rho_I = \rho(z_{(0.7)})$]. Vertical gradients are calculated between ρ_i ($i = 1, 2, \ldots, I$) and ρ_0,

$$G_{Di} = -\frac{\rho(z_{(0.1)}) - \rho_i}{z_{(0.1)} - z_i}, \quad \rho_i = \rho(z_i), i = 1, 2, \ldots, I \tag{2}$$

Their median

$$G_D = \text{Median}\{G_{D1}, G_{D2}, \ldots, G_{DI}\} \tag{3}$$

is used to represent the *characteristic gradient* for the pycnocline, and it is simply called the pycnocline gradient (G_D). Similarly, the same procedure id used to obtain the thermocline gradient (G_T). The annual mean (G_T, G_D) maps

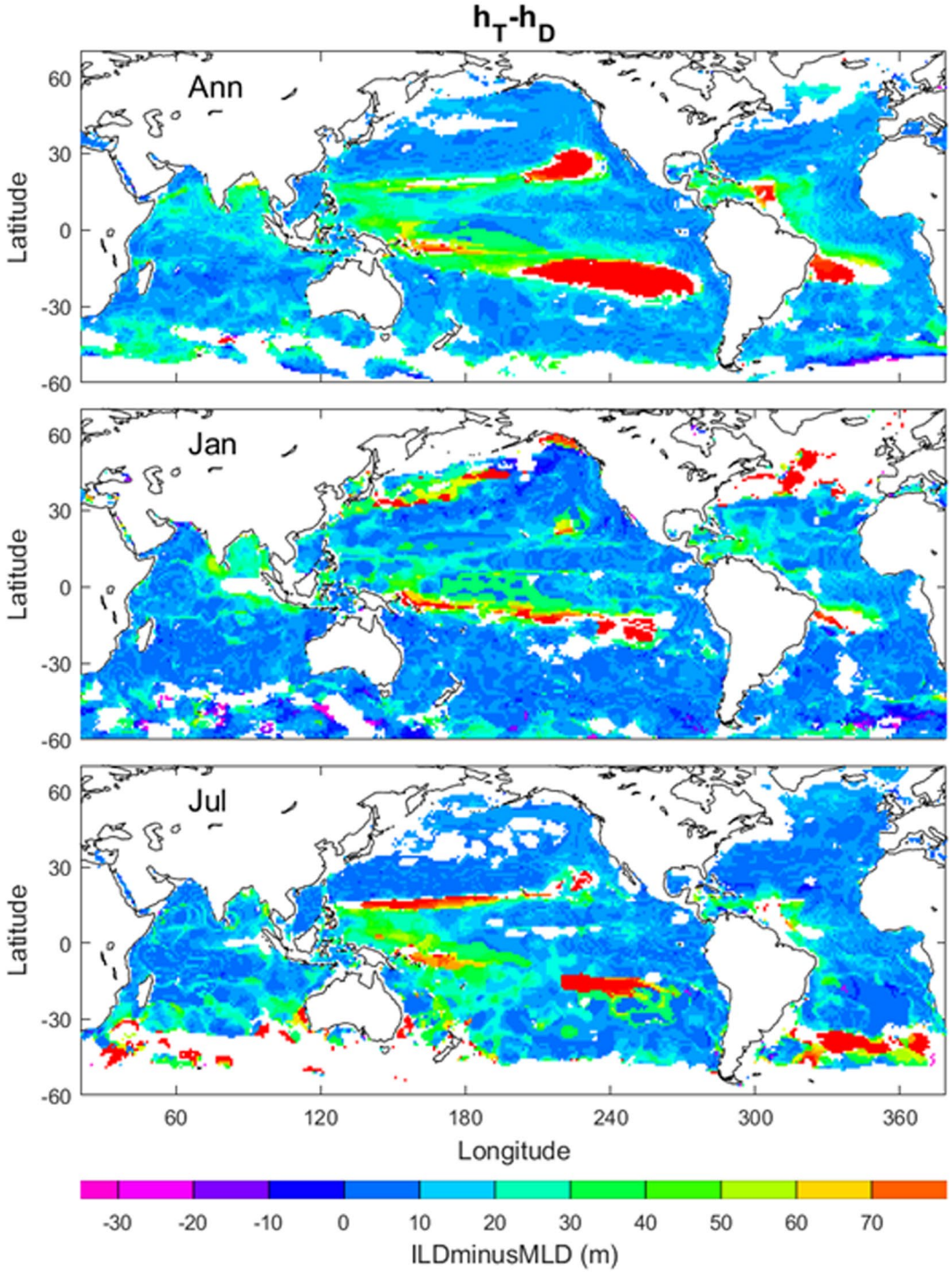

Fig. 6 Annual, January, July mean ($h_T - h_D$) maps. The areas in open oceans with the white colour indicate low quality of the identification. The land and the areas of low quality of the identification are represented by white with the land enclosed by black curves (i.e., coasts).

show gradients are stronger in tropical regions (20°S–20°N) than middle and high latitudes. In the low latitudes, the gradients are stronger in the eastern than western Pacific and Atlantic. The January and July mean (G_T, G_D) maps show stronger seasonal variability in the northern hemisphere than in the southern hemisphere (Fig. 3).

ELG for determining MLD (ILD). Go back to the original profile shown in Fig. 2. Let the number of the data points between z_1 and $z_{(0.7)}$ be N_g, and let $N = [\log_2(N_g)]$ with the bracket indicating the integer part of the real number inside. N is much smaller than N_g. Starting from z_1, the $(N+1)$ exponential leap-forward gradients (ELGs) are calculated at depth z_k [between z_1 and $z_{(0.7)}$] (Fig. 4)

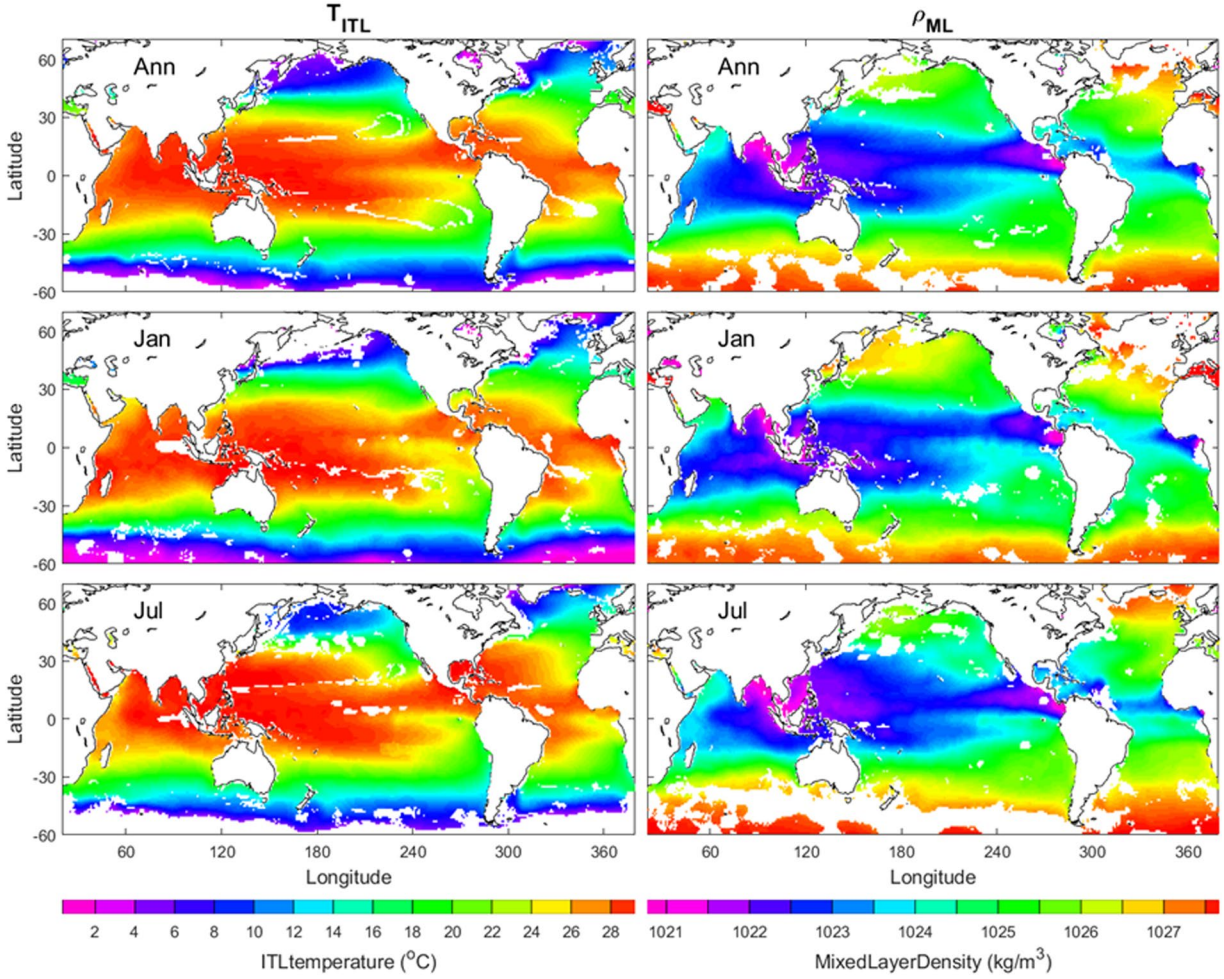

Fig. 7 Annual, January, July mean (T_{ITL}, ρ_{ML}) with the left panels for T_{ITL} and the right panels for ρ_{ML}. The land and the areas of low quality of the identification are represented by white with the land enclosed by black curves (i.e., coasts).

$$D_n\rho(z_k) = \frac{\rho(z_k) - \rho(z_{k+2^n})}{z_k - z_{k+2^n}},\ n = 0, 1, 2, \ldots, N \tag{4}$$

where D_n is the difference operator: $D_1\rho(z_k) = [\rho(z_k) - \rho(z_{k+2})]/(z_k - z_{k+2})$, $D_2\rho(z_k) = [\rho(z_k) - \rho(z_{k+4})]/(z_k - z_{k+4})$, etc. . The averaged value among $(N+1)$ gradients $[D_0\rho(z_k), D_1\rho(z_k), \ldots, D_N\rho(z_k)]$ is computed by

$$\widetilde{G}_D(z_k) = -\frac{\sum_{n=0}^{N} D_n T(z_k)}{N+1} \tag{5}$$

which represents the gradient effectively at the depth z_k with capability to filter out noises in the gradient calculation[28,29].

Since $\widetilde{G}_D(z_k) \approx 0$ if z_k in the mixed layer; $\widetilde{G}_D(z_k) = G_{pyc}$ if z_k in the pycnocline, it is reasonable to use (an order of smaller gradient in mixed layer than in pycnocline),

$$\widetilde{G}_D(z_k)/G_{pyc} < 0.1, \quad z_k \text{ in mixed layer} \tag{6}$$

to identify h_D (similarly h_T). The annual mean (h_T, h_D) maps (upper panels in Fig. 5) show deeper ILD and MLD in the western than eastern Pacific and Atlantic in low latitudes (30°S–30°N). The January and July mean (h_T, h_D) maps (middle and lower panels in Fig. 5) show stronger seasonal variability with deeper (h_T, h_D) in January (July) in the northern (southern) hemisphere, deeper (h_T, h_D) in the Gulf Stream and Kuroshio areas in January, deeper (h_T, h_D) in east of Southern Pacific in July, and deeper h_T in the Circumpolar current west of the Drake Passage in July.

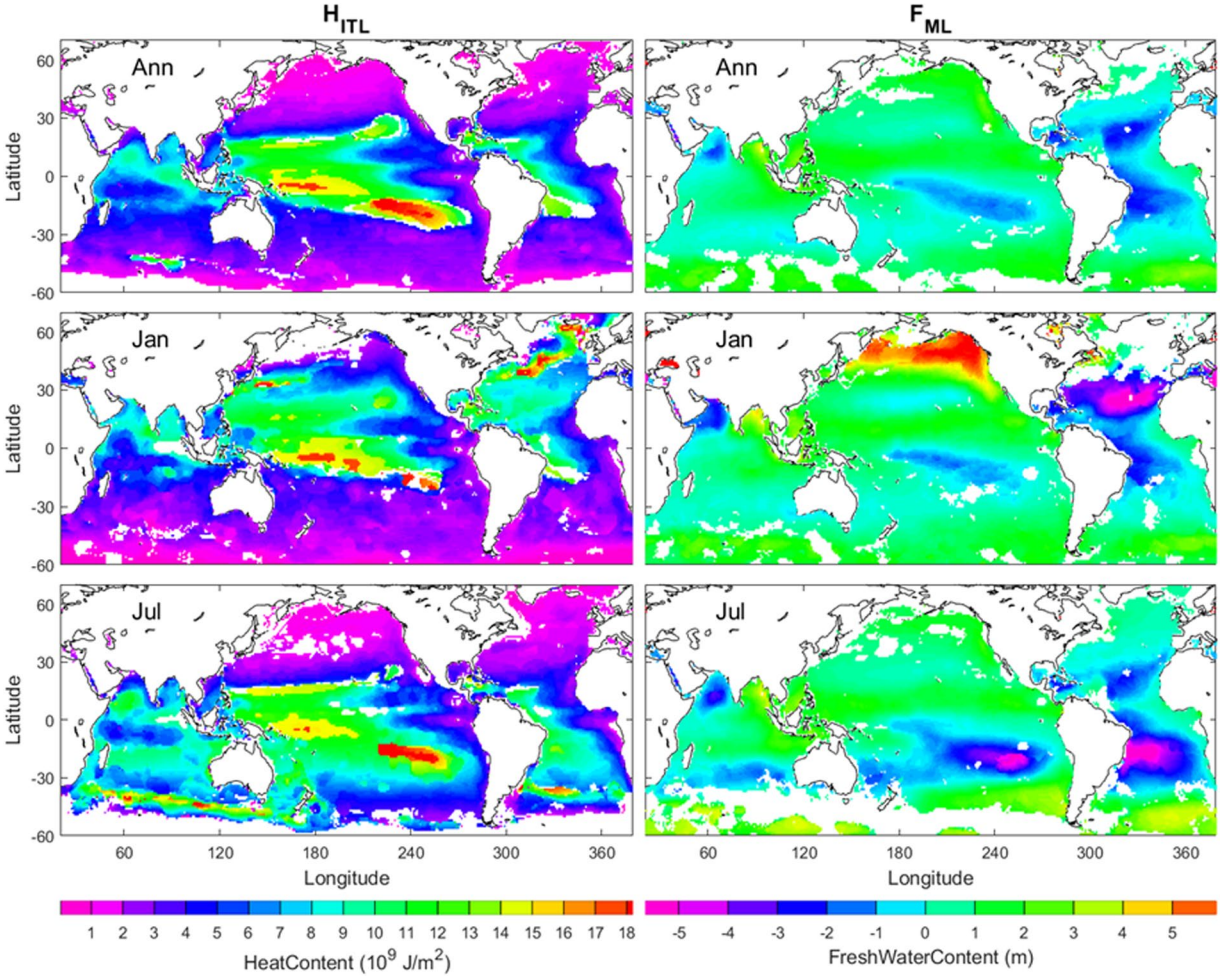

Fig. 8 Annual, January, July mean (H_{ITL}, F_{ML}) maps with the left panels for H_{ITL} and the right panels for F_{ML}. The land and the areas of low quality of the identification are represented by white with the land enclosed by black curves (i.e., coasts).

Barrier (Compensated) layer. Barrier layer occurs if $h_T > h_D$, with the barrier layer depth (BLD) of $(h_T - h_D)$. Compensated layer occurs if $h_T < h_D$, with the compensated layer depth (CLD) of $(h_D - h_T)$[15]. We generate the difference $(h_T - h_D)$ data to identify the occurrence of barrier or compensated layer, with annual, January, July mean $(h_T - h_D)$ maps. The occurrence of barrier layer is much often than the occurrence of compensated layer with the ratio of 23562/170 (~139) from the annual data. Evident barrier layer occurs in the extra-tropical (20°–30°N, 10°–28°S) eastern Pacific Ocean, tropical (0°–8°S) western Pacific west of New Guinea, and low latitudinal Brazilian coast in the annual mean; Kuroshio and Gulf Stream extension regions in January; and south (30°–40°S) Southern Atlantic Ocean in July (Fig. 6).

ITL temperature and ML density. The ITL temperature is the vertically averaged temperature from the sea surface ($z = 0$) down to ILD ($z = -h_T$),

$$T_{ITL} = \frac{1}{h_T} \int_{-h_T}^{0} T dz \tag{7}$$

The ML density is the vertically averaged ρ from the sea surface ($z = 0$) down to MLD ($z = -h_D$)

$$\rho_{ML} = \frac{1}{h_D} \int_{-h_D}^{0} \rho dz \tag{8}$$

Figure 7 shows the annual, January, July mean (T_{ITL}, ρ_{ML}). Note that (T_{ITL}, ρ_{ML}) are not the same as the sea surface temperature and density. The variables (T_{ITL}, ρ_{ML}) follow the ocean mixed layer dynamics. For example, T_{ITL} satisfies following equation due to the ITL layer heat balance[46]

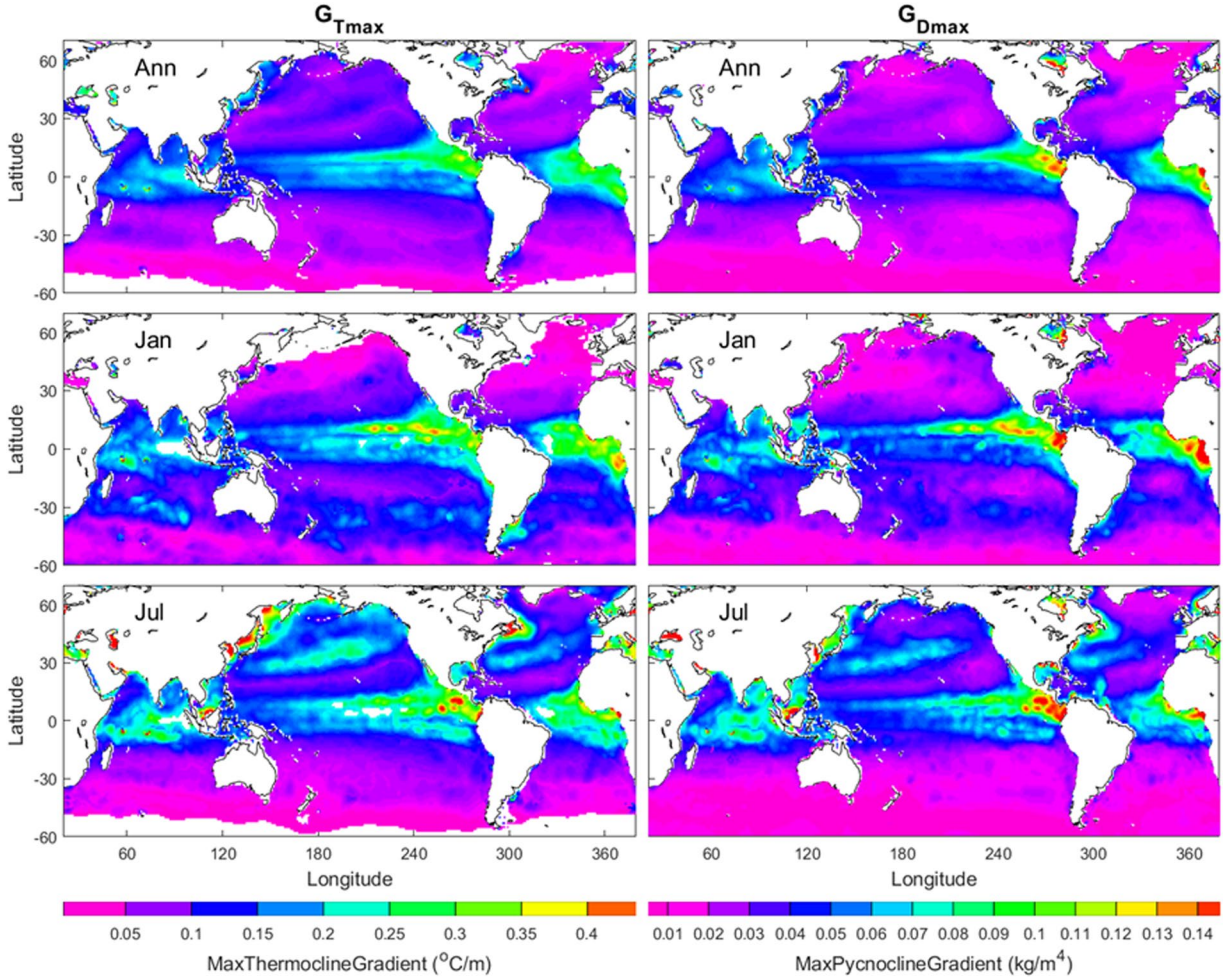

Fig. 9 Maps of annual, January, July mean (G_{Tmax}, G_{Dmax}).

$$h_T\frac{\partial T_{ITL}}{\partial t} + h_T\mathbf{V}\bullet\nabla T_{ITL} + \Lambda w_e\Delta T + \nabla\bullet\int_{-h_T}^{0}\hat{\mathbf{V}}\hat{T}dz = \frac{Q_0 - Q_{-h_T}}{\rho c_p}, \quad \Delta T \equiv T_{ITL} - T_{th} \tag{9}$$

where $\mathbf{V}$ is the vertically averaged horizontal velocity from the surface to $-h_T$; $\hat{\mathbf{V}}$ and $\hat{T}$ are deviation from the vertical average; Q_0 is the net surface heat flux adjusted for the penetration of light below the ITL; Q_{-h_T} is the vertical turbulent diffusion at the base of ITL; w_e is the entrainment velocity; Λ is the Heaviside unit function taking 0 if $w_e < 0$, and 1 otherwise due to the second law of the thermodynamics; ΔT is the temperature difference between ITL and thermocline; T_{th} is the thermocline temperature.

ITL heat content and ML freshwater content. The ITL heat content is the integrated heat stored in the layer from the sea surface ($z = 0$) down to ILD ($z = -h_T$),

$$H_{ITL} = c_p\int_{-h_T}^{0}\rho(z)T(z)dz \tag{10}$$

where $c_p = 3{,}985\,\mathrm{J\,kg^{-1}\,°C^{-1}}$, is the specific heat for sea water. The annual, January, and July mean H_{ITL} show that H_{ITL} is higher in low latitudes than in middle and high latitudes and is higher in the western than eastern Pacific and Atlantic Oceans within the low latitudes (left panels in Fig. 8). The ITL heat content H_{ITL} represents the warming/cooling of the ITL only. However, the fixed-depth OHC represents warming/cooling of ITL, combined ITL-part of thermocline, or combined ITL-thermocline-part of deep layer depending on the selection of depth with various warming trends. For example, the warming trend is estimated as $0.64 \pm 0.11\,\mathrm{W\,m^{-2}}$ for OHC (0–300 m) in 1993–2008[32], and $0.20\,\mathrm{W\,m^{-2}}$ for OHC (0–700 m) in 1955–2012[36]. The global ocean ITL warming rate from 1970 to 2017 is identified as ($0.14\,\mathrm{W\,m^{-2}}$)[47].

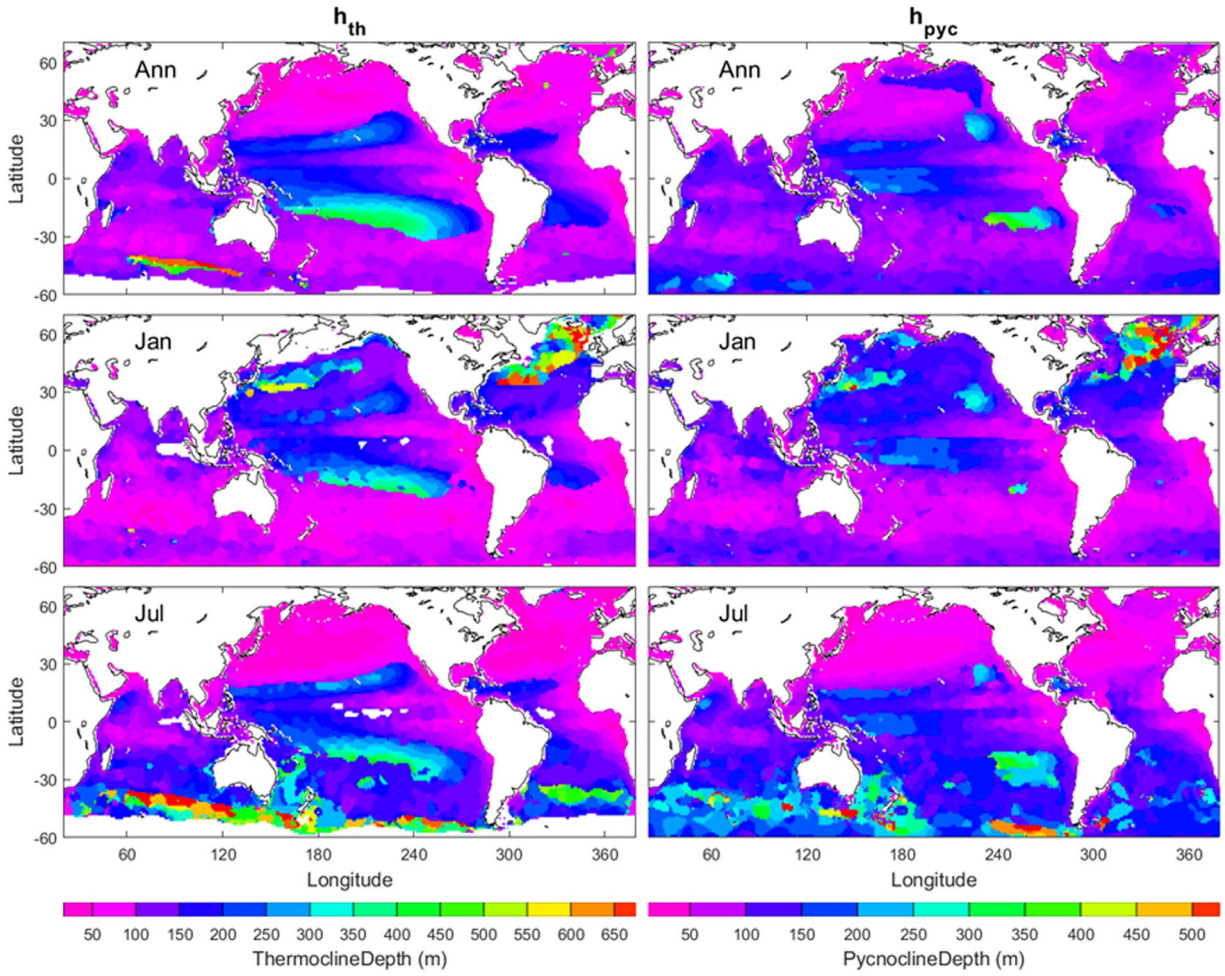

Fig. 10 Maps of annual, January, July mean (h_{th}, h_{pyc}) with the left panels for h_{th} and the right panels for h_{pyc}. The land and the areas of low quality of the identification are represented by white with the land enclosed by black curves (i.e., coasts). The "slice-like" feature may be caused by the vertical resolution of 25 m from 100 m to 500 m depth in the WOA18.

The ML freshwater content (F_{ML}) is the integrated heat stored in the layer from the sea surface ($z = 0$) down to MLD ($z = -h_D$)[48],

$$F_{ML} = \int_{-h_D}^{0} \left[1 - \frac{S(z)}{S_{ref}}\right] dz, \quad S_{ref} = 35\,\text{psu} \tag{11}$$

The annual mean F_{ML} is mostly positive in the North Pacific Ocean, the Southern Ocean, the Indian Ocean except the Arabian Sea, and mostly negative in the Atlantic Oceans. Four evident negative F_{ML} areas are in the Arabian Sea, central and eastern South Pacific Ocean between equator to 25°S, the North Atlantic Ocean from the Caribbean Sea to the west coast of Spain, and the South Atlantic Ocean between the equator to 30°S (upper right panel in Fig. 8).

Seasonal variability is evident with stronger positive F_{ML} in the North Pacific Ocean, weaker negative F_{ML} in the South Pacific Ocean, stronger negative F_{ML} in the North Atlantic Ocean, weaker negative F_{ML} in the South Atlantic Ocean in January than in July (central and lower right panels in Fig. 8). Strongest positive F_{ML} (>5 m) occurs in the northern (35° N–60° N) North Pacific Ocean in January. Strongest negative F_{ML} (<−5 m) occurs in the central to eastern subtropical North Atlantic Ocean in January, and in the eastern subtropical South Pacific Ocean and the subtropical South Atlantic Ocean in July. Note that the ML freshwater content (F_{ML}) dataset is also new since it is different from the fixed-depth freshwater content.

Thermocline and pycnocline. Maximum gradients can be obtained from vertical gradients of (T, ρ) between $z_{(0.1)}$ and $z_{(0.7)}$ calculated by Eq. (2),

$$G_{T\max} = \text{Max}(G_{T1}, G_{T2}, \ldots, G_{TI}), \quad G_{D\max} = \text{Max}(G_{D1}, G_{D2}, \ldots, G_{DI}) \tag{12}$$

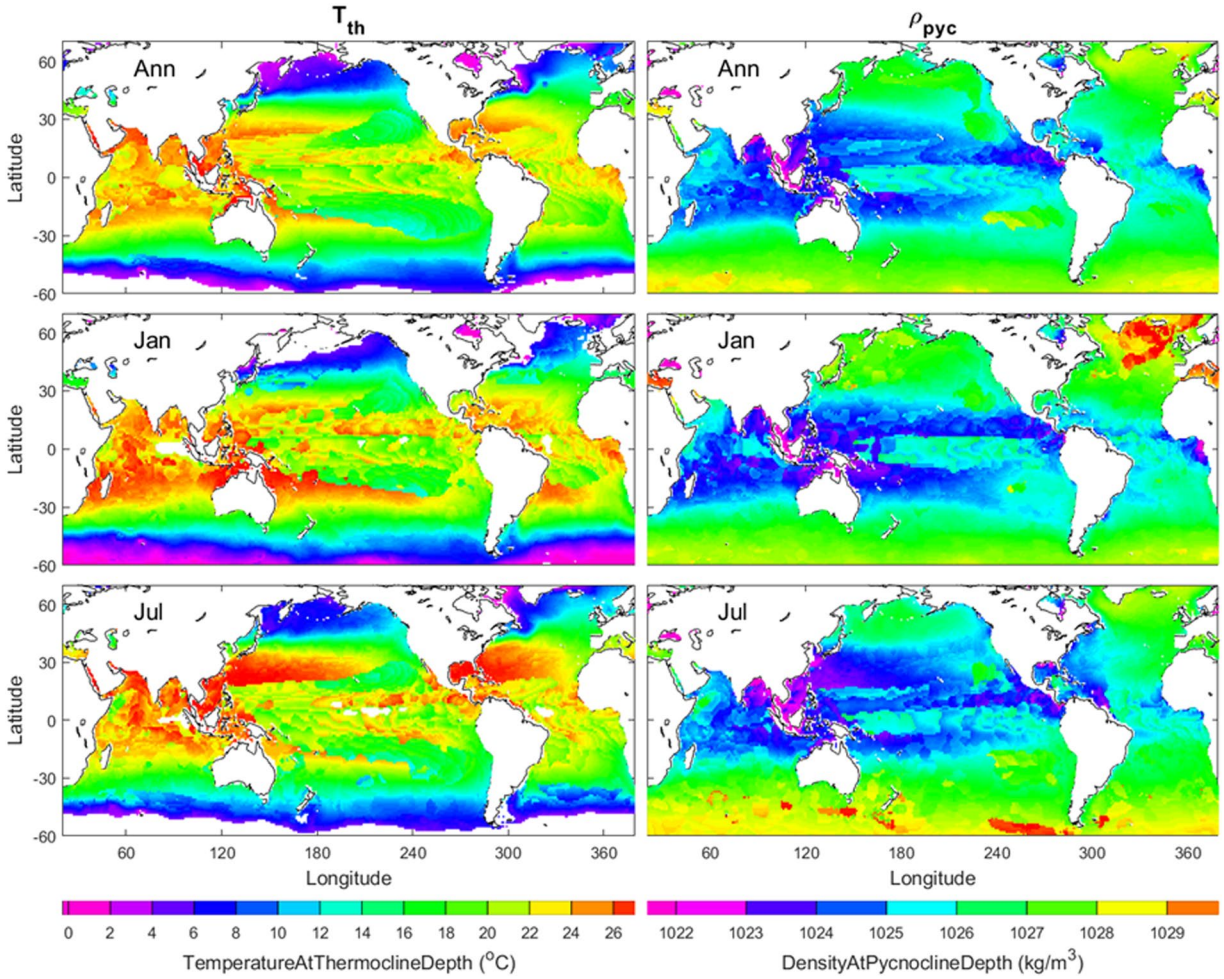

Fig. 11 Maps of annual, January, July mean (T_{th}, ρ_{pyc}) with the left panels for T_{th} and the right panels for ρ_{pyc}. The land and the areas of low quality of the identification are represented by white with the land enclosed by black curves (i.e., coasts). The "slice-like" feature may be caused by the vertical resolution of 25 m from 100 m to 500 m depth in the WOA18.

which are used to represent the depths of thermocline (h_{th}) and pycnocline (h_{pyc})

$$G_{Ti}(z_i = -h_{th}) = G_{T\max},\ G_{Di}(z_i = -h_{pyc}) = G_{D\max} \tag{13}$$

The temperature at the thermocline depth is defined as the thermocline temperature

$$T_{th} = T(z = -h_{th}) \tag{14}$$

The density at the pycnocline depth is defined as the pycnocline density

$$\rho_{pyc} = \rho(z = -h_{pyc}) \tag{15}$$

Annual mean G_{Tmax} and G_{Dmax} have evident latitudinal variability with large values in the tropical regions and small values outside the tropical region (upper two panels in Fig. 9). In the tropical Pacific, Atlantic, and Indian Oceans, the large values are in the central and eastern parts. Seasonal variability of (G_{Tmax}, G_{Dmax}) is strong in the Northern Hemisphere and weak in the Southern Hemisphere. In the Northern Hemisphere, the difference is small between annual and January mean (G_{Tmax}, G_{Dmax}) but large between annual and July mean (G_{Tmax}, G_{Dmax}). In July, very strong (G_{Tmax}, G_{Dmax}) occur in the northeastern Asian and American coastal regions, and strong (G_{Tmax}, G_{Dmax}) appear in the Kuroshio extension and Gulf Stream regions (middle and lower panels in Fig. 9).

Annual mean h_{th} and h_{pyc} have evident spatial variability with large values in low latitudes (30°S–30°N) with shallowing of (h_{th}, h_{pyc}) in the eastern Pacific and Atlantic Oceans (upper two panels in Fig. 10). Seasonal variability of (h_{th}, h_{pyc}) is strong in middle and high latitudes with deep (h_{th}, h_{pyc}) in (30°N–60°N) in January and deep (h_{th}, h_{pyc}) in (30°S–60°S) in July (middle and lower panels in Fig. 10).

Annual mean T_{th} and ρ_{pyc} have evident spatial variability with warm T_{th} and light ρ_{pyc} in low latitudes (30°S–30°N) and cold T_{th} and dense ρ_{pyc} in middle and high latitudes (30°S–60°S, 30°N–60°N) (upper two panels

Thermohaline Parameter	Mean	Standard Deviation	Skewness	Kurtosis
Isothermal Layer Depth (m)	51.4	34.2	2.16	11.3
Mixed Layer Depth (m)	38.7	19.1	0.98	3.95
Mean Thermocline Gradient (°C/m)	0.0638	0.0501	1.49	5.17
Mean Pycnocline Gradient (kg/m^4)	0.0212	0.0179	2.36	12.3
Barrier Layer Depth (m) (Total #: 23562)	14.3	22.4	3.70	18.6
Compensated Layer Depth (m) (Total #: 170)	11.3	12.5	3.33	15.2
Isothermal Layer Heat Content (10^9J/m^2)	4.30	3.80	1.41	4.39
Isothermal Layer Temperature (°C)	19.0	8.02	−0.595	2.02
Maximum Thermocline Gradient (°C/m)	0.0742	0.0580	1.61	6.98
Thermocline Depth (m)	99.4	85.2	2.97	15.1
Temperature at Thermocline Depth (°C)	15.6	6.62	−0.624	2.27
Mixed Layer Freshwater Content (m)	0.0932	0.987	−0.310	2.90
Mixed Layer Density (kg/m^3)	1024.60	1.74	−0.636	8.10
Maximum Pycnocline Gradient (kg/m^4)	0.0246	0.023	3.4	27.3
Pycnocline Depth (m)	73.9	39.4	2.20	13.1
Density at Pycnocline Depth (kg/m^3)	1025.97	1.56	−2.17	25.7

Table 2. Statistical characteristics of thermohaline parameters derived from the annual mean (T, ρ) profiles.

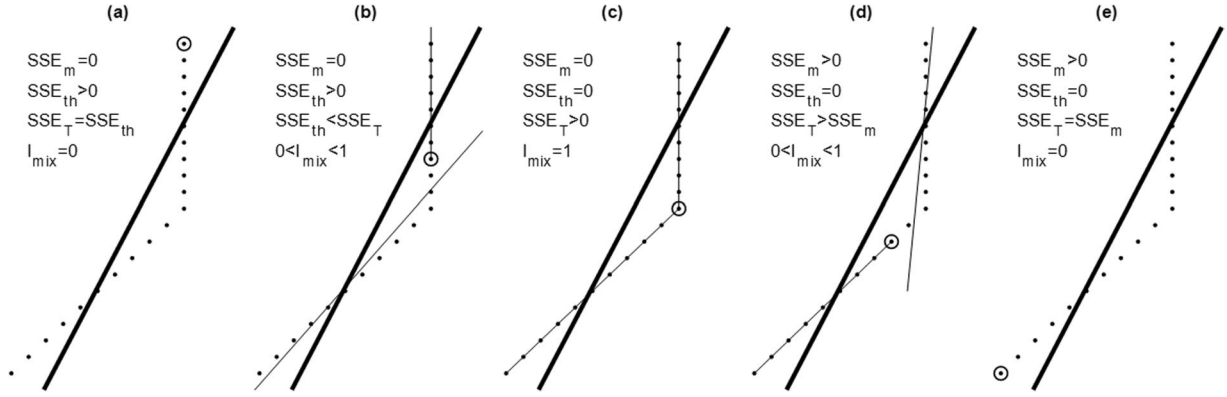

Fig. 12 The i-index to represent the quality of the ITL depth determination. (**a**) ITL not existence or ITL in existence, but not identified ($h_T = 0$). (**b**) Identified ITL depth shallower than the real ITL depth. (**c**) Perfectly identified ITL depth. (**d**) Identified ITL depth deeper than the real ITL depth. (**e**) Identified ITL depth at $z_{(0.7)}$, i.e., thermocline not existence. Here, the ITL depth is marked by a circle. The thick line in each figure is a single linear function fitted to the temperature profile data.

in Fig. 11). Seasonal variability of (T_{th}, ρ_{pyc}) is generally weak except in subtropical oceans where warmer T_{th} (light ρ_{pyc}) in the Sothern Hemisphere in January and in the Northern Hemisphere in July. Furthermore, cold T_{th} and dense ρ_{pyc} appears in North Atlantic Ocean near Greenland in January (middle and lower panels in Fig. 11).

Global statistics of the thermohaline parameters. The identified thermohaline parameters are on the grid points. We use the standard Matlab codes to calculate with area-weighted mean, standard deviation, skewness, and kurtosis for each parameter. Table 2 shows the statistical characteristics of thermohaline parameters derived from the annual mean (T, ρ) profiles. These values can be treated as the overall climatological values of global thermohaline parameters, such as 51.4 m for isothermal layer depth, 38.7 m for mixed layer depth, 14.3 m for barrier layer depth, 99.4 m the thermocline depth, 73.9 m for the pycnocline depth, 0.0638 °C/m for thermocline gradient, and 0.0212 kg/m^4 for pycnocline gradient.

Data Records

This global ocean climatology of thermohaline parameter dataset is publicly available at the NOAA/NCEI data repository as a NetCDF file, which includes data citation, dataset identifiers, metadata, and ordering instructions. The dataset is located at the NOAA/NCEI website (https://doi.org/10.25921/j3v2-jy50).

Technical Validation

From the identified parameters (h_T, G_T) and (h_D, G_D), fitted temperature and density profiles [$\hat{T}(z_k)$, $\hat{\rho}(z_k)$] can be constructed using zero gradient in the ITL and ML, and linear gradient (G_T, G_D) in the thermocline and pycnocline,

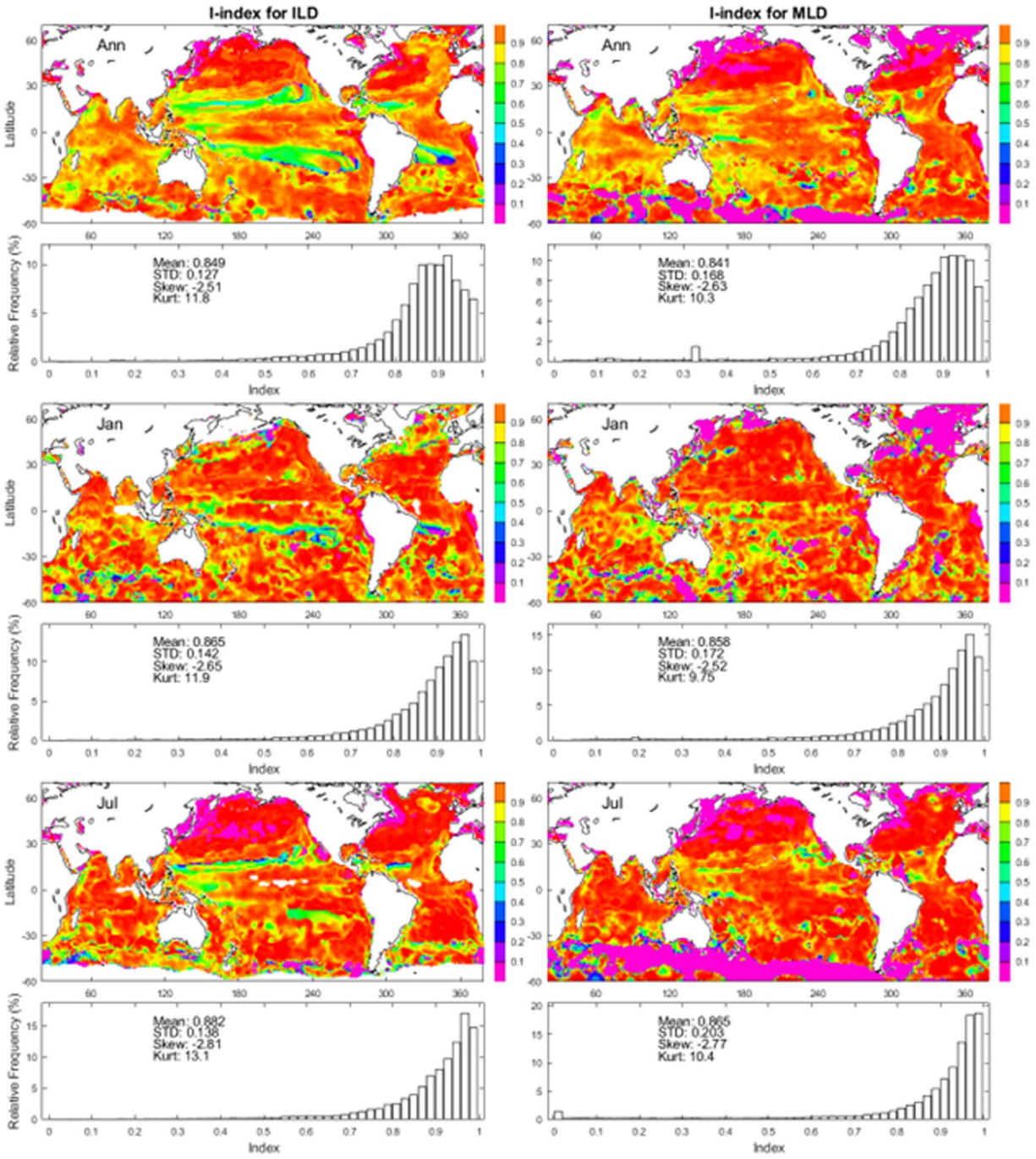

Fig. 13 Maps and histograms of the i-index to determine (h_T, G_T) from the annual, January, July mean temperature profiles (left panels) and (h_D, G_D) from the annual, January, July mean density profiles (right panels). Here, top panels are for the annual mean; middle panels are for the January mean; and lower panels are for the July mean.

$$\hat{T}(z_k) = \begin{cases} T_{ITL} & \text{for } z_k \geq -h_T \text{ (in ITL)} \\ T_{ITL} + (z_k + h_T)G_T & \text{for } z_k < -h_T \text{ (in thermocline)} \end{cases} \tag{16a}$$

$$\hat{\rho}(z_k) = \begin{cases} \rho_{ML} & \text{for } z_k \geq -h_D \text{ (in ML)} \\ \rho_{ML} - (z_k + h_D)G_D & \text{for } z_k < -h_T \text{ (in pycnocline)} \end{cases} \tag{16b}$$

The methodology is evaluated through the comparison between the fitted profiles $[\hat{T}(z_k),\ \hat{\rho}(z_k)]$ and the WOA18 profiles $[T(z_k),\ \rho(z_k)]$.

We take T-profile for illustration. The fitted $\hat{T}(z_k)$ profile is represented by two lines with the first one (near zero gradient in the ITL) from the top to the ITL base (circles in Fig. 12) and the second one (non-zero gradient in the thermocline) from the ITL base to the bottom of the profile. The quality in determination of (h_T, G_T) is identified by the sum of the square error (SSE) between the fitted $\hat{T}(z_k)$ and WOA18 $T(z_k)$ at the WOA18 vertical levels (see Table 1),

$$SSE = \sum_k [\hat{T}(z_k) - T(z_k)]^2 \tag{17}$$

Such double-gradient fitting has errors in the ITL represented by the sum of the square error in ITL (SSE_{ITL}) and in the thermocline represented by the sum of the square error in thermocline (SSE_{TH}). The whole T-profile data fitted to a single linear function (thick lines in Fig. 12) represents the maximum error since it disregards the existence of ITL and thermocline. Such a maximum error is represented by the total sum of the square error (SSE_T). The identification index (called i-index) is defined by[29]

$$I_{ITL} = \sqrt{1 - (SSE_{ITL} + SSE_{TH})/SSE_T}\,. \tag{18}$$

If an ITL exists but is not identified ($h_T = 0$) (Fig. 12a), $SSE_{ITL} = 0$, $SSE_{TH} = SSE_T$; which gives $I_{ITL} = 0$. If the identified h_T is shorter than the real one (Fig. 12b), $SSE_{ITL} = 0$, $SSE_{TH} > 0$, $SSE_{TH} < SSE_T$; which leads to $0 < I_{ITL} < 1$. If the identified h_T is the same as the real one (Fig. 12c), $SSE_{ITL} = 0$, $SSE_{TH} = 0$, $SSE_T > 0$; which makes $I_{ITL} = 1$. If the identified h_T is deeper than the real one (Fig. 12d), $SSE_{ITL} > 0$, $SSE_{TH} = 0$, $SSE_{TH} < SSE_T$; which leads to $0 < I_{ITL} < 1$. If the identified h_T reaches the bottom of the thermocline (Fig. 12e), $SSE_{TH} = 0$, $SSE_{ITL} = SSE_T$; which gives $I_{ITL} = 0$.

Thus, the value of I_{ITL} represents the quality of the determination of h_T from T-profile: (a) $I_{ITL} = 1$ for no error (Fig. 12c), (b) $I_{ITL} = 0$ for 100% error with no ITL identified but actual existence of ITL (Fig. 12a) or identified h_T at the bottom of the thermocline (Fig. 12e,c) $0 < I_{ITL} < 1$ for identified h_T shallower than the actual h_T (Fig. 12b) and for identified h_T deeper than the actual h_T (Fig. 12d).

The annual, January, July mean maps and histograms of the i-index for ILD show high quality of the method. The global average i-index is 0.849 for the annual mean, 0.865 for the January mean, and 0.882 for the July mean (left panels in Fig. 13). The annual, January, July mean maps and histograms of the i-index for MLD show quality of the method. The global average i-index is 0.841 for the annual mean, 0.858 for the January mean, and 0.865 for the July mean (right panels in Fig. 13).

Code availability

The MATLAB codes to determine these thermohaline parameters from WOD18 (T, S) were published in the two related papers[28,29], and can be obtained at https://doi.org/10.1007/s10872-017-0418-0. The main program is Thermohaline.m, along with four Matlab functions: validate.m, getgradient.m, ELGCore.m, and Iindex.m (see the Technical Validation Section). The function validate.m is employed to identify if (T, ρ) profiles having double gradient structure. The function 'getgradient.m' is used to calculate the vertical gradient. The function ELGCore.m is used to get (h_T, G_T) or (h_D, G_D) from T-profile or ρ-profile. The function Iindex.m is used to calculate the i-index (I_{ITL}) for the error estimation. Then, the code 'Thermohaline.m' generates other thermohaline parameters such as ITL heat content, mixed layer fresh-water content, maximum thermocline gradient, thermocline depth, temperature at thermocline depth, maximum density gradient, pycnocline depth, density at pycnocline depth, barrier layer depth, and compensated layer depth. Since the calculation is local for individual (T, S) profile pair, interested readers may use our MATLAB codes to analyse any (T, S) profiles to get the derived thermohaline data with quality identification (i.e., i-index).

Received: 22 December 2022; Accepted: 13 June 2023;
Published online: 24 June 2023

References

1. Shay, L. K. Upper ocean structure: responses to strong atmospheric forcing events. In *Encyclopedia of Ocean Sciences* (ed. Steele, J. H.) 192–210 (Elsevier Ltd. 2010).
2. Chu, P. C., Garwood, R. W. Jr. & Muller, P. Unstable and damped modes in coupled ocean mixed layer and cloud models. *J. Mar Syst.* **1**, 1–11 (1990).
3. Chu, P. C. Generation of low frequency unstable modes in a coupled equatorial troposphere and ocean mixed layer. *J. Atmos. Sci.* **50**, 731–749 (1993).
4. Chu, P. C. & Garwood, R. W. Jr. On the two-phase thermodynamics of the coupled cloud-ocean mixed layer. *J. Geophys. Res.* **96**, 3425–3436 (1991).
5. Monterey, G. I. & Levitus S. *Seasonal Variability of Mixed Layer Depth for the World Ocean*, 1–5, 87 figs (NOAA Atlas NESDIS 14, 1997).
6. Lukas, R. & Lindstrom, E. The mixed layer of the western equatorial Pacific Ocean. *J. Geophys. Res.* **96**, 3343–3357 (1991).
7. Sprintall, J. & Tomczak, M. Evidence of the barrier layer in the surface layer of the tropics. *J. Geophys. Res.* **97**, 7305–7316 (1992).
8. Pailler, K., Bourles, B. & Gouriou, Y. The barrier layer in the western tropical Atlantic Ocean. *Geophys. Res. Lett.* **26**, 2069–2072 (1999).
9. Kara, A. B., Rochford, P. A. & Hurlburt, H. E. Mixed layer depth variability and barrier layer formation over the North Pacific Ocean. *J. Geophys. Res.* **105**, 16783–16801 (2000).

10. Chu, P. C., Liu, Q. Y., Jia, Y. L. & Fan, C. W. Evidence of barrier layer in the Sulu and Celebes Seas. *J. Phys. Oceanogr.* **32**, 3299–3309 (2002).
11. Qu, T. & Meyers, G. Seasonal variation of barrier layer in the southeastern tropical Indian Ocean. *J. Geophys. Res.* **110**, https://doi.org/10.1029/2004JC002816 (2005).
12. Agarwal, N. *et al.* Argo observations of barrier layer in the tropical Indian Ocean. *Adv. Space Res.* **50**, 642–654 (2012).
13. Vissa, N. K., Satyanarayana, A. N. V. & Prasad Kumar, B. Comparison of mixed layer depth and barrier layer thickness for the Indian Ocean using two different climatologies. *Internat. J. Climatol.* **33**, 2855–2870 (2013).
14. Yan, Y., Li, L. & Wang, C. The effects of oceanic barrier layer on the upper ocean response to tropical cyclones. *J. Geophys. Res.* **122**, 4829–4844 (2017).
15. de Boyer Montegut, C., Madec, G., Fisher, A. S., Lazar, A. & Iudicone, D. Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *J. Geophys. Res.* **109**, C12003 (2004).
16. Dong, S., Sprintall, J., Gille, S.T. & Talley, L. Southern Ocean mixed-layer depth from Argo float profiles. *J. Geophys. Res.* **113**, https://doi.org/10.1029/2006JC004051 (2008).
17. Sprintall, J. & Roemmich, D. Characterizing the structure of the surface layer in the Pacific Ocean. *J. Geophys. Res.* **104**, 23,297–23,311 (1999).
18. Suga, T., Motoki, K., Aoki, Y. & Macdonald, A. M. The North Pacific climatology of winter mixed layer and mode waters. *J. Phys. Oceanogr.* **34**, 3–22 (2004).
19. Brainerd, K. E. & Gregg, M. C. Surface mixed and mixing layer depths. *Deep Sea Res., Part A* **9**, 1521–1543 (1995).
20. Wyrtki, K. The thermal structure of the eastern Pacific Ocean. *Dstch. Hydrogr. Zeit., Suppl. Ser. A* **8**, 6–84 (1964).
21. Defant, A. *Physical Oceanography* Vol. 1 (Pergamon Press, 1961).
22. Thomson, R. E. & Fine, I. V. Estimating mixed layer depth from oceanic profile data. *J. Atmos. Oceanic Technol.* **20**, 319–329 (2003).
23. Chu, P. C., Fralick, C. R., Haeger, S. D. & Carron, M. J. A parametric model for Yellow Sea thermal variability. *J. Geophys. Res.* **102**, 10499–10508 (1997).
24. Chu, P. C., Wang, Q. Q. & Bourke, R. H. A geometric model for Beaufort/Chukchi Sea thermohaline structure. *J. Atmos. Oceanic Technol.* **16**, 613–632 (1999).
25. Chu, P. C., Fan, C. W. & Liu, W. T. Determination of sub-surface thermal structure from sea surface temperature. *J. Atmos. Oceanic Technol.* **17**, 971–979 (2000).
26. Chu, P. C. & Fan, C. W. Optimal linear fitting for objective determination of ocean mixed layer depth from glider profiles. *J. Atmos. Oceanic Technol.* **27**, 1893–1898 (2010).
27. Chu, P. C. & Fan, C. W. Maximum angle method for determining mixed layer depth from seaglider data. *J. Oceanogr.* **67**, 219–230 (2011).
28. Chu, P. C. & Fan, C. W. Exponential leap-forward gradient scheme for determining the isothermal layer depth from profile data. *J. Oceanogr.* **73**, 503–526 (2017).
29. Chu, P.C. & Fan, C. W. Global ocean synoptic thermocline gradient, isothermal-layer depth, and other upper ocean parameters. *Nature Sci. Data*, **6**, 119, https://www.nature.com/articles/s41597-019-0125-3 (2019).
30. Chu, P. C. & Fan, C. W. Global ocean thermocline gradient, isothermal layer depth, and other upper ocean parameters calculated from WOD CTD and XBT temperature profiles from 1960-01-01 to 2017-12-31. NOAA National Cnters for Environmental Information Accession 0173210, https://doi.org/10.25921/dgak-7a43 (2018).
31. Chu, P. C. Global upper ocean heat content and climate variability. *Ocean Dyn* **61**, 1189–1204, https://doi.org/10.1007/s10236-011-0411-x (2011).
32. Lyman, J. M. *et al.* Robust warming of the global upper ocean. *Nature* **465**, 334–337 (2010).
33. Gouretski, V., Kennedy, J., Boyer, T. & Köhl, A. Consistent near-surface ocean warming since 1900 in two largely independent observing networks. *Geophys. Res. Lett.* **39**, L19606, https://doi.org/10.1029/2012GL052975 (2012).
34. Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J. & Heimbach, P. Global reconstruction of historical ocean heat storage and transport. *PNAS* **116**, 1126–1131 (2019).
35. Willis, J. K., Roemmich, D. & Cornuelle, B. Interannual variability in upper ocean heat content, temperature, and thermosteric expansion on global scales. *J. Geophys. Res.* **109**, C12036, https://doi.org/10.1029/2003JC002260 (2004).
36. Levitus, S. *et al.* World ocean heat content and thermosteric sea level change (0-2000 m), 1955–2010. *Geophys. Res. Lett.* **39**, L10603, https://doi.org/10.1029/2012GL05116 (2012).
37. Balmaseda, M. A., Trenberth, K. E. & Källén, E. Distinctive climate signals in reanalysis of global ocean heat content. *Geophys. Res. Lett.* **40**, 1,754–1,759 (2013).
38. Cheng, L., Trenberth, K. E., Palmer, M. D. & Zhu, J. Observed and simulated full-depth ocean heat-content changes for 1970–2005. *Ocean Sci.* **12**, 925–935 (2016).
39. Abraham, J. P. *et al.* A review of global ocean temperature observations: implications for ocean heat content estimates and climate change. *Rev. Geophysics* **51**, 450–483 (2013).
40. Bindoff, N. L. *et al.* "Observations: oceanic climate change and sea level. In: *Climate Change, The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. & Miller, H.L. (eds.)], Cambridge, United Kingdom and New York, NY, USA Cambridge University Press, (2007).
41. Kara, A. B., Rochford, P. A. & Hurlburt, H. E. Naval Research Laboratory Mixed Layer Depth (NMLD) Climatologies. Defense Technical Information Center, Accession Number: ADA401151, https://apps.dtic.mil/sti/citations/ADA401151 (2002).
42. Bretherton, F. P., Davis, R. E. & Fandry, C. B. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. *Deep-Sea Res* **23**, 559–582, https://doi.org/10.1016/0011-7471(76)90001-2 (1976).
43. Evensen, G. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn* **53**, 343–367 (2003).
44. Chu, P. C., Fan, C. W. & Margolina, T. Ocean spectral data assimilation without background error covariance matrix. *Ocean Dyn.* **66**, 1143–1163, https://doi.org/10.1007/s10236-016-0971-x (2016).
45. Mixed layer depth, isothermal layer depth, barrier layer depth, and other upper ocean thermohaline parameters, total of 17, derived from climatological gridded vertical profiles from the World Ocean Atlas 2018 (NCEI Accession 0205198) *NOAA National Centers for Environmental Information* https://doi.org/10.25921/j3v2-jy50 (2019).
46. Stevenson, J. W. & Niiler, P. P. Upper ocean heat budget during the Hawaii-to-Tahiti shuttle experiment. *J. Phys. Oceanogr.* **13**, 1894–1907 (1983).
47. Chu, P. C. & Fan, C. W. World ocean thermocline weakening and isothermal layer warming. *MDPI Applied Sci.* **10**(22), 8185, https://doi.org/10.3390/app10228185 (2020).
48. Chu, P. C. & Fan, C. W. Temporal and spatial variability of global upper ocean freshwater content. *Mod. App. Ocean & Pet. Sci.* **1**(5), MAOPS.MS.ID.000123, https://doi.org/10.32474/MAOPS.2018.01.000123 (2018).

Acknowledgements
We thank Dr. Alexandra Grodsky and Dr. Donald Collins for their outstanding efforts to publish this dataset at the NOAA/NCEI website, the NOAA/NCEI for providing the WOA18 data, as well as two anonymous reviewers and Dr. Alireza Foroozani (Associate Editor) for constructive comments, encouragement, and guidance for revising the manuscript.

Author contributions
P.C.C. developed the method, designed the project, conducted the data quality control, and prepared the Data Descriptor. C.W.F. developed the codes, and helped prepare the Data Descriptor.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to P.C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.