Intercomparison and validation of the mixed layer depth fields of global ocean syntheses

Article

Accepted Version

Toyoda, T., Fujii, Y., Kuragano, T., Kamachi, M., Ishikawa, Y., Masuda, S., Sato, K., Awaji, T., Hernandez, F., Ferry, N., Guinehut, S., Martin, M. J., Peterson, K. A., Good, S. A., Valdivieso, M. ORCID: https://orcid.org/0000-0002-1738-7016, Haines, K. ORCID: https://orcid.org/0000-0003-2768-2374, Storto, A., Masina, S., Kohl, A., Zuo, H., Balmaseda, M., Yin, Y., Shi, L., Alves, O., Smith, G., Chang, Y.-S., Vernieres, G., Wang, X., Forget, G., Heimbach, P., Wang, O., Fukumori, I. and Lee, T. (2017) Intercomparison and validation of the mixed layer depth fields of global ocean syntheses. Climate Dynamics, 49 (3). pp. 753-773. ISSN 0930-7575 doi: https://doi.org/10.1007/s00382-015-2637-7 Available at https://centaur.reading.ac.uk/52363/

Published version at: http://link.springer.com/article/10.1007%2Fs00382-015-2637-7
To link to this article DOI: http://dx.doi.org/10.1007/s00382-015-2637-7

Publisher: Springer

Title

Intercomparison and validation of the mixed layer depth fields of global ocean syntheses

Authors

Takahiro Toyoda, Yosuke Fujii, Tsurane Kuragano, and Masafumi Kamachi (Meteorological Research Institute, Japan Meteorological Agency (MRI/JMA), Tsukuba, Japan)

Yoichi Ishikawa (Center for Earth Information Science and Technology, Japan Agency for Marine-Earth Science and Technology (CEIST/JAMSTEC), Yokohama, Japan)

Shuhei Masuda, and Kanako Sato (Research and Development Center for Global Change (RCGC), JAMSTEC, Yokosuka, Japan)

Toshiyuki Awaji (CEIST/JAMSTEC, Yokohama, Japan; Kyoto University, Kyoto, Japan)

Fabrice Hernandez (Institut de Recherche pour le Développement (IRD), Toulouse, France; Mercator Océan, Ramonville Sant-Agne, France)

Nicolas Ferry (Mercator Océan, Ramonville Sant-Agne, France)

Stéphanie Guinehut (Collecte Localisation Satellites (CLS), Ramonville Sant-Agne, France)

Matthew Martin, K. Andrew Peterson, and Simon A. Good (Met Office (UKMO), Exeter, UK)

Maria Valdivieso, and Keith Haines (National Centre for Earth Observation (NCEO), Department of Meteorology, University of Reading (U-Reading), Reading, UK)

Andrea Storto (Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Bologna, Italy)

Simona Masina (CMCC; Istituto Nazionale di Geofisica e Vulcanologia (INGV), Bologna, Italy)

Armin Köhl (Universität Hamburg (U-Hamburg), Hamburg, Germany)

Hao Zuo, and Magdalena Balmaseda (European Centre for Medium-Range Weather Forecasts

(ECMWF), Reading, UK)
Yonghong Yin, Li Shi, and Oscar Alves (Centre for Australian Weather and Climate Research, Bureau of Meteorology (BOM), Melbourne, Australia)
Gregory Smith (Environment Canada, Québec, Canada)
You-Soon Chang (Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration (GFDL/NOAA), Princeton, New Jersey, USA; Kongju National University, Kongju, South Korea)
Guillaume Vernieres (Science System and Applications, Inc., Lanham, Maryland, USA; Global Modeling and Assimilation Office, National Aeronautics and Space Administration Goddard Space Flight Center (GSFC/NASA/GMAO), Greenbelt, Maryland, USA)
Xiaochun Wang (Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, California, USA)
Gael Forget, and Patrick Heimbach (Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA)
Ou Wang, Ichiro Fukumori, and Tong Lee (Jet Propulsion Laboratory (JPL), California Institute of Technology, Pasadena, California, USA)

Corresponding author

Takahiro Toyoda, Oceanography and Geochemistry Research Department, Meteorological Research Institute, Japan Meteorological Agency, 1-1 Nagamine, Tsukuba, 305-0052, Japan
E-mail: ttoyoda@mri-jma.go.jp

Abstract

Intercomparison and evaluation of the global ocean surface mixed layer depth (MLD) fields estimated from a suite of major ocean syntheses are conducted. Compared with the reference MLDs calculated from individual profiles, MLDs calculated from monthly mean and gridded profiles show negative biases of 10-20 m in early spring related to the re-stratification process of relatively deep mixed layers. Vertical resolution of profiles also influences the MLD estimation. MLDs are underestimated by approximately 5-7 (14-16) m with the vertical resolution of 25 (50) m when the criterion of potential density exceeding the 10-m value by 0.03 kg m^{-3} is used for the MLD estimation. Using the larger criterion (0.125 kg m^{-3}) generally reduces the underestimations. In addition, positive biases greater than 100 m are found in wintertime subpolar regions when MLD criteria based on temperature are used. Biases of the reanalyses are due to both model errors and errors related to differences between the assimilation methods. The result shows that these errors are partially cancelled out through the ensemble averaging. Moreover, the bias in the ensemble mean field of the reanalyses is smaller than in the observation-only analyses. This is largely attributed to comparably higher resolutions of the reanalyses. The robust reproduction of both the seasonal cycle and interannual variability by the ensemble mean of the reanalyses indicates a great potential of the ensemble mean MLD field for investigating and monitoring upper ocean processes.

Keywords

ocean reanalysis, mixed layer depth, Ocean Reanalyses Intercomparison Project (ORA-IP), data assimilation, ocean general circulation model, isothermal layer depth

1 Introduction

The ocean surface mixed layer (ML), directly communicating with the atmosphere, transmits surface heat, freshwater and momentum fluxes to the interior ocean, which forces the ocean circulation (e.g., Pedlosky 1996). On the other hand, sea surface temperature (SST), which provides the boundary condition for the atmosphere, is determined through the ML processes. Also, heat content in the ML is one of the primary forcing factors of the atmospheric circulation in some cases (e.g., Shey et al. 2000). Since the ML depth (MLD) is a relevant physical parameter for describing the dynamic nature of the ML, it is of great value to quantitatively determine the spatio-temporal variation of the MLD in the global ocean for better understanding the ocean circulation and air-sea interaction.

The Ocean Reanalyses Intercomparison Project (ORA-IP) was initiated to evaluate global ocean syntheses produced in several research and operational centers (Balmaseda et al. 2015). These syntheses include both analyses that use observations only and analyses that combine ocean models and observations through data assimilation methods (referred to as "reanalyses" in this study). It is important to evaluate their strength and weakness in various aspects in order to understand the extent to which these products can be used to monitor the state of the ocean, initialize climate prediction and understand oceanic physical processes and in order also to identify priorities for new developments (e.g., Lahoz and Errera 2010). Furthermore, the signal-to-noise ratio inferred from an ensemble of these products can improve the understanding of the robustness of oceanic physical processes represented by these reanalyses (e.g., Lee et al. 2009).

MLD is selected as one of the important indices for the ORA-IP (Toyoda et al. 2014) in addition to heat and salt content, steric height, sea level, surface heat fluxes, depth of the 20 degree isotherm and sea ice. In the present study, monthly mean global MLD time series are estimated and intercompared from 19 syntheses (2 observation-only analyses and 17 reanalyses). Following a brief description of MLDs in Sect. 2, we first investigate the observation-only analyses focusing on errors in estimating MLDs in Sect. 3. The ensemble mean of the reanalyses is also examined since it can have a better fidelity in some regions if model errors in the individual reanalyses cancel out through the ensemble averaging approach. Intercomparisons of all the syntheses/reanalyses are provided in Sect. 4. The findings are summarized in Sect. 5.

2 Data

2.1 Definition of MLD

For the MLD definition, density criteria (e.g., Levitus 1982) are used in this study, i.e., MLD is defined as the depth where potential density exceeds the 10-m depth value by $\Delta\rho = 0.03$ or 0.125 kg m^{-3} ("MLDr003"/"MLDr0125"), since these 2 criteria are often used (e.g., Hosoda et al. 2010) and our interest is in sensitivity of the MLD estimation to the criterion value. Similarly, isothermal layer depth (ILD) is defined as the depth where potential temperature differs from the 10-m depth value by $\Delta T = 0.2$ºC or 0.5ºC ("ILDt02"/"ILDt05"). These ILDs were sometimes used as substitutes for MLDs in previous studies since salinity profiles are less numerous than temperature profiles. Hence, for the present intercomparison, both MLDs and ILDs are used as indices for vertical mixing intensity in the upper ocean. Note that the above temperature criterion ($\Delta T = 0.5$ºC), multiplied by the characteristic thermal expansion rate (e.g., 0.24 kg m^{-3} ºC^{-1} at 18ºC and 35 psu), generally correspond to the density criterion ($\Delta\rho = 0.125$ kg m^{-3}).

2.2 Observation-only analyses

Two observation-only analyses archived in the ORA-IP are used in this study, EN3v2a and ARMOR3D (Table 1). EN3v2a analyzed in-situ temperature and salinity (TS) observations; ARMOR3D synthesized satellite-derived sea level anomalies (SLAs) and SSTs in addition to in-situ TS observations. Monthly MLD and ILD time series are calculated from the monthly mean TS fields on the native grids of the individual datasets. Interpolated values on global longitude-latitude grids with one-degree resolution are used for the intercomparison.

In order to evaluate the MLDs/ILDs in the ORA-IP, we use the freely available MLD/ILD datasets of MILA-GPV (Hosoda et al. 2010) and de Boyer Montégut et al. (2004; "deBoyer" hereafter). These data are estimated as the average of MLDs/ILDs deduced from individual TS profiles. In particular, MILA-GPV uses only the Argo profiles without interpolation between grid points, although the spatio-temporal coverage of the dataset is limited. Hence, we use MILA-GPV as a reference for the intercomparison mainly (e.g., Fig. 1). Note that deBoyer provides only the monthly climatological fields for MLDr003 and ILDt02.

Additionally, we use MLDs/ILDs calculated from the monthly TS climatologies of the World Ocean Atlas (WOA) 2009 (e.g., Locarnini et al. 2010). Note that this dataset (derived from TS climatologies) is somewhat similar to EN3v2a and ARMOR3D (derived from monthly TS analyses) but different from MILA-GPV and deBoyer (derived from individual TS profiles).

2.3 Reanalyses

Each of the reanalyses used in this study may have their own systematic error, attributed to ocean general circulation model (OGCM), spatial resolution, surface forcing, ML parameterization, assimilated data and assimilation method adopted in each analysis (Table 2). The ensemble averaging will partially result in compensation of errors, thus decreasing the error of the MLD estimate, but still errors will remain. In addition, the reanalyses can be clustered in several groups: For example, versions of NEMO are used in G2V3, C-GLORS, UR025.4, GloSea5, ORAS4 and ORAP5, while versions of MOM are used in MERRA, ECDA, PEODAS, K7-ODA and K7-CDA; Smoother approaches are adopted in GECCO2, ECCO-NRT, ECCO-v4, K7-ODA and K7-CDA; Coupled models are used in ECDA, K7-CDA and MOVE-C; Relatively high horizontal resolutions are adopted in G2V3, C-GLORS, UR025.4, GloSea5, ORAP5 and ECCO-v4. If similar MLD features are exhibited within the groups, important information for improving the systems can be provided.
As for the observation-only analyses, monthly MLD and ILD time series are calculated from the monthly mean TS fields on the native grids and interpolated onto the common longitude-latitude grids with one-degree resolution.

In addition, the ensemble mean of the 17 model based reanalysis MLDs/ILDs are calculated ("ENSMEAN"; not including EN3v2a and ARMOR3D). Note that these MLDs/ILDs differ from MLDs/ILDs calculated from the ensemble mean TS fields. In order to reduce the influence of the difference in period among the reanalyses (Table 3), the MLD/ILD time series for

ENSMEAN are calculated as follows: Monthly climatologies averaged over the period of 2001-2011 (or longest available during this period hereafter) and monthly interannual anomalies from these climatologies are first computed for the individual ensemble members. Using these climatologies and interannual anomalies, the monthly climatologies and interannual anomalies (for 1948-2012) of ENSMEAN are calculated respectively. The absolute MLD/ILD time series for ENSMEAN are produced as the sum of these monthly climatologies and interannual anomalies.

3 Uncertainties in observation-only analyses

Figure 1 exhibits the zonal-mean monthly MLD/ILD normalized differences of EN3v2a, ARMOR3D, deBoyer, WOA and ENSMEAN from MILA-GPV. (Note that this does not mean that MILA-GPV is true.) The differences between deBoyer and MILA-GPV (MLDr003 and ILDt02) are generally small, since these are comparable datasets that use individual TS profiles. Relatively large differences at high latitudes may possibly result from spatio-temporally limited observations there, especially from the Argo floats. MLDs/ILDs for WOA, EN3v2a and ARMOR3D exhibit biases toward shallower depths. ILDt02s in WOA are 20 to 40% shallower than those in MILA-GPV globally, which is consistent with the result of de Boyer Montégut et al. (2004). They indicated that the global shallow biases are attributed to the fact that MLD/ILD calculated from averaged TS profiles is more strongly affected by profiles from which shallower MLD/ILD are estimated. Therefore, it can be considered that the smaller discrepancies from MILA-GPV for EN3v2a and ARMOR3D than those for WOA, as shown in Fig. 1, are due to their use of the monthly mean TS profiles as opposed to the climatologically averaged TS profiles as in WOA.

As described above, the temperature criterion ($\Delta T = 0.5$ºC) generally correspond to the density criterion ($\Delta\rho = 0.125$ kg m^{-3}). These criteria give similar patterns for each of the observation-only analyses except that large positive biases are seen at about 60ºS-40ºS and 40ºN-60ºN in winter-early spring for the temperature criterion cases. We will discuss these biases later (subsection 3.3).

By using larger values for the criterion ($\Delta\rho = 0.125$ kg m^{-3} and $\Delta T = 0.5$ºC), generally similar patterns to those with smaller values ($\Delta\rho = 0.03$ kg m^{-3} and $\Delta T = 0.2$ºC) are obtained for WOA, EN3v2a and ARMOR3D, respectively, but the amplitudes of the negative biases are much reduced. On the other hand, the positive biases at mid- and high latitudes are enhanced. For ENSMEAN, the discrepancies from MILA-GPV are considerably smaller than those for WOA, EN3v2a and ARMOR3D for each of the criteria. The change in vertical resolution of profiles can be an error source as well as averaging of profiles as indicated by de Boyer Montégut et al. (2004). How these errors differ according to the criterion values is also an important question. In addition, representation of the interannual variability is relevant for climate studies as well as the climatology. These are quantitatively analyzed in the following subsections.

3.1 Errors due to averaging of profiles

de Boyer Montégut et al. (2004) previously revealed that averaging of profiles can lead to underestimations of MLD (shallower biases). In this subsection, we investigate the influence of the time average by comparing the MLD/ILD estimates from monthly mean ("m") and instantaneous ("i") TS profiles ("MLDr003m" and "MLDr003i" and likewise). Since both of the above TS profiles are now provided by the MOVE-G2 experiment, our comparison here focuses on the influence of the time average of profiles. In addition, interannually averaged monthly TS profiles (like climatologies) are also used to estimate the MLDs/ILDs.

Underestimations in the zonal mean of greater than 10 m are seen in case of the monthly mean profiles at mid-high latitudes in March-May (September-December) in the Northern (Southern) Hemisphere (Fig. 2a, b). These are attributed to the re-stratification process of deep wintertime MLDs/ILDs in the Kuroshio Extension region, in the south of the North Atlantic Current and in the Southern Ocean and are generally 10-20 m (Fig. 2c, d). Note that larger biases can be seen in the sea ice region. In addition, use of the climatological profiles (averaged over the 2001-2011 period) results in further underestimation of MLD, especially in the tropics, where TS profiles vary greatly in association with El Nino and Southern Oscillation (not shown).

In the latitudes of 20º-30º, underestimations from this effect are enlarged in March (September) in the Northern (Southern) Hemisphere (e.g., Fig. 2a). A previous study (Takeuchi and Yasuda 2003) identified the MLD shoaling from February to March (from August to September) in a large part of this latitude band in the Northern (Southern) Hemisphere, despite the fact that monthly mean net surface heat flux is cooling the ocean surface. Since they used the averaged profiles (e.g., WOA 1998), the MLD shoaling may partially be explained by the above underestimations brought about by monthly averaging. Note that Takeuchi (2006) discussed the possible effect of variability of surface heat flux within a month by using a simple ML model.

Similar effects can be expected from averaging of profiles over time as above and within a grid cell, when the length scale calculated from the typical advection speed (e.g., $1 \text{ cm s}^{-1} \times 1 \text{ month} \sim 26 \text{ km}$) is comparable to the grid spacing. Hence the above estimation for the effect of averaging over time by using the MOVE-G2 result might be different from that for the effect of averaging within a grid cell. In addition, the impact of temporal averaging of profiles in the estimation of monthly MLD/ILD may be affected by the amount of high frequency variability, which in turn may be affected by horizontal resolution. To address this question we have used G2V3 which has a finer horizontal resolution (1/4º) than MOVE-G2 (1º zonally and 0.3-0.5º meridionally). MLD/ILD estimates from monthly and daily mean TS profiles for an older version of G2V3 have been compared. This comparison generally supports the above-described underestimations of 10-20 m in early spring (not shown). Note that both reanalyses (MOVE-G2 and G2V3) assimilated the satellite-derived SLA observations (Table 2).

It should be noted that profiles from the real observations would have further variability on smaller scales, which cannot be resolved in OGCMs. The averaging of these profiles may cause the underestimation of MLD in the same way as indicated by de Boyer Montégut et al. (2004). Therefore, the broad tendency of larger MLDs in ENSMEAN than in the observation-only analyses as shown in Fig. 1 can be attributed to this effect partly.

Horizontal resolution can affect not only representation of the eddy-scale variability as discussed above but also averaging area of TS profiles for the MLD/ILD estimation. In order to investigate the latter effect, we compare the MLDs (MLDr003m and MLDr0125m) estimated from monthly TS profiles in the MOVE-G2 experiment. Three TS profiles on the one-degree resolution grids are used for the MLD estimation: 1) those interpolated from the MOVE-G2 grids (e.g., "MLDr003m_1x1"), 2) those smoothed by a 9-point filter after the interpolation (e.g., "MLDr003m_3x3") and 3) those smoothed by a 25-point filter after the interpolation (e.g., "MLDr003m_5x5"). Note that the smoothed profiles correspond to the profiles from low resolution analysis. As shown in Fig. 3, shallower MLDs are estimated when the smoothed (low resolution) TS profiles are used. This effect appears mostly in winter, in contrast to the effect of the time averaging (early spring; Fig. 2). Larger-scale smoothing results in greater magnitude of shoaling for both MLDr003m and MLDr0125m. While errors resulting from the smoothing at high latitudes are larger with the larger criterion ($\Delta\rho = 0.125$ kg m^{-3}), errors at mid-latitudes are smaller with the larger criterion. Although various resolutions (about 1/4-1°) are adopted for the reanalyses in this study, a tendency of shallower MLDs for reanalyses with lower resolutions is not seen as shown later (Section 4). Therefore, while horizontal resolution finer than 1° seems not to much influence the MLD estimation, the coarser resolutions (such as 3° and 5°) can largely affect the estimation.

3.2 Effect of vertical resolution

The average vertical resolutions of the observational profiles are 8.2 m, 2.3 m, 19.5 m and 9.4 m for profiling floats, CTD (Conductivity-Temperature-Depth), XBT (eXpendable BathyThermograph) and MBT (Mechanical BathyThermograph) measurements, respectively (de Boyer Montégut et al. 2004), whereas those for mooring arrays are usually about 20 m. On the other hand, the vertical resolutions of TS profiles in the syntheses can be much lower as shown in Fig. 4. For example, the vertical resolution of the WOA data is 25 (50) m at 50-150 (150-300) m depth. The low resolution of TS profiles can also be an error source in the MLD estimation.

Figure 5 shows schematic illustrations for the MLD estimations from high and low resolution profiles. In the high resolution case (Fig. 5a) and low resolution case (1) (Fig. 5b), the estimated MLDs are comparable to the "real" MLD (from the common simultaneous profile). On the other hand, in the low resolution case (2) (Fig. 5c), the estimated MLD is much shallower than the real MLD. Thus, MLD can be underestimated by using a low resolution vertical discretization depending on the relative position of grids to the real MLD. In order to quantitatively assess this effect, we generalize the MLD estimation as in Fig. 5d. Since we assume that we have at least one grid point in the thermocline, which should hold for the most of the regions and resolutions we investigate here, this simplified model does not contain a low-stratified layer beneath the seasonal thermocline. We note that if the thermocline is not resolved with at least one grid point overestimation of MLD can also happen. In fact, larger MLDs are estimated from the averaged profiles than from the individual profiles, in particular with larger criteria, at high latitudes in winter. There, the averaging of the profiles with a relatively weak thermocline results in a weak thermocline being represented in the syntheses, leading to an overestimation of MLD/ILD, as indicated by Noh and Lee (2008).

Here, the estimation error, *e*, is determined by using level spacing, Δz, relative position of the bending point of the profile to the grid, *r*, and vertical density gradient of the seasonal thermocline, $\frac{\partial \rho}{\partial z}$, as

$$e = \begin{cases} \Delta z(1-r) \times \dfrac{\frac{\partial \rho}{\partial z}\cdot \Delta z \cdot r - \Delta\rho}{\frac{\partial \rho}{\partial z}\cdot \Delta z \cdot r} & \text{for } c \equiv \frac{\Delta\rho}{\Delta z} \Big/ \frac{\partial \rho}{\partial z} < r < 1 \\ 0 & \text{for } 1 < r < 1 + c \end{cases} \quad (1)$$

where *c* is the lower limit of *r*, and is given by the relation that the density at a grid point is larger than the reference (10-m) value by $\Delta\rho$ exactly. (Although the true MLD would be the bending point for this profile, we use the "real MLD" based on the density criterion here considering the practical use with rather noisy profiles.) The expected error is then estimated as

$$\bar{e} = \int_{c}^{1+c} e dr \quad (2)$$

Figure 6 exhibits the expected error values depending on the vertical resolution for $\Delta\rho = 0.03$ and 0.125 kg m^{-3}, where Δz of 10 m corresponds to the vertical resolution of typical observations. For $\Delta\rho = 0.03$ kg m^{-3}, the shallower errors at low- and mid-latitudes are within 5-7 (14-18) m with $\Delta z = 25$ (50) m (Fig. 6a). In case of $\Delta\rho = 0.125$ kg m^{-3}, the errors are much smaller (e.g., 3-5 m with $\Delta z = 25$ m) and errors for the observational profiles are less than 1 m (Fig. 6b). In contrast to the errors indicated in subsection 3.1, the distribution of the errors from vertical resolution is rather broad in terms of time and space (e.g., Fig. 6c).

The mean differences of EN3v2a/ARMOR3D from MILA-GPV (over 40ºS-50ºN and the 2001-2011 period) are -11.4 and -8.2 m for $\Delta\rho = 0.03$ and 0.125 kg m^{-3}, respectively. These values are between the above estimations with $\Delta z = 25$ and 50 m (blue and red lines respectively in Fig. 6a, b) and generally consistent with the resolutions of these data at the 100-300 m depth (Fig. 4). Note that the Δz values vary with depth (and per dataset) and also that use of vertical covariances of background errors (or smoothing) in the analyses can make the resolution of the represented vertical variability coarser than the level spacing.

Differences between ENSMEAN and MILA-GPV are generally smaller than differences between EN3v2a/ARMOR3D and MILA-GPV (Fig. 1) as described above. This can be largely attributed to the higher vertical resolutions in the reanalyses, although the possible effect of small scale variability in the real observation data might cause the shallower MLDs in the observation-only analyses partly (subsection 3.1). Figure 4 represents the vertical resolution for each depth and synthesis. The level spacing of the reanalyses is generally less than 20 m at depths important for the ML variation (approximately the upper 200 m), which works well for the relatively small errors in the MLD and ILD fields of ENSMEAN. Note that relatively large ensemble size in this study should contribute to the generally small model errors in ENSMEAN.

3.3 Overestimation of ILDs in the subpolar regions

In addition to the aforementioned underestimations relative to MILA-GPV, overestimation of wintertime ILDs (biased-deep) estimated from the monthly mean TS profiles are seen in the subpolar regions. Figure 7 shows the ILDt05 distributions in March for MILA-GPV, WOA, EN3v2a and ARMOR3D. ILDs deeper than 400 m can be widely seen in the subarctic North Pacific in WOA, EN3v2a and ARMOR3D, while are only seen at a few grid points in MILA-GPV. Similar overestimation occurs in the Southern Ocean in austral winter (not shown).

In these subpolar regions, stratification is mostly determined by the halocline (e.g., Yuan and Talley 1996) and the thermocline is weak especially in winter (e.g., Dodimead 1967). A surface isothermal layer extending to the mesothermal layer (intermediate warm layer) can appear when the dicothermal structure (intermediate cold profile) weakens seasonally as shown, for example, in Fig. 8. However, since temperature in the surface layer changes rather rapidly if resolved with enough temporal resolution, this occurs only during a short period (Fig. 8a). On the other hand, the monthly data represent the occurrence of thick isothermal layer in a whole month, resulting in the overestimation greater than 100 m (Fig. 8b).

3.4 Interannual variability

A comprehensive assessment of the interannual anomaly field on a global scale is rather difficult due to limited independent observations. In this study, we assume white noise for the interannual anomaly field in each of the datasets and thereby investigate correlations of interannual signals between the datasets. Figure 9 presents the zonal mean correlation coefficients for interannual anomaly components in the data rich Argo period, using monthly data (seasonal cycle removed) for 2005-2011, among MILA-GPV, EN3v2a and ARMOR3D. Correlation coefficients (for MLDs in particular) are small at high latitudes presumably due to the limited number of observations there. At low- and mid-latitudes (about 50°S-60°N), values for MLDr003 and ILDt02 generally locate between 0.15 and 0.4 (Fig. 9a, b), whereas those for MLDr0125 and ILDt05 between 0.3 and 0.6 (Fig. 9c, d). Thus the interannual signals are more consistently represented when using the larger criteria ($\Delta\rho = 0.125$ kg m^{-3} and $\Delta T = 0.5$°C).

In addition, correlation coefficients between EN3v2a and ARMOR3D (red) are larger than those between MILA-GPV and EN3v2a/ARMOR3D (light blue/light green) for MLDr0125 and ILDt05 (Fig. 9c, d). This suggests that the signal-to-noise ratio of the interannual anomalies are relatively low in MILA-GPV possibly due to the limited coverage by Argo floats. The correlation coefficients for MLDr0125 between EN3v2a and ARMOR3D (red) are lower by 0.1-0.2 at low latitudes than at mid-latitudes. This reduction of the correlation at low latitudes does not occur for ILDt05. This result implies that salinity analysis in the tropics may not be well constrained by observations. In fact, the lower correlation for MLDr0125 at low latitudes mainly results from relatively low correlation in the western Pacific warm pool region and in the Intertropical Convergence Zone (not shown), where salinity plays an important role in determining the surface-layer stratification due to large freshwater input to the ocean. Note that, although Shi et al. (An assessment of upper ocean salinity content from the Ocean Reanalyses Inter-comparison Project (ORA-IP), submitted to the same issue of Climate Dynamics, 2015) show that the averaged salinities over the 0-700 m depth in this region (156ºE, 8ºN) from the reanalyses are generally consistent with the observations by the TRITON buoy, our result indicates that the vertical salinity distribution in the upper ocean is still uncertain. Figure 10 compares the MLD/ILD time series in the western Pacific warm pool region (150ºE-180º, 5ºS-5ºN). It is demonstrated that MLDr0125 time series (Fig. 10a) are much less consistent with each other than the ILDt05 time series (Fig. 10b), the latter of which are well constrained by relatively rich observations for temperature profiles by mooring arrays in the tropics. Variability in MLDr0125 of ARMOR3D (green line) is relatively weak on both the seasonal and interannual time scales. Although grid-scale correlation between MILA-GPV and EN3v2a is rather lower (Fig. 9c), the area-mean values exhibit similar interannual variations, especially for the period of the correlation analysis (2005-2011).

Zonal mean correlation coefficients between ENSMEAN and other datasets are shown in Fig. 9e, f. For both MLDr0125 and ILDt05, correlation coefficients between ENSMEAN and EN3v2a/ARMOR3D (blue/green) are greater than those between ENSMEAN and MILA-GPV (yellow) at low- and mid-latitudes. This fact also suggests relatively low signal-to-noise ratio of the interannual variability in MILA-GPV as described above. Note that distortion of the monthly MLD distributions by Argo sampling was reported in previous studies (e.g., Juza et al. 2012). On the other hand, the variability in MILA-GPV is by and large consistent with the variability in ENSMEAN in terms of area-mean values, especially for the 2005-2011 period (e.g., Fig. 10).

Correlation coefficients between ENSMEAN and EN3v2a/ARMOR3D (blue/green) are generally greater than those between EN3v2a and ARMOR3D (red) as shown in Fig. 9e, f. This suggests better representation of interannual signals in ENSMEAN, which can be attributed to the use of atmospheric information as surface forcing. Independent validations will ensure the effectiveness of the ensemble use of reanalyses in detecting interannual variability, which awaits future work.

4 Intercomparison of the reanalysis MLDs/ILDs

4.1 Seasonal and interannual variations of MLDs

Seasonal variations of MLDs are basically characterized by the winter- and summer-time MLD features. Following the discussion in subsection 3.1, MLDs in February and August are compared among the syntheses as typical of MLDs in winter and summer respectively with relatively small errors (e.g., MLDs in March are considered to have larger errors). Figures 11 and 12 show that the MLDr0125 discrepancies from MILA-GPV are relatively large in the winter hemisphere on a basin scale for both the observation-only analyses and reanalyses. Distributions for MLDr003, ILDt02 and ILDt05 (figures not shown) are generally similar to those for MLDr0125.

In February (Fig. 11), positive discrepancies are seen in the Kuroshio Extension and recirculation regions in most of the reanalyses, where common biases are known in coarse resolution models (e.g., Hasumi et al. 2010). Similar positive discrepancies are exhibited in the Gulf Stream recirculation region. Positive discrepancies can also be seen in the Arabian Sea and Bay of Bengal for all the renalyses except G2V3 and ECCO-v4. In addition, in many reanalyses, negative discrepancies are seen in the Southern Ocean (in austral summer) such as reported in several studies (e.g., Gnanadesikan et al. 2006; Noh and Lee 2008). Note that MLDs become larger from ORAS4 to ORAP5 in this region probably by incorporating the effects of the wave breaking and Langmuir circulation (Janssen 2012), although other updates of the system (e.g., horizontal resolution) may also have a contribution. Since the above discrepancies are not seen in the observation-only analyses (EN3v2a and ARMOR3D), these can be considered as weakness of the reanalyses that need to be improved.

Negative discrepancies from MILA-GPV are predominant in the North Atlantic Current region for many syntheses including the observation-only analyses. These can be attributed to overestimations of the monthly MLDs by Argo sampling in winter in this region (up to about 50 m) indicated by Juza et al. (2012) at least partly. It is considered that, in the subpolar North Atlantic, both limited observations and poor representations by models affect the large ensemble spread of the syntheses which is defined in this study as the standard deviation from the ensemble mean (e.g., Xue et al. 2012). However, since the absolute MLD values are several times larger in this region than in other regions, normalized ensemble spread values larger than 0.4 generally occur in limited regions near the coast.

Weak negative discrepancies in the tropics for all the syntheses are consistent with our results described in subsection 3.2. Although negative discrepancies in the subarctic North Pacific for ENSMEAN are comparable to those for EN3v2a and ARMOR3D, their amplitudes for the individual reanalyses differ greatly: negative discrepancies larger than the ENSMEAN values in G2V3, GECCO2, ECCO-NRT, ECDA, K7-ODA, K7-CDA, MOVE-G2 and MOVE-CORE; positive discrepancies in C-GLORS, UR025.4 and GloSea5. The overall patterns are similar in C-GLORS, UR025.4 and GloSea5, suggesting that, in this cluster, the choice of model (NEMO3.2), forcing (ERA-interim (ERAi)) and resolution (Table 2) plays a dominant role in determining the MLD. Note that horizontal resolution often influences the tuning of parameterizations such as isopycnal diffusivity that is important to the representation of the stratification in the subpolar regions. Although G2V3, ORAS4 and ORAP5 also use the NEMO model and ERA-interim forcing, a bias correction scheme (Balmaseda et al. 2007) might work to reduce the above positive biases. Note that, although the ERA-interim forcing is also used for ECCO-v4, it is corrected through the 4DVAR approach. Another cluster includes MERRA, ECDA and PEODAS, which commonly use the MOM models and also show similar patterns. Although K7-ODA and K7-CDA also use the MOM3, the assimilation method they use (4DVAR) appears to have stronger effect on the MLD patterns.

In August (Fig. 12), the large MLDs around the Antarctic Circumpolar Current (ACC) region as represented in MILA-GPV are estimated to be smaller in most of the reanalyses. In addition, the amplitudes of the discrepancies are remarkably different between the datasets, and hence the ensemble spread in this region is relatively large. Improvement in representing MLDs in this region is needed in the future. To do so, intercomparisons for sea ice (e.g., Smith et al. 2014) and surface flux (e.g., Valdivieso et al. 2014) might indicate important clues in association with deep convection following the sea ice formation by strong cooling. Further observations in the Southern Ocean, particularly for the sea ice region, are also important, indicated by large differences between MILA-GPV and EN3v2a/ARMOR3D with their signs changing on small scales.

Positive discrepancies in the mid-latitude South Pacific and Atlantic seen in C-GLORS, UR025.4 and GloSea5 (Fig. 12) are similar to those in the North Pacific in boreal winter (Fig. 11). In the South Indian Ocean, positive discrepancies are also seen in most of the reanalyses, which can be attributed partly to a weak representation of the Agulhas retroflection in coarse resolution models as reported by previous studies (e.g., Morioka et al. 2012). Such difficulty in representing the MLD variability is also seen in several reanalyses in the confluent region between the Brazil and Malvinas Currents. The above limitations are also found in the intercomparison of the salinity fields (Alves et al. 2014).

Negative discrepancies in the western tropical Pacific are observed in several reanalyses. Previous studies have reported that precipitation inputs derived from atmospheric reanalysis datasets are much larger than satellite-based estimates in this region (Iwasaki et al. 2014). On the other hand, evaporation for OGCMs is usually estimated from the bulk formula by using simulated SST as in the reanalyses here. The model experiment using the above freshwater fluxes often generates a too strong halocline, thereby leading to negative MLD biases. Assimilation of the recent sea surface salinity observations from satellites is likely to reduce these biases (e.g., Köhl et al. 2014, Toyoda et al. 2015). Furthermore, better results from the reanalyses with ocean-atmosphere coupled models (ECDA and MOVE-C) suggest an advantage of these approaches (e.g., Fujii et al. 2009) as they may eliminate some of the uncertainties associated with precipitation forcing from atmospheric analysis. Obviously, further validation studies for the freshwater fluxes reproduced or corrected in the reanalyses are necessary. This would be addressed in the ORA-IP.

The skillful reproduction of the interannual variability in the ENSMEAN MLDr0125 field is demonstrated in subsection 3.4. Here, we intercompare the correlation of the individual reanalyses with ENSMEAN. Figure 13 shows the distributions of correlation coefficients for the interannual anomalies of MLDr0125 between ENSMEAN and the individual reanalyses/syntheses. Correlation is low at high latitudes as seen for the observation-only analyses (Fig. 9e). At low and mid-latitudes, relatively large correlation can be seen in ORAS4 and MOVE-G2. Since both reanalyses use coarse resolution (about 1°) and the 3DVAR method, higher resolution or sophisticated assimilation method does not always improve the reproduction of the interannual variability. Relatively small correlation coefficients are seen in the Kuroshio Extension, Gulf Stream, Agulhas retroflection and Malvinas Current regions, where discrepancies in mean values from MILA-GPV are also large (Fig. 11, 12). In the regions relevant to El Nino and Southern Oscillation and the Pacific decadal oscillation (e.g., Mochizuki et al., 2010), the correlation coefficients are relatively large for most of the reanalyses, which encourage possible applications of the reanalysis datasets to the studies on these oscillations.

For both observation-only analyses (bottom row in Fig. 13), correlation coefficients at mid-latitudes are higher for MLDr0125 than for ILDt05, as seen in Fig. 9e, f, although these are generally smaller than those between ENSMEAN and the reanalyses. As also seen in Fig. 9e, the correlation coefficients for MLDr0125 are lower at low latitudes than at mid-latitudes. It is shown in Fig. 13 that this arises from relatively low values in the regions of low surface salinity (e.g., the Indonesian maritime continent and the region of the Amazon River plume). In particular, the correlation for MLDr0125 between ENSMEAN and ARMOR3D is remarkably low, consistent with the result shown in Fig. 10. Surface salinity observations from satellites will possibly contribute to the improvements of the analyses in these regions (e.g., Toyoda et al., 2015). Also, observations by the mooring array are required to be maintained in order to keep or enhance the quality of the reanalyses/syntheses.

4.2 Barrier layer thicknesses

The barrier layer is the isothermal layer below the ML that prevents cooling of the ML by entrainment of the underlying waters (e.g., Lukas and Lindstrom 1991). Therefore, the barrier layer thickness (BLT), which is usually deduced from the MLD and ILD, is an important parameter for the surface heat budget in climate studies. In this subsection, BLTs from the syntheses are examined to evaluate the integrated reproduction of the MLDs and ILDs. BLT is defined in this study as difference between MLDr0125 and ILDt05 only when ILDt05 is larger than MLDr0125 (e.g., Maes et al. 2006).

Figure 14 shows the distributions of the BLTs from the syntheses. Results are only displayed for low- and mid-latitudes because of the issues with ILDs at high latitudes (subsection 3.3). Large BLTs present in the western equatorial Pacific and Atlantic and the north-eastern Indian Ocean are represented in MILA-GPV, EN3v2a and ARMOR3D. BLTs in the western Pacific and eastern Indian Oceans are larger in EN3v2a and ARMOR3D than in MILA-GPV. In objective analyses, a large zonal correlation scale may be used owing to fast wave speeds in the tropics. On the other hand, a recent study using the Argo data that the barrier layer develops and ceases on a rather shorter time scale (Katsura and Oka 2014). This fact implies the presence of smaller-scale correlations there due to a confined distribution of intense BLTs as in MILA-GPV.

In addition to the formation of BLTs associated with excess precipitation (and also river discharge) over evaporation in the tropics, another generating mechanism associated with subduction of the salty subtropical waters can be seen around the boundaries between tropics and subtropics. For example, pronounced BLTs around 10°N -20°N in the North Pacific are thought to come from both the North Pacific tropical water (e.g., Suga et al. 2000) and eastern subtropical mode water (e.g., Toyoda et al. 2004). These BLTs are greater in ARMOR3D than in MILA-GPV and EN3v2a in both North and South Pacific. This might result from the use of vertical modes in synthesizing satellite data (e.g., SLA and SST) as well as in-situ TS profiles in ARMOR3D (Table 1).

The reanalyses basically reproduce the above features of the BLT distribution seen in the observational datasets. For example, the important aspects of BLTs in the western tropical Pacific are well represented in ENSMEAN, comparable to MILA-GPV in a quantitative sense. Note that BLTs are relatively small in GECCO2, ECCO-NRT and K7-ODA, which all use the smoother approaches (e.g., 4DVAR and KS). This suggests that the control of the salinity field (such as by adjusting the surface freshwater fluxes) is still challenging in the smoother approach. In contrast, BLTs in the coupled reanalyses (ECDA, K7-CDA and MOVE-C) are quantitatively comparable to those in MILA-GPV and ENSMEAN, which is encouraging for the use of coupled models.

5 Summary and discussion

In the present study, we have investigated the fidelity of a suite of global ocean synthesis products in representing the MLD/ILD fields which are recognized as an important element in the ocean circulation system. These syntheses, including 2 observation-only analyses and 17 reanalyses which assimilate data into models, have been provided by operational and research centers as an international action of the ORA-IP (Balmaseda et al. 2014). First, we compared the observation-only analyses with reference datasets (de Boyer Montégut et al. 2004; Hosoda et al. 2010) that determined MLD directly from individual TS profiles. The purpose is to investigate the errors in estimating MLDs/ILDs unrelated to model errors. Negative biases are seen in the MLDs/ILDs from monthly mean and gridded TS profiles of the above syntheses with respect to those from individual profiles as reported by de Boyer Montégut et al. (2004). It is revealed that these underestimations from the averaging procedure of profiles are associated with a rapid re-stratification process to the relatively deep ML state in early spring and estimated to be approximately 10-20 m. In addition, negative biases are generated depending on the vertical resolutions of profiles, which are distributed more broadly in time and space. When the criterion $\Delta\rho = 0.03$ kg m^{-3} is used, the underestimations from this effect are estimated as 5-7 (14-16) m for the vertical resolution $\Delta z = 25$ (50) m. On the other hand, they are generally much smaller in case of the larger criterion $\Delta\rho = 0.125$ kg m^{-3}. Furthermore, considerable overestimations (greater than 100 m) of the wintertime ILDs from the monthly mean profiles are seen in subpolar regions in association with the mesothermal structure.

Discrepancies between the ensemble mean obtained from the 17 sets of reanalyses and the reference datasets are noticeably smaller than the observation-only analyses in many regions where model errors in the individual reanalyses are mutually canceled out through the ensemble averaging approach. This can be attributed mainly to the higher vertical resolutions of the reanalyses in reproducing the MLDs as well as the large ensemble size. The results (e.g., Fig. 1) show, on the other hand, that there exist a few regions where model errors are not canceled by the ensemble mean (such as the Kuroshio Extension and ACC regions). Such common model errors possibly arise from the coarse horizontal resolutions as reported in previous studies (e.g., Hasumi et al. 2010). Interannual variability is better represented in both the analyses using all available TS observations (including Argo data) and the ensemble mean of the reanalyses than the analysis using the Argo data only, especially when the larger values are used for the criteria ($\Delta\rho = 0.125$ kg m^{-3} and $\Delta T = 0.5^{\circ}$C).

Differences in the individual syntheses were then assessed by intercomparing the winter and summer MLDs/ILDs together with the BLTs. The result shows that, in addition to the consistent features between the reanalyses, differences can also be seen depending on the configurations of the reanalyses. Features seen in the reanalyses with similar configurations offer interesting information toward the improved reanalyses (e.g., the change in mixing parameterization from ORAS4 to ORAP5) and also suggests a limit to the effectiveness of the ensemble mean approach if similar reanalyses are included. At high latitudes, consistency among all the observational and reanalysis datasets is relatively low in terms of both the seasonal cycle and interannual variability. Therefore, observational studies that fully describe the high-latitudinal variability are required as well as further improved modeling and assimilation techniques toward the enhanced syntheses.

Although MLD/ILD data themselves are not assimilated in the syntheses here, the estimated MLDs/ILDs are primarily influenced by the assimilated TS profile data. Therefore, regarding the observational MLDs/ILDs as independent references might yield little extra effect. In addition, smaller biases in a product do not necessarily indicate its superiority in other aspects.

Their magnitude depends on how strongly the observations of TS are used to constrain the model locally in time and space. For example, the smoother approach does not insert local TS corrections but tends to retain the model dynamics while trying to fit the model to the data, which makes an MLD representation close to the observational data rather challenging. Hence, it is important to make an accurate assessment of each synthesis from various aspects, particularly with independent data. However, taking into consideration the fact that the MLD is a key parameter in determining the upper ocean processes, which greatly influences other variables, the validation and intercomparison of the MLD fields of the syntheses in this study are of value for the communities of both model developers and users. In particular, the robust reproduction of both the seasonal cycle and interannual variability of MLD by the ensemble mean of the renalyses indicates a great potential of the ensemble mean MLD field for better investigating and monitoring the upper ocean processes, together with other intercomparison results such as for the surface forcing and heat content variability (e.g., Palmer et al. 2014). Information on uncertainty derived from the ensemble spread should allow further quantitative discussion for results derived from the ensemble mean (e.g., Xue et al. 2012).

Acknowledgements

We thank two anonymous reviewers for their constructive comments. The MILA-GPV dataset (Hosoda et al. 2010) is provided by the RCGC/JAMSTEC from their web site at http://www.jamstec.go.jp/ARGO/argo_web/MILAGPV/. The MLD dataset of de Boyer Montégut et al. (2004) is obtained from his web site at http://www.ifremer.fr/cerweb/deboyer/mld/home.php. The Argo float data are provided by the

NODC/NOAA at their web site http://www.nodc.noaa.gov/OC5/WOD13/. This work was partly supported by the Research Program on Climate Change Adaptation (RECCA) of the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government (MEXT), by the Data Integration and Analysis System (DIAS) of the MEXT, by the joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101), by the UK Public Weather Service Research Programme, and by the European Commission funded projects MyOcean (FP7-SPACE-2007-1) and MyOcean2 (FP7-SPACE-2011-1). During the preparation of this article, our co-author Nicolas Ferry passed away. He was an active and supportive member of the ORA-IP and CLIVAR-GSOP activities.

References

Alves O, Shi L, Wedd R, Balmaseda M, Chang Y, Chepurin G, Fujii Y, Gaillard F, Good S, Guinehut S, Haines K, Hernandez F, Lee T, Palmer M, Peterson KA, Masuda S, Storto A, Toyoda T, Valdivieso M, Vernieres G, Wang X, Yin Y (2014) An assessment of upper ocean salinity reanalyses from CLIVAR GSOP/GODAE system. CLIVAR EXCHANGES 64:11-14

Balmaseda MA, Dee D, Vidard A, Anderson DLT (2007) A multivariate treatment of bias for sequential data assimilation: application to the tropical oceans. QJR Meteorol Soc 133:167–179. doi:10.1002/qj.12

Balmaseda MA, Hernandez F, Storto A, Palmer MD, Alves O, Shi L, Smith GC, Toyoda T, Valdivieso M, Barnier B, Behringer D, Boyer T, Chang YS, Chepurin GA, Ferry N, Forget G, Fujii Y, Good S, Guinehut S, Haines K, Ishikawa Y, Keeley S, Köhl A, Lee T, Martin M, Masina

S, Masuda S, Meyssignac B, Mogensen K, Parent L, Peterson KA, Tang YM, Yin Y, Vernieres G, Wang X, Waters J, Wedd R, Wang O, Xue Y, Chevallier M, Lemieux JF, Dupont F, Kuragano T, Kamachi M, Awaji T, Caltabiano A, Wilmer-Becker K, Gaillard F (2015) The Ocean Reanalyses Intercomparison Project (ORA-IP). J Operational Oceanogr (in press). doi:10.1080/1755876X.2015.1022329

Blanke B, Delecluse P (1993) Variability of the tropical Atlantic ocean simulated by a general circulation model with two different mixed-layer physics. J Phys Oceanogr 23:1363-1388

Chang YS, Zhang S, Rosati A, Delworth TL, Stern WF (2013) An assessment of oceanic variability for 1960–2010 from the GFDL ensemble coupled data assimilation. Clim Dyn 40:775-803. doi:10.1007/s00382-012-1412-2

Chen D, Busalacchi AJ, Rothstein LM (1994) The roles of vertical mixing, solar-radiation, and wind stress in a model simulation of the sea-surface temperature seasonal cycle in the tropical Pacific-Ocean. J Geophys Res 99(C10):20345-20359

Danabasoglu G, co-authors (2014) North Atlantic simulations in coordinated ocean-ice reference experiments phase II (CORE-II). Part I: Mean state. Ocean Modell 73:76-107. doi:10.1016/j.ocemod.2013.10.005

de Boyer Montégut C, Madec G, Fischer AS, Lazar A, Iudicone D (2004) Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. J Geophys Res 109:C12003. doi:10.1029/2004JC002378

Dodimead AJ (1967) Winter oceanographic conditions in the Central Subarctic Pacific. Int North Pacific Comm 999:1-14

Ferry N, Parent L, Garric G, Bricaud C, Testut CE, Le Galloudec O, Lellouche JM, Drévillon M, Greiner E, Barnier B, Molines JM, Jourdain N, Guinehut S, Zawadzki L (2012) GLORYS2V1 global ocean reanalysis of the altimetric era (1993-2009) at meso scale. Mercator Ocean Newsletter 44:28-39.

Fujii, Y, Nakaegawa N, Matsumoto S, Yasuda T, Yamanaka G, Kamachi M (2009) Coupled climate simulation by constraining ocean fields in a coupled model with ocean data. J Clim 22:5541-5557

Gnanadesikan A et al. (2006) GFDL's CM2 global coupled climate models. Part II: The baseline ocean simulation. J Clim 19:675-697. doi:10.1175/JCLI3630.1

Guinehut S, Dhomps AL, Larnicol G, Le Traon PY (2012) High resolution 3D temperature and salinity fields derived from in situ and satellite observations. Ocean Sci 8:845-857. doi:10.5194/os-8-845-2012

Hosoda S, Ohira T, Sato K, Suga T (2010) Improved description of global mixed-865 layer depth using Argo profiling floats. J Oceanogr 66:773-787. doi:10.1007/s10872-866 010-0063-3

Ingleby B, Huddleston M (2007) Quality control of ocean temperature and salinity profiles -

historical and real-time data. Journal of Marine Systems 65:158-175. doi:10.1016/j.jmarsys.2005.11.019

Iwasaki S, Kubota M, Watabe T (2014) Assessment of various global freshwater flux products for the global ice-free oceans. Rem Sens Env, 140:549-561. doi:10.1016/j.rse.2013.09.026.

Janssen PAEM (2012) Ocean wave effects on the daily cycle in SST. J Geophys Res 117:C00J32. doi:10.1029/2012JC007943

Juza M, Penduff T, Brankart JM, Barnier B (2012) Estimating the distortion of mixed layer property distributions induced by the Argo sampling. J Oper Oceanogr 5:45-58.

Katsura S, Oka E (2014) Formation mechanism of winter barrier layer in the subtropical Pacific. 2014 Ocean Science Meeting, 23-28 February 2014, Honolulu Hawaii USA

Köhl A (2014) Evaluation of the GECCO2 ocean synthesis: transports of volume, heat and freshwater in the Atlantic. QJR Meteorol Soc. doi:10.1002/qj.2347

Köhl A, Sena Martins M, Stammer D (2014) Impact of assimilating surface salinity from SMOS on ocean circulation estimates. J Geophys Res 119:5449-5464. doi:10.1002/2014JC010040

Lahoz W, Errera Q (2010) Constituent assimilation. In: Lahoz W, Khattatov B, Menard R (ed) Data Assimilation, Making Sense of Observations. Springer, New York, pp 449-490. doi:10.1007/987-3-540-74703-1

Large WG, McWilliams JC, Doney SC (1994) Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. Rev Geophys 32:363-403

Lee T, Awaji T, Balmaseda MA, Grenier E, Stammer D (2009) Ocean state estimation for climate research. Oceanography 22:160-167. doi:10.5670/oceanog.2009.74

Levitus S (1982) Climatological atlas of the world ocean. NOAA/ERL GFDL, Princeton NJ USA, 173 pp

Locarnini RA, Mishonov AV, Antonov JI, Boyer TP, Garcia HE, Baranova OK, Zweng MM, Johnson DR (2010) World Ocean Atlas 2009, Volume 1: Temperature. In: Levitus S (ed) NOAA Atlas NESDIS 68, US Government Printing Office, Washington DC USA, pp 184

Lukas R, Lindstrom E (1991) The mixed layer of the western equatorial Pacific Ocean. J Geophys Res 96:3343-3357

Maes C, Ando K, Delcroix T, Kessler WS, McPhaden MJ, Roemmich D (2006) Observed correlation of surface salinity, temperature and barrier layer at the eastern edge of the western Pacific warm pool. Geophys Res Lett 33:L06601. doi:10.1029/2005GL024772

Masuda S, Awaji T, Sugiura N, Matthews JP, Toyoda T, Kawai Y, Doi T, Kouketsu S. Igarashi H, Katsumata K, Uchida H, Kawano T, Fukasawa M (2010) Simulated rapid warming of abyssal North Pacific waters. Science 329:319-322. doi;10.1126/science.1188703

Mochizuki T, Ishii M, Kimoto M, Chikamoto Y, Watanabe M, Nozawa T, Sakamoro TT, Shiogama H, Awaji T, Sugiura N, Toyoda T, Yasunaka S, Tatebe H, Mori M (2010) Pacific decadal oscillation hindcasts relevant to near-term climate prediction. Proc Natl Acad Sci USA 107:1833-1837. doi:10.1073/pnas.0906531107

Morioka Y, Tozuka T, Masson S, Terray P, Luo JJ, Yamagata T (2012) Subtropical dipole modes simulated in a coupled general circulation model. J Clim 25:4029-4047. doi:10.1175/JCLI-D-11-00396.1

Noh Y, Kang YJ, Matsuura T, Iizuka S (2005) Effect of the Prandtl number in the parameterization of vertical mixing in an OGCM of the tropical Pacific. Geophys Res Lett 32:L23609. doi:10.1029/2005GL024540

Noh Y, Lee WS (2008) Prediction of the mixed and mixing layer depths from an OGCM. J. Oceanogr 64:217-225. doi:10.1007/s10872-008-0017-1

Pedlosky J (1996) Ocean circulation theory. Springer Berlin Heidelberg, 456 pp. doi:10.1007/987-3-662-03204-6

Palmer M, Balmaseda M, Chang YS, Chepurin G, Fujii Y, Good S, Guinehut S, Hernandez F, Martin M, Masuda S, Peterson KA, Toyoda T, Valdivieso M, Vernieres G, Wang O, Xue Y (2014) GLIVAR-GSOP/GODAE intercomparison of ocean heat content: initial results. CLIVAR EXCHANGES 64:8-10

Shay LK, Goni GJ, Black PG (2000) Effects of a warm oceanic feature on Hurricane Opal. Mon Wea Rev 128:1366–1383

Smith G, Chevallier M, Lemieux JF, Dupont F, Vernieres G, Storto A, Toyoda T, Fujii Y, Chang Y, Valdivieso M, Peterson KA, Ferry N, Hernandez F, Balmaseda MA, Keeley S, Wang X (2014) Preliminary evaluation of sea ice fields form the ocean reanalyses intercomparison project. CLIVAR EXCHANGES 64:32-34

Storto A, Dobricic S, Masina S, Di Pietro P (2011) Assimilating along-track altimetric observations through local hydrostatic adjustment in a global ocean variational assimilation system. Mon Wea Rev 139:738-754. doi:10.1175/2010MWR3350.1

Suga T, Kato A, Hanawa K (2000) North Pacific Tropical Water: its climatology and temporal changes associated with the climate regime shift in the 1970s. Prog Oceanogr 47:223-256

Sugiura N, Awaji T, Masuda S, Mochizuki T, Toyoda T, Miyama T, Igarashi H, Ishikawa Y (2008) Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. J Geophys Res 113:C10017. doi:10.1029/2008JC004741

Takeuchi E (2006) Studies on the wintertime shoaling of oceanic surface mixed layer. PhD thesis, University of Tokyo, 109 pp

Takeuchi E, Yasuda I (2003) Wintertime shoaling of oceanic surface mixed layer. Geophys Res Lett 30:2152. doi:10.1029/2003GL018511

Toyoda T, Awaji T, Ishikawa Y, Nakamura T (2004) Preconditioning of winter mixed layer in the formation of North Pacific eastern subtropical mode water. Geophys Res Lett 31:L17206. doi:10.1029/2004GL020677

Toyoda T, Fujii Y, Kuragano T, Kamachi M, Ishikawa Y, Masuda S, Awaji T, Hernandez F, Ferry N, Guinehut S, Martin M, Peterson KA, Good S, Valdivieso M, Haines K, Storto A, Köhl A, Yin Y, Shi L, Smith G, Chang Y, Vernieres G, Wang X, Wang O, Lee T, Balmaseda M (2014) Mixed layer depth intercomparison among global ocean syntheses reanalyses. CLIVAR EXCHANGES 64:22-24

Toyoda T, Fujii Y, Kuragano T, Matthews JP, Abe H, Ebuchi N, Usui N, Ogawa K, Kamachi M (2015) Improvements to a global ocean data assimilation system through the incorporation of Aquarius surface salinity data. QJR Meteorol Soc (in press). doi:10.1002/qj.2561

Toyoda T, Fujii Y, Yasuda T, Usui N, Iwao T, Kuragano T, Kamachi M (2013) Improved analysis of the seasonal-interannual fields by a global ocean data assimilation system. Theoretical and Applied Mechanics Japan 61: 31-48. doi:10.11345/nctam.61.31

Umlauf L, Burchard H (2003) A generic length-scale equation for geophysical turbulence models. J Mar Res 61(2):235-265. doi:10.1357/002224003322005087

Valdivieso M, Haines K, Balmaseda M, Barnier B, Chang Y, Ferry N, Fujii Y, Köhl A, Lee T, Martin M, Storto A, Toyoda T, Wang X, Waters J, Xue Y, Yin Y (2014) Heat fluxes from ocean and coupled reanalyses. CLIVAR EXCHANGES 64:28-31

Vernieres G, Rienecker MM, Kovach R, Keppenne C (2012) The GEOS-iODAS: Description and evaluation. NASA Tech Rep Series on Global Modeling and Data Assimilation 30:TM-2012-104606, GSFC/NASA, Greenbelt Maryland USA

Yin Y, Alves O, Oke PR (2011) An ensemble ocean data assimilation system for seasonal prediction. Mon Wea Rev 139:786-808. doi:10.1175/2010MWR3419.1

Yuan XJ, Talley LD (1996) The subarctic frontal zone in the North Pacific: Characteristics of frontal structure from climatological data synoptic surveys. J Geophys Res 101:16491-16508

Xue Y, Balmaseda MA, Boyer T, Ferry N, Good S, Ishikawa I, Kumar A, Rienecker M, Rosati AJ, Yin Y (2012) A comparative analysis of upper-ocean heat content variability from an ensemble of operational ocean reanalyses. J Clim 25:6905-6929. dio:10.1175/JCLI-D-11-00542.1

Zuo H, Balmaseda MA, Mogensen K (2014) The ECMWF-MyOcean2 eddy-permitting ocean and sea-ice re-analysis ORAP5. Part 1: Implementation. Tech Rep 736, ECMWF, Reading UK

Figure captions

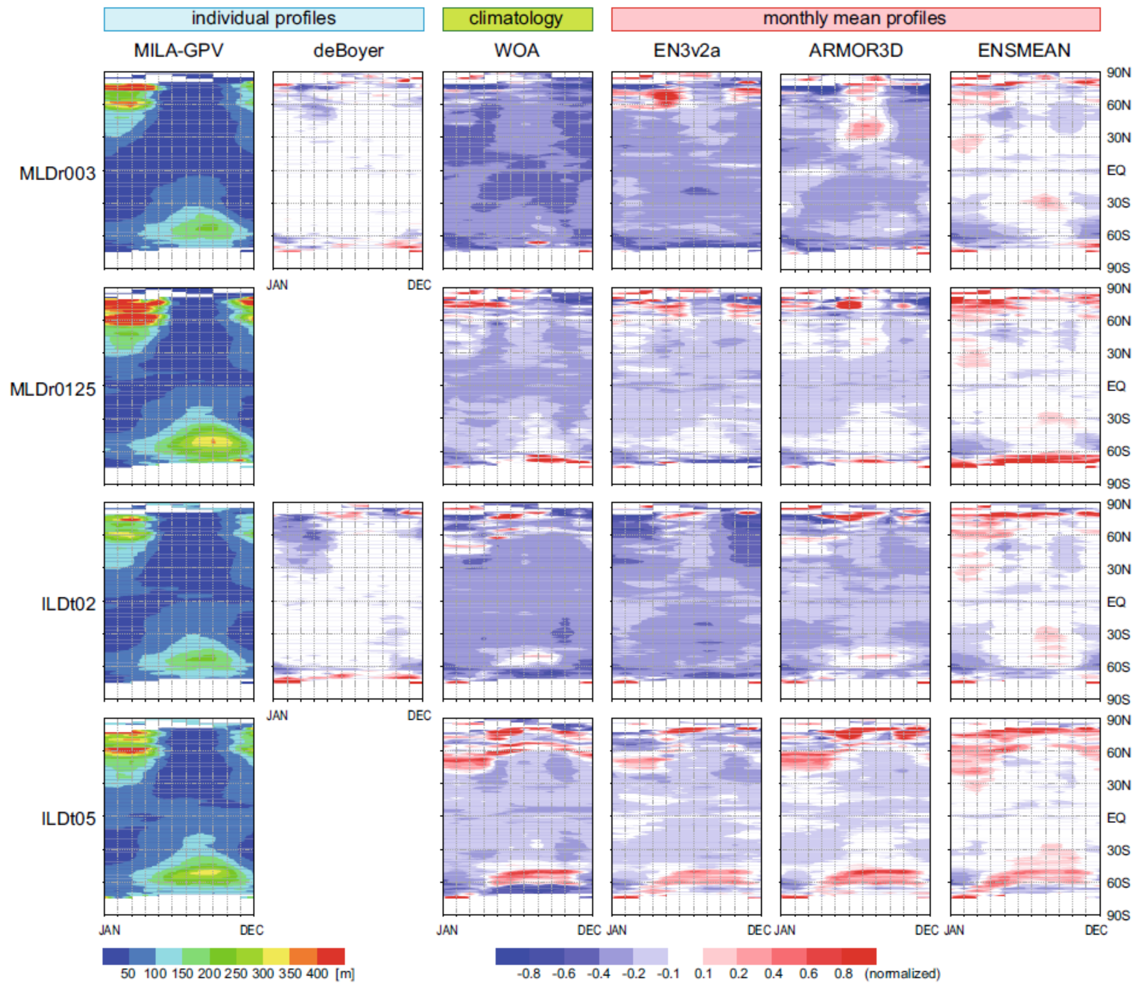


Fig. 1 Zonal mean monthly distributions of MLDr003, MLDr0125, ILDt02 and ILDt05 averaged over 2001-2011 for (left column) MILA-GPV and (others) differences from MILA-GPV, normalized by the MILA-GPV values.

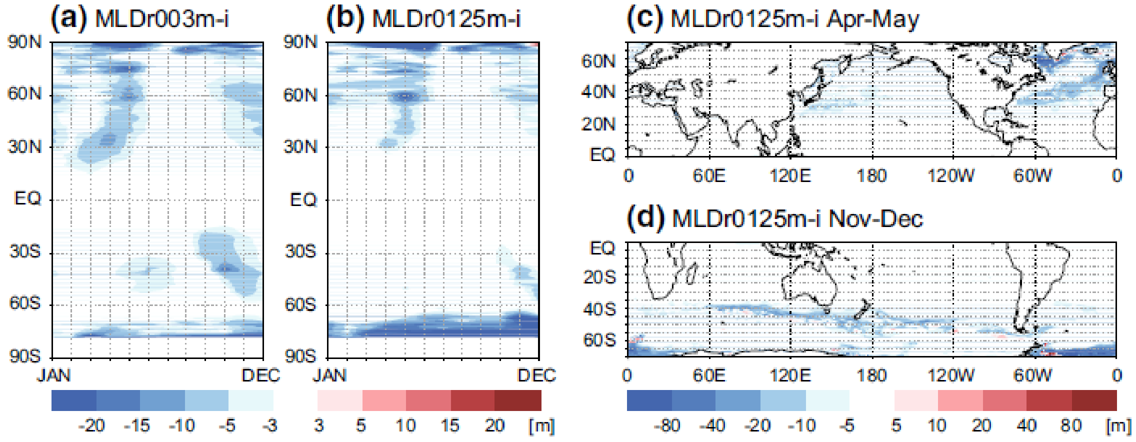


Fig. 2 (a, b) Zonal mean monthly differences MLDr003m−MLDr003i (a) and MLDr0125m−MLDr0125i (b) estimated from MOVE-G2 during the 2001-2012 period. (c, d) Distributions of MLDr0125m−MLDr0125i in April-May (c) and November-December (d).

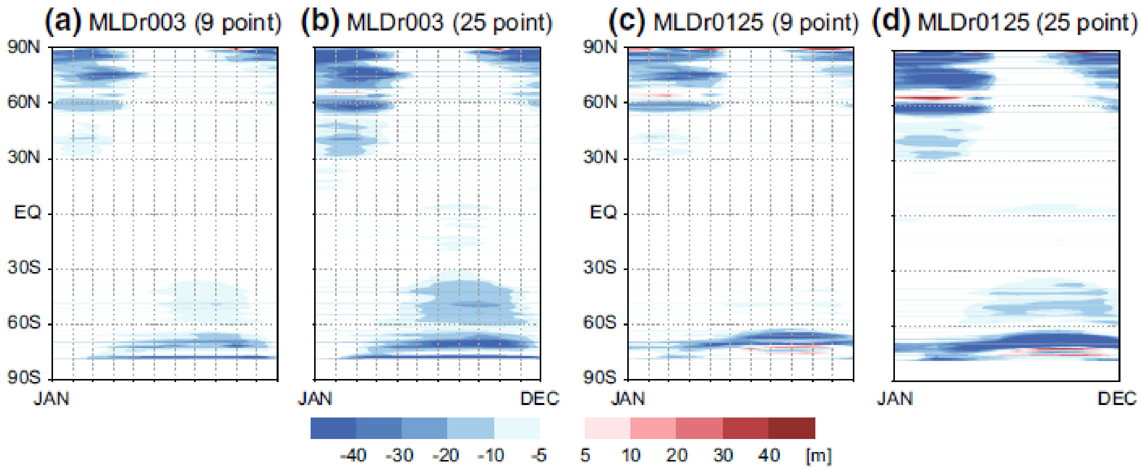


Fig. 3 Zonal mean monthly differences between MLDs from smoothed and unsmoothed TS profiles during the 2001-2012 period in the MOVE-G2 experiment. (a) MLDr003m_3x3−MLDr003m_1x1. (b) MLDr003m_5x5−MLDr003m_1x1. (c, d) Same as (a, b) but for MLDr0125m.

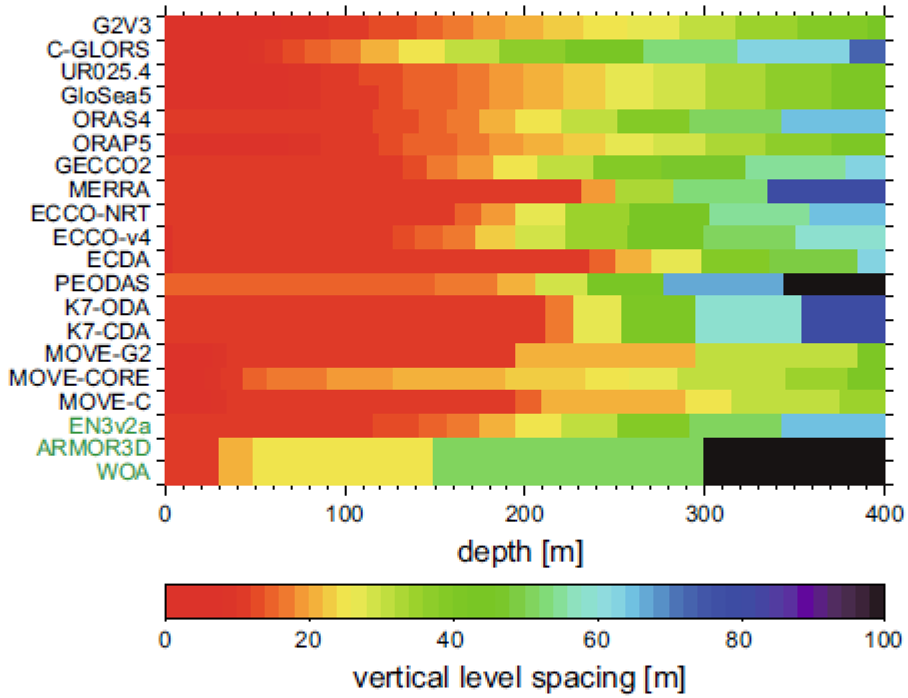
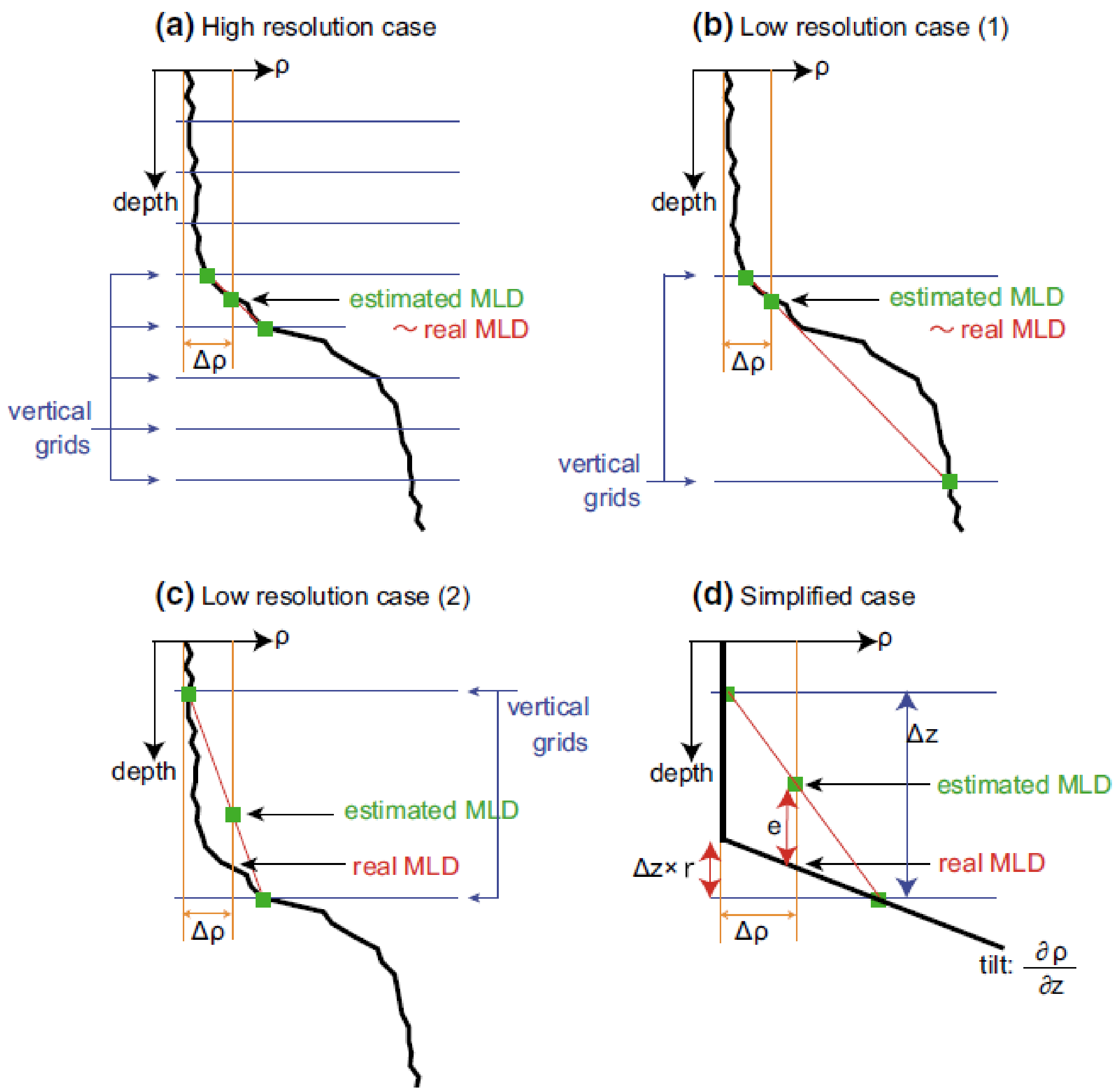


Fig. 4 Vertical resolution of the syntheses in the ORA-IP.



885

886 **Fig. 5** Schematic illustrations of the MLD estimation for (a) high resolution and (b, c) low
887 resolution cases. MLDs are estimated from a common vertical profile of potential density (black
888 line) by using linearly interpolated values between the vertical grids. (d) A simplified sketch
889 showing relationship between error in the MLD estimation and vertical level spacing.

890

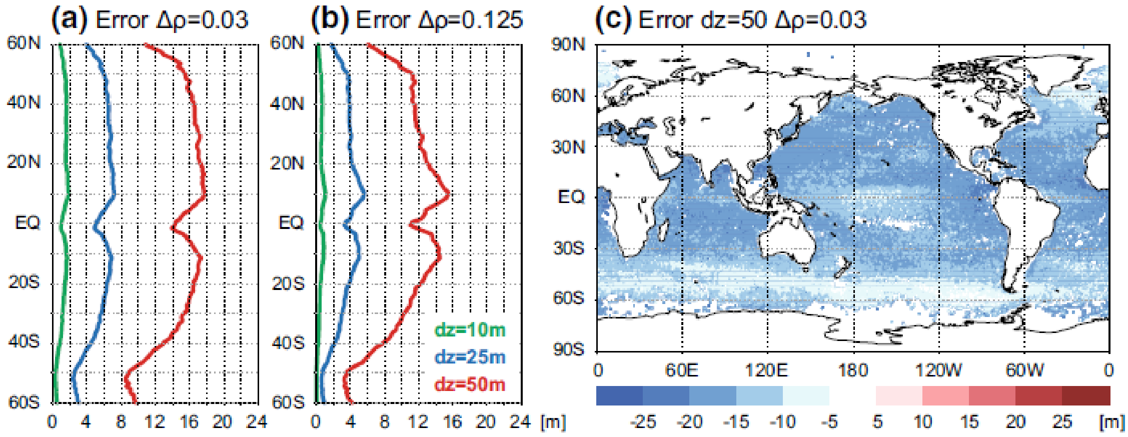


Fig. 6 (a, b) Zonal mean expected values for the underestimation associated with limited vertical resolution of profiles in the MLD estimation by using the criteria, $\Delta\rho = 0.03$ (a) and $\Delta\rho = 0.125$ kg m^{-3} (b). Green, blue and red lines denote the vertical resolutions of Δz =10, 25 and 50 m, respectively. (c) Distribution of the errors when $\Delta\rho = 0.03$ kg m^{-3} and Δz =50 m. $\frac{\partial\rho}{\partial z}$ values at the ML bottom are calculated from the individual Argo profiles over the 2000-2012 period and averaged onto monthly and 1 degree by 1 degree bins for use in Eq. (1). Units are in meter.

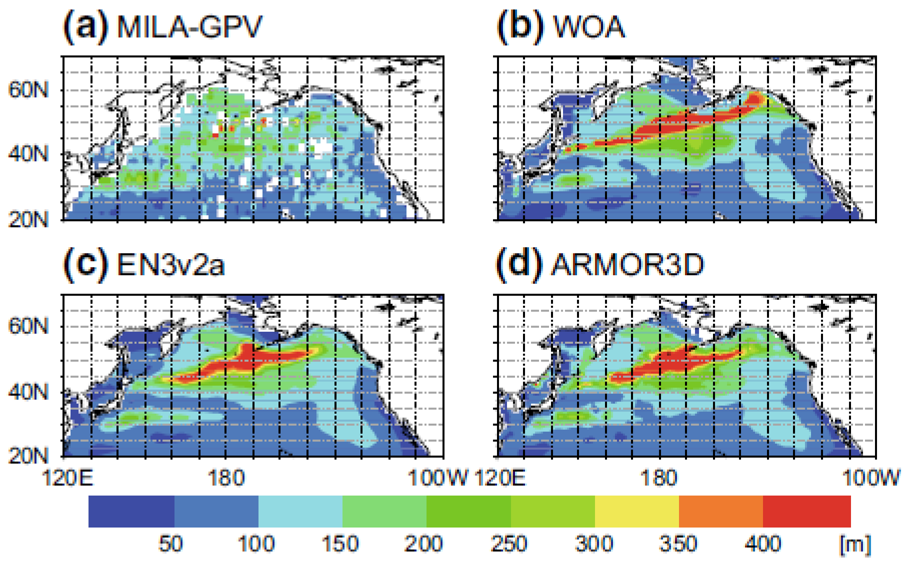


Fig. 7 ILDt05 distributions in March in the North Pacific basin averaged over 2001-2011 for (a)

MILA-GPV, (b) WOA, (c) EN3v2a and (d) ARMOR3D.

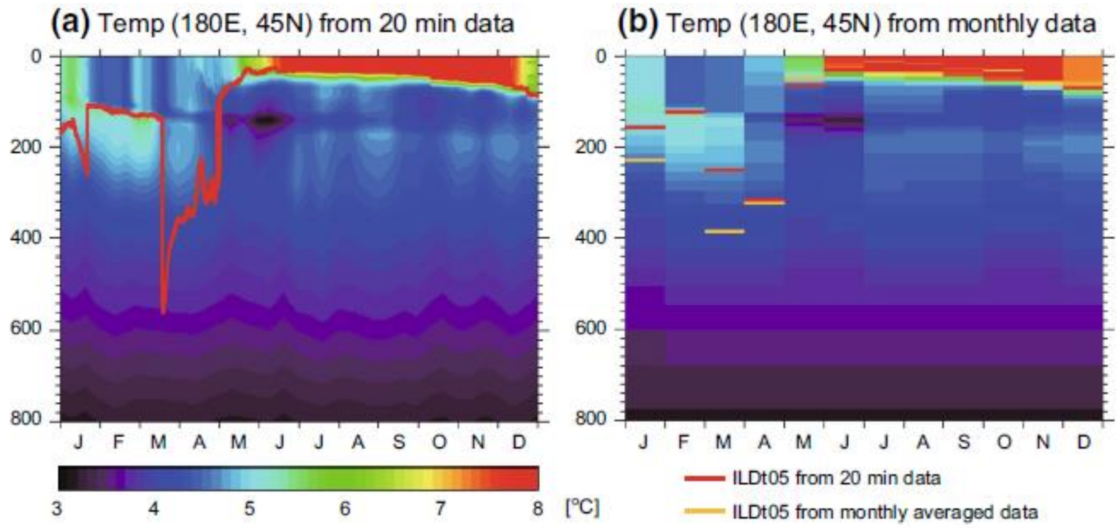


Fig. 8 Time evolution of the vertical temperature profile in the subpolar North Pacific at 180°E, 45°N in 2001 represented (a) by snapshot data with an interval of the model time step (20 minute) and (b) by monthly mean data. Both time series are derived from the MOVE-G2 product. Red (yellow) line indicates ILDt05s calculated from the former (latter) data.

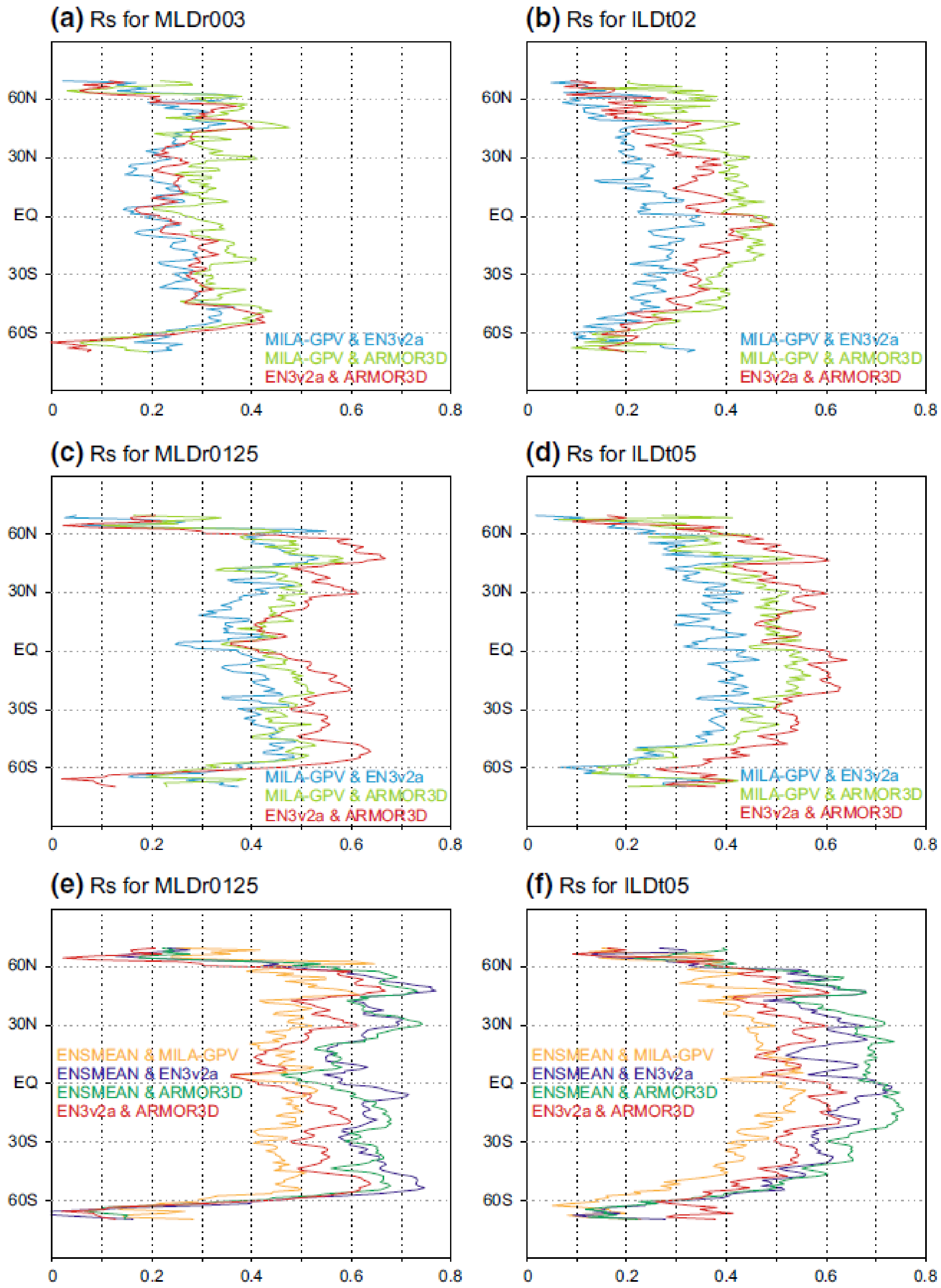


Fig. 9 Zonal mean correlation coefficients of monthly interannual anomalies during 2005-2010 (defined in this study as monthly data over 6 years with the mean seasonal cycle for this period

removed), (a-d) among MILA-GPV, EN3v2a and ARMOR3D and (e, f) between ENSMEAN and MILA-GPV/EN3v2a/ARMOR3D, for (a) MLDr003, (b) ILDt02, (c, e) MLDr0125 and (d, f) ILDt05. The period was chosen because all datasets are available during this period.

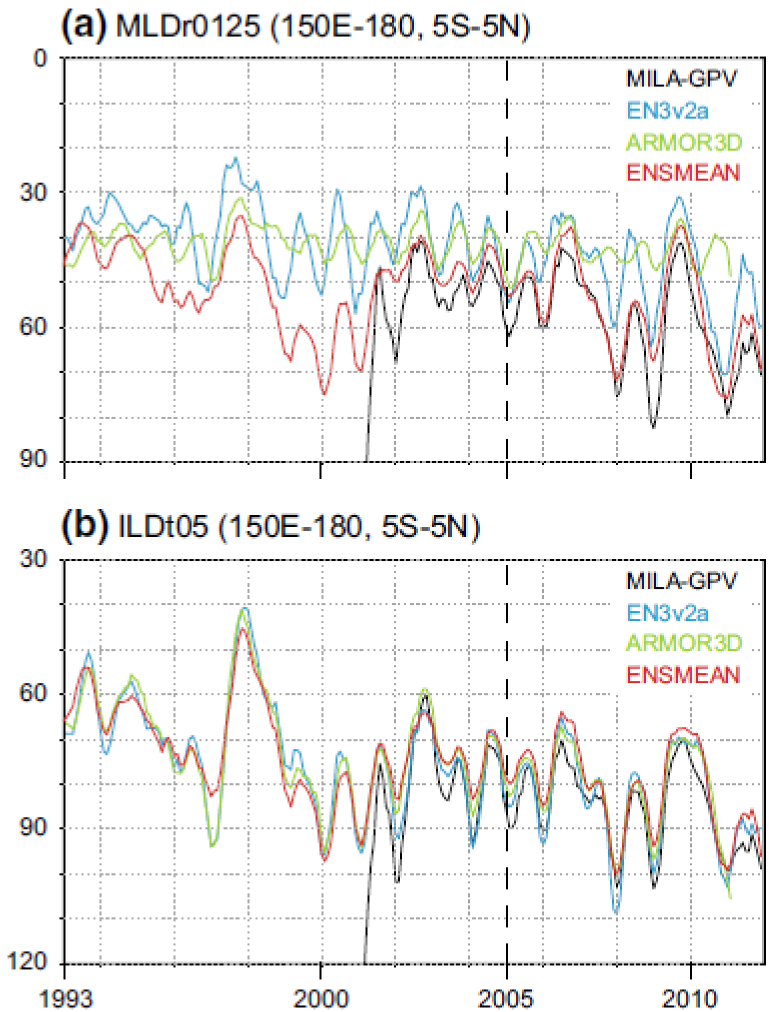


Fig. 10 Time series of MLDr0125s (a) and ILDt05s (b) averaged over the western Pacific warm pool region (150ºE-180º, 5ºS-5ºN) for MILA-GPV, EN3v2a, ARMOR3D and ENSMEAN (black, blue, green and red lines, respectively). 5-month running mean values are plotted.

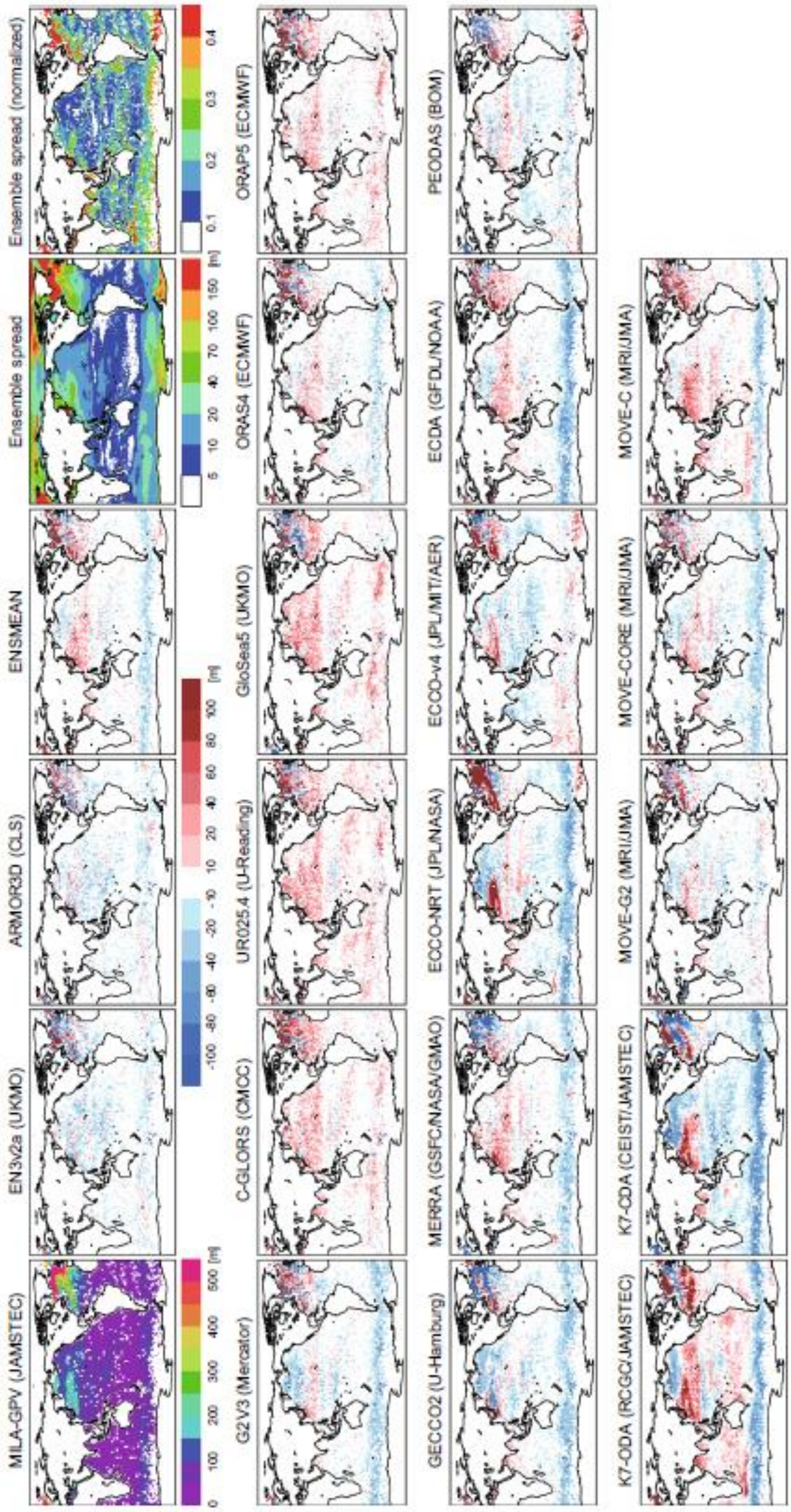


Fig. 11 Distributions of MLDr0125 in February averaged over 2001-2011 for (top row from

left) MILA-GPV, differences between MILA-GPV and either EN3v2a, ARMOR3D or ENSMEAN (e.g., EN3v2a−MILA-GPV), absolute and normalized (by the MILA-GPV values) ensemble spread of the reanalyses and (others) differences between MILA-GPV and the individual reanalyses.

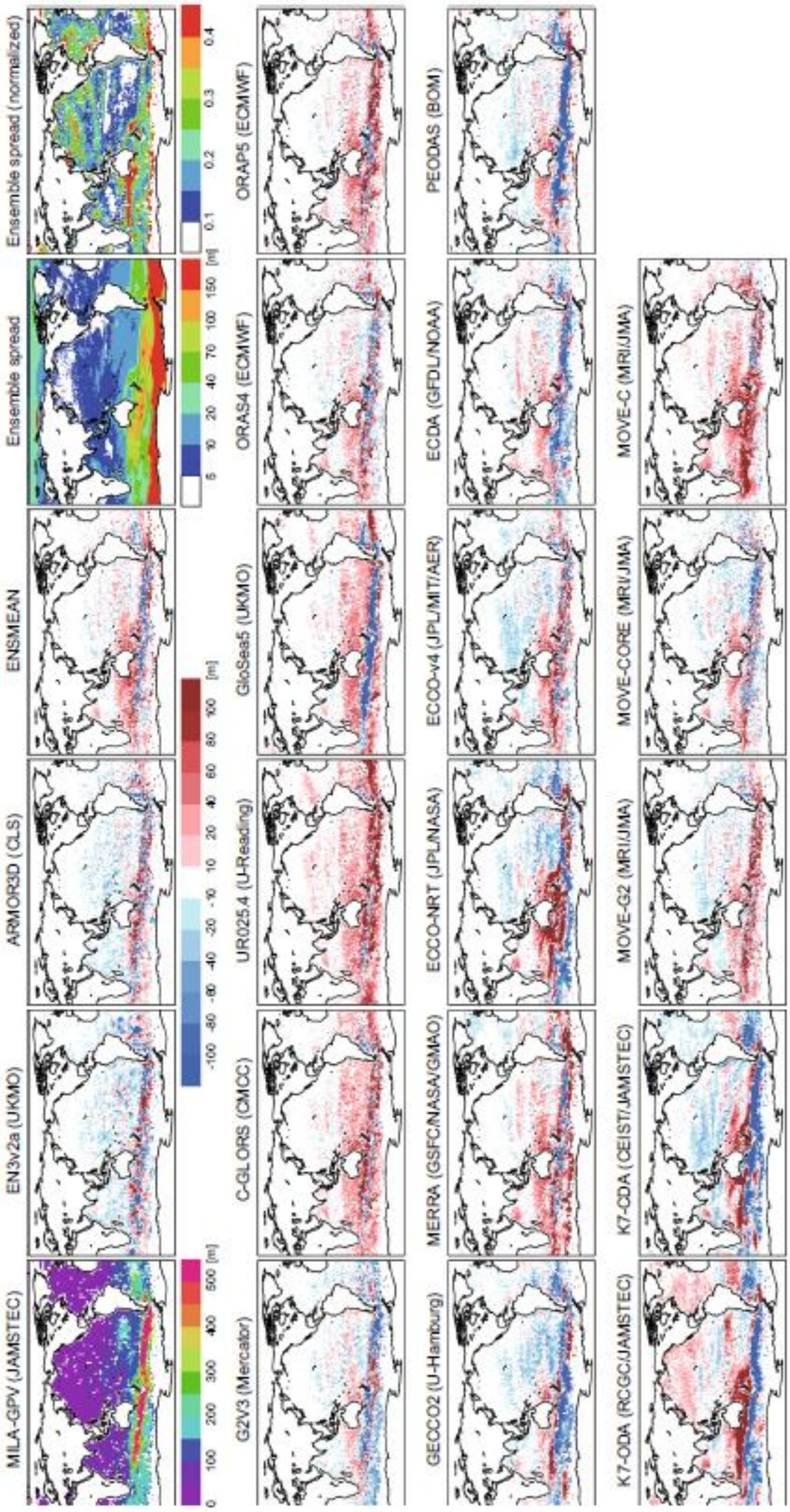


Fig. 12 Same as Fig. 11 but for August.

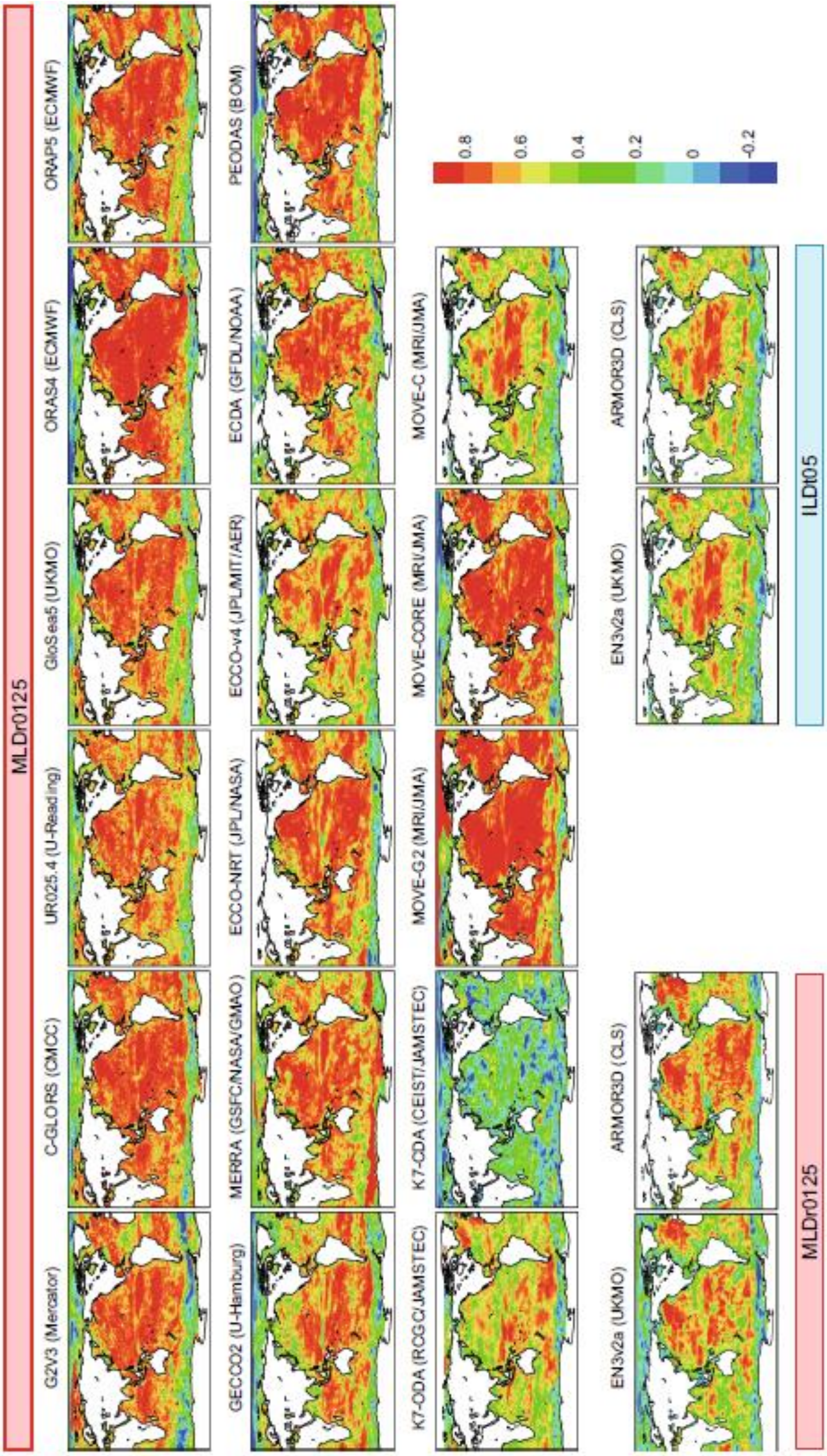

MLDr0125
G2V3 (Mercator)
C-GLORS (CMCC)
UR025.4 (U-Reading)
GloSea5 (UKMO)
ORAS4 (ECMWF)
ORAP5 (ECMWF)
GECCO2 (U-Hamburg)
MERRA (GSFC/NASA/GMAO)
ECCO-NRT (JPL/NASA)
ECCO-v4 (JPL/MIT/AER)
ECDA (GFDL/NOAA)
PEODAS (BOM)
K7-ODA (RCGC/JAMSTEC)
K7-CDA (CEIST/JAMSTEC)
MOVE-G2 (MRI/JMA)
MOVE-CORE (MRI/JMA)
MOVE-C (MRI/JMA)
0.8
0.6
0.4
0.2
0
-0.2
EN3v2a (UKMO)
ARMOR3D (CLS)
EN3v2a (UKMO)
ARMOR3D (CLS)
MLDr0125
ILDt05

Fig. 13 Distributions of correlation coefficients for the interannual anomalies of MLDr0125 over the 2001-2011 period between ENSMEAN and the individual reanalyses (1st-3rd rows) and between ENSMEAN and EN3v2a/ARMOR3D (for both MLDr0125 and ILDt05; bottom row).

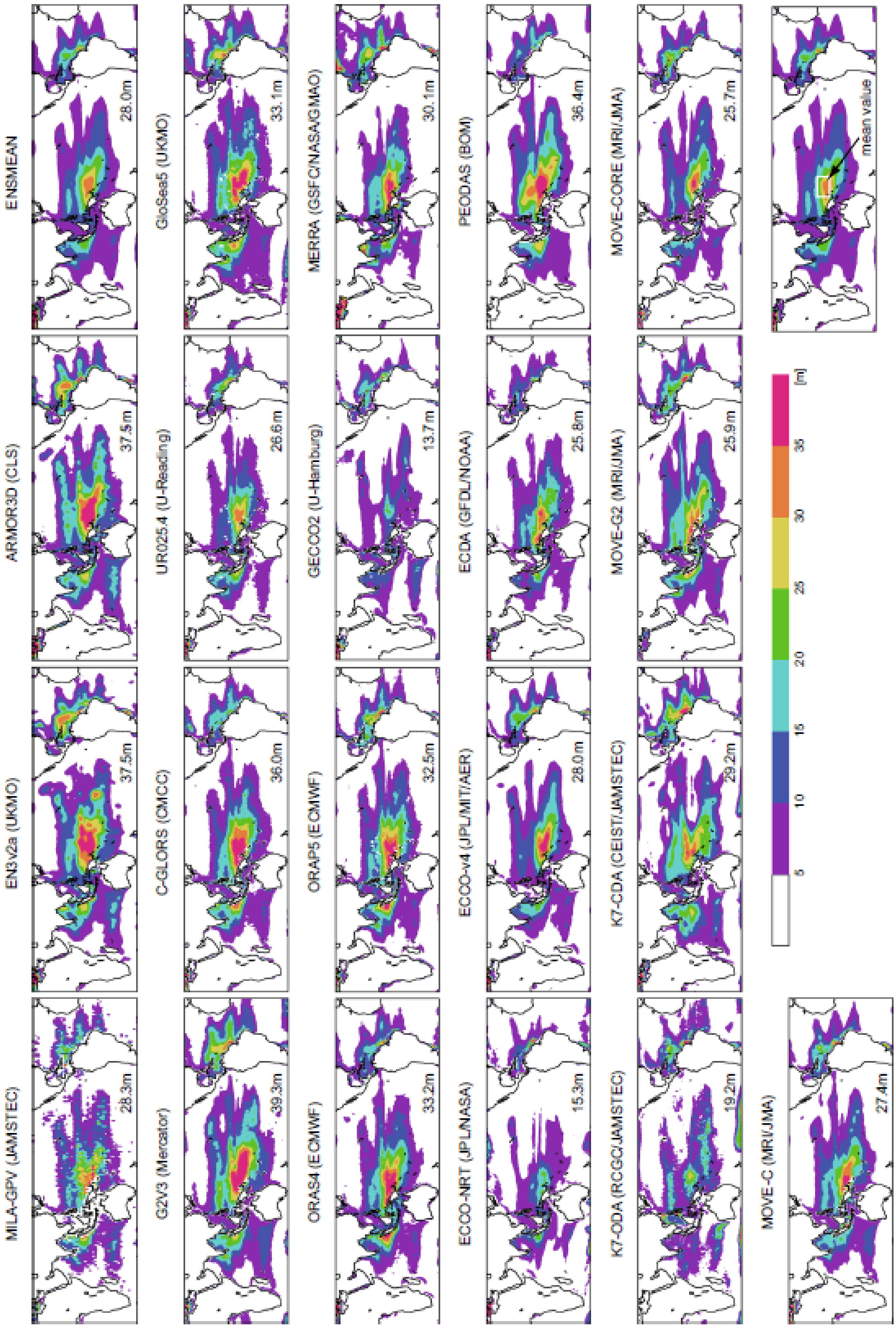


Fig. 14 Distributions of BLT averaged over 2001-2011. BLT value averaged over the western

equatorial Pacific (150ºE-170ºE, 5ºS-5ºN; white box in the right bottom figure) is indicated in each figure.

Table 1 Description of the observation-only analyses in the ORA-IP

Name	Center	Resolution	Assimilationed data	Reference
EN3v2a	UKMO	1°, 30lv	TS	Ingleby and Huddleston (2007)
ARMOR3D	CLS	1/3°, 33lv	TS/SLA/SST	Guinehut et al. (2012)

Table 2 Description of the reanalysis products in the ORA-IP

Name (Center)	OGCM Resolution	Forcing	ML model	Assimilation data	Assimilation Method	Reference
G2V3 (Mercator Océan)	NEMO3.1 1/4°, 75lv	ERAi corrected	Blanke and Delecluse (1993)	TS/SLA/SST/SIC	KF, 3DVAR, Bias,FGAT,IAU	Ferry et al. (2012)
C-GLORS (CMCC)	NEMO3.2 1/2°, 50lv	ERAi corrected	Blanke and Delecluse (1993)	TS/SLA/SST/SIC	3DVAR, FGAT	Storto et al. (2011)
UR025.4 (U-Reading)	NEMO3.2 1/4°, 75lv	ERAi	Blanke and Delecluse (1993)	TS/SLA/SST/SIC	OI, FGAT, IAU	Haines et al. (2012)
GloSea5 (UKMO)	NEMO3.2 1/4°, 75lv	ERAi	Blanke and Delecluse (1993)	TS/SLA/SST/SIC	3DVAR, FGAT, IAU	Blockley et al. (2013)
ORAS4 (ECMWF)	NEMO3.0 1°x(0.3-1)°, 42lv	ERA40, ERAi	Blanke and Delecluse (1993)	TS/SLA/SST	3DVAR, FGAT, Bias, IAU	Balmaseda et al. (2013)
ORAP5 (ECMWF)	NEMO3.4 1/4°, 75lv	ERAi	Blanke and Delecluse (1993)	TS/SLA/SST/SIC	3DVAR, FGAT, Bias, IAU	Zuo et al. (2014)
GECCO2 (U-Hamburg)	MITgcm 1°x(1/3-1)°, 50lv	NCEP-R1 corrected	Large et al. (1994)	TS/SLA/MDT/ SST	4DVAR	Köhl (2014)

MERRA (Ocean) (GSFC/NASA/GMAO)	MOM4 1/2°x(1/4-1/2)°, 40lv	Merra	Large et al. (1994)	TS/SLA/SST/SIC	EnOI	Vernieres et al. (2012)
ECCO-NRT (JPL/NASA)	MITgcm 1°x(0.3-1)°, 46lv	NCEP-R1 corrected	Large et al. (1994)	T/SLA	KF, KS	Fukumori (2002)
ECCO-v4 (JPL/MIT/AER)	MITgcm 0.4°x(0.4-1)°, 50lv	ERAi corrected	Large et al. (1994)	TS/SLA/SST/SIC	4DVAR	Wunsch and Heimbach (2013)
ECDA (GFDL/NOAA)	MOM4 coupled 1°x(0.3-1)°, 50lv	Coupled	Large et al. (1994)	TS/SST	EnKF	Chang et al. (2013)
PEODAS (BOM)	MOM2 2°x(0.5-1.5)°, 25lv	ERA40, NCEP-R2	Chen et al. (1994)	TS/SST	EnKF	Yin et al. (2011)
K7-ODA (ESTOC) (RCGC/JAMSTEC)	MOM3 1°, 45lv	NCEP-R1 corrected	Large et al. (1994)	TS/SLA/SST	4DVAR	Masuda et al. (2010)
K7-CDA (CEIST/JAMSTEC)	MOM3 coupled 1°, 45lv	Coupled	Noh et al. (2005)	TS/SLA/SST	4DVAR	Sugiura et al. (2008)
MOVE-G2 (MRI/JMA)	MRI.COM3 1°x(0.3-0.5)°, 53lv	JRA55 corrected	Noh et al. (2005)	TS/SLA/SST	3DVAR, FGAT, IAU	Toyoda et al. (2013)
MOVE-CORE (MRI/JMA)	MRI.COM3 1°x0.5°, 51lv	CORE2	Umlauf and Burchard (2003)	TS	3DVAR, IAU, Bias	Danabasoglu et al. (2013)
MOVE-C (MRI/JMA)	MRI.COM2 1°x(0.3-1)°, 50lv	Coupled	Noh et al. (2005)	TS/SLA/SST	3DVAR, IAU, Bias	Fujii et al. (2009)

Abbreviations: MDT (mean dynamic topography), KF (Kalman filter), 3DVAR (3 dimensional variational method), Bias (one-step bias-correction algorism), FGAT (first guess at appropriate time), 4DVAR (4 dimensional variational method), EnOI (ensemble optimal interpolation), KS (Kalman smoother), EnKF (ensemble KF), IAU (incremental analysis updates)

Table 3 Variables and duration available for the ORA-IP and contact parson

Synthesis	Variables	Duration	Contact parson
EN3v2a	All	1993-2011	S. Good
ARMOR3D	All	1993-2010	S. Guinehut
G2V3	All	1993-2011	F. Hernandez
C-GLORS	All	1991-2011	A. Storto
UR025.4	MLDr003, MLDr0125, ILDt05	1993-2010	M. Valdivieso
GloSea5	All	1993-Jul. 2012	M. Martin
ORAS4	All	1958-2011	M. Balmaseda
ORAP5	All	1993-2012	H. Zuo
GECCO2	All	1948-Nov. 2011	A. Köhl
MERRA	MLDr0125, ILDt05	1993-2011	G. Vernieres
ECCO-NRT	MLDr0125, ILDt05	1993-2011	O. Wang
ECCO-v4	MLDr0125, ILDt05	1992-2010	X. Wang
ECDA	All	2005-2011	Y.-S. Chang
PEODAS	All	1980-2012	O. Alves
K7-ODA	All	1975-2011	S. Masuda
K7-CDA	All	2000-2006	Y. Ishikawa
MOVE-G2	All	1993-2012	T. Toyoda
MOVE-CORE	All	1948-2007	Y. Fujii
MOVE-C	All	1950-2011	Y. Fujii

Durations submitted to the ORA-IP are sometimes shorter than those of the original reanalyses