Research Papers
Issue RP0279
December 2015

ODA - Ocean modeling and Data Assimilation Division

A Nonlinear Delayed-mode Quality Control Procedure for Climatological Studies using Argo Data in the Pacific Western Boundary Current Region

By **Wenjing Jia,**
Fondazione CMCC

Nadia Pinardi
Department of Physics and Astronomy, Alma Mater Studiorum University of Bologna
n.pinardi@sincem.unibo.it

Dong Wang
Fondazione CMCC

Simona Simoncelli
INGV,
simona.simoncelli@ingv.it

Andrea Storto,
Fondazione CMCC,
andrea.storto@cmcc.it

Simona Masina,
Fondazione CMCC and INGV,
simona.masina@cmcc.it

W.Jia and D.Wang received funding by the Italian Ministry of Education, University and Research and the Italian Ministry of Environment, Land and Sea under the GEMINA project for the OOCC China-Italy collaboration. N.Pinardi and S. Simoncelli were supported by the SeaD ataNet project at INGV. We thank Dr. P. M. Poulain for helpful suggestions.

SUMMARY A nonlinear Delayed-Mode Quality Control (DMQC) methodology is developed to estimate monthly mean climatologies from the large Argo dataset (2005-2012) over the North Pacific Boundary Current region. In addition to the commonly-used linear quality control procedure, which checks for instrumental, transmission and gross value errors, we develop and show the impact of nonlinear, iterative statistical checks on the quality of the resulting climatology. Objective analysis is applied to produce progressively statistical climatological checks to flag out-of-standard-deviation ARGO profiles. The optimal method uses horizontal regional climatological averages defined in five regime-oriented subregions in the Kuroshio areas and the Japan Sea. This new method is shown to produce lower standard deviations in the deep water layers and it is thus capable to reject observations with large representativeness/sampling errors. The resulting monthly mean climatology for the period 2005 to 2012 is shown to be consistent with previous estimates for the region, including the WOA13 observed climatology and re-analysis fields.

Keywords: Kuroshio dynamics, ARGO climatology, non-linear quality control algorithms

1. INTRODUCTION

Featured by mesoscale activities and strong carbon uptake (e.g., Takahashi et al. 2002; Yu and Weller 2007), the western boundary current (WBC) regions play a key role in the ocean heat transport and overturning circulation. Being characterized by a frontal structure and by mesoscale and ring dynamics, the WBC and associated recirculation regions are challenging in terms of observational and modeling requirements for climatological studies. Even with coordinated in-situ and satellite observations it is difficult to depict the three-dimensional WBC synoptic structure (Argo Science Team 2012) because of sampling scheme shortcomings. Furthermore, WBC climatological estimates may be affected by relatively poor representativeness of the observations due to the limited space/time sampling of the data. Representativeness errors normally are defined in data assimilation schemes as the observational error component due to unresolved scales (Janic and Cohn, 2006). Climatological estimates need to filter out high frequency signals from observations, i.e., they cannot consider high space-time variability sampled by the observing system. Thus any in situ observations in a high mesoscale variability region and in particular the WBC areas, is bound to be affected by high representativeness errors if climatological estimates have to be obtained. The distinct dynamics and observational issues warrant specifically designed treatment when computing climatologies for areas characterized by large variability, particularly the WBC regions.

Argo profiling floats have been widely deployed since the 2000s and provide for the first time the capability of monitoring the upper ~2000 m temperature and salinity properties at almost continuous spatio-temporal scales in a cost-efficient and versatile way. Traditional monitoring by XBT and MBT lacks the correct representation of T/S correlations and CTD surveys have large sampling errors. Argo data, with a more uniform sampling and T/S profiles, could serve to better estimate climatologies for different parts of the world ocean. Thus our work concentrates on the estimation of climatologies only from the emerging and ever growing Argo profiles.

The International Argo Program recommends quality control (QC) for Argo data being done in two different steps: (1) real-time QC by automatic screening of errors and spikes etc. and data transmission within 24 hours to Global Data Assembly Centers (GDACs); (2) Delayed-Mode QC (DMQC) with more sophisticated procedures and data transmitted to GDACs every 1-2 years (Barker et al. 2011).

Argo observations have experienced technical problems from the beginning: a pressure sensor drift was detected in 2009 (Argo Steering Team 2011, Barker et al., 2011), a salinity offset due to biofouling was found to affect floats with long lifetimes in different world ocean regions (Oka 2005, Wong et al., 2003, Böhme and Send, 2005, Owens and Wong, 2009), and data transmission errors were documented (Boyer et al. 2013). DMQC procedures are then required to eliminate the floats affected by these instrumental errors.

In addition, datasets undergone QC procedures for instrumental errors may still contain observations that are not representative of the climatological status of the region under study due to sampling issues. For example, an Argo profile could be trapped inside a deep ring for several months and thus give a repetitive measurement of the ring hydrography which is not representative of long term climatology of the region. Thus, efforts are needed to eliminate such representativeness errors, i.e., the errors due to sampling issues of the part of the flow field with high variability. One way is to develop QC procedures that perform a statistical check and flag those profiles with values outside an empirically determined threshold (normally 2 or 3 times the standard deviations) from the climatological mean. After removal of the flagged profiles, the new climatological mean, calculated using the unflagged profiles, may be different from the previous one. Thus, climatology may be changed after each time such statistical check is performed. The new method that we develop and describe will be called nonlinear DMQC hereafter. Our method, in some regards, is similar to the variational quality control, commonly used in meteorological data assimilation (Lorenc, 1988; Anderson and Jarvinen, 1999), which assigns weights to observations based on the departure of the observations from the analysis at a certain stage of the variational minimization, implying nonlinearity. However, variational quality control was conceived for rejecting observations suspicious of gross errors, while nonlinear DMQC aims at flagging out observations non-representative of the climatology. Thus, nonlinear DMQC applies the conservative methodology that observations flagged at some iteration cannot be de-flagged in a later step of the DMQC.

In addition to the statistical check, the new DMQC methodology developed in this paper also uses climatologies defined in subregions of dynamical relevance. An Objective Analysis (OA) technique is used to estimate the climatology for the DMQC check and calculate the standard deviations, iterating this procedure, and eliminating successively the flagged profiles from the previous DMQC step. Climatology is then re-computed at each step after the Argo profiles with values outside two standard deviations from the previous climatological values are eliminated. This nonlinear DMQC check is capable of detecting outlier profiles due to both instrumental and representativeness errors, progressively building up a robust climatological dataset. The climatologies obtained with this method have the potential to advance understandings of key processes of the WBC regions and numerical modeling studies.

In this paper we apply the nonlinear DMQC to the North Pacific WBC region to compute monthly mean climatologies from Argo data. It is not surprising that very limited studies exist on Argo data DMQC for the WBC regions. Most of the previous DMQC studies have been focusing on the large-scale interior open oceans, e.g., Atlantic (Gaillard et al. 2009), Pacific (Zhang et al. 2013), equatorial (Wong et al. 2003), or polar regions (Böhme and Send 2005), or the global ocean (e.g., Owens and Wong 2009; Johnson et al. 2013).

The data and methods for DMQC of the Kuroshio Argo data are presented in section 2, followed by descriptions of the obtained hydrographic profiles and final gridded products in section 3. In section 4 the mapped profiles are compared with other observational climatologies and re-analyses and Section 5 concludes.

2. DATA AND METHODS

a. OBSERVATIONAL DATA AND STUDY REGION

The Argo data used in this study are from the real-time profiling floats during the years 2005-2012 in the study domain (115-145°E, 21-42°N, see Fig. 1) and are accessed from the portal of the Coriolis Data Centre (http://www.coriolis.eu.org). The Argo floats in the study domain were deployed primarily by the Chinese, Japanese, and Korean national programs. The Argo floats are of various types (e.g., APEX, PROVOR, ARVOR, etc.) with different data communication technologies (e.g., Argos, Iridium, etc.; Argo Steering Team 2013). Fig. 1 shows the monthly position of floating profiles in August that is representative of the monthly sampling.

The study region consists of the Kuroshio and Kuroshio Extension (KE) regions (subregions I, III and IV in Fig. 1), a North Pacific Subtropical Countercurrent (STCC) region (subregion II in Fig. 1) and the southern Japan/East Sea (subregion V in Fig. 1). The five subregions have distinct circulation dynamics: subregions I and IV are dominated by the Kuroshio current, subregion II encompasses part of the North Pacific Subtropical Front (NSTF) and subregion III is characterized by the North Pacific

Subtropical Mode Water (STMW) (Kobashi et al. 2006). The northeastern portion of the study domain (northward of 35°N and east of Japan in the open North Pacific, Fig.1) is also considered separately but it is not given a number since it is of minor interest for the climatological characterization of the WBC and its subtropical recirculation area.

The nonlinear DMQC processed Argo profiles from this paper will be compared with other DMQC processed datasets, namely WOD13 (World Ocean Database 2013; Johnson et al., 2013) and GDAC. Subsequently, the gridded climatology derived from the nonlinear DMQC dataset of this paper will be compared with two other gridded datasets, namely WOA13 and RA. WOA13 (World Ocean Atlas 2013) is an observational climatology for the period 2005-2012 (1/4° grid), derived from WOD13 containing CTD casts, XBT, MBT and Argo float measurements (Locarnini et al., 2013, Zweng et al., 2013). RA is the climatology obtained from the average of the 2005-2012 reanalysis produced by Storto and Masina (2013). The RA is on a 1/4° grid and it assimilated both in-situ (including Argo) and satellites observations.

b. NONLINEAR DELAYED-MODE QUALITY CONTROL METHOD

The nonlinear DMQC developed in this paper consists mainly of two steps: the first is the traditional single profile DMQC while the second is statistical and iterative. During the first step, named linear DMQC, each Argo profile is checked for date, time and position duplication, non-monotonic pressure growth with depth, negative values, values of temperature and salinity within a ‘gross range’ ([-5°C, 45°C] for temperature; [25 psu, 45 psu] for salinity) and spikes, as recommended by the standard Argo QC (Argo Data Management Team 2012).

In the second step a climatology check is defined for vertically bin-averaged profiles at standard levels as listed in Table 1. This check is done by flagging profiles that are outside 2 standard deviations from a climatological mean that is computed from the previous step quality checked data. As discussed above, however, the traditional linear QC is ineffective in eliminating non-representative profiles.

The nonlinear DMQC method developed in this paper uses the unflagged data (or 'good data') after the linear step to generate an OA regularly gridded climatology. Vertical climatological means and standard deviations at selected depths (Table 1) are then computed averaging the gridded climatology in each individual subregion in Fig.1. These vertical climatological means and standard deviations are used in the statistical check on the unflagged profiles after the linear QC, flagging some more. Out of this process the remaining unflagged profiles are used again to re-calculate a climatology and evaluate the mean and standard deviations, iteratively. The method progressively flags profiles from new climatological means evaluated without the previously flagged profiles after the statistical check. The sensitivity of the nonlinear DMQC method to different horizontal averaging subregions and number of iterations will be described in section 3.

The OA technique used for the computation of the gridded climatology is from Carter and Robinson (1987). In the present implementation, we use a Gaussian correlation function (Table 2), assuming the target fields are stationary. The optimal OA parameters (i.e., Case 4 in Table 2) are determined from a series of OA sensitivity experiments, by quantifying the OA mapping error (i.e., root mean square differences between mapped variable and sampled variable). In Table 2, the smaller values of correlation function parameters a=200 km (i.e., de-correlation scale) and b=120 km (i.e., e-folding or decay scale) and 1/4°x1/4° grid size are shown to be the optimal choice given the data coverage. To further decrease the uncertainty from OA estimates and increase its statistical reliability, the areas with mapping errors larger than 0.3 are masked out of each OA field and they are not used in the evaluation of the climatological horizontal averages. In previous studies the OA technique has been applied to Argo data at a coarser resolution, e.g., 3° in the Pacific as a whole (Zhang et al. 2013) and 1/2° in the Japan Sea alone (Park and Kim 2013).

3. RESULTS

a. QC-PROCESSED PROFILES

Fig. 2 summarizes the number of Argo profiles before and after the DMQC method in each year (Fig. 2a) and each month of the period 2005-2012 (Fig. 2b). In the entire study region ~89% of raw profiles passed the linear DMQC, whereas in the subregions I through IV the rate grew up to ~95%. In subregion V (i.e., part of Japan Sea) only ~72% of the original profiles passed linear DMQC, likely due to the pressure sensor errors. Fig. 2a shows an overall increase in the quantity of good profiles during 2005-2012 in the whole study region, except a sudden drop in 2010 due to the problematic APEX sensor halt and recall in 2009 (Argo Steering Team 2011; Barker et al. 2011). Since 2010 fewer profiles have been filtered out by the linear DMQC method (Fig. 2a), that is consistent with technical advancements: e.g., the faulty sensors replacement, and the yearly expansion of Iridium technology employed in the floats to reduce data transmission errors (Argo Steering Team 2011). Fig. 2b shows that sufficient Argo data are available after the linear DMQC for monthly climatology analysis.

The monthly climatologies of area-averaged hydrographic profiles over the Kuroshio region before and after the linear and nonlinear DMQC are presented in Fig. 3. The linear DMQC eliminates the errors due to the vertical shifts of the pressure profiles by the aforementioned faulty pressure sensors deployed before 2010 (Barker et al. 2011). The nonlinear DMQC removes the spurious variations of temperature and salinity below 1000 m depth, where such an apparent seasonal cycle should not exist.

The nonlinear DMQC eliminates the remaining variability below 1000 m due to the representativeness observational error discussed in the introduction. The standard deviations of the monthly mean profiles for temperature and salinity are displayed in Fig. 4, for raw data, after linear DMQC, and after nonlinear DMQC respectively. The large standard deviations in the deep layers observed in the raw and linear DMQC profiles for both temperature and salinity are greatly reduced by the nonlinear DMQC.

b. SENSITIVITY EXPERIMENTS TO NONLINEAR DMQC METHOD FREE CHOICES

In this section we discuss the sensitivity of the nonlinear DMQC method to the 'free choices' of the method: the first is the choice of the climatology averaging area where the standard deviations are calculated and the second is the number of iterations of the method. We use the OA gridded fields to show the impact of these choices on the final result, thus defining an optimal nonlinear DMQC method.

Fig. 5 is an example temperature distribution at sea surface, defined as the upper 5 m bin-averaged depth measurements (Table 1) in February. Four climatological check value methods are compared: 1.) the linear DMQC only (without climatology check); 2.) the nonlinear method with a climatology check value and its relative standard deviation taken as the average over 1°x1° bins; 3.) a climatology check value and its relative standard deviation taken as the average over the whole region; and 4.) a nonlinear DMQC using a climatology check value and its relative standard deviation averaged over the five regime-oriented subregions. By comparing with the sea surface temperature field in February after the linear DMQC (Fig. 5a), application of a climatology check with regular 1°x1° grid bins (Fig. 5b) leads to a noisier field with several "bullseyes", leaving too many profiles that are not representative of a large scale climatology, whereas a climatology check over the entire region does not effectively change the field (Fig. 5c). The method using climatology individually defined for each of the five subregions gives the smoothest field (Fig. 5d). In particular, this method is capable of removing also the singularity along the Ryukyu islands. Thus, we call this the optimal nonlinear DMQC method. The optimality of the sub-region based approach comes from its trade-off in terms of resolution. The uniform approach is unable to capture local differences in the ocean mean state and variability in our study region. On the contrary, the grid-box approach requires a large number of observations with uniform coverage, which is unlikely to be achieved in practice, leading to noisy objectively analyzed fields. Similar findings hold for the distribution of sea surface salinity (Figs. 6a-6d).

The second sensitivity experiment is done to check the number of iterations needed for the nonlinear DMQC to ‘converge’. To show the effects we show the salinity distributions at 1000 m (Figs. 6e and 6f) where the successive climatological check iterations are applied once (Fig. 6e) or twice (Fig. 6f). By comparing Figs. 6e and 6f, it is shown that the second-time application of the climatological check removes structures in the field that can be considered noise. For the sake of striking a balance between keeping sufficient data and carrying out adequate climatological DMQC procedures, we believe applying the climatological check twice is the optimal choice for the present case.

4. EVALUATION OF THE COMPUTED CLIMATOLOGY

a. COMPARISON WITH EXISTING DMQC PROFILE DATASETS

The optimal nonlinear DMQC monthly mean vertical profiles are compared with the corresponding Argo profiles processed by two other QC approaches: (1) the WOD13 profiles flagged 0 (“accepted” value, Boyer et al. 2013) using the Johnson et al. (2013)

DMQC method; and (2) the GDACs profiles flagged 1 (“good” value, Argo Data Management Team 2012) based on the DMQC method of Owens and Wong (2009).

Fig. 7 shows the standard deviation of the monthly climatological profiles averaged over the Kuroshio subregion IV from our optimal nonlinear DMQC, WOD13, and GDAC respectively. The nonlinear DMQC method monthly mean climatologies show smaller standard deviations at almost every depth for both temperature (Fig. 7a) and salinity (Fig. 7d) with respect to WOD13 (Figs. 7b and 7e) and GDAC (Figs. 7c and 7f). Similar results are found for the other subregions (not shown).

b. COMPARISON WITH EXISTING GRIDDED CLIMATOLOGIES

The gridded Argo data after the nonlinear optimal DMQC method, thereafter called OA climatology, will be now assessed comparing with two other state-of-the-art gridded datasets, the WOA13 and RA in terms of temperature and salinity vertical and horizontal distributions.

Figs. 8 and 9 show the seasonal evolution of the upper-layer vertical thermohaline characteristics averaged over only the subregion IV and the Kuroshio region (subregions I to IV) from the three datasets: OA, RA and WOA13. OA has more similarities with WOA13 than with RA in all the regions, as it could be expected because both climatologies are computed only from observational data without data assimilation. OA (Figs. 8a and 8b) suggests stronger salinity ventilation of the intermediate layer than RA and WOA13 while OA and WOA13 are similar for temperature and about 20 C warmer than RA. This is consistent with the finding that Argo observed seawater is warmer than that from other-type observations almost everywhere in the global oceans. Moreover, the subsurface saltier band (e.g., 34.8 psu in Fig. 9) in the warm season is better defined in OA (Fig. 9b) than RA (Fig. 9d) and

WOA13 (Fig. 9f), with a larger layer thickness for the intermediate water salinities in the Kuroshio region.

Fig. 10 shows the salinity differences between any two of the three datasets OA, RA and WOA13, at four selected depths: surface, 50 m, 100 m, and 200 m. Generally, better gridded fields would be expected from the combination of multi-type measurements (e.g., Schmid 2005) but here Argo is the dominant dataset and OA is similar to both RA and WOA13 even if it contains only Argo data. The bigger differences mainly occur at the upper 50 m depths (dark colors in Figs. 10a-10f) near the coasts and boundaries (islands) etc., where traditional CTD measurements are twice those of the Argo profiles (Fig.11). The number of Argo observations in OA only accounts for ~40% of the number of the whole in-situ observations in WOA13. The largest differences occur between RA with WOA13 (Figs. 10c and 10f) in the marginal China Seas (e.g., East China, Yellow and Bohai Seas) and the straits (e.g., Tsushima and Taiwan Straits) where traditional data are scarce and RA arguably gives a better interpolation across the shelf than WOA13 because it takes dynamical constraints into account.

Tables 3 and 4 summarize the subregional statistics from the three datasets OA, RA and WOA13 for temperature and salinity at the sea surface where there are usually the largest differences between the different datasets (e.g., Fig. 10) in both cold (February) and warm (August) seasons. The agreement between OA and the other gridded datasets is better in winter than in summer and, in general, the correlations are always larger than 0.6 (statistically significant at the 95% confidence level) except for subregions I, II and III in some seasons.

5. DISCUSSION

A nonlinear DMQC method has been developed and applied to the Kuroshio region and to the 2005-2012 Argo comprehensive dataset. The new nonlinear DMQC method consists of: 1) linear profile QC for gross errors; 2) statistical checks done with climatologies and related standard deviations repeated interatively twice and done using OA mapped fields to generate gridded climatologies at intermediate and final stages of the DMQC procedure. Climatological values and related standard deviations used in the nonlinear DMQC method are shown to be better if defined in regime-oriented subregions I through V instead of a large scale area or in a spatially uniform grid.

The resulting nonlinear DMQC Argo profiles show a reduction of the standard deviations below 1000 m if compared with other DMQC approaches, such as WOD13 and GDAC, indicating the capability of the method to better capture and remove the representativeness measurement errors.

The derived gridded product, so-called OA, was then compared with analogous gridded climatologies, WOA13 and RA. In general the three climatological estimates are consistent and well correlated even if in OA only Argo data are used, i.e. about 40% of the actual overall T-S profile datasets used in WOA13.

In the future the DMQC could be improved by utilizing the complete WOD13 dataset and the procedure could be implemented in the global ocean where subregions of dynamical importance should then be defined to assess the correct standard deviations for quality control statistical checks. Such monthly mean climatologies offer still today the most convenient way to initialize operational forecasting systems and be the reference mean fields for seasonal forecasts.

ACKNOWLEDGEMENTS

W.Jia and D.Wang received funding by the Italian Ministry of Education, University and Research and the Italian Ministry of Environment, Land and Sea under the GEMINA project and OOCC China-Italy collaboration. N.Pinardi and S. Simoncelli were supported by the SeaDataNet project at INGV. We thank Dr. P. M. Poulain for helpful suggestions.

BIBLIOGRAPHY

Anderson, E. and Järvinen, H. (1999), Variational quality control. Q.J.R. Meteorol. Soc., 125: 697–722. doi: 10.1002/qj.49712555416.

Argo Data Management Team, 2012: Argo Quality Control Manual, Version 2.7. Argo Data Management, 47 pp. [Available online at http://www.argodatamgt.org/content/download/341/2650/file/argo-quality-control-manual-V2.7.pdf.]

Argo Science Team, 2012: Report of 4th Argo Science Workshop Meeting and final round table summary (ASW-4), Venice, Italy, 9 pp. [Available online at http://www.argo.ucsd.edu/ASW4_report.pdf.]

Argo Steering Team, 2011: Report of 12th Meeting of the International Argo Steering Team (AST-12), Buenos Aires, Argentina, 123 pp. [Available online at http://www.argo.ucsd.edu/iast12.pdf.]

Argo Steering Team, 2013: Report of 14th Meeting of the International Argo Steering Team (AST-14), Wellington, New Zealand, 121 pp. [Available online at http://www.argo.ucsd.edu/iast14.pdf.]

Barker, P. M., J. R. Dunn, C. M. Domingues, and S. E. Wijffels, 2011: Pressure Sensor Drifts in Argo and Their Impacts. J. Atmos. Oceanic Technol., 28, 1036–1049, doi:10.1175/2011JTECHO831.1.

Böhme, L., and U. Send, 2005: Objective analyses of hydrographic data for referencing profiling float salinities in highly variable environments. Deep-Sea Res. II, 52, 651–664, doi:10.1016/j.dsr2.2004.12.014.

Boyer, T. P., J. I. Antonov , O. K. Baranova, H. E. Garcia, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, T. D. O'Brien, D. Seidov, I. V. Smolyar, M. M. Zweng, 2009: World Ocean Database 2009. S. Levitus, Ed., NOAA Atlas NESDIS 66, 216 pp. [Available online at http://www.nodc.noaa.gov/OC5/indprod.html.]

Boyer, T. P., J. I. Antonov, O. K. Baranova, C. Coleman, H. E. Garcia, A. Grodsky, D. R. Johnson, T. D. O'Brien, C. R. Paver, R. A. Locarnini, A. V. Mishonov, J. R. Reagan, D. Seidov, I. V. Smolyar, and M. M. Zweng, 2013. World Ocean Database 2013. S. Levitus, Ed., A. Mishonov Tech. Ed., NOAA Atlas NESDIS 72, 209 pp. [Available online at http://www.nodc.noaa.gov/OC5/indprod.html.]

Bretherton, F. P., R. E. Davis, and C. B. Fandry, 1976: A technique for objective analysis and design of oceanic experiments applied to Mode-73. Deep-Sea Res., 23, 1B, 559-582.

Bryden, H. L., D. H. Roemmich, and J. A. Church, 1991: Ocean heat transport across 24°N in the Pacifi. Deep-Sea Res., 38, 297-324.

Carter, E. F., and A. R. Robinson, 1987: Analysis model for the estimation of oceanic fields. J. Atmos. Oceanic Technol., 4, 49-74.

Gaillard, F., E. Autret, V.Thierry, P. Galaup, C. Coatanoan, and T. Loubrieu, 2009: Quality control of large Argo data sets. J. Atmos. Oceanic Technol., 26, 337–351.

Johnson, D. R., T. P. Boyer, H. E. Garcia, R. A. Locarnini, O. K. Baranova, and M. M. Zweng, 2013. World Ocean Database 2013 User's Manual. Sydney Levitus, Ed.; Alexey Mishonov, Technical Ed.; NODC Internal Report 22, NOAA Printing Office, Silver Spring, MD, 172 pp. [Available online at http://www.nodc.noaa.gov/OC5/WOD13/docwod13.html.]

Kobashi, F., H. Mitsudera, and S.-P. Xie, 2006: Three subtropical fronts in the North Pacific: Observational evidence for mode water-induced subsurface frontogenesis. J. Geophys. Res., 111, C09033, doi:10.1029/2006JC003479.

Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, C. R. Paver, J. R. Reagan, D. R. Johnson, M. Hamilton, and D. Seidov, 2013: World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 73, 40 pp. [Available online at http://www.nodc.noaa.gov/OC5/indprod.html.]

Lorenc, A. C. and Hammon, O. (1988), Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. Q.J.R. Meteorol. Soc., 114: 515–543. doi: 10.1002/qj.49711448012.

Oka, E., 2005: Long-term sensor drift found in recovered Argo profiling floats. J. Oceanogr., 61, 775–781.

Owens, W. B., and A. P. S. Wong, 2009: An improved calibration method for the drift of the conductivity sensor on autonomous CTD profiling floats by θ-S climatology. Deep- Sea Res. I, 56, 450-457.

Park, J. J., and K. Kim, 2013: Deep currents obtained from Argo float trajectories in the Japan/East Sea. Deep-Sea Res. II, 85, 169–181, doi:10.1016/j.dsr2.2012.07.032.

Schmid, C., 2005: Impact of combining temperature profiles from different instruments on an analysis of mixed layer properties. J. Atmos. Oceanic Technol., 22, 1571–1587, doi: 10.1175/JTECH1785.1.

Storto, A. and S. Masina, 2013: The CMCC Global Ocean Physical Reanalysis System (C-GLORS) version 3.1: Configuration and basic validation. Tech. Rep. CMCC Research Paper RP0211, CMCC/ANS, Bologna, Italy.

Takahashi, T., S. C. Sutherland, C. Sweeney, A. Poisson, N. Metzl, B. Tilbrook, N. Bates, R. Wanninkhof, R. A. Feely, C. Sabine, J. Olaf- sson, and Y. Nojiri, 2002: Global sea-air CO2 flux based on climatological surface ocean pCO2, and seasonal biological and temperature effects. Deep-Sea Res. II, 49, 1601–1622.

Wong, A. P. S., G. C. Johnson, and W. B. Owens, 2003: Delayed-mode calibration of autonomous CTD profiling float salinity data by θ-S climatology. J. Atmos. Oceanic Technol., 20, 308-318.

Yu, L., and R. A.Weller, 2007: Objectively analyzed air-sea heat fluxes for the global ice-free oceans (1981-2005). Bull. Amer. Meteor. Soc., 88, 527-539.

Zhang, C., J. Xu, X. Bao, and Z. Wang, 2013: An effective method for improving the accuracy of Argo objective analysis. Acta Oceanol. Sin., 32, 66–77, doi:10.1007/s13131-013-0333-1.

Zweng, M. M., J. R. Reagan, J. I. Antonov, R. A. Locarnini, A. V. Mishonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, D. R. Johnson, D. Seidov, and M. M. Biddle, 2013: World Ocean Atlas 2013, Volume 2: Salinity. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp. [Available online at http://www.nodc.noaa.gov/OC5/indprod.html.]

The CMCC Foundation Euro-Mediteranean Centre on Climate Change has its registered office and administration in Lecce and other units in Bologna, Venice, Capua, Sassari, Viterbo and Milan. The CMCC Foundation doesn't pursue profitable ends and aims to realize and manage the Centre, its promotion, and research coordination and different scientific and applied activities in the field of climate change study

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici

Table 1. Vertical grid for OA mapping, consistent with the standard levels of the WOD13 (Boyer et al. 2009). Data in the profiles are averaged on the different layers to give an estimate at the depth corresponding to the layer center.

Level	Depth [m]	Layer
1	Surface	5
2	10	10
3	20	10
4	30	10
5	50	30
6	75	30
7	100	30
8	125	30
9	150	50
10	200	50
11	250	50
12	300	100
13	400	100
14	500	100
15	600	100
16	700	100
17	800	100
18	900	100
19	1000	100
20	1100	100
21	1200	100
22	1300	100
23	1400	100
24	1500	250
25	1750	250
26	2000	250

Table 2. Summary of the OA parameter sensitivity experiments in the study domain (21-42°N, 115-145°E). The OA correlation function is defined as:

$$F(x,y) = \left(1 - \frac{\sqrt{x^2 + y^2}}{a}\right) e^{-\frac{(x^2+y^2)}{2b^2}}$$

where *a* is the de-correlation scale, and *b* is the e-folding scale. A maximum of 30 observations, within a radius of 150 km from the estimation grid point, are considered in all experiments.

Case	Grid resolution [°]	a [km]	b [km]
1	1/2 x 1/2	300	180
2	1/2 x 1/2	200	120
3	1/4 x 1/4	300	180
4	1/4 x 1/4	200	120

Table 3. Summary of the five subregions (I, II, III, IV, and V) statistics for comparison of sea surface temperature between OA and RA (or WOA13) in winter (Feb) and summer (Aug). **BIAS** is the spatially averaged difference between OA and RA (or WOA13). **CORR** is the spatial correlation coefficient between OA and RA (or WOA13). **RMSE** is the root mean square error of OA relative to RA (or WOA13), quantifying the absolute differences between two datasets. **N** is the number of data points used for the statistics calculations. All the correlations are statistically significant at the 95% confidence level.

Temperature			BIAS	CORR	RMSE	N
I	**OA-RA**	**Feb**	0.0	0.8	0.6	733
		Aug	0.0	0.2	0.3	723
	OA-WOA13	**Feb**	0.0	0.9	0.5	733
		Aug	0.1	0.5	0.3	723
II	**OA-RA**	**Feb**	0.0	0.7	0.6	736
		Aug	-0.1	0.2	0.4	755
	OA-WOA13	**Feb**	0.0	0.2	0.6	736
		Aug	-0.1	0.5	0.4	755
III	**OA-RA**	**Feb**	0.1	0.9	0.6	1036
		Aug	0.1	0.6	0.3	1033
	OA-WOA13	**Feb**	0.0	0.9	0.5	1036
		Aug	0.0	0.7	0.3	1033
IV	**OA-RA**	**Feb**	0.3	0.9	0.8	1444
		Aug	0.1	0.7	0.4	1442
	OA-WOA13	**Feb**	0.1	0.9	0.7	1444
		Aug	0.1	0.7	0.4	1442
V	**OA-RA**	**Feb**	0.0	0.9	1.0	779
		Aug	0.0	0.8	0.9	913
	OA-WOA13	**Feb**	0.1	0.9	0.9	779
		Aug	0.3	0.8	1.0	913

Table 4. Same as Table 3, but for salinity.

Salinity			DIFF	CORR	RMSE	N
I	**OA-RA**	**Feb**	-0.02	0.7	0.08	652
		Aug	0.03	0.4	0.21	668
	OA-WOA13	**Feb**	0.01	0.8	0.07	652
		Aug	0.04	0.6	0.14	668
II	**OA-RA**	**Feb**	0.02	0.6	0.07	730
		Aug	0.0	0.3	0.11	745
	OA-WOA13	**Feb**	0.02	0.7	0.07	730
		Aug	0.01	0.5	0.10	745
III	**OA-RA**	**Feb**	0.0	0.4	0.06	1037
		Aug	0.01	0.7	0.13	1038
	OA-WOA13	**Feb**	0.0	0.7	0.05	1037
		Aug	0.02	0.9	0.10	1038
IV	**OA-RA**	**Feb**	0.01	0.8	0.05	1438
		Aug	0.12	0.7	0.19	1354
	OA-WOA13	**Feb**	0.04	0.6	0.09	1438
		Aug	0.09	0.7	0.19	1354
V	**OA-RA**	**Feb**	0.04	0.7	0.08	946
		Aug	0.13	0.6	0.29	918
	OA-WOA13	**Feb**	0.09	0.6	0.17	946
		Aug	0.11	0.6	0.26	918

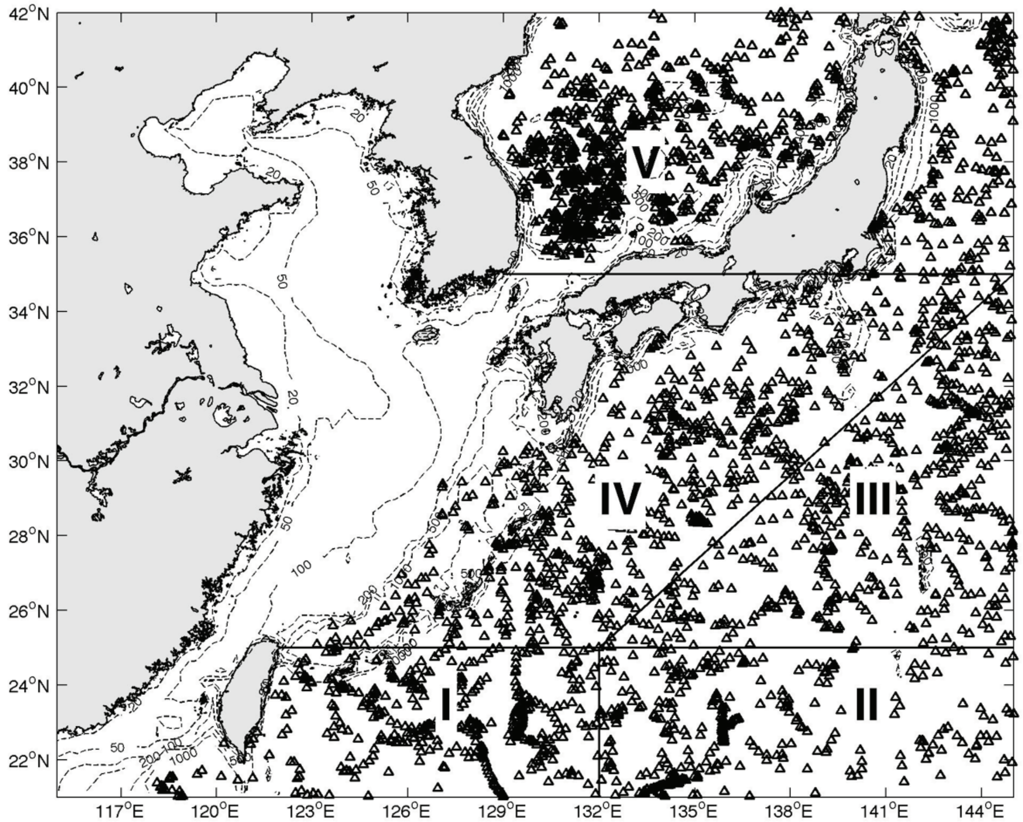

Figure 1. The study domain (115-145°E, 21-42°N) with selected bathymetry contours (20, 50, 100, 200, 500 and 1000 m dashed lines), and the August Argo profile positions (triangles) during the period 2005-2012 and the subregions (I, II, III, IV, and V, separated by thick solid lines).

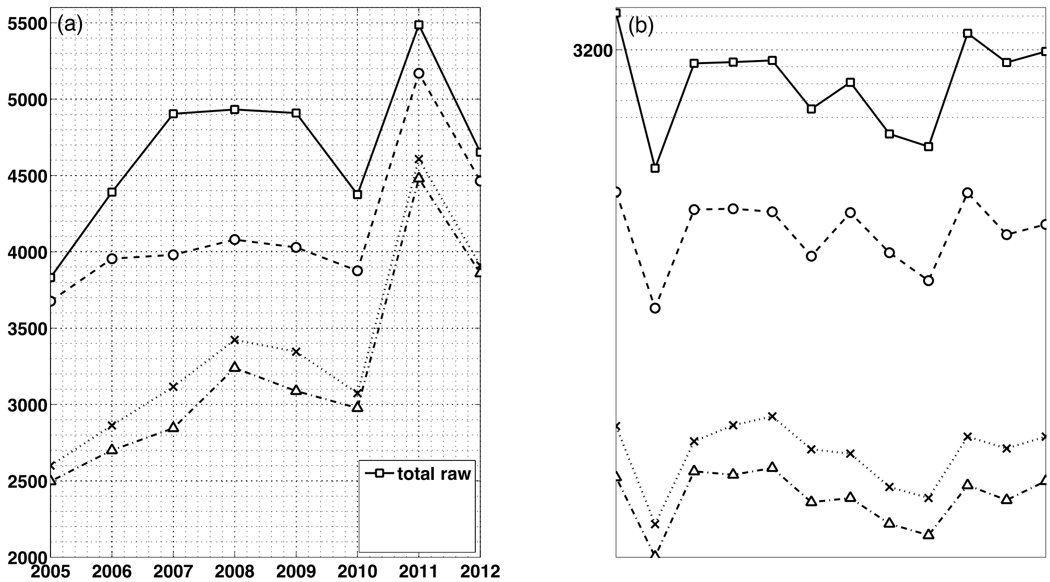

Figure 2. Number of Argo profiles by year (a) and by month (b) during the period 2005-2012 over the entire study domain and subregions I through IV (Fig. 1). Solid line with square (dotted line with cross) represents the number of raw profiles in the whole region of Fig. 1 (only in the subregions I-IV); dashed line with circles (dashed line with triangles) represents the number of profiles after linear DMQC in the total region (only the subregions I-IV).

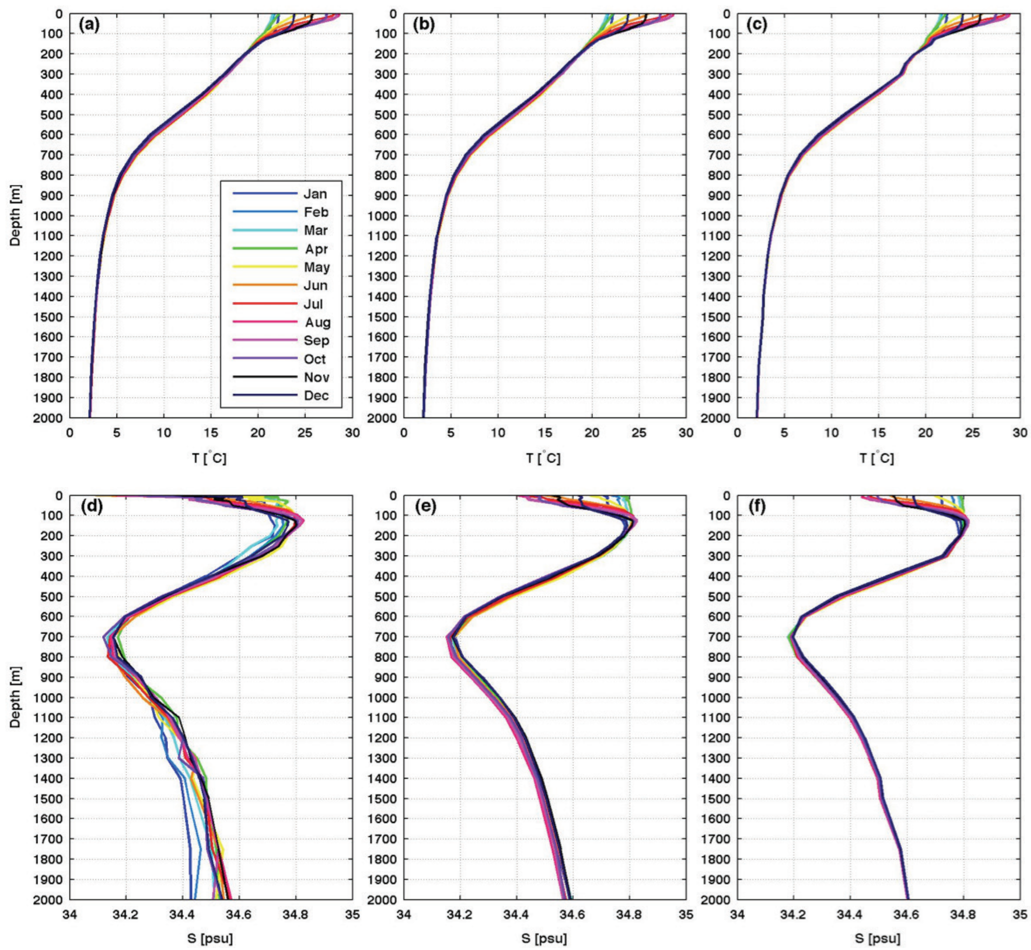

Figure 3. The area-averaged monthly mean Argo profiles (2005-2012) for temperature (a-c) and salinity (d-f) over the subregions I through IV in Fig. 1. Left panel (a, d): raw data; Middle panel (b, e): data after linear DMQC; Right panel (c, f): data after nonlinear DMQC.

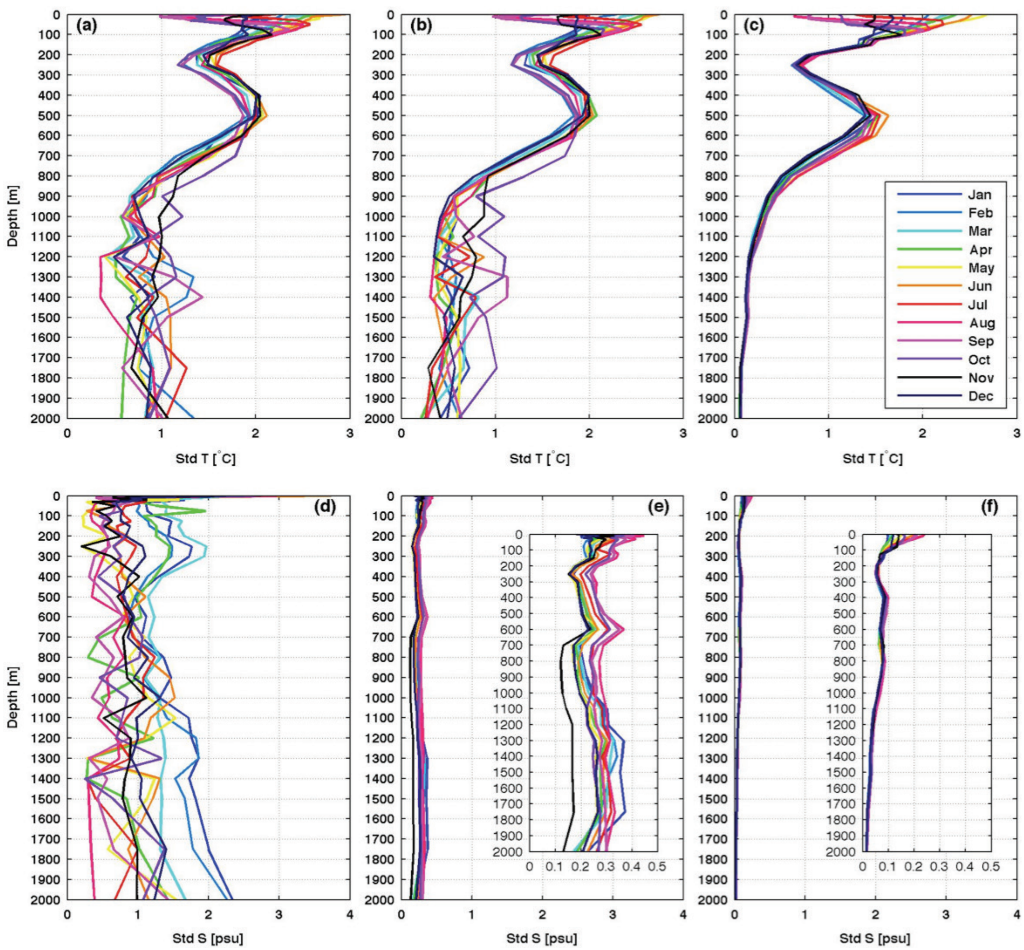

Figure 4. Standard deviation of the Kuroshio area-averaged monthly Argo profiles for temperature (a-c) and salinity (d-f). Left panel (a, d): raw data; Middle panel (b, e): data after linear DMQC; Right panel (c, f): data after nonlinear DMQC.

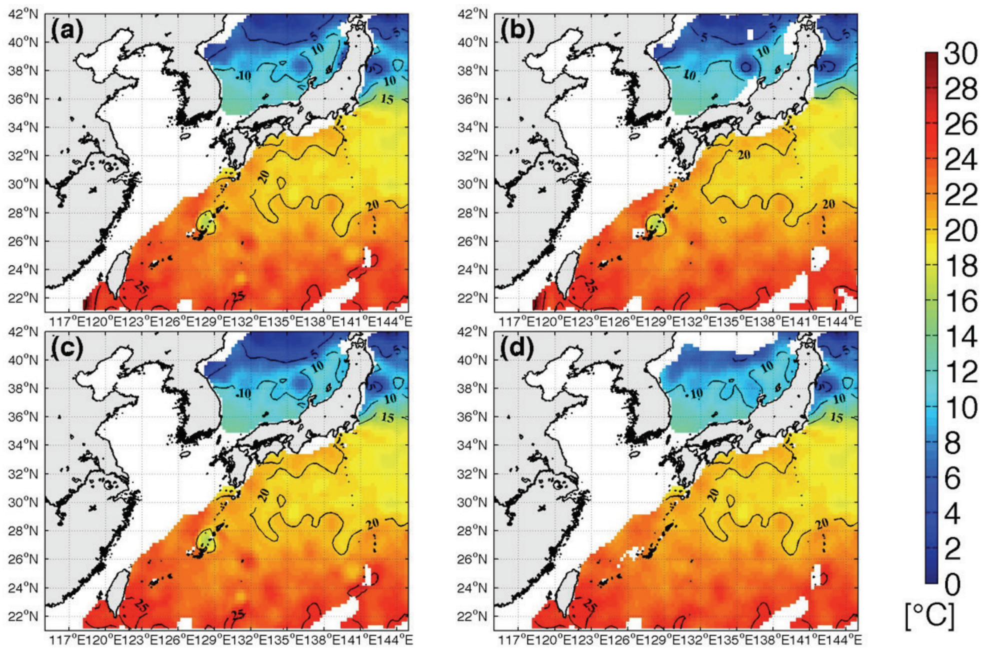

Figure 5. The objective-analyzed (OA) distribution of sea surface temperature in February (2005-2012) from Argo data, with four DMQC methods: (a) linear DMQC only, (b) nonlinear DMQC with a climatological check value given in a regular 1°x1° grid, (c) nonlinear DMQC with a climatological check value for the entire region, and (d) nonlinear DMQC with a climatological check value for I-V subregions. OA errors greater than 30% are masked with white.

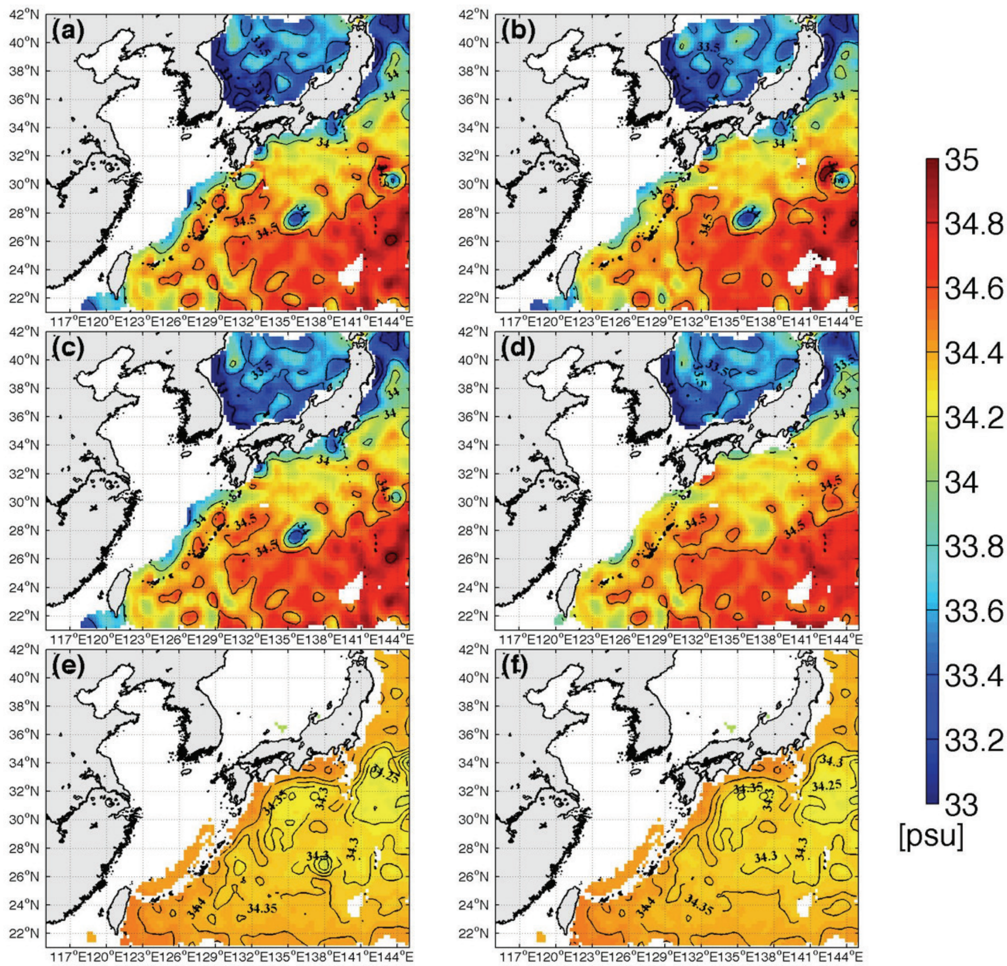

Figure 6. (a)-(d) are the same as Figs. 5 (a)-(d), but for sea surface salinity in August. (e) and (f) are the salinity distributions at the 1000 m depth with the same DMQC method as in (d), but (e) applied the climatology check only once while both (d) and (f) applied the climatology check twice.

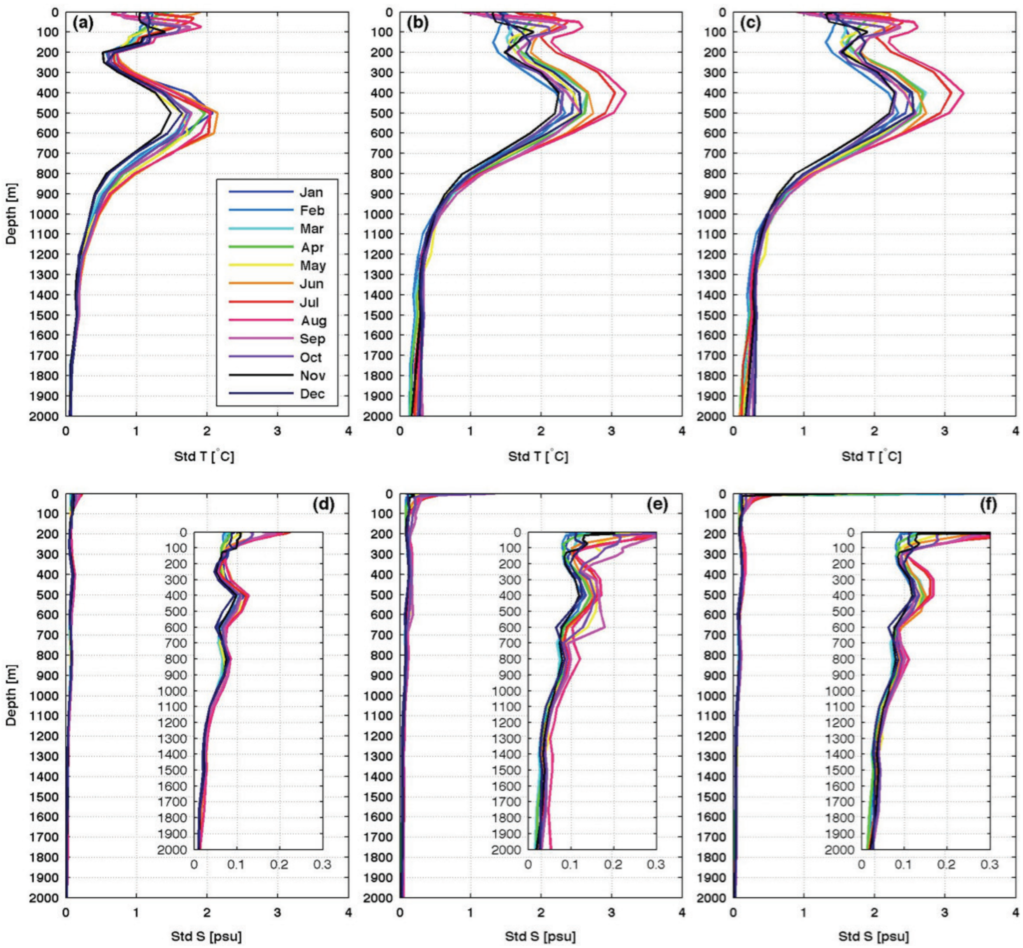

Figure 7. Comparisons of the standard deviation of monthly averaged Argo profiles (2005-2012) over subregion IV (Fig. 1) between the optimal nonlinear DMQC for temperature (a) and salinity (d) and two other DMQC Argo datasets: 1) the WOD13 for temperature (b) and salinity (e), and 2) from the GDACs for temperature (c) and salinity (f).

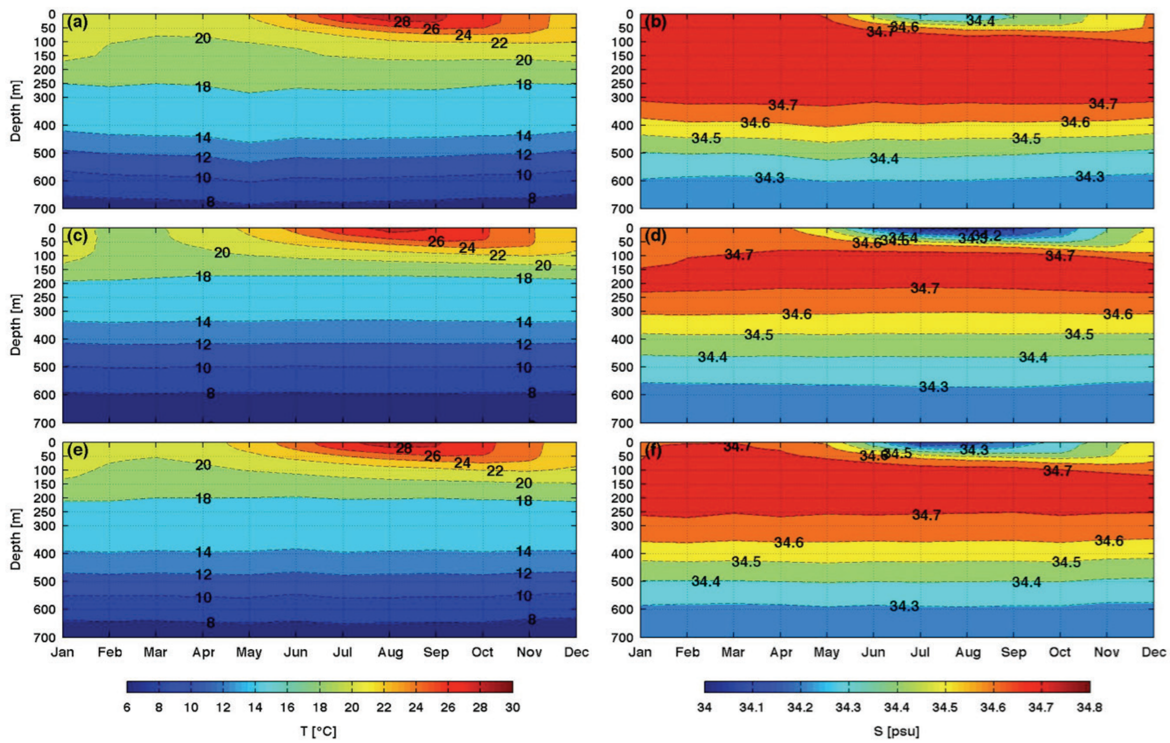

Figure 8. Evaluation of the OA-gridded Argo dataset (2005-2012) subregion IV (Fig.1) averaged seasonal evolution of vertical temperature (a) and salinity (b), with those from two coordinated databases: RA (c: temperature, d: salinity) and WOA13 (e: temperature, f: salinity)

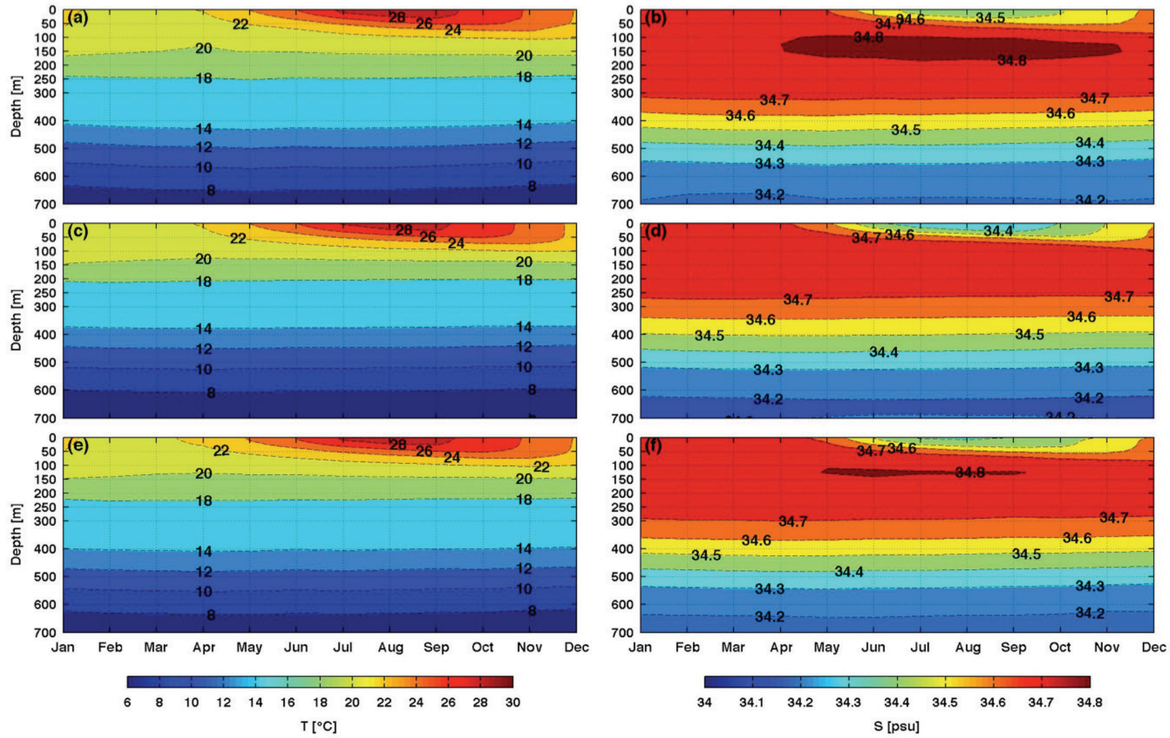

Figure 9. Same as Fig. 8, but for the Kuroshio region (subregions I through IV in Fig. 1).

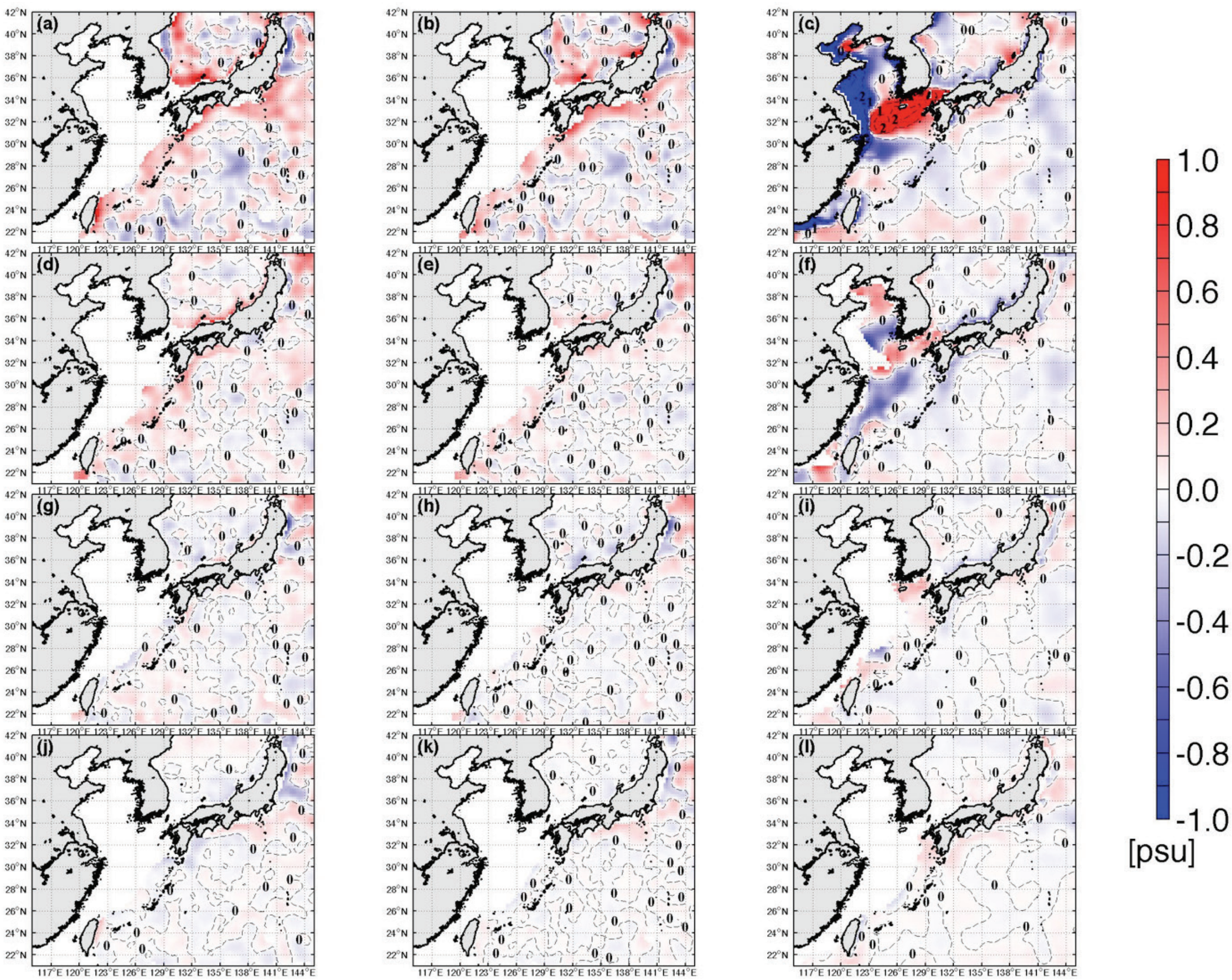

Figure 10. Differences of the salinity distributions in August (2005-2012) between OA and RA (left: a, d, g, and j), between OA and WOA13 (middle: b, e, h, and k), and between RA and WOA13 (right: c, f, i, and l), at the surface (a, b, c), 50 m (d, e, f), 100 m (g, h, i), and 200 m (j, k, l) depths from upper to lower panels, respectively.

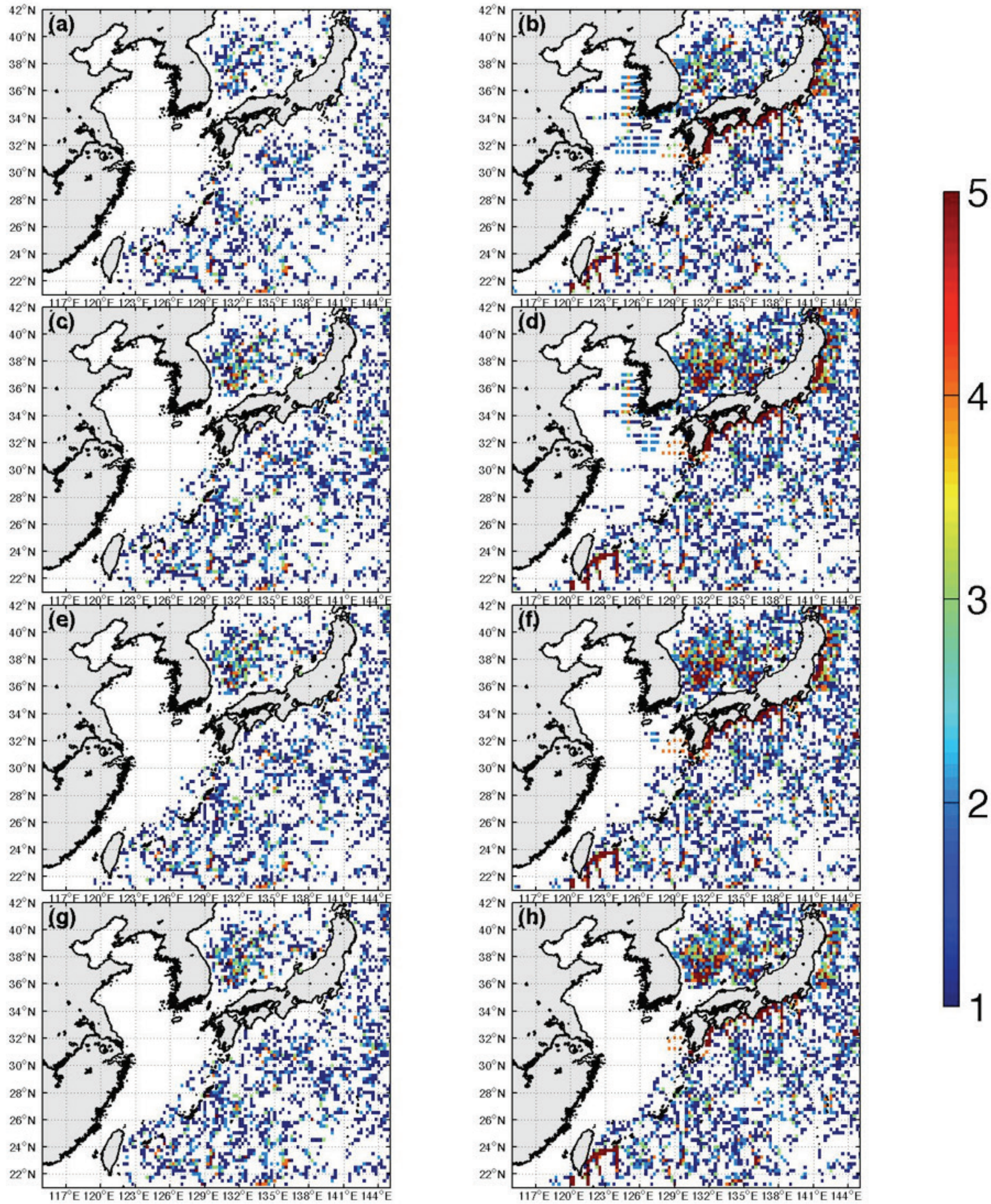

Figure 11. Comparisons of number of observations in each 1/4° grid box between Argo data in OA (left: a, c, e, and g) and all the in-situ measurements used in WOA13 (right: b, d, f, and h) for August (2005-2012) at the surface (a, b), 50 m (c, d), 100 m (e, f), and 200 m (g, h) depths from upper to lower panels, respectively.