An operational objective analysis at INCOIS for generation of Argo value added products

TVS Udaya Bhaskar, M Ravichandran and R Devender

INCOIS Technical Report:
INCOIS-MOG-ARGO-TR-04-2007

Technical Report Report No. : INCOIS-MOG-ARGO-TR-04-2007

An operational Objective Analysis system at INCOIS for generation of Argo Value Added Products

By

TVS Udaya Bhaskar, M. Ravichandran and R Devender

Indian National Center for Ocean Information Services
(Ministry of Earth Sciences, Govt. of India)
Hyderabad

April, 2007

DOCUMENT CONTROL SHEET

01. Report No: **INCOIS-MOG-ARGO-TR-04-2007** Date: **27 April 2007**

02. Title & Sub Title:

An operational Objective Analysis system at INCOIS for generation of Argo Value Added Products.

03. Part No.: — 04. Vol. No.: —

05. Author(s):

TVS Udaya Bhaskar, M Ravichandran and R Devender

06. Originating agency (Group/Project/Entity): **MOG**

07. No. of Pages: **39** 08. No. of figures: **42**

09. No. of references: **12** 10. No. of enclosures/appendices: **4**

11. Abstract (Maximum 100 words):

The system of objective analysis used at Indian National Centre for Ocean Information Services is described. It is a integral part of the Argo data processing system, and designed to operate with minimum of manual intervention. The analysis method, based mainly on the method of McCreary and Kessler is a method of estimating Gaussian weight for all the observation used for estimating the value at grid location. The errors are determined from a comparison of the observation with the estimated value. The analysis system is very flexible, and has been used to analyse many different types of variables.

12. Keywords: **Argo, Objective Analysis, Gridding, Value added products**

13. Security classification: **Unrestricted**

Table of Contents

History of the Document

Version	**Date**	**Comment**
1.0	29/04/2007	Creation of the document.
1.1	15/03/2010	Contents of methodology modified
2.0	17/05/2013	Document modified to reflect the changes to the methodology: error estimation added. Images of errors and analysed fields added
2.0	07/04/2014	Comparison figures with RAMA and OMNI buoys added

TVS Udaya Bhaskar, M Ravichandran, R Devender, 2007, An operational objective analysis at INCOIS for generation of Argo value added products, INCOIS, Technical Report No. INCOIS-MOG-ARGO-TR-04-2007, April 2007.

Abstract

A 10day and monthly gridded datasets of temperature and salinity was produced using data from Argo profiling floats in the Indian Ocean. The generation of gridded product is a integral part of the Argo data processing system, and designed to operate with minimum of manual intervention. The individual profiles are passed through various quality checks before being used for gridded product generation. The analysis method, based mainly on the method of McCreary and Kessler is a method of estimating Gaussian weight for all the observation used for estimating the value at grid location. The errors are determined from a comparison of the observation with the estimated value. The analysis system is very flexible, and has been used to analyse many different types of variables. The dataset is freely available on the Indian National Centre for Ocean Information Services (INCOIS) Live Access Server (las.incois.gov.in).

Keywords: Argo float, objective analysis, Indian Ocean, gridded product.

1. Introduction

In geophysical science one come across a fundamental problem of how to use data collected at finite number of irregular locations, eventually at different times, to estimate value at any desired point of the space or space-time field. The ultimate goal of such process can be a simple visualization of the observed field using various softwares, use of the estimated field as input for numerical models, to mention a few among many other applications. A wide set of techniques have been developed for both diagnostic and prognostic studies/analyses. Since the development of computers in the last decades made possible the automatic implementation of these techniques, they have been referred to as spatial objective analysis (Arteaga, 2002). To put it in a nut shell Objective Analysis (OA) is the process of transforming data from observations at irregularly spaced points into data at points of regularly arranged grid.

Observational datasets of temperature and salinity are necessary for general monitoring the day to day changes in the sea state, initializing the model for better forecast, generation of real time forecasts and also for obtaining information on climate change. Historically the temperature alone data sets dominate the temperature, salinity paired data sets. Many gridded data sets were generated of which many are on climatological scale. For example Levitus has come up with many versions of climatologies (Eg: 1994, 1998, 2001, 2005 and latest of all 2009). White (1995) also constructed monthly mean temperature data sets from surface to subsurface from limited expendable bathythermographs (XBT) and hydrographic stations. The data sets of White has some disadvantages like shallow depth (400m) and coarser horizontal resolution (2.5º x 5º). The continuously updated climatology from Levitus is better both in-terms of maximum depth and horizontal resolution. However these data sets are dominated by the temperature alone profiles. Historically there were no global high-resolution salinity datasets because of non-availability of salinity profiles. This deficiency of salinity profiles were eliminated by the Argo program.

Argo is an internationally coordinated program directed at deploying the global ocean with 3000 profiling floats that measure temperature and salinity. The profiling float sinks after launch to a prescribed pressure level, typically 1000 dbars and after a preset time (typically 10 days) the float dives to 2000 dbars and returns to the surface, collecting temperature and salinity measurements (Argo Science Team, 2001; Ravichandran et al., 2004). On the surface the float transmits the data to ARGOS satellite which are received at the ground station and are analyzed. The target spatial density for the Argo float array is one float per 3º square in the world ocean. In late October 2007, the target of 3000 floats was achieved, providing complete coverage of the world ocean both spatially and temporally for the first time. It is with these data sets, 10 day and monthly mean gridded data sets of temperature and salinity for the Indian Ocean is constructed.

The Objective Analysis (OA) method proposed by Kessler and McCreary (1993) was used to construct a gridded dataset for the Indian Ocean. This method is simple in implementation and often used to estimate grid-point values from observations falling with in radius of influence. The remainder of the document is organized as follows. The theory of objective analysis is dealt in section 2, data and methodology is described in section 3 and section 4 summarises the work.

2. Objective Analysis definition

The basic problem of Objective Analysis is depicted in figure1. At any give instant all one have is irregularly spaced observations (Eg: from Argo profiling floats or ship based CTD along a track) that must be used to provide values for points on a regularly spaced grid. OA in general is the process of interpolating observed values onto the grid points used by the model/analysis in order to define the initial conditions of the atmosphere/ocean (Fred Carr, 1999).

One can question, why isn't this just a simple exercise in mathematical interpolation? The answer to this question are several, viz.,

1. One can use our historical knowledge of oceanic behavior to infer additional information from the data available in the area.

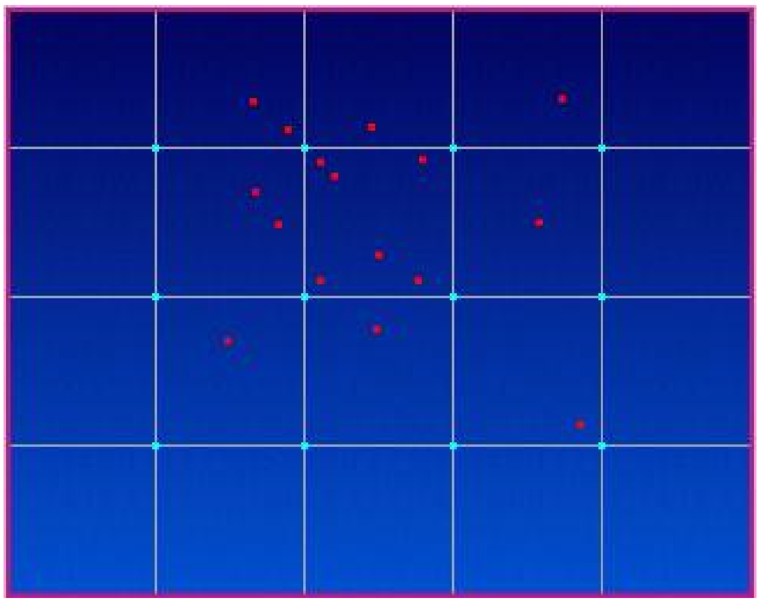


Figure1: Irregularly spaced observations (source: http://www.comet.ucar.edu). Red dots represent observations and blue dots are grid points at which analysis is to done.

2. One can adjust the analysis procedure to filter out scales of motion that can't be forecast by the model being used.

3. One can make use of a first guess field (generally from well know climatology like World Ocean Atlas) or background field provided by an earlier forecast from the same model. The blending of the background fields and the observations in the objective analysis process is especially important in data sparse areas. It allows us to avoid extrapolation of observation values into regions distant from the observation

sites. The background field can also provide detail (such as frontal locations that exist between observations).

Using a background field also helps to introduce dynamical consistency between the analysis and the model.

4. One can also make use of knowledge of the probable errors associated with each observation. One can weigh the reliability of each type of observation based on past records of accuracy.

2.1 Analysis Equation

Figure2 depicts the fundamental OA equation in worded form in order to illustrate the basic principles that contribute to a numerical oceanography analysis. In simplest terms, the OA equation attempts to determine the value of a particular oceanographic variable at a particular grid point (at a particular valid time). In words, the analysis equation can be expressed as shown below.

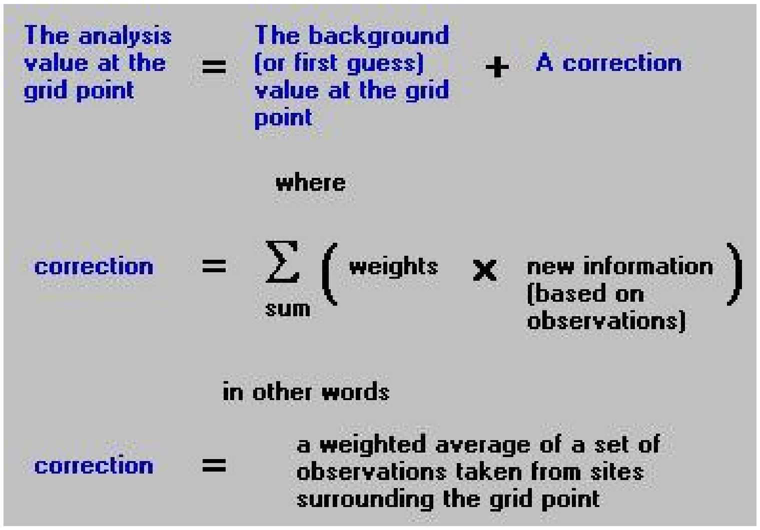


Figure 2: Objective Analysis equation (source: http://www.comet.ucar.edu).

2.2 Importance of Background Field

In the simplest kind of OA scheme, the background values would not be used and the analysis would be based solely on new observations. In this case the equation would become:

$$\text{The analysis value at the grid point} = \sum_{\text{sum}} \left(\text{weights} \times \text{new observations} \right)$$

The observations themselves would be interpolated to the grid point by calculating a weighted average of the data. (One type of weight, for example, is proportional to the distance of the data from the grid point. The farther an observation is from the grid point, the less weight it gets. If a grid point has no nearby observations, the simple scheme described here is in trouble.

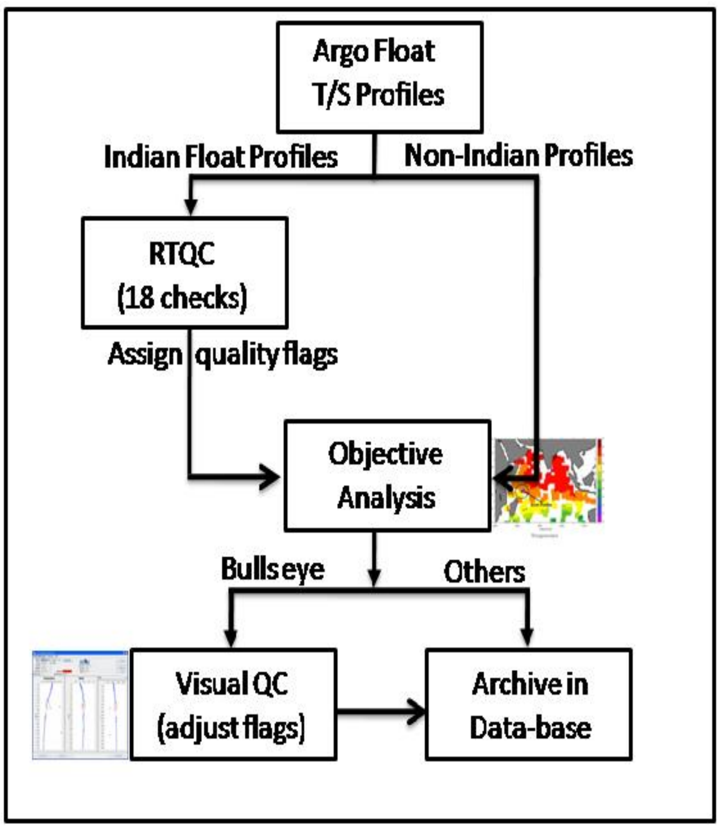


Figure3: Flow diagram showing the modules of three way quality control system implemented at INCOIS for quality control of temperature and salinity data obtained from Argo floats.

3. Data and Methodology

3.1 Data

The data used for generating the gridded product consisted of all the CTD data measured by Argo floats in the Indian Ocean region (30º E – 120º E and 30º S – 30º N). These profiles were obtained from the INCOIS archives which consists of floats deployed by Indian and also other countries in the Indian Ocean. Profile corresponding to non-Indian floats are made available by USGODAE and IFREMER Global Data Assembly Centres (GDACs). All the profiles are passed through a three way quality control system established at INCOIS. Figure 3 shows the flow of the three way quality control system in place at INCOIS for near real time quality control of Argo T/S profiles. The T/S profiles obtained from Argo floats are subjected to three levels of QC as described below:

- **Argo real-time QC:** The first level is the real-time system that performs a set of agreed checks (as prescribed by Argo Data Management Team (Wong et al., 2012)) on all profiles obtained from Argo floats. Real-time data with quality flags assigned are available to users within the 24 hrs timeframe at GDACs and Global Telecommunication Systems (GTS).
- **Objective Analysis:** The second level of QC involves usage of objective analysis method for identification of outlier which appear as bulls eyes.
- **Visual Quality Check:** The third level of quality control involves adjustment of suspicious quality flags set by the Argo RTQC procedures, by visual inspection.

Further details of the three way quality control procedures applied to the Argo profiles can be obtained from udaya bhaskar et al., (2012).

3.2 Data processing

3.2.1 Vertical coordinate transformation

Argo floats measures temperature and salinity against pressure. As we wanted to generated gridded data sets with Z axis in meters, we converted the pressure to depth using the method proposed by Saunders (1981). The relation between depth and pressure is given as

$$Z = \int_0^p \sim dp \,/ \left(g_s + \tfrac{1}{2} \yen' P\right) \qquad (1)$$

where Z is depth, α is specific volume and is reciprocal of density.

$\yen' = 2.226 \times 10^{-6}\ db^{-1}\ m\ s^{-2}$

$g_s = 9.780318(1 + 5.3024 \times 10^{-3}\ Sin^2\varphi - 5.9 \times 10^{-6}\ Sin^2\ 2\varphi)\ [m\ s^{-2}]$

3.2.2 Vertical interpolation and extrapolation

Since all Argo floats do not measure data at same pressure levels, the profiles are linearly interpolated to be used for objective analysis. All the profiles are linearly interpolated to 1 dbar and data corresponding to the Levitus standard depths were extracted to carry out OA for the data at these depths.

3.2.3 Additional checks using climatological data

To eliminate spurious data sets, an additional statistical check is performed with the use of World Ocean Atlas 2001 (conkright et al., 2001). The interpolated Argo profile is overlaid on the climatological mean and three standard deviation envelop. Those data which are falling out of the 3σ from the mean are termed as outlier and are eliminated.

4. Objective Analysis Methodology and Errors

4.1 Analysis Method

For the generation of the gridded product we used the methodology proposed by Kessler and McCreary (1993). The gridding was carried out in two steps as follows:

- First the temperature data Tn (x_n, y_n, z_n, t_n) from each profile n were linearly interpolated to standard depths (1 m from surface to 1000 m) there by creating a modified data set $T^{'}_n$ (x_n, y_n, z_n, t_n). This interpolation was done only when two sample in a profile are with in a selected vertical distances. For the surface data up to 500 meters the vertical distance should be less than 50 m. For deeper depths this should be less than 100 m. In case the vertical distance exceeds the given limits, interpolation is not carried out between these levels and values are filled with undefined values. Here (x_n,y_n) represents the longitude and latitude of the profile, z_n and t_n represents the depth and time of the n^{th} observation respectively.
- Second, once the interpolated profiles is obtained, data corresponding to the Levitus standard depth Z_0 (0, 10, 20, 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000 m) were segregated into different files. For each of the depths, the temperatures $T^{'}_n$ were mapped from irregular grid locations (x_n, y_n, z_n) to regular grid (x_0, y_0, z_0) locations with a grid spacing of 1° X 1° . The value of the gridded temperature $T^{''}$ at each grid point (x_0, y_0, t_0) was estimated by the operation

$$T''(x_0, y_0, t_0) = \frac{\sum_{n=1}^{N_p} T'_n W_n}{\sum_{n=1}^{N_p} W_n} \qquad (2)$$

where N_p is the total number of profiles data falling within the region of influence for a particular grid point. The influence of each of the points around the grid point (x_0,y_0) is controlled by the Gaussian weight function Wn which is in turn based on the distance of the point from then grid point (x_0,y_0). The Guassian weight function W_n in equation (2) is given by

$$W_n(x_n, y_n, t_n) = \exp\left\{-\left[\left(\frac{x_n - x_0}{X}\right)^2 + \left(\frac{y_n - y_0}{Y}\right)^2 + \left(\frac{t_n - t_0}{2}\right)^2\right]\right\} \tag{3}$$

where X = 3º longitude and Y = 3º latitude, time is assumed to be having no influence on the analysis. That is time is assumed to be constant while performing the analysis for the period under consideration (10days or one month) hence third term is zero. The choice of X and Y are based on the design of Argo that there should be one Argo float within one 3ºX3º box in the ocean. Kessler and McCreary (1993) mentioned that this operation is similar to a single iteration of objective mapping as used by Levitus (1982). From equation (3) one get to know that W_n has a non-zero value almost anywhere; so every observation around the grid point (x_0,y_0) no matter how far it is situated will have at least some influence on the analysis. Considering all the points will increase the number of computation, only those points which are within 3ºX3º distance are used and rest of the points are discarded.

In words of Kessler and McCreary, visualising the three dimensional grid (x_0, y_0, z_0) with data points $T'_n(x_n, y_n, z_n)$ scattered irregularly through it, the mapping operation appears as a ellipsoid moving from grid point to grid point averaging the points that fall with in that ellipse. Each data point falls with in summation of several grid points, weighted according to the distance. In the regions of very sparse sampling, a single data point may be the only information for one or several grid points. If no data points fell with in the ellipsoid at a grid point, then that was left blank (Kessler and McCreary, 1993).

4.2 Error estimation and statistical parameters

Once the value of temperature and salinity are estimated at the required grids (analysed value), error in the estimation is calculated based on the analysed value. All the observations falling within the radius of influence from the grid point is used to calculate the RMSE w.r.t to the analysed value at each grid. The RMSE is obtained as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(A - O_i\right)^2} \qquad (4)$$

where A is the analysed value at the grid and O_i s are the observations falling within the radius of influence. Apart from this various statistics like standard deviation, mean value, no of observations are also provided at each grid point. The analysed value along with these statistical parameters will provide enough confidence to the data user.

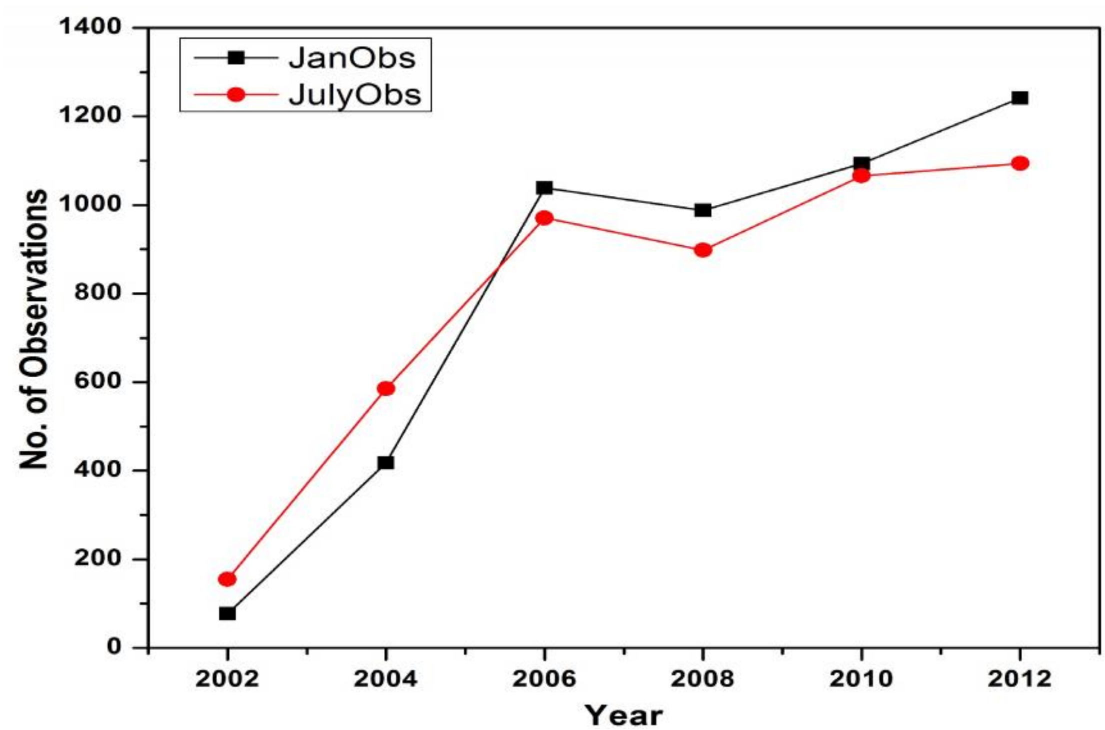


Figure 4. Total number of temperature and salinity observations (at 10 meters) in the Indian Ocean that went into OA for the month of January and July.

5 Data density and interpolation errors

Figure 4 shows the number observations at 10 m for two sample months January and July, that were used for generating the gridded product. Figure 5 shows the spatial

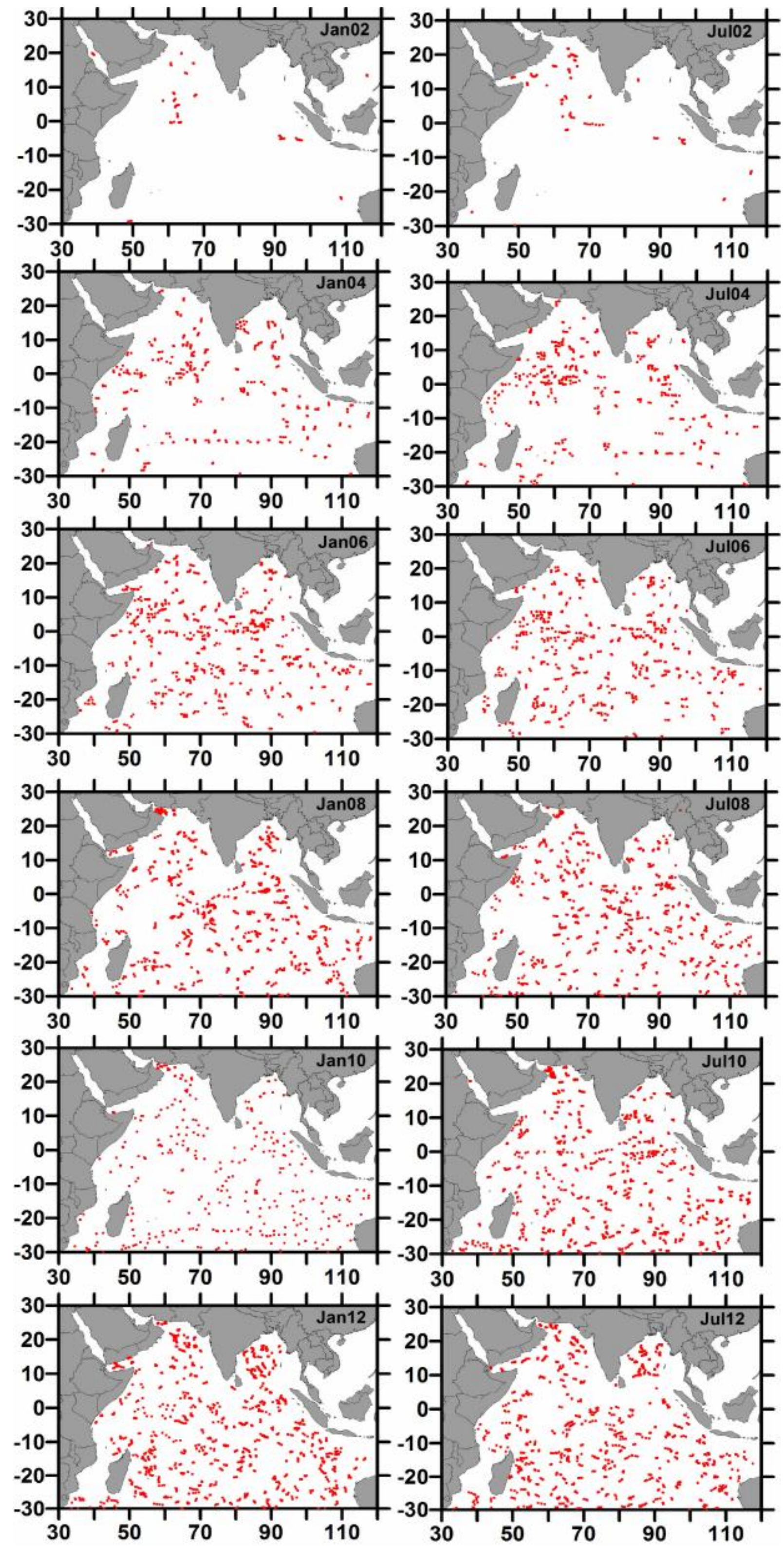


Figure 5. Data density maps for two representative months (January and July) which were used in OA for generation of gridded product.

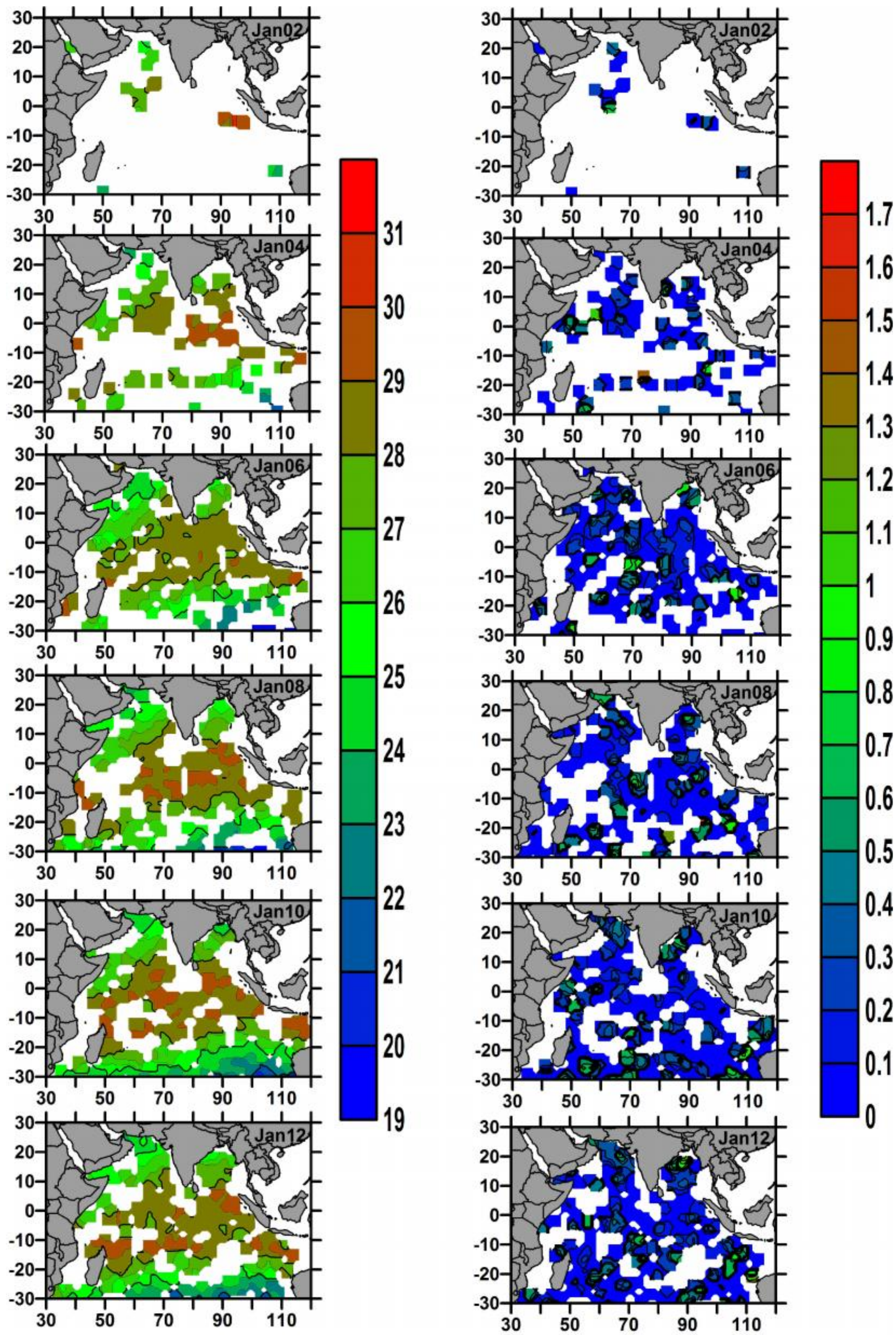


Figure 6a. Maps of interpolated temperature and RMSE for January for different years.

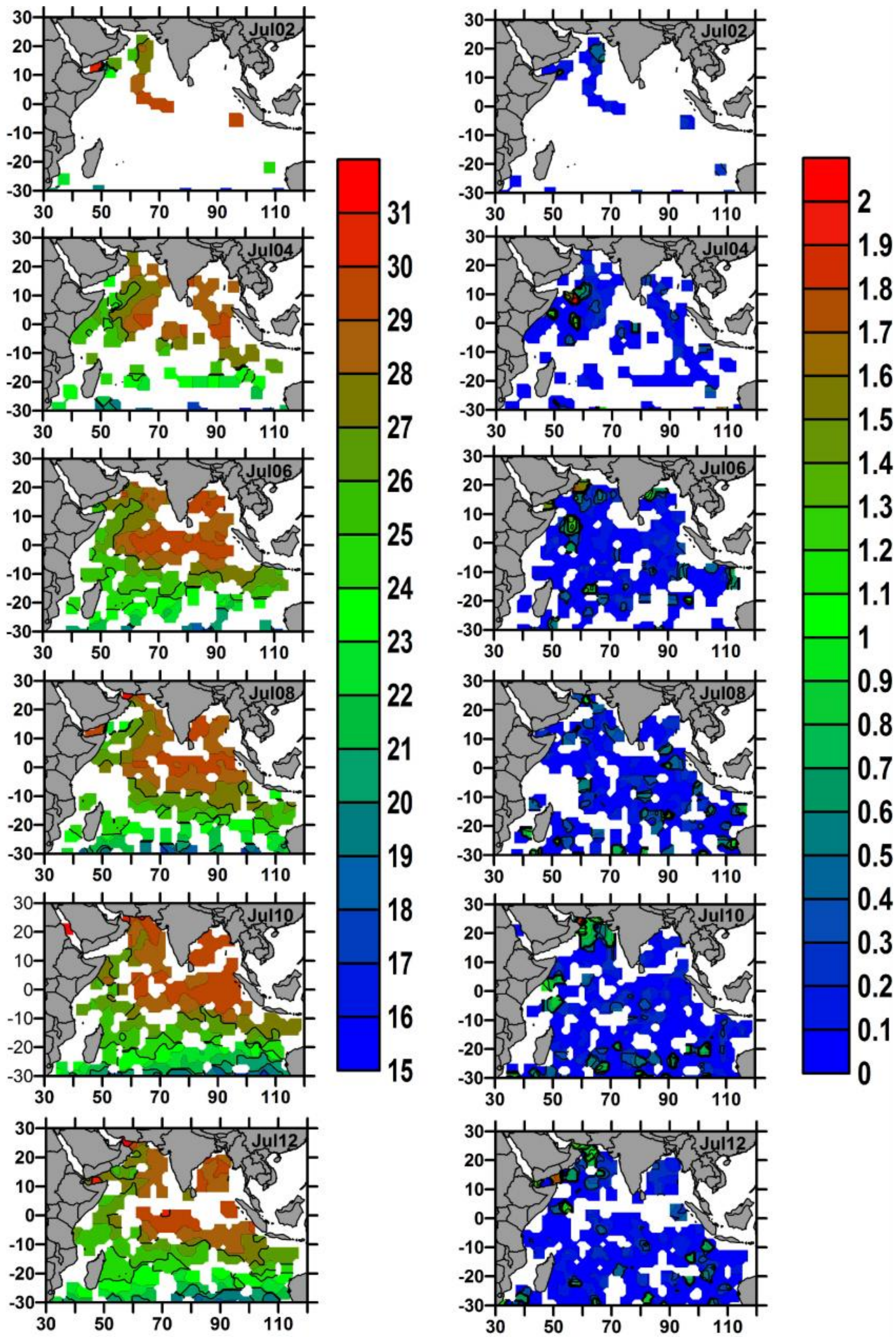


Figure 6b. Same as 6a but for the month of July.

distribution of these temperature and salinity observations. It is clearly evident that the number of observed points are insufficient during the initial years 2002. However starting from 2004, the data is observed to be evenly spread and by 2006 the entire Indian ocean is uniformly measured, thereby making the gridding more reliable. Figure 6 shows the map of temperature and error (RMSE) for the representative months of January and July for different years. Error in all the years are observed to be well below 0.3 except at certain locations. In the objective process, locations where not even a single data is observed in the radius of influence, is left intentionally blank. This was done not to unnecessarily extrapolate the data and create artificial data.

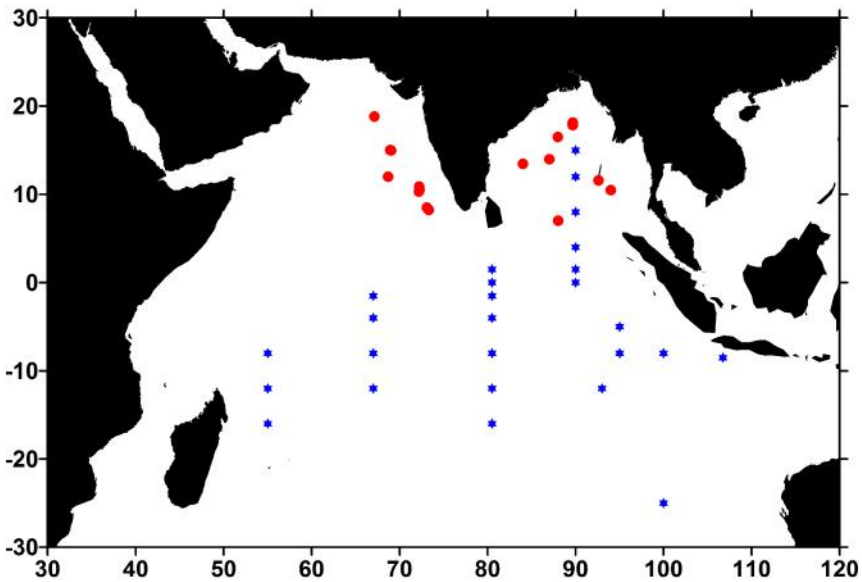


Figure 7. Locations of RAMA (Blue Star) and OMNI (Red closed circles) buoys used for validation of gridded product.

6. Validation of gridded product

To build confidence on the gridded product for its use in scientific studies, it is validated independently with sub-surface observations obtained from RAMA moorings and OMNI Buoys deployed by National Institute of Oceanography. More information about RAMA moorings can be obtained from McPhaden et, al., (2009). Figure 7 shows the locations of the RAMA and OMNI mooring in the tropical Indian Ocean. Under the Indian National Buoy Data program [Venkatesan *et. al.* 2013), National Institute of Ocean Technology (NIOT) under the Ministry of Earth Sciences (MoES), Govt. of India maintains a network of moored buoys (Ocean Moored buoy Network for Northern Indian Ocean (OMNI) buoys) to measure met-ocean parameter in the northern Indian Ocean. These moored buoys can measure met-ocean parameters on an hourly time scale, which provides a unique opportunity to study the upper ocean thermo-haline structure. These buoys transmit the real-time met-ocean data over a satellite network to Indian National Centre for Ocean Information Services (INCOIS), India. Received data undergo quality control procedures, archived and are made available through the Ocean Data Information System (ODIS) system of INCOIS [Shesu et al., 2013]. Subsurface data from these buoys are also used for independent validation of the gridded product.

For validation purpose, RAMA and OMNI buoys data is interpolated to match the temporal scale of Argo gridded product. Series of figures below shows the comparison of Argo gridded product with the OMNI and RAMA buoys. Figure 8 - 13 shows the comparison with OMNI buoys pertaining to Bay of Bengal and Arabian Sea.

6.1 Comparisons with OMNI Buoys

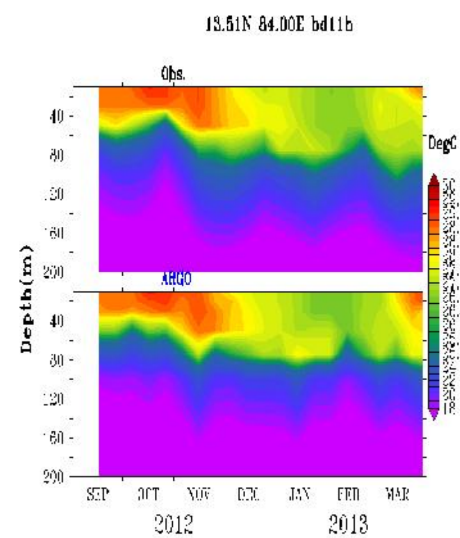
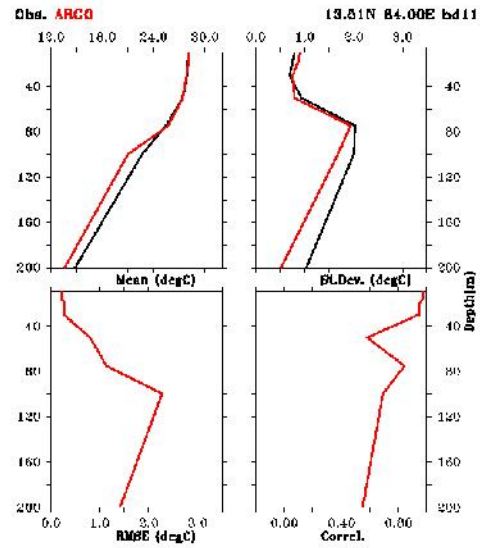


Figure 8(a) section plots of temperature from OMNI buoys (BD11) and Argo gridded product at 13.51 N and 84.0 E.



(b) Statistics viz., Average, Standard Deviation, RMSE, Correlation between the two data sets.

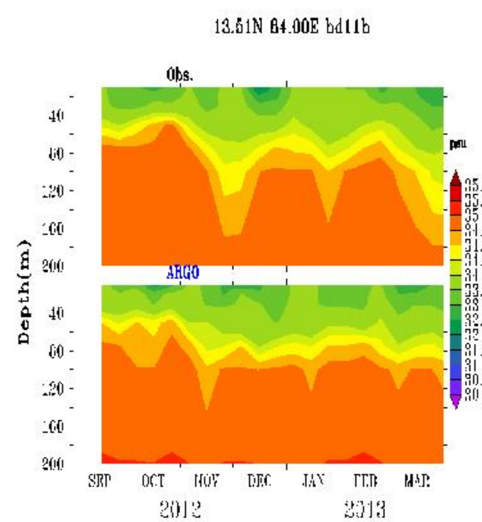
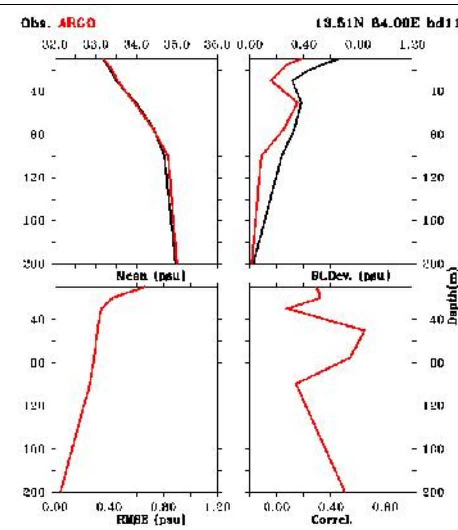
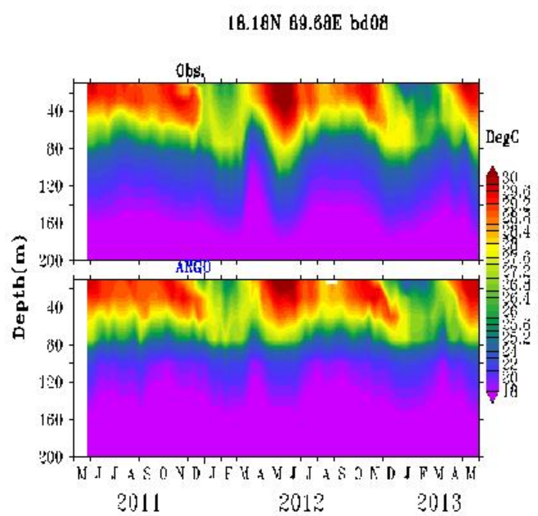


Figure 9(a) Same as Figure 8 but for salinity.



(b) Same as Figure 8(b).



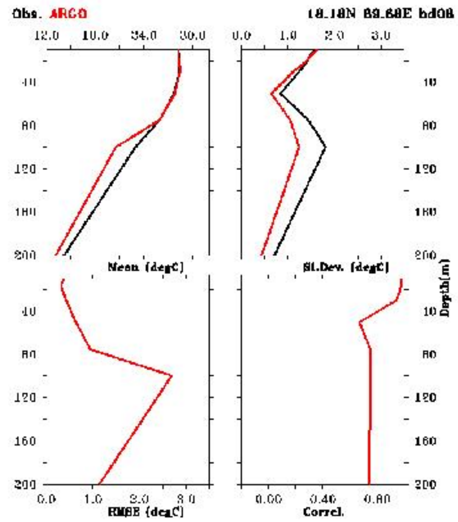
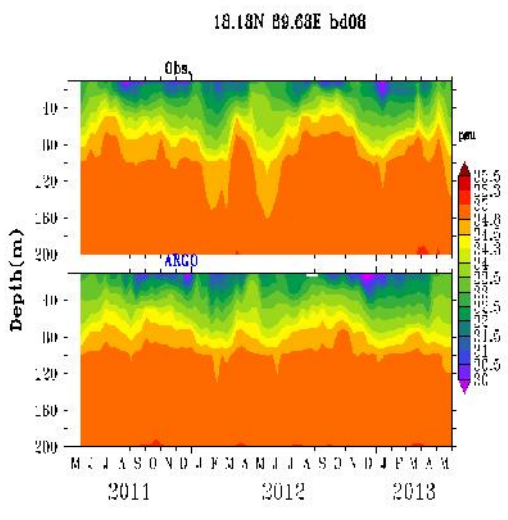


Figure 10(a) Same as Figure 8 at 18.18 N and 89.68 E.

(b) Same as Figure 8(b).



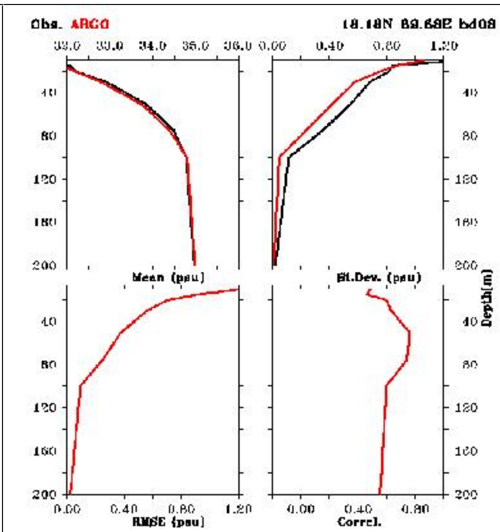
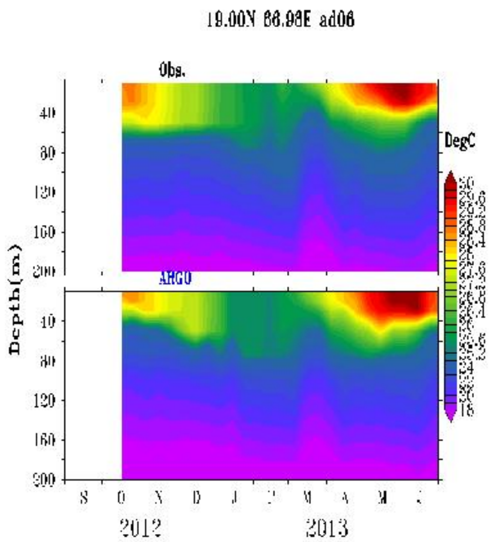


Figure 11(a) Same as Figure 10 but for salinity.

(b) Same as Figure 10(b).



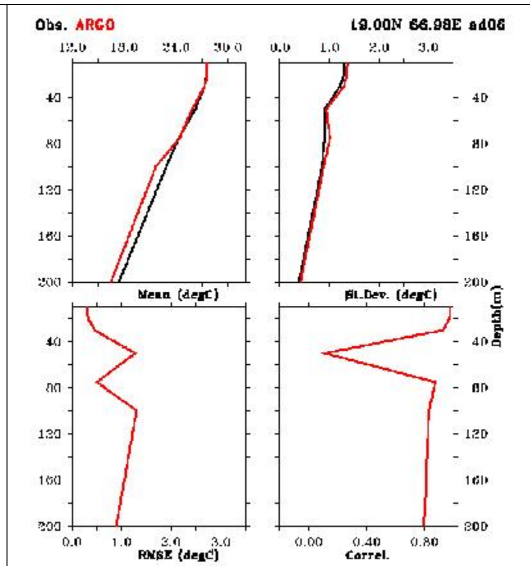
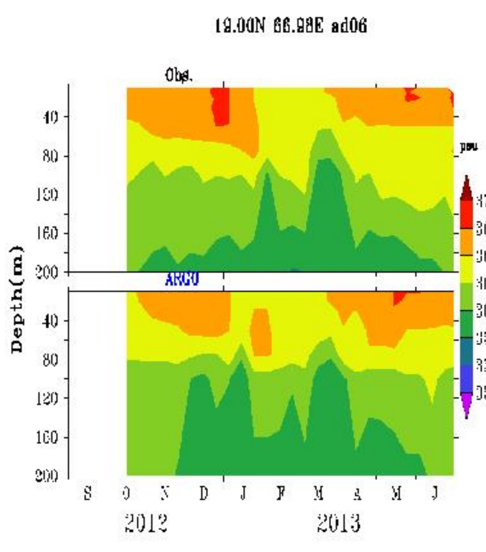


Figure 12(a) Same as Figure 8 at 19.0 N and 66.98 E (AD06).	**(b)** Same as Figure 8(b).



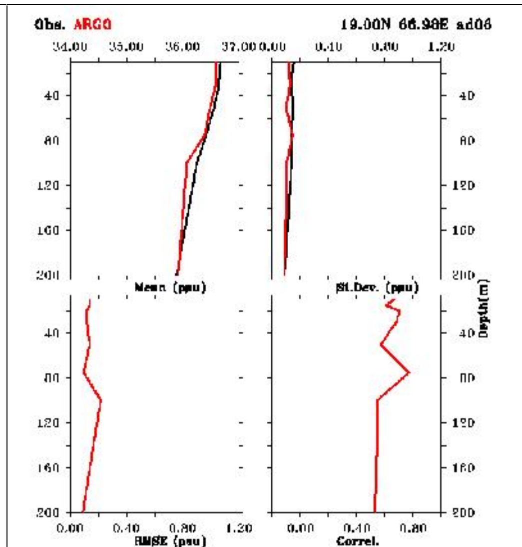


Figure 13(a) Same as Figure 12 but for salinity.	(b) Same as Figure 12(b).

6.2 Comparisons with RAMA Buoys

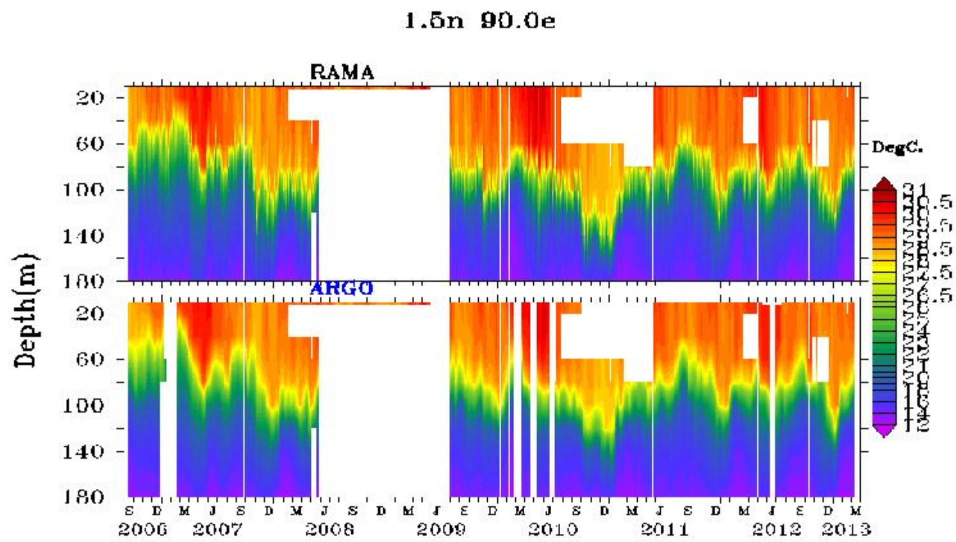


Figure 14(a) Section plots of temperature from RAMA buoy and Argo gridded product at 1.50 N and 90.0 E.

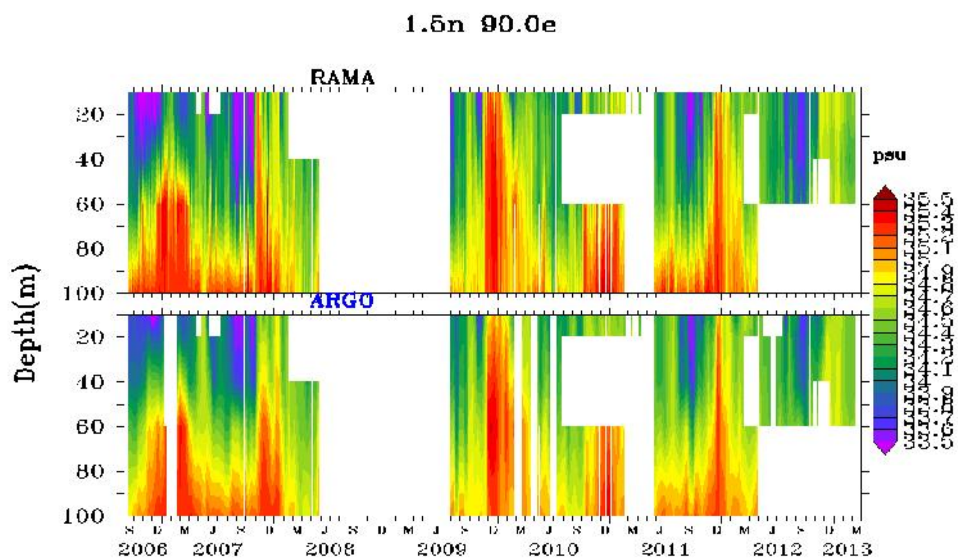


Figure 14(b) Section plots of salinity from RAMA buoy and Argo gridded product at 1.50 N and 90.0 E.

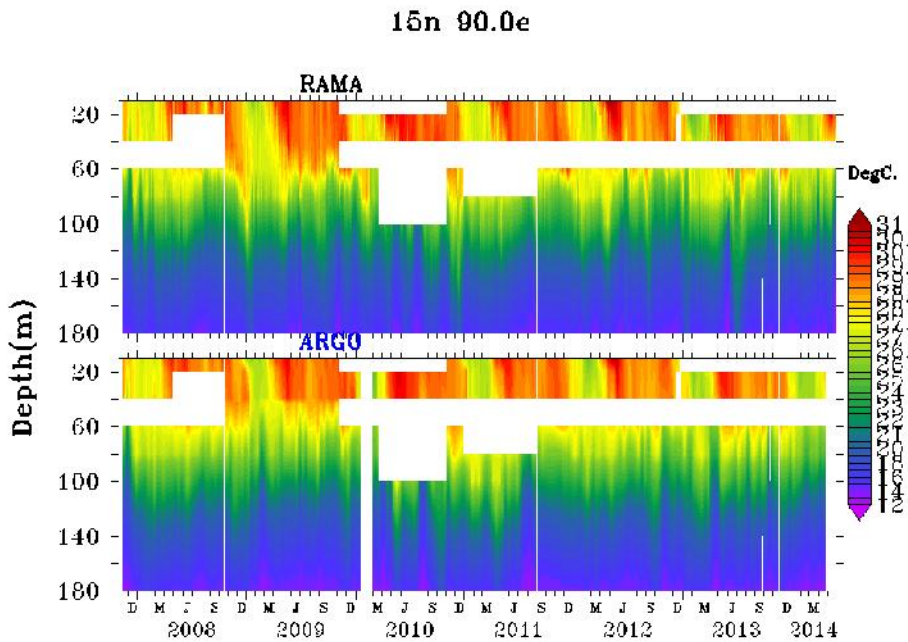


Figure 15(a) Section plots of temperature from RAMA buoy and Argo gridded product at 15.0 N and 90.0 E.

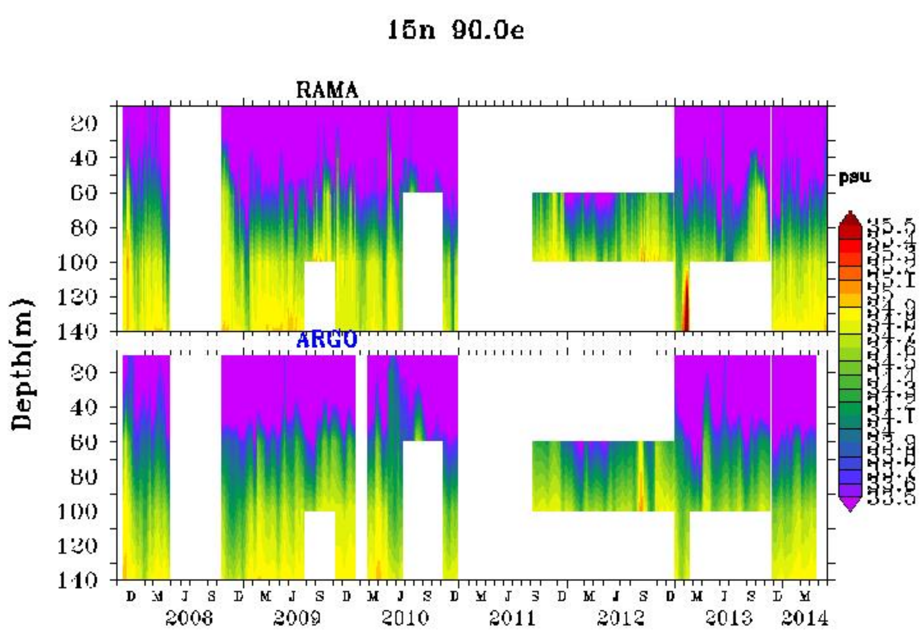


Figure 15(b) Section plots of salinity from RAMA buoy and Argo gridded product at 15.0 N and 90.0 E.

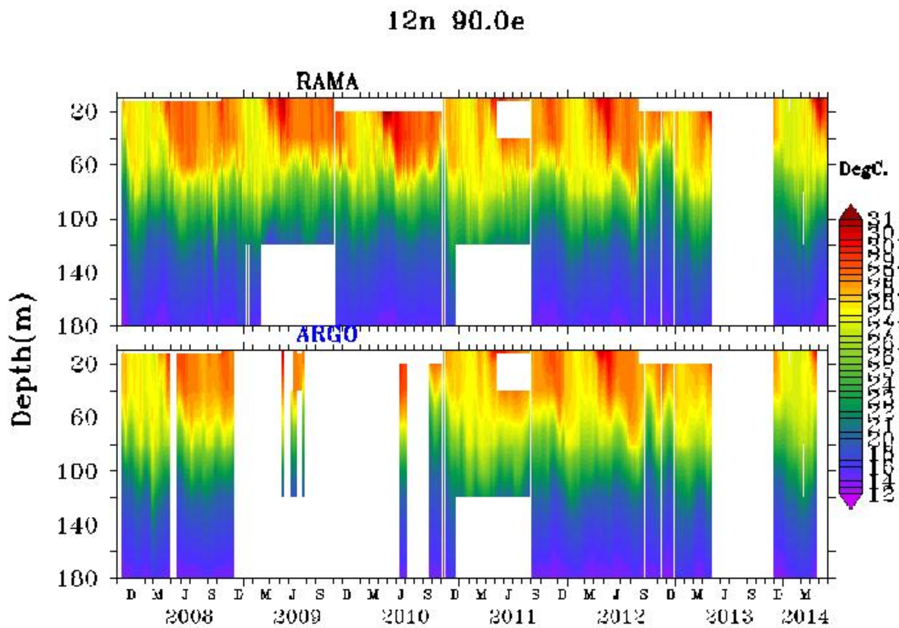


Figure 16(a) Section plots of temperature from RAMA buoy and Argo gridded product at 15.0 N and 90.0 E.

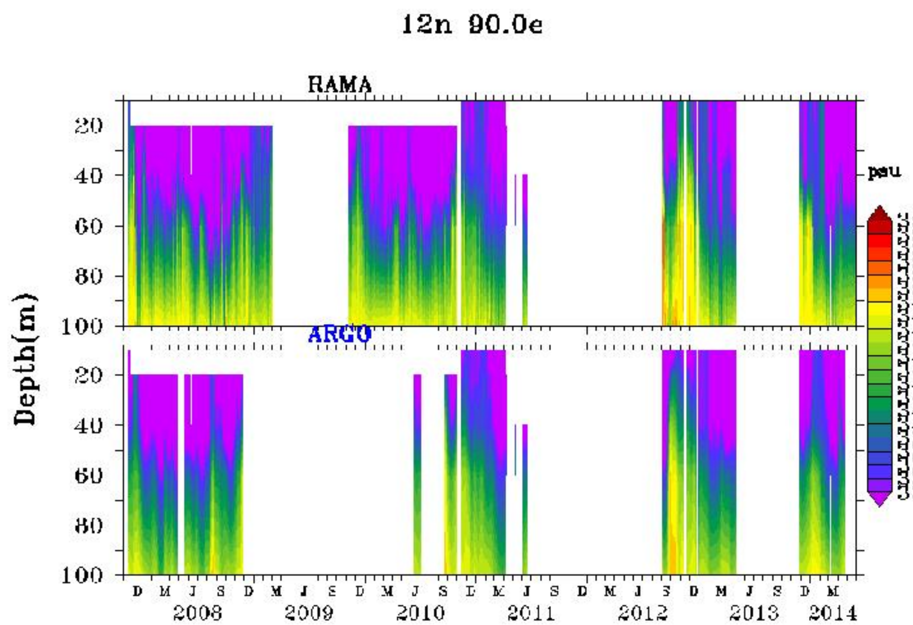


Figure 16(b) Section plots of salinity from RAMA buoy and Argo gridded product at 15.0 N and 90.0 E.

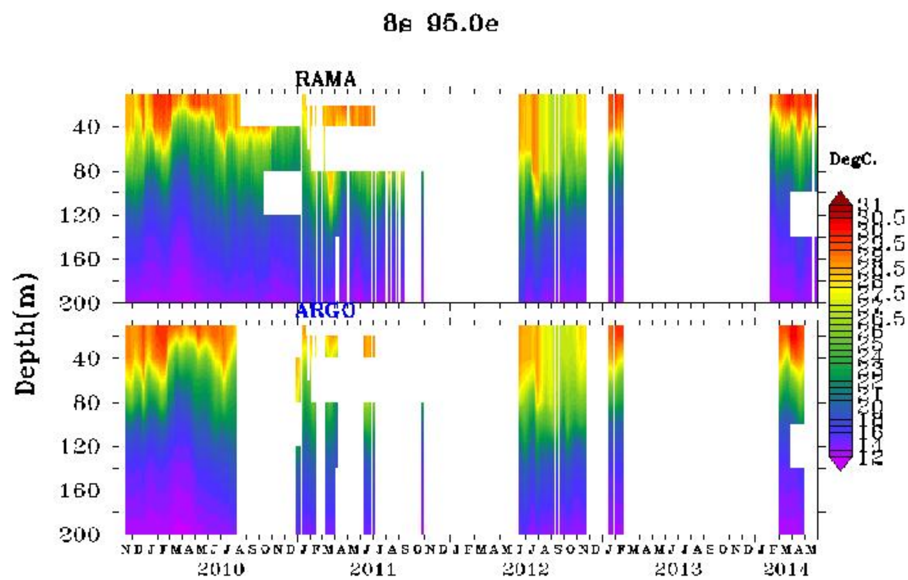


Figure 17(a) Section plots of temperature from RAMA buoy and Argo gridded product at 8.0 S and 95.0 E.

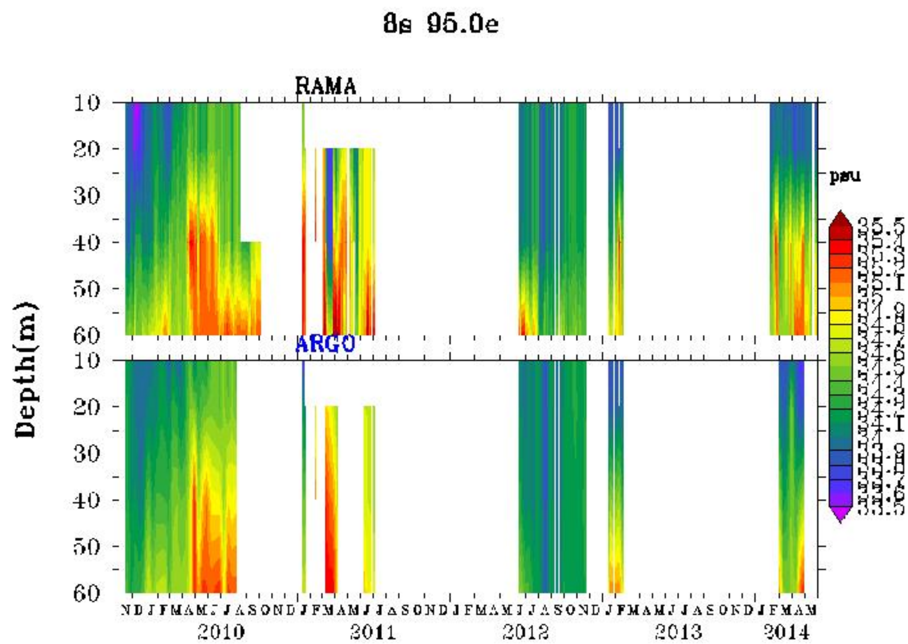


Figure 17(b) Section plots of salinity from RAMA buoy and Argo gridded product at 8.0 S and 95.0 E.

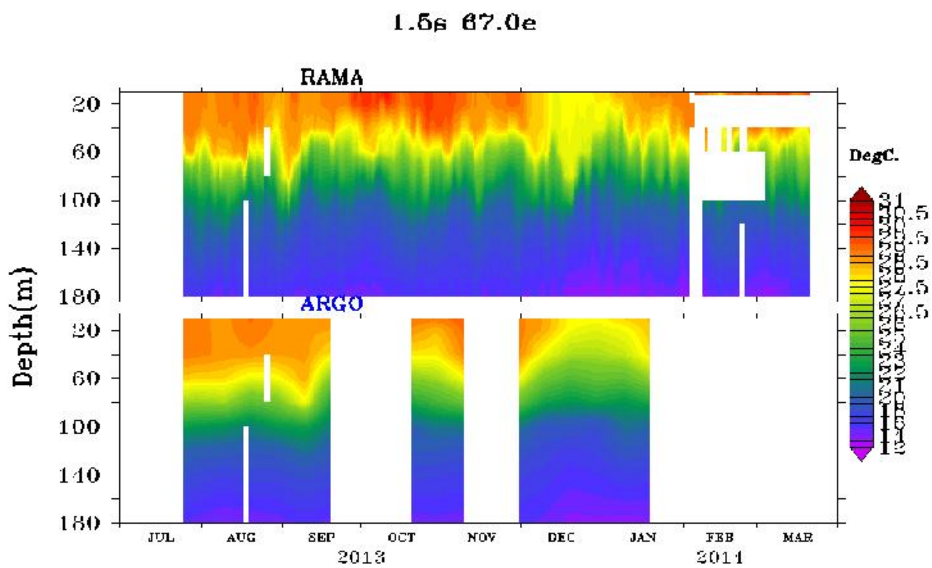


Figure 18(a) Section plots of temperature from RAMA buoy and Argo gridded product at 1.50 S and 67.0 E.

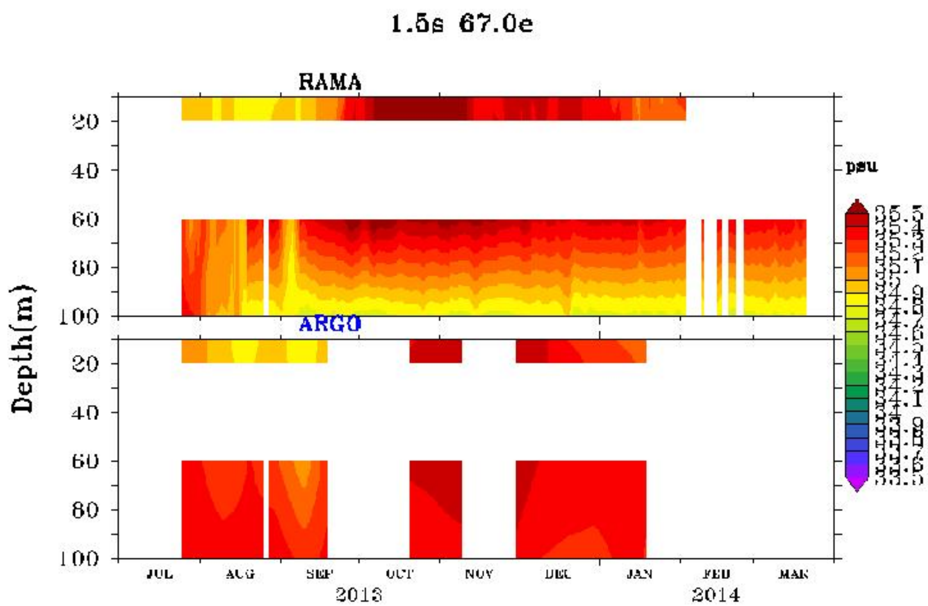


Figure 18(b) Section plots of salinity from RAMA buoy and Argo gridded product at 1.50 S and 67.0 E.

From the figures it can be observed that the gridded product of Argo profiles is highly correlated with the observations from OMNI and RAMA buoys picking most of the variations in temperature and salinity. However data gaps seems to be problem mainly with RAMA buoys and at few locations with Argo gridded product.

7. Summary and conclusions

1º Gridded products of temperature and salinity for the Indian ocean was produced on 10 day and monthly scale using two dimensional objective analysis methodology proposed by Kessler and McCreary (1993) from surface to 2000 m. For the first time 10 day and monthly gridded fields for the Indian Ocean are being generated operationally. The main data sets used for the gridded process are the Argo profiling floats which are deployed by various countries in the Indian Ocean. Data and OA method used for processing are described in detail in this document. Sample monthly Argo value added products (which are made available on the INCOIS web site are given in Appendix -1. Data in NetCDF format is made available on Live Access Server (LAS)

8. Acknowledgements

We express our gratitude to Director, Indian National Centre for Ocean Information Services (INCOIS) for his constant encouragement. We are grateful to William Kessler for clarifying some of the doubts while using the method. Argo data were collected and made freely available by the International Argo Project and the national programmes that contribute to it (http://www.argo.ucsd.edu, http://argo.jcommops.org). The authors wish to acknowledge use of the Ferret program, a product of NOAA's Pacific Marine Environmental Laboratory, for analysis and graphics in this paper.

9. References

- Angélica de León Arteaga, *Spatial Objective Analysis of Oceanographic Variables. An application to the Ebro Delta Shelf/Slope domain (North-Western Mediterranean Sea)*, PhD Thesis, July 2002.
- Bretherton, F.P., R.E. Davis, and C.B. Fandry, 1976: *A technique for Objective Analysis and Design of Oceanographic Experiments applied to MODE-73*, Deep Sea Research, Vol 23, pp. 559-582
- Carter, E.F. and A.R. Robinson, 1987: *Analysis Models for the Estimation of Oceanic Fields*, Jour. of Atmospheric and Oceanic Technology, Vol 4 (1), pp. 49-74.
- Frederick Carr, *Objective Analysis and Data Assimilation*, COMAP Numerical Weather Prediction Symposium, December 1999.
- Gandin, L.S., 1965: *Objective Analysis of Meteorological Fields*, Isreal Program for Scientific Translations, Jerusaleum, 242 pages
- Liebelt, P.B., 1967: *An Introduction to Optimal Estimation*, Addison-Wesley, Reading Mass., 273 pages
- Kessler and McCreary, 1993: *The Annual Wind-driven Rossby Wave in the Subthermocline Equatorial Pacific*, Journal of Physical Oceanography Vol 23, 1192 -1207.
- McPhaden, M.J., G. Meyers, K. Ando,Y. Masumoto, V.S.N. Murty, M. Ravichandran, F. Syamsudin, J. Vialard, L. Yu, and W. Yu, 2009: RAMA: The Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction. *Bull. Am. Meteorol. Soc., 90,* 459-480.
- Shesu RV, TVS Udaya Bhaskar, E Pattabhi Rama Rao, R Devender, T Hemasundar Rao (2013), Open Source Architecture for Web-Based Oceanographic Data Services, Data Science Journal, 12:47 - 55.

- Wong, A., R. Keeley, T. Carval and the Argo Data Management Team. 2004: *Argo quality control manual*, ver. 2.0b, Report, 23 pp.
- Udaya Bhaskar, T.V.S., E. Pattabhi Rama Rao., R. Venkat Shesu., R. Devender. 2012: *A note of three way quality control of Argo temperature and salinity profiles - A semi-automated approach at INCOIS*, International Journal of Earth Sciences and Engineering, Vol 5 (6), pp 1510 - 1514.
- Venkatesan R, Shamji VR, Latha G, Mathew S, Rao RR, Muthiah A, Atmanand MA (2013) In situ ocean subsurface time-series measurements from OMNI buoy network in the Bay of Bengal. Curr Sci 104:1166-1177.

9. Contacts

Comments, questions regarding the Argo gridded dataset and can be directed to any of the authors.

TVS Udaya Bhaskar, M Ravichandran, R Devender

Indian National Centre for Ocean Information Service (INCOIS),

"Ocean Valley", Pragathinagar (BO), Nizampet (SO),

Hyderabad - 500 090

Andhra Pradesh, India.

The authors can also be contacted by phone or e-mail at:

TVS Udaya Bhaskar E-mail: uday@incois.gov.in Tel.: 91 (40) 23886048

M Ravichandran E-mail: ravi@incois.gov.in Tel.: 91 (40) 23886004

R Devender E-mail: devender@incois.gov.in Tel.: 91 (40) 23886099

Link to the web-server on which the gridded product is served at INCOIS is:

http://las.incois.goiv.in/

10. Copyrights

The Argo gridded dataset generated at INCOIS is open for free unrestricted use, copying and distribution specifically for research use. The dataset is a research quality product. Efforts are put to check and delete any spurious data sets. Errors (if any) reported to the authors by users will be corrected and data set will be updated.

Authors request that data when used in publications should be acknowledged in the following form:

"*This study used the data produced at INCOIS [Udaya Bhaskar et al., 2007] Argo gridded dataset of temperature and salinity from Argo profiling float.*"

Reference of this technical paper:

Udaya Bhaskar, T. V. S., M. Ravichandran, and R. Devender (2007), An operational objective analysis system at INCOIS for generation of Argo value added products, Technical Report INCOIS-MOG-ARGO-TR-04-2007, Indian National Centre for Ocean Information Services, Hyderabad, India.

11. Disclaimer

Any use of or dependence by the User on the gridded product or any subset thereof is at the User's own risk and INCOIS shall not be liable for any loss or damage howsoever arising as a result of such use in analysis. The User agrees to assure and hold INCOIS harmless as a result of User's use of or dependence on the gridded data product.

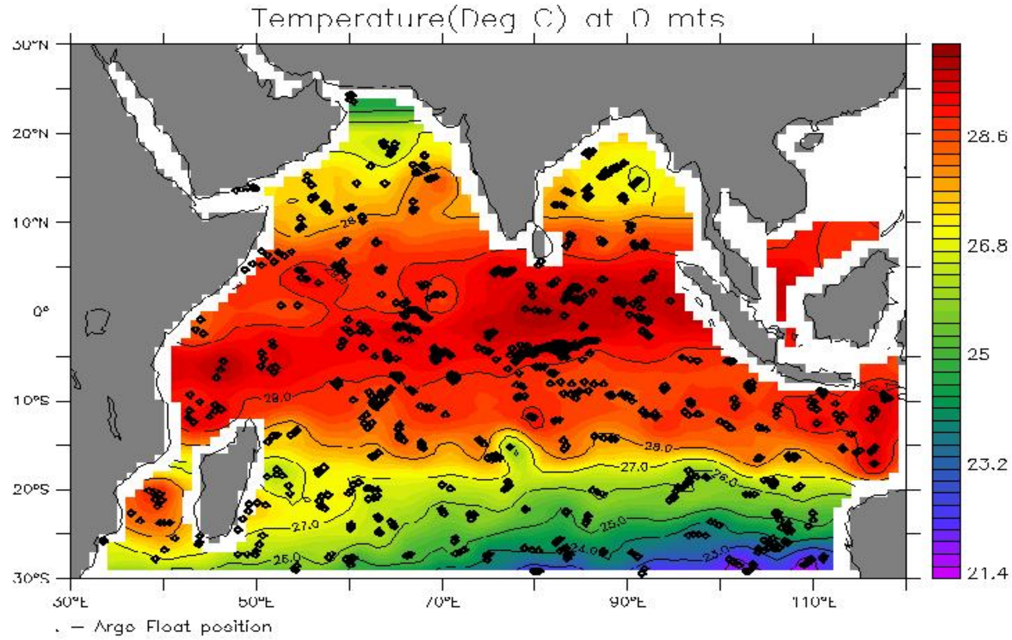
Appendix - 1

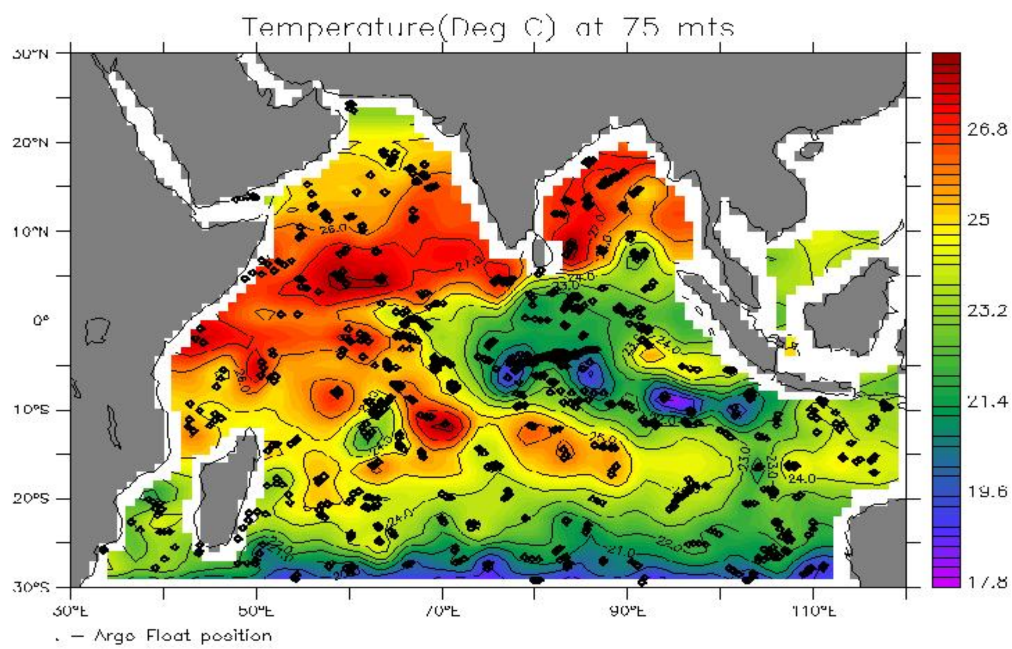
Monthly Argo Value Added Products

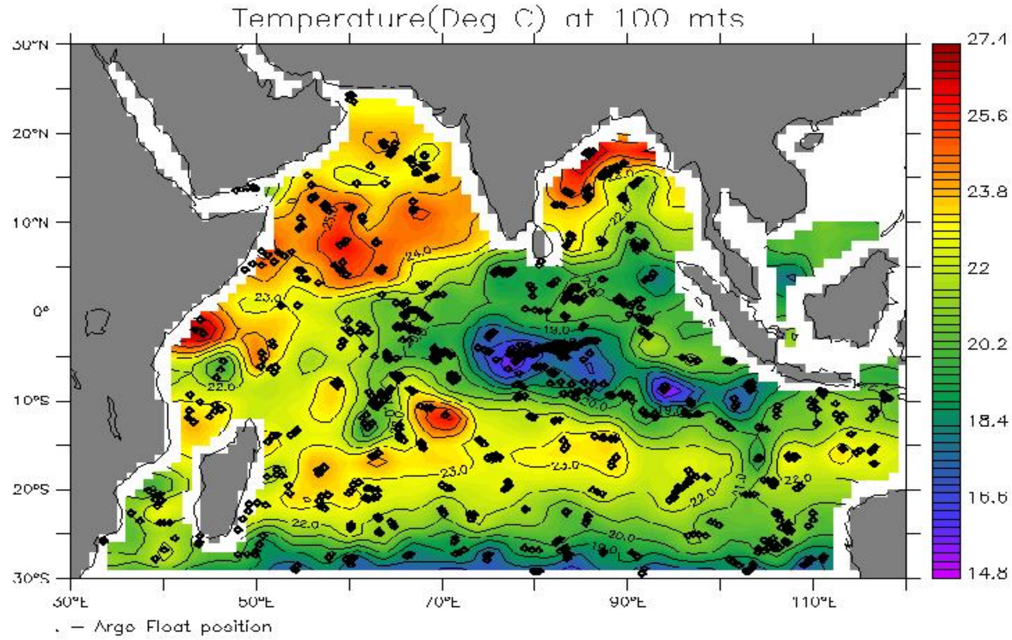
From the monthly gridded temperature and salinity data, maps of temperature and salinity at 0, 75 m, 100m, 200m, 500m and 1000m are displayed. Further this gridded data product is used for generating value added product like:

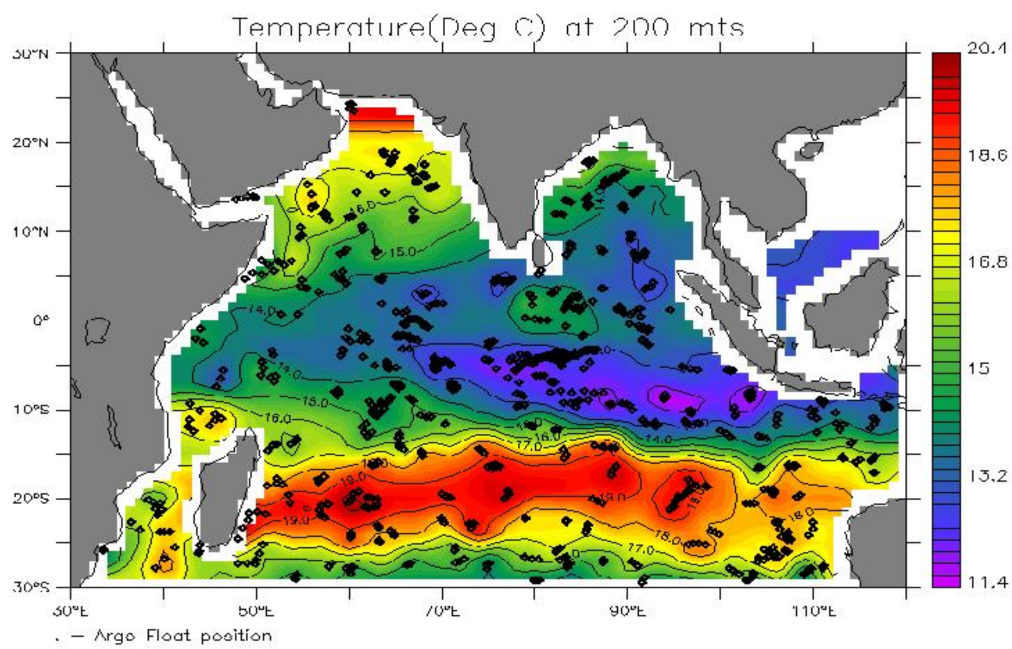
- Geostrophic currents,
- Mixed Layer Depth (MLD) ,
- Isothermal Layer Depth (ILD),
- Heat Content (integrated up to 300m),
- Depth of 26º isotherms,
- Depth of 20º isotherms,
- Dynamic Height.

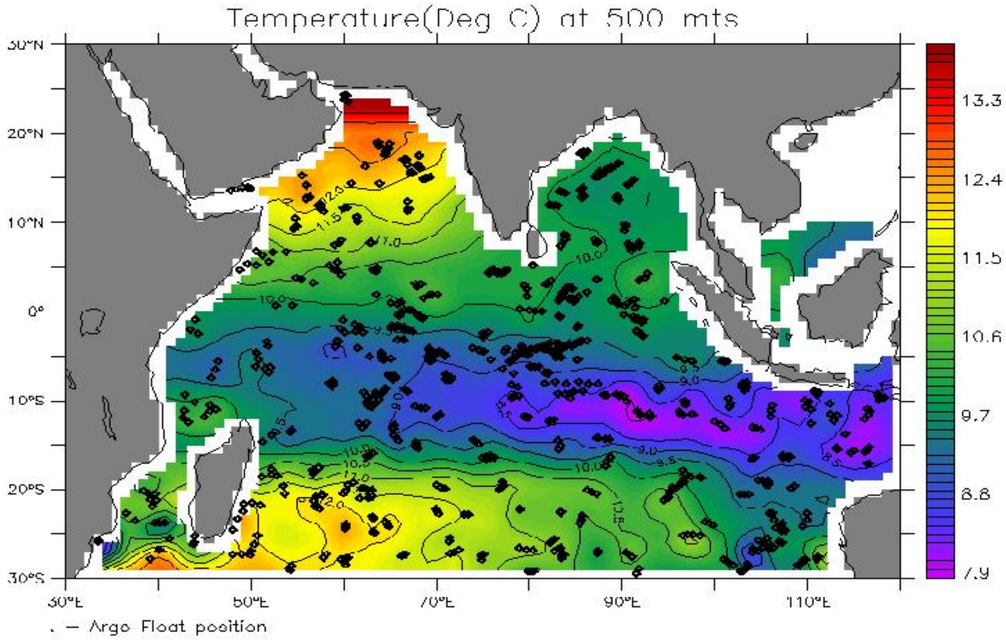
Further this data is converted to Network Common Data Format (NetCDF) and made on the INCOIS Live Access Server (ILAS, las.incois.gov.in). Maps of Argo value added products are also published on INCOIS web site (http://www.incois.gov.in/Incois/argo/products/argo_frames.html). Some sample maps of Argo derived products are shown below:

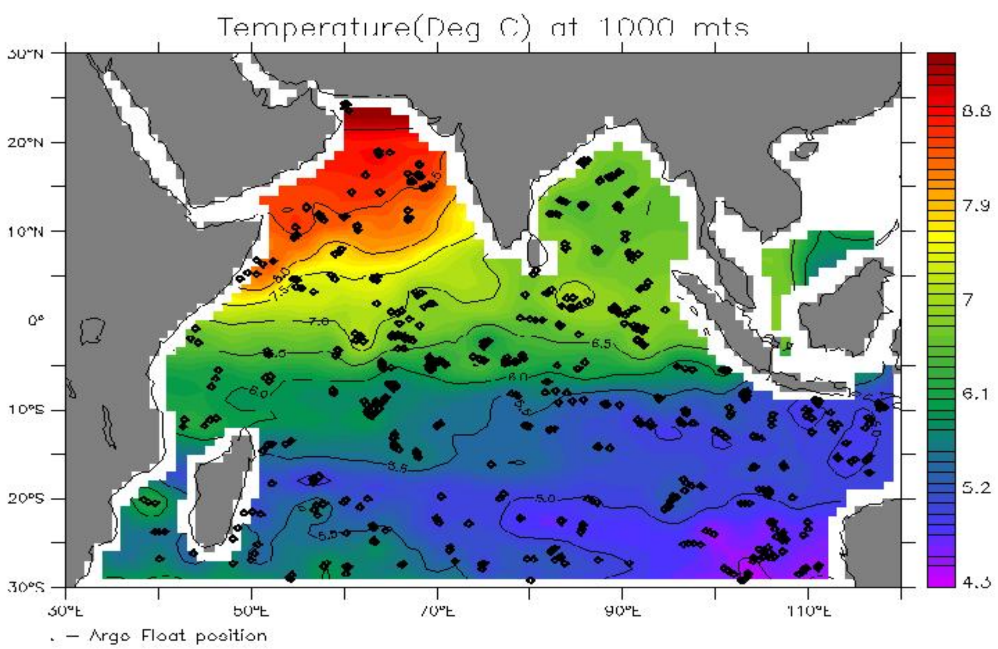

Temperature(Deg C) at 0 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
28.6
26.8
25
23.2
21.4
. — Argo Float position

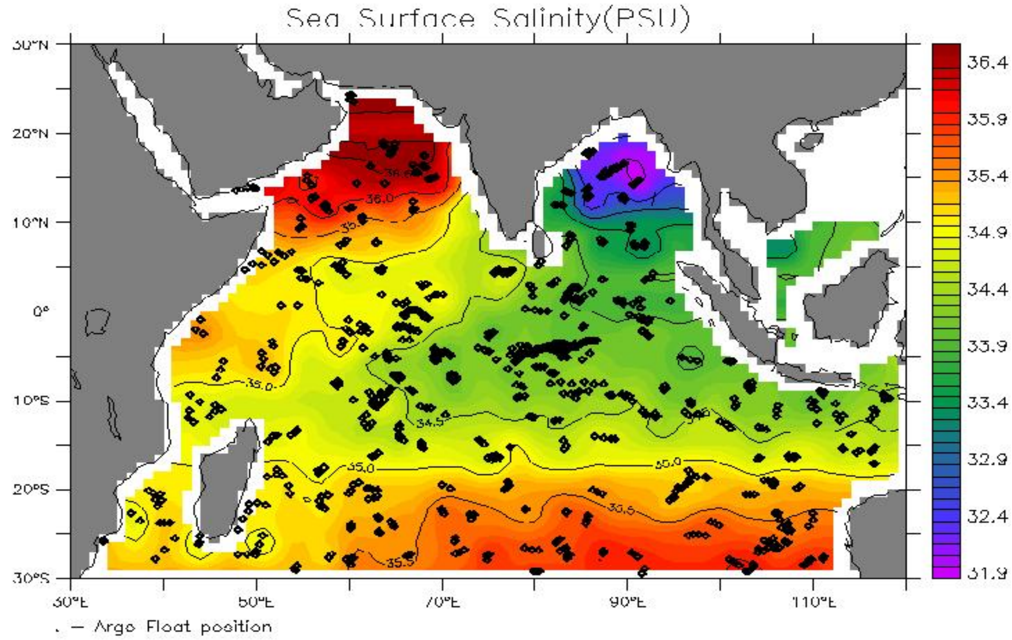

Temperature(Deg C) at 75 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
26.8
25
23.2
21.4
19.6
17.8
. — Argo Float position

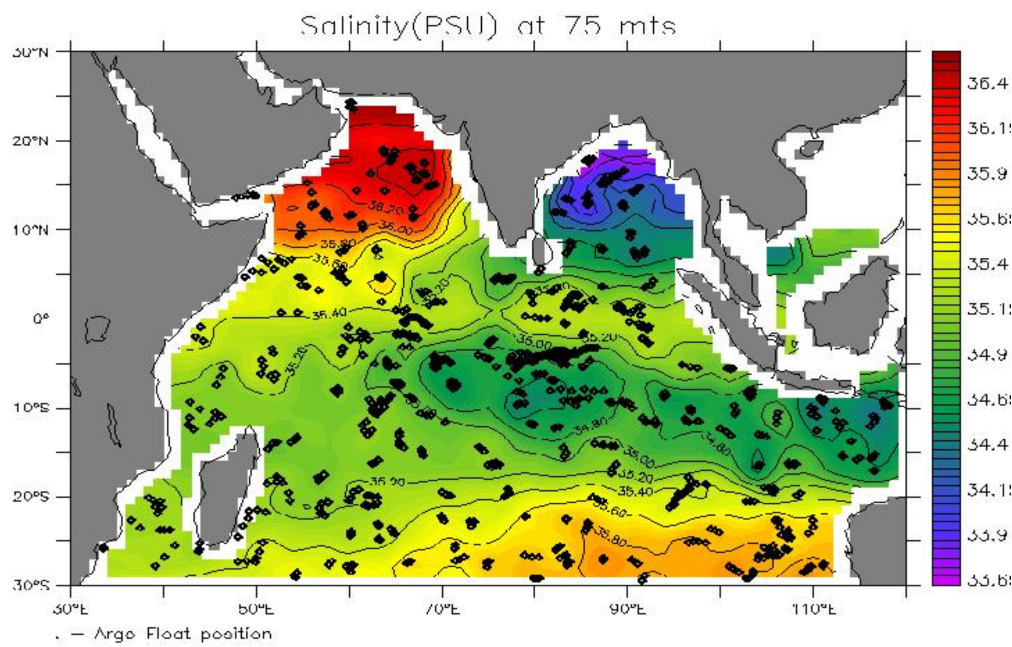

Temperature(Deg C) at 100 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. – Argo Float position
27.4
25.6
23.8
22
20.2
18.4
16.6
14.8

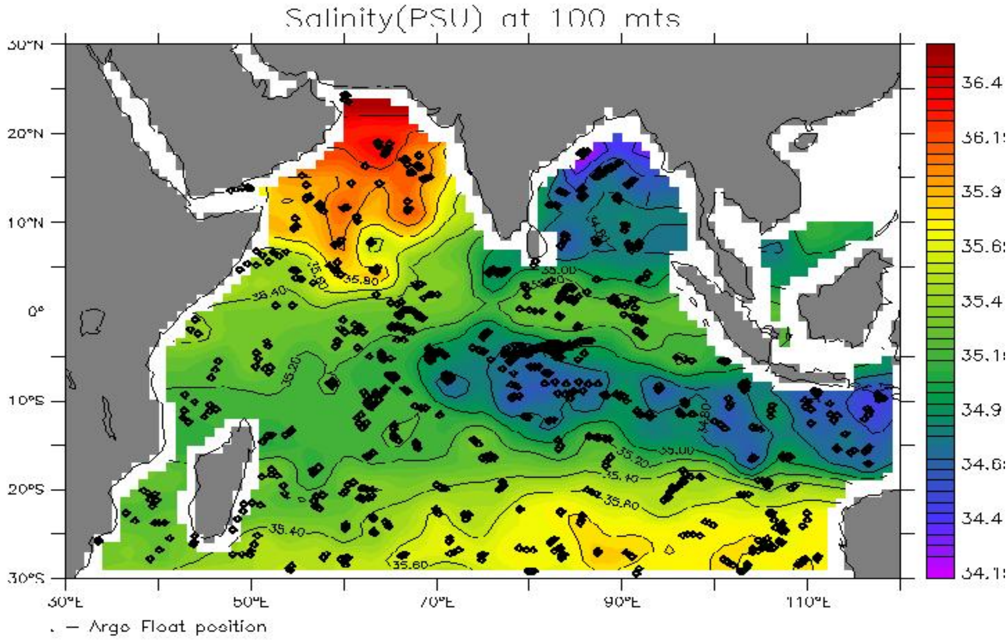

Temperature(Deg C) at 200 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. – Argo Float position
20.4
18.6
16.8
15
13.2
11.4

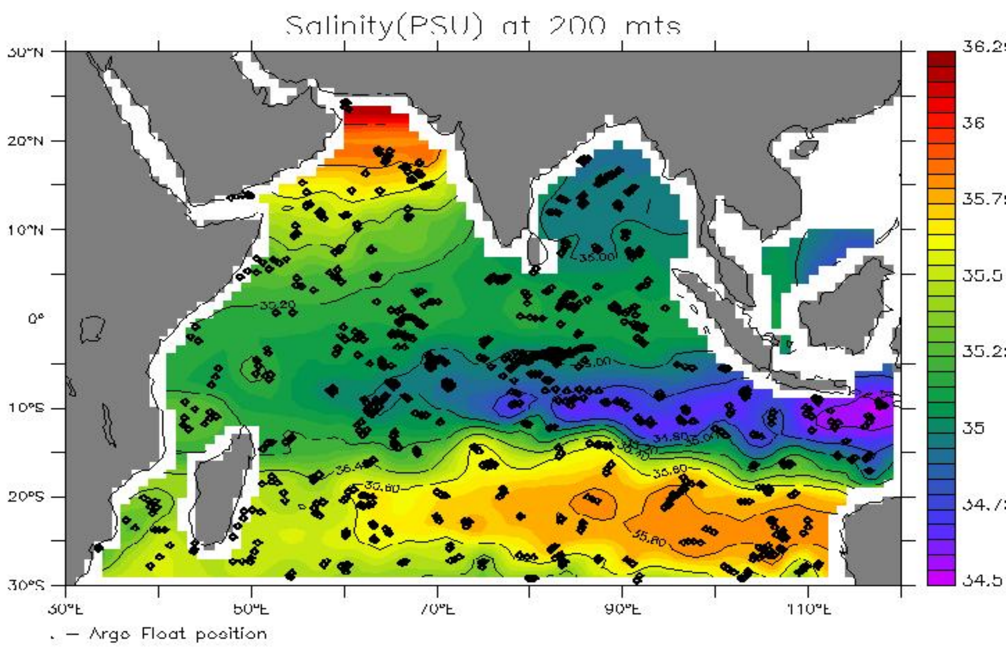

Temperature(Deg C) at 500 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
13.3
12.4
11.5
10.6
9.7
8.8
7.9
. — Argo Float position

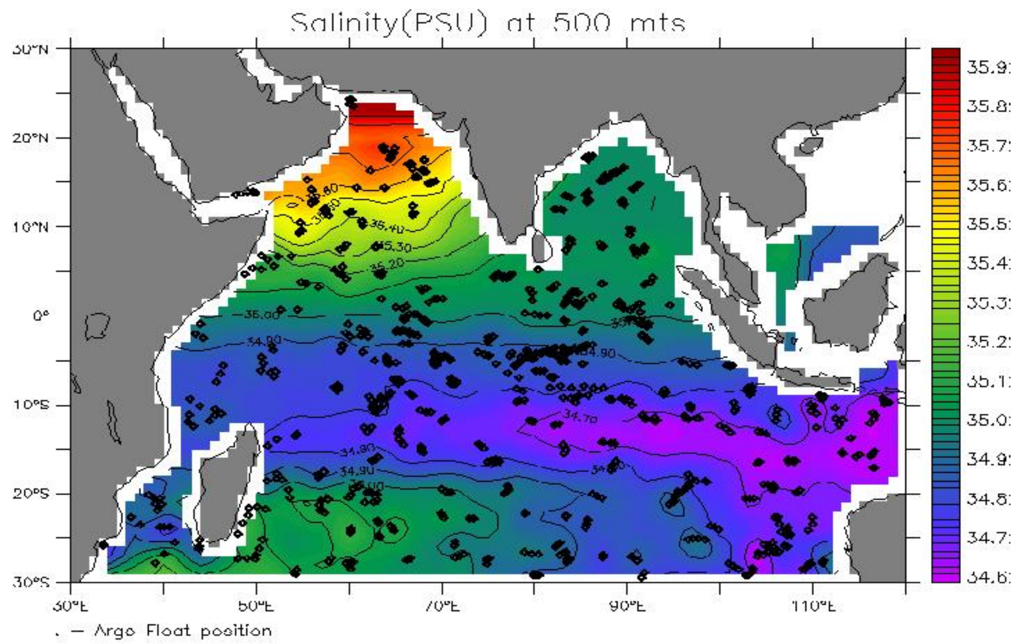

Temperature(Deg C) at 1000 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
8.8
7.9
7
6.1
5.2
4.3
. — Argo Float position

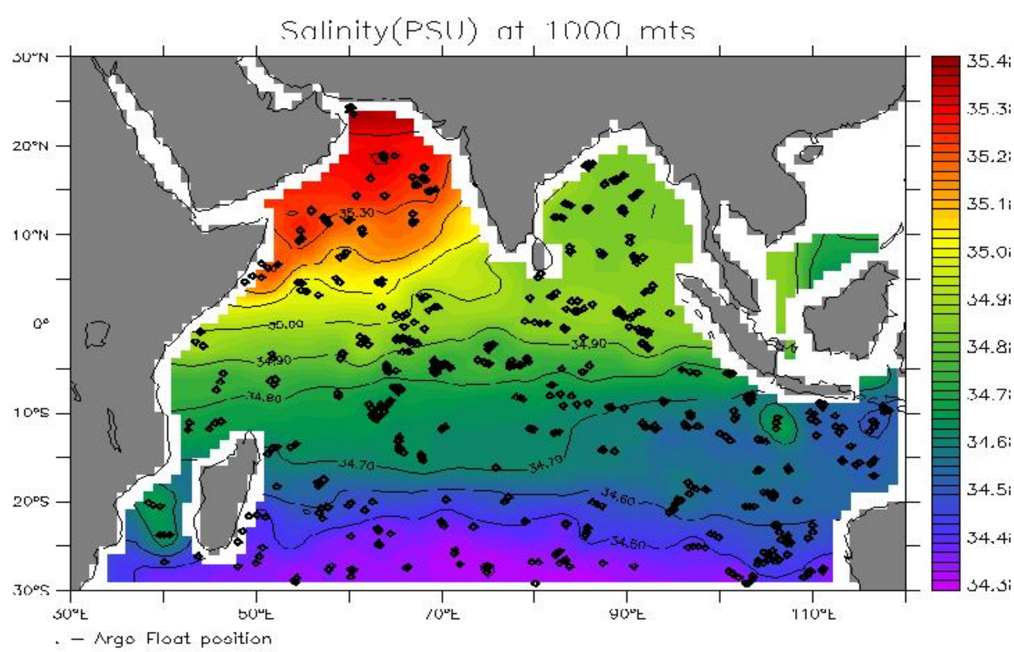

Sea Surface Salinity(PSU)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
36.4
35.9
35.4
34.9
34.4
33.9
33.4
32.9
32.4
31.9
35.0
34.5
35.5
. — Argo Float position

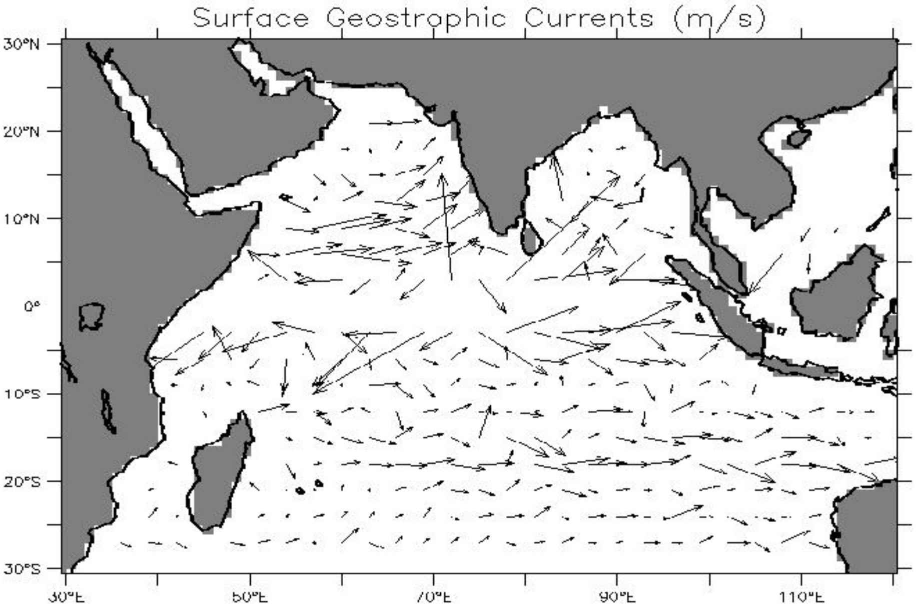

Salinity(PSU) at 75 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
36.4
35.9
35.4
34.9
34.4
33.9
35.40
35.20
35.00
34.80
35.60
35.80
. — Argo Float position

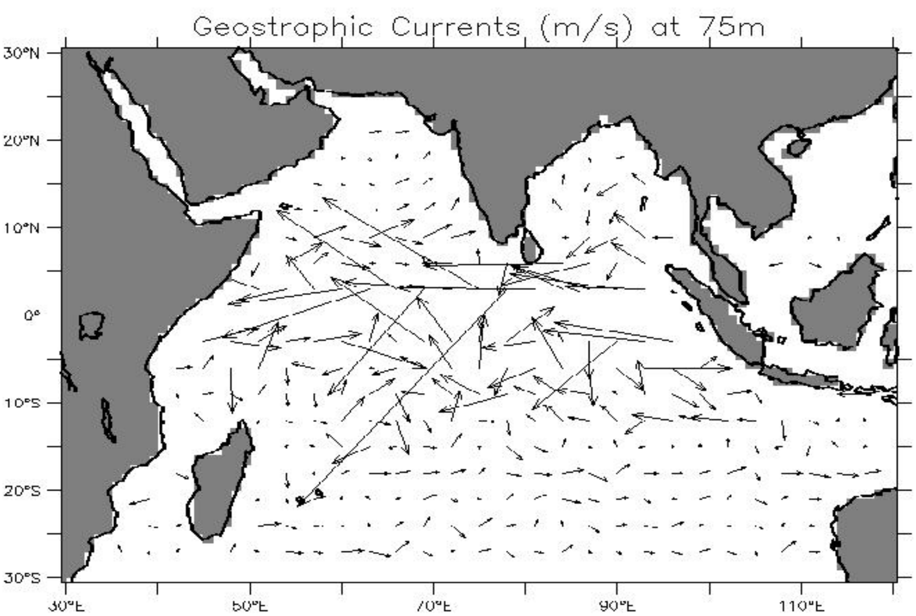

Salinity(PSU) at 100 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
36.4
35.9
35.4
34.9
34.4
. — Argo Float position

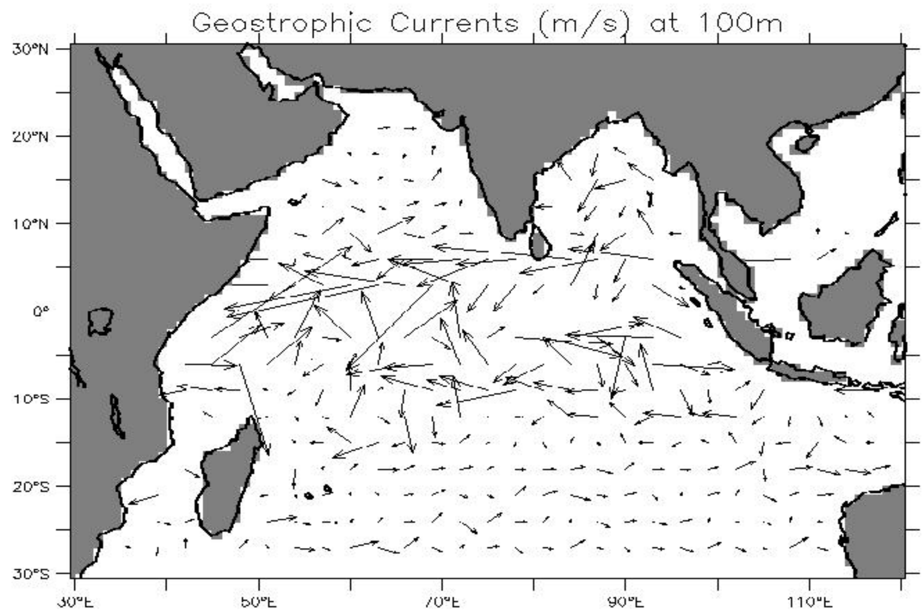

Salinity(PSU) at 200 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
36
35.5
35
34.5
. — Argo Float position

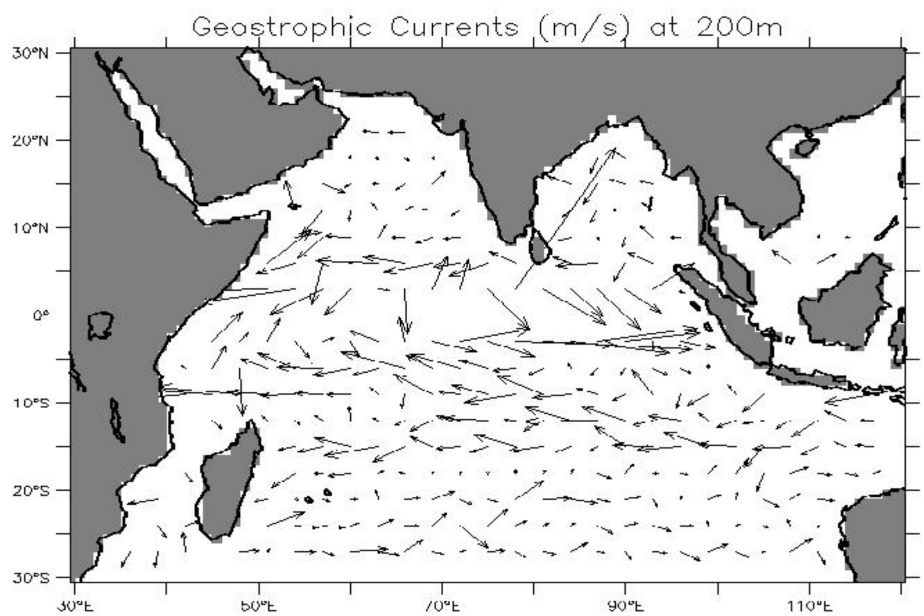

Salinity(PSU) at 500 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. — Argo Float position

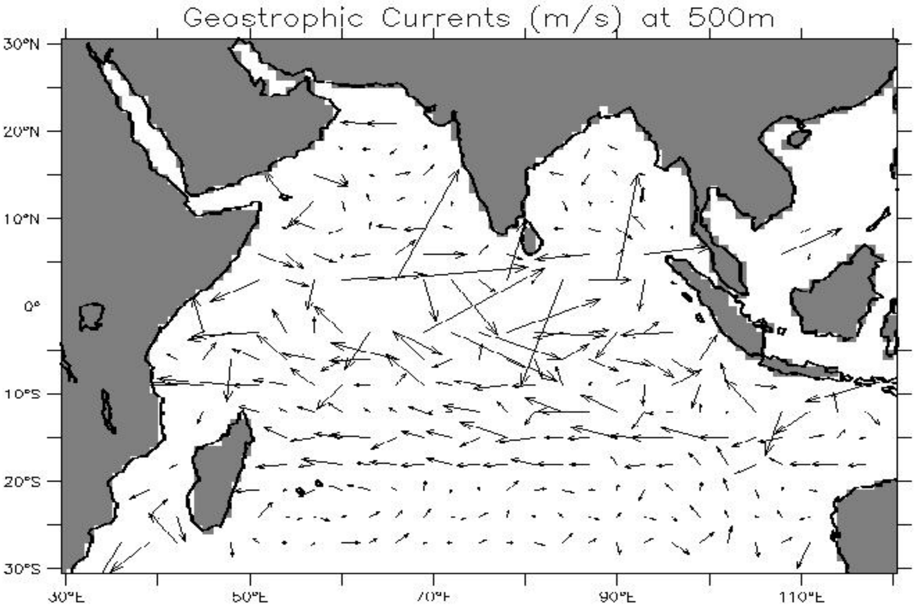

Salinity(PSU) at 1000 mts
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. — Argo Float position

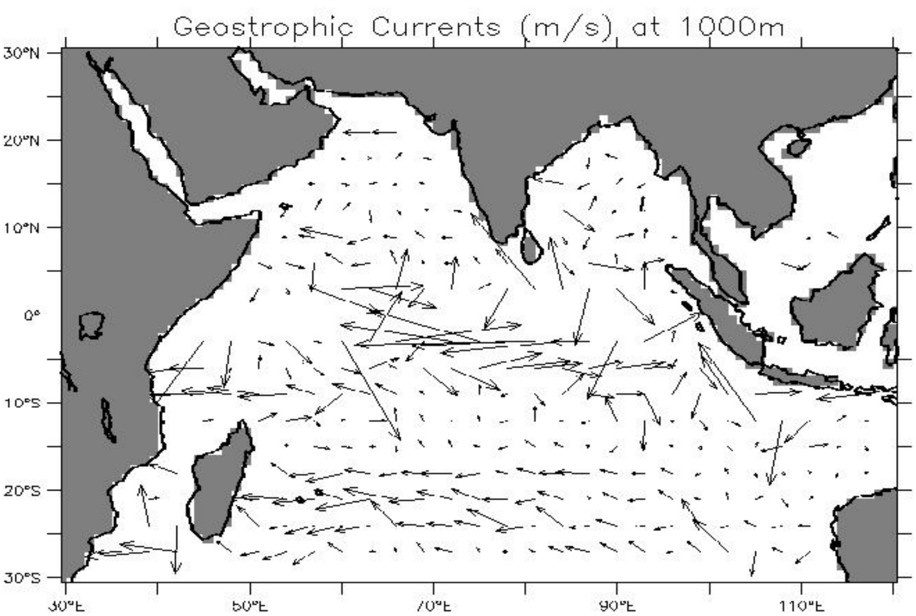

Surface Geostrophic Currents (m/s)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E

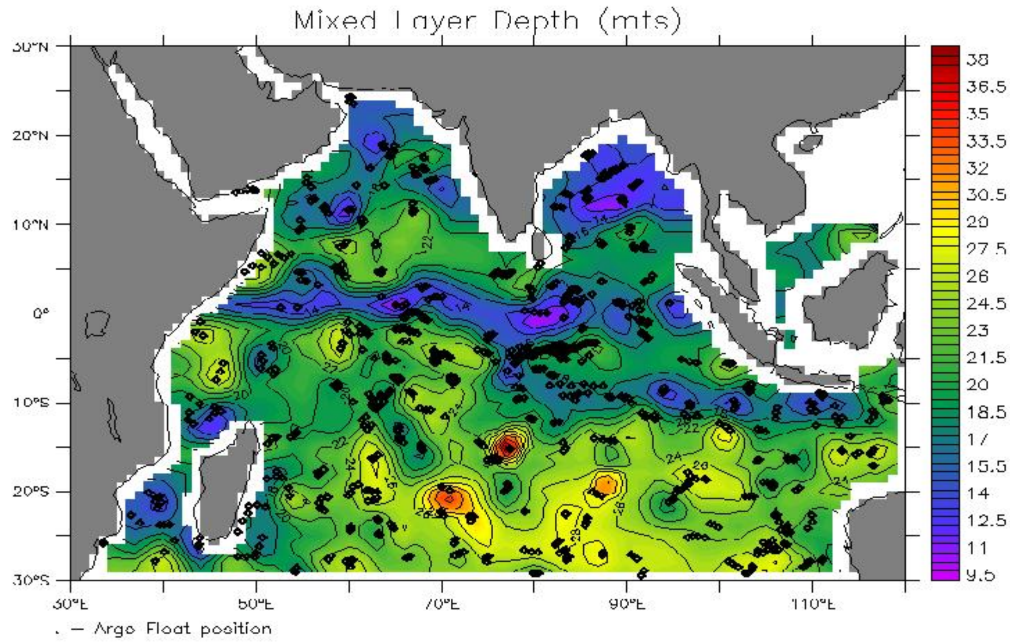

Geostrophic Currents (m/s) at 75m
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E

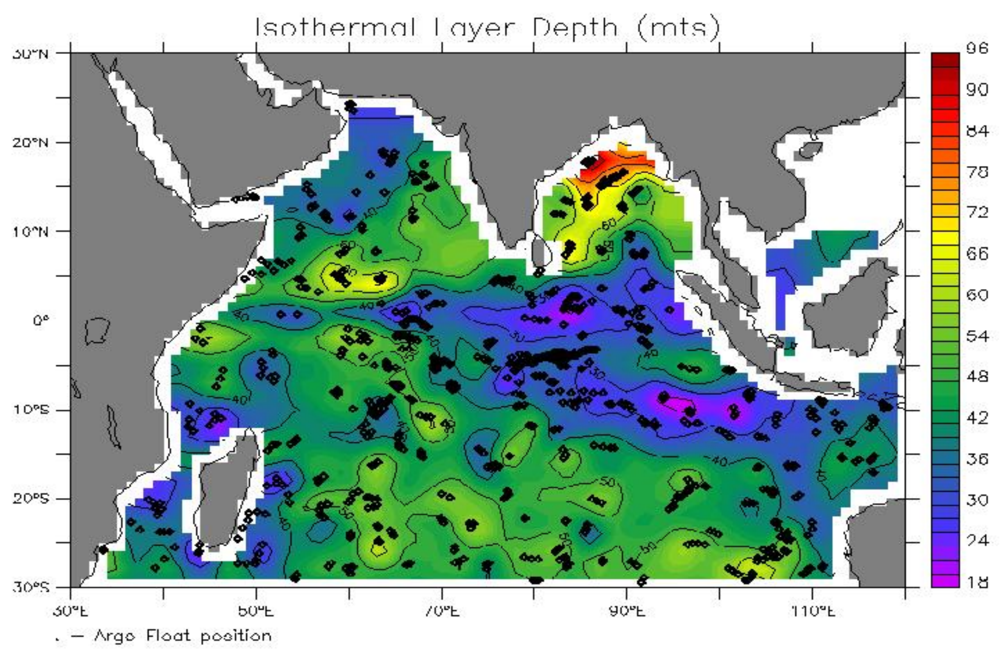

Geostrophic Currents (m/s) at 100m
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E

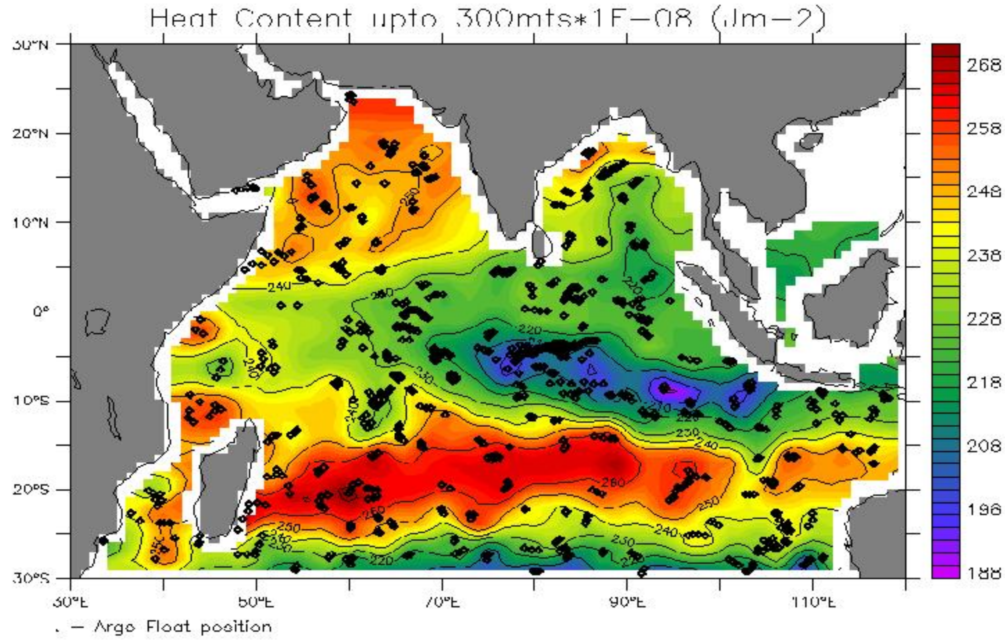

Geostrophic Currents (m/s) at 200m
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E

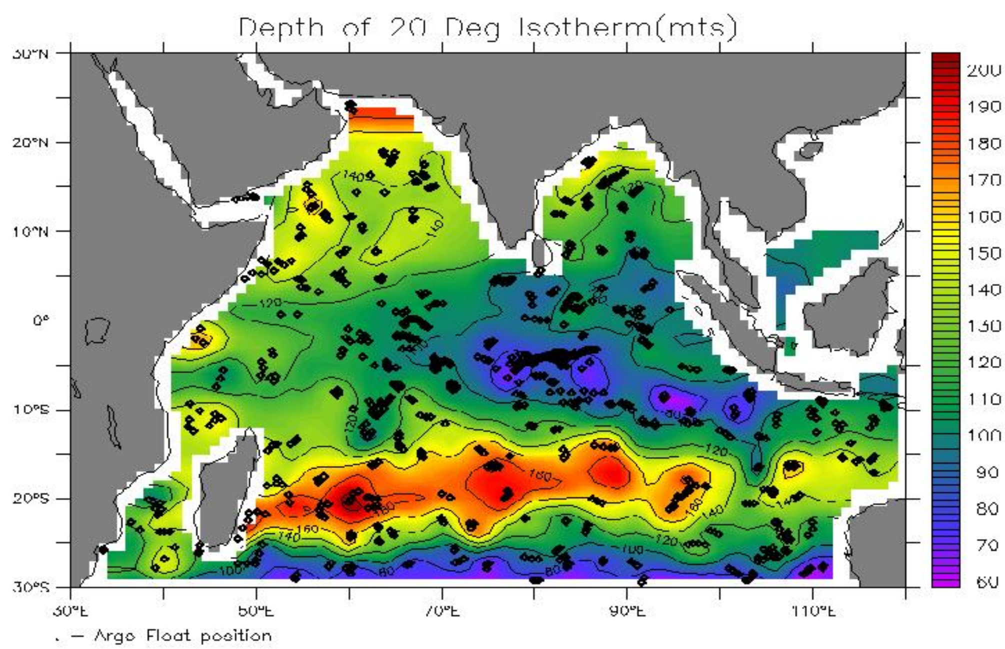

Geostrophic Currents (m/s) at 500m
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E

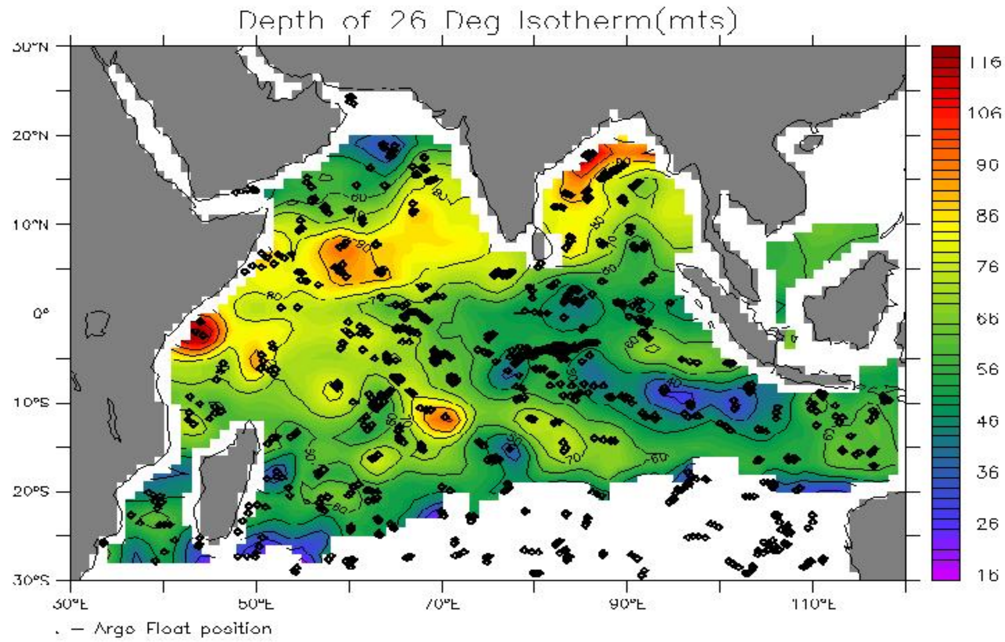

Geostrophic Currents (m/s) at 1000m
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E


Mixed Layer Depth (mts)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. — Argo Float position
38
36.5
35
33.5
32
30.5
29
27.5
26
24.5
23
21.5
20
18.5
17
15.5
14
12.5
11
9.5

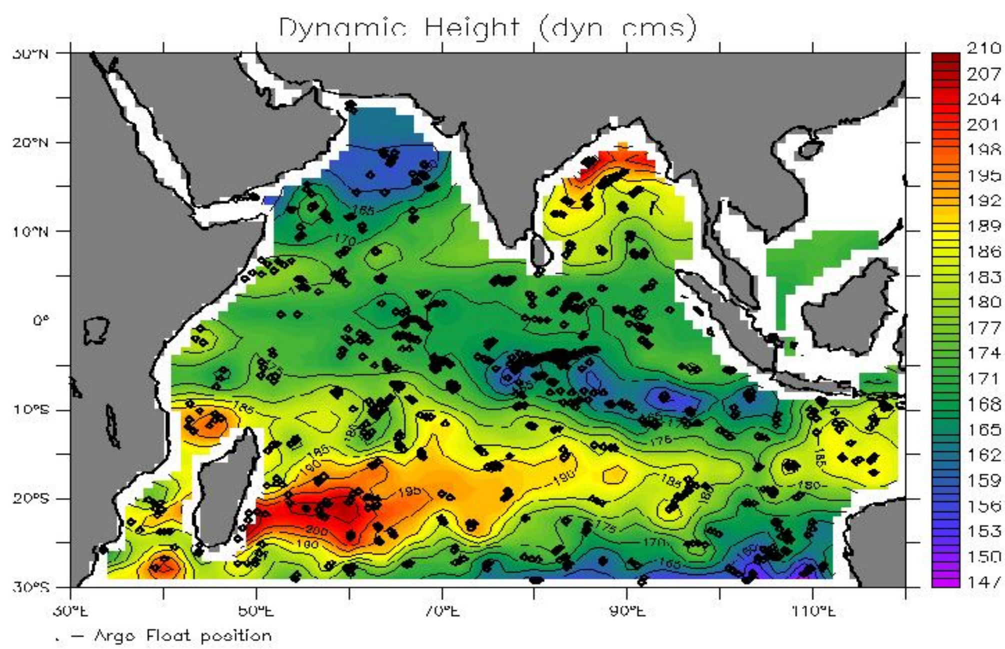

Isothermal Layer Depth (mts)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. — Argo Float position
96
90
84
78
72
66
60
54
48
42
36
30
24
18


Heat Content upto 300mts*1E-08 (Jm-2)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
268
258
248
238
228
218
208
196
188
. – Argo Float position


Depth of 20 Deg Isotherm(mts)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
. – Argo Float position



Depth of 26 Deg Isotherm(mts)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. — Argo Float position
116
106
96
86
76
66
56
46
36
26
16



Dynamic Height (dyn cms)
30°N
20°N
10°N
0°
10°S
20°S
30°S
30°E
50°E
70°E
90°E
110°E
. — Argo Float position
210
207
204
201
198
195
192
189
186
183
180
177
174
171
168
165
162
159
156
153
150
147

Appendix - 2

Derivation of Objective Analysis Equation

(source: *www.taygeta.com/objan/oa_derive.html*)

In this section, we provide a derivation of the **Gauss-Markov** theorem which is the foundation for the method of Objective Analysis which is used in Oceanography and Meteorology for the estimation of fields based upon incomplete and noisy observations.

The derivation

We want to derive an optimal estimate, $\hat{x}$, of a field, *x*, as a linear combination of observation data, θ,

$$\hat{x} = A\theta \tag{1}$$

Our goal is to derive the form of the matrix, *A*, so that the expected mean square difference between the estimated field and the actual field (*x*) is minimized,

$$E[\varepsilon\varepsilon^T] = E[(\hat{x} - x)(\hat{x} - x)^T] = \text{minimum} \tag{2}$$

If we put (1) into (2) and expand, we get

$$E[\varepsilon\varepsilon^T] = E[A\theta\theta^T A^T - x\theta^T A^T - A\theta x^T + xx^T] \tag{3}$$

If we let C_x be the autocorrelation of the field ($E[x\, x^T]$), C_θ be the autocorrelation of the observations ($E[\theta\theta^T]$), and $C_{x\theta}$ be the cross correlation between the field and the observations ($E[x\theta^T]$), then we can write the above as

$$C_\varepsilon = AC_\theta A^T - C_{x\theta}A^T - AC_{x\theta}^T + C_x \tag{4}$$

The next step requires the application of the following matrix identity (proved in the Appendix-3),

$$(A - B\, C^{-1})\, C\, (A - B\, C^{-1})^T - B\, C^{-1}\, B^T = A\, C\, A^T - B\, A^T - (B\, A^T)^T \tag{5}$$

using A in (4) for A in (5), and $C_{x\theta}$ for B as well as C_θ for C, we can reduce (x) to

$$C_\varepsilon = (A - C_{x\theta}C_\theta^{-1})C_\theta(A - C_{x\theta}C_\theta^{-1})^T - C_{x\theta}C_\theta^{-1}C_{x\theta}^T + C_x \tag{6}$$

(note we have also used the fact that $C_\theta = C_\theta^T$).

The matrices C_θ , is an autocorrelation matrix therefore both it and C_θ^{-1} are nonnegative definite (see Appendix-4), therefore

$$(A - C_{x\theta}C_\theta^{-1})C_\theta(A - C_{x\theta}C_\theta^{-1})^T \tag{7}$$

and

$$C_{x\theta}C_\theta^{-1}C_{x\theta}^T \tag{8}$$

are both matrices with positive diagonal elements. This means that the diagonal elements of C_ε are therefore minimized when it is true that,

$$A - C_{x\theta}C_\theta^{-1} = 0 \tag{9}$$

Therefore we have,

$$A = C_{x\theta}C_\theta^{-1} \tag{10}$$

This is the estimator that we are seeking.

$$\hat{x} = C_{x\theta}C_\theta^{-1}\theta \tag{11}$$

Further, we can write down what the expected error for the estimator as,

$$C_\varepsilon = C_x - C_{x\theta}C_\theta^{-1}C_{x\theta}^T \tag{12}$$

Equations (11) and (12) constitute the **Gauss-Markov** estimator for the linear minimum means square estimate of a random variable.

Linear Observations

Upon reflection of equations (11) and (12) a problem arises: the determination of C_θ and $C_{x\vartheta}$ require having the true field values, *x*, but all one can actually observe are the measurement data θ.

In order to account for this important distinction we need to make some assumptions about how the measurements are related to the actual state of the system. We will assume that the observations are a linear function of the actual state plus random noise,

$$\theta_s = Hx_s + v_s \qquad (13)$$

where H is a known matrix that maps the data to the observations and v is the random measurement error or noise. We introduce the index s here in order to be explicit that we are indicating the values at some space-time location s, *where the observations are*, which is not necessarily the location where the estimate $\hat{x}$ of the field is being made.

With this, we can write

$$C_{x\theta} = E[x(Hx_s + v)^T] = C_{xs}H^T + C_{xv} \qquad (14)$$

Applying (13) to the definition of C_θ, gives

$$C_\theta = E[(Hx_s + v)(Hx_s + v)^T] = HC_xB^T + C_{sv}^TB^T + BC_{sv} + C_v \qquad (15)$$

If we suppose that *the actual state and the noise are uncorrelated*, then the terms, C_{xv} and C_{sv} are each zero.

So now we have

$$\hat{x} = C_{xs}H^T[HC_sH^T + C_v]^{-1}\theta \qquad (16)$$

and,

$$C_\varepsilon = C_x - C_{xs}H^T[HC_sH^T + C_v]^{-1}(C_{xs}H^T)^T \qquad (17)$$

If we consider the case where the measurements θ are the same quantity as what we are estimating *x* (i.e. we are using density data to estimate the density field, as opposed to using salinity and temperature to estimate the density field), then *H* is just the identity matrix, so our estimator is,

$$\hat{x} = C_{xs}[C_s + C_v]^{-1}\theta \tag{18}$$

and,

$$C_\varepsilon = C_x - C_{xs}[C_s + C_v]^{-1}C_{xs}^T \tag{19}$$

If the noise is white noise, then C_v is a diagonal matrix and we see that the effect of not having the true state correlations, but estimates of it based upon the observations, is to increase the diagonal elements of the matrix to be inverted by the measurement noise variance.

Appendix-3

The derivation of the **Gauss-Markov** theorem depended upon the matrix identity,

$$(A - BC^{-1})C(A - BC^{-1})^T - BC^{-1}B^T \equiv ACA^T - BA^T - (BA^T)^T \quad (1)$$

This identity is not very intuitive and so we will provide the proof here. This proof depends upon one assumption being true, that $C = C^T$.

Let us define,

$$X \equiv (A - BC^{-1})C(A - BC^{-1})^T - BC^{-1}B^T \quad (2)$$

and,

$$Y \equiv ACA^T - BA^T - (BA^T)^T \quad (3)$$

If we can establish that $X \equiv Y$, then we have our proof.

We start by expanding X,

$$\begin{aligned} X &\equiv (A - B\,C^{-1})\,C\,(A - B\,C^{-1})^T - B\,C^{-1}\,B^T \\ &= (A - B\,C^{-1})\,C\,(A^T - C^{-T}B^T) - B\,C^{-1}\,B^T \end{aligned} \quad (4)$$

and since we have assumed that $C = C^T$, then $C^{-1} = C^{-T}$, so

$$\begin{aligned} X &= (A - B\,C^{-1})\,C\,(A^T - C^{-1}B^T) - B\,C^{-1}\,B^T \\ &= (A\,C - B\,)\,(A^T - C^{-1}B^T) - B\,C^{-1}\,B^T \\ &= ACA^T - BA^T - AB^T + BC^{-1}B^T - BC^{-1}B^T \\ &= ACA^T - BA^T - (BA^T)^T \equiv Y \end{aligned} \quad (5)$$

Appendix-4

In this note we prove that a covariance matrix is non-negative definite.

Consider the product, *H*, that is the product of an arbitrary vector, *a*, and the covariance vector, *x*:

$$H = a^T E[(x-\mu)(x-\mu)^T]a \tag{1}$$

We can re-arrange the above by moving the constant vector *a* inside expectation operator so that we have,

$$H = E[a^T(x-\mu)(x-\mu)^T a] \tag{2}$$

If we define

$$Y \equiv (x-\mu)^T a \tag{3}$$

(which is a random variable because *x* is), then (2) is

$$H = E[\, Y^T Y] \tag{4}$$

This ``squared" quantity is clearly never negative, so that we can conclude that the covariance matrix $E[(x-\mu)(x-\mu)^T]$ is non-negative definite.