Technical Report

Report No.: INCOIS-MOG-TR-01-2009

Delayed Mode Quality control of Indian Argo Floats

By

Sudheer Joseph and M. Ravichandran

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

June, 2009

Document Control Sheet

1. Report No: INCOIS-MOG-TR-01-2009 Date:15-Jun-2009
2. Title: Delayed Mode Quality control of Indian Argo floats
3. Part No.:...........
4. Vol. No.:..........
5. Author(s): Sudheer Joseph and M. Ravichandran
6. Originating agency (Group/Project/Entity): MOG
7. No. of Pages: 32
8. No. of figures: 15
9. No. of references: 10
10. No. of enclosures/appendices: 1- CDROM
11. Abstract (Maximum 100 words):
An observational network array of about 3000 autonomous profiling floats, known as Argo, has been evolving since the year 2000. India is a participant of the International Argo program, and Indian National Centre for Ocean Information Services(INCOIS) has the overall responsibility for implementation of the Indian Argo project, including the Argo regional centre (ARC) for the Indian ocean. One of the prime responsibility of regional data centre is delayed mode duality control (DMQC). The Argo array is delivering large number vertical profiles of temperature, salinity and other parameters from the surface to depths up to 2000 m. While floats are expected to give good measurements of temperature and pressure, salinity measurements sometimes show significant sensor drift with time or offsetsThe calibration procedure is called WJO method in short and OW method, after the recent modification by Owens and Wong [2009]. This report presents the major components of the work carried out at INCOIS using the WJO/OW software to provide best quality data to the end user. This report also presents all floats calibrated at INCOIS and their WJO/OW recommended corrections and the final decision taken about each floats considering the regional oceanography and various other factors (in attached CD-ROM).
12. Keywords: Argo, Salinity sensor, DMQC, Calibration
13. Security classification: Unrestricted

Delayed Mode Quality control of Indian Argo floats

Sudheer Joseph and M. Ravichandran

Indian National Centre for Ocean Information Services, Hyderabad

Contents

1 Introduction & Background 4
2 The Correction method 5
2.1 Applying corrections 10
2.2 Data preparation for DMQC 12
2.3 Visual quality control 13
3 Generation of D-Mode NetCDF files 14
3.1 The Work flow 14
3.2 Results of Experiments carried out for region specific characterization of Salinity variability 16
3.3 Status of Calibrated Indian Floats 20
4 Examples of Diagnostic Plots and Calibrated Float Parameters 21
4.1 Diagnostic Plots 21
5 Recent Improvements in Method of Calibration 26
6 Practical steps in carrying out calibration 28

List of Figures

1 Package Directory Structure 13
2 DMQC Flow Chart 15
3 Location Map for the Test Float 17
4 Calibration Using WOA and Argo 18
5 Calibration using only near by Argo profiles 19
6 Calibration using WOA alone 20
7 Comparison of delayed mode status with other DAC (ref: Coriolis Web site) 21
8 Pre - Calibration(left) & Post Calibration(right) TS diagrams, scatter at close to the bottom indicate clear case of sensor drift, which is rectified in right figure 22
9 Multiplicative (Top) and Additive (Bottom) corrections applied to the time series) 23
10 Salinity anomaly plot: Indicates that salinity sensor started drifting from about 65^{th} profile. 24
11 Salinity anomaly plot: Indicates that salinity sensor started drifting from about 65^{th} profile. 25
12 Pre-calibration & Post-calibration TS diagrams of OW method 27
13 Suggested correction by OW calibration method - Top left panel - suggested correction in terms of conductivity, Bottom left - suggested correction in terms of Salinity. Right panels, depicts quantum of correction applied on 2 isotherms 27
14 Pre-calibration & Post-calibration salinity anomaly plot on potential temperature 28
15 salinity variances on θ levels & Pressure levels. It also shows the 10 float θ levels that are chosen for calibration 29

Summary

An observational network array of about 3000 autonomous profiling floats, known as Argo, has been evolving since the year 2000. India is a participant of the International Argo program, and Indian National Centre for Ocean Information Services(INCOIS) has the overall responsibility for implementation of the Indian Argo project, including the Argo regional centre (ARC) for the Indian ocean. One of the prime responsibility of regional data centre is delayed mode duality control (DMQC). The Argo array is delivering large number vertical profiles of temperature, salinity and other parameters from the surface to depths up to 2000 m. While floats are expected to give good measurements of temperature and pressure, salinity measurements sometimes show significant sensor drift with time or offsets. Unless a float is recovered before the battery fails, recalibration cannot be performed and a remote calibration method is required. A quality control system based on Wong et al. [2001], Owens and Wong [2009] has been set up for the Indian ocean to identify and correct salinity sensor drifts by using historical hydrographic data. In this system an objective mapping method is used that takes the spatial and temporal variations in water mass properties into account. The float salinity data are fitted to the background climatology in potential conductivity space by weighted least squares with a time-varying slope. It is assumed that any conductivity offset changes slowly over time, so that a linear fit of the profile based corrections over the float time series is done. The procedure generates a set of calibrated salinity data with corresponding uncertainties. The calibration procedure is called WJO method in short and OW method, after the recent modification by Owens and Wong [2009]. This report presents the major components of the work carried out at INCOIS using the WJO/OW software to provide best quality data to the end user. This report also presents all floats calibrated at INCOIS and their WJO/OW recommended corrections and the final decision taken about each floats considering the regional oceanography and various other factors (in attached CD-ROM).

1 Introduction & Background

The World Climate Research Programme's (WCRP) Climate Variability and Predictability Experiment (CLIVAR) through its Upper Ocean Panel (UOP) and (Global Ocean Data Assimilation Experiment) GODAE have taken up Argo as one of their pilot programs. Argo program's Objectives may be summarized as

- To provide a quantitative description of the changing state of the upper ocean and the patterns of ocean climate variability from months to decades, including heat and freshwater storage and transport.
- To enhance the value of the Jason altimeter through measurement of subsurface temperature, salinity, and velocity, with sufficient coverage and resolution to permit interpretation of altimeter sea surface height variability.
- For initializing ocean and coupled ocean-atmosphere forecast models, for data assimilation and for model testing.
- To document seasonal to decadal climate variability and to aid our understanding of its predictability. A wide range of applications for high-quality global ocean analyzes is anticipated.

In order to meet the above objectives, Argo data team has prescribed certain quality standards for each of the measured parameters. The target accuracies for temperature and salinity sensors of Argo floats are 0.01^oC and 0.01 PSU [Bhme and Send, 2005] respectively. In general the temperature and pressure are accurate to 0.002^oC and 2.4 db and salinity is expected to be accurate to 0.01 PSU in case of floats without sensor drift [Wong et al., 2001].

Under Indian Argo project, about 168 autonomous CTD profiling floats are deployed in Indian Ocean with a planned replenishment of about 40 floats per Annam. These floats, move freely with ocean currents at fixed parking depths and cycle from parking depth to the sea surface at pre-fixed time intervals . While rising to the surface, floats take measurements of Conductivity(C), Temperature(T) and pressure (P) at pre-fixed depth levels (pressure levels). The float measured data are received at the reception centers across the globe via satellites, before the floats sink back to their parking depths to repeat the cycle.
Due to a multitude of reasons arising from their operation in open oceanic environment, float sensors may develop drift in their accuracies. Bio-fouling is attributed as one of the potential cause for the non-stability of the autonomous CTD profiling floats. During bio-fouling, secretions of the fouling organisms get deposited on the conductivity cell and as a result the measurements made by such conductivity cells tend to give erratic results. However there is another reported cause for sensor inaccuracy arising from the leaching of toxic anti-fouling substances coated on sensors by some manufacturers to deter bio-fouling. In addition to this temporary material attachments which block the free flow of sea-water through the conductivity cells too results erratic float measurements.

2 The Correction method

Wong et al. [2001] strongly base their method on the fundamentals of physical oceanography, which states that the temperature and salinity are conservative in their variability. The two conservative variables of ocean, potential temperature, θ and salinity, S are related to each other by region specific relationships which represent the mean character of the region [Emery and Dewar, 1982]. The $\theta - S$ relationship could be influenced by seasonal or decadal cycles and also by strong eddies which are region specific. Apart from few anomalies, in most of the global oceans, mean $\theta - S$ relationship can be used for estimation of salinity from temperature and pressure. The accuracy of the estimated salinity will depend on extend of spatio - temporal variability of the region and also on the degree of accuracy of the sampling of the region by CTD reference profiles.

A high quality CTD data-base is essential for carrying out a meaningful calibration. As of today Indian Ocean Hydro base (IOHB) [Kobayashi and Suga, 2006] and the recently assembled special data set for Delayed mode QC (version 2009 : `ftp://ftp.ifremer.fr/coriolis/data/DMQC-ARGO/ARGO_for_DMQC/ARGO_for_DMQC_2009V01/ARGO_for_DMQC_2009V01.tar.gz`) forms the basis of the reference data base. We augmented this database adding CTD profiles obtained from different cruises carried out by research institutes and universities in India. The reference data base is specially formatted for the purpose of calibration. To capture most of the water column variability of the world ocean, 54 standard θ levels are chosen from -1°C to 30°C. This was achieved by interpolating the non-uniformly sampled CTD profiles collected during different oceanographic campaigns. A shape preserving spline interpolation scheme [Akima, 1970] is used for carrying out the interpolation in vertical to the standard 54 θ levels. In case of thermal inversions, only salinity on the deepest isotherm was used.

The representative climatological values of salinity at the location of the Argo profile of interest was obtained by vertically interpolating historical salinity data and an objective analysis method. The objective analysis method was based on Gauss-Markov theorem which is suitable for obtaining point wise estimate which is linear and unbiased. The method is optimal in least square sense and it is possible to get an estimate of error [Bretherton et al., 1976, McIntosh, 1990]. This approach ensures that both spatial and temporal variability in climatological $\theta - S$ relationships are taken care. The data covariance is assumed to be Gaussian, and the decay scale is determined by three scale parameters: a longitudinal scale, Lx; A latitudinal scale Ly; and a temporal scale τ. The spatial scales are not equal in north-south and east-west directions which is adopted to take care the influence of the domination of zonal currents over meridional currents in the oceanic interior. Two sets of spatial scales are used by Wong et al. [2001] for ensuring the representation of the large scale field and small scale fields. These are denoted as (Lx_1, Ly_1) and (Lx_2, Ly_2) with default values are set as $(Lx_1 = 20°, Ly_1 = 10°), (Lx_2 = 8° Ly_2 = 4°)$. Considering the geographical restrictions and regional water mass variability scales), these parameters are selected as $(Lx_1 = 4°, Ly_1 = 2°)$ and $(Lx_2 = 2°, Ly_2 = 1°)$ for the application in the Indian ocean. The temporal variability parameter, τ is estimated using the ventilation timescales which is based on the partial pressure of chlorofluorocarbon(CFC-12). A global CFC-12 data sets of World Ocean Circulation Experiment (WOCE)($http://whpo.ucsd.edu$) campaign is used for estimating temporal scale τ at various θ surfaces. For each float profile which is defined by spatio-temporal co-ordinates $(x_\circ, y_\circ, t_\circ)$ on available θ levels of the pre-selected 54 θ levels, CTD data points are selected from the WOA data base which is used as reference data base. This is done by choosing an ellipse with radii Lxi and Lyi with $(x_\circ, y_\circ)$ at the centre. (for representing the large scale field). From this primary set, 600 historical data points are selected for objective mapping with the following criteria. First a random selection of 200 data points are made from the initial elliptical area which is for representation of large scale mean in the data sets chosen. From the remaining data sets another 200 historical data points (x_i, y_i, t_i) with shortest spatial separation factor relative to large length scales, $(x_i - x_\circ)^2/Lx_1^2 + (y_i - y_\circ)^2/Ly_1^2$. The above step is to assure best spatial correlation of the historical data points with the float profile. Finally additional 200 historical data points are selected so that there will be best spatio-temporal closeness with float profile, which is with respect to the small length scales and the temporal scale, which is achieved by solving

$$(x_i - x_\circ)^2/Lx_2^2 + (y_i - y_\circ)^2/Ly_2^2 + (t_i - t_\circ)^2/\tau^2. \qquad (1)$$

The above three criteria makes sure that the choice of the hydrographic data is not biased towards cruise tracks which may have dense sampling along the lines of cruises. Also the historical data points will be closer to the float data in both time and space. So the resulting objective estimate will reflect the mean $\theta - S$ relationship in the region with its spatio-temporal variability. In cases were floats are near to the coast line, the values of (Lx_1, Ly_1) and (Lx_2, Ly_2) has to be reversed to that the longer axis of the ellipse becomes parallel to the coast. This is done as the flow along the coasts will be parallel to the coast and also to avoid land mask coming the selected ellipse.

The objective estimate of salinity, $S^{'}$ at each location at every suitable standard θ level is estimated by solving the below equation.

$$S^{'} = \langle d \rangle + \omega.(d - \langle d \rangle) \qquad (2)$$

where $d = [d_1, ...d_m]$ represents a set of selected historical data on the standard θ levels and $\langle d \rangle$ is the mean of set d. For every data point d_i at (x_i, y_i, t_i), there is a true signal s_i, and noise η_i which includes measurement errors, random processes and natural variability in the ocean which result in deviations from climatology. By assuming $d_i = s_i + \eta_i$, signal variance and noise variance of the data can be estimated, which is incorporated into the coefficient matrix ω. The signal variance is approximated by $(1/m)\Sigma_i(d_i - \langle \mathbf{d} \rangle)^2$ where m is the number of data points on each θ surface. The noise variance is estimated by $(1/2m)\Sigma_i(d_i - d_j)^2$, where d_j is the data point with shortest distance from d_i, on each θ surface. The fundamental assumption in this computation is that the noise is uncorrelated over distance, there exists a uniform variance, and the signal has a longer correlation distance than the data separation.

The coefficient matrix ω in equation (2) is defined as $\omega = \mathbf{C}dg.(\mathbf{C}dd)^{-1}$, where $\mathbf{C}dg$ denotes the data grid covariance matrix and $\mathbf{C}dd$ denotes the data -data covariance matrix. The objective mapping scheme employs a two-stage mapping scheme. In stage one, covariance is assumed to be a function of large-scale spatial separation only, and the decay scale which is assumed to be Gaussian is determined by large spatial scales Lx_i and Ly_i. The mapping is done by solving the following equations.

$$\begin{aligned} \mathbf{C}dd_{ij}(x,y) &= exp\Big\{-\Big[\frac{(x_i - x_j)^2}{Lx_1^2} + \frac{(y_i - y_j)^2}{Ly_1^2}\Big]\Big\}, \\ \mathbf{C}dg_i(x,y) &= exp\Big\{-\Big[\frac{(x_i - x_\circ)^2}{Lx_1^2} + \frac{(y_i - y_\circ)^2}{Ly_1^2}\Big]\Big\}. \end{aligned} \tag{3}$$

By using equation (2) and (3), the reference data is mapped to the location of the float profile, and also to the historical data points themselves. The deviations of the original values and the estimated values at reference data location gives the residual estimate. The primary estimate at the location of the float profile, S_1', is a large-scale estimate without taking care of temporal variability or small-scale features. During the secondary stage of estimation, the residuals from first stage is mapped to the float location using equation (2) but the covariance used in this case is a function of the temporal deviation and small sale deviation. The decay scale is determined using the small spatial scale Lx_2 and Ly_2 together with temporal scale τ.

Second stage estimate uses $\mathbf{C}dg$ $\mathbf{C}dd$ as defined below.

$$\begin{aligned} \mathbf{C}dd_{ij}(x,y,t) &= exp\Big\{-\Big[\frac{(x_i-x_j)^2}{Lx_2^2}+\frac{(y_i-y_j)^2}{Ly_2^2}+\frac{(t_i-t_j)^2}{\tau^2}\Big]\Big\}, \\ \mathbf{C}dg_i(x,y,t) &= exp\Big\{-\Big[\frac{(x_i-x_\circ)^2}{Lx_2^2}+\frac{(y_i-y_\circ)^2}{Ly_2^2}+\frac{(t_i-t_\circ)^2}{\tau^2}\Big]\Big\}. \end{aligned} \tag{4}$$

The second stage take in to account the small scale features and gives lower weight to temporally distant data from the profile of interest. The final objective estimate at the float location is the sum of the primary and secondary stages of mapping, given as $S_f^{'} = S_1^{'} + S_2^{'}$. When there are historical data closer to the profile in time and space, the objective estimate will be of high quality. In cases where the temporal gap between float and historical data exceeds τ, the secondary stage contribution will be less and the final estimate will be relaxed back to the primary stage map and errors are bound to be larger.
During primary stage of mapping, $\mathbf{C}dd$ and $\mathbf{C}dg$ are scaled by the signal variance of the historical data and during second stage both are scaled by signal variance of the residuals. The final error of the objective estimate of salinity at the float profile location is taken from the second-stage mapping:
$\sigma^2_{map}(S_f^{'})$ = signal variance of the residuals **-**$\mathbf{C}dg(x,y,t)^{-1}$
$\mathbf{C}dd(x,y,t)^{-1}\mathbf{C}dg(x,y,t)^T$

2.1 Applying corrections

The float data are corrected by correction factors obtained by fitting the objectively estimated climatological salinity field on the standard θ surfaces by weighted least squares. Potential conductivity is used for application of correction as it is more close to the original measured parameter conductivity and is not influenced by pressure. The equation for correction of potential conductivity for the i^{th} profile of a float takes the form

$$C_i^{'} = r_i C_i + \epsilon_i \tag{5}$$

where C_i is the potential conductivities measured by float, $C_i^{'}$ is the corrected potential conductivity, r_i is the multiplicative correction term, and ϵ_i is the assumed model error. The multiplicative correction term r_i is computed by using standard weighted least squares minimization between float and the climatology. The correction term r_i is solved for individual float profiles F_i. For every profile F_i, r_i is found by minimizing a $2k + 1$ ($k < 0$) profile series of differences between the float potential conductivities and those from climatology. F_i is assumed to be at centre with k profiles available before and k profiles available after F_i. In matrix mode the problem can be presented as:

$$Gm + \epsilon = D \tag{6}$$

where G is the matrix with time series of float conductivities (C_i , s). D the matrix with climatological potential conductivities obtained after objective estimation., m are the model parameters and ϵ are the model errors.

$$\begin{bmatrix} C_{i-k}\ at\ \theta(1) & C_{i-k} \times (-k) \\ \vdots & \vdots \\ C_{i-k}\ at\ \theta(n_{i-k}) & C_{i-k}\ at\ \theta(n_{i-k}) \times (-k) \\ \vdots & \vdots \\ C_i\ at\ \theta(1) & C_i\ at\ \theta(1) \times 0 \\ \vdots & \vdots \\ C_i\ at\ \theta(n_i) & C_i\ at\ \theta(n_i) \times 0 \\ \vdots & \vdots \\ C_{i+k}\ at\ \theta(1) & C_{i+k}\ at\ \theta(1) \times (k) \\ \vdots & \vdots \\ C_{i+k}\ at\ \theta(n_{i+k}) & C_{i+k}\ at\ \theta(n_{i+k}) \times (k) \end{bmatrix} \begin{pmatrix} ri \\ \partial r_i \end{pmatrix} = \begin{bmatrix} C_{i-k}^{'}\ at\ \theta(1) \\ \vdots \\ C_{i-k}^{'}\ at\ \theta(n_{i-k}) \\ \vdots \\ C_i^{'}\ at\ \theta(1) \\ \vdots \\ C_i^{'}\ at\ \theta(n_i) \\ \vdots \\ C_{i+k}^{'}\ at\ \theta(1) \\ \vdots \\ C_{i+k}^{'}\ at\ \theta(n_{i+k}) \end{bmatrix} \tag{7}$$

With the multiplicative correction term, r_i and its time derivative, dr_i a float profile F_i with n_i number of θ levels, the set of equations solving the calibration problem may be represented as above in matrix format. As climatologically estimated profiles are with variable uncertainties at different depth levels, they impart varying influence on calibration constants. This is addressed by defining a diagonal matrix, $\mathbf{W}$ of dimension $\Sigma n_i \times \Sigma n_i$, where the diagonal elements are chosen to be the reciprocal of the mapping error variance corresponding to the potential conductivities in $\mathbf{D}$. Thus $\mathbf{W} = diag[\sqrt{1/\sigma^2_{map}(C')}]$. This step ensures that θ surfaces where $\theta - S$ relationship are more stable are used preferentially for calibration. Here the equation -(6) can be re-written as

$$\mathbf{G}^{'}\mathbf{m} + \epsilon = \mathbf{D} \tag{8}$$

where $\mathbf{G}^{'}\mathbf{m} = \mathbf{WG}$ and $\mathbf{D}^{'} = \mathbf{WD}$.

2.2 Data preparation for DMQC

For each float, the original float data has to be arranged in matrix form, with each column being a profile, in ascending order (i.e. column 1 contains the first profile of the float, column 2 contains the second profile, etc.), in the following variable names:

DATES (1xn, in decimal year, e.g. 10 Dec 2000 = 2000.939726)
LAT (1xn, in decimal degrees, -ve means south of the equator, e.g. 20.5S = -20.5)
LONG (1xn, in decimal degrees, from 0 to 360,)
PRES (mxn, dbar, from shallow to deep, e.g. 10, 20, 30 ...)
TEMP (mxn, in-situ IPTS-68)
SAL (mxn, PSS-78)
PTMP (mxn, potential temperature referenced to zero pressure,

use SAL in PSS-78 and in-situ TEMP in IPTS-68 for calculation, e.g. sw_ptmp.m)
PROFILE_NO (1xn, this goes from 1 to n, PROFILE_NO is the same as CYCLE_NO in the Argo NetCDF files)
The extra spaces in the columns are filled up with NaNs to make up the matrices. Bad data should also be denoted by NaNs. Save the matrices in a ".mat" file in Matlab in the appropriate sub directory in /data/float_source/. There should be one ".mat" file for each float. Presented below is a typical directory structure for Argo calibration system.

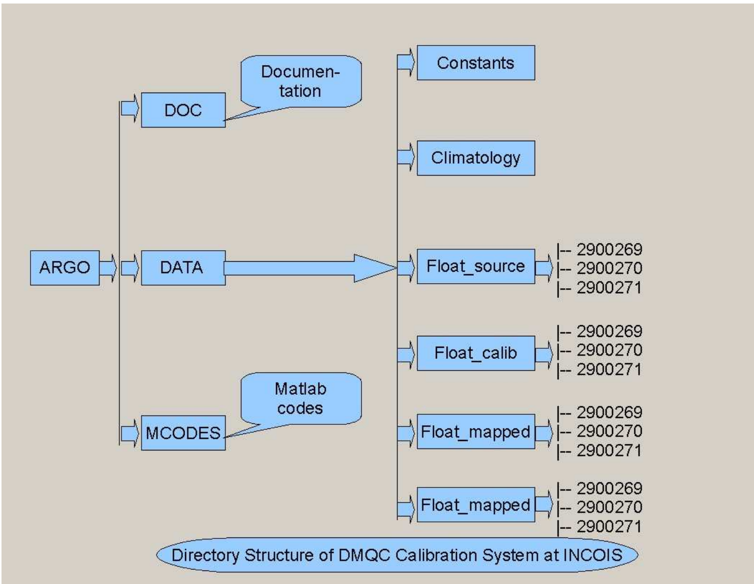

Figure 1: Package Directory Structure

2.3 Visual quality control

The DMQC calibration system of Wong et al. [2001] can only take care of systematic errors occurring in float sensor due to different factors which are described elsewhere. However float data will have different kinds of data errors which need to be corrected before subjecting the float time series to calibration. In order to carryout pre-DMQC quality control, a customized version of the Scrips Institution's (written by John Gilson) matlab GUI package is used at INCOIS. The GUI allows the user to have plots of measured parameters in different modes and the commonly used plots are 'TEMP vs Depth', 'PRES vs Depth', 'SAL vs Depth', 'SAL vs TEMP' etc. The parameters also are compared with the near by floats data as well as the near by historical CTD data before subjecting them to calibration. If any spikes are found at any level of the data those are flagged and are not used for calibration purpose.

3 Generation of D-Mode NetCDF files

Delayed mode calibrated data has to be presented to the user community in Net Common Data Format or NetCDF. The individual float profiles are packed in to individual NetCDF files and are disseminated with additional vectors of calibrated data along with the un- calibrated data vectors.

3.1 The Work flow

Figure 2 presents the flow-chart of the DMQC procedure which is adopted Internationally. Internationally Argo program is envisaged as funded within different scientific programs under different Principal Investigators(PI's) and each PI is responsible for the quality control of Argo data.

Each country has National data assembly centers(DAC) where the PI's deposit the real time and delayed mode quality controlled data from where the data is uploaded to the Global Data Assembly Centers(GDAC). As far as India is concerned

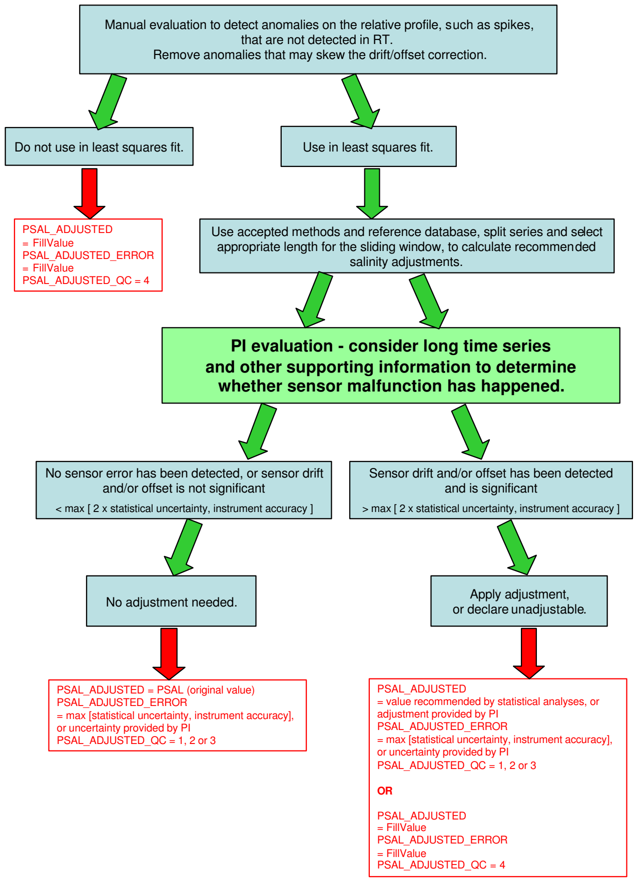

Figure 2: DMQC Flow Chart

the entire Argo program is coordinated by INCOIS and director INCOIS is the formal Principal Investigator of the Indian Argo project.

3.2 Results of Experiments carried out for region specific characterization of Salinity variability

The Indian Ocean is considered as one of the most complex oceanic region which requires special attention for any oceanographic calibration effort. This is due to the fact that the Indian Ocean is less scientifically explored than the other basins resulting sparse data record for calibration purpose and strong variability existing here due to monsoonal influence. Considering this, and the difficulties involved in getting good calibration results from the region, the variability in upper layers was examined [Joseph and Freeland, 2005] as a first step towards understanding the modes of variability and the results obtained demonstrated that there is a strong spatial dependance on temporal variability of salinity in upper layers of the Arabian Sea. To further the understanding of the temporal window effect on calibration, the first workshop on delayed mode quality control held at Scrips Institute of Oceanography Santiago, USA suggested to carry out experiments involving different combinations of Argo data and historical data so that the influence of temporal variability to the calibration effort can be assessed. In order to assess the efficacy of the reference data base and to analyze the response of calibration system to the addition of recent data we conducted experiments involving both original reference data and also with adding recent Argo data to it. The results are summarized below.

Criteria used for addition of Argo data sets to the reference data base:

- Argo data with quality flag of 1 for all the parameters
- Profiles which are deeper than 900 dbar
- Three parameters (P, T and S) from each profile were plot-

ted and visually examined for any apparent problems by checking waterfalls, P vs T, P vs S, T vs S, contour plots and deeper salinity time series with 2x standard deviation bars.

- Profiles with long vertical sampling gaps at deeper layers were avoided. In Indian Ocean, floats are deployed with different profiling strategies. Many are having shallow regular profiling and occasional deep profiling. This is a set back for building up reference data base as only deeper profiles can contribute to better calibration. However all the profiles which are deeper than 900 dbar were selected for inspection of other obvious problems. A Matlab routine is made which look for floats falling in a given WMO grid and generate decision support graphs and the reference data sets in the required format (xxxxxxx_prof.nc)

Experiments :-

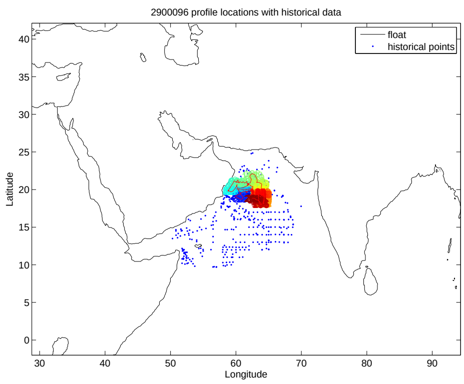

Figure 3: Location Map for the Test Float

After preparation of the reference data sets by inclusion of Argo data, WJO was run under 3 conditions for floats which had obvious salinity drifts.

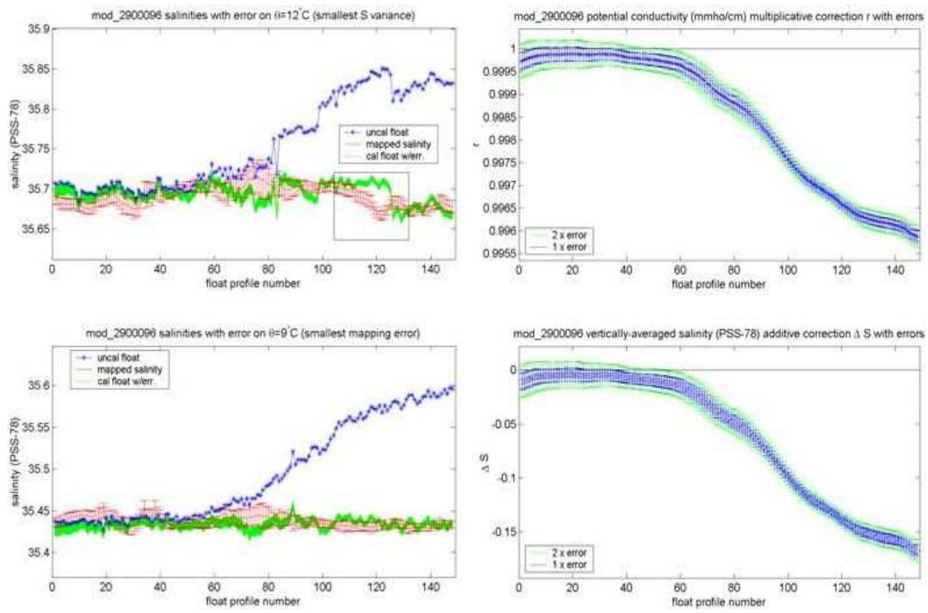

Figure 4: Calibration Using WOA and Argo

- Experiment - I is a run with WOA 2001 based historical CTD data sets with additional good quality Argo reference data set. (In figure the red curves with error bar shows the distribution of background salinity and its error distribution, blue is the original Argo data with out correction and green is the corrected salinity with width proportional to the total error involved in correction procedure). This run resulted in much reduction of the calibration error bar (Fig. 2a and Fig. 2b) and produced an apparently better calibration. This is due to the reduction in average temporal window of the reference database used for calibration.

- Experiment - II

  In this experiment all the old CTD data were flagged so that they were not used by the calibration routine. Only good quality Argo profile data handpicked from the near by

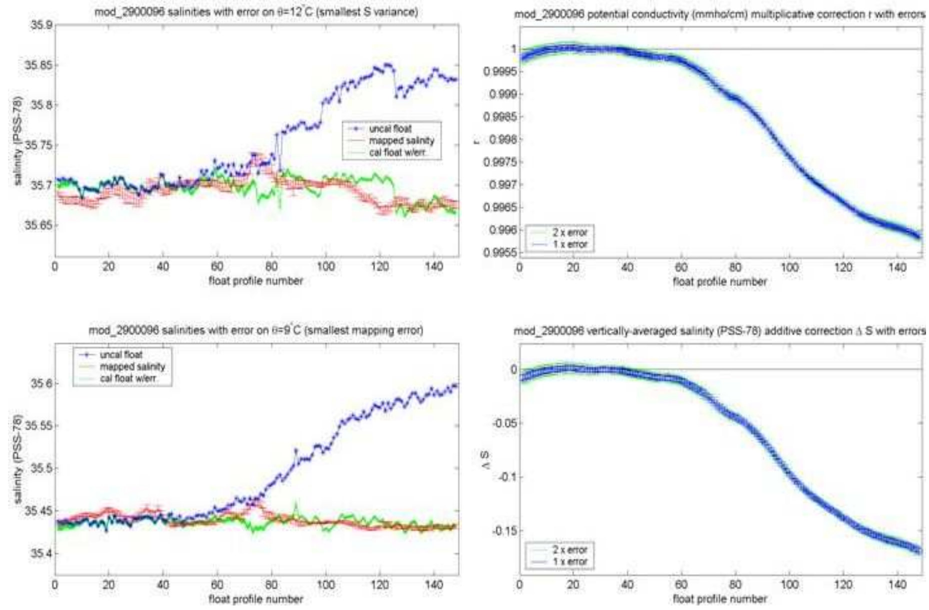

Figure 5: Calibration using only near by Argo profiles

Argo floats are used for this run. This run demonstrate the significance of recent CTD reference data for calibration purpose. With only Argo profiles the width of error bar has considerably reduced and it can produce high quality calibration results. However one has to take care of the building up of errors due to faulty Argo profiles which needs specific care before adding Argo profile to the reference data base.

- Experiment - III is carried out using the default reference data present in the WJO package which is basically based on world ocean atlas 2001. The reference data base is clearly inferior as it contains relatively older data only. As a result a very broad error bar is produced by the calibration software with high uncertainty.

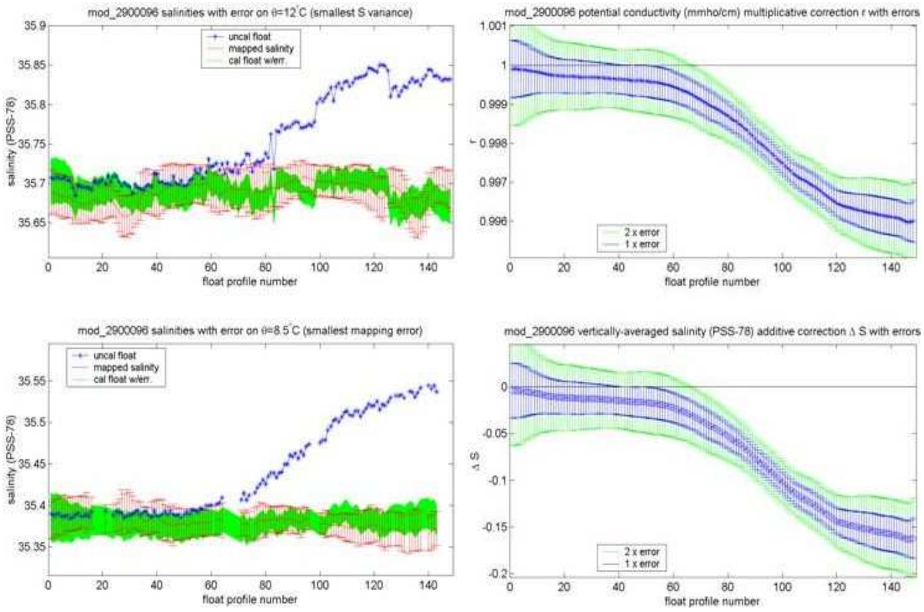

Figure 6: Calibration using WOA alone

3.3 Status of Calibrated Indian Floats

As on 14 May 2008, there are 15665 INCOIS float profiles existing in Global Data Assembly Center at Ifremer, France. Among this 8720 are D files (Files which were subjected to Delayed Mode QC). There are 3620 profiles which are not aged for DMQC and remaining 3325 profiles are to be subjected to DMQC.

Comparison with International scenario:

In spite of the complexities involved, considerable progress has been achieved in delayed mode quality control of the Indian Floats by adopting different strategies at difficult float locations and by interaction with international experts. Figure 7 present the comparison of Indian status with different data assembly centers across the globe.

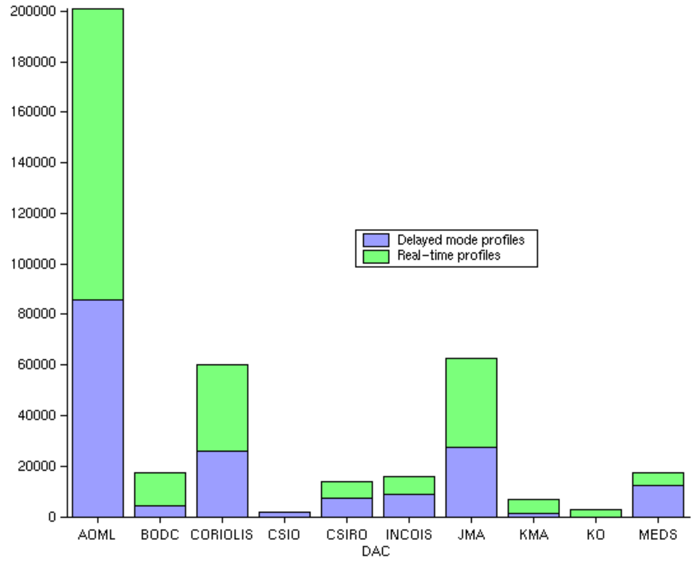

Figure 7: Comparison of delayed mode status with other DAC (ref: Coriolis Web site)

4 Examples of Diagnostic Plots and Calibrated Float Parameters

In this section examples of diagnostic plots of calibrated floats and their netcdf variables are presented for users aimed at helping them in correct usage of Delayed mode files.

4.1 Diagnostic Plots

The Delayed mode software generates 6 diagnostic plots which are aimed at the DMQC operator to make decisions on the kind of corrections to be adopted for calibration. The naming of the graphics outputs is of the convention "$float\_id\_1, ...6.ps$".

Figure 1 - form the software provides the location map of the

float time series and the reference data density in the float location. This figure is very essential to make judgment of how valid the correction suggested by the software, by deciphering the size and shape of eclipse from which the reference data is drawn and also providing insights on how many reference data points are available to the software for calibration(example - Figure 3 on page no 14).

Figure 2 - is(Figure - 8 basically a pre - calibration TS diagram which plots the float salinity against the standard potential temperature levels. This figure helps in identification of any apparent drift at deeper layers where one expects TS convergence.

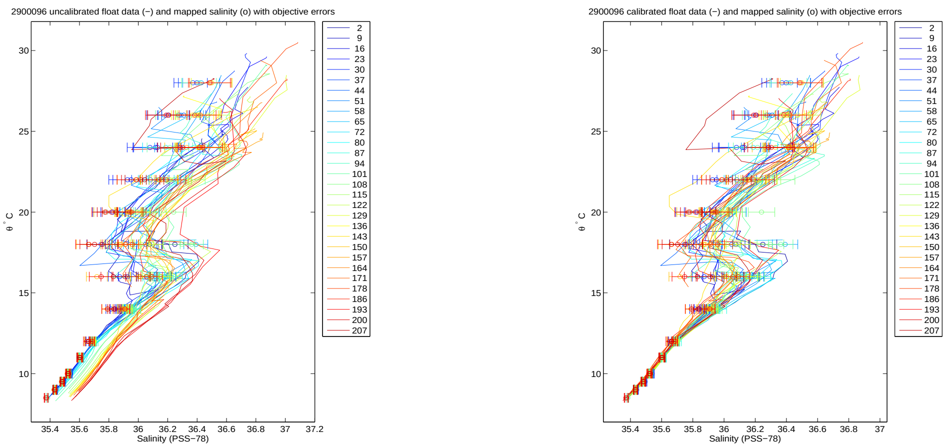

Figure 8: Pre - Calibration(left) & Post Calibration(right) TS diagrams, scatter at close to the bottom indicate clear case of sensor drift, which is rectified in right figure

Figure 3 - Third figure(Figure - 9) of the calibration diagnostics output provides temporal change in calibration correc-

tions applied to the data along with standard errors.(Top panel presents the multiplicative corrections carried out to the conductivities and bottom panel presents the correction in additive correction in terms of salinity.)

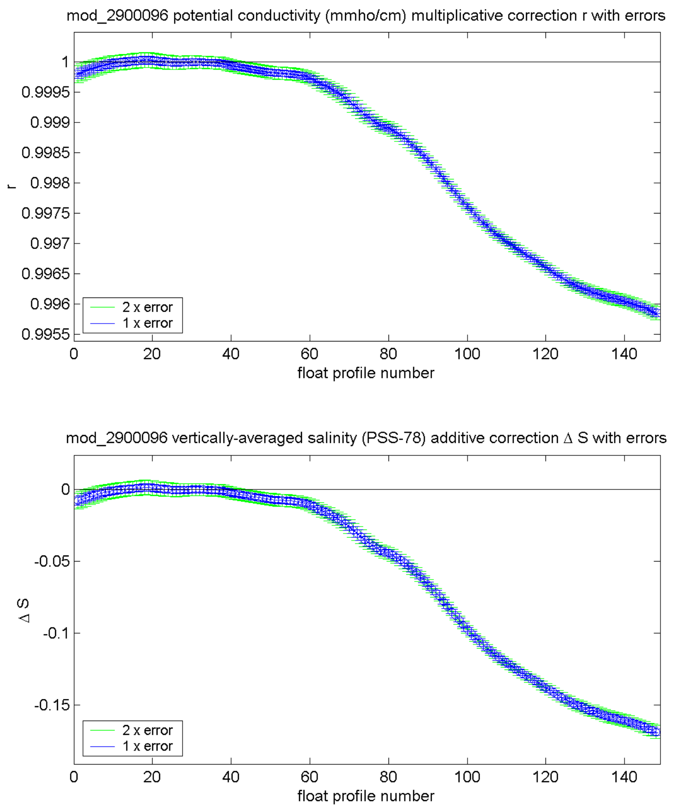

Figure 9: Multiplicative (Top) and Additive (Bottom) corrections applied to the time series)

Figure 4 - presents a post - calibration TS diagram which plots the float salinity against the standard potential temperature levels. This figure helps in assessing the impact made by the calibration software on the data subjected to calibration. In Figure -8 (right panel) improvement in TS convergence is discernible.

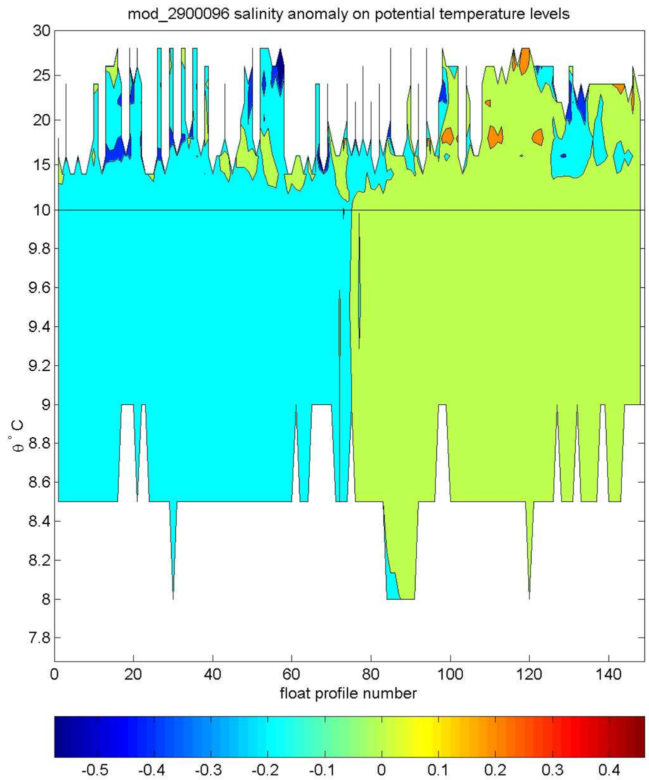

Figure 10: Salinity anomaly plot: Indicates that salinity sensor started drifting from about 65^{th} profile.

Figure 5 - Fourth output figure presents an anomaly map of salinity with cycle number on x axis and potential temperature on the vertical axis. This figure aids in detecting artificial sensor drift by observing salinity anomalies at different vertical levels. Salinity at each theta level will vary either due to genuine sensor drift or due to water mass changes or due to oceanic changes with time. From the plot sensor drift can be easily detected as a change affecting all the levels across the water masses from top to bottom.

Figure 6 - Presents the average correction applied to the float

profiles from top to bottom(Top panel) and also the most stable theta level(Bottom panel), which received the highest vertical weight in the least square fit. This plot demonstrate the evolution of the float salinity drift along with the applied correction with error bars of correction(width of the green broad line) historical data error(red lines) and also the original data with out applying correction(Blue line)

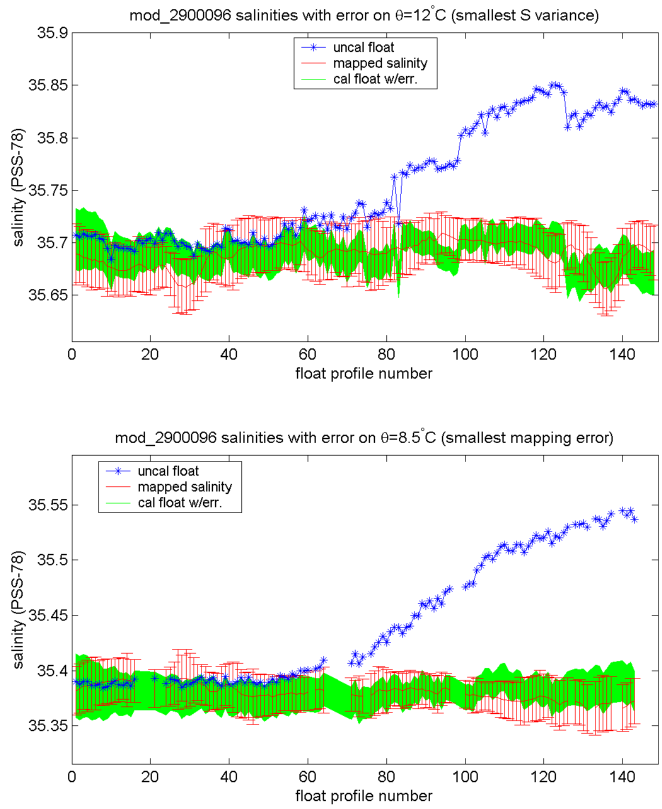

Figure 11: Salinity anomaly plot: Indicates that salinity sensor started drifting from about 65^{th} profile.

A delayed mode netcdf file (D file) is having essentially the same structure as that of a real time file (R file). But in R files the "vectors" starting with "ADJUSTED_" will be vacant or filled with fill values. Also in R files the variables with "_ADJUSTED_ERROR" tag will be left vacant and are filled only during delayed mode processes.

5 Recent Improvements in Method of Calibration

In order to address the specific problems of highly varying oceanic environments, such as north Atlantic, Bhme and Send [2005] developed a method which involves objective mapping of float data taking potential vorticity in to account. In this method the float measurements of each profile are compared to the mapped salinity in potential conductivity space by weighted least-squares, giving one correction for each profile. Very recently, Owens and Wong [2009] evolved a superior calibration methodology, by integrating best features of BS and WJO methods which is named as OW method. In OW method the break points for the calibration at distinct portions of a salinity time series for the lifetime of a float are statistically determined and suggested. This allows the choice of the break points to be more objective. The major Improvements are:

- Improvement in objective mapping method,by introducing a minimum error estimate of mapped values
- Piece-wise linear least-squares fit to the time varying potential conductivity correction factor
- Statistical test based choice of brake points

The revised software with improvements has been configured for Indian Ocean and calibration of floats were successfully carried

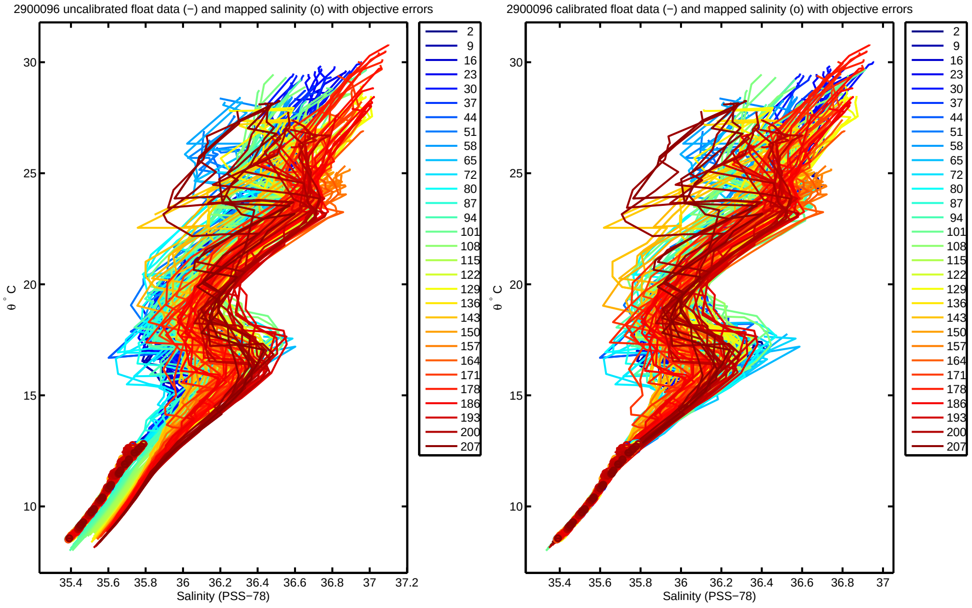

Figure 12: Pre-calibration & Post-calibration TS diagrams of OW method

out. A suitable case for exploring the latest improvements is presented below with graphics and explanations.

6 Practical steps in carrying out calibration

1. Visual Quality check and correction using SIO GUI. look for levels or profiles with,

    - Spikes in Temp, Sal or PRES escaped from RTQC,
    - Thermal inversions
    - Apply thermal mass correction [Johnson et al., 2007]

2. Generating mat files from multiple 'R' netcdf files. Structure of Mat file example below

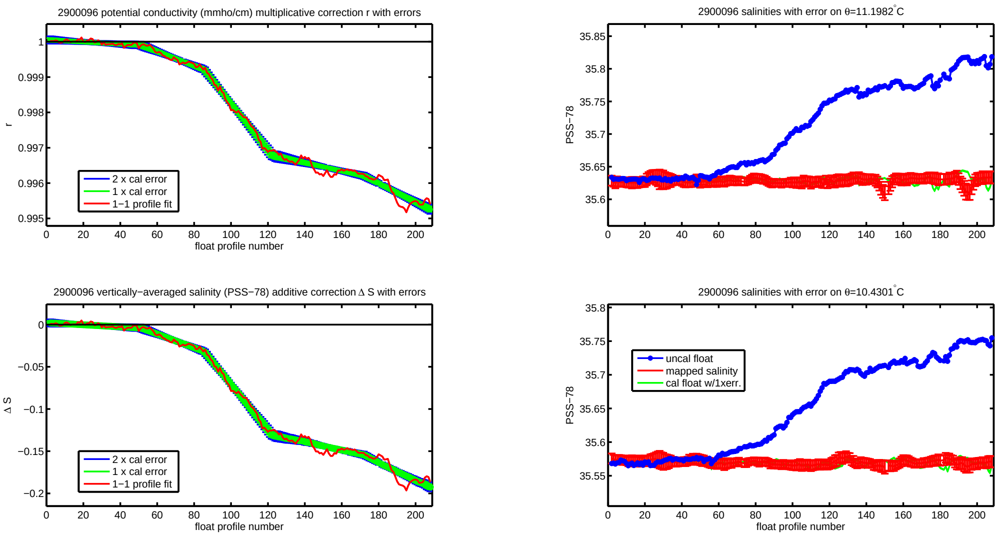

Figure 13: Suggested correction by OW calibration method - Top left panel - suggested correction in terms of conductivity, Bottom left - suggested correction in terms of Salinity. Right panels, depicts quantum of correction applied on 2 isotherms

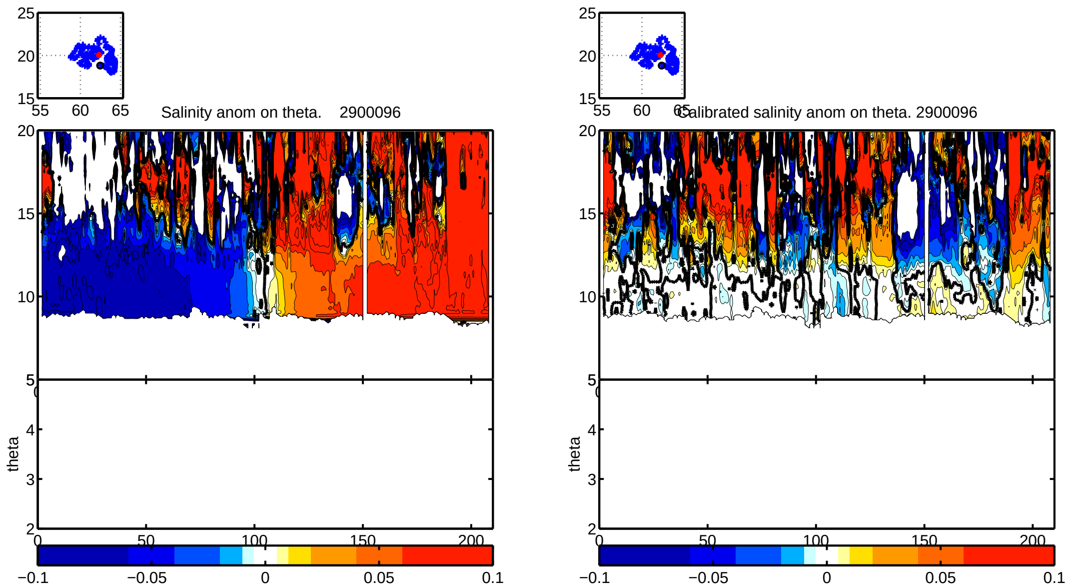

Figure 14: Pre-calibration & Post-calibration salinity anomaly plot on potential temperature

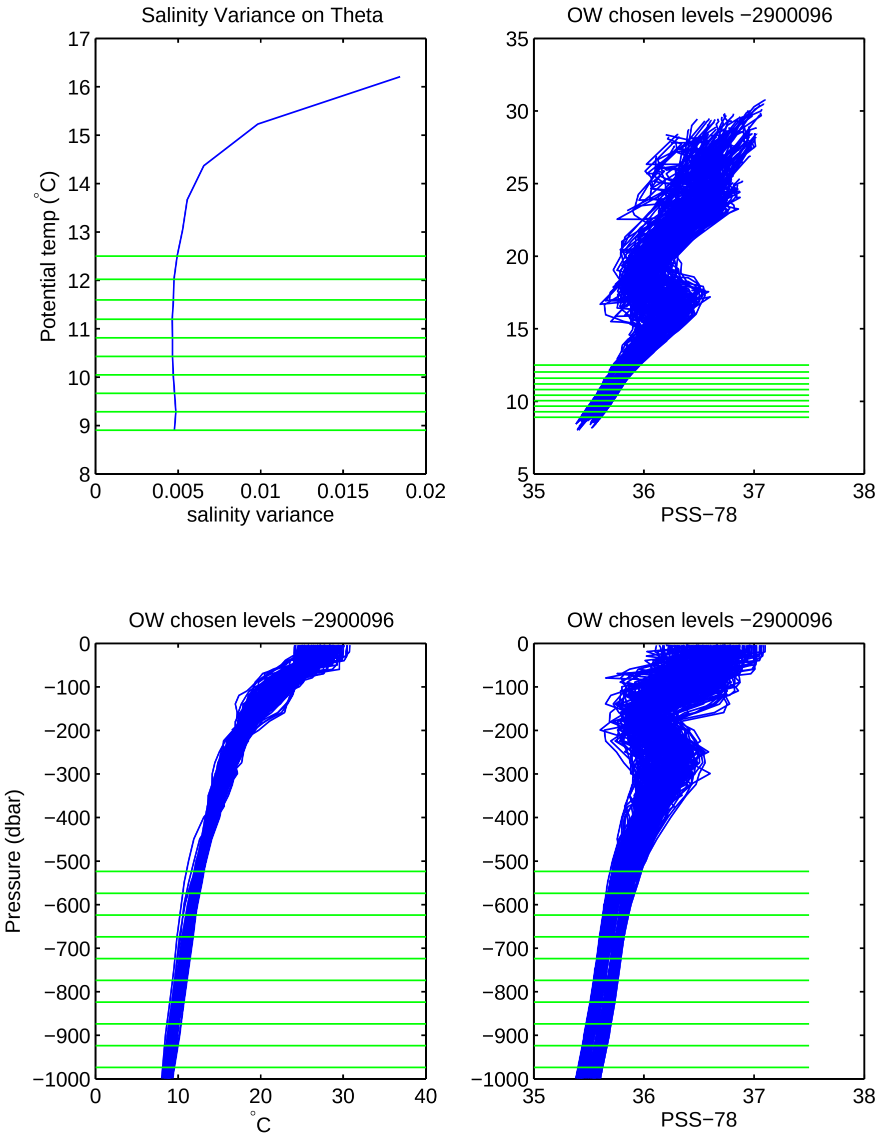

Figure 15: salinity variances on θ levels & Pressure levels. It also shows the 10 float θ levels that are chosen for calibration

$>> load\ \ 2900096$
$>> whos$

Name Size Bytes Class Attributes

DATES 1x205 1640 double

LAT 1x205 1640 double

LONG 1x205 1640 double

PRES 56x205 91840 double

PROFILE_NO 1x205 1640 double

PTMP 56x205 91840 double

SAL 56x205 91840 double

TEMP 56x205 91840 double

WMO_ID 1x1

3. Setting up floats and directories for calibration package.

   – Make directories with names as individual float IDs under data − > float_source, data − > float_mapped, data − > float_calib & data − > float_plots

   – Load matlab workspace with variables: float_names={2900096} and float_dirs={2900096/} & run calibration by invoking "ow_calibration"

4. Actions based on diagnostic graphs

   Graph -I :- The location map of the float with drift over time, embedded in selected background climatology (Figure - 1) ;

   : If the float has drifted through a very large oceanic area with different water masses, then the time series need to be split in to 2 sets.

   : If float has moved near to coast the major and minor axis of the ellipse from which the reference data has been drawn need to be interchanged.

: Care should be taken that the data ellipse covers only areas of ocean basin having similar water mass structure. For example a float near to coast in Arabian Sea should not have reference data drawn from Bay of Bengal as the ellipse may extend to other basin, in such cases limits of the ellipse or the orientation of the ellipse need to be modified.

Graph -II & IV :- If pre-calibration TS diagram has "Bulls eye" levels or profiles, they need to be set to NaNs. Post calibration TS- curve is expected to be tighter than the pre-calibration TS curves. (Figure - 12) . In other cases the float may have samples from spatially different water masses which need to be identified and time series need to be manually split.

Graph -III :- The one and two standard deviations of the float salinity in both conductivity and salinity space is depicted in Figure - 13. The red line represents the mean of the difference between calibrated and un-calibrated salinity. Any abnormal profiles can be located (and need to be verified further for their truth) by detecting events of red line going out of standard deviation envelops.

Graph -V & VII:- These are the pre-calibration and post calibration salinity anomaly plots and the post- calibration anomaly plot is expected to have uniform salinity at deeper part of the ocean where the temporal changes are relatively slow. Where as the pre calibration anomaly plot will have strong gradient in salinity in all levels where the sensor start drifting. (Figure - 14). These two graphs enable one to visualize the correction in total applied to a time series of salinity profiles and redo calibration in case of any abnormal data in post calibration anomaly plot.

Graph -VIII :- Figure - 15 presents salinity variances on θ levels, and pressure levels overlaid with 10 float θ levels that are chosen for calibration. If the calibration is not satisfactory it can be refined by selecting appropriate values for variables *"use_theta_gt", "use_theta_lt", "use_pres_gt", "use_pres_lt", "use_percent_gt" in float_calib/calseries_*.mat .*

5. Writing suitable calibration comments and making final quality controlled D files for upload.
   The calibration comments should reflect the reason for arriving at a conclusion regarding the calibration and it should have the coefficients chosen for arriving at the resulted calibration curve.

Acknowledgments

Director, INCOIS is thanked for the suggested corrections and infrastructural support. Dr. Shailesh Nayak secretary Ministry of Earth Sciences(MOES) is acknowledged for his encouragement and suggestions he made on the first version of this report, during his tenure as Director INCOIS. Dr. Annie Wong from University of Washington, Dr. John Gilson from Scrips Institute of Oceanography and Dr. Brian King from Southampton Oceanographic center are acknowledged for scientific and technical discussions which helped the first author to clarify certain ambiguities regarding calibration details of the floats. National Institute of Ocean Technology (NIOT) Chennai, is acknowledged for float deployments.

References

H. Akima. A new method of interpolation and smooth curve fitting based on local procedures. *J. Assoc. Comput. Mach.*, 17:589–602, 1970.

F. Bretherton, R. Davis, and C. Fandry. A technique for objective analysis and design of oceanographic experiments applied to mode-73. *Deep-Sea Research*, 23:559–582, 1976.

Lars Bhme and Uwe Send. Objective analyses of hydrographic data for referencing profiling float salinities in highly variable environments. *Deep Sea Research Part II: Topical Studies in Oceanography*, 52(3-4):651–664, February 2005. ISSN 0967-0645.

W. Emery and J. Dewar. Mean temperature -salinity, salinity-depth and temperature-depth curves for the north atlantic and the norht pacific. *Progress in Oceanography*, 11:219–305, 1982.

Gregory C. Johnson, John M. Toole, and Nordeen G. Larson. Sensor corrections for sea-bird sbe-41cp and sbe-41 ctds. *Journal of Atmospheric and Oceanic Technology*, 24(6):1117–1130, 2007.

S. Joseph and H.J. Freeland. Salinity variability in the arabian sea. *Geophysical Research Letters*, 32(9):1–4, 2005.

T. Kobayashi and T. Suga. The indian ocean hydrobase: A high-quality climatological dataset for the indian ocean. *Progress in Oceanography*, 68(1):75–114, 2006.

P. McIntosh. Oceanographic data interpolation: Objective analysis and splines. *Journal of Geophysical Research*, 95(C8): 13529–13541, 1990.

W. Brechner Owens and Annie P.S. Wong. An improved calibration method for the drift of the conductivity sensor on autonomous ctd profiling floats by [theta]-s climatology. *Deep Sea Research Part I: Oceanographic Research Papers*, 56(3): 450–457, March 2009. ISSN 0967-0637.

P. S. Wong, Annie, C. Johnson, Gregory, and Owens Brechner W. Delayed-mode calibration of autonomous ctd profiling float salinity data by $\theta - s$ climatology. *Journal of Atmospheric and Oceanic Technology*, 20:308–318, 2001.