Technical Report **Report No. :** ESSO-INCOIS-ODG-TPG-TR-02(2020)

Quality Control of in situ T/S data using α convex hulls

By

R Venkat Shesu, TVS Udaya Bhaskar, E Pattabhi Rama Rao
M. Ravichandran, B. Venkateswara Rao

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

June, 2020

DOCUMENT CONTROL SHEET

Earth System Science Organization (ESSO)
Ministry of Earth Sciences (MoES)
Indian National Centre for Ocean Information Services (INCOIS)

ESSO Document Number: ESSO-INCOIS-ODG-TPG-TR-02(2020)

Title of the report: Quality Control of in situ T/S data using α convex hulls.

Author(s) [Last name, First name]:
R Venkat Shesu, TVS Udaya Bhaskar, E Pattabhi Rama Rao, M. Ravichandran,
B. Venkateswara Rao.

Originating unit
Ocean observations and Data Management Group (ODG), INCOIS

Type of Document:
Technical Report (TR)

Number of pages and figures: 20 and 8.

Number of references: 33

Keywords:
in situ data, Quality Control, Outliers, α Shape, Convex Hulls, Point in Polygon.

Security classification:
Open

Distribution:
Open

Date of publication:
29 April, 2020

Abstract (100 words)
An outlier is an observation that is abnormal compared to its neighbors and lies at an abnormal distance from other values in a random sample from a population. Spatial outliers are objects with distinct features from their surrounding neighbors in space. Detection of spatial outliers helps reveal important and valuable information from large spatial data sets. In oceanographic data, outliers are frequently represented in region, i.e., a group of observations from an instrument which is malfunctioning or

from a float whose sensors have degraded over period after giving valuable data initially. This work discusses a new method of identifying anomalous oceanic in-situ data particularly temperature and salinity (T/S). The proposed method uses gridded fields of T/S obtained from World Ocean Atlas 2013 (WOA13) climatology on 0.25° x 0.25° scale and construct α convex hulls that are used for classifying in-situ data as good or bad. An 'n' sided polygon (convex hull) with least area encompassing all the points is constructed based on the Jarvis March algorithm. Convex hulls are built based on optimal α values fixed, such that the area encompassed by a convex hull is least and either minimal or no points of T/S are left out. The final α value used for building the convex hulls (monthly, seasonal and annual) is arrived after iterative process which are also passed through visual inspections. Extensive sensitivity experiments were done for arriving at the optimal α value such that false positives and true negatives are minimized. It is observed that various types of anomalies in in-situ data viz., spikes, bias, sensor drifts etc, can be identified using this method and classified as good/bad, automatically. The merit of the method lies in its ability to identify majority of outliers in huge oceanographic data automatically with minimal or no human intervention yielding quality in-situ data for research in oceanography.

Table of Contents

Abstract

An outlier is an observation that is abnormal compared to its neighbors and lies at an abnormal distance from other values in a random sample from a population. Spatial outliers are objects with distinct features from their surrounding neighbors in space. Detection of spatial outliers helps reveal important and valuable information from large spatial data sets. In oceanographic data, outliers are frequently represented in region, i.e., a group of observations from an instrument which is malfunctioning or from a float whose sensors have degraded over period after giving valuable data initially. This work discusses a new method of identifying anomalous oceanic in-situ data particularly temperature and salinity (T/S). The proposed method uses gridded fields of T/S obtained from World Ocean Atlas 2013 (WOA13) climatology on 0.25° x 0.25° scale and construct α convex hulls that are used for classifying in-situ data as good or bad. An 'n' sided polygon (convex hull) with least area encompassing all the points is constructed based on the Jarvis March algorithm. Convex hulls are built based on optimal α values fixed, such that the area encompassed by a convex hull is least and either minimal or no points of T/S are left out. The final α value used for building the convex hulls (monthly, seasonal and annual) is arrived after iterative process which are also passed through visual inspections. Extensive sensitivity experiments were done for arriving at the optimal α value such that false positives and true negatives are minimized. It is observed that various types of anomalies in in-situ data viz., spikes, bias, sensor drifts etc, can be identified using this method and classified as good/bad, automatically. The merit of the method lies in its ability to identify majority of outliers in huge oceanographic data automatically with minimal or no human intervention yielding quality in-situ data for research in oceanography.

1. Introduction

Enormous economic and social value is exhibited by the oceans that surround our earth. Hence there is a need for continuous monitoring through various mean viz., in situ and satellite observations. Ocean observations are of paramount importance for many applications including the monitoring of climate and the environment on seasonal-to-interannual-to-decadal time scale. There are many benefits that can be derived from these observations. The benefits from observations include increased efficiency of operations at sea, improved safety to personnel and reduced damage to the environment. In particular, quality weather and ocean state forecasts demands the availability of operational ocean observations. In spite of its increasing importance and pervasive impact, very little is known and understood about the ocean.

Since long, many observational platforms are devised to map the ocean. Number of instruments and techniques are used to gather observational data from the ocean (https://www.ncei.noaa.gov/news/world-ocean-database-profiles-ocean). These include:

- CTD—a rosette of instruments to measure conductivity, temperature, and depth
- Moored buoys—anchored platforms at fixed locations
- Drifting buoys—platforms subject to ocean currents
- Profiling floats—platforms drifting at predetermined subsurface water levels
- Plankton tows—nets or bottles that are cast from a ship
- Undulating oceanographic recorders—instruments towed behind vessels
- Gliders—low-powered autonomous underwater vehicles (AUVs)
- Autonomous pinniped bathythermographs—instruments attached to elephant seals
- Expendable Bathythermographs (XBT)—probe used once and dropped through water column with small wires to transmit temperature back to ship

The number of observations collected over the global ocean (centrally archived at NCEI, USA) is shown in the Figure 1. Starting from the Bottle data, Mechanical Bathy Thermographs (MBT) to the recent Argo floats, there has been a continuous evolution in the observational systems. The eXpendable Bathy Thermographs (XBT) were first deployed around in 1966 and they replaced MBTs in most ocean observations measurement programs. XBT probes are torpedo-shaped devices attached to a spool of copper wire. The instrument can be launched over the side of a moving ship, from an airplane, or from a submarine. Temperature is estimated by measurements of the resistance in a semi-conductor (called a thermistor). After recording, the information is sent back to the command unit over the copper wire. Depth is calculated as a function of time since launch using a manufacturer-supplied equation. Most of the XBT data are collected along the ships of opportunity in operational mode to provide a continuous record of temperature profile data along repeated transects, now known as the Global XBT Network (Goni et al, 2019).

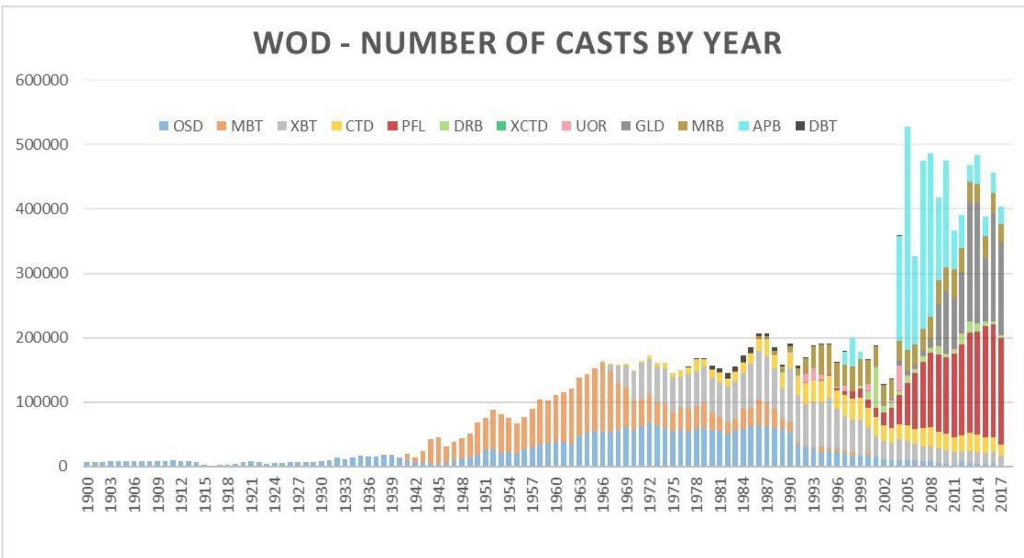


Figure 1. Oceanographic data archived in the World Ocean Database (WOD). Courtesy of NOAA/NCEI World Ocean Database.

A CTD is a device whose primary function is to detect how the conductivity and temperature of the water column changes relative to depth it is being lowered. By the help of conductivity of seawater, salinity can be derived using the temperature and pressure at the same water depth. The depth is obtained from the pressure measurement by the way of calculating the density of water from the temperature and the salinity. CTDs are attached to larger metal frame named rosette, which hold water-sampling bottles that are used to collect water at different depths. Knowledge obtained from CTD devices can provide a more detailed understanding of the ocean water's characteristic through the entire water column, which is crucial for understanding the physics involved.

The XBTs are abundant but is confined along ships of opportunity lines. CTDs are of high resolution and quality and gather data up to deeper depth than XBT, but are observed only along a specified scientific cruises. To overcome this an autonomous method of observing the ocean through profiling floats named Argo was devised. Argo is an international program aimed at deploying 3000 instruments in the global ocean. Two primary types of floats missions have evolved since the initial float design as described in Davis et al., (1992). In the first mission, the float sinks to a prescribed pressure (typically 1000 dbars) and cycles with a prescribed time period (typically 10 days) and in the second, the floats are programmed to perform drifting and profiling at different depths. In November 2007, the global Argo array of profiling floats has reached its initial target of 3000 operating floats worldwide. The array now provides for the first time a global monitoring of ocean temperature and salinity (T/S) data in real time (Gould et al. 2004).

The primary objectives of in situ observations are 1) to provide the continuous changing properties of upper-ocean, 2) initialization of ocean and coupled forecast models, and 3) provide input to global ocean analyses. These in situ data sets are used in studies related to climate change, sea level rise, Ocean Heat Content (Nagamani et al., 2016) and its relation to cyclone intensity (Jangir et al., 2016), mixed layer processes (Udaya Bhaskar et al, 2007), assimilation into various ocean models, generation of monthly gridded products (Udaya Bhaskar et al., 2007) and in generating improved climatologies (Chatterjee et al., 2012). The MBTs are limited to 300 meters depth with a accuracy of 0.1 °C. The XBT probes can observe up to 400, 800 or 1500 meters depth depending on type of probe with a accuracy of 0.2 °C. With CTD profiles of temperature and salinity can be measured up to the bottom. The accuracy of temperature from CTD is 0.005 °C, while the accuracy of conductivity is 0.003 S/m. All these studies warrant that the data from these instruments viz., MBT, XBT, CTD, Argo floats etc be of high quality as the results might be biased owing to instrument errors. All these errors are to be recognized and eliminated or flagged before the data is put to use.

Data from XBT though frequent in time, is confined to shipping routes. CTD data are confined to specific scientific programs. It is only after the beginning of Argo program, geographic distribution of oceanographic T/S profile data has become more uniform over the global ocean. Oceanographic data is widely used in operational running of the ocean models and preparation of climatologies. Different countries deploy different type of ocean observation platforms. XBTs are manufactured by Sippican, USA and TSK, Japan. CTD are mounted with sensors from SeaBird, USA. Argo floats deployed by various countries are manufactured by different manufacturers (Viz., Teledyne, USA; NKE France, Ninja, Japan; SOLO, Scripps Institute of Oceanography among others) who use different types of pressure sensors (Druck, Amtek, Paine etc) and CTD sensors (SeaBird, RBR, FSI etc). Each country has the privilege of setting the T/S profile resolution as per their requirement. With the involvement of different measurement methods, difference in instrumentation and differences in data handling, there can be ample scope for some bad data getting unnoticed. All these data sets have sets of quality control (QC) procedures prescribed by the groups that are extensively working with them. The XBT data is quality controlled using the procedures prescribed in CSIRO cookbook (Bailey et al., 1994). The CTD data is quality controlled using IOC Manual and guides No. 26. For the data from Argo floats, the Argo Data Management Team (ADMT) has proposed 19 quality control steps (Wong et al., 2020). Even after availability of all these quality control methods, various organizations/institutions employ additional QC procedures for performing additional quality checks on these data. With all these efforts, there can be still some scope for existence of erroneous data. As it would be a cumbersome process huge amount of data individually and pin point these bad data, various methods of handling these erroneous data are developed.

Some additional QC methods prescribed to handle the in situ data sets include visual quality control for XBT data (Thadathil et al., 2001), T/S profiles from CTD following UNESCO manuals and guide 26 (1993). In the case of Argo floats, apart from the QC prescribed by ADMT, a secondary quality control methods based on scientific analysis called Delayed Mode Quality Control (DMQC) is also applied following Wong et al.,

(2003) and Owens et al., (2009). There are also some independent methods of QC envisaged by various operational agency involved in processing and archival of ocean in situ data. Guinehut et al., (2009) proposed a quality control method based on satellite altimetry for identifying the salinity sensor degradation. Cabanes et al., (2013) described the methods of quality control of all in situ ocean observation data archived at Coriolis Centre. A three way QC method was proposed by Udaya Bhaskar et al., (2012) to perform QC of heterogeneous data archived at INCOIS including the visual quality control (Udaya Bhaskar et al., (2013)). Still some pros and cons exists with most of these methods. The current work describes a new method proposed to augment the existing procedures, where in the entire temperature and salinity profile can be treated for its quality, instead of using individual observations/records.

Outliers are inherent to any type of datasets. An outlier is an in situ data or observation that is quite unusual when compared to the points in its neighbors (both spatially and temporally) and lies at an anomalous distance from other values, from a random sample derived from a population. Outliers in spatial data are those objects which have distinct features (eg: bull eye like structure) when compared to their neighbors in immediate surroundings. These spatial outliers sometimes helps in revealing important and valuable information from a set of large spatial data. Natural events like cyclones, Tsunamis, storm surges etc can result in spatial outliers. Data from faulty sensors owing to malfunctioning, degradation over a period of time, lacking calibrations etc can also result in outliers in in-situ data sets.

2. Data and Methods

For this work data from various ocean observation platforms are employed. XBT and XCTD data were obtained from odis.incois.gov.in web site (Shesu et al., 2013). CTD data were obtained from different projects funded by Ministry of Earth Science. Argo T/S profiles data are downloaded from the Argo data viewer developed at INCOIS (Geetha et al., 2011). Details of the data sets archived at INCOIS were given in table 1. All the direct raw observation of XBT, XCTD and CTD were taken, while for Argo floats, data sets that have been passed through the real-time quality control procedures as proposed by ADMT were considered.

Table1: List of sub-surface T/S profiles archived at INCOIS.

S.No	WMOID	First observation	Last observation
1.	XCTD	2009	Till date
2.	Argo (Indian & other countries)	2001	Till date
3.	CTD (Sagar Kanya, Sagar Nidhi, Sagar Sampadha	1999	
4.	XBT	1990	Till date

2.1 α Shapes, convex hulls

In computational geometry, α-shape, is a family of piecewise linear simple curves in the Euclidean plane associated with the shape of a finite set of points. The concept of alpha shape was first defined by Edelsbrunner, Kirkpatrick & Seidel (1983). The alpha-shape associated with a set of points is a generalization of the concept of the convex hull, i.e. every convex hull is an alpha-shape but not every alpha shape is a convex hull. Figure 2 gives an illustration of convex hull, α shape and minimal spanning tree. When α is zero, the shape attains the form of minimal spanning tree and when α is infinity the shape turns out to be a convex hull. One need to fine tune the α such that the resultant figure is neither convex hull nor minimal spanning tree. With this understanding α is fine tuned and closed polygons were built based on the observations of (i) Temperature Vs Depth, (ii) Salinity Vs Depth and (iii) Temperature Vs Salinity. Once the α shapes are built they are used for detecting and eliminating anomalous profiles either partly of fully.

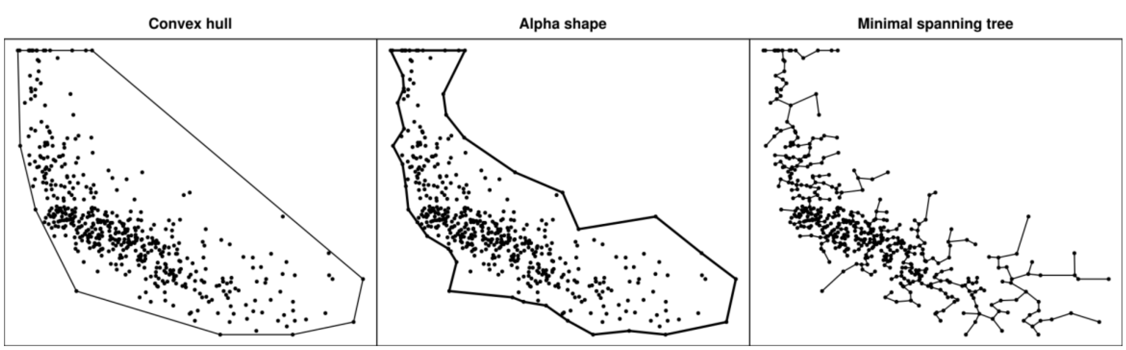


Figure 2. Convex hulls, Alpha shapes and minimal spanning trees associated with a set of points.

The T/S data from World Ocean Atlas 2013 on 0.25° x 0.25° resolution was taken and the α shapes of T/S were build for each month, season and annual scale. An 'n' sided polygon (α shape) with least area encompassing all the points is constructed using the Jarvis March algorithm (Jarvis, 1973). The α convex hulls is built based on an optimal α values fixed, such that the area encompassed by the convex hull is least and either minimal or no points of T/S are left out during its construction. The final α value used for building the convex hulls (monthly, seasonal and annual) is arrived after an exhaustive iterative process which are also passed through visual inspections to minimize the points being left out of the polygon. Extensive sensitivity experiments were done for arriving at the optimal α value such that false positives and true negatives are minimized. Figure 3 shows the sensitivity experiments done with α and the resulting points falling out of α shape.

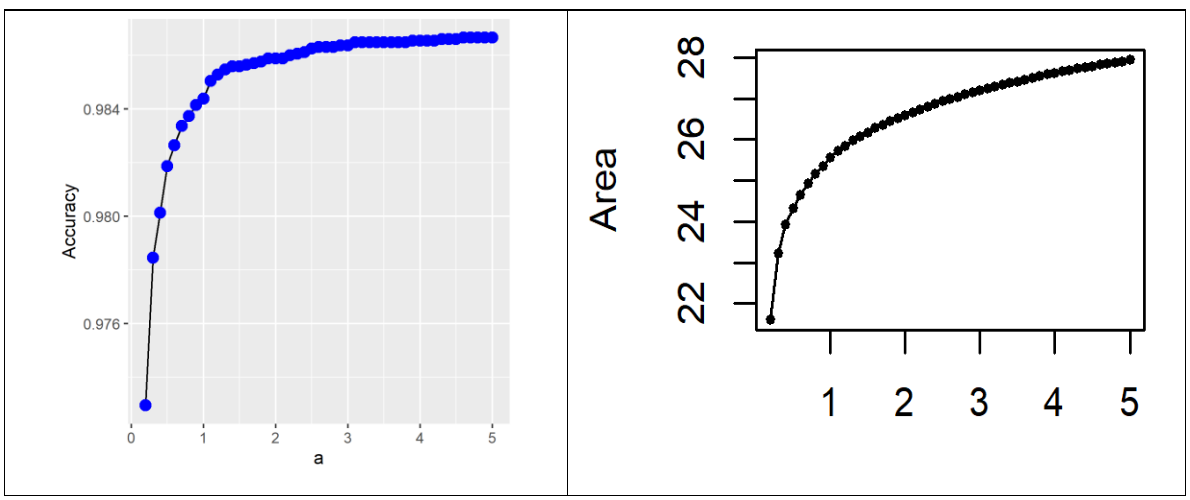


Figure 3. (a) Number of points falling out of α shape with increasing values of α. (b) Area of the α shape with increasing value of α.

From the figure one can observe that as α increases, the number of points falling out or left out are decreased. After certain value of α, there seems to be no changes in the number of points left out and the curve looked more or less flat. Using such sensitivity experiments α values were derived for all the months, seasons and annual data sets. These α values are then used to construct the α convex hulls which were eventually used to detect anomalous T/S profiles obtained from various in situ sources. The methods for classification of the good and bad profiles is similar to what is proposed by Udaya Bhaskar et al (2017) and as described below.

2.2 Implementation methodology

The principle of α convex hull and Point In Polygon (PIP) implemented using Ray Casting Algorithm (Shimrat 1962) are used to identify good vs bad T/S profiles. The steps for application of the proposed method is as follows:

1. The trajectory of observed Argo/XBT/CTD temperature and salinity profiles are obtained and region of interest (ROI) is obtained.
2. Use World Ocean Atlas (2013) temperature and salinity profile data corresponding to this ROI and build a T/S α convex hull with least area encompassing the mean and 2*standard deviation fields of temperature and salinity profile fields.
3. Subsequently the PIP algorithm is used to check if the observed temperature and salinity profile (obtained in step 1) falls within or outside this climatological T/S α convex hull.
4. Set the quality flags as good(bad) for data falling within(outside) the α convex hull there by identifying erroreneous profile data as shown in Figure 4.

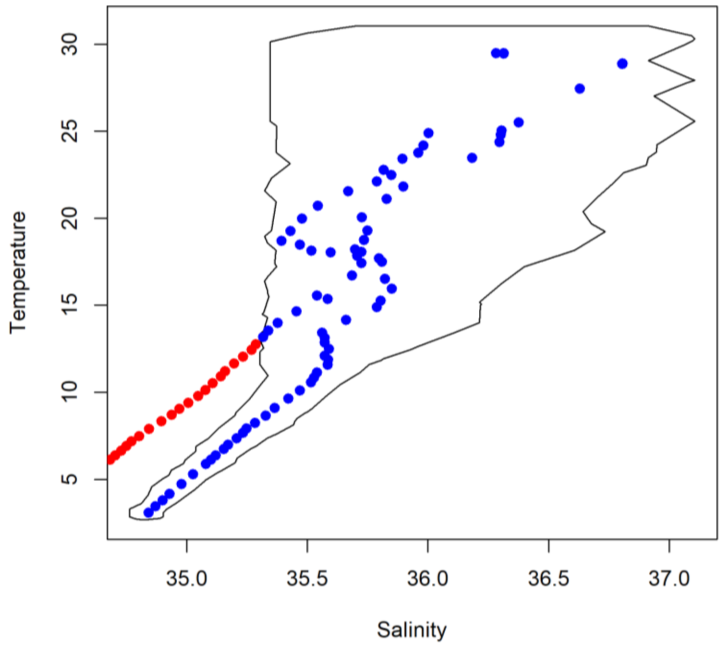


Figure 4. α shape (convex hull) of annual T/S profiles from WOA13. Good profile falling with in α shape is shown in good color and anomalous profile falling out side α shape is shown in red color.

The climatology used for the proposed method is the gridded fields of temperature (Locarnini et al., 2013) and salinity (Zweng et al., 2013) and their standard deviation fields obtained from World Ocean Atlas 2013 (WOA13) of US National Oceanographic Data Centre (NODC). These climatological mean and standard deviation fields are then used to build polygons (α convex hulls) of Temperature Vs Salinity, Temperature Vs Depth and Salinity Vs Depth. The process of interpolating T/S profiles in situ data to standard depths was eliminated, as full length profiles are considered instead of standard depth as proposed in earlier work (Udaya Bhaskar et al., 2017). Jarvis March (1973) algorithm which is also popularly called as gift wrapping algorithm was employed for building these α convex hulls corresponding to months, seasons and annual. The T/S profiles from in situ platforms are checked to see if they are falling within the corresponding α convex hull obtained from WOA13 climatology for that corresponding month, season and annual. PIP algorithm was employed to see if the profile data is falling inside or outside. As with the previous method, this method also has a complexity of O(nh) where *n* is the number of points and *h* is the number of points on the α convex hull.

For more details on how the algorithm was used for building the α convex hull for individual standard depths, kindly refer Udaya Bhaskar et al., (2017).

3. Validation of the proposed method

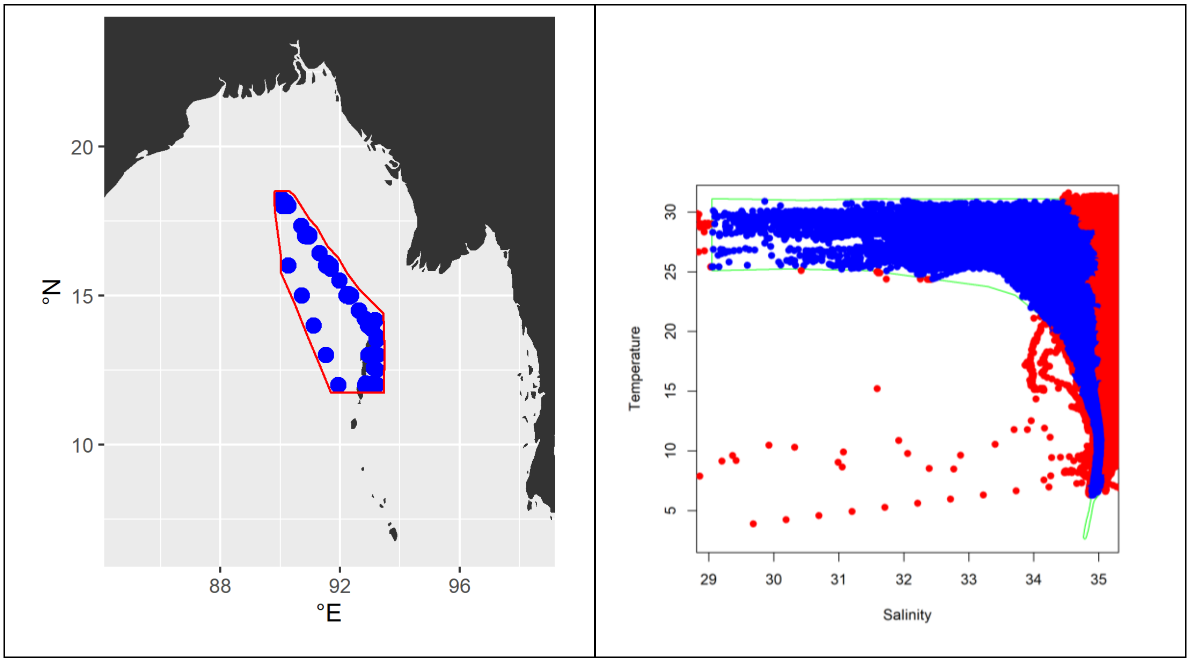
In most cases the sensors which measure pressure and temperature are robust and it is the salinity sensors is vulnerable to changes and often degrade due to bio-fouling (Eg: Argo floats; Wong et al, 2003). Some often encountered problems with salinity sensors mounted on Argo floats are offsets (both fresh and saltier), Tri-Butyl Tin Oxide (TBTO) problem which generally result in freshening, salty drifts after set of cycles etc. Unlike in earlier study by Udaya Bhaskar et al., (2017) the quality of both T/S profiles were checked using the α convex hulls in a single go. The trajectory of Argo float, XBT/XCTD along ships of opportunity and CTD from a scientific cruises were all used and the region of interest (ROI) is obtained. All the climatological T/S profiles corresponding to these ROI are extracted from WOA13 and α convex hulls of Temperature Vs Salinity, Temperature Vs Depth and Salinity Vs Depth are constructed. The observed profiles from Argo floats, XBT/CTD and CTD are overlaid on these α convex hulls. PIP algorithm is used to check if the observed profiles fall outside the α convex hull or not. Based on the data falling outside, the instruments are suspected to have a problem (offset, drift, bias, spike etc). Because the climatology is generated using large number of observations spanning decades, a instrument measured profile is suspected to have a problem, if it fall outside the n-sided α convex hull which is constructed with a mean ± 2*standard deviations.

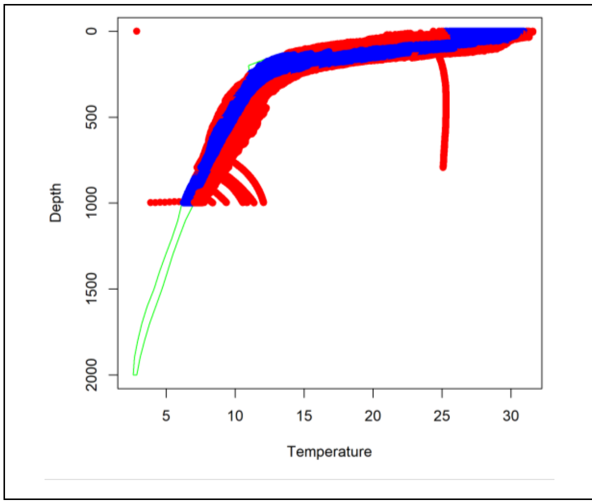
To demonstrate the robustness of the proposed method, 4 typical cases are chosen from XCTD, Argo, CTD from cruise and XBT which represent different problems. The details of the data chosen for the validation are given in the table 2.

S. No	WMOID	First observation	Last observation	Total observations	Total Number of Record	Spurious Records	Percentage of good/bad

					s		
1.	XCTD	08/02/2009	24/02/2015	801	625284	477042	24% / 76%
2.	Argo	25/12/2011	19/12/2014	110	17764	909	95% / 5%
3.	CTD	19/09/2000	17/10/2000	34	51228	536	99% / 1%
4.	XBT	01/01/2010	31/12/2014	466	313141	17614	95% / 5%

Table 2. Details of the ocean observations chosen for validation of α convex hull.





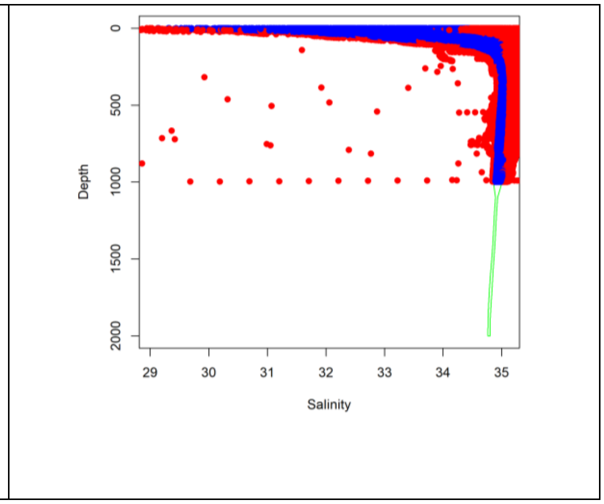


Figure 5. (a) Region of Interest of XCTD data chosen. (b) Polygon with green lines indicate α convex hull based on annual climatological T/S data. (c) Polygon with green lines indicate α convex hull based on annual climatological Temperature Vs Depth data. (d) Polygon with green lines indicate α convex hull based on annual climatological Salinity Vs Depth data. Blue data indicate points within α convex hull and red data indicate points outside it.

Figure 5a shows the locations of XCTD data collected in Bay of Bengal. The climatological T/S profiles falling with in region of interest (figure 5a) are obtained from annual WOA13 climatology and α convex hull is built. The measured XCTD profiles are then checked for their quality using α convex hull. Point-in-Polygon algorithm is used to check good and bad T/S profiles. All the good T/S profiles falling inside the polygon are indicated in blue color the bad data falling outside the polygon are indicated in red color. As can be seen many anomalous profiles can be observed to fall outside the convex hull indicating as bad which can be out rightly eliminated. The details of the spurious records are given in the table 2. This method nearly rejected 76% of total temperature measurements from the XCTD.

Figure 6a shows the trajectory of an Argo float. This float profiles are observed to have drift in the salinity sensor only for few cycles from 16th - 22nd. This type of drift in salinity can sometime happen due to presence of worm into the conductivity pipe. With the worm getting washed out the salinity sensor tends to come back to normalcy and the profiles are observed to good. These erroreneous profiles are clearly marked as bad as they are observed to fall outside the Temperature Vs Salinity α convex hull and depicted in red color in Figure 6b. The floats observations are found to come back to normalcy after cycle 22. The bad salinity profiles are clearly observed in Salinity Vs Depth α convex hull. Few outliers are seen in Temperature Vs Depth α convex hull. The details of the spurious records rejected by the method are given in the table 2. This method nearly rejected 5% of total temperature and salinity measurements by the Argo float.

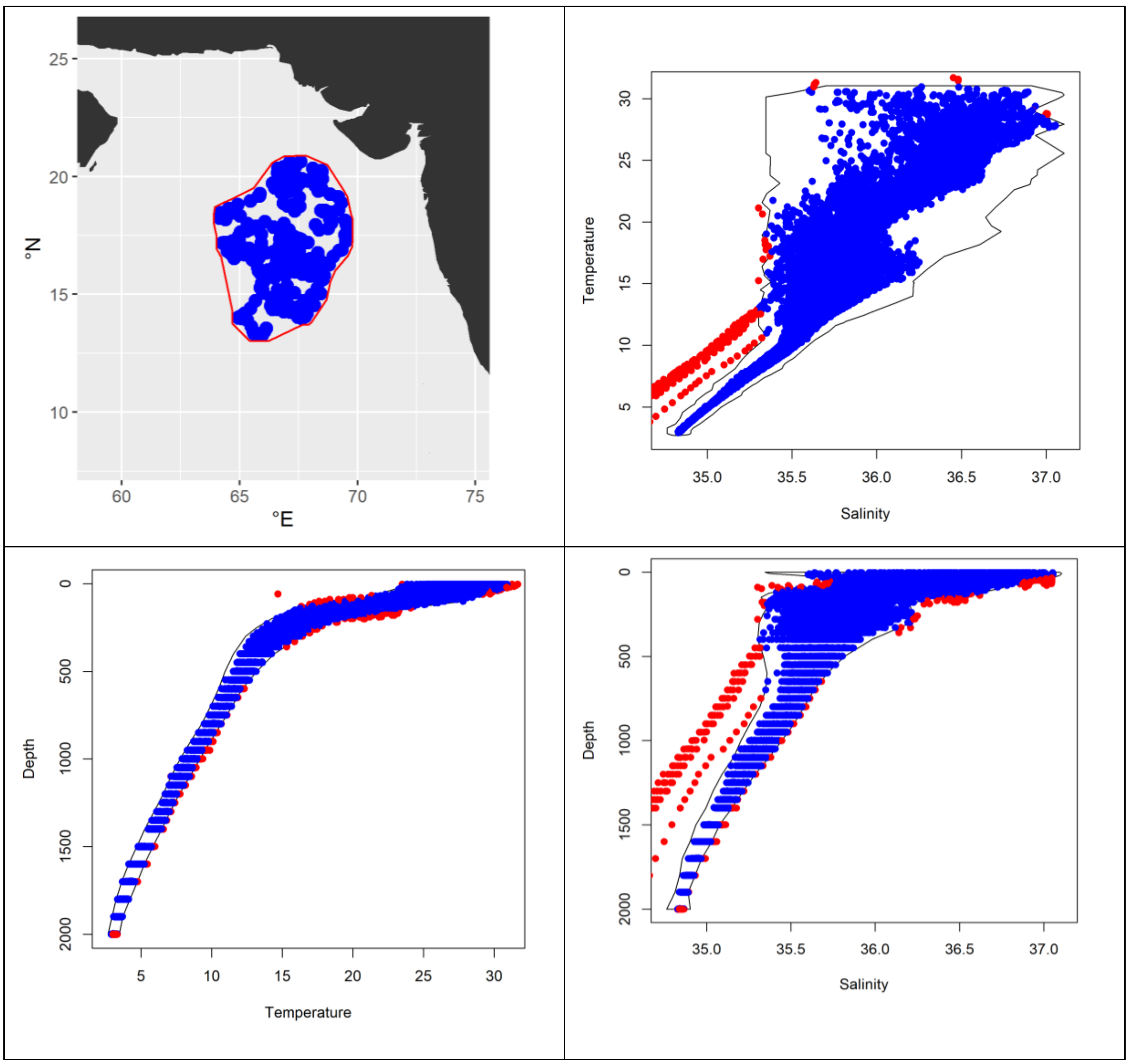


Figure. 6. (a) Trajectory of Argo float with WMOID 2901340 during its life time. (b) Polygon with black lines indicate α convex hull based on annual T/S data. (c) Polygon with green lines indicate α convex hull based on annual T Vs Depth data. (d) Polygon

with green lines indicate α convex hull based on annual S Vs Depth data. Blue data indicate points within convex hull and red data indicate points outside the polygon.

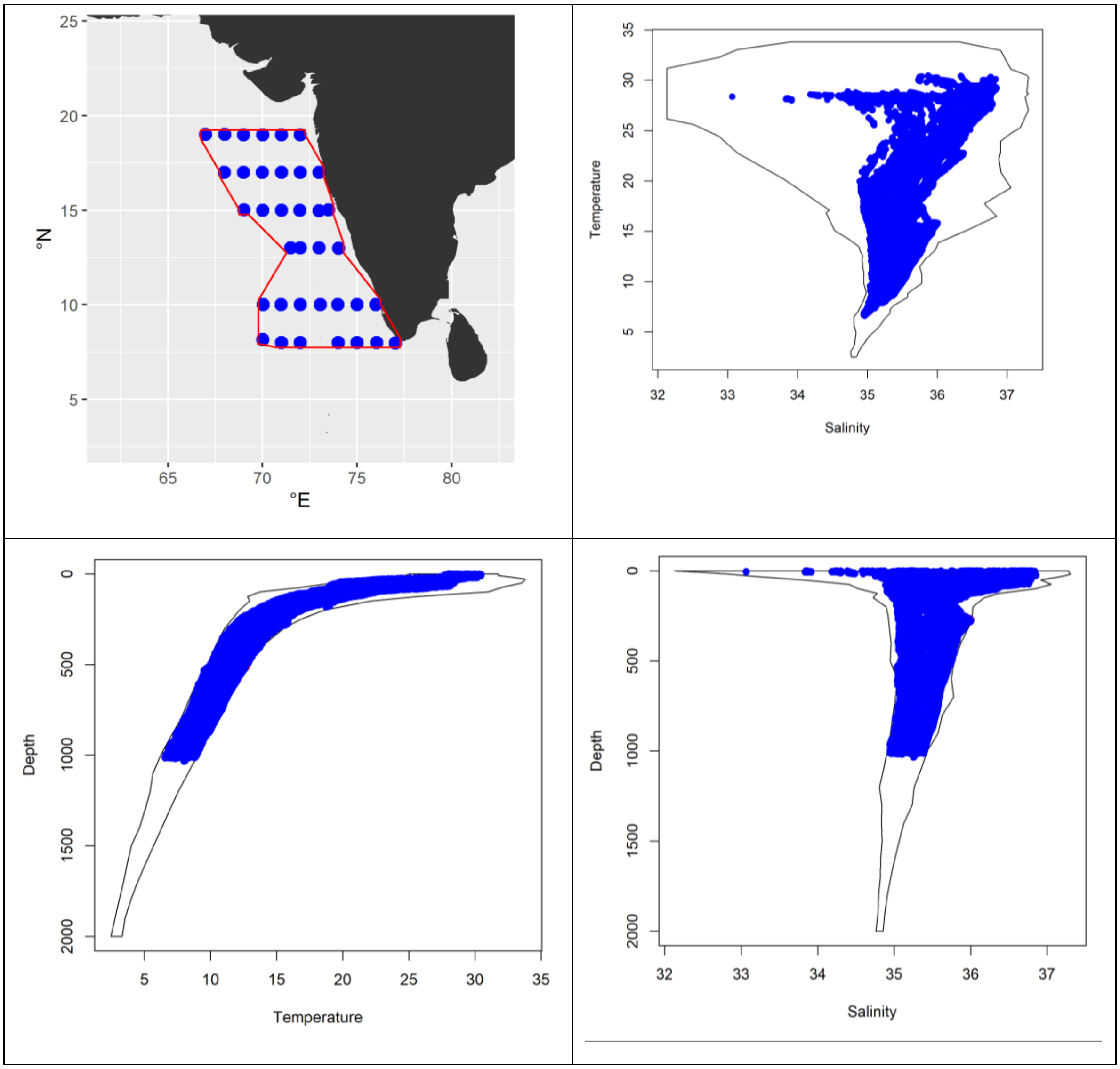


Figure. 7. (a) Location of CTD collected during cruise by FORV Sagar Sampadha. (b) Polygon with black lines indicate α convex hull based on annual climatological T/S data. (c) Polygon with black lines indicate α convex hull based on annual climatological T Vs Depth data. (d) Polygon with green lines indicate α convex hull based on annual climatological S Vs Depth data. Blue data indicate points within α convex hull and red data indicate points outside it.

Figure 7a shows the locations of CTD data collected in Arabian Sea onboard Fisheries Ocean Research Vessel, Sagar Sampadha. Data along different transects are collected starting from 19th Sep, 2000 till 17th Oct, 2000 during cruise SS-188. The climatological T/S profiles encompassing the region of CTD data collected (figure 7a) are obtained from annual WOA13 climatology and α convex hulls are built. The measured CTD profiles are then checked for their quality using these α convex hulls. One can observe that all the measured T/S profiles lie inside the polygon indicating that the observations are all of good quality. This is the biggest advantage of the proposed method as the quality of the entire lot of data is spelt out in a single go without the tedious process of checking each and every profile individually. The details of the spurious records rejected by the method are given in the table 2. This method rejected as lows as 1% of total temperature and salinity measurements from the CTD casts. This proves the robustness of the data from CTD.

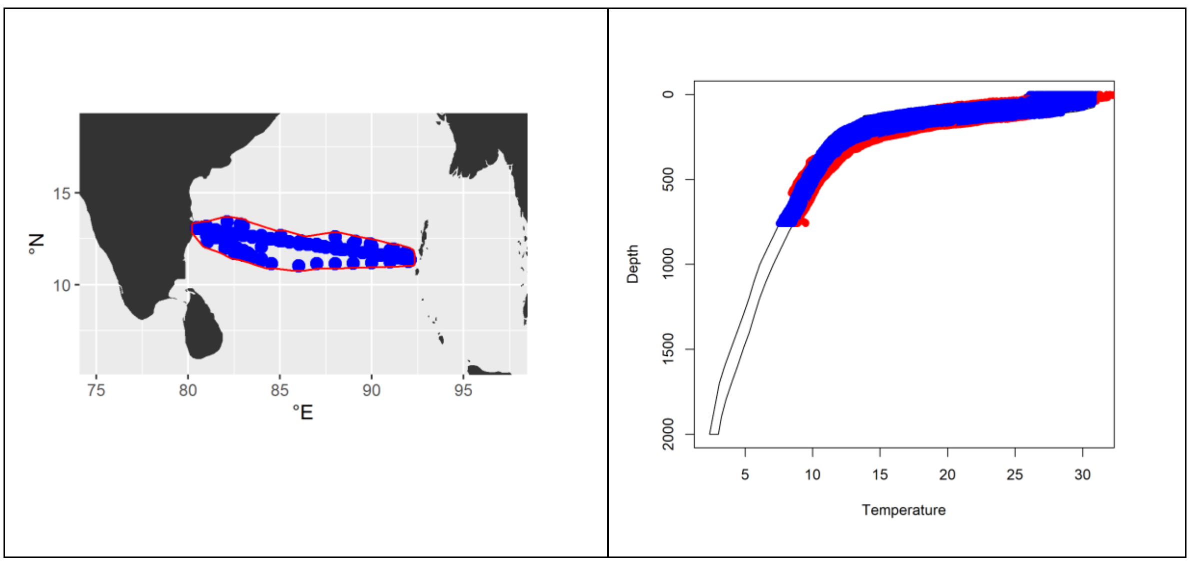


Figure. 8. (a) Location of XBT collected during various cruise from Chennai to Port Blair. (b) Polygon with black lines indicate α convex hull based on annual climatological T Vs Depth data. Blue data indicate points within α convex hull and red data indicate points outside it.

Figure 8a shows the locations of XBT data collected in Bay of Bengal fortnightly along routine shipping line from Chennai to Port Blair. Data along different transects collected during from 19th Sep, 2000 till 17th Oct, 2000 is used for checking their quality. Unlike other examples, here we cannot use the T/S α convex hull as XBT does not measure salinity data. The climatological Temperature profiles encompassing the region of XBT data collected (figure 8a) are obtained from annual WOA13 climatology and α convex hulls of Temperature Vs Depth are built. The measured XBT profiles are then checked for their quality using these α convex hulls. Few temperature profiles are observed to lie outside the polygon indicating that the observations are erroreneous. This shows that the method can be generalized for different instruments which measure temperature alone and both T and S. The details of the spurious records rejected by the method are given in the table 2. The details of the spurious records rejected by the method are given in the table 2. This method nearly rejected 5% of total temperature measurements by XBT probes.

4. Discussions and Conclusions

An improved method for performing quality control of different types of in situ data from observational platforms is proposed. The method is an improvement to the method proposed by Udaya Bhaskar et al., (2017) where in convex hulls are built using climatological data for individual standard depths. To use these climatological convex hulls corresponding to standard depths, all the in situ data are to be interpolated to these standard depth and compared. Interpolations might introduce some errors and also they go un-noticed if any spike or bad data in the profiles are observed between these standard depths. To overcome this, it is proposed to build α convex hulls of whole profile and check the quality of observed temperature and salinity profiles against them. Accordingly α convex hulls of Temperature Vs Salinity, Temperature Vs Depth and Salinity Vs Depth are built using World Ocean Atlas 2013 data and the observed profiles from instruments like XBT, XCTD, CTD and Argo floats are checked for their quality. As in Udaya Bhaskar et al., (2017) Jarvish March (1973) and Point In Polygon algorithms are used for building the α convex hulls and checking for profile falling inside or outside. Most of the issues with observational data like spike, drift, offsets etc, associated with the ocean instruments are identified. For the sample data tested it is observed that outlier are more for XCTD and least for CTD with percentage of rejection being 76%, 5%, 1% and 5% for XCTD, Argo, CTD and XBT respectively. As data from CTD is known to be more reliable, the low return of outliers suggests that the method is robust and can be used for detection of outlier successfully. As the number of data sets from ocean observational platforms are increasing, this method serves to treat all of them in bulk and return only those profiles which are to be checked manually. As manual checking of all the individual profiles is tedious, this method can be used to obtain set of anomalous data which can be further tested visually by experts in the field. For better results, this methods can be augmented with other methods in use like altimetry based QC and objectively analyzed based QC. This method also urges updates to the reference climatology as it forms the backbone of the method.

Acknowledgements

The authors thank the Director, INCOIS for encouragement and for providing necessary facilities to carry out the work. Argo data is made freely available by the Argo community.

References

Akima, H (1970) A new method of interpolation and smooth curve fitting based on local procedures, Journal of Association of Computers and Machines, 17, 589-602.

Jangir, B., Swain, D., Udaya Bhaskar, TVS (2016) Relation between Tropical Cyclone Heat Potential and Cyclone intensity in the North Indian Ocean, ***SPIE***, Vol 9882.

Badouel, Didier (1990) An Efficient Ray-Polygon Intersection, Graphics Gems (Andrew S. Glassner, ed.), Academic Press, pp. 390-393.

Bailey, R., A. Gronell, H. Phillips, E. Tanner, and G. Meyers (1994) Quality Control Cookbook for XBT Data (Expendable Bathythermograph Data). Version 1.1. Australia, CSIRO, 37pp. (CSIRO Marine Laboratories Report: 221). http://hdl.handle.net/11329/127.

Cabanes, C., A. Grouazel, K. von Schuckmann, M. Hamon, V. Turpin, C. Coatanoan, F. Paris, S. Guinehut, C. Boone, N. Ferry, C. de Boyer Montégut, T. Carval, G. Reverdin, S. Pouliquen, and P. Y. Le Traon (2013) The CORA dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements, Ocean Science, 9, 1-1.

Chatterjee, A., D. Shankar, S.S.C Shenoi, G.V. Reddy, G.S Michael, M. Ravichandran, V.V. Gopalkrishna, E.P. Rama Rao, T.V.S Udaya Bhaskar and V.N. Sanjeevan (2012) A new atlas of temperature and salinity for the North Indian Ocean, J. Earth Syst. Sci. 121, No. 3, June 2012, pp. 559–593.

Chong-Wei Huang and Tian-Yuan Shih (1997) On the complexity of point-in-polygon algorithms, Computers and Geosciences, Vol 23 (1), pp 109 - 118.

Davis, R.E., D.C. Webb, L.A. Regier, and J Dufour (1992) The Autonomous Lagrangian Current Explorer, Journal of Atmospheric and Oceanic Technology, 9, 264 - 285.

Edelsbrunner H, Kirkpatrick DG, Seidel R (1983) On the Shape of a Set of Points in the Plane, IEEE Trans. Inform. Theory, 29(4), 551-559. ISSN 0018-9448.

Geetha, G., T.V.S Udaya Bhaskar, E. Pattabhi Rama Rao (2011) Argo data and products of Indian Ocean for low bandwidth users, International Journal of Oceans and Oceanography Vol 5 (1), 1-8.

Goni GJ, Sprintall J, Bringas F, Cheng L, Cirano M, Dong S, Domingues R, Goes M, Lopez H, Morrow R, Rivero U, Rossby T, Todd RE, Trinanes J, Zilberman N, Baringer M, Boyer T, Cowley R, Domingues CM, Hutchinson K, Kramp M, Mata MM, Reseghetti F, Sun C, T.V.S Udaya Bhaskar and Volkov D (2019) More Than 50 Years of Successful Continuous Temperature Section Measurements by the Global Expendable

Bathythermograph Network, Its Integrability, Societal Benefits, and Future. *Front. Mar. Sci.* 6:452. doi: 10.3389/fmars.2019.00452.

Gould, J., D. Roemmich, S. Wijffels, H. Freeland, M. Ignaszewsky, X. Jianping (2004) Argo profiling floats bring new era of in situ ocean observations. Eos Trans. Amer. Geophys. 179, 190–191. doi: 10.1029/2004EO190002.

Jarvis, R. A., (1973) On the identification of the convex hull of a finite set of points in the plane, *Information Processing Letters* **2**: 18–21. doi:10.1016/0020-0190(73)90020-3

Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, C. R. Paver, J. R. Reagan, D. R. Johnson, M. Hamilton, and D. Seidov, 2013. World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 73, 40 pp.

Manual and Guides 26 - Manuals of Quality Control Procedures for Validation of Oceanographic Data. Prepared by: CEC:DG-XII, MAST and IOC: IODE, SC-93/WS-19 UNESCO, (1993).

Nagamani, P.V., M. M. Ali, G. J. Goni, T.V.S. Udaya Bhaskar, J. P. McCreary, R. A. Weller, M. Rajeevan, V. V. Gopala Krishna and J. C. Pezzullo (2016) Heat content of the Arabian Sea MiniWarm Pool is increasing. Atmos. Sci. Let. DOI: 10.1002/asl.596.

Owens, W.B., Wong, A (2009) An improved calibration method for the drift of the conductivity sensor on autonomous CTD profiling floats by y–S climatology, Deep-Sea Research I, 56, pp 450 - 457.

Roemmich, D., and Coauthors (2001) Argo: The global array of profiling floats. Observing the ocean in the 21st century, C.J. Koblinsky and NR Smith, Eds., Melbourne Bureau of Meteorology, 604 pp.

Venkat Shesu, R., T.V.S Udaya Bhaskar, E. Pattabhi Rama (2013) Open Source Architecture for Web-Based Oceanographic Data Services, Data Science Journal Vol 12, 47-55.

Shimrat, M., (1962) Algorithm 112, Position of Point Relative to Polygon, CACM, p. 434.

Stephanie Guinehut, Christine Coatanoan, Anne-Lise Dhomps, Pierre-Yves Le Traon, and Gilles Larnicol (2009) On the Use of Satellite Altimeter Data in Argo Quality Control. Journal of Atmospheric and Oceanic Technology, **26**, 395–402.

Thadathil, P., A.K. Ghosh, J.S. Sarupria, and V.V. Gopalakrishna (2001) An interactive graphical system for XBT data quality control and visualization. Comp. Geosci. 27, 867–876. doi: 10.1016/S0098-3004(00)00172-2.

Udaya Bhaskar, T.V.S., M. Ravichandran, R. Devender (2007) An operational objective analysis system at INCOIS for generation of Argo Value Added Products, Tech. Rept, INCOIS-MOG- ARGO-TR-04-2007.

Udaya Bhaskar, T.V.S., D. Swain, and M. Ravichandran (2007) Mixed layer variability in Northern Arabian Sea as detected by an Argo float, Ocean Science Journal 42 (4), 241-246.

Udaya Bhaskar, T.V.S., E. Pattabhi Rama Rao, R. Venkat Shesu and R. Devender (2012) A Note on 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.

Udaya Bhaskar, T.V.S., R. Venkata Seshu, E. Pattabhi Rama Rao, R. Devender (2013) GUI based interactive system for Visual Quality Control of Argo data, Indian Journal of Geo-Marine Sciences 42 (5), 580-586.

Udaya Bhaskar, T.V.S., R. Venkata Shesu, T.P. Boyer, E. Pattabhi Rama Rao (2017) Quality control of oceanographic in situ data from Argo floats using climatological convex hulls, MethodsX Vol 4, 469-479.

Wong, A. P. S., G. C. Johnson, and W. B. Owens (2003) Delayed-mode calibration of autonomous CTD profiling float salinity data by theta-S climatology, Journal of Atmospheric and Oceanic Technology, **20**(2), 308-318.

Annie Wong, Robert Keeley, Thierry Carval and the Argo Data Management Team (2020). Argo Quality Control Manual for CTD and Trajectory Data. http://dx.doi.org/10.13155/33951.

Zweng, M.M, J.R. Reagan, J.I. Antonov, R.A. Locarnini, A.V. Mishonov, T.P. Boyer, H.E. Garcia, O.K. Baranova, D.R. Johnson, D.Seidov, M.M. Biddle, 2013. World Ocean Atlas 2013, Volume 2: Salinity. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp.