PERGAMON

Computers & Geosciences 27 (2001) 867–876

COMPUTERS & GEOSCIENCES

An interactive graphical system for XBT data quality control and visualization☆

Pankajakshan Thadathil*, Aravind K. Ghosh, J.S. Sarupria, V.V. Gopalakrishna

National Institute of Oceanography, Dona Paula, Goa 403 004, India

Received 3 August 1999; received in revised form 20 April 2000; accepted 1 June 2000

Abstract

A PC-based system has been developed for quality control and visualization of expendable bathy thermograph (XBT) data archived at the Indian Oceanographic Data Centre. The system, coded in Visual C++, is user interactive and runs on Windows-95 platform. Quality control module of the system incorporates various quality norms/checks and has two levels; inventory and data levels. Inventory level checks are applied for land–sea position, speed of the vessel, invalid date/time, duplicates and station sounding. Station sounding check is performed based on the ETOPO bathymetry file having $5' \times 5'$ spatial resolution. Although the QCS is developed for quality control and visualization of the XBT data, it could be used for inventory level quality checks of any general oceanographic data. Data quality module involves tests for XBT-specific errors such as surface transient, temperature inversions, fall rate and depth reversal. This level also involves visual inspection of the profiles for identifying and correcting/flagging of features caused by wire stretch, wire break, bowing and nub in the mixed layer. Provision is given to compare individual XBT profiles with neighboring stations and also with $1^\circ \times 1^\circ$ monthly climatologies. Quality flags are assigned for each inventory and depth fields. Since data exchange (national and international), under the IODE system, stipulates standard quality flags, the Integrated Global Ocean Observing System recommended flags are applied in the system. Station having data with erroneous or doubtful flags are sent to error bin, which can be accessed by privileged users for possible correction and subsequent modifications in the data base. The data visualization module has options for data queries, selection and graphical presentations, like vertical and horizontal distribution of temperature.

Keywords: Quality flags; Surface transient; Fall rate error; Temperature inversion ; Visual C++

1. Introduction

Expendable bathy thermograph (XBT) is the most simple, economic and easy to operate among various instruments available to measure upper ocean thermal data. The XBT was developed to provide an upper ocean temperature-profiling device that could be operated while the ship is underway. The XBT data base is the largest upper ocean temperature data base available presently to any one in the world and is widely used by the oceanographic community. Since research ship time is expensive for long-term measurements and also increasing demand for synoptic measurements, usage of XBTs in oceanographic studies grew fast under the auspices of Tropical Ocean and Global Atmosphere (TOGA) and World Ocean Circulation Experiment (WOCE) programs. XBT is a free-falling temperature probe, which is connected to ship-board computer by a thin cable that breaks off as soon as the depth range of the probe is passed. To achieve free-fall, independent of the ship's motion, the data transfer cable is constructed from a fine copper wire with feed spools in both, the sensor probe and the launching canister. At each stage

☆Code available from server at http://www.iamg.org/CGEditor/index.htm.

*Corresponding author.

E-mail address: pankaj@csnio.ren.nic.in (P. Thadathil).

0098-3004/01/$ - see front matter
PII: S0098-3004(00)00172-2

of the probe's fall, depth of the probe is deduced from the fall rate equation supplied by the manufacturer. Prior to 1980s, XBT recording units used to be of analogue type involving manual digitization to get temperature versus depth data. The PC-based Mark 12 XBT data acquisition system revolutionized upper ocean temperature measurements. As the probe falls through the water column, the MK-12 system display and store the temperature data at depth intervals of ~0.6 m in the onboard computer.

Indian Oceanographic Data Centre (IODC) archives XBT data from Indian Ocean, collected under the national and international programs. The Quality Control System (QCS) discussed here is developed based on our experience in processing the IODC XBT archive, mainly from the tropical Indian Ocean.

Being different from conventional in situ measurements (operational while ship is underway, no depth sensor), XBT data is susceptible to errors. In addition to the XBT-specific errors, it is likely to have the general errors existing in oceanographic station data (UNESCO, 1988), like location error, duplicate records, spikes, profile inconsistency, etc. The XBT specific errors are:

(a) Surface temperature error caused by start up transient in the digitizer electronics (Roemmich and Cornuelle, 1987).
(b) Depth error caused by improper fall rate equation supplied by the manufacturer (Hanawa et al., 1995; Thadathil et al., 1999).
(c) Unrealistic inversion like feature in surface layer temperature caused by wire stretch/pull or pinhole leak. Wire stretch/pull happen when the wire is not freely released from the feed spool (Bailey et al., 1993; Thadathil et al., 1999).
(d) Random errors caused by wire break/pull, pinhole leak in the wire, and bowing effect (Bailey et al., 1993).

Most of these errors, either XBT-specific or general, could be identified and corrected (or flagged for the type of errors) if the data originators/users are aware of these possible source/nature of the errors. However, often, either the originator or the user may not be aware of such errors. Therefore, a user interactive and computer-based QCS, which incorporates all such checks/norms as mentioned above, would be quite a useful tool for quality control of the XBT data. The QCS system discussed in this paper is developed with this objective and to help the data originators/users to identify such errors and to apply the possible correction in both on-line as well as delayed mode of data processing. Although the QCS is developed for quality control and visualization of the XBT data, it could also be used for inventory level quality (Section 2.1) checks of any general oceanographic data. The QCS, involving various modules is discussed in Section 2 and the quality codes considered are discussed in Section 3. The data visualization features are given in Section 4, followed by discussion in Section 5.

2. QCS-modules

An outline of the PC-based QCS with various modules is given in Fig. 1. Source code of the system is written in C and Visual C++ and runs on Windows-95 platform. Mainly, the system consists of two modules, one for quality control and the other for visualization.

2.1. Inventory-level checks

Quality control procedures involved in this sub-module are checks for duplicate records, position, station sounding, date and time. Fig. 2 depicts the QCS window for inventory level checks.

2.1.1. Duplicates

Check for duplicate records is carried out to avoid occurrence of same record more than once. Here, we have considered a record as duplicate if the match is exact, where all the fields in a particular record are identical to those of a previous one. An inexact duplicate will have at least one difference. If a duplicate is

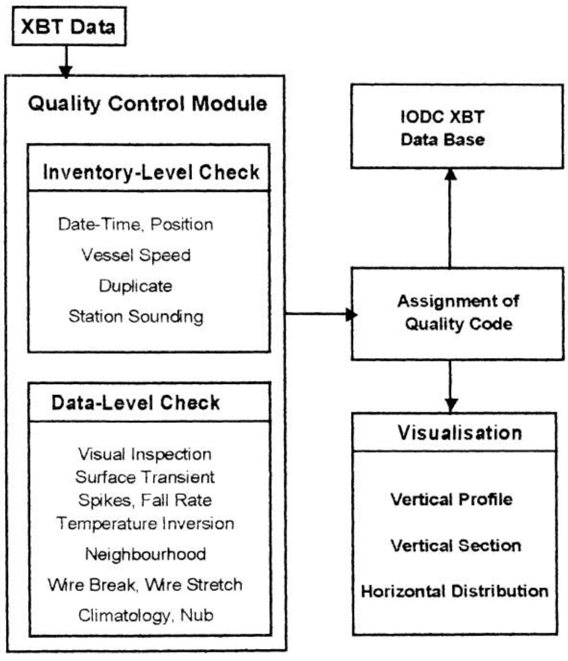


Fig. 1. Outline of QCS, depicting quality control and visualization modules.

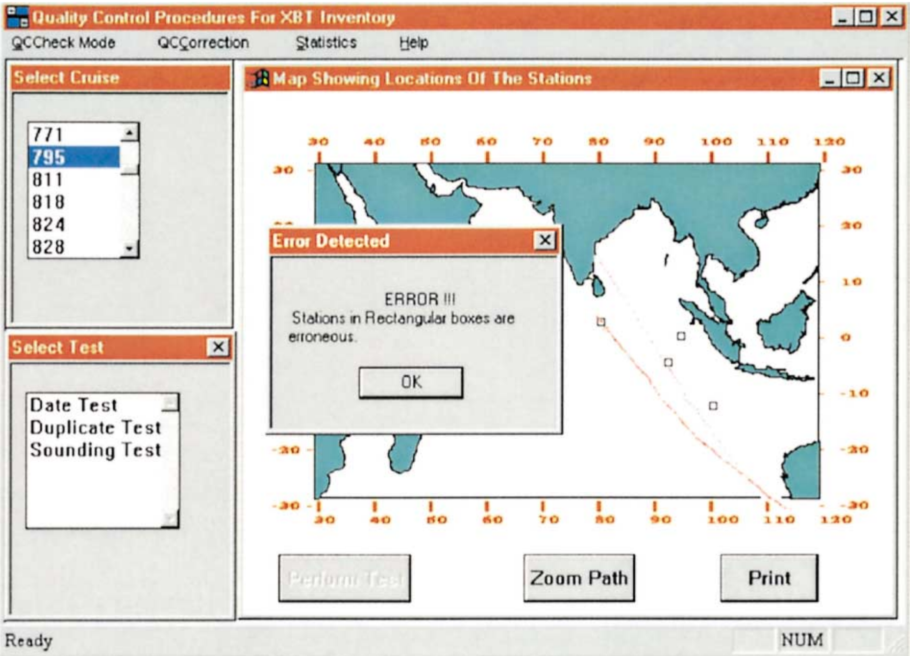


Fig. 2. QCS window for inventory level quality checks. Error message in message window is outcome of position test performed for cruise 795 (see selected cruise). Black square boxes are flagged stations, failing position check based on speed of vessel.

identified as an exact match then one of the records is deleted.

2.1.2. Position check

Gross error in station position in the whole XBT archive is identified by plotting the station position along with land boundary. Stations falling on land are obviously erroneous. Positions of such stations are corrected by seeing the cruise tracks. Even if there is no station falling on land, there may be error in positions. In order to identify such errors, individual cruise track is regenerated from the IODC data base and then compared it with the original cruise track from the cruise reports. Position check is also applied by verifying the numeric of latitude and longitude fields.

2.1.3. Position check based on speed of the vessel

This check is performed on oceanographic stations from a particular cruise. The distance between two neighboring stations is compared with the speed of the vessel and the time taken between the two stations. When the distance covered between the stations appears to be unrealistic, based on maximum possible speed of the vessel (~15 knots), the position is flagged for speed test failure (Fig. 2). The position is not flagged for speed test failure when the time taken is exceeding the required time as such cases are common when ships are stopped for longer period at a particular station.

2.1.4. Station sounding

Station sounding (depth to bottom) is checked based on the ETOPO bathymetry files, having a spatial resolution of $5' \times 5'$. Station sounding is compared with the ETOPO sounding for the same latitude and longitude and flagged as doubtful if the difference exceeds 100 m. This check is found to be valid mainly for open ocean stations.

2.1.5. Date-time

Here, the invalid components of the date-time fields are identified. For example, dates having numeric range other than 1–31 is invalid. Similar criteria are used for other components in the date-time field.

2.2. Data-level checks

Quality checks incorporated in this module are: surface transient, visual inspection, spikes, temperature inversion, fall rate, neighborhood, climatology and other common XBT malfunctions. Fig. 3 depicts the QCS window for data level checks.

2.2.1. Surface transient

Surface temperature error due to start-up transient in the digitizer electronics was reported by Roemmich and Cornuelle (1987). They observed that the magnitude, sign and duration of this transient vary from instrument

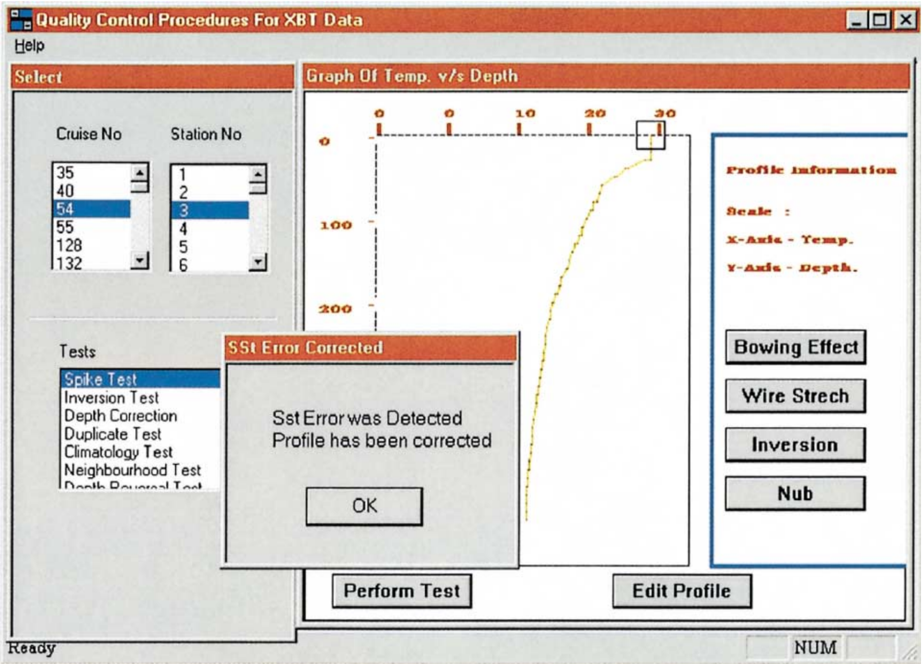


Fig. 3. QCS window for data level quality check. Window depicts outcome of surface transient test (SST) for cruise 54 and station 3. Station profile is plotted with flag (black square) and error message for SST failure. To correct obviously erroneous values there is an '*edit profile*' button. Sample temperature profiles having bowing effect, wire stretch, inversion, and nub could be selected by clicking on respective buttons in order to check existence of such features in current temperature profile.

to instrument and even from cast to cast. One way to overcome the start-up transient problem is to flag the data up to the transient depth for poor quality (Bailey et al., 1993). Instead of flagging the data up to the transient depth, Thadathil et al. (1999), using the simultaneous CTD–XBT data set, suggested a method to minimize the error in XBT surface temperature (XST). When the XST and bucket temperature difference is exceeding 0.2°C and also when the bucket temperature is not reported, it is suggested to take the 5 m XBT-temperature (5-XT) as the XST. This extrapolation procedure of 5-XT as XST was validated using the controlled XBT–CTD data set by Thadathil et al. (1999) and the same is used in the system for surface transient error.

2.2.2. Visual inspection

A provision is given for graphical presentation of temperature profile of the XBT data for a preliminary quality check. Individual station profiles can be selected either randomly or cruise-wise and station-wise (Fig. 3). Numerical values of the profile (depth and corresponding temperature) are also displayed in an adjacent window and the particular value of the profile, where the cursor is, is also highlighted. This window also provides an edit facility for possible correction of the profile where it is found obviously erroneous.

2.2.3. Spikes

Spikes in temperature profiles are caused by small-scale geophysical and instrumental noise. However, in certain cases, the spikes can also be caused by instrument malfunction and when it exceeds a limit (tolerance) the profile should be edited for abnormal spikes. A sample spike is given in Fig. 4. Here, we have used the following formula for detecting and editing such abnormal spikes and the tolerance 'ΔT' is kept as 1°C.

$$\Delta T = [T_m - (T_{m+1} + T_{m-1})/2] - [(T_{m-1} - T_{m+1})/2], \quad (1)$$

where T_{m-1} represents the temperature where the spike starts and T_{m+1} represents the temperature where the spike ends. T_m is the maximum temperature of the spike.

2.2.4. Temperature inversion

Inversion in an XBT profile may be a real one or due to wire stretch or pinhole leak in the cable (Bailey et al., 1993). Inversion-like features may also appear in the surface layers due to some unknown reasons. In certain regions, especially in frontal zones where the warm water meets cold water, such as the Kuroshio and

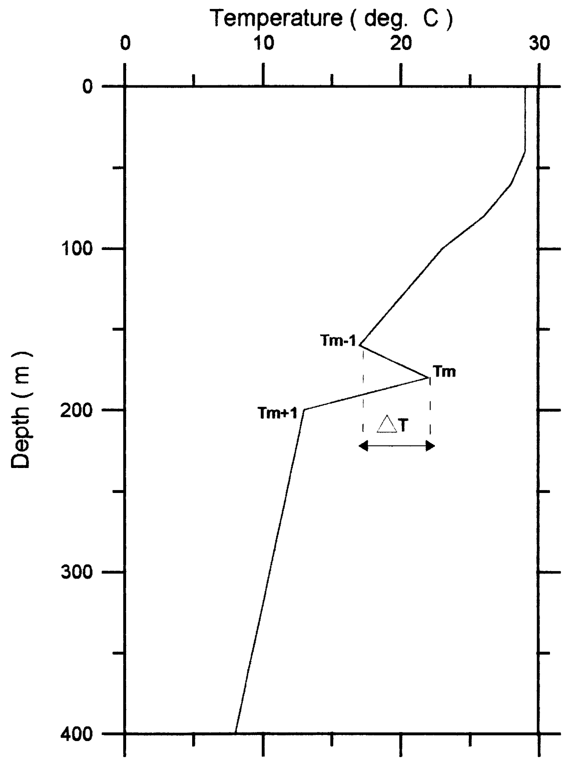


Fig. 4. Sample plot of temperature profile with spike. Spike is defined by temperatures; T_{m-1}, T_m and T_{m+1}. Here, $\Delta T = T_m - T_{m-1}$.

Oyashio, temperature inversions are not unusual features (Nagata, 1979). Though such strong thermal fronts were not reported from Indian Ocean, temperature inversions were reported in many earlier studies (Panakala Rao and Sastry, 1981; Rao et al., 1983; Prabhakara Rao et al., 1987; Thadathil and Ghosh, 1992). For distinguishing real inversion from inversion-like feature we have used the region, range and time of inversion occurrence in the Indian Ocean using the IODC CTD archives and then an inversion filter is applied to find the doubtful inversion. Thadathil et al., 1999 observed that in the Indian Ocean, surface layer temperature inversion occurs mostly in winter season and the inversion layer thickness has a significant linear relation with the surface temperature (Fig. 5). The linear relation is as follows:

$$D = \mathrm{abs}(-\mathrm{SST} + 28.4)/0.05, \tag{2}$$

where D is the inversion layer thickness, shown in Fig. 5. After finding the 'D' from SST (sea surface temperature), it is compared with actual inversion layer thickness of the profile and the inversion is flagged as doubtful after keeping the tolerance of 10 m.

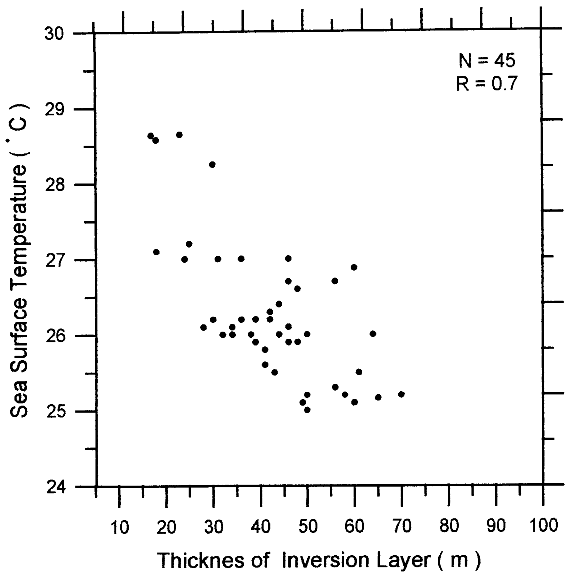


Fig. 5. Plot between sea surface temperature and inversion layer thickness.

2.2.5. Fall rate

In the XBT data, the depth is deduced from the fall rate of the probe, supplied by the manufacturer. However, it is observed that the depth deduction from the fall rate has significant error (~25 m at a depth of 800 m). Hanawa et al. (1995), herein after referred as HN-95, critically examined the XBT depth errors of T5, T6 and T7 probes based on a large set of simultaneous XBT–CTD data from different oceanic regions of the world. They proposed a common depth-time equation for T5, T6 and T7 probes that reduces the mean error to less than 1 m in the depth range of 0–800 m. Since the HN-95 data set was not having any representation from Indian Ocean, Thadathil et al. (1998), using the controlled XBT–CTD data set, considered the use of HN-95 equation for XBT data from Indian Ocean. They found that the HN-95 equation is valid for XBT data from the Indian Ocean as well (Fig. 6). Here, we incorporate the following HN-95 equation for depth correction.

$$\mathrm{ND} = 1.0417D - 75.906(1 - (1 - 0.0002063D)^{1/2}), \tag{3}$$

where ND gives the new depth and D is the original depth.

2.2.6. Neighborhood

In a particular cruise, if the stations are not wide apart (not more than 200 km), the temperature profiles from the neighboring stations are not likely to differ drastically, especially from open ocean, where majority of the XBT data is collected. Here, we have given provisions to compare the station profile with the

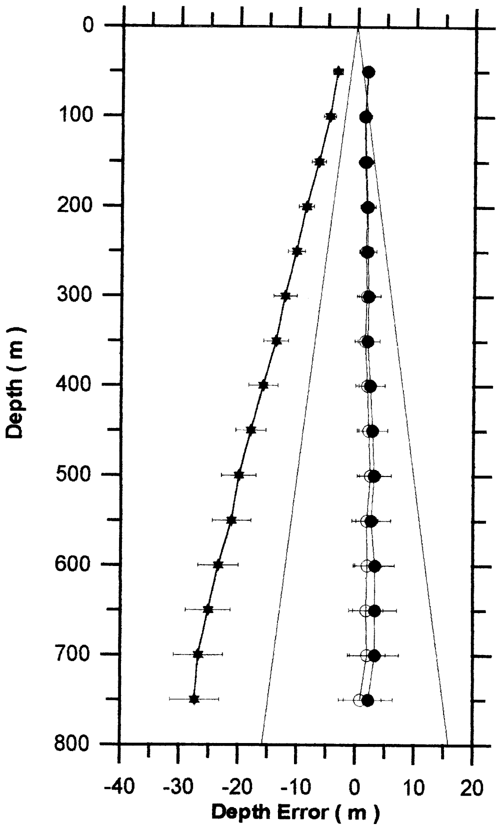


Fig. 6. XBT fall rate error, after Thadathil et al. (1998). Curve with '$*$' represents depth error existing in XBT data. Whereas curve with open circle represents residual depth error after correction using fall rate equation of Thadathil et al., 1998, filled circle represents same for fall rate equation of Hanawa et al., 1995.

preceding and succeeding station's profiles for quality check. Each profile is plotted in distinct colors.

2.2.7. Climatology

For comparing the XBT profile with climatological temperature profiles, we have considered the monthly mean climatology of $1^\circ \times 1^\circ$ grid from the Levitus data set. Here, a particular profile with particular position and time is compared with the corresponding profile from Levitus climatology.

2.2.8. Common XBT malfunctions

The known, major XBT malfunctions are bathysystem leakage or cusping, bathysystem bowing or mixed layer bowing, Sippican MK-9 processor malfunctions or sticking bit problems, wire stretch, wire break, and high-frequency interference. Bailey et al. (1993) reported the above common XBT malfunctions with specimen profiles of such malfunctions and the criteria for accepting/rejecting a profile, identified with such malfunctions. Here, in the visual inspection sub-module of the system, we provide sample profiles with the known malfunctions and associated features in the profile, for immediate reference and quality check.

Table 1
IGOOS quality codes applied in system

Code	Quality description
0	No quality control (QC) has been performed on this element.
1	QC has been performed: element appears to be correct.
2	QC has been performed: element appears to be inconsistent with other elements.
3	QC has been performed: element appears to be doubtful.
4	QC has been performed: element appears to be erroneous.
5	The value has been changed as result of QC.
6	Reserved
7	Reserved
8	Reserved
9	The value of element is missing.

3. Quality codes

Since data exchange (national and international), under the IODE system, stipulates standard quality codes, the Integrated Global Ocean Observing System (IGOOS) recommended codes are applied in the system. The IGOOS quality codes considered in the QCS are given in Table 1. Station having data with erroneous or doubtful codes are sent to error bin, which can be accessed by privileged users for possible correction and subsequent modifications in the data base.

4. Data query and visualization

This module of the QCS has utilities for selecting XBT data in a number of ways, depending on user requirements. It is designed basically to support such application oriented graphical presentations as vertical temperature profile, vertical sections and horizontal distributions. Vertical profile of temperature of a particular station from a particular cruise can be visualized by selecting the station from the cruise and station menus. Selection of temperature profiles from a particular region and month is also made possible either by clicking the mouse on the station coverage window or by entering the coordinates in the latitude–longitude dialog box (Fig. 7a). Vertical profiles of all stations selected will be displayed in different colors, representing different seasons (Fig. 7b).

For horizontal distribution maps, selection can be made for a particular geographic region by manually

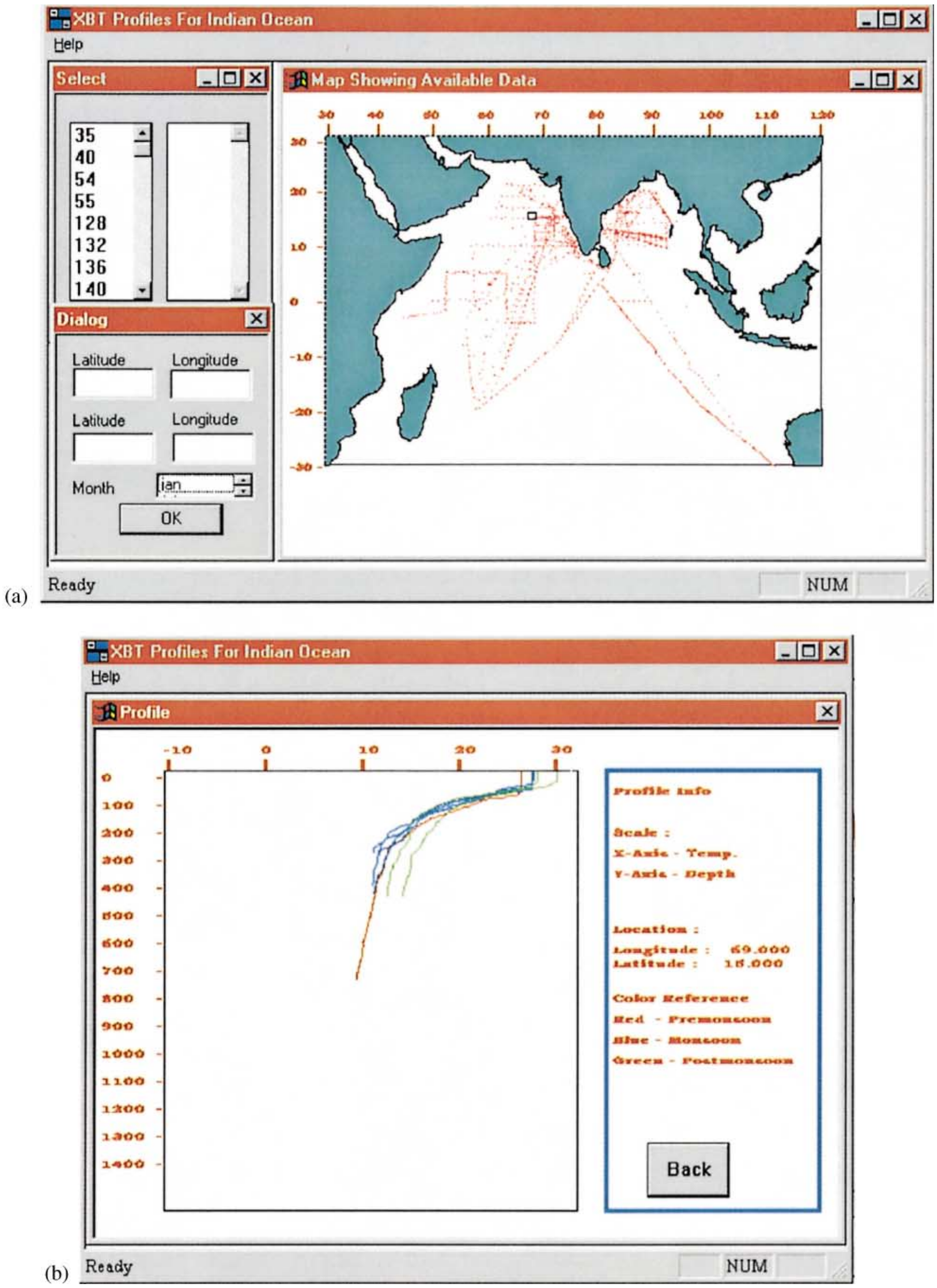


Fig. 7. (a) QCS window for selecting temperature profile/profiles. Selection could be based on cruise and station, latitude–longitude coordinates from dialog box or directly selecting area from coverage map. (b) Represents selected profiles for selected area (black rectangle given in coverage map).

entering the coordinates limits of the desired area. Geographic selection is also possible by clicking the mouse in the station coverage map. Such selections can be combined with time fields such as month/season, years, etc. Once the data selection is completed, the horizontal distribution maps can be prepared for different depths as per the choice of user. While plotting the horizontal distribution maps, the user can select for simple temperature contours or for shaded contours (Figs. 8a and b). For contouring the selected array of temperature values the marching square method is employed. If a particular selection has insufficient data or no data for contouring, the user is prompted about it with a message window. Vertical sections are again

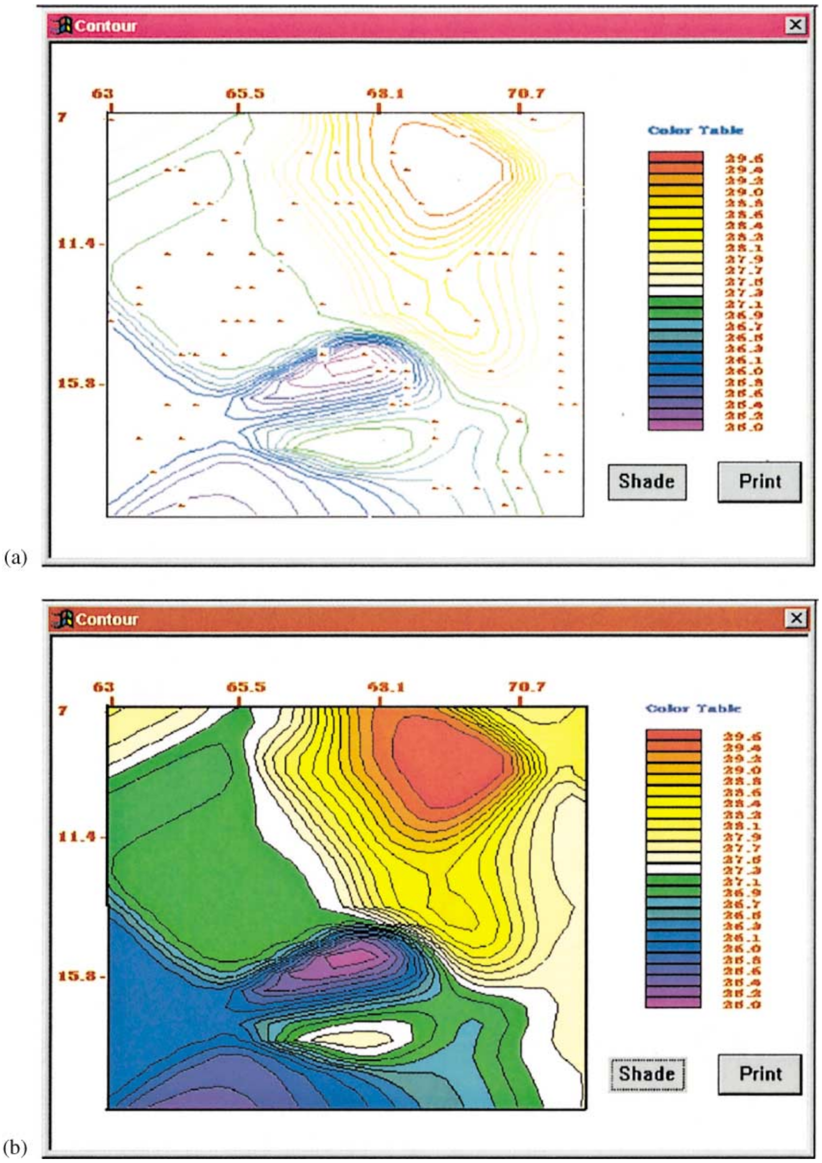
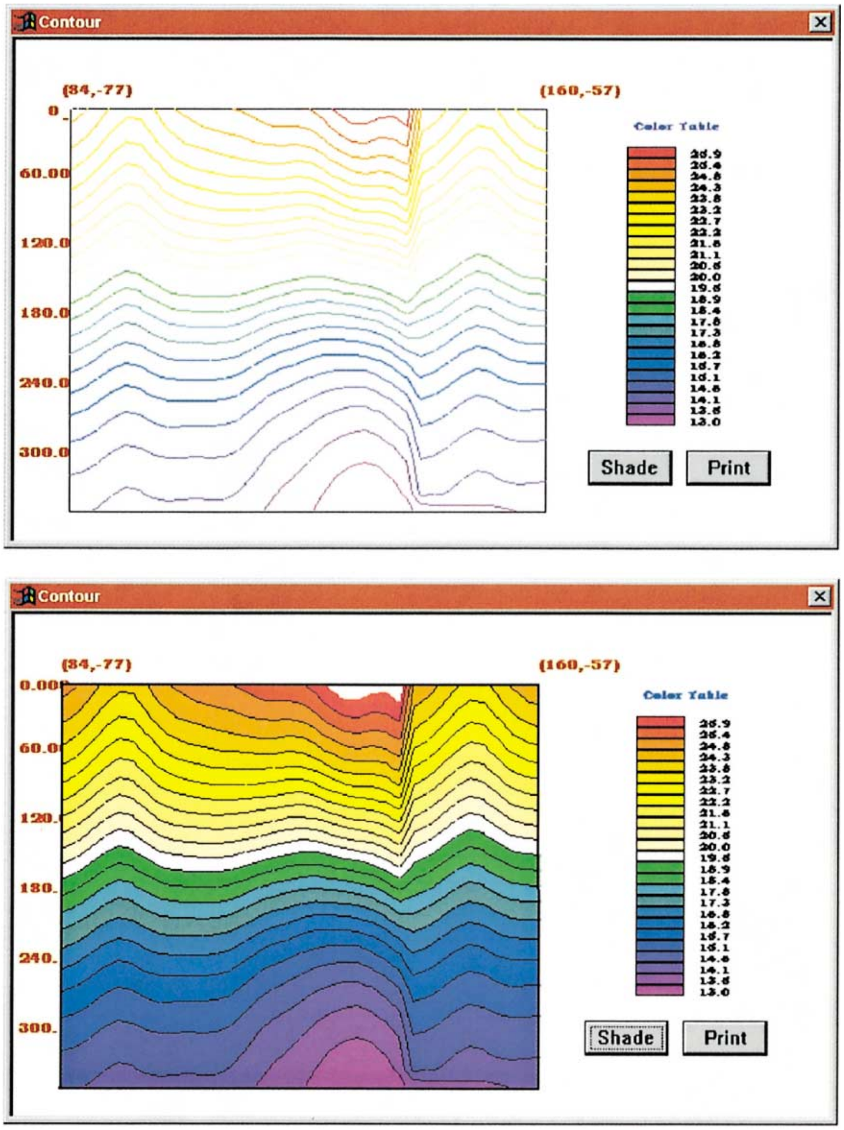


Fig. 8. (a) Contours of horizontal distribution of temperature at surface for selected region from station coverage map. Horizontal distribution maps could be plotted for selected depth, contour intervals and colors from dialog box. (b) Same as in Fig. 8a but for shaded contours.

temperature contours along a particular oceanographic section. Contouring features for vertical sections are similar to the horizontal distribution maps and are displayed in Figs. 9a and b.

5. Discussion

For upgrading any data base from its primitive level of archive to a higher level of value added data-set, it is required to apply stringent QC measures. Such measures have to be persistently followed from the very early stages of data collection till it forms an integral part of a reliable data base. While developing this software for XBT quality control and visualization, effort is made to incorporate quality measures for all known sources of XBT malfunctions, thus enabling for quality data collection, processing and exchange. Beneficiaries of this software are data originators, users and data managers, who may not be aware of XBT-specific error sources in a comprehensive way presented here. By incorporating quality measures applicable to general



(a)

(b)

Fig. 9. (a) Contours of vertical section of temperature for selected section from station coverage map. Vertical section could be plotted for selected contour intervals and colors. (b) Same as in Fig. 9a but for shaded contours.

oceanographic station data along with XBT-specific quality measures the system serves as a comprehensive package with respect to QC. The data visualization module will serve as an additional tool for QC.

Since the system is PC-based and runs on Windows-95 platform, it can be installed on an onboard PC and could also be used for on-line quality control. The window-based application gives better visual effect. Though the system is basically designed for Indian Ocean data archive, it can be used globally except for the inversion filter existing in the QC module.

In all stages of the quality checks, the general philosophy adopted is that the data value is changed only when the value is found to be obviously erroneous. In all other cases only the appropriate codes (given in Table 1) are assigned without changing the original values. Though the QCS is automated, the operator should be cautious in each stage of the QC, especially when special features appear in a profile. In the case of surface layer inversion, though the linear relation between SST and inversion layer thickness can be used as an inversion filter, it is suggested to make use of the profiles from neighboring stations or repeat cast (if any) in delayed mode of QC. In the case of on-line QC, whenever doubtful features appears it is advisable to go for a repeat cast and confirm the existence of the special feature. QC for surface transient and fall rate does not require much judgement by QC operator. Though the station sounding check, based on ETOPO bathymetry sounds well, it is found that the checks are more reliable to open ocean stations compared to coastal stations.

Acknowledgements

We gratefully acknowledge Dr. E. Desa, Director, National Institute of Oceanography for his encouragement in undertaking this study. The NIO contribution number is 3563. The authors are also grateful to their colleagues, from Physical Oceanographic Division (especially the TOGA group), for their efforts put in while collecting the XBT data under the national programme. We thank Mr. Bob Keeley, Marine Environmental Data Services (MEDS), Canada, for providing the ETOPO bathymetry data and also for his valuable suggestions on the quality control aspects presented in this study.

References

Bailey, R.A., Groneu, A., Philips, H., Meyers, G., Tanner, E., 1993. CSIRO cookbook for quality control of expendable bathythermograph (XBT) data, CSIRO Report, Canberra, Australia, 74pp.

Hanawa, K., Rual, P., Bailey, R.J., Sy, A., Szabados, M., 1995. A new depth equation for Sippican or TSK T-7, T-6 and T-4 expendable bathythermographs (XBT). Deep-Sea Research 43 (8), 1423–1451.

Nagata, Y., 1979. Shallow temperature inversions in the Pacific Ocean near Japan. Journal of the Oceanographical Soceity of Japan 35, 141–150.

Panakala Rao, D., Sastry, J.S., 1981. Circulation and distribution of some hydrographical properties during the late winter in Bay of Bengal, Mahasagar. Bulletin of the National Institute of Oceanography 14, 1–15.

Prabhakara Rao, B., Ramesh Babu, Chandra Mohan, P., 1987. Seasonal and diurnal variability of thermal structure in coastal waters of Visakhapatnam. Proceedings of the Indian Academy of Sciences (Earth Planetary Sciences) 96, 69–79.

Rao, R.R., Rao, D.S., Murthy, P.G.K., Joseph, M.X., 1983. A preliminary investigation on the summer monsoonal forcing on the thermal structure of upper Bay of Bengal during Monex-79. Mausam 32, 92–95.

Roemmich, D., Cornuelle, B., 1987. Digitization and calibration of the expendable bathythermograph. Deep-Sea Research 34, 299–307.

Thadathil, P., Ghosh, A.K., 1992. Surface layer temperature inversion in the Arabian Sea during winter. Journal of Oceanography 48, 293–304.

Thadathil, P., Ghosh, A.K., Muraleedharan, P.M., 1998. An evaluation of XBT depth equations for Indian Ocean. Deep-Sea Research 45 (4–5), 819–827.

Thadathil, P., Ghosh, A.K, Pattanaik, J., Ratnakaran, L., 1999. A quality control procedure for surface temperature and surface layer inversion in the XBT data archive from the Indian Ocean. Journal of Atmospheric and Oceanic Technology 16, 171–173.

UNESCO, 1988. IOC-manual and guides-3, Guide to operational procedures for the collection and exchange of IGOS data, 68pp.