

A QUALITY CONTROL PROCEDURE FOR SEAWATER TEMPERATURE DATA

Aravind Ghosh K. and T. Pankajakshan

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

ABSTRACT

A three level quality check has been developed for seawater temperature data. The standards used in the procedure for the quality check are the characteristic property of vertical temperature distribution, watermass property and a standard temperature climatology derived from historical data. In each level, the quality flags are assigned for the inversion station, abnormal intracruise variability and climatologically abnormal station. The first two checks are applied to the Indian National Oceanographic Data Centre (INODC) MBT data files.

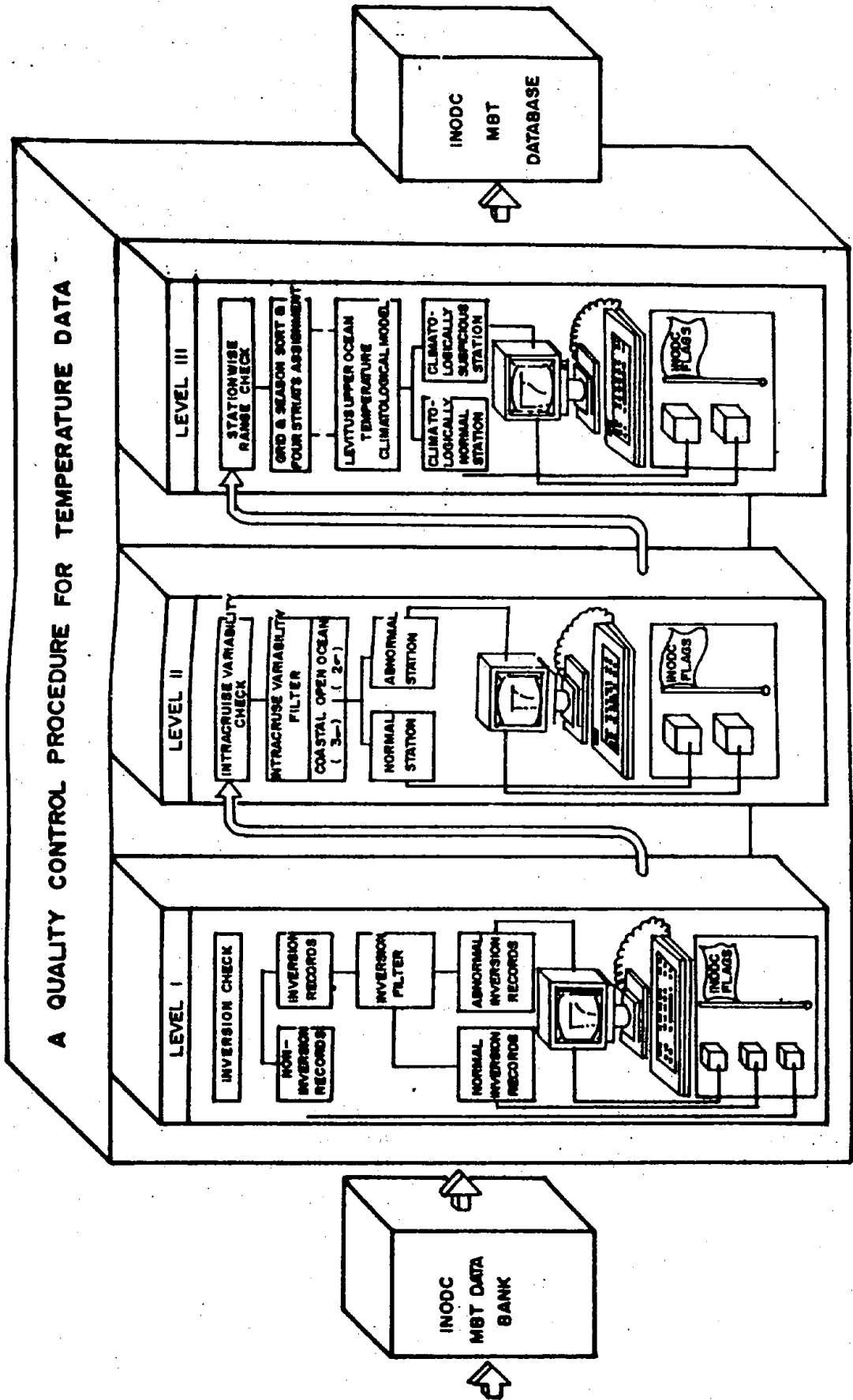
INTRODUCTION

The basic purpose of any quality control procedure for oceanographic data is to eliminate erroneous data from original data. When the value of the erroneous data falls within the range of normal distribution of a property it is difficult to distinguish the erroneous data from the original data. Whereas if the value of the erroneous data falls out of the range of normal distribution, it is comparatively easy to locate. In the latter case the errors appear as spikes in the distribution (vertical/horizontal) profiles. On the basis of the origin, the errors can be classified into three categories. They are physical errors, coding errors and transmitting errors.

Physical errors arise from the malfunctioning of instruments and become visible. This type of errors originate onboard a platform. Coding errors are primarily caused by men, in case they do not encode a measurement according to the rules or in case they insert a wrong value. Transmitting errors appear during transmission of data from one place to another, or if the data changes from one storage medium to another, or because of miscommunication.

Considering all these sources of errors, it is not practically possible to develop a quality control procedure, which can take care of all errors. When they fall out of the range of normal distribution values, they can be corrected to a certain extent, if the data is treated by an expert, who is well aware of the sources of errors and also the ranges of the physical values. It is on principle not possible to identify these errors definitely, due to lack of the original data. But with some simple statistical checks against temperature limits, spikes, vertical stability etc., a lot of obviously erroneous data can be discarded and also the doubtful data can be flagged. In the present paper we have discussed one of such quality checks applied to the INODC bathythermograph data files.

A QUALITY CONTROL PROCEDURE FOR TEMPERATURE DATA



METHOD

Quality check to any data/product is always based on some standards. The standards we used for the quality check of temperature data are the following:

- a) The characteristic property of vertical temperature profile (decrease of temperature with the increase in depth).
- b) Watermass property: An abrupt change of values of consecutive stations of a particular cruise is an uncommon feature except in thermal fronts of high horizontal gradient. Such fronts are not reported from the area where the data has been collected.
- c) A standard temperature climatology derived from historical data.

All these three standards are incorporated in the quality check at three different levels, namely Level-1, Level-2 and Level-3 (Fig. 1). So a particular station temperature data will be undergoing all these three checks.

Level-1: This level uses the characteristic property of vertical temperature profile. The general tendency of vertical temperature profile is, a decrease of temperature with increasing depth. Of course, a subsurface inversion may also be present in some records. Therefore, in Level-1 the first step of operation is, sorting out the inversion records and putting the inversion flag. The inversion records are then passed through inversion filter ($0.5^{\circ}\text{C}/\text{m}$). For the entire data set, it is found that inversion confines to the upper 100 m layer. Hence the limit of inversion depth was taken as 100 m. While passing the temperature records through this filter, it is observed that all the data records have a free passage. To find out the limiting value of inversion by using the stability criteria, the salinity values are not available for the same stations where BT data are available.

After filtering out the inversion records the program takes the temperature data records coming through the inversion filter to the monitor to check for any spikes present in the profile. If any spikes are present an online correction is applied. So in Level-1 the program does the inversion flagging and an online correction for suspicious data.

Level-2: In this level the standard used for checking is the intracruise variability. Within a particular cruise, covering a time scale of the order of a month, abrupt changes in temperature of consecutive stations are considered as suspicious data. The data set for which this quality check is applied comprises the temperature data collected on board R.V. *Gaveshani* & O.R.V. *Sagar Kanya*. The distance between consecutive stations have a minimum of ~ 50 km and the maximum of ~ 100 km. Since the data is collected from the Arabian Sea and the Bay of Bengal, a region having no strong thermal fronts, the temperature between consecutive stations (with a distance of roughly hundred kilometres) may not differ substantially. This low variability between consecutive stations may be attributed to the property of watermass. Instead of taking the station to station difference we have taken a seven station moving average for the statistical convenience.

In the present check the program will take the first seven stations of a particular cruise and find out the average and standard deviation of the temperature at zero depth. The quality check starts from fourth station onwards. The average temperature of the first seven stations will be subtracted from the fourth station's temperature at zero depth. If this value is more than $3\sigma / 2\sigma$ depending on the station in coastal or open ocean, the fourth station's temperature at zero depth will be flagged accordingly. After this the program will apply the same procedure for the remaining depths. Once the entire depth is over the program leaves back the first station and takes the eighth station of the particular cruise and does the same procedure. The limitation of this seven station moving average procedure is that the first and last three stations of a particular cruise will not be under going the second level check. By adaptation of 3σ or 2σ the filter used in this check has been kept wider so that only those records showing too abnormal value would be flagged. The provision is made to keep this "abnormal intracruise variability flag" with each depth value. The separation of coastal and open ocean has been done on the basis of 200 m depth contour.

Level-3: In this level the quality check is based on a temperature climatology, derived from historical data. The area of data flow is divided into one degree square grids and the historical temperature data available in each grid is sorted monthly and the average profile for all the twelve months are assigned to each grid. Now for a new station temperature profile the program will derive three basic straits such as the SST, depth of 20°C isotherm and depth of maximum gradient and then these values will be compared with the values of historical profile. If the difference is greater than 3σ or 2σ depending on coastal or open ocean situation, the station will be flagged accordingly. This check could not be applied to the INODC MBT data files due to the non-availability of a proper set of historical data.

ACKNOWLEDGEMENT

The authors are grateful to Dr. B.N. Desai, Director, NIO for his keen interest and encouragement for developing this quality control procedure. They are also thankful to Sri R.M.S. Bhargava for his valuable suggestions. Thanks are also due to other Divisional colleagues.