Project: WP15

Real Time Quality Control of

Current measurements

First release:
17 April 2010

Authors for deliverable:
Thomas Hammarklint, SMHI, Sweden
Dimitris Kassis, HCMR, Greece
Henning Wehde, IMR, Norway
Lesley Rickards, BODC, United Kingdom

Table of contents

1. INTRODUCTION

MyOcean is the implementation project of the GMES Marine Core Service, aiming at deploying the first concerted and integrated pan-European capacity for Ocean Monitoring and Forecasting (http://www.myocean.eu.org). The project objective is to observe, analyze and forecast the oceans at global and regional (European Seas) scales in order to provide a monitoring service for marine environment and security. Based on the approach on combining space and in-situ observations and their assimilation into 3-D simulation models, the MyOcean Service aims to provide the best information available on the global and regional ocean. This information includes temperature, salinity, currents, ice extent, sea level and primary ecosystems. Its target applications are marine safety, marine resources, climate and seasonal forecasting as well as marine and coastal environment.

An important step within the MyOcean project is to uniform existing Real Time Quality Control (RTQC) and quality assurance procedures of the different nations involved. As the MyOcean service is thought to be available at any time and open to anyone, an agreement in good Real Time Quality Control (RTQC) methods and procedures is vital to guarantee high data quality distributed to users via international exchange. The agreement on the implementation of uniform RTQC procedures have the severe potential to overcome the inconsistency within the existing datasets provided actually by the international community.

This document is based upon previous experiences and has the following objectives:
"To ensure the data consistency within a single data set and within a collection of data sets and to ensure that the quality and errors of the data are apparent to the user who has sufficient information to assess its sustainability for a task." (IOC/IODE, Manuals and Guides 26, 1993).

One of the various tasks of the MyOcean project - the Work Package (WP) 15 – deals with the scientific and technical validation of In Situ-TAC (Technical Assembly Centres) products and forms the frame of this document the RTQC procedure for current measurements. WP15 aims to perform operational quality control (QC) of global and regional products as well as to lead scientific assessment validation activities with regional responsibilities. Beside global scale products, regional specifications are performed in the Arctic, the Black Sea, the North-western Shelves, the Baltic Sea, the South-western Shelves and the Mediterranean Sea. It follows therewith the EuroGOOS regional approach, with establishing regional alliances.

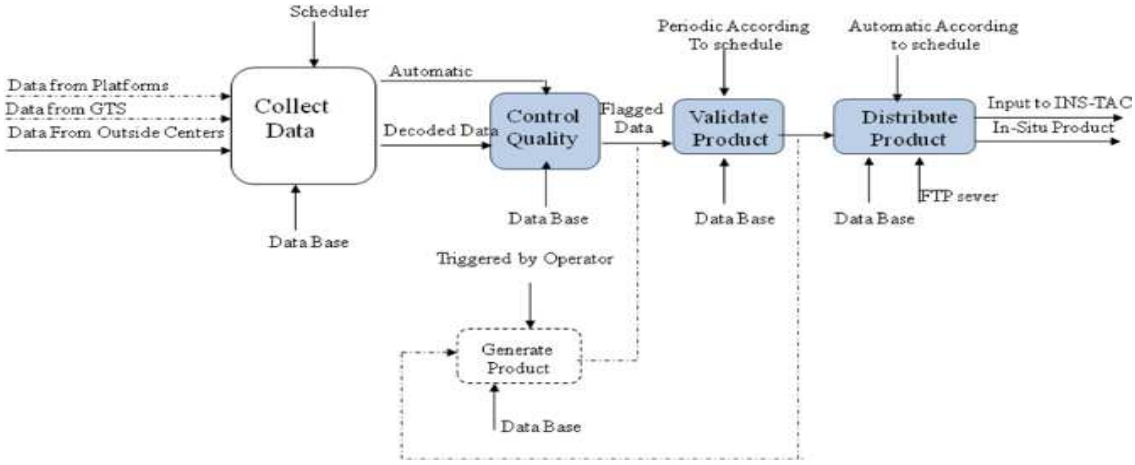

Figure 1: Functions to be implemented by an In-Situ Tac component (meeting report: MYO-INS-MR-2009-03-30)

The different functions to be implemented by the global and regional components of the In-Situ Tac are summarized in Figure 1. The focus of this document lies on one of the central part of the functions, i.e. RTQC of the collected data for the parameter current. The validation procedure includes the delayed mode quality control of the data and will be specified in another guideline.

The document is organized as follows. Section 2 will specify Metadata information. In section 3, automatic RTQC procedures will be listed for different types of measurements, i.e. profile data and time series. Then the QC flagging scales will be defined.

Finally, different procedures of extended QC methods will be discussed which can be applied within the validation procedure (Figure 1). The proposals for RTQC of current data given within this document are based on IOC/IODE (1993), Mersea (2005) and SeaDataNet (2007) QC procedures. As an example of how the current data will be processed, the RTQC scheme for the Baltic area (BOOS) is given below.

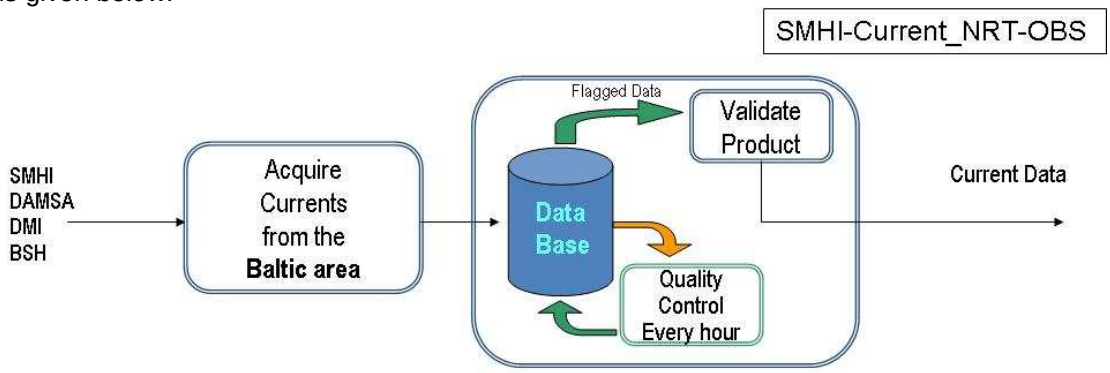

Figure 2. Production Line in BOOS for Current data, which will be maintained by SMHI. Current data are measured at four institutes (SMHI, DAMSA, DMI and BSH). SMHI acquires the data from the four institutes, performs quality control and validation of the product and finally distribute the product to the MyOcean Information System (MIS).

2. METADATA

Detailed Metadata are needed to guideline those involved in the collection, processing, QC and exchange of data. The quality controlled data set requires any data type (profiles, time series etc.) to be accompanied by key background information. The following information is required (MyOcean, 2010):

1. **Position** of the measurement (latitude, longitude, depth).
2. **Date** of the measurement (data and time in UTC or clearly specified local time zone).
3. **Method** of the measurement (e.g. instrument types)
4. **Specification** of the measurement (e.g. station numbers, platform code, name of the data distribution center).
5. **Owner** of the measurement (name and institution of the data originator for traceability reasons).
6. **Processing** of the measurement (e.g. details of processing and calibration already applied, algorithms used to compute derived parameters).
7. **Comments** on measurement (e.g. problems encountered, comments on data quality, references to applied protocols).

2.1 MyOcean Parameter Dictionary Codes

short_name	myo_code	units	Valid_min	valid_max	standard_name	long_name
CurDir	HCDT	deg	0	360	current_direction	DIRECTION_REL._TRUE_NORTH
CurSp	HCSP	m/s	0	10	current_speed	HORIZONTAL_CURRENT_SPEED
CurEast	EWCT	m/s	-10	10	current_east	CURRENT_EAST_COMPONENT
CurNorth	NSCT	m/s	-10	10	current_north	CURRENT_NORTH_COMPONENT

Table 1. The parameters presently included in the MyOcean parameter dictionary which relate to current meter measurements. This can be extended as necessary.

2.2 Quality Control Flags

Quality control flags are applied by the Production Unit during the RTQC procedure. It is the responsibility of final users to read and use these flags to be sure that only the required quality data are used. The quality controlled data will be freely shared and used for various applications in the marine environment. Thus, after the RTQC procedure, an extensive use of flags to indicate the data quality is vital since the end user will permit the selection of the data based on quality flags amongst other criteria. It is important to note that from this scheme, the codes 0, 1, 4 and 9 are mandatory to apply after the RTQC procedure (marked in red).

Code	Meaning
0	No QC was performed
1	Good data
2	Probably good data
3	Bad data that are potentially correctable
4	Bad data
5	Value changed
6	Below detection limit
7	In excess of quoted value
8	Interpolated value
9	Missing value
A	Incomplete information

Table 2: Quality flag scale. Codes marked in red are mandatory after the RTQC procedure.

2.3 Guidance to users

Code	User action
0	Should not be used without a quality control made by the user.
1	Data has passed all applied quality checks. Can be used safely without further analyses.
2	Data may be good for some applications but the user should verify this.
3	Data might be useful in a delayed mode after correction by the data centre.
4	Measurements should be rejected. Do not use this data.
8	Data derived from interpolation, data is not sufficient to mark as code 1 or 2.
9	Data is missing.

The QC flag value assigned by a test cannot override a higher value from a previous test. For example a value with QC flag '4' (bad data) or '3' (bad data that are potentially correctable) is ignored by the quality control tests.

2.4 Current stations in Europe

The measurement of currents in Europe is mostly done with bottom-mounted ADCP. Data is processed and transmitted in real-time via a surface buoy.

Figure 3. Current stations delivering data in real-time in the Baltic and North Sea (from http://www.seprise.eu, real-time products maintained by SMHI).

Figure 4. Current stations delivering data in real-time in the Mediterranean Sea and along the Atlantic coast-line (from http://www.seprise.eu, real-time products maintained by SMHI).

3. REAL TIME QUALITY CONTROL: AUTOMATIC CHECKS

One central part of the functions to be implemented by an in-situ TAC is the control of incoming decoded measurements (Figure 1). Since at this step data should be available in real time, the QC during that process is limited and automated. In the following, automated RTQC will be described for moored current measurements (SeaDataNet, 2007).

3.1 RTQC for vertical profiles: Moored ADCP

The Acoustic Doppler Current Profiler (ADCP) measures current direction in 3 dimensions. As opposed to the average current meter, an ADCP can measure current speeds and direction at varying depths using a principal known as the Doppler Shift. Automated tests for vertical profiles are presented here, i.e. current measurements from moored ADCP. The checklist and example information below shows the information to be used to ensure that the data are adequately described. Further, missing values or bad/strange values will be flagged as missing data (flag 9).

RTQC1. Platform identification

A test to match a platform against known platforms will be made. Data from unknown platforms will not be distributed.

RTQC2. Impossible date test

The test requires that the observation date and time from the profile data be sensible.

- Year until the current year
- Month in range 1 to 12
- Day in range expected for month
- Hour in range 0 to 23
- Minute in range 0 to 59

This check is making sure that we have a valid date/time, but we also test that the actual date/time of the observation correlates to which data/time that is expected.

Action: If any one of the conditions is failed, the date should be flagged as bad data.

RTQC3. Impossible location test

The test requires that the observation latitude and longitude from the profile data be sensible.

- Latitude in range −90 to 90
- Longitude in range −180 to 180

A test to check if the expected position remains the same within a small tolerance will be performed. If latitude and longitude is transmitted together with the new observations, the test detect weather the buoy is moored or not. If latitude and longitude is not transmitted and/or data is missing or out range for a longer period, an automated warning message will be sent.

Action: If either latitude or longitude fails, the position should be flagged as bad data.

RTQC4. Position on land test

The test requires that the observation latitude and longitude from the profile measurement be located in an ocean. Use can be made of any file that allows an automatic test to see if data are located on land. The test will also detect if the mooring is drifting by comparing to its theoretical position.

Action: If the data cannot be located in an ocean, the position should be flagged as bad data.

RTQC 5. Global range test

This test applies to Table 1. Valid values are given by the columns valid_min and valid_max.
Action: Individual values that fail these ranges should be flagged as bad data.

- Current direction in range 0° to 360°.
- Current speed in range 0 m/s to 10 m/s.

Action: If a value fails, it should be flagged as bad data.

RTQC6. Regional range test

This test applies to only certain regions of the world where conditions can be further qualified.
Action: Individual values that fail these ranges should be flagged as bad data.

Current direction should be in range 0° to 360°. Ot herwise the value will be flagged as bad data.

For current speed the ranges needs to accommodate all of the expected extremes encountered in

different regions:

Baltic Sea
- Current speed in range 0 m/s to 3 m/s.

North Sea
- Current speed in range 0 m/s to 10 m/s.

Atlantic coast-line
- Current speed in range 0 m/s to 5 m/s.

Mediterranean
- Current speed in range 0 m/s to 3 m/s.

Current speeds outside these ranges will be flagged as bad data.

RTQC7. Spike test

A spike is a point in the data series which has an anomalous value outside of the surrounding range. This algorithm is used on the current speeds:

Test value = | V2 − (V3 + V1)/2 | − | (V3 − V1) / 2 | ,

where V2 is the measurement being tested as a spike, and V1 and V3 are the values above and below. The V2 value is flagged when the value exceeds 1 m/s.

Action: Values that fail the spike test should be flagged as bad data.

RTQC8. Stuck value test

For profiles this test looks for current speed at consecutive depths within a profile at one point in time. The rate of change (gradient) of the current speed should exceed 0.01 m/s per meter in the profile.

Action: Values that fail this test are considered as probably bad (flag 2).

3.2 RTQC for time series

Automated tests for time series are presented here. Recommended tests for time series have been chosen based on RTQC of SeaDataNet (SeaDataNet, 2007). Specifications are given if tests differ from those already described in section 3.1.

RTQC1. Platform identification

RTQC2. Impossible date test

RTQC3. Impossible location test

RTQC4. Position on land test

RTQC5. Global range test

RTQC6. Regional range test

RTQC7. Spike test

RTQC8. Stuck value test

Additionally, this test can be performed for time series of current data. For time series the test check that the value may not remain constant compared with a number of previous values (3 hours). This is done both for current direction and speed values.

Action: If this occurs, all of the values of the affected variable should be flagged as bad data (flag 4).

RTQC9. Rate of change in time

The aim of the check is to verify the rate of the change in time. It is based on the difference between the current value with the previous and next ones. Failure of a rate of the change test is ascribed to the current data point of the set.

Action: Current speed values are flagged as bad data (flag 4) if:

$$|V_i - V_{i-1}| + |V_i - V_{i+1}| \leq 2*(2*\sigma_V),$$

where V_i is the current speed value of the parameter, V_{i-1} is the previous and V_{i+1} the next one. σ_V is the standard deviation of the examined parameter. If the one parameter is missing, the relative part of the formula is omitted and the comparison term reduces to $2*\sigma_V$. The standard deviation is calculated from the first month of significant data of the time series.

4. REFERENCES

IOC/IODE, 1993: IOC Manuals and guides No.26: Manual of quality control procedures for validation of oceanographic data

Mersea, 2005: In-situ real-time data quality control.

SeaDataNet, 2007: Data quality control procedures, Version 0.1, 6th Framework of EC DG Research.

MyOcean, 2010: Real time quality control of temperature and salinity measurements.