



United Nations
Educational, Scientific and
Cultural Organization



Intergovernmental
Oceanographic
Commission

Manuals and Guides 54

Ocean Data Standards

Volume 3

Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine Meteorological Data

UNESCO



**OCEAN DATA
STANDARDS**



Ocean Data Standards

Volume 3

Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine Meteorological Data

UNESCO 2013

IOC Manuals and Guides, 54, Volume 3
Version 1
18 April 2013

For bibliographic purposes this document should be cited as follows:

Paris. Intergovernmental Oceanographic Commission of UNESCO. 2013. *Ocean Data Standards, Vol.3: Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine Meteorological Data*. (IOC Manuals and Guides, 54, Vol. 3.) 12 pp. (English.)(IOC/2013/MG/54-3)

© UNESCO 2013

Printed in France

TABLE OF CONTENTS

	Page
1. BACKGROUND	1
2. DATA QUALITY FLAGGING	1
3. INTENDED AUDIENCE	2
4. JUSTIFICATION	2
5. THE QUALITY FLAG SCHEME.....	3
5.1 Primary Level.....	3
5.2 Secondary level.....	4
6. BIBLIOGRAPHY	5

ANNEXES

ANNEX A: Selection of representative examples of existing QC flag schemes

ANNEX B: Examples of mapping the IODE quality flag scheme to existing quality flag schemes.

1. BACKGROUND

The value of standards for the management and exchange of data has always been acknowledged. In the oceanography and marine meteorology domain, there have been many efforts to develop common standards and data frameworks for processing data and information but these have never been widely adopted by the community.

IODE (International Oceanographic Data and Information Exchange) and JCOMM (Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology) recognized that, although there were mechanisms to facilitate coordinated ocean data exchange, these had not resulted in the degree of agreement on a wide range of matters that were needed in order to allow the easy exchange and interoperability of collected data. In 2008, the joint IODE/JCOMM Forum on Oceanographic Data Management and Exchange Standards established the Ocean Data Standards Pilot Project (*Intergovernmental Oceanographic Data and Information Exchange. 2010*).

One of the objectives of this Project is to initiate discussions on a limited set of topics for which it is felt that broad agreement is possible and to achieve broad agreement and commitment to adopt key standards related to ocean data management and exchange to facilitate exchange between data centres and contributing programmes. A second objective is to establish an internationally recognized process for submitting proposed standards and their acceptance by the ocean community.

The recommended standards that are produced by this process are intended primarily for the use of the oceanographic and marine meteorological community. After recommendation, their use will be widely encouraged within IOC and other programmes.

2. DATA QUALITY FLAGGING

The objective of data quality control of oceanographic data is 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 suitability for a task. Data quality flags provide the user of the data with clear information about actions taken to change the original data.

The procedure for flagging data values to indicate their quality, reliability, or checks which have been carried out, or altering values after checking, filling in data gaps, etc., can vary from project to project, and between different laboratories and data centres.

The purpose of the recommended IODE Quality Flag Standard is to define a common set of quality flags that can be used by data centres and projects.

This document describes a two-level quality flag scheme (QF) that will facilitate the exchange and integration¹ of multi-disciplinary oceanographic and marine meteorological data. The first, or primary, level defines the data quality flags only, while the secondary level complements the first level by providing the justification for the quality flags, based on quality control tests or data processing history.

3. INTENDED AUDIENCE

The scheme is intended for individual researchers, research groups or oceanographic/marine meteorological data centres who manage and/or exchange oceanographic/marine meteorological data.

The scheme does not require existing programs or projects to change their quality flag systems, but aims to provide a scheme for the exchange of data between existing programmes. It may also serve as a quality flag scheme for new projects and programmes.

4. JUSTIFICATION

Quality flag schemes are used to record results of quality control and quality assessment checks and enable users to filter data based upon known quality criteria.

If the proposed scheme was in place, it would enable users to (i) merge different data sets, (ii) retain previous quality information, (iii) add new information on data quality and processing history, and (iv) make informed decisions to accept or reject data depending on quality requirements for the particular application or research.

When data from different sources are exchanged and combined in one data base, existing information on quality of data can be lost because different quality flag schemes are used by different data centres and there is rarely one-to-one mapping.

It is important to include detailed information on quality test results if such information exists, and one-level quality flag schemes do not support this capability.

Advantages of this two-level scheme:

- Small and fixed number of unambiguous flags at the primary level that can be justified by the details in the second level;
- Primary-level flag values are numeric and ordered such that increasing quality flag values indicate a decreasing level of quality. This supports the identification of all data that meet a minimum quality level and assignment of quality flags to calculated parameters;
- The scheme is universal; it can be applied to all types of data enabling exchange and integration of multi-disciplinary data;

¹ "Integration" in this context covers the combination of data from two or more sources into one database as well as using the combined database for the development of data products.

- Existing QF schemes can be mapped to the proposed scheme with no information loss. This is specifically true when information on the applied tests is delivered by data providers;
- Data sets with different QF schemes can be merged into one data set, preserving all existing quality flags and making it possible to apply new quality tests and add the results.

The flag scheme design was based upon an extensive review of existing quality flag schemes (See Annex A). None of the reviewed schemes met all advantages stated above. A detailed comparison between 15 widely used flagging schemes is available from http://odv.awi.de/fileadmin/user_upload/odv/misc/ODV4_QualityFlagSets.pdf

5. THE QUALITY FLAG SCHEME

A two-level quality flag scheme is proposed.

5.1 Primary Level

The first or primary level is composed of five quality values and their definitions (Table 1).

Table 1: Primary level

Value	Primary-level flag short name	Definition
1	Good	Passed documented required QC tests
2	Not evaluated, not available or unknown	Used for data when no QC test performed or the information on quality is not available
3	Questionable/suspect	Failed non-critical documented metric or subjective test(s)
4	Bad	Failed critical documented QC test(s) or as assigned by the data provider
9	Missing data	Used as place holder when data are missing

Note: The quality of verified "Good" (flag 1) is considered higher (smaller flag value) compared to "Not evaluated" (flag 2), as the latter could turn out to be of any quality from good to bad, once the quality checks have been performed. Consequently, the neutral "Not evaluated" (flag 2) is placed between verified "Good" and verified "Questionable/suspect".

The flagging scheme can be applied to any type of data.

The Primary Level is intended for data users that need only basic data quality flags.

The primary level flags are such that increasing flag values indicate decreasing data quality. This is an important property that facilitates data quality filtering and/or processing, including inheritance of quality flag values for derived variables. The quality of a calculated value inherits the lowest quality qualifier of the variables used in the calculation. For example, when we calculate density from temperature (T) and salinity (S), then if T is of “good” quality and S is of “unknown” quality, then density should inherit the “unknown quality”.

Mapping

Mappings for existing and future programmes, including those in Annex A, will be invited and maintained on the ODS web site (<http://www.oceandatastandards.org>) (See Annex B for examples).

5.2 Secondary level

The secondary level complements the primary level flags by reporting the results of specific QC tests performed and data processing history.

The secondary level content varies in number and description and is chosen by those who implement the scheme, representing information on the applied quality tests (e.g., excessive spike check, regional data range check) and data processing history (e.g., interpolated values, corrected values).

Table 2: An example of quality control tests and data processing history

Example quality control tests / data processing history (description)
Globally impossible value
Monthly climatology standard deviation test
Excessive spike check
Excessive offset/bias when compared to a reference data set
Excessive data uncertainty
Unexpected X/Y ratio (e.g., chemical stoichiometry or property-property X to T, S, density, among others)
Excessive spatial gradient or pattern check (“bullseyes”)
Below detection limit of method
Interpolated value (not measured)
Data offset corrected value relative to a reference data
Expert review

The secondary level tests and their results can be specified as needed. While providing the secondary level information is not mandatory, it is highly recommended that the secondary level be used to explain fully the primary level flags. As shown in the example below, the results of many quality tests can be represented by values.

Example: implementation of secondary level. In the example below, the test is identified by a code and the possible outcomes of the test are represented in “values” and their “meanings”.

Test: *World Ocean database 2009 (WOD2009) flags for entire cast as a function of variable*

Test description: *WOD cast test is represented by a one digit number corresponding to one or a combination of two or more quality control tests as described in the WOD2009 documentation (Johnson et al. 2009).*

Values: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

Meanings: 0= accepted cast, 1= failed annual standard deviation check, 2= two or more density inversions (Levitus, 1982 criteria), 3= flagged cruise, 4= failed seasonal standard deviation check, 5= failed monthly standard deviation check, 6= failed annual and seasonal standard deviation check, 7= “bullseye” from standard level data or failed annual and monthly standard deviation check, 8= failed seasonal and monthly standard deviation check, 9= failed annual, seasonal, and monthly standard deviation check.

Mapping

Existing and future programmes will be invited to provide their Secondary Level Table (including chosen codes, description of quality control tests / data processing history, values and meanings) which will be maintained on the ODS web site (<http://www.oceandatastandards.org>).

6. BIBLIOGRAPHY

Johnson, D.R., T.P. Boyer, H.E. Garcia, R.A. Locarnini, O.K. Baranova, and M.M. Zweng, 2009. World Ocean Database 2009 Documentation. Edited by Sydney Levitus. NODC Internal Report 20, NOAA Printing Office, Silver Spring, MD, 175pp. Available at http://www.nodc.noaa.gov/OC5/WOD09/pr_wod09.html

ANNEX A

Selection of representative examples of existing QC flag schemes

Examples are provided starting from a simple QC flag scheme and increasing through multiple levels of complexity. Additional examples can be viewed on http://odv.awi.de/fileadmin/user_upload/odv/misc/ODV4_QualityFlagSets.pdf

ODV generic quality flags

Code	Description
0	Good
1	Unknown
4	Questionable
8	Bad

QARTOD quality flags – proposed, not yet definitive

Code	Description
0	Quality not evaluated
1	Bad
2	Questionable/suspect
3	Good
9	Missing data

OceanSITES quality flags – proposed, not yet definitive

Code	Description
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
7	Nominal value
8	Interpolated value
9	Missing value

GTSP quality flags

Code	Description
0	No quality control has been assigned
1	QC was performed; appears to be correct
2	QC was performed; probably good

3	QC was performed; appears doubtful
4	QC was performed; appears erroneous
5	The value was changed as a result of QC
9	The value is missing

SeaDataNet quality flags

Code	Description
0	No quality control
1	Good value
2	Probably good value
3	Probably bad value
4	Bad value
5	Changed value
6	Value below detection
7	Value in excess
8	Interpolated value
9	Missing value
A	Value phenomenon uncertain

BODC

Code	Description
<	Below detection limit
>	In excess of quoted value
A	Taxonomic flag for affinis (aff.)
B	Beginning of CTD down/up cast
C	Taxonomic flag for confer (cf.)
D	Thermometric depth
E	End of CTD down/up cast
H	Extrapolated value
I	Taxonomic flag for single species (sp.)
K	Improbable value, unknown QC source
L	Improbable value, originators QC
M	Improbable source, BODC QC
N	Null value
O	Improbable value, user QC
P	Trace/calm
Q	Indeterminate
R	Replacement value
S	Estimated value
T	Interpolated value
U	Uncalibrated
W	Control value

X	Excessive difference
---	----------------------

NODC WOD (observed levels)

Code	Description
0	Accepted value
1	Range outlier
2	Failed inversion check
3	Failed gradient check
4	Observed level “bullseye” flag and zero gradient check
5	Combined gradient and inversion checks
6	Failed range and inversion checks
7	Failed range and gradient checks
8	Failed range and questionable data checks
9	Failed range and combined gradient and inversion checks

ANNEX B

Examples of mapping the IODE quality flag scheme to existing quality flag schemes.

See examples for ODV (Table B.1), WOCE (Table B.2) water sample quality flag definitions (http://cchdo.ucsd.edu/manuals/pdf/90_1/chap4.pdf), GTSP (IOC) (Table B.3), and SDN (Table B.4) below.

Table B.1. An example of mapping the ODV scheme to the proposed quality flag scheme.

ODV scheme		Proposed quality flag scheme			Comments
Flag code	Flag description	Primary-level flag code	Primary-level flag description	Secondary-level flag description (held in a code table)	
0	Good	1	Good	Expert review	Unless the exact list of quality checks is provided
1	Unknown	2	Not evaluated, not available or unknown	Expert review	Unless the exact list of quality checks is provided
4	Questionable	3	Questionable	Expert review	Unless the exact list of quality checks is provided
8	Bad	4	Bad	Expert review	Unless the exact list of quality checks is provided

Table B.2. An example of mapping the WOCE water sample scheme to the proposed quality flag scheme.

WOCE water sample		Proposed quality flag scheme			Comments
Flag code	Flag description	Primary-level flag code	Primary-level flag description	Secondary-level flag description (held in a code table)	
1	Sample for this measurement	9	Missing data	Sample was collected, but	

	was drawn from water bottle, but analysis not received			analysis not received due to unknown reason	
2	Acceptable measurement	1	Good	Expert review	Unless the exact list of quality checks is provided
3	Questionable measurement	3	Questionable	Expert review	Unless the exact list of quality checks is provided
4	Bad measurement	4	Bad	Expert review	Unless the exact list of quality checks is provided
5	Data were expected to be measured, but the observation is missing due to sample loss, contamination, etc.	9	Missing data	Sample was collected, but the observation is missing due to sample loss, contamination, etc.	
6	Mean of replicate measurements	2	Not evaluated, not available or unknown	Mean of replicate measurements	Because no information was provided on the quality of the replicate measurements, this can only be mapped to 2 in the new scheme. However, once checks are applied and the data are considered good then the primary flag can be changed to 1 and secondary flags are added.
7	Manual chromatographic peak measurement	2	Not evaluated, not available or unknown	Manual chromatographic peak measurement	

8	Irregular digital chromatographic peak integration	2	Not evaluated, not available or unknown	Irregular digital chromatographic peak integration	
9	Sample not drawn for this measurement from this bottle	9	Missing data	Sample not collected for this measurement	

Table B.3. An example of mapping the GTSP scheme to the proposed quality flag scheme.

GTSP scheme		Proposed quality flag scheme			Comments
Flag code	Flag description	Primary-level flag code	Primary-level flag description	Secondary-level flag description (held in a code table)	
0	No quality control has been assigned	2	Not evaluated, not available or unknown	Sample collected but QC tests were not applied	
1	QC was performed; appears to be correct	1	Good	Expert review	Unless the exact list of quality checks is provided
2	QC was performed; probably good	1	Good	Expert review	Unless the exact list of quality checks is provided
3	QC performed; appears doubtful	3	Questionable	Expert review	Unless the exact list of quality checks is provided
4	QC performed; appears erroneous	4	Bad	Expert review	Unless the exact list of quality checks is provided
5	The value was changed as a result of QC	1	Good	Changed value; Expert review	Once checks are applied, secondary flags are added.
9	The value is missing	9	Missing data	Not reported	

Table B.4. An example of mapping the SDN scheme to the proposed quality flag scheme.

SDN scheme		Proposed quality flag scheme			Comments
Flag code	Flag description	Primary-level flag code	Primary-level flag description	Secondary-level flag description (held in a code table)	
0	No quality control	2	Not evaluated, not available or unknown	Sample collected but QC tests were not applied	
1	Good value	1	Good	Expert review	Unless the exact list of quality checks is provided
2	Probably good value	1	Good	Expert review	Unless the exact list of quality checks is provided
3	Probably bad value	3	Questionable	Expert review	Unless the exact list of quality checks is provided
4	Bad value	4	Bad	Expert review	Unless the exact list of quality checks is provided
5	Changed value	1	Good	Changed value; Expert review	Once checks are applied, secondary flags are added.
6	Value below detection	4	Bad	Detection limit	
7	Value in excess	4	Bad	Excess limit	
8	Interpolated value	1	Good	Interpolated value	
9	The value is missing	9	Missing data	Not reported	
A	Value phenomenon uncertain	3	Questionable	Expert review	Unless the exact list of quality checks is provided

**Intergovernmental Oceanographic
Commission (IOC)**

United Nations Educational, Scientific and
Cultural Organization
1, rue Miollis, 75732 Paris Cedex 15, France
Tel: + 33 1 45 68 39 83
Fax: +33 1 45 68 58 12
<http://ioc.unesco.org>

IOC Project Office for IODE

Wandelaarkaai 7/61
8400 Oostende, Belgium
Tel: +32 59 34 21 34
Fax: +32 59 34 01 52
<http://www.iode.org>