DATA PRODUCT SPECIFICATION FOR GLOBAL RANGE TEST

Version 1-01
Document Control Number 1341-10004
2012-05-23

Consortium for Ocean Leadership
1201 New York Ave NW, 4th Floor, Washington DC 20005
www.OceanLeadership.org

in Cooperation with

University of California, San Diego
University of Washington
Woods Hole Oceanographic Institution
Oregon State University
Scripps Institution of Oceanography
Rutgers University

Document Control Sheet

Version	Date	Description	Author
0-01	2011-11-04	Initial draft	M. Lankhorst
0-02	2011-11-17	Updated format and modified for scalar instead of vector inputs	S. Webster
0-03	2011-12-14	Renamed to Data Product Specification and updated to match Data Product Spec Outline.	S. Webster
0-04	2012-01-06	Updated based on comments from focused review.	S. Webster
0-05	2012-03-28	Removed Appendix B Reference Tables because QC look-up tables will be kept separate from the DPS.	S. Webster
0-06	2012-04-16	Re-inserted App. B as exemplar table, following request from CI. Addressed focused review comments.	M. Lankhorst
0-07	2012-04-20	Updated exemplar QC lookup table and moved to Test Data section 4.6.	S. Webster
0-08	2012-04-24	Updated exemplar lookup table	M. Lankhorst
0-09	2012-04-25	Minor updates.	S. Webster
1-00	2012-05-22	Initial Release	E. Chapman
1-01	2012-05-23	Formatting, copy edits	E. Griffin

Signature Page

This document has been reviewed and approved for release to Configuration Management.

OOI Chief Systems Engineer: ______________________________

Date:2012-05-22

This document has been reviewed and meets the needs of the OOI Cyberinfrastructure for the purpose of coding and implementation.

OOI CI Signing Authority: ______________________________

Date: 2012-05-22

Table of Contents

1 Abstract

This document describes the OOI Global Range Test automated quality control algorithm used on various OOI data products. This automated algorithm generates flags for data points according to whether they fall within a given range. This range is thought of as a universally valid range, hence the name “global.”

2 Introduction

2.1 Author Contact Information

Please contact Matthias Lankhorst (mlankhorst@ucsd.edu) or the Data Product Specification lead (DPS@lists.oceanobservatories.org) for more information concerning the algorithm and other items in this document.

2.2 Metadata Information

2.2.1 Data Product Name

n/a

2.2.2 Data Product Abstract (for Metadata)

n/a

2.2.3 Computation Name

The name for this quality control algorithm is

- Global Range Test (GLBLRNG)

2.2.4 Computation Abstract (for Metadata)

The OOI Global Range Test quality control algorithm generates a QC flag for a data point indicating whether it falls within a given range.

2.2.5 Instrument-Specific Metadata

n/a

2.2.6 Synonyms

n/a

2.2.7 Similar Algorithms

OOI will use another range test as well, the Local Range Test, which utilizes site-specific and, possibly, time-varying ranges by which the data are judged.

2.3 Instruments

This algorithm is applied to OOI data products as per the table in the appendix.

2.4 Literature and Reference Documents

DCN 1342-000xx	**Instrument-specific Processing Flow documents** contain flow diagrams detailing all of the specific algorithms (product, QA and calibration, QC) necessary to compute all data products from the instrument at all levels of QA and QC and the order that the algorithms must be applied

2.5 Terminology

2.5.1 Definitions

n/a

2.5.2 Acronyms, Abbreviations and Notations

General OOI acronyms, abbreviations and notations are contained in the Level 2 Reference Module in the OOI requirements database (DOORS). There are no other acronyms, abbreviations, or notations for this document.

2.5.3 Variables and Symbols

x,`x`	input scalar data value
lim,`lim`	limits defining the valid range of input values
qc,`qcflag`	output scalar QC flag

3 Theory

3.1 Description

The algorithm output indicates whether the data value falls into a given range (qc=1). If a data value does not (qc=0), the user should assume that this data value is of bad quality and should not be used for scientific analysis. The algorithm output provides a simple means for the user to select which data to use.

3.2 Mathematical Theory

Let $\lim \in \mathbb{R}^2$ such that lim(1)<lim(2).

Then, define $f: \mathbb{R} \to \{0,1\}, x \mapsto qc$ such that:

qc=0 for x<lim(1) or x>lim(2),
qc=1 otherwise.

3.3 Known Theoretical Limitations

n/a

3.4 Revision History

n/a

4 Implementation

4.1 Overview

The algorithm is a simple application of the operators > and <. The max lim(2) and min lim(1) values to use with this function for various data products are given in reference table in Appendix B. Note that the MatLab code in Appendix A accepts vector inputs but the OOI implementation is written for scalar input/output.

4.2 Inputs

`dprd_in`: Input data product. 7-digit identified.
`x`: Input data. Scalar double.
`lim`: Limits defining valid range. Array of class double, size 2-by-1 or 1-by-2. (From QC lookup table.)
`dprd_flag`: Data product that is to be flagged based on the input (From QC lookup table.)

4.3 Processing Flow

Implement the code given in Appendix A.

This code includes error handling.

4.4 Outputs

The outputs of the conductivity compressibility computation are

- `qcflag`: Scalar of class logical. The convention is that data within the valid range are flagged "1", and those outside as "0".

The metadata that must be included with the output are

- The ranges ("`lim`") used for the computation
- An identifier/link that relates the output to the data product (`dprd_flag`) that the QC flags describe

4.5 Computational and Numerical Considerations

n/a

4.6 Code Verification and Test Data Sets

The algorithm code will be verified using the test data set provided, which contains inputs and their associated correct outputs. CI will verify that the algorithm code is correct by checking that the algorithm pressure output, generated using the test data inputs, is identical to the test data output.

Table 1:Test Data Set

x	**lim**	**qcflag**
9	[10 20]	0
10	[10 20]	1
16	[10 20]	1
17	[10 20]	1
18	[10 20]	1
19	[10 20]	1
20	[10 20]	1
25	[10 20]	0

In addition to the output (qcflag), the metadata from the QC Lookup table must be included with the output.

Table 2: Example Global Range Lookup Table

Data Product In (dprd_in)	Data Product Flagged (dprd_flag)	Units (of dprd_in)	Min Value (lim(1))	Max Value (lim(2))	Reference	Reference Person	Reference Date
CONDWAT	CONDWAT	S/m	0	66000	IOC 2010	A. Smith	2012/01/01
PRESWAT	PRESWAT	Pa	0	58842000	IOC 2010	A. Smith	2012/01/01
TEMPWAT	TEMPWAT	degrees C	-2	40	IOC 2010	A. Smith	2012/01/01
PRACSAL	PRACSAL	unitless	0	42	IOC 2010	A. Smith	2012/01/01
DENSITY	DENSITY	kg/m^3	1000	1100	IOC 2010	A. Smith	2012/01/01
OPTATTS	OPTATTS	1/m	0.001	10	TBD	B. Smith	2012/01/01
OPTABSN	OPTABSN	1/m	0.001	10	TBD	B. Smith	2012/01/01
OPTPARW	OPTPARW	µmol photons/ m^2*s	0	19999	TBD	C. Smith	2012/01/01
TRHPHRS	TRHPHRS	Ohm, raw voltage	0		TBD	D. Smith	2012/01/01
TRHPHCC	TRHPHCC	µM	0	210	TBD	D. Smith	2012/01/01
TRHPHEH	TRHPHEH	mV	0	500	TBD	D. Smith	2012/01/01
TRHPHTE	TRHPHTE	degrees C	-2	300	TBD	D. Smith	2012/01/01
WAVSTAT significant wave height T_{sig}	WAVSTAT significant wave height T_{sig}	m	0	30	TBD	Lankhorst	2012/04/24
WAVSTAT significant wave height T_{sig}	WAVSTAT Mean direction D of waves	m	0.5	999	TBD	Lankhorst	2012/04/24
WAVSTAT significant wave height T_{sig}	WAVSTAT direction spread DS of waves	m	0.5	999	TBD	Lankhorst	2012/04/24
VELPROF horizontal velocities u, v	VELPROF horizontal velocities u, v	m/s	0	4	TBD	Lankhorst	2012/04/24
VELPROF “Error velocity”	VELPROF horizontal velocities u, v	m/s	0	0.1	TBD	Lankhorst	2012/04/24
VELPROF “Percent good threshold”	VELPROF horizontal velocities u,v	%	20	100	TBD	Lankhorst	2012/04/24

Note that this table is for example purposes only and some/all values may not be correct. The official QC lookup tables are kept separately from the DPS and, at the time of writing, do not exist in their final form.

Appendix A Example Code

The following routine is example code run under MatLab:

```
% DATAQC_GLOBALRANGETEST   Data quality control algorithm testing
%      if measurements fall into a user-defined valid range.
%      Returns 1 for presumably good data and 0 for data presumed bad.
%
% Time-stamp: <2010-07-28 15:16:00 mlankhorst>
%
% USAGE:   out=dataqc_globalrangetest(dat,validrange);
%
%          out: Boolean, 0 if value is outside range, else 1.
%          dat: Input dataset, any scalar, vector, or matrix.
%               Must be numeric and real.
%          validrange: Two-element vector with the minimum and
%               maximum values considered to be valid
%
% EXAMPLE:
%
%     >> x=[17 16 17 18 25 19];
%     >> qc=dataqc_globalrangetest(x,[10 20])
%
%     qc =
%
%          1     1     1     1     0     1
%
%

function out=dataqc_globalrangetest(dat,datlim);

  if ~isnumeric(dat)
    error('DAT must be numeric.')
  end

  if ~all(isreal(dat(:)))
    error('DAT must be real.')
  end

  if ~isnumeric(datlim)
    error('VALIDRANGE must be numeric.')
  end

  if ~all(isreal(datlim(:)))
    error('VALIDRANGE must be real.')
  end

  if length(datlim)~=2
    error('VALIDRANGE must be two-element vector.')
  end

  datlim=[min(datlim(:)) max(datlim(:))];

  out=(dat>=datlim(1))&(dat<=datlim(2))
```