PML | Plymouth Marine Laboratory

Listen to the ocean

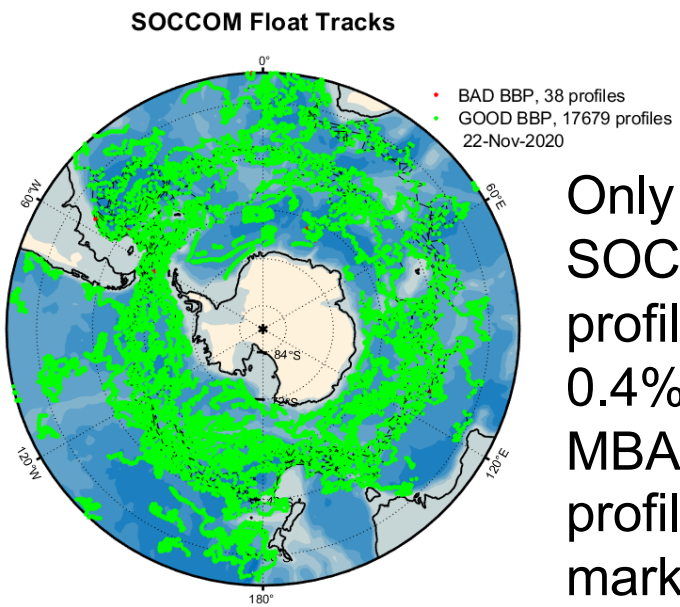
RTQC of BBP data

Giorgio Dall'Olmo
with inputs from Tanya Mauer, Catherine Schmechtig, Raphaëlle Sauzède

MBARI BBP Real-Time Processing

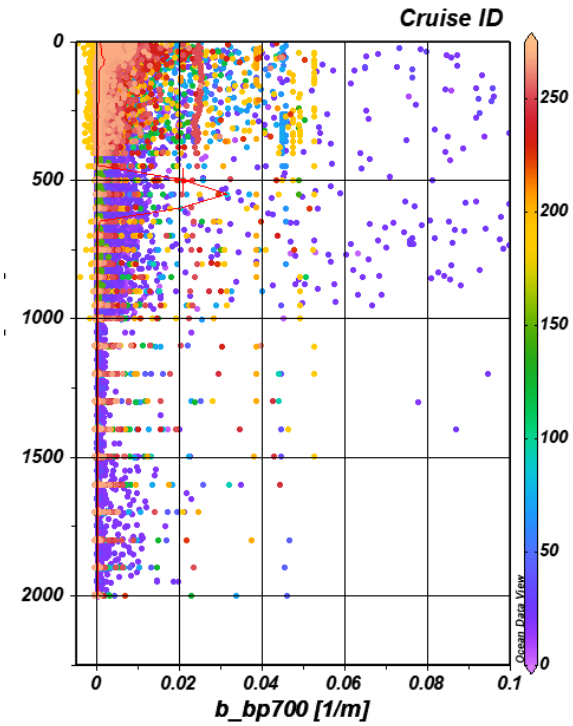
Gross range checks only
BBP700 in range [-0.000025 0.1] m^{-1} (BGC-Argo quality control Manual for BGC data)
BBP532 in range [-0.000005 0.1] m^{-1} (BGC-Argo quality control Manual for BGC data)

Proposal: MBARI wants to populate BBP700_ADJUSTED with BBP700 in REAL-TIME

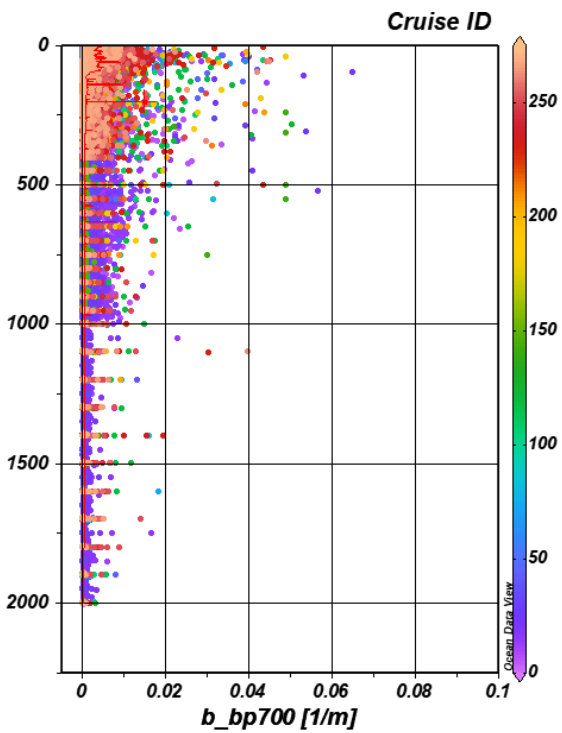


Only 0.2% of SOCCOM profiles (and 0.4% of all MBARI profiles) are marked bad for BBP

No DMQC:



With DMQC:

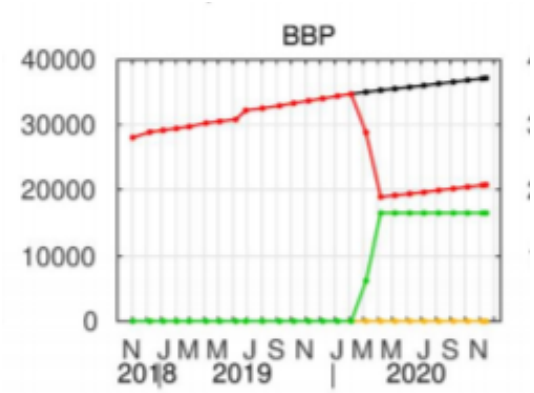


Q: What is DMQC?

LOV: Pushing BBP700 published data into Coriolis

For Floats dead before the 1st January 2019 (Bellacicco et al., 2019) ~ 94 floats (CTS4 and CTS5)
=> production of BD files (routines available on github : https://github.com/qjutard/chl_bbp_ttt)

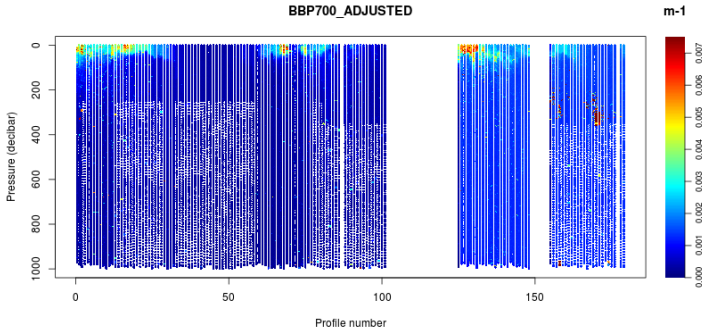
- ✓ From greylist BBP700_ADJUSTED_QC => 3,4
- ✓ Not in range test [-0.000005 ; 0.1] => BBP700_ADJUSTED_QC=4
- ✓ Report BBP700 into the BBP700_ADJUSTED (except for 1 with an OFFSET, 4 more with an offset but not in BD)
- ✓ BBP700_ADUSTED_ERROR =MAX(3counts*SCALE, 20%(BBP700)) => Will be improved with Giorgio's work
- ✓ Final check of the time series (routines available on github : https://github.com/qjutard/time_series_plot)



Example for 6901485, profile 171

- ✓ SCIENTIFIC_CALIB_COMMENT : BBP700 delayed mode adjustment following the work done by M. Cornec in **Bellacicco et al. 2019** (http://dx.doi.org/10.1029/2019GL084078)
- ✓ SCIENTIFIC_CALIB_COEFFICIENT : BBP700_OFFSET = -0.001026
- ✓ SCIENTIFIC_CALIB_EQUATION : BBP700_ADJUSTED = BBP700-BBP700_OFFSET

Courtesy of Henry, BBP at Coriolis February 2020



Q: DMQC LOV == DMQC MBARI?

RTQC for BBP

Problem

- We do not have well defined RTQC and DMQC methods for BBP
- Two untested (AFAIK) tests in the “BGC-Argo QC Manual for BGC data”

Objective

- Generate a **consistent BBP dataset**
- Target audience: **non-expert users**
- Develop automatic RTQC tests for BBP700 (for now)
- Test the tests on the existing BGC-Argo profiles

PROPOSED TESTS FOR RTQC OF BBP700

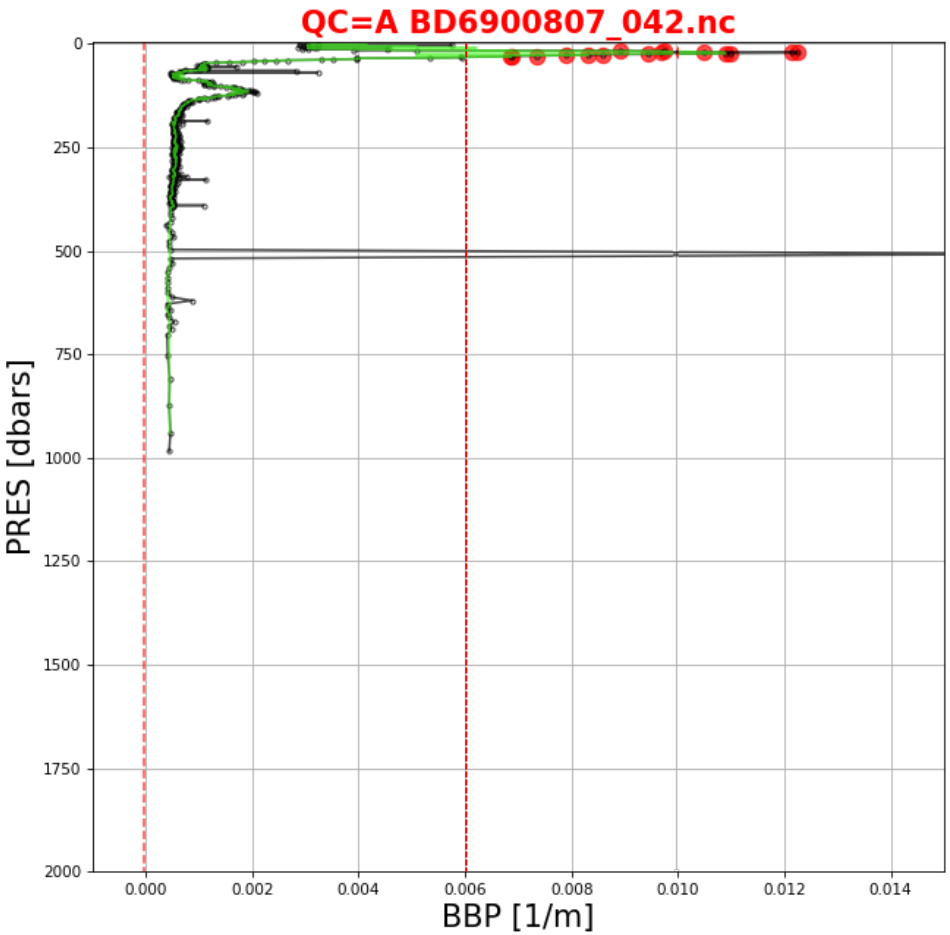
RTQC test that only flags some points

GLOBAL-RANGE TEST

Objective:
To detect and flag values of BBP700 that are outside the expected range.

What is done:
Check that:
BBP700 is within `[0, 0.006]` m^{-1}

QC flag if test fails 3



Comments

- Mostly flags spikes (could be added back)
- Negative profiles
- Previous min value (-0.000025 m^{-1}) was physically unrealistic
- Previous max value (0.1 m^{-1}) was greater than saturation for some instruments

RTQC test that flags entire profile

NOISY-PROFILE TEST

Objective:
To detect and flag profiles of BBP that are affected by noisy data.

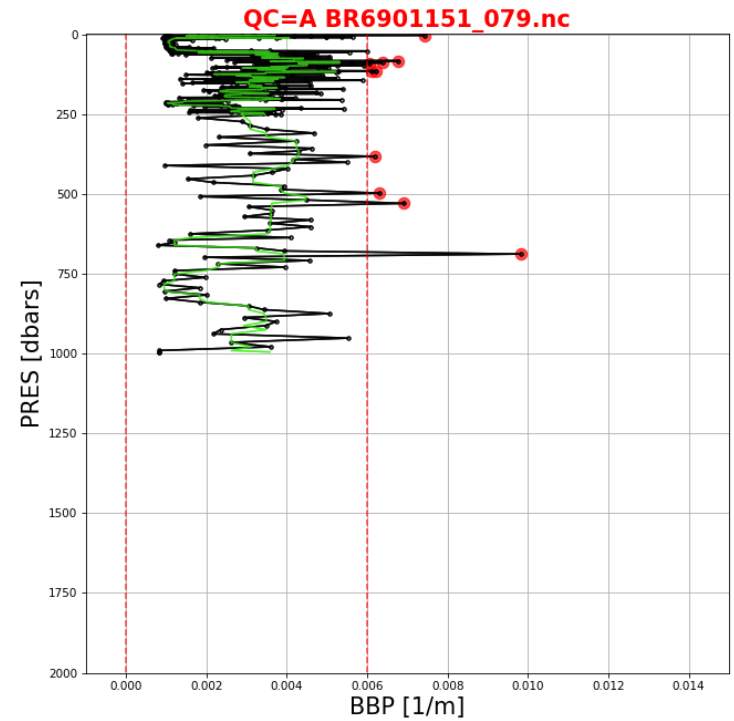
What is done:
Compute
`res = abs(BBP700-medfilt1(BBP700, 31))`

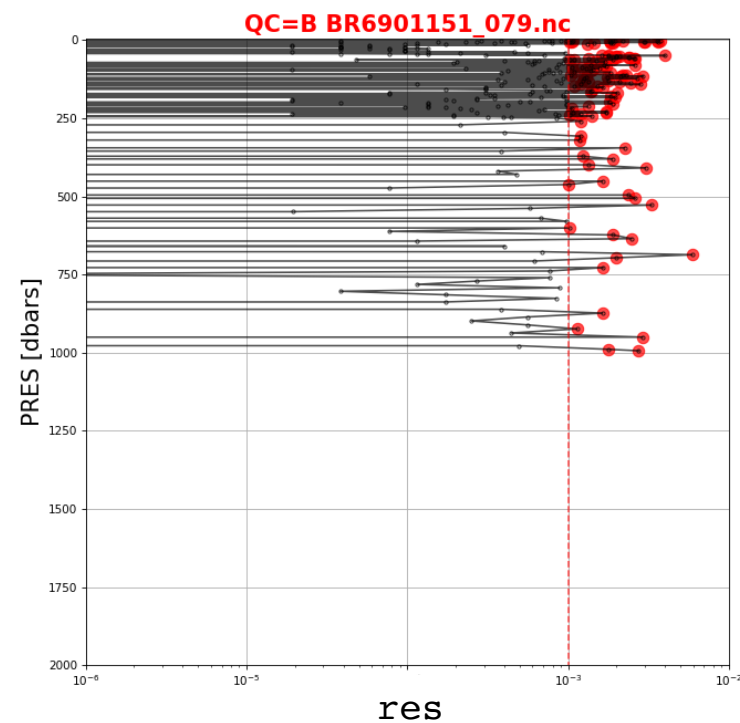
Flag profiles where at least `15%` of the profile has `res > 0.001` m^{-1}.

QC flag if test fails 3

Comments
- Medfilt1 as recommended by Christina for CHLA (variable filter window)

Based on `diff(PRES)`
Where `diff(PRES)<1, win = 11`
Where `1≤diff(PRES)≤3, win = 7`
Where `diff(PRES)>3, win = 3`





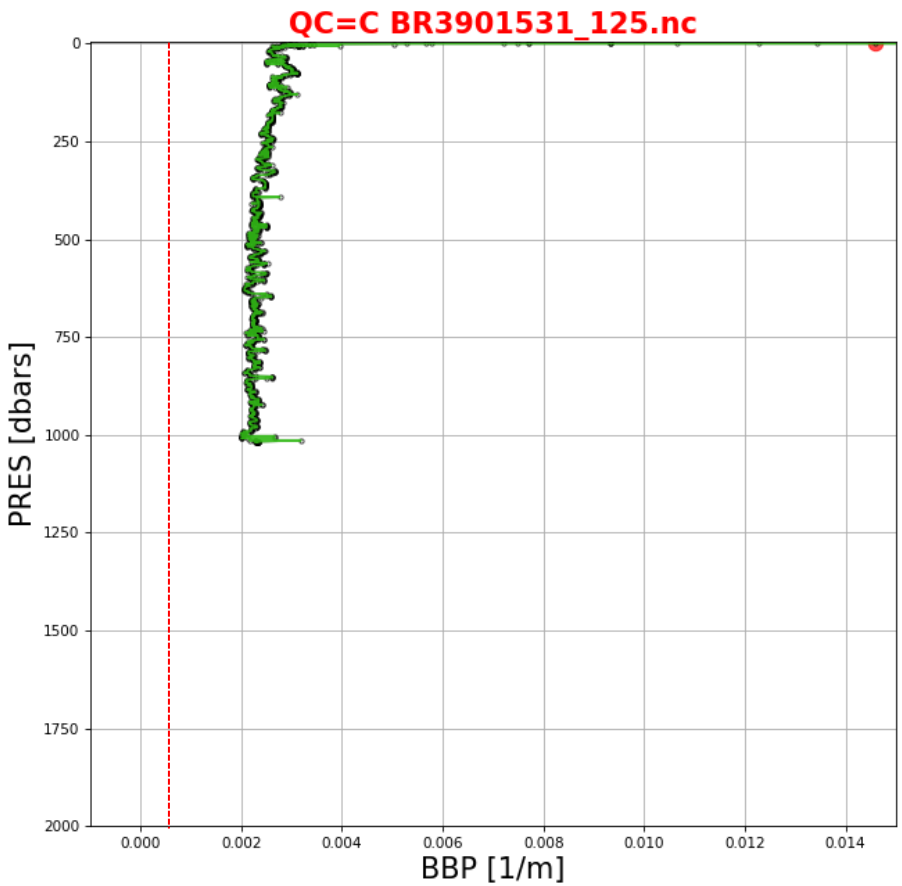
RTQC test that flags entire profile

HIGH-DEEP-VALUE TEST

Objective:
To detect and flag profiles of BBP that have anomalously high values at depth.

What is done:
Check if `median(BBP700)` below 800 dbars is above a threshold of `0.0005` m^{-1} (this is half the value typical for surface bbp in the oligotrophic ocean, smoothed bbp data at depth are expected to be lower than this value).

QC flag if test fails 3



Comments

- It could indicate multiple problems: biofouling, bad sensor, grounding, wrong calibration coefficients, etc.
- If profile is shallower than 800 dbars, test is not performed

RTQC test that flags entire profile

MISSING-DATA TEST

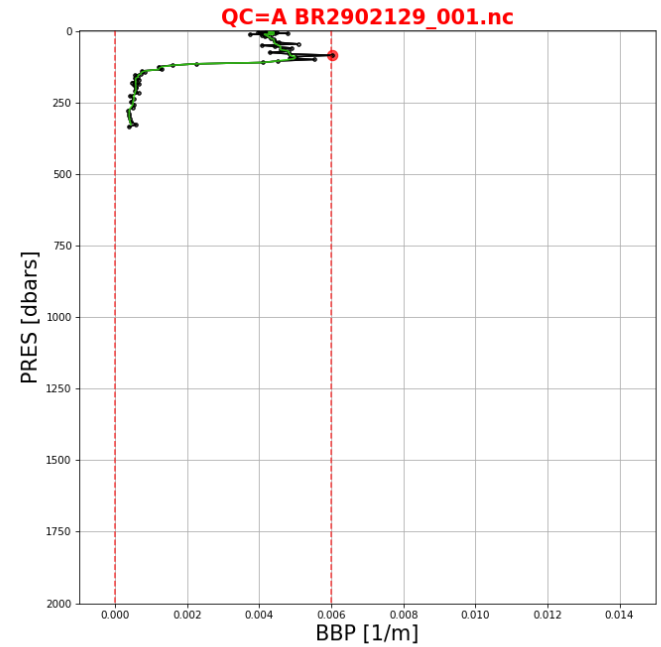
Objective:
To detect and flag profiles of BBP that have a large fraction of missing data. This test can also detect profiles that are too shallow with respect to the Argo mission.

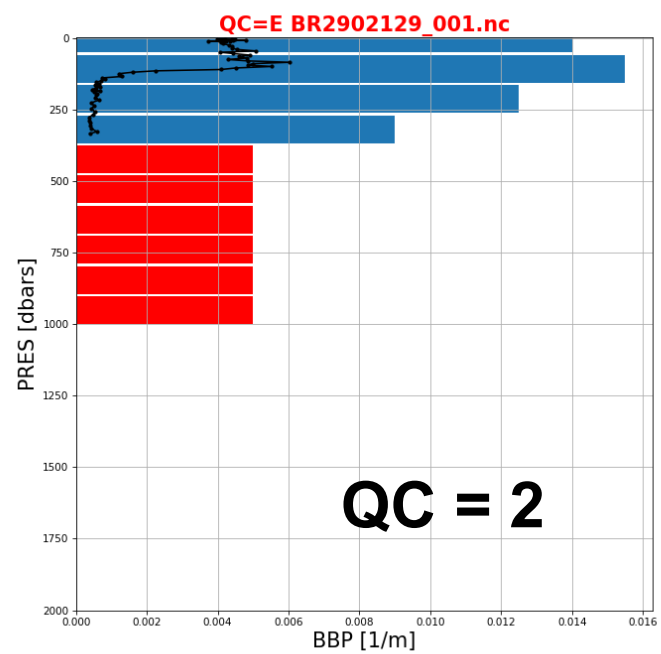
What is done:
Ensure that we have at least 1 measurements every ~100 dbars in the upper 1000 dbars.
Flag entire profile.

QC flag if test fails
2 if shallow profile
3 if less than 1 data point in bins from 0 to 1000 dbars
4 if only data within 1 bin

Comments

- `bins = [ 50, 156, 261, 367, 472, 578, 683, 789, 894, 1000 ]`





RTQC test that flags entire profile

MISSING-DATA TEST

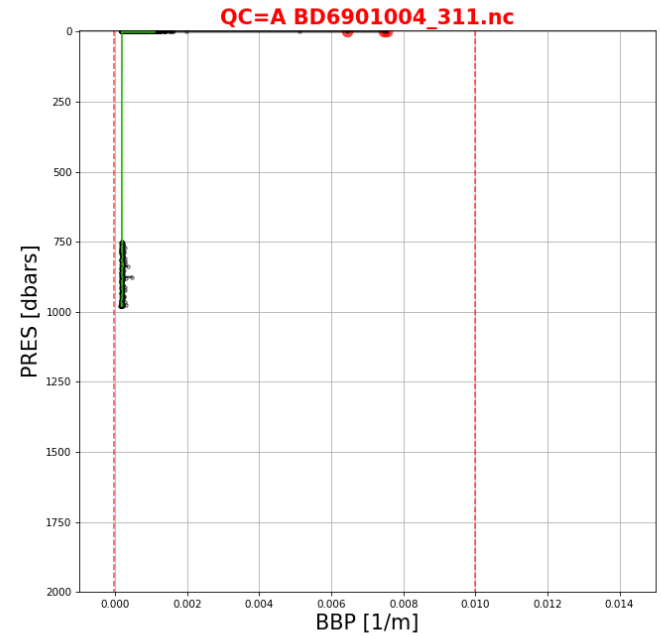
Objective:
To detect and flag profiles of BBP that have a large fraction of missing data. This test can also detect profiles that are too shallow with respect to the Argo mission.

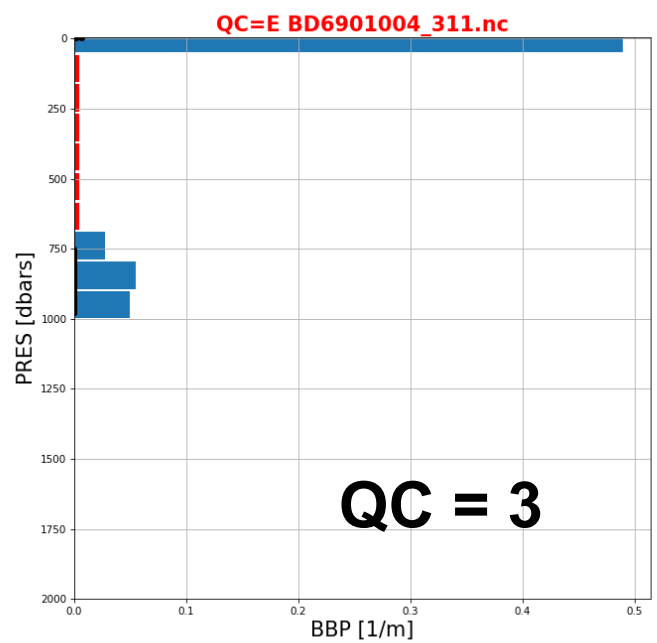
What is done:
Ensure that we have at least 1 measurements every ~100 dbars in the upper 1000 dbars.
Flag entire profile.

QC flag if test fails
2 if shallow profile
3 if less than 1 data point in bins from 0 to 1000 dbars
4 if only data within 1 bin

Comments

- `bins = [ 50, 156, 261, 367, 472, 578, 683, 789, 894, 1000 ]`





RTQC test that flags entire profile

MISSING-DATA TEST

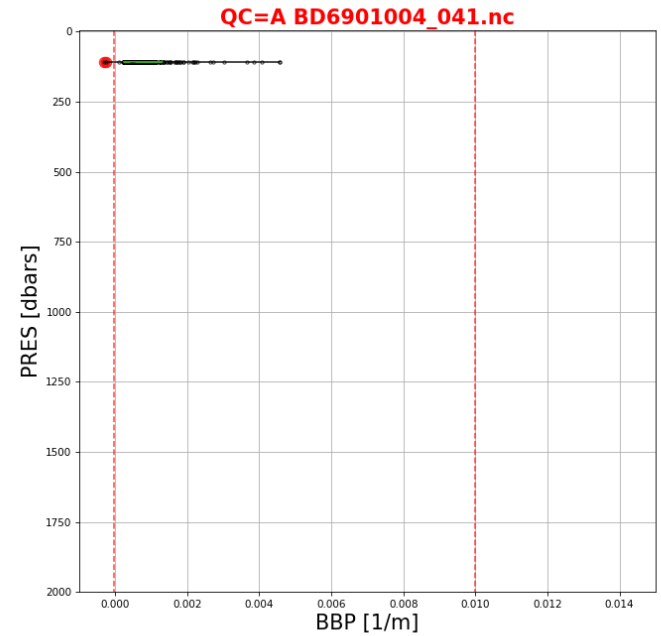
Objective:
To detect and flag profiles of BBP that have a large fraction of missing data. This test can also detect profiles that are too shallow with respect to the Argo mission.

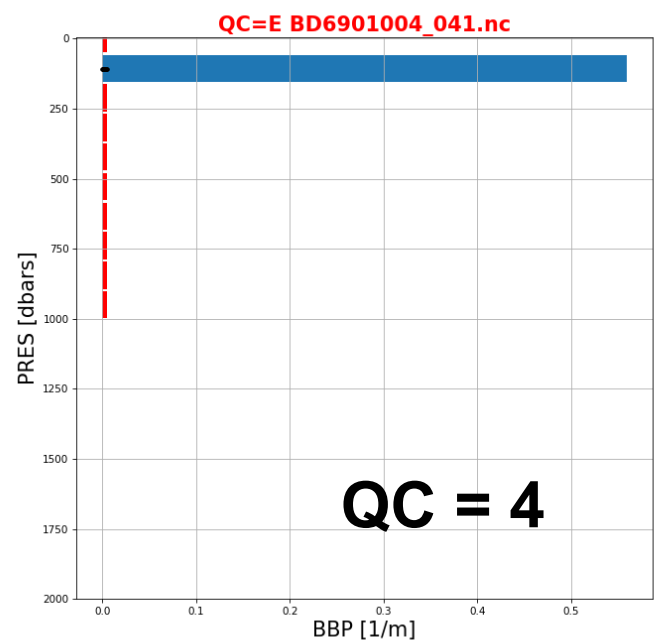
What is done:
Ensure that we have at least 1 measurements every ~100 dbars in the upper 1000 dbars.
Flag entire profile.

QC flag if test fails
2 if shallow profile
3 if less than 1 data point in bins from 0 to 1000 dbars
4 if only data within 1 bin

Comments

- `bins = [ 50, 156, 261, 367, 472, 578, 683, 789, 894, 1000 ]`





Current results

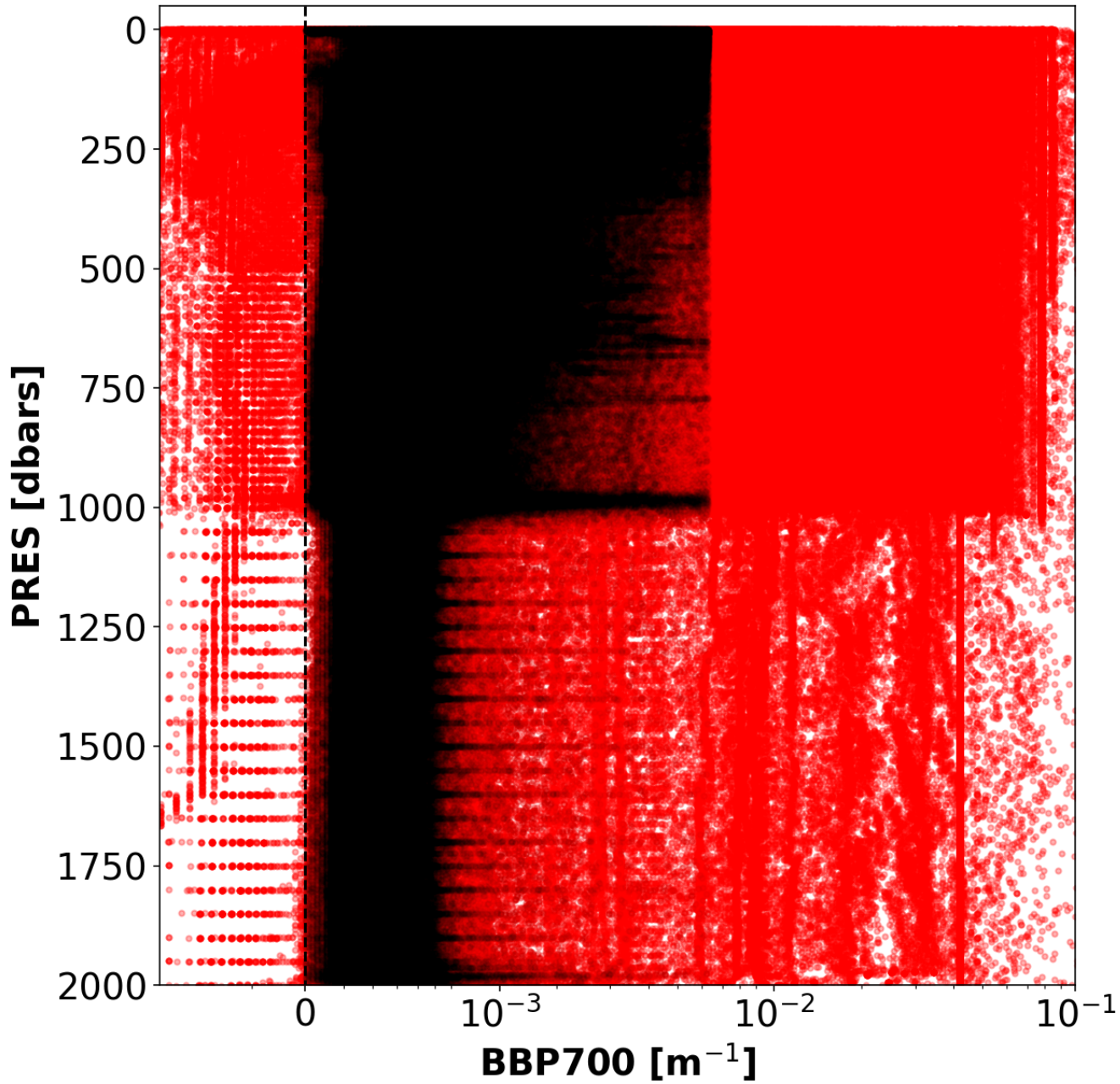
- Tested (visual checks) on ~50 floats from different DACs
- Applied to all 538 floats, 81,745 profiles

- Plot every 100 points

- RED: flagged
- BLACK: not flagged

Global Range: 3.8%
Noisy: 0.6%
High Deep Val: 3.7%
Missing data: 0.7%

Not flagged: 91.1%

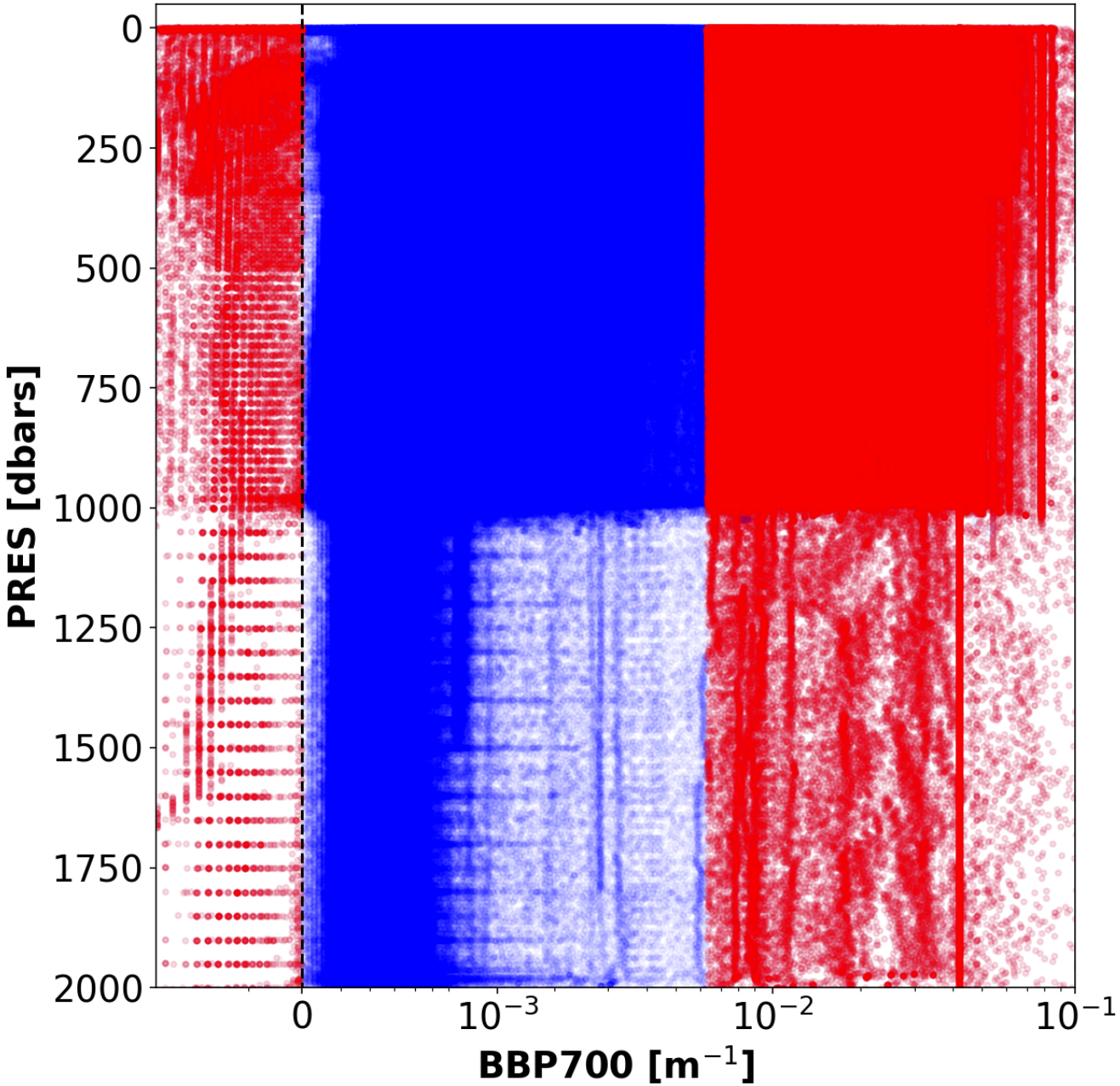


Current results

- Applied to all floats
- Plot all data

- BLUE: QC=0
- RED: failing Global range

Global Range: 3.8%

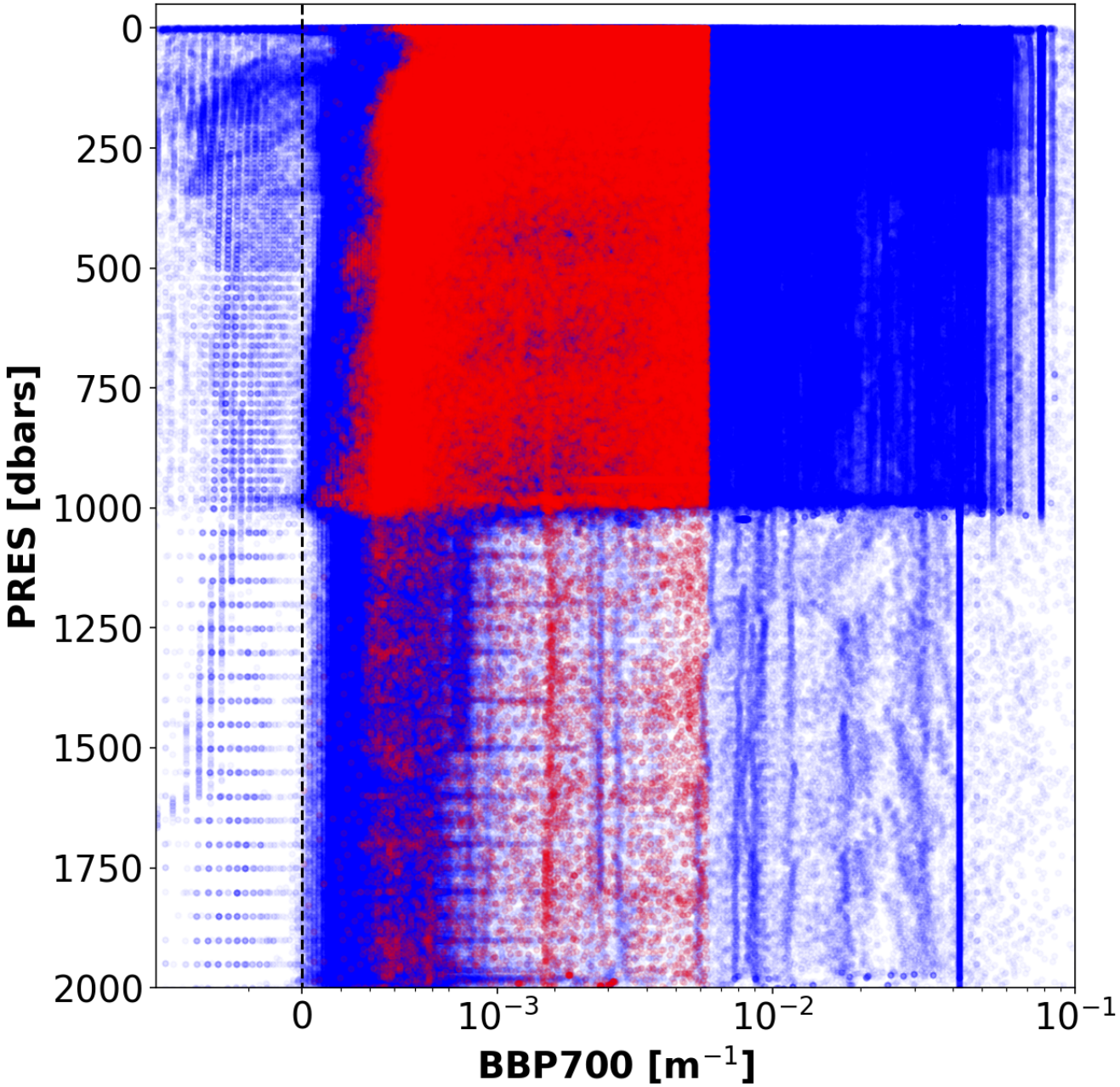


Current results

- Applied to all floats
- Plot all data

- BLUE: QC=0
- RED: failing Noisy profile

Noisy profile: 0.6%

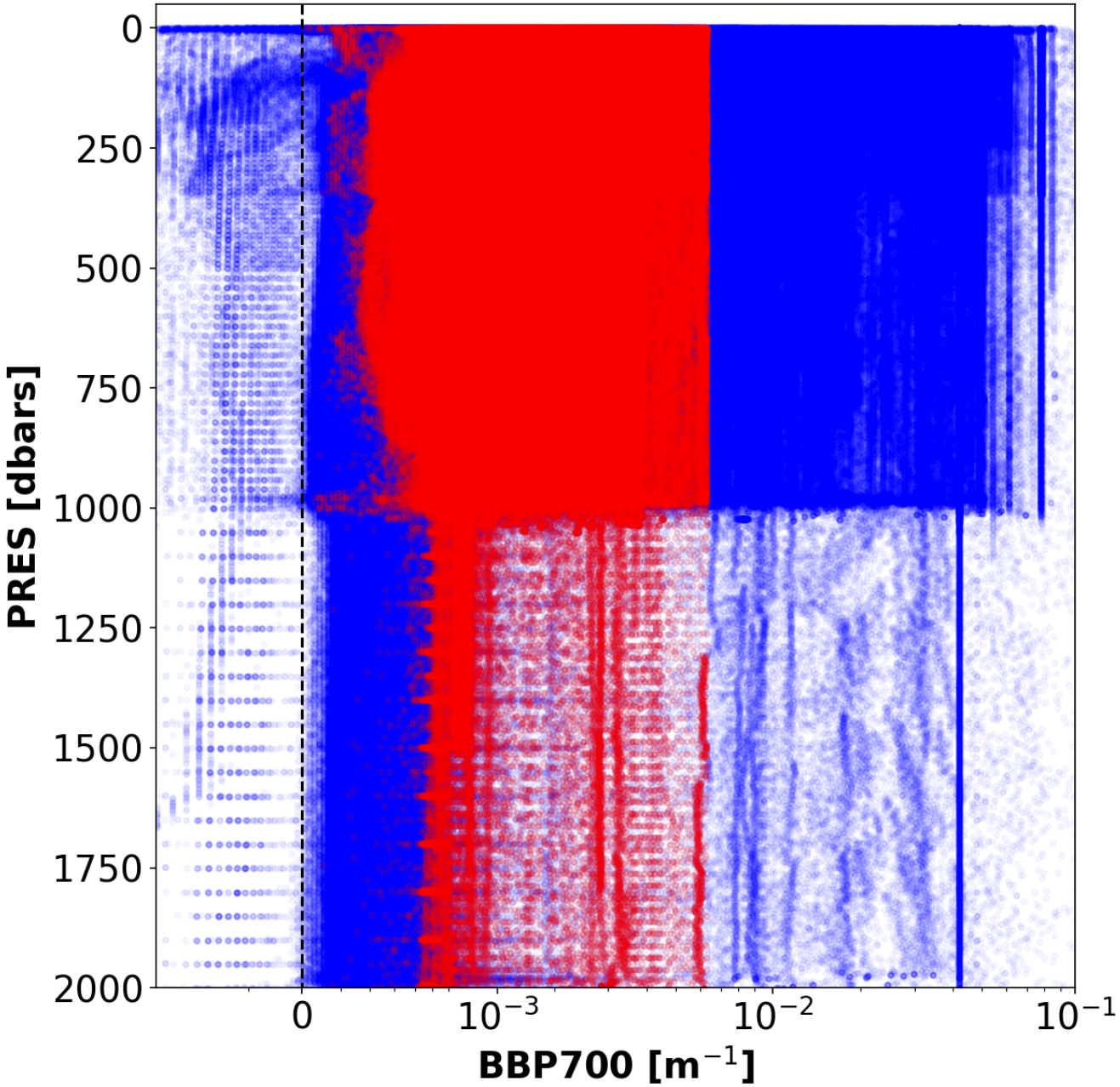


Current results

- Applied to all floats
- Plot all data

- BLUE: QC=0
- RED: failing High Deep Val

High Deep Val : 3.7%

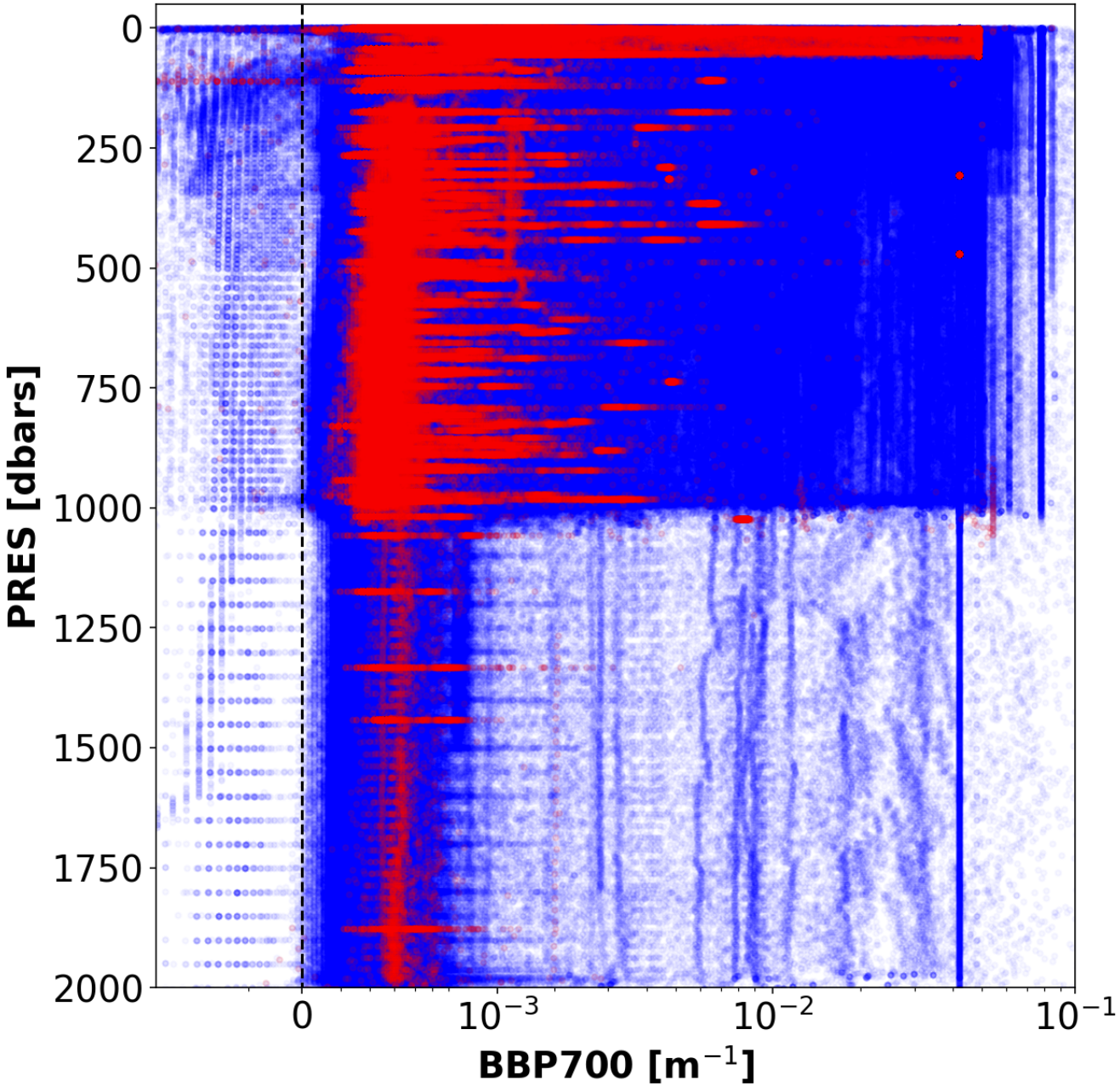


Current results

- Applied to all floats
- Plot all data

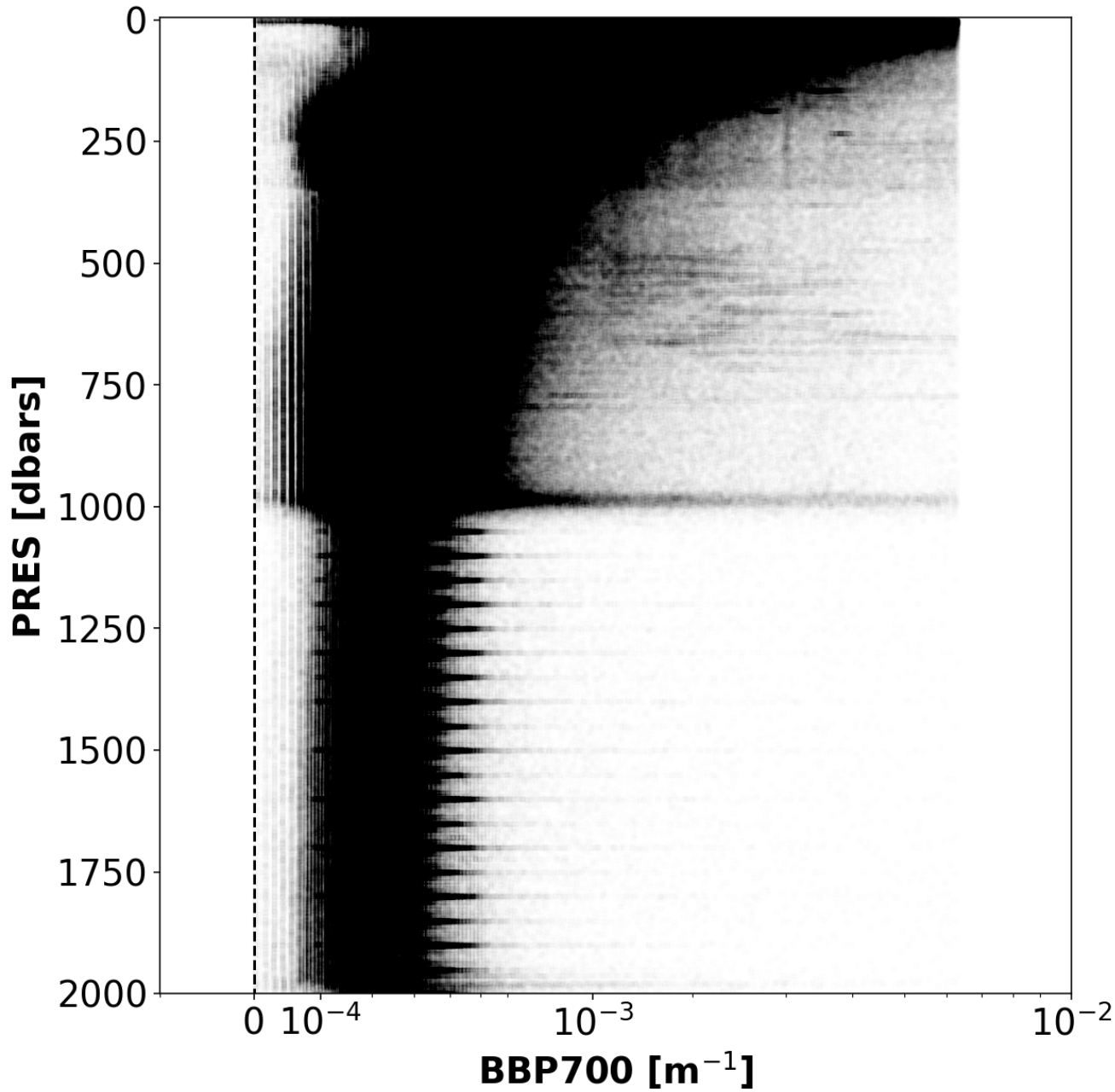
- BLUE: QC=0
- RED: failing Missing data

Missing data (QC>2): 0.7%



Current results

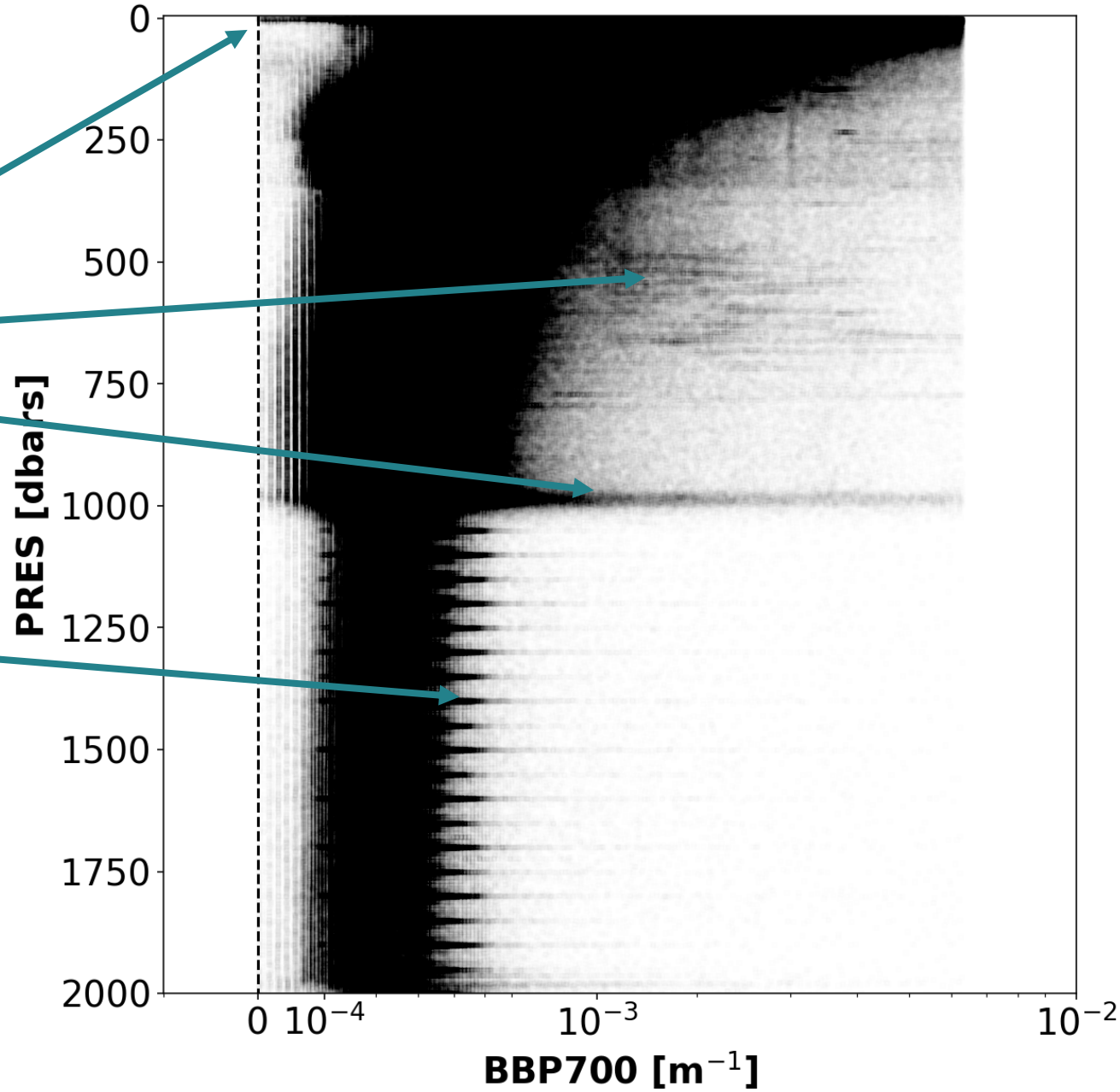
- Tested (visual checks) on ~50 floats from different DACs
- Applied to all 538 floats, 81,745 profiles



Global Range:	3.8%
Noisy profile:	0.6%
High Deep Val:	3.7%
Missing data (QC>2):	0.7%
Not flagged:	91.1%

Additional tests needed

- SURFACE HOOK
- ANIMAL SPIKES
- DEEP HOOK
- STEP CHANGE
- STUCK VALUE
- GROUNDED
- …



Recommendations

The objective of the QC procedures is to screen the raw data to generate a **globally-consistent dataset for non-expert users**

Unless we are willing to reprocess, before DMQC-ing existing data:

- Let's **agree on a set of RTQC tests** (here you have a starting point and some suggestions for additional ones)
- Let's **agree on a common set of DMQC procedures** (I have some suggestions, but see also Raphaëlle's presentation)