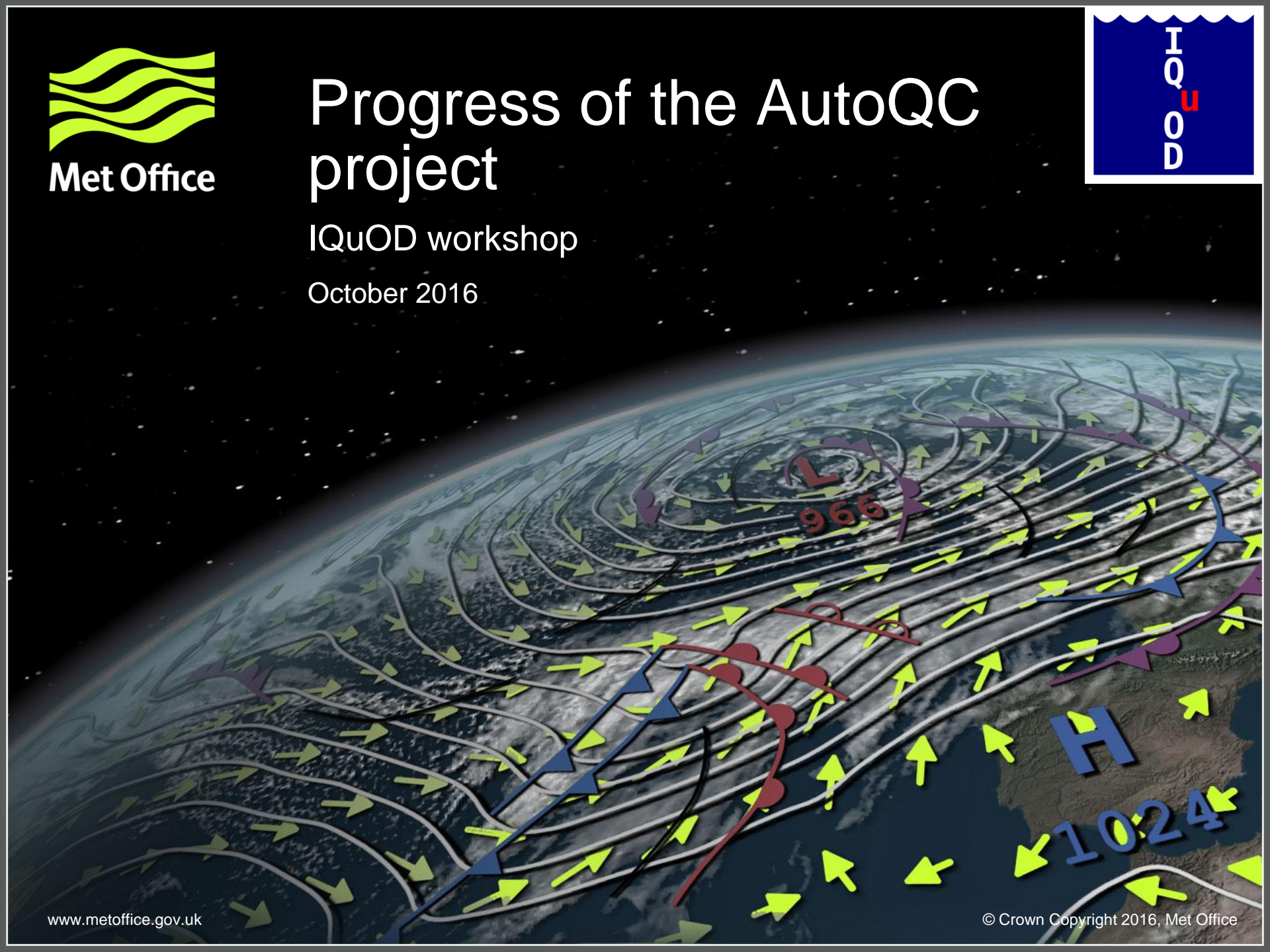
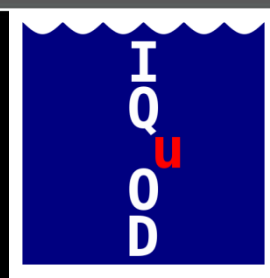




Progress of the AutoQC project

IQuOD workshop

October 2016



What are we trying to do?

Benchmark the automatic quality control tests in use around the world to find an optimum set

The optimum set will be used to:

- 1) Generate an 'AutoQC' IQuOD data release
- 2) Feed profiles to enhanced quality control methods

Background

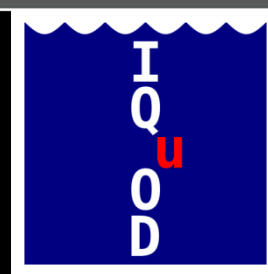
- Different groups around the world have their own quality control (QC) systems (often including both automatic and manual QC tests).
- Due to the large number of profiles in WOD we need good automatic QC as manual QC is not possible for all the data.
- We don't know which are the most effective automatic tests.
- There are smaller datasets with high level of QC performed that can be used to benchmark the automatic tests.

Our approach

- Create an open source repository of Python implementations of as many QC checks as possible on GitHub.
- Wrapped around these is code that reads each test dataset and runs each quality control check on each profile in our test datasets.
- The best combination of tests will be found by comparing to the data provider's QC results.
 - ✓ Avoids difficulties associated with groups running tests themselves;
 - ✓ Modular – easy to add new tests;
 - ✓ Collaboration – working this way has allowed us to benefit immensely from the efforts of Bill Mills and Gui Castelao (see CoTeDe on GitHub);
 - ✓ Tested – comprehensive code testing is performed;
 - ✓ The software should be viewed as IQuOD products – free Python QC and WOD ASCII reading software that anyone can use in their own systems.



Met Office



How to find our code

<https://github.com/IQuOD>

WodPy

*General purpose
Python reader for
WOD ASCII
format data*

To install:

```
pip install wodpy
```

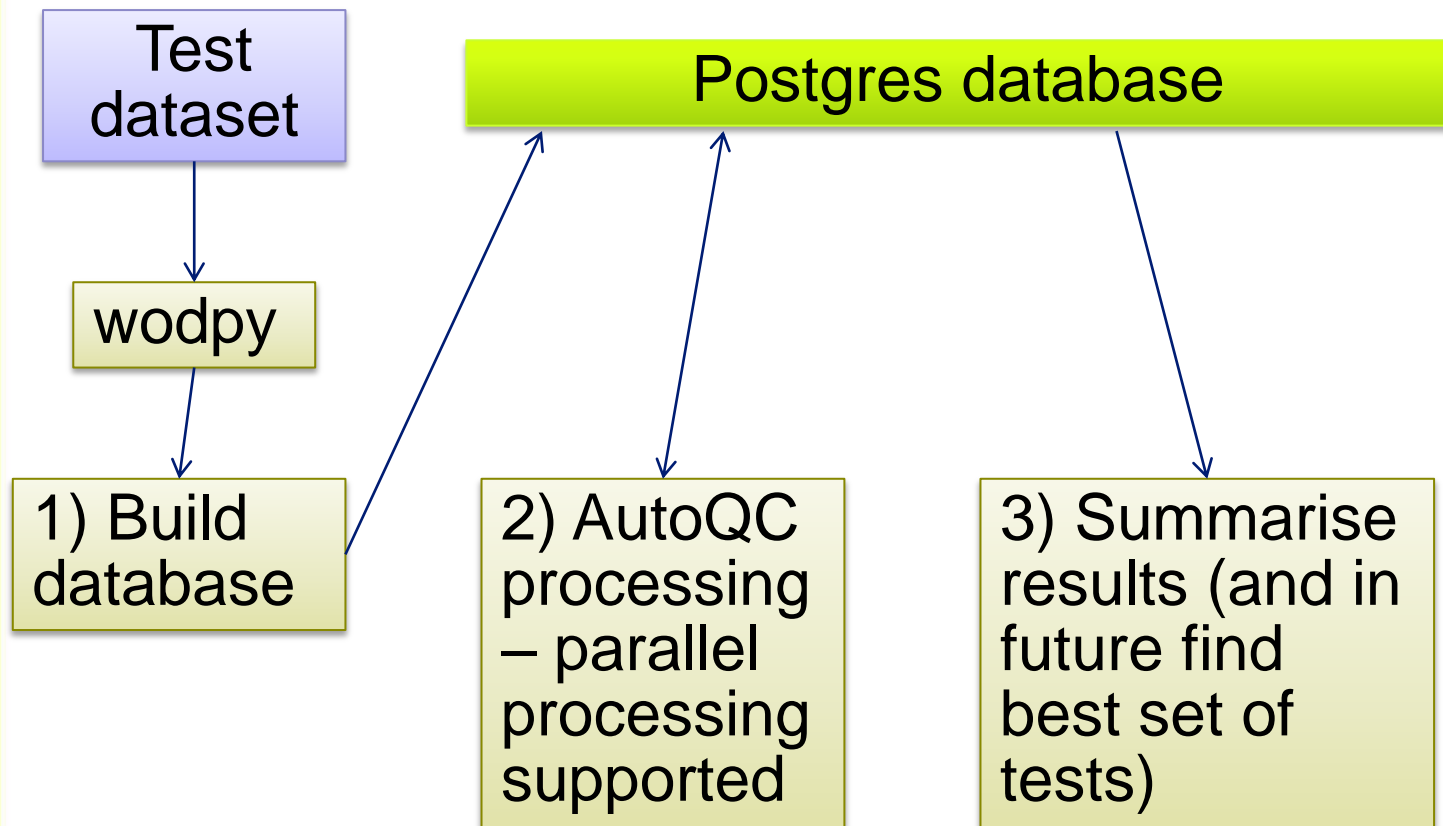
AutoQC

*Our QC code and
benchmarking
software*

Also available as a docker
image:

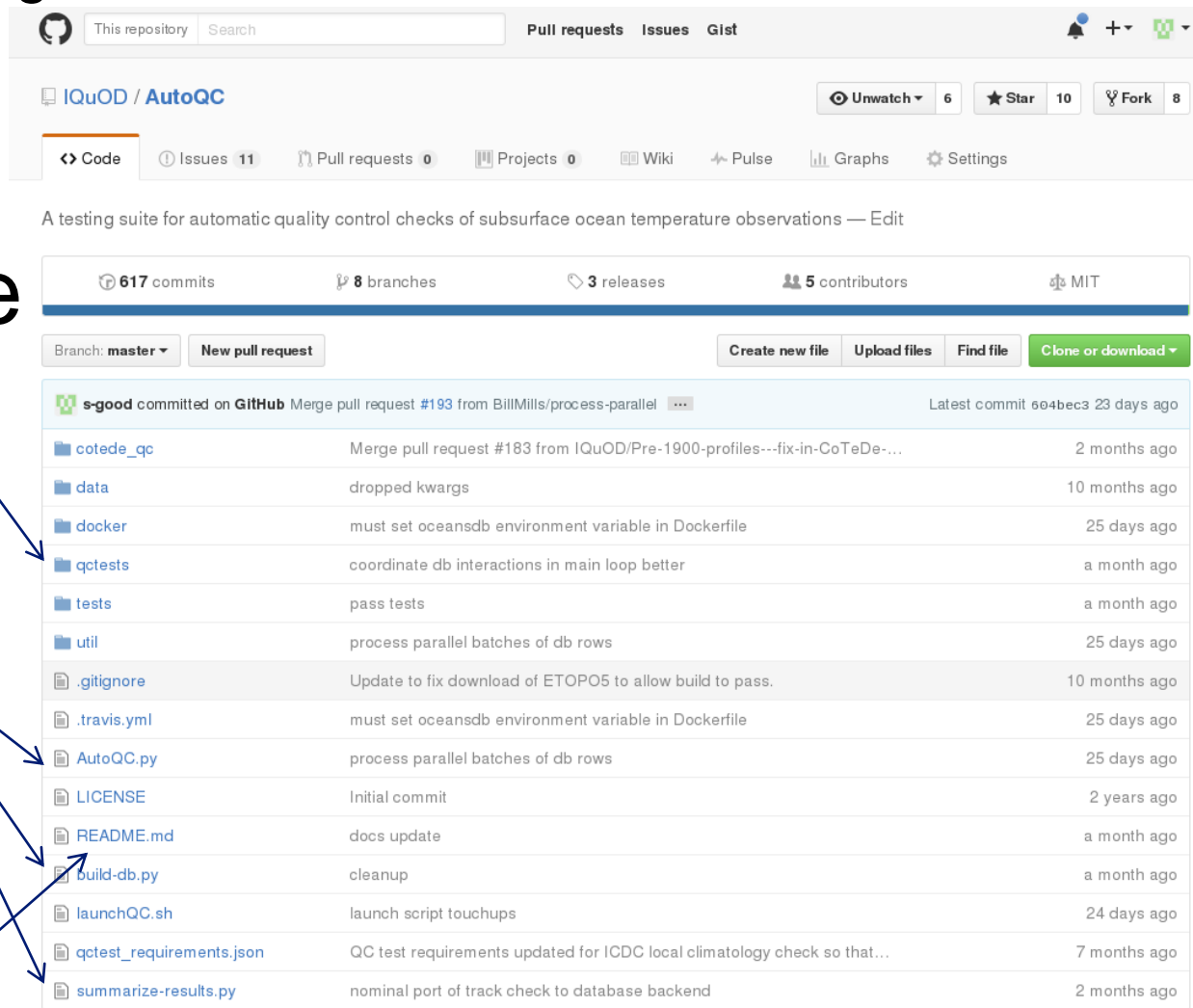
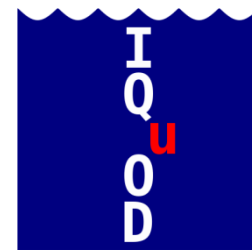
```
iquod/autoqc
```

The AutoQC processor



Code repository:

<https://github.com/IQuOD/AutoQC>



The screenshot shows the GitHub repository page for IQuOD/AutoQC. The repository has 617 commits, 8 branches, 3 releases, and 5 contributors. The commit history is as follows:

Commit	Message	Time
s-good	Merge pull request #193 from BillMills/process-parallel	23 days ago
cotede_qc	Merge pull request #183 from IQuOD/Pre-1900-profiles---fix-in-CoTeDe...	2 months ago
data	dropped kwargs	10 months ago
docker	must set oceansdb environment variable in Dockerfile	25 days ago
qctests	coordinate db interactions in main loop better	a month ago
tests	pass tests	a month ago
util	process parallel batches of db rows	25 days ago
.gitignore	Update to fix download of ETOPO5 to allow build to pass.	10 months ago
.travis.yml	must set oceansdb environment variable in Dockerfile	25 days ago
AutoQC.py	process parallel batches of db rows	25 days ago
LICENSE	Initial commit	2 years ago
README.md	docs update	a month ago
build-db.py	cleanup	a month ago
launchQC.sh	launch script touchups	24 days ago
qctest_requirements.json	QC test requirements updated for ICDC local climatology check so that...	7 months ago
summarize-results.py	nominal port of track check to database backend	2 months ago

QC test code

Main programs

Instructions

50 QC tests implemented

Argo_global_range_check	CoTeDe_location_at_sea_test
Argo_gradient_test	CoTeDe_rate_of_change
Argo_impossible_date_test	CoTeDe_spike
Argo_impossible_location_test	CoTeDe_tukey53H_norm
Argo_pressure_increasing_test	EN_background_available_check
Argo_regional_range_test	EN_background_check
Argo_spike_test	EN_constant_value_check
CSIRO_constant_bottom	EN_increasing_depth_check
CSIRO_depth	EN_range_check
CSIRO_long_gradient	EN_spike_and_step_check
CSIRO_short_gradient	EN_spike_and_step_suspect
CSIRO_surface_spikes	EN_stability_check
CSIRO_wire_break	EN_std_lev_bkg_and_buddy_check
CoTeDe_Argo_density_inversion	EN_track_check
CoTeDe_GTSPP_WOA_normbias	ICDC_aqc_01_level_order
CoTeDe_GTSPP_global_range	ICDC_aqc_02_crude_range
CoTeDe_GTSPP_gradient	ICDC_aqc_04_max_obs_depth
CoTeDe_GTSPP_profile_envelop	ICDC_aqc_05_stuck_value
CoTeDe_GTSPP_spike_check	ICDC_aqc_06_n_temperature_extrema
CoTeDe_Morello2014	ICDC_aqc_07_spike_check
CoTeDe_WOA_normbias	ICDC_aqc_08_gradient_check
CoTeDe_anomaly_detection	ICDC_aqc_09_local_climatology_check
CoTeDe_digit_roll_over	WOD_gradient_check
CoTeDe_fuzzy_logic	WOD_range_check
CoTeDe_gradient	loose_location_at_sea

Performance metrics

- As a first step we aim to find profiles containing any bad data.
 - Will serve as a 'do not use' list for a first IQuOD AutoQC dataset.
 - As input to enhanced quality control.
- Each profile has a True (rejected) or False (not rejected) flag from each QC test.
- These are compare to the 'truth' flags to give:
 - True positive rate;
 - False positive rate;
 - True negative rate;
 - False negative rate.

Performance metrics

Very unrepresentative sample of bad data!

NAME OF TEST	FAILS	TPR	FPR	TNR	FNR
Argo_global_range_check	1624	63.6%	0.0%	100.0%	36.4%
Argo_gradient_test	2240	87.7%	0.0%	100.0%	12.3%
Argo_impossible_date_test	0	0.0%	0.0%	100.0%	100.0%
Argo_impossible_location_test	0	0.0%	0.0%	100.0%	100.0%
Argo_pressure_increasing_test	0	0.0%	0.0%	100.0%	100.0%
Argo_regional_range_test	12	0.0%	1.1%	98.9%	100.0%
Argo_spike_test	679	26.6%	0.0%	100.0%	73.4%
CSIRO_constant_bottom	25	0.2%	1.7%	98.3%	99.8%
CSIRO_depth	3680	99.8%	100.0%	0.0%	0.2%
CSIRO_long_gradient	2917	97.4%	37.9%	62.1%	2.6%
CSIRO_short_gradient	2405	92.7%	3.2%	96.8%	7.3%
CSIRO_surface_spikes	2923	96.8%	39.8%	60.2%	3.2%
CSIRO_wire_break	2243	87.8%	0.0%	100.0%	12.2%
CoTeDe_Argo_density_inversion	0	0.0%	0.0%	100.0%	100.0%
CoTeDe_GTSPP_WOA_normbias	2447	92.4%	7.7%	92.3%	7.6%
CoTeDe_GTSPP_global_range	1632	63.7%	0.4%	99.6%	36.3%
CoTeDe_GTSPP_gradient	2049	80.2%	0.0%	100.0%	19.8%
CoTeDe_GTSPP_profile_envelop	2300	89.7%	0.6%	99.4%	10.3%
CoTeDe_GTSPP_spike_check	713	27.8%	0.2%	99.8%	72.2%
CoTeDe_Morello2014	2987	97.1%	44.9%	55.1%	2.9%
CoTeDe_WOA_normbias	2224	87.0%	0.2%	99.8%	13.0%
CoTeDe_anomaly_detection	3376	97.8%	77.7%	22.3%	2.2%
CoTeDe_digit_roll_over	2234	87.4%	0.1%	99.9%	12.6%
CoTeDe_fuzzy_logic	2503	95.2%	6.3%	93.7%	4.8%
CoTeDe_gradient	2111	82.6%	0.0%	100.0%	17.4%

Performance metrics

Very unrepresentative sample of bad data!

NAME OF TEST	FAILS	TPR	FPR	TNR	FNR
CoTeDe_location_at_sea_test	2	0.0%	0.2%	99.8%	100.0%
CoTeDe_rate_of_change	2388	88.9%	10.3%	89.7%	11.1%
CoTeDe_spike	158	6.2%	0.0%	100.0%	93.8%
CoTeDe_tukey53H_norm	254	9.9%	0.1%	99.9%	90.1%
EN_background_available_check	161	5.8%	1.1%	98.9%	94.2%
EN_background_check	2382	93.0%	0.6%	99.4%	7.0%
EN_constant_value_check	5	0.2%	0.0%	100.0%	99.8%
EN_increasing_depth_check	3	0.1%	0.0%	100.0%	99.9%
EN_range_check	1623	63.5%	0.0%	100.0%	36.5%
EN_spike_and_step_check	313	12.3%	0.0%	100.0%	87.7%
EN_spike_and_step_suspect	2264	88.5%	0.2%	99.8%	11.5%
EN_stability_check	0	0.0%	0.0%	100.0%	100.0%
EN_std_lev_bkg_and_buddy_check	454	16.5%	2.8%	97.2%	83.5%
EN_track_check	87	0.5%	6.6%	93.4%	99.5%
ICDC_aqc_01_level_order	0	0.0%	0.0%	100.0%	100.0%
ICDC_aqc_02_crude_range	2322	90.7%	0.4%	99.6%	9.3%
ICDC_aqc_04_max_obs_depth	1	0.0%	0.0%	100.0%	100.0%
ICDC_aqc_05_stuck_value	58	2.3%	0.0%	100.0%	97.7%
ICDC_aqc_06_n_temperature_extrema	122	4.8%	0.0%	100.0%	95.2%
ICDC_aqc_07_spike_check	175	6.8%	0.0%	100.0%	93.2%
ICDC_aqc_08_gradient_check	2144	83.5%	0.9%	99.1%	16.5%
ICDC_aqc_09_local_climatology_check	2603	95.0%	15.6%	84.4%	5.0%
WOD_gradient_check	2336	90.4%	2.3%	97.7%	9.6%
WOD_range_check	2310	90.4%	0.0%	100.0%	9.6%
loose_location_at_sea	1	0.0%	0.1%	99.9%	100.0%

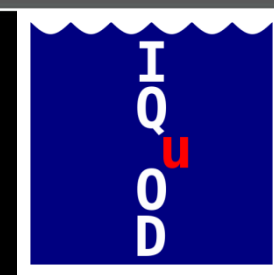
Summary

- ✓ 50 QC tests included in the AutoQC framework (16 are wrappers of CoTeDe tests)
<https://github.com/castelao/CoTeDe>
- ✓ AutoQC benchmarking code is ready to run.
 - ✓ Now using a postgres database to avoid memory issues.
- ✓ Tim has extracted/converted to WOD ASCII the test datasets.
- ✓ Waiting on virtual machine with docker install to be set up at Met Office to process them.
 - ✓ Too many tests and too much data for Amazon server to be used?
- ✓ Some work already done on developing software to find the optimum set.



Met Office

Progress over the last year against actions



- ✓ 1. Tim to extract test datasets by May. Argo DM, QuOTA, North Sea and seal tag data extracted. WOCE dataset could be added if additional data needed.
- ✓ 2. Bec/Ann to identify the 4 months in Quota that have every profile visually QC'd.
- ✓ 3. Bill and Simon to assist with CSIRO (Bec) and ICDC (Viktor) tests. Implementation either with Bill's help or do it yourself in this repo: <https://github.com/IQuOD/AutoQC>.
- ? 4. Ann, Bec, and anyone else interested to provide examples of profiles that are flagged in expert/manual QC. Examples to be sent to Bill Mills and Guilherme Castelao.
- ✓ 5. Simon, Bill + all to implement as many QC tests as possible (May).
- ✗ 6. Simon, Bill to use test datasets to determine performance of tests and the best set to use, and check consistency across datasets (June).
- ✗ 7. Simon, Bill to distribute files containing results of applying the best set of tests to interested people (Viktor, Ann, others?) to calculate statistics/check that the results look sensible (in time for the next IQuOD workshop – September).
- ✗ 8. Simon, Bill + other contributors to draft paper (September onwards).
- ✗ 9. Regarding the AutoQC benchmarking work – the geographic distribution of false positives should be investigated.
- ✓ 10. There may be computational and storage resources available from Amazon –perhaps this should be investigated further?
- ✓ 11. Post the AutoQC “Quick Start” guide to the IQuOD website (Simon/Bec).
- ? 12. Investigate the appropriate license for the IQuOD dataset. Eg: creative commons. Will have to work with NCEI requirements (Simon/Bill/Tim).

Where next?

- Aim to run processing over the next 1-2 months.
 - Raw results could be posted on a website if there is any interest?
- Then find optimum QC set by Christmas.
 - + reject XBT data shallower than 4m?
 - Lead to new version of IQuOD dataset.
 - Publish results.
- Enhanced QC to find which levels within profiles are bad.
 - Look at the AutoQC results to determine how good the automatic tests are?
 - Develop AutoQC to benchmark level QC?
 - Crowdsourcing?
 - Expert QC?
 - Machine learning?



Met Office

