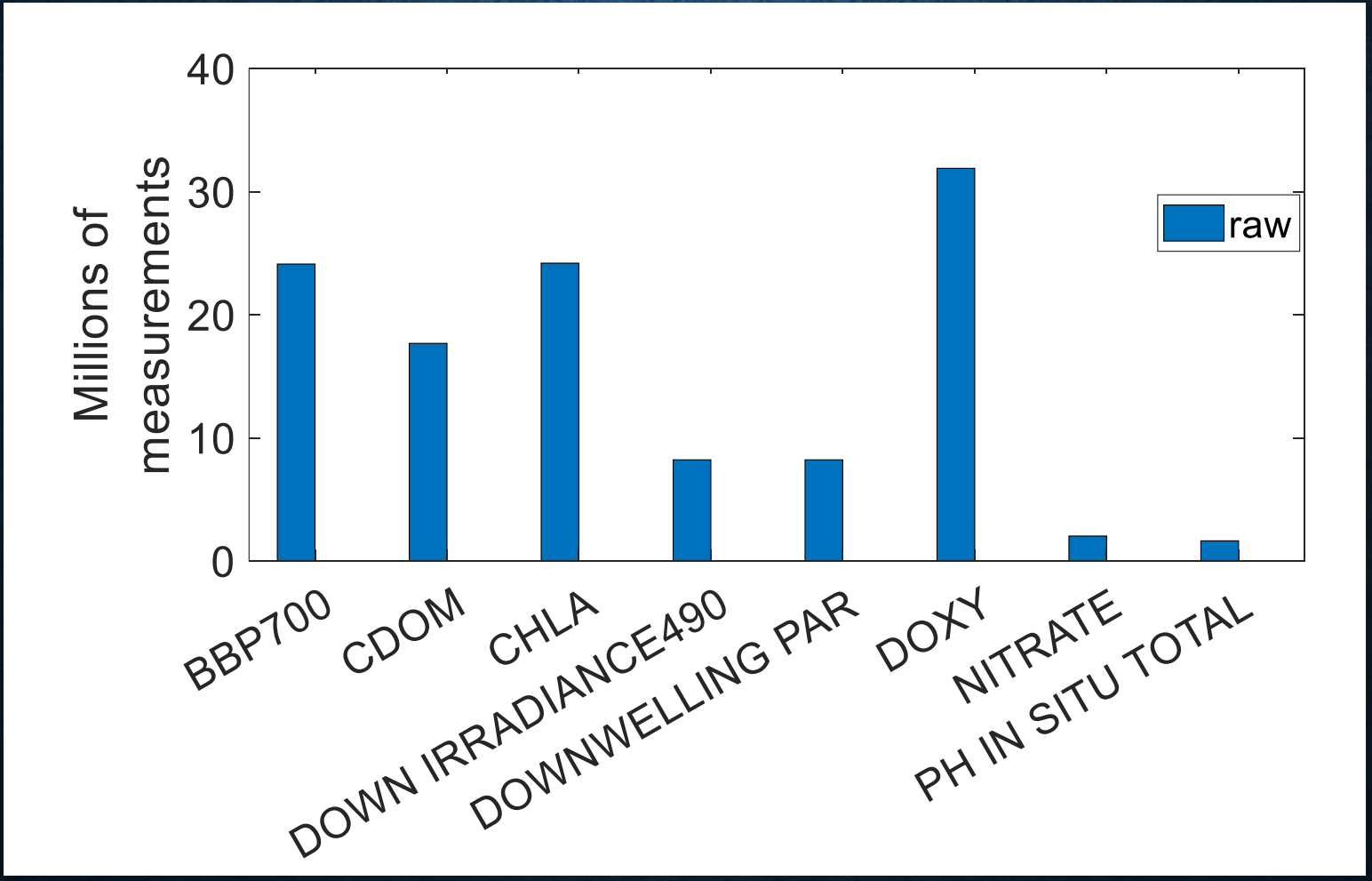
How can we increase the amount of adjusted data at the GDACS

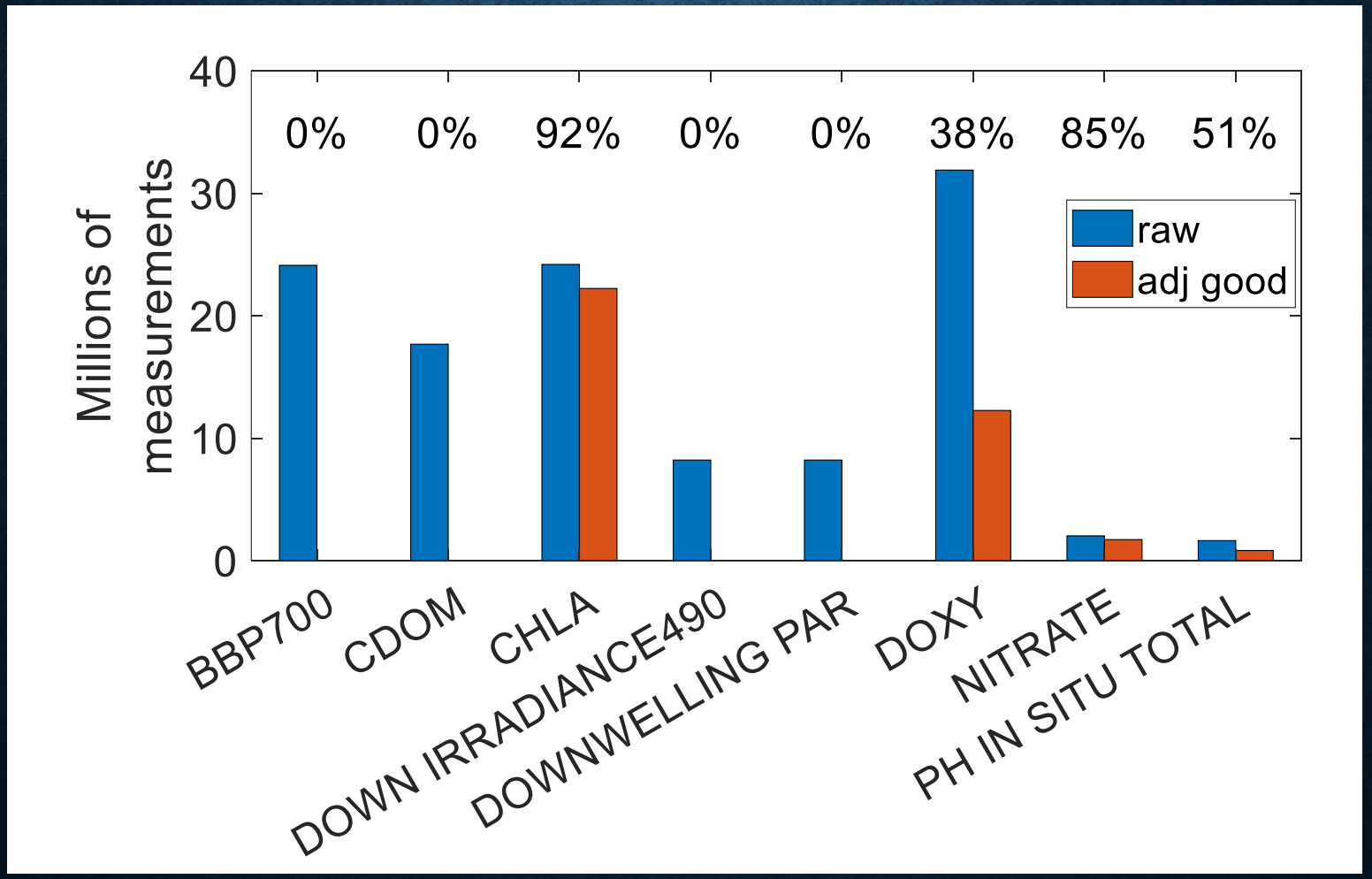
Josh Plant Ken Johnson & Tanya Maurer, MBARI

ADMT20, October 13-18, 2019
Laboratoire d'Océanographie de Villefranche-sur-Mer

BGC Argo has produced over 118,000,000 BGC measurements!



But….only a small amount of good quality data is propagated to the "_ADJUSTED" parameters



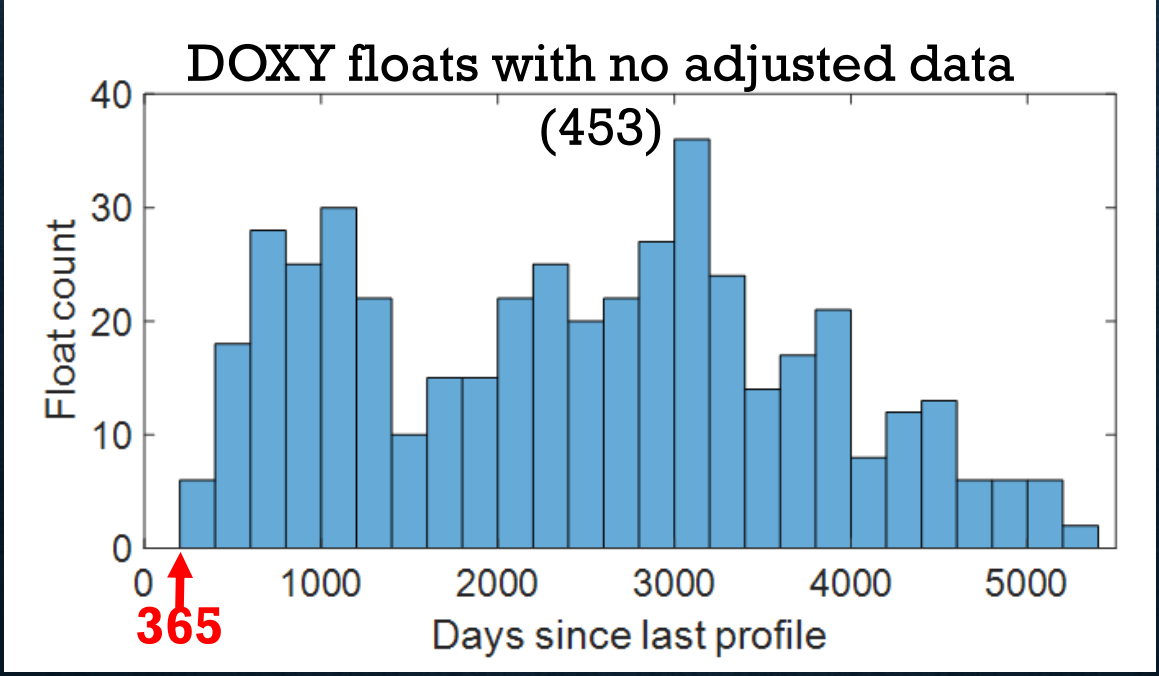
Several reasons why

- Off-line corrected data bases are not propagated to the GDACs
- Undetermined definition for “adjusted” (BBP, CDOM)
- Premature sensor failure (pH)
- Resources lacking to generate adjusted data

From the Argo users manual ……

1.7 Real-time and Delayed mode data

….scientists apply other procedures to check data quality and **the target is for these data to be returned to the global data centres within 6 to 12 months**. These constitute the delayed mode data (DM).



The good news

- Most of the CHL data (92%) has been adjusted
- Many DACs are now producing DOXY_ADJUSTED for recent data
- Older floats need to be corrected next!

Wondering if it would be possible to propose an action item:

Have 90% DOXY data older than a year processed to delayed mode by ADMT21?

Two main hurdles

- Off-line corrected data bases need to be propagated to the GDAC

- Resources lacking to generate the data

How can we help?

- Code sharing
  MBARI uses Matlab, SAGEO2 GUI
  https://github.com/SOCCOM-BGCArgo/ARGO_PROCESSING
- SAGEO2 GUI help
- Data processing for other dacs
- If air cal isn't possible use the gain value we produce for all floats as a first step?

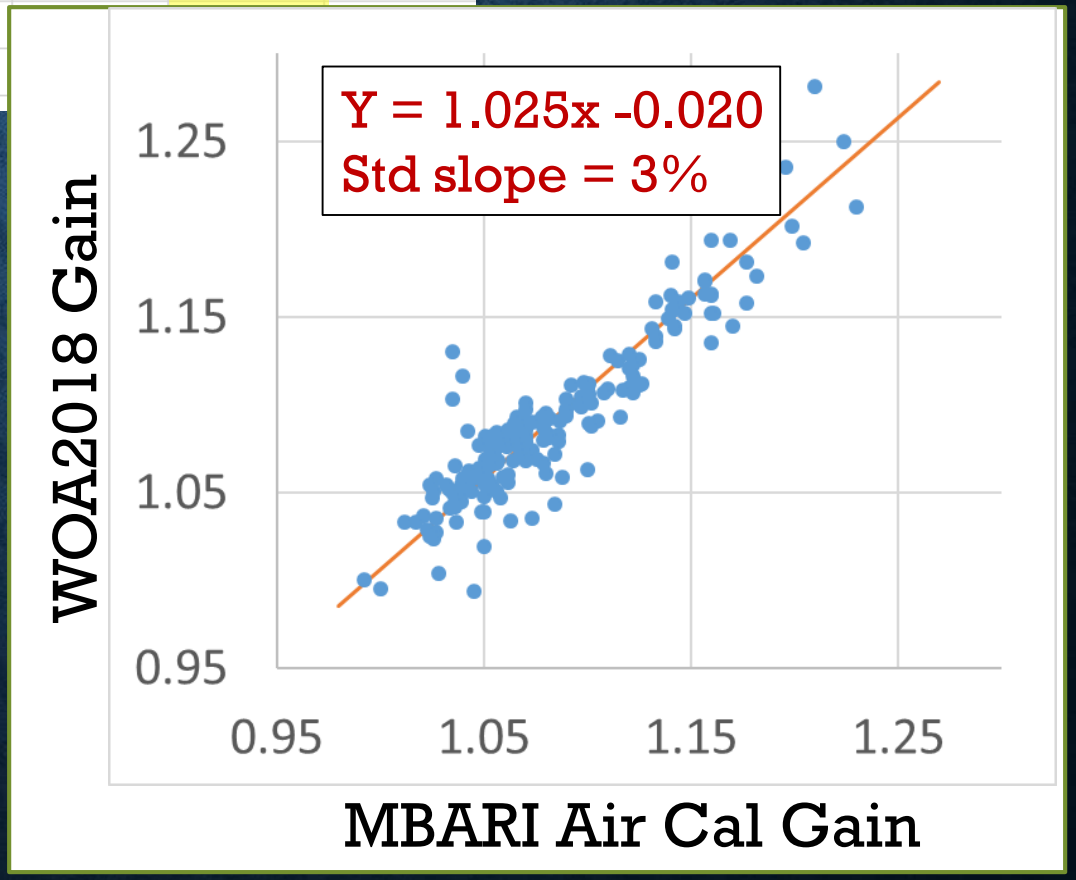
DOXY_ADJUSTED = DOXY * Gain

1	//DOXY BGC Argo gain report				
2	//File created on 13-Oct-2019 05:40:51 by jplant				
3	//Created with argoO2vsWOA				
4	//Only DOXY data used in comparison - not DOXY_ADJUSTED				
5	//All data was extracted from the Sprof merged synthetic profile files				
6	//GAIN = [WOA O2 %sat / DOXY O2 %sat] in the upper 20.0 meters				
7	**DAC**	**PI**	**WMO**	**gain**	**std gain**
8	aoml	STEPHEN RISER	1900722	1.118	0.023
9	aoml	STEPHEN RISER , KENNETH JOHNSON	1901378	1.097	0.023
10	aoml	STEPHEN RISER , KENNETH JOHNSON	1901379	1.025	0.013
11	aoml	BRECK OWENS			
12	aoml	BRECK OWENS			

Gain list from MBARI

WOA gain compares well with Air cal

DOXY_ADJUSTED_ERROR ~ 3%



Thoughts / questions ?

Generating adjusted data at MBARI

Why do we need to adjust BGC float data?

Sensors are not perfect (O_2, NO_3, pH) yet but ….

"Argo data are intended to be research-quality and include estimates of data quality and accuracy"
(from Argo User's Manual)

Raw BGC Argo data should always be used with caution!
"… the accuracy of these biogeochemical data at their raw state is not suitable for direct usage in scientific applications"
(from Argo Quality Control Manual for Dissolved Oxygen)

What kind of data adjustments are required?

	Gain	Offset	Drift
Salinity		X	X
Biooptics	X	X	?
O_2	X		X
NO_3^-		X	X
pH		X	X

- **O_2** optodes suffer from storage drift (0-20%) & deployment drift (-1.9 to +1.4%/ yr) in gain
- **NO_3** drifts and offsets result from dirty optics, reduction in light throughput
- **pH** drifts and offsets result from changes to sensor reference potential overtime

SOCCOM data processing

Automatic QC

MBARI
Monterey Bay Aquarium Research Institute

Data Assembly

Raw data processing
Factory calibrations
Automated QC flagging
Automated QC adjustments

Science quality data to user!

Rules

Manual QC

NCEP
NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION

Empirical Algorithms
LIR, MLR, CANYON

Adjustment & Validation "SAGE" GUIs

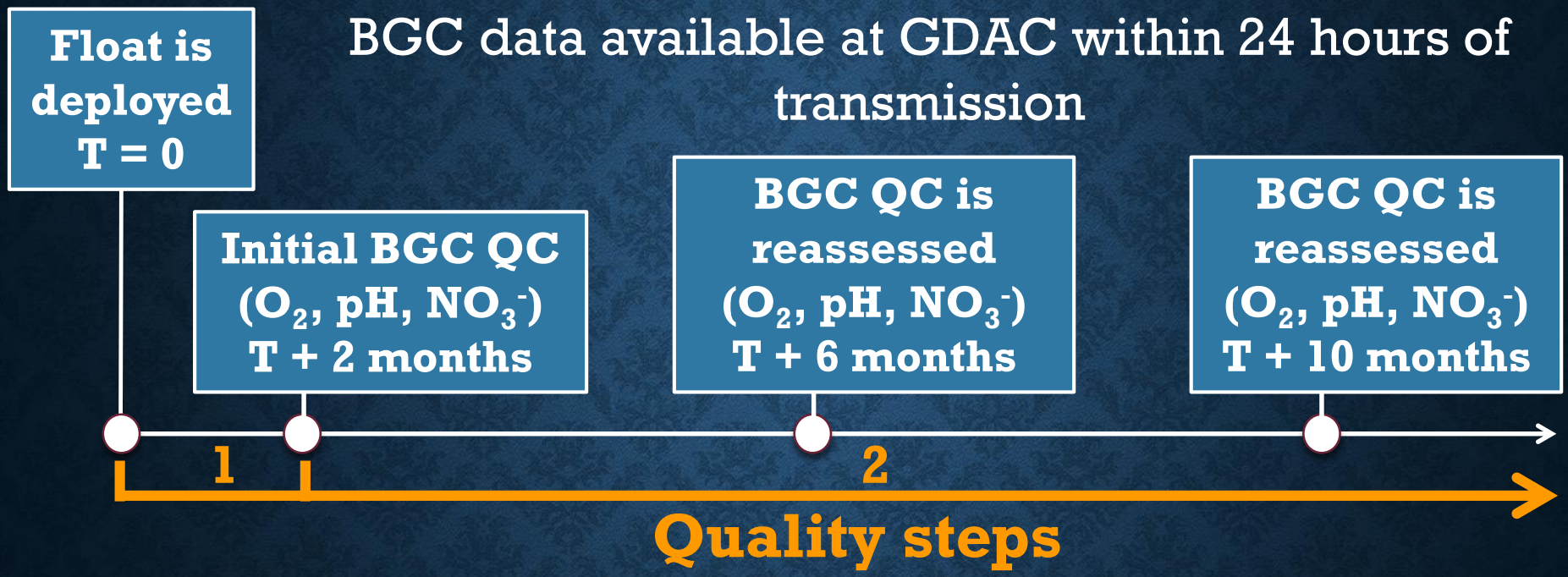
GLODAP
Global Ocean Data Analysis Project
CO_2 ^{14}C ^{13}C NO_3 PO_4 H_4SiO_4

WOA2013
NOAA
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
U.S. DEPARTMENT OF COMMERCE

Bottle data

Informal Audits

Single float manual adjustment timeline



Step 1	QC flags assigned based on range, spike & grey list checks All BGC data flagged questionable at best (RAW)
Step 2	After 2 months 1st data adjustments occur (ADJUSTED) **Living data set** - data flags re-assessed quarterly Every 5th float cycle float is completely reprocessed
Level 3	Final data adjustments after end of float life

Manual QC process – "Step 2"

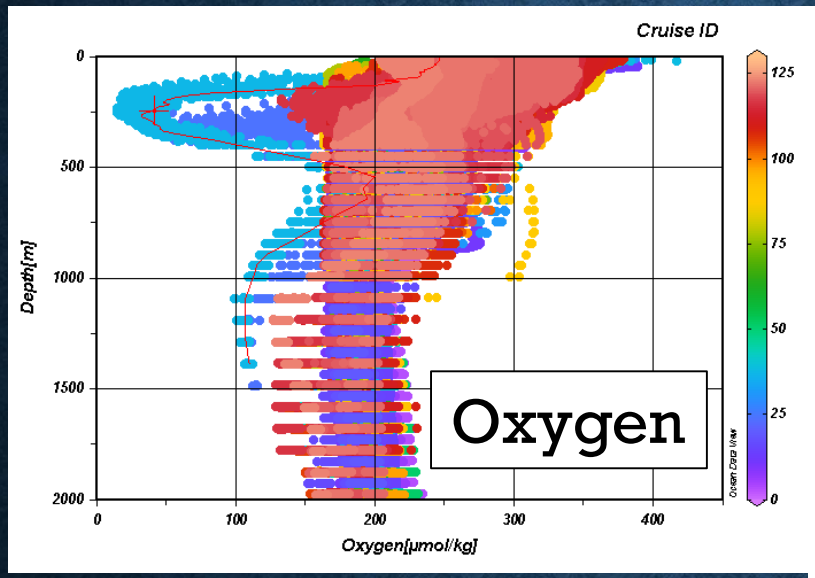
- Presently user interaction required to generate rules
- Rules defined automatically
  (Change point detection & BIC)
- Rules saved & fed to automated process (D mode)
- Rules automatically applied until next data reassessment (A mode)

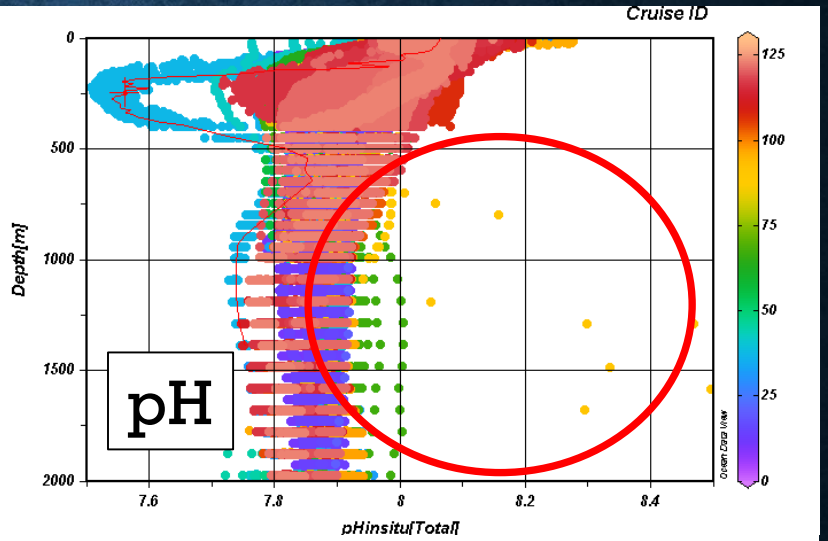
<u>Two manual tasks:</u>

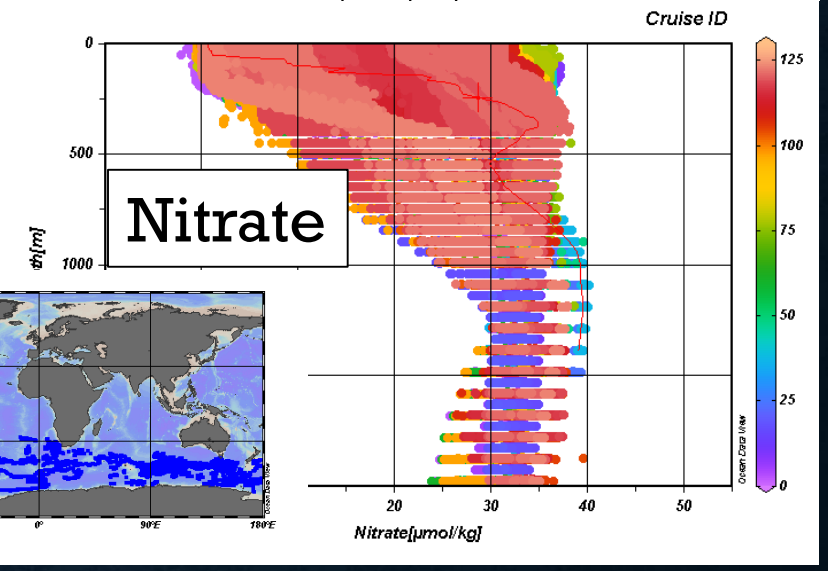
- Visual observation for outliers & odd behavior
- Get data adjustments using reference data sets in SAGE-O2 & SAGE GUIs (generate rules)

Visual observation for outliers

★ Look at your data set!







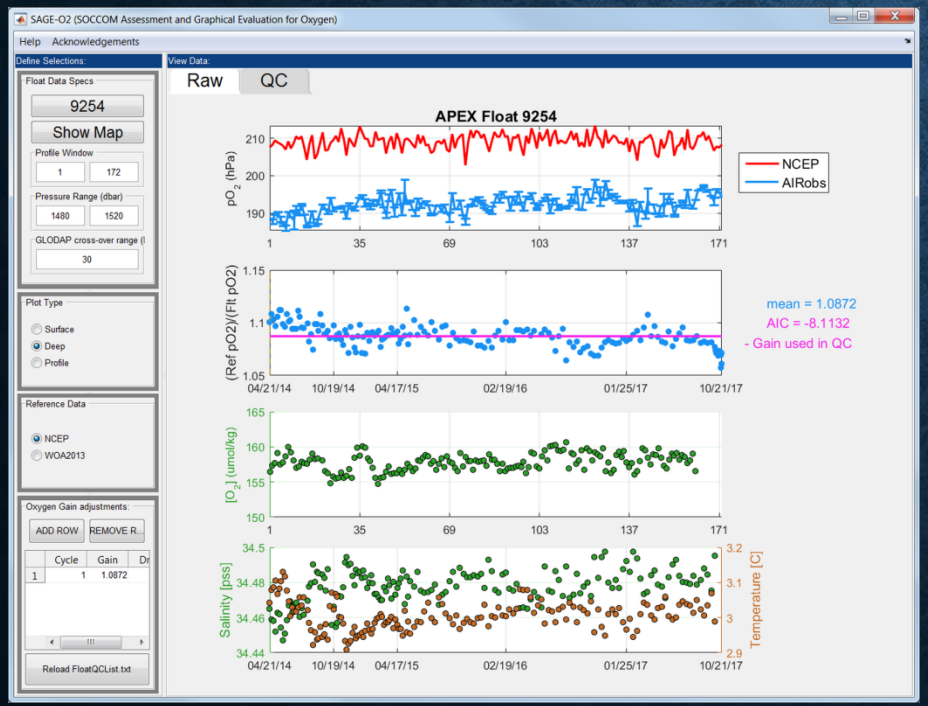
- Simple visualization tools are useful – ie ODV
- Multi-parameter context can help distinguish anomalies from real events

Data adjustment using SAGE-O2 & SAGE GUI's

- Matlab GUIs developed at MBARI
- Used to assign gain, drifts and offsets in float data sets
- Float data compared / corrected to reference data sets
- Free code: https://github.com/SOCCOM-BGCArgo

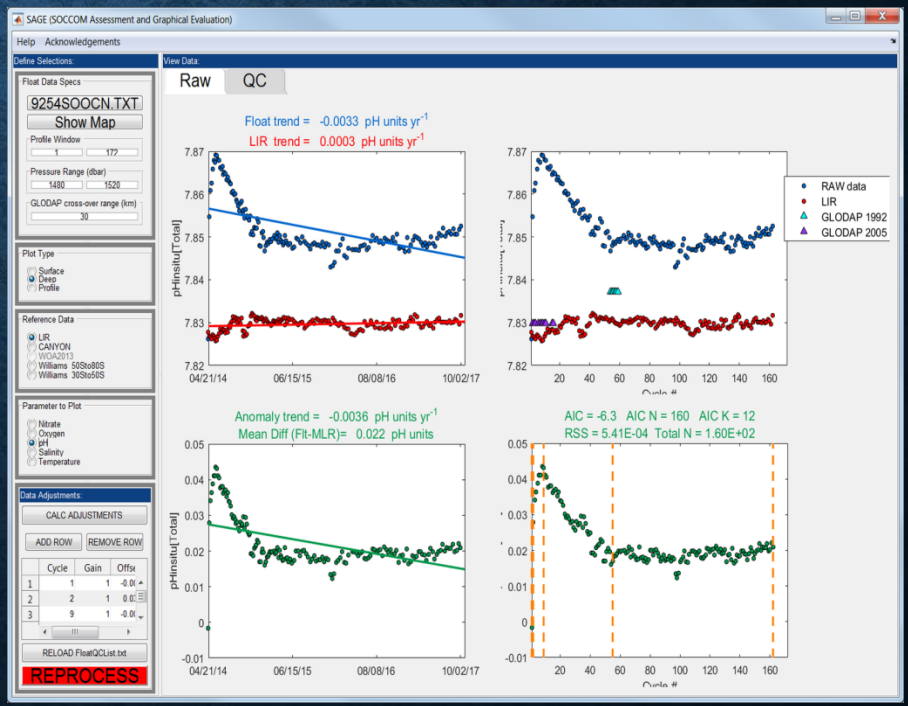
SAGE-O2

SOCCOM Assessment and Graphical Evaluation for Oxygen



SAGE

SOCCOM Assessment and Graphical Evaluation (for pH & NO_3)



Data adjustment using SAGE-O2 & SAGE GUI's

Correcting Oxygen with SAGE O2

O_2 1st parameter to adjust → influences pH and NO_3 adjustments

Adjustment process:

Method 1: Air-calibrate
→Measure air O_2 → convert to pO2 → compare to atmospheric reanalysis (NCEP) *Johnson et al, 2015*

$$(O_2)_{corr} = G \times (O_2)_{raw} \qquad G = \frac{\sum_i^n g_i}{n}$$
$$g_i = pO_2/pO_{2,optode}$$
$$pO_2 = (P_{NCEP} - p_{H_2O}) \times 0.20946$$

Method 2: Shipboard data
→Compare float data to shipboard sample profiles

$$G = \frac{\sum_i^n \left([O2]_{bottle}/[O2]_{float}\right)_i}{n}$$

Method 3: WOA2013
→Compare float data at the surface to World Ocean Atlas climatology (%sat) *Takeshita et al, 2013*

$$G = \frac{\sum_i^n \left(\%sat_{WOA}/\%sat_{float}\right)_i}{n}$$

A first step:

Increasing the quantity of DOXY_ADJUSTED DATA at the GDAC should be a priority!

- Good oxygen data needed for pH & Nitrate corrections
- B-traj files needed for routine air calibration
- If air cal not possible use gains generated from WOA

Undetermined definition of adjusted for BBP CDOM
What makes data “adjusted” ?
Sensor data quality generally good (maintains calibration response)
Accuracy of derived quantities less good
Should BBP & CDOM get automatically propagated to adjusted fields?

Oxygen gains from woa
Compare plot with MBARI air gains

Undetermined definition for “adjusted”

Example: BBP, CDOM

- Data quality generally very good!
  Sensors maintain factory calibration (mostly)
- Conversion to derived parameter (i.e. POC) less accurate (this is a different issue though)
- Community decision

Consider populating adjusted fields for BBP, CDOM, (more?) which pass automated tests?

Premature sensor failure

Example: pH

- Young sensor – still improving reliability

- Improved designs by MBARI & SBE in the works