

# U.S. GEOTRACES PACIFIC TRANSECT DATA MANAGEMENT

Cyndy Chandler  
BCO-DMO WHOI

*US GEOTRACES Pacific Zonal Transect Planning Workshop*

Woods Hole, MA USA

26 April 2013



# TOPICS

1. What is BCO-DMO?
2. How does BCO-DMO relate to US GEOTRACES?
- ~~3. Important concepts in data management ...~~
4. US GEOTRACES data management
5. Lessons learned from 2010 & 2011 North Atlantic Transect cruises

# NSF OCE FUNDED RESEARCH DATA



Biological & Chemical Oceanography Data Management Office

[bco-dmo.org](http://bco-dmo.org) for NSF OCE project data  
current projects, and legacy data from large  
coordinated research programs  
(e.g. US GLOBEC and US JGOFS)



[bco-dmo.org](http://bco-dmo.org)



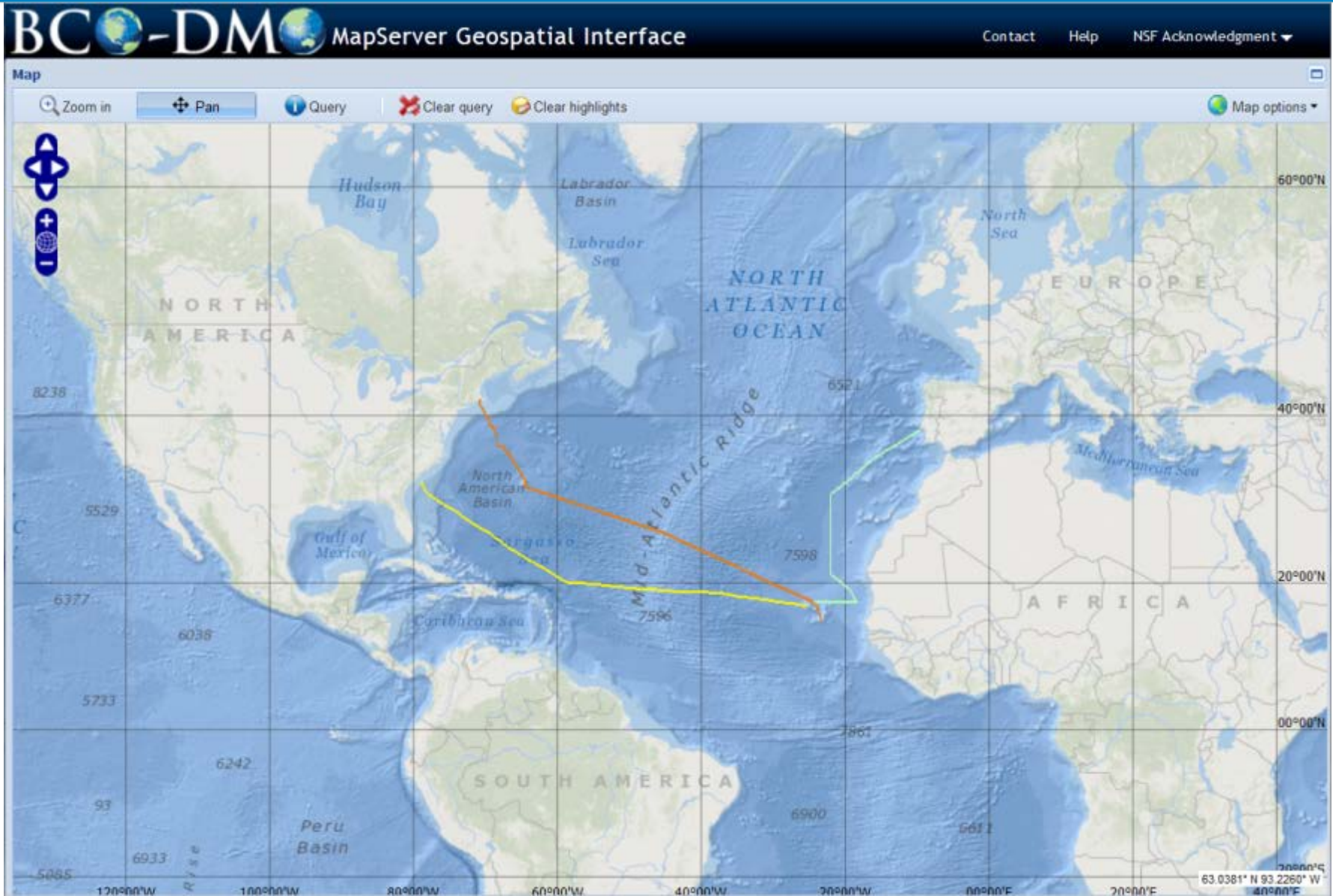
# BCO-DMO AND U.S. GEOTRACES?

BCO-DMO staff members provide data management support for investigators funded by the Biological or Chemical Oceanography sections of NSF Ocean Sciences or OPP ANT – *at no additional cost to the project*

This represents a new model for data management; a long-term commitment to community-wide data management as opposed to a project specific data management office.

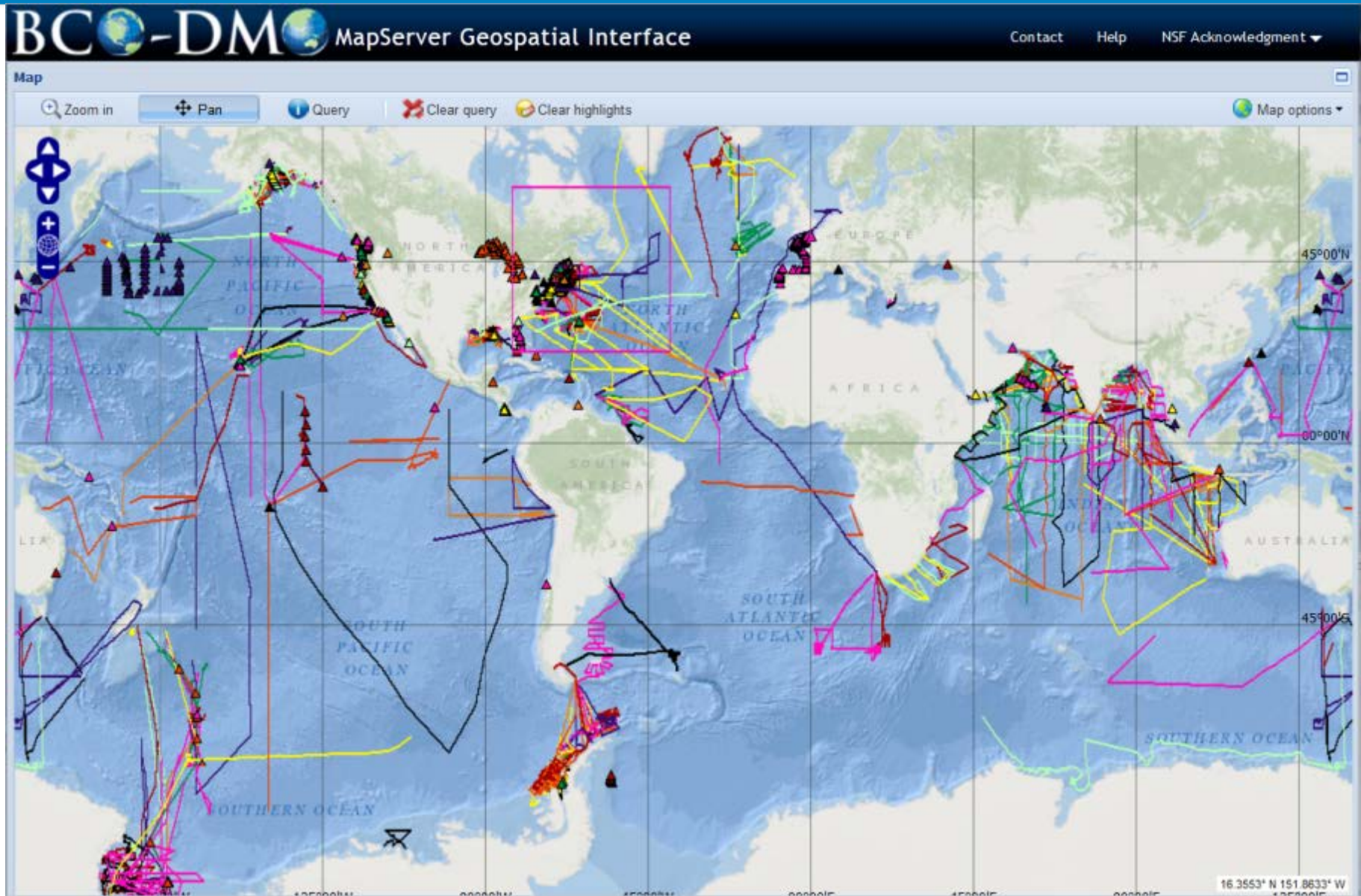


# US GEOTRACES NAT DATA





# NSF OCE RESEARCH PROJECT DATA



# BCO-DMO AND US GEOTRACES

## BCO-DMO: what we do ...

- ❖ Provide data management support  
<http://bcodmo.org/resources>
  - Best Practice Guidelines Manual
  - Data Management Plan
  - How to contribute data
- ❖ Make data and metadata available
  - Restricted or public access as appropriate
- ❖ Ensure final archive of data in appropriate National Data Center (NODC)

# U.S. GEOTRACES DATA MANAGEMENT

based on recommendations from GEOTRACES DMC and SSC

Pacific Section cruise:

❖ Chief Scientist

- point of contact for BCO-DMO
- Cruise report
  - » Cruise metadata including participant list
- Navigation data (X,Y,T cruise track data)
- Scientific sampling event log
- final data inventory (list of expected data sets)

❖ phased reporting of data sets from cruise

- cruise track (nav data) event log and basic hydrography (CTD and Rosette base data); sampling ID log
- bathymetry, ADCP, underway and meteorological
- measurements from on board investigator teams



# OCEAN DATA FACILITY (ODF) AT SIO

ODF personnel are an excellent resource for hydrographic cruise planning

Recommendation:  
include representatives from ODF in pre-cruise planning workshops

*(Chandler & German, 2011)*



# GEOTRACES DATA TIMELINE

## **Data/Metadata Submission (timeline):**

- As soon as a cruise is organized: submit pre-cruise metadata to GEOTRACES IPO and BCO-DMO. (complete form)
- Within one week of cruise completion (Chief Scientist):
  - Submit Post-cruise metadata forms (update pre-cruise)
  - Submit electronic versions (scanned or original) of event log and sample log sheets, [and copies of the bridge log]
  - Submit copy of ROSCOP/CSR form or equivalent cruise report

# GEOTRACES DATA TIMELINE

## **Data/Metadata Submission (timeline):**

- Within 6 months of end of cruise:
  - Chief scientist submits final cruise report
  - Chief Scientist does cruise status review with BCO-DMO
  - Data and metadata for shared ancillary parameters (e.g., nutrients) submitted to BCO-DMO
  - Submit CTD and underway data (both raw and processed files; sensor information and calibration) to BCO-DMO

# GEOTRACES DATA TIMELINE

## **Data/Metadata Submission (timeline):**

As soon as possible, within 2 years of data generation:

- Submit all data sets and accompanying metadata to GDAC
- GDAC: In most cases, data will be submitted initially to a national data centre (DAC or GEOTRACES Data Assembly Center, e.g., BCO-DMO). BCO-DMO is responsible for submitting US GEOTRACES data to BODC.



# GEOTRACES DATA FLOW

BCO-DMO contributes US GEOTRACES data to the Data Portal at BODC



GEOTRACES International Programme data are managed by BODC/NERC

<http://www.bodc.ac.uk/geotraces/>



## GEOTRACES International Data Assembly Centre

GEOTRACES ([www.geotraces.org](http://www.geotraces.org)) is an international programme which aims to improve our understanding of biogeochemical cycles and large-scale distribution of trace elements and their isotopes (TEIs) in the marine environment. The global field programme will run for at least a decade and will involve cruises in all ocean basins run by a variety of nations.

Planning has involved scientists from around 30 countries. GEOTRACES is expected to become the largest programme to focus on the chemistry of the oceans and will improve our understanding of past, present and future distributions of TEIs and their relationships to important global processes.

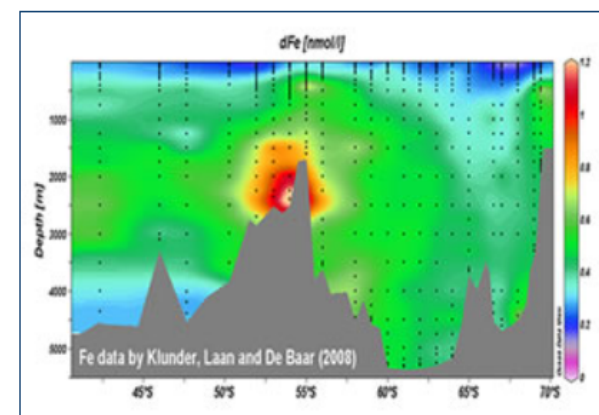
### GEOTRACES mission is:

***To identify processes and quantify fluxes that control the distribution of key trace elements and isotopes in the ocean, and to establish the sensitivity of these distributions to changing environmental conditions.***

Our aim as the GEOTRACES International Data Assembly Centre (GDAC) is to provide the data management to promote data sharing and collaboration between research groups and to ensure data are made widely accessible for long-term use.

To find out more follow these links

- [Introduction](#) — A non-technical insight into the main goals and themes of the GEOTRACES programme.
- [Benefits](#) — A brief description of the long term benefits of the programme.
- [Role](#) — The role of the International Data Management Office.



Deep ocean section of dissolved (<0.2 micron filtered) iron (Fe) at the zero meridian in the Antarctic Ocean. Data collected during expedition ANT 24-3 (2008) aboard icebreaker POLARSTERN in context of the International Polar Year GEOTRACES program. The very low dissolved Fe in surface waters and throughout the water column at 67-68° South is consistent with the overall limitation of Antarctic ecosystems due to lack of essential trace element Fe for biota. © Maarten Klunder. [Enlarge image](#)

<http://www.bodc.ac.uk/geotraces/>

# IMPORTANT LINKS

GEOTRACES data portal at BODC

<http://www.bodc.ac.uk/geotraces/>

BCO-DMO is the US GDAC  
(US GEOTRACES Data Assembly Center)

<http://bco-dmo.org/>

US GEOTRACES data managers at BCO-DMO  
[info@bco-dmo.org](mailto:info@bco-dmo.org)



# U.S. GEOTRACES DATA MANAGEMENT

Important to understand:

- ❖ station identification system  
(e.g., super, deep and shallow stations)
- ❖ event identification system; event log entries
- ❖ robust metadata records to support shared use of data, and future unanticipated use
  - sampling and analytical protocols
  - quality assurance and control procedures

# LESSONS LEARNED

US GEOTRACES North Atlantic Transect  
lessons learned, based on feedback from NAT  
Chief Scientists, PIs and BCO-DMO staff

Bob Groman

Nancy Copley

Dicky Allison

Terry McKee

Danie Kinkade

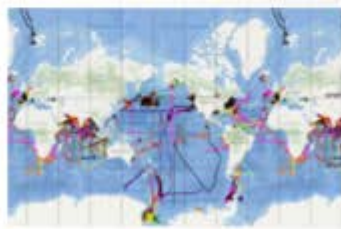
Steve Gegg

Shannon Rauch

## DATABASE

Welcome	
Programs	24
Projects	227
Deployments	1671
Datasets	6351
Instruments	299
Parameters	1300
People	1301
Affiliations	328
Funding	45
Awards	571

## GEOSPATIAL ACCESS



## Project: U.S. GEOTRACES North Atlantic Transect

Acronym: U.S. GEOTRACES NAT

Program: U.S. GEOTRACES [U.S. GEOTRACES]

URL: Project Web Site

URL: Project Data Site

Start date: 2010-10

End date: 2010-12

Geolocation: Subtropical western and eastern North Atlantic Ocean

► **Description:**

### Additional Project Information

► **Funding**

▼ **Datasets**

Dataset
GT10 - As AP Surface Transects
GT10 - CDOM
GT10 - Cruise Event Log
GT10 - Cruise Report
GT10 - Cruise Tracks
GT10 - CTD - GT-C Bottle
GT10 - CTD - GT-C Cast Sheets
GT10 - CTD - GT-C Profiles
GT10 - CTD - GT-C Sample Logs
GT10 - CTD - ODF/SIOR Bottle
GT10 - CTD - ODF/SIOR Cast Sheets
GT10 - CTD - ODF/SIOR Profiles
GT10 - CTD - ODF/SIOR Sample Logs
GT10 - CTD - PUMP Sample Logs



Total of 80 data sets so far  
bco-dmo.org



# LESSONS LEARNED (NAT)

Metadata:

Event number

Station

Cast

GEOTRACES SAMPLE NUMBER

Methods and reference samples

Metadata are essential for accurate and efficient data interpretation and use.



# LESSONS LEARNED (NAT)

Courtesy of everyone from US GEOTRACES  
North Atlantic Transect cruises ...

- Every instrument deployment is an event and must be entered in the event log
- Report event numbers and GEOTRACES sample numbers with the data

# SAMPLE IDENTIFICATION

Browser window: /RESTRICTED/GEOTRACES/NorthAtlanticTransect/CTD\_ODF\_Bottle\_GT10 -- Level 2 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

CTD - ODF/SIOR Bottle | BCO... x /RESTRICTED/GEOTRACES/N... x /RESTRICTED/GEOTRACES/N... x +

data.bco-dmo.org/jg/serv/RESTRICTED/GEOTRACES/NorthAtlanticTransect/CTD\_ODF\_Bottle\_GT10.brev ☆ ⓘ ↻ Google

Most Visited Getting Started Latest Headlines WHOI Internal BCO-DMO BCO data US JGOFS OCB Google Scholar Webster's

## /RESTRICTED/GEOTRACES/NorthAtlanticTransect /CTD\_ODF\_Bottle\_GT10 ---- Level 2

Directory Documentation Download & Other Operations

Level 0 Next Level

# Version: 02 August 2011  
# PIs: Jenkins, et al  
# KN199-04 ODF/SIOR CTD BOT  
#

=====

STNNBR	CASTNO	GEOTRC_EVENTNO
1	1	2008

=====

EXPCODE SECT\_ID

316N20101015

=====

GEOTRC\_SAMPNO

STNNBR	CASTNO	GEOTRC_SAMPNO	TIME	LATITUDE	LONGITUDE
1	1	5523	28	38.3269	-9.6625

=====

GEOTRC_SAMPNO	T_W	BTO_DATE	BTL_TIME	BTL_LAT	BTL_LON	BTMDEPTH	CTDPRS	CTDDEPTH	CTDTMP
5523		20101016	327	38.3292	-9.6655	2819	804.2	796.3	12.81
5522		20101016	323	38.329	-9.6652	2819	963.2	953.5	12.44
5521		20101016	319	38.3289	-9.665	2819	1048.3	1037.6	12.52
5520		20101016	315	38.3287	-9.6648	2819	1198.7	1186	12.31
5519		20101016	310	38.3285	-9.6646	2819	1348.8	1334.2	11.31

8 NTS-8 2

Woods Hole: 46°F Fri: 59°F Sat: 59°F

# GEOTRACES SAMPLE NUMBER

The unique identifier that enables you to connect data from different investigators.

For example ...

Iron speciation from Kristen Buck (BIOS)

Nanomolar Nutrients from Greg Cutter (ODU)

Cobalt data from Mak Saito (WHOI)

... event numbers indicate that  
all were sampled from Go-Flo bottles.

Data can be merged easily by matching the  
GEOTRC\_SAMPNO.

# LESSONS LEARNED (NAT)

The event number and GEOTRACES sample number enable you to answer this important question, all by yourself ...  
“when was my sample taken?”





# LESSONS LEARNED (NAT)

BCO-DMO data managers added the **GEOTRACES SAMPLE NUMBERS** (and other key fields) to the NAT data sets.

It is highly desirable to include event, station, cast, and **GEOTRACES** sample numbers in every data set contributed to BCO-DMO.

# LESSONS LEARNED (NAT)

Importance of event log and base bottle file:

- the event log and the base bottle data file (Niskin and GoFlo) enable merging of different datasets;
- it is imperative that the event log and bottle files be accurate;
- QC review is required;
- access should not be restricted.

# LESSONS LEARNED (NAT)

Consistent naming of common fields is highly desirable; BCO-DMO chose these for NAT data:

station\_GEOTRC

cast\_GEOTRC

event\_GEOTRC

sample\_GEOTRC (GEOTRACES sample #)

sample\_bottle\_GEOTRC (sample bottle #)

bottle\_GEOTRC (Niskin or GoFlo bottle)

depth\_GEOTRC\_CTD (depth from CTD P)

# LESSONS LEARNED (NAT)

## Depth and/or Pressure

- Know the difference!
- If a column is labeled depth, be certain that is what those numbers are.
- Perhaps reporting both would help?



# LESSONS LEARNED (NAT)

Units: (Greg Cutter)

- Report data as “per unit analyzed”
- e.g., if analysis results are pmol/kg then data should be reported in pmol/kg
- Conversion to per liter can be done using salinity from the bottle data file and lab temperature recorded during analysis (linked by the event number and GEOTRACES sample number).



# LESSONS LEARNED (NAT)

## Quality flags:

- Many investigators reported these with the NAT data
- But there was a lot of variation
- Flag definitions were not specified

Recommendation: adopt a common flag scheme for use in GEOTRACES

Ocean Data Standards volume 3

[http://www.iode.org/mg54\\_3](http://www.iode.org/mg54_3)



# QUALITY FLAGS

## QF flag scheme – primary level flags

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

The secondary level flags are optional, but can be used to represent details of quality assessment and control or data processing history.

# LESSONS LEARNED (NAT)

## Reference Materials: (Ken Bruland)

- Comparison with GEOTRACES Reference Station Samples (e.g. SAFe) and/or internal laboratory references must be reported with the metadata for each dataset
- Trace element reference concentrations are essential for inter-comparison

## Reference:

<http://www.geotraces.org/science/intercalibration>

# THANK YOU

## Questions ?



*Woods Hole, Massachusetts, USA*



[bco-dmo.org](http://bco-dmo.org) & [info@bco-dmo.org](mailto:info@bco-dmo.org)

32 of 31