

Building a Database for a Coastal Carbon Synthesis Project

Chris Osburn

**Dept. of Marine, Earth, and Atmospheric
Sciences**

NC State University

Coastal Synthesis Workshop, 12 DEC 2010



Acknowledgements

- NASA A.28 ROSES 2008 Program
 - John Haynes
 - Duane Armstrong
- Cyndy Chandler – BCO-DMO
- DOMSynth Team: Tom Bianchi, Bob Chen, Paula Coble, Eurico D'Sa
 - Xuchen Wang, Sam Harlow, Rick Smith, Jennifer Dickson-Brown



Talk Overview

- Background on the project
- Our approach to accumulating and re-formatting data
- Challenges we faced
- What we will do next

Project Overview: What We Wanted to Do

- Build a publicly accessible, geospatially-referenced database of CDOM absorption, DOC concentration, and oceanographic variables for the northern Gulf of Mexico
- Collate different pieces of information useful to scientists and decision makers
 - Light attenuation in estuarine and coastal waters
 - Terrestrial carbon flux into the ocean
- This amounted to a “data rescue” : compiling disparate data sets from multiple cruises into one “megacruise”
- Deliverables: a database linked to product generation

Two Examples of Databases

Data in a spreadsheet on my computer

Microsoft Excel - working peli data.xls

FileEditViewInsertFormatToolsDataWindowHelpAdobe PDF

Type a question for help

startmatlab putmatrix getmatrix evalstring getfigure

85%

A2	Station name										
A	B	C	D	E	F	G	H	I	J	K	L
1	Site ID	Atchafalaya	Atchafalaya	Atchafalaya	Atchafalaya	Atchafalaya	Atchafalaya	Atchafalaya	Atchafalaya	Atchafalaya	Atchafalaya
2	Station name	GOM01	GOM02	GOM02A	GOM02B	GOM02C	GOM03	GOM03A	GOM03B	GOM03C	GOM04
3	Date	20070507	20070508	20070508	20070508	20070508	20070508	20070508	20070508	20070508	20070508
4	Latitude	29.34817	28.205	28.205	28.205	28.205	28.94172	28.94172	28.94172	29.2125	29.24682
5	Longitude	-91.405	-91.8813	-91.8813	-91.8813	-91.8813	-91.6403	-91.6403	-91.6403	-91.5245	-91.4945
6	GMT	2112	1150	1226	1226	1226	1901	1901	1901	2138	2220
7	Depth	0	0	0	0	40	0	0	7	0	0
8	Temperature	25	23.4	NaN	23.5	21.5	24	24.4	24.1	29.1	27.5
9	Salinity	14.23	35.9	35.1	35.1	35.3	33	32.6	33.5	29.9	25.2
10	D.O.	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	Chl-a	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	DOC	260.2	58.8	59	59.3	58.9	108.1	108	NaN	145.9	192
13	250	27.839	1.596	1.624	1.602	1.790	4.006	4.107	3.226	4.410	14.055
14	251	27.529	1.555	1.581	1.555	1.744	3.939	4.051	3.172	4.345	13.895
15	252	27.237	1.514	1.542	1.517	1.705	3.877	4.000	3.119	4.280	13.737
16	253	26.955	1.476	1.504	1.479	1.667	3.820	3.955	3.074	4.221	13.588
17	254	26.675	1.444	1.469	1.447	1.630	3.761	3.908	3.034	4.162	13.436
18	255	26.415	1.411	1.431	1.411	1.594	3.711	3.863	2.988	4.104	13.286
19	256	26.161	1.377	1.399	1.379	1.559	3.658	3.821	2.949	4.054	13.140
20	257	25.891	1.346	1.368	1.346	1.526	3.607	3.781	2.912	4.002	12.992
21	258	25.631	1.317	1.339	1.317	1.496	3.556	3.742	2.873	3.953	12.850
22	259	25.372	1.285	1.310	1.288	1.467	3.508	3.703	2.838	3.896	12.707
23	260	25.100	1.257	1.280	1.257	1.438	3.458	3.660	2.801	3.845	12.568
24	261	24.841	1.229	1.252	1.229	1.409	3.407	3.618	2.765	3.795	12.434
25	262	24.574	1.198	1.221	1.201	1.381	3.357	3.576	2.723	3.742	12.297
26	263	24.300	1.171	1.194	1.171	1.351	3.306	3.531	2.686	3.688	12.159
27	264	24.030	1.146	1.166	1.143	1.323	3.256	3.486	2.647	3.635	12.018
28	265	23.753	1.114	1.137	1.114	1.291	3.197	3.435	2.607	3.579	11.872
29	266	23.481	1.086	1.106	1.086	1.263	3.143	3.385	2.569	3.524	11.730
30	267	23.192	1.056	1.076	1.056	1.233	3.090	3.337	2.526	3.466	11.585
31	268	22.889	1.025	1.048	1.025	1.202	3.034	3.286	2.487	3.404	11.431
32	269	22.593	0.997	1.020	0.997	1.174	2.977	3.233	2.447	3.345	11.276
33	270	22.299	0.964	0.990	0.964	1.144	2.924	3.177	2.405	3.289	11.121
34	271	22.002	0.934	0.962	0.934	1.114	2.867	3.121	2.359	3.230	10.967

DataSubmissionTemplate

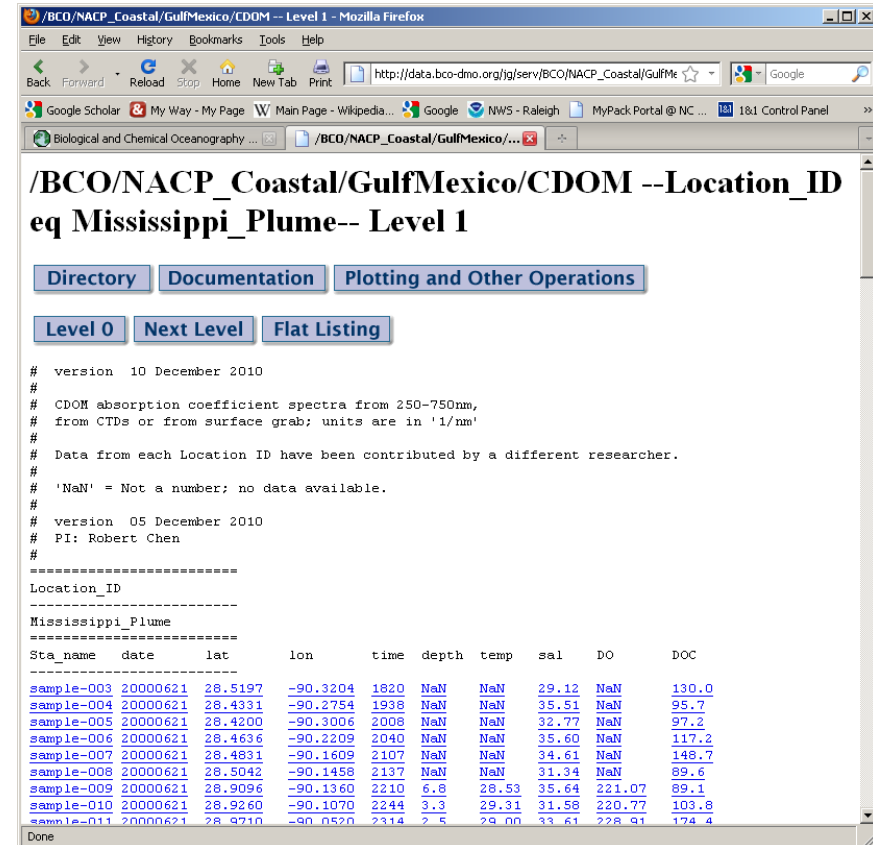
Header

times

Draw

AutoShapes

Data accessible via the Internet



The screenshot shows a web browser window displaying a page titled "/BCO/NACP Coastal/GulfMexico/CDOM --Level 1 - Mozilla Firefox". The page contains a table of data for CDOM absorption coefficient spectra from 250-750nm, organized by Location ID and Station name. The table includes columns for Sta_name, date, lat, lon, time, depth, temp, sal, DO, and DOC. The data is presented in a structured format with headers and footers.

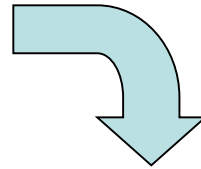
Sta_name	date	lat	lon	time	depth	temp	sal	DO	DOC
sample-003	20000621	28.5197	-90.3204	1820	NaN	NaN	29.12	NaN	130.0
sample-004	20000621	28.4331	-90.2754	1938	NaN	NaN	35.51	NaN	95.7
sample-005	20000621	28.4200	-90.3006	2008	NaN	NaN	32.77	NaN	97.2
sample-006	20000621	28.4636	-90.2209	2040	NaN	NaN	35.60	NaN	117.2
sample-007	20000621	28.4831	-90.1609	2107	NaN	NaN	34.61	NaN	148.7
sample-008	20000621	28.5042	-90.1458	2137	NaN	NaN	31.34	NaN	89.6
sample-009	20000621	28.9096	-90.1360	2210	6.8	28.53	35.64	221.07	89.1
sample-010	20000621	28.9260	-90.1070	2244	3.3	29.31	31.58	220.77	103.8
sample-011	20000621	28.9710	-90.0520	2314	2.5	29.00	33.61	228.91	174.4

“How Can I Go from Left to Right?”

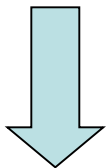
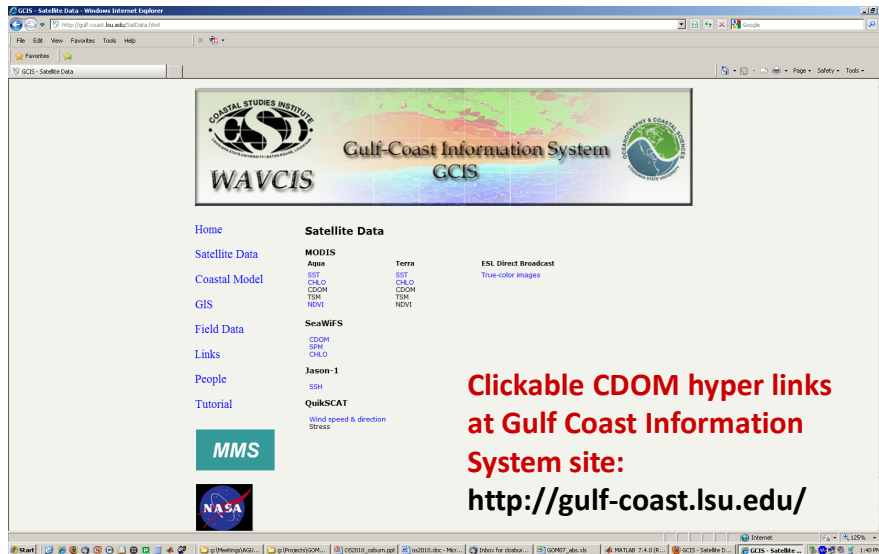
Database Design

- Hosted by BCO-DMO
- Contributions from PI's
- Contributions from scientific community
- Key oceanographic and biogeochemical data on CDOM

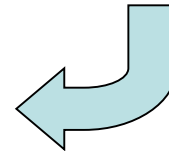
Community Contributions of Data



Format and Upload
contributed data

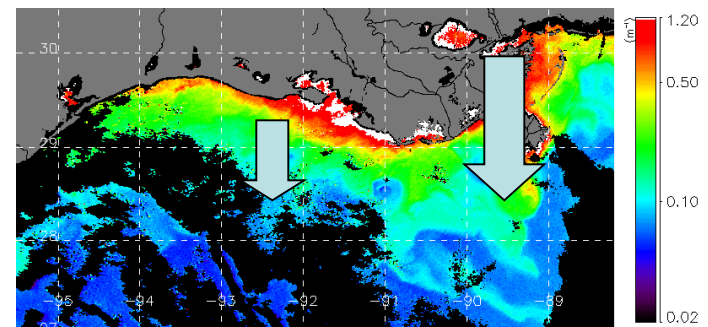


Validate algorithms
relating satellite
reflectance to
surface CDOM
absorption



Link Databases

Produce: Surface CDOM
absorption maps at 412 nm
(example: SeaWiFS imagery for
19 Sep 2005)



Defining the Database

- How do you want to view and access your data?
 - Date? Month? Year?
 - Depth?
 - Salinity?
 - Position?
- Consulted end-users
 - Decided on the most critical data needed to accomplish your goals
 - We talked with resource managers in addition to a variety of scientists and a real database person (Cyndy Chandler)
 - Charles Kovach (FL DEP)
 - Judy Ott (Charlotte Harbor NEP)
 - Carlos Del Castillo (APL, Johns Hopkins)
 - What *functionality* do you want to have?

Defining the Database

- Agreed on a data submission structure and made *everyone* adhere to it
 - Editing is time consuming!
- Used the BCO-DMO platform and data metadata documents
 - Answers to questions you haven't yet thought of...

Metadata – Go back to the Basics

- Tells the who, what, when, where, and how
- Ensures that users can know acquisition conditions
 - Instrument model and operating parameters
 - Clear definition of units
- Serves as a way to QA/QC the data

Good teachable tool for students!

Collating the Data: Create a Template

The screenshot displays a Microsoft Excel spreadsheet titled "DataSubmissionTemplate.xls". The spreadsheet is designed for data submission and includes a list of instructions in column A and a corresponding data entry area in columns B through V.

Instructions (Column A):

- 1. **NOTES, Do Not Fill In This Column**
- 2. Keep less than 16 characters
- 3. Keep less than 12 characters
- 4. format as YYYYMMDD
- 5. Decimal latitude
- 6. Decimal longitude (indicate W w/ '-')
- 7. GMT, in decimal time format
- 8. meters
- 9. deg. Centigrade
- 10. PSU or unitless
- 11. specify, mg/L preferred
- 12. specify, ug/L preferred
- 13. micromolar
- 14.
- 15. NOTE: If values are lacking, please
- 16. type in NaN
- 17.

Data Entry Area (Columns B-V):

The data entry area is organized into columns corresponding to the instructions. The first column (B) contains site IDs from 250 to 295. The subsequent columns (C-V) are currently empty, providing space for data entry.

Columns B-V Headers:

- B: Site_ID
- C: Station_name
- D: Date
- E: Latitude
- F: Longitude
- G: GMT
- H: Depth
- I: Temperature
- J: Salinity
- K: D.O.
- L: Chl-a
- M: DOC
- N:
- O:
- P:
- Q:
- R:
- S:
- T:
- U:
- V:

The spreadsheet is displayed in the Microsoft Excel application window, showing the standard menu bar (File, Edit, View, Insert, Format, Tools, Data, Window, Help) and the status bar at the bottom.

Some formatting guidelines

- Structure the data in the same way
- Ensure at minimum:
 - “xyzt” data:
 - x = lat
 - y = long
 - z = depth
 - t = time
- Time in YYYYMMDD, which can be separated into Year and Month
- ‘NaN’ = for no values (e.g., in Matlab)
 - Don’t leave any blanks!
- Official cruise ID (or equivalent) if known

Example: Our CDOM Database

Biological and Chemical Oceanography Data Management Office - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Back Forward Reload Stop Home New Tab Print <http://osprey.bcodmo.org/project.cfm?flag=view&id=29&s> Google

Google Scholar My Way - My Page Main Page - Wikipedia... Google NWS - Raleigh MyPack Portal @ NC ... 1&1 Control Panel

Biological and Chemical Oceanog... My Way - My Page

BCO-DMO

Biological and Chemical Oceanography Data Management Office

Home Data Resources About Us Contact

Database

- Welcome
- Programs 16
- Projects 131
- Deployments 806
- Instruments 205
- Datasets 4047
- People 1045
- Affiliations 228
- Funding 34
- Parameters 1101

Data Access

- Geospatial access

Project: Gulf of Mexico NACP-OCB Coastal Synthesis

Acronym: GMx NACP-OCB

Programs:

- Ocean Carbon & Biogeochemistry [OCB]
- North American Carbon Program [NACP]
- NACP-OCB Coastal Synthesis [NACP-OCB Coastal]

URL: Project Web Site

Start date: 2008-06

End date:

Geolocation: global coastal zones

► **Description:**

More information about project Gulf of Mexico NACP-OCB Coastal Synthesis

► **Funding**

▼ **Datasets associated with Gulf of Mexico NACP-OCB Coastal Synthesis**

Dataset

GMx_CDOM

► **Platform deployments associated with Gulf of Mexico NACP-OCB Coastal Synthesis**

► **People associated with Gulf of Mexico NACP-OCB Coastal Synthesis**

©2008 Biological and Chemical Oceanography Data Management Office.
Funded by the U.S. National Science Foundation

Done

Biological and Chemical Oceanography Data Management Office - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Back Forward Reload Stop Home New Tab Print <http://osprey.bcodmo.org/dataset.cfm?id=13650&flag=view> Google

Google Scholar My Way - My Page Main Page - Wikipedia... Google NWS - Raleigh MyPack Portal @ NC ... 1&1 1&1 Control Panel

Biological and Chemical Oceanograp...

BCO-DMO

Biological and Chemical Oceanography Data Management Office

Home Data Resources About Us Contact

Database

- Welcome
- Programs 16
- Projects 131
- Deployments 806
- Instruments 205
- Datasets 4047
- People 1045
- Affiliations 228
- Funding 34
- Parameters 1101

Data Access

- Geospatial access

Dataset: GMx_CDOM

[Get Data](#)

Project: Gulf of Mexico NACP-OCB Coastal Synthesis

Validated: yes

Data version: 10 December 2010

Version date: Dec 10, 2010

► **Current state:** preliminary and in progress

▼ **Description:** Absorption spectra and DOC concentrations

Geospatial Synthesis of Chromophoric Dissolved Organic Matter Distribution in the Gulf of Mexico

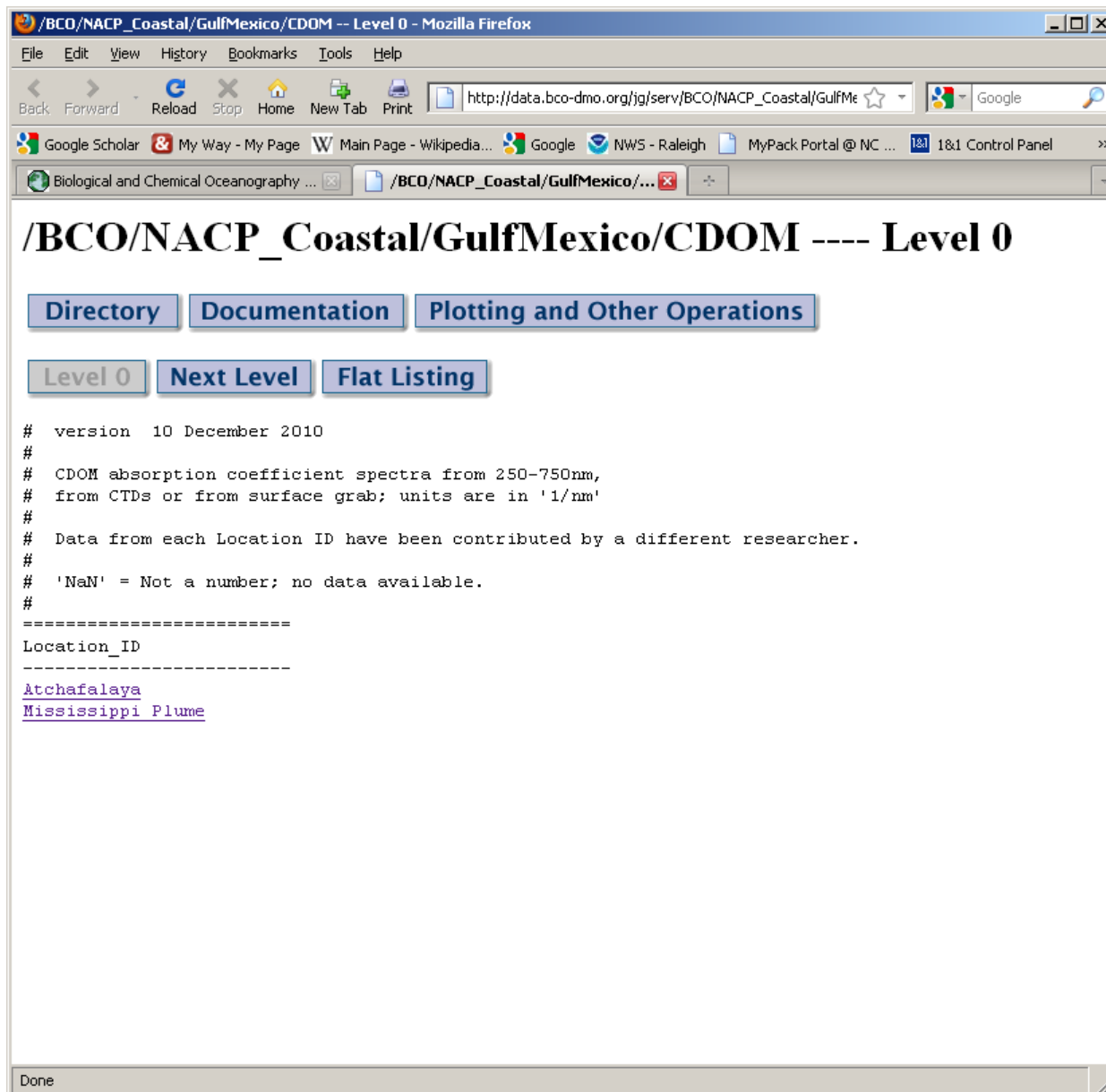
PI: Christopher Osburn, Dept. Marine, Earth, and Atmospheric Sciences, NC State University

Co-PIs: Thomas Bianchi, Texas A&M University, Bob Chen, University of Massachusetts-Boston; Paula Coble, University of South Florida; Eurico D'Sa, Louisiana State University

This dataset is considered an affiliated program of NACP >> see the [entry at the NACP site](#).

This data set will be a synthesis of bio-optical data on chromophoric dissolved organic matter (CDOM) in the coastal regions the Gulf of Mexico - an immediate goal of the NACP and OCB programs. Absorption spectra and DOC concentrations from the Organic Matter Cycli project will be reported. In addition to in situ absorption spectra and DOC concentrations some study locations may also report fluorescence (estimate of chlorophyll a concentration), plus CTD and dissolved oxygen measurements

Done



/BCO/NACP_Coastal/GulfMexico/CDOM -- Level 1

File Edit View History Bookmarks Tools Help

Back Forward Reload Stop Home New Tab Print http://data.bco-dmo.org/jg/serv/BCO/NACP_Coastal/GulfMexico/ Google

Google Scholar My Way - My Page Main Page - Wikipedia... Google NWS - Raleigh MyPack Portal @ NC ... 1&1 Control Panel

Biological and Chemical Oceanography ... /BCO/NACP_Coastal/GulfMexico/...

/BCO/NACP_Coastal/GulfMexico/CDOM --Location_ID eq Atchafalaya-- Level 1

[Directory](#) [Documentation](#) [Plotting and Other Operations](#)

[Level 0](#) [Next Level](#) [Flat Listing](#)

```
# version 10 December 2010
#
# CDOM absorption coefficient spectra from 250-750nm,
# from CTDs or from surface grab; units are in '1/nm'
#
# Data from each Location ID have been contributed by a different researcher.
#
# 'NaN' = Not a number; no data available.
#
# version: 03 September 2010
# PI: Chris Osburn
#
=====
Location_ID
-----
Atchafalaya
=====
Sta_name  date      lat      lon      time  depth  temp  sal  DO  DOC
-----
GOM01     20070507  29.3482  -91.4050  2112  0      25     14.23  NaN  260.2
GOM02     20070508  28.2050  -91.8813  1150  0      23.4   35.9   NaN  58.8
GOM02A    20070508  28.2050  -91.8813  1226  0      NaN    35.1   NaN  59
GOM02B    20070508  28.2050  -91.8813  1226  10     23.5   35.1   NaN  59.3
GOM02C    20070508  28.2050  -91.8813  1226  40     21.5   35.3   NaN  58.9
GOM03     20070508  28.9417  -91.6403  1901  0      24     33.0   NaN  108.1
GOM03A    20070508  28.9417  -91.6403  1901  0      24.4   32.6   NaN  108
GOM03B    20070508  28.9417  -91.6403  1901  7      24.1   33.5   NaN  NaN
GOMS30    20070508  29.2125  -91.5245  2138  0      29.1   29.9   NaN  145.9
```

Done

BCO/NACP_Coastal/GulfMexico/CDOM -- Level 2 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Back Forward Reload Stop Home New Tab Print http://data.bco-dmo.org/jg/serv/BCO/NACP_Coastal/GulfMe

Google Scholar My Way - My Page Main Page - Wikipedia... Google NWS - Raleigh MyPack Portal @ NC ... 1&1 Control Panel

Biological and Chemical Oceanography ... /BCO/NACP_Coastal/GulfMexico/...

/BCO/NACP_Coastal/GulfMexico/CDOM --Location_ID eq Atchafalaya,Sta_name eq GOM01-- Level 2




[Directory](#) [Documentation](#) [Plotting and Other Operations](#)

[Level 0](#) [Next Level](#) [Flat Listing](#)

```
# version 10 December 2010
#
# CDOM absorption coefficient spectra from 250-750nm,
# from CTDs or from surface grab; units are in '1/nm'
#
# Data from each Location ID have been contributed by a different researcher.
#
# 'NaN' = Not a number; no data available.
#
# version: 03 September 2010
# PI: Chris Osburn
#
# GOM01
# Date created: 26-Aug-2010 17:53:28
=====
Location_ID
-----
Atchafalaya
=====
Sta_name  date      lat      lon      time  depth  temp  sal   DO     DOC
-----
GOM01     20070507  29.3482  -91.4050  2112  0      25    14.23  NaN    260.2
=====
Wavelength  a_lambda
-----
250          27.8391
251          27.5289
252          27.2374
```

Done

GIS Interface Output from BCO-DMO

 **Biological and Chemical Oceanography Data Collection**  

Welcome to the MapServer Geospatial Interface.

Available Data

Simple search | Advanced search

FeSynth (13) 1

Historical (11)

IMBER-US (1)

LTER

NACP (2)

Select Project(s): 2

GMx: NACP-OCB (2)

GMx: NACP-OCB (2)

AMT

ANACONDAS (1)

Select Deployment(s): 3

PE07-32

PEJun2000

RB-07-05

Show selected

Show and zoom

- OR -

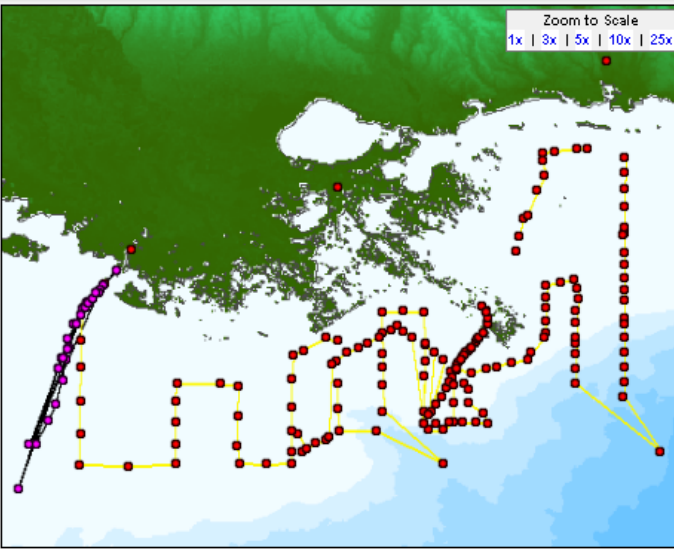
Show all

Show Data 4

Interactive Map

Click on map-it obs. sites to view more info. Hold down the Z key to lock in on an obs. site.

Zoom to Scale: 1x | 3x | 5x | 10x | 25x



28°09'42" N 89°52'24" W

Scale-it legend: 1 - 2 3 - 10 11 - 100 101 - 1000 above 1000

☒ ☒ PEJun2000 GMx_CDOM (170/198 mapped) [WFS](#) | [URL](#) | [View status](#)

☒ ☒ PE07-32 GMx_CDOM (31/47 mapped) [WFS](#) | [URL](#) | [View status](#)

Map

PEJun2000 GMx_CDOM CDOM

30 35 56 N 88 35 56 W

Available on-line data

Choose axes to graph or get from Sta_name sample-124.

X Y

a_lambda Wavelength


[graph it](#) [get it](#) [WFS](#)

Data Layers

☒ Show bathy?

☐ Show grid?

Map Size



Map Focus

☐ Atlantic

☒ Pacific

OGC Info

[GetCapabilities](#)

[Map WMS](#)

[Show Data WFS](#)

[Show Command Log](#)



version 0.600

Metadata [Clear](#)



Show Data Results

Click on a link for more info.

Green links indicate mapable data are available.

PE07-32  

GMx_CDOM



PEJun2000  

GMx_CDOM

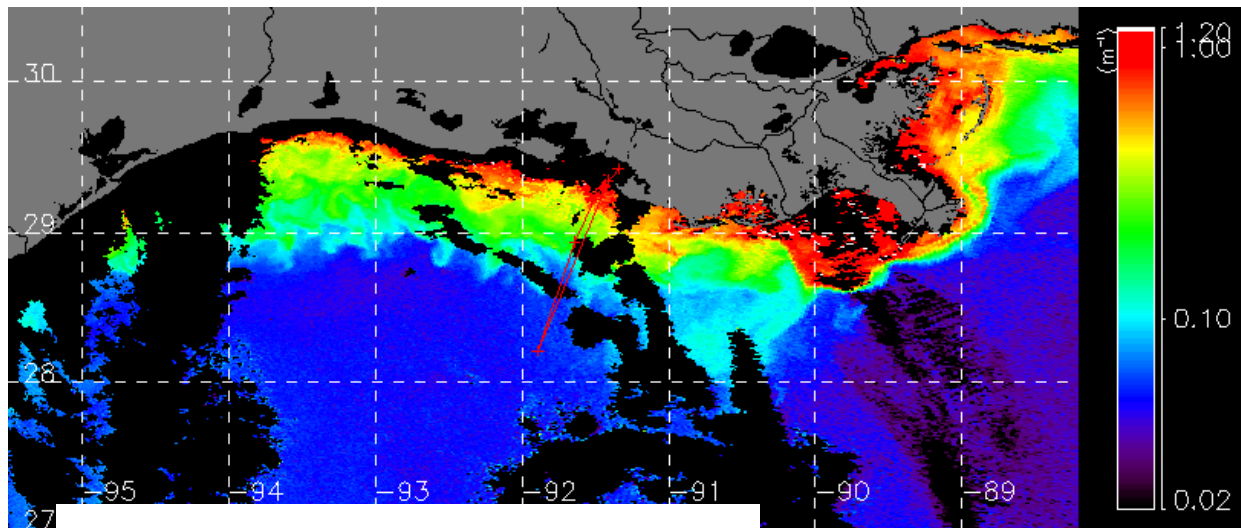
5

GMx_CDOM

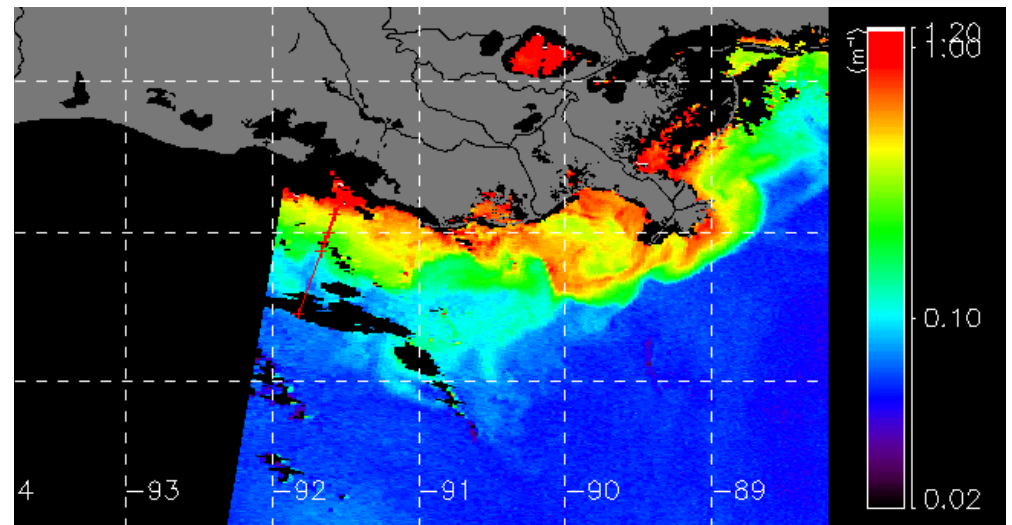
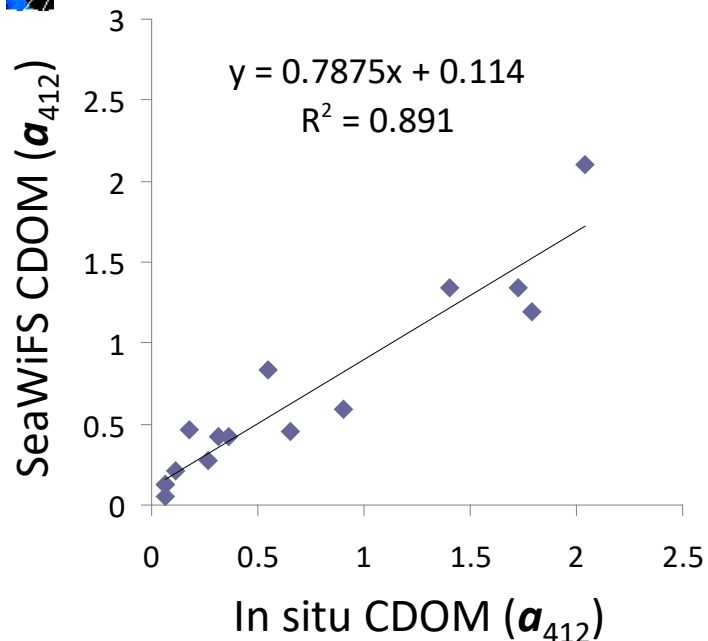
Available data

GMx_CDOM   [map it](#) 6

Some Initial Results



SeaWiFS satellite-derived CDOM absorption at 412 nm for 8 May 2007 (TOP) and 9 May 2007 (BOTTOM RIGHT), using the D'Sa et al. 2006 algorithm. The performance of the algorithm is shown (BOTTOM LEFT). Red line shows the ship track on these images.



Where We Will Go Next...

- Continue satellite cal/val and generation of light attenuation maps
 - (Gulf Coast Information System; Eurico D'Sa, LSU)
- Open data call
- Data sharing activity (Ken Keiser, UAH)
 - Generate Data Casting (GeoRSS) feeds and KML files of the data and data products;
 - User friendly product that is easily accessible by wireless device
 - Recommendations from Deepwater Horizon responders
- Synthesis paper for CDOM and DOC fluxes into the Gulf of Mexico
 - Develop hypotheses, identify knowledge gaps

Recommendations

- Consult with your end-users as much as possible
 - What *functionality* do you want to have?
- Develop reasonable data admission guidelines and adhere to them
 - Serve as QA/QC
- Utilize BCO-DMO metadata guidelines to shape your submission protocols
 - Implore the use of metadata documentation
- Contact: Chris Osburn, closburn@ncsu.edu