

*U.S. Integrated Ocean Observing System (IOOS®)*

# Improving Aquatic Biological/Ecological Observing Data Integration and Dissemination

---

*A Multi-Agency, Multi-Partners Effort to reconcile  
Data/Metadata Standards and Enable Access to  
Integrated Aquatic  
Obs*

*Hassan Mou*

NOAA EDMC

---

**IOOS BDP and Bio. Core Var.**

---

**ATN Obs**

- TOPP to Ocean models
  - POST to Fisheries models
- 

**Acoustic Obs**

---

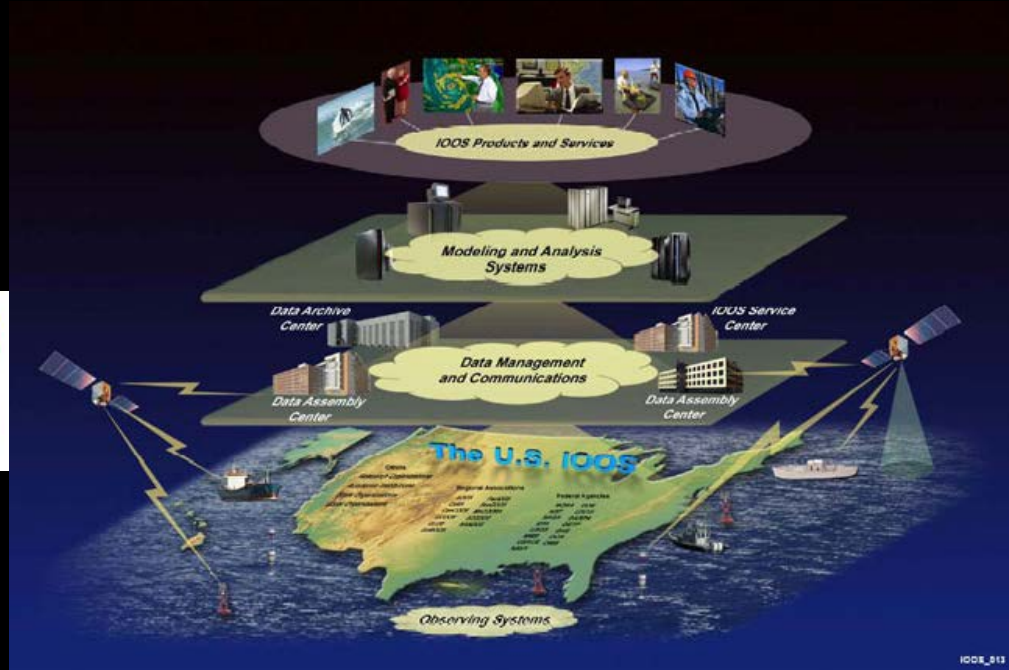
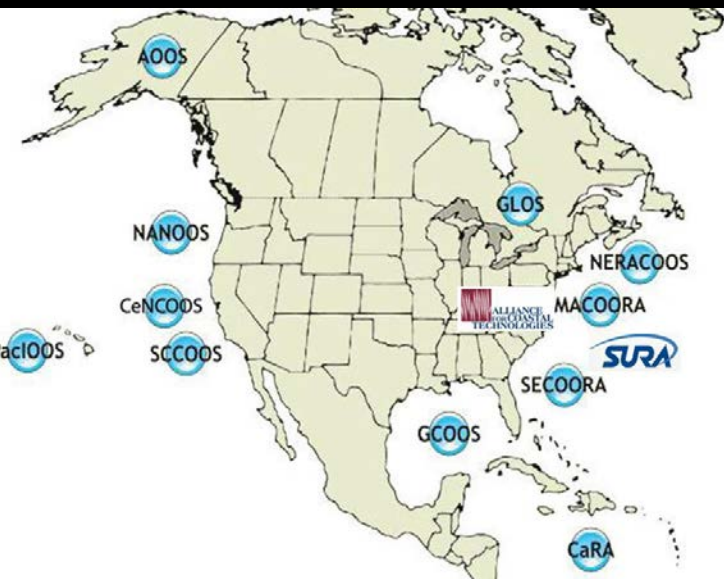
**Looking Forward**



# U.S. IOOS: Program Overview

## Global Component

## Coastal Component (EEZ to the head of tide)



## 7 Goals, 1 System

- Improve predictions of **climate change and weather**
- Improve the safety and efficiency of **maritime operations**
- Improve forecasts of **natural hazards**
- Improve **homeland security**
- Minimize **public health risks**
- Protect and restore **healthy coastal ecosystems**
- Sustain living **marine resources**

## Enhances science and improves decision making



US IOOS BDP ATN TOPP POST ACOUSTIC CHALLENGES

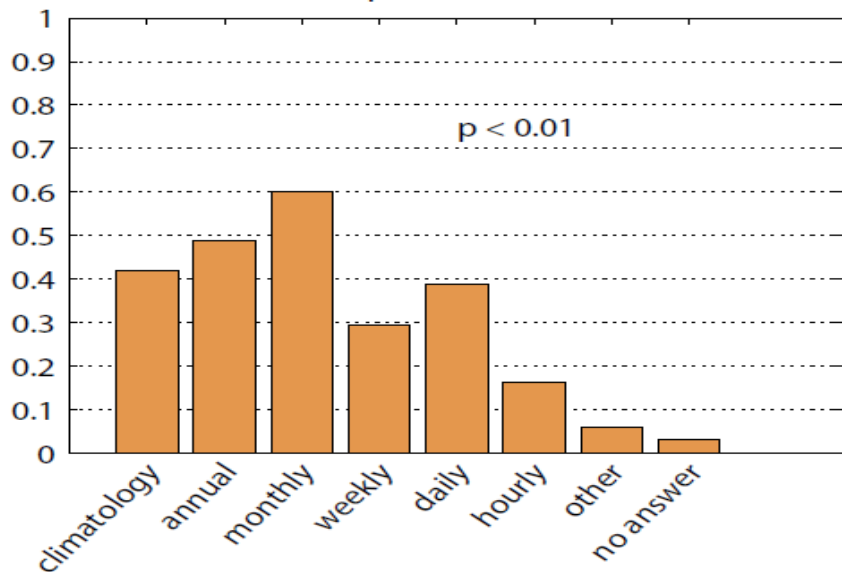


# IOOS<sup>®</sup> Core Variables

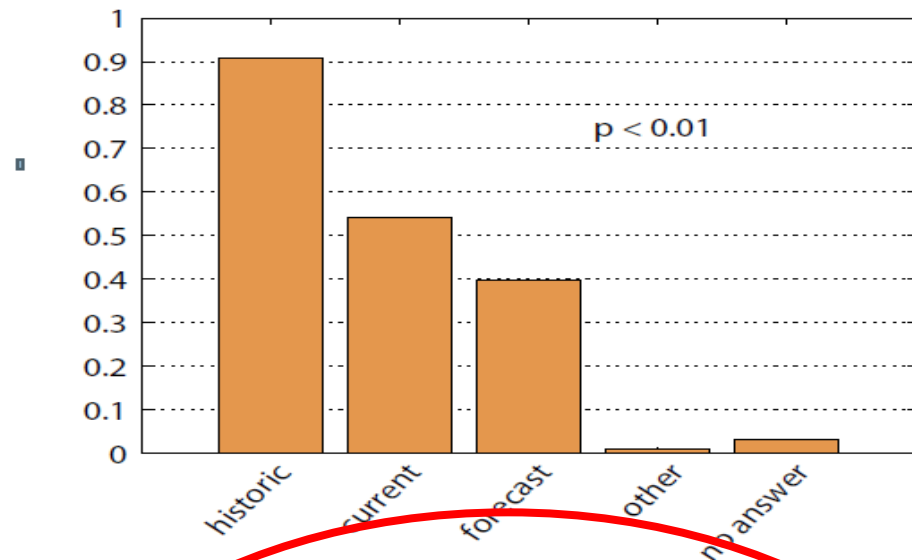
1. **Temperature**
2. **Salinity**
3. **Water level**
4. **Currents**
5. **Surface Waves**
6. **Surface Winds**
7. **Ocean color**
8. Dissolved oxygen
9. pH
10. pCO<sup>2</sup>
11. Heat flux
12. Bottom character
13. Pathogens
14. Bathymetry
15. Ice distribution
16. Contaminants
17. Stream flow
18. Dissolved nutrients
19. Optical properties
20. Total suspended matter
21. Colored dissolved organic matter
- 22. Fish species**
- 23. Fish abundance**
24. Zooplankton species
25. Phytoplankton species
26. Zooplankton abundance

*list could be expanded based on identified needs...*

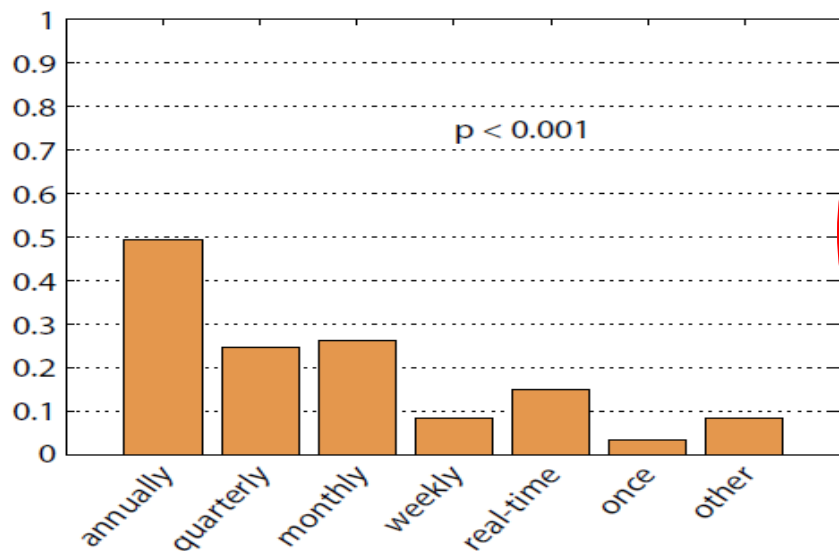
Which scales are you most interested in?  
Temporal resolution



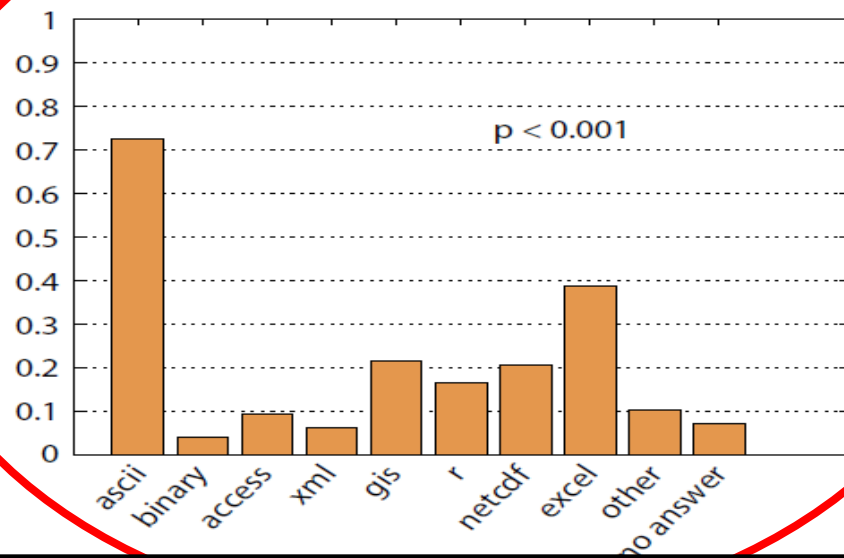
Which scales are you most interested in?  
Time Frame



How frequently will you require the products to be delivered?  
Regular (e.g., monitoring, advice, reporting) — please specify below



Which file formats are most convenient for you?



# Metadata Standards, Content and Access

**Metadata:** information used to document and interpret data

## Metadata Contents

- Purpose
  - Identification / Attribution
  - Discovery
  - Understanding
  - Referral
  - History / QA
  - (Details of Mapping to Terminology)

## Content Standard

FGDC is the Primary Format

- ISO and ncISO Supported
- Multiple Interface Methods

## Metadata Access

FGDC Document of Record

Tools transform FGDC for other formats and methods of access; tools by project team and external.

ncISO on ERDDAP

Native FGDC xml via http://

Converted ISO xml via http://

ERDDAP  
Easier access to scientific data

ERDDAP > info > CRED\_PMNM\_F

Variables and Attributes

Type	Variable Name	Attribute Name	Data Type
attribute NC_GLOBAL	cdm_data_type		String
attribute NC_GLOBAL	Conventions		String
attribute NC_GLOBAL	Eastemmost_Easting		double
attribute NC_GLOBAL	geospatial_lat_max		double
attribute NC_GLOBAL	geospatial_lat_min		double
attribute NC_GLOBAL	geospatial_lon_max		double
attribute NC_GLOBAL	geospatial_lon_min		double
attribute NC_GLOBAL	geospatial_units		String
attribute NC_GLOBAL	history		String
attribute NC_GLOBAL	infoUrl		String
attribute NC_GLOBAL	institution		String

FGDC & ISO; html, xml, text

Credit to OBIS-USA, PIFSC, NPS, PMNM and PacIOOS team





# Links to IOOS BDP data services?

- Data Terminology and exchange standards

- *Data Model XML*

[http://www.ioos.gov/schema/ioosbiology/0.9/ioosbiologicalterminology\\_vpoint9.xsd](http://www.ioos.gov/schema/ioosbiology/0.9/ioosbiologicalterminology_vpoint9.xsd)

- *Terminology Schema*

[http://www.ioos.gov/schema/ioosbiology/0.9/ioos\\_biological\\_terminology\\_vpoint9d.xml](http://www.ioos.gov/schema/ioosbiology/0.9/ioos_biological_terminology_vpoint9d.xml)

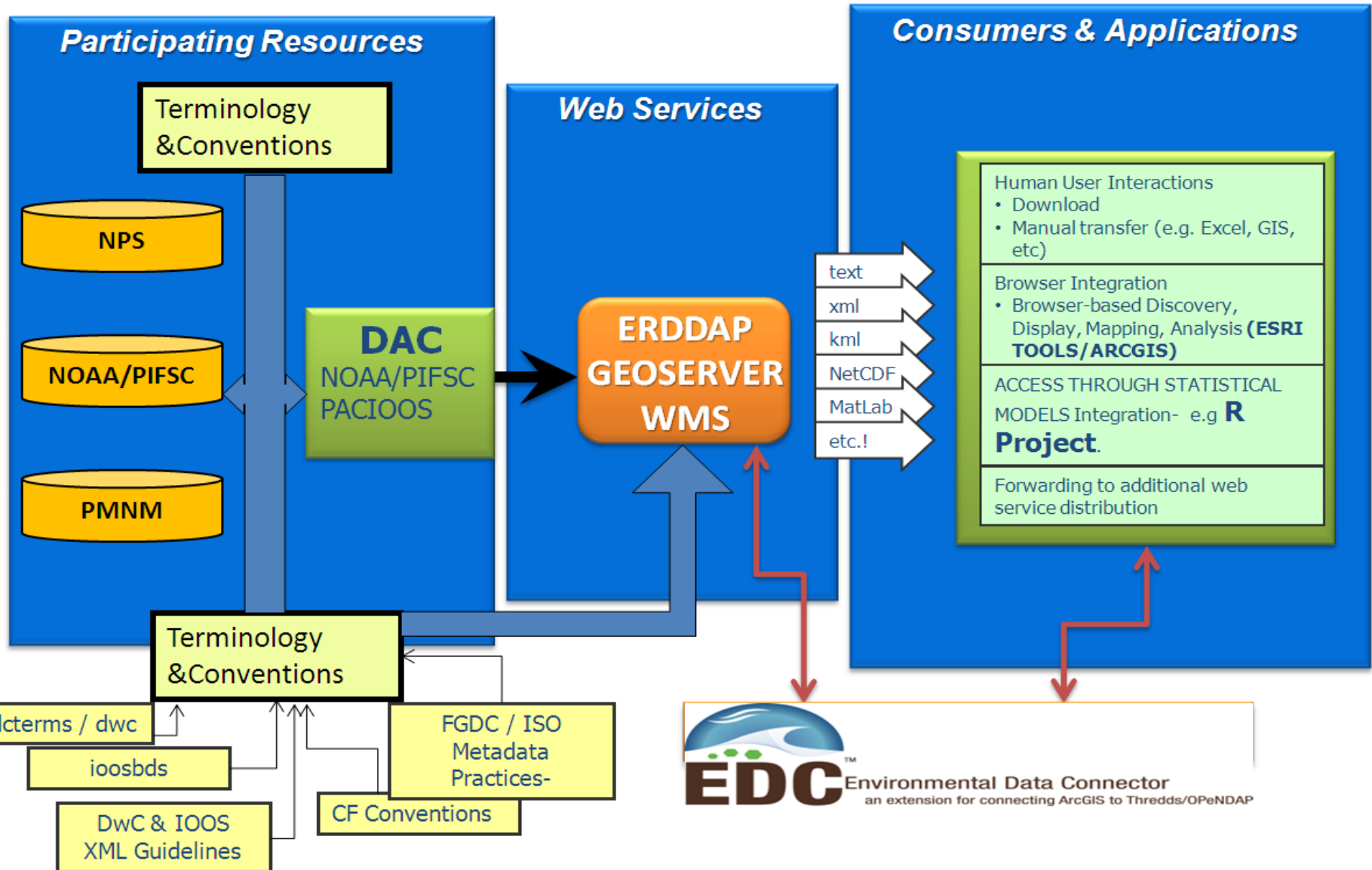
- Data Access services

<http://oos.soest.hawaii.edu/erddap/info/index.html?page=1&itemsPerPage=1000>

Metadata Standards

[http://oos.soest.hawaii.edu/cgi-bin/get\\_metadata.pl?id=PACN\\_FISH\\_TRANSECT&format=fgdc](http://oos.soest.hawaii.edu/cgi-bin/get_metadata.pl?id=PACN_FISH_TRANSECT&format=fgdc)

# PILOT IN PACIOOS



- CF terms were applied at the field level for Lat, Long and Time fields.

There is a reason for doing this.

integrate Biological and other environmental data.

HOW ABOUT NETCDF files ?

ERDDAP -> NetCDF



## Software

Software/Technology Solutions

⊕ OILMAP™

SARMAP™

CHEMMAP™

Environmental Data Server™

AIRMAP™

HYDROMAP™

MUDMAP™

SIMAP™

WQMAP™

⊕ ASA In-house Tools

⊕ ArcGIS® Tools

## GIS & Data Integration

## ASA Software Downloads

## Software

### Downloads : Freeware Catalogue

	Name	Size	Description
<a href="#">Download</a>	<b>Environmental Data Connector MATLAB</b>	106 MB	The Environmental Data Connector (EDC) extension uses a Java-based browser to allow users to connect to THREDDS/SOS/ERDDAP data servers. The connector leverages existing components from the Unidata libraries so that users can filter large amounts of data in space and time. The data and metadata is then converted to into a convenient Matlab struct object and loaded into the workspace for analysis.
<a href="#">Download</a>	<b>Environmental Data Connector R (Windows Package)</b>	86.1 MB	The Environmental Data Connector (EDC) extension uses a Java-based browser to allow users to connect to THREDDS/SOS/ERDDAP data servers. The connector leverages existing components from the Unidata libraries so that users can filter large amounts of data in space and time. The data and metadata is loaded into the R workspace as a named list for analysis.
<a href="#">Download</a>	<b>Environmental Data Connector R (32-bit Macintosh Package)</b>	86.5 MB	The Environmental Data Connector (EDC) extension uses a Java-based browser to allow users to connect to THREDDS/SOS/ERDDAP data servers. The connector leverages existing components from the Unidata libraries so that users can filter large amounts of data in space and time. The data and metadata is loaded into the R workspace as a named list for analysis.



# Links to IOOS BDP data services?

- Data Terminology and exchange standards

- *Data Model XML*

[http://www.ioos.gov/schema/ioosbiology/1\\_0/ioos\\_biological\\_terminology20120608v1point0.xsd](http://www.ioos.gov/schema/ioosbiology/1_0/ioos_biological_terminology20120608v1point0.xsd)

- *Terminology Schema*

[http://www.ioos.gov/schema/ioosbiology/1\\_0/ioos\\_biological\\_terminology20120608v1point0.xml](http://www.ioos.gov/schema/ioosbiology/1_0/ioos_biological_terminology20120608v1point0.xml)

- Data Access services

<http://pacioos-mapserver2.ancl.hawaii.edu/erddap/index.html>

Metadata Standards

[http://oos.soest.hawaii.edu/cgi-bin/get\\_metadata.pl?id=PACN\\_FISH\\_TRANSECT&format=fgdc](http://oos.soest.hawaii.edu/cgi-bin/get_metadata.pl?id=PACN_FISH_TRANSECT&format=fgdc)



# A True Partnership effort

- ❖ U.S. Census of Marine Life (*Michael Feldman*)
- ❖ NOAA Fisheries
  - Pacific Islands Fisheries Science Ctr. (*Rusty Brainard , Anntte DesRochers, Troy Kanemura et al*)
  - Southwest Fisheries Science Ctr. (*Bob Simmons, Roy Mendelssohn*)
  - Office of Science and Technlgy. (*Jim Sargent, Becky Shuford*)
- ❖ NOAA Sanctuaries/Monument (*Randy Kosaki, David Graham, Steve Turner*)
- ❖ University of Hawai'i
  - PacIOOS (*Jim Potemra, Chris Ostrander*)
  - HIMB (*Jo-Ann Leong, Erik Franklin*)
- ❖ USGS and USA-OBIS (*Mark Fornwall, Phillip Goldstein*)
- ❖ NOAA IOOS Program (*Charly Alexander, Hassan Moustahfid*)



# The WORLD of Standards and the World of Interoperability

Others  
e.g. Bio  
DwC  
Dublin Core  
DiGGiR  
Etc.

CF  
NETCDF  
SOS  
TDS  
...

## Dwc Archive

<http://www.gbif.org/informatics/standards-and-tools/publishing-data/data-standards/darwin-core-archives/>

## GBIF IPT

<http://www.gbif.org/informatics/infrastructure/publishing/>

# Looking Forward

- 2011- Reconciled a community data content standards for Bio. Core Var. (i.e. Fish species and fish abundance) as required by IOOS community.
- 2011- Pilot implementation in PacIOOS with three datasets (NOAA. Fisheries, NPS, PMNM) link to fisheries models.
- 2012-13 Testing extensibility of these Stds to other bio/eco. data and in other geographies - GCOOS and SECOORA.

## **GCOOS-RA**

NOAA. Fisheries data (Trawl data, fish, crabs, shrimp etc )  
<http://barataria.tamu.edu/erddap/info/index.html?page=1&itemsPerPage=1000>

## **SECOORA-RA**

FWC FIM, MARMAP data (Different Taxa: Fish, Shrimp, plankton, etc.)

ERDDAP is deployed and waiting for connect data

<http://129.252.139.124/erddap>

# MAPPING ORIGINAL DATA TO IOOS BDP TERMS

IOOS Biological Mapping Document for CAGES Texas data  
 E. Belmonte, P. Goldstein  
 Original Draft: August 15, 2012  
 Last Updated: September 5, 2012

Find the definitions of the following IOOS Biology terms here:

[http://www.ioos.gov/schema/ioosbiology/1\\_0/ioos\\_biological\\_terminology20120608v1point0.xml](http://www.ioos.gov/schema/ioosbiology/1_0/ioos_biological_terminology20120608v1point0.xml)

See also:

<http://www.ioos.gov/dmac/biology/welcome.html>

and <http://www.ioos.gov/dmac/biology/welcome.html>

IOOS Biology Term	Source Field in CAGES Texas data or Literal (literals in quotes)	Comments
<del>modified</del>	"2012-08-01"	This is the date on which we received the data. Is there a better date to use? Publication date listed in metadata (2011-05-05)? <u>I would suggest the date you received the data. I need to revise the metadata since changes have been made to the database.</u>
<del>verbatimModified</del>		
<del>higherInstitutionCode</del>	"NOAA;NOAA Fisheries Service;SEFSC; Fishery Ecology Branch"	Confirm values. Do we want "DOC" for Department of Commerce at the top? Other Science Centers have used DOC at the top. <u>Use of DOC at the top is probably a good idea.</u>  For this and the other <del>institutionCode</del> -related terms, below: has the usage of NMFS name been changed? In the documentation provided with CAGES data the organization is referred to as "NOAA Fisheries Service, formerly National Marine Fisheries Service, Fishery Ecology Branch." <u>Frankly I've always found the interchangeable use of NOAA Fisheries and NMFS confusing. I'll have to get back to you on which is the most current and correct usage.</u>
<del>institutionCode</del>	"NOAA Fisheries Service"	Confirm value; also see above
<del>ownerInstitutionCode</del>	"NOAA Fisheries Service"	Confirm value; also see above
<del>collectionCode</del>	"CAGES"	Confirm value; we can also sub-divide the collection and identify as "CAGES Texas". <u>We can subdivide if you wish. I think it would be a good idea to subdivide the database if you do. Subdivision of the database is necessary in order to accommodate more data in the future. We are approaching the limits of what Access can handle as more data comes in.</u>





- FY13- Finalize a reference implementation plan to implement IOOS BDP services and improve access to IOOS Bio. Core Var. (i.e. fish species and fish abundance etc.).
- Note that these Bio data services can be used for any TAXA Not only FISH from Bacteria to Whales that have presence absence and abundance data
- Also Life history data (Length, Age Weight etc)

(Address IOOS DMAC concerns related to SOS vs ERDDAP, CF vs Darwin Core etc.)

Questions?

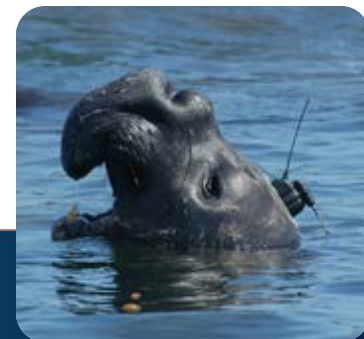
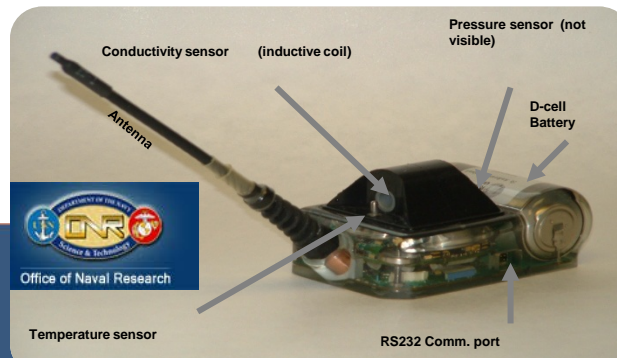
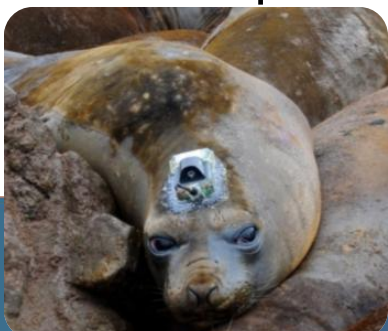
# Animal Telemetry Obs.

- Two projects to improve access to animal telemetry observations and reconcile animal telemetry data standards:

**Project 1:** enabling ocean modelers access to historic physical oceanographic observations collected from CTD sensors on marine animals (e.g. e-seals and sharks) placed by Stanford University's Tagging of Pacific Predators (TOPP) Program; ONR, NAVOCEANO, NCEP, TOPP

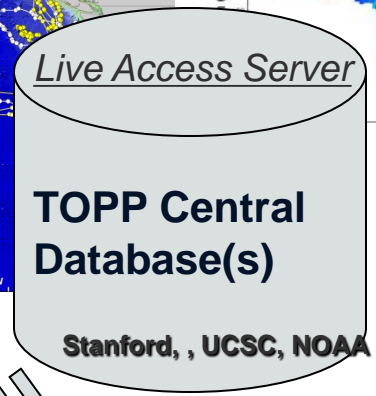
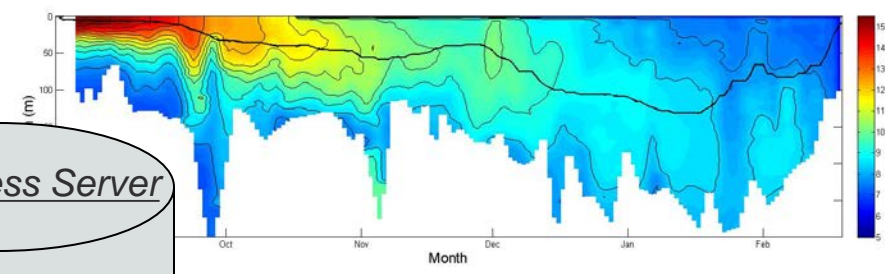
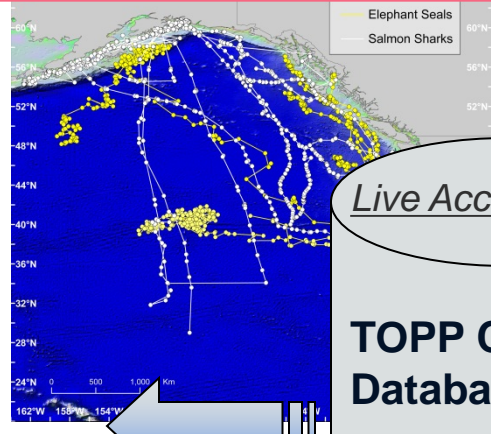
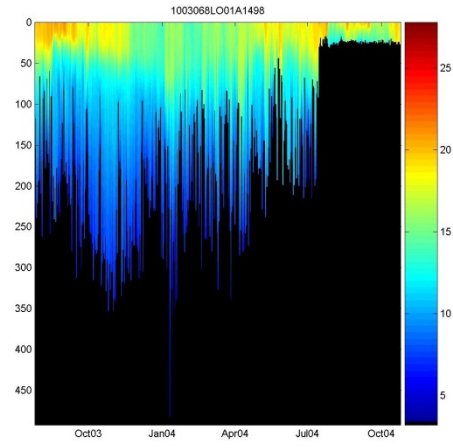
**Status- completed.**

NAVOCEANO – ABS data are sufficiently accurate and high quality to be useful element to fill existing observation gaps in under-sampled ocean regions (i.e. boundary currents, ocean fronts) and to improve operational ocean models (Frank Bub, NAVOCEANO personal comm.)

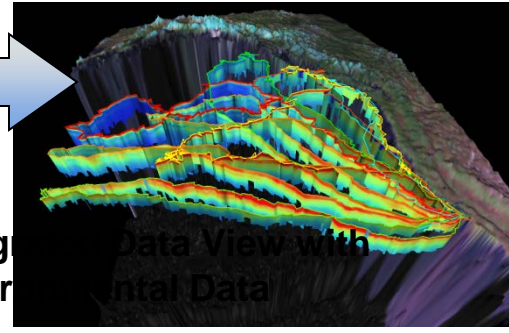


# TOPP Oceanographic Data Servers: Developing Capacity for Spatial-Temporal Habitat Mapping & Delivery of Ocean Observation Data

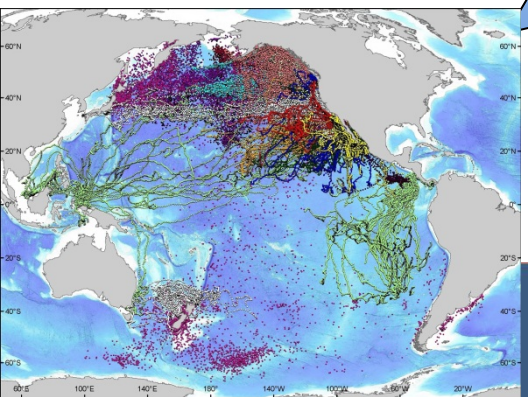
## Animal Sampled Environment



## Animal Oceanography



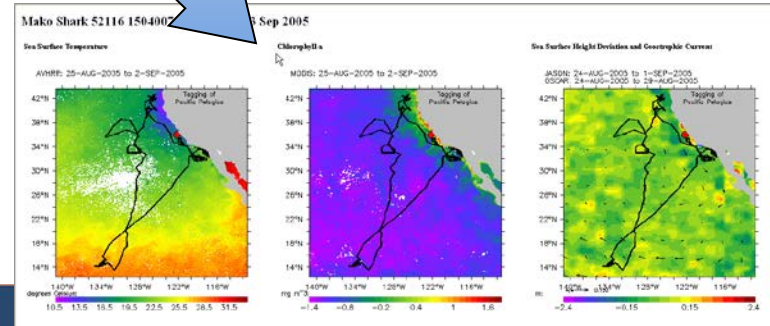
## Predator Movement



## Time Series Data Meta Data

Date	Time	Location	...
2003-10-01	00:00	155.00N, 155.00W	...
2003-10-01	01:00	155.00N, 155.00W	...
2003-10-01	02:00	155.00N, 155.00W	...

## Integ Envir

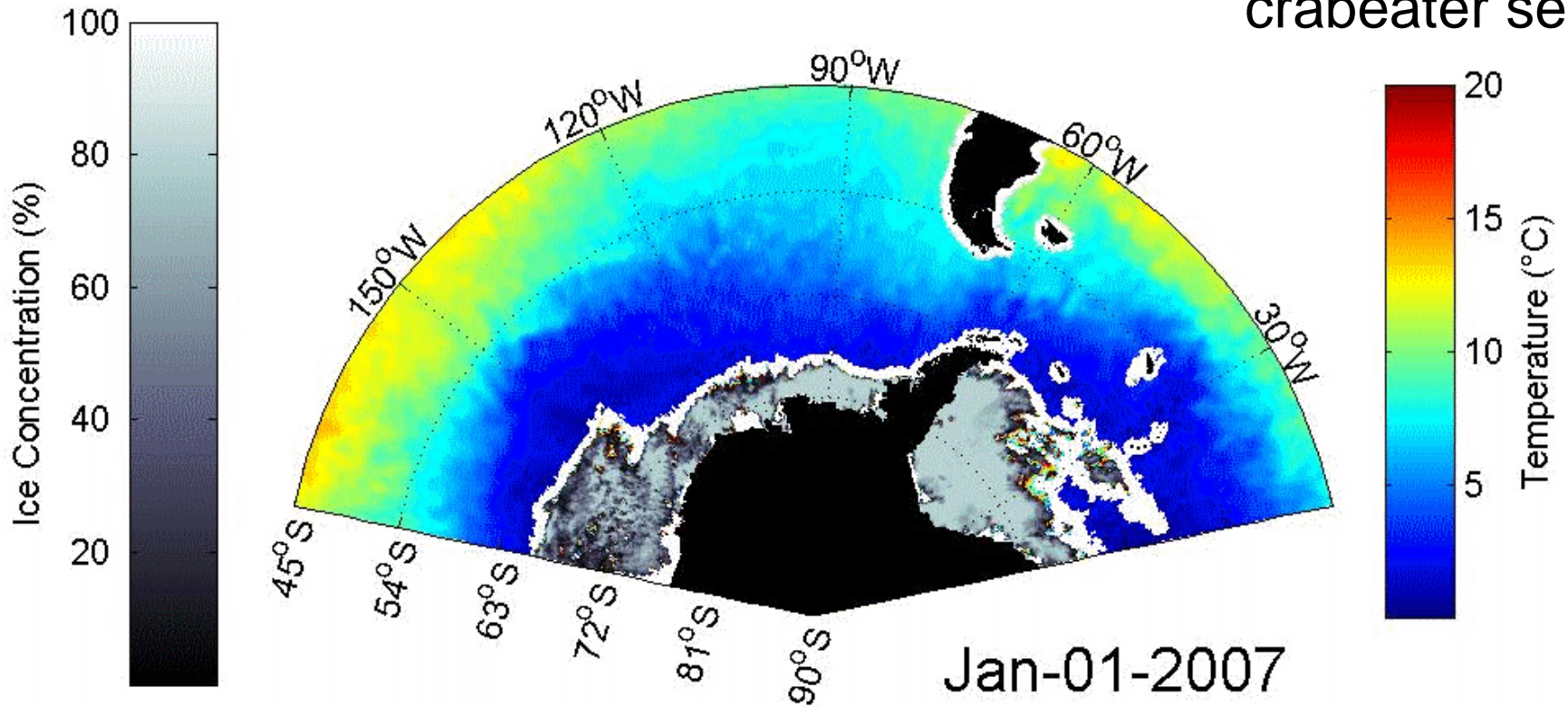


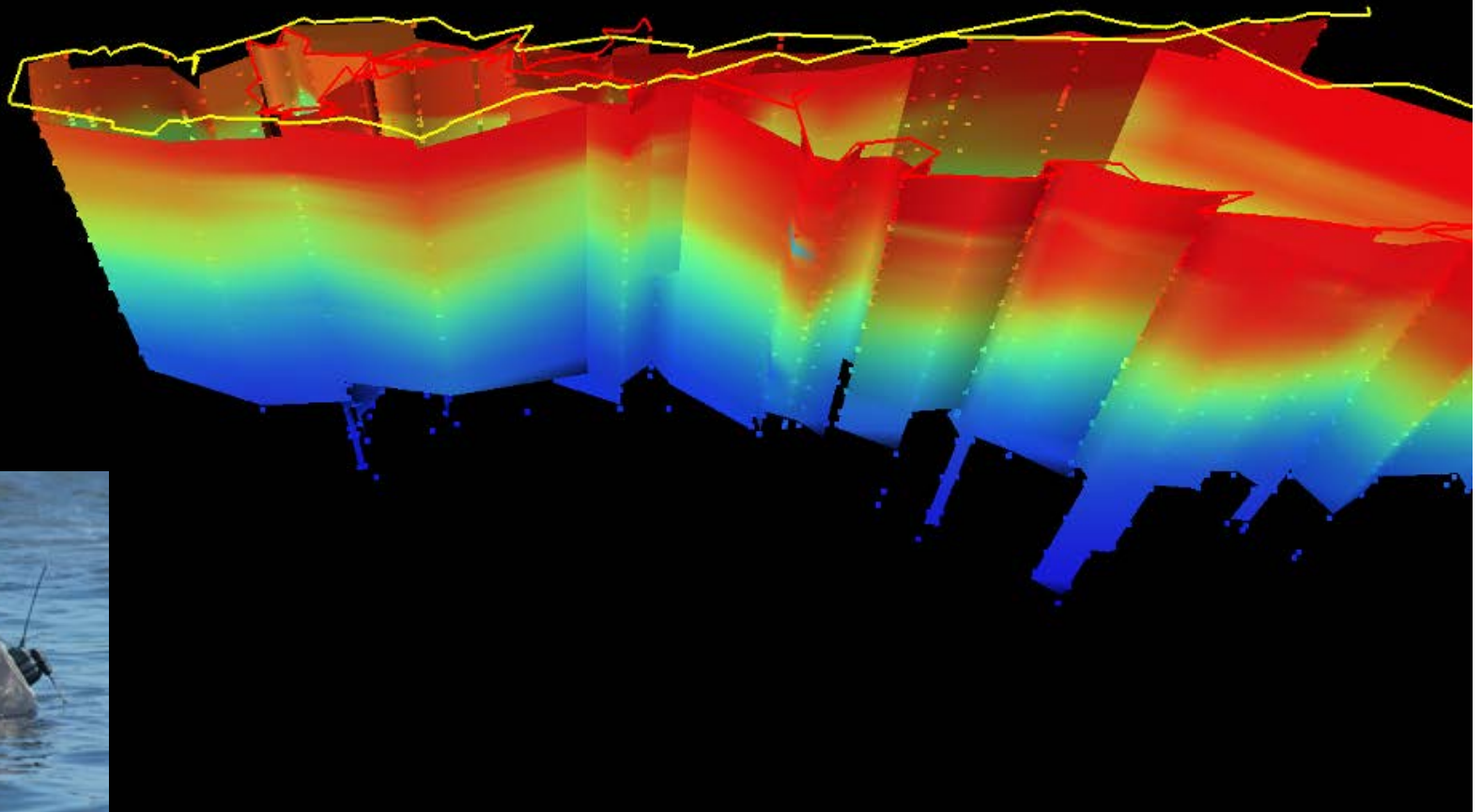


elephant seal

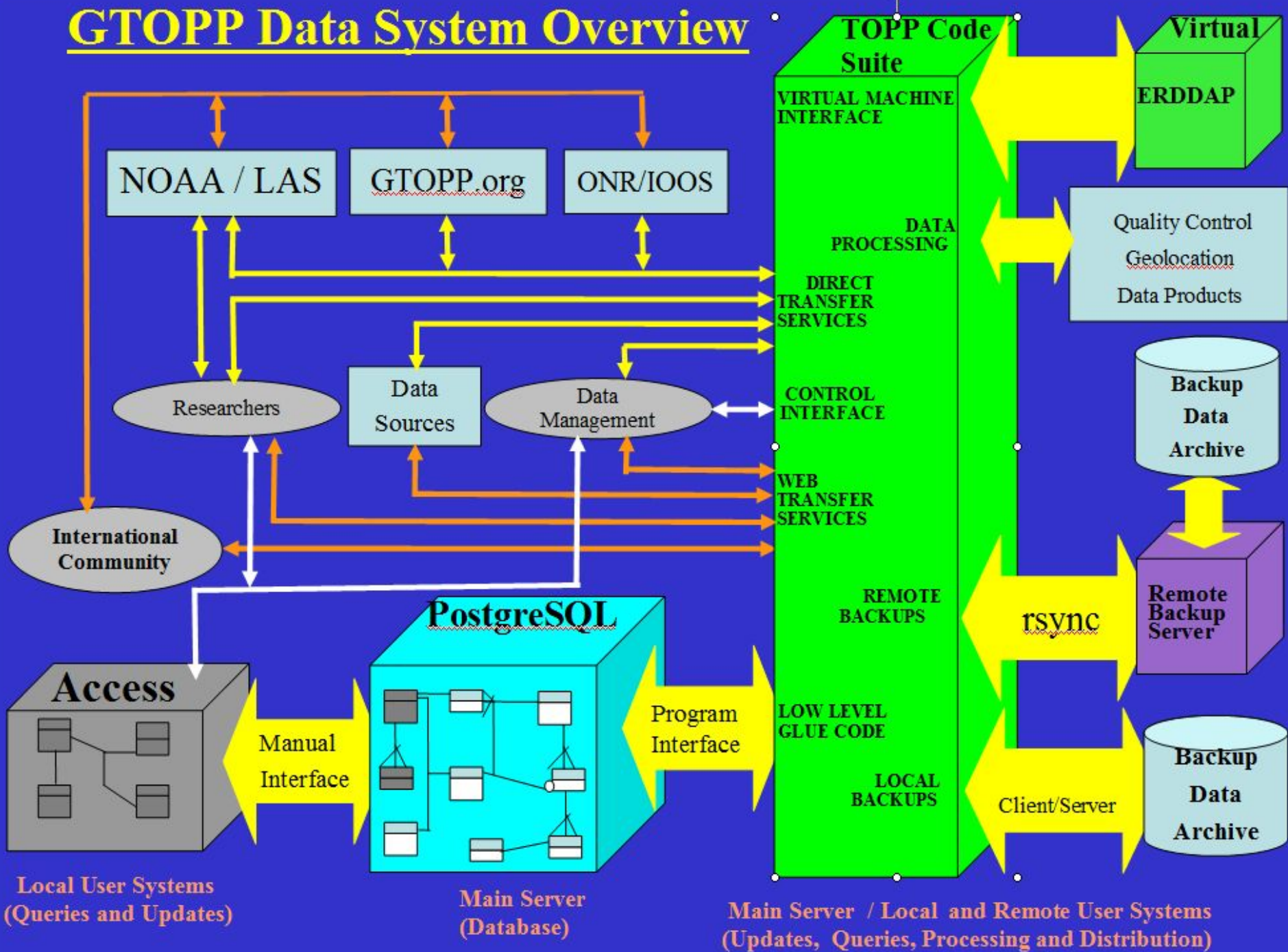


crabeater seal





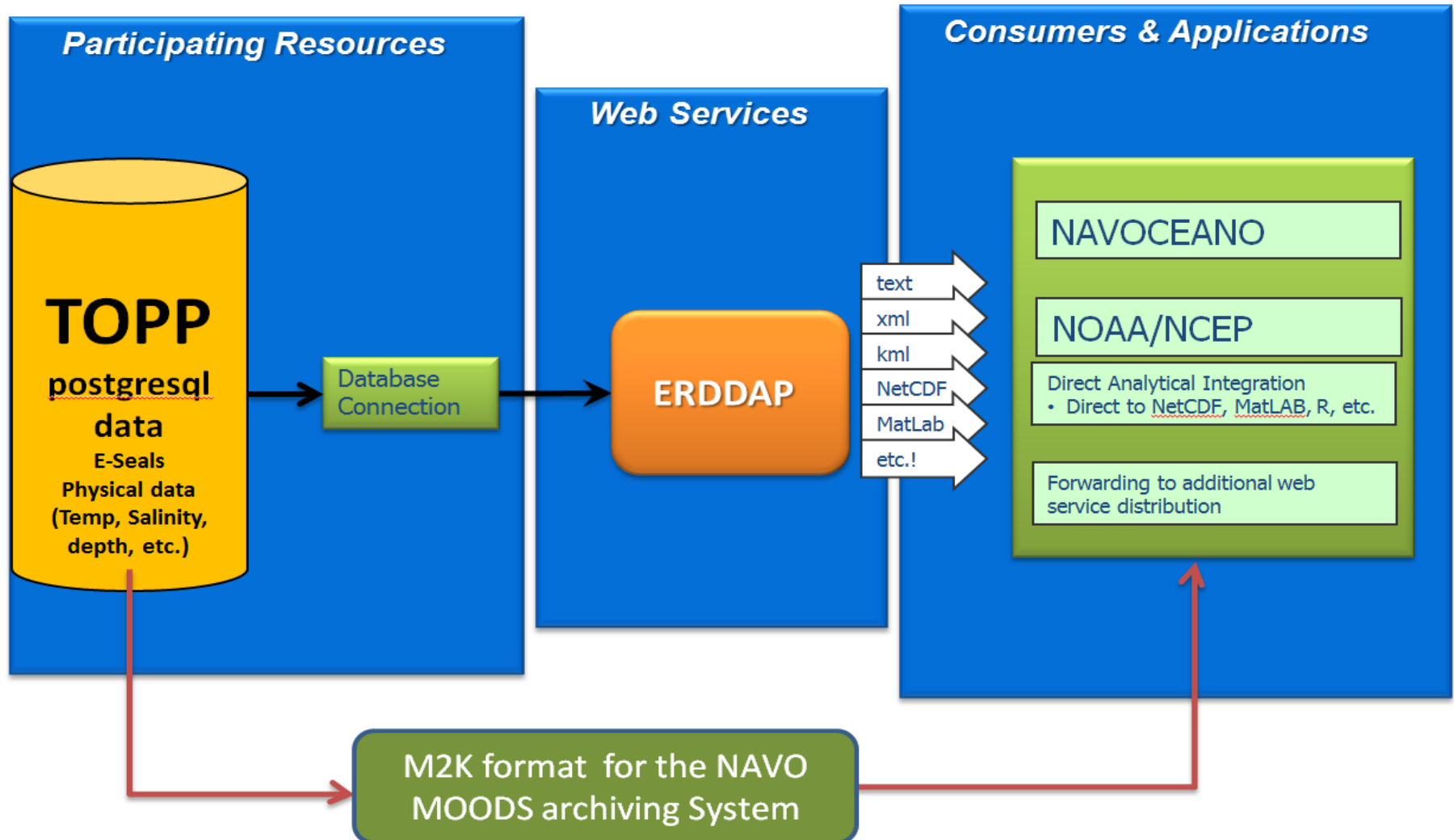
# GTOPP Data System Overview



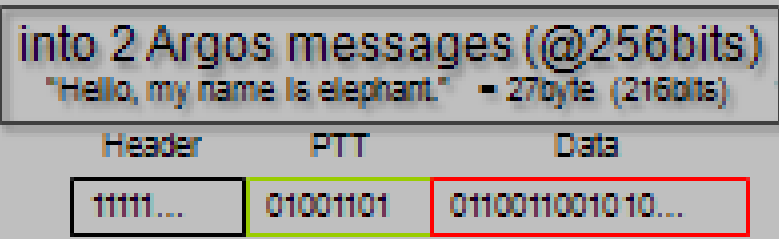
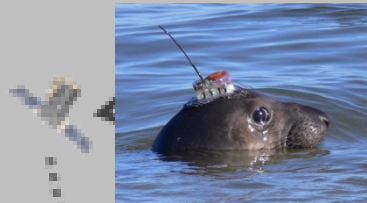
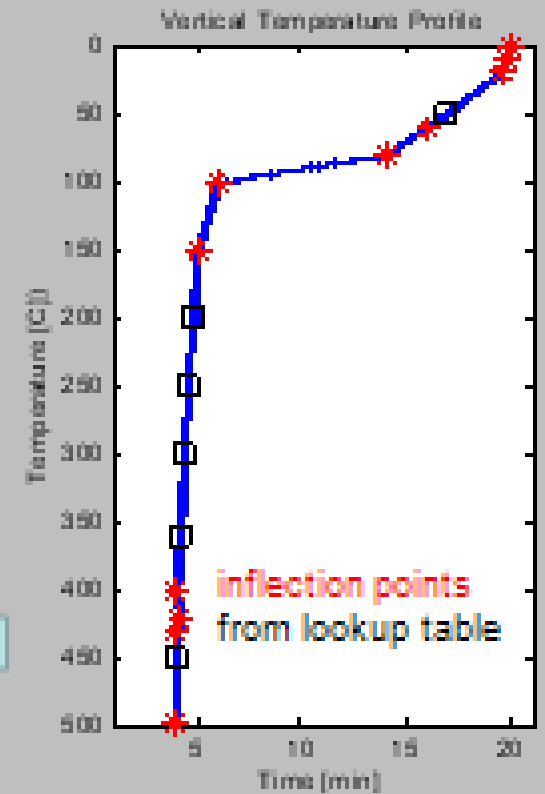
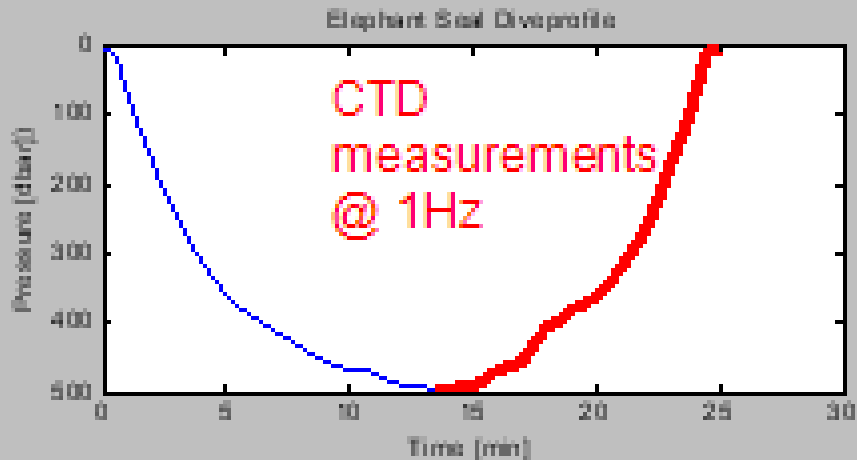
# TOPP Historical Data

<http://dataxfer.stanford.edu:8080/erddap/index.html>

**System Design Diagram**  
**to improve access to TOPP Animal Borne Sensors Physical data**



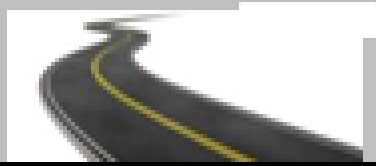
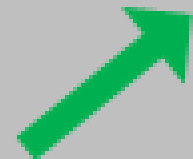




DMQC column in archive  
 (including QC flags and metadata, e.g offset)



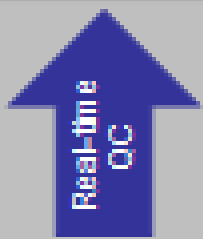
user  
 US NAVY  
 NAVOCEANO



Global Telecommunication System  
 GTS



Rawdata





## – Project 2:

With the Pacific Ocean Shelf Tracking project (POST), NANOOS and other RAs, OTN etc. To reconcile data content and access standards +Descriptive Metadata from acoustic sensors “tags” placed in endangered and threatened fish species such as salmon and green sturgeon.

**Status: We are in the system concept/design phase.**

# EXTENT OF THE POST ACOUSTIC RECEIVER ARRAY

DEPTH  
(METERS)



RECEIVER LINES



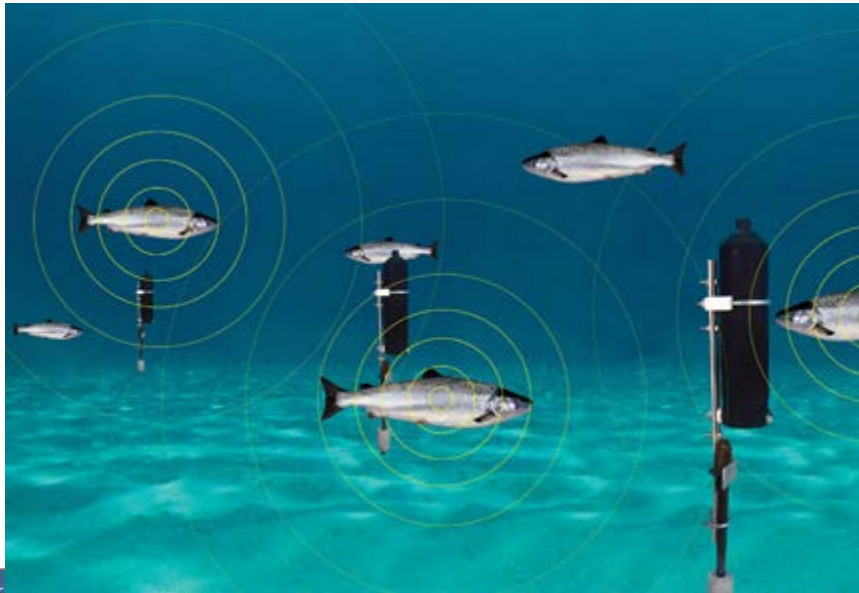
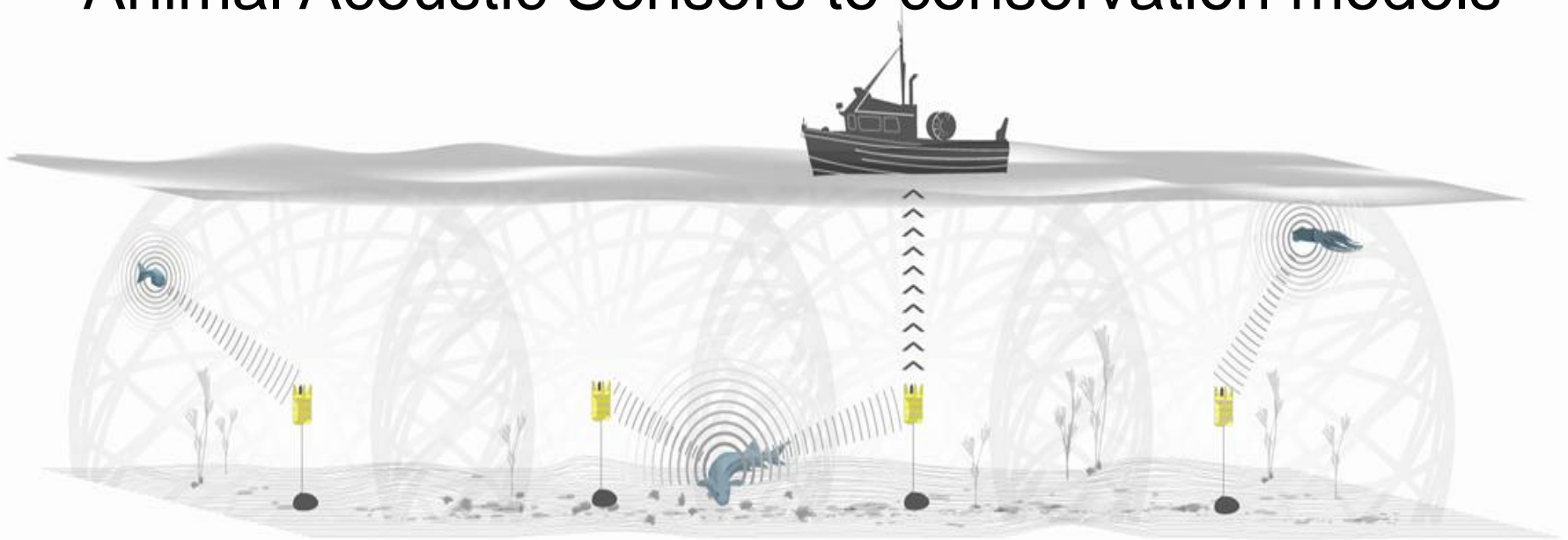
MAJOR RIVERS



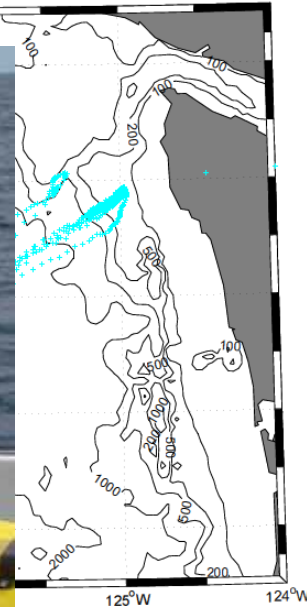
*POST is public,  
non-profit,  
community-supported*

- *465 receivers*
- *10 million detections*
- *16,000 tags*
- *18 species*

# Animal Acoustic Sensors to conservation models



# Seagliders: for testing, targeted searches, investigating temporary phenomena



# Bioprobes

Ecosystem-based Management



- 3) I've been to . . .
- 4) It was cold...
- 5) I met...
- 6) I ate...

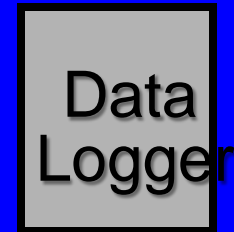
## Fully Integrated Tags

- Send/Receive POST codes
- Geolocation
- Physics/Chemistry
- Acoustic download

1) I'm here . . .



2) Where have you been?



300 m



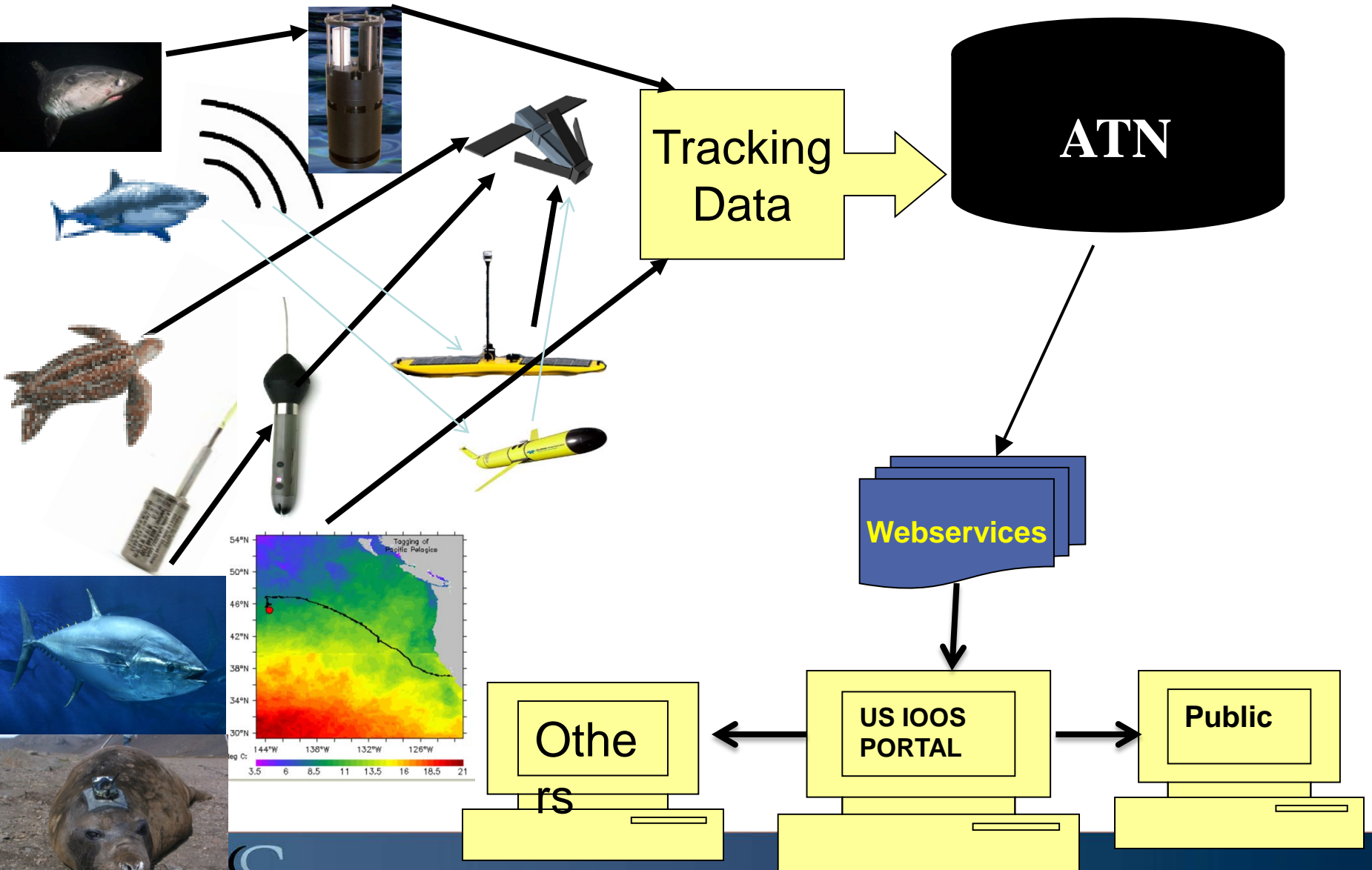
EMILIO PPT

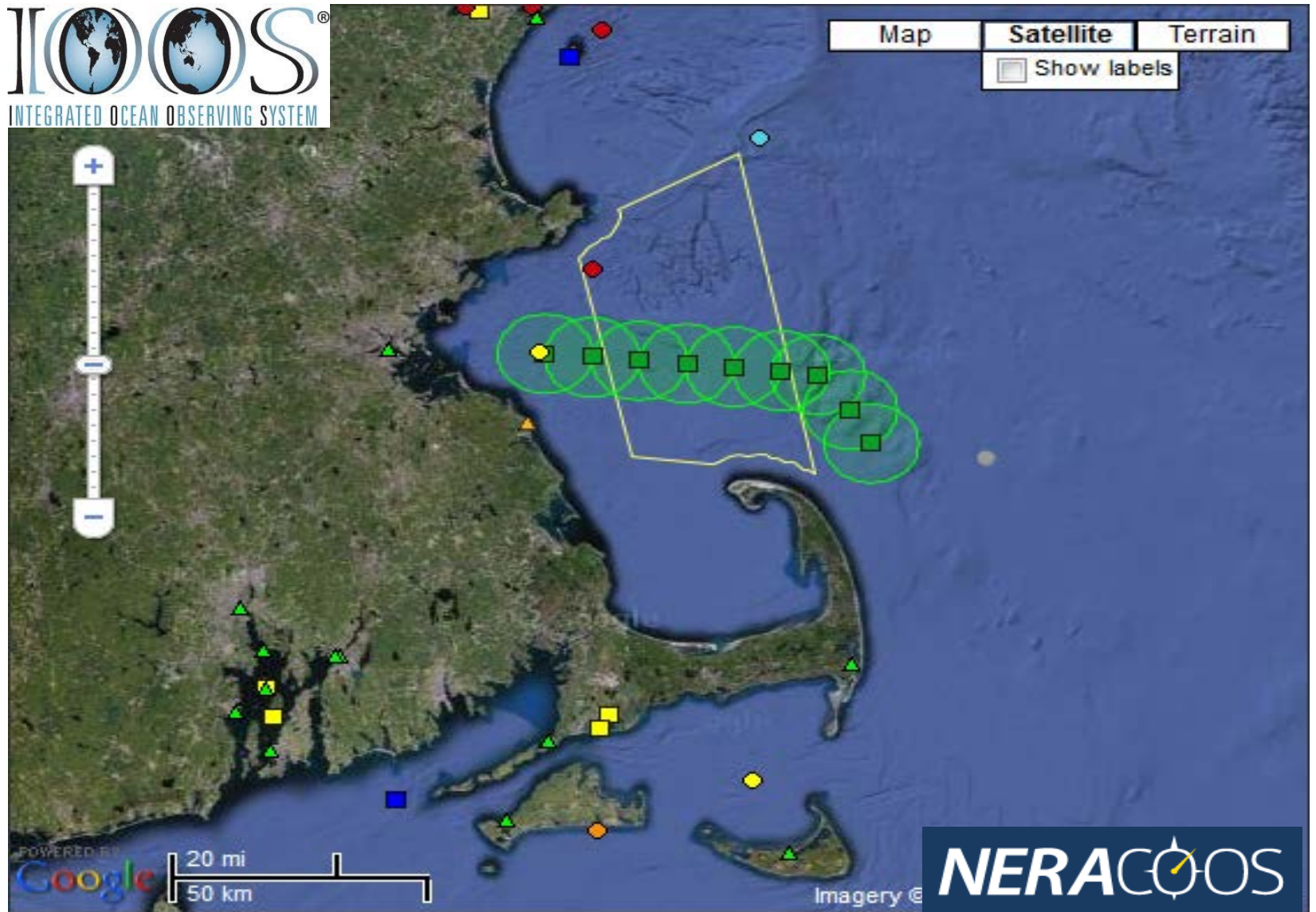
# FY13?

- Expand engagement with Animal Telemetry Community potential activities
  - Establish National Data Server or Data Assembly Center for US ATN data
  - Improve Sharing capability.
  - Make animal telemetry data and products available to community
- Implement ICES stds for Fisheries Hydro-acoustic in one or two NOAA acoustic datasets- potential regions NERACOOS, NANOOS, SECOORA and GCOOS-RA
- **We may define new activities after IOOS Summit**



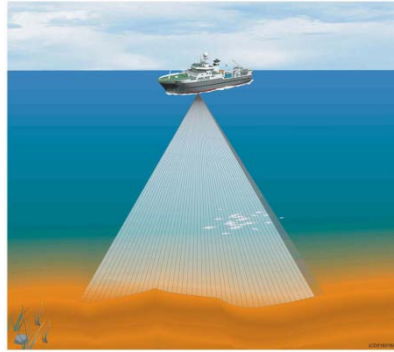
# US IOOS ATN Data Management System Must Ingest the Diverse Animal Tag Platforms



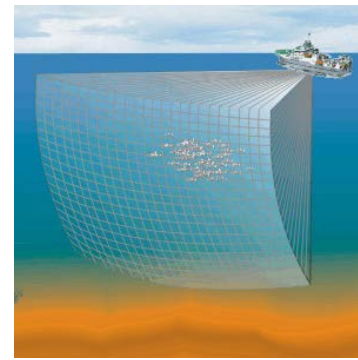


# NOAA Fisheries Independent Surveys- using Hydro-acoustic technology

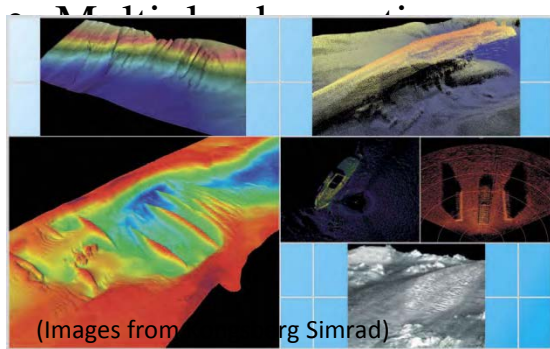
- A single observation can range from cm's to km's in space and tenths to tens of seconds in time



**Kongsberg Simrad ME70 Multibeam Sonar**  
What makes this novel is the capability to simultaneously collect water column and seafloor calibrated backscatter

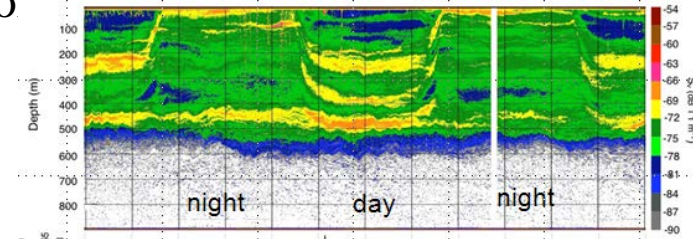
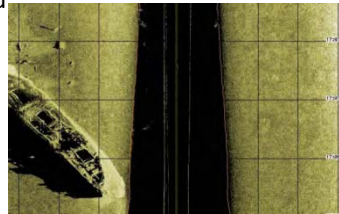


**Kongsberg Simrad MS70 Multibeam Sonar**



range over 1000's of km and months

Multibeam sonar and side scan images of the sea bed and objects on the sea bed



(Horne et al. 2010) 48-hour echogram with time on the x-axis, left to right vs. depth (900 m).

- Acoustic measurements are made from many platforms
  - Vessels, AUVs, ROVs, b



NOAA FRV Delaware II



Retractable keel with EK60 transducers on NOAA FSV HB Bigelow



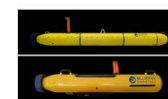
(Image from H. Singh, WHOI)



(Image from NOAA SWFSC)



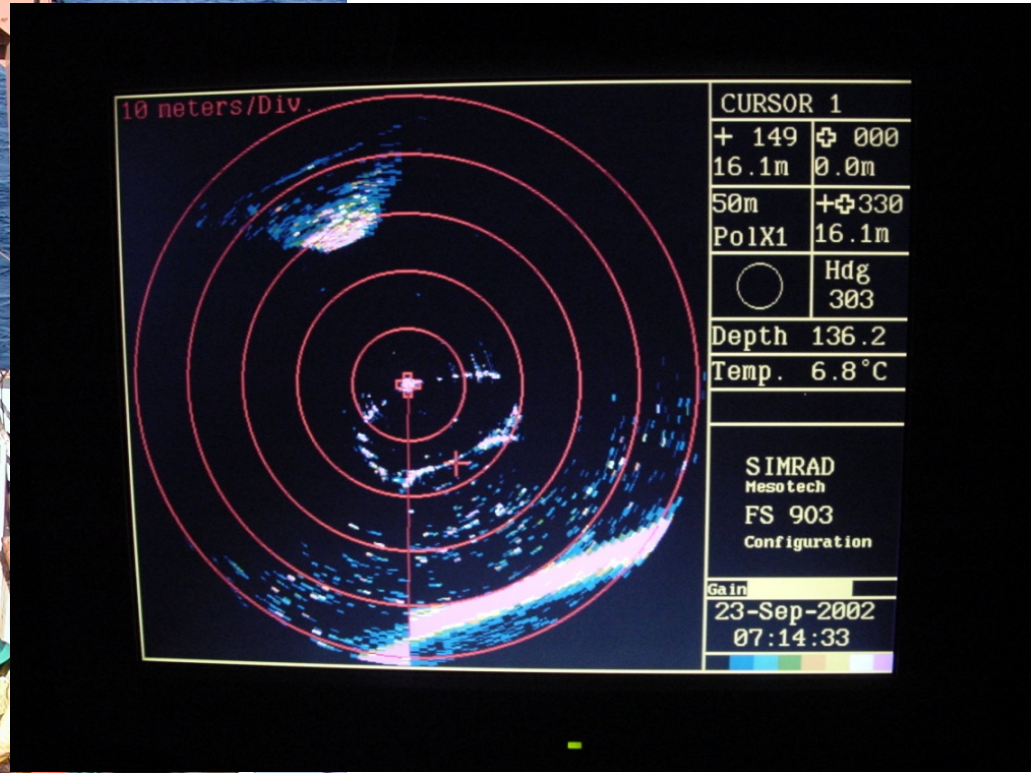
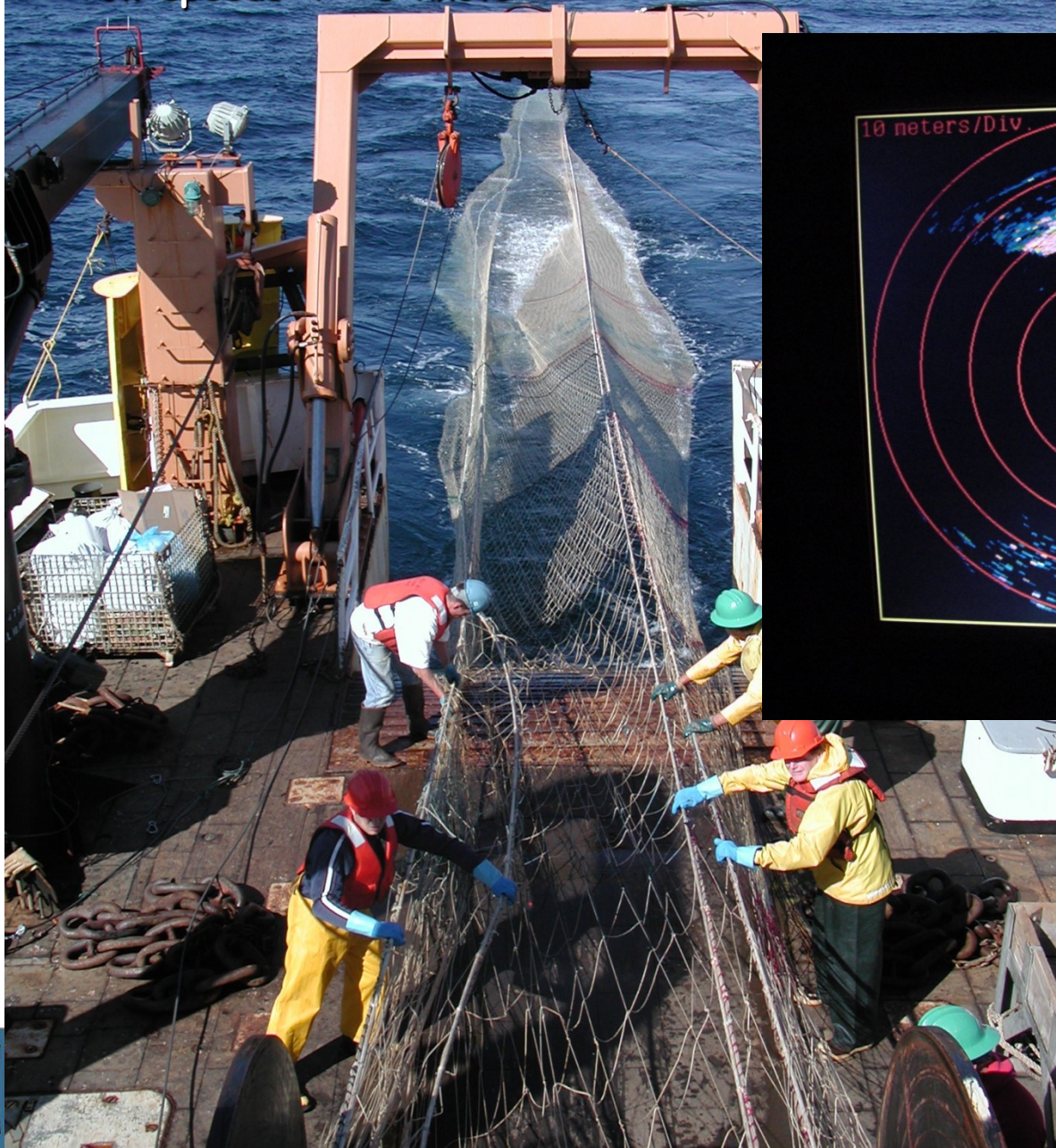
(Image from Kongsberg Simrad)



(Image from Bluefin Robotics)

# Verification and Biological Sampling: Midwater Trawl

Tow speeds 4 - 5 knots



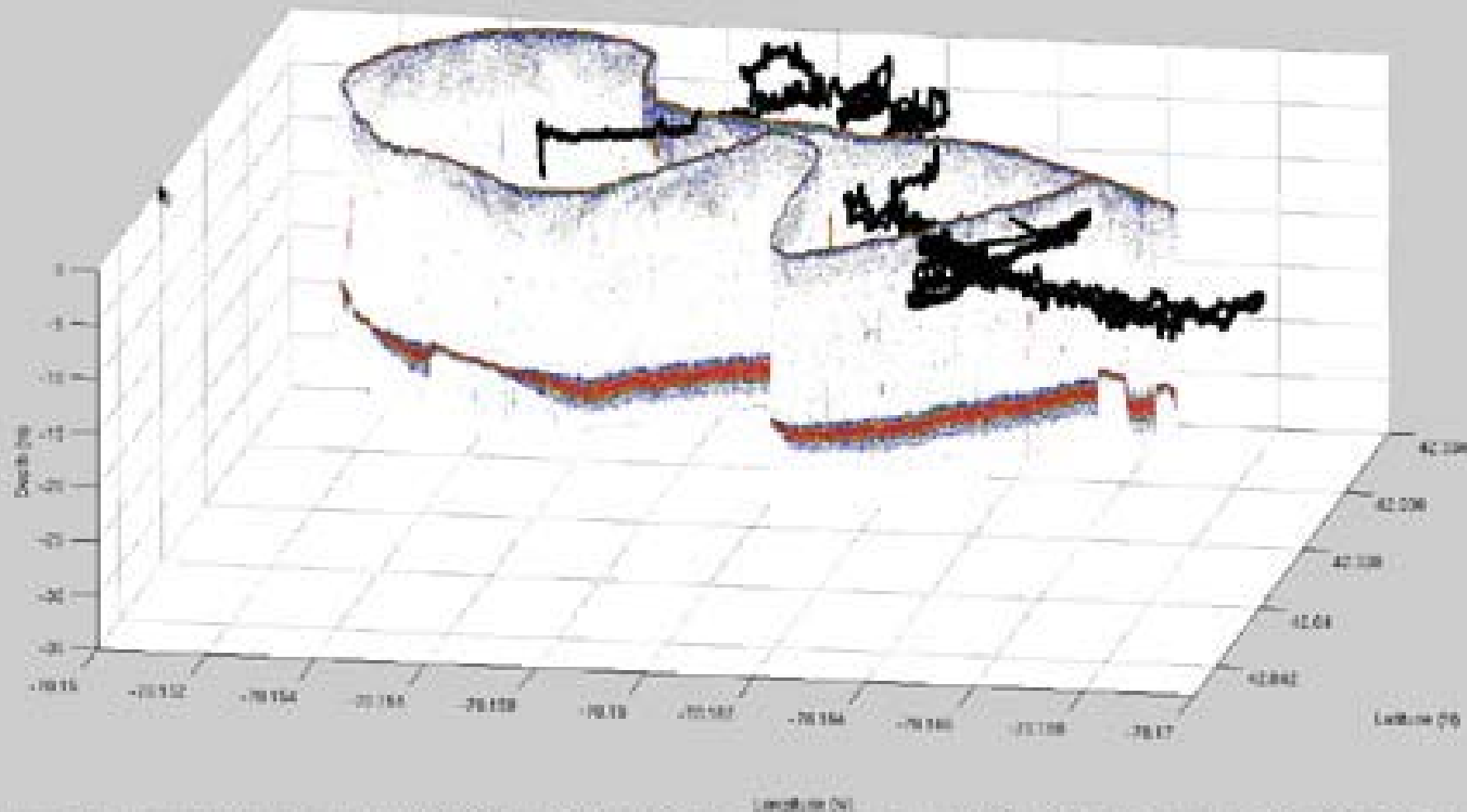


Figure 1. Acoustic echogram (710 kHz) from Cape Cod Bay overlaid with tagged path of a right whale (black line). The right whale spent nearly all its time within the upper 5 m of the water column. Blue areas in the echogram represent patches of calanoid copepods with numerical densities  $> 1000$  animals /  $m^3$  while the red line is the seafloor. The track of the tagged right whale is pseudo-georeferenced so the horizontal uncertainty is on the order of 50 m.

# Scientific Operations during NEFSC Fisheries Acoustic Surveys

FRV Delaware II

FS903 Trawl Monitoring System  
ITI Trawl Monitoring System

Trawling

Underwater  
video

CTD

Plankton Sampling

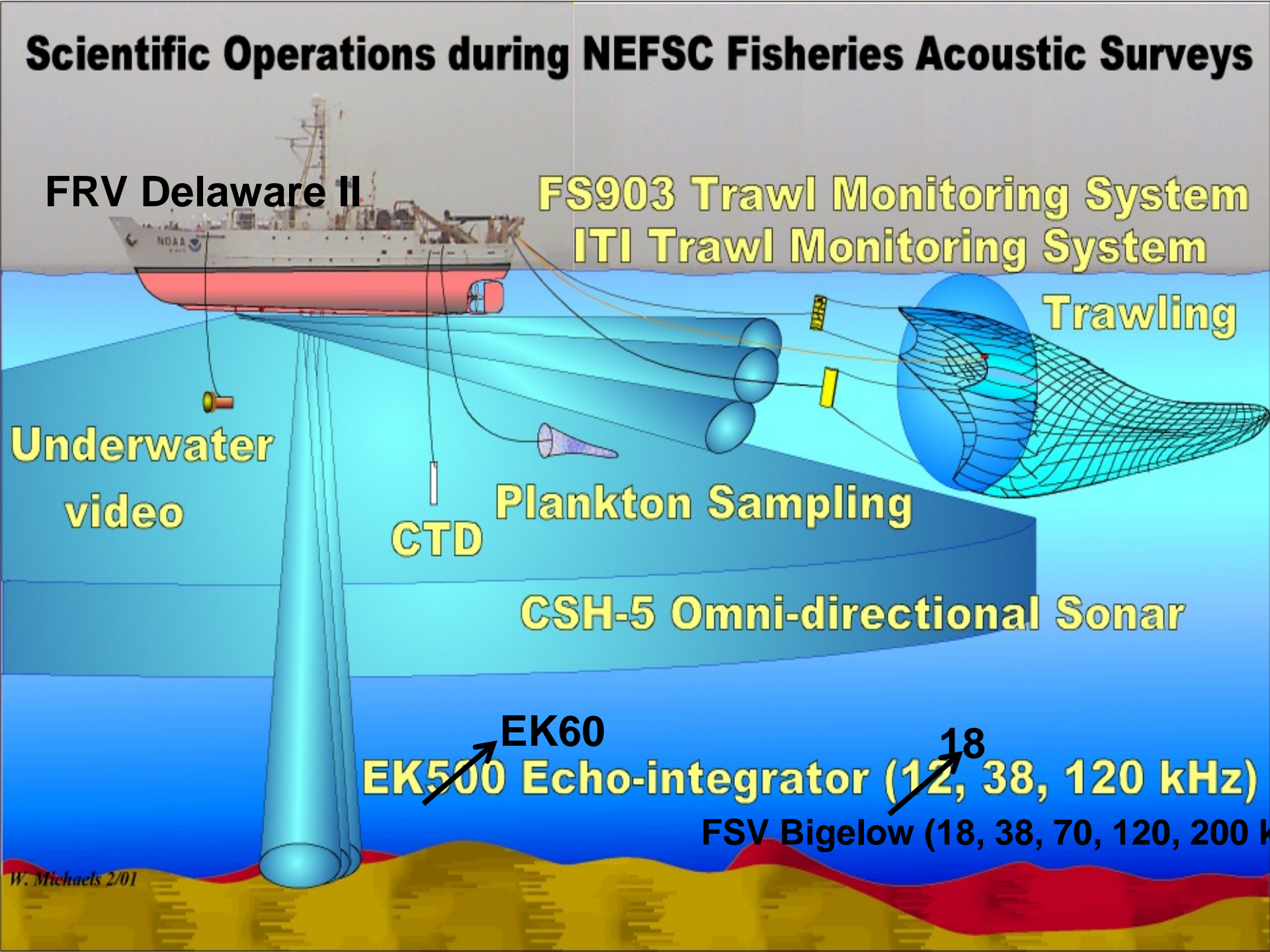
CSH-5 Omni-directional Sonar

EK60

EK500 Echo-integrator (12, 38, 120 kHz)

FSV Bigelow (18, 38, 70, 120, 200 kHz)

18



Which data model fit this type of data?

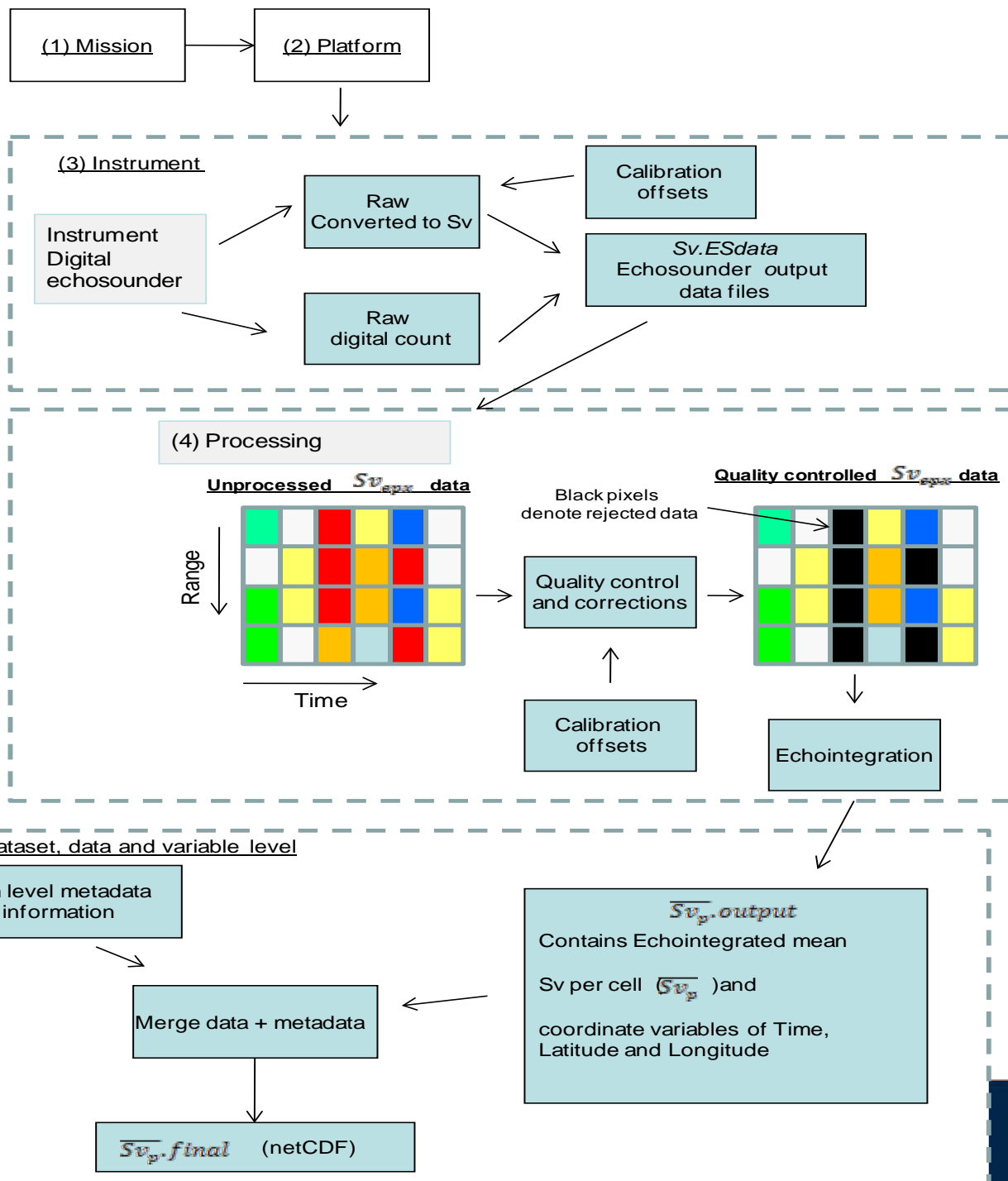
Sds developed within ICES *International Council for the Exploration of the Sea* community

NetCDF (Network Common Data Form) may be a better format for storage of final acoustic output data.

**Sv** (Acoustic Volume Backscatter) final data file is well suited to array-oriented acoustic data.

Further, metadata contained in the NetCDF files can be automatically harvested to facilitate ready discovery of the data.

# Data flow schema for production of mean Sv data with associated metadata levels





## Example. Data variable for volume backscatter and ancillary data quality variable

Sv

Size: 120x2154

Dimensions: DEPTH,TIME

Datatype: double

Attributes:

name = 'Sv'

long\_name = 'mean\_volume\_backscatter\_coefficient'

units = 'm-1'

\_FillValue = 1e+04

valid\_min = 0

valid\_max = 1

ancillary\_variables = 'Sv\_quality\_control'

Sv\_quality\_control

Size: 120x2154

Dimensions: DEPTH,TIME

Datatype: int8

Attributes:

long\_name = 'quality flag for mean\_volume\_backscatter\_coefficient'

quality\_control\_set = 1

quality\_control\_conventions = 'IOOS standard flags'

\_FillValue = 99

valid\_min = 0

valid\_max = 9

flag\_values = [0 1 2 3 4 5 6 7 8 9]

flag\_meanings = 'No\_QC\_performed, Good\_data, Probably\_good\_data,

Bad\_data\_that\_are\_potentially\_correctable, Bad\_data Value\_changed, Not\_used, Not\_used,

Not\_used, Missing\_value