Role of NASA's SeaBASS repository for the legacy of the EXPORTS field biogeochemical measurements

*Inia M. Soto Ramos[1,2], Christopher W. Proctor[1,3], **Ivona Cetinic**[1,2], Susanne Elizabeth Craig[1,2]*

NASA Goddard Space Flight Center[1], USRA[2], Science Systems and Applications, Inc.[3]

Plankton, Aerosol, Cloud, ocean Ecosystem

What is EXPORTS?

- EXport Processes in the Ocean from Remote Sensing (EXPORTS) is a large-scale NASA-led and NSF co-funded field campaign
- Aims to collect extensive dataset to develop and test numerical predictive and satellite-data diagnostic models to understand the export, fate, and carbon cycle impacts of global ocean net primary production

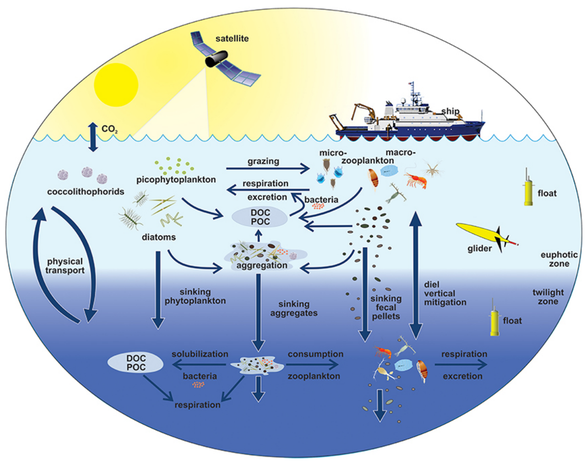
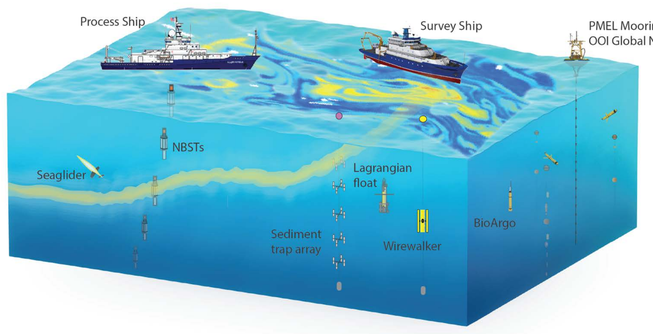


Diagram from Siegel et al., 2016. https://doi.org/10.3389/fmars.2016.00022

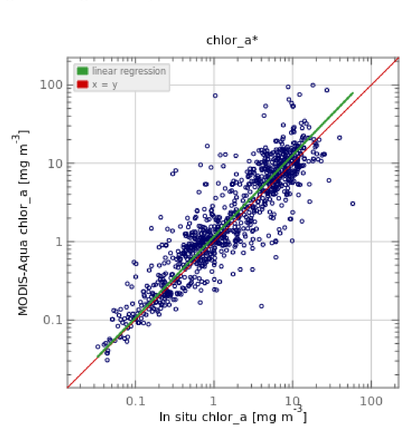
EXPORTS Assets and Collaborations

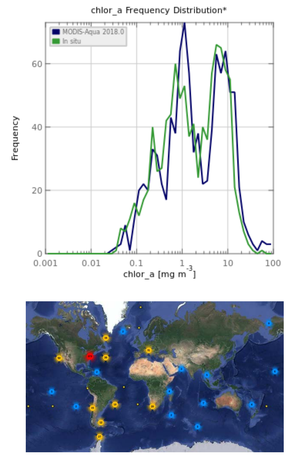
- Two field campaigns: North Pacific (2018)[1], North Atlantic (2020)[2]
- A process ship (R/V Revelle[1] and R/V Atlantis[2]) measures rates and time series of stocks following a Lagrangian float
- A survey ship (R/V Sally Ride[1] and RRS Discovery[2]) provides spatial information of ocean optics & biogeochemistry
- Gliders and other autonomous instruments
- Bio-Argo floats and moorings
- Sediment traps



What is SeaBASS?

- The SeaWiFS Bio-optical Archive and Storage System (SeaBASS) is NASA's repository for in situ oceanographic datasets
- Its data and services support Ocean Color satellite missions with the ongoing ground-truth comparisons needed for the validation of global geophysical measurements





EXPORTS Data Types

Data type	Examples of the data generated during EXPORTS
Optics	In-water and above-water radiometry, hyperspectral particle absorption and attenuation coefficient, Volume Scattering Function (VSF), particle size distribution (PSD)
Particle characterization	Enumeration, sizing and classification of particles (e.g., taxonomy and functional groups, biomass) using microscopy, flow cytometry, mocness/zooscan, UVP, LISST, and others!
Stocks	Biogeochemical measurements: DOC, DIC, POC, PIC, PN, DON, nutrients, HPLC, pH, alkalinity, trace element data, 234Th
Proxies	Proxies are synthesized observations created by mapping one variable onto another (e.g., POC derived from optical backscatter)
Context	Include all physical / sensor measurements from the EXPORTS assets (i.e., ships, autonomous vehicles, floats, moorings, and remote sensing observations) such as temperature, salinity, Chl fluorescence, ocean currents, dissolved oxygen and acoustics
Rates	Primary productivity (GPP, NPP, NCP), bacterial and community respiration, zooplankton respiration using Electron Transport System (ETS), bacterial production, grazing, aggregation and sinking rates and nutrient uptake rates
EXPORTS Pathways	Particle fluxes, distribution of sinking particles, respiration rates of sinking particles, aggregate dynamics and coagulation modeling, zooplankton migratory distributions

SeaBASS Data Management

- SeaBASS provides data in ASCII format
- A Digital Object Identifier (DOI) is registered for each experiment
- Data are reviewed for general metadata and compliance before being archive
- Datasets used for algorithm development and validation go through additional intense QA/QC
- Standardized machine-readable keywords enable automated processing.
- Matlab and Python codes to read the data files are available at the SeaBASS website.
- /fields and /units identify every column in the data block and are also standardized, such as:

Field name	Units	Description
Chl_a	mg/m^3	HPLC Chlorophyll a
POC	mg/m^3	Particulate organic carbon

Example SeaBASS File

```
/begin_header
/identifier_product_doi=10.5067/SeaBASS/EXPORTS/DATA001
/received=20190920
/investigators=Dennis_Hansell,Chelsi_Lopez
/affiliations=University_of_Miami
/contact=dhansell@miami.edu
/experiment=EXPORTS
/cruise=EXPORTSNP
/station=Ocean_Station_Papa
/data_file_name=EXPORTS-EXPORTSNP_DOC_20180815_R1
/documents=Masterfile_SallyRide_Bottlefile_R1.xlsx
/data_type=bottle
/data_status=final
/start_date=20180814
/end_date=20180909
/start_time=08:02:00[GMT]
/end_time=07:06:00[GMT]
/water_depth=4200
/missing=-999
/delimiter=comma
/instrument_manufacturer=Shimadzu
/instrument_model=TOC_L
/fields=R2R_Event,DOC_L,DOC_L_quality,TDN,TDN_quality,depth
/units=none,umol/L,none,umol/L,none,m
!COMMENTS
!Quality flags ...
/end_header
SR1812-SE-20180814.0744.001,47.29,1,44.68,1,502.94
SR1812-SE-20180814.0744.001,-999,-999,-999,-999,502.21
```

Metadata headers: required and optional headers contain key information about the dataset.

Fields and units are predefined by SeaBASS.

Comments are optional. User can provide additional information for the data.

Data separated by comma, space, or tab. ASCII format!

SeaBASS & EXPORTS Challenges!

1. **Data diversity:** Over 300 new data parameters never used previously in SeaBASS!

Photo credits: Colleen Durkin

Solution: An ID system was created to give users flexibility in naming new parameters while minimizing the overhead of creating new fields. For example: A dataset with 18 types of particle fluxes: Flux_particles_1id, flux_particles_2id ...flux_particles_18id. IDs are defined in a metadata header.

2. **Taxonomic data and images:** EXPORTS will generate Terabytes of taxonomic data using optical and imaging systems.

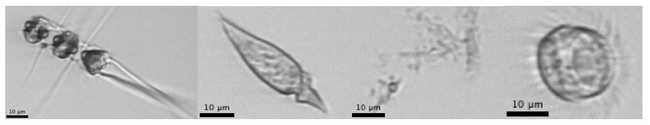


Photo credits: Nils Haentjens, Emmanuel Boss, Lee Karp Boss

Solution: Data is distributed by ECOTAXA and long-term archived at SeaBASS. Imagery associated will be bundled in an associated tar file with the main dataset. Associated files can be downloaded at user-discretion and files can be linked per dataset or data point. Taxonomic identification will require a recognized ID system such as WoRMS AphiaID. For example:

Namespace_manual	Identification_manual	Biotic_group	Abundance
aphiaid	254453	Themisto_pacifica	1

3. **Protocols and collection methods:**

- New field names were developed to convey important details about underlying methods (e.g., NPP vs. rate_13C_uptake_bottle_24h.)
- New standardized checklists documenting the methods for collection and processing are now required for data compliance.

Achievements:

- A total of **2353 EXPORTS data files** available at SeaBASS!
- Available EXPORTS data from 95 different data types and 24 PI's.
- Data includes oceanographic, optical, and biogeochemical data, phytoplankton and zooplankton taxonomy, primary productivity, marine snow catcher, wirewalker and sediment traps!
- More data from NP is expected in the next few months and more data will be collected in the North Atlantic (April-May 2020)

Summary:

- EXPORTS challenged SeaBASS data handling and processing system, and positively impacted data management and future improvements to the repository
- EXPORTS is providing an extensive dataset that will not only satisfy the goals of the project but serve as validation data for future ocean color missions such as PACE.
- SeaBASS is continuing to collaborate with BCO-DMO

Contact Information:

SeaBASS Questions: Inia Soto Ramos (Inia.M.SotoRamos@nasa.gov, Christopher W. Proctor christopher.w.proctor@nasa.gov

EXPORTS Questions: Ivona Cetinic ivona.cetinic@nasa.gov, David Siegel david.siegel@ucsb.edu

EXPORTS website: https://oceanexports.org/

SeaBASS website: https://seabass.gsfc.nasa.gov/