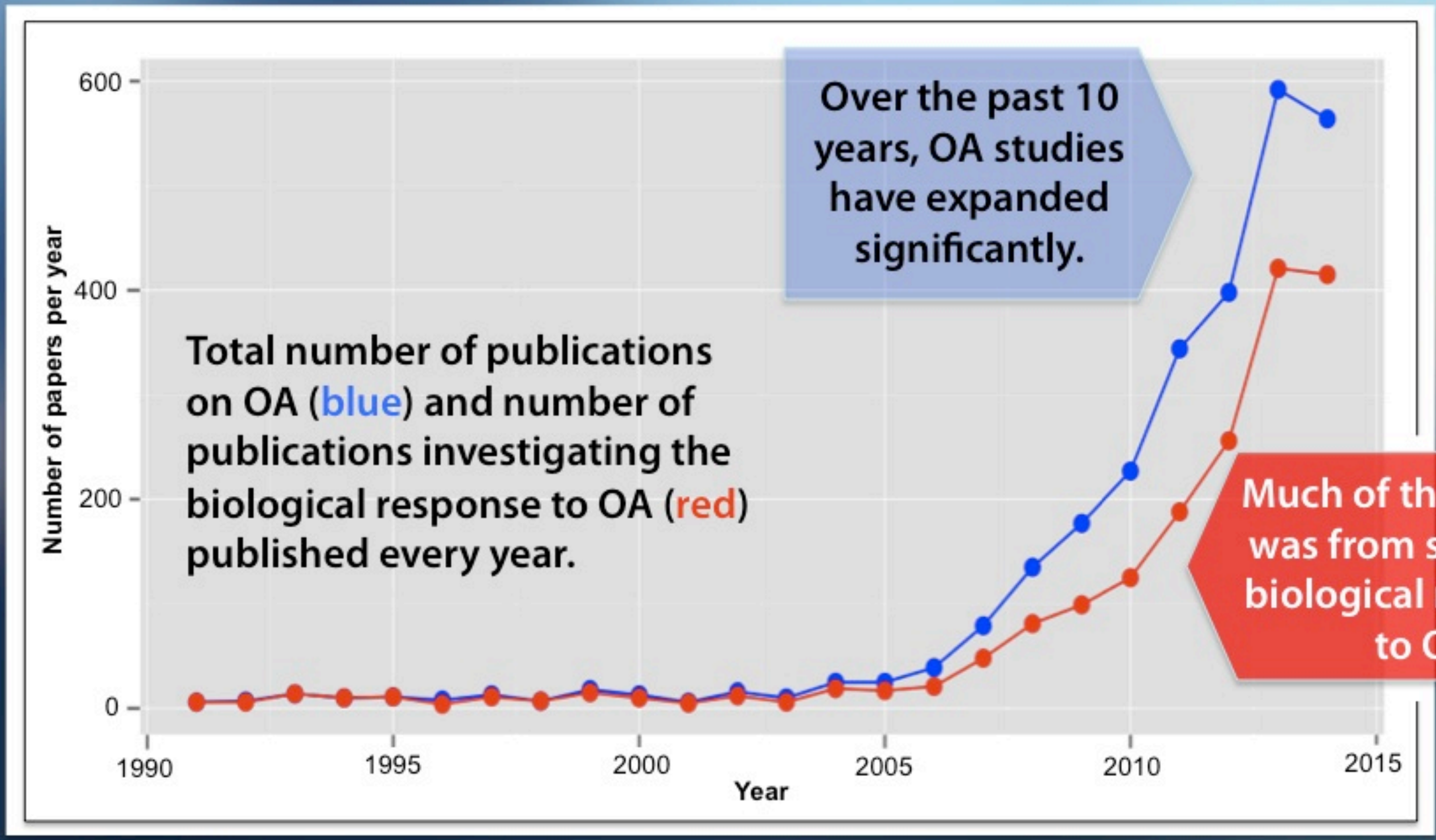
ASLO | LIMNOLOGY AND OCEANOGRAPHY e-Lectures

How to Document Ocean Acidification Data

Li-Qing Jiang
Krisa M. Arzayus
Jean-Pierre Gattuso
Hernan E. Garcia
Cynthia Chandler
Alex Kozyr
Yan Yang
Rob Thomas
Brian Beck
Tobias Spears

ASLO e-Lectures

The need for an OA metadata template



Gattuso & Hansson (2011)

There is an urgent need to develop an OA metadata template to facilitate archiving and access to these data.

Jiang, et. al

doi:10.1002/loe2.10004

What is metadata?

Meta- is a prefix that means "an underlying description."

Meta information of a car	
Year:	2015
Make:	Ford
Model:	Fusion
Color:	Silver
Engine:	2.0T
Transmission:	6-spd

Similarly, meta information of data, commonly called **metadata,** is a description of data, or **data about data**. It is defined as structured information that describes, explains, and locates a data set.

Jiang, et. al

doi:10.1002/loe2.10004

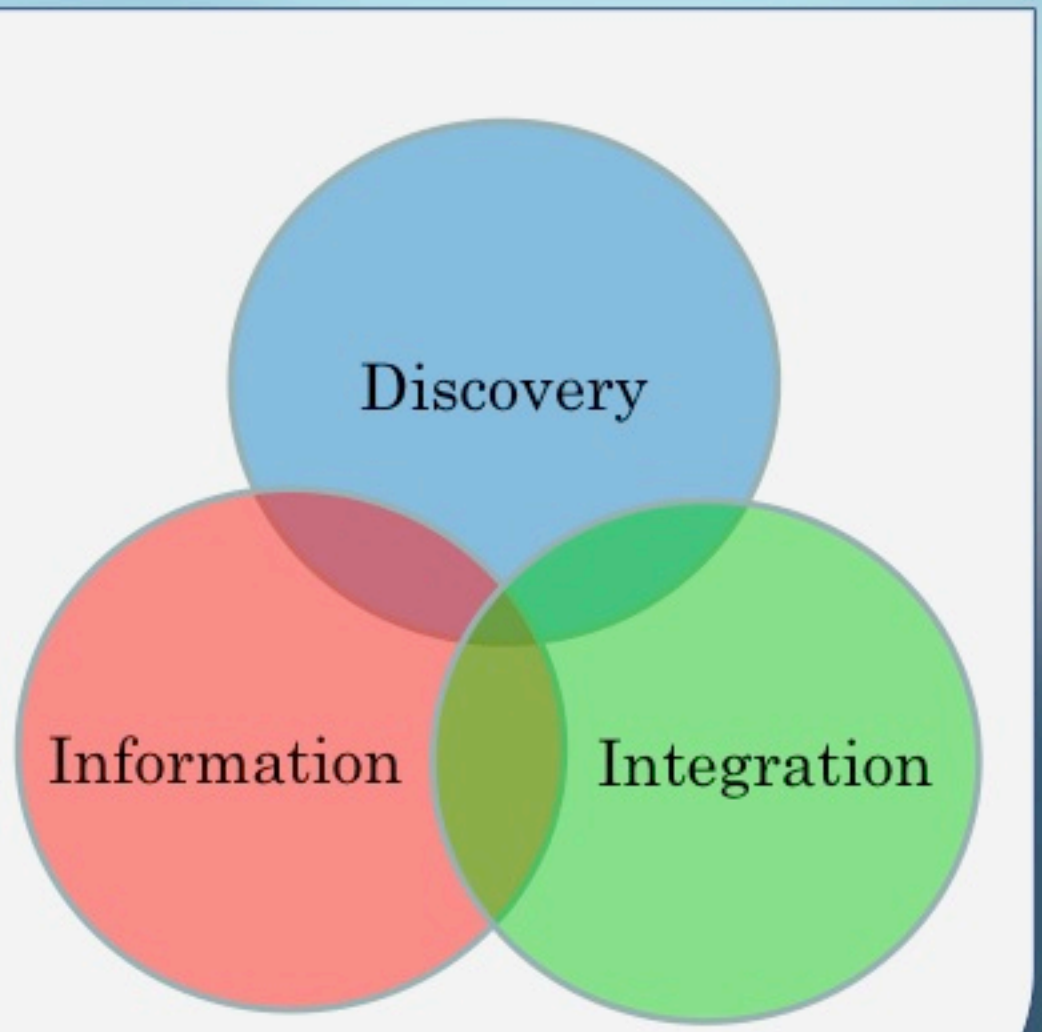
ASLO e-Lectures

Role of metadata in data management

Metadata is critical to **data discovery** by enabling the data to be found through relevant criteria.

Metadata helps to **document information** about the data sets in consistent and standard ways.

Metadata plays an important role in **facilitating data interoperability**, and **integrating legacy data**.



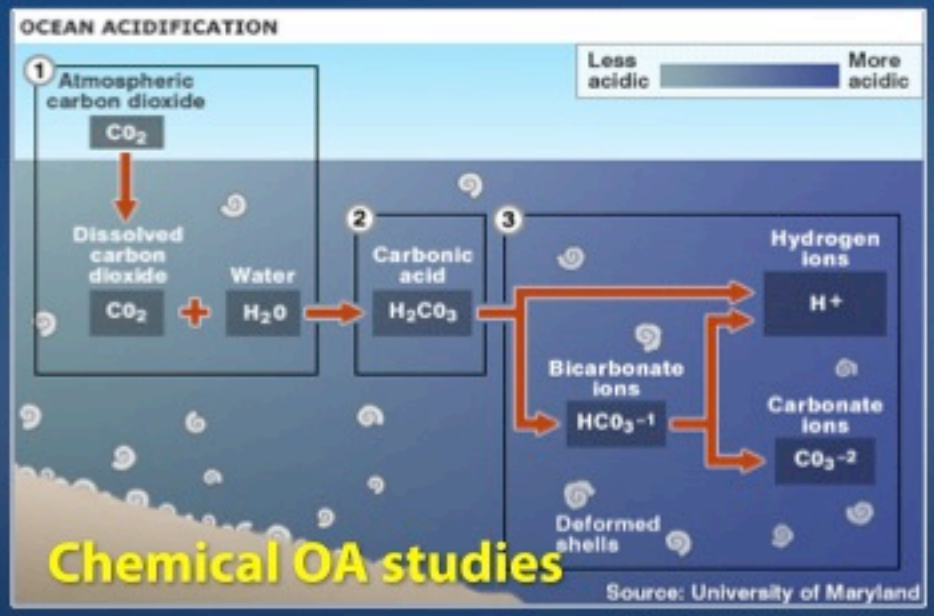
Jiang, et. al

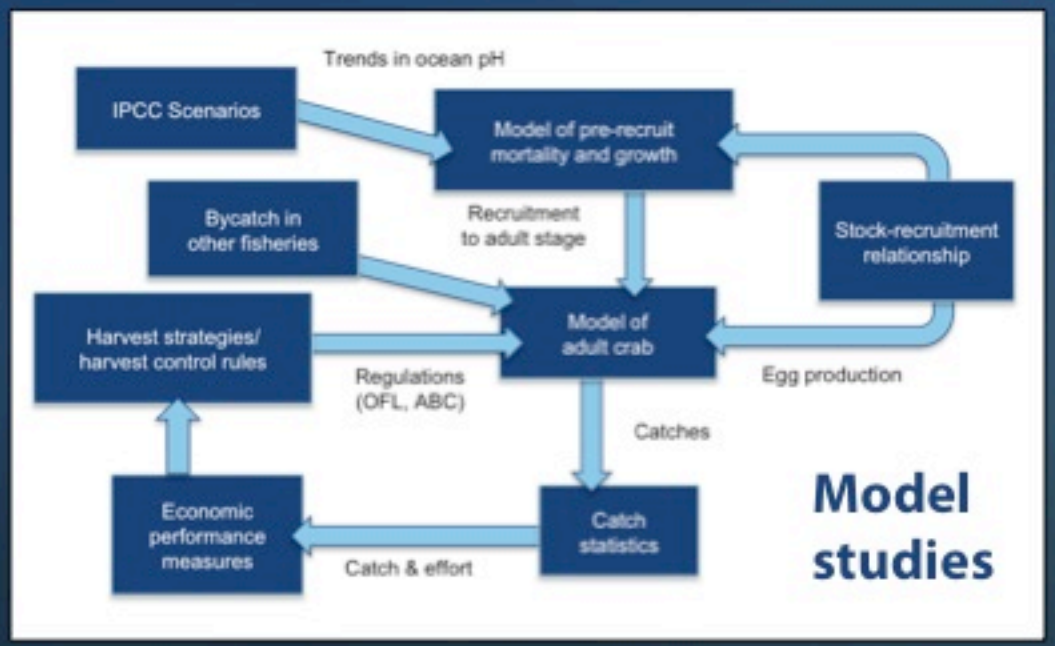
doi:10.1002/loe2.10004

A universal template for OA studies

OA covers a wide range of subject areas, including chemical OA studies, biological monitoring, physiological response experiments, model studies, etc.

If a metadata template can be constructed to apply to many types of OA data sets, the OA data management effort will be much more effective.





Jiang, et. al

doi:10.1002/loe2.10004

OA metadata components

Title

Investigators

Abstract

Type of study

Temporal coverage

Spatial coverage

Geographic names

Location of organism collection

Platforms (e.g., research vessels)

Variable metadata cluster

Publications describing the data set

Supplementary information

1	No	Metadata element name
2	1	Submission Date
3	2	Identification no. of related data sets
4	3	Investigator-1 name
5	4	Investigator-1 institution
6	5	Investigator-1 address
7	6	Investigator-1 phone
8	7	Investigator-1 email
9	8	Investigator-1 researcher ID
10	9	Investigator-1 ID type (ORCID, Researcher ID, etc.)

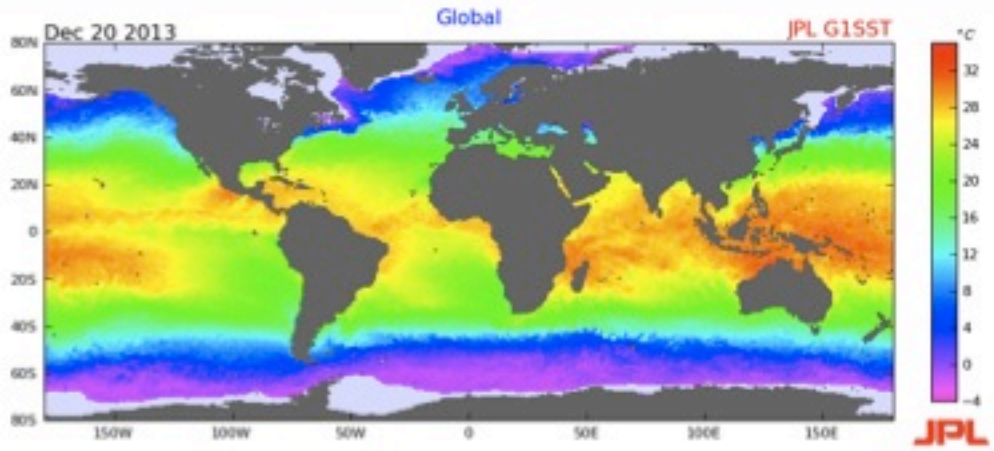
41	40	Spatial reference system
42	41	Geographic names
43	42	Location of organism collection
44	43	Funding agency name
45	44	Funding project title
46	45	Funding project ID (Grant no.)
47	46	Research projects
48	47	Platform-1 name
49	48	Platform-1 ID
50	49	Platform-1 type
51	50	Platform-1 owner
52	51	Platform-1 country

67	66	References
68	67	Supplemental information
69	68	DIC: Variable abbreviation in data files
70	69	DIC: Observation type

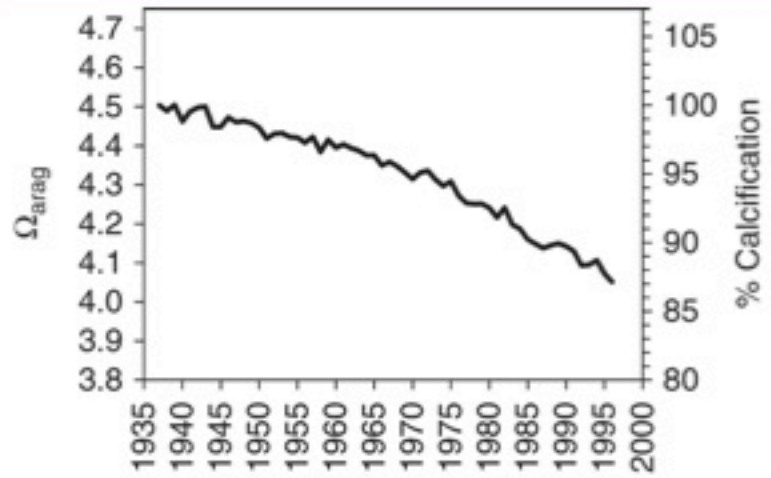
doi:10.1002/loe2.10004

"Variable" as the focal point of the template

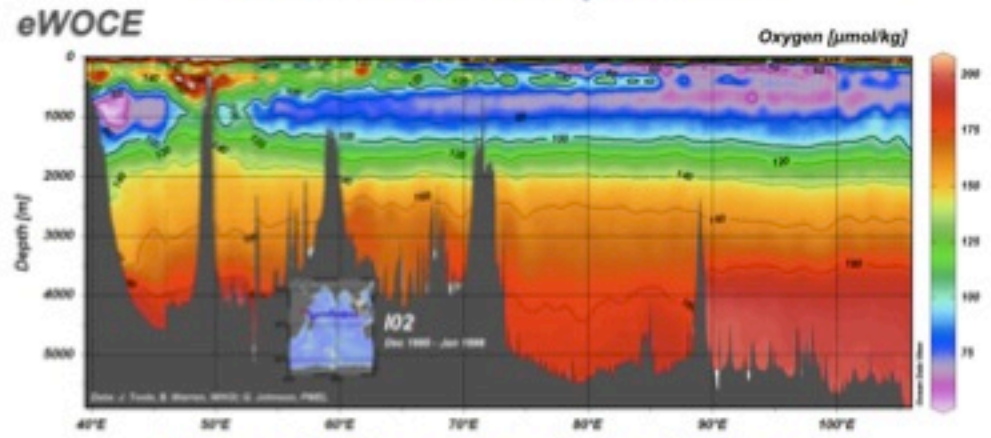
The term "Variables" (or Parameters) refers to the observed or derived properties of a study (e.g., temperature, oxygen, carbonate mineral saturation state, and growth rate, etc.).



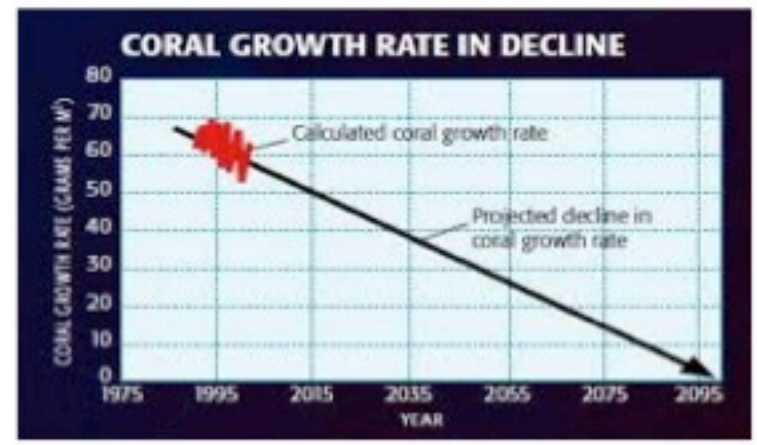
Sea surface temperature



Saturation state



Vertical distribution of oxygen



Growth rate

Schlitzer, R., Electronic Atlas of WOCE Hydrographic and Tracer Data Now Available, Eos Trans. AGU, 81(5), 45, 2000.

Variables are treated as the focal point of the entire metadata template, because we expect them to be the single, most important metadata elements that would be used as search terms to locate an OA data set.

doi:10.1002/loe2.10004

Variable metadata cluster

Root element	Sub-elements
Variable (or Parameter)	**Variable name**
	Unit
	Observation type
	In-situ/manipulation/response variable
	Manipulation method
	Measured or calculated
	Calculation method and parameters
	Sampling instrument
	Analyzing instrument
	Detailed sampling and analyzing info
	Uncertainty
	Method reference (citation)
	Biological subject
	Life stage of the biological subject
	Researcher name
	Researcher institution

Metadata sub-elements of a variable are organized around the variable itself to form a "variable metadata cluster"

This section can be repeated as many times as needed until all of the variables are added to the metadata record.

(The bold elements will be described in more detail on the following slides).

doi:10.1002/loe2.10004

Variable names

Variable abbreviation

The abbreviation or formula of a variable in the data files.

tCO_2
NO_3^-
Ω_{ar}

Full variable name

The detailed variable name collected from the data producers.

Total carbon dioxide,
Nitrate,
Saturation state of aragonite

Controlled vocabulary name

The corresponding term in the controlled vocabulary table for this variable.

Dissolved inorganic carbon,
Nitrate,
Aragonite saturation state

doi:10.1002/loe2.10004

Observation types

The way a variable was captured in relation to its observational context

NOAA

Surface underway

A series of data points along a path at the surface of a water body with monotonically increasing times.

Surface water pCO_2 measured from a voluntary observing ship

Jiang, et. al

doi:10.1002/loe2.10004

ASLO e-Lectures

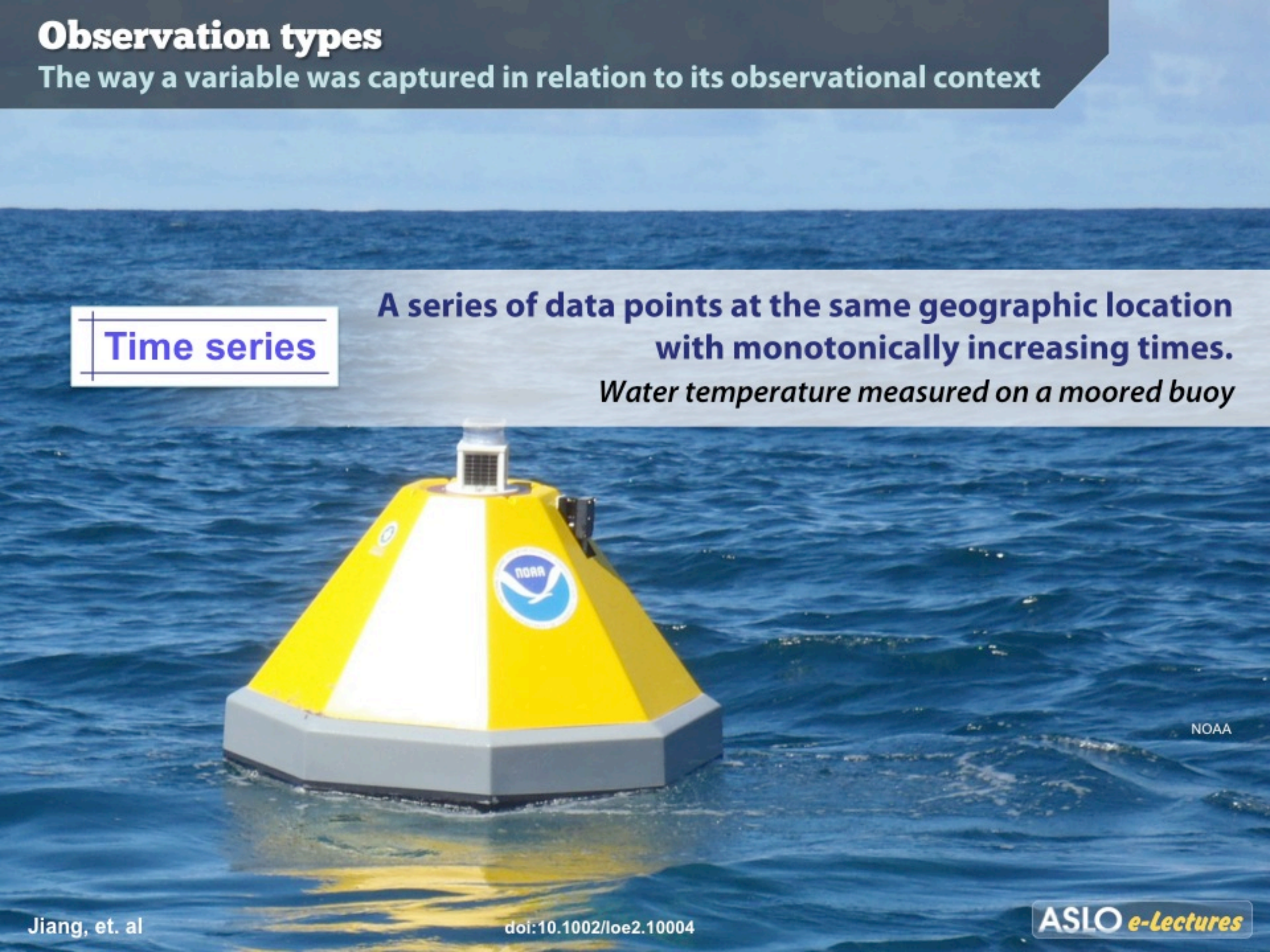
Observation types

The way a variable was captured in relation to its observational context

Time series

A series of data points at the same geographic location with monotonically increasing times.

Water temperature measured on a moored buoy



NOAA

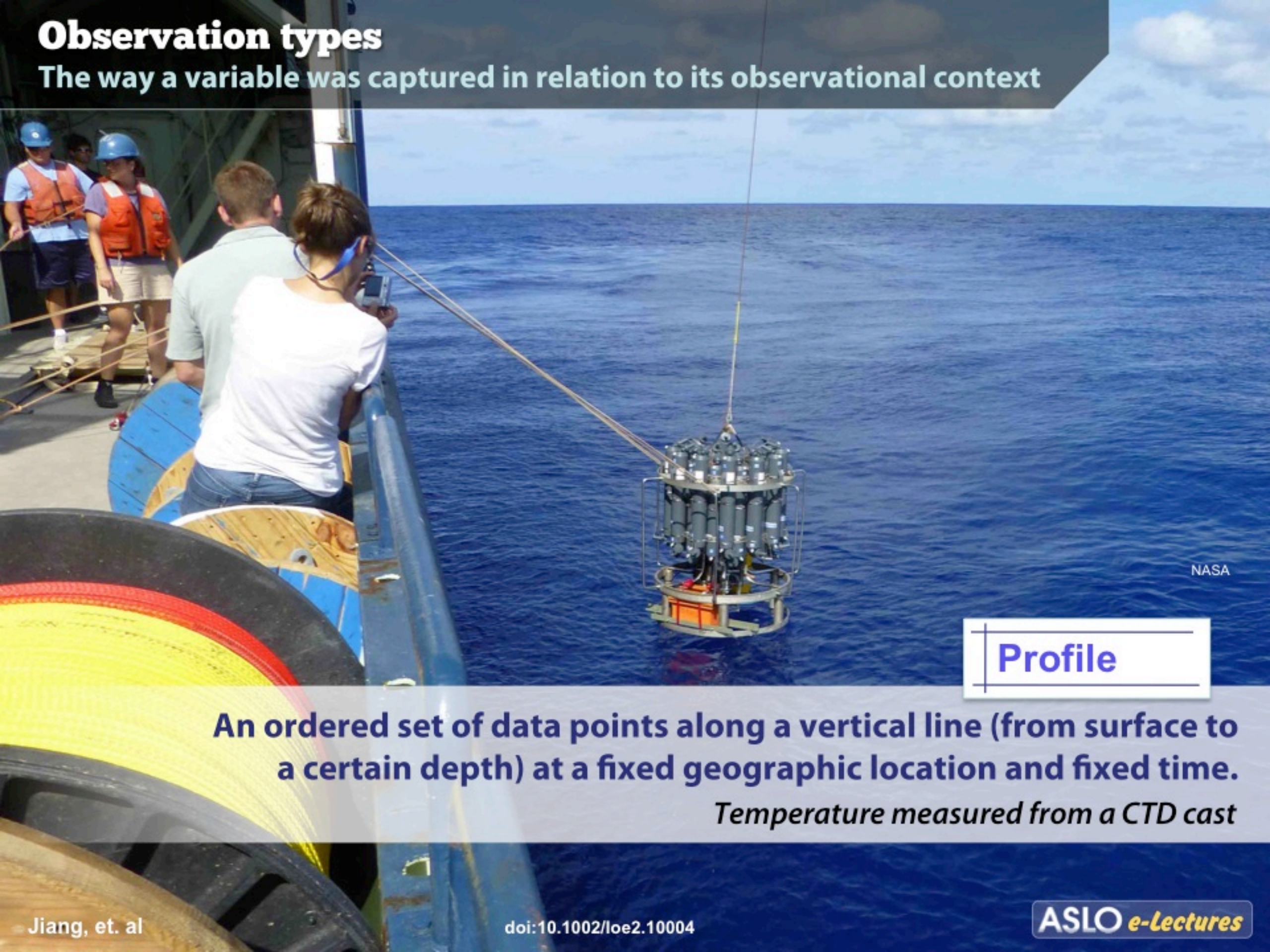
Jiang, et. al

doi:10.1002/loe2.10004

ASLO e-Lectures

Observation types

The way a variable was captured in relation to its observational context



NASA

Profile

An ordered set of data points along a vertical line (from surface to a certain depth) at a fixed geographic location and fixed time.

Temperature measured from a CTD cast

Jiang, et. al

doi:10.1002/loe2.10004

ASLO e-Lectures

Observation types

The way a variable was captured in relation to its observational context

Mote Marine Laboratory

Laboratory experiment

Perturbation experiments in enclosed systems with natural or modified assemblages under modified environmental conditions.

Experimental study to understand how shells of an organism are affected by OA

Jiang, et. al

doi:10.1002/loe2.10004

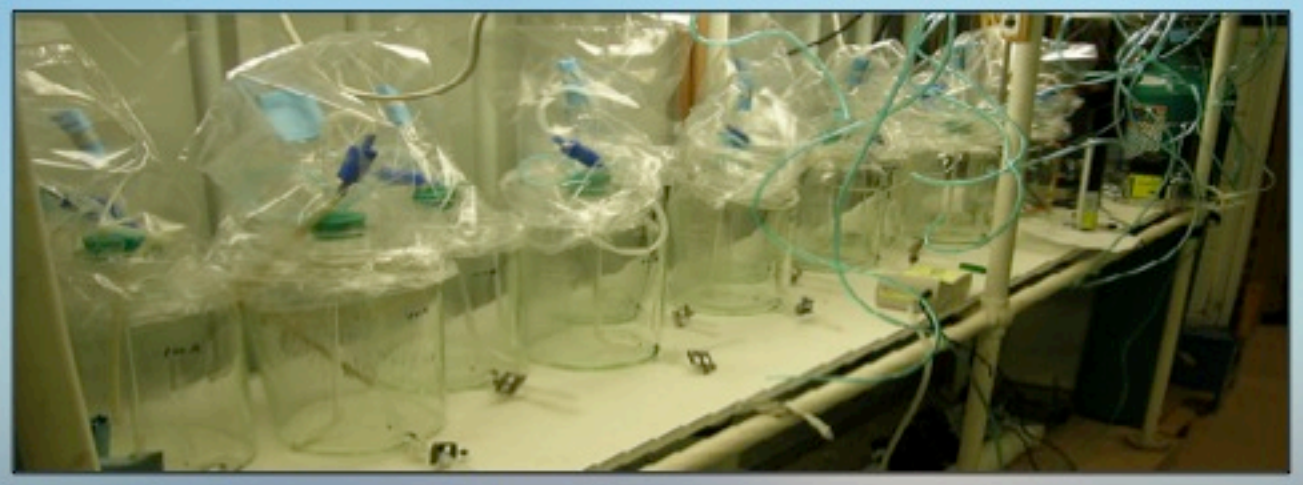
ASLO e-Lectures

In-situ/Manipulation/Response

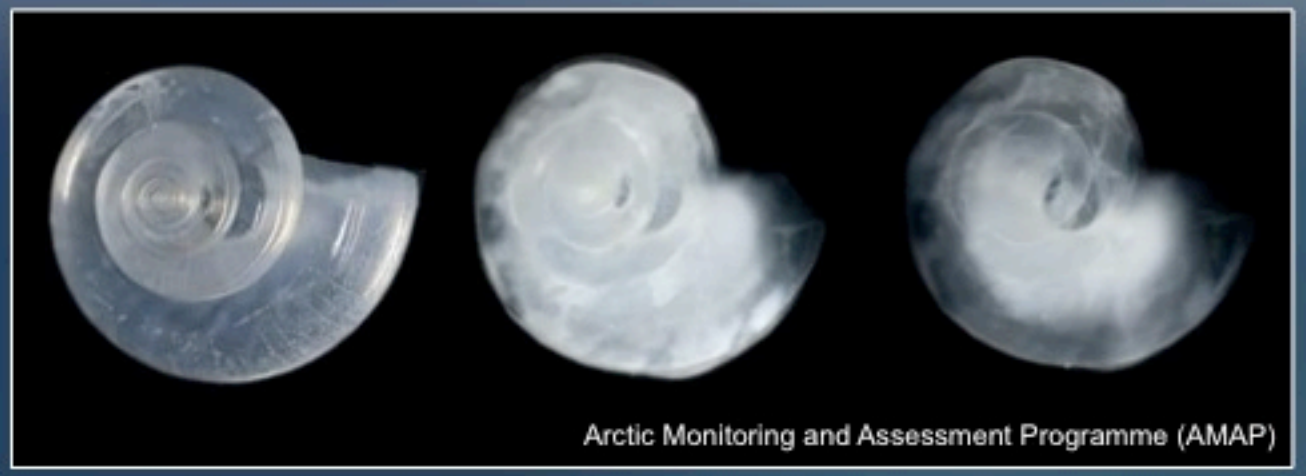
In ocean acidification (OA) studies, variables often fall into these categories:


NOAA

In-situ: Some variables are monitored directly in the field, or collected from the field and measured in the lab later. They fall into the category of in-situ observations.



Manipulation condition: Carbon-related variables, e.g., pH, partial pressure of carbon dioxide (pCO_2), etc. are often manipulated to simulate future OA conditions.


Arctic Monitoring and Assessment Programme (AMAP)

Response variable: Organism related variables, e.g., shell dimensions, calcification rate, growth rate, and larval survival rate, are monitored to understand the responses of the organisms to ocean acidification.

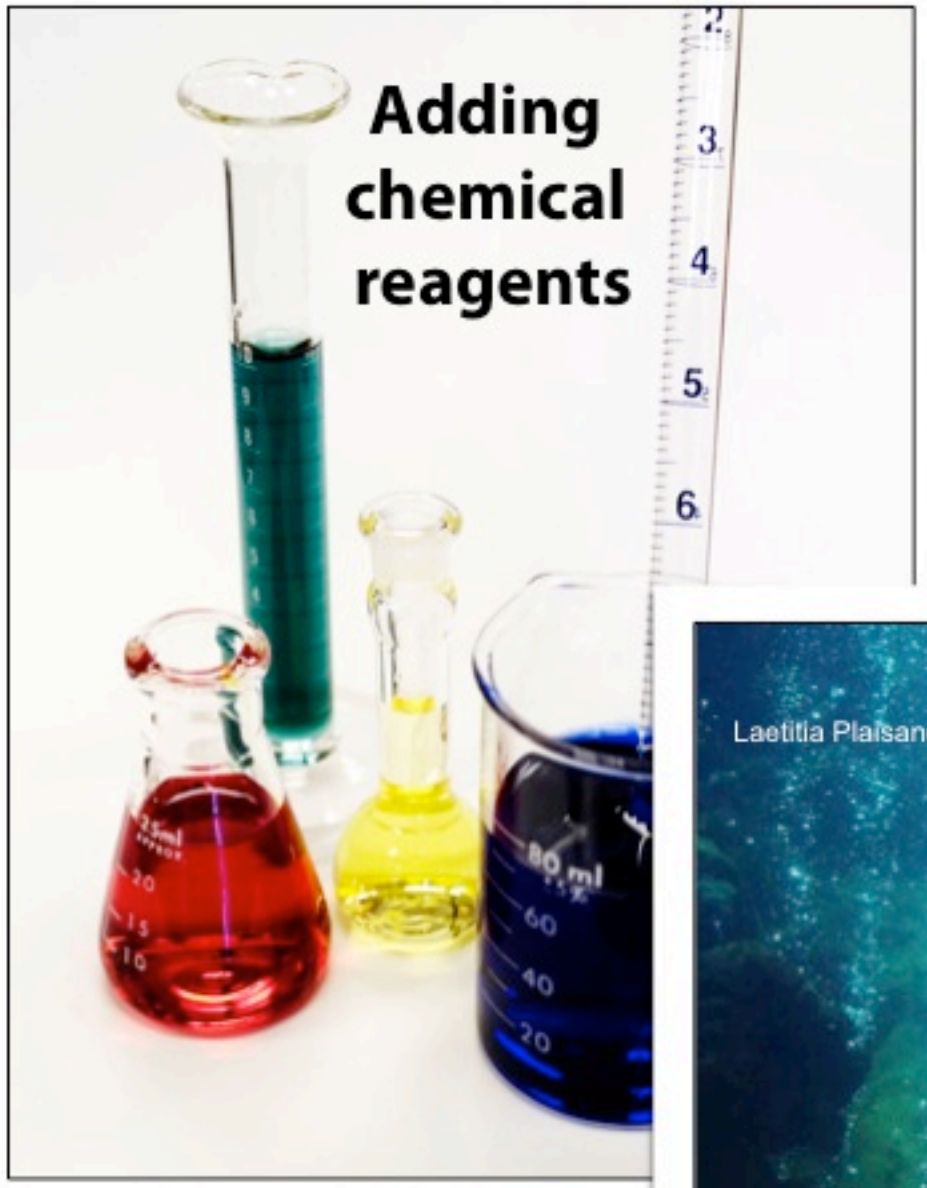
Jiang, et. al

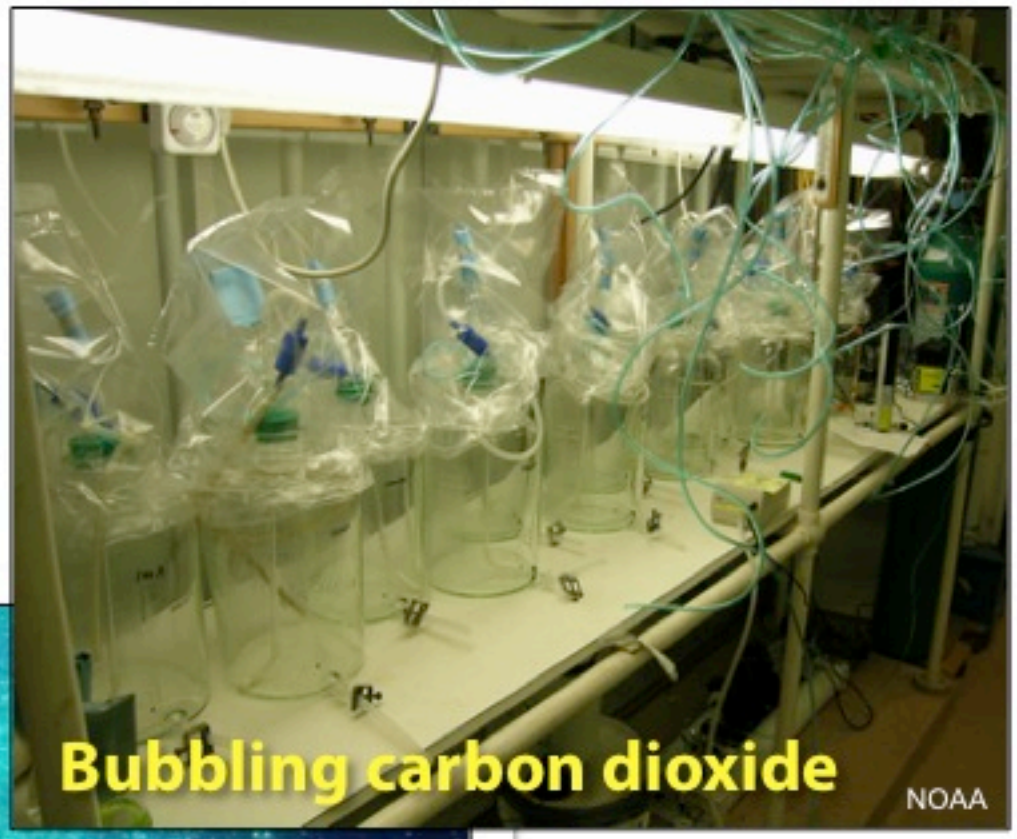
doi:10.1002/loe2.10004

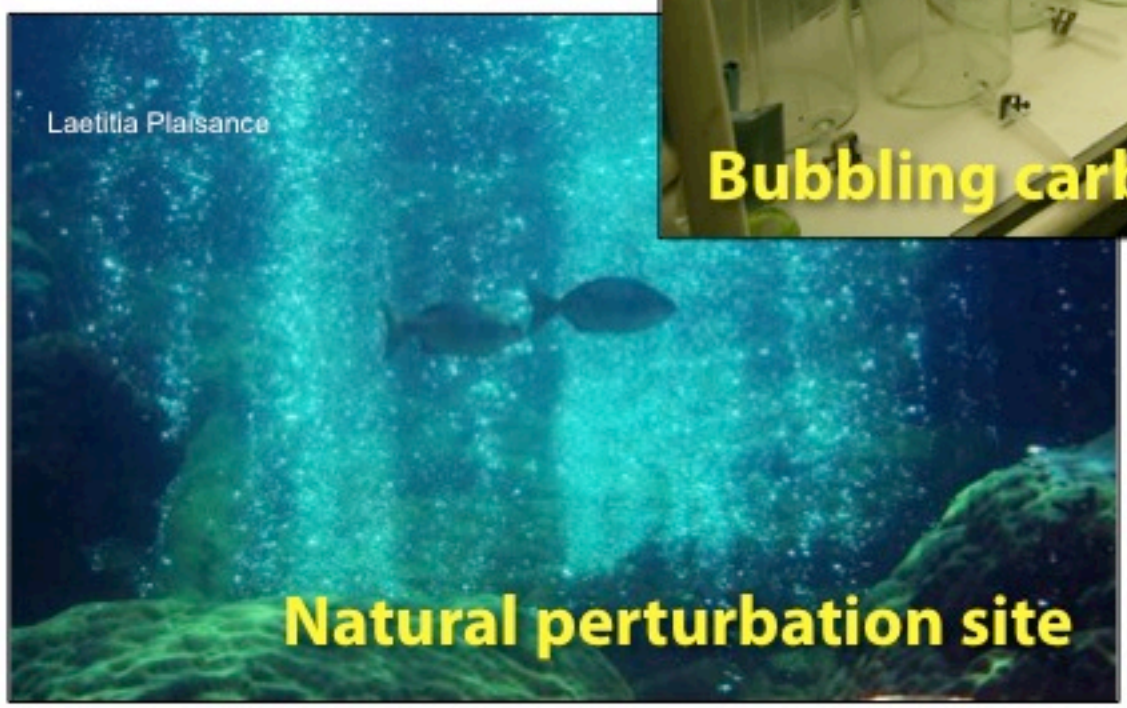
ASLO e-Lectures

Manipulation method

For carbon related variables, their manipulation methods; can be recorded in the field labeled **"Manipulation method"**.



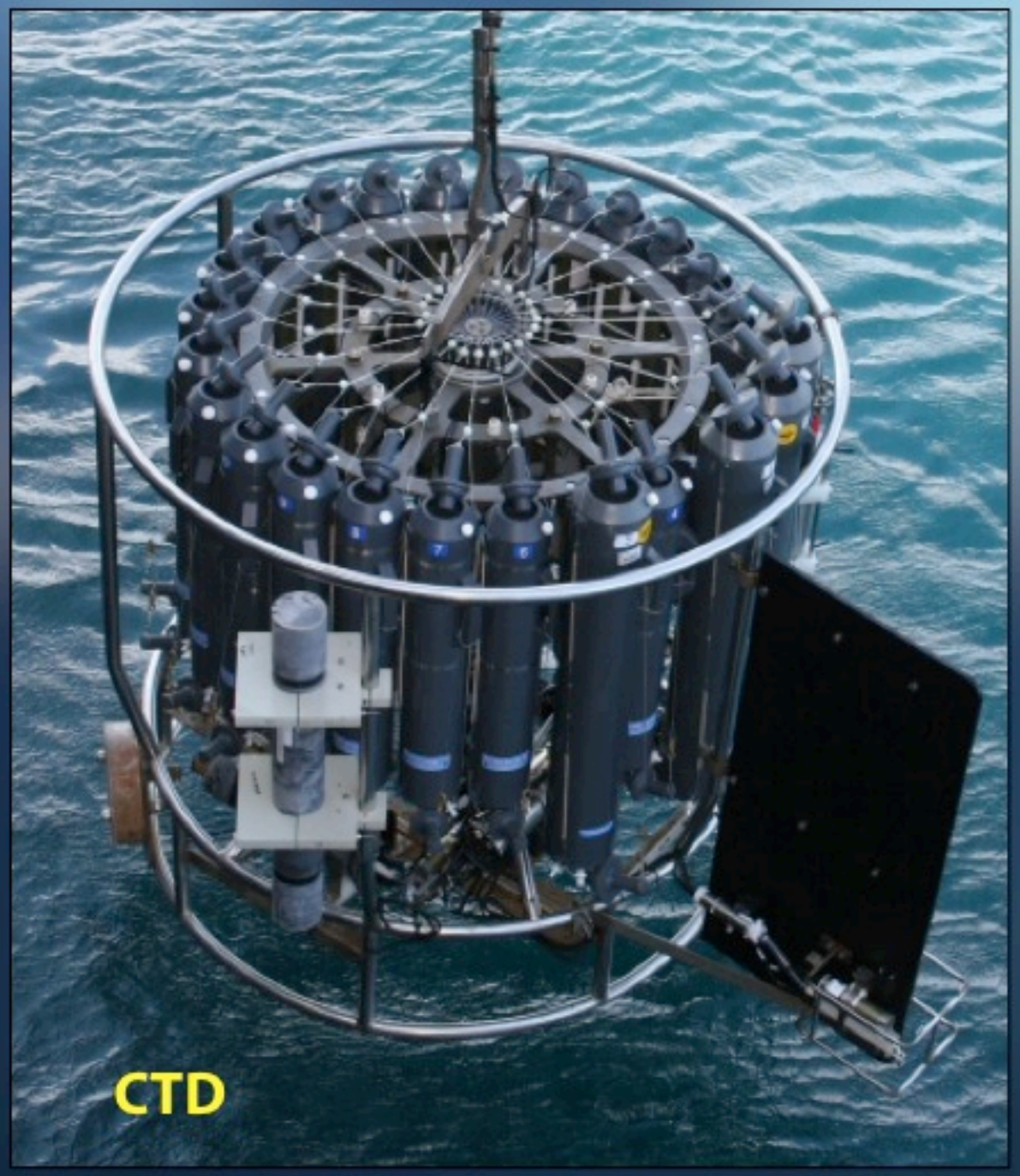




doi:10.1002/loe2.10004

Sampling instrument

Instruments that are used to collect water samples or deploy sensors are here defined as sampling instruments.



Flow through pump

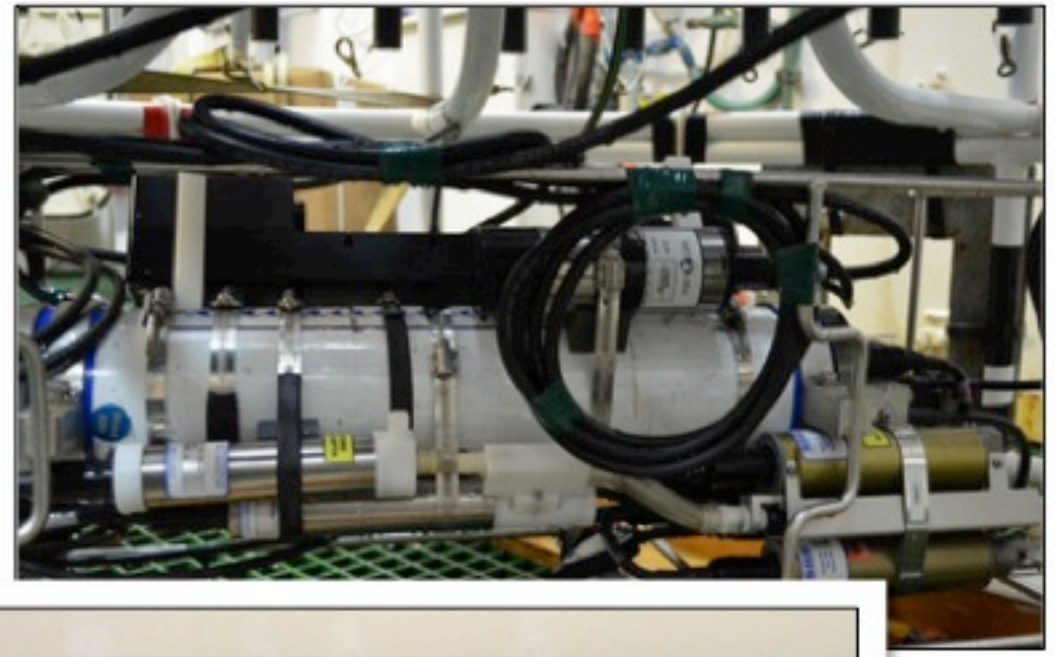
Jiang, et. al

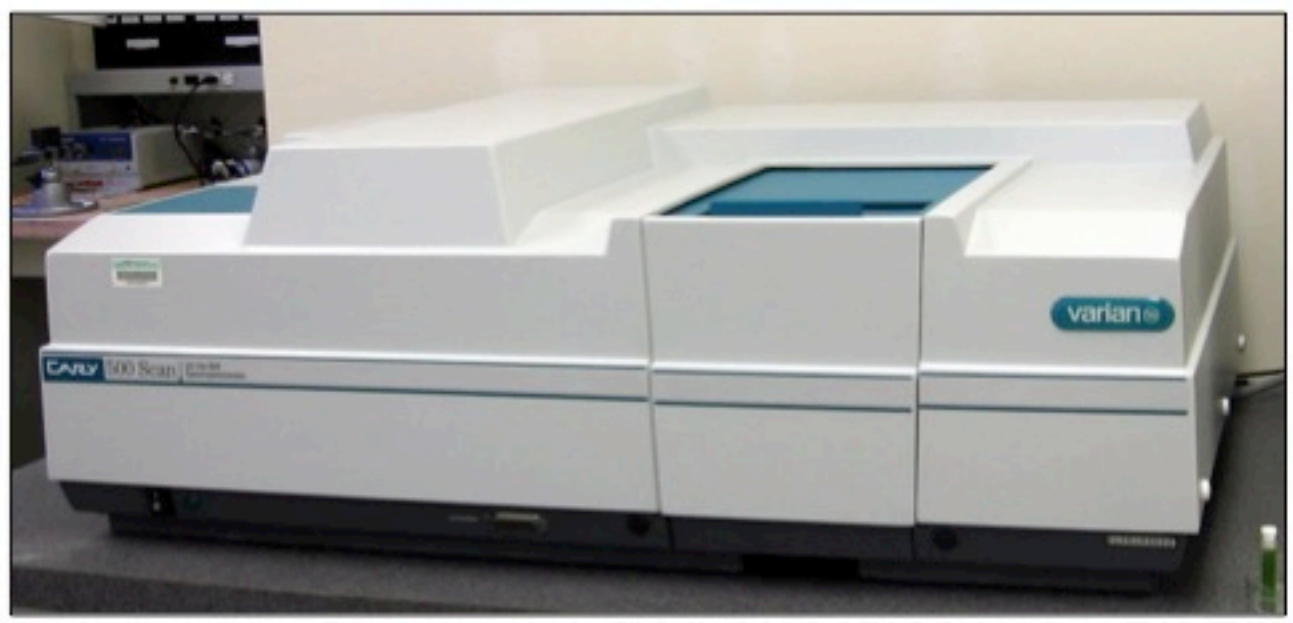
doi:10.1002/loe2.10004

ASLO e-Lectures

Analyzing instrument

Instruments that are used to analyze water samples collected with the sampling instruments, or sensors that are mounted on the sampling instrument to measure some variables of the water.





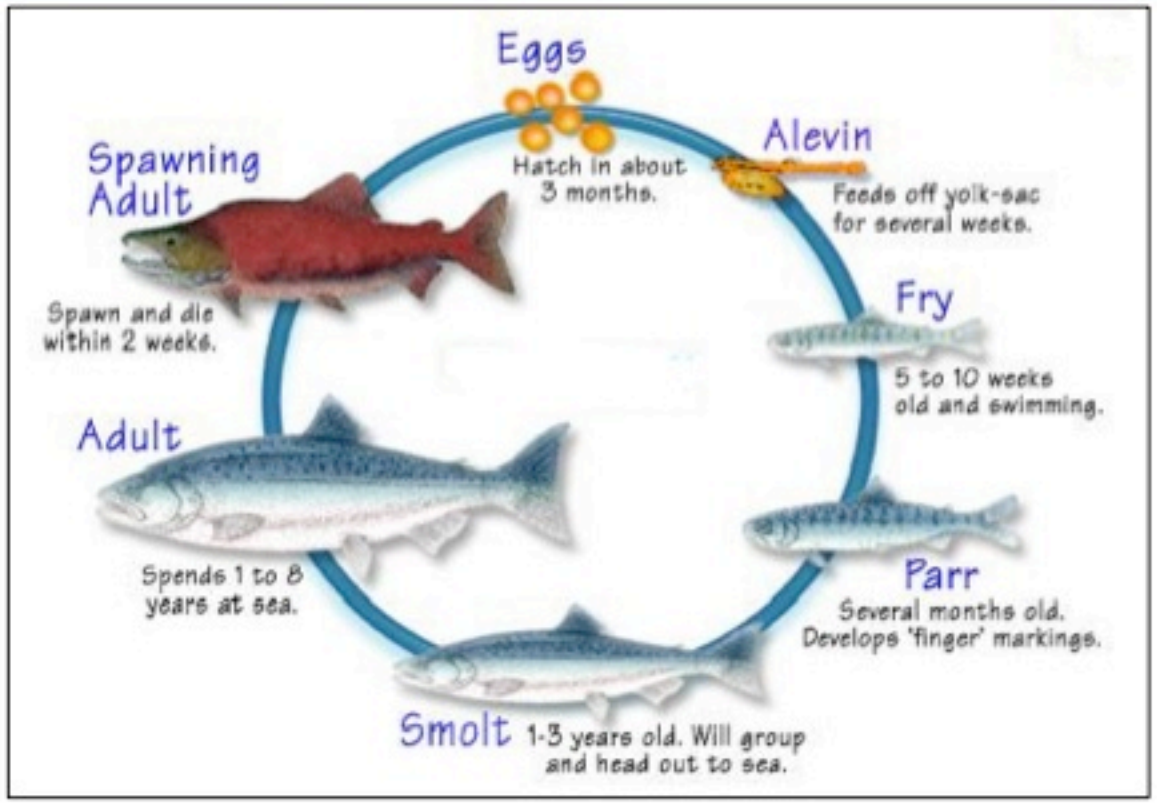
Jiang, et. al

doi:10.1002/loe2.10004

Biological Subject + Life stage

In biological studies, many of the response variables are related to an organism or a biological community.

For example, the variable "Larval survival rate" is not detailed enough without mentioning either the specific organism or the biological community on which the larval survival rate was studied.



Biological subject is where users identify an organism or a biological community, to which the observation of the variable applies.

Life stage of the biological subject is used to document development stage of the species. Input to this field can be eggs, larvae, juveniles, adults, etc.

doi:10.1002/loe2.10004

Spatial Coverage

TYPES OF STUDY:

CTD profile; discrete sampling;

TEMPORAL COVERAGE:

START DATE: 8/12/2011 END DATE: 8/30/2011

SPATIAL COVERAGE:

NORTH BOUND: 48.38

WEST BOUND: -127.55 EAST BOUND: -117.75

SOUTH BOUND: 31.95

GEOGRAPHIC NAMES:

U.S. West Coast California Current System; North Pacific Ocean;

LOCATION OF ORGANISM COLLECTION:

PLATFORMS:

Wecoma (ID: 32WC);

Bounding box

Geographic names

For biological studies, the *bounding box coordinates* and *geographic names* are used to document the location of the water collection.

doi:10.1002/loe2.10004

Spatial Coverage

TYPES OF STUDY:

CTD profile; discrete sampling;

TEMPORAL COVERAGE:

START DATE: 8/12/2011 END DATE: 8/30/2011

SPATIAL COVERAGE:

NORTH BOUND: 48.38

WEST BOUND: -127.55 EAST BOUND: -117.75

SOUTH BOUND: 31.95

GEOGRAPHIC NAMES:

U.S. West Coast California Current System; North Pacific Ocean;

LOCATION OF ORGANISM COLLECTION:

PLATFORMS:

Wecoma (ID: 32WC);

Bounding box

Geographic names

Location of organism collection

"Location of organism collection" is used to document where the organisms are collected.

doi:10.1002/loe2.10004

Platforms

Vessels or other vehicles on which the research is conducted

A platform is often a research vessel, but it could also be a glider, Argo, or satellite, or fixed devices such as moored buoys and towers.

Jiang, et. al

doi:10.1002/loe2.10004

ASLO e-Lectures

EXPOCODE, Section, Cruise ID

EXPOCODE consists of the four-digit *International Council for the Exploration of the Sea* (ICES, http://vocab.ices.dk) ship code and the date of the first day of the cruise in the YYYYMMDD format.

EXPOCODE	32P020130821
Cruise ID	MT901, A16N_2013
Section	A16N

Examples

Cruise ID is the particular ship cruise number, or other alias for the cruise.

Section is the identification code for a research cruise section or leg.

Jiang, et. al

doi:10.1002/loe2.10004

Template availability

The template is archived in the NOAA Institutional Repository

NOAA

NOAA Headquarters, SSMC-3
1315 East-West Hwy,
Silver Spring, MD

http://ezid.cdlib.org/id/doi:10.7289/V5C24TCK

doi:10.7289/V5C24TCK

NOAA Institutional Repository Accession number:
ocn881471371.

Template files

Submission form in Excel spreadsheet

	A	B	C	D
1	No	Metadata element name	Your input	Help reference no.
2	1	Submission Date	4/15/2014	1
3	2	Identification no. of related data sets	nnnnnnn	2
4	3	Investigator-1 name	First M. Last	3.1
5	4	Investigator-1 institution	Name of the Investigator's affiliated institution	3.2
6	5	Investigator-1 address	123 xxxx Rd, City, State 12345	3.3
7	6	Investigator-1 phone	123-456-7890	3.4
8	7	Investigator-1 email	First.last@email.com	3.5
9	8	Investigator-1 researcher ID		3.6
10	9	Investigator-1 ID type (ORCID, Researcher ID, etc.)		3.7
11	10	Investigator-2 name	First M. Last	3.1
12	11	Investigator-2 institution	Name of the Investigator's affiliated institution	3.2
13	12	Investigator-2 address	123 xxxx Rd, City, State 12345	3.3
14	13	Investigator-2 phone	123-456-7890	3.4
15	14	Investigator-2 email	First.last@email.com	3.5
16	15	Investigator-2 researcher ID		3.6
17	16	Investigator-2 ID type (ORCID, Researcher ID, etc.)		3.7
18	17	Investigator-3 name	First M. Last	3.1
19	18	Investigator-3 institution	Name of the Investigator's affiliated institution	3.2
20	19	Investigator-3 address	123 xxxx Rd, City, State 12345	3.3
21	20	Investigator-3 phone	123-456-7890	3.4
22	21	Investigator-3 email	First.last@email.com	3.5
23	22	Investigator-3 researcher ID		3.6
24	23	Investigator-3 ID type (ORCID, Researcher ID, etc.)		3.7
25	24	Data submitter name	First M. Last	4.1
26	25	Data submitter institution	Name of Data submitter's affiliated institution	4.2
27	26	Data submitter address	123 xxxx Rd, City, State 12345	4.3
28	27	Data submitter phone	123-456-7890	4.4
29	28	Data submitter email	First.last@email.com	4.5
30	29	Data submitter researcher ID		4.6
31	30	Data submitter ID type (ORCID, Researcher ID, etc.)		4.7
32	31	Title	The effect of ocean acidification on otolith condition and growth of juvenile scup, *Stenotmus chrysops*	5

doi:10.1002/loe2.10004

Template files

Instruction file

OA metadata template (instructio

Section	Metadata elements			
1	Submission date			
2	Identification no. of related data sets			
3	Investigator (repeat as needed)	Name		
		Institution		
		Address		
		Phone		
		Email		
		researcher ID		
		ID type		
4	Data Submitter	Name		
		Institution		
		Address		
		Phone		
		Email		
		researcher ID		
		ID type		
5	Title			
6	Abstract			
7	Purpose			
8	Temporal coverage	Start date (YYYY-MM-DD)		
		End date (YYYY-MM-DD)		
9	Bounding box coordinates	West bound longitude		
		East bound longitude		
		North bound latitude		
		South bound latitude		
10	Spatial reference system			
11	Geographic names			

ns)

help ref. No	Brief Descriptions
1	Submission Date (for this record only, automated time-stamp may be used for the published metadata record).
2	If you've previously submited a related data set to an archive before, and want to link the archive packages together, please write down all related data package identification numbers here.
3.1	Full name of the investigator (First Middle Last).
3.2	Affiliated institution of the investigator (e.g., Woods Hole Oceanographic Institution).
3.3	Address of the affiliated institution of the investigator.
3.4	Phone number of the investigator.
3.5	Email address of the investigator.
3.6	We recommend to use person identifiers (e.g. ORCID, ResearcherID, etc.) to unambiguously identify the investigator
3.7	Please indicate which type of person identifier is recorded in the above row.
4.1	If applicable, full name of the individual submitting the data to a data center or archive facility.
4.2	Affiliated institution of the data submitter (e.g., Woods Hole Oceanographic Institution).
4.3	Address of the affiliated institution of the data submitter.
4.4	Phone number of the data submitter.
4.5	Email address of the data submitter.
4.6	We recommend to use person identifiers (e.g. ORCID, ResearcherID, etc.) to unambiguously identify the investigator
4.7	Please indicate which type of person identifier is recorded in the above row.
5	Provide a descriptive title for the data set.
6	A narrative summary of the data set.
7	The intentions with which the data set is developed.
8.1	Start date of the first measurement. Please use ISO date format (YYYY-MM-DD).
8.2	End date of the last measurement. Please use ISO date format (YYYY-MM-DD).
9.1	Westernmost longitude of the sampling (decimal degrees, negative for Western Hemisphere longitude).
9.2	Easternmost longitude of the sampling (decimal degrees, negative for Western Hemisphere longitude)
9.3	Northernmost latitude of the sampling (decimal degrees, negative for Southern Hemisphere latitude)
9.4	Southernmost latitude of the sampling (decimal degrees, negative for Southern Hemisphere latitude)
10	A spatial reference system or coordinate reference system defines a specific map projection, as well as transformations between different spatial reference systems. WGS 84 is the reference coordinate system used by the Global Position System.
11	Names of the geographic area where the data collection takes place, e.g., Gulf of Mexico, Baltic Sea, etc.

Jiang, et. al

doi:10.1002/loe2.10004

ASLO e-Lectures

Summary

This e-Lecture illustrates how to document ocean acidification data by explaining major components of an OA metadata template.

Variables and their metadata sub-elements (Variable metadata clusters) are treated as the focal point of the template.

Information about how to access the metadata template files is also stated.

More details about how to document OA data can be found at:
Jiang, L.-Q. S. A. O'Connor, K. M. Arzayus, and A. R. Parsons. 2015. *A metadata template for ocean acidification data*. Earth Syst. Sci. Data, 7, 117-125. *(link)*

doi:10.1002/loe2.10004

Acknowledgements

The template development efforts were funded by the National Oceanic and Atmospheric Administration (NOAA) Ocean Acidification Program (OAP).

Dr. A. Rost Parsons, NOAA National Center for Environmental Information (NCEI), provided guidance for the template development.

We thank Philip Goldstein (University of Colorado) for his contribution to the metadata template development.

We also thank Chris Chambers, Thomas Hurst, Chris Long, Annette DesRochers, Molly Timmers, Nina Bednaršek, Derek Manzello, Renee Carlton, and Jessica Morgan (NOAA), as well as David Kline (Scripps Institution of Oceanography) for their comments on the template.

doi:10.1002/loe2.10004

Author affiliations

Li-Qing Jiang Cooperative Institute for Climate and Satellites-Maryland, Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA.

Krisa M. Arzayus National Centers for Environmental Information, National Oceanographic and Atmospheric Administration, Silver Spring, Maryland, USA.

Jean-Pierre Gattuso Sorbonne Universités, UPMC Univ Paris 06, CNRS-INSU, Laboratoire d'Océanographie de Villefranche, F-06230 Villefranche-sur-mer, France, *and* The Institute for Sustainable Development and International Relations, Sciences Po, F-75007 Paris, France.

Hernan E. Garcia National Centers for Environmental Information, National Oceanographic and Atmospheric Administration, Silver Spring, Maryland, USA.

Cynthia Chandler The Biological and Chemical Oceanography Data Management Office, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA.

Alex Kozyr Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

Yan Yang State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen, Fujian Province, China.

Rob Thomas British Oceanographic Data Center, Liverpool, UK.

Brian Beck Cooperative Institute for Climate and Satellites-Maryland, Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA.

Tobias Spears Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada.

doi:10.1002/loe2.10004

Published by John Wiley & Sons, Inc.

Li-Qing Jiang et al.

How to Document Ocean Acidification Data

doi:10.1002/loe2.10004

Received November 2015; revised version accepted June 2016