

THE OFFICIAL MAGAZINE OF THE OCEANOGRAPHY SOCIETY

Oceanography

CITATION

García, H.E., C. Cosca, A. Kozyr, E. Mayorga, C. Chandler, R.W. Thomas, K. O'Brien, W. Appeltans, S. Hankin, J.A. Newton, A. Gutierrez, J.-P. Gattuso, L. Hansson, M. Zweng, and B. Pfeil. 2015. Data management strategy to improve global use of ocean acidification data and information. *Oceanography* 28(2):226–228, <http://dx.doi.org/10.5670/oceanog.2015.45>.

DOI

<http://dx.doi.org/10.5670/oceanog.2015.45>

COPYRIGHT

This article has been published in *Oceanography*, Volume 28, Number 2, a quarterly journal of The Oceanography Society. Copyright 2015 by The Oceanography Society. All rights reserved.

USAGE

Permission is granted to copy this article for use in teaching and research. Republication, systematic reproduction, or collective redistribution of any portion of this article by photocopy machine, reposting, or other means is permitted only with the approval of The Oceanography Society. Send all correspondence to: info@tos.org or The Oceanography Society, PO Box 1931, Rockville, MD 20849-1931, USA.

Data Management Strategy to Improve Global Use of Ocean Acidification Data and Information

By **Hernan E. Garcia, Cathy Cosca,
Alex Kozyr, Emilio Mayorga,
Cynthia Chandler,
Robert W. Thomas, Kevin O'Brien,
Ward Appeltans, Steve Hankin,
Jan A. Newton, Angelica Gutierrez,
Jean-Pierre Gattuso,
Lina Hansson, Melissa Zweng,
and Benjamin Pfeil**

Ocean acidification (OA) refers to the general decrease in pH of the global ocean as a result of absorbing anthropogenic CO₂ emitted in the atmosphere since preindustrial times (Sabine et al., 2004). There is, however, considerable variability in ocean acidification, and many careful measurements need to be made and compared in order to obtain scientifically valid information for the assessment of patterns, trends, and impacts over a range of spatial and temporal scales, and to understand the processes involved. A single country or institution cannot undertake measurements of worldwide coastal and open ocean OA changes; therefore, international cooperation is needed to achieve that goal. The OA data that have been, and are being, collected represent a significant public investment. To this end, it is critically important that researchers (and others) around the world are easily able to find and use reliable OA information that range from observing data (from time-series moorings, process studies, and research cruises), to biological response experiments (e.g., mesocosm), data products, and model output.

Two major factors currently limit access to OA data. First, the data reside in many geographic locations in varying formats, are subject to different data quality procedures, and are described using different vocabularies and metadata (descriptions of information relevant to the data; e.g., methods and chemical standards used). Second, the data are not made available using interoperable online services that would allow data queries across different networks, making it difficult to discover and access data from a single, “one-stop” location. In addition, researchers may need to request permission to use the data from a particular data source, and then convert the data into a common digital format such as Network Common Data Form (NetCDF) in order to make comparisons and syntheses.

Data managers from a wide range of institutions are developing ideas for a much more coordinated global OA data management system to improve public access to observations, experiments, and data synthesis products and to enable human-to-machine and machine-to-machine data transactions across different networks. The proposed system will enable different data network services to communicate and exchange OA data in a consistent manner. It will also facilitate data integration, for example, combining multidisciplinary observational in situ and remotely sensed data into a single data stream. The goal is to meet the current information needs of both researchers and policymakers with minimal cost while maximizing the long-term benefits to scientific research.

This data management strategy is closely aligned to the Global Ocean Acidification Observing Network (GOA-ON), based on already existing OA data-gathering activities (repeat hydrographic surveys, time-series stations, moorings, float and glider observations, and volunteer observing ships), with additional observations to cover biological impacts. GOA-ON was initiated in 2012 by scientists from 28 countries, recognizing that its success depends

on an ambitious vision to improve the international coordination of OA data management (Newton et al., 2014).

A common OA data access Web data portal (a “one-stop shop”) will benefit from adoption of uniform best practices for scientific management of the data. The Web portal needs to accept and deliver data (and associated documentation) from data centers, data publishing locations, academic institutions, and other data partners. Figure 1 illustrates the scientific stewardship data life cycle for making globally distributed data of known quality available through a user-friendly search and access data portal. This portal should enable easy discovery, access, long-term archiving, and visualization of the wide range of observations and their associated data products. The data system architecture needs

to build on existing data management practices employed at national and international levels, emphasizing measurement collection, end-to-end data management (data life cycle from collection of data up to information, data products, and long-term archiving), and research aligned with the goals of the ocean acidification community. However, the proposed system is not intended to solve all of the problems and challenges associated with managing OA data, and much work (and goodwill) is needed to fully implement such a system.

Serving these diverse globally distributed data will require a cooperative approach between scientists and data managers. Open and full sharing of scientific observational data and associated metadata will be needed in order to study the ocean as a system, to provide timely

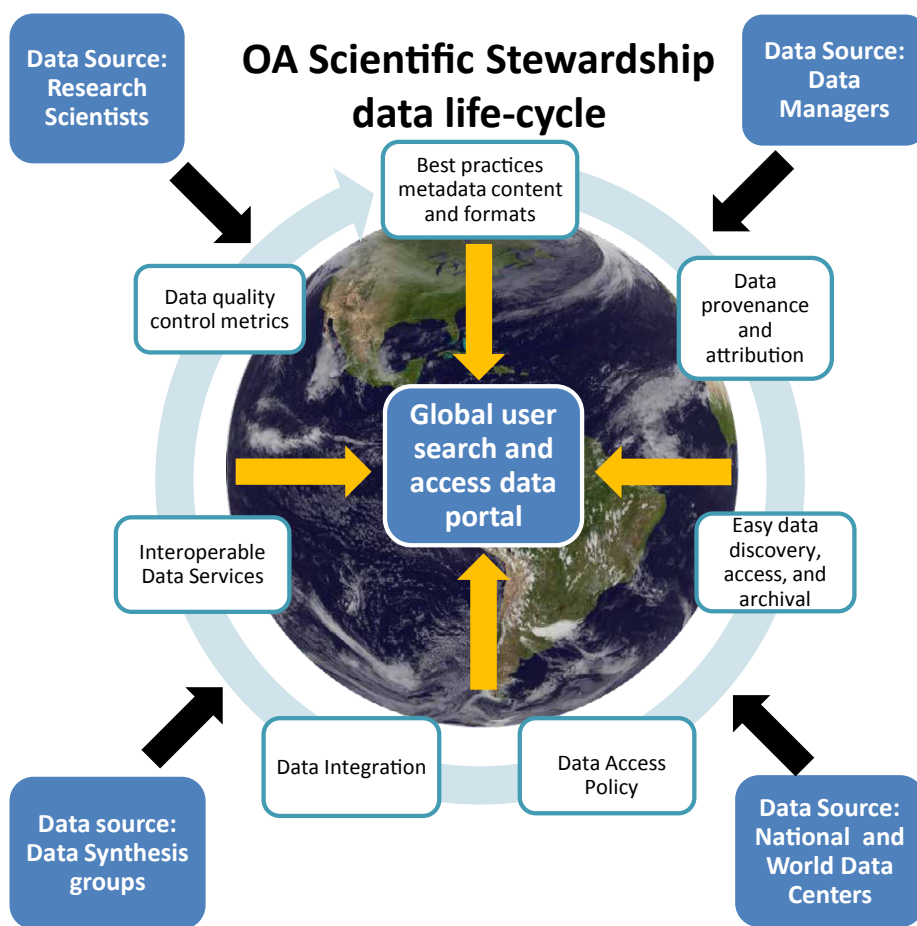


FIGURE 1. Conceptual representation of scientific stewardship for ocean acidification data, showing how globally distributed data of known quality can be made discoverable and accessible through a data portal via interoperable data services.


and reliable forecasts and warnings, to conserve and manage marine resources, and to support informed economic and regulatory decisions.

Detailed metadata are necessary for valid and consistent use of OA data by the scientific community now and into the future (Dickson et al., 2007; Riebesell et al., 2010). For instance, information is needed on measurement protocols, whether Certified Reference Materials were used, pH scales (National Bureau of Standards, free, total, or seawater), reporting units and nomenclature of measured and calculated variables, data reproducibility, and other relevant data such as temperature and salinity.

Clear policies will be needed that address timely availability and sharing of data, including the use of unique and citable digital identifiers to ensure appropriate attribution, provenance-tracking, and recognition of the scientists and project programs responsible for collecting and analyzing the data. While scientific journals increasingly require archiving of data that support scientific results, the data are not always easily found nor are they made to conform to common digital formats that would allow their wider re-use. Quality control of the data is the responsibility of the scientists collecting the data and research groups assembling data synthesis products. The main purpose here is to design a system that serves the critical needs for data and information discovery and access in a coordinated manner.

When creating an appropriate data-sharing framework for OA data, it is critical to learn from, and build on, the data stewardship experiences of other international efforts. Examples of such efforts include WOCE (World Ocean Circulation Experiment), EPOCA (European Project on Ocean Acidification), CLIVAR (Climate and Variability), GLODAP (Global Ocean Data Analysis Project), and SOCAT (Surface Ocean CO₂ Atlas). These are examples of high-quality data compilation products that generally address specific ocean observational and scientific

communities. To help ensure wide dissemination of the data and information, it will be important to coordinate closely with international programs and initiatives such as OA-ICC (Ocean Acidification International Coordination Centre), IODE (International Oceanographic Data and Information Exchange of the Intergovernmental Oceanographic Commission of UNESCO), International Ocean Carbon Coordination Project (IOCCP), GEO (Group of Earth Observations), and others.

The OA data management actions outlined here provide an adaptable coordination strategy based on collaboration between scientists, data centers, and data publishers, and promote the adoption of common data retrieval protocols and best practices for OA metadata production and long-term preservation across political boundaries. Given the importance of assessing OA variability and ecosystem responses at local, regional, and global levels, and at daily to interannual time scales, a well-coordinated and internationally agreed upon OA data management and access strategy can be considered an essential component of environmental security and global resilience management. 

REFERENCES

- Dickson, A.G., C.L. Sabine, and J.R. Christian. 2007. Guide to best practices for ocean CO₂ measurements. *PICES Special Publication* 3:1–191, http://cdiac.ornl.gov/ftp/oceans/Handbook_2007/Guide_all_in_one.pdf.
- Newton, J.A., R.A. Feely, E.B. Jewett, P. Williamson, and J. Mathis. 2014. *Global Ocean Acidification Observing Network: Requirements and Governance Plan*. First Edition, 60 pp. http://goa-on.org/docs/GOA-ON_plan_print.pdf
- Riebesell, U., V.J. Fabry, L. Hansson, and J.-P. Gattuso, eds. 2010. *Guide to Best Practices for Ocean Acidification Research and Data Reporting*. Luxembourg: Publications Office of the European Union, 260 pp., <http://www.epoca-project.eu/index.php/guide-to-best-practices-for-ocean-acidification-research-and-data-reporting.html>.
- Sabine, C.L., R.A. Feely, N. Gruber, R.M. Key, K. Lee, J.L. Bullister, R. Wanninkhof, C.S. Wong, D.W.R. Wallace, B. Tilbrook, and others. 2004. The oceanic sink for anthropogenic CO₂. *Science* 305:367–371, <http://dx.doi.org/10.1126/science.1097403>.

AUTHORS. Hernan E. Garcia (hernan.garcia@noaa.gov) is Oceanographer, National Oceanic and Atmospheric Administration (NOAA), National Centers for Environmental Information (National Oceanographic

Data Center), Silver Spring, MD, USA. Cathy Cosca is Physical Scientist, NOAA Pacific Marine Environmental Laboratory, Seattle, WA, USA. Alex Kozyr is Scientist, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN, USA. Emilio Mayorga is Senior Oceanographer, University of Washington Applied Physics Laboratory, Seattle, WA, USA. Cynthia Chandler is Information Systems Associate, Biological & Chemical Oceanography Data Management Office, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. Robert W. Thomas is Data Scientist, British Oceanographic Data Centre, Liverpool, UK. Kevin O'Brien is a software engineer for the University of Washington/Joint Institute for the Study of the Atmosphere and Ocean (JSIAO), Seattle, WA, USA. Ward Appeltans is Program Specialist, International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission of UNESCO, Oostende, Belgium. Steve Hankin is Computer Scientist, NOAA Pacific Marine Environmental Laboratory, Seattle, WA, USA. Jan A. Newton is Senior Principal Oceanographer, University of Washington Applied Physics Laboratory, Seattle, WA, USA. Angelica Gutierrez is Hydrologist, NOAA, National Weather Service, Silver Spring, MD, USA. Jean-Pierre Gattuso is Research Professor, Centre National de la Recherche Scientifique/ Université Pierre et Marie Curie - Paris 6, Villefranche-sur-Mer, France. Lina Hansson is Project Officer, Ocean Acidification International Coordination Centre, International Atomic Energy Agency Environment Laboratories, Monaco. Melissa Zweng is Oceanographer, NOAA, National Centers for Environmental Information (National Oceanographic Data Center), Silver Spring, MD, USA. Benjamin Pfeil is Senior Engineer, University of Bergen, Bergen, Norway.

ARTICLE CITATION

Garcia, H.E., C. Cosca, A. Kozyr, E. Mayorga, C. Chandler, R.W. Thomas, K. O'Brien, W. Appeltans, S. Hankin, J.A. Newton, A. Gutierrez, J.-P. Gattuso, L. Hansson, M. Zweng, and B. Pfeil. 2015. Data management strategy to improve global use of ocean acidification data and information. *Oceanography* 28(2):226–228, <http://dx.doi.org/10.5670/oceanog.2015.45>.