Integrated Observations and Informatics Improve Understanding of Changing Marine Ecosystems

Abigail Benson  
US Geological Survey

Cassandra M. Brooks  
University of Colorado

Gabrielle Canonico  
National Oceanographic and Atmospheric Administration

Emmett Duffy  
Smithsonian Institution

Frank Muller-Karger  
University of South Florida, carib@usf.edu

See next page for additional authors

Follow this and additional works at: https://scholarcommons.usf.edu/msc_facpub

Part of the Life Sciences Commons

Scholar Commons Citation
Benson, Abigail; Brooks, Cassandra M.; Canonico, Gabrielle; Duffy, Emmett; Muller-Karger, Frank; Sosik, Heidi M.; Miloslavich, Patricia; and Klein, Eduardo, "Integrated Observations and Informatics Improve Understanding of Changing Marine Ecosystems" (2018). Marine Science Faculty Publications. 749.  
https://scholarcommons.usf.edu/msc_facpub/749
Integrated Observations and Informatics Improve Understanding of Changing Marine Ecosystems

Abigail Benson1*, Cassandra M. Brooks2, Gabrielle Canonico3, Emmett Duffy4, Frank Muller-Karger5, Heidi M. Sosik6, Patricia Miloslavich7,8 and Eduardo Klein8

1 U.S. Geological Survey, Lakewood, CO, United States, 2 Environmental Studies Program, University of Colorado Boulder, Boulder, CO, United States, 3 National Oceanographic and Atmospheric Administration, U.S. Integrated Ocean Observing System, Silver Spring, MD, United States, 4 Tennebaum Marine Observatories Network, Smithsonian Institution, Edgewater, MD, United States, 5 Institute for Marine Remote Sensing, College of Marine Science, University of South Florida, St. Petersburg, FL, United States, 6 Woods Hole Oceanographic Institution, Woods Hole, MA, United States, 7 Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia, 8 Departamento de Estudios Ambientales, Universidad Simón Bolívar, Caracas, Venezuela

Marine ecosystems have numerous benefits for human societies around the world and many policy initiatives now seek to maintain the health of these ecosystems. To enable wise decisions, up to date and accurate information on marine species and the state of the environment they live in is required. Moreover, this information needs to be openly accessible to build indicators and conduct timely assessments that decision makers can use. The questions and problems being addressed demand global-scale investigations, transdisciplinary science, and mechanisms to integrate and distribute data that otherwise would appear to be disparate. Essential Ocean Variables (EOVs) and marine Essential Biodiversity Variables (EBVs), conceptualized by the Global Ocean Observing System (GOOS) and the Marine Biodiversity Observation Network (MBON), respectively, guide observation of the ocean. Additionally, significant progress has been made to coordinate efforts between existing programs, such as the GOOS, MBON, and Ocean Biogeographic Information System collaboration agreement. Globally and nationally relevant indicators and assessments require increased sharing of data and analytical methods, sustained long-term and large-scale observations, and resources to dedicated to these tasks. We propose a vision and key tenets as a guiding framework for building a global integrated system for understanding marine biological diversity and processes to address policy and resource management needs. This framework includes: using EOVs and EBVs and implementing the guiding principles of Findable, Accessible, Interoperable, Reusable (FAIR) data and action ecology. In doing so, we can encourage relevant, rapid, and integrative scientific advancement that can be implemented by decision makers to maintain marine ecosystem health.

Keywords: ocean observing, integrated assessments, marine ecosystems, data sharing, essential ocean variables, essential biodiversity variables, FAIR data, action ecology

Abbreviations: CBD, Convention on Biological Diversity; EBV, Essential Biodiversity Variable; EML, Ecological Metadata Language; EOV, Essential Ocean Variable; FAIR, Findable, Accessible, Interoperable, Reusable; GBIF, Global Biodiversity Information Facility; GEO, Group on Earth Observations; GOOS, Global Ocean Observing System; IOC, Intergovernmental Oceanographic Commission; MarineGEO, Marine Global Earth Observatory; MBON, Marine Biodiversity Observation Network; OBIS, Ocean Biogeographic Information System; SDG, Sustainable Development Goals; UNESCO, United Nations Educational, Scientific and Cultural Organization.
INTRODUCTION

Marine ecosystems provide a wealth of services that sustain and enrich human life (Barbier, 2017; Salomon and Dahms, 2018). The international community recognizes the importance of these services and the threats facing them, as shown by numerous initiatives to maintain the health of marine ecosystems, including for example the CBD Strategic Plan for Biodiversity (United Nations Environment Programme [UNEP], 2010), the United Nations SDGs (Lu et al., 2015; Anderson et al., 2017), and the upcoming United Nations Decade of Ocean Science for Sustainable Development (UNESCO-IOC, 2017). These initiatives seek to empower nations to achieve broad societal benefits, including mitigating the effects of climate change and protecting the biodiversity of our global oceans and coastal areas.

The success of these global initiatives demands identifying essential variables that can be measured widely in standard ways to provide large-scale frameworks to address global ocean conservation priorities (Malone T.C. et al., 2014; Müller-Karger et al., 2018). EOVs (Miloslavich et al., 2018) and EBVs (Pereira et al., 2013) can provide high level guidance to scientists and managers for monitoring, biology, biodiversity, and ecosystems and can help to inform policy initiatives for maintaining ecosystem health. The Essential Variables were developed to assist the observing community focus their monitoring efforts around measurements necessary for understanding biological change in a sustained, operational manner, which can then be incorporated into models.

Policy makers need sound, timely, and transdisciplinary science to feed into assessments for marine conservation and sustainable development initiatives (Malone T. et al., 2014; Geijzendorffer et al., 2016; Miloslavich et al., 2018). Progress toward global goals can be tracked and met by collecting appropriate data to establish some baseline conditions for reference, and by monitoring change relative to those conditions. Answers to policy reporting needs (e.g., the ten targets of SDG 14 relating to marine systems) and scientific questions posed by EOVs and EBVs can be achieved through data-intensive investigations. Not limited to one researcher or one project, the integration of data from multiple sources is key to understanding ecology in the globally connected ocean. But, it is not enough to simply collect those data; the data need to be shared broadly and, when feasible, in real time.

The FAIR Guiding Principles (Wilkinson et al., 2016) provide a roadmap for data management that facilitates data reuse and sharing and provides a platform for integrative scientific discovery and policy applications.

In conjunction with the FAIR data sharing principles, action ecology (White et al., 2015) provides an additional framework to support policy using science. Action ecology is an approach focused on incorporating transdisciplinary input, closing the gap between findings and implementation, utilizing the best available technology, and providing policy-ready recommendations – all of which are imperative to the success of a system for marine biological observations.

The needs and goals outlined above require broad societal investment in monitoring biodiversity, new social and cultural attitudes toward data sharing, and adherence to FAIR data principles and best practices. In this paper, we offer suggestions for ways to address these challenges and present a vision for an integrated system of marine biological observations and informatics to monitor change in marine ecosystems.

A BRIEF HISTORY

Under the auspices of the IOC of the UNESCO, the GOOS has developed a framework to address operational needs on ocean physical, biogeochemical, and biological information.Modeled after the Framework for Ocean Observing (Lindstrom et al., 2012), the GOOS framework is intended to be nimble in addressing the needs of the ocean observing community. The information made available using the GOOS framework is required by industries (e.g., fisheries, shipping trade, mining), defense, and environmental agencies around the world. Partners in the process to develop the GOOS framework have included national observing programs, the World Meteorological Organization, the UN Environment Programme, and the International Council for Science. Increasingly, the GOOS framework is the basis for assessing ocean health.

Until recently, GOOS provided advice mostly in the disciplines of physics, climate, and biogeochemistry (Miloslavich et al., 2015, 2018). The GOOS Physics Panel, formed in 1995, was built on the expertise and advice of the Ocean Observations Panel for Climate, while the GOOS Biogeochemistry Panel was built on the International Ocean Carbon Coordination Project. These groups defined their EOVs based on specific scientific and societal requirements driven mostly by climate change and the need for weather forecasts (Lorenzoni and Benway, 2013; Bauer et al., 2015) as well as in support of the requirements of the Global Climate Observing System in meeting the needs of the United Nations Framework Convention on Climate Change (Houghton et al., 2012; Bojinski et al., 2014).

The GOOS Biology and Ecosystems Panel (GOOS BioEco) was established in 2015 to identify biological EOVs based on their societal relevance and technical and scientific scalability (Miloslavich et al., 2018). The IOC had supported initiatives such as the Coastal Ocean Observations Panel which drafted the coastal module for GOOS and an initial set of common variables (United Nations Environment Programme [UNEP], 2002, 2005). Another supported initiative was an IOC technical expert workshop convened in 2013 which covered biological observations with a broader oceanographic scope (UNESCO-IOC, 2014). Within the realm of national or regional observing programs, the Australian Integrated Marine Observing System (Integrated Marine Observing System [IMOS], 2015; Lara-Lopez et al., 2016), the United States Integrated Ocean Observing System (National Ocean Council, 2016) and the Southern Ocean Observing System (Constable et al., 2016) have incorporated biological measurements.

The GEO established the Biodiversity Observation Network (GEO BON), which in 2016 formed a working group to address
the marine theme (the MBON; Duffy et al., 2013; Muller-Karger et al., 2014) and to develop sets of marine EBVs (Pereira et al., 2013). These marine EBVs are complementary to the GOOS biological EOVs (Muller-Karger et al., 2018).

The OBIS began as the data repository for the Census of Marine Life (Grassle, 2000) and has since expanded to become the most complete platform for global marine biological observations. As part of the IOC Intergovernmental Oceanographic Data and Information Exchange program, OBIS plays a key role in promoting and facilitating the sharing of marine biological observation data across the world.

Like OBIS, the GBIF is a data sharing and integration mechanism but it is for all taxa worldwide. Its formation was recommended by the Organization for Economic Cooperation and Development's Megascience Forum in 1999 and GBIF was officially established in 2001. Initially, GBIF focused on natural history collections (Guralnick et al., 2007) but has expanded its coverage to include observations from a range of projects and sampling methodologies.

Taken together, the organizations and projects described above represent a range of activities happening in the marine biological observing space. Efforts dedicated to collaboration and coordination among the entities involved can work to avert duplication of effort, ensure prudent use of resources, and encourage international cooperation.

**INFORMATION SHARING PRINCIPLES**

To build an integrated, end-to-end marine biological observing system, data must adhere to the FAIR data principles, which includes using standards for both data and metadata (Wilkinson et al., 2016), improving calibration and developing common understanding. Metadata must adhere to established standards like the EML and those provided by the International Organization for Standardization.

The Darwin Core standard (Wieczorek et al., 2012) used by OBIS, GBIF, and other national or international scale data aggregation projects like the Atlas of Living Australia, VertNet, and iDigBio, has become the most widely used standard for biological diversity data. Darwin Core was conceived to allow observations of species to be integrated across multiple sources, digitally accessible, and discoverable, such that existing data can be reused for broad-scale analysis (Wieczorek et al., 2012). It evolved from the Dublin Core standard and provides a simple and open set of terms for the exchange and integration of biological observation data (Wieczorek et al., 2012).

The Darwin Core standard provides a framework for building systems that are both syntactically and semantically interoperable. It provides syntax standards for data and file structure and format and a basis for communities to build common semantics through vocabularies and authorities in biological taxonomy and other aspects of the data model. The syntactic interoperability means that different systems and users can expect to find the data elements in a specific format (Darwin Core Archives, a set of CSV files with a manifest (meta.xml) specifying how the files are structured and the relationships between them Wieczorek et al., 2012) as well as data elements relying on the Darwin Core term definitions used to build the CSV files. Semantic interoperability ensures that the end user understands the meaning of each data element and the definitions have been vetted by the user community providing an authoritative resource. Syntactic and semantic interoperability are crucial pieces to effective data sharing (Harvey et al., 1999).

Implementing the FAIR Guiding Principles requires that digital objects be documented, peer reviewed, and managed like manuscripts. By adhering to these principles, we can accelerate the rate of scientific progress through data-intensive scientific inquiry at broad spatial and temporal scales. Similarly, analytical workflows and methods need to be documented and shared, along with the entire provenance trace (Reichman et al., 2011).

Reusing best practices in the form of technology and analytical workflows, as well as reusing data, is the paradigm that can help us maximize the value of data and produce global-and national-scale, policy-ready products. Working toward that goal, initiatives such as MBON (Muller-Karger et al., 2018), Smithsonian-led MarineGEO, Ocean Tracking Network, U.S. Animal Telemetry Network, and Long Term Ecological Research Network, are building communities of practice, but work remains in garnering support for data sharing and adhering to FAIR data principles.

**SOCIETAL FORCING FACTORS**

Strategically, the most straightforward and cost-effective ways to accomplish integration of data streams and developing useful products is by strengthening existing networks and programs and through improved international collaboration (Royal Society, 2011). This promotes scalability, from national to regional to global. Such efforts are currently underway, including, notably through an agreement between GOOS, MBON, and OBIS with corresponding information flow objectives (Figure 1).

Built from the Framework for Ocean Observing system model (Lindstrom et al., 2012), the GOOS-MBON-OBIS collaboration seeks to develop a Community of Practice that can inform policy makers and managers about the present and future state of global marine biodiversity. By leveraging systems that are already in place, we can increase the pace of knowledge building and dissemination.

Authors have suggested that the major barrier to sharing data, analytical tools, workflows, and algorithms in ecology is cultural, rather than technological (Jones et al., 2006; Reichman et al., 2011). Although we are seeing an increase in data sharing initiatives and policy mandates, many biological observations, including historical observations, may not be regularly shared (Tenopir et al., 2011; Fecher et al., 2015) or readily available. While attitudes toward data sharing are improving, the perceived risks and barriers can continue to impede progress (Tenopir et al., 2015). Moreover, broad societal investment through public support of monitoring biodiversity and ecological changes is also needed. The combination of these
FIGURE 1 | Information flow diagram as part of the GOOS-MBON-OBIS collaboration agreement, (http://iobis.org/2016/12/15/goosgeobonobis/), helping to work toward an international integrated system for marine biological observations. The aim of this collaboration is to build a globally coherent, consistent, and coordinated sustained global ocean observing system (GOOS) to assess the state of the ocean’s biological resources and ecosystems.

Two factors is essential to stimulate data sharing by the ecological community.

One solution to advancing FAIR data principles could be for funding agencies to be more prescriptive and supportive in their requirements for data standardization and sharing. Many funding agencies now require a data management plan to be created, however, many researchers feel they do not have the resources or capacity to meet this end (Tenopir et al., 2011). Funds and resources could be set aside specifically to facilitate the data management aspect of projects, and outcomes of the data management plan tracked by the funding agency. Otherwise, the plan can end up being more notional than tangible.

Fortunately, the data community is seeing an increase in institutional policy that supports FAIR practices. High level government memoranda like the United States (Office of Science and Technology Policy [OMB], 2013) open data policy have led to United States Government bureau guidance and policies like the U.S. Geological Survey Fundamental Science Practices survey manual requirements (United States Geological Survey [USGS], 2015) and the National Oceanic and Atmospheric Administration Environmental Data Management Committee Procedural Directives (National Oceanic and Atmospheric Administration [NOAA], 2015a,b,c, 2017). Similarly, international initiatives like the Open Data Charter which has been adopted by 54 national and local governments including Australia, Argentina, Mexico, and United Kingdom (Open Data Charter, 2015) are bolstering the FAIR principles.

Another pressing need, particularly to allow for more real-time management, is to reduce the time between data collection and data availability within open access data systems like OBIS and the GBIF. Achieving this demands the prioritization of developing technologies that lead to automation of measurements, quality assurance/quality control, cost reduction, and put machine-readability at the forefront. A proper data system will be centralized or at least interconnected within a user-friendly portal. This can then readily be accessed to inform policy, support management practices, answer scientific questions and fulfill international reporting obligations. Additionally, citizen science, when appropriately supervised and quality controlled, presents an opportunity to increase observations of marine biological diversity and processes while increasing societal investment in biological monitoring (Cigliano et al., 2015).

In short, an integrated system for marine biological observations and informatics requires an investment in capacity development at all levels, from data collection to data analysis to product creation and data distribution, storage, and maintenance.

BUILDING THE SYSTEM

Many pieces of an integrated marine biological observation and informatics system are already in place. Next, science
communities must build better connections between the pieces to ensure a much greater impact.

**Key Tenets for an Integrated Marine Biological Observing and Informatics System**

1. Data on a minimum set of essential variables collected using comparable methods across time and multiple sites (EOVs and EBVs);
2. Products and modeling informed by national and global reporting needs, and assessments and indicators to help guide the implementation of monitoring and protocols;
3. Data that fit the FAIR principles (findable, accessible, interoperable, and reusable);
4. Analytical algorithms, tools, and workflows that are shared and include provenance;
5. All pieces of the system adhering to the principles of action ecology.

To build an integrated system, we suggest following a cyclical architecture (Figure 2), as opposed to a straight line where the ends are disconnected. In this way, each piece is connected to and helps to inform the other parts of the system. Monitoring and protocols can be informed by the assessments, indicators, and modeling that would themselves be influenced by essential

**FIGURE 2** | An integrated system would allow for different pieces of the system to influence and guide one another.
variables. Problems of broad societal relevance can be addressed by monitoring programs that collect data used to answer those problems. Moreover, decision-makers require data and indicators to assist them with their reporting needs (Geijzendorffer et al., 2016). In this integrated system, data could be mobilized and standardized, as advocated by integrated species data systems such as OBIS and GBIF. Both of these systems use the Darwin Core Standard, so researchers and decision-makers can query across datasets collected at the local level to gain understanding at larger scales.

While these tenets are vital to the success of a global integrated marine observing system, we must be careful that they do not stifle innovation. Allowing guiding principles or the need for certain products to become too prescriptive will be suffocating to the energy and enthusiasm needed to tackle these broad scale societal problems. Research projects are a vital part of the biological ocean observing landscape and should continue exploring innovative questions and methodologies while implementing common standards when possible. As we see in Figure 2, a dynamic process enables the incorporation of new methodologies and ideas. Evidence of this process is visible in the newly added ocean sound EOV and the emerging microbe and benthic invertebrate EOVs. Similar processes could be advantageous for each piece of the integrated circular system; the system cannot be static.

**Standardized Data Collection**

When data across multiple sites are collected with shared protocols, researchers can more readily draw conclusions across ecosystems and spatial scales. For example, MarineGEO has been able to determine predation rates across international sites between different habitats because it has a relatively simple standardized protocol (Duffy et al., 2015). Monitoring and associated protocols will vary across the globe according to capacity to spend time and resources on monitoring efforts and on aligning protocols with best practices and FAIR principles. Wider availability of standardized protocols and the use of best practices (Pearlman et al., 2017; Muller-Karger et al., 2018) could help to mitigate some of these resource discrepancies.

Global Ocean Observing System has been instrumental in developing the EOVs that serve as guidance for data necessary to monitor changes in biodiversity (Miloslavich et al., 2018). Using EOVs, as well as EBVs developed by GEO, as high-level guides for products, modeling, and monitoring can help fill data gaps and reduce data redundancy by exposing areas of data deficiencies spatially, temporally, and taxonomically.

**Addressing Global Reporting Needs**

Policy-makers and resource managers use data, assessments, and indicators to report on and implement global policy initiatives (Geijzendorffer et al., 2016). Thus, considering those requirements is important when building global and national products and conducting monitoring. While this may seem intuitive, stakeholders beyond the scientific community may not be considered in the research process. The principles of action ecology (White et al., 2015) provide a framework for steps to take. In particular, designing projects that are transdisciplinary and therefore include a variety of stakeholders in the scientific process can help ensure global and national reporting needs are being met.

Globally integrated scientific advancement in pursuit of assessments and indicators relies on not only sharing data and analytical workflows but also on collecting the necessary data, as indicated by essential variables. If policy-makers and resource managers need answers to questions, but the data to answer those questions are not being collected, the entire system will not function. By using the EOVs and EBVs as a guide for global reporting needs, monitoring and the concomitant protocols can ensure desired data are available.

**Making Data FAIR**

To implement the FAIR guiding principles, ecological data need to be shared and in conformance with documented standards. The marine biological community needs to be able to find and access the data using search methods they are familiar with using both human- and machine-readable methods. To incorporate these data into products, indicators, or assessments, the data must be interoperable with data already on hand. Finally, the data need to be usable for current projects and reusable for projects yet to be determined.

Ocean Biogeographic Information System and GBIF serve as examples of data standardization and accessibility by requiring both data and metadata standards, Darwin Core and EML, respectively, serving data via human- and machine-readable methods, and integrating datasets from multiple sources.

A significant recent advance in making marine biological data more interoperable is OBIS's development of a new data schema inside the Darwin Core standard (Event-Data schema, De Pooter et al., 2017). This new schema allows the description of the sampling methodology, sampling apparatus, units, and any additional biological measurement or observation associated with each biodiversity occurrence record to be added to the integrated database. Using a standardized vocabulary, the user will have the capability to extract biodiversity records obtained with similar methodologies, making it more comparable at the global and national scale.

Another important component in making sure data are FAIR is assuring data, or at least metadata, are archived for future generations of scientists to discover and use. Utilizing permanent archives like the National Centers for Environmental Information can ensure that data are accessible for many years to come. In addition, data aggregators and data users should ensure adequate attribution back to data originators.

**Analytical Algorithms, Tools, and Workflows Are Accessible**

Catalyzing the progress of this global system goes beyond sharing data to sharing analytical algorithms, tools, and workflows and including provenance. Provenance is including the what, how, when, where, and why about the data and in analyses (Ma et al., 2014). Provenance allows end users to look at a product and trace the history of how it came to be and any of the decision points made in producing it. Sharing analytical algorithms, tools,
and workflows also has the benefit of improving transparency and the scientific process as well as credibility (Reichman et al., 2011; Ma et al., 2014). The intellectual capital contained in these types of products is unrealized potential for increasing the rate of scientific discovery.

**Adhering to Action Ecology Principles**

Inherent to all the work occurring in the realm of marine biological observing, is an integrated system of marine biological observations and informatics can contribute to policy development and management actions. To maintain healthy ecosystems, new scientific findings can be rapidly incorporated into assessments, indicators, and marine policy. Therefore, it is important to include a diverse, transdisciplinary stakeholder group at every stage of the integrated end-to-end system and especially during development of indicators and/or assessments and modeling. Making sure the system is technology-driven provides a means for engaging with the public, fostering societal investment, and catalyzing the progress of the marine biological and informatics system. Thus, adhering to the principles of action ecology ensures monitoring and marine observations are quickly incorporated in policy and management actions.

**CONCLUSION**

The scientific community is poised to respond to the needs of users of ocean information. Developing this information requires an integrated circular system of marine biological observations and informatics to support and strengthen understanding of change in marine ecosystems. The opportunity exists to strengthen existing programs and systems by increasing collaboration and coordination and leverage the outcomes to build comparable capacities worldwide. Using essential variables as a frame of reference assures that research efforts will be relevant to decision makers by implementing the principles of action ecology to drive their development. Finally, researchers can increase the speed of scientific discovery and incorporation in decisions by sharing all parts of the scientific process including data and workflows using standards and best practices and demanding they are FAIR.

**AUTHOR CONTRIBUTIONS**

The idea for the manuscript was invited by the editor. AB led the writing of the manuscript in close collaboration with CB. All authors contributed critically to the drafts and gave final approval for publication.

**ACKNOWLEDGMENTS**

We thank T. Malone and A. Knap for the invitation to contribute our ideas to this topic. We also thank the two reviewers and editor for their comments, which strengthened our manuscript.

**REFERENCES**


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Benson, Brooks, Canonicos, Duffy, Muller-Karger, Sosik, Miloslavich and Klein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.