University of Wollongong

Research Online

2017

A three-dimensional mapping of the ocean based on environmental data

Roger G. Sayre
United States Geological Survey

Dawn J. Wright
Esri

Sean P. Breyer
Esri

Kevin A. Butler
Esri

Keith Van Graafeiland
Esri

See next page for additional authors

Follow this and additional works at: https://ro.uow.edu.au/eispapers1

A three-dimensional mapping of the ocean based on environmental data

Abstract

The existence, sources, distribution, circulation, and physicochemical nature of macroscale oceanic water bodies have long been a focus of oceanographic inquiry. Building on that work, this paper describes an objectively derived and globally comprehensive set of 37 distinct volumetric region units, called ecological marine units (EMUs). They are constructed on a regularly spaced ocean point-mesh grid, from sea surface to seafloor, and attributed with data from the 2013 World Ocean Atlas version 2. The point attribute data are the means of the decadal averages from a 57-year climatology of six physical and chemical environment parameters (temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate). The database includes over 52 million points that depict the global ocean in x, y, and z dimensions. The point data were statistically clustered to define the 37 EMUs, which represent physically and chemically distinct water volumes based on spatial variation in the six marine environmental characteristics used. The aspatial clustering to produce the 37 EMUs did not include point location or depth as a determinant, yet strong geographic and vertical separation was observed. Twenty-two of the 37 EMUs are globally or regionally extensive, and account for 99% of the ocean volume, while the remaining 15 are smaller and shallower, and occur around coastal features. We assessed the vertical distribution of EMUs in the water column and placed them into classical depth zones representing epipelagic (0 m to 200 m), mesopelagic (200 m to 1,000 m), bathypelagic (1,000 m to 4,000 m) and abyssopelagic (>4,000 m) layers. The mapping and characterization of the EMUs represent a new spatial framework for organizing and understanding the physical, chemical, and ultimately biological properties and processes of oceanic water bodies. The EMUs are an initial objective partitioning of the ocean using long-term historical average data, and could be extended in the future by adding new classification variables and by introducing functionality to develop time-specific EMU distribution maps. The EMUs are an open-access resource, and as both a standardized geographic framework and a baseline physicochemical characterization of the oceanic environment, they are intended to be useful for disturbance assessments, ecosystem accounting exercises, conservation priority setting, and marine protected area network design, along with other research and management applications.

Publication Details

Authors

Roger G. Sayre, Dawn J. Wright, Sean P. Breyer, Kevin A. Butler, Keith Van Graafeiland, Mark J. Costello, Peter T. Harris, Kathleen L. Goodin, John M. Guinotte, Zeenatul Basher, Maria T. Kavanaugh, Patrick N. Halpin, Mark E. Monaco, Noel A. Cressie, Peter Aniello, Charles E. Frye, and Drew Stephens

A Three-Dimensional Mapping of the Ocean Based on Environmental Data

By Roger G. Sayre, Dawn J. Wright, Sean P. Breyer, Kevin A. Butler, Keith Van Graafeiland, Mark J. Costello, Peter T. Harris, Kathleen L. Goodin, John M. Guinotte, Zeenatul Basher, Maria T. Kavanaugh, Patrick N. Halpin, Mark E. Monaco, Noel Cressie, Peter Aniello, Charles E. Frye, and Drew Stephens

ABSTRACT. The existence, sources, distribution, circulation, and physicochemical nature of macroscale oceanic water bodies have long been a focus of oceanographic inquiry. Building on that work, this paper describes an objectively derived and globally comprehensive set of 37 distinct volumetric region units, called ecological marine units (EMUs). They are constructed on a regularly spaced ocean point-mesh grid, from sea surface to seafloor, and attributed with data from the 2013 World Ocean Atlas version 2. The point attribute data are the means of the decadal averages from a 57-year climatology of six physical and chemical environment parameters (temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate). The database includes over 52 million points that depict the global ocean in x, y, and z dimensions. The point data were statistically clustered to define the 37 EMUs, which represent physically and chemically distinct water volumes based on spatial variation in the six marine environmental characteristics used. The aspatial clustering to produce the 37 EMUs did not include point location or depth as a determinant, yet strong geographic and vertical separation was observed. Twenty-two of the 37 EMUs are globally or regionally extensive, and account for 99% of the ocean volume, while the remaining 15 are smaller and shallower, and occur around coastal features. We assessed the vertical distribution of EMUs in the water column and placed them into classical depth zones representing epipelagic (0 m to 200 m), mesopelagic (200 m to 1,000 m), bathypelagic (1,000 m to 4,000 m) and abyssopelagic (>4,000 m) layers. The mapping and characterization of the EMUs represent a new spatial framework for organizing and understanding the physical, chemical, and ultimately biological properties and processes of oceanic water bodies. The EMUs are an initial objective partitioning of the ocean using long-term historical average data, and could be extended in the future by adding new classification variables and by introducing functionality to develop time-specific EMU distribution maps. The EMUs are an open-access resource, and as both a standardized geographic framework and a baseline physicochemical characterization of the oceanic environment, they are intended to be useful for disturbance assessments, ecosystem accounting exercises, conservation priority setting, and marine protected area network design, along with other research and management applications.

INTRODUCTION

Ecosystems are a central focus of many current research and policy questions, including: (1) What are the impacts of climate change on ecosystems? (2) Which ecosystems are vulnerable to climate and other perturbations (e.g., invasive species, land and sea use)? (3) Which ecosystems should be targeted for conservation? (4) What are the economic and social values of ecosystem goods and services? and (5) What role do ecosystems play in global food and environmental security

(Liu et al., 2007, 2015)? Fundamental knowledge of the types and locations of global ecosystems is necessary to address these questions, yet that knowledge is generally lacking.

The development of a new global ecosystems map, including terrestrial, freshwater, and marine domains, was therefore commissioned by the intergovernmental Group on Earth Observations (GEO, https://www.earthobservations.org), a consortium of over 100 nations seeking to advance Earth observation approaches for addressing societal challenges related to hazards, food, water, energy, and the environment. The new global ecosystems maps were to be derived from data rather than from expert opinion or socio-political considerations, and they were to be based on the physical environmental

features that are understood to influence the distribution of species.

Ecosystems are geographically identifiable areas where the interactions of organisms with their physical environments produce differences in biotic diversity, trophic structure, and flows of energy and materials between living and nonliving components of the system (Odum, 1971). On land, variations in climate, landform, and substrate establish the environmental potential that controls primary production and species distributions, acknowledging that evolutionary history is also an important element of biogeography (Bailey, 1996, 2014; Kottek et al., 2006; Holt et al., 2013). Responding to the GEO commission request for a standardized, robust, and practical global map of terrestrial ecosystems, a new map of global ecological land units (ELUs) was developed from an integration of climate, landform, lithology, and land cover (Sayre et al., 2014). We now present a similar environmental stratification approach for extending the global ecological units map into the ocean through the delineation of global ecological marine units (EMUs).

There are notable differences between mapping terrestrial and mapping marine ecological units. First, the terrestrial ELUs were mapped as two-dimensional (2D) entities using a raster data surface. Marine ecosystems, however, are fundamentally understood as both 2D (e.g., sea surface and seafloor) and three-dimensional (3D; e.g., water column) entities (e.g., Li and Gold, 2004; Wright et al., 2007). EMUs would ideally need to be mapped using 3D data points representing a volumetric mesh (e.g., Heinzer et al., 2012; Reygondeau et al., 2017) and visualized as 2D and 3D objects. Second, the characteristics of the physical environment that influence the distribution of species and ecosystems are different between terrestrial and marine environments. While the ELUs were identified as distinct combinations of bioclimate, landform, lithology, and vegetation, those

elements of terrestrial ecosystem structure do not apply as such in the ocean, with the exception of seafloor landforms. The abiotic controls on the distribution of marine biota (e.g., temperature, as in Beaugrand et al., 2013), and nutrients, as in Longhurst, 2007) were identified as analogs to the terrestrial environmental characteristics. Third, the ocean is a fluid in motion, and ocean ecosystem conditions can be influenced by processes far from their location via ocean circulation. Marine ecosystems are therefore generally understood as more spatiotemporally dynamic than their terrestrial counterparts.

Considerable research in quantitative water mass analysis has produced a robust and standardized terminology for describing water masses with respect to their origins, properties, and locations (Tomczak, 1999). The ocean's hydrographic structure has been described with evolving complexity, beginning with the seminal work of Sverdrup (1942) and including the early delineation of global water masses by Emery and Meincke (1986) based on temperature and salinity properties. Expanding on the temperature and salinity relationships, more recent quantitative water mass characterizations are multidimensional and feature the use of tracer distributions to identify water mass origins. For example, Gebbie and Huybers (2011) comprehensively identified the surface origin of points in the ocean interior at 33 depth levels using climatological and isotope ratio data and tracer path analysis.

In addition to origin-based quantitative water mass analysis, another common approach to identifying oceanic water bodies involves subdivision of ocean regions or volumes based on differences in their physical, chemical, and biological properties (Sherman et al., 2005; Longhurst, 2007; Spalding et al., 2007; Reygondeau et al., 2017). While various marine oceanic region maps exist (Table 1), few are global in extent, are representative of the entire water column in three dimensions, and were derived from quantitative analysis of data. Three-dimensional, globally comprehensive subdivisions of the ocean are particularly lacking. The quantitative, 3D analysis of Gebbie and Huybers (2010, 2011) identified seven surface water source regions of global ocean water, but did not map different regions by depth. Reygondeau et al. (2017) objectively subdivided the Mediterranean Sea into 63 biogeochemical regions in a true 3D analysis using data and biologically meaningful criteria to separate the water column vertically into epipelagic, mesopelagic, bathypelagic, and seafloor zones. However, there has not been a purely quantitative, unsupervised approach to partitioning the entire global ocean water column from aspatial statistical clustering of global ocean data, even though an unprecedented amount of global marine environmental data is now available (e.g., the 2013 World Ocean Atlas of Locarnini et al., 2013; Zweng et al., 2013; Garcia et al., 2014a, 2014b).

We approached the challenge of aggregating comprehensive marine environmental data through statistical clustering, building on the efforts of previous authors. For example, Harris and Whiteway (2009) used a multivariate statistical method with six biophysical variables (depth, seabed slope, sediment thickness, primary production, bottom-water dissolved oxygen, and bottom temperature)

TABLE 1. Existing maps of ocean regions.

NAME	GEOGRAPHIC SCOPE	BASIS
Countries' Exclusive Economic Zones (EEZs) (United Nations, 1982)	Global, Coastal	Political
Large Marine Ecosystems (LMEs) (Sherman et al., 2005)	Global, Coastal	Management areas
Marine Ecoregions of the World (MEOWs) (Spalding et al., 2007)	Global, Coastal	Expert-derived biogeography (realms, provinces) and management units (ecoregions)
Fisheries and Agricultural Organization (FAO) Major Fishing Areas (FAO, 2016)	Global	Rectangular fishery statistical assessment regions
International Council for the Exploration of the Sea Ecoregions (ICES, 2004)	Regional (Northeast Atlantic)	Large ecosystem and fishery management areas
International Hydrographic Organization Seas and Oceans (IHO, 2002)	Global	Geographically named areas
Ecoregions of the Oceans and Continents (Bailey, 2014)	Global	Expert recommended regions
Global Open Ocean and Deep Seabed (GOODS) Biogeographic Characterization (UNESCO, 2009)	Global, Benthic and Pelagic	Expert recommended regions
Deep-Sea Provinces (Watling et al., 2013)	Global, Benthic	Expert-derived revision of GOODS based on literature review
Biogeochemical Provinces (Longhurst, 2007)	Global	Satellite ocean color
Seafloor Map (GSFM) (Harris et al., 2014)	Global	Expert geomorphological feature extraction using 30 arc-second bathymetry data
Deep-Sea Seascapes Map (Harris and Whiteway, 2009)	Global	Multivariate analysis of seabed morphology and sediments

to objectively classify the entire ocean floor into 53,713 separate polygons comprising 11 different categories. The 11 categories had mean polygon sizes ranging from 1,000 km² to 22,000 km² and were restricted to the seafloor. Reygondeau et al. (2017) statistically clustered data from the Mediterranean Sea into distinct biogeochemical regions within biologically meaningful, predetermined depth zones. To our knowledge, the present study is the first to objectively classify the entire global ocean water column simultaneously across all depths into discrete regions based on comprehensive statistical clustering of physical and chemical environmental data from all points in the World Ocean Atlas (WOA)-derived ocean mesh.

Most oceanography and marine biology textbooks include diagrams that divide the ocean into depth zones (e.g., Figure 1). Shallower, sunlit depths at or near the surface are usually presented as the photic layer, which extends to the general limit of light penetration (99% of incident light) at a depth of about 200 m (Stal, 2016). Beneath this depth, photosynthesis is largely lacking (Costello and Breyer, 2017). This same depth zone to 200 m is also commonly referred to as the epipelagic zone. Beneath this zone at depths commonly understood as between 200 m and 1,000 m is the mesopelagic zone, where organismal respiration is higher relative to deeper areas (Costello and Breyer, 2017). Although the 1,000 m depth boundary is arbitrarily defined, Proud et al. (2017) objectively subdivided the mesopelagic region into distinct subzones using organismal echolocation data, and Reygondeau et al. (2017) use flux of particulate organic carbon to determine biologically meaningful boundaries for the mesopelagic layer. Deeper, darker, and colder zones are usually presented as bathyl, abyssal, and hadal zones. Although depth boundaries for these regions are largely arbitrary and can vary from text to text, they are meant to describe the biogeochemical variation that is correlated with depth. It is

possible that as depth changes, variation in temperature and chemical composition creates distinct ecological zones represented by different ecological communities. However, this concept has never been objectively tested using data at a global scale. With few exceptions (Oliver and Irwin, 2008; Hardman-Mountford et al., 2009; Harris and Whiteway, 2009; Kavanaugh et al., 2014; Schoch et al., 2014; Reygondeau et al., 2017), most existing marine maps and zonation systems are derived from supervised classification and thus are influenced by the perspectives of their authors (Costello, 2009). To test whether recognizable boundaries exist vertically and horizontally in the global ocean, we clustered 3D ocean cells into groups using an unsupervised classification of physical and chemical environmental variables.

METHODS

We used the complete set of variables from the 2013 World Ocean Atlas data set, version 2 (Locarnini et al., 2013; Zweng et al., 2013; Garcia et al., 2014a, 2014b) as our source of physical and chemical environmental data for defining the ocean mesh and subsequently modeling the ecological marine units. The WOA

data set is a compendium of data from a variety of ocean research and monitoring programs over the past five decades. It is an authoritative 57-year climatology that contains over 52 million points, hereafter referred to as the ocean mesh. Each point is attributed with values for temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate, and all WOA values are corrected for the effect of pressure on each variable. The WOA has a horizontal spatial resolution of ¼° × ¼° for temperature and salinity, and 1° × 1° for oxygen, nitrate, phosphate, and silicate. In the vertical dimension, points are located at variable depth intervals, ranging from 5 m increments near the surface to 100 m increments at depth. A total of 102 depth zones extend to 5,500 m. The depth intervals are as follows: 5 m (from 0 m to 100 m), 25 m (from 100 m to 500 m), 50 m (from 500 m to 2,000 m), and 100 m (from 2,000 m to 5,500 m). The deepest points for which data are available do not necessarily represent the actual depth of the water column because the 5,500 m lower limit of the WOA data is approximately half of the maximum depth of the ocean (Jamieson, 2011). However, the 5,500 m lower limit does substantially exceed the mean depth (3,682.2 m)
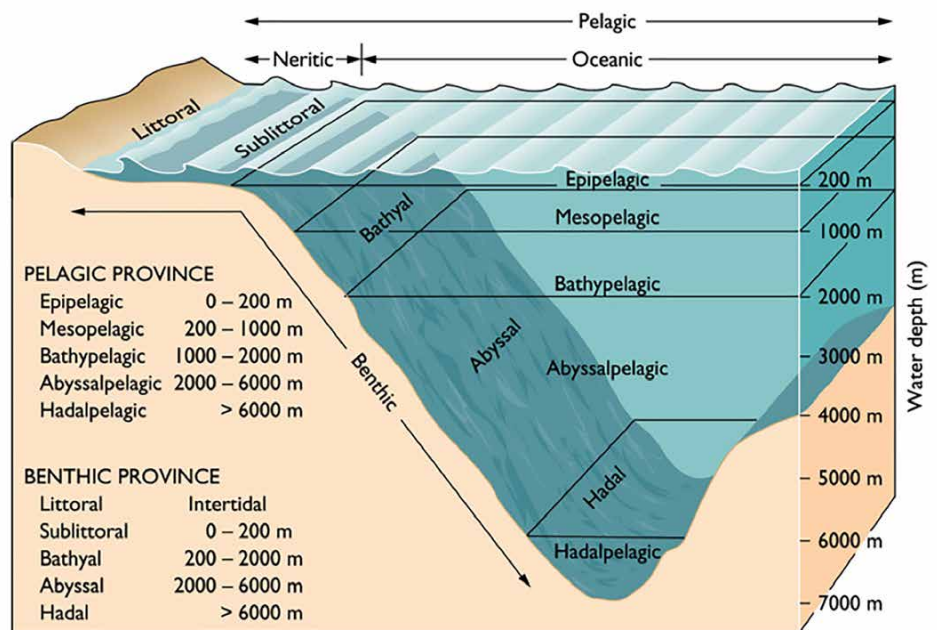

FIGURE 1. Traditional oceanographic notions of vertical zonation in the ocean. *Modified from Pinet (2009).*

of the ocean as reported in the review by Charette and Smith (2010). The WOA data are water-column variables, but seafloor geomorphology may also be significant in influencing both these variables and species ecology. The first global digital map of seafloor geomorphic features is now available (Harris et al., 2014).

Temporally, the WOA archive is available in seasonal, annual, and decadal resolutions. Seasonal data are not available for all points in the mesh, many of which may not have been visited regularly over the 57-year period. Moreover, data from polar regions, typically collected only during warmer summer months when access to ice-bound regions is easier, may under-report true salinity values. Decadal values of the WOA represent the average of the annual mean values for the parameters, themselves derived from the seasonal data. We used the 57-year record of the parameters, which are provided in the WOA database as archival means, derived from the decadal averages. The modeled EMUs therefore represent average distributions of the volumetric regions over the past 57 years.

We constructed an ocean point mesh as a 3D spatial data structure that holds the WOA data in its highest available spatial resolution of ¼° × ¼° (~27 km × 27 km at the equator) in the horizontal dimension. While the temperature and salinity data are available at this resolution, the other four variables (dissolved oxygen, nitrate, phosphate, and silicate) have a coarser native resolution (1° × 1°) and were therefore downscaled to the ¼° resolution to reconcile all data to a common working horizontal resolution. This downsampling was accomplished by subdividing the 1° × 1° by depth-interval rectangular box cuboid into sixteen ¼° × ¼° by depth-interval cuboids and assigning the original attribute values of the parent cuboid's centroid to the center points of all of the ¼° subdivisions. In this piecewise-constant re-meshing, we assume that the attributes of the parent cuboid are uniform throughout the cuboid's volume. This is similar to the universal assumption in vector-based GIS that the attributes of a vector polygon are uniform throughout the polygon's extent. Statistical and nonstatistical

downscaling of coarser-resolution data such as global climate model (GCM) data to finer-resolution data is a common practice in global change modeling (Hall, 2014), and it is also the basis for pansharpening of multiresolution imagery (Vivone et al., 2015).

The data matrix can be conceptualized as columnar stacks of cells whose centroids define the point mesh (Figure 2). In areas where the deepest (5,500 m) WOA data points did not reach the seafloor, the bottom of the mesh was simply extended downward to the seafloor for visualization, without interpolating additional data points. The mesh spacing matched the WOA data matrix and allowed for the structuring and symbolization of data as columnar volumes (or other shapes) that can be queried by ranges of values, and can be spatially analyzed via proximity algorithms and multivariate statistical clustering. We first constructed an "empty" ocean mesh using the 52,487,233 WOA point locations, and then attached the WOA attribute data to those points. The water column was bounded at the top by the sea surface, and at the bottom by the seafloor, as defined in the geomorphic map of Harris et al. (2014). The set of cells intersecting or nearest to the global shoreline (or ice masses) defined the horizontal extent of the water column.

We statistically clustered the points in the mesh in order to identify environmentally distinct regions in the water column. The clustering was blind to both the depth of the point and the thickness of the depth interval at that point's vertical position in the water column. The "big data" nature of the clustering of the entire ocean volume required sophisticated spatial data processing and functionality. The clustering was implemented using SAS software (©2015 SAS Institute Inc; SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA). The ArcGIS platform was utilized for subsequent geospatial assessment and visualization of
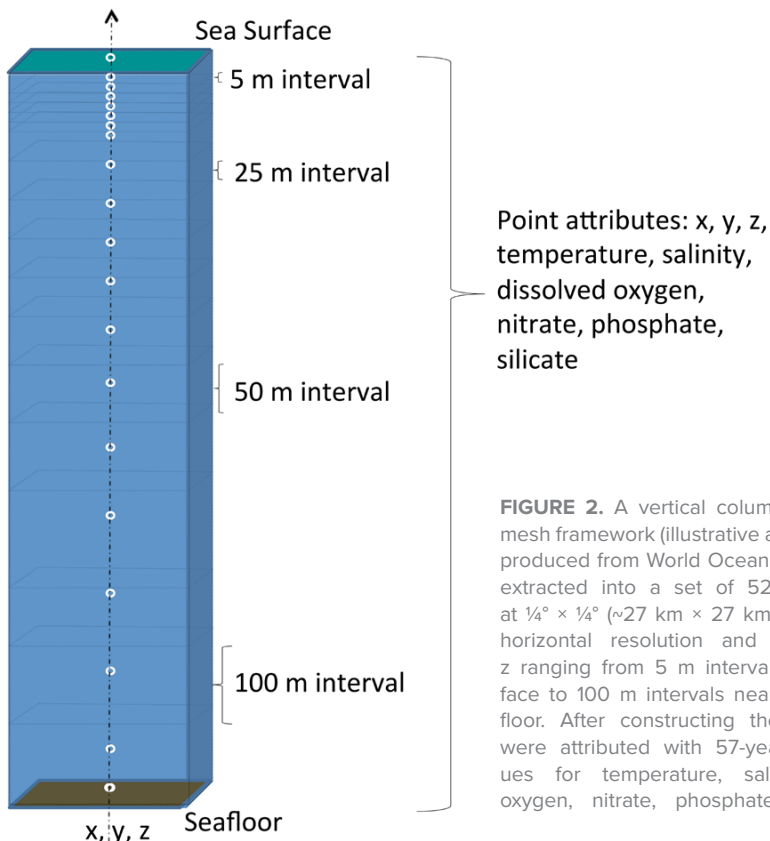
FIGURE 2. A vertical column of the ocean mesh framework (illustrative and not to scale), produced from World Ocean Atlas 2013 data extracted into a set of 52,487,233 points at ¼° × ¼° (~27 km × 27 km at the equator) horizontal resolution and variable depth z ranging from 5 m intervals near the surface to 100 m intervals near the deep seafloor. After constructing the mesh, points were attributed with 57-year average values for temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate.

the clusters. We utilized a *k*-means clustering algorithm to identify the physical and chemical structure of the water column. The *k*-means algorithm determines *k* centroids in the data and clusters points by assigning them to the nearest centroid. Of hundreds of clustering algorithms available, the *k*-means approach is the most widely used due to its simplicity, versatility, extensibility, data handling ability, and generally robust performance (Jain, 2010), although it is sensitive to initial placement of cluster centers (Celebi et al., 2013). While concurrent implementation and integration of complementary clustering approaches has been advocated for ocean partitioning (Oliver et al., 2004; Reygondeau et al., 2017), this multi-algorithm approach was outside the scope of our globally comprehensive and data intensive analysis.

Our statistical approach was prototyped on a subset (97,329 points) of the global point mesh representing the ocean volume off the US West Coast out to the Exclusive Economic Zone (EEZ). The successful identification of known hydrographic features (e.g., the Mendocino Ridge and Fracture Zone off the northern California coast) in the prototype exercise provided initial assurances that the clustering approach would be sensitive to environmental gradients, and that scaling up to global clustering was warranted. We therefore implemented the clustering globally on all cells (>52 million points), with all variables included.

All the WOA variables were standardized to a mean of zero and a standard deviation of one to establish a common basis for comparison between variables of disparate units and value ranges and to promote relative equal weightings of the inputs to the clustering (Milligan and Cooper, 1988). After standardization, a Pearson's correlation analysis of the six inputs was implemented to identify colinearity among variables. To determine the optimal number of clusters that would best represent the collective variation in the input data, clustering of all the WOA points with all

six variables was executed in repeated sequential runs, where the number of clusters produced was incremented by one with each run, starting with two clusters, and ending with 100 clusters. The optimum cluster number was determined by inspection of the behavior of the pseudo F-statistic (Calinski and Harabasz, 1974; Milligan and Cooper, 1985) across the iterations. The pseudo-F statistic is the ratio of between-cluster variance to within-cluster variance. Larger values of pseudo-F indicate "tight" (i.e., low within-cluster variance) and "well separated" (i.e., high between-cluster variance) clusters. A plot of this statistic against the number of clusters should show local peaks of the pseudo-F value at potential cluster-number optima. We did not extend the clustering beyond 100 clusters because there was a clear overall decline in pseudo-F values as the number of clusters increased. Local peaks representing relatively high pseudo-F values were found at 17, 28, 37, and 50 clusters (Figure 3). We explored summary statistics and the horizontal and vertical spatial distributions at each of the four local peaks. At 37 clusters, a strong peak was observed prior to a relatively

sustained decline in the pseudo-F curve (Figure 3), which we interpreted as a point where additional clustering is less likely to reduce the within-cluster variation. Thus, the 37-cluster solution was the basis for our partitioning of the water column, and resulted in the 37 EMUs described below.

Following the depth-blind statistical clustering, basic descriptive statistics (mean, minimum, maximum, and standard deviation) were produced for the six deterministic parameters (temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate) for each EMU. The unit-middle depth for each EMU was also calculated as the median depth of all the points allocated into an EMU.

We then labeled the clusters using the naming criteria (Table 2) of the Coastal and Marine Ecosystem Classification Standard (CMECS), a Federal Geographic Data Committee standard for the United States (FGDC, 2012). The CMECS labels for the EMUs begin with their depth zone assignments based on their median depths, followed by a concatenation of the CMECS descriptors for temperature, salinity, and dissolved oxygen. The CMECS framework does not include
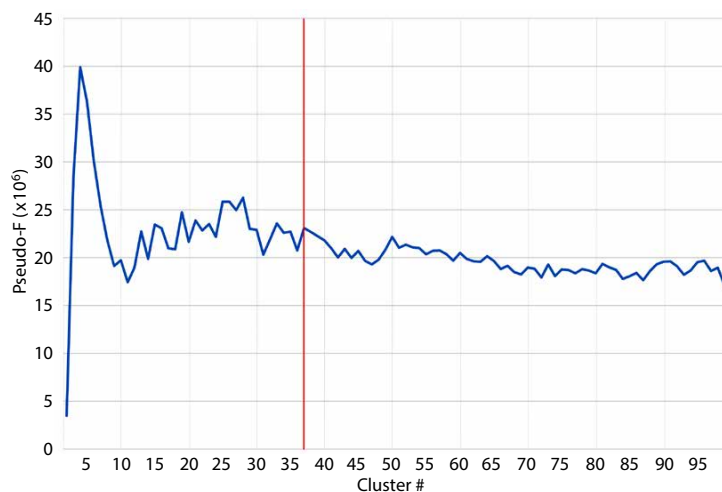
FIGURE 3. A plot of the pseudo-F statistic (y axis) against the requested number of clusters (x axis) in successive iterations from 2 to 100 clusters, incremented by one for each successive iteration. The red vertical line at 37 clusters shows a strong peak prior to a relatively sustained decline in the curve of the pseudo-F statistic, which we interpret as a stopping-point where additional clustering does not significantly reduce within-cluster heterogeneity (Calinski and Harabasz, 1974; Milligan and Cooper, 1985). We therefore chose the 37-cluster solution to represent the number and distributions of global EMUs.

TABLE 2. Depth and physicochemical properties (column 1) and corresponding depth zones and regime units of the Coastal and Marine Ecosystem Classification Standard (CMECS; FGDC, 2012; column 2) and of Ecological Marine Units (EMUs; column 3). CMECS regimes do not exist for nitrate, phosphate, and silicate. They were therefore adopted from EMU terms developed for these three variables, based on assessment and classification of the approximately 52 million observations for each nutrient into three relative classes (high, medium, and low).

DEPTH (m)	CMECS MARINE OCEANIC WATER COLUMN LAYER	EMU WATER COLUMN LAYER
0 to <200	Epipelagic	Shallow
200 to <1,000	Mesopelagic	Moderate Depth
1,000 to <4,000	Bathypelagic	Deep
≥4,000	Abyssopelagic	Very Deep

TEMPERATURE (°C)	CMECS TEMPERATURE REGIME	EMU TEMPERATURE REGIME
20 to <30	Warm to Very Warm	Warm
10 to <20	Moderate to Cool	Cool
5 to <10	Cold	Cold
0 to <5	Very Cold	Very Cold
≤0	Frozen/Superchilled	Superchilled

SALINITY (dimensionless)	CMECS SALINITY REGIME	EMU SALINITY REGIME
>30	Euhaline to Hyperhaline	Normal Salinity
18 to 30	Lower to Upper Polyhaline	Low Salinity
<18	Oligohaline to Mesohaline	Very Low Salinity

DISSOLVED OXYGEN (ml/L)	CMECS OXYGEN REGIME	EMU OXYGEN REGIME
≥8	Highly Oxic to Very Oxic	High Oxygen
4 to <8	Oxic	Moderate Oxygen
2 to <4	Hypoxic	Low Oxygen
0.1 to <2	Severely Hypoxic	Very Low Oxygen
<0.1	Anoxic	No Oxygen

NITRATE (μM)	CMECS NITROGEN REGIME (adopted from next column)	EMU NITROGEN REGIME
>30	High Nitrate	High Nitrate
10 to 30	Medium Nitrate	Medium Nitrate
<10	Low Nitrate	Low Nitrate

PHOSPHATE (μM)	CMECS PHOSPHATE REGIME (adopted from next column)	EMU PHOSPHATE REGIME
>5	High Phosphate	High Phosphate
2.5 to 5	Medium Phosphate	Medium Phosphate
<2.5	Low Phosphate	Low Phosphate

SILICATE (μM)	CMECS SILICATE REGIME (adopted from next column)	EMU SILICATE REGIME
>100	High Silicate	High Silicate
50 to 100	Medium Silicate	Medium Silicate
<50	Low Silicate	Low Silicate

standard names and value ranges for nitrate, phosphate, and silicate. For these three variables, labels corresponding to high, medium, and low nutrient concentrations in a relative sense were determined from assessment of the observed distribution of values. These three nutrient descriptors were added to the four CMECS descriptors for a total of seven descriptors in the CMECS names. The sequence of presentation of the descriptors used in the CMECS names is: depth, temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate. While depth was not used as a clustering variable, it is a key descriptor in the CMECS classification and an important variable for considering ocean depth zones, and was therefore included in the CMECS labeling. As an illustrative example of the CMECS nomenclature, an EMU cluster might be named Bathypelagic, Very Cold, Euhaline, Severely Hypoxic, High Nitrate, Medium Phosphate, and Medium Silicate.

We then developed a separate and parallel label, the EMU name (Table 2), to simplify the CMECS terminology. The EMU equivalent of the CMECS cluster described above is Deep, Very Cold, Normal Salinity, Very Low Oxygen, High Nitrate, Medium Phosphate, and Medium Silicate. The CMECS and simplified EMU names reflect the properties of the EMU, not its location in the ocean. Naming the EMUs based on their chemical and physical properties is both accurate and "classification neutral" in the sense that the label is purely descriptive in a compositional sense (Sayre et al., 2014).

Finally, in addition to the CMECS and EMU compositional names, we also developed a set of EMU volumetric region names that describe both their geographic distributions in the ocean and their vertical positions in the water column. In the field of quantitative water mass analysis, water masses are generally associated with or defined by their formation regions (Tomczak, 1999; Emery, 2001) or surface water sources (Gebbie and Huybers, 2010, 2011). Because we

have not identified the source geographies for our EMUs, we avoid calling them water masses herein. We instead describe the total volumetric distribution of an EMU as a volumetric region rather than as a water mass. Each EMU volumetric region name contains both a geographic descriptor and a CMECS depth zone class (epipelagic, mesopelagic, bathypelagic, and abyssopelagic) based on the median depth of the EMU. Examples of EMU volumetric regions include Antarctic and Subantarctic Bathypelagic, Mediterranean and Red Seas Mesopelagic, and Arctic and Labrador Sea Epipelagic.

RESULTS

The strength of the relationship between the standardized variables and the resulting EMU configurations was either strong (>90 to <95% confidence) or statistically significant (≥ 95% confidence) for each of the six input variables: temperature ($R^2 = 0.95$), salinity ($R^2 = 0.97$), dissolved oxygen ($R^2 = 0.91$), nitrate ($R^2 = 0.96$), phosphate ($R^2 = 0.96$), and silicate ($R^2 = 0.94$). A strong correlation (>.8) among the three nutrient inputs (nitrate, phosphate, silicate) was observed, suggesting that they may have had a slightly disproportionate influence on the clustering. However, we did not remove any of the nutrient variables and then re-cluster in order to ensure that regional variation in any of the six input variables could influence the clustering outcome. The ratio of between-cluster variance to within-cluster variance ($R^2/(1 - R^2)$) was: temperature, 17.47; salinity, 12.89; dissolved oxygen, 10.72; nitrate, 24.42; phosphate, 22.11; and silicate, 15.12. These results indicate that the six parameters contributed strongly and approximately equally to the identification of the clusters.

Clustering the entire set of points in the global mesh yielded 37 mutually exclusive clusters (EMUs) that are volumetric regions of relative compositional homogeneity. For a listing of EMUs described using CMECS and EMU terminology, and the names of the EMU volumetric regions, see Appendix 1 (available in the online supplementary materials). Maps of the EMUs, along with descriptive statistics on their environmental characteristics, are found in Appendix 2 (also available in the online supplementary materials).

The EMUs are presented in a number of 2D slices at several depths in Figure 4. A plot of EMU area against depth, shown in Figure 5, characterizes the vertical position of the EMUs in the water column and the depth at which the maximum and minimum horizontal distributions occur can be easily visually interpreted. This graph also shows the vertical distribution of the EMUs with respect to the CMECS boundaries for epipelagic (0 m to 200 m), mesopelagic (200 m to 1,000 m), bathypelagic (1,000 m to 4000 m), and abyssopelagic (>4,000 m) zones. The number labels in each EMU represent the EMU number for cross-referencing to the EMU names and maps in Appendices 1 and 2. These EMU number labels have been placed vertically at a depth corresponding to the median unit-middle of the EMU.

The global maps of the EMUs listed in Appendix 2 show the maximum global horizontal extent of the clusters looking vertically from above, as well as the
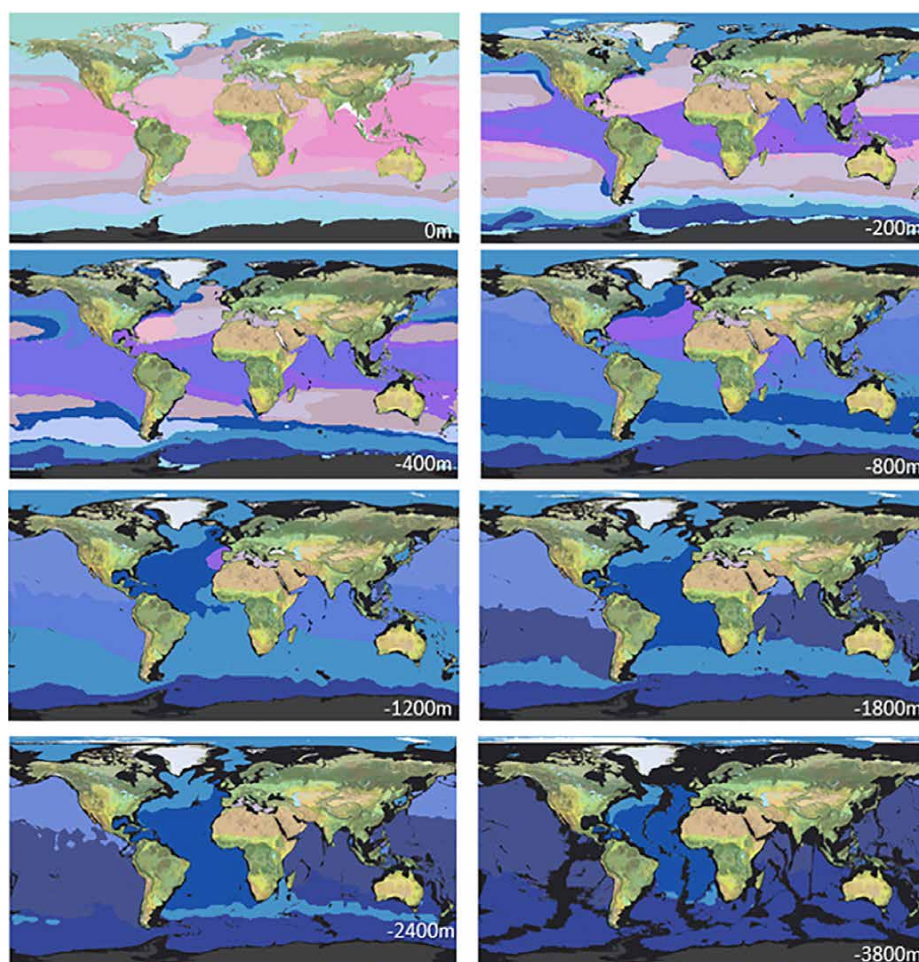

FIGURE 4. The global distribution of EMUs at eight depth intervals. EMUs represent physically and chemically distinct volumetric regions based on combined temperature, salinity, oxygen, and nutrient gradients. While a total of 37 EMUs were statistically determined, a number of them are small, localized, and shallow, and are not discernible in these depth-layer maps. Black indicates regions shallower than the depth at that layer. Major hydrographic features like Northern and Southern Hemisphere gyre systems and coastal upwelling-based westward flow of water from western continental margins are evident, particularly at shallower depths (upper left and right panels). Colors reflect mean EMU temperatures, with pink colors representing warmer EMUs and blue colors representing colder EMUs.

thickness of the EMU at any location. Interacting visually with true volumes in 3D, especially with such a large resource, is a software challenge, but it is possible with commercially available virtual globe-based visualization software, as shown in Figure 6. Although the EMUs are continuous data surfaces, Figure 6 shows that they are more easily visualized as stacked bands on horizontally separated cylinders.

The size and complexity of the EMU resource, a de facto example of big data with over 52 million attributed points in three dimensions, potentially presents barriers to its efficient use. To help mitigate this challenge, we developed an open-access web application, EMU Explorer (http://livingatlas.arcgis.com/emu) that permits real-time query and visualization of both the points and the EMUs by anyone with Internet access. The application shows the vertical profile

of the water column from sea surface to seafloor at any user-selected surface location, and it returns the values for the physical and chemical properties of the points in the column. It also identifies the EMUs associated with any point in the vertical profile and provides the EMUs' descriptive statistics.

DISCUSSION
EMU Geographic Distributions

Twenty-two of the EMUs are large, with essentially global or large regional distributions, while the 15 others are small, shallow, and coastal, and collectively represent only about 1% of the ocean volume. They generally have lower salinities than the other EMUs, and are found where mixing of fresh and saline waters is occurring (e.g., the Baltic Sea and northern, ice-occurring regions). While suitable for global-scale stratification of the

open ocean into large volumetric regions, the ¼° horizontal spatial resolution of the ocean mesh may not be sufficient to completely resolve the finer resolution, ecologically meaningful coastal systems. We therefore consider these very small, yet statistically derived coastal clusters as likely indicators of coastal and estuarine EMUs that need to be further clarified.

Latitudinal distribution patterns of EMUs are observed (Figure 4 and Appendix 2), with EMUs occurring in latitudinal biomes that include polar and subpolar regions (e.g., EMUs 14, 19, 23, 25 and 31), temperate regions (e.g., EMUs 3, 8, and 37), subtropical and tropical regions (e.g., EMUs 11, 24, 26, 33), and equatorial regions (e.g., EMUs 10, 18). Bimodal latitudinal distributions (Northern and Southern Hemispheres) are observed (e.g., EMUs 8, 11, 21, and 24). Some EMUs are associated with physiographic features, such as the Mendocino Ridge and Fracture Zone, situated near the southern boundary of EMU 30. Parts of some EMUs are located in the discharge regions of major rivers like the Amazon and Congo (EMUs 18 and 24). The Mediterranean and Red Seas were clustered into a single unit (EMU 9), consistent with Longhurst's (2007) identification of a single biogeochemical province for the Mediterranean. Reygondeau et al. (2017), however, objectively identified 63 three-dimensional, management-appropriate subdivisions of the Mediterranean, maintaining that while global analyses are useful for macro-scale comparisons of ocean regions, local management strategies and policies will require appropriately scaled geographic assessment and accounting units.

One very large, deep, circumglobal cluster (EMU 13) is observed in the Pacific and Indian Oceans but is all but absent in the Atlantic, consistent with the recognition of a Circumpolar Deep Water (CDW) mass by Emery and Meincke (1986). Likewise, EMU 29 is similar to the Arctic Deep Water (ADW) and North Atlantic Deep Water (NADW) units of those authors. Although some
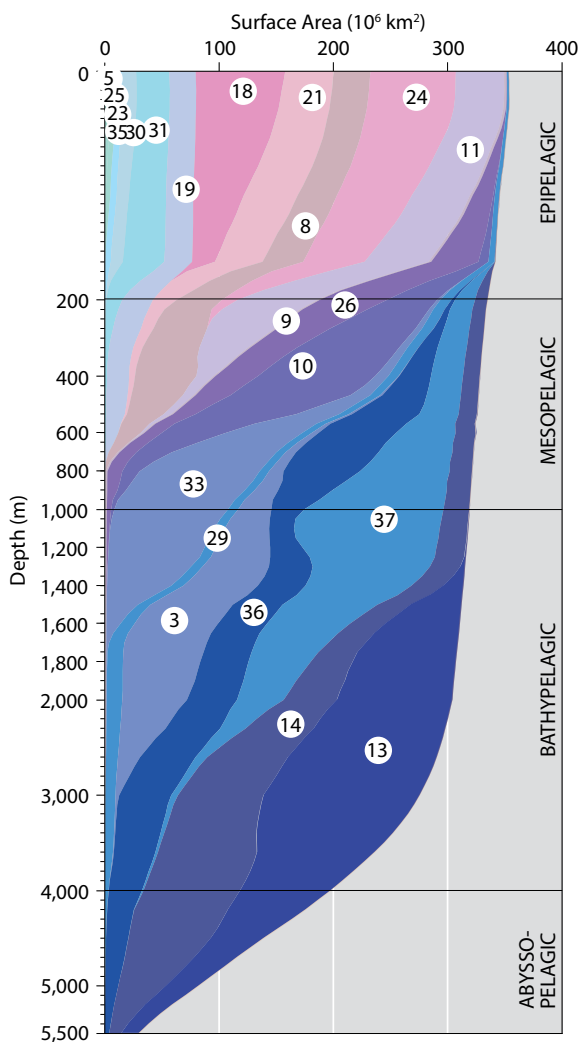

FIGURE 5. EMU distributions by depth. The two-dimensional global area (km²) at any depth is shown for the 22 EMUs that comprise 99% of the ocean volume. The horizontal boundary lines separating the depth zone classes are as described in the Coastal and Marine Ecosystem Classification Standard (CMECS), the Federal Geographic Data Committee (FGDC) standard for the United States (FGDC, 2012). The EMU number labels (see Appendices 1 and 2 in the online supplementary materials for names, maps, and descriptions of the EMUs) are placed at the median unit-middle depth for each EMU. Although the EMUs are not uniformly distributed into the CMECS depth zones, strong vertical separation is evident, with many small EMUs in the upper water column and fewer larger EMUs in the middle and lower water columns. Pink colors indicate warmer EMUs, and blue colors indicate colder EMUs.

of the EMUs are similar to water masses described by Emery and Meincke, in other instances there is less resemblance, which is to be expected, given that the EMUs were derived from six compositional properties rather than from the temperature- and salinity-derived units of Emery and Meincke.

The surface-occurring EMUs (upper left panel of Figure 4) can be compared with the surface-derived biogeochemical provinces (BGCPs) of Longhurst (2007). As mentioned above, the Mediterranean Sea was identified as a single unit (EMU 9) without subdivision in both classifications. The Mediterranean and Red Seas EMU was placed in the mesopelagic class because its median unit-middle depth (302 m) is between 200 m and 1,000 m, but its vertical distribution is throughout the water column. Other similarities between the EMUs and Longhurst's BGCPs are apparent. Both classification systems identify obvious latitudinal banding separating the Antarctic, Subantarctic, and Southern Hemisphere tropics and equatorial regions. EMU 18 (North Pacific Subtropical and Equatorial Indian Epipelagic) closely approximates the distributions of Longhurst's North Pacific Equatorial Countercurrent Province and Pacific Equatorial Divergence Province. Several of Longhurst's provinces in the Arctic and Subarctic regions (e.g., Boreal Polar Province, Atlantic Arctic Province, Atlantic Subarctic Province, North Pacific Epicontinental Sea Province) correspond visually with EMUs 5 (Arctic Epipelagic), 23 (Arctic and Labrador Sea Epipelagic), and 30 (North Pacific and Beaufort Sea Epipelagic). The latitudinal demarcation between Longhurst's two Indian Ocean provinces (Indian South Subtropical Gyre Province and Indian Monsoon Gyres Province) is a strong latitudinal delineation between EMUs 11 (Northern Subtropical and Southern Subtropical Epipelagic), 18 (North Pacific Subtropical and Equatorial Indian Epipelagic), and 24 (Tropical Pacific, Tropical Indian, and Equatorial Atlantic Epipelagic). The Longhurst system does

not include bimodal distributions of provinces in both Northern and Southern Hemispheres as was obtained with some EMUs (e.g., EMU 8, Subantarctic, North Atlantic, and North Pacific Epipelagic). Overall, considerable visual correspondence is observed between Longhurst's BGCPs and the surface-occurring EMUs, and a more quantitative comparison is merited.

Another set of ocean-surface regions that can be compared to the surface-occurring EMUs is that of Gebbie and Huybers (2011), who identified seven global surface water masses: Antarctic, North Atlantic, Subantarctic, North Pacific, Arctic, Mediterranean, and Tropics, representing the formation

regions of ocean waters. The boundaries separating these seven source regions are also present in the surface-occurring EMUs. It is evident that an aggregation of EMUs into the Gebbie and Huybers formation regions would be very "clean," resulting in minimal splitting of EMUs across formation region boundaries.

Neither the Large Marine Ecosystems of Sherman et al. (2005) nor the Marine Ecoregions of the World of Spalding et al. (2007) address open-ocean pelagic ecosystems, so comparisons between them and the geographic extent of the EMUs are not feasible. The interpretive, expert-derived subdivisions of the Global Open Ocean and Deep Seabed (GOODS) Biogeographic Classification (UNESCO,
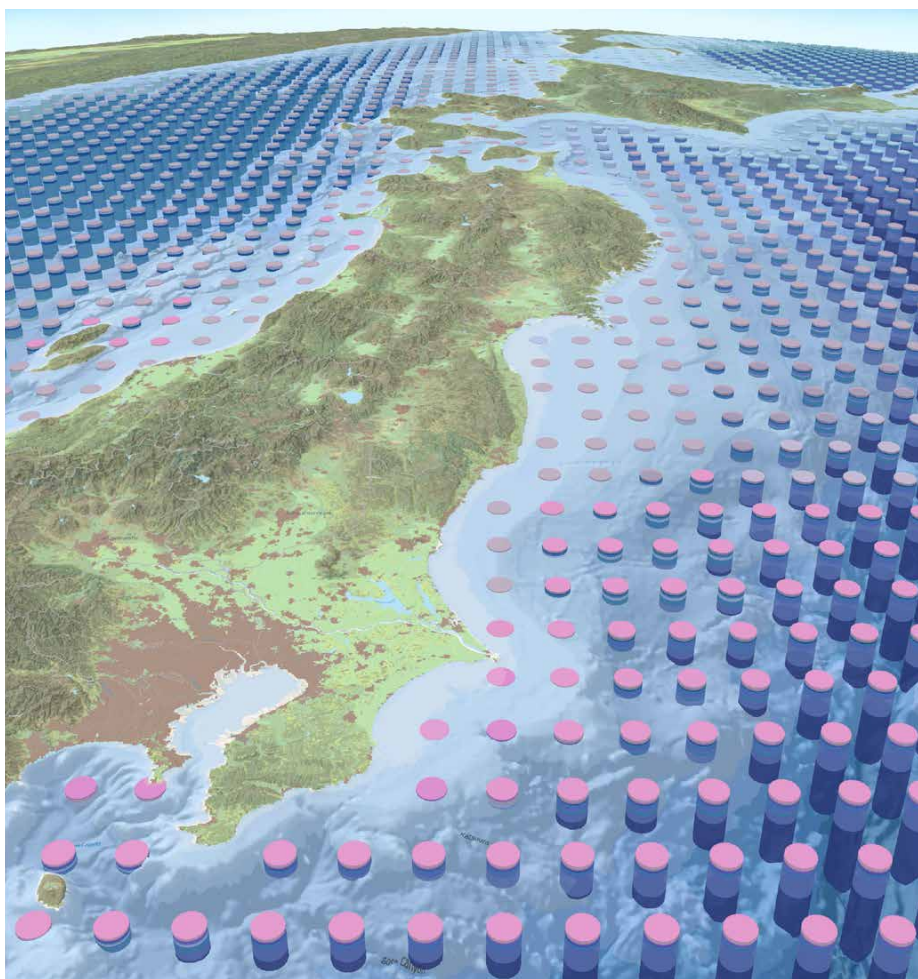

FIGURE 6. Example of the visualization approach taken to represent the EMUs in three dimensions, mapped over space. The region shown is off the eastern coast of Japan. Although the EMUs are mapped as a continuous surface, representing them in three dimensions is facilitated by the use of stacked cylinders, where each color band on a cylinder is an EMU. In the coastal zone, EMUs are single or few, and shallow, whereas offshore there are more and deeper EMUs. Surface temperature gradients are also apparent between the pink (warmer) EMUs in the south, and the blue (colder) EMUs in the north.

2009) are similarly difficult to compare with the quantitative, statistically derived EMUs presented here.

EMU Depth Distributions

The number of EMUs is highest at or near the surface, and decreases with depth (Figures 4 and 5). Although we used the CMECS criteria and the median unit-middle value to classify the EMUs into epipelagic, mesopelagic, bathypelagic, or abyssopelagic zones, Figure 5 shows that those depth class assignments, while informative, are also imperfect. Many EMUs are distributed across depth zone boundaries, with some having most of their distribution in one zone, but with vertical extensions into upper or lower depth zones as well. EMUs vary considerably in water column position, thickness, and horizontal area at varying depths. No EMUs were classified as abyssal as there were no EMUs with a median unit-middle depth >4,000 m. However, the distributions of many EMUs extend beyond the 4,000 m bathypelagic/abyssopelagic boundary. Acknowledging some overlap, the EMUs appear to be better separated, visually, at depths approximately corresponding to water column positions from 0 m to 200 m (upper water column), 200 m to 2,000 m (middle water column), and >2,000 m (lower water column). Although the EMUs are

currently classified into CMECS depth zones using standardized criteria, an attempt to statistically separate EMUs into depth classes and associated quantitative determination of the depth boundaries for those groupings appears warranted. In addition, the EMUs should be assessed for their depth-based relationship to physical properties like light attenuation limits (Stal, 2016) and biologically mediated phenomena like respiration (Costello and Breyer, 2017) and carbon flux (Reygondeau et al., 2017). We suggest that the depth boundaries of the vertical zones and the environmental factors controlling the vertical separation of the pelagic ocean need additional analysis and further clarification.

Biogeography and EMUs

Biogeographic regions are delineated from an analysis of species distribution data. The number and types of taxonomic groups represented (Fontaine et al., 2015), as well as the relative focus on endemism (Briggs and Bowen, 2012), can vary widely. Although not quantitatively assessed, we made a preliminary visual comparison (Figure 7) of the EMUs' spatial distributions and the distributions of 30 marine biogeographic realms developed from statistical clustering of species-distribution data, based on recent work of author Costello. The

30 realms were obtained from 2D clustering of occurrence records representing over 65,000 species from the Ocean Biogeographic Information System (OBIS, http://www.iobis.org) database, and they reflect global patterns of species endemicity. Initially, we note correspondence between realms and EMUs (e.g., realms 2, 5, 7, and 30) in some areas, but in other cases a relationship is less apparent (e.g., realms 18, 21, and 22). In another global assessment of 11,567 species occurrences representing 13 taxonomic groups, Tittensor et al. (2010) identified concentrations of coastal and open-ocean species in the eastern Pacific and in mid-latitudinal belts, respectively. The mapping of the EMUs will allow improved characterization of the horizontal and vertical distributions and the chemical and physical natures of species-rich regions.

Limitations and Future Work

We recognize limitations in our work related to both temporal scaling dimensions and parameters selected for the clustering. The WOA data offer several native temporal resolutions (seasonal, annual, and decadal) that we did not exploit. As our aim was to use long-term historical average values for the point locations, we used the 57-year mean values for the six parameters to map EMU
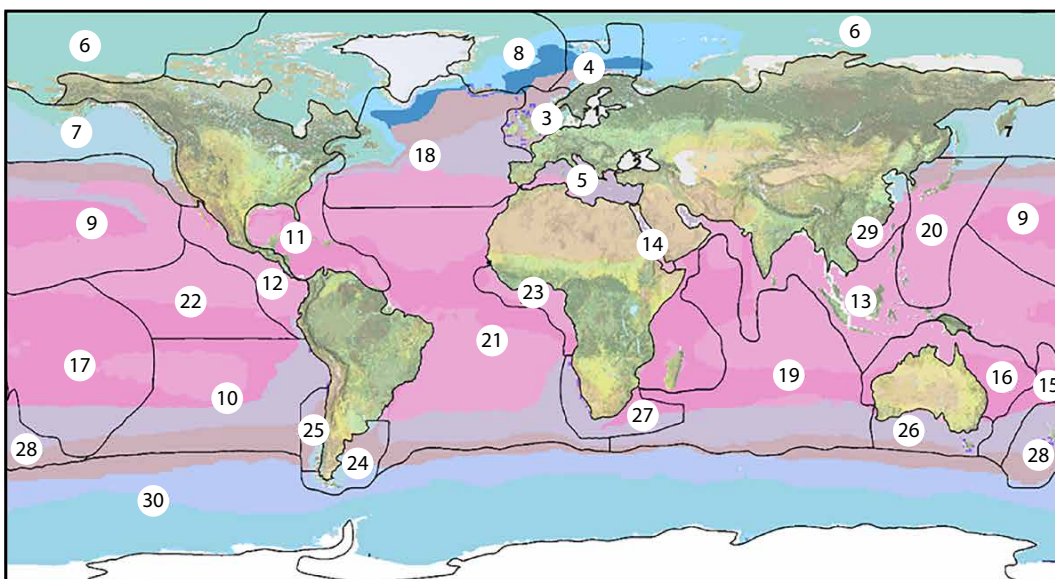

FIGURE 7. Relationship between surface-occurring EMU distributions (colors) and marine biogeographic realms (numbered, outlined polygons), from recent work of author Costello. Spatial congruence between biogeographic realms and surface-occurring EMUs is apparent for some realms (e.g., 5, 7, 26, 30) but not for others (e.g., 18, 21, 22).

extents and locations, and this approach has been successful in mapping ocean regions. However, it prohibits assessment of temporal variability and trends. Recent work shows that it would be possible to construct temporally sequenced EMU distribution maps (e.g., Oliver and Irwin, 2008; Reygondeau et al., 2013; Kavanaugh et al., 2014). We now have a framework for that assessment and are planning the development of seasonal, annual, and decadal characterizations of oceanic water masses. However, the computational requirements for six variables increase by orders of magnitude when contemplating temporal variations for over 52 million points. This is currently a big-data challenge (Gallagher et al., 2015; Alder and Hostetler, 2015; Coro et al., 2016; Wright, 2016), but as spatial processing technologies evolve, these kinds of analyses will be rendered less computationally intense than they are at present.

The clustering of oceanic data to derive EMUs was based on the six variables in the WOA data set. The addition of other variables would likely influence the oceanic partitioning we present here. The inclusion of data on particulate organic carbon (POC), carbonate contents, and ocean current patterns might influence the clustering results. POC plays a crucial role in the marine and global carbon cycle and is a primary component of oceanographic food webs (Buesseler et al., 2007). POC flux was one of the parameters used to subdivide the Mediterranean Sea into 63 biogeochemical regions (Reygondeau et al., 2017), where it was used to quantitatively separate the mesopelagic layer from the bathypelagic layer. Variability in the vertical flux of POC is important for understanding the main pathways by which organic carbon is formed in ocean surface waters via photosynthesis and then transferred to the deep ocean where it may be sequestered (Lutz et al., 2007). Similarly, the carbonate chemistry of the ocean is ecologically important, as the persistence of ocean acidification is likely to have implications for many surface and pelagic ecosystems and

communities (Sherman, 2014; Wallace et al., 2014; Thresher et al., 2015). Finally, variables associated with ocean currents, such as flow direction and magnitude, may substantially influence EMU characteristics and distributions. We plan to pursue the addition of these attributes to the ocean mesh and to study their effects on EMU distributions in future statistical clustering analyses.

We also plan to enrich the EMU resource by combining the EMU data with other data layers. This will result in the creation of new 2D layers for the sea surface and the seafloor. For example, for the seafloor, we have combined the bottom-occurring EMUs with the seafloor physiographic regions and features of Harris et al. (2014) in order to evaluate the influence of seafloor geomorphology on the water-column structure above it. We also have combined a 13-year average ocean color value data set (chlorophyll *a* from the NASA Aqua-MODIS sensor) to our surface-occurring EMUs, and plankton abundance characteristics are now available for the surface data points and the surface-occurring EMUs. We intend to continue adding associative attributes from other globally available resources, and we are exploring the relationship between EMUs and established temporally dynamic climatological classifications (Oliver and Irwin, 2008; Kavanaugh et al., 2016).

While we are calling these 37 volumetric regions EMUs, we acknowledge that their true ecological character has not yet been established. Their derivation from entirely physicochemical data, and their similarity to widely recognized global surface waters, lends validation to the EMUs as physically and chemically distinct volumetric regions. We call them ecological in the general sense that depth, temperature, salinity, oxygen, and nutrients are known to be important in structuring biotic distributions (Longhurst, 2007; Oliver and Irwin, 2008), and because microbial processes shape nutrient and oxygen distributions throughout the water column (Kavanaugh et al.,

2016), but we have not documented the relationship between environmental variation and species diversity. As a first step, we are currently undertaking a more quantitative assessment of this relationship between physically and chemically distinct regions in the ocean and species biogeography, the results of which will facilitate a deeper understanding of the true ecological nature of the EMUs. For example, we are exploring the cross-indexing of OBIS species records and EMUs, and we expect that subsequent versions of EMUs will not only contain species records as attributes but may also change geographically to be more reflective of marine organism distributions.

Finally, we recognize the opportunity to use finer-resolution data to create a more refined mapping of EMUs at regional and local scales, as was demonstrated by Reygondeau et al. (2017). Moreover, we plan to elucidate the coastal and estuarine units in greater detail in an independent development of a set of ecological coastal units (ECUs), which will be undertaken along the entire global shoreline. We are working on the development of a new global shoreline vector extracted from satellite imagery (30 m spatial resolution) and attributed with environmental characteristics as the spatial framework for the planned development of a global set of ECUs.

CONCLUSIONS

The present EMU mapping effort is an objective partitioning of the global ocean into environmentally distinct volumetric region units using an aspatial clustering exercise where clusters were not constrained into particular ocean regions and the cluster sites were selected by homogeneity in physical and chemical parameters only, blind to both depth and location. We have developed a new classification scheme for 37 compositionally varying marine volumetric regions. Their properties are listed using both standardized CMECS descriptors and new EMU terminology. EMU volumetric region names are compiled by combining a

geographic descriptor and a depth zone term (Appendices 1 and 2).

The aim of this work was to produce a new global characterization and detailed data set of marine environments as a resource for biogeographic assessments, impact studies, biodiversity priority setting, and ecosystem accounting, management, and research. The mapping of global ecological land units (ELUs), sponsored by the GEO commission, has now been extended to the ocean. We partitioned the global ocean water column into 37 physically and chemically distinct volumetric regions using available data from the 57-year World Ocean Atlas data set. The global map of EMUs is an initial, unsupervised, statistical classification approach to mapping ocean environmental structure in three dimensions. Based on this methodology, we can re-cluster the ocean using additional or different deterministic variables if desirable, and also cluster the global ocean in different time intervals ranging from seasonal to annual to decadal to explore the temporal geographies of EMUs. The existence of the EMUs has the potential to facilitate research on the extent to which environmental drivers control biotic distributions. The EMU data we have created allow for characterization of the physical and chemical environment contained in marine protected areas, fishing grounds, or other marine geographies. As an open-access resource, the EMU data are available to scientists, managers, and the interested public. Future work will enrich the EMU resource by adding additional attributes to the ocean mesh and developing a finer resolution ecological coastal unit (ECU) product along the global shoreline. ◉

REFERENCES
Alder, J.R., and S.W. Hostetler. 2015. Web based visualization of large climate data sets. *Environmental Modelling & Software* 68:175–180, https://doi.org/10.1016/j.envsoft.2015.02.016.

Bailey, R.G. 1996. *Ecosystem Geography*. Springer-Verlag, New York, NY, 204 pp.

Bailey, R.G. 2014. *Ecoregions – The Ecosystem Geography of the Oceans and Continents,* 2nd ed. Springer, New York, NY, 180 pp.

Beaugrand, G., I. Rombouts, and R.R. Kirby. 2013. Towards an understanding of the pattern of biodiversity in the oceans. *Global Ecology and Biogeography* 22:440–449, https://doi.org/10.1111/geb.12009.

Briggs, J.C., and B.W. Bowen. 2012. A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography* 39:12–30, https://doi.org/10.1111/j.1365-2699.2011.02613.x.

Buesseler, K.O., C.H. Lamborg, P.W. Boyd, P.J. Lam, T.W. Trull, R.R. Bidigare, J.K.B. Bishop, K.L. Casciotti, F. Dehairs, M. Elskens, and others. 2007. Revisiting carbon flux through the ocean's twilight zone. *Science* 316(5824):567–570, https://doi.org/10.1126/science.1137959.

Calinski, T., and J. Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics* 3(1):1–27, https://doi.org/10.1080/03610927408827101.

Celebi, M., H. Kingravi, and P. Vela. 2013. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications* 40(1):210–215, https://doi.org/10.1016/j.eswa.2012.07.021.

Charette, M.A., and W.H.F. Smith. 2010. The volume of Earth's ocean. *Oceanography* 23(2):112–114, https://doi.org/10.5670/oceanog.2010.51.

Coro, G., C. Magliozzi, A. Ellenbroek, K. Kaschner, and P. Pagano. 2016. Automatic classification of climate change effects on marine species distributions in 2050 using the AquaMaps model. *Environmental and Ecological Statistics* 23:155–180, https://doi.org/10.1007/s10651-015-0333-8.

Costello, M.J. 2009. Distinguishing marine habitat classification concepts for ecological data management. *Marine Ecology Progress Series* 397:253–268, https://doi.org/10.3354/meps08317.

Costello, M.J., and S. Breyer. 2017. Ocean depths: The mesopelagic and implications for global warming. *Current Biology* 27:113–119, https://doi.org/10.1016/j.cub.2016.11.042.

Emery, W.J. 2001. Water types and water masses. Pp. 1,556–1,567 in *Encyclopedia of Ocean Sciences*. J.H. Steele, S.A. Thorpe, and K.K. Turekian, eds, Academic Press, San Diego, CA, 3,399 pp.

Emery, W.J., and J. Meincke. 1986. Global water masses: Summary and review. *Oceanologica Acta* 9(4):383–391.

FAO (Food and Agriculture Organization). 2016. Coordinating Working Party Handbook of Fishery Statistical Standards, Section H: Fishing Areas for Statistical Purposes. FAO Fisheries and Aquaculture Department, http://www.fao.org/fishery/cwp/handbook/h/en.

FGDC (Federal Geographic Data Committee). 2012. *Coastal and Marine Ecological Classification Standard Version 4.0.* Report FGDC-STE-18-2012, Federal Geographic Data Committee Secretariat, US Geological Survey, Reston, Virginia, 343 pp.

Fontaine, A., R. Devillers, P.R. Peres-Neto, and L.E. Johnson. 2015. Delineating marine ecological units: A novel approach for deciding which taxonomic group to use and which taxonomic resolution to choose. *Diversity and Distributions* 21(10):1,167–1,180, https://doi.org/10.1111/ddi.12361.

Gallagher, J., J. Orcutt, P. Simpson, D. Wright, J. Pearlman, and L. Raymond. 2015. Facilitating open exchange of data and information. *Earth Science Informatics* 8:721–739, https://doi.org/10.1007/s12145-014-0202-2.

Garcia, H.E., R.A. Locarnini, T.P. Boyer, J.I. Antonov, O.K. Baranova, M.M. Zweng, J.R. Reagan, and D.R. Johnson. 2014a. *World Ocean Atlas 2013 version 2 (WOA13 V2), Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation.* S. Levitus, ed., and A. Mishonov, technical ed., NOAA Atlas NESDIS 75, NOAA National Centers for Environmental Information, Silver Spring, MD, 27 pp.

Garcia, H.E., R.A. Locarnini, T.P. Boyer, J.I. Antonov, O.K. Baranova, M.M. Zweng, J.R. Reagan, and D.R. Johnson. 2014b. *World Ocean Atlas 2013 version 2 (WOA13 V2), Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate).* S. Levitus, ed., and A. Mishonov, technical ed., NOAA Atlas NESDIS 76, NOAA National Centers for Environmental Information, Silver Spring, MD, 25 pp.

Gebbie, G., and P. Huybers. 2010. Total matrix intercomparison: A method for determining the geometry of water-mass pathways. *Journal of Physical Oceanography* 40:1,710–1,728, https://doi.org/10.1175/2010JPO4272.1.

Gebbie, G., and P. Huybers. 2011. How is the ocean filled? *Geophysical Research Letters* 138, L06604, https://doi.org/10.1029/2011GL046769.

Hall, A. 2014. Projecting regional change. *Science* 346(6216):1,461–1,462, https://doi.org/10.1126/science.aaa0629.

Hardman-Mountford, N.J., T. Hirata, K. Richardson, and J. Aiken. 2009. An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. *Remote Sensing of Environment* 112:3,341–3,352, https://doi.org/10.1016/j.rse.2008.02.016.

Harris, P.T., M. Macmillan-Lawler, J. Rupp, and E.K. Baker. 2014. Geomorphology of the oceans. *Marine Geology* 352:4–24, https://doi.org/10.1016/j.margeo.2014.01.011.

Harris, P.T., and T. Whiteway. 2009. High seas marine protected areas: Benthic environmental conservation priorities from a GIS analysis of global ocean biophysical data. *Ocean & Coastal Management* 52:22–38, https://doi.org/10.1016/j.ocecoaman.2008.09.009.

Heinzer, T.J., M.D. Williams, E.C. Dogrul, T.N. Kadir, C.F. Brush, and F.I. Chung. 2012. Implementation of a feature-constraint mesh generation algorithm within a GIS. *Computers and Geosciences* 49:46–52, https://doi.org/10.1016/j.cageo.2012.06.004.

Holt, B.G., J. Lessard, M.K. Borregaard, S.A. Fritz, M.B. Araújo, D. Dimitrov, P. Fabre, C.H. Graham, G.R. Graves, K.A. Jønsson, and others. 2013. An update of Wallace's zoogeographic regions of the world. *Science* 339(6115):74–78, https://doi.org/10.1126/science.1228282.

ICES (International Council for Exploration of the Seas). 2004. *Report of the Regional Ecosystem Study Group for the North Sea.* April 5–7, 2004, ICES CM 2004/ACE:06, International Council for Exploration of the Seas, Copenhagen, Denmark.

IHO (International Hydrographic Organization). 2002. *International Hydrographic Organization Publication S-23: Limits of Oceans and Seas, draft 4th ed.* International Hydrographic Organization, Paris, Monaco, https://www.iho.int/mtg_docs/com_wg/S-23WG/S-23WG_Misc/Draft_2002/Draft_2002.htm.

Jain, A.K. 2010. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters* 31(8):651–666, https://doi.org/10.1016/j.patrec.2009.09.011.

Jamieson, A.J. 2011. Ecology of deep oceans: Hadal trenches. *eLS*, https://doi.org/10.1002/9780470015902.a0023606.

Kavanaugh, M.T., B. Hales, M. Saraceno, Y.H. Spitz, A.E. White, and R.M. Letelier. 2014. Hierarchical and dynamic seascapes: A quantitative framework for scaling pelagic biogeochemistry and ecology. *Progress in Oceanography* 120:291–304, https://doi.org/10.1016/j.pocean.2013.10.013.

Kavanaugh, M.T., M.J. Oliver, F.P. Chavez, R.M. Letelier, F.E. Muller-Karger, and S.C. Doney. 2016. Seascapes as a new vernacular for pelagic ocean monitoring, management and conservation. *ICES Journal of Marine Science* 73(7):1,839–1,850, https://doi.org/10.1093/icesjms/fsw086.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. World map of the Köppen-Geiger climate classification updated. *Meteorologische Zietschrift* 15(3):259–263.

Li, Z., and C. Gold. 2004. Multi-dimensional geospatial technology for geosciences. *Computers and Geosciences* 30(4):321–323, https://doi.org/10.1016/j.cageo.2003.09.010.

Liu, J., T. Dietz, S.R. Carpenter, M. Alberti, C. Folke, E. Moran, A.N. Pell, P. Deadman, T. Kratz, J. Lubchenco, and others. 2007. Complexity of coupled human and natural systems. *Science* 317(5844):1,513–1,516, https://doi.org/10.1126/science.1144004.

Liu, J., H.A. Mooney, V. Hull, S.J. Davis, J. Gaskell, T. Hertel, and J. Lubchenco, K.C. Seto, P. Gleick, C. Kremen, and S. Li. 2015. Systems integration for global sustainability. *Science* 347(6225), https://doi.org/10.1126/science.1258832.

Locarnini, R.A., A.V. Mishonov, J.I. Antonov, T.P. Boyer, H.E. Garcia, O.K. Baranova, M.M. Zweng, C.R. Paver, J.R. Reagan, D.R. Johnson, and others. 2013. *World Ocean Atlas 2013 version 2 (WOA 2013 V2), Volume 1: Temperature*. S. Levitus, ed., and A. Mishonov, technical ed., NOAA Atlas NESDIS 73, NOAA National Centers for Environmental Information, Silver Spring, MD, 40 pp.

Longhurst, A. 2007. *Ecological Geography of the Sea*, 2nd ed. Academic Press, San Diego, CA, 542 pp., https://doi.org/10.1016/B978-012455521-1/50000-0.

Lutz, M.J., K. Caldeira, R.B. Dunbar, and M.J. Behrenfeld. 2007. Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. *Journal of Geophysical Research* 112, C10011, https://doi.org/10.1029/2006JC003706.

Milligan, G.W., and M.C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50(2):159–179, https://doi.org/10.1007/bf02294245.

Milligan, G.W., and M.C. Cooper. 1988. A study of standardization of variables in cluster analysis. *Journal of Classification* 5:181–204, https://doi.org/10.1007/BF01897163.

Odum, E.P. 1971. *Fundamentals of Ecology*, 3rd ed. W.B. Saunders, Philadelphia, PA, 574 pp.

Oliver, M.J., S. Glenn, J.T. Kohut, A.J. Irwin, O.M. Schofield, M. Moline, and W.P. Bissett. 2004. Bioinformatic approaches for objective detection of water masses on continental shelves. *Journal of Geophysical Research* 109, C07S04, https://doi.org/10.1029/2003JC002072.

Oliver, M.J., and A.J. Irwin. 2008. Objective global ocean biogeographic provinces. *Geophysical Research Letters* 35, L15601, https://doi.org/10.1029/2008GL034238.

Pinet, P.R. 2009. *Invitation to Oceanography,* 5th ed. Jones and Bartlett Publishers Inc.

Proud, R., M.J. Cox, and A.S. Brierley. 2017. Biogeography of the global ocean's mesopelagic zone. *Current Biology* 27:113–119, https://doi.org/10.1016/j.cub.2016.11.003.

Reygondeau, G., C. Guieu, F. Benedetti, J. Irrison, S. Ayata, S. Gasparini, and P. Koubbi. 2017. Biogeochemical regions of the Mediterranean Sea: An objective multidimensional and multivariate environmental approach. *Progress in Oceanography* 151:138–148, https://doi.org/10.1016/j.pocean.2016.11.001.

Reygondeau, G., A. Longhurst, E. Martinez, G. Beaugrand, D. Antoine, and O. Maury. 2013. Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles* 27(4):1,046–1,058, https://doi.org/10.1002/gbc.20089.

Sayre, R., J. Dangermond, C. Frye, R. Vaughan, P. Aniello, S. Breyer, D. Cribbs, D. Hopkins, R. Nauman, W. Derrenbacher, and others. 2014. *A New Map of Global Ecological Land Units: An Ecophysiographic Stratification Approach*. Association of American Geographers, Washington, DC, 46 pp.

Schoch, G.C., D.M. Albert, and C.S. Shanley. 2014. An estuarine habitat classification for a complex fjordal island archipelago. *Estuaries and Coasts* 37(1):160–176, https://doi.org/10.1007/s12237-013-9622-3.

Sherman, K. 2014. Adaptive management institutions at the regional level: The case of Large Marine Ecosystems. *Ocean and Coastal Management* 90:38–49, https://doi.org/10.1016/j.ocecoaman.2013.06.008.

Sherman, K., M. Sissenwine, V. Christensen, A. Duda, G. Hempel, C. Ibe, S. Levin, D. Lluch-Belda, G. Matishov, J. McGlade, and others. 2005. A global movement toward an ecosystem approach to marine resources management. *Marine Ecology Progress Series* 300:275–279.

Spalding, M.D., H.E. Fox, G.R. Allen, N. Davidson, Z.A. Ferdaña, M. Finlayson, B.S. Halpern, M.A. Jorge, A. Lombana, S.A. Lourie, and others. 2007. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience* 57(7):573–583, https://doi.org/10.1641/B570707.

Stal, L.J. 2016. The euphotic realm. Pp. 209–225 in *The Marine Microbiome*. L.J. Stal and M.S. Cretoiu, eds, Springer International Publishing, Switzerland.

Sverdrup, H.U., M.W. Johnson, and R.H. Fleming. 1942. *The Oceans: Their Physics, Chemistry and General Biology*. Prentice-Hall, Englewood Cliffs, New York, 1,087 pp.

Thresher, R.E., J.M. Guinotte, R.J. Matear, and A.J. Hobday. 2015. Options for managing impacts of climate change on a deep-sea community. *Nature Climate Change* 5:635–639, https://doi.org/10.1038/nclimate2611.

Tittensor, D.P., C. Mora, W. Jetz, H.K. Lotze, D. Ricard, E.V. Berghe, and B. Worm. 2010. Global patterns and predictors of marine biodiversity across taxa. *Nature* 466(7310):1,098–1,101, https://doi.org/10.1038/nature09329.

Tomczak, M. 1999. Some historical, theoretical, and applied aspects of quantitative water mass analysis. *Journal of Marine Research* 57:275–303, https://doi.org/10.1357/002224099321618227.

UN (United Nations). 1982. Article 57: Breadth of the Exclusive Economic Zone. *United Nations Convention on the Law of the Sea of 10 December 1982,* United Nations Secretary General, Paris, France.

UNESCO (United Nations Educational Scientific and Cultural Organization). 2009. *Global Open Oceans and Deep Seabed (GOODS): Biogeographic Classification.* UNESCO-IOC Technical Series 84, United Nations Educational and Scientific Council Intergovernmental Oceanographic Commission, Paris, France, 96 pp.

Vivone, G., L. Alparone, J. Channussot, M. Dalla Mura, A. Garzelli, G. Licciardi, R. Restaino, and L. Wald. 2015. A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing* 53(5):2,565–2,586, https://doi.org/10.1109/TGRS.2014.2361734.

Wallace, R.B., H. Baumann, J.S. Grear, R.C. Aller, and C.J. Gobler. 2014. Coastal ocean acidification: The other eutrophication problem. *Estuarine, Coastal, and Shelf Science* 148:1–13, https://doi.org/10.1016/j.ecss.2014.05.027.

Watling, L., J. Guinotte, M.R. Clark, and C.R. Smith. 2013. A proposed biogeography of the deep ocean floor. *Progress in Oceanography* 111:91–112, https://doi.org/10.1016/j.pocean.2012.11.003.

Wright, D.J., ed. 2016. *Ocean Solutions, Earth Solutions,* 2nd ed. Esri Press, Redlands, CA, 500 pp.

Wright, D.J., M.J. Blongewicz, P.N. Halpin, and J. Breman. 2007. *Arc Marine: GIS for a Blue Planet.* Esri Press, Redlands, CA, 202 pp.

Zweng, M.M., J.R. Reagan, J.I. Antonov, R.A. Locarnini, A.V. Mishonov, T.P. Boyer, H.E. Garcia, O.K. Baranova, D.R. Johnson, D. Seidov, and M.M. Biddle. 2013. *World Ocean Atlas 2013 version 2 (WOA13 V2), Volume 2: Salinity*. S. Levitus, ed., and A. Mishonov, technical ed., NOAA Atlas NESDIS 74, NOAA National Centers for Environmental Information, Silver Spring, Maryland, 39 pp.

AUTHORS
Roger G. Sayre (rsayre@usgs.gov) is Senior Scientist for Ecosystems, U.S. Geological Survey (USGS), Reston, VA, USA. **Dawn J. Wright** is Chief Scientist, **Sean P. Breyer** is Program Manager ArcGIS Content, **Kevin A. Butler** is Spatial Statistics Product Engineer, and **Keith Van Graafeiland** is Product Engineer, all at Esri, Redlands, CA, USA. **Mark J. Costello** is Associate Professor, Institute of Marine Science, University of Auckland, Auckland, New Zealand. **Peter T. Harris** is Managing Director, GRID-Arendal, Arendal, Norway. **Kathleen L. Goodin** is Chief of Staff, Conservation Science Division, NatureServe, Arlington, VA, USA. **John M. Guinotte** is Fish and Wildlife Biologist, U.S. Fish and Wildlife Service, Denver, CO, USA, formerly at the Marine Conservation Institute, Seattle, WA, USA, during this study. **Zeenatul Basher** is Research Technologist, USGS, Reston, VA, USA. **Maria T. Kavanaugh** is Research Associate, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. **Patrick N. Halpin** is Associate Professor of Marine Geospatial Ecology, Nicholas School of the Environment, Duke University, Durham, NC, USA. **Mark E. Monaco** is Acting Director, Center for Coastal Monitoring and Assessment, National Ocean Service, National Oceanic and Atmospheric Administration, Silver Spring, MD, USA. **Noel Cressie** is Distinguished Professor, National Institute for Applied Statistics Research Australia, University of Wollongong, Wollongong, Australia. **Peter Aniello** is Geospatial Scientist, Sandia National Laboratory, Albuquerque, NM, USA. **Charles E. Frye** is Chief Cartographer, and **Drew Stephens** is Ocean Industry Manager, both at Esri, Redlands, CA, USA.