Ecological Marine Units as a Framework for Collaborative Data Science and Knowledge Discovery

Dawn Wright[1], Roger Sayre[2], Sean Breyer[1], Kevin Butler[1], Keith VanGraafeiland[1], Mark Costello[1], Kathy Goodin[1], Maria Kavanaugh[1], Noel Cressie[1], Zeenatul Basher[1], Peter Harris[1], and John Guinotte[1]

[1]Affiliation not available
[2]US Geological Survey

November 22, 2022

Abstract

We present a data-derived, ecosystem mapping approach for the global ocean as commissioned by the Group on Earth Observations (GEO) and as a contribution to the Marine Biodiversity Observation Network (MBON). These ecological marine units (EMUs) are comprised of a global point mesh framework, created from over 52 million points from NOAA's World Ocean Atlas with a spatial resolution of 1 by 1 degree (27 x 27 km at the equator) at 44 varying depths and a temporal resolution that is currently decadal. Each point carries attributes of chemical and physical oceanographic structure (temperature, salinity, dissolved oxygen, nitrate, silicate, phosphate) as likely drivers of many marine ecosystem responses. We used a k-means statistical clustering algorithm to identify physically distinct, relatively homogenous, volumetric regions within the water column (the EMUs). Backwards stepwise discriminant analysis determined if all of six variables contributed significantly to the clustering, and a pseudo F-statistic gave us an optimum number of clusters worldwide at 37. A major intent of the EMUs is to support marine biodiversity conservation assessments, economic valuation studies of marine ecosystem goods and services, and studies of ocean acidification and other impacts. As such, they represent a rich geospatial accounting framework for these types of studies, as well as for scientific research on species distributions. To further benefit the community and facilitate collaborate knowledge building, data products are shared openly and interoperably via www.esri.com/ecological-marine-units. This includes provision of 3D point mesh and EMU clusters at the surface, bottom, and within the water column in varying formats via download, web services or web apps, as well as generic algorithms and GIS workflows that scale from global to regional and local. Work is in progress to delineate EMUs at finer spatial and temporal resolutions and to include ocean currents and various biodiversity observations. A major aim is for the ocean science community members to move the research forward with higher-resolution data from their own field studies or areas of interest, with the original EMU project team assisting with GIS implementation (especially via a new online discussion forum), and hosting of additional data products as needed.

Ecological Marine Units

as a Framework for Collaborative Data Science and Knowledge Discovery

OD24B-2728 DOI: 10.5670/oceanog.2017.116

Dawn J. Wright[1,2], Roger G. Sayre[3], Sean Breyer[1], Kevin A. Butler[1], Keith VanGraafeiland[1], Mark J. Costello[4], Kathy Goodin[5], Maria Kavanaugh[2], Noel Cressie[6], Zeenatul Basher[3], Peter T. Harris[7], John M. Guinotte[8]

[1] Environmental Systems Research Institute (aka Esri), Redlands, CA 92373, USA, dwright@esri.com; [2] College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR; [3] US Geological Survey; [4] University of Aucklund, New Zealand; [5] NatureServe; [6] University of Wollongong, Australia; [7] GRID-Arendal, Norway; [8] US Fish and Wildlife Service

We present a data-derived, ecosystem mapping approach for the global ocean as commissioned by the Group on Earth Observations (GEO) and as a contribution to the Marine Biodiversity Observation Network (MBON). These ecological marine units (EMUs) are comprised of a global point mesh framework, created from over 52 million points from NOAA's World Ocean Atlas with a spatial resolution of 1 by 1 degree (~27 x 27 km at the equator) at 44 varying depths and a temporal resolution that is currently decadal. Each point carries attributes of chemical and physical oceanographic structure (temperature, salinity, dissolved oxygen, nitrate, silicate, phosphate) as likely drivers of many marine ecosystem responses. We used a k-means statistical clustering algorithm to identify physically distinct, relatively homogenous, volumetric regions within the water column (the EMUs). Backwards stepwise discriminant analysis determined if all of six variables contributed significantly to the clustering, and a pseudo F-statistic gave us an optimum number of clusters worldwide at 37.

A major intent of the EMUs is to support marine biodiversity conservation assessments, economic valuation studies of marine ecosystem goods and services, and studies of ocean acidification and other impacts. As such, they represent a rich geospatial accounting framework for these types of studies, as well as for scientific research on species distributions.

To further benefit the community and facilitate collaborate knowledge building, data products are shared openly and interoperably via www.esri.com/ecological-marine-units. This includes provision of 3D point mesh and EMU clusters at the surface, bottom, and within the water column in varying formats via download, web services or web apps, as well as generic algorithms and GIS workflows that scale from global to regional and local.

Work is in progress to delineate EMUs at finer spatial and temporal resolutions and to include ocean currents and various biodiversity observations. A major aim is for the ocean science community members to move the research forward with higher-resolution data from their own field studies or areas of interest, with the original EMU project team assisting with GIS implementation (especially via a new online discussion forum), and hosting of additional data products as needed.

Methods

The EMU resource is comprised of a global point mesh framework, created from 52,487,233 points from the NOAA World Ocean Atlas; spatial resolution is ¼° by ¼° by varying depth; temporal resolution is currently decadal; each point has x, y, z, as well as six attributes of chemical and physical oceanographic structure (temperature, salinity, dissolved oxygen, nitrate, silicate, phosphate) that are likely drivers of many ecosystem responses.

We implemented a k-means statistical clustering of the point mesh to to identify physically distinct, relatively homogenous, volumetric regions in the water column. Calculation and inspection of the behavior of the pseudo F-statistic suggested the identification and mapping of 37 environmentally distinct 3D regions (candidate 'ecosystems') within the water column.

	DEPTH	TEMPERATURE	OXYGEN CONTENT	SALINITY	CHEMICAL CONTENT Nitrate	Silicate	Phosphate
UPPER LEVELS	5	Superchilled	Highly Oxic	Low	Low	Low	Low
	23	Superchilled	Highly Oxic	Normal	Low	Low	Low
	25	Superchilled	Highly Oxic	Normal	Low	Low	Low
	19	Cold	Oxic	Normal	Medium	Low	Low
	9	Moderate to Cool	Oxic	Normal	Low	Low	Low
	11	Moderate to Cool	Oxic	Normal	Low	Low	Low
	8	Moderate to Cool	Oxic	Normal	Medium	Low	Low
	35	Superchilled	Oxic	Normal	Low	Low	Low
	31	Superchilled	Oxic	Normal	Medium	Medium	Low
	30	Very Cold	Oxic	Normal	Medium	Low	Low
	18	Warm to Very Warm	Oxic	Normal	Low	Low	Low
	21	Warm to Very Warm	Oxic	Normal	Low	Low	Low
	24	Warm to Very Warm	Oxic	Normal	Low	High	Low
MID LEVELS	26	Moderate to Cool	Hypoxic	Normal	Medium	Low	Low
	37	Very Cold	Oxic	Normal	High	Medium	Low
	29	Very Cold	Oxic	Normal	Medium	Low	Low
	10	Cold	Severely Hypoxic	Normal	High	Low	Low
	3	Very Cold	Severely Hypoxic	Normal	High	High	Medium
	33	Very Cold	Severely Hypoxic	Normal	High	Medium	Medium
LOWER LEVELS	13	Very Cold	Hypoxic	Normal	High	High	Medium
	14	Very Cold	Oxic	Normal	High	High	Low
	36	Very Cold	Oxic	Normal	Medium	Low	Low

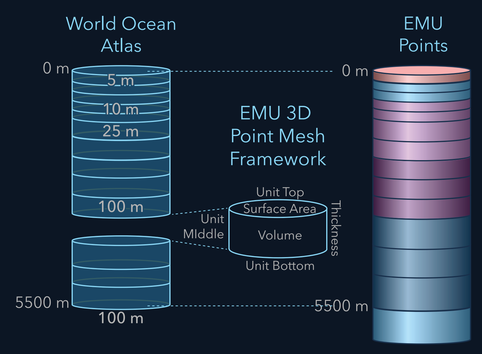
Feature Attributes

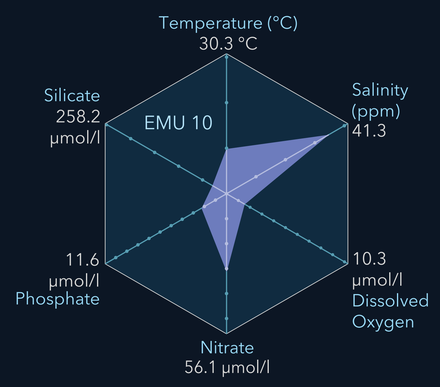
Depth Level
Temperature
Salinity
Dissolved Oxygen
Nitrate
Silicate
Phosphate

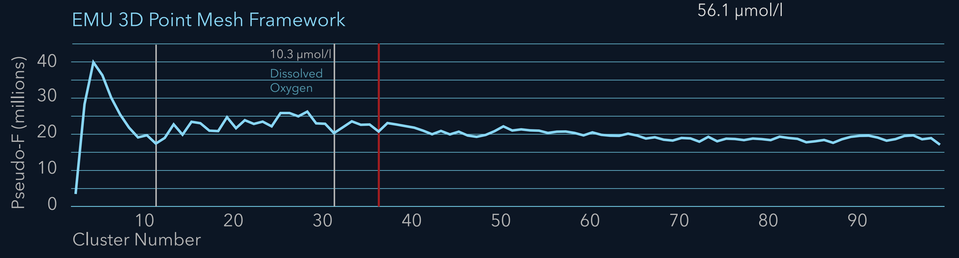
MODIS ocean color

PointID
QuarterID
UnitTop (m)
UnitMiddle (m)
UnitBottom (m)
Thickness (m)
ThicknessPos (m)
EMUID
EMU Name

GeomorphologyBase
GeomorphologyFeatures
SurfaceArea
Volume, SpecialCases





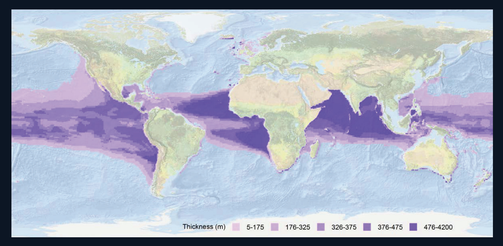


A plot of the pseudo-F statistic (y axis) against the requested number of clusters (x axis) in successive iterations from 2 to 100 clusters, incremented by one for each successive iteration. The red vertical line at 37 indicates an optimum number of clusters.

Results

The result is a standardized, rigorous, and ecologically meaningful set of ocean ecosystem units which may be used as a basemap alongside an organization's own GIS overlays for climate change impacts studies, biodiversity priority-setting, economic and social valuation studies, research, and marine spatial planning.

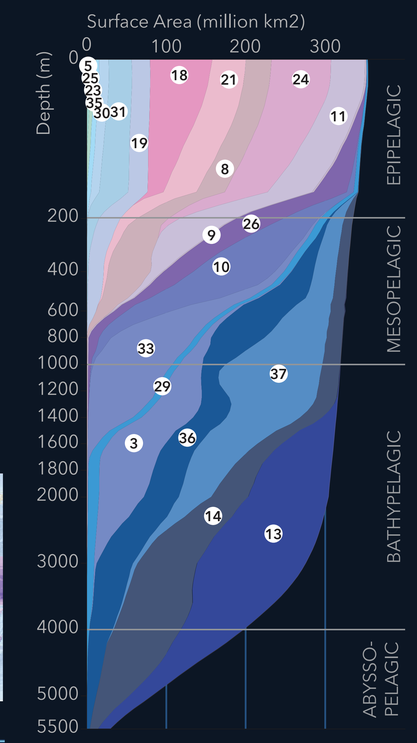
The two-dimensional global area (km2) at any depth is shown for the 22 EMUs that comprise 99% of the ocean volume (while the remaining 15 are small, shallow, and coastal, and collectively represent only about 1% of the ocean volume). The horizontal boundary lines separating the depth zone classes correspond with the Coastal and Marine Ecosystem Classification Standard (CMECS). The global distribution of EMUs is also shown at 8 depth intervals.



EMU 10 Summary Statistics

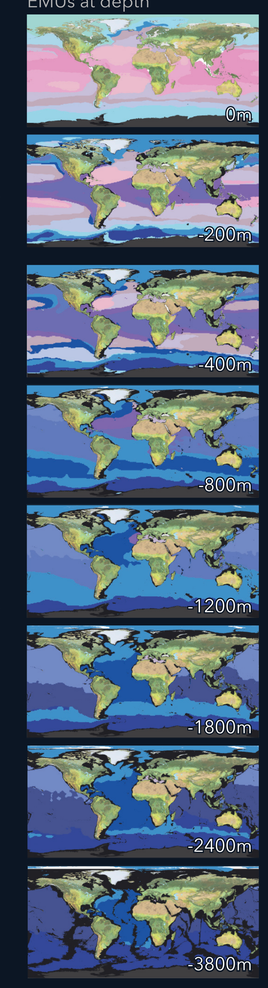
	Minimum	Mean	Maximum	Standard Dev.
Temperature (°C)	4.66	9.83	24.2	2.26
Salinity (unitless)	33.63	34.78	36.94	0.3
Dissolved Oxygen (µmol/l)	0.03	1.56	3.55	0.87
Nitrate (µmol/l)	9.87	30.84	43.71	4.2
Phosphate (µmol/l)	1.26	2.28	3.36	0.31
Silicate (µmol/l)	6.75	31.96	96.02	12.38

EMU water volume (km3)	45,669,990.26
Percent of EMU to Global	3.34%
Unit middle median (m)	390.89
Thickness mean (m)	339.64

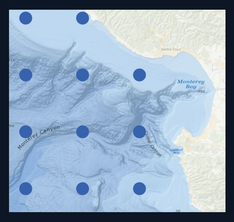


Technical name
- Bathypelagic
- Very Cold
- Euhaline
- Hypoxic
- High Nitrate
- Medium Phosphate
- High Silicate

Common Name
- Deep
- Very Cold
- Normal Salinity
- Low Oxygen
- High Nitrate
- Medium Phosphate
- High Silicate



A Solution for local hi-res EMU's



Global EMUs

... are built on ¼ degree grid from the World Ocean Atlas

... have135 vertical cells

The shallow area has 3 EMUs

The deeper area has 5 EMUs

There are no EMUs in the bay

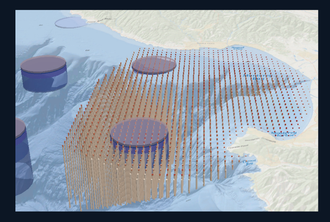
Monterey Bay Hi-Rez EMUs

... are built on a user-defined 500m grid from the World Ocean *Database*.

... have 87,000 vertical cells

... have 5m spacing to 500m depth

... have 100m spacing to >500m depth



Workflow