ocean_data_tools: A MATLAB toolbox for interacting with bulk freely-available oceanographic data

Laur Ferris[1]

1 Virginia Institute of Marine Science

DOI: 10.21105/joss.02497

Software

- Review
- Repository
- Archive

Editor: Kristen Thyng

Reviewers:

- @kakearney
- @castelao

Submitted: 03 July 2020
Published: 30 October 2020

Statement of Need

`ocean_data_tools` simplifies the process of extracting, formatting, and visualizing freely-available oceanographic data. A wealth of oceanographic data (from research cruises, autonomous floats, global ocean models, etc.) is accessible online. However, many oceanographers and environmental scientists (particularly those from subdisciplines not accustomed to working with large datasets) can be dissuaded from utilizing this data because of the overhead associated with determining how to batch download data and format it into easily-manipulable data structures. `ocean_data_tools` solves this problem by allowing the user to transform common oceanographic data sources into uniform structure arrays, call general functions on these structure arrays, perform custom calculations, and make graphics.

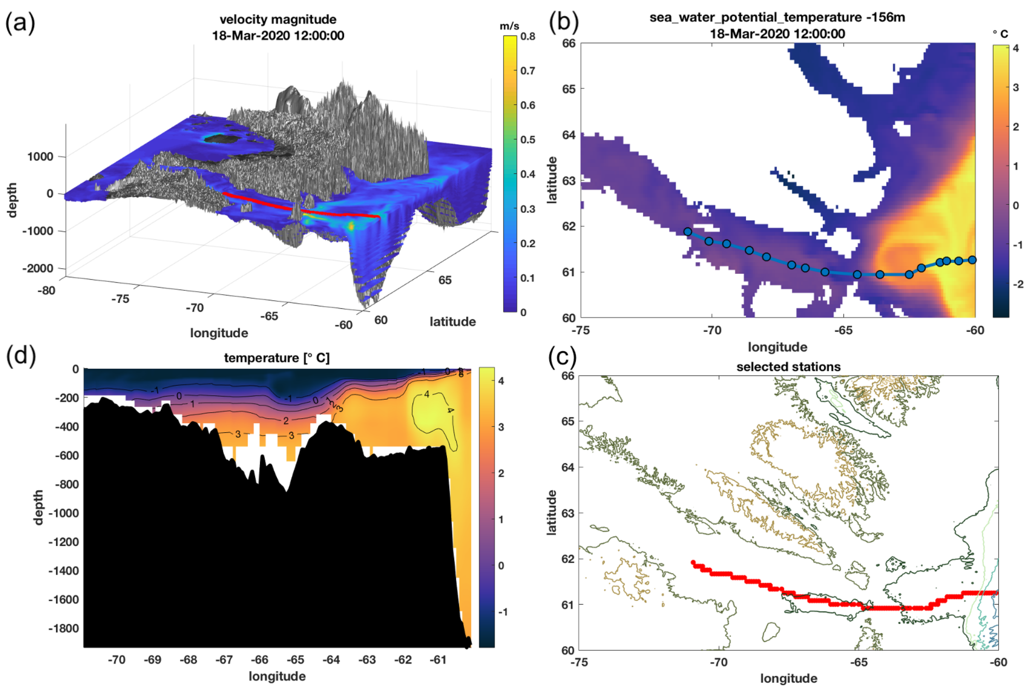

Figure 1: Building a virtual cruise from the Operational Mercator global ocean analysis and forecast system at 1/12 degree with 3D bathymetry (Smith & Sandwell, 1997). Showing (a) a 3D velocity plot created using `model_domain_plot`, (b) virtual cruise selection using `transect_select`, and `model_build_profiles`, (c) coordinates of the resulting uniform structure array, and (d) a temperature section plotted using `general_section` with `bathymetry_section`. Three of the subplots use colormaps from cmocean (Thyng, Greene, Hetland, Zimmerle, & DiMarco, 2016).

Summary

Structure arrays, the common currency of `ocean_data_tools`, are more user-friendly than the native data storage underlying many of the datasets because they allow the user to neatly group related data of any type or size into containers called fields. Both the structure array and its fields are mutable, and data is directly visible and accessible in the Matlab workspace (unlike NetCDF which requires a function call to read variables). Matlab was chosen as the language of choice for this toolbox because it is already extensively used within the oceanographic community. It is also a primary language for much of the community, which is important because this toolbox aims to lower the barrier to entry for using the growing variety of freely-available field- and model-derived oceanographic datasets.

The workflow of `ocean_data_tools` is to build uniform structure arrays (e.g. `argo`, `cruise`, `hycom`, `mercator`, `woa`, `wod`) from raw datasets and call general functions on these structure arrays to map, subset, or plot. Functions with the `\_build` suffix load raw data into uniform structure arrays. Structure arrays are compatible with all `general_` functions, and serve to neatly contain the data for use with custom user-defined calculations or other toolboxes such as the commonly-used Gibbs-SeaWater (GSW) Oceanographic Toolbox (McDougall & Barker, 2011). One application of the `\_build` feature is to create virtual cruises from model output Figure 1. The user draws transects on a map (or passes coordinates as an argument) to build vertical profiles from model data. This may be used as a cruise planning tool, to facilitate comparison of observations with model output, or to support decision-making in underwater glider piloting (using model forecasts to inform ballasting or adjust flight for ocean currents). Some `ocean_data_tools` functions employ `nctoolbox` (Schlining, Signell, & Crosby, 2009).

Table 1: Current ocean_data_tools data sources.

Data Source	DOI, Product Code, or Link
Argo floats	doi:10.17882/42182
Smith & Sandwell bathymetry	doi:10.1126/science.277.5334.1956
IOOS Glider DAC	https://gliders.ioos.us/
MOCHA Climatology	doi:10.7282/T3XW4N4M
HYbrid Coordinate Ocean Model	https://hycom.org
CMEMS Global Ocean 1/12° Physics Analysis and Forecast	GLOBAL_ANALYSIS_FORECAST_ PHY_001_024
GO-SHIP hydrographic cruises	https://www.go-ship.org/
World Ocean Atlas 2018	https://www.ncei.noaa.gov/products/ world-ocean-atlas
World Ocean Database	https://www.ncei.noaa.gov/products/ world-ocean-database

There are several high-quality ocean and/or climate related Matlab toolboxes such as Climate Data Toolbox for Matlab (Greene et al., 2019), those part of SEA-MAT: Matlab Tools for Oceanographic Analysis, and Gibbs-SeaWater (GSW) Oceanographic Toolbox (McDougall & Barker, 2011). However, there are no other documented and designed-to-be-shared toolboxes filling the same data exploration niche as this one. `ocean_data_tools` is unique in encouraging the user to invoke a variety of freely-available data into their exploration and does not expect the user to provide privately-collected measurements or privately-generated model output. It connects users to specific, well-documented data sources Table 1. `ocean_data_tools` has already been used for data exploration in support of scientific publications (Bemis, Tyler, Psomadakis, Ferris, & Kumar, 2020, p. @Crear:2020). This toolbox is built for extensibility; the objective is to welcome contributors and continuously add support for additional datasets such as Remote Sensing Systems products and European Centre for Medium-Range Weather

Forecasts (ECMWF) products. The source code for `ocean_data_tools` has been archived to Zenodo with the linked DOI: (Ferris, 2020).

Acknowledgements

The Virginia Institute of Marine Science (VIMS) provided financial support for this project. I am grateful to Donglai Gong for ongoing mentorship. I thank the many organizations providing freely-available data to the oceanography community including (but not limited to) Argo, the HYCOM consortium, the Copernicus Programme, the International Global Ship-based Hydrographic Investigations Program (GO-SHIP), and the National Oceanic and Atmospheric Administration (NOAA). I also thank the two reviewers for helpful feedback, especially Kelly Kearney for her insightful suggestions. The figure was generated using E.U. Copernicus Marine Service Information. This paper is Contribution No.3960 of the Virginia Institute of Marine Science, William & Mary.

References

Bemis, K. E., Tyler, J. C., Psomadakis, P. N., Ferris, L. N., & Kumar, A. B. (2020). Review of the Indian Ocean spikefish genus Mephisto (Tetraodontiformes: Triacanthodidae). *Zootaxa*, *4802*. doi:10.11646/zootaxa.4802.1.5

Crear, D. P., Watkins, B. E., Saba, V. S., Graves, J. E., Jensen, D. R., Hobday, A. J., & Weng, K. C. (2020). Contemporary and future distributions of cobia, Rachycentron canadum. *Divers Distrib.*, *00*. doi:10.1111/ddi.13079

Ferris, L. N. (2020).*Zenodo*. doi:10.5281/zenodo.4151538

Greene, C. A., Thirumalai, K., Kearney, K. A., Delgado, J. M., Schwanghart, W., Wolfenbarger, N. S., Thyng, K. M., et al. (2019). The Climate Data Toolbox for MATLAB. *Geochemistry, Geophysics, Geosystems*. doi:10.1029/2019GC008392

McDougall, T. J., & Barker, P. M. (2011). Getting started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox, SCOR/IAPSO WG127, ISBN 978-0-646-55621-5.

Schlining, B., Signell, R., & Crosby, A. (2009). nctoolbox. *GitHub repository*. GitHub. Retrieved from https://github.com/nctoolbox/nctoolbox

Smith, W. H., & Sandwell, D. T. (1997). Global sea floor topography from satellite altimetry and ship depth soundings. *Science*, *277*. doi:10.1126/science.277.5334.1956

Thyng, K. M., Greene, C. A., Hetland, R. D., Zimmerle, H. M., & DiMarco, S. F. (2016). True colors of oceanography. *Oceanography*, *29*. doi:10.5670/oceanog.2016.66