

Seascape Ecology

Lab 07 - Data Sources

Ryan Reisinger

2025-11-17

Introduction

Today we'll take a look at some ways to get environmental data for your seascape analyses. These are only a few options, there are many out there; finding them requires a bit of research: Google and patience are your friends. Also, keep an eye out for linked datasets in any papers you read. Right at the end of this tutorial, I have listed some datasets that students used in previous years.

This step (finding data) can sometimes be difficult and frustrating. Go easy on yourself: stay patient and ask for help, from your classmates, me and Josh, or the internet.

Ask me and Josh about different datasets that you might want to use. Usually every new dataset requires a bit of 'wrangling' to get it to work. Be patient with that step—it's an integral part of data science!

Bathymetry data

Using the marmap package

Refer back to Lab 04 where you practiced using the `marmap` package (<https://cran.r-project.org/package=marmap>) to download bathymetry data. You could also download data manually from NOAA (<https://www.ncei.noaa.gov/products/etopo-global-relief-model>) or GEBCO (https://www.gebco.net/data_and_products/gridded_bathymetry_data/).

Environmental data

Copernicus Marine

<https://data.marine.copernicus.eu/products>

How to work with files from Copernicus Marine

<https://marine.copernicus.eu/services/user-learning-services/using-r-work-copernicus-marine-data>

This tutorial by David March has a step-by-step video and the associated materials are available at: <https://atlas.mercator-ocean.fr/s/dCTPqL77Ydposfi>. The NetCDF (.nc) file that David downloads (`download.file()` function) and then reads in (`nc_open()` function) is hosted on his Github page (and is also contained in the training materials), but is originally downloaded directly from Copernicus Marine. For example, you can look at all the products with a global spatial coverage at: <https://data.marine.copernicus.eu/products?facets=areas%7EGlobal+Ocean>

Once you have selected a data product you are interested in, click on the 'data access' button in the toolbar on the left, which load a more detailed page with different download options. You will need to register for a free account first. You could, for example, click on the 'Browse' links in the 'Files' column to access the folder and

raw files directly using your browser (e.g., https://data.marine.copernicus.eu/product/GLOBAL_ANALYSISFORECAST_BGC_001_028/files) or 'Form' links in the 'Subset' column to enter parameters (like spatial and temporal extent) and then download the resulting .nc file. This webpage has more details about how to use the Graphical User Interface to download data: <https://help.marine.copernicus.eu/en/articles/8078281-how-to-download-copernicus-marine-products-using-our-graphical-user-interfaces> (probably the easiest option).

You can explore the dataset visually by clicking on the map icon at top right.

Global Ocean 1/12° Physics Analysis and Forecast updated Daily

Home > Marine Data Store > Product

Data access and mapping services

There are multiple ways to download data from this product:

- If you prefer a graphical tool, click on above.
- To subset data in time and/or space, choose MOTU.
- If you use an OPeNDAP client such as *netCDF4/xarray* (Python), *ferret*, or *MATLAB*, choose OPeNDAP or ERDDAP.
- To download raw files, choose FTP.
- To request maps from QGIS or similar tools, use our [Web Mapping Service \(WMS\)](#).

Dataset	MOTU	OPeNDAP	ERDDAP	FTP	WMS
Daily mean fields from global ocean physics analysis and forecast updated daily	Link	Link	-	Link	Link
Daily mean fields from global ocean physics analysis and forecast updated daily (4)	Link	Link	-	Link	Link
Daily mean fields from global ocean physics analysis and forecast updated daily (5)	Link	Link	-	Link	Link

You could also explore and download products through their MyOcean Pro visualisation tool: <https://data.marine.copernicus.eu/viewer/expert>.

As David says in his tutorial, there is an R package called RCMEMS (<https://github.com/markpayneatwork/RCMEMS>) to download these files directly through R, but it is fairly complex to set up, since you need a Python installation on your machine. Only for advanced users...

Remember, you should read and understand the product overview (e.g., https://data.marine.copernicus.eu/product/GLOBAL_ANALYSISFORECAST_PHY_001_024/description) and you will need to give these details in a report (like your workflow assessment, see the marking rubric) or paper. You can cite the product using the DOI (for example, <https://doi.org/10.48670/moi-00016>).

Once you have worked through David's tutorial, try with your own data product that you choose from Copernicus Marine.

NOAA

ERDDAP


NOAA provides its gridded datasets through a server called ERDDAP.

You can see the datasets (what we called 'products' above) here: <https://coastwatch.pfeg.noaa.gov/erddap/griddap/index.html?page=1&itemsPerPage=1000>. You can use that page to get the data by clicking the 'data' link in the leftmost column for a given product (the column is called 'Grid DAP Data'), choosing the temporal and spatial extent, selecting .nc in the file type dropdown, and then clicking 'Submit'. You can do



an advanced search on the page to find what you're looking for (<https://coastwatch.pfeg.noaa.gov/erddap/search/advanced.html?page=1&itemsPerPage=1000&protocol=griddap>).

For example, I went to the search page, typed “chlorophyll global” in the “Full Text Search for Datasets” box, then clicked on the ‘data’ link for the result with this title: “Chlorophyll-a, Aqua MODIS, NPP, L3SMI, Global, 4km, R2022 NRT, 2003-present (Monthly Composite)”. As for Copernicus Marine above, it’s important to read the product details. You can get that information by clicking on the ‘background’ link in the ‘Background Info’ column. This link should take you directly to the product data access page: https://coastwatch.pfeg.noaa.gov/erddap/griddap/erdMH1chlamday_R2022NRT.html (and the background info. is at <https://oceandata.sci.gsfc.nasa.gov/>, but it’s not so informative in this case).

Take note of the dataset ID found near the top of the page and also in the website URL: ‘erdMH1chlamday_R2022NRT’. You can use this ID later in the `rerddap` package. I left the time dimension as-is, and I left the latitude and longitude as-is. This should give me the latest available global file. Select ‘nc’ in the File type dropdown, and click ‘Submit’.

**ERDDAP**
Easier access to scientific data

ERDDAP > griddap > Data Access Form

Dataset Title: **Chlorophyll-a, Aqua MODIS, NPP, L3SMI, Global, 4km, R2022 NRT, 2003-present (Monthly Composite)**  

Institution: NASA/GSFC OBPG (Dataset ID: erdMH1chlamday_R2022NRT)

Information: [Summary](#) | [License](#) | [FGDC](#) | [ISO 19115](#) | [Metadata](#) | [Background](#) | [Files](#) | [Make a graph](#)

Dimensions	Start	Stride	Stop	Size	Spacing
<input checked="" type="checkbox"/> time (Centered Time, UTC)	2022-11-16T00:00:00Z	1	2022-11-16T00:00:00Z	9	30 days 15h 0m 0s (uneven)
<input checked="" type="checkbox"/> latitude (degrees_north)	89.97916	1	-89.97917	4320	-0.04166667 (even)
<input checked="" type="checkbox"/> longitude (degrees_east)	-179.9792	1	179.9792	8640	0.04166667 (even)

Grid Variables (which always also download all of the dimension variables)

☒ chlorophyll (Chlorophyll Concentration, OCI Algorithm, mg m-3)

File type: [\(more information\)](#)

.nc - Download a NetCDF-3 binary file with COARDS/CF/ACDD metadata.

Just generate the URL:

[\(Documentation / Bypass this form\)](#)

(Please be patient. It may take a while to get the data.)

Be patient once you have clicked submit, it takes a while. This generates a pretty big file (~150 MB) that you can download. You may wish to change the spatial extent (the latitude and longitude) so the download is smaller. I changed mine to -20 - +20 latitude and -10 - +10 longitude for this example.

Notice you can also generate a link by clicking ‘Just generate the URL’. In my case this gave:

[https://coastwatch.pfeg.noaa.gov/erddap/griddap/erdMH1chlamday_R2022NRT.nc?chlorophyll1%5B\(2022-11-16T00](https://coastwatch.pfeg.noaa.gov/erddap/griddap/erdMH1chlamday_R2022NRT.nc?chlorophyll1%5B(2022-11-16T00)

We can read the file into R using the `raster` package.

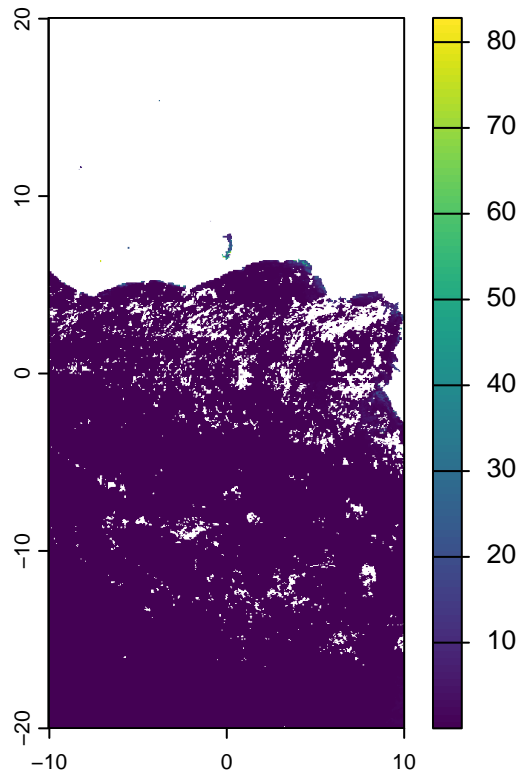
If you can’t get the above step to work, both this file and the one downloaded in the next section are uploaded on Blackboard. You should save them in a folder called ‘resources’ inside your working directory (or modify the path names, removing the `resources` part).

```
library(raster)

## Loading required package: sp
library(terra)

## terra 1.8.70
chl <- terra::rast("./resources/erdMH1chlamday_R2022NRT_2b95_1e81_5998_U1670403378531.nc")

# Plot to check
plot(chl)
```



```
# We could also read it in directly from the URL:
# Note that you will need to create a folder (in your working directory)
# called 'resources'. Or, you can modify the 'destfile =' argument,
# taking out the "resources/" part.
my_file <- download.file(url = "https://coastwatch.pfeg.noaa.gov/erddap/griddap/erdMH1chlamday_R2022NRT_2b95_1e81_5998_U1670403378531.nc",
                        destfile = "./resources/my_chl.nc")

# And read it in
chl <- terra::rast("./resources/my_chl.nc")

# Plot it
plot(chl)
```

A more detailed tutorial is provided here: <https://coastwatch.gitbook.io/satellite-course/tutorials/r-tutorial/1.-how-to-work-with-satellite-data-in-r>

We can also get data from the ERDDAP server using the **rerddap** package (<https://docs.ropensci.org/rerddap/>). The CoastWatch tutorial above includes rerddap tutorials, and rerddap has two vignettes:

<https://cloud.r-project.org/web/packages/rerddap/vignettes/rerddap.html>

https://cloud.r-project.org/web/packages/rerddap/vignettes/Using_rerddap.html

There are also some general tutorials like the following:

https://cran.r-project.org/web/packages/heatwaveR/vignettes/OISST_preparation.html

Ocean Color Data

<https://oceandata.sci.gsfc.nasa.gov/>

You can use the NASA Ocean Color L3 browser: <https://oceancolor.gsfc.nasa.gov/l3/>

Other data sources

If you encounter a product you're interested in using, speak to me or Josh about how to access and use the data for your assessment.

These are some of the datasets that Seascapers have used previously:

- Animal tracking data from Movebank (refer to the animal movement lab for links to Movebank tutorials)
- Coral reef monitoring data from AIMS
- Northern hemisphere sea-ice concentration
- Data on coral beaching from [here](#) and [here](#)
- Sea surface temperature data from Copernicus
- Data on the potential habitat of sardines

Often, papers will publish the data they used to do their work.

Remember, you need to fully cite your data source, and you'll need to report its characteristics (like spatiotemporal extent and resolution). Refer the the marking rubric for the assignment.

Practice

Start playing around with these data sources, and start thinking about what you want to do for your assessment. Josh and I will walk around to answer questions, and next week is a drop-in session where you can bring your ideas and/or datasets for help.