NWEM ERDDAP Tutorial Webinar

Seth Travis
setht1@uw.edu

October 26, 2023

With support from NANOOS
Schedule

• ERDDAP Tutorial
  • NWEM ORCA Moorings
  • ERDDAP – What is it?
  • Demonstrate NWEM ERDDAP to see data availability
  • Discuss NWEM Data Products
  • Discuss ERDDAP programming language interfaces

• NWEM Data Processing
  • Discuss NWEM Data Workflow
  • Discuss processing steps for each data product
NWEM ORCA Moorings

- Northwest Environmental Moorings (NWEM) maintains a series of permanent moorings through Puget Sound, with the goal of providing real-time and long-term environmental monitoring.

- We call these mooring ORCA buoys (Oceanic Remote Chemical Analyzer):
  - Consists of a winch-driven profiling system with a sensor package.
  - Sensor package contains a CTD, as well as a variety of chemical and optical sensors.
Buoy Locations

- ORCA1: Twanoh
- ORCA2: Hoodsport
- ORCA3: Hansville
- ORCA4: Dabob Bay
- NPBY1: Pt. Wells
- NPBY2: Carr Inlet
Typical Cast

• Each cast samples at 4 Hz, with the CTD package moving at a target speed of 25 cm/s
  • A sample is taken every ~6 cm when package is moving
  • For a 100 m cast, this results in ~1600 samples for just the downcast portion of each cast
## Data Coverage

<table>
<thead>
<tr>
<th>Buoy Name</th>
<th>Buoy Depth</th>
<th>Number of Individual Casts (as of 2023-09-12)</th>
<th>First Cast Date</th>
<th>Number of Years of Data Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORCA1: Twanoh</td>
<td>35 m</td>
<td>36,215</td>
<td>2005-Jan-19</td>
<td>18</td>
</tr>
<tr>
<td>ORCA2: Hoodsport</td>
<td>120 m</td>
<td>24,645</td>
<td>2005-Oct-19</td>
<td>18</td>
</tr>
<tr>
<td>ORCA3: Hansville</td>
<td>100 m</td>
<td>15,304</td>
<td>2005-Nov-18</td>
<td>18</td>
</tr>
<tr>
<td>ORCA4: Dabob Bay</td>
<td>105 m</td>
<td>7,161</td>
<td>2010-Jun-07</td>
<td>13</td>
</tr>
<tr>
<td>NPYB1: Pt. Wells</td>
<td>100 m</td>
<td>9,129</td>
<td>2010-Apr-10</td>
<td>13</td>
</tr>
<tr>
<td>NPYB2: Carr Inlet</td>
<td>47 m</td>
<td>24,950</td>
<td>2010-Oct-13</td>
<td>13</td>
</tr>
<tr>
<td>Total Casts</td>
<td></td>
<td><strong>117,404</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Target of at least 1 cast per day
  - This can be higher during the summer, when we have more time for solar to recharge the batteries
- Each cast can consist of 1200-4000 samples, depending on buoy depth
ERDDAP - What is it?

• ERDDAP is a data server platform to distribute & visualize data
  • Allows for temporal and spatial subsetting of data
  • Allows for multiple data file formats (ex. netCDF, csv, JSON, mat, etc.)
  • Visualization tools let you explore the data and save the resulting images in your preferred output (png, pdf, etc)
  • Server is “RESTful”, meaning that the URL used to access the data completely defines the output

• Overall, ERDDAP is a data platform that allows for direct user data exploration, and automated machine access
ERDDAP – Resources

• Coast Watch ERDDAP Tutorial – Best Resource: https://coastwatch.gitbook.io/satellite-course/tutorials/erddap-tutorial
ERDDAP – Web Interface

Front page of the NWEM ERDDAP lists:

• Description of the NWEM program

• Link to the data manual

• Access to the data sets

NWEM ERDDAP: https://nwem.apl.washington.edu/erddap/index.html
ERDDAP – Web Interface

Front page of the NWEM ERDDAP lists:

- Description of the NWEM program
- Link to the data manual
- Access to the data sets

NWEM ERDDAP: https://nwem.apl.washington.edu/erddap/index.html
ERDDAP – Web Interface

Front page of the NWEM ERDDAP lists:

- Description of the NWEM program
- Link to the data manual
- Access to the data sets

NWEM ERDDAP: https://nwem.apl.washington.edu/erddap/index.html
ERDDAP – Web Interface

Front page of the NWEM ERDDAP lists:

• Description of the NWEM program
• Link to the data manual
• Access to the data sets

NWEM ERDDAP:
https://nwem.apl.washington.edu/erddap/index.html
ERDDAP – Web Interface

List of all Data Products:

• Title & Dataset ID
ERDDAP – Web Interface

List of all Data Products:

- Title & Dataset ID
- Data Access Form
  - TableDAP or GridDAP options
  - TableDAP: A list of points with values associated with each Example: Excel Tables
  - GridDAP: Variables on a fixed data grid (with axis variables) Example: Maps of sea surface temperature
List of all Data Products:

- Title & Dataset ID
- Data Access Form
- TableDAP or GridDAP options
- Data Visualization
## NWEM Data Products

### ERDDAP > Search

Do a Full Text Search for Datasets:
- Pu Wells [Search]

11 matching datasets, with the most relevant ones listed first.
(Or, refine this search with Advanced Search 🍂)

<table>
<thead>
<tr>
<th>Grid</th>
<th>DAP Data</th>
<th>Subset</th>
<th>Table DAP Data</th>
<th>Make A Graph</th>
<th>WS Files</th>
<th>Title</th>
<th>Summary</th>
<th>FGDC, ISO Metadata</th>
<th>Background Info</th>
<th>RSS</th>
<th>E mail</th>
<th>Institution</th>
<th>Dataset ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L1 Profile Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L1_profiles</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L2 Pressure Gridded Data - 0.25 dbar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L2_gridded_025</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L2 Pressure Gridded Data - 1.00 dbar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L2_gridded_100</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L3 Climatology - Depth Gridded - 0.25 meter (Incomplete)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L3_depth_climatology_025</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L3 Climatology - Depth Gridded - 1.00 meter (Incomplete)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L3_depth_climatology_100</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L3 Pressure Gridded Data - 0.25 dbar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L3_pressure_gridded_025</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L3 Pressure Gridded Data - 1.00 dbar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L3_pressure_gridded_100</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L3 Depth Gridded Data - 0.25 meter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L3_depthgridded_025</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L3 Depth Gridded Data - 1.00 meter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L3_depthgridded_100</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L4 Anomaly - Depth Gridded Data - 0.25 meter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L4_anomaly_025</td>
</tr>
<tr>
<td>data</td>
<td>graph</td>
<td>files</td>
<td>NPBY1 - Pt Wells - L4 Anomaly - Depth Gridded Data - 1.00 meter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest Environ.</td>
<td>npby1_L4_anomaly_100</td>
</tr>
</tbody>
</table>

The information in the table above is also available in other file formats (.csv, .htmlTable, .ix, .json, .jsonCSVv1, .jsonCSV, .jsonKVP, .mat, .nc, .ncsav, .tqv, .xhtml) via a RESTful web service.
NWEM Data Products

- **L1: Profile Data**
  - Contains every point taken by the CTD package during a cast
  - Contains all the IOOS QARTOD quality control tests performed for every point

- **L2: Pressure-Gridded Data**
  - The downcast portion of each cast is interpolated onto a uniform pressure grid
  - Gridding is done at 1.00 dbar and 0.25 dbar resolution

- **L3: Depth-Gridded Data**
  - Pressure-gridded casts are converted into hydrostatic depths, and re-gridded to uniform depth grids
  - Gridding is done at 1.00 m and 0.25 m resolution
NWEM Data Products

• **L1: Profile Data**
  - Contains every point taken by the CTD package during a cast
  - Contains all the IOOS QARTOD quality control tests performed for every point

• **L2: Pressure-Gridded Data**
  - The downcast portion of each cast is interpolated onto a uniform pressure grid
  - Gridding is done at 1.00 dbar and 0.25 dbar resolution

• **L3: Depth-Gridded Data**
  - Pressure-gridded casts are converted into hydrostatic depths, and re-gridded to uniform depth grids
  - Gridding is done at 1.00 m and 0.25 m resolution
NWEM Data Products

- **L1: Profile Data**
  - Contains every point taken by the CTD package during a cast
  - Contains all the IOOS QARTOD quality control tests performed for every point

- **L2: Pressure-Gridded Data**
  - The downcast portion of each cast is interpolated onto a uniform pressure grid
  - Gridding is done at 1.00 dbar and 0.25 dbar resolution

- **L3: Depth-Gridded Data**
  - Pressure-gridded casts are converted into hydrostatic depths, and re-gridded to uniform depth grids
  - Gridding is done at 1.00 m and 0.25 m resolution
NWEM Data Products

• L3: Climatology
  • Climatological means and standard deviations for each sensor variable are created
  • Climatologies are only considered “complete” if we have at least 6 full years of data over the time period used (2010-2022)
  • Only 3 buoys have “complete” climatologies: ORCA1 – Twanoh, ORCA2 – Hoodsport, and NPBY2 – Carr Inlet
  • Climatologies are made for pressure- and depth-gridded data, at all resolutions: 1.00 dbar and 0.25 dbar for pressure-gridded; 1.00 m and 0.25 m for depth-gridded

• L4: Anomaly
  • Depth-gridded data with the climatological mean for the cast day-of-year removed
  • At 1.00 m and 0.25 m resolution
NWEM Data Products

• L3: Climatology
  • Climatological means and standard deviations for each sensor variable are created
  • Climatologies are only considered “complete” if we have at least 6 full years of data over the time period used (2010-2022)
  • Only 3 buoys have “complete” climatologies: ORCA1 – Twanoh, ORCA2 – Hoodsport, and NPBY2 – Carr Inlet
  • Climatologies are made for pressure- and depth-gridded data, at all resolutions: 1.00 dbar and 0.25 dbar for pressure-gridded; 1.00 m and 0.25 m for depth-gridded

• L4: Anomaly
  • Depth-gridded data with the climatological mean for the cast day-of-year removed
  • At 1.00 m and 0.25 m resolution
ERDDAP – NWEM Demonstration

Let’s walk through some examples of downloading data from the NWEM ERDDAP server:

https://nwem.apl.washington.edu/erddap/index.html
ERDDAP – Download via Data URL

- ERDDAP creates data URLs which can be used to download data

- By understanding the format of these URLs, they can be manipulated to download the data of your choosing

- Example Download From TableDap:
  https://nwem.apl.washington.edu/erddap/tabledap/orca2_L1_profiles.csv?cast_number%2Ctime%2Csample_time%2Csea_water_pressure%2Csea_water_temperature&time%3E=2023-10-24T19%3A00%3A08Z&sea_water_temperature_qc_agg%3C=3
TableDAP: Deconstruct the Example:

https://nwem.apl.washington.edu/erddap/tabledap/orca2_L1_profiles.csv?cast_number%2Csample_time%2Csea_water_pressure%2Csea_water_temperature&time%3E=2023-10-24T19%3A00%3A08Z&sea_water_temperature_qc_agg%3C=3

- Base url: https://nwem.apl.washington.edu/erddap
- ERDDAP Data Type: tabledap
- dataset_id: orca2_L1_profiles
- Desired file format: csv
- Selected Variables: {cast_number, sample_time, sea_water_pressure, sea_water_temperature}
- Constraints: {time%3E=2023-10-24T19%3A00%3A08Z, sea_water_temperature_qc_agg%3C=3}
ERDDAP – Additional Resources

- Coast Watch ERDDAP Tutorial – Best Resource: [https://coastwatch.gitbook.io/satellite-course/tutorials/erddap-tutorial](https://coastwatch.gitbook.io/satellite-course/tutorials/erddap-tutorial)

- Python – Direct Interface Client:
  - erddapy: [https://ioos.github.io/erddapy/](https://ioos.github.io/erddapy/)

- R – Direct Interface Client:
  - rerddap: [https://cran.r-project.org/web/packages/rerddap/index.html](https://cran.r-project.org/web/packages/rerddap/index.html)

- Matlab – It may be possible if you have the Mapping Toolbox
Any questions?

Seth Travis: setht1@uw.edu
L1: Profile Data
- Contains every point taken by the CTD package during a cast
- Contains all the QARTOD tests performed for every point

L2: Pressure-Gridded Data
- The downcast portion of each cast is interpolated onto a uniform pressure grid
- Gridding is done at 1.00 dbar and 0.25 dbar resolution

L3: Depth-Gridded Data
- Pressure-gridded casts are converted into hydrostatic depths, and re-gridded to uniform depth grids
- Gridding is done at 1.00 m and 0.25 m resolution

L3: Climatology
- Climatological means and standard deviations for each sensor variable are created
- Climatologies are only considered “complete” if we have at least 6 full years of data over the time period used (2010-2022)
- Only 3 buoys have “complete” climatologies: ORCA1 – Twanoh, ORCA2 – Hoodsport, and NPBY2 – Carr Inlet
- Climatologies are made for pressure- and depth-gridded data, at all resolutions: 1.00 dbar and 0.25 dbar for pressure-gridded; 1.00 m and 0.25 m for depth-gridded

L4: Anomaly
- Depth-gridded data with the climatological mean for the cast day-of-year removed
- At 1.00 m and 0.25 m resolution
Data Initial Workflow

Use Archived Raw HEX Data from Buoy

L1: Individual Casts

L2: Pressure-Gridded Data
(1.00 dbar & 0.25 dbar)

L3: Depth-Gridded Data
(1.00 m & 0.25 m)

L3: Pressure-Gridded Climatologies
(1.00 m & 0.25 m)

L3: Depth-Gridded Climatologies
(1.00 m & 0.25 m)

L4: Anomaly Data
(Depth-gridded only)

Data is added to ERDDAP database
Data Ingestion Workflow

Download Raw HEX Data from Buoy

L1: Individual Casts

L2: Pressure-Gridded Data (1.00 dbar & 0.25 dbar)

L3: Pressure-Gridded Climatologies (1.00 m & 0.25 m)

L3: Depth-Gridded Data (1.00 m & 0.25 m)

L3: Depth-Gridded Climatologies (1.00 m & 0.25 m)

L4: Anomaly Data (Depth-gridded only)

Data is added to ERDDAP database
Typical Cast

• Each cast samples at 4 Hz, with the CTD package moving at a target speed of 25 cm/s
  • A sample is taken every ~6 cm when package is moving
  • For a 100 m cast, this results in ~1600 samples for just the downcast portion of each cast
Typical Cast

• Sensor calibrations applied to each cast

• Sensor alignment performed on every cast
  • Needed to ensure sensor measurement corresponds to correct water parcel (ex. Salinity spiking)

• All these parameters are stored in the individual netCDF files for each cast
Data QA/QC

• For every cast, IOOS-standard QARTOD tests are performed, assigning a QC flag value to every point
  • 1 = good, 3 = questionable, 4 = bad, 9 = untested
  • QARTOD tests performed:
    • Spike test, flat-line test, gross-range test, climatology test
    • Check NWEM ORCA Data manual for the thresholds we use

• We perform a few additional tests that we’ve developed in-house. If any of these fail, the whole cast “fails”
  • Ensure that CTD port is not clogged by making sure T-S space on downcast is same as on upcast
  • Ensure that non-CTD based sensors aligned on up- and downcasts
  • Check for density inversions in the dataset
  • Make sure that pressure data is good
Data QA/QC

- T-S Alignment Test:
  - This test checks to ensure that the temperature-salinity signature on the up- and downcast are the same

Correctly aligned
Data QA/QC

- **T-S Alignment Test:**
  - Pump port for CTD sampling can get clogged.
  - When this happens, the values that the CTD samples look realistic, but seem odd.
  - By testing if the up- and downcast have the approximately the same temperature-salinity (T-S) signal, we can check to see if the CTD is functioning correctly.
Data QA/QC

- Density Inversion Test:
  - We check for any density inversions that occur in a profile
  - Small density inversions could be detected, and shouldn’t be discounted

- We look for large density inversions, or inversions that occur over a significant portion of the water column
- If inversions are detected, we assume the whole profile is suspect, and flag as bad
Pressure-Gridding

- When gridding the data, we first smooth the data with a rolling centered mean with a window width of 1-grid size
  - Ex. 0.25 dbar smoothing window for 0.25 gridded-data
- Bad data (QC_flag = 4) is excluded from the smoothing
- The smoothed data is then interpolated onto the fixed pressure grid
Pressure-Gridding

- QC flags are propagated through by checking all QC flags with +/-0.5 grid-size
  - Ex. For 1.00 dbar gridding, the point at 10.0 dbar uses data between 9.5-10.5 dbar
- If all QC flags within the window = 4, then the new QC flag = 4
- If there is no data within the window, the new QC flag = 3
- If the QC flag for >25% of points in the window = 3, then the new QC flag = 3
- For all other cases, assume the data is good, and the new QC flag = 1
Gridded Data
Depth-Gridding

• Pressure is roughly equivalent to depth, but not exactly

• Depth is a more natural coordinate for many users

• To convert our pressure-gridded data to depth-gridded data, we:
  • Calculate a hydrostatic depth from our pressure-gridded for every cast: \( d(p) = \int \delta p / \rho(p)g \)
  • Interpolate the hydrostatic depth onto a fixed depth grid
Climatologies

- Climatological means and standard deviations for each sensor variable are created

- Time period used is 2010 – 2022; Only consistent time period across all buoys

- Climatologies are only considered “complete” if we have at least 6 full years of data

- Only 3 buoys have “complete” climatologies: ORCA1 – Twanoh, ORCA2 – Hoodsport, and NPBY2 – Carr Inlet

- Climatologies are made for pressure- and depth-gridded data, at all resolutions: Pressure: 1.00 dbar and 0.25 dbar; Depth: 1.00 m and 0.25 m
Climatologies

1. First, we average all gridded casts taken on any given day
   • Prevents biasing against days with higher sampling, such as during the summer

2. Find the day-of-year for every cast

3. Average the all data by z-level (pressure or depth) and day-of-year
Climatologies

4. Perform a 2-D convolution smoother

- Using a 2\(^{nd}\) order Lanczos filter, with a 6-week time-dimension half-width and 2.5 z-units (dbar or m) z-dimension half-width

- Use periodic boundaries in the time-dimension (i.e., smoother loops around from December to January)
Climatology Data

ORCA Twanoh @ 5 meters

ORCA Twanoh @ 25 meters
Anomalies

• Anomaly data is only calculated from depth-gridded data

• To find anomalies, we take the day-of-year for any given cast, and subtract of the climatological mean for the same day-of-year
Anomaly Data
Data Access

• All data has been posted to an ERDDAP server, hosted on the NWEM server: [https://nwem.apl.washington.edu/erddap/index.html](https://nwem.apl.washington.edu/erddap/index.html)
  • All buoys have L1-L4 data products, with 1.00 and 0.25 gridding for L2-L4 products (11 products per buoy)
  • Immediately after data processing, all data is added to the ERDDAP data directory
  • The data directory refreshes itself:
    • Every 15 seconds for changed files
    • Every 15 minutes for added files

• Typical time from completing sampling cast to posting on ERDDAP is <6 hours; mostly limited by how often we are able to connect to buoys
Any questions?

Seth Travis: setht1@uw.edu