Example Data Evaluation Procedure

The following is a generalized procedure that can be used to evaluate a dataset of interest. This is more a list of suggested steps and questions, rather than a hard-fast process. How much (or how little) evaluation is necessary, really depends dataset, its provenance, and the requirements of your scientific study. Feel free to adapt this to meet the needs of the data and analysis you wish to perform.

1. Identify the platform(s), instrument(s) and dataset(s) you wish to evaluate.
2. Identify one or more time periods you wish to focus on. Possible ranges could include:
    - An entire deployment
    - Deployment change-overs, where two deployments can be compared with each other during the overlap period
    - Glider flybys
    - Dates of shipboard sampling.
        - OOI cruise data typically includes CTD/DO/Fluorometry, plus chemical analyses: DIC/alkalinity, nitrate, oxygen, Chl-a, and salinity in some cases.
    - Significant events (e.g. storms, earthquakes)
3. Download the data and related metadata from the system, including:
    - Deployment information
    - Included parameters
    - Sensor metadata and vocabulary
    - Instrument calibrations
    - Related cruise info
    - Annotations
4. Plot the full dataset, e.g. as time series, depth profiles, profile time series, and/or relevant specialized plots (e.g. T/S)
    - Do the gross structure & value ranges look correct?
        - If helpful, create a reference map of the dataset to make sure location data is correct and/or to plot the glider track for reference.
    - Are there offsets between deployments?
    - Are there points that need to be filtered out (e.g. outliers or bad data)?
    - Are there any issues with the dataset as a whole?
5. Plot the dataset for one or more periods of interest (e.g. one day, one profile)
    - Do the gross structure & value ranges look correct?
    - Is the plot reasonable for the time/date/location for the selected parameter?
    - Are there any issues with the dataset specifically during the period of interest?
6. Are there any related shipboard or environmental datasets you can include in your review?
    - Are there any related cruise reports available?
    - Are there any relevant CTD casts or analytical (water sample) data available?

    - Are there any additional datasets from nearby profilers, gliders, moorings or satellite instruments during the time period of interest?
7. Compare the dataset with shipboard data or other environmental data (e.g. cruise, satellite or neighboring buoy/glider data)
    - Do the gross structure & value ranges look correct?
    - Is the plot reasonable for the time/date/location for the selected parameter?
    - Are there any issues with the dataset compared to the reference dataset(s)?
8. If possible, compare the downloaded dataset with the available raw data
    - Are there any issues with the raw data files or the dataset itself?
    - If the raw data looks okay, are there any noticeable errors that occurred during data processing?
9. Summarize notable results or actions needed
    - Are there any issues with the dataset or metadata that need to be corrected to produce usable data?
    - Are there any issues with the dataset or metadata that should be corrected to meet community standards?
    - Are there any anomalies in the dataset (good or bad) that should be noted?
    - Are there any annotations that should be added to the dataset / data system?
    - Are there any additional issues with the metadata?