

Data Management Plan

David Nicholson, Samar Khatiwala and Rachel Stanley

I. Types of data

Our proposed research will include both a compilation of previously made field and laboratory measurements as well as producing numerical model output as the result of model simulations. Observational data will include the abundance of noble gases (He, Ne, Ar, Kr and Xe) as well as N₂ dissolved in seawater samples. The noble gas data will be coupled with relevant hydrographic data such as temperature and salinity. Noble gas data utilized in this work was/will be produced by a number of investigators including previous work of PIs Nicholson and Stanley. In total, the observational database will encompass over 1000 sample locations (with each location typically measured in duplicate or triplicate). This proposal will be responsible for working with a number of investigators from several institutions to gather existing observational data and compile them into a common database to be used in model investigations. Our research will also produce modeling tools, for example, matrices that represent the circulation of certain ocean general circulation models (OGCMs) in addition to simulated fields of noble gas distributions.

II. Observational Data and Metadata

We will be working with the Biological and Chemical Oceanography Data Management Office (BCO-DMO) located at the Woods Hole Oceanographic Institution (WHOI). The BCO-DMO was established to serve the data management needs of biological and chemical oceanographers. With the help of the staff at BCO-DMO we will develop a common data format and metadata standards for noble gas observations based on the BCO-DMO “Data Management Best Practices” document. Data we will seek to include will be latitude, longitude, depth, date, time, temperature, salinity, pressure and gas abundances. Metadata will include such items as the associated investigators, projects, ship/cruise, hydrographic cast and Niskin bottle numbers, sample methodology, analytical lab and method and method precision.

The responsibility and control of original data will be in the hands of the individual investigators of the original data producing projects. We will provide a streamlined, planned and consistent framework for these investigators to share data. The common data format will create a single access point for a number of independent project datasets. The BCO-DMO utilizes the JGOFS/GLOBEC distributed, object-oriented, relational data management system software that will make these data accessible through any web browser. To enhance ease of use amongst a wide range of users, including observationalists and modelers, data will be available in a number of common formats including text, Matlab and NetCDF file formats. Through the BCO-DMO website datasets can be selected either through a text-based or geospatial interface.

III. Model Data and Metadata

A scientific goal of the proposal is to produce a new parameterization for air-sea gas transfer including the effects of bubbles. We will make these parameterizations available as functions in a number of languages commonly used in scientific computing, including MATLAB, FORTRAN and Python. Model simulations of noble gases and other tracers

calculated using circulation estimates from three ocean models as well as forcing fields will also be archived. All data will be stored in NetCDF format for distribution to interested scientists. We will work with the BCO-DMO to archive these results and make them publicly available once published.

In addition, transport matrices computed from the ocean component of the CESM 1.0 and CM2.1 models will also be made available via a website maintained at LDEO for this purpose (<http://www.ldeo.columbia.edu/~spk/Research/TMM/tmm.html>) .

IV. Long Term Data Archiving

The BCO-DMO provides solutions for short to medium term data management. In the long term, our data will be transferred and archived at the NOAA National Ocean Data Center (NODC).