Data management plan

Collaborative Research: Combining Theory and Observations to Constrain Global Ocean Deoxygenation

Takamitsu Ito[1], Curtis Deutsch[2], and Matthew C. Long[3]

[1]School of Earth and Atmospheric Science, Georgia Institute of Technology
[2]School of Oceanography, University of Washington
[3]Climate & Global Dynamics Laboratory, National Center for Atmospheric Research

Data Types

The modeling data created by this project will be output from the Community Earth System Model (CESM). As such, we will abide by the CESM Data Management Plan, available here: http://www.cesm.ucar.edu/management/docs/data.mgt.plan.2011.pdf. Output will be produced from (1) Development Simulations and (2) Production Simulations. Development simulations entail runs conducted to evaluate model behavior during the development cycle, including for the purposes of tuning and testing performance. Data from these runs does not typically have much use beyond completion of developments and will thus not be archived or made publically available. Data from Production Simulations is what will be used to conduct the proposed research; these data will be made publicly available and archived.

Research products other than outputs from the CESM runs include data analysis products derived from the World Ocean Database 2013 and the computer programs that are used to generate such data product. Data products include interpolated observations for the quantification and visualization of model-data comparison, and spatially averaged variables for further quantitative analyses. Such data products will be archived and be made publically available through the Biological and Chemical Oceanography Data Management Office (BCO-DMO).

Data Format

The model output data will be in netCDF format, and the file size will range from 1 GB to 250 GB. The data files will be named and structured using the following convention:
 CaseName.ComponentName.OutputType.Date.nc
See here for additional information: http://www.cesm.ucar.edu/models/cesm1.0/filename_conventions_cesm.html.
Data analysis products other than the outputs from the CESM will also be stored in the netCDF format, and the file size will be less than 1 GB. The data analysis software will be in commonly used language such as Fortran, Python and MATLAB.

Metadata

File level metadata for each data file will be recorded automatically through the built-in header section of the netCDF file, which is also compliant to the Climate and Forecast (CF) metadata convention.

Access to Data and Data Sharing Practices and Policies

During the project lifetime, all data generated during the project will be stored and available for project use and for other interested parties on a request basis that recognizes proprietary access to the project members. The final results of the project will be moved into the public domain and be made publicly available through the Earth System Grid (https://www.earthsystemgrid.org). While public access is free of charge, a registration process with the Earth System Grid will be required. This is to help in tracking the data's distribution and usage.

Any data analysis products other than the outputs from the CESM and the data analysis software developed during this project will be contributed to BCO-DMO, and we will work with the BCO-DMO staff to manage such data.

Policies for Re-Use, Re-Distribution

Those who are interested in using the project's final data can obtain access via Earth System Grid as described in the above section. Users are expected to cite the project and the corresponding data according to the citations that will be established by the project team.