DATA MANAGEMENT PLAN

1.   **Types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project**

This project will generate laboratory and field data, software (global chemical transport model code and results), and curriculum materials.

2.   **Standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);**

We will follow best practices for recording and archiving laboratory data. All results of laboratory and field experiments and analyses will be stored electronically, including the raw data from analysis and from radiochemical counting. Where instruments generate paper data output, this will also be stored for five years. These data will include quality assurance information such as blanks, calibration curves, replicates and related data. Mason's research group participates regularly in national or international intercalibration exercises, such as the GEOTRACERS Intercalibration, and will participate in a planned methylmercury intercalibration exercise currently being developed by Lamborg at WHOI. All data generated will be stored centrally at the institution at which it was generated. Data will be backed up on computers as well as multiple CDs, flash drives, and/or external hard drives, to ensure redundancy and protection of data integrity. These various storage media are kept in multiple locations and will be properly labeled to ensure clarity. GEOS-Chem model code will be written in the FORTRAN computer language, and model results are produced and archived in GEOS-Chem standard.bpch format which is readable using the publically-available GAMAP code package. All code will be documented and manual information generated using the ProTex automated documentation package.

3.   **Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;**

GEOS-Chem is a community model and as such, code generated as part of this project will be shared with other model users as it is contributed and incorporated into the standard model code. We have no specific privacy, confidentiality, security or intellectual property concerns relevant to this project. The raw data from the laboratory and field experiments are not normally posted on websites in order to avoid issues of inaccuracy since these have not yet been peer-reviewed. However, data will be available to all legitimate colleagues and institutions upon request, as we have done in the past. Once the data is published in the peer-reviewed literature, then the data will be made available publically. These published data are readily accessible by means of databases that mine the published literature, and we intend to continue that policy. Published papers will thoroughly describe all experimental protocols and conditions. Because the PIs believe in openness, transparency, and reproducibility of the scientific process, they intend to share any primary data and relevant samples with colleagues at other institutions, upon

reasonable request and as possible, and also will comply with any regulations from NSF. This project will not generate data that will impact upon national security nor do the PIs produce data that will infringe upon the personal privacy and confidentiality of any individual.

4.   Policies and provisions for re-use, re-distribution, and the production of derivatives; and

We have no current plans for re-use, re-distribution or the production of derivatives.

5.   Plans for archiving data, samples, and other research products, and for preservation of access to them.

Prof. Sunderland will be responsible for modeling data/software and curriculum materials. Model results and code will be archived as part of standard backups of the computer cluster used by this project in the Harvard School of Engineering and Applied Sciences. Results will be archived along with other model products, including analyzed data and data analysis code or methods. These will be archived along with publications in an internal Wiki site of the Harvard Atmospheric Chemistry Modeling Group, which Sunderland co-leads (http://acmg.seas.harvard.edu/index.html). This practice will ensure access to these materials is maintained regardless of the presence of key personnel for 3+ years after the project is complete. The research groups of Mason and Fisher, for example, keep electronic copies of all primary raw data generated in their respective laboratories or at sea. Secondary data generated through data manipulation and extrapolation will also be stored at a central location at the institution where it is generated. The University of Connecticut is in the initial stages of developing a university-wide data management facility and this will be used when completed. We will also investigate whether the data generated by these studies are suitable to be archived in the BCO-DMO (Biological and Chemical Oceanography Data Management Office) database, and if it is possible to do so, we will submit our data and comply with their recommendations with regard to data formatting and metadata generation in all instances where relevant.