Data Management Plan

1. Data types to be produced during research project
Three main data types will be produced during the project:
  i.   Model output from the ROMS shelf-scale simulations: Model fields are 4-dimensional (X, Y, Z, and T) and include physical variables (e.g., velocity, temperature, salinity) and melt rates under included ice shelves. Model output is in NetCDF format and CF 1.0 compliant (see Section 2).
  ii.   Model output from the MITgcm fjord-glacier scale simulations: Model fields are 4-dimensional (X, Y, Z, and T) and include physical variables (e.g., velocity, temperature, salinity) and melt rates. All model output is in NetCDF format and CF 1.0 compliant.
  iii. Observational data: These data have already been collected (co-PI Domack) starting in the 1980's and include high-resolution CTD and mutlibeam bathymetry data.

2. Standards for data and metadata format and content
ROMS is compliant with the NetCDF Climate and Forecast (CF) 1.0 metadata conventions. As such, all input (grid, surface forcing and boundary conditions) and output (physical and biological fields) files associated with ROMS are already following an accepted standard for processing and sharing model information. CF conventions are increasingly gaining acceptance and have been adopted by a number of projects and groups as a primary standard. The conventions define metadata that provide a definitive description of what the data in each variable represents, and the spatial and temporal properties of the data. This enables users of data from different sources to decide which quantities are comparable, and facilitates building applications with powerful extraction, re-gridding, and display capabilities (http://cf-pcmdi.llnl.gov). The MITgcm model fields follow similar compliance standards as ROMS.

Data quality for the historic CTD and multibeam data will be ensured by co-PI Domack. Much of the multibeam data are already archived at National Geophysical Data Center (NGDC), while the CTD data have been processed and quality controlled for use in a different scientific study (on sedimentological studies in the wAP). For submission of the unarchived hydrographic data, we will follow metadata standards from the National Oceanographic Data Center (NODC). Metadata files for all data sets (historic data and summary model files) will be sent to the Antarctic Master Directory, via the USAP Data Coordination Center (http://www.usap-data.org/) in the form of a Directory Interchange Format (DIF) entry. We will follow guidelines for DIF generation from the Global Change Master Directory website (http://gcmd.gsfc.nasa.gov/).

3. Data archiving and access
All model and observation files will be archived locally at each institution (shelf-modeling: ODU; fjord-modeling: UO; data compilations: UCSC and USF) on multiple backups in the form of redundant hard drives (RAID systems available to the PIs). There is no centralized repository for oceanographic or cryospheric model data. If requested, access to the raw, unprocessed, model output will be provided via contact with the lead PI. All model output will, in principle, be readily available for access and sharing as soon as is reasonably possible after processing and will be preserved for at least five years beyond the award period, as required by NSF guidelines. Both model codes, with extensive documentation, are open source and have large, active online communities, available at http://mitgcm.org and http://www.myroms.org, for the MITgcm and ROMS models, respectively.

Observational data will be submitted as soon as possible *during* the project and all PIs will work to ensure open access of all data types before the end of the project period, including appropriate metadata. The PIs have a track record consistent with open collaborations, and in line with that, data and processing codes will be made available through PI websites and by specific request of any colleague. All PIs have included funds for the above efforts in their budgets.