Data Management Plan

The PIs will comply with the data management and dissemination policies described by the NSF Division of Ocean Sciences Sample and Data Policy. The PIs will work with the Biological & Chemical Oceanography Data Management Office (BCO-DMO) to make all data publicly available within two years.

Pre-Cruise Planning

Pre-cruise planning will be done via teleconferencing and pre-cruise meetings. Detailed plans for final station locations, instrument deployments, a water sampling strategy, and sample analysis procedures will be written as a science implementation plan for each cruise. All sampling logs will be recorded in digital format in Excel spreadsheets. Soon after completion of the cruise, a post-cruise assessment report will be prepared. The original underway data will be contributed by the vessel operator to the UNOLS central data repository.

Description of Data Types

Data that will be collected from the deployed instrumentation during the cruises is summarized in Table 1. Physical collections are limited to samples collected with a deployable Rosette that will be filtered, and analyzed for particulate content.

Table 1. Summary of instrumentation deployed as part of this project proposal and expected data format.

Measurement	**Instruments**	**Data format for dissemination**
Turbulence dissipation rates, current profiles	Shear probes, ADCP	CSV data files
Inherent optical properties	Filtered and unfiltered ac-9s, backscatter, chlorophyll meter	ASCII and CSV data files
Water samples	Rosette	CSV data files of contents, microscope images .jpgs
Temperature and salinity	CTD	CSV data files
Small-particle breakup strength	LISST-100X iDAS	CSV and Matlab data files
Large-particle breakup strength	HoloSea iDAS	CSV and Matlab data files of particle size distributions, image .jpgs of individual aggregates, CSV data files of aggregate morphology statistics

Data collected during laboratory experiments will include particle size distribution measurements, microscopic images of particle aggregates, and turbulence information. This data will be maintained in Excel and Matlab data files and image files (jpgs or tiffs) and added to the BCO-DMO repository.

Data outputs from the theoretical model development, including data used for validation, modelling outcomes, and source codes will be made available through BCO-DMO using PDF, Matlab, and Excel file formats.

Data and Metadata Formats and Standards

Field observations will be stored in ASCII files, which can be easily read by different software packages. Field data will include cruise name and number, date, time, latitude, longitude, cast number, depth, wind conditions, and sea state as appropriate. Metadata for the breakup measurements will include pump setting of the iDAS breakup systems, along with time and initial condition of the recording and any visual observations of the particle plumes as appropriate. Diagrams in PDF format of the iDAS systems, the theory behind their operation, and their orientation in the water during sampling will be provided. Metadata will

be prepared in accordance with BCO-DMO conventions and will include detailed descriptions of the collection and analysis procedures.

Data Storage and Access during the Project
During the cruise, data will be stored on personal computers and external hard drives. Duplicate copies of all data will be maintained and stored in separate locations. Upon return from the field, all data will be immediately uploaded to a secure network server that is backed up to a tape archiving system for recovery. Data will be shared using the PI's Box enterprise account and ScholarSphere. Box is a secure, cloud-based file storage and sharing service. All files stored using this service are 256-bit AES encrypted and password protected. ScholarSphere is an open access repository service administered by the University Libraries and Information Technology Services and allows open-access sharing of data and supplementary material between researchers.

All data generated as part of the experimental and theoretical investigations will be maintained and stored using secure network servers. Files will be shared between institutions using the above mentioned Box and ScholarSphere services.

Mechanisms and Policies for Access, Sharing, Re-Use, and Re-Distribution
Immediately after completion of the research cruises, underway data and metadata will be submitted to the Rolling Deck to Repository (R2R) project. All data sets produced from this work will be made available through the BCO-DMO repository within two-years of the date of collection or generation. The PIs will work with the BCO-DMO data managers to make the project data available online in compliance with the NSF OCE Sample and Data Policy. No access restrictions will be implemented once submitted to the public repositories.

Data produced by this project may be of interest to oceanographers, biologists, climate scientists, and engineers. We will adhere to the standards, policies, and provisions for data and metadata submission, access, re-use, distribution, and ownership as described by the BCO-DMO Terms of Use.

Model Code
Model code supporting publications will be archived and made publicly available through a Github website, maintained by the PI's University. Code will be fully documented using frameworks such as Doxygen or Sphinx.

Plans for Archiving
All data will be archived through R2R and the BCO-DMO. The PIs will work with these programs to ensure data is archived appropriately with complete documentation.

Roles and Responsibilities
Each PI will be responsible for sharing his/her data among the project participants in a timely manner through the resources described above. The lead PI will oversee all in-lab experimental data generation and field data measurements. The secondary PI will lead the theoretical model development and implementation. The lead PI will coordinate the overall data management and sharing process (including any data obtained by unfunded collaborators) and will submit the project data to the BCO-DMO.