

Data Management Plan

In accordance with the NSF data and information policies, all data generated from this project will be submitted to established open archives, described below, to ensure that data is permanently archived and accessible to the scientific community. Processed, calibrated, corrected data will be submitted to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) for CTD data, chemical parameters, abundance, and biomass analyses; raw sequence data will be submitted to NCBI SRA. In accordance with NSF Division of Ocean Sciences Data and Sample Policy all data will be uploaded to public archives within two years of data generation. Data formats and standards as well as metadata formats will follow the guidelines established by BCO-DMO and NCBI SRA as appropriate for each dataset. Further information is detailed below in five sections governing the collection, preservation, and distribution of data.

1) Types of data to be collected:

Samples will be collected from the California Current ecosystem during the proposed field campaign outlined in the *Project Description*. Our analyses will generate depth profiles of the following data:

Table 1. Project data generated and archival database

	Abbreviation	Parameter	Data acquisition method	Archival database
Hydrographic data	T	temperature	CTD sensor	BCO-DMO
	S	salinity	CTD sensor	BCO-DMO
	O ₂	dissolved oxygen	CTD sensor	BCO-DMO
	Chl	chlorophyll fluorescence	CTD sensor	BCO-DMO
	PAR	photosynthetically available radiation	CTD sensor	BCO-DMO
Energy rates	ΔATPa	ambient light ATP phosphorylation rate	32PO ₄ incubation	BCO-DMO
	ΔATPd	dark ATP phosphorylation rate	32PO ₄ incubation	BCO-DMO
	R	Respiration rate	O ₂ drawdown	BCO-DMO
Energy stocks	[ATP]	ATP concentration	HPLC	BCO-DMO
	[TAG]	triacylglycerol concentration	HPLC-MS	BCO-DMO
Community	Structure	16S/18S DNA	Illumina sequencing	NCBI
	Function	16S/18S DNA	paprica analysis	NCBI
Biomass rates	3H-BP	3H-leucine bacterial production	3H-Leu incubation	BCO-DMO
	32P-lipid	phospholipid turnover rate	32PO ₄ incubation	BCO-DMO
	14C-PP	primary production	14C-bicarbonate incubation	BCO-DMO
Biomass stocks	CC	cell counts	flow cytometry	BCO-DMO
	POC	particulate organic carbon	elemental analyzer	BCO-DMO
	PN	particulate nitrogen	elemental analyzer	BCO-DMO
	PartP	particulate phosphorus	hydrolysis and spectrophotometry	BCO-DMO
	[P-lipid]	membrane phospholipid concentrations	HPLC-MS	BCO-DMO
	[Chl]	size fractionated chlorophyll	fluorometry	BCO-DMO
Nutrients	DIN	dissolved inorganic nitrogen	nitrate + nitrite by autoanalyzer	BCO-DMO
	DIP	dissolved inorganic phosphorus	MAGIC and spectrophotometry	BCO-DMO
	Si	dissolved silicate	autoanalyzer	BCO-DMO
	DOC	dissolved organic carbon	total organic carbon with N analyzer	BCO-DMO
	DON	dissolved organic nitrogen	total organic carbon with N analyzer	BCO-DMO
	DOP	dissolved organic phosphorus	hydrolysis and spectrophotometry	BCO-DMO

2) Standards for data and metadata format:

Before the cruise we will develop a sample plan that includes station locations, water budgets, incubator use, shipboard sample processing and sample storage. During the cruises we will keep detailed logs of sampling events and generate a table of samples collected, logging their collection location, date,

time, CTD cast number (when appropriate), sample treatment, and state of processing or preservation (processed shipboard, stored frozen at $-20\text{ }^{\circ}\text{C}$ or $-80\text{ }^{\circ}\text{C}$, etc). Immediately following the cruises we will write a cruise report detailing the samples collected, data generated, and plans for sample processing. The table of samples generated during the cruise and cruise report will be used to track the processing of each sample and keep the critical metadata associated with the ultimate data generated.

During sample collection necessary blanks and background samples will be collected, including killed controls for biological rate measurements, to enable proper quantification of background signal and assess potential contamination. Quantification of biomass stocks and nutrient concentrations will be calculated from instrument signals using response factors (signal/mole compound) determined by analyses of standard curves, prepared using commercially available purified individual analytes at a range of concentrations. Analysis of standard curves will be done in triplicate to quantify instrument precision. Processed data will report values greater than the limit of quantification and indicate where samples were analyzed but signal was below the limit of quantification.

Sequence data will be uploaded to NCBI SRA as demultiplexed fastq files following standard procedures. For the environmental samples, all relevant hydrographic data (T, S, O₂, Chl, PAR) will be included with the submission as metadata. Care will be taken to ensure consistent sample names between the BCO-DMO and NCBI SRA submissions to allow all other data to be linked. To improve data discovery we will ensure that the NCBI BioProject accession number is included with the BCO-DMO submission.

3) Policies for access and sharing including data security:

The cruise report and sample log will be stored on both a cloud platform (a shared Google Drive with unlimited storage, provided by UC San Diego at no cost to the project) and redundant back-ups on local computer hard drives at both Scripps and RSMAS. The cloud platform sample log will be password protected and will facilitate synchronous access for teams at both RSMAS and Scripps, and will automatically log version histories and provide backup data retrieval. Throughout sample processing and analysis each researcher will maintain a detailed laboratory notebook, following standard lab procedures, which will be kept in the respective labs and will be retained for long-term record keeping. Sample processing protocols will identify critical sample information that will be digitally transferred from the lab notebooks either to a cloud-based database or through digital scans of physical lab notebooks depending on the data type to keep important metadata associated with the final data products.

Processed, calibrated data in final form will be available to the scientific community following publication by submitting the data to archival databases, BCO-DMO and NCBI.

4) Policies for re-use and re-distribution:

Data will be used to write articles for publication in peer-reviewed scientific journals. Full information about the experimental conditions, treatments, analysis methods, and data generated will be included in the papers, providing tables of data where appropriate. Opportunities to provide full datasets as archived online supplementary information (i.e. downloadable spreadsheets) will be utilized when possible to make data available to the scientific community. It is important to note that any data submitted as online supplementary information will also be archived with BCO-DMO or NCBI SRA as appropriate.

5) Plans for archiving data and other research products:

Soon after the completion of the cruise, the original underway data will be contributed by the vessel operator to the UNOLS central data repository at <http://www.rvdata.us/catalog/> managed by the Rolling Deck to Repository (R2R) project. Also, R2R will ensure that the original underway measurements will be archived permanently at NODC and/or NGDC as appropriate for the data type.

After data has been vetted for all necessary corrections, processing, and quality control, the water column data that we generate, along with appropriate metadata, will be contributed to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) and NCBI, making the datasets permanently available online to the scientific community.