Data Management Plan: Oceanic Remote Chemical Analyzer (ORCA) moorings, part of the Northwest Environmental Moorings Lab (NWEM)

I.   Type of data and information created

1. What data will you collect or create in the research?
Contextual statement describing what data are collected and relevant URL (IOOS Certification, f 2)
The Oceanic Remote Chemical Analyzer (ORCA) moorings are autonomous profiling systems that collect near real-time water property and atmospheric data in Puget Sound, Washington State. Data are transmitted from the moorings to a server housed at the University of Washington in Seattle where they are processed, quality controlled, and disseminated.

2. What data types will you be creating or capturing?
Water property data are collected for water temperature, salinity, pressure, dissolved oxygen concentration, chlorophyll concentration, turbidity, nutrient concentrations, pH, and current velocity. The program also collects atmospheric data, including wind speed and direction, air temperature, barometric pressure, humidity, and PAR.

3. How will you capture or create the data?
Describe how the data are ingested (IOOS Certification, f 2)
Data are collected using multiple in situ sensors, then compiled into files on the buoy controllers. Data files are transmitted via cellular links to the server and redundantly stored on the buoy controller until file transfers are verified.

Describe how data are managed (IOOS Certification, f 2)
The data are managed on a dedicated server at the University of Washington. Once transferred from the buoys data are processed, plotted, compiled, and posted to the web using custom scripts. The data are stored in both ASCII and MatLAB formats. Server hard drives are mirrored automatically. Backup to a local external hard drive occurs weekly, with monthly backup to an external hard drive kept offsite. All post-processed profile data are automatically harvested by NANOOS for ingestion and archival into the NANOOS-APL Data Assembly Center (DAC) database (Mayorga et al, 2010) soon after the data files have been updated. This database integrates ORCA data with other regional marine datasets and serves as a distribution point for access via the NANOOS NVS and also the US IOOS network.

Near real-time ORCA data for the previous month are available for visualization and direct download via NVS (Mayorga et al, 2010) immediately after ingestion into the NANOOS-APL DAC database.

1

Data Management Plan: Oceanic Remote Chemical Analyzer (ORCA) moorings, part of the Northwest Environmental Moorings Lab (NWEM)

The NANOOS-APL DAC database is a relational database running on PostgreSQL and taking advantage of the PostgreSQL PostGIS geospatial add-on. It is hosted at APL-UW in a robust server with automatic daily and staggered database dumps and backups. This system is regularly analyzed and enhanced in support of NANOOS' operational and data-archival mission.

Describe the data quality control procedures that have been applied to the data. (IOOS Certification, f 3)

A combination of automated and human quality control procedures are used, with a technician providing the overall data quality oversight, flagging of bad data streams, assurance of automated data processing, sensor calibration, and flow of data streams to the web.

As a NANOOS Observing System provider, we follow industry best practices and manufacturer guidance where applicable, to calibrate, operate, and maintain the equipment used in this effort, and will provide documentation of this upon request.

Further, we maintain equipment inventories, shipping logs, and instrument maintenance history logs, as appropriate, that are available upon request.

4. **If you will be using existing data, state that fact and include where you got it. What is the relationship between the data you are collecting and the existing data?**
   N/A

Data Management Plan: Oceanic Remote Chemical Analyzer (ORCA) moorings, part of the Northwest Environmental Moorings Lab (NWEM)

II.   Expected schedule for data sharing

Adheres to the NOAA Data Sharing Procedural Directive. The System is an operational system; therefore the RICE should strive to provide as much data as possible, in real-time or near real-time, to support the operation of the System. (IOOS Certification, f 4)

Once data have been acquired and processed the complete data set is available online (near-real time, approximately 20 minutes after the data are transmitted), with the posted caveat that quality control is an on-going process.

1. **How long will the original data collector/creator/principal investigator retain the right to use the data before opening it up to wider use?**
    N/A

2. **How long do you expect to keep the data private before making it available? Explain if different data products will become available on different schedules (Ex: raw data vs processed data, observations vs models, etc.)**
    N/A

3. **Explain details of any embargo periods for political/commercial/patent reasons? When will you make the data available?**
    N/A

4/18/2017

Data Management Plan: Oceanic Remote Chemical Analyzer (ORCA) moorings, part of the Northwest Environmental Moorings Lab (NWEM)

III. Standards for format and content

1. **Which file formats will you use for your data, and why?**
   How can the information be accessed? (IOOS Certification, f 2)
   Data are available in multiple file formats.
   ASCII – Individual profiles and weather data are stored in standard text files.
   MatLAB – Profiles are ingested into MatLAB for processing and creating plots. Several MatLAB objects are created in the process, including a full data set for each buoy location as well as yearly data matrices.

2. **What file formats will be used for data sharing?**
   All of the Above.

3. **What metadata/ documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?**
   Buoy name, data type (meteorological or hydrological), date, and cast number (if applicable) are stored in the file name for both ASCII and MatLAB files. ASCII files contains header information for each data column. A header file is distributed with each MatLAB file and includes definitions for each data variable.

4. **What contextual details (metadata) are needed to make the data you capture or collect meaningful?**
   Data quality control processes are such that flagged data are excluded from the total data set posted online. Lists of flagged data, as well as the original data set, are available upon request.

5. **How will you create or capture these details?**
   Flagged data are stored in processing scripts.

6. **What form will the metadata describing/documenting your data take?**
   ASCII text files, available upon request.

7. **Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage)**
   N/A

Data Management Plan: Oceanic Remote Chemical Analyzer (ORCA) moorings, part of the Northwest Environmental Moorings Lab (NWEM)

IV.   **Polices for stewardship and preservation**

1.   **What is the long-term strategy for maintaining, curating and archiving the data?**
Points of contact- Individuals responsible for the data management and coordination across the region (CV's attached); (IOOS Certification f 1.i)
Jan Newton - Principal Investigator, Principal Senior Oceanographer
206-543-9152
janewton@uw.edu

John Mickett – Principal Investigator, Senior Oceanographer
206-897-1795
mickett@uw.edu

Wendi Ruef – Research Scientist/Engineer
206-221-6760
wruef@uw.edu

Identify the procedures used to evaluate the capability of the individual (s) identified in subsection 997.23(f)(1) to conduct the assigned duties responsibly. (IOOS Certification, f 1.iii)
The University of Washington has a process in place for personnel evaluation. All personnel listed have received excellent evaluations.

2.   **Which archive/repository/database have you identified as a place to deposit data?**
Documents of the RICE's data archiving process or describes how the RICE intends toarchive data at the national archive center (e.g., NODC, NGDC, NCDC) in a manner that follows guidelines outlined by that center. Documentation shall be in the form of a Submission Agreement, Submission Information Form (SIF) or other, similar data producer-archive agreement (IOOS Certification, f 6).
Data are archived for 60 days on the Northwest Association of Networked Ocean Observing Systems server. Data will be archived annually at the National Centers for Environmental Information (NCEI).

3.   **What procedures does your intended long-term data storage facility have in place for preservation and backup?**
Local redundant HDD storage at the University of Washington lab, with redundant offsite HDD storage. The NANOOS-APL DAC database is hosted at APL-UW in a robust server with automatic daily and staggered database dumps and backups.

4.   **How long will/should data be kept beyond the life of the project?**
Data are indefinitely stored.

Data Management Plan: Oceanic Remote Chemical Analyzer (ORCA) moorings, part of the Northwest Environmental Moorings Lab (NWEM)

5. **What data will be preserved for the long-term?**
   All data are publicly available and preserved.

6. **What transformations will be necessary to prepare data for preservation / data sharing?**
   Raw data are decoded and formatted, analyzed and quality controlled.

7. **What metadata/ documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?**
   Header files are stored with all MatLAB files.

8. **What related information will be deposited?**
   All processing scripts and configuration files.

4/18/2017

Data Management Plan: Oceanic Remote Chemical Analyzer (ORCA) moorings, part of the Northwest Environmental Moorings Lab (NWEM)

V.   Procedures for providing access

1. **What are your plans for providing access to your data? (on your website, available via ftp download, via e-mail, or another way)**
   Describe how data are distributed including a description of the flow of data through the RICE data assembly center from the source to the public dissemination/access mechanism. (IOOS Certification, f 2 and 4)
   Data are available for download directly from the ORCA website (nwem.ocean.washington.edu/data) after obtaining a password (website provides email link to obtain password; there are no restrictions on obtaining a password). Recent (last 60 days) data are also available for immediate download from the NANOOS Visualization System (NVS, nvs.nanoos.org).

2. **Will any permission restrictions need to be placed on the data?**
   ORCA data are freely available for public use. Password login is required for downloading the full data set, though there are no restrictions on obtaining a password.

3. **With whom will you share the data, and under what conditions?**
   Data are publicly available.

4. **Will a data sharing agreement be required?**
   In general, a data sharing agreement will not be required. However, data should be properly acknowledged.

5. **Are there ethical and privacy issues? If so, how will these be resolved?**
   N/A

6. **Who will hold the intellectual property rights to the data and how might this affect data access?**
   The funding agency & the University of Washington.

VI.   Previous published data