Data Management Plan: Coastal Margin Observation and Prediction (CMOP) SATURN network
(Taken from NOAA Data Sharing Template and adapted for IOOS Certification)

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 Science And Technology University Research Network (SATURN; www.stccmop.org/saturn) collects near real-time biogeochemical environmental data mostly in the Columbia River estuary and surrounding coastal waters of Oregon and Washington. 10 SATURN stations collect interdisciplinary data; of these, 7 are fully managed by CMOP, while CMOP only captures data for the remaining 3. A varying number of stations (historically as many as 18, currently 8) collect physical oceanographic data; all are fully managed by CMOP. Several stations collect data from multiple water levels, and one station collects high resolution vertical profiles.

2. **What data types will you be creating or capturing?**
   The program captures meteorological data (wind speed and direction, barometric pressure, air temperature) and oceanographic data temperature, salinity, current speed and direction, water pressure, chlorophyll fluorescence, turbidity, colored dissolved organic material fluorescence, quantum yield, nitrate, pH, pCO2, conductivity, dissolved oxygen, phycoerythrin fluorescence). Data is primarily collected at fixed locations, but some data is collected from mobile underwater platforms (gliders) and vertical profiling platforms. The glider has sensors to measure pressure, temperature, conductivity, salinity oxygen, cdom, chlorophyll and backscatter and is capable of estimated depth-averaged velocity.

3. **How will you capture or create the data?**
   Describe how the data are ingested (IOOS Certification, f 2.)
   The data are collected by several redundant pathways.
   1. The majority of our buoy data are transmitted via wireless network.
   2. Glider data is transmitted via Iridium network. (Glider is currently offline. Redeployment is currently unscheduled, pending additional glider-specific funding.)
   3. For some instruments, data is recorded internally, and is available upon recovery.

   Describe how data are managed (IOOS Certification, f 2.)
   All data (except real-time glider data) is initially stored on mirrored servers at the Astoria field station in ASCII files. Data is immediately copied to the main CMOP servers in Beaverton, OR, where it is processed and initial QA is applied. Data is stored in ASCII files and in a PostgreSQL database, and in netcdf files. Data is backed up to tape storage quarterly, and netcdf files are archived with NCEI monthly.

Data Management Plan: Coastal Margin Observation and Prediction (CMOP) SATURN network
(Taken from NOAA Data Sharing Template and adapted for IOOS Certification)

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

Initial real-time data QA involves manual flagging of suspect data, and automated flagging of data that fails range tests. Implementation of QARTOD data flagging for variables is planned in 2018.

Quality control measures for stations vary by station design. Several stations that are located at docks are visited weekly for maintenance and calibration with deionized, oxygenated water. Less accessible stations are visited roughly quarterly for instrument calibration checks and maintenance. Available instruments for chlorophyll, cdom, nitrate, and oxygen are cross calibrated in the lab roughly quarterly. Water samples are taken at several stations for laboratory analysis of chlorophyll, CDOM, and nitrate and those results are used for instrument calibration.

QA/QC procedures are described further here:
http://www.stccmop.org/books/qaqc_information/

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

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, processed, and quality controlled, CMOP makes the complete data set available. (Near-real time, approximately 30 minutes after the data are transmitted)

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. Rare exceptions may occur for specific data collected under contract.

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.)**

Data Management Plan: Coastal Margin Observation and Prediction (CMOP) SATURN network

(Taken from NOAA Data Sharing Template and adapted for IOOS Certification)

Privacy: N/A in general. Rare exceptions may occur for specific data collected under contract. Raw data are available in real-time. Processed data are available as soon as processing is completed; schedule varies with the data stream, and funding availability.

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

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)
   SATURN shares data in a variety of file formats.
   1. FM 65 XML - Used for the real-time data push to the NDBC. FM 65 format is described here http://www.ndbc.noaa.gov/decode.shtml.
   2. NetCDF - A self-describing, machine-independent data format that support the creation, access, and sharing of array-oriented scientific data, available from the SATURN site http://data.stccmop.org:8080/thredds/archive.html
   3. ASCII - Text file that are easily read and parsed by people and programs via the web, available from the CMOP website (e.g. http://www.stccmop.org/datamart/observation_network/fixedstation?id=saturn03, Inventory tab)

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?**
   The netcdf files generated for the NCEI archive contain metadata meeting the NCEI gold standard for netcdf file documentation.

4. **What contextual details (metadata) are needed to make the data you capture or collect meaningful?**
   Metadata provided in the NCEI archive files includes instrument type, physical location, time reference, quality assessment flags and flag definitions, program information, and data generator contact information

5. **How will you create or capture these details?**
   SATURN metadata is generated by querying our 'archive' PostgreSQL database and passed through python code for generating NCEI compliant netcdf files.

Data Management Plan: Coastal Margin Observation and Prediction (CMOP) SATURN network
(Taken from NOAA Data Sharing Template and adapted for IOOS Certification)

6. **What form will the metadata describing/documenting your data take?**
   SATURN's data sets are described by detailed metadata in a number of formats:
   - ISO 19115 XML - NetCDF via NCEI archive
   - HTML - CMOP website has documentation http://www.stccmop.org

7. **Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage)**
   NCEI gold standard netcdf file metadata and ISO 19115 metadata are both accepted standards and mandated by the US Federal Government.

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)
   Antonio Baptista - Employee 29 years, Principal Investigator/Program Manager
   baptista@ohsu.edu

   Michael Wilkin - Employee 21 years, Field Program Manager
   wilkin@stccmop.org

   Charles Seaton - Employee 17 years, Data Manager
   seatonc@ohsu.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)
   N/A

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 to archive 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.).
   National Centers for Environmental Information (NCEI) is the federal archive repository. Historic data from SATURN stations are archived monthly and available at NCEI (http://www.nodc.noaa.gov/access/index.html). The archive process was established with the NCEI Submission Information Form (https://goo.gl/AmX8F8).

Data Management Plan: Coastal Margin Observation and Prediction (CMOP) SATURN network

(Taken from NOAA Data Sharing Template and adapted for IOOS Certification)

3. **What procedures does your intended long-term data storage facility have in place for preservation and backup?**
Local redundant HDD storage at the Astoria field station, Beaverton servers, tape backup and NCEI.

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

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?**
NetCDF files have complete metadata and quality control flags.

8. **What related information will be deposited?**
Quality control narrative documents will be deposited with NCEI.

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)
SATURN Access to Data http://data.stccmop.org:8080/thredds/archive.html
Access to realtime and historical data via NCEI compliant netcdf files will be provided through the CMOP thredds server. (Currently, access to realtime data is provided in the form of netcdf files that are not NCEI compliant, but that will change in the near future)

2. **Will any permission restrictions need to be placed on the data?**
SATURN data and products are freely available for public use, unless otherwise noted. Acknowledgment of use is encouraged. Unless otherwise noted on the CMOP web site, and as appropriate by the use, the preferred acknowledgment includes a link to the CMOP homepage and a citation be made to:

Data Management Plan: Coastal Margin Observation and Prediction (CMOP) SATURN network
(Taken from NOAA Data Sharing Template and adapted for IOOS Certification)

Baptista AM, Seaton C, Wilkin M, Riseman S, Needoba JA, Maier D, Turner PJ, Karna T, Lopez JE, Herfort L et al.. 2015. Infrastructure for collaborative science and societal applications in the Columbia River estuary. Frontiers of Earth Science. 9(4):659-682.

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

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?**
   Unless otherwise noted, the collecting university or agency holds the intellectual property rights

VI. Previous published data

The overarching publication that describes the system and its data is:

Baptista AM, Seaton C, Wilkin M, Riseman S, Needoba JA, Maier D, Turner PJ, Karna T, Lopez JE, Herfort L et al.. 2015. Infrastructure for collaborative science and societal applications in the Columbia River estuary. Frontiers of Earth Science. 9(4):659-682.

In addition, consult: http://www.stccmop.org/publications