Data Management Plan: Ferry-based Monitoring Program (FMP)

(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)
   Since May 2010, Washington State Department of Ecology's (Ecology) Ferries for Science Monitoring Program collects near real-time geo-referenced environmental data from Puget Sound and the eastern Strait of Juan de Fuca. These data are collected during daily transits of the private passenger ferry vessel, *Victoria Clipper IV*, as it travels between Seattle, WA and Victoria, B.C. Ecology collaborates with the Applied Physics Laboratory at University of Washington and with Integral Consulting, Inc. Data are transmitted from the ferry vessel to a Digital Ocean cloud server for data processing and dissemination.

2. **What data types will you be creating or capturing?**
   The program captures hydrographic data for temperature, salinity, and bio-optical measurements of chlorophyll fluorescence, turbidity, and colored dissolved organic matter at 5-sec intervals, via the VCIV ferry vessel.

3. **How will you capture or create the data?**
   Describe how the data are ingested (IOOS Certification, f 2.)
   The data are generated using a Turner Designs C3 optical fluorometer, Teledyne Citadel thermosalinograph, and global positioning system (GPS). The current method for data ingestion differs from the method used before July 2014. Both methods are described below.

   May 2010 to January 2014:
   1. Coordinates and measurements were transmitted from sensors to an external data logging system. The logging system was located on board the ferry vessel and included a laptop and a serial bus to accommodate multiple data I/O cables. The laptop had proprietary software and Python scripts to retrieve and store data into daily files, one for each sensor (Level 0 files).
   2. Once a day, a scheduled task ran software routines on a secure network server at Ecology headquarters and contacts the external data logging system and laptop to retrieve daily files for storage and subsequent data processing. The server was backed up daily following agency policy.
   3. Daily files from each sensor (Level 0) were processed into daily, geo-referenced files (Level 1) using MATLAB scripts.
   4. In July 2014, data retrieval, storage and data management activities were transferred to a server owned by a cloud computing service provider, DigitalOcean. All historical Level 1 files were grouped into months and posted

onto the Digital Ocean server. Level 0 files remain on Ecology's designated network server.

July 2014 to present:

1. All sensor measurements are transmitted to external data loggers located on board the ferry vessel. Each logger has a Secure Digital (SD) memory card with Wi-Fi capability. Daily files, unique for each sensor (Level 0), are stored on the cards.
2. Once a day, a scheduled task runs software scripts on the DigitalOcean server and contacts the external data loggers' SD cards through an internet connection provided via an on-board cellular router with a static IP, then retrieves the files for storage and subsequent data processing. When a card becomes full, the oldest file is removed as the latest file is written onto the card.
3. For this time period, original uploaded files, Level 0 (copies of uploaded files), and Level 1 files are stored on the DigitalOcean server as NetCDF files.

All data:

1. Level 2 NetCDF files are processed from Level 1 files by spatially gridding data using a system created by Ecology's Water Quality Program for assessing marine water quality in order to determine 303(d) listings under the federal Clean Water Act. Each grid is approximately 1 km. by 1 km.
2. Level 3 NetCDF files are generated after all QC testing has been conducted, and include a final dataset with errors flagged and overall data quality annotated in the NetCDF file as appropriate.

Describe how data are managed (IOOS Certification, f 2.)

May 2010 to January 2014 data that remain on Ecology's designated, secured network server are managed by Ecology and archived according to agency policy.

July 2014 to present Level 0 data and all Level 1 and Level 2 data are managed on the DigitalOcean server account by Integral Consulting, Inc. This includes processing Level 0 files into Level 1 and Level 2 NetCDF files.

File inventory of data on both Ecology's server and the DigitalOcean server are managed by Ecology.

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

Data from the sensors are run through a set of automated and manual (staff-driven) quality control procedures based on criteria developed from QARTOD guidance (https://ioos.noaa.gov/ioos-in-action/oceanic-optics/), sensor manufacturer recommendations, in conjunction with location specific and climatology tests to ensure that measurements meet and pass these tests. If any sensor measurements appear to be

unusual or fail QC criteria, they are investigated, and corrected if legitimate conditions allow. All data are flagged ”pass” or ”fail”. In addition, routine data reviews of statistical summaries and plots are conducted. Results and conclusions from these data reviews also generate further QA/QC information for the data.

Ecology also conducts instrument-specific testing and control. These are described in the program’s Quality Assurance Monitoring Plan: https://fortress.wa.gov/ecy/publications/SummaryPages/1503115.html. An agency Standard Operating Procedure (in process) describes the tests more specifically.

File Error Handling:
In the event that one or more sensor data files are not uploaded, Ecology staff will investigate to determine a possible cause such as power interruptions, dislodged equipment, poor data transmissions, or failing equipment due to age and environmental exposure. A remedy will be conducted as soon as feasible.

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: Ferry-based Monitoring Program (FMP)

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

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.)

Data are being disseminated on the DigitalOcean server: http://138.68.225.121/VictoriaClipper30/. All data are currently provisional and will be flagged with QA/QC codes as they are reviewed.

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

Data Management Plan: Ferry-based Monitoring Program (FMP)

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

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 provided in two file formats:
   1. ASCII – Text file that are easily read and opened in standard software. Raw Level 0 files are in this format.
   2. NetCDF – A Network Common Data Format with data oriented into arrays and with metadata, including data attributes. Level 1 and Level 2 files use this format for standardization with other NetCDF-formatted scientific and spatial data available online and for data accessibility.

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?**
   Metadata are provided within each NetCDF file and use CF-1.6 conventions. They are also provided in a THREDDS data aggregate located at: http://138.68.225.121:8080/thredds/VC_WaterQuality_Catalog.html and through an OPeNDAP data access form on the same THREDDS catalog.

4. **What contextual details (metadata) are needed to make the data you capture or collect meaningful?**
   Current metadata are stored as global attributes in Level 1 and Level 2 NetCDF files, whether downloaded as a single file or generated through the THREDDS data aggregate. Metadata provide information on:
   1. Company – Identifies who generated the file and manages data flow.
   2. Conventions – Standard definition of metadata to describe data within NetCDF files/queries.
   3. Sensor ID – Manufacturer and model of sensor. This excludes the GPS whose coordinates are merged with other sensor's measurements for Level 1 data.
   4. Sensor serial number – Serial number of sensor whose measurements are in the data file.
   5. Creator – Contact person for data source.

5. **How will you create or capture these details?**
   Automated scripts that process Level 0 data into Level 1 and then Level 2 generate metadata as global attributes in each NetCDF file.

6. **What form will the metadata describing/documenting your data take?**

Metadata will be CF-compliant and within each NetCDF file and any queries generated through OPeNDAP.

7. **Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage)**
   Metadata will follow CF-1.6 conventions which is the standard recommended by Unidata, the maintainer of NetCDF software, libraries, and information.

Data Management Plan: Ferry-based Monitoring Program (FMP)
(Taken from NOAA Data Sharing Template and adapted for IOOS Certification)

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)
Washington State Department of Ecology:

Christopher Krembs, Ph.D. – Employee 8 years, Senior Oceanographer
360-407-6675
christopher.krembs@ecy.wa.gov

Julia Bos – Employee 17 years, Marine Monitoring Coordinator
360-407-6674
julia.bos@ecy.wa.gov

Suzan Pool – Employee 6 years, Marine Ferry Monitoring Technician
360-407-7287
suzan.pool@ecy.wa.gov

Integral Consulting, Inc.:

Brandon Sackmann, Ph.D. – Managing Scientist
360-705-3534 ext. 412
bsackmann@integral-corp.com

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)
Ecology has a process in place for personnel evaluations. These evaluations are on file with Ecology's Human Resources office. Ecology determines that collaborators are qualified to meet the data management needs of the Ferries for Science Monitoring Program.

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.).
At this time, data are stored on the DigitalOcean server. Plans are being developed to archive data with NCEI.

Data Management Plan: Ferry-based Monitoring Program (FMP)

(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?**
   Preservation and back up procedures are managed by Integral Consulting, Inc. and consists of local redundant storage on a secure server as well as storage on DigitalOcean. For future NCEI archiving, we expect to archive data with NCEI monthly, guided by the procedures laid out by the NANOOS SATURN (OHSU CMOP) archiving project and corresponding NCEI Submission Agreement. It is anticipated that data will be submitted as NetCDF trajectory files following the NCEI 2.0 Templates.

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

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

6. **What transformations will be necessary to prepare data for preservation / data sharing?**
   Raw data are decoded and formatted, gridded, 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 will have CF-compliant metadata and quality control flags.

8. **What related information will be deposited?**
   Currently, none. In the near future, supplementary information on naming conventions processing routines and quality assurance products will be included.

Data Management Plan: Ferry-based Monitoring Program (FMP)

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

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)

   Ferry-based Monitoring Program access to data

   1. Ecology's Marine Water Quality Monitoring website: http://www.ecy.wa.gov/programs/eap/mar_wat/index.html
   2. DigitalOcean cloud computing provider for uploaded, Level 0, Level 1, and Level 2 files: http://138.68.225.121/VictoriaClipper30/
   3. THREDDS data catalog for Level 1 and eventually Level 2 files: http://138.68.225.121:8080/thredds/VC_WaterQuality_Catalog.html
      a. Data are filed by dates and are in NetCDF format.
      b. Files are divided by folders which identify sensor that data are from:
         i. Optical fluorometer.
         ii. Thermosalinograph.
      c. Level 1 files are posted on the catalog daily after Level 0 data streams are merged and geo-referenced on the DigitalOcean server.
      d. Level 2 files are not yet being posted on THREDDS, but plans are in place for this to occur.
      e. The THREDDS data aggregate can be accessed through:
         i. OPeNDAP using 2 different methods:
            (1) Data access form. The data access form provides an option for getting ASCII or binary data. It also provides a way to select criteria of one or more parameters. An example of a data access form for the data aggregate of optical fluorometer data is at: http://138.68.225.121:8080/thredds/dodsC/VCIV_L1_C3_2300185/VCIV_L1_C3_2300185.ncml.html
            (2) Original user-generated query to obtain the entire data set by using a data URL in software such as Python, MATLAB, R, etc.
         ii. HTTPServer for downloading a single file.

2. **Will any permission restrictions need to be placed on the data?**

   Data are available for public use with request to reference and acknowledge source. When referenced, please provide:

   1. Standard html, link to Ecology's Marine Water Quality Monitoring webpage for accessing ferry-based data with the following reference:

Data courtesy of Washington State Department of Ecology, http://www.ecy.wa.gov/programs/eap/mar_wat/data.html.

2. Offline references, with appropriate form from the recommended acknowledgments below:
    a. Short form (figure captions, etc.)
    "…data from Washington State Department of Ecology."
    b. Longer form (in text)
    "…data were furnished by Washington State Department of Ecology's Ferries for Science Monitoring Program."
    c. Full form (acknowledgements at conclusion of papers, etc.)
    "…data were furnished by Washington State Department of Ecology's Ferries for Science Monitoring Program in collaboration with University of Washington Applied Physics Laboratory and with Integral Consulting, Inc."

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?**
Washington State Department of Ecology, University of Washington Applied Physics Lab and Integral Consulting, Inc. through an inter-agency agreement and contractual agreement.

VI. Previous published data

Coastal and Estuarine Research Federation 2015:

Bos, J., S. Albertson, C. Krembs, S. Pool, C. Falkenhayn Maloy, and B. Sackmann. 2015. Ecosystem Thresholds in Surface Waters of the Salish Sea using Continuous Measurements from Ferry Sensors. Poster presented at Coastal Estuarine and Research Federation 2015 Conference, Portland, Oregon. Washington State Department of Ecology Publication No. 15-03-041.
https://fortress.wa.gov/ecy/publications/documents/1503041.pdf

Pool, S.S., C. Krembs, J. Bos, and B. Sackmann. 2015. Physical, Chemical, and Biological Conditions during *Noctiluca* Blooms in an Urban Fjord, Puget Sound. Poster presented at Coastal Estuarine and Research Federation 2015 Conference, Portland, Oregon. Washington State Department of Ecology Publication No. 15-03-040.
https://fortress.wa.gov/ecy/publications/documents/1503040.pdf

Salish Sea Conference 2014:

Falkenhayn Maloy, C., C. Krembs, S. Pool, J. Bos, L. Hermanson, S. Helgath, J. Thomson, W. Deppe, and B. Sackmann. 2014. Using Ferries for Marine Water Quality Monitoring in the Salish Sea. Poster presented at Salish Sea 2014 Conference, Seattle, Washington. Washington State Department of Ecology Publication No. 14-03-017.
https://fortress.wa.gov/ecy/publications/documents/1403017.pdf

Ocean Sciences 2014:

Sackmann, B., C. Krembs, S. Pool, J. Bos, M. Keyzers, T. Khangaonkar. 2014. Eyes Over Puget Sound: Producing Validated Satellite Products to Support Rapid Water Quality Assessments in Puget Sound. Poster presented at Ocean Sciences 2014 Meeting, Honolulu, Hawaii.
http://www.eposters.net/pdfs/eyes-over-puget-sound-producing-validated-satellite-products-to-support-rapid-water-quality.pdf

Eyes Over Puget Sound, monthly online publication by Ecology:

http://www.ecy.wa.gov/programs/eap/mar_wat/surface.html