IOOS HF Radar Network **May 2016**

Cloud-based Storage of Range Series Data Files

The need for storage of the range series data files is driven by their utility in

- recreating radial files,
- creating new versions of radial files. This may be in response to, for example, a need to re-analyze HF radar data for tsunami detection or to create a “research quality” or “best of” data set.

The need for robust data archiving has been recognized by the IOOS HF radar community and the IOOS HF Radar Technical Steering Committee for several years. Additionally, discussions at the Radiowave Operators Working Group (ROWG) have sought solutions to fulfill this need.

Nearly all of the approximately 30 institutions operating IOOS HF radars currently store their range files. However, those methods vary from institution to institution. The IOOS Program Office has contracted with Axiom Data Science via AOOS to provide this capability.

Project Objective

The intent of the project is to provide a simple and low-cost method for the IOOS HF radar network operators to archive range series data files and to make them more readily available to users for reprocessing. It should provide a robust and convenient capability for all operators to use. This would ensure that all range files that currently exist and those that are being continually created would be safely stored. There will also be redundancy provided by the Axiom cloud-based workspace.

Eventually, it would provide a single location from which the National Centers for Environmental Information (NCEI) (formerly National Oceanographic Data Center and others) could transfer range series files. A methodology that complies with NCEI protocols to implement that transfer is planned for later years.

Implementation

In response to a request from the IOOS Program Office, volunteers from the IOOS HF radar community formed a working group (WG) to explore various methods to carry out this archiving. Crucially, the WG has spent a significant amount of time to try to determine what metadata and other file types may be need to be stored to ensure that the range series file archiving is comprehensive.

Implementation: Range Series and Related Files

The range series files should be arranged in a directory structure consistent with the following:

/AFFILIATION/SITE/RangeSeries/YYYY/MM/DD

Where AFFILIATION is the regional affiliation, and SITE is the 3 or 4 character site name. Site configuration files used to configure the HF radar during data collection should also be included. On the SeaSonde, examples of these include:

AnalysisOptions.txt

Phases.txt

Site_BARR.log (or other operator note files)

SeaSondeAcquisition.plist

SeaSondeController.plist

MeasPattern.txt

Header.txt

These should be arranged in the following directory structure:

/AFFILIATION/SITE/Configs/YYYYMMDD-YYYYMMDD

where the dates, YYYYMMDD, specify the time period during which the files were applied. Associated files that give the antenna pattern measurements (APMs) and other information related to the processing of the APM also need to be included. These files are:

Loop files

TimeSeries

GPS data associated with signal source

Optionally: site configs (Radial Configs), operator notes, TRAK files from processing

These should be put in a sub-folder:

SITE/APM/YYYYMMDD/

where the date, YYYYMMDD, refers to the date that the APM was conducted (rather than applied to processing).

Implementation First Phase: Historical File Storage

Given the enormous amount of data, (at least for today's internet bandwidth) that already exists in historical range series files, the first step will be for each HF radar operator to transfer their existing range series files and related files to an external hard drive in the directory structure given in the previous section.

These hard drives will then be shipped to Axiom who will then upload the data to the Axiom cloud workspace.

Estimates for data storage of the range files vary from 7 to 30 GB per month per site (approximately 11 TB to 49 TB per year for the entire 135 HFR network). Nearly all sites have been operating for more than 5 years and some as long as 15 to 20 years.

Implementation Second Phase: Ongoing Monthly File Storage

The project also seeks to begin monthly storage of range files and related files from individual radar sites. The expectation is that the Axiom workspace will provide a way for each operator to upload their range files on a monthly basis. The second phase may actually occur, for some operators, in parallel with the first phase. For example, some operators may have so many sites and range files currently stored that getting them onto the external hard drives may require many man-hours of effort spread over months. At the same time, the operator could begin the monthly upload of range files. The implementation includes an automated way of syncing range series files. Axiom workspace supports file sync with rsync, a utility that is supported by a number of operating systems. This will allow scripted, unattended file transfer from the operators to Axiom, which, once properly configured, should require little effort from the operators to sync to the archive.

Other Notes:

While out of the scope of the current project, it would be constructive to have software for automatically organizing data into the appropriate folders. This would be made available via the ROWG emailing list, or from a publicly available code repository such as https://github.com/rowg.