The Impact of Reprocessing Efforts on the Mid-Atlantic's Surface Current Product

Teresa G. Updyke
Center for Coastal Physical Oceanography
Old Dominion University
Norfolk, VA, USA
garner@ccpo.odu.edu

Hugh J. Roarty, Michael Smith, and Laura Nazzaro
Rutgers University
New Brunswick, NJ, USA
hroarty@marine.rutgers.edu

***Abstract*— The Mid-Atlantic Regional Association Coastal Ocean Observing System (MARACOOS) has undertaken a reprocessing effort for their 6-kilometer gridded surface current product which utilizes data from the network of long-range High Frequency (HF) radar systems. Radar data from January 2017 to December 2017 have been re-examined and reprocessed. It is worthwhile to consider how errors in the real-time processing may be corrected by reprocessing and what improvements may be realized in the total vector map product by such an effort.**

Keywords—high frequency radar; surface currents; Mid-Atlantic; quality control

I. Introduction

A network of 17 long-range High Frequency (HF) radar stations contribute radial data to a MARACOOS 6-kilometer gridded surface current data product (Fig. 1). The 5 MHz network covers the 1,000 km of Mid Atlantic Bight Shelf from Cape Hatteras to Cape Cod. The near real-time version of the data product is computed from radial data that is subject to several quality assurance (QA) and quality control (QC) procedures [1,2].

These procedures include quality assurance methods that are conducted on the radar hardware and site visits at least once every six months. Operators periodically check the settings that extract first order sea echoes (or Bragg echoes) from the rest of the radar Doppler spectra. This extracted data is processed to radial data. The goal is to include all first order echo, but to exclude any noise or interference as much as possible. Antenna calibrations are performed typically once a year or when site diagnostics indicate that the calibration file needs to be updated. These calibrations are important for accuracy in "direction finding", the process which places radial vectors into bearing bins within a range ring on the radial grid.

Quality control of radial data involves routine remote inspection of the radial data and system diagnostics. For QC purposes, radial vectors are plotted with the blue/red colormap where blue indicates vectors that are travelling towards the radar and red vectors indicating currents that are travelling away from the radar, consistent with redshift and blueshift from electromagnetic Doppler phenomenon. We utilize the 25-hour mean radial map and a weekly plot of average radial velocity and radial vector count as quick diagnostics for station health. These diagnostics are similar to those of previous researchers. If a station or data type (ideal or measured) is not in agreement with surrounding stations, the operator begins an inspection of the system to look for problems. We have also found that a consistent average radial bearing [2] is an indication of a properly operating station and if this measurement has a step change or becomes erratic then that is an indication of a failure somewhere within the system.

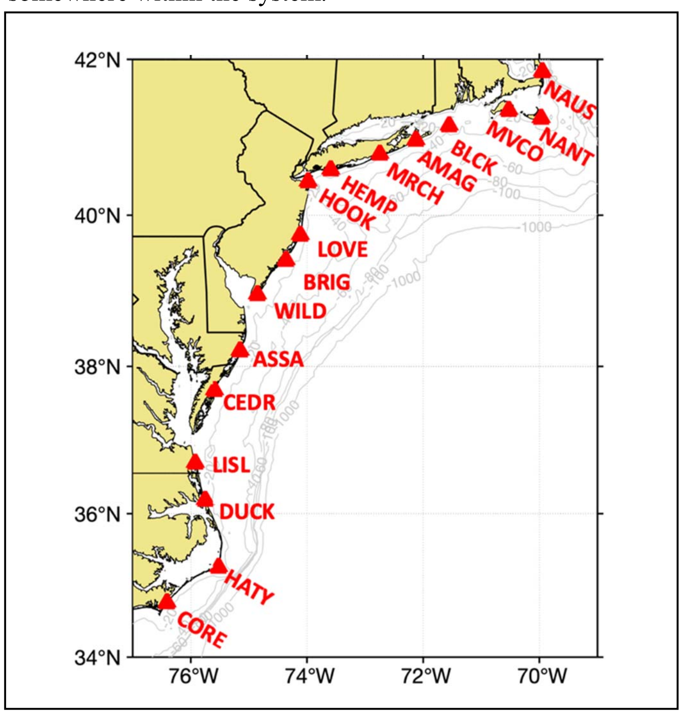

Fig. 1. Locations of the 5 MHz High Frequency radar stations that contribute to the MARACOOS surface current product.

Despite the routine use of QA and QC procedures, the real-time product carries the potential for errors that might be corrected in post-processing. For example, events may occur

This work was funded by NOAA Award Number NA16NOS0120020 "Mid-Atlantic Regional Association Coastal Ocean Observing System (MARACOOS): Preparing for a Changing Mid Atlantic". Sponsor: National Ocean Service (NOS), National Oceanic and Atmospheric Administration (NOAA) NOAA-NOS-IOOS-2016-2004378 / CFDA: 11.012, Integrated Ocean Observing System Topic Area 1: Continued Development of Regional Coastal Ocean Observing Systems.

that invalidate a calibration pattern or damage a radar and there is typically some delay between the time of the event and when the data is pulled from the real-time processing. This delay is often a reason to reprocess a portion of the radial maps from spectra or to choose a different pattern than the one that was used in real-time total processing.

II. Methods

A. Data Review & Radial Reprocessing

The first step in the reprocessing effort was to check for data gaps in the real-time product and collect the missing hourly radial files into the appropriate directories at the data assembly center (DAC) if those files could be found. If the radial files were not found, it was sometimes possible to generate the files from Doppler spectra. A more common problem was that a site communication problem and/or delay in file transfer excluded a radial file from real-time processing, but the file later became available. The real-time software is able to handle up to a 7 day delay in data transfer latency.

The next step involved a review of the logs, diagnostics, radial distributions and averages from each station to check for any events that might signal a change in radial quality. Radial maps surrounding those events were reviewed to check for velocities that did not appear to be realistic or were inconsistent with nearby data or data from neighboring stations. In some cases, data from a station needed to be excluded from total vector processing for a period of time. In other cases, applying configuration changes and reprocessing from spectra significantly improved the radial maps. Radial maps processed using a measured pattern were preferred over maps processed using the assumption of an ideal pattern if a suitable pattern measurement was available [3].

Diagnostics we found particularly useful included sea echo amplitudes and phases, signal to noise, noise floor, and average radial bearing. For example, a step change in the sea echo amplitude of the receive antenna often signals a change in the antenna pattern. In this case, spectra might be reprocessed to radials using a pattern that was measured after the step change. A step change in average radial bearing with no coinciding change in antenna bearing can also signal an error in the configuration that might be corrected in reprocessing. Data from stations reporting low signal to noise and/or high background noise diagnostics for long periods of time were often indicative of equipment failures and were excluded from processing.

Weekly average and distribution plots were reviewed by the QC team (a group including operators, technicians and scientists). Plots for ideal pattern and measured pattern radial maps were reviewed side-by-side. Angular gaps in the weekly radial distribution plots or anomalously placed radials in the average maps were indications that a pattern type might not suitable or that maps might need to be reprocessed from spectra.

B. Radial Quality Control

A major component of the reprocessing effort focused on the implementation of QARTOD quality control (QC) tests [4,5]. Version 1.0 of the QARTOD Manual for Real-Time Quality Control of High Frequency Radar Surface Current Data describes several QC tests that may be performed at different levels of radar data processing including tests for the spectra, radial component and total vector processing stages. The present study has focused on the tests for radial data. The North Carolina station radials had an additional radial metric QC test applied [6]. The QARTOD radial tests that have been applied are listed in Table 1 along with threshold values that were chosen to implement the tests. The QC06 and QC09 test flags apply to the entire file and are reported in the header metadata. In this case, a failure flag means that none of the radials in file are included in total vector computations. The QC07, QC08 and QC10 apply to individual radial vectors within the file and are reported in extra columns for each row of radial data contained in the file.

The syntax test requires that the following metadata be present in the file: file type LLUV, site code, timestamp, site coordinates, antenna pattern type and time zone. Other requirements include 1) the file name timestamp must match the timestamp reported within the file, 2) radial data tables (Lon, Lat, U, V, ...) must not be empty 3) radial data table columns stated must match the number of columns reported for each row 4) site location must be within range: $-180 \leq \text{Longitude} \leq 180$ $-90 \leq \text{Latitude} \leq 90$ and 5) time zone must be Greenwich Mean Time.

TABLE I. QARTOD RADIAL QC TESTS APPLIED

Test Code	Radial QC Test List		
	Test Name	***Suspect Flag***	***Fail Flag***
QC06	Syntax	N/A	see text
QC07	Max Threshold	N/A	velocity > RSPDMAX *RSPDMAX = 300 cm/s*
QC08	Valid Location	N/A	*VFLG = 128*
QC09	Radial Count	RCMIN[a] >= count <= RCLOW[a]	count < RCMIN[a]
QC10	Spatial Median	N/A	velocity > CURLIM *RCLIM=2.1 cells, ANGLIM = 10 degrees, CURLIM[b] = 30 or 50 cm/s*

a. RCMIN and RCLOW are site dependent thresholds.

b. Stations LISL, DUCK, HATY, CORE use 50 cm/s. All others use 30 cm/s.

The radial count test will flag a radial file if it contains less than a minimum number of radial vectors (RCMIN). The RCMIN threshold is site specific and dependent on the number of radial grid cells that are available given 40 range cells, five degrees of bearing resolution and omitting any cells that are invalid (e.g. over or behind land). RCMIN is defined as 10% of the available radial grid cells "rounded" to the nearest 25. RCLOW is defined as 30% of the available radial grid cells "rounded" to the nearest 25.

Each of the QARTOD tests were converted into Python code and are assembled in a GitHub repository (https://github.com/rucool/codar_processing). The assembled radial data is then run through the QC code and new radial files with QC metadata and QC flags are generated.

C. Total Vector Generation

Measured pattern radials were chosen as the preferred radial type for most radar stations for the 2017 reprocessing. DUCK, HATY, and CORE contributed ideal pattern radials. Ideal pattern radials were also used for BRIG (July-Sept) and NANT (Aug-Sep) stations.

Radial vectors that received failure flags for one or more of the QC tests were excluded from total vector processing. After radial reprocessing and radial filtering based on QARTOD test flags, an updated "best" set of radials were used to recompute total vectors maps. Two sets of totals were computed: one set using an unweighted least squares (UWLS) method and the other set using an optimal interpolation (OI) method [7,8]. In this paper, we focus on the UWLS product. At least three radial vectors and a minimum of two contributing radar stations were required to compute a total vector. The search radius for total vector processing was set to 10 kilometers. Vectors with GDOP total error values greater than 1.25 were removed from the vector maps. Computations were performed using the HFR-Progs MATLAB toolbox. An online summary of the reprocessing effort is available at this url: https://marine.rutgers.edu/~michaesm/reprocessed/index.html

D. Evaluation

An analysis of the QARTOD radial QC flag data has indicated which tests flag the most radials and where failure flags occur most often within radar coverage areas.

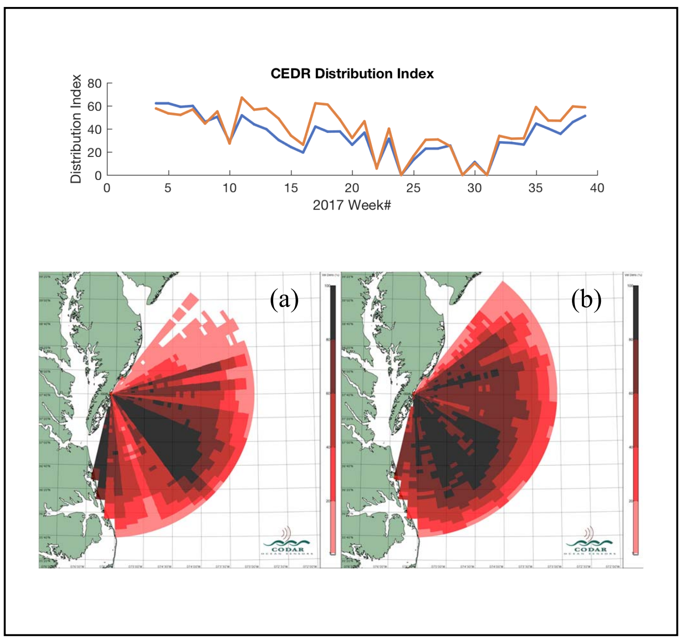

Fig. 2. Top panel: Weekly radial distribution index for real-time (blue) and reprocessed (orange) radials at the CEDR radar station. Bottom panels: Weekly radial distributions at CEDR station using (a) real-time and (b) reprocessed radial maps for April 24 – Apr 30 2017.

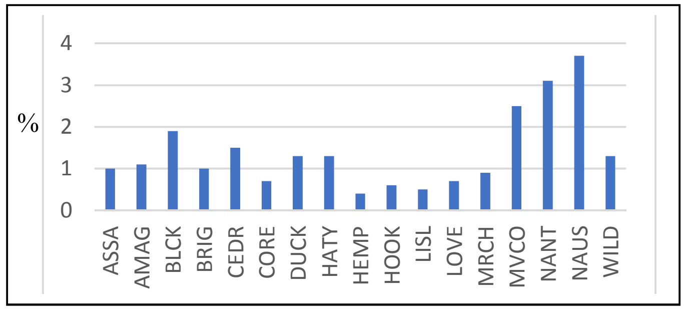

Fig. 3. On average, the percent of radials in a file that fail the spatial median QC test. Only failures for radials at valid locations were considered for this chart.

Finally, the real-time and reprocessed total vector map products were compared to assess the overall impact of the effort. Complex correlations [9] and root mean square differences were calculated between the real-time velocity time series and the reprocessed velocity time series. The impact of reprocessing on monthly averages was also investigated.

III. Results

A. Real-time Radials vs Reprocessed Radials

In 2017, radials were reprocessed from spectra for several radial sites to apply updated patterns and this improved data coverage. For example, CEDR station radials from January 26, 2017 through September 25, 2017 were reprocessed with a pattern that was measured on September 25, 2017. Fig. 2 compares a weekly radial distribution index for the real-time measured pattern maps and the reprocessed measured pattern maps. The weekly index is the number of radial grid cells containing data at least 80% of the time divided by the total number of radial grid cells expressed as a percentage. The total number of grid cells was the count of all cells that contained at least one radial vector during the week. No cells over land were included. The weekly distributions of reprocessed radials often showed improved radial coverage according to this metric. Week 17 of 2017 (Apr 24 to Apr 30) is one example of a dramatic improvement. The bottom panels of Fig. 2 compare the distributions for that week. The reprocessed radials also produced average velocity maps containing less outliers.

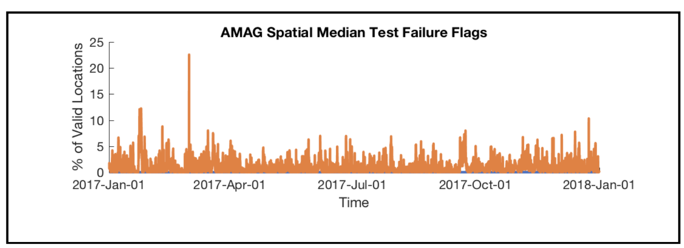

Fig. 4. Percent of radials in each hourly file that were flagged by the spatial median QC test at the AMAG station.

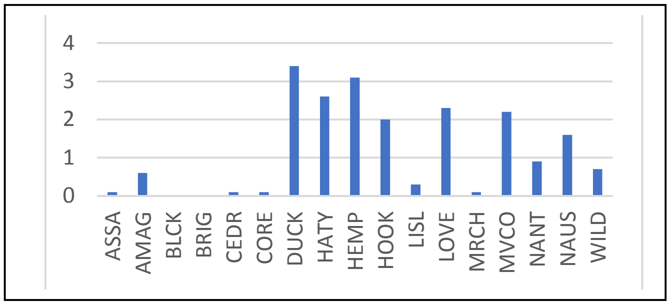

Fig. 5. Percent of available radial files that failed the radial count QC test.

The North Carolina stations (DUCK, HATY and CORE) had been producing radial maps in near-real-time on site using radial metric QC [5] since August 2017 but those versions, called the QCD versions, of the radial maps were not used for total generation in the real-time totals data product until late in 2017. In the reprocessing effort, QCD versions were used for the entire year.

Filling in data gaps due to missing or delayed files resulted in the inclusion of 7,270 more files in the reprocessed data set for 2017.

B. Radial Quality Control Flag Statistics

The real-time product did not make use of radial QC flags. It did however, exclude radials in invalid locations which were identified by codes produced by the manufacturer software.

The syntax test failed for a small set of files at six sites coinciding (surprisingly) with the 2017 switch from Eastern Daylight Time to Eastern Standard Time. Timestamps in the file headers of four hourly files were offset by one hour from the times given in the filenames. Data at all stations were collected in UTC time.

The maximum velocity QC test was put into practice for this reprocessing effort. However, in 2017, radar stations had a limit set on the maximum velocity that was allowed to pass from the spectra to radial stage of processing and this maximum was less than the maximum of 300 cm/s chosen for the QC test. Therefore, this test failed no radials. In future processing, the velocity limits will no longer be set at the radar station or they will be set at a much higher level in this earlier processing stage so that the QC max threshold test flags will become effective and will provide flag statistics that may be analyzed.

On average, the spatial median test failed between 0 and 4% of the radials (at valid overwater locations) in hourly files (Fig. 3). NAUS had the highest average of 3.7%. At most stations, the hourly time series show that the percent of radials with fail flags usually varies within the 0-7% range and includes some spikes up to 12% or 25%. Fig. 4 is an example for a single station. Sudden shifts in amounts of failures often coincide with changes to site configuration settings.

For the spatial median test, maps were created to highlight the most frequently flagged locations at each station. Highlighted bearing spokes near coast or edges of coverage are common in these maps. Most radials in the reprocessed data set were measured pattern radials and the measured radial type is more likely to show a pattern that lines up along a bearing or set of bearings since errors may concentrate along bearing lines. The arcs as well as partial spokes in mid to far ranges are signs of the test removing ionospheric interference. Close ranges at all bearings were more frequently flagged at NANT and MVCO. Maps for stations near the Gulf Stream showed that the test was often flagging the Gulf Stream gradient. In order to minimize this erroneous flagging, the current difference threshold at those sites (DUCK, HATY, CORE and LISL) was increased from 30 cm/s to 50 cm/s. The problem was not completely eliminated but flag counts decreased as a result.

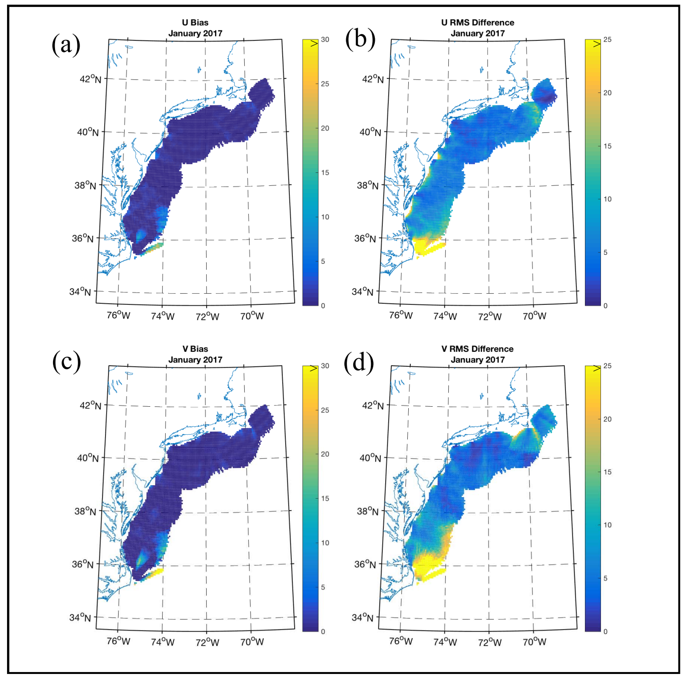

Fig. 6. Biases between the real-time and reprocessed surface currents for the U and V components of velocity are shown in (a) and (c) respectively. Unbiased root mean square differences between the real-time and reprocessed surface currents for the U and V components are shown in (b) and (d) respectively.

Fig. 5 shows radial count file failures given as a percent of files that failed out of a total number of available radial files for 2017. 1794 radial files were excluded based on the radial count test.

C. Real-time versus Reprocessed Total Vector Products

A comparison of surface current maps for the month of January shows that the bias, or difference between means, for U and V components is below 10 cm/s for most of the coverage area (Fig. 6). The unbiased root mean square differences are typically less than 15 cm/s. In the south, in locations where radial data was added, where radial metric QC was applied, and where currents are stronger due to the presence of the Gulf Stream, there are greater biases as well as greater unbiased root mean square differences. Fig. 7 shows the monthly average vectors at grid locations with 60% data availability after removing data with GDOP > 1.25 for a southern section of the Mid-Atlantic. Data coverage is greater in the reprocessed product and reveals the Gulf Stream currents. Also, a number of suspect vectors near the coast with strong average currents directed offshore have been eliminated in the reprocessed product.

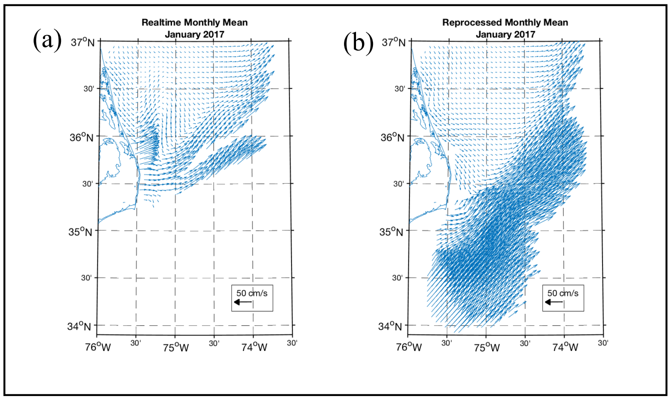

Fig. 7. January 2017 average current vectors for the (a) real-time and (b) reprocessed surface current products.

IV. Discussion

Observing system networks should weigh the potential benefits that may result from reprocessing HF radar data with the time commitment. The initial radial review relied heavily on the expertise of the QC team to make judgment calls and reprocess radials from spectra when deemed appropriate. This is a time-consuming step, but the impact on the radials and totals can be quite significant. The application of radial QC flags is comparatively more objective and efficient. Once QC test thresholds are assigned, this process is completely automated. However, QC flags may not catch some errors that would be easily noticed by an operator looking at a radial map.

Many types of problems can be addressed by reprocessing from spectra. Cable swaps are a good example of a mistake that is easy to correct in reprocessing and one that is extremely important to correct because it can affect a large area in a significant way. Cable swaps can even generate maps with radials going the opposite direction of the true ocean current.

First order determination and direction finding are crucial steps in the creating the radial map and yet they are also parts of the process that can be dramatically altered for reprocessing. For example, if new algorithms are developed that improve removal of ionospheric interference from spectra, software running those algorithms could be utilized for reprocessing. This has the potential to make substantial improvements to data collected years ago and provide better quality data sets for researchers.

Differences between real-time and reprocessed data will be greatest in hourly plots and daily average plots. Weekly or monthly averages of total vector data and averages over large spatial areas will not differ as much. A researcher who would like to use surface current data to inform a study involving shorter time scales and/or a smaller geographic area may see significant benefit to using a reprocessed data product.

V. Conclusions

This paper has presented initial findings on the impact of reprocessing HFR data and has shown that significant differences occur between real-time and reprocessed products. However, much more may be done to show the impact on data quality. Future analysis will compare the real-time and reprocessed products to other data sources.

Acknowledgment

We thank the University of North Carolina and the Coastal Studies Institute for the use of HF radar data from the North Carolina radar stations. Sara Haines applied radial metric QC to radial data from UNC stations and kindly provided us with those reprocessed radial files for this effort. Thanks to the radar operators who helped maintain the long-range radar systems in 2017: Ethan Handel, Rich Arena, Mike Muglia, Patterson Taylor, Nicolas DeSimone and Anthony Whipple. Josh Kohut provided guidance on how to approach this reprocessing effort and helped determine our initial steps in the effort. Josh Kohut and Travis Miles also helped review and annotate hourly maps of reprocessed totals.

References

[1] H. Roarty, M. Smith, J. Kerfoot, J. Kohut, and S. Glenn, "Automated quality control of High Frequency radar data," in *Oceans, 2012*, 2012, pp. 1-7.

[2] H. Roarty, L. Palamara, J. Kohut, and S. Glenn, "Automated quality control of high frequency radar data II," in *OCEANS 2016 MTS/IEEE Monterey*, 2016, pp. 1-3.

[3] J. Kohut and S. Glenn. "Improving HF radar surface current measurements with measured antenna beam patterns." Journal of Atmospheric and Oceanic Technology., vol. 20(9), pp. 1303-1316. 2003.

[4] Manual for Real-Time Quality Control of High Frequency Radar Surface Current Data: a Guide to Quality Control and Quality Assurance for High Frequency Radar Surface Current Observations. Version 1.0. Integrated Ocean Observing System (U.S.), U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Integrated Ocean Observing System, Silver Spring, MD, 2016.

[5] Manual for the Use of Real-Time Oceanographic Data Quality Control Flags. Version 1.1. Integrated Ocean Observing System (U.S.), U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Integrated Ocean Observing System, Silver Spring, MD, 2017.

[6] Haines, S., H. Seim and M. Muglia. "Implementing Quality Control of High-Frequency Radar Estimates and Application to Gulf Stream Surface Currents," Journal of Atmospheric and Oceanic Technology, vol. 34(6), pp. 1207-1224. 2017,

[7] Kim, S. Y., E. J. Terrill, and B. D. Cornuelle, "Mapping surface currents from HF radar radial velocity measurements using optimal interpolation," J. Geophys. Res., vol. 113(C10), 2008.

[8] Kohut, J., Roarty H.J, Randall-Goodwin, E., Glenn, S. and C. S. Lichtenwalner, "Evaluation of two algorithms for a network of coastal HF radars in the Mid-Atlantic Bight," Ocean Dynamics, vol. 62, pp. 953–968. 2012.

[9] Kundu, P. "An analysis of inertial oscillations observed near Oregon coast," Journal of Physical Oceanography, vol. 6(6), pp 879-893, 1976.