**Report on Historical data files QA/QC**

<table>
<thead>
<tr>
<th>Version of the report</th>
<th>Changes made by</th>
<th>Nature of changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>V2</td>
<td>M. Chifflet &amp; A. Rubio</td>
<td>Spatial coverage details</td>
</tr>
<tr>
<td>V3</td>
<td>A. Rubio &amp; M. Chifflet</td>
<td>+ figures</td>
</tr>
<tr>
<td>V4</td>
<td>A. Rubio &amp; M. Chifflet</td>
<td>New version after reflagging</td>
</tr>
<tr>
<td>VR2020_12</td>
<td>L. Solabarrieta &amp; A. Rubio</td>
<td>Update 2020</td>
</tr>
</tbody>
</table>

System: EUSKOOS  
Data set: [Totals/Radials] Total  
Period: 12-Jan-2009 – 12-Jul-2020

**INFO ON QA/QC Settings and Calibration**

- %%% QC info for all the period 12-Jan-2009 – 31-oct-2019
- OceanSITES quality flagging for GDOP threshold QC test. Threshold set to 2.
- OceanSITES quality flagging for Data density threshold QC test. Threshold set to 3 radials.
- OceanSITES quality flagging for Velocity threshold QC test. Threshold set to 1.2 m/s.
- OceanSITES quality flagging for variance threshold QC test. Test not applicable to Direction Finding systems. The Temporal Derivative test is applied. Threshold set to 1.2 m/s.

- %%% Calibration info for all the period 12-Jan-2009 – 12-Jul-2019
- MATX : 2018-12-04T00:00:00Z; HIGE: 2017-06-12T00:00:00Z

Calibration information is missing for some files. But annual/biannual calibration campaigns have been carried out and processed for this system.

**RESULTS OF HIST DATA INSPECTION**

**General comments:**

We have noticed 11 periods to be controlled and/or reflagged, occurring in: 2009, 2010, 2014, 2015, 2016, 2018, 2019

- 2009: unstable year
- 2010: no available data until mid-May
- 2013: no data from mid-October to end-November
- 2019: unstable year

<table>
<thead>
<tr>
<th>year</th>
<th>General comment</th>
<th>Periods to be reflagged</th>
<th>Reason for new flagging</th>
<th>Sugg. Flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Unstable year</td>
<td>24-Jan-2009 – 25-Jan-2009</td>
<td>high mean current values</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30-Aug-2009 – 07-Sep-2009</td>
<td>low data availability</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23-Dec-2009 – 31-Dec-2009</td>
<td>low data availability</td>
<td>3</td>
</tr>
<tr>
<td>2010</td>
<td>No data available until 12-May-2010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td>03-Feb-2014 – 11-Feb-2014</td>
<td>high mean and std current v</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29-Dec-2014 – 31-Dec-2014</td>
<td>low data availability</td>
<td>3</td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td>01-Jan-2015 – 16-Jan-2015</td>
<td>low data availability</td>
<td>3</td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td>08-Nov-2016 – 29-Nov-2016</td>
<td>low data availability</td>
<td>3</td>
</tr>
</tbody>
</table>
From 2009 to 2018, the spatial/temporal coverage of the system is quite stable. In 2019, the available data are scarce, and the system has collapsed for several periods. In 2020, data availability is continuous again but with a noisy period from May-Jul 2020.

The annual temporal mean for most of the analyzed years is a cyclonic pattern with higher current velocities over the shelf slope. There are some years, for example 2020, when the temporal mean for Jan-Jul does not show a clear cyclonic pattern.

### Spatial Coverage vs. Temporal coverage: objective of USCG 80-80% data availability

<table>
<thead>
<tr>
<th>Period</th>
<th>General comments</th>
<th>Nb. analysed hours</th>
<th>80%-80% obj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>90.64 % spatial availability 80% of time</td>
<td>5530</td>
<td>y</td>
</tr>
<tr>
<td>2010</td>
<td>82.61 % spatial availability 80% of time</td>
<td>5145</td>
<td>y</td>
</tr>
<tr>
<td>2011</td>
<td>98.33 % spatial availability 80% of time</td>
<td>8676</td>
<td>y</td>
</tr>
<tr>
<td>2012</td>
<td>100 % spatial availability 80% of time</td>
<td>8764</td>
<td>y</td>
</tr>
<tr>
<td>2013</td>
<td>100 % spatial availability 80% of time</td>
<td>7453</td>
<td>y</td>
</tr>
<tr>
<td>2014</td>
<td>100 % spatial availability 80% of time</td>
<td>7297</td>
<td>y</td>
</tr>
<tr>
<td>2015</td>
<td>98.03 % spatial availability 80% of time</td>
<td>8551</td>
<td>y</td>
</tr>
<tr>
<td>2016</td>
<td>100 % spatial availability 80% of time</td>
<td>7270</td>
<td>y</td>
</tr>
<tr>
<td>2017</td>
<td>100 % spatial availability 80% of time</td>
<td>8586</td>
<td>y</td>
</tr>
<tr>
<td>2018</td>
<td>100 % spatial availability 80% of time</td>
<td>8544</td>
<td>y</td>
</tr>
<tr>
<td>2019</td>
<td>0 % spatial availability 80% of time</td>
<td>5612</td>
<td>n</td>
</tr>
<tr>
<td>Jan-Jul 2020</td>
<td>48.97% spatial availability 80% of time</td>
<td>4254</td>
<td>n</td>
</tr>
</tbody>
</table>

### Annex I Applied QA/QC tests

<table>
<thead>
<tr>
<th>QC Flag Variable name</th>
<th>Short name</th>
<th>Short description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td>Syntax</td>
<td>Syntax check: this test will ensure the proper formatting and the existence of all the necessary fields within the total NetCDF file. This test is performed on the NetCDF files and it assesses the presence and correctness of all data and attribute fields and the correct syntax throughout the file. This test is performed by the European HFR Node before pushing data to the distribution platforms.</td>
</tr>
</tbody>
</table>
**Data Density Threshold:** this test labels total velocity vectors with a number of contributing radials bigger than the threshold with a “good data” flag and total velocity vectors with a number of contributing radials smaller than the threshold with a “bad data” flag.

**Velocity Threshold:** this test labels total velocity vectors whose module is bigger than a maximum velocity threshold with a “bad data” flag and total vectors whose module is smaller than the threshold with a “good data” flag.

**Variance Threshold:** this test labels total vectors whose temporal variance is bigger than a maximum threshold with a “bad data” flag and total vectors whose temporal variance is smaller than the threshold with a “good data” flag. This test is applicable only to Beam Forming (BF) systems. Data files from Direction Finding (DF) systems will apply instead the “Temporal Derivative” test reporting the explanation “Test not applicable to Direction Finding systems. The Temporal Derivative test is applied.” in the comment attribute.

**Temporal Derivative:** for each total bin, the current hour velocity vector is compared with the previous and next hour ones. If the differences are bigger than a threshold (specific for each grid cell and evaluated on the basis of the analysis of one-year-long time series), the present vector is flagged as “bad data”, otherwise it is labelled with a “good data” flag. Since this method implies a one-hour delay in the data provision, the current hour file should have the related QC flag set to 0 (no QC performed) until it is updated to the proper values when the next hour file is generated.

**GDOP Threshold:** this test labels total velocity vectors whose GDOP (Geometrical Dilution Of Precision) is bigger than a maximum threshold with a “bad data” flag and the vectors whose GDOP is smaller than the threshold with a “good data” flag.

### Annex II QC Flags

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No QC was performed</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>Good data</td>
<td>All real-time QC tests passed.</td>
</tr>
<tr>
<td>2</td>
<td>Probably good data</td>
<td>-*</td>
</tr>
<tr>
<td>3</td>
<td>Bad data that are potentially correctable</td>
<td>These data are not to be used without scientific correction.*</td>
</tr>
<tr>
<td>4</td>
<td>Bad data</td>
<td>Data have failed one or more of the tests.</td>
</tr>
<tr>
<td>5</td>
<td>Value changed</td>
<td>Data may be recovered after transmission error.</td>
</tr>
<tr>
<td>6</td>
<td>Not used</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>Nominal value</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>Interpolated value</td>
<td>Missing data may be interpolated from neighbouring data in space or time.</td>
</tr>
<tr>
<td>9</td>
<td>Missing value</td>
<td>-</td>
</tr>
</tbody>
</table>

*These two are to be used after examination of the hist data sets and exchanges with the data provider

### Annex III Figures for the QA/QC tests

Fig A – Temporal series of the spatial average of the current velocity module (top panel), its standard deviation (middle panel) and the grid points of the total coverage (bottom panel). Black dots are the values obtained considering all the data in the domain, in green those considering only data with QC flag =1 (good data).
Fig B - Temporal series of the QC flags for all the grid nodes with data.

Fig C - Maps of the mean velocity module and the mean value of QC flags for the target year (left column) and their standard deviations (right column) for the target year.

Fig D - Spatial (x-axis) vs. temporal (y-axis) coverage 80/80 annual metric. Allows to check if the system has reached the goal of providing surface currents over the 80% of the area during 80% of the time.

Fig E – Map of the % of availability of data in each grid point and contour showing the area of temporal availability >80%

Fig F - Mean surface current maps for the indicated systems and periods. The means are computed in the area of 80% temporal coverage for the target year.
Period: 2009

A

B
**HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage**  
*12/01/2009 to 31/12/2009*

- 90.6355% data points available 80% of the time  
- Number of hours analyzed = 5530  
- (USCG 80-80 metric for data availability)
HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
12/01/2009 to 31/12/2009

HFR-EUSKOOS: HFR Surface current average [m/s]
12/01/2009 to 31/12/2009
Period: 2010

A

B
**HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage**
25/03/2010 to 31/12/2010

82.6087% data points available 80% of the time
Number of hours analyzed = 5145
(USCG 80-80 metric for data availability)
C

mean velocity module (m s\(^{-1}\))

std velocity module (m s\(^{-1}\))

mean gdop QC flag

std gdop QC flag

mean overall QC flag

std overall QC flag

D

HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2011 to 31/12/2011

98.3278% data points available 80% of the time
Number of hours analyzed = 8676
(USCG 80-80 metric for data availability)
HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
01/01/2011 to 31/12/2011

HFR-EUSKOOS: HFR Surface current average [m/s]
01/01/2011 to 31/12/2011
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2012 to 31/12/2012

100% data points available 80% of the time
Number of hours analyzed = 8764
(USCG 80-80 metric for data availability)
Period: 2013

A

B
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2013 to 31/12/2013

100% data points available 80% of the time
Number of hours analyzed = 7453
(USCG 80-80 metric for data availability)
Period: 2014

A

B
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2014 to 31/12/2014

100% data points available 80% of the time
Number of hours analyzed = 7297
(USCG 80-80 metric for data availability)
HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
01/01/2014 to 31/12/2014

HFR-EUSKOOS: HFR Surface current average [m/s]
01/01/2014 to 31/12/2014
Period: 2015

A

Mean velocity modulus - 2015

3rd velocity modulus - 2015

Number of road data - 2015

B

Overall QC Flag

G10p QC Flag

D10n QC Flag

Caped QC Flag

Vart QC Flag
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2015 to 31/12/2015

98.0328% data points available 80% of the time
Number of hours analyzed = 8551
(USCG 80-80 metric for data availability)
HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
01/01/2015 to 31/12/2015

HFR-EUSKOOS: HFR Surface current average [m/s]
01/01/2015 to 31/12/2015
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2016 to 31/12/2016

100% data points available 80% of the time
Number of hours analyzed = 7270
(USCG 80-80 metric for data availability)
HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
01/01/2016 to 31/12/2016

HFR-EUSKOOS: HFR Surface current average [m/s]
01/01/2016 to 31/12/2016
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2017 to 31/12/2017

100% data points available 80% of the time
Number of hours analyzed = 8586
(USCG 80-80 metric for data availability)
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage  
01/01/2018 to 31/12/2018

100% data points available 80% of the time  
Number of hours analyzed = 8544  
(USCG 80-80 metric for data availability)
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2019 to 17/12/2019

0% data points available 80% of the time
Number of hours analyzed = 5612
(USCG 80-80 metric for data availability)
Period: Jan-Jul 2020
D. HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
09/01/2020 to 12/07/2020

48.9764% data points available 80% of the time
Number of hours analyzed = 4254
(USCG 80-80 metric for data availability )

E. HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
09/01/2020 to 12/07/2020