# Report on HFR - EUSKOOS Historical data files QA/QC

Data provider information:

- contributors name: Anna Rubio, Lohitzune Solabarrieta, Ivan Manso-Narvarte

- contributors contact: arubio@azti.es, lsolabarrieta@azti.es, imanso@azti.es

- acknowledgements: These data were collected and made freely available by the Copernicus project and the programs that contribute to it. Data collected by Euskalmet and processed by AZTI within EuskOOS and JERICO-NEXT projects.

System: **EUSKOOS**
Sites: **HIGE**
Data set: **Radials**
Data source: Radials from the provider
Period: 01-Jan-2009 – 31-Dec-2022
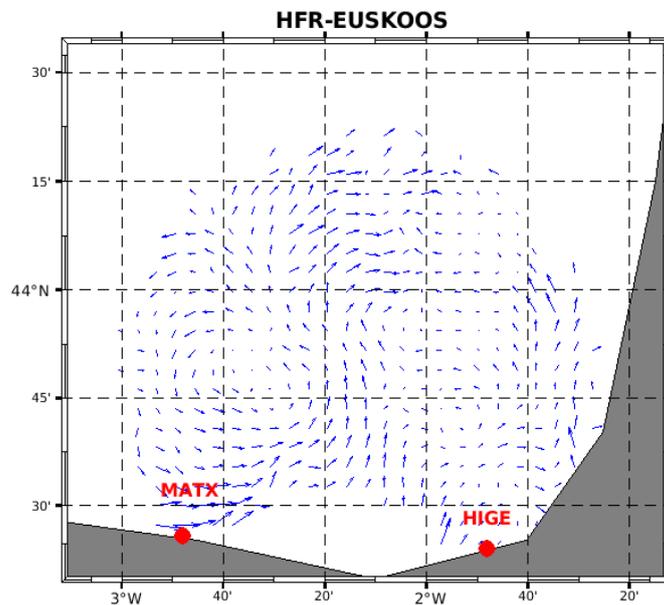
RESULTS OF HIST DATA INSPECTION

General comments:

The antenna has been functioning for more than 11 years. The general functioning has been almost continuous, expect for some weeks/months of non-functioning gaps, for example during 2010, 2015 & 2019.

The AVRB_QC thresholds is set to [210-320] degrees and it should be checked by the provider in order to guarantee the best functioning of this QC, as it is generating "bad" data for example during 2009, 2016 & 2017, when the data do not look bad. The provider has already been informed about this.

The antenna has been calibrated every two years. But the calibration patterns used for each time varies for some punctual and short periods, as for May 2014, when the data were calibrated with the pattern generated on 2013-11-28 but there are 2 days (May 16 to May 18) when the pattern used is for 2009-04-30 :

%%% Calibration info for time: 14-May-2014 12:00:00
2013-11-28T00:00:00Z

%%% Calibration info for time: 16-May-2014 12:00:00
2009-04-30T00:00:00Z

%%% Calibration info for time: 18-May-2014 12:00:00
2013-11-28T00:00:00Z

The provider has been informed about this.

year	General comment	Periods to be reflagged	Reason for new fagging	Sugg. Flag
After exchanges with the provided the following periods where reflagged:				
Year	General comment	Periods to be reflagged	Reason	New Flag

The data provider has been informed about the quality flag =4 for some periods. We suspect that it could be related to the threshold limits of the AVRB_QC. It´s modification could change the quality flag of these periods periods.

During the analyzed period, the spatio/temporal 80-80% objective for data availability has not been reached. There are two reasons that affects to this:
- The AVRB_QC flag: the modification of the limits could improve the spatio/temporal coverage
- The periods when the antenna has not been functioning decrease the yearly data availability.

But it is worth to mention the really high spatio/temporal coverage for 2011, 2012, 2013, 2021 & 2022 despite the previously mentioned affections.

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

Period	General comments	Nb. analysed hours	80%-80% obj.
2009	0 % spatial availability 80% of time	309*	n
2010	0.48 % spatial availability 80% of time	209*	n
2011	54.52 % spatial availability 80% of time	362*	n
2012	68.70 % spatial availability 80% of time	335*	n
2013	46.98 % spatial availability 80% of time	316*	n
2014	38.04 % spatial availability 80% of time	355*	n
2015	0 % spatial availability 80% of time	358*	n

2016	0 % spatial availability 80% of time	351*	n
2017	0 % spatial availability 80% of time	346*	n
2018	14.81 % spatial availability 80% of time	337*	n
2019	25.61 % spatial availability 80% of time	245*	n
2020	0 % spatial availability 80% of time	334*	n
2021	57.58 % spatial availability 80% of time	309*	n
2022	69.98 % spatial availability 80% of time	353*	n

*A subsample dataset (1 hour per 8 available hours) has been analysed for plotting purposes

Annex I Applied QA/QC tests

QC Flag Variable name	Short name	Short description
-	**Syntax**	**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.
AVRB_QC	**Average Radial Bearing**	**Average Radial Bearing**: this test labels the entire data file with a 'good_data" flag if the average radial bearing of all the vectors contained in the data file lies within a specified margin around the expected value of normal operation. Otherwise, the data file is labeled with a "bad_data" flag. The value of normal operation has to be defined within a time interval when the proper functioning of the device is assessed. The margin has to be set according site-specific properties. This test is applicable only to DF systems. Data files from BF systems will have this variable filled with "good_data" flags (1) and the explanation "Test not applicable to Beam Forming systems" in the comment attribute.
CSPD_QC	**Velocity Threshold**	Velocity **Threshold**: this test labels radial 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.
MDFL_QC	**Median Filter**	**Median Filter**: for each source vector, the median of all velocities within a radius of <RCLim> and whose vector bearing (angle of arrival at site) is also within an angular distance of <AngLim> degrees from the source vector's bearing is evaluated. If the difference between the vector's velocity and the median velocity is greater than a threshold, then the vector is labeled with a "bad_data" flag, otherwise it is labeled with a "good_data" flag.
OWTR_QC	**Over Water**	**Over water**: This test labels radial vectors that lie on land with a "bad data" flag and radial vectors that lie on water with a "good data" flag.

RDCT_QC	Radial Count	**Radial Count**: test labeling the entire data file having a number of radial velocity vectors bigger than the threshold with a "good data" flag and data file having a number of radial velocity vectors smaller than the threshold with a "bad data" flag.
TIME_QC	Temporal Derivative	**Temporal** Derivative: for each radial 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.
VART_QC	Variance Threshold	**Variance** Threshold: this test labels radial 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.
QCflag	Overall QC	

Annex II QC Flags

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

*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 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%

* The "Fig C" is not missing; these letters have been assigned on purpose, to be in accordance with total figure names (where there are also Fig C and Fig F, which are no relevant in the radial case)

Period: 2009

A

B

D
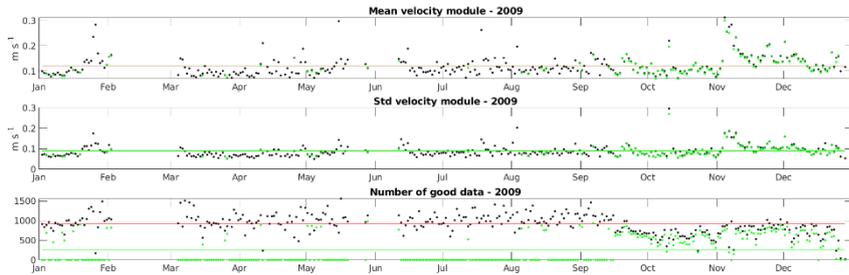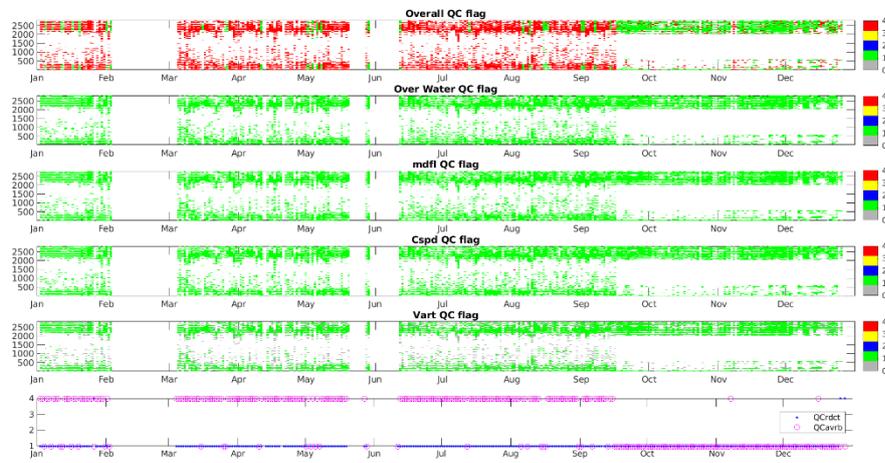
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2009 to 27/12/2009

0% data points available 80% of the time
Number of hours analyzed = 309
(USCG 80-80 metric for data availability)
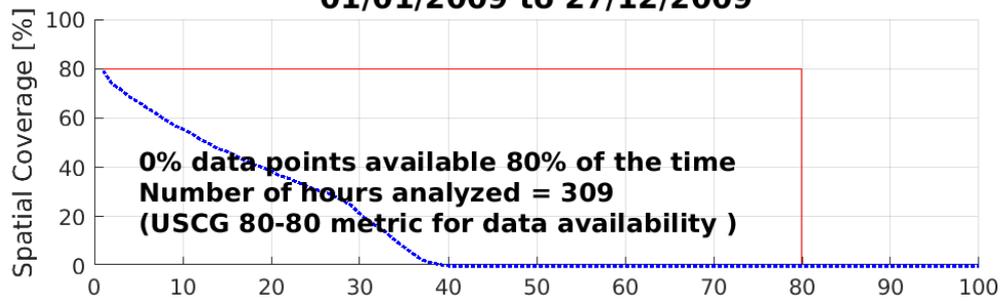
E HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
01/01/2009 to 27/12/2009

Period: 2010

A

Mean velocity module - 2010
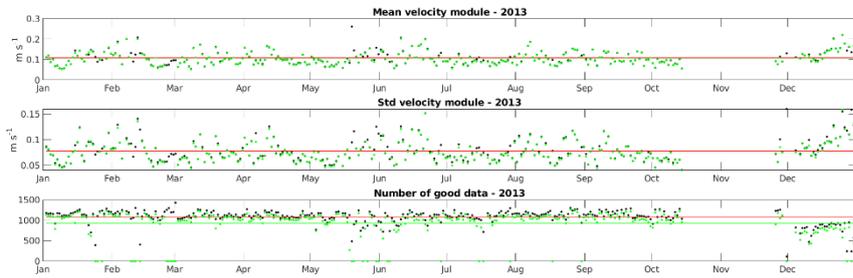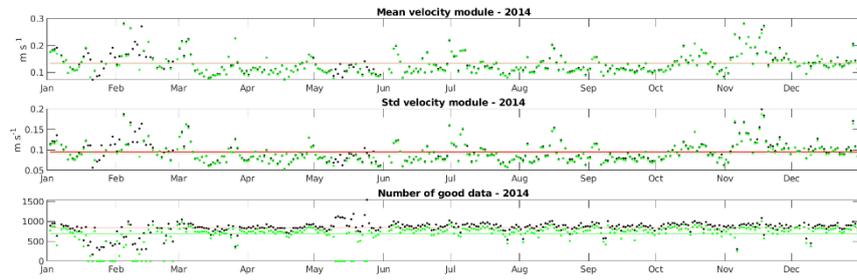
Std velocity module - 2010

Number of good data - 2010

B

Overall QC flag

Over Water QC flag

mdfl QC flag

Cspd QC flag

Vart QC flag

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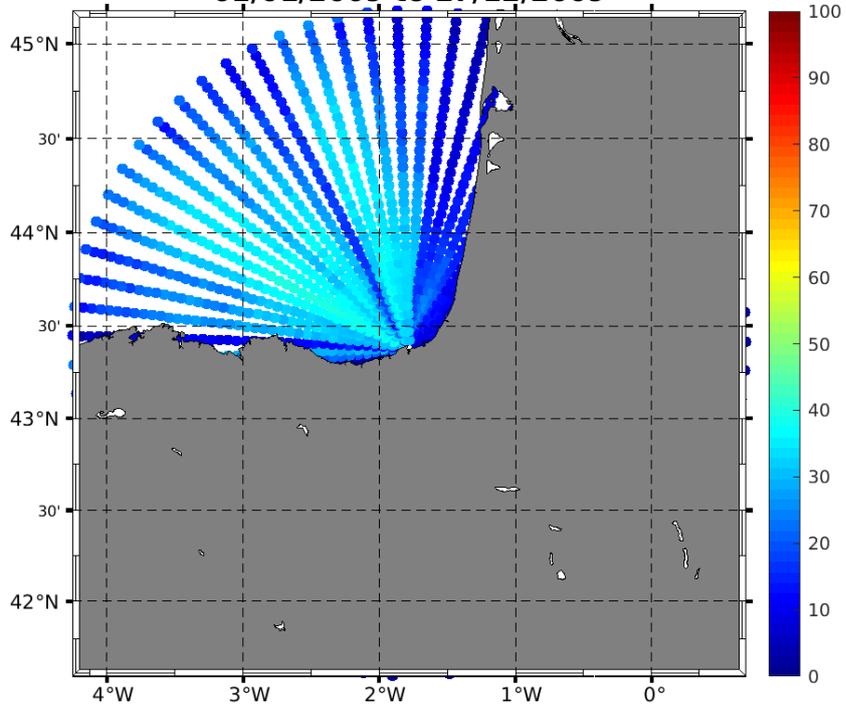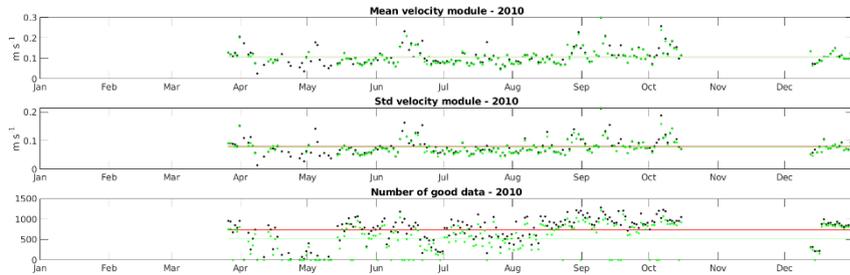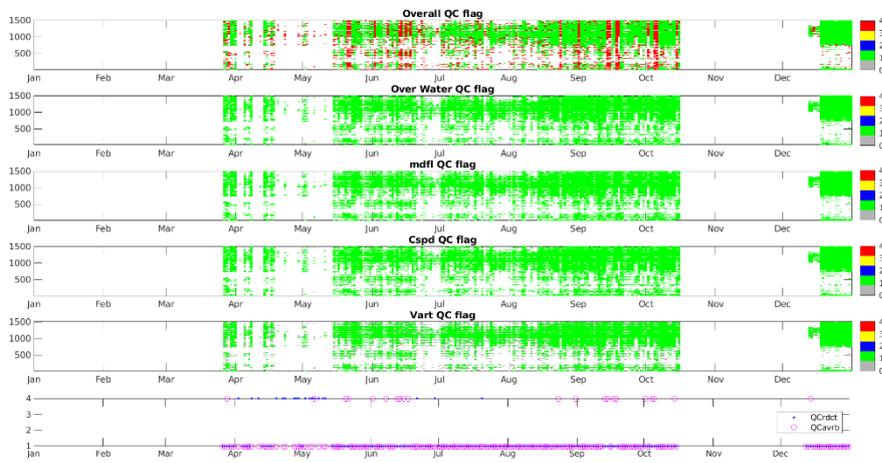
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
25/03/2010 to 31/12/2010

0.4851% data points available 80% of the time
Number of hours analyzed = 209
(USCG 80-80 metric for data availability)
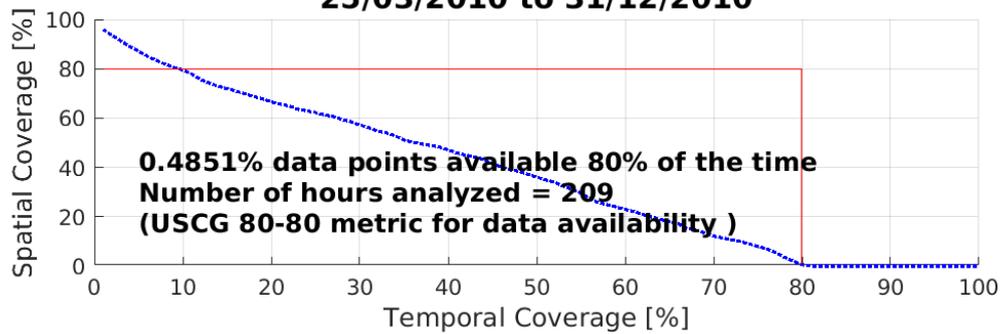
E | **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%) 25/03/2010 to 31/12/2010**
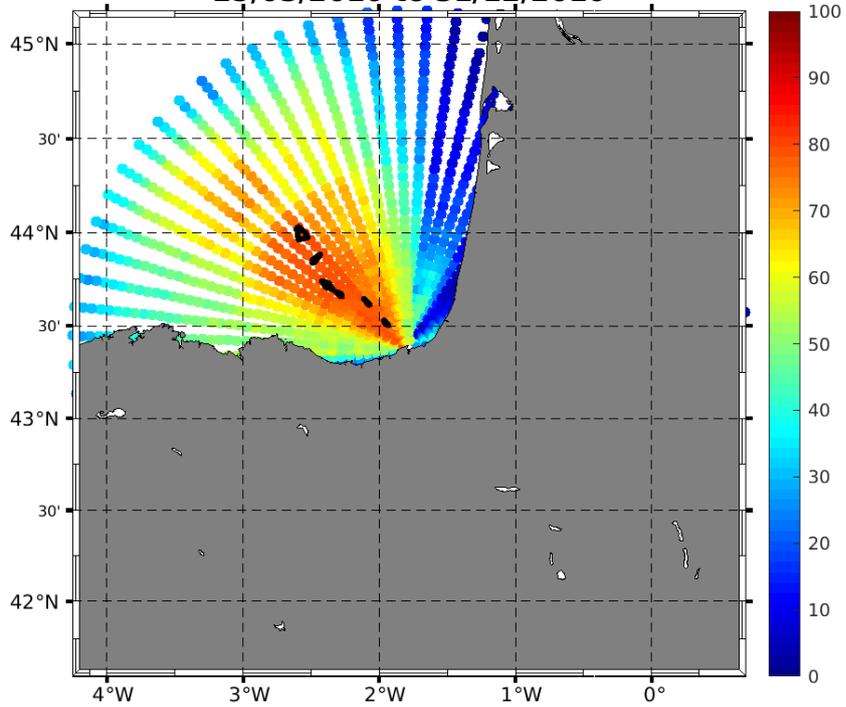
Period: 2011

A

B

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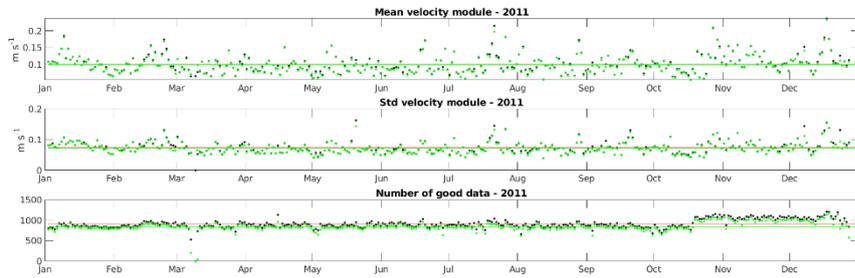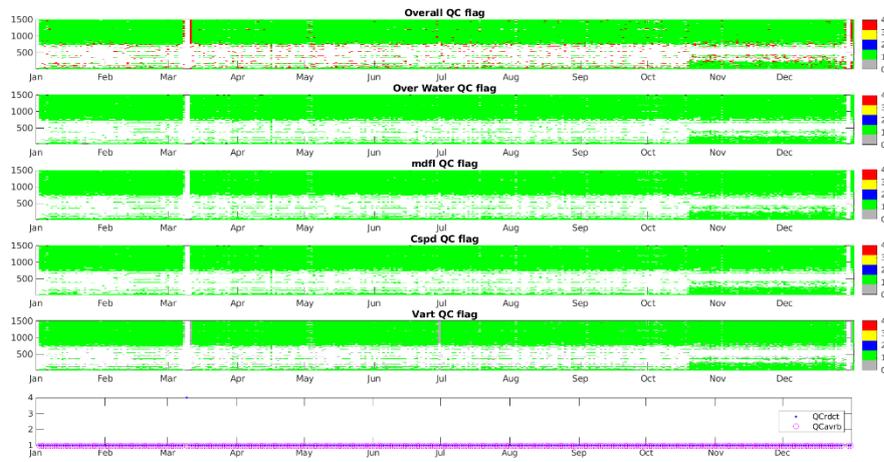
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2011 to 31/12/2011

54.5196% data points available 80% of the time
Number of hours analyzed = 362
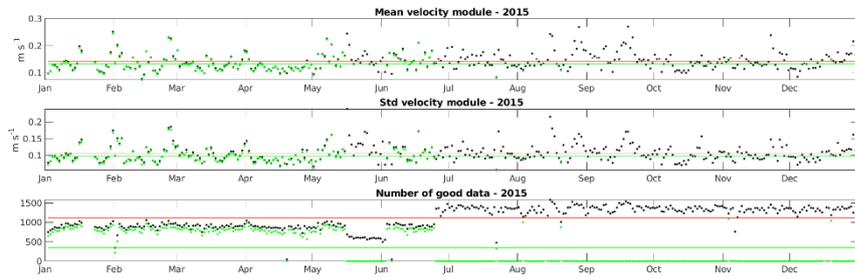**(USCG 80-80 metric for data availability)**
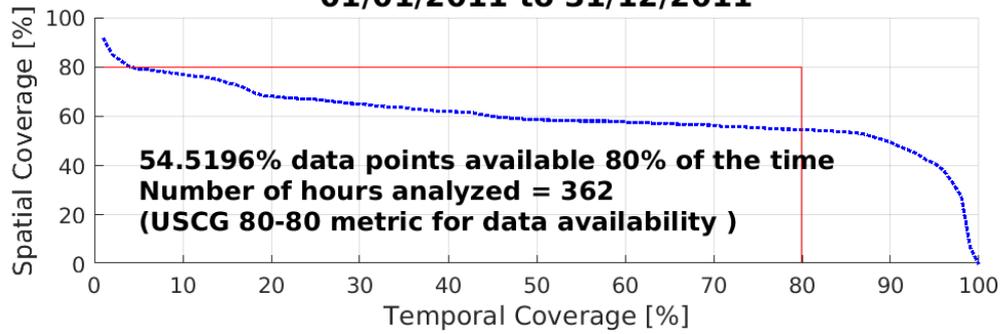
E **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)**
01/01/2011 to 31/12/2011

Period: 2012

A

Mean velocity module - 2012

Std velocity module - 2012

Number of good data - 2012

B

Overall QC flag

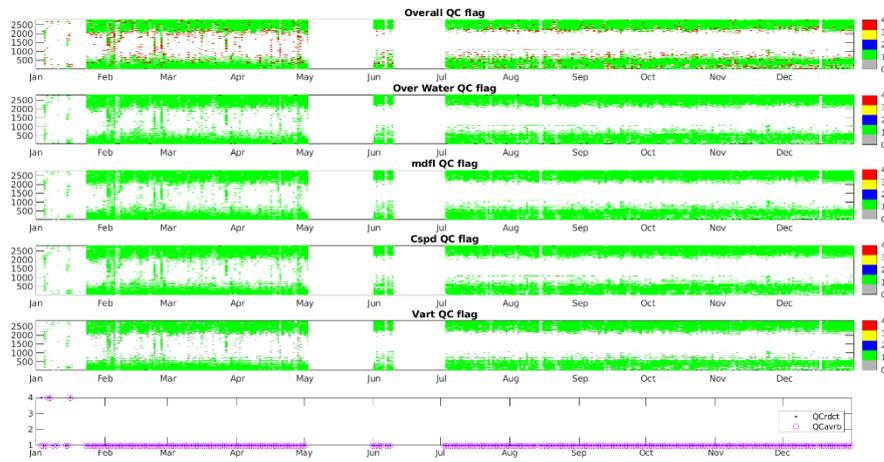Over Water QC flag

mdfl QC flag

Cspd QC flag

Vart QC flag

QCrdct
QCavrb

D
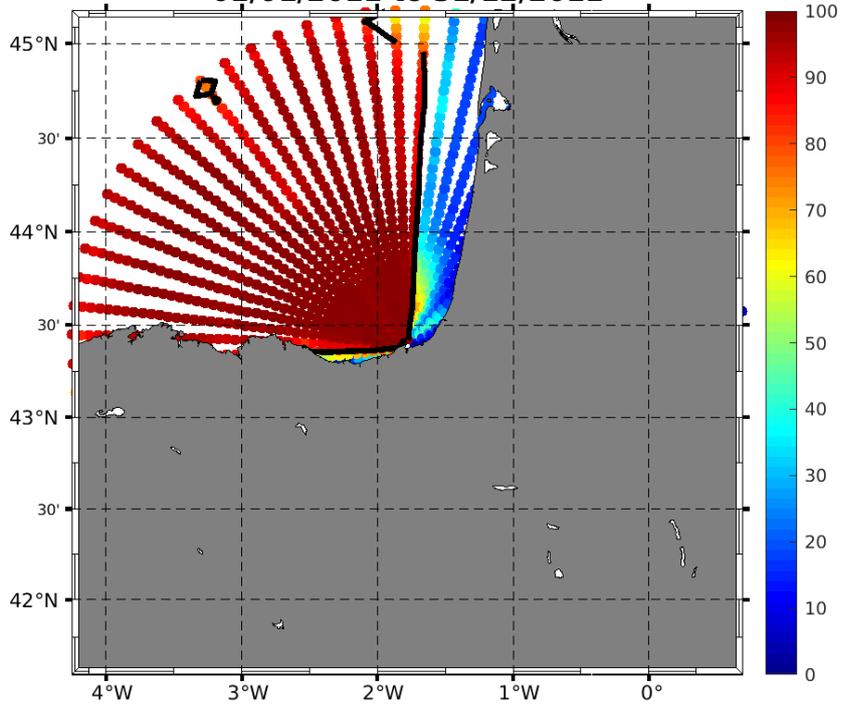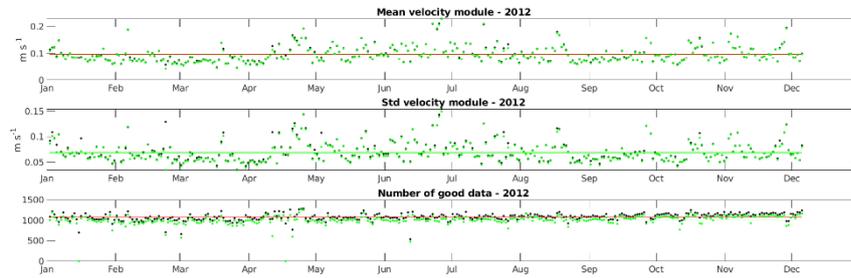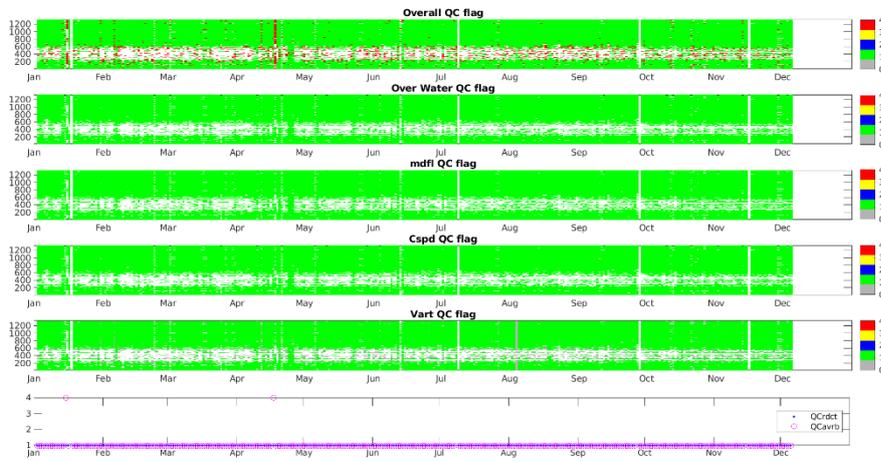
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage 01/01/2012 to 04/12/2012

68.7075% data points available 80% of the time
Number of hours analyzed = 335
(USCG 80-80 metric for data availability)
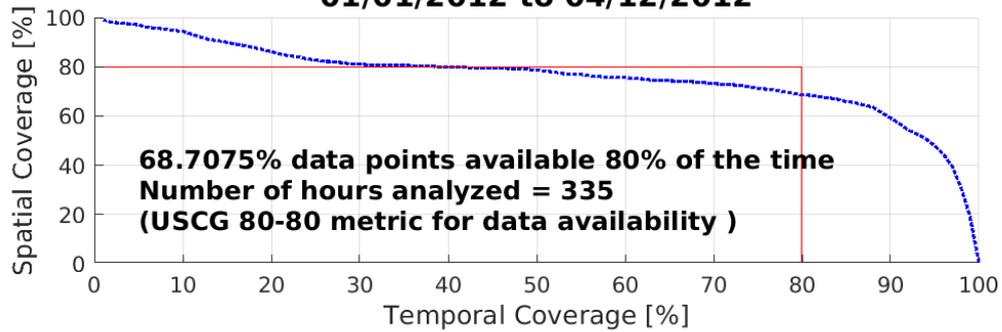
E | **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)**
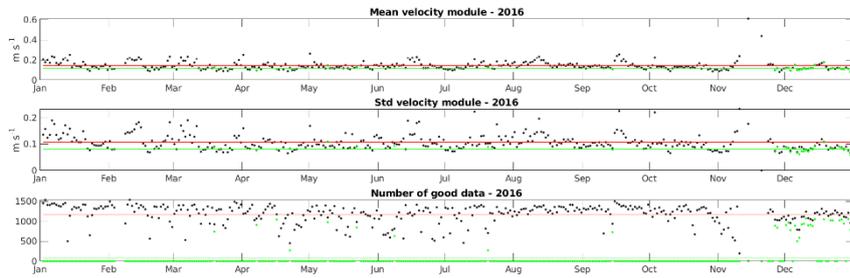**01/01/2012 to 04/12/2012**

Period: 2013

A

Mean velocity module - 2013

Std velocity module - 2013

Number of good data - 2013

B

Overall QC flag
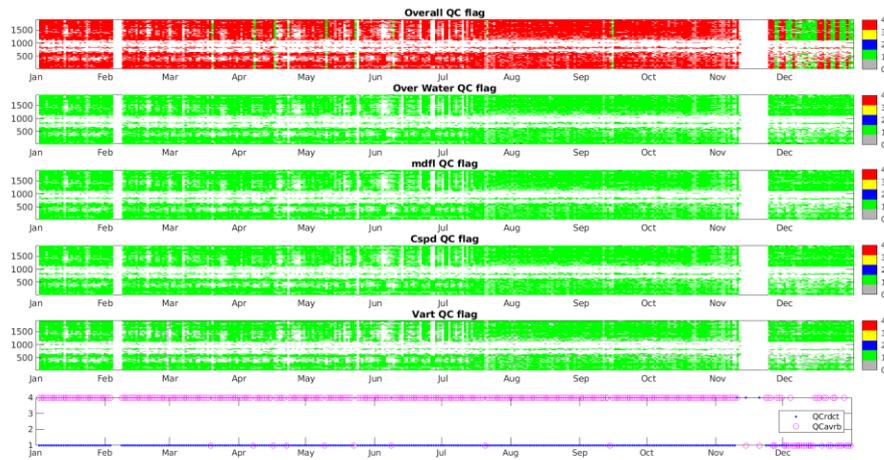
Over Water QC flag

mdfl QC flag

Cspd QC flag

Vart QC flag

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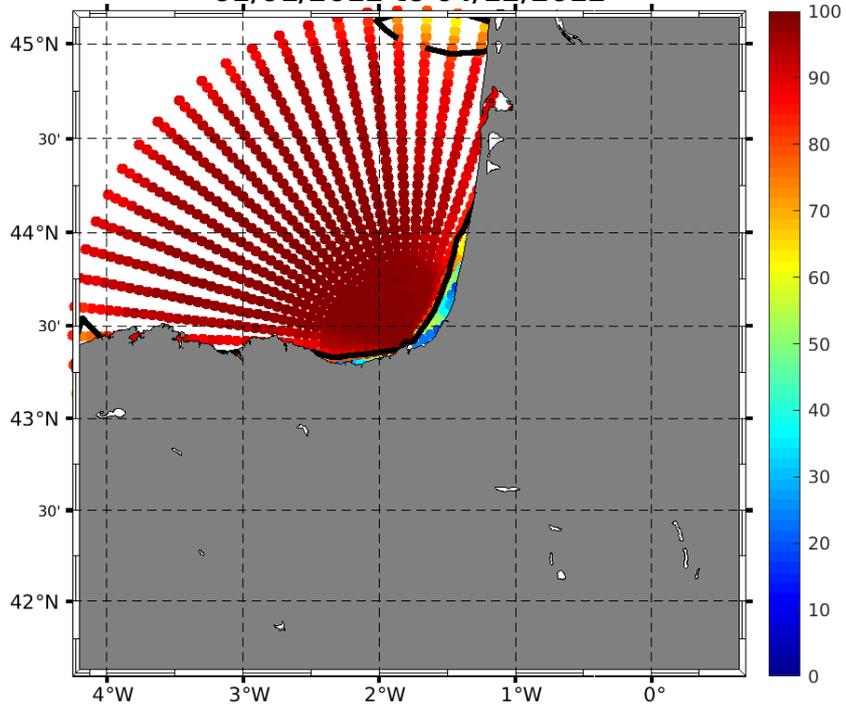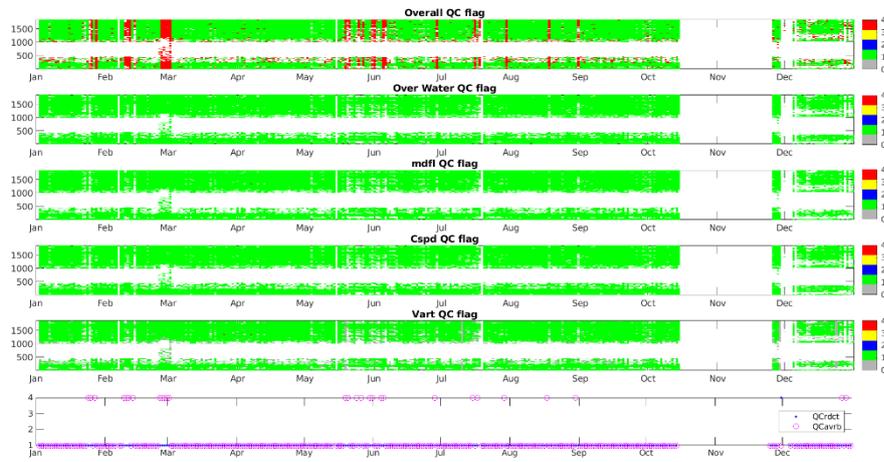
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2013 to 31/12/2013

46.9774% data points available 80% of the time
Number of hours analyzed = 316
(USCG 80-80 metric for data availability)
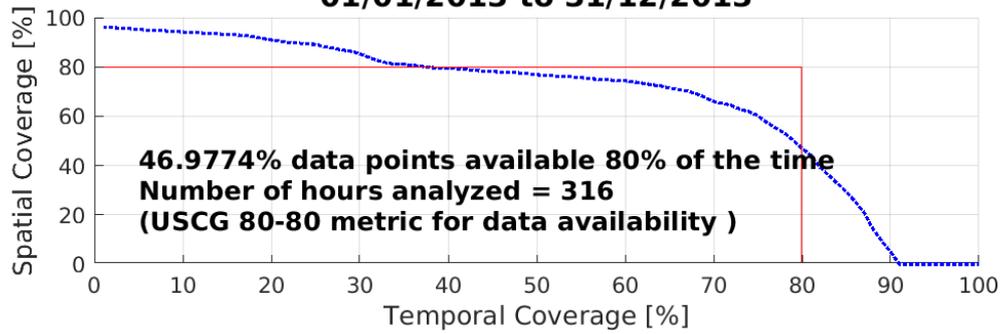
E | **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%) 01/01/2013 to 31/12/2013**

Period: 2014

A

Mean velocity module - 2014
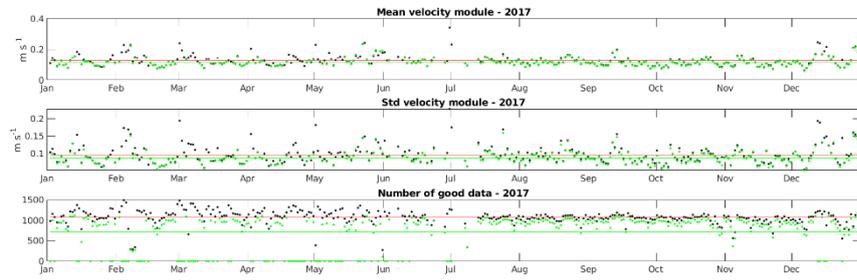
Std velocity module - 2014

Number of good data - 2014

B

Overall QC flag

Over Water QC flag

mdfl QC flag

Cspd QC flag

Vart QC flag

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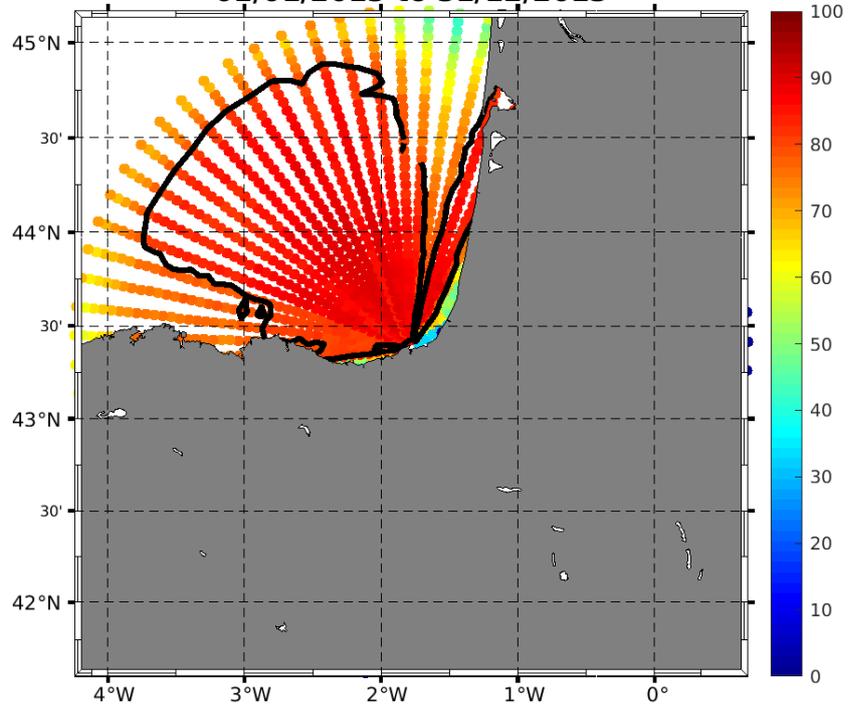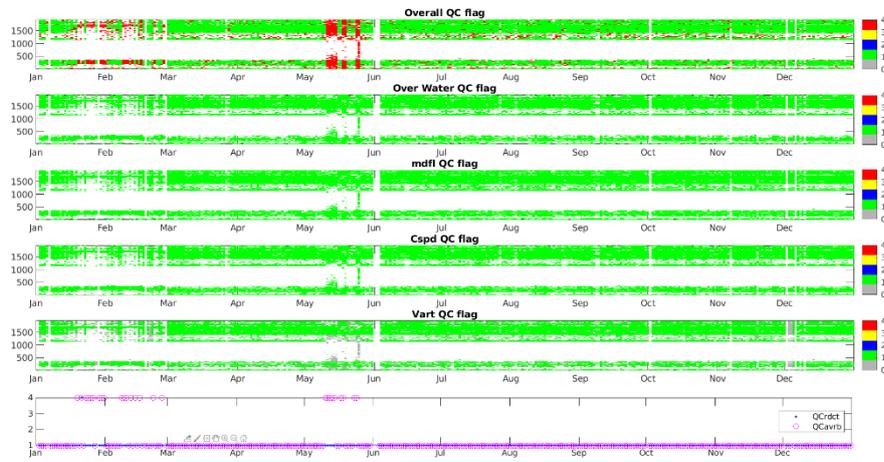
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2014 to 31/12/2014

38.0363% data points available 80% of the time
Number of hours analyzed = 355
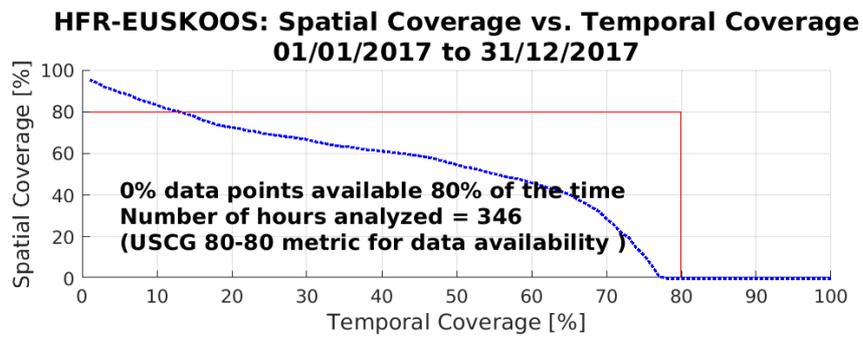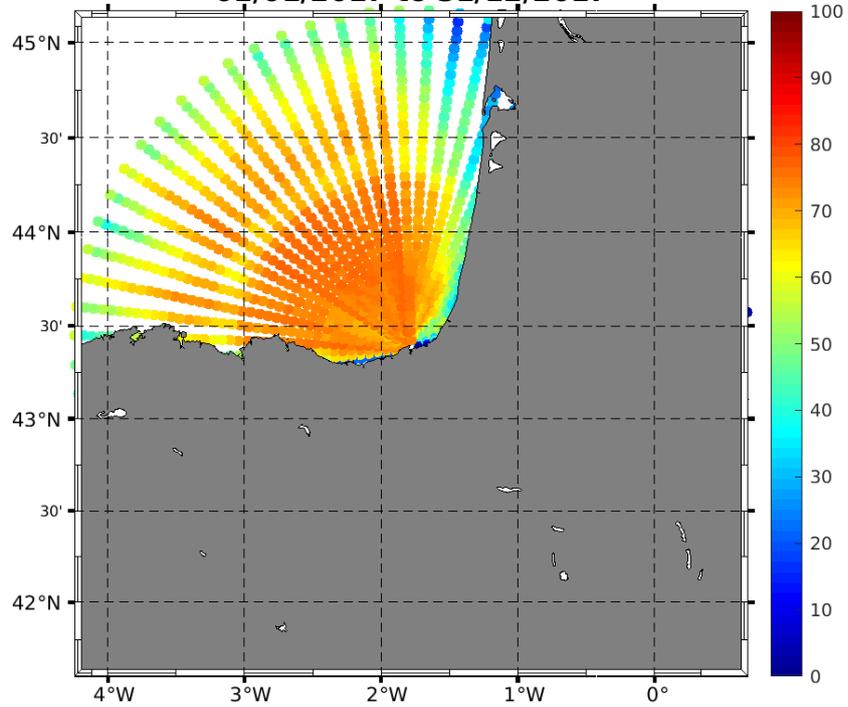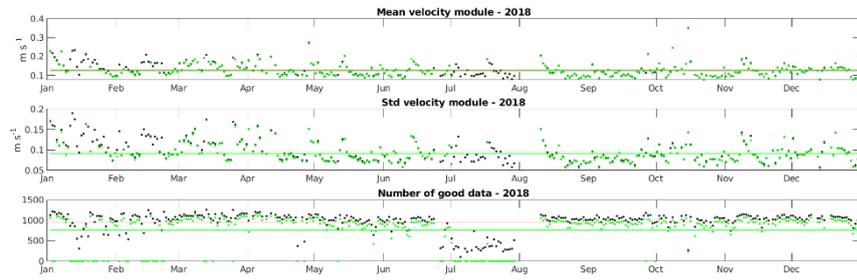**(USCG 80-80 metric for data availability)**
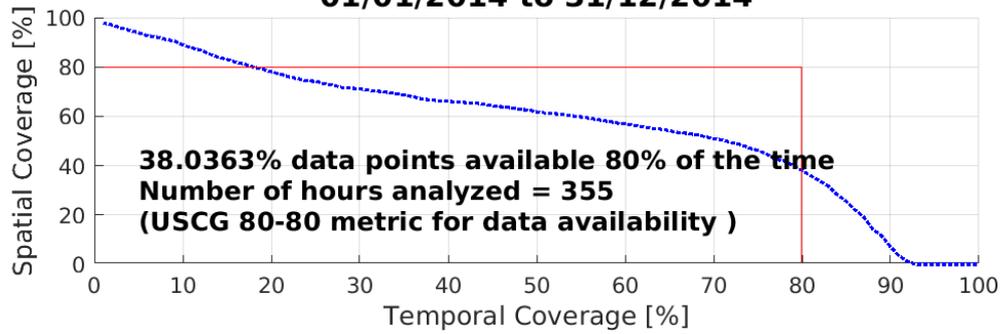
E   **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)**
01/01/2014 to 31/12/2014

Period: 2015

A

Mean velocity module - 2015

Std velocity module - 2015

Number of good data - 2015

B

Overall QC flag

Over Water QC flag

mdfl QC flag

Cspd QC flag

Vart QC flag

D
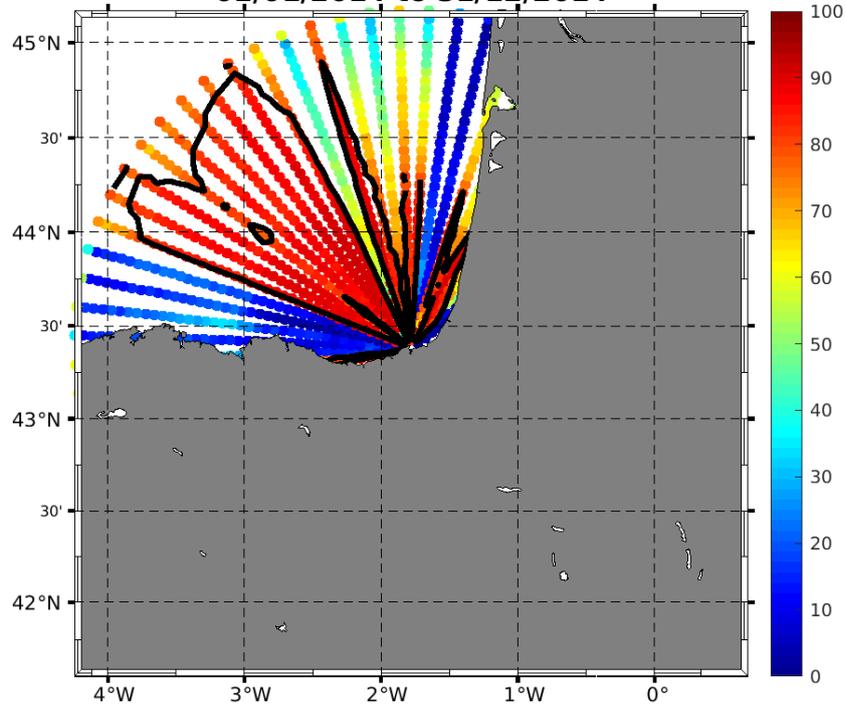
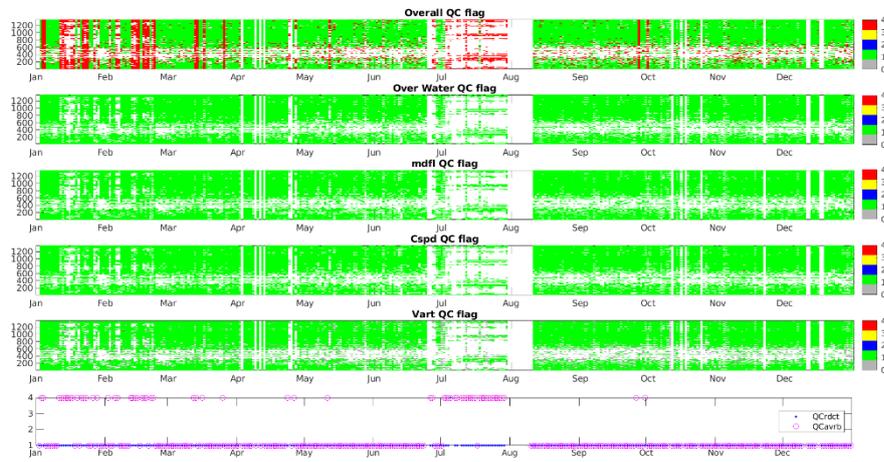**HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage 01/01/2015 to 31/12/2015**

0% data points available 80% of the time
Number of hours analyzed = 358
(USCG 80-80 metric for data availability)
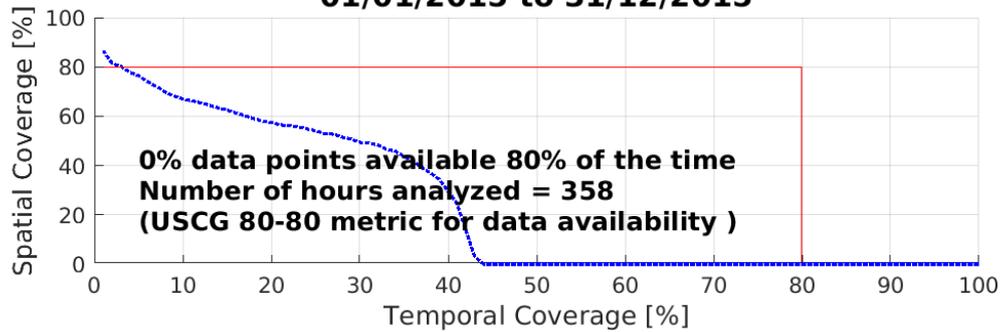
E | **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)**
01/01/2015 to 31/12/2015
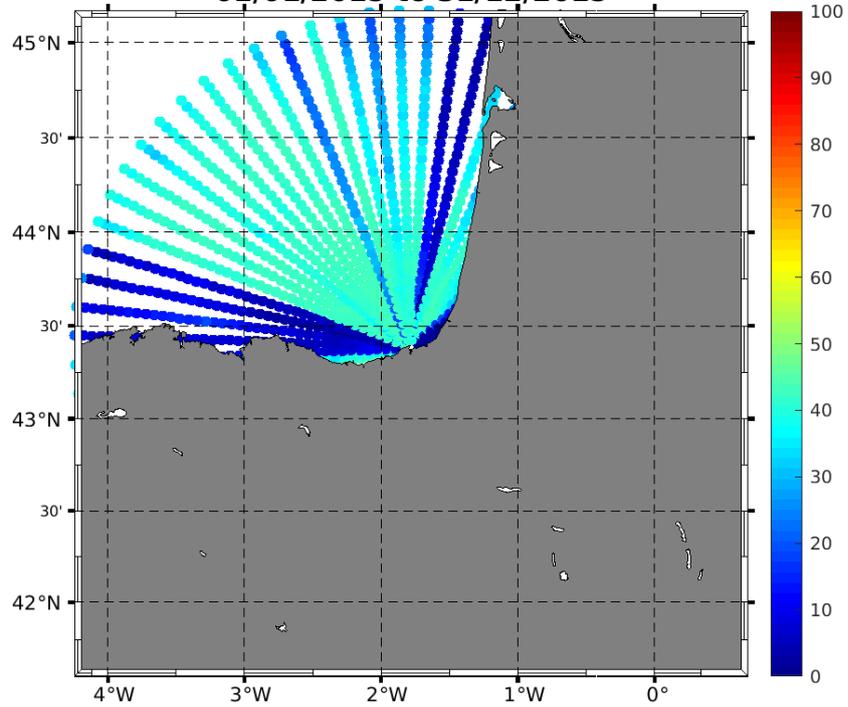
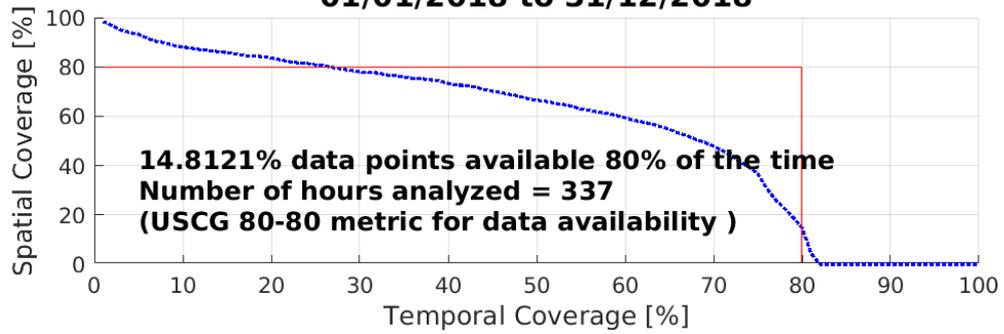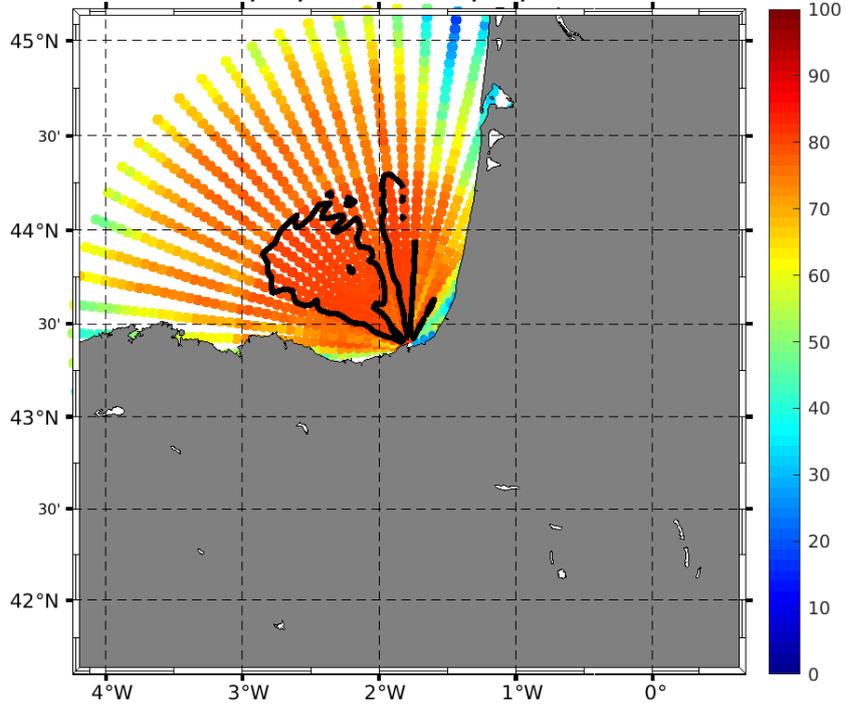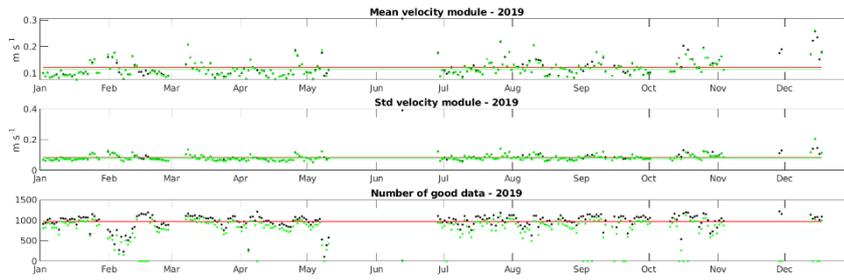| Period: 2016	
A	Mean velocity module - 2016 / Std velocity module - 2016 / Number of good data - 2016
B	Overall QC flag / Over Water QC flag / mdfl QC flag / Cspd QC flag / Vart QC flag
D	HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage 01/01/2016 to 31/12/2016
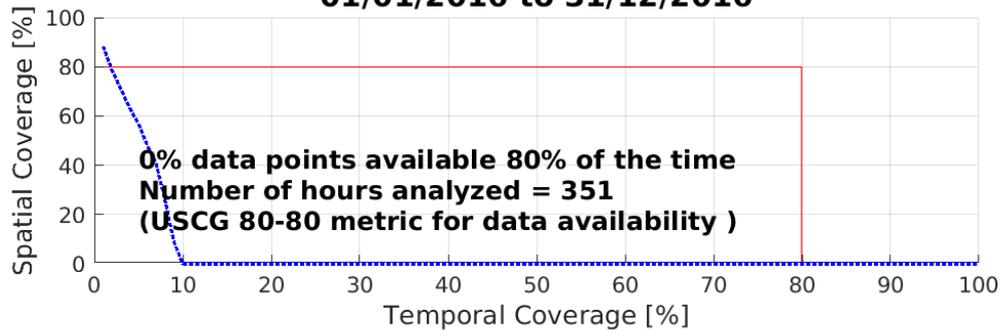
E   HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
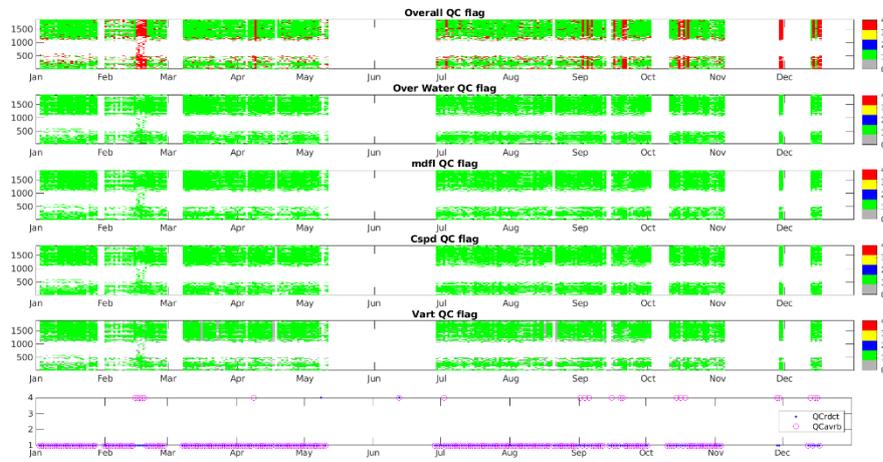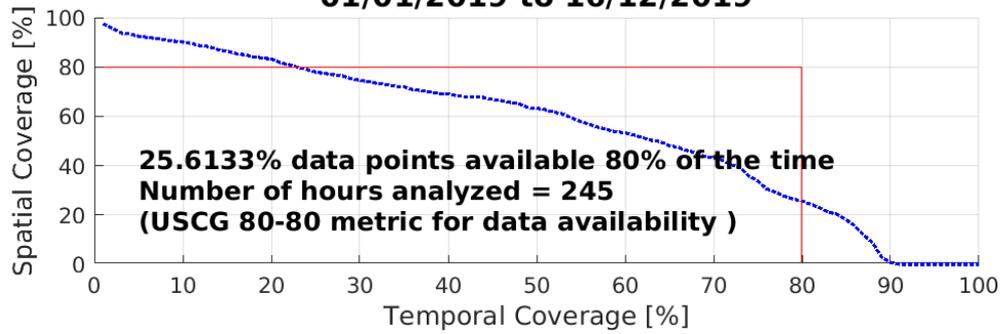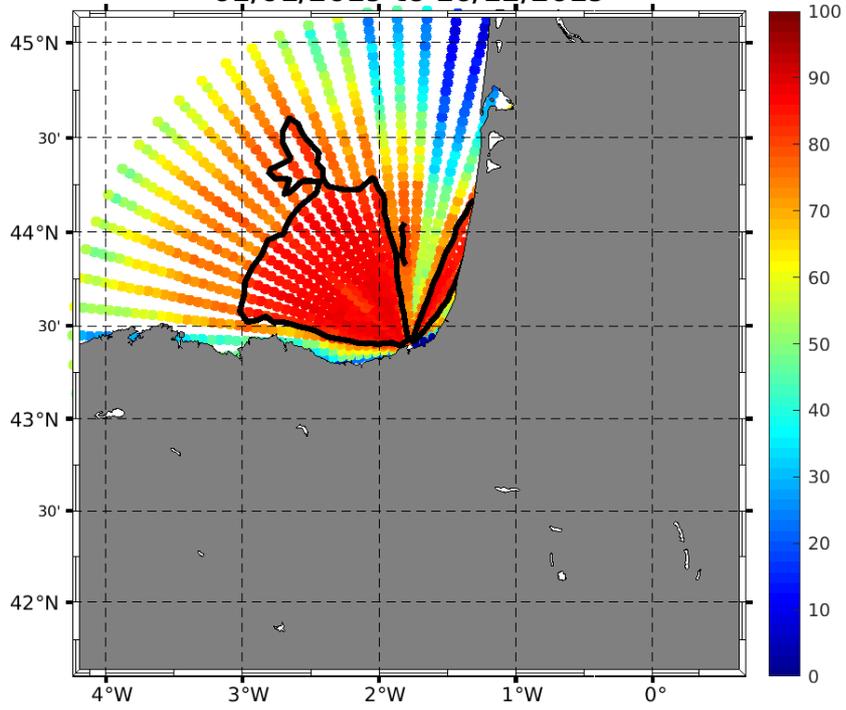**01/01/2016 to 31/12/2016**

Period: 2017

A

B

D

HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2017 to 31/12/2017

0% data points available 80% of the time
Number of hours analyzed = 346
(USCG 80-80 metric for data availability)
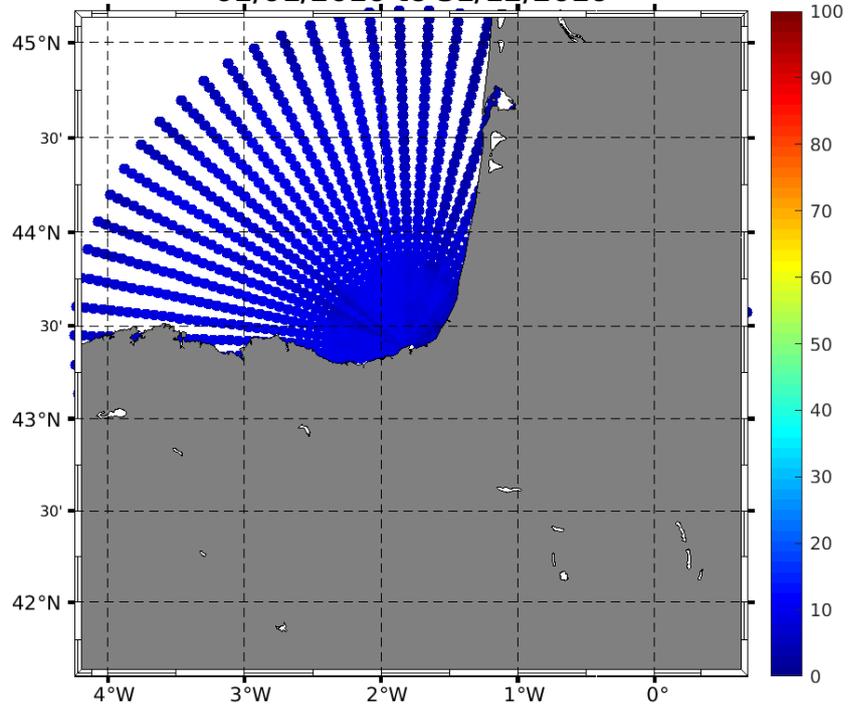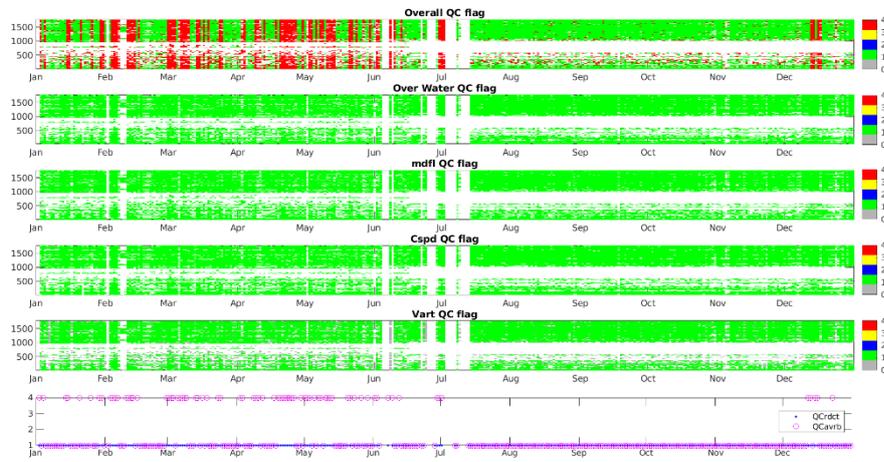
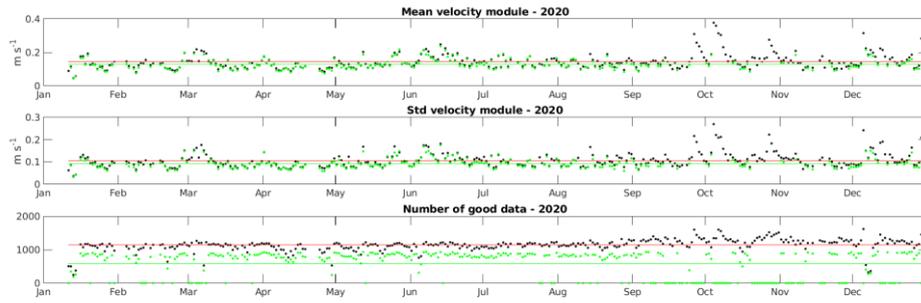**E** | **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)**
01/01/2017 to 31/12/2017

Period: 2018

A

B

D

HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2018 to 31/12/2018

14.8121% data points available 80% of the time
Number of hours analyzed = 337
(USCG 80-80 metric for data availability)

E   HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
01/01/2018 to 31/12/2018

Period: 2019

A

B

D

HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2019 to 16/12/2019

25.6133% data points available 80% of the time
Number of hours analyzed = 245
(USCG 80-80 metric for data availability)

E **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)**
01/01/2019 to 16/12/2019

Period: 2020

A

B
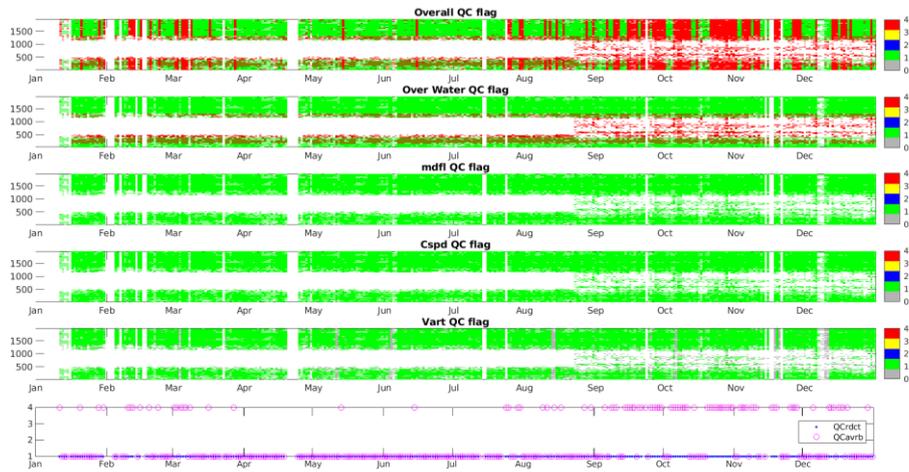
D | **HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage**
10/01/2020 to 31/12/2020

0% data points available 80% of the time
Number of hours analyzed = 334
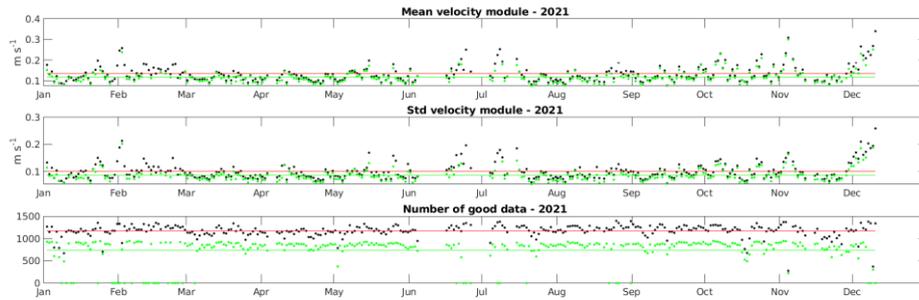(USCG 80-80 metric for data availability)
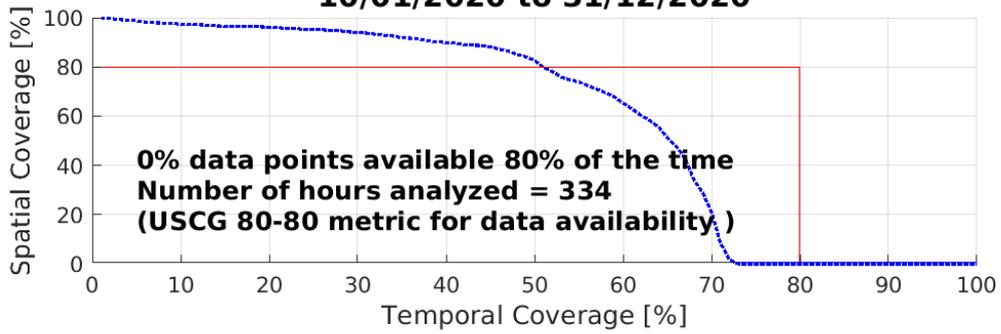
E | **HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)**
10/01/2020 to 31/12/2020

Period: 2021

A

Mean velocity module - 2021

Std velocity module - 2021

Number of good data - 2021

B

Overall QC flag
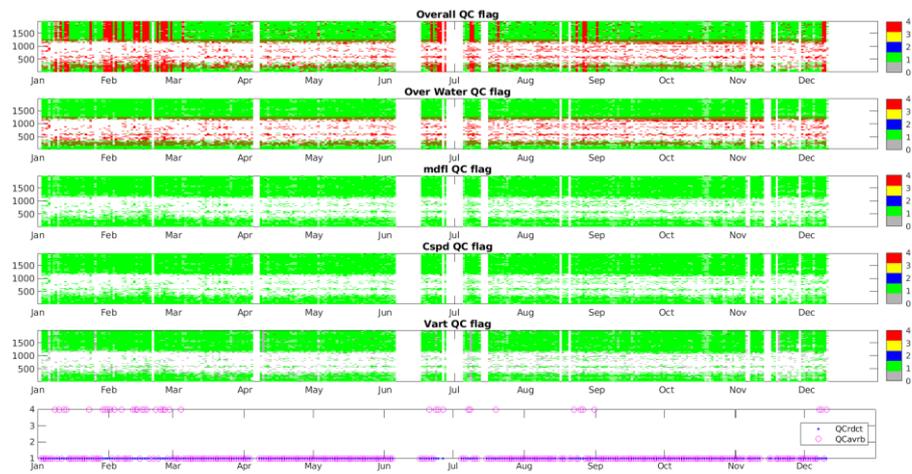
Over Water QC flag

mdfl QC flag

Cspd QC flag

Vart QC flag

D
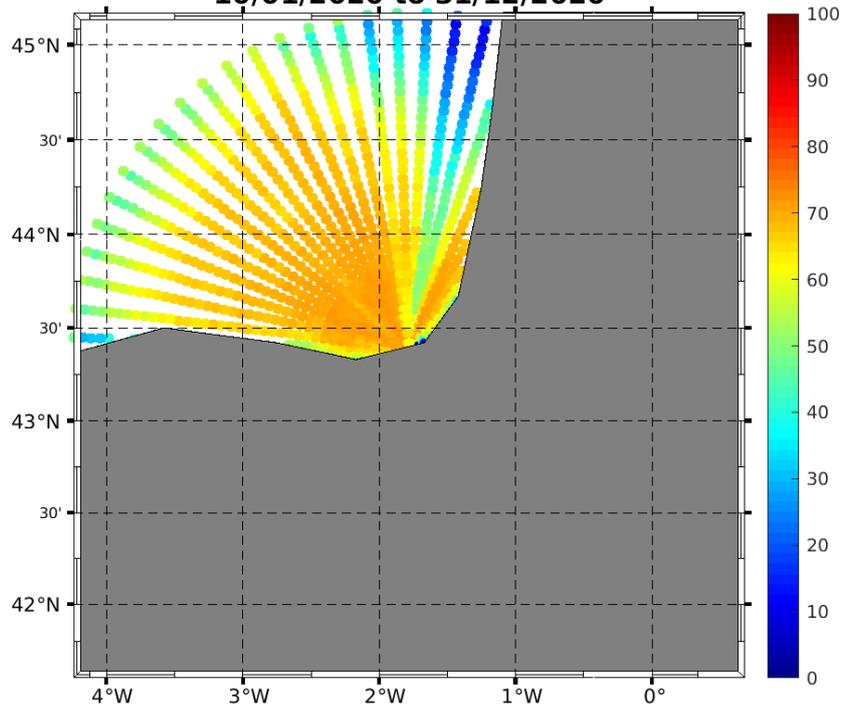
HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
01/01/2021 to 09/12/2021

57.582% data points available 80% of the time
Number of hours analyzed = 309
(USCG 80-80 metric for data availability)

E
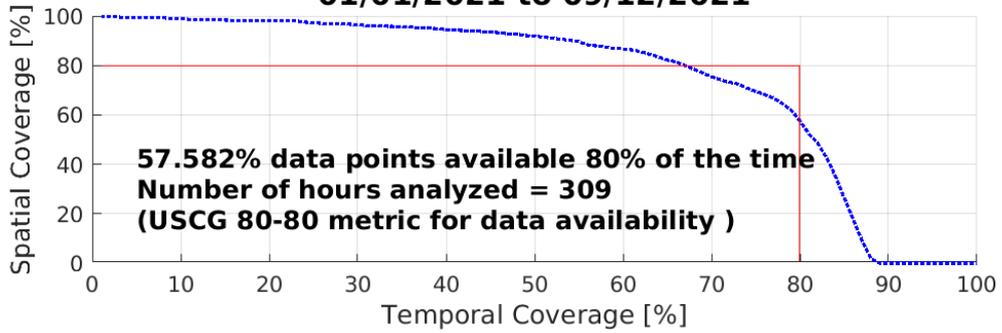
HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
01/01/2021 to 09/12/2021

A

Mean velocity module - 2022

Std velocity module - 2022

Number of good data - 2022

B

Overall QC flag

Over Water QC flag

mdfl QC flag

Cspd QC flag

Vart QC flag

D

HFR-EUSKOOS: Spatial Coverage vs. Temporal Coverage
04/01/2022 to 31/12/2022

69.9795% data points available 80% of the time
Number of hours analyzed = 353
(USCG 80-80 metric for data availability)

E

HFR-EUSKOOS: Percent Total Vector Coverage (contour showing >80%)
04/01/2022 to 31/12/2022