

Report on Historical data files QA/QC

Version of the report	Changes made by	Nature of changes
VR2021_04	L. Solabarrieta & A. Rubio	Report generation

System/Site: Lisboa/EPSM
Data set: [Totals/Radials] Radials
Period: 18-Sep-2012 – 21-Jul-2020

INFO ON QA/QC Settings and Calibration for info time: 18-Sep -2012 – 21-Jul-2020

%%% AVR_B_QC info:

OceanSITES quality flagging for Average Radial Bearing QC test. Thresholds set to [235-300] deg.

%%% CSPD_QC info:

OceanSITES quality flagging for Velocity Threshold QC test. Threshold set to 1.2 m/s.

%%% DEPH_QC info:

OceanSITES quality flagging for depth coordinate.

%%% MDFL_QC info:

OceanSITES quality flagging for Median Filter QC test. Threshold set to 5 km, 30 deg, 1 m/s,

%%% owtr_QC info:

OceanSITES quality flagging for Over-water QC test.

%%% POSITION_QC info:

OceanSITES quality flagging for position coordinates

%%% RDCT_QC info:

OceanSITES quality flagging for Radial Count QC test. Thresholds set to 200 vectors.

%%% TIME_QC info:

OceanSITES quality flagging for temporal coordinate.

%%% VART_QC info:

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 m/s.

%%% Calibration info for time: 18-Sep-2012 12:00:00

2012-09-10T00:00:00Z

%%% Calibration info for time: 13-Aug-2014 12:00:00

2014-08-07T00:00:00Z

RESULTS OF HISTDATA INSPECTION

General comments:

The antenna has been functioning during more than 8 years, with just few non-functioning periods (days to weeks). The longer gaps appeared in the plots are related to the Range/bearing change in the raw data, clarified in the next paragraph.

Changes in the Bearing Angle and Range Resolution has been observed in the raw dataset. For this reason, the available Temporal Aggregated (TA) file covers just the period from 27-Aug-2014 to 21-07-2020 where some periods have been extracted, to compile the netcdf file characteristics, despite the provider has provided radial data from 18-Sep-2012 – 21-Jul-2020.

The antenna has been calibrated 2 times.

The AVR_B_QC thresholds is set to [235-300] degrees and this is generating the flag=4 for the whole dataset. As the introduction of the system parameters is a responsibility of the data provider, they will be informed about it, to update the AVR_B_QC thresholds with proper values in the web form. Most of the data will be therefore reflagged with this adequation.

During the analyzed period the measured data number, the mean velocity and std values do not show any suspicious data but all the dataset will be checked again, once the AVR_B_QC thresholds are adequated.

Year	General comment	Periods to be reflagged	Reason for new fagging	Sugg. Flag

The AVR_B_QC threshold values are the main reason for the 0% of the spatio-temporal data availability, shown in the next table. Once the AVR_B_QC is corrected, the rest of the flags will be checked again

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

Period	General comments	Nb. analysed hours*	80%- 80% obj.
2012	0% spatial availability 80% of time	99	n
2013	0% spatial availability 80% of time	302	n
2014	0% spatial availability 80% of time	117	n
2015	0% spatial availability 80% of time	337	n
2016	0% spatial availability 80% of time	358	n
2017	0% spatial availability 80% of time	365	n
2018	0% spatial availability 80% of time	362	n
2019	0% spatial availability 80% of time	365	n
Jan-Jul 2020	0% spatial availability 80% of time	202	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

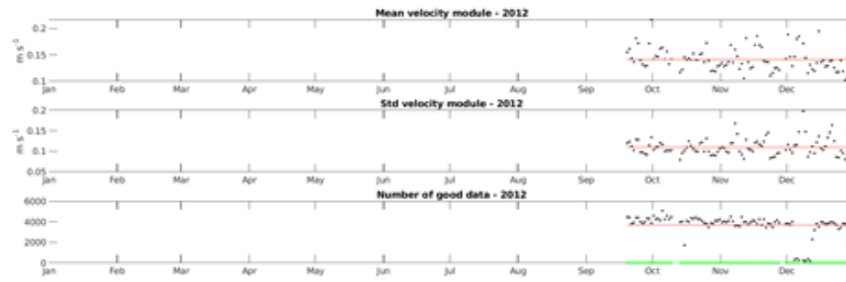
Fig D - Spatial (x-axis) vs. temporal (y-axis) coverage 80/80 annual metric. Allows to check if the system has reached the goal of providing surface currents over the 80% of the area during 80% of the time.

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

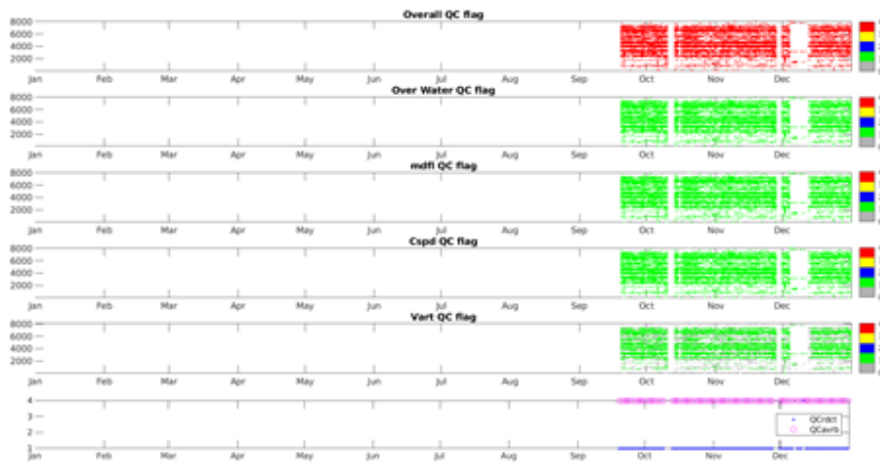
* The “Fig C” is not missing; these letters have been assigned on purpose, to be in accordance with total figure names (where there are also Fig C and Fig F, which are no relevant in the radial case)

Period: 2012

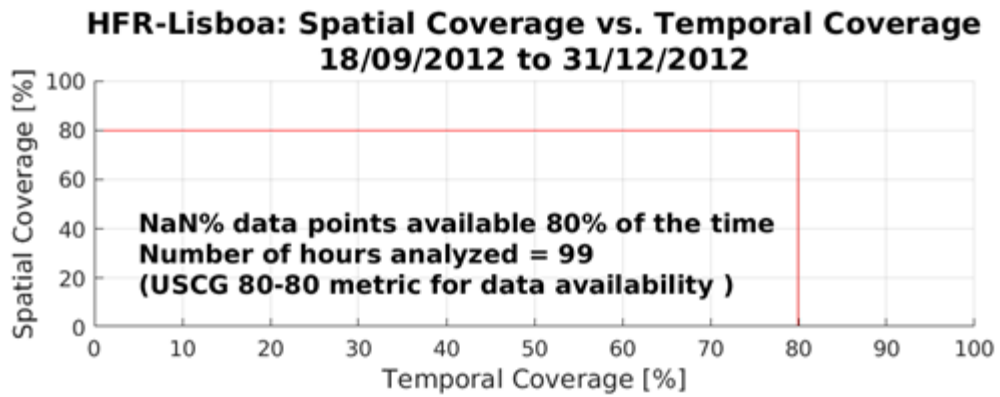
A



B

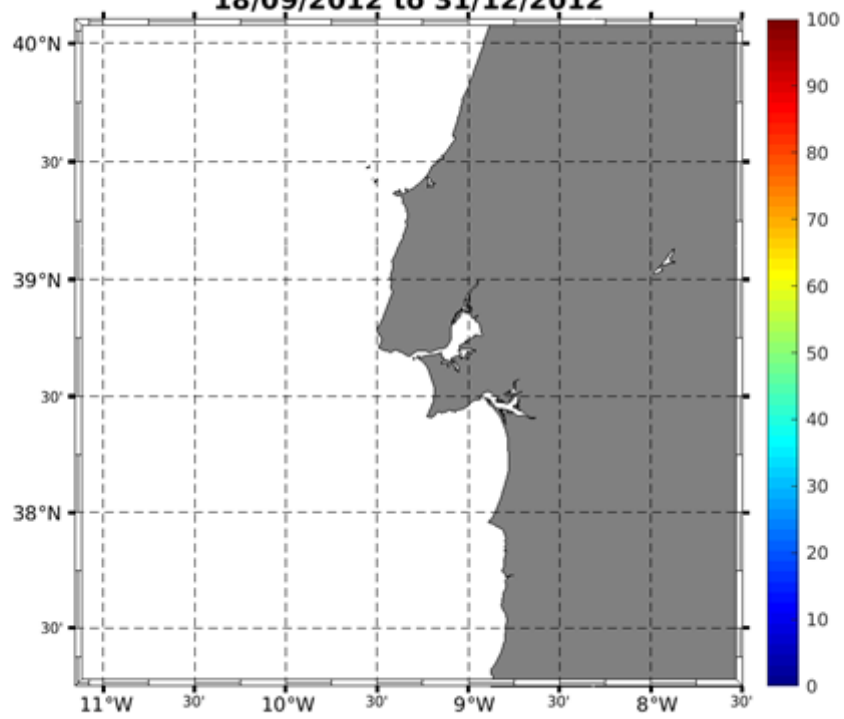


D



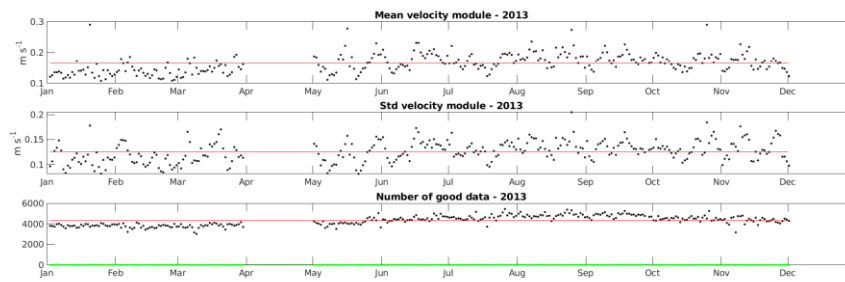
E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
18/09/2012 to 31/12/2012**



Period: 2013

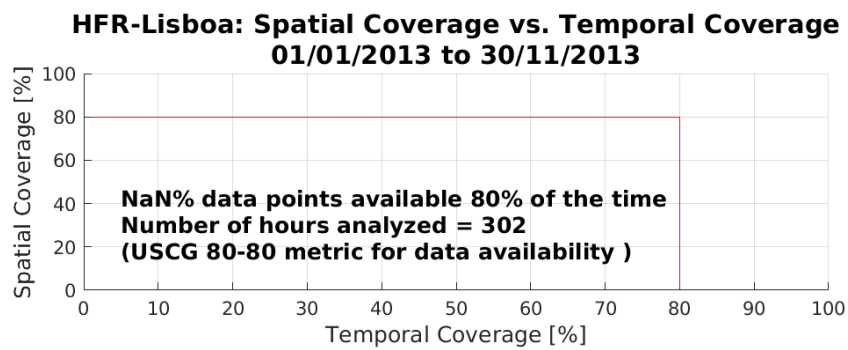
A



B

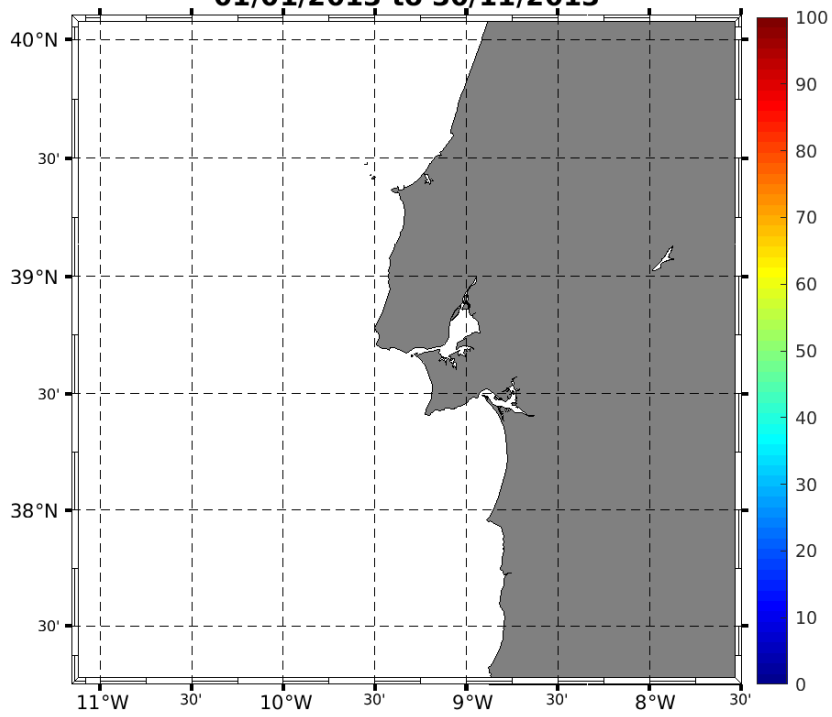


D



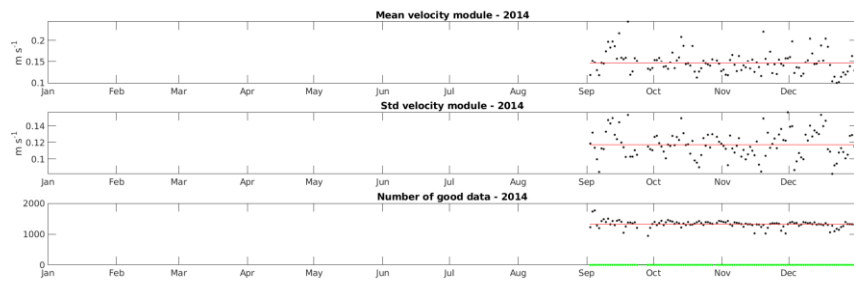
E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
01/01/2013 to 30/11/2013**

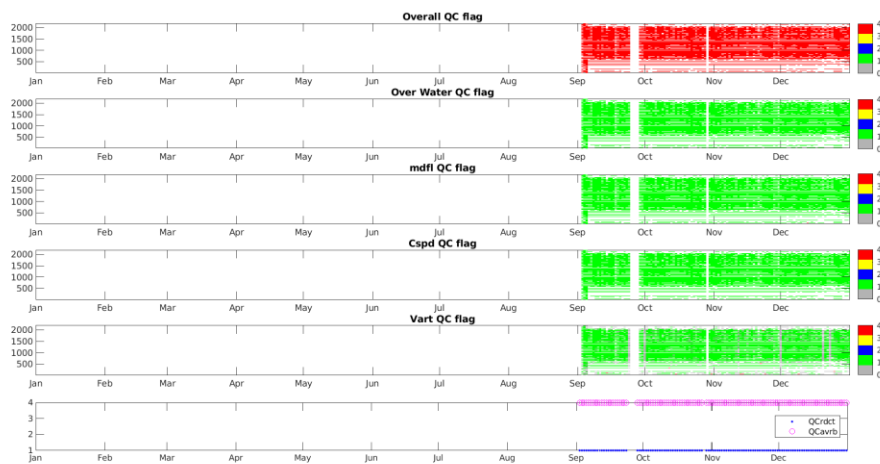


Period: 2014

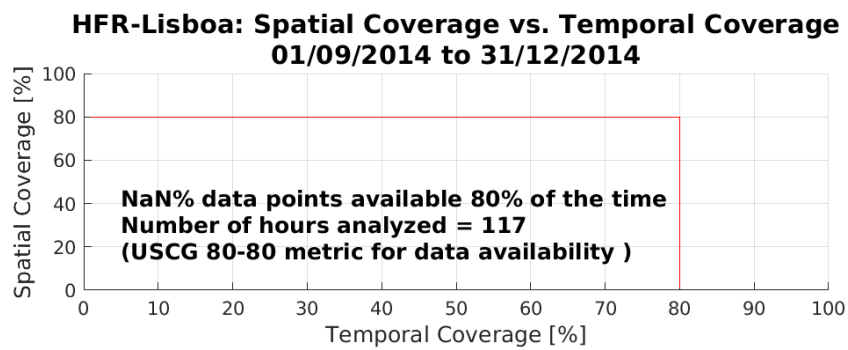
A



B

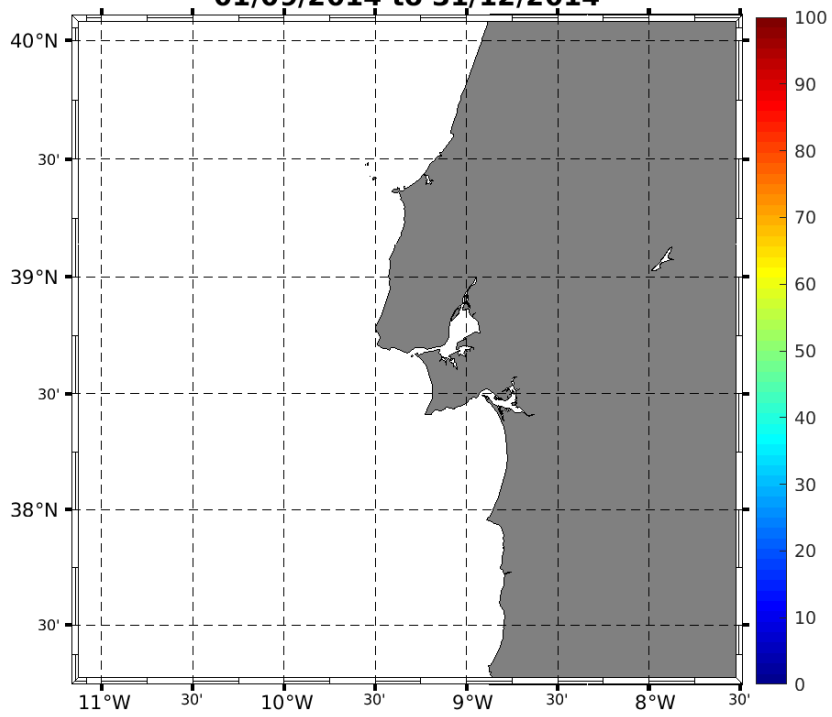


D



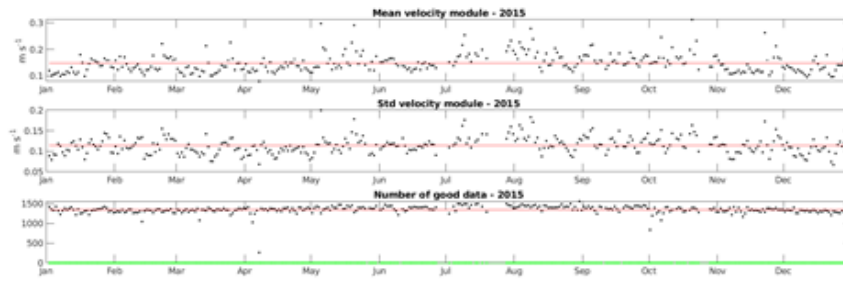
E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
01/09/2014 to 31/12/2014**

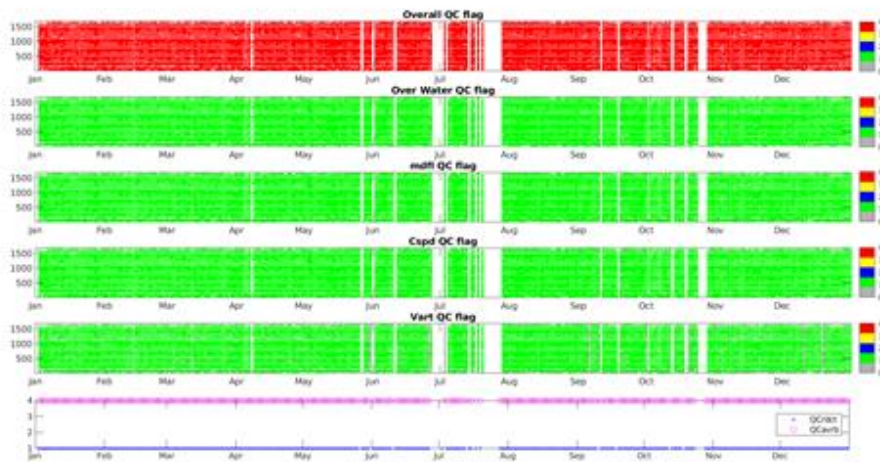


Period: 2015

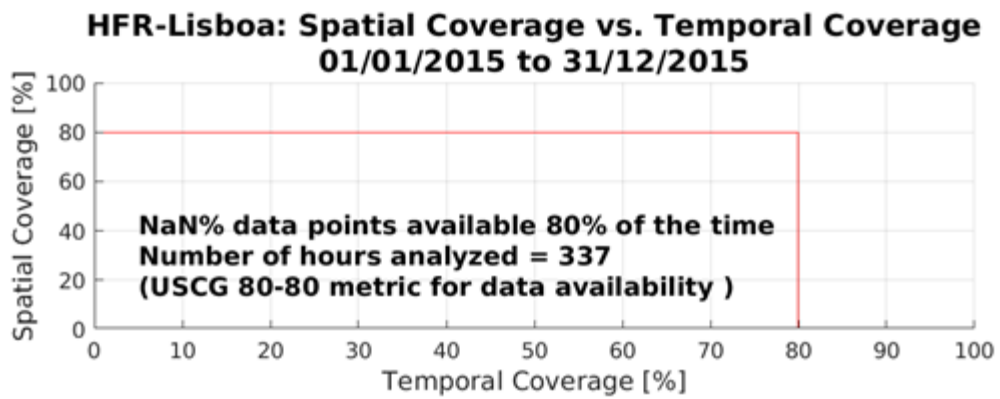
A



B

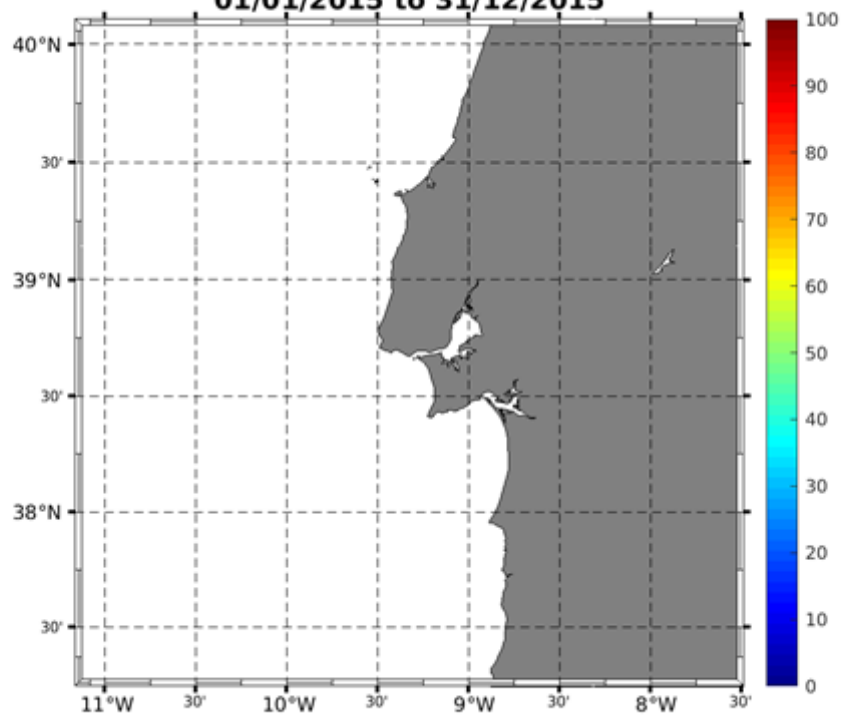


D



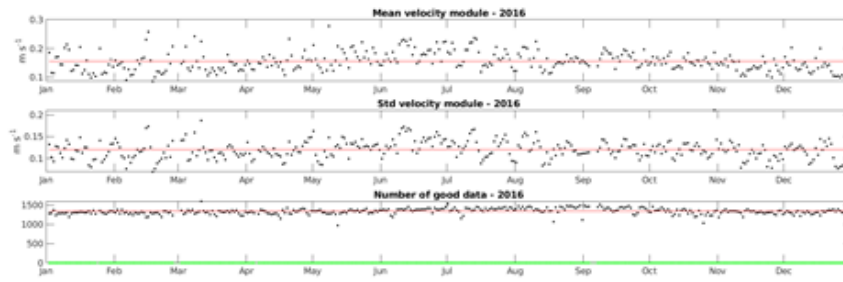
E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
01/01/2015 to 31/12/2015**

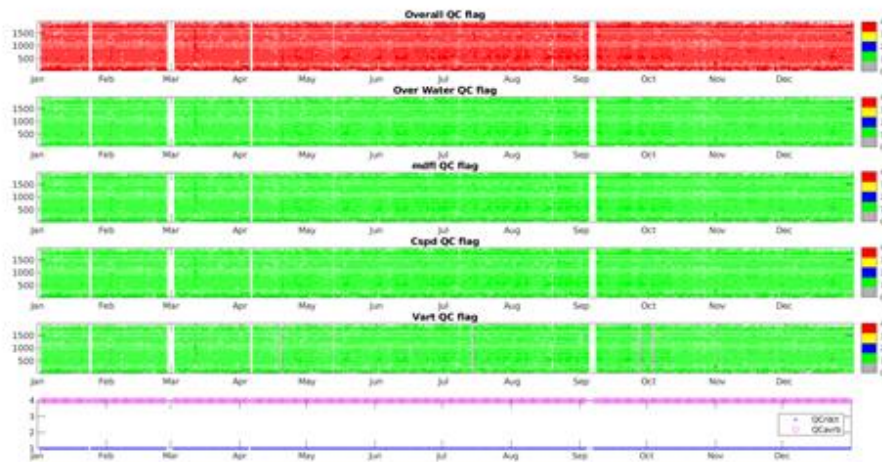


Period: 2016

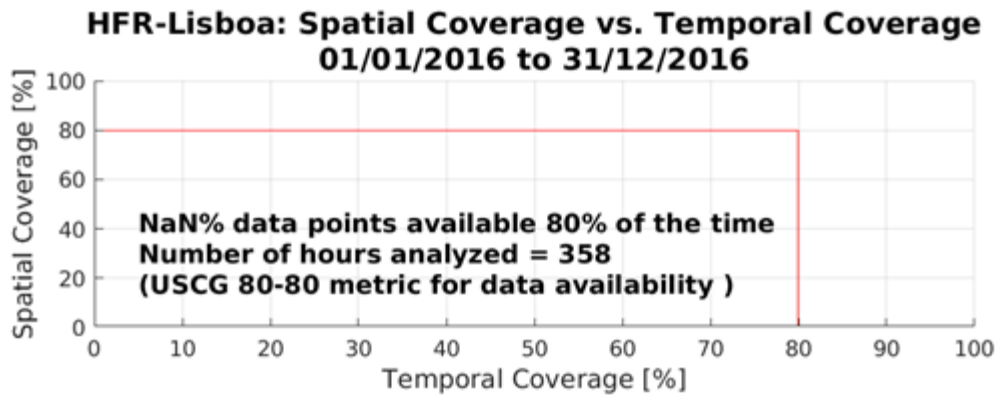
A



B

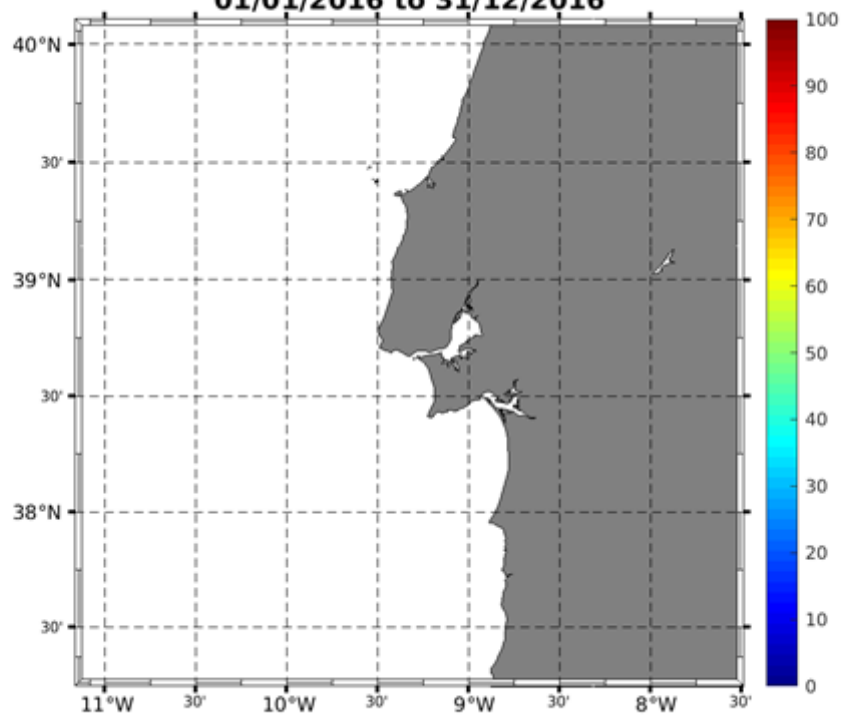


D



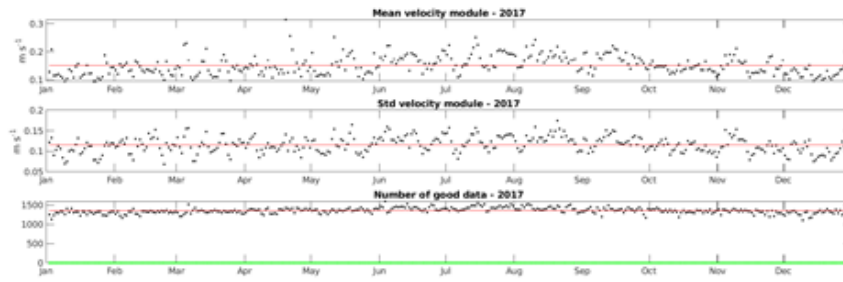
E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
01/01/2016 to 31/12/2016**

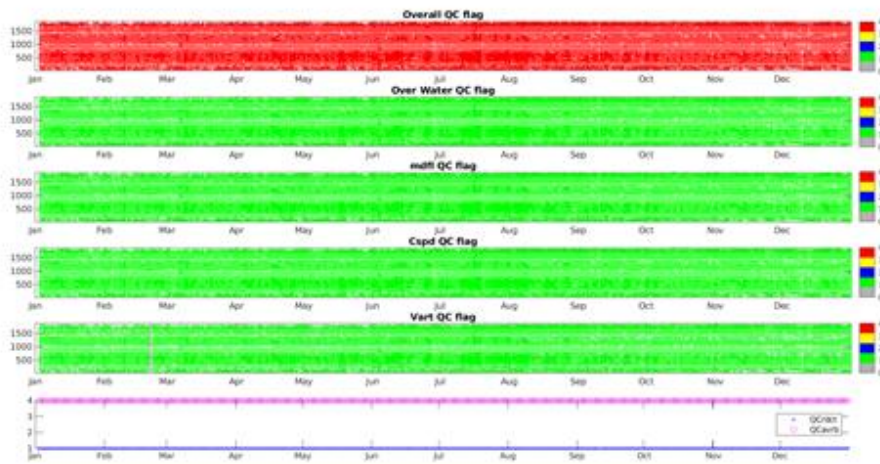


Period: 2017

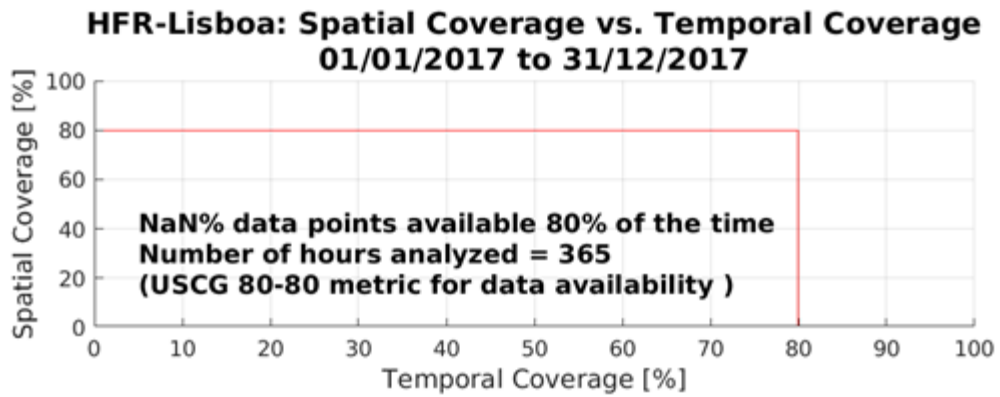
A



B

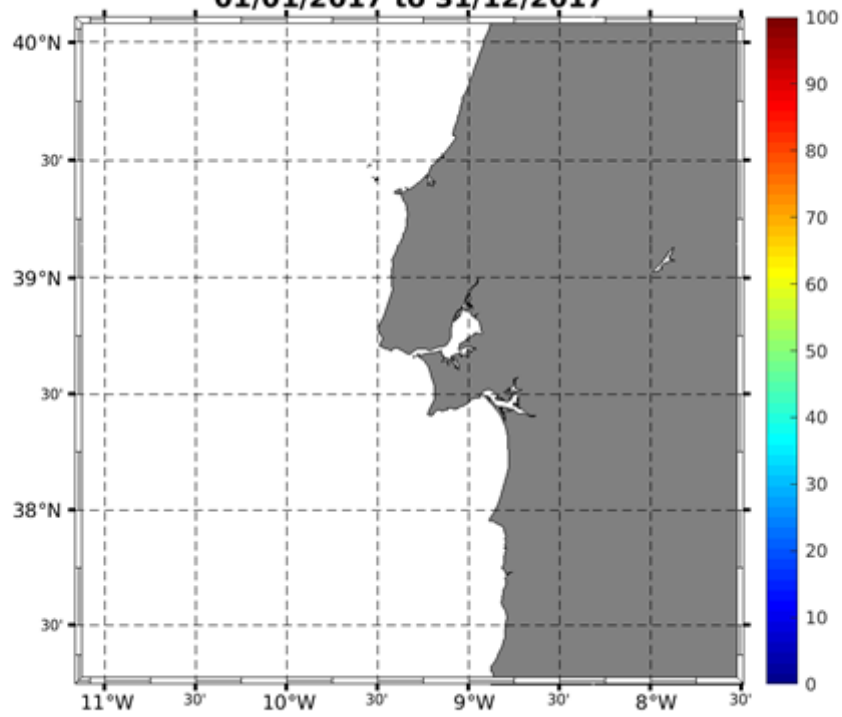


D



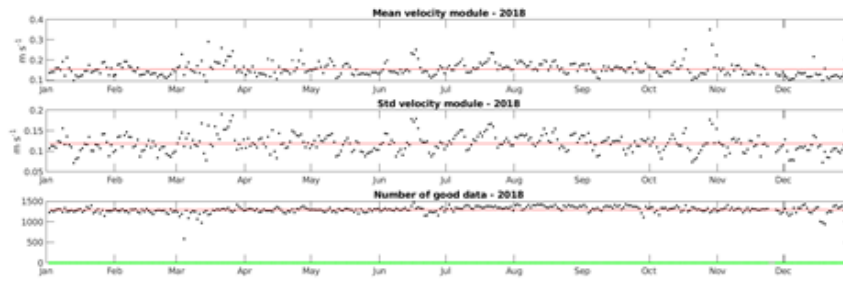
E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
01/01/2017 to 31/12/2017**

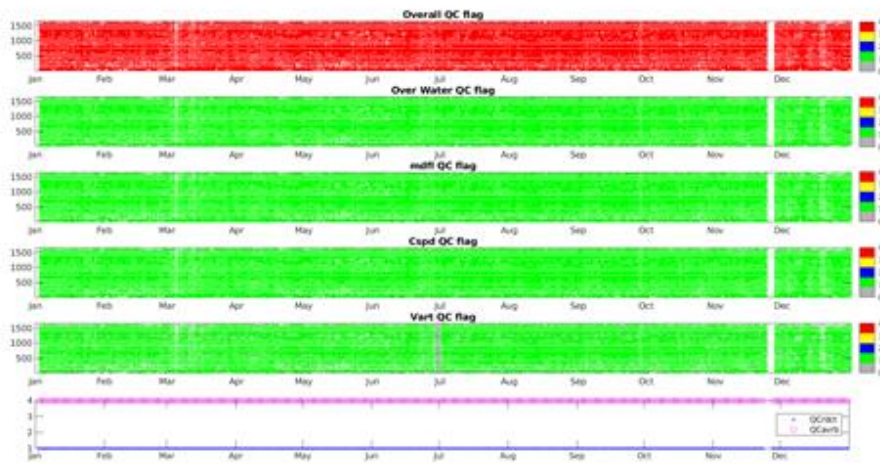


Period: 2018

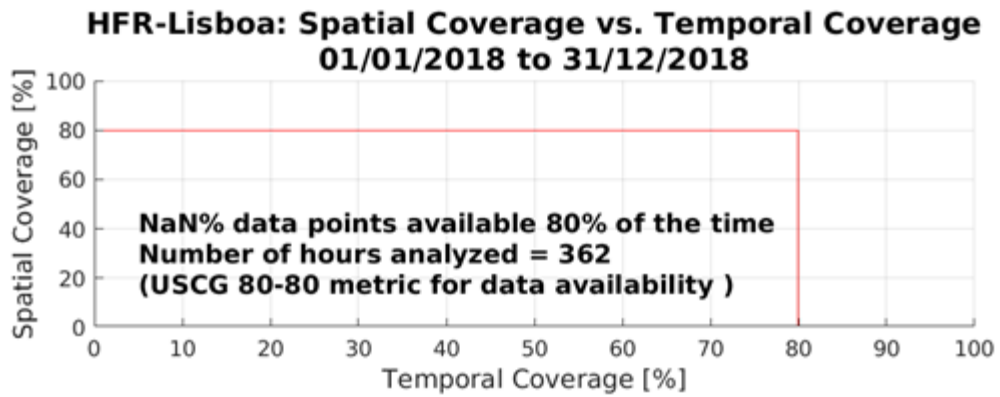
A



B

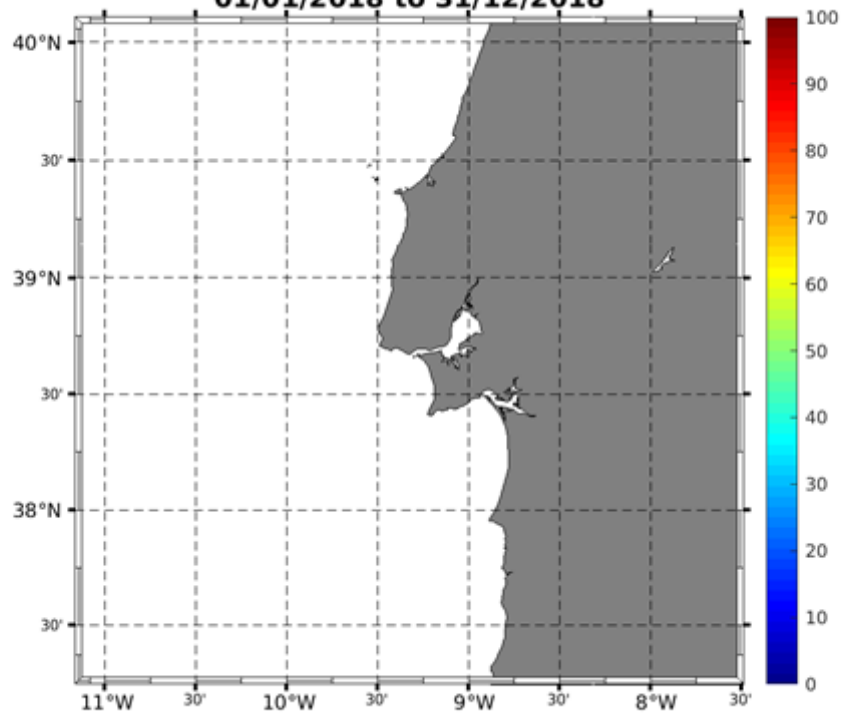


D



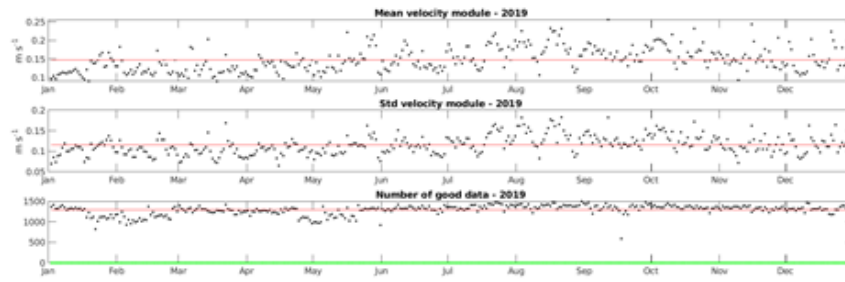
E

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

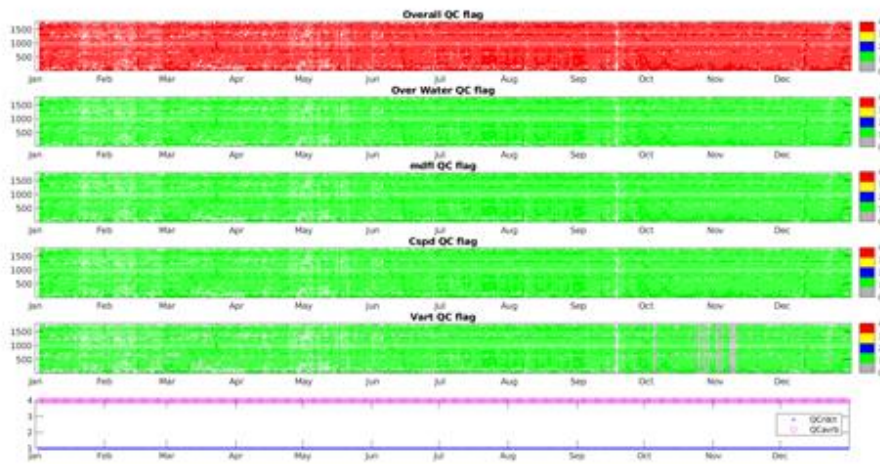


Period: 2019

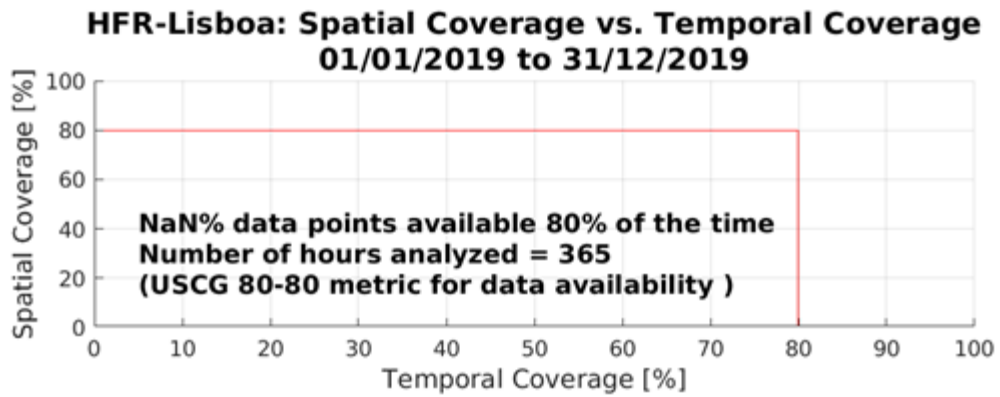
A



B

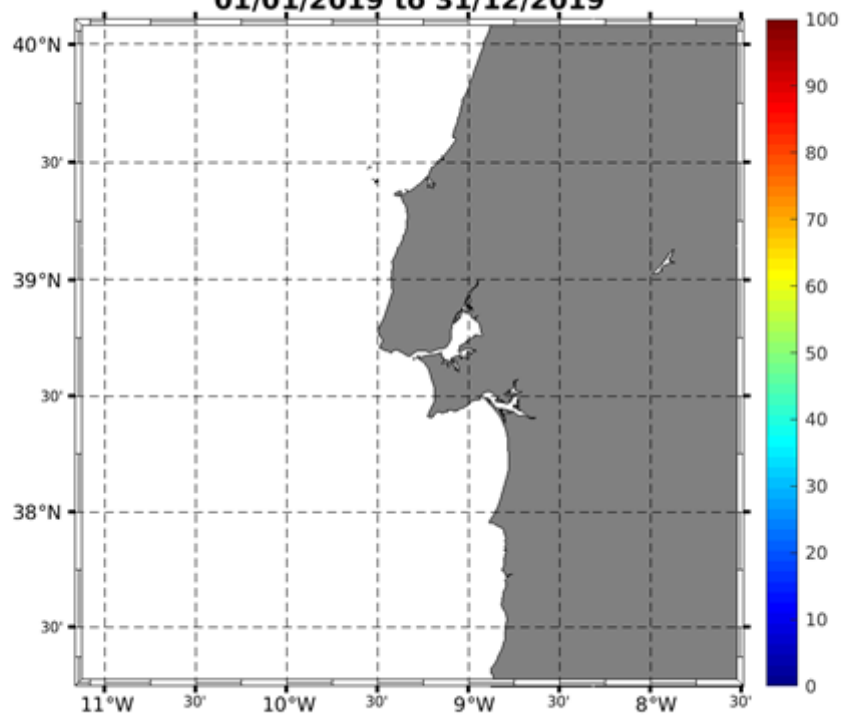


D



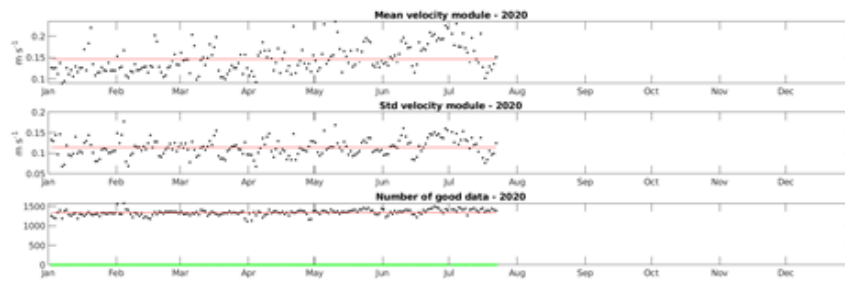
E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
01/01/2019 to 31/12/2019**

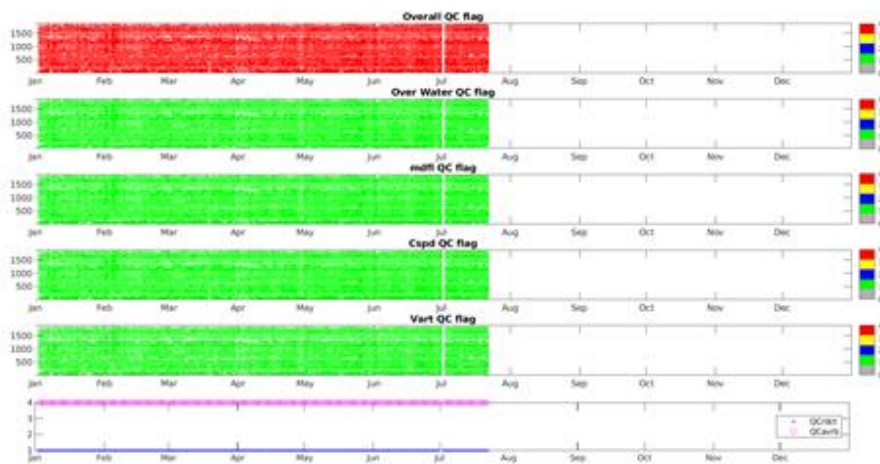


Period: Jan-Jul 2020

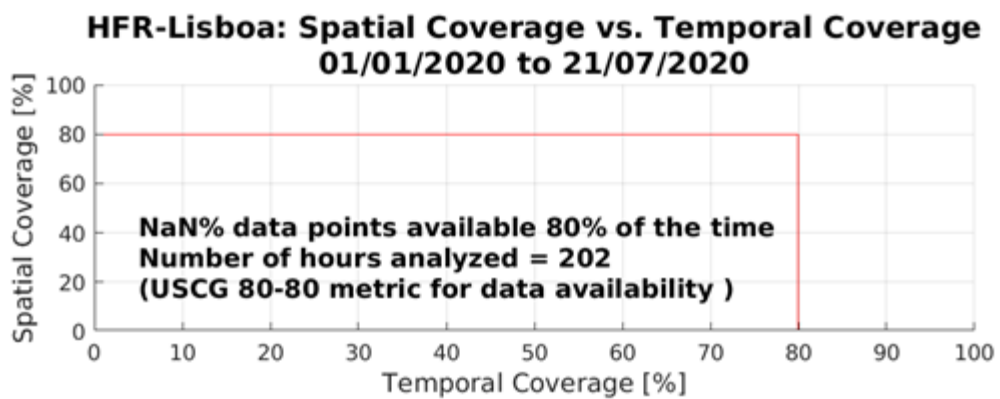
A



B



D



E

**HFR-Lisboa: Percent Total Vector Coverage (contour showing >80%)
01/01/2020 to 21/07/2020**

