Report on Historical data files QA/QC

Version of the report	Changes made by	Nature of changes
VR2020_12	L. Solabarrieta & A. Rubio	First complete version of the report

---

System: US – Puerto Rico Virgin Islands
Data set: Totals
Period: 01-Jan-2019 to 31-Jul-2020

INFO ON QA/QC Settings and Calibration

%%% QC info for time: 01-Jan-2019 to 31-Jul-2020
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 5 m/s.
OceanSITES quality flagging for variance threshold QC test. Threshold set to 1 m2/s2.

%%% Calibration info for time: 01-Jan-2019 to 31-Jul-2020
Not available for US Systems

RESULTS OF HIST DATA INSPECTION

General comments:

The total surface current data from the US systems are directly harvested from https://hfrnet-tds.ucsd.edu/thredds/catalog.html and they are transformed to comply the EU standard Netcdf format by the EU-HFR Node. They may contain less information than the EU files, namely the details of the radial sites used for the totals generation.
Summary figures are included in this document.
The reader is referred to the original source for more details

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

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

Period	General comments	Nb. analysed hours	80%-80% obj.
Jan-Dec 2019	*20.04%* spatial availability 80% of time	1742*	n**
Jan-Jul 2020	*7.55%* spatial availability 80% of time	1009*	n**

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

**the 80%-80% objective has been reached in some coastal areas (see figures E)

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.
DDNS_QC	Data Density Threshold	**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.
CSPD_QC	Velocity Threshold	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.
VART_QC	Variance Threshold	**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.
TIME_QC	Temporal Derivative	**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_QC	GDOP Threshold	**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.
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).
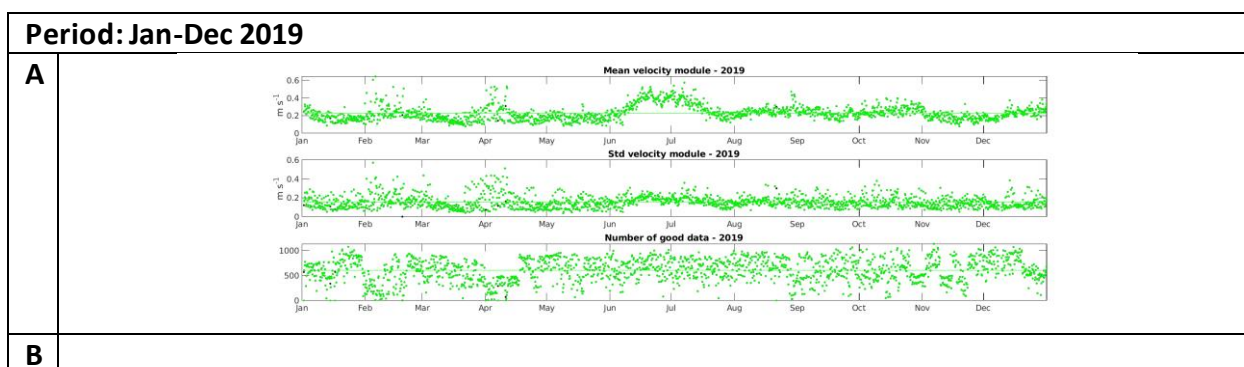
Fig B - Temporal series of the QC flags for all the grid nodes with data (not shown for US systems).

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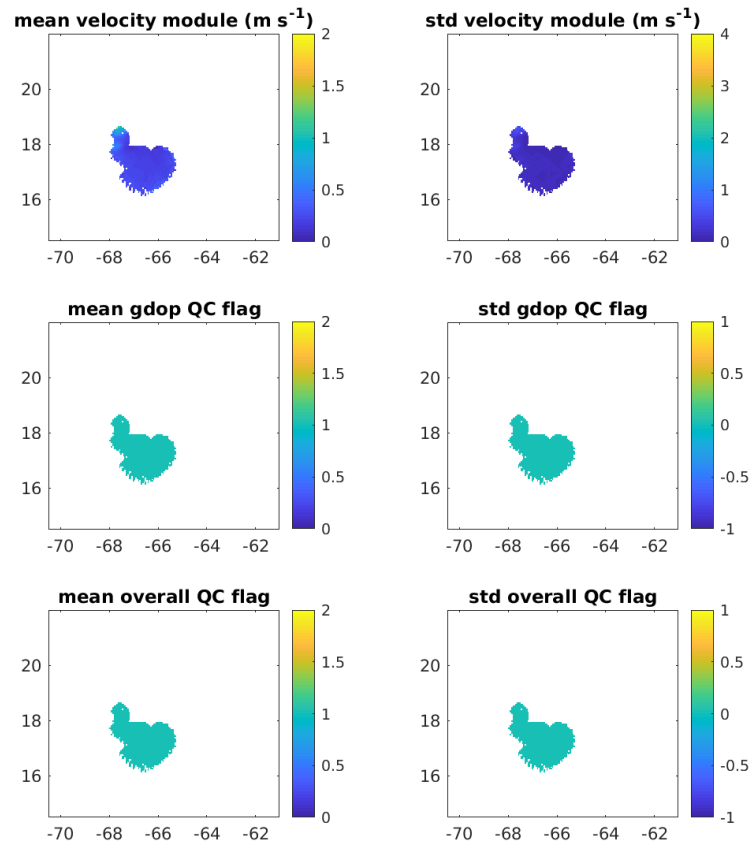
Fig D - Spatial (x-axis) vs. temporal (y-axis) coverage 80/80 annual metric. Allows to check if the system has reached the goal of providing surface currents over the 80% of the area during 80% of the time.

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.
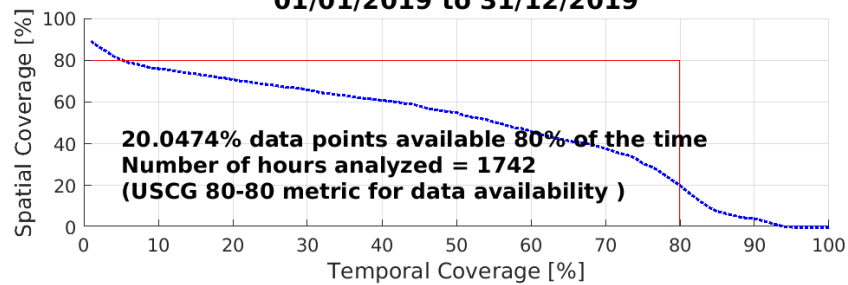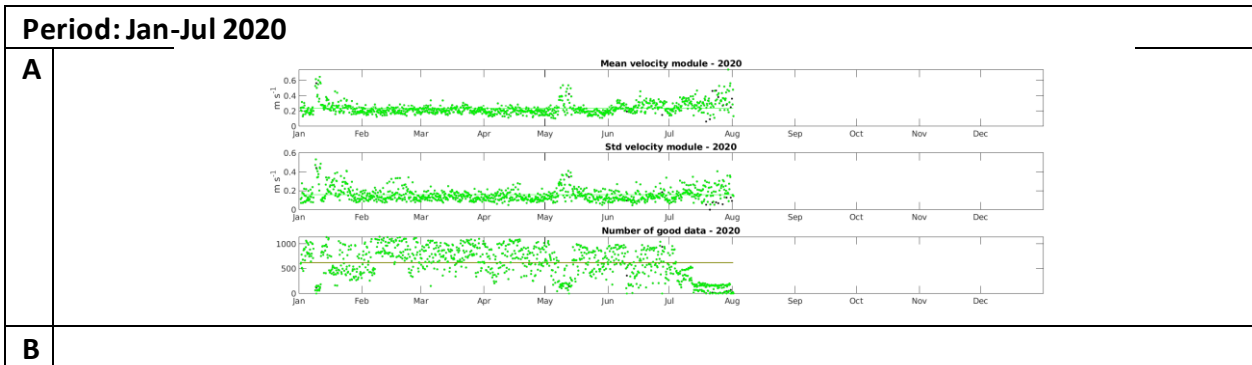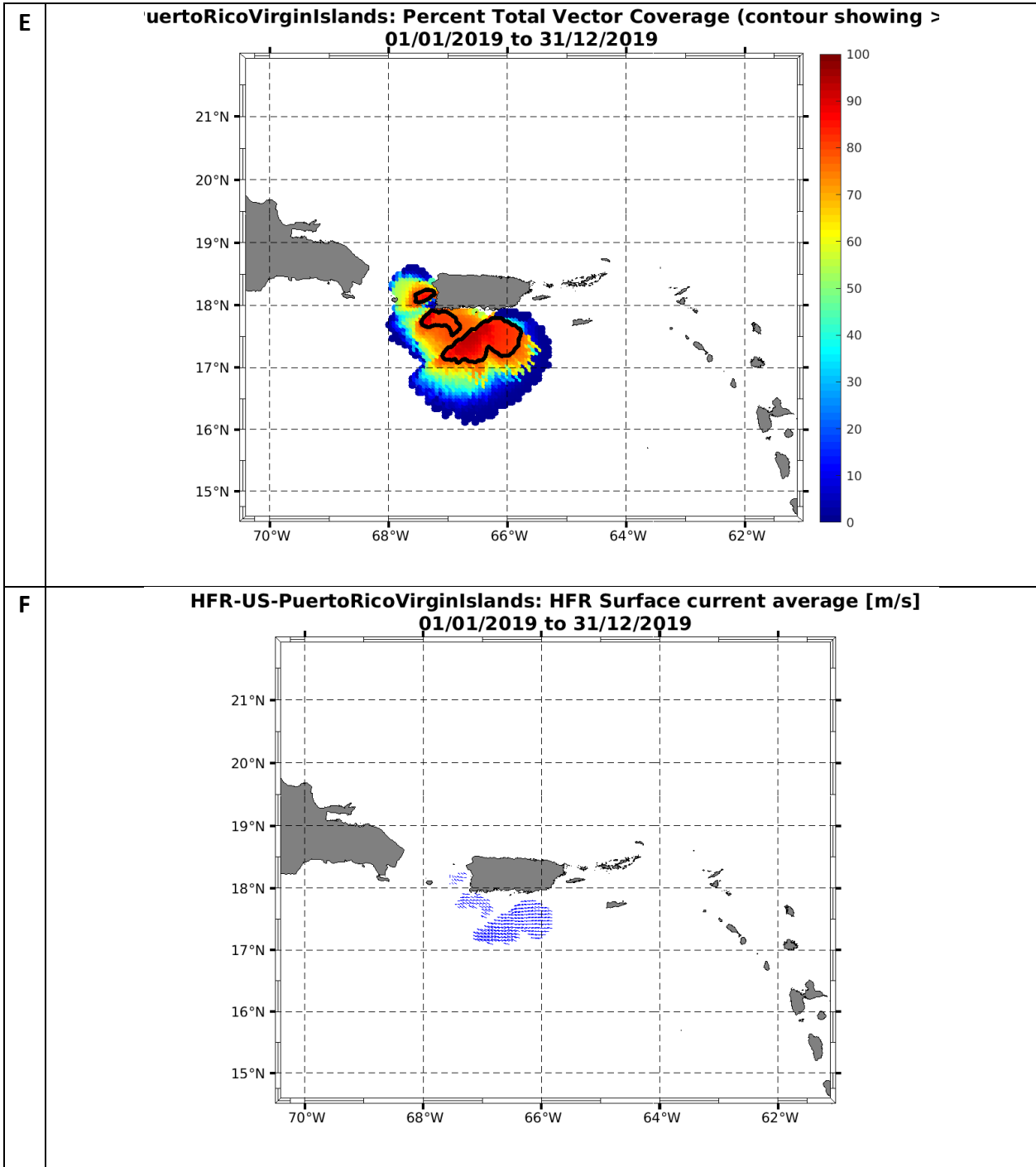
Period: Jan-Dec 2019	
A	
B	

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

FR-US-PuertoRicoVirginIslands: Spatial Coverage vs. Temporal Cove
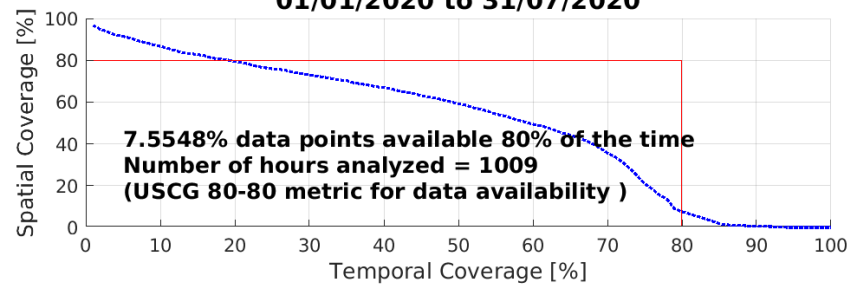**01/01/2019 to 31/12/2019**

20.0474% data points available 80% of the time
Number of hours analyzed = 1742
(USCG 80-80 metric for data availability)

Spatial Coverage [%]

Temporal Coverage [%]

| E | **PuertoRicoVirginIslands: Percent Total Vector Coverage (contour showing >**
01/01/2019 to 31/12/2019 |
|---|---|
| | |
| F | **HFR-US-PuertoRicoVirginIslands: HFR Surface current average [m/s]**
01/01/2019 to 31/12/2019 |
| | |

Period: Jan-Jul 2020

A	
B	

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

'R-US-PuertoRicoVirginIslands: Spatial Coverage vs. Temporal Cove 01/01/2020 to 31/07/2020

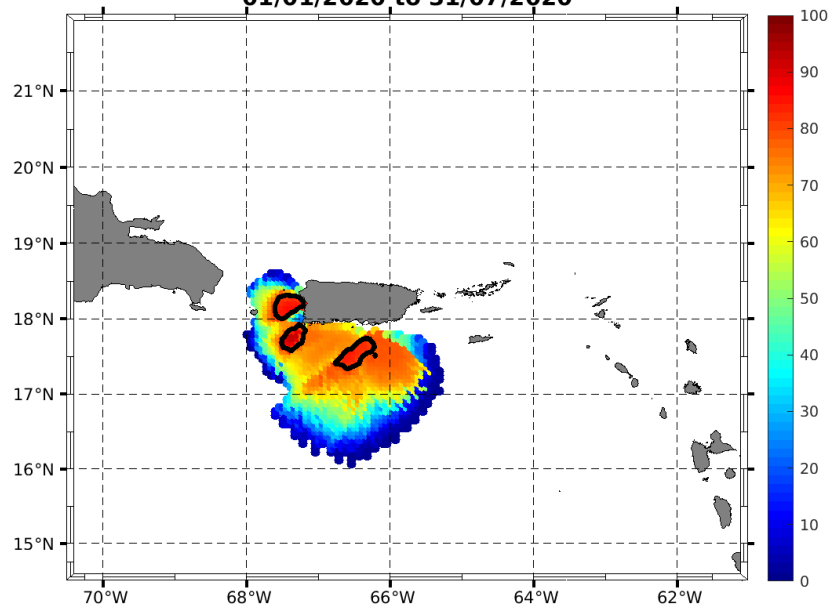**7.5548% data points available 80% of the time**
Number of hours analyzed = 1009
(USCG 80-80 metric for data availability)

E

PuertoRicoVirginIslands: Percent Total Vector Coverage (contour showing >
01/01/2020 to 31/07/2020

F

HFR-US-PuertoRicoVirginIslands: HFR Surface current average [m/s]
01/01/2020 to 31/07/2020