Climate Change Initiative+ (CCI+) Phase 1
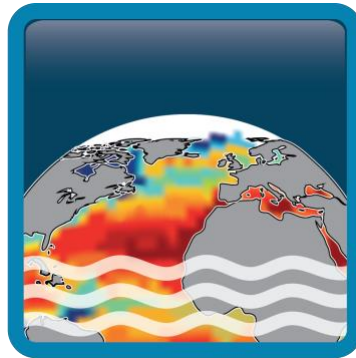# Sea Surface Salinity

Product Validation and Algorithm Selection Report (PVASR)

Customer: **ESA**

Ref.: ESA-CCI-PRGM-EOPS-SW-17-0032

Version: v3.1

Ref. internal: AO/1-9041/17/I-NB_v2r1

Revision Date: 14/09/2021

Filename: SSS_cci-D2.1-PVASR-v3.1.docx

Deliverable code: D2.1

Signatures

	Name	Signature	Date
Author	Fabrice Bonjean		14/09/2021
Author	Jacqueline Boutin		14/09/2021
Reviewed by	Jean-Luc Vergely		
	Rafael Catany		
	Roberto Sabia (Technical Officer)		
APROVED by	Jacqueline Boutin (Science Leader)		14/09/2021
	Nicolas Reul (Science Leader)	<signature>	<date>
	Rafael Catany (Project Manager)		14/09/2021
Accepted by	Susanne Mecklenburg (Technical Officer)	<signature>	<date>

Diffusion List
Sea Surface Salinity Team Members
ESA (Susanne Mecklenburg, Roberto Sabia)

Amendment Record Sheet

	Document Change Record	
Date / Issue	**Description**	**Section / Page**
15/07/2019 / v1.0	Update template and review formatting (v1.0) – First issue to ESA	Whole document
03/12/2019 / v1.1	Add reference documents	Section 1.3.2 / page 1
03/12/2019 / v1.1	Add Definitions (new section 2) terms relevant to this document	Section 2 / page 3
03/12/2019 / v1.1	Added Figure 1	Section 2 / page 4
03/12/2019 / v1.1	Improved description of the Description of the algorithm & ancillary data tested during the round robin	Section 5
28/05/2020/ v2.0	Update all comparisons with v2.3 and 2.4 CCI+SSS	Sections 4 to 8
28/05/2020/ v2.0	Add comparisons with ship binned tracks in the northern Atlantic Ocean and with Tara circle experiment in the Arctic Ocean	Section 4
28/05/2020/ v2.0	Add some comparisons with ISAS maps	Section 6.2
24/06/2020 / v2.1	Added ESA's feedback and suggestions (minor changes to previous version)	N/A
26/07/2021/ v3.0	Update all comparisons with v3.2 and 2.3 CCI+SSS	Sections 4 to 8
14/09/2021 / v3.1	Updated template and final delivery to ESA	N/A

Table of Contents

Acronyms

AD	Applicable Document
Aquarius	Aquarius NASA/SAC-D sea surface salinity mission
ATBD	Algorithm theoretical basis documents
CCI	The ESA Climate Change Initiative (CCI) is formally known as the Global Monitoring for Essential Climate Variables (GMECV) element of the European Earth Watch programme
CCI+	Climate Change Initiative Extension (CCI+), is an extension of the CCI over the period 2017–2024
E3UB	End-to-end ECV Uncertainty Budget
EASE	Equal-Area Scalable Earth (EASE) Grid
ECMWF	European Centre for Medium Range Weather Forecasts
ESA	European Space Agency
ISAS	In Situ Analysis System
LEGOS	Laboratoire d'Etudes en Géophysique et Océanographie Spatiales
LOCEAN	Laboratoire d'Océanographie et du Climat, Expérimentations et Approches Numériques
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
OTT	Ocean Target Transform
PIRATA	Prediction and Research Moored Array in the Tropical Atlantic
PMEL	Pacific Marine Environmental Laboratory
PSD	Product Specification Document
PUG	Product User Guide
PVASR	Product Validation and Algorithm Selection Report
RD	Reference document
RFI	Radio Frequency Interference
RMSD	Root Mean Square Differences
RR	Round Robin
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture and Ocean Salinity
SoW	Statement of work
SSS	Sea Surface Salinity
SSSOS	Sea Surface Salinity Observation Service
SST	Sea Surface Temperature
URD	User Requirement Document
WP	Work package
WS	Wind Speed

List of Figures

List of Tables

1 Introduction

1.1 Executive Summary

This document holds the Product Validation and Algorithm Selection Report (**PVASR**) prepared by the CCI+SSS team, as part of the activities included in the [WP200] of the Proposal (Task 2 from SoW ref. **ESA-CCI-PRGM-EOPS-SW-17-0032**).

1.2 Purpose and Scope

This document collects the information of the 3rd Round Robin algorithm comparison exercise, intended to verify the performance of the CCI+SSS v3.2 product, and to evaluate progress with respect to the preceeding v2.3 version that was previously evaluated and selected in PVASR 2020.

1.3 References

1.3.1 Applicable Documents

ID	Document	Reference
AD01	CCI+ Statement of Work	SoW
AD02	Product User Guide (PUG)	PUG
AD03	User Requirement Document (URD)	SSS_cci-D1.1-URD-v2
AD04	Product Specification Document (PSD)	SSS_cci-D1.2-PSD-v2
AD05	Algorithm Theoretical Baseline Document	SSS_cci-D2.3-ATBD_L3_L4-v3

1.3.2 Reference Documents

ID	Document	Reference
RD01	Alory G., T. Delcroix, P. Téchiné, D. Diverrès, D. Varillon, S. Cravatte, Y. Gouriou, J. Grelet, S. Jacquin, E. Kestenare, C. Maes, R. Morrow, J. Perrier, G. Reverdin and F. Roubaud, 2015. The French contribution to the Voluntary Observing Ships network of Sea Surface Salinity. Deep Sea Res., 105, 1-18, doi:10.1016/j.DSR.2015.08.005.	
RD02	Reverdin, G. and Alory, G. (2018) "Monthly binned sea surface salinity, temperature, and density in the North Atlantic subpolar gyre." The French Sea Surface Salinity Observation Service (SSS OS). doi: 10.6096/sss-bin-nasg.	
RD03	Robert R. Sokal & Rohlf, F. James, 1936- joint author (1981). Biometry the principles and practice of statistics in biological research (2d ed). San Francisco W. H. Freeman	

ID	Document	Reference
RD04	X. Yin, J. Boutin, P. Spurgeon, Analysis of biases between measured and simulated SMOS brightness temperature over ocean, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* doi: 10.1109/JSTARS.2013.2252602, 2013.	
RD05	Bureau International des Poids et Mesures, Guide to the Expression of Uncertainty in Measurement (GUM), JCGM 100:2008, 2008. Available online at http://www.bipm.org/en/publications/guides/gum.html	
RD06	Boutin, J., Y. Chao, W.E. Asher, T. Delcroix, R. Drucker, K. Drushka, N. Kolodziejczyk, T. Lee, N. Reul, G. Reverdin, J. Schanze, A. Soloviev, L. Yu, J. Anderson, L. Brucker, E. Dinnat, A.S. Garcia, W.L. Jones, C. Maes, T. Meissner, W. Tang, N. Vinogradova, B. Ward (2016b), Satellite and In Situ Salinity: Understanding Near-surface Stratification and Sub-footprint Variability, Bulletin of American Meteorological Society, 97(10), doi: 10.1175/BAMS-D-15-00032.1.	
RD07	Vinogradova, N., Lee, T., Boutin, J., Drushka, K., Fournier, S., Sabia, R., Stammer, D., Bayler, E., Reul, N., Gordon, A., Melnichenko, O., Li, L., Hackert, E., Martin, M., Kolodziejczyk, N., Hasson, A., Brown, S., Misra, S., & Lindstrom, E. (2019). Satellite Salinity Observing System: Recent Discoveries and the Way Forward. Frontiers in Marine Science, 6(243), 23p. Publisher's official version : https://doi.org/10.3389/fmars.2019.00243.	

1.3.3 Document outline

The **PVASR** is structured as follows:

Section 2 provides the definition of some terms used in the document.

Section 3 presents an overview of the carried out tasks and of the comparison results.

The main and following part of the document is composed of the description of the Round Robin (RR) methodology (section 4), where in section 4.1 we describe the in-situ data used in the RR tests (ship tracks and PIRATA moorings), in section 4.2 the colocation methodology between in-situ and satellites for moorings and ship tracks, and in sections 4.3 and 4.4 the metrics and plots used in the RR tests.

We then present the CCI+SSS satellite data products that are evaluated and compared to in situ data in section 5.

In section 6, we show some results of the RR exercise (section 6.1), some independent comparisons using the ISAS dataset (section 6.2), a summary of the most salient comparison features between the two products, and open issues (section 6.3). We enumerate some perspectives for future RR in section 7.

2 Definitions

We provide below definitions, taken from [RD 05], and considerations adapted to the round robin exercise for evaluating the algorithm for satellite SSS products, that have been adopted throughout this document:

Measurand: particular quantity subject to measurement, in our case, the salinity, defined as the relative amount of salt dissolved in sea water (corresponding to gram of salt per kilogram of sea water) at the sea surface.

Error: result of a measurement minus a true value of the measurand. Since the 'true' value of the measurand is not known, the 'true' value of the error is unreachable.

Uncertainty: parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand. Uncertainty of measurement comprises, in general, many components. In the case of RR, since measurements are validated by comparisons with measurements in the fields, 'experimental standard deviations' classically evaluated from the statistical distribution of the results of series of measurements realized in the same conditions, cannot be estimated. Hence, in the case of RR, the uncertainty is evaluated from assumed probability distributions of the measurand derived, with some uncertainty, from in situ measurements.

In [RD 05], 'it is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion'. In the case of satellite radiometric measurements, the absolute calibration of the SSS is not well known and important differences between the various satellite SSS come from the different systematic corrections that are applied. As a consequence, we will distinguish between 'uncertainties associated with systematic effects' (that can be quantified by a bias - see below), from the 'uncertainties associated with random errors' coming from the noise of the measurements (linked to the radiometric resolution), from errors that are not well characterized given the present knowledge of the sources of errors.

Discrepancy: The difference between the data product and the validation value.

(Relative) Bias: The mean value of the discrepancy.

Validation: The process of assessing by independent means the quality of the data products derived from the system outputs.

Precision: The difference between one result and the mean of several results obtained by the same method, i.e. reproducibility (includes non-systematic errors only).

Observational errors: Observational errors are the ones corresponding to the precision of the instruments, plus when available, the ones due to inaccurate absolute calibration. The precision of in situ SSS is generally considered to be less than 0.01 for an individual measurement but absolute calibration of merchant ships TSG can be as large as 0.1 for a given transect. For satellite SSS, the absolute calibration error is usually unknown, the precision is on the order of 0.4 - 0.6 for individual SSS in warm regions as retrieved from Aquarius, or from SMOS and SMAP respectively. These observational errors are reduced at level 3 and level 4 according to the number of satellite passes occurring in the same pixel over one week, by roughly a factor $\sqrt{2}$ for Aquarius, and a factor 2 to 3 for SMOS and SMAP. Since an absolute reference is usually not available, what is provided in the products is an **observational uncertainty** (see E3UB report).

Sampling errors: According to [RD 07], sampling errors arise when one data type does not represent a process (or scale) that the other does, e.g., due to the differences in their spatial and/or temporal samplings. The "expected" differences, i.e. the low bound at which two estimates are allowed to differ, are called in the following **sampling uncertainties**.

Satellite SSS: Sea Surface Salinity within the first centimetre of the sea surface, by nature integrated over a surface that depends on the radiometer characteristics and on the data processing.

In-situ SSS: Near Surface Salinity measured at several cm to several meter depth (see Figure 1).
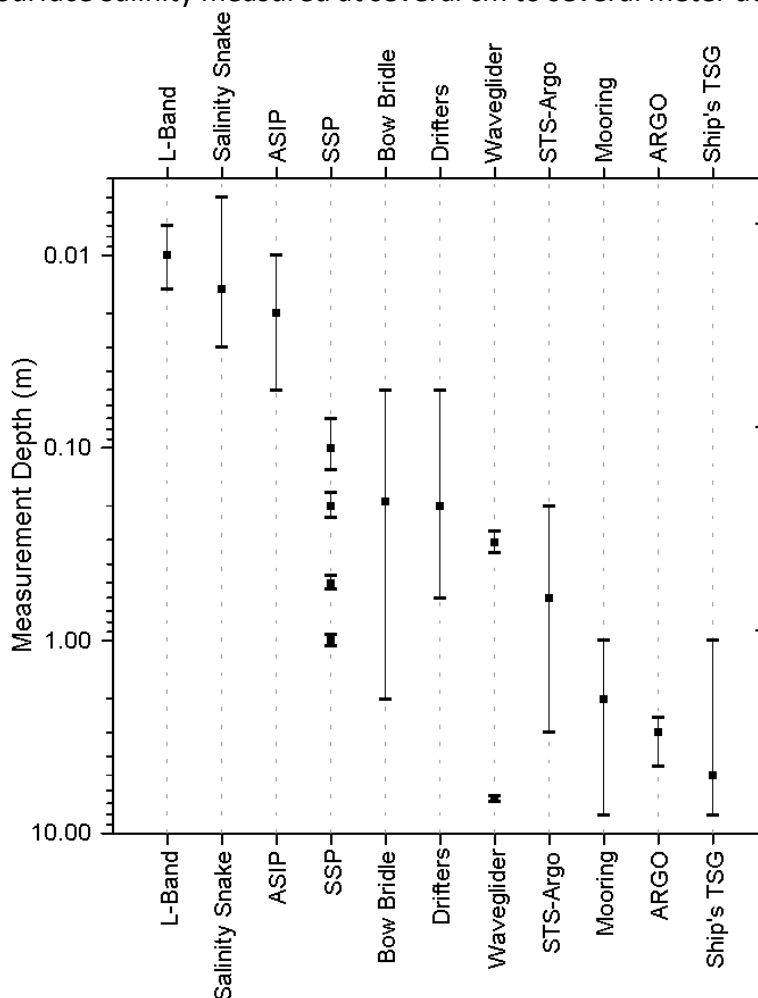
Figure 1: Scale portraying the typical depth at which near-surface salinity is measured by various sensors/platforms. The small squares show the average measurement depth and the capped lines show the range for that average. For profiling platforms (ASIP, Bow Bridle, STS-Argo, Argo) the range represents the variability of the top-most point in the profile. For platforms with standardized configurations that measure at fixed depths (Salinity Snake, SSP, Wave Glider) the mean and range of each sensor at a particular depth are shown. For platforms where there are multiple sensor configurations (drifters, mooring, shipborne TSG) or that sample at different depths depending on the specifics of the platform, the range of measurement depths across all platforms is shown. Radiometric penetration depths were calculated using the Stogryn (1997) relationship and show penetration depths at 1.43 GHz over the salinity range of 20 pss to 38 pss and temperature range of −2 ºC to 35 ºC (where the "mean" value shown in the figure is for 20 ºC and 35 pss). (Figure taken from [RD 06])

3 Overview

We use three metrics for assessing the performances of the algorithms/products. They are designed to characterize uncertainties coming from three different types of errors that are handled differently in the various satellite processing:

- **M1: The robust standard deviation between satellite SSS and in situ SSS**: this characterizes random errors that are expected from measurements noise, from errors that are not well characterized given the present knowledge of the error sources, etc… Statistical robustness is strengthened by the use of the median rather than the mean calculation since the median is much less influenced by extreme values and outliers.

- **M2: The bias between satellite SSS and in situ SSS**: this characterizes systematic errors that are expected from e.g. radiometer calibration issues, land-sea contamination, sun contamination, etc…

- **M3: The coefficient of determination (square of correlation coefficient) between satellite SSS and in situ SSS:** this is indicative of the signal to noise ratio and is very sensitive to overly stringent filtering or smoothing of extremum values (e.g. low SSS in river plumes).

- **M4: The centered reduced variable**: its statistical distribution properties inform on the appropriateness of CCI L4 SSS uncertainties.

The significance level of the difference between metrics derived from different products is evaluated using classical statistical tests as described in section 6.1. These significance tests take into account estimates of the 'observational uncertainties' and 'sampling uncertainties'.

The RR tests of the ongoing CCI+SSS project include comparisons of salinity retrievals against in-situ observations in the Atlantic Ocean. We use two in situ data sets: 1) from the PIRATA moorings, and 2) from the repetitive ship tracks across the Atlantic Ocean.

We compute metrics between CCI+SSS data and in-situ SSS over several periods of time coinciding with the time spanned by the different satellite missions; this includes the total period which comprises all periods up to present and that is currently secured by SMOS. This is in order to compare the same metrics with respect to the satellite data availability during the constellation time (SMOS + Aquarius + SMAP SSS).

We co-locate satellite and in-situ data with a satellite centered methodology. Results of this RR exercise show that the new version 3.2 of CCI+SSS yields better results globally in comparison to the preceding product, for all the periods (Table 4). We observe significant improvements with respect to v2.3 especially in high latitudes and near coasts at mid-latitudes.

4 Round robin methodology

In this section we present the description of the RR test methodology. The aim of these tests is to provide some metrics to compare and validate different satellite products. In this third version, we do not prioritize metrics for algorithm selection, as only one new CCI product is candidate to succeed to the previous 2.3 version. The algorithm team tested another version, 3.1, but since preliminary comparisons showed degradation with respect to version 2.3, this is not reported here.

We introduce a new coefficient based on a centered variable reduction that considers the CCI SSS uncertainty in its calculation and that is intended to validate this uncertainty.

As in the previous PVASR, we decided to focus on Atlantic Ocean, because this is an area 1) of contrasted SSS regimes with very high and stable SSS in the subtropics and very low SSS near the Amazon and Congo plumes, 2) with some regions strongly affected by RFIs, and 3) completely independent of the SMOS Ocean Target Transformation region (RD04: Yin et al. 2013) where SMOS Tbs are a posteriori calibrated.

This region is well monitored by numerous in-situ measurements along repetitive ship tracks crossing maximum and minimum SSS regions, or by moorings, allowing to monitor the data quality over the whole satellite period.

4.1 In situ data

4.1.1 Ship tracks

We use regular merchant ship tracks along AX20 and AX11 lines between Europe and South America, from ORE SSS data base (http://sss.sedoo.fr/). Note that results presented in this report come from datasets that have not been updated in the source database since 2019.

AX20 salinity measurements come mainly from Toucan and Colibri merchant ships. AX11 measurements come from Cap San Lorenzo, Rio Blanco and Santa Cruz merchant ships (**RD01**). AX20 and AX11 areas are delimited using linear boundaries, defined as $Lat = a.Lon + b$, with a and b coefficients shown in Table 1.

A transect is considered as AX20 if the majority of its points below 45°N are between lines 1 and 2 (Figure 2).

A transect is considered as AX11 if the majority of its points below 45°N are between lines 2 and 3.

We consider only points below 45°N to determine if a transect is AX20 or AX11, because these two lines overlap in northern latitudes.

Only points included in these areas, more than 40 km away from coasts and below 45° north, are taken into account in the statistics.

We only use ship tracks that extend in latitude at least from 20°N to 35°N.

The quality check is used to keep only good and probably good data.

Considering all these measurements, we get between 6 and 9 good quality ship transects for AX11 and AX20 per year, except for 2019 (Figure 3).

Salinity measurements are collected at 5 meters depth for all AX20 tracks and at 10 meter depth for AX11 transects.

Figure 2: Area coverage of ship tracks AX20 and AX11. The numbers in brackets refer to the geographical limits we defined to attribute ship tracks to AX20 or AX11 dataset (equations coefficients set in Table 1).
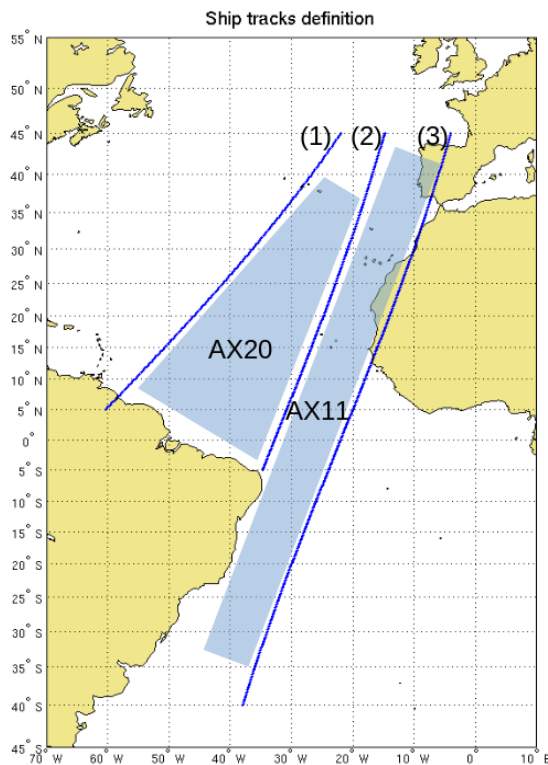
	a	b
Line 1	1.04	67.8
Line 2	2.5	82
Line 3	2.5	55

Table 1: linear boundaries used to delimit AX11 and AX20 ship tracks used in the RR tests
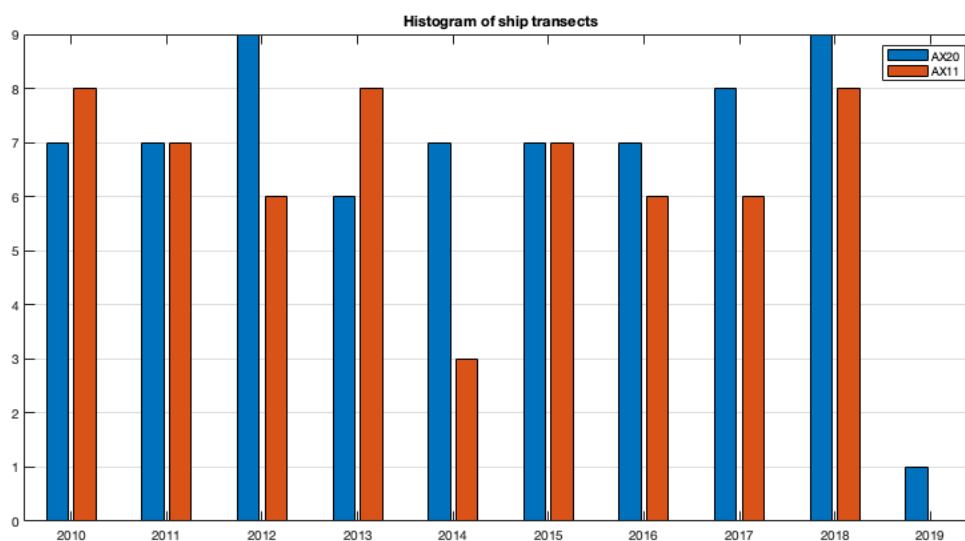
Figure 3: Histogram of ship transects

4.1.2 PIRATA

The second dataset we use is the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA). PIRATA is composed of 18 moorings in tropical Atlantic, between 20°S and 20°N (Figure 4). These moorings provide daily measurements of salinity. We use a dataset from 2010 to 2019 (the database covers 2020, however the CCI+SSS v2.3 that is used in RR ends in 2019). We use delayed mode data from PMEL (http://www.pmel.noaa.gov/gtmba/pmel-theme/atlantic-ocean-PIRATA), corrected from biases and drifts by Gilles Reverdin (LOCEAN) and Elodie Kestenare (LEGOS) (personal communication). This dataset is available from 2010 to 2019

For our tests, we only use data collected at a depth smaller than five meters.
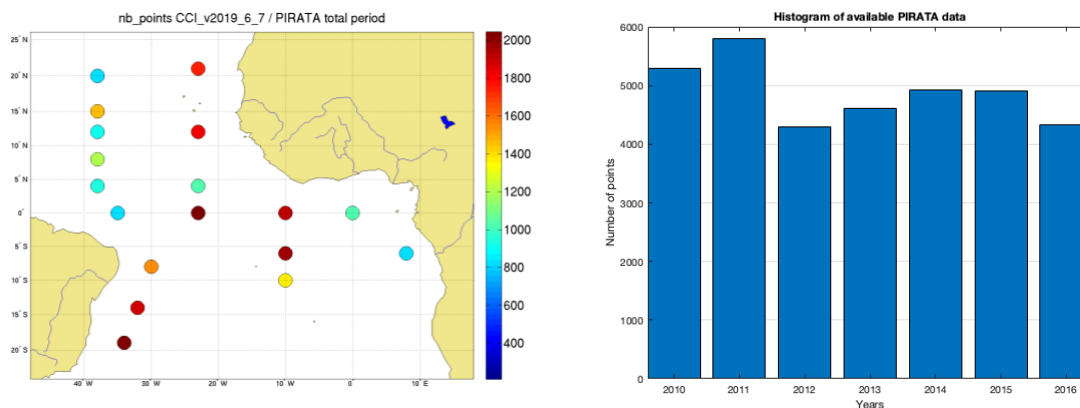
Figure 4: Number of points from 2010 to 2016, PIRATA moorings

4.1.3 Monthly binned ship tracks in North Atlantic

From the previous RR version onwards, we have added new datasets in northern latitudes. The first dataset is a binned SSS that covers regularly sampled ship-of-opportunity lines BAX01 and BAX02, maintained by Gilles Reverdin and Gaël Alory. Grid boxes are shown in Figure 5. The binned data were made available by the French Sea Surface Salinity Observation Service (http://www.legos.obs-mip.fr/observations/sss/). These data are available from 1993 to 2019.
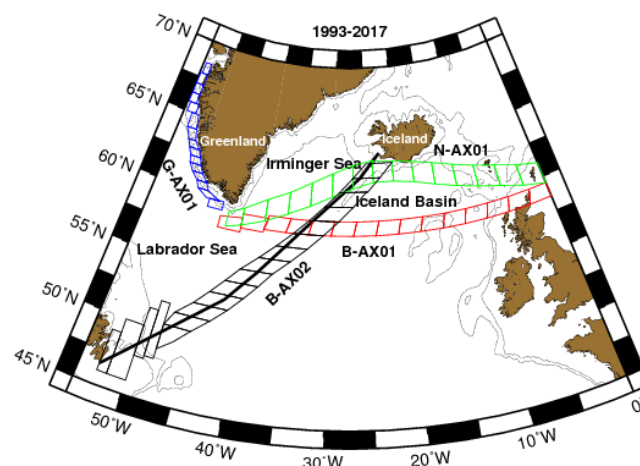
Figure 5: grid boxes provided along lines B-AX01 (red), B-AX02 (black), (Other transects are not used in our study).

4.2 Colocation methodology

Our colocation methodology is centered on satellite measurements. We adapt the in-situ measurements in order to get an equivalent representativity of data to the satellite product we study.

4.2.1 Moorings

Spatial colocation: we compare mooring data to the nearest neighbour satellite pixel (i.e. given the spatial resolution of the satellite grids, at less than 0.125° for SMOS, 0.1° for SMAP and 0.5° for Aquarius measurements).

Temporal colocation: mooring data are averaged over the satellite data temporal averaging period (T) (e.g. 7-days or one month for CCI comparisons depending on the product resolution) at each available day of the period. We use a mean weighted with a gaussian distribution (sigma=T/4). An example is provided in Figure 6.
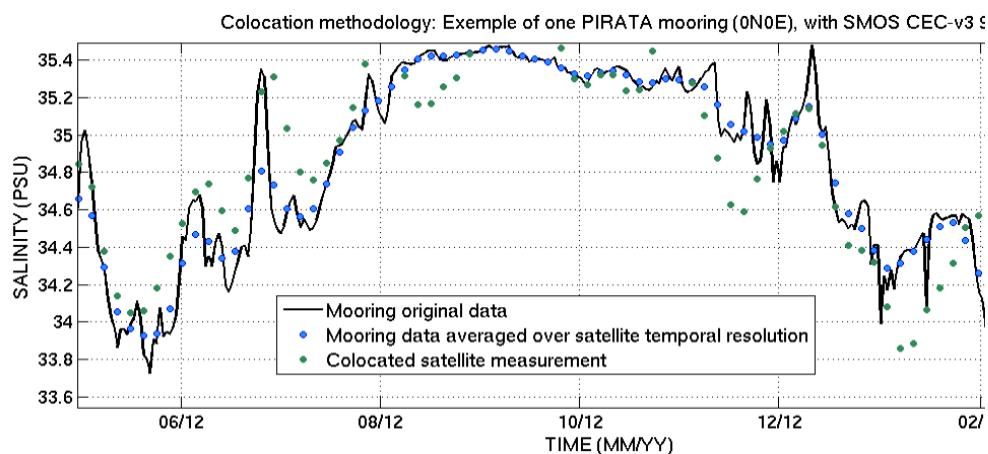
Figure 6: Moorings colocation methodology example.

4.2.2 Ship tracks

Spatio-temporal colocation: Ship measurements are smoothed along the track, over the spatial resolution of the satellite, using a gaussian weight.

The standard deviation in the gaussian function is set to one quarter of the spatial resolution (95% of the weight within a radius of half the spatial resolution).

This smoothing is limited to 2 hours before and 2 hours after the central point, which is usually much longer than the time spent by the ship to cross the pixel.

Then, we average ship measurements taken successively over the same satellite pixel.

Only pixels with more than 4 available in-situ measurements are used for the statistics.

The resulting colocations are compared to the nearest corresponding satellite pixel in time.

An example is provided in Figure 7.

4.2.3 Monthly binned ship tracks in northern Atlantic

To compare satellite products to the monthly binned ship tracks dataset, satellite data are averaged over one month and in boxes described in Figure 5.
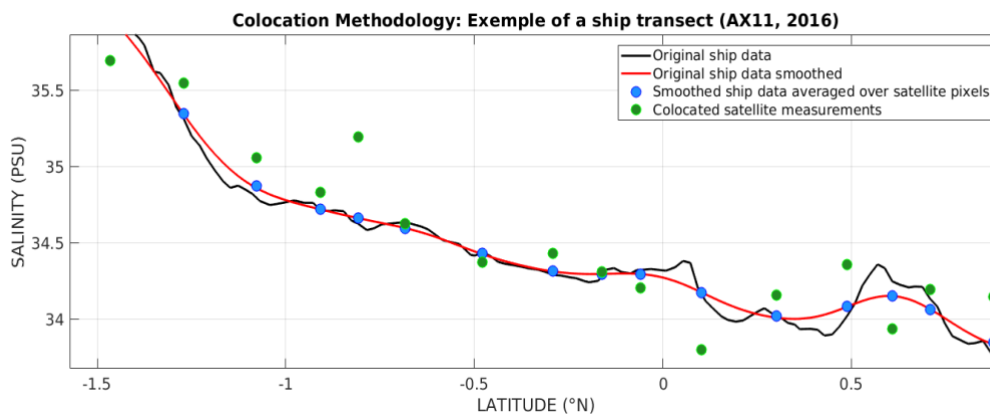
Figure 7: Ship track colocation methodology example

4.3 Metrics

For selecting the algorithm, the metrics introduced in section 3 are computed as described below (*horizontal bars indicate the mean over a set of measurements)*. The significance tests of the difference between metrics derived from different products are described in section 6.1.

- Standard deviation of the differences (*std diff*):

$$std\_diff = \sqrt{\overline{(SSS_{in\text{-}situ} - SSS_{satellite})^2} - \overline{(SSS_{in\text{-}situ} - SSS_{satellite})}^2}$$

- Robust standard deviation (*std diff rob*):

$$\frac{median(|(SSS_{in\text{-}situ} - SSS_{satellite}) - median(SSS_{in\text{-}situ} - SSS_{satellite})|)}{0.6745}$$

- Bias:

$$bias = \overline{SSS_{in\text{-}situ} - SSS_{satellite}}$$

- Coefficient of determination (r^2) of the linear regression y between in-situ SSS and satellite SSS:

$$r^2 = max\left[0,1 - \frac{\sum(Y - y)^2}{\sum(Y - \overline{Y})^2}\right], \text{ with } Y = SSS_{in\text{-}situ} - SSS_{satellite}$$

- Standard deviation of the reduced centered difference (*std diff cr*).
  The SSS difference is divided by the time and space varying uncertainty magnitude as given below:

$$\Delta SSS_{cr} = \frac{\Delta SSS}{sat_{uncertainty}}$$

for each measurement point, where sat$_{uncertainty}$ is the satellite uncertainty at that point (random uncertainty estimated from the L4 generation).

The std is calculated using this scaled and non-dimensional formulation in place of $SSS_{in\text{-}situ} - SSS_{satellite}$. Variable centering is applied as part of the std diff formulation above.

- Robust std of the reduced centered difference (*std diff cr rob*):
  As above using the median calculation.

In these equations, $SSS_{\text{in\_situ}}$ corresponds to the salinity of the in-situ measurements, after the colocation processing described in section 4.2.

$SSS_{satellite}$ corresponds to the salinity sensed by satellite.

In addition, we compute the *rmsd*:

$$rmsd = \sqrt{(SSS_{\text{in-situ}} - SSS_{satellite})^2}$$

We do not use rmsd for evaluating the algorithm as this coefficient is related to the bias and std diff that are already considered:

$$rmsd = \sqrt{std\_diff^2 + bias^2}$$

4.4 Figures

We use three kinds of representations for displaying our RR tests.

- Maps

For moorings we compute these six metrics considering every colocation point during a given laps of time, at each mooring location.

Ship transects are divided into bands of three degrees of latitude. We compute statistics for each band, considering every colocation point during the studied time laps. In order to highlight the evolution between CCI versions, we plot CCI version 3.2 and differences between CCI+SSS v3.2 and version 2.3.

Maps are used to see in which areas strong differences between CCI product and reference are observed.

- Scatterplots

Statistics are computed for CCI products and for a reference satellite product. Results are analysed in the form of a scatterplot comparing results of CCI products vs the reference product.

In every scatterplot we plot the diagonal in blue and a linear fit in red (except for biases, where the red line corresponds to zero).

Scatterplot for AX20 and AX11 are in two separated figures.

For ship tracks, colours correspond to the mean latitude of the latitudinal bands we defined to calculate statistics.

- Time series and Hovmoller diagrams

To examine the temporal evolution of the products we used time series and Hovmoller plots at selected (range of) locations.

For PIRATA we plot the evolution of each mooring, in comparison to the corresponding colocated satellite data (CCI and reference).

For merchant ship data we compute global statistics for each available ship transect and we plot the evolution of the metrics in time.

5 Description of the algorithms & ancillary data tested during the round robin exercise

In this third RR test installment, we tested 2 versions of CCI+SSS products which allow to identify progress. We report here comparisons between CCI+SSS v2.3 and v3.2. We present here comparisons obtained with CCI weekly fields after having removed pixels flagged with SSS, land and ice contamination quality checked. In order to ensure fair comparisons, we only consider satellite pixels when available in both versions. Hence, the number of collocated points are identical for both products.

Only in situ measurements farther from land than a satellite footprint are considered.

Period Name	Reference	Temporal resolution	Spatial resolution	Frequency of data available	Period of study
SMOS Total	CCIL4v3.2 7 flg2 CCIL4v2.3 7 flg2	7 days	50km	1 day	2010/01/12 2021/05/26
SMAP	CCIL4v3.2 7 flg2 CCIL4v2.3 7 flg2	7 days	50km	1 day	2015/05/01 2021/05/26
Aquarius	CCIL4v3.2 7 flg2 CCIL4v2.3 7 flg2	7 days	50km	1 day	2011/06/01 2015/04/01

Table 2: Original products compared to CCI in the RR tests

Notes:

- monthly CCI fields have also been compared with in situ SSS but we do not show the results here as comparing pointwise SSS measurements with monthly means is less meaningful than with weekly means. Nevertheless, we do not observe anomalous behaviour of monthly fields, neither when comparing them with punctual in situ measurements nor when comparing them with *In situ* Analysis System (ISAS) SSS monthly fields.
- The ISAS-CCI comparisons that we briefly show in this document are only intended to check that no anomalous behaviour occurs at global scale while a scientific and deepened assessment is left to the validation team. A complete description of ISAS SSS fields is given in the validation plan and in the PVIR document. We used ISAS salinity at 5m depth, and comparisons between CCI monthly fields and ISAS SSS appear in the discussion section (section 6.3).

Climate Change Initiative+ (CCI+)
Phase 1

**Product Validation and
Algorithm Selection Report**

Ref.: **ESA-CCI-PRGM-EOPS-SW-17-0032**
Date: 14/09/2021
Version : v3.1
Page: 25 of 41

6 Algorithm/Product evaluation

6.1 Results of the round robin exercise

In order to highlight the differences between the two most recent CCI products, we report the summary of the comparison metrics in the tables below. We coloured the table cells to indicate in green improved statistics with respect to the previously released product and in red degraded statistics. When products have equivalent statistics, colours are set to white. See Table 3 for a more precise explanation of the colouring method.

Products	Case 1	Case 2	Case 3
CCIL4**v3.2** 7 flg2			
CCIL4**v2.3** 7 flg2			
	V3.2 better than v2.3	**V3.2 worse than v2.3**	**V3.2 equivalent to v2.3**

Table 3: Meaning of colours in result tables.

To define the colours, we look at the level of significance of the differences between the metrics computed with both products. Along the rows corresponding to the latest v3.2 product, **green means that v3.2 is significantly better than v2.3, white means that v3.2 is equivalent to v2.3, red means that v3.2 is significantly worse than v2.3.** We also include summarized results from the previous RR exercise comparing CCI L4 v2.3 and v1.8. Metrics are defined as follows :

- *Bias:*

The mean difference between CCI SSS and in situ SSS is called 'bias' in the statistics reported below. However, this difference includes mean error coming from the in situ measurements, from the different space and time variability sampled by both data sets (representativity error) and from CCI SSS. In order to estimate the significance of the 'bias' in term of mean error coming from the CCI SSS, we estimate the uncertainty coming from the in situ measurements and from the representativity of each set of measurements.

We first consider a systematic observational uncertainty on each ship transect equal to the one specified in ship files, $E_{ship\_transect}$. We also consider the sampling uncertainty coming from the natural variability of SSS within a satellite pixel and within one week, E_{nat}; we derive an order of magnitude of E_{nat} from the standard deviation of in situ SSS within 50 km (Figure 8). The resulting uncertainty on the ship-satellite SSS coming from ship SSS observational error and sampling error is roughly estimated as:

$$E_{bias} = \sqrt{\frac{\overline{E_{ship\_transect}}^2}{N_{transect}} + \frac{\overline{E_{nat}}^2}{N_{ind}}}$$

With $E_{ship\_transect}$ the systematic observational uncertainty over each transect,
$N_{transect}$ the number of transects
E_{nat} the sampling uncertainty related to SSS natural variability
N_{ind} the number of independent colocation points.
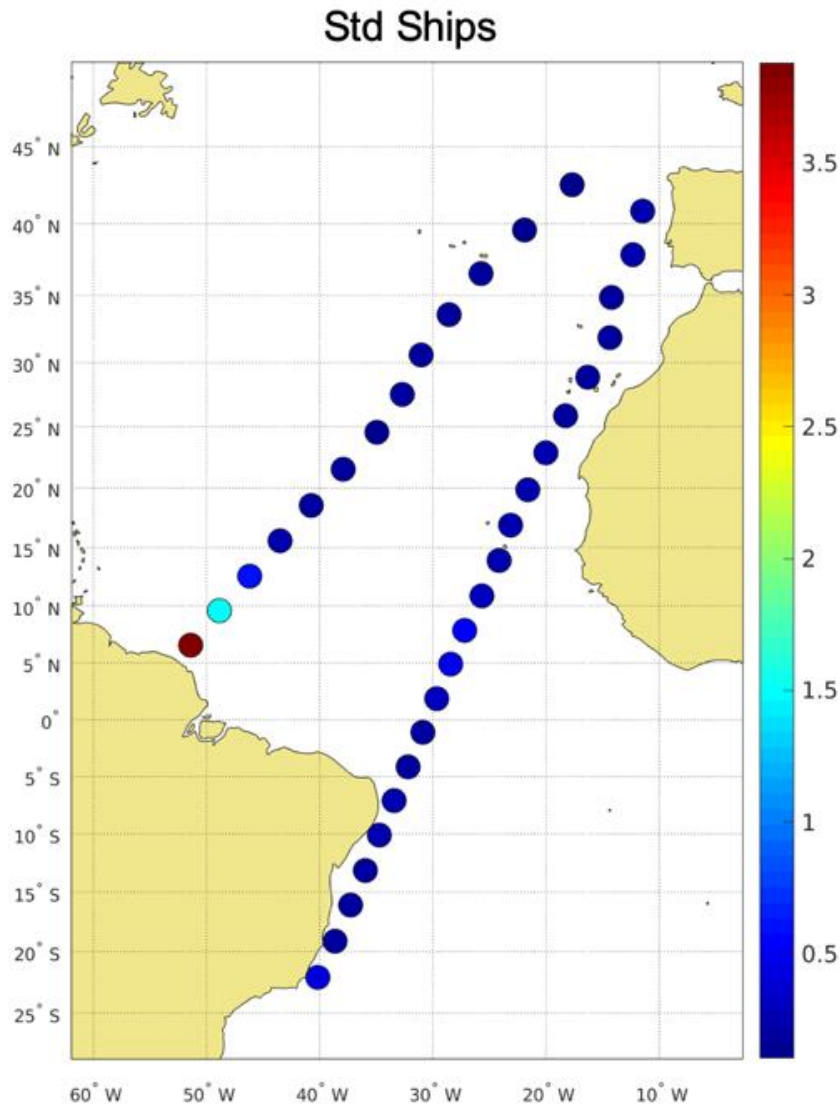
Figure 8 Average std of in-situ measurements over 50km along ship tracks.

Far from coast, according to Figure 8, the sampling uncertainty due to natural variability is smaller than 0.1 so that $E_{ship\_transect}$ dominates the total uncertainty. It is on the order of 0.1 pss for each transect; considering the number of averaged transects $N_{transect}$, we use as threshold for significant bias:

$$B_{thresh} = \sqrt{\frac{0.1^2}{N_{transect}}}$$

It varies in the range 0.009pss for the total period to 0.016pss for the SMAP period.

On another hand, at less than 100 km from coast, B_{thresh} is dominated by the sampling uncertainty. Considering Figure 8, a natural variability of 0.5 pss seems a reasonable order of magnitude for B_{thresh}.

For PIRATA moorings, we consider an uncertainty of the order of 0.1 pss, so $B_{thresh} = \sqrt{\frac{0.1^2}{N_{moorings}}}$, with $N_{moorings}$ the number of moorings.

- ***Std diff and robust std diff:***

We test the significance of the differences between *std diff* of CCI and original products by using a Fisher-Snedecor test, with a threshold of 5% of significance.

- ***r²***

We test the significance between correlation coefficients of CCI (r_{CCI}) and original products (r_{orig}). If r_{orig} is near 0, we can use a t-test; if not, we have to pass the correlation coefficients through Fisher transformation, and we use the test described in Biometry (**RD03**) page 585. We use a threshold of 5% of significance.

We define 'better' , 'worse' and 'equivalent' statistics using the following criteria:
- *Bias*
  - ➢ Product A better than product B if $|bias_A| \leq |bias_B| - B_{thresh}$
  - ➢ Product A equivalent to product B if $|bias_A - bias_B| < B_{thresh}$
  - ➢ Product A worse than product B if $|bias_A| \geq |bias_B| + B_{thresh}$
- *Std diff and robust std diff*
  - ➢ Products A and B are equivalent if the difference of *std diff* is not significant
  - ➢ Product A better than product B if the difference of *std diff* is significant, and $std\_diff_A \leq std\_diff_B$
  - ➢ Product A worse than product B if the difference of *std diff* is significant, and $std\_diff_A \geq std\_diff_B$
- *r²*
  - ➢ Products A and B are equivalent if the difference of r^2 is not significant
  - ➢ Product A better than product B if the difference of r^2 is significant, and $r2_A \geq r2_B$
  - ➢ Product A worse than product B if the difference of r^2 is significant, and $r2_A \leq r2_B$

Climate Change Initiative+ (CCI+)
Phase 1

**Product Validation and
Algorithm Selection Report**

Ref.: **ESA-CCI-PRGM-EOPS-SW-17-0032**

Date: 14/09/2021

Version : v3.1

Page: 28 of 41

Tables

The tables below summarizes the results and some figures are added. The complete set of figures is shown in Section 8.

Period	flag	Product	N	σ	σcr	σ^*	σcr^*	bias	r^2
SMOS-Total	PIRATA	CCIL4v3.2 7 flg2	5365	0.20	1.10	0.15	1.03	0.00	0.94
		CCIL4v2.3 7 flg2	5365	0.21	0.88	0.16	0.83	0.01	0.93
		CCIL4v2.3 7	3772	0.22		0.16		0.01	0.93
		CCIL4v1.8 7	3771	0.20		0.15		0.02	0.94
	High Latitude Ships	CCIL4v3.2 7 flg2	2752298	0.32	1.78	0.24	1.66	0.03	0.84
		CCIL4v2.3 7 flg2	2752298	0.34	1.35	0.25	1.22	0.05	0.82
		CCIL4v2.3 7	3412	0.9		0.44		-0.05	0.06
		CCIL4v1.8 7	3376	0.9		0.44		-0.04	0.06
	Total Atlantic Ships	CCIL4v3.2 7 flg2	71910	0.21	1.26	0.15	1.12	-0.03	0.91
		CCIL4v2.3 7 flg2	71910	0.23	1.07	0.16	0.94	-0.02	0.91
		CCIL4v2.3 7	71763	0.26		0.16		-0.02	0.89
		CCIL4v1.8 7	74837	0.39		0.16		-0.01	0.87
	Atlantic Ships dist coast <400km Lat<45°	CCIL4v3.2 7 flg2	20202	0.28	1.46	0.17	1.27	-0.03	0.83
		CCIL4v2.3 7 flg2	20202	0.29	1.28	0.19	1.13	-0.03	0.81
		CCIL4v2.3 7	20197	0.38		0.19		-0.02	0.82
		CCIL4v1.8 7	22216	0.65		0.22		0.01	0.83
	Atlantic Ships dist coast >400km Lat<45°	CCIL4v3.2 7 flg2	51708	0.19	1.18	0.14	1.07	-0.03	0.94
		CCIL4v2.3 7 flg2	51708	0.19	0.98	0.15	0.88	-0.02	0.93
		CCIL4v2.3 7	51566	0.20		0.15		-0.02	0.93
		CCIL4v1.8 7	52621	0.21		0.14		-0.01	0.93

Table 4: Results of RR tests over total period (NB: except for High Latitude, the number of points N corresponds to the number of points over the EASE grid that is oversampled at 25km resolution; the number of independent pixels is roughly ½ this number of points. σ: std of difference; σcr: std diff for reduced centered variables; σ∗ and σcr∗: robust coefficients (see text); r2: coefficient of determination. Note: for information, we also included previous 2020 results obtained with the CCIL4v2.3 weekly product with respect to the CCIL4v1.8 weekly product. The flag and data sampling used for those comparisons were different than those used for this PVASR installment.

Period	Flag	Product	N	σ	σ_{cr}	σ^*	σcr^*	bias	r^2
SMAP	PIRATA	CCIL4v3.2 7 flg2	2604	0.19	1.12	0.15	1.03	0.01	0.95
		CCIL4v2.3 7 flg2	2604	0.21	0.87	0.16	0.83	0.01	0.94
		CCIL4v2.3 7	965	0.22		0.16		0.03	0.93
		CCIL4v1.8 7	965	0.21		0.15		0.03	0.94
	High Latitude Ships	CCIL4v3.2 7 flg2	1343914	0.28	1.62	0.20	1.46	0.13	0.87
		CCIL4v2.3 7 flg2	1343914	0.29	1.13	0.21	1.01	0.15	0.87
		CCIL4v2.3 7	1356	0.95		0.43		0.1	0.04
		CCIL4v1.8 7	1320	0.96		0.44		0.1	0.03
	Total Atlantic Ships	CCIL4v3.2 7 flg2	28750	0.22	1.31	0.15	1.17	-0.01	0.92
		CCIL4v2.3 7 flg2	28750	0.22	1.04	0.15	0.91	0.00	0.91
		CCIL4v2.3 7	28772	0.26		0.15		0.00	0.91
		CCIL4v1.8 7	29597	0.43		0.16		0.03	0.88
	Atlantic Ships Dist coast <400km Lat<45°	CCIL4v3.2 7 flg2	8771	0.25	1.32	0.16	1.24	0.00	0.86
		CCIL4v2.3 7 flg2	8771	0.25	1.08	0.16	0.99	0.02	0.86
		CCIL4v2.3 7	8760	0.35		0.16		0.03	0.90
		CCIL4v1.8 7	9502	0.69		0.20		0.07	0.86
	Atlantic Ships Dist coast>400km Lat<45°	CCIL4v3.2 7 flg2	19979	0.20	1.30	0.14	1.13	-0.01	0.93
		CCIL4v2.3 7 flg2	19979	0.21	1.02	0.15	0.88	-0.01	0.93
		CCIL4v2.3 7	20012	0.22		0.15		0.00	0.92
		CCIL4v1.8 7	20095	0.23		0.15		0.01	0.92

Table 5: : RR results for the SMAP period.

Climate Change Initiative+ (CCI+)
Phase 1

Product Validation and
Algorithm Selection Report

Ref.: **ESA-CCI-PRGM-EOPS-SW-17-0032**
Date: 14/09/2021
Version : v3.1
Page: 30 of 41

Period	flag	Product	N	σ	σ_{cr}	σ^*	σcr^*	bias	r^2
Aquarius	PIRATA	CCIL4v3.2 7 flg2	1965	0.20	1.12	0.14	1.03	0.01	0.94
		CCIL4v2.3 7 flg2	1965	0.21	0.91	0.14	0.82	0.03	0.94
		CCIL4v2.3 7	1972	0.21		0.14		0.03	0.94
		CCIL4v1.8 7	1972	0.20		0.14		0.03	0.94
	High Latitude Ships	CCIL4v3.2 7 flg2	1194521	0.30	1.71	0.23	1.59	-0.06	0.87
		CCIL4v2.3 7 flg2	1194521	0.33	1.39	0.24	1.23	-0.02	0.83
		CCIL4v2.3 7	718	1.02		0.14		-0.42	0.14
		CCIL4v1.8 7	1479	0.90		0.47		-0.10	0.06
	Total Atlantic Ships	CCIL4v3.2 7 flg2	31394	0.20	1.17	0.14	1.10	-0.04	0.93
		CCIL4v2.3 7 flg2	31394	0.22	1.05	0.15	0.96	-0.02	0.91
		CCIL4v2.3 7	27779	0.27		0.15		-0.02	0.88
		CCIL4v1.8 7	29037	0.38		0.16		-0.02	0.87
	Atlantic Ships Dist coast <400km Lat<45°	CCIL4v3.2 7 flg2	8206	0.26	1.32	0.17	1.23	-0.02	0.85
		CCIL4v2.3 7 flg2	8206	0.29	1.25	0.18	1.14	-0.01	0.81
		CCIL4v2.3 7	7266	0.42		0.18		0.00	0.76
		CCIL4v1.8 7	8132	0.63		0.23		0.01	0.82
	Atlantic Ships Dist coast >400km Lat<45°	CCIL4v3.2 7 flg2	23188	0.17	1.10	0.13	1.05	-0.04	0.95
		CCIL4v2.3 7 flg2	23188	0.18	0.97	0.14	0.90	-0.03	0.94
		CCIL4v2.3 7	20513	0.18		0.14		-0.03	0.94
		CCIL4v1.8 7	20905	0.20		0.14		-0.03	0.94

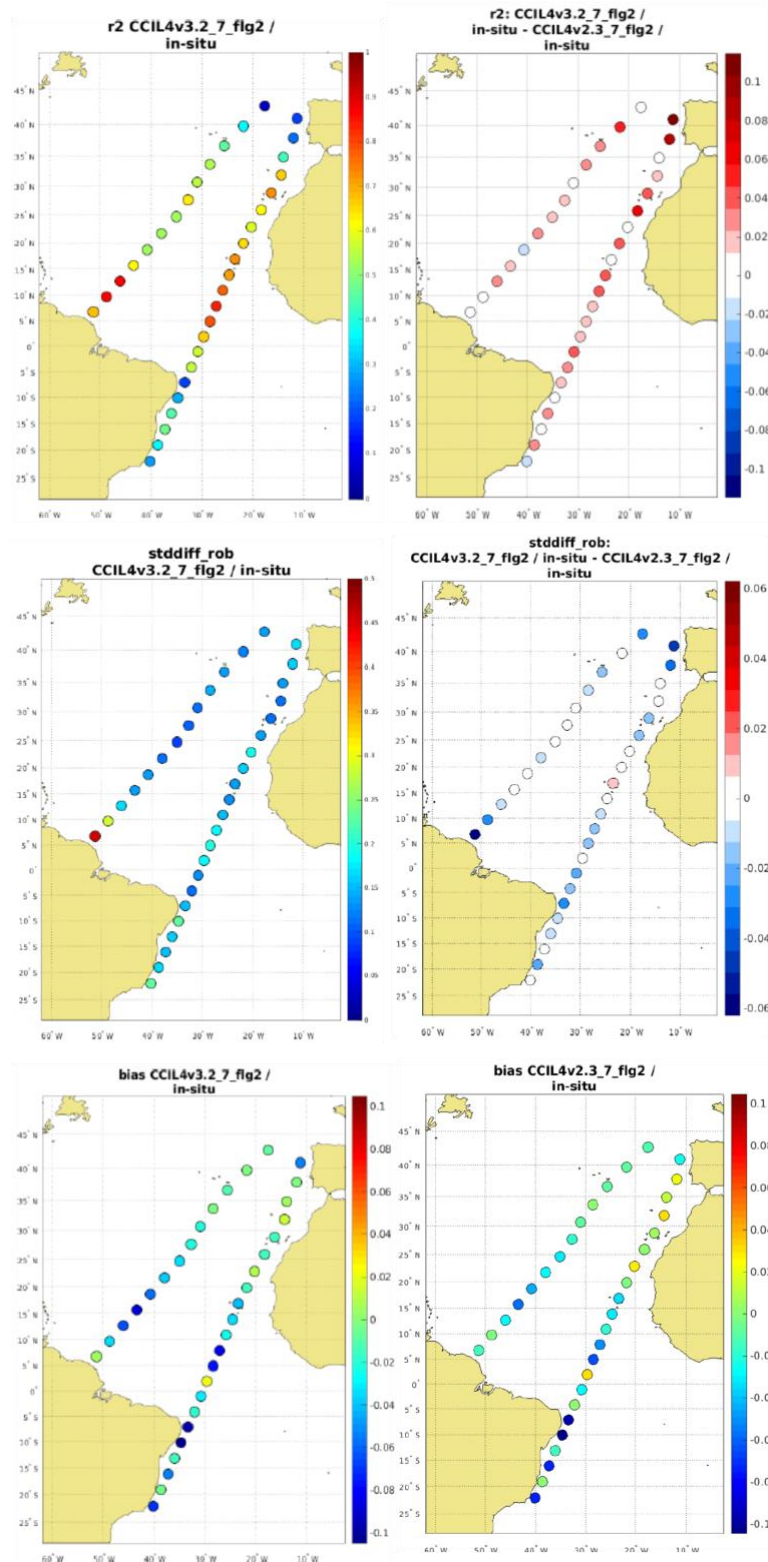*Table 6: RR results for the Aquarius period.*

<u>Maps</u>

Figure 9: Results of the comparisons with ship tracks in Atlantic during global period. Column 1 : CCI v3.2 metric; Column 2 : CCI V3.2metric minus CCI v2.3 metric, except for Line 3: CCI v2.3 bias. Line 1 : r2; Line 2 : robust std diff; Line 3 : bias.
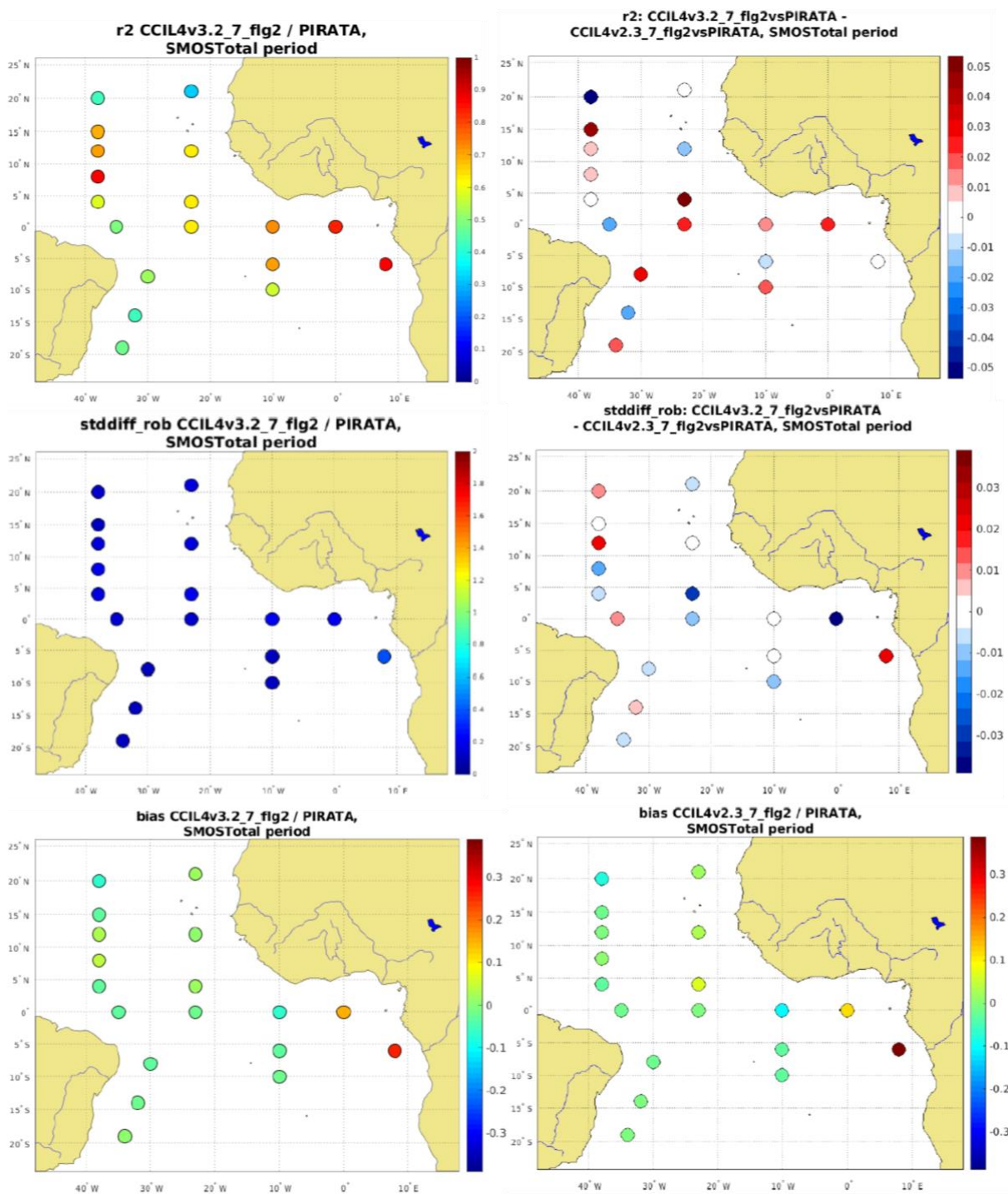
Figure 10: Comparisons with PIRATA moorings, global period. Column 1: CCI v3.2 metric; Column 2: CCI v3.2 metric minus CCI v2.3 metric, except for Line 3: CCI v2.3 bias. Line 1 : r2; Line 2: robust std diff; Line 3: bias.
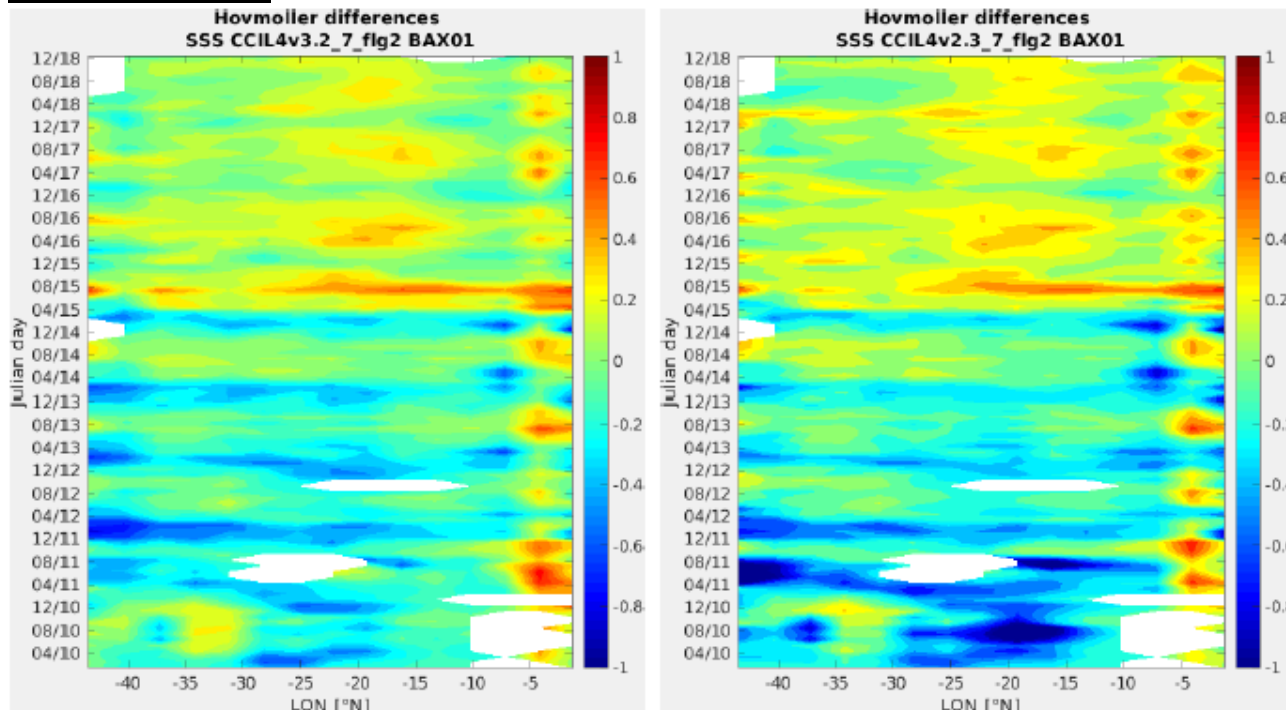
Hovmoller diagrams

Figure 11: Hovmoller diagrams of SSS difference in high northern latitudes between CCI L4 and BAX01 ship data. Left: v3.2; Right: v2.3. (See also Figure 5).
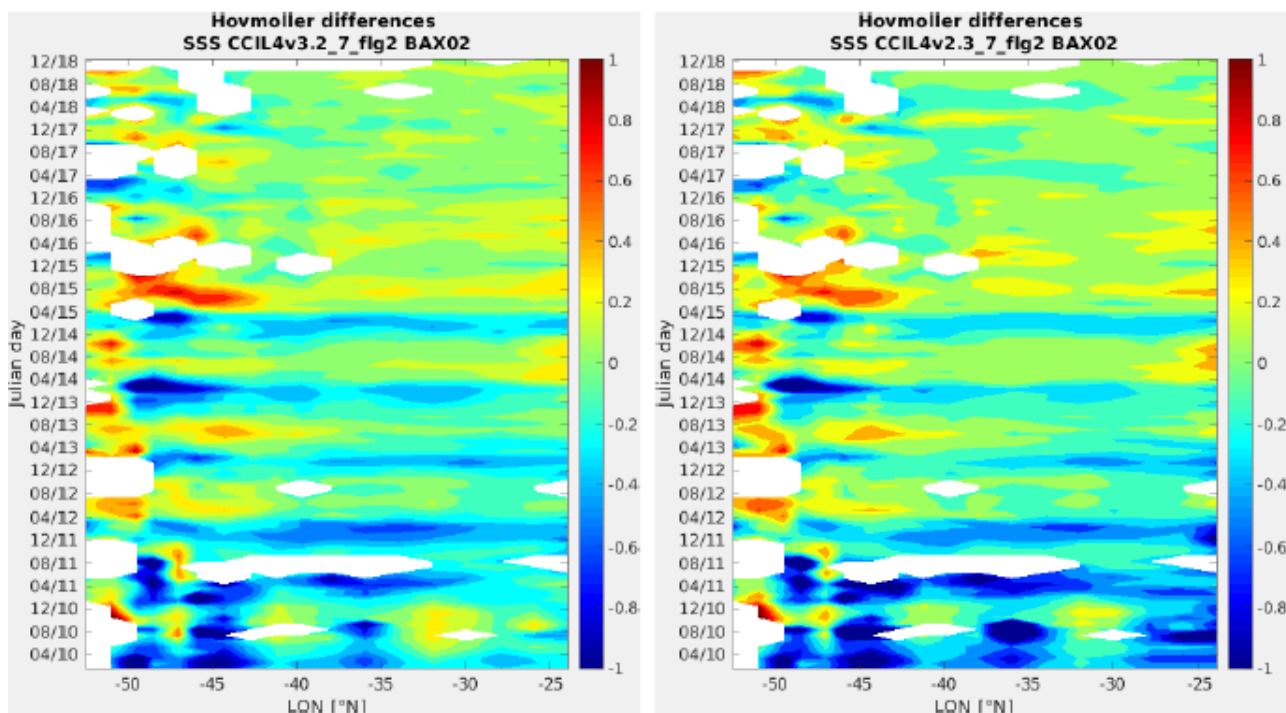

Figure 12: Hovmoller diagrams of SSS difference in high northern latitudes between CCI L4 and BAX02 ship data. Left: v3.2; Right: v2.3. (See also Figure 5).
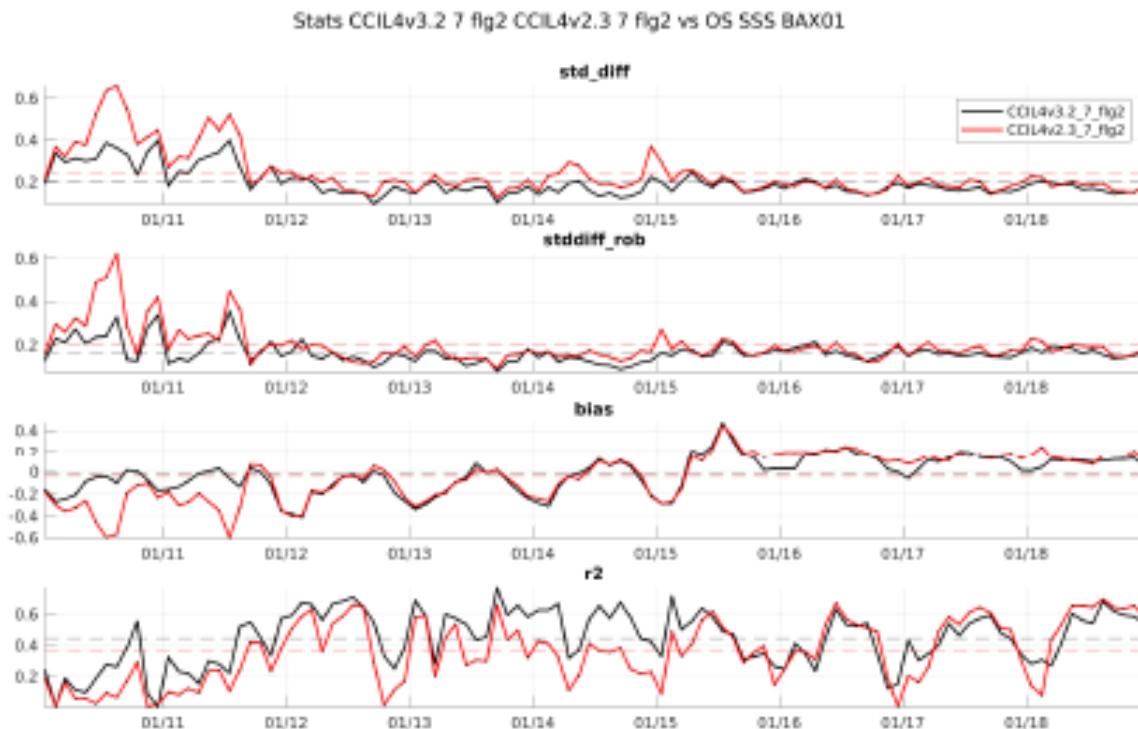
Time series

Figure 13: Temporal variations of the various metrics with respect to the BAX01 ship data. (Black: v3.2; Red: v2.3).
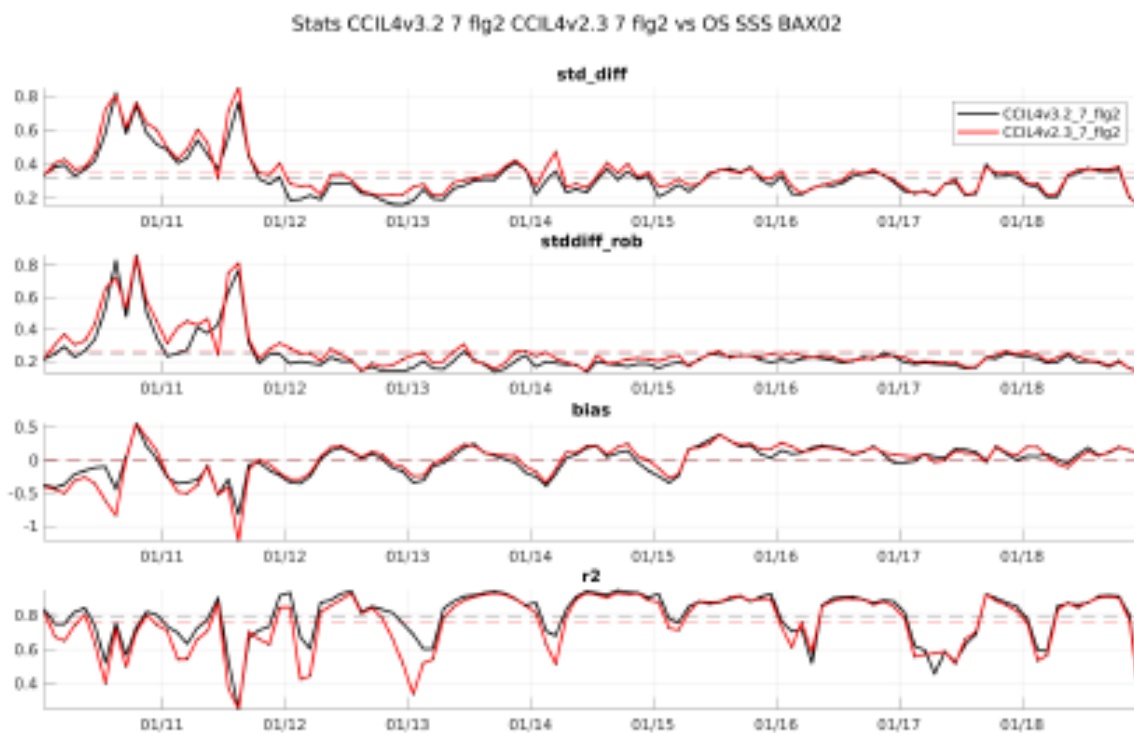

Figure 14: Temporal variations of the various metrics with respect to the BAX02 ship data. (Black: v3.2; Red: v2.3).

Table 4 to Table 6 summarize the results of the RR tests.

Overall, the CCI L4 version 3.2 yields significantly better or equivalent results with respect to the previous CCI L4 v2.3 product.

As a reminder, a general improvement was also found in the former PVASR 2020 where the CCI L4 versions 2 and 1 and the individual products (SMOS, SMAP and Aquarius L3) were all evaluated. For total Atlantic Ocean, the CCI products were in better agreement with the in-situ data than all the original products taken separately or simply averaged. For information, we also included former RR 2020 results obtained with the CCI v1.8 product in the tables. The main salient result was that v1.8 SSS was too smoothed, causing degraded metrics in strong variability regions like along the Atlantic ship transects south of 45°N within 400km from coast. This exaggerated smoothing also yielded v1.8 metrics in non variable regions that could sometimes be artificially enhanced.

From a basic standpoint, even if improvements are not always significant they are nethertheless systematic from v2.3 to v3.2: r^2 is increased, stddiff is decreased. σ_{cr} is also systematically larger than 1 with v3.2 ; it was not the case with v2.3, indicating that the estimated SSS uncertainties used in v2.3 were too large (σ_{cr} higher than 1 is expected, given that we neglect the representativity uncertainty).

Differences between CCI versions such as v3.2 and v2.3 are generally much smaller than when we compare different products (as seen in the previous PVASR). In Figure 9 and Figure 10Figure 9, to highlight the differences between CCI versions we plot differences between metrics (*std diff* and r^2) obtained with the two versions (right panels) instead of metrics of both products themselves. The metrics obtained with the latest product v3.2 are also displayed on the left panels.

Over the ship tracks in the Atlantic (see Figure 9), and over the whole period, *std diff rob* decreases by 0.01 to up to 0.06 off the Portugal coast and off the Orenoque mouth. r^2 rises almost everywhere by 0.01 to up to 0.1 off the Portugal coast.

Statistics computed with PIRATA moorings rarely show strictly significant improvement, although all coefficients show slightly better performance. Correlation is significantly larger (from 0.94 to 0.95) with the new v3.2 during the total and SMAP periods.

When comparing to ship tracks in high northern latitudes (Table 4), results are significantly improved for versions 3.2 with respect to version 2.3, but statistics in this cold area remain worse than at lower latitudes. The time variations of the differences with the ship data and of the various metrics clearly show improvement of v3.2 with respect to v2.3 (Figure 11 to Figure 14).

The results during the Aquarius mission show similar improvement as for the total period, except for slight, albeit significant, degradation of the bias with respect to ship data.

SMOS and SMAP during their overlapping period both contribute to improvement with respect to PIRATA.

6.2 Products evaluation summary

Comparisons with ship SSS indicate that v3.2 SSS is in better agreement with ship SSS than v2.3 almost everywhere. The main improvements are observed:

➢ In high northern latitudes. Statistics with respect to ship data are systematically better for version 3.2 .

➢ In the land vicinity. This is especially true, near the Iberian Peninsula, *robust std diff* (*r2*) is smaller (larger) for v3.2 than for v2.3, and near Orenoque plume, versions 3.2 and 2.3 get better *std diff rob*.

➢ Where the SSS uncertainties, too large in v2.3, become consistent with the differences observed between L4 CCI v3.2 and in situ measurements.

The improvements are not so obvious in comparisons with PIRATA moorings, as they depend on the mooring location. This is possibly because the spatial representativity is better taken into account with ship data than with punctual moorings, or that the improvement is less clear in some PIRATA moorings regions.

6.3 Open issues and discussion

The improved comparison with in situ measurements yielded by CCI L4 v3.2 is attributable to major changes in SMOS v3.2:

- Better RFI filtering; this leads to reduced biases especially in the high northern latitudes at the beginning of the period.

- Improved characterization of the uncertainties. In previous versions, the uncertainties on SSS from individual sensors were overestimated and uncorrectly propagated into the L4 analysis, leading to overestimated L4 SSS uncertainties, particularly in strong variability areas. Example of CCI L4 SSS uncertainties is illustrated on Figure 15.
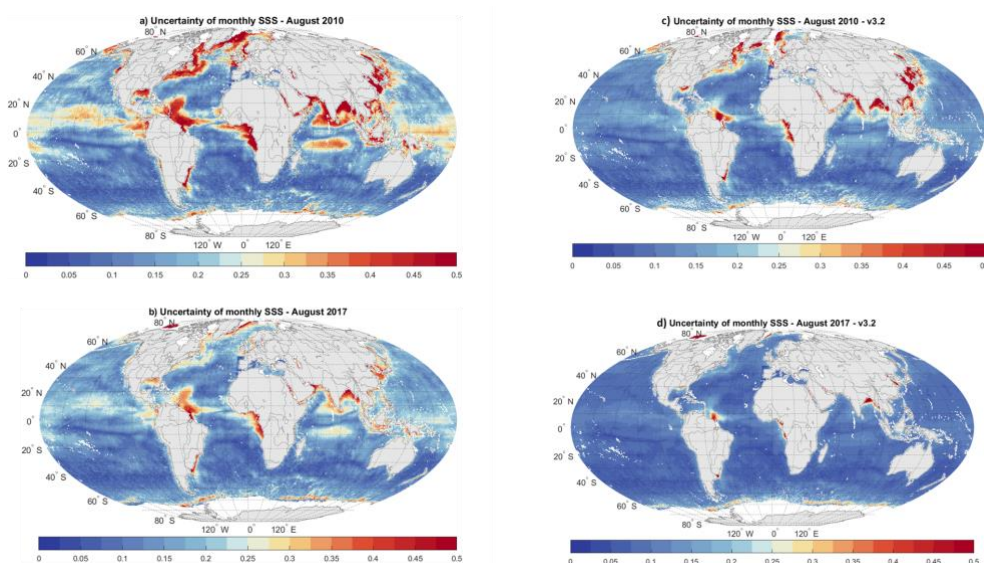

Figure 15: CCI L4 SSS uncertainties a) and b) with version 2.3; c) and d) with version 3.2. a), c) SMOS period, August 2010; b), d) SMOS+SMAP period, August 2017.

In addition to the validation performed with in situ measurements in the northern Atlantic Ocean and in the tropics, we also did some comparisons using ISAS SSS over the global ocean, which allow to further interpret the ships and moorings SSS comparisons. Beyond 1000km from coast (Figure 16), we observe a bias decrease and an improved SSS stability especially in the southern hemisphere after 2015. In particular, the SSS negative bias after 2017 diseappears. This is mainly attributable to changes in the SMOS L2 processing: the OTT in the southeastern Pacific is now derived using ISAS SSS instead of a climatological SSS, which improves stability. This likely is also an effect of the use of the same model version (ERA5) auxiliary parameters during the whole period instead of the ECMWF forecast model which versions evolve regularly. However, significant biases remain before 2015, mainly due to Aquarius seasonal latitudinal biases (Kao et al. 2018 and see below) that are not adjusted in our processing.
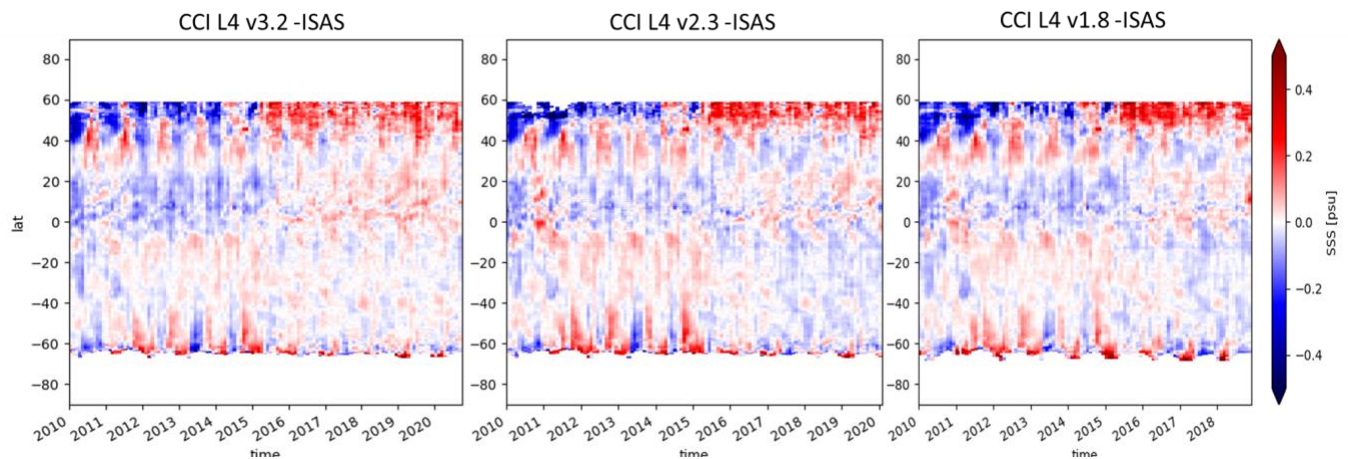

Figure 16: CCI L4 SSS minus ISAS SSS further than 1000km from coast. Left v3.2, middle, v2.3, right, v1.8.

The slight but systematic improvement of the standard deviation of the differences seen in ship comparisons is also observed at global scale in ISAS comparisons both spatially and temporally (Figure 17, Figure 18) at high latitudes, in regions affected by RFIs, in low SST conditions (improved dielectric constant), or for high wind speed (filtering of high winds since v2.3), and in the tropics as expected by the rain rate correction since v3.2. The improved representativity uncertainties in v3.2 also yield reduced std difference in regions with very low SSS variability like in the southern Pacific Ocean. This metric has to be interpreted with caution as the spatio-temporal smoothing of ISAS over radii of ~300km and ~1month blurs the small scale variability, especially in highly dynamic areas, so that it should not be trusted in the vicinity of land (Figure 17 has been done only with pixels further than 1000km from land). From version 1 to 3, the outliers filtering has been refined entailing reduced biases, in particular close to the ice edge in the high latitudes of the southern hemisphere, and especially from v1.8 and later versions (Figure 16).
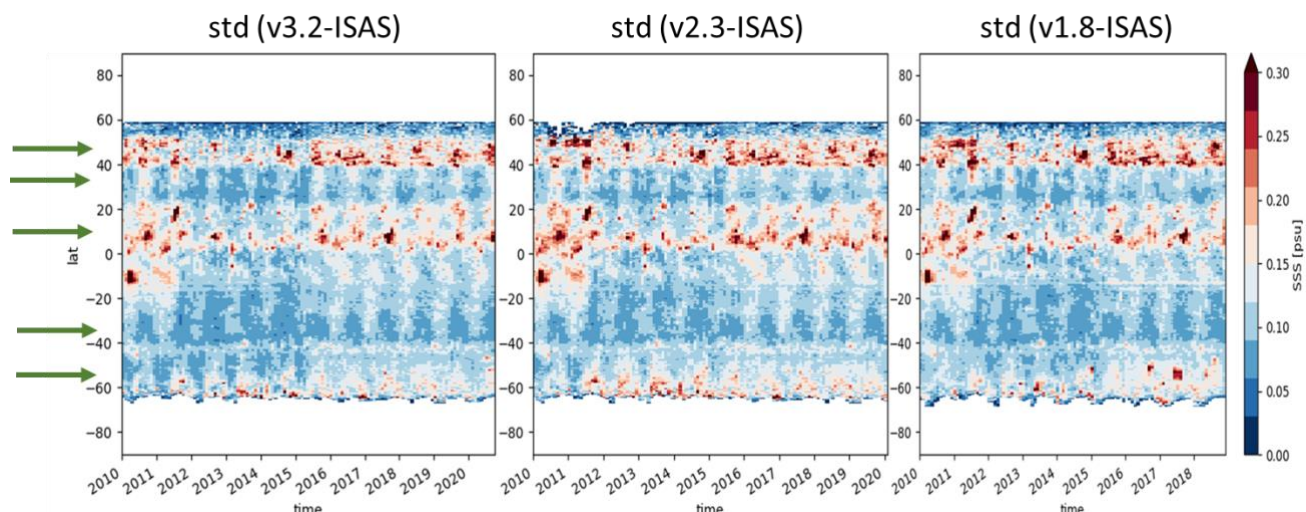
Figure 17: Hovmoller diagram of the standard deviation of the differences between monthly CCI L4 SSS and ISAS SSS for v3.2 (left), 2.3 (middle) and 1.8 (right), further than 1000km from land. Green arrows indicate latitudinal bands with noticeable improvements. (color scale is from 0 to 0.30pss). This metric is indicative of the stability of the differences across all longitudes at each time step.
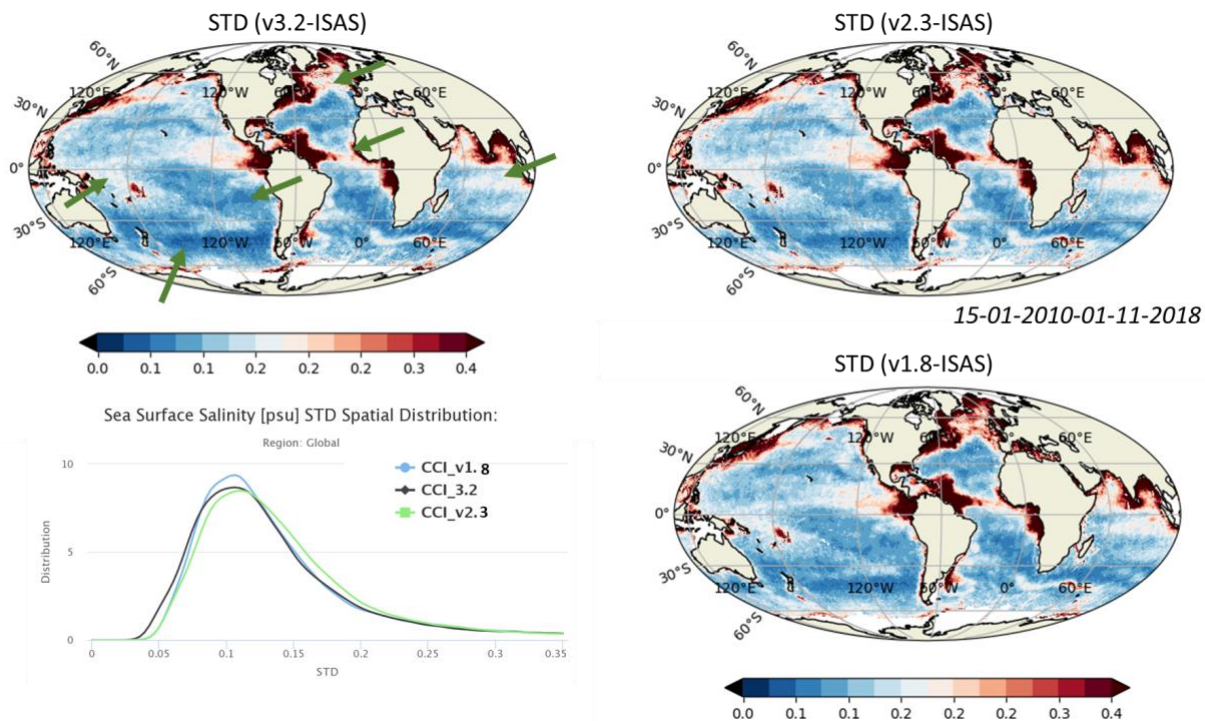

Figure 18: Standard deviation of the differences between monthly CCI L4 SSS and ISAS SSS for v3.2 (top left), 2.3 (top right) and 1.8 (bottom right). Histogramm of the standard deviations differences (bottom left). Green arrows indicate regions with noticeable improvements. (color scale is from 0 to 0.35pss). This metric is indicative of the temporal stability of the differences at each grid point.

Besides these improvements, some flaws remain. In particular, we notice some seasonal variation of the difference CCI+SSS – in situ SSS in the high latitudes, particularly in the north and during the Aquarius period (Figure 16). This flaw slightly decreased from v1.7 to v3.2 but remains a major issue.

In order to analyse the causes responsible for this flaw, we show on Figure 19 the latitudinal-temporal profiles of the differences for CCI L4 and each instrument (notice that only SMOS SSS are corrected for latitudinal seasonal biases in the CCI processing).
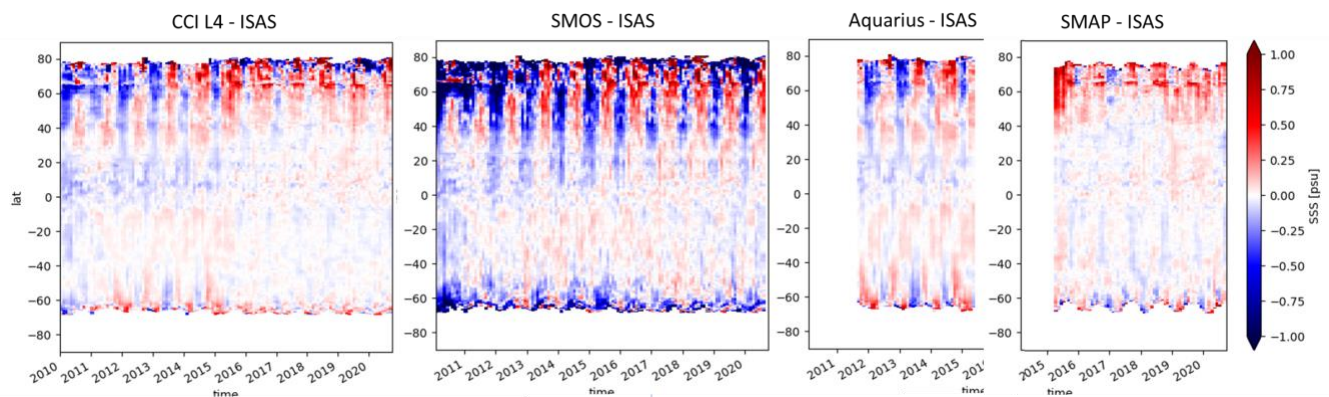

Figure 19: Satellite SSS minus ISAS SSS. From left to right: CCI L4 v3.2, SMOS SSS after CCI v3 debiasing, Aquarius SSS after CCI v3 debiasing (the seasonal variation of the biases are of similar magnitude as in Kao et al. (2018)), SMAP SSS after CCI v3 debiasing.

- For SMOS SSS reprocessed in the frame of CCI v3:
  - ➤ In the northern hemisphere, north of ~10°N, SMOS SSS are systematically biased towards low values in boreal winter, and these biases also vary interannually so that our seasonal latitudinal correction cannot fully correct for them.
  - ➤ Above ~45°N, very few pixels are found beyond 1000km from coast, hence we have to relax the criteria for estimating the latitudinal seasonal biases to pixels 600km away from coast. Even with this criteria, the bias is estimated with mostly all pixels in the Atlantic ocean, and we observe that the bias estimates are very sensitive to the pixel locations we choose (Pacific or Atlantic Ocean; not shown).

- For Aquarius SSS: As shown by Kao et al (2018), seasonal latitudinal biases that were not corrected for in CCI+SSS phase 1 remain in Aquarius SSS. They are very well phased with SMOS SSS biases and hence with the biases in CCI L4 SSS, both in northern and southern hemispheres. It will be important to correct for them in phase 2.

- For SMAP SSS: Even though SMAP SSS are less biased seasonally and latitudinally, we notice biases at very high northern latitude. Their seasonal variation is in phase with SMOS and Aquarius biases (negative bias in boreal winter). This would deserve more attention for future developments.

7 Conclusion and future work

In conclusion, we found improved comparisons of CCI L4 v3.2 SSS with respect to in situ SSS, and more realistic CCI L4 v3.2 SSS uncertainties. This is the combined result of 1) the upgraded SMOS processing, 2) the ameliorated SMOS, SMAP and Aquarius SSS optimal interpolation, 3) the improved uncertainty propagation and estimates of representativity uncertainty.

Nevertheless, some issues remain to be tackled during CCI+SSS phase 2:

- In the high northern latitudes, seasonal latitudinal biases and land-sea (and/or ice-sea) contamination jointly affect all SMOS ocean pixels. Hence, the current method which separates the determination of the two (or three) types of contamination does not allow to remove all the contaminations. The methodology should be adapted to deal with both contaminations in the future phase.

- The interannual variation of SMOS SSS biases after systematic bias correction has not been fully removed with the use of ISAS for OTT computation and the use of ERA5 auxiliary parameters. The recent release of SMOS L1/L2 v7, now operationally run, is subject to different biases that still include some interannual variability. To which extent the methods used with SMOS v6 could be successfully applied to SMOS v7 should be explored. One issue is whether the full SMOS field of view (FOV) or only a fraction (e.g Alias Free FOV) should be considered without loosing too much SSS precision. The ice contamination impacting SMOS v7 more strongly than v6, especially on ascending orbits in the Southern hemisphere, should also be considered with great care.

- The ERA5 auxiliary parameters have been used instead of ECMWF forecasts without changing the direct models. The consequence of this choice should be further evaluated.

- Aquarius seasonal and latitudinal biases have been found to play a major role in the degradation of biases before 2015. Hence they should be corrected before the data are incorporated in CCI L4.

- A new version of SMAP (RSS v5) is planned to be released in Fall 2021. Together with SSS, it will provide SSS uncertainties which could be interesting to include in the CCI processing.

The metrics chosen to discriminate between various CCI versions could also be improved to ease and complement diagnostics. For example:

  - Classify the metrics depending on the natural SSS variability. For instance, this would allow to more clearly detect some exaggerated smoothing , spuriously yielding stddiff diminution (increase) in low (high) variability SSS regions.

  - In addition to validating SSS within pixels that are common to two versions, separately validate SSS within pixels present in only one version (e.g coastal pixels in v3.2).

  - Error budget closure: we could investigate how the comparison statistics between in situ and CCI SSS fare with the ones expected from the uncertainties on radiometric measurements, on radiometric modelling, on in situ measurements, and on representativity. For instance, if the rms of difference is equal or below the quadratic sum of the uncertainties, it would mean that the CCI SSS cannot be further improved without incorporating other external information.

8 Annexe: Round Robin Tests Report PVASR wkly flag 2

See pdf document below

Adobe Acrobat
Document