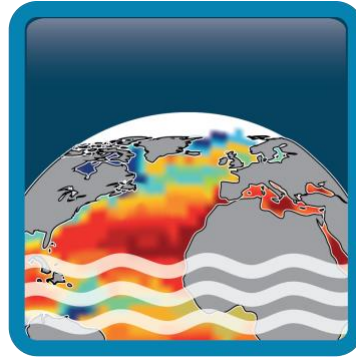# Climate Change Initiative+ (CCI+) Phase 1

# Sea Surface Salinity

**Product Validation and Intercomparison Report** (PVIR)
**Customer:** ESA
Ref.: ESA-CCI-PRGM-EOPS-SW-17-0032

Version: **v2.0**
Ref. internal: **AO/1-9041/17/I-NB_v1r2**
Revision Date: **11/11/2020**
Filename: **SSS_cci-D4.1-PVIR-v2_i1.docx**
Deliverable code: **D4.1**

# Signatures

|  | Name | Signature | Date |
|---|---|---|---|
| **Lead Author** | Adrien Martin (NOC) |  | 11-11-2020 |
| **Approved By** | Jacqueline Boutin (Science Leader) LOCEAN | *[signature]* | 11-11-2020 |
|  | Nicolas Reul (Science Leader) IFREMER |  | 11-11-2020 |
|  | Rafael Catany (Project Manager) ARGANS | *[signature]* | 11-11-2020 |
| **Accepted By** | Craig Donlon (Technical Officer) ESA |  |  |

| Diffusion List |
|---|
| Sea Surface Salinity Team Members |
| ESA (Craig Donlon, Paolo Cipollini) |

# Amendment Record Sheet

| Date / Issue | Description | Section / Page |
|---|---|---|
| | **Document Change Record** | |
| Date / Issue | Description | Section / Page |
| 15-Oct-2019/v1r1 | Template | Whole document |
| 24-Dec-2019/v1r2 | Final draft distributed to the team involved in validation | Whole document |
| 10-Jan-2020/v1.0 | Final Distribution to ESA | |
| 7-Apr-2020/v1.1 | The document has been updated using ISAS at 5 m depth, using only data with PCTVAR<95%. Comparison between ISAS and CCI has been proceeded on CCI grid not to affect CCI uncertainty estimate. | Section 4 / 27-54 |
| | Old section 5.4.1 about comparison between CCI and CMEMS in river plume has been removed from the PVIR and included in the CAR. | Section 5.4.1 / p 71 |
| | section 5.5 about case study 5: comparing salinity variability between observations and models has been added to the document (originaly was in the CAR) | Section 5.5 / p.74 |
| | Four figures from Pi-MEP match-up mooring report representing SSS power spectra have been included in Annex C. | Annex C / p.115 |
| | Update of the executive summary | Section 2 / p. 19-21 |
| | A sentence to summarize what specific improvements of SMOS in the Bay of Bengual has been added. | Section 5.3 / p. 64 |
| | The document has been reviewed for spelling/grammar/typos errors | Whole document |
| 2-Nov-2020/v2.0 | Version 2 final draft distributed to the team | Whole document |
| 11-Nov-2020/v2.0 | Final distribution to ESA | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

## Table of Contents

# List of figures

# List of tables

# 1 Introduction

## 1.1 Purpose and scope

The purpose of this document (D.4 Product Validation and Intercomparison Report, PVIR, document version v2.0) is to describe the results of the validation of the Sea Surface Salinity (SSS) products obtained during the ESA CCI+ SSS project when compared with other data sources. The PVIR is a requirement of the Statement of Work (Task 3 SoW ref. ESA-CCI-PRGM-EOPS-SW-17-0032). The PVIR contains a list of all reference datasets used for validation of each SSS product.

In this report are assessed, the level 4 and level 3 (ascending, descending, combined ascending plus descending) (1) monthly and (2) weekly products. The products are based on a temporal optimal interpolation of SSS data measured by SMOS, Aquarius-SAC and SMAP satellite missions. All products are gridded on an equal area EASE 2 grid with a grid resolution of ~25 km.

## 1.2 Structure of the document

This document is composed of six sections:

Section 1 introduces the purpose and scope of the document. Section 2 provides an executive summary of the results presented. Section 3 presents the data and methods used for the systematic validation presented in Section 4. Supplementary material is provided in Annex A.

## 1.3 Applicable Documents

| PSD | Product Specification Document | SSS_cci-D1.2-PSD-v1r6 |
|---|---|---|
| PUG | Product User Guide | SSS_cci-D4.3-PUG-v1.1 |
| PVP | Product Validation Plan | SSS_cci-D2.5-PVP-v1.1 |
| SoW | CCI+ Statement of Work | SOW |

Table 1 – Applicable documents (as seen in CCI+SSS website, http://cci.esa.int/salinity)

## 1.4 Reference Documents

| ID | Document | Reference |
|---|---|---|
| RD01 | Product Validation Plan | |
| RD02 | Pi-MEP consortium, March 2019; Match-up database Analyses report, CCI-L4-ESA-MERGED-OI-V1.5-MONTHLY Argo Global Ocean: pimep-mdb-report_GO_cci-l4-esa-merged-oi-v1.5-1m_argo_20190315.pdf | |
| RD03 | G. Reverdin, S. Morisset, L. Marié, D. Bourras, G. Sutherland, B. Ward, J. Salvador, J. Font, Y. Cuypers, L.R. Centurioni, V. Hormann, N. Koldziejczyk, J. Boutin, F. D'Ovidio, F. Nencioli, N. Martin, D. Diverres, G. Alory & R. Lumpkin (2015). Surface salinity in the North Atlantic subtropical gyre during the STRASSE/SPURS summer 2012 cruise. Oceanography 28 (1): 114-123 | |
| RD04 | N. Hoareau, A. Turiel, M. Portabella, J. Ballabrera-Poy & J. Vogelzang (2018). Singularity Power Spectra: A Method to Assess Geophysical Consistency of Gridded Products - Application to Sea-Surface Salinity Remote Sensing Maps. IEEE Transactions on Geosciences and Remote Sensing 56, 5525-5536 | |
| | Stammer et al., 2020, How good do we know ocean salinity and its changes?, submitted to Progress in Oceanography | Stammer et al., 2020 |
| | Boutin et al 2016 , 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 | Boutin et al 2016 |
| | Supply, A., J. Boutin, J.-L. Vergely, N. Kolodziejczyk, G. Reverdin, N. Reul, and A. Tarasenko (2020), New insights into SMOS sea surface salinity retrievals in the Arctic Ocean, Remote Sensing of Environment, 249, 112027, https://doi.org/10.1016/j.rse.2020.112027. | Supply et al, 2020a |
| | Supply, A., J. Boutin, G. Reverdin, J.-L. Vergely, and H. Bellenger, 2020: Variability of satellite sea surface salinity under rainfall. In: Satellite Precipitation Measurement, V. Levizzani, C. Kidd., D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, F. J. Turk, Eds., Springer Nature, Cham, Advances in Global Change Research, 69, 1155-1176, https://doi.org/10.1007/978-3-030-35798-6_34. | Supply et al, 2020b |

*Table 2 – Reference documents*

## 1.5   Acronyms

CAR         Climate Assessment Report

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

CDR         Climate Data Record

CMEMS       Copernicus Marine Environmental Monitoring Service

CMIP        Coupled Model Intercomparison Project

CMUG        Climate Modelling User Group

CRDP        Climate Research Data Package

CRG         Climate Research Group

DARD        Data Access Requirements Document

EASE-2      Cylindrical Equal Area Scalable Earth grid 2.0

ECMWF       European Centre for Medium Range Weather Forecasts

ECV         Essential Climate Variable

FRM         Fiducial Reference Measurements

ISAS        In-Situ Analysis System

ISDB        in situ database (of Fiducial Reference Measurements and satellite measurements)

MDB         Match-up DataBase

Pi-MEP      Pilot Mission Exploitation Platform

PMP         Project Management Plan

PSD         Product Specification Document

PUG             Product User Guide

PVIR            Product Validation and Intercomparison Report

PVP             Product Validation Plan

QA4EO           Quality Assurance Framework for Earth Observation

RFI             Radio Frequency Interference

SISS            Satellite and In situ [Working Group]

SMAP            Soil Moisture Active Passive [mission of NASA)

SMOS            Soil Moisture and Ocean Salinity [satellite of ESA]

SoW             Statement of Work

SSS             Sea Surface Salinity

TSG             ThermoSalinoGraph

UCR/CECR        Uncertainty Characterisation Report (formerly known as the Comprehensive Error Characterisation Report)

UNFCCC          United Nations Framework Convention on Climate Change

URD             User Requirements Document

# 2 Executive Summary

## 2.1 Sea Surface Salinity products

The products validated are:

For L4:

```
ESACCI-SEASURFACESALINITY-L4-SSS-MERGED_OI_Monthly_CENTRED_15Day_25km-
xxxxxxxx-fv2.3
ESACCI-SEASURFACESALINITY-L4-SSS-MERGED-OI-Weeky-CENTRED-1Day-25km-xxxxxxx-
fv2.3
```

For L3:

```
ESACCI-SEASURFACESALINITY-L3C-SSS-
SMOSSMAPAQUARIUS_A_Monthly_Centred_15Day_25km-xxxxxxxx-fv2.3
ESACCI-SEASURFACESALINITY-L3C-SSS-
SMOSSMAPAQUARIUS_D_Monthly_Centred_15Day_25km-xxxxxxxx-fv2.3
ESACCI-SEASURFACESALINITY-L3C-SSS-
SMOSSMAPAQUARIUS_Monthly_Centred_15Day_25km-xxxxxxxx-fv2.3
ESACCI-SEASURFACESALINITY-L3C-SSS-SMOSSMAPAQUARIUS_A_Weeky-CENTRED-1Day-
25km-xxxxxxxx-fv2.3
ESACCI-SEASURFACESALINITY-L3C-SSS-SMOSSMAPAQUARIUS_D_Weeky-CENTRED-1Day-
25km-xxxxxxxx-fv2.3
ESACCI-SEASURFACESALINITY-L3C-SSS-SMOSSMAPAQUARIUS_Weeky-CENTRED-1Day-25km-
xxxxxxxx-fv2.3
```

Full description of the dataset can be found in the Product User Guide (PUG). The products follow recommendations of the Product Specification Document (PSD).

## 2.2 Main results

- In situ reference data are Argo floats upper salinity measurement between 0 m and 10 m, which is the dataset providing the most complete spatio-temporal coverage over the globe;
- Argo and CCI product pairwise match-up database (25km, 7.5 days) is regridded on an Equal-Area EASE-2 grid at 175km resolution; bi-weekly;
- Need to take robust estimator (based on the data distribution: median, standard deviation estimated from a ratio to IQR) to be robust to non-normal distribution and fairly representative of the behaviour of more than 50% of the observations;
- No systematic bias against reference data (see summary for PiMEP match-up report in section 2.3 below for more details);
- Global precision against reference gridded data is of 0.15 pss (see details in summary for PiMEP match-up report in section 2.3 below);

- CCI version 2 products show similar performance than v1 but is one year longer and provide access to individual satellite and passes;
- Coherent variability between CCI and in situ data
  - more coherent small-scale high-frequency variability for CCI;
  - coherent annual amplitude signal;
  - larger amplitude in the inter-annual variability for CCI;
- Good agreement between CCI and reference data, including long-term stability, differences within +-0.05 pss for latitudinal band between [40°S-20°N];
- Remaining seasonal oscillation of CCI SSS differences against reference:
  - CCI are fresher/saltier in Winter/Summer than reference;
  - Amplitude is maximum at high latitudes (40°-60°) and can exceed 0.1 pss peak-to-peak;
  - Amplitude is stronger for L3 SMOS;
- CCI SSS is lower than reference data in the beginning of the time series (2010) up to 2012 with an amplitude up to 0.1 pss;
- CCI data in the Arctic and Southern Ocean have not been properly validated as there are limited suitable in situ references;
- Uncertainty provided in CCI product are in good agreement with observations (within +-25%), excepted for Aquarius for which provided uncertainty is too low;

## 2.3 Main results from Pi-MEP match-up reports

- No global bias against Argo except for filtered collocations where:
  - SSS less than 33 pss (CCI saltier by 0.06 pss);
  - Mixed layer depth shallower than 20m (CCI saltier by 0.04 pss);
  - SSS higher than 37 pss (CCI fresher by 0.03 pss);
- Global precision (robust standard deviation; pairwise difference) against Argo of 0.16 pss
  - Decreasing to 0.13 pss for optimal region (>800 km from the coast; area with temporal standard deviation smaller than 0.2 pss);
  - Increasing to 0.25 pss for area closer than 150 km from the coast or with SSS < 33 pss, but these numbers also contain differences due to the different type of sampling by Argo and satellite;
  - Increasing to 0.2 pss for area characterised by one of the following conditions: rain and low wind; mixed layer depth <20m; area with temporal standard deviation >0.2pss; SST < 5°C;
- Comparison with other 29 satellite SSS products against Argo
  - CCI products have the best precision (and no bias) except for Aquarius L4 IPRC v5 products.
  - Same precision for the monthly and weekly products.
- Good agreement between the observed CCI product SSS power spectra and mooring for the two averaging period (weekly and monthly).

## 2.4 Recommendations and caveats to use CCI+SSS dataset

CAVEATS

- Products have not been fully optimised for some issues encountered at very high latitudes (i.e. ice, RFI, biases due to land-sea contamination).
- The criteria for flagging data close to land (including islands) are conservative and likely to be too restrictive in places.
- There is a systematic global underestimation (0.1 pss) of SSS starting at the beginning of the dataset, and gradually disappearing at the end of 2010.
- There is a seasonal varying bias (~0.1, peaking in the middle of the year) at high latitude.

# 3 Validation: Data & Methods

This section describes the Data and Methods used for the main validation results given in section 4.

Following PVP [RD1] recommendations, the reference dataset used for product validation consists of:

- In situ measurements of close-to-surface (<10 m) Argo from Pi-MEP

The reasons for this choice of reference dataset are as follows:

- In the list of acceptable Fiducial Reference Measurements (FRM) referred to in PVP [RD1], the Argo dataset has been selected as it is the only dataset to provide regularly an almost complete coverage of global open water ocean. The temporal distribution from 2010 is also homogeneous [Pi-MEP – RD2].

In the following, Argo dataset is described with its collocation criteria along with the gridded method and the method to estimate uncertainties and representativeness errors. A summary of the spatial representativeness error of in situ measurement, as described in the PVP [RD1], is given here. Finally, quality metrics to assess CCI products are presented.

The weekly products are using monthly fields on which observed variability is added, therefore the focus of this summary report is on monthly fields, unless mentionned otherwise. We report in this document validation against other in situ dataset (TSG, drifters, mamal measurements) which are provided on the Pi-MEP platform. Description of these datasets and full validation of Weekly and Monthly fields are available on the Pi-MEP platform:

- https://pimep.ifremer.fr/diffusion/analysis/
- ftp://ftp.ifremer.fr/ifremer/cersat/pimep/diffusion/analysis/

## 3.1 Dataset description

### 3.1.1 Argo & gridded method

The Argo floats used for validation have been taken from Pi-MEP where quality control checks have been made. The text below is an extract of the detailed description of the Argo dataset and of the collocation (Match-ups Data Base - MDB) with CCI+SSS products.

Argo is a global array of 3,000 free-drifting profiling floats that measures the temperature and salinity of the upper 2000 m of the ocean. This allows continuous monitoring of the

temperature and salinity of the upper ocean, with all data being relayed and made publicly available within hours after collection. The array provides around 100,000 temperature/salinity profiles per year distributed over the global open water oceans at an average of 3-degree spacing. Only Argo salinity and temperature float data with a quality index set to 1 or 2 and data mode set to real time (RT), real time adjusted (RTA) or delayed mode (DM) are considered in Pi-MEP. Argo floats that may have problems with one or more sensors appearing in the grey list maintained at the Coriolis/GDACs are discarded. Furthermore, Pi-MEP provides an additional list of ∼1000 "suspicious" Argo salinity profiles that are also removed before analysis. The upper ocean salinity and temperature values recorded between 0 m and 10 m depth are considered as Argo sea surface salinities (SSS) and sea surface temperatures (SST). These data were collected and made freely available by the international Argo project and the national programs that contribute to it [Argo (2000)].

The Argo MDB is produced from the previously described cleaned Argo dataset. For the monthly CCI+SSS product, the match-up temporal window radius is 7.5 days around the central date of each satellite time step (bi-weekly, monthly averaged), and 12.5 km for the spatial window radius for each grid nodes centre of a 25 km spatial resolution product. If several satellite pixels are found to meet these criteria, the final satellite SSS match-up point is the closest in time from the in situ data measurement date. The final spatial and temporal lags between the in situ and satellite data are stored in the MDB files. A wide range of collocalised auxiliary information are also provided in the MDB.

All the data are freely available as NetCDF files at:

- https://pimep.ifremer.fr/diffusion/data/
- ftp://ftp.ifremer.fr/ifremer/cersat/pimep/diffusion/data/

The Argo/CCI pairwise MDB is regridded on the CCI grid subsampled by a factor 7 both in latitude and longitude. It corresponds to a 175 km Equal Area EASE grid. The same biweekly temporal sampling as CCI monthly product is conserved. The median value for each grid point of all pairwise MDB values is taken.

## 3.2 Uncertainty validation

To validate satellite uncertainty estimates, the approach is to compare the distribution of the difference of satellite SSS minus reference SSS ($\Delta SSS = CCI - ref$). In an ideal scenario, the $\Delta SSS$ standard deviation equals the satellite uncertainty ($\Delta \sigma_{sat}$):

$$\sigma_{\Delta SSS = CCI-ref} = \Delta \sigma_{sat}$$

However, as stated in the PVP [RD1] the geophysical variability of reference SSS data over the time-space scale of remote sensing products depends not only on the particular spatial

resolution and time window defining the remote sensing products, but also on the region at which this variability is estimated (inter-regional variability being quite significant [RD3]). Consequently, the $\Delta SSS$ standard deviation is a combination of both the satellite SSS uncertainty and the uncertainty in the reference SSS ($\Delta\sigma_{ref}$):

$$\sigma_{\Delta SSS} = \sqrt{\Delta\sigma_{sat}^2 + \Delta\sigma_{ref}^2}$$

In the reference uncertainty all the following terms are included:

- $\Delta\sigma_{meas.}$ : Measurement uncertainty (direct instrument error);
- $\Delta\sigma_{space}$ : Spatial representativeness error (difference in spatial sampling of a point measurement versus a surface measurement defined by a grid cell);
- $\Delta\sigma_{time}$ : Time representativeness error;
- $\Delta\sigma_{vertical}$ : Vertical representativeness error (difference in depth of the measurements).

The reference uncertainty corresponds to the following combination:

$$\Delta\sigma_{ref} = \sqrt{\Delta\sigma_{space}^2 + \Delta\sigma_{time}^2 + \Delta\sigma_{vertical}^2 + \Delta\sigma_{meas.}^2}$$

In the following, we assume the measurement uncertainty to be negligible ($\Delta\sigma_{meas,} = 0$). This is true at first order as we consider all poor measurements to have been discarded with the quality control and filtering methods applied by Pi-MEP.

The vertical representativeness error, will be discussed in section 4.3. Although sometimes important, it is neglected for now ($\Delta\sigma_{vertical} = 0$). The time representativeness error, although sometimes important (e.g. river plumes), is not considered for now ($\Delta\sigma_{time} = 0$). Argo measurements have been selected in a +-7.5 days range around the central date of each satellite time step with a 30 days/monthly running mean. The spatial representativeness error is the only remaining reference uncertainty considered in this uncertainty assessment. This error is fully described in the PVP [RD1], a summary is provided below. The spatial power spectra of SSS consistently exhibits a spectral slope of -2.4 ($S(k)=\beta \, k^{-2.4}$) in a range going from a few kilometres to basin scale (~10,000 km) [RD4]. The variance contained between the spatial frequency $k_L$ and $k_l$ (respectively, between the scales l and L, with l<L) is given by the double integral:

$$S^2(k_L, k_l) = \iint\limits_{k_L < k < k_l} d\boldsymbol{k} \; S(\boldsymbol{k}) = B \int_{k_L}^{k_l} kdk \; k^{-2.4}$$

Assuming three spatial scales: $g$ for the ground truth measurements, $r$ for the remote sensing product and $L$ for the basin scale, $g \ll r \ll L$, $\sigma_0 = \sigma(r)$ the standard deviation of SSS contributed by all scales as measured by remote sensing, we obtain the following relationship:

$$\mathrm{D}s^2(g,r) \approx s_0^2 \left(\frac{r}{L}\right)^{0.4} \approx s^2(\mathrm{r}) \left(\frac{r}{L}\right)^{0.4}.$$

Assuming $L = 5000$ km, with $r = 25$ km for the SSS product, the spatial representativeness is estimated as follow:

$$\Delta\sigma_{space} = \sigma_0 * 0.35$$

With $\sigma_0 = \sigma(r)$ the CCI SSS field standard deviation in time for each grid cell.

### 3.3 Quality metrics

Two types of quality metrics have been used throughout this document:

- Standard statistics: **mean** and **standard deviation (std)**. It assumes the central limit theorem can be relied on to produce normally distributed estimates;
- Robust statistics based on ranking which are robust against deviation from a normal distribution assumption: **median** and a robust standard deviation (**std***) scaled from the InterQuartile Range (IQR) by a factor 27/20 assuming a normal distribution.

As recommended in the PVP [RD1], statistics with less than 30 samples have been discarded (except mentionned otherwise, different threshold used are 9 and 3). For readability, the number of figures has been restricted and limited, when necessary, to the robust statistics (median and robust standard deviation based on IQR) which are more representative of the majority of the distribution.

# 4 Validation of products, stability, resolution and product uncertainty estimates

In this section, we present a systematic validation with a focus on the CCI L4 version 2 product. It will be compared to the CCI L4 v1 and to the produced L3 (ascending, descending, combined) products when these data are different from L4. Section 4.1 describes the accuracy and precision of the products; section 4.2 analyses their stability and section 4.3 analyses the in situ vertical representativeness error. The temporal effective resolution and uncertainty are respectively assessed in section 4.4 and 4.5.

## 4.1 Accuracy & Precision

### 4.1.1 Products validation for L4

SSS are presented in the top panels of **Error! Reference source not found.** centred on 15th of January 2015 for the CCI+SSS monthly product and all Argo profiles top measurements. Although the two subplots of the top panel are difficult to compare, Argo profiles are point wise measurements and CCI provides an SSS field, there is a good agreement between the two sets of observations. The satellite derived product enables accurately mapping the gradient which is difficult with Argo point measurements. The subplots in middle panel represent the temporal median of CCI and of the gridded Argo taken from MDB. There is very good agreement in the resolved patterns between the two fields. In the gridded Argo MDB field some areas are not sampled enough (less than 30 grid points) particularly in some coastal areas affected by river plumes (e.g. Amazon, Congo rivers), strong boundary currents (e.g. Gulf Stream) or enclosed seas (e.g. Caribbean Sea, maritime continent), and in the open ocean in the middle of the subtropical gyres or at high-latitude (Arctic and Southern oceans). The temporal variability observed by CCI and Argo is represented on the subplots at the bottom of **Error! Reference source not found.** using the robust standard deviation. The high variability regions (e.g. in the vicinity of the Amazon and Congo plumes, in the Indian ocean, the ITCZ, or the Gulf stream) are well observed in both the CCI and Argo measurements. However, the high variability observed at high latitudes (e.g. Brazil-Malvinas Convergence Zone, Agulhas return current, Gulf Stream) with Argo floats is not totally reproduced by the CCI products. Part of this SSS variability might occur at finer spatial resolution than sampled by the satellites (<50 km) and this effect is expected to be more present at high latitude were mesoscale is at finer scale than at low latitudes.
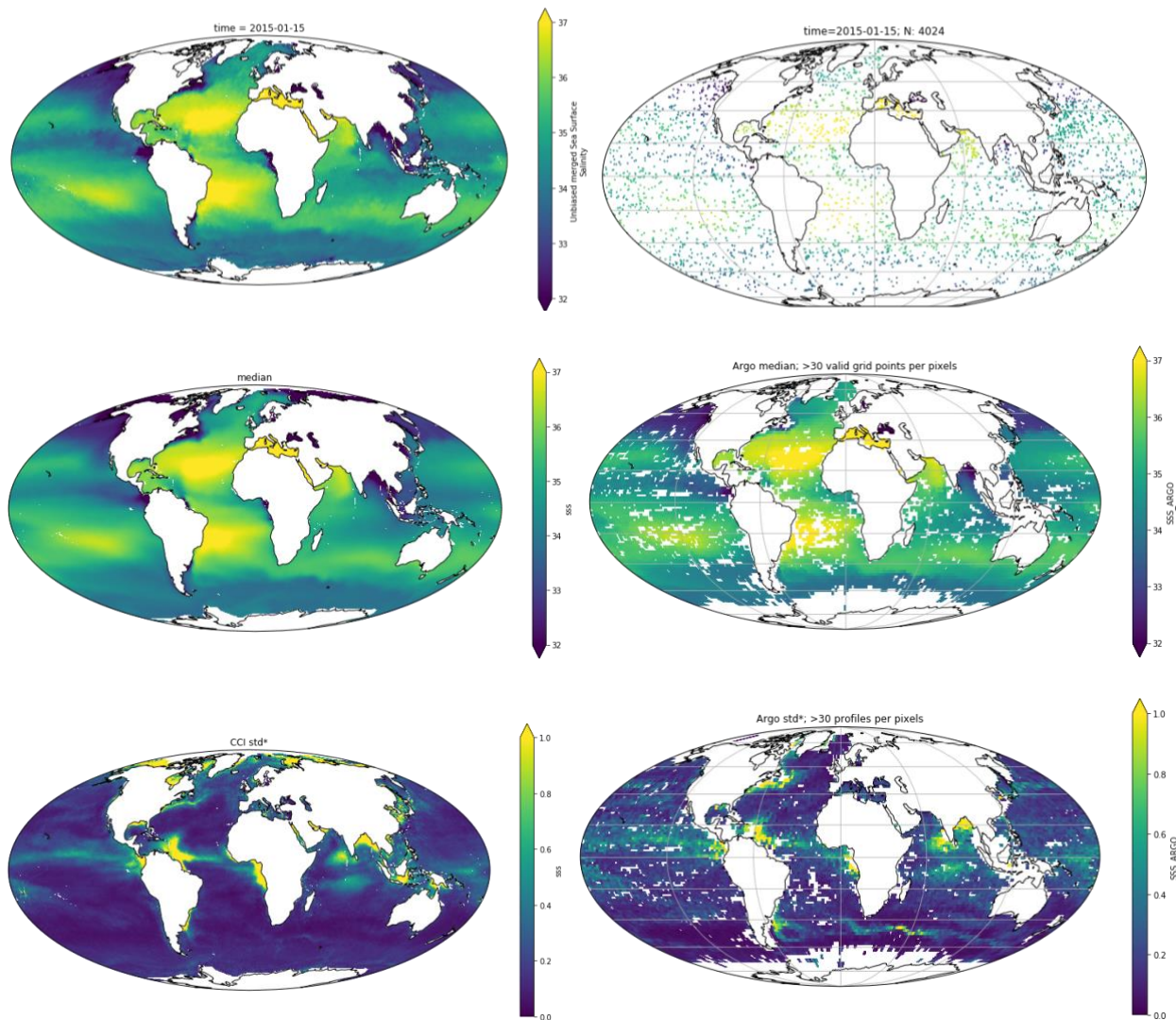
Figure 1: (left) L4 CCIv2 monthly, (right) Argo of (top) SSS fields for the 15th of January 2015; (middle) median field computed on the full time series; (bottom) robust standard deviation. All four tops subfigures share the same colour bar.
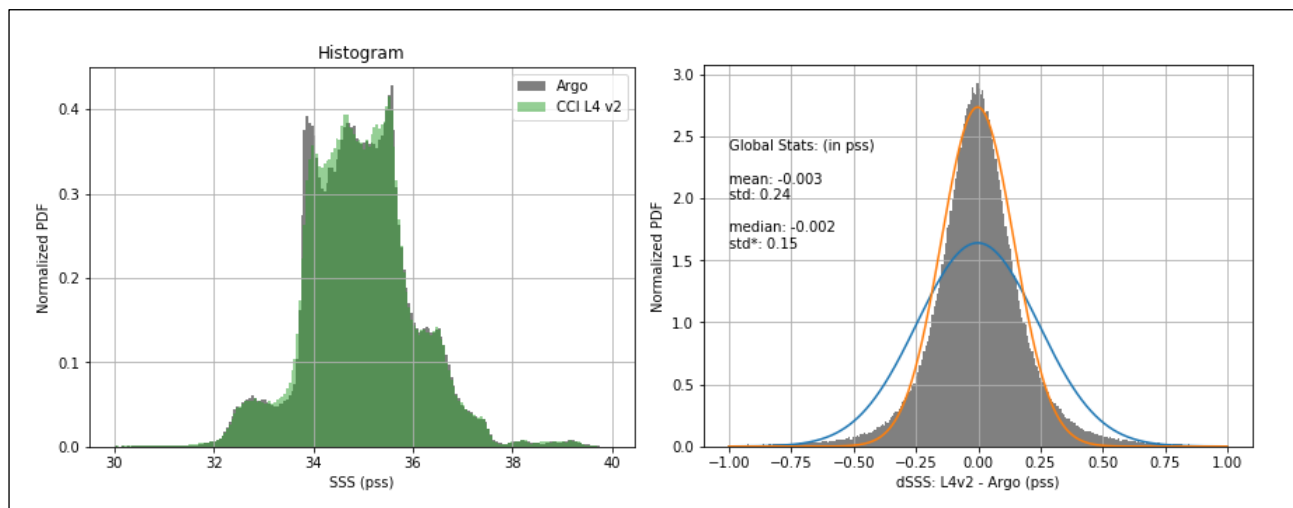
*Figure 2: Histogram of all pairwise gridded data (left) Argo SSS in grey and CCI L4 v2 in green; (right) CCI L4 v2 minus Argo difference, (blue line) normal pdf using computed mean and std, (orange curve) normal pdf using computed median and robust std.*

The distributions of the gridded pairwise CCI L4 and Argo SSS are very similar (Figure 2-left) over the full range of SSS from 30 pss to 40 pss. The distribution of the gridded pairwise CCI difference against Argo (Figure 2-right) highlights the absence of systematic bias (not significant at 5%), and a dispersion of 0.15 pss using the robust std (0.24 pss for the usual std). This difference between the robust and usual std is due to the non-normal distribution of the data difference (longer tails). The gridded pairwise measurements of Argo and CCI present a square of the Pearson correlation coefficient ($R^2$) of 97%.

Pi-MEP statistics against Argo DM for L4 v1.8 and v2.3, in addition to those for L3 products, are summarised in Table 4 and Table 5 (with PiMEP C1 criteria – no rain, moderate wind, SST>5°C, distance to coast > 800km).

Pi-MEP statistics of CCI L4 v2.3 against TSG, drifters and mammal's data with the C1 criteria are reproduced in Table 3. CCI data are adjusted on the mean of ISAS 2015 (up to 2015) and ISAS NRT (v6.2) on the period 2011-2018. Whereas TSG and drifters data are included in ISAS NRT, they are fully independent from ISAS 2015 (Gaillard et al., 2015). ISAS 2015 only uses Argo, marine mammals, ITP and TAO-Triton-Pirata-Rama data.
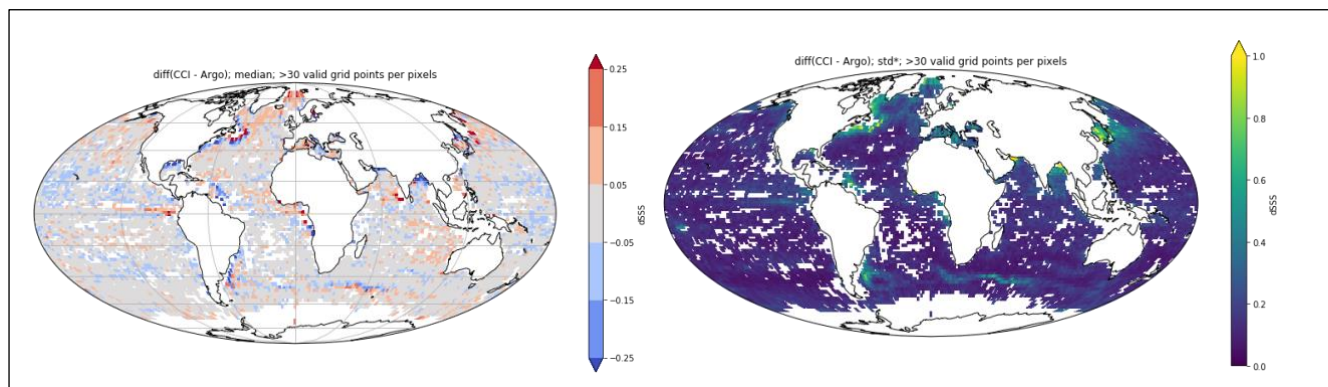
*Figure 3 (left) Temporal median and (right) temporal robust standard deviation of gridded pairwise SSS differences between CCI and Argo.*

To further assess the agreement between datasets, Figure 3 presents the temporal average (median) of the gridded pairwise SSS differences between CCI and Argo. At large scale (open ocean), the median difference is within +-0.05 pss and the robust std difference is below 0.2 pss. There is no large scale systematic spatial difference versus Argo. In the central Pacific Ocean, CCI is slightly fresher (blue) than Argo and it tends to be the opposite for high-latitude in the Northern hemisphere. Closer to the coast, major river plumes appear fresher (blue) in CCI. Close positive/negative differences are observed in the Gulf stream and Agulhas return current where meanders are common, suggesting differences between Argo and satellite sampling and spatial representativeness (pointwise versus 50km pixel). These higher discrepancies between CCI and in situ are also visible in the spread, with a temporal robust standard deviation of the differences higher than 0.4 pss at these fronts, in coastal areas and river plumes (Amazon plume, Bay of Bengal, …).

An assessment of the high-frequency variability is under review in (Stammer et al., 2020). Figure 4 compares an ensemble of in situ observations (details in the publication) with the respective CCI L4 product for different time scale. It shows that the higher temporal resolution of the satellite data leads to more meaningful small-scale high-frequency variability than in situ data. The magnitude of variability is approximately 1.5 times higher for CCI+SSS than for the ensemble in situ salinity, and larger differences between the spatial patterns are observable in the Gulf Stream region, Amazon outflow, eastern tropical Pacific, north-eastern Indian Ocean and around the maritime continent.
The standard deviation of the band-pass-filtered SSS data reflects the annual amplitude with large values in the ITCZ region, again off the major rivers and frontal regions. Here, the satellite and in situ fields show corresponding patterns, varying only slightly in magnitude. Larger differences between CCI+SSS and in situ variability spatial patterns and magnitudes can be observed in their low-frequency variability, in that interannual variability patterns are much larger for CCI+SSS than for the in situ data, especially in the central Indian Ocean, the major river outflow regions and the Northern North Atlantic. Higher amplitudes are also present in the equatorial Pacific Ocean in regions of strong vertical stratification, where the top surface variability is better sampled by satellite than in situ, as detailed in section 4.3.

On the other hand, many of the pattern have direct relations to physical mechanism such as run-off from large rivers.
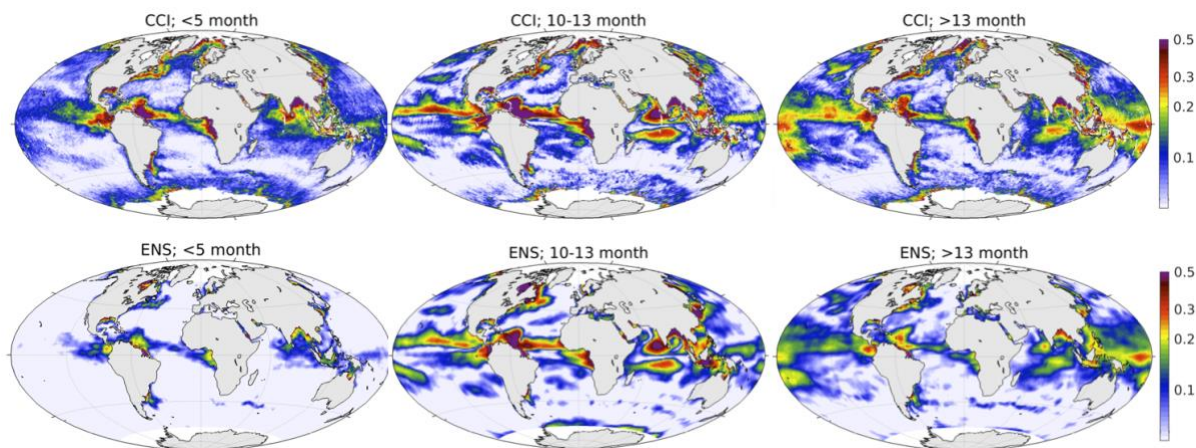


*Figure 4: STD of high-pass (<5 months) filtered time series of the CCI+SSS data and high-pass (<5 months) filtered time series of the ensemble in situ data. Middle and right column show the STD of the band-passed (10-13 months) filtered, and low-passed (>13months) filtered CCI+SSS and ensemble in situ data*

*Table 3: Statistics of CCI L4 v2.3 30dr against in situ data for the global ocean applying criteria C1 (only pairs where RR=0mm/h, 3<U<12m/s, SST>5°C, distance to coast > 800km). From PiMEP*

| Insitu database | Nb | median | mean | std* | std | R2 |
|---|---|---|---|---|---|---|
| Argo | 303937 | 0.00 | 0.00 | 0.13 | 0.16 | 0.97 |
| mammal | 3023 | -0.06 | -0.05 | 0.19 | 0.23 | 0.87 |
| drifter | 835159 | -0.02 | 0.01 | 0.13 | 0.22 | 0.98 |
| tsg-legos-dm | 442522 | -0.02 | -0.02 | 0.15 | 0.18 | 0.96 |
| tsg-gosud-research-vessel | 321856 | 0.02 | 0.02 | 0.13 | 0.18 | 0.97 |
| tsg-gosud-sailing-ship | 161867 | 0.02 | 0.03 | 0.14 | 0.16 | 0.98 |
| tsg-samos | 1260556 | 0.03 | 0.06 | 0.18 | 0.25 | 0.94 |
| tsg-legos-survostral | 37796 | 0.02 | 0.01 | 0.15 | 0.16 | 0.54 |
| tsg-ncei-0170743 | 87367 | 0.11 | 0.10 | 0.13 | 0.16 | 0.93 |
| tsg-polarstern | 28686 | -0.02 | -0.01 | 0.13 | 0.17 | 0.97 |

### 4.1.2 Products validation differences for Aquarius, SMAP, SMOS L3 products

In this subsection we will look at differences between CCI L3C v2 Aquarius, SMAP or SMOS products and CCI L4 v2. L3C products are simple grid averages of individual L2 satellite SSS after having applied systematic corrections.

The temporal mean differences between CCI L3C SSS products and L4v2 do not highlight regions higher than 0.05 pss except for the difference against SMOS where differences exceed 0.05 pss at high latitude and around Asia (Figure 5-top-left). The spread of L3C products versus L4 is generally contained below 0.1 pss excepted for SMAP and Aquarius in area of strong natural SSS variability or close to sea ice, and for SMOS, the spread exceeds 0.5 pss for most of the the Northern hemisphere in the vicinity (about 1,000 km) of the Asian, European and North-American continents, as well as in the vicinity of Antarctic (Figure 5-top-right).

The difference between ascending and descending passes does not show systematic differences higher than 0.05 pss for Aquarius and SMAP, but presents differences higher than 0.1 pss around the continents in the Northern hemisphere for SMOS (Figure 5-middle). Aquarius, does not highlight strong differences between ascending and descending orbits, but some regions appear as white in Figure 5-bottom relating to an absence of data in these regions during some period of the Aquarius time series. As a side note, the first time step of Aquarius ascending and descending data are corrupted and should only be used from the 15th of September 2011 onwards.

If one looks at the individual distributions of the gridded pairwise CCI L3C with Argo for each satellite, then none are able to reproduce the SSS distribution as measured by Argo floats (Figure 6), but CCI L4 combined products with a temporal OI are in very good agreement with Argo (Figure 2-left).

**Climate Change Initiative+ (CCI+) Phase 1**
Product Validation and Intercomparison Report

Ref.: ESA-CCI-PRGM-EOPS-SW-17-0032
Date:     11/11/2020
Version : v2.0
Page: 28 of 45
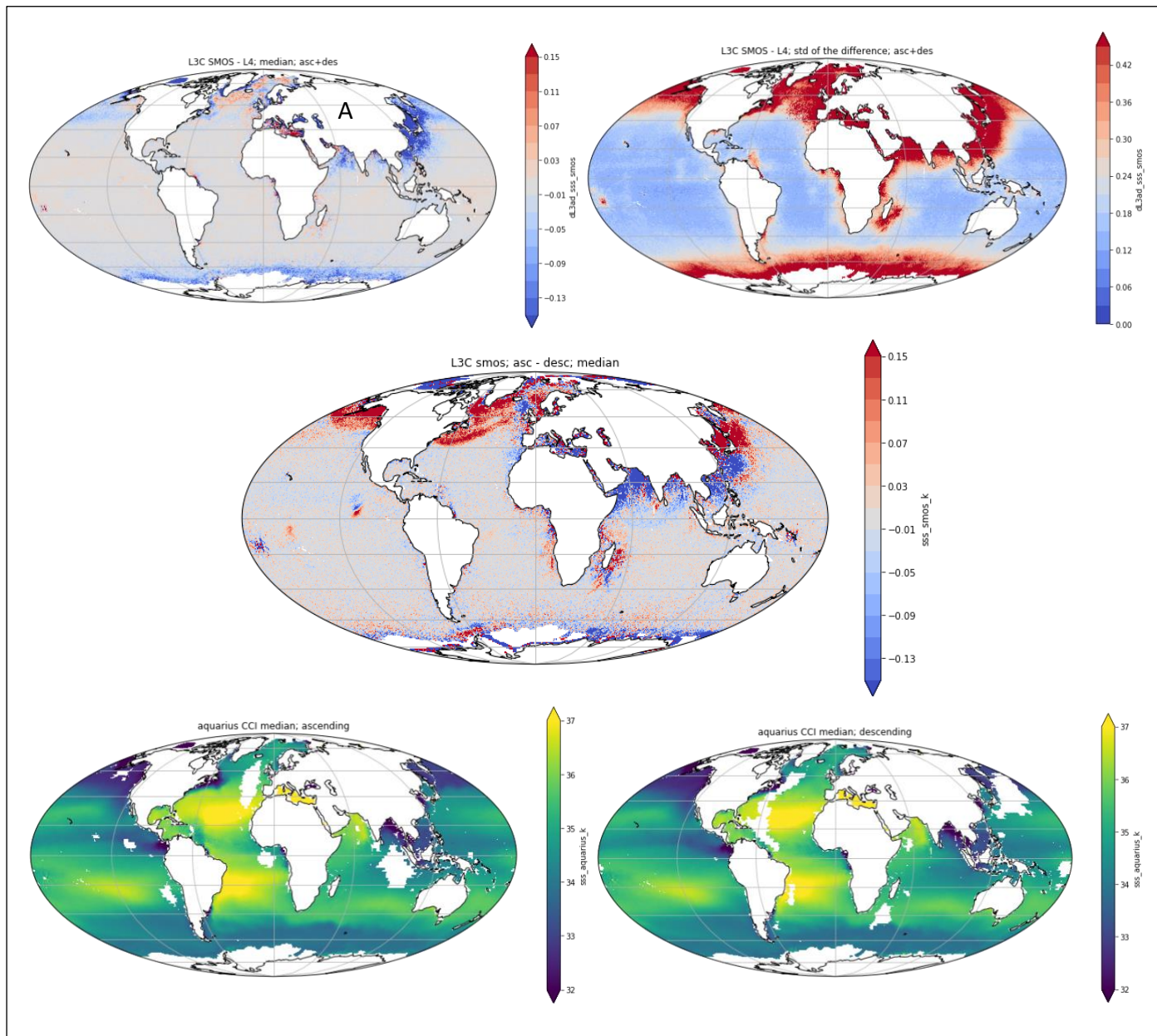
*Figure 5: (top-left) temporal median of the CCI L3C SMOS difference against CCI L4 v2; (top-right) same but for the temporal std of the difference; (middle) temporal median of the CCI L3C SMOS difference between Ascending and Descending passes; (bottom) temporal median of CCI L3C Aquarius SSS for (bottom-left) ascending and (bottom-right) descending passes.*
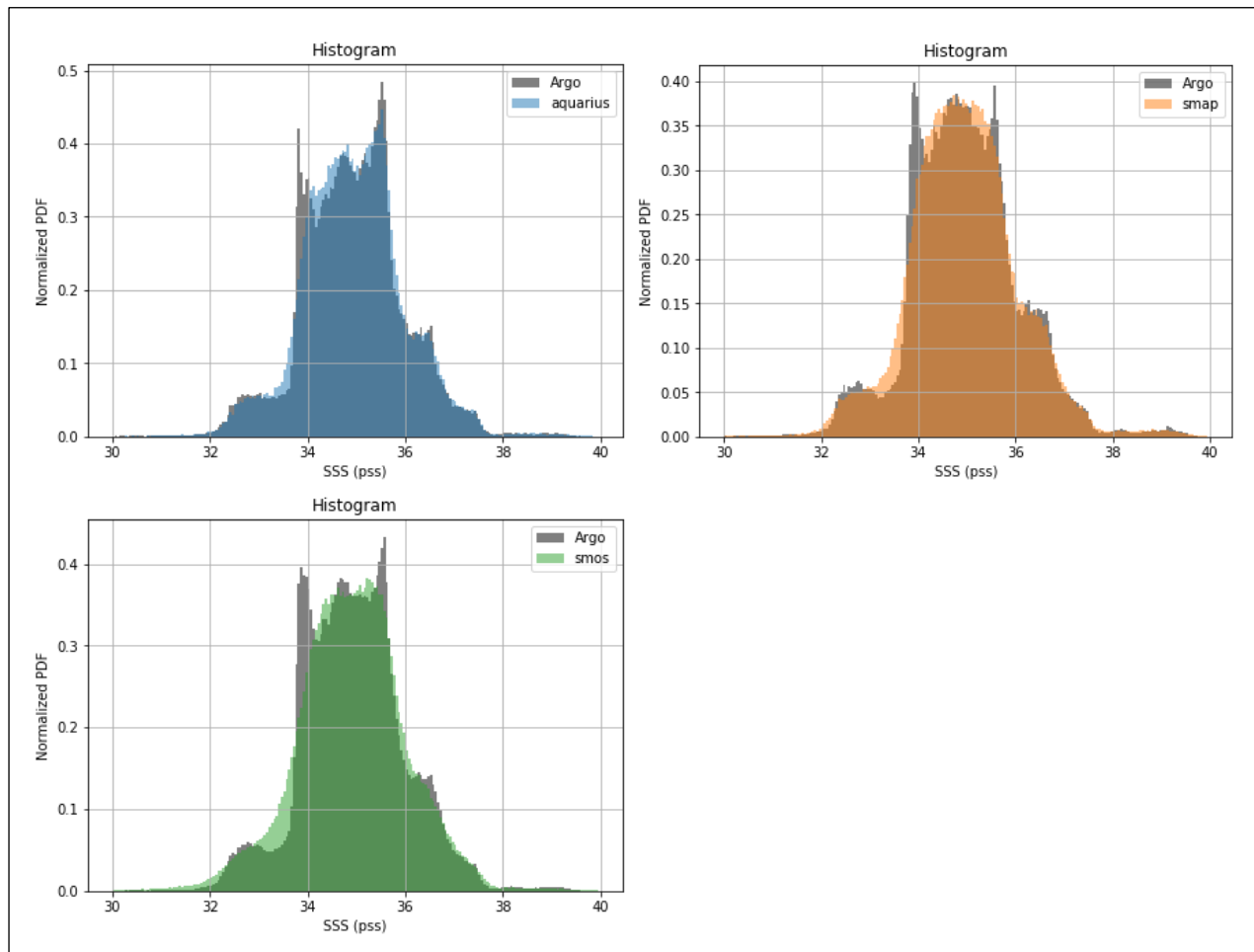
*Figure 6: Normalised PDF of all CCI/Argo pairwise gridded data for (grey) Argo, (blue) Aquarius, (orange) SMAP and (green) SMOS.*

*Table 4: Statistics of CCI products against Argo data for the global ocean. From PiMEP.*

| Satellite | Nb | median | mean | std* | std | R2 |
|---|---|---|---|---|---|---|
| | | | | | | Weekly running average |
| cci-l4-esa-merged-oi-v1.8-7dr | 809250 | 0.00 | 0.00 | 0.17 | 0.29 | 0.97 |
| cci-l4-esa-merged-oi-v2.3-7dr | 910889 | 0.00 | 0.00 | 0.17 | 0.28 | 0.97 |
| cci-l3c-esa-aquarius-v2.3-7dr | 327936 | 0.01 | 0.01 | 0.21 | 0.29 | 0.94 |
| cci-l3c-esa-smap-v2.3-7dr | 434331 | 0.00 | 0.00 | 0.34 | 0.42 | 0.95 |
| cci-l3c-esa-smos-v2.3-7dr | 873717 | -0.01 | -0.04 | 0.48 | 0.80 | 0.72 |
| | | | | | | Monthly running average |

**Climate Change Initiative+ (CCI+)**
**Phase 1**
Product Validation and
Intercomparison Report

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

Date: 11/11/2020

Version : v2.0

Page: 30 of 45

| cci-l4-esa-merged-oi-v1.8-30dr | 802009 | 0.00 | 0.00 | 0.17 | 0.30 | 0.97 |
| cci-l4-esa-merged-oi-v2.3-30dr | 895494 | 0.00 | 0.00 | 0.16 | 0.27 | 0.97 |
| cci-l3c-esa-aquarius-v2.3-30dr | 335811 | 0.01 | 0.02 | 0.17 | 0.28 | 0.95 |
| cci-l3c-esa-smap-v2.3-30dr | 443645 | 0.00 | 0.00 | 0.21 | 0.34 | 0.97 |
| cci-l3c-esa-smos-v2.3-30dr | 868609 | -0.01 | -0.03 | 0.28 | 0.53 | 0.90 |

*Table 5: Statistics of CCI products against Argo data for the global ocean applying criteria C1 (only pairs where RR=0mm/h, 3<U<12m/s, SST>5°C, distance to coast > 800km). From PiMEP.*

| Satellite | Nb | median | mean | std* | std | R2 |
|---|---|---|---|---|---|---|
| | | | | | Weekly running average | |
| cci-l4-esa-merged-oi-v1.8-7dr | 272928 | 0.00 | 0.00 | 0.14 | 0.16 | 0.97 |
| cci-l4-esa-merged-oi-v2.3-7dr | 309400 | 0.00 | 0.00 | 0.14 | 0.17 | 0.97 |
| cci-l3c-esa-aquarius-v2.3-7dr | 114333 | 0.01 | 0.01 | 0.17 | 0.20 | 0.96 |
| cci-l3c-esa-smap-v2.3-7dr | 151973 | 0.00 | 0.00 | 0.30 | 0.32 | 0.89 |
| cci-l3c-esa-smos-v2.3-7dr | 305341 | 0.00 | -0.01 | 0.38 | 0.46 | 0.80 |
| | | | | | Monthly running average | |
| cci-l4-esa-merged-oi-v1.8-30dr | 269747 | 0.00 | 0.00 | 0.13 | 0.17 | 0.97 |
| cci-l4-esa-merged-oi-v2.3-30dr | 303937 | 0.00 | 0.00 | 0.13 | 0.16 | 0.97 |
| cci-l3c-esa-aquarius-v2.3-30dr | 115006 | 0.01 | 0.01 | 0.14 | 0.17 | 0.97 |
| cci-l3c-esa-smap-v2.3-30dr | 153716 | 0.00 | 0.00 | 0.18 | 0.21 | 0.95 |
| cci-l3c-esa-smos-v2.3-30dr | 301967 | 0.00 | -0.01 | 0.22 | 0.26 | 0.93 |

**Climate Change Initiative+ (CCI+) Phase 1**
Product Validation and Intercomparison Report

Ref.: ESA-CCI-PRGM-EOPS-SW-17-0032
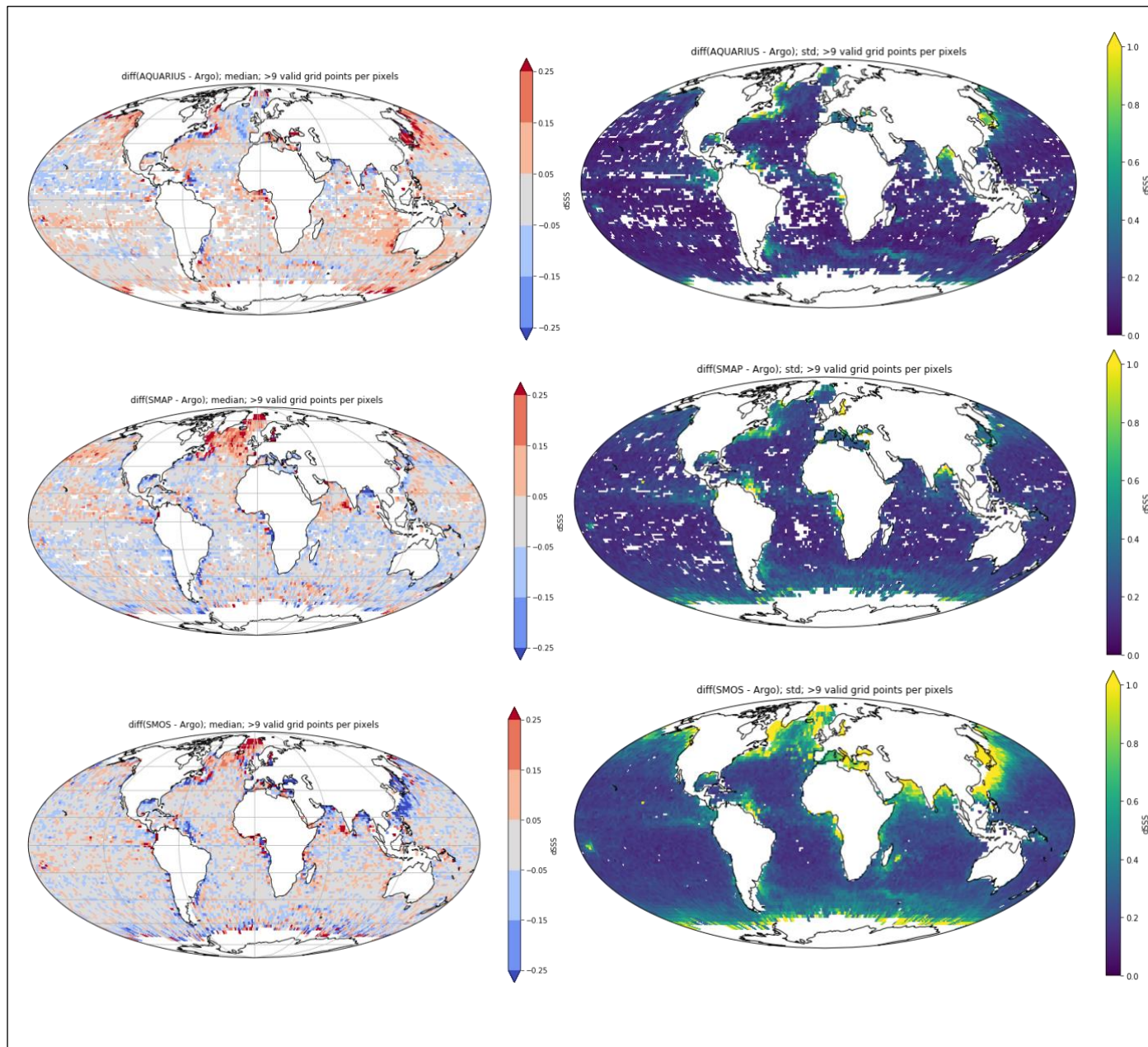Date:    11/11/2020
Version : v2.0
Page: 31 of 45

*Figure 7: (left) median of; (right) usual std; of the gridded pairwise difference between CCI L3 combined asc+desc products and Argo for (top row) Aquarius; (second row) SMAP; (last row) SMOS.*

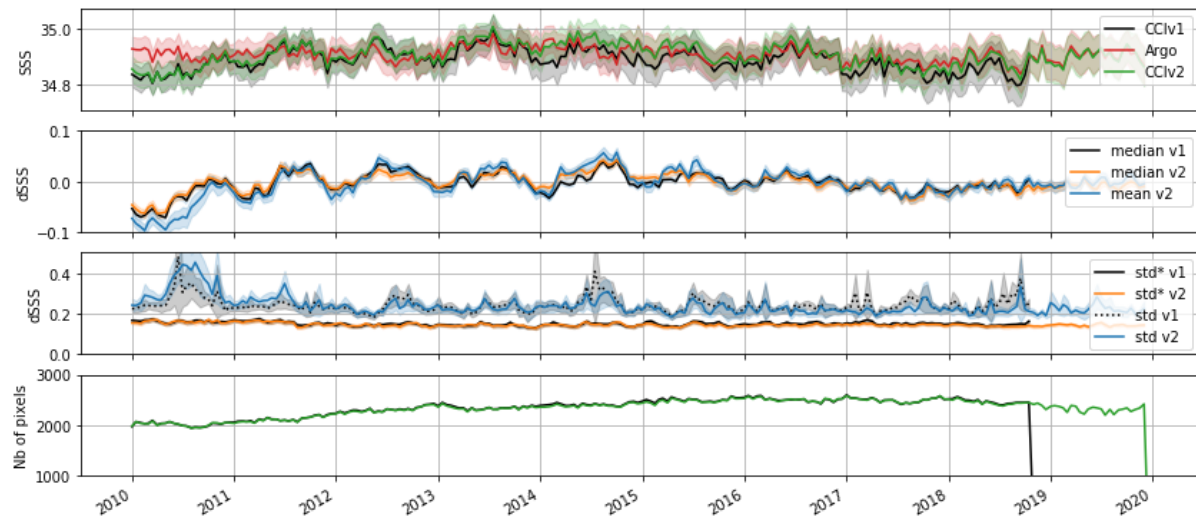## 4.2 Time series stability: intra-annual & long-term stability



*Figure 8: (1st panel): SSS mean of gridded pairwise Argo in red and L4 CCI measurements in black for v1 and green for v2; (2nd panel) Average of; (3rd panel) standard deviation of; the gridded pairwise SSS difference between CCI and Argo. Blue and black dashed lines represent (2nd panel) the mean (3rd panel) standard deviation. Orange and solid black lines represent (2nd panel) the median and (3rd panel) the robust standard deviation. The shading indicates the 95% confidence interval.*

The time series in Figure 8 represent the temporal evolution of gridded pairwise measurements of Argo and CCI (L4 v1 and v2) and their differences. The mean SSS temporal variability represented on the top panel shows good agreement between CCI and Argo with a mean around 34.9 pss. The beginning of the period in 2010 highlights a lower value for CCI than Argo of less than 0.1 pss. The two middle panels represent the gridded pairwise differences of CCI with Argo for average difference (mean and median); and dispersion (usual std and robust std). The global, temporal difference remains within ±0.05 pss. There is a small but appreciable global seasonal cycle with a minimum at the beginning of each year. The amplitude decreases with time, in particularly since 2015. The dispersion, as estimated by the robust std of the difference, stays relatively constant over the full time series between 0.13 pss and 0.16 pss. However, the standard deviation presents some peaks in the middle of the year (Northern Summer) suggesting more extreme values in the tail of the distribution. Globally, CCI L4 version 1 and 2 are very similar except version 2 is one year longer and version 2 includes Level 3 independent satellite data.

The same Argo and CCI L4 v2 pairwise gridded data, as in Figure 8, are reproduced in Figure 9 to simplify comparisons with Level 3 data. From the top panel, representing the global mean SSS, the main feature is L3 SMOS has a higher variability than that observed by Argo and other CCI products (i.e. L4 and L3 for Aquarius and SMAP data). Differences between SMOS and L4 or Argo during the first six months in 2010 exceed 0.1 pss. The second panel represents the median pairwise gridded difference between CCI products and Argo. The small trend in 2010, at the beginning of the time series is similar for L4 and SMOS (median in this panel versus mean in the top panel as the latter average is more affected by outliers from RFI). The global

**Climate Change Initiative+ (CCI+)**
**Phase 1**
Product Validation and
Intercomparison Report

Ref.: ESA-CCI-PRGM-EOPS-SW-17-0032
Date:    11/11/2020
Version : v2.0
Page: 33 of 45

seasonal cycle with a minimum at the beginning of each year is visible for L3 Aquarius, L4 and L3 SMOS data but its amplitude is stronger for the later. With the inception of SMAP data in February 2015, the amplitude of L4 differences decrease. In this second period, SMOS oscillation amplitude is smaller than before 2015. SMAP oscillation is smaller than for SMOS, but they are in phase opposition, which perhaps explain why L4 stay more stable in this second period.
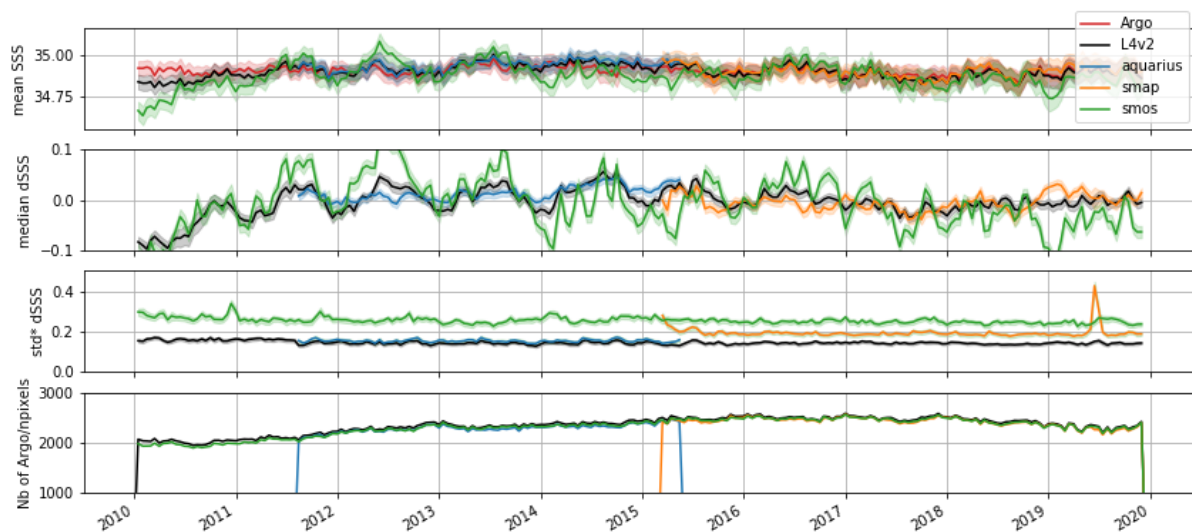


*Figure 9 (1st panel): SSS mean of gridded pairwise Argo and CCI measurement; (2nd panel) Median of; (3rd panel) robust standard deviation of; the gridded pairwise SSS difference between CCI and Argo. (4th panel) number of valid gridded pairwise Argo and CCI values. The curves are for pairwise gridded data with Argo of (black) CCI L4 v2 SSS and (red) Argo, and other curves in (colour) are for pairwise gridded data of L3C Aquarius, SMAP and SMOS L3C, asc. + desc., data.*

Concerning the precision (robust std of the difference), L4 is constant over the full time series between 0.13 pss and 0.16 pss, even in the period with only SMOS data, and is better than any other product taken independently. For L3 products, the best precision against Argo is for Aquarius (0.14-0.17 pss), then for SMAP (0.18-0.21 pss) and SMOS (0.24-0.28 pss). The peak in precision of SMAP corresponds to the period during which SMAP was in safe mode (19 June 2019 to 23 July 2019). For both SMOS and SMAP the std* is slightly higher at the beginning of their time series.

The temporal variability of the pair-wise CCI/Argo differences have been further assessed using latitude-time (Hovmöller) plots over the global ocean (Figure 10) for the L4 and L3 products. Most of the pixels appear in grey, indicating differences are below an absolute difference of 0.05 pss (confidence interval at 95% of 0.13 pss for 9 samples and a standard deviation of 0.2 pss). It appears there are significant oscillating signals with stronger amplitudes at higher latitude for all analysed CCI products (L4, L3C). The amplitude is of the same order of magnitude for L4, L3 Aquarius and L3 SMAP, but is stronger for L3 SMOS. For all cases, the oscillation is in phase opposition between Northern and Southern hemisphere. This means the CCI data are fresher in winter than Argo. The first year, 2010, indicates CCI L4

and L3C SMOS are fresher than Argo particularly in the Northern hemisphere at high latitudes. Early on in the time series, CCI is fresher than Argo at all latitude particularly for the first 6 months. A more quantitative estimate is provided in Figure 12.

The spatial representation of seasonal climatology of the gridded pairwise difference (Figure 11), calculated using the median for each season over the full time series, highlights fresher CCI L4 in the Northern hemisphere in Winter (DJF) and Spring (MAM) but saltier in Summer (JJA) and Fall (SON). This is particularly clear around Japan and in the northern North-Atlantic, which are also regions characterized with intermittent strong RFIs. A seasonal spatial signature is less pronounced in the Southern hemisphere. Some local seasonal differences are visible close to the coast, generally related to river plumes, potentially associated to vertical stratification (see details in section 4.3). The seasonal average spatial patterns for L3 Aquarius and SMOS are very similar to L4, but with higher amplitude for SMOS. SMAP has a low amplitude and different pattern, possibly reflecting the better RFIs sorting on SMAP (Figure 18 in Annex).
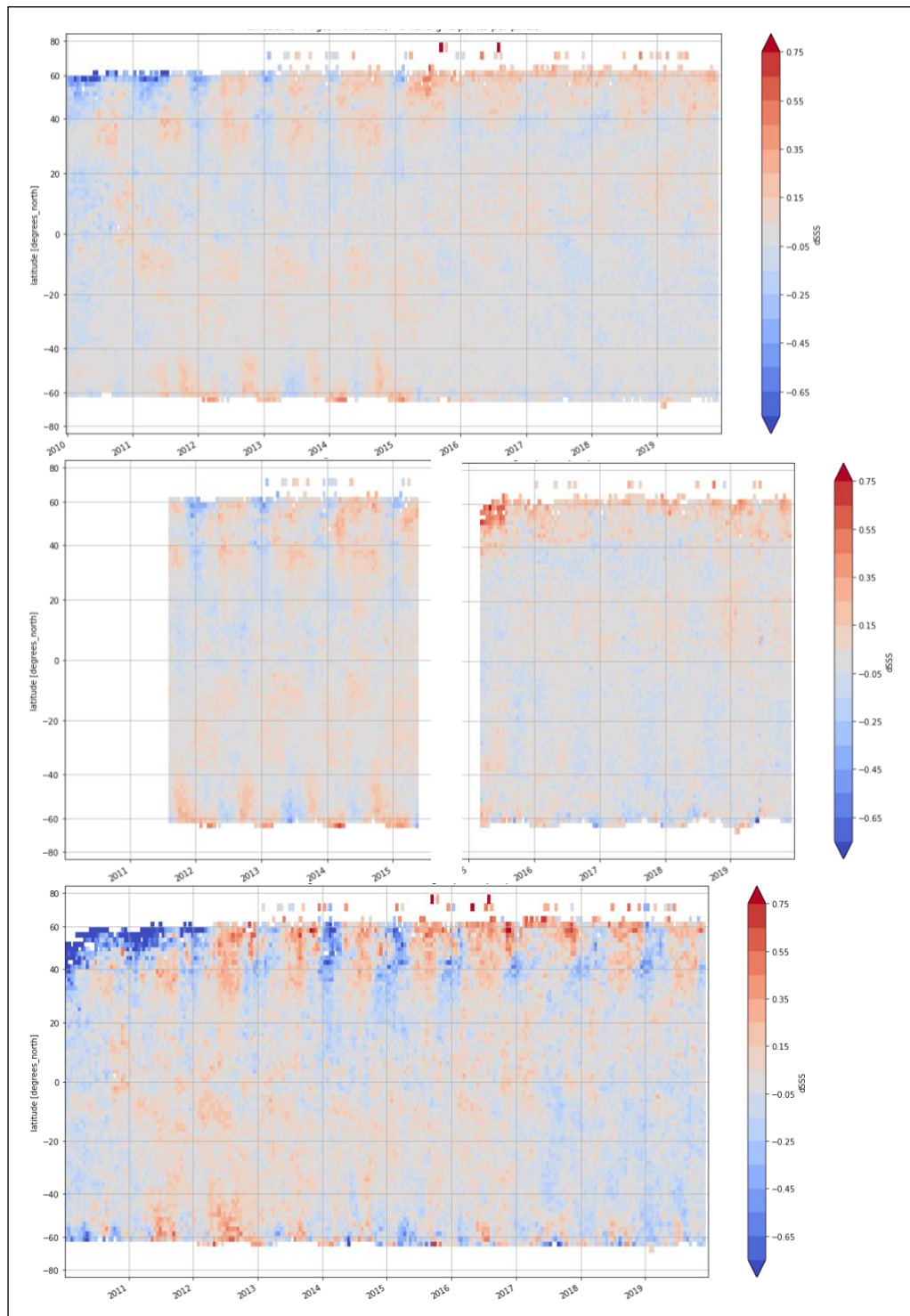
*Figure 10: Global latitude-time Hovmöller of the gridded pairwise CCI difference with Argo for (top) L4 v2; (middle-left) L3C Aquarius; (middle-right) L3C SMAP; (bottom) L3C SMOS. Each pixel represents the median value when there are more than 9 observations. Otherwise no value is shown. All sub-figures share the same colour bar.*
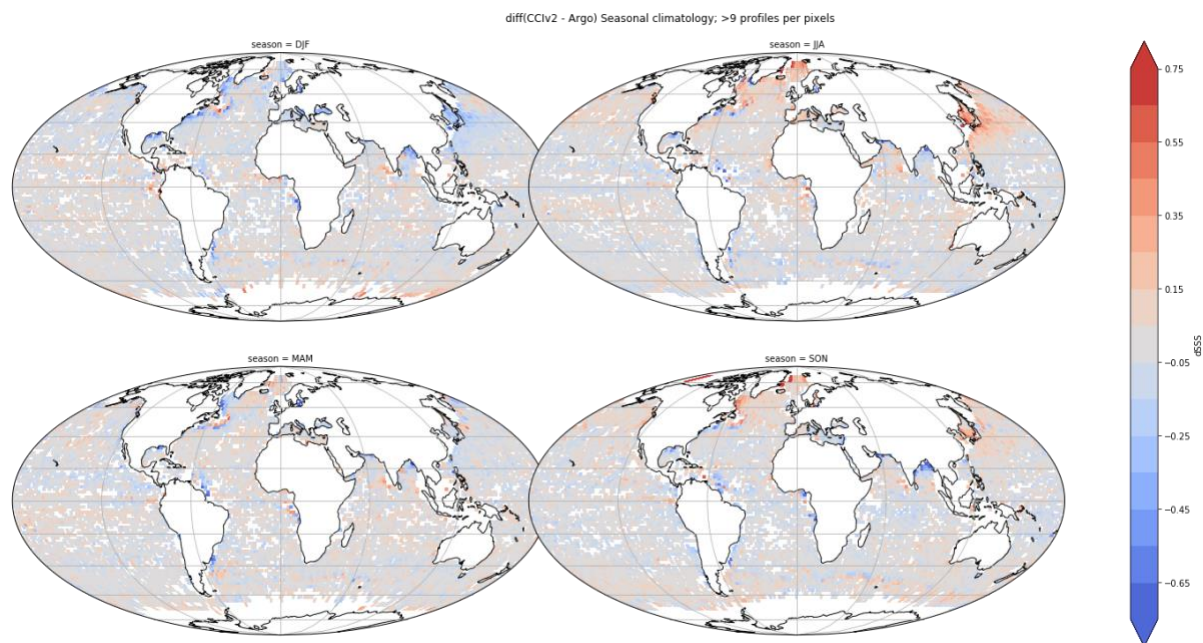
*Figure 11: Seasonal climatology of the gridded pairwise CCI L4 difference with Argo calculated using the median. Only pixels with more than 9 valid points are represented.*
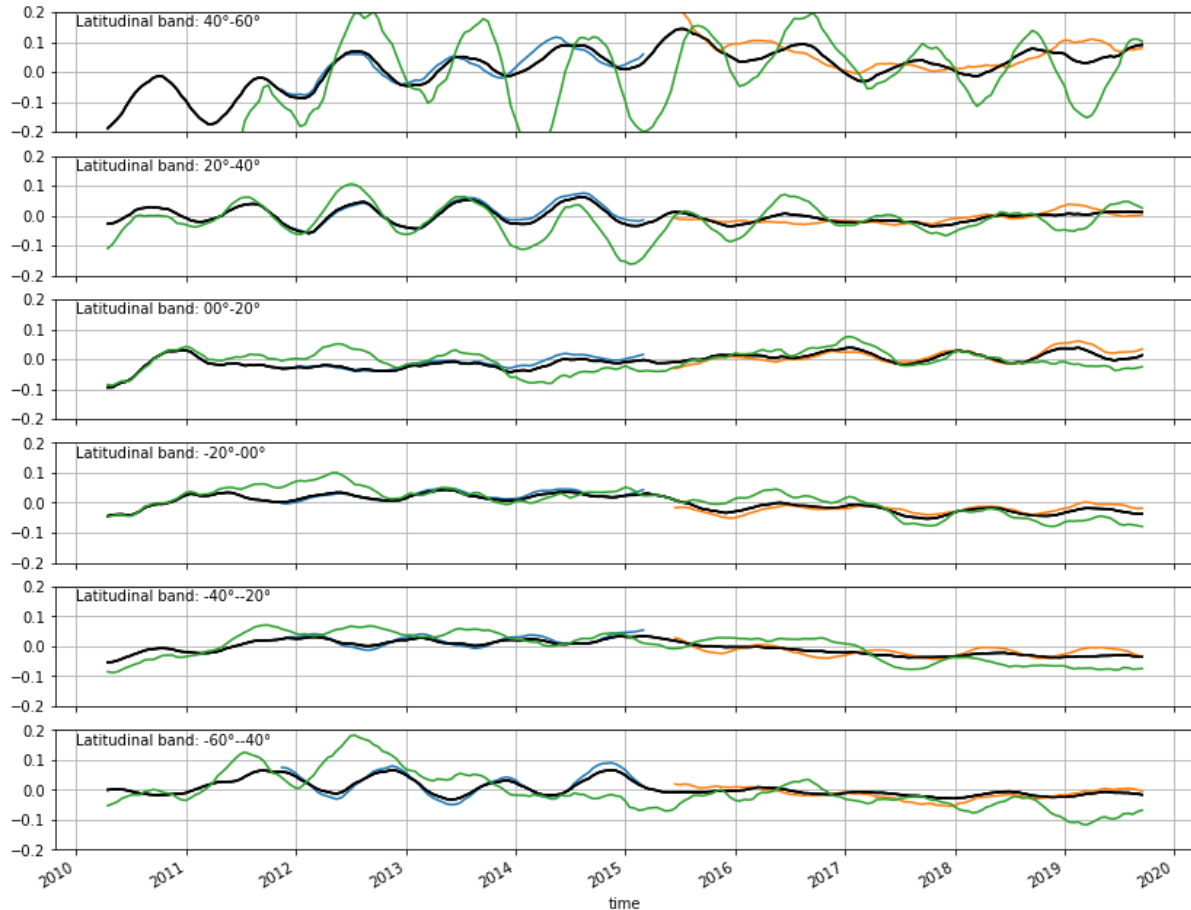
*Figure 12:Latidunal band (20° wide) median of the gridded pairwise SSS difference between CCI and Argo from (top) to (bottom) of [40°N;60°N] to [40°S;60°S]. A yearly rolling average is further applied to the data. Curves are in (black) for L4v2; (green) L3 SMOS; (blue) L3 Aquarius; (Orange) L3 SMAP.*

Figure 12 represents 20° latitudinal median averaging of the gridded pairwise CCI (L4 and L3 Aq., SMAP and SMOS) / Argo differences and applying a yearly running average. The seasonal oscillation of the differences is strongly marked for the band 40˚N-60˚N (top panel). The seasonal cycle amplitude for this band is maximum for L3 SMOS with an amplitude generally exceeding 0.2 pss; for L4 and L3 Aquarius the amplitude is lower at about 0.1 pss. Variability in SMAP differences are above 0.1 pss but without a clear seasonal cycle at these latitudes. There is a positive trend of 0.2 pss in L4 and L3 Aquarius from 2010 to Autumn 2015 for 40˚N–60˚N.

Other latitudinal bands have smaller seasonal cycles in the difference. The other most significant band is the band 20°-40°N. The amplitude of the oscillation of 0.1 pss is similar for L4 and L3 SMOS and Aquarius data. Whereas the amplitude decreases with the inception of SMAP for L4, it only decreases in SMOS in 2017. For this band, SMAP and SMOS differences tend to be in phase opposition, although amplitude in SMAP oscillation is smaller.

Although the band 40°S-60°S presents strong interannual variability there is no long-term seasonal cycle except during the Aquarius period for L4 and L3 Aquarius.

In terms of long-term stability, bands 60˚S-40˚S; 40˚S-20˚S; 20˚S-0˚; 10˚S-10˚N (not shown); 0˚-20˚N; 20˚N-40˚N stay within ±0.04 pss range for L4. Differences are similar for Aquarius and SMAP, but are larger for SMOS. All bands tend to have a positive trend for about the first 18 months. The variation in long term stability is maximum for band 40˚N-60°N as discussed above. Although very small, long term variations of the difference is opposite between the 20°S-0° and 0°-20°N bands.

Applying latitude-time (Hovmöller) and latitudinal band averaging by ocean basin (Pacific, Atlantic, Indian) gives slightly different results. After excluding data in 2010, there is a good agreement between CCI L4 and Argo for both seasonal differences and long-term stability where the latter remains within +-0.05 pss for Pacific and Atlantic Oceans between 40°S and 20°N and for Indian Ocean between 40°S and 10°S. Although differences between CCI and Argo are particularly strong in the Atlantic north of 40°N with consecutive minimum/maximum exceeding 0.3 pss before 2015, the amplitude decreases after 2015 staying just below 0.3 pss. In the Pacific, north of 40°N, the amplitude is below 0.3 pss up to end 2012 and decreases to 0.1 pss at the end of the time series. However, the oscillations are more pronounced in the Pacific (>+-0.1 pss) than in the Atlantic (generally +- 0.05 pss with peak at 0.1 pss) for the latitudinal band between 20°N and 40°N. In the Pacific for this band the amplitude of the difference decreases after 2015. Results by latitudinal band do not change when excluding data closer to 800 km to the coast.

## 4.3 In situ vertical representiveness error

The skin depth of satellite measurements depends on the wavelength; at a frequency of 1.4GHz, the skin depth is 1cm. In most situation, this depth is expected to represent well the first meters of the upper ocean, but significant differences between the surface ocean and a few meters depth have been observed in fresh regions either for a few hours after a rainfall (typically 1 to 5 hours, depending on the wind conditions) (Boutin et al. (2016); Supply et al. 2020), or in river plumes where large differences can be found between the first meter and a few meters depth (e.g. Supply et al. 2020).

In order to get a global distribution of this vertical representativeness error, we calculate the gradient for each Argo profile between an acquisition at 5m and 10m. We use the same grid as for the pairwise comparison and take the median value of this gradient for each cell (time and space). The seasonal climatology of this gradient in salinity is represented in Figure 13 highlighting most of the ocean does not show much gradients between 5m and 10m, excepted in areas with strong freshwater fluxes (river plumes, ITCZ, Labrador current, …). As expected, the surface at 5m is nearly always fresher than the based salinity at 10m. Surface is saltier only for very specific period as in the Mediterranean Sea in summer but for very small values. Strongest gradient in salinity concerns mainly the tropics and for all seasons with

typical values higher than 0.02 pss/m. If we extrapolate linearly the gradient from 5m to the surface, we could expect differences due to the vertical sampling exceeding 0.1 pss. In Summer, vertical gradients appear in the Northern Hemisphere in the vicinity of the western boundary current (Gulf Stream and Kurushio).

These results suggest the SSS measured by satellite would be fresher than the one measured in situ by Argo. However, this effect is an order of magnitude less than the one observed on Figure 11 so that the mean difference CCI minus Argo on Figure 11 does not show the same pattern.
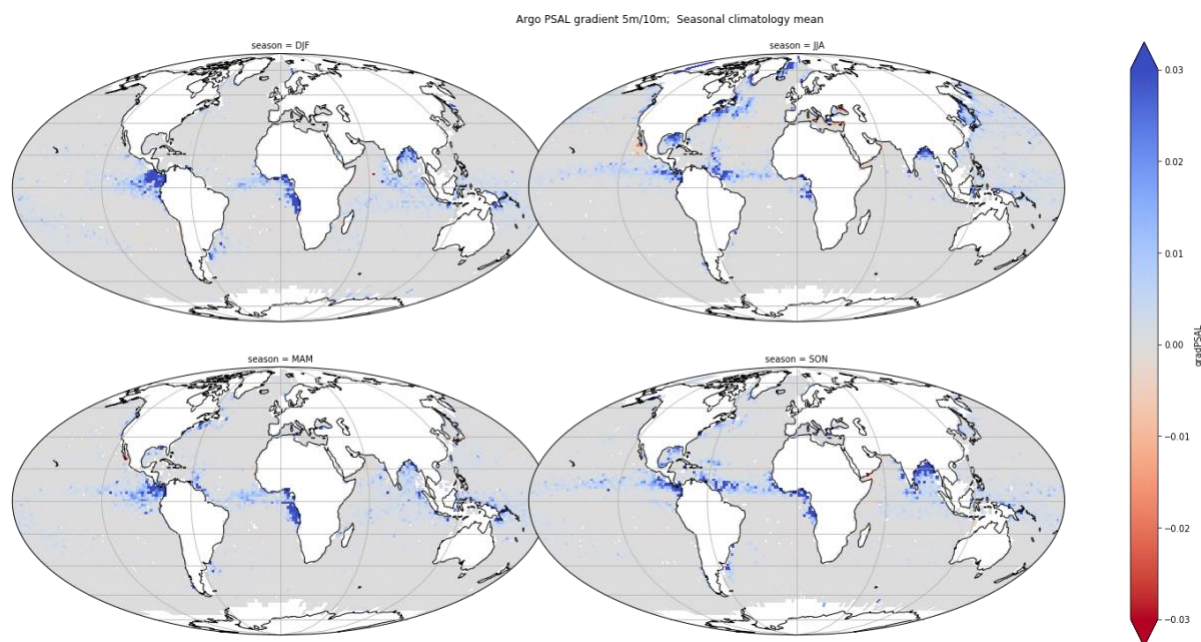


*Figure 13: Salinity gradient (in pss/m) derived from Argo between 5m and 10m. Gradient are gridded on the same grid as used for the pairwise difference (bi-weekly; 175 km).*

## 4.4 Temporal & spatial effective resolution

The temporal power spectrum of the Weekly (Figure 14) and Monthly (Figure 15) CCI L4 and L3 products shows as expected a decrease at the Nyquist frequency (respectively 14 days and 60 days). For the weekly products, whereas the merged L4 products power spectrum is very low for frequency higher than a week, they stay relatively high, with rebound, for the individual satellite products (Aquarius, SMAP or SMOS). The fact than the L4 merged product does not catch these higher frequency signals suggests the individual signals from L3 are not coherent and might be due to acquisition artefacts.

CCI L4 and CCI L3 Aquarius power spectrum are well aligned with the one from the moorings, but this is not the case for CCI L3 SMAP and SMOS products which might be explained by stronger noise floor of these satellites. On contrary, for the monthly products, all L4 and L3 products are well aligned with the one from the moorings.
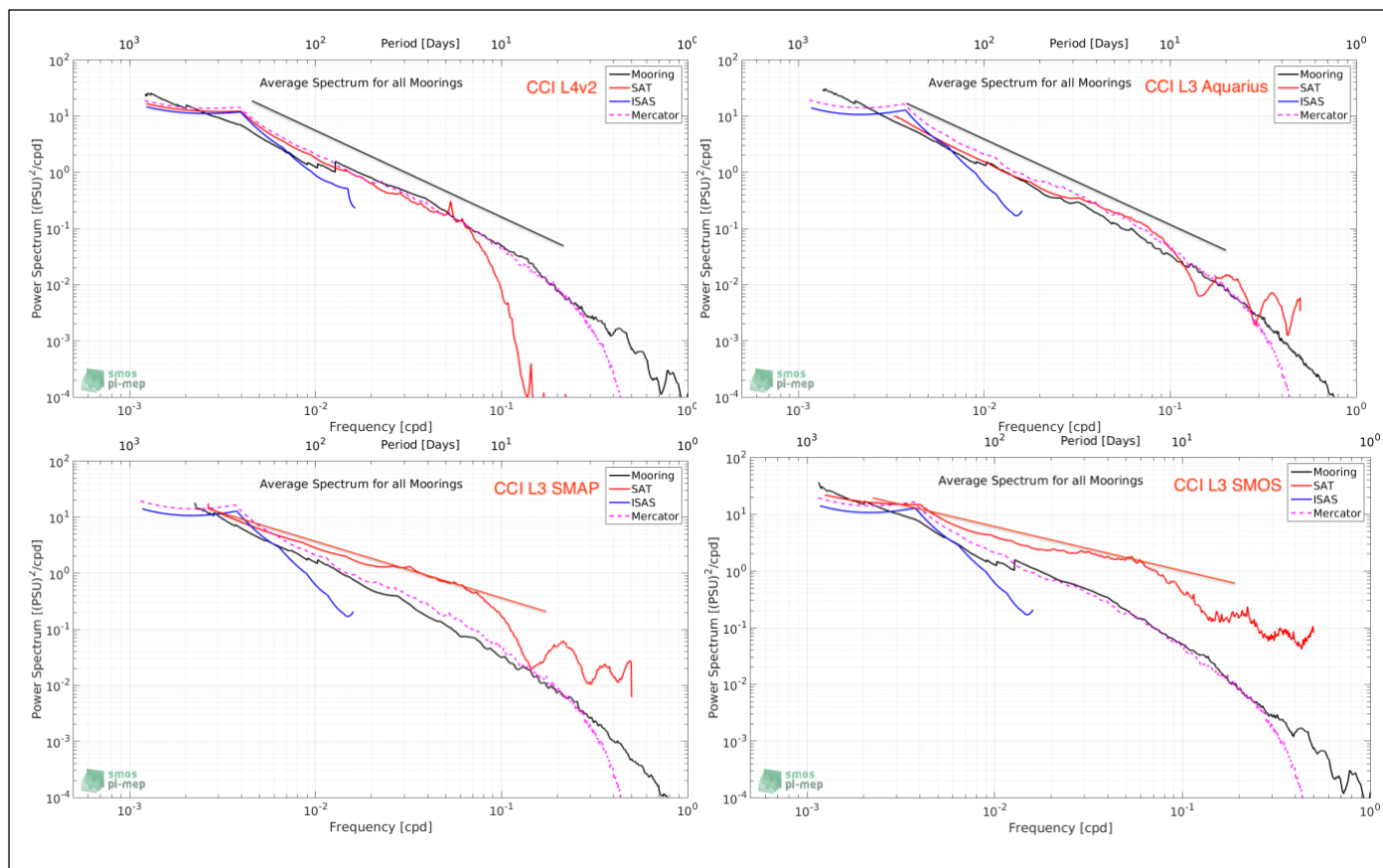


*Figure 14 : Average power spectrum of SSS from (black) moorings, (red) CCI weekly products stated in legend, (blue) ISAS, (pink) Mercator; from PiMEP. Straight lines have been added to ease the interpretation.*
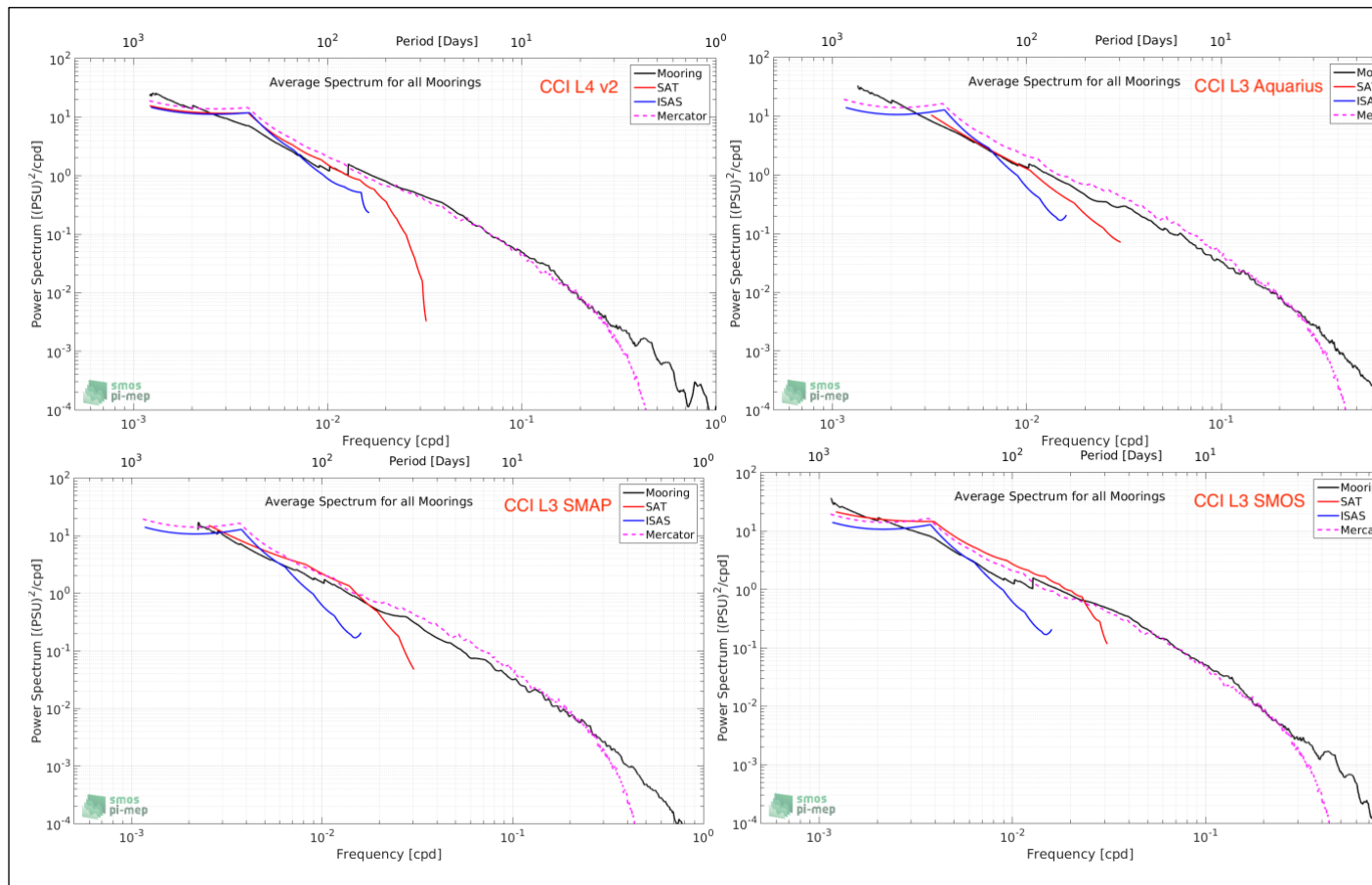
*Figure 15: same as above but for the Monthly products.*

## 4.5 Uncertainty

As explained in section 3.2 above, we will follow two approaches to validate satellite uncertainty estimates:

- Normalise the dSSS by the uncertainty with a centred reduced variable and analysed its variation compared to a theoretical behavior of a random and normalised variable of mean 0 and std 1.
- Compare the dSSS distribution with the uncertainty estimates.

For both cases, we will considered the satellite uncertainty $\Delta\sigma_{sat}$ alone or the total uncertainty which combines the satellite uncertainty with the reference uncertainty itself which includes the representiveness error ($\sigma_{tot} = \sqrt{\Delta\sigma_{sat}^2 + \Delta\sigma_{ref}^2}$).
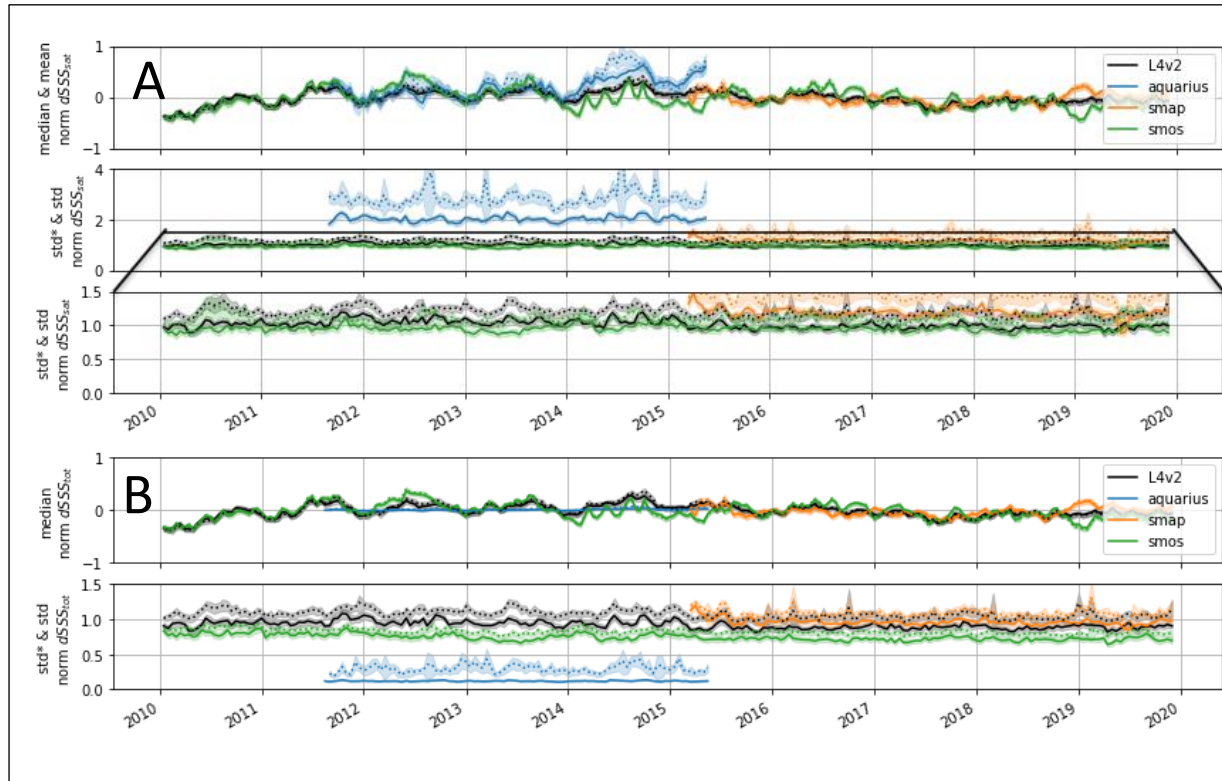
### 4.5.1 Normalised SSS



*Figure 16 : Time series of the normalised SSS normalised using (A) the satellite uncertainty; (B) the total uncertainty combining the sat and reference uncertainty. (1st row for each panels) represent (solid line) the median and (dashed line) the mean. (last row for each panels) represent (solid line) the robust std and (dashed line) the usual std. (middle row of top panel) is a zoom out of the top panel last row. (colours) are for the L4v2 and L3C Aquarius, SMAP and SMOS data.*

The normalised SSS represented in (Figure 16 A and B bottom rows) shows the standard deviation is closed to a value of 1 (excepted for L3C Aquarius), confirming a good estimate of the uncertainty (satellite or total) depending if we use the robust or usual std. For Aquarius, if one considers only the satellite uncertainty (Figure 16 A-middle row, blue line), the standard deviation exceeds a value of 1 indicating the uncertainty is strongly underestimated. On the contrary, of one considers the total uncertainty, accounting for the spatial representativeness error, it is strongly overestimated. Calculations for the spatial representativeness error are done with a spatial resolution of 25km. Results would be worse if we use the coarser real spatial resolution of Aquarius.

In detail, L3 SMAP uncertainty is slightly underestimate using the satellite uncertainty (Figure 16 A-bottom, orange line) but is correct once accounting the spatial representativeness error (Figure 16 B-bottom, orange line). L3 SMOS uncertainty is correct if one considers only the satellite uncertainty (Figure 16 A-bottom, green line), but is overestimated once accounting the representativeness error (Figure 16 B-bottom, green line). Surprisingly, L4v2 uncertainty

is correct whatever if the satellite or total uncertainty is used with the robust or usual std of each side of a value of one.

The average (median, mean) time series of the normalised SSS (Figure 16 A/B 1st row) does not inform on the uncertainty. For each point, we have about more than 2000 observations (cf Figure 8 last row), leading to a theoretical variability (std) of the normalised SSS of 0.02 which is much higher than what we observe.
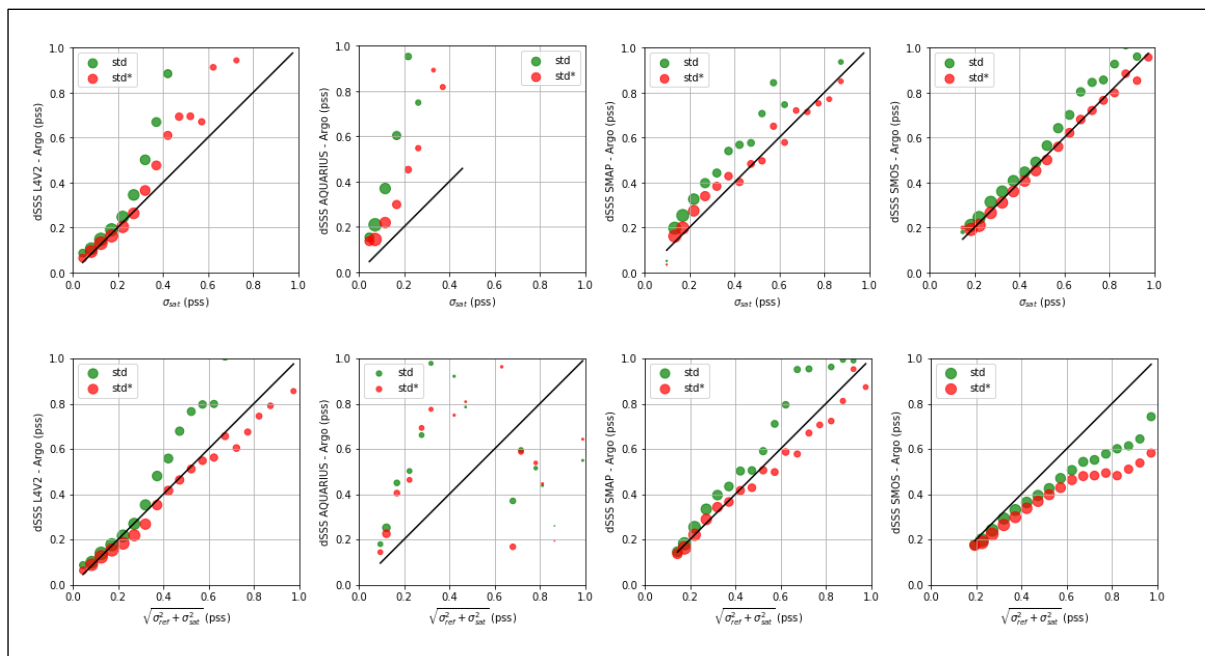
## 4.5.2 Compared SSS distribution



*Figure 17: measured standard deviation (green and red dots) for resp. usual and robust std; of the gridded pairwise CCI/Argo difference for each uncertainty bins. (top) using satellite uncertainty; (bottom) using total uncertainty - sat + ref. (column from left to right) for L4v2, L3 Aquarius, SMAP, SMOS. The size of the circle indicates the number of data.*

The gridded pairwise differences are binned by uncertainty bin of 0.05 pss wide (Figure 17) and computed over the full time series for each product. All products, except for Aquarius, have their uncertainties close to a one-to-one relation between the observed and estimated uncertainty. Aquarius uncertainty is strongly underestimated (factor 2 or more) for values of uncertainty below 0.5 pss and this whatever which uncertainty estimates is used (satellite or total). When using the total uncertainty, Aquarius uncertainty is overestimated particularly for the bins with total error between 0.6 and 0.8 pss where there is some data (dots size).

CCI L4, L3 SMAP and SMOS, observation of the uncertainty is within +-25% for depending which of the standard deviation metrics is used (usual or robust std).

*Page Intentionally Blank*
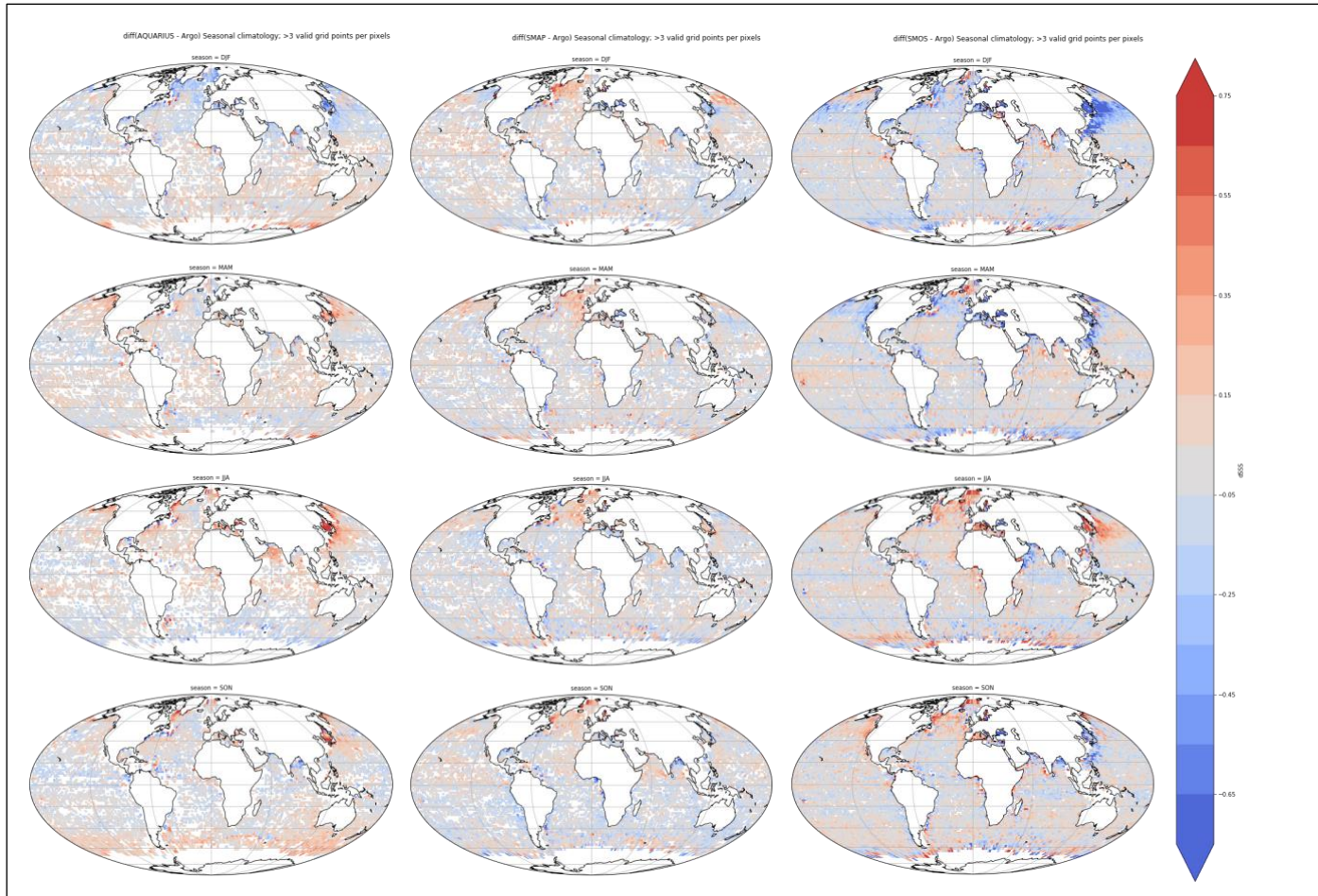
# Annex A: Supplementary material



*Figure 18 : Seasonal climatology of the gridded pairwise CCI L3C (from left to right: Aquarius, SMAP, SMOS) difference with Argo calculated using the median. Only pixels with more than 3 valid grid points are represented.*