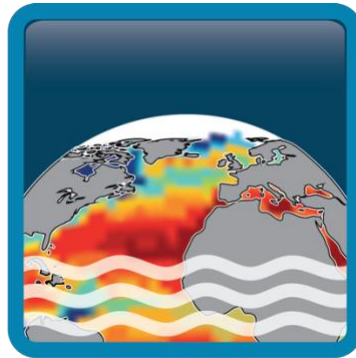


Climate Change Initiative+ (CCI+) Phase 2

Sea Surface Salinity



[D4.1] Product Validation and Intercomparison Report (PVIR)

Customer: ESA

Ref.: ESA-EOP-SC-AMT-2021-26

Version: v4.0

Ref. internal: 4000123663/18/I-NB

Revision Date: 21/01/2024



Signatures

Authors	Adrien Martin	NOC/NOVELTIS	
	Sébastien Guimbard	OceanScope	
	Gilles Reverdin	LOCEAN	
	Nicolas Kolodziejczyk	LOPS	
	Phoebe Hudson	NOC/UoE	
	Alexandre Supply	LOPS	
	Ramiro Ferrari	NOVELTIS	
Reviewed By	Chris Banks	NOC	
Approved By	Jacqueline Boutin (Science Leader)	LOCEAN	
	Nicolas Reul (Science Leader)	IFREMER	
	Rafael Catany (Project Manager)	ARGANS	
Accepted By	Roberto Sabia (Technical Officer)	ESA	

Diffusion List
Sea Surface Salinity Team Members

Amendment Record Sheet

Date / Issue	Description	Section / Page
	For previous amendment record sheet, please consult v3	
21/01/2024	Delivery to ESA for version 4	New document

Table of Contents

1	Introduction	1
1.1	Scope of this document	1
1.2	Structure of the document	1
1.3	Applicable Documents.....	1
1.4	Reference Documents	1
2	Executive Summary	3
2.1	Sea Surface Salinity products	3
2.2	Main results	3
2.3	Main results from Pi-MEP match-up reports.....	4
2.4	Recommendations and caveats to use CCI+SSS dataset	4
3	Validation: Data & Methods	5
3.1	Fiducial Reference Measurements	5
3.1.1	Level of independence to CCI products.....	5
3.2	In situ Dataset.....	6
3.2.1	Argo match-ups database: Pi-MEP MDB.....	6
3.2.2	Reference Dataset: Gridded MDB	7
3.3	Uncertainty validation	8
3.4	Quality metrics.....	9
4	Validation of Products, Stability, Resolution and Product Uncertainty Estimates.....	11
4.1	Accuracy & Precision	11
4.1.1	Global products validation for L4	11
4.1.2	Global products validation differences for Aquarius, SMAP, SMOS L3 products and weekly L4 products 17	
4.1.3	North Atlantic & Arctic Case Studies	19
4.2	Time series stability: intra-annual & long-term stability.....	31
4.3	In Situ Vertical Representiveness Error.....	34
4.4	Temporal & spatial effective resolution	35
4.4.1	Temporal effective resolution	35
4.4.2	Spatial effective resolution: Assessment of mesoscale features in Tropical Atlantic.....	37
4.4.3	Power Density Spectrum (PDS)	41
4.5	Uncertainty.....	43
4.5.1	Normalised SSS	43
4.5.2	Compared SSS Distribution.....	44
4.5.3	Estimation of mismatch error using GLORYS 1/12°	45

List of figures

Figure 1: (top-left) L4 CCIv4 monthly SSS for the 15th of January 2015, (top-right) Argo salinity measurements from the Pi-MEP MDB for the same month. (middle/bottom-left) temporal median/robust standard deviation of L4 CCIv4 SSS and (middle/bottom-right) temporal median/robust standard deviation of the gridded MDB SSS. Diagonal hatching indicates regions with less than 30 points over the full time-series. ----- 12

Figure 2 L4 CCIv4 monthly SSS for the 15th of January 2015 for Antarctic (top) and Arctic (bottom), (right) Argo salinity measurements from the Pi-MEP MDB and (left) temporal median/robust standard deviation of L4 CCIv4 SSS. Diagonal hatching indicates regions with less than 30 points over the full time-series.----- 13

Figure 3: Histogram of all pairwise gridded MDB (left) Argo SSS in grey and CCI L4 v4 SSS in green; (right) Gridded MDB of the CCI L4 v4 minus Argo difference, (blue line) normalized probability function (PDF) using computed mean and standard deviation, (orange curve) normalized PDF using computed median and robust standard deviation. ----- 13

Figure 4 (top-left) Temporal median and (top-right) temporal robust standard deviation of gridded pairwise SSS differences between CCI and Argo. (on the left) A moving window of 2 pixels in longitude and latitude is applied to the median in order to highlight statistical significance (at 95%) which are indicated with dots. (on the right) Robust standard deviation calculated with less than 30 valid grid points in time are hatched. (bottom) differences of the temporal median differences between CCI and Argo between version 4 and version 3 using the comparison dataset. ----- 15

Figure 5: Temporal mean of the difference satellite – Argo for (1st row) Aquarius; (2nd row) SMAP; (3rd row) SMOS. Taken from Pi-MEP reports. ----- 19

Figure 6 {left} Distribution of data in 2021 (colour represents SSS). (right) sub-domains selected as a function of hydrography and shelf break distance to coast. ----- 21

Figure 7 left time series of salinity in different bins as a distance to coast in (up left) south-west Greenland, and bottom (down left) south Greenland (blue in situ data; red V4.4; green a SMOS only LOCEAN products). Both are off the shelf (the lower closer to shelf break). Notice very low salinity in September-October 2021 on both. To the right, average profiles as a function of distance to coast (four sub-regions: West, South-west, South, and South-east Greenland). Upper panel average salinity of V4.4; lower panel, average difference CCI-in situ. ----- 21

Figure 8 Statistics of the time series as a function of distance to coast for four different sub-regions. Top panel root mean square difference of the in situ time series illustrating usual large decrease from shelf to ocean interior, except in W where the largest values are found in the middle of Denmark Strait (150-200 km to coast) due to the presence of large water mass fronts. Middle panel, rms differences between V4.4 and in situ. Lower panel, squared Pearson correlation coefficient between V4.4 and in situ. ----- 22

Figure 9 Distribution of CTD profiles between 2010 and 2022 from ICES data base (magenta, www.ices.dk), UDASH (green, <https://glodap.info/>) and Argo (red, <https://argo.ucsd.edu/>). --- 24

Figure 10 Comparison between SSS from ICES CTD database and CCI+SSS v4.4 (left), and res1_nows (right). (top) Scatter of SSS collocations as a function of the month around the year (colour). (middle) scatter of SSS differences as a function of SSS with mean (red dots) and

standard deviation envelope (black dots). (bottom) Monthly values of linear fitting error between satellite product and in situ data (blue), bias and STANDARD DEVIATION of the difference (red) and correlation (black). ----- 25

Figure 11 (upper panel) Tracks of Saildrones 1036 and 1037 in the Chukchi Sea and Beaufort Gyre (Supply et al., 2022). (lower panel) Timeseries of SSS from CCI+SSS v4.4 (red), res1_nows (blue) and Saildrone (black) 1036 (left) and 1037 (right).----- 26

Figure 12: GLORYS 1/12° v1 Salinity vertical transect (top) in 2016, (bottom) in 2019. Transect are obtained along red transect through in situ data represented on map in bottom left for each subpanel with the SSS from JPL SMAP as a background. In situ data are overlaid with black rings. On top of each subpanel, satellite SSS data along that transect are represented for (bottom to top) L3 JPL SMAP v5, L3 LOCEAN SMOS Arctic v1.1, L3 BEC SMOS Arctic+ v3.1, RSS SMAP v4, CCI v4.4 products with different flags applied. Satellite data are made semi-transparent when flags are applied, JPL SMAP excludes data where the provided uncertainty is > 1 pss; CCI_SSS_QC excludes SSS_QC flags, CCI_LSC_QC excludes Land Sea flags, CCI_ISC_QC excludes Ice Sea flags, CCI_SSS_QC2 excludes all flagged data. This figure has been adapted from Figure 3 in [Hudson et al., 2023].----- 28

Figure 13: (1st panel): mean SSS of gridded MDB of pairwise Argo in red and L4 CCI in black for v4 and in orange and green for comparison v3 and v4; (2nd panel) Average of; (3rd panel) standard deviation of; the gridded MDB of the pairwise SSS difference between CCI and Argo. Solid lines represent (2nd panel) the median (3rd panel) the robust standard deviation. Dashed lines represent (2nd panel) the mean (3rd panel) the standard deviation. The shading indicates the 95% confidence interval. (4th panel) number of valid gridded MDB values.----- 31

Figure 14: Global latitude-time Hovmöller of the gridded MDB of the pairwise CCI difference with Argo for (top) L4 v3; (bottom) L4 v4. Each pixel represents the median value after a moving window over 2 pixels in latitude and time. Data which are significantly different from 0 (at 95%) are indicated with dots. All sub-figures share the same colour bar. ----- 32

Figure 15: Global latitude-time Hovmöller of the gridded MDB of the pairwise CCI difference with Argo for (top) L4 v3 comparison dataset; (middle) L4 v4 comparison dataset; (bottom) Absolute difference between the two subplots above. Green indicates improvement towards zero; Red indicates degradation away from zero. Each pixel represents the median value after a moving window over 2 pixels in latitude and time. Data which are significantly different from 0 (at 95%) are indicated with dots. All sub-figures share the same colour bar. ----- 33

Figure 16: Seasonal climatology of the gridded pairwise CCI L4 difference with Argo calculated using the median. A moving window average of 2 x 2 pixels in longitude and latitude have been applied to increase the number of sampled, hence the significativity. Pixels, which are statistically significant (at 95%) are indicated with dots. ----- 34

Figure 17: Seasonal Salinity gradient (in pss/m) derived from Argo at 5 m and 10 m. Gradient are gridded on the same grid as used for the pairwise difference (bi-weekly; 175 km). ----- 35

Figure 18 : Average power spectrum of SSS from (black) moorings, (red) CCI Weekly products, (blue) ISAS, (pink) Mercator; (top) for CCI L4 Weekly v3.2; (bottom) for CCI L4 Weekly v4.4. from Pi-MEP. ----- 36

Figure 19: Average power spectrum of SSS from (black) moorings, (red) CCI Weekly products, (blue) ISAS, (pink) Mercator; (A) CCI L4 Monthly v3.2; (B) CCI L4 Monthly v4.4; (C) CCI L3C Monthly SMOS; (D) CCI L3C Monthly Aquarius; (E) CCI L3C SMAP. from Pi-MEP. ----- 37

Figure 20: CCI+SSS on 30 June 2011 with 93 TSG transects in the Subtropical North Atlantic (dashed) and 26 TSG transect in the Tropical Atlantic. ----- 38

Figure 21: a) Density spectra from from 88 collocated TSG (black); CCI+SSS v2.31 (dashed red); CCI+SSS v3.1 (solid red) SSS transects in Subtropical North Atlantic. Vertical thick black bar is the level of confidence at 95%. b) Coherency between the TSG and CCI+SSS SSS transects. Dashed line is the level of significance at 95%. c) Density spectra from from 26 collocated TSG (black)/CCI+SSS(red) SSS transects in Tropical Atlantic. Vertical thick black bar is the level of confidence at 95%. d) Coherency between the TSG and CCI +SSS SSS transects. Dashed line is the level of significance at 95%. ----- 40

Figure 22: Regions as defined in [Hoareau et al. 2018] where the PDS are calculated. Regions are NATL for North Atlantic Ocean; SPURS for Salinity Processes in the Upper ocean Regional Study; ITCZ for Inter-Tropical Convergence Zone; SPAC for South-Pacific ocean; STP for South-Tropical Pacific ocean; ARC for Agulhas Return Current. ----- 41

Figure 23: Power density spectra (PDS) in the 6 regions defined [Hoareau et al., 2018] and presented in Figure 22: Regions as defined in [Hoareau et al. 2018] where the PDS are calculated. for (in red) 30 days V3 (in black), 30 days V4 (in blue) days V4 products calculated from 2010 to 2020. A grey line, representing slope of -2.4, is provided for reference. ----- 42

Figure 24 : Time series of the normalised SSS normalised using (A) the satellite uncertainty; (B) the total uncertainty combining the satellite and reference uncertainty. Top row for each panel represents (solid line) the median and (dashed line) the mean. Bottom row for each panel represents (solid line) the robust standard deviation and (dashed line) the classic standard deviation. Colours are for the L4v4 and the comparison dataset for L4v3 and L4v4. ----- 44

Figure 25: measured standard deviation (green and red dots) for classic and robust standard deviation respectively; of the gridded pairwise CCI/Argo difference (dSSS) for each uncertainty 0.05 bin. (top) using satellite uncertainty; (bottom) using total uncertainty - satellite + reference (column from left to right) for L4v4, comparison dataset L4v4, comparison dataset L4v3. The size of the circle indicates the number of data. ----- 45

Figure 26: (top) Temporal median and (bottom) temporal robust standard deviation of the estimated sampling mismatch using GLORYS. This sampling mismatch estimates is obtain from the difference between GLORYS averaged over 50km, 30 days and GLORYS sampled at Argo time and position (horizontal and vertical). The colour scale is zoomed by 40% compared to the colour scale in Figure 4. ----- 46

Figure 27: (top) mean and (bottom) standard deviation of the observed difference between CCI and Argo per bins of expected mismatch using GLORYS with the different resampling strategy represented in colour (see legend). For clarity, on the top figure, only the most significant plots are presented. ----- 47


List of tables

Table 1 Applicable documents (as seen in CCI+SSS website, http://cci.esa.int/salinity)-----	1
Table 2 Reference documents -----	1
Table 3: Statistics for Pi-MEP MDB comparison dataset for v3 and v4 using exactly the same Argo floats colocation. Cf Pi-MEP reports for the details about the conditions criteria. -----	14
Table 4: Statistics of CCI L4 v4.4 30dr against in situ data for the global ocean applying criteria C1 (only pairs where RR=0 mm/h, $3 < U < 12$ m/s, SST > 5°C, distance to coast > 800km). From Pi-MEP -----	16
Table 5: Statistics of CCI L4 and L3C (SMOS, Aquarius, SMAP) products against Argo data for the global ocean. From Pi-MEP. -----	17
Table 6: Statistics of CCI L4 and L3C (SMOS, Aquarius, SMOS) products against Argo data for the global ocean applying criteria C1 (only pairs where RR=0mm/h, $3 < U < 12$ m/s, SST > 5°C, distance to coast > 800km). From Pi-MEP.-----	17
Table 7 Statistics (STANDARD DEVIATION of the differences, bias, RMSD and R ²) from comparison of CCI+SSS products v4.4 and res1_nows with Saldrones 1036 and 1038.-----	27
Table 8: Statistics of Reanalysis and satellite SSS products against in situ at 10m depth for the period a) 2010-2020 and b) 2015-2020-----	30

Acronyms (Check from SOW)

AD	Applicable document
ADB	Actions database
AMOC	Atlantic Meridional Overturning Circulation
ATBD	Algorithm theoretical basis documents
BRO	Brochure
CLIC	Climate and Cryosphere
DIR	Directory
DS	Dataset availability
DS-UM	Dataset user manual
DVP	Development and validation plan
EC RTD	European Commission Directorate General for Research and Innovation
EDS	Experimental dataset
EMI	Electromagnetic Interference
EO	Earth Observation
EOEP	Earth Observation Envelope Program
ESA	European Space Agency
FR	Final review
FWF	Freshwater fluxes
GCOS	Global Climate Observing System
IAR	Impact assessment report
ITT	Invitation to tender
IPP	Year of Polar Prediction
KO	Kick-off
MR	Monthly report
MTR	Mid-term review
MV-TN	Modelling and validation technical note
NDVI	Normalized Difference Vegetation Index
PAR	Preliminary analysis report
PGICs	Peripheral glaciers and ice caps
PM	Progress meeting
PMP	Project management plan
RD	Reference document
RB	Requirements baseline
SAR	Synthetic Aperture Radar
SIAR	Scientific and impact assessment report
SMOS	Soil Moisture and Ocean Salinity
SoW	Statement of work
SR	Scientific roadmap
SSS	Sea Surface Salinity

SST	Sea Surface Temperature
TDP	Technical data package
TDS	Training Data Set
TN	Technical note
VIR	Validation and intercomparison report
VR	Validation report
WCRP	World Climate Research Programme
WP	Work package
WS	Workshop minutes
WWRP	World Weather Research Programme
AD	Applicable document
ADB	Actions database
AMOC	Atlantic Meridional Overturning Circulation
ATBD	Algorithm theoretical basis documents
BRO	Brochure
CLC	Climate and Cryosphere
DIR	Directory
DS	Dataset availability
DS-UM	Dataset user manual
DVP	Development and validation plan
EC RTD	European Commission Directorate General for Research and Innovation
EDS	Experimental dataset
EMI	Electromagnetic Interference
EO	Earth Observation
EOEP	Earth Observation Envelope Program
ESA	European Space Agency
FR	Final review
FWF	Freshwater fluxes
GCOS	Global Climate Observing System
IAR	Impact assessment report
ITT	Invitation to tender
IPP	Year of Polar Prediction
KO	Kick-off
MR	Monthly report
MTR	Mid-term review
MV-TN	Modelling and validation technical note
NDVI	Normalized Difference Vegetation Index
PAR	Preliminary analysis report

	<p align="center">Climate Change Initiative+ (CCI+) Phase 2</p> <p align="center">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 1 of 63</p>
--	---	--

1 Introduction

1.1 Scope of this document

This document holds the Product Validation and Intercomparison Report (PVIR) prepared by the CCI+SSS team, as part of the activities included in WP400 of the Proposal (Task 4 from SoW ref. ESA-EOP-SC-AMT-2021-26).

This report contains an assessment of the CCI version 4 products for both Level 3 and Level 4 products for weekly and monthly time periods. The products are based on a temporal optimal interpolation of SSS data measured by SMOS, Aquarius-SAC and SMAP satellite missions.

1.2 Structure of the document

This document is composed of four 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

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

PSD	Product Specification Document	SSS_cci_PHASE#02-D1.2-PSD-v3.0
PUG	Product User Guide	SSS_cci_PHASE#02-D4.3-PUG-v4.0
PVP	Product Validation Plan	SSS_cci_PHASE#02-D2.5-PVP-v2.0
SoW	CCI+ Statement of Work	SOW

1.4 Reference Documents

Table 2 Reference documents

ID	Document	Reference
RD01	Product Validation Plan	SSS_cci_PHASE#02-D2.5-PVP-v2.0
RD02	In-situ database Analyses Report. Pi-MEP consortium. June 2023; Match-up database Analyses report, CCI L4 ESA GLOBAL MERGED-OI V4.4-MONTHLY Argo Global Ocean: pimep-mdb-report_GO_cci-l4-esa-merged-oi-v4.4-30dr_argo_20230615.pdf	
RD03	Guimbard, S.; Reul, N.; Sabia, R.; Herlédan, S.; Khoury Hanna, Z.E.; Piollé, J.-F.; Paul, F.; Lee, T.; Schanze, J.J.; Bingham, F.M.; Le Vine, D.; Vinogradova-Shiffer, N.; Mecklenburg, S.; Scipal, K.; Laur, H. The Salinity Pilot-Mission Exploitation Platform (Pi-MEP): A Hub for Validation and Exploitation of Satellite Sea Surface Salinity Data. <i>Remote Sens.</i> 2021 , <i>13</i> , 4600. https://doi.org/10.3390/rs13224600	



**Climate Change Initiative+ (CCI+)
Phase 2**

**Product Validation and
Intercomparison Report**


Ref.: ESA-EOP-SC-AMT-2021-26

Date: 1/21/2024

Version : v4.0

Page: 2 of 63

ID	Document	Reference
RD03	G. Reverdin, S. Morisset, L. Marié, D. Bourras, G. Sutherland, B. Ward, J. Salvador, J. Font, Y. Cuyppers, 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. <i>Oceanography</i> 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. <i>IEEE Transactions on Geosciences and Remote Sensing</i> 56, 5525-5536	Hoareau et al., 2018
RD05	ATBD	SSS_cci_PHASE#02_D2.3-ATBD-v4.0
RD06	End-to-End ECV Uncertainty Budget - E3UB	SSS_cci_PHASE#02_D2.3-E3UB-v4.0
RD07	Product Validation Plan version #1	SSS_cci-D2.5-PVP-v1.1
RD08	GLORYS 1/12° Reanalysis	
RD05	Product Validation and Intercomparison Report version 2	
	Stammer et al., 2020, How good do we know ocean salinity and its changes? <i>Progress in Oceanography</i> , vol. 190, p. 102478, 2021, doi: https://doi.org/10.1016/j.pocean.2020.102478	Stammer et al., 2020
	Boutin et al 2016, Satellite and In Situ Salinity: Understanding Near-surface Stratification and Sub-footprint Variability, <i>Bulletin of American Meteorological Society</i> , 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, <i>Remote Sensing of Environment</i> , 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: <i>Satellite Precipitation Measurement</i> , V. Levizzani, C. Kidd., D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, F. J. Turk, Eds., Springer Nature, Cham, <i>Advances in Global Change Research</i> , 69, 1155-1176, https://doi.org/10.1007/978-3-030-35798-6_34 .	Supply et al, 2020b
	Supply, A., Boutin, J., Kolodziejczyk, N., Reverdin, G., Lique, C., Vergely, J.-L., & Perrot, X. (2022). Meltwater lenses over the Chukchi and the Beaufort seas during summer 2019: From in situ to synoptic view. <i>Journal of Geophysical Research: Oceans</i> , 127, e2021JC018388. https://doi.org/10.1029/2021JC018388	Supply et al., 2022
	Hudson, P. A., Martin, A., Josey, S., Marzocchi, A., and Angeloudis, A.: Drivers of Laptev Sea interannual variability in salinity and temperature, <i>EGUsphere</i> [preprint], https://doi.org/10.5194/egusphere-2023-1403 , 2023.	Hudson et al., 2023

	<p align="center">Climate Change Initiative+ (CCI+) Phase 2</p> <p align="center">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 3 of 63</p>
--	---	--

2 Executive Summary

2.1 Sea Surface Salinity products

The products validated are:

CCI Level 4 version 4 products (referred in the following either as v4.4 or v4) for Weekly and Monthly averaged products. The products are distributed in three grids:

Using a regular 0.25° grid for the global ocean:

```
ESACCI-SEASURFACESALINITY-L4-SSS-GLOBAL-
MERGED_OI_7DAY_RUNNINGMEAN_DAILY_0.25deg-xxxxxxx-fv4.4.nc
ESACCI-SEASURFACESALINITY-L4-SSS-GLOBAL-
MERGED_OI_Monthly_CENTRED_15Day_0.25deg-xxxxxxx-fv4.4.nc
```

Using an equal area (EASE2) grid for the Northern and Southern poles:

Northern Hemisphere

```
ESACCI-SEASURFACESALINITY-L4-SSS-POLAR-
MERGED_OI_7DAY_RUNNINGMEAN_DAILY_25kmEASE2-NH-xxxxxxx-fv4.4.nc
ESACCI-SEASURFACESALINITY-L4-SSS-POLAR-
MERGED_OI_Monthly_CENTRED_15Day_25kmEASE2-NH-xxxxxxx-fv4.4.nc
```

Southern Hemisphere

```
ESACCI-SEASURFACESALINITY-L4-SSS-POLAR-
MERGED_OI_7DAY_RUNNINGMEAN_DAILY_25kmEASE2-SH-xxxxxxx-fv4.4.nc
ESACCI-SEASURFACESALINITY-L4-SSS-POLAR-
MERGED_OI_Monthly_CENTRED_15Day_25kmEASE2-SH-xxxxxxx-fv4.4.nc
```


In section 4.1.2 p.17, CCI Level 3 v4 concatenated products for SMOS, SMAP and Aquarius are used:

```
ESACCI-SEASURFACESALINITY-L3C-SSS-
SMOSSMAPAQUARIUS_Monthly_Centred_15Day_25km-20220801-fv4.4.nc
```

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 dataset derived from Argo floats upper salinity measurements (see details in section 3.1 below);
- CCI v4 data available in the Arctic and Southern Ocean (see details in section 4.1.3 for case studies validation in North-Atlantic and in the Arctic);
- No systematic biases against reference data found (see details in summary for Pi-MEP match-up report in section 2.3 below for more details);
- Global precision against reference gridded data is of 0.14;

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 4 of 63</p>
--	---	--

- CCI version 4 products show better performance than version 3, in particular is less affected by seasonal undulation at high latitude;
- Degraded performance of the CCI v4 in the North-Atlantic from April to August 2015, linked to known issue in SMAP v5.0.
- Remaining seasonal oscillation of CCI SSS differences against reference before 2015:
 - Amplitude is maximum in the Northern Hemisphere;
- CCI SSS is lower than reference data at the beginning of the time series (2010) with an amplitude up to 0.1 pss;
- Mesoscale features in Tropical Atlantic are coherent between CCI v3, v4 and TSG transects up to a wavelength of 250-300km (features of ~150km);
- Uncertainties provided in CCI product are in good agreement with observations (within 30%).


2.3 Main results from Pi-MEP match-up reports

- No global bias found against Argo except for filtered collocations where:
 - SSS less than 33 pss (CCI saltier by 0.05 pss);
 - SSS higher than 37 pss (CCI fresher by 0.05 pss);
- Global precision (robust standard deviation; pairwise difference) against Argo of 0.15 pss
 - Decreasing to 0.12 and 0.13 pss for optimal region (>800 km from the coast; area with temporal standard deviation smaller than 0.2 pss);
 - Increasing to 0.2 pss for area characterised by one of the following conditions: rain and low wind; area with temporal standard deviation >0.2pss; closer than 150 km to the coast; SSS < 33 pss;
- Systematic improvements of v4 against v3 (fraction of variance unexplained improved globally by 10%) with particular improvements for regions: close to the coast (< 150km), with SST < 5°C or with SSS > 37 pss.
- Good agreement between the observed CCI SSS product power spectra and moorings for the two averaging period (weekly and monthly).

2.4 Recommendations and caveats to use CCI+SSS dataset

CAVEATS

- 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 pss) in the Northern Hemisphere which decreases after 2015.
- In the North Atlantic, between April and August 2015, CCI v4 data are saltier by more than 0.2 pss, this is due to known issues in SMAP v5.0 which have been ingested for CCI v4 generation. This will be corrected for next CCI version.
- Products are in general of better quality after 2015 due to the inclusion of SMAP data and reduced contaminations (RFI, sun) on SMOS data.

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 5 of 63</p>
--	---	--

3 Validation: Data & Methods

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

3.1 Fiducial Reference Measurements

According to the GEO/CEOS Quality Assurance for Earth Observation (QA4EO), Fiducial Reference Measurements (FRM) are a suite of independent, fully characterized, and traceable ground measurements for the validation of satellite SSS. FRM must be characterised (and documented) for the property for which they are a reference; at a level commensurate with the application; temporally stable over the period of use; a value must be SI traceable or community agreed; and must be accompanied by a procedure on use.

In this report, all validation datasets are provided and documented by Pi-MEP [RD03] using community standards with quality control checks. In the following subsection we will define the level of independence against the CCI products.

3.1.1 Level of independence to CCI products

CCI L4v4 products are a combination of SMOS, Aquarius and SMAP products. Each product is adjusted to a base dataset to correct for unresolved bias. This section summarises the various adjustments to base, to define these bases and their underlying *in situ* dataset. For further details, please refer to the ATBD [RD05] and E3UB [RD06].

CCI L4v4 products are adjusted pixel-by-pixel to get the same quantile over the full-time series (~12 years) than the quantile (typically 50% for the median, but in river plumes it goes to 80%, see [RD05] for details) of a base. The base is defined as a combination of ISAS 2017 (for the period 2011-2017) and ISAS NRT (for the period 2018 to October 2022). Adjustments are excluded for the 1st year (in 2010, due to bias in SMOS data at the beginning). There is an additional latitudinal adjustment to correct for seasonal biases based on the same base.


The temporal median over the full period of the base is taken as the prior for the objective analysis used to produce the L4.

SMOS, SMAP and Aquarius L2 products used for the generation of the CCI L4v4 products are also calibrated. SMOS is calibrated in the Eastern Pacific using the same base (ISAS 2017 and NRT). Aquarius and SMAP are calibrated to output from the HyCOM numerical model with assimilation (assimilating Argo among other).

The main base used for calibration has the following differences:

- ISAS 2017: contains the following *in situ* datasets: Argo, CTD (ICES), moorings. It does not contain TSG or drifters.
- ISAS NRT: contains all observations included in Coriolis including TSG and drifters.

These corrections do not influence the interannual or longitudinal variability of the product.

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 6 of 63</p>
--	---	--

3.2 In situ Dataset

Following PVP [RD1] recommendations, the reference dataset used for product validation is based on:

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

The reasoning for this choice of reference dataset is 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, the Argo dataset is described with its collocation criteria along with the gridding method and the method to estimate representativeness errors and validate uncertainties. A summary of the spatial (horizontal) 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 use monthly fields on which observed variability is added, therefore the focus of this summary report is on monthly fields, unless mentioned otherwise. We report in this document on validation against other *in situ* dataset (TSG, drifters, measurements using mammals) which are provided from Pi-MEP. A description of these datasets and full validation of Weekly and Monthly fields are available from:

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

3.2.1 Argo match-ups database: Pi-MEP MDB

The Argo floats used for validation have been taken from Pi-MEP [RD02, RD03] where quality control checks have already been completed. 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 measure oceanic variables including temperature and salinity of the upper 2000 m. This allows continuous monitoring of the temperature and salinity of the upper ocean, with all data being relayed and made publicly available within hours of collection. The array provides around 100,000 temperature and salinity profiles per year distributed over the global open water oceans at an average of 3-degree spacing. These data were collected and made freely available by the international Argo project and the national programs that contribute to it [Argo (2000)]. 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 included in Pi-MEP. All measurements from Argo floats that may have problems with one or more sensors and appear 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).

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 7 of 63</p>
--	---	--

The Argo MDB is produced from the previously described cleaned quality controlled Pi-MEP 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). The spatial window is 12.5 km radius for each node at the centre of the 25 km grid. If more than one satellite pixel meets these criteria then the final satellite SSS match-up point is the closest in time from the *in situ* measurement. The spatial and temporal offsets between the *in situ* and satellite data are stored in the MDB files alongside a wide range of colocalised auxiliary information.

There are three separate Argo MDB products from Pi-MEP:

- the baseline MDB covering the full CCI L4v4 time series;
- two MDB specific for the comparison between CCI L4v3 and CCI L4v4. For these two MDBs, only Argo floats collocated with both CCI v3 and v4 products are used. Given CCI v3 stops in 2021, it is per construction the case for these two MDBs. In v3, there was nearly no data in polar area, hence per construction, in this comparison dataset, there is nearly no data in polar areas for the “MDB v4 comparison dataset”.

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 baseline MDB is used as it is in all Pi-MEP reports, and it will be specified in the text if the context is different from that described in §3.2.2 below.

3.2.2 Reference Dataset: Gridded MDB

Typically, Argo floats sample the water column every 10 days. In order to study some specific processes with strong temporal variability, some Argo floats profile the water column up to every 2 hours. The floats tend to be located in areas where SSS variability is strong and artificially increase the number of collocations in these specific areas. Consequently, it degrades comparisons against satellite based SSS. To solve this issue, following PVP [RD01-section 3.3.5] recommendations, a Monte Carlo approach is used. A sample is randomly selected from each grid cell (biweekly; 25km Equal Area EASE2 grid) and is repeated a multiple number of times (here nine times). All validation metrics are computed for each subsample (e.g. difference satellite and *in situ*). The multiple subsamples are aggregated together by calculating the median. The spatial sampling is subsampled by a factor of seven in both latitude and longitude to ensure sufficient data for a given pixel over the full time series and get significative statistics. Each grid point corresponds to a 175 km Equal Area EASE grid.

This approach is applied for the baseline MDB covering the full CCI L4 v4 time series for gridded MDB and is used as the reference and is referred simply as “v4.4” or “v4” for short. The same method is applied for the two MDB specific for the comparison between CCI L4v3 and L4v4. They are referred respectively as “compv3” and “compv4”.



3.3 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 ΔSSS temporal standard deviation equals the pixel-based 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 ΔSSS 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 ground-truth 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 assuming gaussian distribution and independent errors:

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

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.

In section 4.5.3, we estimate the full representativeness error (aka mismatch error) in space (horizontal and vertical) and time using the 1/12° daily GLORYS numerical reanalysis [RD08].

The vertical representativeness error from Argo floats measurements is discussed in section 4.3.

In section 4.5.1 and 4.5.2, the vertical representativeness error is neglected ($\Delta\sigma_{vertical} = 0$). The time representativeness error, although sometimes important (e.g. river plumes), is not considered ($\Delta\sigma_{time} = 0$). 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 horizontal spatial representativeness error is the only remaining reference uncertainty considered in this



uncertainty assessment. This error is fully described in the PVP [RD07], 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 ($\sim 10,000$ 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:

$$S^2(k_L, k_l) = \iint_{k_L < k < k_l} d\mathbf{k} S(\mathbf{k}) = B \int_{k_L}^{k_l} k dk k^{-2.4}$$

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:

$$DS^2(g, r) \approx S_0^2 \left(\frac{r}{L}\right)^{0.4} \approx S^2(r) \left(\frac{r}{L}\right)^{0.4}.$$

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 temporal standard deviation in time for each grid cell.

3.4 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 assumed 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.

PVP [RD1] recommends to discard points with less than 30 samples. In the following, two approaches have been followed. If one can estimate significance of an hypothesis (e.g. bias), values which are significant at 95% are indicated with dots. For the standard deviation, where there is no related hypothesis, points calculated with less than 30 samples are shown with hatching. The mean and median are considered significant for values higher than 1.96 (at 95%) of the standard errors of mean and median respectively. The confidence interval estimates for the mean and median follow the same approach. The confidence interval estimates for the standard deviation and the robust standard deviation based on IQR use a random resampling of the data (python `astropy.stats.bootstrap` method). 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.

Standard Error of the Mean is estimated following:



**Climate Change Initiative+ (CCI+)
Phase 2**

Product Validation and
Intercomparison Report

Ref.: ESA-EOP-SC-AMT-2021-26

Date: 1/21/2024


Version : v4.0

Page: 10 of 63

$$sem = \frac{\sigma}{\sqrt{N}}$$

Standard Error of the Median is estimated following:

$$sem_{median} = \frac{\sigma}{\sqrt{\frac{2(N+2)}{\pi}}}$$

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 11 of 63</p>
--	---	---

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 4 product. It is compared to the CCI L4 v3 and to the v3 L3 (ascending, descending, combined) products. Section 4.1 describes the accuracy and precision of the products, including case studies at high latitude in North Atlantic and in the Arctic; section 4.2 analyses their stability and section 4.3 analyses the in situ vertical representativeness error. The effective temporal and spatial resolution are assessed in section 4.4. The quality of uncertainty estimation is assessed in section 4.5.

4.1 Accuracy & Precision

4.1.1 Global products validation for L4

SSS are presented in the top panels of Figure 1 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 as 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 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 MDB over the full time period [2010-01 to 2022-10]. There is very good agreement in the resolved patterns between the two fields. In the gridded MDB field, some areas are insufficiently sampled (less than 30 grid points) particularly in some coastal areas affected by river plumes (e.g. Amazon) or enclosed seas (e.g. Caribbean Sea, maritime continent, Mozambique channel), 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 lower subplots of Figure 1 using the robust standard deviation. The high variability regions (e.g. in the vicinity of the Amazon and Congo plumes, Northern Indian Ocean, the ITCZ, or the Gulf Stream) are distinct in both the CCI and gridded MDB. However, the high variability observed at high latitudes (e.g. Brazil-Falkland/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 pronounced at high latitude where the mesoscale is at finer scale than at low latitudes.

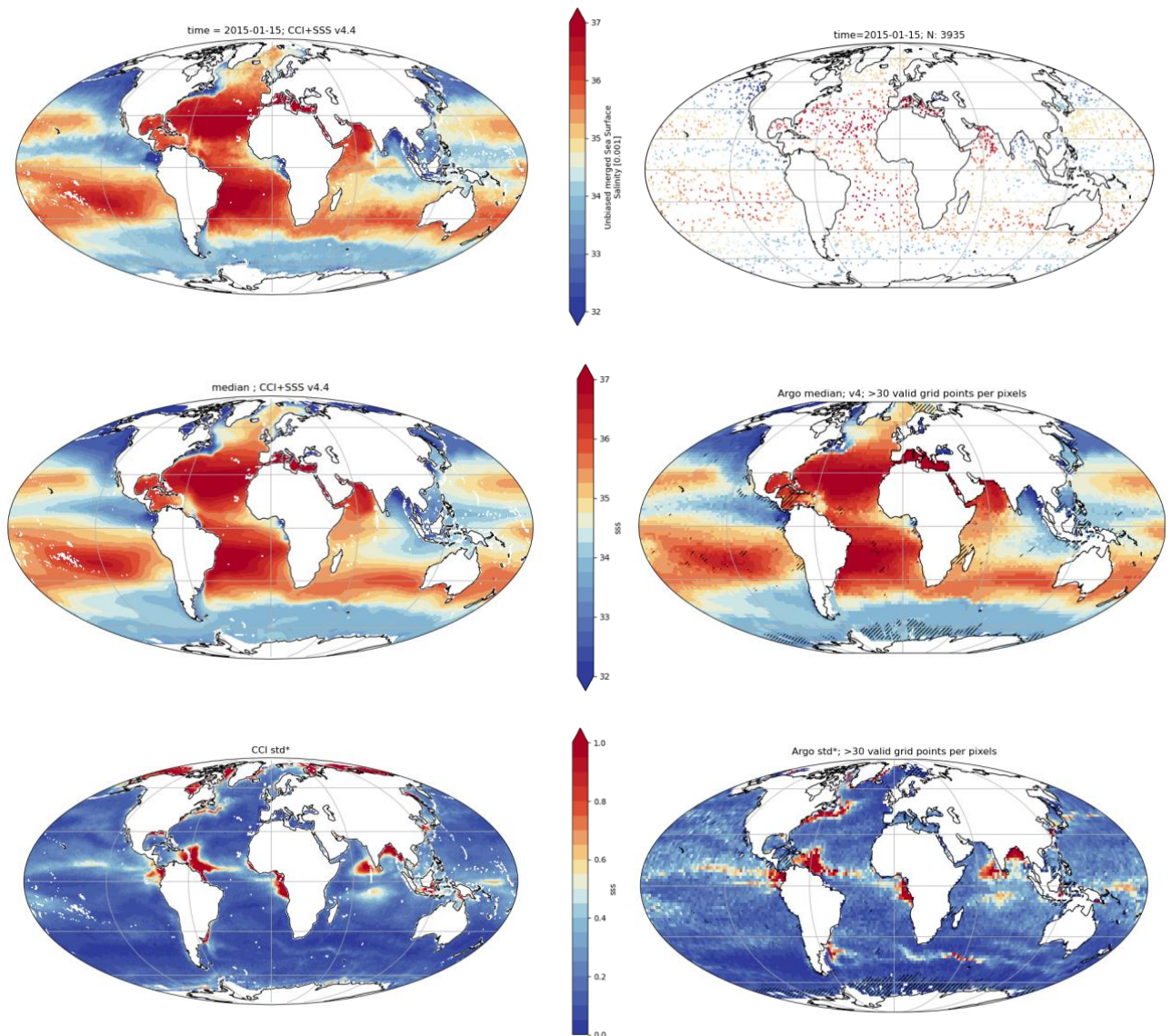


Figure 1: (top-left) L4 CCIv4 monthly SSS for the 15th of January 2015, (top-right) Argo salinity measurements from the Pi-MEP MDB for the same month. (middle/bottom-left) temporal median/robust standard deviation of L4 CCIv4 SSS and (middle/bottom-right) temporal median/robust standard deviation of the gridded MDB SSS. Diagonal hatching indicates regions with less than 30 points over the full time-series.

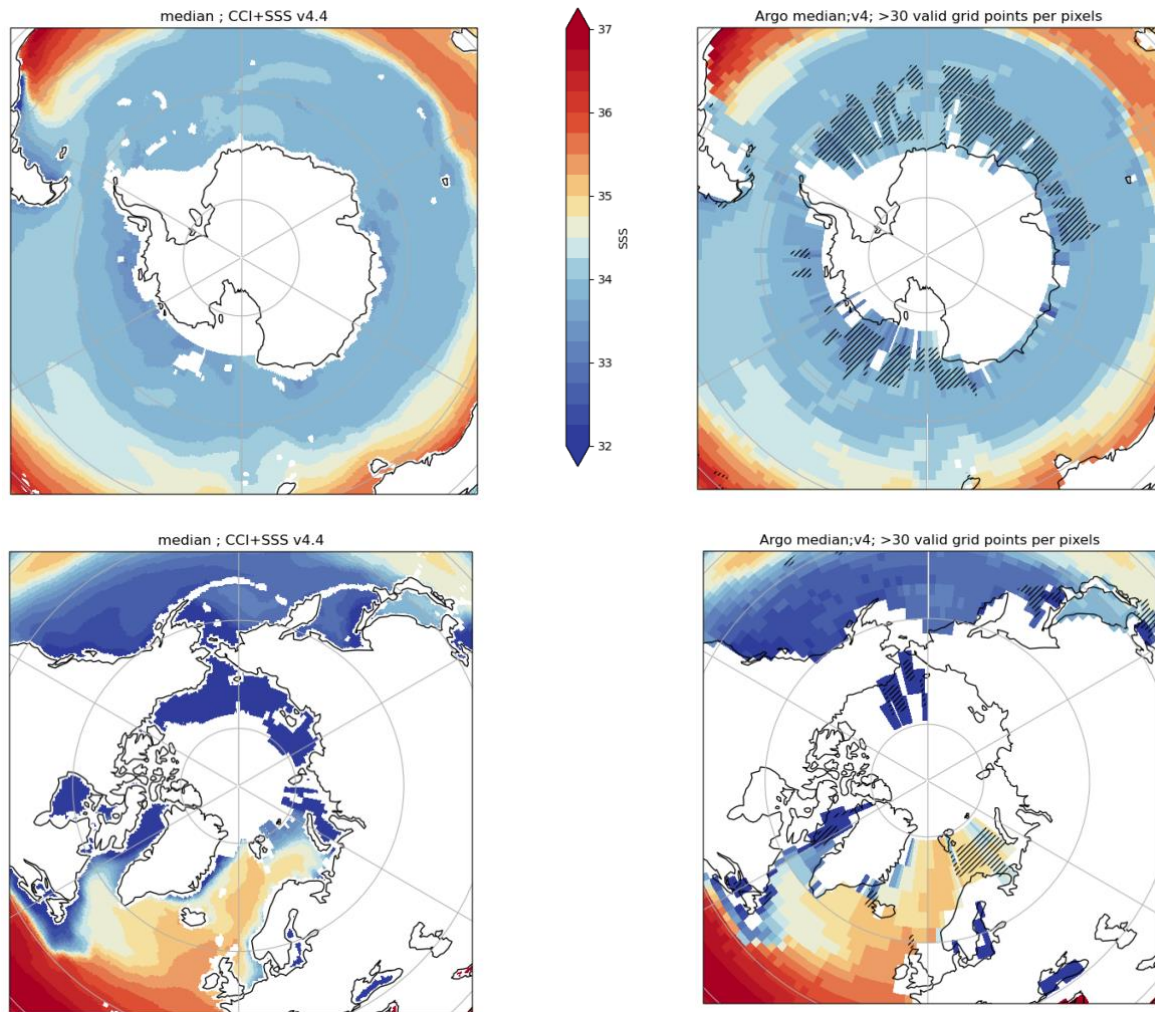


Figure 2 L4 CCIv4 monthly SSS for the 15th of January 2015 for Antarctic (top) and Arctic (bottom), (right) Argo salinity measurements from the Pi-MEP MDB and (left) temporal median/robust standard deviation of L4 CCIv4 SSS. Diagonal hatching indicates regions with less than 30 points over the full time-series.

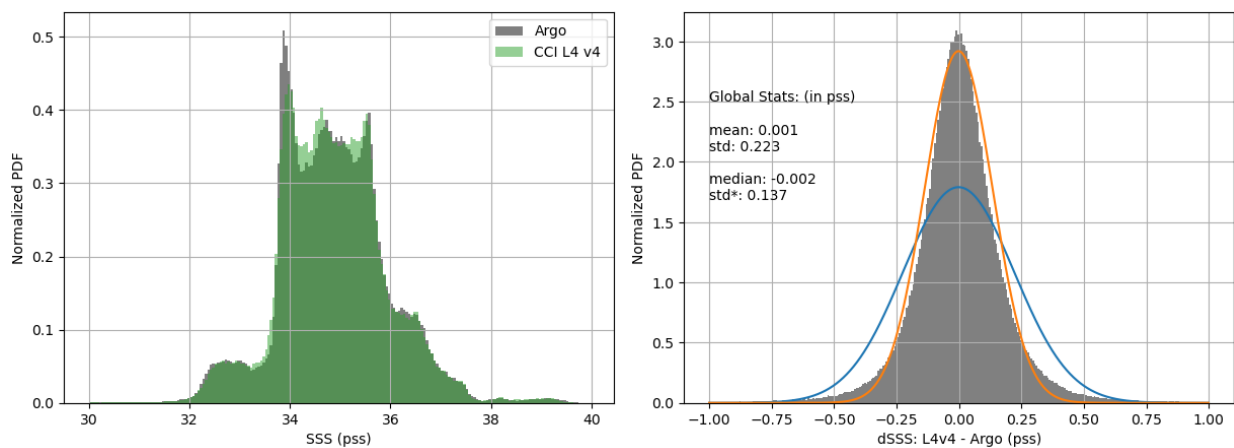


Figure 3: Histogram of all pairwise gridded MDB (left) Argo SSS in grey and CCI L4 v4 SSS in green; (right) Gridded MDB of the CCI L4 v4 minus Argo difference, (blue line) normalized probability function (PDF) using computed mean and standard deviation, (orange curve) normalized PDF using computed median and robust standard deviation.

The distributions of the gridded MDB of pairwise CCI L4 and Argo SSS are very similar (Figure 3-left) over the full range of SSS from 30 pss to 40 pss. The peak in the distribution ~ 34 pss is well resolved for CCI L4 v4 and compared well with the distribution peaks of Argo. The distribution of the gridded MDB of pairwise CCI difference against Argo (Figure 3-right) highlights the absence of bias ($< |0.01|$ pss), and a robust standard deviation of 0.14 pss (0.22 pss for the classic standard deviation). The difference between the robust and classic standard deviations 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 0.97.

A comparison of Pi-MEP using the exact same Argo floats colocation for v4 and v3 Monthly products are summarised in Table 3. V4 is generally 0 to 0.01 pss lower than v3 with a small bias low. There is an improvement on nearly all dispersion metrics (robust and classic standard deviation) and all correlation indicated in green in Table 3. There is particular improvements for three specific conditions:

- [C7a: distance to coast < 150 km] with particular improvements on the classic standard deviation;
- [C8a: SST $< 5^\circ\text{C}$] with particular improvements on the classic standard deviation and the square correlation (explained variance);
- [C9c: SSS > 37 pss] with particular improvements on the robust standard deviation and the square correlation coefficient (explained variance), despite a degradation on the classic standard deviation.

The explained variance (R^2) is very high with values around 0.97 for the two versions. The fraction of variance unexplained ($FVU = 1 - R^2$) improves by 10% between v3 and v4 when considering all the colocations between v3, v4 and the Argo floats. The FVU increases by more than 25% for the three conditions highlighted above ie C7a, C8a and C9c. For C8b (SST in $[5-15]^\circ\text{C}$), C5 (WOA2013 $\text{std} < 0.2$ pss) and C9b (SSS in $[33-37]$ pss) FVU increases by more than between 10% to 20%.

Table 3: Statistics for Pi-MEP MDB comparison dataset for v3 and v4 using exactly the same Argo floats colocation. Cf Pi-MEP reports for the details about the conditions criteria.

Condition	Nobs	Median		Mean		Std*		std		R2	
		V3.2	V4.4	V3.2	V4.4	V3.2	V4.4	V3.2	V4.4	V3.2	V4.4
all	626618	0.00	-0.01	0.00	0.00	0.15	0.15	0.22	0.21	0.968	0.971
C1	221560	0.00	-0.01	0.00	0.00	0.13	0.12	0.15	0.15	0.973	0.974
C2	428601	0.00	-0.01	0.00	-0.01	0.14	0.14	0.21	0.20	0.973	0.974
C3	5605	0.02	0.02	0.02	0.02	0.19	0.19	0.25	0.25	0.956	0.957
C5	486004	0.00	-0.01	0.00	-0.01	0.14	0.13	0.19	0.17	0.976	0.979

C6	117400	0.00	0.00	0.00	0.00	0.22	0.22	0.34	0.33	0.946	0.948
C7a	35578	-0.02	-0.02	-0.03	-0.02	0.22	0.22	0.39	0.35	0.981	0.985
C7b	257766	0.00	-0.01	0.00	0.00	0.17	0.16	0.26	0.25	0.955	0.959
C7c	332704	0.00	0.00	0.00	0.00	0.13	0.13	0.16	0.16	0.971	0.973
C8a	48268	0.01	0.00	0.01	0.00	0.17	0.17	0.28	0.24	0.683	0.750
C8b	145161	0.01	0.01	0.01	0.01	0.15	0.14	0.23	0.21	0.959	0.965
C8c	432387	-0.01	-0.01	-0.01	-0.01	0.15	0.14	0.21	0.21	0.968	0.970
C9a	27831	0.05	0.05	0.08	0.08	0.20	0.19	0.40	0.40	0.968	0.968
C9b	574536	0.00	-0.01	0.00	-0.01	0.15	0.14	0.21	0.20	0.940	0.946
C9c	24251	-0.04	-0.05	-0.05	-0.06	0.14	0.15	0.23	0.20	0.911	0.929

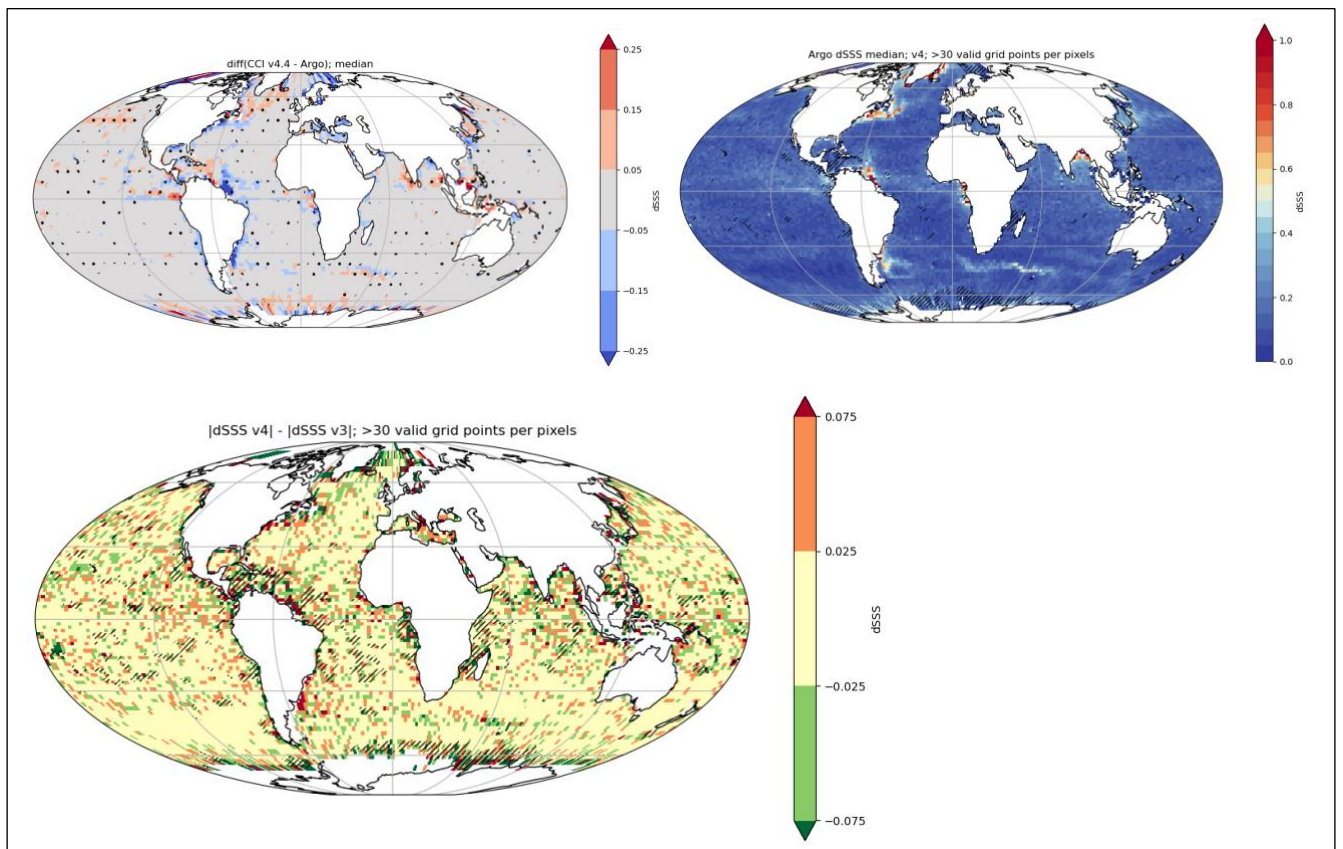


Figure 4 (top-left) Temporal median and (top-right) temporal robust standard deviation of gridded pairwise SSS differences between CCI and Argo. (on the left) A moving window of 2 pixels in longitude and latitude is applied to the median in order to highlight statistical significance (at 95%) which are indicated with dots. (on the right) Robust standard deviation calculated with less than 30 valid grid points in time are hatched. (bottom) differences of the temporal median differences between CCI and Argo between version 4 and version 3 using the comparison dataset.



To further assess the agreement between datasets, Figure 4 presents the temporal average (median) and the robust standard deviation of the gridded MDB v4 for the differences between CCI and Argo. At large scale (open ocean), the median difference is within ± 0.05 pss and the robust standard deviation difference is below 0.2 pss. There is no large scale systematic spatial difference versus Argo. Small portions of the North Pacific present significant salty bias higher than 0.05 pss. The other significant differences are in the vicinity of strong SSS gradients or strong current. Close 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 measurement versus 50 km 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, ...). , Figure 4-bottom represents the improvements in green and degradation in orange of the median difference against Argo between version 4 and version 3, using the gridded MDB comparison dataset (compv4 and compv3). This figure highlights there is no systematic and large-scale difference and most of the ocean appears yellow, i.e. no difference higher than ± 0.025 pss.

Pi-MEP statistics of CCI L4 v4 against TSG, drifters and mammal's data with the C1 criterion are reproduced in Table 4. As highlighted in section 3.1.1, the time series of CCI SSS in each pixel is calibrated by comparing a quantile of the statistical distribution of CCI SSS with the one of ISAS SSS (ISAS 2017 up to 2017 and ISAS NRT afterwards). Whereas TSG and drifters data are included in ISAS NRT, they are fully independent from ISAS 2017. Statistics of CCI data against TSG and drifters are in the same range as for Argo. Whereas some datasets present biases in absolute term higher than 0.05 pss (in order: saildrone, tsg-ncei), tsg-legos-dm, tsg-gosud, tsg-polarstern, and drifters are within ± 0.01 pss. The robust standard deviation of the difference between CCI and in situ measurements is the same between Argo and Drifters (0.12 pss) but stay close to the Argo performance (< 0.15 pss) for tsg-legos-dm, tsg-gosud, tsg-legos-survostral and tsg-polarstern. The explained variance (R2) is higher than Argo for, in order, drifters, tsg-polarstern, tsg-gosud-sailing-ship, tsg-legos-dm confirming the good performance of CCI L4v4 against independent measurements.

Table 4: Statistics of CCI L4 v4.4 30dr against in situ data for the global ocean applying criteria C1 (only pairs where $RR=0$ mm/h, $3 < U < 12$ m/s, $SST > 5^\circ C$, distance to coast > 800 km). From Pi-MEP

	Nobs	Median	Mean	STD	RMS	R2	STD*
argo	361151	0	0	0.16	0.16	0.973	0.12
tsg-legos-dm	632189	-0.01	-0.01	0.17	0.17	0.975	0.14
tsg-gosud-research-vessel	309804	0	0.01	0.17	0.17	0.972	0.13
tsg-gosud-sailing-ship	187451	-0.01	0	0.15	0.15	0.979	0.13
tsg-samos	1429283	0.03	0.07	0.3	0.31	0.912	0.16

mammal	13347	0.03	0.04	0.2	0.2	0.889	0.17
drifter	753755	-0.01	0.02	0.22	0.22	0.984	0.12
tsg-legos-survostral	52500	0.03	0.02	0.14	0.14	0.584	0.13
tsg-ncei-0170743	113600	0.09	0.09	0.17	0.19	0.918	0.16
tsg-polarstern	38867	0	0	0.15	0.15	0.981	0.13
saildrone	122273	0.14	0.19	0.37	0.42	0.765	0.28
tsg-csic-utm	162347	0.02	0.02	0.21	0.21	0.948	0.16
ices	109	-0.03	0.09	0.41	0.42	0.475	0.11

4.1.2 Global products validation differences for Aquarius, SMAP, SMOS L3 products and weekly L4 products

In this subsection we will look at differences between CCI L3C v4 Aquarius, SMAP or SMOS products and CCI L4 v4. L3C products are simple grid averages of individual L2 satellite SSS after applying systematic corrections. Globally, using all collocation with Argo or only the ones corresponding to the best-case criteria (C1: only pairs where $RR=0\text{mm/h}$, $3 < U < 12\text{m/s}$, $SST > 5^\circ\text{C}$, distance to coast $> 800\text{km}$), there is no systematic bias (bias lower than 0.03 pss in absolute value, see Table 5 and Table 6). Spatially, there is the presence of biases, represented in Figure 5, with e.g. positive bias in the Southern Ocean for Aquarius and SMAP, whereas it is negative for SMOS; SMAP and SMOS present negative biases in the Mediterranean and Nordic seas; in the Eastern Pacific SMAP tends to be bias positive whereas Aquarius is bias negative.

In term of global statistics (Table 5 and Table 6), performances of the dispersion against Argo (std and std*) are very similar for the CCI L4v4 7 days, 30 days and L3C Aquarius. Performances are lower for SMOS than SMAP if we take all the collocations (Table 5) but the performances are inverted if collocations are selected for the best-case criteria (C1, see Table 6).

Table 5: Statistics of CCI L4 and L3C (SMOS, Aquarius, SMAP) products against Argo data for the global ocean. From Pi-MEP.

Satellite	Nb	median	mean	std*	std	R2
Weekly running average						
cci-l4-esa-merged-oi-v4.4-7dr	1158756	0	0	0.15	0.23	0.98
Monthly running average						
cci-l4-esa-merged-oi-v4.4-30dr	1139048	0	0	0.15	0.24	0.975
cci-l3c-esa-aquarius-v4.4-30dr	329171	0.01	0.01	0.16	0.24	0.957
cci-l3c-esa-smap-v4.4-30dr	679071	-0.02	-0.01	0.28	0.38	0.969
cci-l3c-esa-smos-v4.4-30dr	1113950	-0.01	-0.03	0.27	0.44	0.897

Table 6: Statistics of CCI L4 and L3C (SMOS, Aquarius, SMOS) products against Argo data for the global ocean applying criteria C1 (only pairs where $RR=0\text{mm/h}$, $3 < U < 12\text{m/s}$, $SST > 5^\circ\text{C}$, distance to coast $> 800\text{km}$). From Pi-MEP.



**Climate Change Initiative+ (CCI+)
Phase 2**

Product Validation and
Intercomparison Report

Ref.: ESA-EOP-SC-AMT-2021-26

Date: 1/21/2024

Version : v4.0

Page: 18 of 63

Satellite	Nb	median	mean	std*	std	R2
Weekly running average						
cci-l4-esa-merged-oi-v4.4-7dr	367514	0	0	0.13	0.15	0.974
Monthly running average						
cci-l4-esa-merged-oi-v4.4-30dr	361151	0	0	0.12	0.16	0.973
cci-l3c-esa-aquarius-v4.4-30dr	106932	-0.01	0	0.13	0.15	0.975
cci-l3c-esa-smap-v4.4-30dr	213007	-0.01	0	0.23	0.25	0.928
cci-l3c-esa-smos-v4.4-30dr	359593	0	0	0.21	0.24	0.938

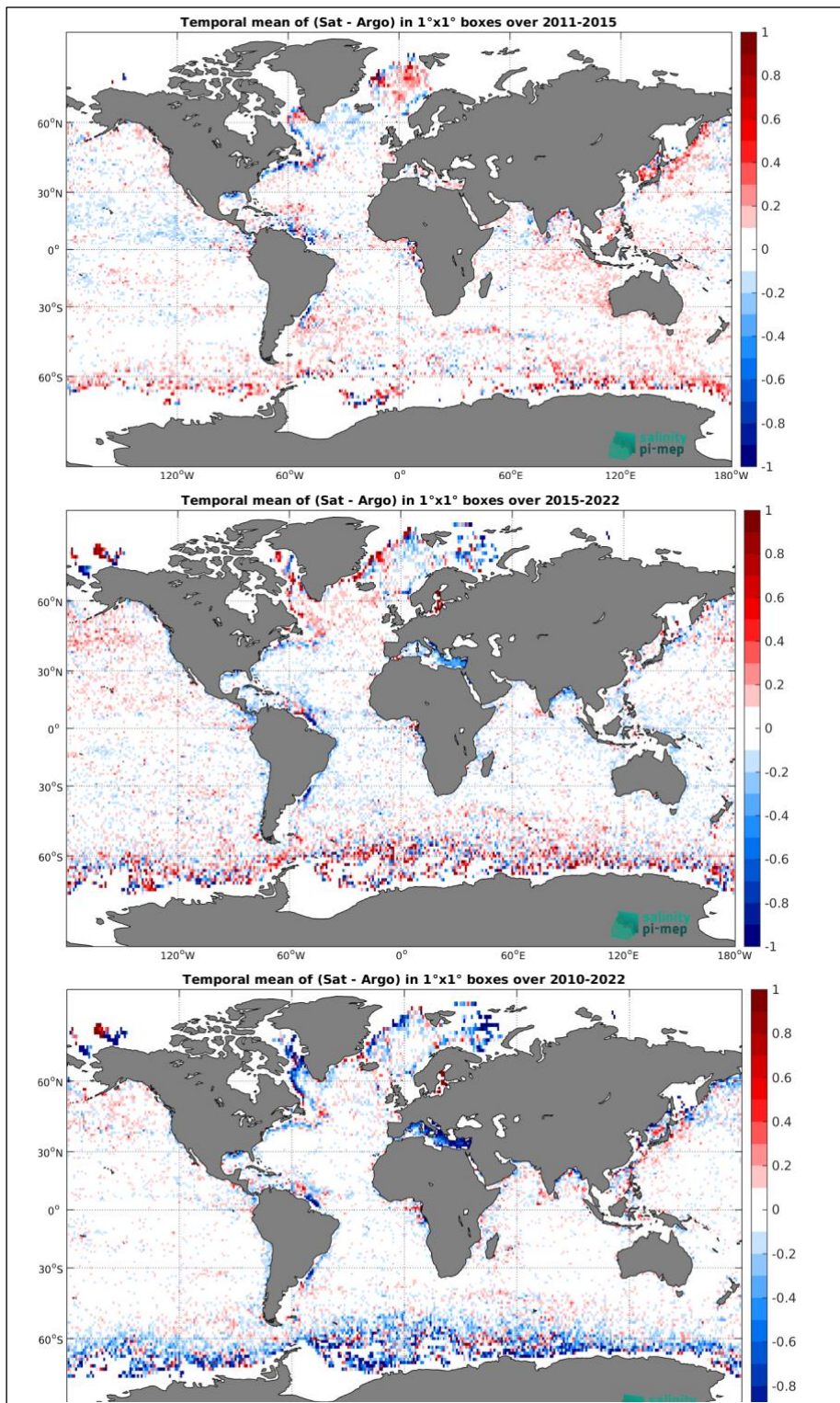


Figure 5: Temporal mean of the difference satellite – Argo for (1st row) Aquarius; (2nd row) SMAP; (3rd row) SMOS. Taken from Pi-MEP reports.

4.1.3 North Atlantic & Arctic Case Studies

The following sections (§4.1.3.1 - §4.1.3.3) provide three independent case studies focussing on specific regions within the Arctic. The first focuses at the region around Southern Greenland; the

second at using ICES database (in the Nordic and Barents Seas) data base and Saildrones (Beaufort and Chukchi Seas); and the third considers river input in the Laptev Sea. Polar regions are especially difficult for satellite measurement due to both sea ice contamination (signal to noise ratio) and the lower sensitivity of L-band in cold waters.

4.1.3.1 Arctic Case Study: Vicinity of Southern Greenland (G. Reverdin, F. Bonjean, L. Kilian, J. Boutin, LOCEAN, J.-L. Vergely, ACRI-ST)

Relevance for climate studies

Remote sensing observations as well as in situ data suggest a large variability in the Labrador Sea at the mesoscale with outflow events from the Greenland shelf that might affect oceanic convection and formation of intermediate and deep waters in the key region of the North Atlantic subarctic. A large variability is also observed on the Greenland shelf, which is poorly monitored and might be influenced both by export of Arctic water and freshwater melt from Greenland icesheet. This is a challenging area for validation not just due to the issues discussed above but also the influence of the icy Greenland land mass and intermittent radio frequency interference (RFI; prior to 2015).

To complement the data gathered from Pi-MEP, we accessed and validated other data sources: additional TSG data from German and Canadian research vessels (Meteor, MSM, Amundsen), Danish coastguard vessel (Hauge Koch); data from salinity drifters (from NIOZ and WHOI); from seals deployed near Newfoundland (Canada); as well as A-XCTD and CTD surveys from OMG and glider data (NOC). We also corrected or removed some biased data (mostly TSGs) present in the Pi-MEP archive. This resulted in a large increase of data in particular close to Greenland south of 70°N and nearby waters (cf the example of data in 2021 below). Most validated and adjusted data were transmitted to Pi-MEP, and additional data (not yet incorporated or public, such as drifters deployed in 2022) will be later included.

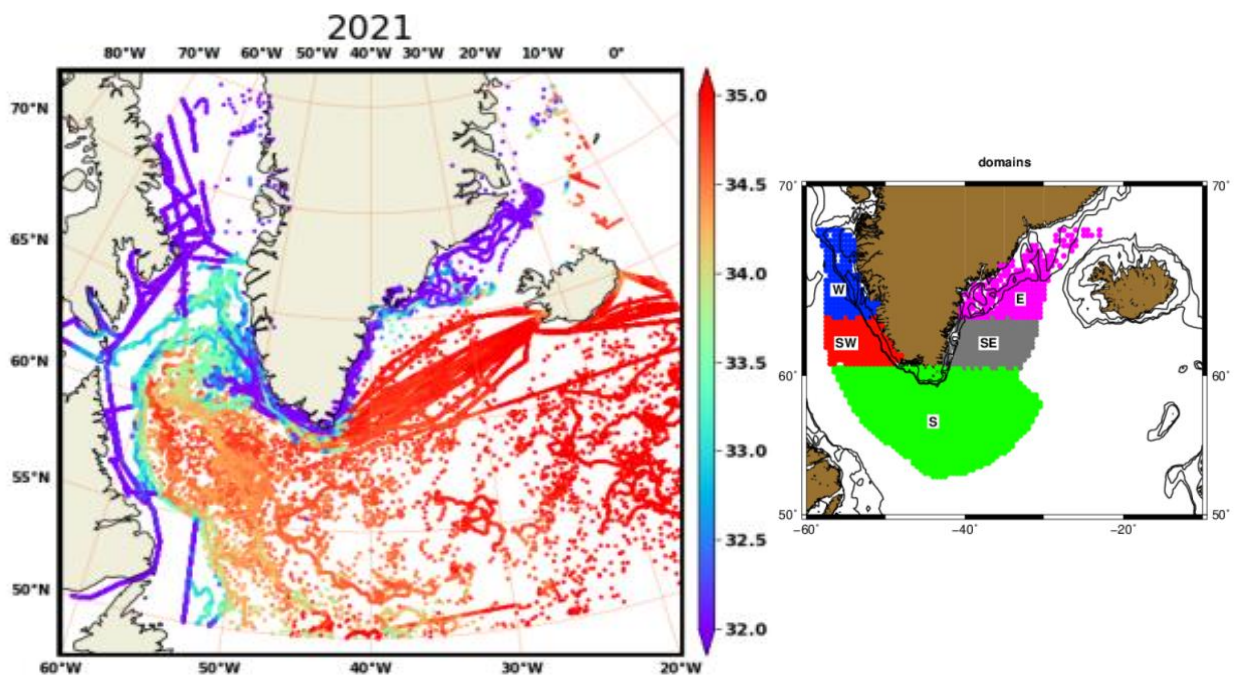


Figure 6 (left) Distribution of data in 2021 (colour represents SSS). (right) sub-domains selected as a function of hydrography and shelf break distance to coast.

The resulting coverage in 2010-2022 is sufficient to leave very few ocean areas on a 50 km scale devoid of data, albeit often with very poor temporal sampling. Notice also that sampling on shelves is mostly during late spring to late fall (except along southwestern Greenland, where it takes place all year round).

This nonetheless allows some attempts to establish statistics based on collocations between spatially and time reduced *in situ* data and measurements from the CCI V4.4 product. We present these statistics only when there are at least 3 collocations. To have a meaningful comparison, we averaged the co-located data over time and space. Based on what we know of the circulation and shelf break distance, we separated the area around Greenland into different domains with distance from the coast (and shelf break) as the main criteria. Time series in each of these sub-domains were then produced (we only retain values when they are at least two independent collocated data in a given bin).

Figure 7 illustrates the possibility to create time series of collocated data that seem reasonable, but present large biases (right panel) close to the coast that originate from the climatological fields used to adjust the mean of V4.4 over 2010-2022. There are also weak systematic (negative) biases just offshore of the shelf-break as well, and in the interior of the gyres (both Labrador and Irminger Seas) which result from the spatial smoothing of the ISAS product.

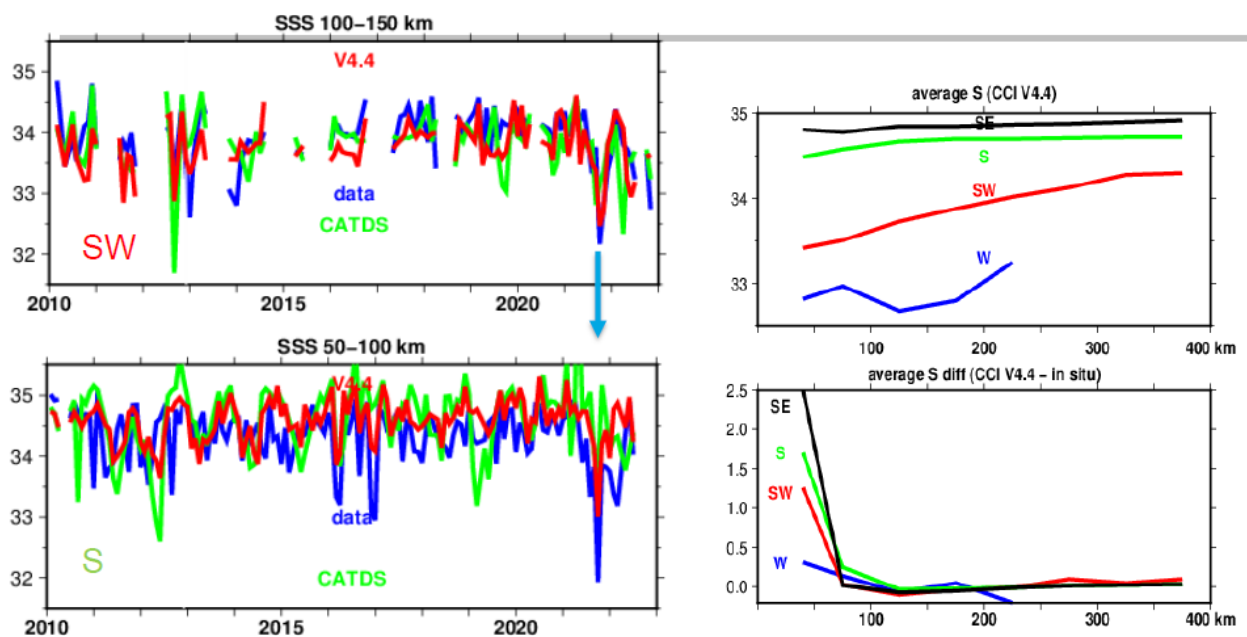


Figure 7 left time series of salinity in different bins as a distance to coast in (up left) south-west Greenland, and bottom (down left) south Greenland (blue *in situ* data; red V4.4; green a SMOS only LOCEAN products). Both are off the shelf (the lower closer to shelf break). Notice very low salinity in September-October 2021 on both. To the right, average profiles as a function of distance to coast (four sub-regions: West, South-west, South, and South-east Greenland). Upper panel average salinity of V4.4; lower panel, average difference CCI-*in situ*.

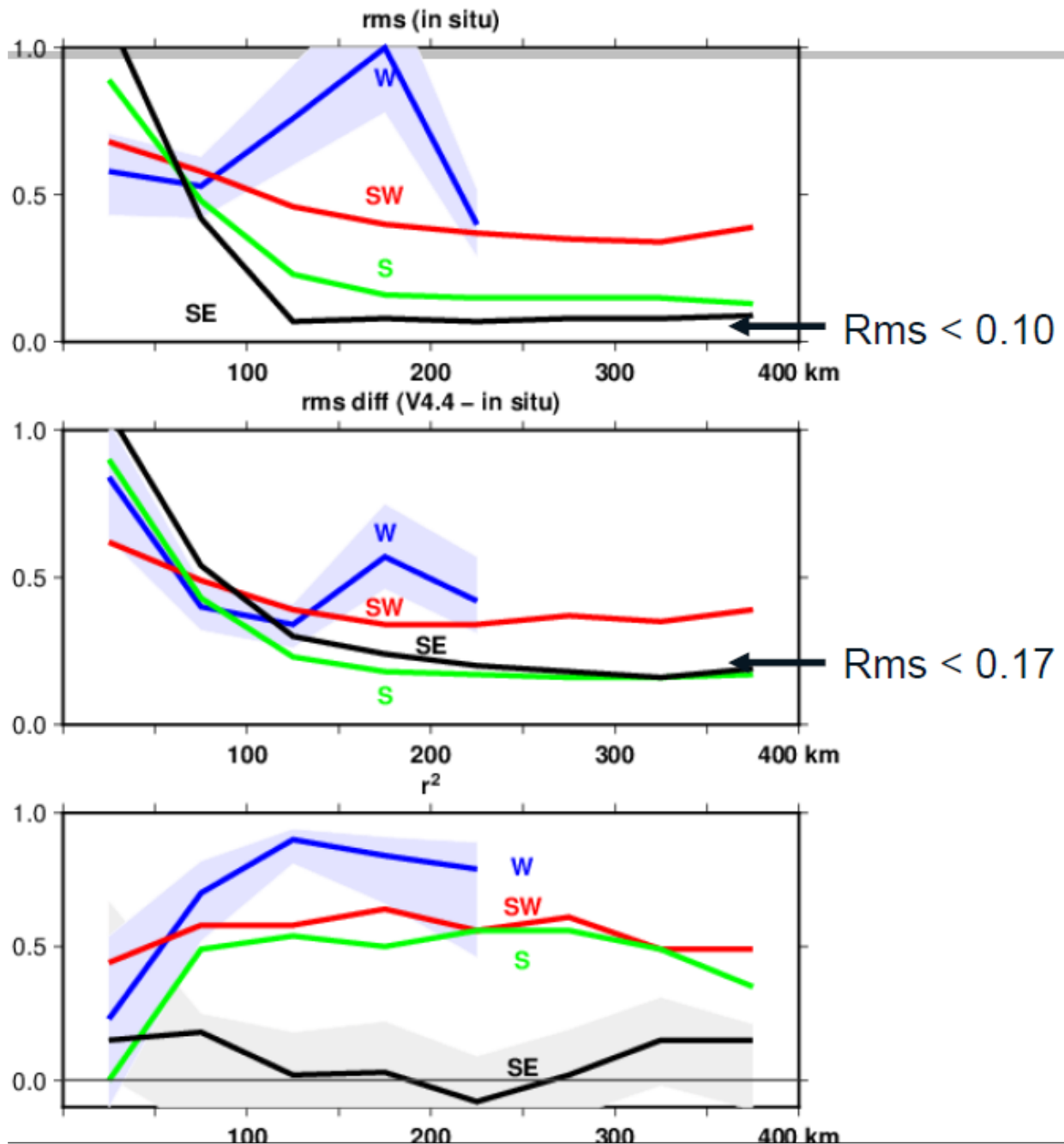



Figure 8 Statistics of the time series as a function of distance to coast for four different sub-regions. Top panel root mean square difference of the in situ time series illustrating usual large decrease from shelf to ocean interior, except in W where the largest values are found in the middle of Denmark Strait (150-200 km to coast) due to the presence of large water mass fronts. Middle panel, rms differences between V4.4 and in situ. Lower panel, squared Pearson correlation coefficient between V4.4 and in situ.

We investigated the statistical properties of the collocated time series presented in Figure 8 (left). In these statistics, the rms differences are never less than 0.15-0.17 (in the interior gyres South and South-east Greenland). Such values are expected from the errors in the product (but a full investigation of the error budget has not been undertaken), and their relative low values indicate that the product has been able to mitigate the worst RFI data that were plaguing earlier versions in particular south or south-east of Greenland, as well as rather well merging data of the different satellite missions. In SE (Irminger Sea), this is still larger than the actual variability portrayed by

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 23 of 63</p>
--	---	---

in situ data. Thus, it results in very low (non-significantly differing from 0) correlations in this region. Elsewhere, correlations between collocated V4.4 and *in situ* time series are significant. We also found that positive correlation coefficients increase (and rms differences decrease) if one selects only the ‘warmer/ice free’ season (June to November) and even more so if the comparison is limited to the period since June 2015 when there are both SMOS and SMAP data. The regional data set on the other hand is too limited to identify any possible bias in May-September 2015 associated with SMAP data.

4.1.3.2 Arctic Case Study: Comparison with ICES (Nordic Seas and Barents Sea) data base and Saildrones (Beaufort and Chukchi Sea) - Nicolas Kolodziejczyk, Alexandre Supply, Jean-Luc Vergely, Jacqueline Boutin

Relevance for climate studies

The Nordic and Barents Seas are key regions for the Atlantic Overturning Circulation (AMOC) branch entering the Arctic region. The warm and salty water coming from the subtropical Atlantic are transformed by intense air-sea fluxes and mixing to flow below the fresher and cold Arctic layer north of the Fram Strait and the Polar Front in the Barents Sea. The Barents Sea is a region for ice export from the Nansen Basin associated with strong freshwater input north of the Polar Front. Under climate change, the Barents Sea is one of the hot spots of rapid change in Arctic Ocean conditions. Detection of SSS signatures from melting and SSS gradients from satellite L-Band is also challenging because of the narrow range of SSS in these regions (~1-2 pss) in addition to the cold water and sea ice contamination discussed above. In this region, SSS variability is poorly monitored and understood from intraseasonal to interannual scales, and at the mesoscale.

In the Canadian basin, the Beaufort Gyre is the largest reservoir of freshwater in the surface layer of the Arctic contrasting with the much more saline waters observed on the neighbouring Chukchi Plateau. The seasonal formation and retreat of the sea ice produce also contrasts with SSS features with a strong impact on surface air-sea interactions and ocean surface dynamics. Monitoring the SSS at intra-seasonal scale and mesoscale is crucial to better understanding of the freshwater cycle in the Arctic Ocean.

Validation Dataset

In this case study, the CCI+SSS monthly products v4.4 and res1_nows (res1_nows being an exploratory dataset for the Arctic) projected on the Northern Polar grid are compared with data from ICES CTD database (<http://www.ices.dk/>). The Argo data comparison is already implemented in the framework of Pi-MEP. The UDASH data present few data in the sea ice free regions (Figure 9). In addition, comparisons are provided in the Beaufort Gyre and Chukchi Sea from two Saildrones operated during Summer 2019 (previously described in Supply *et al.*, 2022).

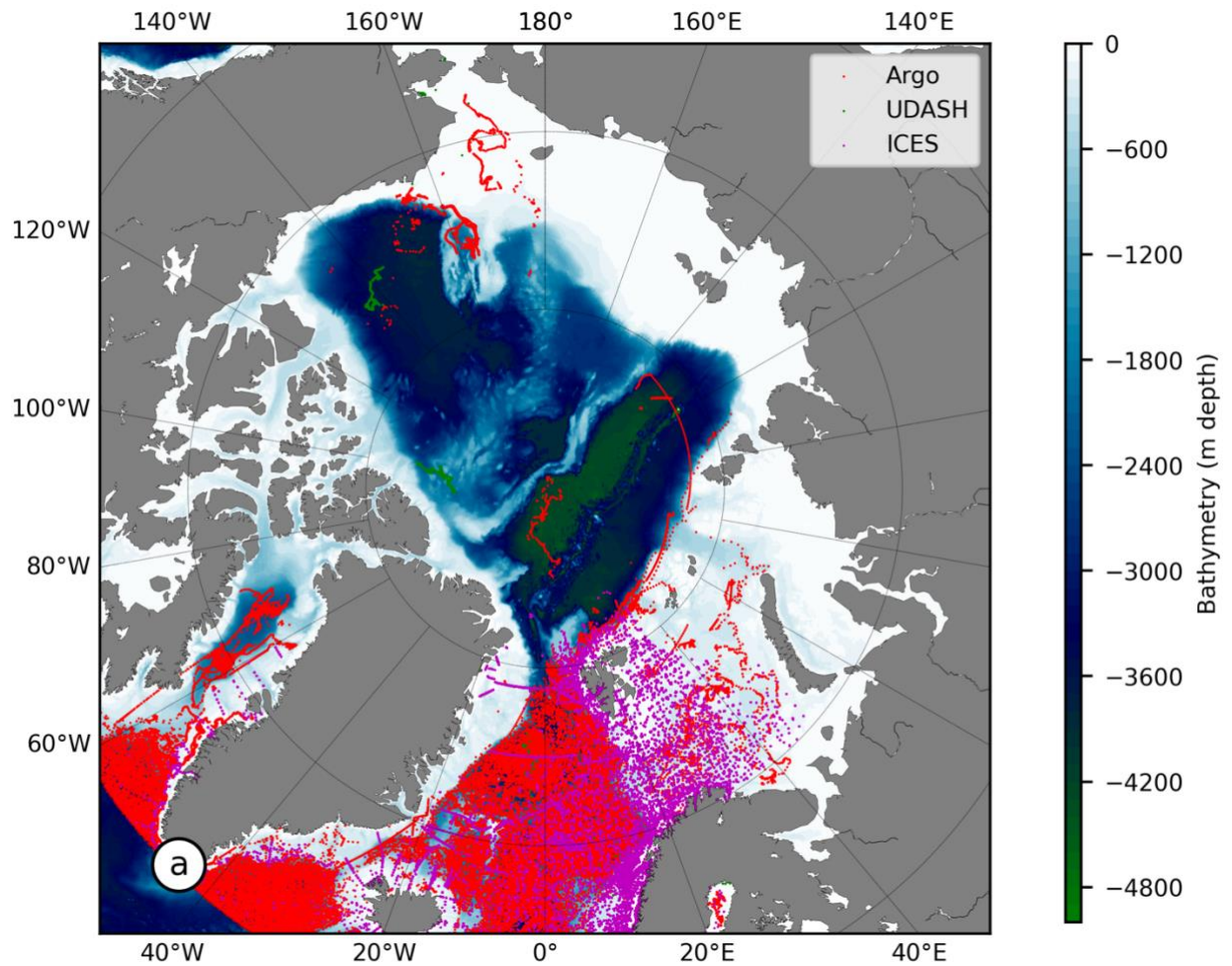


Figure 9 Distribution of CTD profiles between 2010 and 2022 from ICES data base (magenta, www.ices.dk), UDASH (green, <https://glodap.info/>) and Argo (red, <https://argo.ucsd.edu/>).

Validation in the Atlantic Sector of the Arctic

In the Atlantic Sector of the Arctic, the direct comparison of the monthly CCI+SSS v4.4 (res1_nows) product with SSS from the ICES *in situ* dataset shows a correlation around 60% (51%). In comparison to res1_nows product, the v4.4 version of the CCI product shows less scatter (Figure 10, top row) and a reduced bias for low salinities (below 34 pss ; Figure 10, middle row) and during the summer (Figure 10, top row). The best agreement between CCI+SSS products is reported from June to November with correlations larger than 60%. In spite, of the large SSS variability during this period due to sea ice melt and the associated intensification of the freshwater flux, the standard deviation of the difference is lower for the v4.4 products (~0.5-0.6 pss) compared to the res1_nows (~1 pss). This reflects an effective mitigation of the noise in v4.4 products.

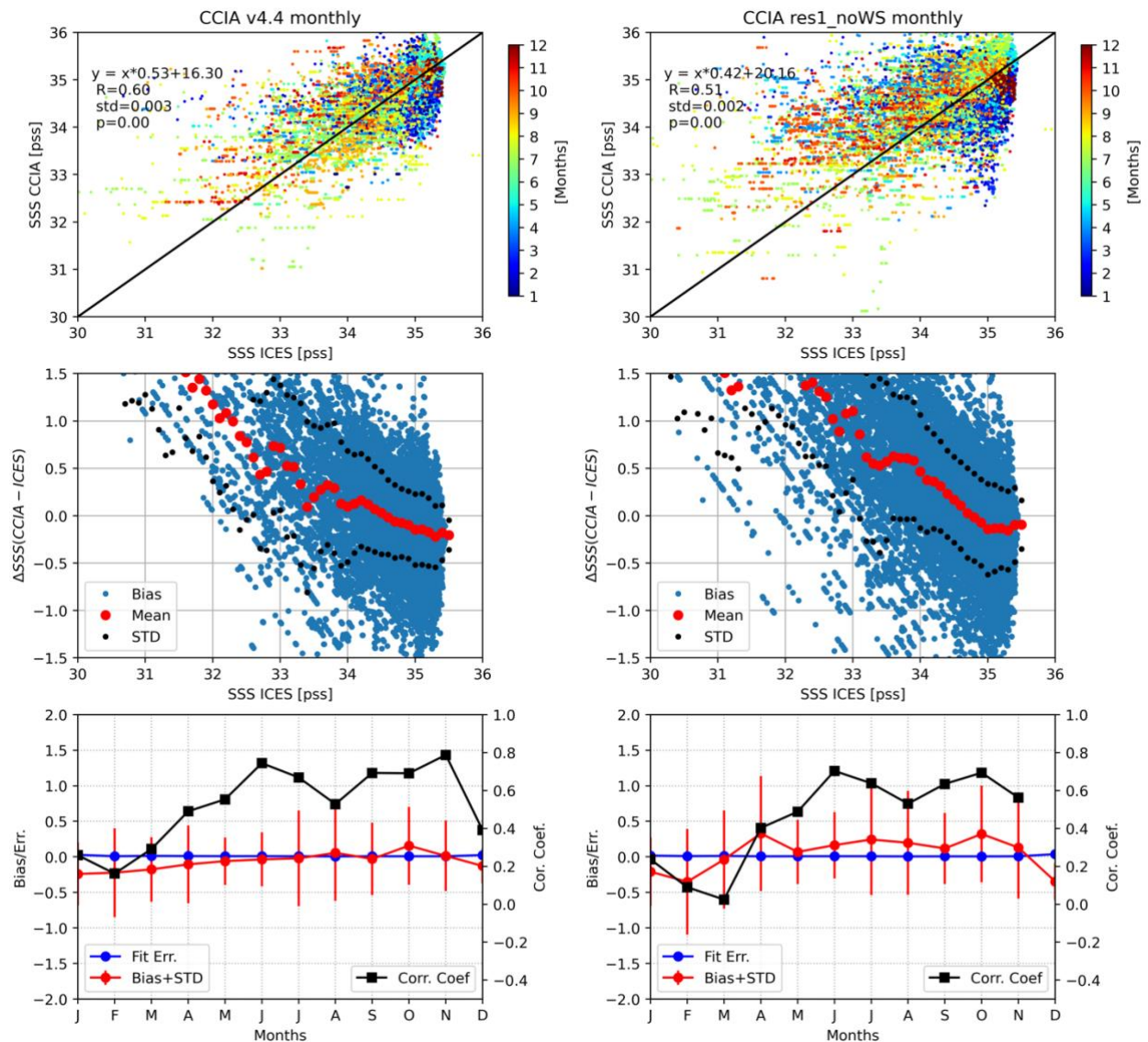


Figure 10 Comparison between SSS from ICES CTD database and CCI+SSS v4.4 (left), and res1_nows (right). (top) Scatter of SSS collocations as a function of the month around the year (colour). (middle) scatter of SSS differences as a function of SSS with mean (red dots) and standard deviation envelope (black dots). (bottom) Monthly values of linear fitting error between satellite product and in situ data (blue), bias and STANDARD DEVIATION of the difference (red) and correlation (black).

Validation in the Canadian Basin

In the Canadian Basin, the comparison between CCI+SSS products and Saildrone SSS shows good agreement to reproduce the dynamics of the SSS signal seen by the Saildrones. In particular, events around day 18 for Saildrone 1036 and day 38 for Saildrone 1037. In contrast, close to the sea ice edge and between days 80 and 120, the v4.4 product appears biased toward fresher. The res1_nows appears slightly positively biased during this period, but better represents the variability over the sailing period next to the ice edge.

Saildrones SSS

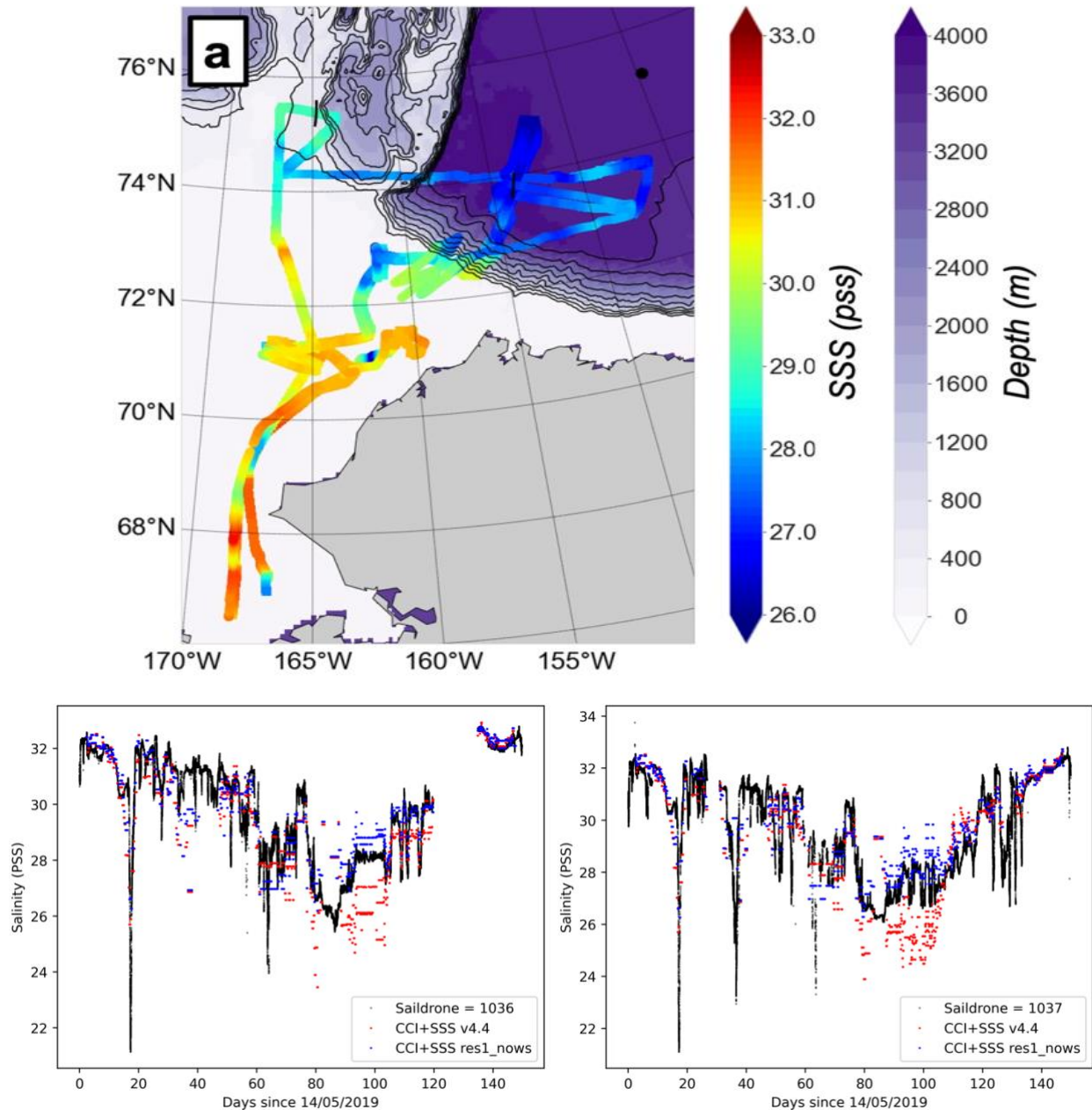


Figure 11 (upper panel) Tracks of Saildrones 1036 and 1037 in the Chukchi Sea and Beaufort Gyre (Supply et al., 2022). (lower panel) Timeseries of SSS from CCI+SSS v4.4 (red), res1_nows (blue) and Saildrone (black) 1036 (left) and 1037 (right).

In the Canadian Basin, the comparison of SSS from Saildrones compared to CCI products shows better performance of the res1_nows configuration with RMSD around 1.00 (RMSD~1.3 for the v4.4), smaller bias, in absolute value, around 0-0.25 (bias~0.5 for v4.4), but with comparable correlations ($R^2=0.85$). This suggest slightly noisier and more biased v4.4 SSS in this region during summer 2019 (Table 7).

Table 7 Statistics (STANDARD DEVIATION of the differences, bias, RMSD and R²) from comparison of CCI+SSS products v4.4 and res1_nows with Saildrones 1036 and 1038.

	Saildrone 1036		Saildrone 1037	
	V4.4	res1_nows	V4.4	res1_nows
STD_diff	1.13	1.00	1.26	1.03
Bias	-0.55	-0.01	-0.41	+0.25
RMSD	1.26	1.00	1.33	1.06
R²	0.85	0.85	0.85	0.85

4.1.3.3 Arctic Case study: Interannual variability in the Laptev Sea (Phoebe Hudson, Adrien Martin)

Eurasian Rivers provide a quarter of total fresh water to the Arctic, maintaining a persistent fresh layer that covers the surface of the Arctic Ocean. This freshwater export controls Arctic Ocean stratification, circulation, and basin-wide sea ice concentration. The Lena River supplies a large volume of runoff and plays a key role in this system, as runoff outflows into the Laptev Sea as a particularly shallow plume. Hudson et al. (2023) confirms zonal wind is the dominant driver of the interannual variability of the path of the Lena river plume, hence the SSS in the Laptev Sea. Eastward winds confine the plume nearshore and westward winds drive the plume offshore. Comparison with the rare *in situ* data gives better agreement against satellite-based SSS measurements (correlation coefficient: $r > 0.8$) than reanalysis products ($r > 0.7$). Please refer to Hudson *et al.* (2023) for full product references.

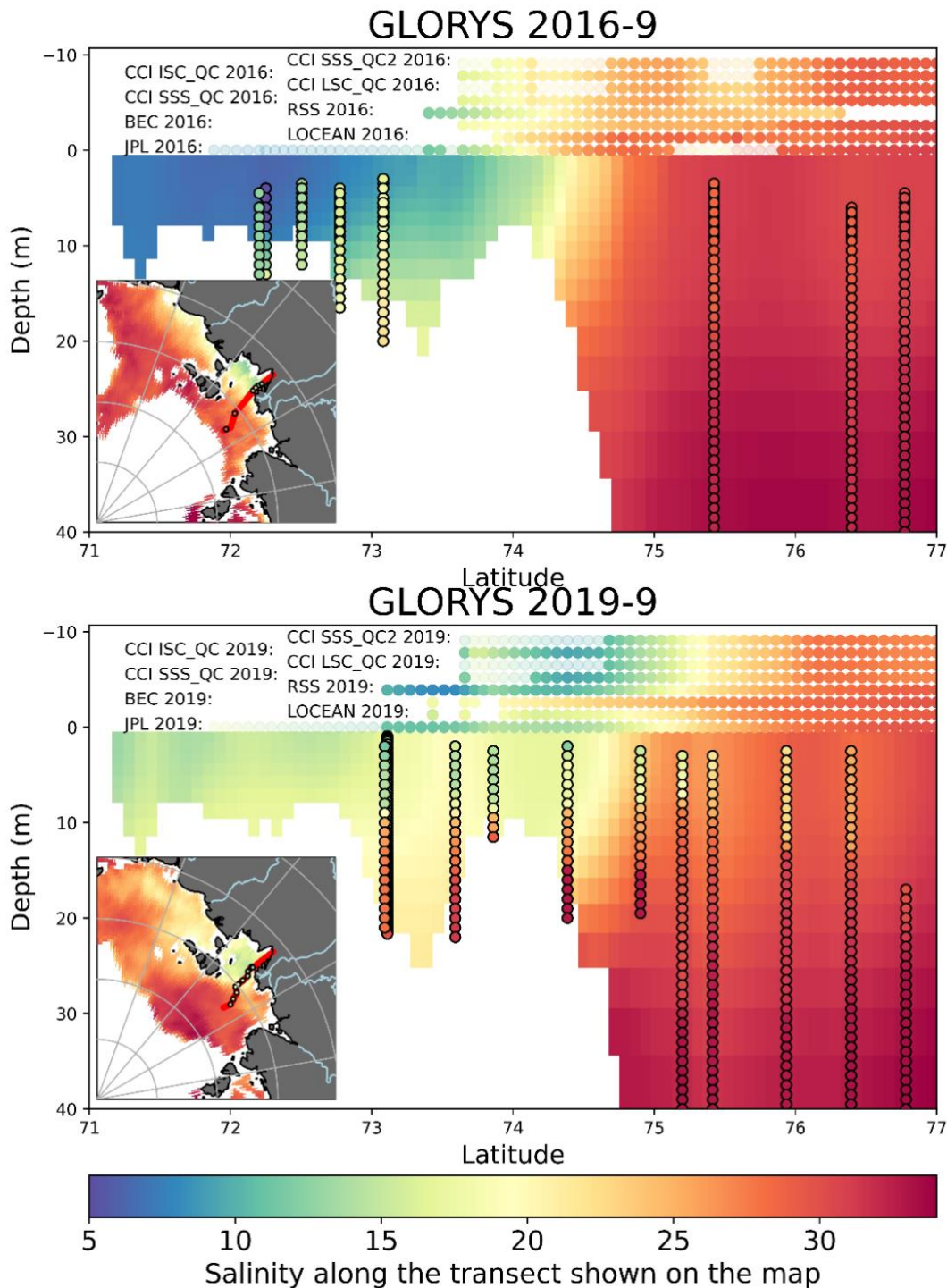



Figure 12: GLORYS 1/12° v1 Salinity vertical transect (top) in 2016, (bottom) in 2019. Transect are obtained along red transect through in situ data represented on map in bottom left for each subpanel with the SSS from JPL SMAP as a background. In situ data are overlaid with black rings. On top of each subpanel, satellite SSS data along that transect are represented for (bottom to top) L3 JPL SMAP v5, L3 LOCEAN SMOS Arctic v1.1, L3 BEC SMOS Arctic+ v3.1, RSS SMAP v4, CCI v4.4 products with different flags applied. Satellite data are made semi-transparent when flags are applied, JPL SMAP excludes data where the provided uncertainty is > 1 pss; CCI_SSS_QC excludes SSS_QC flags, CCI_LSC_QC excludes Land Sea flags, CCI_ISC_QC excludes Ice Sea flags, CCI_SSS_QC2 excludes all flagged data. This figure has been adapted from Figure 3 in [Hudson et al., 2023].

The years 2016 and 2019 stand out as having notably different patterns of Laptev SSS (Figure 12). In 2016, the freshest salinities are coastally confined and do not travel far off the continental

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 29 of 63</p>
--	---	---

shelf. In 2019, the freshest salinities travel considerably further offshore and extend over most of the Western Laptev and East Siberian Sea.

Despite the strong overall similarity between gridded products, notable differences are visible between *in situ* data and satellite products with GLORYS SSS. In 2019, the fresh layer appears to extend further offshore in *in situ* data than in GLORYS. LOCEAN, JPL, RSS and CCI (products versions are described in Figure 12 caption) appear to capture this extended plume better, but still do not capture the full extent visible in *in situ* data. This difference is likely primarily due to the temporal mismatch between the September monthly mean GLORYS and satellite products with *in situ* data collected in late September 2016 and early October 2019.

Of all the products considered here, CCI and LOCEAN products capture particularly consistent patterns of interannual variability and have strongest correlations with *in situ* data (Table 8 , $r > 0.8$) over the full period from 2010 to 2020. Over the SMAP period, correlations are the strongest for pure SMAP products (JPL and RSS) and for the CCI product. When limiting the data to a common dataset for comparison, we get the same conclusions, but with better RMSD against *in situ* linked to smaller SSS variability in the known flagged regions.

Test on CCI flagging are visible in Figure 12 and Table 8 (see descriptions of the flags in Figure 12 caption) and highlights improvements of the statistics (both increased correlation and decreased RMSD) when the Ice (isc) flag is raised.

Table 8: Statistics of Reanalysis and satellite SSS products against in situ at 10m depth for the period a) 2010-2020 and b) 2015-2020

a)	Period 2010-06 to 2020-01; Observation at 10m depth					
	All observations			Common observations		
	N _{obs}	RMSD	r	N _{obs}	RMSD	r
glorys	1676	1.90	0.77	299	2.61	0.53
bec	396	2.21	0.76	299	2.06	0.67
locean	406	2.07	0.84	299	1.71	0.80
cci	379	2.06	0.87	299	1.95	0.83
cci_qc	373	2.07	0.83	299	1.95	0.83
cci_isc_qc	322	2.03	0.89	299	1.95	0.83
cci_lsc_qc	361	1.96	0.85	299	1.95	0.83
cci_qc2	303	1.96	0.82	299	1.95	0.83
b)	Period 2015-04 to 2020-01; Observation at 10m depth					
	All observations			Common observations		
	N _{obs}	RMSD	r	N _{obs}	RMSD	r
glorys	222	3.16	0.80	54	3.55	0.70
jpl	124	2.98	0.88	54	2.10	0.92
jpl_ultr	100	1.85	0.92	54	2.10	0.92
rss	67	2.77	0.93	54	2.17	0.89
bec	133	2.90	0.79	54	3.11	0.74
locean	132	2.53	0.86	54	2.43	0.87
cci	114	2.70	0.90	54	3.17	0.86
cci_qc	110	2.73	0.85	54	3.17	0.86
cci_isc_qc	100	2.66	0.91	54	3.17	0.86
cci_lsc_qc	108	2.63	0.88	54	3.17	0.86
cci_qc2	92	2.63	0.82	54	3.17	0.86

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

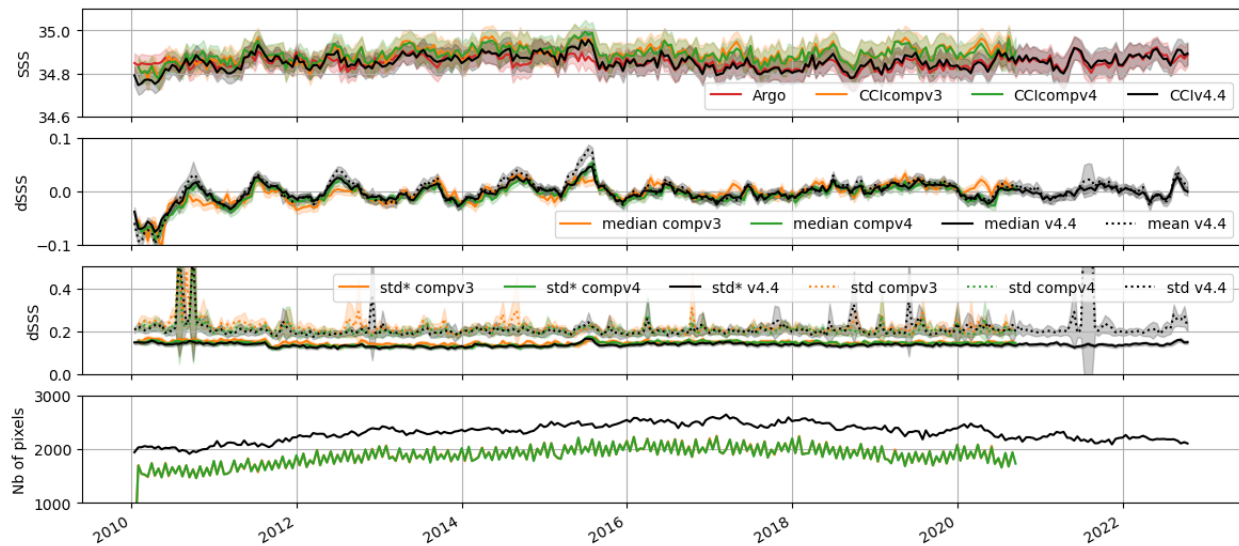


Figure 13: (1st panel): mean SSS of gridded MDB of pairwise Argo in red and L4 CCI in black for v4 and in orange and green for comparison v3 and v4; (2nd panel) Average of; (3rd panel) standard deviation of; the gridded MDB of the pairwise SSS difference between CCI and Argo. Solid lines represent (2nd panel) the median (3rd panel) the robust standard deviation. Dashed lines represent (2nd panel) the mean (3rd panel) the standard deviation. The shading indicates the 95% confidence interval. (4th panel) number of valid gridded MDB values.

The time series in Figure 13 represents the temporal evolution of gridded MDB of the pairwise measurements of CCI L4 (v4, compv3 and compv4) and Argo and their differences. The mean SSS temporal variability represented on the top panel shows good agreement between CCIv4 (CCIv4.4 in the label) and Argo (collocated against v4) with a mean around 34.9 pss. The beginning of the period in 2010 highlights a lower value for CCIv4 than Argo of less than 0.1 pss. In April to August 2015 there is an increase of v4 against Argo, which is linked to known issue in SMAP v5.0 in North Atlantic. There are no significant differences between compv3 and compv4.

The two middle panels represent the gridded MDB of the pairwise differences of CCI with Argo for average difference (mean and median); and dispersion (classic standard deviation and robust standard deviation). 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 2016. Similarly, to the previous panel, strong differences are observed at the beginning of the period in 2010 and in April to August 2015. There is significant differences between compv3 and compv4 between mid-2010 and 2013, between April to August 2015 and at the beginning of 2020.

The dispersion, as estimated by the robust standard deviation of the difference, stays relatively constant over the full time series between 0.13 pss and 0.16 pss. There is a diminution of the robust standard deviation in mid-2011 with Aquarius and a local peak mid-2015 corresponding to the known issue with SMAP v5.0 in North-Atlantic. The classic standard deviation presents some peaks suggesting more extreme values in the tail of the distribution. Globally, the three CCI L4 versions are very similar. The number of valid pixels in the gridded MDB for each time step (bi-weekly) is indicated in the bottom panel. It slightly increases in time, directly link to the increase of the number of Argo profiles from 2010 to 2016.

The temporal variability of the gridded MDB of the pairwise CCI/Argo differences is further assessed using latitude-time (Hovmöller) plots over the global ocean (Figure 14) for the L4v4 products. About half of the pixels are not significantly different from zero (at 95%). There are significant oscillating signals with stronger amplitudes at higher latitude. There is a strong difference in mid-2015 between 40°-80°N linked to the known issue with SMAP v5.0 between April and August 2015. The first 6 months in 2010, indicates CCI are fresher than Argo particularly in the Northern hemisphere and at high latitudes. There is a small trend from fresher (blue) to saltier for latitudes between 0°N and 30°N. A symmetric trend exists in the southern hemisphere.

Figure 15 is similar to Figure 14, but using the gridded MDB comparison dataset for v3 and v4, and showing at the bottom the improvements (in green) and degradation (in orange) between v3 and v4. This latter figure highlights the strong improvement in the correction of the seasonal cycle in the difference between CCI and Argo which was strongly present in the Northern Hemisphere up to 2015 in v3. It also improves seasonally at high latitude in both hemispheres. Some degradations appear sporadically at high latitude and at the beginning of the time series at low latitude (+-20°).

The spatial representation of seasonal climatology of the gridded MDB v4 difference (Figure 16), does not highlight strong, significant and large-scale differences. It is calculated using the median for each season over the full time series. Around Japan and in the northern North-Atlantic, CCI L4v4 is fresher in Winter (DJF) and saltier in Summer (JJA). A seasonal spatial signature is less pronounced and significant in the Southern hemisphere. Some local seasonal differences are visible close to the coast, generally related to river plumes, potentially associated with vertical stratification (see details in section 4.3).

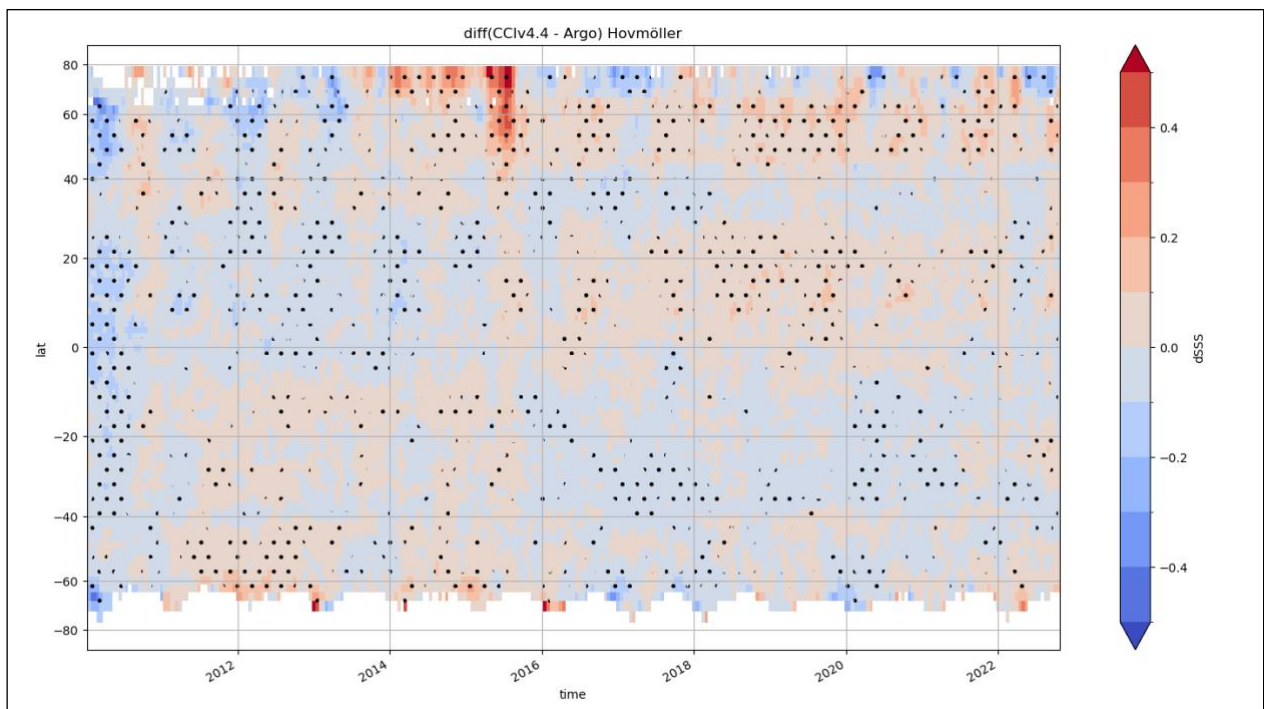


Figure 14: Global latitude-time Hovmöller of the gridded MDB of the pairwise CCI difference with Argo for (top) L4 v3; (bottom) L4 v4. Each pixel represents the median value after a moving window over 2 pixels in latitude and time. Data which are significantly different from 0 (at 95%) are indicated with dots. All sub-figures share the same colour bar.

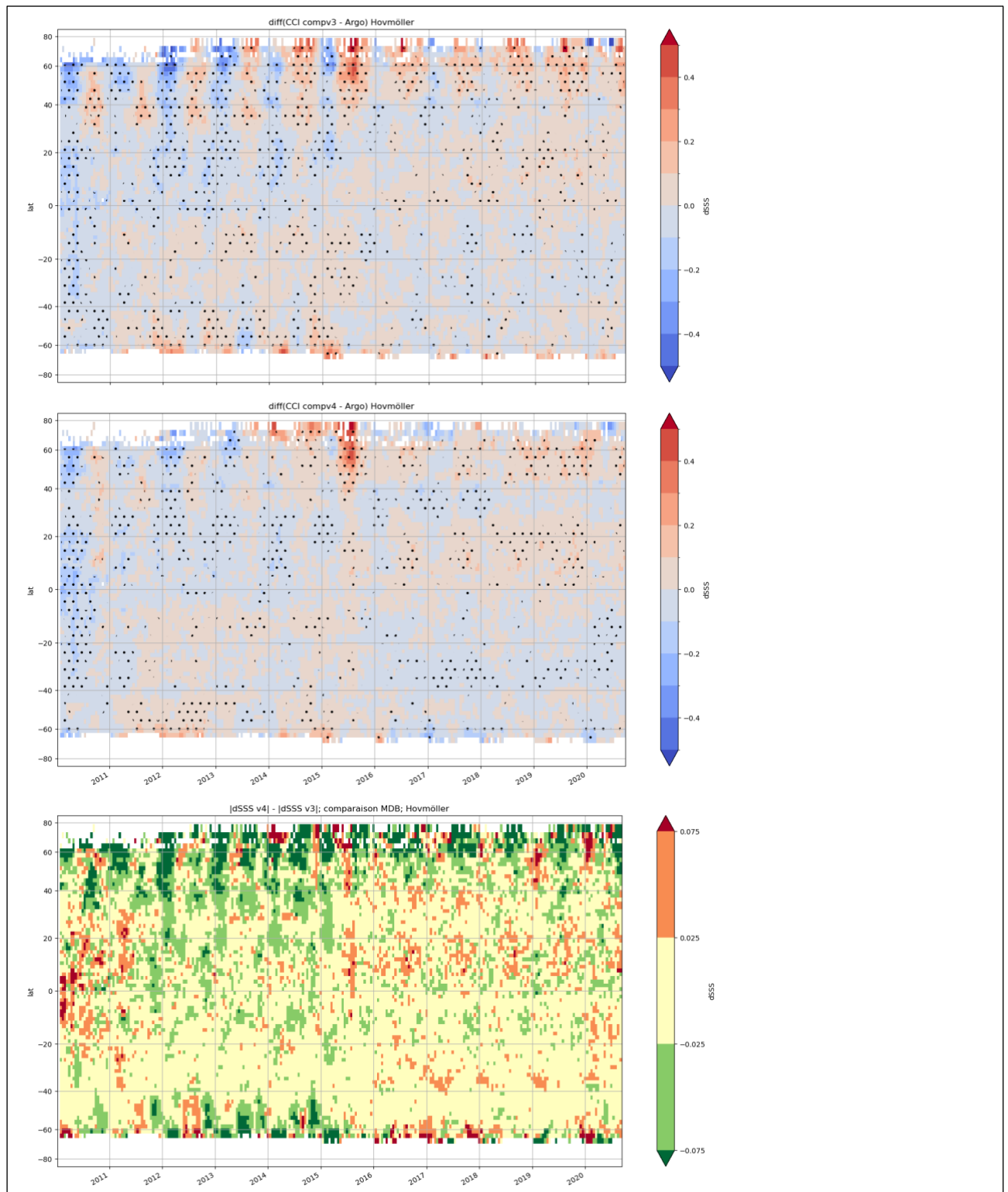


Figure 15: Global latitude-time Hovmöller of the gridded MDB of the pairwise CCI difference with Argo for (top) L4 v3 comparison dataset; (middle) L4 v4 comparison dataset; (bottom) Absolute difference between the two subplots above. Green indicates improvement towards zero; Red indicates degradation away from zero. Each pixel represents the median value after a moving window over 2 pixels in latitude and time. Data which are significantly different from 0 (at 95%) are indicated with dots. All sub-figures share the same colour bar.

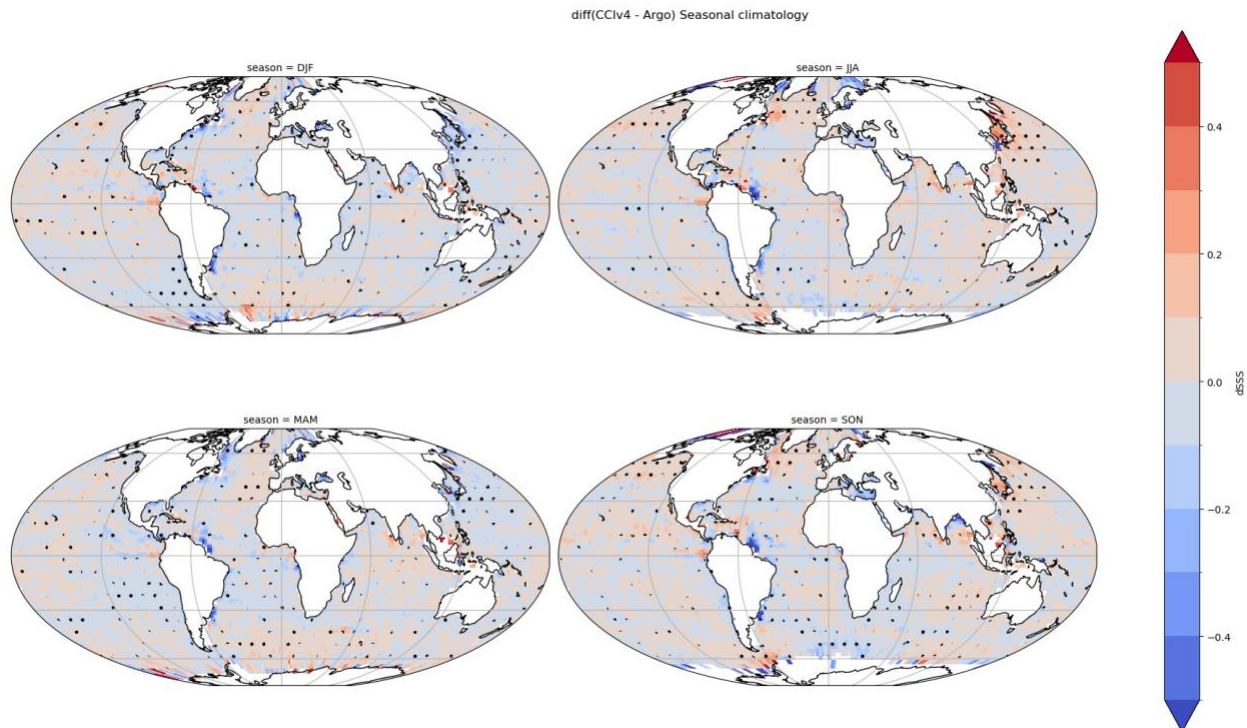


Figure 16: Seasonal climatology of the gridded pairwise CCI L4 difference with Argo calculated using the median. A moving window average of 2 x 2 pixels in longitude and latitude have been applied to increase the number of sampled, hence the significance. Pixels, which are statistically significant (at 95%) are indicated with dots.

4.3 In Situ Vertical Representiveness Error

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

In order to get a global distribution of the vertical representiveness error, we calculate the gradient for each Argo profile between an acquisition depth of 5 m and 10 m. We use the same grid as for the pairwise comparison and take the median value of this gradient for each cell (in time and space). The seasonal climatology of this gradient in salinity is represented in Figure 17 highlighting that most of the ocean does not show a noticeable gradient between 5m and 10m, except in areas with strong freshwater fluxes (e.g. river plumes, ITCZ, Labrador current, ...). As expected, the value at 5 m is usually fresher than the salinity at 10 m. The surface is saltier only for very specific areas and periods such as the Mediterranean Sea in Summer, but only a low apltitude. The strongest gradients in salinity relate to the tropics in all seasons with typical values higher than 0.02 pss/m. If we linearly extrapolate these high gradients from 5 m to the surface, we might suspect differences due to the vertical sampling exceeding 0.1 pss. In Summer, vertical gradients appear in the Northern Hemisphere in the vicinity of western boundary currents (Gulf Stream and Kuroshio).

These results suggest the SSS measured by satellite would tend to be fresher than the one measured *in situ* by Argo. However, this effect is an order of magnitude less than the seasonal one observed on Figure 16 so that the mean difference of CCI minus Argo in Figure 16 does not reflect the patterns below.

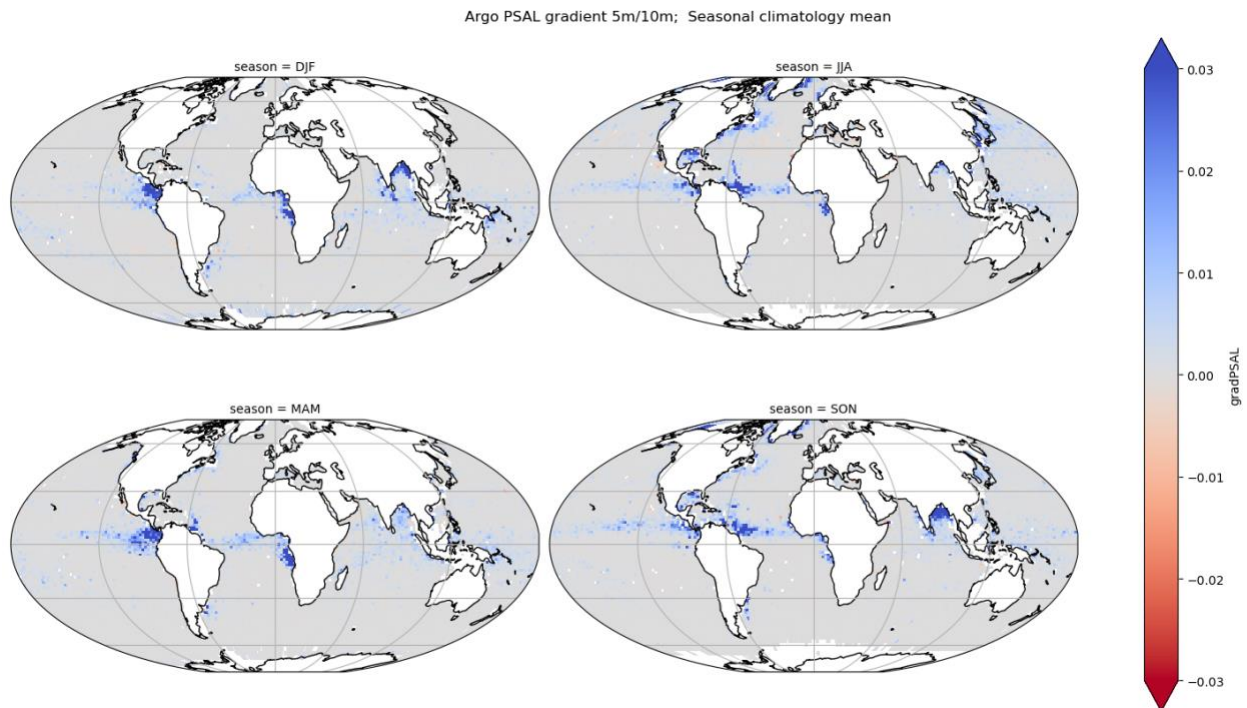


Figure 17: Seasonal Salinity gradient (in pss/m) derived from Argo at 5 m and 10 m. Gradient are gridded on the same grid as used for the pairwise difference (bi-weekly; 175 km).

4.4 Temporal & spatial effective resolution

4.4.1 Temporal effective resolution

The average temporal power spectra of SSS from all moorings and CCI collocations from the Pi-MEP MDB are represented in Figure 18 for the weekly products and in Figure 19 for the monthly products. ISAS optimal interpolation SSS and the SSS from the Mercator numerical circulation model are also shown. CCI L4 Weekly and Monthly products shows as expected a decrease at the Nyquist frequency (respectively 14 days and 60 days).

For the Monthly products (Figure 19), we have the power spectra for CCI L4 v3.2, v4.4 and for CCI L3C SMOS, Aquarius, SMAP. SMAP power spectra is fully aligned with Mercator, and higher than the moorings without decrease at the Nyquist frequency. Aquarius power spectra decrease and depart from the mooring power spectra from 100 days. SMOS power spectra is very similar to the one from CCI L4 but slightly higher.

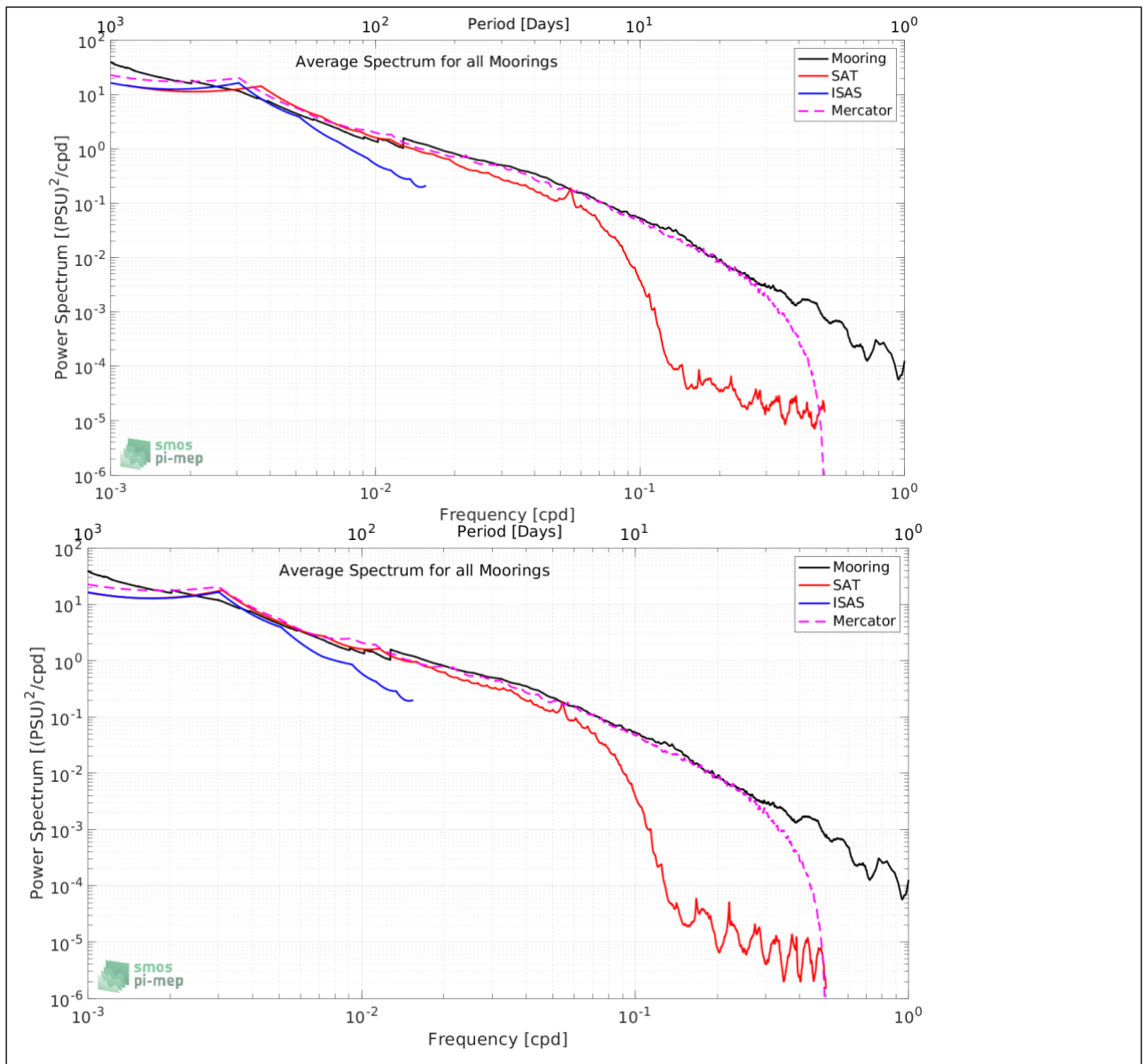


Figure 18 : Average power spectrum of SSS from (black) moorings, (red) CCI Weekly products, (blue) ISAS, (pink) Mercator; (top) for CCI L4 Weekly v3.2; (bottom) for CCI L4 Weekly v4.4. from Pi-MEP.

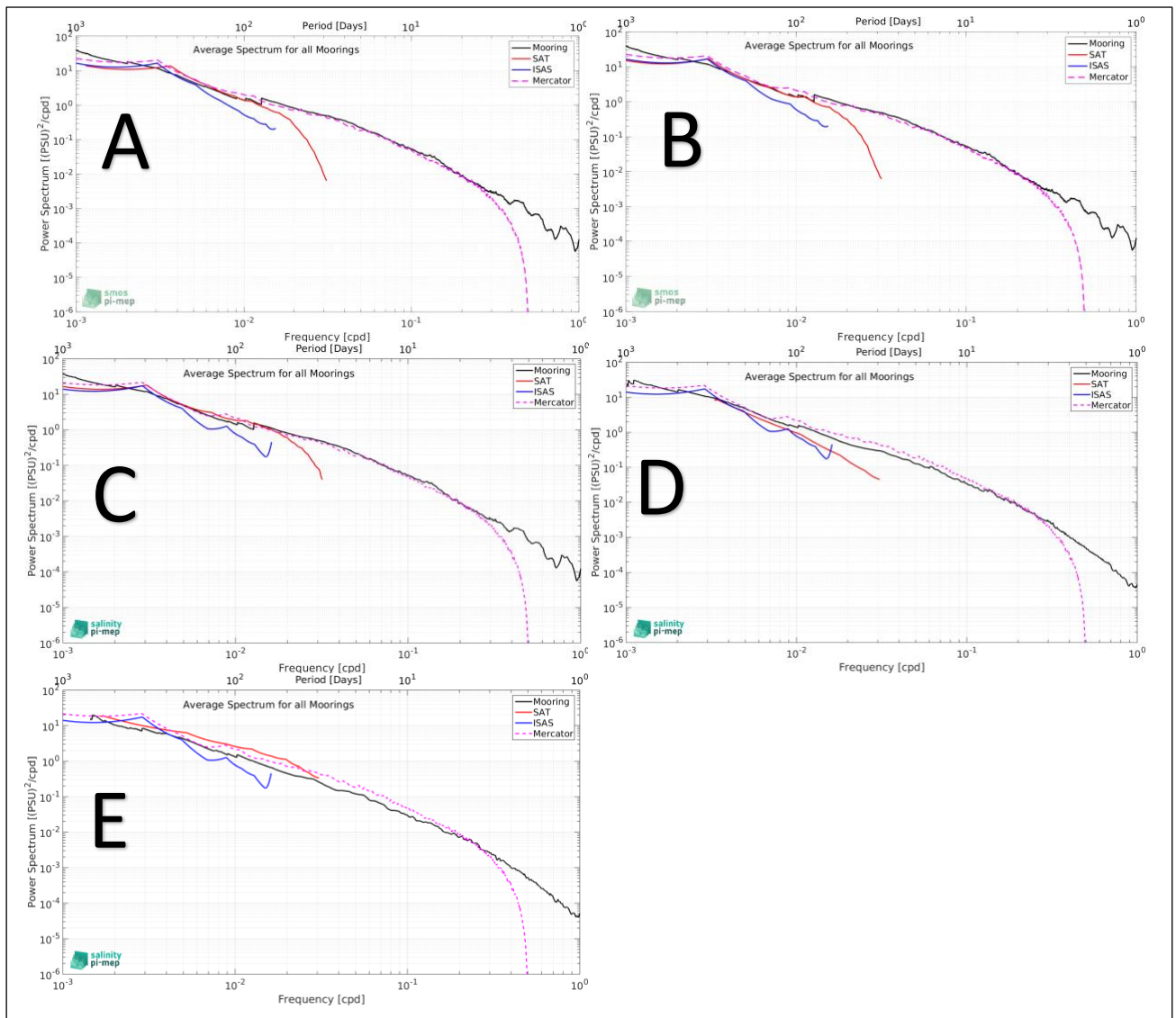


Figure 19: Average power spectrum of SSS from (black) moorings, (red) CCI Weekly products, (blue) ISAS, (pink) Mercator; (A) CCI L4 Monthly v3.2; (B) CCI L4 Monthly v4.4; (C) CCI L3C Monthly SMOS; (D) CCI L3C Monthly Aquarius; (E) CCI L3C SMAP. from Pi-MEP.

4.4.2 Spatial effective resolution: Assessment of mesoscale features in Tropical Atlantic

The surface mixed layer thermohaline structures at meso-scale to submesoscale (smaller than the local radius of deformation, *Chelton et al., 1998*) are ubiquitous features in the global ocean. They contribute to horizontal and vertical heat and salt exchange and vertical re-stratification (*Fox-Kemper et al., 2005*). They have a global impact on ocean circulation and climate since they contribute to the cascade of energy from large scale toward the smallest scales of diffusive mixing (*Callies and Ferrari, 2013*). Eventually, they have a major impact on bio-geochemistry and ecosystems. The submesoscale processes are characterized by very intense vertical velocities that allow strong exchanges of carbon, oxygen and nutrient between the surface and subsurface ocean (*Lévy and Martin, 2013*).

Until early 2010, satellite capabilities for observing surface thermohaline variability have mainly relied only on the observation of Sea Surface Temperature (SST), resolving small scale features such as 10 km (*Kilpatrick et al., 2015*). In contrast, synoptic images of Sea Surface Salinity (SSS) were not available and *in situ* SSS at high resolution are only available from a few high-resolution sections from Thermosalinograph (TSG) surveys from ships of opportunity, repeated transects or cruise campaign (*Kolodziejczyk et al., 2015b*). Since 2010, thanks to ESA SMOS mission, then NASA Aquarius and SMAP missions, 4-7 days global maps of SSS at resolution between 40-100 km are now available permitting observation of larger mesoscale features in subtropical and tropical region (*Reul et al., 2014; Kolodziejczyk et al., 2015a*).

To verify the effective capability of the new CCI-SSS products v4.4 and v3.2 (7 days) to monitor the large mesoscale features of SSS in the subtropical and tropical regions, the CCI SSS were systematically co-localized and compared with TSG SSS along existing repeated transects in the Subtropical North Atlantic and Tropical Atlantic. An effective metric to assess the SSS horizontal variance and scale content of both products is to compute the spectra and coherency spectra between TSG SSS and CCI-SSSv4.4 and v3.2 (*Boutin et al., 2018*).

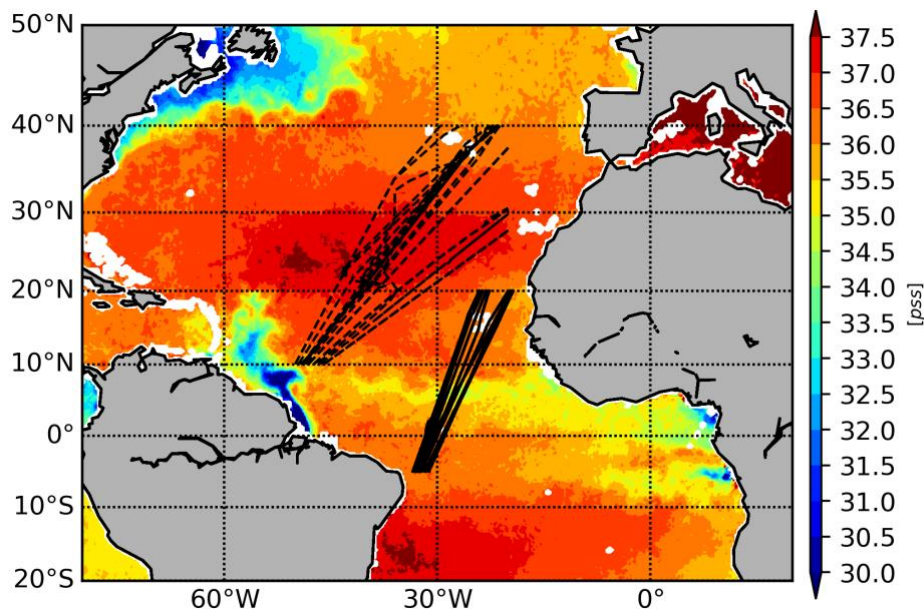



Figure 20: CCI+SSS on 30 June 2011 with 93 TSG transects in the Subtropical North Atlantic (dashed) and 26 TSG transect in the Tropical Atlantic.

SSS TSG transects were collected from ships of opportunity (representative of salinity at 10 m depth), resolving horizontal SSS features at around 2-3 km (*Alory et al., 2015*). Two regions were chosen for the present study (Figure 20): i) the North Atlantic subtropical SSS maximum (50-20°W/10-40°N), where 88 transects between 2011-2016 are available; and ii) the Tropical Atlantic (40-10°W/5°S-20°N) where 26 transects between 2014-2016 are available. Individual transects were visually inspected and suspicious transects were discarded. To reduce uncertainty due to noisy individual spectrum from each individual transect, spectra were averaged for both regions.

The horizontal SSS coherency spectra refers to the coherency of the SSS horizontal variability between the co-located TSG SSS and CCI+SSS products, *i.e.* the level of correlation of

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 39 of 63</p>
--	---	---

the SSS signal for a given wavelength range. This allows the assessment of the actual capability of CCI+SSS products to observe mesoscale features (>50 km) from the noise and spurious SSS contamination.

In the Subtropical North Atlantic (Figure 21a), the horizontal variance spectrum for CCI+SSS versions 4.4 and 3.2, alongside the TSG spectra, indicates a good match despite a marginal reduction in energy across the 50-1000 km wavelength range. This similarity, particularly in the slope within the same wavelength span, implies that the variance of mesoscale features within the CCI+SSS products may be somewhat smoothed. Interestingly, the coherency spectra (Figure 21b) exhibit a quasi-linear decrease from large scale (coherency>0.75 for wavelength > 1000 km) to mesoscale (coherency~0.30 for wavelength ~ 300 km). The significance at 95% is lost for wavelength below 200 km. This suggests that wavelengths smaller than 300 km are poorly represented in the CCI+SSS product. The CCI+SSSv3.2 spectrum exhibits slightly less coherency than the CCI+SSSv4.4 spectrum, this is likely due to slightly noisier SSS v3.2 fields. Overall, these results are consistent with a previous study on investigating the SMOS LOCEAN CEC L3 product (*Boutin et al., 2018*) in the same region, however, with a slightly improved coherency for CCI+SSS product. As such, no significant differences from the previous PVIR-CCI+SSS report are reported.

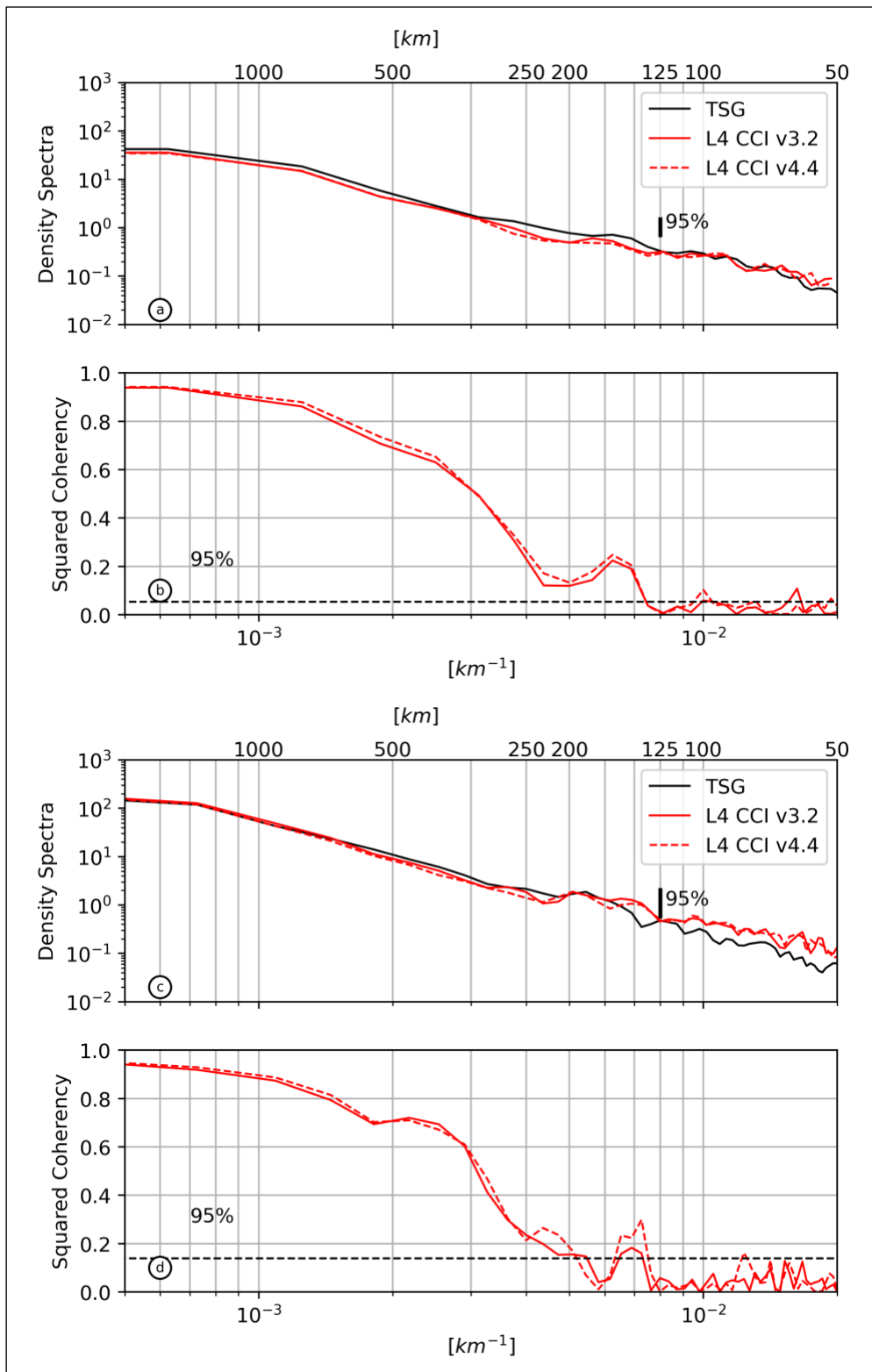


Figure 21: a) Density spectra from from 88 collocated TSG (black); CCI+SSS v2.31 (dashed red); CCI+SSS v3.1 (solid red) SSS transects in Subtropical North Atlantic. Vertical thick black bar is the level of confidence at 95%. b) Coherency between the TSG and CCI+SSS SSS transects. Dashed line is the level of significance at 95%. c) Density spectra from from 26 collocated TSG (black)/CCI+SSS(red) SSS transects in Tropical Atlantic. Vertical thick black bar is the level of confidence at 95%. d) Coherency between the TSG and CCI +SSS SSS transects. Dashed line is the level of significance at 95%.

In the Tropical Atlantic (Figure 21c), the spectra from TSG and CCI+SSS display similar patterns in terms of variance levels and slopes. Additionally, these spectra demonstrate a high coherence level (coherency > 0.5) for wavelengths above 300 km, as illustrated in Figure 21d. However, at wavelengths below 250 km, a noticeable decline in coherency suggests that the CCI+SSS product struggles to accurately resolve scales smaller than 125 km. This represents a minor enhancement compared to the initial CCI+SSS version 1, as documented in the previous PVIR report. It's also noted that there are no marked differences between versions 3.2 and 4.4 of the CCI+SSS product.

In conclusion, in the subtropical Atlantic, the CCI+SSSv4.4 and v3.2 products can resolve wavelengths of the order of 300 km. This wavelength corresponds to horizontal mesoscale features of the order of about 150 km (gradient, eddy). However, the level of coherency between TSG SSS horizontal variability and CCI+SSS drops rapidly at the mesoscale. In the tropics, the level of coherency remains high up to 300 km wavelengths, then drops dramatically.

The loss of coherency at smaller horizontal wavelengths could be explained by i) the limiting resolution of SSS satellite mission (>50 km), ii) remaining noise and artifacts in the CCI+SSS data, and iii) smoothing from the objective analysis procedure of the CCI+SSS products. Nevertheless, it is worth pointing out that inconsistency between instantaneous and point-wise measurements from the TSG data and co-localized CCI+SSS products (7 days, 25 km) may be responsible for a shift and lag between TSG SSS measurements and CCI+SSS products resulting in loss of coherency for the smaller and faster SSS mesoscale structures.

4.4.3 Power Density Spectrum (PDS)

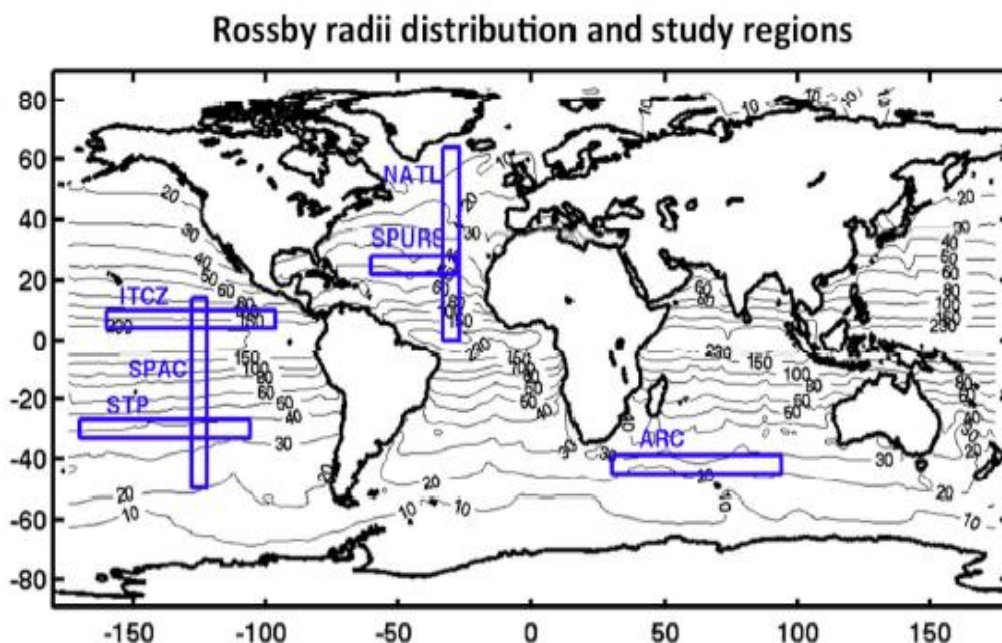


Figure 22: Regions as defined in [Hoareau et al. 2018] where the PDS are calculated. Regions are NATL for North Atlantic Ocean; SPURS for Salinity Processes in the Upper ocean Regional Study; ITCZ for Inter-Tropical Convergence Zone; SPAC for South-Pacific ocean; STP for South-Tropical Pacific ocean; ARC for Agulhas Return Current.

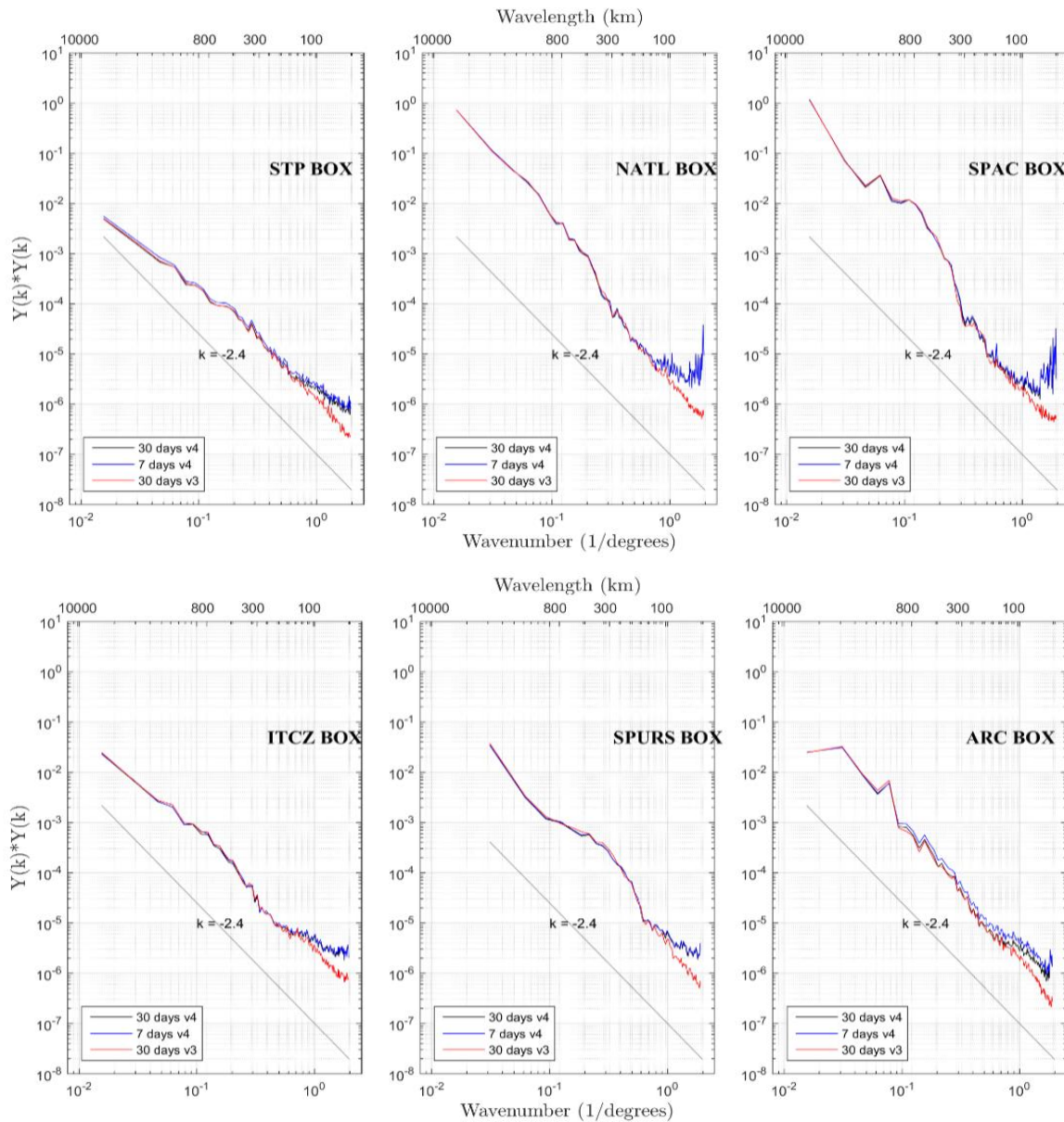



Figure 23: Power density spectra (PDS) in the 6 regions defined [Hoareau et al., 2018] and presented in Figure 22: Regions as defined in [Hoareau et al. 2018] where the PDS are calculated. for (in red) 30 days V3 (in black), 30 days V4 (in blue) days V4 products calculated from 2010 to 2020. A grey line, representing slope of -2.4, is provided for reference.

Figure 23 represents the Power Density Spectra in the six regions defined in Figure 22 for CCI L4v4 Monthly and Weekly products and for CCI L4v3 Monthly products (regridged on a $\frac{1}{4}^{\circ}$ degree). We notice a good match in PDS between v3 and v4 Monthly products from large scales up to scales of about 100km. PDS for v4 Weekly products follow the same curve as v4 Monthly except for the Agulhas Return Current (ARC) box with more power for the Weekly products. Whereas v3 is following a slope with $k=-2.4$ without any saturation, v4 for the Weekly and Monthly products fold around spatial scale of 100km.

	<p style="text-align: center;">Climate Change Initiative+ (CCI+) Phase 2</p> <p style="text-align: center;">Product Validation and Intercomparison Report</p>	<p>Ref.: ESA-EOP-SC-AMT-2021-26 Date: 1/21/2024 Version : v4.0 Page: 43 of 63</p>
--	---	---

4.5 Uncertainty

As explained in section 3.3 above, we will follow two approaches to validate satellite uncertainty estimates:

- Normalise the dSSS by the uncertainty with a centred reduced variable and analyse the variation compared to a theoretical behavior of a random normalised variable with mean of zero and standard deviation of one.
- Compare the dSSS distribution with the uncertainty estimates.

For both cases, we will consider the satellite uncertainty $\Delta\sigma_{sat}$ alone or the total uncertainty which combines the satellite uncertainty with the reference uncertainty itself which includes the

horizontal representiveness error ($\sigma_{tot} = \sqrt{\Delta\sigma_{sat}^2 + \Delta\sigma_{ref}^2}$).

4.5.1 Normalised SSS

In this section we look at normalising the SSS differences (dSSS; satellite and Pi-MEP) by the uncertainties (either just satellite or satellite with reference uncertainties). In an ideal situation, with perfect estimation of the uncertainties, the normalised standard deviation would be one.

Figure 24 represents a time series of the normalised dSSS using (A) the CCI provided satellite uncertainties and (B) the quadratic mean of the satellite uncertainties plus the reference uncertainties described in the methods. The normalised SSS represented in (Figure 24A bottom rows) shows a standard deviation between one and two when the normalisation uses just the satellite uncertainties. Inclusion of the Aquarius is clearly seen in the timeseries with a step following launch in June 2011. The end of mission for Aquarius (June 2015) is not as obvious, perhaps due to the proximity (in time) to the launch of SMAP (April 2015). In mid-2015, there is a blob in std which is linked to the known issues from SMAP v5.0 seen above. Over the period 2010 to 2015, compv4 is closer to one than compv3 with no sensible difference afterwards. The robust standard deviation always gives slightly lower estimates compared to the classic standard deviation due to longer tails in the distributions. The higher values observed here are linked to the reference uncertainties linked to differences between pointwise measurements and pixel-averaged measurements. With the total uncertainties in (Figure 24B bottom), including the reference representativeness error, standard deviation are closer to one, even slightly below one for the SMOS-only period (01/2010-06/2011).

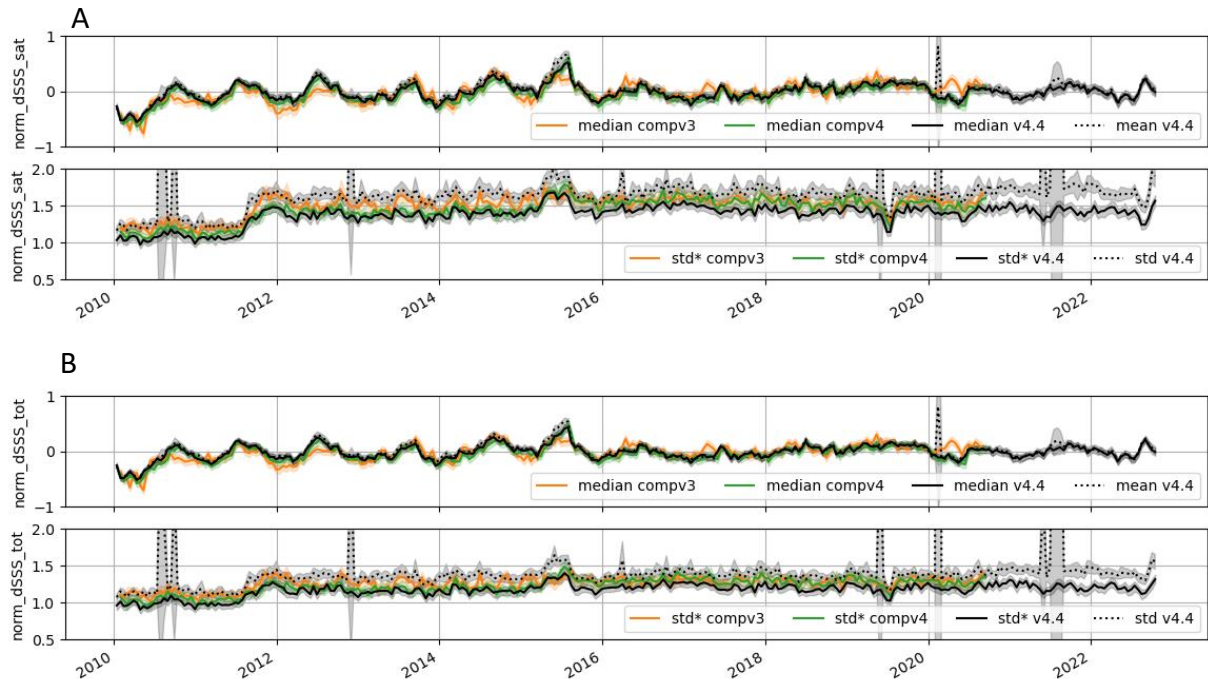


Figure 24 : Time series of the normalised SSS normalised using (A) the satellite uncertainty; (B) the total uncertainty combining the satellite and reference uncertainty. Top row for each panel represents (solid line) the median and (dashed line) the mean. Bottom row for each panel represents (solid line) the robust standard deviation and (dashed line) the classic standard deviation. Colours are for the L4v4 and the comparison dataset for L4v3 and L4v4.

The average (median, mean) time series of the normalised SSS (Figure 24 A/B 1st row) does not provide information about the uncertainty. For each point, we have more than ~2000 observations (cf Figure 13 last row), leading to a theoretical variability (standard deviation) of the normalised SSS of 0.02 which is much higher than what we observe.

4.5.2 Compared SSS Distribution

The gridded MDB of the pairwise differences are now binned in uncertainty 0.05 pss bins (Figure 25) and computed over the full time series for each product (CCIV4, compv4 and compv3). The top row is based on satellite uncertainty whereas the bottom row reflects total uncertainty. Ideally the standard deviation (classic or robust) and the dSSS should follow a one-to-one relationship.

If one only takes into account the satellite uncertainties (Figure 25-top), then the observed standard deviation exceeds the satellite uncertainties but follows a nearly linear relation up to a satellite uncertainty of 0.4 pss. On the contrary, if one adds the representativeness error (Figure 25-bottom), all products are closer to the one-to-one relation up to a total uncertainty of about 0.4 pss.

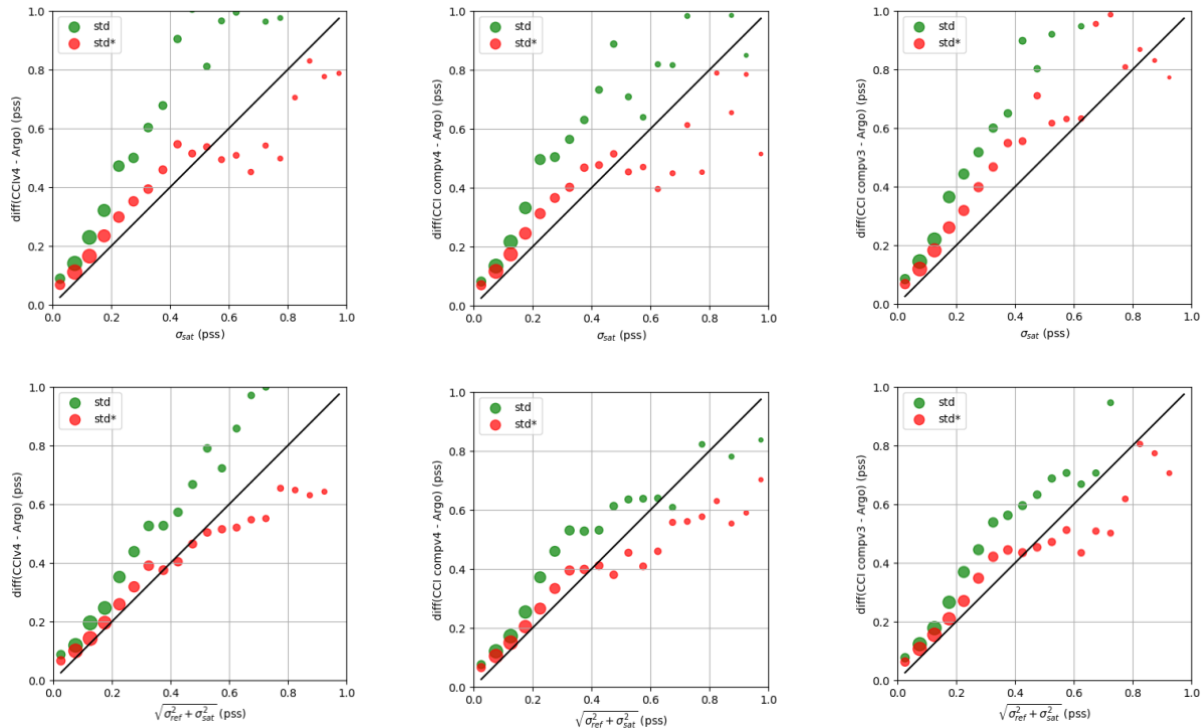


Figure 25: measured standard deviation (green and red dots) for classic and robust standard deviation respectively; of the gridded pairwise CCI/Argo difference (dSSS) for each uncertainty 0.05 bin. (top) using satellite uncertainty; (bottom) using total uncertainty - satellite + reference (column from left to right) for L4v4, comparison dataset L4v4, comparison dataset L4v3. The size of the circle indicates the number of data.

4.5.3 Estimation of mismatch error using GLORYS 1/12°

In this subsection, we use the Pi-MEP MDB to estimate the mismatch error using output from the GLORYS 1/12° daily model. GLORYS salinity is taken at the time, horizontal position and at two depths (0.5m depth and in situ depth) of the in-situ measurements (here Argo). To simulate coarse acquisition from satellite, average GLORYS values are taken within a radius of 25/50 km of GLORYS SSS (0.5m depth) and within 7/30 days around Argo float locations/time using a spatio-temporal averaging.

The difference of salinity between GLORYS resampled (50 km, 30 days, surface) and GLORYS at Argo depth presents some systematic differences (Figure 26) with negative bias, similar to the ones observed in Figure 4, close to large rivers plumes like Amazon, Niger, Mississippi, Bay of Bengal or for some western boundary currents like the Gulf Stream or the Brazil current. Positive bias, similar to the ones observed in Figure 4, are present in the Falkland/Malvinas current and continues following the circumpolar current up to the Kerguelen Plateau and in the North-westerns part of the Atlantic. The robust standard deviation of the mismatch difference highlights the same areas as mentioned earlier, i.e. areas characterized by strong current and/or strong SSS gradients such as river plumes, which are the same as the ones observed in Figure 4. For the full dataset, the mismatch bias is close to zero (-0.002 pss) with a standard deviation of 0.11 pss (robust standard deviation of 0.05 pss), which is about half the difference between CCI and Argo.

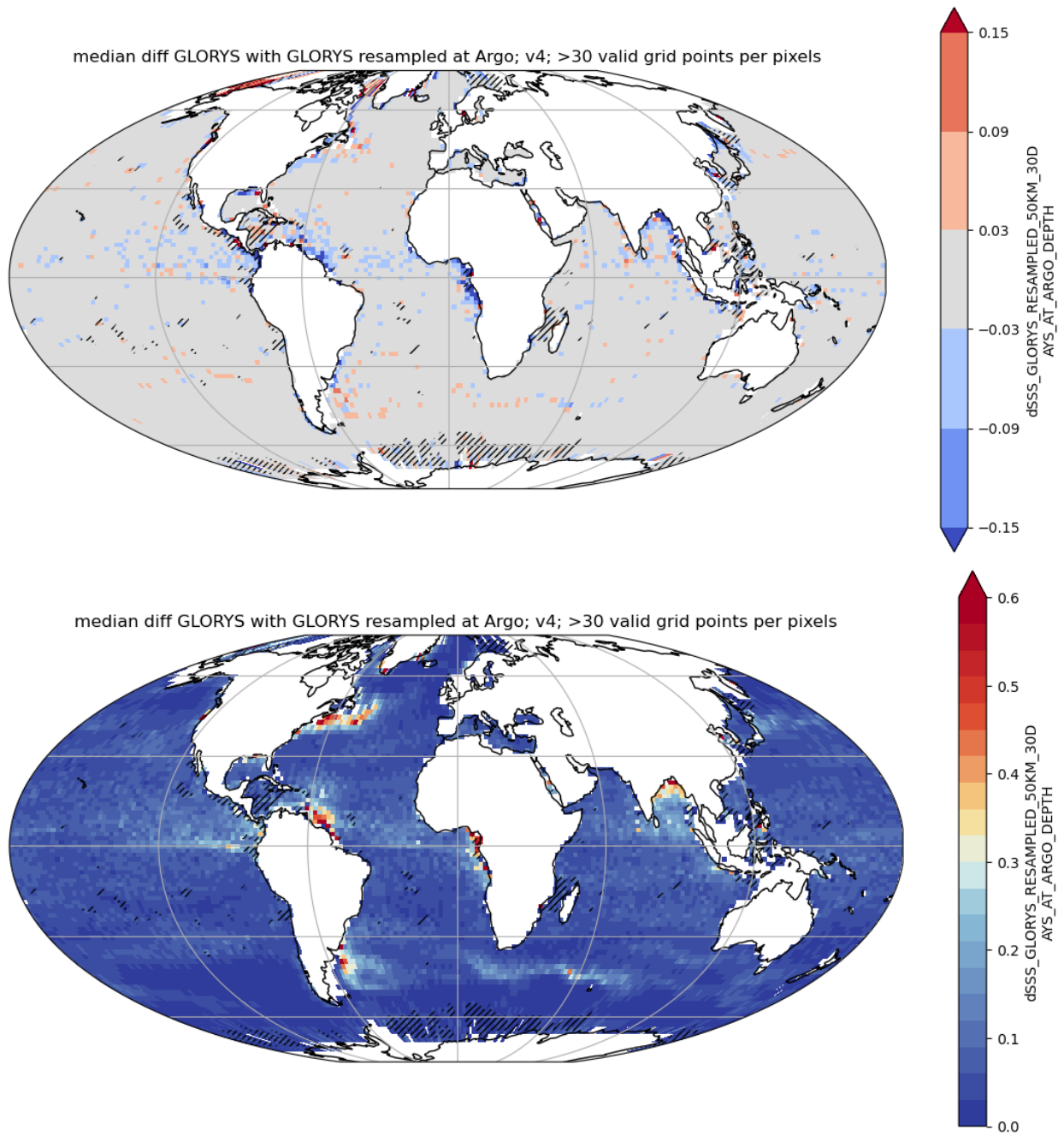


Figure 26: (top) Temporal median and (bottom) temporal robust standard deviation of the estimated sampling mismatch using GLORYS. This sampling mismatch estimates is obtain from the difference between GLORYS averaged over 50km, 30 days and GLORYS sampled at Argo time and position (horizontal and vertical). The colour scale is zoomed by 40% compared to the colour scale in Figure 4.

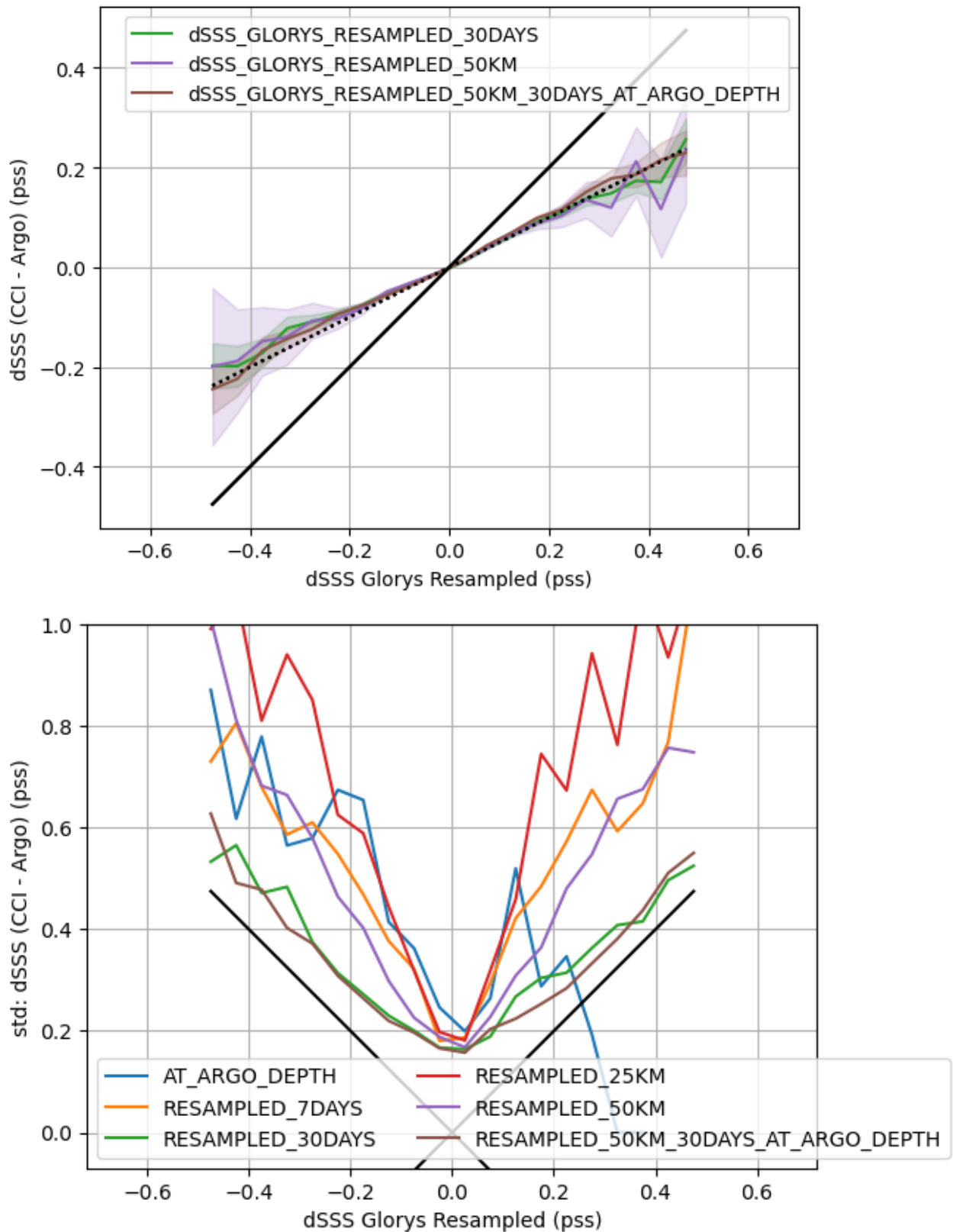


Figure 27: (top) mean and (bottom) standard deviation of the observed difference between CCI and Argo per bins of expected mismatch using GLORYS with the different resampling strategy represented in colour (see legend). For clarity, on the top figure, only the most significant plots are presented.



The comparison of the observed difference between CCI and Argo with the one expected by GLORYS (Figure 27-lower panel) highlights a systematic difference which is explained by the sampling mismatch between surface and point measurements (for GLORYS resampled at 50km, 30days or both). This difference explained about half the difference observed between CCI and Argo and is linear on the range +/-0.5 pss in the expected mismatch. If one resampled over 25km or 7 days (not shown), it follows about the same linear relation for the range +/-0.1 pss and is not significant after 0.2 pss. If one only estimates the vertical mismatch, without taking into account the spatio-temporal mismatch, the difference in the observation is significantly different from zero only for a difference GLORYS SSS minus GLORYS at Argo depth of -0.15 pss to 0 pss and it is anti-correlated with the expectation (not shown).

If one corrects for this systematic mismatch based on GLORYS and a 50 km, 30 days, Argo at depth sampling: i.e. $CCI - Argo - ratio * (GLORYS_{resampled} - GLORYS_{Argo})$, we obtain a standard deviation of the corrected difference between CCI and Argo of 0.206 pss (robust standard deviation of 0.132 pss), which is best for a ratio of 0.5.

As highlighted e.g. in [Thouvenin-Masson et al., 2022], and in Figure 26 the sampling mismatch between surface and point measurements play a significant role in the uncertainty. Figure 27-right represents the observed standard deviation of the difference CCI with Argo as function of the expected mismatch using GLORYS. It shows an expected vertical symmetry at zero for all resampling estimates excepted the one taking only the difference between the surface and the depth in blue. Mismatch estimates for resampling at "30 days" in green and "50km, 30 days, Argo depth" in marron follow a one-to-one relation suggesting these spatio-temporal scale are relevant to explain the observed variability. The other curves (see legend for resampling strategy) are much steeper, suggesting they only explain one portion of the observed variability. Part of the unexplained variability could come from the smallest scale which are not resolved or badly resolved by GLORYS 1/12° daily outputs. In [Thouvenin-Masson et al., 2022], in order to take the GLORYS unresolved small scales into account the variability was increased by a factor 1.20.

To conclude this subsection, we highlight in addition to the increase uncertainty due to the sampling mismatch, there is a systematic mismatch we can correct for based on a numerical model, which improves the performance of the CCI, Argo difference from 0.223 pss to 0.206 pss (0.137 pss to 0.132 pss for robust standard deviation).

End of document