Aquarius/SMAP sea surface salinity optimum interpolation analysis

Oleg Melnichenko[1], Peter Hacker[2], James Potemra[2], Thomas Meissner[3], and Frank Wentz[3]

[1] International Pacific Research Center, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, Hawaii
[2] Hawaii Institute of Geophysics and Planetology, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, Hawaii
[3] Remote Sensing Systems, Santa Rosa, California

IPRC Technical Note No. 7

May 7, 2021

Table of Content

1. Introduction

The purpose of this note is to describe a new global sea surface salinity (SSS) product. The product combines observations from NASA's Aquarius/SAC-D and Soil Moisture Active-Passive (SMAP) satellite missions into a continuous and consistent multi-satellite SSS data record.

The new dataset covers the period from September 2011 to the present. The beginning segment, from September 2011 to June 2015, utilizes data from the Aquarius satellite and is based on Optimum Interpolation analysis (OI SSS; Melnichenko et al., 2016). The analysis is produced on a 0.25-degree grid at a 4-day interval and uses a dedicated bias-correction algorithm to correct the satellite retrievals for large-scale biases with respect to in-situ data. The time series is continued with the SMAP satellite-based SSS data provided by Remote Sensing Systems (RSS). SMAP SSS fields are produced from Level-2 (swath) data using the OI algorithm. To ensure consistency and continuity in the data record, SMAP SSS fields are further adjusted using a set of optimally designed spatial filters to reduce small-scale noise and, at the same time, to ensure that the data record is consistent across the scales. For the overlap period (April-May 2015), the data from the two satellites are averaged together to ensure a smooth transition from one dataset to another. Measurements from ESA's Soil Moisture and Ocean Salinity (SMOS) satellite are used to fill the gap in SMAP observations during June-July 2019, when the SMAP satellite was in a safe mode and did not deliver scientific data.

The consistency and accuracy of the new SSS dataset have been evaluated against in situ salinity from Argo floats and moored buoys. The mean root-mean-square difference (RMSD) between the Aquarius/SMAP OI SSS dataset and concurrent in-situ data globally is around 0.19 psu. The product bias is around zero.

2. Major processing steps and the algorithm flow

Figure 1 shows a schematic diagram of the Aquarius/SMAP sea surface salinity optimum interpolation analysis (OI SSS)

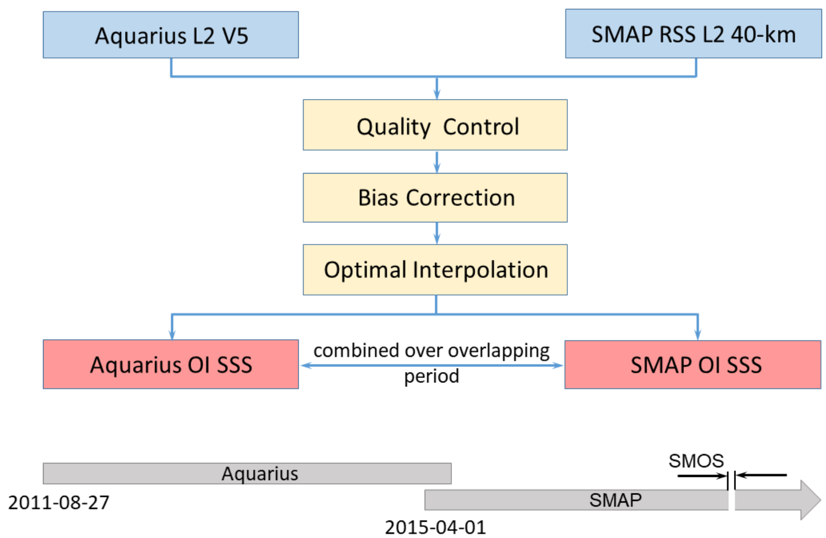

Figure 1. Schematic diagram of the Aquarius/SMAP sea surface salinity optimum interpolation analysis (OI SSS)

The input data are Level-2 SSS retrievals from the Aquarius and SMAP satellites. SMOS data are used to replace SMAP observations during 19 June – 24 July, 2019, when the SMAP satellite was not operational. The input data are described in Section 3.1. The first step in the processing algorithm is to check the Level-2 data for quality. Quality control flags and other related information (e.g., SST, surface winds) provided in the data files are used for this purpose. The quality control procedures are described in Section 3.2. The next step in data processing consists of a large-scale adjustment of the satellite SSS data relative to in-situ data. The algorithm corrects the satellite data for only persistent (time-mean) biases which are determined independently for Aquarius and SMAP observations. The bias correction algorithms are described in Section 3.3. Finally, gridded SSS fields are obtained with the Optimal Interpolation algorithm described in Section 4. Over the overlapping period (April-May 2015), the data from the Aquarius and SMAP satellites are combined together to ensure a smooth transition from one dataset to another. The output (Section 5) is a sequence of SSS fields processed with a unified algorithm.

3. Satellite SSS data, data quality control and correction for satellite biases

3.1 Satellite SSS data

3.1.1. Aquarius SSS data

The Aquarius/SAC-D satellite mission provided observations of SSS from August 2011 to June 2015. The satellite was positioned on a polar sun-synchronous orbit crossing the equator at 6 pm (ascending) and 6 am (descending) local time with a repeat cycle of 7 days. The Aquarius instrument consisted of three microwave radiometers that generated three beams at different angles relative to the sea surface. The beams had elliptical footprints on the sea surface (76 x 94 km, 84 x 120 km, and 96 x 156 km) aligned across a ~390-km-wide swath (Figure 2a). The emission from the sea surface, measured by the radiometers as an equivalent brightness temperature (Tb), was converted to SSS, subject to corrections for various geophysical effects. A detailed description of the Aquarius/SAC-D satellite mission and the Aquarius instrument can be found in Le Vine et al. (2007).

The Aquarius observations of SSS have been obtained from Level-2 (L2) version 5.0 (end-of-mission) Aquarius data produced by the NASA Goddard Space Flight Center's Aquarius Data Processing System (ADPS). The L2 data files, distributed by the Physical Oceanography Distributed Active Archive Center (PO.DAAC) of the Jet Propulsion Laboratory (JPL), contain retrieved SSS, navigation data, ancillary fields, quality flags, and other related information such as surface winds. The data are structured as a sequence of files, each corresponding to one orbit of Aquarius. An orbit is defined as starting when the satellite passes the South Pole. Individual observations along each orbit consist of a sequence of data points sampled at a 1.44-second (~10 km) interval. Each individual observation represents the average salinity in the upper 1-2 cm layer and over a ~100 km footprint (Le Vine et al., 2007; Lagerloef et al., 2008). A detailed description of Aquarius data can be found in the Aquarius User Guide (Aquarius Dataset Version 5.0). The retrieval algorithm is descried in Meissner et al. (2017, 2018).

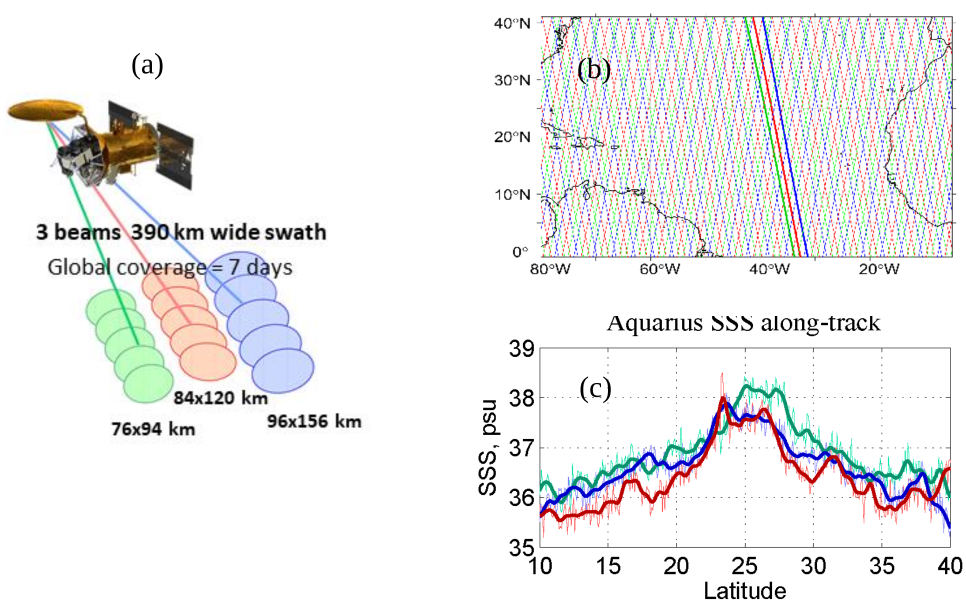

Figure 2. (a) Aquarius measurement geometry. (b) Example pattern of Aquarius ground tracks over the North Atlantic over a 7-day period. Colors indicate the three Aquarius beams. Ascending passes are from southeast to northwest. (c) Example of Level-2 SSS (three beams; 390-km-wide swath) passing through the subtropical North Atlantic on September 14, 2012 (thick lines in (b)). Thin curves – raw data; thick curves – smoothed with a running Hanning filter of half-width of ~60 km (approximately half-width of the Aquarius footprint). Colors indicate the three Aquarius beams.

An example of Aquarius L2 SSS data is shown in Figure 2. Figure 2 demonstrates that there are at least two types of errors in the SSS retrievals. One significant source of error is the accuracy of individual measurements along the satellite tracks. The instrument noise is essentially ‘white’ in nature and can be suppressed by filtering the data along track such as shown in Figure 2c (heavy lines). Of much greater concern are differences between the three beams, which can be as large as 0.5-0.8 psu and appear to be correlated over large distances along the satellite tracks. The inter-beam biases are likely a manifestation of residual geophysical corrections.

3.1.2. SMAP SSS data

NASA's SMAP satellite, launched on January 31, 2015, started collecting SSS observations in April 2015, overlapping with Aquarius observations for about three month period (April-June 2015). The satellite is positioned on a polar sun-synchronous orbit crossing the equator at 6 pm (ascending) and 6 am (descending) local time with a repeat cycle of 8 days. Similar to Aquarius, the measurement principle of SMAP is based on the response of the L-band sea surface brightness temperature (Tb) to SSS. The measuring instrument is a large rotating antenna which provides Tb observations within approximately 1000-km wide swath with nominal resolution of about 40-km and a near global coverage in 3-4 days (Figure 3a).

SMAP observations of SSS have been obtained from Level-2 version 4.0 SMAP data produced by the Remote Sensing Systems (RSS; www.remss.com). SSS is retrieved on a 0.25^{o} Earth grid using the 40-km spatial resolution Backus Gilbert optimum interpolation from the original Level-1 footprint measurements (Meissner et al., 2019). The L2 data files contain retrieved SSS (variable ‘sss_smap_40km’), navigation data, ancillary fields, quality flags, and other related information such as surface winds, sea surface temperature, etc. Each file corresponds to one orbit of SMAP. A detailed description of SMAP data and the retrieval algorithm can be found in Meissner et al. (2018, 2019). An example of SMAP L2 SSS data is shown in Figure 3b.

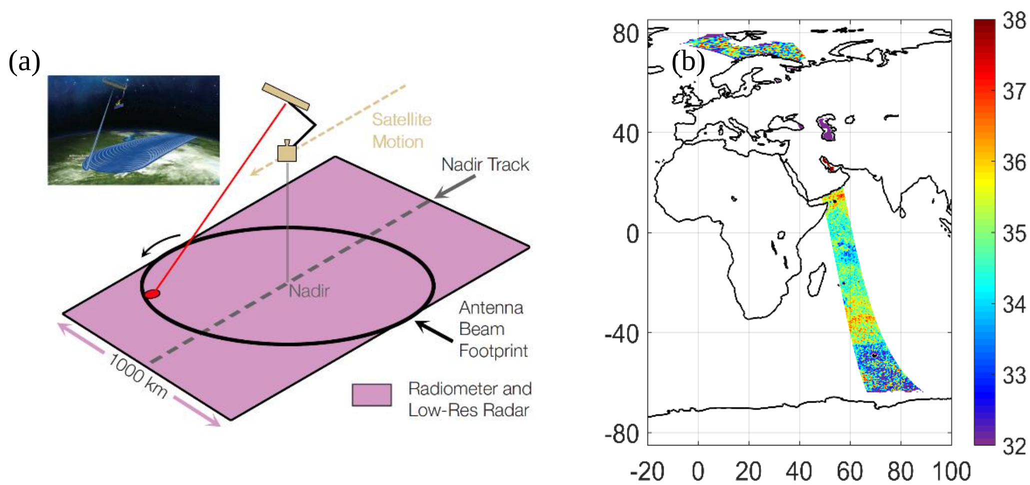

Figure 3. (a) SMAP measurement geometry. Credit: NASA. (b) Example of SMAP Level-2 SSS for the orbit passing through the Indian Ocean on April 1, 2015.

3.1.3. SMOS SSS data

In the current version of the multi-satellite SSS dataset, SMOS data are used only to fill the gap in SMAP observations during the period from June 19 to July 23, 2019, when the SMAP satellite went into a safe mode and data collection was disrupted.

The SMOS satellite, launched on November 2, 2009, operates on a sun-synchronous polar orbit crossing the equator at 6 am (ascending) and 6 pm (descending) local time. The measuring instrument is MIRAS (Microwave Imaging Radiometer using Aperture Synthesis), a two-dimensional L-band interferometric radiometer, which consists of an array of 69 receivers arranged in a Y-shape structure. The instrument provides measurements of Tb in an approximately 1000-km wide swath with spatial resolution of ~45 km and revisit time of 3-5 days.

SMOS observations of SSS have been obtained from the SMOS Level-2 SSS data products generated by version 662 of the Level 2OS Operational Processor (L2OS). SSS is retrieved on a 25-km Equal-Area Scalable Earth (EASE) grid from Tb recorded by the MIRAS radiometer. The L2 data are structured as a sequence of files, each file containing half-orbit data (from pole to pole). The files contain retrieved SSS, navigation data, ancillary fields (surface winds, sea surface temperature, etc.), quality flags, and other related information. A detailed description of SMOS data and the retrieval algorithm can be found in SMOS L2 OS ATBD, https://earth.esa.int/documents/10174/1854519/SMOS_L2OS-ATBD. The data are available from the ESA SMOS online dissemination service at https://smos-diss.eo.esa.int/oads/access/.

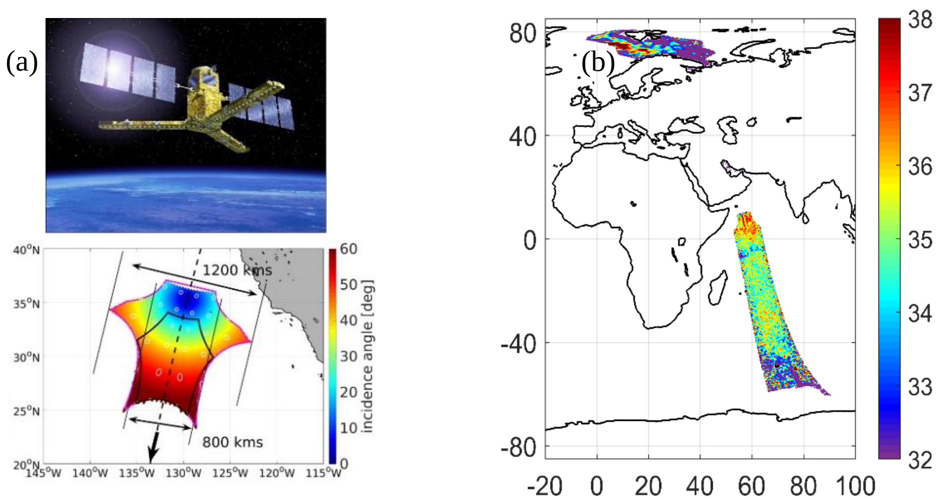

Figure 4. (a) SMOS measurement geometry. Bottom panel shows the shape of a Tb image as reconstructed from SMOS observations. Colors show incidence angle. Credit: ESA and Reul et al. (2020). (b) Example plot of SMOS Level-2 SSS for the orbit passing through the Indian Ocean on August 26, 2011.

An example of SMOS L2 SSS data is presented in Figure 4. Individual measurements along the orbit are very noisy. Retrievals near the edge of the swath typically have larger uncertainties due to a smaller number of observations and contaminations from various sources (Reul et al., 2017). SMOS SSS observations are also subject to significant large-scale biases which arise due to contamination from land or sea-ice emission, which are visible as far as 1000 km from the coastline or sea-ice edge, radio frequency interference (RFI), uncertainties in the retrieval algorithm, and other sources (Boutin et al., 2018). Biases also depend on the orbit orientation. Ascending orbits typically have smaller biases than descending orbits, particularly during January-March and October-December periods (Reul et al., 2017).

3.2. Data quality control

Level-2 data from each of the three satellites were first processed to identify and quality control (Q/C) issues. This Q/C process was different for each sensor, as described below.

3.2.1. Aquarius SSS data

Observations are discarded if the bits for any quality checks are set: 7 (direct solar flux contamination), 8 (reflected solar flux contamination), 9 (sun glint), 12 (non-nominal navigation), 13 (radiometer telemetry), 14 (roughness correction failure), 16 (pointing anomaly), 17 (brightness temperature consistency), 19 (radio-frequency interference (RFI)), and 21 (reflected radiation from Moon or Galaxy). In the case of flags 19 and 21, the data are excluded from the analysis if the conditions indicated by the flags are either moderate or severe. For other flags, only severe conditions are taken into account. Also excluded from the analysis are data points that are contaminated by land (land fraction > 0.01), sea ice (sea ice fraction > 0.0025), sampled during high wind (wind speed > 18 m/s) and/or in cold water (SST < 0^{o}C). A detailed description of the Aquarius quality flags including recommended thresholds can be found in the Aquarius User Guide (Aquarius Dataset Version 5.0).

3.2.2. SMAP SSS data

Observations are discarded if the bits for any of the following quality checks are set: the sun glint (bit 5 in Q/C flag, Table 4 in Meissner et al., 2019), moon glint (bit 6), reflected galactic radiation (bit 7), and Tb consistency (bit 10). Also excluded from the analysis are data points that are contaminated by land (gain weighted land fraction > 0.008 or land fraction in 3-dB footprint >0.0005), sea ice (sea ice fraction > 0.0025), sampled during high wind (wind speed > 18 m/s) and/or in cold water (SST < 0^{o}C). A detailed description of the SMAP quality flags can be found in Meissner et al. (2019).

3.2.2. SMOS SSS data

Observations are discarded if the following Quality Flags are set (Table 4-20 in SMOS Level 2 and Auxiliary Data Products Specifications document): outside range (bit 2), high retrieval sigma (bit 3), poor fit quality (bits 4 and 5), sun glint (bit 7), moon glint (bit 8), high galactic noise (bit 9), low number of measurements (bit 13), too many outliers (bit 14), and high Marquardi increment (bit 15).

Observations are discarded if the following Science Flags are set (Table 4-21 SMOS Level 2 and Auxiliary Data Products Specifications document): high TEC gradient (bit 3), grid point with maximum extend of sea ice (bit 4), high ice concentration (bit 5), suspect ice in grid point (bit 6), and high rain rate (bit 7). Also excluded from the analysis are data points that are near the edge of the swath (x_swath > 400 km), near a coastline (distance to the nearest coast < 75 km), measured during high wind (wind speed > 15 m/s) and/or in cold water (SST < 0^{o}C). A detailed description of the SMOS quality flags can be found in the SMOS Level 2 and Auxiliary Data Products Specifications document (https://earth.esa.int/documents/10174/1854583/SMOS_L2_Aux_Data_Product_Specification).

3.3. Bias correction

The next step in data processing consists of a large-scale adjustment of the satellite data relative to in-situ data. Only static (time-mean) biases are taken into account.

Generally, the bias-adjusted satellite observations S_{adj} are determined from the retrieved values S_{obs} as

$$S_{adj} = S_{obs} - \Delta S \,, \tag{1}$$

where the bias ΔS is determined by interpolating the bias fields into the locations of the satellite measurements. Specific procedures, however, are slightly different for Aquarius, SMAP and SMOS data.

3.3.1. Aquarius SSS data

Analysis of long time series of Aquarius SSS data indicate that satellite retrievals have large-scale biases relative to in-situ observations (Kao et al., 2018). The causes of the biases in Aquarius SSS data are only partially understood, but may be related to SST-dependent errors in the dielectric constant and the model for atmospheric absorption, which are part of the retrieval algorithm (Meissner et al., 2017; 2018).

In the OI SSS analysis, the large-scale biases in satellite SSS are corrected with respect to in-situ salinity data. To adapt to the Aquarius measurement geometry, the bias fields are constructed on a repeat track basis. To construct the bias fields, satellite observations along each repeat track are averaged over a 3-year period from September 2011 through August 2014 and compared to in-situ salinity averaged over the same period. The in-situ salinity, which we regard as the “ground truth” at large spatial scales, is a compilation of four Argo-based products. These products are:

- APDRC of the University of Hawaii Argo-derived salinity product (http://apdrc.soest.hawaii.edu/projects/Argo/data/gridded/On_standard_levels/index-1.html);
- Scripps Institution of Oceanography Argo-derived salinity product (http://sio-argo.ucsd.edu/RG_Climatology.html; Roemmich, D. and J. Gilson, 2009);
- Met Office Hadley Center objective analysis from the profile data, version EN.4.2.1 (http://hadobs.metoffice.com/en4/index.html; Good et al., 2013); and
- ISAS-15 salinity gridded fields (http://www.seanoe.org/data/00412/52367/; Kolodziejczyk et al., 2017).

The average of the four products is assumed to better represent the 'ground truth' (at large spatial scales) provided that the mapping errors of the products are not correlated. The 'ground truth' was interpolated into the ground track locations; thus, there are two bias fields, one for ascending and one for descending ground tracks. The bias fields are shown in Figure 5 (note that the bias fields are constructed on a specific (irregular) grid, which corresponds to the ground track segments). Correcting for the large-scale satellite biases on a repeat track basis separately for each of the three Aquarius beams helps eliminate residual inter-beam biases which otherwise persist even after applying a multi-year average.

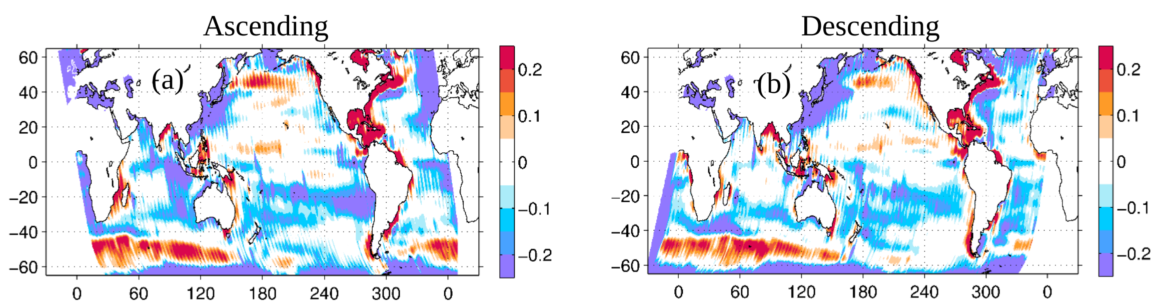

Figure 5. Mean spatial bias correction fields (psu) for Aquarius ascending (a) and descending (b) SSS data.

3.3.2. SMAP SSS data

Persistent large-scale biases have been present in all versions of SMAP SSS data and are characterized as biases that manifest in the long-term mean. To construct the bias field, satellite observations at each grid point were averaged over a 3-year period from April 2015 through March 2018 and compared to the 'ground truth' averaged over the same period. The "ground truth" was assessed as a compilation of the following four Argo-based products:

- APDRC of the University of Hawaii Argo-derived salinity product (http://apdrc.soest.hawaii.edu/projects/Argo/data/gridded/On_standard_levels/index-1.html);

- Scripps Institution of Oceanography Argo-derived salinity product (http://sio-argo.ucsd.edu/RG_Climatology.html; Roemmich and Gilson, 2009);
- Met Office Hadley Center objective analysis from the profile data, version EN.4.2.1 (http://hadobs.metoffice.com/en4/index.html; Good et al., 2013); and
- JAMSTEC MOAA GVP global gridded salinity product produced by optimal interpolation of all available observations including ARGO (Hasoda et al., 2008; http://www.jamstec.go.jp/ARGO/argo_web/argo/?page_id=223&lang=en).

The SMAP bias field is shown in Figure 6. As there are almost no ascending-descending differences, the bias field is the same for ascending and descending data.

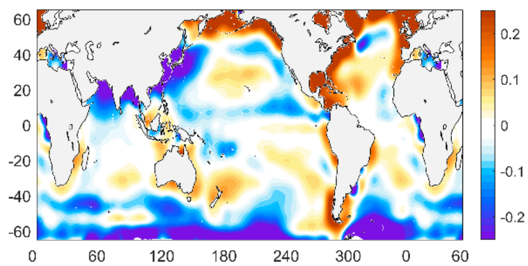

Figure 6. Mean spatial bias correction field (psu) for SMAP SSS data.

3.3.2. SMOS SSS data

As stated in section 3.1.3, SMOS SSS data are used only to fill the gap in SMAP observations from June 19 to July 23, 2019. Because the biases in SMOS satellite observations are of a very complex nature and because of a relatively short duration of the period over which SMOS data are used in the analysis, the bias correction is not applied explicitly to SMOS observations.

Instead, the impact of the large-scale biases in SMOS SSS data is reduced using a multi-scale filtering approach (Sakurai et al., 2005). Specifically, the SSS observations collected over a 4-day period are split into two layers, a large-scale part and a high-resolution part (including noise), by applying a low-pass filter. The filter is a 2D Hanning window with cut-off wavelength adjusted to match the spatial resolution of a first guess filed. The large-scale part is discarded and the high-resolution part, extracted from SMOS SSS observations, is used to correct the first guess field in the OI algorithm.

3.4. Filtering

The final step in data preparation consists of additional filtering to remove outliers and reduce noise.

3.4.1. Aquarius SSS data

Aquarius SSS data are filtered along track as described in Melnichenko et al. (2014). The filter is a subsequent application of a 5-point Median filter and 6-point Hanning filter, which has been found to perform quite efficiently to considerably reduce high-frequency instrument noise, yet preserve the ocean signal from over-smoothing. An example is presented in Figure 1b. According to the degree of filtering, the SSS data are then sub-sampled every third point along track.

3.4.2. SMAP SSS data

A standard statistical test based on the standard deviation (STD) is applied to SMAP SSS data. Data points are rejected if the SSS anomalies, determined relative to the first guess, exceed 5, 4 and 3 STDs in the areas of low variability, STD<0.1, moderate variability, 0.1<STD<0.2, and high variability, STD>0.2 psu, respectively. The geographical distribution of the standard deviation of SSS for this analysis is obtained from weekly time series of Aquarius OI SSS fields for the period from September 2011 to June 2015 (first segment in the Aquarius/SMAP data record).

3.4.2. SMOS SSS data

SMOS SSS data are processed in the same way as SMAP SSS data (sec. 3.4.2).

4. Optimum Interpolation (OI) algorithm

4.1 General description

The interpolation expression for OI with N observations can be written as (Bretherton et al., 1976; Le Traon et al., 1998):

$$\hat{S}_x = S_x^0 + \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}^{-1} C_{xj} (S_i^{obs} - S_i^0)\,, \qquad (2)$$

where $\hat{S}_x$ is the interpolated value (or estimate) at the grid point $\mathbf{x}$; S_x^0 is the forecast (or "first guess") value at the grid point $\mathbf{x}$; S_i^{obs} is the measured value at the observation point i: $S_i^{obs} = S_i + \varepsilon_i$, where ε_i is random measurement error; S_i^0 is the forecast value at the observation point i; $\mathbf{A}$ is the $N \times N$ covariance matrix of the data

$$A_{ij} = < (S_i - S_i^0)(S_j - S_j^0) > + < \varepsilon_i \varepsilon_j >; \qquad (3)$$

and $\mathbf{C}$ is the joint covariance of the data and the field to be estimated

$$C_{xj} = < (S_x - S_x^0)(S_j - S_j^0) > . \qquad (4)$$

In (3) and (4), it is assumed that the errors and the field are not correlated.

The OI analysis is determined relative to the first guess field, which is assumed to be a good approximation of the true state. The estimate and the observations are then equal to the first guess plus small increments. In this way, the grid point analysis consists of interpolation of the first-guess field to the observation points followed by interpolation of the differences between the observed and first-guess values back to the grid point. The grid point analysis is completed by adding the analysis increment to the first guess.

4.2. Specifics

The OI method assumes that the first guess and statistics of the field to be analyzed are known a priory. These parameters are the following.

4.2.1. First guess

The first guess fields are assessed from a compilation of four Argo-derived products (monthly-mean SSS fields). The products are the same as have been used to evaluate the satellite biases (see section 3.3.1 and 3.3.2 for a list of products). An example for the first week of September 2011 is presented in Figure 7. The Argo-derived SSS fields are chosen because they are independent of the analysis of the satellite data and provide an unbiased estimate of the first guess.

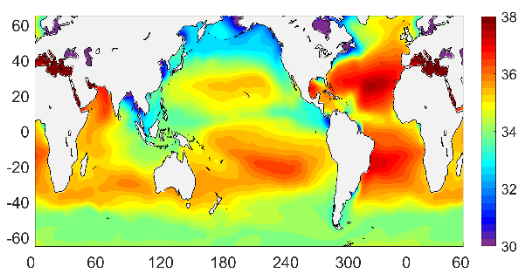

Figure 7. First guess SSS field for the week September 2-8, 20011.

4.2. Signal statistics

The normalized spatial covariance of SSS anomalies is described by the Gaussian function of the form

$$C(r_x, r_y, t) = \exp(-r_x^2 / R_x^2 - r_y^2 / R_y^2 - t^2 / T^2)\,, \qquad (5)$$

where r_x and r_y are spatial lags in the zonal and meridional directions, respectively, t is time lag, R_x and R_y are the zonal and meridional correlation scales, and T is the correlation time scale. This particular form of the correlation structure is chosen because the associated spectrum is positive everywhere and because the resulting covariance matrixes are always positive definite (Weber and Talkner, 1993), which is a strict requirement on the choice of a possible analytical form of the correlation function in the OI analysis (Bretherton et al., 1976).

The zonal and meridional correlation scales in Eq. (5) are allowed to vary with latitude. The meridional scales have been determined by fitting the Gaussian model to the sample covariances estimated in 10° latitude bins from the Aquarius L2 data as described in Melnichenko et al. (2014). Based on the observed structure (Figure 8), the latitudinal dependency of R_y [km] is modeled by the following functional form

$$R_y(y) = 14\exp(-(y-4)^2/225)) + 92, \qquad (6)$$

where y is latitude in degrees. Thus, the meridional scales are somewhat larger in the tropics (106 km at 4°N) than at high latitudes (92 km).

The zonal correlation scales at mid- and high latitudes are set to equal the meridional scales, while in the tropics they are scaled to represent the zonal elongation of correlation as follows

$$R_x(y) = R_y(y)(0.5\exp(-(y-4)^2/56.25)+1). \qquad (7)$$

Near the equator, the aspect ratio R_x/R_y equals 1.5 ($R_x =$ 160 km at 4°N) and gradually decreases toward higher latitudes (Figure 8). Poleward of about 20°, the correlation function (5) becomes isotropic ($R_x = R_y = 92$ km). We note, however, that our assumptions of the zonal correlation scales are somewhat arbitrary and are mostly based on previous observational studies (e.g., Delcroix et al., 2005; Reverdin et al., 2007).

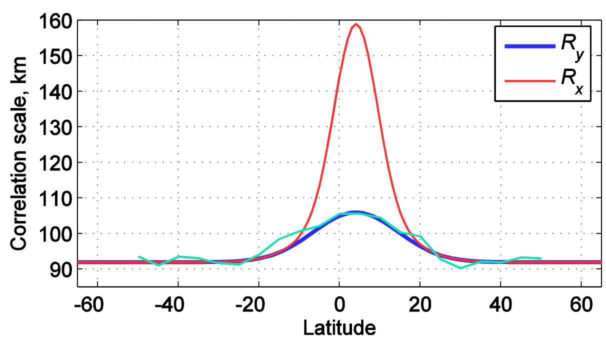

Figure 8. Meridional (blue) and zonal (red) correlation scales applied in the OI SSS analysis. The green curve shows the along track correlation scales determined by fitting the Gaussian model to the sample covariances estimated in 10° latitude bins from the Aquarius L2 data as described in Melnichenko et al. (2014).

The temporal correlation scale is set to T =7 days. This provides a smooth map-to-map transition while preventing the time series from over-smoothing.

4.3. Error statistics

4.3.1. Aquarius SSS data

Analysis of Aquarius along track SSS data reveals that there are long-wavelength errors, referred to here as inter-beam biases, which are correlated over long distances along the satellite tracks. To incorporate statistical information on these errors into the OI scheme, the following error covariance model for the Aquarius data is introduced (Melnichenko et al., 2014):

$$< \varepsilon_i \varepsilon_j >= \delta_{ij}\sigma_w^2 + \sigma_L^2$$ -if data points i, j are on the same track and beam, and in the same cycle, and

$$< \varepsilon_i \varepsilon_j >= \delta_{ij}\sigma_w^2$$ -otherwise,

where δ_{ij} is the Kronecker delta, σ_w^2 is the variance of the uncorrelated (white) noise, and σ_L^2 is the variance of the long-wavelength (along-track) error.

Given prior filtering of Aquarius L2 SSS data (section 3.4.1), the variance of the white noise is assumed to be 10% of the signal variance, independent of the geographical location. The long-wavelength error correlation structure is represented by the exponential function of the form

$$C_L(l) = \exp(-l / R_L) \tag{8}$$

where l is the along track separation distance and $R_L = 500$ km is the exponential decay scale. The estimate of R_L is obtained by fitting the curve (8) to the inter-bean bias statistics evaluated by comparison of the covariances of the inter-beam differences for Aquarius and ancillary SSS data as described in Melnichenko et al. (2014).

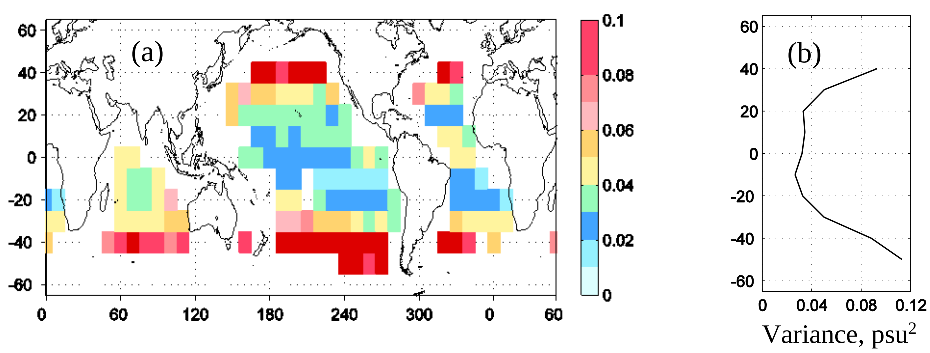

Figure 9. (a) Variance of long-wavelength error (psu^2) in 20º longitude x 20º latitude boxes and (b) the zonal average of the variance.

The variance of the long-wavelength error varies with latitude from about 0.04 psu^2 in the tropics to 0.1 psu^2 at high latitudes (Figure 9). Following the latitudinal changes in both the error and signal variances (not shown in figure), the ratio of the error variance to the signal variance, η, is approximated by the following analytical curve

$$\eta = 2*(1-\exp(-y^2/400))/1.43)+0.3\,. \qquad (9)$$

Thus, the relative variance of the long-wavelength error is set to vary from 30% in the equatorial region, where the signal variance is large, to more than 150% at high-latitudes, where the error variance is large.

4.3.1. SMAP SSS data

The error covariance matrix consists of one part, $<\varepsilon_i\varepsilon_j>=\sigma_w^2$, and represents uncorrelated errors. The variance of the uncorrelated error is assumed to be 50% of the signal variance, independent of the geographical location.

4.3.1. SMOS SSS data

Similar to SMAP, random errors in SMOS SSS data are assumed to be uncorrelated. The variance of the uncorrelated error is assumed to be 50% of the signal variance, independent of the geographical location.

4.4. Implementation

The OI SSS analysis is performed on a 0.25^o longitude x 0.25^o latitude grid at a 4-day interval starting from September 2011. The OI SSS analysis is run in a local approximation; that is, only data points in a smaller sub-domain around the analysis grid point are used. The radius of the sub-domain is defined to be 4 times the spatial correlation scale and ± 7 days, which allows for accommodating both the signal and error correlation (in the case of Aquarius data; section 4.3.1). The local approximation also helps reduce the effect of spatial inhomogeneity in the signal and error statistics (Weber and Talkner, 1993). Likewise, to reduce the computational load, SMAP and SMOS data are grouped into 4-day intervals as shown schematically in Figure 10. At each time step t=t_0, the OI SSS analysis utilizes data points at t=t_0-4 days, t=t_0, and t=t_0+4 days.

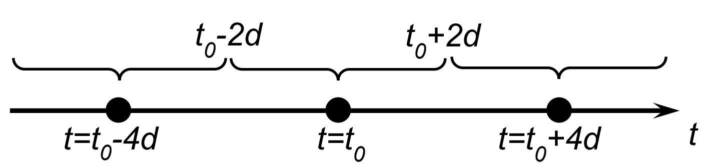

Figure 10. SMAP and SMOS data processing. Observations are grouped into 4-day intervals and time averaged within each group. OI SSS analysis at time t=t_0 utilizes data points at t=t_0-4 days, t=t_0, and t=t_0+4 days.

Additionally, SMAP (SMOS) OI SSS fields are produced in two steps. The observations at each time interval (Figure 10) are subsampled every other data point (even and uneven indexes) and two SSS maps are produced from the subsampled data using the OI algorithm. The final map is then the average of the two. This procedure helps reduce noise in the SSS maps and also match the spatial resolution of SMAP (SMOS) OI SSS analysis to that of Aquarius. In this regard, a relatively low level of noise in SMAP (SMOS) SSS maps is achieved at the expense of reduced spatial resolution.

4.5. Uncertainty estimate

The uncertainty is estimated empirically by comparing weekly SSS maps with concurrent Argo buoy data. Argo buoy salinity measurements in the near-surface layer (depth <10 m) are assumed to represent in-situ SSS. The error statistics are computed by comparing buoy measurements for a given week with SSS values at the same locations obtained by interpolating the corresponding SSS maps. Uncertainties are estimated in 8^o-longitude x 8^o-latitude bins as the RMSDs between the SSS maps and the corresponding buoy data. The coarse resolution RMSD map is then interpolated onto the analysis grid to provide estimates of the uncertainty. In the areas lacking buoy observations, the uncertainties are estimated by extrapolating from adjacent regions. The estimated uncertainty includes the so-called sampling error which arises due to unresolved small-scale SSS variability (see Sec. 5.2)

5. Global OI SSS fields

5.1. Spatial coverage and resolution

Figure 11 presents example plots of Aquarius and SMAP OI SSS in several regions. The plots in the center (Figures 11a, b) are global and show the product spatial coverage. Aquarius/SMAP OI SSS dataset covers the full global ocean including the Arctic and Antarctic in the areas free from ice. The coverage includes coastal areas and marginal seas, such as the South China Sea and the Gulf of Mexico, but does not include internal seas, such as the Mediterranean and the Baltic Sea, which require special treatment.

The resolution capabilities of the Aquarius/SMAP OI SSS analysis can be inferred from the regional maps shown in Figure 11. In particular, Figures 11c (Aquarius) and 11d (SMAP) show zooms on a large area in the North Atlantic. Among the many features represented in the maps is a frontal structure associated with the Gulfstream and Gulfstream Extension, which separates low-salinity slope water from salty Sargasso Sea. The front extends further north into the Labrador with local salinity gradients as large as 1 psu/100 km. Both the Aquarius and SMAP maps show a double front near where the Gulfstream separates from the coast (the area is circled in the figures).

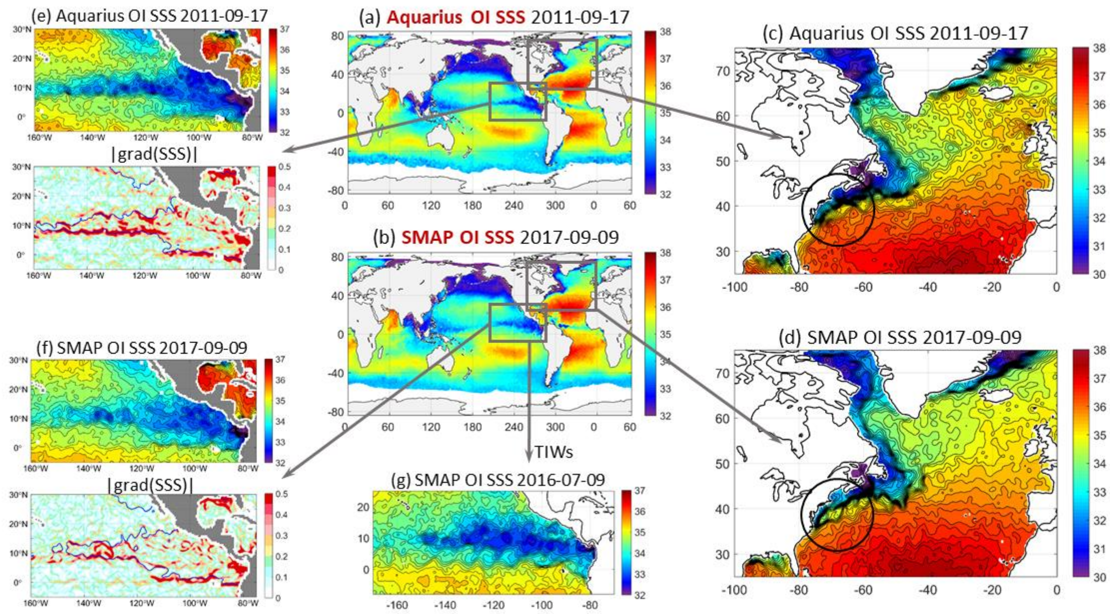

Figure 11. Example plots of Aquarius/SMAP OI SSS.

Another example is in the eastern tropical Pacific. Figures 11e (Aquarius) and 11f (SMAP) show the SSS field for one week in September 2011 and September 2017, respectively. September is time of year when the Intertropical Convergence Zone (ITCZ) is at its northernmost annual position and the eastern Pacific fresh pool (EPFP), delineated by the 34 psu isohaline, reaches its greatest westward extent at around 170^{o}W (Melnichenko et al., 2019). The resolution capabilities of the OI SSS maps are illustrated by the corresponding maps of the SSS gradient. Embedded in a fine-scale structure of SSS fronts in the region is a sharp SSS front along the southern boundary of the low salinity belt along ~10°N. Another prominent feature is a sharp front along the equator, which originates from the coast of South America and extends westward to nearly 120°W, delineating the boundary between the fresh pool (SSS < 34 psu) to the north and saltier water (SSS > 34 psu) to the south. The gradients across the front are as large as 0.7psu/100 km.

Finally, Figure 11g shows the SSS signature of Tropical Instability Waves (TIWs) in the eastern tropical Pacific, clearly seen as cusp-like features between ~0° and 5°N with wavelength of ~1,000 km (~10° of longitude). The waves have a dominant period of about 30 days and propagate westward at a speed of about ~0.5 m s^{-1} (not shown in figure). These examples show that the spatial resolution of the Aquarius/SMAP OI SSS analysis allows observing large mesoscale features and fronts.

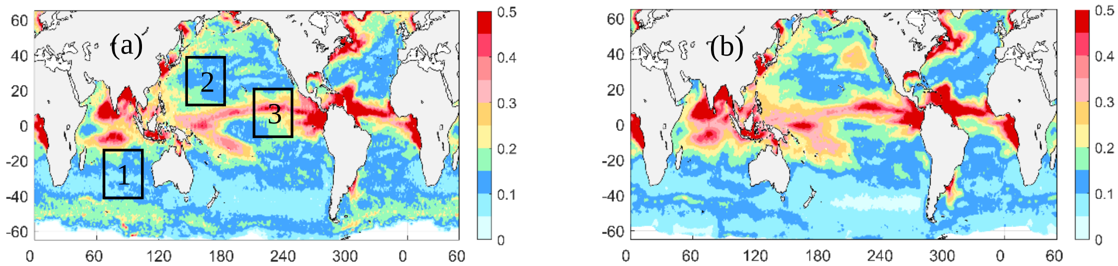

Figure 12. The standard deviation of SSS computed from weekly time series of OI SSS fields for the period (a) September 2011 through May 2013 and (b) June 2015 through November 2020. Black boxes show regions selected for the spectral analysis.

To quantify the magnitude and spatial distribution of SSS variability in the OI SSS analysis, Figure 12 shows the standard deviation of SSS computed from the time series of weekly OI SSS maps over two periods. The first period is from September 2011 through May 2015 and consists of Aquarius OI SSS maps (Figure 12a). The second period is from June 2015 through November 2020 and consists of SMAP OI SSS maps (Figure 12b). Comparing the maps of the standard deviation, we can see that the two time periods, based on different satellite data, are consistent by this metric. Apparently, there are no spurious jumps or differences in the SSS variance between the two segments. Several regions stand out as having the largest SSS variability: the rainy belts (low local salinity) associated with the ITCZ in the North Pacific and North Atlantic (standard deviations around 0.3-0.5 psu), the South Pacific convergence zone, the eastern equatorial Pacific, the tropical Indian Ocean, the western boundary current regions of the Kuroshio and Gulfstream, the Southern Ocean, as well as the areas near outflows of major rivers, such as the Amazon. Maximum values of the SSS standard deviation, exceeding 1 psu, are observed in the far eastern equatorial Pacific, the Bay of Bengal, and the western part of the tropical North Pacific. Apart from the regions of high SSS variability, the standard deviation of SSS has typical values of around 0.1-0.15 psu.

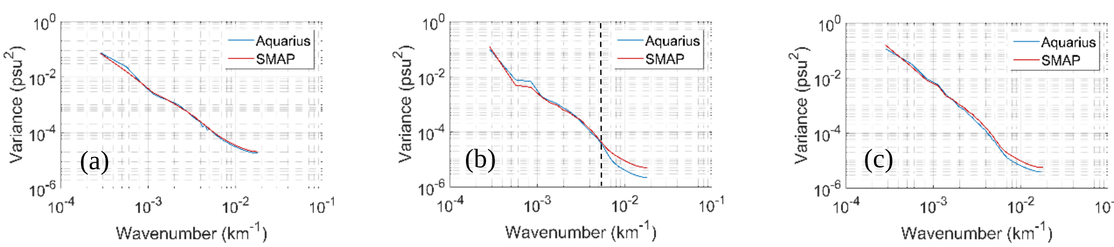

Figure 13. Meridional wavenumber spectra of SSS computed from the weekly Aquarius (blue) and SMAP (red) OI SSS fields in (a) South Indian Ocean (box 1), (b) western North Pacific (box 2), and (c) eastern tropical Pacific (box 3). The vertical dashed line in (b) corresponds to wavelength of 560 km.

The consistency between the two segments in the Aquarius/SMAP OI SSS dataset is further verified by comparing the associated wavenumber spectra. Examples for a few selected regions are shown in Figure 13. The examples demonstrate that the Aquarius and SMAP OI SSS spectra are very similar in shape with nearly equal distribution of variance. For wavenumbers higher than about 0.002 km^{-1}, the spectra start to quickly roll off, from which we conclude that the effective resolution of the product is about 500-600 km in terms of a wavelength (length scale larger than about 125 km).

5.2. Validation

Salinity from Argo buoy observations in the near-surface layer are used to estimate the error statistics for the OI SSS analysis. The Argo buoy network provides quasi-random geographical distribution of about 1100 in-situ salinity measurements for each week. Only measurements shallower than 10 m depth and flagged as good from each Argo profile are used in this analysis. The error statistics for the OI SSS analysis are calculated by comparing buoy measurements for a given week with SSS values at the same locations obtained by interpolating the corresponding SSS maps.

Figure 14 shows the time series of the bias (average of the differences between the product and buoy data over all buoy locations) and root-mean-square difference (RMSD) of the Aquarius/SMAP OI SSS analysis evaluated against concurrent Argo buoy observations. The product yields the time-series of the global bias oscillating around zero (Figure 14a). The standard deviation of the weekly biases is 0.008 psu. The RMSD between the OI SSS analysis and concurrent buoy data is oscillating around 0.2 psu (Figure 14b). The mean RMSD of the analysis over the period September 2011 – November 2019 is 0.19 psu.

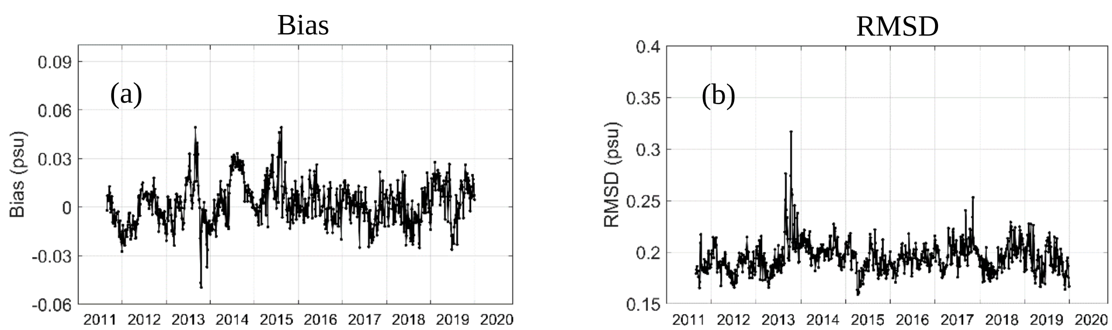

Figure 14. (a) Weekly mean differences and (b) RMSD between Argo buoy data and Aquarius/SMAP OI SSS analysis. The error statistics are computed by comparing Argo buoy measurements for a given week with SSS values at the same locations obtained by interpolation of the corresponding OI SSS maps.

The effect of the bias correction and a smooth transition from Aquarius to SMAP in the Aquarius/SMAP OI SSS dataset can be seen in the zonally averaged differences between the weekly OI SSS maps and the corresponding buoy data shown in Figure 15. The zonally averaged biases were calculated weekly by averaging these statistics over 5-degree latitude bands. The bias distribution for the OI SSS fields shows nearly zero bias throughout the 10-year period of comparison. Very small residual biases (<0.05 psu), varying with a seasonal cycle, can be observed at high latitudes during 2011-2015 and in the tropical belt during 2015-2020.

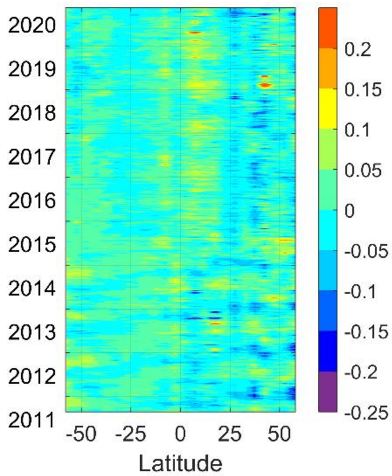

Figure 15. Latitude-time distribution of the zonally averaged differences (psu) between the weekly OI SSS maps and the corresponding Argo buoy data. The error statistics were computed by comparing Argo buoy observations for a given week with SSS values at the same locations obtained by interpolation of the corresponding SSS maps. The zonally averaged biases were computed by averaging these statistics over 5-degree latitude bands.

The geographical distribution of the RMS error for the OI SSS analysis is shown in Figure 16 separately for the two time periods, covering different satellite data. The RMS error was computed in 8-degree spatial bins from the differences between the weekly SSS maps and the corresponding in-situ observations. The bin size was selected to ensure an adequate number of collocations (>100) in each bin. Figure 16 demonstrates that the error statistics are very similar for the two segments of the analysis.

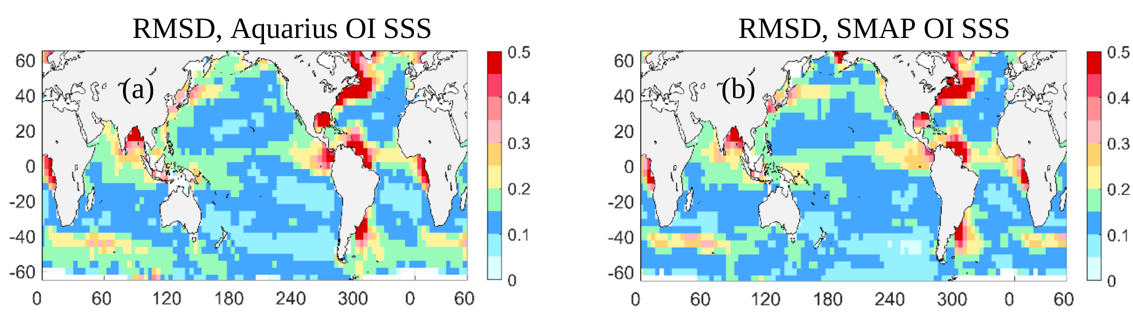

Figure 16. Geographical distribution of the RMS differences (psu) between the weekly OI SSS analysis and in-situ buoy data over the period (a) September 2011 through May 2013 (Aquarius OI SSS) and (b) June 2015 through November 2020 (SMAP OI SSS). The error statistics are computed in 8^{o} bins by comparing Argo buoy measurements for a given week with SSS values at the same locations obtained by interpolation of the corresponding OI SSS maps.

The largest RMS errors, exceeding 0.2 psu, are found in the regions of strong variability in SSS (see Figure 12), such as along the North Pacific and North Atlantic ITCZ, the South Pacific convergence zone, the Gulfstream, and near the outflows of major rivers, such as the Amazon in the North Atlantic. However, these relatively large discrepancies between the satellite SSS maps and buoy data are not necessarily due to errors in satellite observations or errors in the mapping procedure. Large RMS differences between the mapped SSS and in-situ observations can be due to

1) Strong vertical gradients of salinity in the near-surface layer, such that salinity at ~5 m depth, sampled by a typical Argo buoy, differs significantly from the surface salinity, sampled by satellites. Such conditions are frequently observed in the tropics, particularly in the rainy belts associated with the ITCZ (Henocq et al., 2010).
2) Unresolved small-scale variability. In the presence of strong SSS gradients, the difference between a point measurement by a buoy and the area averaged SSS sampled by a satellite (or the grid cell of the OI SSS analysis) can readily exceed 0.2 psu (Vinogradova and Ponte, 2013).
3) Unresolved temporal variability (Vinogradova and Ponte, 2012).
4) Errors in the SSS maps.
5) All of the above.

The histogram distribution of the differences between the buoy data and OI SSS analysis over the whole period from September 2011 to November 2020 is shown in Figure 17a. The OI SSS estimates have an overall good agreement with the buoy data such that the histogram of the differences is very narrow. About 55% of the differences are smaller than 0.1 psu and more than 80% are smaller than 0.2 psu. The number of outliers, defined as the differences larger than 0.5 psu, is less than 3%. Their geographical distribution is shown in Figure 17b. The majority of 'outliers' are located in the areas of strong variability in SSS (see Figure 12), generally consistent with the distribution of sampling error (Vinogradova and Ponte, 2013; their Figure 2).

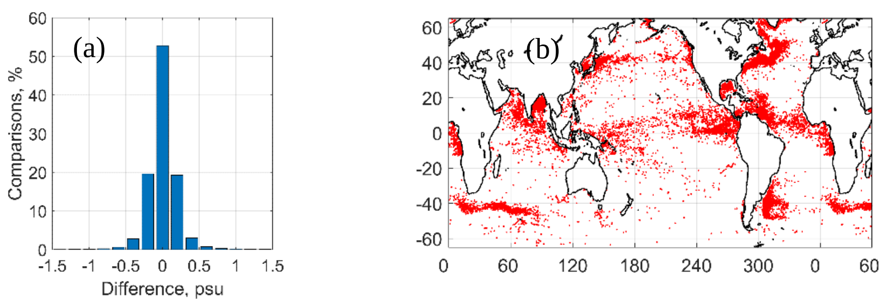

Figure 17. (a) Statistics of the differences between Argo buoy data and Aquarius/SMAP OI SSS analysis. (b) Locations of 'outliers', defined as the differences large than 0.5 psu. The error statistics are computed by comparing Argo buoy measurements for a given week with SSS values at the same locations obtained by interpolation of the corresponding OI SSS maps.

To further verify the consistency in the new dataset, the OI SSS fields are compared to time series of in-situ SSS at moored stations from the global tropical moored buoy array (https://www.pmel.noaa.gov/gtmba/). An example is presented in Figure 18. The time series show a smooth transition from one satellite to another without spurious jumps and trends. Few stations showed suspiciously large differences and RMSDs (Figure 19). Their visual inspection, however, showed (not shown in figures) that the large differences were rather due to errors in buoy measurements than errors in the satellite data.

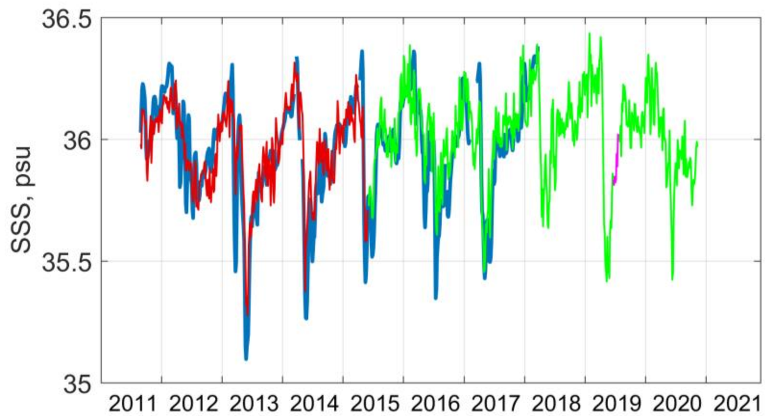

Figure 18. Time series of moored buoy SSS and Aquarius/SMAP OI SSS at buoy location 10ºW, 6ºS. The blue curve shows buoy measurements at 1-m depth (measurements are not available after 2017). The time series is smoothed with a 7-day running mean. The red and green curves are Aquarius and SMAP OI SSS, respectively. The period filled with the SMOS data is shown by magenta.

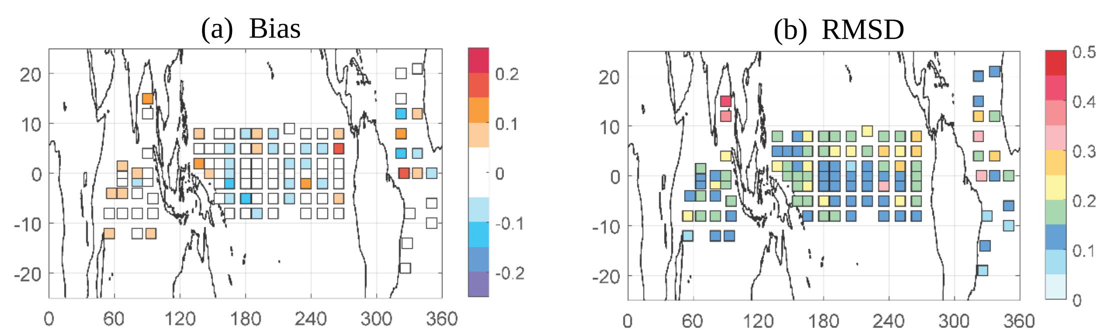

Figure 19. (a) Mean differences and (b) RMSD between moored buoy observations and the Aquarius/SMAP OI SSS. The error statistics were computed by comparing buoy observations for a given week (running mean) with SSS values at the same locations obtained by interpolation of the corresponding SSS maps.

Despite significant effort taken to reduce errors and biases in the satellite SSS data, the resulting OI SSS maps may still have errors and biases, both globally and regionally. **We encourage users to continue to validate the Aquarius/SMAP dataset and quantify**

noise and errors in their research. We are especially interested in receiving feedback on the utility of the product and potential issues relevant to noise and biases especially from regional studies, which may have access to higher quality and enhanced resolution in-situ data sets.

6. Access to the data

The Aquarius/SMAP OI SSS analysis can be accessed from the APDRC webpage http://apdrc.soest.hawaii.edu/ either through the Live Access Server or OPeNDAP.

Digital data of the weekly OI SSS analysis (netCDF files) are also available at http://iprc.soest.hawaii.edu/users/oleg/oisss/GLB/Aquarius_SMAP_OISSS/.

7. Copyright and terms of use

Use of the dataset should be acknowledged as follows:
"Aquarius/SMAP combined dataset is produced by International Pacific Research Center (IPRC) of the University of Hawaii, in collaboration with Remote Sensing Systems (RSS), Santa Rosa, CA, and sponsored by the NASA Ocean Salinity Science Team.

Reference to this technical paper:
Melnichenko, O., P. Hacker, J. Potemra, T. Meissner, and F. Wentz: Aquarius/SMAP sea surface salinity optimal interpolation analysis, IPRC Technical Note No. 7, May 7, 2021, 24p.

Comments, questions regarding the Aquarius OI SSS dataset and requests for the data can be directed to

Oleg Melnichenko
Email: oleg@hawaii.edu
Tel: 1-808-956-0747

8. Acknowledgements

This project was supported by the National Aeronautics and Space Administration (NASA) Ocean Salinity Science Team through grant NNX17AK06G and grant 80NSSC20K1338. The Argo data were collected and made freely available by the International Argo Program and the national programs that contribute to it (https://www.ocean-ops.org). The Argo Program is part of the Global Ocean Observing System. Global Tropical Moored Array data were retrieved from the GTMBA Project Office of NOAA/PMEL (https://www.pmel.noaa.gov/tao/drupal/disdel/). The authors acknowledge the many constructive dialogues with the members of the NASA Ocean Salinity Science Team.

10. References

AQUARIUS USER GUIDE, Dataset Version 5.0. Guide Version: 8.0, November 27, 2017. JPL D-70012 AQ-010-UG-0008. JPL URS CL#: 17-5944. Document accessed 2018-05-03 at http://dx.doi.org/10.5067/DOCUM-AQR01.

Bretherton, F. P., R.E Davis, and C.B. Fandry, 1976: A technique for objective analysis and design of oceanographic experiments applied to MODE-73, *Deep Sea Res.*, **23**, 559-582.

Boutin, J., J.L. Vergely, S. Marchand, F. D'Amico, A. Hasson, N. Kolodziejczyk, N. Reul, G. Reverdin, and J. Vialard (2018), New SMOS Sea Surface Salinity with reduced systematic errors and improved variability, *Remote Sens. Environ.*, 214, 115-134.

Delcroix, T., M.J. McPhaden, A. Dessier, and Y. Gouriou, 2005: Time and space scales for sea surface salinity in the tropical ocean, *Deep-Sea Res. I*, **52**, 787-813.

Gandin, L.S., 1965: *Objective Analysis of Meteorological Fields*, 242 pp, Israel Program for Scientific Translation, Jerusalem.

Good, S.A., M.J. Martin, and N.A. Rayner (2013). EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, *J. Geophys. Res. Oceans*, 118, 6704-6716, doi:10.1002/2013JC009067.

Henocq, C., J. Boutin, F. Petitcolin, G. Reverdin, S. Arnault, and P. Lattes, 2010: Vertical variability of Near-Surface Salinity in the Tropics: Consequences for L-Band Radiometer Calibration and Validation, *J. Atmos. Oceanic Technol.*, **27**, 192-209.

Hosoda, S., T. Ohira, and T. Nakamura, 2008: A monthly mean dataset of global oceanic temperature and salinity derived from Argo float observations, *JAMSTEC Rep. Res. Dev.*, Vol. 8, 47-59.

Kao, H.-Y., G.S.E. Lagerloef, T. Lee, O. Melnichenko., T. Meissner, and P. Hacker, 2018: Assessment of Aquarius Sea Surface Salinity, *Remote Sensing*, 10(9), 1341, doi:10.3390/rs/10091341.

Kolodziejczyk Nikolas, Prigent-Mazella Annaig, Gaillard Fabienne (2017). ISAS-15 temperature and salinity gridded fields. SEANOE. http://doi.org/10.17882/52367.

Lagerloef, G., F.R. Colomb, D. LeVine, F. Wentz, S. Yueh, C. Ruf, J. Lilly, J. Gunn, Y. Chao, A. deCharon, G. Feldman, and C. Swift, 2008: The Aquarius/SAC-D Mission: Designed to meet the salinity remote-sensing challenge, *Oceanography*, **20**, 68-81.

Le Traon, P.Y., F. Nadal, and N. Ducet, 1998: An improved mapping method of multisatellite altimeter data, *J. Atmos. Oceanic Technol.*, **15**, 522-534.

Le Vine, D.M., G.S.E. Lagerloef, F.R. Colomb, S.H. Yueh, and F.A. Pellerano, 2007: Aquarius: An instrument to monitor sea surface salinity from Space, *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 2040-2050.

Meissner, T., F.J. Wentz, and D.M. Le Vine, 2017: Aquarius Salinity Retrieval Algorithm End of Mission ATBD, report number 120117, Remote Sensing Systems, Santa Rosa, CA, 113 pp. Available online at http://images.remss.com/papers/tech_reports/2017/Meissner_AQ_ATBD_EOM_final.pdf

Meissner, T., F.J. Wentz, and D.M. Le Vine, 2018: The Salinity Retrieval Algorithms for the NASA Aquarius Version 5 and SMAP Version 3 Releases, Remote Sensing 10, 1121, doi:10.3390/rs10071121.

Meissner, T., F. J. Wentz, A. Manaster, and R. Lindsley, 2019: Remote Sensing Systems SMAP Ocean Surface Salinities [Level 2C, Level 3 Running 8-day, Level 3 Monthly], Version 4.0 validated release. Remote Sensing Systems, Santa Rosa, CA, USA. Available online at www.remss.com/missions/smap.

Melnichenko, O., P. Hacker, N. Maximenko, G. Lagerloef, and J. Potemra, 2014: Spatial Optimal Interpolation of Aquarius Sea Surface Salinity: Algorithms and Implementation in the North Atlantic, *J. Atmos. Oceanic Technol.*, 31, 1583-1600.

Melnichenko, O., P. Hacker, N. Maximenko, G. Lagerloef, and J. Potemra, 2016: Optimum interpolation analysis of Aquarius sea surface salinity, *J. Geophys. Res. Oceans*, 121, 602-616, doi:10.1002/2015JC011343.

Reul, N.; Grodsky, S.A.; Arias, M.; Boutin, J.; Catany, R.; Chapron, B., D'Amico, F.; Dinnat, E.; Donlon, C.; Fore, A.; Fournier, S.; Guimbard, S.; Hasson, A.; Kolodziejczyk, N.; Lagerloef, G.; Lee, T.; Le Vine, D.M.; Lindstrom, E.; Maes, C.; Mecklenburg, S.; Meissner, T.; Olmedo, E.; Sabia, R.; Tenerelli, J.; Thouvenin-Masson, C.; Turiel, A.; Vergely, J.L.; Vinogradova, N.; Wentz, F.; Yueh, S., 2020: Sea surface salinity estimates from spaceborne L-band radiometers: An overview of the first decade of observation (2010–2019), *Remote Sens. Environ.*, 242, 111769, doi:10.1016/j.rse.2020.111769.

Reul, N., M. Arias, P. Spurgeon, et a., 2017: Read-me-first note for SMOS Level 2 Sea Surface Salinity data products, ESA, 10 May 2017, available online at https://earth.esa.int/documents/10174/1854503/SMOS-Level-2-Ocean-Salinity-v662-release-note.

Reverdin, G., E. Kestenare, C. Frankignoul, and T. Delcroix, 2007: Surface salinity in the Atlantic Ocean (30^{o}S-50^{o}N), *Prog. Oceanogr.*, **73**, 311-340.

Roemmich, D. and J. Gilson, 2009: The 2004-2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Prog. Oceanogr*., 82, 81-100.

Sakurai, T., K. Yukio, and T. Kuragano, 2005: Merged satellite and in-situ data global daily SST, *Proceedings: 2005 IEEE International Geoscience and Remote Sensing Symposium*, 2005, IGARSS'05, 2005, 2606-2608.

Vinogradova, N. T., and R.M. Ponte, 2012: Assessing temporal aliasing in satellite-based surface salinity measurements, *J. Atmos. Oceanic Technol.*, 29, 1391-1400, doi:10.1175/JTECH-D-11-00055.1.

Vinogradova, N. T., and R.M. Ponte, 2013: Small-scale variability in sea surface salinity and implications for satellite-derived measurements, *J. Atmos. Oceanic Technol.*, **30**, 2689-2694.

Weber, R.O., and P. Talkner, 1993: Some Remarks on Spatial Correlation Function Models, *Mon. Wea. Rev.*, **121**, 2611-2617.