Envisat and CryoSat-2 Satellite Altimetry Product Handbook

Gridded and along-track Sea Level Anomalies
for the Southern Ocean
2002.07 - 2018.10

Nomenclature: dot_all_30bmedian_geoid_3sig.nc
Issue: 1.1
Date: 10 Mar 2025

Chronology Issues			
Issue	Date	Created by	Reason for change
1.0	27/02/2024	Oana Dragomir	Creation of the document following the layout of SSALTO/DUACS Experimental Product Handbook (https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/regional/antarctic-sea-level-heights.html)
1.1	10/03/2025	Oana Dragomir	expanded the description of the processing and added references to the corresponding scripts

List of Acronyms	
DOT	Dynamic Ocean Topography
MDT	Mean Dynamic Topography
MSS	Mean Sea Surface
SLA	Sea Level Anomaly
SSH	Sea Surface Height

1. Introduction 5
2. Access and Citation 5
3. Product Description 6
 3.1. Versioning 6
 3.2. Geographical characteristics 6
 3.3. Temporal availability and grid characteristics 6
 3.4. Nomenclature 6
 3.5. Format 6
4. Processing into a Gridded monthly combined product 7
 4.1. Along-track processing 7
 4.2. Offsets 12
 4.3. Geoid 15
 4.4. Merging Satellites 16
5. Validation 20
 5.1. Aviso+ altimetry 20
 D1a_cmems_processing.py, D1b_validation_aviso.py 20
 5.2. Bottom Pressure recorders 20
 D2_validation_bpr_DPS.py, D2_validation_bpr_DPSdeep.py, D2_validation_bpr_myrtlec.py 20
6. Citations 21
7. Contacts 21
Bibliography 22

1. Introduction

This document describes a Southern Ocean data product providing sea level height to the south of 50S including over the ice covered regions. The latter is possible by using observations from the leads (i.e. openings in the sea ice). The gridded product combines data collected during two satellite missions monitored by the European Space Agency: Envisat and CryoSat-2. Along-track data are cleaned and processed to obtain a merged gridded product of monthly averages of Dynamic Ocean Topography (DOT) that span the period from July 2022 to October 2018.

The files made available are the following:

- **dot_all_egm08_30median_3sig.nc**
- **dot_all_goco05c_30median_3sig.nc**
- **dot_all_eigen6v4s_30median_3sig.nc**

Multiple gridded products are produced due to using three different geoid products as a reference (present in the nomenclature):
- GOCO05c (Pail et al, 2016)
- EGM2008 (Pavlis et al, 2012)
- EIGEN64v4 (Foerste et al, 2014)

The data are binned on a **0.5 degree latitude x 1 degree longitude grid** and have a 300km radius Gaussian filter applied.

Only the technical details are presented here to keep it brief. A more detailed description including satellite specifications can be found in Chapter 1 of the PhD thesis.

A detailed description of the processing of along-track data is provided in Section 4.
Access and citation guidelines are presented in Section 2 and the product details are described in Section 3. Section 4 and 5 contain the processing and validation methods, respectively. Some examples of where this product has been used and recommendations for improvement are shown in Section 6.

2. Access and Citation

This altimetry product was created as part of the PhD thesis “Dynamics of the Subpolar Southern Ocean response to Climate Change” by Oana Dragomir.
When using this Envisat/CryoSat-2 product please cite:

Dragomir, Oana Claudia (2024) Dataset supporting the University of Southampton doctoral thesis "Dynamics of the subpolar southern ocean response to climate change". University of Southampton doi:10.5258/SOTON/D3006 [Dataset]

Git (for scripts only): https://github.com/oana-dragomir/PhD_altimetry
For the raw altimetry data please contact (need to check with UCL if we can make these public):

3. Product Description

3.1.Versioning

First and only version, expected to be superseeded by the AVISO product in ca. 2025. I don't expect this product to be reviewed or maintained, although it might be possible that it will be extended in time by other PhD students. The more varied the products and methods, the better …

3.2. Geographical characteristics

elevation to the south of 50S
references to the geoid

3.3. Temporal availability and grid characteristics

The data span the period from July 2022 to October 2018 and are binned on a 0.5 degree latitude x 1 degree longitude grid.
and have a 300km radius Gaussian filter applied

3.4. Nomenclature

The nomenclature is as follows (and filenaming combinations):
parameter_satellite_BinThreshold&Statistic_geoid_GaussianFilter.nc

parameter	satellite	bin threshold and statistic	geoid	Gaussian filter (sigma)
dot	env cs2 all (env+cs2)	30median 30mean (not used)	egm08 goco05c eigen6v4s	3sig 2sig (not used)

N.B.: We denote SSHA (sea surface height anomaly) the anomaly relative to a Mean Sea Surface Height (MSS) provided with our data. After the offset corrections are applied (see the thesis), the SSH is referenced to the geoid to obtain the Dynamic Ocean Topography (DOT = SSH - geoid). A Mean Dynamic Topography (MDT) is obtained by computing a time average in every grid cell for a defined time period, ideally an integer number of years. The anomaly wrt to this surface is the Sea Level Anomaly: SLA = DOT - MDT.
We distinguish between these variable names solely to ensure consistency and clarity when explaining the processing steps.

3.5. Format

2.5.4.1 Dimensions
Along-track products
the defined dimensions are:
- time: number of measurements in current file

Gridded products:

2.5.4.2 Data Handling Variables
2.5.4.3 Attributes

4. Processing into a Gridded monthly combined product

4.1. Along-track processing

Along-track Level 2 (Bouffard, 2018) sea surface height (SSH) data referenced to the WGS84 ellipsoid are obtained from the Centre of Polar Observations and Modelling (CPOM) and the University College London.

The altimetry data span the period from November 2002 to October 2018 and were collected during two satellite missions monitored by the European Space Agency (ESA): Envisat, which was active between May 2002 and March 2012, and CryoSat-2 which was launched in April 2010 and it still operational at the moment of writing.

- ENVISAT data are available from July 2002 to March 2012.
- CryoSat-2 underwent 6 months of on-orbit testing during which the data density is lower. As a result, we start our analysis from November 2010. The data processed in this study end in October 2018.

Point-wise data are grouped by calendar months and stored in separate text files (.elev). The variables included in every file are shown in the Table 1 below.

The Antarctic Mean Sea Surface (MSS UCL04) provided is computed by combining a high

Field	Comments
Surface Type	0: Unknown, 1: Ocean, 2: Lead, 3: Floe
Validity	0: invalid, 1: valid
Source Packet ID	(not used)
Block Number	(not used)
Time	Days since 1/1/1950
Lat	
Lon	
Elevation/SSH	
MSS	
Peakiness	(not used)
Backscatter	(not used)
Percentage Ice Concentration	
Sea Ice Type	0: Unset, 1: Open Water, 2: FYI, 3: MYI, 4: Ambiguous
Confidence in Sea Ice Type	0: Not Processed, 1: Computation Failed, 2: Use with Care, 3: Acceptable, 4: Good, 5: Excellent
Sigma of Fit to Specular Echo	(not used)
Error in Fit to Specular Echo	(not used)

Table 1. List of variable names present in the original along-track data files from CPOM/UCL.

frequency component based on an average from a complete 369-day cycle of CryoSat-2 data (from 28th January 2011 to 31st January 2012) with a low frequency component from one month of data (March 2011) (Dotto, 2019).

A range of geophysical and atmospheric corrections were previously applied to the satellite data. These corrections include dry and wet tropospheric, ionospheric, inverse barometric, dynamic atmospheric, ocean tides, ocean loading tide, long-period equilibrium ocean tide, solid Earth tide, and geocentric polar tide.

PhD/PhD_scripts/aux_func/

aux_2a_combined_coastlines_plot.py

- create a wrapper to plot the coastline and the ice shelves to the south of 50S
- the **coastline product** used in all subsequent plots is made up of:
    > Basemap coastline north of 60S (GSHHG: https://www.soest.hawaii.edu/pwessel/gshhg/)
    > Antarctic Digital Database (ice shelves+coast:https://www.add.scar.org/); cite Paul Holland and Clement Vic for the data in PhD_data/land_masks/holland_vic

aux_2b_coastlines_gridded_land_mask.py

- the goal was to create a gridded land mask to apply it to the gridded product to mask any grid cells that I might have missed in the processing that are inside the land contour
- I seem to have misplaces the script that created this so I only have a this file which creates coastline_nested_lists.pkl; the file that it should create is land_mask_gridded_50s.nc

Topography/coastline
aux_regrid_GEBCO_topog.py -> coarse_gebco_p5x1_latlon.nc
bathy_mask_1km.nc

PhD/PhD_scripts/A_altimetry/A_along_track/

A0_text2nc_files_ENV.py, A0_text2nc_files_CS2.py

- iterate through every .elev file
- select valid (valid=1) ocean (surf=1) and lead (surf=2) data
- remove outliers where:
    > percentage Sea Ice Concentration (SIC) < 0 (concentration is a +ve quantity)
    > Confidence in Sea Ice Type (CSIT < 4) [keep only data where csit 4=good & 5=excellent]
    > Sea Ice Type (SIT) < 1 (keep all types, outliers have values<1)
- discard |SSHA| > 3 m, where SSHA = SSH - MSS
- assign track numbers and directions (ascending[1]/descending[-1])

For CS2, also assign a label to every point indicating inside which type of surface/retracker** they are found (1=LRM, 2=SAR, 3=SARIn)
CS2 geographical masks used: **version 3.8 (https://earth.esa.int/eogateway/instruments/siral/description#geographical-mask-mode)

- save monthly data as .nc files (`PhD_data/altimetry_cpom/1_raw_nc/`)

Tip - see if you can use Dask to process quicker

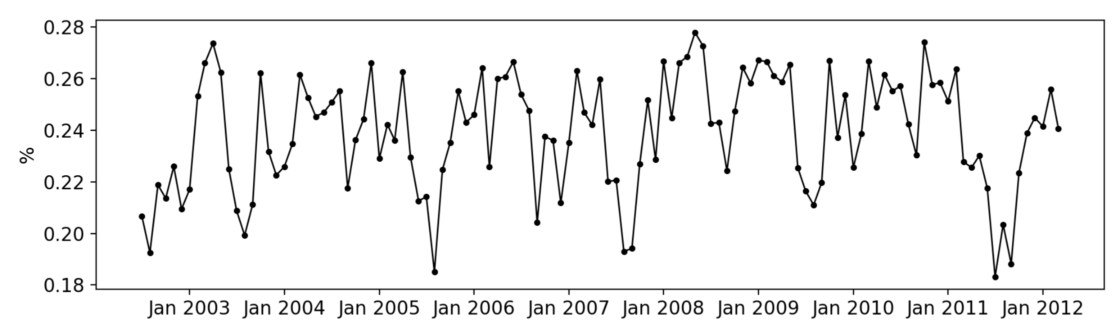
A1_distance_to_land.py

- this adds another field in every monthly .nc file ("distance_m") with the distance in metres along the ellipsoid to the nearest coastline point

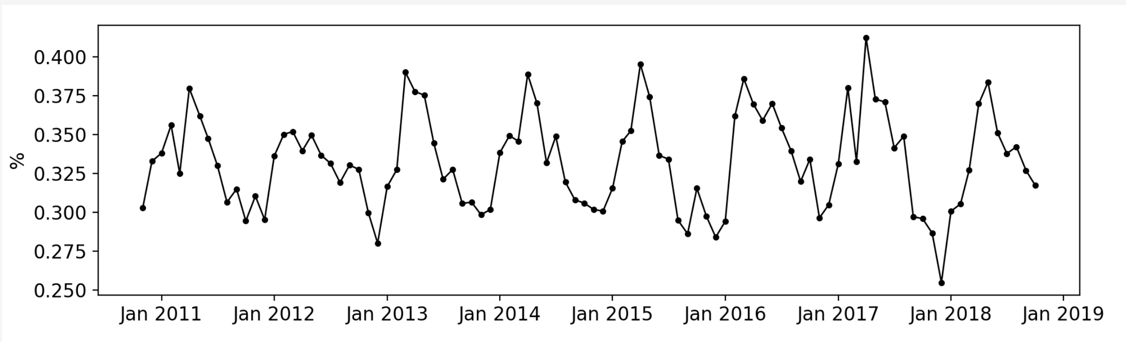
  **the package I'm using to compute the distance is not compatible with >python3.7 so an alternative should be found; or skip this altogether because it takes ages to run and use a gridded mask instead because there is an insignificant percentage of points inside contours/on land; another workaround is to install Cython which then allowed me to install geo-py with python 3.9

A1_fig_number_of_spotSSH.py

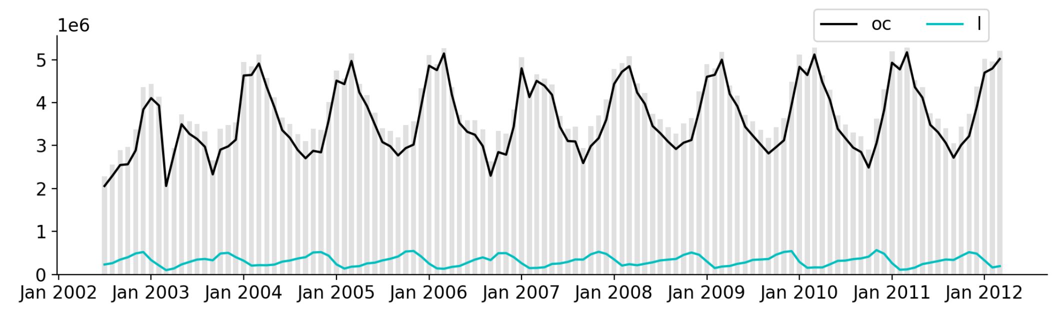
> the aim is to discard point-wise data that are less than **10 km** away from the nearest coastline; This is less than 0.5% of the number of data in a month so it's not really relevant (a bit more relevant for CS2 which makes sense given better resolution near the coast)



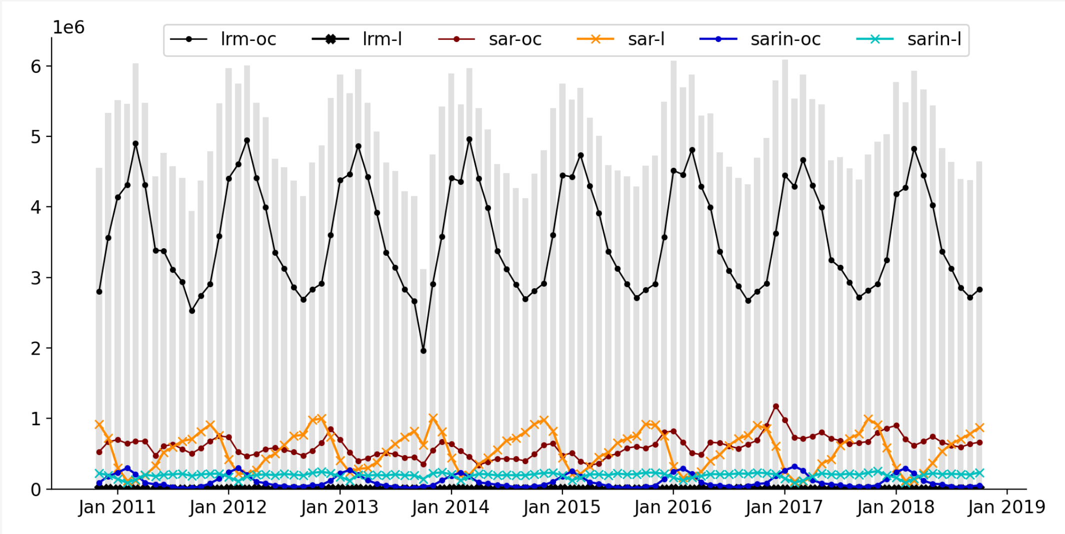
Envisat; Percentage of points that are less than 10 km away from the nearest coastline relative to the total number of pointwise data in every month



Same as above but for CS2



Envisat
grey: total number of pointwise data in every month
Ocean/lead data are in black/teal.



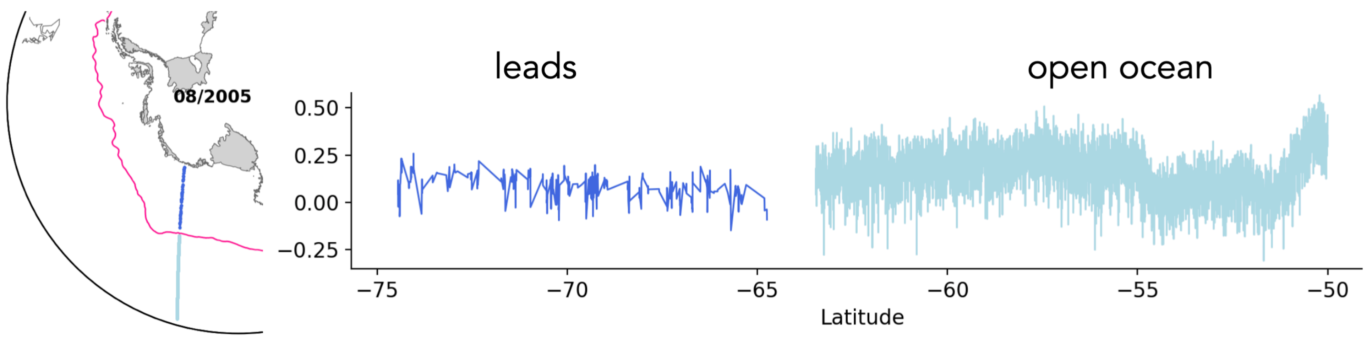
CryoSat-2

grey: total number of pointwise data in every month
black dotted: LRM-ocean
see legend

The total number of data points is split into the type of surface and retrackers

Tips:
1. Could probably skip this step and only apply the gridded land mask at the end since there is a very small percentage of points inside the land contour.
2. Could try applying a filter to the along-tracks and see how that changes results [from what I remember it doesn't affect the gridded product significantly]

A1_fig_single_along_track.py



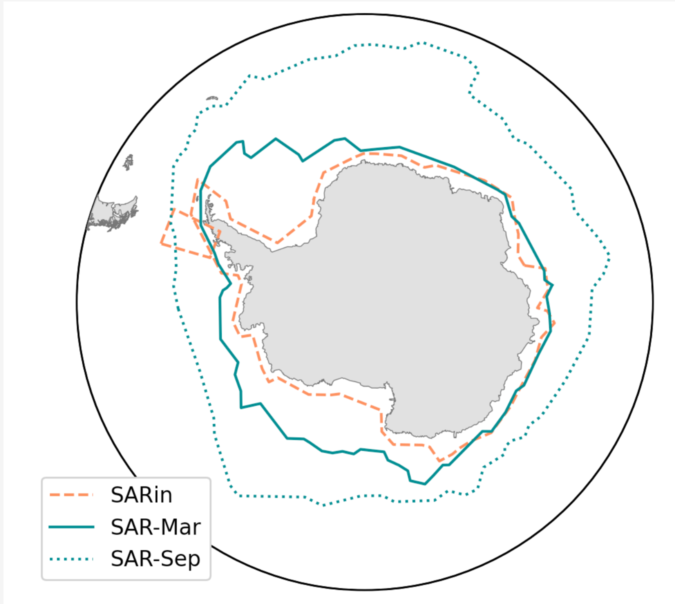
Difference between sampling in the open ocean and in the leads - Envisat

A1_fig_tracks_all.py
- figures with the along-track data for every month; I suggest using A3_bin_numpts.py instead because the binned data is quicker to plot
- use this to look at the data density for every month

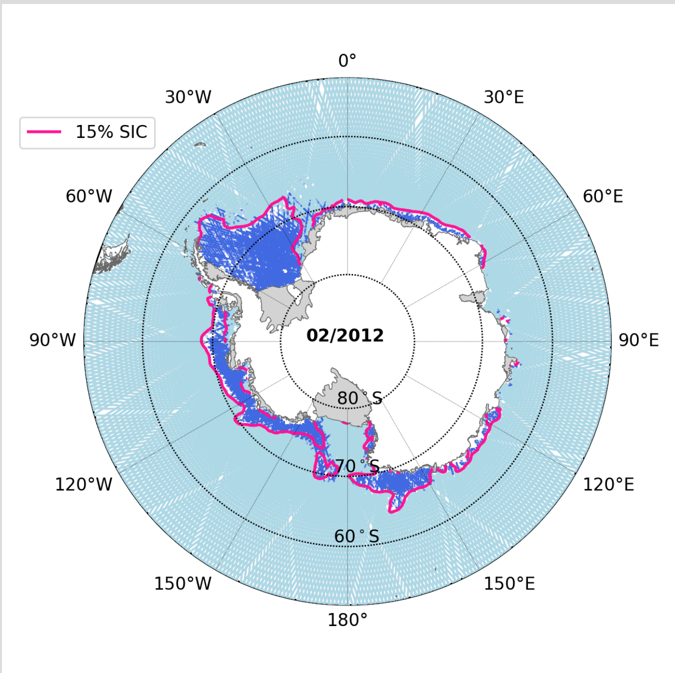
A2_fig_tracks_cs2_modes_mask.py	

plot of the CS2 retracker outlines (version 3.8) as downloaded from https://earth.esa.int/eogateway/instruments/siral/description#geographical-mask-mode

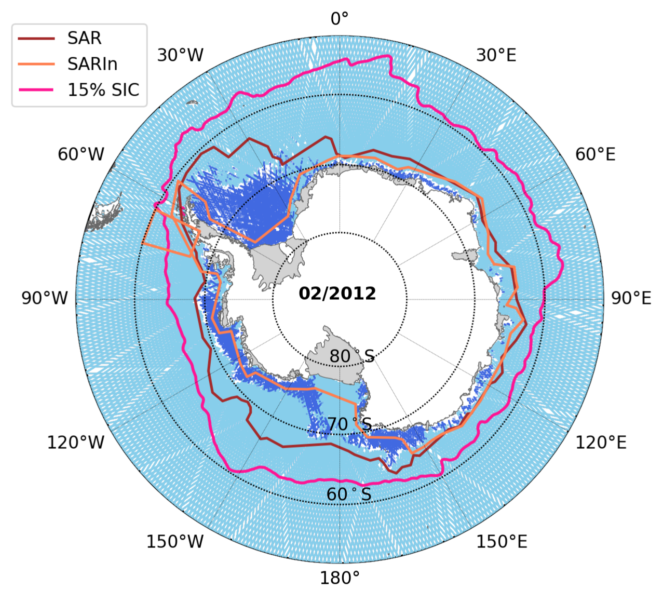
> SARin mask is static
> SAR varies monthly with Sea Ice extent but it does no change from year to year (!this might change in different versions)



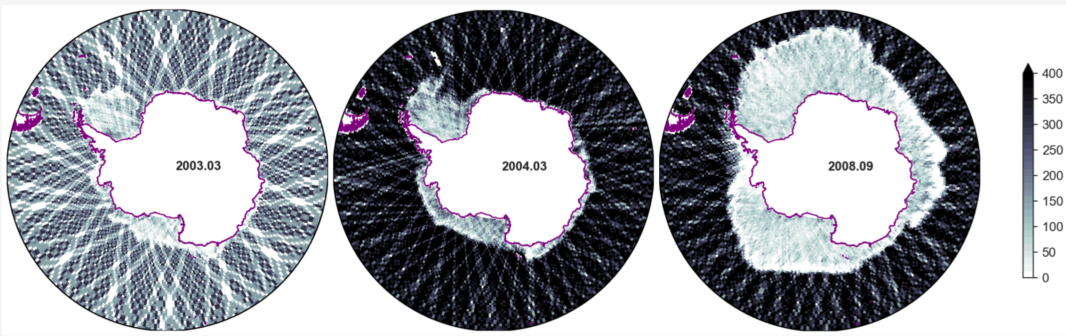
A2_fig_tracks_env_ol.py



A2_fig_tracks_cs2_modes.py

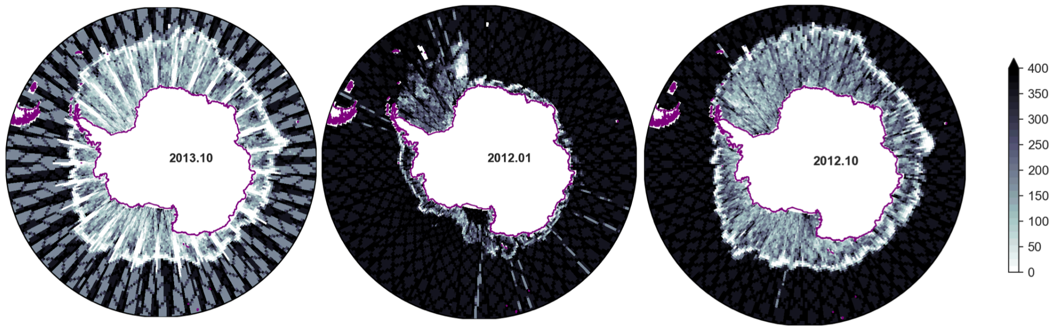


A3_bin_numpts.py



ENV
examples of data density after binning the along-track data

Envisat was rebooted in March 2003, May and June 2006 (Kaczmarska, 2012) which explains the lower data coverage.



Same as above but for CS2; unsure what happened in 2013.10 but there is a lower data coverage

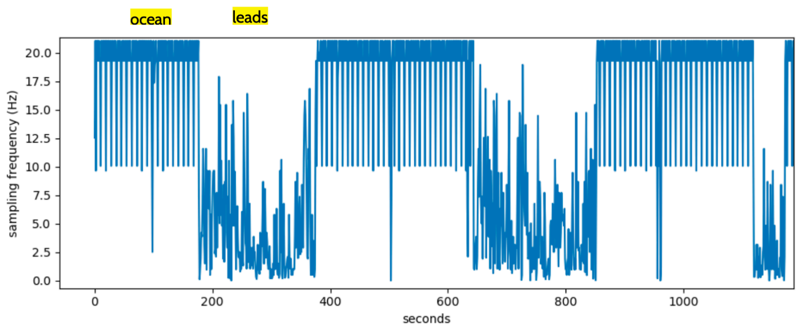
- We looked at plots of the density of along-tracks in every month and decided to leave out certain months (eg. **5-6.2002 of ENV** and **7-10.2010 of CS2**) due to low data coverage.

Sampling frequency

Sampling frequency

- number of measurements/s
ENV : ~ 20 Hz in open ocean
< 10 Hz over leads

```
> dt = time[1:] - time[:-1]
> sec = dt[dt!=0]*24*3600
> a = np.arange(1, len(time))
> indx = a[dt!=0]
> count = indx[1:]-indx[:-1]
> count = np.hstack((indx[0], count))
> freq = count/sec
```



4.2. Offsets

The aim is to inspect the spatial and temporal distribution of the offsets between O-L and between the three retrackers/surfaces for CS2. My approach to the latter is slightly different to Tiago Dotto and Jack Hooley ,because I first isolate the O-L surfaces and correct for that offset, and then I split the along track data in the 3 surfaces (LRM, SAR, SAR-In) and compute the offsets between SAR-SAR-In and LRM-SAR and level to LRM.

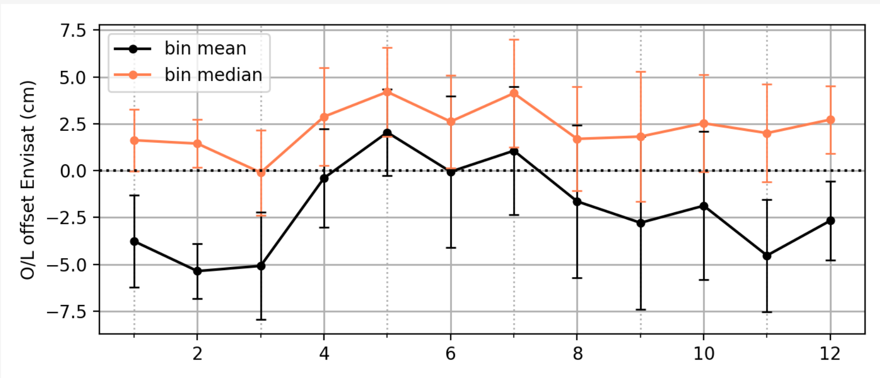
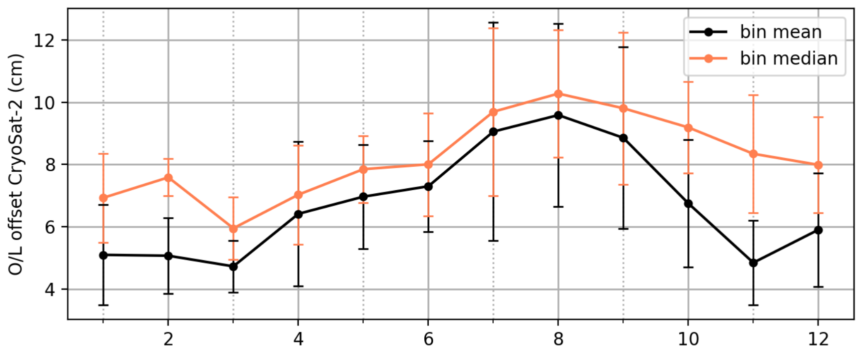
PhD_scripts/A_altimetry/B_offsets/
B1_bin_ssha_OL_files.py
- bin ocean/lead data separately for both ENV and CS2; save in one file the gridded ocean and lead data together with the number of points in each bin

grid/bin size:
-- compute autocorrelation length-scale to determine grid size (based on Tiago's analysis and a quick check this is roughly < 50km)
- or use what other papers have found that autocorrelation length scale is ~50 km hence use a grid cell of **1 deg lon x 0.5 deg lat**

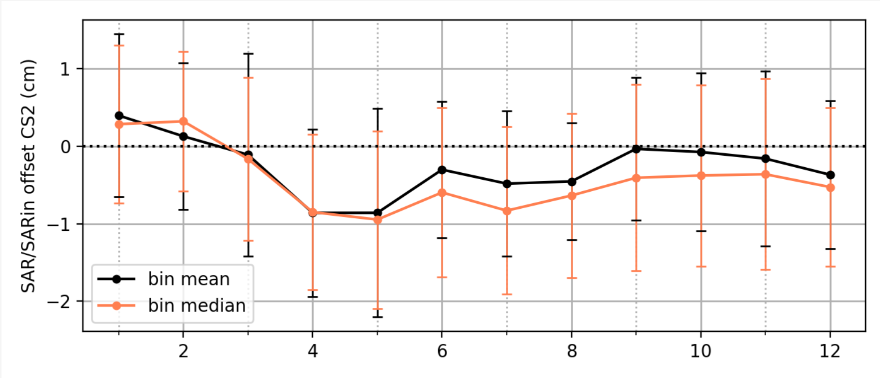
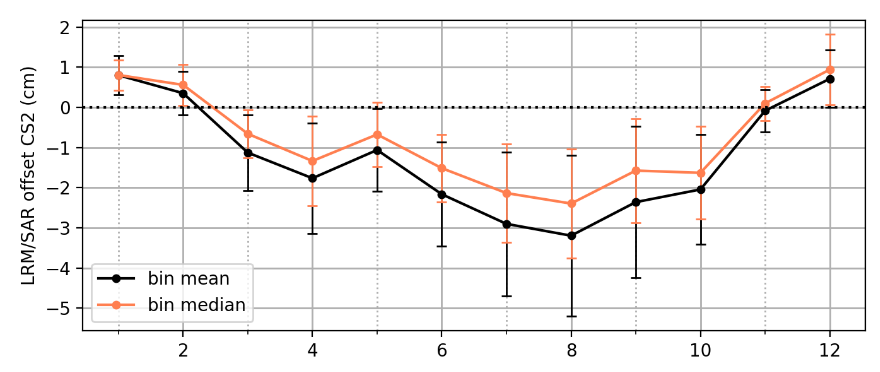
B2_OL_analysis.py
- set a **threshold** - min number of points in every bin is 30
- subtract gridded lead from ocean and check time/spatial patterns of the offset

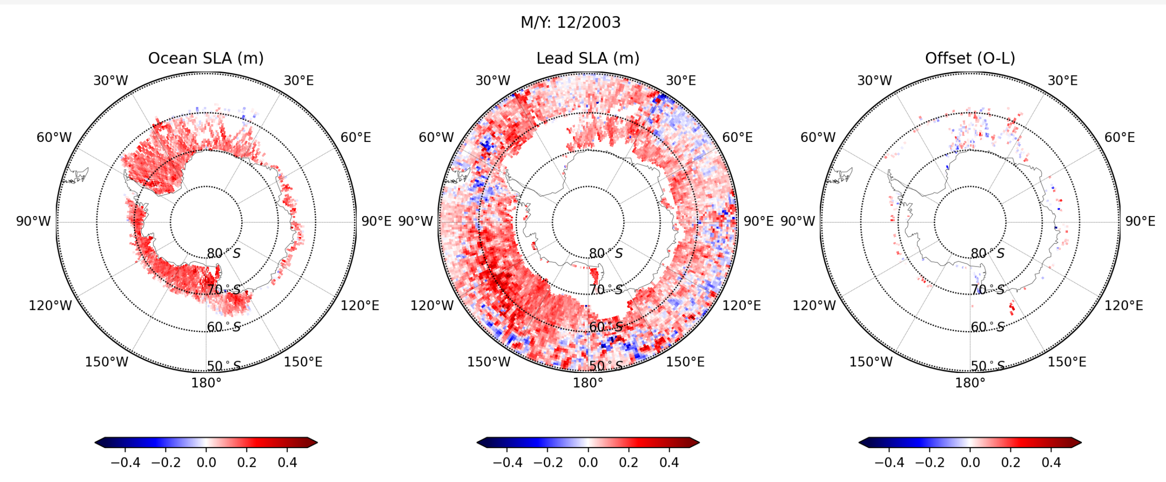
- compute **monthly area-weighted averages** of the bias and do a spread of residuals or weighted stDev

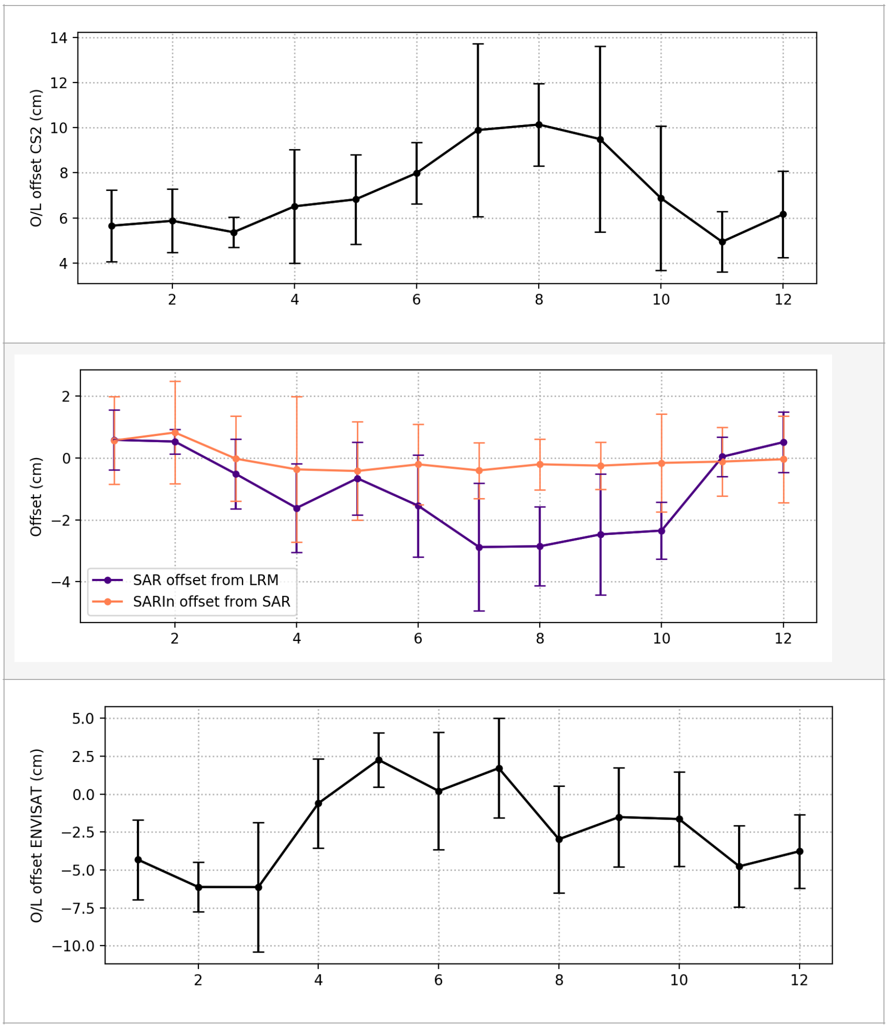
- **compute monthly climatology that we apply to the along-track leads**
    - we tested different stats, what we picked is plotted in the next section

B2_OL_plots.py	O-L monthly climatology offsets (median)
ENV & CS2 - compute monthly area-weighted mean offset (weights are normalized) - compute monthly climatology StDev - weighted Stdev (sqrt(weighted_mean((offset-sw_mean)^2)	 

B3_bin_ssha_SAR_cs2.py
- applied to CS2 data only
- correct for the O/L offset using the monthly climatology (median) described above
- split along-track data into SAR, SARIn, bin the three surfaces, and store everything in a file

B3_SAR_plots_file.py	
CS2 - correct along-track leads and then compute SAR-SARIn offset - compute SAR-SARIn mclim and the spread of monthly area-weighted means	
B4_bin_ssha_LRM_cs2	- correct Leads, correct SARIn - split into LRM and SAR(In) and save file
B4_LRM_plots_file.py	
CS2 - correct along-track leads and then correct SARIn to level with SAR - treat SAR and SARIn together [SAR(In)] - compute LRM-SAR(In) mclim and spread of monthly area-weighted means	

B5_fig_stereoplot_offset.py

gridded Ocean and lead data and the offset in the overlapping bins



summary

> compute the offset as a **monthly climatology** (with standard deviation of the weighted means averaged over every month) to correct the along-track (SSH) data
> each satellite is referenced to LRM
> the bin statistic can be mean or median; we consider all combinations but use the **median** as the bin statistic in the end
> geoid products tested: EGM2008 and GOCO05c (and also EIGEN6s4v2); a more thorough comparison can be made but we choose GOCO05c because it is an improvement over and it is used in other papers processing similar data (i.e. Armitage et al. (2016, 2018))

From here on, we reference the SSH data to the geoid which gives us Dynamic Ocean Topography (DOT). This parameter is more suited for examining the dynamics.
But first, we describe how the geoid is obtained and processed.

4.3. Geoid

Path: **PhD/PhD_data/geoid**

Download the geoid product from: http://icgem.gfz-potsdam.de/home
Geoid products used: **EGM2008, GOCO05c** [accessed: 12 Nov 2018], **EIGEN6v** [accessed: 19 Nov 2021]

The download is done in two files (I think because of memory restrictions):
- lat: -80 to -48
- lon: 0.025 to 179.975 and -180 to 0 (in the file the convention is 180 to 360)
- grid size: 0.025
- filenames:
    > EGM2008_0p025_179p975_lon_m80_m48_lat.gdf
    > EGM2008_0_m180_lon_m80_m48_lat.gdf

Processing:
1. Cut and paste the header in a different file (.txt)
2. Change extension of original data files to .txt (EGM2008_0_180_lon.txt, EGM2008_180_360_lon.txt)

3. use GMT to convert .txt files to .nc
    [N.B.: installed gmt with conda]

```
>> gmt xyz2grd EGM2008_0_180_lon.txt -Gegm1.nc-R0.025/179.975/-80/-48 -I0.025[d]/0.025[d]
>> gmt xyz2grd EGM2008_180_360_lon.txt -Gegm2.nc -R180/360/-80/-48 -I0.025[d]/0.025[d]
```

4. combine the two .nc files into one (egm2008.nc) using the script 1_concatenate_ncfiles_neg.py; the _neg is added because the altimetry uses -180/180 convention compared to 0/360 that the geoid used so this is adjusted to match the altimetr altimetry data lon runs from -180 to 180; geoid data lon runs from 0 to 360 - so use the script to also change the reference to match the altimetry data; this helps for the next step

5. Interpolate the gridded geoid data to the along-track locations stored in the files in data/altimetry_cpom/1_raw_nc_lonlat/ using C0_crop_lonlat.py
- use the scripts: at_geoid_interp_CS2.gmt and at_geoid_interp_ENV.gmt (adjust for geoid name)
    In the terminl run:

```
>> cd ../PhD_data/geoid/
>> chmod +x ./at_geoid_interp_CS2.gmt
>> ./at_geoid_interp_CS2.gmt
```

and repeat for the other satellite.

4.4. Merging Satellites

C0_crop_lonlat.py
- - use along-track data to extract lon/lat in separate text files to be used for interpolating the geoid at those locations

C1_correct_bin_dot_env.py, C2_correct_bin_dot_cs2.py
- correct the along-track SSH for each satellite separately
- discard values that are less than 10 km away from the nearest coastline

- subtract geoid (egm and goco05): DOT = SSH - geoid
- only keep values where |DOT| < 3 m

- bin DOT for each satellite discarding bins that have less than **30 points;** bin statistic can be either mean or median; we compute both but here we choose **median**

- interpolate data to fill missing values (nearest neighbour) and apply a 2D Gaussian filter (sigma=150 km)

C3_intersat_offset.py
> > > compute intersatellite offset

> overlap period: November 2010 - March 2012
> using the gridded DOT from Envisat and CS2:
- compute mean dynamic topography (MDT) from the Envisat data over the overlap period:
MDT_env = mean of DOT_env during 11.2010-03.2012

- compute anomalies in ENV and CS2 relative to MDT_env:
SLA_env = DOT_env - MDT_env
SLA_cs2 = DOT_cs2 - MDT_env

- for every month, compute (maps of) differences between SLA_env and SLA_cs2; this produces **maps of monthly SLA differences**
- compute the **median** of these SLA differences in every grid cell (the mean gives a more spatially uniform residual) [see figures below]
- compute the area-weighted average of the median of the SLA differences [except where the map values are negative] - this is the **intersatellite offset value [3.86 cm]** that is then applied to the binned CS2 data

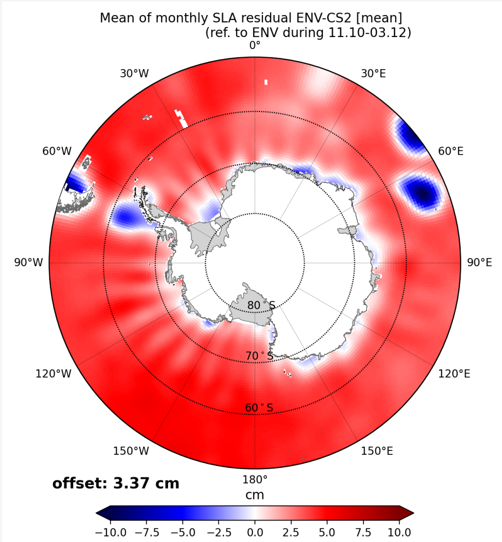
** the blue blobs in the figure should ideally not be there but I couldn't find a way to remove them and I just discarded them from the area-weighted average

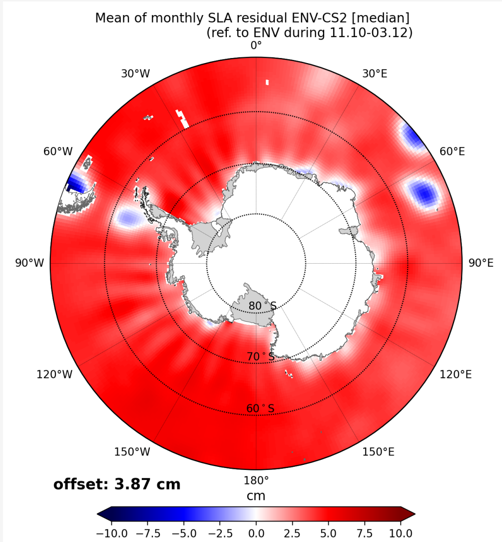
C3b_RMS.py
- compute a map of the root mean square between ENV and CS2 for the overlap period after applying the intersatellite offset to CS2 (see figures below in the green colourmap)

C4_combined_dot
- finally, the monthly gridded ENV DOT and the corrected CS2 DOT are averaged in the overlap period to merge the two satellites

	mean of residual @ bin	median of residual @ bin

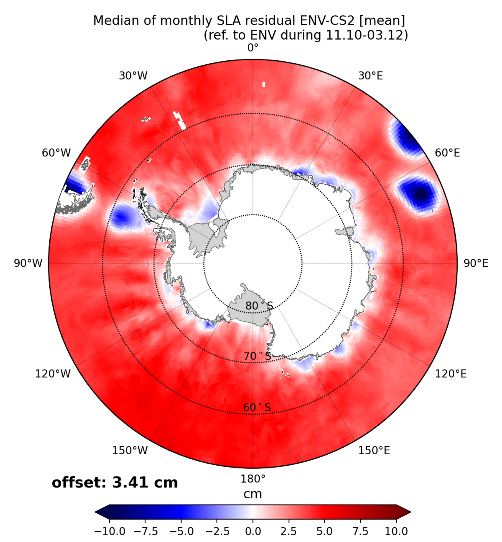
_mean_egm08

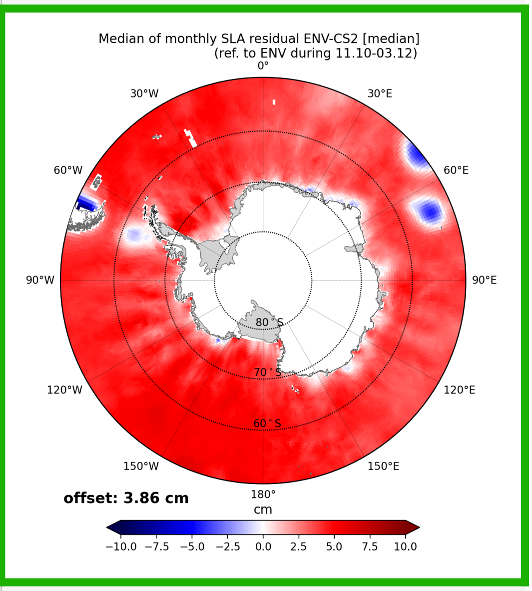




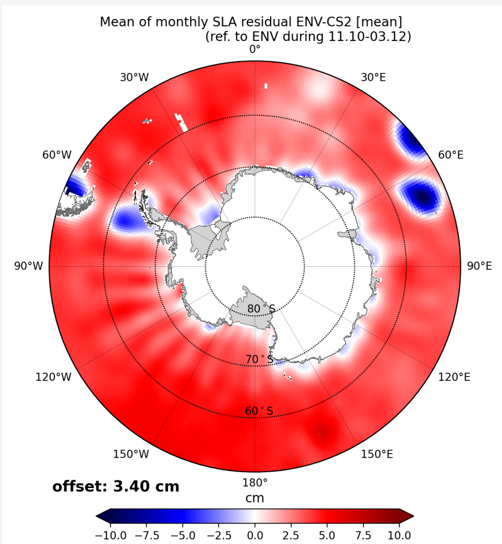
_median_egm08

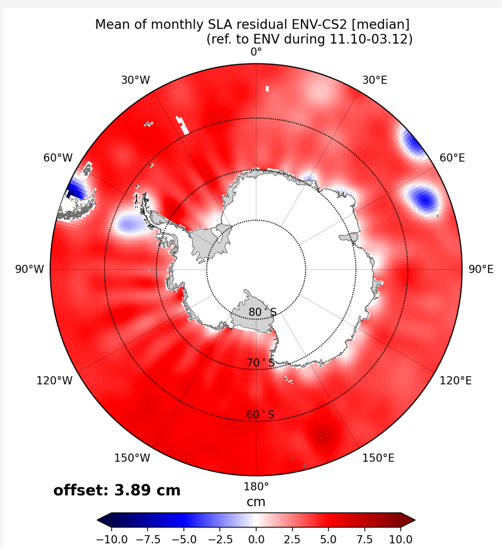
green contour = method used for the egm08 merged product





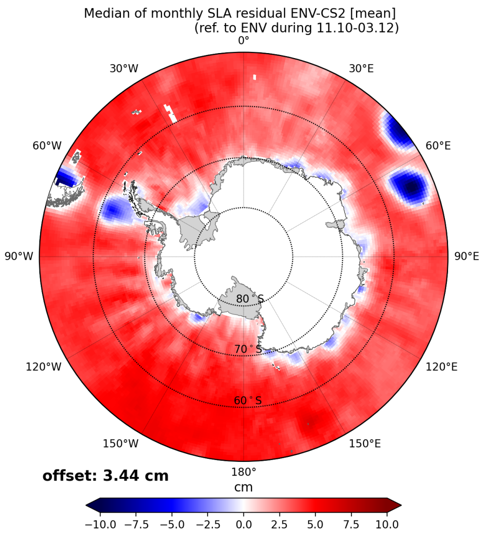
_mean_goco05c

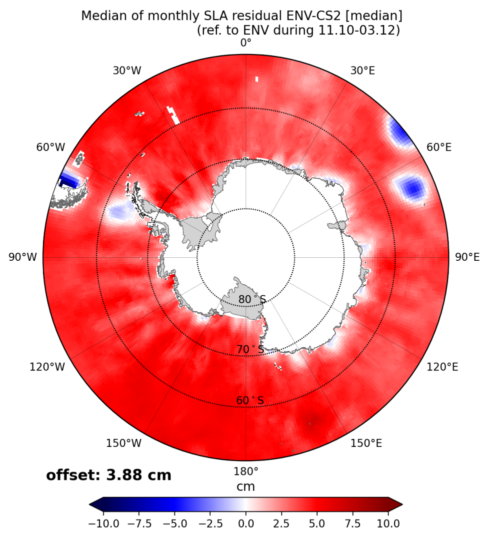




_median_goco05c

black outline = = method used for the goco05c merged product



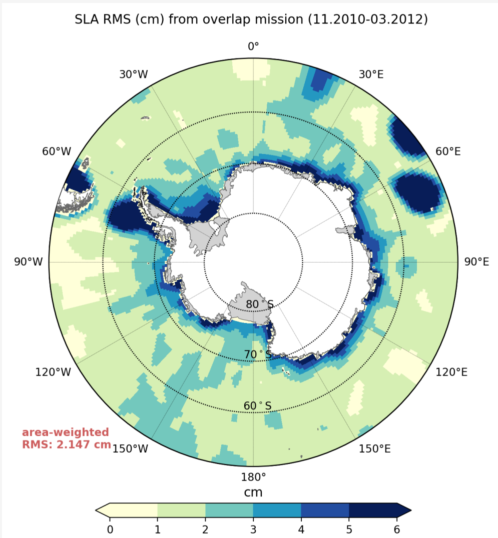


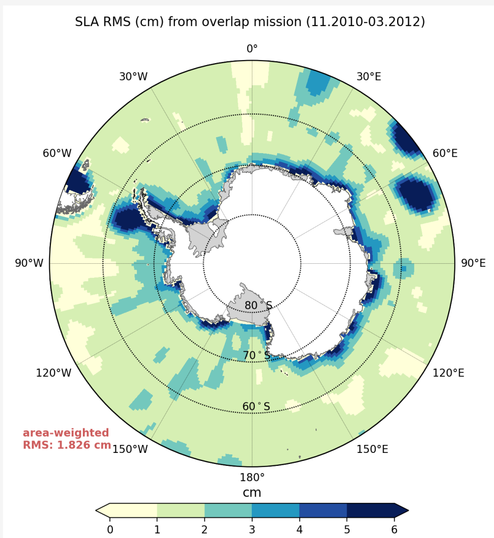
RMS of residual SLA after correcting CS2 DOT (binned, 30bmedian) with the area-weighted average of the map with the median SLA difference (3.86 cm)

geoid: egm08

N.B.: correcting with the mean intersat values and/or using the mean as the bin stat gives a higher RMS of residuals for both geoids; egm08 is just slightly better but probs not significantly

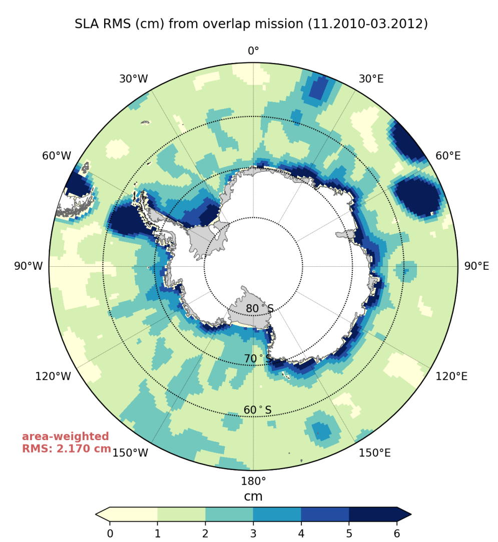
-> use median as the bin stats and the median-based offset

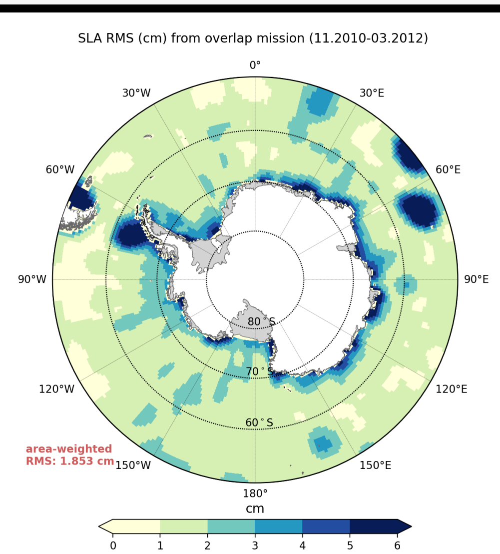




RMS of residual SLA after correcting CS2 DOT (binned, 30bmedian) with the area-weighted average of the map with the median SLA difference (3.88 cm)

geoid: goco05c



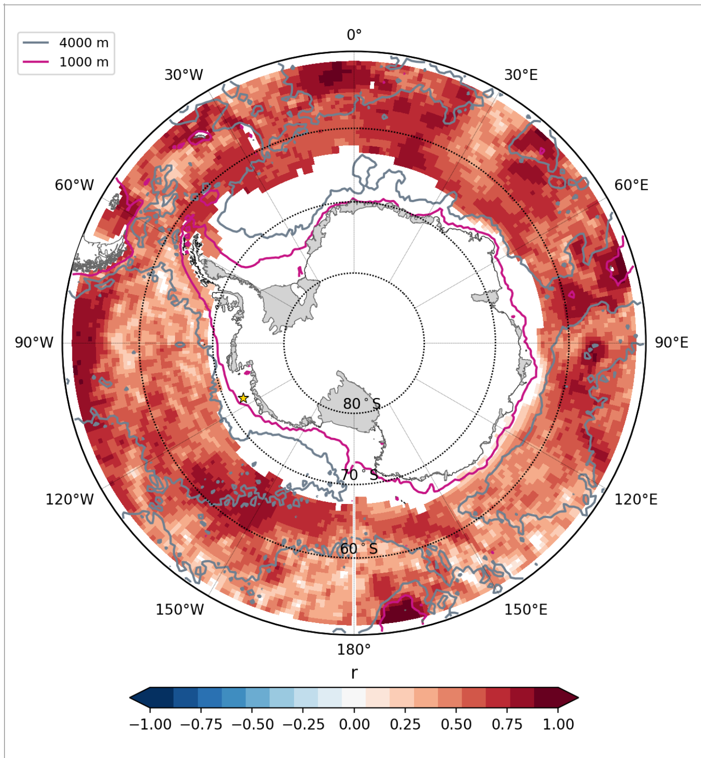


5. Validation

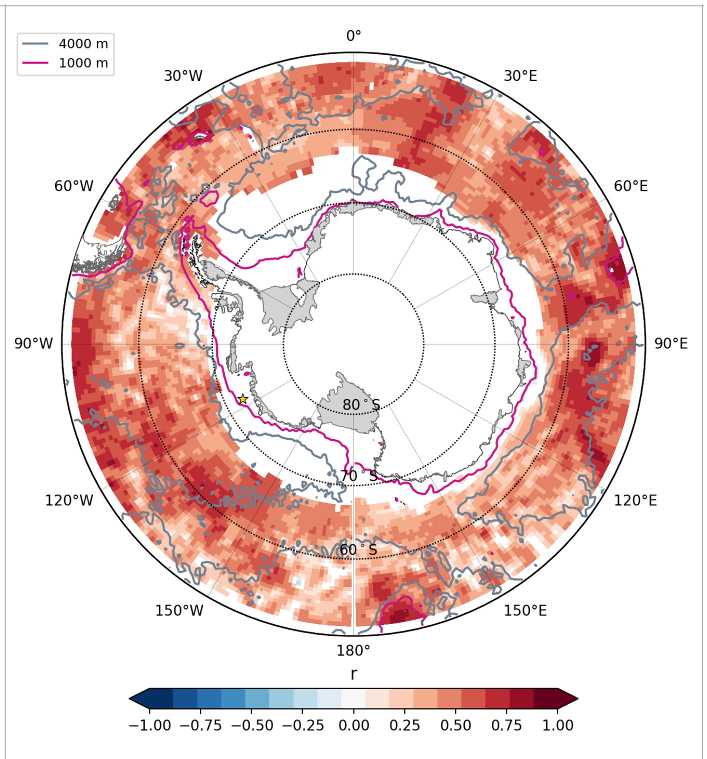
5.1. Aviso+ altimetry

D1a_cmems_processing.py, D1b_validation_aviso.py

> CLS/CMEMS: monthly gridded ADT data (re-gridded onto our product's grid)
- discard bins where values are missing more than 70% of the time
- compute correlation maps where there are at least 12 points in the time series
- period: 10.2002-10.2018
- plots below are where p-val <0.1
- detrend both datasets and compute correlation maps of multi-year anomalies
- maps are noisy because the CMEMS resolution is higher (I think); smoothing the CMEMS product might help



geoid: egm2008



geoid GOCO05c

5.2. Bottom Pressure recorders

D2_validation_bpr_DPS.py, D2_validation_bpr_DPSdeep.py, D2_validation_bpr_myrtlec.py

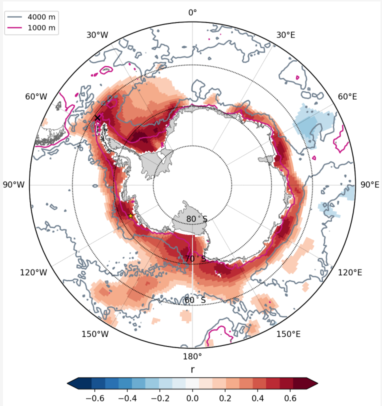
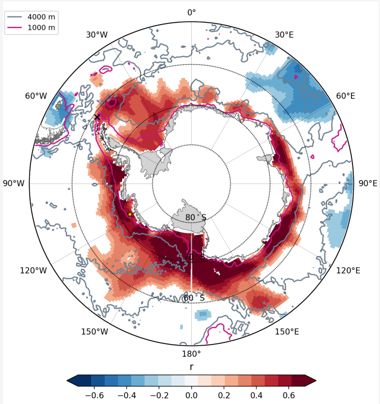
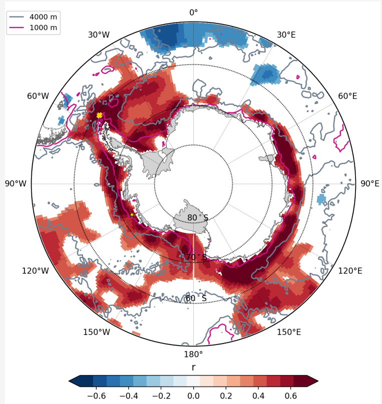
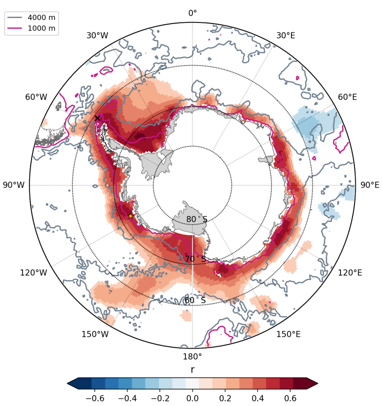
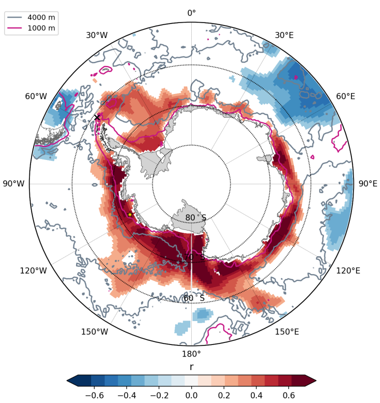
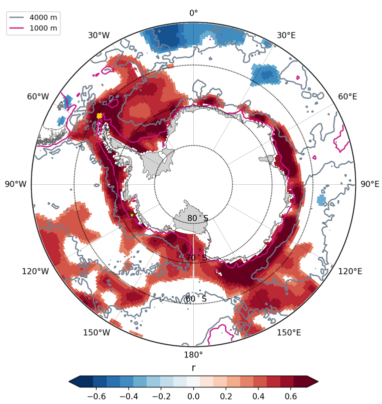
- BPR data: hourly bottom pressure from 3 recorders in Drake Passage: https://www.psmsl.org/data/bottom_pressure/table.php

> use field 'residual after removing tidal signal and drift'

> remove flagged data

> compute monthly averages of residual pressure after removing tidal prediction and drift (in **mbar**)

TIP: ideally should download the daily data because manipulation is easier but because we are using monthly means it does not make a difference

	DP south	DP south deep	Myrtle C
egm2008			
GOCO05c			
DOT and bottom pressure anomalies are detrended - not significant difference between geoid products, although EGM2008 seems to have higher correlations on the shelf (but by a small margin) **> the circumpolar signal on the shelf suggests that either product captures the dominant mode of variability there**			

6. Citations

- Thesis:
- undergraduate
- Aditya, Jenny papers, amundsen paper

6. Recommendations
- error calculation
-

7. Contacts

For more information, please contact: oana.dragomir.17@gmail.com

The user service is also interested in user feedback; questions, comments, proposals, requests are much welcome.

Bibliography

Pavlis, N. K., S. A. Holmes, S. C. Kenyon, and J. K. Factor (2012), The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), J. Geophys. Res., 117, B04406, doi:10.1029/2011JB008916.

Foerste, Christoph; Bruinsma, Sean.L.; Abrykosov, Oleh; Lemoine, Jean-Michel; Marty, Jean Charles; Flechtner, Frank; Balmino, Georges; Barthelmes, Franz; Biancale, Richard (2014): EIGEN-6C4 The latest combined global gravity field model including GOCE data up to degree and order 2190 of GFZ Potsdam and GRGS Toulouse. GFZ Data Services. https://doi.org/10.5880/icgem.2015.1

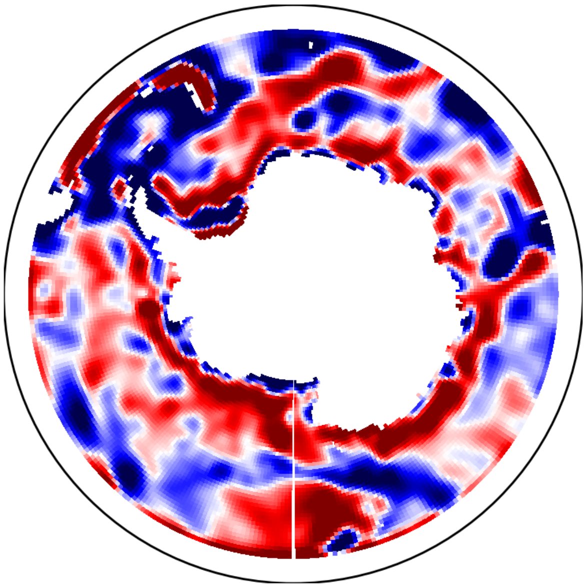
Pail, Roland; Gruber, Thomas; Fecher, Thomas; GOCO Project Team (2016): The Combined Gravity Model GOCO05c. GFZ Data Services. https://doi.org/10.5880/icgem.2016.003

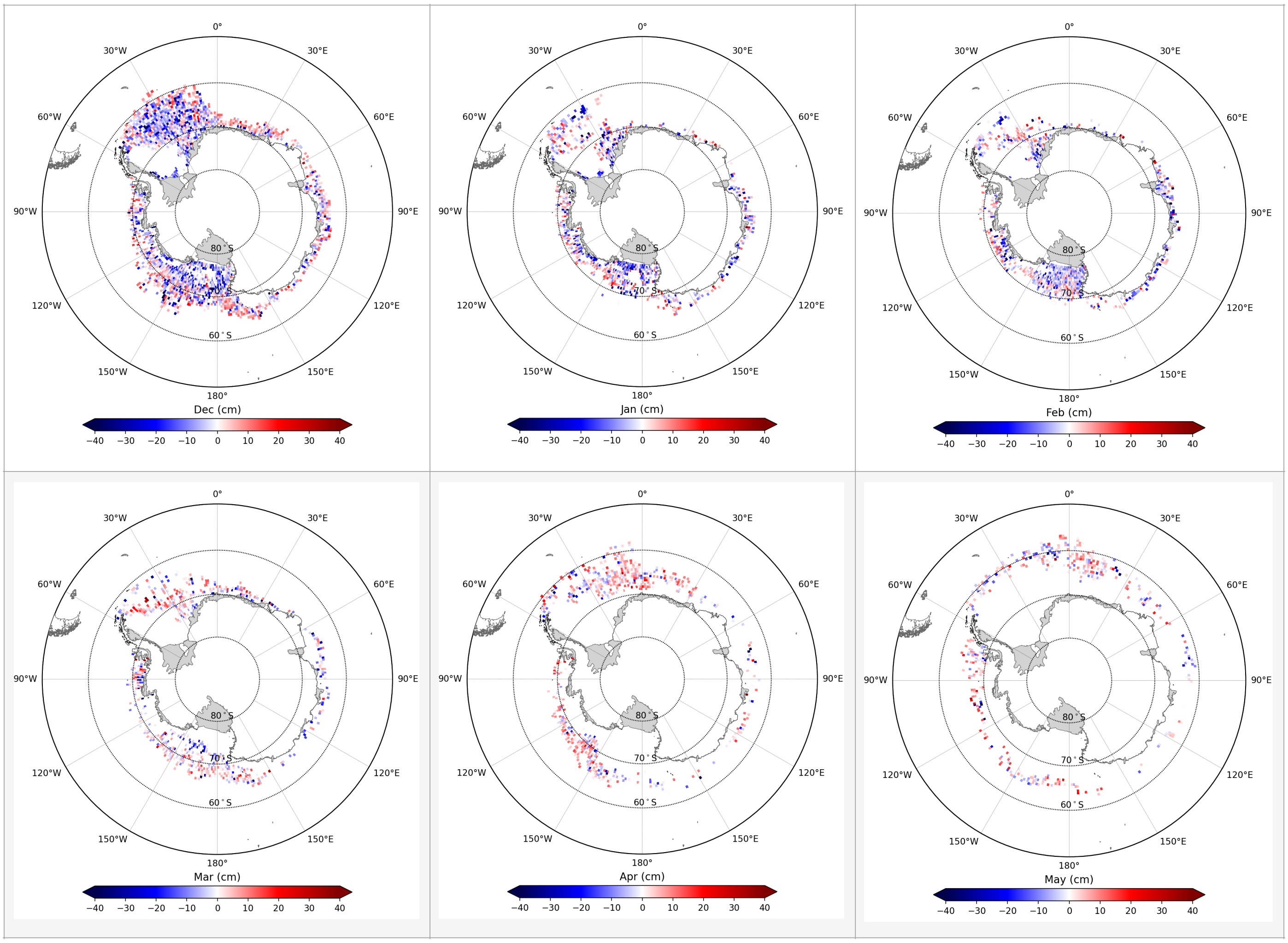
Extra

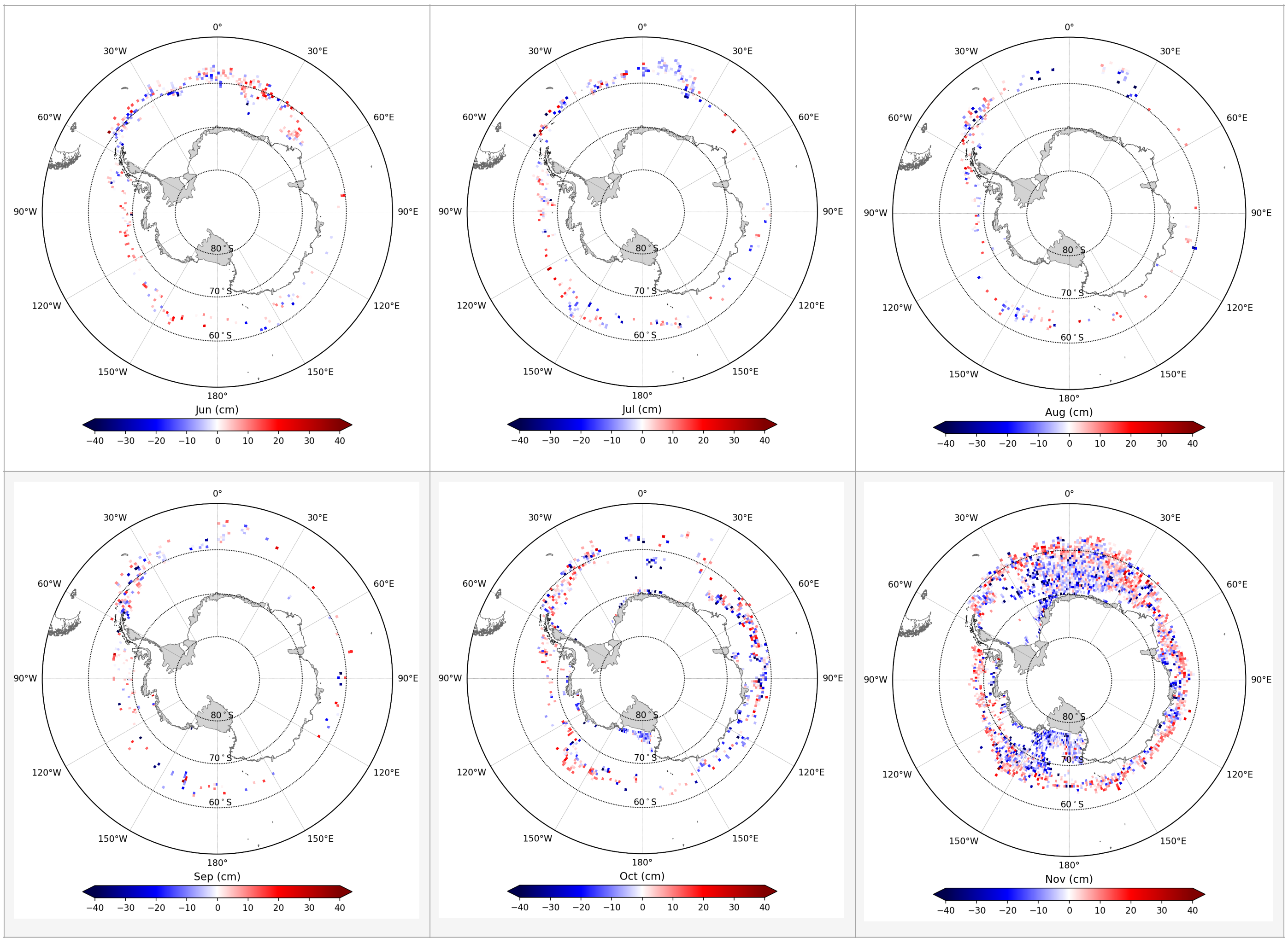
- compute mean of ssh, lat, lon, mss in every grid cell and then interpolate to the centre of the grid (bin-interp)
- there is a bias in both env and cs2 data when looking at difference maps between the bin-interp method and the simple bin average (see right). this difference pattern seems to be consistent with time. Is it just due to the interpolation? But how come that it does not affect the offset for cs2?

- mean value of the difference [bin-interp minus bin] 1.77 to 2.5 cm

Probably best to use just the median in every bin with the threshold for the number of points in a cell







Average O-L offset using mean bin avg for ENV for every month. Is it time mean of all O/L offsets in a month?