Technical note: ODYSEA L2 simulated datasets

Lucile Gaultier

Submitted by:

OceanDataLab (ODL), France

Chronology

Issue	Date	Change record	Author
1.0	Sept 04, 2024	Initial document	Lucile Gaultier

Contents

1 Introduction

This note documents the L2 ODYSEA dataset simulated using the statistical simulator odyseamulator-L2 and the JPL MitGCM LLC2160 - C1440 fully coupled simulation as an input.

The following will give an overview of the input model (section 2), the statistical simulator (3 and a description of the output L2-like data (4).

2 Input model

The LLC2160-C1440 simulation comprises a C1440 configuration of the Goddard Earth Observing System (GEOS) atmospheric model, with 7-km horizontal grid spacing and 72 vertical layers, coupled to a LLC2160 configuration of the Massachusetts Institute of Technology general circulation model (MITgcm) with 2–4-km grid spacing and 90 vertical levels. The ocean model includes tidal forcing. The C1440-LLC2160 simulation has been integrated for 14 months, starting from prescribed initial conditions on January 20, 2020. Hourly atmospheric and oceanic model output of all model variables are available. More information are available at `https://www.nas.nasa.gov/SC21/research/project16.html`

For easier access to the data, the .data binary files originally stored in faces (one face for one region of the globe) have been converted into a global regular 4 km grid and saved in netcdf4. Each netcdf file contains one day of data so 24 hourly time steps and no depth as only the surface data have been included. Each variable is stored separately in directories, except for vectors (current, wind and stress) as both variables are included in the same netcdf file.

The variables 10m wind and surface current have been used as an input to the odyseamulator-L2:

- the ocean current variables are originally stored in the orientation of the grid. They have been rotated in the netcdf files and are now oriented in agreement with the longitude /latitude grid. Surface current are provided in llc2160_daily_latlon_SSC directory. Each file follows the pattern llc2160_[year]-[month]-[day]T000000_SSU-SSV.nc. The eastward current componant is named SSU and the northward componant is named SSV.
- The wind is provided at 10 m heigth in the longitude / latitude orientation. Fields are provided in llc2160_daily_latlon_wind directory. Each file follows the pattern llc2160_[year]-[month]-[day]T000000_geo5_u10m-geo5_v10m.nc. The eastward current componant is named geo5_u10m and the northward componant is named geo5_v10m.

3 Statistical simulator description

The odyseamulator-L2 is a statiscal simulator that directly computes L2 ODYSEA-like observation using OGCM (Ocean General Circlation Model) data as an input for the ocean and wind variables. The noise is randomly computed following the caracteristics provided by the JPL and CNES development team. The Odyseamulator was initially forked from Alexander Wineteer ODYSEA simulator available on github (`https://github.com/awineteer/odysea-science-simulator`). Only the generation of the L2 swath relies on the forked code swath sampling module, it has been adapted to handle more diverse sources of orbits.

The computation of the L2 ODYSEA like data consists of several steps:

1. Generation of the L2-swath sampling
2. Interpolation of the Current and wind data from the OGCM on the swath and projection on the radial directions
3. Compute errors on the swath as a function of the geometry and wind

4. Retrieve northward and eastward current from the radial ODYSEA-like fore and aft observation using an optimal interpolation

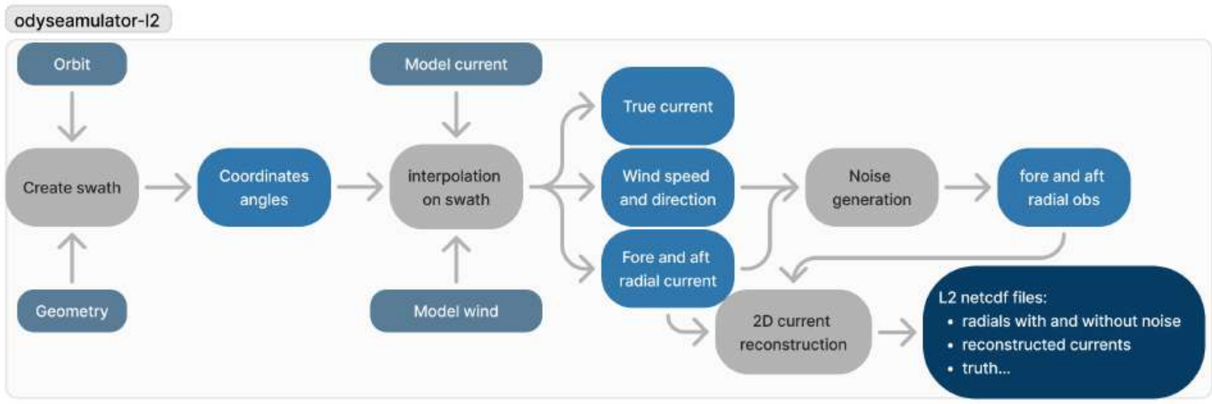


Figure 1: Schemtic view of the simulator

3.1 Generation of L2 swath

The module swath_sampling.py handles the generation of the swath. Several orbits have been made available with several format. The x, y, z and time coordinates are read from this file. If the information to cut the orbit into passes is not provided, the splitting is recomputed. The orbit is then interpolated at the resolution of the L2 swath (5 km along track) and the across track coordinate are projected also with a 5 km cross track direction. The sampling of the swath (5km x 5km) is configurable and thus can be easily modified using the parameter file.

The ODYSEA L2 like data are available for the following orbits:

orbit	height (km)	swath width (km)	look angle	Configuration
orbit 800 km	800	1497	41º	E2, G
orbit 590 km	590	1486	49º	F

Table 2: Sumary of the orbits used to simulate ODYSEA-L2. The orbits and configuration parameters have been provided by CNES.

For the following steps, a loop on passes have been implemented to process each pass (ascending or descending) separately.

The following variables will be saved in the netcdf file at this stage:

lon	longitude
lat	latitude
sample_time	Time from the beginning of the simulation
along_track	Along track distance from start in km
cross_track	Across track distance from nadir in km
encoder_fore	Angle (º) counter-clockwise from the bearing of the satellite for the fore look
encoder_aft	Angle (º) counter-clockwise from the bearing of the satellite for the aft look
azimuth_fore	Angle (º) counter-clockwise from the North for the aft look
azimuth_aft	Angle (º) counter-clockwise from the North for the aft look
bearing	Direction of travel of the satellite platform (º)
radial_angle_fore	Angle (º) clockwise from the North for the fore look
radial_angle_aft	Angle (º) clockwise from the North for the aft look

Table 3: Summary of the variables produced during the construction of the swath and saved later in the netcdf file

3.2 Interpolation of input OGCM data

The wind and current data are interpolated in space and time on the swath grid. Note that the time is constant in the cross track direction. The wind speed and direction are computed from the northward and eastward componants. for each pixel, the fore and aft radiale ocean velocities are retrieved projecting the eastward and northward componants using the encoder angle in the pixel.

The following variables are produced at this stage:

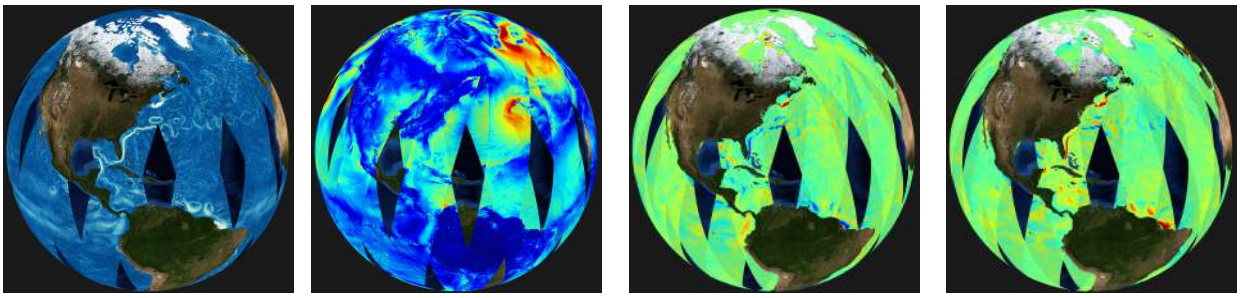
u_model	Eastward velocity from model (Truth) (m/s)
v_model	Northward velocity from model (Truth) (m/s)
ur_nonoise_aft	Radial velocity without noise for the aft look (m/s)
ur_nonoise_fore	Radial velocity without noise for the fore look (m/s)
wind_speed	Wind velocity norm (m/s)
wind_direction	Wind velocity direction from North (º)

Table 4: Summary of the variables available after the interpolation of model data

3.3 Computation of errors

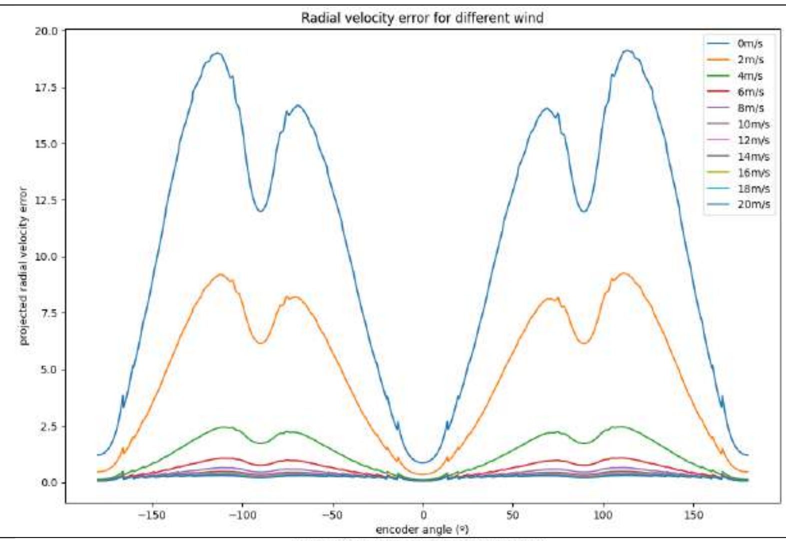
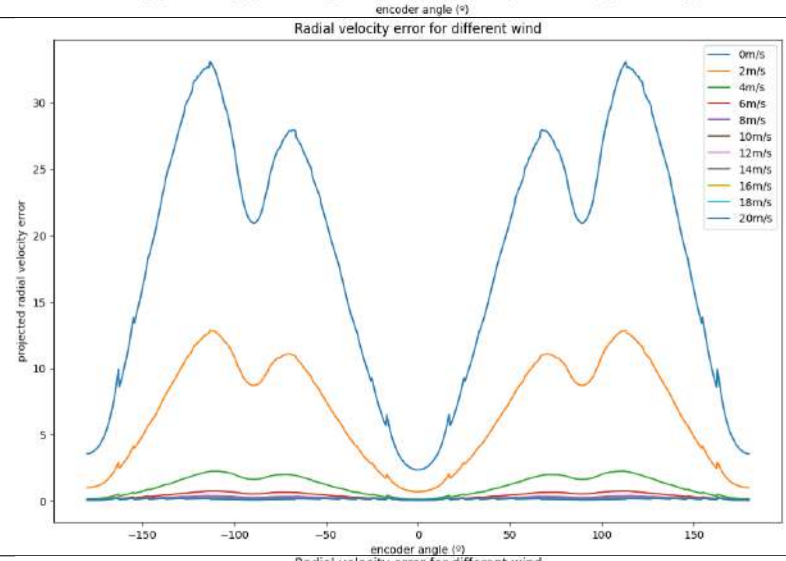
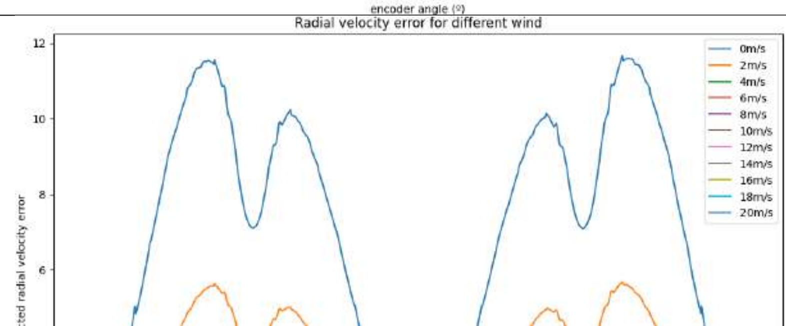
The errors are computed randomly using look up tables provided by the development team. The errors varies with the encoder angle and the wind speed and direction. For each pixel, the standard deviation is retrieved from the look up table using the wind speed and direction computed in the previous step and the encoder angle for the fore and aft direction. This values is used to pick a random sample from a Gaussian distribution.

Three lookup tables have been provided by the team, corrseponding to the three studied configuration, the level of noise depends on the orbit height, look up angle and the performance of the instrument.



(a) Current from Model (b) Wind speed from Model (c) Aft radial look (d) Fore radial look

Figure 2: Model data interpolated on swath, projection of the current on the fore and aft look view for the E2 configuration

Configuration	Error as a function of encoder for different wind speed
E2	
F	
G	

The Fore and Aft Radial errors are added the the fore and aft current to produce L2 like ODYSEA radial currents on the swath:

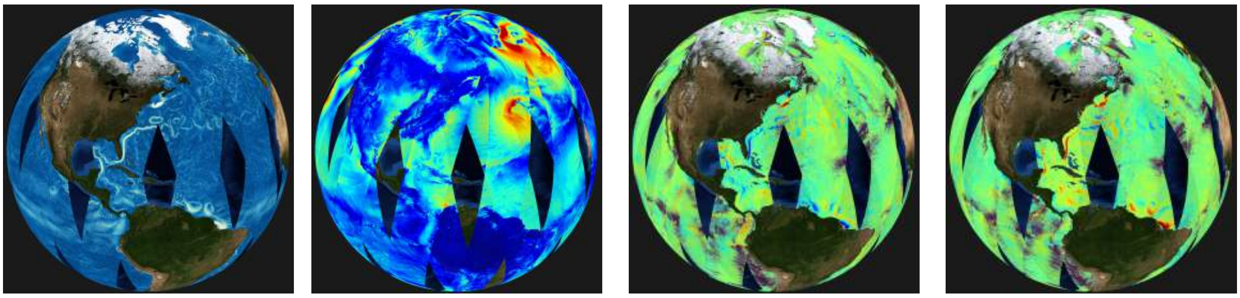
$$ur_{fore} = error_{fore} + ur\_nonoise_{fore}$$

$$ur_{aft} = error_{aft} + ur\_nonoise_{aft}$$

The following variables are added to the netcdf file:

ur_aft	Radial velocity for the aft look (m/s)
ur_fore	Radial velocity for the fore look (m/s)

Table 6: Summary of the variables available after the interpolation of model data



(a) Current from Model (b) Wind speed from Model (c) Aft radial look with noise (d) Fore radial look with noise

Figure 3: Model data interpolated on swath, projection of the current on the fore and aft look view

3.4 Eastward, Northward Current reconstruction

Processing radial components of different angles to a field of vectors follows the same methodology used for HF radars, with radial components combined into vectors. A wide range of techniques has been used for HF radars, as reviewed by Kim et al. (2008), from UnWeighted Least Squares, to Optimal Interpolation. Here we use a bi-variate Weighted Least Square (WLS) approach that is similar to the one described in Kim et al. (2008) and has been implemented to stay close to the measurement data. Let's call Ur_{fore} and Ur_{aft} the radial fore and aft componants computed previously If at least two radial observations with a different line of sight are available, one can apply the following formula to get the eastward (u) and northward (v) components of the velocity vector:

$$\begin{bmatrix} u \\ v \end{bmatrix} = (H^T R^{-1} H)^{-1} H^T R^{-1} \begin{bmatrix} Ur_{fore} \\ Ur_{aft} \end{bmatrix} \tag{1}$$

where R is the covariance matrix of observation error, here assumed diagonal. R-1 is a weight matrix, W. The diagonal terms are set to vary as a function of prescribed uncertainty for each L2B observation, according to observation distance from the grid point, view angle and wind. In this case, if we keep only the two observations on the pixel and assume that the fore and aft errors are equivalents so we can set the matrix R as 1 . H is the observation operator, transforming the state vector to equivalent Line-of-Sight observations, therefore writing:

$$H = \begin{bmatrix} cos(\theta_{fore}) & sin(\theta_{fore}) \\ cos(\theta_{aft}) & sin(\theta_{aft}) \end{bmatrix} \tag{2}$$

where θ_{fore} and θ_{aft} are respectively the fore and aft radial angles as described in section 3.1. Along with the (u, v) retrieval, an uncertainty is computed following the formal error formula of weighted least-squares:

$$E = (H^T R^{-1} H)^{-1}$$

Where E is a (2x2) matrix representing the error covariance of the velocity vector. With R = 1, the previous formulae are simplified and an analytical solution can be computed.

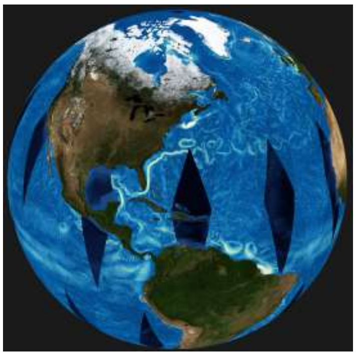
A simple rotation using the bearing angle (platform direction) enables to compute the current in the along track and across track directions.

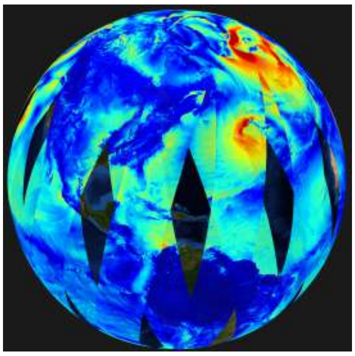
The retrieval of u, v is applied ont the total radial velocities (with noise) as well as on the radial velocities with no noise in order to check the mapping error.

After this final step the following variables are added to the netcdf file

ur_eastward	Eastward velocity reconstructed from ur_fore and ur_aft (m/s)
ur_northward	Northward velocity reconstructed from ur_fore and ur_aft (m/s)
ur_nonoise_eastward	Eastward velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)
ur_nonoise_northward	Northward velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)
ur_ac	cross track velocity reconstructed from ur_fore and ur_aft (m/s)
ur_al	Along track velocity reconstructed from ur_fore and ur_aft (m/s)
ur_nonoise_ac	Cross track velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)
ur_nonoise_al	Along track velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)
u_model_ac	cross track velocity from model (truth) (m/s)
u_model_al	Along track velocity from model (truth) (m/s)

Table 7: Summary of the variables available after the interpolation of model data


(a) Current from Model (truth)


(b) Wind speed from Model


(c) Reconstructed current (norm) from the Aft and Fore look

Figure 4: Reconstruction of current aligned with the longitude/latitude grid. Therer is a strong impact of low wind on the observation of current.

4 Output data description

L2 ODYSEA like data are available for three configuration with different orbits and instrumental performance, the following table summarizes the difference between the different configurations:

Configurations	E2	F	G
Orbit height (km)	800	590	800
Look Angle (º)	41	49	41
Rotation speed (rpm)	14.78	15.02	14.78
Swath width (km)	1497	1486	1497
PRI	$15.\ 10^{-4}$	16.510^{-4}	$16.5\ 10^{-4}$
Beam width	0.15	0.11	0.11

Table 8: Summary of the different configurations used to produce this dataset. E2 and G have similar orbits, F and G have similar instrumental noise.

For each of these configurations, one year is available for the entire globe.

4.1 Nomenclature

Each directory contains data for one configuration, directory is labeled `odysea_l2_global_mitgcm_conf[configuration]`. Each directory contains 10464 files for configurations E2 and G and 10949 files for orbit F.

Each netcdf file contains one pass (ascending or descending). Files are labelled: `odysea_l2_global_mitgcm_conf[configuration]_c[cycle]_p[pass].nc`

WARNING: Note that the splitting into cycle is not correct and there fore eventhough all data are there the repeat of each cycle is not correct. It only impact your study if you are looking to a specific pass and cycle by looking at a file name.

4.2 Netcdf file content

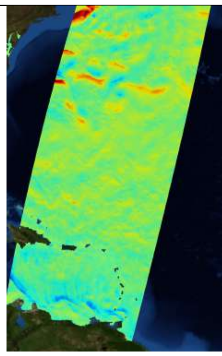
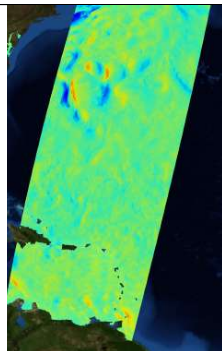
The following tables are a summary of the tables shown previously in the simulator description (section 3). It describes the content of the netcdf files with illustration when it is relevant.

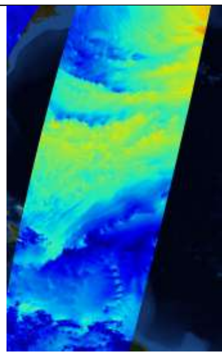
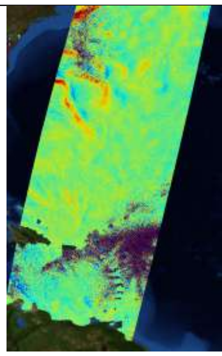
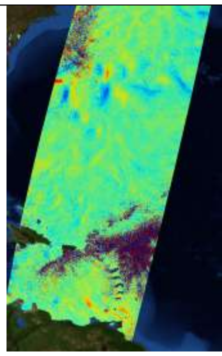
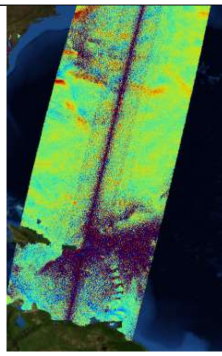
4.2.1 Geometry

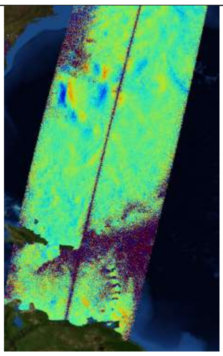
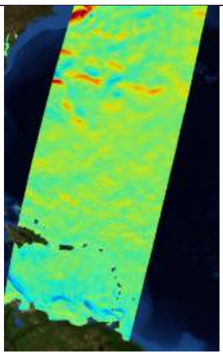
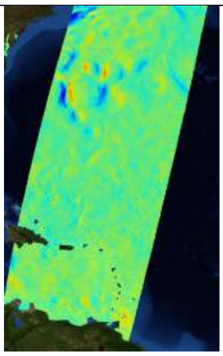
lon	longitude
lat	latitude
sample_time	Time from the beginning of the simulation
along_track	Along track distance from start in km
cross_track	Across track distance from nadir in km
encoder_fore	Angle (^{o}) counter-clockwise from the bearing of the satellite for the fore look
encoder_aft	Angle (^{o}) counter-clockwise from the bearing of the satellite for the aft look
azimuth_fore	Angle (^{o}) counter-clockwise from the North for the aft look
azimuth_aft	Angle (^{o}) counter-clockwise from the North for the aft look
bearing	Direction of travel of the satellite platform (^{o})
radial_angle_fore	Angle (^{o}) clockwise from the North for the fore look
radial_angle_aft	Angle (^{o}) clockwise from the North for the aft look

Table 9: Geometry and coordinate variables

4.2.2 Geophysical variables

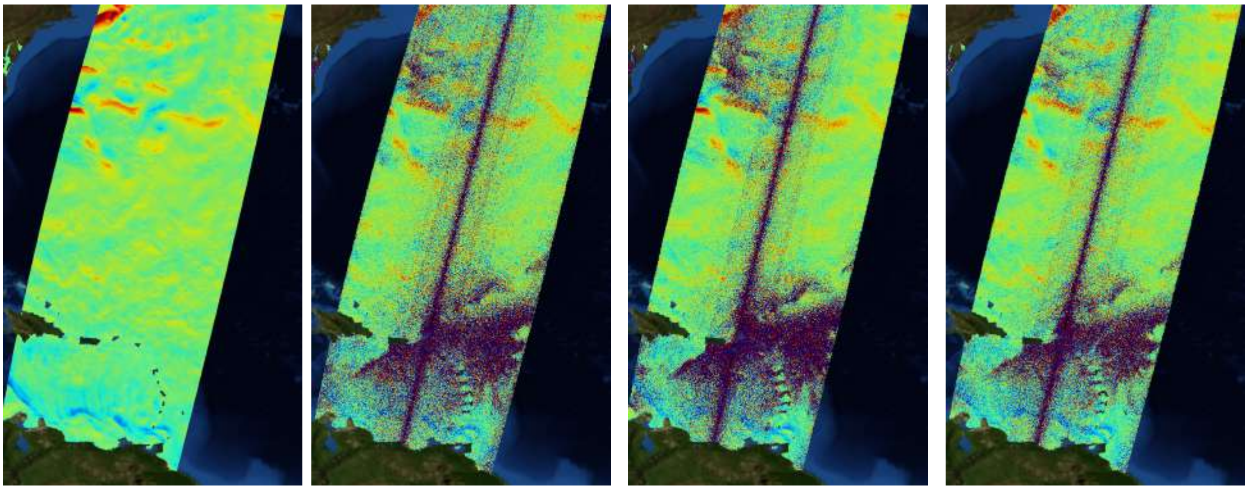
u_model	Eastward velocity from model (Truth) (m/s)	
v_model	Northward velocity from model (Truth) (m/s)	
ur_nonoise_aft	Radial velocity without noise for the aft look (m/s)	

ur_nonoise_fore	Radial velocity without noise for the fore look (m/s)	
wind_speed	Wind velocity norm (m/s)	
wind_direction	Wind velocity direction from North (º)	
ur_aft	Radial velocity for the aft look (m/s)	
ur_fore	Radial velocity for the fore look (m/s)	
ur_eastward	Eastward velocity reconstructed from ur_fore and ur_aft (m/s)	

ur_northward	Northward velocity reconstructed from ur_fore and ur_aft (m/s)	
ur_nonoise_eastward	Eastward velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)	
ur_nonoise_northward	Northward velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)	
ur_ac	cross track velocity reconstructed from ur_fore and ur_aft (m/s)	
ur_al	Along track velocity reconstructed from ur_fore and ur_aft (m/s)	
ur_nonoise_ac	Cross track velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)	
ur_nonoise_al	Along track velocity reconstructed from ur_nonoise_fore and ur_nonoise_aft (m/s)	

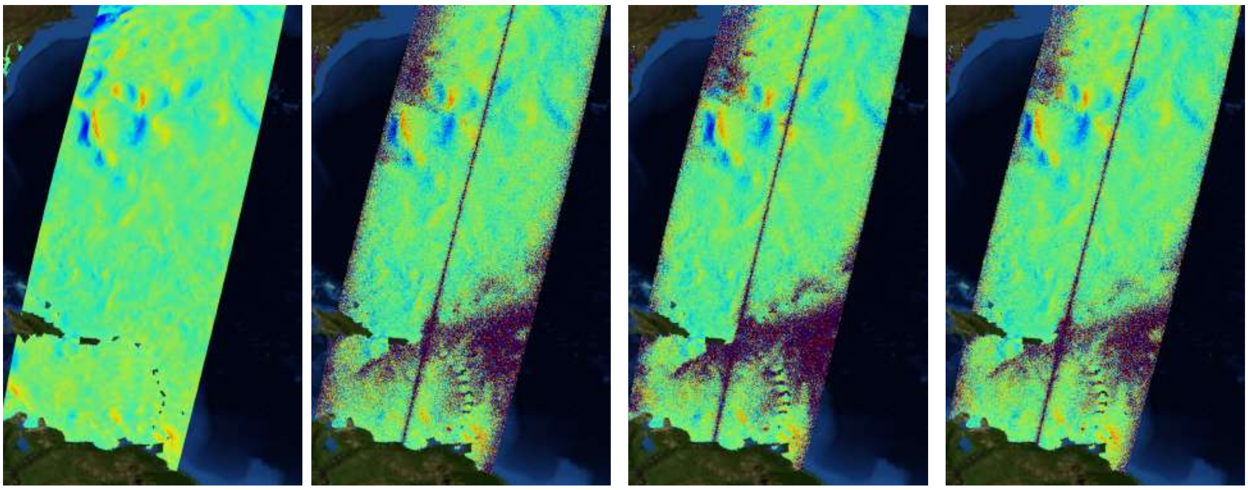
u_model_ac	cross track velocity from model (truth) (m/s)	
u_model_al	Along track velocity from model (truth) (m/s)	

Below is a comparison of the reconstructed speed for the three configuration and comparison to the truth to see the impact of the choises made for the different configuration on the level of noise:



(a) Model (truth) (b) reconstruction from configuration E2 (c) reconstruction from configuration F (d) reconstruction from configuration G

Figure 5: Eastward component of the truth (model, a), from configuration E2 (b), F (c) and G (d)



(a) Model (truth) (b) reconstruction from configuration E2 (c) reconstruction from configuration F (d) reconstruction from configuration G

Figure 6: Northward component of the truth (model, a), from configuration E2 (b), F (c) and G (d)

4.3 Known issues

There are some known issues that we did not have to tackle because of the time constraint. Most studies should not be impacted, still we are working to resolve them as soon as possible

- the splitting into cycle is not correct and there fore eventhough all data are there the repeat of each cycle is not correct. It only impact your study if you are looking to a specific pass and cycle by looking at a file name.
- Nan values are presents in eastward, northward, across track and along track fields at the IDL (-180, 180 line)
- The discretisation of the noise can be seen in the data when looking at very small variations (looks like boxes)