User's Guide
NASA-SSH: JPL Sea Surface Height Anomalies, Version 1

Joshua K. Willis, Séverine Fournier, Emmy Killett, Kevin Marlis and Julie Sanchez
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Abstract

Satellite observations of sea level, or sea surface height, have been a mainstay of NASA's Earth Science Program for over 30 years. An essential variable for climate science, oceanography, and a wide variety of operational uses, measurements of the sea level remain highly valuable. The aim of this data release is to provide improved sea level data products based on radar satellite observations from the early 1990s to the present.

These products will be continuously updated and provided with a latency about two weeks, always based on the most accurate and up-to-date processing available. At present, two classes of data products will be released:

1) an along-track product with once-per-second sampling, broken into daily files for ease of handling, and
2) a simple gridded sea level product, calculated using the daily along-track files.

These products are referred to as "NASA-SSH", and are the most accurate and up-to-date sea level data products provided by NASA, based on radar altimeter satellite missions. They consist of observations from multiple satellites, all homogenized and provided as anomalies relative to a common time mean.

To assist in analysis of these data, a set of ocean and lake basins has been defined. Each sea level data point is assigned a number, corresponding to a specific ocean basin, lake or land mass. The basin masks, their numbers, and a table defining basin connections are also provided as part of this data release.

Please cite these data in the following manner (with the DOI that corresponds to each product):

Willis, J. K., S. Fournier, K. Marlis, E. Killett and J. Sanchez, 2024, NASA-SSH: JPL Sea Surface Height Anomalies, Version 1. PO.DAAC, CA, USA. Dataset accessed [YYYY-MM-DD] at [DOI example: https://doi.org/10.5067/NSREF-AT0V1]

Table of Contents

1. Along Track Daily Files

These files contain observations of sea level, or sea surface height, measured by radar from satellite altimeters. The radar on each satellite is nadir pointing and measures an area 5 to 10 km in diameter directly below the satellite, returning a measurement about once per second. This type of data is referred to as “along-track” data. Variables in the files contain latitude, longitude, date/time, sea surface height and several flags for each of these observations, and are organized into “daily” files, where a single file contains all the measurements from a single satellite on a given day.

Native along-track data from altimetry satellites are usually delivered from the space agencies in files that are organized into passes and cycles. The ground track of a full orbit is composed of an ascending pass and a descending pass (Figure 1). An ascending pass contains data recorded as the satellite travels from the southernmost turning point of its orbit to the northernmost turning point (and vice versa for the descending pass).

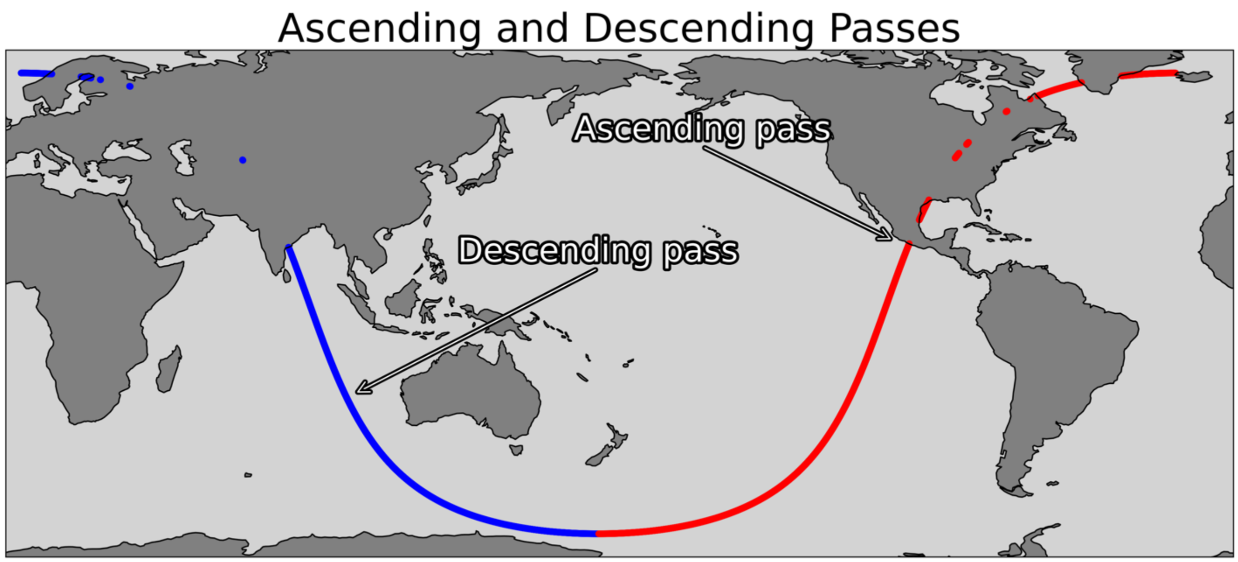

Figure 1: Illustration of a descending and an ascending pass.

The orbital periods of the major "reference" altimetry satellites are all slightly less than two hours, so each pass contains slightly less than an hour of data. Due to Earth's rotation, the satellite does not pass over the same ground track after each orbit. In particular, the reference orbit was chosen so that the ground track repeats after roughly ten days (Figure 2). (This cycle length is different for satellites in orbits which have different altitudes or inclinations than the reference orbit.)

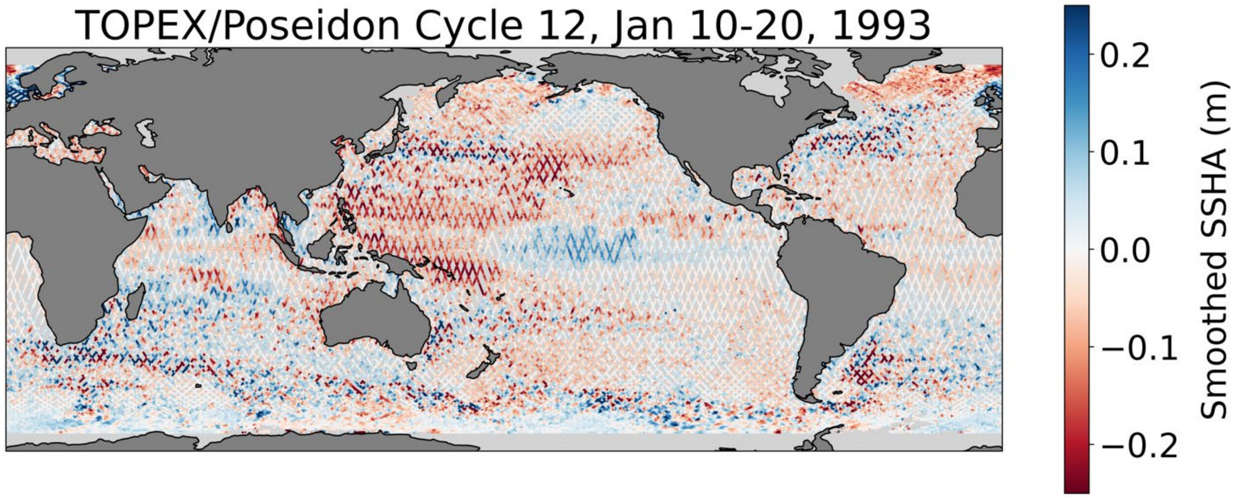

Figure 2: Map of the smoothed SSHA from 10 daily files from January 10 to January 20, 1993.

When using reference satellite data sources distributed as one file per pass, such as Sentinel-6, each daily file is constructed by combining ~25 pass files. When using reference satellite data sources distributed as one file per cycle, such as Goddard Space Flight Center (GSFC), the daily files are constructed by splitting each cycle file into ~10 separate daily files (Figure 3). Both source datasets are distributed by PO.DAAC.

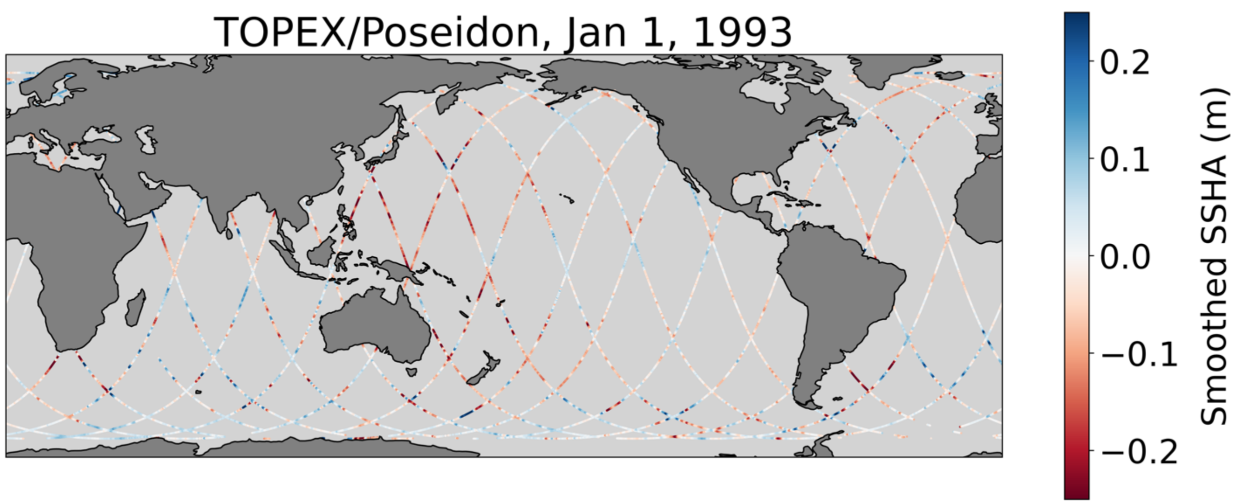

Figure 3: Map of the smoothed SSHA from 1 daily file on January 1, 1993.

Each daily file contains sea surface height anomalies (SSHA) relative to a precalculated "mean sea surface". For example, an SSHA value of zero indicates that the altimetry measurement at that time and location exactly matches the precalculated mean sea surface. The mean sea surface accounts for the persistent effects of spatial variations in the static gravitational field, ocean currents, tides, salinity variations, atmospheric pressure variations, and other long-term oceanographic phenomena. The daily file SSHA values are relative to the DTU21 mean sea surface (Andersen et al. (2023); Andersen (2022)), which was calculated using altimetry data from 1993-2012 (Figure 4).

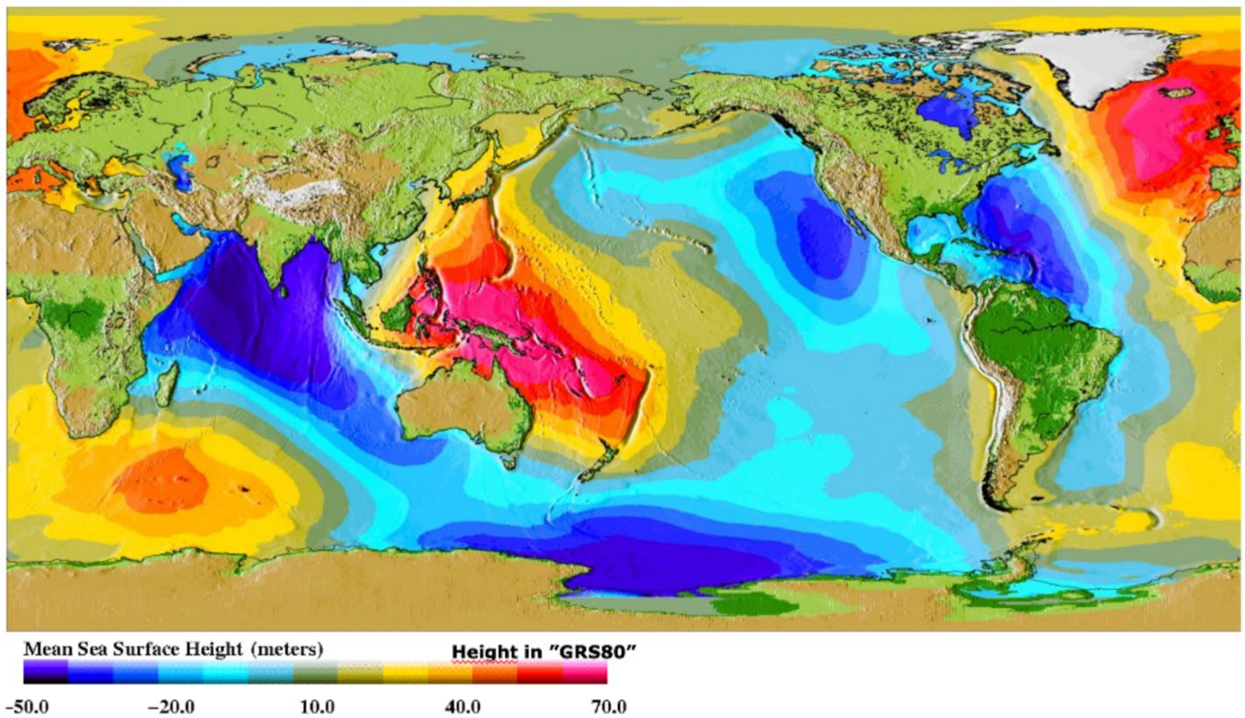

Figure 4: Map of the DTU21 mean sea surface.

Basin flags that have been developed classify the data points as representing land or one of many ocean basins, inland seas, gulfs, bays, lakes, channels, sounds, straits, inlets, bights, or fjords (Figure 5).

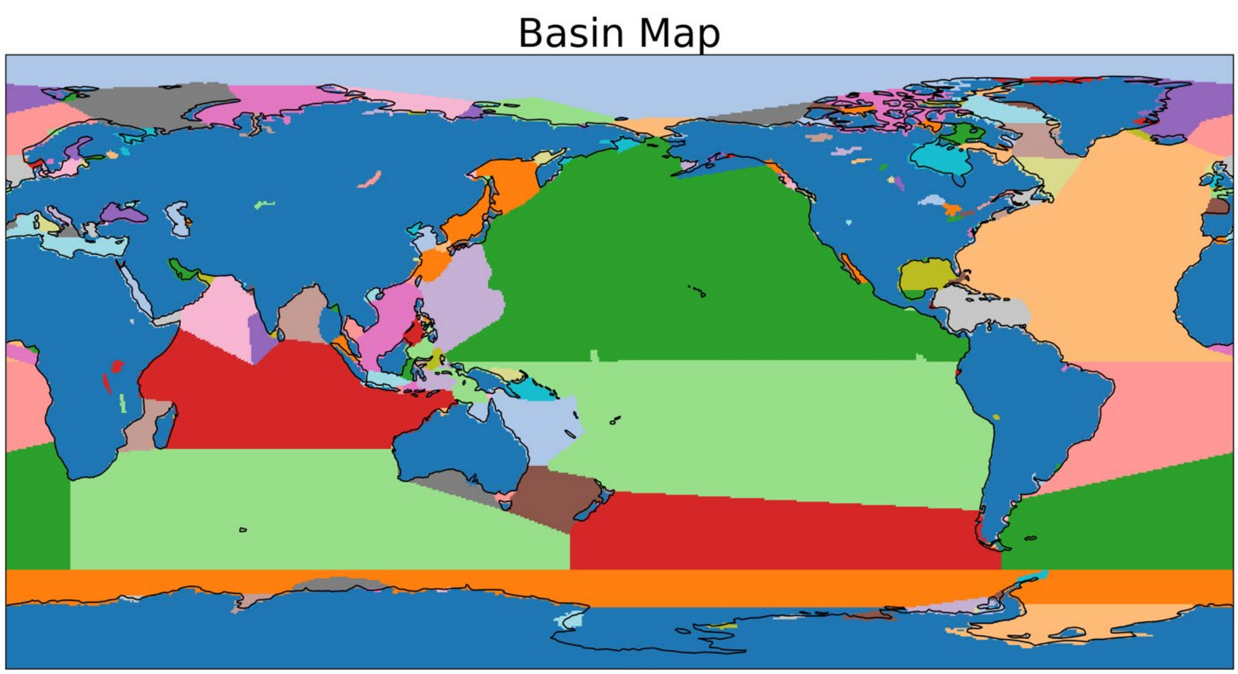

Figure 5: Map of the different basins, color coded.

The basin polygons were derived using products from free vector and raster map data Natural Earth. Ocean basins were derived from version 5.1.0 of the 1:10m Marine Areas, and Lakes were derived from version 5.0.0 of the 1:50m Lakes and Rivers polygons.

Only the 26 largest lakes and inland seas were retained, as most smaller lakes are not typically sampled by traditional nadir altimeters. Some marine basin polygons were joined or split to simplify grouping altimeter data by regions where it is expected to be geographically correlated. For example, southern sections were split from the South Pacific and South Atlantic Oceans to simplify separation of these basins across the South American Peninsula. See Section 4 for additional information about the basin masks and polygons.

An along-track “median filter” was developed in order to flag and ignore outlier points from the original along-track data. The median filter detects and flags points which are more than 5 standard deviations away from a 15-point along-track running median. Prior to the calculation of the along-track median filter, the following source flags were also used to ignore erroneous observations. These flags were originally generated by each satellite data source:

GSFC flags:

- 0: Both corresponding points in neighboring cycles are > 50 cm different
- 1: Radiometer Observation is Suspect
- 2: Attitude Out of Range
- 3: Sigma0 Ku Band Out of Range
- 4: Possible Rain Contamination
- 5: Sea Ice Detected
- 9: Any Applied SSHA Correction Out of Limits

PO.DAAC Sentinel-6 flags:

- range ocean nr qual
- surface classification flag
- rad water vapor qual
- rain flag
- sig0 ocean nr
- swh ocean nr
- ssha nr

A new "NASA flag" is then created which uses the median filter, source flags, and basin flags to indicate which points should be used in future analysis steps (nasa_flag = 0) or not used (nasa_flag = 1). To mitigate problems with sea ice which are not sufficiently addressed by the steps described above, data at latitudes outside of +/- 60 degrees with an SSHA absolute value greater than 1.2 meters are assigned nasa_flag = 1. Similarly, data in the Gulf of Suez are also assigned nasa_flag = 1 due to issues with some tide or other corrections there.

The SSHA data are copied to a new array called ssha_smoothed. Where nasa_flag = 1, the ssha_smoothed values are set to "Not A Number" (NaN). The ssha_smoothed array is then smoothed by a 19-point Gaussian-like normalized filter specific to that satellite's speed (Figure 6). This filter linearly interpolates over NaNs that aren't on the edge of the window. It accounts for NaNs on the edges of the window by mirroring the NaNs to the other side of the 19-point window so the filter remains symmetric.

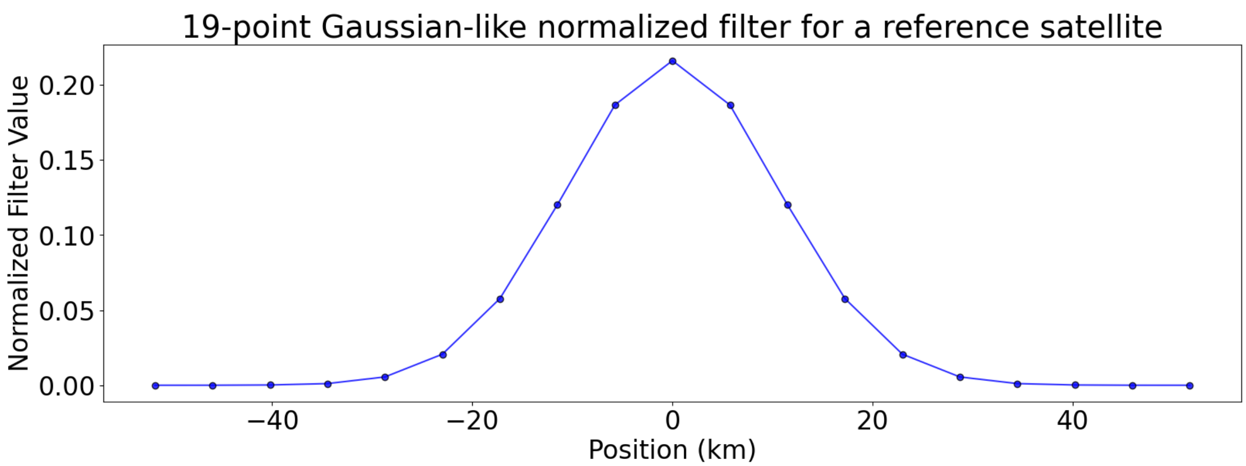

Figure 6: 19-point Gaussian-like normalized filter.

A satellite often crosses over its own recent ground track. If a "crossover" happens within about a week after the last measurement, the ocean surface will be at nearly the same height because oceanographic processes are too slow to significantly move the ocean surface (when averaged over the large altimetry "footprint" where the radar beam reflects off the ocean, and corrected for fast-changing phenomena such as tides and atmospheric loading). As a result, large numbers of altimetry measurement differences at crossover points can be used to detect and correct systematic “orbit” error. Following the method of Le Traon and Ogor 1998, the orbit error has been reduced in the daily file altimetry data. Although these errors are typically referred to as “orbit error”, they actually include any errors or unmodeled phenomena that have a spatial scale of approximately one pass, and time scale of approximately 1 hour. Applying the orbit error correction reduces the RMS variability of the crossovers by a variance of about 2.3 cm, with a slightly higher reduction in the early years of the TOPEX/Poseidon record, and a slightly smaller reduction in the most recent Sentinel-6 MF record (See Figure 7).

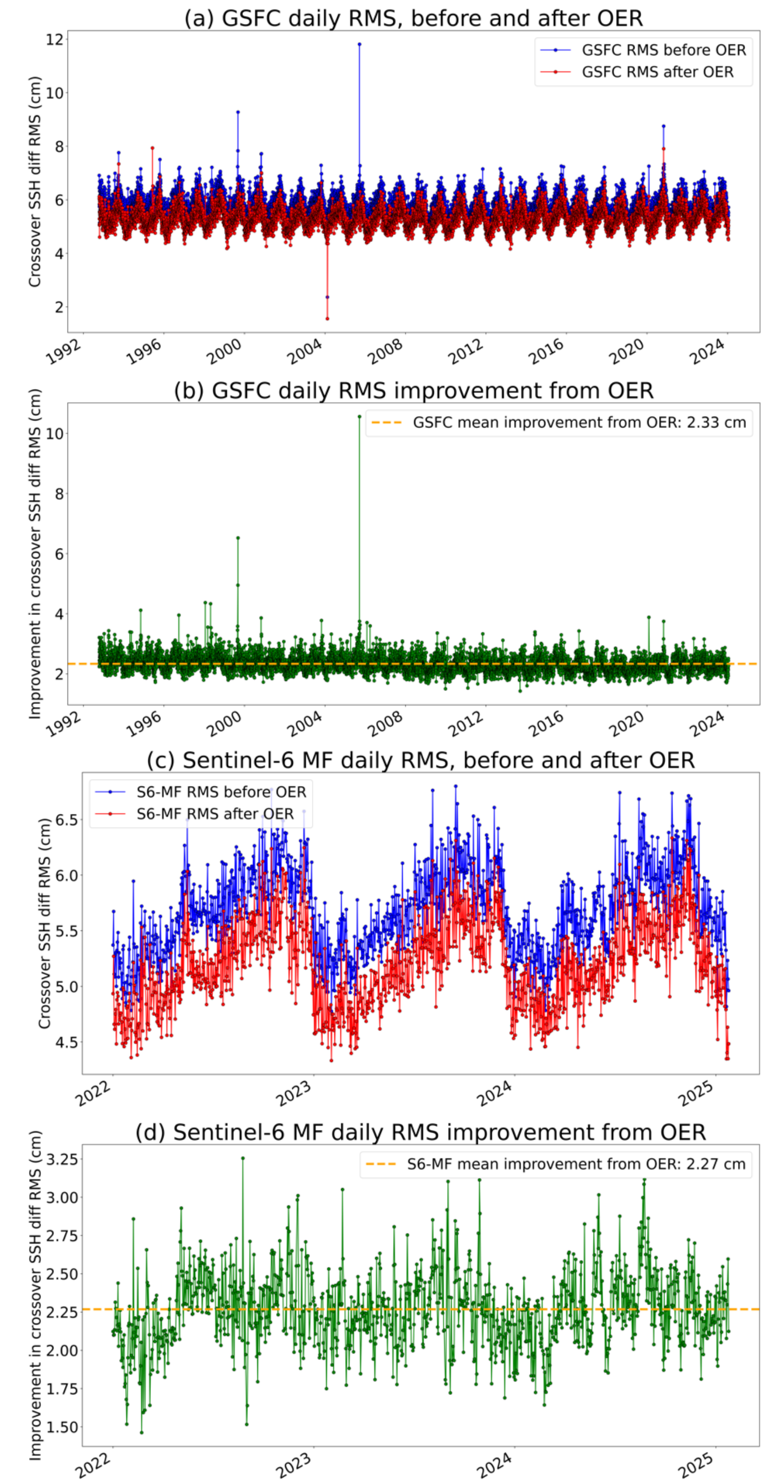

Figure 7: RMS of crossovers, grouped into daily bins. The crossovers are computed using passes within a 10-day period of a given pass. The OER correction is computed by fitting a polynomial to the crossover data (top panel shows GSFC crossovers, bottom panel shows Sentinel-6 crossovers). The reduction in RMS variability is shown by comparing data before and after orbit error reduction is applied (subplots a,c), along with the variance of the difference in RMS before and after (subplots b,d).

While the orbit error reduction (OER) algorithm reduces or removes most large errors that occur on a roughly once per pass basis, it is limited by design to only consider crossovers with absolute differences of less than 30 cm. This means that occasionally very large errors are not corrected. For this reason, remaining bad passes are identified by examining the mean and RMS of SSHA values of crossovers within a 10-day window for each pass. A bad pass is identified if either of these criteria are true:

- Mean > 0.1 meters with at least 15 SSHA crossover values in the pass.
- RMS > 0.27 meters with at least 25 SSHA crossover values in the pass.

Table 1: List of the variables included in the daily files, and their description.

Variable	Summary
ssha	Sea surface height anomalies, relative to mean sea surface DTU21. Select valid data points by only loading points with nasa_flag = 0.
ssha_smoothed	Smoothed sea surface height anomaly values computed using a 19-point filter. The nasa_flag is applied prior to filter and should not be used to remove points from this field.
time	UTC time in seconds since 1990-01-01 00:00:00.
latitude	Latitude in degrees
longitude	Longitude in degrees east (0 to 360)
basin_flag	Basin ID number mapping each observation to a geographic basin
basin_names_table	Table mapping basin ID numbers to basin names
pass	Pass number
cycle	Cycle number
dac	The dynamic atmospheric correction (DAC) is an additive correction that has been applied to remove atmospheric effects. Subtract this field from ssha or ssha_smoothed to un-apply this correction.
oer	Orbit Error Reduction array. Add this variable to ssha and ssha_smoothed to reduce orbit error.
median_filter_flag	Median filter flag is set to 0 for good data, 1 for data that fail a 5 standard deviation filter relative to a 15-point along-track median. See documentation for details.

nasa_flag	Quality flag to be used for ssha, (already applied to ssha_smoothed). nasa_flag is set to 0 for data that should be retained, and 1 for data that should be removed.
source_flag	Data flags provided by the satellite data source, some of which are used to calculate nasa_flag. See documentation for details.

2. Simple Grids

In order to create a more versatile and intuitive data product for sea surface height (SSH), along-track observations of sea level are interpolated onto a uniform latitude/longitude grid using a very simple, weighted-average, gridding technique. These “simple grid” files contain a 2-D map of sea surface height measurements from satellite radar altimeters.

The radar measures an area 5 to 10 km in diameter directly below the satellite, returning a measurement about once per second. This type of data is referred to as “along-track” data, but gridded data are preferred for many analyses in climatology and oceanography because they simplify time series analysis and the averaging process helps prevent biases due to missing data.

Each simple grid file is constructed using 10 days (the approximate time it takes for the reference satellite’s ground track to repeat) of along-track data centered on the date of the simple grid file. Each simple grid file is constructed by averaging data points with coverage as shown in Figure 8:

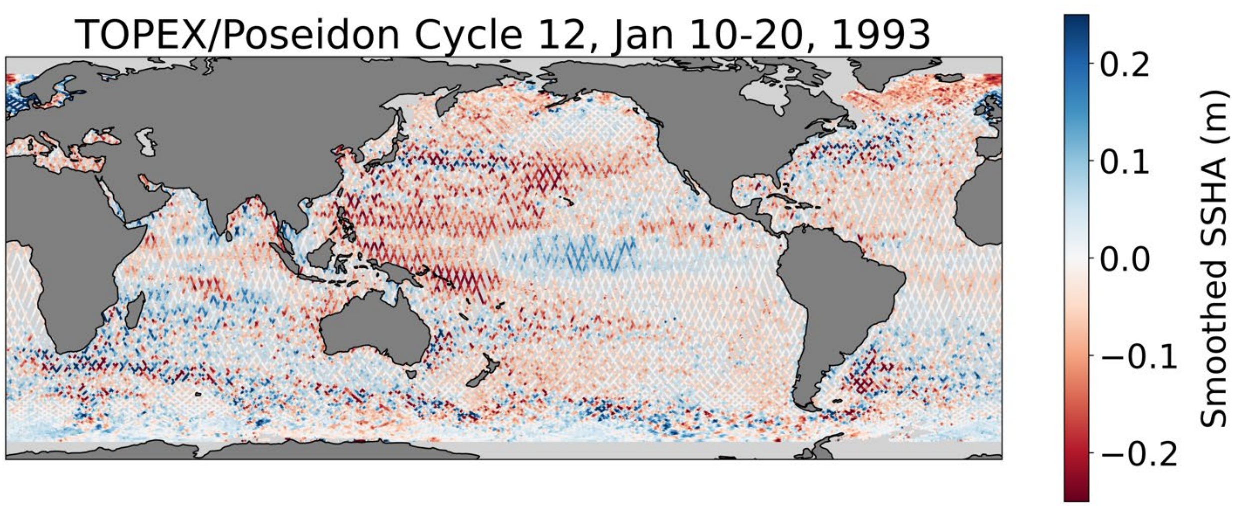

Figure 8: Map of the smoothed SSHA from 10 daily files from January 10 to January 20, 1993.

The simple grid file contains the following interpolated estimate of sea surface height, for approximately the same time period (Figure 9):

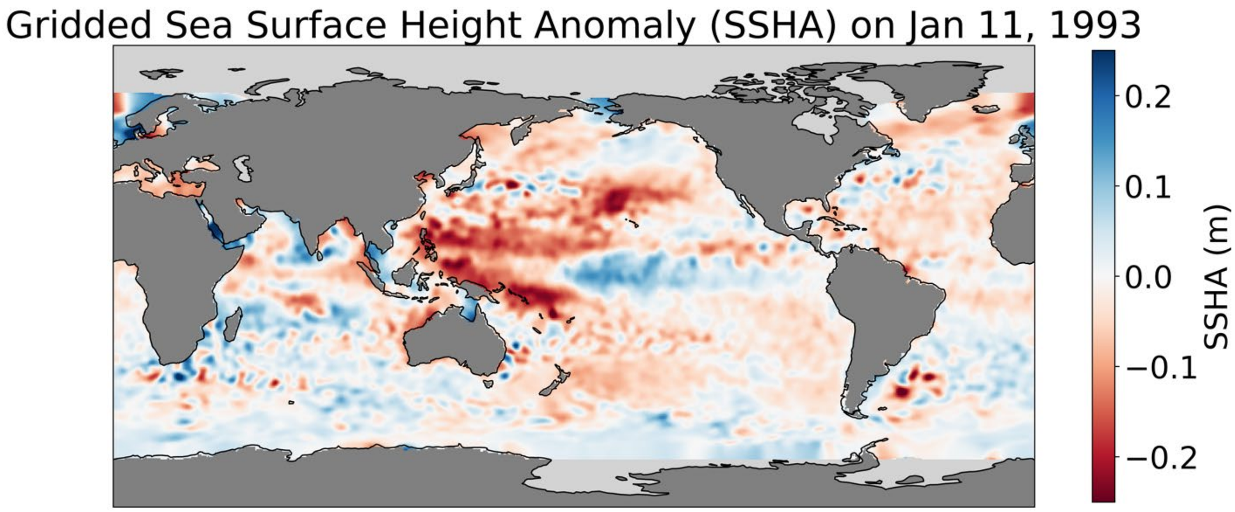

Figure 9: Map of the gridded SSHA over 10 days.

A simple grid file is created to represent the sea surface height every 7 days (every Monday). Since 10 days of data are used to produce a simple grid every 7 days, maps from simple grid files separated by only 7 days typically have some data in common so they are not completely independent. However, the large-scale patterns of sea level change very little in 7 days, so we do not expect this to result in significant systematic errors in the grids. In addition, future versions of these grids will include data from additional satellites, making the 7-day resolution more appropriate.

The along-track altimetry data were resampled onto a regular 0.5 x 0.5-degree grid using the resample_gauss function from the Python pyresample library. This function applies a Gaussian weighting scheme to resample the data, where the influence of each data point decreases with distance from the grid point. Each grid point averages over a maximum of 500 altimetry data points (using more than 500 points does not significantly change the results) that were taken within 600 km of the grid point while respecting basin connection rules (see below in Section 3). The width of the Gaussian weighting function is set to 100 km, and the variable “counts” shows how many along-track data points were averaged to create each gridded average value (Figure 10).

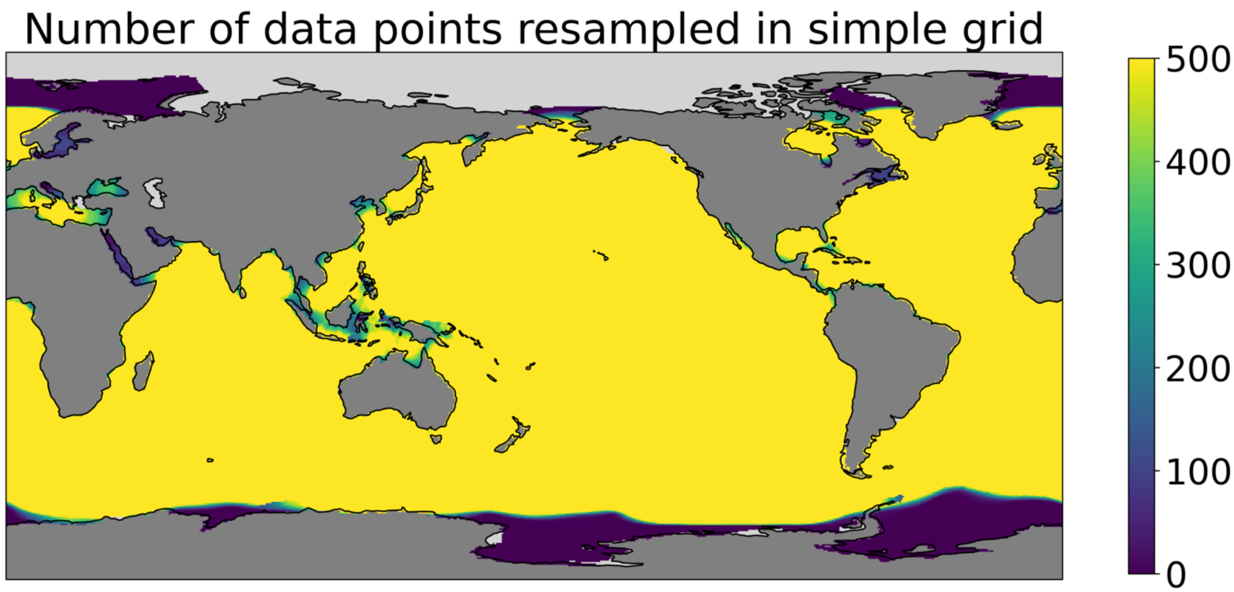

Figure 10: Map of the number of along-track data points used to compute each grid SSHA value.

Basin flags that have been developed classify the along-track data points as representing land or one of many ocean basins, inland seas, gulfs, bays, lakes, channels, sounds, straits, inlets, bights, or fjords (Figure 11).

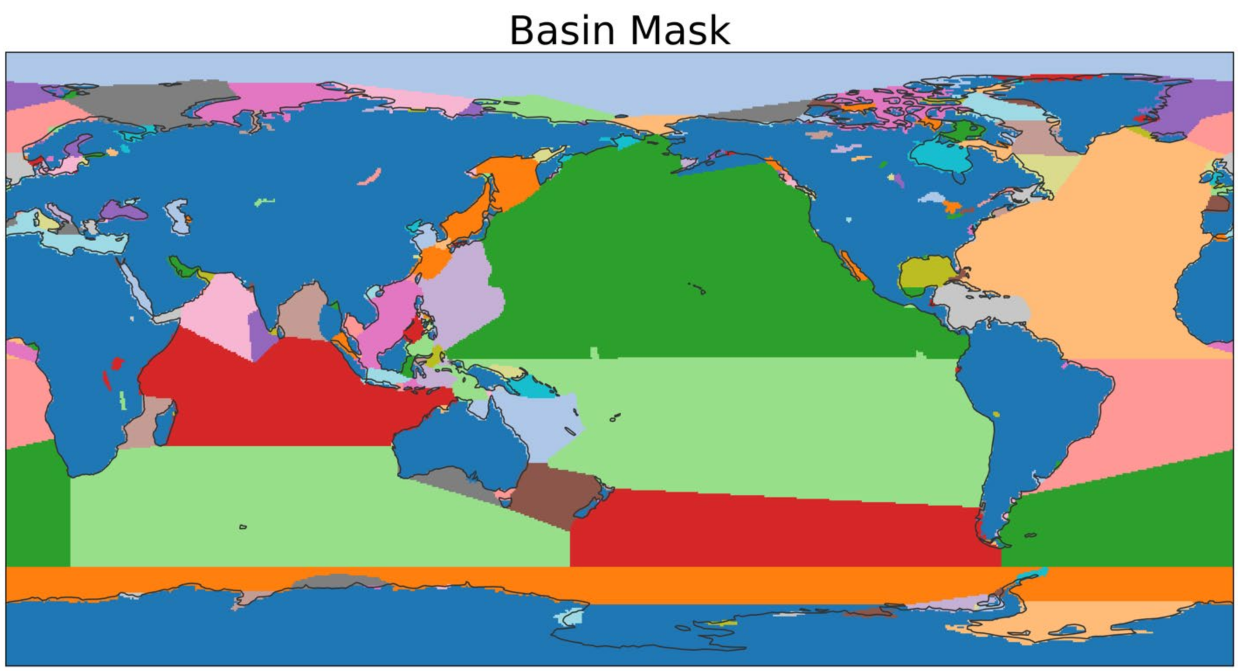

Figure 11: Map of the different basins, color coded.

The basin polygons were derived using products from free vector and raster map data Natural Earth. Ocean basins were derived from version 5.1.0 of the 1:10m Marine Areas, and Lakes were derived from version 5.0.0 of the 1:50m Lakes and Rivers polygons.

Only the 26 largest lakes and inland seas were retained, as most smaller lakes are not typically sampled by traditional nadir altimeters. Some marine basin polygons were joined or split to simplify grouping altimeter data by regions where it is expected to be geographically correlated. For example, southern sections were split from the South Pacific and South Atlantic Oceans to simplify separation of these basins across the South American Peninsula.

In addition to the polygons themselves, a table listing connections between polygons is also provided. This allows users to select observations in regions that are likely to be correlated over time scales of days to weeks or longer. Each polygon carries a unique numerical identifier (the Arctic Ocean is 1, the Southern Ocean is 2, etc.). In the simple grid files, the variable “basin_flag” associates each grid point with its basin numerical identifier, and the variable “basin_names_table” lists the name of that basin. For each identifier, the connection table lists all the other identifiers that polygon is connected to. When averaging data points to create an average value in the simple grid file at some location, only points that are in that basin or a connected basin are used. For example, consider basins connected to the North Pacific Ocean (shown below in Figure 12). Since the Caribbean is not connected to the North Pacific Ocean, the simple grid value directly to the west of Panama does not include any data points in the Caribbean even if those points are within 600 km.

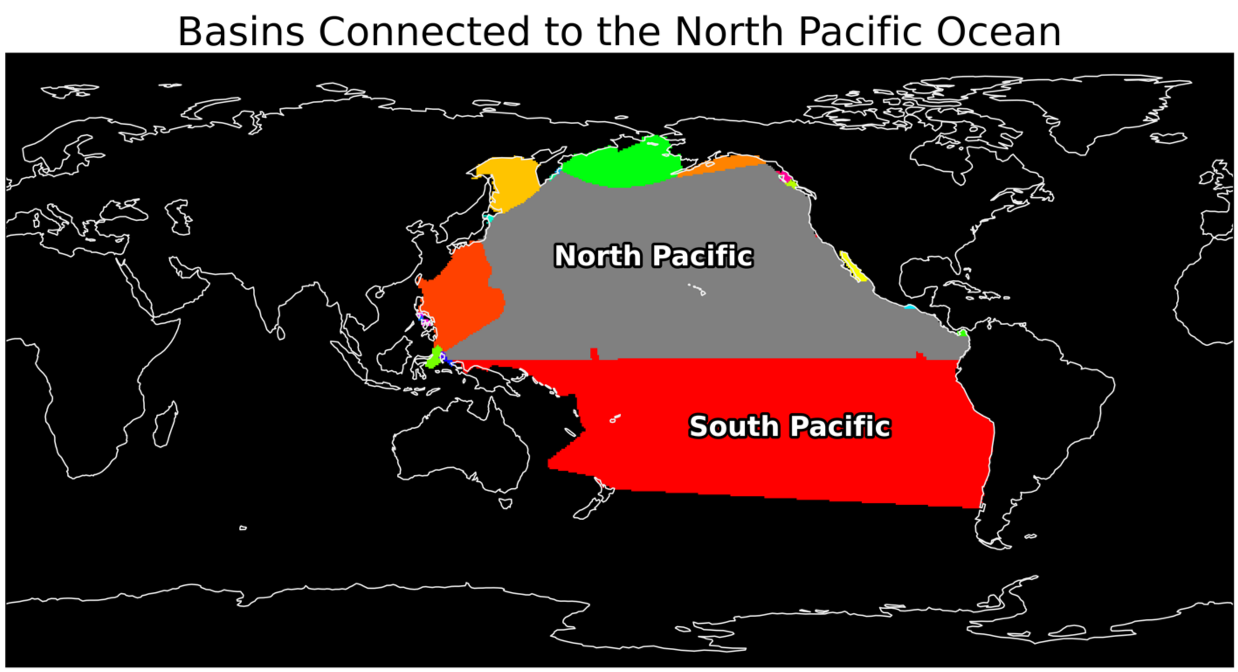

Figure 12: Map of all the basins connected to the North Pacific basin.

Similar to the daily file along-track data, each simple grid file contains values of sea surface height anomalies (SSHA) relative to a precalculated "mean sea surface". An SSHA value of zero indicates that the altimetry measurement at that time and location exactly matches the precalculated mean sea surface. The mean sea surface accounts for the persistent effects of spatial variations in the static gravitational field, ocean currents, tides, salinity variations, atmospheric pressure variations, and other long-term oceanographic phenomena. The simple grid SSHA values are relative to the DTU21 mean sea surface (Andersen et al. (2023); Andersen (2022)), which was calculated using altimetry data from 1993-2012 (Figure 13).

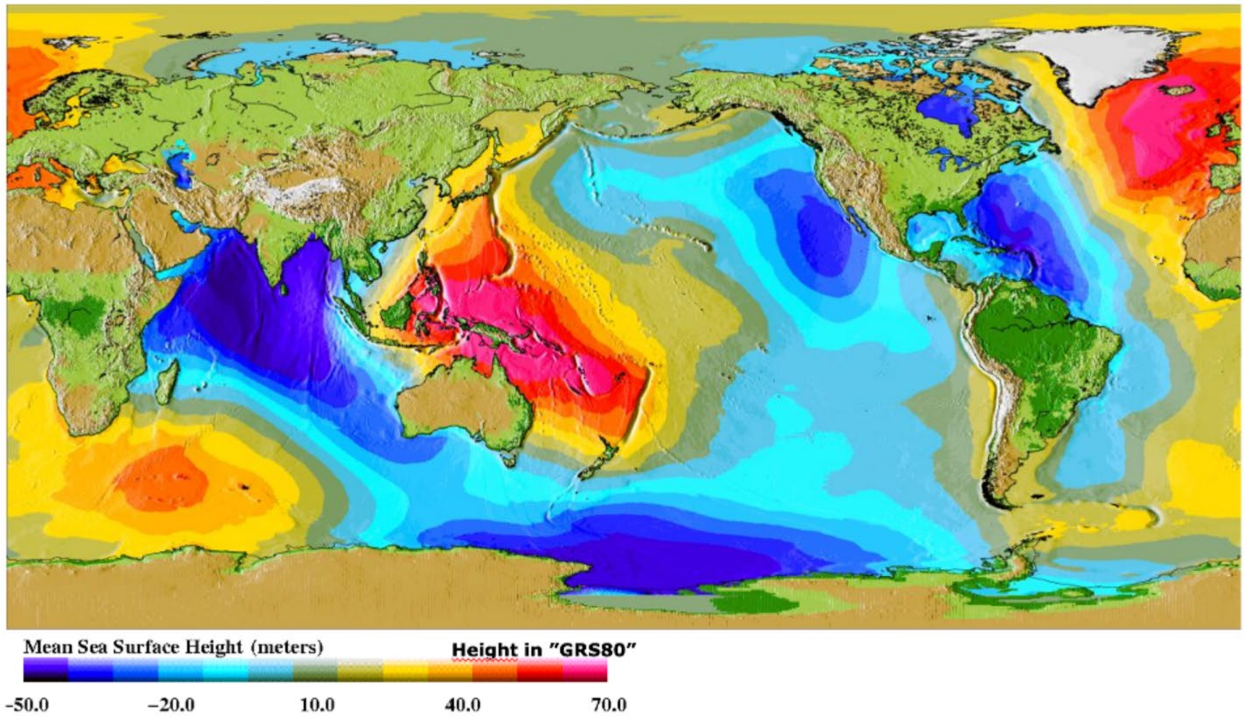

Figure 13: Map of the DTU21 mean sea surface.

Table 2: List of the variables included in the simple grid files, and their description.

Variable	Summary
ssha	Gridded averages of sea surface height anomalies relative to mean sea surface DTU21.
time	UTC time in seconds since 1990-01-01 00:00:00.
latitude	Latitude in degrees (-89.75 to 89.75 every 0.5 degrees)
longitude	Longitude in degrees east (0.25 to 359.75 every 0.5 degrees)
basin_flag	Basin ID number mapping each grid point to a geographic basin
basin_names_table	Table mapping basin ID numbers to basin names
counts	Number of data values used in weighting each element in SSHA (maximum count = 500).
source_files (attribute)	Along-track daily files used to create this simple grid. See daily file documentation for details.

3. Basin Masks

This collection of files contains geographically registered polygons defining the shape of ocean basins and large lakes across the globe, designed for use in gridding and analysis of satellite-based observations of sea level.

The polygons were derived using products from free vector and raster map data Natural Earth (https://naturalearthdata.com). Ocean basins were derived from version 5.1.0 of the 1:10m Marine Areas:

https://www.naturalearthdata.com/downloads/10m-physical-vectors/

and Lakes were derived from version 5.0.0 of the 1:50m Lakes and Rivers polygons:

https://www.naturalearthdata.com/downloads/50m-physical-vectors/

Only the 26 largest lakes and inland seas were retained, as most smaller lakes are not typically sampled by traditional nadir altimeters. Some marine basin polygons were joined or sometimes split in order to simplify grouping altimeter data by regions where it is expected to be geographically correlated. For example, southern sections were split from the South Pacific and South Atlantic Oceans to simplify separation of these basins across the South American Peninsula.

In addition to the polygons themselves, a table listing connections between polygons is also provided. This allows users to select observations in regions that are likely to be correlated over time scales of days to weeks or longer. Each polygon carries a unique numerical identifier (the Arctic Ocean is 1, the Southern Ocean is 2, etc…). For each identifier, the connection table lists all of the other identifiers that polygon is connected to. This is used in the NASA-SSH gridding process to down-select data used to estimate sea level at a specific location.

Both the ocean and lake polygons themselves, and the connection table can be easily visualized in Google Earth (or other geographic mapping software) using the KMZ file provided. The connection table provides sets of polygons connected to each individual feature. For example, the Arctic Ocean polygon is connected to the Beaufort Sea, the Greenland Sea, the Barents Sea, etc. These can be easily visualized by turning on subsets of features in the Basin Connections folder within the KMZ file.

Files contained in this dataset include:

- basin_files.tar.gz – a tar gzip file that contains a .dbf, .prj, .shx, and .shp Shape file that can be loaded into a geographic mapping program such as QGIS or Google Earth. This contains the polygon definitions, including their names.
- basin_name_table.txt – an ascii text file containing a list of all the basin ID numbers and simplified version of the basin names, separated by a colon ":". A few basins were created for this dataset and do not have common geographic names. These are given names based on their ID number for example "Feature ID: 240".
- basin_connection_table.txt – an ascii text file containing a list of all the basin ID numbers that are geographically connected to a given basin ID. The basin ID number in question is listed first on each row, and the connected ID numbers follow a colon ":", in a comma separated list.

- NASA-SSH Basins.kmz – This KMZ file contains all of the polygon definitions, along with the set of polygons for each basin that shows which basins it is connected to. If loaded into Google Earth, it will create a folder in the Google Earth "Places" panel called "NASA-SSH Basins". Below this, two subfolders will be created, one will be called "All Basin Polygons" and will contain all of the basins polygons colored red. Clicking on any one of the polygons will show the Basin ID number and name of this polygon. The second subfolder is called "Basin Connections" and contains a list of subfolders, one for each Basin ID. These can be turned on 1 at a time and will show a given basin polygon and all of the polygons it is connected to.

4. Homogenization of Data

In order to combine observations from different satellites into a single, homogeneous data set, several steps are required. These include:

1) Ensuring a consistent reference frame is used for orbit calculations across each satellite
2) Ensuring consistent geophysical corrections are applied, including wet/dry tropospheric corrections, ionospheric correction, dynamic atmospheric corrections
3) Ensuring consistent tidal models are applied for the solid earth, ocean, pole tide, load tide, long-period tide, and internal tide
4) Ensuring consistent sea state bias corrections are applied
5) Removal of a consistent, high-resolution time-mean sea surface height field
6) Removal of mission-to-mission biases

At present, the along-track data set consists of a continuous record built using only reference altimeter missions: TOPEX/Poseidon, Jason-1, Jason-2, Jason-3 and Sentinel-6 Michael Freilich. For the majority of the data record, the along-track data is obtained from the multi-mission data set produced by Goddard Space Flight Center (GSFC): Integrated Multi-Mission Ocean Altimeter Data for Climate Research Version 5.2 (Beckley et al., 2024; https://podaac.jpl.nasa.gov/dataset/MERGED_TP_J1_OSTM_OST_CYCLES_V52). Observations from this data set are used to create the NASA-SSH along-track daily files, through the last available cycle of data in this record. Beyond that, data from the Sentinel-6 Michael Freilich (S6-MF) satellite are used (https://podaac.jpl.nasa.gov/Sentinel-6) to bring the record to within about two weeks of present day.

The above homogenization steps are carried out as part of the generation of the original GSFC data product. However, for the NASA-SSH product, position-dependent correction was applied to give sea surface height anomalies relative to an updated mean sea surface, DTU21 (Andersen (2022); https://data.dtu.dk/articles/dataset/DTU21_Mean_Sea_Surface/19383221) in order to be consistent with data from the Sentinel-6 mission.

For the Sentinel-6 along-track data, sea surface height anomalies from the S6-MF product were selected to be consistent with GSFC data for steps 1) – 5) listed above. To level the S6-MF data with the GSFC data, simple grids (computed as detailed in Section 2 on "Simple Grids") were calculated for a 6-month overlap period ending with the most recent complete cycle provided in the GSFC data. The grids were globally averaged over the ocean between 66° N & S to compute a time series of the global mean difference between the S6-MF and GSFC data. This was averaged over the 6-month period in order to provide an offset that brings the recent S6-MF data in line with the processed GSFC data.

Beyond this, the along-track data undergoes quality control, smoothing, orbit error correction and removal of bad passes before product release, as discussed in Section 2.

5. References

Andersen, Ole Baltazar, Stine Kildegaard Rose, Adili Abulaitijiang, Shengjun Zhang, and Sara Fleury. "The DTU21 global mean sea surface and first evaluation." *Earth System Science Data* 15, no. 9 (2023): 4065-4075.

Andersen, Ole Baltazar (2022). DTU21 Mean Sea Surface. Technical University of Denmark. Dataset. https://doi.org/10.11583/DTU.19383221.v1

Beckley, B.; Zelensky, N.P.; Holmes, S.A.; Lemoine, F.G.; Ray, R.D.; Mitchum, G.T.; Desai, S.; Brown, S.T.. 2024. Integrated Multi-Mission Ocean Altimeter Data for Climate Research Version 5.2. Ver. 5.2. PO.DAAC, CA, USA. Dataset accessed at https://doi.org/10.5067/ALTCY-TJA52

Le Traon, P.-Y., and F. Ogor (1998), ERS-1/2 orbit improvement using TOPEX/POSEIDON: The 2 cm challenge, *J. Geophys. Res.*, 103(C4), 8045–8057, doi:10.1029/97JC01917.

The preparation of this document and the data it describes was carried in part out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).