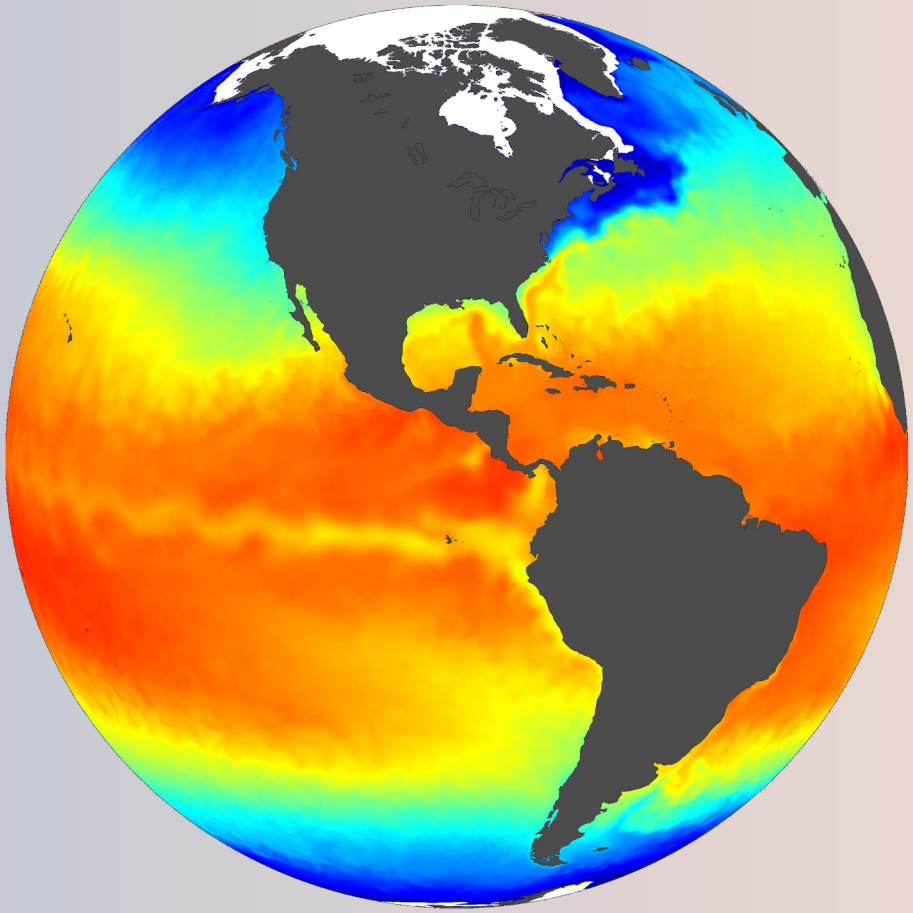


# GHRSST Data Discovery and Cataloguing Service : A federated access to all sea surface temperature data


Jean-François Piollé[1], Edward M. Armstrong[2], Erwan Bodéré[1], Dominique Briand[1],

[1]Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), France.
Jean.Francois.Piolle@ifremer.fr
[2]Jet Propulsion Laboratory, California Institute of Technology, United States of America.
edward.m.armstrong@jpl.nasa.gov


***To provide operational users and the science community with the SST measured by the satellite constellation***

Jet Propulsion Laboratory
California Institute of Technology

GHRSST
*Group for High Resolution Sea Surface Temperature*

*IGARSS 2023*

*Pasadena, CA*

*19 July 2023*

WE1.R16: SST Applications and the GHRSST

# Motivation

- Provide a data management infrastructure based on open services to decentralize the production, storage, dissemination and discovery of GHRSST data products
  - Allow new GHRSST data producers to independently maintain the curation of their datasets
  - Allow existing GHRSST data producers to play “as is” in the new environment. No fundamental changes needed.
  - Minimize stovepipe dependencies
  - Improve user experience and data discovery

# GHRSST Distribution (Current)

GHRSST pioneered a Regional/Global Task Sharing Framework (R/GTS) which uses a scientifically and technically feasible strategy to acquire existing SST data products, add additional information and create a new generation of products in a common format.

The data management from production, distribution and archiving is “stove piped” across a few entities

# GHRSST Distribution (Current)

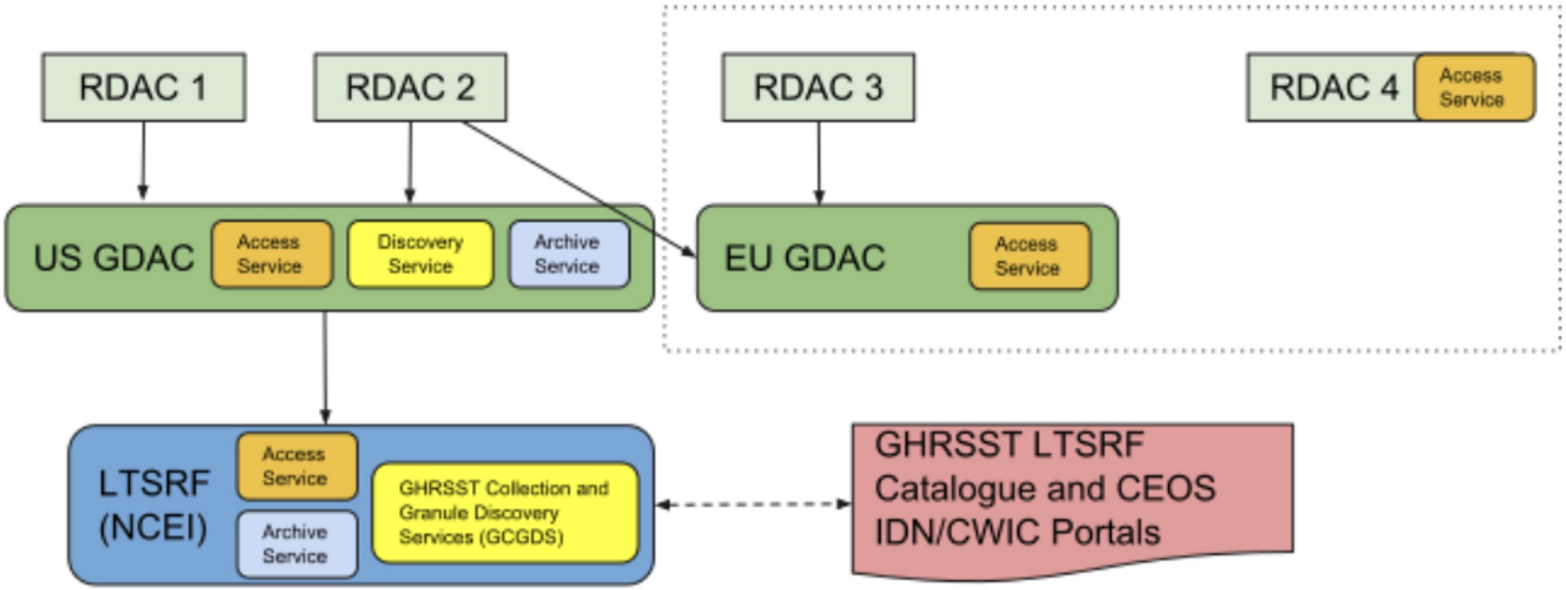


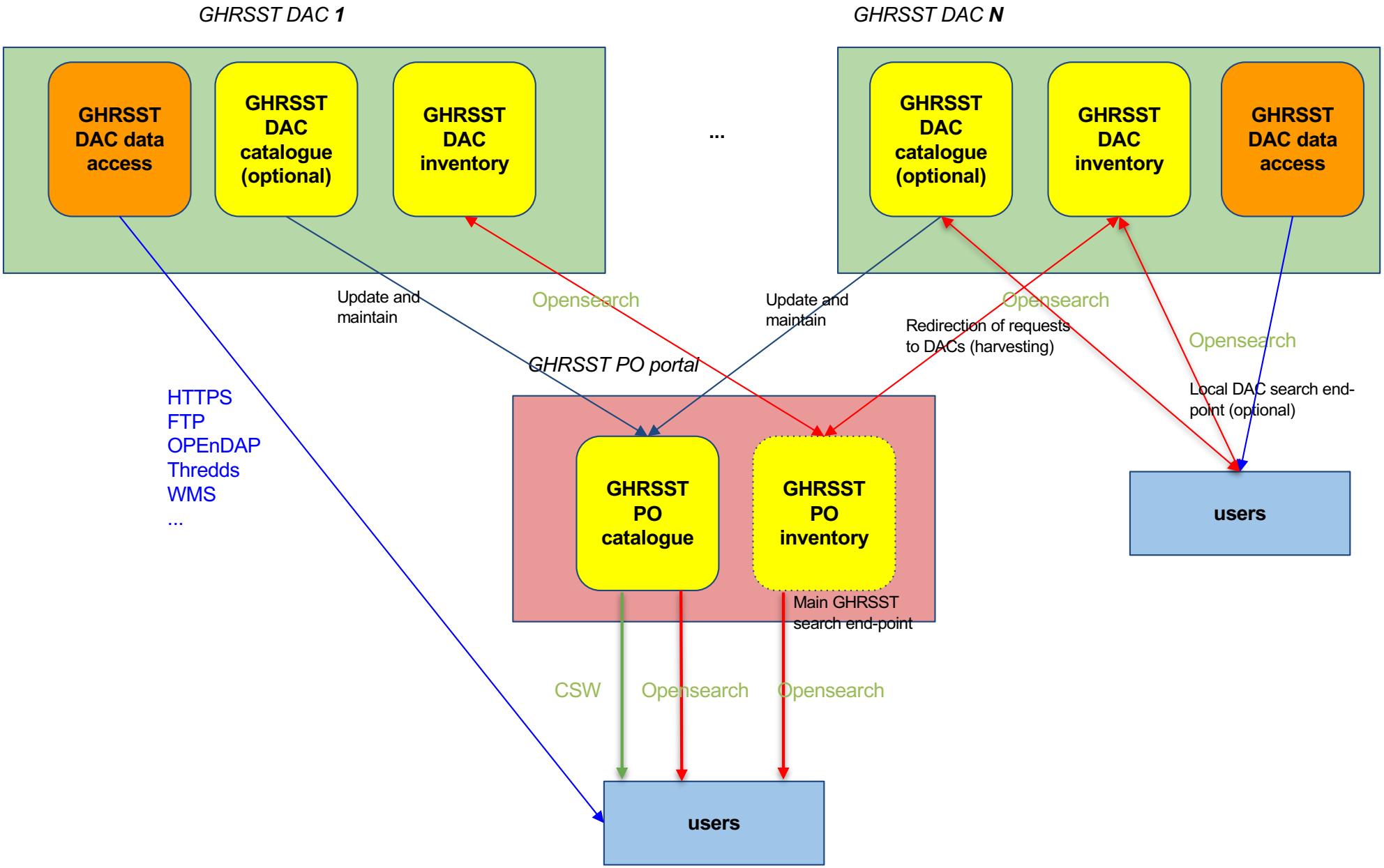

**Figure 1: The initial (2006 to 2019) R/G TS framework. Data and metadata from data producers (RDACs) flow first to a GDAC (such as JPL GDAC or the less comprehensive EU GDAC). After 30 days this flow is also ingested by the NOAA LTSRF. The JPL GDAC has the most comprehensive metadata catalogue. LTSRF has the most comprehensive data archive. All GHRSST metadata are also ingested by the NASA CEOS CWIC repository.**

# GHRSST Distribution (Evolved)

A new system, with decentralisation of data ingestion and distribution, was designed in 2022 to prepare GHRSST for future growth and facilitate the integration of new data producers (Data producers, GDPs and Distributing centres, DACs).

https://www.ghrsst.org/about-ghrsst/task-teams/task-team-on-evolution-of-the-regional-global-task-sharing-r-g-ts-tt/

# R/G TS evolved data discovery, search and access system



# Core design and operations principles

- A public catalog (website) accessible with established open protocols for discovery all GHRSST data products regardless where they reside and are managed
  - Data producers (and DACs)  maintain dataset (e.g., collection, product)  level metadata to ISO 19115-3 specifications in the central catalog
  - Data Assembly Centers maintain a granule level Opensearch search API for granule discovery as part of the dataset level metadata record
  - A harmonized and federated Opensearch service has been implemented with connectors for each flavour of end-point
- Consumers search the catalog for datasets and query for federated Opensearch granule endpoints (HTTPS, FTP, OPeNDAP, S3 etc.) if desired

# GHRSST Dataset metadata profile (½)

| TITLE | Definition | Use | Content example |
|---|---|---|---|
| **Description** | | | |
| Title | Title of the described resource | 1 | *Brazil/Tropical Atlantic High-Resolution Sea Surface Temperature Gridded Level 4 Daily Analysis* |
| Abstract | Abstract of the described resource | 1 | *This L4 SST product is produced at ultra-high resolution (UHR) on a 0.02 x 0.02 degree grid...* |
| Project | Project name(s) linked to the resource | 0:n | *GHRSST* |
| Collection ID | Resource identifier (linked to producer for exemple) | 0:1 | *CER-SST-BRA-1D-002-ODY-MGD* |
| DOI | Digital Object Identifier | 0:1 | *10.17882/52804* |
| Instrument | Instrument name | 1:n | *AVHRR* |
| Platform | Platform name | 1:n | *METOP-A* |
| Level | Production level | 1 | *L4* |
| Acquisition pattern | Acquisition type | 1 | *composite* |
| Compositing | Composition method if any | 1 | *Optimal interpolation* |
| Latency | Delay from acquisition to distribution | 0:1 | *less than 24 hours* |
| Begin date | Date of first observation available | 1 | *08/29/2010* |
| End date | Date of last observation available (empty if still in production) | 0:1 | |
| Temporal Resolution | Time between two observations | 1 | *1 day* |
| Spatial Resolution | Spatial resolution | 1 | *0.02 degree* |
| Projection | Geographic projection | 1 | *ETRS89 (EPSG:4258) - Equirectangular projection* |
| Geographic area | Geographic area covered with the observation | 1 | *Tropical Atlantic* |
| Geographic bounding box | Bounding box of the geographic area | 1 | *westBoundLongitude -75.00*<br>*EastBoundLongitude> -15.00*<br>*southBoundLatitude -25.00* |
| | | | *southBoundLatitude 25.00* |
| Main instrumental or geophysical parameters | Keyword from an internal thesaurus | 0:n | *Ocean Temperature* |
| Keywords (GCMD) | Keyword from GCMD thesaurus | 0:n | */Ocean Temperature/Sea Surface Temperature* |
| ***Contacts*** | | | |
| Point of contact (2 fields) | Information about resource point of contact(s) | 0:n | Institution /Name / e-mail / URL |
| | | | *Helpdesk Cersat* / *cersat@ifremer.fr* / *https://cersat.ifremer.fr* |
| Principal Investigator (2 fields) | Information about the PI | 0:n | Institution / Name / e-mail / URL |
| | | | *Emmanuelle Autret* / *eautret@ifremer.fr* / *https://www.ifremer.fr/lops/* |
| Producer (2 fields) | Who is in charge of producing the data | 0:n | Institution / Name / e-mail / URL |
| | | | *Ifremer/CERSAT* / *cersat@ifremer.fr* / *https://cersat.ifremer.fr* |
| Distributor (2 fields) | Who is in charge of distributing the data | 0:n | Institution / Name / e-mail / URL |
| | | | *Ifremer/CERSAT* / *cersat@ifremer.fr* / *https://cersat.ifremer.fr* |
| Funder (2 fields) | Who is the funder | 0:n | Institution / Name / e-mail / URL |
| | | | *ESA* / / *https://esa.int* |
| ***Access and Usage*** | | | |
| Helpdesk (2 fields) | Information about the Helpdesk | 0:1 | Institution /Name / e-mail |
| | | | *HelpDesk CERSAT* / *cersat@ifremer.fr* |
| Access policy | Access policy from a list of restrictions | 1 | *OtherRestrictions* |
| Usage policy | Usage policy | 1 | |
| Required citation | Required citation | 1 | *These data are produced for ESA/Medspiration project and were obtained from the Centre de Recherche et d'Exploitation Satellitaire (CERSAT), at IFREMER, Plouzané (France)* |
| Distribution Format | Name of the format | 0:1 | *Netcdf* |

# GHRSST Dataset metadata profile (2/2)

| | | | | | |
|---|---|---|---|---|---|
| Format Version | Version of the format | 0:1 | *3* | | |
| Format Convention | Convention name (from a list) | 0:n | *CF-1.4* | | |
| Format Amendment | Amendment number | 0:n | | | |
| ***History*** | | | | | |
| Status | Status of the dataset (active/completed) | 1 | *active* | | |
| Updates (2fields) | Description /date of an update | 0:n | Information | Date | |
| | | | | | |
| Issues (3 fields) | Description / temporal interval of an issue | 0:n | Information | Start Date | End Date |
| | | | *No acquisition. calibration manoeuvre* | *2021/01/12 06:23UTM* | *2021/01/17 08:45UTM* |
| Versions (3 fields) | Date description and number of a new version | 0:n | Version Information | Version Number | Date |
| | | | *New sensors addedd.* | *V5.4* | *2020/01/03* |
| ***Discover the resource*** | | | | | |
| Overview | Preview | 0:n | *png, jpeg file* | | |
| ***Direct download*** | | | | | |
| | | | | | |
| FTP | FTP information | 0:1 | eftp.ifremer.fr/osisaf/data/amsr2 | | |
| HTTPS | HTTPS information | 0:1 | https://ifremer.fr/data/osisaf/amsr2 | | |
| THREDDS | THREDDS information | 0:1 | tds0.ifremer.fr/osisaf/amsr2 | | |
| Cloud | Public cloud access information | 0:n | | | |
| **Service** | | | | | |
| WMS | WMS server link | 0:1 | wms.ifremer/fr/osisaf/ams2 | | |
| OPeNDAP | OPeNDAP link | 0:1 | opendap.ifremer.fr/osisaf/amsr2 | | |
| **Documents** | | | | | |
| User guide | Link(s) to user guide(s) | 0:n | | | |
| Processing and validation | Link(s) to Processing and validation document(s) | 0:n | | | |
| Other document(s) | Link(s) to other document(s) | 0:n | | | |

# Implementation of Central catalogue

- Initial work as Pilot Project circa 2021-2023
- Profile defined from GHRSST L2P/L3/L4 products based on ISO 19115-3, based on different catalogue examples (e.g., PO.DAAC, NOAA, Ifremer/CERSAT)
- GHRSST central catalogue populated from existing catalogues (e.g., NASA CMR) whenever possible (tested in Pilot Project)
  - Customized ingest/import scripts
  - Mapping GHRSST Dataset Id (defined by producer) to Opensearch queries
  - Harmonize the inputs to have the same specification & content: most relevant and most common metadata extracted to defined the GHRSST common metadata model
  - Included cloud access endpoints if available
- Currently undergoing review and editing

GHRSST
GROUP FOR HIGH RESOLUTION SEA SURFACE TEMPERATURE

Group for High Resolution Sea Surface Temperature

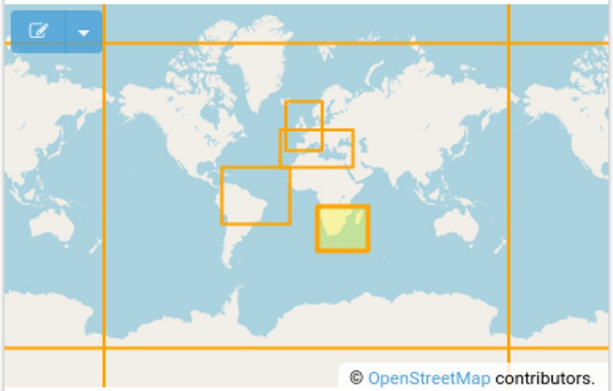
Search...

HOME | QUICK START GUIDE | LATEST SST MAP | PROJECT OFFICE | GHRSST DATA & SERVICES | RESOURCES | OUTREACH

# GHRSST CATALOGUE

SIGN IN

Search ...



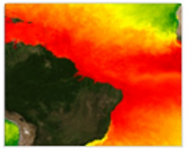

- Platform
- Sensor
- Level
- Latency
- Spatial Resolution
- Producer
- Distributor
- Geographic area

Reset filters

Results 1 to 6 on 6 : 30 by page

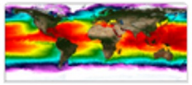
Sort by : Title

**ODYSSEA Brazil/Tropical Atlantic High-Resolution Sea Surface Temperature Gridded Level 4 Daily ...**



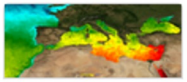
Project(s): Medspiration, GHRSST
Platform(s): METOP-A, METOP-B, NOAA-18, NOAA-19, AQUA, Envisat, MSG, GCOM-W
Instrument(s): AVHRR, AVHRR, AVHRR, AVHRR, MODIS_A, AATSR, SEVIRI, AMSR2
Parameters(s): Ocean Temperature
Temporal resolution: 1 day(s) ...

**ODYSSEA Global Sea Surface Temperature Gridded Level 4 Daily Analysis**



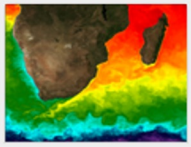
Project(s): MyOcean, GHRSST
Platform(s): METOP-A, METOP-B, NOAA-18, NOAA-19, AQUA, Envisat, TRMM, AQUA, GOES-11, GOES-12, NOAA-17, MSG, GCOM-W
Instrument(s): AVHRR, AVHRR, AVHRR, AVHRR, MODIS_A, AATSR, TMI, AMSRE, Imager, Imager, AVHRR, SEVIRI, AMSR2 ...

**ODYSSEA Mediterranean Sea High-Resolution Sea Surface Temperature Gridded Level 4 Daily Analysi...**



Project(s): GHRSST, Medspiration
Platform(s): METOP-A, METOP-B, NOAA-18, NOAA-19, AQUA, Envisat, NOAA-17, MSG, GCOM-W
Instrument(s): AVHRR, AVHRR, AVHRR, AVHRR, MODIS_A, AATSR, AVHRR, SEVIRI, AMSR2 ...

**ODYSSEA South-Africa/Agulhas Atlantic High-Resolution Sea Surface Temperature Gridded Level 4 Daily...**



Project(s): Medspiration, GHRSST
Platform(s): METOP-A, METOP-B, NOAA-18, NOAA-19, AQUA, Envisat, MSG, GCOM-W
Instrument(s): AVHRR, AVHRR, AVHRR, AVHRR, MODIS_A, AATSR, SEVIRI, AMSR2
Parameters(s): Ocean Temperature
Temporal resolution: 1 day(s) ...

# Responsibilities and Goverance

**GHRSST data producers** are responsible for:

- The compliance of their products to GHRSST GDS specifications. Run the format checker and provided result when submitting for publication in catalogue.
- **Primary editors**: Creating, editing and keeping up to date the description of the datasets they produce for GHRSST.

**GHRSST data assembly centres** are responsible for:

- **Secondary editors**: Creating, editing and keeping up to date the description of the data access services they offer for the datasets they host and distribute for GHRSST => **liaise with primary editors**.

**Review Board** validate the submitted entries and publish them

# Status of the GHRSST Catalogue

The initial content was harvested and populated from PO.DAAC catalogue in 2022

All harvested catalogue entries to be reviewed by producers and DACs and published : opportunity to fix obsolete metadata or add missing metadata

Current list of reviewed & validated datasets include: OSI SAF, EUMETSAT, NOAA, Ifremer and (some) ABoM datasets

Catalogue public since end of June - no public advertising (i.e., the “soft” release) - and work in progress to review and add entries

https://www.ghrsst.org/ghrsst-data-services/for-sst-data-producers/ghrsst-catalogue/

# Opensearch search request example

*https://opensearch-ghrsst.ifremer.fr/granules.raw*?**datasetId**=AVHRR_SST_METOP_B-OSISAF-L2P-v1.0&**startPage**=0&**count**=100&**timeStart**=2020-01-01T00:00:00Z&**timeEnd**=2020-01-01T23:59:59Z&**geoBox**=-180.0,-90.0,180.0,90.0

- No assumption on where the data are served (here both at PO.DAAC and Ifremer)
- Single access point
- Results from both DACs are returned at once
- Returned as XML or JSON

# Demo Tutorial (Notebook)