



National Aeronautics and Space
Administration

Goddard Earth Science Data Information
and Services Center (GES DISC)

README Document for

Reprocessing and Goddard Satellite-based Surface Turbulent Fluxes (GSSTF) Data Set for Global Water and Energy Cycle Research

Goddard Earth Sciences Data and Information Services Center (GES DISC)

<http://disc.gsfc.nasa.gov>

NASA Goddard Space Flight Center

Code 610.2

Greenbelt, MD 20771 USA

Last Revised **10/28/2011**

Prepared By:

Andrey Savtchenko

Name

GES DISC

GSFC Code 610.2

10/28/2011

Date

Name

External Organization

Reviewed By:

Reviewer Name

GES DISC

GSFC Code 610.2

Date

Revision History

<i>Revision Date</i>	<i>Changes</i>	<i>Author</i>

Table of Contents

1. INTRODUCTION	5
1.1. DATASET DESCRIPTION.....	5
1.2. ALGORITHM BACKGROUND	5
1.3. DATA DISCLAIMER.....	6
2. DATA ORGANIZATION	6
2.1. HDF-EOS5 PRODUCTS RESULTING FROM GSSTF2C.....	6
2.2. FILE NAMING CONVENTION	8
2.3. FILE FORMAT AND STRUCTURE.....	10
2.4. KEY SCIENCE DATA FIELDS	10
2.5. SCIENCE AREA	11
3. DATA CONTENTS	12
3.1. DIMENSIONS	12
3.2. GLOBAL ATTRIBUTES.....	12
4. OPTIONS FOR READING THE DATA	13
4.1. COMMAND LINE UTILITIES	13
4.2. TOOLS FOR SIMPLE VISUALIZATION AND FILE CONTENT VIEW.	15
5. DATA SERVICES.....	16
6. MORE INFORMATION.....	16
7. ACKNOWLEDGEMENTS.....	17
REFERENCES.....	17

1. Introduction

This document provides basic information for using Goddard Satellite-based Surface Turbulent Fluxes (a.k.a. GSSTF2c) Data Set products.

The GSSTF2c consists of products generated for the focus on Global Water and Energy Cycle Research. The global water cycle's provision of water to terrestrial storage, reservoirs, and rivers rests upon the global excess of evaporation to precipitation over the oceans. Variations in the magnitude of this ocean evaporation excess will ultimately lead to variations in the amount of freshwater that is transported (by the atmosphere) and precipitated over continental regions. The air-sea fluxes of momentum, radiation, and freshwater (precipitation – evaporation) play a very essential role in a wide variety of atmospheric and oceanic problems (e.g., oceanic evaporation contributes to the net fresh water input to the oceans and drives the upper ocean density structure and consequently the circulation of the oceans). Information on these fluxes is crucial in understanding the interactions between the atmosphere and oceans, global energy, and water cycle variability, and in improving model simulations of climate variations.

1.1. Dataset Description

In contrast to previous versions, the GSSTF Version 2c contains only one set of combined data, "Set1".

The GSSTF, Version 2c, daily fluxes have first been produced for each individual available SSM/I (Special Sensor Microwave Imager) satellite tapes (e.g., F08, F10, F11, F13, F14 and F15). Then, the Combined daily fluxes are produced by averaging (equally weighted) over available flux data/files from various satellites. These Combined daily flux data are considered as the "final" GSSTF, Version 2c, and are stored in this HDF-EOS5 collection.

1.2. Algorithm Background

The bulk flux model (based on the surface layer similarity theory, Chou, 1993) used for performing the newly produced/future production (GSSTF2c/GSSTF3) is essentially the same as that for GSSTF2 (Chou et al., 2003) and GSSTF2b (Shie et al. 2010). Similar to GSSTF2 and GSSTF2b, GSSTF2c requires the same methodology and same kinds of input data such as the surface/10-m wind speeds (U), total precipitable water (W), bottom-layer (500 m) precipitable water (WB), SST, 2-m air temperature (T_a), and sea level pressure (SLP).

The daily 1^o GSSTF2b turbulent fluxes were derived from the SSM/I surface winds and surface air humidity, as well as the 2-m air and sea surface temperatures of the NCEP-NCAR reanalysis, using a bulk aerodynamic algorithm based on the surface layer similarity theory. Transfer coefficients, which reflect

efficiency of the vertical transportation of momentum, heat, and moisture flux, are a non-linear function of the vertical gradient in wind speed, temperature, and water vapor near the surface, and hence are affected by the stability of the surface air.

The coefficients increase with decreasing wind when the 10-m winds are less than 3 m s^{-1} : The GSSTF2 transfer coefficients (for 10-m winds up to $< 18 \text{ m s}^{-1}$) are in close agreement with those of Zeng et al. (1998), Fairall et al. (1996), and Renfrew et al. (2002).

1.3. Data Disclaimer

Data should be used with care and proper citations, i.e., properly indicating your applications with, e.g., "using the combined 2001 data file of Version 2c" or "using the 2001 F13 data file".

2. Data Organization

2.1. HDF-EOS5 products resulting from GSSTF2c

The original GSSTF2c collections are in binary format, and their organization is described in:

<ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxda2c/Shie&Hilburn-2011.pdf>
<ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxda2c/README-Data>
<ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxda2c/README-General>
<ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxda2c/README-Monthly-mean>
<ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxda2c/README-Climatology>

The MEaSUREs project at GES DISC, however, converted all data into HDF-EOS5 format and reorganized it into separate daily and monthly files, in a manner consistent with the on-going Earth Observing System (EOS) missions such as Aura, Aqua, and Terra. The monthly, seasonal and yearly climatologies are also in HDF-EOS5 format, in separate files. I.e. the daily files are now easily searchable and downloadable by data day (section 5).

The essential meaning of HDF-EOS5 is that the data are now in a standard "Grid" format. The eight major variables are grouped into one separate grid, named "SET1". Further, the "minor" variable - total precipitable water - is grouped together with the eight major variables in the corresponding SET1. Thus the grid contains 9 variables. This organization is identical for all daily and monthly files, apart from the model reanalysis. The "common", NCEP/DOE Reanalysis II, data are stored in separate files with one Grid containing the four "common" variables. The climatologies (monthly, seasonal, and yearly) also contain one SET1 grids, but in addition are also containing the four NCEP variables in a separate grid.

The “individual tapes” representing the individual SSM/I satellites are stored in separate collections and daily files, one day per file. They contain one grid that is named on the satellite name (F08, F10, F11, F13, F14, and F15). The grid contains the nine variables (8 major + 1 minor) that go into the computation of the final nine “combined” retrieved variables.

All data within these HDF-EOS5 are easily recognizable as 360x180 arrays, representing global map, at 1x1 deg grid cell size. The endian-ess of the remote storage and users’ local machines become irrelevant. The concept of headers and offsets, typical for the binary format, disappears and all that matters are Grid and data field names that are very easy to list out using command line utilities. Examples are listed in section 4.1.

This data organization results in 13 data types, with Short Names listed below. The Short Name is the first string in all filenames, and it is also stored inside the files as a global file attribute “ShortName”.

A. Daily:

There are total of Eight (8) daily data types.

GSSTF
GSSTF_NCEP
GSSTF_F08
GSSTF_F10
GSSTF_F11
GSSTF_F13
GSSTF_F14
GSSTF_F15

1) GSSTF

It has 1 grid, "SET1". Every grid has 9 parameters, the 8 "major" plus the "minor" total precipitable water.

2) GSSTF_NCEP

It has one grid, "NCEP", with 4 parameters. The original were the "common" binary files, that is the NCEP/DOE Reanalysis II.

3)GSSTF_Fxx

These are the individual satellites, where Fxx is one of the following (F08, F10, F11, F13, F14, and F15). The HDF-EOS5 files have only one grid which takes the name of the individual satellite. Every grid has 9 parameters, the 8 "major" plus the "minor" total precipitable water.

B. Monthly:

There are two Monthly data types in he5.

GSSTFM

GSSTFM_NCEP

1) GSSTFM

It has four (2) grids: SET1 and SET1_INT, where "INT" stands for interpolated. Each of these has 9 (nine) parameters: the 8 "major" plus the "minor" total precipitable water.

2)GSSTFM_NCEP

This is the "common" monthly NCEP/DOE Reanalysis II. It has one grid with four parameters.

C. Climatology

There are three (3) climatological data types: Monthly, Seasonal, and Yearly in HDF-EOS5:

GSSTFMC

GSSTFSC

GSSTFYC

The climatologies (monthly, seasonal, and yearly) also contain the SET1 grid, but in addition are also containing the four NCEP/DOE Reanalysis II variables in a separate grid, "NCEP".

Summary of All Data Types:

GSSTF

GSSTF_NCEP

GSSTF_F08

GSSTF_F10

GSSTF_F11

GSSTF_F13

GSSTF_F14

GSSTF_F15

GSSTFM

GSSTFM_NCEP

GSSTFMC

GSSTFSC

GSSTFYC

2.2. File Naming Convention

The file naming convention for the non-climatological HDF-EOS5 files produced at GES DISC for the GSSTF2c project is as follows:

ShortName.vv.yyyy.mm.dd.he5

Where:

- ShortName = one of the following Data Types:

GSSTF
GSSTF_NCEP
GSSTF_F08
GSSTF_F10
GSSTF_F11
GSSTF_F13
GSSTF_F14
GSSTF_F15
GSSTFM
GSSTFM_NCEP
GSSTFMC
GSSTFSC
GSSTFYC

- vv = 2c for this release
- yyyy = data year
- mm= data month
- dd = start date for the data
- he5 = commonly accepted extension for HDF-EOS5 files.

Filename example for the daily “combined” turbulent fluxes retrieval, for November 1, 2000:

GSSTF.2c.2000.11.01.he5

Climatologies have slightly different file names that reflect the Month (for monthlies), and the range of Months (for seasonal), and the range of years for which the climatology is built. The yearly climatology has one file only.

Example file name for Monthly climatology for November:

GSSTFMC.2c.Nov.1988_2008.he5

Example file name for seasonal climatology for September-November:

GSSTFSC.2c.Sep_Nov.1988_2008.he5

2.3. File Format and Structure

GSSTF2c files are in HDF-EOS5 format, which is an extension of the Hierarchical Data Format Version 5 (HDF-5), developed at the National Center for Supercomputing Applications <http://www.hdfgroup.org/>. This extension facilitates the creation of a standard Grid data structures that are relevant to this project. All data arrays represent global map of size 360x180 (columns x rows) grid cells.

Thanks to the HDF-EOS5 format, the **path** to the data fields inside the file is always standardized. For the Grid format, it is always of the form:

`/HDFEOS/GRIDS/<grid_name>/Data Fields/<data_field_short_name>`.

For example, the sensible heat flux in all relevant GSSTF2c files in HDF-EOS5 will be found under this path:

`/HDFEOS/GRIDS/SET1/Data Fields/H`

Thus, users need to know only the `grid_name` and the `data_field_short_name`. The rest of the path is standard.

2.4. Key Science Data Fields

In the tables below, the `data_field_short_name` is the one that is used in the HDF-EOS5 path inside the files to uniquely identify a data field. The long name is an attribute to the particular data field that serves as a short description of the field.

Each **SET1** grid (including interpolated, SET1_INT) contains the following 9 science data fields:

data_field_short_name	Data field long_name (units)
"E"	'latent heat flux' (W/m**2)
"STu"	'zonal wind stress' (N/m**2)
"STv"	'meridional wind stress' (N/m**2)
"H"	'sensible heat flux' (W/m**2)
"Qair"	'surface air (~10-m) specific humidity' (g/kg)
"WB"	'lowest 500-m precipitable water' (g/cm**2)
"U"	'10-m wind speed' (m/s)
"DQ"	'sea-air humidity difference' (g/kg)
"Tot_Precip_Water"	'total precipitable water' (g/cm**2)

Each NCEP Grid contains the following 4 data fields:

Data field short name	Data field long_name (units)
"SST"	'sea surface skin temperature' (C)
"Psea_level"	'sea level pressure' (hPa)
"Tair_2m"	'2m air temperature' (C)
"Qsat"	'sea surface saturation humidity' (g/kg)

2.5. Science Area

The global water cycle's provision of water to terrestrial storage, reservoirs, and rivers rests upon the global excess of evaporation to precipitation over the oceans. Variations in the magnitude of this ocean evaporation excess will ultimately lead to variations in the amount of freshwater that is transported (by the atmosphere) and precipitated over continental regions. The air-sea fluxes of momentum, radiation, and freshwater (precipitation – evaporation) play a very essential role in a wide variety of atmospheric and oceanic problems (e.g., oceanic evaporation contributes to the net fresh water input to the oceans and drives the upper ocean density structure and consequently the circulation of the oceans).

Information on these fluxes is crucial in understanding the interactions between the atmosphere and oceans, global energy, and water cycle variability, and in improving model simulations of climate variations.

3. Data Contents

3.1. Dimensions

All data fields in the GSSTF2c data types represent global maps of size 360x180 (columns x rows) grid cells. This corresponds to 1x1 degree cylindrical map representation. There only two dimension names in the grid structure, Xdim and Ydim, that address the longitude, and latitude, respectively.

3.2. Global Attributes

In addition to SDS arrays containing variables and dimension scales, global metadata is also stored in the files. Some metadata are required by standard conventions, some are present to meet data provenance requirements and others as a convenience to users of **GSSTF2c** products. Under the HDF-EOS5 standards, the global (file) attributes are grouped under the following path:

/HDFEOS/ADDITIONAL/FILE_ATTRIBUTES

Global Attribute	Type	Description
BeginDate	Character string	Begin time of data
EndDate	Character string	End time of data
ShortName	Character string	Short name of the data type (collection) from section 2.1
LongName	Character string	Long name of the collection
CollectionDescription	Character string	Description of the collection

Apart from the global (file) attributes, every data field has its own “local” attributes. Presently there are only three local attributes:

_FillValue
long_name
Units

_FillValue=-999 is preserved from the binary originals for the entire GSSTF2c collection. The long_name and the Units were described above and can be found in the tables of sec. 2.3.

4. Options for Reading the Data

4.1. Command Line Utilities

The HDF Group lists a number of HDF5 command line tools on their website:

http://www.hdfgroup.org/products/hdf5_tools/index.html#h5dist

Among them, the one that is the most useful for previewing HDF5 content is **h5dump**. Since the HDF-EOS5 is just a standard-structured HDF5, the entire functionality of h5dump is fully preserved. H5dump is the tool to use to have a very first look at any HDF5-formatted file, even if no documentation describing data is available. The tool by itself does not need extensive documentation, it is very straightforward to use and its all options can be conveniently listed at the command line using the help option “-h”:

```
h5dump -h
```

As of the current naming setup of the data fields, a simple list request “-n” will produce the following output, where full paths to groups and individual data fields (dataset) are obvious:

```
h5dump -n GSSTF.2c.1998.12.19.he5
```

```
HDF5 "GSSTF.2c.1998.12.19.he5" {
FILE_CONTENTS {
group /
group /HDFEOS
group /HDFEOS/ADDITIONAL
group /HDFEOS/ADDITIONAL/FILE_ATTRIBUTES
group /HDFEOS/GRIDS
group /HDFEOS/GRIDS/SET1
group /HDFEOS/GRIDS/SET1/Data Fields
dataset /HDFEOS/GRIDS/SET1/Data Fields/DQ
dataset /HDFEOS/GRIDS/SET1/Data Fields/E
dataset /HDFEOS/GRIDS/SET1/Data Fields/H
dataset /HDFEOS/GRIDS/SET1/Data Fields/Qair
```

```

dataset /HDFEOS/GRIDS/SET1/Data Fields/STu
dataset /HDFEOS/GRIDS/SET1/Data Fields/STv
dataset /HDFEOS/GRIDS/SET1/Data Fields/Tot_Precip_Water
dataset /HDFEOS/GRIDS/SET1/Data Fields/U
dataset /HDFEOS/GRIDS/SET1/Data Fields/WB
group /HDFEOS INFORMATION
dataset /HDFEOS INFORMATION/StructMetadata.0

```

It is easy to distinguish the Set1 group, and the associated data fields that are always under the group "Data Fields". The dataset "StructMetadata.0" is written by the HDF-EOS5 libraries, during the process of the Grid structures creation. It can also be inspected by more advanced users, to validate dimensions and data organization (note, some outputs can be very long, so they are abbreviated here for clarity):

```
h5dump -d "/HDFEOS INFORMATION/StructMetadata.0" GSSTF.2c.1998.12.19.he5
```

...

```

GROUP=GridStructure
  GROUP=GRID_1
    GridName="SET1"
    XDim=360
    YDim=180
    UpperLeftPointMtrs=(-180000000.000000,-90000000.000000)
    LowerRightMtrs=(180000000.000000,90000000.000000)
    Projection=HE5_GCTP_GEO
    GROUP=Dimension
      OBJECT=Dimension_1
        DimensionName="Xdim"
        Size=360
      END_OBJECT=Dimension_1
      OBJECT=Dimension_2
        DimensionName="Ydim"
        Size=180
      END_OBJECT=Dimension_2
    END_GROUP=Dimension
  GROUP=DataField
    OBJECT=DataField_1
      DataFieldName="E"

```

...

The option -d can be used to list numerical datasets, too:

```
h5dump -d "/HDFEOS/GRIDS/SET1/Data Fields/E" GSSTF.2c.1998.12.19.he5
```

...

```
(75,165): 196.857, 191.538, 191.316, 184.728, 184.373, 198.597, 188.133,
```

(75,172): 178.673, 177.775, 181.975, 169.355, 146.068, 124.323, 112.992,
(75,179): 102.341, -999, 84.1494, 72.4783, 73.2529, 77.3465, 76.4261,
(75,186): 80.3385, 84.5448, 76.2291, 58.7598, 26.8386, -999, -999, -999,
...

where numbers in parentheses indicate (row, column).

4.2. Tools for simple visualization and file content view.

Among these, **HDFView** is one of the simplest to use and install on your local desktop. A strength of HDFView is that it is good for both, HDF4 and HDF5, formats and it will open files in both formats seamlessly for the user. Another strength is that file content is presented graphically in a very friendly fashion – all groups are presented as folders where user can easily drill down the file hierarchy, which is especially useful for HDF5. Numerical data sets can be viewed as spreadsheets, and as images (multi-dimensional datasets). Simple data manipulations are possible. For download and more information, follow this link:

<http://www.hdfgroup.org/hdf-java-html/hdfview/index.html>

There are many other free tools for HDF viewing.

Panoply, developed at the Goddard Institute for Space Studies (GISS), is compliant with NetCDF Climate and Forecast (CF) Metadata Convention that is gaining popularity. A strength of the tool is that data can be previewed “remotely” over the network – i.e. user can preview file content of HDF files stored on a remote site, without downloading them. Panoply is available from GISS:

<http://www.giss.nasa.gov/tools/panoply/>

The two major commercial data languages, **IDL** and **MatLAB**, are coming with their libraries that fully support all HDF formats. The HDFEOS group has added simple recipes on their web site in particular for GSSTF data in the HDF-EOS5:

http://hdfEOS.org/zoo/index_openGESDISC_Examples.php

5. Data Services

GES DISC provides basic temporal and advanced (event) searches through its search and download engine, **Mirador**:

<http://mirador.gsfc.nasa.gov/>

Mirador offers various download options that suit users with different preferences and different levels of technical abilities. Users can start from a point where they don't know anything about these particular data, its location, size, format, etc., and quickly find what they need by just providing relevant keywords, like "GSSTF2c", or "flux", or "turbulence", or just "Heat".

As the project progresses, more services will be added that will allow options for subsetting, format conversion, and previews using open source software such as OPeNDAP and Panoply.

6. More Information

All GSSTF2c data types are published to the Global Change Master Directory:

<http://gcmd.nasa.gov>

This is a centralized depository of climate data information that is catalogued by popular keywords, which facilitate data discovery. Since the Directory is visited by a large number of people working on a broad range of research topics, it is an excellent forum to popularize data collections. It also serves as source of data type information for Mirador.

7. Acknowledgements

The creators for GSSTF2c are:

Shie, C.-L., K. Hilburn, L. S. Chiu, R. Adler, I-I Lin, E. Nelkin, and J. Ardizzone, 2011: The Goddard Satellite-Based Surface Turbulent Fluxes Dataset --- Version 2c (GSSTF 2c) [global (grid of $1^\circ \times 1^\circ$) daily air-sea surface fluxes from July 1987 to December 2008], distributed via Goddard Earth Sciences (GES) Data and Information Services Center (DISC), October 2011.

References

Chou, S.-H., 1993: A comparison of airborne eddy correlation and bulk aerodynamic methods for ocean-air turbulent fluxes during cold-air outbreaks. *Bound.-Layer Meteor.*, 64, 75-100.

Chou, S.-H., E. Nelkin, J. Ardizzone, R. M. Atlas, and C.-L. Shie, 2003: Surface turbulent heat and momentum fluxes over global oceans based on the Goddard satellite retrieval, version 2 (GSSTF2). *J. Climate*, 16, 3256-3273.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, 1996: Bulk parameterization of air-sea fluxes for Tropical Ocean Global Atmosphere Coupled Ocean-Atmosphere Response Experiment. *J. Geophys. Res.*, 101(C2), 3747-3764.

Renfrew, I. A., G. W. K. Moore, P. S. Guest, and K. Bumke, 2002: A comparison of surface layer and surface turbulent flux observations over the Labrador Sea with ECMWF analyses and NCEP reanalyses. *J. Phys. Oceanogr.*, 32, 383-400.

Shie, C.-L., L. S. Chiu, R. Adler, I-I Lin, E. Nelkin, and J. Ardizzone, 2010: The Goddard Satellite-Based Surface Turbulent Fluxes Dataset --- Version 2b (GSSTF 2b) [global (grid of $1^\circ \times 1^\circ$) daily air-sea surface fluxes from July 1987 to December 2008] distributed via *Goddard Earth Sciences (GES) Data and Information Services Center (DISC)* in October 2010.

Zeng, X., M. Zhao, and R. E. Dickinson, 1998: Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. *J. Climate*, 11, 2628-2644.

