

# **TOOLS AND DATA SERVICES FROM THE NASA EARTH SATELLITE OBSERVATIONS FOR CLIMATE APPLICATIONS**

**Gilberto A. Vicente and the NASA GES/DISC/DAAC Science Support Team**

NASA Goddard Earth Sciences (GES), Data and Information Services Center (DISC)  
Distributed Active Archive Center (DAAC), Greenbelt, MD, 20071, USA  
vicente@daac.gsfc.nasa.gov

## **ABSTRACT**

The NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) is one of the major Distributed Active Archive Centers (DAAC's) archiving and distributing remote sensing data from the NASA's Earth Observing System (EOS). This virtual data portal provides convenient access to Atmospheric, Oceanic and Land datasets and value added products from various current NASA missions and instruments as well as heritage datasets from AIRS/AMSU/HSB, AVHRR, CZCS, LIMS, MODIS, MSU, OCTS, SeaWiFS, SORCE, TOMS, TOVS, UARS and TRMM. In addition, the GES DISC has developed various value-adding processing services for searching and ordering data products and access to many others NASA DAAC's. The data is available online either via FTP or from archives searchable by type, dataset and parameter. Other services include data tool aids to help users to fully utilize and manipulate DAAC HDF and HDF-EOS data products. Some services allow direct data and product access and manipulation in the GES DISC own servers as the Interactive Online Visualization and Analysis Infrastructure (Giovanni). This user-friendly Web interface to data and products customized to many disciplines contains analysis and data visualization tools, suitable for publication standards all without having to download data. There are also desktop applications that users run on their local machines such as the HDFLook and the Near-line Archive Data Mining (NADM), a web portal that enables registered users to submit and execute data mining algorithm codes on the GES-DISC data pool. The generated mined data products can then be transferred via the network to the user.

## **1. INTRODUCTION**

With the launch in 1999 of the first Earth Observing System (EOS) satellite Terra, the National Aeronautics and Space Administration (NASA) began the first satellite-based observing system to offer integrated measurements of the Earth's processes. It consists of a science component and a data system supporting a coordinated series of polar-orbiting and low-inclination satellites for long-term global observations of the land surface, biosphere, solid Earth, atmosphere, and oceans. The largest scientific data system in the world -- the Earth Observing System Data and Information System -- currently collects environmental measurements from more than 30 satellites, such as NASA's Terra, Aqua, Aura and ICESat.

The NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC) is one of eight NASA Earth Science Enterprise (ESE) Distributed Active Archive Centers (DAAC's) that offers Earth science data, information, and services to research scientists, applications scientists, applications users, and students. The GES DAAC is the home (archive) of ESE's Ocean Color, Hydrology, Atmospheric Chemistry and Dynamics, and Land Biosphere data and information, as well as data and information from other related disciplines. With all these capabilities, just the GES DISC processes more than 1 TB of data a day. Overall archive volume is currently 1.75 PB (PetaByte =  $10^{15}$  Byte). Peak daily numbers are: 3.3 TB for processing, 1.7 TB for archiving (we don't archive some intermediate products) with daily

distribution to the public at peak 1.75 TB. Our record distribution to the public in number of granules is 62000. Total volume is 1.75 PB.

This unprecedented volume creates new never before observed challenges with EOS data usability. Since the EOS data products have been originally optimized for efficient processing not distribution (most data files are too big – up to 1 GB; single product daily volume – tens of GB, with multiple (up to hundreds or even thousands) channels or parameters packed together), the GES DISC has developed various Value-added Services such as the Near-line Archive Data Mining (NADM) system, On-demand and On-the-fly subsetting (FTP-based) approaches and Gridded data on-line visualization and analysis (Giovanni) system.

## **2. SATELLITE AND INSTRUMENTS**

Currently the GES-DISC-DAAC provides online and archived data from 15 NASA satellites and instruments: AIRS/AMSU/HSB, AVHRR, CZCS, LIMS, MODIS, MSU, OCTS, SeaWiFS, SORCE, TOMS, TOVS, UARS and TRMM. A quick description of them and access information follows.

### **2.1 AIRS - Atmospheric Infrared Sounder**

AIRS is a high spectral resolution spectrometer with 2378 bands in the thermal infrared (3.7 - 15.4  $\mu\text{m}$ ) and 4 bands in the visible (0.4 - 1.0  $\mu\text{m}$ ) aboard the second Earth Observing System (EOS) polar-orbiting platform, EOS Aqua. In combination with the Advanced Microwave Sounding Unit (AMSU) and the Humidity Sounder for Brazil (HSB), AIRS constitutes an innovative atmospheric sounding group of visible, infrared, and microwave sensors. AIRS data are generated continuously and global coverage are obtained twice daily (day and night) on a 1:30pm sun synchronous orbit from a 705-km altitude. For processing convenience, the data is divided into 6-minute granules (the smallest unit of data products).

<http://disc.gsfc.nasa.gov/AIRS/intex.shtml>

## 2.2 **AVHRR - Advanced Very High Resolution Radiometer**

The Pathfinder AVHRR Land data sets are global data mapped to an 8 km equal area projection. Parameters produced include reflectances and brightness temperatures derived from the five-channel cross-track scanning AVHRR aboard the NOAA Polar Orbiter 'afternoon' satellites (NOAA-7, -9, -11 and -14), along with a derived Normalized Difference Vegetation Index (NDVI), cloud and quality control indicators, and ancillary data. The data are being produced from the NOAA Global Area Coverage (GAC) Level 1B data spanning a period of more than 20-years (1981-2001). There is a permanent data gap in 1994 between the middle of September to the end of December due to satellite failure. September 30, 2001 is the last day of data processed. NASA/HQ is currently reviewing future 8 km NDVI AVHRR processing strategies.

<http://disc.gsfc.nasa.gov/data/dataset/AVHRR/>

## 2.3 **CZCS - Coastal Zone Color Scanner**

The CZCS data set is a global ocean color data set. In most oceanic regions, the color of the ocean is primarily determined by the abundance of phytoplankton and their associated photosynthetic pigments. As phytoplankton pigment concentration increases, ocean color gradually shifts from blue to green. CZCS utilizes this characteristic of ocean color for the observation of pigment distribution in the oceans. The CZCS was launched on the Nimbus-7 satellite in October 1978. During the 91 months of its operation (October 1978 - June 1986), CZCS acquired nearly 68,000 images, each covering up to 2 million square kilometers of ocean surface.

<http://disc.gsfc.nasa.gov/data/dataset/CZCS/>

## 2.4 **LIMS - Limb Infrared Monitor of the Stratosphere**

The LIMS instrument measured vertical profiles of temperature, geopotential height, and mixing ratios of ozone (O<sub>3</sub>), nitrogen dioxide (NO<sub>2</sub>), water vapor (H<sub>2</sub>O), and nitric acid (HNO<sub>3</sub>). The LIMS instrument was launched on the Nimbus-7 satellite and was operational for about seven months from 25 October 1978 until May 28, 1979. Data is available from the DAAC either electronically via ftp, or they can be ordered on CD-R media as a seven disk set.

<http://disc.gsfc.nasa.gov/data/dataset/LIMS/>

## 2.5 **MODIS – Terra/Aqua - Moderate Resolution Imaging Spectroradiometer**

The MODIS is one of many Earth-observing instruments on board the EOS Terra satellite launched in December 1999 and Aqua launched in May 2002. It provides images of land, ocean and atmosphere in 36 spectral bands (from 0.4µm to 14.5µm) with spatial resolutions of 250m (bands 1-2), 500m (bands 3-7) and 1000m (bands 8-36).

<http://disc.gsfc.nasa.gov/MODIS/index.shtml>

## 2.6 **MSU – Microwave Sounding Unit**

Data from the MSU LIMB93 processing is available in two formats, namely, the Hierarchical Data Format (HDF) and a native format originally used by the data producer. In both cases, separate files are provided for the Lower Tropospheric Temperature (LTT), the Upper Tropospheric Temperature (UTT), the Lower Stratospheric Temperature (LST) and the Oceanic Precipitation (OP). Each HDF file contains a single year of daily grids at 2.5 degree resolution. The native format files contain all 16 years of data in single files (1979-1994), except for precipitation which is separated by year as in the case of the HDF version of the data. The data and read programs are available for both data formats via anonymous FTP.

<http://disc.gsfc.nasa.gov/data/dataset/MSU/>

## 2.7 [OCTS](#) - Ocean Color and Temperature Sensor

The OCTS operated on the Japanese "Midori" remote-sensing satellite from November 1996 to June 1997. The OCTS data set has been processed into the same format as SeaWiFS data files and has similar applications as SeaWiFS data, primarily intended to quantify global ocean bio-optical properties. OCTS data therefore provide information on phytoplankton distribution, sediments, and related quantities for the mission period, which occurred just prior to the beginning of the SeaWiFS mission.

<http://daac.gsfc.nasa.gov/data/dataset/OCTS/>

## 2.8 [SeaWiFS](#) - Sea-viewing Wide Field-of-view Sensor

The SeaWiFS Ocean Color Project provides quantitative data on global ocean bio-optical properties to the Earth science community. Subtle changes in ocean color signify various types and quantities of marine phytoplankton (microscopic marine plants), the knowledge of which has both scientific and practical applications. The SeaWiFS data is available through the DAAC starting November 1996.

The SEAWIFS Land data set contains vegetation products including 4.63 km resolution 16-day composited continental data, and 0.25 degrees monthly global gridded data. The continental data are at the same spatial and temporal resolution as the MODIS composited vegetation indexes to facilitate comparison studies between the two instruments. Parameters produced include reflectances from bands 2 (blue), 4 (blue-green), 6 (red), and 8 (near IR), the Vegetation Index (VI), the Enhanced Vegetation Index (EVI), and the Normalized Difference Vegetation Index (NDVI).

<http://daac.gsfc.nasa.gov/data/dataset/SEAWIFS/>

[http://daac.gsfc.nasa.gov/data/dataset/SEAWIFS\\_LAND/](http://daac.gsfc.nasa.gov/data/dataset/SEAWIFS_LAND/)

## 2.9 [SORCE](#) - Solar Radiation and Climate Experiment

The SORCE mission measures the solar radiation incident at the Earth's atmosphere. Data products are available containing the total solar irradiance, and solar spectral irradiance (1 to 35 nm and 115 to 2000 nm), both at 6 hourly and daily averages. SORCE data is available through the DAAC starting February 2003.

<http://daac.gsfc.nasa.gov/data/dataset/SORCE/>

## 2.10 [SSI](#) - Surface Solar Irradiance

The SSI dataset was derived for the SeaWiFS Project to study the temporal fluctuations of total solar radiation and Photosynthetically Active Radiation (PAR) on time scales relevant to phytoplankton physiology. This data also provides information on global cloud coverage and cloud optical depth useful as input to radiative transfer models, initialization of global climate models and energy budget studies. The SSI data set provides a global view of solar energy incident at the surface for both the full spectrum (.4 - 2.2 microns) and the photosynthetically active region of the spectrum (.4 - .7 microns). The major atmospheric inputs to the model are obtained from the International Satellite Cloud Climatology Project (ISCCP) DX data. The SSI data are mapped to a 1/2 degree rectangular grid at 3 hourly, daily and monthly time resolutions.

<http://daac.gsfc.nasa.gov/data/dataset/SSI/>

## 2.11 [TOMS](#) - Total Ozone Mapping Spectrometer

The TOMS data is the only source of high resolution global information about the total ozone content of the atmosphere. The data set is important in studies involving upper air dynamics and chemistry on both short (inter-annual) and long (decadal) time scales. TOMS is the primary means by which regional (e.g., polar) and global ozone patterns have been monitored on a daily basis since 1978. TOMS instruments have flown on four satellites to date: Nimbus-7 (November 1978 to May 1993), Meteor-3 (August 1991 to December 1994), Earth Probe (July 1996 to present), and ADEOS (September 1996 to June 1997).

<http://daac.gsfc.nasa.gov/data/dataset/TOMS.v7/>

#### 2.12 [TOVS](#) - TIROS Operational Vertical Sounder

The TOVS instrument package provides information on temperature and humidity profiles, total ozone, clouds and radiation on a global scale. The Pathfinder processing of the data products using fixed algorithms provides a wealth of climatological information. TOVS has been carried aboard NOAA polar orbiting weather satellites since 1978, and continues to the present.

<http://daac.gsfc.nasa.gov/data/dataset/TOVS/>

#### 2.13 [UARS](#) – Upper Atmosphere Research Satellite

The goal of the UARS is to understand the chemistry, dynamics, and energy balance above the troposphere, as well as the coupling between these processes and between regions of the atmosphere. The UARS platform provides near global (-80° to +80°), simultaneous, coordinated measurements of atmospheric internal structure (trace constituents, physical dynamics, radiative emission, thermal structure, density) and measurements of the external influences acting upon the upper atmosphere (solar radiation, tropospheric conditions, electric fields).

<http://daac.gsfc.nasa.gov/data/dataset/UARS/>

#### 2.14 [TRMM](#) - Tropical Rainfall Measuring Mission

The TRMM is a joint endeavor between NASA and Japan's National Space Development Agency. It is designed to monitor and study tropical rainfall and the associated release of energy that helps to power the global atmospheric circulation, shaping both global weather and climate. TRMM data and products are available in the GES-DISC-DAAC starting December 1997.

<http://disc.gsfc.nasa.gov/precipitation/instruments.shtml>

### 3. RESEARCH AREAS

The GES-DISC-DAAC structure offers a cross links among the Earth science disciplines, NASA remote sensing instruments, derived data parameters, and the applications. Our areas of expertise include Atmospheric Dynamics, Field Experiments, Interdisciplinary Sciences, Geomorphology, Hydrology, Land Biosphere and Ocean Color. Here are a few examples:

#### 3.1 Precipitation and Hydrology

The Precipitation Data and Information Service Center (PDISC) at the GES-DISC-DAAC is designed to facilitate overall access and use of NASA related precipitation data sets. The site begins with data and products set selection and navigate the user through documentation, tools, science applications, instruments and links that directly contributes to the user's benefit.

The PDISC support the following precipitation data and products:

1. TRMM only data and precipitation products
2. TRMM combined with other satellite products
3. Precipitation ancillary data
4. Precipitation data and products visualization and manipulation tools

<http://disc.gsfc.nasa.gov/precipitation/index.shtml>

#### 3.2 Atmosphere Dynamics

The GES-DISC-DAAC designed two atmospheric data sets portals to facilitate overall access, visualization and use of atmospheric composition and dynamics data sets in carrying out the atmospheric

chemistry studies and in the studies of the dynamics of Polar Regions that play crucial role in the ozone hole formation.

The atmospheric composition data sets includes Aura [HIRDLS MLS](#), and [OMI](#), (the EOS chemistry mission launched July 15, 2004), Upper Atmosphere Research Satellite (UARS) and the Total Ozone Mapping Spectrometer (TOMS) products, as well as data sets from other missions dedicated to studying the chemistry, dynamics and solar interactions of the upper atmosphere. It also has data sets from the Solar Radiation and Climate Experiment (SORCE) (launched January 25, 2003).

The atmospheric dynamics data sets include data from MODIS, AIRS/AMSU, TOVS and data from the Goddard Modelling Assimilation Office (GMAO).

<http://disc.gsfc.nasa.gov/atmosphere/index.shtml>

### **3.3 Ocean Color and Ocean Color Time Series**

Oceans data from the GES-DISC-DAAC has been constantly used for numerous scientific investigations in the oceanic realm. Starting with the CZCS data archive as an unprecedented source of data for biological oceanography, the archive of oceanographic data grew with the data from SeaWiFS, OCTS, and MODIS sea surface temperature, a basic data element for studies in physical oceanography. Data from other instruments, such as TRMM and the AIRS, provide data that can be used to investigate the energetic of ocean-atmosphere interaction.

We address four main oceanographic topics:

1. Does the activity of phytoplankton at the ocean surface influence the formation of clouds? If so, how does this happen, and how might it affect weather and climate?
2. Is there a connection between desert dust storms and harmful or toxic blooms of algae? How strong is this connection?
3. How does the passage of a hurricane cause physical and biological changes at the ocean surface?
4. Is oceanic productivity – the amount of organic carbon produced by phytoplankton – increasing over time, or decreasing? How might changes in oceanic productivity affect fish and fisheries?

<http://disc.gsfc.nasa.gov/oceans/index.shtml>

### **3.4 Radiances and Upper Atmospheres**

The GES-DISC-DAAC has been archiving and distributing Radiance data, and serving science and application users of these data for over 10 years now. The user-focused stewardship of the radiance data from the AIRS, AVHRR, MODIS, SeaWiFS, SORCE, TOMS, TOVS, TRMM, and UARS instruments exemplifies the GES DAAC tradition and experience. Radiance data include raw radiance counts, onboard calibration data, geolocation products, radiometric calibrated and geolocated-calibrated radiance/reflectance.

The GES DAAC effort concentrates in building bridges between multi-sensor data and the effective scientific use of the data, with an emphasis on the heritage of the science products. The intent is to inform users of the existence of this large collection of radiance data suggesting starting points for cross-platform science projects. In addition the GES-DISC-DAAC is now pioneering methods to function as a research partner, providing data, analysis tools, and expert advice to assist and augment the efforts of scientists around the world.

<http://disc.gsfc.nasa.gov/radiance/index.shtml>

## 1. Interdisciplinary Data

The data collections available through the GES-DISC-DAAC are from a variety of sources and have been integrated to help support Interdisciplinary Research. Brief descriptions of the collections are presented below:

### 1.1.1 *CIDC - Climatology Interdisciplinary Data Collection*

The CIDC contains global climatology data at various temporal and spatial resolutions. The data sets in this collection have been acquired from numerous sources and made available by the GES DAAC in an integrated form to support researchers interested in interdisciplinary science. The CIDC consists of monthly mean data from the fields of meteorology, atmospheric, land surface, oceanic, cryospheric, biospheric, solar and remote sensing sciences. Over 70 physical parameters have been selected from over 24 important climate data sets. Most parameters are presented as IEEE Float 4-byte words on a 1x1 degree grid.

The CIDC data are available on CD-ROM and by FTP. Included on the CD-ROMs are over 70 monthly mean global climate parameters, three heuristic scenarios dealing with monsoons, El Nino events and global warming and the Grid Analysis and Display System.

[http://disc.gsfc.nasa.gov/interdisc/cidc/cidc\\_overview.shtml](http://disc.gsfc.nasa.gov/interdisc/cidc/cidc_overview.shtml)

### 1.1.2 *SIESIP - Seasonal to Inter-annual Earth Science Information Partner*

The SIESIP Pentad & Decade Data Collection is an expansion of the CIDC. This data collection consists of climate parameters at a 5-day and 10-day mean temporal resolution. The creation of this dataset is one of the GES DAAC's contributions to the SIESIP project.

[http://disc.gsfc.nasa.gov/interdisc/siesip\\_pentads.shtml](http://disc.gsfc.nasa.gov/interdisc/siesip_pentads.shtml)

### 1.1.3 *GEDEX - Greenhouse Effect Detection Experiment*

The GEDEX data collection contains over 60 data sets with parameters relevant to greenhouse effect research (surface and upper air temperature, solar irradiances, radiation budget, clouds, and greenhouse gases). Many data sets are available for a 10 year period spanning the 1980's. Depending upon the data set, the coverage is global, regional or local.

<http://disc.gsfc.nasa.gov/interdisc/gedex.shtml>

### 1.1.4 *GMAS – Global Modeling and Assimilation Office*

The GEOS-1 Multiyear Assimilated dataset produced by the GMAS at the GSFC by using an atmospheric general circulation mode I (GCM) provides dynamically consistent, global and gridded atmospheric data for use in climate research. With a resolution of 2.0x2.5 latitude longitude by 20 levels at 6 and 3 hour intervals, the dataset contains various parameters such as upper air heights, winds, temperature, and moisture as well as numerous derived quantities.

<http://daac.gsfc.nasa.gov/data/dataset/DAO/>

## 4. TOOLS FOR DATA/PRODUCTS ACCESS AND MANIPULATION

The GES-DISC-DAAC provides a variety of methods for searching and ordering data products that includes the search and order data Web-Based Hierarchical Ordering Mechanism (WHOM) Interface, the EOS Data Gateway (EDG) and access to many others NASA DAAC's. The data is available online either via FTP or from archive with special services for data search by types, datasets and parameters. Real-time search and order for all data is also available from the NASA's ESE.

The GES-DISC-DAAC also provides many data tools as aids to help users fully utilize and manipulate DAAC HDF and HDF-EOS data products. These include data services that manipulate data on the DAAC servers as the Interactive Online Visualization and Analysis Infrastructure (Giovanni) and desktop applications that users run on their local machines such as the HDFLook. The later is a multifunctional GES DAAC data processing and visualization tool for MODIS (Radiometric and Geolocation, Atmosphere, Ocean and Land data) and AIRS L1B data. Giovanni on the other hand provides a Web interface to data, analysis, and data visualization tools suitable for publication standards, all without having to download data.

Other data/product manipulation tools supported by the GES-DISC-DAAC include the Open Source Project for a Network Data Access Protocol (OPeNDAP), formerly Distributed Oceanographic Data System users and the Near-line Archive Data Mining (NADM). NADM is a web portal that enables registered users to submit and execute data mining algorithm codes on the EOS-DAAC data pool. The generated mined data products can then be transferred via the network to the user.

<http://disc.gsfc.nasa.gov/tools/index.shtml>

## **5. CONCLUSION**

This paper is a quick overview of the satellite data, products, tools and services provided by the NASA GES-DISC. We describe 15 satellite and instruments which data and products are distributed through our services and 5 research areas of interest: Precipitation and Hydrology, Atmosphere Dynamics, Ocean Color and Ocean Color Time Series, Radiances and Upper Atmosphere and Interdisciplinary Data. We also describe various tools and services designed to assist in the use and distribution of the immense data/product archived from the NASA's EOS and a wide spectrum of options intended to shift the processing and data management burden from the user to the GES-DISC.

Our main goal is to make more cost-effective to users the utilization of the GES-DISC existing data management and processing infrastructure for their own satellite data and product needs. This allows scientists to concentrate on science, while the GES-DISC handles the data management and data processing at a lower cost. That is possible because the GES DISC has the capability, infrastructure and personnel that supports a wide spectrum of options, from simple data support to sophisticated online analysis tools.

For more information visit the NASA GES-DISC-DAAC main site located at <http://disc.gsfc.nasa.gov/>.