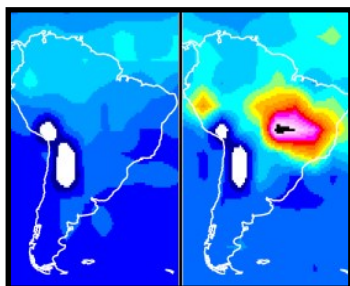


Mission Critical: A Changing Climate

HDF software is still the software of choice—The power and flexibility to spur new discovery.



Remote sensors track Carbon Monoxide from aboard the NASA satellite down to its source

The year is 1999; the space agency that had landed a spacecraft on Mars and deployed a telescope to observe the far-away galaxies was inaugurating what was arguably its most critical mission.

[NASA](#) was launching [Terra](#), the first of an eventual convoy of satellites that together would gather a comprehensive picture of Earth’s global environment. Terra was the size of a small school bus and was originally equipped with five different scientific sensor instruments. Its orbit was coordinated with those of other satellites then circling Earth’s poles so that

when data from each were combined, they yielded a complete image of the globe. With each orbit, the picture grew richer.

Another constellation of satellites would be added between 2002 and 2004. They measured different biological and physical processes and revolutionized science’s understanding of the intricate connections between Earth’s atmosphere, land, snow and ice, ocean, and energy balance. In time, all the combined data would reveal, among other things, the climate trajectory of our planet.

When NASA was searching for the data format with the power and flexibility to assemble those portraits, they chose HDF. “Earth as a system is interdisciplinary. That’s why the data format needed to be flexible,” said Martha Maiden, who was the program manager at NASA Goddard when the space agency turned its attention to Earth. “It wasn’t a format but a system of formats,” she said. “It could handle any kind of data and scale.”

When NASA launched the flagship satellite *Terra*, they chose HDF as the data format system for their critical, new mission: collecting the first comprehensive pictures of our planet’s changing climate.

“These were the largest datasets at the time,” said NASA’s Hampapuram Ramapriyan. “They had volume, variety, and velocity before the Big Data people invented the 3 V’s.”

In the span of five years, the field of earth science had gone from *data poor* to *data rich*. Back then, nearly a terabyte of data per day was transmitted from the [Earth Observing System](#), as the mission was called, to NASA’s data archive centers and then re-distributed to thousands of scientists around the world.

In [fiscal year 2015](#), the system added **16 terabytes per day to 12 data archive centers**. Now totaling over 14 petabytes, the data encompasses **9,400 different data products**. In 2015, **32 terabytes per day** were redistributed to more than **2.4 million end users** worldwide. The flotilla of satellites monitoring our planet has grown to 18.

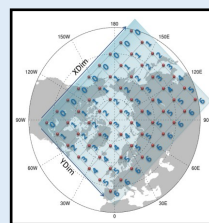
Today, HDF is still the data format system of choice. “Our challenge is variety. With HDF we can use one library and provide it to people to read and write a lot of our products,” said Ramapriyan. “We are behind the software. It’s not something that’s going away any time soon.” [See NASA’s Worldview.](#)

HDF-EOS GRIDS, SWATHS AND POINTS

HDF can represent data in countless ways. EOS created a standard way to organize objects, called “[HDF-EOS](#).” **HDF-EOS expresses how climate scientists think about their data.**

HDF-EOS defines “geolocation data types,” allowing the scientist to query a file’s contents by earth coordinates and time. An example is a “grid data type.” An HDF-EOS grid represents locations on a rectangular grid, based on an earth projection, such as UTM. Other types include “swath” (scan lines from a satellite instrument), and

“point” (irregularly spaced locations, such as buoys).



A grid

EOS instruments collect a variety of measurements into many different “products.” EOS

products include the original data from satellites, and products derived by various analyses. In all, **EOS has more than 9,400 different data products.**

HDF-EOS files also include metadata, essential for scientists to find, use, and understand the data.



A swath

EOS metadata standards assure consistency and completeness, a critical need for software and algorithms to be able to operate effectively on the data.

Software is provided to simplify reading and writing HDF-EOS, and to guarantee consistency. C and Fortran libraries are available, and applications such as MATLAB and IDL have been adapted to understand HDF-EOS files.

The EOS project is a good example of a community or application domain standardizing on a data format. Similar HDF-based standards exist in areas as varied as Bathymetry, Oil and Gas, and medical imagery.

[HDF-EOS website](#)