NASA EOSDIS Earth Observing System Data and Information System

Published on *EOSDIS - Earth Data Website* (https://earthdata.nasa.gov)

Home > User Resources > Sensing Our Planet > The NOAA/NASA Pathfinder Program

The NOAA/NASA Pathfinder Program [1]

by Annette Varani
Published in 1998

In the decade since the EOS Program Office at NASA Headquarters joined with the National Oceanic and Atmospheric Administration (NOAA) to fund the Pathfinder Program, the business of data management has been transformed -- and not solely by computational progress.

> In addition to a rich data yield, the Pathfinder Program's value in pioneering data management issues will be measured in the EOS era ahead.

The Pathfinder Program's primary scientific aims were to make sure that key remote sensing data sets significant to global change research were scientifically validated, consistently processed and made readily available to researchers at minimal cost. That those goals are realized is evidenced by the Pathfinder data sets now found in abundance at the DAACs and other data centers. Yet, in addition to a rich data and information yield, the Program's extended value in pioneering data management issues is still to be appreciated. The first EOS satellite, with its unprecedented observational capacities, will test standards evolved by Pathfinder teams for processing and handling large data sets.

"Through the Pathfinder process, we began to realize just what 'large data sets' meant. Everybody talked gigabytes and terabytes, but not until they started having to process them did the reality of storage, processing times, and multiple parallel processing come into play," said James Dodge, coordinator of interdisciplinary science at NASA Headquarters' research division in the office of Earth Science.

"A terabyte is a million megabytes. The EOS platforms are going to generate at least a terabyte a day of data, and probably another terabyte in products. It's a lot of data to handle on a routine basis," he said.

The Pathfinder program identified four long time-series data sets from existing archives for reprocessing: the Advanced Very High Resolution Radiometer (AVHRR) data, the TIROS Operation Vertical Sounder (TOVS) data, Geostationary Operational Environmental Satellite (GOES) data, and Special Sensor Microwave/Imager (SSM/I) data. The following snapshot of the first project processed by the Pathfinder Program -- development of the AVHRR Land and Polar data sets -- foreshadows challenges in scale and logistics that will continue to be hurdles of EOS-era data streams.
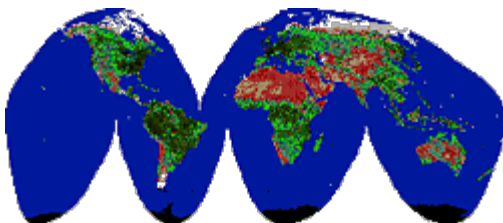
AVHRR vegetation index for May 11-20, 1998 (Image courtesy of Goddard Space Flight Center)

Initially, NASA supported development of four types of land products and a sea surface temperature (SST) product from the AVHRR. The instrument's high resolution presented tremendous data volume issues. The Global Land 1 km AVHRR data developed at EROS Data Center under the leadership of Principal Investigator Jeff Eidenshink, remote sensing scientist with the USGS, proposed development of an index showing the vigor and density of green vegetation around the world. These Pathfinder data would meet an urgent need for information to help monitor Northern Hemisphere and tropical forest resources, and was called for to inform new algorithm development for a next-generation instrument planned for launch on the Terra spacecraft in 1999. But

of immediate value, processing the AVHRR Land Pathfinder resulted in an unparalleled collaboration between AVHRR data collectors around the world.

"We had to engineer a system that could process about 10 gigabytes of data a day," Eidenshink said, "but the problem was that there was no single ground station receiving the data." To address acquisition, Eidenshink and his colleagues formed an international network of nearly 40 ground receiving stations -- 23 worldwide in addition to NOAA local area coverage (LAC) recorders -- that teamed to send data to EROS every two to three months, Eidenshink said.

The EDC DAAC became the central repository for all these data, with attendant voluminous ingest and cataloging duties. The AVHRR product generation team secured the satellite record by sending a complete copy of all data to the European Space Agency, in effect creating a duplicate archive. Nearly 5,000 individual images -- approximately 3.2 terabytes in all -- were acquired during the first 30 months of the project.

Once the acquisition phase of the project was underway, the Pathfinder team turned to product generation. Their focus was formation of a global 10-day vegetation index. Vegetation indices are based on the reflectivity of plants imaged by the sensor. Plant pigments absorb visible light to start photosynthesis, appearing dark in the blue and red regions of the electromagnetic spectrum, showing energy take-up, whereas plants reflect and look bright in the near-infrared region, unlike bare ground or water which are very dark in the infrared spectral regions. By mathematically comparing the differences between energy absorbed (indicated in visible channels), with the amount of scattering (measured by infrared channels), scientists can detect the percentage of vegetation in an AVHRR image and minimize the contributions of other surface cover.
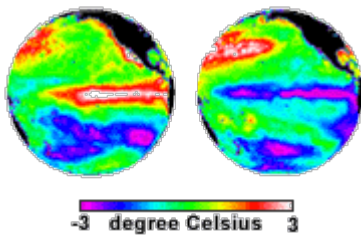
Sea surface temperature anomalies for El Niño and La Niña (Image courtesy of Jet Propulsion Laboratory)

To achieve global coverage, the researchers developed a new data management procedure that combined data from six receiving stations, eliminated overlapping segments and improved data quality in the process by replacing dropped and bad scan lines with sound data, while reducing the number of data units to be handled. Eidenshink calls the method "orbital stitching."

"It was a megasystem," he said. "We got to the point where we can process one day's data every day." Along the way, the project succeeded in developing processing standards for calibration, atmospheric correction, and geometric registration of the AVHRR data. Collaborating and receiving science guidance was key, Eidenshink said.

"From the beginning, the EDC DAAC was necessary," Eidenshink said. "DAAC funds were used to acquire the data, Pathfinder funds were used to pay for processing, the DAAC distributed the data and products to the science community. The DAAC science advisory team worked to keep momentum on the project and helped keep linkages between AVHRR and Terra," Eidenshink said.

As the value of the Pathfinder program began to be felt within the scientific community, a 1994 NASA Research Announcement (NRA) re-competed the 1990 product generation awards, and solicited proposals for eight new Pathfinder efforts that would strengthen and extend existing product generation. Among proposals awarded funding were projects to produce 1.25 and 5 km AVHRR grids over the North and South Poles.

Here, the invaluable connections established by Eidenshink's product generation team fed the polar effort. Acting as a clearing house for data from Fairbanks and Tromso, Eidenshink's team staged those data along with data streams from Antarctic stations to the National Snow and Ice Data Center (NSIDC) in Boulder, perfecting an automated data transfer system in the process.

Ted Scambos, a research scientist with the National Snow and Ice Data Center, and co-PI with James Maslanik on the AVHRR Polar Pathfinders, said "We saw a need for establishing as long-term a baseline as possible indicator of conditions of the snow pack, such as temperature and albedo. Such a baseline had already been established for the passive microwave data sets, with a nice record going back 25 years, but at much coarser

resolution. That kind of record doesn't exist in the visible and thermal channels," he said.

The Polar Pathfinders proposed development of a detailed calibrated record of temperature, ice motion, and albedo in the polar regions, extending from the early 1980s.

"We learned a lot going through the process," Scambos said. "Initially we had conservative objectives in terms of how much data we were going to be able to process. Four years ago it seemed a much more daunting task to handle terabytes of data than it does now."

Other improvements included refining the processing algorithm, Scambos said. "We got to know the data better and learned where the weak spots were in data product generation. For instance, telling clouds apart from the ice surface was always a problem spot. It's very difficult to come up with good cloud algorithms over snow and ice. When you've got liquid water clouds over snow and ice, the problem is not severe, but with ice-particle clouds over a dry snowy surface, the distinction becomes difficult. A cold snowy surface and a cirrus cloud are very similar in just about every respect.

"We've made strides on that mostly by looking for day-to-day surface changes. On a daily basis, the true surface doesn't change much, but of course clouds pass by with an attendant big change in texture, temperature, and albedo, so we can eliminate pixels on that basis.

"These data are going to be helpful in comparison with field studies on localized areas. Compared to coarser resolution data offered by passive microwave sensors, the AVHRR is more appropriate for regional ice sheet, rather than hemisphere-wide, monitoring," Scambos said.

Landsat 7 image of Iceberg B10A (Image courtesy of U.S. Geological Survey)

Scambos and Maslanik teamed with Polar Pathfinder data producers working with different instrument data to package products in the same grid projection.

"In effect, this collaboration created a data set that would be the equal of the EOS Terra platform, in the decade leading up to the Terra launch, because we organized data from the same day, in the same grid, from several different sensors, from several different satellites in a common format that's easy to compare. It offers the possibility of deriving unique synthetic products, comparing, say, passive microwave emission with true skin temperature and albedo variations with variations in the polarization in the passive microwave," Scambos said.

Both Eidenshink and Scambos look forward to the next generation radiometer to AVHRR, the Moderate Resolution Imaging Spectroradiometer (MODIS) scheduled to launch aboard Terra in winter 1999.

Eidenshink was a member of the MODIS land science team, and was the only participant who brought experience processing large data sets, and the systems and science issues raised by users, to the table. Similar requirements for MODIS are expected, he said.

"Our land product generation team provided raw data to the MODIS algorithm development team, and they were able to use our 1 km AVHRR data as a test data set." Another fallout from his team's work, Eidenshink said, was the ability to advise French teams on ground processing the data from the Spot 4 satellite launched in March 1998, also equipped with a vegetation instrument.

A new NASA Research Announcement concluded in September 1999. This time, James Dodge says, the thrust of proposals will be research. "We don't want to just spend another three years extending the data sets without some serious analysis. We are not a data organization, we're a research organization. We solve a need for having data to study, in order to get an idea of the time and space variabilities of climate parameters. This next set of proposals will be focused on analysis, on the meaning of the variabilities," he said.

"Overall, the Pathfinder Program was a basis from which to begin the research that will lead to EOS studies with data from better sensors," Dodge said. "The Pathfinder Program was a good ramp-up. It was big but it's certainly nowhere near as big as what researchers have coming when Terra starts broadcasting its data.

"The DAACs have been extremely helpful and useful in this process and very cooperative. Sometimes our scientists came in with challenging requirements, but the DAACs figured out ways to meet them."

Source URL: https://earthdata.nasa.gov/featured-stories/featured-research/noaanasa-pathfinder-program

Links:
[1] https://earthdata.nasa.gov/featured-stories/featured-research/noaanasa-pathfinder-program

Feedback