NOAA/NASA AVHRR Oceans Pathfinder

Sea Surface Temperature Data Set

User's Guide

Version 1.2

September 27, 1995

Jorge Vazquez

Andy Van Tran

Rosanna Sumagaysay

Elizabeth Smith

Mike Hamilton

Click here for Picture (JPL)

Click here for Picture (PO.DAAC)

Jet Propulsion Laboratory

California Institute of Technology

TABLE OF CONTENTS

1.0 INTRODUCTION

1.1 Scope

This document describes the production, quality assurance, archive, and methods of acquiring and using the Advanced Very-High Resolution Radiometer (AVHRR) Ocean Pathfinder sea-surface temperature (SST) data set (hereafter, Pathfinder SST). It briefly discusses the AVHRR instrument and National Oceanic and Atmospheric Administration (NOAA) satellite platforms, the history of satellite-derived SST, and the refinement of the nonlinear correction algorithm and cloud clearing techniques used in reprocessing AVHRR level-1B global-area coverage (GAC) data to produce the Pathfinder SST data set. It provides information on the various Pathfinder SST products and details of the distribution of the data by the Earthe Observing System Data and Information System (EOSDIS) Physical Oceanography Distributed Active Archive Center (PO.DAAC) at the Jet Propulsion Laboratory (JPL).

1.2 Project Description

The Earth Observing System (EOS), centerpiece of NASA's Mission to Planet Earth, is expected to generate about 2 trillion bytes of new data per day. The Pathfinder projects were initiated as preparation for the projected volume of data that will be returned from EOS. The mandate of the Pathfinder SST task is to produce, validate, and evaluate a long time-series of AVHRR-derived SST as a precursor to EOS data sets, and for use in climatologic investigations and modeling. Historically, radiometer data have been used with several algorithms to produce estimates of SST (see McClain et. al, 1985 for a review). Multichannel sea-surface temperatures (MCSST) have been computed from AVHRR radiances operationally since 1981.

As part of the Pathfinder SST task, a detailed re-analysis of the calibration procedures for the NOAA AVHRR based on thermal vacuum test data was performed by researchers at the University of Miami (Brown et al., 1993). This analysis has been done using the multi-channel instruments on board the NOAA-7, NOAA-9, and NOAA-11 satellites (see Table 1). New procedures were derived which improve the overall calibration accuracy as well as an increase in the number of temperature retrievals by an approximate factor of two.

NASA's JPL is tasked with reprocessing historical AVHRR data over the oceans to produce a prototype satellite SST database for global climate studies. This will provide a consistent SST time series of greater than 10 years, from 1981 to the present, with known statistics across the various AVHRR platforms. Comprehensive calibration and validation information are included, and access to the data set is through the use of new technologies in storage and retrieval.

2.0 SATTELLITE AND INSTRUMENT

A more detailed description of the NOAA-series satellites, the AVHRR instrument, and the AVHRR Global Area Coverage (GAC) Level-1B data can be found in the Polar Orbiter Users Guide (Kidwell, 1991), which can be obtained from the National Oceanic and Atmospheric Administration's National Environmental Satellite and Data Information Service (NOAA/NESDIS), and from which portions of the following information is reproduced.

2.1 Platform Description

Each of the series of NOAA satellites operates in a near-polar, sun-synchronous orbit. The orbital period is ~102 minutes, giving 14.1 orbits per day. Because the number of orbits/day is not an integer, the suborbital tracks do not repeat daily, although the local solar time of the satellite's passage is essentially unchanged for any latitude (Kidwell 1991). The 110.8[[ring]] cross-track scan equates to a swath width of about 2700 km. This swath width is greater than the 25.3[[ring]] separation between successive orbital tracks, and provides overlapping coverage. Table 1 lists the operational dates and nominal ascending node equator-crossing times for all platforms, although the data described in this User's Guide are currently derived only from the NOAA 7,9 and 11 satellites which carry the 5-channel instrument. The advent of the 5-channel multiwindow AVHRRs provide a better correction for atmospheric attenuation than the previous 4-channel correction methods (see McClain, 1985).

Table 1

NOAA-Series Satellites

(from the *Polar Orbiter Users Guide*, Kidwell 1991)

SATELLITE	DATES OF OPERATIONAL DATA COLLECTION	EQUATOR CROSSING TIME
TIROS-N NOAA-6 NOAA-7*	OCT 19, 1978-JAN 30,	15:00 LST 19:30 LST
NOAA-8 NOAA-9* NOAA-10	1980 JUN 27, 1979 - MAR	14:30 LST 19:30 LST
NOAA-11* NOAA-12	5, 1983 AUG 19, 1981 -	14:30 LST 19:30 LST
	JUN 7, 1986 JUN 20, 1983	14:30 LST 19:30 LST
	- OCT 31, 1985 FEB 25,	
	1985 - NOV 7, 1988 NOV	
	17, 1986 - present NOV	
	8, 1988 - present MAY	
	14, 1991 - present	

* used for SST processing

2.2 AVHRR Instrument Description

Each of the NOAA polar-orbiting satellites have carried an AVHRR as one of three sensors aboard the spacecraft. AVHRR was designed for multispectral investigations of meteorological, oceanographic, and hydrologic parameters, measuring emitted and reflected radiance in four or five spectral bands (Table 2), spanning the visible portion of the spectrum to the thermal infrared. Coverage is global, twice daily, at an instantaneous field of view (IFOV) of ~1.4 milliradians, giving a ground field of view of ~1.1 km at nadir for a nominal altitude of 833 km.

Table 2

AVHRR Spectral Bands

	Channel Position um (from 1 to 5)				
Platform	1	2	3	4	5
Tiros-N	0.55-0.90	0.725-1.10	3.55-3.93	10.3-11.3	
NOAA-7,-9,-11,-12	0.55-0.68	0.725-1.10	3.55-3.93	10.3-11.3	11.5-12.5
NOAA-6,-8,-10	0.55-0.68	0.725-1.10	3.55-3.93	10.3-11.3	

The AVHRR has a cross-track scanning system which uses a mirror rotating at 360 RPM about an axis parallel to the Earth. The instrument is designed to maintain a constant operating temperature for the IR detectors and provide a signal-to-noise ratio (SNR) of 3:1 at 0.5% albedo. Each AVHRR scan views Earth for 51.282 milliseconds, during which time each channel of the analog data output is digitized. Scans occur at the rate of 6 per second, and the sampling rate of the AVHRR sensors is 39,936 samples per second per channel. During a scan, the detectors view an internal target, cold space, and the external scene. The temperature of the internal target is monitored, and space is assumed to have a black-body temperature of 3[[ring]]K. In this way, a simple two-point linear calibration is done internally (Schwalb, 1978). The nonlinear modification to this calibration is achieved at the time of postprocessing, and takes into account sensor nonlinearities, measurement of internal target temperature, calculation of target radiance, internal reflections and emissions, etc.

2.3 Level - 1B Data

The Global Area Coverage (GAC) data are subsampled from the original 1 km AVHRR resolution to approximately 4 km, recorded internally, and downlinked daily. These data are the starting point for the Pathfinder SST processing. Level-1B data are defined as radiometrically-corrected and calibrated data in physical units at full instrument resolution as acquired. To produce the NOAA GAC Level-1B data, the Level-0 (unprocessed) instrument data are quality controlled, assembled into discrete data sets, and have calibration and Earth location information appended. Data are then stored as full orbits consisting of both ascending (daytime) and descending (nighttime) data.

3.0 ALGORITHM AND PROCESSING

The history of SST computation from AVHRR radiances is discussed at length by McClain et al. (1985). Briefly, radiative transfer theory is used to correct for the effects of the atmosphere on the observations by utilizing "windows" of the electromagnetic spectrum where little or no atmospheric absorption occurs. Channel radiances are transformed (through the use of the Planck function) to units of temperature, then compared to *a-priori* temperatures measured at the surface. This comparison yields coefficients which, when applied to the global AVHRR data, give estimates of surface temperature which have been nominally accurate to 0.3[[ring]] C.

Recently, the AVHRR thermal vacuum test data have been examined in detail (Brown et. al., 1993) in order to quantify drift in the calibration coefficients of the channels. Through this work, the nonlinear SST algorithm first developed by Walton (1988) has been modified with a time-dependent term. Processing has been further modified by dividing the earth into three regimes of atmospheric water vapor. Regression coefficients are computed independently for each of these regimes, to compensate for the well-known limitations of AVHRR SST retrievals in tropical areas, which are an artifact of high humidity.

A detailed description of the Pathfinder SST data processing is attached as Appendix A.

3.1 Computation of Sea-Surface Temperature

The AVHRR Level-1B sensor counts in the visible channels (1 and 2) are first converted to Rayleigh-corrected radiances and then to optical depth for use in removing the effects of the atmosphere and viewing and illumination geometry. Channels 3-5 are transformed to units of "brightness temperature", using the Planck black body function and a newly-determined (Brown et al., 1993) correction for sensor calibration non-linearity in the longer-wavelength channels. The algorithm used is essentially the nonlinear SST (NLSST: Walton, 1988), with a modification for sensor calibration drift with time. The algorithm is also conditioned in three regimes of atmospheric water vapor, and separate regression coefficients are applied. The form of the algorithm is:

SST = a + b*T4 + c*(T4-T5)*Tsurf + d*(sec([[theta]])-1)*(T4-T5) + e*lifetime (1)

where [[theta]] is the zenith angle of the instrument, and T4 and T5 are the brightness temperatures from AVHRR channels 4 and 5, respectively, as determined using the procedure outlined in NOAA Technical Report NESDIS 69. Tsurf is an *a-priori* estimate of the SST. It is calculated after a spatial interpolation to the 9km grid of the weekly, 1-degree optimum interpolated SST analysis produced by Dr. Richard Reynolds of NOAA/NESDIS (Reynolds and Smith, 1994). The spatial interpolation used is a bilinear interpolation of the 4 closest neighboring points surrounding the 9km grid point. The empirical coefficients a, b, c, d, and e were determined through a multiple-regression of AVHRR radiances with a database of *in-situ* temperatures, measured using moored and drifting buoys. These coefficients are calculated over a yearly time scale. In order to be considered a match, the pixel location and *in-situ* measurement must differ by no more than 0.1 degree spatially, and temporally by no more than 30 minutes. The following is a table of the coefficients, including the benchmark period. As future coefficients are added they will be included via World Wide Web for the project (http://podaac.jpl.nasa.gov/sst). Each line is for a different range of brightness temperature differences for channel 4 - channel 5 or T4 - T5.

Table 3

Coefficients

(NOAA-9: Valid for 1987-1988)

a	b	c	d	e	T4 - T5
1.600	0.930	0.149	0.085	-0.029	< 0.7
2.202	0.953	0.089	0.673	-0.086	>=0.7 & <1.8
4.470	0.868	0.090	0.439	-0.176	>=1.8

(NOAA-11: Valid for November 1988 - May 1991)

a	b	c	d	e	T4 - T5
1.011	0.934	0.120	1.118	0.026	< 0.7
1.336	0.945	0.087	0.984	0.003	>=0.7 & <1.8
3.250	0.869	0.078	0.805	0.009	>=1.8

New coefficients will be calculated and evaluated by the science working team during 1995. These coefficients will be based on a monthly time scale. Only the time period for calculating the coefficients themselves will change and **not** the algorithm. Thus significant changes in the data are not anticipated. Any changes in the processing will be reported through the FTP site or the homepage world wide web (*http://podaac-www.jpl.nasa.gov/sst/*).

An integral part of the NOAA/NASA AVHRR Oceans Pathfinder SST data set is validation. To aid in this process, a large validation data set has been compiled and is being distributed with the satellite-derived SST fields. Podesta et al (1995) describes a multi-satellite, multi-year database of AVHRR and high-quality in situ SST match-ups. The Pathfinder Matchup Data Base (PFMDB) is distributed in yearly files. It contains all the parameters necessary to validate the Pathfinder SST fields. The matchups also facilitate the application and validation of other SST algorithms to the original data. These databases are available through the JPL PO.DAAC Data Availability version 1-94. An important point to be made is that if better coefficients become available, especially during periods of high aerosol content, reprocessing might occur. People who want automatic updates of the of Pathfinder processing can add themselves to the mailing list accessible through the PO.DAAC homepage URL (*http://podaac-www.jpl.nasa.gov*).

3.2 Initial Assignment of Quality Flags and Cloud Detection

Temperature retrievals as detailed in section 3.1 are determined for all available pixels. Several tests are then performed to assign an estimate of the quality of each retrieval, in the form of a flag with four possible values. The "satellite" test uses a channel 4-5 threshold to detect how warm a pixel is, combined with a spatial homogeneity test. In addition, daytime images are subjected to a channel 3-5 difference threshold in order to elimimate sunglint problems. The "Reynolds" test is a comparison of the initial temperature retrieval to the weekly, 1-degree optimally interpolated SST field produced by Dr. Richard Reynolds of NOAA/NESDIS (Reynolds and Smith, 1993). The Reynolds data set combines sea surface temperature from the AVHRR instrument with in-situ temperatures to create the optimally interpolated SST field. If a pixel passes all of these tests, it is considered "best", and assigned a quality flag of 3. Passing the Reynolds test but failing the satellite test generates a 2 (or "mediocre" quality), failing the Reynolds test and passing the satellite test generates a 1, and failing all tests gives a quality flag of zero (indicating the pixel is most likely contaminated by a cloud). The following is a detailed scenario of how these flags are assigned.

Level 1 testing (Satellite tests)

These tests involve doing comparisons between different channels.

1) If channels 3 - 5 differ between -2 and 6 degrees the test passes and the pixel is good

2) If channel 4 < 35 degrees centigrade and channel 5 < 35 degrees centigrade the test passes and the pixel is good, otherwise the pixel is labelled as bad and test fails

3) If channel 4 abs(max-min) values are greater than a 2 degree threshold with neighboring pixels within a 3x3 box the pixel is bad and test fails.

4) If channel 5 abs(max-min) values are greater than a 2 degree threshold with neighboring pixels within a 3x3 box the pixel is bad and test fails.

5) If the satellite zenith angle > 55 degrees the test fails.

Level 2 testing (Comparison with Reynolds climatology)

1) If the calculated temperature from the satellite is greater or less than 2 degrees from the Reynolds interpolated value the test fails.

The flag values are then assigned based on the two levels of testing.

The following quality flags are assigned based on the previous scenario.

A value of 3 (good) is assigned if both level 1 and 2 testing pass . A 2(mediocre) is assigned if any of the level 1 testing fails but passes the second channel 5 (max-min) test , zenith angle test, and the level 2 testing (Reynolds comparison).

A value 1(bad) is assigned if the satellite tests all pass but level 2 testing fails (Reynolds).

A value 0(bad) is assigned if any of the satellite tests fail and the level 2 testing fails.

Declouding is further effected through the creation of composite images over three weeks before and after the target week, and means computed from these. The composite means are used to fill a central weekly mean image which contains the day being declouded (if the central mean image is missing values, and if there is a mean pixel of sufficient quality). If the weekly means from week(n-1) or week(n+1) cannot be used to fill empty values in the central (week n) mean, a spatial interpolation is done. The completely-filled weekly image is then compared to the daily image, and simple thresholding is used to indicate partial or complete cloudiness. Repeating this process generates a cloud mask for every day of data.

3.3 Binning and Mapping

The Pathfinder SST data are processed in an equal-area grid (Appendix C) based on one developed by the International Satellite Cloud Climatology Project (ISCCP). The bin size is approximately 9.28 km on a side, which gives 5,940,422 bins over the globe. An advantage of this grid is that it can be easily combined into grids with different zonal resolutions because the number of bins per row is always an integer. Since the GAC data were originally sampled at approximately 4 km resolution, bin values are averages. The number of retrievals which were averaged into each bin is a standard data product, and can be obtained to correctly perform any special weighted averaging.

After processing, the data are remapped into an equal-angle projection (4096x2048 rectangular grid), in order to facilitate visualization and extraction of regional subsets. The data are stored in flat binary and Hierarchical Data Format (HDF). Several products are generated from these data, including 18 km and 0.5 [[ring]] equal-angle projections. See Appendix A for details of the binning procedure.

4.0 QUALITY ASSURANCE

A semi-automated quality-assurance (QA) scheme has been developed which examines AVHRR SST retrievals for temporal and spatial consistency, and for areas in each image which may be missing retrievals due to instrument or processing errors. This is carried out in a three-part statistical post-processor, followed by a visual inspection. The automated portion of the analysis serves to guide an operator in the visual inspection phase, greatly reducing the time necessary to characterize spurious findings. The data quality information thus gained is passed to the end-user in the form of a QA report for each image file. The QA analysis provides a qualitative and quantitative description of anomalies found in the data.

4.1 Temporal Analysis

Phase 1 is a time-series examination, using a centered running mean over a window 21 days in length and comparing each retrieval (cloudfree pixels with highest-quality SST estimates) to this mean and a threshold. This examination will potentially generate 3 flag values. If the retrieval is outside an envelope defined by the running mean plus-or-minus 2[[ring]]C, it is flagged as anomalous (too high or too low). The 2[[ring]]C threshold was selected to be consistent with that used in the autoprocessing. If the retrieval was the only cloudfree day in the 21-day window, it is flagged for a secondary spatial examination, since there was effectively no time series with which to compare. The length of the window was determined after a sensitivity analysis, which determined the effect of varying this length on the number of flags generated in each category. As the window increases, it becomes less probable that a retrieval will be the only one that was free of clouds. The 21-day length was selected as it is close to a temporal mesoscale in most of the world ocean, and a mean determined over this period will smooth over large changes in temperature that are reasonable with respect to physical oceanography. This then avoids setting spurious flags that must be cleared by visual inspection. Further, the number of flags generated as a function of window-length (from the sensitivity analysis) begins to asymptote at 21 days, so that using a larger window will not result in significantly fewer flags being set.

4.2 Spatial Analysis

The result of the phase 1 time series examination is that many flags are set, mainly a result of retrievals being the only cloudfree day within their 21-day window. Most of these retrievals are perfectly reasonable temperatures, however since there was no time series with which to compute a running mean, there is no way to know this without further examination. Phase 2 is therefore a spatial test, which compares each previously-flagged retrieval to a spatial mean plus-or-minus a 1.5[[ring]] threshold. A sensitivity analysis for this test examined the effect of varying the temperature threshold, as well as different spatial radii for computing the local mean. Since we are examining an oceanographic quantity, we use the minimum bin radius (1.5 pixels = 13.5 km) to avoid biasing the mean where there may be a large change in temperature over a small distance (for example on the edge of a front, Gulf Stream, etc.) If the value of a flagged retrieval is within the accepted range of a spatial mean computed in this fashion, the flag is cleared. If the pixel is still out of range, it is left flagged for visual inspection.

4.3 Automated Detection of Missing Data/Dropped Orbits

The processed temperature data are examined by another automated portion of the process, which examines each image for missing orbits and misprocessed data. Each image is split into four quadrants, and the frequency distribution of missing data are computed. If the number of pixels without valid temperature or landmask in a quadrant exceeds 5% of the total, the quadrant is flagged as missing an orbit. This information is checked against the NOAA tables of known data dropouts for each instrument, and if it is determined that missing portions are due to processing errors, the data are queued for reprocessing. Information on missing data is contained within the QA report for each image.

4.4 Visual Inspection by an Operator

Using the previous tests to guide the search, a visualization package is used which allows an operator to browse an image in segments, toggle the display of flagged data locations, and display scalable zoomed subsets of the image in detail. The zoomed subsets are displayed along with a histogram of temperature distributions in the subset area, a contour map overlay of the locations that remain flagged, and a display of all temperatures in the row or column in which the cursor is located. This allows the user to zoom in on problem areas and carefully examine the area around flagged pixels, and make a determination as to whether the retrieval should remain flagged or not. If the flag is to remain set, the operator adds the location of the pixel to a file. This file is later merged with the processing flags. The operator also makes notes as to anomalous patterns which may be apparent in the gross appearance of the image. For example, missing orbits which are not due to instrument malfunction or nonoperation, linear features which may be due to the suturing-together of scan lines, single pixels marked as cloudy which are not near any other masked pixels, and large temperature gradients near feature boundaries are all noted.

4.5 Warnings on Usage

At this point, few users have experience with this newly processed product, so its accuracy is still an open issue. Some past restrictions using the MCSST data may apply.

1) The measurement is one of skin, and not bulk temperature (Schuluesso et al., 1990).

2) Atmospheric water vapor partly affects the retrieval, but no independent water vapor data sets are used (Emery et al, 1994).

3) Most successful uses of past MCSST data concentrated on identifying spatial temperature gradients (Gulf Stream fronts, etc.) rather than absolute temperature values. However, our calibrations are aimed at making calibrations that are consistent in time.

4) Clouds in any one image can be minimized by taking the warmest pixel at a fixed location over all images within one week. The logic is that clouds are "cold", and they move much farther in one week than ocean features.

Your experience with this data is valuable to the NOAA/NASA AVHRR Oceans Pathfinder Project and ask that you inform the team as to what you find about its appropriateness, accuracy, etc.

5.0 DATA SET DETAILS

The NOAA/NASA AVHRR Oceans Pathfinder SST data are distributed in a variety of resolutions, projections, and temporal averages, to accomodate researchers with varying processing capabilities and needs. Each data product is produced as both ascending (daytime) and descending (nighttime) images. These products are produced as daily composites, which are defined as spatial bins of all temperature retrievals at a maximum resolution of 9 km. Auxiliary information include quality and sampling data, as well as simple statistics. From the daily products, weekly, monthly, and yearly composites are formed. It is important to remember that composites and data products will be made available as they are produced. The WWW URL (http://podaac.jpl.nasa.gov/sst) will always give the current status of data production. A detailed description of the products follow.

Example of Monthly Composite of Sea Surface Temperature from NOAA/NASA AVHRR Oceans Pathfinder SST Data Set

Click here for Picture

5.1 Equal Area Product

The equal area product is based on a gridding scheme where the number of bins per longitude is dependent on the latitude. For details of this binning scheme refer to Appendix A. One of the data sets generated for distribution is a 9km equal-area product with 6 different bands or extractable parameters describing the sea surface temperature in a given bin. These are distributed as HDF files, and are approximately 120 MB in size. The equal-area files are also available, upon special request, to users who are familiar with the DSP language. The sum squared SSTs and number of observations per bin are included for proper resampling, should a researcher have special spatial or temporal requirements. Pixel quality and mask bits are determined during processing, and based on a variety of tests. Details of the pixel quality and mask bits and the method of their determination is given in section 3.2. The 6 bands included in an equal-area product in Table 4.

Table 4

Extractable Parameters for an Equal Area Product

```
  BAND    PARAMETER                DESCRIPTION
    1      Bin Number               A unique number assigned to a particular
                                   bin based on the equal-area grid. This
                                   bin_number then is associated with a
                                   specific geographical or latitude, longitude
```

		coordinate.
2	# of Observations/Bin	Because the 9km bins are based on an average of 4km Level-1B data, this parameter indicates the number of observations that went into the average of each bin.
3	Pixel Quality	A quality flag generated during processing, which indicates the quality of the temperature estimate at each pixel. Values can be between 0 and 3 inclusive, depending on a series of statistical tests and comparisons with other sources of data (see section 3.2).
4	Mask Bits	This band contains different image masks that are used, such as cloud or ice masks.
5	Sum SST	For a given 9km bin this number is the sum of the sea surface temperature values in that bin. This number, along with the number of observations per bin, can then be used to derive the average sst value.
6	Sum Squared SST	For a given 9km bin this number is the sum of the squared sst values, to be used in computing higher-order statistical moments.

5.2 9km Equal Angle All SST

A significant part of the processing is in mapping the equal-area product into a format suitable for image display and analysis. Thus, the 9km equal-angle product consists of the mapped equal area grids onto an equal-angle grid with an image size of 4096 x 2048. This product contains all pixels regardless of data quality flag and will be available in the Hierarchical Data Format (HDF) developed by the National Center for Supercomputing Applications. In addition to the daily day/night data, an on-line compositing system is being developed by the Pathfinder team to allow individuals to create their own composites from 2 days to a year. See section 6.3 for more details. The product contains 3 bands or image planes of data:

```
              Data Set:  9km Equal Angle All Sea Surface Temperature
            Image Size:  4026 x 2048 (DAILY)
             Data Size:  ~24.7  MBytes  (~5.3 MBytes compressed)
                Format:  HDF
# Extractable Parameters:  3 Bands
                Band 1:  Sea Surface Temperature: Value of retrieval
                Band 2:  Pixel Quality: Flag Value between 0 and 3 as
                         defined in Appendix B and Section 3.2
                Band 3:  Number of Observations Per Bin: Number of SST
                         values that were averaged from the 9km bin.
```

5.3 9km Equal Angle Best SST

Same dimensions as the 9km equal-angle product except only those SST values with a pixel quality of 3 are kept. This product is produced in the HDF format but in addition is also available in a raw binary image format. It contains two bands or image planes of data:

```
              Data Set:  9km Equal Angle Best Sea Surface Temperature
            Image Size:  4026 x 2048 (DAILY)
             Data Size:  ~16.5  MBytes (~1.2 MBytes compressed)
                Format:  HDF and Raw Binary
# Extractable Parameters:  2 Bands
                Band 1:  Sea Surface Temperature: Value of retrieval
                Band 2:  Number of Observations Per Bin: Number of SST
                         values that were averaged from the 9km bin.
```

5.4 18km Equal Angle All SST

This is the same as section 5.2 except the spatial bin size is now 18km instead of 9 km, so that the dimensions of the image are 2048 x 1024. This product is available in the HDF format with 3 bands or image planes of data:

```
               Data Set:  18km Equal Angle All Sea Surface Temperature
             Image Size:  2048 x 1024 (DAILY)
              Data Size:  ~6.3  MBytes (~1.8MBytes compressed)
                 Format:  HDF
# Extractable Parameters:  3 Bands
                 Band 1:  Sea Surface Temperature: Value of retrieval
                 Band 2:  Pixel Quality: Flag Value between 0 and 3 as
                          defined in Appendix B and Section 3.2
                 Band 3:  Number of Observations Per Bin: Number of SST
                          values that were averaged from the 18km bin.
```

5.5 18km Equal Angle Best SST

This product is identical to section 5.3 except that the data have been averaged from 9 km to a spatial bin size of 18km, giving a dimension of 2048 x 1024. This product is available in the HDF or raw binary formats with the same 2 bands or image planes:

```
               Data Set:  18km Equal Angle Best Sea Surface Temperature
             Image Size:  2048 x 1024 (DAILY)
              Data Size:  ~4.2  MBytes (~406 KBytes compressed)
                 Format:  HDF and Raw Binary
# Extractable Parameters:  2 Bands
                 Band 1:  Sea Surface Temperature: Value of retrieval
                 Band 2:  Number of Observations Per Bin: Number of SST
                          values that were averaged from the 18km bin.
```

5.6 All SST Equal Angle 0.5 Degree (54km) SST

The purpose of the 0.5 (54 km) images is to allow the researcher to quickly view the data and decide if the data are suitable for his/her purposes. In addition these data are also available for scientific studies of global scale ocean phenomena. The data consist of a 0.5 degree product, 720 x 360. As in previous sections (5.2, 5.4) this data set contains all pixels and not just those with a quality flag of 3. It is also available in the standard day or night daily, monthly and yearly composites. It is available in the HDF format and consists of the 1 band of SST values and 2 additional bands, one of which is the number of observations per bin and the other is the quality flag.

```
               Data Set:  54km Equal Angle All Sea Surface Temperature
             Image Size:  720 x 360 (DAILY)
              Data Size:  ~777.6  KBytes
                 Format:  HDF
# Extractable Parameters:  3 Bands
                 Band 1:  Sea Surface Temperature: Value of retrieval
                 Band 2:  Pixel Quality: Flag Value between 0 and 3 as
                          defined in Appendix B and Section 3.2
                 Band 3:  Number of Observations Per Bin: Number of SST
                          values that were averaged from the 54km bin.
```

5.7 Best SST Equal Angle 0.5 Degree (54km) SST

This data set is the same as section 5.6 except that only the best SST values of quality flag 3 are kept.

```
                Data Set:  54km Equal Angle Best Sea Surface Temperature
              Image Size:  720 x 360 (DAILY)
               Data Size:  ~518.0  KBytes
                  Format:  HDF and Raw Binary
 # Extractable Parameters:  2 Bands
                  Band 1:  Sea Surface Temperature: Value of retrieval
                  Band 2:  Number of Observations Per Bin: Number of SST
                           values that were averaged from the 54km bin.
```

*All the above data, starting from the daily files, are available for subsetting and compositing through the on-line homepage URL (http://podaac-www.jpl.nasa.gov/sst).

5.8 Matchup Database

A large validation data set is being distributed with the NOAA/NASA AVHRR Oceans Pathfinder global SST fields (Podesta et al, 1995). This data set is a multi-satellite, multi-year compilation of AVHRR and in situ SST data matchups, and is called the Pathfinder Matchup Data Base (PFMDB).

The PFMDB is organized into several files, by year and by satellite. It includes quality-controlled in situ SST data from both moored and drifting buoys. Table 5 shows the sources of in situ SST data included in the PFMDB:

Table 5

Source of In Situ SST in PFMDB

```
BUOY TYPE:                SOURCE:
  Moored Buoys             - U.S. National Data Buoy Center (via NODC) - Japan
                          Meteorological Agency - TOGA/TAO Project NOAA Pacific
                          Marine Environment Laboratory
 Drifting Buoys            -NOAA Atlantic Oceanographic and Meteorological
                          Laboratory Canadian Marine Environmental Data Service
```

Most of the initial in situ data compilation and quality control was done in collaboration with Dr. Charles McClain and his research group at NASA's Goddard Space Flight Center.

AVHRR data were extracted at the times and locations of the in situ observations. The maximum temporal separation between the satellite retrieval and the in situ observation was required to be 30 minutes, in order for the pair to be considered a "match". Spatially, the satellite retrieval and in situ observation were required to be within approximately 10km of each other to be considered a match. The result of this matching process is a series of records which contain both satellite-derived and in situ observations. The quantities included in the PFMDB are in Table 6.

The filter code indicates which specific match-ups were used in the estimation of the SST algorithm coefficients. Only matchups with a filter code of 1 (25-30%) were used in the algorithm development. The rest of the matchups are ideally suited for validation because they were not used in the algorithm coefficient estimation.

Table 6

Quantities in PFMDB

```
Quantities in PFMDB
 In situ date&time
Satellite data&time
```

```
Satellite latitude&longitude
Average PRT temperature
Solar&Satellite zenith angles
Sun glint index
Emissivity for Ch3-5
Central value of 5x5 pixel box, Ch1-5
Min value of 5x5 pixel box, Ch1-5
Max value of 5x5 pixel box, Ch1-5
Average value of 5x5 pixel box, Ch3-5
PRT temps. 1-4
Buoy latitude&longitude, Buoy ID
In Situ SST
Filter Code
```

6.0 ARCHIVE AND DATA ACCESS

The NOAA/NASA AVHRR Oceans Pathfinder SST data are available through the PO.DAAC at the JPL. Because the processing of the level-1B AVHRR data to SST is also at JPL, the data can be browsed during the processing phase nearly as soon as the images are complete and have been checked for quality. The data may be accessed using a WWW browser, *http://podaac-www.jpl.nasa.gov*, such as Netscape, Mosaic or Lynx, downloaded using anonymous ftp, or by making a request through electronic mail (or by telephone) to the staff at the PO.DAAC. See Appendix B.

6.1 Using a Web Browser to Access the Data

Using a web browser tool, the http protocol may be used to access the NOAA/NASA AVHRR Oceans Pathfinder SST data as the images are produced, as well as learn more about the NOAA/NASA AVHRR Oceans Pathfinder SST project, JPL, and NASA. The uniform resource locator (URL) for the NOAA/NASA AVHRR Oceans Pathfinder SST homepage is *http://podaac-www.jpl.nasa.gov/sst*.

The homepage contains details of the AVHRR instrument, an overview of the NOAA/NASA AVHRR Oceans Pathfinder SST project, a description of the image resolutions, file formats, and projections that are available, and a user may browse the latest SST files added to the site. This is a dynamic process, and new files are added each day at the approximate rate of 20 days of data (day and night) per day of processing. Selecting an image to browse will display the image with colortable as well as the metadata header (see section 7.5 Metadata). This includes information on which satellite platform and AVHRR instrument were used to collect the channel radiances, calibration information, start and end times of the data collection, and other details.

The web site also has an on-line order form, so that a researcher may acquire any or all of the NOAA/NASA AVHRR Oceans Pathfinder SST data through the PO.DAAC. Upon completing and submitting the order form, an electronic mail confirmation of the order will be sent back, with an order reference number. The data will arrive through U.S. Mail on the selected media (8mm tape, DAT, etc.).

6.2 Regional Subsetting and Extracting Using WWW

As an additional service to the user an on-line regional subsetting capability has been included and is accessible through the NOAA/NASA AVHRR Oceans Pathfinder homepage. This capability allows the user to extract regional data from the global data set. The user selects a region using the maximum latitude, minimum latitude, maximum longitude and minimum longitude and this spawns an automatic on-line subsetting at JPL. When the subsetting is completed the user is informed through an e-mail with the location of the extracted data sets. The extracted data set may then be accessed via FTP. URL: *http://podaac-www.jpl.nasa.gov/sst/subset.html*

6.3 Compositing of Data Sets Using WWW

In addition, an on-line compositing capability has been included and is accessible through the NOAA/NASA AVHRR Oceans Pathfinder homepage. This capability allows the user to temporally composite data from the global daily data set into 2-days, 7-days, monthly to yearly composites. When the compositing is completed the user is informed through an e-mail with the location of the data composites. The composites may then be accessed via FTP.

6.4 Downloading the Data Using Anonymous FTP

An anonymous ftp may be used to obtain the NOAA/NASA AVHRR Oceans Pathfinder SST data. Connect to *podaac.jpl.nasa.gov* using ftp, and enter 'anonymous' for a user name. Please use your full e-mail address for a password, as you will be placed on the NOAA/NASA AVHRR Oceans Pathfinder SST mailing list. The directory structure as of the date of this document is contained in Appendix B. At the ftp prompt change directory to: *ftp>cd pub/archive/avhrr/*

There four main sub directories below *avhrr/* directory. They are *data/, document/, matchups/, and software/.*

document: this directory contains avhrr documentation such as the users guide in post script and a readme file.

matchups: this directory contains matchup databases.

software: this directory contains routines for reading the HDF data files, dumping an HDF image to a flat binary (byte, no header) formatted file, and a package written in IDL for browsing the full resolution images in detail. Use of the IDL routines and package requires that you have IDL version 3.5 or better installed on your system, and use of the FORTRAN routines requires that you acquire, compile, and install the HDF library (available via anonymous ftp from ftp.ncsa.uiuc.edu).

data: directory consists of the following three major products: Please note that the products referred by their product numbers (i.e. *product050*) are located below the *pub/* directory.

all_pixel or product050 -> Global Equal - Angle All Pixel SST

best_sst or product051 -> Global Equal - Angle Best SST

half_degree or *product053* -> 0.5 Degree Spatial Resolution Global SST

all_pixel/hdf_09: contains 9km (4096 x 2048) equal-angle (rectangular) projections which contain all retrievals. This includes temperatures, clouds, bad data, etc.; virtually everything before the data quality and cloud screening processes mask the retrievals of dubious quality. These are provided for researchers who may want to develop cloud detection procedures of their own, or use the full set of retrievals for some other purpose.

all_pixel/hdf_18: contains 18 km (2048 x 1024) equal-angle projections. These data are included as they are the same resolution as the MCSST data set, and can be used as a replacement for ongoing research without the necessity of rebinning by the researcher. These files contain images with all retrievals.

all_pixel/hdf_54: contains 54 km (720 x 360) equal-angle images.

best_sst/hdf_09: contains the same resolution as the *all_pixel/hdf_09/* directory, but with only the "best" (see section on pixel quality) temperature retrievals.

best_sst/hdf_18: contains the same projection and resolution as the *all_pixel/hdf_18/* directory but with clouds and poor-quality temperature retrievals masked.

best_sst/hdf_54: are the same images, with poor quality retrievals and clouds masked.

best_sst/raw_09: contains the same values as *best_sst/hdf_09/* but in raw binary image format.

Each of the ftp directories contains the same time-span of data; only the resolution of the images and the pixel quality varies between them. It is important to remember that the FTP site will always be updated as more data is produced.

6.5 Data Availability Through Information Management System

Data may also be ordered through the EOSDIS Version 0 Information Management System (IMS). For more information contact podaac@podaac.jpl.nasa.gov

7.0 READING AND USING DATA SETS

7.1 Format and Image Size

The JPL PO.DAAC is distributing these data in the Hierarchical Data Format (HDF). HDF was developed at the National Center for Super computing Applications at the University of Illinois at Urbana-Champain. The PO.DAAC is supplying FORTRAN drivers to read and write HDF data, however these must be linked with the HDF subroutine library. The library is available via anonymous ftp at ftp.ncsa.uiuc.edu (141.142.3.135). We also supply example IDL code for reading and writing HDF data; if you use IDL the HDF library is already included, however IDL version 3.5 or better must be installed at your site.

The files themselves are daily images of sea surface temperature data. Values range from 0 to a possible maximum of 255. Values of 0 refer to missing data or cloud cover. The format of the file consists of a byte array of dimension 4096 x 2048 for a 9km spatial resolution data set. Thus the files can be simply read by declaring a byte array of dimension (4096,2048) and reading the data into that array. The byte values can then be scaled into the appropriate sea surface temperature by using the following y-intercept and slope values.

y-intercept=-3.0

slope=0.15

and

SST=pixelvalue*slope + y-intercept

where SST is in degrees Celsius. For the 18km data the dimension of the byte array is s (2048,1024). The dimensions for the 54 km data is (720,360). See section 5.0 for data set details.

7.2 Read Software for Raw Images

The raw image files can then be read using the following programs written in either FORTRAN or IDL. These programs are available under the FTP site *podaac.jpl.nasa.gov* using the anonymous login.

FORTRAN PROGRAM

```
program read_raw

C

C Small bit of code to read a ns X n1 byte image
```

```
C mkh 11/94

C

parameter(ns=4096,n1=2048)

byte bimage(ns,n1)

open(unit=11,file='test.raw',access='direct',

> status'old',recl=ns)

do 10 iline=1,n1

read(11,rec=iline)(bimage(i,iline),i=1,ns)

10 continue

close(11)

C

C Your analysis code should go here.

C

end
```

IDL PROGRAM

```
;

;This program is designed to read only global AVHRR images.

;User must provide xsize and ysize

;Example rd_raw,'88001b09da-gdm.raw',4096,2048

;Written by: Andy V. Tran

pro read_raw,file,img,xsize,ysize

close,25

openr,25,file

img = bytarr(xsize,ysize)

readu,25,img

close,25

end
```

In the FORTRAN program a word of warning must be added. For the open statement the parameter for the record length "recl" is specified in bytes in Sun Unix, but on SGI Unix the "recl" parameter is specified in word length. Thus for an SGI machine the statement:

open(unit=11,file='test.raw',access='direct',, status'old',recl=ns)

should be replaced with

open(unit=11,file='test.raw',access='direct',, status'old',recl=ns/4).

7.3 Read Software for HDF Images

The hdf image files can be read using the following sample program written in IDL. This program is available under the FTP site *podaac.jpl.nasa.gov* using the anonymous login.

; **READ_HDF** written by Andy V. Tran

; Andy@grumpy.jpl.nasa.gov

; THIS will work only with IDL version 3.5 or Higher

; User don't have to specify plane

pro read_hdf,file,image,r,g,b,plane=plane

if n_elements(plane) eq 0 then begin

dfr8_restart

endif

dfr8_getimage,file,image,pal

r = bytarr(256)

g = bytarr(256)

b = bytarr(256)

w = n_elements(pal)

if(w gt 1) then begin

r(*) = pal(0,*)

g(*) = pal(1,*)

b(*) = pal(2,*)

endif

end

7.4 Regional Extraction and Subsetting Programs

Several programs in both IDL and FORTRAN have been written to extract regional data from the global image. The inputs to such programs include the geographic coordinates or latitudes and longitudes of the desired area. In addition a subsetting capability has been implemented through the WWW homepage (*http://podaac-www.jpl.nasa.gov/sst/*) which allows the user to do regional extraction. The following are example programs in IDL and FORTRAN which also do the regional extraction. These programs are currently available through the anonymous FTP site (podaac.jpl.nasa.gov) under the software directory. The subsetting is a different algorithm in the WWW homepage and is not based on either of the following programs. The point of providing these programs is to give the user as much flexibility as possible in doing regional extraction.

IDL EXTRACTION PROGRAM

```
pro extract, x,xlatmn,xlatmx,xlonmn,xlonmx,xext,i180

;

;convert from lat,lon coordinates to pixel coordinates

; input: image file and maximum, minimum latitudes and longitudes

; for region to extract

;x: byte array containing image data

; output: extract image file xext

;i180: parameter that controls whether want -180 to 180 or 0 to 360

; coordinate system

; : = 0 0 to 360

; : = 1 -180 to 180

if i180 eq 0 then xlon1=0.

if i180 eq 1 then xlon1=-180.

xlat1=-90.

xlon1=-180.

delta=4096./360.

iymin=fix((xlatmn-xlat1)*delta)

iymax=fix((xlatmx-xlat1)*delta)

ixmin=fix((xlonmn-xlon1)*delta)

ixmax=fix((xlonmx-xlon1)*delta)

print,ixmin,ixmax,iymin,iymax

nxdim=(ixmax-ixmin+1)
```

```
nydim=(iymax-iymin+1)

xext=bytarr(nxdim,nydim)

if i180 eq 1 then x180=bytarr(2048,1024)

if i180 eq 1 then x180(0:1023,*)=x(1024:2047,*)

if i180 eq 1 then x180(1024:2047,*)=x(0:1023,*)

if i180 eq 1 then x=x180

xext(0:nxdim-1,0:nydim-1)=x(ixmin:ixmax,iymin:iymax)

end
```

FORTRAN EXTRACTION PROGRAM

```
subroutine extract(x,xlatmn,xlatmx,xlonmn,xlonmx,xext,i180)

c

c *****convert from lat,lon coordinates to pixel coordinates

c*****input: image x (raw no header) and maximum, minimum latitudes c*****and longitudes

c*****for region to extract

c*****output: extracted image file xext

c

c i180: parameter that controls whether want -180 to 180 or 0 to 360

c coordinate system

c : = 0 0 to 360

c : = 1 -180 to 180

c

byte x(4096,2048),xext(4096,2048),x180(4096,2048)

if (i180 .eq. 0) xlon1=0.

if (i180 .eq. 1) xlon1=-180.

xlat1=-90.

delta=4096./360.

if i180 eq 0 then

do 60 j=2049,4096
```

```
do 61i=1,2048

x180(j-2048,i)=x(j,i)

61 continue

60 continue

do 70 i=1,2048

do 71 j=1,2048

x180(j+2048,i)=x(j,i)

71 continue

70 continue

endif

iymin=fix((xlatmn-xlat1)*delta)+1

iymax=fix((xlatmx-xlat1)*delta)+1

ixmin=fix((xlonmn-xlon1)*delta)+1

ixmax=fix((xlonmx-xlon1)*delta)+1

print,ixmin,ixmax,iymin,iymax

nxdim=(ixmax-ixmin+1)

nydim=(iymax-iymin+1)

ix=0

do 100 j=ixmin,ixmax

ix=ix+1

iy=0

do 101 i=iymin,iymax

iy=iy+1

if (i180 .eq. 1) xext(ix,iy)=x180(j,i)

if (i180 .eq. 0) xext(ix,iy)=x (j,i)

101 continue

100 continue

end
```

7.5 Metadata

The following is an example of the metadata taken from an Best SST 09km HDF file. Each HDF image or file has associated with it the metadata which is contained in a header. This particular file has two bands of data including the number of observations per bin and the best sea surface temperature value, see section 5.0. The metadata may be found by accessing the WWW through the Netscape/Mosaic homepage URL *http://podaac-www.jpl.nasa.gov/sst*/.

Scientific Brows-o-rama

Datasets

There are **2 datasets** and **32 global attributes** in this file.

Available datasets :

Dataset **AVHRR Oceans Pathfinder Equal Angle 4096 x 2048** has rank 2

with dimensions **[4096, 2048]**. The dataset is composed of signed 8-bit

integers. It has the following attributes :

Attribute scale_factor has the value : **0.150000**

Attribute scale_factor_err has the value : **0.000000**

Attribute add_offset has the value : **-3.000000**

Attribute add_offset_err has the value : **0.000000**

Attribute calibrated_nt has the value : **20**

Attribute Slope has the value : **0.150000**

Attribute Intercept has the value : **-3.000000**

Data set or band # 2:

Dataset number of observation **4096 x 2048** has rank 2 with dimensions

[4096, 2048]. The dataset is composed of signed 8-bit integers. It has the

following attributes :

Attribute Band Name has the value : **number of observation**

Global attributes :

Attribute Title has the value : **AVHRR Oceans Pathfinder Equal Angle**

Attribute Data Center has the value : **NASA/PO.DAAC AVHRR Oceans**

Pathfinder

Attribute Station has the value : **NOAA/NESDIS**

Attribute Mission has the value : **AVHRR Oceans Pathfinder**

Attribute Mission Characteristics has the value : **NOAA/NASA AVHRR OCEANS PATHFINDER**

Attribute Sensor has the value : **NOAA POLAR ORBITER DATA**

Attribute Sensor Characteristics has the value : **National Oceanic and Atmospheric Administration Polar Orbiter**

Attribute Product name has the value : **EQUAL ANGLE MAP HDF**

Attribute Quality type has the value : **Best SST**

Attribute Binning period has the value : **DAILY**

Attribute Replacement flag has the value :

Attribute Processing control has the value :

Attribute Processing log has the value :

Attribute Input files has the value :

Attribute Data start time has the value : **1987004051305**

Attribute Data end time has the value : 1**987004051305**

Attribute Data processing time has the value : **1995011170838**

Attribute Start year has the value : **1987**

Attribute End year has the value : **1987**

Attribute Start day has the value : **4**

Attribute End day has the value : **4**

Attribute Start millisec has the value : **51305**

Attribute Number of rows has the value : **2048**

Attribute Number of columns has the value : **4096**

Attribute Map projection has the value : **Equirectangular projection**

Attribute Parameter name has the value : **Sea Surface Temperature**

Attribute Orbit has the value : **10633.000000**

Attribute Maximum Latitude has the value : **89.956055**

Attribute Minimum Latitude has the value : **-89.956055**

Attribute Maximum Longitude has the value : **179.956055**

Attribute Minimum Longitude has the value : **-179.956055**

8.0 FREQUENTLY ASKED QUESTIONS

1. What is the naming convention used for the data sets?

An example file name is 88020h09da-gdm.hdf, where:

88 - year

020 - day

h - data type (h=hdf, p=postage stamp (or

equal-area projection),

d=dsp (byte with header

and trailer), b=binary

(byte, no header), and

x=extract/subsetted data).

09 - resolution (ground field-of-view, km)

d - daily (w=weekly, m=monthly)

a - ascending (daytime; d=descending (nighttime))

gdm - good declouded mean (highest-quality retrievals;

adm=all declouded mean)

hdf - file format (hdf; dsp, bin=binary)

2. Where can I learn more about the HDF data format?

Use a web browser (e.g. Mosaic) and make an http connection to the URL *http://www.ncsa.uiuc.edu /SDG/Software/HDF/HDF-FAQ.html*. This document describes the details of the format, new features, visualization and analysis tools (both free and commercial), how to obtain source code, how to make a bug report, etc. Documentation and technical specifications of the HDF data format can be obtained by using anonymous ftp to ftp.ncsa.uiuc.edu. The following are example programs on how to read and write HDF using the Interactive Data Language (IDL). Similar programs can be used in FORTRAN. These programs may be downloaded from the anonymous FTP site on *podaac.jpl.nasa.gov* under the software directory.

; **READ_HDF** written by Andy V. Tran

; Andy@grumpy.jpl.nasa.gov

; THIS will work only with IDL version 3.5 or Higher

```
; User don't have to specify plane

pro read_hdf,file,image,r,g,b,plane=plane

if n_elements(plane) eq 0 then begin

dfr8_restart

endif

dfr8_getimage,file,image,pal

r = bytarr(256)

g = bytarr(256)

b = bytarr(256)

w = n_elements(pal)

if(w gt 1) then begin

r(*) = pal(0,*)

g(*) = pal(1,*)

b(*) = pal(2,*)

endif

end

; WRITE_HDF written by Andy V. Tran

; Andy@grumpy.jpl.nasa.gov

; THIS will work only with IDL version 3.5 or Higher

; User doesn't have to specify plane

pro write_hdf,file,image,r,g,b,plane=plane

pal = bytarr(3,256)

pal(0,*) = r

pal(1,*) = g

pal(2,*) = b

if n_elements(plane) eq 0 then begin

plane = 1L

dfr8_setpalette,pal
```

dfr8_putimage,file,image

endif else begin

dfr8_setpalette,pal

dfr8_addimage,file,image

print,plane

endelse

end

3. How do I report a problem with the data, or contact someone at the JPL PO.DAAC?

Problems with the data, placing an order for data (without using the WWW interface), and questions of a general nature should be directed to the PO.DAAC User Services (podaac@podaac.jpl.nasa.gov). Technical questions should also be mailed to the same address, and will be spooled to the appropriate science personnel for attention.

4. What does "best" sst mean, as opposed to "all"?

For more information, see the section in the Users Guide describing the quality flags. Briefly, the best SST fields only contain data that have a quality flag of 3 (highest confidence in quality) corresponding to the temperature retrievals. All other values are flagged as missing, with pixel values set to 0. The all pixel data has values for all the pixels in the image, and thus includes quality flags of 1, 2, and 3. The assignment of these flags indicates that a temperature retrieval failed one or more of the processing quality tests, which are comparisons to climatologic means and examination of the radiance data in certain atmospheric windows as a check for clouds and high atmospheric aerosols.

5. How do I obtain Pathfinder data via ftp?

You can obtain compressed, binary (byte) data with no header via anonymous ftp using the following guide:

yourprompt% ftp podaac.jpl.nasa.gov

Name (podaac:yourlogin) anonymous

Password: you@your.mail.address

This will connect you with the podaac.jpl.nasa.gov anonymous ftp server which will contain the following directories specifically for NOAA/NASA AVHRR Oceans Pathfinder products:

+pub+archive+avhrr +documents

+software

+matchups

+landmask

+data +best_sst +hdf_09 +daily +1985 +ascending+<files>

+ + + + +descending+<files>

+ + + +1986 ->Same as 1985

+ + + +1987 ->Same as 1985

+ + + +1988 ->Same as 1985

+ + + +1989 ->Same as 1985

+ + + +1990 -> Same as 1985

+ + + +1991 -> Same as 1985

+ + +monthly -> Same as daily

+ + +weekly -> Same as daily

+ + +yearly -> Same as daily

+ +hdf_18 -> Same as hdf_09

+ +hdf_54 -> Same as hdf_09

+ +raw_09 -> Same as hdf_09

+all_pixel -> Same as best_sst

+half_degree +all_hdf_54 +daily +1985 +ascending +<files>

+ + + +descending +<files>

+ + +1986 -> Same as 1985

+ + +1987 -> Same as 1985

+ + +1988 -> Same as 1985

+ + +1989 -> Same as 1985

+ + +1990 -> Same as 1985

+ + +1991 -> Same as 1985

+ +monthly -> Same as daily

+ +weekly -> Same as daily

+ +yearly -> Same as daily

+best_hdf_54 -> Same as all_hdf_54

(Directory structure using product numbers)

+pub+product050 -> Same as /pub/archive/avhrr/data/all_pixel/

+pub+product051 -> Same as /pub/archive/avhrr/data/best_sst/

+pub+product053 -> Same as /pub/archive/avhrr/data/half_degree/

6. How much data is currently available?

Currently (as of September 1995) there is data from 1987-1991 with plans for processing from 1981 to the present. It is anticipated that by the end of the calendar year for 1995 data will be processed from 1985 to the present.

7. What products are available?

As of September 1995 data products available include the daily images for 1987-1991 with work being done to process the monthly and yearly products. It is anticipated that by the end of September-October time frame monthly products will be available for the 1987-1991.

8. Will this data be on a CDROM?

Yes. Monthly composite data for the same time periods as the daily fields will be available on CDROM.

9. What about landmask?

A global landmask at 9km, 18km and 54km resolutions will be available through the anonymous FTP site.

10. How do I convert Pathfinder values to SST's?

The pathfinder 9km data files are 4096x2048, 18km are 2048x1024, and 54km are 720x360 with each value equal to one byte. The range of values should be from 0 to 255. Once you have read in the data file, you may convert to temperature using:

temperature (degrees C) = bytevalue*0.15 - 3.0.

For further FAQs see the home page under *http://podaac-www.jpl.nasa.gov/sst/*.

9.0 REFERENCES

Brown O.B., J.W.Brown and R.Evans, 1985. Calibration of AVHRR infrared observations, *J. Geophys.. Res.*, **90 (C6),** 11667-11677.

Brown J. W., O. B. Brown, and R. H. Evans, 1993. Calibration of AVHRR Infrared channels: a new approach to non-linear correction, *J. Geophys.. Res.*, **98 (NC10)**, 18257-18268.

Emery, W., Y. Yu, and G. Wick, 1994. Correcting infrared satellite estimates of sea surface temperature for atmospheric water vapor attenuation, *J. Geophys. Res.*, **99 (C3)**, 5219-5236.

Evans, R. H., Shenoi, S. H., Podesta, G. P., 1994 (in prep), *A report on the exploratory analysis of pathfinder sea surface temperature retrieval algorithms*, University of Miami.

Evans, R. H., 1995 (in prep), Science Working Group Report.

JPL Physical Oceanography Distributed Active Archive Center (PO.DAAC) Data Availability, Version 1-94, JPL Publication **90-49**, rev. 5.

Kidwell, K., 1991. NOAA Polar Orbiter User's Guide. NCDC/NESDIS, National Climatic Data Center, Washington, D.C..

McClain E. P., W. G. Pichel, and C. C. Walton, 1985. Comparative performance of AVHRR based multichannel sea surface temperatures, *J. Geophys.. Res.*, **90**, 11587-11601.

McMillin, L. M., and D. S. Crosby, 1984. Theory and validation of the multiple window sea surface temperature technique, *J. Geophys.. Res.*, **89(C3)**, 3655-3661.

NOAA Technical report NESDIS 1989: Non linearity corrections of the thermal infrared channels of the Advanced Very High Resolution Radiometer: assessment and recommendations.

Podesta, G.P., S. Shenoi, J.W.Brown, and R.H. Evans, 1995. AVHRR Pathfinder Oceans Matchup Database 1985-1993 (Version 18), draft technical report of the University of Miami Rosenstiel School of Marine and Atmospheric Science, June 8, 33pp.

Reynolds, R.W. , 1993. Impact of Mt. Pinatubo aerosols on satellite-derived sea surface temperatures, *J. .Climate*, **6**, 768-774.

Reynolds, R. W. and T. S., Smith, 1994. Improved global sea surface temperature analyses, J. *Climate*, **7**, 929-948.

Schluessel, P., W.J. Emery, H. Grassl and T.Mammen, 1990. On the Skin-Bulk Temperature Difference and its Impact on Satellite Remote Sensing of Sea Surface Temperature, *J.Geophys.Res.*, **95**, 13341-13356.

Stowe, L. L., E. P. McClain, R. Carey, P. Pellegrino, G. G. Gutman, P. Davis, C. Long, and S. Hart, 1991. Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data, *Adv. Space Research*, **3**, 51-54.

Walton, C., C., 1988. Nonlinear multichannel algorithms for estimating sea surface temperature with AVHRR satellite data, *J. Appl. Meteor.*, **27,** 115-124.

Wick, G.A. and W. Emery, 1992. A Comprehensive Comparison between Satellite-Measured Skin and Multichannel Sea Surface Temperature, *J. Geophys.. Res.*, **97(C4)**, 5569-5595.

10.0 ACKNOWLEDGMENTS

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. We gratefully acknowledge funding by the Earth Observing System, Data and Information System, NASA Headquarters Code YD, Dr. Martha Maiden, Program Manager. The project is a joint NOAA/NASA program to reprocess a long-time series of sea surface temperature data suitable for global scale ocean studies.

APPENDIX A - PROCESSING DETAILS

*OVERVIEW OF PRODUCT PROCESSING

The generation of the NOAA/NASA AVHRR OCEANS PATHFINDER SST products is a multi-step process. Specific details are outlined in Appendix E, however a brief synopsis is presented here.

Level-1B data are first ingested from optical disk, then converted from to a standard image format. Data are then navigated from line/pixel(image) coordinates to latitude/ longitude coordinates and subsetted. Next the non-linear correction algorithm adjusts for the calibrations of the AVHRR channels, and SST is computed from predetermined regression coefficients which are specific to 3 regimes of water vapor.

Once SST retrievals are determined, the data subsets are binned to produce the 9.28 km equal-area orbitals for both ascending and descending nodes and an initial quality flag is assigned to each retrieval. After the orbitals for an entire day have been completely processed, they are composited to a single daily file.

The next phase is declouding. This is effected through the creation of composite images over 3 weeks, from each of which a mean is computed. These are then used to fill a central weekly mean image which contains the day being declouded (if the central mean image is missing values). If the weekly means from week(n-1) or week(n+1) cannot be used to fill empty values in the central (week n) mean, a spatial interpolation is done. The completely-filled weekly image is then compared to the central daily, and simple thresholding is used to indicate partial or complete cloudiness. This process generates a cloud mask for every day of data.

**INGESTION

Processing of data is done with DSP, a software system supporting oceanographic

satellite and image processing developed by the remote sensing group at the University of Miami/RSMAS. Level-1B data is a sensor level data set which consists of the raw sensor data organized as one scan line per record with quality control information, calibration coefficients for each channel, and earth locations for selected data spots appended to each scan line. The data are read in from optical disk and is converted to standard image format. A standard image format consists of 1200 scan lines of data read in from the medium. This is done with a program called GET_SCAN that locates particular points in each pass. These are the times of the poles crossings for NOAA-9 satellite at +90 (north pole) crossing and -90 (south pole) crossing. The splitting points (north/south crossings), chops the pass up into pieces (1200 scan line files or standard image format). The files become accessible for navigation.

**NAVIGATION/SECTOR

Adjustments and calibrations of the mapping from line/pixel(image) space to latitude/longitude space occurs next. (Refer to DSP U.G.)To determine the actual location of the line/pixel image, time and attitude parameters are corrected using navigation ephemeris files to get a comparable match-up of the coastlines that are visible from the actual image to a reference map outline. This generates a more refined set of parameters most importantly the earth location of the image in latitude and longitude values. Sectors are extracted data specified by a latitude/longitude center taken from the ingested files. The selection of the latitude and longitude center depends on the accuracy of the navigation process. The navigation file is updated for this piece or sector. It uses the same file but modifies the ingested file. For subsequent processing, the volume of data from a sector is far smaller compared to a typical ingested file. This is desirable in terms of efficient processing.

**APPLICATION OF NON-LINEAR CORRECTION ALGORITHM

A non-linear correction algorithm (See Algorithm above) to calculate the Sea Surface Temperature is applied to the modified ingested files.

**BINNING

After the SST retrievals are determined two binning processes occurs next. Since the Global Area Coverage (GAC) data were originally sampled at approximately 4km resolution, data are binned to an approximate 9.28 km bin size averages called spatial binning which gives 5,940,422 bins over the glove. An advantage of this grid is that it can be easily combined into grids with zonal resolutions because the number of bins per row is always an integer. The number of retrievals which were averaged into each bin is a standard product, and can be obtained to correctly perform any special weighted averaging. The spatial binning produces ascending and descending binned data including day splitting.

Time binning begins when files that cover one orbit are spatially binned. These spatial binned files are composited into two files descending and ascending Equal Area Binned Orbitals.

**EQUAL AREA DAILY

When all of the Equal Area Binned Orbitals are completed for a particular day, they are then composited to a single daily file in both ascending and descending called Equal Area Daily that includes 6 channels information..

1) Bin Number

2) Number of Observations per bin

3) Pixel Quality

4) mask bits

5) sum_sst for a given 9km bin

6) sum squared sst for a given 9km bin.

**DECLOUDING

In the declouding process a weekly ascending and descending files are generated from the Equal Area Daily files of the given week. These weekly files will be used to create the weekly reference map. A reference file or reference map is created by Pathtiming or time binning the previous week(n-1) and the next week(n+1). The pathtiming process does a comparison between channel differences of channels 4 and 5 with neighboring pixels (refer to time binner and scenario diagram) in a 21 day moving reference, 7 days before and 7 days after. It

takes the quality flag values and compares them with one another. The best quality value out of the three weeks and its corresponding SST value is assigned to the reference file of week(n). If two of the weeks have the same best quality value then the average value of their corresponding SST is calculated and assigned to the reference file of week(n) along with the quality flag value.

Next the reference file goes through a data filling process. If the pixel contains a quality value of less than a 3 (3 being the best SST) then a linear interpolation is done to a 5x5 pixel array of the particular pixel. Then the new quality value replaces the original value of the particular pixel.

Once the modified reference map is created, the Equal Area Daily Files of week(n) are then cloud-masked by thresholding against the reference map on a pixel-by pixel basis, using separate reference maps for ascending and descending data. The output will be the same daily files in the weekly directory, but with modification dates, reflecting the cloud-masking time.

**REMAPPING

The first output product, Equal Area Declouded Daily 9km File is written to disk and is remapped to an Equal-Angle projection to generated 72 different products for distribution (see product list). After process, the data are remapped into an equal-angle project(rectangular grid), 4096X2048 pixels, in order to facilitate visualization and extraction of regional subsets. The data are stored as Hierarchical Data Format (HDF) files, and as flat binary.

APPENDIX B - CONTACT POINTS, SCIENCE WORKING GROUP AND FTP DIRECTORY

STRUCTURE

Members of the Pathfinder AVHRR Oceans Science Working Group

Name	Affiliation
Peter Cornillon, Chair	University of Rhode Island Graduate School of Oceanography
Robert Evans	University of Miami, Rosenstiel School of Marine and Atmos. Sciences
Gene Feldman	NASA/GSFC
Richard Legeckis	NOAA/NESDIS
Richard Reynolds	NOAA/NWS
Charles Walton	NOAA/NESDIS

For further information on the data set contact the following persons:

Susan Digby

podaac@podaac.jpl.nasa.gov (main point of contact)

M/S 300/323

4800 Oak Grove Dr.

Pasadena, Ca. 91109

818-354-0151

818-393-2718 (fax)

Jorge Vazquez (same as above address)

818-354-6980

(jv@pacific.jpl.nasa.gov)

Andy Van Tran

(andy@grumpy.jpl.nasa.gov)

Rosanna Sumagaysay

(rosanna@bashful.jpl.nasa.gov)

Elizabeth Smith

(lizsmith@ccpo.odu.edu)

Mike Hamilton

(mkh@albedo.com)

General questions about the data sets should be sent to:

podaac@podaac.jpl.nasa.gov

Homepage URL: http://podaac-www.jpl.nasa.gov

FTP site: podaac.jpl.nasa.gov

login: anonymous

password: your complete email address

This will connect you with the podaac.jpl.nasa.gov anonymous ftp server which will contain the following directories specifically for NOAA/NASA AVHRR Oceans Pathfinder products:

The following is the ftp directory structure. The structure itself will change as more data sets are produced. For example as additional years are produced the appropriate subdirectories will be added. Please see following page.

(Directory structure using product numbers)

+pub+product050 -> Same as /pub/archive/avhrr/data/all_pixel/

+pub+product051 -> Same as /pub/archive/avhrr/data/best_sst/

+pub+product053 -> Same as /pub/archive/avhrr/data/half_degree/

+pub+archive+avhrr +documents

+software

+matchups

+landmask

+data +best_sst +hdf_09 +daily +1985 +ascending+<files>

+ + + + +descending+<files>

+ + + +1986 ->Same as 1985

+ + + +1987 ->Same as 1985

+ + + +1988 ->Same as 1985

+ + + +1989 ->Same as 1985

+ + + +1990 -> Same as 1985

+ + + +1991 -> Same as 1985

+ + +monthly -> Same as daily

+ + +weekly -> Same as daily

+ + +yearly -> Same as daily

+ +hdf_18 -> Same as hdf_09

+ +hdf_54 -> Same as hdf_09

+ +raw_09 -> Same as hdf_09

+all_pixel -> Same as best_sst

+half_degree +all_hdf_54 +daily +1985 +ascending +<files>

+ + + +descending +<files>

+ + +1986 -> Same as 1985

+ + +1987 -> Same as 1985

+ + +1988 -> Same as 1985

+ + +1989 -> Same as 1985

+ + +1990 -> Same as 1985

+ + +1991 -> Same as 1985

+ +monthly -> Same as daily

+ +weekly -> Same as daily

+ +yearly -> Same as daily

+best_hdf_54 -> Same as all_hdf_54

APPENDIX C - GRIDDING SCHEME

Introduction

This document describes the equal-area gridding scheme proposed by the RSMAS Remote Sensing Group for the binned sea surface temperature fields produced by the AVHRR Pathfinder project. The same approach is being adopted for SeaWIFS binned ocean color products. The gridding scheme is based on that adopted by the International Satellite Cloud Climatology Project (ISSCP).

This document does not motivate the need for an equal area grid for oceanographic products. Such motivation can be found in a paper by W. Rossow and L. Gardner (Selection of a map grid for data analysis and archival, Journal of Climate and Applied Meteorology, 1984, 23: 1253-1257). Furthermore, this document describes only the design of the proposed equal-area grid, and does not discuss other related topics such as rules for spatially or temporally combining observations into the equal-area bins.

Overview

The gridding scheme proposed consists of rectangular bins or tiles, arranged in zonal rows. A compromise between data processing and storage capabilities, on one side, and the potential geophysical applications of satellite data, on the other side, suggest that a suitable minimum bin size would be approximately 8-10 km on a side.

In the scheme proposed here, the tiles are approximately 9.28 km on a side. This size (9.28 km) was chosen because (a) it has approximately the desired minimum resolution, and (b) it results in 2160 zonal rows of tiles from pole to pole (i.e., 1080 in each hemisphere). This particular number of rows (2160) has some advantages which will be discussed in more detail below. Because the total number of rows is even, the bins will never straddle the Equator (i.e., there will be an equal number of rows above and below the Equator). This avoids possible situations where the Coriolis factor is zero. This is a characteristic that numerical modellers expect from any gridding scheme adopted.

The total number of approximately 9-km bins is 5,940,422. The bins or tiles are arranged in a series of zonal rows; the number of tiles per row varies. The rows immediately above and below the Equator have 4320 tiles. This number is derived by dividing the perimeter of the Earth at the Equator by the standard tile size (i.e., 2[[pi]]Re/9.28), where Re is the equatorial radius of the Earth (Re = 6378.145 km). The number of tiles per row decreases approximately as a cosine function as the rows get closer to each pole (rigorously, there should be an adjustment for ellipticity of the Earth, as the equatorial radius decreases progressively to the smaller polar radius; this adjustment is not applied in the current implementation). At the poles, the number of tiles is always three. This special situation will be discussed in detail below. The number of tiles per row as a function of latitude is shown on Figure 1.

Figure 1. Number of 9.28 km tiles per zonal row as a function of latitude (North or South). The number of tiles is 4320 at the Equator and decreases to 3 at the poles.

Click here for Picture

The number of bins in each zonal row is always an integer. To ensure an integer number of bins, the width of each bin (the size of a bin along a parallel, or x-length) must vary slightly from row to row. However, the bins are always 9.28 km long along the meridians. That is, only one of the bin dimensions changes. The size of the bins at each zonal row is established in the following manner. First, a preliminary value for the number of tiles (Np) at a given latitude (L) is computed as

Np = 2[[pi]]r / X,

where X is the x-size of a bin at the Equator (9.28 km) and r is the radius of the circle produced by slicing the Earth with a plane parallel to the Equator at latitude L. The radius r can be calculated as

r = Req cos(L),

where Req is the equatorial radius of the Earth. If the fractional part of Np is greater or equal than 0.5, then Np is rounded up to the nearest integer (i.e., the final number of tiles will be the integer portion of Np plus one), otherwise Np is rounded down (the final number of tiles is the integer portion of Np. Once the final integer number of tiles along a row is calculated, the X-size of the tiles must be adjusted. This is done by dividing the perimeter of the row (2[[pi]]r) by the integer number of tiles. The result is the x-length of a tile (width) for a given row.

Because the x-length of the tiles is adjusted to ensure an integer number at each row, the "equal area" characteristics of this binning scheme are not rigorously preserved. However, variations in tile size are negligible throughout most of the globe, and only become relevant at very high latitudes, where there are fewer tiles per row and, thus, any adjustments are more noticeable. As soon as the number of tiles increases with distance from the poles, the difference between tile sizes rapidly becomes practically unnoticeable. To provide an idea of the magnitude of the fluctuations in tile size, the worst possible case occurs when half a tile remains "uncovered" after filling a zonal row with an integer number of tiles. Once a row has 100 bins (approximately 16 rows, or 148 km from the poles), the worst possible difference between the actual tile x-length and the standard x-length is of the order of 0.5% (i.e., half a tile's length redistributed among about 100 tiles). For a tile of about 9 km a side, this represents a difference in the x-length of about 45 m. Through a similar calculation, a row with 50 bins (about 80 km away from the poles) has a 1% variation with respect to the standard bin size.

The gridding scheme described here has an extremely useful feature: the number of 9.28 km tiles in each hemisphere (1080) is divisible by many numbers (e.g., 2,3,4,5,6) and therefore it is extremely easy to generate an integer number of rows at many useful spatial resolutions. For instance, 12 rows of 9.28 km tiles can be combined to generate zonal bands of approximately one degree (one degree of latitude is equal to 111.12 km; 12 bins would form a band 111.20 km wide). Another example is the use of 30 rows of to generate zonal bands of approximately 2.5deg. (a typical output resolution of atmospheric circulation models).

The poles

Both the North and South poles are special cases in the gridding scheme presented here. The pole areas are always covered by three tiles, shaped like pie sectors. While the meridional size of the polar bins (the y-length) will be the usual 9.28 km, the length of the bins along the arc of the sectors will be slightly larger. Neglecting sphericity, the area encompassed by the last row of tiles is [[pi]]X2, where X = 9.28 km. If we express the area of the circle as a rectangle of height X, the remaining dimension is [[pi]]X. If we divide the perimeter by three (to yield three tiles), each tile will have dimensions X by [[pi]]X/3 (approximately 1.05X). That is, the bases of the triangular polar tiles are about 5% larger than the x-length of the equatorial tiles.

APPENDIX D - ACRONYMS

AVHRR Advanced Very High Resolution Radiometer

EOS Earth Observing System

FTP File Transfer Protocol

FAQ Frequently Asked Question

GAC Global Area Coverage

HDF Hierarchical Data Format

IMS Information Management System

JPL Jet Propulsion Laboratory

MCSST Multi-Channel Sea Surface Temperature

NASA National Aeronautics and Space Administration

NOAA National Oceanic and Atmospheric Administration

PO.DAAC Physical Oceanography Distributed Active Archive Center

QA Quality Assurance

RSMAS Rosenstiel School of Marine and Atmospheric Sciences

SST Sea Surface Temperature

SWG Science Working Group

WWW World Wide Web

PFMDB Pathfinder Matchup Database