The Past, Present, and Future of the AVHRR Pathfinder SST Program

Kenneth S. Casey, Tess B. Brandon, Peter Cornillon, and Robert Evans

SEDAR-PW6-RD41

6 June 2014

The original publication is available at www.springerlink.com.
(http://www.springerlink.com/content/t2l4275xk08t5m51/)
Casey,K.S., T.B. Brandon, P. Cornillon, and R. Evans (2010).
"The Past, Present and Future of the AVHRR Pathfinder SST
Program", in Oceanography from Space: Revisited, eds. V.
Barale, J.F.R. Gower, and L. Alberotanza, Springer. DOI:
10.1007/978-90-481-8681-5_16.

The Past, Present, and Future of the AVHRR Pathfinder SST Program

Kenneth S. Casey[1], Tess B. Brandon[1], Peter Cornillon[2], and Robert Evans[3]

[1] NOAA National Oceanographic Data Center, Silver Spring, MD, USA
[2] University of Rhode Island Graduate School of Oceanography, Narragansett, RI, USA
[3] University of Miami Rosenstiel School of Marine and Atmospheric Science, Miami, FL, USA

Abstract. With origins dating back to 1990, the Advanced Very High Resolution Radiometer (AVHRR) Pathfinder Sea Surface Temperature (SST) Program has experienced a 20-year history of reprocessing space-based observations to create accurate, consistent, climate data records. Both scientific and programmatic aspects of this history are reviewed and summarized in this paper, along with a review of the currently available Pathfinder SST data. In addition, a look forward to the next generation of Pathfinder currently under development is presented.

1. Introduction

With today's modern satellite sensors, many oceanic parameters are capable of being observed from space. Of these, only sea surface temperature (SST) from the Advanced Very High Resolution Radiometer (AVHRR) series on board the NOAA polar-orbiting satellites enjoys an unbroken, nearly 30-year long history of observation from the same class of instrument. These AVHRR observations date back to 1981, when the first five-channel AVHRR flew on board the NOAA-7 platform. The AVHRR series continues to fly today, and will continue perhaps another ten years into the future on board the remaining NOAA polar orbiting satellites and the European METOP platforms.

Early in the life of this series of instruments, NASA Headquarters conceived the Pathfinder program to support the needs of the U.S. Global Change Research Program (King and Greenstone, 1999). These Pathfinder efforts were to provide unprecedented access to large volumes of consistently processed satellite datasets, in advance of the launch of the

V. Barale, J.F.R. Gower and L. Alberotanza (eds.), *Oceanography from Space, revisited*.
© Springer Science+Business Media B.V. 2010

NASA Earth Observing System (EOS) satellites. Climate-quality datasets would be generated and made available, providing valuable experience in reprocessing, archiving, and distributing large satellite data sets. Thus, these Pathfinder efforts would "find the path" forward in preparation for the ambitious EOS series of instruments and missions that were to follow.

In October of 1990, NOAA and NASA signed a Cooperative Agreement to commence work on three joint NOAA/NASA Pathfinder projects. These joint projects would use existing NOAA datasets, focusing on AVHRR Global Area Coverage (GAC) data, the Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) data, and the Geostationary Operational Environmental Satellite (GOES) data. Several months later, in 1991, a project involving Special Sensor Microwave/Imager (SSM/I) data, which was archived by NOAA under an agreement with the U.S. Navy, became the fourth NOAA/NASA Pathfinder. In 1992 a NASA/EPA/USGS Landsat Pathfinder project was formed, as was the first NASA-only Pathfinder, using Scanning Multichannel Microwave Radiometer (SMMR) data. Funding for these activities came mainly from NASA.

Each of the Pathfinder projects would form a Science Working Group (SWG) that would determine the specific products to be created based on scientific need, identify community-consensus algorithms to generate those products, and make recommendations on product validation, storage, maintenance, and required data services. A NASA Distributed Active Archive Center (DAAC) would then make the data available, as would the EOS Data and Information System (EOSDIS).

Because the AVHRR data stream was perceived to be of value to several different areas in the Earth sciences, three product science working groups were formed under the AVHRR SWG umbrella: an atmosphere SWG, a land SWG and an oceans SWG (OSWG in the following). The official goals of the AVHRR Pathfinder projects, as presented on 08 January 1992 by Mary James of NASA Goddard Space Flight Center to a group that included the project chairpersons, were:

- To produce long term global datasets for research, modeling, and trend analysis. Datasets were to be consistently processed with the best available community consensus algorithms.

- To produce multiple geophysical products from a common input stream using integrated processing concepts.

- To assemble a consistent, low maintenance, readily accessible archive for AVHRR data with browse and on-line access capabilities.

In December of 1994 a Pathfinder NASA Research Announcement was released, marking the transition from the first phase to a second phase of the Pathfinder datasets. The NRA emphasized building long-term datasets and addressing consistency issues that arise when linking together multiple sensors over time. This second phase ended in 2001, which marked the formal end of NASA support for the NOAA/NASA Pathfinder efforts. However, because of the successes obtained and momentum built over the previous decade, some of these efforts continued operating under different mechanisms. For example, the AVHRR Pathfinder Atmosphere (PATMOS) project continued as a NOAA National Environmental Satellite Data and Information Service (NESDIS) effort (Stowe *et al*., 2002).

The OSWG met for the first time in February of 1991. As with the PATMOS project, the Pathfinder SST activities also continued with the goal of building on the successes of the previous decade. In the case of Pathfinder SST, the program continued through the efforts of the NOAA/NESDIS National Oceanographic Data Center (NODC), beginning in 2002. Details on the history, current status, and future directions of Pathfinder SSTs constitute the remainder of this paper.

2. History of the Pathfinder SST Program

Reprocessing of AVHRR data to create long, accurate, and consistent SST records was one component of the overall NOAA/NASA Pathfinder Program. To remain consistent with terminology, the activities that took place between 1990 and 2001 are referred to as the "NOAA/NASA AVHRR Oceans Pathfinder SST Project." During that period, NASA was the primary sponsor of the effort. Beginning in 2002 NODC became involved and a variety of NOAA funding sources began supporting the efforts. Beginning in 2009, new NASA funding supported some specific aspects of the program. The overall set of activities, including both the early 1990-2000 and the 2002 and beyond periods is termed simply the "Pathfinder SST Program" (see Table 1).

The AVHRR Ocean Pathfinder SST Project was an outgrowth of the work of the NASA funded SST Archive Science Working Group (SASWG) formed in early 1987. The objective of this group was "to determine the needs of the scientific community for SST fields and the possibility of meeting these needs with existing sources of data." The SASWG met several times over the next two years and on 1 June 1989 delivered its final report detailing two useful classes of SST products and the steps necessary to produce these fields. One class of products, with fine spatial resolution and high relative accuracy, would address the needs of feature-related studies and the other, with high absolute accuracy, would address the needs of heat flux related studies. The SASWG final report identified the AVHRR GAC data stream as appropriate for the generation of both SST products outlined in the report.

Table 1. Summary of Overall Pathfinder SST Program Periods.

The Pathfinder SST Program		
NOAA/NASA AVHRR Oceans Pathfinder SST Project (Versions 1-4)	NOAA-supported Pathfinder SST Program (Version 5)	Version 6 (in preparation)
1990-2001	2002-2008	2009-2010

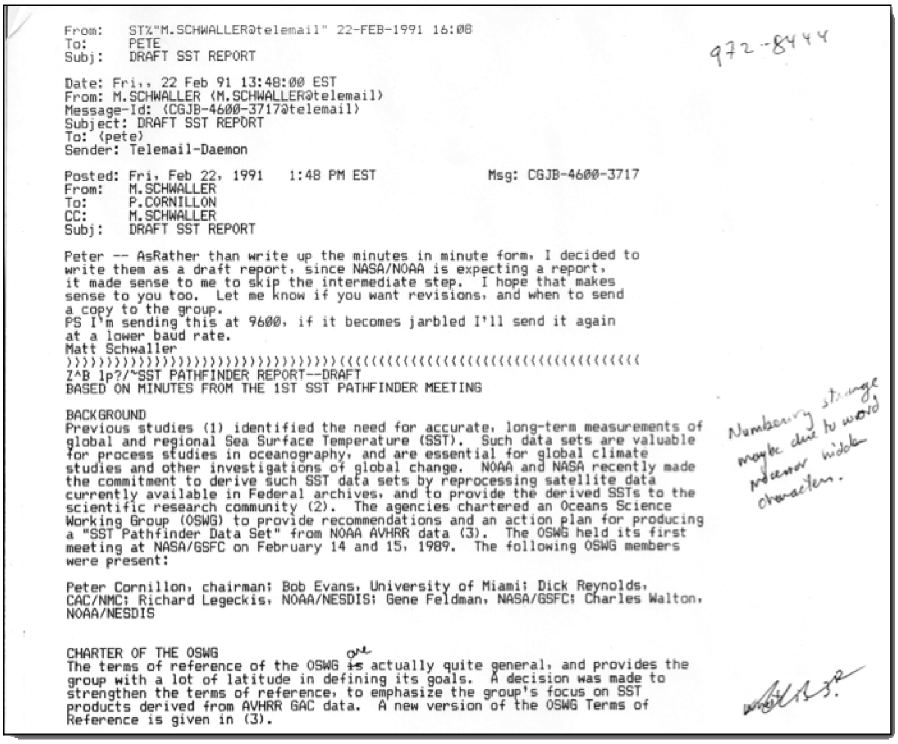
Following the SASWG report, Stan Wilson (then of NASA Headquarters but currently with NOAA/NESDIS) assembled a subgroup of the SASWG plus several others within NASA with expertise in the processing of GAC data to explore the actual production of the desired fields. This group included a representative from the terrestrial science community in addition to those from the oceanographic community. Two observations emerged from this meeting: first, much of the required GAC data existed only at NOAA so a collaboration with NOAA would be required as part of any reprocessing effort and second, other Earth science communities could benefit from ready access to a complete global GAC data set. Based on these observations and the desire to learn how to handle large, multidisciplinary, open-ended data sets for the upcoming EOSDIS, Stan Wilson initiated a data transcription effort and a Cooperative Agreement with NOAA. The Cooperative Agreement, entitled “Early EOSDIS Pathfinder Data Set Activity”, was signed on 15 October 1990. This agreement was the formal start of the Pathfinder Program, as described in Section 1.

However, despite the Cooperative Agreement, the available funding, and the desire of the NOAA staff (Bud Booth and Levin Lauriston) to support the nascent Pathfinder group with the data transcription, little progress was made over the next 13 months. To determine the nature of the problem, in November 1990 Peter Cornillon and Otis Brown met with John Knauss, who was the NOAA Administrator at that time. In that meeting it became clear that the delays resulted form a misunderstanding within NOAA of a comment made by Knauss. When shown what NASA was doing with satellite data, Knauss made a statement to the effect, “Why can’t NOAA do these things?” This statement was misinterpreted by NOAA staff to mean that the data were not to be provided to others and that NOAA was to undertake the processing steps. Knauss had not intended his statement to be interpreted this way and clarified the situation with his staff after the meeting with Cornillon and Brown. The data transcription began in earnest and progress continued steadily on the Pathfinder Program from that point forward.

The Pathfinder OSWG first met in February of 1991 at Goddard Space Flight Center in Greenbelt, Maryland, USA. Peter Cornillon of the University of Rhode Island chaired the OSWG, which included members from NOAA, NASA, and universities (see Appendix). The OSWG met approximately annually over the next four years continually evolving and improving the algorithms and reprocessing techniques. During this same period, the chairmen of the Pathfinder groups met approximately annually to coordinate the work of the different Pathfinder projects and subgroups met when needed to deal with data transcription and calibration issues. To give a sense of the technical challenges faced by the OSWG to simply manipulate the large volumes of data, Figure 1 shows the first part of an e-mail sent to Cornillon on 22 February 1991 containing a draft report from the first OSWG meeting. Note the sender’s concern that the text might be garbled because he had sent it at 9600 baud, which was considered to be very fast at the time. Much of the information emerging from the various groups was distributed by fax or over the Internet at extremely slow data rates by today’s standards.

Sea surface temperature retrieval algorithms proposed by the OSWG were tested and implemented under the direction of Robert Evans at the University of Miami where the data were processed from Level 0 (L0)/Level 1b (L1b) through to Level 3 (L3) equal-area datasets on an integerized sinusoidal grid. These data were then sent to the NASA/Caltech JPL Physical Oceanography Distributed Active Archive Center (PO.DAAC) where they were converted to equal-angle, regularly gridded datasets in HDF4-Raster format and provided to users. The first four versions of Pathfinder SST were created at a reduced resolution of 9.28 km, due to computational resource limitations of the 1990s.

NASA funding for the initial effort ended in 2001-2002 and in 2002 NODC reinvigorated the Pathfinder SST activities with the goal of producing a Version 5 dataset at the full GAC resolution. Modest levels of NOAA funding from various sources including NODC base resources and the NOAA Coral Reef Conservation Program kept the Pathfinder SST Program alive from 2002 through 2006. During this period, the Version 5 data were produced at the University of Miami and then sent to NODC for final quality assurance, archiving, and provision to users. NODC provided the Version 5 data to the PO.DAAC as well, to maintain continuity for users of earlier versions of the Pathfinder SST data who had grown accustomed to accessing it from the PO.DAAC systems. From 2007 through 2009, the NOAA Scientific Data Stewardship Program supported efforts to transition the reprocessing from the University of Miami to NODC, where it could be supported over the long term.



```
From:   ST%"M.SCHWALLER@telemail" 22-FEB-1991 16:08
To:     PETE
Subj:   DRAFT SST REPORT

Date: Fri., 22 Feb 91 13:48:00 EST
From: M.SCHWALLER (M.SCHWALLER@telemail)
Message-Id: (CGJB-4600-3717@telemail)
Subject: DRAFT SST REPORT
To: (pete)
Sender: Telemail-Daemon

Posted: Fri, Feb 22, 1991   1:48 PM EST              Msg: CGJB-4600-3717
From:   M.SCHWALLER
To:     P.CORNILLON
CC:     M.SCHWALLER
Subj:   DRAFT SST REPORT

Peter -- AsRather than write up the minutes in minute form, I decided to
write them as a draft report, since NASA/NOAA is expecting a report,
it made sense to me to skip the intermediate step.  I hope that makes
sense to you too.  Let me know if you want revisions, and when to send
a copy to the group.
PS I'm sending this at 9600, if it becomes jarbled I'll send it again
at a lower baud rate.
Matt Schwaller
))))))))))))))))))))))))))))))))))((((((((((((((((((((((((((((((((((((
Z^B 1p?/~SST PATHFINDER REPORT--DRAFT
BASED ON MINUTES FROM THE 1ST SST PATHFINDER MEETING

BACKGROUND
Previous studies (1) identified the need for accurate, long-term measurements of
global and regional Sea Surface Temperature (SST).  Such data sets are valuable
for process studies in oceanography, and are essential for global climate
studies and other investigations of global change.  NOAA and NASA recently made
the commitment to derive such SST data sets by reprocessing satellite data
currently available in Federal archives, and to provide the derived SSTs to the
scientific research community (2).  The agencies chartered an Oceans Science
Working Group (OSWG) to provide recommendations and an action plan for producing
a "SST Pathfinder Data Set" from NOAA AVHRR data (3).  The OSWG held its first
meeting at NASA/GSFC on February 14 and 15, 1989.  The following OSWG members
were present:

Peter Cornillon, chairman; Bob Evans, University of Miami; Dick Reynolds,
CAC/NMC; Richard Legeckis, NOAA/NESDIS; Gene Feldman, NASA/GSFC; Charles Walton,
NOAA/NESDIS

CHARTER OF THE OSWG
The terms of reference of the OSWG is actually quite general, and provides the
group with a lot of latitude in defining its goals.  A decision was made to
strengthen the terms of reference, to emphasize the group's focus on SST
products derived from AVHRR GAC data.  A new version of the OSWG Terms of
Reference is given in (3).
```

Fig. 1: Excerpt from a 22 February 1991 email to Peter Cornillon, demonstrating the slow data transmission rates available at the time. Despite the sender's concerns, the transmission was not garbled.

Beginning in 2009, new NASA funding augmented the continued NOAA support, with the goal of adding two new components to the Pathfinder SST Program. The first addition was the ability to generate error estimates, a critical component necessary to make the future Pathfinder Version 6 data compliant with the standards of the Group for High Resolution SST (GHRSST, Donlon *et al*., 2007). The second addition was to enable Pathfinder reprocessing of not just AVHRR GAC data, but also a large collection of 1 km resolution Local Area Coverage (LAC) and High Resolution Picture Transmission (HRPT) data from the AVHRR instruments. By the end of 2010, the GAC reprocessing system is expected to be functional at NODC and the LAC/HRPT at the University of Rhode Island. Long-term stewardship, archiving, and provision of the Version 6 data to users will take place at NODC.

3. Earlier Pathfinder SST Versions

Over its twenty-year history, the Pathfinder SST Program has produced five distinct versions and is actively developing the sixth. The Pathfinder SST algorithm is based on the Non-Linear SST (NLSST) algorithm (Walton *et al*., 1998) and was originally chosen by the SWG because of the algorithm's adequate performance, operational nature, and its acceptance by the scientific user community (Kilpatrick *et al*., 2001). The basic form of the algorithm is

$$SST_{sat} = a + bT_4 + c(T_4 - T_5)SST_{guess} + d(T_4 - T_5)[\sec(\theta) - 1], \quad (1)$$

where

SST_{sat} = the satellite-derived SST estimate,
T_4 and T_5 = brightness temperatures in the 10.8 and 11.4 μm AVHRR bands (channels 4 and 5, respectively),
SST_{guess} = a first-guess SST value,
θ = the satellite zenith angle, and
a, b, c, and d = coefficients estimated from regression between collocated and coincident *in situ* and satellite measurements.

Numerous modifications were made over time to improve the performance of the algorithm. A review of the algorithm evolution from Version 1 through 4 is provided below, based on information provided in an early version of the Pathfinder Users Manual (Vazquez *et al*., 1995) and Casey and Cornillon (1999). For more details, with an emphasis on Pathfinder Version 4, see Kilpatrick *et al*. (2001). Version 5 and Version 6 are discussed in more detail in the following sections.

Version 1

The first version of Pathfinder implemented several key enhancements to the NLSST algorithm. The form of the NLSST equation was slightly modified to include a time dependent term. Extra care was taken with the first steps in which the digital counts are converted to radiance and brightness temperatures. The first of these steps involves the linear transformation of counts to radiance based on the space-view and sensor base plate onboard calibration information. Next, a non-linear correction factor was applied based on pre-launch calibration data. Finally, lookup tables based on the sensor's operating temperature were used to convert the channel 4 and 5 radiances to brightness temperatures.

Pathfinder Version 1 also included improvements to the navigation of the satellite observations over previous AVHRR SST data products. These improvements focused on the clock drift and spacecraft and sensor attitude. Clock drifts caused uncertainties in the along-track position and were corrected using a database of satellite clock time and Earth time offsets determined by comparing precise time measurements recorded at the University of Miami Domestic Communication Satellite (DOMSAT, the communication satellite NOAA used to transfer AVHRR data) station with the time information in HRPT passes received at Wallops Island, VA, and Fairbanks, AK. Averaging over multiple geographic locations and times along the orbital track mitigated attitude errors.

Further improvements introduced in Pathfinder Version 1 were implemented during the next phase of reprocessing, in which SST values are calculated. Three sets of algorithm coefficients (*a, b, c,* and *d* in Eq. 1) were generated using the satellite-*in situ* matchup database: one for high atmospheric water vapor regimes where the $T_4 - T_5$ brightness temperature difference is greater than or equal to 1.8 ºC, one for low water vapor regimes where the $T_4 - T_5$ brightness temperature difference is less than 0.7 ºC, and one for moderate levels of water vapor where the $T_4 - T_5$ brightness temperature difference is greater than or equal to 0.7 ºC and less than 1.8 ºC. In Version 1, the coefficients were calculated on an annual basis, using 12 months of matchup data.

Finally, an extensive set of quality tests were implemented in Version 1, including checks for gross cloud contamination, uniformity tests, satellite zenith angle tests, stray sunlight tests, and a check against a reference SST field based on the Reynolds Optimally Interpolated SST (OISSTv1; Reynolds and Smith, 1994). Pixels on the edges of a scan line or in the first or last scan line of an orbital piece were excluded, and those subject to sun glint were also identified. Based on the combined results of these many tests, Version 1 Pathfinder SSTs were then assigned a quality level of between 0 (worst) and 4 (best). Pathfinder Version 1 data spanning 1987 to 1993 were released to the public beginning in 1995.

Version 2

A second version of Pathfinder was developed to better correct for temporal changes observed in the match-up statistics. The time-dependent term was removed and coefficients were generated on a monthly instead of annual basis. The coefficients were calculated using a temporally weighted five-month running window and were found to better correct the temporal changes than the time-dependent term used in the NLSST equation in Version 1. The central month is given a 100% weight, the adjacent months an 80% weight, and the months at the ends of the window a 50% weight. This version of the Pathfinder data set was never released to the public.

Version 3

In the third version of Pathfinder SST, two sets of algorithm coefficients (*a, b, c,* and *d* in Eq. 1) were generated using the satellite-*in situ* matchup database instead of the three sets used in Version 1. The two sets were found to better reduce the overall bias between satellite and buoy SST. One set was for high atmospheric water vapor regimes where the $T_4 - T_5$ brightness temperature difference is greater than or equal to 0.7 ºC and the other was for low water vapor regimes where the $T_4 - T_5$ brightness temperature difference is less than 0.7 ºC. To avoid discontinuities, in cases where the $T_4 - T_5$ brightness temperature difference was between 0.5 ºC and 0.9 ºC the SST was calculated using a weighted combination of SST values determined using both sets of coefficients. Pathfinder Version 3 data were first released to the public in 1997 and spanned only 1991-1994. They were quickly superseded by Pathfinder Version 4, which was released only a short time later.

Version 4

In Version 4, a more robust algorithm coefficient scheme was introduced, and decision trees were implemented to better eliminate cloudcontaminated pixels. Quality levels were expanded to 0 to 7 to allow for greater user flexibility in determining trade-offs between data coverage and data quality. Version 4 data became available in 1997 and eventually covered 1985 to 2003. The Version 4 data marked the culmination of the NOAA/NASA AVHRR Oceans Pathfinder period.

4. Current State of Pathfinder SST: Version 5

Beginning in 2002, the Pathfinder SST Program entered a new phase targeted at producing an even more accurate, consistent, and finer resolution SST data set. By this time, the term "Climate Data Record" had emerged and the Pathfinder SST Program renewed its efforts toward delivering the SST climate data record. Improvements in spatial resolution, the land mask, and the way sea ice information was used in the quality level determinations were implemented. The 1-degree resolution OISSTv1 first-guess and reference SST field used in earlier Pathfinder versions was replaced with a newer version of the product, the OISSTv2 (Reynolds *et al*., 2002) For the earliest part of the AVHRR record from late 1981 through 1984, a new 25 km Daily Optimally Interpolated SST (DOISST, Reynolds *et al*., 2007) was used. A more quantitative analysis of the improvements achieved along with a description and evaluation of new SST climatologies derived from Pathfinder Version 5 is available in Casey *et al*. (2009).

By 2009, production, distribution, and long-term stewardship of the Pathfinder Version 5 SST data had become a joint University of Miami and NODC effort. The University of Miami processed the data from L0/L1b to L3, including a transformation from the integerized sinusoidal grid to the regularly gridded data in HDF4-Scientific Data Set format. These data were then transferred to NODC, where they were subjected to additional quality assurance involving browse image generation, visual inspection of every image, and comparison against the HadSST2 (Rayner *et al*., 2006) *in situ* dataset. Any problems identified were then analyzed, and corrections implemented at the University of Miami where the data were regenerated before re-inspection at NODC. Once cleared through the quality assurance steps, the data were then archived at NODC and provided to users through its http, ftp, and OPeNDAP services. Access to the entire Pathfinder Version 5 collection is available on line[1]. An example image, generated from a climatological mean of Pathfinder Version 5 SST, is shown in Figure 2.

---

[1] at http://pathfinder.nodc.noaa.gov

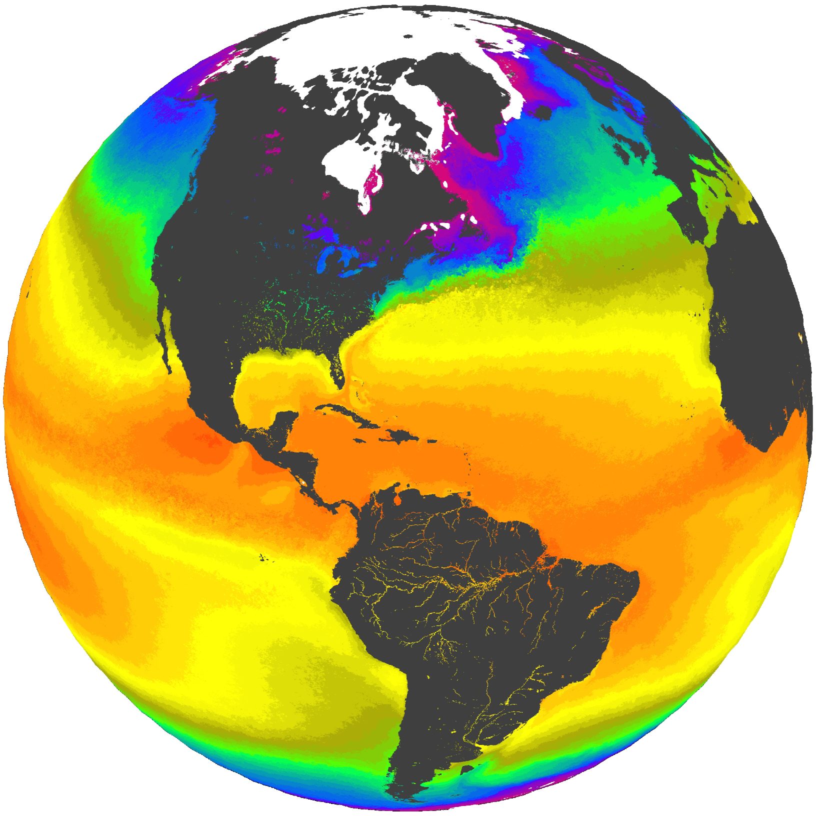


Fig. 2. Climatological mean SST for Week 50, based on Pathfinder Version 5 data.

The following steps provide the details on the current state of the Pathfinder processing system. Boldface items highlight the data and information provided to the processing system.

Step 1: Ingestion, calibration, and navigation of GAC data

1.1. Calibrate and convert AVHRR digital counts for channels 1 through 5 to radiances

1.1.1. Obtain **AVHRR channels 1 through 5 radiometer count data** from NOAA and University of Miami collections of GAC data.

1.1.2. For channels 1 and 2, use **pre-launch calibration coefficients** to perform a linear counts-to-radiance conversion, followed by a correction for temporal changes using **sensor decay rate data** and then a correction for inter-satellite differences using **inter-satellite standardization data to a NOAA-9 reference**, both of which use Libyan desert target area data.
1.1.3. For channels 3, 4, and 5 use both the above pre-launch calibration data and **onboard blackbody (space view and sensor base plate) data** to perform a non-linear counts-to-radiance conversion.

1.2. Navigation, Clock, and Attitude Corrections

1.2.1. Apply satellite clock corrections using **Earth time offset data** based on University of Miami DOMSAT recorded HRPT files.
1.2.2. Apply attitude corrections made using **coastline comparison data**.
1.2.3. At this point, navigated, calibrated albedos/brightness temperatures are available for all five channels. Note that channels 1 and 2 are used in the Pathfinder Matchup Database decision trees [see Kilpatrick *et al*. (2001), Figure 7] and channel 3 is used only in assignment of a quality indicator (see step 2.4.1).

Step 2: SST Calculation

2.1. Convert channel 4 and 5 brightness temperatures to SST in ºC using the Pathfinder algorithm (Equation 1), which requires a set of **monthly coefficients**.
2.2. These coefficients are derived using the **Pathfinder Buoy Matchup Database,** a set of *in situ* SST observations and collocated AVHRR data. The *in situ* data consist mainly of drifting buoys, but during 1981-1984 bias-corrected ship-based observations are also used.
2.3. The calculation of SST in step 2.1 also requires a first-guess SST field. This first-guess field is the **OISSTv2** product for 1985-2009 and the **DOISST** for 1981-1984. Note that Versions 1-4 of Pathfinder used OISSTv1.
2.4. Quality Flag Assignment
    2.4.1. A Channel 3, 4, and 5 brightness temperature test is performed. The brightness temperatures were calculated in step 1.1.3.
    2.4.2. The viewing angle is then evaluated using a **satellite zenith angle** check.

2.4.3. Next, a reference field comparison check is made against the OISSTv2/DOISST used in step 2.3.
2.4.4. A stray sunlight test is then performed which requires information on whether the data in question are to left or right of nadir.
2.4.5. An edge test is performed next, which checks the location of the pixel within a scan line and the location of the scan line within the processing piece (a "piece" is a subset of an entire orbit file).
2.4.6. Then, a glint test is performed which requires a **glint index** calculated according to the Cox and Munk (1954) formulation.
2.4.7. Finally, these steps are all combined into an overall quality flag assignment for each pixel.

Step 3: Spatial Binning

3.1. The GAC pixels, converted to SST in Step 2, are then binned into an equal-area integerized sinusoidal grid.
3.2. To reduce discontinuities along the date line, a data-day based on a spatial data-day definition (Podestá, 1995) is used.
3.3. A **land mask** is applied to the data, identifying pixels that fall on land. In Pathfinder Version 5, an improved mask based on a 1 km resolution Moderate Resolution Imaging Spectroradiometer (MODIS) dataset (MOD12Q1) derived by the USGS Land Processes Distributed Active Archive Center is used[2]. In earlier versions of Pathfinder, the land mask was based on a Central Intelligence Agency database of navigation hazards, but no citation or further information is available.

Step 4: Temporal Binning

4.1. Step 4 begins by temporally binning the spatially binned pieces from Step 3 into a single daytime and single nighttime file for each day. In the case of overlapping satellite passes, only the highest quality pixels available for each are used. If there is more than one contributing pixel at the highest quality, these values are averaged. In Pathfinder Version 5, temporally binned files are also created for 5-day, 7-day, 8-day, monthly, and yearly periods.

[2] see http://www-modis.bu.edu/landcover/userguidelc/lc.html for more info

4.2. An internal Pathfinder reference check comparison is then made to an internal 3-week Pathfinder comparison field. A **sea ice mask** based on **weekly SSM/I and sea ice information in the OISSTv2** is used to exclude pixels from the computation of the internal reference field. Note that sea ice information is not used in this way in the earlier versions of Pathfinder.
4.3. The SST, quality, and related fields are reformatted from equal-area to equal-angle for distribution and archiving in HDF4-Scientific Data Set format. Older Pathfinder versions were distributed in HDF4 Raster format.

Step 5: Quality Assurance, Archiving, and Distribution

5.1. Utilizing checksums to ensure file integrity, the HDF4-SDS files are then acquired by NODC from the University of Miami and browse graphics are generated.
5.2. Each browse graphic is then visually inspected and problematic data are sent back to the University of Miami for correction and reprocessing.
5.3. After passing the visual inspection, an *in situ* comparison against the HadSST2 data set (Rayner *et al*., 2006) is conducted. Again, any errors detected are investigated and, if the data are subjectively confirmed to be bad, they are returned to the University of Miami for correction and reprocessing.
5.4. After passing the visual inspection and *in situ* comparisons, the Pathfinder data are formally archived by NODC.
5.5. The result of these steps is the Pathfinder Version 5 SST product. The data are distributed by NODC[3] and made available to the PO.DAAC for redistribution.

Pathfinder Version 5 data were first released by NODC to the public in April of 2003. By December of 2003, data for 1985-2001 were available. Over the next several years, additional data were added incrementally. In April of 2009, Pathfinder data for 1981-1984 were made available for first time in the Pathfinder SST Program history. In prior versions, the lack of adequate *in situ* matchups and suitable first-guess SST field made production of those early data from the NOAA-7 platform impossible. However, these limitations were overcome through the use of not just drifting buoys in the matchup database, but also by including bias-corrected ship observations. In addition, working iteratively, the DOISST, which itself uses

[3] http://pathfinder.nodc.noaa.gov

Pathfinder SST data, was extended to include 1981-1984. The DOISST then served as the first guess and reference SST field for this period. By 2010, Pathfinder Version 5 data were available for 1981-2009.

5. Future Directions in Pathfinder SST: Version 6

Despite the improvements implemented over the years, problems still remain in Pathfinder Version 5. These errors are being corrected, and implemented in a new, Version 6 of Pathfinder. Version 6 will improve upon Version 5 in several notable ways. First, known errors in the land mask will be corrected, and Version 6 will utilize the DOISST throughout the record. Use of the finer resolution DOISST is resulting in improved ability of the quality procedures to properly identify cloudy pixels and to retain good pixels in the vicinity of coasts and strong gradient regions. Sea ice information will also be used to a greater extent in Version 6, and will be provided as part of the data set. Many spatial and seasonally varying biases evident in the Pathfinder Version 5 comparisons with HadSST2 will also be minimized through the use of new coefficients generated using a latitudinal band scheme.

Another problem in Version 5 manifests itself as a jump in SST every 18 lines in the regularly gridded data. While not clearly visible in a single image, edge detection algorithms that examine sequences of data identified this problem. The problem is related to the mapping procedures by which the equal-area integerized sinusoidal bins in which the data are generated are mapped into the equal-angle, uniform grid space in which the data are distributed. This problem has been resolved through the use of a refined mapping procedure that will be used in Version 6.

Significantly, Version 6 will also conform to the newest GHRSST data format, data content, and metadata requirements currently being developed, known as the GHRSST Data Specification Version 2 (GDS v2). The GDS requires each SST pixel value to have an associated bias and uncertainty error. These error estimates are being developed for Pathfinder Version 6 using an error hypercube approach which partitions the matchup database into a multi-dimensional lookup table. Pathfinder Version 6 will also be in netCDF and will contain Climate and Forecast metadata attributes. For the first time, Pathfinder SST data will be made available in not just collated L3 files, but also at the L2 swath and uncollated L3 processing levels.

Another addition to Version 6 will be to include a large collection of HRPT and LAC data collected at stations around the world. These full-resolution, 1 km data sets will provide key inputs to ultra-high resolution analysis systems and other fine scale applications. HRPT passes covering the east and west coasts of the United States, the western Pacific off of Japan, and the waters around Australia, have already been identified for inclusion into the Pathfinder Version 6 processing system and other HRPT station data are being actively sought.

6. Conclusion

For 20 years, the ongoing efforts of the AVHRR Pathfinder SST Program have contributed, and continue to contribute to a wide range of marine applications. To date, over 65,000 unique users have accessed AVHRR Pathfinder Version 5 SST data and applied them to a diverse set of applications and research topics. While it is difficult to accurately quantify the use of Pathfinder SST data in the scientific literature due to varying ways in which the data have been acknowledged or cited, searches of ScienceDirect and the Web of Science indicate that number to be in the hundreds and perhaps thousands. These studies include climate research areas such as SST trend analysis (*e.g.*, Casey and Cornillon, 2001), numerical modelling applications (*e.g.*, Shu *et al.*, 2009), and the use of Pathfinder data as the basis for higher-level SST products (*e.g.*, Reynolds *et al.*, 2007; Marullo *et al.*, 2007). A wide range of ecosystem-related studies have also relied on Pathfinder SST data (*e.g.*, Somoza *et al.*, 2008; Halpern *et al.*, 2008).

Many other scientific studies, too numerous to summarize, have utilized Pathfinder SST across a wide range of disciplines and topics. Direct applications of the Pathfinder data to various societal benefits are also numerous. For example, Pathfinder data have been used to characterize humpback whale distributions to reduce ship collisions, improve marine protected area design and placement, plan optimal routing for yacht races, evaluate performance of ocean and numerical weather prediction models, and establish baselines for integrated ecosystem assessments and marine spatial planning projects.

The improvements coming in Version 6 will enable an even greater range of science and applications, enhance the compatibility of Pathfinder data with numerous other GHRSST-compliant data streams, and will result in a more accurate, consistent, and useful climate data record for SST.

Acknowledgements

The authors wish to thank their respective institutions for support during the preparation of this manuscript and to Jorge Vazquez and Ed Armstrong of the NASA Physical Oceanography DAAC for their long-term efforts to support the Pathfinder SST community of users. Thanks also go to the many NASA and NOAA program and project managers who kept the AVHRR Pathfinder SST Program alive over the course of the last two decades, including Eric Lindstrom and John Bates. Finally, special thanks go to the last three Directors of the NOAA National Oceanographic Data Center, most recently Margarita Conkright Gregg and before her, Zdenka Willis and Lee Dantzler. All three deemed the Pathfinder effort sufficiently worthy to contribute significant NODC resources over the years. Without these contributions, the history of Pathfinder SST would have come to an end circa 2001 and would have fallen well short of making the most of the available AVHRR record.

References

Casey KS, Cornillon P (1999) A comparison of satellite and *in situ* based sea surface temperature climatologies. J Climate 12:1848-1863

Casey KS, Cornillon P (2001) Global and regional sea surface temperature trends. J Geophys Res 14(18):3801-3818

Casey KS, Evans R, Kilpatrick K, Kearns EJ, Brandon TB, Reynolds RW, Vazquez-Cuervo J (2009) Toward the sea surface temperature climate data record from space: An improved Advanced Very High Resolution Radiometer Pathfinder dataset. In preparation for J Climate

Cox C, Munk W (1954) Measurement of the roughness of the Sea surface from photographs of the sun's glitter. J Optical Soc Amer 44(11):838-850

Donlon C, Robinson I, Casey KS, Vazquez-Cuervo J, Armstrong E, Arino O, Gentemann C, May D, LeBorgne P, Piollé JF, Barton I, Beggs H, Poulter DJS, Merchant CJ, Bingham A, Heinz S, Harris A, Wick G, Emery B, Minnett P, Evans R, Llewellyn-Jones D, Mutlow C, Reynolds RW, Kawamura H, Rayner N (2007) The Global Ocean Data Assimilation Experiment High-resolution Sea Surface Temperature Pilot Project. Bull Amer Meteorol Soc 88(8):1197-1213 doi:10.1175/BAMS-88-8-1197

Halpern BS, Walbridge S, Selkoe KA, Kappel CV, Micheli F, D'Agrosa C, Bruno JF, Casey KS, Ebert C, Fox HE, Fujita R, Heinemann D, Lenihan HS, Madin EMP, Perry MT, Selig ER, Spalding M, Steneck R, Watson R (2008) A global map of human impact on marine ecosystems. Science 319(5865):948-952 doi: 10.1126/science.1149345

Kilpatrick KA, Podestá GP, Evans R (2001) Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. J Geophys Res 106(C5):9179-9197

King MD, Greenstone R, eds (1999) Earth Observing System (EOS) Reference Handbook. National Aeronautics and Space Administration, EOS Project Science Office, 355 pp

Marullo S, Nardelli BB, Guarracino M, Santoleri R (2007). "Observing the Mediterranean Sea from space: 21 years of Pathfinder-AVHRR sea surface temperatures (1985 to 2005): re-analysis and validation. Ocean Science 2(2):299-310

Podestá GP (1995) Case studies for SeaWiFS calibration and validation, part 3, SeaWifs global fields: What's in a day? NASA Technical Memorandum, no. 10456, 27 pp

Rayner NA, Brohan P, Parker DE, Folland CF, Kennedy JJ, Vanicek M, Ansell T, Tett SFB (2006) Improved analyses of changes and uncertainties in sea surface temperature measured *in situ* since the mid-nineteenth century: the HadSST2 data set. J ClImate 19:446-469

Reynolds RW, Smith TM (1994) Improved global sea surface temperature analysis using optimum interpolation. J Climate 7:929–948

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved *in situ* and satellite SST analysis for climate. J Climate 15:1609-1625

Reynolds, R.W., T.M. Smith, C. Liu, D.B. Chelton, K.S. Casey, and M.G. Schlax (2007) Daily high-resolution-blended analyses for sea surface temperature." J Climate 20(22):5473-5496

Somoza, RD, Kampel M, Rudorff FD, Sousa RB, Cobas S (2008) Lobster (Panulirus argus) captures and their relation with environmental variables obtained by orbital sensors for Cuban waters (1997-2005). Brazilian J Oceanography 56(3):225-237

Stowe LL, Jacobowitz H, Ohring G, Knapp KR, Nalli NR (2002) The Advanced Very High Resolution Radiometer (AVHRR) Pathfinder Atmosphere (PATMOS) Climate Dataset: Initial Analyses and Evaluations." J Climate 15(11):1243-1260

Shu YQ, Zhu J, Wang DX, Yan CX, Xiao XJ (2009) Performance of four sea surface temperature assimilation schemes in the South China Sea. Continental Shelf Res 29(11-12):1489-1501.

Vazquez JV, Tran AV, Sumagaysay R, Smith E, Hamilton M (1995) NOAA/ NASA AVHRR oceans Pathfinder sea surface temperature data set user's guide version 1.2. See http://pathfinder.nodc.noaa.gov/AVHRR_Pathfinder_Oceans_UsersGuidev1.2_27Sept1995.pdf

Walton CC, Pichel WG, Sapper JF (1998) The development and operational application of non-linear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. J Geophys Res 103:27,999-28,012

Appendix

Original Members of the AVHRR Pathfinder Oceans SWG

Peter Cornillon, University of Rhode Island (Chair)
Robert Evans, University of Miami
Gene Feldman, NASA
Richard Legeckis, NOAA
Charles McClain, NASA
Richard W. Reynolds, NOAA
Charles Walton, NOAA