

**Processing Framework and Match-up Database
MODIS Algorithm
Version 3**

By

Robert H. Evans
University of Miami
Miami, FL 33149-1098
April 30, 1999

TABLE OF CONTENTS

PREFACE	6
1.0 INTRODUCTION	7
1.1 Algorithm and Product Identification.....	7
1.2 Algorithm Overview.....	7
1.3 Document Scope.....	7
1.4 Applicable Documents and Publications	7
2.0 OVERVIEW AND BACKGROUND INFORMATION	7
2.1 Experimental Objective	7
2.2 Historical Perspective	8
3.0 DESCRIPTION OF ALGORITHM.....	8
3.1 Introduction based on AVHRR-Oceans Pathfinder	8
Matchup Databases	8
3.1.1 Global matchup databases.....	9
3.1.1.1 MODIS Matchup Databases	11
3.1.1.2 SST Matchup Database.....	11
3.1.1.3 Time Coordinates	13
3.1.1.4 Generation of a satellite data extraction list.....	14
3.1.1.5 Extraction and subsetting Level 1a fields	15
3.1.1.6 Matchup of the <i>in situ</i> and satellite quantities.....	17
3.1.1.7 Quality control and cloud flagging.....	18
3.1.1.8 Ocean color matchup database	18
3.1.1.9 Quantities in the oceancolor matchup record.....	19
3.1.1.10 Merging for the Matchup process.....	22
3.1.2 MATHEMATICAL ASPECTS OF THE ALGORITHM	22
3.1.2.1 Cloud filtering in the AVHRR Matchup database	22
3.1.2.3 Transition to MODIS	27
3.1.3 Variance or uncertainty estimates.....	28
3.1.3.1 Observation frequency for useful <i>In situ</i> - satellite matchups.....	28
3.1.3.2 IR Matchup database.....	29
3.1.3.2.1 Ocean Color Matchup databases	40
3.1.3.3 Error sources	42
3.2 Practical Considerations	44
3.2.1 Overview	44
3.2.2 Programming/Procedural Considerations.....	44

3.2.2.1 Overview	44
3.2.2.2 SST and ocean color processing requirements	45
3.2.2.4 Overall ocean processing time for the Level-2 radiances	45
3.2.2.5 Data Volume	46
3.2.2.5.1 Volume of data to be transferred to Miami	48
3.2.2.5.2 Total volume	48
3.2.2.6 Network Capacity	48
3.2.3 Calibration Validation	48
3.2.3.1 Introduction	48
3.2.3.2 Satellite Field and Analysis Techniques	49
3.2.3.3 Temporal comparison of temperature fields	52
3.2.3.4 Comparison of reference climatology fields	53
3.2.4 Quality Control and Diagnostics	56
3.2.4.2 Potential quality effects of stray light	58
3.2.4.3 Quality control of <i>in situ</i> observations	59
3.2.4.4 Ocean Processing Level-2 Pixel Quality control flags	63
3.2.4.4.1 SST atmospheric correction processing flags	66
3.2.4.4.2 Additional SST Flags To Be Added	67
3.2.4.4.3 Ocean color atmospheric correction processing flags	68
3.2.4.5 Ocean Processing granule/product level Quality control flags	74
3.2.5 Exception Handling	75
3.2.6 Data Dependencies	75
3.2.7 Output Product	76
4.0 Constraints, Limitations, Assumptions	76
5.0 REFERENCES	76

LIST OF TABLES

Table 1. Continuous Pathfinder time coordinate values each month, 1986-1989	14
Table 2. Quantities included in the distributed Pathfinder matchup databases	16
Table 3. <i>In situ</i> quantities tentatively identified for inclusion in the Ocean color matchup database	20
Table 4. Satellite quantities tentatively identified as needed for ocean color algorithm development	21
Table 5. Quantities to be extracted at the time of Level1 subsetting to allow subset processing to level 1b and level 2 products	21
Table 6. Validation test	26
Table 7. Buoy retrieval statistics for fixed mooring and NOAA-9 AVHRR	29
Table 8. M-AERI cruise times and locations	31
Table 9. Summary Statistics for M-AERI Matchups	35
Table 10. MODIS response function	38
Table 11. Percent difference of SeaWiFS – Buoy nLw’s MOBY time series	42
Table 12. Processing times for L2 and L3 Oceans products	45
Table 13. Summary of File sizes for SST and ocean color Archived products	46
Table 14. Table showing MODIS product number, names and ESDT relationships	47
Table 15. The climatologies assembled for inter-comparisons	54

LIST OF FIGURES

Figure 1. Residual for NOAA-14 Pathfinder Matchup database 1995-97 by latitude band..	10
Figure 2. Geographic distribution of Pathfinder SST matchups for NOAA-14 1997.....	???
Figure 3. Prune classification tree NOAA-9.....	25
Figure 4. SST estimates and the skin temperature measured by MAERI for the 1998 GASEX cruise.....	33
Figure 5. The SST mean and standard deviation of the difference from the reference MAERI SST.....	36
Figure 6. The relationship between water vapor content as estimated from the SSM/I instrument versus the Pathfinder-MAERI.....	37
Figure 7. Seasonal correction function: modeled SST4 residuals versus day of year.	39
Figure 8. Simulated Band 22 – Band 20 versus water vapor.....	40
Figure 9. Time series of SeaWiFS and MOBY buoy nLw’s.....	41
Figure 10. Pathfinder SST image for 1991, week 12.....	50
Figure 11. Reynolds blended analysis for 1991, week 12.....	51
Figure 12. Pathfinder - Reynolds difference map.	52
Figure 13. Four year presentation of Pathfinder-Reynolds difference for zonally averaged week 12, years 1990 to 1993, Noaa-11.	52
Figure 14. Shows the zonally averaged differences for week 38.....	53
Figure 15. The average ascending MPFSST minus Reynolds’ OISST [top panel] and the average TOMS Aerosol Index [bottom] for the month of June over the years 1988-1993.....	56
Figure 16. Pathfinder channel 3 - channel 4 differences (MODIS Ch20-Ch31) for two 1 week period, 1986 week 28, upper panels, and week 41, lower panels.....	59
Figure 17. OA Pathfinder SST - Buoy SST for all buoys for 1992 to 1994.	60
Figure 18. Mean OA Pathfinder SST - Buoy SST per buoy, histogrammed for all buoys for 1992 to 1994.	61
Figure 19. Plot of Buoy and OA Pathfinder SST, upper panel; OA Pathfinder – buoy SST, middle upper panel, OA Pathfinder - Buoy SST histogram, lower middle panel, and buoy trajectory, lower panel.	62
Figure 20. Plot of Buoy and OA Pathfinder SST, upper panel; OA Pathfinder – buoy SST, middle upper panel, OA Pathfinder - Buoy SST histogram, lower middle panel, and buoy trajectory, lower panel.....	63
Figure 21. Panels A-L MODIS results for converted SeaWifs data.....	65

APPENDICES

Appendix 1 Equal-area gridding scheme.....	A1.1
A1.1 Introduction.....	A1.1
A1.2 Overview.....	A1.1
A1.3 The poles.....	A1.4
A1.4 Binning software.....	A1.4

Appendix 1: LIST OF FIGURES

Figure A1-1. Number of 9.28 km tiles per zonal row as a function of latitude.....	2
--	----------

Appendix 2 Data Day Definition.....	A2.1
A2.1 Introduction.....	A2.1
A2.2 A 24-hour data day.....	A2.2

A2.3 A spatial data-day definition.....	A2.4
A2.4 How is the beginning of a data-day defined?	A2.7
A2.5 Advantages of the Spatial Definition of a Data-Day	A2.9
A2.6 Other Issues.....	A2.14

Appendix 2: LIST OF TABLES

Table A2-1.	Beginning times of fifteen data-days for descending orbits.....	9
--------------------	---	----------

Appendix 2: LIST OF FIGURES

Figure A2-1.	Descending NOAA-11 tracks for a 24-hour data-day	2
Figure A2-2.	Locations of the boundaries of 24-hour data	4
Figure A2-3.	NOAA-11 descending orbits for a spatially-defined data-day.....	5
Figure A2-4.	AVHRR nadir track and scan lines.....	6
Figure A2-5.	Latitude of crossing of the 180° meridian	8
Figure A2-6.	NOAA-11 descending orbits for spatially-defined data-day.....	11
Figure A2-7.	Data-day beginning on July 27 1992 15:14:00 UTC, segments	12

Preface

This Algorithm Theoretical Basis Document (ATBD) describes our current working model for construction of a matchup data base for calibration of the ocean visible and infrared bands of the MODIS sensor and integration of the ocean product algorithms into a cooperative group of programs. While effort has been made to make this document as complete as possible, the reader should understand that this version of the document is a snapshot of ongoing work, *i.e.*, the algorithm development is an evolving process.

As will be seen from reading the document, there are areas that still require research effort before finalization. In particular this effort depends on continued availability of algorithm descriptions and codes from the participating ocean investigators. The results described in this document are based on ongoing joint development and tests associated with the NASA/NOAA Pathfinder AVHRR Oceans activity and SeaWiFS projects. Experience gained with these efforts is directly assisting development of the MODIS comparison database with respect to design, testing and implementation.

1.0 Introduction

Document focuses on implementation of a Level-2 framework incorporating the Level-2 algorithms of the MODIS Ocean Team investigators (Abbott, Brown, Carder, Clark, Esaias, Evans, Gordon, Hoge). The implementation first addresses the conversion of L1 counts into calibrated radiance or reflectance, conversion of these at satellite radiances into brightness temperatures or water leaving radiances L_w , and generation of the products by executing the individual product algorithms.

1.1 Algorithm and Product Identification

The calibration data set produced by the algorithm will be labeled version 2. This is a level 2 product with EOSdis product number 3303; it is MODIS product number 34, labeled Calibration Data.

1.2 Algorithm Overview

This algorithm is being developed on the MODIS Ocean Team Computing Facility (*MOTCF*) for use in the EOS Data and Information System (*EOSdis*) core processing system. The individual ocean product algorithms described in the respective MODIS Ocean Team PI ATBD's will be combined into a structure that provides easy integration into the MODIS TLMCF leading to final porting into the MODAPS environment.

Initial focus is Level-2 conversion of Level-1 counts/reflectance into water leaving reflectance for the visible and brightness temperatures for the infrared. Algorithms described in the ATBD's for Gordon and Brown are used to describe considerations affecting the development of appropriate match-up data bases, integration of their algorithms into a Level-2 program and subsequent comparison of the algorithm results with the in-situ observations. Considerations for generating Level-3 products from the Level-2 are presented in Appendices 1 and 2. The Level-3 fields provide an important validation component of the algorithms integrated in the Level-2 program.

1.3 Document Scope

This document describes the basis for the Calibration database, gives the structure of the current version 1 algorithm, discusses implementation dependencies on other observing streams, and describes validation needs.

1.4 Applicable Documents and Publications

MODIS Proposal, 1990, Processing and Calibration for Ocean Observation with EOS/MODIS, Robert H. Evans

MODIS Execution Phase Proposal, 1991, Processing and Calibration for Ocean Observation with EOS/MODIS, Robert H. Evans

2.0 Overview and Background Information

2.1 Experimental Objective

The goals of this proposal are twofold: generation of a matchup database that will serve as a basis for algorithm validation and integration of algorithms developed by the MODIS Ocean Team (MOT) into a coherent processing package. Validation of the water leaving radiances will be effected through application of a match-up database that includes appropriate satellite and *in-situ* measurements. Initial uses of the matchup data base include testing of color and SST algorithm performance as a function of viewing geometry (satellite and solar zenith angle), space (latitude and longitude), time, a variety of geophysical conditions (*e.g.* wind speed, ozone, water vapor concentration) and state of the sensor (*e.g.* sensor temperature, mirror position, orbit). Previous applications of this approach include characterizing CZCS and SeaWiFS ocean color and AVHRR infrared long-term sensor drift and development and evaluation of candidate SST algorithms for the ocean Pathfinder project.

Generation of the matchup database and integration of the MODIS ocean team algorithms utilizes the SCF developed at the University of Miami/RSMAS. The facility presently includes DEC ALPHA workstations and a multi-processor Compaq and SGI computers. Sufficient processing power is available to generate MODIS ocean products at full resolution for a reasonable fraction of the world's oceans. However, rather than processing the full resolution data set, a subset of the data will be analyzed to test calibration and validate algorithms.

2.2 Historical Perspective

Heritage programs provide the basis for MODIS algorithm. A program developed for the Pathfinder ocean SST product forms the framework for analyzing AVHRR derived SST, algorithm development and validation, and application of the matchup database. Development for ocean color algorithms is based on experience gained in transition from CZCS to SeaWiFS algorithms. Development of the SeaWiFS program involves integration of algorithms generated by H. Gordon (atmospheric correction) and K. Carder (chlorophyll) with additional MODIS algorithms provided by F. Hoge, D. Clark and M. Abbott. In addition, data validity tests are being developed through collaboration of the SeaWiFS CAL/VAL team and our group. The SeaWiFS and Pathfinder programs have been converted to the MODIS implementations and coded using the FORTRAN 90 language and EOS coding standards.

3.0 Description of Algorithm

3.1 Introduction based on AVHRR-Oceans Pathfinder Matchup Databases

Algorithm development supported for this effort includes generation of a matchup database (MDB) that consists of spatially and temporally matched satellite and *in situ* observations and application of this database to assist in MOT algorithm development and validation. The following sections describe current efforts where a MDB has been created for NOAA AVHRR instruments and used to analyze the behavior of AVHRR SST retrieval equations. This methodology has been extended to include ocean color

observations obtained from the SeaWiFS sensor. These two data sources provide equivalent wavebands for all MODIS ocean observations with the exception of the 4 micron infrared channels and the fluorescence channel. In addition to providing experience, the near equivalence of the present sensors provides a comparison database for MODIS observations.

As part of the Oceans component of the AVHRR Pathfinder project sponsored by NOAA and NASA, “matchup” data bases are compiled, combining *in situ* surface temperature (SST) measurements (and, in some cases, other *in situ* environmental variables) and AVHRR data corresponding to the times/locations of *in situ* measurements. The purpose of the matchup data bases is the development, testing and validation of Pathfinder SST algorithms.

Extension to include population of an *in situ* portion of the ocean color database became possible in late Fall, 1996 following the initial test deployment of the MOBY instrument by D. Clark. This data set has been combined with SeaWiFS observations to compile an initial ocean color MDB.

3.1.1 Global matchup databases

Experience in producing and evaluating the matchup database (MDB) has been gained using the NOAA AVHRR 4-km Global Area Coverage (GAC) observations. The database has been applied in the following areas: generation of SST retrieval equation coefficients for the AVHRR instruments, evaluation of spatial and temporal trends, quality control of the *in situ* validation data set and comparison with radiative transfer simulations. Figure 1 presents an example of the application of the PFMDB to examine the bias in satellite SST retrievals for AVHRR Pathfinder. Satellite– buoy residuals are presented by 20-degree latitude bands for NOAA-14, 1995-1997. The dotted lines show the $\pm 0.5\text{C}$ range and the dark box for each point shows the temperature spread for 50% of the retrievals. Each point represents a one-month collection within the latitude band. The outer braces are approximately 3 standard deviations while the remaining bars show individual retrievals that exceed 3 std. Deviations and likely represent unresolved ‘clouds’. The white bar within the black box is the median of the data. For most points the 50% box falls well within the $\pm 0.5\text{C}$ line. A notable exception occurs during the end of 1997 for the equatorial band. This period is shown in the blue insert and contains a number of buoys that were deployed in the eastern tropical Atlantic when Saharan dust was present. The effect of the dust is a noticeable depression in the retrieved satellite SST. The Pathfinder approach for the most part yields high quality SST retrievals. Exceptions include unresolved clouds and aerosols. With respect to the buoys the Pathfinder SST is 0.02C and the standard deviation is 0.53C . Pathfinder retrieval equation coefficients are determined using regression of satellite brightness temperatures against buoy temperatures for cloud free observations from the matchup database.

NOAA-14 Pathfinder-BuST Residuals

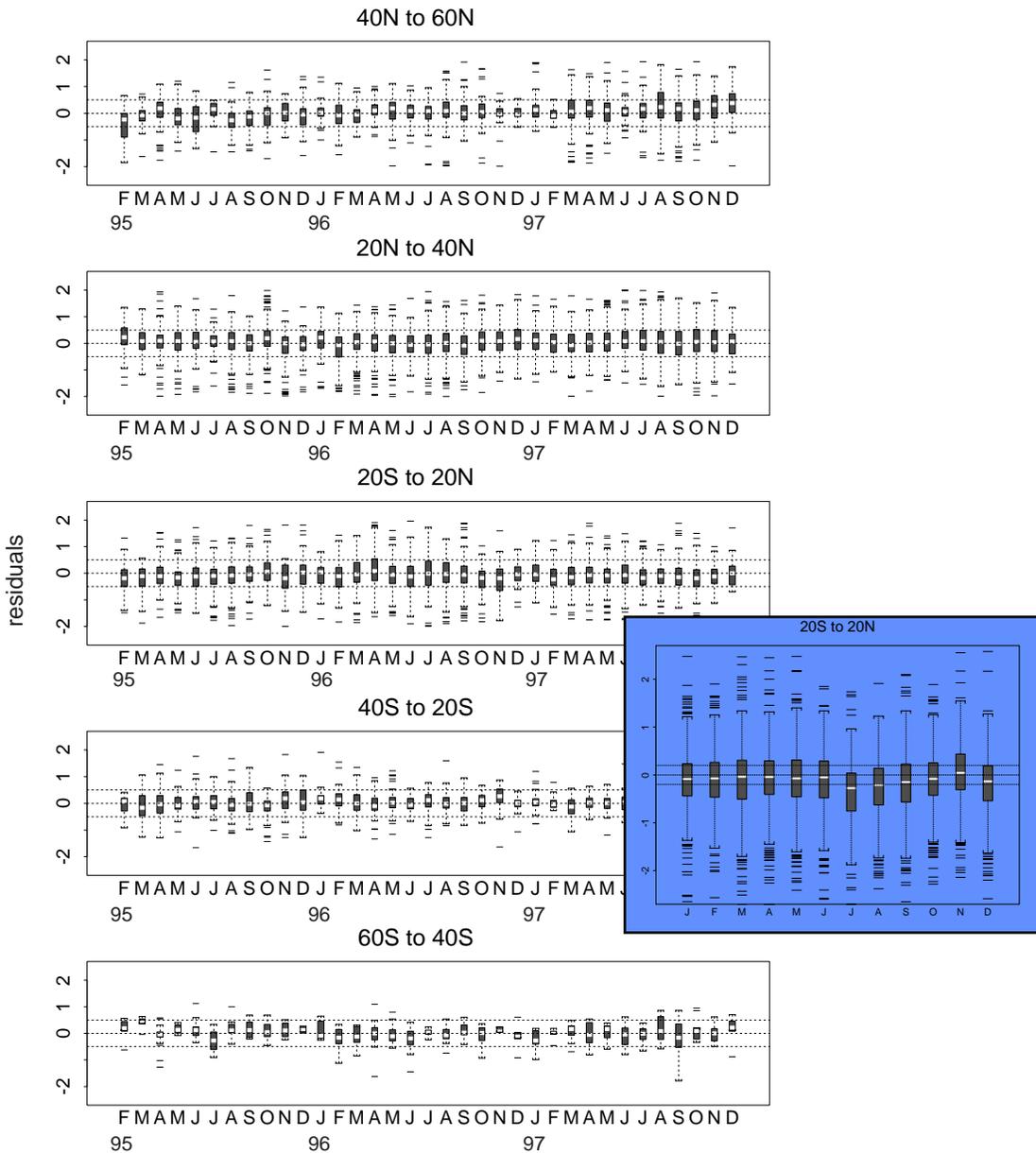


Figure 1. Residual for NOAA-14 Pathfinder Matchup database 1995-97 by latitude band. Insert shows expanded view of 1997 residuals in the equatorial latitude band.

The matchup databases for MODIS are based on procedures developed for the AVHRR MDB. Spatial resolution of the database will be 1-km rather than the 4-km available with AVHRR GAC observations. The initial MODIS MDB might be forced to continue using the present 4-km resolution as a fall back option should either a limitation in subsetting capability or data delivery capability impact access to the global Level-1

MODIS observations. The present intent is to generate the MDB at Miami using global, full resolution L1A files obtained from the Goddard DAAC.

3.1.1.1 MODIS Matchup Databases

Two matchup database products; MOD18_L2_Bouy and MOD28_L2_Bouy, will be developed for the MODIS sensor. These databases will contain temporally co-located *in-situ* measurements and subsetted Level-1a and Level-2 satellite data for ocean color and sea surface temperature, respectively. Each of these databases contains both a "real-time" and a retrospective component.

The "real-time" component is a preliminary database containing *in situ* records that were available electronically on the day that the level-1 granule is processed at MODAPS. A month of daily "real-time" matchups are assembled into a single file and should be available for distribution shortly after the close of each month. This component of the database is relatively small and typically contains moored buoy locations representing only 20% of *in situ* records collected globally on a given day. The limited spatial coverage of the "real-time" component requires a second, retrospective approach.

The retrospective component contains *in situ* data sources from both moored and fixed buoys, and for ocean color, shipboard measurements that are gathered over a 1 year time period from sources not available in "real-time". The corresponding year's archived Level-1a data is then searched to extract the satellite information. A year of retrospective matchups are split into individual months and should be available for distribution within the first quarter of the following year. Details of the processing and factors, which may affect the timing of the distribution of these products, are presented below.

Compilation of the Matchup databases:

There are four main common steps involved in the compilation of the Matchup databases:

- *Obtaining and reformatting of *in situ* data
- *Extraction of the subsetted satellite quantities
- *Matchup of the *in situ* and satellite quantities
- *Quality control and flagging of cloud contaminated data

These common aspects will first be discussed in the context of the SST Matchup database. Differences associated with the Ocean color Matchup database will be discussed separately in a later section.

3.1.1.2 SST Matchup Database

Obtaining and Reformatting of *in situ* data:

The SST *in situ* data originate from two main sources, moored and drifting buoys. The *in situ* sources for the MODIS sensor will be the same as those presently used in the Pathfinder Oceans Matchup Database(PFMDB). The agencies supplying the data are:

Moored buoys:

US National Data Center (NDBC)

Japan Meteorological Agency

TOGA/TAO Project office(NOAA Pacific Marine and Environmental Laboratory)

Drifting buoys:

AOML(NOAA Atlantic and Oceanographic and Meteorological Laboratory)

MEDS(Canadian Marine and Environmental Data Service)

The retrospective component of the database obtains data from the above sources via FTP or tape directly from the archival agency after the data has been subjected to various stages of quality control and made available to the general science community. In 1995 200,000 *in situ* records compiled from these sources were used in the AVHRR MDB. These agencies recently have been reducing their network of buoys over the last few years. If this trend continues we estimate that 300,000 records per year will be contained in the retrospective component of the Sea Surface temperature Matchup Database for the MODIS sensor the current distribution of cloud free 1997 Matchups are presented in Figure 2.

NOAA-14 1997 Matchups

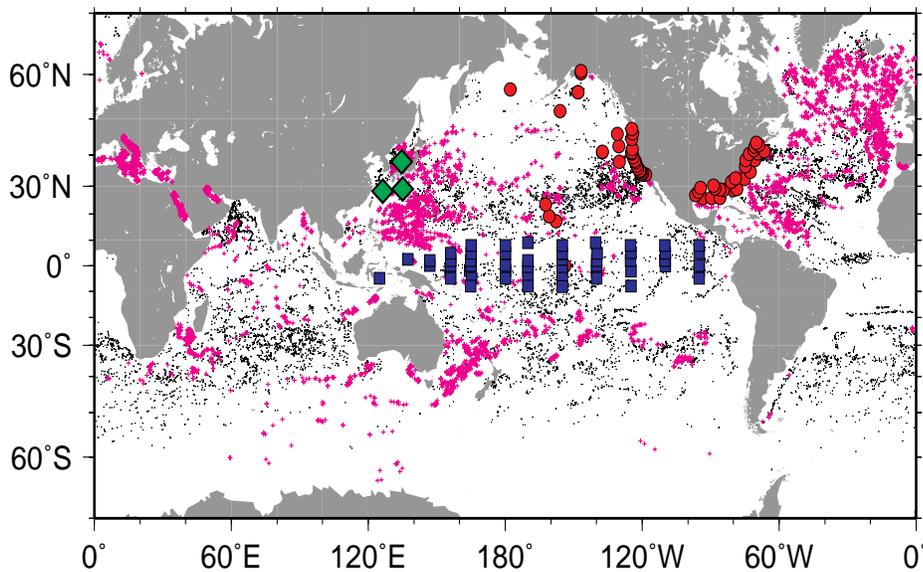


Figure 2. Geographic distribution of Pathfinder SST matchups for NOAA-14 1997. Only those matchups flagged as "cloud free" are plotted. The dots indicate the NDBC moored buoys, the diamonds are the Japanese moored buoys, triangles are the UK buoys and the squares are the TOGA/TAO buoys. Small pink points indicate AOML drifters, whereas small crosses denote MEDS drifters. NATO drifters are indicated as small yellow dots.

Many of the agencies post a portion of the daily data collected from their buoy network on the GTS system for real-time distribution. The U.S. Navy (NAVOCEANO) monitors the GTS and daily assembles a file containing data posted by all of the above sources. The Navy file is then released and electronically transferred to Miami RSMAS the following morning. In 1995 the retrospective and real-time components of the AVHRR MDB contained 772 and 571 unique buoy IDs, respectively. It is important to note that in the retrospective component a buoy reports 6-8 times a day. In contrast, buoys in the real-time component report 1-2 a day.

In situ data compiled from all of the sources are reformatted into a standard format and very general quality control is applied to the data. The quality control includes testing that the position is not on land, SST value is within a reasonable range, and valid date and time values are present. Drifting buoys present in the retrospective component are separated by unique ID's and the speed and acceleration of each buoy is checked. The buoys in the real-time component are continually monitored by unique ID for the same quality issues. If a buoy is found to have failed the quality control tests from the preceding day it will be removed from the satellite extraction list for the incoming day. After passing quality control tests, the original platform IDs (usually WMO or ARGOS ID) are modified to facilitate the identification of the various *in situ* SST sources once all the matchups are assembled by adding a source dependent constant to the original ID. Buoys IDs in the real-time component are not modified and remain as the WMO ID. To facilitate the matchup process of the *in situ* and satellite records the dates and times of the *in situ* SST reports are converted to a continuous time coordinate. The Gregorian dates and UTC time of the *in situ* data are included in the Matchup Database for convenience of the users.

3.1.1.3 Time Coordinates

To facilitate the matchup process, dates and times of both the satellite and *in situ* data are converted to a continuous time coordinate, "seconds since January 1, 1981," here referred to as "Pathfinder seconds". Table 1 shows the values of Pathfinder seconds for the beginning of each month in 1982-89. The values correspond to day 1 of each month at 00:00:00 GMT. These values can be subtracted and then the actual date can be obtained through a few simple calculations.

Table 1. Continuous Pathfinder time coordinate values for the beginning of each month, 1986-1989.

	1986	1987	1988	1986
Jan	157,766,400	189,302,400	220,838,400	252,460,800
Feb	160,444,800	191,980,800	223,516,800	255,139,200
Mar	162,864,000	194,400,000	226,022,400	257,558,400
Apr	165,542,400	197,078,400	228,700,800	260,236,800
May	168,134,400	199,670,400	231,292,800	262,828,800
Jun	170,812,800	202,348,800	233,971,200	265,507,200
Jul	173,404,800	204,940,800	236,563,200	268,099,200
Aug	176,083,200	207,619,200	239,241,600	270,777,600
Sep	178,761,600	210,297,600	241,920,000	273,456,000
Oct	181,353,600	212,889,600	244,512,000	276,048,000
Nov	184,032,000	215,568,000	247,190,400	278,726,400
Dec	186,624,000	218,160,000	249,782,400	281,318,400

3.1.1.4 Generation of a satellite data extraction list

Due to the time consuming nature of the satellite data extraction process for the retrospective component of the database, we have developed a procedure to exclude those times and locations for which there was not a satellite pass within ± 30 minutes from the *in situ* measurements. The procedure involves the computation of the Time of Closest Approach (TCAP) with respect to a given point and time. This procedure relies on orbital routines provided by D. Baldwin (Univ. of Colorado). The TCAP filter significantly reduces the time required for the extraction process. The TCAP filter outputs records containing buoy ID, latitude, longitude, time, orbit #, and slant distance for *in situ* records which fall within the position/time window. These records are then assembled into individual day extraction lists. *In situ* records which occur within -30 minutes of the start of day N or $+30$ minutes of the end of day N are placed on the extraction list of day N. For MODIS, the time criterion will match the MODIS time to the closest available buoy time, typically this will be from the 'morning' NOAA satellite.

The building of extraction lists for the real-time component is different than that of the retrospective component due to several constraints. The satellite data subsetting and extractions will be done in near real-time with the Level 1 processing thereby eliminating the lengthy archive search time present in the retrospective component. In addition, the total number of *in situ* records available at the time of Level-1 processing is significantly less. The decrease in both the volume of data and search time precludes the need for TCAP filtering. A timing constraint still remains in regard to the receipt of the *in situ* data used to build extraction lists and the processing of Level1 granules. The

real-time *in situ* file for day N is received on day N+1. Typically an *in situ* file for day N contains records for a 3-4 day time window centered on day N. In the AVHRR MDB this constraint is handled by filling the extraction list for day N with records from *in situ* files received for day N to N+7 which fall on day N. The extraction list for day N is submitted for extraction on day N+7. For the Pathfinder AVHRR, we are able to hold a rolling 7-10 day temporary archive of GAC data on-line to facilitate extraction. We assume that the on line storage of MODIS Level 1 granules will limit completeness of the real-time extraction list. We are planning to fill the extraction list with position/time from *in situ* drifting and fixed buoy records received at the time of required extraction list close. The extraction list for day N will then be supplemented with known fixed buoy positions that we expect to receive for that data day after list close. Analysis of the Navy *in situ* files shows that 80 and 90% of the *in situ* records posted to the GTS are received within 4 and 7 days, respectively of *in situ* collection.

The compilation of all of the *in situ* data and all processing relating to the Matchup databases will occur at the Miami/RSMAS SCF. These products will then be delivered to the DAAC for insertion in the EOS/MODIS data server.

3.1.1.5 Extraction and subsetting Level 1a fields

Satellite data from the level 1a granules will be extracted for various size boxes centered at each *in situ* location. Once the granule to be extracted has been identified, the geographical position of the *in situ* point must be inverted to granule coordinates of scan line and pixel. The initial extraction subset includes the level-1a counts, information needed to convert the level 1a to level-1b, pixel geometry, and other ancillary data. For SST this includes the level 1a pixel counts, cloud mask, reference climatology, geolocation, aerosol, and calibration files. It is anticipated that 2,000 positions will be extracted a day.

The extraction box size for 25 positions will be 1000*1000 pixels, 75 positions at 512*512 pixels, and 1900 at 100*100 pixel boxes. The larger sized boxes will correspond to special long term study sites where it will be desirable to follow evolution of ocean or atmospheric features in the area..

The initial extraction subset is then processed to level-1b (brightness temperatures) and other quantities needed for algorithm development and evaluation. Summary statistics will then be determined for a 5x5 pixel box centered on the *in situ* position. The summary statistics for each of the relevant channels (bands 20, 22, 23, 26, 31 and 32) of the satellite record includes the central pixel value, mean and median of all values inside the box, and the minimum and maximum values within the box. Also included in the satellite record will be information relevant to sensor calibration, and the slopes and intercepts for the counts-to-radiance conversions. Using the geometric information, a sun glint index is calculated to assist in the identification of pixels contaminated by glint. An example of the fields included in the satellite record of the present AVHRR MDB is shown in Table 2.

Table 2. Quantities included in the distributed Pathfinder matchup databases. The variable names given in the last column (“Code”) are used in the “Record filtering and flagging” section.

Field	Field Description	Units	Code
1	In situ date (Gregorian, YYMMDD)	yr,mo,day	yymodd
2	In situ time (UTC, HHMMSS)	hr,min,sec	hhmmss
3	Satellite time (Pathfinder coordinates)	seconds	stime
4	Latitude of center pixel	Decimal deg.	slat
5	Longitude of center pixel	Decimal deg.	slon
6	Average PRT temperature	°C	prt
7	Solar zenith angle	Degrees	solz
8	Satellite zenith angle	Degrees	satz
9	Sun glint index	sr ⁻¹	glnt
10	Emissivity , channel 3	—	em3
11	Emissivity , channel 4	—	em4
12	Emissivity , channel 5	—	em5
13	Central value of 5x5 pixel box, channel 1	—	ch1
14	Central value of 5x5 pixel box, channel 2	—	ch2
15	Central value of 5x5 pixel box, channel 3	°C	ch3
16	Central value of 5x5 pixel box, channel 4	°C	ch4
17	Central value of 5x5 pixel box, channel 5	°C	ch5
18	Median of 5x5 pixel box, channel 1	—	med1
19	Median of 5x5 pixel box, channel 2	—	med2
20	Median of 5x5 pixel box, channel 3	°C	med3
21	Median of 5x5 pixel box, channel 4	°C	med4
22	Median of 5x5 pixel box, channel 5	°C	med5
23	Minimum value of 5x5 pixel box, channel 1	—	min1
24	Minimum value of 5x5 pixel box, channel 2	—	min2
25	Minimum value of 5x5 pixel box, channel 3	°C	min3
26	Minimum value of 5x5 pixel box, channel 4	°C	min4
27	Minimum value of 5x5 pixel box, channel 5	°C	min5
28	Maximum value of 5x5 pixel box, channel 1	—	max1
29	Maximum value of 5x5 pixel box, channel 2	—	max2
30	Maximum value of 5x5 pixel box, channel 3	°C	max3
31	Maximum value of 5x5 pixel box, channel 4	°C	max4
32	Maximum value of 5x5 pixel box, channel 5	°C	max5
Table 2. continued			

Field	Field Description	Units	Code
33	Average value of 5x5 pixel box, channel 1	—	av1
34	Average value of 5x5 pixel box, channel 2	—	av2
35	Average value of 5x5 pixel box, channel 3	°C	av3
36	Average value of 5x5 pixel box, channel 4	°C	av4
37	Average value of 5x5 pixel box, channel 5	°C	av5
38	PRT 1 Temperature	°C	prt1
39	PRT 2 Temperature	°C	prt2
40	PRT 3 Temperature	°C	prt3
41	PRT 4 Temperature	°C	prt4
42	Central value channel 1 (same as field 13)	—	cm1
43	Central value channel 2 (same as field 14)	—	cm2
44	Central value w/ emissivity correction, ch. 3	°C	cm3
45	Central value w/ emissivity correction, ch. 4	°C	cm4
46	Central value w/ emissivity correction, ch. 5	°C	cm5
47	Calibration slope, channel 3	$\text{mW m}^{-2} \text{cm}^{-1}$ $\text{sr}^1 \text{count}^{-1}$	slope3
48	Calibration slope, channel 4	see Field 47	slope4
49	Calibration slope, channel 5	see Field 47	slope5
50	Calibration intercept, channel 3	$\text{mW m}^{-2} \text{cm}^{-1}$ sr^1	intcp3
51	Calibration intercept, channel 4	see Field 50	intcp4
52	Calibration intercept, channel 5	see Field 50	intcp5
53	Time of in situ SST, Pathfinder coordinates	Seconds	btime
54	Buoy latitude	Decimal deg.	blat
55	Buoy longitude	Decimal deg.	blon
56	Buoy ID	—	bid
57	In situ SST	°C	bsst
58	Delta-SST (First-guess sat. SST minus in situ SST)	°C	sst1
59	Filter code (1 or 2)	—	pass

In the AVHRR MDB only the summary statistics for the original subset level-1a processed to level-1b 5x5 pixel box is retained in the satellite portion of the record. For the sensor it may be desirable to make the original subset level-1b pixel data available and include a file pointer in the matchup database. This would enable end users of the database to derive contextual information.

3.1.1.6 Matchup of the *in situ* and satellite quantities

The next step in the generation of the matchup databases is to temporally match *in situ* records against satellite extractions. To limit the variability introduced by the time separation between the two data sources (Minnett 1991), the absolute difference between the time of the *in situ* report and the time at which that location was viewed by the satellite (matchup time window) is restricted to a maximum of 30 minutes or to the closest available *in situ* record where these data are obtained via another satellite. *In situ* records that do not fall within the stipulated time window will be rejected. Furthermore, the real-time satellite records that may have data for predicted *in situ* times that were not received, will also be rejected. *In situ* records, which pass the temporal matchup, must subsequently pass a spatial test. A maximum distance of 0.1° in latitude and longitude will be allowed between the *in situ* location and the location of the central pixel in the extraction box. The matchup procedure for TOGA/TAO buoys will be slightly different from those used for other data sources. These buoys differ from other sources because the reported SST values represent averages over longer time periods. The reported SSTs for the TOGA/TAO buoys are the average of six measurements taken every 10 minutes; the reporting time is the end of each averaging period. The *in situ* time for the center of the matchup window and reported in the matchup database will correspond to the center of the *in situ* averaging period. The output of the matchup process is a series of records that contain both satellite-derived and *in situ* derived data. In a small number of cases, the satellite part of the matchup record is the same for two *in situ* records collected a short time apart. For instance, some NDBC buoys report data every 30 minutes. In these cases the same satellite extraction may satisfy the matchup window for two consecutive *in situ* observations. Only the matchup record with the closest *in situ* and satellite time will be retained.

3.1.1.7 Quality control and cloud flagging

A large number of the matchup records will have obvious problems such as gross cloud contamination and should not be used for algorithm development. A series of cloud tests will be developed by the University of Miami RSMAS to identify high quality matchups records which can be confidently used in algorithm development and coefficient estimation. Presently we are planning to distribute all matchup records and include a quality flag indicating if the record passed the Miami cloud tests. This will enable end users the ability to develop other cloud tests if desired. A description of the present development of new cloud filters and coefficient estimation by the University of Miami for the Pathfinder program can be found in Appendix A. The final stage of the database is to assemble matchup records into individual monthly files that will be delivered to the DAAC for distribution. The cloud flags developed by Menzel will be included once their procedures are validated.

3.1.1.8 Ocean color matchup database

The overall scheme of the process used to create the ocean color matchup database; MOD18_buoy will be the same as that for the SST matchup database described above.

The main difference between the two databases is the source and structure of the ocean color *in situ* data and the number of satellite quantities to be included in the final matchup record. The *in situ* data compiled from various sources include drifter and moored buoys, and shipboard along-track and profile measurements of bio-optical quantities. In contrast to the SST measurements, the *in situ* bio-optical quantities in the database will have both a vertical and horizontal measurement for a given geographical location when available. This three dimensional nature of the ocean color data will require that a matchup record for a given time and geographical location be separated into multiple files which share at least one field in common so as to link the *in situ* information. The matchup files submitted to the DAAC for distribution will be flat files. The University of Miami will create a relational database for these files, which will allow custom queries to create unique views of the quantities present across the matchup records. The team members and the general bio-optical scientific community is presently being solicited for input on the *in situ* and satellite quantities which should be included in this database. Sources and quantities that have been tentatively identified are discussed below.

Data sources:

The real-time component will consist of data from the MOBY moored buoy and the MOCE cruises located off Hawaii being delivered by Dennis Clark and associates. Several other bio-optical moored buoys are planned by a number of different Agencies and these will be added as they become available. We are estimating that 20 moored and drifting buoys will eventually be available real-time.

The retrospective component will consist of *in situ* data available from the calibration-validation effort and other Agencies that include bio-optical measurements in their respective field programs. Sources tentatively identified include the JGOFS field program, Bermuda Atlantic Time Series (BATS), Hawaii Ocean Times series (HOTS), and the Japanese YBOM in support of OCTS. Cruise data from individual investigator program will be included as they are identified. In the event that the SEABASS-SIMBIOS bio-optical database in support of SEAWIFS is still available it will also be included as an *in situ* data source.

3.1.1.9 Quantities in the oceancolor matchup record

In situ matchup record:

Table 3 shows the *in situ* quantities that have been identified from the MOCE datasets for inclusion in matchup record. Many of these quantities are measured by several different instruments using different wavebands and depth resolutions. Where appropriate the matchup record will contain the same quantity measured or derived by several different techniques or instruments. For example the water leaving radiance (Lw) in the MOCE dataset is measured by both the Biospherical MER and the MOS-SIS high resolution spectrometer. Data from both instruments will be included in the matchup record. A field with instrument and technique flags will need to be added for many of the quantities.

Table 3. *In situ* quantities tentatively identified for inclusion in the Ocean color matchup database. Many of these quantities consist of both a depth and spectral measurement.

<i>In Situ</i> quantity	Code
date	date
time	time
latitude	lat
longitude	lon
depth	Z
water- leaving radiance	Lw
normalized water- leaving radiance	nLw
diffuse attenuation downwelling	Kd
diffuse attenuation upwelling	Ku
optical depth	1/Kd
reflectance	R
remote sensing reflectance	R1
Photosynthetically available radiation	PAR
transmission	%t
beam attenuation coefficient	c
chlorophyll fluorescence	ChlFl
aerosol optical thickness	AOT
chlorophyll concentration (fluoremetric)	chl
chlorophyll concentration (hplc)	chla-hplc
Chlorophyll fluorescence	chlF
Phycobilin pigment concentration	PUB,PEB
Coccolith concentration	cocco
total particulate absorption	aph
detrital absorption	dph
dissolved organic matter concentration	DOC
total suspended solids	ses
air temperature	airT
windspeed	ws
relative humidity	%rh
Sea surface temperature	sst
instrument/technique flags	flag

Satellite matchup record:

The satellite portion of the matchup record will include the statistical summary for the level 1b data derived from the level-1a extraction subset for all of the visible channels and any other channels, calibration, and ancillary data identified for algorithm development. In addition the level 1b subset will be processed to level-2 and a statistical summary of the level 2 products will be included. It is not planned to include all 36 level-2 products in the matchup record. Only those level-2 products (i.e., Lw's + others) identified as required for algorithm development will be present. Other intermediate quantities used or derived during Level-2 processing which are important in algorithm development may also be include. A list of satellite quantities presently identified as important for algorithm development is presented in Table 4.

Table 4. Satellite quantities tentatively identified as needed for ocean color algorithm development.

Satellite quantities	Processing level
pixel counts	level 1a (bands 8-16 ,5-7,26)
calibration information	level 1b (bands 8-16 , 5-7,26)
pixel geometry	level 1b (bands 8-16 , 5-7,26)
reflectance	level 1b (bands 8-16, 5-7,26)
aerosol reflectance (La)	-
Aerosol optical thickness (tau)	-
Aerosol model used	-
Lw	Level 2 (bands 8-16, 5-7,26)
chlorophyll concentration	level 2
Coccolith concentration	level 2
Fluorescence line height	level 2
Dissolved organic matter	level 2
Phycobin concentrations	level 2
Sea surface temperature	level 2
various quality flags	Level 1 and 2

Ancillary matchup record fields:

The addition of the level-2 data in the ocean color matchup record requires that the level-1a extraction and subsetting process include all information needed as input to the level-1b and level-2 processing PGE's. The data types required as output of the matchup extraction and subsetting process is shown in Table 5. As suggested in the SST matchup database, it may be desirable to make the level 1b and level-2 subsetted pixel data available for contextual purposes and include a file pointer in the matchup record.

Table 5. Quantities to be extracted at the time of Level1 subsetting to allow subset processing to level1b and level 2 products.

Subsetted products and quantities
Level 1a raw pixel counts MOD01
calibration information Level1b
geolocation MOD03-L1a
cloud mask MOD06
aerosol MOD 04
3 week L2 ocean color reference field
SST MOD28
Ancillary meteorology and ozone

3.1.1.10 Merging for the Matchup process

Merging of the ocean color *in situ* and satellite records in the matchup process will be the same as described above for SST. The matchup record will include the satellite data *in situ* measurements, and a quality flag indicating if the record passed our cloud flagging technique. It is planned that a single monthly matchup file will be delivered to the DAAC in an HDF format. However due to the complexity of the *in situ* record multiple files may be required.

3.1.2 MATHEMATICAL ASPECTS OF THE ALGORITHM

3.1.2.1 Cloud filtering in the AVHRR Matchup database

A large number of matchup records have obvious problems (for instance, gross cloud contamination) and cannot be used for algorithm testing or coefficient estimation. For that reason, a sequence of tests is implemented in the PFMDB to identify most cloud-contaminated matchups. The tests include two stages: (a) a set of initial tests common to all satellites (b) a decision tree involving test derived separately for data from each sensor. The initial tests include absolute thresholds for brightness temperature or radiance values in channels that intend to exclude very anomalous values that may result from digitizer errors. Spatial uniformity tests are intended for further cloud identification. These tests involve thresholds for differences between the minimum and maximum values within the extraction box for a given channel. Early versions of the PFMDB included cloud-flagging tests that had been defined after extensive interactive examination of the data. Although these tests fulfilled their objective of excluding cloud-contaminated matchups from the algorithm estimation process, they were overly conservative, rejecting a large amount of potentially usable matchups. Furthermore, the selection of tests were repeated for matchups from each AVHRR, as calibration changes from sensor to sensor could potentially invalidate the use of the same tests for all matchups. For these reasons, a new methodology was developed in Version 19 of the PFMDB for the second stage (b) of the cloud-flagging step. This methodology is based on the tree models described by Clark and Pregibon (1992) and by Venables and Ripley (1994). There are two main advantages in the new method: the cloud-flagging tests are selected objectively, and the number of potentially useful matchups rejected is lower.

Briefly, the classification trees are based on binary recursive partitioning, whereby a data set is successively split into increasingly homogeneous subsets. In the present context, tree models can find the best way to predict membership in one of two groups ("cloud-contaminated" or "cloud-free") as a function of a set of predictor variables which may contain information about cloud contamination (*e.g.*, differences in brightness temperatures between channels).

The process of growing a Tree based model for use in classifying cloud-free ("good") and cloud-contaminated ("bad") matchup records involves several steps.

1. Initial filtering of PFMDB for quality control (reasonableness).
2. Define category membership of record as "good" or "bad".
3. Creation of training and validation datasets.
4. Selection of variables to be used in the Tree model.

Initial filters:

These tests contain thresholds for the channel brightness temperatures to ensure that the satellite data had reasonable geophysical values and thereby remove records that contain digitization errors. Our experience has been that biased final satellite SST values are obtained from the central pixel located in non-homogeneous extraction boxes. We therefore included a homogeneity test (max-min <0.7 of 5x5 extraction box) for both AVHRR channels 4 and 5.

Define category membership:

Category membership of a record was determined from the residual of the buoy SST minus a First Guess SST. The First Guess SST may be calculated using operational algorithms (*i.e.* NOAA-NESDIS NLSST) or coefficients developed from other cloud flagging routines such as described earlier. A "good" record is defined as having an $\text{abs}(\text{residual}) \leq 2^{\circ}\text{C}$, records while an $\text{abs}(\text{residual}) > 2^{\circ}\text{C}$ is defined as "bad". The category membership is dependent on having an unbiased estimate of the First Guess SST. The median of the "good" residuals must be close to zero to use the First Guess SST for membership definition. If the algorithm used for the First Guess SST results in a biased estimate, alternative formulations for the First Guess SST must be found.

Training and Validation sets:

The training set was created by randomly selecting 1/3 of the night time records from the matchup database. The selection process included a probability weighting function so that the training set would contain approximately equal numbers of both "good" and "bad" records. The use of only nighttime records was chosen to minimize the possible bias in threshold values for the splits as a result of solar contamination. The goal was to create a generalized tree, which could be used to classify both day and night, records in regards to cloud contamination only. The training set was used to develop the Tree based model locating the important binary splits. The remaining 2/3 of the nighttime records were used as the validation set. The validation set provided independent data on which to verify the generality and misclassification rate of the tree model developed from the training set.

Tree variables:

Spectral quantities, which have been previously shown to indicate cloud contamination, were used as variables in the tree based model. The spectral variables used in the tree model were homogeneity (max-min) for AVHRR channels 3, 4, and 5, channel difference for ch3-ch5, ch4-ch3, and ch4-ch5, and linear model values for channels 3 and 5 which correspond to MODIS channels 20, 31 and 32. The limitations of the software (Splus) to identify splits based on optimal linear combinations of variables required the pre-calculation of these linear models using least square regression techniques. The pre-calculations were done for channel 3 as a function of channel 4, channel 3 as a function of channel 5, and channel 5 as a function of channel 4. The coefficients for each of the functions were determined by least square regression using the channel data from all nighttime matchup records defined as having a "good" membership. Finally, the satellite zenith angle was collapsed into 4 groups (0-29, 30-39, 40-49, >50 degrees) and was also included as a variable in Tree estimation.

Tree estimation:

The training set was then recursively partitioned on the above 10 variables and allowed to run to completion. The complete tree for NOAA-9 had 89 terminal nodes and was probably over fitted to the training set. Recall that the recursive binary partitioning technique attempts to make the membership of each terminal node as homogeneous as possible. To limit this over fitting of the tree model to the training dataset, the full tree model was then pruned (Figure 3) to create a more parsimonious tree but still retain a reasonable misclassification rate (8%). The pruned tree model was then used to classify the records in the Validation set. The tree predicted classification for the Validation set was then compared to the actual membership (Table 6). The results from the validation test verify that the pruned tree model was not over fitted to the training set and the misclassification rate remained low. The next step was to repeat verification process using daytime records. This test demonstrated that the final tree model was robust with a misclassification rate of 8.8%. This tree model can then be used to classify NOAA-9 matchup records of unknown quality. Matchup records classified by the tree as "good" are then used in coefficient estimation for the Miami Pathfinder SST algorithm (MPFSST).

Figure 3. Prune classification tree NOAA-9. Boxes represent the terminal nodes of the tree. Number of records classified in each terminal node is shown. The purity of the records contained in the node is shown in parenthesis. Symbols used in the splits represent brightness temperatures; T35= ch3-5, t45=ch3-ch5, T3=ch3.

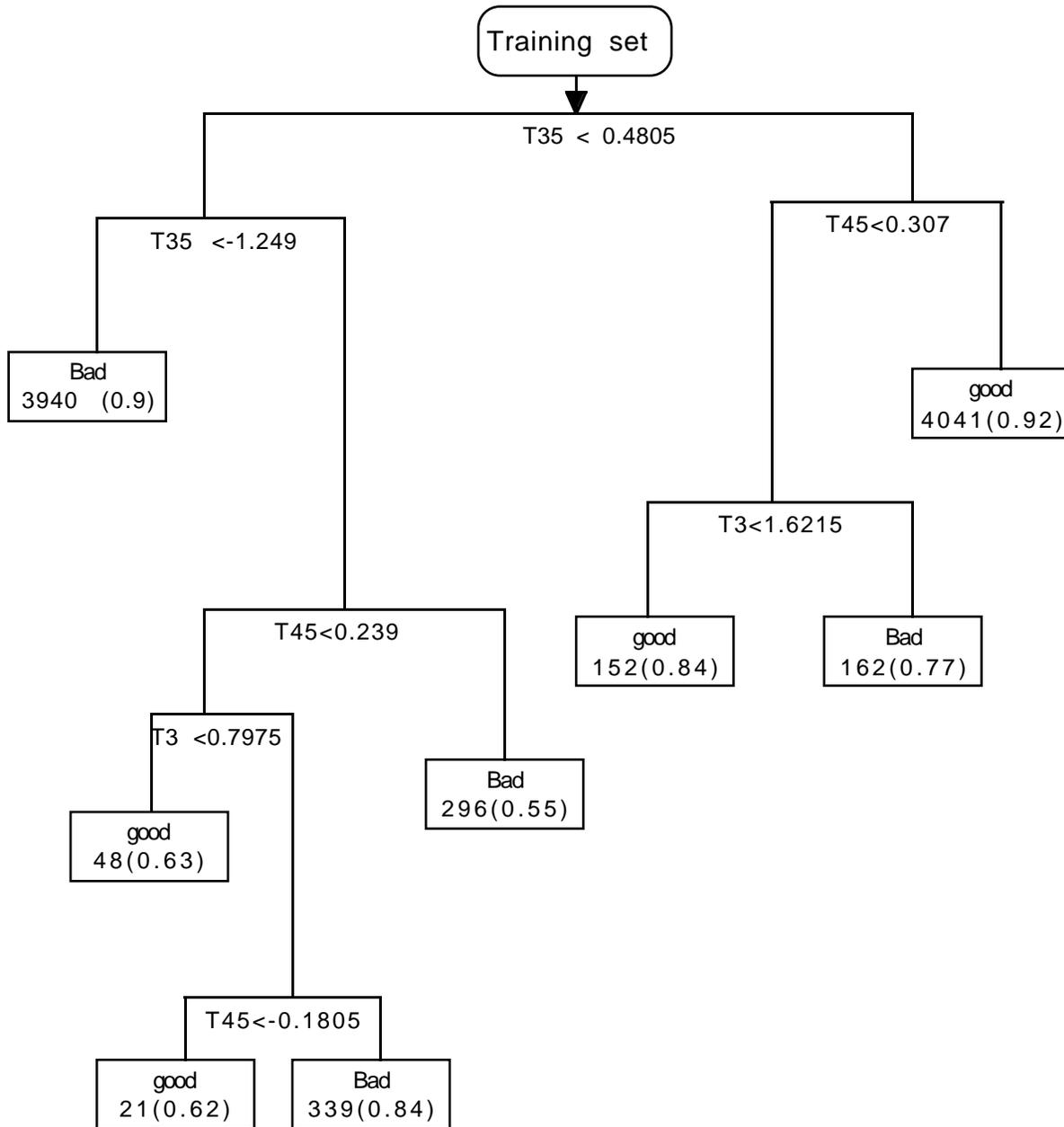


Table 6. Validation test. Cross tabulation of actual classification of record from residuals and predicted classification from pruned tree model. Total misclassification rate is 13% for the validation dataset.

classification category	predicted Bad	predicted good	actual classification
Bad	2522	252	2775
Good	1089	6411	7500
Total	3611	6411	10275

For MODIS we plan to use information from both the decision tree technique and the MODIS cloud mask product to assess the quality of the matchup records. All “cloud-free” and “cloud-contaminated” matchup records will be distributed in the MDB, the matchup record includes a quality flag indicating if the record passed the set of Miami cloud tests. This will enable end users the ability to develop other cloud tests if desired. The final stage of the database is to assemble matchup records into individual monthly files and the addition of ancillary information from other sources. In the PFMDB this includes the Wentz columnar water vapor concentration from SSMI, Reynolds OI weekly climatology, and water column depth from the etopo5 earth topography database. The MODIS ocean color MDB will include additional ancillary data sets for example; aerosol index, aerosol optical depth, and climatology information developed from SeaWiFS and other sensors for a given records time and location.

3.1.2.2 MPFSST Coefficient estimation

Coefficients are determined for a two piece algorithm based on the different atmospheric regimes. The regime is defined by the channel 4-5 difference (T45). A set of coefficients are determined for each month (N) using matchup records from a 5 month (N-2 to N+2) moving window. Coefficients are estimated by an iterative process involving robust linear regression techniques. Coefficients are determined in a 3 steps process.

1. Use robust regression procedures to determine initial coefficient values for the SST equation.
2. Define robustness weights
3. Perform weighted least squared regression procedures based on the robustness weight.

Robust regression procedure:

A robust regression procedure was used to determine the initial coefficients for the Pathfinder SST equation.

$$\text{Pathfinder SST} = C1 + (C2 * \text{ch4}) + C3 * (\text{ch4} - \text{ch5}) * \text{bsst} + C4 * (\text{ch4} - \text{ch5}) * (1 / \cos(\text{satZ})) - 1$$

Where C1:C4 = regression coefficients, bsst = buoy SST, ch= channel brightness temperature, satz= satellite zenith angle.

A robust regression is made using matchup records from each atmospheric regime which were classified as "good" by the cloud flagging routine. The robust procedure is less sensitive to remaining outliers such as those caused by cloud contamination and errors in the buoy SST. The residuals from the robust regression are then used to determine weights for the subsequent least square regression.

Robustness Weights:

The weights used for the least squared regression were the product of the residual and monthly time weighting factors. Residual weights are determined by applying a Bi square function to the residuals of the robust regression. The cut off for the Bi-squared function was set at 6*MAD (median of the absolute value of the residuals). The monthly time weighting factors are set at 0.5 , 0.8, 1.0, 0.8, 0.5 for months N-2 to N+2 when estimating coefficients for month N.

Weighted Least square linear regression:

Coefficients are then estimated for each T45 regime (ch4-ch5 <0.7 and ch4-ch5 > 0.7) for month N. The weighted least squared regression uses the matchup records from the 5 month time window centered on month N and classified as being in the given T45 regime. Only records with weights greater than zero are included in the regression.

3.1.2.3 Transition to MODIS

The MODIS calibration activity presented in this document builds on the work by the MCST. We expect to receive information on the relative calibration of the detectors for each band, calibration dependency on scan angle and the absolute calibration.

Additional information that influences calibration is detector linearity and spectral stability of the filters together with platform and sensor parameters such as navigation, pointing and temperature. The algorithms and tables necessary to normalize outputs of within band detectors and correct to mirror position produced by the MCST will be used in this effort. The prime point of departure will be application of the final calibration coefficients that this algorithm utilizes to produce absolute calibration. We will utilize the vicarious method to adjust the coefficients to achieve agreement between the atmospherically corrected $[L_w(i)]_N$ and the corresponding *in situ* observations. In particular, we will utilize the results obtained from MCST study of the MODIS solar calibrations to test trends obtained from the vicarious method.

Method of application:

Get Level-1 counts. Apply MCST algorithms for normalizing detectors, linearity, mirror correction (scan angle, mirror side, polarization), time correction (sensor drift, temperature). It is expected that most of these corrections will utilize lookup tables. The final conversion of counts to radiance will utilize correction coefficients determined from use of the vicarious method.

The initial version of these tables will be based on pre-launch thermal vacuum test data. After launch methodology tested and validated in the SeaWiFS program will be utilized to update MODIS pre-launch coefficients. This sensor initialization is described in Gordon's ATBD. Finally, if the SeaWiFS, ADEOS/OCTS or COLOR sensors are operational, we will be able to compare retrievals from each of these sensors. These steps are briefly summarized below and defined more fully in the following section: 1)Apply pre-launch calibration data. 2)Obtain radiances, use satellite location and pointing information. 3)Locate satellite observation corresponding to *in-situ* observation. 4)Apply quality control, cloud test procedures. 5a)Visible - Use Gordon's sensor initialization procedure to obtain initial set of calibration coefficients. 5b)Infrared - Use thermal vacuum test data to obtain counts to radiance conversion. 6)Apply appropriate atmospheric correction procedure to obtain geophysical quantity (*e.g.* radiance, temperature). 7)Compare satellite and *in-situ* quantities as a function of observation parameters. Adjust calibration coefficients to obtain agreement.

3.1.3 Variance or uncertainty estimates

3.1.3.1 Observation frequency for useful *In situ* - satellite matchups

Application of the match-up database focuses on changes in sensor performance at time scales longer than six months. The expected scatter in the satellite- *in situ* residuals will be large with respect to changes that will likely occur during normal sensor operations. A number of samples spaced over an extended time period is needed to establish the temporal trends. Table 7 shows the fraction of useful retrievals available for two buoys extracted from the NOAA-9 MDB. If the useful fraction of 25% is indicative of the fraction expected for the MODIS timeframe, then the frequency of MODIS observation of a given buoy location, 1 every 2 days, will yield approximately 90 samples/year. These will be distributed in scan angle by the orbit repeat frequency, every 4 - 8 days depending on magnitude of scan angle yielding 10 - 20 sample/year for a given buoy for a range of scan angles. Data from a number of buoys is required to increase data density and to develop error statistics given the uncertainties present in both the *in situ* and satellite observations. Where multiple buoys are available, the satellite becomes a "transfer standard" permitting relative calibration checks between *in situ* platforms. This approach has been used to check buoy calibrations across retrieval and redeployment events and to determine when drifting buoys have exceeded their useful lifetime.

Table 7. Buoy retrieval statistics for fixed mooring and NOAA-9 AVHRR

Time	Total Observations over 4 years	Cloud-free	% Useful
Buoy 1-all	1731	390	23%
day	868	191	22%
night	863	199	23%
Buoy 2-all 1	686	526	29%
day	955	250	26%
night	831	276	33%

The smaller daytime percentage reflects a loss of coverage due to sunglint. The AVHRR sensor collects full earth coverage within a 24 hour period. Even with this extensive coverage, approximately 75% of the *in-situ* observations do not contribute useful matchups with the satellite due to clouds and sun glint. Matchup retrievals obtained from MODIS will have a lower percentage of useful observations due to the restricted scan coverage where 2 days are required to observe the entire surface of the earth. Thus visible wavelength matchup database will be sparsely populated. These data together with the more extensive data sets obtained during cruise periods form the test data sets that will be used to quantify algorithm/sensor performance.

Other investigators (e.g. Abbott, Lewis) have discussed the possibility of deploying optical drifters. Data provided by this type of instrument have the potential of greatly increasing the number of in-situ observations available to characterize the MODIS sensor and algorithm performance. While these observations will not provide the same level of calibration and continuity as the fixed buoy instruments, they provide a greater range of geographical and geophysical observations than the limited set of fixed buoy observations. In a relative sense, changes in retrieved radiance can be tested against parameters such as scan angle, time, location. Availability of a large number of *in situ* sources provides a dispersed set of reference information that permits changes in retrieved radiance to be traced to the sensor, algorithm, or in-situ location. Brown in his ATBD discusses the difference between skin and bulk SST. Skin temperature is measured using a downward looking radiometer deployed on ships or buoys. While there are likely to be more IR than visible instruments, the overall situation of a limited number of locations observed by radiance instruments will be true for IR observations as well as visible. Data describing the larger geophysical setting for the most part is provided by fixed or drifting buoys measuring bulk temperature with a thermistor mounted at a depth of one to several meters. This larger data set again provides a context check for the limited number of more accurate radiometric measurements.

3.1.3.2 IR Matchup database

AVHRR pathfinder M-AERI Matchups:

The reference *in situ* sea surface temperatures were measured using the Marine-Atmosphere Emitted Radiance Interferometer (M-AERI). This remote-sensing instrument was originally developed at the Space Science and Engineering Center (SSEC) of the University of Wisconsin-Madison to acquire high spectral resolution measurements of atmospheric emitted radiance as part of the Atmospheric Radiation Measurement (ARM) initiative by the Department of Energy. The M-AERI is a passively-sensing, infrared radiometric interferometer. Radiance emitted in, or scattered into, its field of view is reflected into the body of a Bomem interferometer by a mirror angled at 45 degrees to its axis of rotation. Incoming radiation is then directed to a two-detector sandwich composed of InSb and HgCdTe which, when combined, give measurements in the approximate $500\text{-}3000\text{ cm}^{-1}$ (approx 2.5-20 μm) range. The mirror, which is gold-plated to minimize possible polarization effects, rotates to provide scans of both sea and sky at complementary angles to nadir and zenith, respectively. An end-to-end calibration of the measurements is accomplished by having the mirror rotate to view two black bodies, one at 60 °C and the second at ambient temperature, both before and after each sequence of sea and sky views. The black bodies are traceable to a NIST standard thermometer. Results from a recent RSMAS workshop show that the M-AERI is accurate to within 0.01 °C when directed at a NIST black body at 30 °C.

The M-AERI sea surface temperature estimates were subject to procedures and quality controls designed to produce the cleanest possible dataset. Care was taken to position the instrument so that a clear view of the sea was obtained that was not affected by the bow wave while the ship was in motion. Data were not collected during periods of precipitation or at times when sea spray could have contaminated the optics. Nevertheless, variable cloud conditions, sun glint, spray, etc. produced some spurious SST estimates and it was necessary to remove these outliers from the M-AERI SST record. First, data from times when the instrument was covered during periods of spray or precipitation were eliminated, both by manual extraction and by removing all data points for which the corresponding air and sea surface temperatures differed by less than 0.005 °C. Secondly, any M-AERI data that lay outside a +/- 3 °C range from the Reynolds' OISST estimate were eliminated. Thirdly, those M-AERI SST estimates that differed from the ship's TSG measurements outside of an established range were excluded. This range was determined by empirical results from an earlier M-AERI deployment which established that the difference $T_{\text{skin}} - T_{\text{bulk}}$ between the skin and the bulk temperatures under a range of typical wind and sea conditions could be expected to be $-1.8\text{ }^{\circ}\text{C} \leq T_{\text{skin}} - T_{\text{bulk}} \leq 0.5\text{ }^{\circ}\text{C}$. To establish even more strict limits for the purposes of this study, the threshold values for the quality control were taken to be 97 % of this range $-1.75\text{ }^{\circ}\text{C}$ to $0.49\text{ }^{\circ}\text{C}$. And lastly, the standard deviation of each air and sea surface temperature estimate were used to eliminate those measurements where the standard deviation of the air temperature estimate exceeded 0.06 °C or that of the sea surface temperature estimate exceeded 0.09 °C. These procedures were enacted with the hope of producing the cleanest SST dataset possible for comparison to the Pathfinder (and other) SSTs.

Table 8. M-AERI cruise times and locations

Cruise Name	Ship	Year	Begin Day	End Day	Area of Study
Combined Sensor Program (CSP)	NOAAS Discoverer	1996	78	103	Equatorial Western Pacific
Hawaii-New Zealand transect (HNZ)	R/V Roger Revelle	1997	272	286	Central Pacific Meridional Section 24 °N Section
Section 24 °N Section (24N)	NOAAS Ronald H. Brown	1998	8	55	Zonal Section along 24 °N in North Atlantic
GASEX (GSX)	NOAAS Ronald H. Brown	1998	127	188	Mid-latitude North Atlantic
Florida- Panama- Oregon Transit (FPO)	NOAAS Ronald H. Brown	1998	196	210	Florida to Panama to Oregon Transit
North Water Polynya study (NOW)	CCGS Pierre Radisson	1998	150	203	Baffin Bay, Arctic Polynya

The M-AERI locations and times from the cruises listed in Table 8 were used to identify those AVHRR orbits that would provide data coincident and collocated with the M-AERI instrument. The Pathfinder data were mapped at 4 km resolution. For each mapped orbital scene, the closest pixels, of all quality pixels within 4 km and 90 minutes, to the M-AERI instrument were extracted. (Note that it is possible that an individual pixel may have been compared with more than one M-AERI SST record. Quality means that the pixel must have obtained a Pathfinder v4.2 quality flag of 6 or greater (from a range of 1 to 7): the pixel has passed range checks ± 2 from the 3 week weighted average Reynolds fields), and strict homogeneity tests (the maximum

difference between high and low values of the channel 4 and 5 within a 3-by-3 box centered on the target pixel must be less than 0.7 °C (Podesta et. al. 1998).

Some ancillary information was also assembled to aid in the interpretation of the comparison. For each pixel, daily Special Sensor Microwave Imager (SSM/I) water vapor values were averaged via bilinear interpolation to its location to provide some independent idea of the atmospheric water vapor content. Each research vessel's thermosalinograph (TSG) data, with SST computed every 30 seconds, provide a bulk estimate of the SST which can be compared via linear interpolation in time to the M-AERI SST all along the whole cruise track, not just at coincident and contemporary Pathfinder points (these data are not yet available for the NOW98 cruise, necessitating that some statistics be quoted both inclusive and exclusive of this cruise). The values of the weekly Reynolds OISST were also extracted for each target pixel via bilinear interpolation from the 1° fields; these values can also be compared to M-AERI SST along the whole cruise track. The assembly of these other estimates for comparison with the M-AERI SSTs allows the M-AERI-Pathfinder comparisons to be placed in a more familiar context.

The SST from the M-AERI/Pathfinder points from the 1998 GASEX cruise are plotted in Figure 4, which also display the time and location of the M-AERI/Pathfinder points, as well as the continuous records from the TSG and interpolated values from the OISST. Note the scarcity of the M-AERI/Pathfinder matchups as compared to the M-AERI/TSG and M-AERI/OISST data - the NOAA polar orbiting satellites will pass overhead twice a day, and at those times it is necessary for the M-AERI to be in a location of clear skies in order to register as a good M-AERI/Pathfinder comparison.

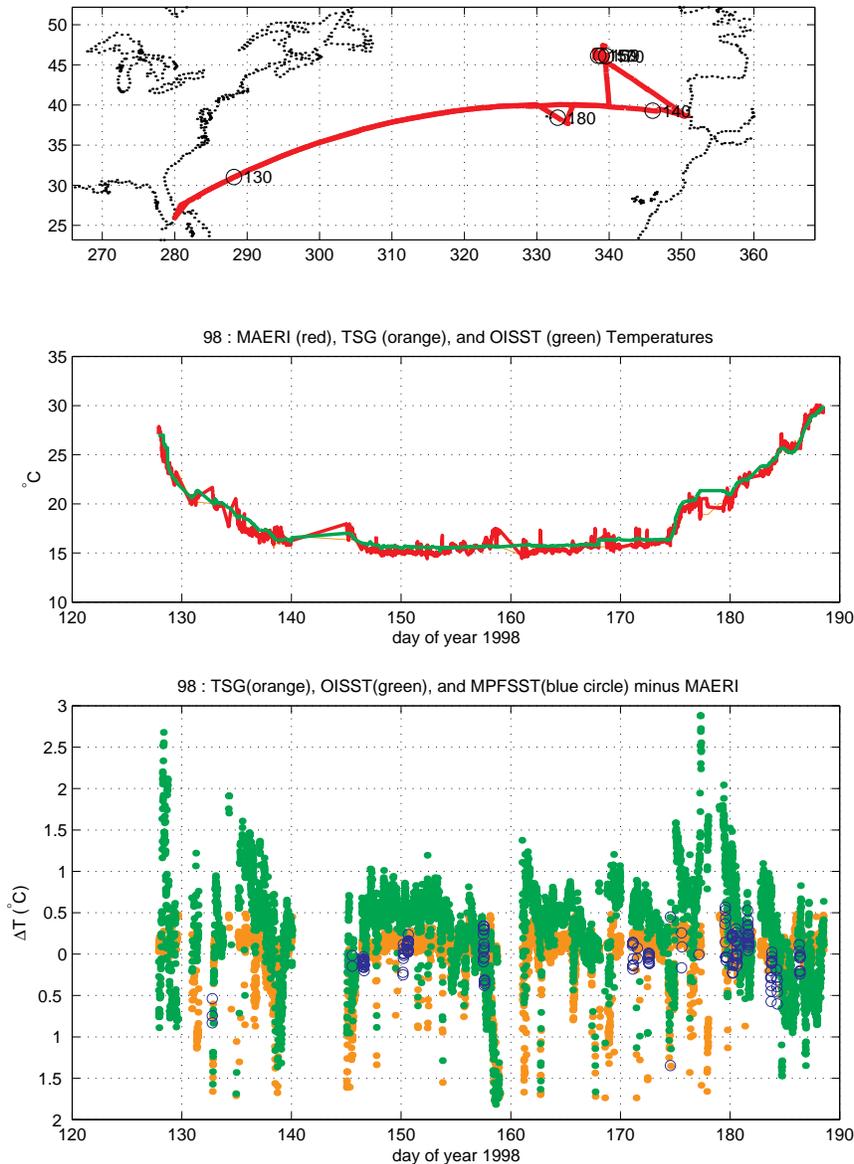


Figure 4 The cruise track and day of year [top panel], continuous MAERI (red), TSG (orange), and OISST (green) sea surface temperature estimates [middle], and the differences between the TSG, OISST, and Pathfinder (blue) SST estimates and the skin temperature measured by MAERI [bottom] for the 1998 GASEX cruise. Note the sparsely of Pathfinder/MAERI matchups due to the twice-daily overflight of the AVHRR instrument coupled with the need for cloud-free conditions to produce quality Pathfinder SSTs for accurate comparisons to MAERI data.

The statistics of the results from the sst, using the M-AERI SST as the reference temperature, from each cruise and as a whole are presented in Table 9 and are depicted graphically in Figure 5. Note that the mean difference, combining all good records from the mid-latitude cruises, between the Pathfinder SST and the M-AERI SST is $0.06\text{ }^{\circ}\text{C}$, with a standard deviation of $0.29\text{ }^{\circ}\text{C}$. Inclusion of the noisier NOW data increase these estimates to $0.13 \pm 0.37\text{ }^{\circ}\text{C}$. The results are biased towards the results from the GASEX and NOW cruises, since the number of data collected during these extended cruises is far greater than the number of data collected during the other cruises.

These results compare favorably with the previously quoted results from the climatological study of Casey and Cornillon 1999 and with the estimates from a similar, though much more exhaustive in space in time, comparison of NOAA-14 MPFSST to PMD buoy/drifter data which provides a mean difference of 0.02 and a standard deviation of 0.53C. The ships' TSGs are closest to the M-AERI SST; those times where there are substantial differences are tied to skin/bulk SST differences at times of high insolation and small wind mixing. Next best is the Pathfinder 4 km resolution estimate, at least 10% worse than the TSG with an additional bias of approximately 0.5 °C. The Reynolds' OISSTs have the largest errors, not at all surprising given the averaging and smoothing inherent in those estimates which introduces error into those weekly estimates.

Table 9. Summary Statistics for M-AERI Matchups

Cruise Description	SST Difference	Mean	Standard Deviation
CSP 1996, N = 23 (1112 total)	TSG - M-AERI	0.07(0.04)	0.10 (0.20)
	OISST - M-AERI	0.20 (0.09)	0.32 (0.45)
	MPFSST- M-AERI	0.16	0.20
HNZ 1997, N = 6 (726 total)	TSG - M-AERI	0.10 (0.14)	0.05 (0.19)
	OISST - M-AERI	0.04 (-0.13)	0.08 (0.49)
	MPFSST- M-AERI	-0.03	0.25
24N 1998, N = 16 (1833 total)	TSG - M-AERI	0.22(0.17)	0.07(0.13)
	OISST - M-AERI	0.05(0.08)	0.42 (0.41)
	MPFSST- M-AERI	0.03	0.18
GASEX 1998, N = 168 (5104 total)	TSG - M-AERI	0.02(0.02)	0.30(0.32)
	OISST - M-AERI	0.32(0.30)	0.47(0.56)
	MPFSST- M-AERI	-0.01	0.25
FPO 1998, N = 47 (1244 total)	TSG - M-AERI	0.14(0.06)	0.19(0.29)
	OISST - M-AERI	0.85(0.37)	0.86(0.71)
	MPFSST- M-AERI	0.27	0.40
NOW 1998 (Arctic), N = 176 (4251 total)	TSG - M-AERI	NA (NA)	NA (NA)
	OISST - M-AERI	-0.79 (-1.11)	0.57 (0.82)
	MPFSST- M-AERI	0.24	0.44
Total, all data, N = 436 (total 14277)	OISST - M-AERI	-0.08 (-0.18)	0.82 (0.89)
	MPFSST - M-AERI	0.13	0.37
Total, excluding NOW data, N = 260 (total 10015)	TSG - M-AERI	0.06 (0.06)	0.26 (0.28)
	OISST - M-AERI	0.38 (0.21)	0.58 (0.56)
	MPFSST- M-AERI	0.06	0.29

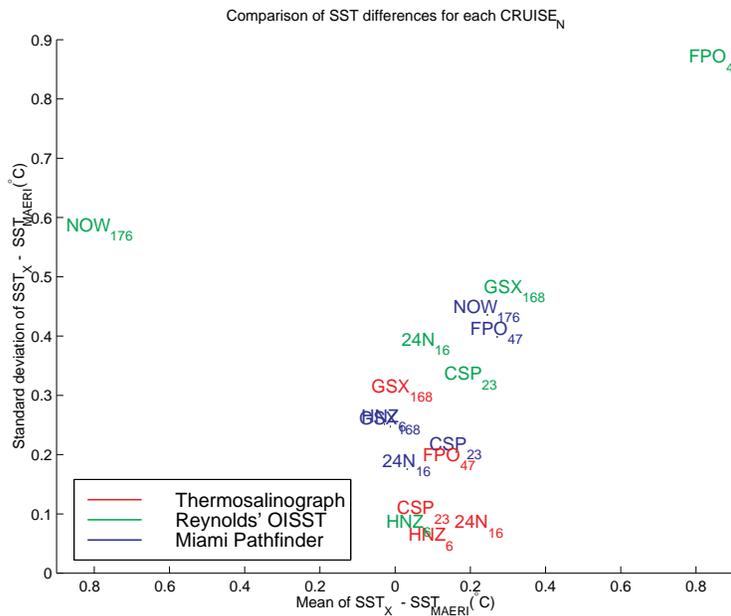


Figure 5 The mean (abscissa) and standard deviation (ordinate) of the difference of the various SST estimates from the reference MAERI SST. The cruise abbreviation (see Table 8) is centered on the point with the number of observations for that cruise appearing as a subscript. The TSG data are generally most accurate, followed by Pathfinder SST and lastly Reynolds' OISST. The OISST outliers are the result of poor OISST boundary conditions in the Arctic (NOW) and along the west coast of North America (FPO).

Although the number of M-AERI/Pathfinder points is relatively small, these results suggest that the Miami Pathfinder algorithm is much more accurate than has been estimated by previous studies - at least for those atmospheric and oceanic conditions sampled by these 6 cruises. The fact that Pathfinder SSTs are nearly as good as the thermosalinographs of these research vessels is very encouraging for global SST studies using AVHRR data. The larger error exhibited by the Reynolds' OISST fields gives a good indication of the minimum temperature difference necessary for meaningful interpretation of that difference when making Pathfinder SST vs. Reynolds' OISST comparisons, which is often done since the Reynolds' fields are global and are easily available.

The results of the comparisons from the Arctic NOW98 cruise have enhanced error for a number of reasons. The lack of the TSG data and a meaningful Reynolds' OISST field hindered the M-AERI quality control effort as MAERI outliers are more difficult to identify. More important than this is the fact that the Pathfinder algorithm may not perform well in the Arctic due to a lack of *in situ* buoy/drifter data in that latitude range with which to calculate the appropriate Pathfinder coefficients, a poor first guess field provided by the Reynolds' OISST average, and the nearby presence of sea ice may adversely affect the AVHRR retrievals.

Given the small size of the current M-AERI dataset, the accurate portrayal of the effects of clouds, water vapor, and aerosols on the Pathfinder SST retrievals will have to wait

for future M-AERI deployments in a greater range of atmospheric conditions. Nevertheless it is instructive to show a few of these relationships. Figure 6 shows the effect of water vapor for oblique scan angles. While all points are considered high quality with regard to the pixels' homogeneity, the Pathfinder/M-AERI difference versus the SSM/I water vapor estimate suggests a tendency for the large scan angle (>45°, quality level 6) to underestimate the SST with increasing integrated water vapor content. While there are not enough independent data to accurately model this relationship in this study, this effect has also been noted by (Kumar 1999).

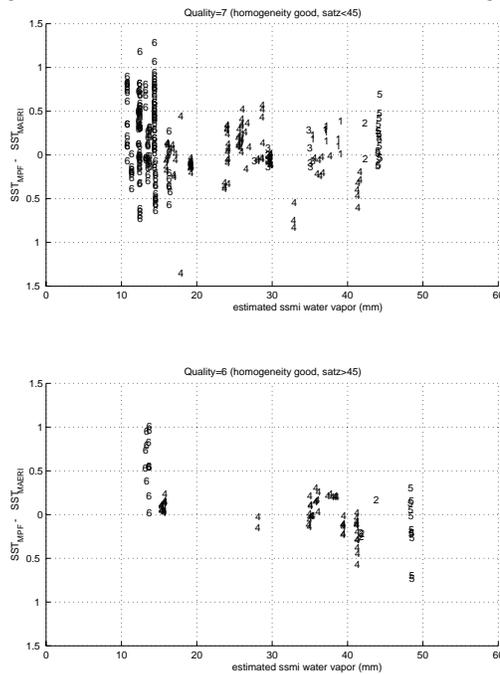


Figure 6 The relationship between water vapor content as estimated from the SSM/I instrument versus the Pathfinder-MAERI difference for quality level 7 pixels [top panel] and quality level 6 pixels [bottom]. The level 6 pixels differ from the level 7 pixels only in that they are derived from observations with an oblique (>45 degrees) viewing angle. An oblique viewing angle leads to a longer path length through the atmosphere which makes the SST estimate more susceptible to the effects of atmospheric water vapor, resulting in a trend towards negative residuals with increasing water vapor for the level 6 pixels. The numbers in the plots are keyed to the cruises as follows: 1=CSP, 2=HNZ, 3=24N, 4=GSX, 5=FPO, 6=NOW. See Table 8 for the full description of these cruises.

Simulated MODIS IR Sea-surface temperature Matchups:

The MODIS will produce two SST products; a product using the 11 and 12um bands (SST) and a second product using the 3 and 4um bands (SST4). During the past 6 months in conjunction with Drs. Peter Minnett and Richard Sikorski we have created a

simulated matchup database for the MODIS IR channels using radiosonde data, a modified Rutherford -Appleton IR-spectral model updated with recently supplied Clough continuum spectra for vapor optical properties, and the latest MODIS spectral response characteristics to produce channel data. This simulated matchup database is being used to develop at-launch algorithm coefficients for the SST algorithm and evaluate current and alternative formulations of the SST algorithms in light of anticipated channel cross talk and calibration uncertainties.

The modified RAL was used with a global dataset of 761 marine and coastal radiosondes to simulate satellite-viewed brightness temperatures (BTs) for the currently available response functions (RSRs) Table 10 for MODIS AM-1 3-micron and 4-micron IR bands (B20, B22, and B23). New algorithms were developed for SST4 retrieval for a zenith viewing angle using the 3 and 4 um bands, and new relationships were observed for retrieval of total column water vapor.

Table 10. MODIS response function

Band	Center width(nm)	Bandwidth(nm) From 1% to 1%
B20	3788.2	182.6
B22	3971.9	88.2
B23	4056.7	87.8

We based the algorithm on a simple channel difference method. We applied a seasonal correction (Figure 7), and a latitudinal band correction. Initial regressions showed a strong zenith angle dependence, which may be included that in future algorithms.

Algorithm (equivalent to channel plus channel difference):

$$SST4 = a + b * B22 + c * B23 + f(x)$$

(note: B23 may be replaced by B20 if warranted, using the appropriate set of coefficients. At present the advantage of the B22:B23 pair over the B20:B22 pair is quite small.)

Seasonal term (based on solar declination):

$$f(x) = m * \cos(2 * 3.14159 * (x + n) / 365) + p$$

Definitions:

a,b,c,m,n,p are coefficients estimated separately for each of 3 latitudinal bands based distance from the equator.

x(northern hemisphere)=days after 173 (summer solstice)

x(southern hemisphere)=days after 357 (winter solstice)

B20 = MODIS Band 20

B22 = MODIS Band 22

B23 = MODIS Band 23

for leap years, standard year days = leap year days *365/366

Residuals (reference SST versus algorithm derived SST4) showed an RMS of 0.269 and 0.285 ,respectively for bands 22/23 and 20/22 formulations. Analyses during algorithm development revealed that certain band differences are a good proxy for total column water vapor. Plotting the regression residuals vs. radiosonde total vapor, the relationship is best for B20-B22 = 1 to -1 degrees versus 0 to 6 g/cm2 vapor. It is similar, but noisier (especially drier atmospheres), for B20-B23 = 2 to -0.5 degrees versus 0 to 6g /cm2 vapor. The B22-B23 difference shows virtually no slope vs. vapor load, and is noisy for drier atmospheres.

Figure 7 Seasonal correction function: modeled SST4 residuals versus day of year. Dots- simulated matchups residuals prior to addition of seasonal correction to the algorithm; Squares- fitted seasonal correction function is shown.

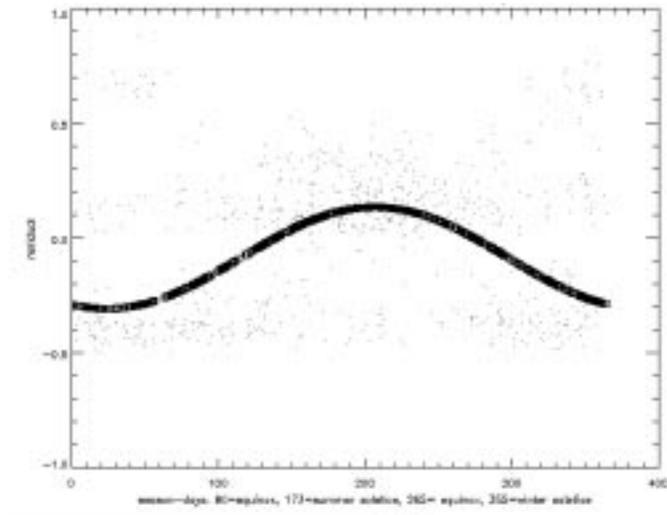
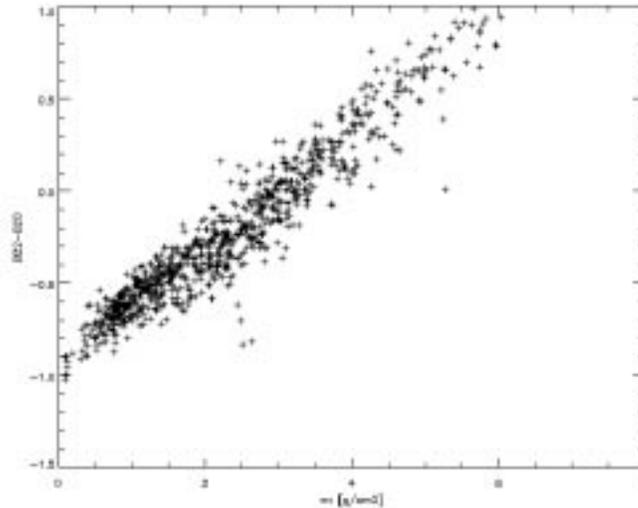


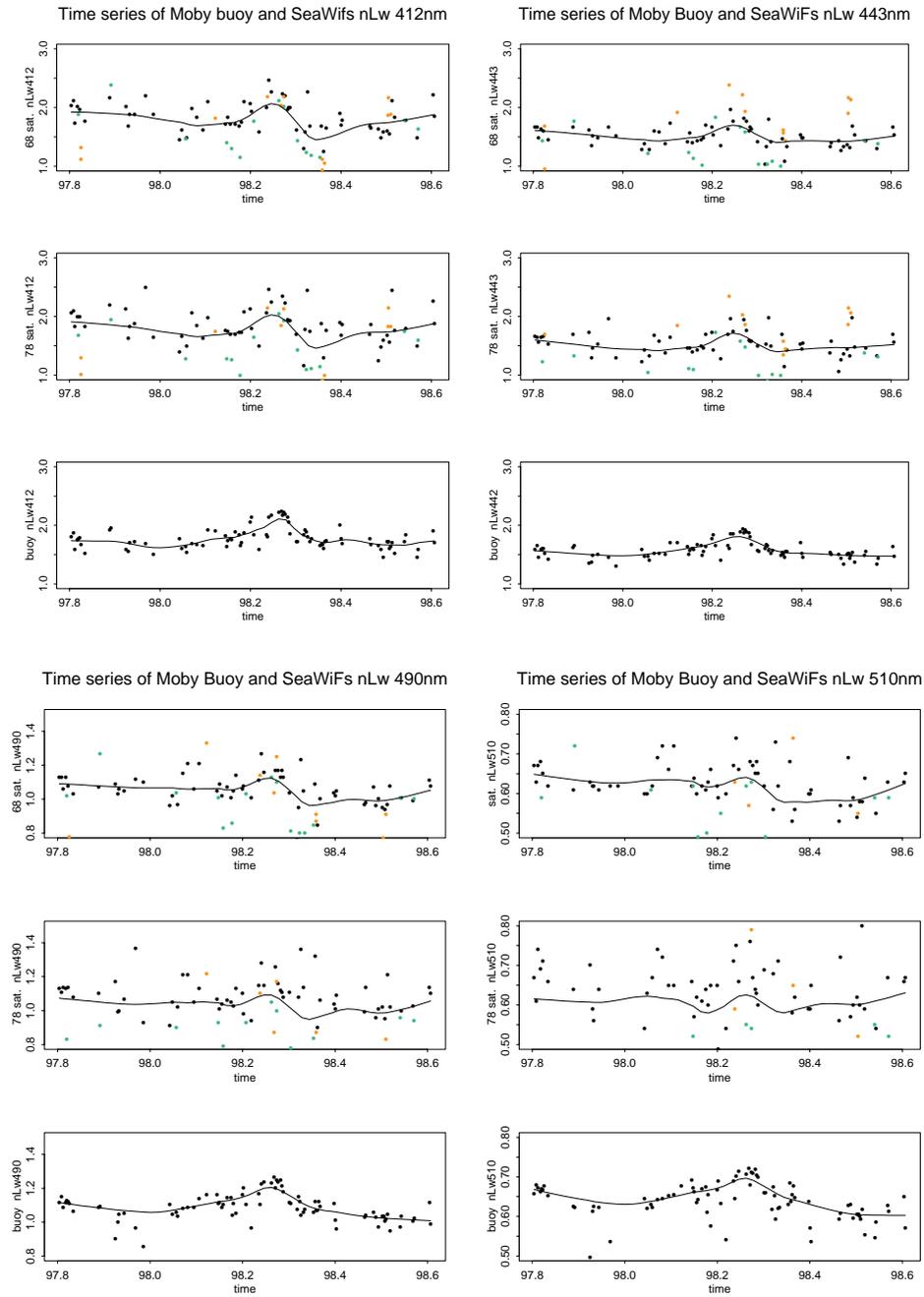
Figure 8 simulated Band 22 – Band 20 versus water vapor



3.1.3.2.1 Ocean Color Matchup databases

An ocean color matchup database developed from SeaWiFS satellite data and Dennis Clark's MOBY buoy is presented below. In this application changes were made to our local SeaWiFS processing stream to incorporate recent developments in calibration and atmospheric model selection in collaboration with Howard Gordon at the University of Miami Physics department. These changes included new calibration values in channels 1-7 (generally on the order of a 1% lower) and the addition of a new oceanic atmospheric model. This new atmospheric model assumes only a single mode of large particles associated with breaking waves. The current maritime atmospheric models contain two modes, big and small particles. We implemented the above changes in our processing stream and extracted the satellite information for a 3x3 km box located over the MOBY Hawaii location to examine the impact of these changes on the accuracy of the atmospheric correction and nLw retrieval. We also evaluated the use of SeaWiFS bands 7 and 8 versus bands 6 and 8 for the atmospheric correction. Figure 9 shows the comparison of the retrieved satellite water leaving radiances is made using the MOBY buoy deployed by Dennis Clark off the island of Lanai, Hawaii. A time series of cloud free satellite retrievals for each of the SeaWiFS bands are presented for both the 6/8 and 7/8 band pair based atmospheric correction and for the buoy. The observations are color coded to reflect the observation pathlength; black 1-2, green 2-3, and yellow >3 atmospheres. Each line is a loess fit to the observations. There is less scatter for the 6/8 correction as well as lower scatter for the shorter pathlength observations.

Figure 9 Time series of SeaWiFS and MOBY buoy nLw's using both 6/8 and 7/8 Atmospheric correction Processing. The observations are color code to reflect the observation pathlength; black 1-2, green 2-3, and yellow >3 atmospheres. Each line is a loess fit to the observations.



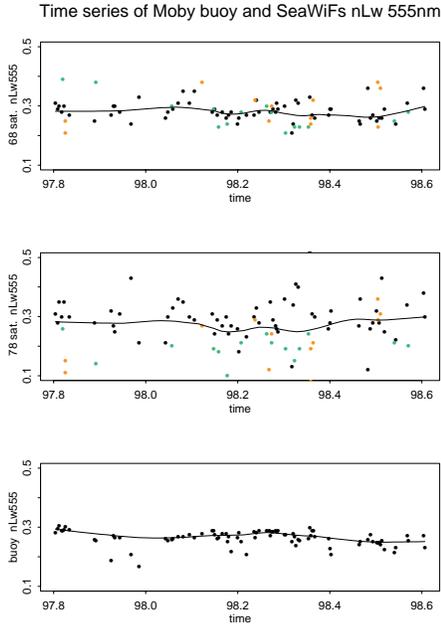


Table 11 presents the median and standard deviations for the 6/8 and 7/8 retrievals and buoy observations relative to the MOBY loess fit. The comparisons are similar to the MOCE initialization results; bias is order of 1% and standard deviation, 14%, is approximately twice that seen with the MOCE (ship based observations). The MOBY observations have order of 6% standard deviation relative to the loess fit to the MOBY observation.

Table 11. Percent difference of SeaWiFS – Buoy nLw’s MOBY time series

Band		processing			median Moby nLw
		6/8	7/8	buoy	
555 nm	median	3.7%	5.6%	0.1%	0.268
	std	17.2%	31.3%	8.2%	
510 nm	median	-3.6%	-2.6%	-0.3%	0.643
	std	25.0%	28.0%	6.4%	
490 nm	median	-4.2%	-4.4%	-0.8%	1.088
	std	13.8%	17.8%	5.2%	
442 nm	median	-3.3%	-4.6%	-1.0%	1.555
	std	15.9%	18.2%	6.4%	
412 nm	median	-0.6%	0.2%	-1.3%	1.731
	std	16.1%	19.0%	8.1%	

3.1.3.3 Error sources

Error calculations for the MODIS ocean color atmospheric correction approach are presented in H. Gordon’s ATBD and will not be reproduced here. Similar calculations

are provided by O. Brown for the SST algorithms. Application of the matchup database will be directed towards trying to differentiate between errors due to the atmospheric correction process, sensor characterization and *in situ* measurements. We will utilize comparisons between expected cross scan behavior for the 500 nm bands and observed change to indicate change in mirror reflectivity or polarization effects. Figures are presented later showing behavior of the long wavelength AVHRR IR bands as a function of scan angle and latitude. Comparisons of similar fields obtained from SeaWiFS, AVHRR and MODIS will be used in conjunction with ancillary data to identify locations and conditions where sensor and algorithm performance is satisfactory. Where not, the data sets will help identify potential sources of error.

The MOCEANS strategy consists of two-tiered QA of Oceans processing. The top level QA is done at the MOCEANS processing facility to detect the gross errors in processing. Data products will be checked for completeness of the level2 and level3 production by visual examination of the 40 level 3 daily products to determine if data is missing in the L3 file. The Q/A analyst must trace the problem to determine where in the processing stream the drop out occurred. Approximately 6 months post-launch activities are anticipated to move toward zero-order science QA tasks. These include: checks and trending performance of pixel level QA and metadata summary flag performance and values. Checks will be made to determine data consistency with climatology comparison/correlation with other MODIS products. These checks generally focus on verifying “reasonableness” of the data and identifying location of gross algorithm failure.

The second tier of the quality assurance plan is Post processing Science QA done at RSMAS SCF and MOCEAN team member institutions. Science QA procedures will examine both pixel and global context with the goal of understanding differences due to instrumental, code/algorithm, geophysical, and biological effects. Primary output of this effort will be revisions in the criteria and thresholds used to define and set run-time pixel level flags, and rules for using pixel flags at level 2 to control acceptability for binning level 2 pixels into level 3 fields and establishing the confidence flags.

A hierarchy of Science QA procedures have been defined with complex QA procedures applied less frequently than more simple QA procedures. Science QA procedures will be applied most intensively where processing QA failures occur and where algorithm understanding can predict expected problems. They will also be applied to a regular sample of each data product in a spatially and temporally stratified manner coincident with expected natural variations in geophysical and atmospheric phenomena. MOCEAN products will require examination of global or partial samples of the globe and time series will be extracted from some products for trending analyses.

Science QA procedures will examine both pixel and global context with the goal of understanding differences due to instrumental, code/algorithm, geophysical, and biological effects. Pixel level context: Science quality flags and product specific criteria and thresholds Global context: SeaWiFS, AVHRR same day retrieval, Climatologies

developed from both MODIS and other sensors, MODIS atmosphere products, last weeks MODIS product.

Coordination of end-to-end MOCEAN QA of all MOCEAN products will be performed periodically at common locations distributed across the globe. This will be useful for resolving data dependency problems expected immediately after launch and after algorithm updated and will provide a mechanism to verify that the configured algorithms behave in the same way as the scientific algorithms. Common locations will include those selected under the MOCEANS validation plan. QA long-term validation and comparison with in-situ observations will occur at the RSMAS SCF with the Match-up Data Bases to assign confidence/uncertainty limits of the products.

3.2 Practical Considerations

3.2.1 Overview

Processing time for the ocean color algorithms is dominated by the atmospheric correction step. Experience using the SeaWiFS test data set has provided an opportunity to evaluate potential mechanisms to minimize processing time. One method that has been incorporated is to retain knowledge of the aerosol models chosen to be appropriate at the previous pixel and assume that the same models work at the current pixel. This assumption is tested, and if found to be true, results in a considerable savings since only two rather than twenty aerosols need to be evaluated. Programs have been developed and submitted (Version 2) to the MODIS team (SDST) for compliance testing and delivery to the MODAPS. Programs have been coded using C and Fortran 90 using EOS toolkit support routines and HDF-EOS file structures. There are two Level-2 routines, one for SST and another for ocean color. Routines have been integrated from Brown for SST, Gordon for atmospheric correction, Carder for chlorophyll and other Case II water products, Abbott for fluorescence line height and chlorophyll fluorescence efficiency, Hoge for PUB and PEB, Clark for various Case I water properties. Binning programs have been developed to generate Level-3 equal area files for each of the products. These files can be binned into desired space and time resolutions required to meet both EOS standard and investigator specific resolutions. Finally a mapping program is available to produce gridded products.

3.2.2 Programming/Procedural Considerations

3.2.2.1 Overview

Processing estimates are currently based on Version 2 ocean color and SST PGEs executing on SGI Origin R10000 200Mhz processors using IRIX 6.5, EOS toolkit calls and HDF-EOS file structures and AVHRR or SeaWiFS->MODIS Level-1b input data. Granule processing times are averaged based on processing a week of real satellite observations and averaging to obtain processing times for a day's worth of granules. The Version 2 programs have integrated the latest versions of the atmospheric correction and product algorithms supplied by the MODIS ocean team investigators.

3.2.2.2 SST and ocean color processing requirements

Processing times for SST and ocean color Level-2 and Level-3 programs are presented in Table 12.

Processing estimates forecast in the Version 0-2 ATBD allowed 5 100 megaflop processors and 12 100 megaflop processors respectively for the SST and ocean color algorithms. The present estimates are consistent with the original forecast since the 288 processor hours required for Level-2 ocean color can be delivered by 16 processors.

3.2.2.4 Overall ocean processing time for the Level-2 radiances

The processing estimates presented in previous ATBD's utilized a test data set that assumed no clouds or land. In addition, the at-sensor radiances/reflectances presently are not well matched to the atmospheric correction algorithms and consequently did not take advantage of provisions made in the codes to minimize processing time by utilizing pixel to pixel coherence in choice of potential atmospheric correction models. The previous Level-2 ocean color processing times were overstated by as much as a factor of three.

Revised estimates are presented in Table 12. These estimates are based on SeaWiFS and AVHRR data that was converted to the MODIS L1B format and processed with the MODIS Version 2 Oceans code. These values give a more realistic estimate of the processing times required to produce a day of MODIS oceans Level 2 and Level 3 data.

Table 12. Processing times for L2 and L3 Oceans products

Process	CPU hours per data-day produced	number of processors
Ocean Color Level 2 and space binning	20	8
Sea Surface temperature Level 2 and space binning	20	2
Level 3 40 daily binned products and 3 map resolutions	20	3
De-cloud 40 daily products and creation of 3-week reference	20	3

Our goal has been to improve processing capacity through a combination of algorithm efficiency improvements and upgraded hardware to achieve a capability to process reduced resolution (4km) MODIS observation at a rate that exceeds 10 data days/day. This capability will permit changes such as algorithm enhancements or updated calibration to be easily verified using long time series of global observations (> several

years) prior to submitting the updates to the MODIS team for consideration. In particular, if we can process high resolution global data at a rate of 1 data day/day, we will meet the goal to process 4km reduced resolution data at a rate > 10 data days/day. Finally the algorithms are continually reviewed to determine where computation efficiencies can be achieved.

3.2.2.5 Data Volume

Table 13 presents a list of the products that will be produced by the Level 2 and Level 3 PGEs for ocean color and SST.

Table 13. Summary of File sizes for SST and ocean color Archived products.

	Files/day	GB/file	GB/day
Level 2			
Level 2 OC QFlags	144	0.19790	28.50
Level 2 OC Data	432	0.15667	67.68
Level 2 SST Qflags	288	0.12369	35.62
Level 2 SST Data	288	0.04398	12.67
Level 3 Daily Binned 4.6 km			
Ocean Color 36 prods (mean, sd,n,qual, flags)	36	0.62000	22.32
SST (day, night, 4m, 11m) mean, sd,n,qual, flags)	4	0.64000	2.56
Level 3 Weekly Binned, 4.6 km			
Ocean Color	5	0.62000	2.79
SST	1	0.64000	0.32
MAPS EOS grid (13 products x 8 fields and 27 products * 7 fields)			
4.6 daily	293	0.13400	21.80
36 km daily	293	0.00220	0.33
1 degree daily	293	0.00030	0.05
4.6 km weekly	37	0.13400	2.73
36 km weekly	37	0.00220	0.05
1 degree weekly	37	0.00030	0.01
Total	2186		197

Table 14. Table showing MODIS product number, names and ESDT relationships

MODIS Prod ID	Field Description	DAAC L2-ESDT	DAAC L3 binned ESDT	DAAC L3 maps ESDT			
MOD 18	QC	MODOCQC	MODOCQC				
MOD 18	nLw, 412	MODOCI2	MODOCX where time resolution x=R,D,W R= 3-week reference D= daily W=weekly	MOXTD where space resolution x= 04,36,1D 4km=04 36km=36 1 degree= 1D and parameter T=M,S,N,Q,F,1,2,3 M=mean S= standard deviation N=count Q=quality flag F=common flag 1=flag 1 2=flag 2 3=flag3			
MOD 18	nLw, 443						
MOD 18	nLw, 490						
MOD 18	nLw, 531						
MOD 18	nLw, 551						
MOD 18	nLw, 667						
MOD 18	nLw, 678						
MOD 37	Tau aerosol,865						
MOD 37	epsilon (765/865)						
MOD 37	Aerosol model 1						
MOD 37	Aerosol model 2						
MOD 39	epsilon for clear water, 531						
MOD 19	CZCS Pigment				MODOCL2A		
MOD 19	Chlorophyll_MODIS						
MOD 19	Total pigment - case 1						
MOD 20	Fluorescence Line Height						
MOD 20	Fluorescence Baseline						
MOD 20	Fluorescence Efficiency						
MOD 23	Suspended Solids Conc.						
MOD 25	Pig. conc. in coccolith. blooms						
MOD 25	coccolith concentration						
MOD 25	calcite concentration						
MOD 26	Diffuse Attenuation (K_490)						
MOD 31	Phycoerythrobilin (PEB)						
MOD 31	Phycourobilin (PUB)						
MOD 21	Chlorophyll_a (SeaWiFS)	MODOCL2B					
MOD 21	Chlorophyll a (semianalytic)						
MOD 22	Instantaneous PAR						
MOD 22	Absorb. Radn by Phyto. (ARP)						
MOD 24	Gelbstoffe absorption coef.						
MOD 36	Chl. a. abs						
MOD 36	total absorption,412						
MOD 36	total absorption,443						
MOD 36	total absorption,488						
MOD 36	total absorption,412						
MOD 36	Total absorption,551						
MOD 28	QC	MOD28QC					
MOD 28	SST (skin, bulk), Day	MOD28L2	MOD28X				
MOD 28	SST (skin,bulk), Day, 4 m		where x=				
MOD 28	SST (skin,bulk), Night		R,D,W				
MOD 28	SST(skin, bulk), Night, 4 m						

3.2.2.5.1 Volume of data to be transferred to Miami

Data volumes that will be required in Miami have been computed using various approaches. With the advent of MODAPS processing the Miami SCF will require a complete set of L1A, geolocation, cloud mask, and ancillary data for parallel processing and algorithm development and a subset of the DAAC L2 and L3 archived data. Due to the large volume of data ~130GB per day several options were explored to facilitate this transfer. The current plan assumes that the Miami SCF will receive from the DAAC reduced L1a files (PGE71,filename:MOD01_L1ASS) which contain only the 20 bands required for ocean processing (ch5-16, ch20, ch22, ch23, ch26, ch27, ch29, ch31, ch32). The geolocation and cloudmask will be recreated at the Miami SCF on a dedicated workstation which will reduce both bandwidth and archive costs. This will require that the land/water mask and the level-3 snow and ice products also be transferred.

3.2.2.5.2 Total volume

Total volume requirements will range between 30 and 100 gigabytes/day depending on available DAAC and MODAPS system services.

3.2.2.6 Network Capacity

The previous analysis suggests 50 gigabytes will be exchanged each day which would require a sustained data rate of approximately 5 megabits/sec. Assuming a link efficiency of 25% to 50%, a link speed of 10 to 20 megabits/sec is necessary. The link to Miami will provide data for three MODIS investigators: Gordon, Brown and Evans. The link will also support access to the TLMCF and for exchange of data and products between the DAAC, TLFC, MOTCF and other ocean SCF's. Present network access is provided by aDS3 (45 megabits/sec) circuit via vBNS to NISN. To support this bandwidth requirement for the larger volume, the external network consisting of vBNS/Internet2 will be upgraded at the Miami SCF in the spring of 1999. This network will increase capacity from 45 to 155 megabits with a sustaining throughput on the order of 2 megabytes/sec over a 24 hr period required to transfer L1A and selected L2 and L3 products for QA analysis.

3.2.3 Calibration Validation

3.2.3.1 Introduction

The matchup database provides a basis for comparison of coincident satellite and *in situ* observations. Experience gained from application of this type of data during the CZCS mission is presented as a guide for a MODIS strategy. The procedures presented below will be modified to reflect experience gained from SeaWiFS where both the *in situ* (buoy and ship) and satellite sensor data closely reflect the data flow that will be available for MODIS. Data included in the *in situ* portion of the matchup database will be calibrated and validated by the data suppliers. The calibration time series will be compared with similar calculations produced by the MCST using the MODIS on board calibrators.

Finally the time series will be referenced to the lunar and solar time series produced by the MCST.

The vicarious method will be augmented by using space and time series comparisons assembled using the 4km reduced resolution global and selected area 1km data sets to enable testing of long term calibration trends and evaluation of new algorithms to be performed. As CPU performance increases, we expect to be able to run data at rates of 10-20 subsetted data days/day on global data sets and thus have the ability to easily produce extensive time series that will form the basis for comparison of effects of application of calibration or algorithm changes. These data sets would be available to test other algorithms dependent on MODIS ocean water leaving radiances.

3.2.3.2 Satellite Field and Analysis Techniques

Comparisons to assess algorithm performance are not limited to results obtained using the MDB. The following sequence of figures demonstrates a series of comparisons based on analyzing fields obtained from applying the geophysical algorithm to global fields. This sequence is based on processing AVHRR infrared retrievals for the period 1987 through 1993 as part of the AVHRR Ocean Pathfinder project. Use of global fields permits us to address some of the comments raised by the ATBD review where the reviewers recognized the limited temporal and geographical coverage afforded by the *in situ* data sets and their inability to completely address algorithm performance throughout the satellite data set.

A multi-year time series of satellite fields has been assembled and compared to itself in time as well as to a 'blended analysis' produced by R. Reynolds of the NOAA Climate Analysis Center. Reynolds uses a combination of ship and buoy *in situ* observations to correct NOAA satellite observations using an Optimal Interpolation technique. The Pathfinder fields use only buoy observations to determine algorithm coefficients and no further constraints are applied. Due to the filters applied while constructing the MDB, approximately 90% of the buoy observations are removed from the MDB leaving the remaining buoy observations as an independent comparison data set. These observations are used to establish local, regional, and temporal algorithm behavior by using an Objective Analysis technique to compute complete fields (cloud fill) using only satellite data. The OA'ed fields are then compared with the buoy observations (not shown).

Figure 10 shows one week (1991-week12) of NOAA-11 Pathfinder SST mapped to a standard equal angle 1 degree grid. Figure 11 shows the corresponding field produced by the Reynolds blended analysis. These weekly fields form the basis for much of the analysis. Figure 10 shows coverage typical of a weekly period. Work with shorter time period fields must be able to deal with a nominal 25% useful areal coverage or requires some form of gap filling procedure such as the Reynolds Optimal Interpolation or Mariano's Objective Analysis programs. Figure 10 present the basic, non-interpolated, satellite retrievals for 1991, week 12. The field is composed of all ascending 4km AVHRR GAC (Global Area Coverage) that passed the cloud tests during the week. The

data is assembled as a series of daily files and mapped to an equal area 9km bin projection. The daily files are summed in space and time to create the 1 degree equal angle weekly map corresponding to the Reynolds analysis.

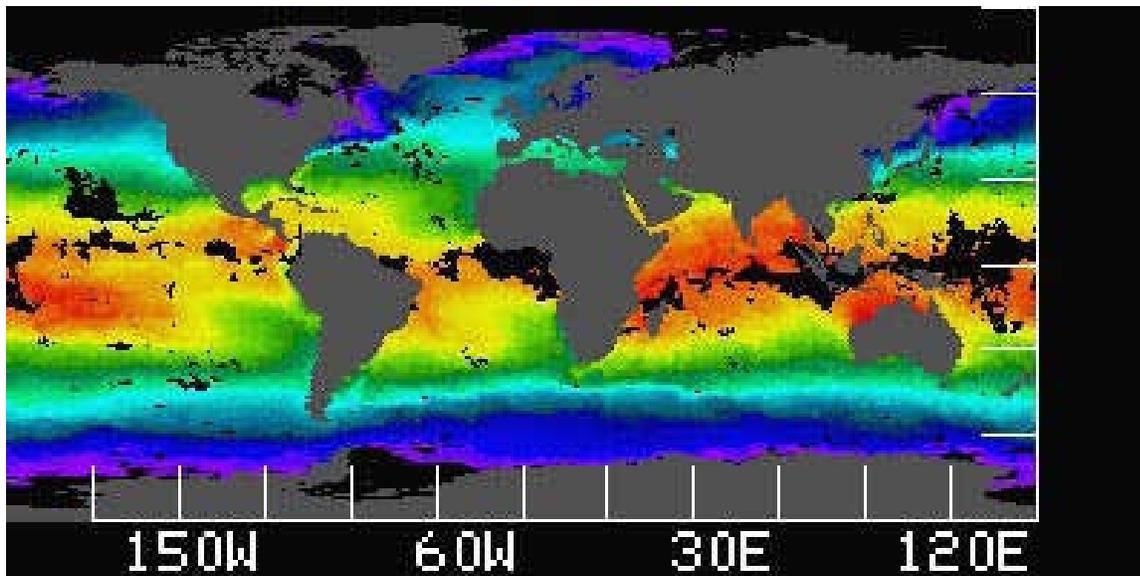


Figure 10. Pathfinder SST image for 1991, week 12. Black areas were continuously cloudy for the week; gray regions are land. Blues areas are cold SST while red regions are warm.

The Reynolds blended analysis, Figure 11, is generated from NOAA satellite retrievals for the week. The retrievals are sorted into 1 degree bins and averaged. An Optimal Interpolation step that incorporates ship and drifting buoy data is used to correct the satellite retrievals and interpolate through regions that had no retrievals. Polar region and area south of the equator are sparsely represented in the ship data and are more subject to biases introduced by a particular ship or represent only the NOAA satellite data.

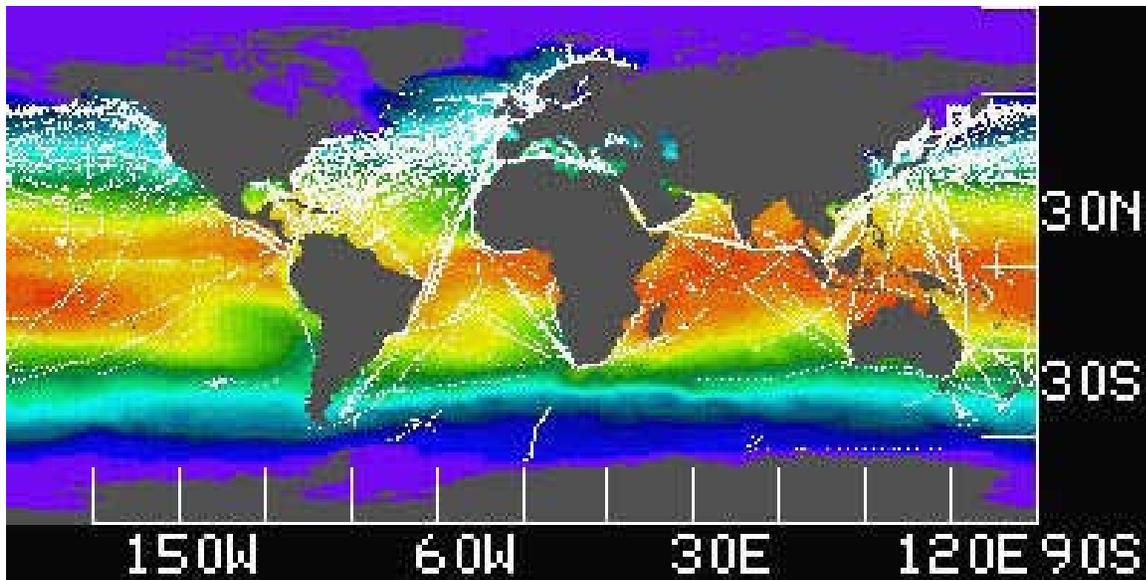


Figure 11. Reynolds blended analysis for 1991, week 12. White areas are locations of ship observations used as the reference for the Optimal Interpolation correction of the NOAA SST retrievals.

Figure 12 present the difference map between the Pathfinder and Reynolds analysis for week 12 of 1991. The gray areas, differences $< 0.33\text{C}$, represent the regions of agreement between the two products. The regions where differences are larger are seasonal. Research is underway to try and establish how the differences are partitioned due to factors such as the process of developing each product, the nature of the satellite SST retrieval process and geophysical considerations such as ocean stratification, distribution of atmospheric water vapor, wind...

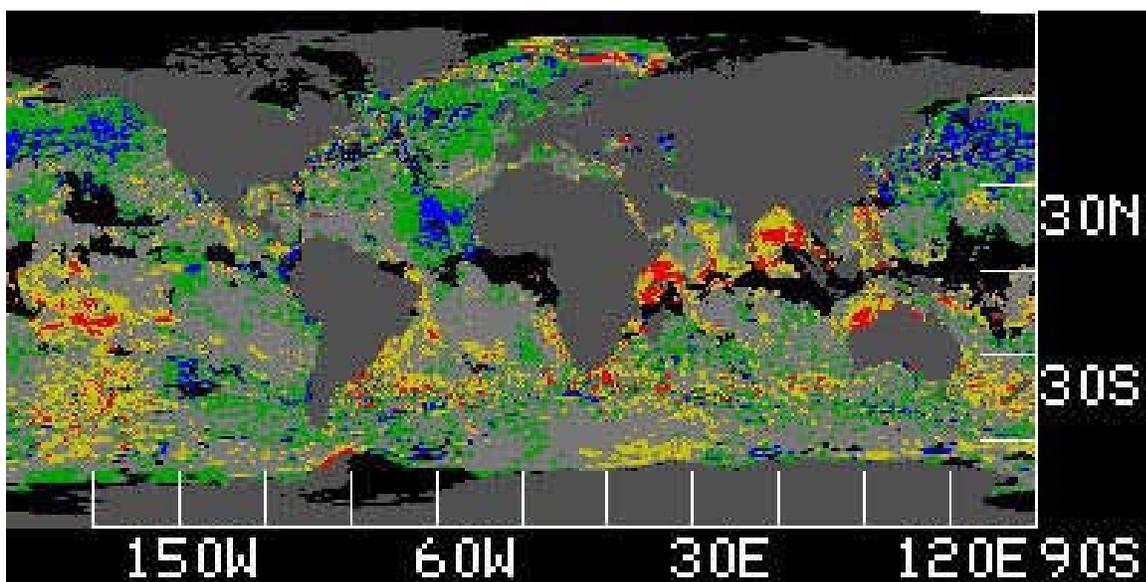


Figure 12. Pathfinder - Reynolds difference map. Light gray regions show areas where the Pathfinder and Reynolds fields differ by less than 0.33C. Green and yellow areas are 0.33 to 1C cold and warm respectively while the blue and red areas show cold and warm departures > 1C. A significant fraction of the regions of large discrepancy lies in areas with little to no ship data.

3.2.3.3 Temporal comparison of temperature fields

Time provides another dimension to analyze the performance of the satellite retrieval algorithms. The interannual behavior of the Pathfinder-Reynolds differences seen in the following two plots (Figures 11 and 12) of one week for each of four years, 1990-1993. Figure 13 shows the zonally averaged differences for week 12. The small jumps seen in each of the curves in Figure 11 is 0.15C or slightly more than the least count digitization for the long wavelength channels (10 and 11 μ) of the AVHRR sensor. For most of the plot, data for the four years falls within ± 1 count quantifying the stability of the retrievals between the two products, each referenced to a different *in situ* data set. The systematic behavior for the four years results suggests other sources of SST (e.g. wider geographical distribution) will be required to address these differences.

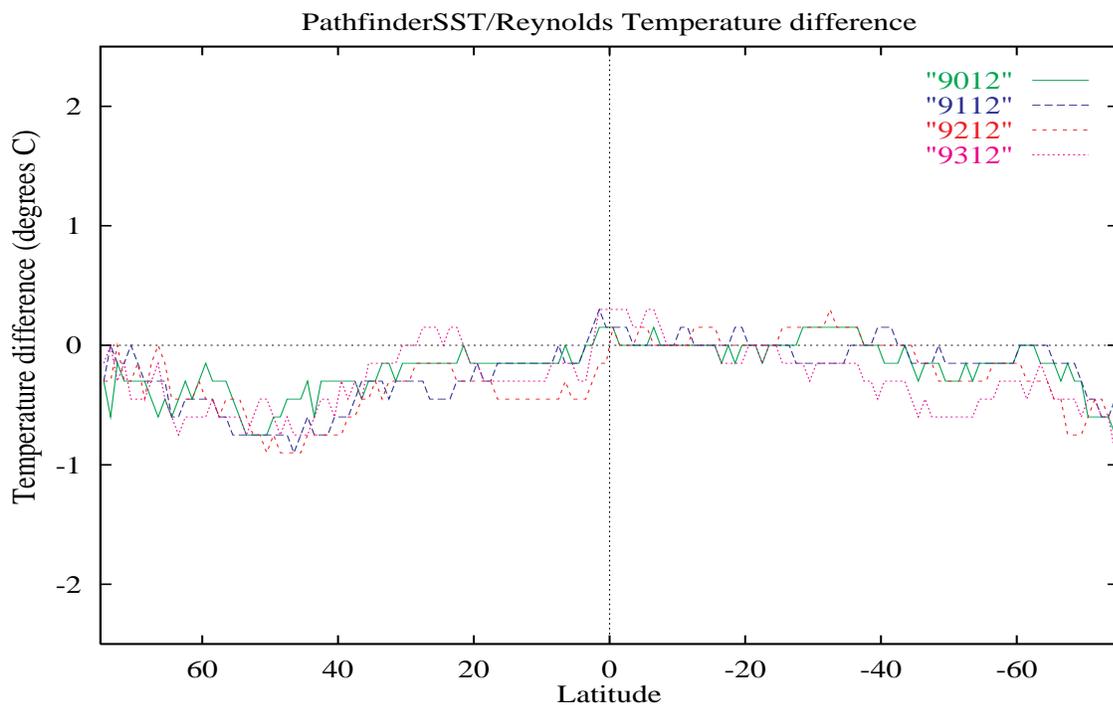


Figure 13. Four year presentation of Pathfinder-Reynolds difference for zonally averaged week 12, years 1990 to 1993, NOAA-11. Small vertical excursions in the plot are at the digitization limit of the AVHRR sensor. Temperature patterns for the four years do not show temporal drift. Consistent differences between the Pathfinder and Reynolds fields will be analyzed and

compared with other data sets such as wind, water vapor, sensor state, *in situ* observations to help determine the source of the discrepancies.

Figure 14 shows the zonally averaged differences for week 38. The curve for 1991 shows the latitudinal range influenced by the Mt. Pinatubo aerosols, 30N to 25S. Again data for the four years is tightly grouped with more agreement seen in the northern hemisphere than in the southern. The peak of the aerosol influence is seen at 10N, the latitude of the volcano.

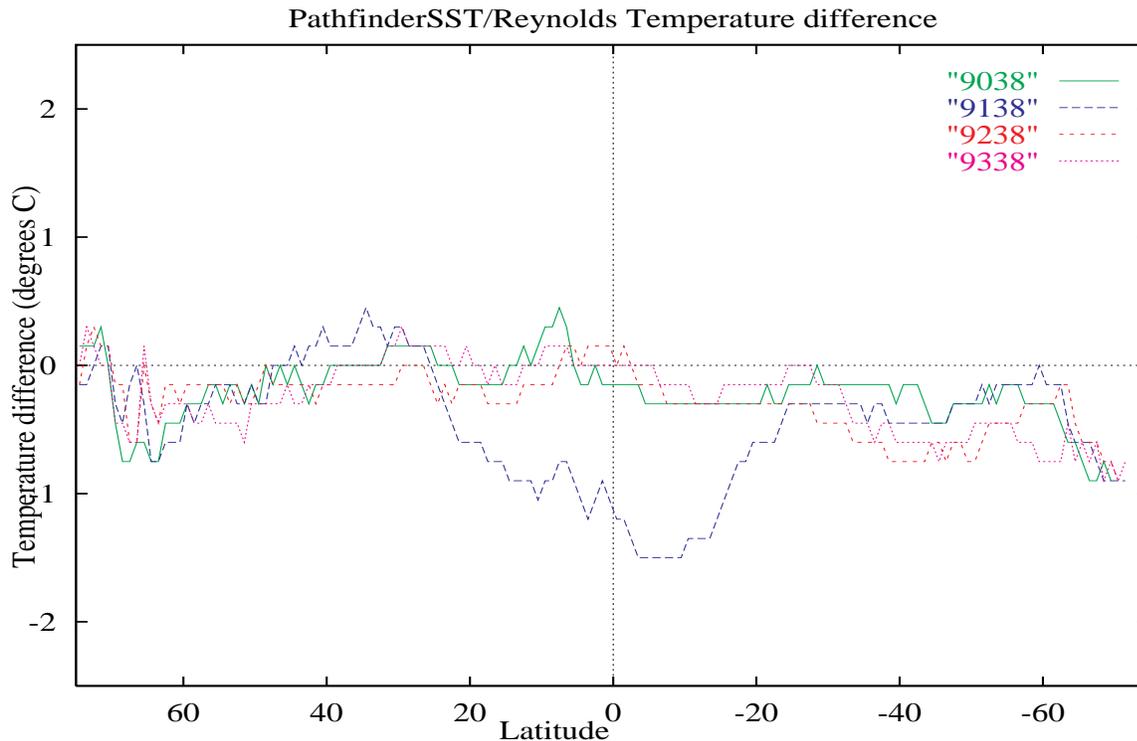


Figure 14. Same presentation as Figure 13 for week 38 showing the effect of the Mt. Pinatubo aerosols on the Pathfinder algorithm in 1993

3.2.3.4 Comparison of reference climatology fields

The evaluation of the global Pathfinder SST fields requires that a standard global SST field be available for comparison. Without an accurate reference field, or at least a reference field with known limits in accuracy, it is not possible to make meaningful comparisons. Such comparisons are necessary to identify SST discrepancies and to compare with other fields (aerosols, water vapor) in order to make improvements in the Pathfinder SST algorithm.

There are several global SST fields that are widely available for comparative purposes, including the Pathfinder, GOSTA, Reynolds OISST, and World Ocean Atlas 1994 (WOA94) datasets. A recent study by Casey and Cornillion 1999 compared a Pathfinder SST climatology with other SST climatologies to historical *in situ* surface SST measurements from research vessels which are found in the WOA94 database (Levitus

1994). They found that a climatology assembled from the 9 km Version 4.1 Pathfinder SST product had the lowest global standard deviation (1.45 C) as compared to the GIST 1 degree (1.55 C), Reynolds 1 degree (1.58 C), WOA94 1 degree climatology (1.58 C), and GOSTA 5 degree (2.07 C) products. Casey and Cornillon 1999 took these results to mean that the Pathfinder fields are more accurate than the fields used in the other climatologies. However, the use of global means and standard deviations, and even such statistics taken from zonal band ensembles, does not indicate those specific areas where a particular climatology may experience problems that cause it to deviate significantly from reality. Also, the use of monthly averages rather than annual means will be able to identify those processes that occur on some shorter frequencies.

The climatologies used in this study are listed in Table 15. The MPFSST, AOML, TOMS, and SSM/I climatologies were assembled and computed for this study; the other climatologies have been assembled by other researchers for other purposes. The Pathfinder Ascending (MPFSST--A) and Descending (MPFSST--D) climatologies were produced by averaging the daily 1 degree resolution SST estimates for each month, for the years 1988 to 1993. The AOML climatology was assembled from the National Oceanic and Atmospheric Administration's Atlantic Oceanic and Meteorological Laboratory's database of surface drifters. The data were binned in 1 degree bins and a splined surface under tension was fit to average between the (usually) sparse measurements. The TOMS aerosol index fields for each month from these years were bilinearly interpolated to 1 degree from their original 1 by 1.25 degree resolution, but other than this subsampling procedure they were not spatially smoothed. The SSM/I fields were assembled by assembling a composite of the daytime and nighttime values on a 1 degree grid.

Name	Quantity	Source	Smoothed?
MPFSST-A	SST	AVHRR	No
MPFSST-D	SST	AVHRR	No
GOSTA	SST	Ship Reports	Yes
AOML	SST	Surface Drifters	Yes
OISST	SST	Sips, Buoys, AVHRR	Yes
SuperObs Ship	SST	Ships	Yes
SuperObs Buoy	SST	Buoys	Yes
TOMS	Aerosol Index	TOMS	No
SSM/I	Water Vapor	SSM/I	No

Table 15. The climatologies assembled for inter-comparisons.

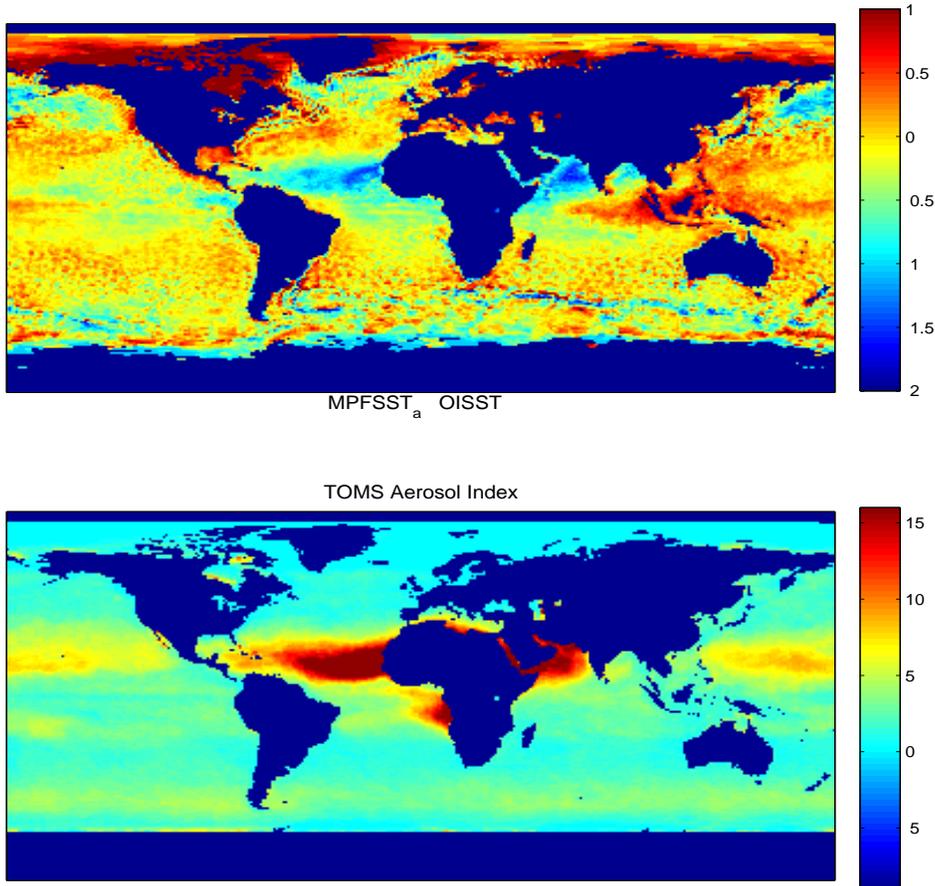
The inter-comparisons of these monthly 1 degree climatologies revealed that the GOSTA, AOML, and both SuperObs fields cannot provide an adequate global reference field, mostly due to the (unconstrained) interpolation to vast undersampled areas. This problem is alleviated in the "blended" OISST fields by using AVHRR SSTs to constrain the fields between in situ observations.

The OISST fields appear to be the most promising as a reference field for large scale climatological Pathfinder comparisons. However, the best comparisons were not accomplished through the differencing of the average fields, rather the most realistic comparisons are accomplished through the averaging of the differences between the fields. It are these "averages of MPFSST - OISST" that were used to compare to the aerosol and water vapor estimates.

Figure 15 depicts the average of the MPFSST - OISST differences for the month of July and the average TOMS aerosol index for the same month. The correlation between the high MPFSST-OISST difference and the high aerosol index in the southern North Atlantic and Arabian Sea is striking, and appears to be the result of Saharan dust which is often present in the atmosphere at that time of year. That the OISST fields are able to resolve the temperature deficit in that area despite relatively sparse observations and an AVHRR SST that was plagued by the same aerosols as the Pathfinder estimates is encouraging for future comparisons. Of course, of equal importance is why other areas (e.g. the western North Pacific) do not show as striking a correlation. Work is continuing on the comparisons of both the TOMS and the SSM/I water vapor fields with the MPFSST - OISST differences. The spatial correlations between these fields will be quantified in order to best identify those areas and conditions that produce errors in the Pathfinder SST fields.

Figure 15. The average ascending MPFSST minus Reynolds' OISST [top panel] and the average TOMS Aerosol Index [bottom] for the month of June over the years 1988-1993. Strong correlations between the fields exist over areas that typically have Saharan dust in the atmosphere. Other areas with high Aerosol Index values do not always have a corresponding SST deficit.

1988 1993 Average for Month 06



3.2.4 Quality Control and Diagnostics

A number of automated quality control mechanisms have been included in the Pathfinder SST processing path. During Level-2 calculation a series of tests, shown in Section 3.1.2, is employed that utilize spatial and spectral properties to determine suitability of the sensor retrieval. The calculated SST is next compared with a climatology local to the time period, *e.g.* a SST field produced from an OA analysis of the last week's SST. Finally a longer space-time series is used to produce a better estimate of the expected SST and its variance. Results from each of the tests are used to set a flag. A rule base has been formulated for AVHRR Pathfinder that uses the flags, pointing angles, and SST variance to determine the expected quality of a given retrieval.

These tests have been employed in the Pathfinder processing and have resulted in approximately 25% of the satellite observations classified as good and the remainder placed in a lower quality category. Of the 2×10^7 retrievals calculated for each SST field, approximately 100 bad retrievals are classified as good. Further study is required to determine the size of the good fraction that is misclassified as bad. Each pixel in the result's field has a companion status field. The status field is used in subsequent processing such as preparing daily to long term global maps. Each map location contains data ranging in quality from bad to good permitting location of persistent cloud contaminated areas.

Additional tests can be performed such as generating difference maps of daily field - reference climatology, zonal and meridional sections. Each map location will have an expectation of acceptable variance. The processing software can perform these tests and report results to a Q/A monitor. The expected outcome of a discrepancy is a report to the SCF for further analysis. These tests provide quality testing for mapped global fields.

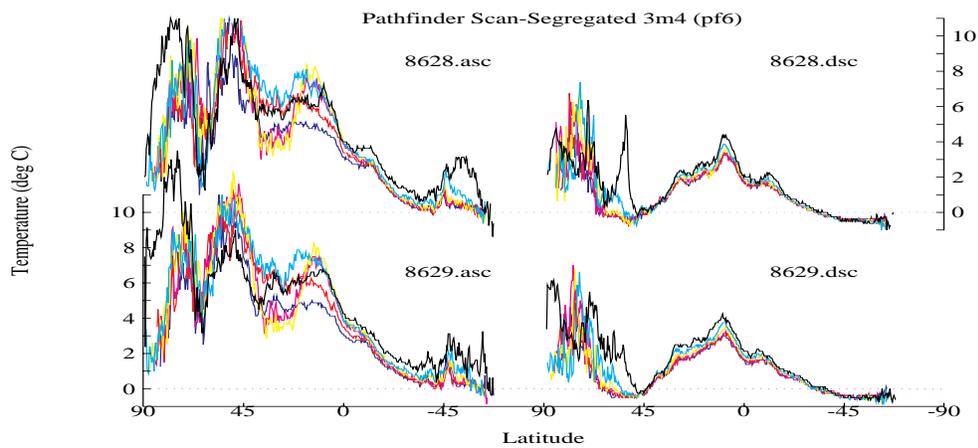
The time series generated from the matchup data base also provides a quantitative test on the combined sensor and algorithm performance. These tests provide only a limited analysis of product and sensor stability when utilized on a single retrieval (day) basis. Results obtained from analysis of longer time series will be required to differentiate the various failure modes. For example, the sensor could undergo undiagnosed calibration shifts with time or the algorithm's correction for scan angle might be incomplete. A suitable number of samples is required to develop a suitable set of tests given the expected retrieval noise.

Validation of MODIS retrievals will be based on comparisons with *in situ* data obtained from ship and buoy observations. Assuming continuation of the present international buoy deployment efforts, order of 500-1000 drifters should be available to provide reference for SST. Brown intends to augment this data set with a number of ship base radiometers. Clark will deploy optical buoys and we expect that several investigators will deploy optical drifters. These measurements provide absolute or relative reference data depending on the ability to establish and maintain calibration. In addition, comparison will be made using consistency between fields in space and time. Fields derived from other satellite sources (AVHRR, SeaWiFS, multiple MODIS instruments) provide additional comparison opportunities.

The sensor calibration time series obtained from the vicarious method will be compared with the record developed by the MCST. The calibration record developed by MCST using MODIS sensor sources will be the only method of tracking orbit to orbit sensor stability. The vicarious and matchup database will provide long term trend analysis and calibration, validation for the retrieved water leaving radiances for time periods longer than ~six months. Products produced at the MOTCF for a subset of the MODIS observations will be compared with the standard products produced at the PGS to verify proper operation of the algorithm.

3.2.4.2 Potential quality effects of stray light

An important application of the quality control analysis is to understand the range of observation conditions that yield useful surface retrievals. Figures 16a and b are based on first producing global day and night fields of various satellite measurements. In the following Figures 16a and b, the channel 3 - channel 4 (MODIS channels 20 and 31) difference is calculated and the results stratified by 10 degree satellite zenith angle bins. Daily fields are produced and subsequently combined into weekly fields. The weekly fields are then zonally averaged to show variation in the channel 3-4 with latitude for selected weekly periods. Each of the figures presents daytime (ascending on left) and nighttime (descending on right) with satellite zenith angle range discriminated by color. Channel 3, 3.7μ band, contains both emissive and reflective solar energy during daytime observations increasing the channel 3-4 difference. Nighttime retrievals should show only the effects of radiance emitted from the ocean surface and intervening atmosphere. The large spikes that occur in the northern latitudes (poleward of 45 degrees) during mid-year (week 28) for large satellite zenith angles (>50 deg) suggest that energy is being received by the detectors from sources other than the surface. Later in the year for week 41, smaller abrupt transitions are seen for angles >40 deg. In each case we would need to restrict retrievals for time and locations that are not subject to these problems. An added benefit is the production of a climatology for the channel difference that permits quality testing of the individual retrievals to determine whether a given retrieval is valid.



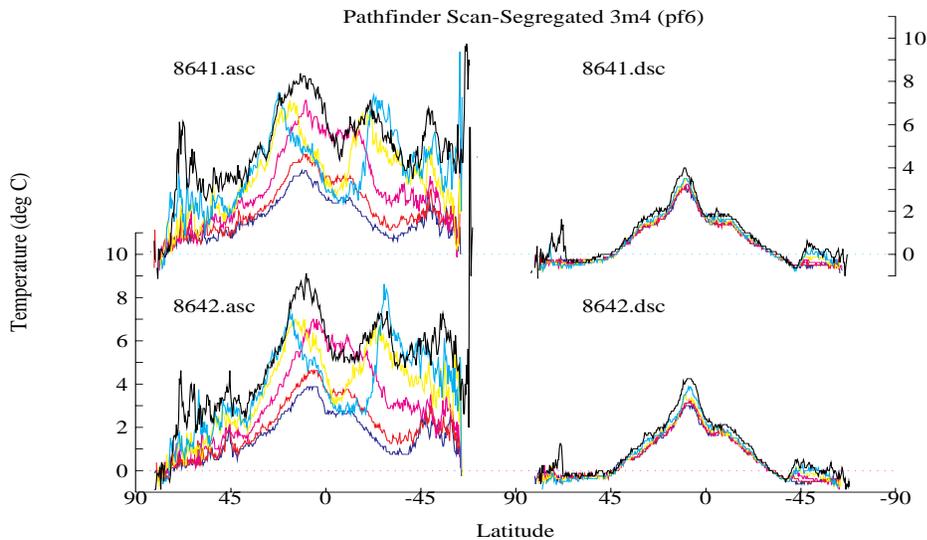


Figure 16. Pathfinder channel 3 - channel 4 differences (MODIS Ch20-Ch31) for two 1 week period, 1986 week 28, upper panels, and week 41, lower panels. Day orbits presented on left(asc) and night orbits on right (dsc). Data processed as daily, 9km, equal area fields, time binned into 7 day week periods, and then zonally averaged. Colors represent satellite zenith angle ranges, dark blue, 0-10 deg; red, 10-20; pink 20-30, yellow 30-40; light blue 40-50; black >50. Large differences for day passes shows influence of reflected sunlight in channel 3. Abrupt increases in channel difference poleward of 45 deg at high satellite zenith angles suggests stray light problems. Finally increasing channel difference near the equator reflects 'average' effects of path radiance and serves as a baseline for possible algorithm option selection and for data quality tests.

3.2.4.3 Quality control of *in situ* observations

A fundamental basis of both the SST and ocean color calibration and validation activities based on comparison of satellite retrievals with corresponding *in situ* observations. Since the *in situ* observation are subject to error and calibration drift, especially for drifting buoys, we have developed a preliminary approach to identifying buoys with potentially faulty measurements. Figures 17 and 15 are based on processing satellite retrievals using the equations and quality filtering based on the previously presented matchup process. The resulting daily satellite fields are then objectively analyzed to remove data gaps due to cloud or instrument problems. Residuals are computed for the individual OA Satellite SST - buoy SST and histogrammed by year.

In Figure 17, histograms are presented showing a gaussian-like distribution of residuals principally centered between $\pm 1C$ with biases $< 0.2C$. Neglecting partitioning the residuals by various criteria, e.g. satellite and solar zenith angle, time, location, ancillary data..., the histograms present an indication of the performance of the algorithm.

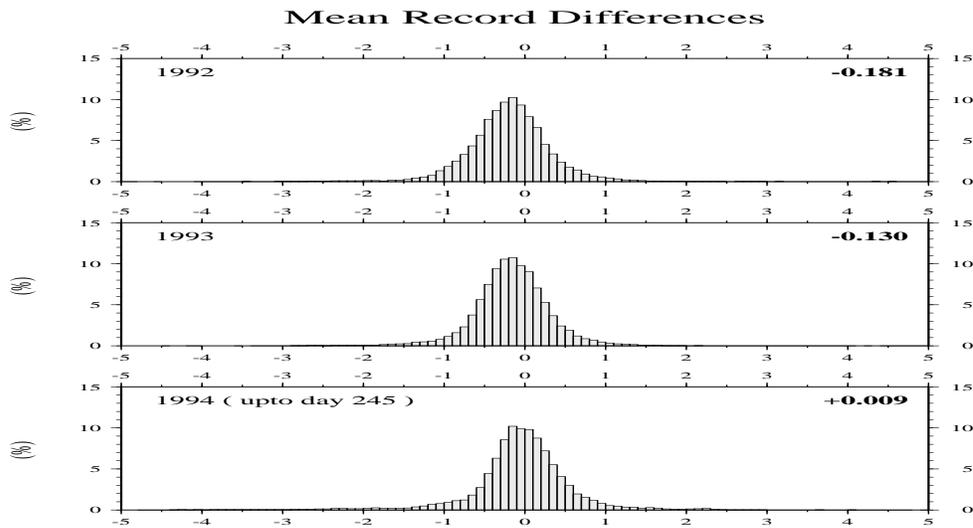


Figure 17. OA Pathfinder SST - Buoy SST for all buoys for 1992 to 1994. Daily Pathfinder SST fields are Objectively Analyzed to fill data voids (clouds). Figure 15 restructures the data by computing residuals on a buoy by buoy basis.

Here we assume that the satellite SST field provides a transfer reference between buoys in (bias) is present, the residuals are distributed about zero or begin to depart at some point during the buoy lifetime. Most average residuals fall within $\pm 0.5\text{C}$ but some buoys report average temperatures that depart from the satellite field by more than 1C . Buoys that exhibit this characteristic must be individually examined to determine if this behavior is due to buoy calibration or anomalous atmospheric conditions, *e.g.* unusual atmospheric moisture profiles or aerosols.

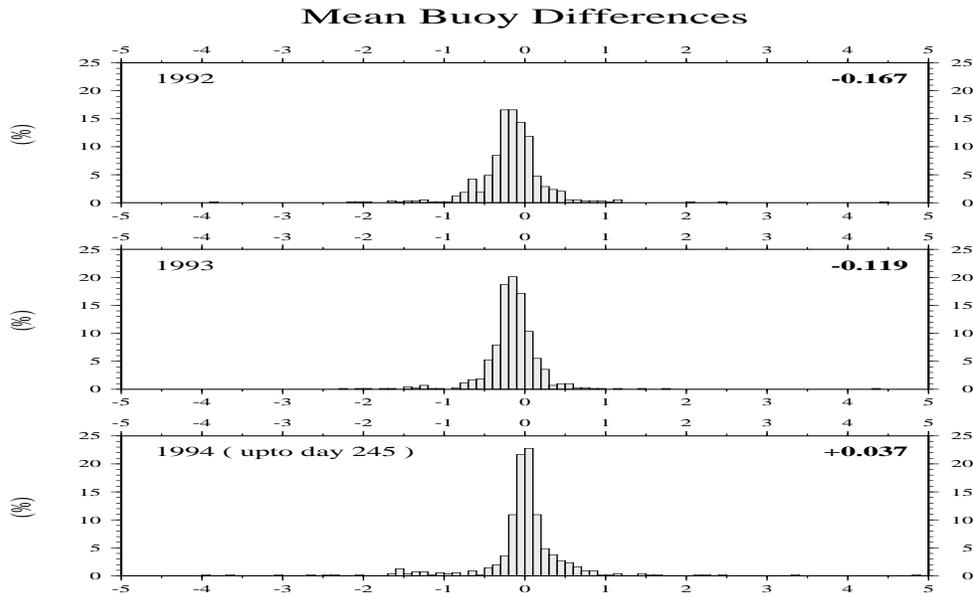


Figure 18. Mean OA Pathfinder SST - Buoy SST per buoy, histogrammed for all buoys for 1992 to 1994. Daily Pathfinder SST fields are Objectively Analyzed to fill data voids (clouds). Departures from 0 are indicative of possible buoy calibration error.

The next two figures (19 and 20) present a time series of the comparison of OA satellite SST and buoy SST. Figures 19 and 20 show the behavior of two buoys, one considered “good” and another “bad”. The four panels for each figure show from top down SST time series, satellite-buoy residuals by retrieval and with an 8 day filter to minimize effects of satellite zenith angle, a histogram of the residuals for this buoy and finally the trajectory of the buoy. This particular buoy (Fig. 19) compares well with the satellite SST for most of the record. The record also helps localize the time and location of maximum residual, near day 180 and thus points to the specific satellite fields that can be further analyzed to determine if the buoy is located near a front or eddy and thus subject to errors introduced by navigation error.

1992 aoml.noaa_11.asc.GLB 531

x

229 -1.716 -0.096 2.263 (0.644)

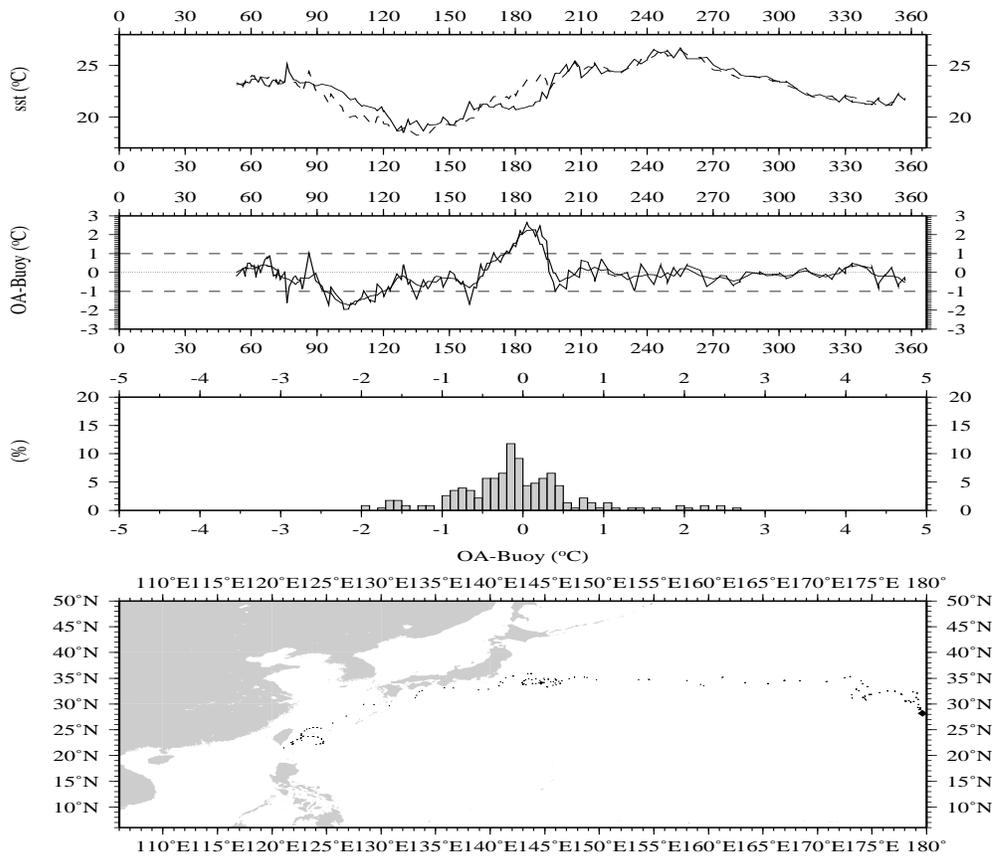


Figure 19. Plot of Buoy and OA Pathfinder SST, upper panel; OA Pathfinder – buoy SST, middle upper panel, OA Pathfinder - Buoy SST histogram, lower middle panel, and buoy trajectory, lower panel. Statistics for the lifetime of a ‘good’ buoy where residuals oscillate about the zero line for the life of the buoy.

Figure 20 shows a buoy where a consistent offset is seen throughout the buoy record. The type of trend is likely a result of buoy miss-calibration. Care must be taken to insure that the offsets are not an indication of presence of aerosols or other atmospheric conditions. Correlation with other satellite or ancillary fields can assist in this process.

1992 aoml.noaa_11.asc.GLB 1368

x

185 -1.718 -1.120 -0.709 (0.241)

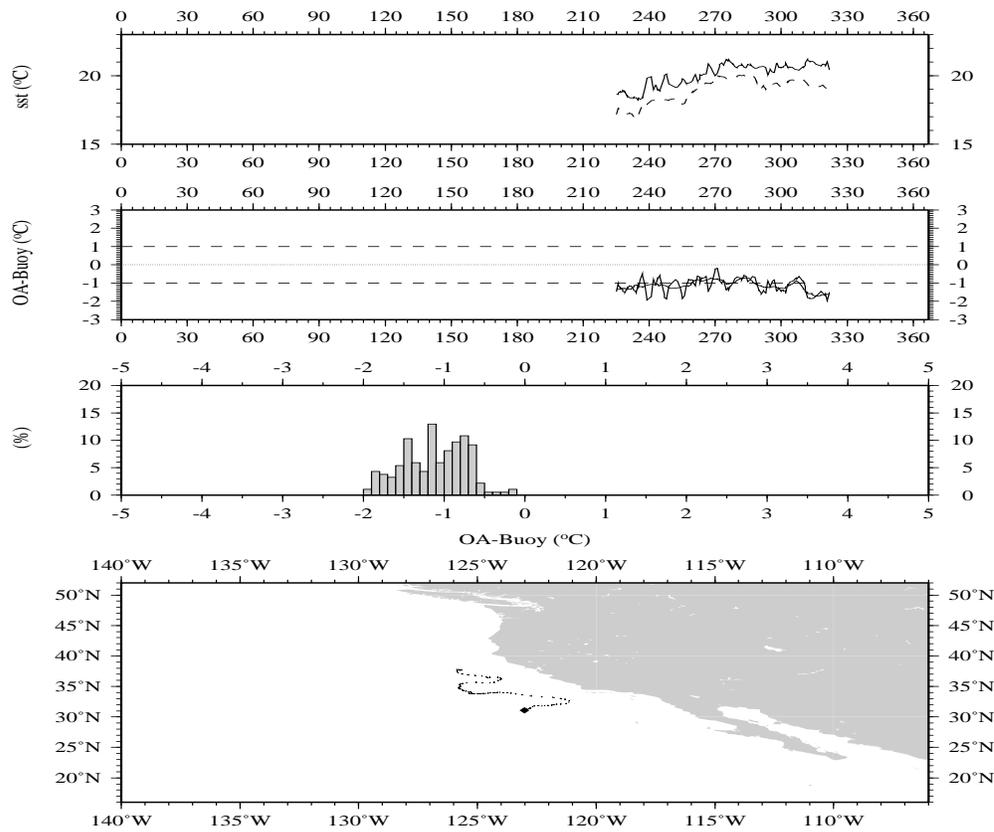


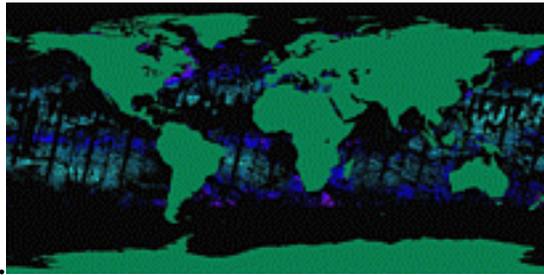
Figure 20. Plot of Buoy and OA Pathfinder SST, upper panel; OA Pathfinder – buoy SST, middle upper panel, OA Pathfinder - Buoy SST histogram, lower middle panel, and buoy trajectory, lower panel. Statistics for the lifetime of a 'bad' buoy where residuals oscillate about the zero line for the life of the buoy. Consistent offset of the OA Pathfinder SST and the buoy SST indicate a buoy calibration (bias) problem or possible unusual atmospheric condition.

3.2.4.4 Ocean Processing Level-2 Pixel Quality control flags

The sample images presented in Figure 21 panels A-L were produced from MODIS V2 at launch algorithms using SeaWiFS converted to MODIS level 1 data format. Two days of SeaWiFS data for July 2 and 3rd 1998 were processed to create these daily images. This has enabled us to test the complete MODIS oceans processing chain from Level 1b ->Level 2 calibrated parameters and space binning ->Level 3 time binned and mapped products. We are currently evaluating the ranges of values produced and the associated pixel level quality flags to verify the at launch algorithms. The quality control flags for all of the MODIS ocean products are presented below. These Flags are used to control data binning and are important diagnostic tools for understanding the accuracy of the data.

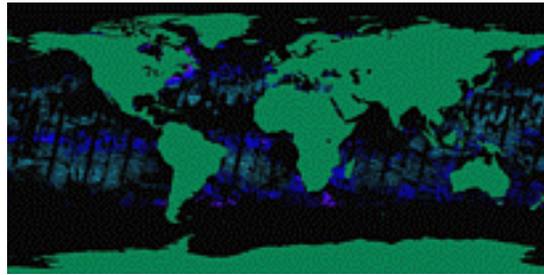
July 2 and 3rd 1998

A.



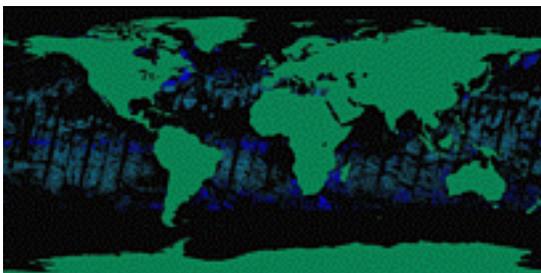
nLw 412nm

B.



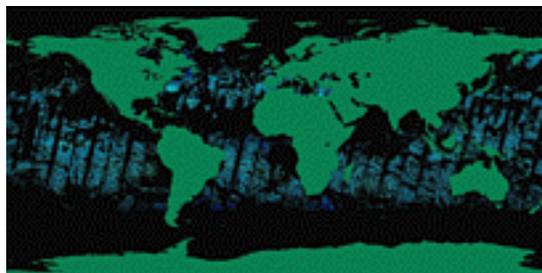
nLw 443nm

C.



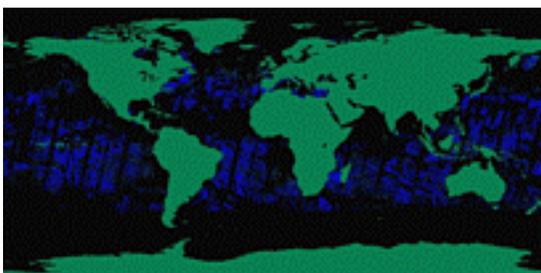
nLw 488nm

D.



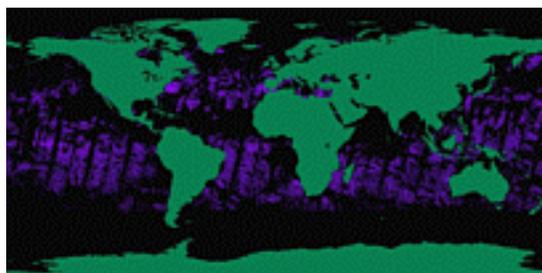
nLw 531nm

E.



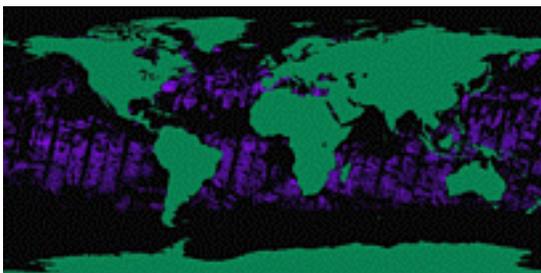
nLw 551nm

F.



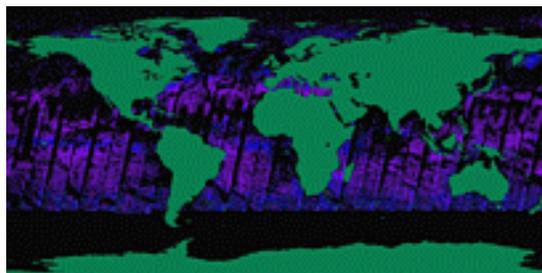
nLw 667nm

G.



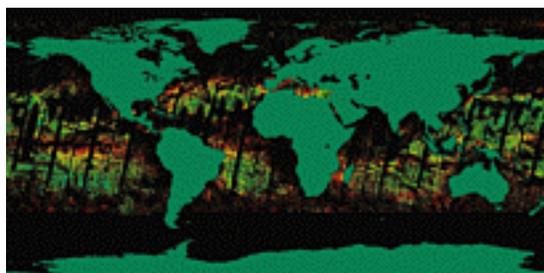
nLw 678nm

H.



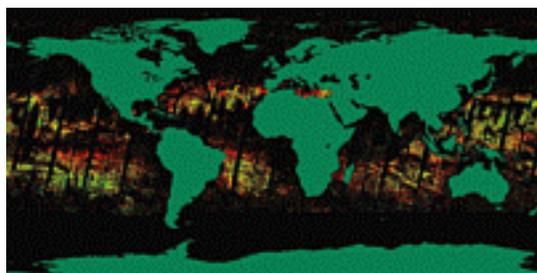
K490nm

I.



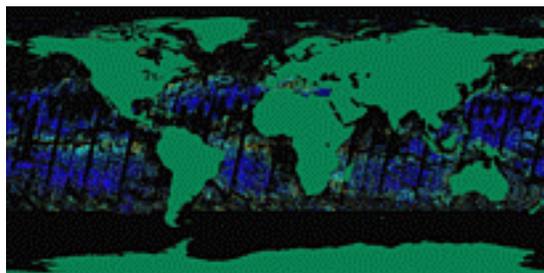
MODIS Chlorophyll

J.



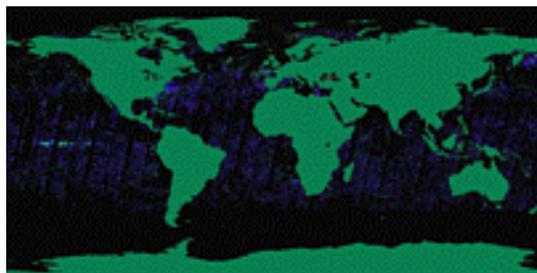
pigment_c1 total

K.



Suspended solids

L.



Coccolith Calcite concentration

Figure 21 panels A-L. MODIS results for converted SeaWiFS data

A series of quality flags have been defined to aid in the interpretation and selection of data. The flags indicate ranges of observational parameters where product calculation could be subject to increased error, *e.g.* large satellite or solar zenith angles, atmospheric conditions near the limit of retrieval capability, *e.g.* proximity to clouds, high aerosol or water vapor loading or specific in water conditions, *e.g.* low chlorophyll leading to low fluorescence, coccolithophorid blooms with high in water scattering. By testing combinations of conditions, the suitability of particular product retrieval for use in qualitative or quantitative applications can be determined.

The matchup database has been used for SST to define tests to help differentiate clear and cloudy pixels. When the additional spectral channels provided by MODIS become available, it is hoped that further tests can be defined that will help identify pixels subject to aerosol.

This section describes the MOCEAN QA storage protocol. The protocol includes both pixel level QA results stored in science data sets (SDS) generated at the time of execution (non-searchable) and granule/product level ECS QA metadata (Automatic Quality flag, Operational quality flag, Science Quality flag, text explanations, QA Percent Missing) which are searchable.

Pixel level QA parameters common to all products are:

- common flags: 1 byte for each pixel:

Bit	Description	Result
0	Pixel not processed	0=processed, 1=not processed
1	Atmospheric Correction	0=successful,1=failed (due to atmos_corr,bad ancillary data, and/or sun glint)
2	Satellite Zenith Angle	1=too large, 0=ok
3	Solar Zenith Angle	1=too large, 0=ok
4	Shallow water	1=shallow
5	Sun_Glint - Glint > threshold	
6	Supp_Data - Invalid or missing ancillary data	
7	Land	

- 4 bytes from Cloud Mask flags

Pixel QA parameters unique for each product consist of two types; flags which contain information on specific tests, and a quality level which summarize the results of various flags for each pixel in context for a given product. Below is listed the quality level and science QA flags for each MOCEAN level 2 and 3 product

3.2.4.4.1 SST atmospheric correction processing flags

MODSST flag names:

-quality - 1 byte

0-1 Mandatory Qual for sst band:

0 (good) if general bits are ok, and input Lw's are ok, and 2-7 are ok

1 (questionable) if any of: shallow, large zenith angles, or flags 2-7

2 (cloud) if any input radiances are negative and saturated

3 (bad, other than cloud) if any input radiances are negative and

not saturated, or Atmos Corr (bit 1 common flag) failed, or land

2-3 Mandatory Qual for sst4 band:

0 (good) if general bits are ok, and input Lw's are ok, and 8-13 are ok

- 1 (questionable) if any of: shallow, large zenith angles, or flags 8-13
- 2 (cloud) if any input radiances are negative and saturated
- 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 common flag) failed, or land
- 4-7 Spare

-flags - 2 bytes

- 0 =0 if sst input radiance (20,31,32) are ok
- 1 =0 if sst4 input radiance (20,22,23) are ok
- 2 31/32 uniformity test 1
- 3 31/32 uniformity test 2
- 4 31/32 zenith angle 1
- 5 31/32 zenith angle 2
- 6 31/32 tree test
- 7 31/32 sst diff from reference
- 8 20/31/32 brightness temps ok
- 9 22/23 uniformity test 1
- 10 22/23 uniformity test 2
- 11 22/23 zenith angle 1
- 12 22/23 zenith angle 2
- 13 22/23 tree test
- 14 22/23 sst different from reference
- 15 20/22/23 brightness temps ok

3.2.4.4.2 Additional SST Flags To Be Added

Recent instrument characterization results suggest MODIS instrument specific tests that will augment the algorithm specific tests. Both the need for such tests and the specific nature of the tests can be determined only after the MODIS instrument characterization is completed and the data studied by the MCST. Implementation of these tests will occur by late Spring or early Summer of 2000. These flags will be applied at the band and pixel level. Additional flag topics presently under consideration:

Polarization within limits

Spectral cross talk problem
Spatial cross talk problem
potential near field scattering
potential far field scattering
Lw < 0 or not correctable - Lt out of range

Additional bits for instrument/band specific problems that emerge from testing time histories and field analyses and climatologies will be employed once the data becomes available. Some of the fields initially can be populated using results from heritage instruments while others must await the launch of MODIS. Two examples are included:

Channel calibration changes - compare with history
| Lw-reference | out of range

Two additional spectral test are being implemented for the SST product:
Cloud tests from the Menzel cloud product will be incorporated, these include:

Cirrus cloud test
Low level cloud or fog

We expect that once the combination of spectral channels on MODIS is available, a number of factors that influence both ocean color and SST retrievals can be tested.

Expected examples include:
White caps
Absorbing aerosols
Atmospheric correction model carry over
Cirrus clouds
Delta sst vs delta chl comparison, possible absorbing aerosol test

3.2.4.4.3 Ocean color atmospheric correction processing flags

MODCOL Flags:

MODCOL2 - 3 bytes QA parameters:

-quality levels - 1 byte

bits

0-1 Mandatory Quality for all of Gordon's nLw bands:

0 (good) if general bits are ok, and flag bits 1-12 are ok

1 (questionable) if any of: shallow, large zenith angles, or flags 9-12

2 (cloud) if any input radiances are negative and saturated

3 (bad, other than cloud) if any input radiances (all 9) are negative

and not saturated, or Atmos Corr (bit 1 above) failed, or land

2-3 Mandatory Quality for Carder's clear water epsilon band

0 (good) if general bits are ok, and input Lw's are ok, and flag 14 ok

1 (questionable) if any of: shallow, large zenith angles, or flag 14

2 (cloud) if any input radiances are negative and saturated

3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land

4-7 Spare

-flags - 2 bytes

bit

- 0 Cloudy - Albedo @ 865 > threshold 1=cloudy 0= clear
- 1 Bad_Lw - One or more bands missing
- 2 Bad_Lw - Any LwXXX <= 0.
- 3 Bad_Lw - Any band counts < 0
- 4 Atmos_Corr - Questionable polarization correction/mirror reflectance
- 5 Atmos_Corr - Gordon aerosol failure
- 6 Atmos_Corr - Epsilon out of range (< lower limit or > upper limit)
- 7 Atmos_Corr - Any LaXXX <= 0.
- 8 Atmos_Corr - Invalid Raleigh scattering data
- 9 nLw550_low - Calculated nLw550 is too small
- 10 Cocco - Coccolithophorid radiance exceeds threshold
- 11 TurbidCase2 - Actual_rrs555 > Turbid_rs555
- 12 Hi_la865 - Calculated La865 is too large
- 13 input Lw's for Carder's clear water epsilon band (11,13) are ok
- 14 lo_eps - epsilon < threshold
- 15 absorbing aerosol 1=present 0= absent

MODOCL2A - 6 bytes of QA parameters:

-quality level - 3 bytes

Bits

- 0-1 Mandatory Quality for Clark's pig_c band
- 0 (good) if general bits are ok, and input Lw's are ok
 - 1 (questionable) if any of: shallow, large zenith angles
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land
- 2-3 Mandatory Quality for Clark's pig_total and chl_modis band
- 0 (good) if general bits are ok, and input Lw's are ok
 - 1 (questionable) if any of: shallow, large zenith angles
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land
- 4-5 Mandatory Quality for Clark's susp_solid band
- 0 (good) if general bits are ok, and input Lw's are ok
 - 1 (questionable) if any of: shallow, large zenith angles
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land
- 6-7 Mandatory Quality for Clark's k490 band
- 0 (good) if general bits are ok, and input Lw's are ok
 - 1 (questionable) if any of: shallow, large zenith angles
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land
- 8-9 Mandatory Quality for Abbott's fl_baseline band:
- 0 (good) if general bits are ok, and input Lw's are ok, and 7-11 are ok
 - 1 (questionable) if any of: shallow, large zenith angles, or flags 7-11
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land

- 10-11 Mandatory Quality for Abbott's FLH bands:
- 0 (good) if general bits are ok, and input Lw's are ok, and 7-11 are ok
 - 1 (questionable) if any of: shallow, large zenith angles, or flags 7-11
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land
- 12-13 Mandatory Quality for Abbott's Chlor Fluorescence efficiency band:
- 0 (good) if general bits are ok, and input Lw's are ok, and 7-11 are ok
 - 1 (questionable) if any of: shallow, large zenith angles, or flags 7-11
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 common flag) failed, or land
- 14-15 Mandatory Quality for Gordon's `cocco_pig_c`, `cocco_conc`, `calcite_conc` bands:
- 0 (good) if general bits are ok, and input Lw's are ok, and 13-15 are ok
 - 1 (questionable) if any of: shallow, large zenith angles
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land, or flags 13-15
- 16-17 Mandatory Qual for Hoge's `peb` and `pub` bands:
- 0 (good) if general bits are ok, and input Lw's are ok, and 17-21 are ok
 - 1 (questionable) if any of: shallow, large zenith angles
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 common flag) failed, or land, or flags 17-21
- 18 Cloudy - Albedo @ 865 > threshold 1=cloudy 0= clear
- 19-23 Spare

-flags - 3 bytes

- 0 =0 if pig_c input Lw's (9,12) are ok
- 1 =0 if Chl_A input Lw's (9,10,11,12) are ok
- 2 =0 if pig_total input Lw's (9,10,11,12) are ok
- 3 =0 if susp_solid input Lw's (9,10,11,12) are ok
- 4 =0 if k490 input Lw's (9,12) are ok
- 5 =0 if Abbott's Fluor baseline input Lw's (13,15) are ok
- 6 =0 if Abbott's FLH input Lw's (13,14,15) are ok
- 7 =0 if Fluorescence efficiency input Lw's (8-13) are ok
- 8 FLH_Range - FLH out of range
- 9 L748_High - L748 > L667
- 10 L678_Base - L678 peak below baseline
- 11 chlflag - Chlor < 2.0 mg/m³
- 12 chlbad - Invalid chlor input
- Gordon Cocco:
- 13 =0 if cocco input Lw's (flag 9,12) are ok
- 14 LoRadiance - value below lower bound of lookup table
- 15 HiRadiance - value above upper bound of lookup table
- 16 InvalidEntry - invalid data in lookup table
- Hoge:
- 17 Hoge's peb,pub input Lw's (8-12) are ok
- 18-22 range_iop_flags[5]: 5 separate flags; one for each IOP to flag if any IOP is outside of an expected realistic range. *** If any one of these flags is set, output IOPs are suspect.
- 23 Cloudy - Albedo @ 865 > threshold 1=cloudy 0= clear

MODOCL2B- 3 bytes of QA parameters:

-quality level - 1 byte

bits

- 0-1 Mandatory Quality for Carder's chlor_a_3,ag400,aphi675,atot_mod*:
0 (good) if general bits are ok, and input Lw's are ok, and 2-9 are ok

- 1 (questionable) if any of: shallow, large zenith angles, or flags 2-9
- 2 (cloud) if any input radiances are negative and saturated
- 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 above) failed, or land, or flag 1
- 2-3 Mandatory Quality for Carder's ipar and arp bands:
 - 0 (good) if general bits are ok, and input Lw's are ok, and 1,11 ok
 - 1 (questionable) if any of: shallow, large zenith angles, or flag 11
 - 2 (cloud) if any input radiances are negative and saturated
 - 3 (bad, other than cloud) if any input radiances are negative and not saturated, or Atmos Corr (bit 1 common flag) failed, or land, or flag 1
- 4-7 Spare

-flags - 2 bytes

Carder Chlorophyll:

- 0 =0 if input Lw's (8-13) for Carder's bands are ok
- 1 neg_rrs_flag - one or more rss are less than or equal to zero
- 2 low_412_flag - rss[0](412) less than thresh_412
- 3 low_555_flag - rss[4](555) less than thresh_555
- 4 default_flag - using default chlorophyll model
- 5 chl_inconsistent_flag - calc chl exceeds chl_incon._thresh
- 6 chl_quality_flag - (currently unused)
- 7 hi_scatter_flag - (currently unused)
- 8 blend_flag - aph_mod between .03 and .06 chlor. blend
- 9 package_flag - chl package or unpackage

Carder PAR:

- 10 =0 if ipar,arp input Lw's (8-13) are ok
- 11 hi_windspeed - wind speed > threshold
- 12 Cloudy - Albedo @ 865 > threshold 1=cloudy 0= clear
- 13-15 Spare

Productivity Flags

CHLOROPHYLL FLAGS:	# bits	position
Chlorophyll a Field Id (0-5)	4	0-3
Chlorophyll a Quality Index (0=Good,15=Bad,1-14=Ugly)	4	4-7
High-Variance Class	1	8
Low-Sampling Rate (n < _____)	1	9
Open Ocean Flag (1=ocean, 0=inland waters)	1	10
Coastline in Bin	1	11
Depth < 30 meters	1	12
High Average Aerosol	1	13
Low Average Radiance	1	14
High Cloud Frequency	1	15

(added to running annual average field in V2)

SUBPRODUCT FLAGS:	# bits	position
Function Id (0-15)	4	0-3
Quality Index (0=Good,15=Bad,1-14=Ugly)	4	4-7
No Chlorophyll a	1	8
No Chlorophyll a Uncertainty	1	9
No Sea Surface Temperature	1	10
No Sea Surface Temperature Uncertainty	1	11
No Mixed-Layer Depth	1	12
No Mixed-Layer Depth Uncertainty	1	13
No Average Daily PAR	1	14
No Average Daily PAR Uncertainty	1	15

3.2.4.5 Ocean Processing granule/product level Quality control flags

The MOCEANS QA protocol will make use of the ECS QA metadata to communicate and summarize the status of QA procedures. The MOCEAN QA protocol makes use of the **Science quality flag** and associated text explanation. These flags are set at the granule/product level.

Science quality flag can hold the following values.

Not investigated- which is the default value set by the PGE at execution.

Being investigated - indicating that MODAT is examining the granule or product and suspects a problem

Failed - indicates the granule failed QA analysis the problem has been solved and the granule or product may now be reprocessed.

Passed- indicates passed or inferred pass of science post run time QA procedures

Science quality text explanation – will briefly summarize results from the QA analysis.

3.2.5 Exception Handling

Exception handling for ocean processing is relatively simple, products are computed at each pixel if inputs are valid. Each investigator has defined specific tests appropriate to the algorithm. Results are produced and flagged as appropriate. The flags in turn are used when Level-3 products are produced. When multiple quality levels are present, pixels are binned at the highest available quality level and data at any lower quality levels are discarded for each pixel. Summary statistics are retained in the output granule to complement the pixel levels flags carried for each separate product. It is expected that all ancillary data will be pre-screened prior to Level-2 calculation and alternatives selected if missing or erroneous data is detected. If substantial delay is experienced in receiving validated ancillary data, suitable climatologies will be used for initial reprocessing and a subsequent processing run will be required once proper ancillary data becomes available. SeaWiFS has scheduled four levels of reprocessing separate from calibration run to accommodate updates in ancillary fields, calibration and algorithm improvements. The time frame for these reprocessing efforts ranges from several days to a year or more.

3.2.6 Data Dependencies

Data dependencies have been minimized for pre-launch and for initial post-launch processing. All ancillary data used in processing is based on presently available NOAA or NASA data products. Transition to EOS products will be undertaken once the respective products have been validated and global products become routinely available in a timely manner.

Inputs to the current matchup data base are given in Section 3.1.1 in Tables 2, 3, and 4. Contents of the matchup database will be expanded to include data from Brown's *in situ* observation program and the optical moorings deployed by Clark. The external data sets specified by the MOT PI's will be included to correspond with the matchup observations.

3.2.7 Output Product

Output matchup database entries for SST and ocean color for this product are vectors composed of the retrieved SST value, input calibrated radiances and derived brightness temperatures for each channel, ancillary data, flags which quantify the cloud screening results, latitude, longitude and time. A description of the vector components and data types is given in Tables 2 and 3 in Section 3.1.1. These entries will be updated to reflect *in situ* and ancillary data either collected or required by the MOT PI's. This update will begin with receipt of the ATBDs from the MOT PI's.

4.0 Constraints, Limitations, Assumptions

The matchup database will contain data for both "good" and "bad" retrieval times, locations. No *a priori* assumptions will be made as to use of quality flags for retention of samples within the database. This approach is taken since the data quality criteria will likely change as the satellite retrieval algorithms evolve. The individual MOT PI's will determine algorithm constraints.

5.0 REFERENCES

- André, J.-M. and A. Morel, Atmospheric corrections and interpretation of marine radiances in CZCS imagery, revisited, *Oceanologica Acta*, **14**, 3-22, 1991.
- Austin, R. W. and T.J. Petzold, Remote sensing of the diffuse attenuation coefficient of sea water using the Coastal Zone Color Scanner, in *Oceanography from space*, edited by J. R. F. Gower, pp. 239-256, Plenum Press, New York, 1981.
- Ball Aerospace Division, Boulder CO, Development of the Coastal Zone Color Scanner for Nimbus-7, Volume 2 -- Test and performance data, Final Report F78-11, Rev. A NASA Contract NAS5-20900, May 1979.
- Bowman, K. P. and A. J. Krueger, A global climatology of total ozone from the Nimbus 7 total ozone mapping spectrometer, *J. Geophys. Res.*, **90**(D5), 7967-7976, 1985.
- Bricaud, A. and A. Morel, Atmospheric corrections and interpretation of marine radiances in CZCS imagery: Use of a reflectance model, *Oceanologica Acta*, **7**, 33-50, 1987.
- Casey, K. S. and P. Cornillon, 1999. "A Comparison of Satellite and In Situ Based Sea Surface Temperature Climatologies", *Journal of Climate*. In press.
- Clark, D. K., Phytoplankton algorithms for the Nimbus-7 CZCS, in *Oceanography from Space*, edited by J. R. F. Gower, pp. 227-238, Plenum Press, New York, NY, 1981.
- Cox, C. and W. Munk, Measurements of the roughness of the sea surface from photographs of the sun's glitter, *J. Opt. Soc.Am.*, **44**, 838-850, 1954.

- Evans, R. H. and H. R. Gordon, CZCS "System Calibration:" A retrospective examination, Submitted to the *Journal of Geophysical Research --- Oceans*
- Feldman, G. C., N. Kuring, C. Ng, W. Esaias, C. R. McClain, J. Elrod, N. Maynard, D. Endres, R. Evans, J. Brown, S. Walsh, M. Carle and G. Podestá. Ocean Color: Availability of the Global Data Set, *EOS Trans. Amer. Geophys. Union*, **70**, 634-641, 1989.
- Fraser, R. S. and Y. J. Kaufman, Calibration of satellite sensors after launch, *Appl. Optics*, **25**, 1177-1185, 1986.
- Gordon, H. R., Removal of atmospheric effects from satellite imagery of the oceans, *Appl. Optics*, **17**, 1631-1636, 1978.
- Gordon, H. R., A preliminary assessment of the Nimbus-7 CZCS atmospheric correction algorithm in a horizontally inhomogeneous atmosphere, in *Oceanography from Space*, edited by J. R. F. Gower, pp. 257-266, Plenum Press, New York, NY, 1981a.
- Gordon, H. R., Reduction of error introduced in the processing of coastal zone color scanner-type imagery resulting from sensor calibration and solar irradiance uncertainty, *Appl. Optics*, **20**, 207-210, 1981b.
- Gordon H. R., Calibration requirements and methodology for remote sensors viewing the oceans in the visible, *Remote Sens. Environ*, **22**, 103-126, 1987.
- Gordon, H. R., Radiometric considerations for ocean color remote sensors, *Appl. Optics*, **29**, 3228-3236, 1990.
- Gordon, H. R., Some studies of atmospheric optical variability in relation to CZCS atmospheric correction, NOAA National Environmental Satellite and Data Information Service, Final Report Contract No. NA-79-SAC-00714, February 1984.
- Gordon, H. R., J. W. Brown, O. B. Brown, R. H. Evans and D. K. Clark, Nimbus-7 Coastal Zone Color Scanner: reduction of its radiometric sensitivity with time, *Appl. Optics*, **22**, 3929-3931, 1983.
- Gordon, H. R., J. W. Brown and R. H. Evans, Exact Rayleigh scattering calculations for use with the Nimbus-7 Coastal Zone Color Scanner, *Appl. Optics*, **27**, 862-871, 1988.
- Gordon, H. R., O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker and D. K. Clark, A semi-analytic radiance model of ocean color, *J. Geophys. Res.*, **93D**, 10909-10924, 1988.
- Gordon, H. R. and D. J. Castaño, The Coastal Zone Color Scanner atmospheric correction algorithm: Multiple scattering effects, *Appl. Optics*, **26**, 2111-2122, 1987.
- Gordon, H. R. and D. K. Clark, Atmospheric effects in the remote sensing of phytoplankton pigments, *Boundary-Layer Meteorology*, **18**, 299-313, 1980.

- Gordon, H. R. and D. K. Clark, Clear water radiances for atmospheric correction of Coastal Zone Color Scanner imagery, *Appl. Optics*, **20**, 4175-4180, 1981.
- Gordon, H. R., D. K. Clark, J. W. Brown, O. B. Brown, R. H. Evans and W. W. Broenkow, Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison between satellite determinations and Coastal Zone Color Scanner estimates, *Appl. Optics*, **22**, 20-36, 1983.
- Gordon, H. R., D. K. Clark, J. L. Mueller and W. A. Hovis, Phytoplankton pigments derived from the Nimbus-7 CZCS: initial Comparisons with surface measurements, *Science*, **210**, 63-66, 1980.
- Gordon, H. R. and A. Y. Morel, *Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review*, Springer-Verlag, New York, 1983, 114 pp.
- Gordon, H. R. and M. Wang, Surface roughness considerations for atmospheric correction of ocean color sensors. 1: The Rayleigh scattering component, *Appl. Optics*, **31**, 4247-4260, 1992a.
- Gordon, H. R. and M. Wang, Surface roughness considerations for atmospheric correction of ocean color sensors. 2: Error in the retrieved water-leaving radiance, *Appl. Optics*, **31**, 4261-4267, 1992b.
- Hansen, J. E. and L. D. Travis, Light scattering in planetary atmospheres, *Space Science Reviews*, **16**, 527-610, 1974
- Hofmann, D. J., Aerosols from past and present volcanic emissions, in *Aerosols and Climate*, edited by P. V. Hobbs and M. P. McCormick, pp. 195-214, A. Deepak Publishing, Hampton, VA, 1988.
- Hooker, S. B., W. E. Esaias, G. C. Feldman, W. W. Gregg and C. R. McClain, *SeaWiFS Technical Report Series: Volume 1, An overview of Sea WiFS and Ocean Color*, NASA Technical Memorandum 104566, July 1992.
- Hovis, W. A., D. K. Clark, F. Anderson, R. W. Austin, W. H. Wilson, E. T. Baker, D. Ball, H. R. Gordon, J. L. Mueller, S. Y. E. Sayed, B. Strum, R. C. Wrigley and C. S. Yentsch, Nimbus 7 coastal zone color scanner: System description and initial imagery. *Science*, **210**, 60-63, 1980.
- Hovis, W. A., J. S. Knoll and G. R. Smith, Aircraft measurements for calibration of an orbiting spacecraft sensor, *Appl. Optics*, **24**, 407-410, 1985.
- Kenizys, F. X., E. P. Shettle, W. O. Gallery, J. H. Chetwynd, L. W. Abreu, J. E. A. Selby, S. A. Clough and R. W. Fenn, Atmospheric Transmittance/Radiance: The LOWTRAN 6 Model, Air Force Geophysics Laboratory, Hanscomb AFB, MA 01731, AFGL-TR-83-0187, 1983, NTIS AD A137796.

- Kilpatrick, K.A., G.P. Podesta, and R.H. Evans, 1998. AVHRR Pathfinder Oceans Sea Surface Temperature Global Area Coverage Processing Version 4.0. University of Miami/Rosenstiel School for Marine and Atmospheric Science technical report.
- Koepke, P., Vicarious satellite calibration in the solar spectral range by means of calculated radiances and its application to Meteosat, *Appl. Optics*, **21**, 2845-2854, 1982.
- Kumar, A., 1999. AVHRR Pathfinder: algorithm performance. International Scientific Symposium on Biogeochemistry of the Arabian Sea: Synthesis and Modeling. Poster presentation.
- Mie, G., Beiträge zur Optik trüber Medien, speziell kolloidalen Metall-lösungen, *Ann. Phys.*, **25**, 377-445, 1908.
- Mueller, J. L., Nimbus-7 CZCS: Confirmation of its radiometric sensitivity decay rate through 1982, *Appl. Optics*, **24**, 1043-1047, 1985.
- Mueller, J. L., Nimbus-7 CZCS: Electronic overshoot due to cloud reflectance, *Appl. Optics*, **27**, 438-440, 1988
- Neckel, H. and D. Labs, The solar radiation between 3300 and 12500 Å, *Solar Phys.*, **90**, 205-258, 1984.
- Salomonson, V. V., W. L. Barnes, P. W. Maymon, H. E. Montgomery and H. Ostrow, MODIS: Advanced facility instrument for studies of the earth as a system, *IEEE Geosci. Rem. Sens.*, **27**, 145-152, 1989.
- Shettle, E. P. and R. W. Fenn, Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties, Air Force Geophysics Laboratory, Hanscomb AFB, MA 01731, AFGL-TR-79-0214, 1979.
- Slater, P. N., S. F. Biggar, R. G. Holm, R. D. Jackson, Y. Mao, M. S. Moran, J. M. Palmer and B. Yuan, Reflectance- and radiance-based methods for the in-flight absolute calibration of multispectral sensors, *Remote Sens. Environ.*, **22**, 11-37, 1987.
- Smith, R. C. and W. H. Wilson, Ship and satellite bio-optical research in the California Bight, in *Oceanography from Space*, edited by J. F. R. Gower, pp. 281-294, Plenum, New York, NY, 1981.
- Vigroux, E., Contribution a l'étude expérimentale de l'absorption de l'ozone, *Ann. Phys. Paris*, **8**, 709-762, 1953.
- Viollier, M., Radiance calibration of the Coastal Zone Color Scanner: A proposed adjustment, *Appl. Optics*, **21**, 1142-1145, 1982.
- Viollier, M. and B. Sturm, CZCS data analysis in turbid coastal water, *J. Geophys. Res.*, **89**(D4), 4977-4985, 1984.

Wang, M., Atmospheric Correction of the Second Generation Ocean Color Sensors, 1991, Ph.D. Dissertation, University of Miami, Coral Gables, FL, 135pp.

Appendix 1 Equal-area gridding scheme

Equal-area gridding scheme implemented for MODIS

A1.1 Introduction

This document describes the equal-area gridding scheme proposed by the RSMAS Remote Sensing Group for the binned ocean fields produced for MODIS. The same approach has been adopted for AVHRR Ocean Pathfinder SST and 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 SeaWiFS or other 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. These considerations are presented in Appendix 3.

A1.2 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 and can accommodate bin sizes of 1 and 4 km as well. 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, 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., $2R_e/9.28$), where R_e is the equatorial radius of the Earth ($R_e = 6378.145$ 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 A1-1.

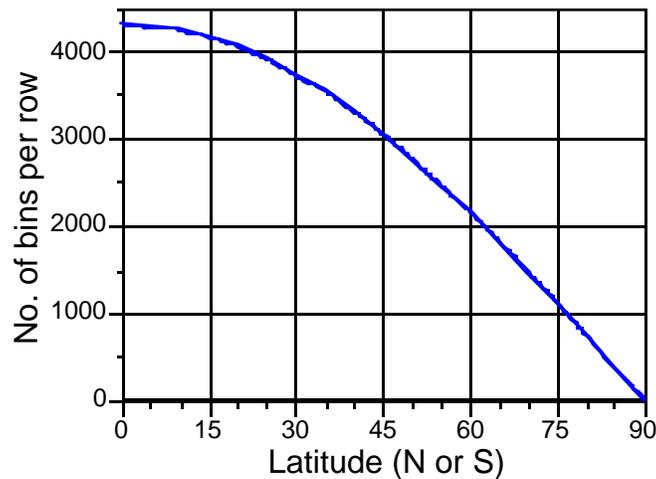


Figure A1-1. Number of 9.28 km tiles per zonal row as a function of latitude (North or South). The number of tiles is 4320 at the Equator and decreases to 3 at the poles.

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. The bins, however, 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 (N_p) at a given latitude (L) is computed as

$$N_p = 2r / X,$$

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

$$r = R_e \cos(L),$$

where R_e is the equatorial radius of the Earth. If the fractional part of N_p is greater or equal than 0.5, then N_p is rounded up to the nearest integer (i.e., the final number of tiles will be the integer portion of N_p plus one), otherwise N_p is rounded down (the final number of tiles is the integer portion of N_p). 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 ($2r$) 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 2.5° (a typical output resolution of atmospheric circulation models).

A1.3 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 X^2 , where $X = 9.28$ km. If we express the area of the circle as a rectangle of height X , the remaining dimension is X . If we divide the perimeter by three (to yield three tiles), each tile will have dimensions X by $X/3$ (approximately 1.05 X). That is, the bases of the triangular polar tiles are about 5% larger than the x-length of the equatorial tiles.

A1.4 Binning software

Several routines have been developed to perform the principal transformations required for binning and mapping data, such as converting latitudes and longitudes into bin numbers. Other routines perform the inverse transformation, that is, given a bin number they return a latitude and longitude corresponding to the centroid of that bin. These routines use a common initialization routine that must be executed prior to calling the conversion routines.

Appendix 2 Data Day Definition

A2.1 Introduction

The basic products generated by both the AVHRR Pathfinder and the SeaWiFS projects are global daily fields of geophysical quantities such as sea surface temperature and chlorophyll concentration. This definition is proposed for use with global fields generated from MODIS products. The daily fields will be the basis of subsequent temporal compositing into weekly and monthly products. One basic question, however, is: what constitutes “a day’s worth” of data? This is the question we address in this document.

The need for a consistent definition of a data-day is only really relevant for the production or analysis of *global* data fields. If one is dealing with a limited area (although, in this case, “limited” means anything less than global, and can encompass entire ocean basins), one takes advantage of the fact that the satellite sensors usually sample a region at *approximately* the same time(s) every day. In this way, data separated by approximately 24-hour periods can be assigned to different data-days (a further separation into daytime and nighttime fields can be made with the AVHRR). Analyses of the resulting daily data fields will introduce a minimal amount of temporal aliasing, as the difference in sampling times is of the order of a couple of hours over an approximate repeat cycle of a few days.

In contrast, when daily global satellite data fields are to be constructed, a consistent definition of a data-day needs to be adopted. This definition should be easy to implement in practice and should minimize temporal aliasing and discontinuities in the resulting products. In the following paragraphs we explore some of the alternatives.

A2.2 A 24-hour data day

The most obvious definition of a data-day is a 24-hour period. For instance, a daily field would encompass all the data collected between 00:00:00 UTC (or any other arbitrary start of the day) and 23:59:59 UTC. This definition is simple, intuitive, and extremely easy to implement. Its negative aspects, however, become apparent when one considers the orbital characteristics of the spacecraft on which the sensors of interest are, or will be, flown.

To illustrate the problem, we present a plot of nadir tracks for the NOAA-11 spacecraft (Figure A3-1). To simplify the visualization we only display descending tracks (i.e., the spacecraft is flying from north to south). The NOAA descending tracks correspond to nighttime data, although, in the case of SeaWiFS, the descending tracks will correspond to the daytime data (the only data archived for this sensor, other than special calibration measurements).

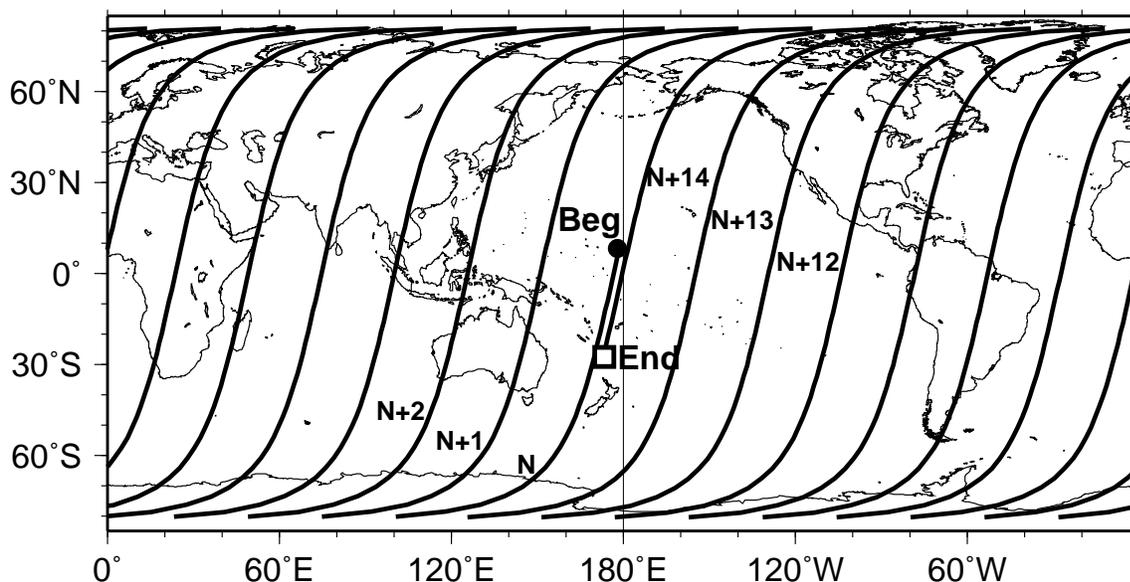


Figure A2-1. Descending NOAA-11 tracks for a 24-hour data-day beginning on July 26 1992 15:22:00 UTC. The data-day begins at the point labeled “Beg” and ends at the square labeled “End”. The first orbit after the beginning of the data-day is labeled “N”, and subsequent orbits are “N+1”...“N+14”.

For comparison with subsequent cases, we choose to begin the 24-hour data-day on July 26, 1992 at 15:22:00 UTC, when the nadir track intersects the 180° meridian (Marked

“Beg” on Figure A3-1). The descending orbit immediately after the beginning of the data-day is labeled N. Subsequent descending tracks pass to the west, and are offset by a distance of about 25 degrees of longitude at the Equator. The swaths viewed by the AVHRR in consecutive orbits have an increasingly larger overlap with latitude. This means that areas at intermediate and high latitudes may be sampled twice or more during a data-day (ignoring, for the time being, the ascending orbits). When an area is sampled in two consecutive descending orbits, measurements will be separated by about an hour and a half. Unless one is concerned with features with very small scales, it is probably safe to assume that the ocean fields will not change significantly between consecutive passes, thus temporal aliasing should be negligible.

The NOAA polar platforms that carry the AVHRR have an orbital period of approximately 102 minutes. The actual period depends on the spacecraft altitude, and, therefore, will be slightly different for each NOAA spacecraft. The orbital period may also vary with time, as the altitude of a satellite changes. Given an orbital period of about 102 minutes, the number of revolutions that the NOAA spacecraft will complete in a 24-hour period is approximately 14.12. The last descending orbit of the 24-hour data day is labeled N+14. It is apparent from Figure A3-1 that there is along-track overlap between descending tracks N and N+14. The areas in which there is overlap will have been sampled twice (ignoring the smaller overlap between consecutive orbits) in a data-day, and the output will be the average of measurements taken almost 24 hours apart. The 24-hour definition of a data-day, therefore, may result in temporal aliasing in areas near the beginning and end of the 24-hour period, due to the inclusion in a given day of overlapping orbit tracks sampled almost 24 hours apart.

A second problem of the temporal definition of a data-day is the existence of areas on the global fields with large temporal discontinuities in sampling times, even though they may be spatially contiguous. For instance, in Figure A3-1, one can see descending track N+14, the last track of the data-day. To the north of that track (i.e., over the Arctic Ocean north of Alaska), data are contributed by track N+1 and, possibly, N+2. These two tracks, however, were sampled near the beginning of the data-day, more than 20 hours before track N+14. The daily fields, then, will contain large temporal discontinuities along the boundaries between data swaths from tracks N+14 and N+1. If there is overlap between the two swaths, data collected far apart in time may be averaged, once again introducing potential aliasing. Similar problems occur in the area

south of track N (south of New Zealand), which is sampled by tracks N+13 and N+12 much later in the day.

The aliasing and temporal discontinuity effects are further complicated by the fact that the locations where they occur change in time. Figure A3-2 shows the locations along the nadir tracks of the boundaries between 24-hour data-days for a 10-day period beginning on July 26, 1992 (for the NOAA-11 spacecraft). The dot labeled “1” corresponds to the beginning of the cycle on July 26, 1992 at 15:22:00 UTC. The dot labeled “2” indicates the beginning of the second 24-hour data-day, and so forth. The shift in the location of the daily boundaries is a direct result of the difference between the 24-hour data-day and the shorter time it takes the spacecraft to complete a number of revolutions that would ensure global coverage.

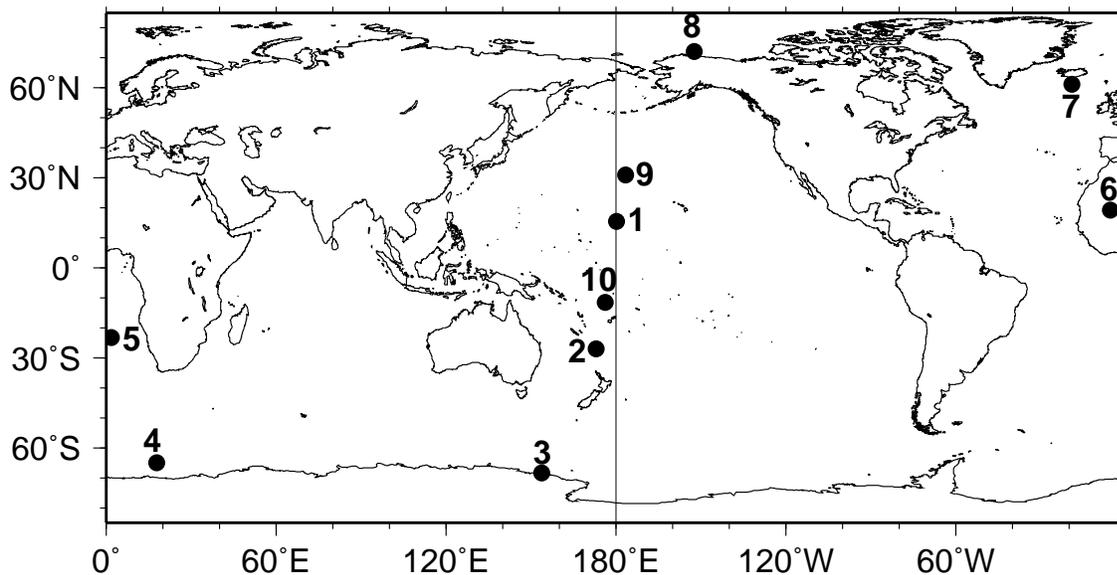


Figure A2-2. Locations of the boundaries of 24-hour data days for a 10-day period beginning on July 26, 1992 15:22:00 UTC (dot labeled “1”).

A2.3 A spatial data-day definition

Because of the problems associated with a temporal definition of a data-day, we explored the implications of adopting a spatial definition. In this case, the boundary between data-days is not defined by time but, instead, by a fixed geographic reference. A similar criterion is commonly used for designating orbit numbers in several

spacecraft: the orbit number usually is incremented upon crossing the Equator. For the initial investigations, we selected the 180° meridian as the boundary between data-days.

Figure A2-3 shows NOAA-11 nadir tracks for a spatially-defined data day. Because the nadir tracks crosses the reference line several times during a day, one of the crossings must be selected as the beginning of a data-day. An operational definition of this is presented below. For this discussion, we define the day to begin on July 26, 1992 at 15:22:00 UTC, when the spacecraft crosses the 180° meridian flying from north to south (i.e., at the same time at which the 24-hour data-day shown on Figure A3-1 started). The first descending track of the day is labeled N.

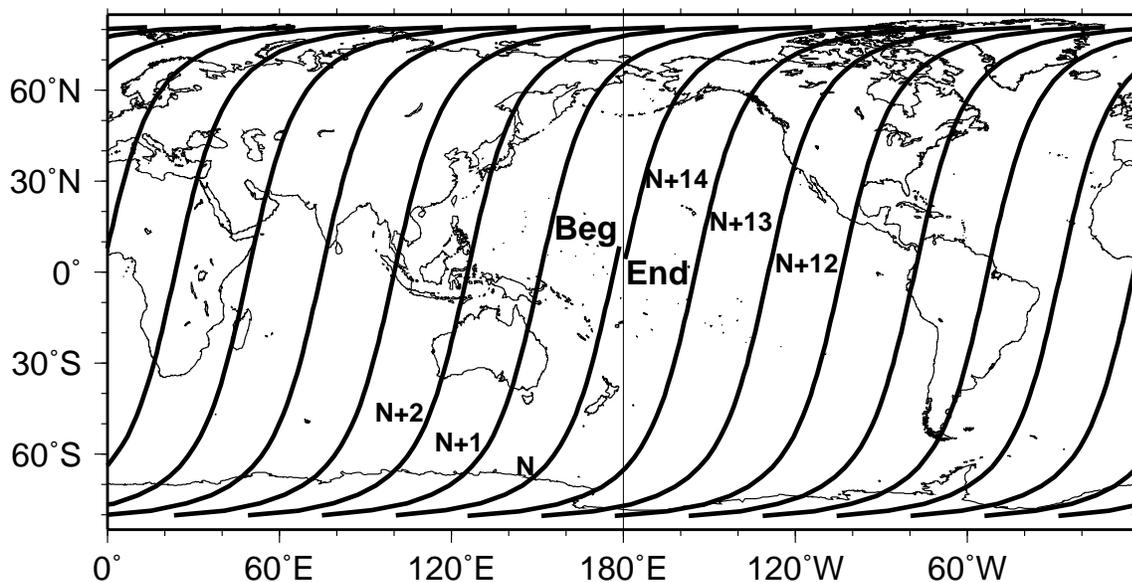


Figure A2-3. NOAA-11 descending orbits for a spatially-defined data-day beginning on July 26 1992, 15:22:00 UTC. At this time, the nadir track crosses the 180° meridian.

As almost 24 hours worth of data are required to ensure global coverage, we define the data-day as ending when the nadir track crosses the 180° meridian during revolution N+14. This happens approximately on June 27, 1992 at 15:14:00 UTC. The most immediate observation, then, is that a spatial definition results in a data-day that does not necessarily correspond to a 24-hour day: in this case the data-day is approximately 23 hours and 52 minutes long. This figure is only approximate for two reasons. In the first place, it is sometimes necessary to include an additional revolution in order to ensure global coverage (that is, the last orbit of the day would be N+15). Secondly, the spatial definition is applied on a pixel-by-pixel basis. That is, pixels along the same scan

line on a given orbit can be assigned to different days depending on whether they are on one side or the other of the 180° meridian.

Figure A3-4 illustrates the pixel-by-pixel assignment of data to a given day. The figure shows the sampling pattern of the AVHRR onboard NOAA-11 between 15:12:00 and 15:32:00 UTC on July 26, 1992 (i.e., ± 10 minutes from the start of the data-day at 15:22:00 UTC). The scan lines shown on Figure A3-4 are separated by one minute (in a one-minute interval there are 360 LAC scans or 120 GAC scans). Pixels along a given scan line that are located east of 180° are assigned to day N. If pixels along the same scan line are west of 180°, those pixels are assigned to the following day (N+1). It is apparent from Figure A3-4 that, even before the nadir track crosses the 180° meridian, pixels are already being assigned to day N+1. Conversely, after the nadir track has crossed the reference meridian (at 15:22:00 UTC), pixels east of the meridian are still being allocated to day N. It is this allocation mechanism that makes it difficult to define precisely the duration of a data-day.

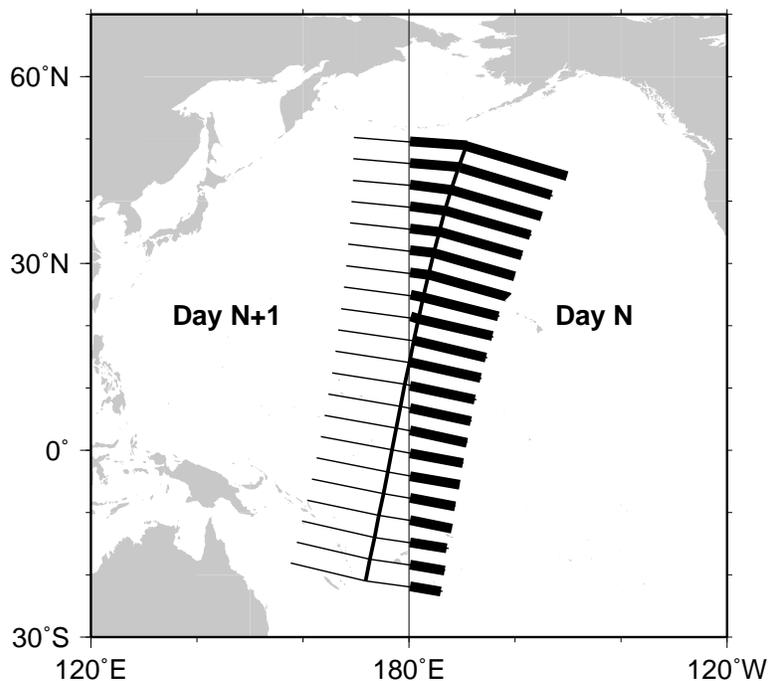


Figure A2-4. AVHRR nadir track and scan lines for a 20-minute period between 15:12:00 and 15:32:00 UTC on July 26, 1992. Pixels to the east of the 180° meridian (marked in a thicker line) get assigned to data-day N, whereas the pixels to the west of the meridian correspond to data-day N+1.

A2.4 How is the beginning of a data-day defined?

How is the spatial definition of a data-day implemented in routine processing of global satellite data fields? The first step is to define a meridian that will serve as the reference for the data-day definition. The 180° meridian used in the previous examples is a good alternative, as this choice minimizes differences between actual dates and the dates assigned to the data-days. As the spatial data-days are not 24-hours long, a suitable naming convention will have to be established.

A second step in defining a data-day is to decide which of the descending (or ascending) crossings of the reference meridian will mark the beginning of the descending (or ascending) data-days. As mentioned above, there are several (usually seven to nine) descending crossings of the reference meridian in a day; the same is true for ascending orbits. This is illustrated in Figure A3-5, which shows the latitude of descending crossings of the 180° meridian as a function of time for the NOAA-11 spacecraft, beginning on July 26, 1992; a 10-day span is shown. Most of the crossings (shown as dots) take place at high latitudes, and one or two crossings per day occur at tropical to intermediate latitudes.

Any of the crossings of the 180° meridian shown on Figure A3-5 can be potentially selected as the one marking the beginning of a data-day for descending and ascending orbits. For operational purposes, we propose the following definition: *a data-day for descending orbits is defined to begin at the descending crossing of the 180° meridian closest to the Equator.* A similar definition can be applied to ascending crossings, yielding data-days for ascending orbits. Such definition is the easiest to implement because there is always only one crossing in a day that fulfills the condition (although consecutive crossings may sometimes have very similar absolute latitudes of intersection, one on the Southern hemisphere, and the other on the Northern hemisphere).

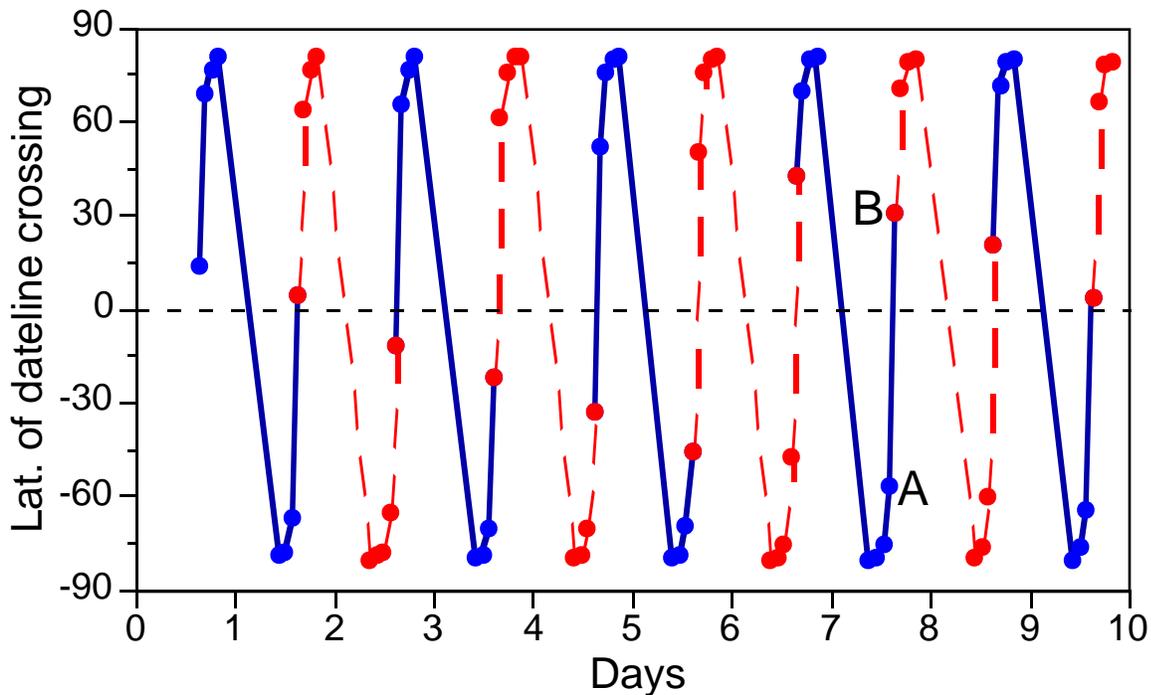


Figure A2-5. Latitude of crossing of the 180° meridian for NOAA-11 descending orbits. Data shown for a 10-day period beginning on 26 July 1992, 15:22:00 UTC. The alternating solid and dashed lines indicate consecutive data-days.

The alternating solid and dashed lines in Figure A3-5 indicate consecutive data-days. Initially, the latitude of the data-day beginning seems to follow a regular progression to the south. For instance, the first two data-day boundaries in Figure A3-5 are on the Northern hemisphere, and the next four are progressively further south on the Southern hemisphere. Note, however, that the southward progression is interrupted near the end of data-day 6 (the point labeled A). In this case, the next descending crossing (point labeled B) is actually closer to the Equator, so the data-day is extended until this next crossing (located in the Northern hemisphere). That is, the data-day is slightly longer (one more revolution) in this case. The southward progression of the crossings subsequently resumes. For ascending orbits, the progression of the latitude of the crossings is reversed, that is, it occurs from south to north.

Table A2-1 contains a list of start times of descending data-days for a 15-day period beginning on July 26, 1992, as well as the latitude at which the crossing of the 180° meridian occurs.

Date	Beginning time	Latitude of 180° crossing
07/26/92	15:22:04	13.85
07/27/92	15:13:39	0.67
07/28/92	15:05:14	-12.52
07/29/92	14:56:34	-24.81
07/30/92	14:47:28	-35.52
07/31/92	14:37:51	-44.37
08/01/92	15:43:42	39.33
08/02/92	15:34:24	29.34
08/03/92	15:25:35	17.60
08/04/92	15:17:06	4.63
08/05/92	15:08:42	-8.64
08/06/92	15:00:08	-21.29
08/07/92	14:51:11	-32.58
08/08/92	14:41:43	-41.99
08/09/92	15:47:31	41.94

Table A2-1. Beginning times of fifteen data-days for descending orbits , NOAA-11 spacecraft. The latitude of the 180° meridian crossing is also shown.

We must stress that, because of the pixel-by-pixel allocation described above, parts of the field will include data collected both before and after the times listed in Table A3-1. Notice the jump in the southward progression of crossing latitudes (e.g., from July 31 to August 1), which is associated with a slightly longer data-day.

A2.5 Advantages of the Spatial Definition of a Data-Day

In previous sections we proposed a spatial definition for a data-day, together with an objective definition for the temporal “beginning” and “end” of such a data-day. So far, however, we have not discussed the advantages or disadvantages of the proposed definitions.

Some problems associated with a temporal definition of the data-day were the potential presence of aliasing and large temporal discontinuities, and the fact that the day boundaries changed with time. The spatial definition avoids temporal changes in the location of boundaries, as the boundary is fixed (e.g., the 180° meridian). Furthermore, because there is no overlap of swaths at the beginning and end of a data-day, the spatial

definition reduces the aliasing resulting from averaging data sampled almost 24 hours apart. The presence of large temporal discontinuities among adjacent areas is still present, however.

The large temporal discontinuities identified on Figure A3-1 north of Alaska and south of New Zealand are still present in Figure A3-3. It is clear that the large temporal discontinuities occur in two places near the meridian that defines the separation between data-days. The first place is the area south of the first track of the data-day and west of the reference line. The second area with discontinuities occurs north of the last track of the data-day, east of the reference line. In addition to the large temporal discontinuities between adjacent swaths, when the swaths overlap at higher latitudes once again data will be averaged that were sampled far apart in time. Elsewhere on the global fields, any given track is surrounded by tracks sampled one orbital period (about 100 minutes) earlier or later.

The presence of temporal discontinuities or the averaging of data collected at very different times may not be too important for many applications, although users should certainly be made aware of the occurrence of these events. In other situations, however, the temporal discontinuities may cause significant problems. Examples of such applications may be the estimation of the translation speed of certain features, or the computation of fluxes.

In order to limit the large meridional temporal discontinuities near the data-day boundary, the short track segments north and south of the first and last tracks of the data-day could simply be eliminated (e.g., parts of N+1, N+2, N+12 and N+13). This approach is illustrated in Figure A3-6, which shows descending tracks between July 27 1992, 15:14:00 UTC and July 28 1992, 15:05:00 UTC (the data-day following the one shown on Figure A3-3). The map is now centered at 0°, rather than at 180° as in Figure A3-3. Note that the nadir tracks for which segments were eliminated seem to end a bit before or after the 180° line. This is because positions were predicted at one-minute increments.

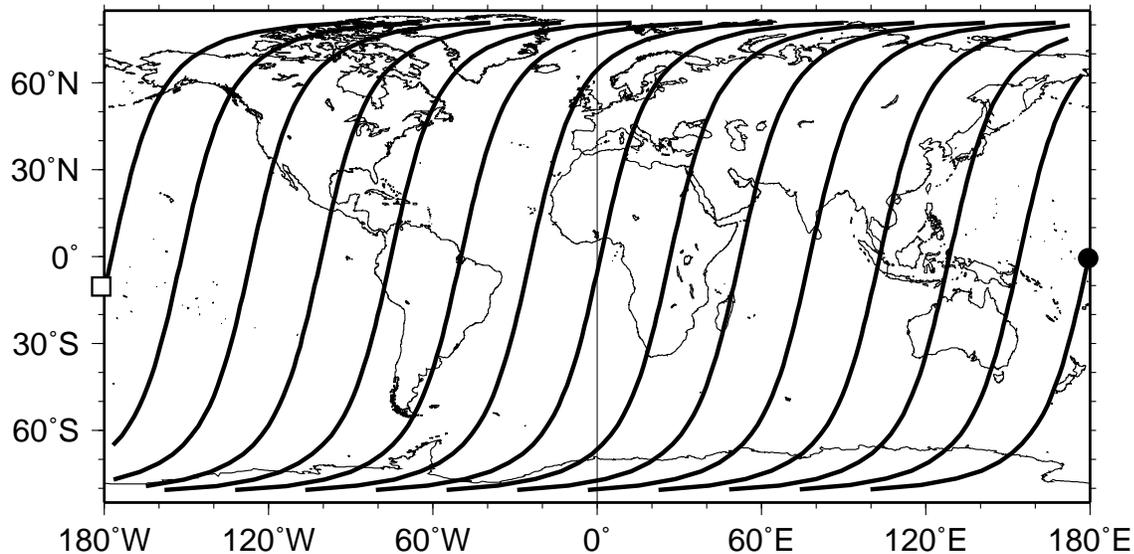


Figure A2-6. NOAA-11 descending orbits for spatially-defined data-day beginning on July 27, 1992 15:14:00 UTC. Segments that introduce large north-south temporal discontinuities (see text) are excluded.

The elimination of segments may result in areas not being sampled (e.g., upper left and lower right corners of the map). These gaps may possibly be filled by the wide swath of the first and last tracks of the data-day (tracks N+14 and N in the north and south, respectively). However, the size of the gaps is a function of the latitude of the reference line crossing which defines the beginning of the data-day. As shown above (Figure A3-5), this latitude changes with time, moving north and south approximately between 60°N and 60°S. When the crossing is further north, the gap to the south of the first track will be larger. Conversely, when the crossing is further south, the gap north of the last track will get larger. We propose that two additional swaths be added at each end of the data-day, in order to replace the eliminated segments. Experience has shown that two additional swaths are enough to fill each of the gaps and ensure complete coverage. The added swaths would be temporally continuous with the first and last tracks of each data-day, thus eliminating the problems of temporal discontinuities. An operational scheme would involve the following steps:

1. The times of the beginning and end of a spatially-defined data-day are found following the definition suggested above. These times will be referred to as the “beginning” and “end” of the data-day.
2. Data east of the 180° meridian and collected up to 216 minutes (about two orbits) after the beginning of the data day will be excluded. Data west of the 180° meridian

sampled up to 216 minutes before the end of the data-day will be similarly excluded. The net result of these actions is similar to the elimination of segments shown on Figure A3-6.

3. To ensure full coverage, data collected up to 216 minutes before the beginning of the data-day and west of the 180° meridian are added to the data-day. This fills the gap to the south of the first track of the day. Data collected up to 216 minutes after the end of the data-day and east of 180° are also added. These data fill the gap north of the last track of the data-day. The end result is illustrated on Figure A3-7. Note that only the descending (AVHRR night or ascending day) portions of the extra orbits are included in the fields.

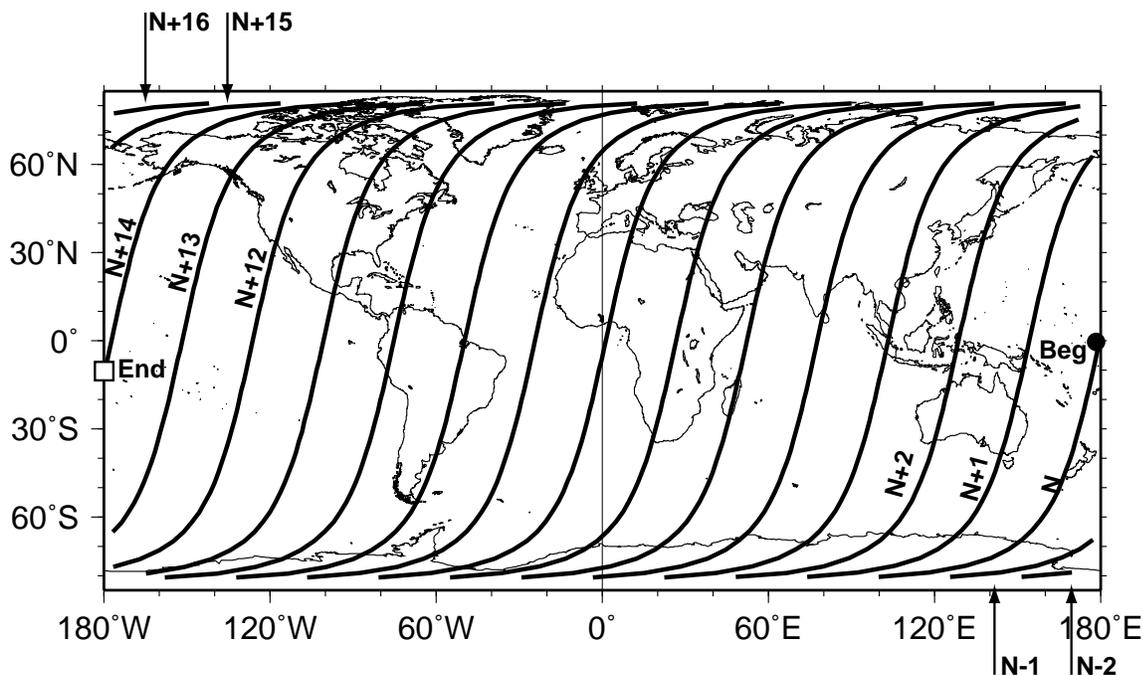


Figure A2-7. Data-day beginning on July 27 1992 15:14:00 UTC, showing the addition of four segments (indicated by arrows) in order to minimize temporal discontinuities. The first track sampled after the estimated beginning time of the day (“Beg”) is track N. The two segments to the south correspond to the two previous orbits (N-1, N-2). The last track before the estimated end time of the data-day (“End”) is track N+14. The two segments to the north correspond to the next two orbits (N+15, N+16).

Figure A2-7 shows the descending orbits for the data-day beginning approximately on July 27 1992, 15:14:00 UTC. The gaps shown on Figure A3-6 have been filled by the addition of four short segments, indicated by arrows on Figure A3-7. Note that these segments have been sampled before (N-1, N-2) and after (N+15, N+16) the times

estimated for the beginning and end of this data-day (see Table A3-1). However, because the added segments are close in time to orbits N and N+14, the large temporal discontinuities have been eliminated. The segments excluded from this data-day are the first portion of tracks N+1 and N+2, east of 180°, and the last portion of tracks N+12 and N+13, west of 180°.

Admittedly, it is somewhat difficult to grasp the methodology proposed. To facilitate comprehension, we may present a simple analogy. Envision a continuous strip chart on which the continents are drawn. Above the chart recorder there is a clock showing UTC time and date. As the chart moves from left to right, a pen draws descending tracks, one at a time. The speed of the chart movement is appropriate to ensure that the nadir track's latitude and longitude corresponding to any given UTC time are correct. That is, the nadir tracks should look similar to those on Figures A3-6 and A3-7.

Suppose we position the chart so that the pen is just crossing the 180° meridian near the Equator on July 27 1992. The time shown by the clock should be about 15:14:00 UTC. We then allow the chart recorder to run for almost 24 hours, until a track crosses the 180° meridian again very close to the Equator. The time should be about 15:04:00 UTC on July 28, 1992. If we cut the chart along the two 180° meridians drawn (left and right), the tracks on the chart should look exactly like Figure A3-6. As in Figure A3-6, there will be some gaps in the coverage. On the right side of the chart, there is a gap south of the first track (N) of the day. This gap should have been filled by the last portion of tracks N+12 and N+13, which have been drawn to the left of the 180° meridian on the left side of the chart. These lines, however, were eliminated when we cut the chart along the left 180° line. Similarly, the gap north of the last track of the day should have been filled by the initial portions of tracks N+1 and N+2. These segments were drawn east of the 180° meridian on the right side of the plot. However, as we cut along the 180° line on the right, these segments were excluded. It is apparent, then, that the chart recorder analogy reproduces the action of eliminating tracks which cause large temporal discontinuities, as the end result looks exactly like Figure A3-6. Let us see if we can fill the gaps in the global fields using the same chart recorder analogy.

Suppose that we do not start the chart recorder at 15:14:00 UTC on July 27 1992 but, rather, we move the chart backwards and start about 216 minutes earlier. If we start the recorder then, a few additional tracks (e.g., N-1 and N-2) will be drawn before the nadir track of orbit N crosses the 180° meridian at 15:14:00 UTC (defined as the temporal

beginning of the data-day). The southern segments of tracks N-1 and N-2 will fall west of the 180° meridian, filling the gap previously existing in the south. We let the recorder run up to 216 minutes past the time estimated for the end of the day (July 28 1992, 15:04:00 UTC) and, again, a few additional tracks will be drawn. If the last track of the day is N+14, the northern portions of tracks N+15 and N+16 will fill the northern gap. Once we have allowed the recorder to run for about 24 hours plus the additional 216 minutes on either end, we take a pair of scissors and cut the chart along both 180° meridians. That is, we are applying the spatial pixel-by-pixel assignment of data to a given data-day. The end result should look exactly like Figure A3-7. Finally, we could envision running the recorder for long periods and repeatedly cutting the long chart along the 180° meridians. Each of the maps would correspond to one data-day.

When discussing the elimination of segments that caused large temporal discontinuities, we could have given the impression that the data in these segments would be unused, and therefore wasted. However, if one follows the analogy presented above, it is easy to see that the data will not be deleted but, rather, they will be assigned to the previous or the following data-days. For instance, the northern portions of tracks N+1 and N+2 (not labeled) in Figure A3-6 would be plotted to the east of the right 180° meridian on the chart. When we cut the chart, these portions get assigned to the previous data-day, which begins on July 26 1992, 15:22:00 UTC. In the same way, the southernmost portions of tracks N+12 and N+13 are plotted to the west of the left 180° meridian, thus being assigned to the next data-day after the chart is cut along the meridian. The end result of the scheme proposed is a daily global field where all parts of a field are temporally separated from adjacent areas by, at most, one orbital period.

A similar scheme can be implemented for ascending tracks. The definitions of the temporal beginning and end of an ascending data-day were discussed above. The chart recorder analogy can also be similarly formulated for ascending nadir tracks.

A2.6 Other Issues

An aspect that we have not discussed so far is that at both the extreme north and south of the fields, data from several tracks will be averaged for a data-day. At high latitudes, the spacecraft is flying almost in an east-west direction and, thus, the scan lines have a north-south orientation. For instance, there are seven to nine passes a day at high

latitudes (see Figure A3-5). Near the 180° meridian, where some of the passes are excluded at high latitudes, as described above. In other high latitude regions, however, the fields will contain the average of several passes. This should not have too many consequences on ocean products, as the areas affected will be mostly on land in the southern hemisphere and under permanent ice cover in the northern hemisphere.

One final issue that needs to be pointed out is that the spatial scheme proposed above will result in temporal discontinuities in areas that straddle the reference line. Suppose that an oceanographer is studying an area of the North Pacific Ocean encompassed between 150°W and 150°E, straddling the 180° line. If the oceanographer obtains a global field for a given data-day, he/she must realize that the portion of the study area west of 180° has been sampled much earlier than the portion to the east. Again, this may not be relevant for some research, but it could be in some cases. A solution would be to place the reference line elsewhere, for instance along 0°, but there will always be some location where areas on either side of the line will be sampled far apart in time. Alternatively, a user might obtain product fields for two consecutive data-days and paste the appropriate portions. Going back to the Pacific example presented above, the eastern part of the study area would be extracted from data-day X and the western part would be taken from day X+1. This can be accomplished without reprocessing and without introducing any spatial or temporal aliasing.