MODIS QUARTERLY REPORT
- September 1996 -

UNIVERSITY OF MIAMI
RSMAS/MPO

DR. ROBERT H. EVANS

NAS5-31362

==========

Due to the interlocking nature of a number of projects, this and subsequent reports will contain coding to reflect the funding source. MODIS funded activities are designated with an M, SeaWiFS with an S, and Pathfinder with a P. There are several major sections within this report; Database, client/server, matchup database, and DSP support.

A. NEAR TERM OBJECTIVES
B. OVERVIEW OF CURRENT PROGRESS
C. FUTURE ACTIVITIES
D. PROBLEMS

A. NEAR TERM OBJECTIVES

A.1 MODIS Objectives (M)
A.1.1. Continue to develop and expand the processing environment

a. increase computational efficiency through concurrent operations
b. determine and apply more efficient methods of data availability for processes

A.1.2. Begin extensive testing using global CZCS and AVHRR GAC data with database processing to test the following:

a. algorithm capability
b. machine and operating system stability
c. functionality required for the processing and analysis environment

A.2 SeaWiFS Objectives (S)
A.2.1. Continue testing of processing methodology.
A.2.2. Continue to develop relationship between database and in-situ environment.

A.3 Pathfinder Objectives (P)
A.3.1. Expand matchup database as applicable.
A.3.2. Continue testing of methodology.
A.3.3 Train and integrate new personnel into Matchup Database processing scheme.

A.4 DSP Objectives (M)

A.4.1. Continue testing of processing methodology.
A.4.2. Continue to expand the number of sites supported.
A.4.3. Expand the supported hardware/software platforms

B. OVERVIEW OF CURRENT PROGRESS

B.1 Automatic Processing Database (P)

B.1.1 Processing -Pathfinder

B1.1.1 July Processing

A pfv3 version was processed for the last 7 weeks in 88, and all of 89 and 90.

A weeks worth of binned files were provided to JPL for verification of processing.

A new set of scan-segregated files was produced.

B1.1.2 August Processing

A run of the new version, pfv4, was begun. A new "speed record" was established: 450 day's worth of GAC data was processed in five days without interruption.

Shortly after the period 87001 - 88315 was processed, electrical work at the school caused major problems in the DLT machines holding the input GAC data, requiring that the tapes be scanned for file content. This slowed the processing.

A new pfv4 was started, with new coefficients. The time period weeks 8601 to 8624 were calculated.

B1.1.3 September Processing

The new pfv4 run was continued, finishing 86, 87 and 88 through week 46 (the end of NO9).

When calculations started on 85, it was discovered that the input GAC files on the DLT machine were not correct, and 85 will be read in again from laser disk. This proceeded through September.

B.1.2 Development Activities

B.1.2.1 July Development

The only development work this month were minor changes to command files. The most recent Pathfinder processing method was transferred to both JPL and the University of Rhode Island.

B.1.2.2 August Development

(Should be covered elsewhere) A new method of determining quality at the pathnlc level was tested for NO9. This version of the pathfinder algorithm is to be called pfv4.

Usage of any files on the SGI machine, andrew was eliminated from the processing stream, and all input and out file locations were changed to disks mounted only on the DEC machines. (Due to a bug in the most recent XFS filesystem, use of some the andrew's disks by the processing machines was causing the SGI to crash.)

B.1.2.3 September Development

Development this month centered on the post-processing procedures (producing maps and time series products form the binned files), as well as setting up a data extraction stream, that will retrieve processed files, make several maps, and extract point data for time series construction.

Modifications were begun on the data spooling procedures, to streamline them, increase their efficiency, and possibly use them to construct a scan database, but the changes are only preliminary.

B.2 Processing Systems Status (M)

B.2.1 MODIS Version 1

B.2.1.1 MOCEAN IO Library

Integration of the mocean IO library into the version 1 code base was completed and included in the 9/17 delivery to SDST. The library now supports reading and writing of MODIS ocean level 2 and level 3 files and the MODIS geolocation file.

Due to the size of the global L3 files, the library was enhanced to utilize HDF external data files. The current HDF version (4.0) supports single files up to a maximum of 2 GB. External data files provide a mechanism for distributing data across multiple files. Unfortunately, this type of HDF file will not be supported by the Release. B Data Server.

B.2.1.2 HDF-EOS

HDF-EOS V1 was successfully integrated into the mocean IO library. The library now supports reading and writing to HDF-EOS v1 swath structures.

B.2.2 Build Procedure Enhancements

The GSFC DAAC's comments and recommendations following the Beta software delivery were reviewed. Based on this feedback the following enhancements were made and included in Sept. 17th delivery:

- All code was delivered in the directory structure requested by DAAC and SDST.

- Only F90, F77 and C code was delivered (no rat FORTRAN or mice tables).

- The legacy libraries used by MODIS code were separated and stripped down to include only routines used by MODIS.

- Shared libraries (.so) were converted to standard archive files (.a).

- All of ocean's pmake files were rewritten in standard make.

- Conditional build logic for separate platforms (such as VMS) was removed.

B.2.3 Version 1 HDF file formats

B.2.3.1 Level 2 EOS Swath

The Sept. 17 delivery included level 2 files in EOS swath format which is now supported as one of the standard data formats in mocean IO library.

B.2.3.2 Level 3 File Formats

MODIS Ocean Level 3 files HDF file formats which utilized HDF external data elements were proposed and submitted to SDST in August. External data elements allow data to be separated into any number of physical files which appear as a single HDF file. Using this feature the 2 GB HDF file size limit was maintained. Unfortunately, this scheme presents problems to the Release. B Data Server and this has forced a redesign of level 3 file formats.

The present proposal calls for each MODIS Ocean product to be stored in a separate Level 3 file of approximately 500 MB. This design increases substantially the number of files managed by the Data Server. This issue has been submitted to ECS. If this design also proves problematic, another alternative is to divide the products among 12 files, instead of 36, in order to reduce Data Server overhead.

B.2.4 Near Term Objectives

Integrate SDP Toolkit metadata handling routines into the Nov. Version 1 delivery.

Resolve outstanding issues with level 3 files. Meetings with Hughes and SDST are scheduled around the MODIS Science Team meeting in October.

Write scripts for PGEs to combine level 2, ocean color and SST (MOD_PR18 and MOD_PR28), and level 3 space binning (MOD_PR18E). This will reduce the number of PGE activations required by MODIS Ocean processing by 576 per day (number of level 1 granules/day x 2).

B.3 AVHRR SST retrospective processing

B.3.1 Near Term Objectives

Continue work on improving system durability by automating the process of redistributing work assigned to a machine that is no longer available

to the processing system.

Bring the programs and procedures under control of a source code management system (CVS).

B.4 Systems/COTS

B.4.1 Autosys

Platinum still has not certified it's client for IRIX 6.2. Platinum has asked us to participated in a beta software evaluation of the next client release for IRIX should we acquire a license for the SGI Challenge.

B.4.2 Sybase

Sybase Release 10 is not compatible with DUNIX 4.0. We continue to run the dataserver at DUNIX 3.2c for compatibility. Sybase has indicated certification of DUNIX 4.0 is in progress for Release 11.

C.3 ATM Local area networking

Two additional hosts, monstera and ceriman (4 processor Alpha 4100s), were added to the ATM network.

B.3 Matchup Database (P)

During this period, there was a major revision of the way in which cloud-flagging tests for the matchups are handled. Previously, we had implemented a series of very restrictive tests, based on brightness temperature (BT) thresholds, spatial homogeneity constraints, and channel-to-channel BT differences. The filters fulfilled their main purpose: to ensure that only a limited amount of cloud-contaminated matchups were used in the estimation of SST algorithms.

Although the cloud tests were mainly developed for the matchup database, they were also temporarily implemented in the processing of Pathfinder image products, while a consensus Pathfinder cloud-flagging scheme was adopted. It was observed that the end result of the cloud tests in the image products was that many pixels were being flagged as cloud-contaminated when in reality they were not.

At the time when we began to process the NOAA-9 matchups (1985–1988), we had to revise the thresholds for the filters. At this time, we tried an approach known as “classification trees”. In this procedure, we start with a training sample, for which whether a matchup is cloud-contaminated or not is supposed to be known. Using the training sample, a decision tree is “grown”, using various satellite quantities (e.g., differences between BTs among various channels). The tree procedure finds successive splits that maximize the separation between the two groups of matchups (“cloudy” and “not cloudy”).

The classification tree was implemented for the NOAA-9 matchups. The end result was that 60% more matchups were retained than with the earlier filters, while at the same time the proportion of cloud-contaminated pixels that were flagged as good was similar to that obtained with the older filters. The next step is to try this approach with NOAA-11 and NOAA-9 gap data.

Another change adopted during this period concerns the estimation of algorithm coefficients via multiple regression. We adopted an iterative procedure in which a robust regression procedure is first used. The robust regression is supposed to be very resistant to outlier points. The residuals from the robust regression are used as weights in a subsequent weighted linear regression step. The end result is that the values of the monthly coefficients are much more stable in time than before.

B.4 MODIS Algorithm Integration (M)

Third Quarter Activities.

B.4.1 Testing:
Tested and delivered V1 MODIS level 2 processing (modcol and modsst).

B.4.2 Modifications/Additions to DSP:
Change $NAMSIZ so that we have room for the MODIS path and file names.
Add prologs to MODIS source files.
Renamed programs and libraries so that the MODIS versions and the DSP versions have unique names.
Rearranged routines so that modules used by more than one program are in a library.
ATMCORSHR: New library which contains all of the routines and include files which are common to both modcol and modsst (except the routines which have to do with binning which are in binshr). Move splitandappend to binshr since it is now used by the binning routines also. get_image_name wasn't being used. Changes to get rid of f90 compiler warnings. Fix MODIS headers. ifdef out DSP environment stuff.
BINSHR: New library which contains all of the binning routines and include files which the MODIS programs share. This also includes the routines which read the global land and shallow water masks. splitandappend moved from atmcorshr since the binners use it. parsefilestring is no longer used. Fix check for directory separator in splitandappend.
Put all include files which are used by more than one program in modinc.
MODIS executables have .exe extension.
Fix all makefiles to use MODIS environment variables.
Add -n32 support. Add f90 support.
Change pcf logical numbers to be in the 300,000 range for MODIS oceans.
Fix .pcf files for latest toolkit version.
Fix use of PGS_SMF_* toolkit error handling routines.
Clean up include files so they look ok in f90 files.

B.4.3 Problems fixed:
MODCOL: Include proper PGS error message file; output calibration information as band attributes; fix calculation of the time of the input scan.

Implement complete algorithm for MOD31. Increase arrays to leave room for vagaries of input data. Correct string declaration. Scantime is now a single value (was two element vector before). l1b_bands_read is now an input from main program. Add secnds compatibility routine (for testing). Initial CFE/enhanced FLH code (Abbott). Output PEB and PUB to L2 file. Fix pel_* routines to work with SeaWiFS data converted to MODIS format. Add check for invalid input values. Take out old SeaWiFS and DSP stuff to make it easier to add MODIS prologues. Read HDF versions of bit masks using mocean routines instead of DSP files. Remove unused code. Scantime is an 8 byte real now. Get rid of blanks from the input mask/flag names. Change fold_arith_limit to 3134. Define Ftiltr and include string.h. Add attributes to output files to make the ncdump output look pretty. Fix MODIS error messages. Fix MODIS headers. Fix wang2.f to make f90 happy. Use swath file type.

MODSST: Get input files from .pcf file instead of assuming they are in $DSPROOT/cal. Include proper PGS error message file. Split is no longer in the mice table. Get input file name and path from pcf file. Start to add two more bands of input. Add another SST band to the output file. Allow short or real input pixels (using ifdef's); add second channel 5 uniformity check; add second SST output band; start to add two more input bands; a lot of little changes to do things more like modcol (and SeaWiFS) does; output other bands in QC file; write out band attributes; use pathfinder version 3 - 2 piece algorithm and coefficients. Adjust format statements to make radiance values readable. Add to documentation of USE_LAMBDA. Finish code for USE_LAMBDA (radiances in micro-meter vs cm^{-1}). Add more diagnostic printouts. Put radiation constants in a common block for etbpsub.rat. Add reftyp to specify type of reference file. New routine to read Reynolds weekly SST data directly instead of converting it to a DSP image. Use date of data to make sure correct reference file is being used. Add input of MODIS bands 21,22,23, and 27; write out raw and bright for bands 22 and 23 to the QC file; add option for another type of reference file - Reynolds weekly data (not converted to DSP).
Use swath type file organization. Increase FOPTS -Olimit. Fix format statements to output logicals as integers. Take out references to ingest library. Fix attributes. Fix error messages. Don't allow DSP version of oi reference file. Add check for invalid input values.

MODLIB/MOCEAN: Added support for multi-file format (BSX_MF). Added support for HDF-EOS swath format (ESW_2D and ESW_MF). Use SMF routines instead of printing to stderr. Allow multiple bands to a single subordinate file.

OISST: Add option to read weekly file (formatted reals).

MSBIN: Fix debug ifdef. Get file organization from image open. Add attributes; fix close routine - add missing braces for if statement. Add long name attribute to output file; fix band number calculation. Read attributes from input file; fix max number of output bands; fix extraction of orbit number from input file; fix references to multi-dimensioned arrays. Change pcf logical numbers to be in 300,000 range for MODIS oceans; add attributes to output file; fix check for successful output file creation. Fix use of proc_con and proc_log. Take out DSP stuff. Fix attributes. Change pcf files to fix template pcf stuff, fix order of L3 output files, use 4km resolution.

MTBIN: Make things more standard for V1 delivery. Fix use of HDFLIB stuff; add rules to make writetest - Warner's test routine. Remove old and temporary files. Add option to subtract data in one file from the files

being summed, but it is not completely tested yet. Fix routine which reads input data. Create output file names for subordinate files. Fix handling of flag bits. Add attributes; fix extraction of orbit number; fix referencing of multi-dimensioned arrays. Add attributes; fix string reference (use " instead of ' to make sure it is null terminated); fix format statement to make f90 happy; fix calculation of weight (wtpix) to make f90 happy. Add attributes; don't use subordinate files. Fix use of proc_con and proc_log. Fix attributes.

MMAP: Increase number of input bands; fix record read; get file organization from image open. Fix extraction of orbit number from input file. Fix use of encode to use EOS instead of 0. parsefilestring is now in binshr. Now have two pcf files for MOD28 instead of just one. Rename pcf files; fix use of proc_con and proc_log. Fix attributes in output file.

RATF90: Add trailing null character to strings that are expected to be null terminated. Some sources don't quite match what was last checked out.

MSSTSHR5: Only use trname for VMS. Remove all unused stuff. Get rid of relative paths. Remove old, unused stuff; don't reference routines which are in other libraries. Fix error messages. Move raygetpol.inc to $DSPROOT/inc because it is also used by modsst.

IMG2HDFBIT: new program to Convert DSP image mask into a bit mask hdf file. Used for land and shallow water masks. Add attributes to output file to make ncdump pretty.

MCOLSHR8: Get rid of unused stuff. len_str is now in atmcorshr. Routines from other libraries are no longer in this library. Move get_climatology.* from mcolshr7 which we no longer need. Update prolog to make checker happy.

B.4.3 Richard Sikorski has been utilizing SCF elements to model possible errors in MODIS sensors due to uncertainties in the pre-launch characterization of the spectral sensitivity of thermal IR band sensitivity bands. RAL radiance model has been used with each generation of spectra to determine whether these filter functions allow the achievement of the level one goals of MODIS for SST

B.5 DOCUMENTATION ACTIVITIES

Modified and update all prologues for new science algorithms to satisfy requirements of the automated standards and guidelines checker utility used to verify coding standards of the version 1 code delivery. Created prologues for all subroutines and functions located in the heritage code libraries. Updated document containing all ancillary and output file descriptions and version 1 packing list. Revised flow diagrams and volume and load estimates to reflect changes for version 1.

Matchup databases:
Continued processing matchups for 1994 and 1995 pathfinder sea surface temperature. Expanded the Pathfinder matchup database to include 1984 and 1985. Continuing to collect in-situ buoy data for 1996. Began exploring new techniques to identify cloud contaminated pixels in the GAC extractions contained in the matchup data base. This technique uses decision trees to develop new tests for flagging cloud contaminated satellite data for exclusion during coefficient estimation. These new tests should allow dramatic

expansion of the matchup database while still maintaining stringent quality control of the data contained in the matchup records.

The preliminary design phase of two new matchup data bases processes was initiated. The first is a real-time matchup process for sea surface temperature. We began obtaining real-time in-situ buoy data from the Navy. This data was analyzed for quality and source of these datum for inclusion in the matchup process. Scripts were then developed to automate the quality checking and reformatting of the real time in-situ data. The initial design phase was begun to automate the extraction of real-time satellite data to be included in the "real-time" matchup process. The second database design is the ocean color matchups for use with MODIS. We are in the process of designing a querible relational database system which will contain in-situ ocean color data in both the vertical and horizontal and satellite extractions of L1 and L2 data for the time and space locations of the in-situ records. We have identified 21 in-situ measurements which should be included in this database and are presently soliciting input from MODIS team members and other potential users of this system during the Design phase.

B.6 Team Activities

C. FUTURE ACTIVITIES

C.1 Database Future Work

C.2 Client/Server Future Work

C.3 Pathfinder (P)

C.3.2 Continue algorithm tests and Pathfinder-Reynolds comparisons.

C.4 MODIS (M)
C.4.1. Delivery of prototype ocean code to MODIS team.
C.4.2. Work with team to update product algorithms.
C.4.3. Work with Hughes on processing rules/scenarios.

C.5 SeaWiFS (S)
C.5.1 Continue testing of Gordon's algorithms and its interaction with HDF ancillary routines.
C.5.2 Continue timing tests with CZCS and SeaWiFS algorithms.

D. PROBLEMS

D.1 Database Problems

None listed separately

D.2 Client/Server Problems

None listed separately

D.3 Matchup Database Problems

None listed separately

D.4 DSP Related Problems

None listed separately