MODIS SEMI-ANNUAL REPORT
- JUNE 1993 -

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, Pathfinder with a P, and Headquarters with an H. 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 (H)

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 (S)

B.1.1 Operational Testing

B.1.1.1 January Testing

In January, the command files used to load data from the optical disk to the VMS data staging disks were tested and integrated into the operation. The command files GET-SUEGAC.VMS, GET-SUEGAC.BUILD and PUT-SUEGAC.VMS check the available disk space on the staging disk, and if room is available, copies one day's worth of files from the optical disk.

Once this was in place, the system could run in more or less operational mode. That is, the processing began to proceed more smoothly, and less direct intervention was needed. Some problems that before had seemed minor now needed attention. For example, occasionally the client/server connections would fail, and one or more records would not be added for a given input file. Previously, the processing had proceeded step-by-step, so this would be caught. Now, an extra error trap had to be added to the ADDRECGAC batch job. If not all required records are added, the input file is renamed

into an error directory instead of the spool area, the process control records are altered accordingly.

As the month progressed, the processing became somewhat smoother, with fewer errors and stops. The hardware and software were put in place to back the PST files up onto DAT tape, and a new set of disks made available for that task. At this point, the creation of daily maps from the daily time-binned files and to backing up procedure is still manual. Changes are planned that will also automate these steps. The transfer rate for data to DAT was a function of the channel and the protocol. TCP transfers achieved rates of 920 KB/sec. with Ethernet and 1.5 MB/sec using FDDI. NFS transfers ran at similar rates over both channels, approximately 60-90 KB/sec.

In the early part of the month, the files were taking a day or more to process one day's worth of passes, but by the end of the month, multiple days were begin processes in one day. By the end of January, days 88265 through 88305 from NOAA-9 had been processed at the rate of 7 days/day.

B.1.1.2 February Testing

The AUTOPROC system, which had processed days 88260 through 88312 in January, processed 88180 through 88259 of NOAA-9 AVHRR GAC data. Global daily day and night files were produced, and DSP images of "all" and "good" data were made for days 88235 through 88311. A preliminary threshold was used to eliminate more cloud-contaminated data. This procedure consisted of warmest-pixel compositing of day data over a 30-day interval, then a pixel-by-pixel comparison with the daily day data, marking as bad any pixels that differs from the reference pixel by more than 2 degrees. This seemed to produce promising results, so the method will be explored further.

During the testing, the database operations became the limiting factor in processing, so the use of indexes on database tables was begun, to speed up information retrieval.

Various problems and stops were encountered and resolved. Three of the problems were:

1. Difficulties making the RPC client/server connection were diagnosed and cured.

2. Use of indices confused some procedures; records were no longer being retrieved in the correct order. 'SORT' clauses were added to the affected database queries.
3. The DSP utility CALLER stopping has not yet been entirely cured, but is being investigated.

B.1.1.3 March Testing

Days 88001 through 88070 of global GAC AVHRR data were processed. The automatic processing consisted of:

1. Entry of control records into the database, moving the input file from the VMS disk to the UNIX disk.
2. Ingest a 1200-line piece of a file.
3. Apply the RSMAS atmospheric correction and preliminary cloud masking to the piece.
4. Space bin the piece to the 9.8 km equal-area binning scheme, separating the day and night sections of the pass.
5. TIMEBIN the pieces into a daily day or night file (as appropriate).

Further processing was done manually, and consisted of a preliminary attempt at further cloud masking. The steps for this consisted of:

1. Making warmest-pixel composites of seven daily files to form weekly composites.
2. For a given week, warmest-pixel compositing the previous week (n-1), that week (n), and the following week (n+1), to create a 3-week reference file.
3. Comparison of each of the daily files with the reference file on a pixel-by-pixel basis, and marking any pixels that differed from the reference by more than two degrees C as a cloud pixel.

This procedure was performed on the day passes of days 88008 through 88063. New weekly composites were also created from these declouded day files. Reference maps (4096 X 2048, cylindrical projection) were made for three cases: all pixels, good-only pixels before declouding and good-only pixels after declouding, for visual comparison.

B.1.1.3 April Testing

In April, GAC AVHRR data from 88001 through 88050 were processed again, to test a number of changes in the automatic processing system. The major changes tested were:

1. JOB TRIGGERING

A new, more flexible method of triggering processing jobs was tested in the operational system. Previously, much of the triggering was treated in the interface as 'special cases', which were hard-coded in the db_Report process. THREE fields were added to the PROCESS_CONTROL table in the database, REC_TO_TRIGGER, TRIGGER_CLASS and CLASS_TO_TRIGGER. When one job is to trigger one other job (for example, one Ingest/Atmos. Corr job triggers a SpaceBin job for one piece), the REC_TO_TRIGGER is set to the process control record that is to be triggered. Use of the REC_TO_TRIGGER field looks like:

PROCESS_CONTROL TABLE (fragment 1)

RECORD	RECIPE	STATUS	REC_TO_TRIGGER
12345	GAC_SPACEBIN	HOLD	-1 (-1 => none to trigger)
12346	GAC_INGATCOR	EXECUTING	12345

When db_Report marks the INGATCOR job as 'FINISHED,' it also checks the REC_TO_TRIGGER. If this is >0, that PROCESS_CONTROL record is then marked as 'SUBMITTED' for processing.

PROCESS_CONTROL TABLE (fragment 2)

RECORD	RECIPE	STATUS	REC_TO_TRIGGER
12345	GAC_SPACEBIN	SUBMITTED	-1 (-1 => none to trigger)
12346	GAC_INGATCOR	FINISHED	12345

The use of the 'CLASS' triggering is a little more complicated. Consider the short example, where one file is broken into 3 pieces, then recombined. The initial database fragment would be:

PROCESS_CONTROL TABLE (fragment 3)

RECORD	MAIN_ LINK	RECIPE	STATUS	TRIGGER_CLASS	CLASS_TO_TRIGGER
10001	101	GAC_PTA	HOLD	TIMEBIN	NONE
10002	101	GAC_PTD	HOLD	TIMEBIN	NONE
10003	101	GAC_SPACEBIN	HOLD	NONE	TIMEBIN
10004	101	GAC_INGATCOR	SUBMITTED	NONE	NONE

```
10005   101  GAC_SPACEBIN
                          HOLD NONE         TIMEBIN
10006   101  GAC_INGATCOR
                          SUBMITTED NONE              NONE
10007   101  GAC_SPACEBIN
                          HOLD NONE         TIMEBIN
10008   101  GAC_INGATCOR
                          SUBMITTED NONE              NONE
```

(The MAIN_LINK refers to the input file.) Assume the INGATCOR jobs have finished, as well as two of the three spacebin jobs. The database would now show:

PROCESS_CONTROL TABLE (fragment 4)

```
MAIN_   LINK         STATUS     TRIGGER    CLASS_TO
RECORD  RECIPE                          CLASS        _TRIGGER

10001   101  GAC_PTA     HOLD TIMEBIN     NONE
10002   101  GAC_PTD     HOLD TIMEBIN     NONE
10003   101  GAC_SPACEBIN
                          FINISHED    NONE TIMEBIN
10004   101  GAC_INGATCOR
                          FINISHED    NONE NONE
10005   101  GAC_SPACEBIN
                          FINISHED    NONE TIMEBIN
10006   101  GAC_INGATCOR
                          FINISHED    NONE NONE
10007   101  GAC_SPACEBIN
                          EXECUTING NONETIMEBIN
10008   101  GAC_INGATCOR
                          FINISHED    NONE NONE
```

When db_Report then marks the third SPACEBIN job as 'FINISHED,' and sees that there are no single records to trigger, it checks for any 'SUBMITTED' or 'EXECUTING' records. If there are none, the CLASS_TO_TRIGGER is retrieved, in this case, the TIMEBIN class. Then, all records with this TRIGGER_CLASS are marked 'SUBMITTED' and ready to be processed, in this case, the GAC_PTA GAC_PTD jobs. The database would then show:

PROCESS_CONTROL TABLE (fragment 5)

MAIN_RECORD	LINK RECIPE	STATUS	TRIGGER CLASS	CLASS_TO _TRIGGER
10001	101 GAC_PTA	SUBMITTED	TIMEBIN	NONE
10002	101 GAC_PTD	SUBMITTED	TIMEBIN	NONE
10003	101 GAC_SPACEBIN	FINISHED	NONE	TIMEBIN
10004	101 GAC_INGATCOR	FINISHED	NONE	NONE
10005	101 GAC_SPACEBIN	FINISHED	NONE	TIMEBIN
10006	101 GAC_INGATCOR	FINISHED	NONE	NONE
10007	101 GAC_SPACEBIN	FINISHED	NONE	TIMEBIN
10008	101 GAC_INGATCOR	FINISHED	NONE	NONE

2. DATA-DAY DEFINITION

In conjunction with the DATA-DAY definition that has been developed for the global AVHRR GAC processing, a new table has been added to the database, DAYSTART, and has been integrated into the automatic processing. The table contains the beginning and ending date/time for each yearday. The information is retrieved for the new version of the SPACEBIN program (covered in a different section).

3. NEW SST ALGORITHM

A second set of processing recipes was developed to test a new, non-linear correction, for comparison with the previous calculation. Adjustments were made to programs and command files to more easily define the type of processing desired. These changes were numerous, but generally minor in nature, so will not be listed here. As with the data-day definition, the specifics of the new algorithm will be covered elsewhere.

During the operational testing, a number of minor problems were encountered. Two examples of these are:

1. While testing the new SPACEBIN program, program errors occurred, but were masked by the inability of the UNIX operating system to create an error 'core' file, due to lack of

disk space on a particular disk. All command files were then changed to allow any core files to be written to one of the large, NFS usr disks.

2. The db record-adding program refused to process three input files, claiming that channel information was not present in the header record. It turned out that in the channel section of the header, the letter 'o' was mistakenly present where a zero was required. The regular TIROS ingester, as well as the record-adder were modified to account for this occasional problem.

B.1.1.5 May Testing

A new, non-linear SST calculation was implemented in May, and days 88001 - 88071 were re-run with the new algorithm. In addition, a new strategy for spacebinning the data was used, that is intended to split a pass at the beginning/end of a data-day. This data-splitting is not yet correct, and is being worked on. Further modifications on both the SST algorithm and the spacebin are planned.

B.1.1.3 June Testing

In June, the automatic processing system was tested on four DEC ALPHA computers (Orange, Enuka, Kelso and Lychee) and the multiprocessor Silicon Graphics machine (Modis), using global GAC AVHRR data. Three Alphas (Enuka, Kelso and Lychee) are used for processing the orbital input files through sectorization, ingest, atmospheric correction, space binning and timebinning, producing two SST files for each orbit (ascending and descending data). The SGI machine is used to timebin orbits together to create daily files, and the fourth ALPHA will be used for additional cloud masking and data backup. Many test runs were made throughout the month to adjust to the new, multi-machine environment. The major adjustments needed in the system will be covered in the development section.

The present job flow is:

Each alpha:
init copies input file to local disk
ingest/atmospheric correction on pieces
1 asc & 1 dsc timebin puts orbit back together
clean job deletes intermediate files, triggers copy1 and next init
Then, on MODIS:
copy1 job copies from alphas to MODIS

When all input files from one day have been processed, a 'daily1' job for TWO DAYS AGO is triggered (to be sure all has been done) daily1 timebins the asc and dsc orbit files into asc and dsc daily files, triggers the 'copy2' job

Finally, on orange (this is next to modify)
copy2 copies the daily pst's from MODIS to orange, release the weekly1 job
The weekly1 job will composite the daily files into weekly files.
The weekly2 job will create the reference file.

B.1.2 Development

B.1.2.1 January Development

In late January, the executable versions of the AUTOPROC software were installed and tested on a VMS and UNIX computer at Jet Propulsion Laboratory, and the staff there is currently learning how to operate the system.

Work was begun on a generic template client program, which will allow the ease addition of service functions to the AUTOPROC system. A number of these tools already exist, such as programs to recover from a computer crash, to reset the database after an error in processing, and to move records between tables, but these currently work outside the scope of the client/server. When they are used, the work request and record addition programs must suspend operation, to avoid lock conflict in the database. The template client is nearly ready.

The entry procedures, work request and triggering will be revised to allow for the use of multiple computers.

B.1.2.2 February Development

Testing of new programs or procedures has become more challenging. Given the existing environment, operational processing must be halted, the new system installed and tested, and the operational system reinstalled. To expedite the process, a second computer was made available for testing of the development copy of the AUTOPROC system. This 'APTEST' database system was installed

and tested, to insure that it did not interfere with the operational system.

Additional procedures are currently being developed that will facilitate processing at sites with multiple machines, and to automatically map and backup output from the processing.

B.1.2.3 March Development

A new job triggering scheme is being developed that will increase the flexibility to enhance existing or develop new processing threads. In addition to triggering single jobs, the capability has been added to define and trigger classes of jobs. For example, in the old method, completion of all space bin jobs for a particular pass would trigger the release of ALL jobs marked with a process-status of "HOLD," so only the TIMEBIN jobs could be marked "HOLD," and any follow-on jobs (CLEAN for the GAC processing) had to be marked differently, and triggered in a separate manner. To be able to trigger these jobs, specific sections in the database interface were used. In the new scheme, each "class" of jobs is assigned a "trigger_class," as well as a "class_to_trigger". Thus, the spacebin jobs have a "class_to_trigger" of "TIMEBIN," that refers to both the day (GAC_PTD) and night (GAC_PTN) TIMEBIN jobs.

Other additions were made to the database interface, but the change in triggering was the most significant. These changes were implemented in the test database, but have not yet been installed in the operational system.

B.1.2.3 April Development

Data Overlap
When the satellite switches data recorders (once per pass), there is usually some overlap between the data received, as one recorder is turned on before the other is stopped. This duplicate data will be eliminated. A new field has been added to the MAIN database record, giving the PASS_END of that file. The record-adder will search for this overlap, and adjust the beginning scan line of the data to eliminate the overlap.

New Initial Procedures
Previously, files were entered into the system one-by-one. A stand-alone VMS command procedure polled the UNIX staging

disk, and when room was present, one file was moved into the GAC_ADDREC directory. The batch job ADDRECGAC checked this directory. When a file was present, the records were added to the database. The file was then copied to the UNIX staging disk, and the INGATCOR records were marked as 'SUBMITTED' for processing. While this method was working for a single UNIX processing computer, it needed revision allow for the use of multiple processing computers. A new initial recipe, GAC_INIT is being developed that will allow multiple computers to use and process in parallel.

New QC Procedures

The current automatic processing is completed when the daily ASC and DSC PST files have been produced. We have been experimenting with first-cut ways to improve the cloud detection, using a combination of warmest-pixel compositing and thresholding to eliminate bad data. Currently, these methods are being performed by hand after the automatic processing has completed for a particular day.

Processing recipes are begin developed that will add these QC procedures to the automatic processing stream, as well as other actions, such as automatic mapping and data extraction.

B.1.2.5 May Development

Changes were made in both the SST calculation and the space bin programs.

Overlap elimination

The database and its interface were modified to store information about the overlap between passes. When the satellite switches from one recorder to another, there is usually an overlap, generally about five minutes long, when both recorders record the same data. The database now stores the time of the last scan for a pass, in addition to the first time. When directed to do so, the automatic processing system automatically retrieves the last time of the previous pass, and starts the ingestion at this time, instead of the beginning of the pass.

MCP Modifications

During May, some the DECStation 5000's were upgraded with DEC ALPHA processors. The host processing program, MCP was

modified to run on these ALPHAS. Many changes were made to the command files used to start, stop, and monitor the processing, to accommodate this expansion of capabilities. Such changes as the naming of log files, definition of local processing directories using UNIX environmental variables, modification of DSP procedure files used in the processing, and restriction of certain sections of the job stream to single computers were also made, and other changes planned. The job stream itself has been modified, and other changes are currently being made. Currently, the job stream consists of:

Old Stream	New Stream	Status
	NLC_INIT	Complete
GAC_INGATCOR	NLC_INGATCOR	Complete
GAC_SPACEBIN	NLC_SPACEBIN	Complete
GAC_PTD	NLC_PTA	Complete
GAC_PTN	NLC_PTD	Complete
GAC_CLEAN	NLC_CLEAN	Complete
	NLC_COPY1	Planned
	NLC_DTA	Planned
	NLC_DTD	Planned
	NLC_COPY2	In progress
	NLC_COMP1	Planned
	NLC_FLTR1	Planned
	NLC_REF	Planned
	NLC_DCLD	Planned
	NLC_COMP2	Planned
	NLC_MAP_CLEAN2-BCKUP	Planned

These processing recipes perform the following tasks:

NLC_INIT: (Submitted when all record have been added to the database.) This recipe copies the input file to the local host that will perform the processing, restrict processing through

'NLC_COPY1' to that computer, and marks all NLC_INGATCOR records associated with the input file as SUBMITTED for processing.

NLC_INGATCOR: Ingests the piece, completes the navigation information, and performs the SST atmospheric correction. There is one record for each piece in the pass, and the completion of each piece triggers the NLC_SPACEBIN record for that piece.

NLC_SPACEBIN: Converts the satellite-perspective SST file to an equal-area, 9-km bin file (or files). The ascending and descending data are placed in separate files, and if the data is "near" the beginning/end of a data-day, it also creates a separate file, splitting the data according to the day in which it should be placed. When all spacebins for a given pass have been completed, the first timebins jobs (NLC_PTA & NLC_PTD) ARE SUBMITTED FOR PROCESSING.

NLC_PTA & NLC_PTD: These jobs timebin the pieces from the spacebins into orbit into single files for ascending and descending data. (A third file may be produced if the orbit is near a data-day end.) Completion of both of these jobs triggers the CLEAN1 job.

NLC_CLEAN: Deletes intermediate and input files, then triggers the COPY1 job.

NLC_COPY1: This job runs only on the computer designated for the daily timebins. It copies the files from the host that performed the previous jobs, and deletes the local host's files. When all files from a particular day have been processed, the next timebins for two days before are triggered. (The two day-lag is designed to assure that all files are present before further processing.)

NLC_DTA & NLC_DTD: This recipe will timebin all ascending and descending files for a day.

NLC_COPY2: This job copies the daily timebinned files to the computer that will perform the QC and backup procedures. Previously, these functions were performed by hand, but

command files are being designed to automatically composite, filter, cloud-mask and backup the daily files.

NLC_COMP1: This recipe will create weekly warmest-pixel composites.

NLC_FLTR1: This job will filter the weekly files.

NLC_REF: This job will use the weekly filtered files to produce a weekly reference file.

NLC_DCLD: This recipe will threshold the daily files against the weekly reference, marking pixels as 'probably clouds' those which differ from the reference by a particular value.

NLC_COMP2: This job will produce a new set of weekly composite files.

NLC_MAP_CLEAN2-BCKUP: This recipe will produce any DSP image maps requested, any other products from the daily files, and back up the daily, weekly, and reference files onto DAT tape.

B.1.2.3 June Development

Due to the need to define local directories for each machine, two new subroutines were written to extract information from the database, and new tables were defined. A new table was added to the database, LOCALDIRS, and the new subroutine get_localdirs is used to define the output directories for each individual computer.

Some new command files were written to monitor the processing on the ALPHAS, and more are planned to allow startup, selective shutdown and more monitoring from a single machine, eliminating the need to log in to each machine.

An error in the day start/yearday definition was fixed.

Operational testing on the ALPHAS exposed some bottlenecks in the processing system. The response time of the client/server and of the database itself is becoming the limiting factor in the processing. Steps were taken, and more are planned, to increase the throughput.

To reduce the number of clients requesting service from the database, the size of each piece to be processed was increased form 1200 scan lines to 2400 lines, reducing the number of pieces for each orbit from 10-11 to 5-6. As two records are entered for each piece (the ingest/atmospheric correction step and the spacebin step), this will significantly reduce the number of clients. Also, The "INIT" job (which copies the input file from the VMS staging disk to the local UNIX disk) was changed to run a single time, and be started by the 'CLEAN' job, which removes an input file that has been processed. More changes such as this are planned.

This change involved modification of the MCP program, which previously ran continuously. Now, for selected process classes, it will run only once, then exit.

An analysis of the type and number of queries was made, to begin to identify places where more efficient strategies can be used. For example, one very common query paired a search for a particular satellite and sensor with one or two other field searches. These queries were changed by searching on a new field, SATSEN_CODE, replacing a complicated character-string search on at lease two fields with a single integer search. Further efficiency was introduced by bringing the SATSEN table into memory upon server initialization, eliminating the repeated queries on that table. More changes such as this are also planned.

We are also starting to experiment with the use of indexes on the database table to improve response.

When problems occur in the processing, such as a CALLER crash, the single disk that had held the ALPHAS log files might fill. To avoid this, the ALPHA log files were located in each computer's local disk.

B.2 Client/Server Status (S)

B.2.1 JANUARY

VDC was customized for our environment. VDC can now run on our computer network, including VMS and mutliple vendor Unix machines. More than 9800 RTE runs were completed.

B.2.2 FEBRUARY

VDC modification has been completed; more than 25000 runs have been made in three batches to process Gordon/Wang's rte data. Results of these runs have been provided to Gordon and Wang. The vdc process is being extended to CZCS data for testing purposes.

Modifications are being made to the space binner and time binner to accommodate Janet Campbell's request for the statistics.
Preliminary changes have been made and testing begun to gather comparison data.

B.2.3 MARCH
Client/Server development focused on three areas during this period.

1. VDC was customized for our environment. VDC can now run on our computer network, including VMS and multiple vendor Unix machines.

2. With VDC modification completed; more than 25000 runs were made in three batches to process Gordon/Wang's rte data. Results of these runs have been provided to Gordon and Wang. The VDC process is being extended to CZCS data for testing purposes.

3. Modifications are being made to the CZCS SPACE BINNER, TIME BINNER and MOSAIC to accommodate Janet Campbell's request for the statistics and will provided the basis for SeaWIFS processing. Preliminary changes have been made and testing begun to gather comparison data; the changes included:

    1. Modification of gsfcbin9.mice to obtain czcsbin.mice as the CZCS space binner.
    2. modified spacetime.mice and mosaic9.mice to obtain CZCS time binner.
    3. modified mosaic9.mice to accommodate the space binner/time binner modifications.

B.2.4 APRIL
Client/server efforts have focused on the modifications CZCS space binner and time binner to accommodate the new statistical routines proposed by Project. Further modifications included similar changes to the mosaic and map programs.

These changes were made to the SeaWIFS processing routines also.

Preliminary testing was begun to compare the results of the existing and modified processing schemes. A working version of the new processing regime is anticipated to be available in the May time frame.

B.2.5 MAY

Client/server efforts have continued on the modifications to both CZCS and SeaWiFS space binners and time binners to accommodate the new statistical routines proposed by Project. The space binners are based on the PATHBIN.program which itself has been evolving during this period. as a resuls, multiple versions of these binners have been created to address or include changes made to PATHBIN. Further modifications reflect similar changes to the MOSAIC and MAP programs.

Preliminary testing was begun to compare the results of the existing and modified processing schemes. A working version of the new processing regime is anticipated to be available in the June time frame due to the multiple changes included in the programs.

B.2.4 June

Client/Server development focused on three areas during this period.

1. VDC was customized for our environment. VDC can now run on our computer network, including VMS and multiple vendor Unix machines.
2. With VDC modification completed; more than 35000 runs were made in four batches to process Gordon/Wang's rte data. Results of these runs have been provided to Gordon and Wang. The VDC process is being extended to CZCS data for testing purposes.
3. Modifications are being made to obtain the CZCS/SeaWIFS SPACE BINNER, TIME BINNER, MOSAIC and MAP to accommodate Janet Campbell's request for the statistics. Preliminary changes have been made and tested.

The changes included:

1. Modification of pathbin.mice to obtain czcsbin.mice as the CZCS/SeaWIFS space binner.
2. modified spacetime.mice/pathtime.mice and mosaic9.mice to

obtain CZCS/SeaWIFS time binner.
3. modified mosaic9.mice to accommodate the space binner/time binner modifications.
4. modified map9.mice to accommodate the space binner/time binner modifications.

B.3 Matchup Database (P)

B.3.1 Definition of a "data-day"

The basic products to be 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. 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? During this period we examined this question.

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 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 a repeat cycle of a few days.

When daily global satellite data fields are to be constructed, one needs to adopt a consistent definition of a data-day. This definition should be easy to implement in practice and should minimize temporal aliasing in the resulting products. During the past month,

we have explored some of the alternative definitions; a brief document discussing our findings is in preparation.

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. The 24-hour definition of a data-day, however, 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. The effects of temporal aliasing are further complicated by the fact that the locations where aliasing may occur change in time.

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. We selected the 180° meridian as the boundary between data-days. This seems to be a better alternative. There is some potential for temporal aliasing, but they can be mitigated by substituting segments of tracks sampled almost 24 hours apart with corresponding segments from either the previous or following data-day.

The various data-day definitions will be illustrated more thoroughly in a document currently under preparation, which may be converted into a paper for submission to a scientific journal. The preferred definition of the data-day was implemented and incorporated to the binning/mosaicing programs.

B.3.2 SSM/I Global Atmospheric Water Vapor Fields

The potential incorporation to the matchup database of data from the Special Sensor Microwave/Imager (SSM/I), a microwave radiometer flying on the DMSP spacecraft was examined. SSM/I geophysical

data tapes produced by F. Wentz were obtained from JPL's PO-DAAC and transcribed to optical disk, to facilitate future access. Software was developed to read the SSM/I data files and extract values for a given time/space range. Testing revealed that the initial version of the software had an error that resulted in only a small proportion of the data being retrieved. These errors were corrected and we began producing SSM/I global water vapor fields with a double purpose: (a) to extract water vapor values at the times/locations of AVHRR and in situ matchups, and (b) to support future quality control of Pathfinder SST fields.

The first step in producing the global water vapor fields was to extract the valid water vapor values from the daily SSM/I files that had been archived in optical disk at Miami. The water vapor values were then subsampled into a global 0.5° grid, using a block-median filter. The daily SSM/I files do not cover the oceans completely, as there are gaps in the sampling. To fill these gaps, global water vapor fields for a given day were constructed using information from that day, as well as the previous and following days. Data from adjacent days were weighted using a Gaussian filter. Daily SSM/I global water vapor fields were produced for the first six months of 1988.

The next step was to develop procedures to extract SSM/I water vapor values at the times/locations of existing matchups. This step is very time-consuming, so we tested using the smaller East Coast HRPT matchup database. The evaluation of the usefulness of water vapor information is currently in progress.

B.3.3 Additions and Modifications to Matchup Databases

During this period, a series of additional fields was incorporated into the matchup databases. Examples of the additions are individual baseplate temperatures (previously only the average of the four measurements was included). Also, brightness temperatures for channels 3, 4 and 5 corrected for emmissivity changes dependent on scan angle were added. Finally, a series of changes was introduced to the calibration procedures, which influenced the brightness temperature values included in the databases. The current version of the matchup databases (Version 17) incorporates the consensus Pathfinder nonlinearity correction (documented in a paper to appear shortly).

B.3.4 Additional In Situ Data

During this period we made efforts to update the available in situ SST observations. Contact was made with the Japanese Meteorological Agency, which kindly supplied data from moored buoys around Japan for 1991-1992. NOAA's Atlantic Oceanographic and Meteorological Laboratory provided drifting buoy data up to October 1992. Because the impending move of the NODC might have delayed considerably the delivery of in situ SST data from the NDBC moored buoys, special arrangements were made with NODC to retrieve buoy data over the network.

B.3.5 Matchup Database Release (P)

The first public release of Pathfinder matchup data bases took place during this period. The released matchup data bases (Version 17 in our version control system) were delivered to the Physical Oceanography DAAC at NASA's JPL. They were also sent to investigators at NOAA/NESDIS (C. Walton). We released two matchup data bases: (a) global matchups for the NOAA-9 spacecraft, January through November 1988, and (b) the "experimental" matchup data base including data only for moored buoys off the US east coast and Gulf of Mexico and NOAA-9 AVHRR HRPT data from the University of Miami archive, January 1985 to November 1988. The data were filtered to eliminate matchups contaminated by clouds; this step reduced significantly the volume of the matchup data sets. The criteria used in data filtering was described in documentation accompanying the released matchup data bases.

B.3.6 Current Matchup Database Status and Planned Activities

We are currently extracting NOAA-9 data from the GAC disks for 1987. Once the extractions are completed, we will build the global database for that year. Next, we will begin extractions from the HRPT archive for 1989,1990 and 1991. This will give us a three-year East Coast database for NOAA-11, which will be used to derived proto-algorithms for that satellite. We will then continue extractions from the global data.

B.3.7 SST Algorithm Development

We are exploring empirical SST algorithms based on statistical modeling of the association between in situ SST and various

independent variables (for instance, Channel 4 brightness temperature (BT), difference between channels 4 and 5 BTs, satellite scan angle, etc.). The algorithm is estimated through regression procedures. Regression estimates require that the independent variables be linearly related to the independent variable, although that is not usually the case with the matchups. We experimented with modern regression techniques that find empirical transformations required to linearize the association. These techniques were the ACE and AVAS algorithms included in the S-PLUS statistical package. The empirical transformations were used in an exploratory approach, to gain insight on the nature of the non-linearities in each of the independent and dependent variables. This effort is presently ongoing, so no conclusions are available yet.

B.4 DSP Support (H)

B.4.1 Modifications/Additions to DSP:

B.4.1.1 January

Misc. changes for OpenVMS.
Add definition of workspace variable _PATH_SEPARATOR in DSP.
Add Dsp_DeleteFile function to the imageio library.
We changed the way we handle character strings in all Fortran source files.

B.4.1.2 February

Modifications/Additions to DSP:
Add new DSP function "getenv".
Add 'run away' error checking to catch missing ()'s in RATFOR.
Add RECLEN utility.
More library modifications for Sparc with gcc.
Add 'mr.clean' make option.

B.4.1.3 March

Complete new DSP function "getenv".
Complete 'run away' error checking to catch missing ()'s in RATFOR.
Add RECLEN utility.
More library modifications for Sparc with gcc.
Add 'mr.clean' make option.
Add 9k bin bit mask for land masking to PATHBIN.
New programs to convert between image mask and 9k bin bit mask.
New program, PATHMASK, to mask bins from a pst file.
New program, PATHCLOUD, to take cloudy bins out of a pst file.

New program, PATHCOMP, to composite (ala COMPOS) two pst files.
New version of atmospheric correction: PATHNLC.
Add quality checking to PATHBIN to handle the nlc mask bits.
Add day splitting of pieces in PATHBIN.

B.4.1.4 April:

Add handling of Univ. of Tokyo format to PATHSCRIPP ingester.
Allow a subsampling factor larger than 10 in CONVRT.
Add option to DIFF to output zero if either input is outside the min or max values instead of bounding the inputs before subtracting.
Modify MIA2TIFF to not require navigation information in the input image.
Add data day splitting to PATHBIN.

B.4.1.5 May

Modify 9k bin transform routines to allow two simultaneous bin sets.
Add ability to have multiple bands of different pixel sizes and types.
Add URI program REPLACE which replaces a pixel with the specified replacement value if the pixel is less than, or greater than or equal to the threshold value.
Modify DBMAN to not do the database accesses if workspace symbol DBMAN__IGNORE is defined.
Add new TIME_BIN associated data blocks for SeaWifs PST files.
Add new ALOG calibration equation.
Use shorter title strings for the xfbd image plane windows.

B.4.1.6 June

Add new Seaspace format to SCRIPP.
Use ENDIAN_BIG instead of BIG_ENDIAN.
New program 9KLM2IMG to convert a land mask bit array to an image file.
New program CZCSMAP9 to mosaic/remap a czcs 9k pst band.
New program DMPCNT to dump pixel values to an ASCII file.
New program GSFCBIN9 to bin czcs data into 9km bins.

B.4.2 Problems fixed:

B.4.2.1 January:

Misc. fixes to the utility programs.
Fix makefile's to properly handle files in other directories.
Fix error handling in FILEEPS.
Rewrite EXIST in C so that it will work on all systems.
Fix HIST to declare common properly (with the full size).

Fix problem with PATHSCRIPP re NOTIFY_VALID_PRT.
Misc. fixes to makefiles, mostly for VMS.
Properly handle asynchronous reads on Unix.

B.4.2.2 February

Final string fixes in TROUTC.
Fix string handling in the ingest programs.
Fix the makefiles for some of the C programs.
Some makefiles needed a special case for Sun's and gcc.
Fix makefiles for utilities and libraries for VMS.
Misc. changes to MAKE-BSD utility.
Misc. fixes to DSP (fix handling of user directories).
Fix handling of multi-page menus.
"Fix" SLD CHAR command. On SGI, DECstation, and Unix Alpha the user must put the character string in double quotes (which are not printed).
Sun's work with or without the quotes (if present they are not printed).
For VAXstations the string should not be surrounded by quotes (or they will be printed).
Fix string handling in PATHSST.
PATHBIN: Change land mask to a dsp image file; fix string handling. Fix declaration of image pointers in PATHTIME.
PATHMAP: Change subroutine so other programs can use it; fix pixel bounds; fix int/real type conversions.

B.4.2.3 March

More string handling fixes.
Modify 9 km bin mask handling routines so they can be called by programs other than PATHBIN.
Fix EXIST for VMS.
Fix SCRIPP for big endian machines (SGI, SUN).

B.4.2.4 April:

Fix string handling and big endian handling in PATHSCRIPP ingester and ingest library routines.
Fix string handling in PATHTIROS ingester.
Fix CZCS ingester, FB library, XFBD-MOTIF, and MAKE-BSD for OSF/1.
Fix handling of _STATUS workspace variable.
Fix PALSAV for big endian machines.
Change EXAMIN to nicely line up the band number column.
Take out "inquire error" message from EXIST.
Fix string handling to get epsilon_* workspace symbols in COLORSHR.
More string handling changes for 2CHAN.

Fix PATHMAP to not calibrate the quality or mask bands.
Fix PATHTIME to allow an input file with just one line.
More string handling fixes to SLD.

B.4.2.5 May

Fix REMAP corner calculation; change REMAPGRID to also work for REMAPT so we don't need REMAPGRDT anymore.
Fix SLICE to clear the graphic bit if the pixel is outside the range.
Some fixes to NETCDF and CALLER for Sun.
Misc. fixes to XFBD, WRKTLK, and CALLER.
Fix WRKTLK to make sure that GROUP can write to the lock file.
Change CALLER to delete the global sections on exit.
Declare scopy and smove externals.
Change library names to all start with "lib".
Add prototypes for some VMS functions to the FB library.
Fix RTLIB for VMS.
Fix ascending/descending flagging in PATHSST.
Fix error reporting in CDF routine.
Fix PATHBIN's gregorian to julian calculation and PI value; change binning algorithm to map pixels to all apprpriate bins instead of just one; add ability to sum only one band; fix data day splitting at 180; fix ascend/descend splitting; more string handling changes.
Enlarge the XFBD dialog box for 100dpi fonts.
Allow PATHMAP input pst files to have any number of bands; only use quality and mask bands if they are in the input image.
Fix SOLVELPS test for presence of DSP_DIM.
Change argument type of fchunk on SGI's in dsplib.h include file.
Modify IDIR.H include file: subimageheader is now dynamically allocated.
Modify IMGSTDIO.H include file: add caching structure.
Modify NETCDF.H include file: netCDF is no longer distributed.
Modify TYPES.H include file: remove extraneous typedefs.
Modify NEWNAMES include files: add entry for Dsp_PackBand.
Fix XFBD to set curimg in DISPLYCOM.
Fix TRAVEC to read the pmv data lines properly.
Add calibration degradation for channels one and two to NDVI and SCALE8.
Merge PATHTIROS into TIROS and PATHSCRIPP into SCRIPP.

B.4.2.6 June

Fix SCRIPP to check for EOF on Sun and fix byte swapping for Dundee data.
Fix error message in TIROS.

Use "time_t" for type of time variable in VMSFORLIB routine IDATE.
Make the VMSFORLIB routine GETADR output pointer (void *).
Add pointer variable type, PTR, to MIAMI.RAT include file.
Modify RAYGETPOL in COLORSHR5 and in AVHRRSHR5 to define FAST_LOAD to allow system to select either BINARY (if present) or ASCII file format; and fix string handling.
Fix ingesters to use new library names.
Fix VHRR library routines for Alpha 64 bit pointers.
Fix VHRR TIME command decode; declare arrays PTR if they store addresses.
Fix MICE for VMS.
Fix signal handlers type (void) for OSF in ingest library routine RTAPE_LIB.
Fix DSP module to call signal handlers with one argument.
Fix PATHBIN to force all type conversions; fix ability to sum only one band; change 180 deg check again; increment size of AZSIZE; use AZSIZE instead of BSIZE for BOXCODE array size.
Fix SHARP to agree with the new SCRIPP ingester.
Misc fixes to IO library routines (for subimageheaders that are bigger than 512 bytes; flush the scan line cache on closer; add support for cache logic; add support for different data types in different bands).
Change Dsp_CreateOutputImage argument PIXSIZE to be an array.
Fix TRACE: string handling, reads, and decodes; use image size instead of plane size.
Add option to PATHNLC to output all 18 bands or just three: sst, mask1, mask2.
Add "nopad" option to SAVENOHED.
Misc. changes to xfbd.
Change HDF library for alpha/VMS.
Change the INGEST library so it works on Alpha's.
Change PATHTIME to allow any number of summed bands.
Fix an error message in 2CHAN.
Fix an internal read in DBMAN.
Change PATHMAP to use either the original pixel quality (from pathbin) or any new ones (from pathcloud); change BIN_NUMBER output band to 32 bits.
Fix CALLER ATEXIT function for Sun.
Change variable PTR to PNT in SKTLND.
Fix use of LOGICAL command line variables in a few programs.
Fix F2C problem with functions returning a logical value.

B.4.3 Testing:

B.4.3.1 January

None listed separately

B.4.3.2 February

None listed separately

B.4.3.3 March

None listed separately

B.4.3.4 April

Test 'mr.clean' make option.

Test 9k bin bit mask for land masking to PATHBIN.

Test new programs to convert between image mask and 9k bin bit mask.

Test new program, PATHMASK, to mask bins from a PST file.

Test new program, PATHCLOUD, to take cloudy bins out of a PST file.

B.4.3.5 May

None listed separately

B.4.3.6 June

None listed separately

B.5 Direct Project Support

B.5.1 SeaWIFS **(S)**

B.5.1.1 January

i. Bob Evans attended Seawifs meeting at Annapolis, presented SeaWIFS level 2 and 3 processing programs, equal area grid and data day definitions. (Jan. 19-22)

ii. Seawifs discussion with G. Feldman, C. McClain and Cal/Val group refining plans for level 2 and level 3 SeaWIFS Programs. Specifically, adding J. Campbell's MLE statistics to the binning programs, adding Cal/Val HDF routines for HDF file i/o, adding HDF access to the ancillary data fields (relative humidity, atmospheric pressure, wind speed, and ozone). Present ANLY7 routine implementing an early form of pixel level epsilon calculation and Rayleigh-aerosol interaction will be incrementally upgraded to current generation SeaWIFS algorithms and HDF i/o. SeaWIFS test processing has been shifted to use ANLY7 to more closely represent SeaWIFS level 2 machine loading. Dummy channels will be added to the present DSP output files to represent I/O resource requirements. A sequence of related changes then will introduce in the EDGEMASK, SPACEBIN and TIMEBIN programs.

The MOSAIC program will be replaced by a mapping program supporting the various projections offered by the present REMAP program. (Feb. 3)

iii. Have initiated discussions to define changes in the BIN programs and database control programs (VDC and MCP control databases) to support the data day definitions within SeaWIFS.

B.5.1.2 February

a. SeaWIFS: Radiative Transfer Equation Processing

i. Original Level of Effort

Original 10000 runs completed for aerosol

Additional humidity level added, 3000+ runs

ii. Overall Effort

19,000 out of 26,000 runs completed

Progressing at 600 to 800 runs a day

Completion expected within next two weeks

Results to be transferred to Gordon's group to incorporate into Rayleigh-aerosol routines.

b. SeaWIFS: Modifications and Integration

i. Adding J. Campbell's MLE Binning Statistics into Level 3 Programs

ii. Coordinating Developments with C. McClain

HDF I/O Routines

HDF Ancillary Data Field Routine

iii. Level 2 Conversions

Routines to work with CZCS and SeaWIFS Data will use Gordon's new Atmospheric Correction

iv. Integration in March/April Time Frame

v. Testing Beginning in May

B.5.1.3 March

i. Dalu Li testing new binning statistics for Space and Time Binning programs. The operations for CZCS channels 1-4 have been replicated to simulate the computational and data loading for the Seawifs data stream.

ii. The spatial data day coding will be transferred to the Seawifs binning programs. The ANLY program will be modified to incorporate the Gordon's new method for atmospheric correction.

iii. Rayleigh-aerosol code will be integrated into Seawifs ANLY Level 2 program in late April, early May following J Brown's return from cruise.

iv. Dalu Li implementing and testing J. Campbell's Level 3 binning statistics. Coordination meetings have been held with Cal/Val group to clarify questions relating to the binning scheme. The results of these discussions were compiled by M. Darzi and transferred to D. Li for coding.

B.5.1.4 April

i. Robert Evans attended SeaWIFS algorithm meeting, discussed sensor problems and possible remedial activity. He talked with processing and Cal/Val groups concerning navigation, ancillary and L-1b routines. Worked with H. Gordon to define routines necessary for SeaWIFS atmospheric correction.

ii. Routines will be delivered to J. Brown for incorporation into Seawifs ANLY program. Work progresses in testing L-3 binning routines for Seawifs.

iii. . Testing of SeaWIFS L-2 to L-3 routines will be performed using CZCS data for input. The CZCS/SeaWIFS routines will be extended to include additional SeaWIFS channels once the CZCS testing is completed.

iv. University issued PO to Sybase for data base software.

v. Ordered QStarr MastarMind jukebox staging software for magneto-optical jukebox. This permits use of the jukebox as a 30 gigabyte file server.

B.5.1.5 May

i. Jim Brown worked with both cal/val and data processing personnel. They exchanged information to implement HDF ancillary data routines for SeaWiFS/CZCS atmospheric corrections.

ii. Jim transferred several algorithms to Miami.

iii. Jim Brown has implemented Gordon's SeaWiFS algorithms in a trail program to process CZCS input files.

iv. Jim Brown has worked with Miami's in-situ personnel to further develop data collection programs.

B.5.1.6 June

i. Jim Brown is currently working at Goddard directly with his SeaWiFS counterparts.

B.5.2 MODIS **(M)**

B.5.2.1 January

i. Order DEC 3000/400 workstations to enhance the computational capability at RSMAS
ii. Ordered SGI server upgrade 4D/480 to Challenge
iii. Implemented Howard Gordon's RTE calculation on multiple workstations.

B.5.2.2 February

i. Implemented NOAA Non-Linearity Algorithm
ii. Extracted First Three Months of 1988
iii. Investigating Use of SSM/I Data to Improve Cloud Detection
iv. Preparing New Matchup Database
Will contain first three months of 1988
Available for distribution by end of April
v. Processing Remainder of NOAA-9 for 1988
vi. Extraction Rate
5 to 6 days per day
Expected completion within two months

B.5.2.3 March

i. Bob Evans attended MODIS science team meeting and discussed the future work algorithm theoretical basis document (July, 93)
ii. Worked with DEC to try and accelerate delivery of 3000/400 workstations; delivery now expected in early April.
iii. Designed and ordered expansion disk configurations to support MODIS development and processing. Delivery expected in early April.
iv. Angel Li and Jim Brown participated in set up for D. Clark's optical cruise, Jim remained to participate in cruise.
v. Angel Li traveled to GSFC for one week to work with Seawifs personnel to implement HDF capability within the DSP I/O structure.
vi. Completed work on 26,000 RTE simulations to obtain Rayleigh-aerosol interaction coefficients.
vii. Generated prototype algorithm timing tests, results have been sent to P. Ardnoy. Product sizing estimate; results sent to R. Chin

B.5.2.4 April

i. Robert Evans talked with S. Hooker and D. Clark concerning in-situ data base and in-situ data activities.
ii. We received upgrades to DECstations, all DECstations have been replaced by DEC alpha platforms.
iii. We received and installed new magnetic disk space to accommodate Pathfinder and in-situ data base activities.

iv. All Silicon Graphics and DEC computers are now interfaced using FDDI. Recent driver improvements have increased cache file transfers between Alpha machines to 3-7.5 megabytes/sec. NFS transfers run at approximately half this speed.

B.5.2.5 May

i. Information from team meeetings indicates that NASA will provide data to the National Archive three years after acquisition.

B.5.2.6 June

Bob Evans attended a meeting of the MODIS data producers. A simple model of MODIS processing represented as a single large makefile was suggested. A number of the instrument representatives objected to this simplistic approach because it makes management and updating the algorithms tedious and is inconsistent with the concepts of structured programming.

Peter Evans attended a meeting of MODIS data users. Among the topics discussed was the creation of a registry of users who would want to be notified of changes in the databases, processing algorithms, cal/val, etc.

B.5.3 Pathfinder **(P)**

B.5.3.1 January

a. Meetings

i. Bob Evans traveled to JPL to discuss transfer of Pathfinder processing from Miami to JPL. (E. Smith, P. Cornillon, A. Tran, B. Evans). Discussed data day definition and equal area grid, refined definition of data products. (Jan. 28-30)

ii. Bob Evans attended Pathfinder Chairs meeting. Presented Data day definition and equal area grid. Discussed AVHRR pre-launch sensor calibration for IR channels and sensor non-linearity. Presented slides to show implications for science due to AVHRR digitizer behavior. (Feb. 2)

b. Activities

i. Extract 1988 matchup data set for NOAA 9 using NOAA method of non-linear sensor calibration. Presently have first two months of 1988 extracted. These matchup data will be provided to any group interested in generating AVHRR algorithms and cloud filters.

ii. Pathfinder benchmark period for 1988 should begin in July, '93. Processing for data products will be distributed on a monthly basis, processing and distribution will be done by JPL,

matchup database generation and algorithm coefficient testing will be done by Miami.

iii. We have received 20 optical disks for 1987; 30% of the shipment have been duplicated. There are approximately 33 in one year's data.

iv. Have initiated discussions to define changes in the BIN programs and database control programs (VDC and MCP control databases) to support the data day definitions within Pathfinder.

v. The global 1988 test SST data set for both day and night was computed for days 265 through 311 using a 9.28 km. grid. Equal rectangular fields were computed using the test data; the results have been distributed to JPL and URI.

B.5.3.2 February

a. Direct Activities

i. Processing of Days 88180 Through 88311 Completed
   Miami Test Algorithms Used
   88311 was transition date to NOAA-11

ii. Daily Maps Prepared
   Fields shipped to URI and JP

iii. Second Pass Cloud Filter
   Implemented
   Applied to 28 days beginning with 88183
   Will be applied to remaining 1988 test fields and made available to review group

iv. Processing Rate
   6 to 7 data days per day
   Approximately two months per data year

b. Related Activities

i. EOS V0 1 Km Project
   Received and decoded baseband data tapes for April 1992 through November 1992
   Approaching real-time within two weeks
   Data copied to DAT and sent to EROS

ii. CZCS Processing for West Coast Time Series Completed

B.5.3.3 March

i. Extraction of global matchup data base continues, expect to be complete for 1988 NOAA-9 before end of April.

ii. January and February, 1988 processed to test data flow and to check capability of using reference images in a second pass data quality test. At present the second pass quality test is defined as follows:

a) All images for ascending (descending) passes are binned into daily files using data quality flags generated during Level 2 computation. These flags are based on channel differences 4-5, 3-5 and homogeneity for channels 4 and 5. Preference is given to lowscan angle data.

b) daily ascending (descending) files are composited into 1 week intervals selecting the warmest pixel for each week.

c) three weekly files are composited using warmest pixel using the weeks before, including and following the week of interest to create a reference file for the center week.

d) Each of daily files is compared to the reference file on a pixel basis. Pixels differing from the reference file by more than 2 degrees are rejected, i.e. the data quality mask for the pixel is marked as bad data.

e) New weekly files are generated by averaging the weekly files. The number of undetected bad pixels per daily file is reduced order several hundred.

f) Processing of 1988 global data halted pending implementation of new SST algorithm and spatial data day definition in Level 3 binning programs.

B.5.3.4 April

i. Processing Efficiency:

We have been examining the processing path for Pathfinder to improve efficiency. January and February, 1988 were reprocessed and reference fields generated to improve rejection of invalid data. Several approaches were examined:

1. Current Scheme: using combination of day and night fields
2. using 3 week time interval with addition of 2 additional weeks where bad data is still present.

ii. Processing Rates:

We reviewed limitations constraining processing rate; two approaches have been utilized to increase performance:

1. include caching capabilities in DSP I/O library.
2. utilize in core ram disk for intermediate files where software for the particular machine/operating system and a suitable amount of memory exist.

Applying these two steps improved the rate of matchup data extraction from 3+ months/data year to two weeks. Used new approach to run 1987. 1989 extraction in progress.

iii. Testing New Reference Image Method

Matchup Data Base. Data for all months in 1988 has been extracted . A matchup data was built for 1985-1988 for east coast US and Jan-Nov 88 for global.

iv. Database Processing Control

1. Data base process control was extended to remove overlap between L-1b tapes. There is approximately 5 minutes of duplicated data that bridge that end of one GAC record session and the initiation of the following recording session.

2. Data base process control was extended to allow multiple processors from differing vendors to participate in process load sharing.

3. Added test to PATHCLOUD reference program to permit data that had been rejected by the PATHNLC sst program due to homogeneity and channel difference tests to be included in the difference from the reference map is within acceptable limits. A new quality flag has been defined to distinguish this case.

4. Improved data day programming to properly distinguish cases where a file crosses the poles and the data line.

5. Adding caching I/O libraries to DSP. tests will be made to determine the extent of improvement in processing rate.

6. Added capability to Pathfinder processing routines to remove extra test channels for production processing. these channels will be available for special processing or test/validation purposes.

7. We are preparing to test changes, then ship system to JPL for initial production testing.

v. Pathfinder algorithm:

1. Working with 85-88 east coast and Jan-Nov 88 global matchup data sets to generate and test Pathfinder algorithm.

2. Shipped matchup data sets to C. Walton to support NLSST algorithm work.

3. Working with JPL/NSIDC to define ice mask using SSMI data; this is also applicable to SeaWIFS.

B.5.3.5 May

i. Received several NLSST suggested algorithm coefficients from C. Walton for 1988/89.

ii. Implemented several algorithms in anticipation of benchmark processing.

B.5.3.5 June

Activation of DEC Alpha platforms and expansion of an FDDI based network has increased processing capacity from 5-7 data

days/day to around 15/day. A goal has been set to achieve 20/day which is the throughput limit for a single optical jukebox; this corresponds to handling 600 MB/hr.

B.5.4 DSP

B.5.4.1 June

i. Jim Brown has been working with data processing personnel to further development of DSP software.

ii. Dalu Li prepared a number of test runs of SeaWiFS binning programs to validate processing on multiple processing architectures.

B.6 Team Interactions

B.6.1 January through March

Meetings listed as part of projects; no others listed

B.6.2 April

Robert Evans attended the Pathfinder science working group meeting and presented a report of progress to date.

B.6.3 May

Meetings listed as part of projects; no others listed

B.6.4 June

Meetings listed as part of projects; no others listed

C. FUTURE ACTIVITIES

C.1 Database Future Work

C.2 Client/Server Future Work

C.2.1. Creation of a resource manager and a performance monitor.

C.2.2. Expansion of the error handler to provide broader coverage and to integrate into the overall system error recovery scheme.

C.2.3. Continue testing the client/server with CZCS and AVHRR data. This would include the acquisition of a UNIX resident database to run parallel tests.

C.2.4 Continue enhancement of processing efficiency through greater use of concurrent processing.

C.2.5 Examine other processing schemes to determine which elements could be either included or adapted for use within the client/server concept.

C.3 Pathfinder (P)

C.3.1. Continue development of linking processes between *in-situ* and processed satellite data.

C.3.2. Expand the validation dataset.

C.3.3 Continue to work to expand processign capacity toward the 20 datadays/day goal.

C.4 Headquarters (H)

C.4.1. Create tools to assist in results interpolation.

C.4.2. DSP - Fix programs that access the graphics plane to use the navigation from the input image and not the graphics plane.

C.4.3 Refine PATH binning and mosaic pixel quality algorithm to eliminate clouds.

C.4.4 Verify workstation DSP (SGI, SUN, DECstation, VAXstation) by comparing each program's output with the Adage system.

C.5 Modis (M)

C.5.1. Continue working with H. Gordon on an implementation of prototype ocean color atmospheric correction algorithms.

C.5.2 Continue working with D. Clark on *in-situ* database requirements.

D. PROBLEMS

D.1 Database Problems

None listed separately

D.2 Client/Server Problems

D.3 Matchup Database Problems

None listed separately

D.4 DSP And Headquarters Related Problems

None listed separately.