MODIS
SCIENCE DATA SUPPORT TEAM
PRESENTATION

March 20, 1992

AGENDA

Page

1. Action Items 1
2. MODIS Airborne Simulator (MAS) Status 2
3. Cloud Optical Program 7

ACTION ITEMS:

01/03/92 [Team]: Check on the set of software engineering tools available in Code 530 to see if any of these would be of use to the SDST. (Discussions were held previously with Frank McGarry of Code 530. The file dump algorithm and the cloud algorithm were processed using PR:QA when it was available in Code 563.2.) STATUS: Open. Due date 02/14/92.

01/17/92 [Tom Goff]: Have a polished version (with peer review) of the file dump routine ready for the MODIS Science Team Meeting. STATUS: Open. Due date 04/01/92.

02/21/92 [Ed Masuoka]: Talk to Code 930 and find out what tools they have for porting data between computers from different vendors. STATUS: Open. Due date 03/13/92.

02/21/92 [Lloyd Carpenter and Team]: Identify a list of risks associated with porting Team Members' algorithms to the PGS. Prepare these for discussion at the Science Team Meeting. STATUS: Open. Due date 04/01/92.

02/28/92 [Liam Gumley]: Develop a plan to accelerate MAS processing using less of Liam's time. STATUS: Open. Due date 03/20/92.

03/13/92 [Ed Masuoka and Liam Gumley]: Find out what is involved in setting up the catalogue for MAS data. STATUS: Open. Due date 04/03/92.

MODIS Airborne Simulator status (Liam Gumley)

Progress up to 19 March 1992

(1) MAS data processing status

Flight Date	Area covered during flight	Level-0 data received	Processing completed	INS offset fixed
10/31/91	Ames test flight CA/NV	yes	3/3 tracks	yes
11/12/91	Ferry flight CA to TX	yes (subset)	1/1 tracks	no
11/14/91	Coffeyville KS	yes	16/16 tracks	no
11/18/91	Coffeyville KS	yes	Underway	
11/21/91	Coffeyville KS	yes		
11/22/91	Coffeyville KS	yes		
11/24/91	Gulf coast TX/LA	yes		
11/25/91	Coffeyville KS	yes		
11/26/91	Coffeyville KS	yes		
12/03/91	Gulf coast TX/LA	yes		
12/04/91	Gulf coast TX/LA	yes		
12/05/91	Coffeyville KS	yes	29/29 tracks	no
12/07/91	Coffeyville KS	yes		
11/16/91	Ground visible calibration	yes	10481 scanlines (no navigation)	
11/20/91	Ground visible calibration	yes	6078 scanlines (no navigation)	
11/23/91	Ground visible calibration	yes	10281 scanlines (no navigation)	

(2) MAS/INS clock offset correction

I commenced work on a program to estimate as accurately as possible the offset between the MAS and INS clock data, based on the roll information from each system. This program will perform the following functions.

- Ingest MAS and INS roll versus time data for a specified time period,
- Interpolate the MAS and INS roll data to a common time interval using a cubic spline algorithm,
- Compute the cross correlation of the MAS and INS roll data sets, with time lag in a direction specified by the user,
- Interpolate the cross correlation data using a cubic spline algorithm,
- Compute the first derivative of the interpolated cross correlation data,
- Determine the local maximum of the cross correlation data by finding the zero crossing point(s) of the first derivative using a secant algorithm.

The local maximum (possibly confirmed by the user) will be the offset time between the MAS and INS time data. This needs to be determined as accurately as possible so that other instruments which collocate data with the MAS (e.g. lidar) can ensure time concurrency. The software tools necessary to perform each of these steps have already been developed previously, and are currently being developed into a functional main program.

(3) MAS data catalog development

I spoke to several people during the week regarding the process of setting up a MAS data catalog. Most of these people were associated with the Goddard Version-0 DAAC. As a result of these discussions, several points arose.

(a) The NASA Master Directory (MD) is an appropriate place to set up a Directory Service for the MAS. The format for data entry into this service is standard (Directory Interchange Format , DIF) and I have documentation on the format. To compile the necessary information and set it up on the MD apparently takes less than a day. The MD is a menu option in the NSSDC online data information system (NODIS) which may be accessed via TELNET over Internet. The MD also provides 'links' to other services, such as the NASA Climate Data System (NCDS).

(b) The NASA Climate Data System (NCDS) currently offers the type of Guide Service and Inventory Service which we desire for MAS. NCDS accepts both standard metadata entries, and dataset-specific metadata entries. There is a prescribed format for these entries, and it is possible to submit the prepared metadata entries in electronic form.

(c) The Goddard Version-0 DAAC is in the early stages of development, and does not offer the functions we require in the near term. However in the next 6 to 12 months the DAAC will be accepting prototype datasets for archive and distribution. The capabilities of the NCDS will eventually be folded into the V0 DAAC, and the NCDS will cease to exist.

It seems that the NCDS would be an appropriate place to handle the MAS data catalog in the near term. Lessons learned from this experience could then be applied to transferring the data catalog (and possibly distribution) functions for MAS to the V0 DAAC.

(4) Plan for accelerating MAS processing using less of my (Liam) time

With the current manpower and hardware facilities available, it appears that under optimal conditions about one MAS flight can be processed in one working week. This involves the following steps.

(1) Transporting tapes from RDC to GSFC
(2) Ensuring sufficient disk space is available for Level-0 data
(3) Mounting Level-0 tapes
(4) Reading Level-0 tapes onto LTP VAXcluster
(5) Checking Level-0 data for anomalies or problems
(6) Determining the clock offset between the MAS and INS
(7) Determining flight track time limits from INS data
(8) Computing navigation regressions from time limit data
(9) Computing calibration data (batch mode)
(10) Ensuring sufficient disk space is available for Level-1 data
(11) Creating Level-1B datasets (batch mode)
(12) Ensuring sufficient disk space is available on FTP site
(13) Moving Level-1 data to FTP site (batch mode)
(14) Writing Level-1 data summary for FTP site
(15) Notifying users that Level-1 data is available

Most of these steps are not overly complicated, however they are time-consuming. In order to speed up these tasks, while using less of my time, I believe two things need to occur.

(a) Part time data-tech/operator/programmer support

The MAS data processing software was not designed to be an automatic system. Human input is required at many points in the processing sequence, from basic things such as loading magnetic tapes, to more sophisticated tasks such as breaking up the data into straight line flight tracks. To this end, the availability of a support person to run the MAS software would be of great benefit. This person would take over most, if not all of the tasks listed previously. Most of these tasks are of such a nature that they are routine, and only require adjustment/fixing if severe problems occur or new information is encountered. The person would likely be a junior programmer with experience on Unix systems and VAX systems. I believe such a person working on an approximately 50% time basis would have little trouble coming up to speed on MAS processing.

The MAS software still requires upgrading in terms of features and capabilities. Examples are

- Semi-automated determination of MAS/INS clock offset
- Automatic determination of straight line flight tracks
- Automatic flagging of noisy image channels
- Automatic extraction of Level-1B metadata items
- Upgrade to possible new Exabyte Level-0 data format when available
- Upgrade to 50 channel 12 bit data system when available
- Incorporation of MODIS prototype algorithms

The use of a part time programmer for 'operational' MAS processing would free up my time to work on these and other MODIS related projects.

(b) Allocation of sufficient hardware resources

A major complicating factor of the MAS processing is the volume of the data involved. It is only possible to load at most 4 tapes of Level-0 data (about 0.775 GB) on the VAX cluster at any one time. It is also only possible to create about 1 GB of output Level-1B data on the Iris at any one time. Once created, the Level-1B data must be moved to the FTP site, as long as space exists on the FTP site.

The solution to this problem is to have more disk space available for processing. As reported previously (SDST report 02/21/92) it is believed that about 5 GB of dedicated disk space is required for optimum MAS processing. This would be in a system (workstation) that has Exabyte and 9-track magnetic tape devices for Level-0 data ingest and Level-1 data archival/distribution.

The CPU speed of the processing platform is not such a problem. The main overhead in the processing is data I/O, so disk and interface speed is more important than raw CPU speed.

NASA Climate Data System (NCDS)-NCDS

Description:
The NASA Climate Data System (NCDS) is an interactive information system which provides tools to support atmospheric and oceanographic research. NCDS is composed of two subsystems; Data Access and Data Applications. The Data Access Subsystem includes summary descriptions for rapid scanning of data set characteristics, as well as detailed descriptions of data set attributes such as parameters, sources, data processing cycle, quality, access procedures, bibliographic references, etc. In addition, NCDS provides detailed temporal and spatial information about the data sets archived by the system, as well as the capability of extracting subsets of data. The Data Application Subsystem provides tools for browsing, analyzing, and plotting supported data sets.

NCDS supports a variety of data from polar orbiting satellites such as the NIMBUS and NOAA series, as well as the GOES geostationary satellite series. In-situ measurements, such as surface station climatologies, and data gathered by coordinated scientific projects (e.g. First GARP Global Experiment, International Satellite Cloud Climatology Project) are also supported. NCDS supports many NASA-funded scientists in research projects dealing with atmospheric composition, clouds and radiation, global climatologies and oceanography, and solar activity/irradiance.

Project Manager:
Lola M. Olsen
NASA/GSFC
Code 934
Greenbelt, MD 20771
Email: NSInet > NCF::OLSEN
Email: INTERNET > OLSEN@NSSDCA.GSFC.NASA.GOV
Phone: (301) 286-9760
Phone: FTS 888-9760

System Contact:
NCDS User Support Office
NASA/GSFC
Code 934
Greenbelt, MD 20771
Email: NSInet > NCF::NCDSUSO
Email: INTERNET > NCDSUSO@NSSDCA.GSFC.NASA.GOV
Phone: (301) 286-3209
Phone: FTS 888-3209

Access Procedures:
NCDS is located on a VAX 6410 system at NASA's Goddard Space Flight Center and is available for use by NASA-funded and university researchers.

Access to NCDS data information is available through the Master Directory, but use of full NCDS capabilities requires an NCDS account. To request an account for NCDS, call the NCDS User Support Office.

Ordering / Price Policy:
There is no direct charge to approved users of NCDS.

Available Distribution Media:
NCDS supports data in NSSDC Common Data Format (CDF). Data in CDF is accessible for analysis via the NCDS Data Application Subsystem. Several output devices are supported, including roller pen, ink jet, and laser plotters.

Cloud Optical Program and MAS

Thomas E. Goff
19 March, 1992

teg@ltpiris2.gsfc.nasa.gov,
(301) 982-3704
tgoff on GSFC mail

* **Cloud Optical Program Port, Continued** - Mike King's Cloud Optical Program (cldopt.f) is now executing on the SGI ltp iris computer. The port of the execution phase of this program was accomplished by modifying the code to place the large data arrays that had been placed into subroutine argument lists, into a common area. This was necessary because the SGI iris passes subroutine arguments by value instead of address and the large size of these arrays was producing a stack overflow condition. In addition, all necessary variables were explicitly set to zero upon program initiation. A related problem to uninitialized variables is the use of volatile subroutine variables by default. If a variable that is declared in a subroutine and used upon multiple entries to that subroutine, then there is no guarantee that the value of the variable is retained upon subsequent calls. Variables must be declared as type static to prevent their values from being changed when the subroutine is swapped to disk and brought back into a different part of physical memory before being called.

  The original data set was designed to allow the program to terminate upon reading an EOF from the input data stream. The subroutine stubs that I wrote to interface the FTIO routines into the UNIX disk structure detected the EOF condition but could not pass this back to the FORTRAN calling routine via the FORTRAN error return extensions (CALL FWRITE(, ,*25,*26)). The cldopt.f code was modified to bypass this extension. The original program further confused it's internal variables which resulted in a termination due to a user specified IBM run time limit instead of the normal programmatical termination. This was corrected in the input data set to specify only one run case instead of the original two plus runs. The IBM 3081 execution was terminated by the operating system after 54 minutes of execution and produced two output cases. The SGI iris executed one case in 3 minutes. This equates to an approximate order of magnitude decrease in execution time on the iris when compared to the IBM 3081. The output data set that is sent to unit 6 was compared between the IBM and SGI iris runs and found to be identical.

  This porting of the actual code took less than three days. However, the port of the data sets took approximately three weeks. The code is now not only badly structured but also cobbled up.

* **MAS INS Data Plotting** - I have a hacked program on the SGI iris that will read the MAS *.ins data files and produce an output file that can be used as data to the PC based acrospin program. This will produce a real-time display of the 3-D aircraft flight path for any of the MAS flights that can be displayed, rotated, scaled, etc on any PC under user real-time

control. Although this plot can not be used for determining the straight line flight paths, it can be used to visually determine the coincidence of the race track segments and any elevation changes as the fuel load lightens. Ground coverage can be deduced from this display plot. If access to the Apple Mackintosh MacSpin program can be made available, I will be glad to provide a clone of the previous hack to produce MacSpin specific MAS flight track data sets to those parties who have an interest. This display illustrates the coarseness of the INS data and why the INS data needs to be filtered before a flight line can be determined.

* **MAS Graphic Overlays** - The MAS data has been ingested by the Khoros public domain imaging program which has been installed on Virginia Kalb's Sun 4. The data was ingested via my subset program, available on our anonymous ftp site. This program was then modified by Virginia to include a representation of the auxiliary data (lat-long, az-el, etc) as a graphic overlay to the original data. When a better facility for obtaining hard copy of these images is available, we will be able to produce quality assurance products to validate the ground location and other ancillary data. Note that Khoros will only run in a computer that has the X11 rev 4 windowing facility. This capability will be expanded, on a time available basis, to include software tools slanted toward the remote sensing disciplines.

* **The MAS subset Program** - The program subset was executed in two formats: the original hyper slab method and a line at a time method on both the SGI IRIS and a SUN computer. The time to execute the two versions of this program was almost identical, indicating that the NetCDF library functions perform a good job of paging the data file without causing an undue amount of disk accesses.

  The subset program that currently exists on the ftp site is an early version which was made available to users who do not wish to write their own program. Several users have picked up this program and are now using it routinely to produce image data for ingesting into image processing programs. I have a much newer version (3.1 currently) of this program which incorporates a limited data conversion to 8 and 32 bit data in addition to the existing 16 bit. I have been asked by several users to incorporate new capabilities into this program to support (for example) floating point image data. This program is the type of program that can grow into a full conversion program from NetCDF images to any flat file image format. A limited capability to add graphics overlays to the MAS data has been incorporated (see above). Because this program needs to be a flexible program with many enhancements and changes added or modified per individual user request, it is a good candidate for recoding into the C++ language. This would permit a module inheritance to be used to good advantage.

* **The fdump Utility Program** - We have completed our (RDC's) internal review of the fdump program which is being used as a model for the discussion of computer coding standards with the MODIS Team Members. This program is a reasonably complicated program which, in my opinion, pushes the envelope of acceptable coding style and standards. The modules are very large but well delineated by content in the embedded comments. The program has a maximum nesting of four which is as deep as nesting should be.

-- Miscellaneous --

& **VAX Cluster Modem Lines** - 53241 now appears to work consistently.

& **UNIX Mail** - The SGI mail facility succeeded in transferring mail to/from an outside computer on the internet.

& **SLIP Connection** - I have revision 3.5 of SUN's PC-NFS installed on my PC. I will now establish the procedures required to connect to the GSFC NCCS slip server. This is a non-standard implementation due to the existence of the Rohm switching system and may take some time and patience to make fully functional.

& **FORTRAN Source Code Checking** - I have a trial evaluation copy of FORTRAN-LINT arriving by UPS which I would like to install on the SGI ltpiris2 computer. This software requires the X11 rev 4 windowing facility and can be installed when this facility is available. I have given this tape to the ltp facility manager (Ed Masuoka) to be installed on any IRIS when time allows.

& **Current Source Code Checking Facilities** - We plan to be using the LINT and public domain FTNCHEK programs to check our C and FORTRAN code for non-portable or wasteful features. The LINT program on the iris complains of an undefined variable which is defined by the operating system in the < stdio.h > include file. This needs to be exorcised by SGI in the future.

& **LTP Computer Facility** - It would be very useful, especially for the off-site people, to have a synopsis of the ltp computer facility. This could include, but not be limited to, the following:

A list of the computers with the operating system type, the current revision level of the operating system, and if a UNIX based system, which UNIX (BSD or AT&T) was used as the basis for the vendor supplied UNIX clone. Also which windowing or GUI is provided on the computers.

The physical location of the computers and what means of accessibility they have (locked rooms, etc).

A list of available peripherals with locations and how to access these devices ala help files etc. This would include the location of post script, HP pcl, and HPGL devices; and magnetic tape systems with density, format, and network or direct computer access. How do we obtain image B/W or color hard copies, and what software is available for image format conversions?

A list of major software that is available on the various ltp computers. We have interest in publishing tools, word processors and word processing translator capabilities, image processing packages with a list of capabilities (classification techniques, panable graphics overlays, user specifable translation equations, etc), data

representation tools (2 and 3-D plotting, contour plotting, FFT's and similar transforms), spreadsheet and similar columnar data translating programs and editors, library tool kits for math and general data manipulation, etc.

A list of direct contacts for obtaining additional information on networking, operating system and computer architectural problems, software usage, ftp sites, usenet connections for public domain, shareware, and freeware programs, computer system factory support (which systems are supported, by whom, with what level of response), etc.

An informal discussion with Ed Masuoka and Penny Masuoka was held outlining the desirable capabilities that an image processing system in the MODIS era might require. This was also expanded into data visualization tools in general.

c:\modis\status.wp