I. SATELLITE DATA PRODUCTS

The table below illustrates the variables, original resolution, and units of the satellite data which are available for the main two types of GOES data – Imager (IMG) and Sounder (SND).

DATA MODE	VARIABLE	ORIGINAL RESOLUTION	UNITS
IMG	Cloud Albedo	4 km	%
IMG	Cloud Top Pressure	4 km	Pascals
IMG	Insolation	4 km	Watts per Square Meter
IMG	Surface Albedo	4 km	%
SND	Cloud Top Pressure	10 km	Pascals
SND	Skin Temperature	10 km	Kelvin

Both the skin temperature and cloud top pressure retrievals rely on NAM model first guess data. Cloud top pressure (CTP) is available as an Imager or Sounder product. The Imager product is approximately 4 km resolution, but is a simple infrared temperature to pressure retrieval. The Sounder product is at a coarser resolution of about 10 km, but is a combined retrieval method using the IR temperature for lower clouds and CO_2 slicing method for the higher clouds and is therefore a more robust CTP product.

II. SOFTWARE INTRODUCTION

This software system was developed within a csh environment on a Linux cluster using the Portland F90 compiler. We can not guarantee that it will work on all systems but it should be fairly robust. The main program files sat_regrid.f90 and reformat.f90 have been written using modules for all subroutines. To compile these files the script new.csh is used in the respective directories (REGRID for sat_regrid.f90, and REFORMAT for reformat.f90) that splits the program file into its individual subroutines (modules) and then compiles them in the appropriate order. If these scripts do not work on your system or if you prefer "module free" F90 versions we can provide them upon request. Program sat_regrid.f90 reads GOES data files in grib format and averages or interpolates the desired fields to a specified analysis grid and outputs the results in FORTRAN binary files. Program reformat.f90 uses the latter binary output and creates ascii output files that can be used with the pre-processor in MM5. Test data is included in the setup process that can be run as described in section VII.

III. SYSTEM SETUP

For the sake of this discussion the main directory you choose for this work will be called "PROJECT". From the website `http://goes_assim.nsstc.uah.edu` download the file sat_regrid.tar to the PROJECT directory. Extract the files by the following UNIX command:

tar –xvf sat_regrid.tar

.

One of the files extracted is “setup.csh”. The top portion of setup.csh is shown below:

```
#!/bin/tcsh -f
#
#######################################################################
#
# These variables will have to be set to the appropriate values
# for your system .....
#
set COMPILER = "pgf90"
set OPTIONS = "-O2 -byteswapio -Mextend "
set UNIX_LARGE_FILE = "/usr/local/pgi/linux86/5.2/liblf"
#
#######################################################################
#.
```

The user will need to make the appropriate changes for these 3 script variables before running setup.csh. If you are using the Portland compiler then “COMPLIER” and “OPTIONS” will require no changes. The default version of sat_regrid.f90 reads and writes files that are small in size, but if the user makes changes to output large files (depending upon the system, typically around 2 GB) then you will need to link to the respective compiler’s large file library which is “UNIX_LARGE_FILE”. That option is made available but if you are using the default version setting UNIX_LARGE_FILE to a blank should work.

After editing setup.csh for these changes, do the following from the PROJECT directory:

./setup.csh >& setup.LOG

The script setup.csh will create the default environment for running programs sat_regrid.f90 and reformat.f90. Any problems which occur during the setup process will be in the file setup.LOG.

The following are the default directories which are created: DATA_OUT, DATA_ASCII_OUT, GOES_DATA, MISC, REGRID, REFORMAT, RUN, SCRIPTS, WEIGHTS_IMG, WEIGHTS_SND, and WGRIB. DATA_OUT contains the output binary data after sat_regrid.f90 is run, DATA_ASCII_OUT contains the output ascii data after reformat.f90 is run, GOES_DATA is where the tar files downloaded from the website

`http://goes_assim.nsstc.uah.edu`

need to be located, MISC contains some of the “leftover” files from the setup process, REGRID contains the file sat_regrid.f90 and all the other files and scripts required to compile it, REFORMAT contains the file reformat.f90 and all the other files and scripts required to compile it, RUN is where an actual execution of sat_regrid.f90 and reformat.f90 occurs, SCRIPTS contains all the scripts which are needed to run

sat_regrid.f90 and reformat.f90, WEIGHTS_IMG contains the weighting files for the GOES Imager data (~ 4 km original resolution), WEIGHTS_SND contains the weighting files for the GOES Imager data (~ 10 km original resolution), and finally WGRIB contains the source and executable code for the wgrib utility. If you want to be sure to have the most recent version of the wgrib utility you may download it from http://www.cpc.ncep.noaa.gov/products/wesley/wgrib.html.

IV. SAT_REGRID.INP INPUT FILES

During the setup process two input files are created in the RUN directory: sat_regrid.IMG.inp and sat_regrid.SND.inp. They are identical except for the some of the settings that will be discussed below.

The default version of the file sat_regrid.IMG.inp is shown below which is in the default run directory. Your version will be identical except that during the setup process the names of the directories in the namelist section "dir_in" will have been automatically been changed to your values.

```
!
!**********************************************************************
!&dir_in
data_in_dir =  '/nas/rstor3/dotykg/NASA_APPLICATIONS_NEW_TEST/GOES_DATA'
wgt_img_dir=   '/nas/rstor3/dotykg/NASA_APPLICATIONS_NEW_TEST/WEIGHTS_IMG'
wgt_snd_dir=   '/nas/rstor3/dotykg/NASA_APPLICATIONS_NEW_TEST/WEIGHTS_SND'
data_out_dir = '/nas/rstor3/dotykg/NASA_APPLICATIONS_NEW_TEST/DATA_OUT'
script_dir =   '/nas/rstor3/dotykg/NASA_APPLICATIONS_NEW_TEST/SCRIPTS'
wgrib_dir =    '/nas/rstor3/dotykg/NASA_APPLICATIONS_NEW_TEST/WGRIB'
run_dir =      '/nas/rstor3/dotykg/NASA_APPLICATIONS_NEW_TEST/RUN'
/

!  Choices for instrument source are:
!
!  data_mode = 'IMG',  GOES Imager, original resolution of about 4 km
!  data_mode = 'SND',  GOES Sounder, original resolution of about 10 km
!
!  PLEASE SEE DOCUMENTATION FOR FURTHER DETAILS ...

&var_use
read_calb =    .true.
read_ctp =     .true.
read_insl =    .true.
read_salb =    .true.
read_tskin =   .false.
data_mode =    'IMG'
/

&date_range
nyr_beg =      2006
nmo_beg =      07
ndy_beg =      15
nhr_beg =      00
nmin_beg =     0
nyr_end =      2006
nmo_end =      07
ndy_end =      15
nhr_end =      23
nmin_end =     59
/

&analysis_grid
```

```
nx =           164
ny =           128
glat_cen =     +40.0
glon_cen =     -97.0
lat_north_analysis = +45.0
lat_south_analysis = +33.0
dxg =          36.0
analysis_mode =       'AVE'
glat_fname = 'mm5.2006.36km.glat.dat'
glon_fname = 'mm5.2006.36km.glon.dat'
/

&write_out_options
write_raw_satellite =  .true.
/

&weight_options
store_weights =               .true.
max_mem_mb =          250
/

&fine_grid_setup
dxf_in =              50.0
/

!**********************************************************************
!
```

Each of the namelist sections will now be discussed in the order in which they appear in the file.

A. NAMELIST section "dir_in"

The program sat_regrid.f90 is designed for a given run to either process entirely Imager data, or entirely Sounder data. The file sat_regrid.IMG.inp shown above is designed to process Imager data. Therefore, it will require two runs of sat_regrid.f90 to process all the data types for a given analysis grid configuration and averaging/interpolation option choice (discussed later). This is the reason for the two types of input files. For more than one analysis grid and/or more than one choice of averaging/interpolation options, then one or more of the directories will generally need to be changed. The directory setup can only be made clear by a description of the different run types.

Consider first the case of an analysis grid mesh size that is much larger than the satellite resolution. As defined above this version of the sat_regrid.inp file would process GOES Imager data (set by data_mode='IMG' variable in the namelist section "var_use"). The specified analysis grid resolution is set at 20 km (set by the "dxg" variable in the namelist section "analysis_grid"). Whether averaging or interpolation is performed is determined by the value of the namelist variable "analysis_mode" – it must be "AVE" (for averaging) or "INT" (for interpolation). In general, for analysis grid mesh sizes much larger than the satellite resolution (about 4 km for the Imager data and 10 km for the Sounder data) one normally would want to average the satellite data to the analysis grid. Here in the default version analysis_mode = 'AVE' so averaging is performed on the Imager satellite data. Changing only the data_mode to "SND" and then running sat_regrid.f90 again from the same run directory would then process the same set of satellite data for the Sounder data and again use averaging to create the analysis data.

After these two runs the wgt_img_dir and wgt_snd_dir directories will contain the weighting information for both data types for averaging the data for one analysis grid. The output data from these two runs will both be in the same output directory "data_out_dir" if the default value was not changed between the two runs. If now a different run is desired, then in general the directories wgt_img_dir, wgt_snd_dir, and data_out_dir would need to be changed.

Consider now the case of an analysis grid mesh size that is much smaller than the satellite resolution. For this case one will need to interpolate the data to the smaller grid and this is obtained by setting analysis_mode = 'INT'.

For analysis grid resolutions that are between 4 and 10 km one may have to choose averaging for the Imager data and interpolation of the Sounder data by selecting the correct value for the variable analysis_mode.

Once the weighting information has been written to the wgt_img_dir and wgt_snd_dir directories, a file called "DONE" is created in each weight directory. The DONE file includes all the relevant information for the satellite and analysis grid and the analysis_mode values. If one attempts a run which conflicts with the values in the DONE file then sat_regrid.f90 is terminated and one will have to diagnose the problem from the sat_regrid.log file. The DONE file should stop most runs that have not made the correct choices in the sat_regrid.inp file.

The file sat_regrid.SND.inp is identical to the file sat_regrid.IMG.inp except different variables are selected in the "var_use" namelist section and data_mode is set to "SND" to process Sounder data.

B. NAMELIST section "var_use"

As discussed in the last section, the value of data_mode determines whether Imager or Sounder data is to be processed. For Imager data (data_mode='IMG') one can process the specific variables of cloud albedo, cloud top pressure, shortwave insolation, and surface shortwave albedo. These variables can be selected by setting read_calb, read_ctp, read_insl, and read_salb, respectively, to "true". If one selects read_tskin for Imager data it is automatically set to false since that data type and variable combination do not exist. In a similar way for Sounder data (data_mode='SND') one can process the specific variables of cloud top pressure and skin temperature. These variables can be selected by setting read_ctp, and read_tskin, respectively, to "true". If one selects any other data types they are automatically set to false for the same reason as just mentioned.

C. NAMELIST section "date_range"

In this namelist section the beginning and ending year, month, day, UTC hour, and minute can be specified to determine the time window of the desired data.

D. NAMELIST section "analysis_grid"

This section specifies the desired analyis grid. The variables glat_fname and glon_fname are the file names of the 2D latitude and longitude grids, respectively, which define the analysis grid. In this program, the array location (i,j) defines "i" as the west-east coordinate and "j" as the "north-south" coordinate. This is the "transpose" of the MM5 convention where "i" is the north-south coordinate and "j" is the "west-east" coordinate. If the analysis grid is a cross-point grid then the ascii latitude and longitude files (mm5.2006.36km.glat.dat and mm5.2006.36km.glon.dat are the example defaults) would be created from within MM5 by FORTRAN code something like the following to account for the coordinate change:

```
open (unit=99,status='new',form='formatted',file='glat.dat')
do I = 1, ilx
do j = 1, jlx
write (99,'(2i5,f10.3) j, i, xlat(j,i)
end do
end do
close (unit=99)
```

The longitude file would be created in a similar way. These files must either be linked or local to the RUN directory. The variable nx is the number of analysis grid points in the west-east direction, and ny is the number of analysis grid points in the north-south direction. Again, to use the previous example, if within MM5 ilx=128 and jlx=164, then in the input file nx=164 and ny=128. The variable glat_cen is the latitude of the origin or pole of the projection for the analysis grid, and glon_cen is the longitude of the same location. The variable dxg is the horizontal mesh size in km of the analysis grid, and analysis_mode determines whether averaging or interpolation is to be performed. The variables "lat_north_analysis" and "lat_south_analysis" are the two standard parallels for the Lambert-Conformal projection for the analysis grid.

E. NAMELIST section "write_out_options"

By choosing write_raw_satellite=.true. the raw (original) data are output as binary files.

F. NAMELIST section "weight_options"

If store_weights is set to ".true." and the required memory is less than the size in megabytes given by the specified value of max_mem_mb, then the data weights are read into memory. If store_weights is set to ".false." or if the memory requirement exceeds the value of max_mem_mb, then the data weights are read from the weight files each time they are needed. The latter option is of course much slower.

G. NAMELIST section "fine_grid_setup"

The value of dxf_in in units of meters is used to determine a fine mesh size that enables the determination of the portion of the area of a satellite data grid cell which overlaps a given analysis grid cell. Values of 50-100 m are usually sufficient.

V. REFORMAT.F90 INPUT FILE

The default version of the file reformat.inp is shown below which is in the default run directory. Your version will be identical except that during the setup process the names of the directories in the namelist section "dir_in" will have been automatically been changed to your values.

```
!
!**********************************************************************
!

&dir_in
data_out_dir =             '/rstor15/dotykg/NAPP_TEST/DATA_OUT'
data_out_ascii_dir =       '/rstor15/dotykg/NAPP_TEST/DATA_ASCII_OUT'
script_dir =               '/rstor15/dotykg/NAPP_TEST/SCRIPTS'
run_dir =                  '/rstor15/dotykg/NAPP_TEST/RUN'
/

&time_window
jdiff_max =  10
/

&clear_definition
clear_threshold =   66.0
/

&shortwave_cov
shortwave_cov_min = 95.0
/

&ctp_data_selection
ctp_data_mode =            'IMG'
/

!
!**********************************************************************
!
```

Each of the namelist sections will now be discussed in the order in which they appear in the file.

A. NAMELIST section "dir_in"

The only new variable in this section as compared to the same namelist section in the sat_regrid.inp file is the directory data_out_ascii_dir. This is the location of the output ascii data after the program reformat.f90 is run.

B. NAMELIST section "time_window"

The various original satellite data files have time stamps which are slightly different. The value of jdiff_max in minutes considers all satellite files which are with this time range to be the "same time".

C. NAMELIST section "clear_definition"

If the clear sky percentage of an analysis grid cell is defined as 100 – the percentage area covered by defined cloud albedo values, then an analysis grid cell is deemed "clear" if the clear sky percentage is greater than or equal to the clear_threshold value, and in the same manner it is deemed "cloudy" if the clear sky percentage is less than the clear_threshold value.

D. NAMELIST section "shortwave_cov"

During daylight hours when an analysis grid cell is close to outer limits of the satellite data coverage and/or is close to the day-night terminator, then there may only be partial coverage of the shortwave-related products. The percentage shortwave_cov_min insures that for satellite data area coverage of less than this value for an analysis grid cell will result in the shortwave-related products being set to undefined.

E. NAMELIST section "ctp_data_selection"

Since cloud top pressure is available from both the Imager and Sounder sensors, one needs a way to determine which product to use in the reformat.f90 program. If ctp_data_mode='IMG" then Imager cloud top pressures are used, and if ctp_data_mode='SND" then Sounder cloud top pressures are used. If one wants to make a decision to choose depending on data availability then changes would be required to the program reformat.f90.

VI. AQUIRING DATA

From the website `http://goes_assim.nsstc.uah.edu` download the desired satellite data into the directory "GOES_DATA". These files will consist of one or more "master"tar files that in turn contain other "gzipped" tar files that contain the satellite grib files.

VII. RUNNING SAT_REGRID

For a given run directory, once appropriate versions of sat_regrid.IMG.inp, sat_regrid.SND.inp, and reformat.inp are local, execution is done by using the script run.csh in the run directory.

The script run.csh requires three arguments on the command line, with each argument having two choices as shown below:

"img=yes" or "img=no"

"snd=yes" or "snd=no"
"ascii=yes" or "ascii=no".

If "ascii=yes" is desired then one must have "img=yes" and "snd=yes". The argument "img=yes" directs the script to run sat_regrid.f90 to process Imager data according to the choices made in the input file sat_regrid.IMG.inp. The argument "snd=yes" directs the script to run sat_regrid.f90 to process Sounder data according to the choices made in the input file sat_regrid.SND.inp. If "ascii=yes" then the program reformat.f90 is run that uses the binary output of the prior "IMG" and "SND" runs to create ascii output files. The three arguments can be in any order and can be in any mixture of lower-case and upper-case letters.

This example processes Imager and Sounder data and also creates ascii output:

./run.csh "img=yes" "snd=yes" "ascii=yes" >& run.LOG

The quotes are needed when this command is executed from a Unix line prompt.

The input files are set up to process some sample data that automatically exists in the directory "GOES_DATA". The command line above will process all the data using sat_regrid.f90 and reformat.f90.

If one is making a run for the first time in a new run directory (i.e., not the default run directory), then the script run.csh and the input files sat_regrid.IMG.inp, sat_regrid.SND.inp, and reformat.inp must be local. When running sat_regrid.f90 for the first time execution can take at least an hour or more for each data type since all the weighting files are being created. Runs made henceforth for the same analysis grid and averaging/interpolation choice then simply read the weight information and run much quicker (provided the weight information can be stored in memory). In our initial tests for a 200x200 grid with a 20-km grid mesh the creation of the weight files takes about an hour to complete for each data type. After the weight creation phase data processing proceeds on the order of a few minutes per day of satellite data.

VIII. MAKING CHANGES TO PROGRAM FILES

If you are using the default version that is module based, make any changes in the file sat_regrid.f90 in the directory "REGRID" and **NOT** in the "REGRID/split" sub-directory. Compiling the edited file is then performed by ./new.csh in the "REGRID" directory. In the same way make any changes in the file reformat.f90 in the directory "REFORMAT" and **NOT** in the "REFORMAT/split" sub-directory. Compiling the edited file is then performed by ./new.csh in the "REFORMAT" directory

IX. CAUTIONARY NOTES

Undefined values in the output data are set to "-999.0" for real values and –999 for integer values.

The binary data are output in arrays with indices (i,j), where “i” is the west-east direction which increases to the east, and “j” is the north-south direction which increases to the north. If you need a different index arrangement then you will need to make changes to the subroutines “output_data” and “output_raw” in sat_regrid.f90 or make the necessary changes to your read statements in your application program that utilizes the output binary data. The default arrangement is different from the MM5 convention where “i” is the north-south index and “j” is the west-east index.

The original (raw) satellite data are on a Lambert Conformal map projection with standard parallels at 30° N and 60° N. The default version of sat_regrid.f90 assumes that the analysis grid is also a Lambert Conformal map projection. If the user wants a different analysis projection then you will need, at the minimum, to make changes to subroutine “make_analysis_grids” and call your own map projection subroutines in place of “xytoll_lambert_conformal” and “lltoxy_lambert_conformal”.

Although the satellite data exist to portions of the west coast of the United States, these data are derived from the GOES-East satellite and therefore some if not all of the derived variables at the extreme satellite viewing angles are not reliable. It is up to the user to define what subsection of the analysis domain, if applicable, needs to be set to undefined for this reason.

The output binary data files (written by subroutine “output_data” in sat_regrid.f90) include information on the grid used, the variable name, the date, the actual data values, and the data coverage. It is up to the user to use the percentage values of data coverage to determine what analysis data values may have to be set to undefined because of low coverage by the original satellite data. This applies to the situation when the reformat option is not used (i.e., “ascii=no” in the run.csh command line) and the user is utilizing only the binary output data.