

GridSat Data Flow Diagram Documentation

The GridSat product is produced in a manner that is detailed in the GridSat Data Flow Diagram (DFD) and this accompanying document. The purpose of this document is to detail each step and intermediate file during production. There are three sections of this document: Files, Steps and Optional Paths. The Files section details how files are obtained, where they are stored and which processes use them. The Steps section spells out the processing components of each step of the process, whether it be IDL, FORTRAN, BASH scripts or human QC. Finally, Optional Paths are documented. These are either algorithms which lack thorough analysis (e.g., the ISCCP B1U centering algorithm) or are not required, like the IRWVP channel calibration. Optional paths are denoted by dashed arrows and grayed objects in the DFD.

Figure 1 - GridSat Data Flow Diagram

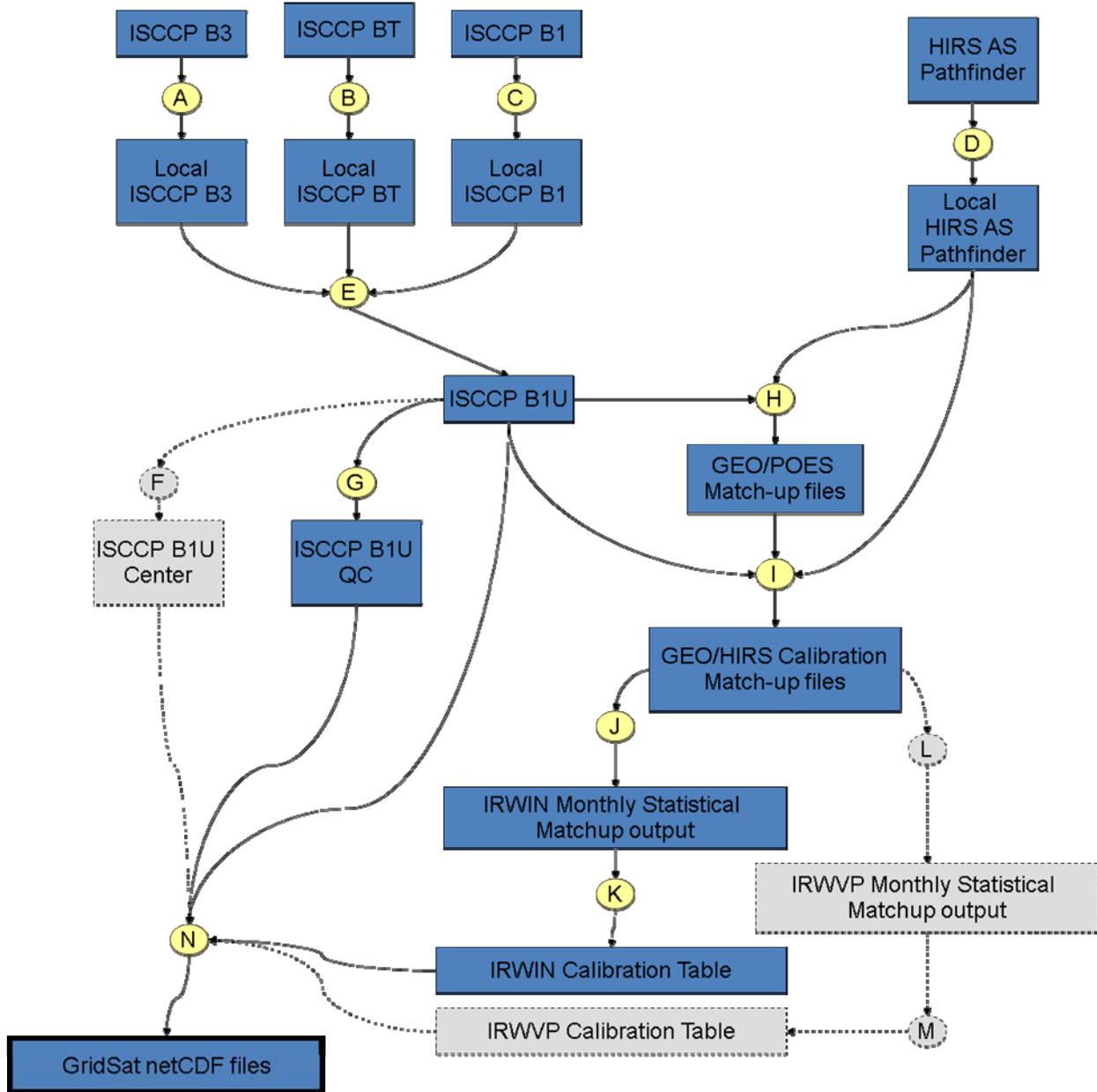


FIGURE 1 - GRIDSAT DATA FLOW DIAGRAM.....	2
FILES	4
File: ISCCP B3	4
File: ISCCP BT	4
File: ISCCP B1	4
File: HIRS AS Pathfinder.....	5
File: Local ISCCP B3	5
File: Local ISCCP BT	5
File: Local ISCCP B1	6
File: Local HIRS AS Pathfinder.....	6
File: ISCCP BIU	6
File: ISCCP BIU QC.....	6
File: GEO/POES match-up files	6
File: GEO/HIRS Calibration Match-up Files	7
File: IRWIN Monthly Statistical Match-up output	7
File: IRWIN Calibration Table	7
FINAL PRODUCT: File: GridSat-B1 netCDF files.....	8
STEPS.....	8
Step A: Obtaining ISCCP B3	8
Step B: Obtaining ISCCP BT	8
Step C Obtaining ISCCP B1	9
Step D: Obtaining HIRS Pathfinder AS	9
Step E: Creating ISCCP BIU	9
Step G: ISCCP BIU QC Algorithm.....	10
Step H: Writing GEO/POES match-up files.....	11
Step I: Performing GEO/HIRS Calibration Match-ups	11
Step J: IRWIN Monthly statistics aggregation.....	11
Step K: IRWIN calibration coefficient calculation.....	12
Step N: Creating GridSat-B1 data	13
OPTIONAL PATHS	15
Step F: ISCCP BIU Centering Algorithm.....	15
File: ISCCP BIU Center.....	15
Step L: IRWVP Monthly Statistics Aggregation.....	15
File: IRWVP Monthly Statistical Match-up output	15
Step M: IRWVP Calibration Coefficient Calculation	16
File: IRWVP Calibration Table	16
APPENDIX A: SUBVERSION FILES	17
APPENDIX B: FILE NAME CONVENTIONS	18
File: GEO/HIRS Calibration Match-up Files.....	18
File: IRWIN Monthly Statistical Match-up output	18
File: IRWIN Calibration Table	18

Files

All data files reside under one primary ISCCP directory. In the following discussion, ISCCPDIR will be used to denote the directory in which all subdirectories are maintained. This variable (ISCCPDIR) is not used by the processing, but the code could be converted to use a generic ISCCPDIR location.

In the following discussion of files and directories:

ISCCPDIR=glaze:/raid1a/raid/nb/kknapp/data/isccp

or symbollicly linked at:

ISCCPDIR=glaze:/home/kknapp/DATA/isccp

File: ISCCP B3

Purpose: These data are required by GridSat inasmuch as they provide ISCCP calibration data to be used in pre-calibrating and producing ISCCP B1U data. However, this is not truly required since data are re-normalized (i.e., calibrated) to HIRS.

Storage: B3 files are maintained in: \$ISCCPDIR/b3/YYYY where YYYY denotes subdirectories by calendar year.

Provider & Format: ISCCP B3 data are produced by GISS. ISCCP B3 data are obtained in their native B3 format. The ISCCP website is: <http://isccp.giss.nasa.gov/> and the B3 format is documented at: <http://isccp.giss.nasa.gov/docs/documents.html>.

Timing: The ISCCP B3 data are generally released with a 6 to 12-month lag. However, this can be longer since ISCCP will wait for a 1-2 year period to fully determine the calibration of a new instrument (e.g., a new Geostationary or reference polar orbiter).

File: ISCCP BT

Purpose: These data are required by GridSat to provide ISCCP calibration data to be used in pre-calibrating and producing ISCCP B1U data.

Storage: BT files are maintained in: \$ISCCPDIR/bt/YYYY where YYYY denotes subdirectories by calendar year.

Provider & Format: ISCCP BT data are produced by GISS. ISCCP BT data are obtained in their native BT format. The ISCCP website is: <http://isccp.giss.nasa.gov/> and the BT format is documented at: <http://isccp.giss.nasa.gov/docs/documents.html>.

Timing: The ISCCP BT data are for corrections to the calibration tables in the original ISCCP B3 files. Therefore, they are only released when problems are discovered with the B3 calibration. These files need to be incorporated in GridSat processing as they become available, but processing should not wait for their provision/creation.

File: ISCCP B1

Purpose: These data are required by GridSat to provide the GEO data used in the GridSat-B1 data.

Storage: B1 files are maintained in: \$ISCCPDIR/b1/YYYY where YYYY denotes subdirectories by calendar year. Realtime ingest of B1 files are maintained by the Archive Branch. As they ingest files, they also copy B1 files to the current server (glaze). These files are stored in \$ISCCPDIR/b1-rt and are generally moved to \$ISCCPDIR/b1/YYYY at the end of the calendar year.

Providers & Format: The ISCCP B1 data are provided by numerous ISCCP Satellite Processing Centers (SPCs), each with their own satellite data format. This data resides on the NCDC HDSS. Current providers include:

- NOAA/NCDC – GOES-East and West positions
- EUMETSAT – Meteosat Europe/Africa and Indian Ocean positions
- China (CMA) – Chinese FY2-C/E position
- Japan (JMA) – Japan/Australian position

Timing: ISCCP B1 providers routinely send data on a daily basis. The sole exception is NCDC, which waits 10 days before sending the data. Thus, NCDC data is often a few days behind.

File: HIRS AS Pathfinder

Purpose: These data are required by GridSat to provide the reference calibration data set for the GEO IR window brightness temperatures.

Storage: Older data were provided by Daren Jackson. Current data are provided by Lei Shi (NCDC). Because the files are outside the ISCCP processing, the files (or symbolic links) do not reside in the ISCCPDIR. HIRS data is currently located in:

glaze:/home/kknapp/DATA/hirs/path/as.

Provider & Format: Lei Shi (NCDC) provides the HIRS Pathfinder All-sky (AS) data. She obtains HIRS Level 1B data from NCDC, processes the data using AAPP into Level 1C data and further processed 1C data to Pathfinder All Sky (AS) data. The format of the data is a flat binary file. Specific details of the format can be determined from the FORTRAN algorithm:

~/for/hirs_b1u/ b1u_HIRS-path-AS_matchup.f90.

Timing: Lei Shi produces these files shortly after the end of the year (generally are available by January). She may sometime set up scripts to have data produced regularly (e.g., daily, monthly, etc).

File: Local ISCCP B3

Format: ISCCP B3 data retain their native format locally. The files are stored in subdirectories by year.

Directory: ISCCPDIR/b3/

File: Local ISCCP BT

Format: ISCCP BT data retain their native format locally. The files are stored in subdirectories by year.

Directory: ISCCPDIR/bt/

H:\documents\projects\GriSat\GridSat Cookbook.doc

File: Local ISCCP B1

Format: ISCCP B1 retain their native format locally. The files are stored in subdirectories by year.

Directory: ISCCPDIR/b1/

File: Local HIRS AS Pathfinder

Summary: Lei Shi generally stores data on glaze. Thus, files are stored in my data directory with symbolic links to her directories.

Format: There is no change to the original data, so formats are the same.

Directory: glaze:/home/kknapp/DATA/hirs/path/as/

File: ISCCP B1U

Purpose: ISCCP B1U files are required because there is no Fortran routine which can read all the various B1 formats. Thus, IDL is used to reformat the data to ISCCP B1U.

Summary: The ISCCP B1U format is an attempt to unify the various B1 formats and navigation styles. It really is an intermediate step since the final reformat of B1U should use the NetCDF (version 3 or 4) format.

Format: Currently, ISCCP B1U are in flat binary format file.

Directory: ISCCPDIR/b1u/

File: ISCCP B1U QC

Purpose: To provide B1U quality information to remove bad scans from the resulting GridSat data.

Summary: These files are produced by a Fortran algorithm located in the *kmod_iscpp.f90* file.

Format: The files are flat binary files with QC information at the scan line level (each scan line has a one flag per channel).

Directory: glaze:/home/raid/kknapp/data/iscpp/b1/

File: GEO/POES match-up files

Purpose: These files provide a means to simply determine temporal match-ups between GEO and LEO data sets.

Summary: There is one GEO/POES match-up file for each synoptic time (00, 03, 06, ..., 21 UTC). All ISCCP B1U files are listed along with any POES data which are within 30 minutes of the synoptic time. The data sets inventoried include: ISCCP B1U, ISCCP B2/AVHRR, HIRS Level 1C (obsolete), HIRS Pathfinder All-Sky, and SSMI SDR files.

Format: These are ASCII files. The order of the data sets remains the same. Essentially, the format is: (number of files for this data set, list of files) x 5. The files are stored using the automount nfs names of the directories such that files can be accessed from any computer in the RSAD cluster.

H:\documents\projects\GriSat\GridSat Cookbook.doc

Directory: glaze:/home/raid/kknapp/data/isccp/match-poes/

File: GEO/HIRS Calibration Match-up Files

Purpose: To store the co-located observations of GEO and HIRS observations for further analysis by other software.

Summary: These are intermediate files produced by Step I which are further analyzed by Steps J and L. These allow quick reanalysis of the calibration without rerunning the full set of matchups. The data are produced by a Fortran90 file. The data are rather large, proportional to the yearly volume of B1 and HIRS data. The yearly sizes are 3, 10, 30 GB for the smallest, average and largest years, respectively. The total size for 1978-2009 is 287 GB.

Format: The format is a flat binary file of variable length and non-intuitive design. The format can be gathered either from the file which produced the data (*~/for/hirs_b1u/ b1u_HIRS-path-AS_matchup.f90*) or the file which reads the data (*~/idl/isccp/calib/ hirs_b1u_month.pro*).

Filename convention: Nnn.GEO.YYYY.MM.DD.HHMM.SPC (see Appendix for abbreviations).

Directory: glaze:/home/raid/kknapp/data/isccp/b1ucal/matchup/ (in subdirectories by year)

File: IRWIN Monthly Statistical Match-up output

Purpose: To provide an intermediate calibration analysis step.

Summary: Because the GEO/HIRS calibration matchup files are so voluminous, the intermediate monthly analysis is stored in these files. The primary file is the stat file which stores the results for each channel (along with some information about the matchup criteria used). These files are then used by Step K to calculate and produce the final calibration table. Also, latlon files are produced which can be used to spatially locate the matchups to analyze geographic distribution (the latlon files are not really necessary in the production of GridSat data). The files are stored in various subdirectories. The *Wu*, *Proportionate* and *Corrected* subdirectories contain data used in the *JTech* publication. GridSat B1 v01r01 uses *b1ucal/monthly/2010-06*

Format: The latlon format is arbitrary flat binary and can be quickly determined from the production file (*~/idl/isccp/calib/cal_b1u_monthly.pro*). The stat files are ASCII files that provide monthly channel data in rows with various statistics by column. The last line in the file provides information about the matchup criteria used to produce the results.

Filename convention: Nnn.GEO.YYYY.MM.stat, Nnn.GEO.YYYY.MM.latlon.c

Directory: glaze:/home/raid/kknapp/data/isccp/b1ucal/monthly/

File: IRWIN Calibration Table

Purpose: To provide a final, versioned calibration table from which corrections may be applied.

Summary: The IRWIN calibration table provides the final calibration adjustment via a slope and offset. The data are produced by *~/idl/isccp/calib/cal_b1u_correct.pro* and used by various algorithms to either analyze the final calibration or produce GridSat netCDF files.

H:\documents\projects\GriSat\GridSat Cookbook.doc

Format: The file format is quite simplistic. Each row represents a separate month of geostationary data. For each month, statistics of the correction are available along with the correction.

Numbering: Each iteration of the calibration increases the version number by two. It is desired that even numbers represent intermediate results during processing and odd numbers are a final release at some final stage of GridSat processing.

Naming convention: b1u.calibration.correction.CHAN.vVV (some early versions have the CHAN missing, thus represent solely the IRWIN calibration).

Directory: glazg:~/idl/isccp/calib/

FINAL PRODUCT: File: GridSat-B1 netCDF files

Purpose: To provide small, inter-operable, observations of weather and climate systems with the least temporal bias.

Summary: This final file represents the data set archived at NCDC and denoted by Data Set Index (DSI) 3648. Documentation on the data is available at: [TBD](#) and the dataset home page is: <http://www.ncdc.noaa.gov/oa/gridsat/>.

Format: The data are stored in netCDF format (<http://www.unidata.ucar.edu/software/netcdf/>) using the CF convention (<http://www.cfconventions.org/>). File metadata are stored in the Unidata Discovery format (<http://www.unidata.ucar.edu/software/netcdf-java/formats/DataDiscoveryAttConvention.html>).

Naming convention: See data set documentation.

Directory: glazg:/home/kknapp/dataset/gridsat-irwin/YYYY/

Steps

Step A: Obtaining ISCCP B3

The ISCCP Central Archive is NCDC, so ISCCP B3 data are obtained from the HAS (HDSS Access System) at NCDC.

Time: Since a large amount of the B3 is already online, updating the data with the most recent from HDSS/HAS takes < 1 week.

Step B: Obtaining ISCCP BT

ISCCP BT data are obtained from the ISCCP ftp site: ftp://isccp.giss.nasa.gov. Structure:

- pub/bt – Old BT files from 1996 (Currently no useful information for GridSat)
- pub/calibration – empty (Currently no useful information for GridSat)
- pub/data/BT – Check this directory for new BT files (i.e., new since the last GridSat update).

mirror_bt.sb mirrors BT directory to the local disk.

Time: Since a large amount of the BT is already online, updating the data with the most recent from FTP takes < 1 day.

Step C Obtaining ISCCP B1

Summary: Because of the data volume, it is easier to have data delivered on a daily basis, rather than yearly obtaining the entire year of data. The data are delivered to glaze by Robert Summers (RSAD/AB) via an automated script that copies data there as it is received on the FTP ingest computer. The data from Robert are stored in *glaze:/home/raid/kknapp/data/isccp/b1-rt* and prior to processing are moved to *glaze:/home/raid/kknapp/data/isccp/b1/YYYY/*.

Time: The ISCCP B1 data can be voluminous. Obtaining more than one year from HAS (through Dave Bowman) can be quite time consuming. The entire record would take more than 1 week to obtain.

Step D: Obtaining HIRS Pathfinder AS

Shi data summary: Lei Shi provides HIRS all sky data on the RSAD server *crystal* in her directories. *b1u_match_poes.pro* now points to this directory.

Jackson data summary: The earliest files were obtained via FTP from Darren Jackson. Due to the volume of HIRS AS data, Darren provided the data in intervals on FTP to be downloaded. The files were downloaded using FTP (and the script *getall*).

Knapp directories: *glaze:/raid1c/nb/kknapp/data/hips/path/as*

Shi directories: *crystal:/net/crystal/san3/hips/hipsas/*

Step E: Creating ISCCP B1U

Summary: This is perhaps one of the more complicated processes in all the GridSat processing, and most likely deserves its own documentation. In lieu of such, a summary is provided here.

Algorithm: The routine *update_b1u_files.pro* produces a list of B1 files which need to be reconverted to B1U. Circumstances when a file should be reconverted include:

- B1U file is missing [this could be problematic if the B1U file is missing because the B1 file read failed]
- Calibration file in B1U was missing, but now one exists (particularly for recent years which ISCCP is in the process of updating)
- Calibration file is different (in the case of an updated ISCCP calibration)
- Calibration file was modified more recently than the B1U file (in the chance that the file was changed without changing the version).

The data are converted to B1U using the IDL algorithm: *~/idl/isccp/b1u/b1_2_b1u.pro*. This file is often driven by the routine *convert_b12b1u_year.pro*. In short, once the input files are in place, the read files are checked for accuracy (particularly after format changes), then this file can be run to produce ISCCP B1U data.

Reading B1 data: As described, there are numerous ISCCP B1 formats and just as many algorithms to read the data. In short, the *idl/isccp/read/b1read.pro* algorithm is an umbrella algorithm which provides the generic interface to ISCCP B1 files. The user inputs the filename to read and a few output options (in the form of keywords) can be selected. The primary output is a *b1u* data structure which is then used to create the *b1u* file (and is subsequently read in using

the *b1uread* algorithms). The *b1read.pro* calls a variety of read algorithms: *b1read_SPC_SAT.pro*. Currently, there are no less than 14 of such files. The SPC refers to the Satellite Processing Center (a.k.a., the data provider). SAT refers to the satellites served by this particular reader. For instance, *b1read_eum_met.pro* reads all EUMETSAT Meteosat B1 data. However, for CSU, there are 5 readers: *b1read_csu_g6.pro*, *b1read_csu_g7.pro*, *b1read_csu_g8.pro*, *b1read_csu_g91012.pro* and *b1read_csu_met3.pro*. The individual read files read in the data in various formats and, for these applications, return the data in the *b1u* data structure.

Calibration: Calibration information is obtained in the following order of precedence of BT first, then B3. First, files are checked for the identical date. If no calibration is found, then B3/BT files for the same day are searched. Finding none, the same month is searched. Finding none, nearby months are searched for BT/B3 files with calibration information. At any point in which a calibration file is discovered, the calibration table is retained as the calibration information. If no calibration information is found during the entire search, then nominal calibration tables are used (e.g., for a satellite without ISCCP B3 data such as SMS in 1981). This is all performed by *idl/isccp/read/b1read_calv2.pro*.

Navigation: In the same way that many file formats exist, so do numerous navigation styles also exist. As part of reading the B1 data, the navigation needs to be handled and is output in one of two-styles: GVAR and Kepler. The GVAR style matches that of more recent GOES files which have GVAR navigation headers in MCIDAS format. Because of the exactness of this format, and relative ease in porting it to a *b1u* data structure, the header was maintained. All other satellites have their navigation normalized to a Kepler style navigation (akin to the pre-GVAR navigation). This means that non-Kepler navigation styles (e.g., Meteosat and GMS) were converted by fitting Kepler elements to each navigation module. Also, quality checks were added for files providing navigation information to determine whether the provided Kepler elements were realistic (this reduces the chance of having non-converging navigation algorithms).

Step G: ISCCP B1U QC Algorithm

Summary: The ISCCP B1U algorithm provides a quality control check at the scan line level for the ISCCP B1U data. The test was developed using IDL, but the final algorithm is the *b1u_get_qc* subroutine in the *for/kmod_isccp.f90* module. The module attempts to read the QC file. If found, it returns the QC information, if not found it calculates the QC, saves it to the QC file and returns the calculated flags. The QC files are only automatically produced if another Fortran routine attempts to read the QC data. However, since the next step in GridSat processing is the *b1u_2_gridsat.pro*, then this needs to be performed.

Code: The *b1qc* algorithm can be launched using the *b1qc* executable. The source merely acts as a driver which calls the *b1u_get_qc* subroutine in the *kmod_isccp* module.

Algorithm: The algorithm performs 4 tests. As tests are tripped, bits are set to one. Therefore, scans with QC=0 means passing all tests. When QC \neq 0, then the failed tests can be determined from the bit values. The four tests are:

1. Test for dropped scans (no variation of data in the scan)
2. Test for faint or noisy lines (based on mean and standard deviation of nearby lines)

H:\documents\projects\GriSat\GridSat Cookbook.doc

3. Test for duplicate lines (by keeping only the 1st instance of a repeated line)
4. Test temporal consistency (for files with temporal information in the scan header)

Code directory: `glaze:~/for/b1qc/`

Compilation: `glaze:~/for/b1qc/Makefile`

Executing:

```
glaze:~ $ b1qc <B1U_filename>
```

Time: This is rather quick and reprocessing through all B1U files should complete in a weekend. However, this is an estimate based on past experience and not very accurate.

Step H: Writing GEO/POES match-up files

Once the ISCCP B1U files and HIRS files are in place, the match-up files are produced by `~/idl/isccp/match_b1u_poes.pro`. This step is as simple as executing:

```
glaze:~ $ ndl "match_blu_poes"
```

Time: This process takes ~30 minutes to complete, but is only run when new files are available (that is, new files which the process inventories).

Step I: Performing GEO/HIRS Calibration Match-ups

Algorithm: The Fortran routine `~/for/hirs_b1u/b1u_HIRS-path-AS_matchup.f90` performs the calibration match-ups. It reads in a GEO/POES matchup file, then performs matchups for each GEO-POES combination available. The output is the GEO/HIRS matchups output for each GEO file/HIRS combination. The process is slow due to the numerous calls to navigate the GEO files.

Scripting: The match-ups are run by scripts. The script `~/scripts/hirs.run_all_matchups` facilitates this by creating `~/scripts/hirs.runall` which can then be run by troll. An example is:

```
glaze:~/scripts $ hirs.run_all_matchups hirs.runall 2006 2007
```

where the first argument is the output script file, the second is the beginning year and the third is the ending year (if this is omitted, then one year is processed). The output from this script can also be shortened to specifically process targeted times.

Troll: `troll.batch_runner` can be used to speed up the calibration matchups. The output from `~/scripts/hirs.run_all_matchups` is used directly via:

```
glaze:~ $ nohup troll.batch_runner 40 hirs.runall 10 &> ...
```

Time: Without troll, this job would be quite slow. Using troll, one year can be processed in about 1 day. And even faster using the new RSAD servers (with 8 processors each)

Step J: IRWIN Monthly statistics aggregation

Reading calibration matchups: The calibration matchups are read using `~/idl/isccp/calib/hirs_b1u_month.pro` which is called by `~/idl/isccp/calib/cal_b1u_monthly.pro`. The `hirs_b1u_month` routine was quite slow until an upper bound on the array sizes were chosen. This routine needs to be monitored to assure that the upper bound (NMAX) is not being exceeded. In short, if this routine crashes, it is likely caused by NMAX being too small.

Monthly statistics algorithm: The instantaneous matchups are aggregated into monthly statistics. This is performed using `~/idl/isccp/calib/cal_b1u_monthly.pro`. For each satellite month, the algorithm reads in all the matchup files, limits the data according to the Proportionate criteria, then outputs 2 types of files: stat and latlon. Stat files provide one line statistical summaries of the monthly matchups. These are the files used to create the calibration coefficients. The latlon files are merely the locations of the points used in the stat file. The stat and latlon files are output to: `~/DATA/isccp/b1ucal/monthly/`. It is often best, however, to move these files to a subdirectory for processing. Latlon files are not used, but were output as diagnostics to analyze the locations of the matchups.

Listing satellite-month combinations: The `~/leo.geo.date` file provides a complete list of all satellite-month combinations which result from the monthly matchups. This file needs to be re-created as new satellite months are added to the matchup library. To do so:

```
glaze:~ $ hirs.idl_monthly_matches
```

This routine completes within an hour.

Scripting: This monthly statistic process can be scripted (then trolled) using `~/scripts/hirs.cal_monthly_make` and the output file: `~/leo.geo.date`. To do so:

```
glaze:~ $ hirs.cal_monthly_make > hirs.cal_monthly.sh
```

(completes in about 12 hours without troll) or using troll:

```
glaze:~ $ nohup troll.batch_runner 8 hirs.cal_monthly.sh 2 &> ...
```

The output script is then run using troll (although, do not select too many parallel runs (here, I chose 8) since the process is largely I/O bound). Also, since fixing `hirs_b1u_month.pro`, the process completes 1 year of matchups in about a day, so it may not be as necessary to use troll.

Time: ~ 1 day to process one year of data.

Step K: IRWIN calibration coefficient calculation

This process is really two steps. First the stat, output files from step J are aggregated and modified. This is then used to calculate the calibration coefficients.

Statistic aggregation: This is accomplished with command line scripts which might eventually be automated. All the following commands are executed from the subdirectory in `~/DATA/isccp/b1ucal/monthly/` to which the data has been moved (e.g., *Proportionate-Update/*)

1. Collect the output statistics for the IRWIN channel:

```
$ cat *stat | gawk '{ if ($2 == 0) print}' > IRWIN.cal
$ cp IRWIN.cal IRWIN.cal.edit
```

2. Edit `IRWIN.cal.edit` and add spaces for columns which have no space between them. This most often happens for the Chi column.

3. Edit `IRWIN.cal.edit` and replace all “GOES9” with “GOE-9”

3. Modify `~/scripts/hirs.fix_cal_file` to point to the appropriate subdirectory and `IRWIN.cal.edit` file. Then run the script. This adds rows to the file for satellite-months which need calibration coefficients, but no matchups exist. This causes the `cal_b1u_correct.pro` routine to add data for these satellite-months by temporal interpolation. The output is `IRWIN.cal.edit.fix`.

H:\documents\projects\GriSat\GridSat Cookbook.doc

4. The data then needs to be sorted by year/month:

```
$ cat IRWIN.cal.edit.fix | sort -k 2 -t . > IRWIN.cal.edit.sorted.fix
```

Coefficient calculation: The IDL algorithm `~/idl/isccp/calib/cal_b1u_correct.pro` uses `IRWIN.cal...` to determine the calibration coefficients. The code should be checked and changed to ensure:

- Correct channel is used: ‘IRWIN’ or ‘IRWVP’
- Correct subdirectory is used (e.g., ‘Proportionate-Update’)
- Correct version number is set (e.g., ‘v6’) such that older tables are not overwritten.

Once these are checked, execute

```
IDL> cal_blu_correct,/showold
```

The calibration table is created and monthly time series plots are output. Analyze time series plots (in `~/work/pic/`) to look for potential errors in the calibration process.

Time: Completes within 1 hour.

Step N: Creating GridSat-B1 data

The GridSat-B1 files are created by `~/idl/isccp/b1u/b1u_2_grisat.pro`. Routine inputs and keywords are used to determine how it is run.

Code changes:

1. When a significantly new version is required, increment variable ‘hursatver’
2. When new calibration tables become available, change variable ‘cfile’

Summary: For each geostationary observation for this time, it is remapped and then stored in the gridded set of variables.

The routine can be run directly from IDL for a given year: e.g.,

```
IDL> blu_2_grisat,1978
```

Or the script `~/scripts/grisat.make_batch.sh` can output `blu_2_grisat` calls using individual files. Just modify the file to match the required years and the output is sent to a new batch file.

Then, this can be launched using troll via:

```
nohup troll.batch_runner 15 grisat.1980-2008.sh 1
```

which takes quite a while.

The grisat data files are placed in:

```
glaze: ~/datasan/grisat/YYYY/
```

These files contain the entire dataset (all variables and channels, etc.). For the CDR project, a subset of this file was created using netCDF operators. A script that retains only the CDR variables was run: `gridsat.make_cdr_files.sh`

This file runs NCO’s ncks command:

```
ncks -4 -L 6 -a -v irwin,satlat,satlon,satrad,satname,b1ufile,calslp_irwin,caloff_irwin,satid_ir
```

H:\documents\projects\GriSat\GridSat Cookbook.doc

This converts files to netCDF4 (-4), compresses variables (-L 6) and only processes certain variables (-v). The data from this step are then sent to *~/datasan/gridsat-irwin/YYYY*.

These 'gridsat-irwin' variables were then archived by the CDR project.

These files were tarred up using: *gridsat.make_cdr_tarballs.sh*

Optional Paths

Step F: ISCCP B1U Centering Algorithm

Purpose: To correct navigation errors based on center pixel and scan which can easily be detected using IR full disk imagery.

Optional: This path is optional because the corrected navigation has not been fully tested on all satellites. The algorithm can be applied to all satellites, but appears to have strange results when navigating, for example, GMS files. The original algorithm was developed for early GOES imagery (GOES-7 and earlier) but should be tested for all other satellites.

Algorithm: The file *idl/isccp/b1u/b1u_finalcenter.pro* runs the internal routine *b1u_get_center* which calculates the corrected center scan and pixel. It does so by processing the IR image (which is present in all ISCCP B1 files). The center pixel can be determined as long as the eastern and western limbs of the Earth are present. If so, a polynomial is fit to the list of center pixels (particularly because the center pixel drifted on early spin-scan radiometers). The algorithm then determines the center scan line if a portion of both hemispheres (northern and southern) are visible. Again, this algorithm needs to be more fully tested before being applied to GridSat processing.

File: ISCCP B1U Center

Purpose: To allow a navigation correction based on an adjusted center scan line and center pixel.

Format: Flat binary file. The format can be gathered from the creating file (*idl/isccp/b1u/b1u_finalcenter.pro*) or the read algorithm (subroutine *b1u_get_center* in *kmod_isccp.f90*).

Naming convention: <ISCCP_B1U>.center

Directory: Same as ISCCP B1U file.

Step L: IRWVP Monthly Statistics Aggregation

Purpose: To provide recalibrated IRWVP data as needed.

Optional: This step is optional since the GridSat-B1 data does not require accurate water vapor calibration. However, it is required, for instance, when processing GridSat-Basin data with calibrated IR water vapor output.

Algorithm: Currently, the recalibration is based on N-17 data comparisons. This recalibration requires an extra step since the correction is non-linear. In short, the routine *cal_b1u_monthly_wv.pro* replaces *cal_b1u_monthly.pro*. It creates matchups (the stat and latlon files) and also outputs the actual brightness temperatures (in Kelvin) matchups to a 'dat' file.

File: IRWVP Monthly Statistical Match-up output

See algorithm description of Step M.

Step M: IRWVP Calibration Coefficient Calculation

Summary: The IRWVP Coefficients are calculated using: *derive_hirs_geo_correction_empirical.pro* using the new 'dat' files. So far, the coefficients are sent only to the screen. So the data are captured and pasted into the final calibration table.

File: IRWVP Calibration Table

Summary: This is the water vapor equivalent to the IRWIN calibration table. The primary difference is that the number of coefficients increases from 2 (a linear correction for IRWIN) to 4 (a non-linear, cosine-based correction). Again, each satellite-month has different calibration coefficients.

Naming convention: Identical to the IRWIN Calibration Table. The latest file is *b1u.calibration.correction.IRWVP.beta* which was produced by copying the screen output from *derive_hirs_geo_correction_empirical.pro* and saving it to this file.

Appendix A: Subversion Files

Files related to the creation of GridSat data are stored in subversion. Via a directory in Knapp's home directory (*~/project_directories/grisat/*), files are synced from separate subversion repositories and copied to gridsat software releases using `gridsat.make_source_tag.sh`

This file uses the files linked in the *~/project_directories/grisat/src/* directory to create source releases. The current tag is available at:

https://conman.ncdc.noaa.gov/svn-repos/gridsat/source/tags/source_v01r01/

Appendix B: File name conventions

Title	Summary	Examples
Nnn	POES satellite number	N16, N17, ...
GEO	GOES Satellite	MET-5, MTS-1, GOE-5, ...
YYYY	Year	1980
MM	Month (01-12)	01, 02, 03, ..., 12
DD	Day of Month (01-31)	01, 02, 03, ..., 31
HHMM	Time of Day (UTC)	0300, 0600, 1130, ...
SPC	Satellite Processing Center	EUM, JMA
CHAN	Geostationary channel	IRWVP or IRWIN
vVV	Version identification	beta, v01, v02, ...

File: GEO/HIRS Calibration Match-up Files

Filename convention: Nnn.GEO.YYYY.MM.DD.HHMM.SPC

File: IRWIN Monthly Statistical Match-up output

Filename convention: Nnn.GEO.YYYY.MM.stat, Nnn.GEO.YYYY.MM.latlon.c

File: IRWIN Calibration Table

Naming convention: b1u.calibration.correction.CHAN.vVV