THEMIS
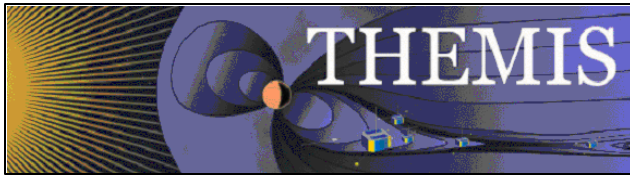# GBO All-Sky Imager L1 Data files Variable Name Definition

THM-SOC-109
January 8, 2007

Timothy Quinn, THEMIS Science Operations

Dr. B. Jackel, THEMIS ASIs

Dr. H. Frey, THEMIS ASIs

Vassilis Angelopoulos, THEMIS PI

Document Revision Record

Rev.	Date	Description of Change	Approved By
1	2006-09-22	Draft	-
2		Signature Version	Signatories
3	2006-09-27		
4	2006-09-28	Lower case variable names	
5	2006-10-04	Final def. of stations, keogram file	
6	2006-10-05	Removed images for data var., include end time	
7	2006-10-16	Removed names, exposures, updated SKTEditor, map added	
8	2006-10-20	Renamed Row/column	
9	2007-01-08	Corrected typo FSMP to FSIM	

Distribution List

Name	Email
Jim Lewis, U.C. Berkeley	jwl@ssl.berkeley.edu
Stuart Harris	sharris@ssl.berkeley.edu
Dr. Eric Donovan, U of Calgary	eric@phys.ucalgary.ca
Dr. Stephen Mende, UCB	mende@ssl.berkeley.edu
Dr. Dave Sibeck, NASA GSFC	david.g.sibeck@nasa.gov

TBD List
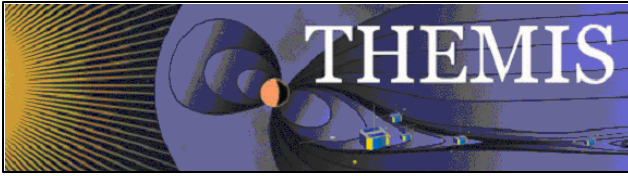
Identifier	Description

Table of Contents
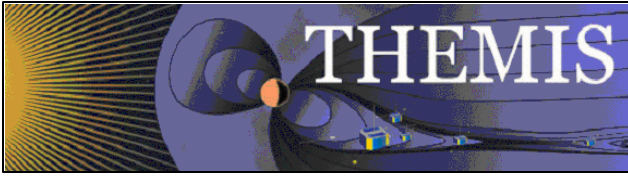
1. Introduction

1.1 *Purpose and Scope.*

THEMIS Level 1 ASI data shall be CDF files. The THEMIS team shall provide to the community additional calibration files and plotting tools that read these files and plot the data in a scientifically useful way (produce physical quantities and plots) and integrate analysis and plotting routines along or with the rest of THEMIS data products (GMAGs and spacecraft). A derivative of these L1 data files shall be created as L2 files and be utilized for SPDF, NSSDC and CDAWeb use, i.e., for further distribution and plotting using nominal CDAWeb capabilities.

THEMIS ASI data are two types:

Daily transmission data are those data that are transmitted to UCB by internet via a main node at the University of Calgary. Those data consist of: Thumbnails, keograms, and hourly full resolution checkout data. The latter two are jpeg images, which are treated as housekeeping data to ensure station-health and are not maintained in CDF format. They are available for viewing on the web in their native jpeg format. The daily and hourly keograms (32 or 256 pixel vertical scans at 1min or 6s cadence) and the thumbnails (at 3s or 6s cadence) are the highest temporal and depth resolution (8 or 16 digitization bits) possible given the available station bandwidth.

Mailed disk data are high-resolution data stored on a local hard disk at the site, which are retrieved by local custodians performing periodically (1-3 months) a disk extraction and shipment. Those data are transmitted to UCB by mail, after they have been downloaded on a server at the central retrieval, archival and redistribution node at the University of Calgary. Those data consist of high spatial resolution images (256x256 pixels) at the highest time resolution possible by the site hardware (3s).

All CDF files are processed at UCB and University of Calgary independently but using common processing code, in order to adhere to the same format and file structure described herein. The same structure is intended for use by future ancillary datasets, as they become available, such that analysis can proceed seamlessly for those stations as well.

The purpose of this document is to define the ASI variable names within the L1 data CDFs.

1.2 *Applicable Documents.*

1. THM_SYS_012_PDMP   THEMIS Project Data Management Plan
2. THM_SOC_101_TIME   THEMIS TIME Definition
3. THM_SOC_108_GMAG_L1_VARNAMES   THEMIS GMAG Variable Name Def's
4. THM_SOC_117_ASI_CALPROC   THEMIS ASI calibration definition

2. General Level 1 Data File Variable Naming Conventions

2.1 *VARNAMES: thg_asx, or thg_asx_ssss; x Î {t, k, f}*

Following an overall principle of naming THEMIS variables using more general to less general descriptors, separated by the underscore character "_", the ASI variables shall be named in lower case like: thg_asx_ssss, where "g" denotes ground based observatory, the "as" denotes All-Sky Imager and ssss is the 4-letter identification of the station (**Table 1**). The character "x" identifies the data type and resolution. x ∈
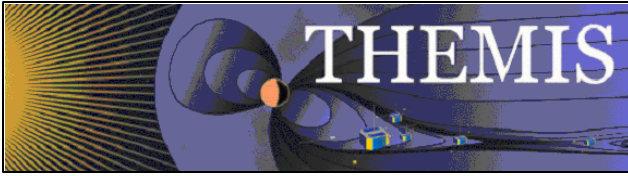
{t, k, f} for {thumbnails, keograms, full image}, respectively. Time and Epoch shall accompany the quantities as specified in THM_SOC_101_TIME.DOC. Figure 1 is a snapshot of a master-cdf for the ast variable. Variables are given in **Table 2.**

Figure 1 SKTEditor snapshot of the required quantities for a single station L1 data ast variable CDF file.
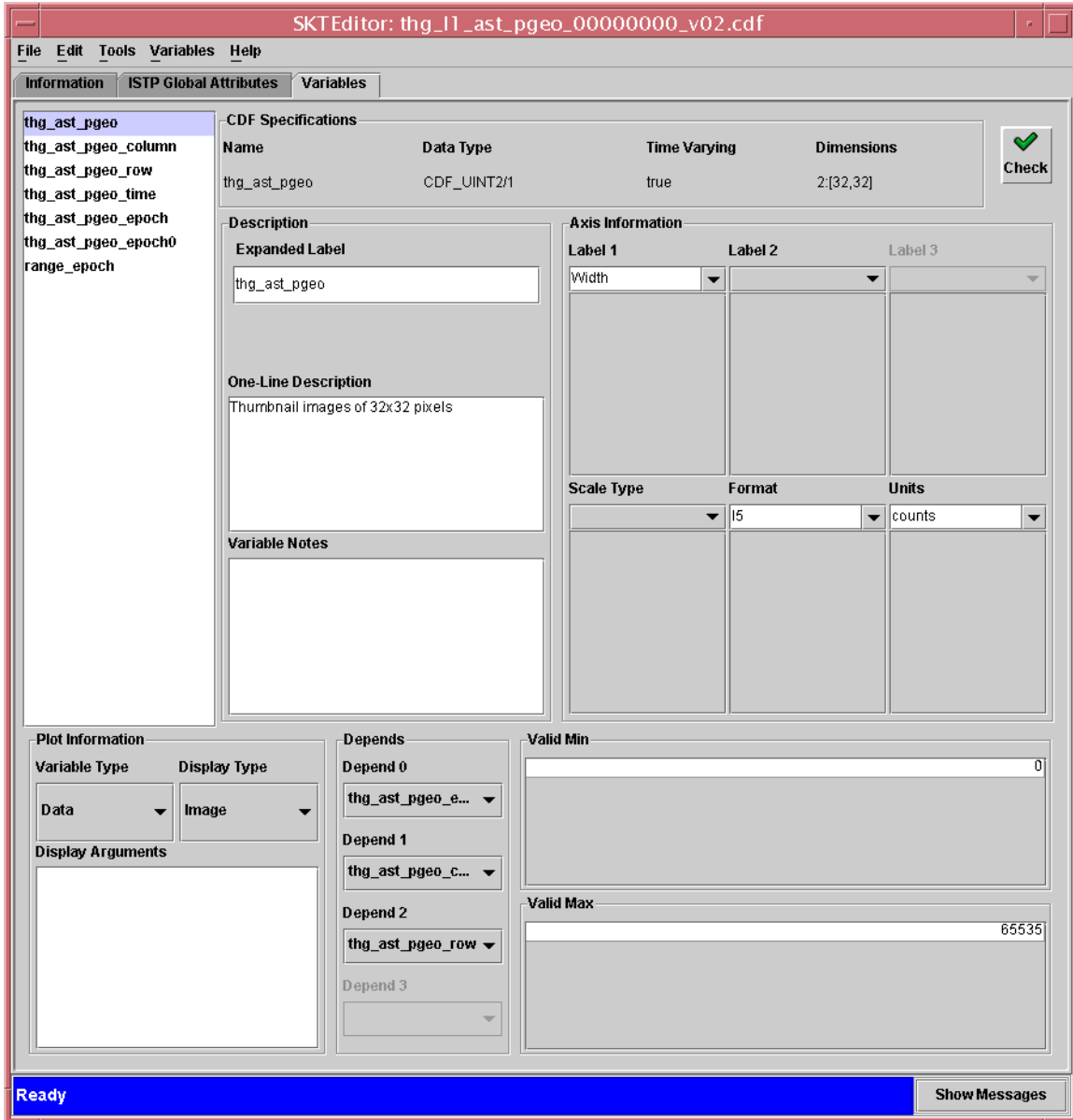
This construct works well with all data from the GBO stations. When further specificity is required in order to ensure optimal file size, such as, e.g., image depth, additional descriptors _yy may be utilized. No additional descriptors are envisioned for ASIs at the time of writing this document.

Ancillary ASI data in CDF format for use by THEMIS shall adhere to the same format with the identifier "o" for "other". E.g., tho_asx_ssss denotes ancillary ASI data from station ssss.
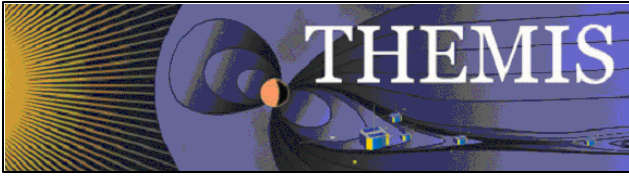
Filenames shall be lowercase derivatives of variable names with the insertion of "l1" characters, the date (and start hr for hourly files, as necessary) and the version number. E.g., the thg_asf_ssss images for hh:00-hh:59.9 UT shall be in thg_l1_asf_ssss_yyyymmdd_hh_v01.cdf (**Table 3**).

Table 1: THEMIS GBO station names and abbreviations starting in west Alaska towards east Canada

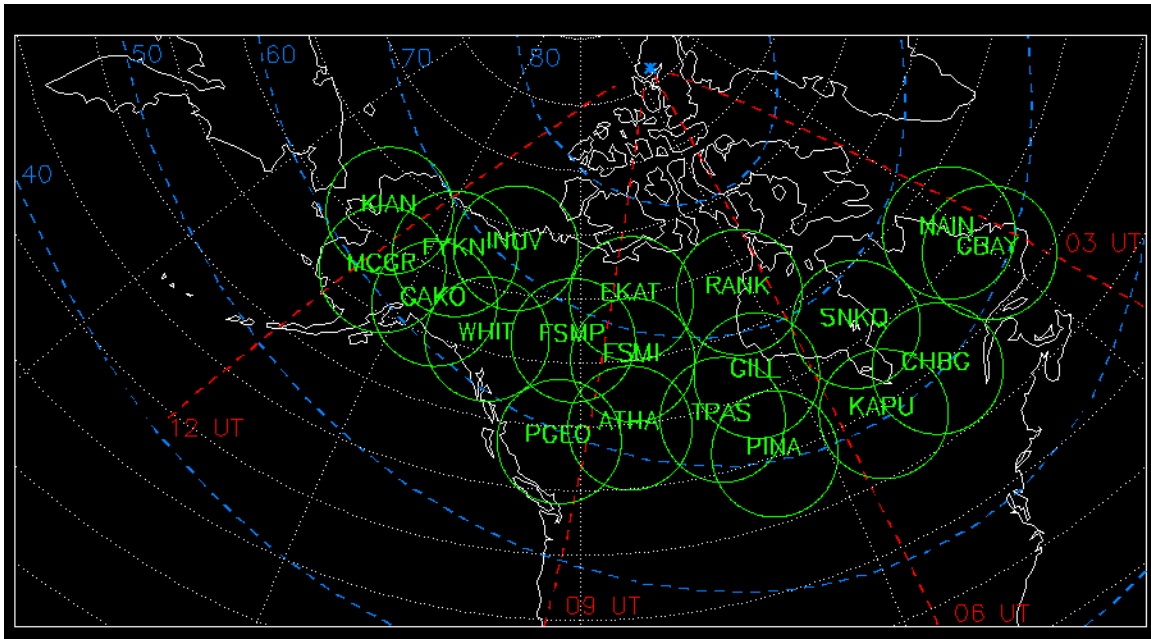|	
Kiana	kian
McGrath	mcgr
Fort Yukon	fykn
Gakona	gako
Inuvik	inuv
Whitehorse	whit
Fort Simpson	fsim
Prince George	pgeo
Ekati	ekat
Fort Smith	fsmi
Athabasca	atha
The Pas	tpas
Rankin Inlet	rank
Gillam	gill
Pinawa	pina
Kapuskasing	kapu
Sanikiluaq	snkq
Chibougamau	chgb
Nain	nain
Goose Bay	gbay

Figure 2: Map of north America with the GBO station names and their approximate fields of view of the all-sky cameras. Magnetic latitudes are shown and labeled in blue. The red lines represent the meridians of local magnetic midnight at 03, 06, 09, and 12 UT.

Table 2. ASI data products and names. All ASI data will be transformed into 16 bits before they are written to level-1 data files.

	L1 VARNAME	Width	Frame Height	Size	Cadence	Sites	Hour	Day Ave	Max	Year 20 sites
		Pixel	Pixel	Bytes	frames/ minute	Per file	Mega Bytes	Mega Bytes	Mega Bytes	(GB)
High T-Res Keograms	thg_ask	1	256	512	20	20	12.3	98.3	295	700
High T-Res Thumbnails	thg_ast_ssss	32	32	2048	10	1	1.23	9.8	29.5	1391
High T-Res Full frames	thg_asf_ssss	256	256	131072	20	1	150	1200	3600	8554

Table 3. ASI file names

	L1 VARNAME	L1 FILENAME
High Time Res Keograms	thg_ask	thg_l1_ask_yyyymmdd_vnn.cdf
High Time Res Thumbnails	thg_ast_ssss	thg_l1_ast_ssss_yyyymmdd_vnn.cdf
High Time Res Full frames	thg_asf_ssss	thg_l1_asf_ssss_yyyymmdd_hh_vnn.cdf

3. Variable names and definitions

3.1 thg_ast data file

This file will contain one day of thumbnail images in their highest time resolution from one station. It can be created about 24-48 hours after data collection. The variables are ISTP-compliant and their names, sizes, and definitions are given in Table 4.

Table 4: Definition of variables for thg_ast files

thg_ast_ssss	UINT2, 32x32x(exposures) array	Individual image frames
thg_ast_ssss_column	UINT2, 32 elements	Column (x-dim, Width) in image
thg_ast_ssss_row	UINT2, 32 elements	Row (y-dim, Height) in image
thg_ast_ssss_time	REAL8, as many as exposures	Exposure start time of frames in seconds since 1970-01-01/00:00:00.000
thg_ast_ssss_epoch	EPOCH	Epoch for exposure start time, virtual variable
thg_ast_ssss_epoch0	EPOCH	EPOCH value of time baseline 1970-01-01/00:00:00.000
range_epoch	EPOCH, 2 elements	Epoch values of first and last data record

3.2 thg_ask data file

This file will contain one day of keogram stripes in their highest time resolution from all available stations. It will be created only after the locally stored data are shipped weeks to months after data collection. The

variables are ISTP-compliant and their names, sizes, and definitions are given in Table 5. All variables will be repeated as many times as stations were operating with ssss being different for each station.

Table 5: Definition of variables for thg_ask files. Every variable with ssss will be repeated for as many stations as were operating.

thg_ask_ssss	UINT2, 1x256x(exposures) array	Individual keogram stripes
thg_ask_stations	BYTE	Number of stations that have provided data for this file
thg_ask_names	CHAR, as many as stations	4-letter Station names
thg_ask_ssss_column	UINT2, 1 elements	Column (x-dim, width) in image
thg_ask_ssss_row	UINT2, 256 elements	Row (y-dim, height) in image
thg_ask_ssss_time	REAL8, as many as exposures	Exposure start time of frames in seconds since 1970-01-01/00:00:00.000
thg_ask_ssss_tend	REAL8, as many as exposures	Exposure end time of frames in seconds since 1970-01-01/00:00:00.000
thg_ask_ssss_epoch	EPOCH	Epoch for exposure start times, virtual variable
thg_ask_ssss_epoch0	EPOCH	EPOCH value of time baseline 1970-01-01/00:00:00.000
range_epoch	EPOCH, 2 elements	Epoch values of first and last data record

3.3 thg_asf data file

This file will contain one hour of full spatial resolution images in their highest time resolution from one station. It can be created only after the locally stored data are shipped weeks to months after data collection. The variables are ISTP-compliant and their names, sizes, and definitions are given in Table 6.

Table 6: Definition of variables for thg_asf files

thg_asf_ssss	UINT2, 256x256x(exposures) array	Individual image frames
thg_asf_ssss_column	UINT2, 256 elements	Column (x-dim, width) in image
thg_asf_ssss_row	UINT2, 256 elements	Row (y-dim, height) in image
thg_asf_ssss_time	REAL8, as many as exposures	Exposure start time of frames in seconds since 1970-01-01/00:00:00.000
thg_asf_ssss_tend	REAL8, as many as exposures	Exposure end time of frameS in seconds since 1970-01-01/00:00:00.000
thg_asf_ssss_epoch	EPOCH	Epoch for exposure start times, virtual variable
thg_asf_ssss_epoch0	EPOCH	EPOCH value of time baseline 1970-01-01/00:00:00.000
range_epoch	EPOCH, 2 elements	Epoch values of first and last data record