Scientific data stewardship of International Satellite Cloud Climatology Project B1 global geostationary observations

Kenneth R. Knapp
NOAA's National Climatic Data Center, 151 Patton Ave,
Asheville, North Carolina, 28801, USA
Ken.Knapp@noaa.gov

Abstract. The International Satellite Cloud Climatology Project (ISCCP) B1 data was recently rescued by NOAA's National Climatic Data Center (NCDC). ISCCP B1 data are geostationary imagery from satellites worldwide which are subsampled to 10 km and 3 hourly resolution. These data were unusable given the disarray of format documentation and lack of software for reading the data files. After developing access software, assessing data quality, and removing infrared window calibration biases, the data have been used for research in studying tropical cyclones and is available for other topics, such as rainfall and cloud cover. This resulted not only in valuable scientific data for weather and climate research but also in important lessons learned for future archiving of scientific data records. The effort also exemplifies principles of scientific data stewardship.

Keywords: satellite, geostationary, data stewardship, calibration, navigation

1 INTRODUCTION

Satellites have recorded radiances reflected from and emitted by the Earth for more than 50 years from low-Earth orbit and 40 years from geostationary orbit. In particular, the Polar Operational Environmental Satellite (POES) and Geostationary Operational Environmental Satellite (GOES) operated by the U.S. National Oceanic and Atmospheric Administration (NOAA) have long histories of observations from inter-consistent sensors. While the initial goal of these satellites was to observe atmosphere, ocean and land processes, their extensive record provides an opportunity to analyze the Earth's climate as well.

In 1983, the International Satellite Cloud Climatology Project (ISCCP) began collecting satellite data from the NOAA POES satellites as well as various meteorological geostationary satellites around the world in an effort to characterize global cloudiness. Due to computing limitations at the time, data were subsampled to ~30 km for processing, but data at a finer scale (~10 km) were archived for future processing. The time has come to rescue and process the latter, given the advances in computing capabilities since 1983 and the need for higher resolution global cloud analysis to support climate studies.

ISCCP B1 data are a collection of measurements from imagers on international meteorological satellites which are sampled to approximately 10 km and 3 hourly. Data were originally archived at NOAA's National Climatic Data Center (NCDC) and, through the efforts described in this article, are now available in a consistent format with well-documented navigation and calibration routines. The following is a summary of the B1 data rescue, quality assessment, calibration comparison and application to geophysical parameters –an end-to-end application of scientific data stewardship.

After a short summary of ISCCP, the B1 data characteristics are described. Data rescue efforts are summarized with an emphasis on navigation algorithm issues and lessons learned from this rescue are also reviewed. Calibration assessments of the visible data are then

Received 8 Aug 2008; accepted 3 Nov 2008; published 19 Nov 2008 [CCC: 19313195/2008/$25.00]

discussed. Finally, we provide a summary of ISCCP B1 uses to date which includes a new objective reanalysis of worldwide tropical cyclone activity.

1.1 The International Satellite Cloud Climatology Project

ISCCP was initially envisioned as a 5-year project but has now continued for more than 25 years. One of the goals of ISCCP is to derive a global cloud climatology to improve our understanding of the Earth's radiation budget [1]. ISCCP accomplishes this goal via a federation of data providers (termed Satellite Processing Centers, SPCs) which supply raw satellite data to the ISCCP Global Processing Center (GPC). During the history of ISCCP, there have been nine such SPCs. Each SPC subsamples full resolution geostationary data to roughly 10 km and 3 hours, thereby creating so-called B1 data. The B1 data have been archived at NCDC but were rarely accessed. Due to the relatively large volume of the B1 data, the GPC processed B2 data, which are a further subsampling of the full resolution data to 30 km and are provided by the SPCs. Thus, no processing, testing or usage of the B1 data occurred between the date of initial archive and 2003.

In 2003, data rescue efforts began at NCDC, but by this time the original B1 data archives had already languished for years with little documentation, access software or quality assessment. Thus, in light of the need to reprocess the entire B1 data set for climate research, it became necessary to assess and correct these deficiencies. The described work produces climate-quality information from the B1 data by compiling documentation, developing input and output software, testing data quality, assessing calibration and making the data accessible to the public. More information on ISCCP is available through its website (http://isccp.giss.nasa.gov/), a function created well after the start of the project since web technology did not yet exist in 1983.

2 ISCCP B1 DATA

In spite of the fact that the ISCCP B1 data set is made up of different sensors on satellites launched by various countries, the observations are similar. The ISCCP B1 data span 1978 to present and are available at 3 hour intervals near the synoptic hour (00, 03, … 21 UTC). The start and end dates (along with primary satellite location) of the ISCCP B1 data record from each satellite are provided in Table 1 and summarized by Fig. 1. The satellite have different operational start and end dates; the dates in Table 1 and Fig. 1 refer to ISCCP B1 provision from the SPC. The Japanese operate the Geostationary Meteorological Satellite (GMS) satellite series located over 140° East which was more recently replaced by the Multi-Function Transport Satellites (MTSAT-1R, abbreviated MTS-1 in Fig. 1). European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) operates the Meteosat (MET) series of satellites at the Prime Meridian (0° East) and more recently an additional location over the Indian Ocean at 63° East (prior to EUMETSAT, the European Space Agency provided data to ISCCP for Meteosat 2 and 3). Lastly, the United States operates the Geostationary Operational Environmental Satellite (GOES) series in two positions spanning the Western Hemisphere: GOES-West and East at 135° and 75° West, respectively. Also, data from the historical U.S. Synchronous Meteorological Satellite (SMS) series have recently become available. As can be seen in Fig. 1, not all geostationary satellites are stationary with respect to their temporal and spatial coverage. The GOES coverage was routinely adjusted when only one satellite was available, such as after the failures of GOES-5 and GOES-6. Likewise, the cooperation of countries led to moves of MET-3, MET-5 and GOES-9 to minimize gaps in satellite coverage.

The annual ISCCP B1 data volume was relatively large by 1983 standards, but has become more manageable with time. The annual data rates, by satellite, are provided in Fig. 2. The rate in the 1980s was just more than 15 GB per year. The increase in the mid-1990s is

Table 1. Details of ISCCP B1 data (where *Current* is as recent as October 20, 2008)

Satellite	Period of Record	Meridian	Original Dynamic Resolution (bits)				
			Vis	IR (μm)			
				3.9	6.7	11	12
GMS-1	03/01/1981 - 06/30/1984	140ºE	6	-	-	8	-
GMS-2	07/01/1983 - 09/27/1984	140ºE	6	-	-	8	-
GMS-3	09/28/1984 - 12/03/1989	140ºE	6	-	-	8	-
GMS-4	12/04/1989 - 06/12/1995	140ºE	6	-	-	8	-
GMS-5	06/13/1995 - 05/21/2003	140ºE	6	-	8	8	8
MTSAT-1R	10/01/2005 – *Current*	140ºE	10	10	10	10	10
SMS-1	01/15/1979 - 04/19/1979	75°W	6	-	-	8	-
SMS-2	04/19/1979 - 08/05/1981	75°W	6	-	-	8	-
GOES-1	12/01/1982 - 05/31/1983	135°W	6	-	-	8	-
GOES-2	04/20/1978 - 12/31/1978	75°W	6	-	-	8	-
GOES-3	11/20/1978 - 03/05/1981	135°W	6	-	-	8	-
GOES-4	09/24/1980 - 11/26/1982	135°W	6	-	-	8	-
GOES-5	08/05/1981 - 07/30/1984	75ºW	6	-	-	8	-
GOES-6	06/01/1983 - 01/19/1989	*Variable*	6	-	-	8	-
GOES-7	02/03/1987 - 01/11/1996	*Variable*	6	-	-	8	-
GOES-8	12/01/1994 - 04/01/2003	75ºW	10	10	10	10	10
GOES-9	01/01/1996 - 07/27/1998	135ºW	10	10	10	10	10
GOES-9	05/22/2003 - 10/31/2005	140ºE	10	10	10	10	10
GOES-10	07/28/1998 - 06/21/2006	135ºW	10	10	10	10	10
GOES-11	06/26/2006 – *Current*	135°W	10	10	10	10	10
GOES-12	04/01/2003 – *Current*	75ºW	10	10	10	10	-
MET-2	09/03/1981 - 08/25/1988	0ºW	8	-	8	8	-
MET-3	08/25/1988 - 06/19/1989	0ºW	8	-	8	8	-
MET-3	09/12/1992 - 12/31/1994	75ºW	8	-	8	8	-
MET-4	07/01/1989 - 02/03/1994	0ºW	8	-	8	8	-
MET-5	02/03/1994 - 02/13/1997	0ºW	8	-	8	8	-
MET-5	07/01/1998 – 04/16/2007	63ºE	8	-	8	8	-
MET-6	02/13/1997 - 06/03/1998	0ºW	8	-	8	8	-
MET-7	06/03/1998 - *Current*	*Variable*	8	-	8	8	-
MET-8	01/01/2005 - 05/10/2007	0ºW	10	10	10	10	10
MET-9	05/10/2007 - *Current*	0°W	10	10	10	10	10
FY-2C	07/22/2005 - *Current*	105ºE	10	10	10	10	10

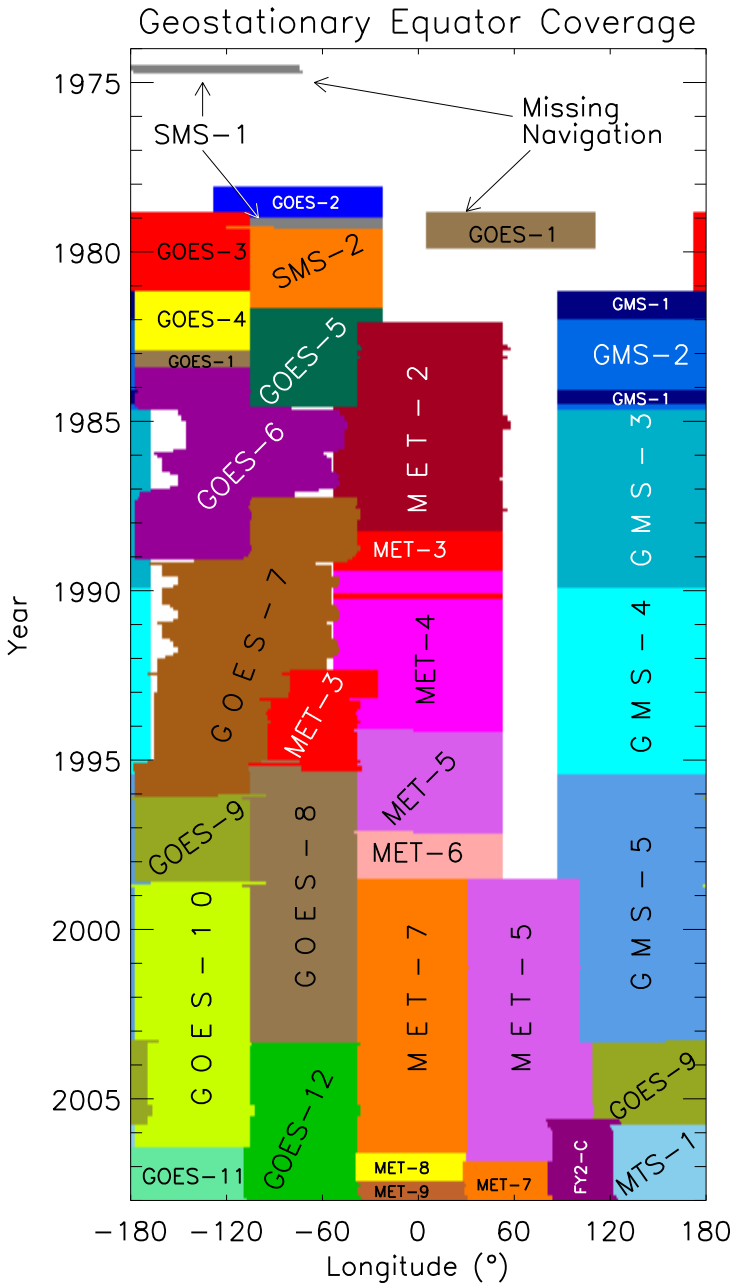


Fig. 1. Time series showing geostationary observations at the Equator.

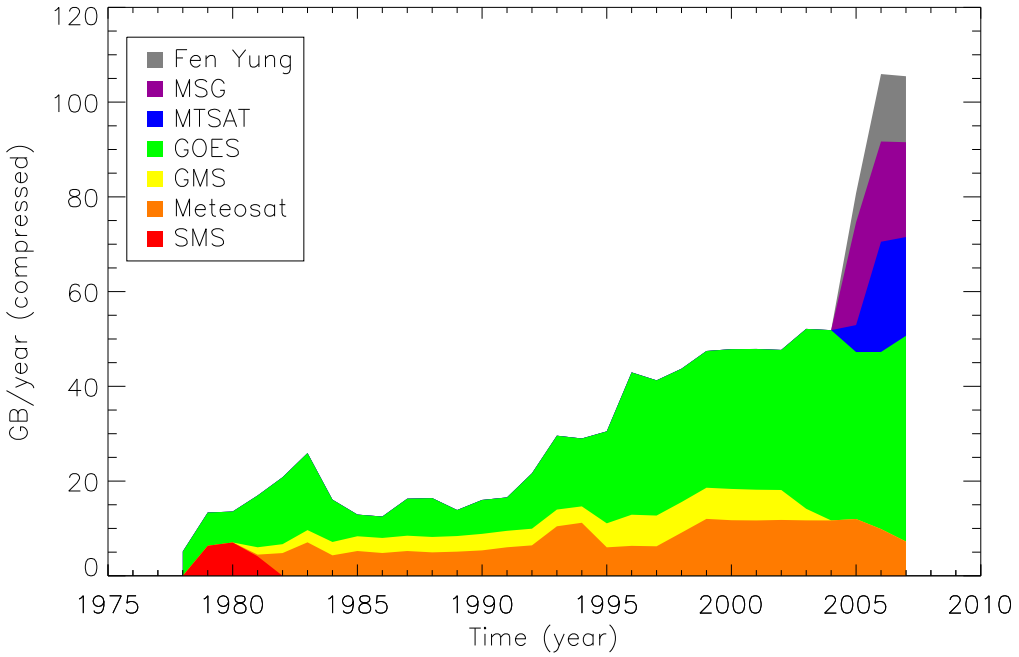


Fig. 2. Annual compressed data volume of the ISCCP B1 data sorted by satellite series.

due to the start of the 5-channel GOES data while the more recent increase in 2004 is caused by the start of other multichannel instruments (Meteosat Second Generation (MSG), MTSAT-1R) and the addition of a new satellite to the constellation (FenYung 2C, FY-2C). Another cause for the recent increase is the move from a 1-byte to a 2-byte data format. The dynamic resolution of the instruments comprising B1 data in 2008 are 10-bit (c.f., Table 1), however the initial ISCCP practice was to unify the dynamic resolution to 1-byte (i.e., 8 bits) to reduce data volume. Recently, this practice was abandoned to maintain the full dynamic resolution of the B1 data. This caused part of the most recent increase in the annual data rates.

While the temporal and dynamic characteristics of the B1 data are rather uniform, the spatial and spectral details vary, and will be described in the following sections.

2.1 Spatial Characteristics

The primary strategy of the spatial sampling was to create a data set which had a field-of-view of about 5 km subsampled to approximately 10 km resolution. Infrared channels were generally sampled to 8 or 10 km on many of the satellites, as seen in Table 2. Visible channels, on the other hand, were first averaged to match the native IR resolution then sampled in the same manner as the IR channels (see Table 3 for details on the sampling). For instance, 8x8 pixel averaging was performed on visible data from the early GOES satellites to match the 8km IR footprint, while the later GOES satellites have spatial averaging of 4x7 (to match the newer IR spatial footprint). Thus, the spatial resolution of the B1 data set is approximately 10 km, with some variation by satellite series.

Also, it should be noted that while geostationary satellites can view a large portion of the Earth, their most glaring weakness is an inability to view the poles. Nonetheless, 90% of the globe can be observed with a view zenith angle of less than 75º when 5 geostationary satellites are operating. As such, when discussions of geostationary data herein mention global coverage, we mean continuity of coverage around the Equator. Observations of the remaining 10% of the Earth are left to polar-orbiting satellites.

Table 2. Spatial sampling of full resolution IR data in creating ISCCP B1 data, where m×n is the number of scan lines (m) by the number of pixels along the scan (n).

Satellite	Original IR Res. (km)	Spatial Sampling	Nadir IR B1 Res (km)
GMS 1-5	5	2×2	10
SMS 1-2	7	1×2	7
GOES 1-7	7	1×2	7
GOES 8-12	4	2×4	8
MET 2-7	5	2×2	10
MET 8-9	3	3×3	9
MTSAT 1R	4	2×2	8
FY2 C	5	2×2	10

Table 3. Spatial averaging and sampling of visible data in creating ISCCP B1 data. Due to over sampling, IR pixels are under sampled in the along scan direction for GOES-8 through 12.

Satellite	Original Visible Res (km)	Spatial Averaging	Spatial Sampling	Nadir B1 Visible Res. (km)
GMS-1-5	1.25	4×4	2	10
SMS 1-2	0.9	8×8	1	7
GOES 1-7	0.9	8×8	1	7
GOES-8-12	1	4×7	2	8
MET 2-7	2.5	2×2	2	10
MET 8-9	3	None	3	9
MTSAT 1R	1	4×4	2	8
FY2 C	1.25	4×4	2	10

The ISCCP B1 spatial coverage of the Earth is depicted in Fig. 1. In short, most of the Earth is sampled for the entire span of the ISCCP record. The only substantial gap is the lack of satellite data over the Indian Ocean. However, the gap was filled in 1998 with the move of Meteosat-5 to 63ºE. This provides global coverage of the visible, infrared window and water vapor channels.

2.2 Spectral characteristics

The ISCCP B1 data are primarily composed of visible (VIS), infrared water vapor (IRWVP) and infrared window (IRWIN) channels (roughly 0.6μm, 6.7μm and 11μm, respectively). The visible and infrared channels are window channels that are more sensitive

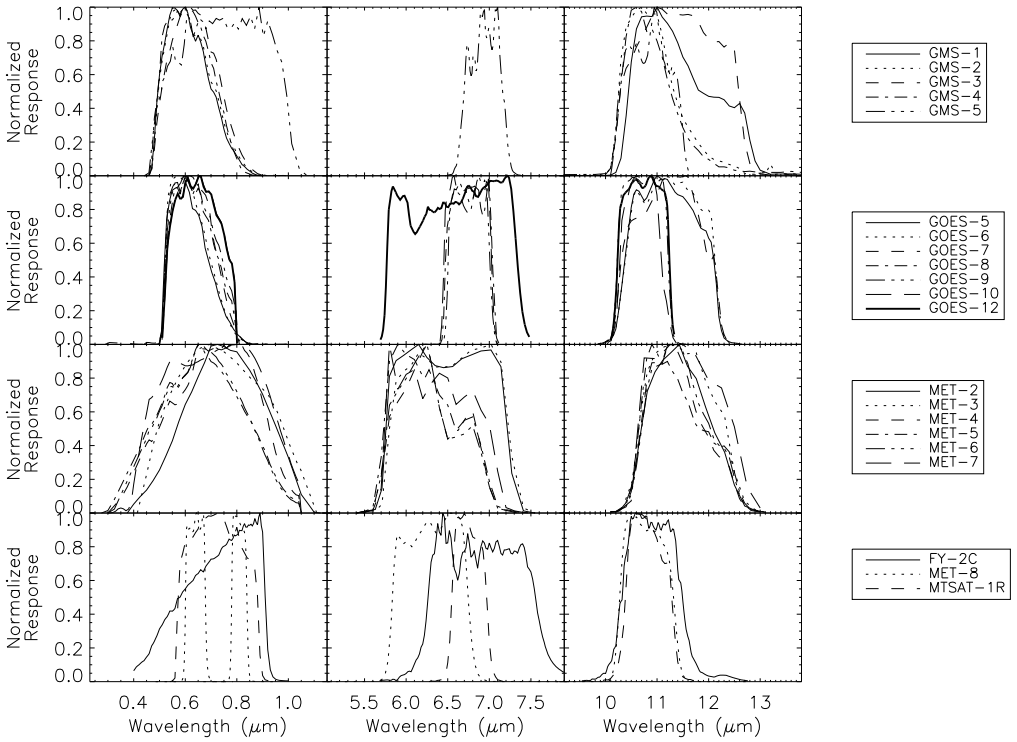


Fig. 3. Spectral response function of the visible (left column), water vapor (center) and infrared window (right) for the satellites in the ISCCP B1 data set.

to the surface than the atmosphere which helps discriminate clouds from clear sky [2]. Conversely, the water vapor channel is mostly opaque and as such, is not used in the ISCCP cloud mask algorithm. However, its discussion is included here since it makes up a large part of the B1 data set and has global coverage since 1998. The spectral responses (obtained from ISCCP SPCs) for each channel from all satellites are shown in Fig. 3. Radiative transfer calculations provide insight into the variability amongst the instruments. The following uses calculations from the Santa Barbara DISORT Atmospheric Radiative Transfer (SBDART) model [3, 4] which are presented in Fig. 4. The following is a discussion of similarities and differences amongst the instruments with consideration toward producing a geostationary-based, temporally-consistent climate record.

The VIS channel is part of the daytime cloud mask algorithm [2]. However, it has other quantitative uses, such as retrieval of surface albedo [5] and aerosol optical depth [6, 7]. Most visible bands (left column of Fig. 3) have peak sensitivity near 0.7 μm with central wavelengths (Fig. 4) between 0.6 and 0.8 μm. The Meteosat Second Generation which comprises Meteosat 8 and 9 has three visible channels. The high resolution visible data are not provided to ISCCP, so the two visible channels for Meteosat 8 in Fig. 4 are the 0.6 and 0.8 μm channels. (Meteosat 9 has nearly identical spectral characteristics as Meteosat 8, so it is not shown). These two channels act differently than the broad visible channels of the other instruments as they were designed to provide information on vegetation. The result is vast differences in some quantities, such as the theoretical reflectance of a conifer forest[1].

[1] Conifer reflectance obtained from the ASTER Spectral Library through the courtesy of the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California. © 1999, California Institute of Technology. ALL RIGHTS RESERVED.

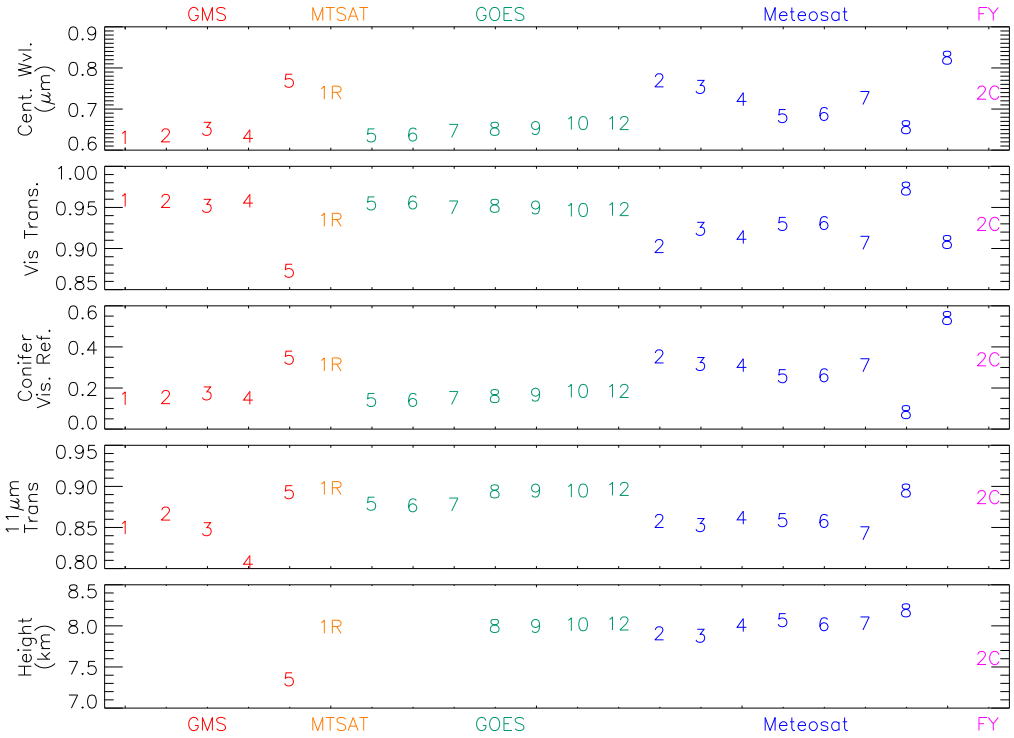


Fig. 4. Various theoretical values for ISCCP B1 channels demonstrating variability between instruments (top to bottom): visible central wavelength (μm), visible transmittance, visible reflectance for a conifer forest, IRWIN transmittance, and altitude (km) of peak responsivity of the IRWVP channel (where available).

Other channels with longer central wavelengths (e.g., GMS-5, MET-2) have less transmittance and larger conifer reflectance. Such variability between satellites does not significantly affect cloud detection but will have larger impacts when retrieving parameters with smaller signals, such as aerosol or surface albedo.

The IRWIN channel is generally useful for detecting clouds both day and night and similarly has both broad and narrow bands. The reason is due to the differences in atmospheric absorption between channels centered at 11 and 12 μm. Most earlier satellites (early GOES, GMS and all Meteosat) have a broad band centered near 11μm which have lower transmittances due to increased absorption near 12μm (IRWIN transmittance in Fig. 4 are calculated using the US Standard atmosphere). A split window method [8] to remove the atmospheric transmittance by locating two narrow bands with different levels of atmospheric absorption (e.g., narrow channels centered at 11 and 12 μm) was adopted on GOES-8 through 11, GMS-5 and other new satellites. Correspondingly, these instruments also have a split window channel located at 12 μm that are not shown here but are available in ISCCP B1 data.

One of the primary purposes of the IRWVP channel is to provide atmospheric motion vectors [9, 10] and it has since been used for atmospheric circulation studies at various temporal and spatial scales [11, 12]. In the ISCCP B1 data set, there is some variation in spectral responses among satellite series but most instruments are sensitive to radiation between 6 and 7 μm. The altitude of peak responsivity (i.e., the peak of the weighting function), $H_{6.7}$, ranges from (aside from GMS-5) 7.8 to 8.1 km, calculated using SBDART with the US standard atmosphere. So while the weighting functions differ, the net effect is

little difference in $H_{6.7}$. ISCCP does not provide the same level of calibration for IRWVP as the VIS and IRWIN channels because it is not used in the ISCCP cloud detection routine and only recently began global coverage (since 1998).

Data for other satellite channels are available in ISCCP B1 data. GOES-8 through 11, MTSAT-1R and FY-2C include near infrared (3.9μm) and the split window (12 μm) channels. Also, all eleven channels from MSG satellites are available in the ISCCP B1 data set. However, since these other channels are not available globally, ISCCP does not provide any further inter-satellite calibration. Nonetheless, the calibration available from the SPCs allows them to be used for regional studies of various atmospheric phenomena.

The VIS, IRWIN and IRWVP channels represent the observations of ISCCP B1 data which are available on all current satellites. While differences between the channels described above do exist, their similarities have allowed calibration toward obtaining geophysical retrievals, such as the ISCCP cloud analysis. However, prior to performing geophysical retrievals using B1 data set, algorithms were needed to decode the data into useful information.

3 ISCCP B1 DATA RESCUE

Given the potential for weather and climate applications from this data set, an effort needed to be made to use these data. However, decoders were not available to read the files nor were algorithms on hand to calculate Earth locations for the pixels. Furthermore, significant gaps existed in the archived period of record. The results of the ensuing data rescue effort produced some significant lessons learned which apply to the long-term archive of any scientific data set.

3.1 Lessons learned

ISCCP, at its inception a 5-year project, did what was then practical to ensure usability of the B1 data and expanded procedures as it became clear the project would be long-lived. Nonetheless, several issues contributed to the failure to properly maintain ISCCP B1 data. The common thread in each issue was that the archived data were never actually used. Data users might have requested documentation of the data and read algorithms, in which case, the archive would have formulated and collated the incoming data and documentation from the various providers to support users. Lastly, the SPCs might have worked with the archive to make sure that the correct files and documents were archived. Alas, there were no actual users of the B1 data from the archive. Neither the SPCs nor the archive ensured data and documentation were properly archived. The ISCCP project focused processing on the B2 data, pushing the B1 data to the "back burner" where it was soon forgotten. While the mere presence of users would not necessarily have saved the data, it might have drawn attention to the state of the archive and kept it from falling into disarray. Regardless, the focus of the archive process should have been maintaining the information, not merely maintaining the data files (which can quickly become obsolete).

The results of the data rescue provide insight toward the proper archiving of data, in particular, toward data stewardship activities which might minimize the loss of such information. Recommendations are detailed by [13] and are summarized here with annotations detailing on how the rescued ISCCP B1 data has minimized the potential of future loss of information. Much of the technology described below (common data formats, metadata standards, etc) were not available when ISCCP began, but will be incorporated into the ISCCP B1 data rescue.

1. **Adopt existing standards where possible, creating new formats only when absolutely necessary** – Standards exist for metadata, file formats, etc. There is little

need to create data in non-standard file formats. To this end, ISCCP B1 data will be available in NetCDF version 4 when complete.

2. **Answer as many of the "Why?" questions in the documentation as possible** – To the extent possible (i.e., given the existing documentation or lack thereof), ISCCP B1 data will be completely documented.
3. **Include descriptive metadata within a data file** – Since the netCDF format will be used, each file will contain full documentation on reading and understanding the file contents.
4. **Make the descriptive metadata as complete as possible** – Again, given the lack of existing documentation on satellite maneuvers, calibration results, etc, the ISCCP B1 data set will be documented to the greatest extent possible.
5. **Verify that all data and documentation are accurate and are available to the users from the archive** – The work to rescue the ISCCP B1 showed that a lack of users increases the chance for poor data support. This can be avoided, in part, by using standard formats.
6. **Minimize the potential for information atrophy** – This issue becomes much more important as the time since the date of archive increases. The NetCDF format will be backwards compatible, however atrophy primarily occurs from a lack of use. It is expected that B1 data will be used for many years to come due to its compact size and global coverage. Therefore, we expect the potential for atrophy to be minimized.

These lessons were learned as a result from attempts to read and navigate the ISCCP B1 data.

3.2 Reading files

Given the lack of ISCCP B1 software decoders and centralized file format documentation, the laborious task of determining data formats and constructing file decoders for data from each satellite and SPC began. Once an algorithm was developed that could decode a file (i.e., read and understand most variables), it was tested on other files since the decoders for some data were often close to the formats from other satellites or SPCs. However in one case, several decoders were needed for the same satellite series from the same SPC due to format changes made by the SPC. Additionally the archive sometimes added to the confusion. For example, the format of a set of files from an SPC was found to be identical to another format barring some sporadically-spaced bytes within the file. Eventually, it was found that the sporadic bytes were tape headers used in writing the data on 9-track tape; when the files were ported to other media, the obsolete tape headers were inadvertently left in the files. Developing decoders for all ISCCP B1 file formats was complicated by these and other such issues.

The means of determining the data formats were as varied as the data:

- The EUMETSAT format was found in the original 1983 *ISCCP Data Management Plan*.
- GPC provided documents which described the Atmospheric Environment Service of Canada (AES) data formats.
- Colorado State University (CSU) provided IDL routines to read their B1 data formats.
- The JMA provided B1 format description documents and translated them to English.
- The University of Wisconsin file format was determined (through trial and error) to be an early version of the Man Computer Interactive Data Access System (McIDAS) Area data files whose documentation was found on a bookshelf in NCDC.
- Formats for the scan line headers of some GOES files were derived from Mode A & AAA documentation.

During this process, data from new satellites began to arrive at NCDC, such as Meteosat-8 and MTSAT-1R, requiring even more decoders to be developed. As of 2008, more than 14 formats have been identified from 28 satellites provided by 8 SPCs. However, reading the files merely allows the data to be imaged for qualitative analysis. It is still necessary to perform image navigation for the data to be used quantitatively, which is the process of calculating the Earth location for each image pixel.

3.3 Navigating files

Image navigation requires three sets of parameters defining: the location of the pixel within the image, the satellite attitude relative to the Earth, and the satellite orbit. The information defining the satellite image was provided by all SPCs in a similar manner (center scan line and pixel, the angular dimensions of the pixel, etc.). The satellite attitude is generally provided via the roll, pitch and yaw angles. The satellite orbit is often defined using Kepler orbital parameters (for more details on satellite navigation, see ref. [14]). However, navigating data for some satellites required unique navigation algorithms.

In particular, the satellite attitude parameters as provided for the GMS series required an algorithm to derive the roll, pitch and yaw. The attitude parameters provided by the SPC could not be used within available navigation algorithms and the source code provided with the data could not be successfully compiled. The Rosetta stone lay in the anchor points provided by the SPC which are a set of line/pixel pairs corresponding to a fixed set of Earth locations (latitude/longitude pairs) for a set of points on the Earth. At the suggestion of ISCCP [Rossow, personal communication, 2004], we derived an algorithm to iteratively compute the roll, pitch and yaw to match the provided anchor points.

The algorithm iterates attitude and orbital parameters to minimize the mean difference between the calculated image pixel for each anchor point and those provided in the file. A measure of the performance of the algorithm is the root mean square (rms) difference between the reported and calculated lines and pixels. The rms difference is provided in Fig. 5 for all satellites requiring this process. The rms in the pixel calculation (commonly, the east-west direction) are generally less than 1 pixel. The rms values in the line calculation (the north-

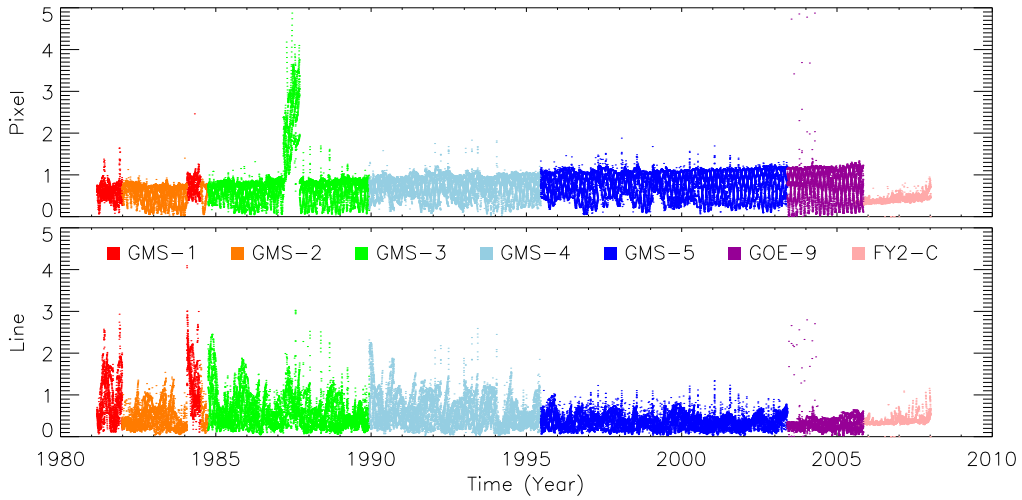


Fig. 5. Root mean square error between calculated and actual anchor points for image scan lines and pixels.

south direction) are a bit larger, but still close to one. These values are in line with the estimated ISCCP navigation uncertainty of ±1 pixel. Interestingly, the errors are much smaller for the FY2-C satellite. Further work should be performed to decrease the rms and understand the cause of the lower rms values for FY2-C.

Another issue in determining navigation was when data were over-determined. In the case of another satellite, three sets of parameters were provided which could be used to define the satellite orbit: Kepler elements, position and velocity, and Chebyshev polynomials. Calculations allow the conversion between the parameters sets. However, in some cases, the results of converting between the parameter sets were inconsistent. For example, an invalid semi-major axis in the Kepler parameters affects the estimated orbit, which could conflict with the position/velocity parameter set. Tests were developed for each parameter set (e.g., do the Kepler parameters define a valid geostationary orbit?) and to inter-compare the consistency between these parameters sets (e.g., does the position from the Kepler parameters at time T_o match the reported position?) to determine which parameter set provided the most accurate description of the satellite orbit.

Navigation accuracy is generally difficult to quantify. Herein, we checked navigation by producing mapped imagery and comparing coastlines in the image with those in a reference map. While tedious, this method suggests that the bulk of the imagery is mapped properly.

3.4 Filling gaps and extending coverage

Two relatively large gaps existed in the data archive. For an unknown reason, GOES-7 data from 1992 through 1994 were missing from the archive. Likewise, 18 months of GOES-8 data were missing in the 1996 to 1997 time frame. Also, various smaller gaps (~1 month or less) have been found. Overall, these gaps accounted for about 5% of the entire B1 record.

In order to fill the data gaps, full resolution GOES data were obtained from NCDC. The full resolution data overlap existing B1 data, such that the results of the averaging and sampling steps were compared with existing B1 files. In general, the visible imagery were averaged to match the IR resolution, and then both were sampled to the B1 resolution, following Tables 2 and 3. Comparisons between the new and original B1 files ensured consistency of the radiometric and navigation quality.

It was clear that once the capability existed to fill gaps in the B1 period of record, this process could also be used to extend it. ISCCP data collection had a somewhat arbitrary start date of 1 July 1983. However, the geostationary satellite operators have archives of data prior to this date. Satellite operators were contacted to obtain B1 data for the historical data. The JMA processed the earliest records of GMS-1 and EUMETSAT processed the earliest Meteosat-2 records. Likewise, NCDC processed the earliest GOES and Synchronous Meteorological Satellite (SMS) records.

These efforts extended the ISCCP B1 data back in time. ISCCP B1 data for SMS and GOES begin in 1978, GMS in 1981 and Meteosat in 1982. Thus it is now possible to extend the ISCCP global cloud climatology back to 1982 (which extends the period of record by 15 months) with regional analysis as early as 1978.

4 CALIBRATION ASSESSMENT

Nearly 30 years after operational geostationary satellites began operation from numerous countries, there are still very few channels in common among all satellites. The mainstay has been the visible and infrared window channels, available on all geostationary meteorological instruments. A recent addition is the infrared water vapor channel. The ISCCP calibration of the infrared window and visible channels is described elsewhere [15, 16]; we assess the ISCCP visible channel calibration below.

4.1 Visible channel – 0.7μm

We perform an independent assessment of the ISCCP visible calibration following [17]. The basis of the method is that the distribution of the energy reflected from the Earth is temporally constant for a given geostationary position (similar to the ISCCP calibration [15]). That is, for a full disk image at local noon where the disk of the Earth is fully illuminated, the difference in radiance between two quantiles is temporally constant (after correcting for mean Earth-Sun distance), following [18].

First, we investigate the change in digital counts (DC) from the ISCCP visible channels. This is the time varying portion, V(t), of Equation 6 in [17] and is calculated as:

$$V(t) = \frac{\mu_o F_o}{DC_{80} - DC_5} \tag{1}$$

where DC_{80} and DC_5 are the digital counts for the 5th and 80th percentile (respectively) from a full-disk image near local noon, μ_o is the cosine of the solar zenith angle for the satellite nadir point and F_o is the solar irradiance integrated over the instrument spectral response (and adjusted to account for eccentricity of Earth's orbit). Thus, a decrease in the instrument responsivity (such as from a contaminated mirror) decreases DC_{80}, decreasing the denominator and thereby increasing V(t). The time series of V(t) is provided for each primary geostationary subpoint in Fig. 6. Each portion of the figure contains data from a separate region of geostationary coverage: GOES-West, GOES-Central, GOES-East, Europe/Africa, Indian Ocean and the Japan/Australia position.

Trends in V(t) represent change in instrument performance, specifically in instrument responsivity. The newer GOES satellites show large degradation in responsivity in their upward trend of V(t) for GOES-8,9,10,12. Such trends have been noted previously [19] and should be removed by applying post-launch calibration coefficients. Other satellites, such as MET-7 show smaller trends. The GMS series was particularly stable in time.

If the assumption made by [17, 18] holds, calibrating the data should eliminate the trends in V(t). To that end, a parameter assumed to be invariant by [17] is defined as, C(t):

$$C(t) = \frac{L_{80} - L_5}{\mu_o F_o} \tag{2}$$

where L_{80} and L_5 are the calibrated radiances at the 80th and 5th percentiles (respectively) from the same image. The denominator accounts for variation in the Sun-Earth distance in instrument differences. The radiances at the 80th and 5th percentiles represent the reflectance of the brighter clouds and darker ocean, respectively. The difference between the two are assumed to be unchanging. The stability of C(t) over time for a given satellite position is a measure of the calibration quality and a test of the assumption that C(t) should be constant. The time series of C(t) is shown in Fig. 7.

The trends in Fig. 6 are mostly removed by the ISCCP calibration as shown in Fig. 7. Notably, the trends disappear for the GOES-East and West positions. The differences between satellites in the series (e.g., the step change between GOES 8 and 12 in Fig. 6) disappears when considering C(t) (c.f. Fig. 7). Nonetheless, the results suggest areas where further calibration correction is required. In particular, the C(t) for GOES-7 in the GOES-East and Central locations is markedly low (0.08 versus 0.28 of other satellites in the series). This was determined to be the result from a discrepancy between the B1 dynamic resolution and the calibration tables produced by ISCCP and can easily be fixed. In other regions, the trends are not zero and will need to be further analyzed. For example, the calibration of the ISCCP B1 MET-2 data prior to 1983 was never performed (because ISCCP began in 1983). We initially

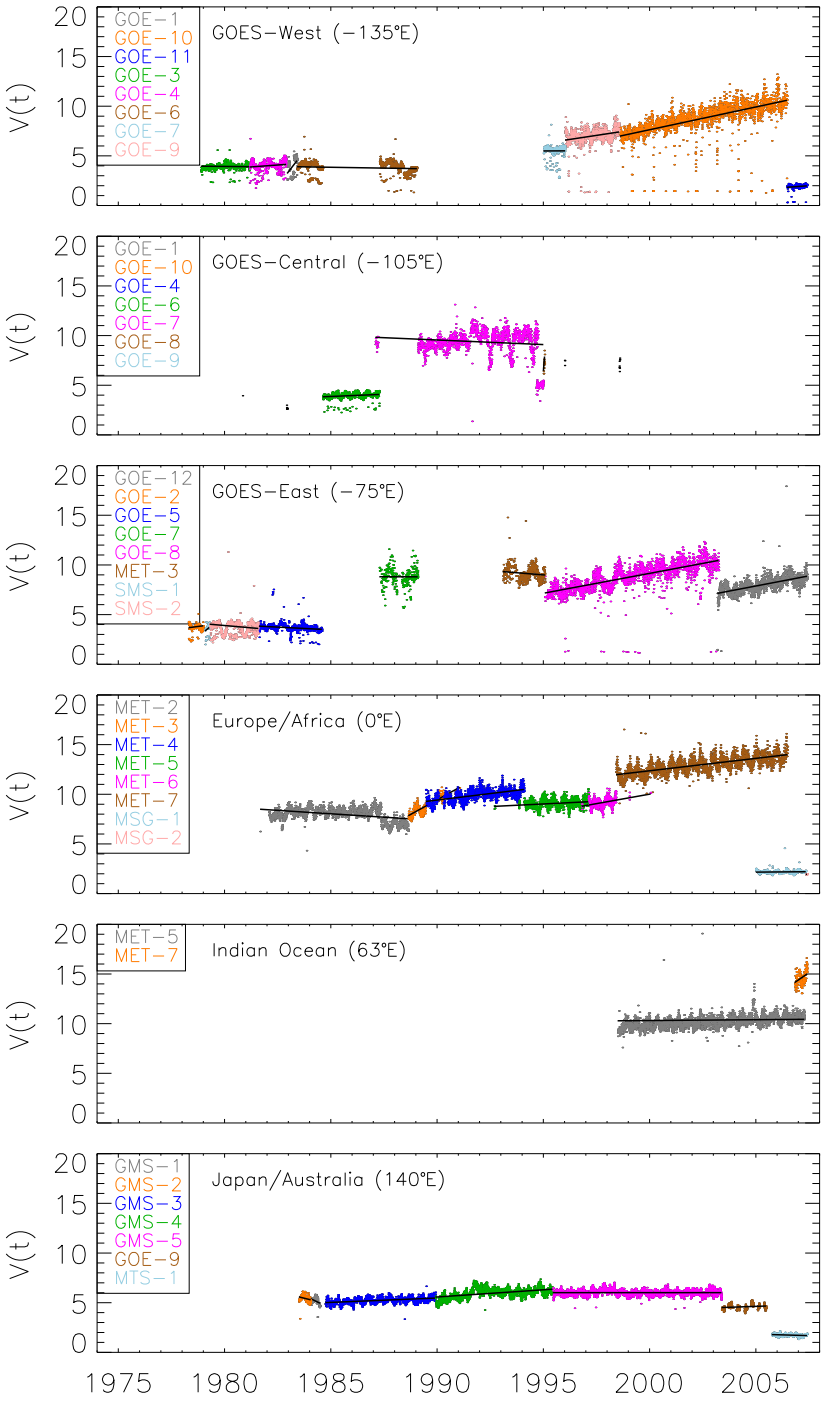


Fig. 6. Time varying calibration, V(t), defined by Equation 1 where solid lines are a linear fit for each satellite.

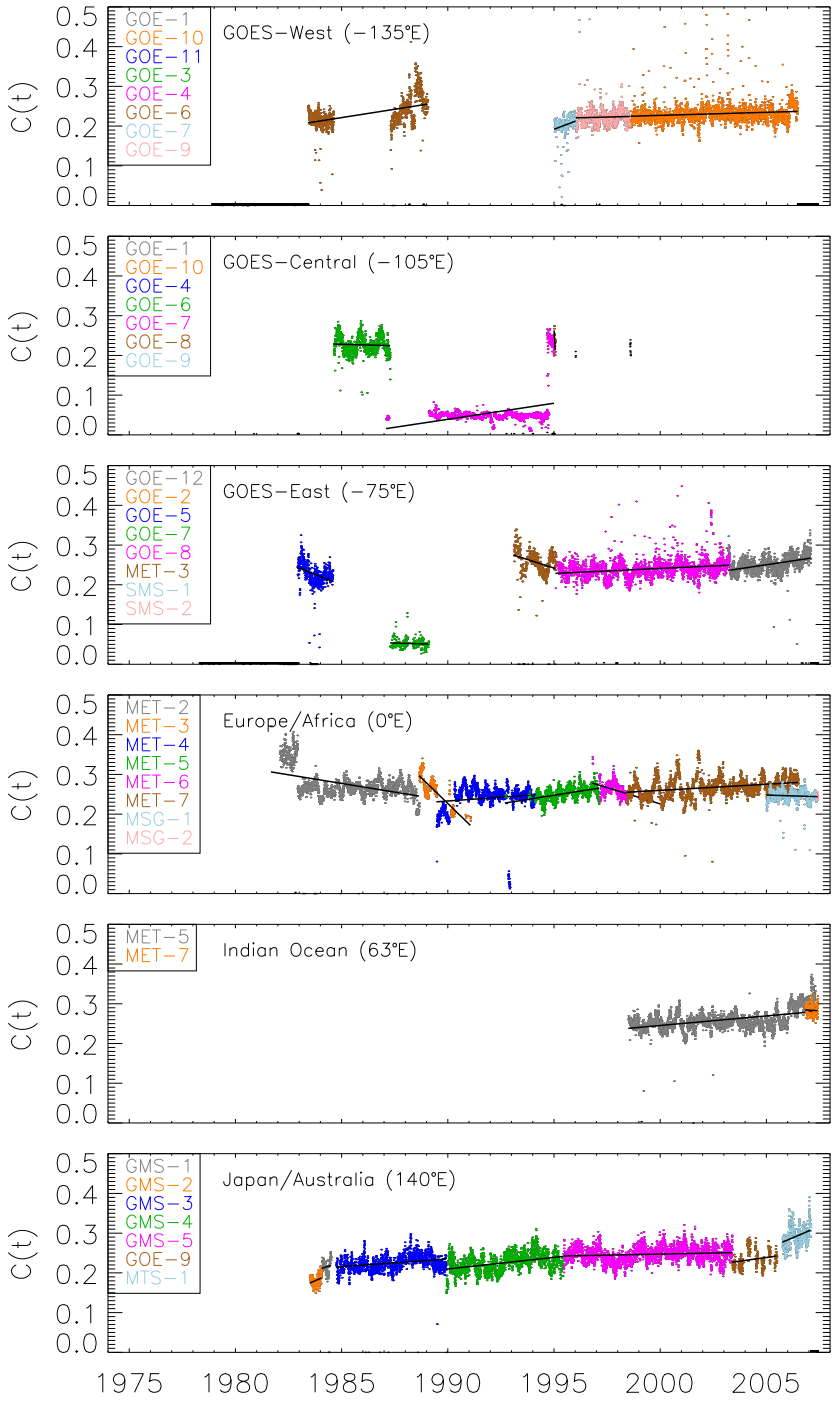


Fig. 7.Time invariant variable, C(t), defined by Equation 2 where solid lines are a linear fit for each satellite.

applied the same calibration to all 1982 MET-2 data as the July 1983 calibration (that is, the first available month of MET-2 ISCCP calibration). This analysis shows that to be inaccurate and that the full ISCCP calibration should be performed on all pre-1983 data. Furthermore, GOES-6 (in the GOES-West slot) seems unstable relative to other satellites as do Meteosats 3 and 4 from 1989 to 1991. Nevertheless, the overall calibration seems quite stable and the method proposed by [17] works well.

4.2 Infrared water vapor channel – 6.7μm

The infrared water vapor channel is not routinely calibrated by ISCCP. However, inter-calibration for the period of record will not be possible until a HIRS inter-satellite bias correction is available due to changes in the HIRS central wavelength, between NOAA-14 and NOAA-15. Once the bias correction is complete, the ISCCP B1 data should be normalized with the aim to provide water vapor data globally from 1998 and regionally beginning in 1982.

4.3 Infrared window channel – 11μm

The infrared window channel was calibrated using the HIRS instrument as a reference [20] as part of this data rescue. B1 data were co-located with HIRS observations. The result was that an error in the ISCCP calibration of AVHRR was found due to a change in the calibration coefficients provided by NOAA. The calibration removed this temporal bias to provide a consistent record of observations for the IRWIN. The ISCCP B1 data are updated annually using this calibration.

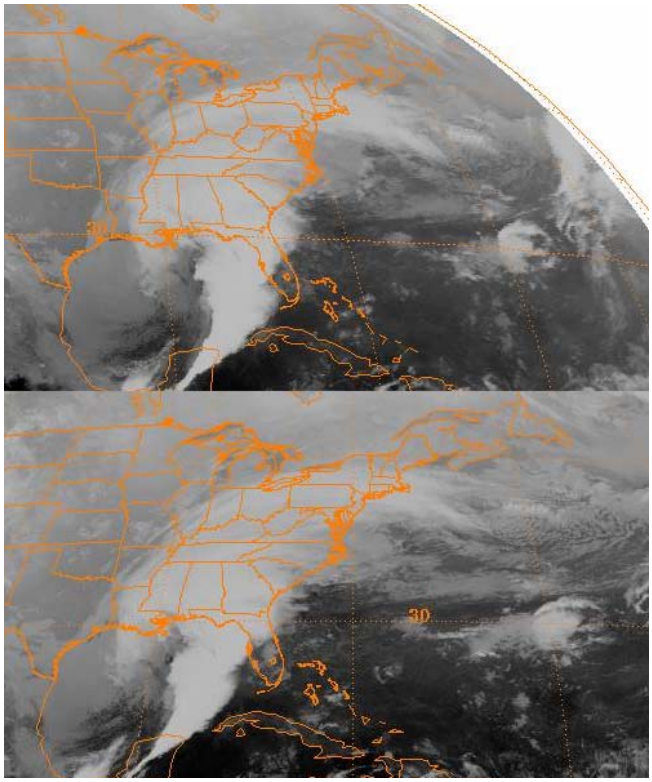
5 SUMMARY

5.1 Data Usage

A unique use of the data arose from an effort to assess data quality. It was found that viewing imagery from the rescued data was a quick means for determining whether the read and navigation algorithms were working properly. To facilitate such analysis, we developed the Global ISCCP B1 Browse System (GIBBS) website[2]. The result allows users to view any B1 image for the three primary channels (visible, infrared window, and infrared water vapor). The imagery are updated daily as data are provided by the SPCs. The result is an unparalleled web resource for browse-able satellite imagery. Nearing one million images, the uses of the website have been numerous. An example image from GIBBS is provided in Video 1 focusing on the "Storm of the Century"– a unique blizzard which developed over the Gulf of Mexico on March 12, 1993 and impacted much of the Eastern United States over the next two days [21]. The linked movie image was produced using consecutive imagery provided by GIBBS. Also, GIBBS imagery have been used to populate Wikipedia[3] with hurricane imagery and to provide imagery for researching case studies, such as South American rainfall variability [22].

ISCCP B1 data has been applied in various ways to hurricane research. In addition to aiding case studies and investigating extratropical transitions of hurricanes [23], IRWIN data over tropical cyclones were gridded to produce the Hurricane Satellite[4] B1 (HURSAT-B1) data set [24]. The 3-hourly coverage of the tropics provided numerous observations of tropical

[2] http://www.ncdc.noaa.gov/gibbs/
[3] http://commons.wikimedia.org/wiki/Image:Hurricane_Javier_1992.jpg
[4] http://www.ncdc.noaa.gov/oa/rsad/hursat/



Video 1. Coincident imagery from GIBBS for the Storm of the Century on
13 March 1993 at 03:00 UTC from
GOES-7 at 100° West (top) and Meteosat-3 at 75° W (bottom) (MPEG, 2.0 MB)

cyclones, such as that shown in Video 2 of Hurricane Andrew. Kossin et al. [25] then apply an objective algorithm to derive hurricane intensity from azimuthally-averaged IRWIN brightness temperatures. The results produced a new intensity record that is free of temporal inconsistencies. In addition to these climate studies, HURSAT-B1 data has been critical in the development of real time tropical cyclone forecasting tools, for example, in constructing inner core wind structure in lieu of aircraft reconnaissance [26] and forecasting eyewall replacement cycles [27].

Hence, the availability of the rescued ISCCP B1 data has piqued weather interest from the most basic user to research scientists. The data have improved our understanding in disciplines as disparate as short-term forecasting and long-term climate assessment. There are many other untapped topics which can benefit from an analysis of the rescued ISCCP B1 data set.

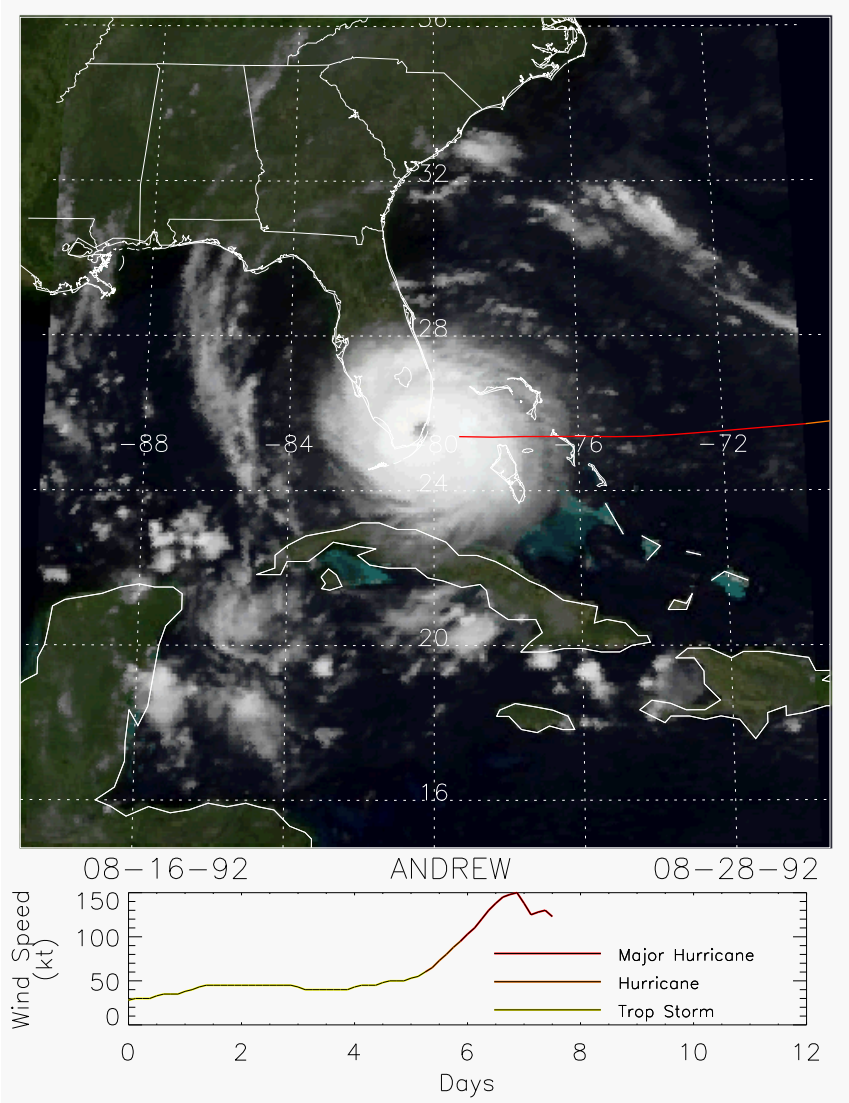
5.2 Future Work

Further work includes incorporating B1 data in future ISCCP reprocessing of cloud products, a data set reformat, further navigation analysis and continued analysis of data calibration.

The ultimate goal of the ISCCP B1 data rescue was to incorporate the data into a future reprocessing of the ISCCP cloud climatology. The result would then provide higher spatial resolution in data sets which depend on ISCCP, such as the GEWEX Surface Radiation Budget (SRB), which could then provide better estimates for applications such as solar energy. With the data rescued, the work now depends on the modifications to the ISCCP software to accommodate the higher resolution.

Currently, the ISCCP B1 data set is still in a flat binary file with full documentation provided by NCDC. However, it is expected that it will soon be ported to netCDF version 4 since it provides the added benefits of: per variable compression, widely–supported, self-describing data format, internal metadata and documentation and possibly serving these data through modern data services such as THREDDS[5].

Further investigation of the ISCCP calibration, particularly visible and water vapor channels is required with comparisons to various reference data sets to maximize the calibration accuracy should be performed.



Video 2. Sample image of HURSAT-B1 data for Hurricane Andrew making landfall in Florida as a Saffir-Simpson Category 5 hurricane (MPEG, 7.3 MB).

[5] http://www.unidata.ucar.edu/projects/THREDDS/

5.3 Synopsis

The rescue of this data set represents true end-to-end data stewardship. The data were rescued where no data read, calibration, or navigation algorithms existed. The data were then assessed for quality and calibration. Small adjustments were made to the ISCCP inter-satellite calibration, due to an error found in the ISCCP calibration. The new homogeneous record has since been used to create a data set centered on historic hurricanes/tropical cyclones allowing a reassessment of trends in global hurricane activity. Such a stewardship process can be used in other fields and this effort may be seen as a model of scientific data stewardship of satellite data.

Acknowledgements

The author acknowledges many involved in ISCCP without whom the ISCCP B1 data set would still be un-useable bytes in the NCDC archive:

- John Bates and others at NCDC, for supporting the research.
- Bill Rossow, Chris Bishop, and Violeta Golea at the ISCCP Global Processing Center who provided much assistance throughout the process.
- The ISCCP Satellite Processing Centers for their support by providing the ISCCP B1 data collection: AES (Environment Canada), CSU (Colorado State University), UWS (University of Wisconsin), JMA, EUMETSAT, ESA (European Space Agency), and CMA (Chinese Meteorological Administration).
- Okuyama Arata and the JMA, for translating ISCCP B1 documentation to English, providing historic (pre-1983) B1 data and filling B1 data gaps.
- Garrett Campbell and Colorado State University, for help in reading and navigating B1 data provided by CSU.
- Ken Holmlund and EUMETSAT, for providing historical B1 data from Meteosat imagery and helping to fill B1 data gaps.

References

[1] W.B. Rossow and R.A. Schiffer, "ISCCP Cloud Data Products," *Bull. Am. Meteorol. Soc.* **72** (1), 2-20 (1991) [doi:10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2].

[2] W.B. Rossow and L.C. Garder "Cloud Detection Using Satellite Measurements of Infrared and Visible Radiances for ISCCP," *J. Clim.* **6** (12), 2341-2369 (1993) [doi:10.1175/1520-0442(1993)006<2341:CDUSMO>2.0.CO;2].

[3] K. Stamnes, S.C. Tsay, W. Wiscombe, and K. Jayaweera, "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media," *Appl. Optic.* **12** (12), 2502-2509 (1988)

[4] P. Ricchiazzi, S. Yang, C. Gautier, and D. Sowle, "SBDART: A research and teaching software tool for plane-parallel radiative transfer in the Earth's atmosphere," *Bull. Am. Meteorol. Soc.* **79** (10), 2101-2114 (1998) [doi:10.1175/1520-0477(1998)079<2101:SARATS>2.0.CO;2].

[5] A. Ipe, N. Clerbaux, C. Bertrand, S. Dewitte, and L. Gonzalez, "Pixel-scale composite top-of-the-atmosphere clear-sky reflectances for Meteosat-7 visible data," *J. Geophys. Res. Atmos.* **108** (D19), 4612 (2003) [doi:10.1029/2002JD002771].

[6] K.R. Knapp, "Quantification of aerosol signal in GOES-8 visible imagery over the U.S.," *J. Geophys. Res. Atmos.* **107** (D20), 10.1029/2001JD002001 (2002) [doi:10.1029/2001JD002001].

[7] K.R. Knapp, T.H. Vonder Haar, and Y.J. Kaufman, "Aerosol optical depth retrieval from GOES-8: Uncertainty study and retrieval validation over South America," *J. Geophys. Res. Atmos.* **107** (D7), 4055 (2002) [doi:10.1029/2001JD000505].

[8] E.P. McClain, W.G. Pichel, and C.C. Walton, "Comparative performance of AVHRR-based multi-channel sea surface temperatures," *J. Geophys. Res. Atmos.* **90**, 11587-11601 (1985) [doi:10.1029/JC090iC06p11587].

[9] N. Eigenwillig and H. Fischer, "Determination of mid-tropospheric wind vectors by tracking pure water vapor structures in Meteosat water vapor image sequences," *Bull. Am. Meteorol. Soc.* **63**, 44-58 (1982)

[10] C.S. Velden, C.M. Hayden, S.J. Nieman, W.P. Menzel, S. Wanzong, and J.S. Goerss, "Upper-tropospheric winds derived from geostationary satellite water vapor observations," *Bull. Am. Meteorol. Soc.* **78**, 173-195 (1997) [doi:10.1175/1520-0477(1997)078<0173:UTWDFG>2.0.CO;2].

[11] B.J. Tian, B.J. Soden, and X.Q. Wu, "Diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere: Satellites versus a general circulation model," *J. Geophys. Res. Atmos.* **109** (D10), D10101 (2004) [doi:10.1029/2003JD004117].

[12] J.J. Bates, X. Wu, and D.L. Jackson, "Interannual Variability of Upper-Troposphere Water Vapor Band Brightness Temperature," *J. Clim.* **9** (2), 427-438 (1996) [doi:10.1175/1520-0442(1996)009<0427:IVOUTW>2.0.CO;2].

[13] K.R. Knapp, J.J. Bates, and B. Barkstrom, "Scientific Data Stewardship: Lessons learned from a satellite data rescue effort," *Bull. Am. Meteorol. Soc.* **88** (9), 1359-1361 (2007) [doi:10.1175/BAMS-88-9-1359].

[14] S.Q. Kidder and T.H. Vonder Haar, *Satellite Meteorology: An Introduction*. 1995, San Diego, CA: Academic Press, Inc. 466.

[15] C.L. Brest, W.B. Rossow, and M.D. Roiter, "Update of radiance calibrations for ISCCP," *J. Atmos. Ocean. Tech.* **14** (5), 1091-1109 (1997) [doi:10.1175/1520-0426(1997)014%3C1091:UORCFI%3E2.0.CO;2].

[16] Y. Desormeaux, W.B. Rossow, C.L. Brest, and G.G. Campbell, "Normalization and Calibration of Geostationary Satellite Radiances for the International Satellite Cloud Climatology Project," *J. Atmos. Ocean. Tech.* **10** (3), 304-325 (1993) [doi:10.1175/1520-0426(1993)010<0304:NACOGS>2.0.CO;2].

[17] C. Rigollier, M. Lefevre, P. Blanc, and L. Wald, "The Operational Calibration of Images Taken in the Visible Channel of the Meteosat Series of Satellites," *J. Atmos. Ocean. Tech.* **19** (9), 1285-1293 (2002) [doi:10.1175/1520-0426(2002)019<1285:TOCOIT>2.0.CO;2].

[18] M. Lefèvre, O. Bauer, A. Iehlé, and L. Wald, "An automatic method for the calibration of time-series of Meteosat images," *Int. J. Rem. Sens.* **21**, 1025-1045 (2000) [doi:10.1080/014311600210416].

[19] K.R. Knapp and T.H. Vonder Haar, "Calibration of the Eighth Geostationary Observational Environmental Satellite (GOES-8) Imager Visible Sensor," *J. Atmos. Ocean. Tech.* **17**, 1639-1644 (2000) [doi:10.1175/1520-0477(1995)076<0165:OOTMS>2.0.CO;2].

[20] K.R. Knapp, "Calibration of long-term geostationary infrared observations using HIRS," *J. Atmos. Ocean. Tech.* **25** (2), 183-195 (2008) [doi:10.1175/2007JTECHA910.1].

[21] P.J. Kocin, P.N. Schumacher, R.F. Morales Jr., and L.W. Uccellini, "Overview of the 12-14 March 1993 superstorm," *Bull. Am. Meteorol. Soc.* **76** (2), 165-182 (1995) [doi:10.1175/1520-0477(1995)076<0165:OOTMS>2.0.CO;2].

[22] M.W. Douglas, J.F. Mejia, N. Ordinola, and J. Boustead, "Synoptic variability of rainfall and cloudiness along the coasts of Northern Peru and Ecuador during the 1997-8 El Niño event," *Mon. Weather Rev.* **In press** (2009) [doi:10.1175/2008MWR2191.1].

[23] M.J. Dickinson, K. Corbosiero, and L.F. Bosart, "The extratropical transitions of eastern Pacific Hurricane Lester and Atlantic Hurricane Andrew (1992): A comparison," *Mon. Weather Rev.* **In preparation** (2009)

[24] K.R. Knapp and J.P. Kossin, "New global tropical cyclone data from ISCCP B1 geostationary satellite observations," *J. Appl. Remote Sens.* **1**, 013505 (2007) [doi:10.1117/1.2712816].

[25] J.P. Kossin, K.R. Knapp, D.J. Vimont, R.J. Murnane, and B.A. Harper, "A globally consistent reanalysis of hurricane variability and trends," *Geophys. Res. Lett.* **34**, L04815 (2007) [doi:10.1029/2006GL028836].

[26] J.P. Kossin, J.A. Knaff, H.I. Berger, D.C. Herndon, T.A. Cram, C.S. Velden, R.J. Murnane, and J.D. Hawkins, "Estimating hurricane wind structure in the absence of aircraft reconnaissance," *Weather and Forecasting* **22**, 89-101 (2007) [doi:10.1175/WAF985.1].

[27] J.P. Kossin and M. Sitkowski, "An objective model for identifying hurricane secondary eyewall formation," *Mon. Weather Rev.* **In press** (2009) [doi:10.1175/2008MWR2701.1].

Kenneth R. Knapp is a meteorologist at NOAA's National Climatic Data Center in Asheville, NC. He received his BS in meteorology and mathematics from Lyndon State College in 1994, and his MS and PhD degrees in atmospheric science from Colorado State University in 1996 and 2000, respectively. His research interests include geostationary satellite data rescue, remote sensing of aerosol and clouds and other climate uses for operational satellite data.