Calibration Memo CAL/GEN/92-008

Calibration Index Files

Ian M George, Bill Pence, Ron S Zellar and M. F. Corcoran
NASA/GSFC
Greenbelt MD 20771

Original Version: 1995 Mar 01
Last update: 2020 Apr 20

`CIFVERSION = 1.1`

SUMMARY

This document discusses the role and format of Calibration Index Files (CIFs) within the Calibration Database (*CALDB*). Important HEASoft *CALDB* software tasks and *CALDB* library are described. We also reference documentation for the installation and maintenance of Calibration Database (*CALDB*) software & data at remote sites.

Intended audience: authors of analysis software needed to calibrate supported science data, and general users of HEASARC data.

LOG OF SIGNIFICANT CHANGES

Release Date	Sections Changed	Notes
1993 Feb 22	—	Original Version
1995 Mar 01	All	Made compatible with `LaTeX2HTML` software
2004 Apr 01	All	Made compatible with tth
2022 Apr 19	3.1.2	MFC: corrected CAL_VSD format to YYYY-MM-DD
2022 Apr 20	all	MFC: general updates

1 INTRODUCTION

In line with IAU and NASA policy, Calibration Index Files (CIFs) (and, all files within the OGIP[1]) are formatted using the Flexible Image Transport System (*FITS*; *eg.* see Wells et al. 1981, Griesen & Harten 1981). The files conform to the approved FITS standard, Version 4.0 from the IAU *FITS* Working Group.

Calibration Index Files serve as a *FITS*-formatted, simple database of valid files stored within a Calibration Database (*CALDB*). Due to various reasons, and as an aid to clarity, calibration data are stored in the *CALDB* as a relatively large number of small files. Each file usually contains a single aspect of calibration. Often, closely related datasets may be combined in a single *FITS* file to minimize disk space and for efficiency of access. In the event of a calibration update, the new calibration information will be written into a new file, which generally employs the same format as the outdated calibration file. Outdated calibration files are maintained in the *CALDB* which allows a *CALDB* user to compare effects of new calibration on older results. The CIF allows a user to locate and retrieve valid calibration data from the *CALDB* for a given high-energy mission & instrument.

The CIF for a given mission & instrument is a *FITS* file containing a single BINTABLE extension which contains specific information for each of the extensions in a (*FITS*-formatted) file in the *CALDB*. The CIF also stores with information on the validity of that file for calibration of data. The *FITS* header of each calibration file extension contains header keywords defined by the HEASARC which specify necessary details concerning the contents, origin, validity times and type of calibration data and applicable calibration parameters (appropriate off-axis angles or temperature ranges, as common examples)

This document describes the format and use of CIFs as implemented in the *CALDB*. Section 2 describes the overall purpose and design of a CIF. Section 3 gives a description of the CIF format. Section 4 lists the mandatory keywords which need to be included for all calibration files to be indexed in a CIF. Section 5 briefly describes important software tasks and routines for creating CIFs, updating CIFs when new calibration files are produced, and accessing data from the *CALDB* using the CIF. Section 6 provides references to other relevant documentation.

2 CALIBRATION INDEX FILES: Design

NASA's High Energy Astrophysics Science Archive Research Center (HEASARC) is dedicated to the proposition that all software should be as mission/detector independent as possible, to enable analysis of data in a multi-wavelength context with a minimal learning curve. This means that physically valid results should be obtainable through standard analysis of high-energy observations by non-expert users who may have minimal knowledge of the specifics of the

[1]The Office of Guest Investigator Services, or OGIP, was the office established at the NASA/Goddard Space Flight Center which managed the HEASARC when the HEASARC was first established in 1990 as NASA's multi-mission archive for X-ray and gamma-ray astronomy.

instrument. The *CALDB* is designed to allow software to apply the appropriate instrumental calibrations for an observation on at a specified time under specific instrumental conditions. To this end, calibration filenames are not hardcoded into the analysis software (a deprecated practice common in the 1980's), but selected by the analysis software from the *CALDB* using standard *CALDB* software tools. In the case of calibration data, this policy obviously gives users & software greater flexibility in choosing which calibration data they wish to use. Although an effort is made to make data calibration and analysis as "turnkey" as possible, it is the ultimate responsibility of the user to ensure any calibration data they use is appropriate and gives physically realistic results. The HEASARC helpdesk is available to users to help answer questions about the data, calibration, or analysis, or to point the user to additional resources.

2.1 Design Requirements

Calibration Index Files (CIF) serve the following functions:

1. To enable the HEASARC to keep track of the numerous calibration files within the archive.
2. To provide users with a concise summary of the available calibration information.
3. To enable analysis software to locate and retrieve required calibration information, either from a *CALDB* installed on a locally-mounted disk or remotely from the HEASARC *CALDB*.

To achieve these purposes, CIFs were designed with the following requirements:

1. The format of the file must be *FITS*.
2. The CIF must contain all the information necessary for software to uniquely locate the required file or files and calibration dataset stored therein (usually as one or more *FITS* extension).
3. The CIFs should be stored in a standard location within the *CALDB* (but can be stored in other locations if necessary).

In order to simplify identification of and access to calibration information by users and software, a CIF must contain location and instrumental parameter information for the calibration files associated with a given instrument.

2.2 *CALDB* Management

A research scientist on staff at the HEASARC serves as *CALDB* Manager and is responsible for maintenance of *CALDB* data, software tools, documentation and web pages, including installation of new CIFS and calibration files when available and announcing availability of new

calibration data to the user community. The HEASARC *CALDB* manager can be contacted at caldbhelp@athena.gsfc.nasa.gov. The name of the current *CALDB* manager at the HEASARC is given on the *CALDB* home page.

Users can manage local installations of the HEASARC *CALDB* themselves. The *CALDB* "supported missions" page has a list of all missions and instruments supported by the HEASARC *CALDB*, along with links to tar files of the calibration data for users to download and install. Updates and management of such local installations must be done by the user.

Management of a public *CALDB* (one accessed by more than a single user) for a given mission/instrument and modification of CIFs should be restricted. If a need arises for experimentation or customization, a *CALDB* can be copied to a user's local disk. To use a customized *CALDB*, a user would need to redefine the `$CALDB` environment to point to the local installation.

2.3 CIF Usage

The CIFs are in *FITS* format and hence accessible to users via standard *FITS* compatible software (like HEASoft and the `astropy.io.fits` python package). Both HEASoft and `astropy.io.fits` accept local and virtual file names, and so both can access information in CIFs stored locally or remotely on the internet.

Users can use standard *CALDB* software (the standalone `caltools` or lower-level subroutines in the `callib` subroutine libraries) to query, identify, manipulate and access CIFs and associated *CALDB* data. Users can also use other FTOOLS or other software packages to inspect, extract & manipulate the data file themselves, if desired.

2.4 Versioning

The *CALDB* is designed to preserve a record of previous calibration data, in order that the effects of updated calibrations can be compared to previous ones. To make such comparisons as easy as possible, the *CALDB* uses a simple versioning system. A given *CALDB* release is specified by a CIF which has a name of `caldb.indxYYYYMMDD`, where YYYYMMDD is the (4-digit) year, (2-digit) month and (2-digit) day corresponding to date of the *CALDB* release. The convention used by the HEASARC is to store these individual CIF files in an `index/` subdirectory of the *CALDB* for a mission/instrument.

Since the standard *CALDB* access software (see Sec. 5) assumes that the current CIF is at the main directory of the *CALDB* for the mission and instrument, for example

`$CALDB/data/<mission>/<instrument>`,

the CIF in the main mission/instrument directory is a relative symbolic link to the latest CIF

in the `index` subdirectory. For example, for the Swift XRT

`$CALDB/data/swift/xrt/caldb.indx` → `index/caldb.indx20210915`

where `caldb.indx20210915` is the most recent version of the CIF as of this writing.

3 CALIBRATION INDEX FILE FORMAT

A Calibration Index File is a FITS file with a null primary array and a single BINTABLE extension. Each row within the extension refers to a single calibration dataset within a single extension in a single calibration file. The values inserted into the various columns of the BINTABLE originate either from the calibration file or are supplied by the installation software (`crcif`, `udcif`) as noted below. As stated above, when taken together each row of the CIF contains all the information for analysis software to identify, locate, and access a required calibration dataset.

The structure of a standard CIF is:

```
 No. Type     EXTNAME      BITPIX Dimensions(columns)      PCOUNT  GCOUNT

  0  PRIMARY                   8     0                          0    1
  1  BINTABLE CIF              8     947(18) 62                 0    1

     Column Name                 Format     Dims       Units     TLMIN  TLMAX
     1 TELESCOP                    10A
     2 INSTRUME                    10A
     3 DETNAM                      20A
     4 FILTER                      10A
     5 CAL_DEV                     20A
     6 CAL_DIR                     70A
     7 CAL_FILE                    40A
     8 CAL_CLAS                    3A
     9 CAL_DTYP                    4A
    10 CAL_CNAM                    20A
    11 CAL_CBD                     630A70
    12 CAL_XNO                     I
    13 CAL_VSD                     10A
    14 CAL_VST                     8A
    15 REF_TIME                    D
    16 CAL_QUAL                    I
    17 CAL_DATE                    10A
    18 CAL_DESC                    70A
```

3.1 The INDEX BINTABLE Extension

The *CALDB* index for a given mission/instrument is located in the first extension of the CIF.

3.1.1 Extension Header

Besides the standard FITS keywords, the header contains the following (mandatory) keywords/values:

- `EXTNAME = 'CIF     '`, the name of the extension
- `CIFVERSN` - the OGIP version number of the FITS format in use, as specified by the most recent version of this document (currently `CIFVERSN = '1.1     '`)

3.1.2 *FITS* BinTable Column Description

Table 1 defines the table columns in the CIF bintable extension (the first extension) of a calibration index file.

Table 1: Definitions of the columns in a Calibration Index File

Column Name	Data Type	Description	Origin
TELESCOP	10-byte CHARACTER string	The name of the telescope or mission to which the calibration data applies. See CAL/GEN/92-011 for a summary of available strings.	TELESCOP keyword in the primary header of the indexed calibration file.
INSTRUME	10-byte CHARACTER string	The name of the instrument on the telescope used for the observation to which the calibration data applies.	The INSTRUME keyword in the primary HEADER of the indexed calibration file.
DETNAM	20-byte CHARACTER string	Specifies the detector to which the calibration applies, if needed. If not needed, set to NONE.	The DETNAM keyword in the appropriate extension header of the indexed calibration file extension.
FILTER	10-byte CHARACTER string	The filter in use for the observation data to which the calibration data applies, if needed. A filter is defined as an optical element placed in front of a detector to alter the response/sensitivity of the detector. Value set to NONE if not needed for calibration.	The FILTER keyword in the extension header of the indexed calibration file extension.
CAL_DEV	20-byte CHARACTER string	Either ONLINE if the data are available on a mounted volume, or OFFLINE if the data have been stored on an unmounted volume. CAL_DEV applies to all extensions in the indexed calibration file.	Defined by the creator of the CIF.

Continued on next page

Table 1 – continued from previous page

Column Name	Data Type	Description	Origin
CAL_CLAS	3-byte CHARACTER string	Either BCF, CPF or PCF if the indexed calibration file is a BASIC CALIBRATION FILE, a CALIBRATION PRODUCT FILE or a PRIMARY CALIBRATION FILE. See CAL/GEN/92-003 for discussion of these classes.	Obtained from the value of the CCLSmmmm keyword in the calibration file
CAL_DIR	70-byte CHARACTER string	Specifies the sub-directory location of the indexed calibration file relative to $CALDB/data/telescope/instrume/CAL_CLAS.	Defined by the creator of the *CALDB*
CAL_DTYP	4-byte CHARACTER string	Specifies whether the calibration data consists of real data (DATA), or virtual data (VIRTUAL)	Defined by the creator of the CIF
CAL_CNAM	20-byte CHARACTER string	Specifies the OGIP codename which describes the type of calibration data in the calibration file extension	Obtained from the CCNMxxxx keyword in the indexed calibration file extension
CAL_CBD	A 9 element array of 70-byte CHARACTER strings	Specifies the instrumental parameter boundaries for which the indexed calibration extension is appropriate	Derived from the CBDnxxxx keyword in the header of the indexed calibration file extension
CAL XNO	2-byte integer	Specifies the extension number of the indexed calibration file extension. An value of 0 (zero) is used to specify the primary image header.	Derived from the extension number in the indexed *FITS* calibration file.
CAL_QUAL	2-byte INTEGER	Specifies the validity of the indexed calibration data. A value of 0 (zero) means that the indexed calibration data are considered valid. A non-zero positive value indicates that the data should generally not be used to calibrate observations.	Defined by the creator of the CIF
CAL_VSD	10-byte character string	Gives the UTC start date when the indexed calibration is valid, in YYYY-MM-DD format.	Specified by the creator of the CIF
CAL_VST	8-byte character string	Gives the UTC start time when the indexed calibration is valid, in HH:MM:DD format.	Specified by the creator of the CIF
REF_TIME	8-byte REAL	The CAL_VSD and CAL_VST date and time converted to MJD	Defined from CAL_VSD and CAL_VST
CAL_DATE	10-byte character string	Date on which the file was installed in the *CALDB*, in YYYY-MM-DD formats	Defined by the creator of the CIF
CAL_DESC	70-byte CHARACTER string	Brief description the the data stored within the indexed extension	Defined by the creator of the CIF

4 MANDATORY KEYWORDS FOR CALIBRATION FILE EXTENSION HEADERS

As noted in Section 3.1.2, in order to include a calibration data file in a CIF, the following keywords are mandatory in the header of any FITS extension containing calibration data:

- `TELESCOP` - the name of the satellite/mission.
- `INSTRUME` - the name of the instrument.
- `DETNAM` - the name of the specific detector (applicable only when the value of the `INSTRUME` keyword is insufficient to uniquely specify the necessary information)
- `FILTER` - the name of the filter in use (not required for instruments without a moveable filter, or calibration datasets for which the filter information irrelevant).
- `CCLS`*xxxx* - the OGIP-class of this calibration file.
- `CDTP`*xxxx* - the code denoting whether the extension contains real or vitual data.
- `CCNM`*xxxx* - the OGIP codename of the extension to be used within CIF to describe the contents (for downstream software). This keyword is not mandatory in the case of PCFs.
- `CBD`*nxxxx* - an array of strings (with *n* arbitrary integers between 1 & 9) giving the parameter limitations of the dataset (*eg.* energy range, off-axis angles *etc.*) used within the CIF to further describe the contents for downstream software (in association with the value of the `CCNM`*xxxx* keyword).
- `CVSD`*xxxx* - the UTC date when this calibration data should first be used.
- `CVST`*xxxx* - the UTC time on the day CVSD*xxxx* when this calibration data should first be used.
- `CDES`*xxxx* - a string giving a brief descriptive summary of this dataset.

where *xxxx* is a number of the form `0001`, `0002`, `0003` *etc.* These keywords are further described, along with their allowed values in CAL/GEN/92-011, available on-line as pdf and html.

We recommend that the TELESCOP and INSTRUME keywords be present in the primary header of the file as well as in each extension. It is the responsibility of those supplying the files for inclusion in the *CALDB* to ensure that all other relevant keywords are present and correct, and that appropriate `COMMENT` and/or `HISTORY` keywords are supplied to identify and describe the dataset. It is also the responsibility of suppliers of calibration datasets to provide all necessary documentation concerning the origin, description, use and limitations of each dataset to the manager of the HEASARC *CALDB*. Send e-mail to caldbhelp@athena.gsfc.nasa.gov for more information.

5 ASSOCIATED SOFTWARE TASKS & ROUTINES

HEASoft `caltools` software includes tasks to create and access CIFs and the information indexed therein. The caltools are freely available to users as source code and pre-compiled binaries for common operating systems and are based on `CFITSIO`, a standard, widely-used library of C and Fortran subroutines for reading and writing *FITS* data files.

5.1 Creating & Modifying Calibration Index Files

There are two main `caltools` tasks used to create and modify calibration index files: `crcif` and `udcif`.

- `crcif` ("create CIF") is used to create a blank Calibration Index File with no rows. By default the CIF's name will be caldb.indx, however this can be changed by using the filename parameter on the command line. It is recommended that this file be renamed to include version information in YYYYMMDD format, for example `caldb.indx20220420`, and the file be placed in the `index` subdirectory of the main directory for the calibration data for the mission/instrument.
- `udcif` ('update CIF')is used to include extensions in a single calibration file in the *CALDB* into a CIF. The routine checks the file for any extensions which include the required *CALDB* keywords (as given in Section 4 above). The required keywords from each valid extension in the calibration file are then extracted and included as a row in the CIF BINTABLE extension. Before the new entry is written to the CIF, all other CIF entries are checked to see if the new entry will duplicate another dataset. One entry in the CIF is a duplicate of another if all calibration parameters (calibration code, instrument, telescope, filter, detector, and calibration boundaries) are the same. If a duplicate is found, the user has the choice to set the quality for the previous entry in the CIF to a value of 5, which means the new entry will be accessed by the CALDB access software instead of the previous entry. In this way, new calibration datasets can replace older, outdated ones in the CIF without losing the history of prior calibrations. Users also have the option of keeping both datasets with a quality value of 0, which can be useful in rare circumstances. If the duplicate entry has the same file name and extension number as the file being added, the user is notified and the update stops.

When updating a CIF using `udcif`, in order to preserve a record of past calibrations, the standard practice is to make a copy of the latest CIF in the `index` directory and update the new copy. For example, suppose there's a new release of the Swift XRT *CALDB*, to be released on 2023/01/31. In the `index` subdirectory directory, the *CALDB* manager should copy the previous version of the CIF to a new CIF with a new YYYYMMDD version in the index directory:

```
% cd $CALDB/data/swift/xrt/index
% cp caldb.indx20210915 caldb.indx20230131
```

The *CALDB* manager should then update the `caldb.indx20230131` file with the new calibration data using `udcif`, and after that's completed change the symbolic link to point to the new file:

```
% cd $CALDB/data/swift/xrt
% ln -fs caldb.indx index/caldb.indx20230131
```

In general, the `udcif` update routine is sufficient for adding updated calibrations and invalidating older ones. There are special circumstances which can arise during *CALDB* maintenance when a user might desire to edit a CIF manually. This might include deleting an existing row from a CIF (not generally recommended), updating the quality flag for a row, correcting the CAL_VSD or CAL_VST values, or other changes. Calibration Index files can be edited with standard *FITS* file editors. The HEASoft tool `fv` provides a convenient way to edit any *FITS* file. Caution should be exercized when manually editing CIFS, however, since the contents and structure of the original CIF must be maintained (and checksums updated) in order to use the CIF to access the appropriate calibration data using standard `caltools`.

5.2 Finding Calibration Data

Users can search a CIF for calibration data for a given mission and instrument using the `quzcif` task in the `caltools`, a command-line interface to a *CALDB*. `quzcif` finds the rows in a CIF which meet the user-specified selection criteria. The filename field (with the complete directory path) and the extension number is printed for each dataset which satisfies the specified criteria.

If the user is accessing the HEASARC *CALDB* using remote access, the full URL of the files is returned. Note that quzcif can download remote files to the current working directory if the `quzcif retrieve` parameter is set to `yes`. For example, the following command would retrieve the response matrices for the NICER XTI using remote access to the HEASARC *CALDB*:

```
% quzcif mission=nicer instrument=xti detector=- filter=- codename=matrix \
          date=now time=now expr=- retrieve=yes
nixtirmfbase20170601v001.fits   2
nixtiref20170601v003.rmf   2

% ls nixtir*
nixtiref20170601v003.rmf nixtirmfbase20170601v001.fits
```

assuming that the `$CALDB` environment variable is set to

https://heasarc.gsfc.nasa.gov/FTP/caldb

5.3 Inquiry subroutines in the Calibration Library (`callib`)

The *CALDB* subroutine library, `callib`, includes low-level subroutines (in FORTRAN) which can be used in user-developed software to access calibration data from a *CALDB*:

- `gtcalf.f90`: This subroutine is used by the `quzcif caltool` task. It returns the location of calibration datasets located in the Calibration Database. Selection of the appropriate calibration data is based on the values of the arguments TELE, INSTR, DETNAM, FILT, CODENAM, STRTDATE, STRTTIME, STPDATE, STPTIME, EXPR passed to the subroutine. These arguments respectively describe the mission or telescope, instrument, detector, filter, type of dataset, start date & time, stop date & time and calibration boundaries for which the returned datasets should be valid. In addition to the arguments explicitly listed here, this routine also uses the values of the environment variables `$CALDB`, and `CALDBCONFIG`. See the Caldb user's guide for details on setting these environment variables. The maximum number of datasets to return is given by the MAXRET argument, which defaults to returning all the valid files in the index. Any datasets which meet the selection criteria are returned through the FILENAM and EXTNO arrays. Each element of the FILENAM array contains the complete system dependent path (including the filename) to the file where the calibration data resides. The corresponding element of the EXTNO array contains the FITS extension number of the calibration data within the file.
- `gtcalf2.f90`: an updated version of `gtcalf.f90` which allows a user to specify a particular version of a CIF to use.

6 INSTALLATION & USE AT REMOTE SITES

Documentation describing the installation, maintenance and use of CIFs is provided in the HEASARC *CALDB* library:

- within the OGIP: see CAL/GEN/92-014 (Zellar & George 1993)
- at remote sites: see CAL/GEN/92-015 (George & Zellar 1993a)

As stated above, in order to reduce disk-space and maintenance requirements, it is strongly recommended that remote users on clustered machines share a single local caldb, and that a single person be responsible for its maintenance.

ACKNOWLEDGMENTS

We thank the numerous people, both inside and outside the OGIP, who have contributed ideas and suggestions.

REFERENCES

Cotton, W.D. & Tody, D., 1992. In preparation.
George, I.M., 1992. *Legacy*, **1**, 56, CAL/GEN/91-001
(available online in pdf and html versions).
George, I.M. & Zellar, R., 1993a. In preparation. (CAL/GEN/92-015)
George, I.M., Zellar, R., & Pence, W., 1993. *OGIP Calibration Memo* CAL/GEN/92-011
(available online in pdf and html versions).
George, I.M., Zellar, R., & Yusaf, R., 1993a. *OGIP Calibration Memo* CAL/SW/93-004
(available online in pdf and html versions).
George, I.M., Zellar, R., & Yusaf, R., 1993b. *OGIP Calibration Memo* CAL/SW/93-005
(available online in pdf and html versions).
Griesen, E.W. & Harten, R.H., 1981. *Astron. Astrophys. Suppl.*, **44**, 371.
Grosbol, P., Harten, R.H., Greisen, E.W. & Wells, D.C., 1988. *Astron. Astrophys. Suppl.*, **73**, 359.
Harten, R.H., Grosbol, P., Griesen, E.W. & Wells, D.C., 1988. *Astron. Astrophys. Suppl.*, **73**, 365.
Mukai, K., 1992. *OGIP Memo* OGIP/92-010.
Pence, W., 1992. *Legacy*, **1**, 14..
Wells, D.C., Griesen, E.W. & Harten, R.H., 1981, *Astron. Astrophys. Suppl.*, **44**, 363.
Zellar, R. & George, I.M., 1992. In preparation. (CAL/GEN/92-014).