OGIP Calibration Memo CAL/GEN/92-019

THE OGIP FORMAT FOR EFFECTIVE AREA FILES

(EARVERSN = 1992a)

Ian M George & Ron S Zellar
Mail Code 668,
NASA/GSFC,
Greenbelt,
MD 20771.

Version: 1995 Jul 31

SUMMARY

Ths document describes the standard format adopted by the HEASARC for the storage of the effective area of an instrument as a function of energy, and position.
Intended audience: primarily HEASARC programmers & hardware teams.

Log of Significant Changes

Release Date	Sections Changed	Brief Notes
1992 Jul 24	First Draft	(within memo CAL/GEN/92-003)
1993 Oct 03	All	Separation from CAL/GEN/92-003
1993 Nov 25	All	Added HDUCLASn info
1994 Jul 24	2.2	Clarified CSYSNAME conventions
1994 Aug 09	All	General Review/up-dates
1995 Jan 19	All	Made compatible with `LaTeX2HTML` software
2004 Apr 01	All	Made compatible with tth

RELATED DOCUMENTATION

The following documents may also be of use:

- *BCF & CPF Calibration File Guidelines*
  CAL/GEN/92-003 (George & Zellar)
- *Calibration Index Files*
  CAL/GEN/92-008 (George, Pence & Zellar)
- *Mandatory FITS Keywords for Calibration Files*
  CAL/GEN/92-011 (George, Zellar & Pence)
- *Virtual Calibration Files*
  CAL/GEN/92-013 (George, Zellar & White)
- *The HEASARC Format for Vignetting Functions*
  CAL/GEN/92-021 (George & Zellar)
- *The HEASARC Format for Obscuration Factors*
  CAL/GEN/92-022 (George & Zellar)

1 Introduction

Historically, the term 'Effective Area' has been used somewhat ambiguously. Within the HEASARC caldb, the term is reserved exclusively to denote that of the telescope mirror/collimator assembly (only), but including any vignetting and obscuration effects applicable for off-axis angles of incidence[1].

1.1 Storage Options

In the general case an Effective Area calibration dataset consists of a 3-dimensional grid, $E_{ff}Area$, with 1 axis giving the photon energy (E), and 2 axes defining the position relative to the optical axis — invariably the off-axis angle (θ_{XMA}) & azimuthal angle (ϕ_{XMA}).

However, there are two basic methods in which this information can be stored within the caldb:

- **Method A**: as a single 3-dimensional dataset giving the Effective Area at each E,θ_{XMA},ϕ_{XMA} grid point, *i.e.* storing $E_{ff}Area(E, \theta_{XMA}, \phi_{XMA})$ directly.
- **Method B**: as a (1-dimensional) list of the **on-axis** Effective Area as a function of energy, $E_{ff}Area(E, 0, 0)$, with the information necessary to correct this data for off-axis positions (*i.e.* due to the effects of Vignetting & Obscuration) stored elsewhere, usually in separate files (see Section 1.5).

Method A provides the clear advantage to users and s/w requiring an off-axis Effective Area that all the necessary information is contained within a single dataset. However, this method has the potential disadvantage that the Total Vignetting function (including the effects of vignetting and obscuration) has been folded in, requiring the entire dataset to be recreated if (say) the Total Vignetting function only is improved in-flight. By the same token Method B has a disadvantage to users and s/w requiring an off-axis Effective Area that additional calibration information is required – namely the Total Vignetting function, $Vignet$, appropriate for the desired off-axis position and energy range, *i.e.*

$$\begin{aligned} E_{ff}Area(E, \theta_{XMA}, \phi_{XMA}) &= E_{ff}Area(E, 0, 0) \\ &\quad \times Vignet(E, \theta_{XMA}, \phi_{XMA}) \end{aligned} \tag{1}$$

However, since the on-axis Effective Area is sometimes known to a higher accuracy than the Total Vignetting function, the advantage of Method B is that the isolation of the latter component means that updates can be made without the necessity of updating the (on-axis) Effective Area BCF. The pros and cons of the two methods are further discussed in Section 1.2.

[1]The result of multiplying the Effective Area of the mirror/collimator assembly by the transmission due to any filters & detector windows along the optical path, and by the efficiency of the detector, is referred to as the Spectral Response of the instrument within the HEASARC caldb

The data formats described in Section 2 are able to accomodate both Methods A& B, with the CBD$nxxxx$ keyword(s) providing downstream software with the necessary information as to whether further calibration i/p (*i.e.* off-axis Vignetting & Obscuration factors) is required.

1.2 Dataset Origins & Storage Recommendations

The construction, format used (within the limitations discussed here) and delivery of the data to the HEASARC (including any updates) is the responsibility of the h/w teams and/or GOF. However, below, are the recommendations of the HEASARC calibration team based on their experience.

General
For completeness and maximum flexibility, in the case of instruments for which the calibration is well understood and stable, the HEASARC would prefer that an effective area dataset is stored using Method A, but that the associated vignetting and obscuration functions are also stored in CCNM$xxxx$ = VIGNET and CCNM$xxxx$ = OBSCFACT datasets.

Pre-launch
Prior to launch, the on-axis effective area is usually accurately measured at a (limited) number of photon energies during ground calibration experiments, and combined with a theoretical model to produce a good estimation of the effective area at all energies within the required band. The effects of vignetting and obscuration at off-axis positions are also measured at a (usually more limited) number of photon energies during such experiments and/or combined with theoretical (*e.g.* ray-tracing) models to produce the off-axis correction factors. Thus in most cases, Method B (Section 1.1) is the natural method of storing the necessary calibration datasets at this time.

Post-launch
The effective area of the optics **only** cannot be measured in-orbit. Instead, observations of standard cosmic sources (*e.g.* the Crab) combined with spectral modelling enables the Spectral Response of the instrument (*i.e.* the effective area of the optics multiplied by the transmission of any filters & windows and by the detector efficiency as a function of energy) to be determined. The results of such calibration observations should be stored as a CCNM$xxxx$ = SPECRESP dataset. However, should such measurements reveal that a discrepancy with previous calibrations which is identified with (or interpreted as) a mis-calibration of the effective area of the optics, h/w & GOF teams are urged to isolate and also supply an updated Effective Area dataset to the HEASARC. Since such discrepancies are often the result of uncertainties in the Vignetting function, Method B (Section 1.1) is recommended for initial in-flight datasets also. Only when the calibration is considered to be well understood and stable, is it recommended that the Effective Area information be reformatted to use Method A. When such a time occurs should be the descision of the h/w & GOF teams, and a new Effective Area BCF using Method A (and any new CCNM$xxxx$ = VIGNET and CCNM$xxxx$ = OBSCFACT BCFs necessary) be delivered to the HEASARC.

1.3 Dataset vs Task Summary

Due to the complexity of the Effective Area of an instrument as a function of energy (as the result of sharp discontinuities due to atomic processes *etc.*), such a calibration dataset is not easily parameterized. Thus, whilst theoretically possible, it is recommended that an Effective Areas dataset is **not** described by a virtual calibration file.

1.4 Software Considerations

Data Files:
In both Methods A & B outlined in Section 1, interpolation between the θ_{XMA}, ϕ_{XMA} grid points is usually required. By default, downstream software will use a simple 2-dimensional **linear** interpolation when calculating the Effective Area between θ_{XMA}, ϕ_{XMA} grid points. Thus the θ_{XMA}, ϕ_{XMA} grid should be of sufficient resolution to enable this to be a reasonable approximation.

As discussed in the general guidelines for calibration files (CAL/GEN/92-003), it is strongly recommended that the energy grid is of sufficient resolution and carefully chosen such that interpolation of this parameter is not required. However, in cases where interpolation is required, as simple 1-dimensional linear interpolation will be performed (which will clearly be inaccurate close to sharp features).

Virtual Files:
Not applicable (see Section 1.3).

1.5 Relationships to Other Calibration Datasets

Downstream s/w should assume further calibration input is required for an Effective Area dataset under the following conditions:

- *condition:*
  an off-axis effective area is required, and CBDn0001 = THETA(0)*unit* and CBDm0001 = PHI(0)*unit* keywords are present (where n and m are integers, and *unit* is any string) indicating the contents of the Effective Area BCF datset is only applicable on-axis (see CAL/GEN/92-003).
  requirement:
  a CCNM$xxxx$ = TVIGNET calibration dataset (or equivalently CCNM$xxxx$ = VIGNET and CCNM$xxxx$ = OBSCFACT datsets) giving the Total Vignetting function at the requested off-axis position.
  alternative:
  If no such calibration sets are available within the CALDB, then either the user should

be informed and the s/w task stopped, or (if appropriate) a CCNM*xxxx* = SPECRESP dataset searched for within the CALDB with the corresponding implications to the subsquent operation of downstream s/w.

An Effective Area dataset is used in the construction of the following calibration datasets:

- A CCNM*xxxx* = SPECRESP dataset, containing the total spectral response of an instrument (see CAL/GEN/92-002 and CAL/GEN/92-002a).

2 Effective Data File Formats

The HEASARC FITS Working Group (HFWG) Header-Data Unit (HDU) keywords and values for this type of dataset are:

- `HDUCLASS = 'OGIP'`
  - the name of the organization that defined this file format.
- `HDUDOC   = 'CAL/GEN/92-019'`
  - the name of the document describing the format (*ie* this document)
- `HDUCLAS`n
  - giving the HDUCLAS hierarchy for this format.
  - `HDUCLAS1 = 'RESPONSE'`
  - `HDUCLAS2 = 'EFF_AREA'`

These are valid for all datasets described in this section, and should be present in the **header of the extension** containing the $E_{ff}Area$ dataset.

2.1 Summary of Effective Area file formats versions

The following versions of file formats for an $E_{ff}Area$ dataset have been defined:

- `HDUVERS  = '1.0.0'` (Section 2.2)
  This format is currently still **VALID**.

2.2 The Effective Area Extension (HDUVERS = '1.0.0')

Description:
One file for each telescope containing a single BINTABLE FITS extension. The BINTABLE only has a single row, using arrays for the 5 necessary columns.

Note: this format was formally known as EARVERSN = '1992a', and is still occasionally refered to as such.

Extension Header
Beyond the standard FITS keywords required, and the HDU keywords/values given in Section 2, the following keywords/values are mandatory:

- `TELESCOP` - the name of the satellite/mission.
  Allowed values are given in CAL/GEN/92-011.
- `INSTRUME` - the name of the telescope mirror/detector of collimator/detectory assembly. Allowed values are given in CAL/GEN/92-011. In the case of collimated detectors, this is usually the name of the detector itself.
- `HDUVERS  = '1.0.0'` - giving the version of the format.
- `TDIM`*nnn* - the number of elements and ordering (see CAL/GEN/92-003) of each multi-dimensional array.
  Only the `EFFAREA` column here (with $nnn = 5$ in the example below).
- *i*`CTYP`*nnn* - The axis labels for dimension i (i = 1, 2, 3) of the `EFFAREA` column. In the example given below, *nnn=5* and
  - `1CTYP5  = 'ENERGY'`
  - `2CTYP5  = 'COORD-1'`
  - `3CTYP5  = 'COORD-2'`

  (see CAL/GEN/92-003 for further details).
- `CREF`*nnn* = The column referencing keyword for each multi-dimensional array.
  Only the `EFFAREA` column here (with $nnn = 5$) giving:
  - `CREF5    = '(ENERG_LO:ENERG_HI,THETA,PHI)'`

  in the example below.
- `CSYSNAME` - the spatial coordinate system in use (see CAL/GEN/92-003)
  (`CSYSNAME = 'XMA_POL'` is assumed in the example below)

and the following keywords/values are mandatory for CIF purposes (see CAL/GEN/92-011):

- `CCLS0001= 'BCF'` - the OGIP class of this calibration file
- `CDTP0001= 'DATA'` - the OGIP class of the data type
- `CCNM0001= 'EFF_AREA'` - the OGIP codename for the contents
- `CBD`n`0001` - the parameter-space limitations of the dataset (see below)
- `CVSD0001` - calibration validity start date
- `CVST0001` - calibration validity start time
- `CDES0001` - a descriptive string of the calibration dataset

and the following are optional to supply further information:

- `EARVERSN= '1992a'` - the OGIP version of the FITS format in use

Data Format:
The data within the extension is organised as a BINTABLE with the following columns:

1. E_{low}, a fixed-length REAL vector (array, each element within which is 4-byte) containing the lower energy bounds of the energy bins.
   The FITS column name is **ENERG_LO**.
   The recommended units are keV.
2. E_{high}, a fixed-length REAL vector (array, each element within which is 4-byte) containing the upper energy bounds of the energy bins.
   The FITS column name is **ENERG_HI**.
   The recommended units are keV.
3. θ_{XMA}, a fixed-length REAL vector (array, each element within which is 4-byte) containing the off-axis angles.
   The FITS column name is **THETA** (but see below)
   The recommended units are arcmin.
4. ϕ_{XMA}, a fixed-length REAL vector (array, each element within which is 4-byte) containing the azimuthal angles.
   The FITS column name is **PHI** (but see below).
   The recommended units are arcmin.
5. $E_{ff}Area$, a fixed-length REAL vector (array, each element within which is 4-byte) containing the effective area at each $E, \theta_{XMA}, \phi_{XMA}$ grid point.
   The FITS column name is **EFFAREA**.
   The order of data storage is $E_{ff}Area(E, \theta_{XMA}, \phi_{XMA})$, where E represents the E_{low} and E_{high} arrary (see below).
   The recommended units are cm^2.

Table 1: Summary of the OGIP format for Effective Areas (EARVERSN = 1992a).

Extension **to** *(filename).(ext)*

HDUCLASS: OGIP
HDUDOC: CAL/GEN/92-019
HDUVERS: 1.0.0
HDUCLAS1: RESPONSE
HDUCLAS2: EFF_AREA
EXTNAME : EFFECTIVE AREA (suggested, not required)
Description: Effective areas (including vignetting and obscurration effects) as a function of energy and off-axis & azimuthal angle.
An alternate spatial coordinate frame may also be used (see text).
Optional columns containing the statistical and systematic error arrays are not shown.
Format: BINTABLE

column 1	2	3	4	5
contents				
Low energy bounds	**High energy bounds**	**Off-axis angles**	**Azimuthal angles**	**Effective Areas**
E_{low}	E_{high}	θ_{XMA}	ϕ_{XMA}	$E_{ff}Area$
format of each column				
4-byte real array	4-byte real array	4-byte real array	4-byte real array	4-byte real array
total number of elements per row				
i	i	j	k	$i \times j \times k$
column name				
ENERG_LO	ENERG_HI	THETA	PHI	EFFAREA

These are summarized in Table 1.

Points to Note & Conventions

- The ordering of the columns is of course arbitrary, however that used here is recommended.
- The rules and conventions concerning the energy grid (E_{low} & E_{high}) given in CAL/GEN/92-003 apply.
- An alternate spatial coordinate frame may be used, in which case
  - the value of the CSYSNAME keyword should be replaced by the appropriate string listed in CAL/GEN/92-003.
  - and/or (if necessary) the THETA & PHI column names replaced by more suitable alternatives if a different coordinate notation is employed. In this case the CSYSNAME keyword is mandatory and should give the column names used (see CAL/GEN/92-003).
- The parameter-space limitations on the dataset involving the following *pname* strings are recommended to be specified via the CBDn0001 keywords (see CAL/GEN/92-003):
  - *pname* = THETA - giving the range of off-axis angle for which the dataset is valid;
  - *pname* = PHI - giving the range of azimuthal angle for which the dataset is valid;

  (or corresponding alternate values of *pname* if a different coordinate notation is employed) along with any other limitations the authors of the dataset consider necessary.
  For CIF purposes (and hence for downstream s/w) datasets which contain only the on-axis Effective Area **must** include CBDn0001 = `'THETA(0)arcmin'` and CBDm0001 = `'PHI(0)arcmin'` keywords (where n and m are integers). The presence of these keyword values will be used to indicate to downstream s/w that additional calibration i/p is required if an off-axis Effective Area is required (see Section 1.4).
- Datasets in which $E_{ff}Area$ is independent of either spatial coordinate should **NOT** contain the corresponding column. It is recommended that a COMMENT card is used within the header to explain this fact to human readers (eg see Section 5.1).
- Alternative physical units are allowed for all columns of the table as long as they conform to the rules given in OGIP/93-001. The same is true for the physical units associated with the CBDn0001 keywords.
- The order of $E_{ff}Area(E, \theta_{XMA}, \phi_{XMA})$ whereby energy parameters changes fastest, and the azimuthal angle parameter slowest (see CAL/GEN/92-003) was chosen to facilitate access for the most common applications: interpolation in θ_{XMA}-space of $E_{ff}Area$ vs E_{low}, E_{high} arrays. This ordering is further confirmed by the value of the mandatory TDIM*nnn* and *i*CTYP*nnn* keywords (where *nnn* is the column number, and *i* the axis number). The rules and conventions governing these keywords are given in CAL/GEN/92-003 (see also Section 5).

- The optional arrays containing the 1σ statistical error associated with each element of $E_{ff}Area$ (if required) should be contained in additional columns named **STAT_MIN** (for the negative error) and **STAT_MAX** (for the positive error).
  Similarly, the optional arrays containing the 1σ fractional systematic error associated with each element of $E_{ff}Area$ (if required) should be contained in additional columns named **SYS_MIN** (for the negative error) and **SYS_MAX** (for the positive error).
  The rules and conventions governing such arrays (if present) are given in CAL/GEN/92-003. These arrays are provided here for completeness, and rarely either provided by the h/w teams or used by downstream s/w.

3 Virtual File Formats & Allowed Standalone Tasks

As noted in Section 1.3, given the difficulty parameterizing the Effective Area in energy-space, it is not recommended that such datasets are stored as Virtual Calibration Files.

4 Related Software

The following list of subroutines/tasks are available:

- FORTRAN subroutine `wtear1.f` (callib)
  writes an EARVERSN = 1992a dataset (Section 2.2)

5 Example FITS headers

Below are several examples of files currently available within the HEASARC caldb. Note that the authors of datasets are encouraged to supply copious COMMENT cards to aide human readers.

5.1 *ROSAT*

Follows is the header from an extension containing an EARVERSN=1992a dataset. The effective area data is stored in column 4 and is a function of 729 energies and 14 off-axis angles, THETA. There is no PHI dependence for this dataset. From the *i*CTYP4 (and also implied by the TDIM4) keyword, it can be seen that ENERGY is the 1st axis/dimension of the EFFAREA data array, and THETA the 2nd axis/dimension.

```
XTENSION= 'BINTABLE'           / binary table extension
BITPIX  =                    8 / 8-bit bytes
NAXIS   =                    2 / 2-dimensional binary table
NAXIS1  =                46712 / width of table in bytes
NAXIS2  =                    1 / number of rows in table
PCOUNT  =                    0 / size of special data area
GCOUNT  =                    1 / one data group (required keyword)
TFIELDS =                    4 / number of fields in each row
TTYPE1  = 'ENERG_LO'           / Lower boundaries of energy bins
TFORM1  = '729E    '           / data format of the field: 4-byte REAL
TUNIT1  = 'keV     '           / physical unit of field
TTYPE2  = 'ENERG_HI'           / Upper boundaries of energy bins
TFORM2  = '729E    '           / data format of the field: 4-byte REAL
TUNIT2  = 'keV     '           / physical unit of field
TTYPE3  = 'THETA   '           / Spatial coord grid: dimension 1
TFORM3  = '14E     '           / data format of the field: 4-byte REAL
TUNIT3  = 'arcmin  '           / physical unit of field
TTYPE4  = 'EFFAREA'            / Effective Area dataset
TFORM4  = '10206E  '           / data format of the field: 4-byte REAL
TUNIT4  = 'cm**2   '           / physical unit of field
EXTNAME = 'EFFECTIVE AREA'     / name of this binary table extension
HDUCLASS= 'OGIP    '           / format conforms to OGIP standard
HDUDOC  = 'CAL/GEN/92-019'     / format definition document
HDUVERS = '1.0.0   '           / Version of family of formats
HDUCLAS1= 'RESPONSE'           / dataset relates to instrument response
HDUCLAS2= 'EFF_AREA'           / dataset is an effective area
CSYSNAME= 'XMA_POL '           / spatial coord system used in this dataset
1CTYP4  = 'ENERGY'             / 1st axis of EFFAREA is energy
2CTYP4  = 'COORD-1'            / 2nd axis of EFFAREA is coord-1
```

```
CREF4   = '(ENERG_LO:ENERG_HI,THETA)'
TELESCOP= 'ROSAT   '           / mission/satellite name
INSTRUME= 'XRT     '           / instrument/detector name
FILTER  = 'NONE    '           / filter in use
COMMENT   Dataset assumed to be independent of PHI
TDIM4   = '(729,14)'           / Ordering of n-d EFF_AREA array
1CTYP4  = 'ENERGY  '           / Axis of 1st dimension of EFF_AREA array
2CTYP4  = 'THETA   '           / Axis of 2nd dimension of EFF_AREA array
EARVERSN= '1992a   '           / OGIP classification of FITS format
HISTORY    Extension written by WTEAR1 1.1.0
COMMENT
COMMENT   The following keywords are required for the OGIP CALDB
CCLS0001= 'BCF     '           / OGIP class of calibration file
CDTP0001= 'DATA    '           / OGIP type of dataset (DATA, TASK etc)
CCNM0001= 'EFFAREA '           / OGIP codename for this type of cal file
CVSD0001= '01/06/90'           / Dataset validity start date (UTC)
CVST0001= '00:00:00'           / Dataset validity start time (UTC, of day CVSD)
CDES0001= 'XRT Effective Area (version 1); 729 energies vs 14 off-axis angles'
CBD10001= 'THETA(0-60.0)arcmin'/ dataset parameter boundary
CBD20001= 'PHI(0-360)deg'      / dataset parameter boundary
CBD30001= 'ENERG(0.0546-3.01)keV'/ dataset parameter boundary
COMMENT
COMMENT   NOTES: 1994 Jul 21       (Ian M George, HEASARC)
COMMENT   ------------------
COMMENT   This dataset was converted to OGIP FITS format by
COMMENT   Ian M George (HEASARC) from the ASCII file
COMMENT          area_b_1.asc
COMMENT   supplied by Steve Snowden (ROSAT GOF, NASA/GSFC).
COMMENT   The area_b_1.asc file consists of the (total) spectral
COMMENT   response for PSPCB. The current dataset was therefore created
COMMENT   by DIVIDING the area_b_1.asc dataset by:
COMMENT      the gas efficiency:            pspc_v1.gas_eff
COMMENT      the window transmission:       pspcb_v1.wind_trans
COMMENT
COMMENT   The area_b_1.asc dataset has also been converted to
COMMENT   an OGIP FITS format as is stored in file
COMMENT      pspcb_v1.spec_resp
COMMENT
END
```

REFERENCES

Information regarding on-line versions of any of the following references with an OGIP Memo number (*ie* documents starting `OGIP/..` or `CAL/..`) can most easily be found via the WorldWide Web by following the links from the URL:
`/docs/heasarc/caldb/caldb_docs_index.html`

Most OGIP Calibration Memos of general community interest will eventually appear as articles in *Legacy*, but are also available on request from The Office of Guest Investigator Programs, Code 668, NASA/GSFC, Greenbelt, MD 20771, USA.

George, I.M. & Zellar, R.S., 1992.
(CAL/GEN/92-003)
George, I.M., Arnaud, K.A., Pence, W. & Ruamsuwan, L., 1992. *Legacy*, **2**, 51.
(CAL/GEN/92-002)
George, I.M., & Arnaud, K.A., 1992.
(CAL/GEN/92-002a)
George, I.M., Zellar, R.S. & Pence, W., 1992.
(CAL/GEN/92-011)
George, I.M., Pence, W. & Zellar, R.S., 1992.
(CAL/GEN/92-008)
George, I.M., Zellar, R.S. & White, N.E., 1992.
(CAL/GEN/92-013)
George, I.M. & Zellar, R.S., 1992.
(CAL/GEN/92-021)
George, I.M. & Zellar, R.S., 1992.
(CAL/GEN/92-022)