eboxdetect

January 27, 2025

Abstract

Sliding box EPIC source searching using both local or global background.

1 Instruments/Modes

Instrument	Mode
EPIC MOS:	IMAGING
EPIC PN:	IMAGING

2 Use

pipeline processing	yes
interactive analysis	yes

3 Description

The task **eboxdetect** is used to perform source detection on individual images or on sets of images from different energy bands and/or different EPIC instruments. If source detection on multiple images is performed, the input images and corresponding exposure images, background images, and detection masks have to be supplied as lists of file names. All input images and accompanying exposure images, background images, and detection masks must have identical orientation and binning and must be supplied in a consistent order. The task is designed for a maximum of five energy bands and three EPIC instruments. It is technically possible to run **eboxdetect** for up to six energy bands, but users are encouraged not to use more than five because of the low count numbers in narrow-band images.

eboxdetect has two different modes of operation: (1) local detection and (2) map detection. **eboxdetect** is first run in local mode. The local-mode source lists serve as input for the task **esplinemap** which creates background maps (one per energy band for each EPIC instrument). They are then used to run **eboxdetect** in map mode. The boolean input parameter `usemap` controls whether the program operates in map mode or in local mode.

1. **Local detection mode:** The purpose of the local detection step is to provide an input list of source positions for the task **esplinemap**, which then constructs a background map from the non-source locations. Source counts are accumulated from a 3x3 or 5x5 pixel window (controlled by the

parameter `boxsize`) and the background is determined from the surrounding 40 (7x7 pixel window minus 3x3) or 56 pixels (9x9 pixel window minus 5x5), respectively. Detection of moderately extended objects (up to several times the PSF size) is achieved by searching the image in up to three consecutive detection runs, each doubling the pixel size. The input image pixel size resampled from the original detector pixels is chosen such that the detection window corresponds to the size of the on-axis PSF. It is thus required that the PSF is at least moderately oversampled (which is the case for both pn and MOS CCD EPIC data). It is recommended to use a detection threshold of `likemin`=8 to provide a complete source list as input for **esplinemap**.

2. **Map detection mode:** Source counts are accumulated from a 3x3 or 5x5 pixels window with the option to use a position and energy dependent PSF weighted filter (matched filter; not yet implemented). Detection of extended objects is again achieved by doubling the pixel sizes in up to three consecutive detection runs. In map detection mode the background is taken from the background maps determined by **esplinemap**, resulting in an improved detection sensitivity as compared to the local detection step. If the map detection run is used as input for the task **emldetect**, it is recommended to use a somewhat lower detection threshold than for **emldetect** itself, e.g. `mlmin`=10 for emldetect and `likemin`=8 for **eboxdetect**.

In both modes, source searching is only performed in the area of the images which is marked by an optional detection mask created by task **emask**. Use of a detection mask is controlled by the boolean input parameter `withdetmask`. If true, one detection mask for each EPIC instrument must be supplied (parameter `detmasksets`). Parameter `withexpimage` controls the optional use of exposure maps as created by task **eexpmap**. The value of the exposure image at the location of each detected source is then used to calculate vignetting- and deadtime-corrected source count rates. If the parameter `withexpimage` is set to true, one exposure image for each energy band of each EPIC instrument must be supplied (parameter `expimagesets`). If no exposure images are supplied, raw count rates are calculated by deriving the exposure information from the good time intervals.

Background-subtracted source counts are calculated by applying correction factors to account for the respective fractions of source counts falling in each source and background area. The respective off-axis angle dependent correction factors are calculated using the medium-accuracy PSF. The following equations describe the PSF correction of source and background counts as implemented in the code of the task.

n = detection box size

Enboxed energy fraction in source box:

$$\alpha = \sum_{n \times n} \mathrm{PSF}$$

Fraction of source counts in background counting area:

$$\beta = \sum_{(n+4)\times(n+4)} \mathrm{PSF} - \sum_{n \times n} \mathrm{PSF}$$

Raw box counts:

$$\texttt{BOX\_CTS} = \sum_{n \times n} \mathrm{image}$$

Raw background map:

$$\texttt{BG\_RAW} = (\sum_{(n+4)\times(n+4)} \text{image} - \sum_{n\times n} \text{image})/((n+4)^2 - n^2)$$

In local detection mode:

PSF corrected and background subtracted source counts:

$$\texttt{SCTS} = \frac{\texttt{BOX\_CTS} - \texttt{BG\_RAW} \cdot n^2}{\alpha - \beta \cdot n^2/((n+4)^2 - n^2)}$$

PSF corrected background map:

$$\texttt{BG\_MAP} = \texttt{BG\_RAW} - \texttt{SCTS} \cdot \beta/((n+4)^2 - n^2)$$

$$err_{\text{bg}} = \sqrt{\texttt{BG\_RAW} \cdot ((n+4)^2 - n^2)}/((n+4)^2 - n^2)$$

Error of source counts:

$$\texttt{SCTS\_ERR} = \frac{\sqrt{\texttt{BOX\_CTS} + (n^2 * err_{\text{bg}})^2}}{\alpha - \beta \cdot n^2/(n+4)^2}$$

In map detection mode:

$$\texttt{SCTS} = (\texttt{BOX\_CTS} - \texttt{BG\_MAP} \cdot n^2)/\alpha$$

The resulting output source table contains one row per input image for each detected source, plus a number of summary rows containing the broad band results for each EPIC telescope and the combined results for all EPIC telescopes taken together. The individual source rows are identified through the column entries `ID_INST` and `ID_BAND` in the output table where `ID_INST` refers to the EPIC instrument (1: PN, 2: MOS1, 3: MOS2, 0: summary row) and `ID_BAND` is the energy band number as defined by the ordering in which the energy bands are given on the command line. An `ID_BAND` value of 0 again refers to the summary information. No summary rows are output if the source detection is only performed on a single input image.

The source table lists statistical errors for both count rates and source positions. Count rate errors are calculated by assuming Poissonian statistics (error = $\sqrt{\text{counts}}$) in both the source and background cells (if **eboxdetect** is run in local mode) and by applying standard error propagation. If run in map mode, the background taken from the spline background maps is assumed to be free of statistical errors. Positional errors are assumed to be equal to the standard deviation of the distribution of the counts in the detection cell. The errors of the derived parameters, such as count rates, fluxes, and source positions in celestial coordinates are derived from the count and image pixel positional errors, respectively.

Following the definition which was e.g. used by the ROSAT mission, detection likelihoods (per energy band and total) are given for each source in the form $L = -\ln p$ where p is the probability of Poissonian random fluctuation of the counts in the detection cell which would have resulted in at least the observed number of source counts. The value of p is calculated using the incomplete Gamma function $P(a, x)$ as a function of raw source counts and raw background counts in the detection box. See Press et al., Numerical Recipes, chapter 6.2 for the calculation of $P(a, x)$.

Table 1: Default band assignments of hardness ratios HR_i for the EPIC instruments

i	n	m
1	1	2
2	2	3
3	3	4

$$\texttt{LIKE} = -\ln p(\texttt{BOX\_CTS}, \texttt{BG\_RAW} \cdot n^2)$$

In the case of simultaneous detection runs over several energy bands, the `LIKE` values from each individual energy band are added and transformed to equivalent single band detection likelihoods using the incomplete Gamma function:

$$L = P(n, \sum_{i=1}^{n} L_i)$$

where n is the number of energy bands. A source is included in the output table if the equivalent single band detection likelihood exceeds the threshold given by the parameter `likemin`.

If detection over several energy bands is performed, up to three hardness ratios are calculated from the source counts in the individual bands. The hardness ratios are defined as follows:

$$HR_i = \frac{B_m - B_n}{B_m + B_n}$$

where B denotes the count rates in energy bands n and m, respectively. n and m are specified in the input parameter `hrdef` for each of the (up to) three hardness ratios. The default band assignment is given in the following table:

The band numbers n and m (output table column `ID_BAND`) are assigned to the individual bands by numbering the corresponding input images in the order in which they are given on the command line. It is therefore important that the ordering of the input images is consistent with the contents of `hrdef` to obtain meaningful hardness ratios.

Input images. With version 4.24, the parameter `obsmode` has been introduced to distinguish between input data from pointing observations (`obsmode`=pointing) and from slew-mode observations (`obsmode`=slew). Input data in detector coordinates are accepted from version 4.25 on and identified by their header keywords CTYPE1/2=DETX/Y. Only one instrument is allowed in this case. The columns `X_IMA`, `Y_IMA` of the output source lists are given in the same coordinate system as the input data. Two additional columns `DETX`, `DETY` hold the detector coordinates of the detections. Position errors are not converted into detector coordinates. If required, users can derive them from the `X_IMA_ERR`, `Y_IMA_ERR` columns. *No PSF correction* is calculated for data in detector coordinates. Use of the defaults – pointing observation, X/Y coordinates – is recommended.

4 Parameters

This section documents the parameters recognized by this task (if any).

Parameter	Mand	Type	Default	Constraints
imagesets	yes	filename	image.fits	

Names of EPIC fits images[1,2]

boxlistset	yes	filename	eboxlist.fits	

Name of output source list

bkgimagesets	no	filename	bkgimage.fits	

Names of spline background maps[1,2] (task **esplinemap**)

expimagesets	no	filename	expimage.fits	

Names of exposure images[1,2] (task **eexpmap**)

detmasksets	no	filename	detmask.fits	

Name of detection mask[3] (task **emask**)

mergedlistset	no	filename	mergedlist.fits	

Name of merged source list (not yet implemented)

usemap	no	boolean	false	

true: map mode; false: local mode

usematchedfilter	no	boolean	false	

Matched filter flag; not supported in current version

withdetmask	no	boolean	false	

Detection mask flag

withexpimage	no	boolean	false	

Exposure image flag

withoffsets	no	boolean	false	

Controls whether the offsets will be read from the merged (eident) source list – obsolete.

boxsize	no	integer	5	3 or 5

Size of detection box (3x3 or 5x5 pixels)

nruns	no	integer	3	1≤param≤4

Number of detection runs (detection box size is doubled after each run)

likemin	no	float	10.0	[1.0<param<50.0]

Minimum detection likelihood

ecf	no	float	1.0	[0.0<param<1000]

Energy conversion factors, given in units of 10^{-11} counts cm^2 erg^{-1}

hrdef	no	integer	1 2 2 3 3 4	0,10

Array of six integer numbers specifying the upper and lower energy band for each of (up to) three hard-

ness ratios

pimin	no	integer	2000	[1<param<30000]

Lower PI channel limits of input images[2] in eV. Up to six different energy bands are allowed per instrument.

pimax	no	integer	4500	[1<param<30000]

Upper PI channel limits of input images[2] in eV. Up to six different energy bands are allowed per instrument.

imagebuffersize	no	integer	640	100<param<10000

Parameter that controls memory requirements for mosaic images (cf. **emldetect** description).

withimagebuffersize	no	boolean	no	

Allow user-defined values of imagebuffersize.

obsmode	no	string	no	pointing—slew

Observation mode of input data: pointing (default) or slew. Parameter introduced with task version 4.24 (2015).

[1] File names must be ordered by instrument and energy band; maximum of six energy bands per instrument.
[2] One per energy band per instrument.
[3] One per instrument.

5 Errors

This section documents warnings and errors generated by this task (if any). Note that warnings and errors can also be generated in the SAS infrastructure libraries, in which case they would not be documented here. Refer to the index of all errors and warnings available in the HTML version of the SAS documentation.

BufferFull *(error)*
> 10000 detections: source buffer overflow

MissingParameter *(error)*
Missing input file name

FileMismatch *(error)*
Inconsistent number of input images

FileMismatch *(error)*
Inconsistent instruments or bands

FileMismatch *(error)*
Inconsistent coordinate systems

TooManyBands *(error)*
Images and / or `pimin`, `pimax` parameters for more than six energy bands provided.

WrongInst *(error)*
Only one instrument allowed for images in detector coordinates

WrongInst *(error)*
Unknown instrument

SizeMismatch *(error)*
Image has wrong size

WrongType *(error)*
Input image has wrong type

MissingAttribute *(warning)*
Keyword is missing in input file
corrective action: Keyword is not copied to output file

NoPSFcorrection *(warning)*
Input images are in detector coordinates.
corrective action: Counts and fluxes in the output source list will not be corrected for PSF losses.

ObsoleteParameter *(warning)*
Parameter withoffsets is obsolete and ignored.
corrective action:

6 Input Files

1. PPS product (from task **evselect**) : EPIC FITS images. One per instrument per energy band.
2. PPS product (from task **eexpmap**) : Exposure images. One per instrument.
3. PPS product (from task **emask**): Detection mask. One per instrument.
4. PPS product (from task **esplinemap**): Spline background maps One per instrument per energy band.

7 Output Files

1. PPS product (to be used by tasks **esplinemap** and **emldetect**) : EPIC **eboxdetect** source list

Table 2: Columns of the output source table

Column name	Unit	Description
BOX_ID_SRC		Source index
ID_INST		Instrument (1:EPN, 2: MOS1, 3: MOS2)
ID_BAND		Index of input images per instrument, =0 for summary row
SCTS	cts	Net source counts corrected for PSF losses
SCTS_ERR	cts	Statistical error of SCTS
BOX_CTS	cts	Raw counts in detection box without background subtraction or PSF corrections.
X_IMA	image pixel	X source position in image pixels
X_IMA_ERR	image pixel	Statistical error of X_IMA
Y_IMA	image pixel	Y source position in image pixels
Y_IMA_ERR	image pixel	Statistical error of Y_IMA
LIKE*		Detection likelihood
BG_MAP	cts/pixel	Background counts at source position
BG_RAW	cts/pixel	Mean counts/pixel in background area, without correction for scattered source counts. For map mode identical with BG_MAP.
EXP_MAP	seconds	Exposure at source position.
FLUX	erg/(cm^2 sec)	Flux calculated from RATE and parameter ecf
FLUX_ERR	erg/(cm^2 sec)	Statistical error of FLUX
RATE	cts/sec	Count rate (SCTS / EXP_MAP)
RATE_ERR	cts/sec	Error of RATE
RA	deg	Right ascension
DEC	deg	Declination
RADEC_ERR	deg	Statistical error of RA and DEC (cf. **emldetect** description)
LII	deg	Galactic longitude
BII	deg	Galactic latitude
HRn	$(1 \leq n \leq 3)$	Hardness ratios 1..3 as defined via parameter hrdef
HRn_ERR	$(1 \leq n \leq 3)$	Statistical error of the hardness ratio HRn
BOX_SIZE	pixel	Size of the detection box in image pixels
EEF		Enboxed energy fraction in the detection box. Calculated from the medium accuracy CAL PSF. Set to 1.0 for rebinned images.
DIST_NN	arcsec	Distance to nearest neighbour.

* Note that column LIKE was named SIGMA in **eboxdetect** versions up to 4.8. The content of LIKE is identical to that of SIGMA in these older versions.

8 Algorithm

```
subroutine eboxdetect

  Rebin images by linear factors of 1, 2, 4 ...
    (number of rebinning steps specified by parameter)

  Loop over rebinned images:
    Loop over image pixels:
```

```
      IF within area marked by detection mask
           (if mask flag is 'false': loop over all image pixels)
      THEN
        Loop over all instruments:
          Loop over all energy bands:

            IF local detect THEN
              Perform simple box search:
                1) Accumulate counts in 3 x 3 or 5 x 5 pixel source box.
                2) Accumulate counts in background area (16 background pixels
                   from 5 x 5 box or 24 background pixels from 7 x 7 box
                   centered on source box)
                3) Calculate statistical probability of excess in source box.
                   Correct for border effects if source/background pixels
                   are at image/detection mask border).
                4) Accumulate combined statistical probability of excesses
                   in all energy bands of all instruments.
            END IF

            IF map detect THEN
              Perform matched filter/box search:
                1) Accumulate PSF weighted counts (PSF centred on 5 x 5 box)
                   and/or non-weighted counts (to handle extended objects)
                   in 5 x 5 pixel box.
                2) Calculate statistical probability of excess with respect
                   to background map. Correct for border effects if
                   source/background pixels are at image/detection mask
                   border.
                3) Accumulate combined statistical probability of excesses
                   in all energy bands of all instruments.
            END IF

          END Loop
        END Loop

        IF combined statistical probability of excess exceeds threshold THEN
           Write source records to EBOXDETECT source table.
           Add fluxes and hardness ratios.
        END IF

      END IF
    END Loop
  END Loop

end subroutine  eboxdetect
```

9 Comments

10 Future developments

- add matched filter option

References