



catcorr

November 4, 2014

Abstract

Catcorr updates the positions of sources in an EPIC source list in an OBSMLI file. It derives these by cross-matching the positions with up to three reference catalogues to find optical or infra-red counterparts and finding the small frame shifts or rotations which optimise the match. These are then applied to all the detected positions. It is designed as a replacement for eposcorr and evalcorr.

1 Instruments/Modes

Instrument	Mode
EPIC	Imaging

2 Use

pipeline processing	yes
interactive analysis	yes

3 Description

The task `catcorr` replaces both `eposcorr` and `evalcorr`, but it requires another task `catprep` to be run first to extract from the three reference catalogues a small list of sources in the field of view of the current observation. This extracted list is a FITS file which can be made available as a data product to any user who might wish to run `catcorr` again with different parameters.

Since `catcorr` is used in the pipeline before any data quality assessment is performed, there is no information on whether the current fields are affected by problems such as very bright central sources, large extended sources, or various telescope artifacts. For this reason `catcorr` includes an initial rather basic attempt to detect an excess of spurious detections by finding whether any of them lie on what appear to be straight lines running through the centre of the field. Such lines are usually composed of mostly spurious detections, and so they are marked internally as of *bad* quality and not used in the cross-matching. Other detections ignored are those with `RADEC_ERR` exactly zero (a defect of source-finding) or those with `RADEC_ERR` over 3.0 arc-seconds, or with `EP_EXTENT` above 0.1 arc-seconds. Unless the number of *good* X-ray detections in the list exceeds the minimum specified by parameter `mingood` then the correction procedure is abandoned for this observation.



The matching procedure uses each of the reference catalogues separately: a statistic is computed which measures the overall goodness of fit between the set of XMM detections and the best matching reference source positions, allowing for the probability of positional coincidences. Only a proportion of the XMM sources will have counterparts and of course this proportion is unknown. Provided a reasonable number, perhaps five, appear to have counterparts, the following procedure appears to be reliable. Given the typical densities of X-ray and optical objects there may be a few false matches but not enough to affect the results significantly. The cross-matching is repeated while the position of the XMM frame is shifted by small increments in RA and DEC and rotated slightly about the centre of the field. A function minimisation routine due to Nelder and Mead (Computer Journal, January 1965, pp 308-313) is used to find the shift and rotation parameters which optimise the overall goodness of fit. This routine is similar in algorithm to routine AMOEBA which appears in *Numerical Recipes*, but the routine used here was originally from Rothamstead Research Establishment.

The reference catalogue which produces the highest maximum goodness of fit is adopted, provided it exceeds the minimum fit (parameter `minfit`) and the resulting absolute offset measured in arc-seconds does not exceed the parameter `maxoffset`. This positional shift and rotation are then applied to all detections in the OBSMLI file, and two new columns called RA_CORR and DEC_CORR are put in the FITS table (or updated if they already exist). The position shifts and standard errors in them, and the goodness of fit, are inserted in the file as FITS headers.

3.1 Statistical method

If r_{ij} is the great-circle distance between an X-ray detection (i) and an optical or infra-red source (j), and σ_{ij} is the associated error, then the probability of match P_{ij} is given as:

$$p_{ij} = \exp -0.5 \left(\frac{r_{ij}}{\sigma_{ij}} \right)^2 \quad (1)$$

while the probability of a coincidence with a optical/infra-red object within a radius r_{ij} when there are n_o optical/infra-red objects in the field of view of radius r_f is given by:

$$q_{ij} = n_o \left(\frac{r_{ij}}{r_f} \right)^2 \quad (2)$$

The net probability is the difference of these, subject to being above zero. Thus the overall goodness-of-fit is computed as:

$$L = \sum_{i=1}^{n_x} \sum_{j=1}^{n_o} \max(0.0, p_{ij} - q_{ij}) \quad (3)$$

Note that there are three components in σ_{ij} which are summed in quadrature:

- The estimated X-ray positional error in the OBSMLI file which is given in the column RADEC_ERR,
- The corresponding positional error in the reference catalogue,
- An additional systematic error determined to be 0.35 arc-seconds.

The procedure maximises L over all possible matching pairs as the frame shifts and rotations are applied.

The documentation for `eposcorr` sets out an analysis of the variance of a similar, but not identical, estimator arising from the positional errors in the input data, and suggests that the minimum acceptable value of L should be 5.1. I am not confident in the validity of this analysis, but extensive trials suggest that a useful positional correction can be obtained when L is about 5 or more.



3.2 Reference catalogues

- SDSS: the Sloan Digital Sky survey data release 8 - positions extracted from the photoPrimary table to a set of text files at Leicester using SQL scripts. This has the deepest sky coverage but only about one third of the sky is covered. When the XMM field is covered by SDSS, its goodness of fit value is almost always the largest.
- 2MASS: the Two Micron All Sky Survey, conducted by CalTech using telescopes in Arizona and Chile. The catalogue covers the entire sky, and contains about 600 million objects with a typical positional accuracy of 0.5 arc-seconds. The main value of 2MASS in this context is at very low galactic latitudes.
- USNO-B1: the US Naval Observatory catalog contains just over one billion objects from measurements of Schmidt plates taken at various observatories over the past 50 years. It covers the entire sky.

3.3 Temporary Astrometric Correction

The implementation of an astrometric correction in catcorr is a temporary measure designed to support the bulk reprocessing of data during 2012/3 to make the 3XMM catalogue. It is expected that in due course this correction will be applied via an update to the calibration system.

A recent analysis by Mike Watson and Andy Read (University of Leicester) has identified a small positional shift which is a function of off-axis angle. The effect is of the order of 0.65 arc-seconds at off-axis angles of 15 arc-minutes, near the edge of the field, in the sense that the real position is closer to the axis than that originally measured by elmdetect. If the task parameter astcorr is set to yes (its default value is no), this correction is applied to the RA, DEC values in the OBSMLI file. The empirical formula is:

$$yoff = c(0) + c(1) * r + c(2) * r**2 + c(3) * r**3$$

Where r is the off-axis angle from the pointing position in arc-minutes, and yoff is then the offset in arcseconds, which is subtracted from the original off-axis angle. The current coefficients of the cubic are:

$$real :: c(0:3) = (/+0.0158736, -0.036454, +0.00942358, -0.000277756 /)$$

4 Parameters

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

Parameter	Mand	Type	Default	Constraints
-----------	------	------	---------	-------------

srclistset	yes	dataset		
------------	-----	---------	--	--

File name for the FITS file of type OBSMLI containing a binary table of X-ray source positions and their errors. This must be an uncompressed FITS file which can be opened for read/write access.

catset	no	dataset	catextract.ds	
--------	----	---------	---------------	--

File name for the FITS binary table containing the external catalogue, which is used as the reference of



the source positions.

mingood	no	integer	10	
----------------	----	---------	----	--

Minimum number of good X-ray detections in the input dataset

minfit	no	real	5.0	
---------------	----	------	-----	--

Minimum overall goodness-of-fit for updating of RA_CORR and DEC_CORR.

maxoffset	no	real	10.0	
------------------	----	------	------	--

Maximum resulting offset in arc-seconds

syserr	no	real	0.1	
---------------	----	------	-----	--

Assumed systematic position error in arc-seconds

maxpofferr	no	real	0.75	
-------------------	----	------	------	--

Maximum position error offset for fit to be accepted in arc-seconds

astcorr	no	bool	no	
----------------	----	------	----	--

Set yes to perform astrometric correction dependent on off-axis angle

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.

alreadyCorrected (*warning*)

The program runs anyway, this is mainly for information

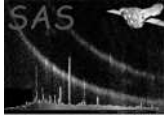
corrective action: The OBSMLI file has already been processed by catcorr or eposcorr

fitsioError (*error*)

The message indicates the location of the error in the software, but it is likely to be a defect in the FITS structure which is the cause.

ibinTooLarge (*error*)

In computing a histogram of position angles a bin is out of range: should not happen please report to the helpdesk

**ibinUnexpectedlyLarge** (*warning*)

This is a result of rounding errors and does not matter

corrective action: In computing a histogram of position angles a bin is out of range

invalidDateObs (*warning*)

This is used only to correct proper motions in the USNO-B1 catalog, and will not have a large effect

corrective action: The DATE-OBS value cannot be decoded

missingExtract (*error*)

Cannot open the catalogue extracts file produced by catprep

missingColumn (*error*)

FITS file does not have a column of required information

missingKeyword (*error*)

FITS table does not have a required header keyword

missingOBSMLI (*error*)

The OBSMLI file specified cannot be opened

obsmliNotTable (*error*)

In this FITS file HDU2 is not binary table as expected.

tooManyPairs (*warning*)

A fit will be produced based on the maximum number of pairs that can be stored

corrective action: There are too many pairs of detections for cross-matching

6 Input Files

1. An OBSMLI file: a PPS product containing a list of X-ray detections with right ascension (RA), declination (DEC). and position error (RADEC_ERR) of each.
2. Catalogue extract: a FITS file generated by `catprep` containing extracts for the current field of view from the reference catalogues, respectively SDSS, 2MASS, and USNO. Each extract is in a separate HDU. If the SDSS survey did not cover the relevant part of the sky the HDU still exists but will have zero rows. The other two catalogues have all-sky coverage so this should not happen.

It is possible for any other catalogue extract to be provided instead of that generated by `catprep`. The format is designed to be compatible with that produced by `eposcorr`. If a user-generated file is provided instead, this needs to contain at least one HDU which is a binary table; it may contain up to three tables in total if information from more than one catalogue is to be provided. Each binary table must contain at least these three columns:

- CAT_RA: the Right Ascension of the catalogue object, degrees.
- CAT_DEC: the Declination of the catalogue object degrees.
- CAT_RADEC_ERR: the 1-sigma error radius of the position, degrees.

The table must also define the EXTNAME keyword (the value is normally the name of the catalogue, but its value is not checked).



7 Output Files

There is no separate output file, the output is inserted in the OBSMLI file which is the principal input. Two columns RA_CORR and DEC_CORR are created (or updated if already present) with corrected celestial positions (degrees) of the X-ray source; if the attempt to find frame shifts is unsuccessful then these columns will be identical to RA and DEC respectively.

A new column, SYSERRCC, is inserted in the table. This is an estimate of the systematic error (in arc-seconds) in celestial positions that arises from the uncertainty in the position and rotation angle of the field of view. This is estimated by `catcorr` from the errors of the fit to the reference sources. Because a field rotation is fitted as well as shifts along the RA and declination axes, the uncertainty is slightly higher for detections further from the centre of the field, so that SYSERRCC is now a column and not just a fixed value for the field as it was in earlier versions of `catcorr`. If `catcorr` is unable to get a satisfactory fit to reference sources then SYSERRCC is set to 1.5 arc-seconds for all detections in the field, this being the value determined empirically which seems to be the best fit to the uncorrected positions.

The following keywords will be inserted in the FITS table (or updated if already present):

- RAOFFSET, DEOFFSET the frame shifts applied (arc-seconds).
- RAOFFERR, DEOFFERR the estimated standard errors in the above (arc-seconds).
- ROT_CORR the frame rotation angle (degrees).
- ROT_ERR the estimated standard error on the above (degrees).
- LIK_HOOD the goodness of fit parameter.
- NMATCHES the number of X-ray/reference catalogue matches used in the fit.
- POSCOROK a logical value set T if the fitting was successful, F otherwise. If the fit is not successful the shift and rotation parameters will be zero.
- ASTCORR is a logical value set F normally but set to T if the `astcorr` parameter has been set to implement the temporary astrometric corrections.

8 Comments

The `catcorr` task is believed to be a modest improvement on `eposcorr` in the following respects:

- `catcorr` uses three reference catalogues, instead of just USNO. The 2MASS infra-red positions are mostly useful at very low galactic latitudes. In parts of the sky covered by the Sloan Digital Sky Survey (SDSS) it has much superior sensitivity and hence a larger number of X-ray counterparts.
- `catcorr` uses an iterative function optimisation algorithm to find the best fit, whereas `eposcorr` evaluated the fit on a relatively coarse grid of positions shifts and rotations, which resulted in some quantisation of the resulting values.
- `catcorr` attempts to detect and ignore certain image artefacts and so is able to find a reasonable cross-match in some cases when `eposcorr` could not.



- `catcorr` uses the positional error estimates provided in the reference catalogues which `eposcorr` did not (because it relied upon `WCStools` for access to catalogue data). In the case of USNO it also uses proper motion values which can be significant for nearby stars, especially when the epoch of the reference catalogue differs significantly from the date of the XMM observation.

The result is that whereas `eposcorr` was able to find improved positions in about 60% of fields, the success rate of `catcorr` is close to 85%.

References