XMM emldetect sensitivity maps for ARCHES

Christian Motch

Observatoire astronomique de Strasbourg

Building sensitivity maps for the 4XMM

Motivation:

- ESA requests ARCHES SEDs for all sources as part of the 4XMM delivery.
- ARCHES SEDs are assembled using the multi-catalogue statistical cross-match tool developped by F.-X Pineau at CDS.

Improvements wrt ARCHES 2015:

- Build SEDs for as many as possible 4XMM sources **=> extend the 'clean' catalogue**
- Improve statistical priors **=> group sources by ranges of count rates** (not by observations of similar exposure time)

=> Computing source surface density in a range of flux requires "**sensitivity" maps.**

Word of caution

- We do not compute actual upper limits, just maps of the pixels for which the count rate yields a DET_ML equal to the XMM catalogue threshold
- Sensitivity (or detection ?) maps applicable to XMM catalogues should take into account the characteristics of the 'Cash statistics' used in the emldetect algorithm with:

  EP_8_DET_ML > 6 (ML limit for inclusion in XMM catalogues)

emldetect (in a nut shell)

Assuming Poissonian statistics a detector with N pixels, with an expected number of events e_i in each pixel and an observed number of events n_i, the probability (likelihood) to exactly have the given realisation is:

$$Pr(n_i, e_i) = \prod_{i=1}^{N} \frac{e_i^{n_i} e^{-e_i}}{n_i!}$$

Cash statistics is derived from a likelihood- ratio test comparing the goodness of fit of two models:

- background only (value given by external task)
- source + background (maximized wrt to source parameters)

$$LR = \frac{\prod_{i=1}^{N} \frac{b_i^{n_i} e^{-b_i}}{n_i!}}{\max(/s)\ \left(\prod_{i=1}^{N} \frac{(s_i+b_i)^{n_i} e^{-(s_i+b_i)_i}}{n_i!}\right)}$$

Wilks's theorem:

-2 ln(*LR*) is distributed as a χ^2 with dof = number of fitted parameters

=> Provides an convenient way to compute the probability that the source is real.

emldetect (in a nut shell)

In practice: $$C = -2\sum_{i=1}^{N}(n_i \ln e_i - e_i)$$

$$\Delta C = C(b) - C(s+b) = -2\ln LR$$ (for one band and one camera)

$$\Delta C_T = \sum_{k=band,camera} (\Delta C_k)$$ (Summing over all bands and cameras)

χ^2_n cumulative distribution function

$$DET\_ML = -\ln(1 - \mathrm{CDF}(\Delta C; \nu))$$

Number of fitted parameters: position (x,y) + countrate in each band.

Modelling emldetect by a simple "Poissonian" statistics

Let's assume that e^{-DET_ML} is <u>identical</u> to the probability of finding the same excess of counts from a Poissonian background fluctuation measured in a small sky area

(Algorithm used by e.g. Carrera et al. 2007, A&A, 469, 27 and Mateos et al. 2008, 492, 51)

Here applied to the broad band 8 (so as to relate it to the selection threshold used in the XMM catalogues)

Modelling emldetect by a simple "Poissonian" statistics

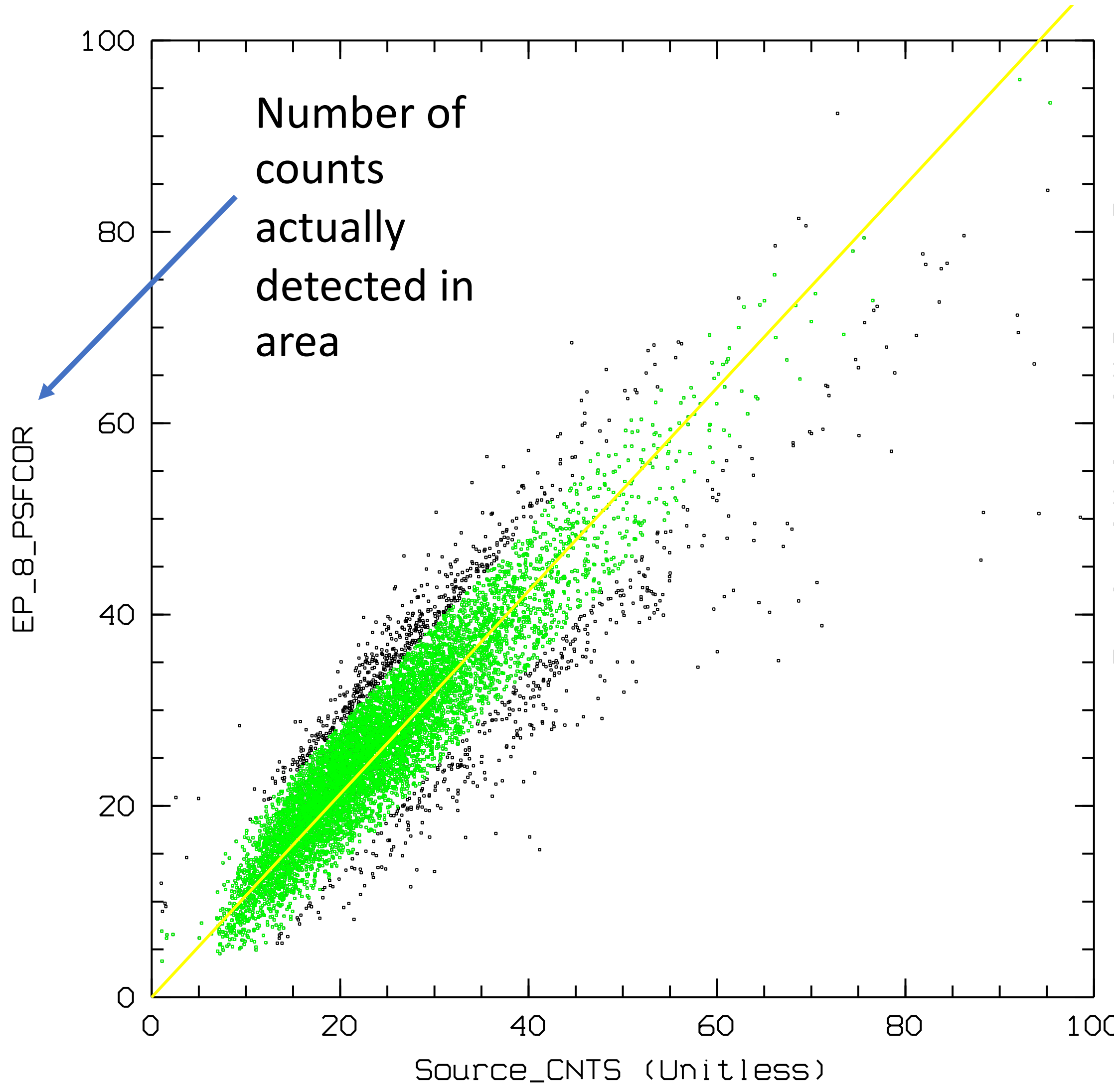
We fit the number of counts for which the Poissonian cumulative distribution function is equal to $e^{(-DET\_ML)}$ to the actual number of counts detected by XMM (EP_8_PSF_COR = psf * EP8_TRUCNTS)

Scan various sky area considered and fraction of psf covered in area.

Force relation to go through 0,0 and have a slope 1

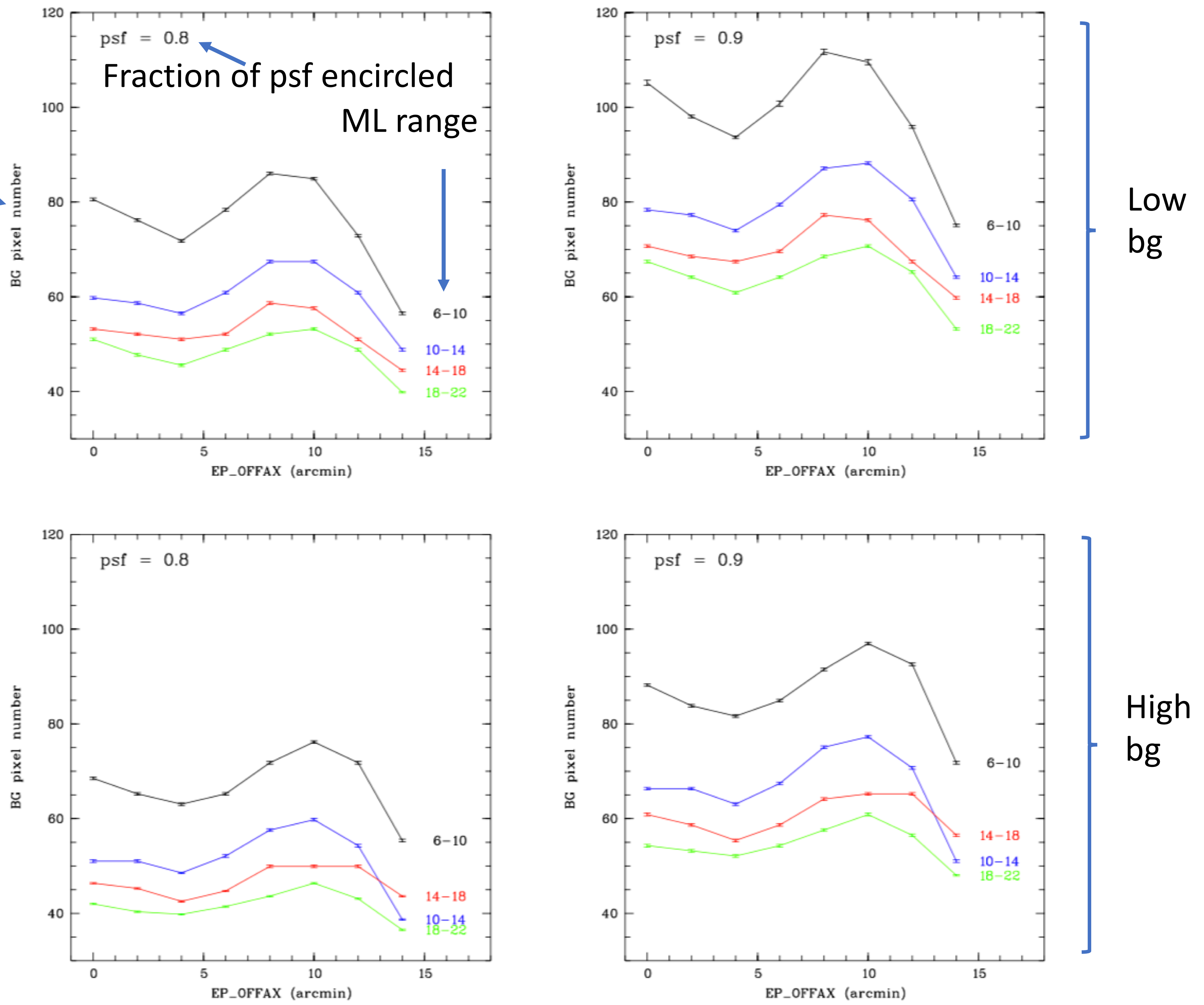
Use only 'good' sources (low background, clean, off_axis < 14', etc)

Here : EP_8_DET_ML < 15



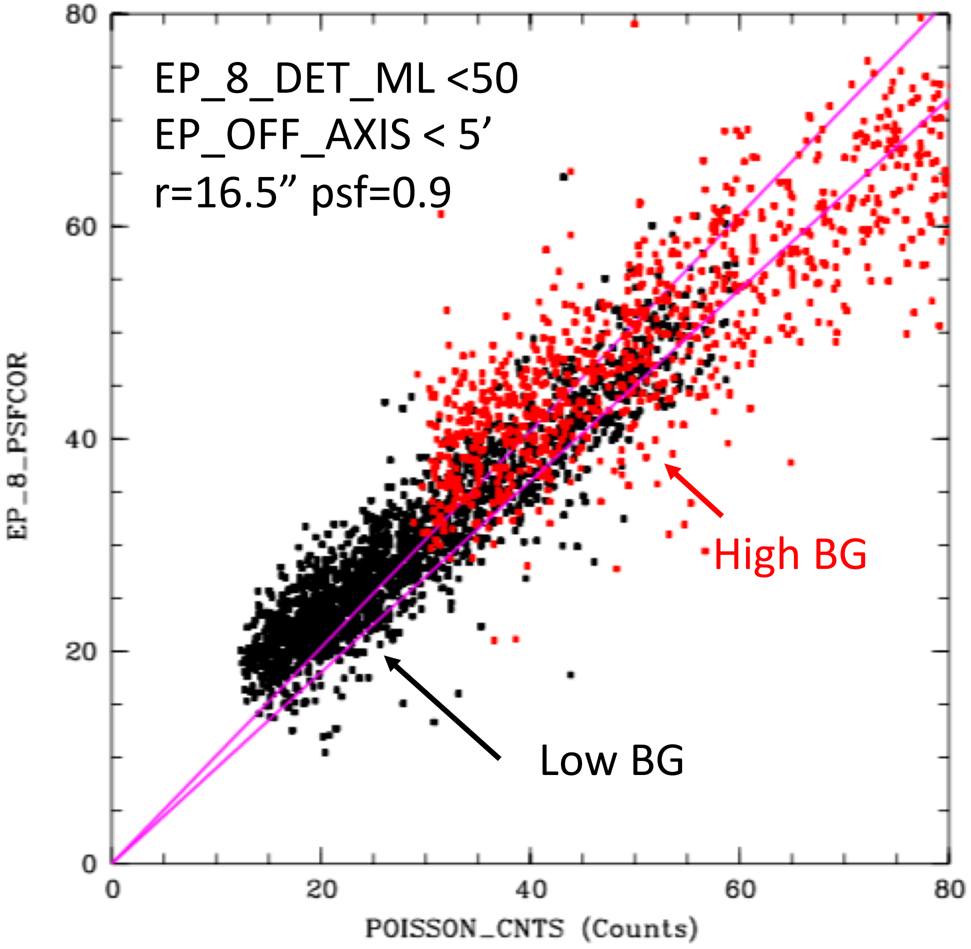
(derived from Poissonian statistics)

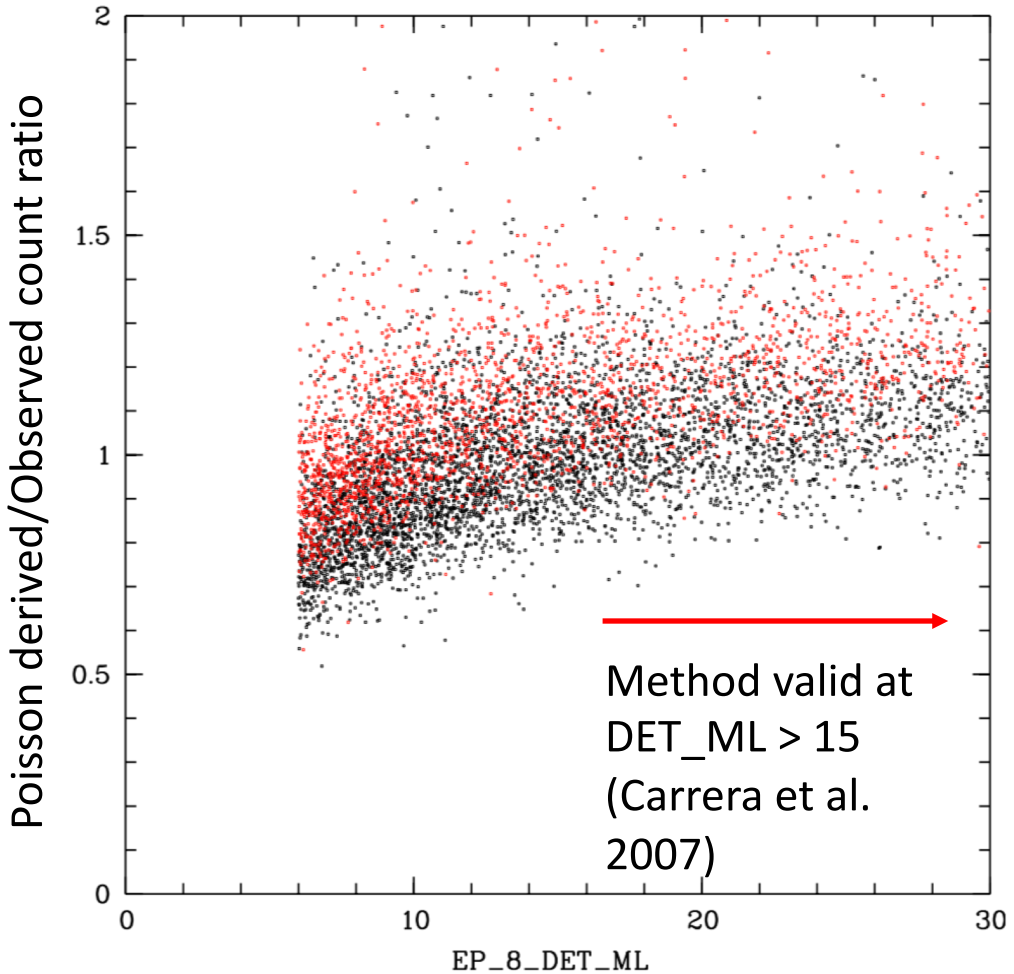
Sky area for which
Detected cnts = cnts(DET_ML)
assuming simple Poissonian
statistics



Problems:

- Fitting process does not yield a unique area/psf fraction solution
- Best fit area/psf values depend on background, offaxis angle and ML





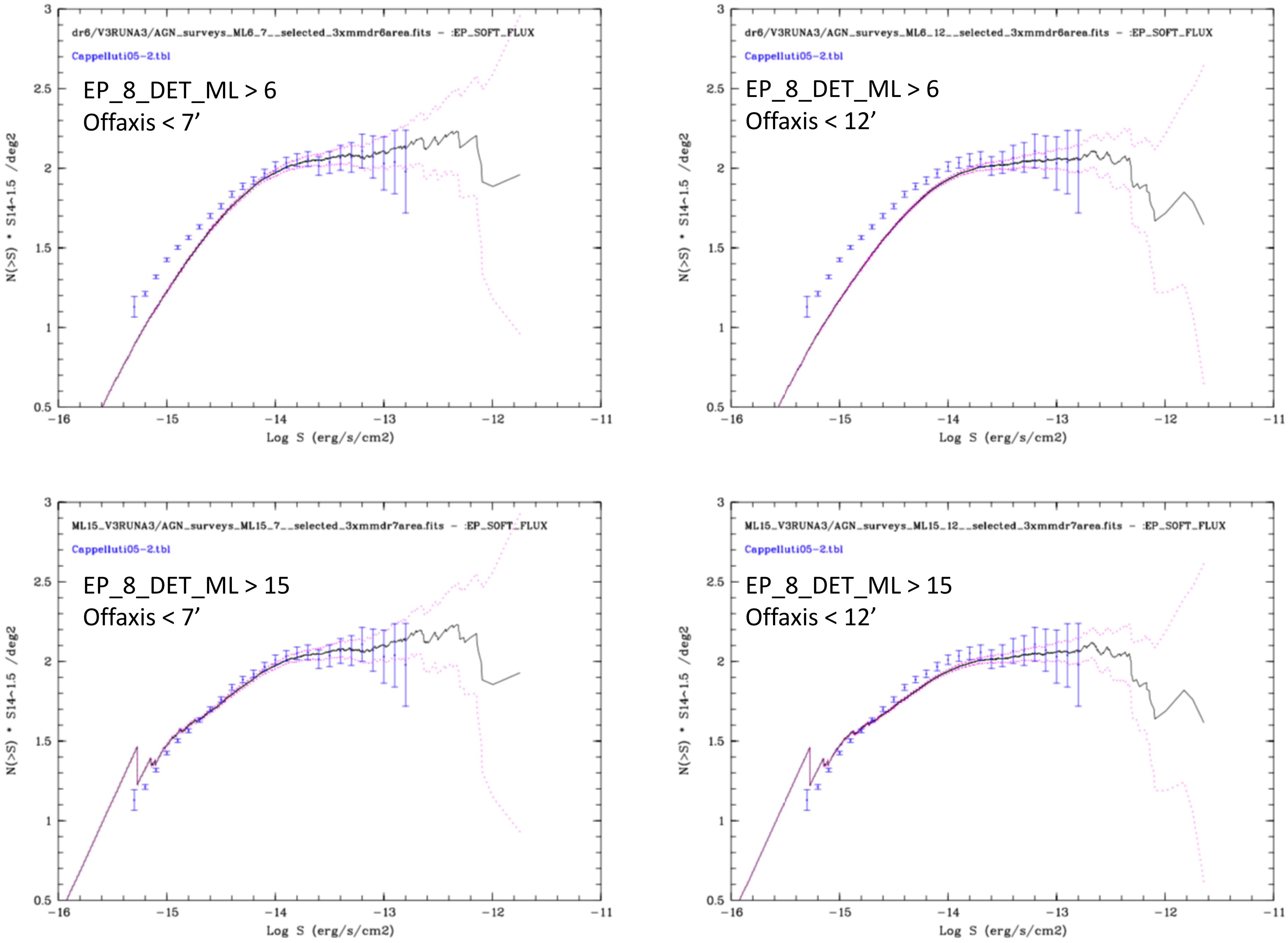
Reconstructing EG logN-LogS in the soft band.

(No correction for Eddington bias)

Psf = 0.9, r=19.7”

Source sample = all large EG surveys:
XXL, XMM LSS, XMM-LSS, XMM-HATLAS, Stripe 82, ”Deep Field”, DEEP, ”Cosmos field”, ”COSMOS FIELD”, CFHTLS or CDFS) with OBS CLASS = 1.

EG logN-logS from Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A , 497, 635



Conclusions

- The area within which “simple Poissonian” statistics properly matches the emldetect statistics varies with DET_ML, offaxis angle and background.
- However, parameter changes are small at DET_ML > 15, thus validating the method used by e.g. Carrera et al. 2007.
- This approach cannot be used to compute sensitivity maps down to the limit of EP_8_DET_ML = 6.

Using the Cash statistics

The basic and idea proposed here is to compute the expected value of EP_8_DET_ML, by rerunning a simplified *emldetect* process centred at each point of the detector.

Recipe:

- Assume a spectral shape as a count rate distribution between different bands and camera.
- For each band and camera, compute expectation of $\Delta C = C(b) - C(s + b)$, sum them up and using the Wilks theorem, derive the corresponding EP_8_DET ML.
- With a zero finding routine, determine the count rate yielding the target EP8_DET_ML

Using the Cash statistics

In each band / camera:

Number of pixels

Bg map

$$E\left(\frac{C(b)}{2}\right) = Nbg - \ln(bg)(S + Nbg) \qquad S = \sum_{i=1}^{N} s_i$$

*Count rate * exposure map * psf distribution*

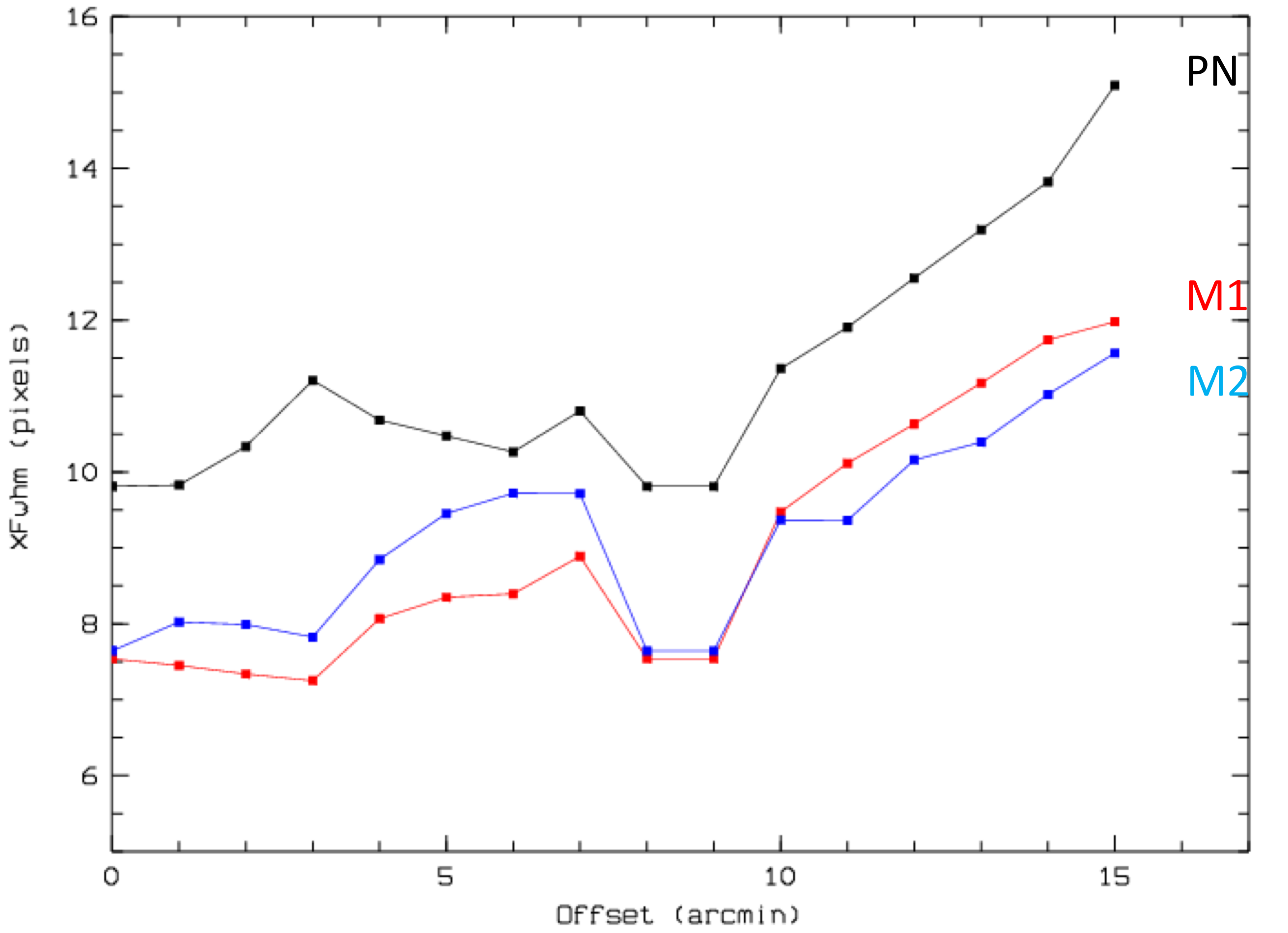
$$E\left(\frac{C(s+b)}{2}\right) = Nbg + S - \sum_{i=1}^{N} \left[(s_i + bg)\ln(bg + s_i)\right]$$

PSFs

Several model tested. Good results can already be obtained using a simple MEDIUM model in *psfgen*

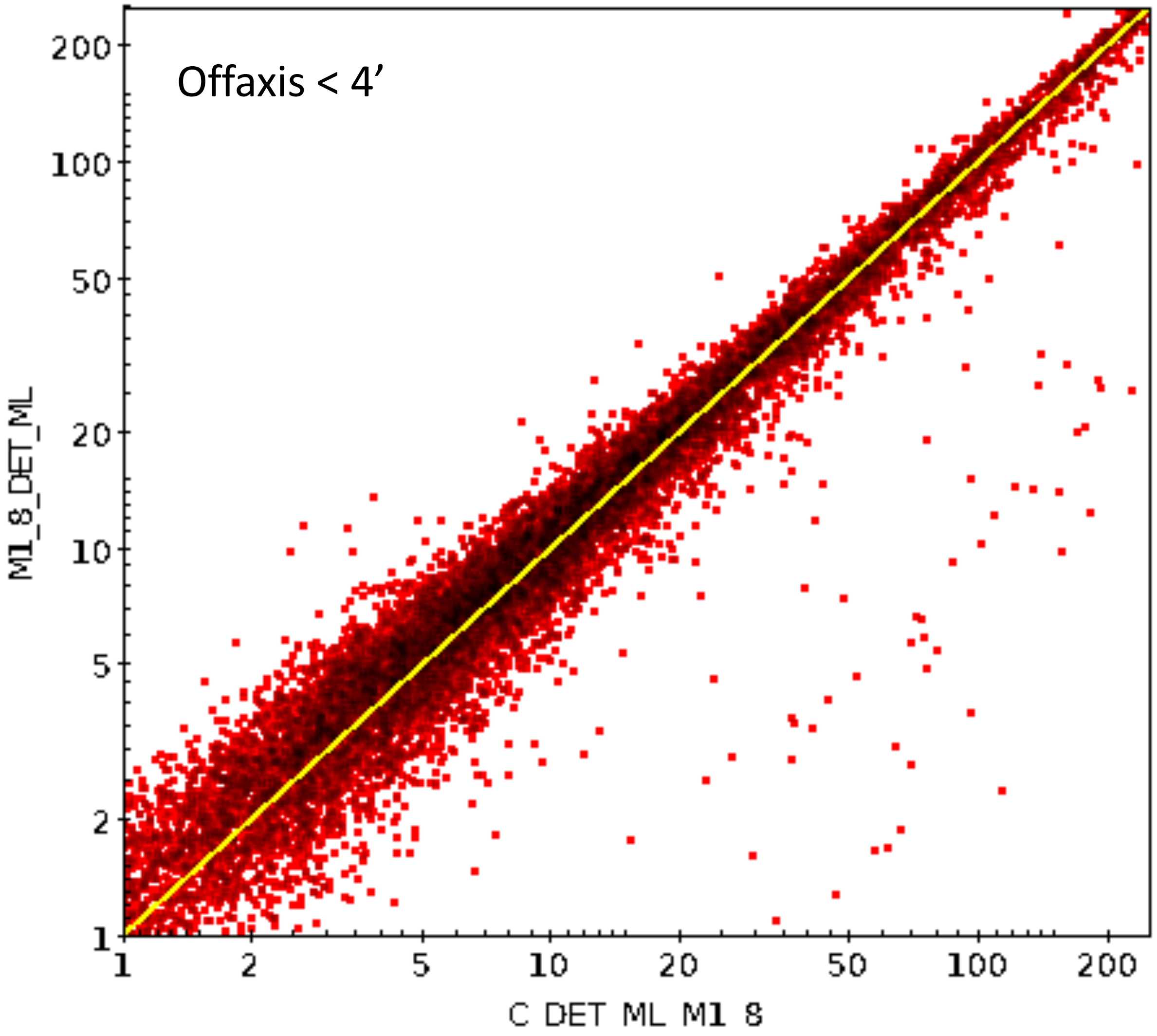
Final code implementation uses the ELLBETA model with specific telescope and band PSFs for offaxis angles 0-14’ with step 1’.

Caveat: azimuthal angle remains constant

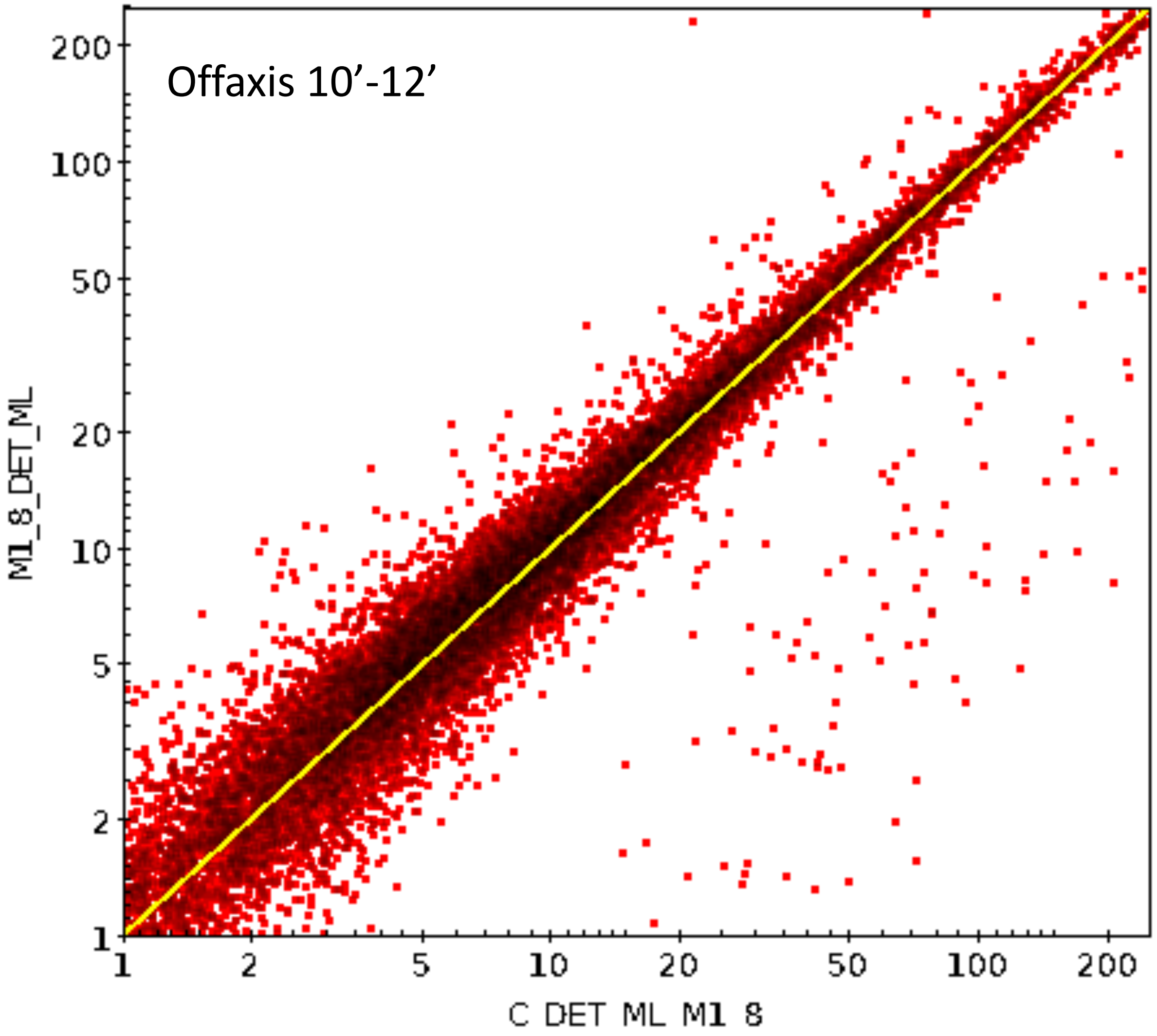


FWHM in the x-direction resulting from 2-D Gaussian fits to the one arcsec pixel size ELLBETA PSFs. Energy band 3 ($< E > = 1500$eV), azimuthal angle = 0 deg

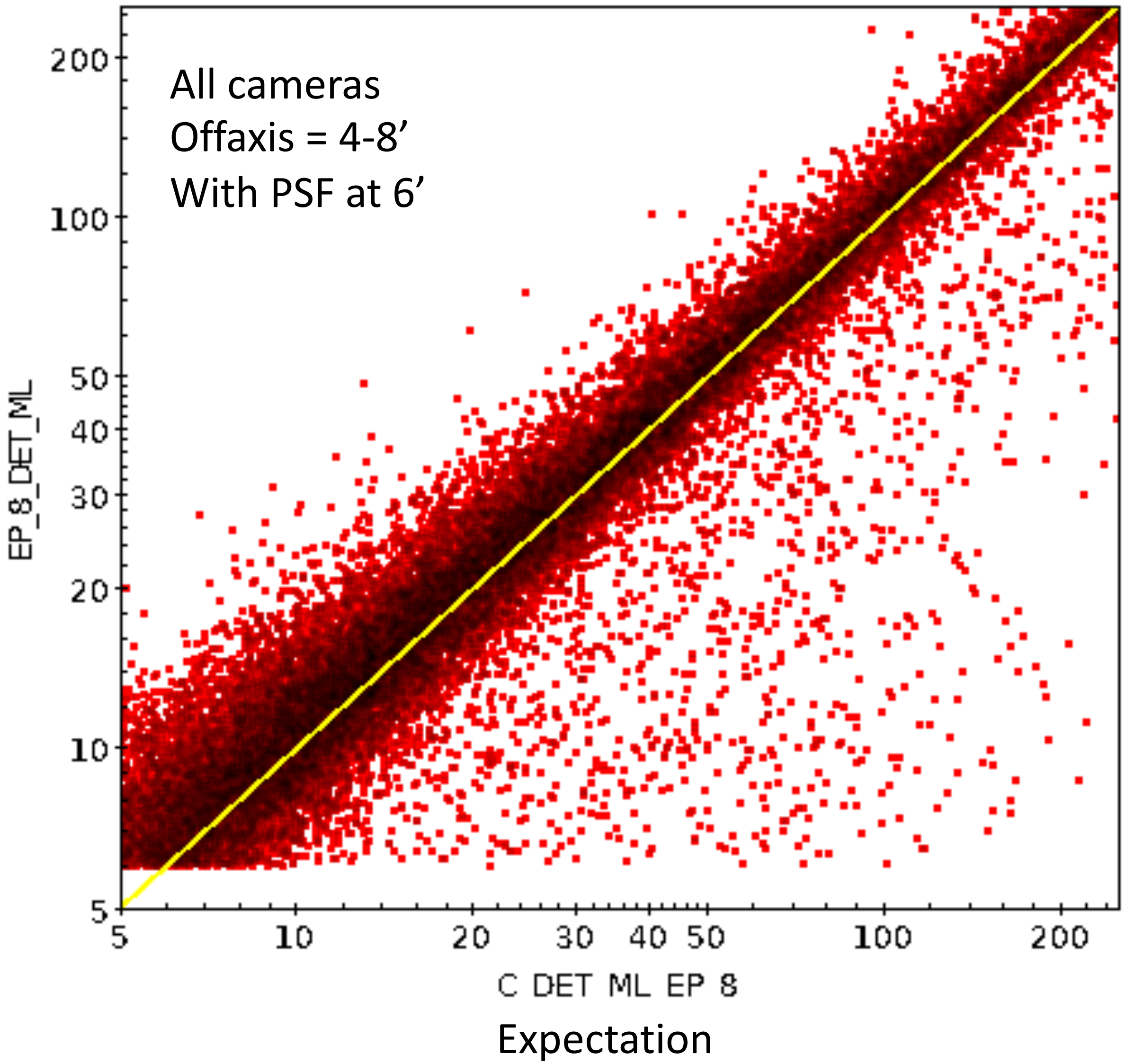
Expected value ok down to very low ML threshold



Expectation



Expectation


All cameras
Offaxis = 4-8’
With PSF at 6’
EP_8_DET_ML
C_DET_ML_EP_8
Expectation

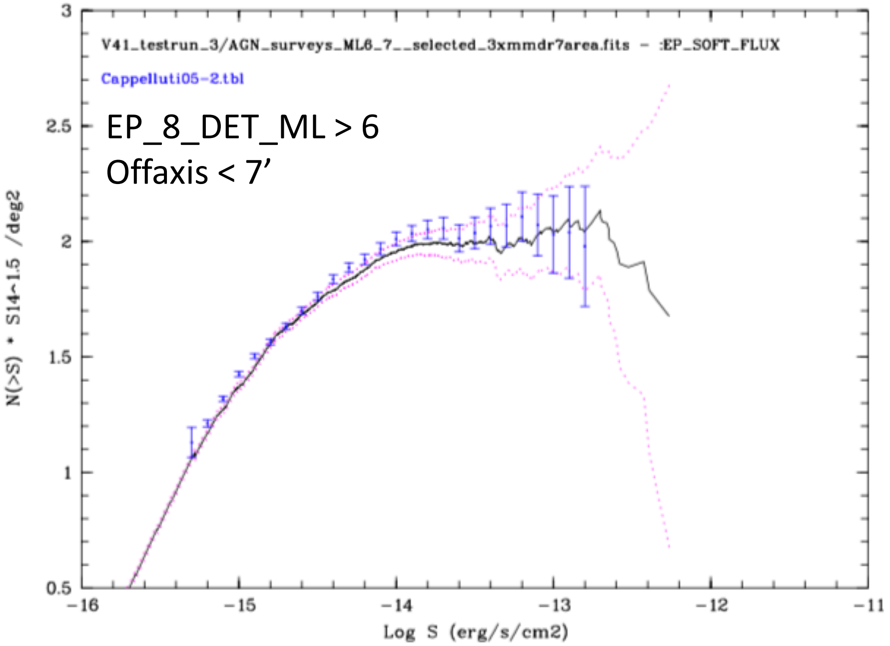
Reconstructing EG logN-LogS in the soft band.

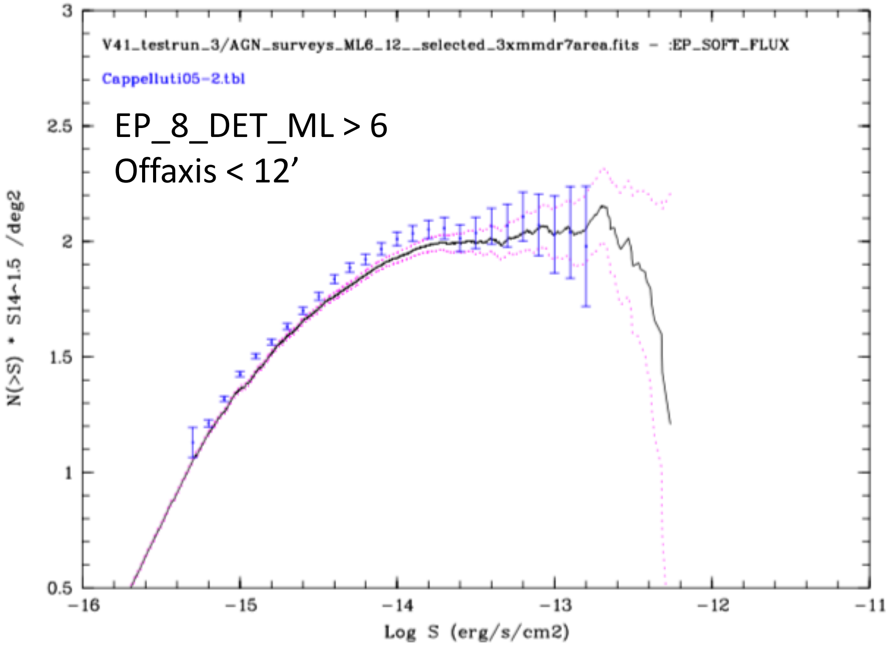
(No Eddington bias)

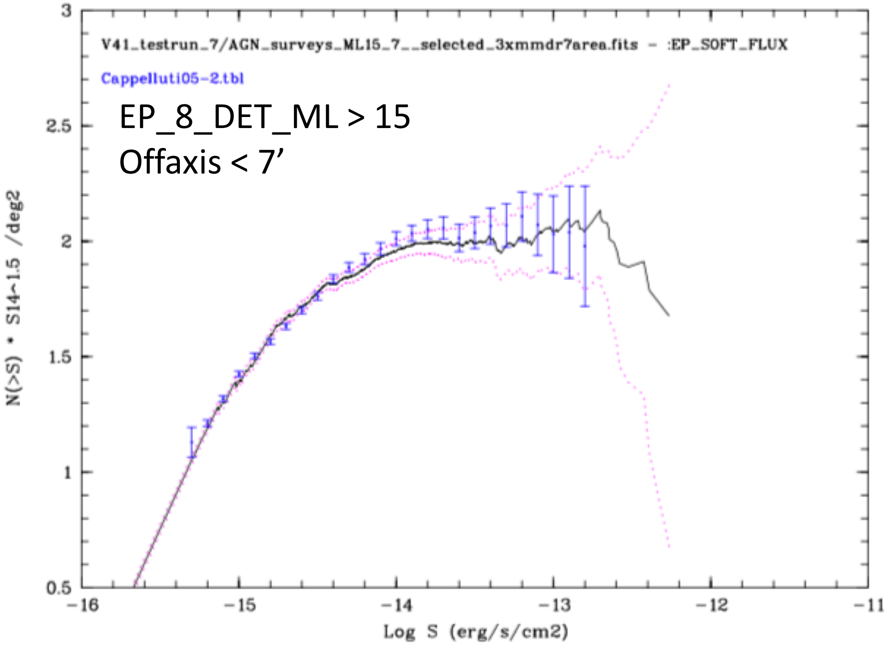
Mean survey count rate spectrum

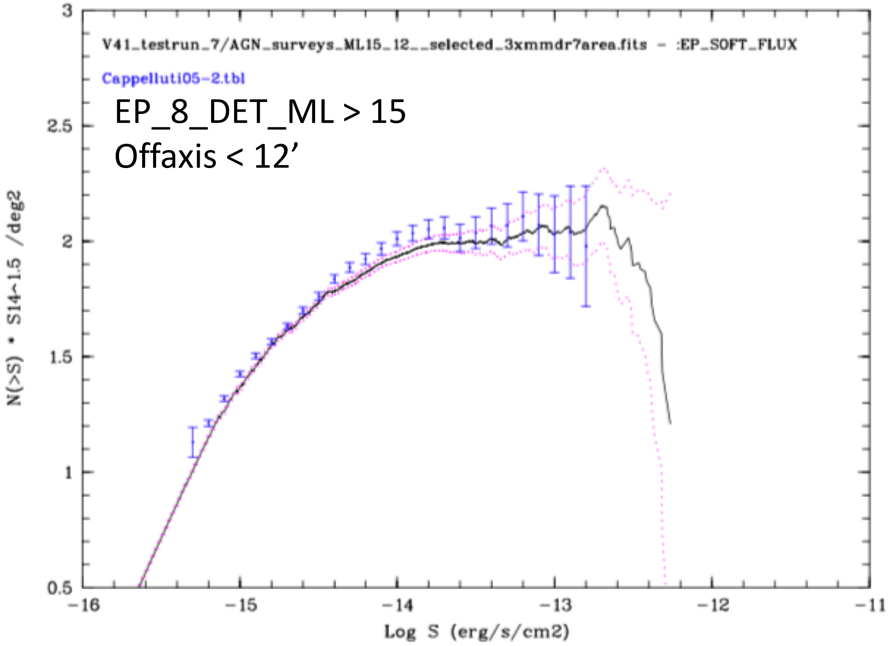
Source sample = all large EG surveys:
XXL, XMM LSS, XMM-LSS, XMM-HATLAS, Stripe 82, ”Deep Field”, DEEP, ”Cosmos field”, ”COSMOS FIELD”, CFHTLS or CDFS) with OBS CLASS = 1.

EG logN-logS from Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A , 497, 635









Conclusions (1)

- There is no “simple Poissonian” model matching the behaviour of the detection maximum lilkelihood provided by emldetect. Areas and psf fractions depend on background, PSF size (off axis angle) and ML range.

  (Relatively good approximations still exist at DET_ML > 15)

- Using the Cash statistics provides much more reliable and stable estimates of the sensitivity (or detection maps ?) relevant to XMM catalogues.
- Cash sensitivity maps are now the baseline for computing source densities to be fed in the ARCHES crossmatch tool.

Conclusions (2)

- Better results by the Cash method are obviously not surprising. Emldetect includes the resemblance of the photon distribution to the PSF in the statistics in contrast to the “simple Poissonian” method. Emldetect also approximates detection probabilities using the Wilks theorem *
- Code needs about 15 min per observation. DR7 processing lasted 2 days on our 96 core server.

* (to be taken with a grain of salt) : one possible worry is the validity of the Wilks theorem. *Preliminary simulations suggest that for sources detected with about 60-80 total counts in all 3 cameras (17 dof), Wilks theorem may overestimate the detection likelihood.*