eROSITA data analysis of clusters of galaxies

Jeremy Sanders
MPE

Cluster properties

- To study cluster physics and cosmology we are interested in a number of cluster properties (luminosity, mass, …)
- However, clusters are not point sources, which makes their analysis challenging
- Often these are integrated quantities (i.e. within some radius) or we may want to study the radial profiles of quantities
- For individual systems we may want to look at the image or map properties

Abell 3266 Sanders et al. 2022

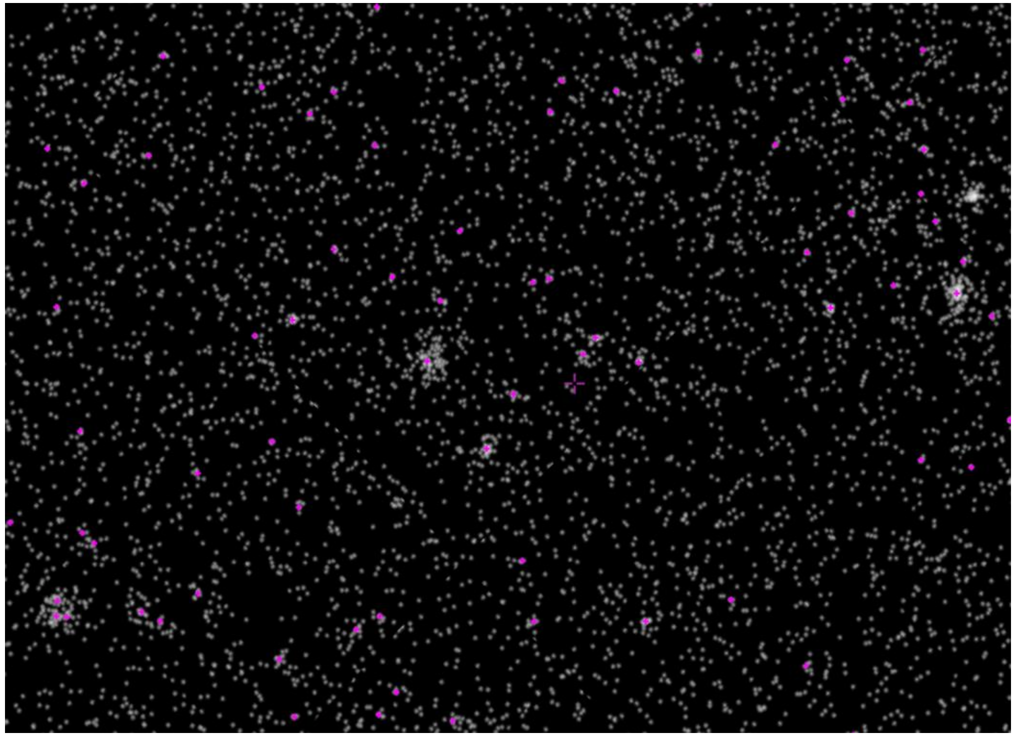
Where to start?

- We might want to look at a set of objects from an existing catalogue
- However, if we're interested in eROSITA selected objects, then we should look at the catalogues produced in the first data release:
  - https://erosita.mpe.mpg.de/dr1/AllSkySurveyData_dr1/Catalogues_dr1/

Main source catalogue (Merloni et al. 2024)

- All detected X-ray sources, including point sources
- Sources detected via box detection (**erbox**), followed by a maximum likelihood fitting procedure (**ermldet**)
- **ermldet** fits objects with point and extended source model
- Extended model is beta model, where surface brightness $S(r) = S_0\,(1 + r^2/r_c^{\,2})^{-1.5}$
- Models convolved with PSF and background added
- Computes detection likelihood of source
- Computes likelihood that the source is extended
- If the source has too low detection likelihood, it is not in the output catalogue
- If the extent likelihood is below threshold, only the point source model results are retained

Example detected sources in DR1



Catalogue contents

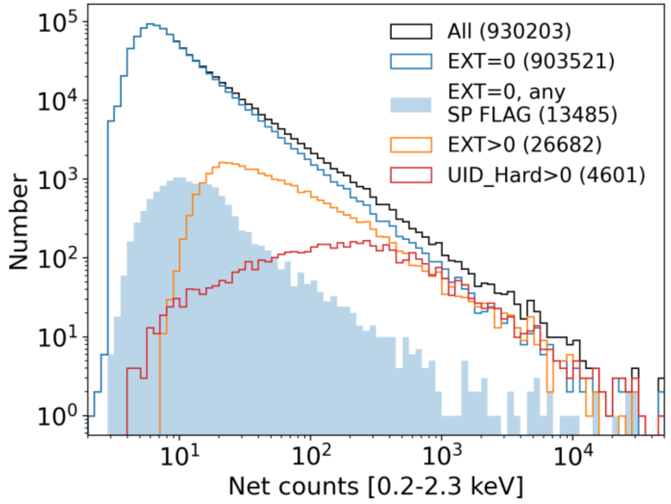
Data Model: eRASS1_Main

https://erosita.mpe.mpg.de/dr1/AllSkySurveyData_dr1/Catalogues_dr1/MerloniA_DR1/eR

Binary Table Caption for HDU1

Name	Type	Unit	Description
IAUNAME	char[23]	None	String containing the official IAU name of the source
DETUID	char[32]	None	String unique detection ID
SKYTILE	int32	None	Sky tile ID
ID_SRC	int32	None	Source ID in each sky tile. Use SKYTILE+ID_SRC to identify the corresponding source products
UID	int64	None	Integer unique detection ID. It equals CatID*10^11+SKYTILE*10^5+ID_SRC, where catID is 1 for the 1B detected Main and Supp catalogs and 2 for the 3B detected Hard catalog
UID_Hard	int64	None	Hard catalog UID of the source with a strong association, or -UID if the association is weak
ID_CLUSTER	int32	None	Group ID of simultaneously fitted sources
RA	float64	deg	Right ascension (ICRS), corrected
DEC	float64	deg	Declination (ICRS), corrected
RA_RAW	float64	deg	Right ascension (ICRS), uncorrected
DEC_RAW	float64	deg	Declination (ICRS), uncorrected
RA_LOWERR	float32	arcsec	1-sigma lower error on RA
RA_UPERR	float32	arcsec	1-sigma upper error on RA
DEC_LOWERR	float32	arcsec	1-sigma lower error on DEC
DEC_UPERR	float32	arcsec	1-sigma upper error on DEC
POS_ERR	float32	arcsec	1-sigma positional uncertainty
RADEC_ERR	float32	arcsec	Combined positional error, raw output from PSF fitting
LII[1]	float64	deg	Galactic longitude

- Important properties include
  - RA, DEC
  - DET_LIKE_0
  - EXT_LIKE
  - EXT (here r_c)
  - ML_CTS_1
  - ML_FLUX_1
  - ML_RATE_1
  - ML_BKG_1



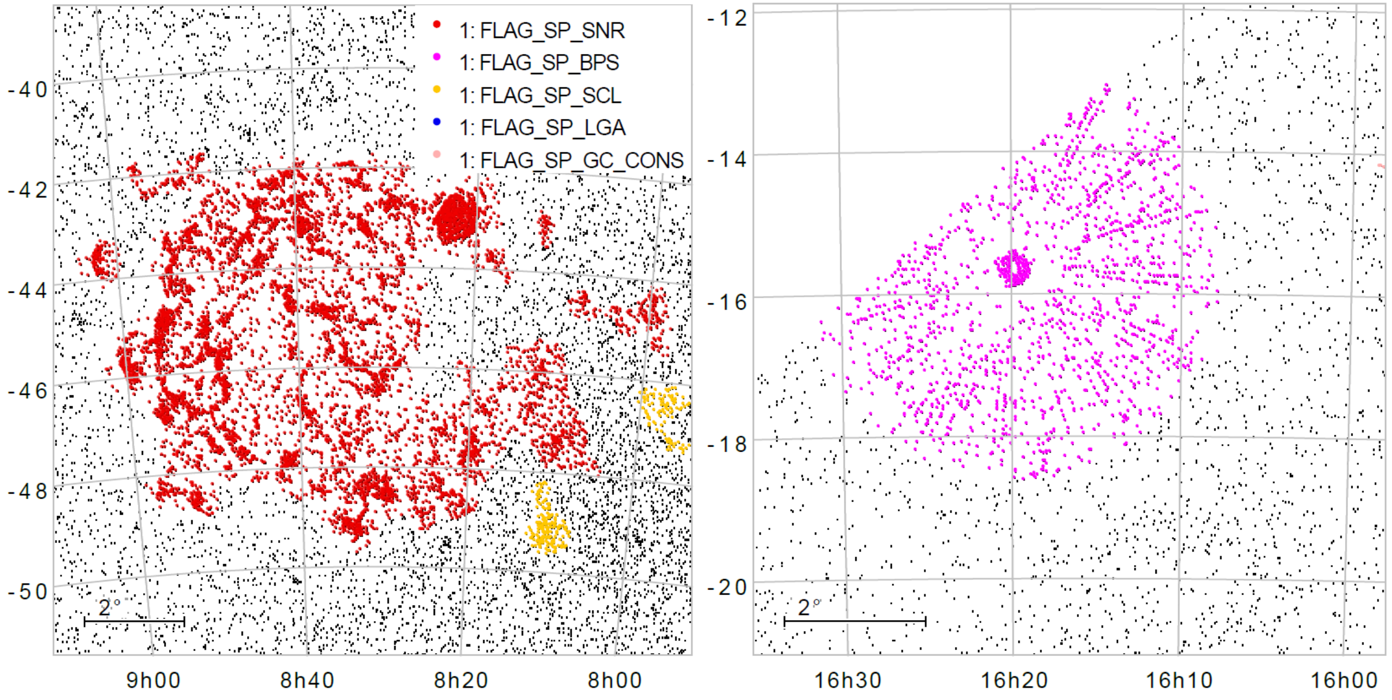
- Be careful using rates, etc, for extended sources as these tend to be unreliable – you should do proper source fitting

https://erosita.mpe.mpg.de/dr1/AllSkySurveyData_dr1/Catalogues_dr1/

Merloni et al. 2024

Catalogue issues

- Bright nearby sources (e.g. nearby clusters) can be split into multiple objects
- Source detection can be affected by bright nearby sources
- Low surface brightness objects can be lost due to background fitting and maximum extent of source fitting



Merloni et al. 2024

Cluster catalogue (Bulbul et al. 2024)

- Merge split nearby sources
- Apply sky mask (e.g. exclude Galactic plane)
- Clean sample using optical confirmation (Kluge et al. 2024)
- Two samples
  - Main sample with extent likelihood > 3
    - 12,247 optically confirmed clusters
  - Purer cosmology sample with extent likelihood > 6
    - 5,259 optically confirmed clusters
- Photometric and literature redshifts in catalogue
- X-ray properties obtained using MBProj2D

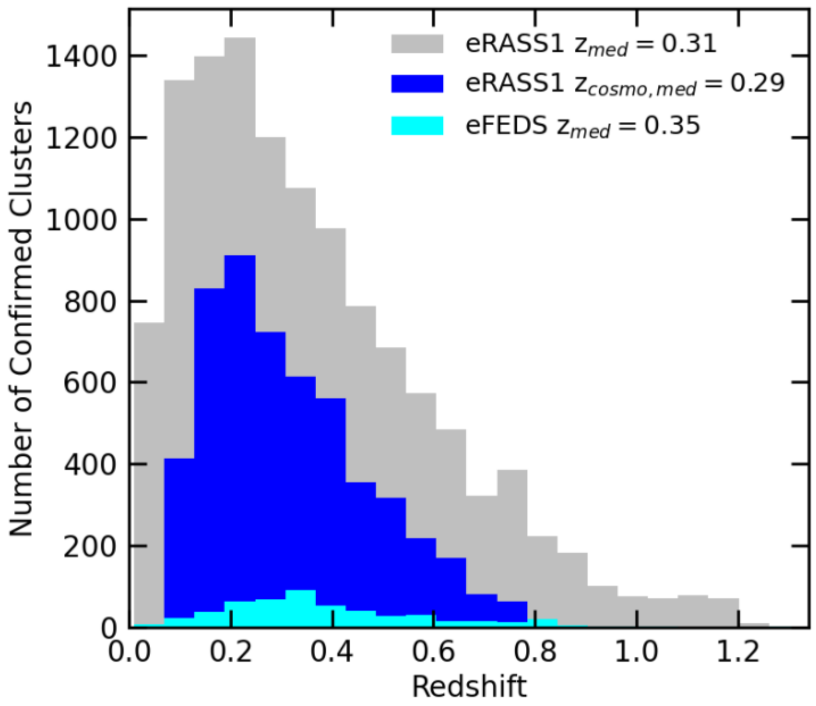


Fig. 6. Redshift distribution of the 12 247 confirmed eRASS1 clusters and groups. Shown in gray is the eRASS1 cluster sample with $\mathcal{L}_{ext} > 3$, compared to those of the cosmology sample in blue, and the redshift distribution of the 477 clusters confirmed in the eFEDS field in cyan (Liu et al. 2022). The median redshift of the eRASS1 cluster catalogs is slightly lower than that of the eFEDS clusters ($z_{med} = 0.35$).

Cluster catalogue on eROSITA website

Data Model: erass1cl_main_v3.2

https://erosita.mpe.mpg.de/dr1/AllSkySurveyData_dr1/Catalogues_dr1/BulbulE_DR1

Binary Table Caption for HDU1

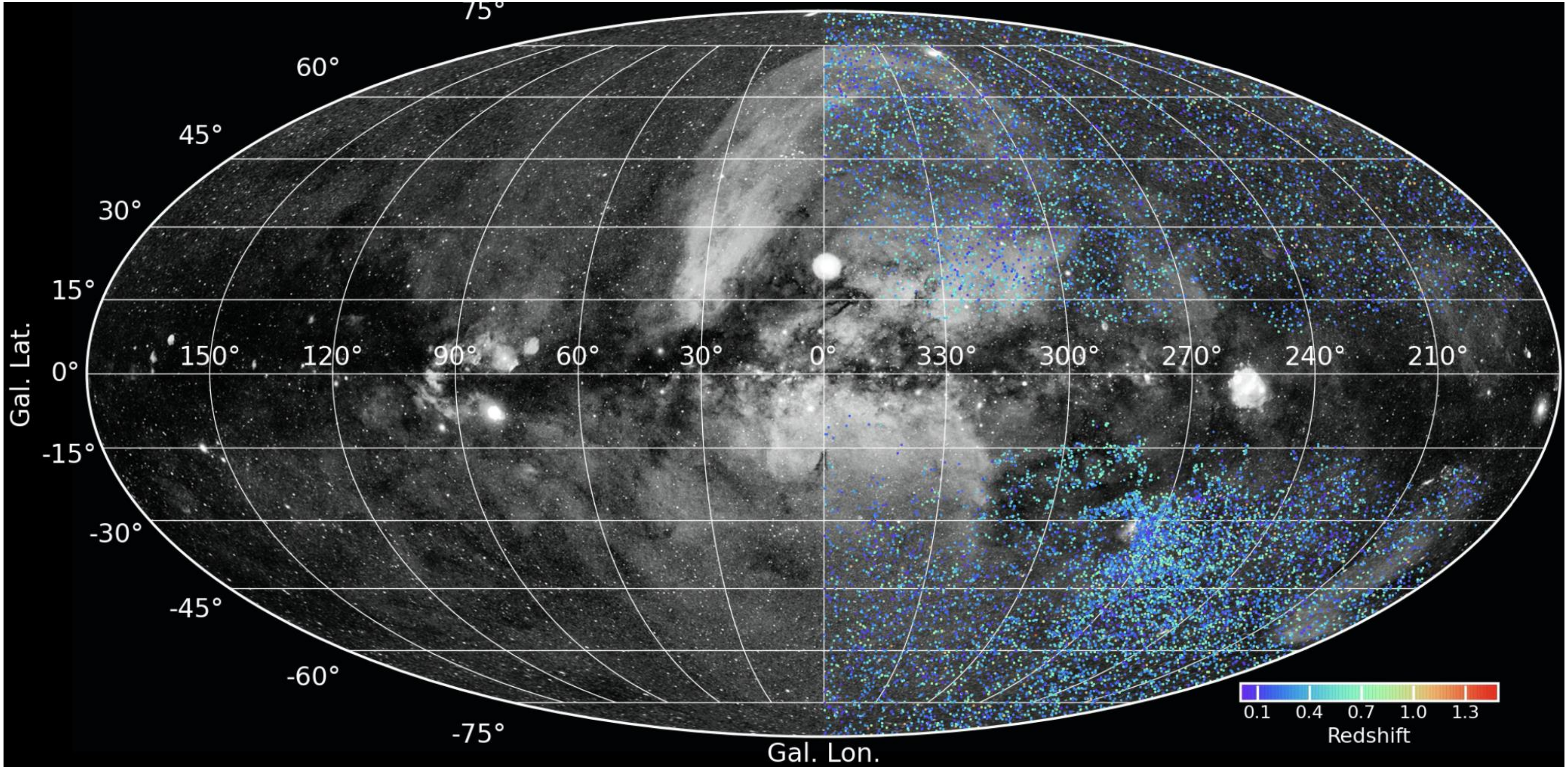
Name	Type	Unit	Description
DETUID	char[32]	None	Unique source ID
NAME	char[23]	None	Source name 1eRASS JHHMMSS.s+DDMMSS
RA	float64	deg	Right ascension (J2000)
DEC	float64	deg	Declination (J2000)
RA_XFIT	float64	deg	Best-fit RA from MBProj2D
DEC_XFIT	float64	deg	Best-fit DEC from MBProj2D
EXT_LIKE	float32	None	Source extent likelihood
DET_LIKE_0	float32	None	Source detection likelihood in 0.2-2.3keV
EXP	float32	None	Average exposure
BEST_Z	float64	None	Best available redshift
BEST_ZERR	float64	None	1 sigma error of BEST_Z
BEST_Z_TYPE	char[16]	None	type of BEST_Z
PCONT	float64	None	Probability of being a contaminant
CR300kpc	float64	s-1	0.2-2.3keV count rate within 300kpc
CR300kpc_L	float64	s-1	1sigma lower limit of CR300kpc
CR300kpc_H	float64	s-1	1sigma upper limit of CR300kpc
CR500	float64	s-1	0.2-2.3keV count rate within R500
CR500_L	float64	s-1	1sigma lower limit of CR500
CR500_H	float64	s-1	1sigma upper limit of CR500
CTS300kpc	float64	None	0.2-2.3keV counts within 300kpc

Important quantities include

- RA_XFIT, DE_XFIT (fitted MBProj2D positions)
- BEST_Z (best redshift)
- L500 (luminosity)
- KT (temperature, if bright)
- MGAS500 (gas mass)
- M500 (total mass)
- R500 (cluster radius)
- PCONT (contaminant probability)

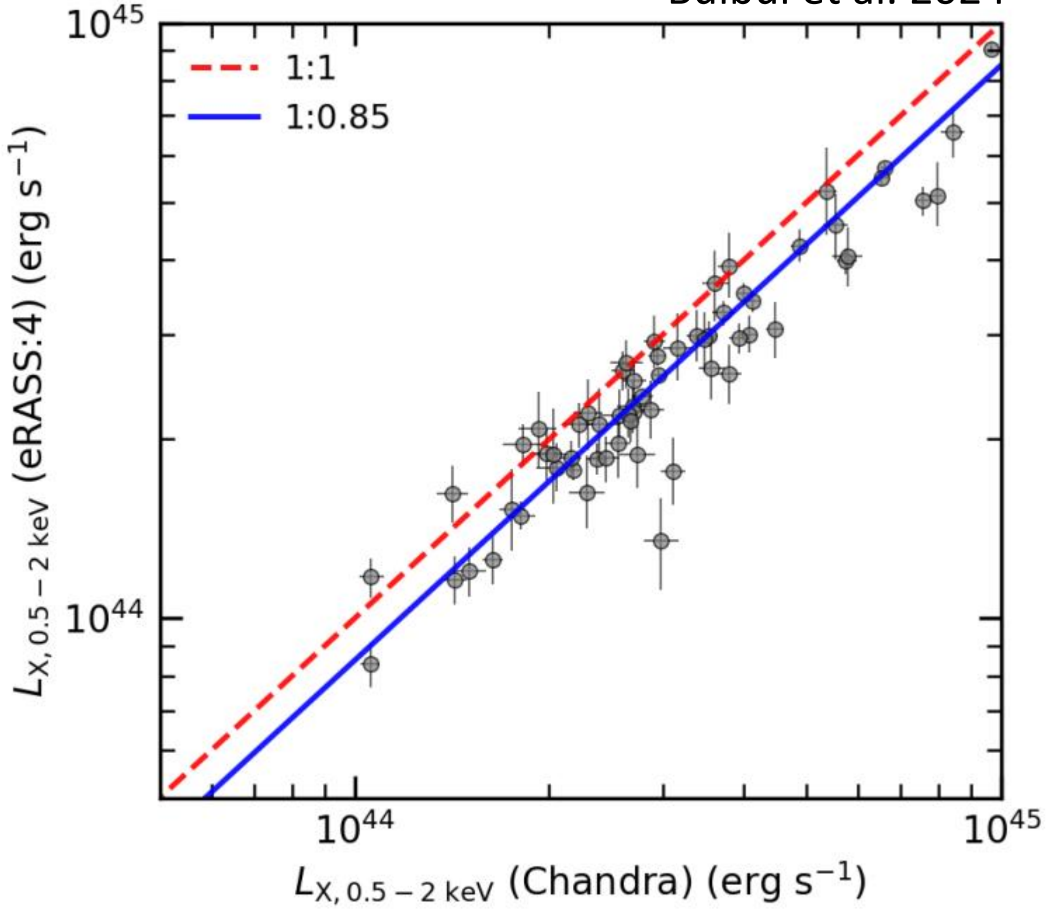
https://erosita.mpe.mpg.de/dr1/AllSkySurveyData_dr1/Catalogues_dr1/

Cluster sky distribution



Data issues: luminosity

Bulbul et al. 2024

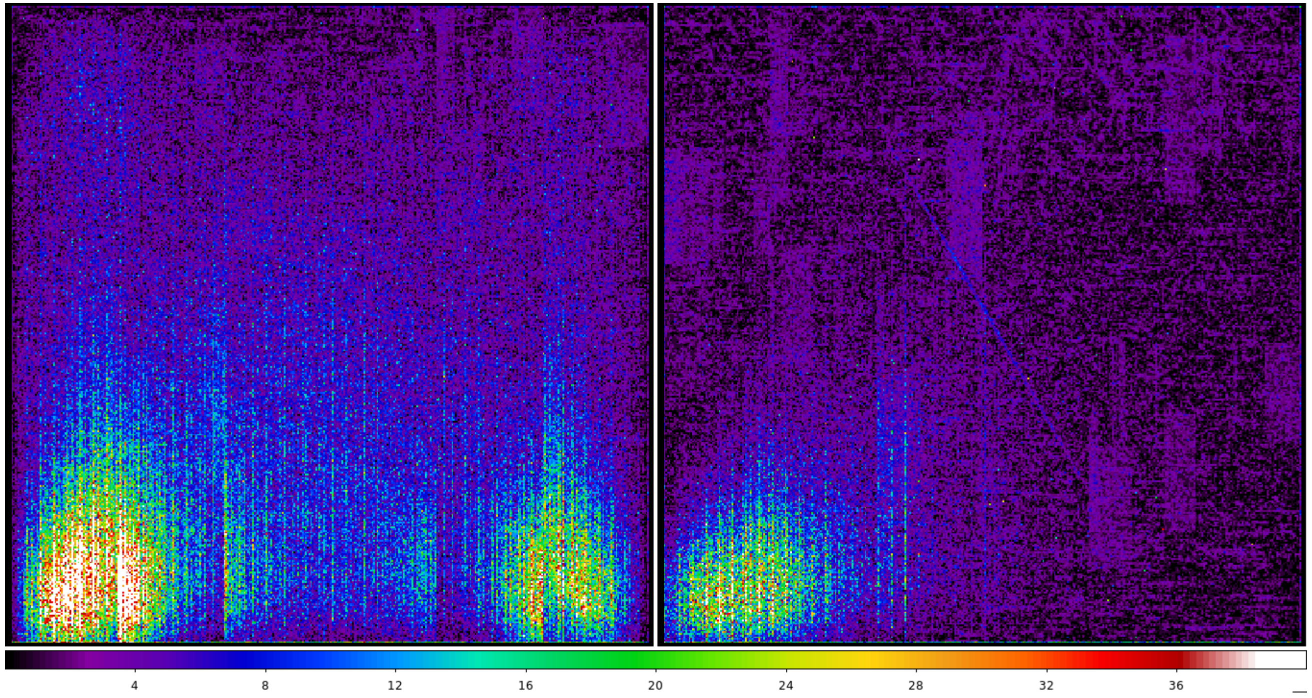


- Comparison of luminosities for SPT-detected clusters between eROSITA and Chandra
- eROSITA luminosities are around 15% lower than Chandra
- Similar for XMM
- Some of this (~6%) is due to photons being lost in the 010 early in the pipeline due to a CCD threshold – corrected in later processing

- In X-rays we don't know the true luminosities, but can just compare different measurements

Data issues: light leak

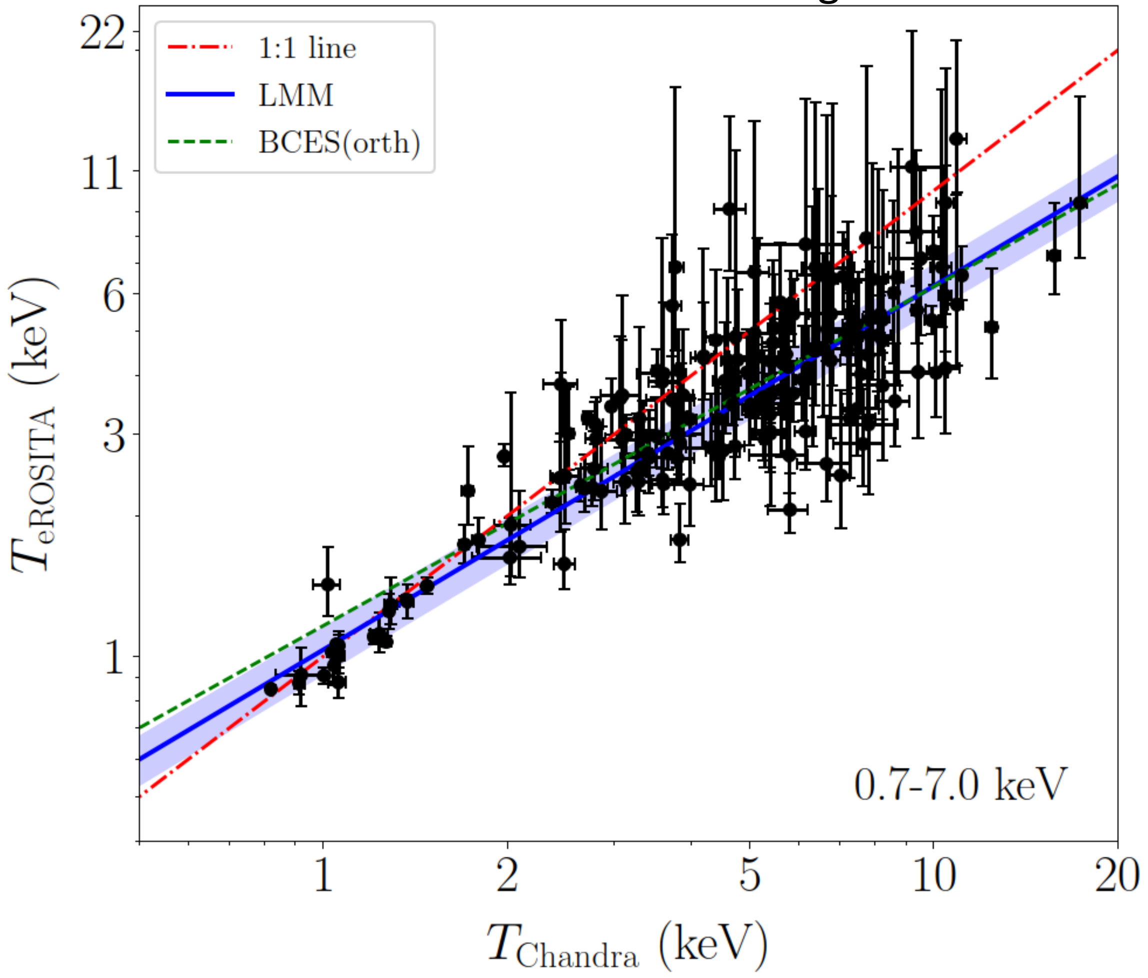
- Two cameras without on-chip filters (TM5+TM7) affected by optical light
- Dependent on solar angle
- Light causes high background rates at low energies
- Energy calibration of these cameras is poor
- OK for imaging above 0.3-0.4 keV
- Procedures to mitigate this under investigation
- Suggestion is not currently to use for spectral analysis



Images of the light leak in the two affected cameras in detector coordinates, TM5 (left) and TM7 (right)

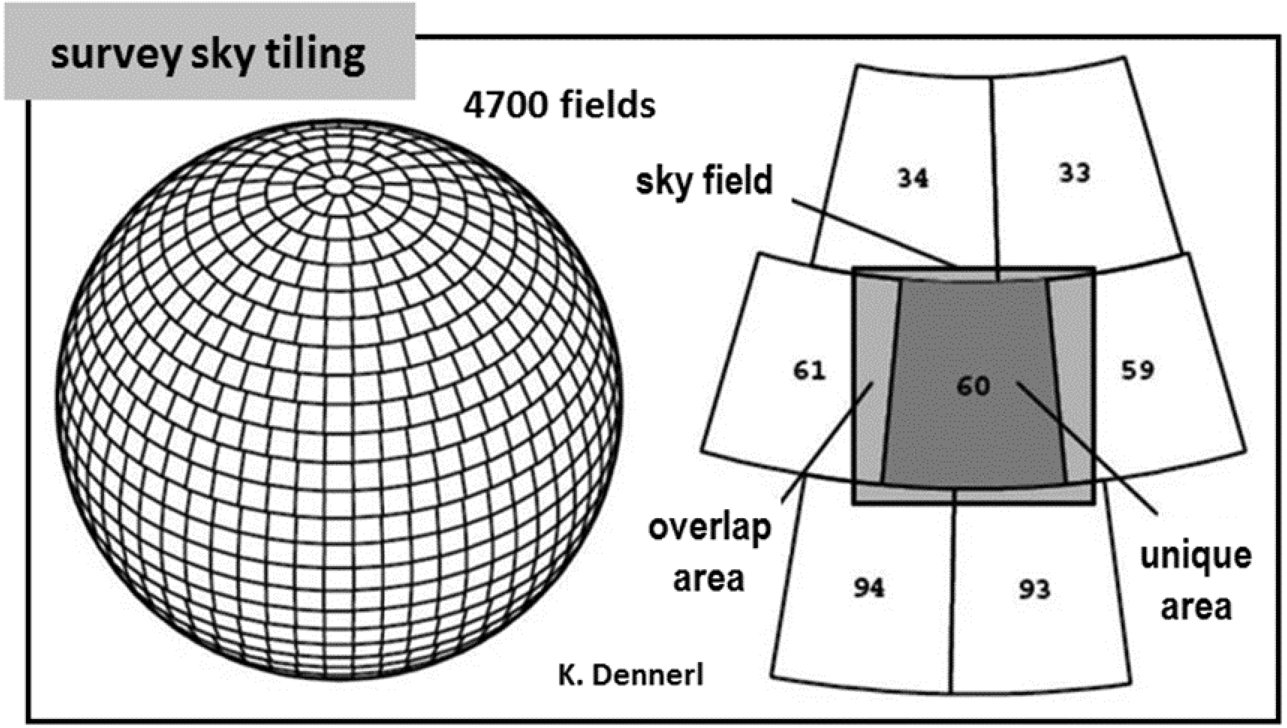
Data issues: high energy calibration

Migkas et al. 2024



- eROSITA does not have a lot of effective area > 2.3 keV
- Cluster temperatures from full band around
  - For 3 keV clusters
    - 20% lower than Chandra/ACIS
    - 14% lower than XMM/EPIC
  - For 10 keV clusters
    - 38% Chandra/ACIS
    - 32% XMM-EPIC
  - Small discrepancy for groups
- We don't know which is correct
- Investigations underway to check various aspects of calibration

eROSITA data



Sky is split into 4700 overlapping sky tiles (3.6°×3.6°), numbered using six digits using RA and Dec

- The sky is split into tiles. The event files and other products are created for each tile.
- Tiles are numbered RRRDDD, where RRR and DDD are three-digits based on RA and Dec in degrees
- Tile defined in a SKYMAPS.fits file obtained from the eROSITA website.

- You can map from position to tile number using this file, or using the page on the website: https://erosita.mpe.mpg.de/dr1/erodat/skyview/skytile_search/

SKYMAP.fits

fv: Binary Table of SKYMAPS_052022.fits[1] in /home/jsanders/data1/erosita/

File Edit Tools Help

Select	SRVMAP	OWNER	RA_MIN	RA_MAX	DE_MIN	DE_MAX	RA_CEN	DE_CEN
	1J	1I	1D	1D	1D	1D	1D	1D
All			deg	deg	deg	deg	deg	deg
Invert	Modify	Modify	Modify	Modify	Modify	Modify	Modify	Modify
1	1000	1	0.000000	360.000000	88.500000	90.000000	0.000000	90.000000
2	20003	1	0.000000	40.000000	85.500000	88.500000	20.000000	87.045181
3	60003	1	40.000000	80.000000	85.500000	88.500000	60.000000	87.045181
4	100003	1	80.000000	120.000000	85.500000	88.500000	100.000000	87.045181
5	140003	1	120.000000	160.000000	85.500000	88.500000	140.000000	87.045181
6	180003	1	160.000000	200.000000	85.500000	88.500000	180.000000	87.045181
7	220003	1	200.000000	240.000000	85.500000	88.500000	220.000000	87.045181
8	260003	1	240.000000	280.000000	85.500000	88.500000	260.000000	87.045181
9	300003	1	280.000000	320.000000	85.500000	88.500000	300.000000	87.045181
10	340003	1	320.000000	360.000000	85.500000	88.500000	340.000000	87.045181
11	11006	1	0.000000	22.500000	82.500000	85.500000	11.250000	84.042970
12	34006	1	22.500000	45.000000	82.500000	85.500000	33.750000	84.042970
13	56006	1	45.000000	67.500000	82.500000	85.500000	56.250000	84.042970
14	79006	1	67.500000	90.000000	82.500000	85.500000	78.750000	84.042970
15	101006	1	90.000000	112.500000	82.500000	85.500000	101.250000	84.042970
16	124006	1	112.500000	135.000000	82.500000	85.500000	123.750000	84.042970
17	146006	1	135.000000	157.500000	82.500000	85.500000	146.250000	84.042970
18	169006	1	157.500000	180.000000	82.500000	85.500000	168.750000	84.042970
19	191006	1	180.000000	202.500000	82.500000	85.500000	191.250000	84.042970
20	214006	1	202.500000	225.000000	82.500000	85.500000	213.750000	84.042970
21	236006	1	225.000000	247.500000	82.500000	85.500000	236.250000	84.042970

Go to: Edit cell:

https://erosita.mpe.mpg.de/dr1/AllSkySurveyData_dr1/

Tile numbers in SRVMAP

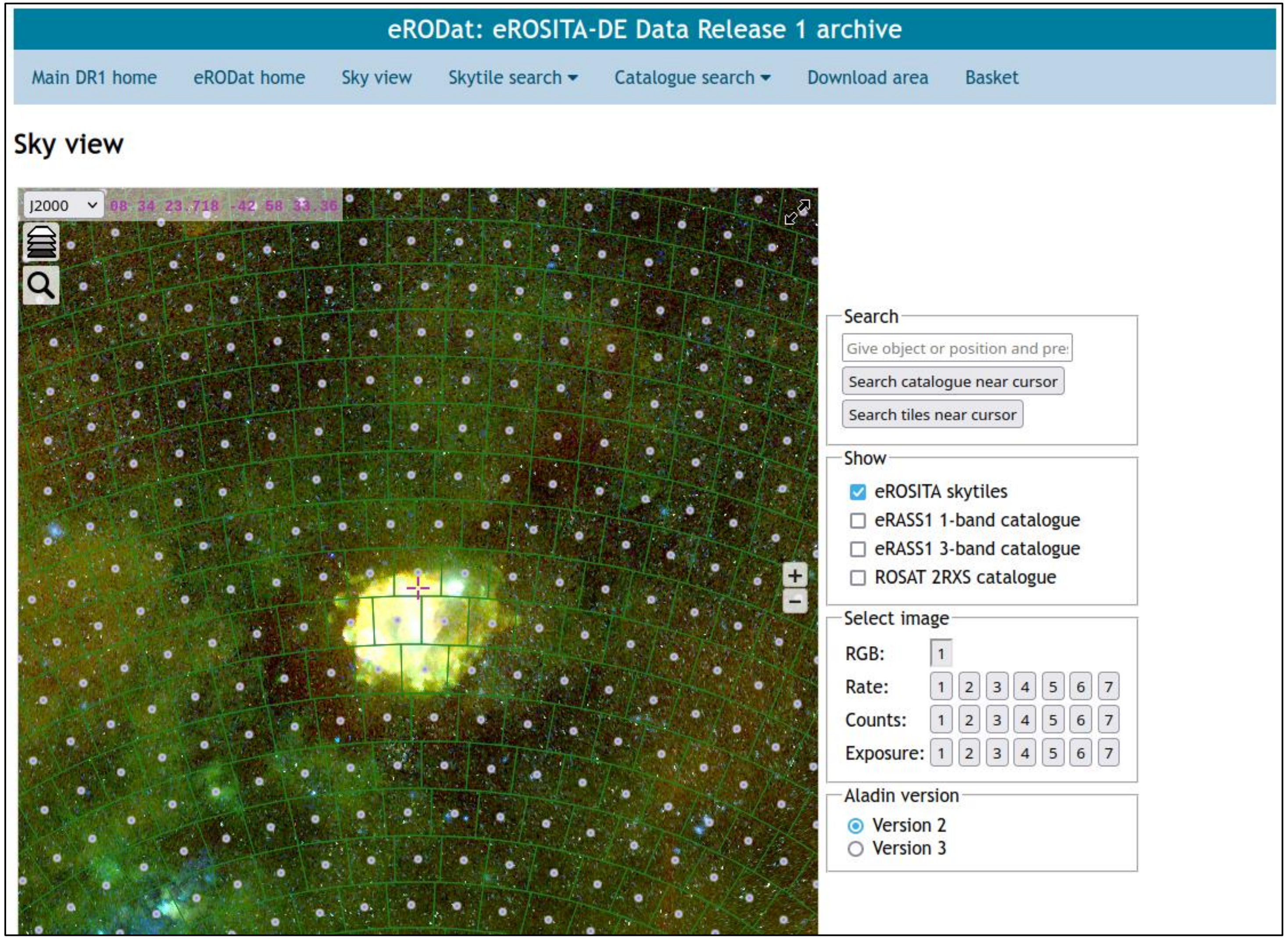
- SRVMAP: Number of tile
- OWNER: Whether a DE, RU or joint tile
- RA_MIN/RA_MAX: RA range
- DE_MIN/DE_MAX: Dec range
- RA_CEN/DE_CEN: Centre

RA/DE MIN/MAX define unique area for tile

Skytiles can also be returned by API:

https://erosita.mpe.mpg.de/dr1/erodat/skyview/skytile_search_api/?RA={RA}&DEC={DEC}&RAD={RADIUS}

Obtaining data



Data can be obtained through eRODat website: https://erosita.mpe.mpg.de/dr1/erodat/

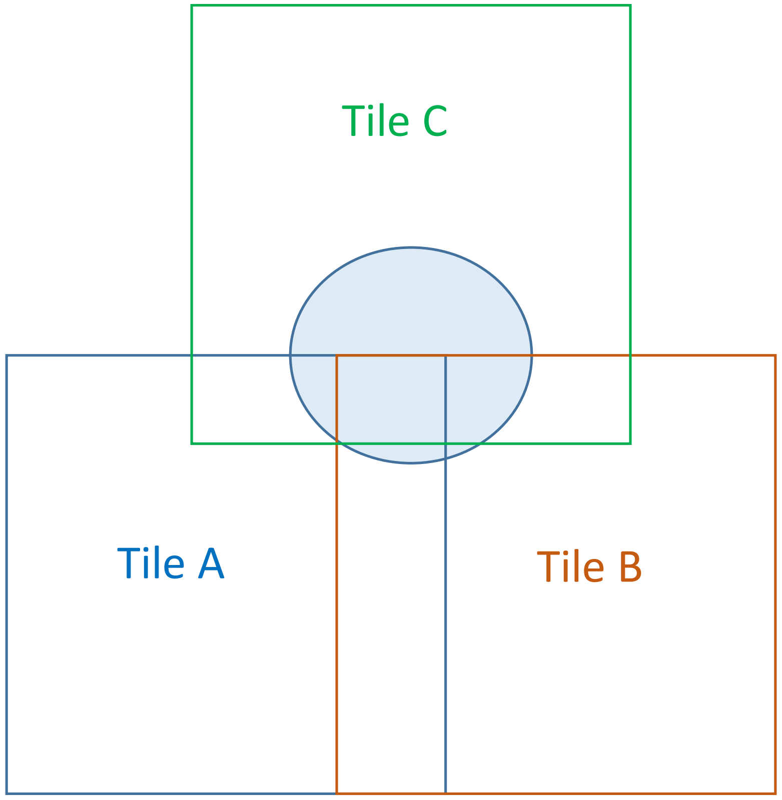
Skytile search will let you find the skytile(s) associated with the cluster

Data can either be downloaded by

- Click on data file directly
- Add to basket
- Download basket
  - As TAR file
  - Via script

- For clusters, the event files are likely the most useful

Merging sky tiles

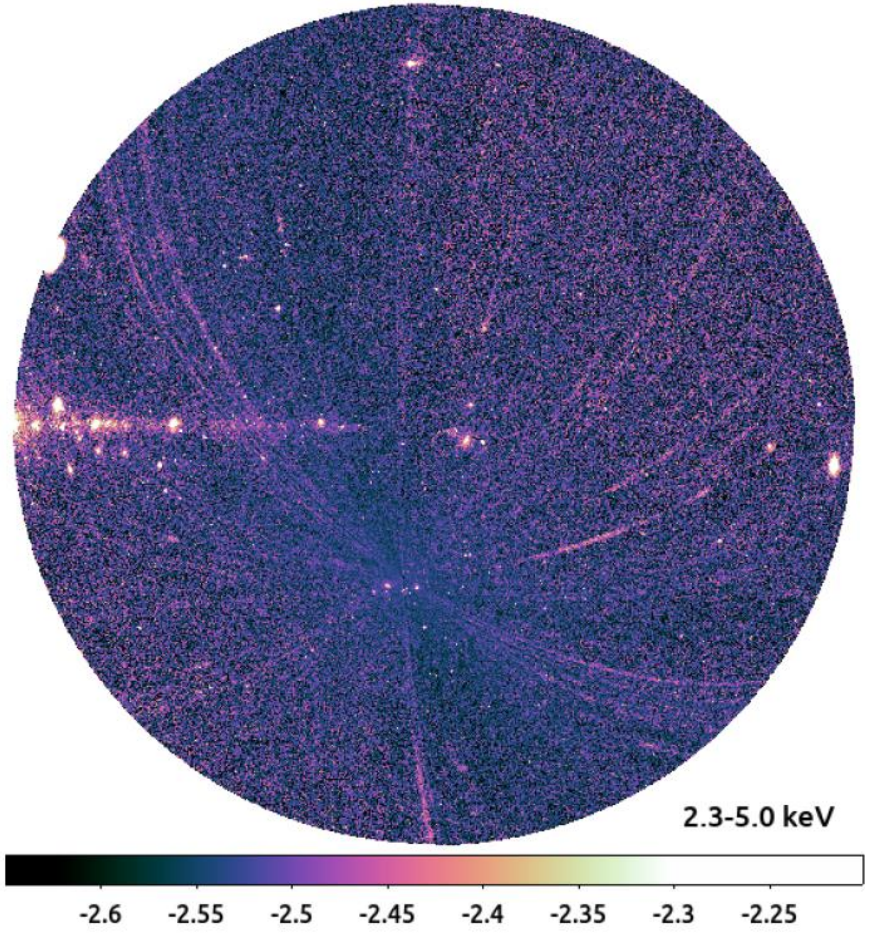


- If your cluster is large or lies near the edge of a tile, you may need to merge the event files to ensure complete coverage
- Duplicate events are removed during merging
- You can use the eROSITA website to find tiles within radius of a point
- Merge, then reproject to reduce geometric distortion from projection

- Merge tiles:
  evtool eventfiles="A.fits B.fits C.fits" outfile=out.fits
- Reproject sky coordinates around point:
  radec2xy file=evt.fits ra0=123.4 dec0=-12.3

Flaring

- There is little flaring of the X-ray background in the eRASS1 survey
- However, the eSASS tool **flaregti** can construct good time intervals
- It aims to optimize the detectability of faint sources and not leave gaps in the survey
- The most important parameter for clusters are the size (how big a source to optimize for) and the energy band
- For clusters and extended objects, a big size such as 100 arcsec is better



log count rate in harder band
(Merloni et al 2024)

Flaring

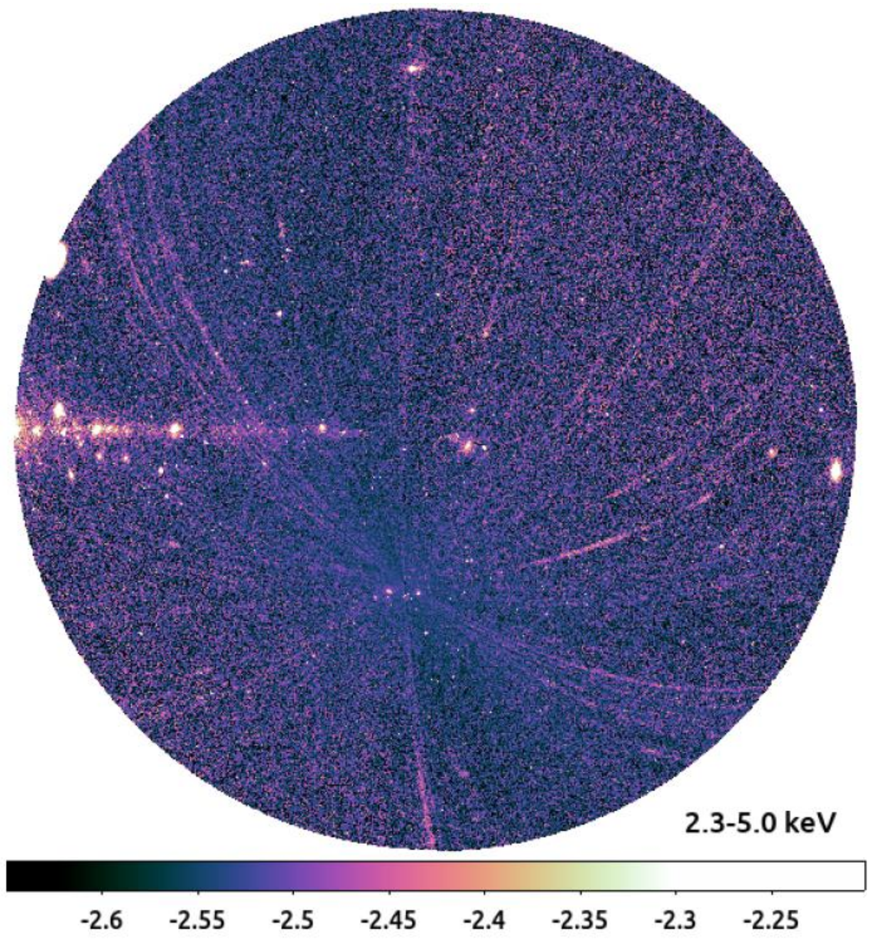
- Create **FLAREGTI** extensions

  flaregti eventfile=evt.fits size=100

- Filter input event file by its **FLAREGTI** extensions

  evtool eventfiles=in.fits outfile=out.fits gti=FLAREGTI

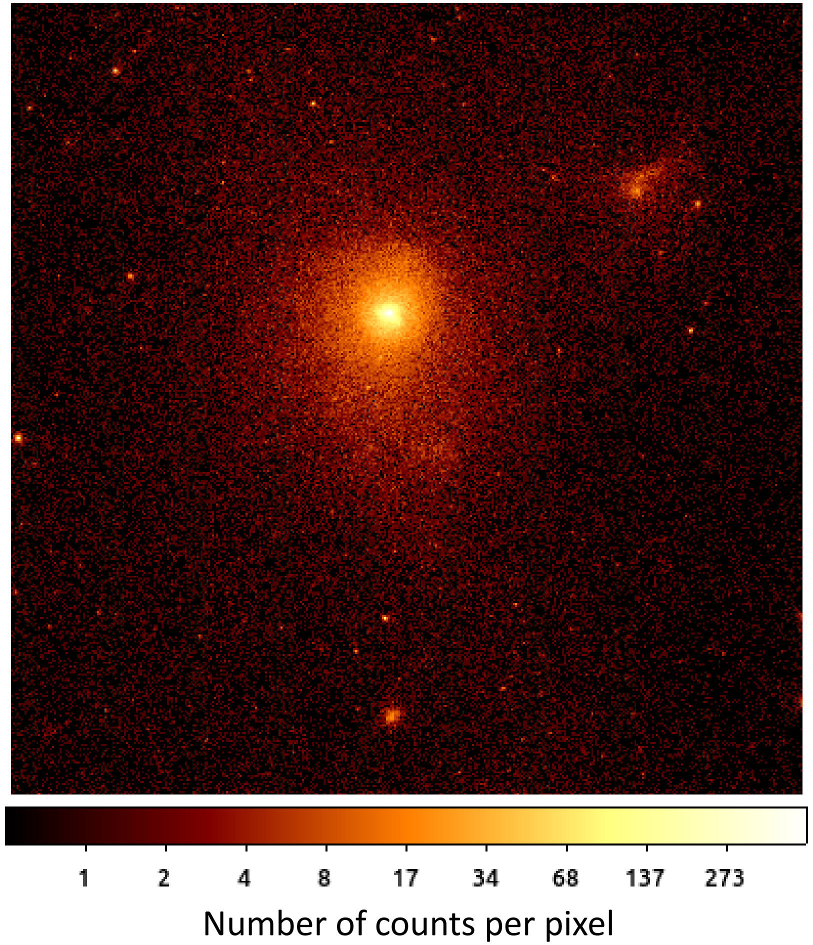
- For large areas, recommended to run on a merged file to get consistent filtering at borders
- Note: if you're using it on a merged event file, then see the known issues in the documentation!
- You should also modify the **gridsize** parameter and and range (**xmin, xmax, ymin, ymax**)



log count rate in harder band
(Merloni et al 2024)

Making images of clusters

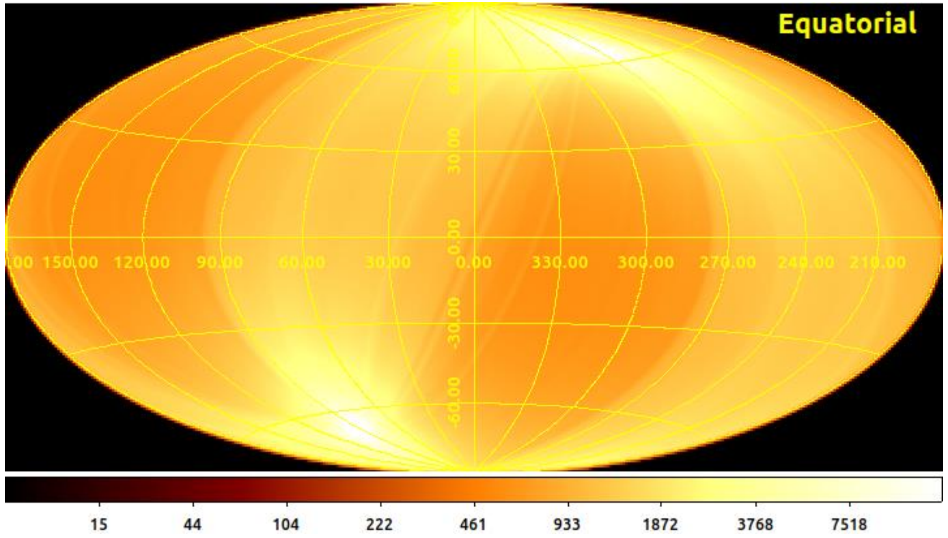
- **evtool infile=event.fits outfile=image.fits emin=0.3 emax=2.3 size=1000 rebin=80 events=no image=yes center_position="0 0"**
- Note: **rebin=80** (4 arcsec pixels) needed for source detection, though **events=yes** also needed for source detection using event-based mode
- **emin** to **emax** is the energy range in keV
- **center_position** optional
- Positions are in X,Y sky coordinates (relative to manual run of **radec2xy**, or centre of tile)
- **size** is in pixels



Exposure maps

- The exposure of the eROSITA survey varies over the sky
- To make a rate image we divide the count image by exposure map to make a rate image
- The merged exposure map (if **weights=yes**), gives the equivalent combined-TM (TM0) on-axis exposure time
- Exposure maps are either vignetted (suitable for objects or XRB), or non-vignetted (suitable for detector/particle background)
  - Vignetting accounts for effective area reduction as source goes off axis
  - Real data is a mixture of vignetted and non vignetted components (though one might dominate)

Full sky eROSITA exposure map

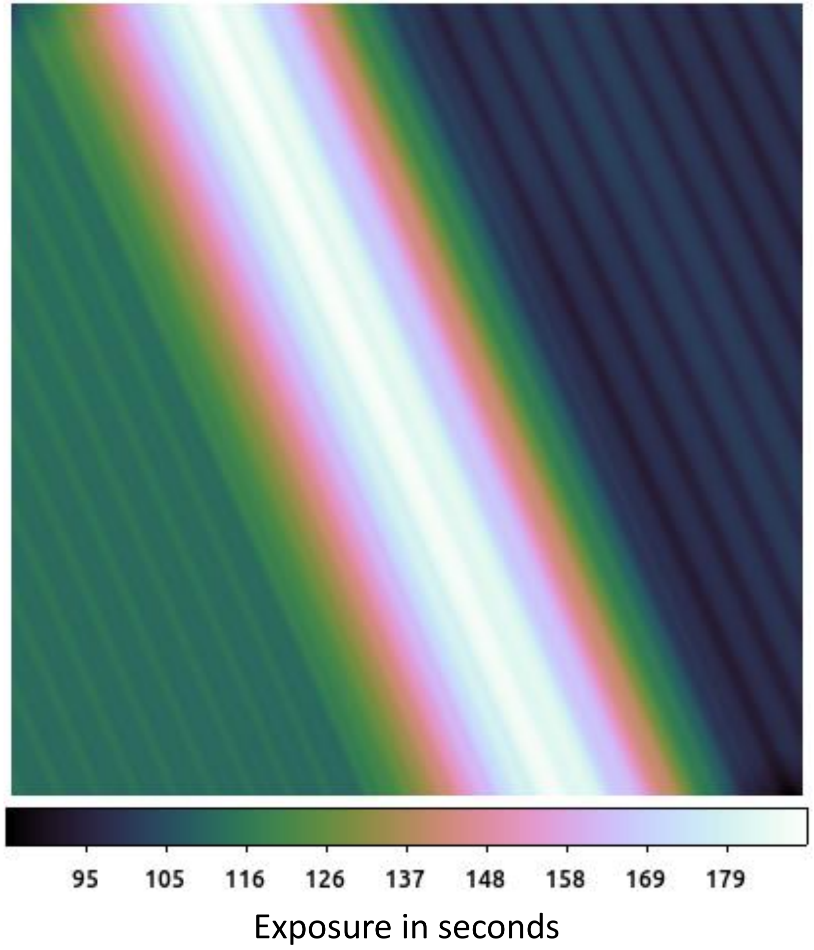


Exposure in seconds

Making exposure maps

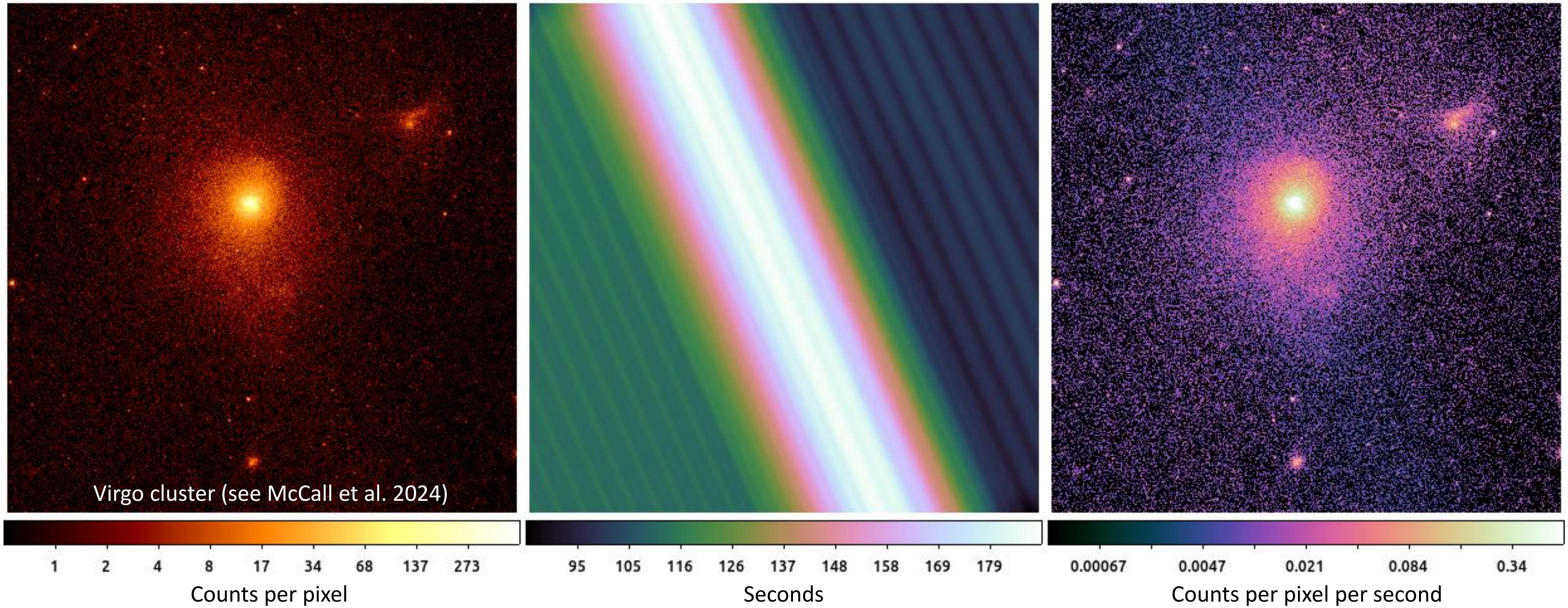
- expmap inputdatasets=input.fits templateimage=image.fits emin=0.3 emax=2.3 withvignetting=yes withdetmaps=yes withmergedmaps=yes mergedmaps=merged.fits

  - **input.fits**: input image or event file (needs eROSITA TM-specific HDUs)
  - **image.fits**: image containing WCS needed in output (can be same as above)
  - Weighting over energy range assumes a powerlaw model
  - **withdetmaps=yes**: *make sure you set this*, as the expmap otherwise includes detector regions filtered out in the data!
  - **merged.fits**: output image



Exposure correction

Count image | Exposure map image | Rate image (count/exposure)



- Calculate rate using, for example, Python or **farith counts.fits expos.fits rate.fits DIV**

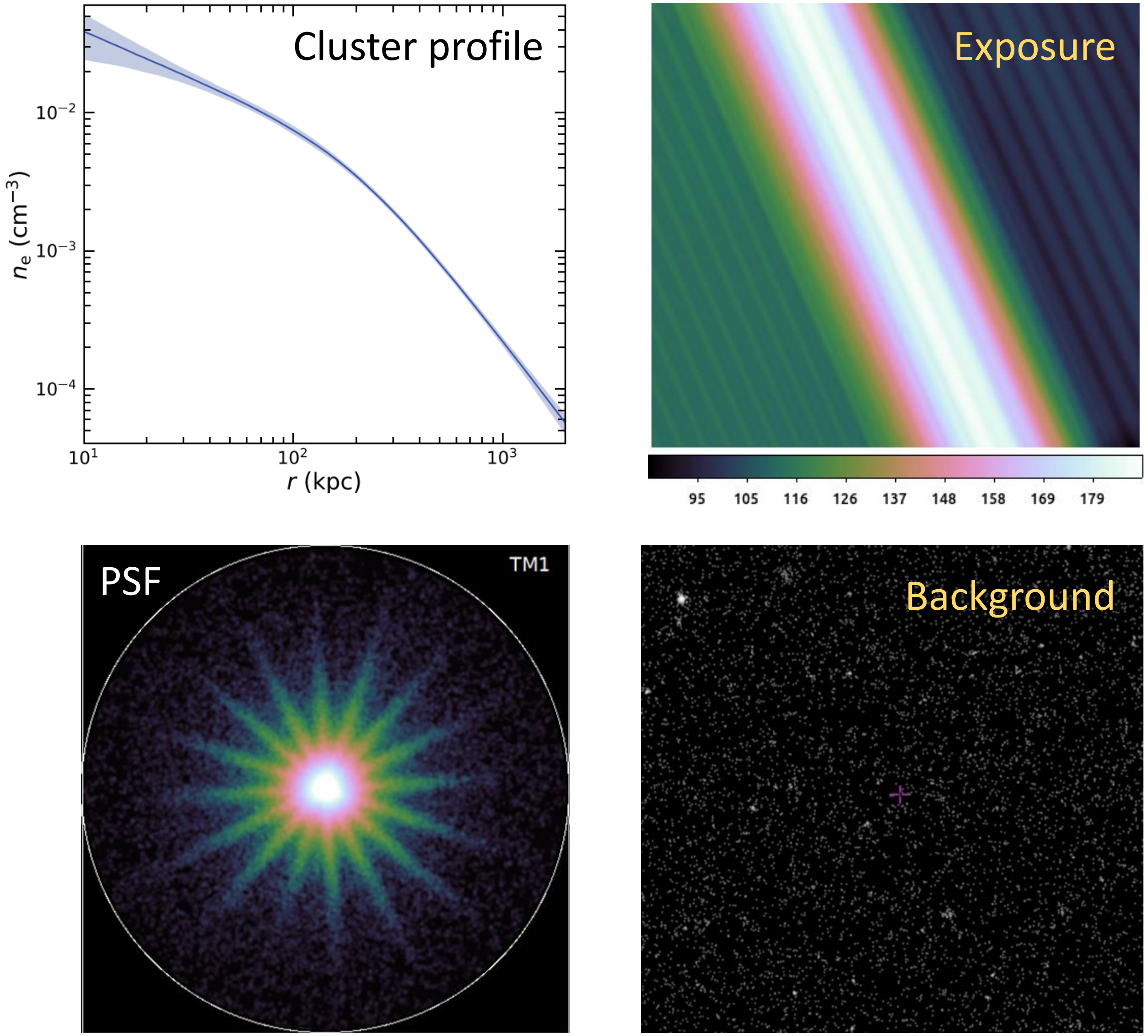
Making RGB images

- I normally use ds9 for making images generally
- Ds9 can also make RGB images
- You should make rate images in three bands (e.g. 0.3-0.6, 0.6-1.0, 1.0-2.3 keV – see the eROSITA effective area curve and model spectra for choosing ranges)
- These can be loaded into ds9 using
  - **`ds9 -rgb -rgb red red.fits -rgb green green.fits -rgb blue blue.fits`**
- It's hard to get a nice looking image. You likely will need to
  - Adjust ranges of each r,g,b channel in ds9 to show a sensible range
  - May need to smooth data (e.g. adaptive smoothing)
  - Adjust contrast/brightness in bands
- More of an art than a science!

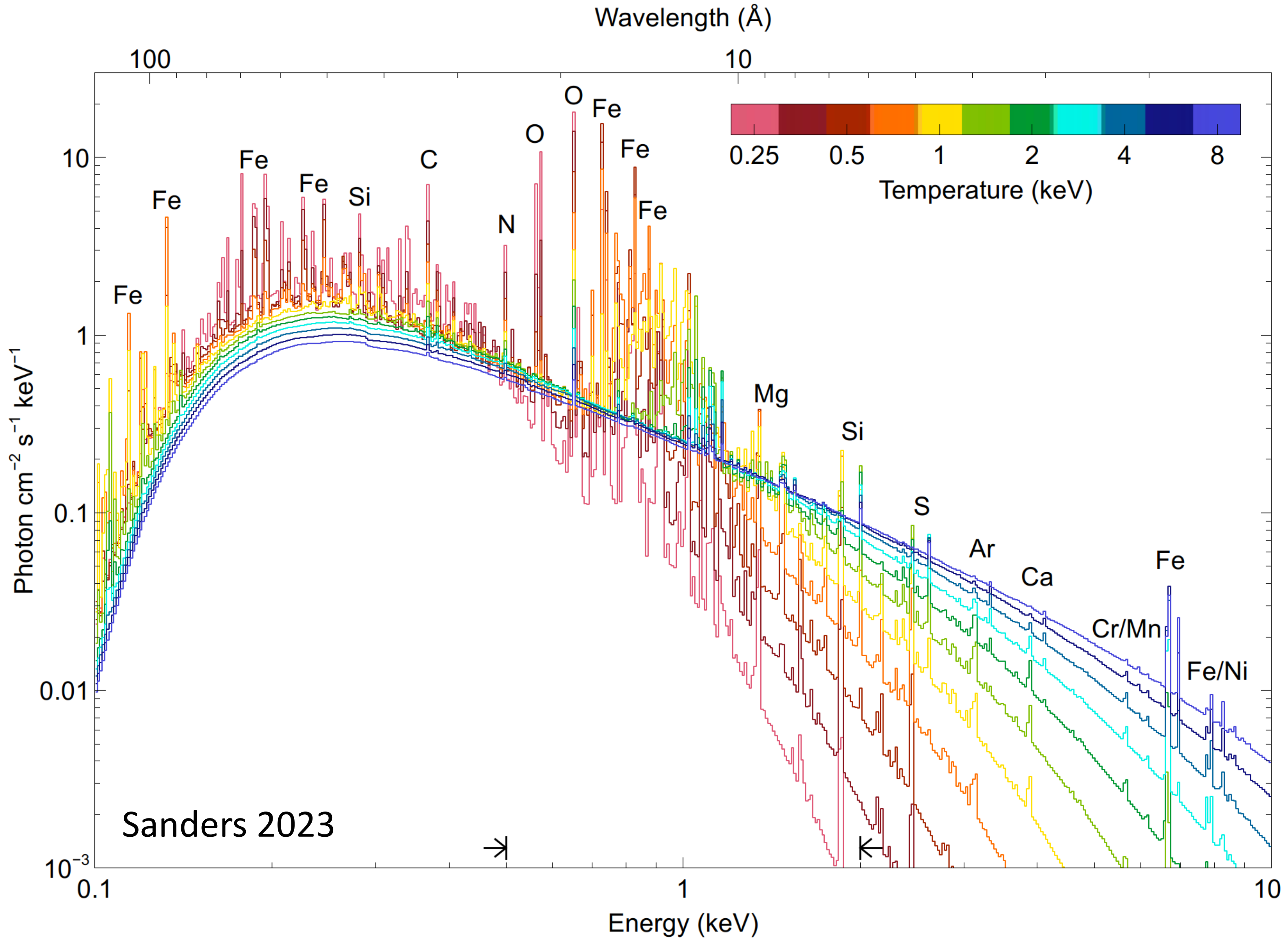
Abell 3266 Sanders et al. 2022

Fitting cluster data

- We want physical properties of eROSITA clusters
- We need to account for
  - Profile of cluster
  - Exposure
  - PSF of telescope
  - Background



Cluster spectra



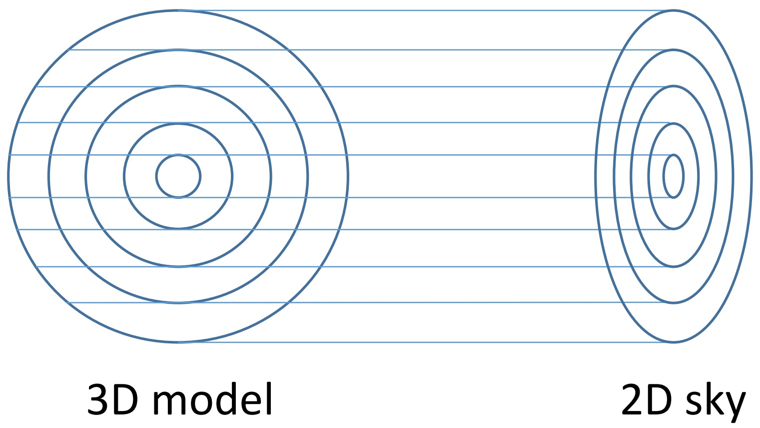
- At high spectral resolution, the spectra of clusters should be complex.
- However, CCD spectral resolution broadens the lines out, particularly at lower energies.
- Many eROSITA clusters only have a low number of counts, so high precision spectral analysis is impossible.
- Full spatial and spectral fitting is very expensive.

Fitting cluster data

- Single spectral fit cannot measure cluster profile
- Fitting spectra in annuli difficult, because
  - PSF mixes spectra in different annuli
  - eROSITA clusters are typically low count
  - Each annulus has few counts
- We want to simultaneously model the whole cluster to obtain parameters
  - Combine imaging and spectroscopy
- MBProj2D is software for modelling a cluster within Python, used in Bulbul et al. (2024)
  - Forward models images of cluster in multiple energy bands

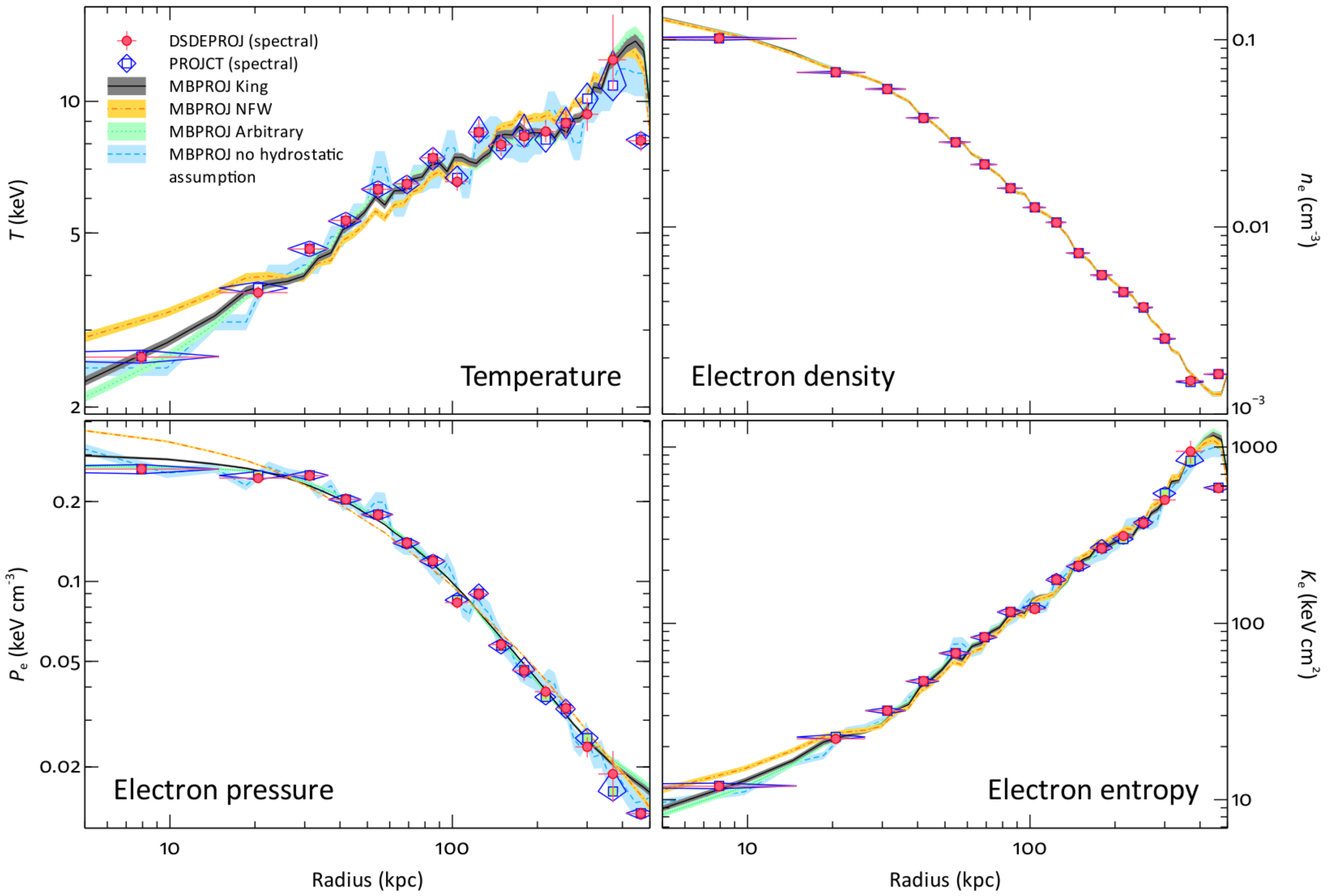
MBProj2D – cluster forward modelling code

- We have 1D radial profiles of gas density, temperature and metallicity
- We know the cluster redshift and Galactic absorption
- Given model parameters, we calculate the 3D emissivity of the cluster
- This is projected in 2D to compare to X-ray images
- By making images and models in different energy bands, we are sensitive to the spectral variation
- PSF and background can be included in model
- Use MCMC to obtain 3D model profiles and uncertainties



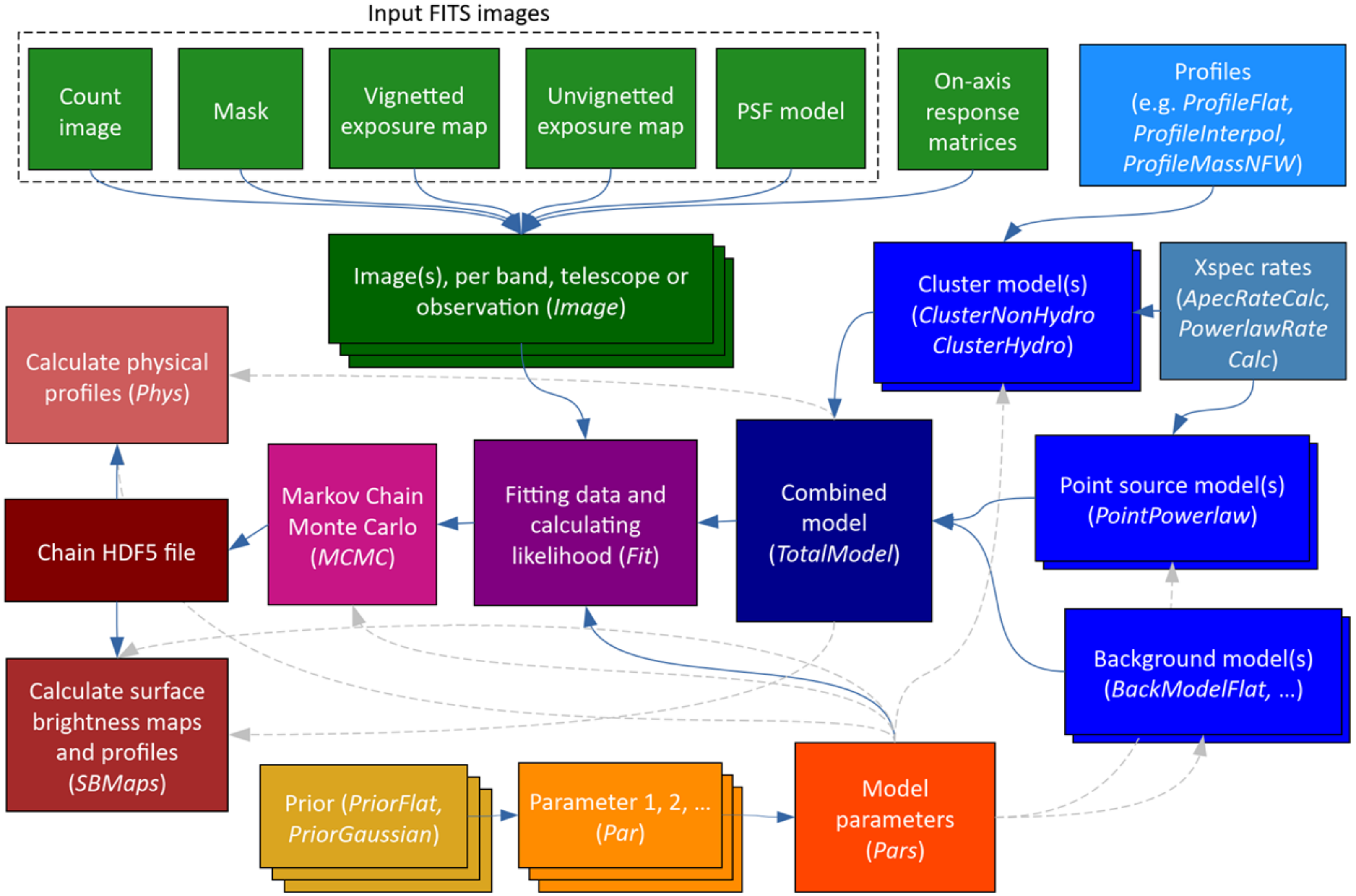
Comparison of methods for PKS 0745

Sanders et al. 2014



- Test of previous version of method, comparing spectral results against multiband imaging.
- Good agreement between two, though you need to carefully choose the bands used in the analysis to be sensitive to the parameters probed.

MBProj2D – block diagram



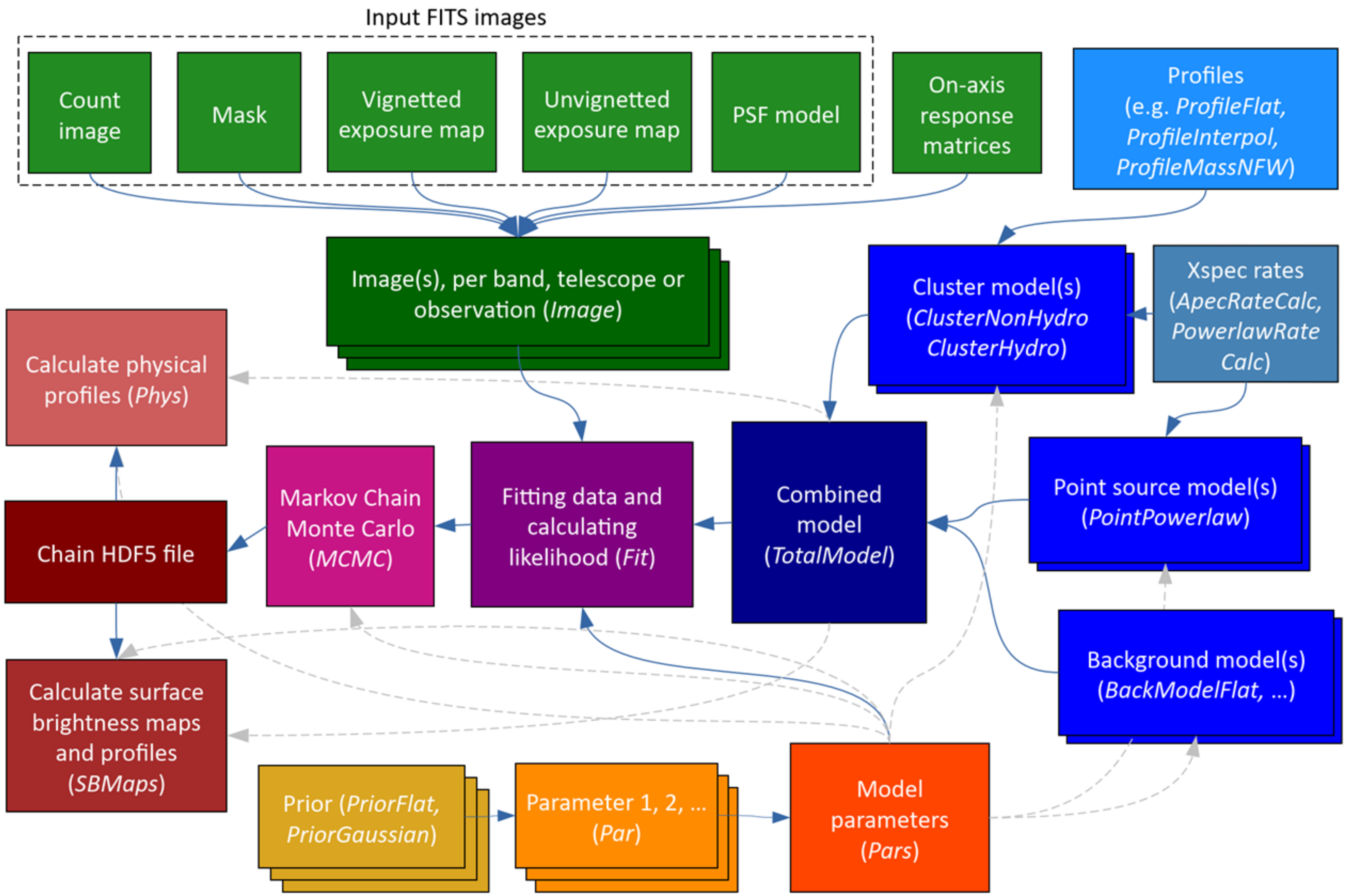
Inputs

- Images (in multiple bands)
- Mask – regions to include in fit (1=use, 0=exclude)
- Exposure maps (in multiple bands)
- PSF models (in each band)
- Response matrices

Model

- Define N clusters and point sources in image
- For clusters, choose model profiles for density, temperature, metallicity
- Choose model for background (e.g. flat)
- Construct total model

MBProj2D – block diagram



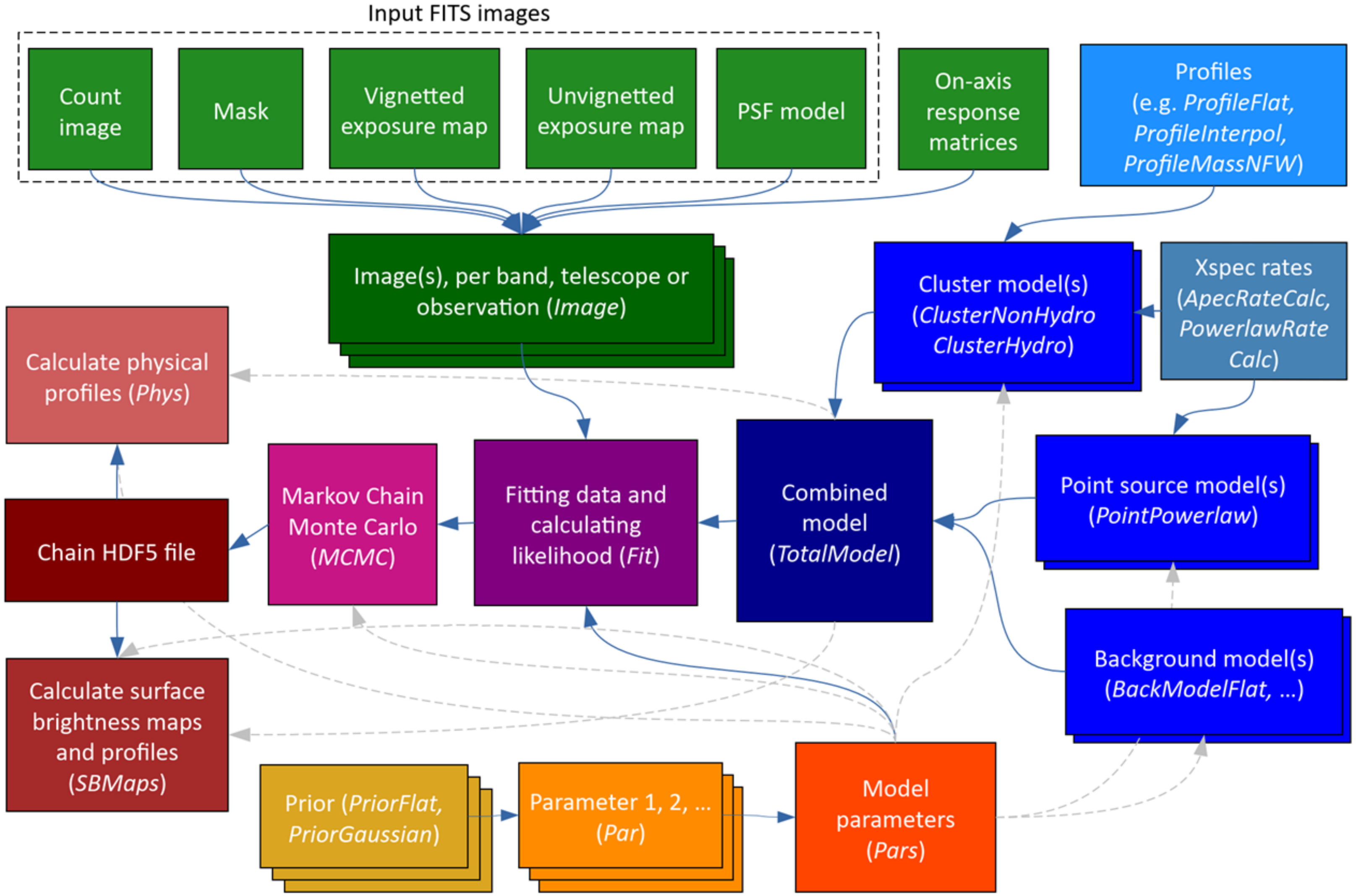
Fitting

- For model components we have parameters
- Choose priors on parameters, if necessary
- Find best fitting model

MCMC

- Use MCMC to get posterior probability distributions on model parameters
- Use emcee sampler by default, choosing number of walkers, burn in and chain length
- Written to a chain output file

MBProj2D – block diagram



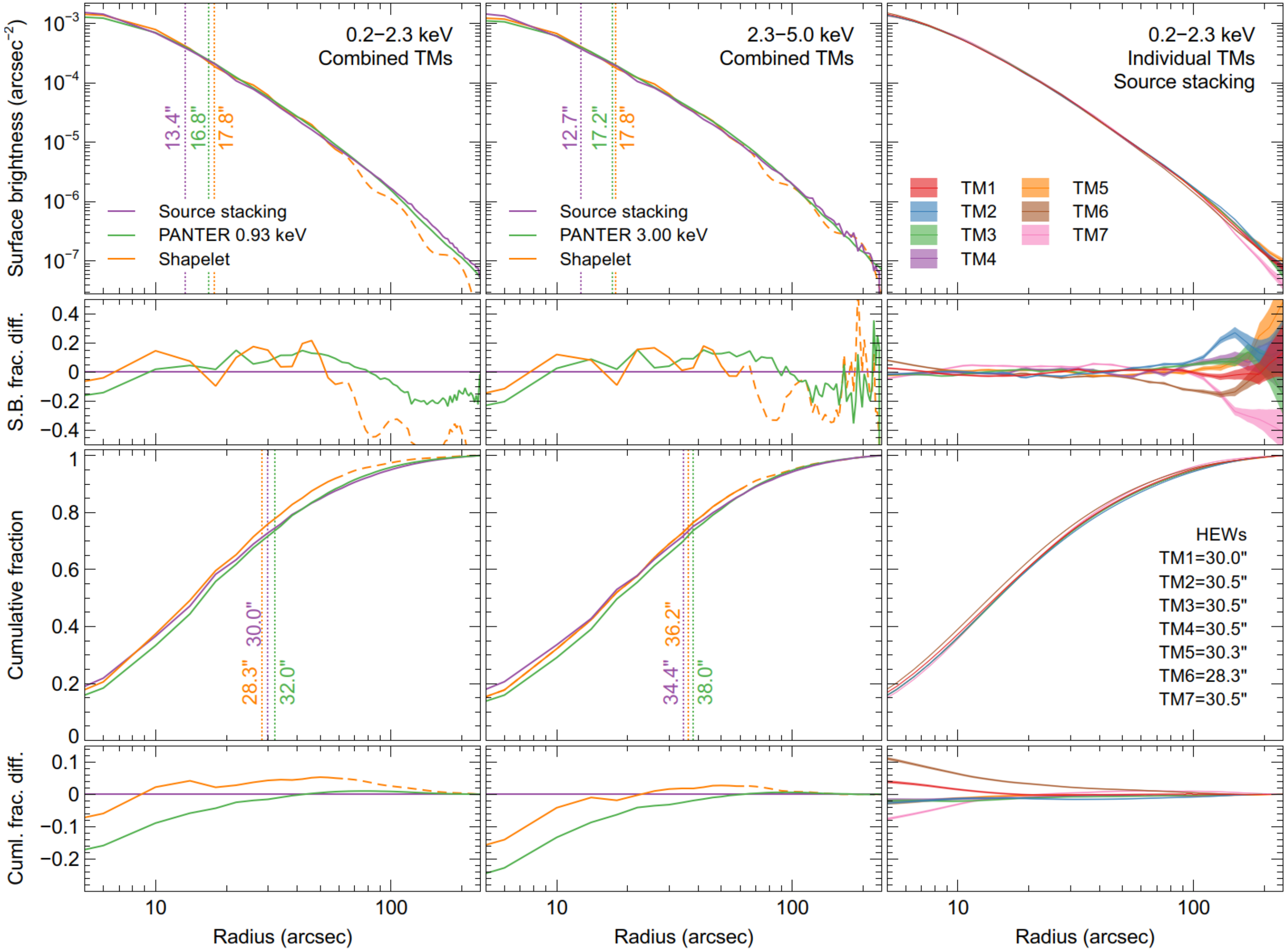
Physical profiles

- We want the physical profiles (e.g. density, pressure, gas mass) from the analysis
- Replay values in MCMC chain, calculating the temperature, density and metallicity profiles each time
- Calculate derived quantities (e.g. pressure) from these
- Look at distribution of physical profiles
- Output written to a FITS or HDF5 file

Inputs

- Mask – define using eROSITA source catalogue
  - Can either fit point sources or mask them out
  - Might want both depending how bright sources are
- Images and exposure maps created using eSASS
  - Choose bands to be sensitive to spectral variations (e.g. 0.3-1.0, 1.0-2.3, 2.3-6.0 keV)
- Response and ARF are standard on-axis eROSITA files ARFs for TM0 (note merged eROSITA exposure is defined for combined TMs=TM0)
- PSF taken from calibration database

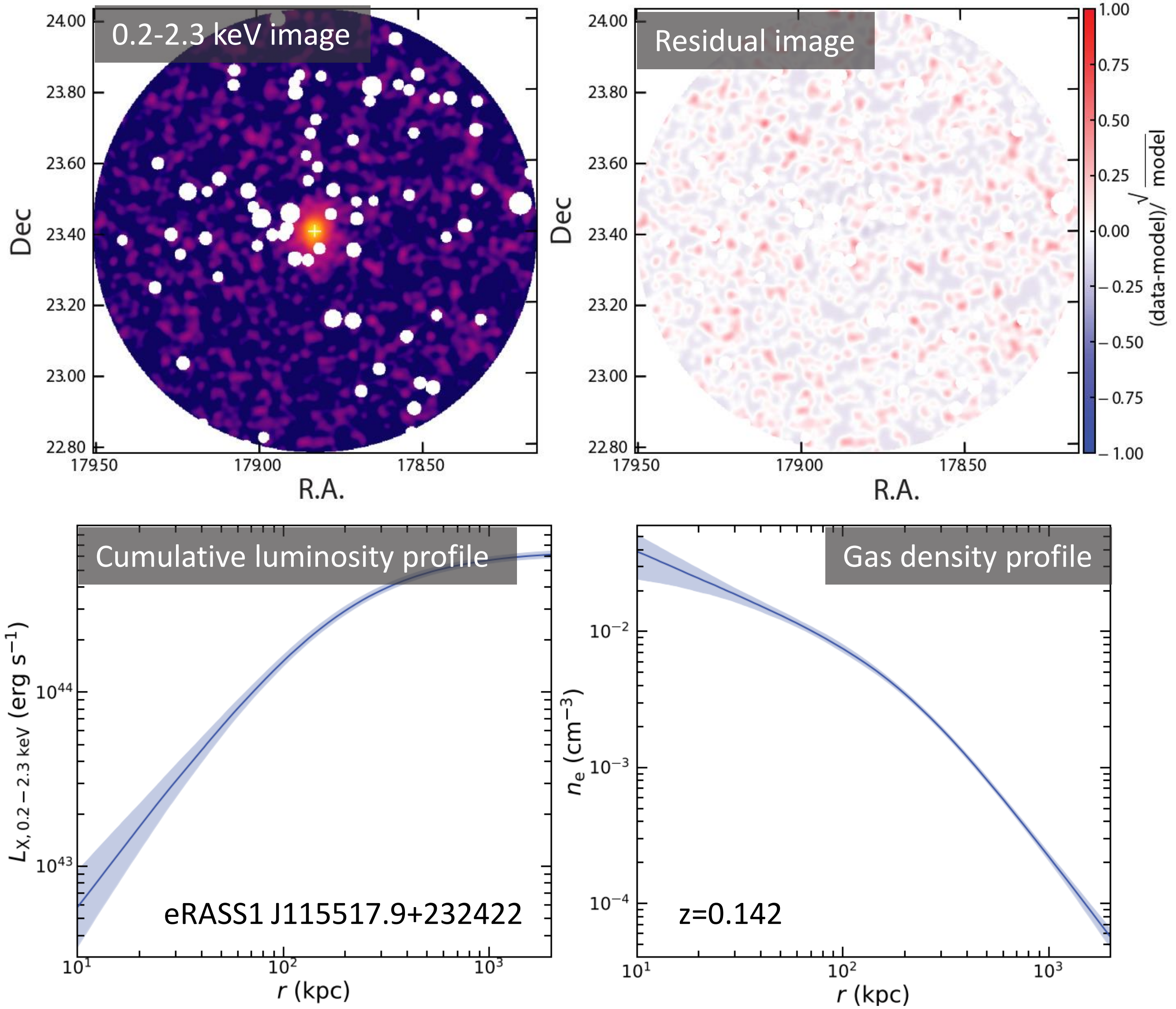
Point Spread Function (PSF)

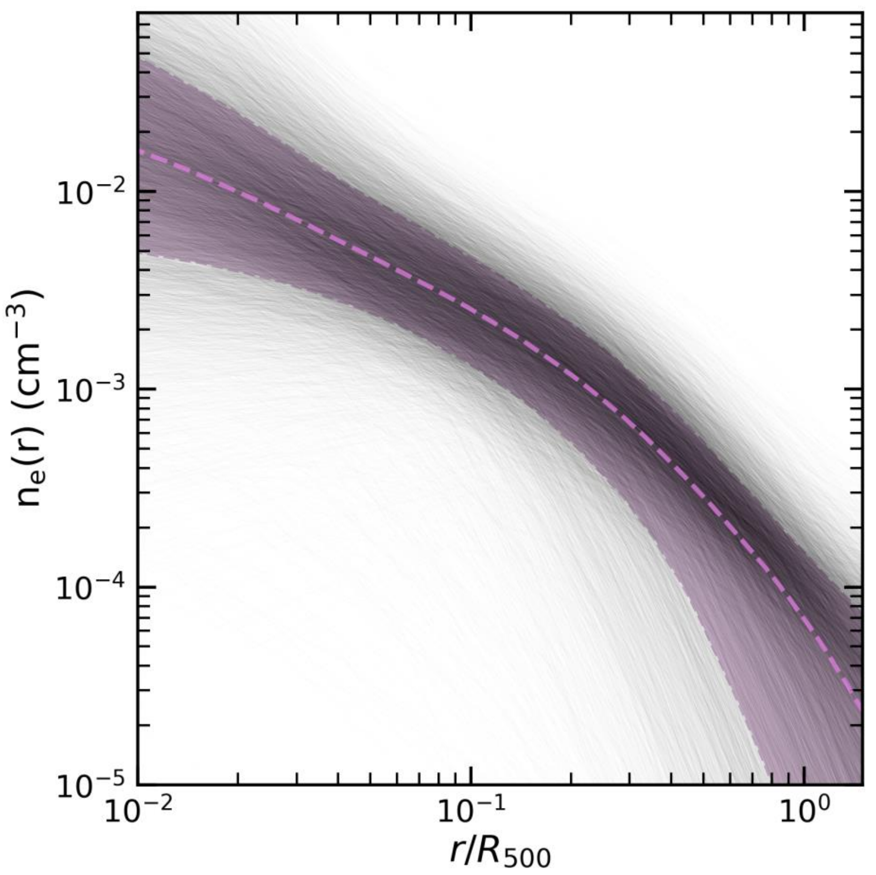


Merloni et al. (2024)

- Due to survey scanning, average PSF is uniform over the sky.
- Currently in CALDB, all TMs have same PSF.
- See calibration database file TMx_2dpsf_19022v03.fits for survey average.

Examples from Bulbul et al. 2024





Density profiles of eRASS1 clusters
(and 1 sigma percentiles)

For more details about MBProj2D

- https://github.com/jeremysanders/mbproj2d
- https://mbproj2d.readthedocs.io/en/latest/
- Sanders et al. 2014 (Feedback, scatter and structure in the core of the PKS 0745-191 galaxy cluster)
- Sanders et al. 2018 (Hydrostatic Chandra X-ray analysis of SPT-selected galaxy clusters)
- Bulbul et al. 2024

- eFEDS example: https://github.com/jeremysanders/mbproj2d/tree/master/examples/eFEDS

Spectral analysis

- If we want to go beyond multiband imaging analysis we need to make spectra
- Spectral analysis of extended sources more complex than point sources in eROSITA
- eSASS task **srctool** makes spectra (and lightcurves) of sources
- For extended sources, we need to care about
  - Extraction region
  - Background
  - PSF
  - Spectral model

https://erosita.mpe.mpg.de/dr1/eSASS4DR1/eSASS4DR1_tasks/srctool_doc.html

Extraction regions

- **srctool** supports some common region types – specified on **srcreg** and **backreg** parameters
- Types: **circle, annulus, ellipse, ellannulus, box, rectangle**
- Components can be subtracted from previous components
- If you need a more general shape, then there is the **mask** region type, where the user can supply a FITS file, where >0 is included, <=0 excluded (needs valid WCS)
- Region position can be ***,*** which is replaced by source position (**srctool** is designed for multiple sources)

Source model

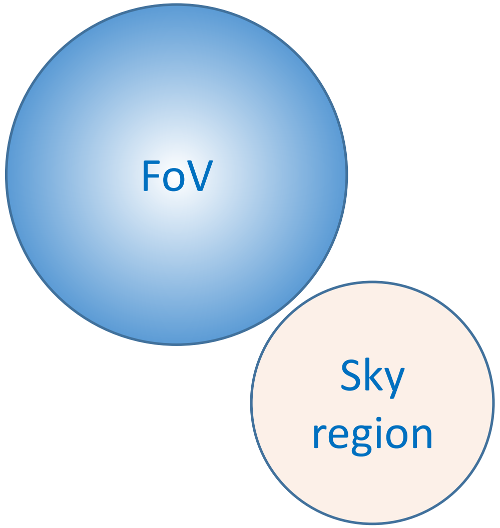
- **srctool** weights regions according to a source model (e.g. to calculate effect of source lost outside FoV, PSF correction or vignetting)
- Options are POINT, GAUSSIAN, BETA, TOPHAT, MAP
- Parameters for BETA (r_c, cut off radius), GAUSSIAN (sigma, cut off radius) and TOPHAT (outer radius)
- MAP type takes a FITS file – recommended for bright nearby systems
- MAP type does not support PSF correction
- However, *srctool PSF correction is not advised for large regions – it is designed for compact sources*
- PSF correction assumes monochromatic source

Background spectra

- **srctool** extracts background spectra with the *same time intervals* as used for the source (i.e. when the source is inside the FoV). This is appropriate for variable point sources.
- For large background regions this is not good
- Therefore, for large sources it is advised to extract backgrounds using a separate run

ARFs, exposures and areas

- In a survey, exposures and effective area is harder to define
- Exposure time in **srctool** is the amount of time the source is within the field of view (after deadtime + GTIs)
- In contrast, it is the ARF takes account for various factors
  - Bad pixels
  - When part of the source is outside the FoV
  - Average vignetting of source
- Take care if giving rates from spectra, as the ARF is an effective one
- In the output spectrum there are useful measures of source area
  - BACKSCAL – average area in deg^2 of the source in the FoV during the exposure
  - REGAREA – geometric area of extraction region
  - RGDMAREA – geometric area where source model>0 in extraction region



Background modelling

- You can extract a local background and subtract it in e.g. Xspec. This is likely ok if the source isn't too large.
- Otherwise, simultaneously model your background and source region with a model consisting of the X-ray background components (e.g. AGN, soft thermal components) and the filter wheel closed (FWC) background
- The appropriate area scaling for non-vignetted model component (e.g. particle background) is BACKSCAL
- The appropriate area scaling for the source model component is REGAREA/RGDMAREA
- If simultaneously modelling, you may need different ARFs for cluster and its background, particularly if the spatial distribution is different!

See e.g. https://erosita.mpe.mpg.de/dr1/AllSkySurveyData_dr1/FWC_dr1/ and Sanders et al. 2022, and session in this school on local Galactic emission

Intracluster medium modelling

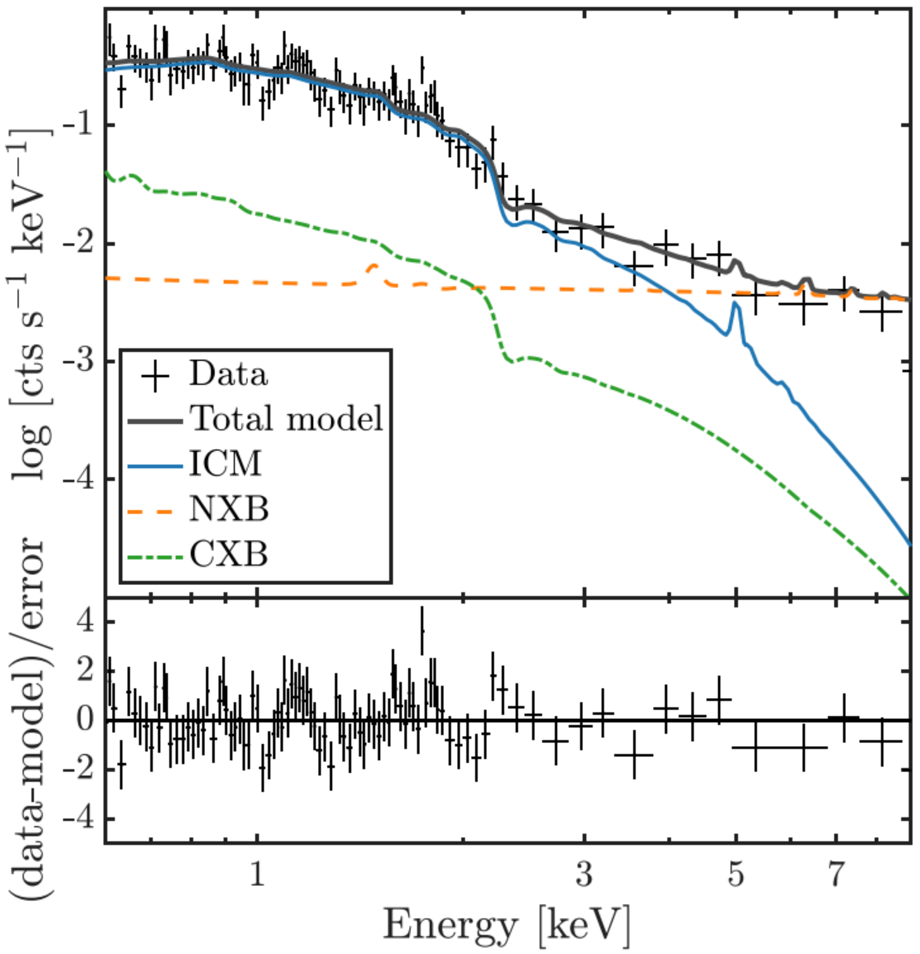
- Typically use APEC component to model thermal emission (SPEX model also available)
  - Temperature
  - Metallicity (often assume 0.3 Solar for fainter clusters)
  - Redshift
  - Normalisation (proportional to n_e^2 dV, integrated over source)

- Also photoelectric model absorption (e.g. PBABS, PHABS)

Running time

- srctool models various effects by sampling spatially and temporally
- The density of these samples greatly affects the runtime
- For large regions, srctool likely too slow with default parameters
- **xgrid** controls spatial grid sampling (8" units) – runtime is $\propto$ **xgrid**$^{-2}$
- **tstep** controls time sampling (seconds) – runtime is $\propto$ **tstep**$^{-1}$
- eROSITA scans at 90" per second
- If the region is large, suggest increasing **xgrid** from 1 or perhaps **tstep** from **0.05**
- Up to 7 cores can be used by setting environment variable **OMP_NUM_THREADS=7**

Example spectral fit



Cluster eFEDS J092121.2+031726 at redshift 0.333

Cluster in the eFEDS survey (Liu et al. 2021)

This example merges all 7 TMs

Temperature is 5.2 keV and soft band luminosity is 2×10^{44} erg s^{-1}.

End of lecture

Next comes the hands on session…