COSMOS PROJECT REPORT

Created	@Dec 9, 2020 5:00 PM
Created By	S Sicong Huang
Last Edited By	S Sicong Huang
Last Edited Time	@Jan 23, 2021 12:37 PM
Stakeholders	
Status	In Review 👀
Type	Technical Spec

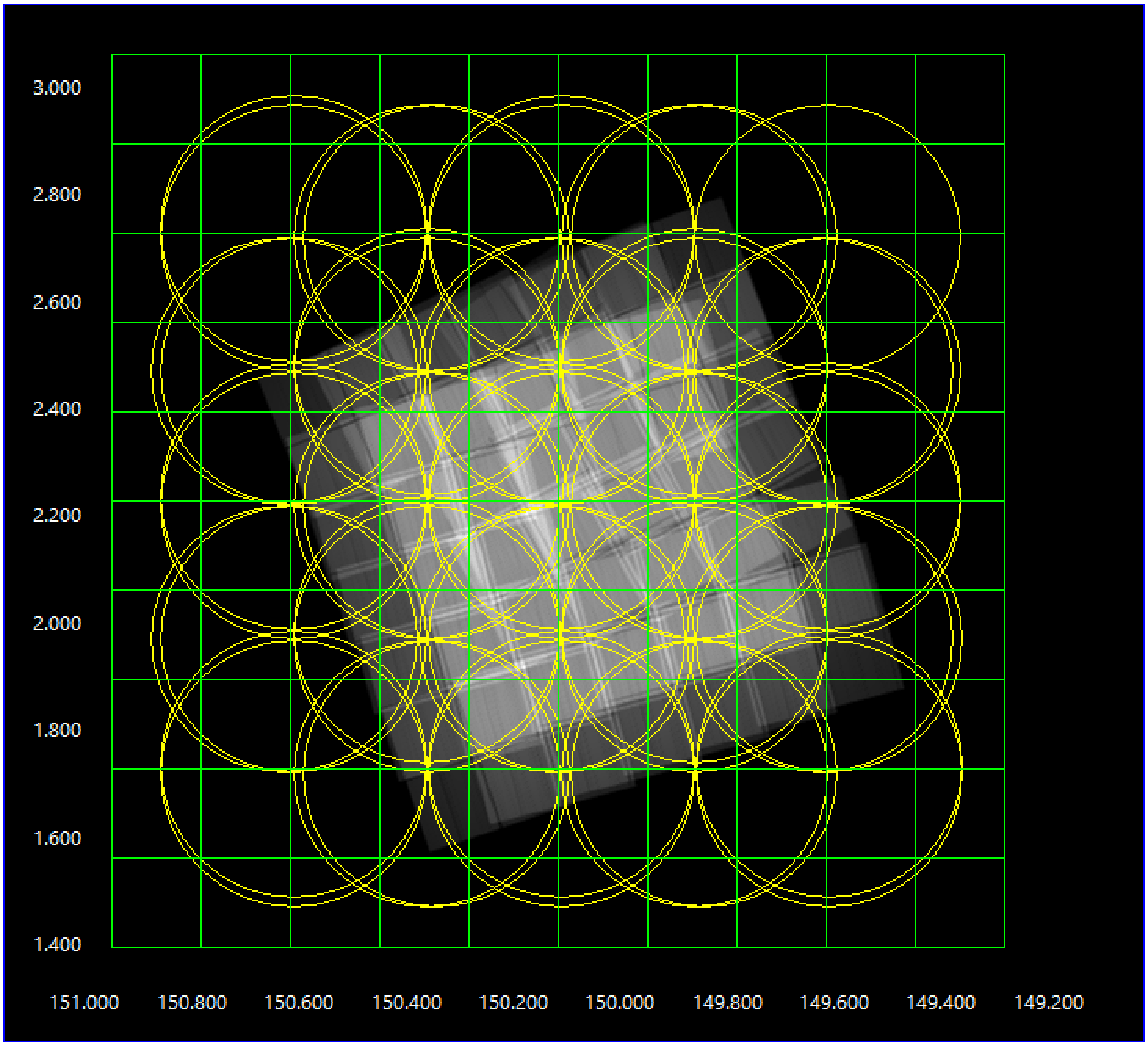
Summary
Background
Goals
Proposed Solution
Data preparation
Notice
Phase 1
Notice
Phase 1 Follow-up Tasks
Phase 2
Notice
Phase 2 Follow-up Tasks
Results
Phase 3
Notice

Summary

This project generates and fits spectra from XMM observations of the COSMOS field to study the potential spatial variation of Emission Measure of Galactic Halo/Circum-galactic medium.

Background

- The full 2 deg^2 COSMOS field is covered entirely by 56 XMM pointings for a total exposure ~ 1.5 Ms.
- To study the spatial variation of the GH/CGM EM, we divided the whole field into 10×10 regions, each with 10' in width and height.



The image of the COSMOS field. The white blocks at the center is Chandra C-COSMOS observations; the yellow circles are XMM-COSMOS pointings; the green boxes are the 10×10 regions that covers the whole field.

- We remove the point sources detected by Chandra and XMM, and generated spectra in each box region.

Goals

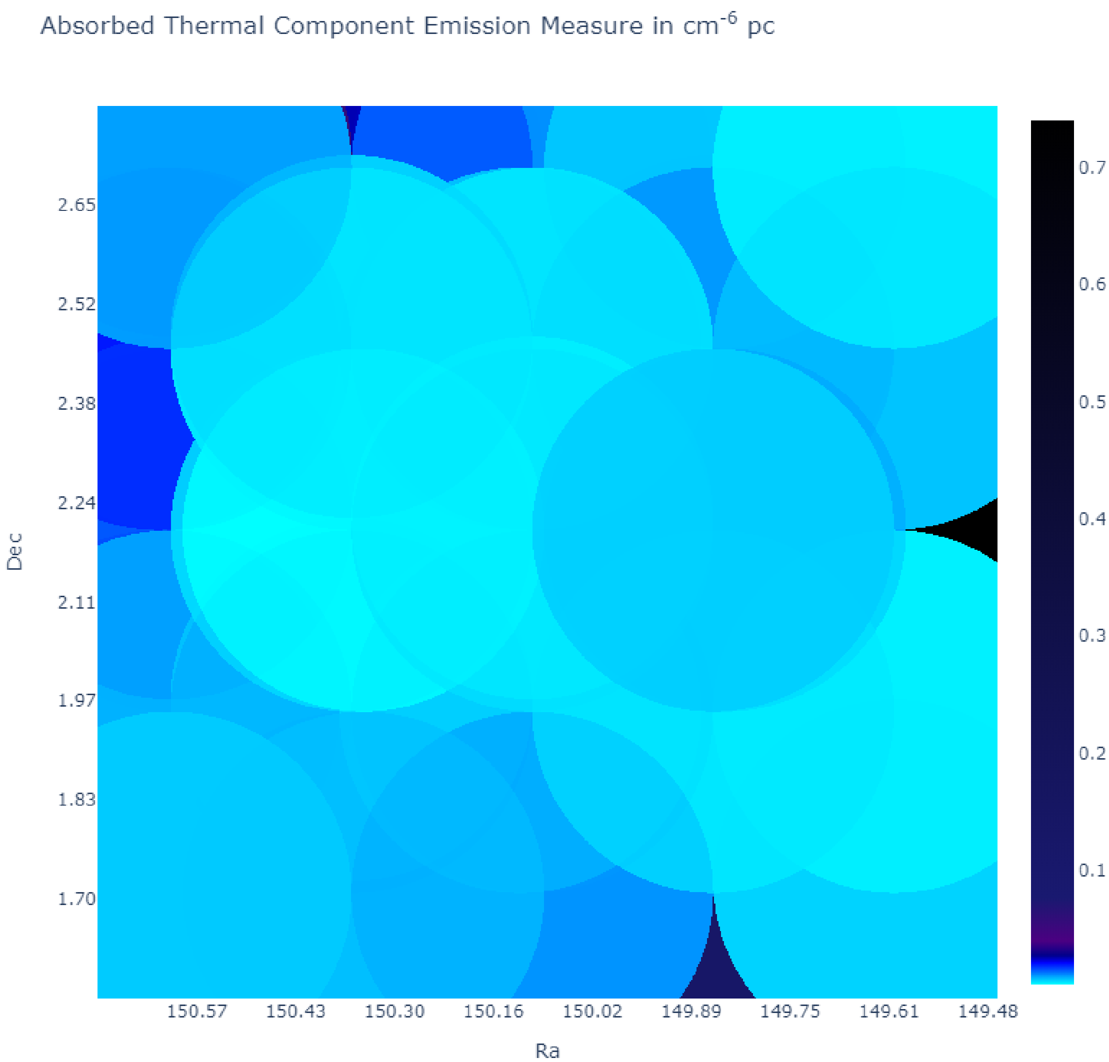
By fitting the spectra in each box region, we can visualize the variation of EM and determine quantitatively if the variation is real or simply due to statistics.

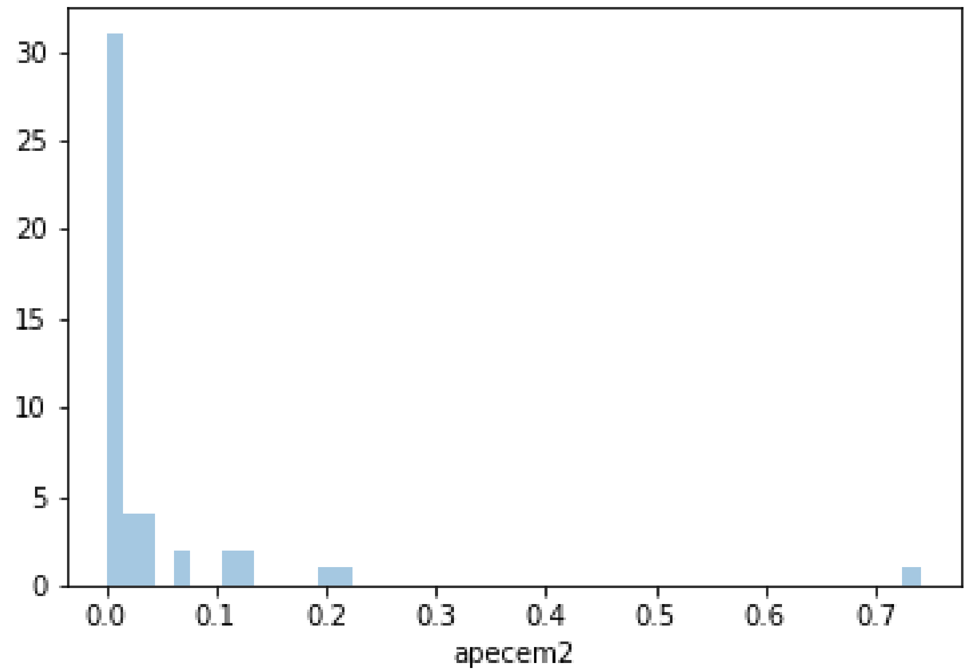
Proposed Solution

Data preparation

Data preparation phase aims to collect and clean data from XMM for spectra generation.

- Generate Chandra source list with 50% power diameters for source removal.
- Inspect and remove observations suffered from heavy soft proton flares.
- Generate spectra for 56 XMM pointings.
- Get temperature and EM for Local hot bubble from ROSAT map.
- Fit spectra in XMM circular pointings.


Absorbed Thermal Component Emission Measure in cm^{-6} pc
Dec
2.65
2.52
2.38
2.24
2.11
1.97
1.83
1.70
Ra
150.57
150.43
150.30
150.16
150.02
149.89
149.75
149.61
149.48
0.7
0.6
0.5
0.4
0.3
0.2
0.1



GH/CGM EM distribution plot.

- EM seems to show clumpiness. Hard to tell from the distribution plot.

Notice

Possible bugs.

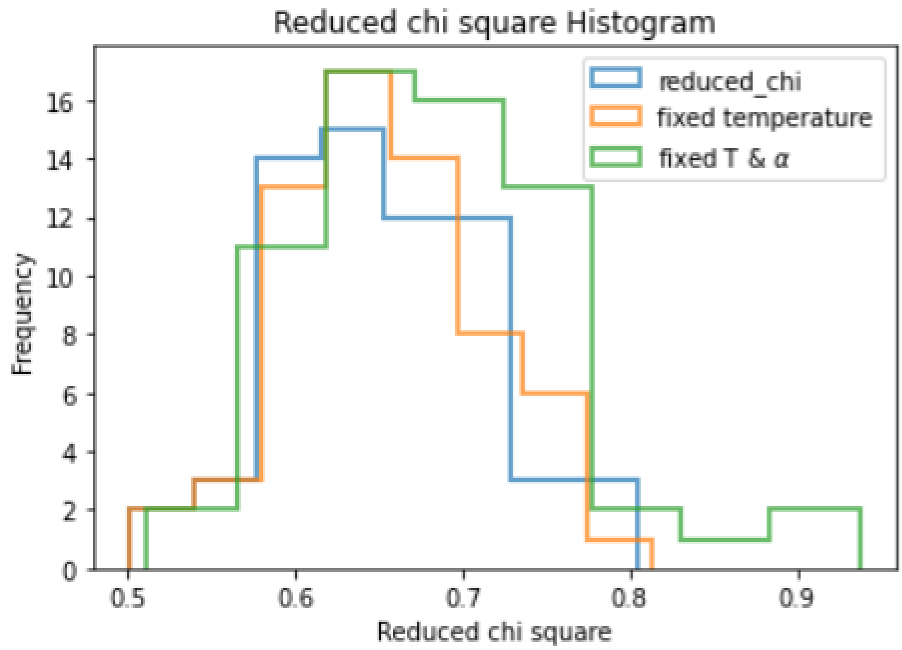
- Unable to generate spectra for observation **0302352401**.
- Observation **0203361601 0203362001 0203362301 0203362501 0302352001 0302352201 0302352301** suffer from heavy proton flares.

Phase 1

Phase 1 aims to generate, co-add spectra, and fit the composite spectrum in each box region (10 arcmin width).

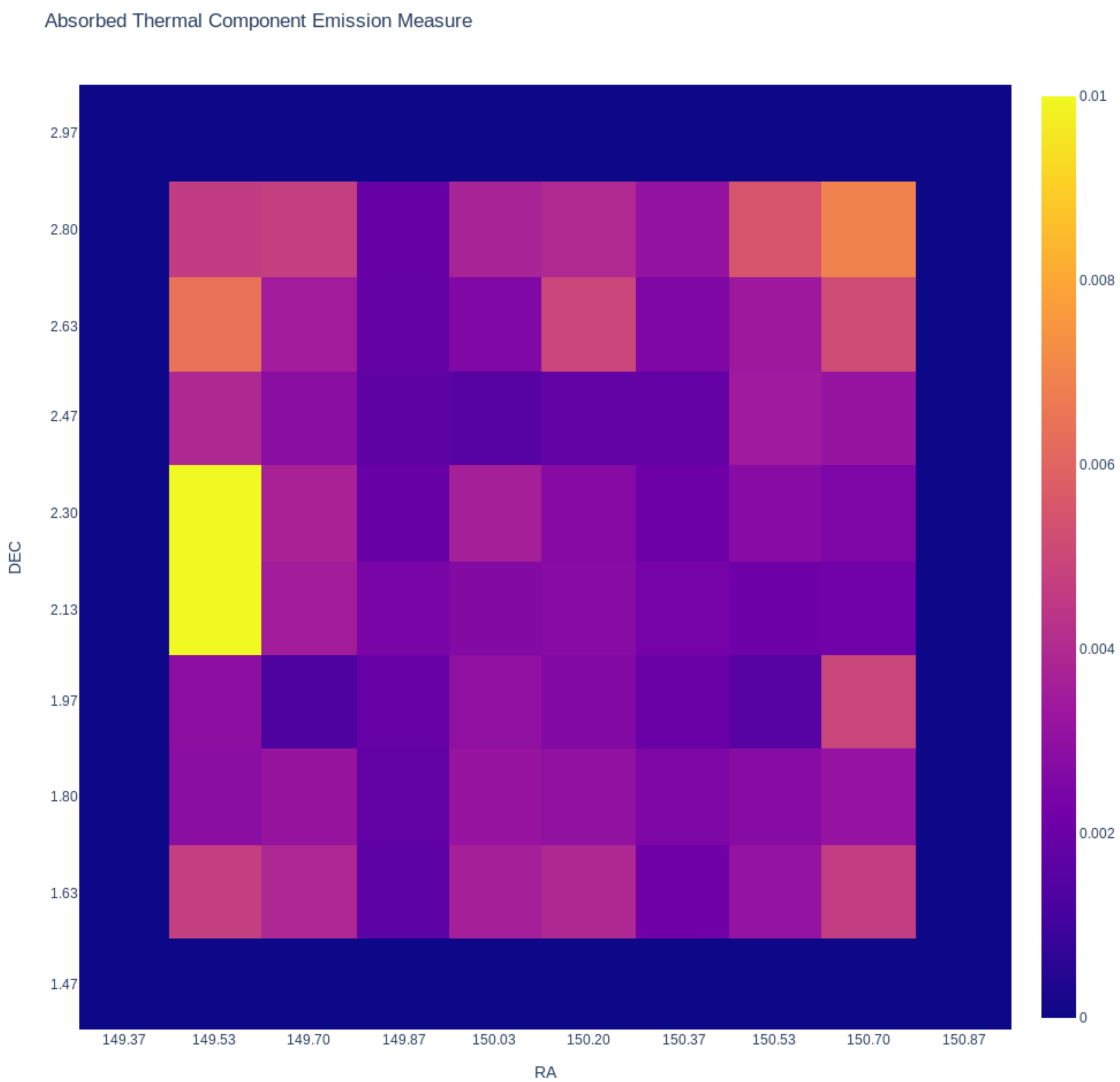
- Use region files in pn_spectra to generate spectra.
- Use mathpha, addarf, addrmf to combine spectra, arfs, rmfs.
- Change the backscale factor of background files according to counts in 10-14 keV .

- Fit spectra in each box region with 1. free parameters for GH and powerlaw, 2. fixed GH temperature and 3. fixed GH temperature and powerlaw index.
- An inspection of the reduced chi square distribution shows that we're overfitting the spectra.

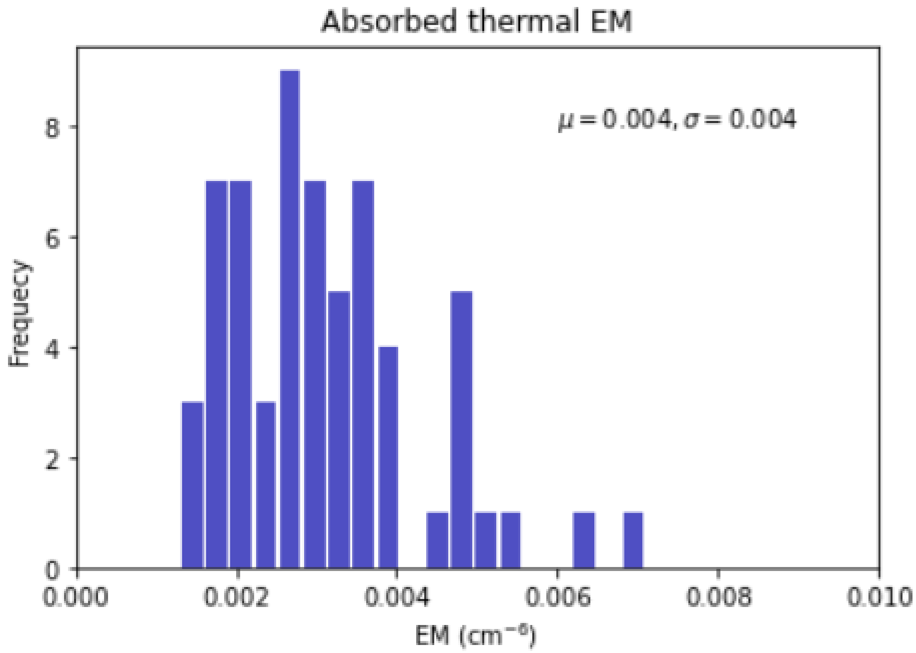


The average of reduce chi squares is 0.65, with a std = 0.058. We're overfitting the spectra, probably due to overestimating the error bars.

- Plot CGM/GH EM heatmap and distribution plot.



The outmost regions are excluded due to low data counts. GH temperatures are free to vary in the fitting process.



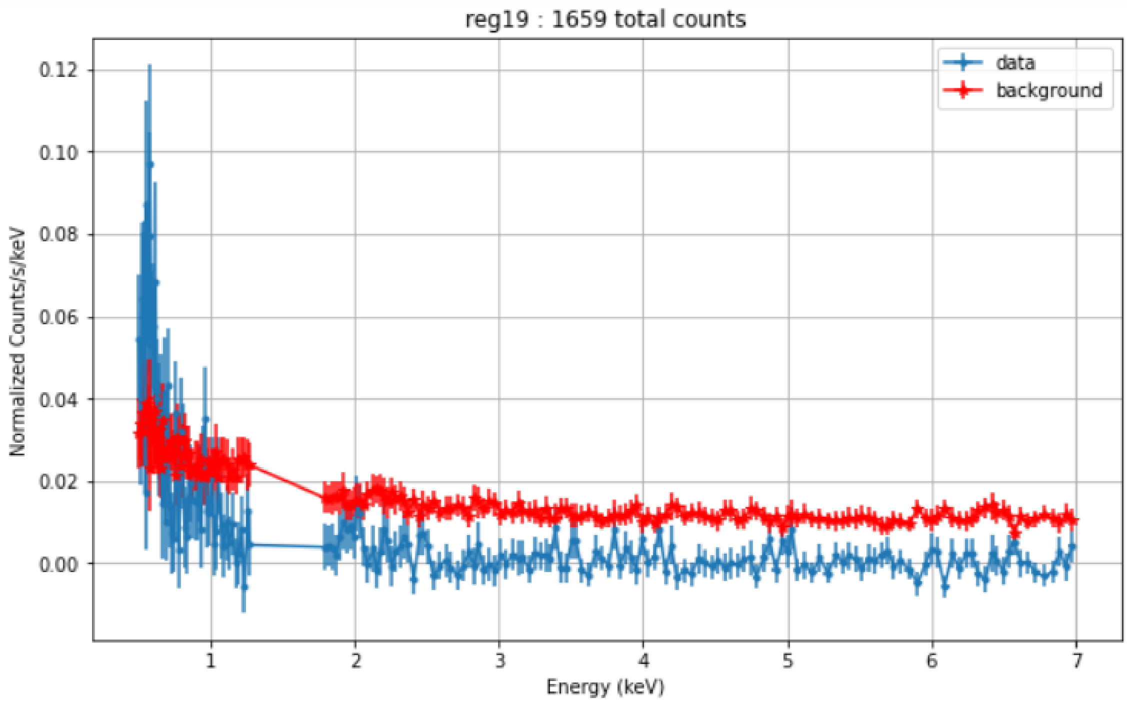
The distribution does not follow a normal distribution.

- Judging from the heatmap, the EM from outer regions seem to be generally higher than central regions. The EM distribution does not follow a strict normal distribution, suggesting a potentially real variation in EM.

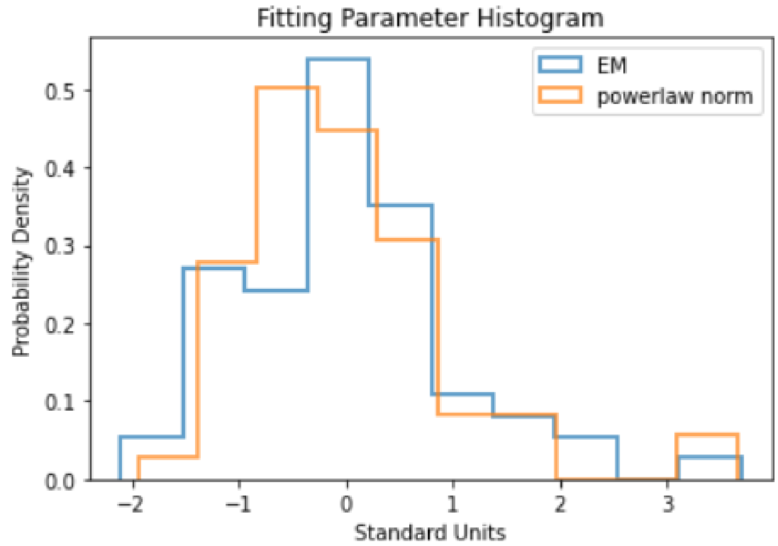
Notice

Possible bugs.

- The reduced chi squares for all fittings are too low, suggesting overfitting. A closer look at the spectra and background spectra suggests this problem can be due to an incorrectly calculated exposure time when using mathpha. Need to double check.



- The individual spectra uses column "count rate", while by setting the parameter 'unit' to 'C' in mathpha, the combined spectra uses column "count" instead. Maybe I need to set the parameter to 'R' and combine spectra in count rate instead, but I need to find out how mathpha adds up count rate.
- For error propagation, mathpha suggests setting properr='yes' and errmeth = 'Gauss', where the errors are calculated using $\sqrt{N}$. Testing suggests doing this does not affect the result significantly.
- Fixing the GH temperature does not improve reduced chi squares, but does make EM distribution more normal-like.



Phase 1 Follow-up Tasks

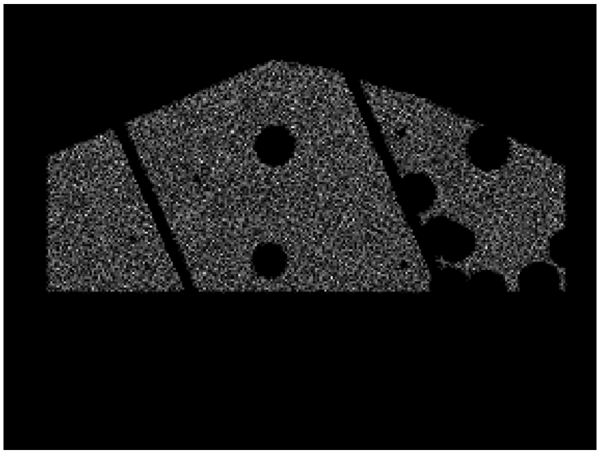
What needs to be done next for phase 1?

- [ ] Change columns in combine spectra to count ratio.
- [x] ~~Check exposure times are calculated correctly.~~
- [x] ~~Check fitting is not biased due to difference in total counts in each region.~~

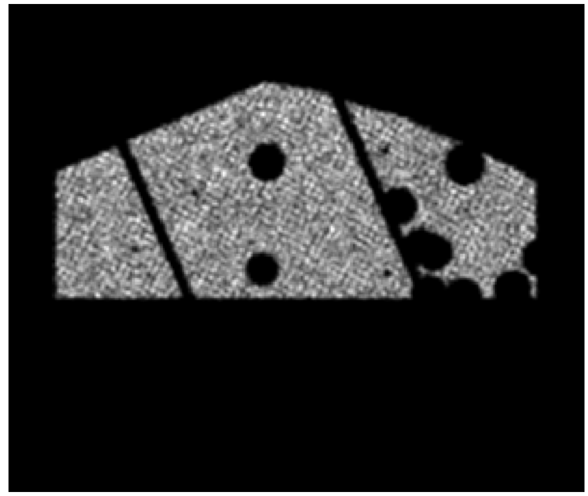
Phase 2

Phase 2 aims to fit the individual spectra from different pointings in the same box region simultaneously, as a workaround to bugs produced in the process of combining spectra. This new fitting strategy allows me to manually inspect and remove bad pointings in each region and can significantly improve reduced chi square statistics.

- The solid angles of different pointings in the same box region are calculated using background images produced by pn_back pnS005-back-im-sky-400-7200.fits. The images are smoothed by a 3×3 pixels Gaussian kernel to fill the pixel gaps in the original image.

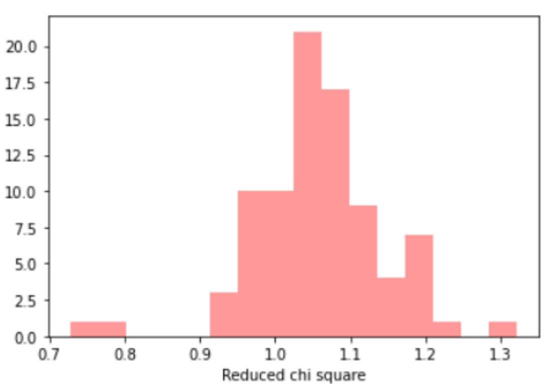


Before smoothing.

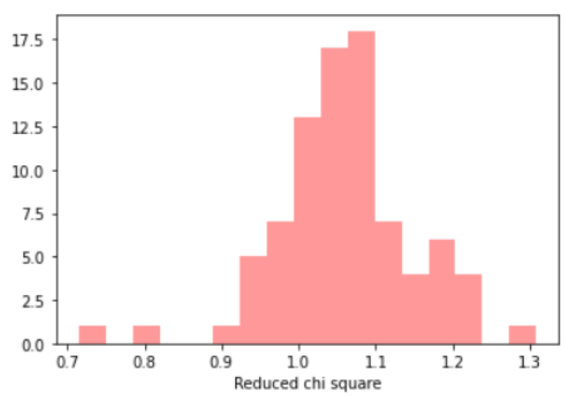


After smoothing.

- For different pointings in the same region, the temperature of LHB are set to the same value, while the normalizations are re-calculated based on their solid angles; the GH/CGM temperature and powerlaw index are linked; the GH/CGM and powerlaw normalizations are also linked, with a scale factor multiplied based on their solid angles.
- The spectra are fitted with the same model again with GH/CGM temperature fixed to mean value (1.48 keV).
- After inspecting and removing bad pointings in each region, the reduced chi squares tend to 1. This is evidence of good fittings.

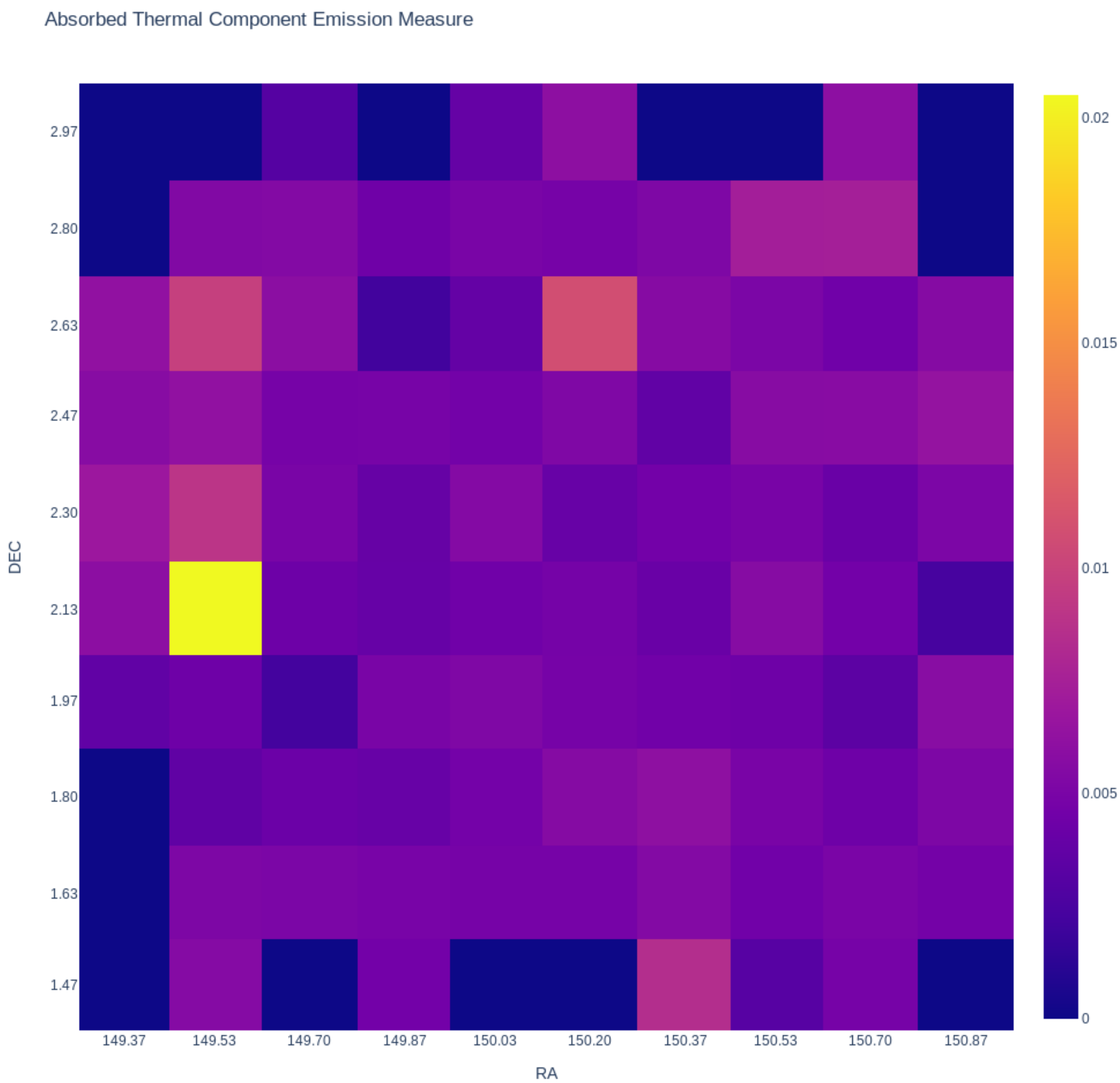


Free GH/CGM Temperature

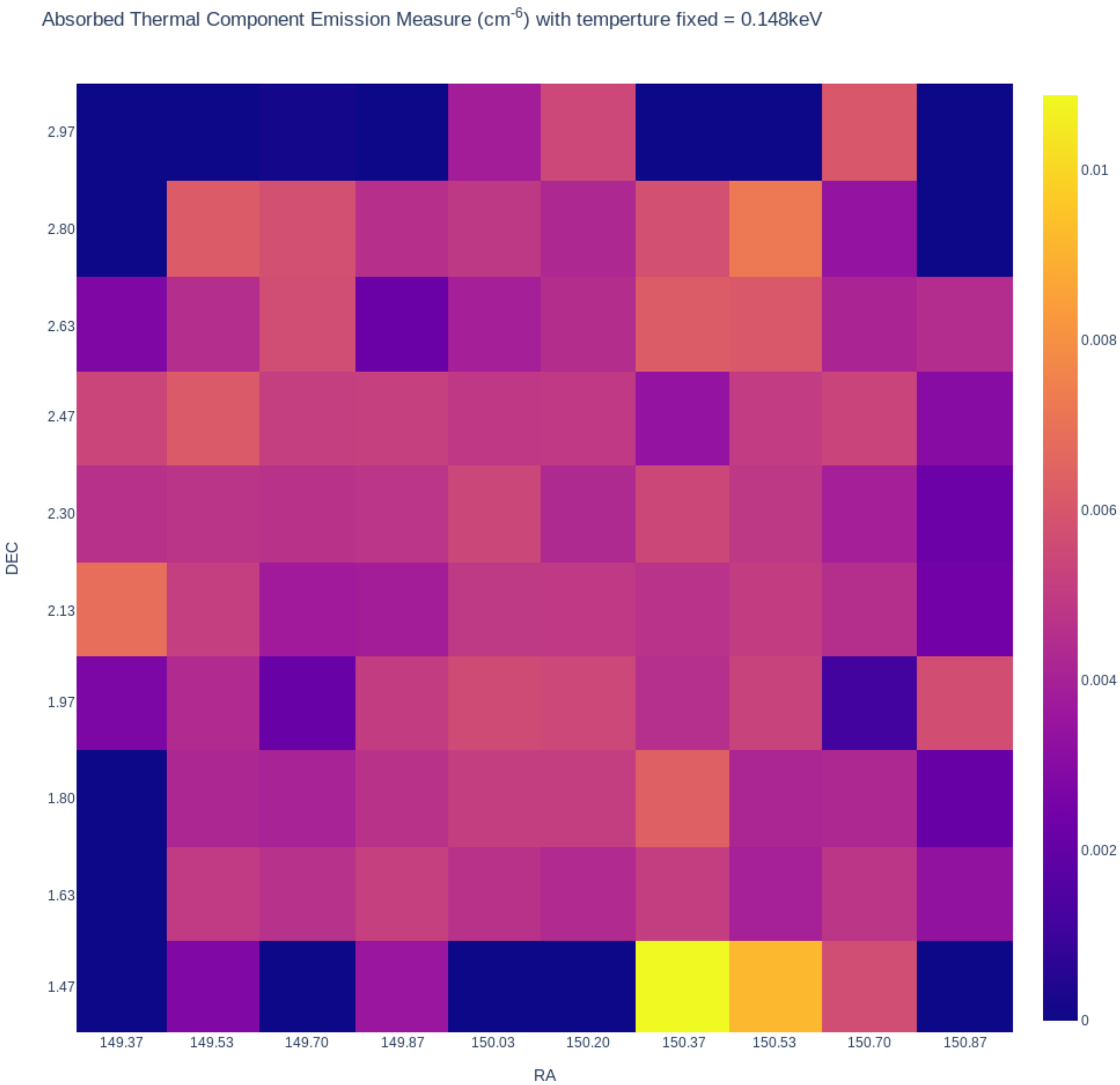


Fixed GH/CGM Temperature

- A heatmap of GH/CGM EM is plotted.

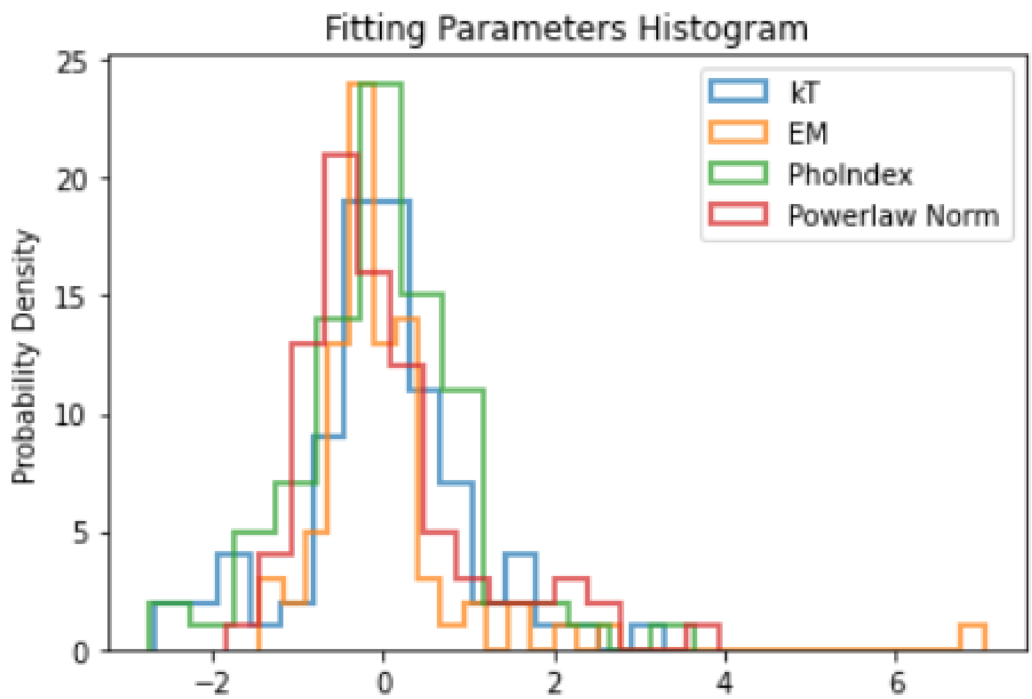


Free GH/CGM Temperature

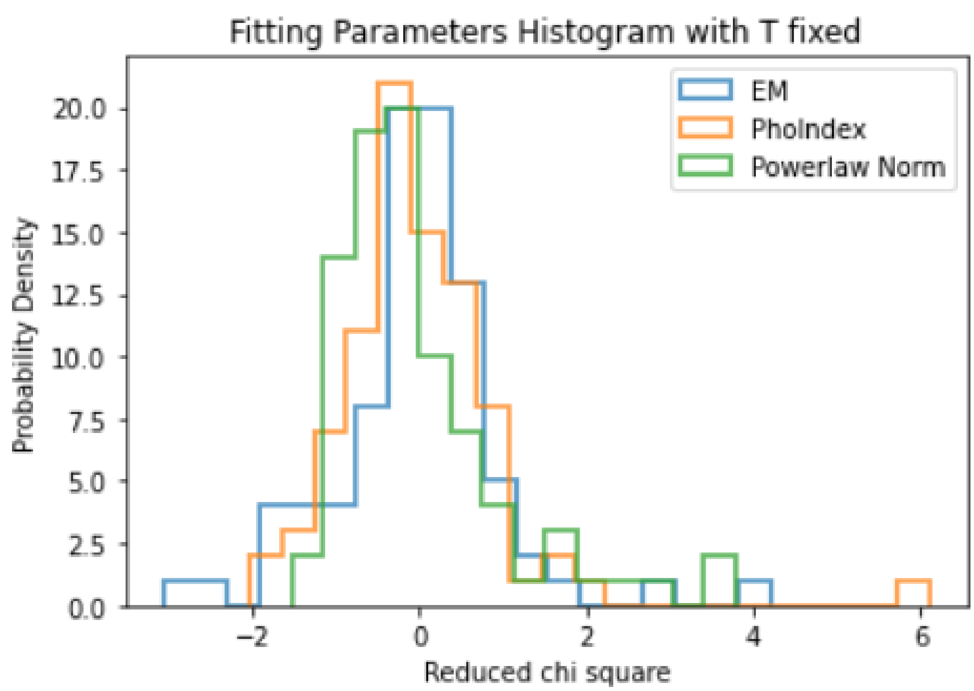


Fixed GH/CGM Temperature

- Histogram of fitting parameters for absorbed thermal component and powerlaw are plotted.



Free GH/CGM Temperature.



Fixed GH/CGM Temperature.

- Table of all fitting parameters

Fitting Parameters with free temperature

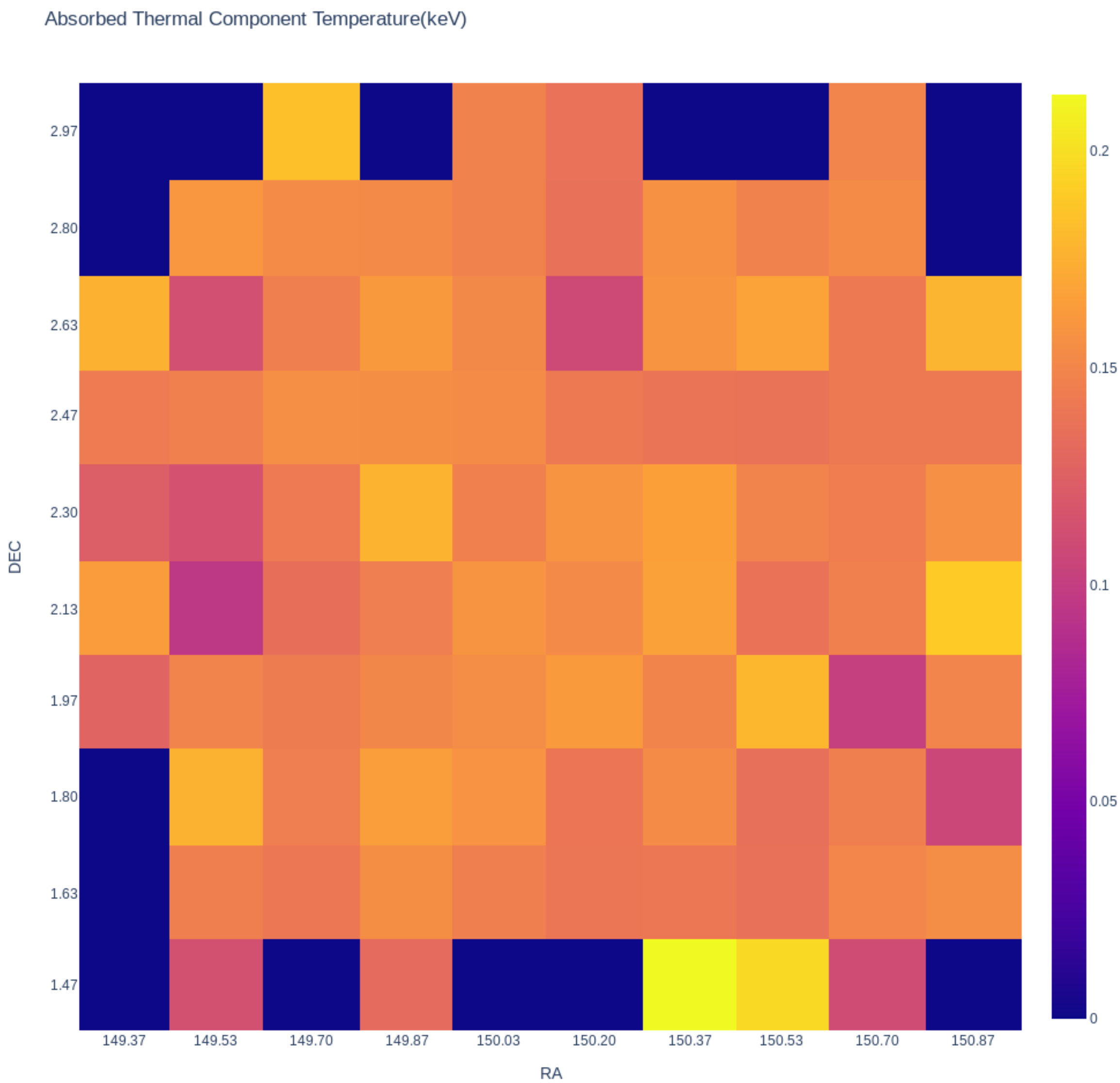
Parameters	Sample Mean	Standard Deviation	Standard Error	combined error
Absorbed Thermal Temperature (keV)	0.15	0.020	0.0022	0.021
Absorbed Thermal EM (cm^-6)	0.0053	0.0022	0.00024	0.0020
Powerlaw Index	2.01	0.52	0.056	0.12
Powerlaw Norm (10^5 photons/keV/cm^2/s)	7.76	3.74	0.41	0.11

Copy of Fitting Parameters with fixed temperature

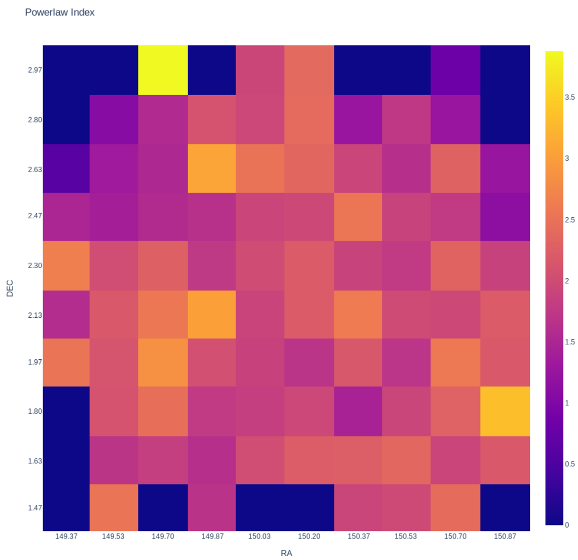
Parameters	Sample Mean	Standard Deviation	Standard Error	combined error
Absorbed Thermal EM (cm^-6)	0.0047	0.0015	0.00016	0.00011
Powerlaw Index	2.04	0.61	0.066	0.045
Powerlaw Norm (10^5 photons/keV/cm^2/s)	7.90	4.01	0.44	0.0079

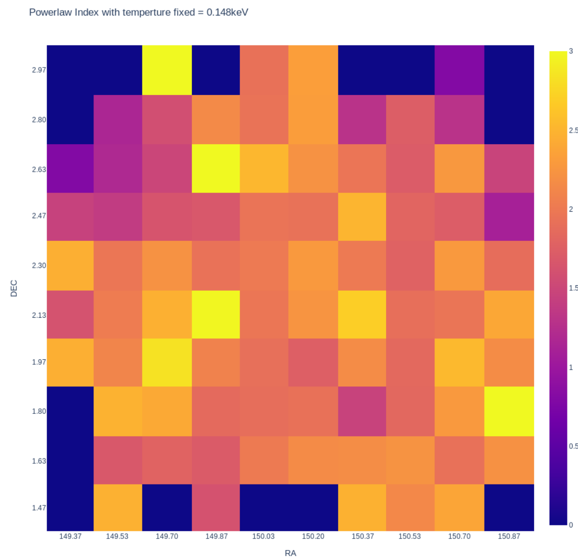
- A visual inspection of the heatmap and histogram suggests that most of the fitting parameters follow the shape of a normal distribution, with most of the values in 3 sigma limit and NO significant spatial variation is present.
- Heatmaps for GH/CGM Temperature, powerlaw index and powerlaw normalization are also plotted.

    ▼ Temperature

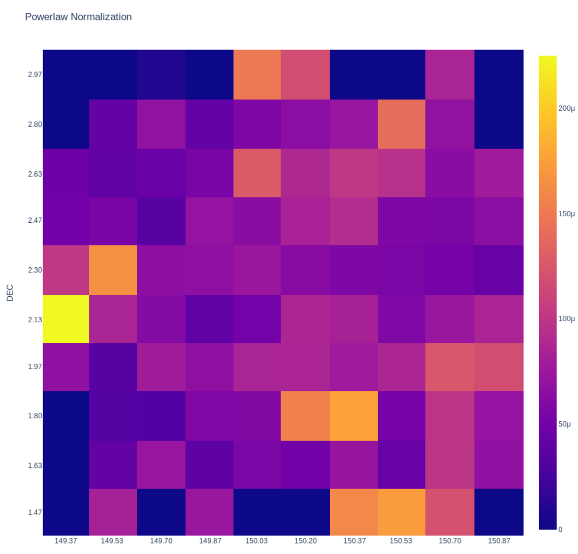


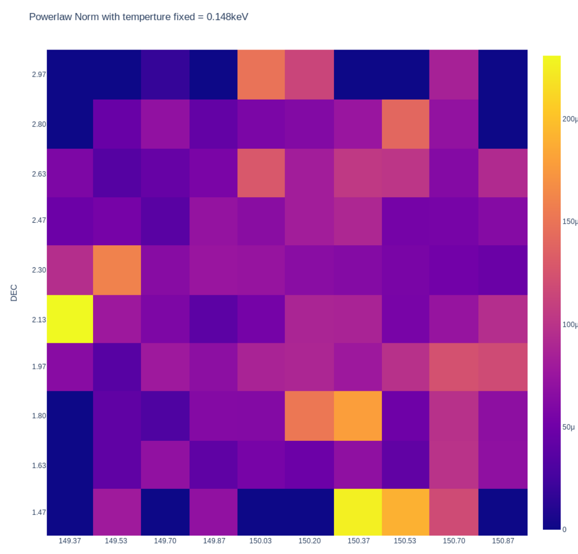
▼ Powerlaw Index





▼ Powerlaw Normalization





Notice

Possible bugs.

- Some potintings have unnormal high count rate. These pointings are labeled as 'bad' and excluded from fittings.

Phase 2 Follow-up Tasks

What needs to be done next for phase 1?

- [ ] Identify outliers.
- [ ] Compare results with updated phase 1.

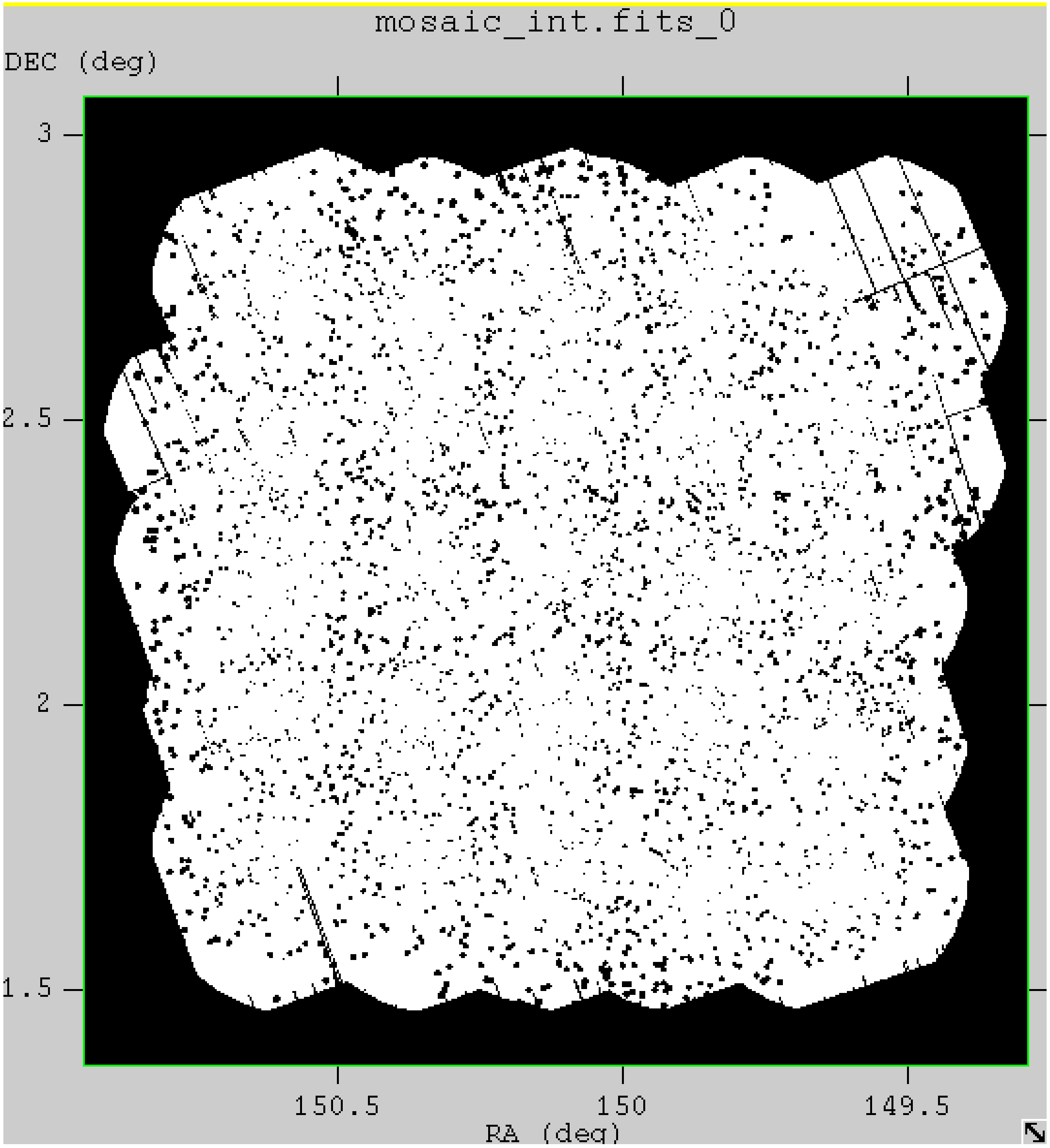
Results

Result from phase 2 suggests that there's no real spatial variation in GH/CGM EM. Everything we see is due to statistical variance.

Phase 3

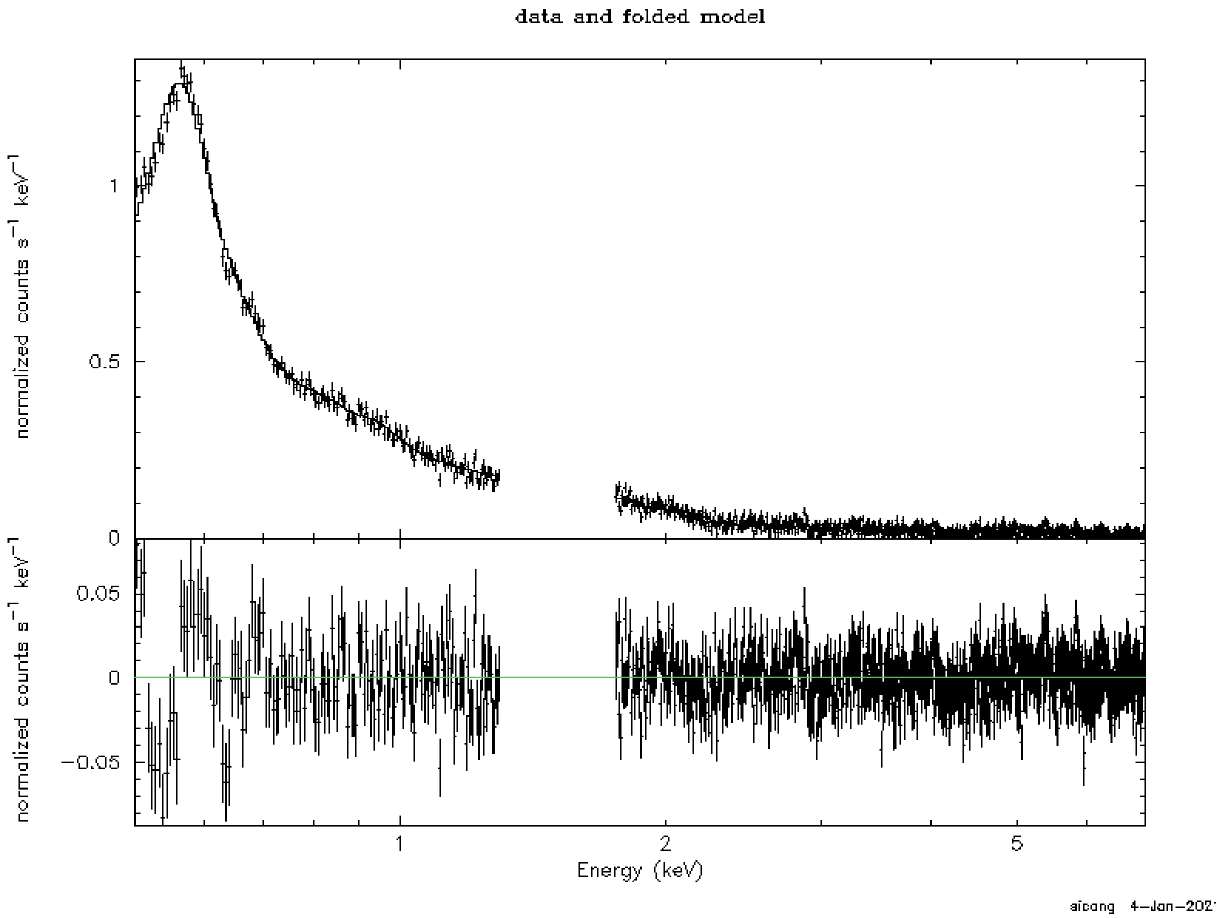
Phase 3 aims to combine spectra from all pointings and fit with an additional Gaussian around 0.9 keV for potential Ne IX emission line.

- The spectra and backgrounds are added using mathpha, while arfs and rmfs are averaged with equal weights. This results in incorrect spectral counts in energy bins. However, this cannot be avoid since the pointings are not aligned but partially overlapped. Thus, the norm for LHB is left free to vary in fitting process so that the model can fit the data.
- Despite the fact that the LHB norm is not frozen so we don't have to calculate it by EM and solid angles, the total solid angle can be obtained from combined cheese map:



Combined cheese map. The total solid angle is 0.00059 sr, the average solid angle is 3.01E-5 sr.

- Fitting with an additional Gaussian results in a reduced Chi-squared = 1.76, while fitting without Gaussian gives Chi-squared = 1.81.



- The Gaussian line is at 0.95 +/- 6.49E-3 keV. The line width is fixed to 0, with Intensity = 0.30 **+/- 0.037 LU**, which is 8.12 sigmas away from 0. Combined with the results from CDFS, the line detection is significant. (LU= photons/s/cm2/sr)

Table 4. Fitting Parameters of Ne IX Component

Parameter	DXB90	Galaxy90	DXB50	Galaxy50
Line Energy(keV)	0.93 ± 0.008	0.95 ± 0.014	0.93 ± 0.0084	0.92 ± 0.031
Normalization($10^{-6} \times photonscm^2s^{-1}$)	1.12 ± 0.21	0.68 ± 0.18	1.33 ± 0.26	0.14 ± 0.08
Flux(LU)	0.17 ± 0.032	0.089 ± 0.024	0.13 ± 0.025	0.088 ± 0.05

Notice

- In order to improve quality of the spectrum, observation **0203361601 0203362001 0203362301 0203362501 0302352001 0302352201 0302352301 0302352401** are removed before combining**.**