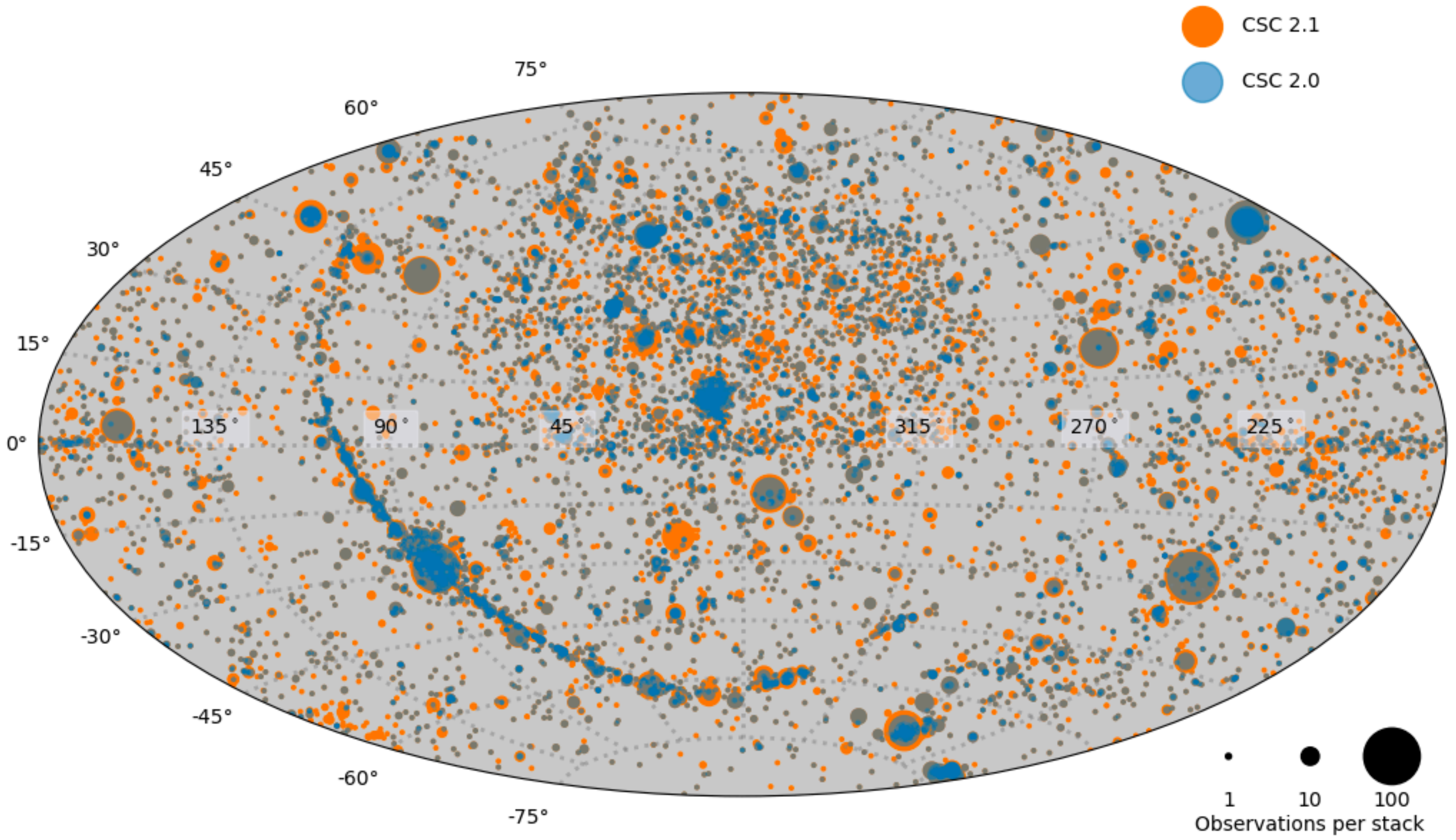
The Chandra Source Catalog Version 2.1

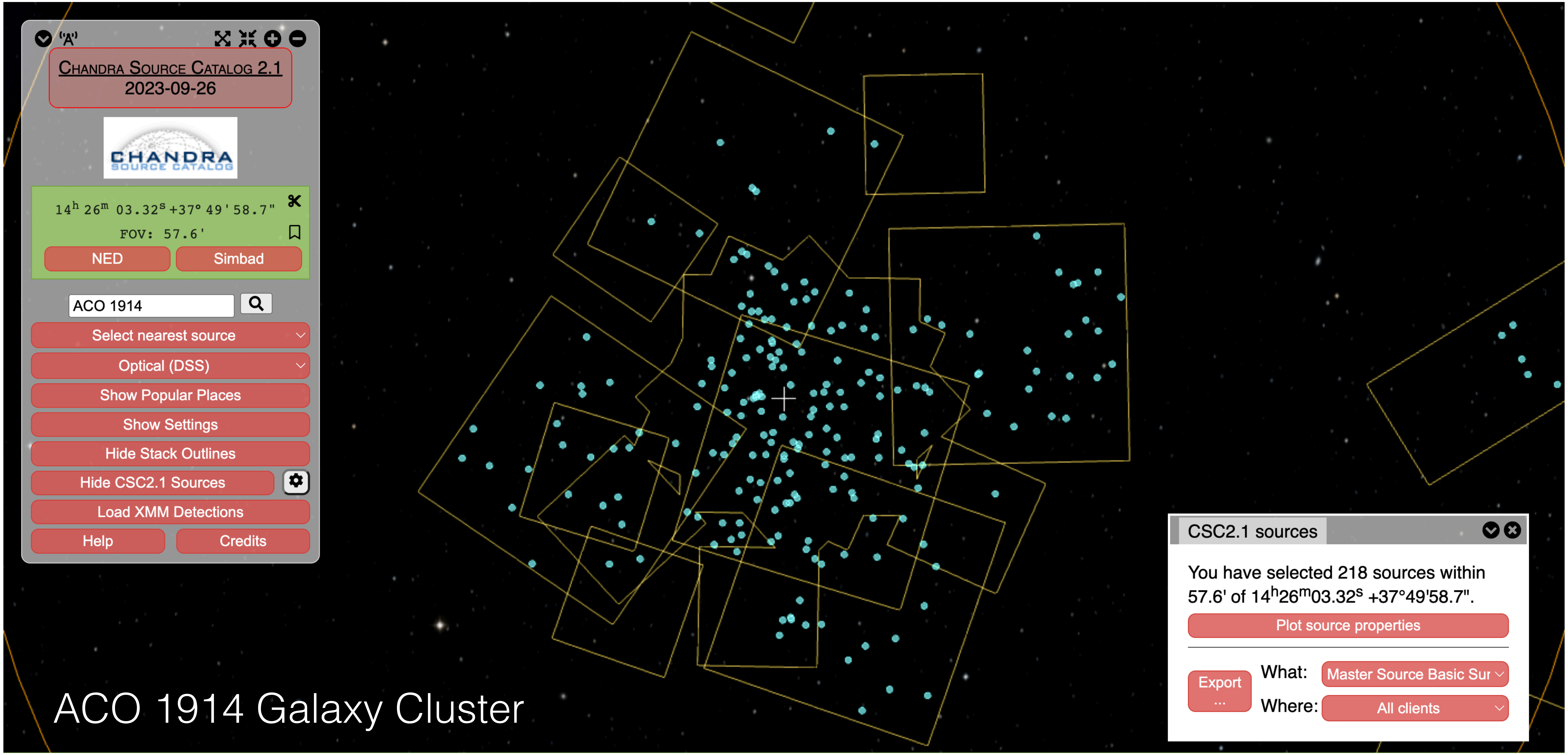
Rafael Martinez-Galarza & the CSC team



- CSC2.1 production data are available via several interfaces. 93% of all stacks have completed processing.
- Only three regions of the sky in processing (Sgr A* by far the largest)
- Science verification being finalized
- Full release expected in January 2024. Will include DOIs

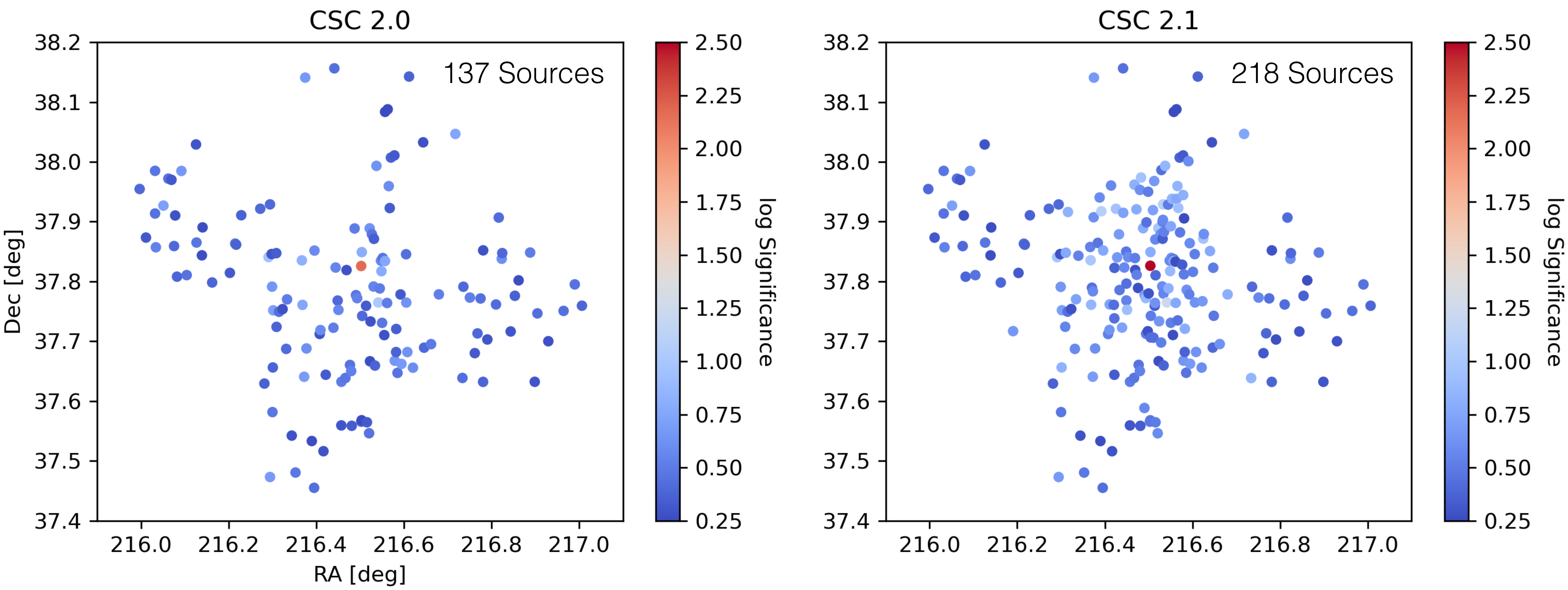
- ACIS + HRC observations that were publicly available on 12/31/2021. Early obs, HRC-S and gratings not included.
- ~390,000 individual sources on the sky, which is over 70k more than in CSC 2.0.
- >40% increase in sky coverage with respect to CSC 2.0.
- Astrometry now tied to the Gaia reference frame.
- Improved aperture photometry algorithm.
- Improved source position fitting algorithm.
- Several interfaces: CSCview, World Wide Telescope, VO-accesible via pyvo, Jupyter notebooks, CfA Nexus.

New stacks in previously covered areas



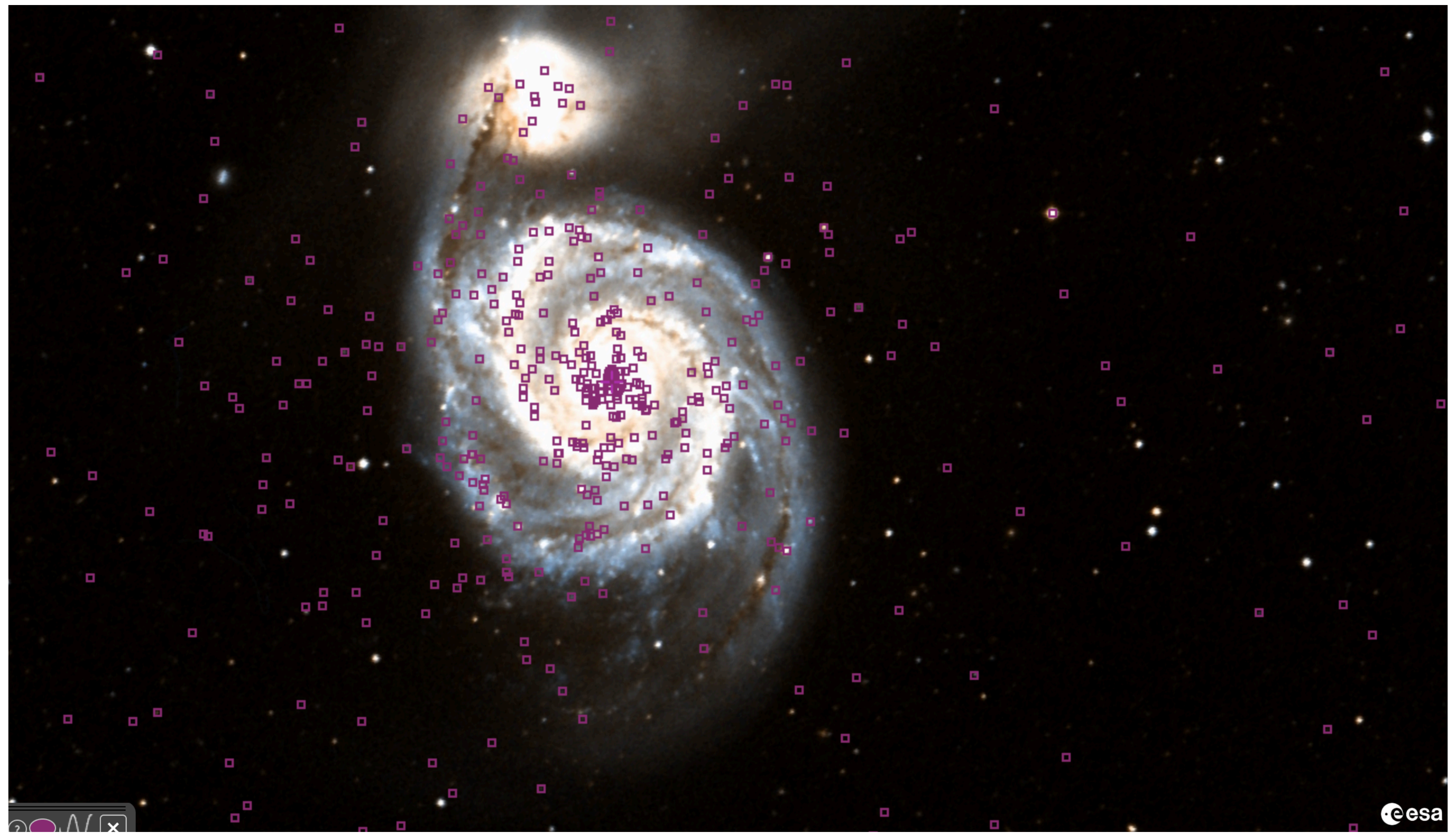
https://cxc.cfa.harvard.edu/csc/wwt21.html#

Improved detectability in CSC 2.1

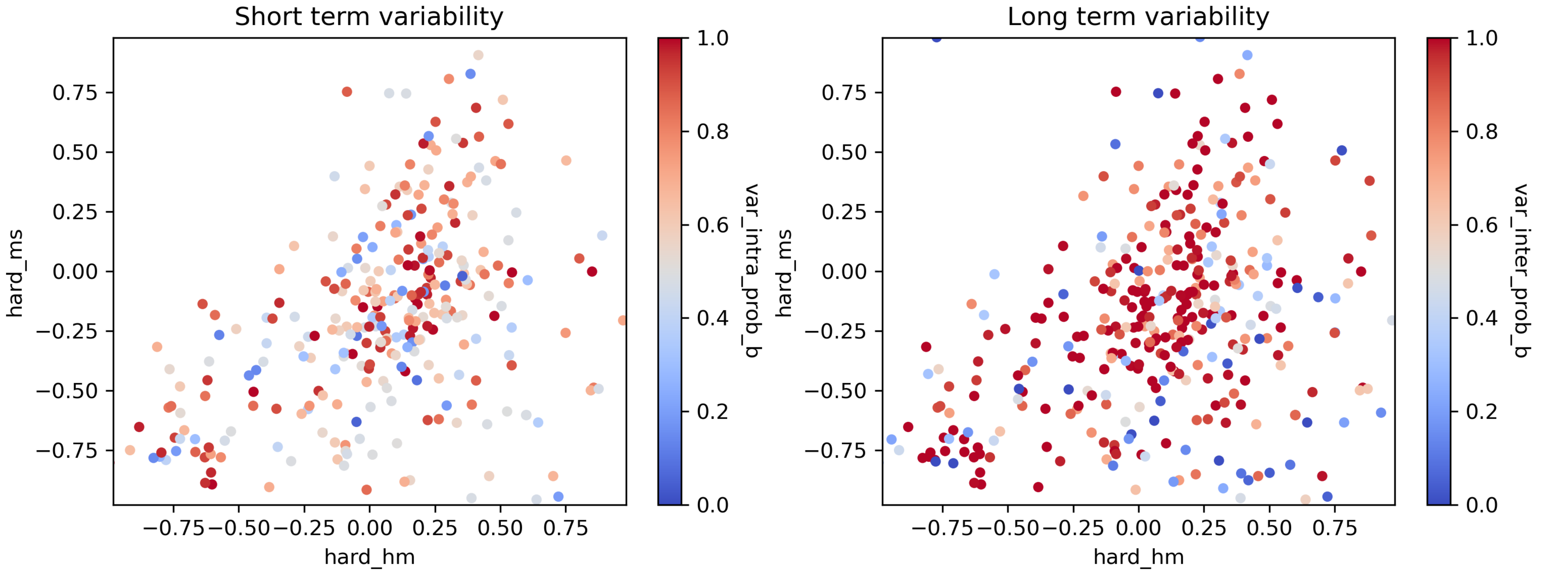


ACO 1914 Galaxy Cluster

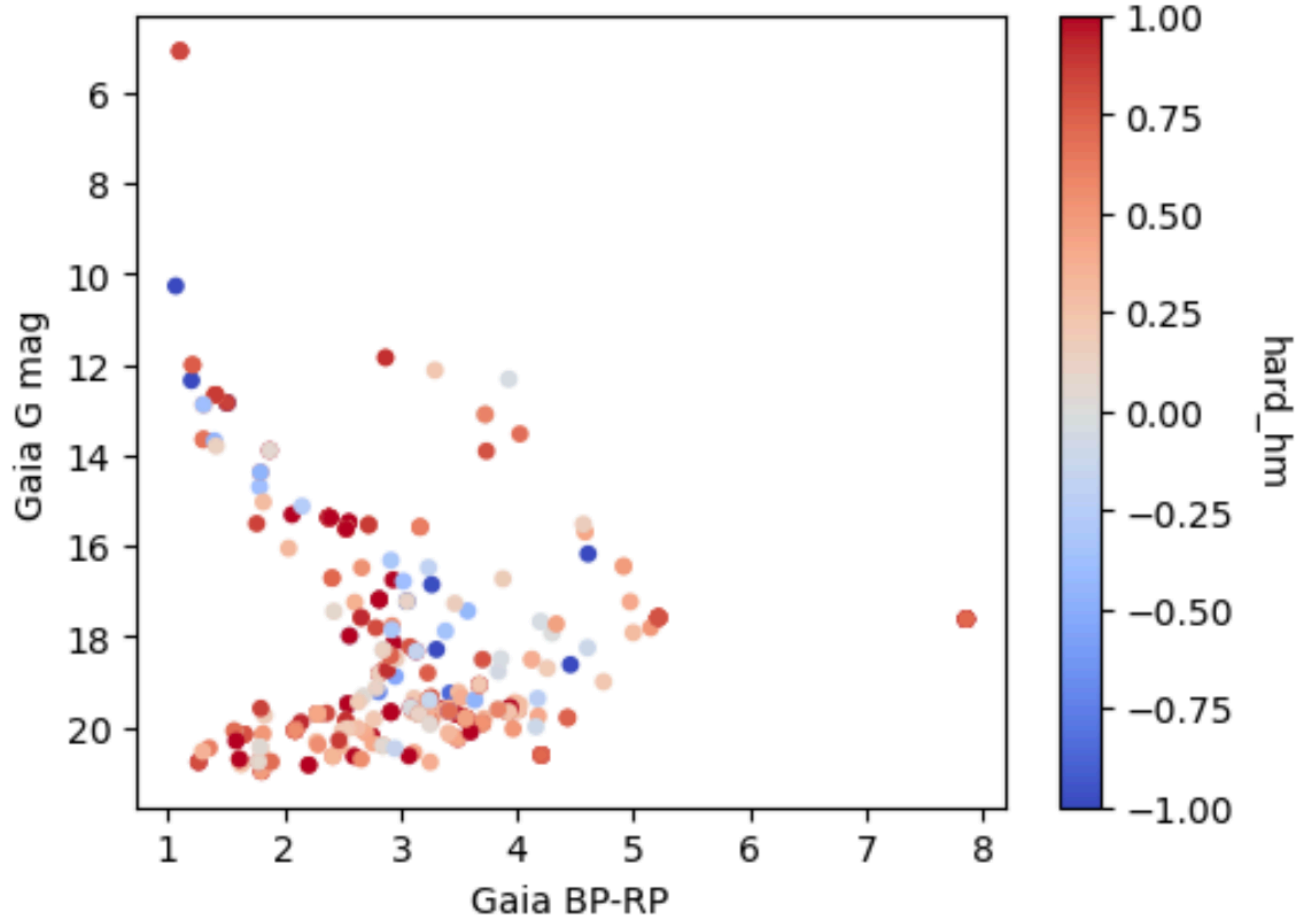
CSC 2.1 Source Properties



- Source properties computed at the per-observation detection, stacked observation detection, and master source level.
- These include astrometric, photometric, spectral, and time variability properties.
- Uniform computation of properties for hundreds of thousands of sources enable data-driven and data mining approaches.

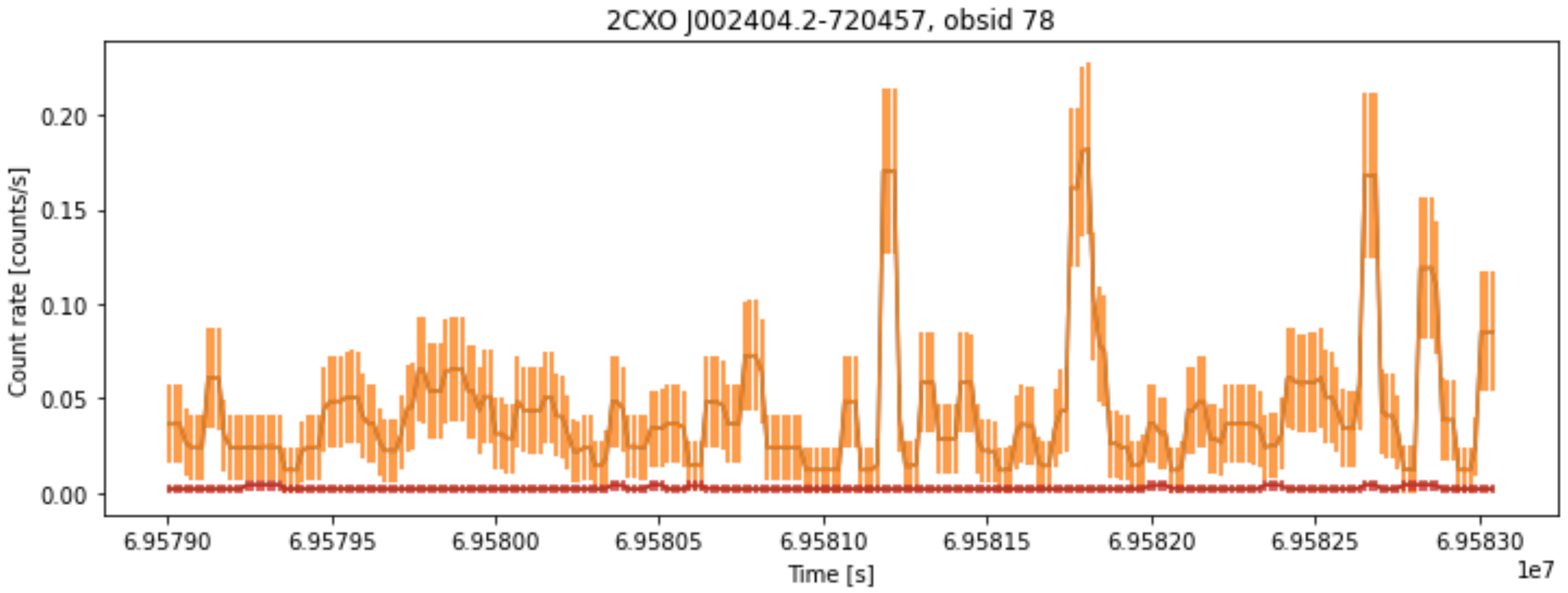


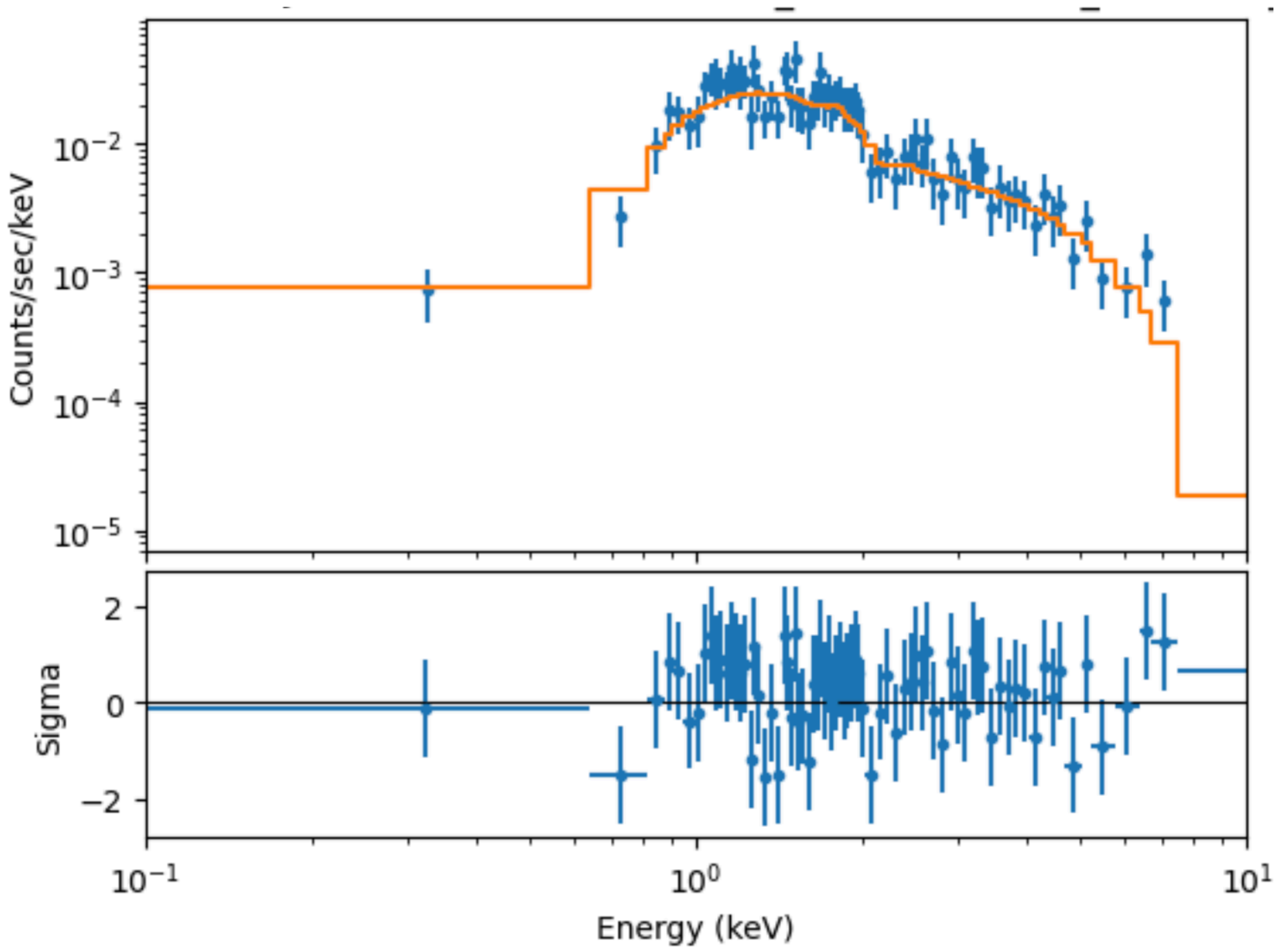
X-ray variability and spectral hardness for M51 sources



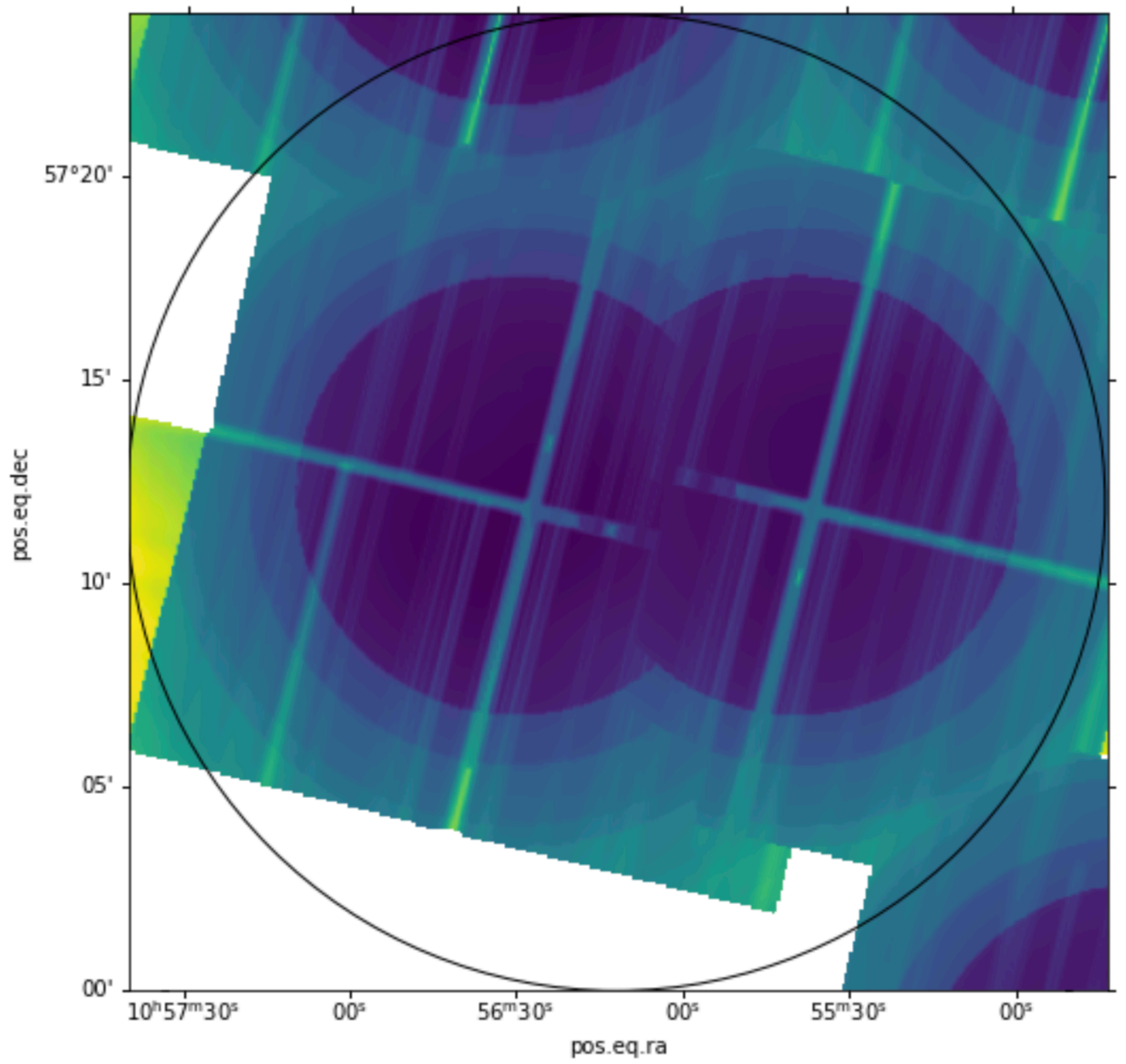
HR diagram of W40 sources

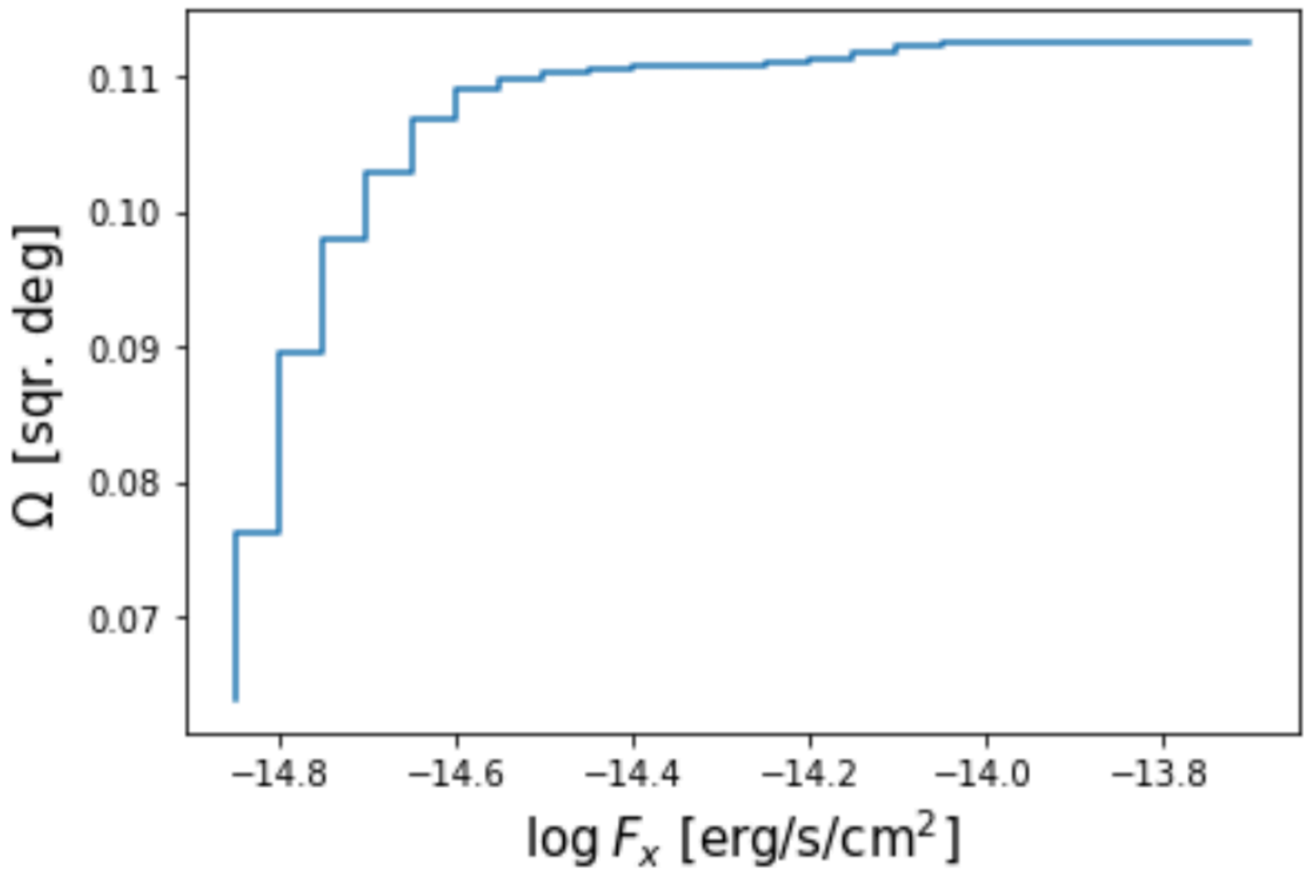
CSC 2.1 Data Products





- Science-ready data products are provided and can be accessed via several interfaces, including Jupyter notebooks and CSCview.
- 45 types of data products, totaling ~22 TB compressed for CSC 2.1
- They include spectra, light curves, Bayesian blocks, background and exposure maps, PSFs, etc.
- Sensitivity maps, for example, allow for the determination of the cumulative sky coverage in flux-limited surveys, as shown below for the Lockman Hole survey.





See https://github.com/juramaga/CSC2_tutorials/

CSC 2.1 processing and release schedule

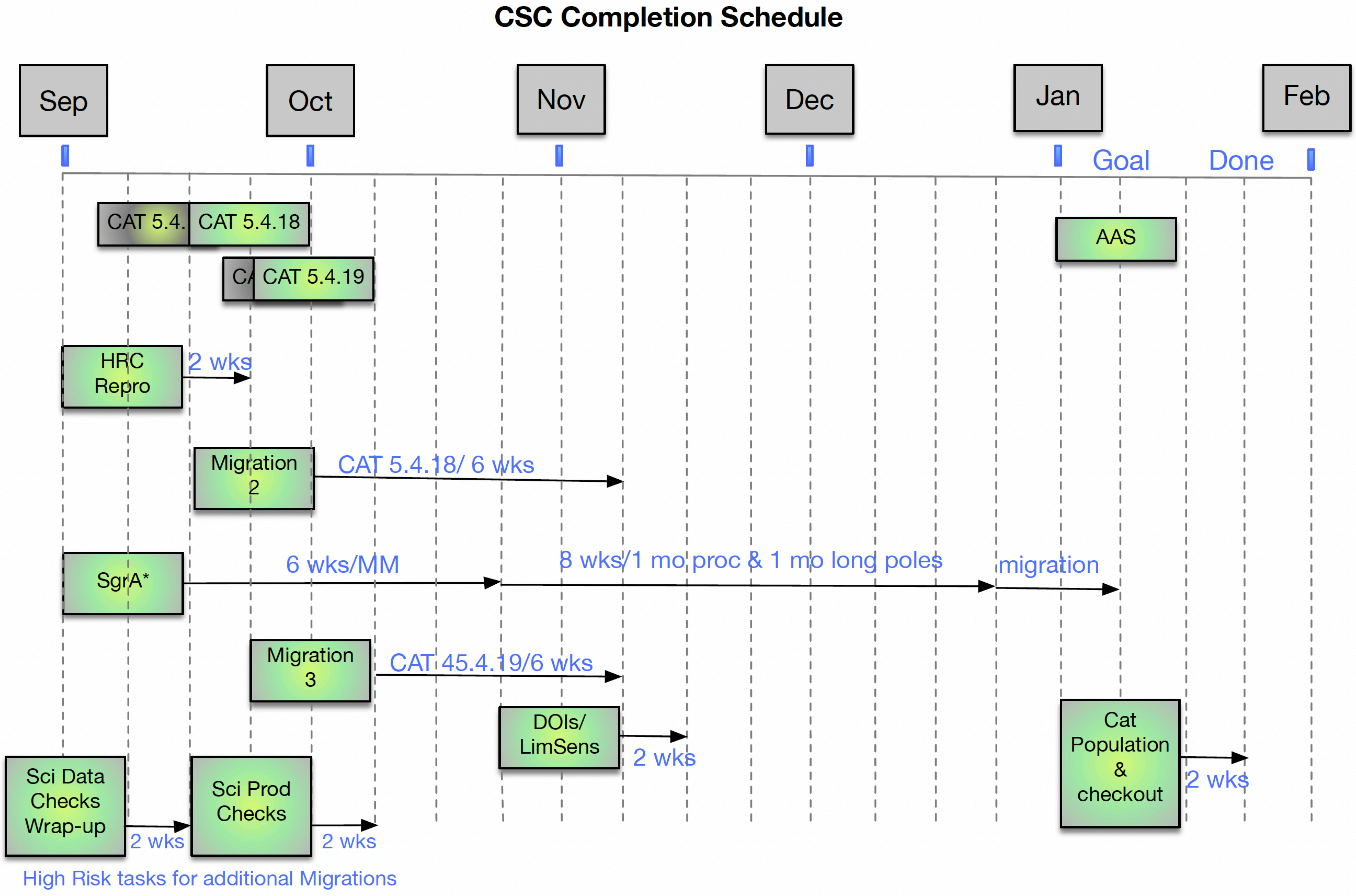
- We have processed 93% of all stacks, and 92% of all observations that will be included in the final release. Processed sources are available to CSC users, on a rolling basis.

- We have completed science checks of the CSC 2.1 properties to make sure that there is statistical consistency with CSC 2.0. Some of these checks lead to further minor processing (migrations, HRC repro).

- Sgr A*, M82 and EO102 are the field still processing. The schedule is mostly driven by the first, which is compute intensive, and requires significant software and science oversight due to a large number of sources.

- SgrA* has passed manual quality assurance (QA) to check for correct detection positions and overall astrometry. We are now processing Master Match in QA. We estimate that computation of source properties will take 8 weeks to complete (worst case scenario).

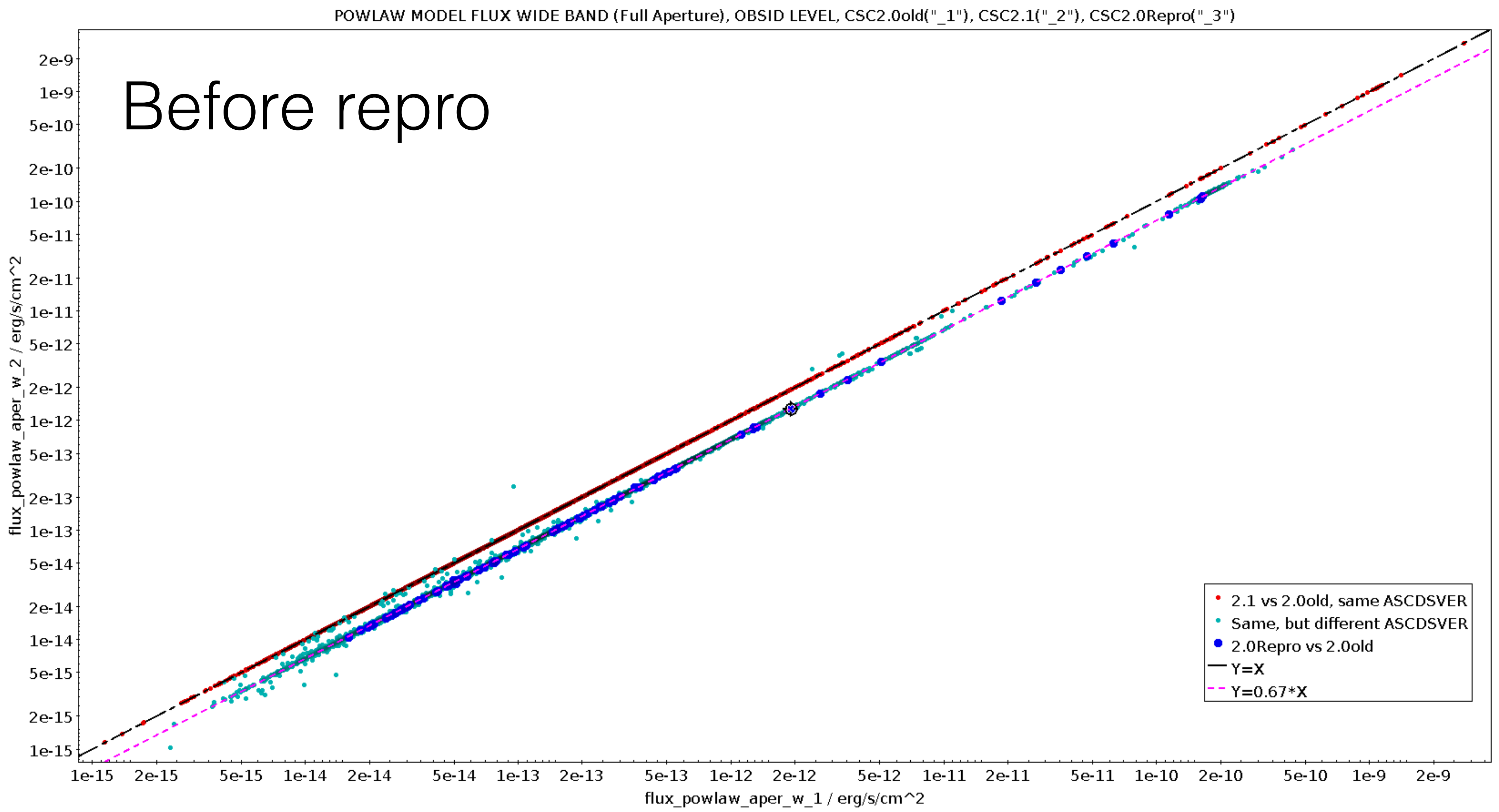
CSC 2.1 processing and release schedule

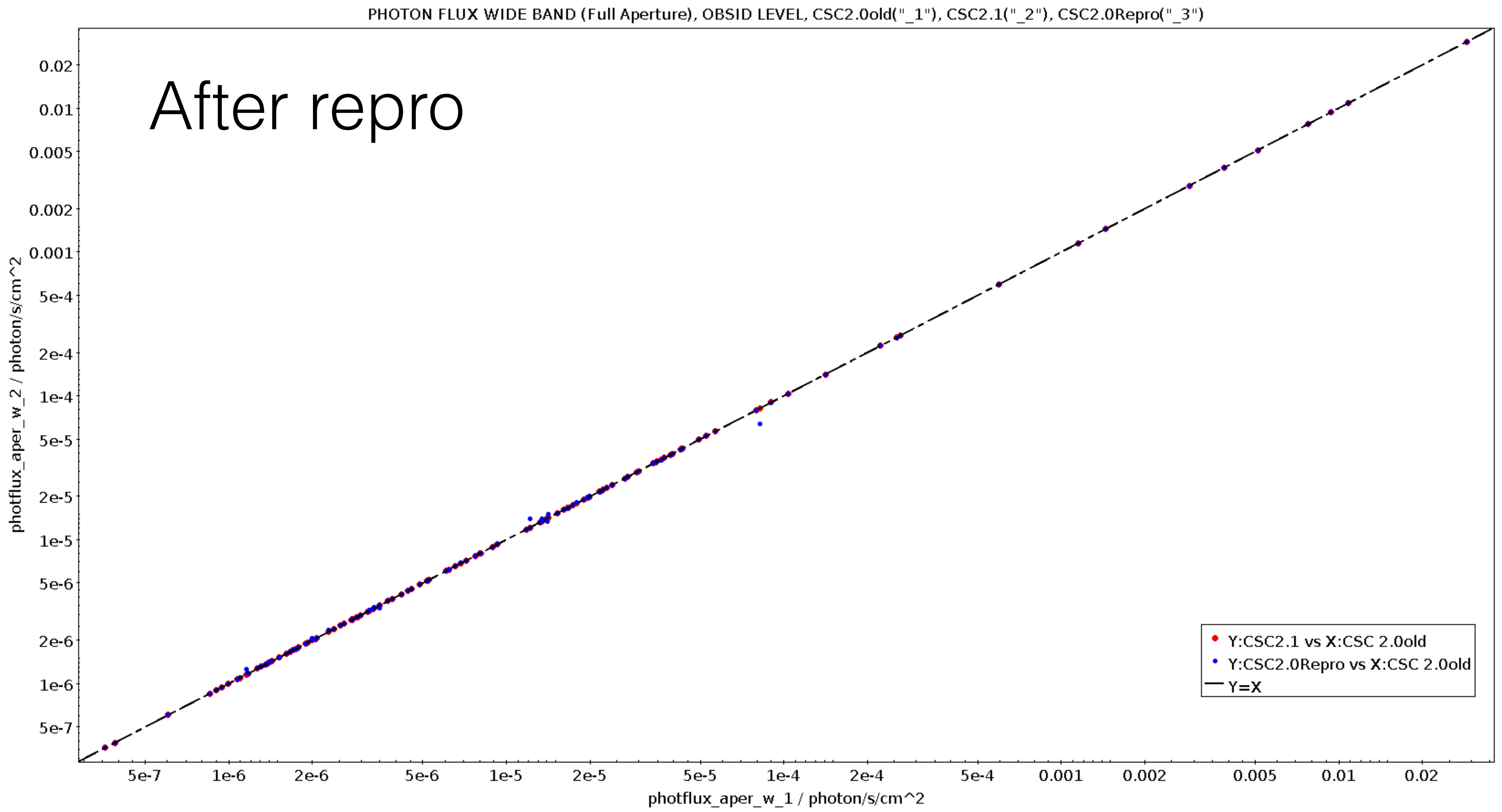


Why are we delayed?

- Level of uncertainty was high when we estimated the completion date at the previous CUC meeting.
  - **Manual Quality Assurance (QA):**
    - Initial estimate for the amount of QAs was based on CSC 2.0 QA numbers + 7 years of data.
    - However, data taken after 2014 has different trends. For example, the amount of large programs that target regions with extended emission (e.g. galaxy clusters), has significantly increased, requiring additional QA.
  - **Computing performance:**
    - We estimated compute cluster efficiency to be ~50% during the QA intensive phase. We were at ~25% due to the manual work in QA.
  - **Science validation/checks:**
    - We have validated the statistical consistency of the catalog properties against the previous versions, but a few issues were identified that required migration/reprocessing

HRC reprocessing



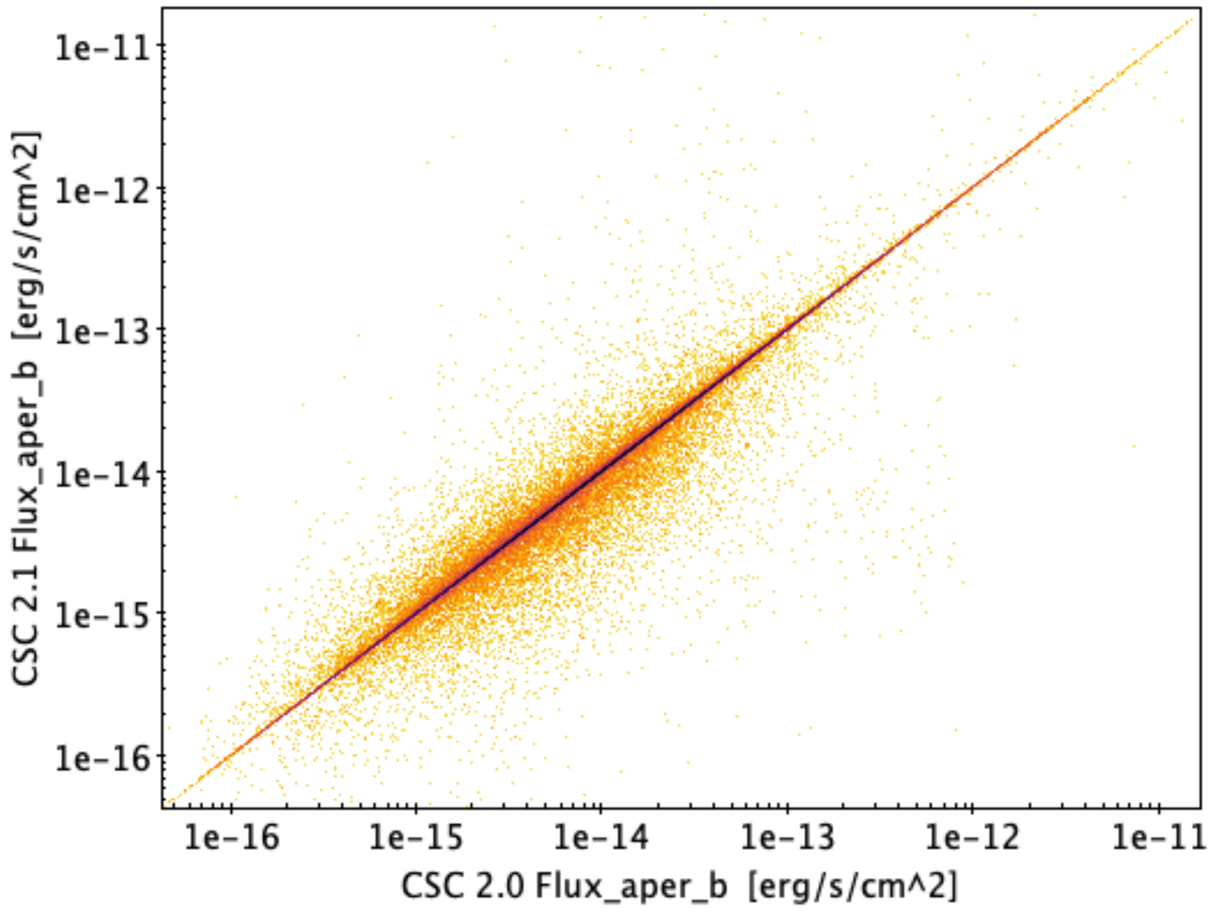


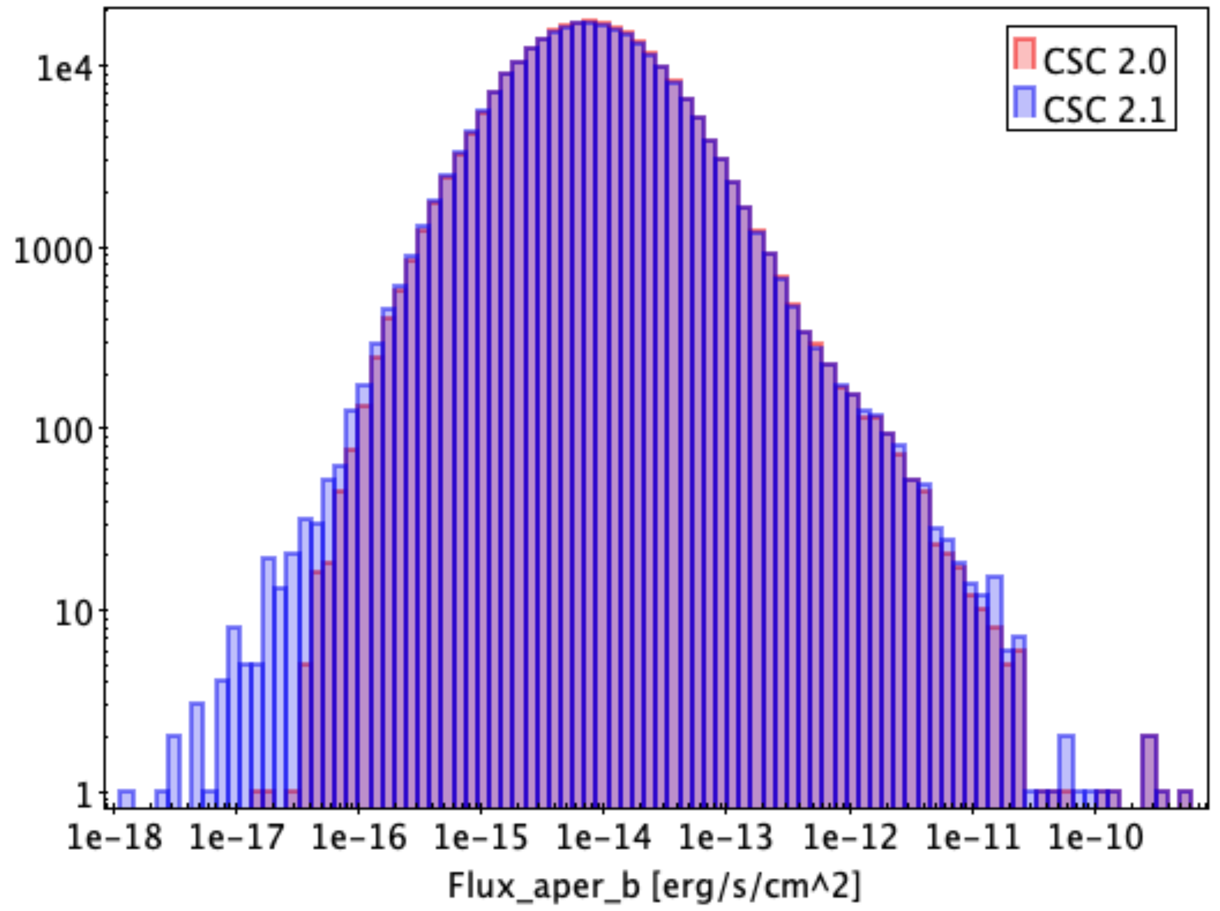
- Science checks revealed that the HRC modelflux values had a systematic shift of ~30%.
- Aperture photometry also showed a systematic shift, but much less severe (5% level, within error bar for a large fraction of detections).
- The issue was due to a faulty CalDB HRC RMF file used for CSC 2.0 processing.
- This faulty RMF was already corrected in the new CALDB that was used to process CSC 2.1 data.
- Thus, a decision was made to repro all unchanged CSC 2.0 HRC detections.

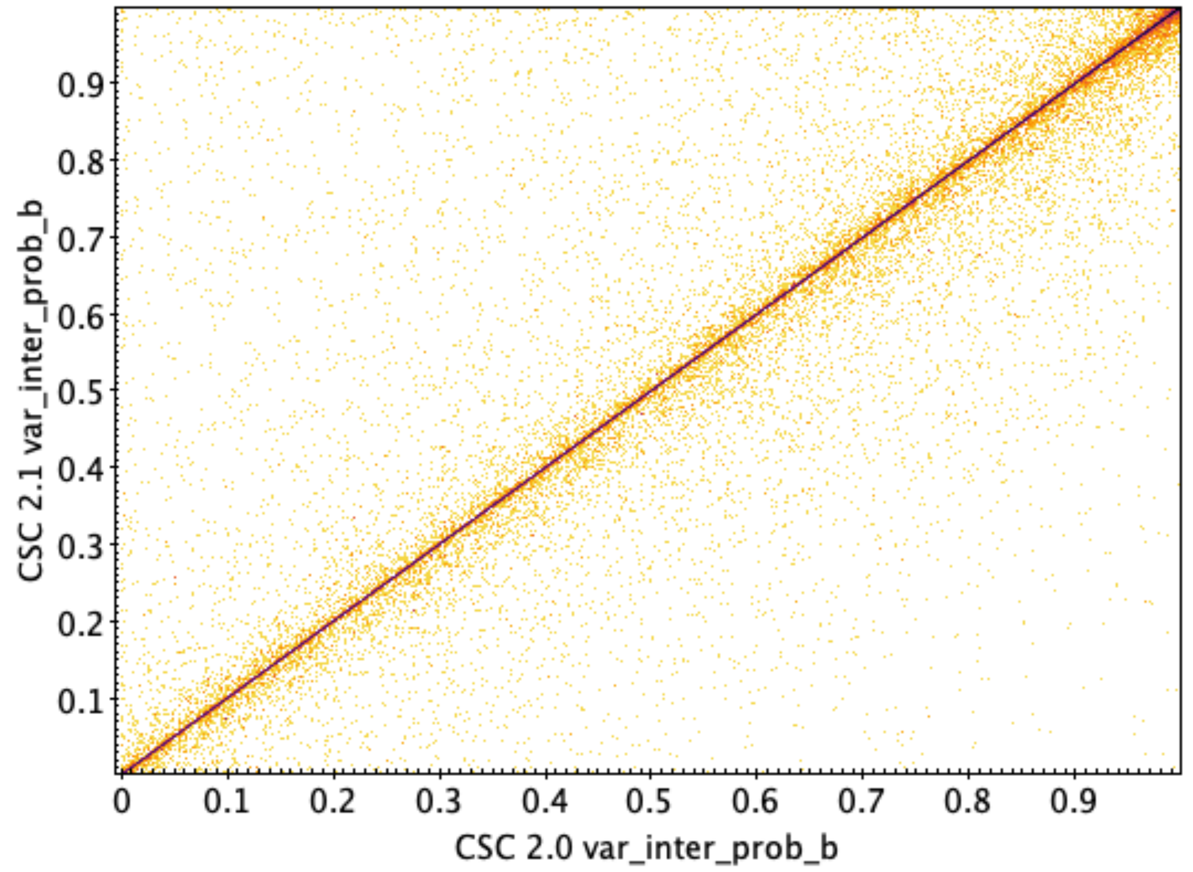
Confidence in the timeline

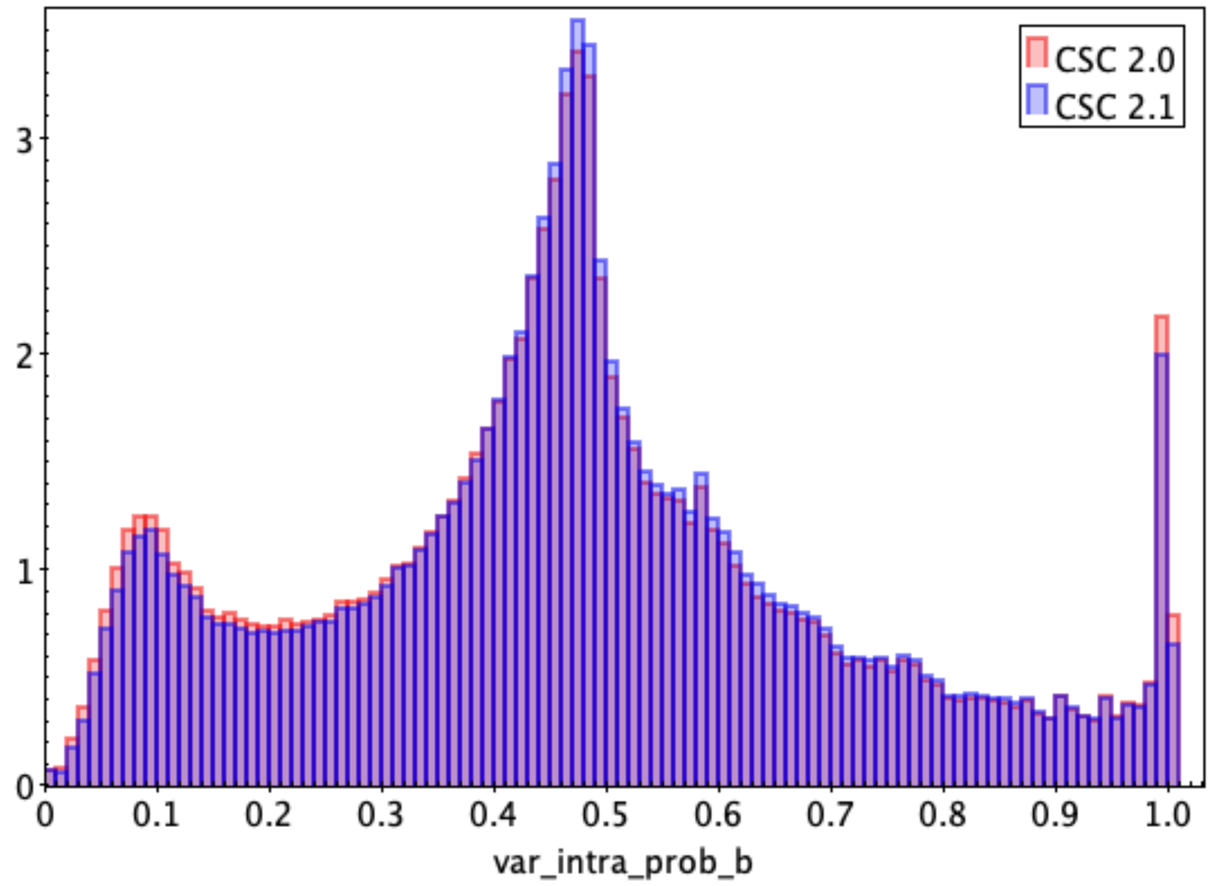
- The two main sources of uncertainty in the schedule have now been completed:
  - **Manual Quality Assurance (QA):** we have completed the QAs for source validation, source position, astrometry, and source matching.
  - **Science validation/checks:** We have validated the statistical consistency of the catalog properties against the previous versions, and resolved issues that we found during validation.
- The timeline is currently mainly driven by the compute-intense derivation of source properties in SgrA*, but a several other things that happen in parallel (e.g. migrations) are also in progress.

CSC 2.1 Science Checks



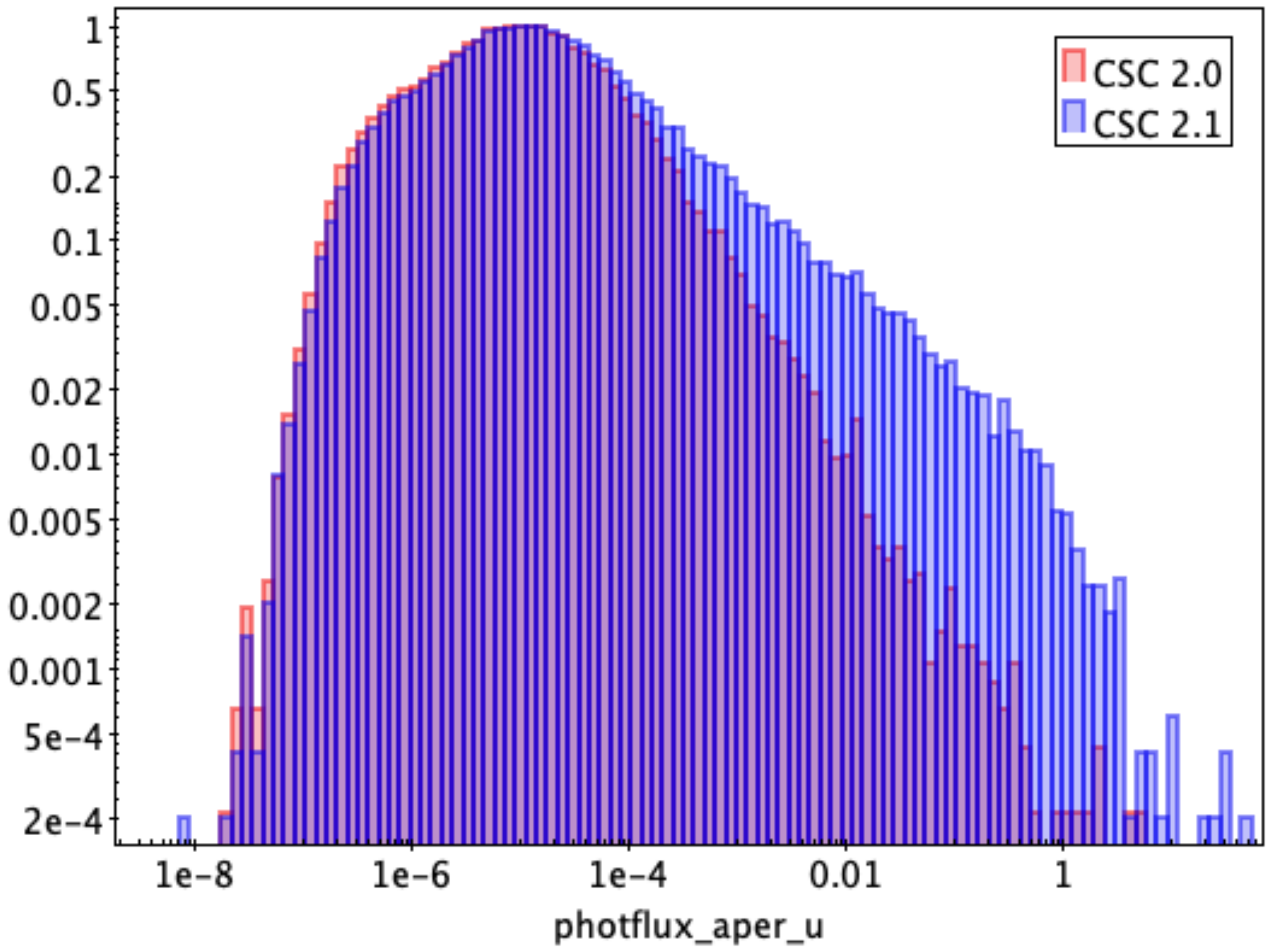




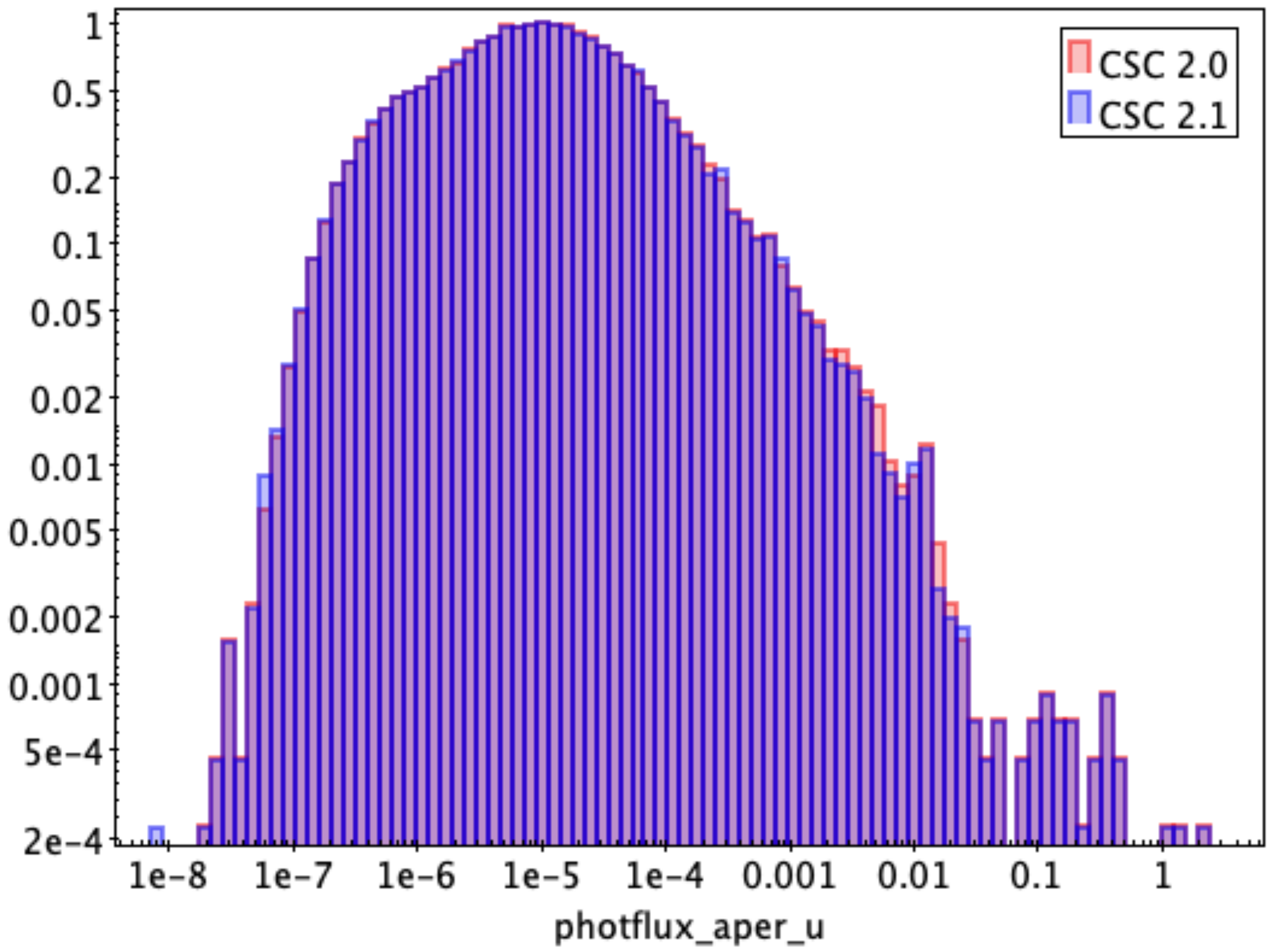


- Statistical consistency between CSC 2.1 and CSC 2.0. Looking at both the properties distributions, and the scatter between CSC 2.0 and CSC 2.1 properties for matching sources.
- The vast majority of properties showed the expected behavior.
- Migration of some properties, in particular a small fraction of source positions, variability flags, and astrometry.
- We have also migrated the variability flag to include inter-observation variability in the variability flag, and updated the var_index thresholds above which a source is considered variable.

All detections



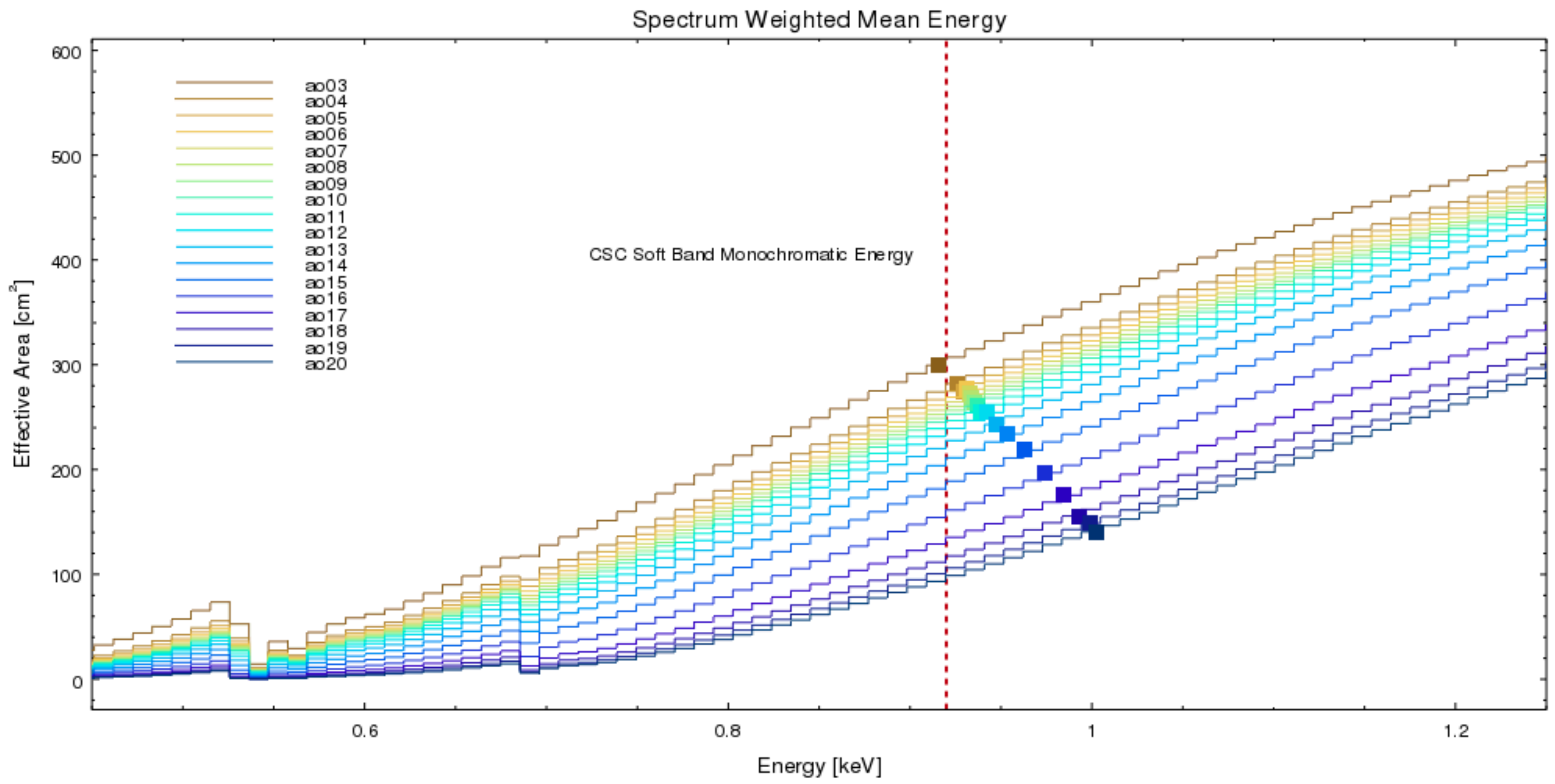
Matched detections

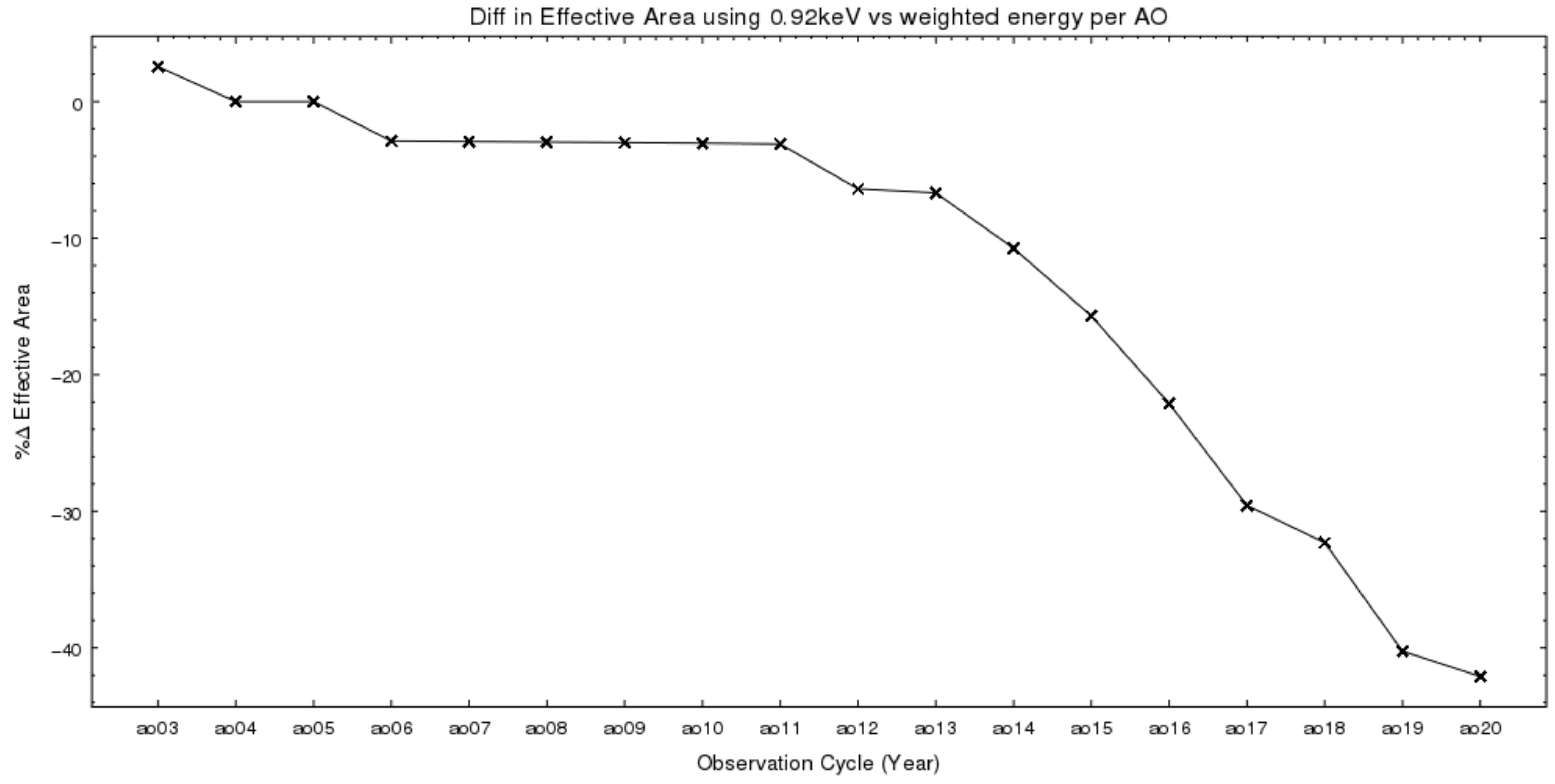


Photon fluxes and contamination

- There is a more prominent tail of high photon fluxes in CSC 2.1 with respect to CSC 2.0.
- This mostly affects the u-band, with a minor effect in the s band.
- The tail is not observed for CSC 2.1 that have a counterpart in CSC 2.0. This indicates that it affects mostly detections in newer observations.
- The issue is due to a drift in effective monochromatic band energy due to a change in low-energy spectral response.
- The additional tail is also not as prominent (or is negligible) in the energy flux distribution, which indicates that CALDB corrections for contamination is effective

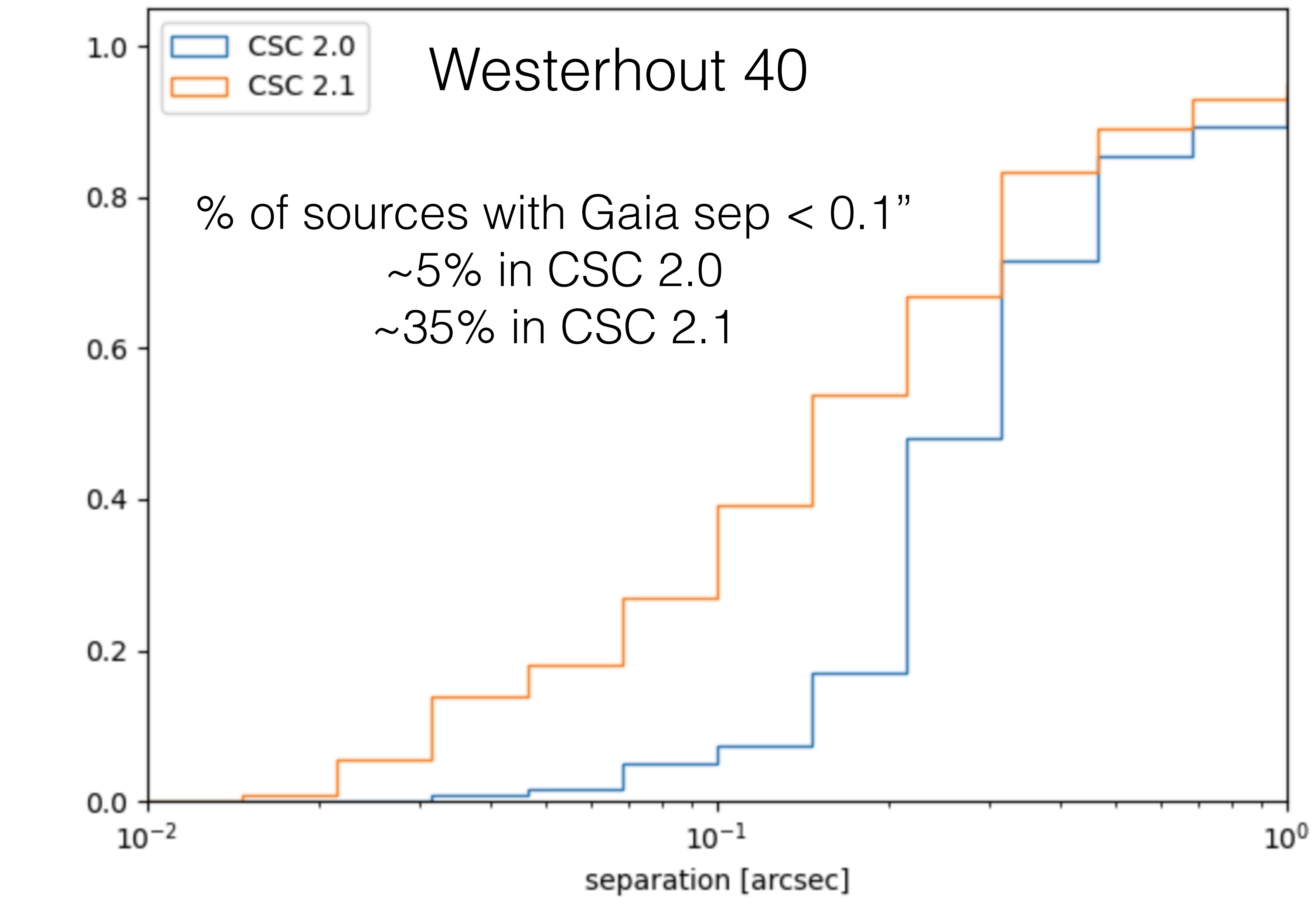
The effect of assuming a monochromatic energy

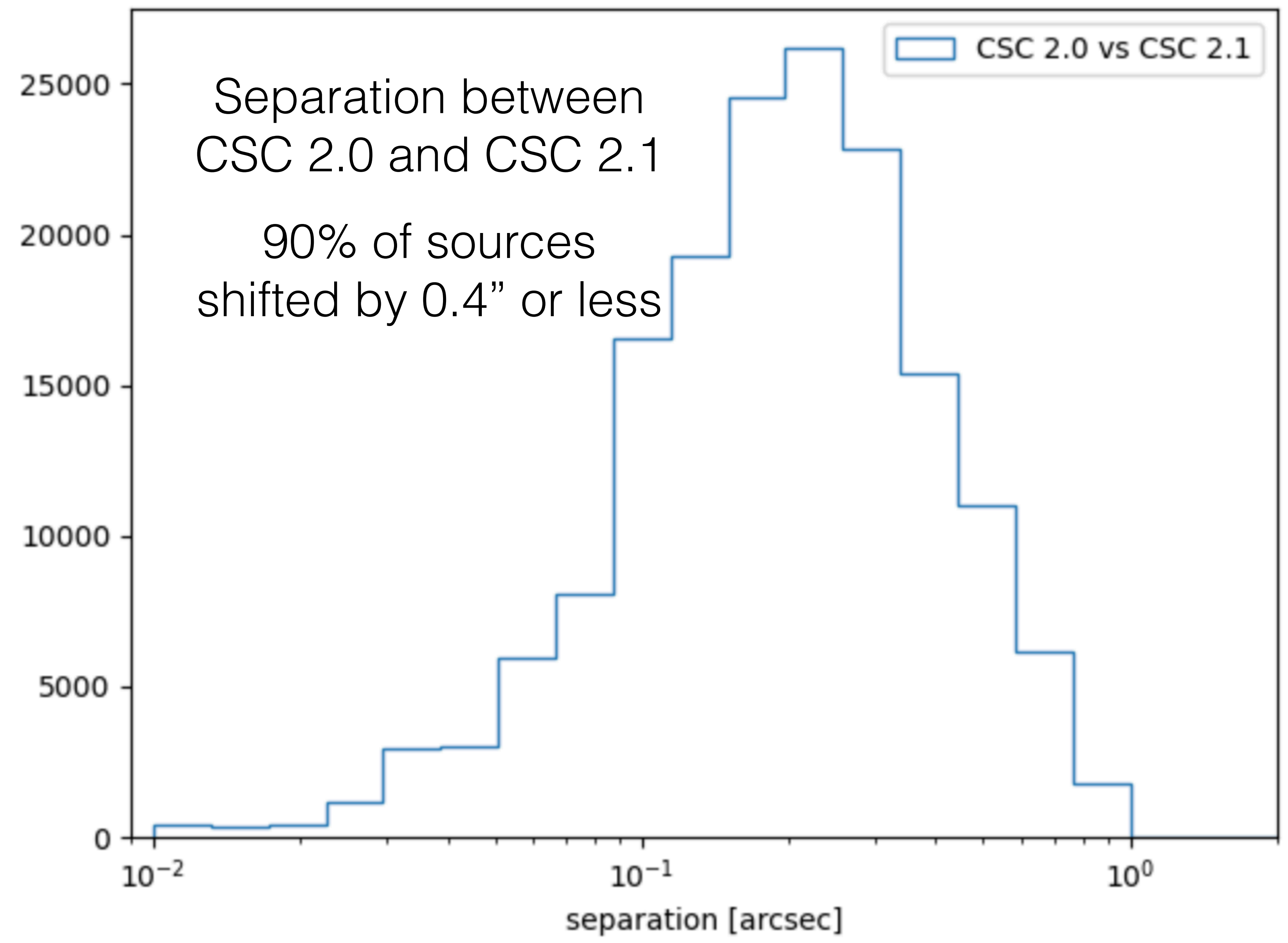




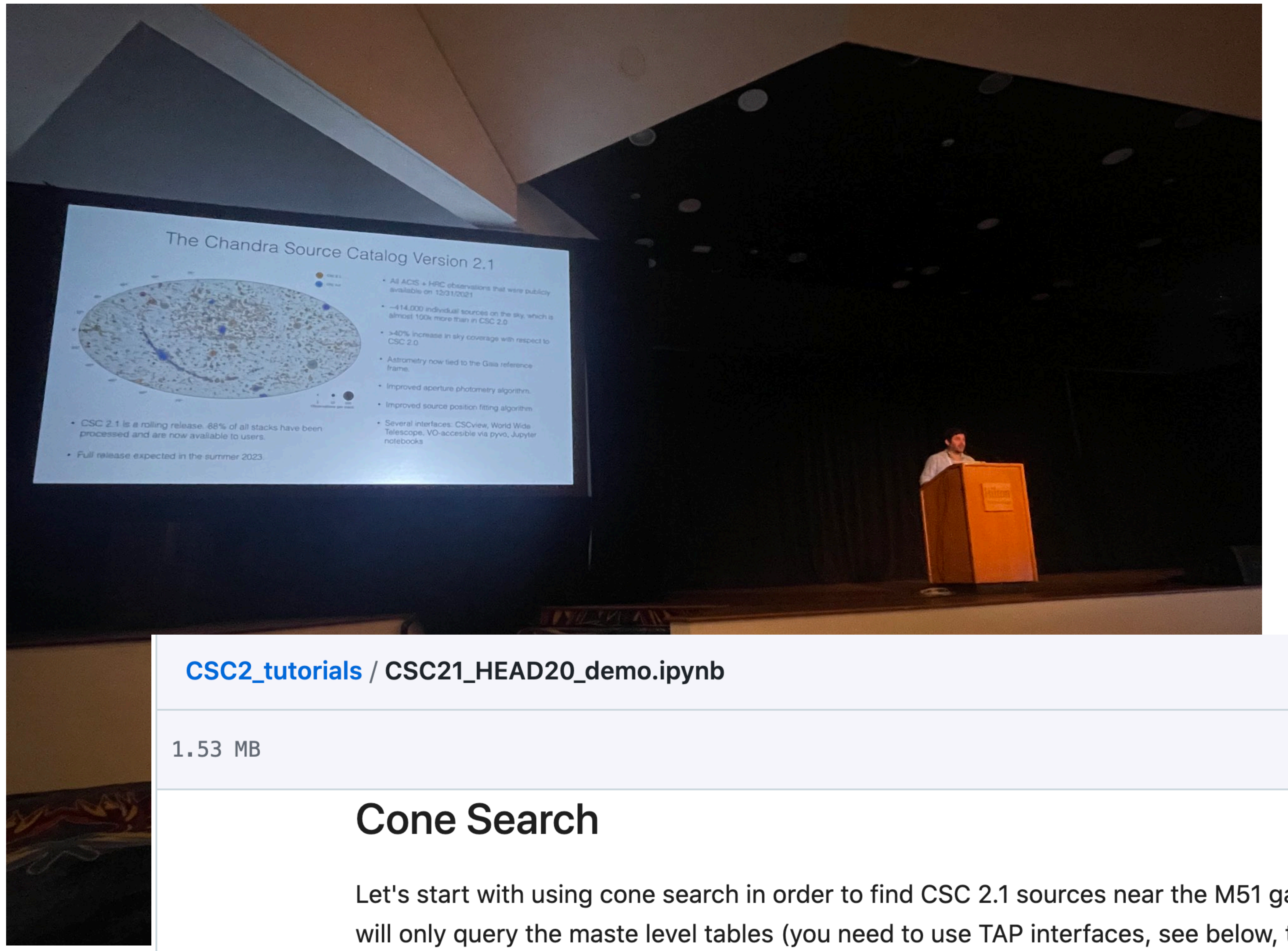
- The srcflux tool is used to compute photon fluxes, net counts, etc.
- We assumed a monochromatic energy to represent spectral response across the energy band to reduce computation effort
- Such assumption degrades as a function of mission time, and significantly affects the effective area in the u and s bands after ~2013.
- This results in photon flux estimates that are higher at later times. This is what we observe in CSC distributions of u band fluxes.
- We will include a correction in a future release of the catalog.

Astrometry improvements





- Anchoring the CSC 2.1 astrometry has resulted in a significant improvement in the absolute position error in the new version.
- The systematic error term in the position error has gone from 0.7” to 0.29” per axis at the 95% confidence level.

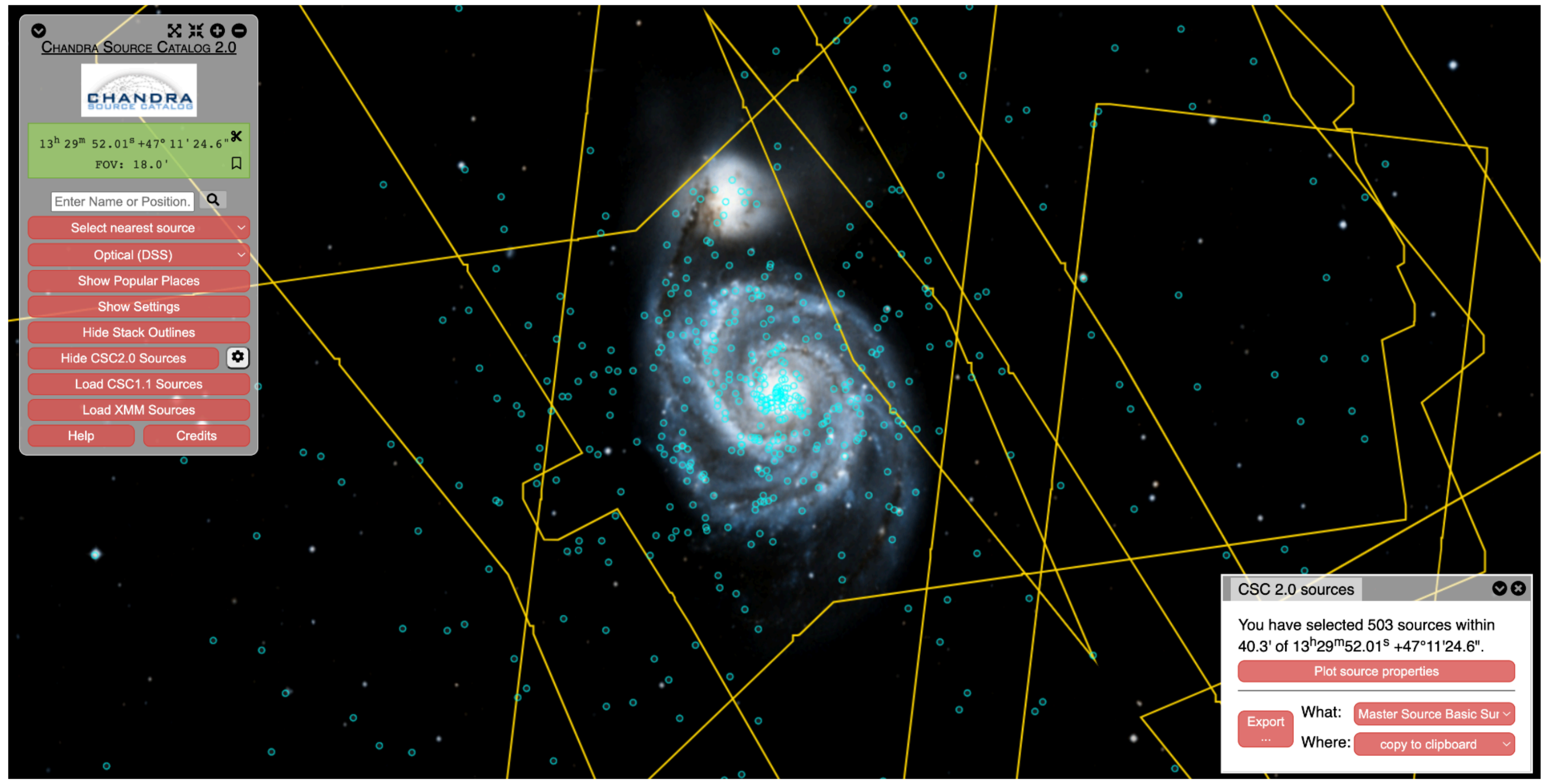


CSC2_tutorials / CSC21_HEAD20_demo.ipynb

1.53 MB

Cone Search

Let's start with using cone search in order to find CSC 2.1 sources near the M51 galaxy. Cone search will only query the maste level tables (you need to use TAP interfaces, see below, to query per-obsid properties). We will use a search radius of 40 arcminutes. We will use the name resolver from Astropy. Here is a visualization of the CSC 2.1 sources in this area, from our World Wide Telescope interface:



Community engagement

- We have been organizing a both hands-on tutorial sessions and science results sessions at the HEAD meetings.
- We plan to keep doing it on a regular basis.
- Jupyter notebook tutorials help in engaging a new generation of data-driven high energy astrophysicists.

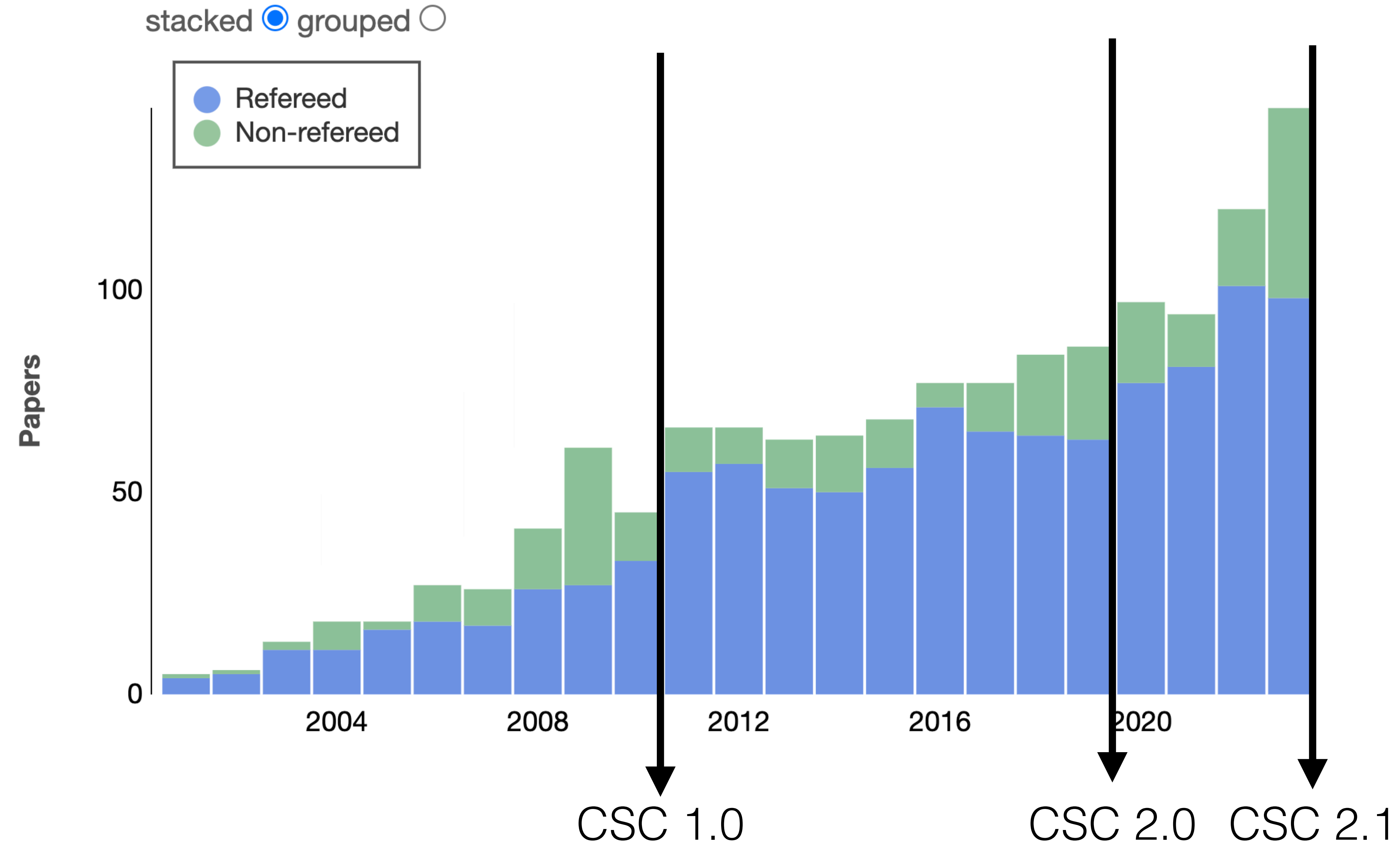
The CDA DOI strategy

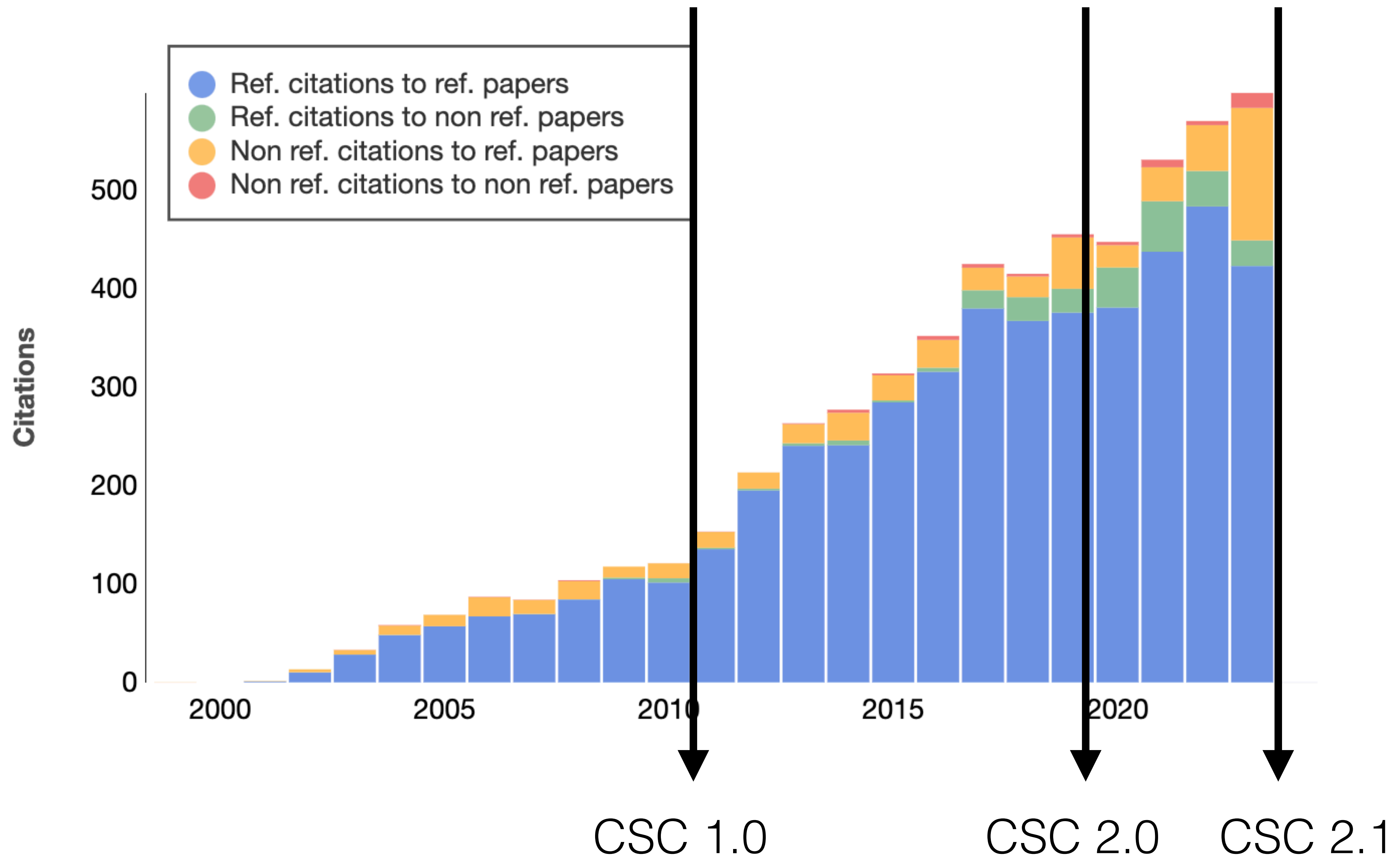
- In order to facilitate citation of Chandra data and increase the reproducibility of scientific results based on Chandra, in 2020 the Chandra Data Archive adopted Datacite Data Object Identifiers (DOIs) for archival observations and catalog data products (to become available with CSC2.1). Recently, we started to mint Chandra Data Collection DOIs (CDC, a.k.a. “paper DOIs”) upon user’s request (https://cxc.cfa.harvard.edu/cda/dois.html).

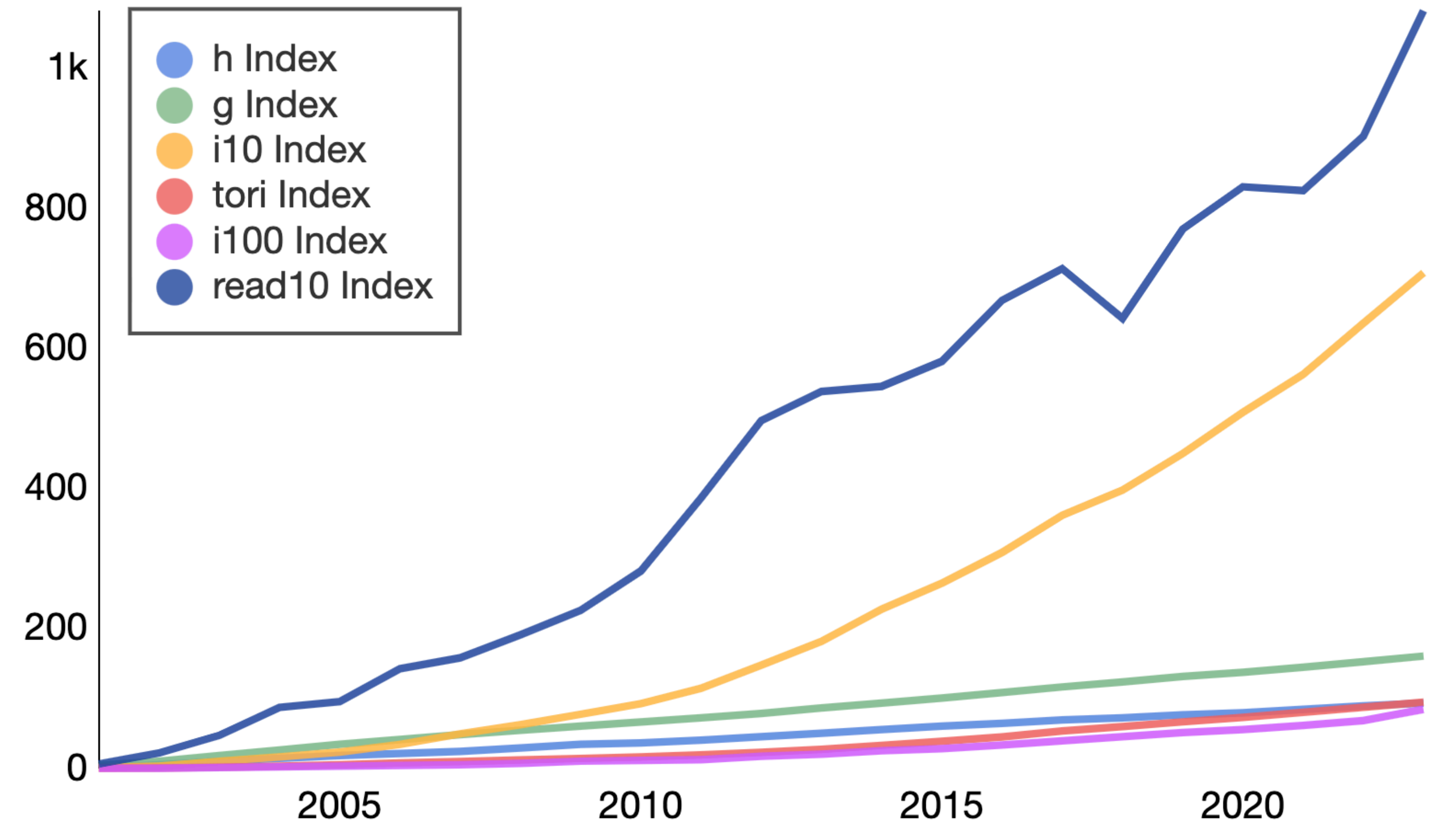
- **Archival observations DOIs**
  - Distinct DOIs for each distinct observation
    - DOI based on *ObsID* (https://doi.org/10.25574/23336)

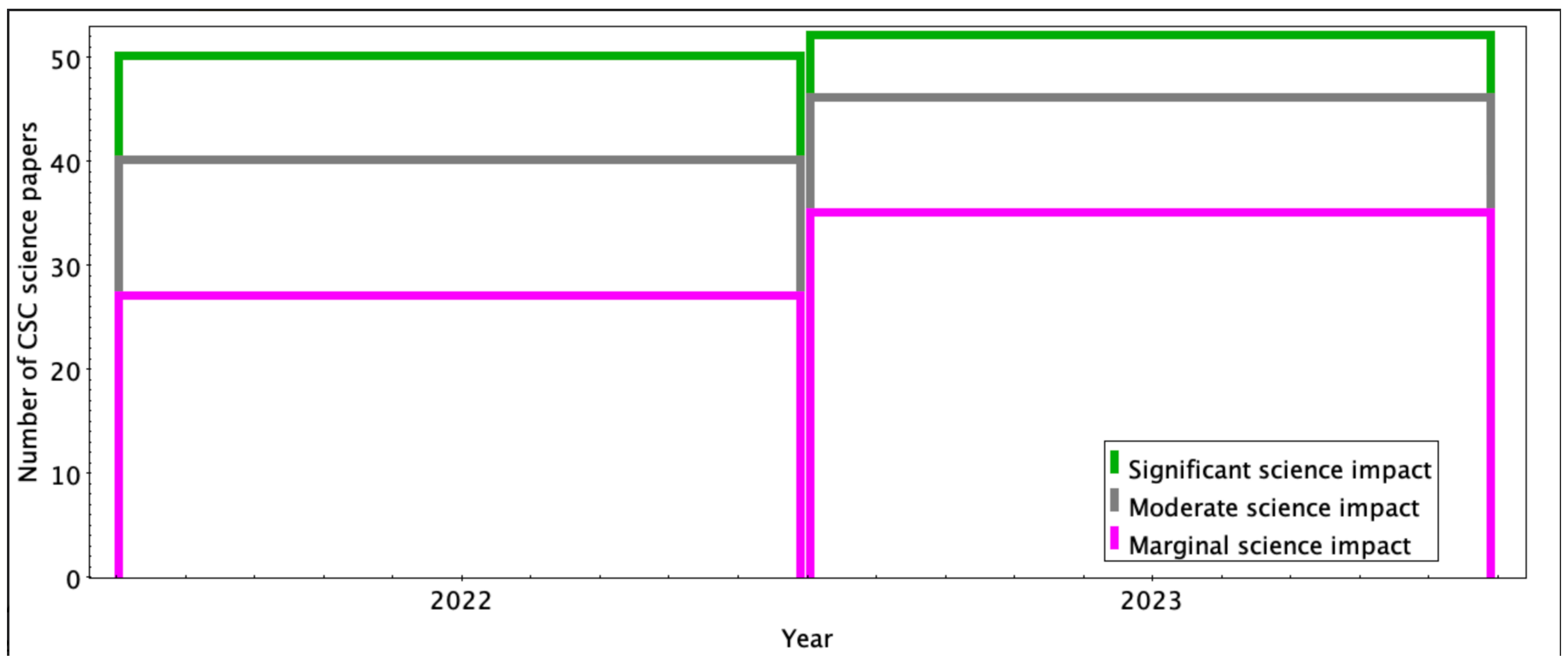
- **Chandra Source Catalog DOIs (starting with CSC2.1)**
  - One single DOI for each major version of the CSC: 10.25574/csc2
  - Distinct DOIs for full-field data products for CSC stacks and observations
    - CSC Stacks: 10.25574/csc2.stk.acisfJ1509217p073311_001
    - CSC ObsIDs: 10.25574/csc2.obs.18225
- **Chandra Data Collection (CDC) DOIs**
  - DOIs expressing an arbitrary collection of Chandra data entities
    - **Usually** associated with a publication
    - Created based on users’ requests

The community usage of the CSC has increased

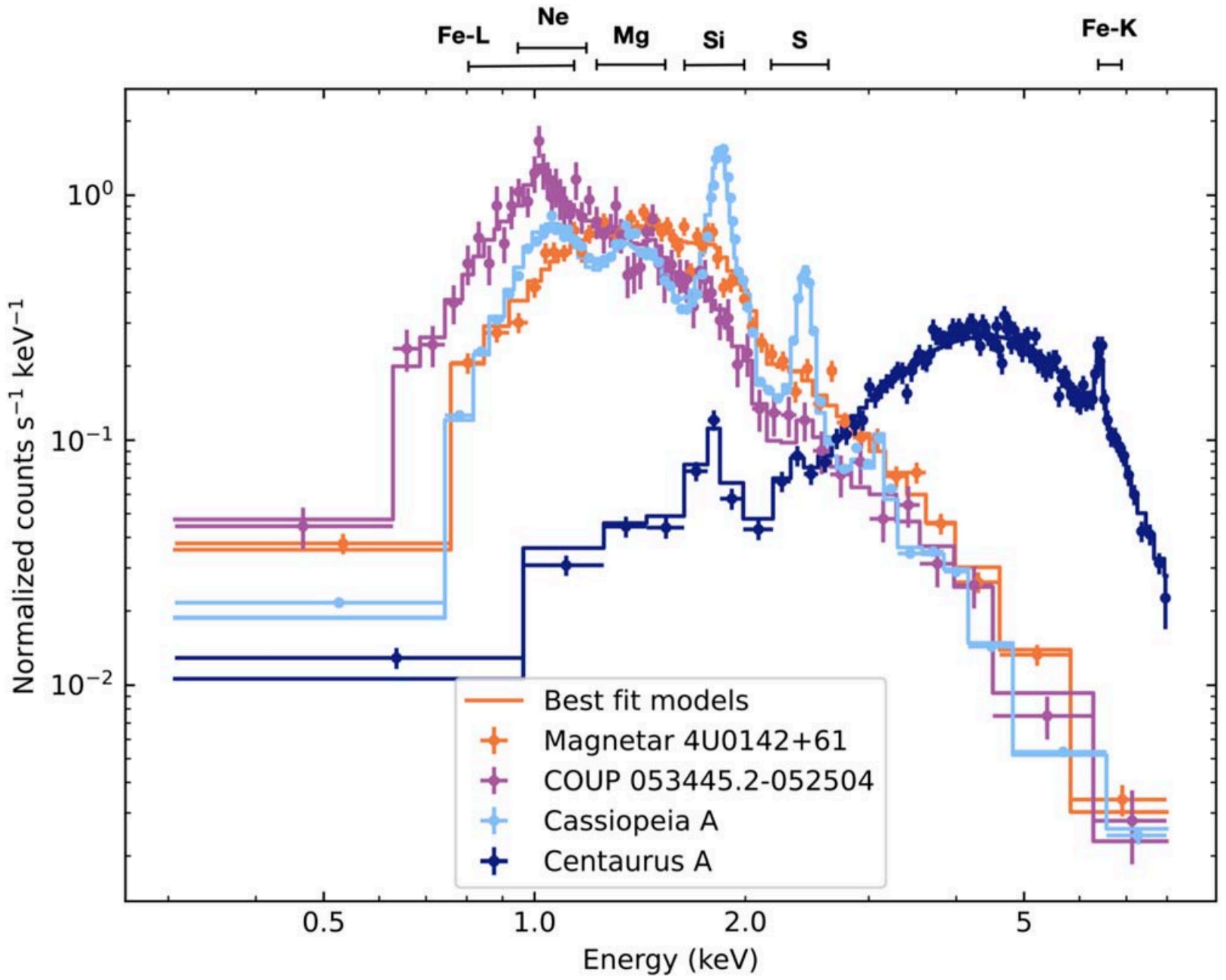


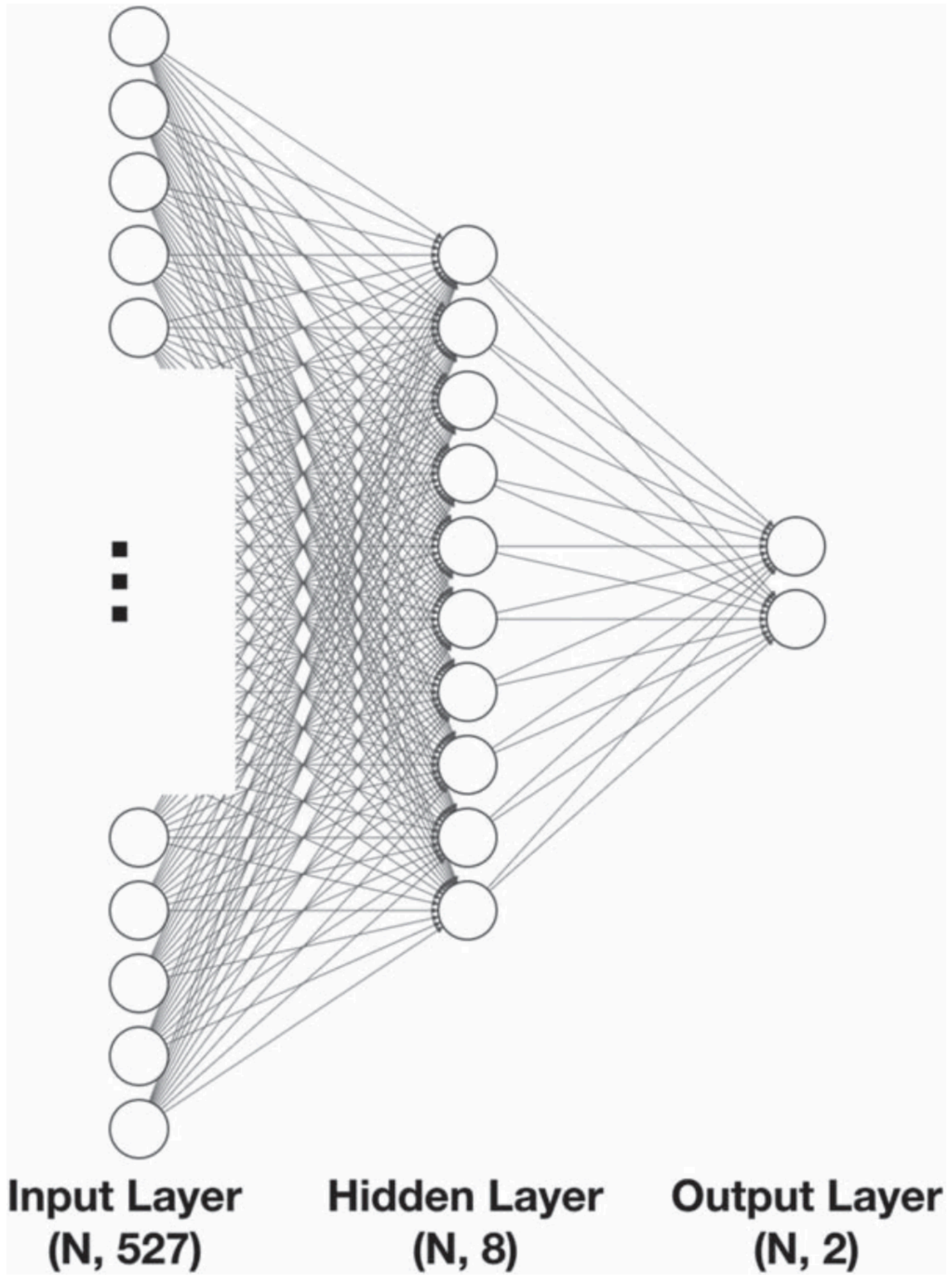


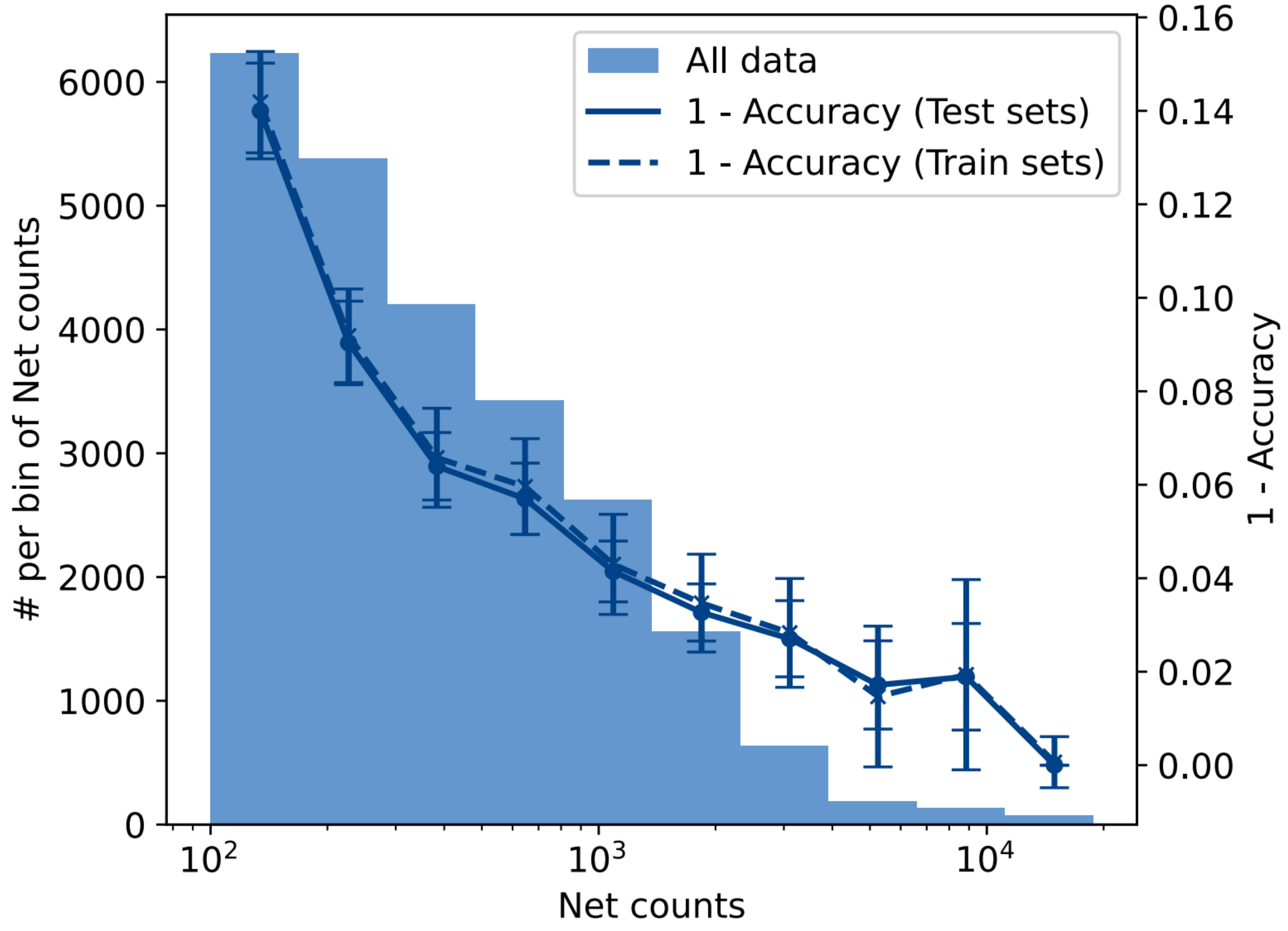




Classification of CSC spectra using Neural Nets (Pavan et al. 2023)

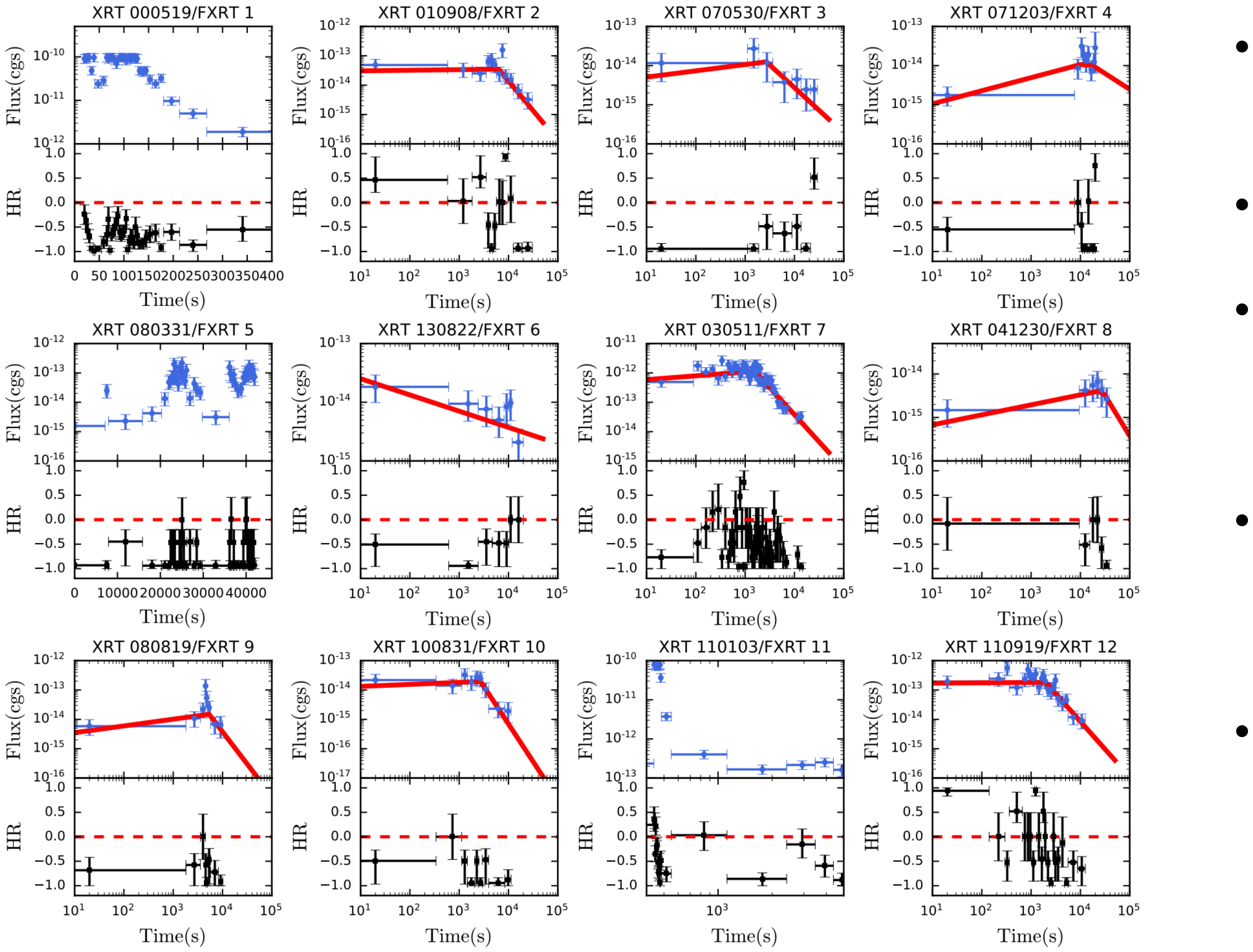






- Authors use CSC data products (spectra) to generate a set of input parameters for spectral simulations.
- They then train a NN to classify between stars and AGNs, using the spectra themselves as the input to the ML algorithm
- They then apply the trained model to both the simulated and observed CSC spectra, and achieve accuracies over 90%
- See also Yang et al. 2022, Chen et al. 2023, Kumaran et al. 2023, Perez et al 2023

Successful systematic searches for FXRTs in the CSC

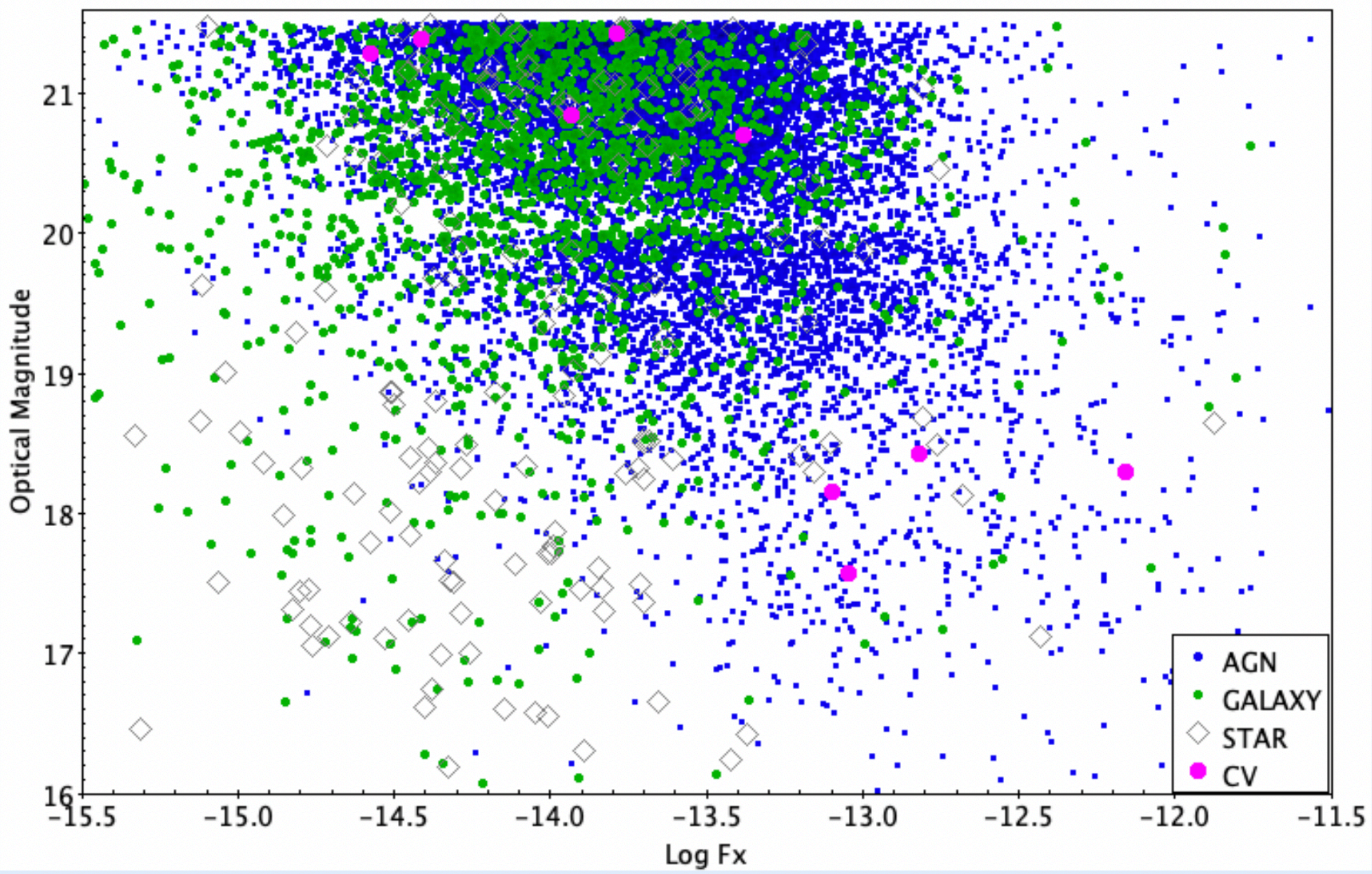


- 14 FXRTs that are consistent with extragalactic origin
- None detected as a sGRB.
- Majority consistent with XRBs, ULXs, via association with optical counterparts, X-ray luminosity
- For at least 3 FXRTs, their lack of optical counterparts, luminosity range consistent with off-axis GRBs, or TDEs
- "…progress here will crucially hinge upon the ability of current and future X-ray observatories to carry out efficient strategies for (onboard) detection and alert generation to trigger follow-up campaigns "

Quirola-Vásquez et al. 2022
Quirola-Vásquez et al. 2023
Dillmann et al. 2023

Chandra, meet SDSS. SDSS spectroscopy of CSC counterparts (Green et al. 2023)

Optical/IR Catalog	Magnitude Range	% Matches
Gaia DR3	$14 < G < 20$	37
Legacy	$14 < (g\|r\|z) < 21.5$	38
PS1	$14 < (g\|r\|i\|z) < 21.5$	16
2MASS	$H <= 14$	8



- The CSC has partnered with SDSS-V to obtain additional optical and infrared spectroscopy for ~40,000 CSC sources.
- This adds to the ~17,000 CSC sources that already have SDSS spectra
- SDSS-V will provide IR spectra for a fraction of CSC galactic plane sources that have infrared but no optical counterparts.
- Most of the X-ray sources matched to H<14 2MASS counterparts will be young stellar objects or X-ray binaries buried by dust in the Galactic plane.
- The Chandra X-ray Center will be releasing to the public catalogs of these matches with basic source properties at

  https://cxc.cfa.harvard.edu/csc/csc_crossmatches.html