The Chandra Source Catalog

G. Fabbiano

Presentation to the CUC - 27 September 2016

Chandra Source Catalog

Mining the High Resolution X-ray Sky

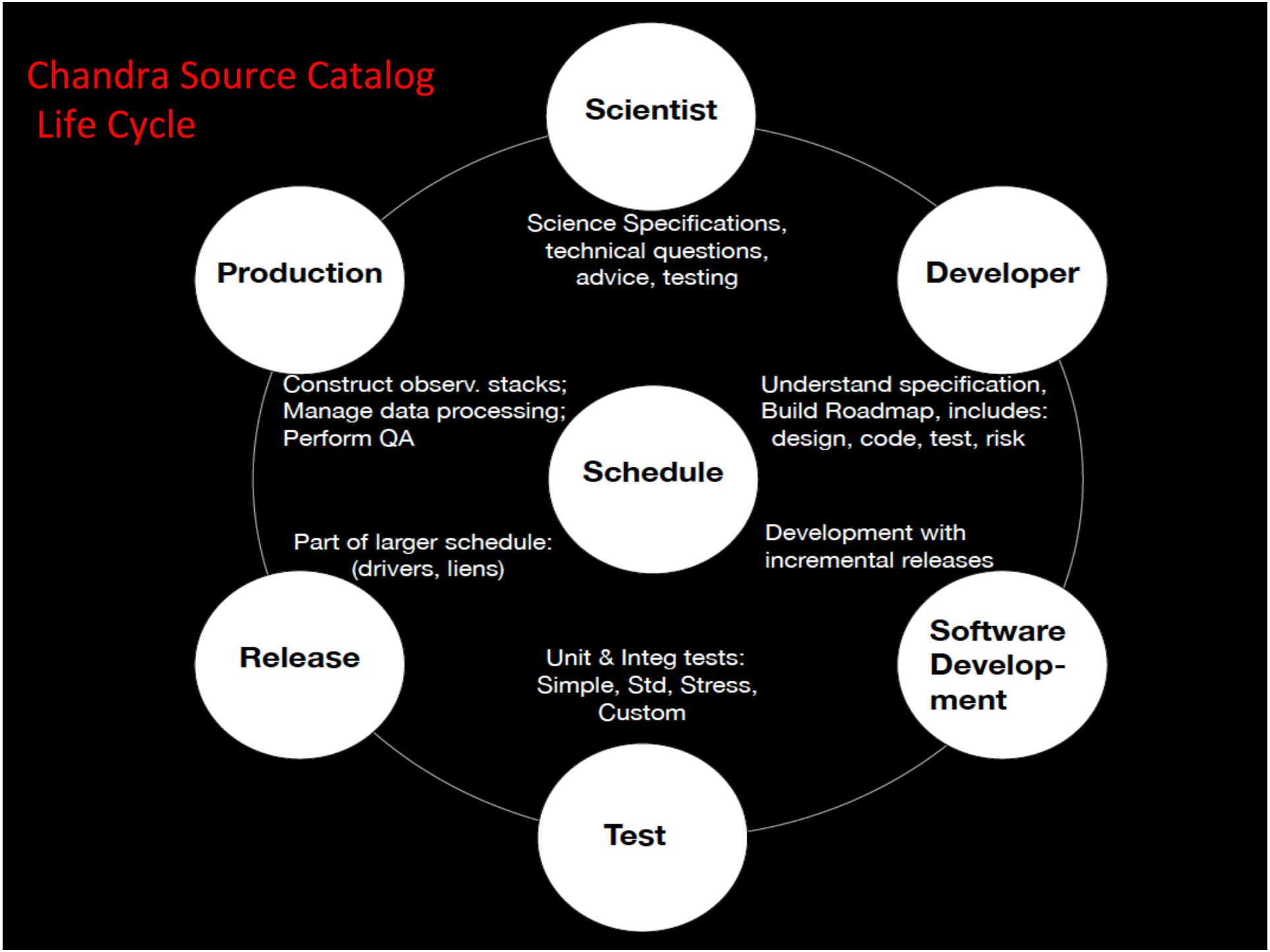
- Highly endorsed by the 2016 NASA Senior Review (SR)
- The **Chandra Legacy Product**
  - CSC2: The first of the co-added legacy releases
    - Data up to 2014
  - To be followed by releases at ~2-3yr intervals
    - More data
    - Software / algorithm upgrades
  - Final release after end of mission
    - Archival data processing
    - Final calibration

- SR also recommends that the CXC provide the **tools to exploit scientifically the CSC**

CSC2 Team & Leads

- The CSC2 is a 10 FTE effort that needs to be fit into ongoing Chandra activities
  - Science (DS + SDS) - Lead: Ian Evans
  - Software (Infrastructure/Pipelines, Tools, Archive) - Manager: Janet Evans
  - Data Processing (QA) – Manager: Joy Nichols
  - Systems - Manager: Durai

Chandra Source Catalog Life Cycle



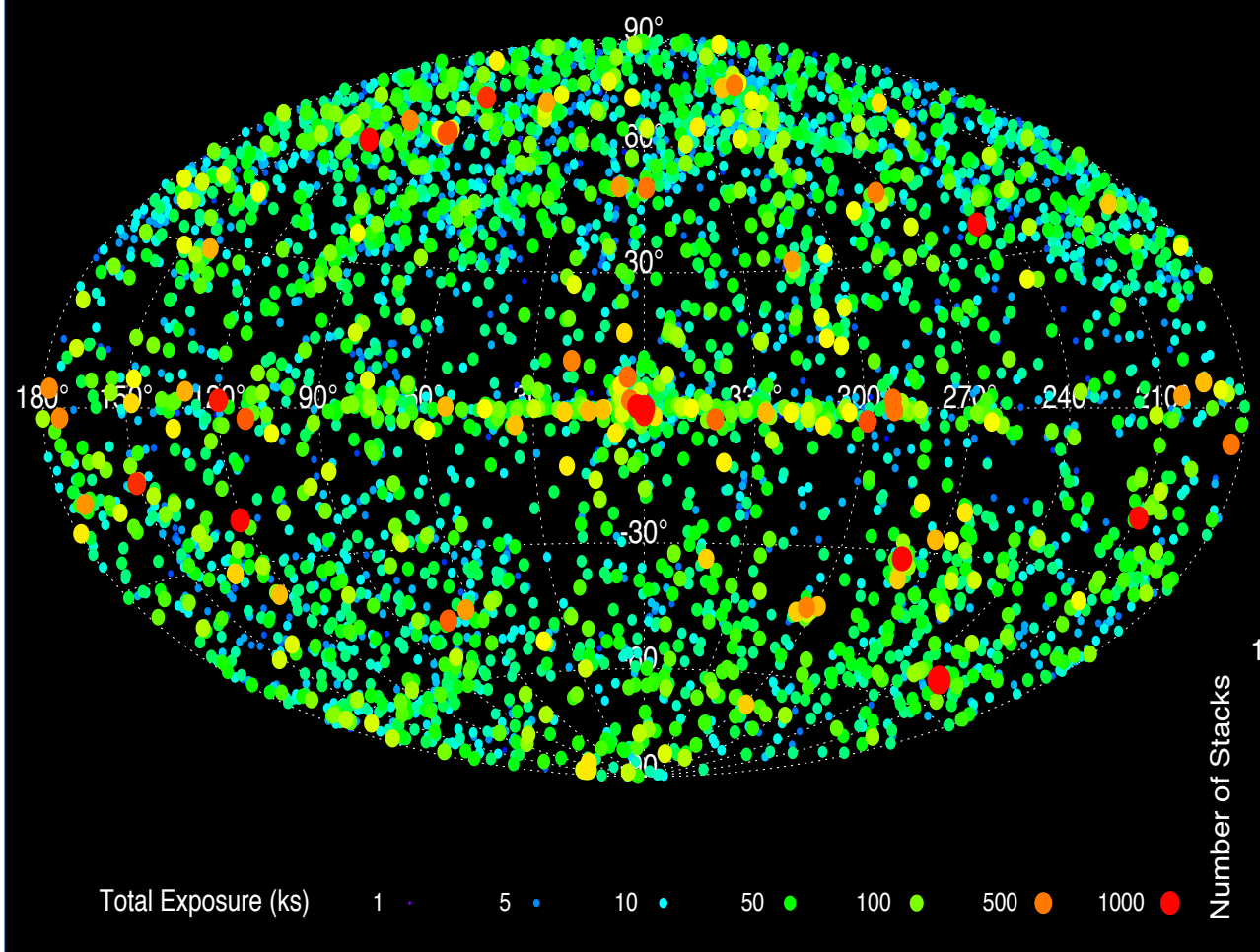
CSC2

- CSC2 is in production
  ~350,000 detections expected, including moderately extended
- Major improvement over CSC1 is co-adding observations
  - Co-added observations maximize sensitivity
    - CSC2 longest exposure stack 5.8 Ms (CDFS) vs CSC1 173.8 ks (Orion)
  - CSC2 doubles total Chandra exposure time used
    - CSC2 t_{exp} = 245.8 Ms vs CSC1 t_{exp} = 104.1 Ms
  - CSC2 doubles sky coverage
    - Current Chandra sky coverage ~785 deg^2 (1.9% of sky)
    - CSC2 sky coverage ~700 deg^2 (1.7%
    - CSC1 sky coverage ~320 deg^2 (0.8% of the sky)

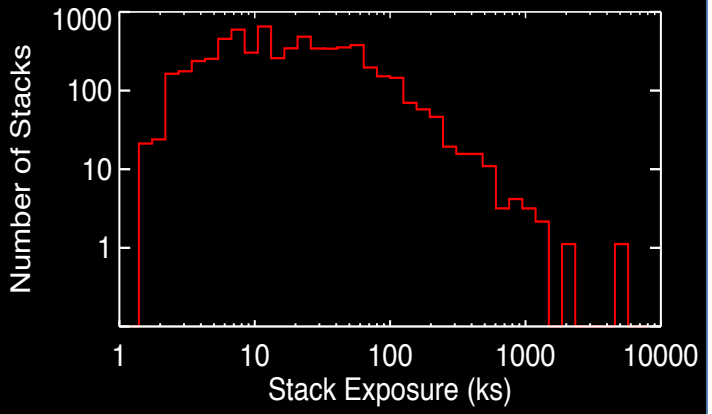
Chandra Source Catalog Release 2

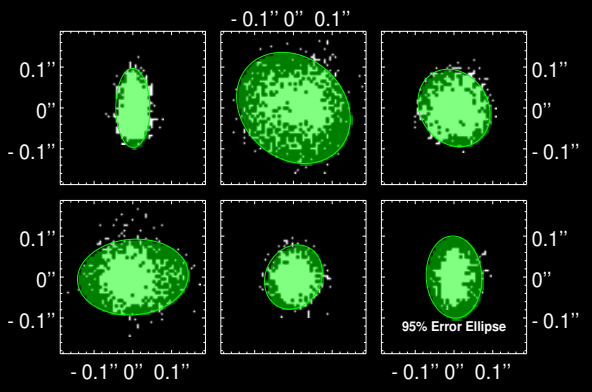
Mining the high-resolution X-ray sky

Source positions, calibrated photons, multi-band X-ray photometry, images, spectra, and light-curves for your analysis

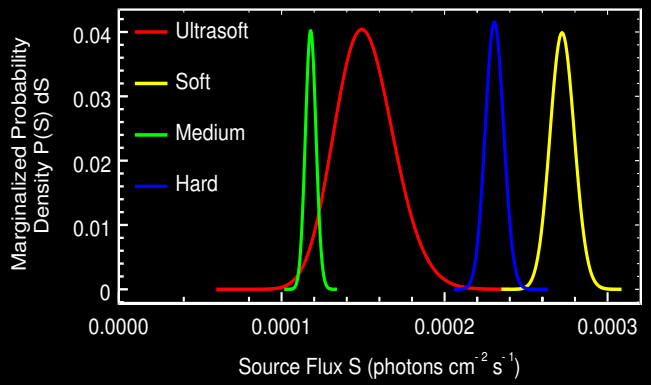


10382 individual observations
7302 stacks (aimpoints < 1’ apart)
245.8 Ms total exposure
5.8 Ms largest stack (CDFS)

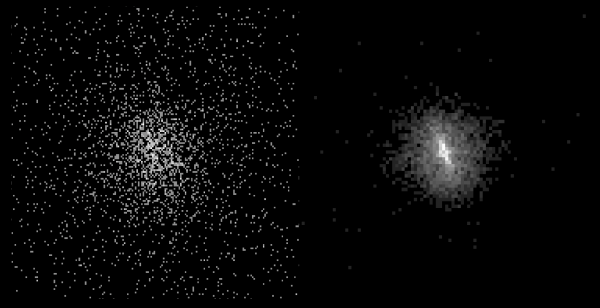




Position error ellipses with position confidence MCMC draws



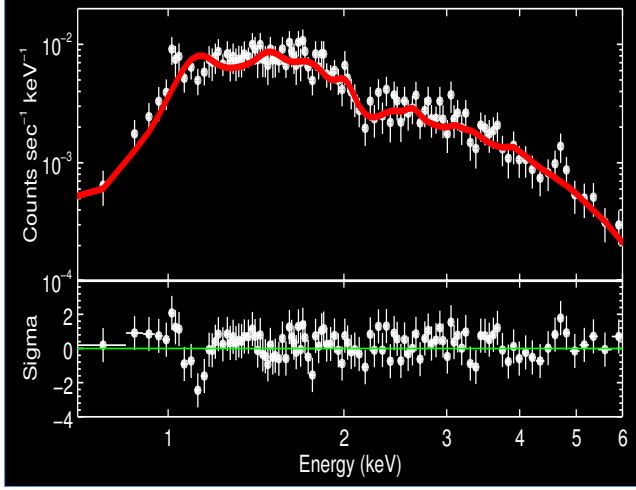
Multi-band X-ray aperture photometry with
Bayesian probability density functions



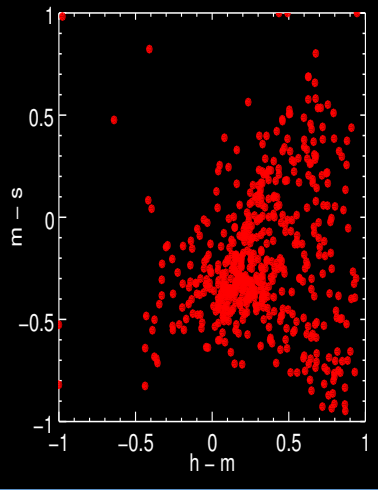
Source extent and local PSF models for every source and energy band

Source properties — all have associated upper and lower confidence bounds

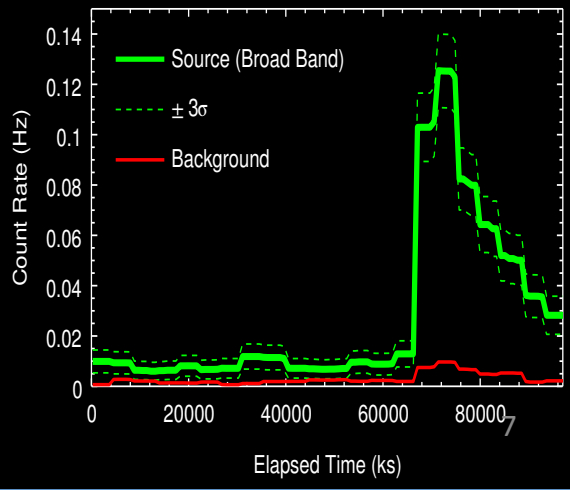
Spectral model fits and fluxes determined using multiple models (>150 cts)



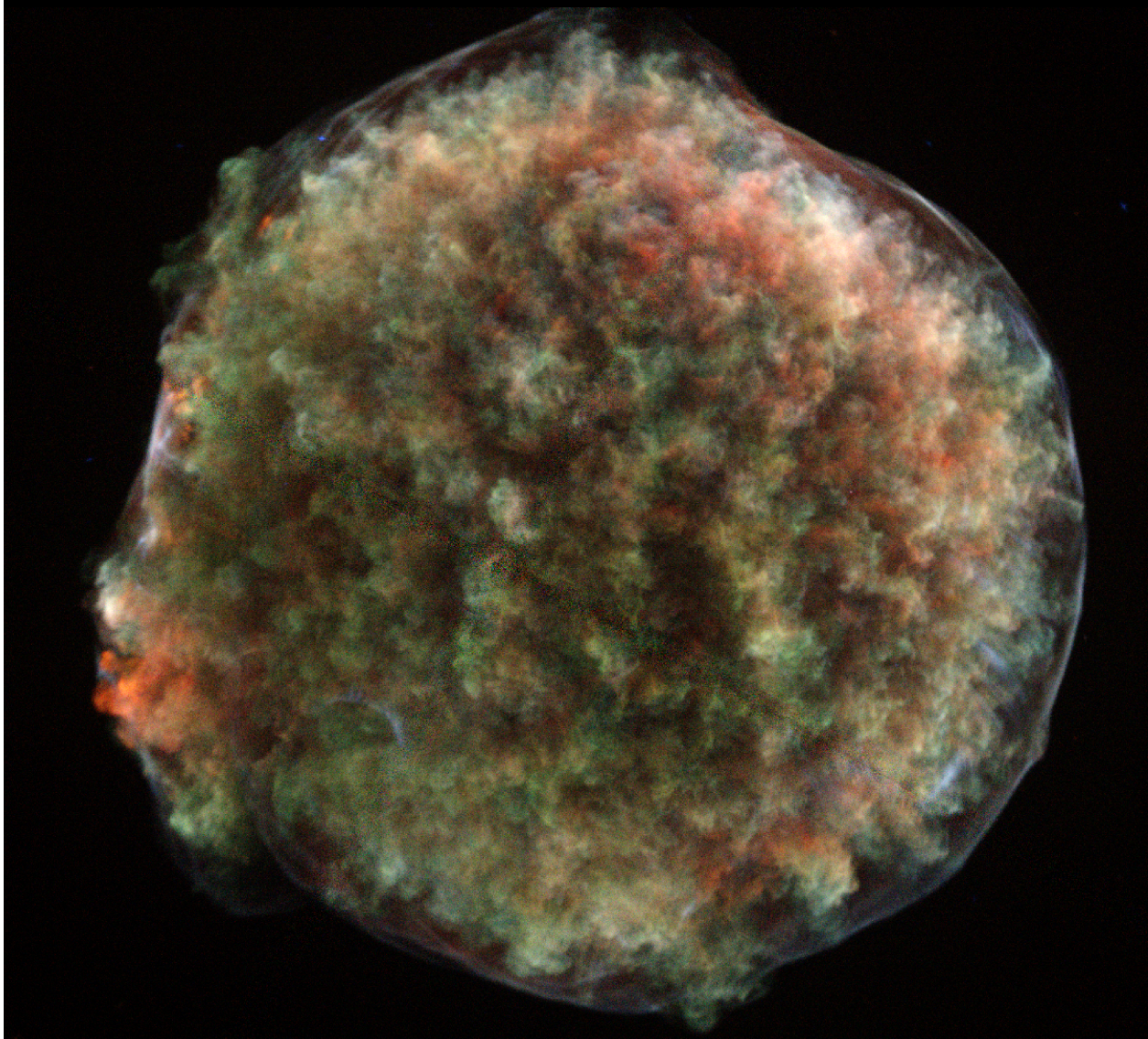
Hardness ratios



Intra- and inter-observation variability measures and light curves



Release 2 of the catalog includes extended X-ray emission in addition to point and compact sources



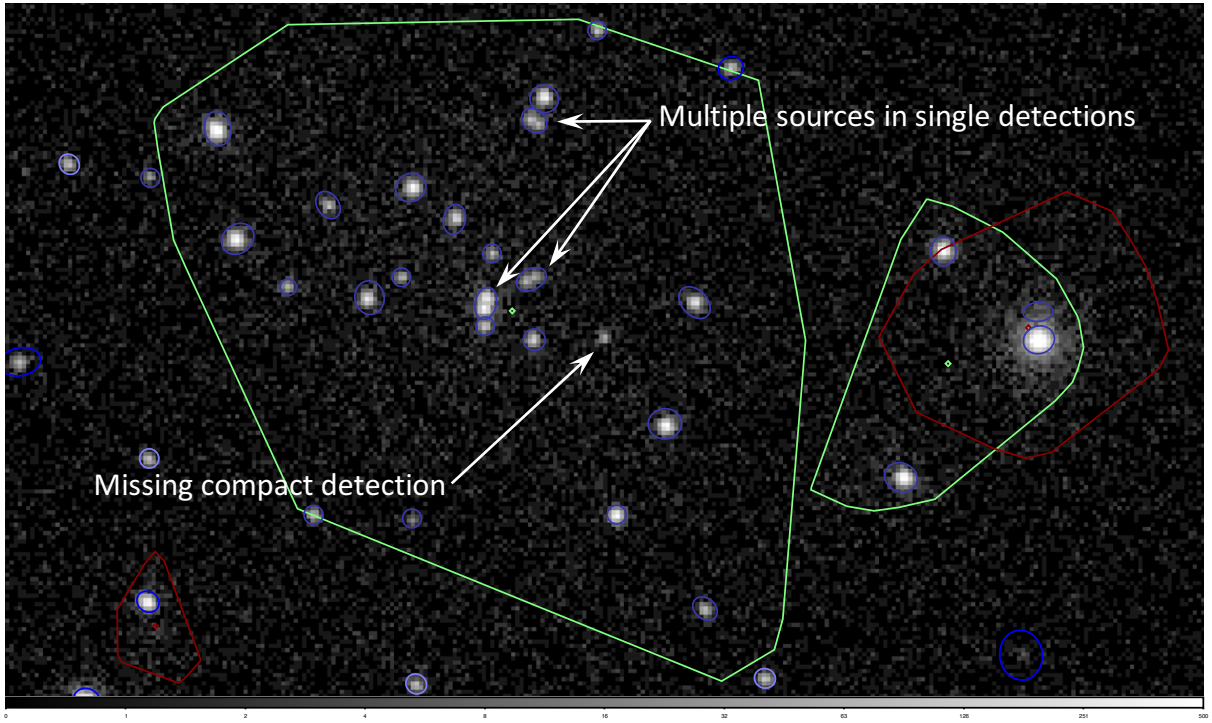
For extended sources (e.g. Tycho's SNR) Photometric properties are integrated over a bounding region aperture

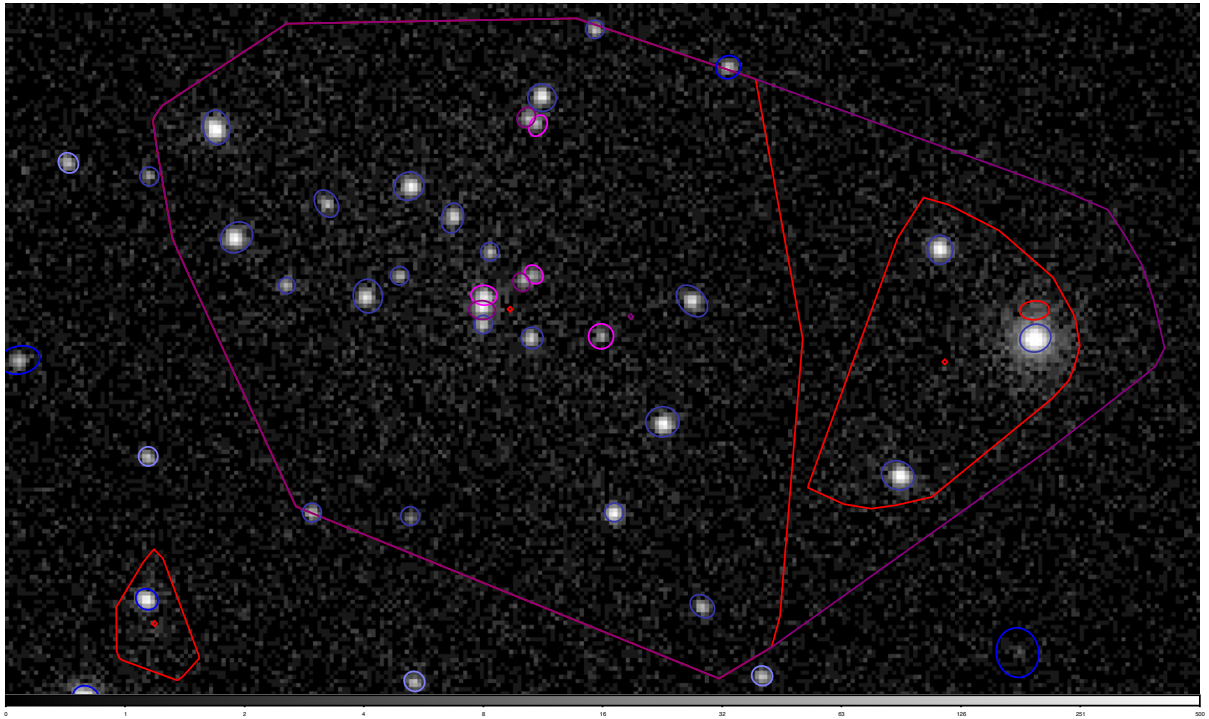
CSC2 Status

- First preliminary release of ~240,000 detections in Summer 2015
- Production run and Quality Assurance (QA) for stacking pipeline processing ongoing
  - ~250,000 detections cleared
  - ~100,000 to go
  - QA is important –cross referencing of QA team recommendations identified need for additional training and partial reprocessing
- In parallel ongoing preparation for next phases

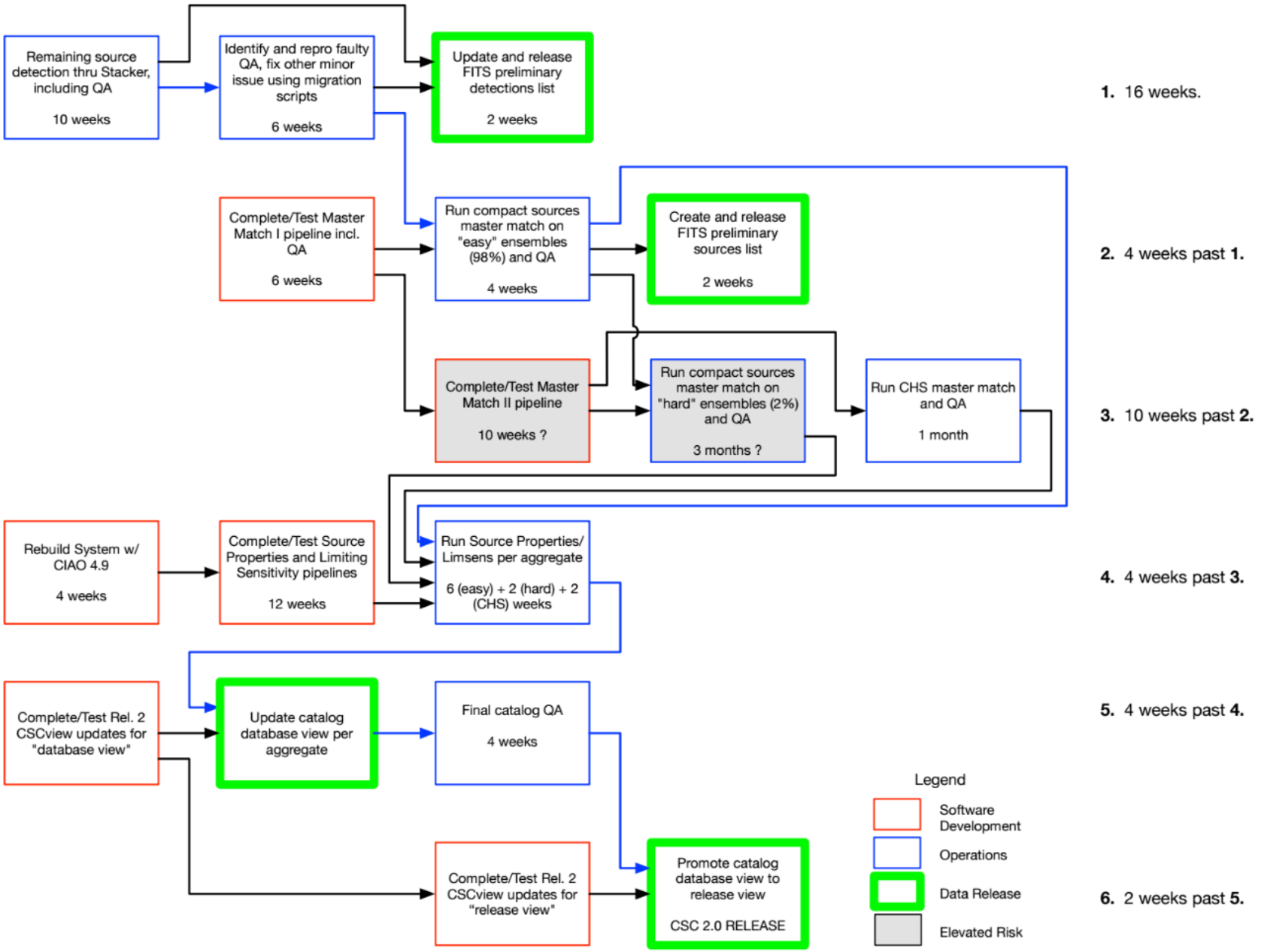
Example of Manual QA:

- Manual QA triggered by software test
- <10% of field affected
- Is the polygon a good representation of the extended emission?
- Are these point detections good for MLE?

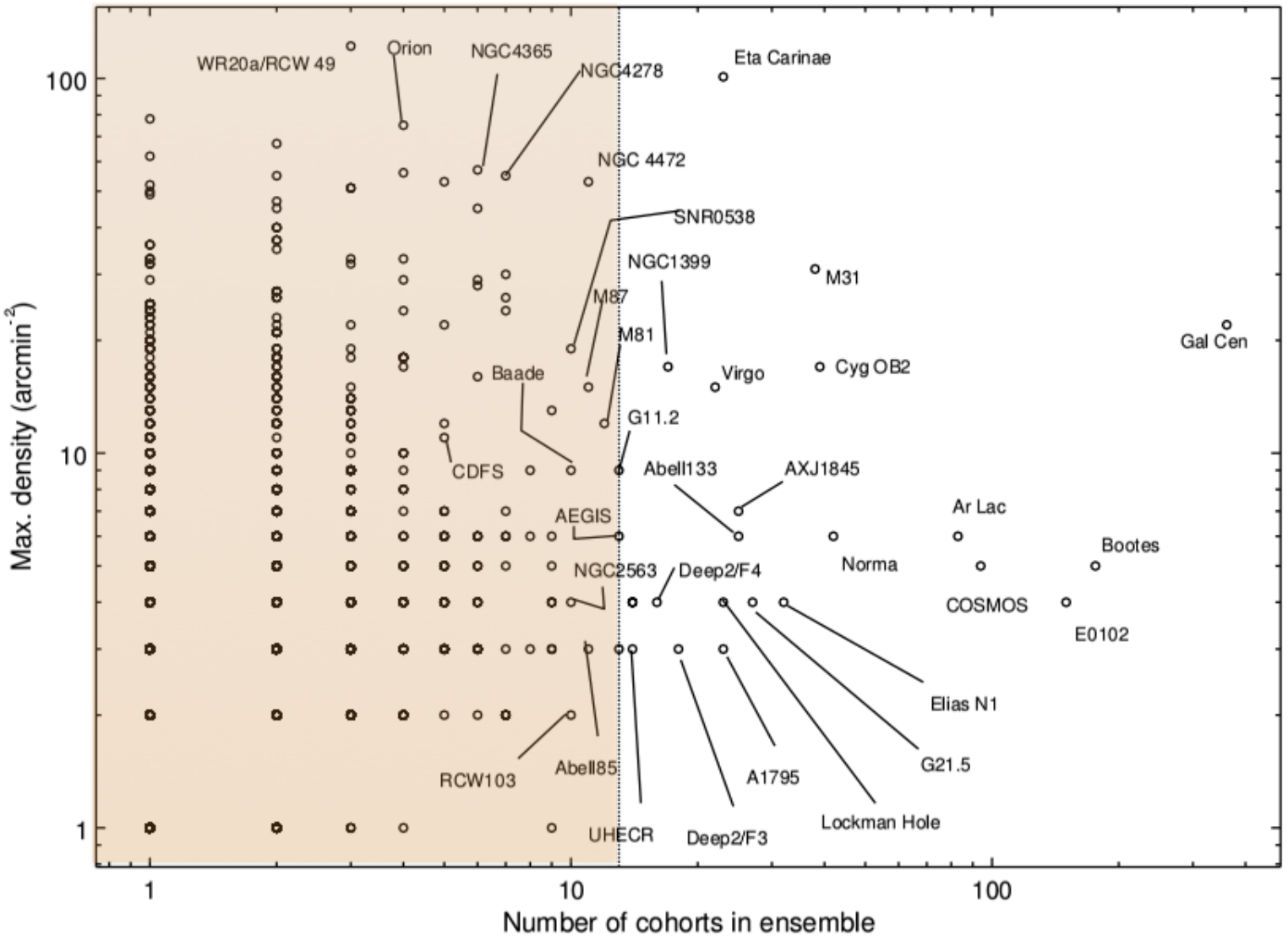




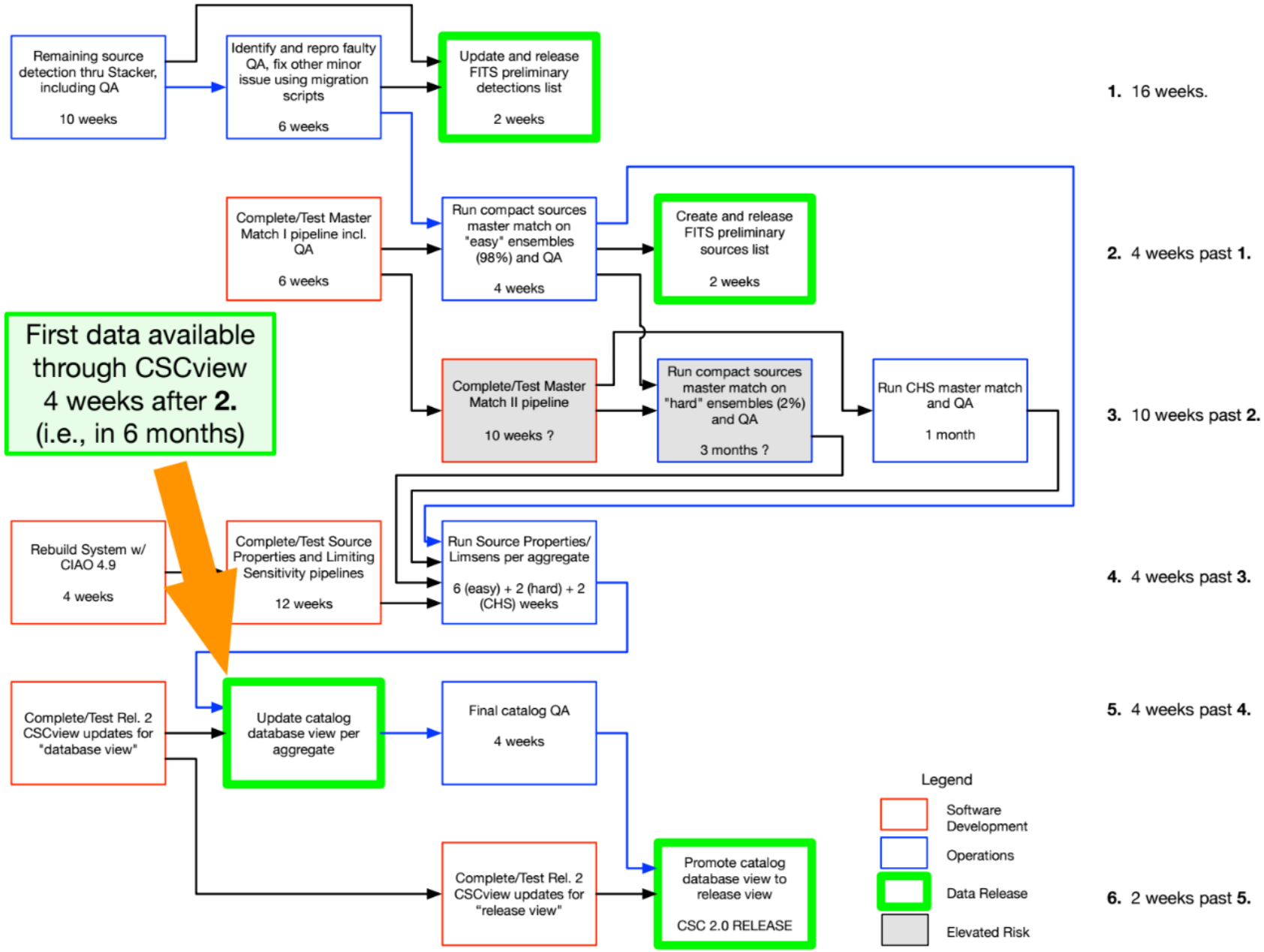
CSC2 Re-plan Schedule: 10 months to final CSC2 release



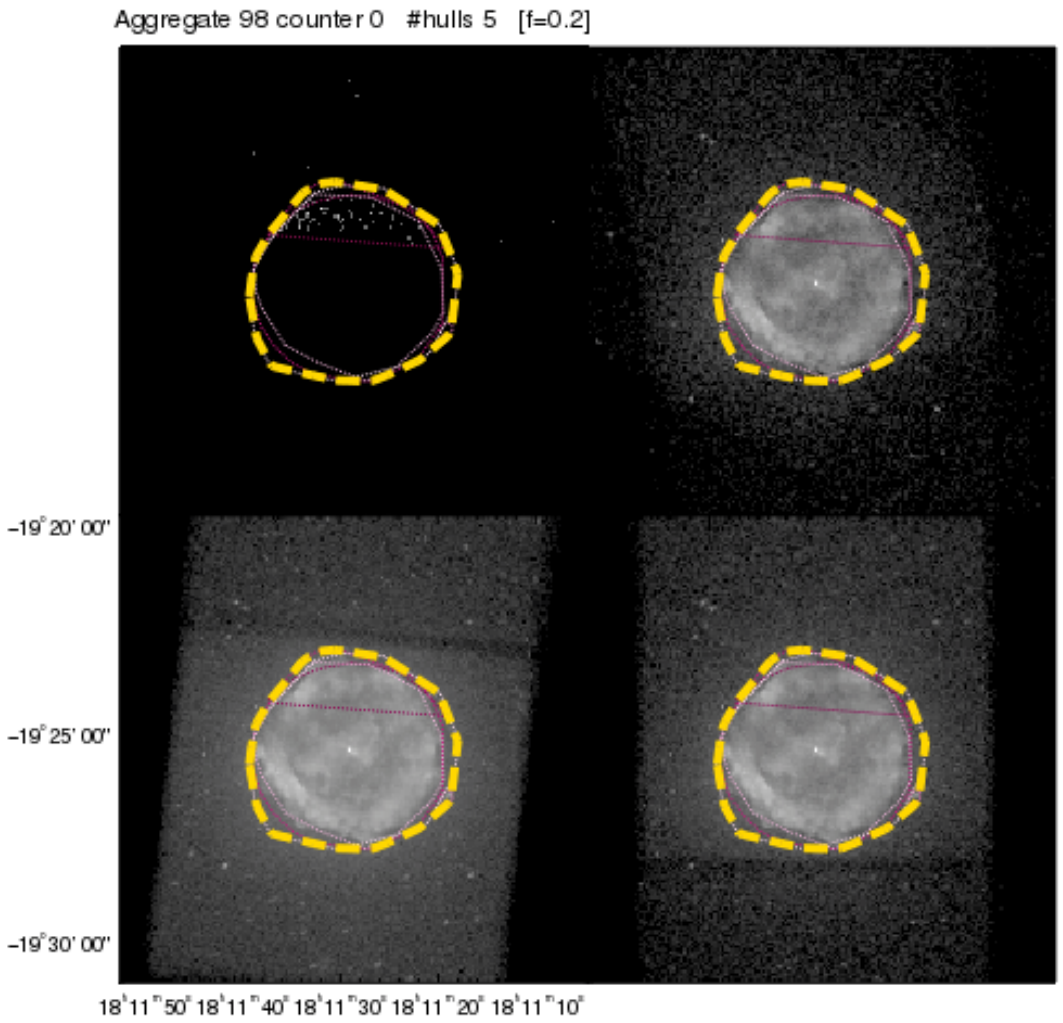
The 98% ‘easy cross-match’ sample

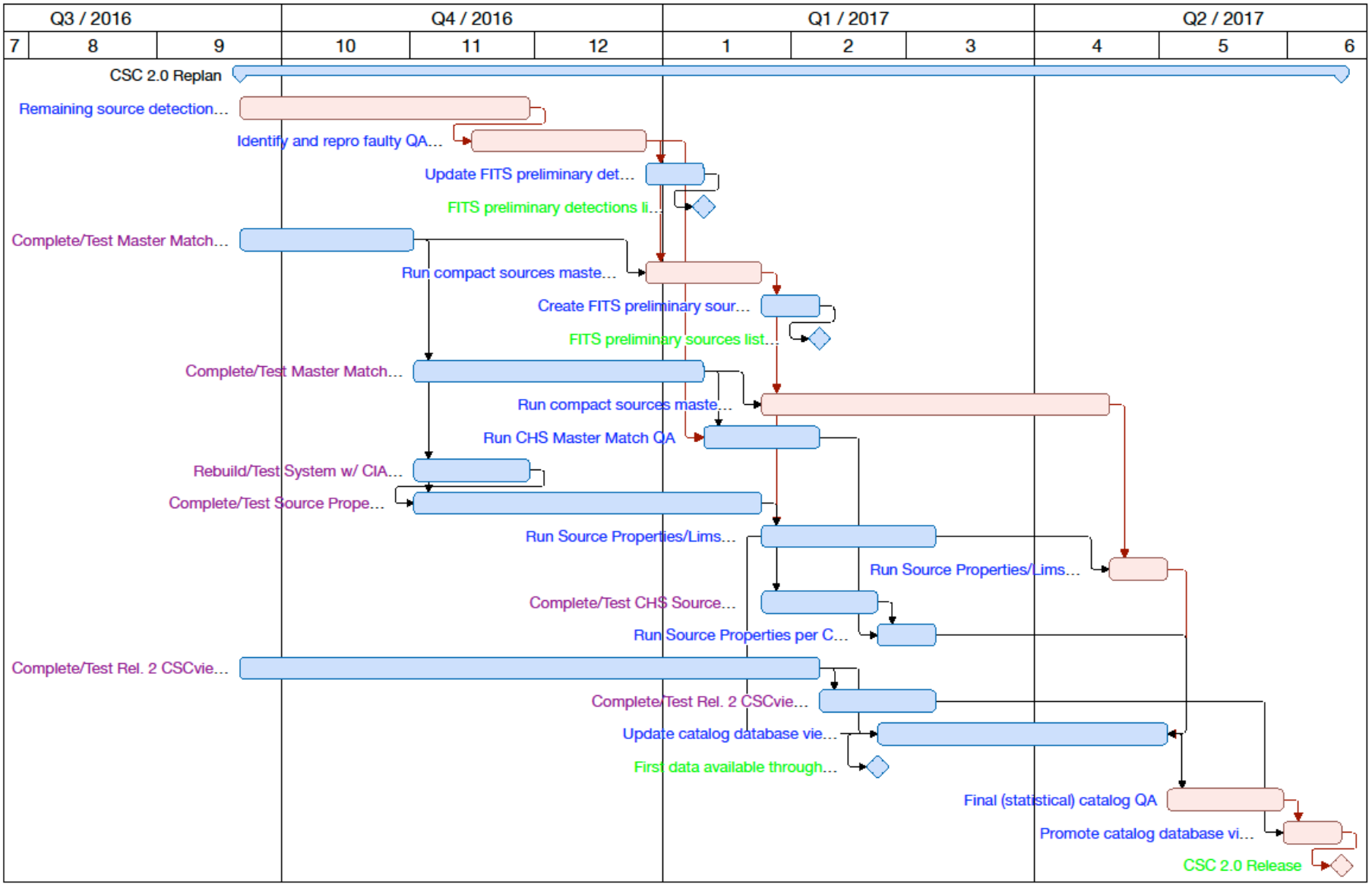


CSC2 Re-plan Schedule: 10 months to final CSC2 release



Convex Hull Source Master Match




Q3 / 2016
Q4 / 2016
Q1 / 2017
Q2 / 2017
7
8
9
10
11
12
1
2
3
4
5
6
CSC 2.0 Replan
Remaining source detection...
Identify and repro faulty QA...
Update FITS preliminary det...
FITS preliminary detections li...
Complete/Test Master Match...
Run compact sources maste...
Create FITS preliminary sour...
FITS preliminary sources list...
Complete/Test Master Match...
Run compact sources maste...
Run CHS Master Match QA
Rebuild/Test System w/ CIA...
Complete/Test Source Prope...
Run Source Properties/Lims...
Run Source Properties/Lims...
Complete/Test CHS Source...
Run Source Properties per C...
Complete/Test Rel. 2 CSCvie...
Complete/Test Rel. 2 CSCvie...
Update catalog database vie...
First data available through...
Final (statistical) catalog QA
Promote catalog database vi...
CSC 2.0 Release

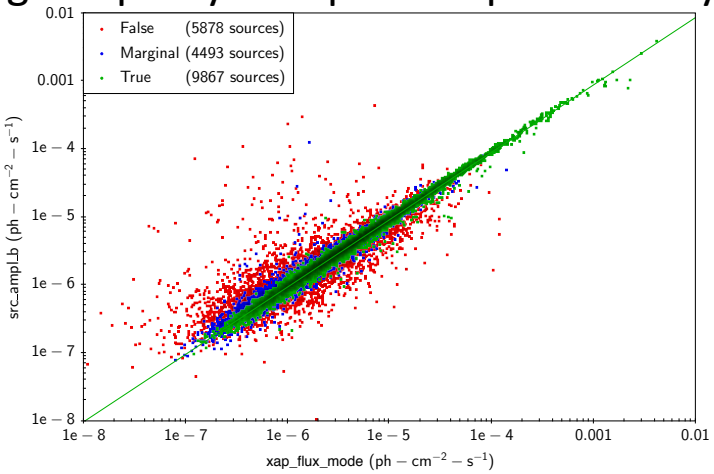
Why do we believe this projection

- The re-plan is based on extensive processing experience
  - ➔ We know how long it takes on average
- We have exposed and addressed remaining risks
  - Manual QA issues
  - ➔ team training and monitoring
  - Cross-match algorithm must be modified for deeper cohorts and denser data sets
  - ➔only 2% of data affected
  - ➔working on alternate approaches for these

DATA PRODUCTS

Preliminary Detection Lists

- A FITS table containing a subset of information for all of the compact detections with preliminary classification TRUE or MARGINAL
  - Detection likelihood and preliminary classification
  - Fitted position and 95% internal-error error ellipse
  - Fitted detection amplitude and confidence interval in 4 bands for ACIS (broad, soft, medium, hard) and 1 band for HRC-I (wide)
    - Amplitude is a good proxy for aperture photometry for TRUE point sources



  - Classification of detection as point or extended
  - Fitted ellipse parameters for extended detections
  - Documentation in the form of column descriptions and caveats is provided with the FITS data product

Release 2.0 Catalog Database Content

- Master Source Properties (X-ray sources on the sky)
  - Source name, position and position errors*, significance*, flags*, extent* (deconvolved), Bayesian block aperture photometry* (photon and energy fluxes, model fluxes), Bayesian block hardness ratios, Bayesian block spectral model fit properties, inter- and intra-observation temporal variability measures*, observation summary
- Stack Detection Properties (preliminary; X-ray detections)
  - Detection identification, position and position errors*, significance*, flags*, extent* (deconvolved), aperture photometry* (apertures, counts, count rates, photon and energy fluxes, model fluxes), stack identification, instrument information, processing information
- Per-Observation Detection Properties (X-ray detections)
  - Detection identification, position and position errors*, significance*, flags*, extent* (deconvolved), aperture photometry* (apertures, counts, count rates, photon and energy fluxes, model fluxes), hardness ratios, spectral model fit properties, intra-observation temporal variability measures*, observation identification, observation pointing, observation astrometry (aspect information), observation instrument configuration, processing information

  * Multiple energy bands

- Note: numerical properties include associated lower and upper confidence limits

Release 2.0 Catalog Data Products

FITS Catalog Data Products

- Per-Observation Full Field Data Products
  - Event list, exposure corrected image*, background image*, exposure map*, adaptively smoothed exposure map*, aspect solution (incl. fine astrometry updates), aspect histogram, bad pixel map, field of view, pixel mask, extended source region polygons* (multiple contour levels)
- Stack Full Field Data Products
  - Event list, exposure corrected image*, background image*, exposure map*, field of view, limiting sensitivity*, merged source detection list
- Per-Observation Source Region Data Products
  - Region definitions, region event list, region image*, local PSF* (~50K counts), region exposure map*, PHA spectrum, ARF, RMF, light curve*, position error MCMC draws*, aperture photometry PDF*
- Stack Source Region Data Products
  - Region definitions, region event list, region image*, region exposure map*, position error MCMC draws*
- Master Source Data Products
  - Bayesian block aperture photometry PDFs*, Bayesian block spectral fits, Bayesian block model fluxes*, Bayesian block hardness ratios, Bayesian block temporal properties*, master light curve*

---

* Multiple energy bands

Release 2.0 Data Access

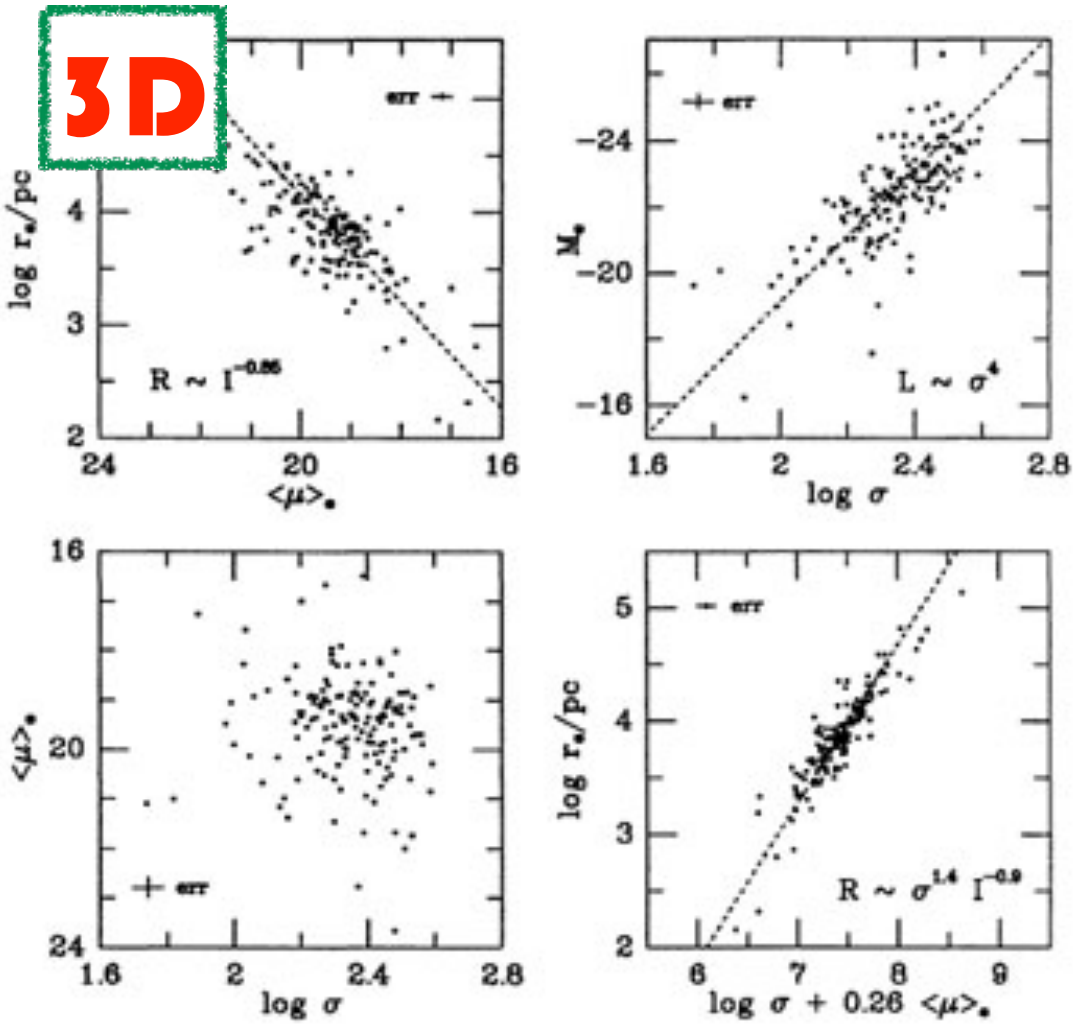
Release 2.0 Interfaces

- CSCview query interface
  - Enhanced to support release 2.0 content
  - Continued support for IVOA standards (including ADQL, SAMP, VOTable)
  - Limiting sensitivities will be reported directly by CSCview (4″×4″ HEALPIX)
  - Improved cross-match algorithm performance and input table capacity
  - Continued access to release 1.1 content (as a user selection)
  - Possible tighter integration with *Iris* (*Sherpa*/IVOA-standards based interactive SED and spectral fitting application) via SAMP
- Virtual Observatory (VO) interfaces
  - Support for Simple Cone Search (SCS), Simple Image Access Protocol (SIAP), and Table Access Protocol (TAP) data access standards
  - Same as release 1.1
- Command Line Interface (CLI)
  - Enhanced to support release 2.0 content, otherwise same as release 1.1
- *New* simple web interface
  - Interactive equivalent to the VO SCS, with results displayed in web browser
- CSC Sky
  - Same as release 1.1; will follow catalog release
- CSC/SDSS cross-match catalog
  - Same as release 1.1; will follow catalog release
  - Additional cross-match catalogs now possible more readily: `master_match` algorithm allows us to perform Bayesian catalog cross-matching *in house*

Catalog Science

- The SR panel emphasized the need to provide tools to the users to exploit the rich CSC data set, by itself and in combination with other multi-wavelength catalogs.
  - We plan to provide cross-match of CSC2 with other catalogs
  - We are investigating the use of Data Mining tools and how best to provide them to users

The emergence of multivariate Astronomy

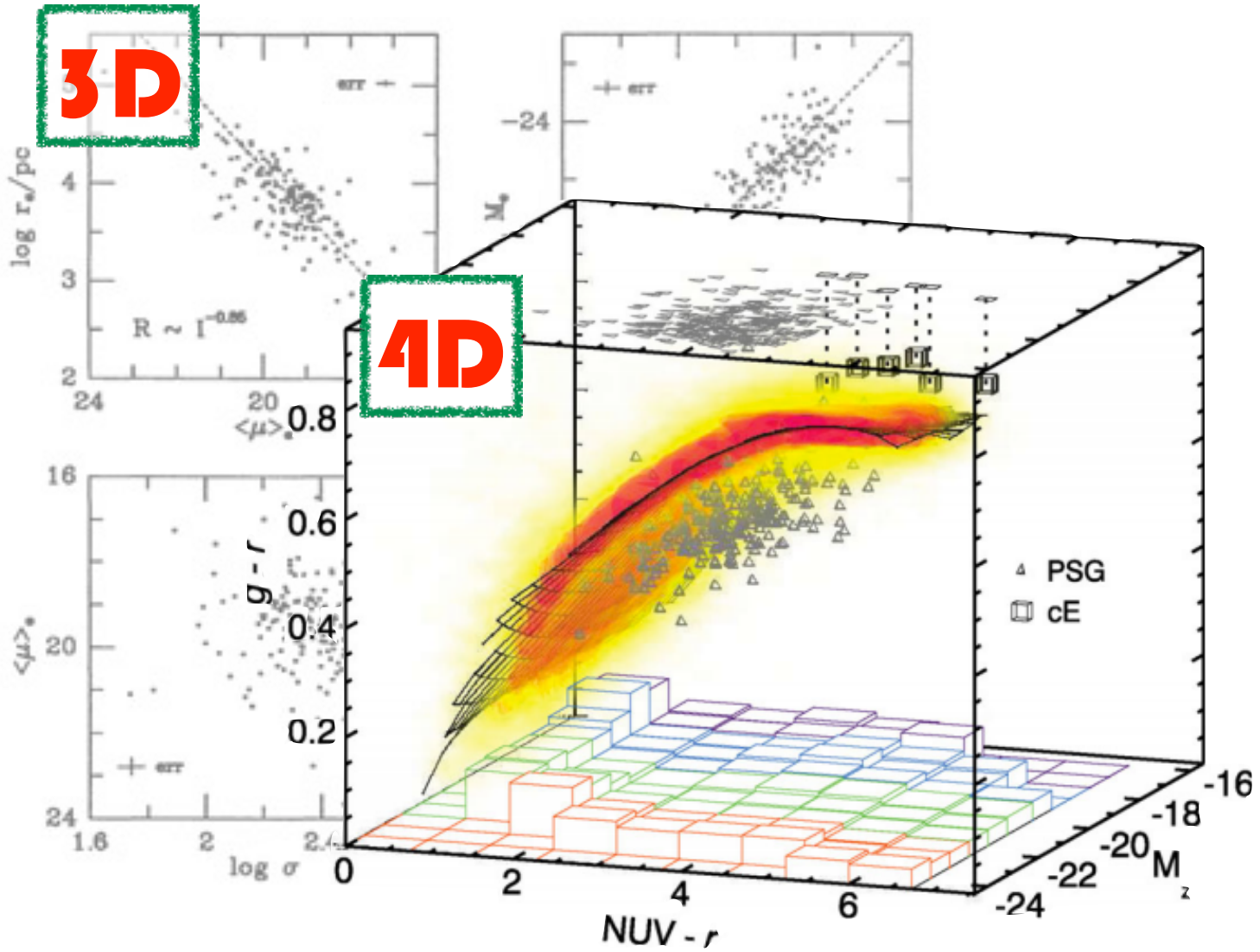


Fundamental plane of elliptical galaxies

(Djorgovski&Davis1987)

Raffaele D'Abrusco

The emergence of multivariate Astronomy

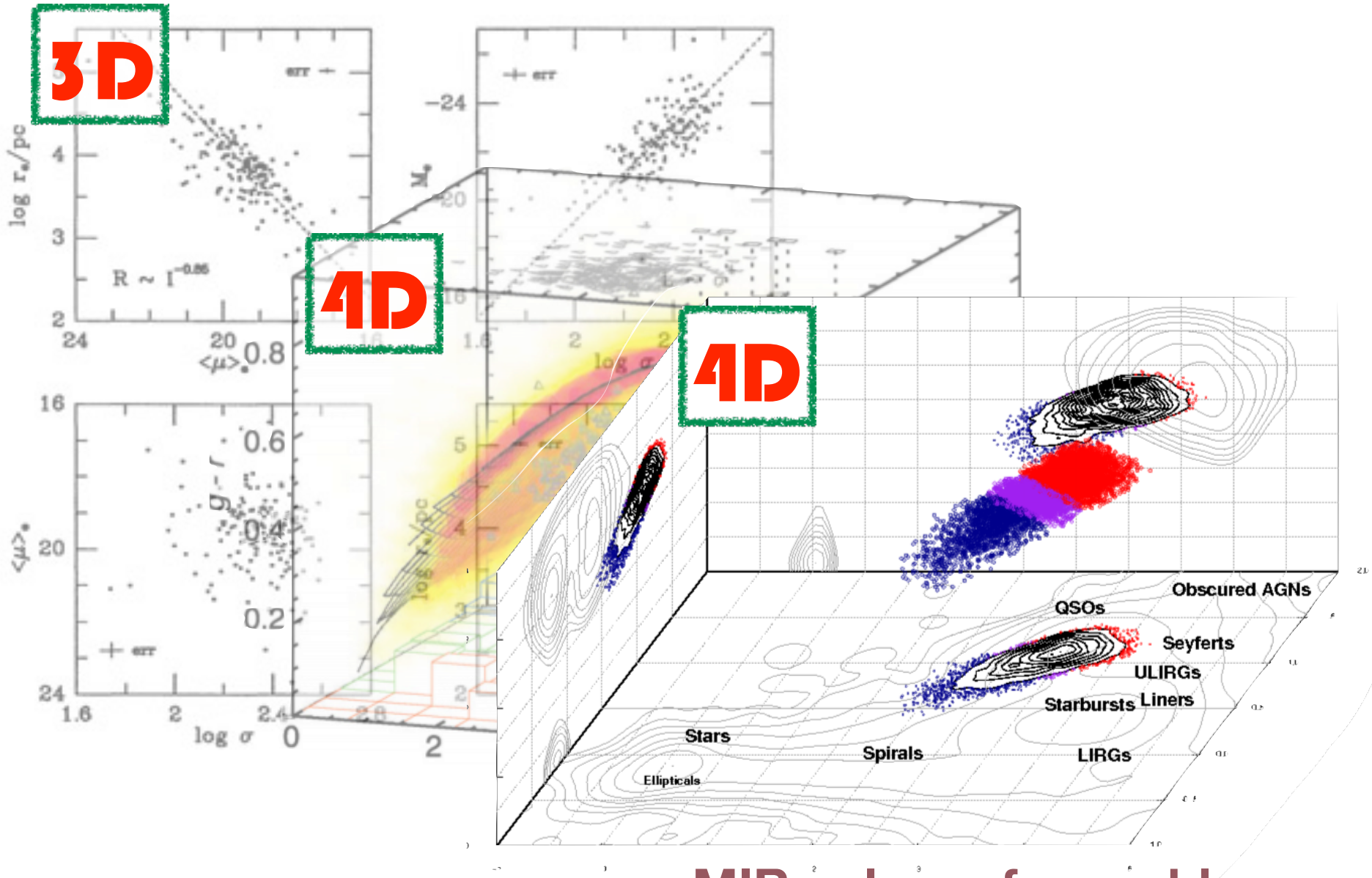


Optical-NUV locus of passive galaxies

(Chilingarian&Zolotukhin2012)

Raffaele D'Abrusco

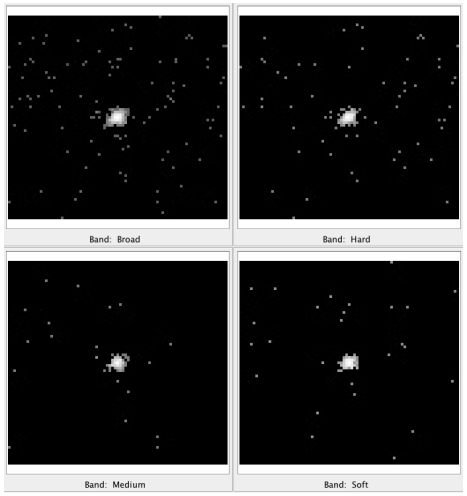
The emergence of multivariate Astronomy

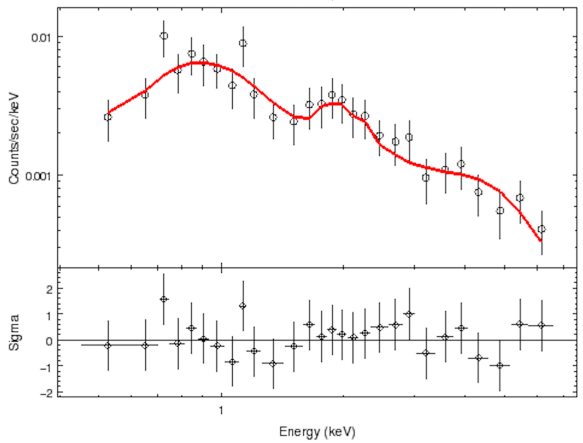


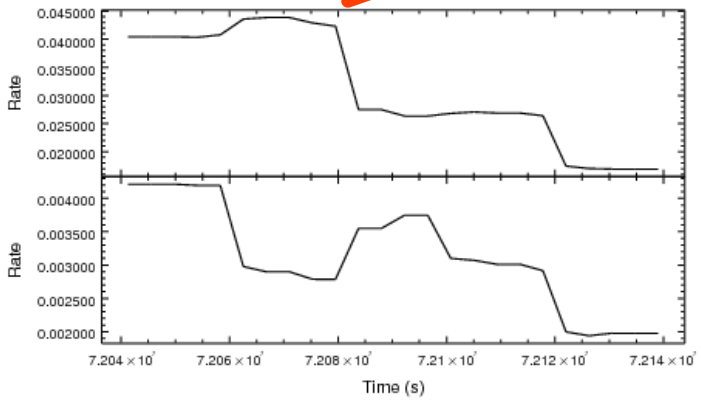
MIR colors of γ-ray blazars
(D'Abrusco+2012)

CHANDRA SOURCE CATALOG

170960	CXO J001300.5-270751	00 13 00.52	-27 07 51.57	0.62953	FALSE	FALSE	26.6051	1.2424E-13	1.19349E-13	1.29181E-1
177343	CXO J001351.5-302500	00 13 51.51	-30 25 00.26	0.320671	FALSE	FALSE	25.6305	2.11018E-13	2.04545E-13	2.17427E-1
82564	CXO J000136.1+130639	00 01 36.10	+13 06 39.50	0.309182	FALSE	FALSE	24.83	1.42236E-13	1.37323E-13	1.47199E-1

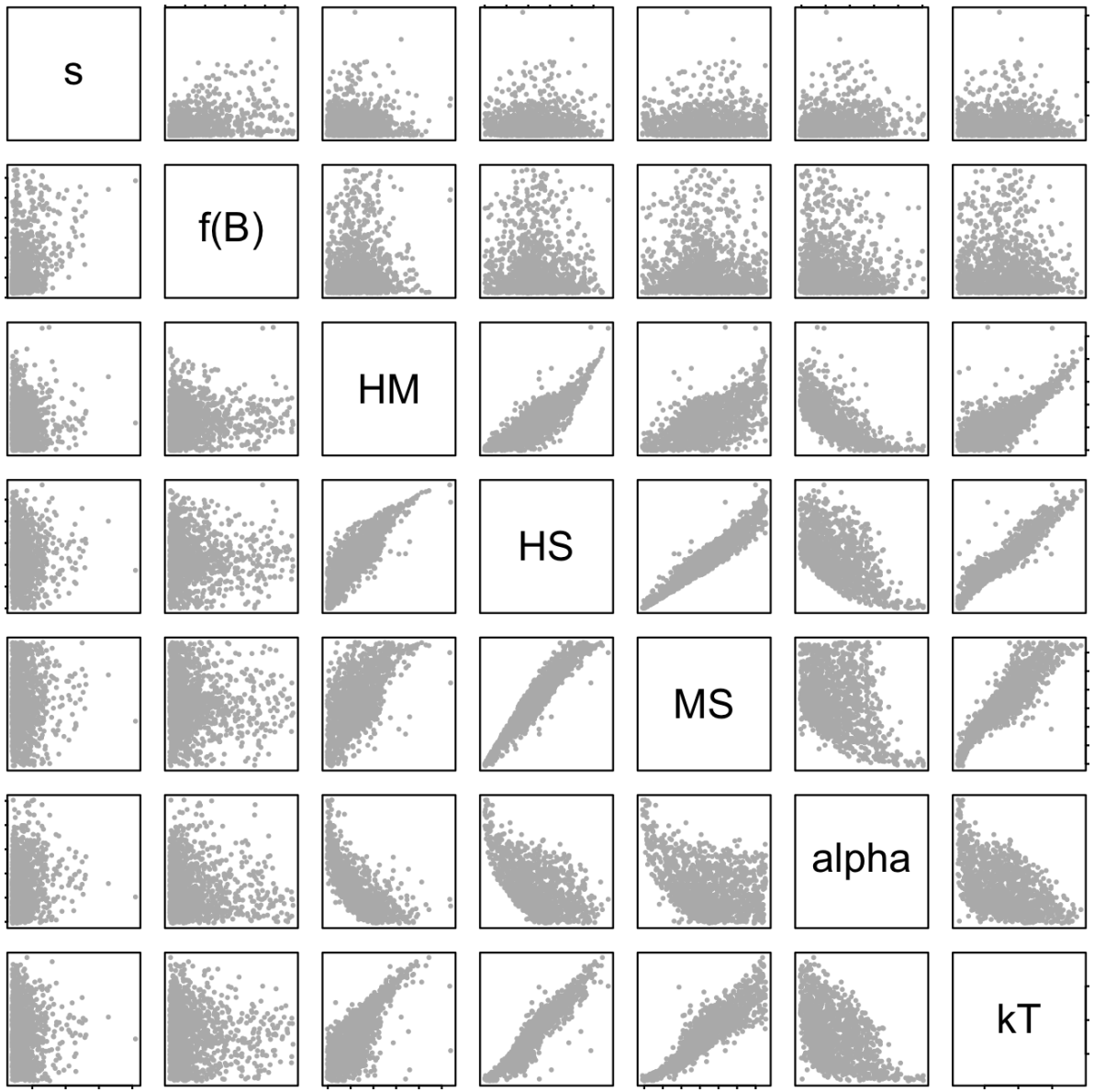






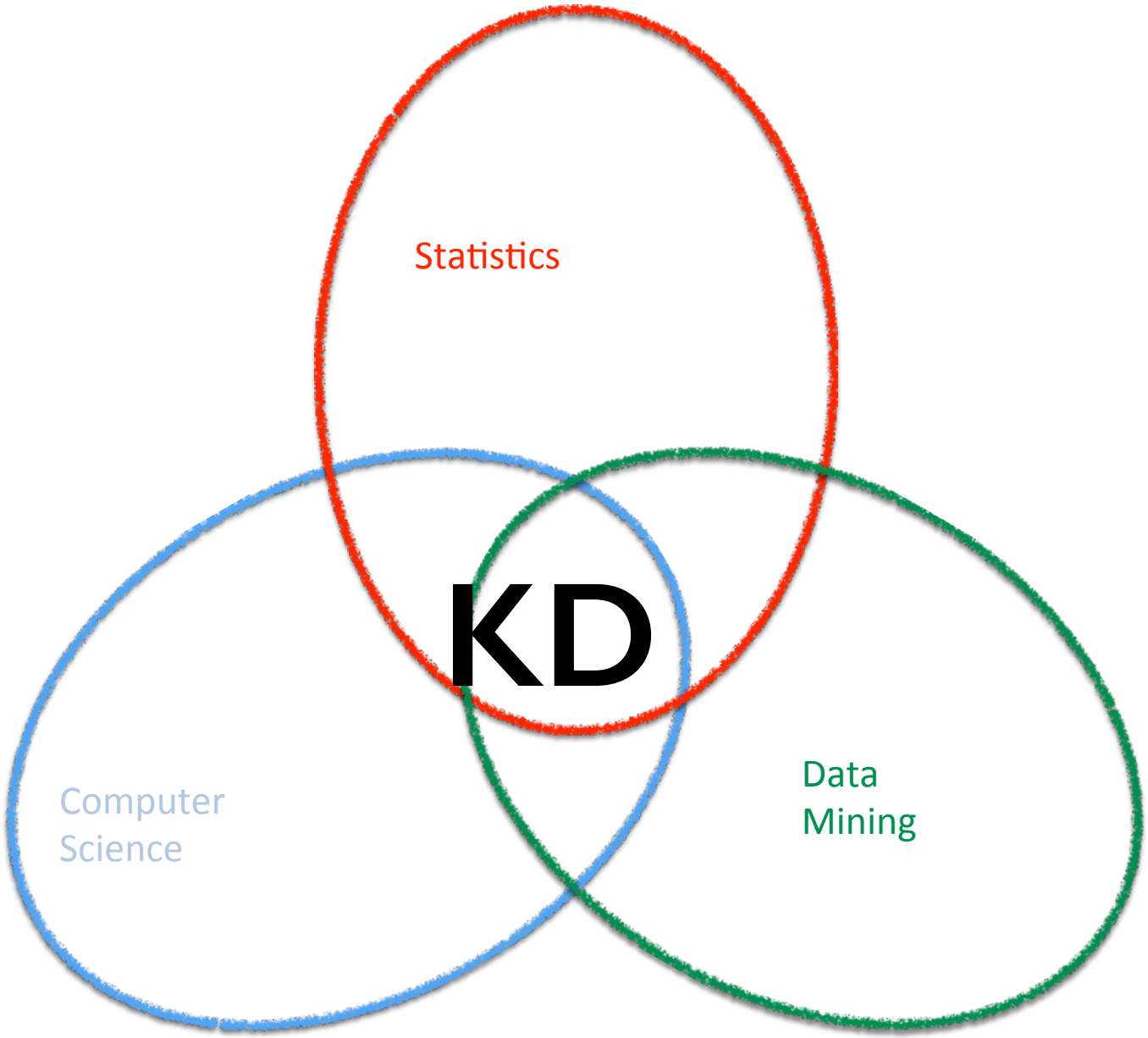
Raffaele D'Abrusco

Exploring High-Dimensional Data



Raffaele D'Abrusco

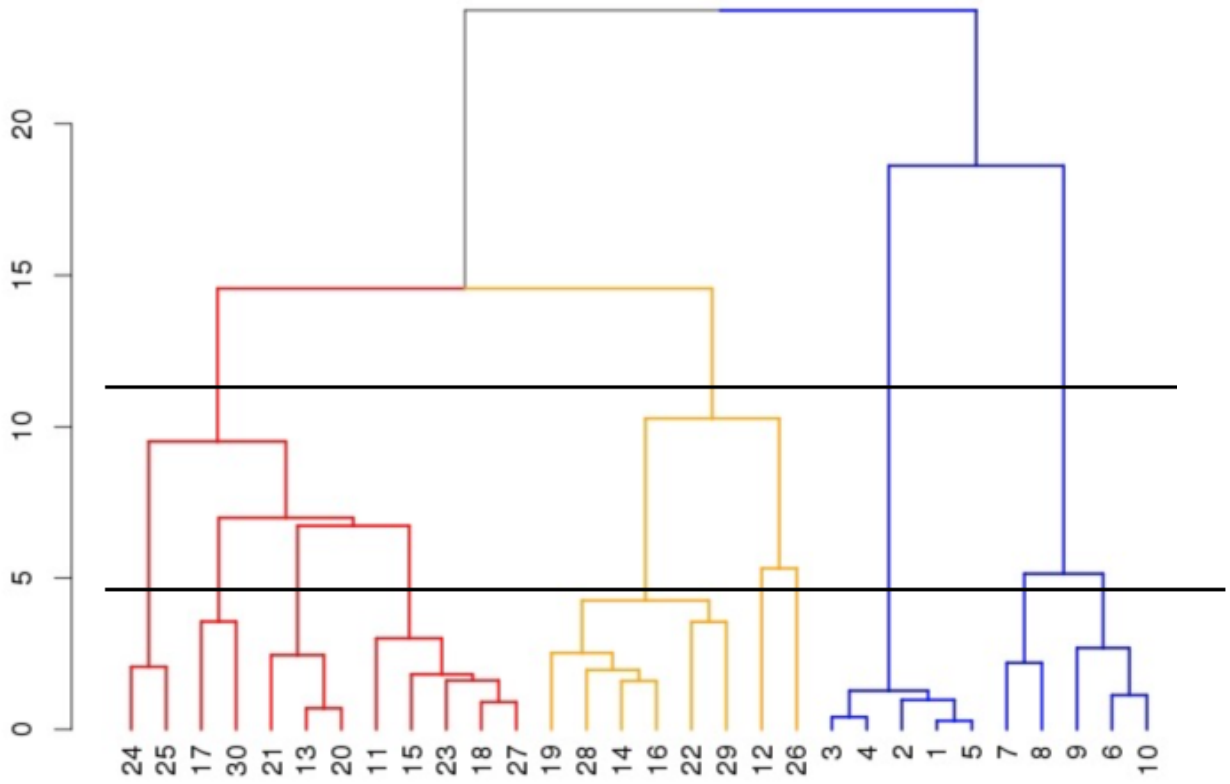
A modern Toolbox: Knowledge Discovery

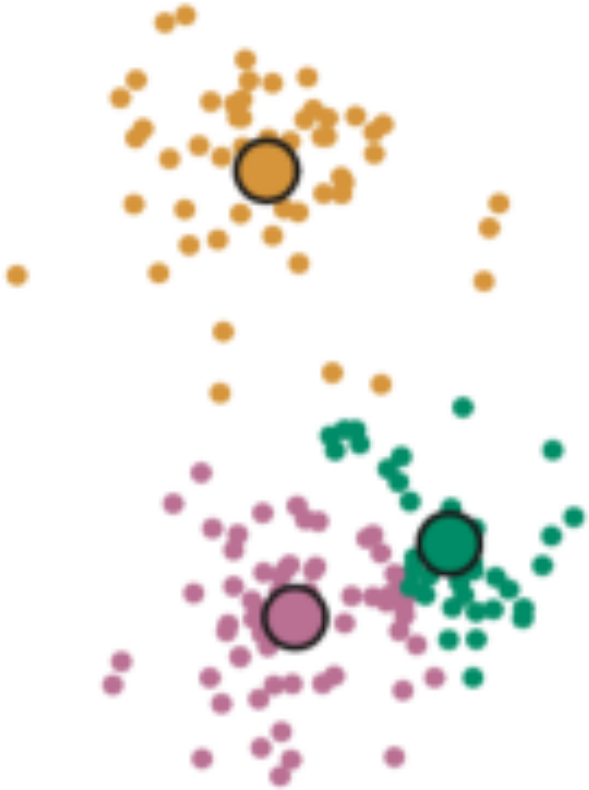


Raffaele D'Abrusco

Learning about the data structure

Finding spontaneous aggregations of source in the *feature* space, the **clusters**

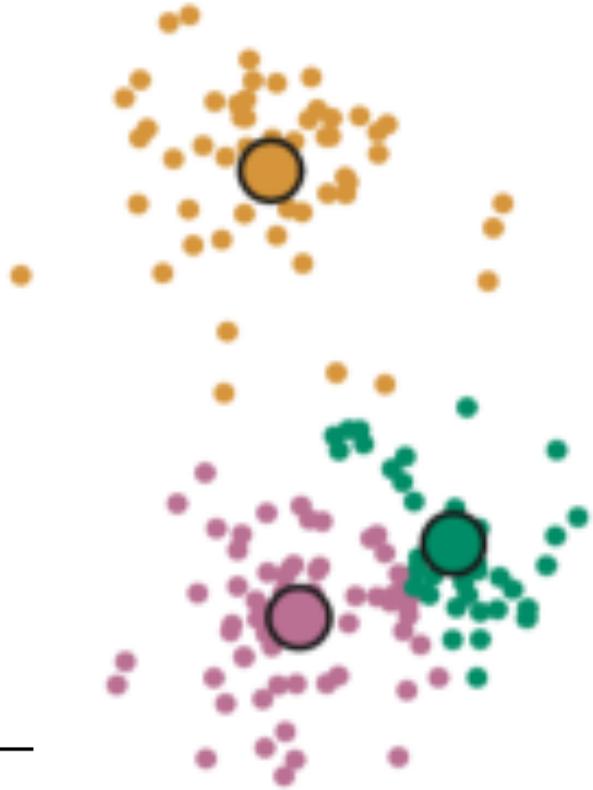


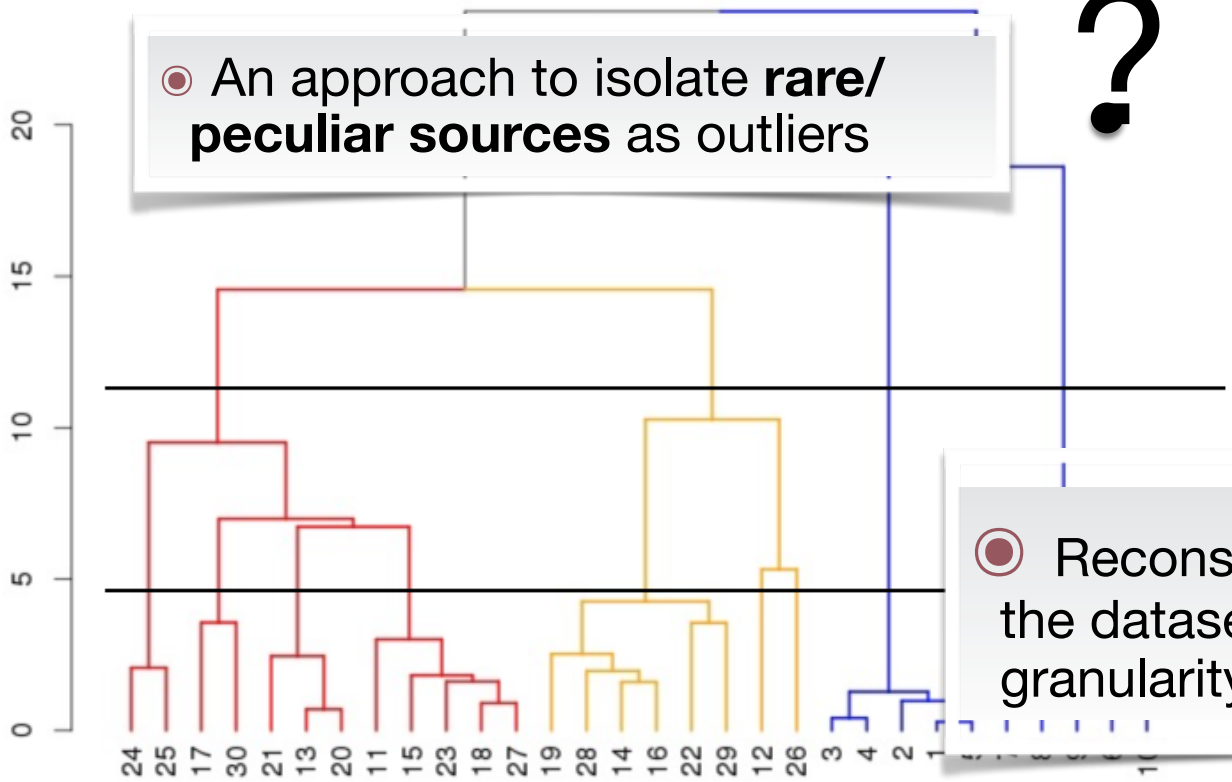


Raffaele D'Abrusco

Learning about the data structure

Finding spontaneous aggregations of source in the *feature* space, the **clusters**





?

- An approach to isolate **rare/ peculiar sources** as outliers

- Reconstructing the **structure** of the dataset at different levels of granularity (i.e. signal to noise)

Raffaele D'Abrusco

Clustering

Classification

Raffaele D’Abrusco

Searching for known sources

- **Ensembles of classifiers** and **priors** improves accuracy with negligible decrease in flexibility

- **Probabilistic classification** can reflect degenerate cases or transitional classes of sources

Raffaele D'Abrusco

Clustering

Classification

Raffaele D’Abrusco

In Summary

- The SR has recognized the CSC as Chandra Legacy Product
  - The CXC tasked to have a path for ongoing Catalog processing updates with releases every 2-3 year and final release after end of mission
  - The CXC tasked with expanding user tool set to include CSC mining tools in a multi-wavelength setting
  - The CXC will review progress with the SR in 2018
- The CXC has done a thorough internal review of catalog status
  - Replan
  - Parallel activities
  - Preliminary data releases
  - Final CSC2 release in 10 mo.
- The CXC will transfer CSC pipelines (L3) to pipeline operations for future releases
  - The Catalog will be another routine data production activity
  - It will follow the well established software and release cycle of other CXC pipelines
- The CXC has begun investigating DM tools and will value CUC advise