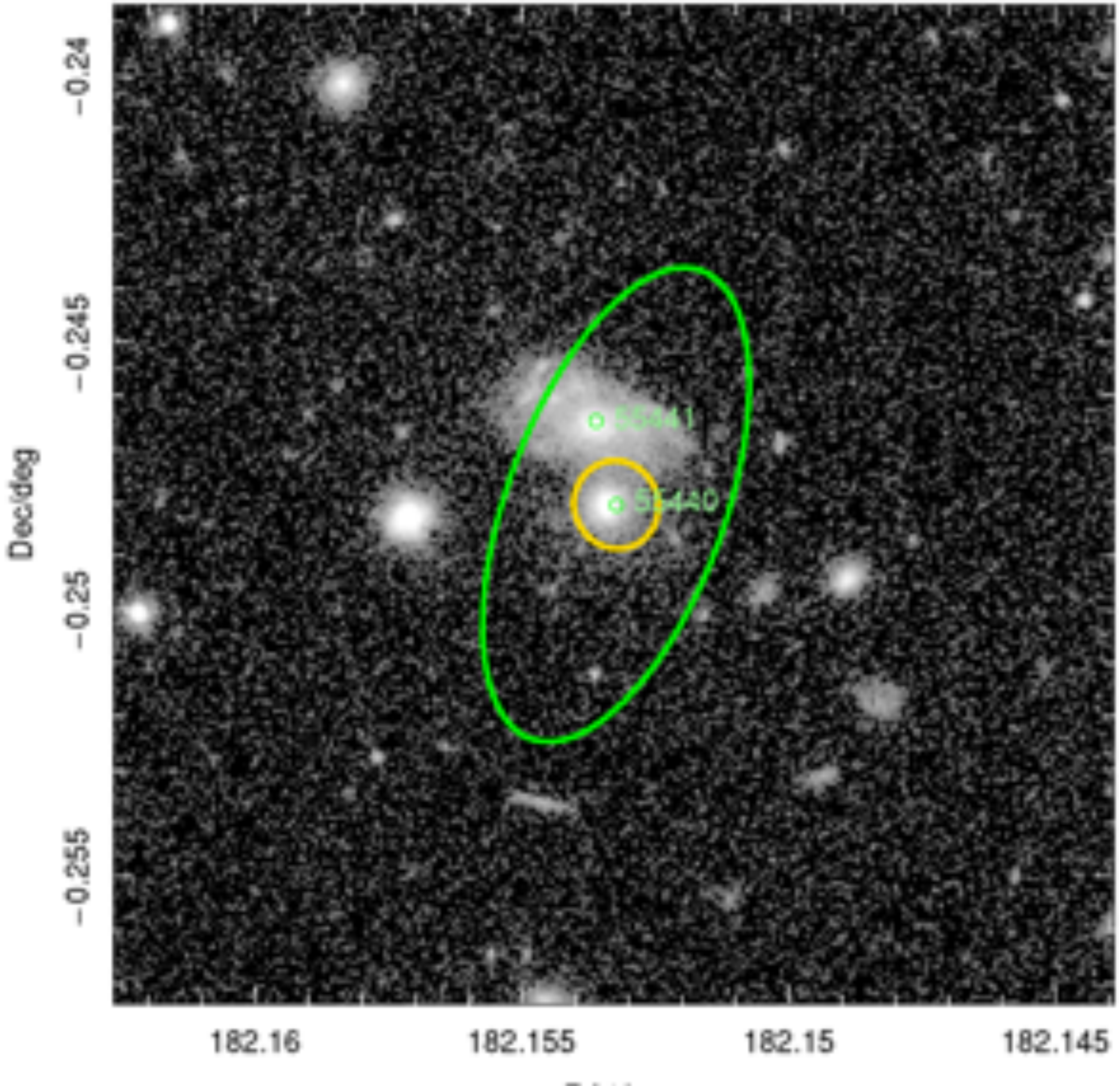
FORM FREE SOURCE EXTRACTION FOR RADIO ASTRONOMY

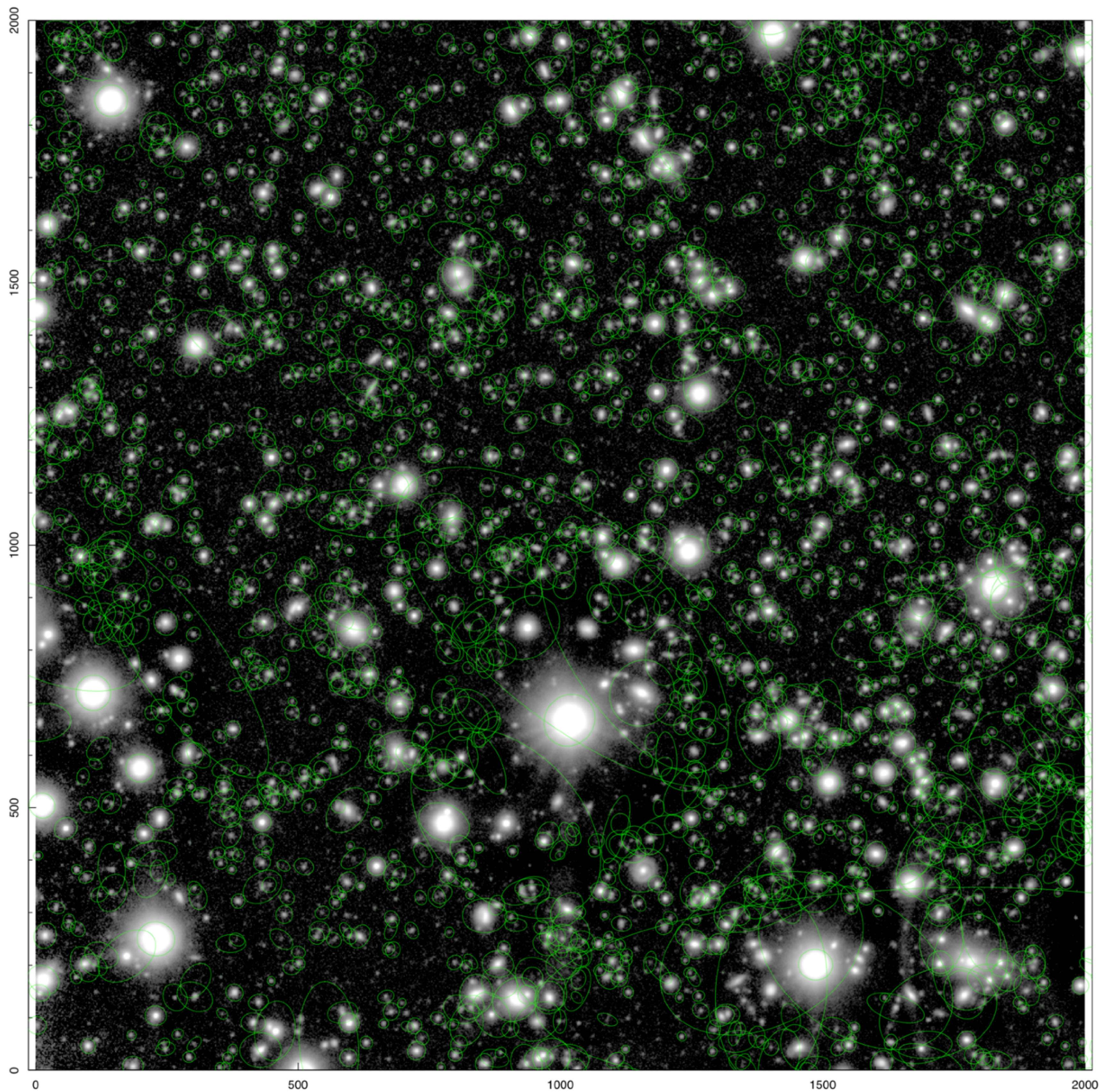
PROFOUND RADIO EXTRACTION

Robotham et al 2018 (GitHub: asgr/ProFound)

GEOMETRY FORCED (E.G. ELLIPTICAL) APERTURES OFTEN FAIL



THE PROBLEM WITH OUR CURRENT (SEXTRACTOR) APERTURES

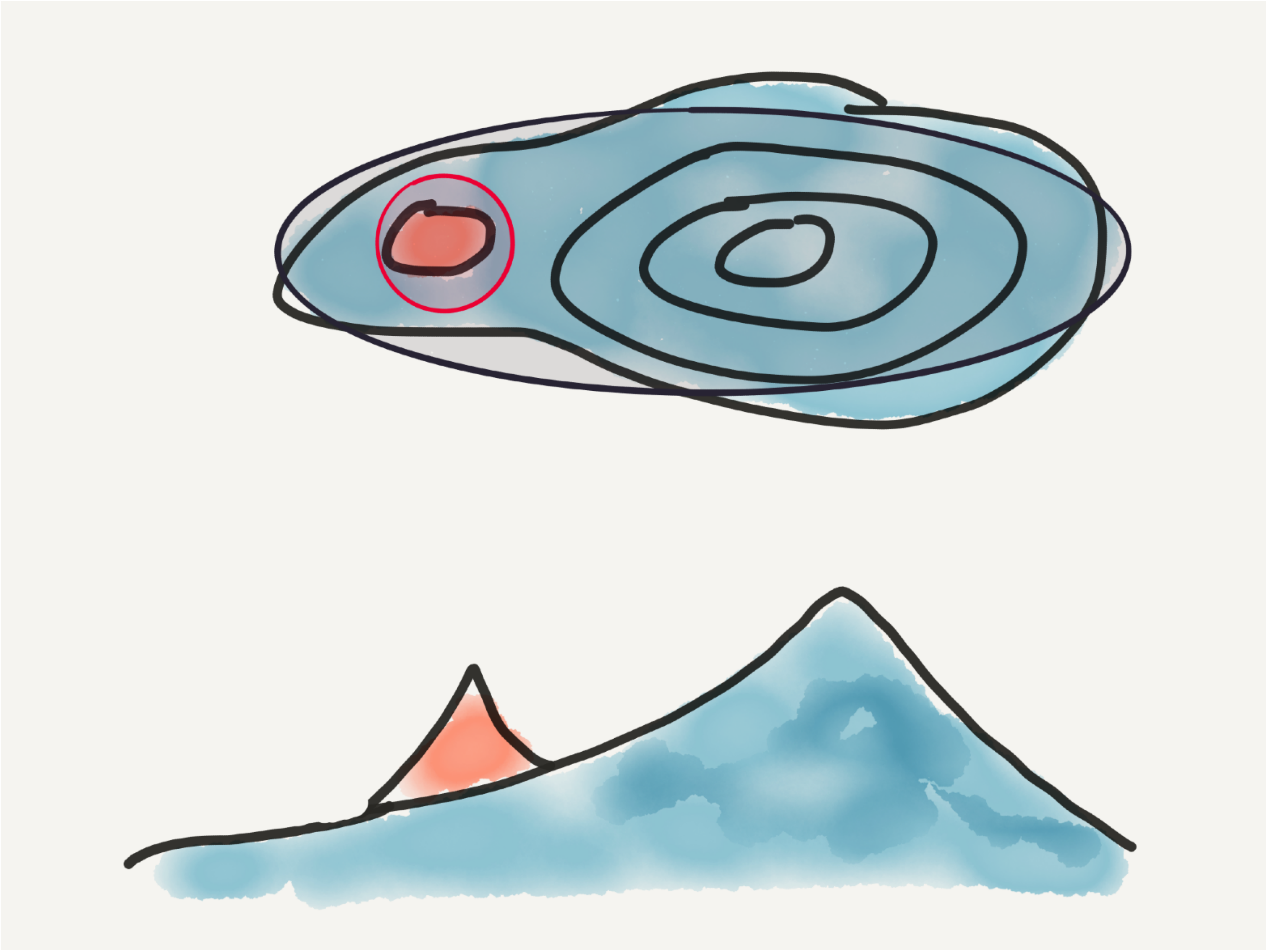


Ultra-VISTA Y-band stack + final “best” release catalogue (~2016)

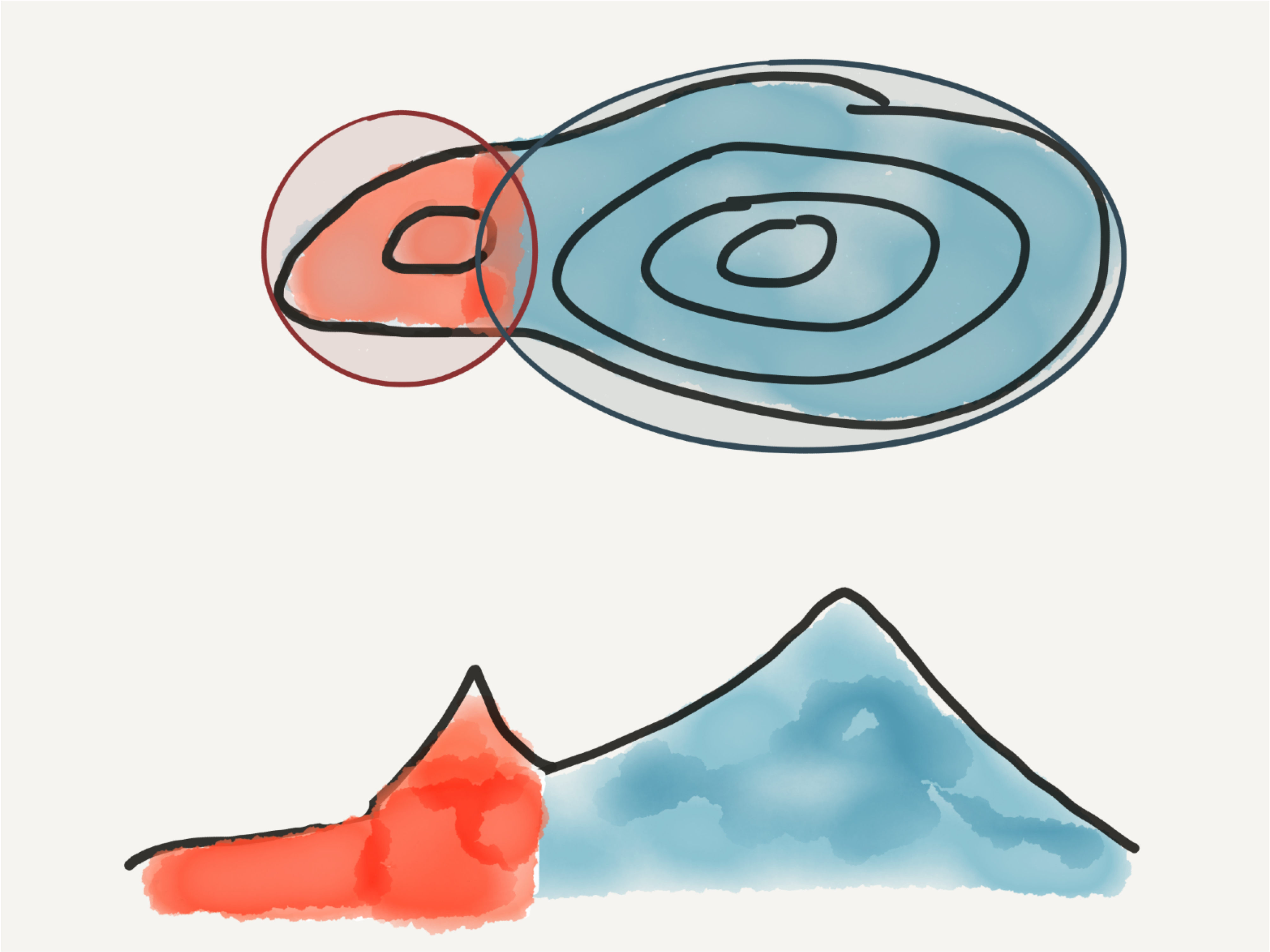
THE NEED TO START AGAIN

- Next generation surveys require high quality input catalogues, and produce too many sources to fix "by hand".
- In short, I started again with the source extraction.
- It was not obvious what improvements might be possible over SExtractor (given how well tested and established it is) but two areas quickly came to light:
  - It does not watershed de-blend optimally (the most common failure we see is due to this).
  - It uses strictly elliptical apertures and then tries to distribute overlapping flux using a number of internal schemes.

SEXTRACTOR TENDS TO CREATE WATERSHED ISLANDS



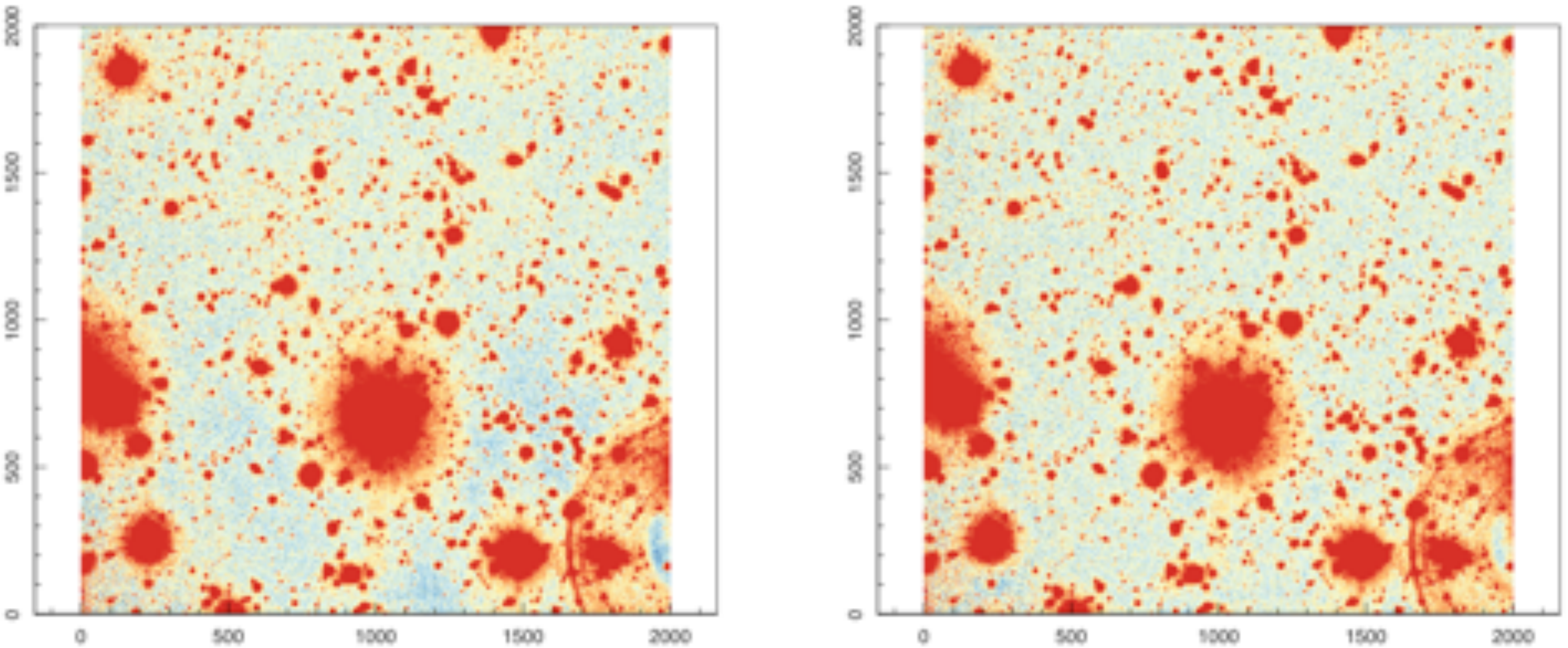
PROFOUND WATERSHEDS THROUGH SADDLE CONTOURS



PROFOUND GETS ROUND THESE ISSUES

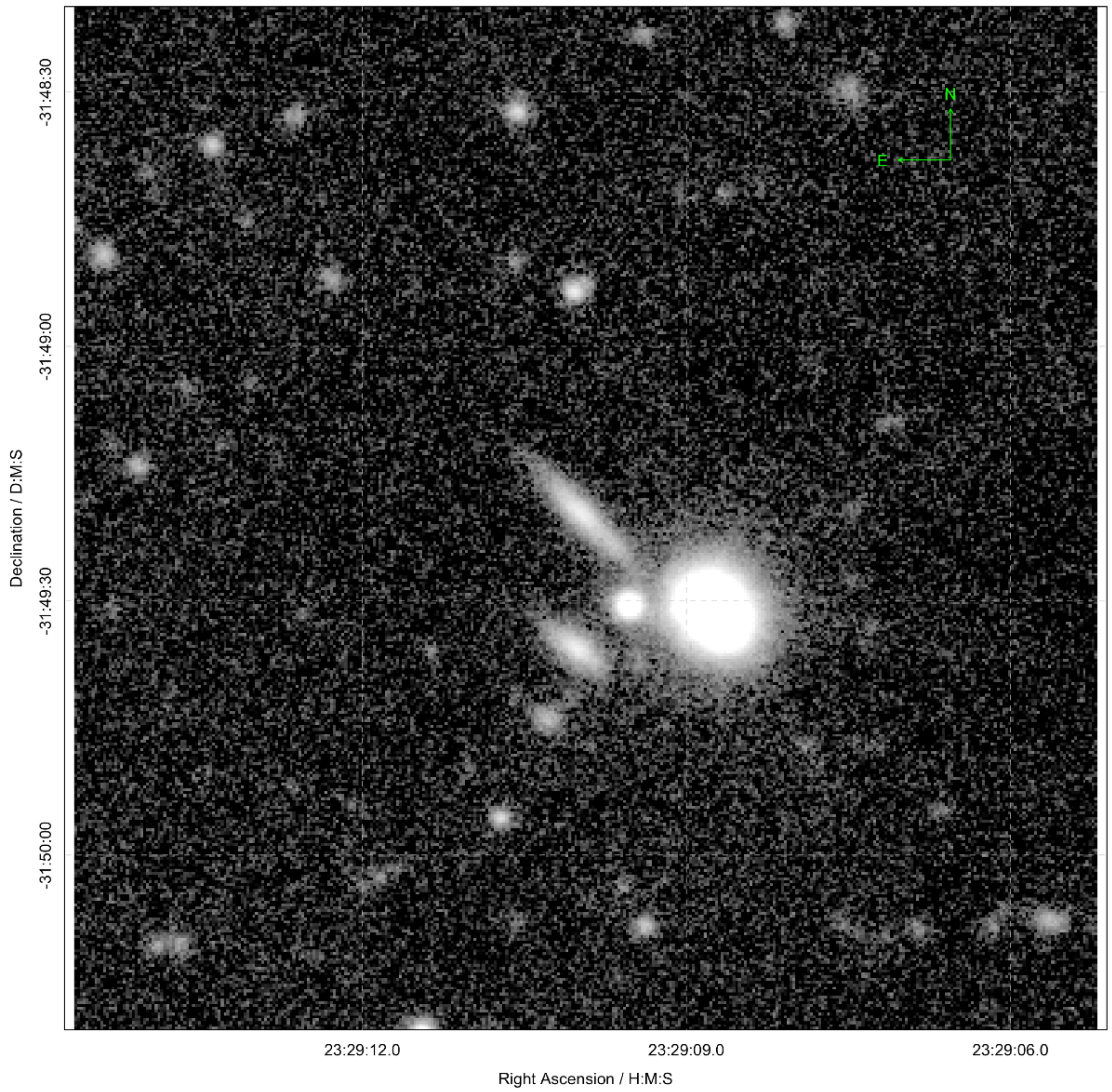
- We use a similar approach to find the initial high S/N image segments:
  - Careful sky subtraction (iterative masking and clipping)
  - Find seed pixel complexes after image filtering.
  - Segments are de-blended to some tolerance (using a different algorithm to SExtractor- non-discretised surface brightness / sky-RMS thresholds and locally adaptive).
  - Segments are grown organically- apertures never used.

PROFOUND USES AN AGGRESSIVE MESH BASED SCHEME

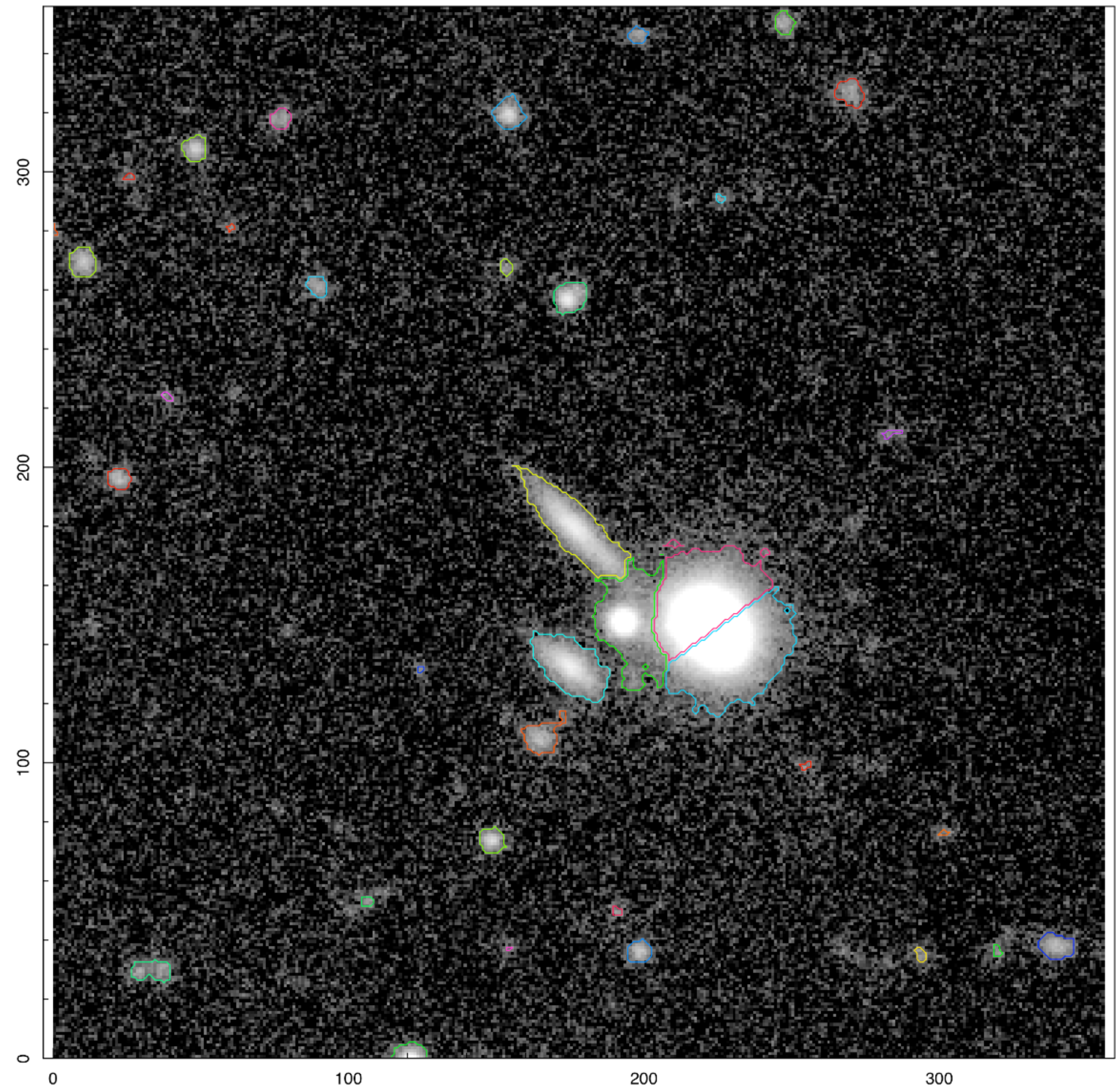


Ultra-VISTA sky versus ProFound sky

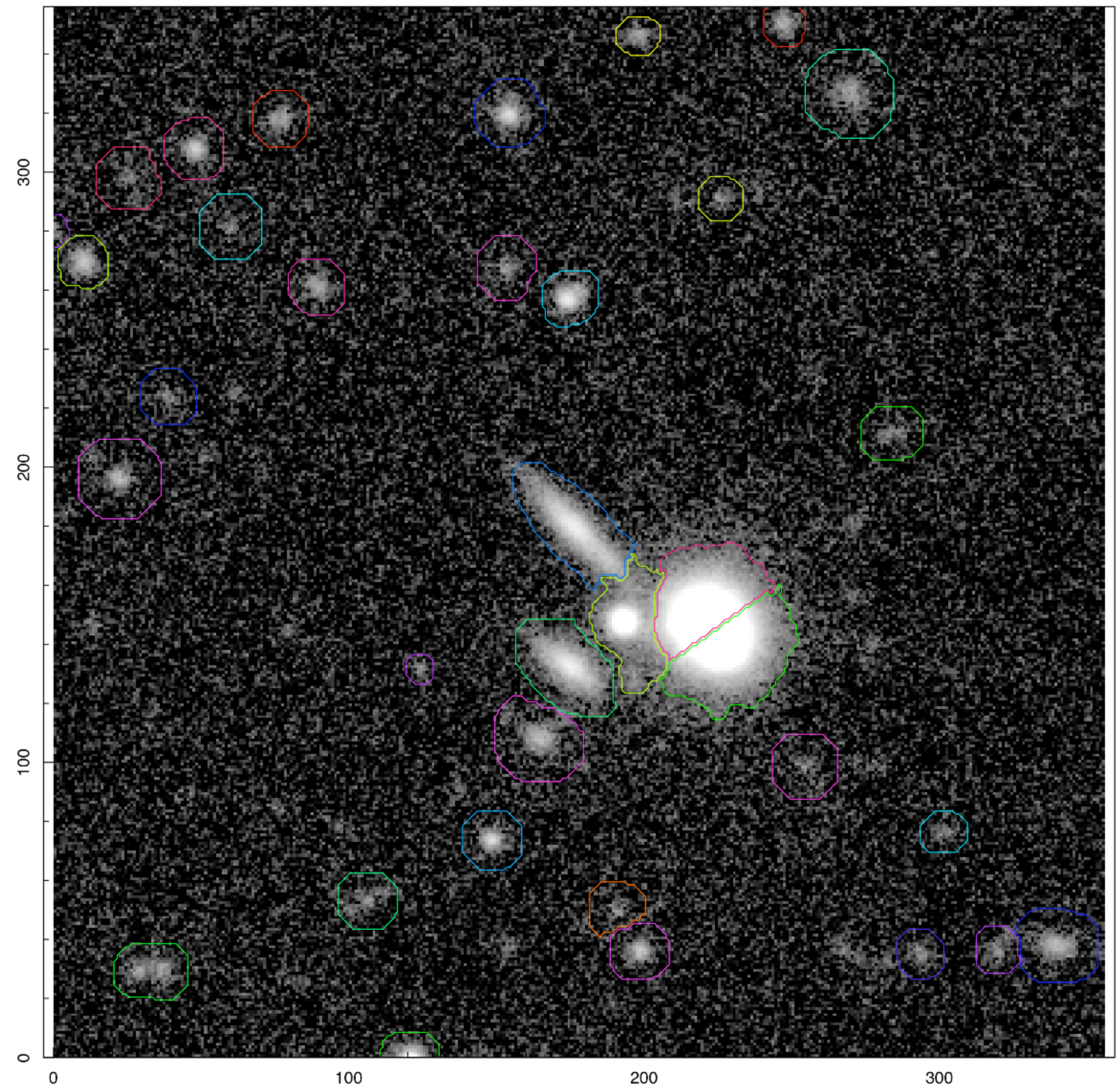
INITIAL VIKING Z-BAND IMAGE



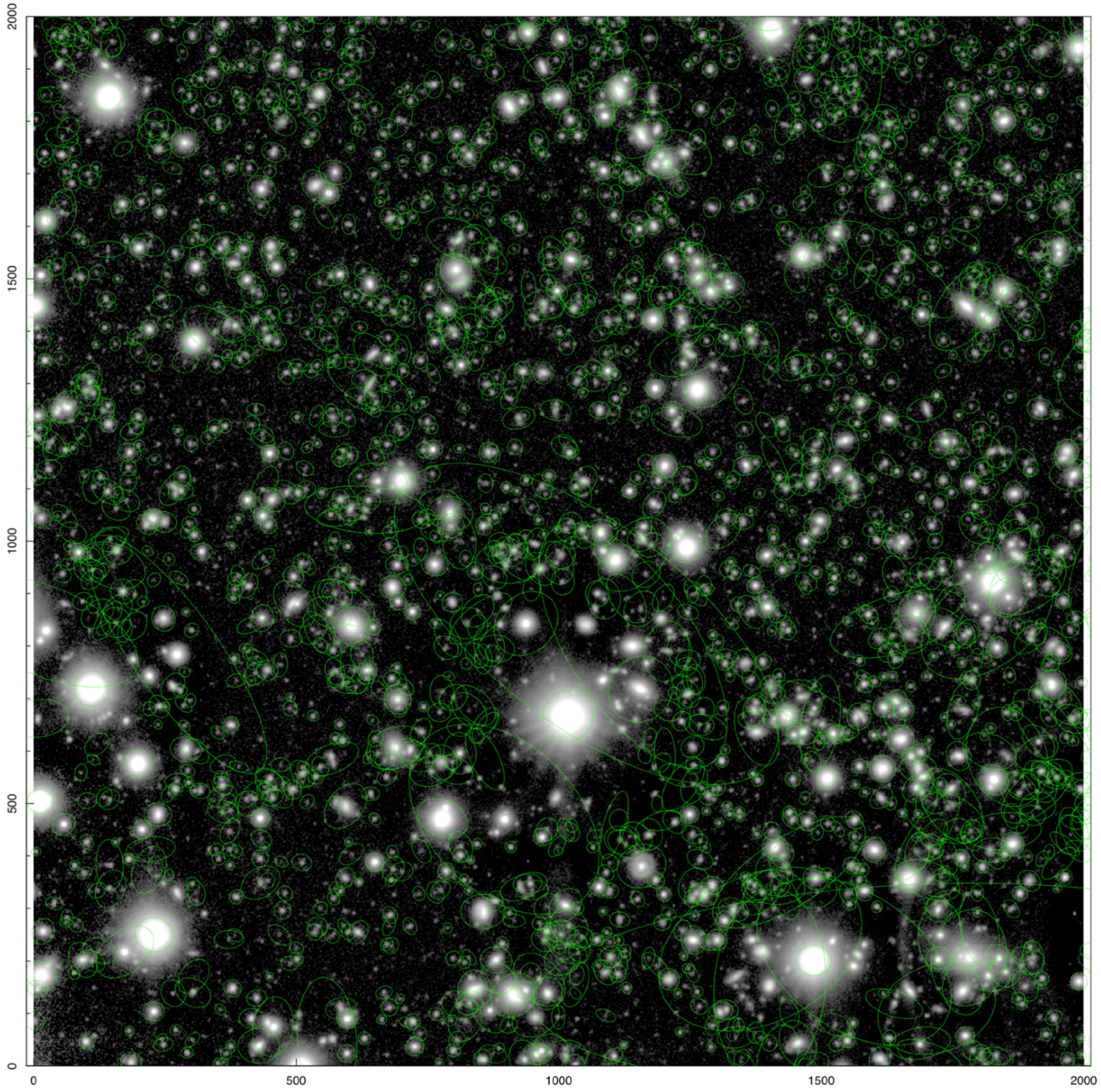
BRIGHT SEGMENT SEED COMPLEXES IDENTIFIED AND DE-BLENDED



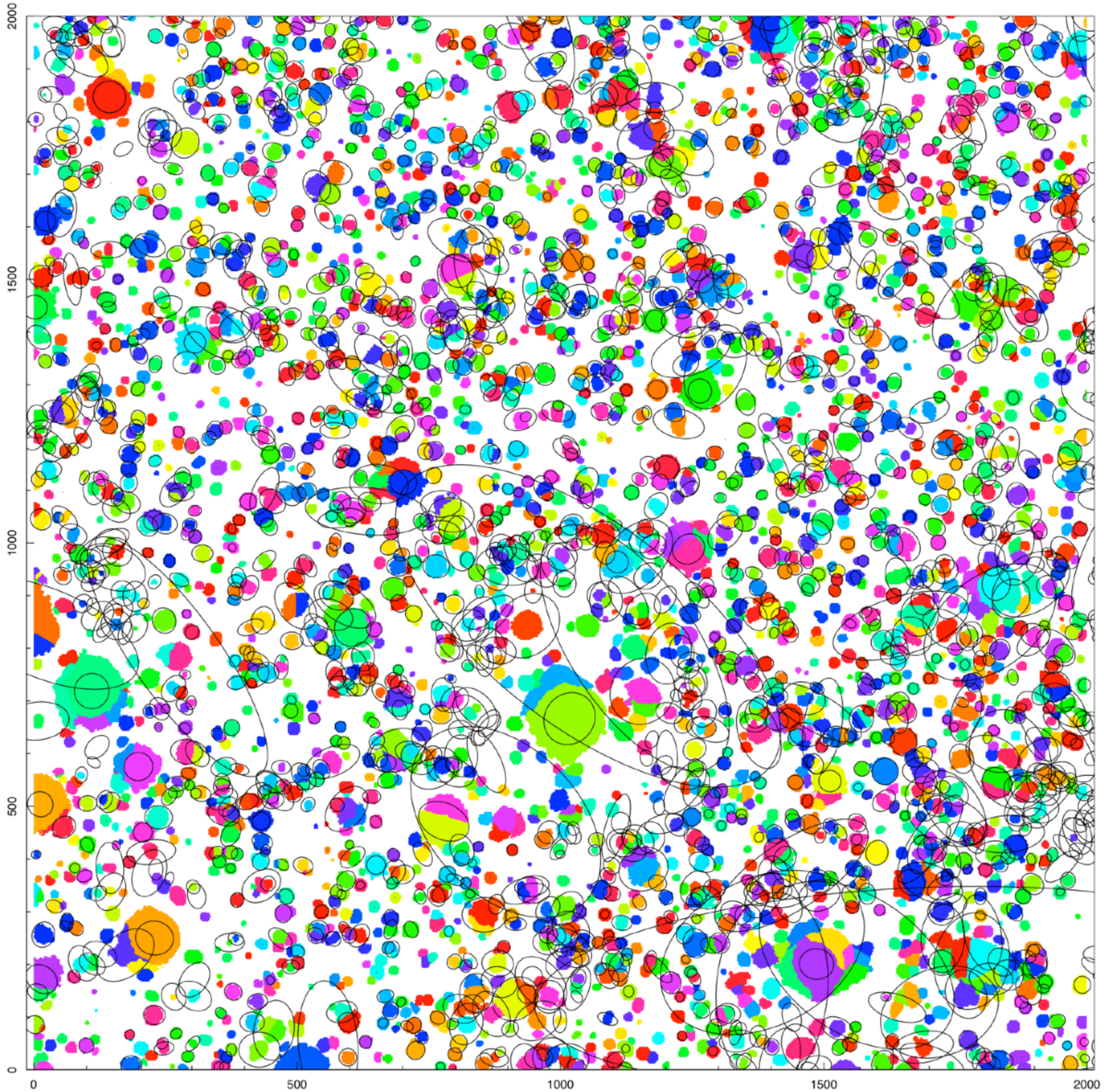
SEGMENTS DILATED UNTIL THE FLUX CONTAINED CONVERGES



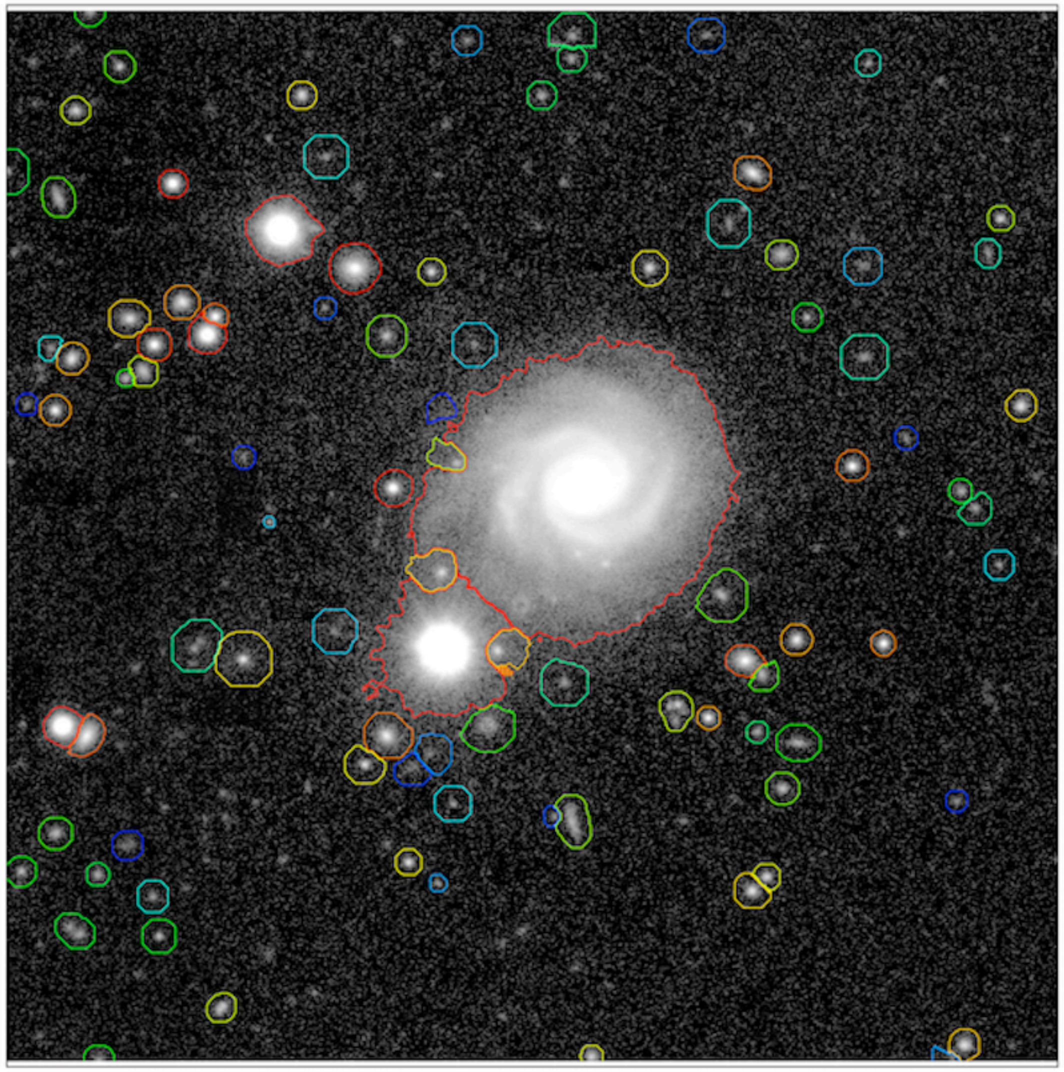
THE BEST TOTAL PHOTOMETRY HAS THE USUAL ISSUES IT SEEMS:

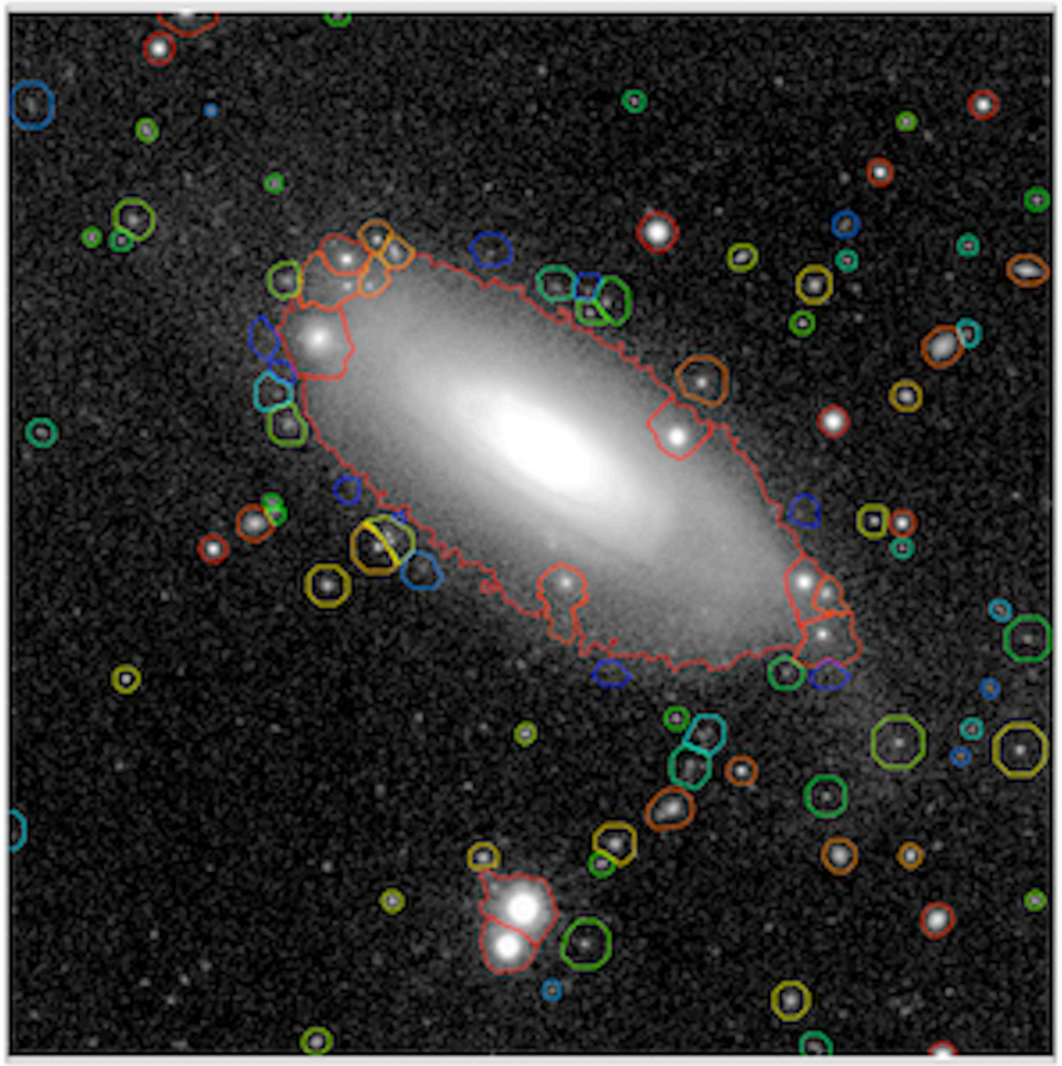


PROFOUND FIXES THE VERY SERIOUS ISSUES NICELY



NEW DE-BLENDER ALSO WORKS WELL ON RESOLVED SOURCES





- This is important for more general classes of problem, where we cannot guess the geometry in advance- i.e. the Universe is not full of smooth elliptical things. E.g. continuum image radio jets etc.

ProFound: Application to Radio Data

Work done by Catherine Hale (arxiv-1902.01440)

Current Radio Source Detectors

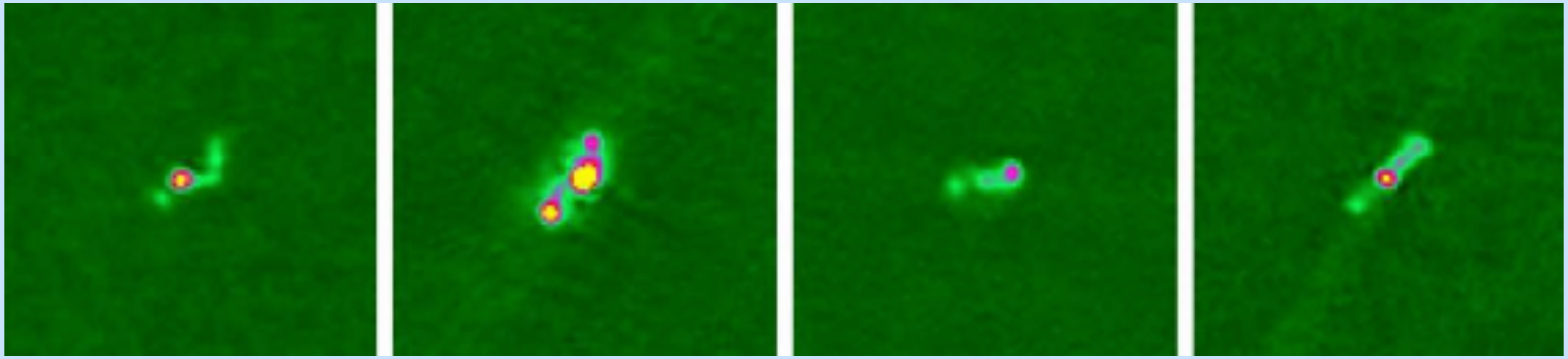
Hale et al 2019, arxiv-1902.01440

- PyBDSF (Mohan+ 2015)
  - Used in e.g. Shimwell+ 2017, Intema+ 2017
- BLOBCAT (Hales+ 2012)
  - Used in e.g. Smolcic+ 2017
- AEGEAN (Hancock+ 2012,2018)
  - Used in e.g. Hurley-Walker+ 2016

Rely on finding bright pixels above the sky and then fitting Gaussian components

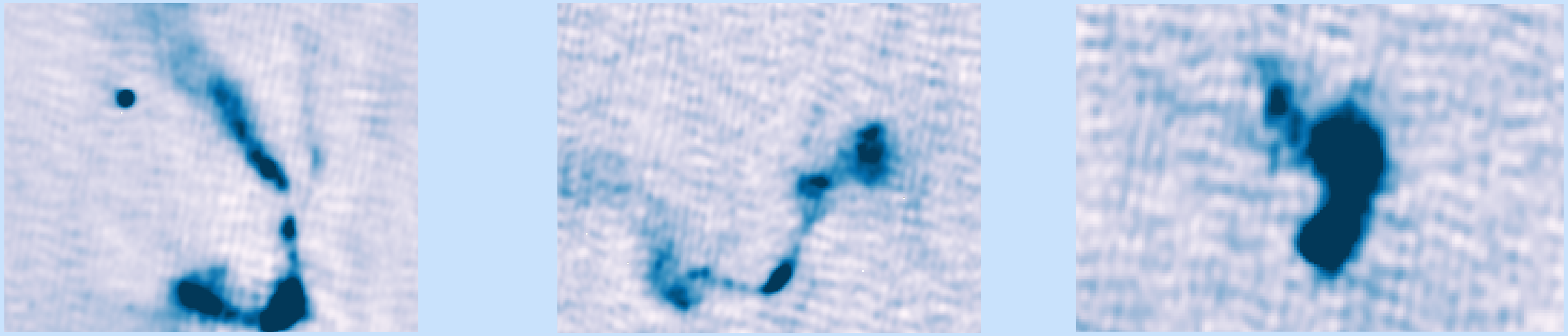
But

VLA
1.4 GHz



Hodge+ 2011

LOFAR
150 MHz



Hale+ submitted

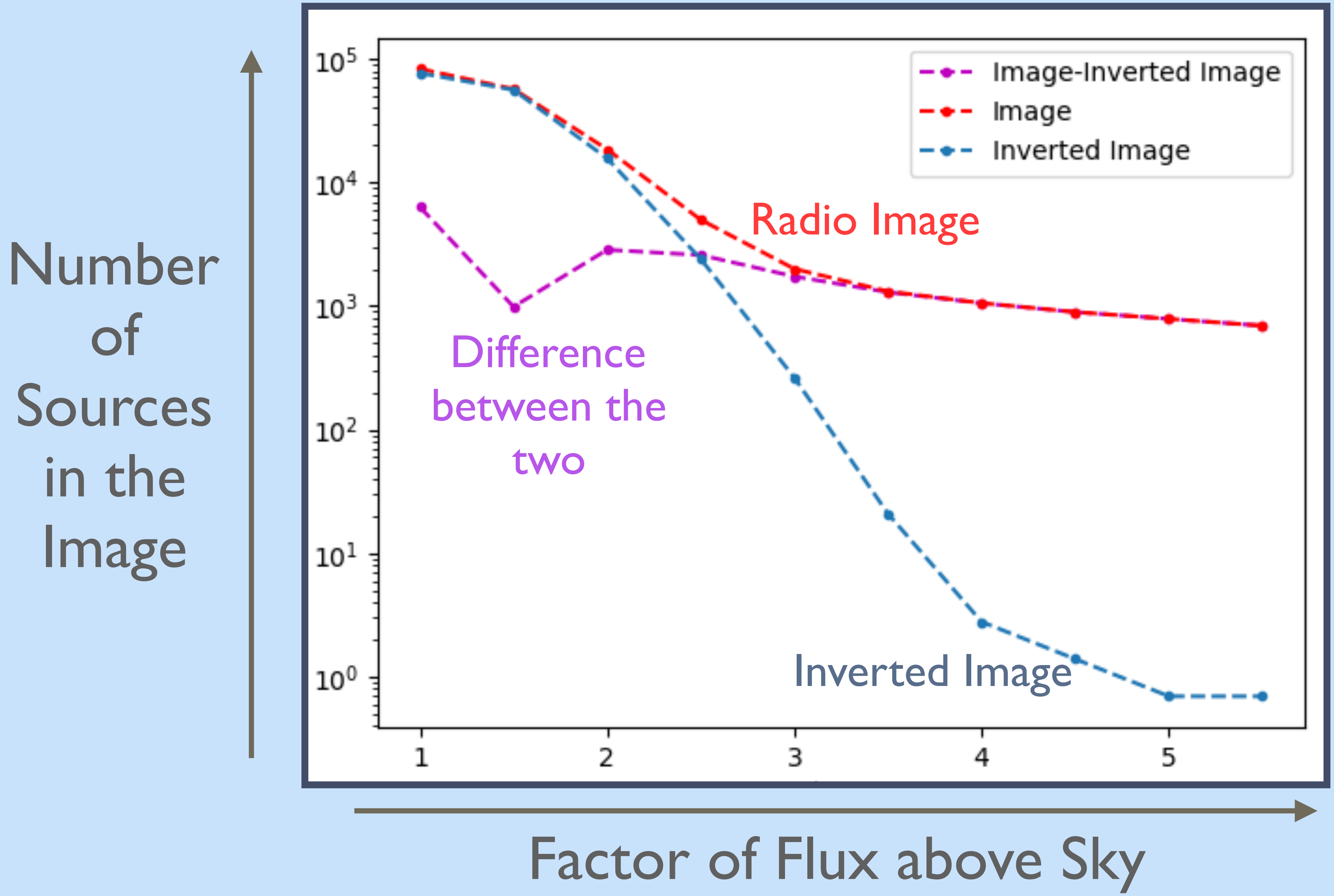
ProFound (Robotham+ 2018)

- It has been used in many **optical/IR studies** (e.g Davies+ 2018, Robotham+ 2018, Turner+ in prep)
- Uses pixel based extraction of fluxes (similar to SExtractor)
- Does not assume any source morphology
- We want to investigate whether it can be used as a **radio source detector**
- Other benefit to ProFound
  - Multi-wavelength capabilities

Possible caveats

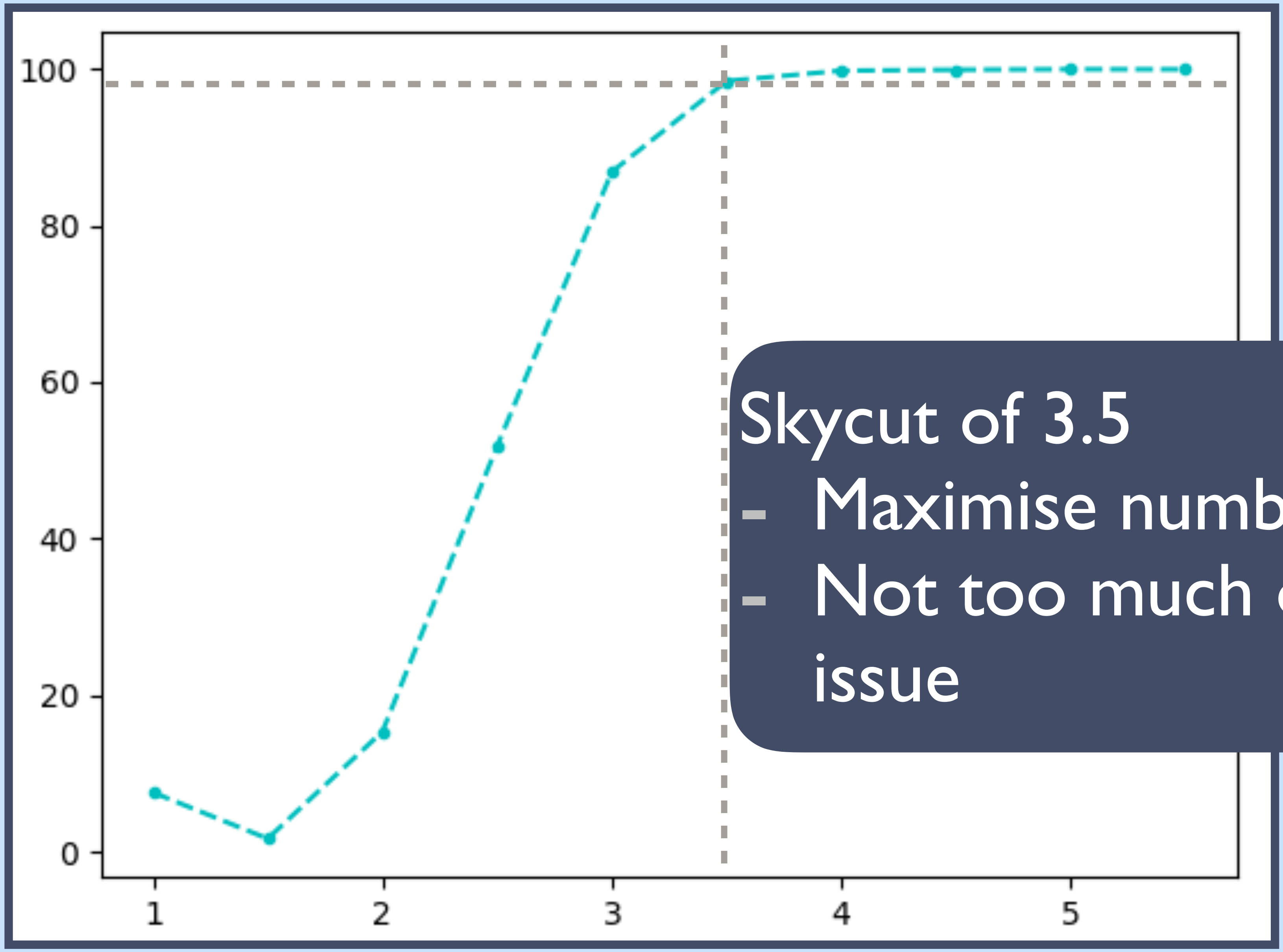
- Noise in radio surveys is **correlated**
  - More likely to detect noise as sources in optical/IR
  - Will need to use higher thresholds → Use false detection rate analysis to work out your threshold

False Detection Rate



False Detection Rate

Percentage of Real Detections



Factor of Flux above Sky

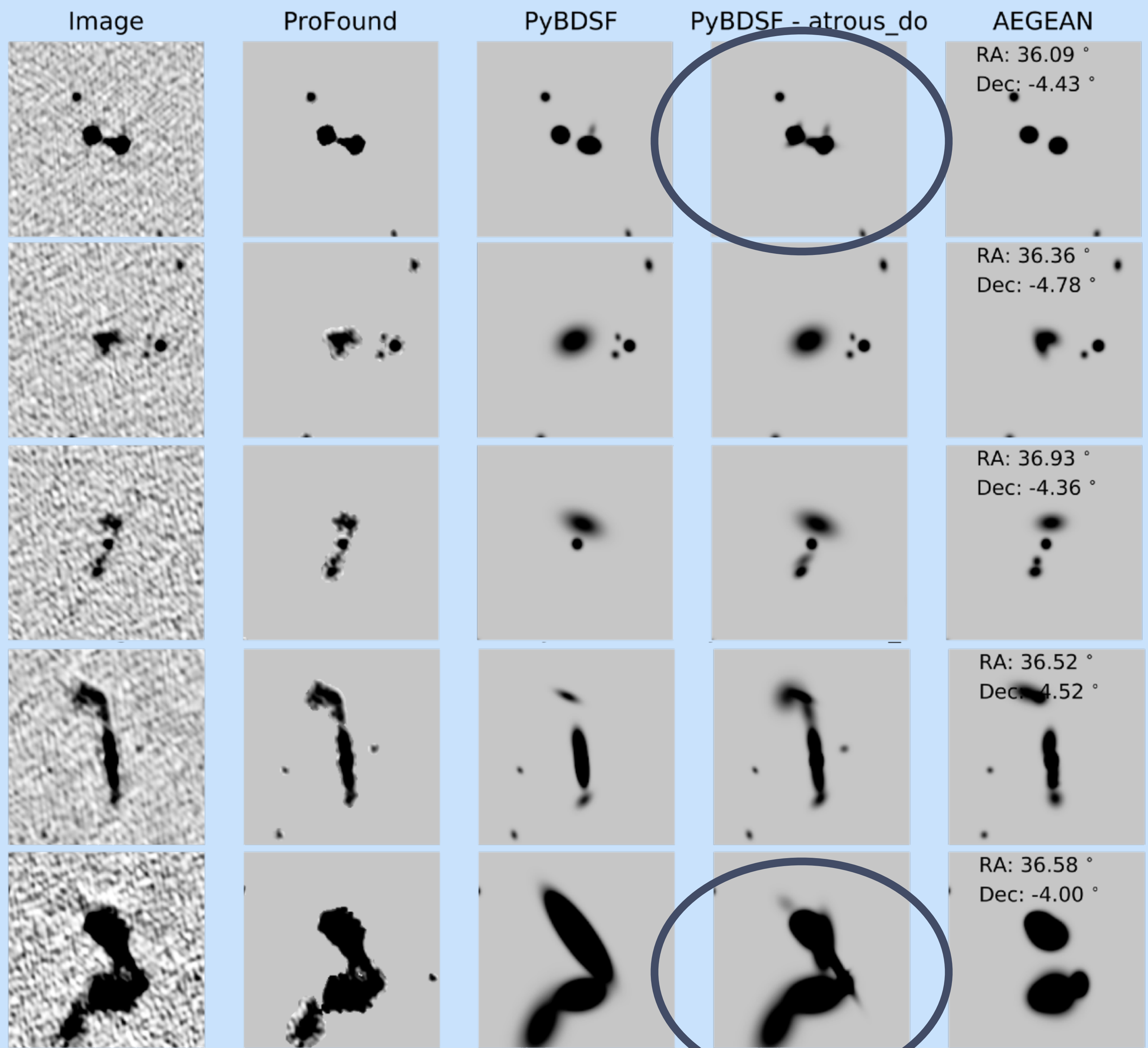
Skycut of 3.5
- Maximise number of Sources
- Not too much of a FDR issue

Methods

- Compared ProFound to both PyBDSF and AEGEAN
- Optimised FDR for each code:
  - ProFound - skycut=3.5
  - PyBDSF - thresh_isl=3, thresh_pix=5
  - AEGEAN - floodclip=4, seedclip=5

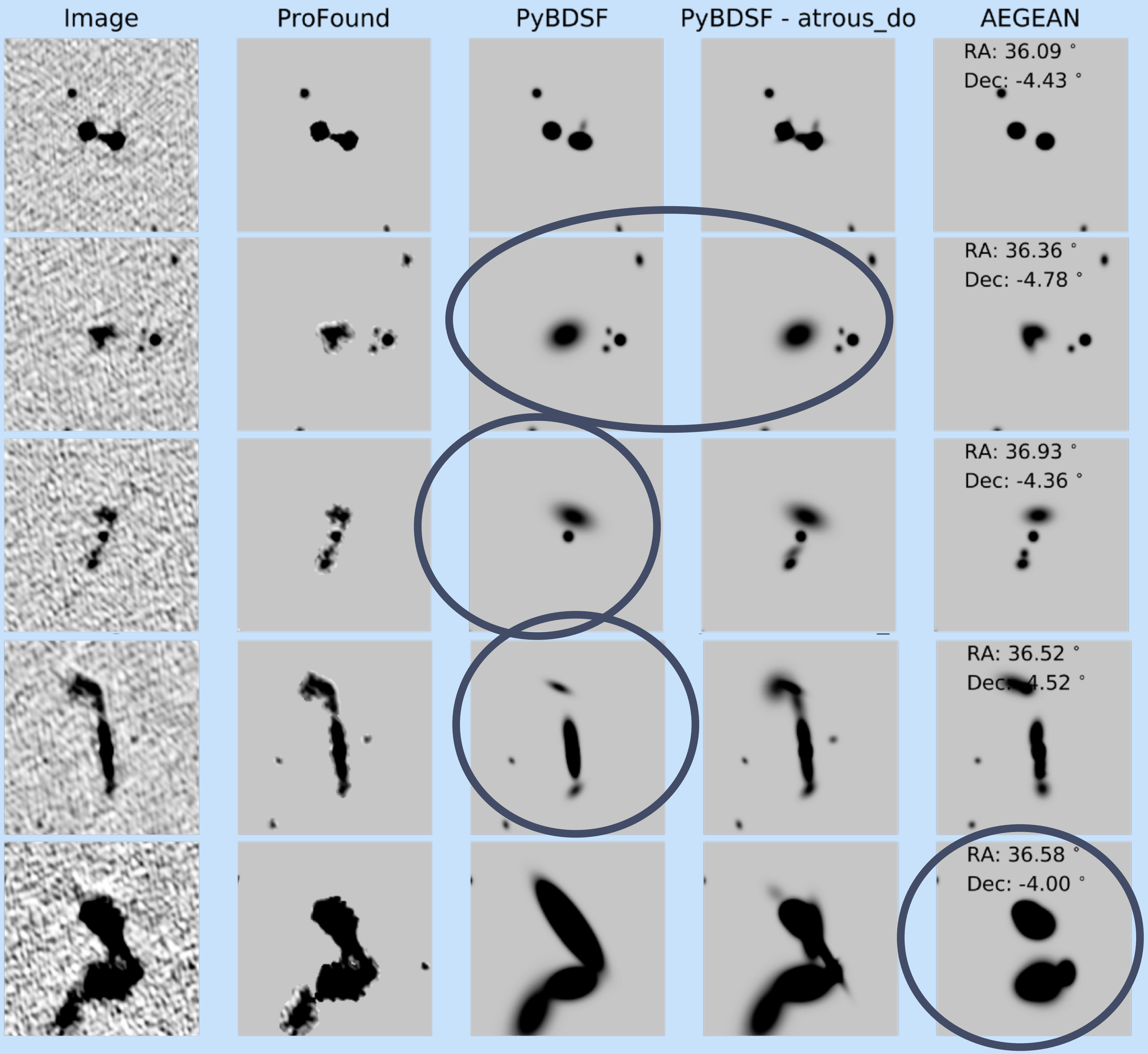
Sources WITH EXTENDED MORPHOLOGY

Streaking of emission due to extra gaussian components "needed"



JVLA-VIDEO Images (Heywood+ in prep)

Sources WITH EXTENDED MORPHOLOGY

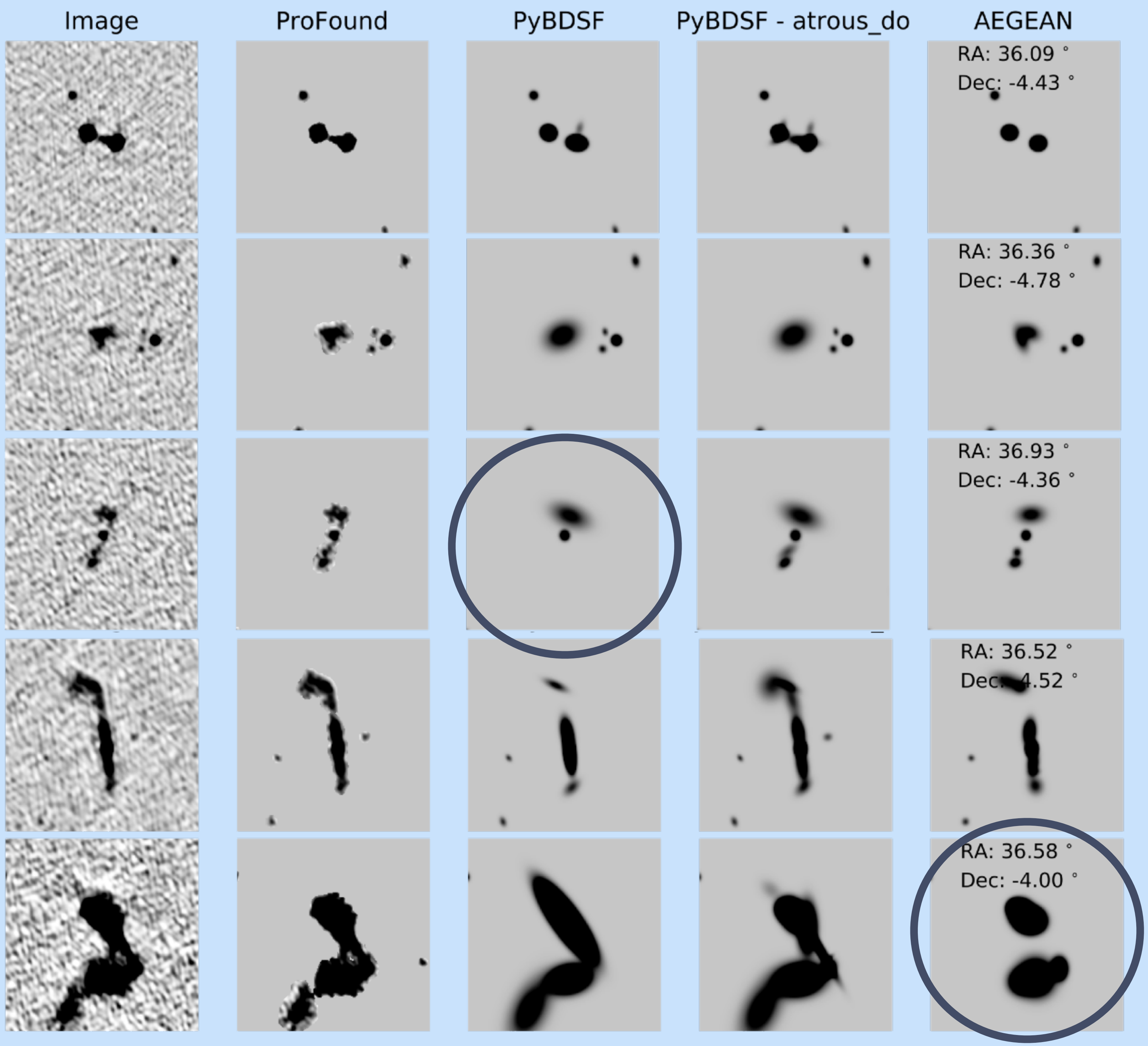


Too simple a model

JVLA-VIDEO Images (Heywood+ in prep)

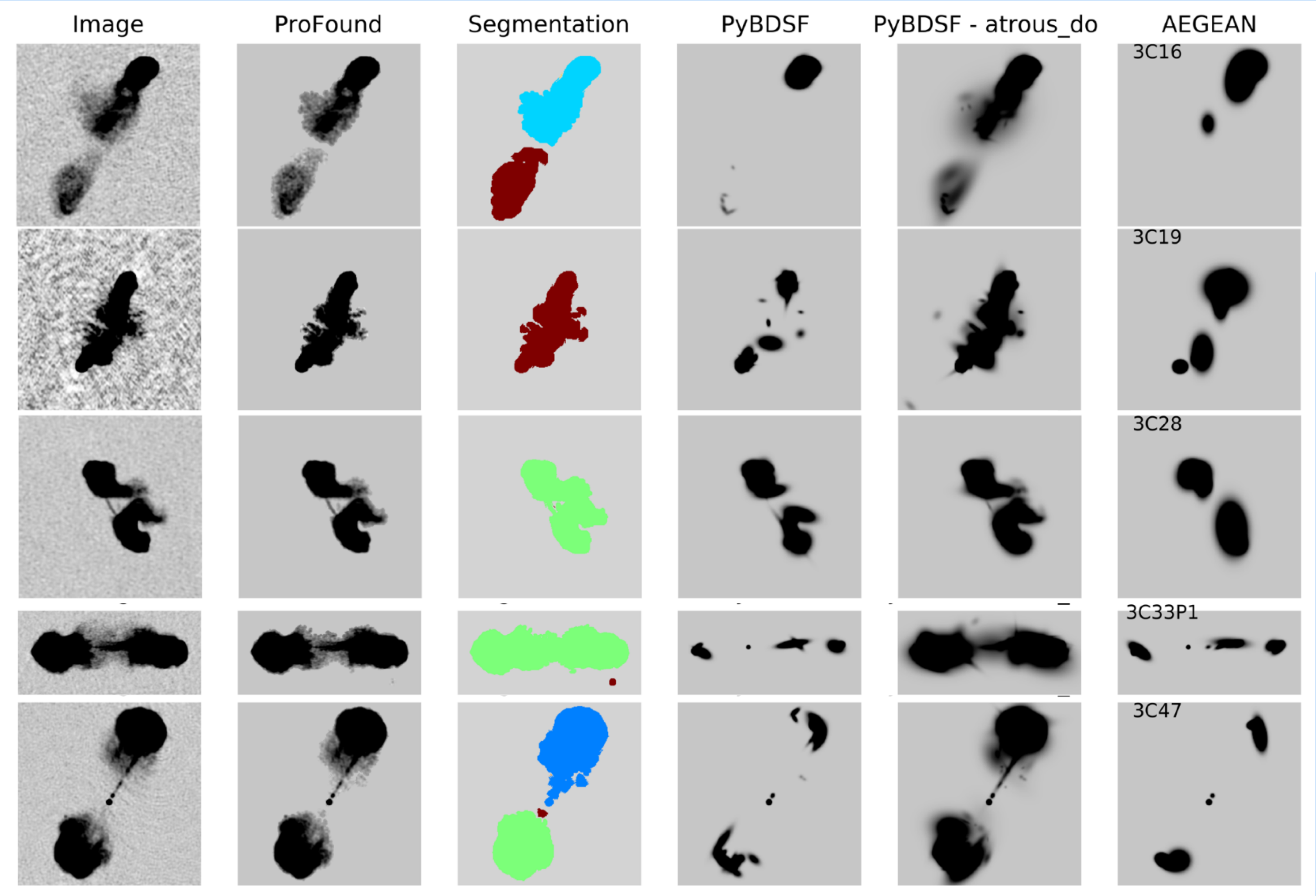
Sources WITH EXTENDED MORPHOLOGY

Missing Flux



JVLA-VIDEO Images (Heywood+ in prep)

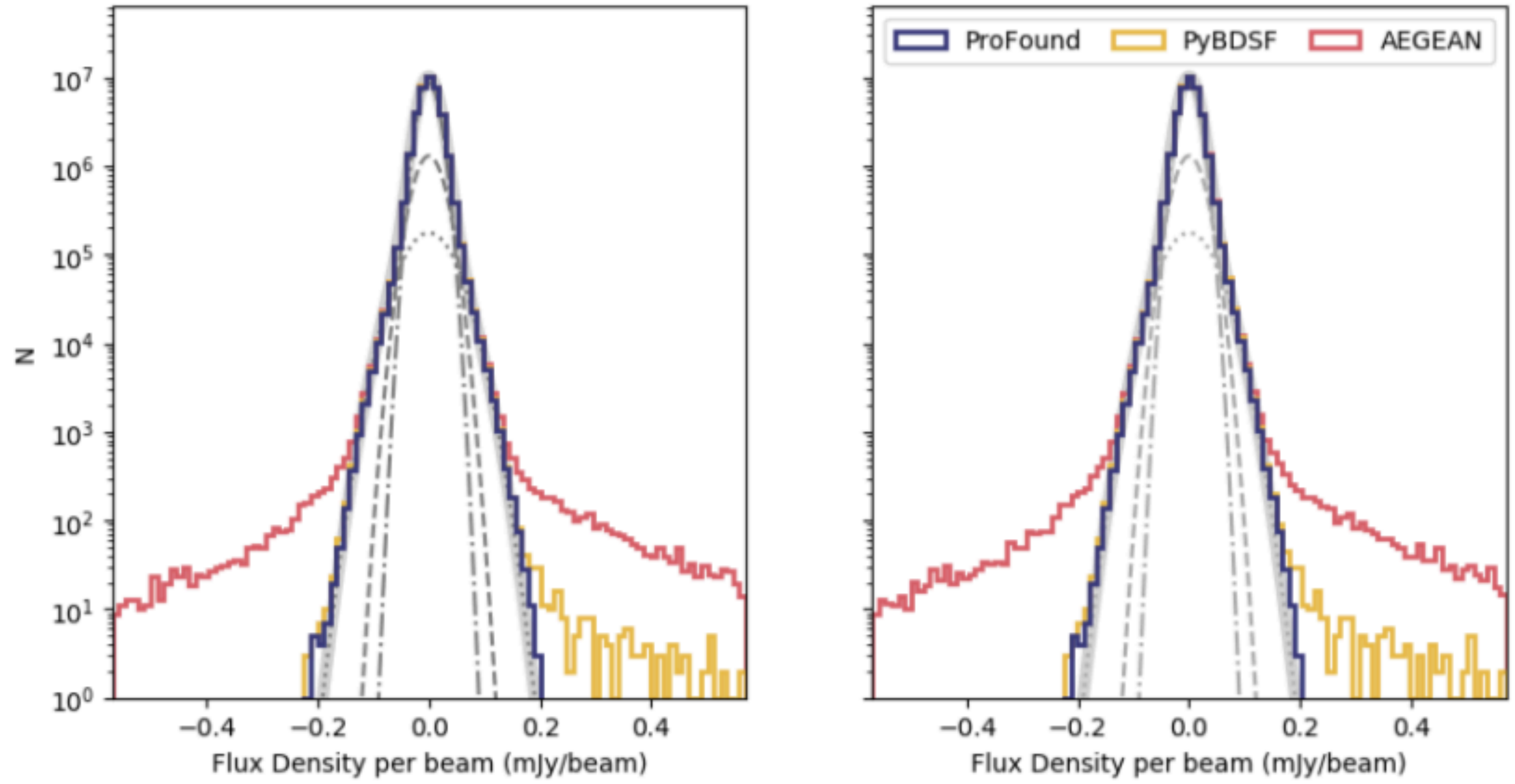
Observations of 3C Sources



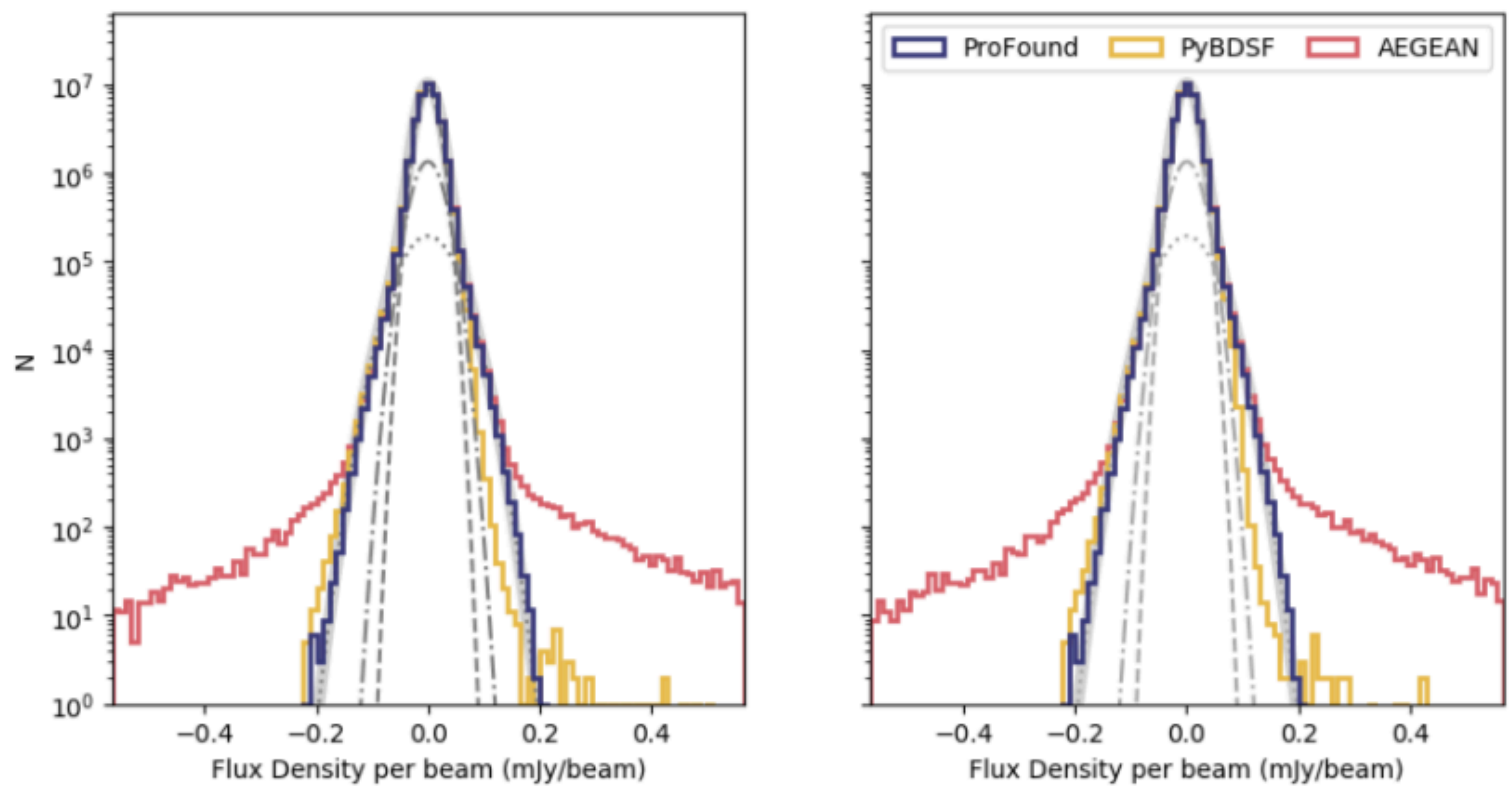
http://www.jb.man.ac.uk/atlas/

Residuals

- **Non-source sky residuals should be Normal around 0.**
- **ProFound behaves well, with no serious positive of negative flux remaining.**
- **PyBDSF slightly under models the true sources (so excess positive flux remaining).**
- **AEGEAN both under and over models the sources, and more aggressively than PyBDSF.**



(a) Without atrous_do used for PyBDSF.



(b) **With atrous_do used for PyBDSF.**

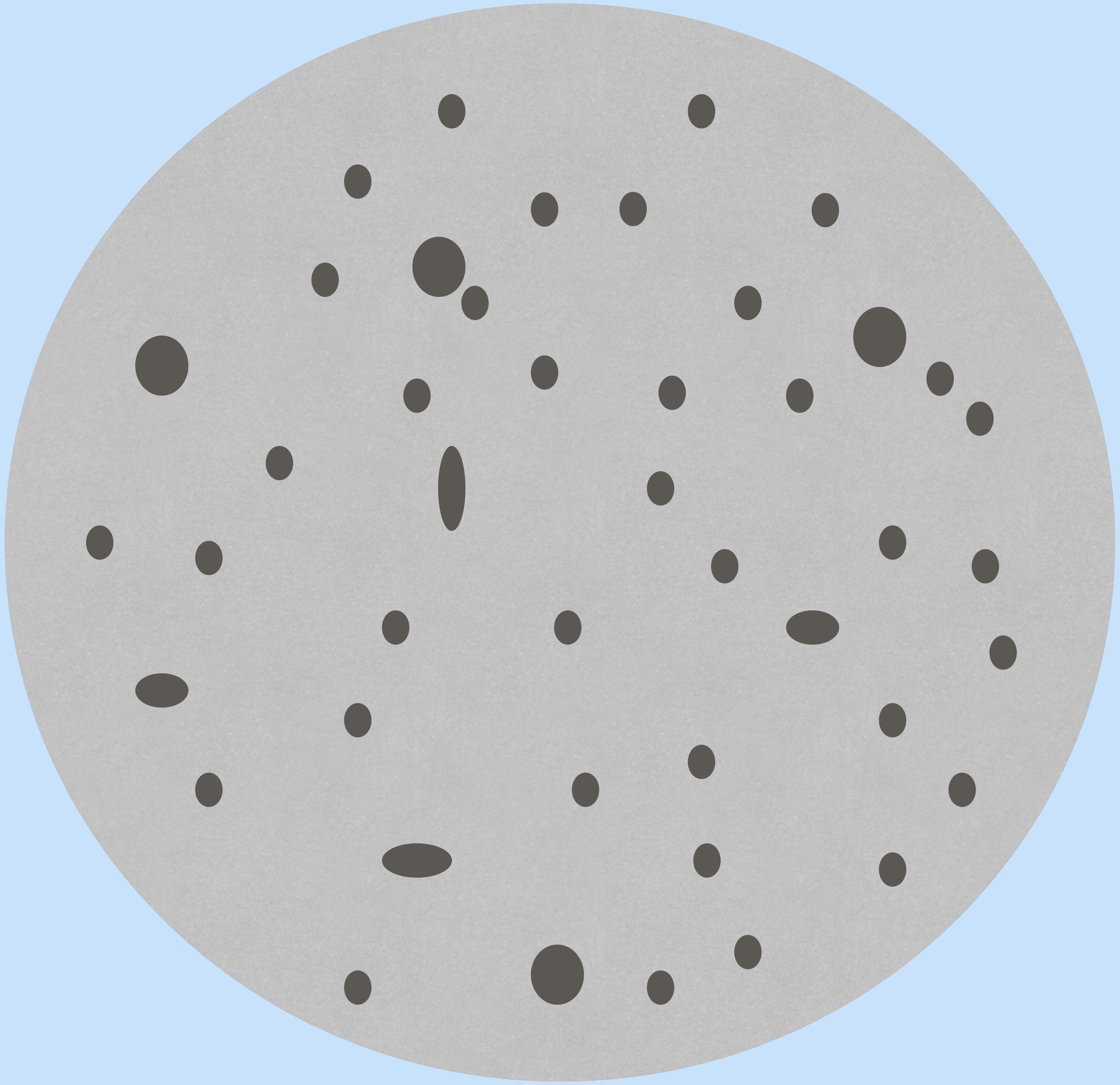
Testing ProFound

- Need to test how it works on known data
- Perform four simulations:
  - Model Gaussian sources
  - Model extended elliptical sources
  - Complex real sources from VLA images
  - Lobed morphology real source from 3C

Gaussian and Elliptical simulations

- Gaussian:
  Realistic sizes from PyBDSF sources and fluxes from Wilman+ 2008 SKA Simulated Skies

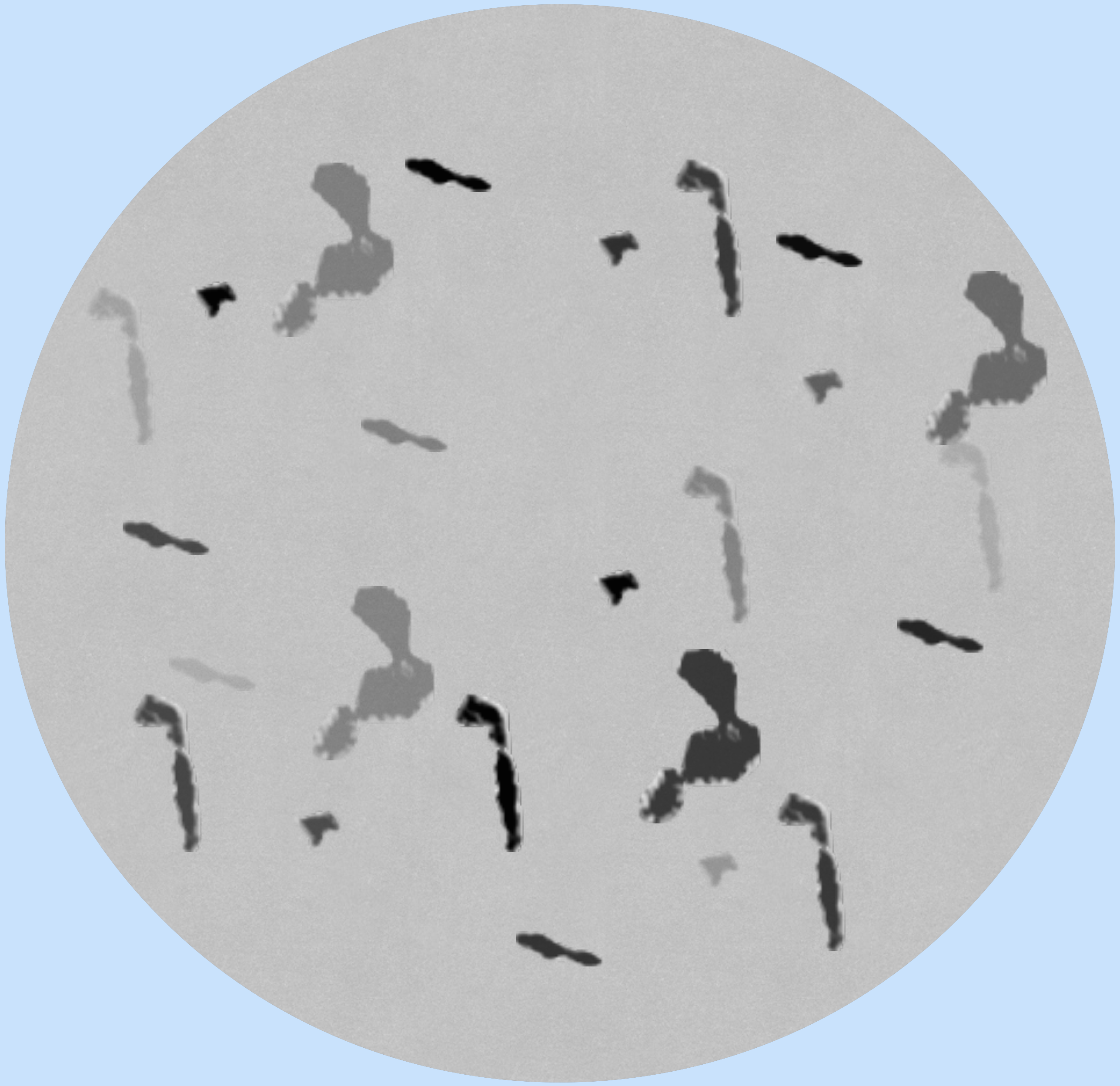
- Elliptical:
  Elliptical components from Wilman+ 2008 convolved with the beam



Complex and Lobed simulations

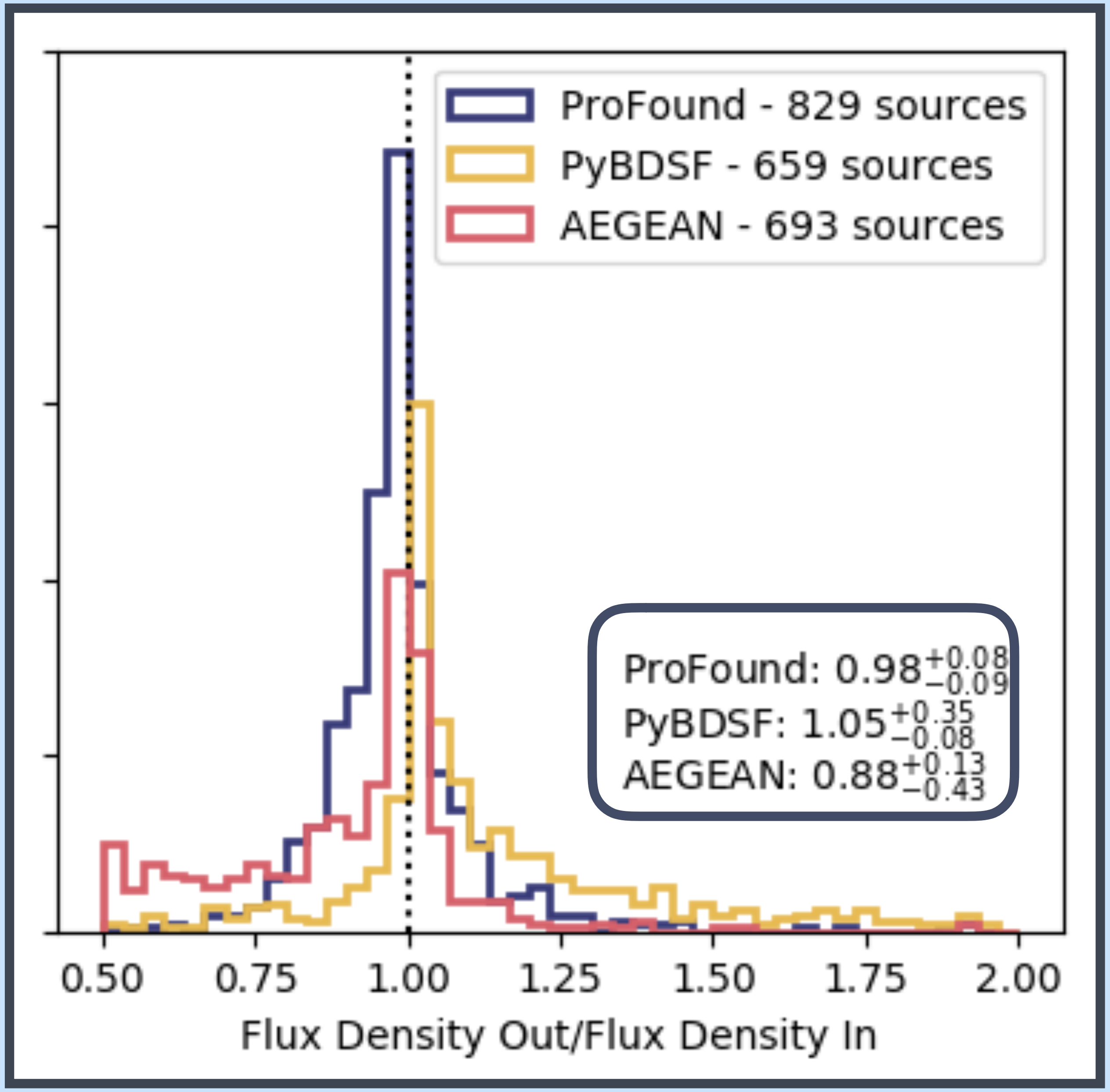
- Complex:
  Use source model of complicated sources from ProFound multiplied by factor (<1)

- Lobed:
  Use multi-component lobed sources from Wilman+ 2008

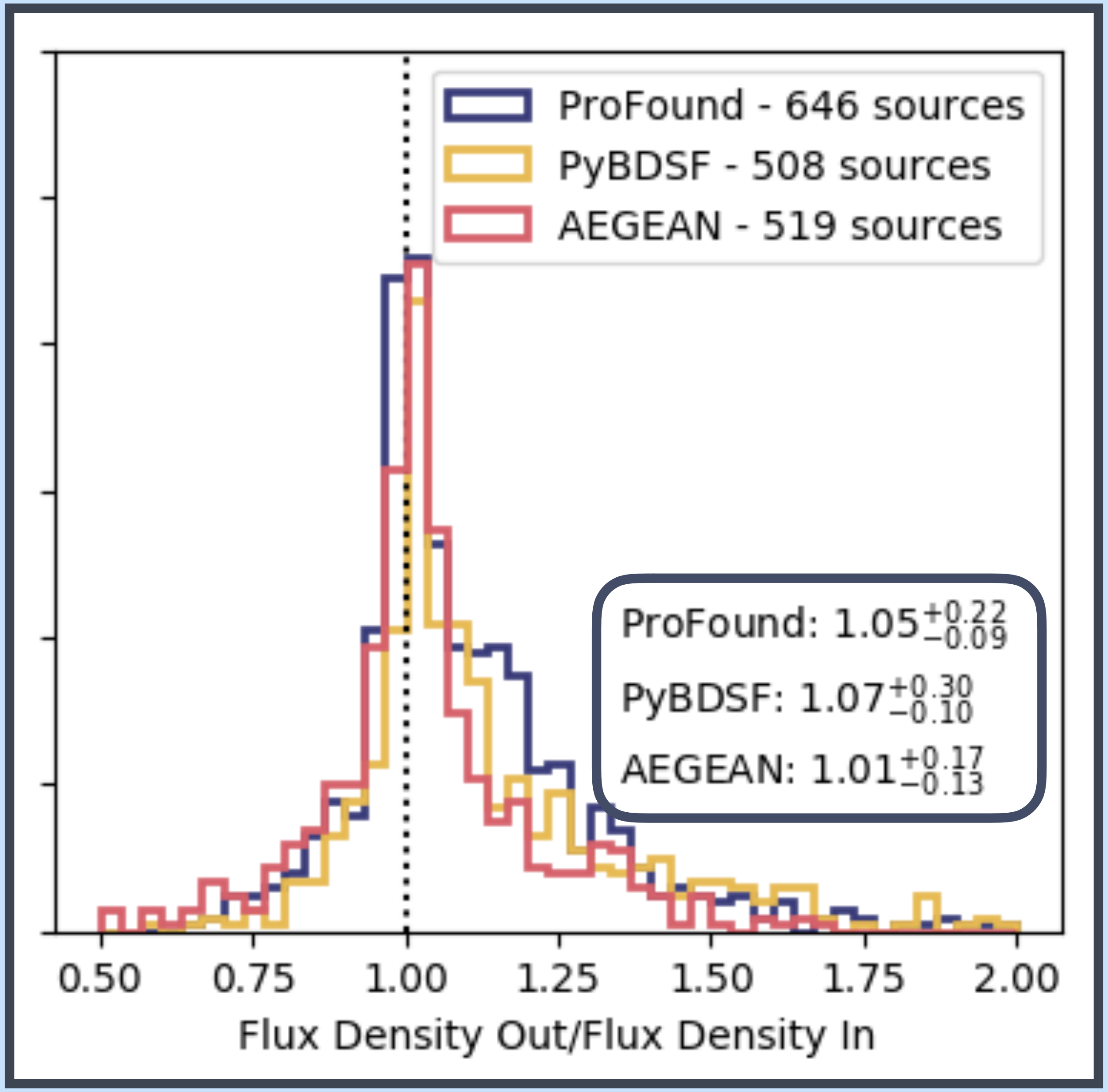


Point Like Sources ✔

Gaussians

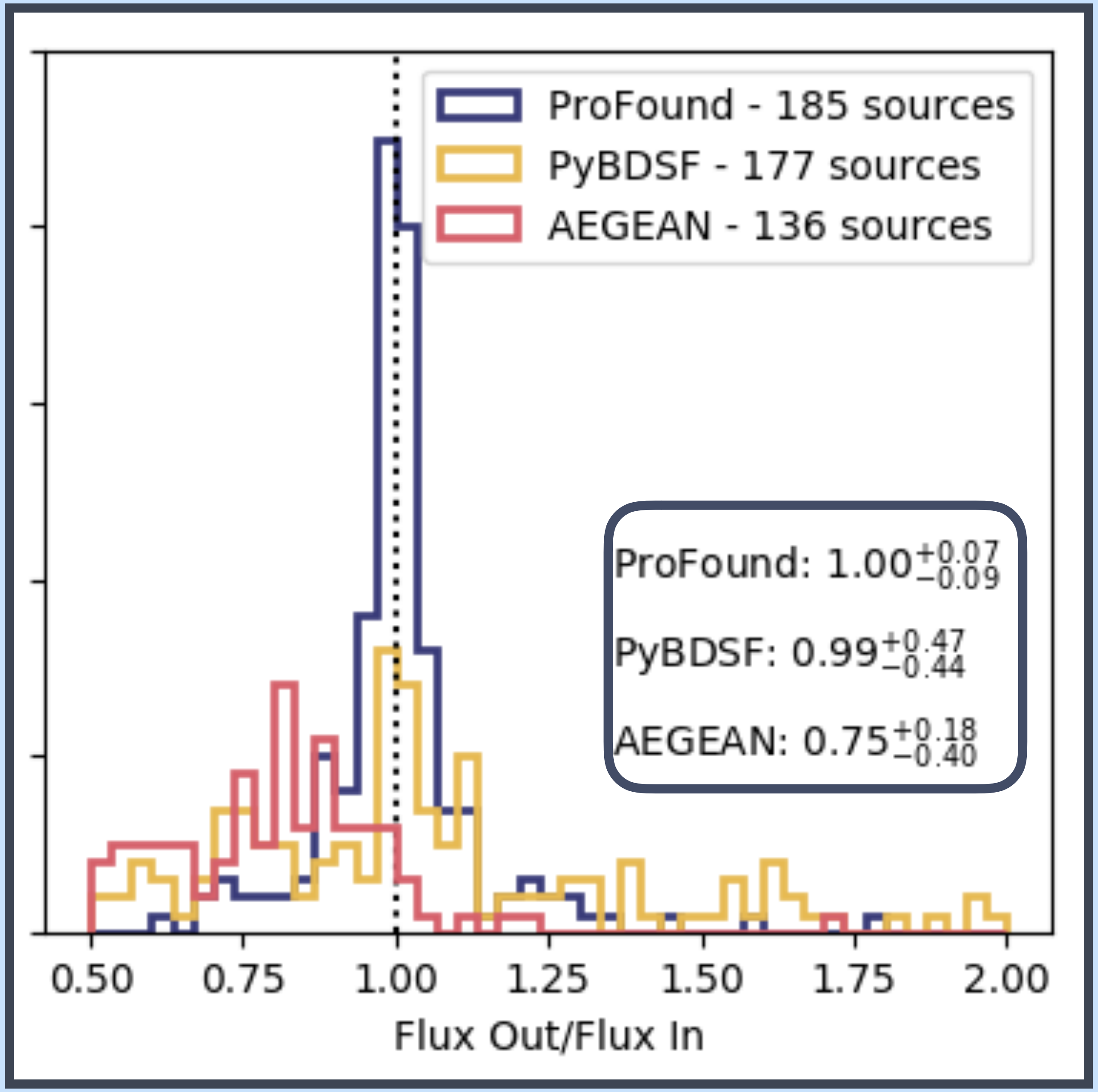


Elliptical

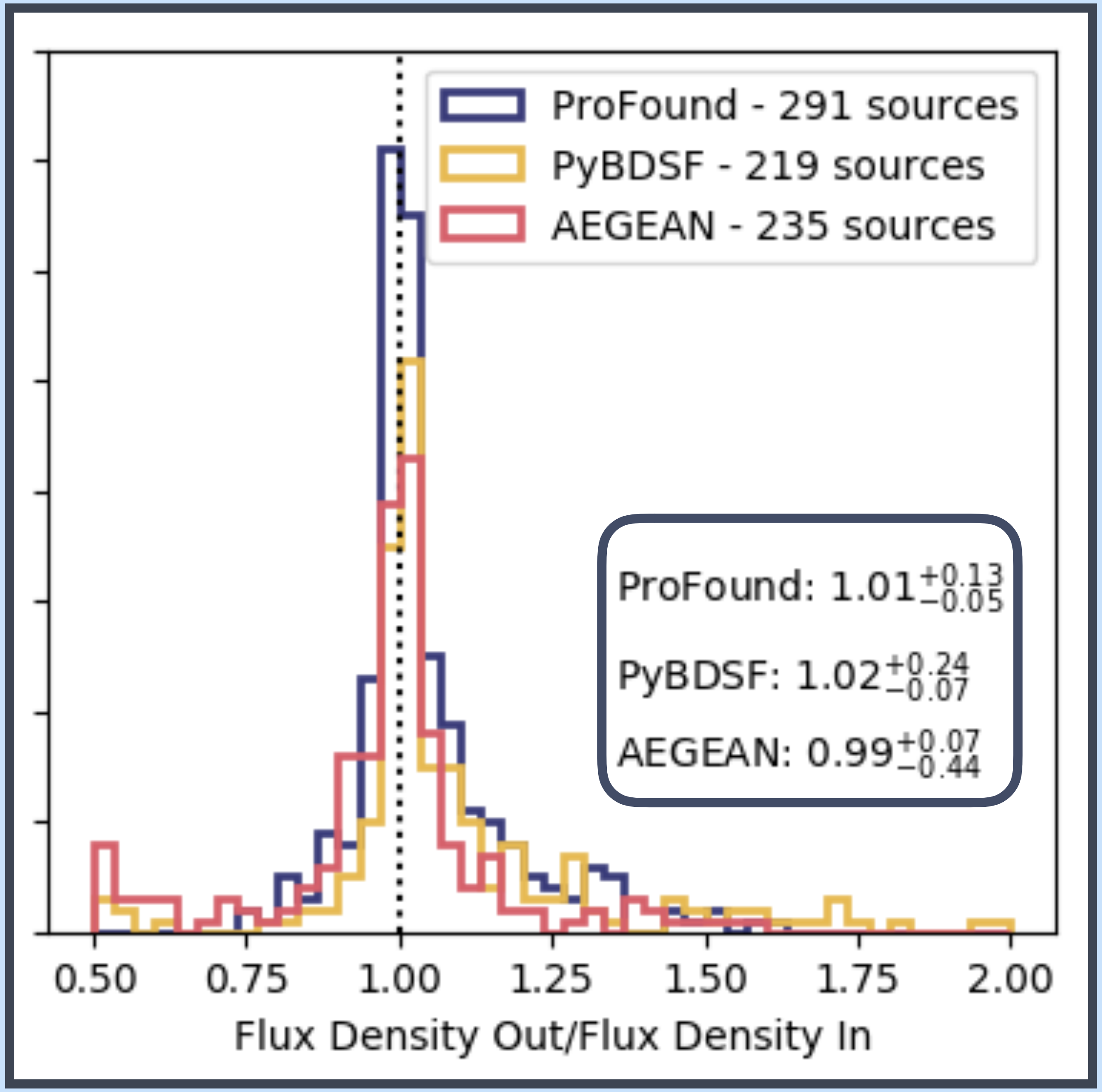


Extended Morphology Sources ✔

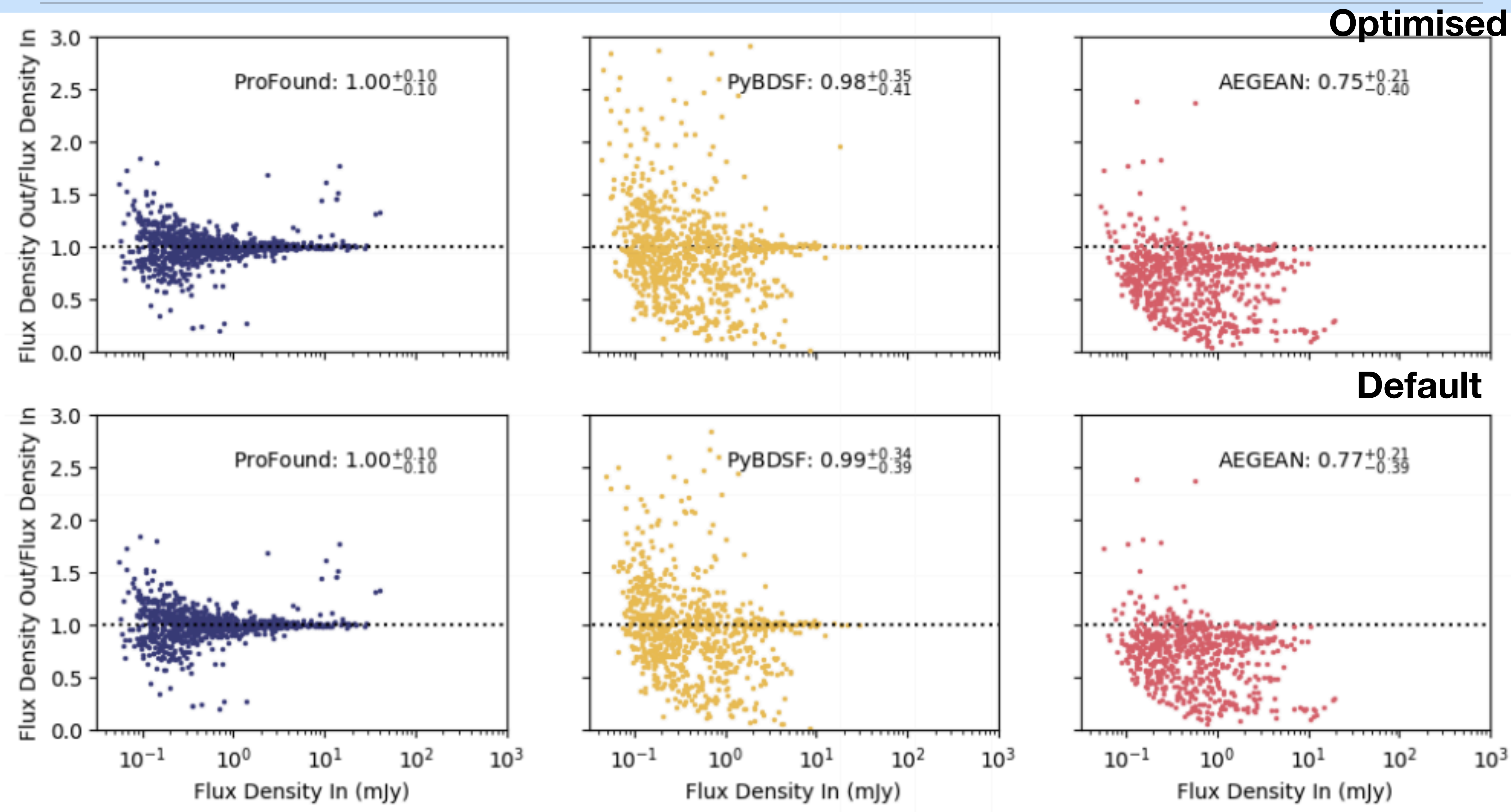
Complex



Lobed



Complex Sources



Advantages of ProFound

Advantages	Disadvantages
Easily finds and models extended structure	Can correct for this
Easy to combine together multiple components	Simple numerical factor
Not reliant on source morphology	
Can be used in multi-wavelength framework	

Conclusions

- ProFound is available on GitHub, and is now being used for a number of large optical and radio surveys Robotham et al 2018 (GitHub: asgr/ProFound)
- ProFound looks like it can be a **really useful radio source** extractor, see Hale et al 2019, arxiv-1902.01440
- It both models **compact sources** and **extended sources** well
- Many **multi-wavelength studies** applications
- Entered into the 1st SKA data challenge. Results are not out yet, but ProFound is clearly performing very well.