# How to build a TraP

## An image-plane transient-finding pipeline

**Tim Staley**
& TraP contributors.

Lunchtime talk, Southampton, Jan 2015

*WWW:* `4pisky.org` , `timstaley.co.uk/talks`

# Outline

# What are we missing?

Image surveys are best for finding

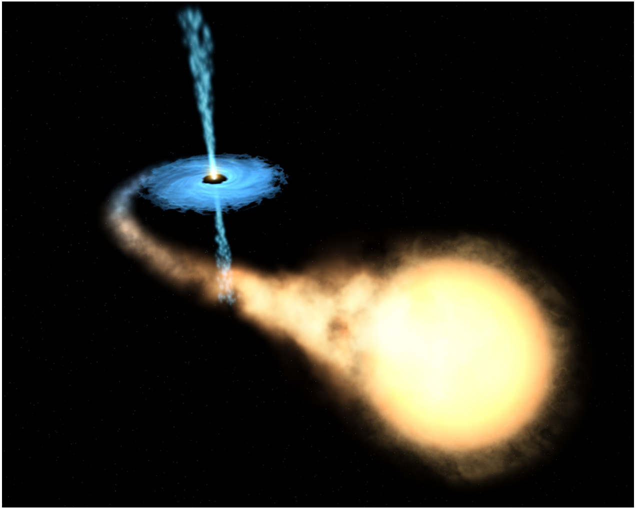
‘slow’ transients

i.e. > 1 second timescale —
excludes regular pulsars, etc.

Such as. . .

# What are we missing?

## Accretion flares



Artist's impression of the microquasar GRO J1655-40. Image credit: NASA/STScI

# What are we missing?

## 'Orphan' gamma-ray burst afterglows

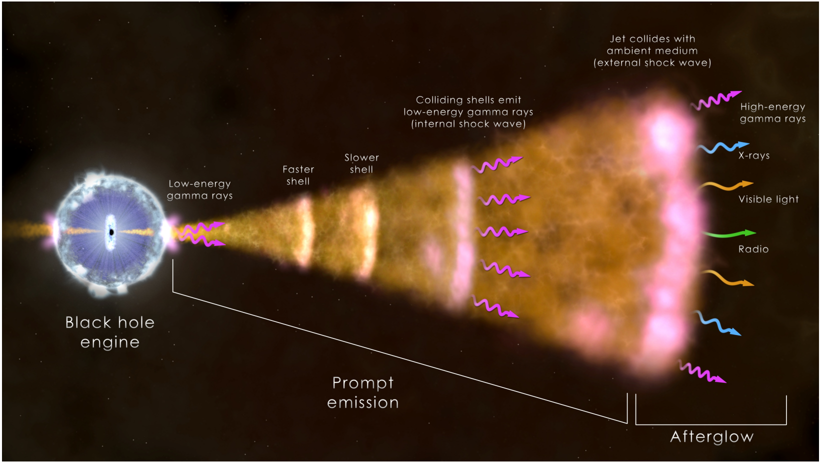



Image credit: NASA's Goddard Space Flight Center

e.g. *Ghirlanda 2014*, `http://adsabs.harvard.edu/abs/2014PASA...31...22G`

# What are we missing?

## Flare-star events

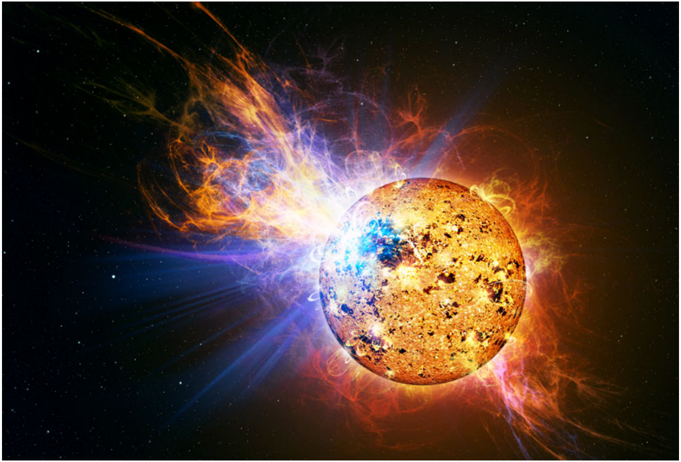


Image credit: Casey Reed/NASA

e.g. *Osten 2010*, `http://ukads.nottingham.ac.uk/abs/2010ApJ...721..785O`

# What are we missing?

Image surveys are best for finding 'slow', i.e. > 1 second timescale transients (excludes regular pulsars, etc).

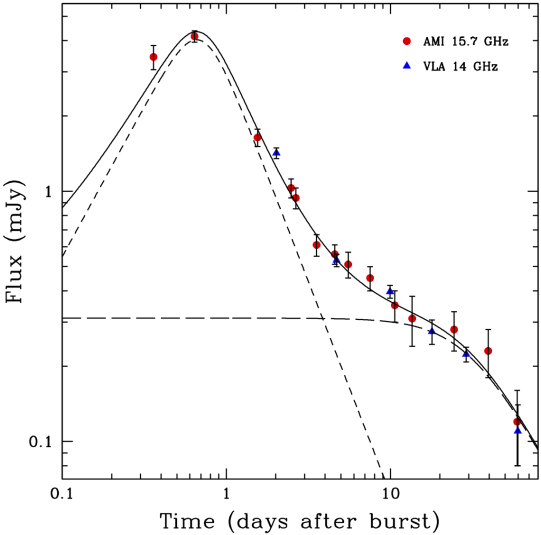
- AGN tidal disruption events
- Compact-object binary flares
- Orphan gamma-ray bursts
- Flare stars
- Nulling and eclipsing pulsars (e.g. J. Broderick et al, in press)
- (The unknown?)

# Proof of concept: ALARRM

## AMI-LA Rapid Response Mode

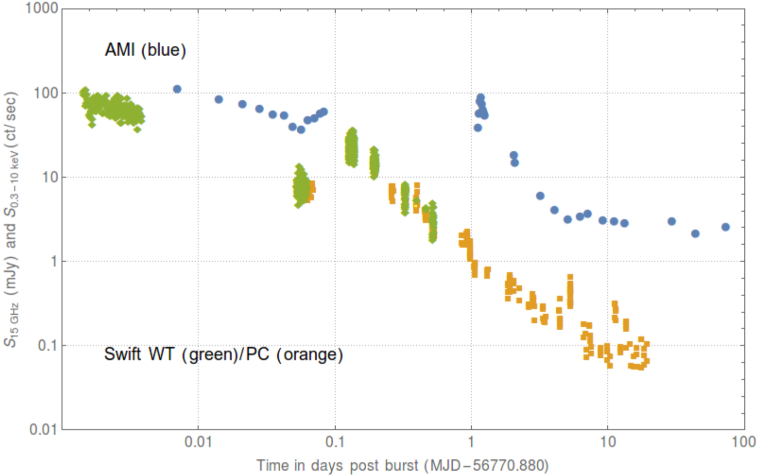
*Staley 2013*, `http://ukads.nottingham.ac.uk/abs/2013MNRAS.428.3114S`

# GRB140327A




*Anderson 2014*, http://adsabs.harvard.edu/abs/2014MNRAS.440.2059A
*van der Horst 2014*, http://adsabs.harvard.edu/abs/2014MNRAS.444.3151V

# DG CVn M-dwarf superflare



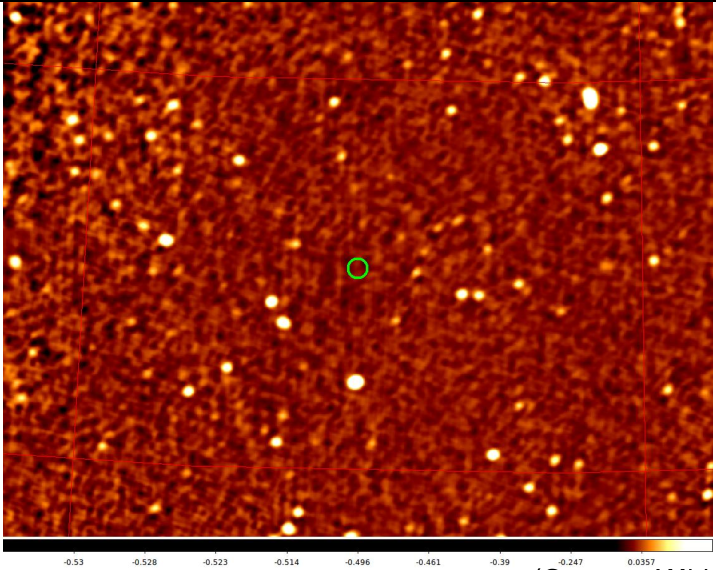

*Fender 2014*, `http://adsabs.harvard.edu/abs/2014arXiv1410.1545F`
*Osten et al (in prep)*

So:

Radio transients are out there.

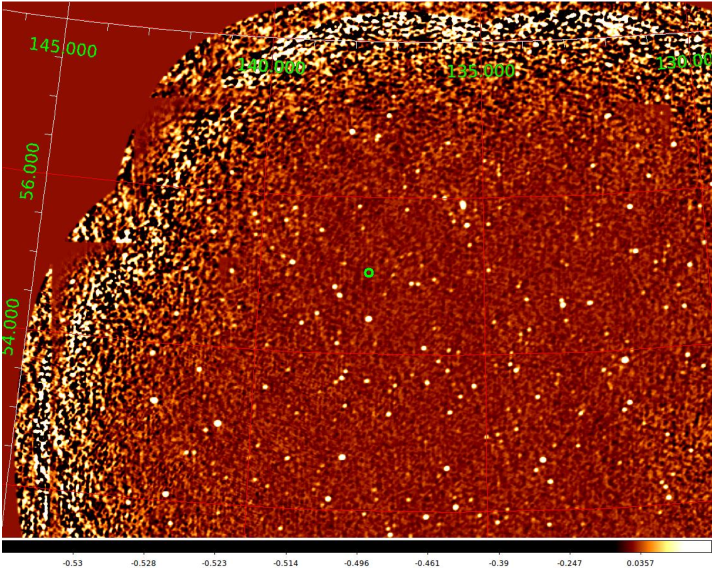
How do we find them directly?

# Bigger fields of view



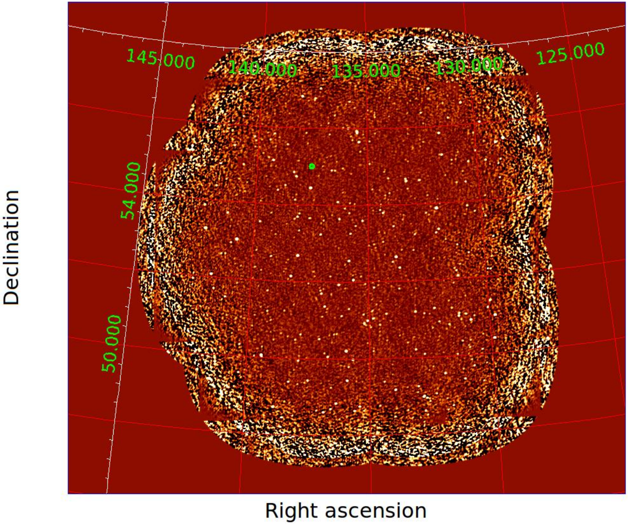

(Green = AMI-LA FoV)

# Bigger fields of view

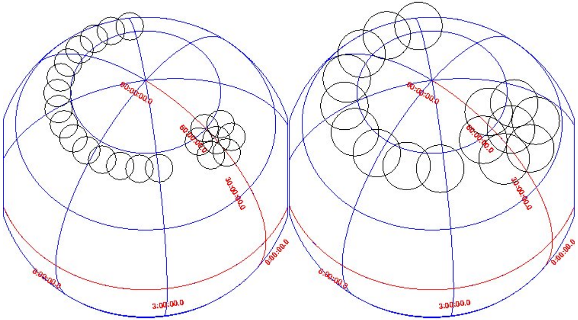


# Bigger fields of view

## LOFAR-RSM footprint



# The LOFAR 'Radio Sky Monitor'




Eight 7-beam tiles in LBAs tiles out entire zenith strip ($\sim 1800\text{deg}^2$ / $\sim \frac{1}{4}$ hemisphere)
Sixteen 7-beam tiles in HBAs for a narrower strip ($\sim 1000\text{deg}^2$ )

# Mini-summary

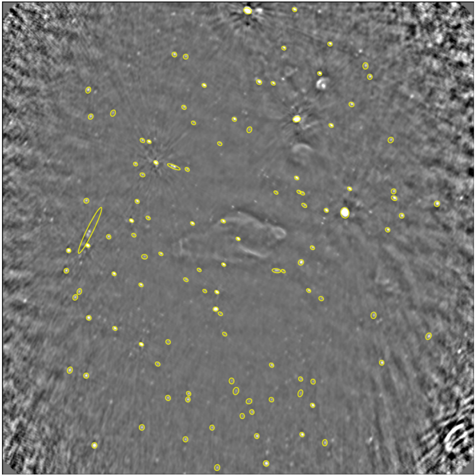
- There are interesting radio-transients waiting.
- Radio sensitivity / field of view is increasing by orders of magnitude.

# Mini-summary

- There are interesting radio-transients waiting.
- Radio sensitivity / field of view is increasing by orders of magnitude.
- $\implies$ Many uninteresting pixels, and a few exciting rare events.
- $\implies$ We need tools to search this data.

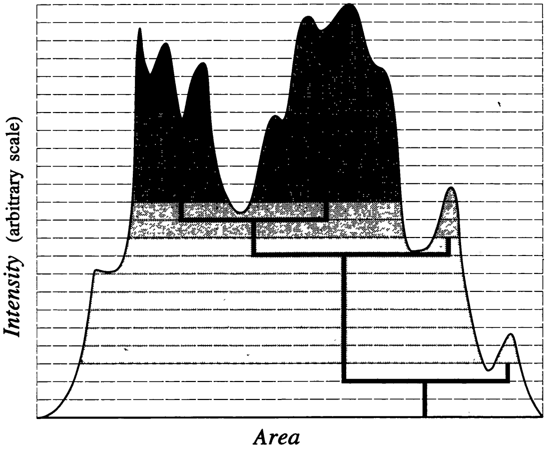
# Outline

# Step 1: PySE



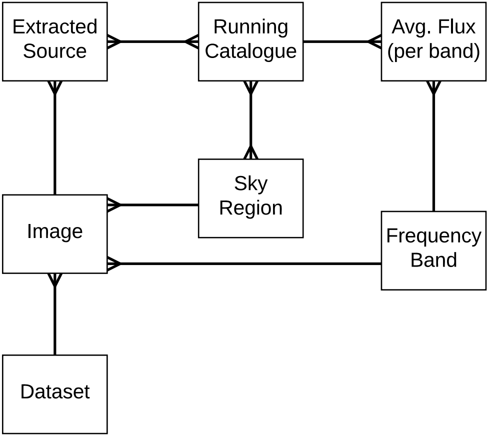
Python Source Extractor - Loosely based around S-Extractor algorithms, but tuned for radio data. Written solely in Python (extensive use of Numpy).

# Sourcefinding algorithms




An illustration of the island deblending method pioneered by *S-Extractor* (Bertin et al 1996).

# Step 2: Load 'extractedsources' into SQL database




NB: Store extractions without cross-matching initially.

# Brief aside on SQL

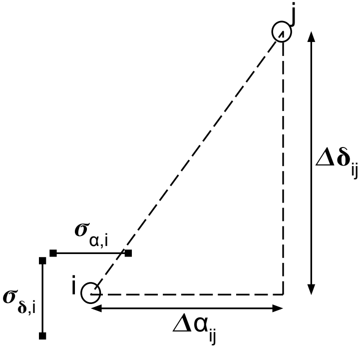
# Brief aside on SQL

- SQL is almost always *the* most efficient tool for searching large, well-parsed datasets.
- Are astronomers missing out?

# Step 3: Cross-match with known sources

## a.k.a. 'association'

First calculate DeRuiter radius for candidate associations:

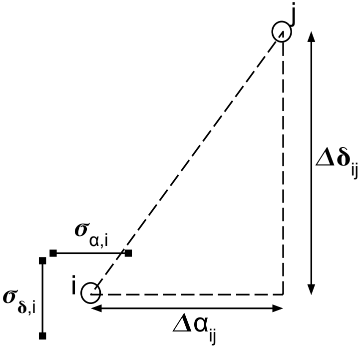



$$r_{ij} = \sqrt{\frac{(\Delta\alpha_{ij})^2}{\sigma_{\alpha,i}^2 + \sigma_{\alpha,j}^2} + \frac{(\Delta\delta_{ij})^2}{\sigma_{\delta,i}^2 + \sigma_{\delta,j}^2}}$$

# Step 3: Cross-match with known sources

## a.k.a. 'association'

First calculate DeRuiter radius for candidate associations:



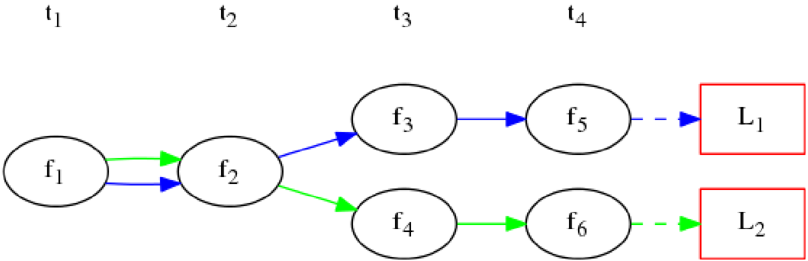

$$r_{ij} = \sqrt{\frac{(\Delta\alpha_{ij})^2}{\sigma_{\alpha,i}^2 + \sigma_{\alpha,j}^2} + \frac{(\Delta\delta_{ij})^2}{\sigma_{\delta,i}^2 + \sigma_{\delta,j}^2}}$$

*Handling meridian-wrap, celestial poles, left as exercise to reader.*

# Step 3: Cross-match with known sources

## a.k.a. 'association'

Mostly we just pick the closest match, and everything works out fine. But we also try to deal with some variable-PSF issues:



# Step 4: Identify bright new transients

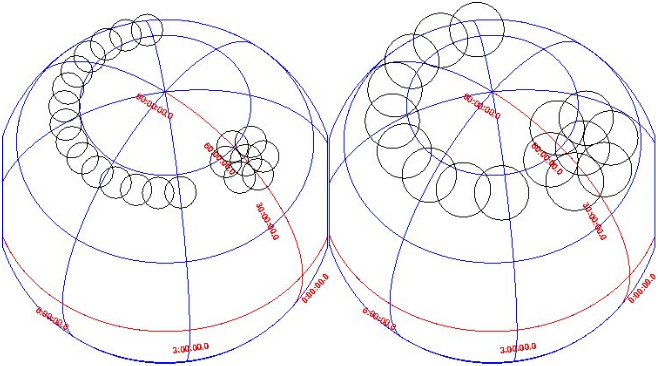
## The problem:

One of the neat features of TraP is that we can tell immediately when a bright new source appears. This may sound trivial initially, but...

# Step 4: Identify bright new transients

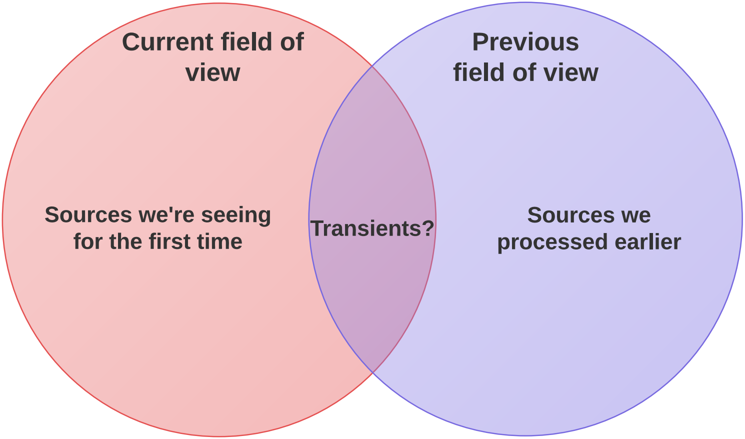
## The problem:

One of the neat features of TraP is that we can tell immediately when a bright new source appears. This may sound trivial initially, but...



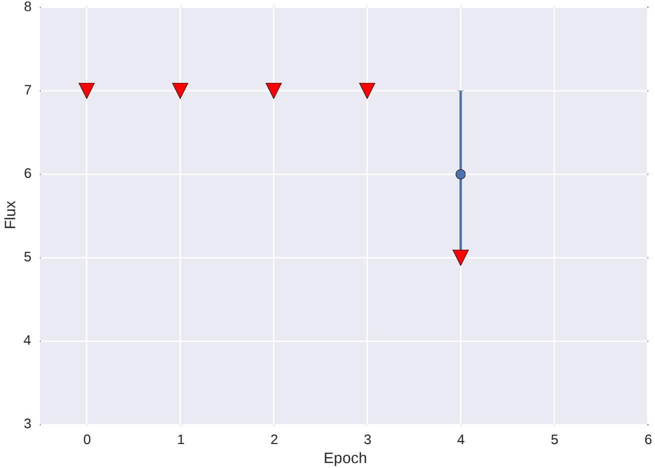
# Step 4: Identify new sources

## Tracking fields of view



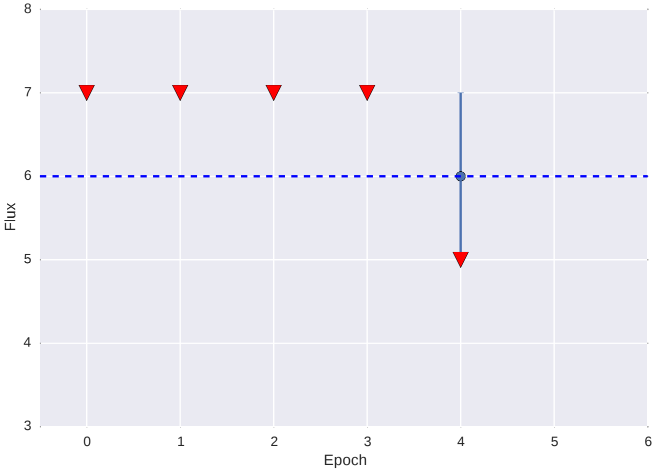
# Step 4: Identify new sources

## Tracking detection limits



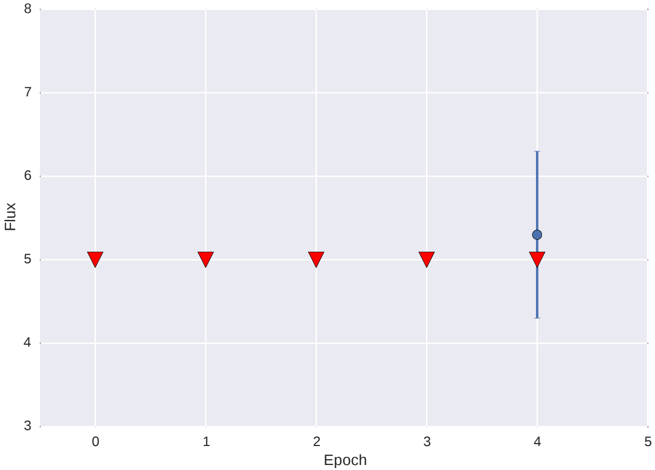
# Step 4: Identify new sources

## Tracking detection limits



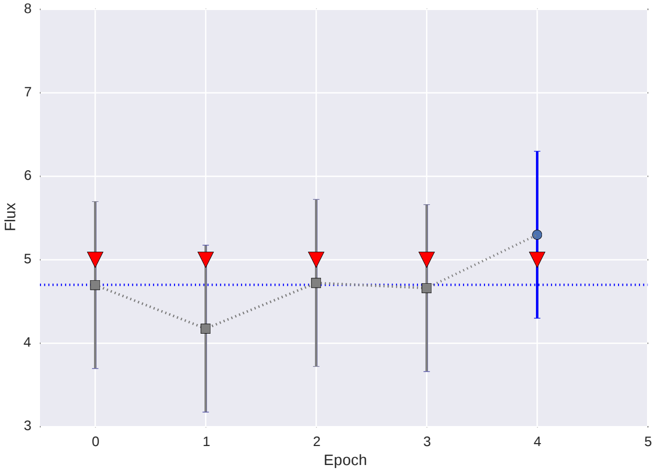
# Step 4: Identify new sources

## Tracking detection limits



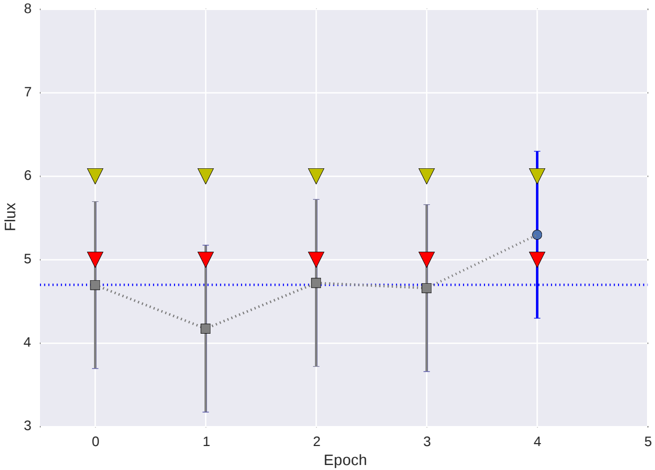
# Step 4: Identify new sources

## Tracking detection limits



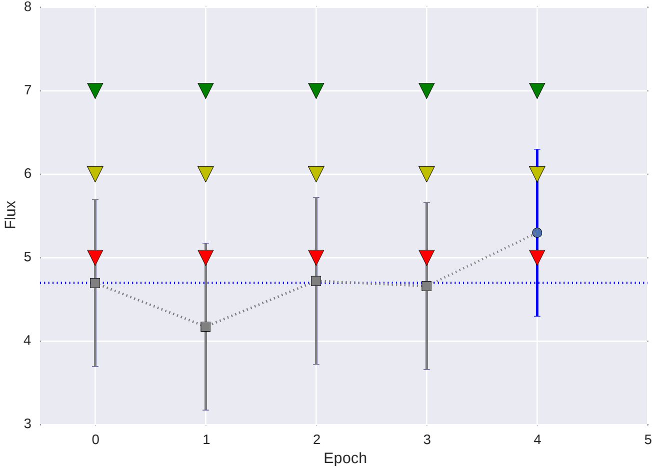
# Step 4: Identify new sources

## Tracking detection limits

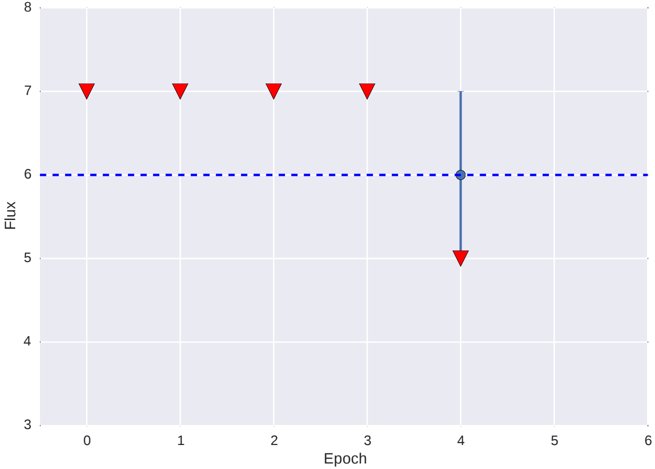


# Step 4: Identify new sources

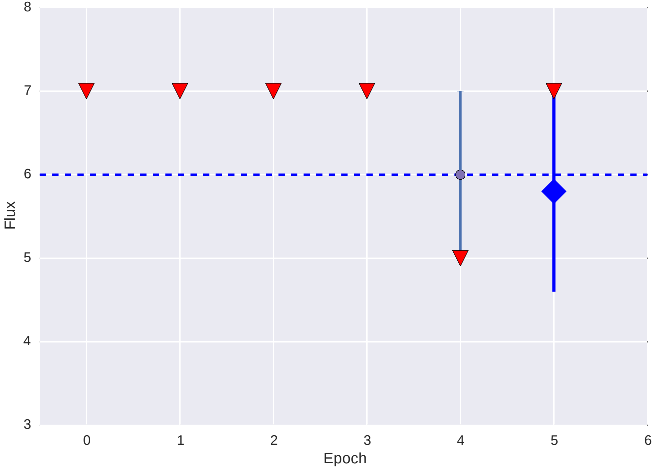
## Tracking detection limits



# Step 5: Force measurement of any missing known sources



# Step 5: Force measurement of any missing known sources



# Step 6: Analyse lightcurves

We keep a running aggregate (i.e. only need to include additional data, no recalculation of previous timesteps) for:

- Regular and weighted mean fluxes, $\mu$ & $\xi$

$$\xi_{x_{N+1}} = \frac{W_{x_N}\xi_{x_N} + w_{x_{N+1}} x_{N+1}}{W_{x_N} + w_{x_{N+1}}} \tag{1}$$

(2)

# Step 6: Analyse lightcurves

We keep a running aggregate (i.e. only need to include additional data, no recalculation of previous timesteps) for:

- Regular and weighted mean fluxes, $\mu$ & $\xi$

$$\xi_{x_{N+1}} = \frac{W_{x_N} \xi_{x_N} + w_{x_{N+1}} x_{N+1}}{W_{x_N} + w_{x_{N+1}}} \quad (1)$$

$$(2)$$

- 'Coefficient of variation', $V = \sigma / \mu$

- Calculate reduced $\chi$-squared value ($\eta$), against fit to straight line at level of weighted-mean $\xi$.

# Outline

# Installation

- TraP can be run on a laptop. But . . .
- Makes heavy use of SQL database (most astronomers not familiar)
- Expected to be run on large datasets

# Installation

- TraP can be run on a laptop. But . . .
- Makes heavy use of SQL database (most astronomers not familiar)
- Expected to be run on large datasets
- Solution: Web-interface. Displays data in user-friendly fashion, works extremely well in server-client model.

# Demo

$\Longrightarrow$ Demo.

# Development

## Facts

- ~20,000 lines of code, ~350 unit tests, 26K lines of docs
- 4 core developers, plus ~4 testers, 3 continents
- Remote, collaborative, development model
- Issue tracking
- Open (going forward)

# Development

## Implications

( / Ruminations)

- Astronomy –> More software intensive
- Getting anything done requires better code re-use
- Core software efforts are larger, require ongoing effort from many contributors.
  (cf `http://astropy.org`!)

# Development

## Implications

( / Ruminations)

- Astronomy –> More software intensive
- Getting anything done requires better code re-use
- Core software efforts are larger, require ongoing effort from many contributors.
  (cf `http://astropy.org`!)
- You don't have to be part of it, **but**, you should be aware of it
- Get to know the latest tools, maybe submit a bugfix here and there if you can
- Do better science, faster (hopefully!)

# Outline

# Different approaches

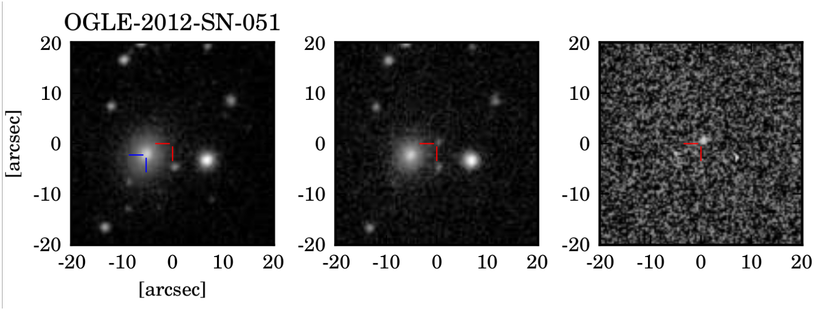
Two main approaches to image-based transient surveys:

- Cataloguing: Extract source representations, store in database, analyze lightcurve catalogue.
- Difference image analysis a.k.a. image subtraction

# Lightcurve cataloguing

- Basic concept easily understood - 'just' glue together source extraction and lightcurve analysis.
- However, blind source extraction typically requires good signal to noise - will miss marginal sources.
- Crowded fields are also a problem.

# Difference image analysis




- Better at picking out faint sources in clean data, *much* better in crowded fields.

*Alard & Lupton 1997*, `http://adsabs.harvard.edu/abs/1998ApJ...503..325A`
*Wyrzykowski et al, 2014*, `http://adsabs.harvard.edu/abs/2014AcA....64..197W`

# Optical survey characteristics

## Which technique should we employ?

Most transient surveys to date are optical:

- Fields often crowded or even confusion limited (best places to look for stellar flares, microlensing)
- PSF usually quite well behaved (smooth)
- Pixel noise usually uncorrelated / varies on a different scale to the PSF (dependent on sampling)
- $\implies$ DIA is usually best approach.

# Radio survey characteristics:
## (Contrast with optical)

- Fields usually quite sparsely populated, at least with current generation of instruments.
- Noise is correlated on beam-width scale.
- Dirty beam / PSF may vary significantly from image to image. May be well modelled, but this depends on system characterisation.
- (Phased arrays e.g. LOFAR) May see artifacts due to side lobes from out-of-field bright sources.
- $\implies$ DIA would cause many false positives, better to stick to high SNR cataloguing.

# Outline

# Summary

TraP:

- Good for (wide-field) sparsely populated surveys.
- Can be used for real-time transient detection.
- Produces catalogue of all sources.

# Summary

TraP:

- Good for (wide-field) sparsely populated surveys.
- Can be used for real-time transient detection.
- Produces catalogue of all sources.
- Server-based / web-interface reduction model well suited to large, challenging datasets.
- Open-source Python / SQL, with comprehensive test-suite and documentation.
- `http://ascl.net/1412.011`