PTFIDE: PTF Image Differencing & Extraction

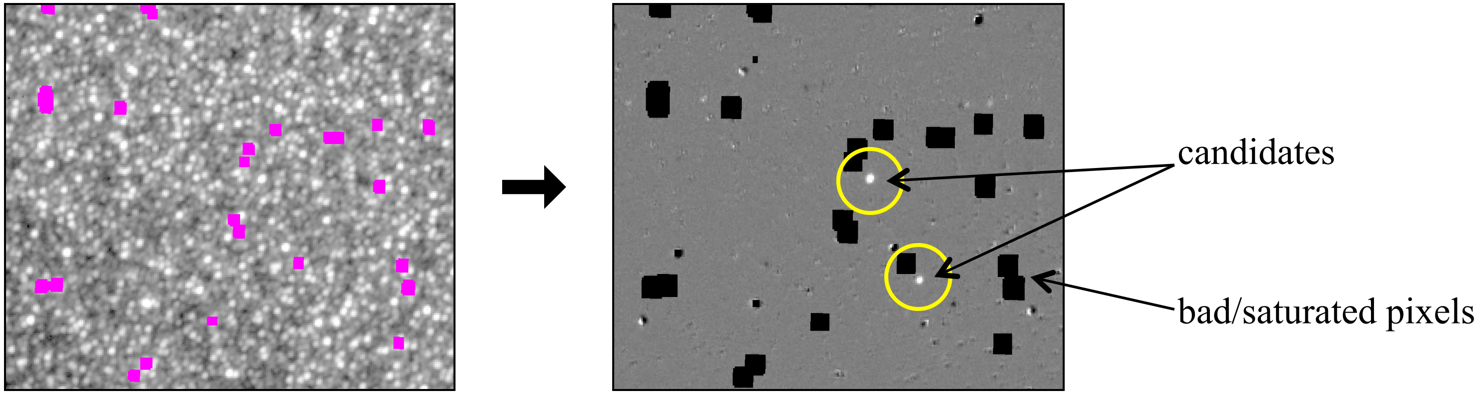
Frank Masci & the iPTF Collaboration

iPTF workshop, August 2013

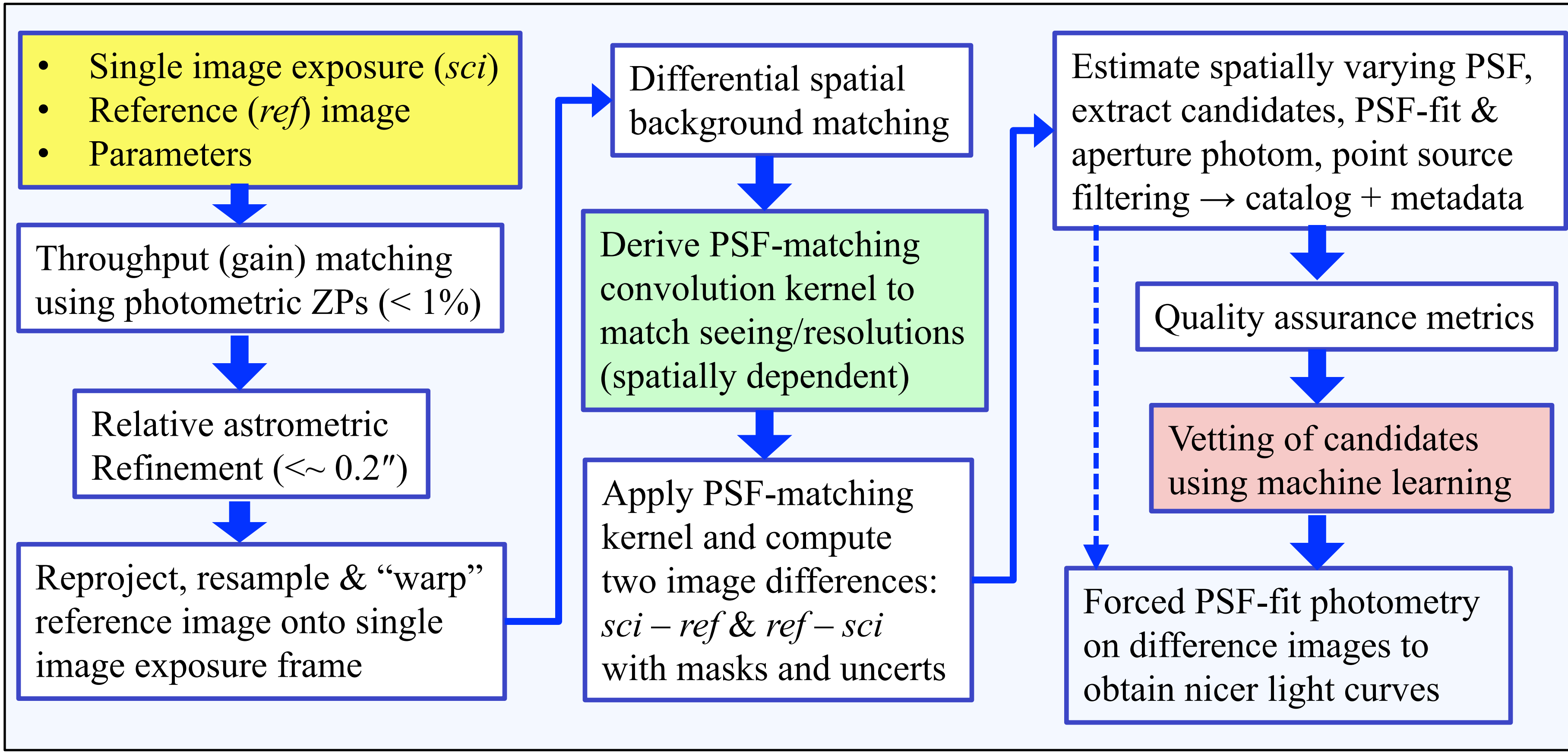
Goals

- Yet another difference-imaging pipeline! To support real-time transient discovery at Caltech.
- **Flexibility:** robust to instrumental artifacts, bad astrometry, adaptable to all seeing, little tuning.
- Operate in a **range of environments:** high source density, complex backgrounds and emission.
- **Generic:** discover transients of all types: pulsating & eruptive variables, SNe, asteroids.
- **Maximize** reliability of candidates and photometric accuracy to streamline vetting process.
- **Preprocessing steps:** “relative” calibration of input images crucial for good difference-imaging.
- Now in the operations environment at IPAC/Caltech to support archival research requests
  - ➢ currently supporting the moving object pipeline to discover asteroids.

near galactic bulge
~ 5 arcmin



PTFIDE processing flow



http://web.ipac.caltech.edu/staff/fmasci/home/miscscience/ptfide-v4.0.pdf

PSF-matching

- An observed image (exposure) can be modeled as:

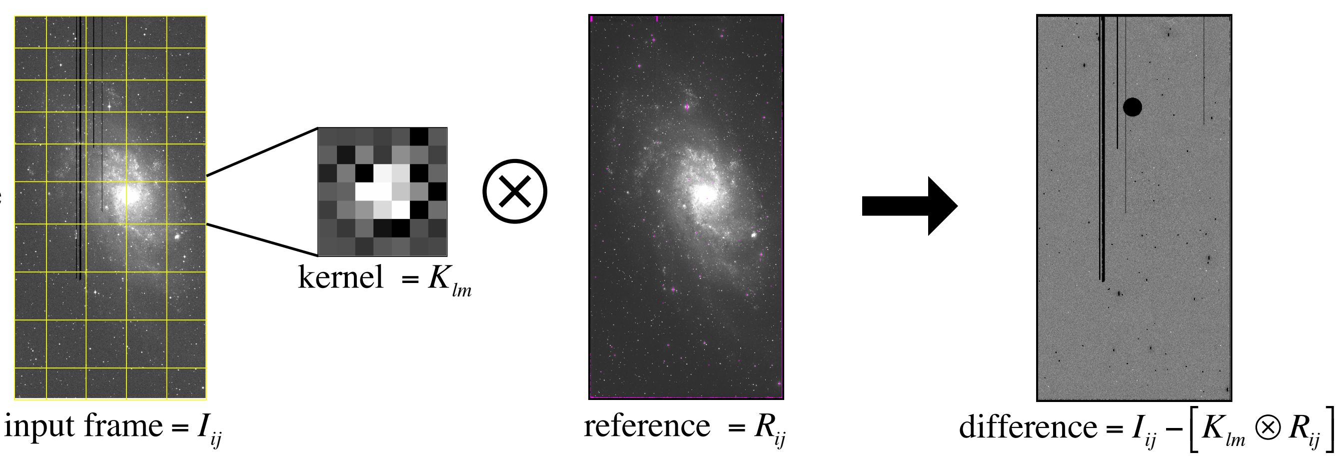
$$I_{ij} = \left[K_{lm} \otimes R_{ij}\right] + dB + \varepsilon_{ij}$$

- PSF-matching entails finding an optimum convolution kernel K by minimizing some cost function:

$$C = \sum_{i,j}\left[I_{ij} - \left(K_{lm} \otimes R_{ij}\right) - dB\right]^2$$

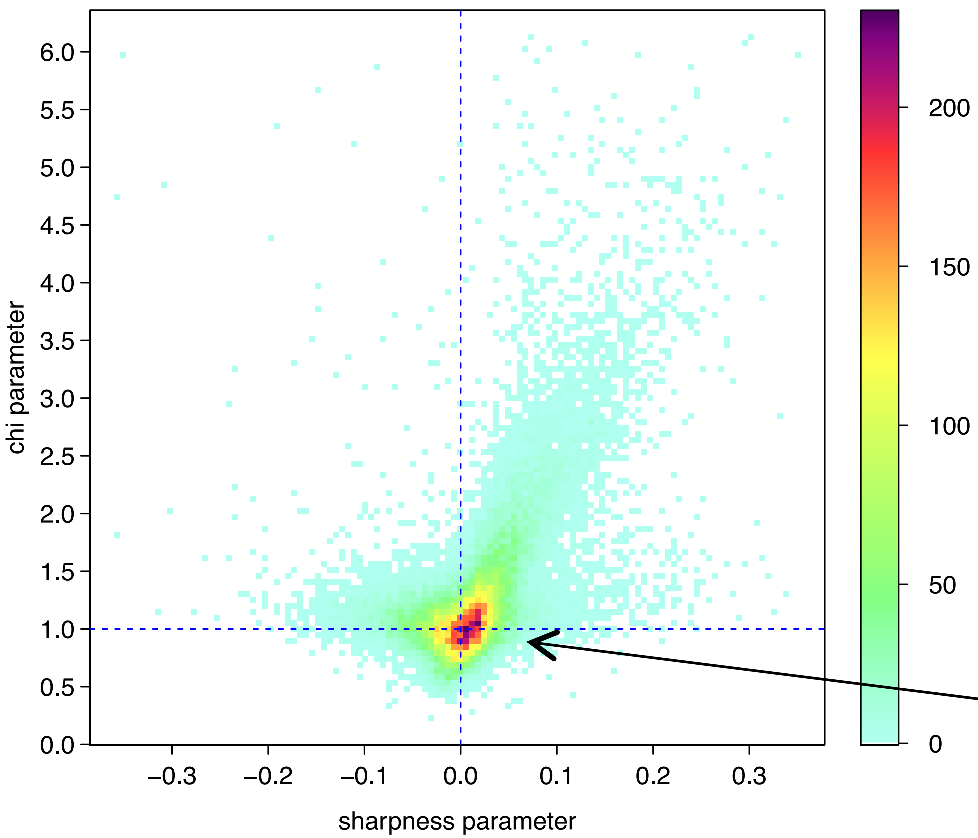
- Traditional method: decompose K into a sum of Gaussian basis functions × by polynomials (e.g., Alard 2000) as implemented in *HOTPANTS*, *ISIS* software. Coefficients are then fit for.
  - ➢ User must specify number of basis functions, Gaussian widths, polynomial orders, including spatial orders.
  - ➢ No rules of thumb to ensure optimality for all images. Hard to tune for a survey – at least for PTF!
- Instead we solve for each of the kernel pixel values K_{lm} (= 7 × 7 parameters) directly via LLS.
  - ➢ Similar to Bramich (2008); more flexible, K can take on more general shape, compensate for bad astrometry.
  - ➢ Since PSF is spatially dependent, we grid images into 5 × 10 overlapping squares, then solve for K in each.

M33 example



Advantages of PSF-fitting for transient photometry

- PSF fitting: better photometric accuracy for moderate to faint fluxes.
- Provides diagnostics to distinguish point sources from glitches (false-positives) in diff. images.
- Maximizes reliability of difference-image extractions since “static” transients are point sources.
- Assumes accurate PSF-estimation (over chip) and image registration prior to differencing.



DAOPhot metrics

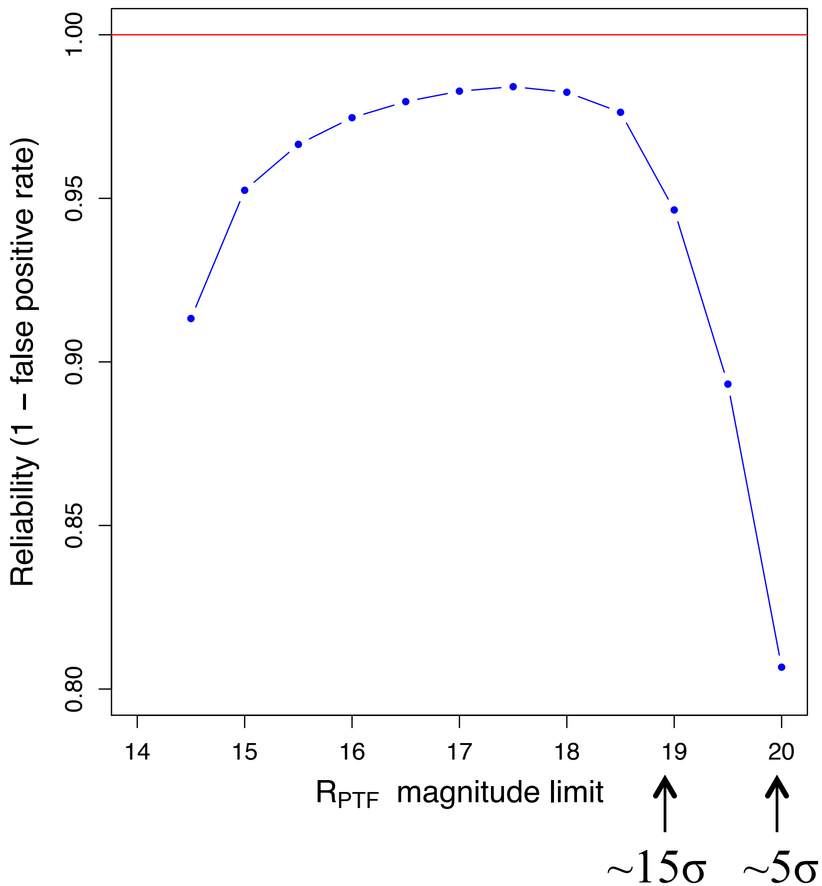
$$sharpness = \mathrm{FWHM}^2_{source} - \mathrm{FWHM}^2_{calibratedPSF}$$

$$\chi = \frac{\left\langle (\text{PSF-fit residuals})^2 \right\rangle^{1/2}}{\left\langle \text{prior 1-}\sigma \text{ uncertainties} \right\rangle}$$

ideal point sources: sharp, chi ~ 0, 1

Performance: real vs. bogus (reliability)

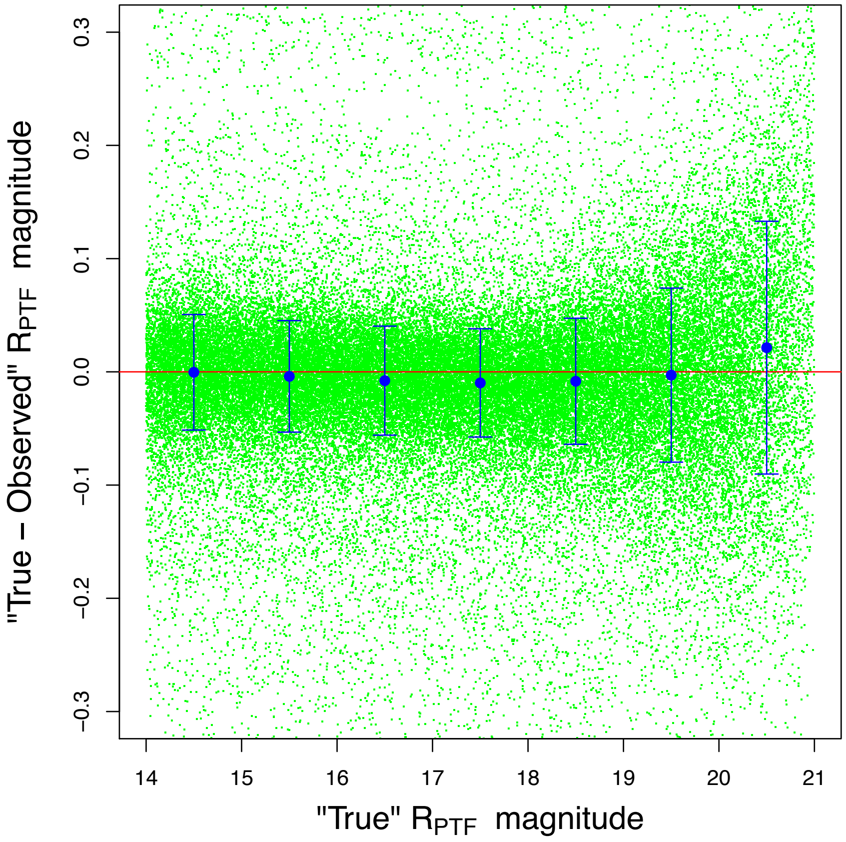
- took ~350 real, moderately dense *R*-band frames, derived spatially-varying PSFs, then simulated point source transients with random positions and fluxes.
- executed PTFIDE to create diff images and extract candidates with **fixed** threshold (S/N = 4) and filter params.



$$R = \frac{\#\,\text{matched to truth}\ (< R)}{\#\,\text{total extracted}\ (< R)}$$

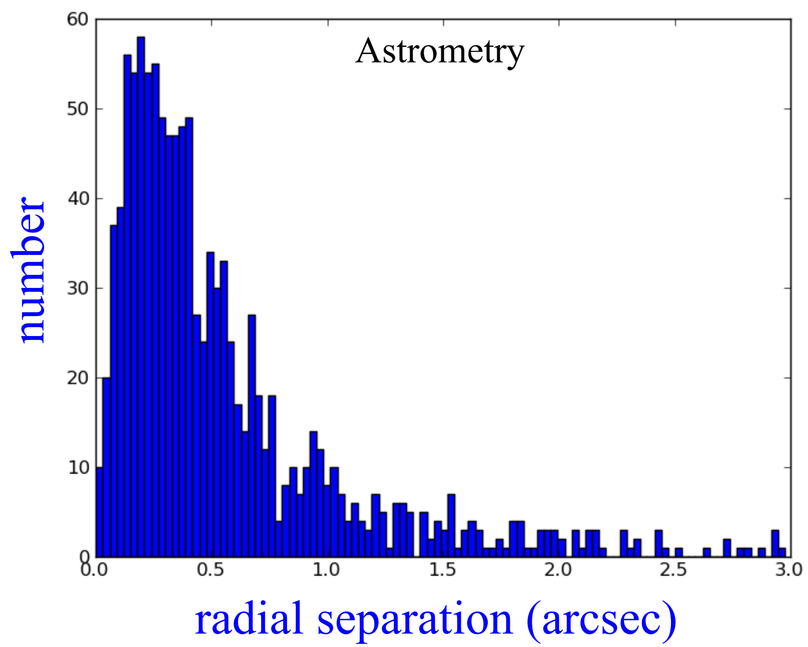
Performance of PSF-fit (AC) photometry

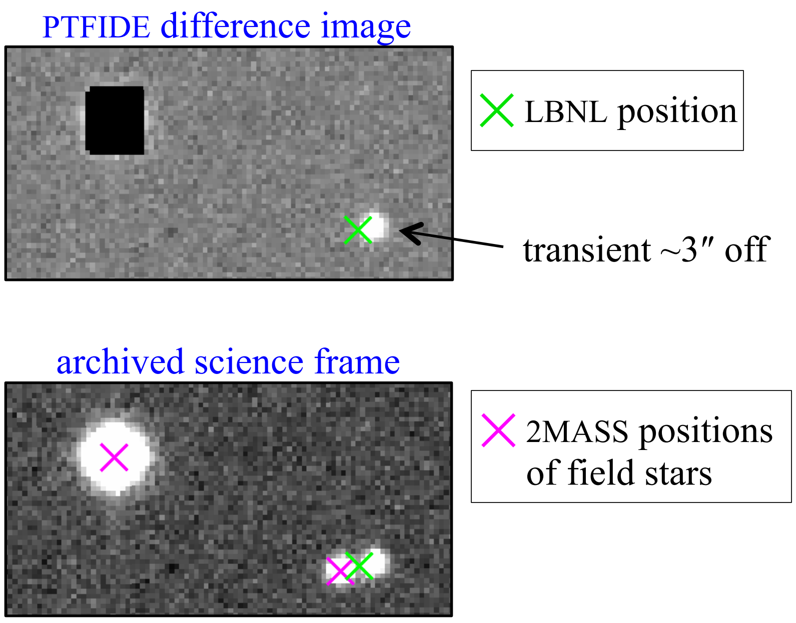
- took ~350 real, moderately dense *R*-band frames, derived spatially-varying PSFs, then simulated point source transients with random positions and fluxes.
- then executed PTFIDE to create diff images and extract candidates.
- difference image (AC) fluxes consistent with truth.



Comparison to transients discovered with LBNL pipeline

- Courtesy: **Alexandra Cong** - California Institute of Technology
  Umaa Rebbapragada - Jet Propulsion Laboratory, California Institute of Technology
- Executed PTFIDE on archival data containing 1549 *R*-band transients discovered: Apr 2009 – Feb 2012.
- Recovered 1182 objects within 3″ of LBNL position; 333 not extracted but do appear in diff images; 34 failed.
- A large fraction of LBNL positions provided to us are not on the actual transient! Analysis needs to be redone.



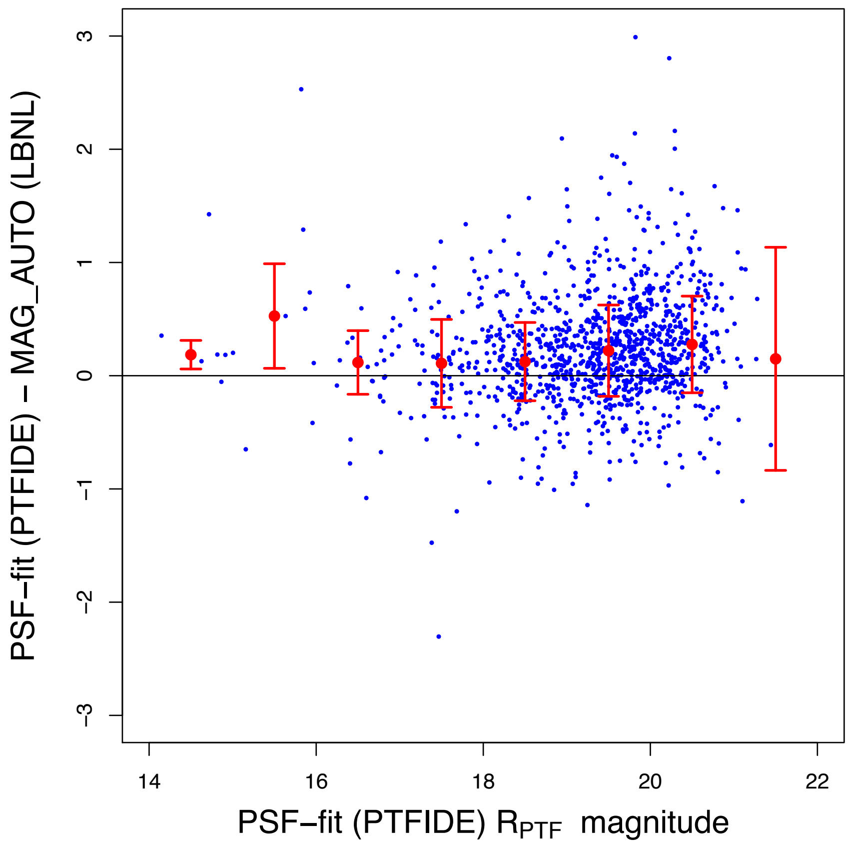


LBNL centroids appear to be from science images => subject to contamination / blending

A portion of the research described in this presentation was carried out at the Jet Propulsion Laboratory under a Research and Technology Development Grant, under contract with the National Aeronautics and Space Administration. US Government Support Acknowledged.

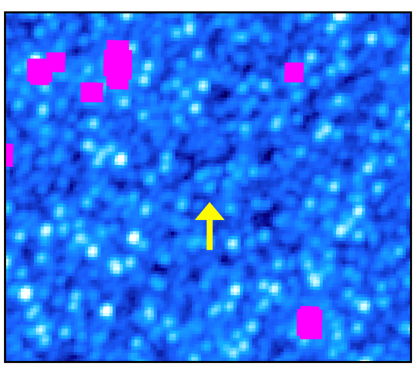
Comparison to transients discovered with LBNL pipeline

Photometry comparison: PSF-fit photometry from PTFIDE vs Kron-like aperture photometry (MAG_AUTO) with SExtractor from LBNL:

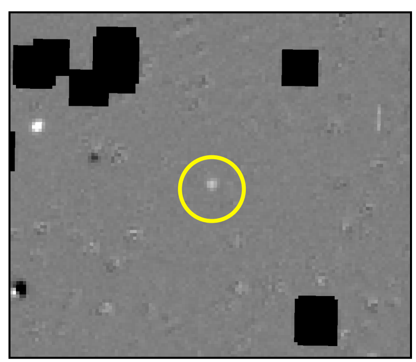


From difference-images to light curves

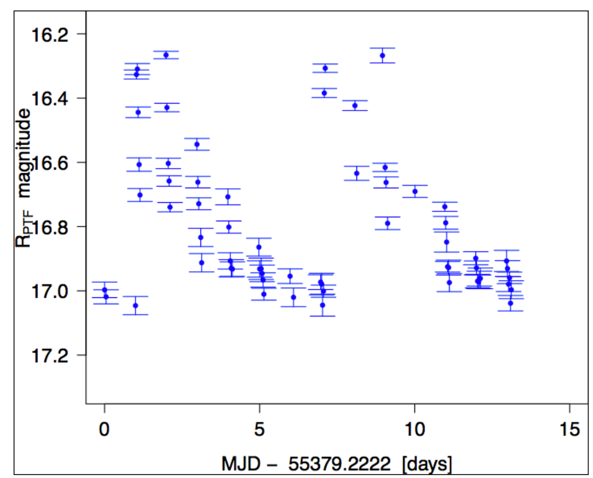
- Difference image photometry and candidate catalogs tied to reference image zero point.
  - derived by gain-matching input frame ZPs initially derived by matching Kron-like aperture phot. to SDSS.
  - calibration infrastructure based on PSF-fit photometry is not yet in place.
- Recall PTFIDE uses a fixed detection threshold (S/N ~ 4) to aid discovery.
- Light curve generation on candidates of interest: use forced PSF-photometry at fixed sky position through stack of difference images with no threshold.
  - enables unbiased measurements down to low S/N; tighter upper limits or better S/N by combining data.
  - implemented as a new pipeline in operations environment.
  - **DC_MAG = 27.0 – 2.5log$_{10}$[DC_flux]** where DC_flux = "AC_flux + RefImg_flux" > *N*-sigma.



galactic bulge field
~ 2.2 arcmin

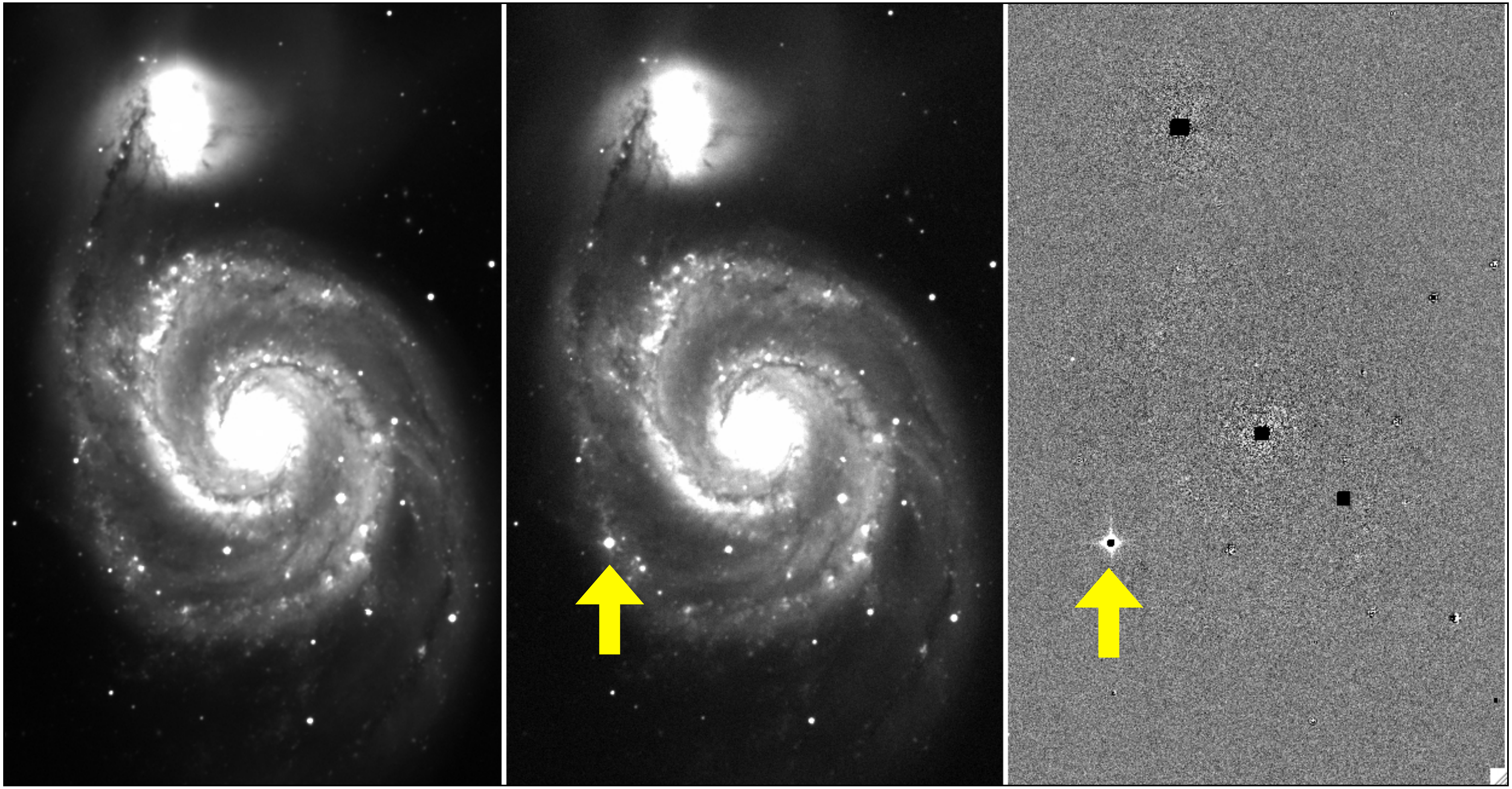


difference image



Cepheid variable (?)

SN 2011dh (PTF11eon) in Messier 51



Reference image = co-add of 20
R exposures (pre-outburst)

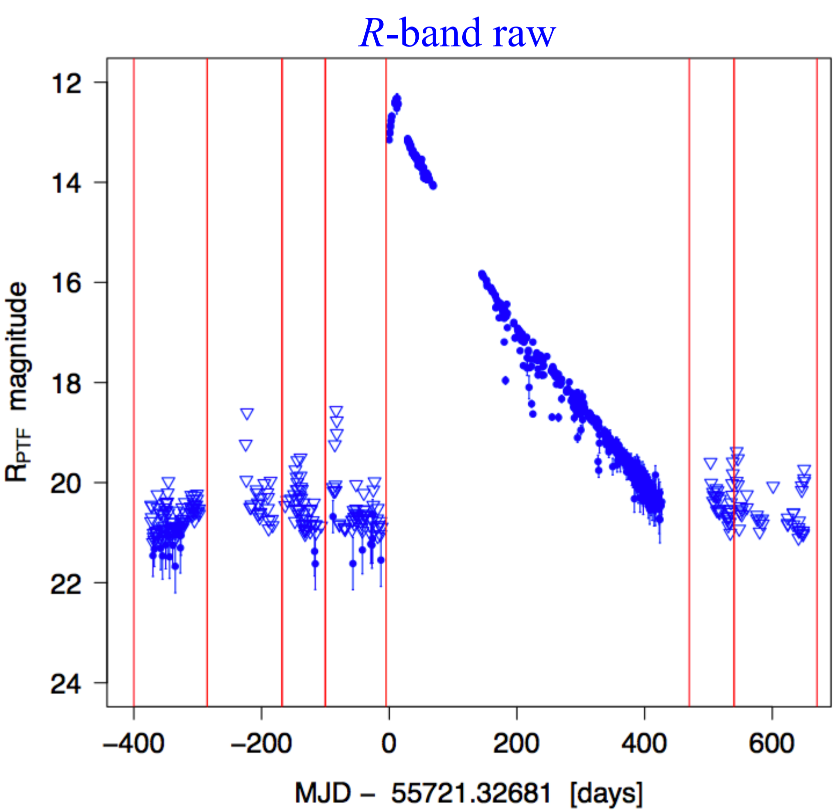
R exposure on June 19, 2011
Type IIb supernova ~ $10^9 L_\odot$

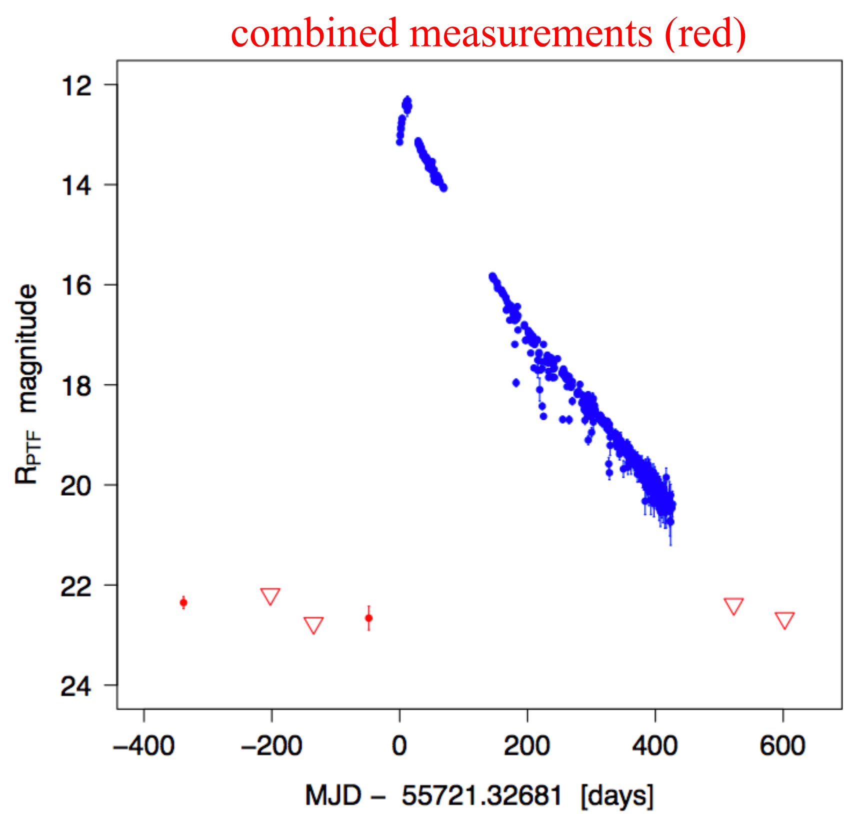
Difference image:
exposure - reference

SN 2011dh *R*-band light-curve from windowed-averaging

- Combine measurements within windows to improve S/N or obtain tighter upper limits on non-detections.
  - ➢ faster than co-adding images!
- **Assumption:** fluxes in a window ~ constant with time.
  - ➢ or can collapse using more complex model based on prior (slope fit)

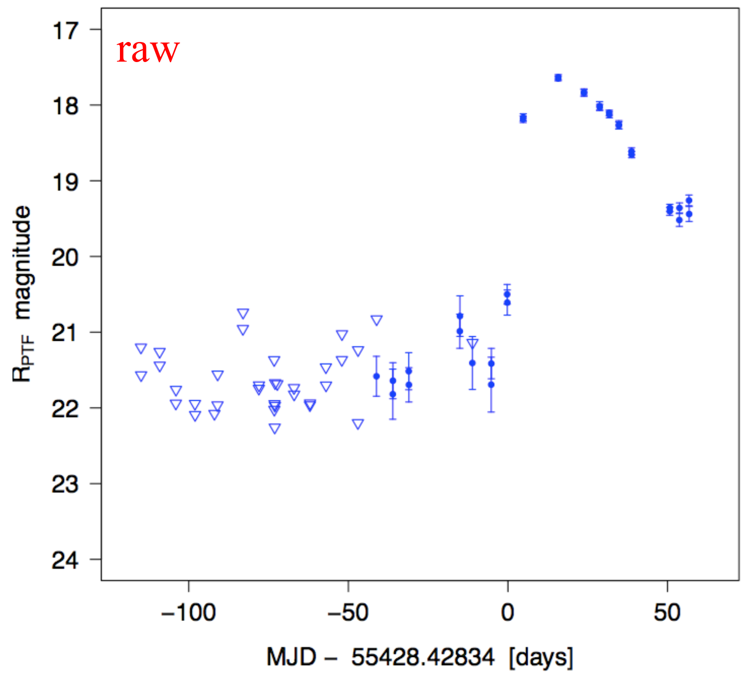
$$f_{coadded} = \frac{\sum_i f_i/\sigma_i^2}{\sum_i 1/\sigma_i^2}$$

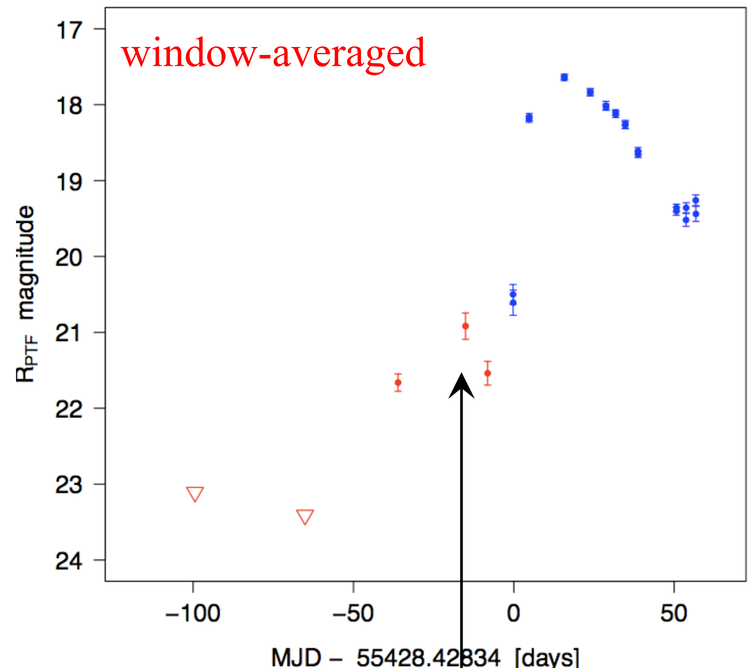


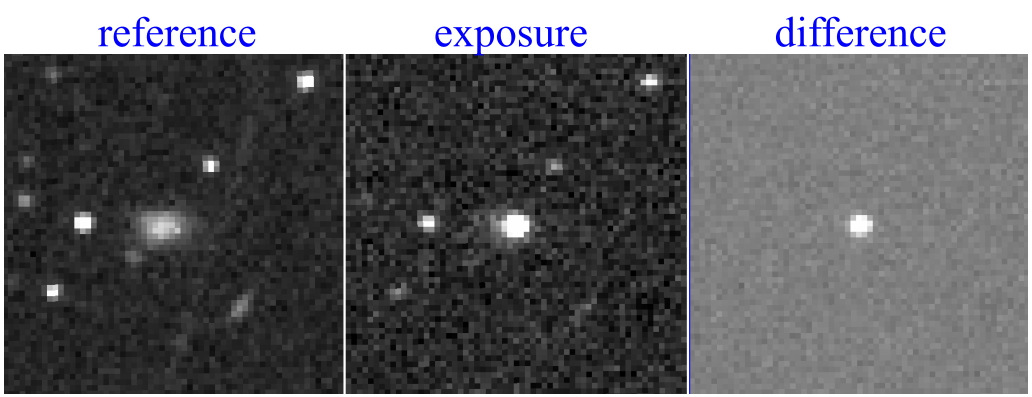


SN 2010mc (PTF10tel)

Type IIn supernova in unknown galaxy at ~ 153 Mpc







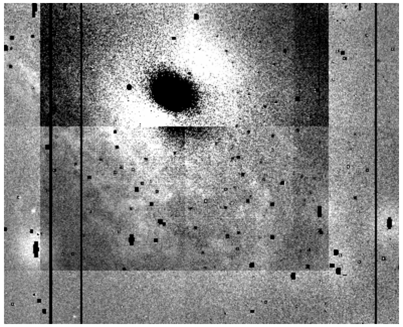
Outburst 40 days before explosion:
Ofek et al. (2013), Nature, 494, 65

Future improvements for PTFIDE

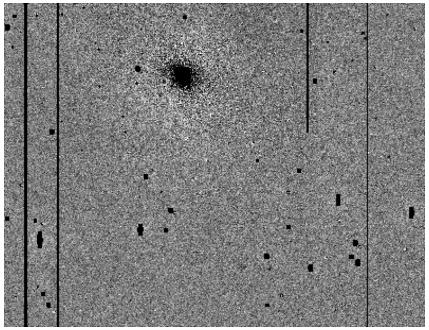
Mostly make PSF-matching more robust:

- Image partitions with complex/extended emission can result in bad diff-img residuals (bad gain matching?): borrow kernel solutions from "good" neighboring partitions or interpolate.

~20′



M31 bulge: BAD difference



M31 bulge: good difference (?)

- Instead of using linear-least squares to estimate PSF-matching kernel *K*, minimize L1-norm:

$$C = \sum_{i,j} \left| I_{ij} - \left(K_{lm} \otimes R_{ij} \right) - dB \right| \qquad \Rightarrow \text{ more robust against outliers}$$

- Regularization tricks to obtain smoother kernel solutions in crowded/noisy image regions (e.g., Becker et al. 2012). Penalize fits that give a high variance for *K* (high second derivative):

$$C = \sum_{i,j} \left[I_{ij} - \left(K_{lm} \otimes R_{ij} \right) - dB \right]^2 + \underbrace{\lambda \nabla^2 K_{lm}}$$

Summary

- A new discovery engine (PTFIDE) is currently in production to support archival research.
  - with forced photometry (post-processing) pipeline for candidates of interest.

- Vetting (real-bogus) infrastructure not yet in place. Validation and testing continues.

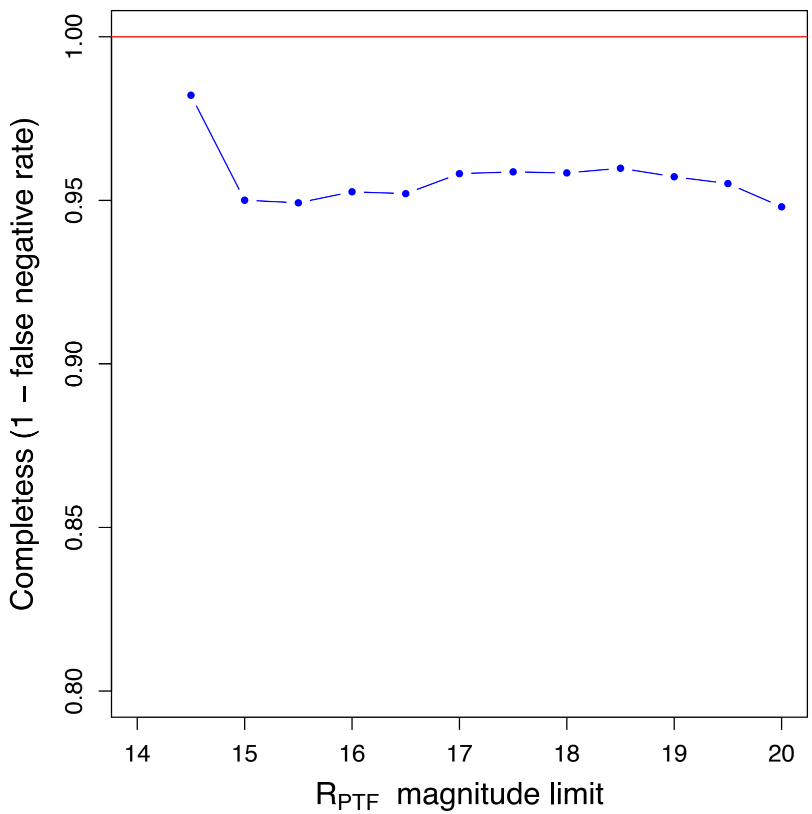
- Good image calibration, reference image quality, flexible PSF-matching are key to obtaining good difference images.

- What matters in the end is the <u>content</u> of the candidate extraction catalogs:
  - use of PSF-fit photometry and associated diagnostics crucial to minimize false positives
  - even if a difference image is not perfect (within random noise), can still proceed

Back up slides

Performance: completeness

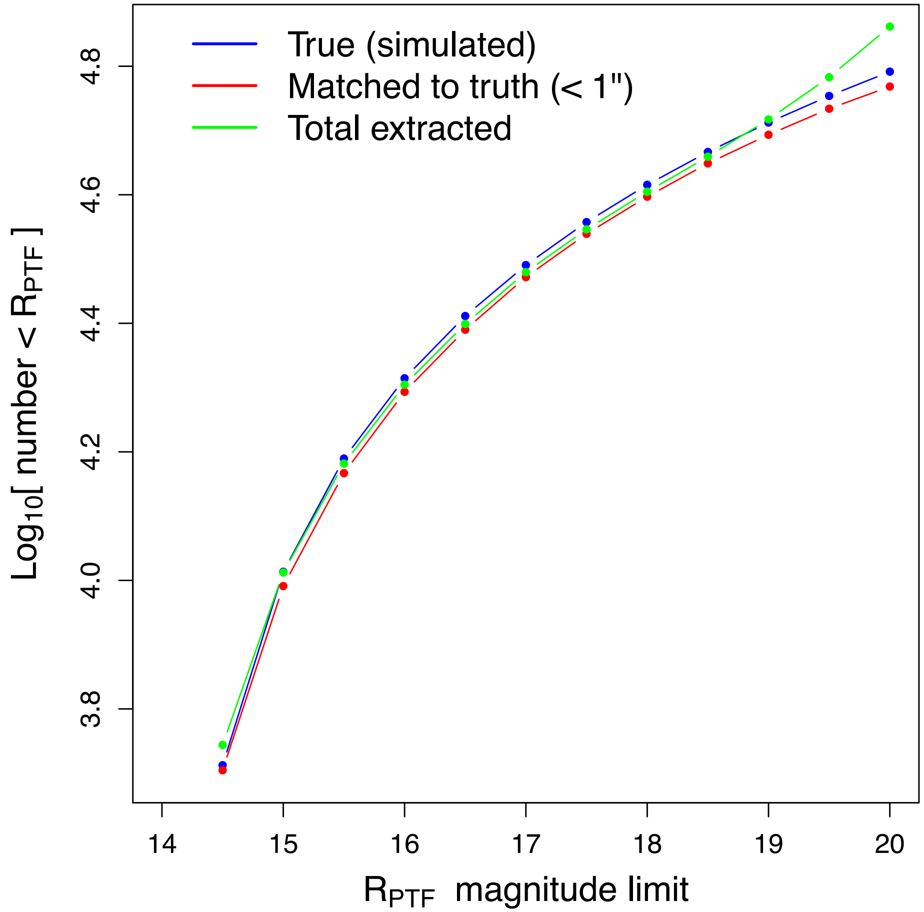
- took ~350 real, moderately dense *R*-band frames, derived spatially-varying PSFs, then simulated point source transients with random positions and fluxes.
- executed PTFIDE to create diff images and extract candidates with **fixed** threshold (S/N = 4) and filter params.



$$C = \frac{\#\,\text{matched to truth}\ (<R)}{\#\,\text{total truth}\ (<R)}$$

Performance: #extractions vs "truth"

- took ~350 real, moderately dense *R*-band frames, derived spatially-varying PSFs, then simulated point source transients with random positions and fluxes.
- executed PTFIDE to create diff images and extract candidates with **fixed** threshold (S/N = 4) and filter params.

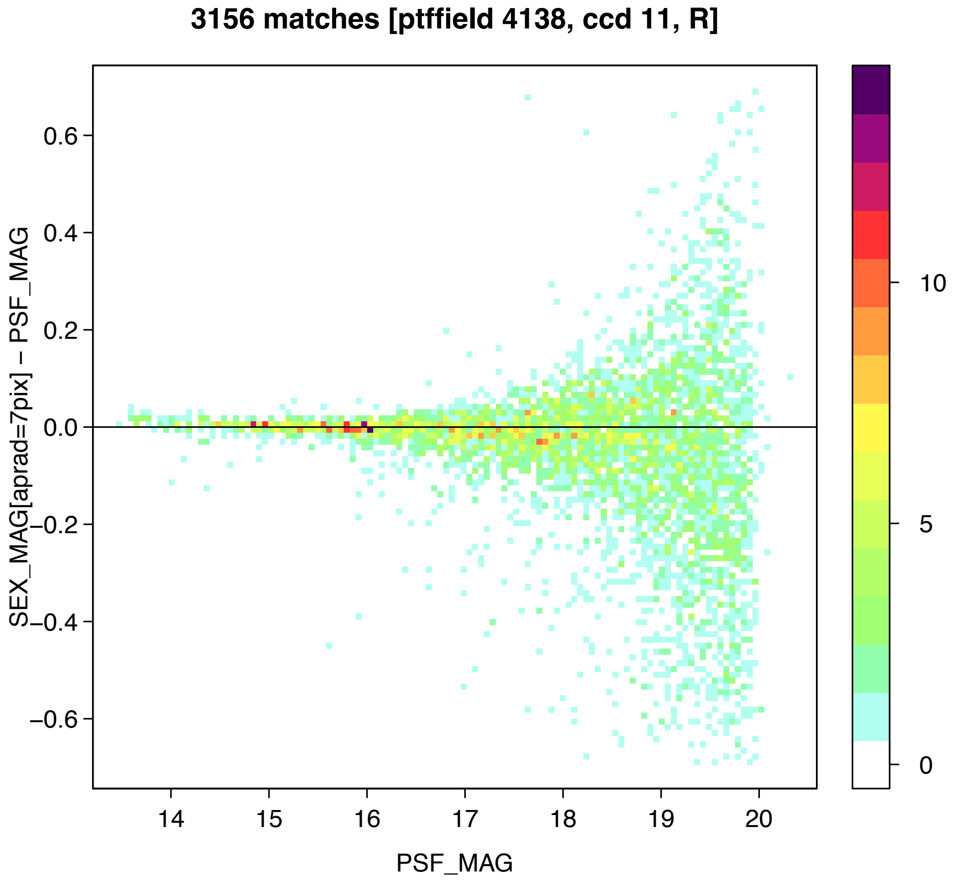


$$R = \frac{\#\,\text{matched to truth}\,(< R)}{\#\,\text{total extracted}\,(< R)}$$

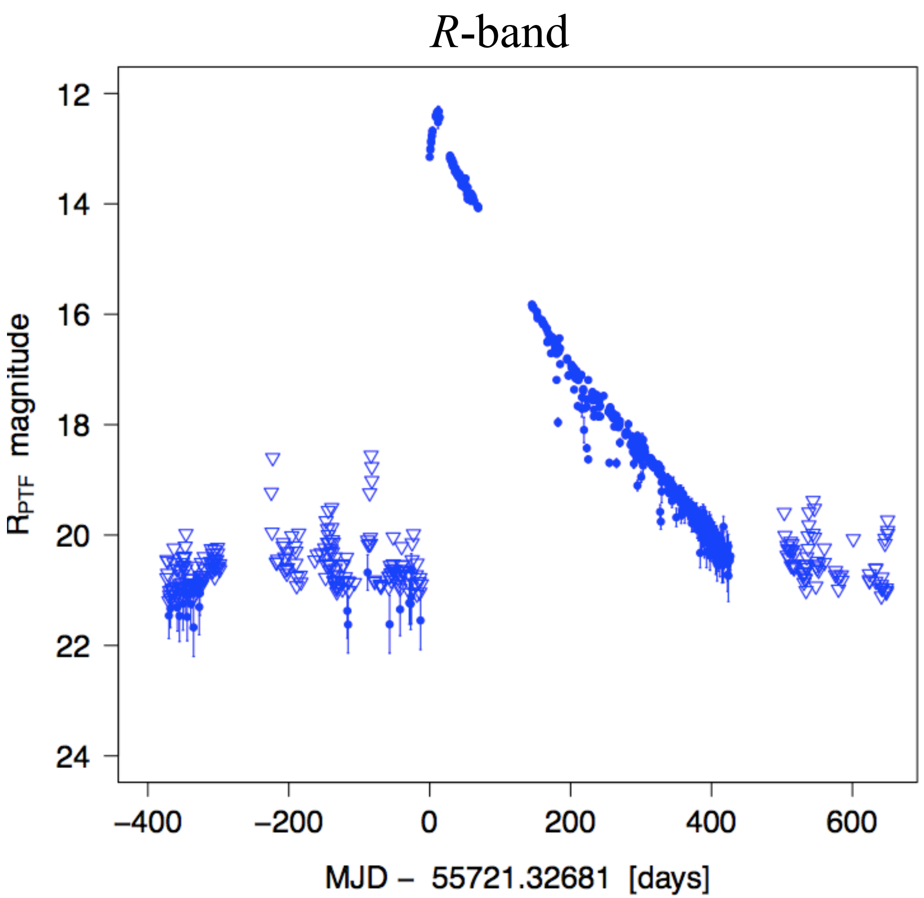
$$C = \frac{\#\,\text{matched to truth}\,(< R)}{\#\,\text{total truth}\,(< R)}$$

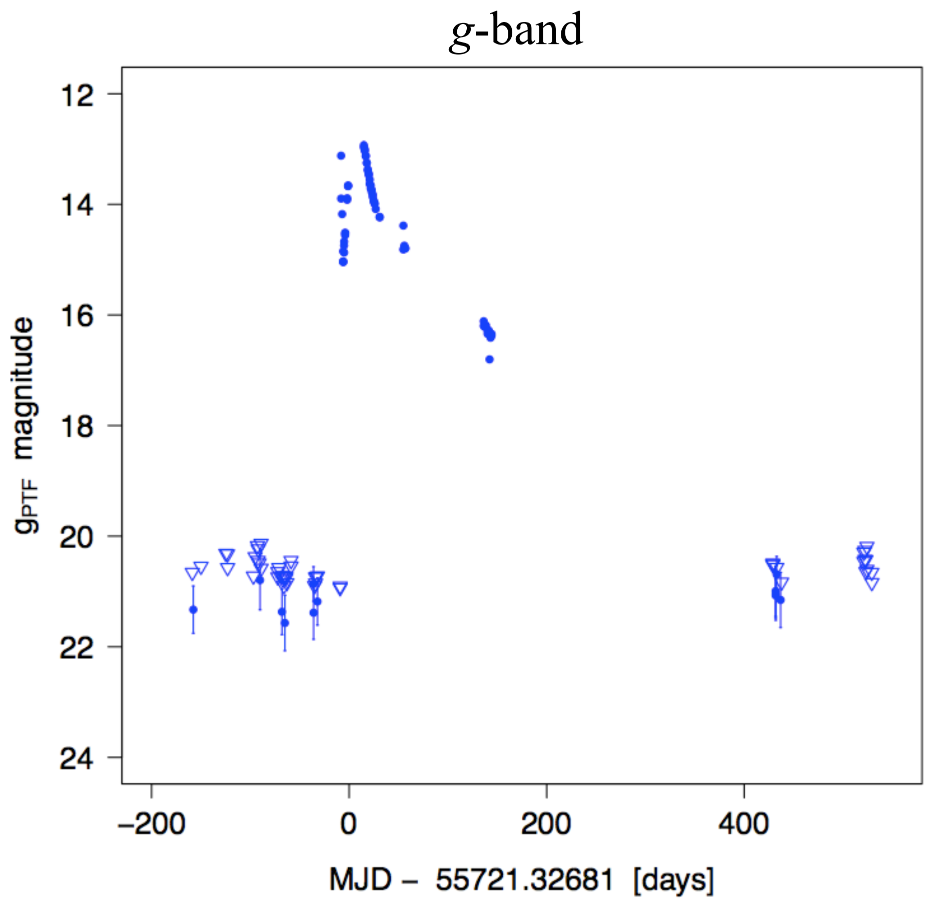
PSF-fit vs SExtractor aperture photometry

- Comparison below is a single science exposure image.
- SExtractor photometry is based on a fixed (relatively large) 7 pixel radius aperture .
- Galaxies filtered; only point sources are compared.



SN 2011dh light curves from PTF difference image photometry



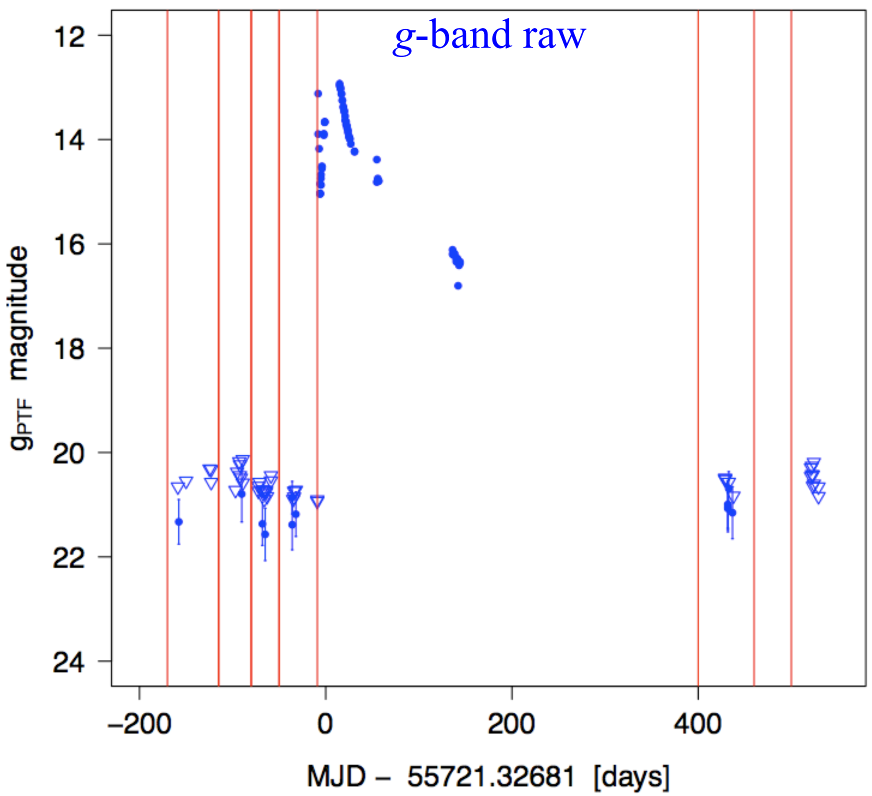


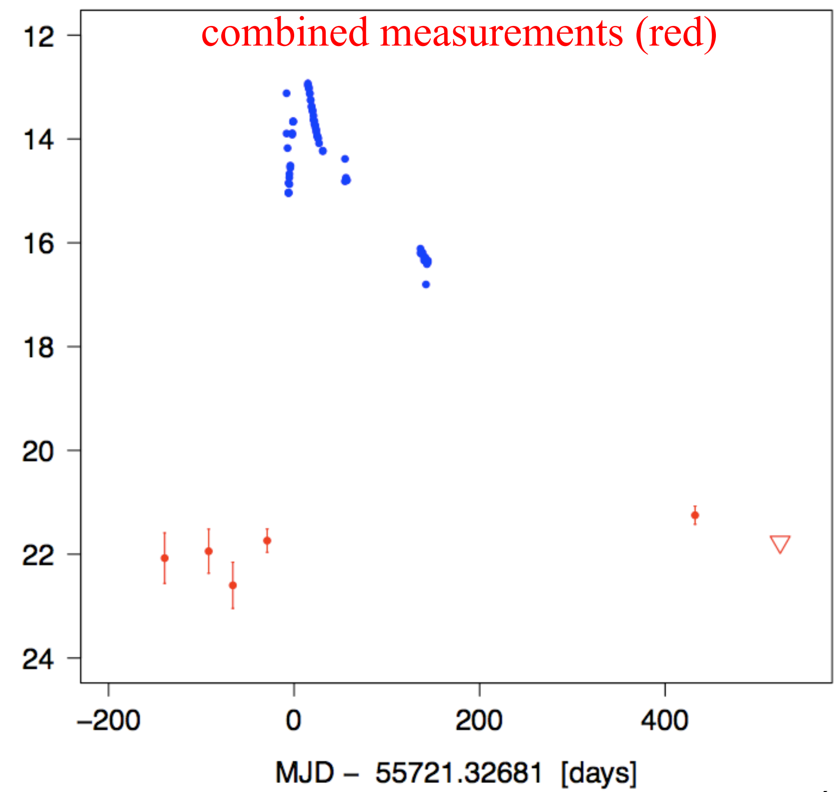
triangles: non-detections shown as 3.5σ upper limits

SN 2011dh *g*-band light-curve from windowed-averaging

- Combine measurements within windows to improve S/N or obtain tighter upper limits on non-detections. Faster than coadding images!
- **Assumption:** fluxes in a window ~ constant with time. Can also collapse using more complex model based on prior (slope fit)
- Reveal any "burst" behavior not seen in lower S/N exposures

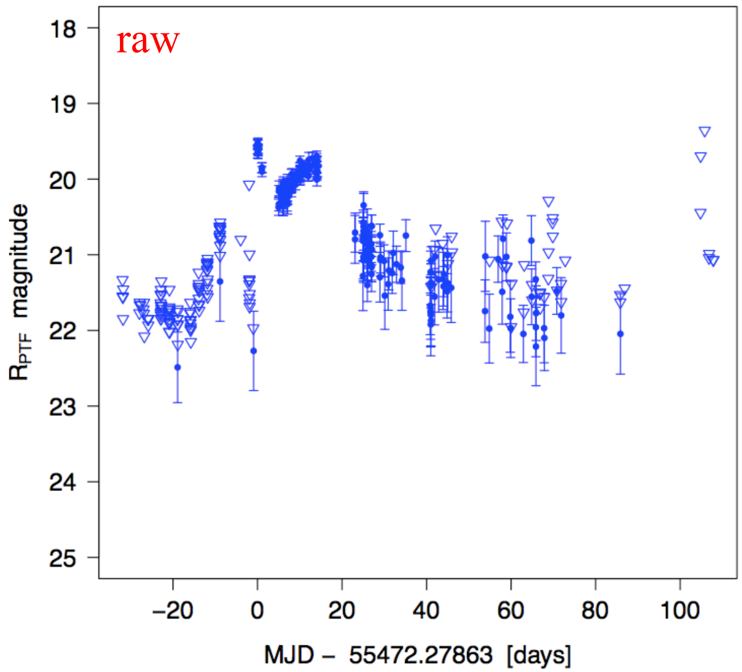
$$f_{coadded} = \frac{\sum_i f_i/\sigma_i^2}{\sum_i 1/\sigma_i^2}$$

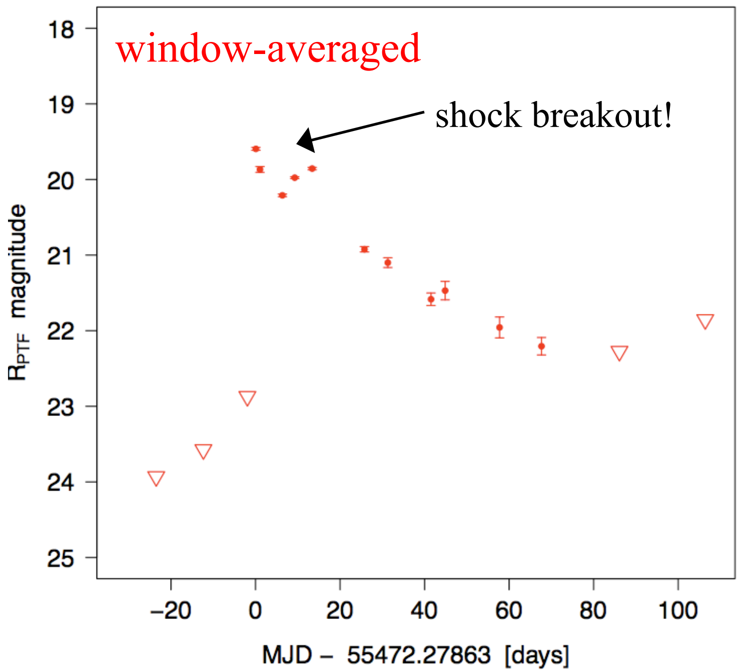


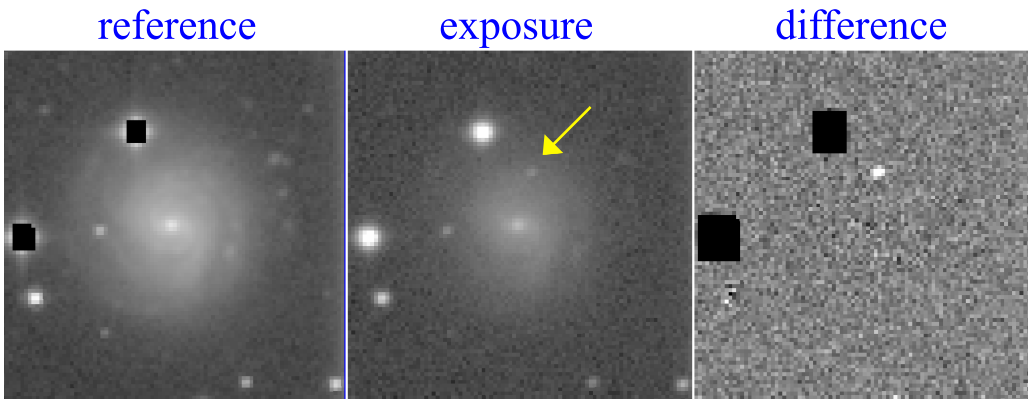


SN PTF10xfh

Type Ic supernova in NGC 717 at ~ 65 Mpc (Yi Cao, private communication)







SN PTF13ai
(or PSN J12541585+0926259)

- Type Ia Supernova in galaxy PGC 43884 (~197 Mpc); discovered Feb 5, 2013
- One of the first to be discovered for iPTF