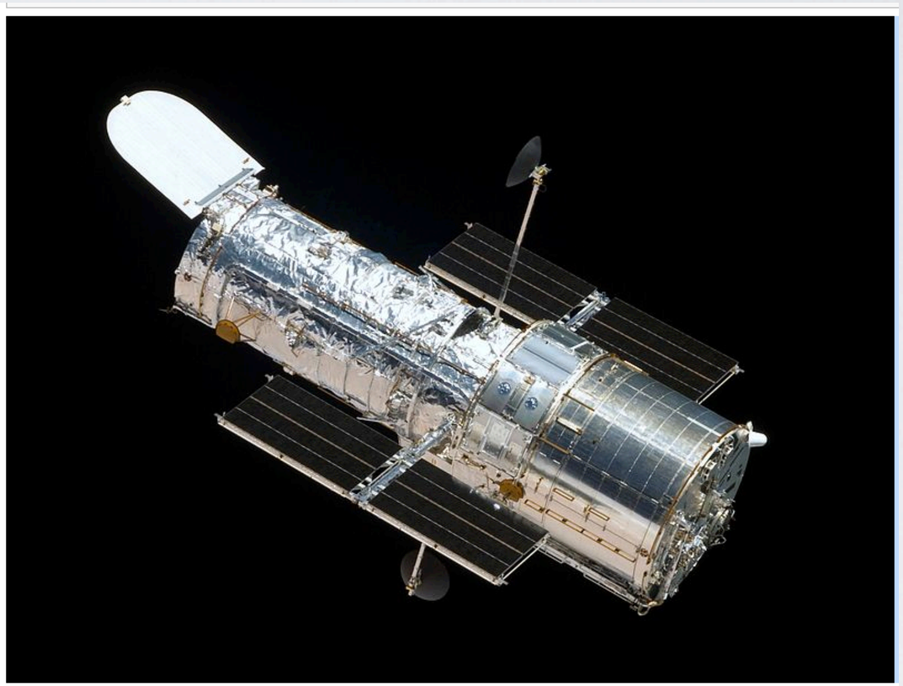
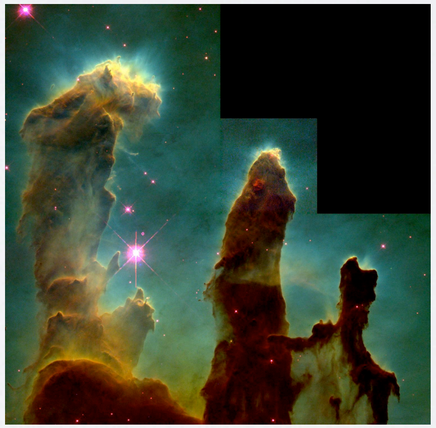
HUBBLE SOURCE CATALOG

Steve Lubow
Space Telescope Science Institute
June 6, 2012

HUBBLE SPACE TELESCOPE

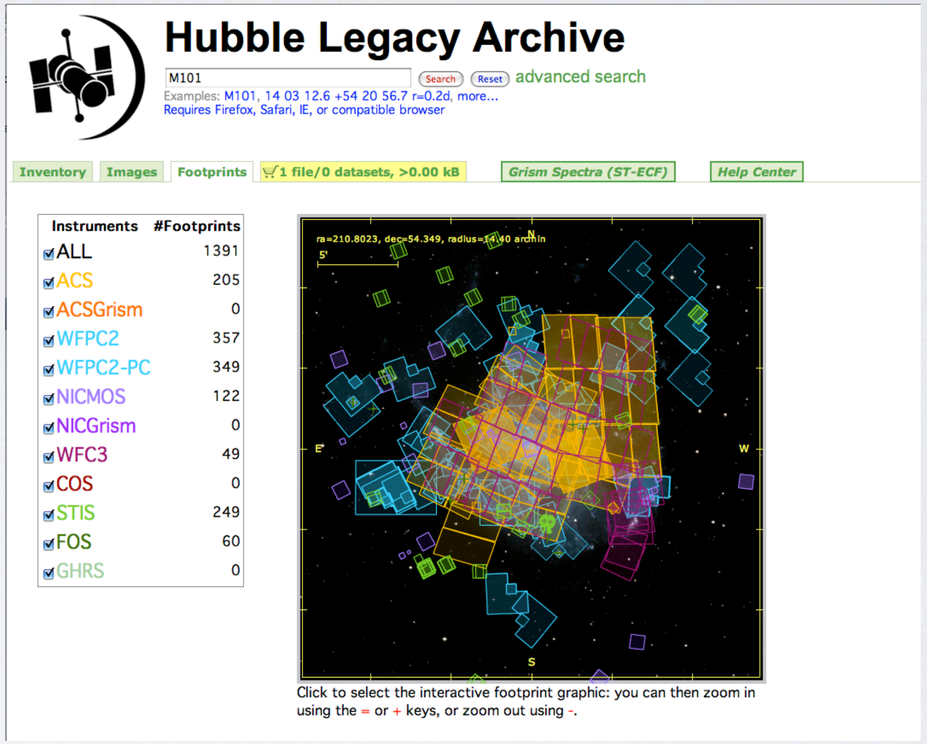
- 2.4 m telescope launched in 1990
- In orbit above Earth's atmosphere
- High resolution, faint sources
- Several cameras. Many filters.
- Taken over 500K science exposures.
- Many individual proposals (> 5000)
- Mission *not* optimized for catalog





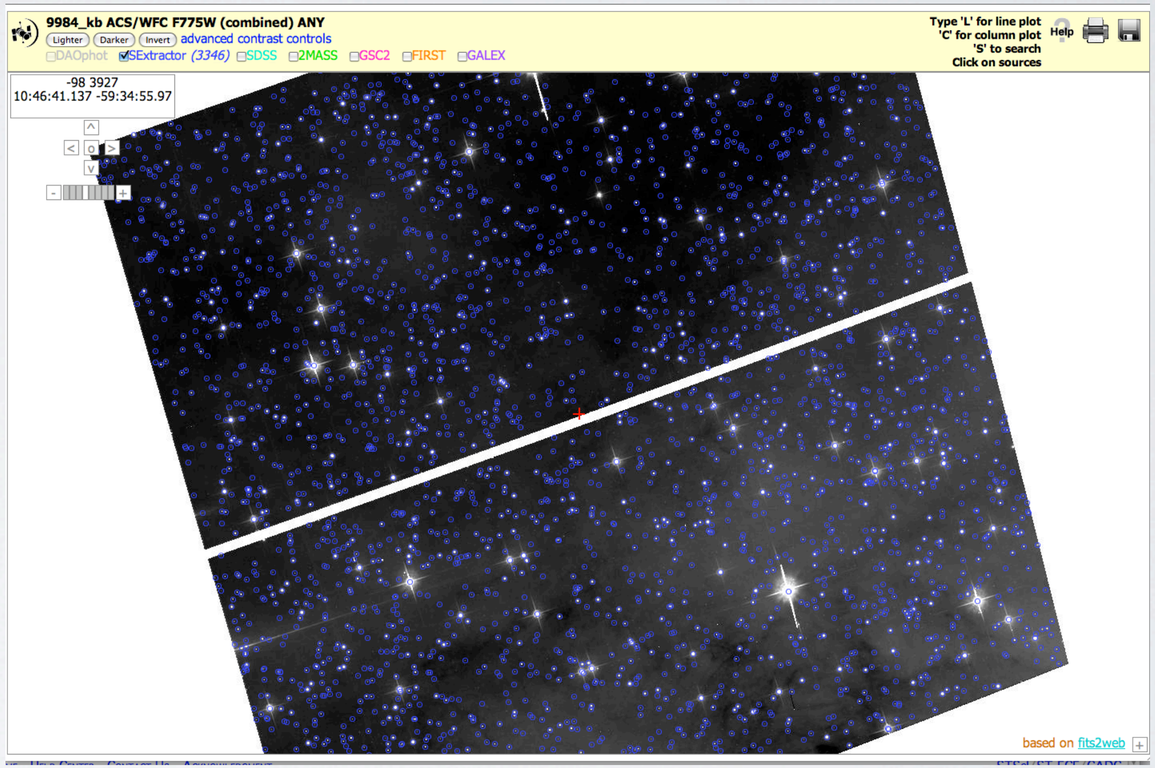
HUBBLE LEGACY ARCHIVE

- Recent web-based archive for Hubble http://hla.stsci.edu/
- Provides search, images, footprints, **`source lists`**



SOURCE LISTS

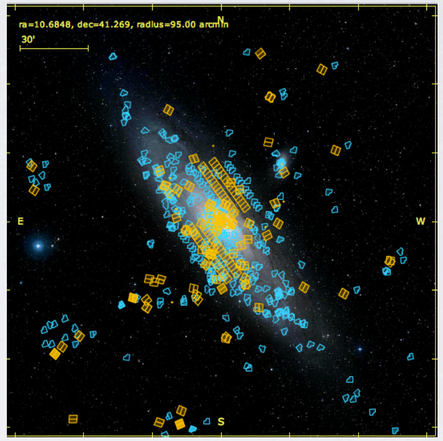
- Images run through Source Extractor software
- Generates properties of sources, including positions stored in a SQL Server database (45 million source detections)
- Want to CROSSMATCH sources at nearly same position in different images
  - Obtain spectral and temporal properties of astronomical objects

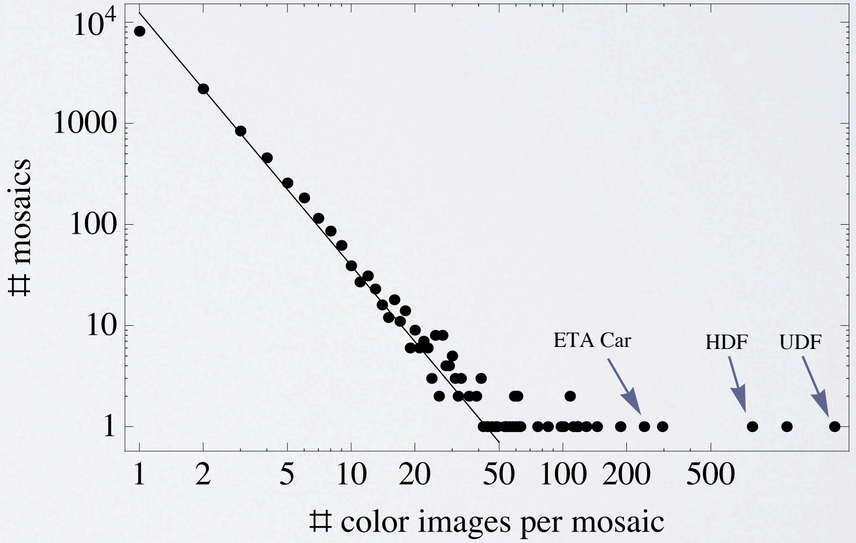


CROSSMATCHING

Budavari & Lubow 2012, Astrophysical Journal, submitted

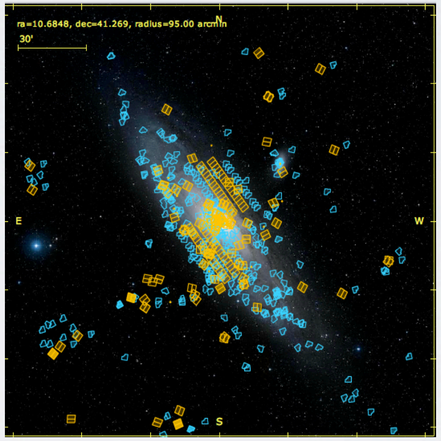
- Challenges
  - Many irregularly placed and timed exposures
  - Sources involve different cameras, filters, and exposure times.
  - Cases with many overlapping exposures (mosaics)
  - Possible false detections

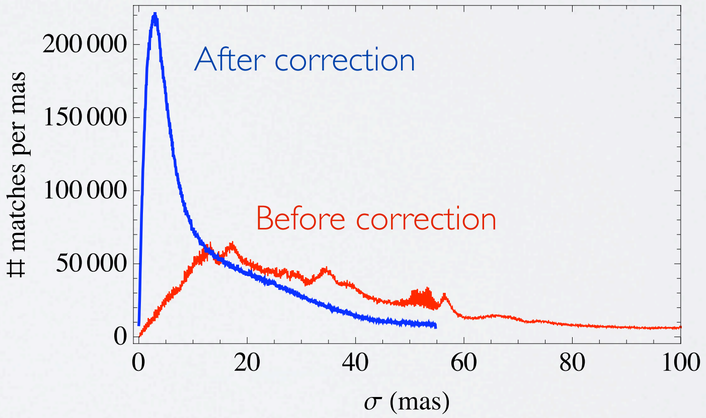




METHODS

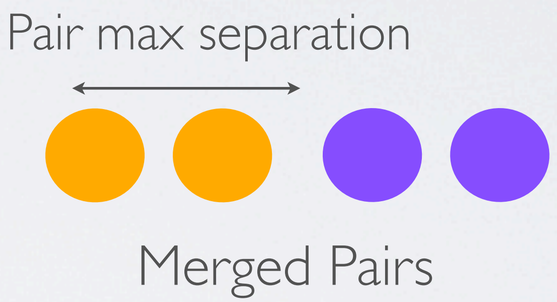
- Determine mosaics by friends of friends
- Determine nearby source pairs with some threshold separation
- Minimize pair separations by adjusting image positions (astrometric correction)
- Select matching sources
- Build catalog





STATISTICAL MATCHING

- Apply FoF for close pairs across images
- Can end up with long chains
- Want to break up matches for best configuration (*chainbreaker* tool)
  - How to measure goodness?
  - Which configs should be considered?



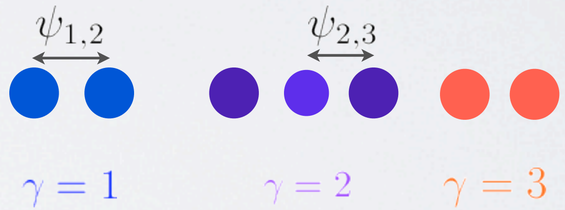
LIKELIHOOD

- Bayesian approach
- Adopt Fisher (Gaussian) distribution of positions for a source i on a unit sphere with some uncertainty σ_i
- Determine ratio of probabilities of some split config relative to a baseline G2 with FoF matching

$$p(D|G) = 2^{n-\Gamma}\left(\prod_{i=1}^{n} w_i \frac{\delta(r_i - 1)}{4\pi}\right)\prod_{\gamma=1}^{\Gamma} \frac{\exp\left(\frac{\sum w_i w_j \psi_{ij}^2}{-4\sum w_i}\right)}{\sum w_i}$$

$$w_i = 1/\sigma_i^2$$

Example $n = 7$ $\Gamma = 3$

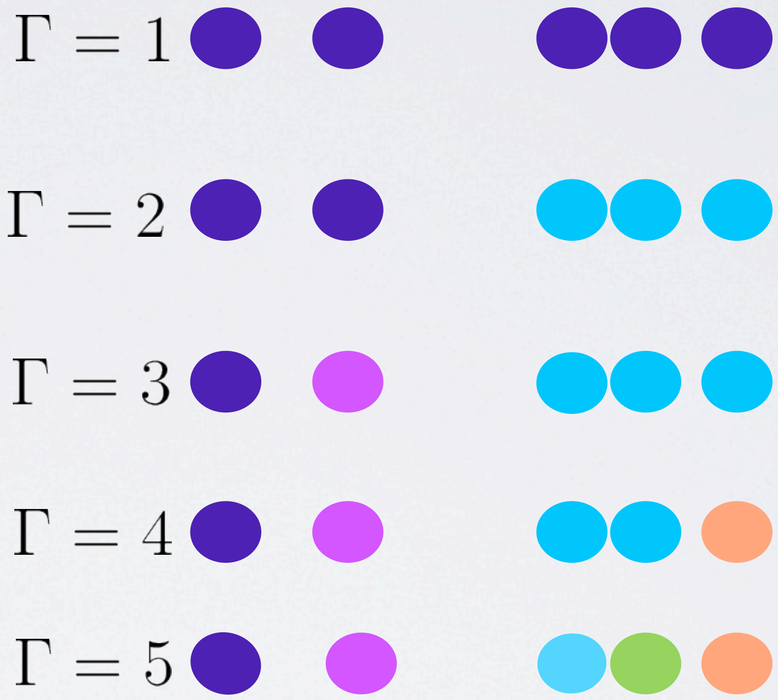


$$B_{12} = \frac{p(D|G_1)}{p(D|G_2)}$$

CONFIGURATIONS

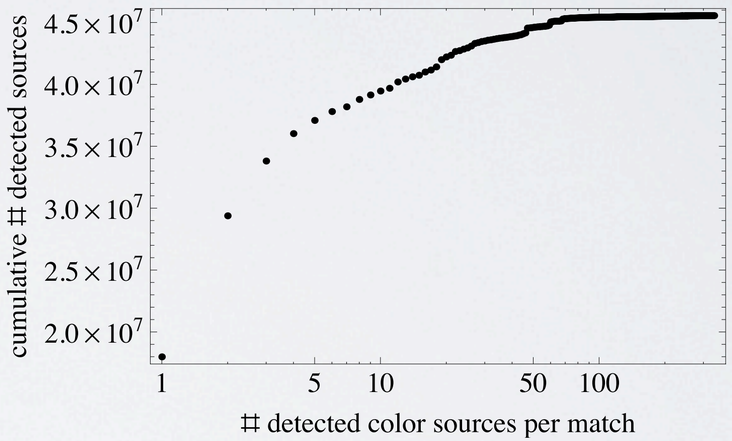
- Cannot consider all possible cases
- Apply greedy algorithm that reduces max separations of pairs in each step
  - Split pair with largest separation
  - Store in DB if likelihood is favorable
  - Spit the split config again
  - Apply above steps to each subset
  - Iterate until split to one source per subset
  - select best config: highest likelihood

Example n=5



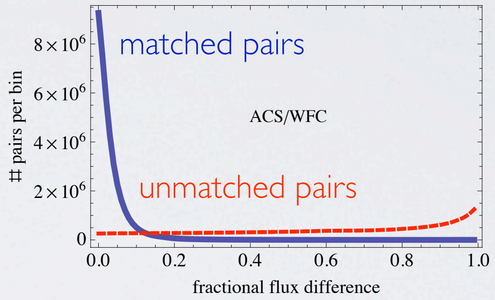
IMPLEMENTATION

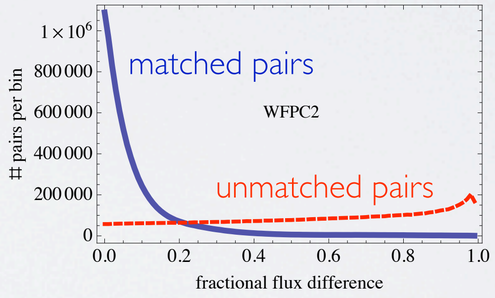
- SQL Server database, including extended stored procedures
- Considered 2 detectors: WFPC2 and ACS/WFC
- Took a few days to build catalog of 60 million sources (attempted 45 million sources for cross matching)

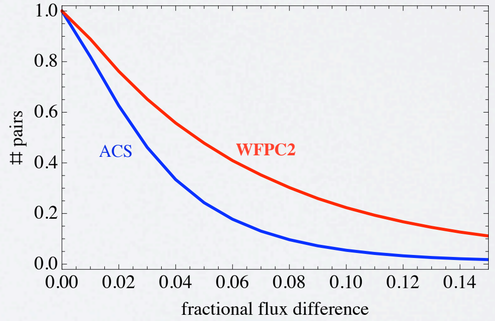


VERIFYING RESULTS

- How can we tell whether matches are good?
- Matches determined only by position
- Select an independent measure: source flux
- Do pairs within matches with same detector and filter have same flux?
- Compare with unmatched case







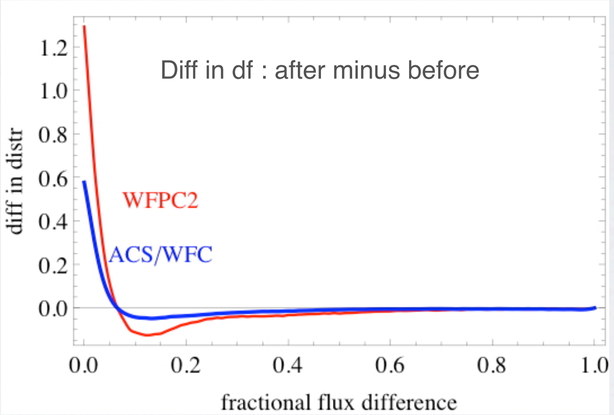
$$f_{i,j} = \frac{|F_i - F_j|}{max(F_i, F_j)}$$

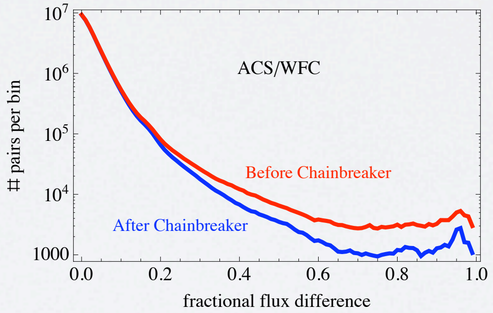
VERIFYING CHAINBREAKING

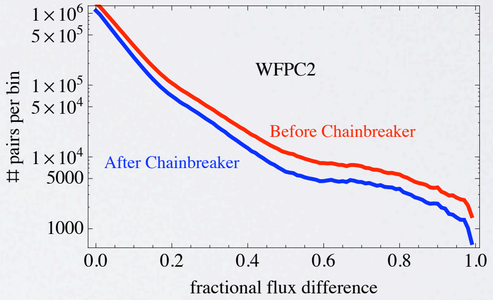
- Verify improvement of chainbreaker
- Compared *f* distribution before and after chainbreaker

$$f_{i,j} = \frac{|F_i - F_j|}{max(F_i, F_j)}$$

- Main effect: reduces tail



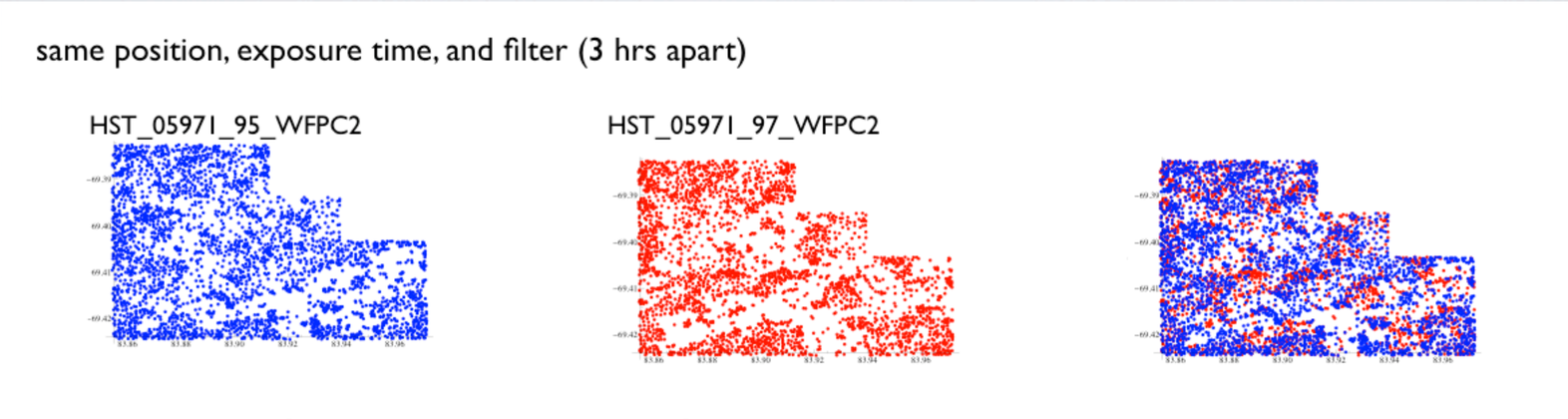




NONDECTIONS

- Procedure for building source lists
  - Build multi-filter image
  - find sources in each filer
- Can then determine nondetections in certain filters at certain positions
- What are the detection limits?

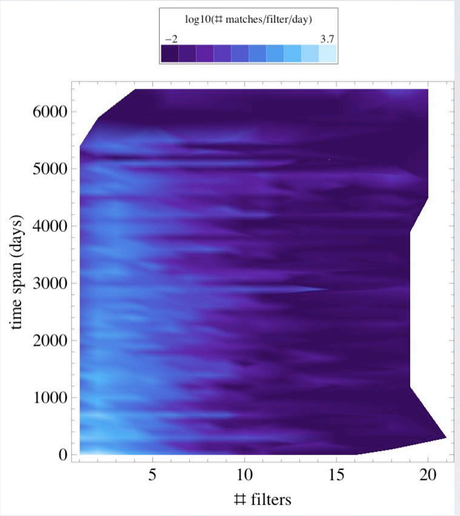
FALSE DETECTIONS



- Source detection software not perfect
- Many cases of false detections
- Find by looking for images with similar positions, exposure times, etc., but different number of detected sources
- Do not expect many false detections in crossmatched sources

SUMMARY

- Built a catalog for HST with spectral/temporal info on astronomical objects
- Cross matching sources from HST involves statistical challenges
  - Build maximal sets of matches
  - Break them up
  - Results improve tail of distribution
- Other challenges:
  - nondetections
  - false detections



Mission Search / Missions / Contacts / STScI / MAST

Columns Help / Archive Status

HLA_CAT Search Results

Edit Query

Display numeric columns graphically using VOPlot

number of rows returned = 100

Match ID	Mem ID	Det	Target Name	Image Name	Match RA (J2000)	Match DEC (J2000)	D	D Sigma	Mode	Aperture	Exp Time	Obs. Start Time	Obs. Stop Time	FluxAper2
17383311	1	Y	NGC5457	hst_08591_27_wfpc2_f547m_wf	14 03 13.56	+54 21 06.9	3.405	4.815	IMAGE	PC1	1600	2001-04-22 05:42:14	2001-04-22 07:22:54	3.565
17383311	2	Y	ANY	HST_9720_02_ACS_WFC_F658N	14 03 13.56	+54 21 06.9	3.405	4.815	ACCUM	WFC	2430	2003-08-27 07:44:47	2003-08-27 08:27:35	1.219
17383311	3	N	NGC5457	hst_08591_27_wfpc2_f656n_wf	14 03 13.56	+54 21 06.9	3.405	4.815	IMAGE	PC1	2100	2001-04-22 07:27:14	2001-04-22 09:26:34	
17383328	1	Y	NGC5457	hst_08591_27_wfpc2_f547m_wf	14 03 14.38	+54 21 10.3	17.107	24.192	IMAGE	PC1	1600	2001-04-22 05:42:14	2001-04-22 07:22:54	7.880
17383328	2	Y	ANY	HST_9720_02_ACS_WFC_F658N	14 03 14.38	+54 21 10.3	17.107	24.192	ACCUM	WFC	2430	2003-08-27 07:44:47	2003-08-27 08:27:35	2.703
17383328	3	N	NGC5457	hst_08591_27_wfpc2_f656n_wf	14 03 14.38	+54 21 10.3	17.107	24.192	IMAGE	PC1	2100	2001-04-22 07:27:14	2001-04-22 09:26:34	
17383274	1	Y	NGC5457-7	HST_9490_02_ACS_WFC_F435W	14 03 12.85	+54 21 00.5	5.366	7.589	ACCUM	WFC	900	2002-11-13 17:20:48	2002-11-13 17:38:05	5.802
17383274	2	Y	NGC5457-7	HST_9490_02_ACS_WFC_F555W	14 03 12.85	+54 21 00.5	5.366	7.589	ACCUM	WFC	720	2002-11-13 17:41:08	2002-11-13 17:55:25	12.011
17383274	3	Y	NGC5457-7	HST_9490_02_ACS_WFC_F814W	14 03 12.85	+54 21 00.5	5.366	7.589	ACCUM	WFC	720	2002-11-13 17:58:19	2002-11-13 18:12:36	27.749
17383274	4	Y	ANY	HST_9720_02_ACS_WFC_F658N	14 03 12.85	+54 21 00.5	5.366	7.589	ACCUM	WFC	2430	2003-08-27 07:44:47	2003-08-27 08:27:35	0.897
17383177	1	Y	M101-POS1	hst_09720_01_wfpc2_f336w_wf	14 03 12.65	+54 20 58.4	36.735	51.951	IMAGE	WFALL-FIX	2400	2004-02-10 00:09:15	2004-02-10 00:52:15	1.330
17383177	2	Y	NGC5457-1	HST_9490_01_ACS_WFC_F435W	14 03 12.65	+54 20 58.4	36.735	51.951	ACCUM	WFC	900	2002-11-15 22:45:54	2002-11-15 23:03:11	6.232
17383177	3	Y	NGC5457-1	HST_9490_01_ACS_WFC_F555W	14 03 12.65	+54 20 58.4	36.735	51.951	ACCUM	WFC	720	2002-11-15 23:06:14	2002-11-15 23:20:31	10.490
17383177	4	Y	NGC5457-1	HST_9490_01_ACS_WFC_F814W	14 03 12.65	+54 20 58.4	36.735	51.951	ACCUM	WFC	720	2002-11-15 23:23:25	2002-11-15 23:37:42	31.482
17383326	1	Y	NGC5457	hst_08591_27_wfpc2_f547m_wf	14 03 14.88	+54 21 10.4	3.882	5.491	IMAGE	PC1	1600	2001-04-22 05:42:14	2001-04-22 07:22:54	14.675
17383326	2	Y	NGC5457	hst_08591_27_wfpc2_f656n_wf	14 03 14.88	+54 21 10.4	3.882	5.491	IMAGE	PC1	2100	2001-04-22 07:27:14	2001-04-22 09:26:34	0.678
17383326	3	Y	ANY	HST_9720_02_ACS_WFC_F658N	14 03 14.88	+54 21 10.4	3.882	5.491	ACCUM	WFC	2430	2003-08-27 07:44:47	2003-08-27 08:27:35	5.905
17383331	1	Y	NGC5457	hst_08591_27_wfpc2_f547m_wf	14 03 14.89	+54 21 11.4	2.878	4.071	IMAGE	PC1	1600	2001-04-22 05:42:14	2001-04-22 07:22:54	11.692
17383331	2	Y	ANY	HST_9720_02_ACS_WFC_F658N	14 03 14.89	+54 21 11.4	2.878	4.071	ACCUM	WFC	2430	2003-08-27 07:44:47	2003-08-27 08:27:35	4.625
17383331	3	N	NGC5457	hst_08591_27_wfpc2_f656n_wf	14 03 14.89	+54 21 11.4	2.878	4.071	IMAGE	PC1	2100	2001-04-22 07:27:14	2001-04-22 09:26:34	