

The Hubble Source Catalog: A Progress Report

Brad Whitmore

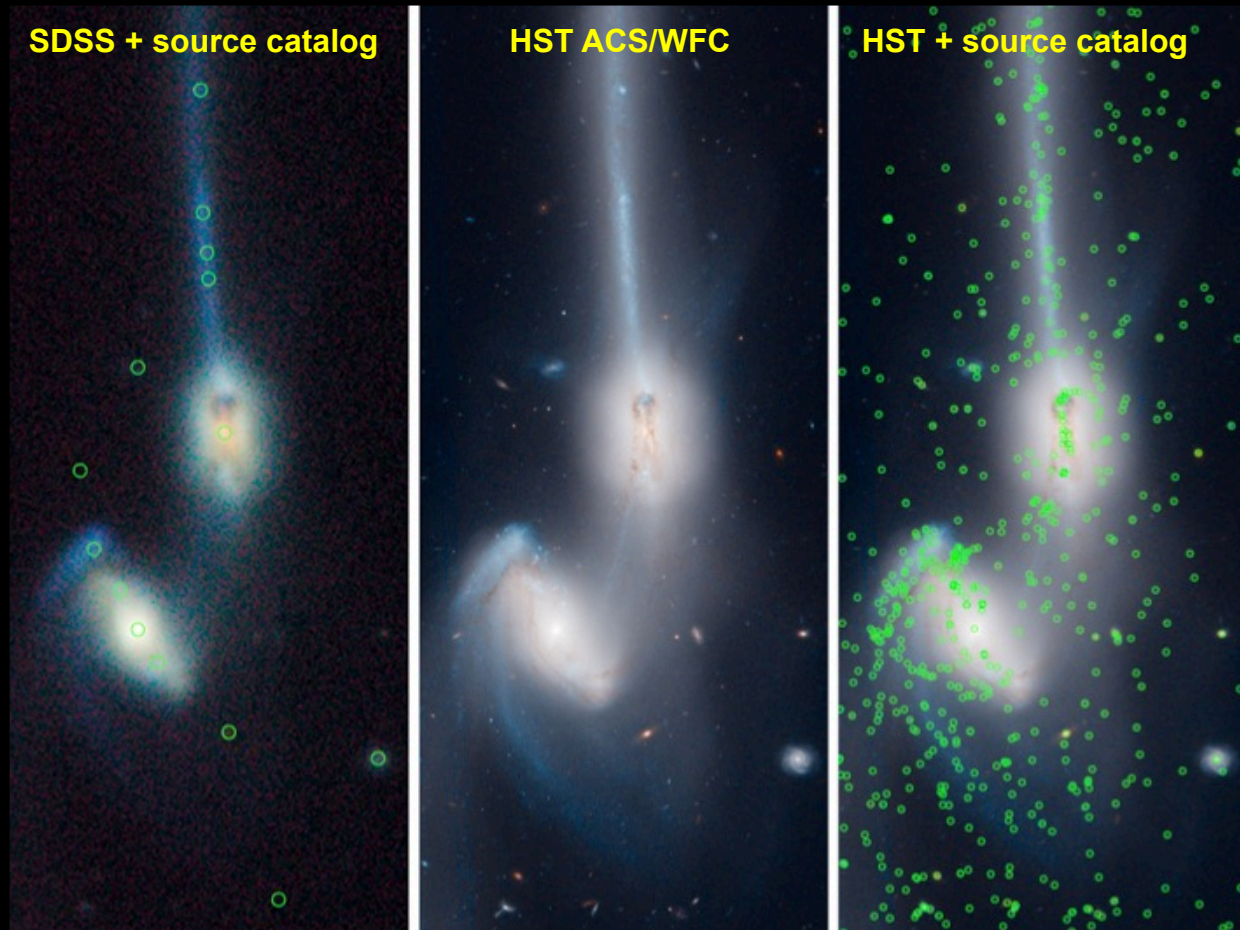
(with Stefano Casertano, Anton Koekemoer,
Steve Lubow, Rick White)

STUC Meeting, Nov. 8, 2012

1. Background and Motivation
2. Methodology: Use Cases > Requirements
3. Progress
4. Timeline

HST Source Catalog

- Easily accessible catalogs of objects are a mainstay of astronomy.
- Example: The SDSS source catalog is a primary catalyst for SDSS science productivity.
- HST Source Catalog
 - Increases the scientific return of the archive for decades to come
 - Fundamental reference for JWST science
 - Complements upcoming surveys (PanSTARRS, LSST)



See April 2012 STUC presentation and white paper for background material.

High Level Science Requirements

1. Provide online (downloadable) search capability for photometry of a target (or list of targets) using **WFPC2, ACS, WFC3** source lists from the HLA (~90 % completeness level, e.g., no source lists for ACS SBC).

[Stretch - **Include other HST instruments?** e.g., NICMOS]

2. Provide averaged magnitudes, number of repeats, and RMS for each filter and instrument combination. Also provide detailed version of form that shows all data.

3. Include **temporal information**.

4. Provide **cross reference** between HSC sources and SDSS, 2MASS, Spitzer, Chandra, GALAX catalogs. Include matches with **existing HST spectral extractions** (e.g., NICMOS and ACS grisms)

[Stretch - Include matches with future HST spectral extractions.]

5. Develop tools for **easy access and quick-look analysis** (e.g., correlations, light curves, ...

[Stretch - Relational search capabilities (e.g., all Sc galaxies from Messier catalog with UBV observations longer than 500 sec).]

6. Provide uncertainties for all measurements (accurate to ~ 25 %).

High Level Science Requirements – cont.

7. **Involve the community** (e.g., small working groups) to insure the final products meet the needs of research scientists.
8. **Document**/advertise the HSC (e.g, algorithms, comparisons with existing standard datasets, **PASP paper**, workshops, ...).
9. Develop **use cases and detailed requirements** (e.g., accuracies, % of artifacts, completeness estimation, ...)
10. Promote the development of value-added catalogs from the HSC (e.g., compact star clusters).

Some questions under discussion:

1. DAOPHOT and/or SExtractor catalogs (i.e., the HLA currently has both, but only SExtractor catalogs are currently used by HSC) ?
2. Output options (e.g., include SQL search) ?
3. Versioning (e.g., do we need to save all old versions of software or just products) ?
4. Degree to which mosaics are incorporated.

How will people use the HSC ?

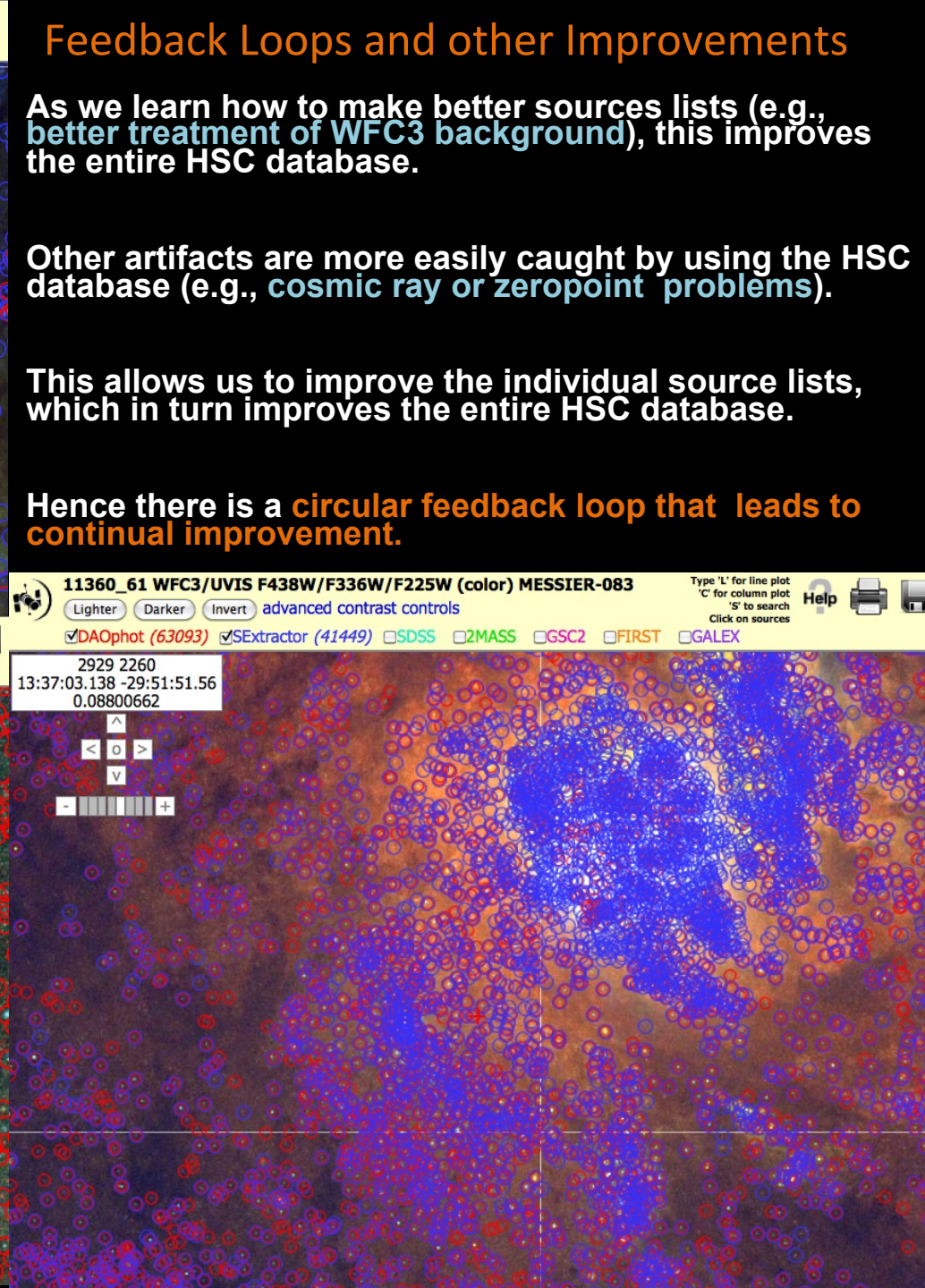
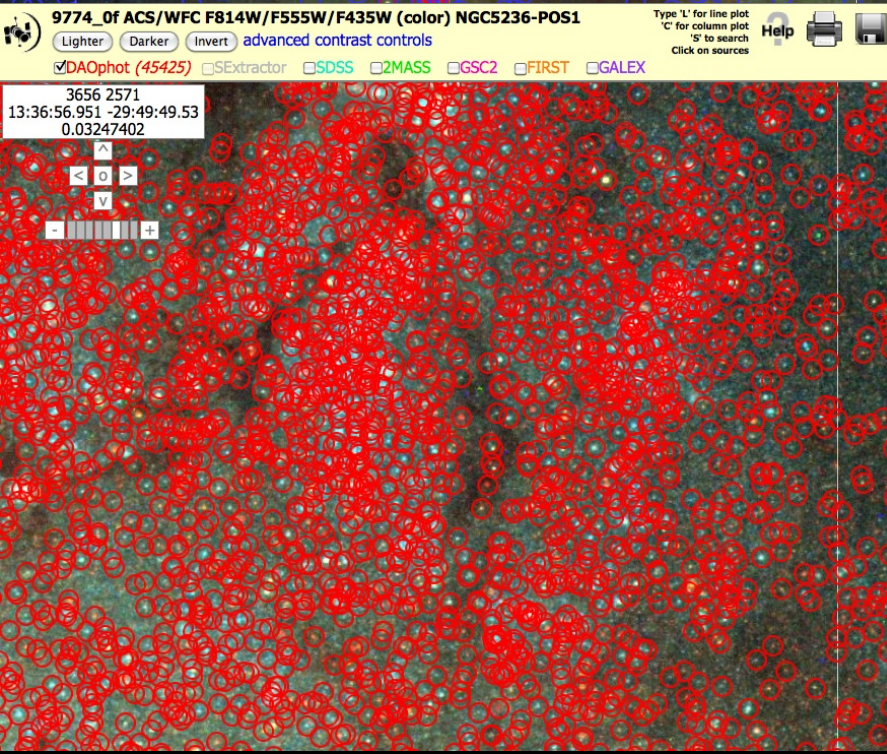
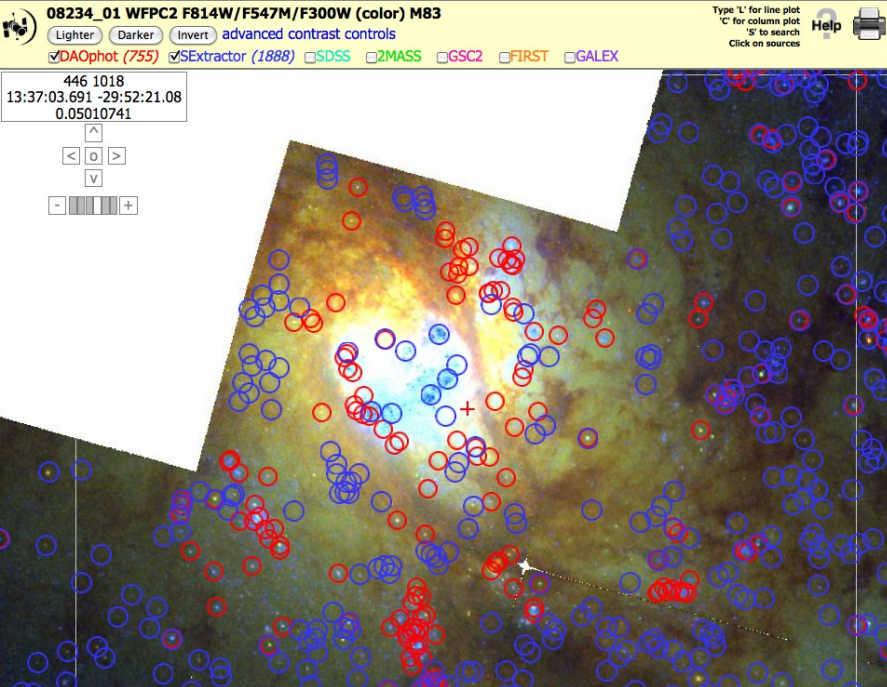
1. What HST data exist for my particular target ? (100 %)
2. Do a sanity check on my own catalog (80 % ?)
3. Make use of existing cross matches with other catalogs (HST HLSP, SDSS, Chandra, Spitzer, spectral extractions, ... (50 % ?)
4. Use it to do my science (10 – 50 % ?)

Caveats:

It is important to note that the HSC will not eliminate the need for astronomers to make more detailed catalogs optimized to **address their own specific science goals in many cases.**

The HSC is designed to be a **general-use catalog** sufficient for many, but not all users.

We are in an early stage (**Beta-1**) of the project. The HSC has a number of known shortcomings that are being addressed for a **Beta-2** release planned for January, and a **Version 1** release in May.



Use Case Categories

1. **Time variability** (e.g., SN, GRBs, light echos, proper motion, transients)
2. **Astrometry** (absolute astrometry, identification of "blind" sources such as ULXs)
3. **Photometry** (e.g., CMDs, color-color, photo-Z's, ...)
4. **Morphology** (of individual objects rather than ensemble - e.g., clumpy galaxies, star vs cluster, ...)
5. **Structure determination** (use ensemble of objects - e.g., Milky Way components, streams, galaxy clusters, ...)
6. **Very large datasets** (e.g., all LMC observations)
7. **Cross-matching catalogs** (e.g., with user-supplied catalog , multiwavelength catalogs, ...)
8. **Spectroscopic cross-matching** (e.g., with ACS and NICMOS grism extractions)

NOTE: HLA data release 7, highlighting search and display of contributed spectroscopic products - Nov. 2

also

"Enhancing the Legacy of HST Spectroscopy" meeting – Nov 15 -16

Example of a detailed version of a use case

Category 3 – Photometry

“Finding LBVs in M83 and similar galaxies” - Brad Whitmore

Reference: Chandar et al. 2010, ApJ, 719, 966

Step 1 - Make **U-B vs. V-I color-color diagrams** for stars (using concentration index [CI] or equivalent to isolate stars from clusters and background galaxies).

Step 2 - Allow definition of selection box in the color-color diagram to isolate LBV candidates. Make a working catalog (various formats allowed) based on objects in the selection box.

Step 3 - Show where these subsets are by **overlaying on image**. Are they associated with star formation regions or are they random.

Step 4 – Edit working catalog (approve/disapprove each object as go or change a selection criteria and see how it affects which objects get selected).

Step 5 - Look for **variability in the photometry** from repeat measurements. Compare/overplot light curves from various filters if available

Step 6 - Identify other Messier galaxies with the needed combination of filters.

The Use Cases **guide the development of a set of tools** and requirements for the project

LBV example - Required **tools**

- Make plots $\frac{1}{2}$ ✓
- Over plot SED models ---
- 3-plane visualization tool $\frac{1}{2}$ ✓
- Make working catalogs $\frac{1}{2}$ ✓
- Plot "light curves" ---
- Relational search capabilities ---

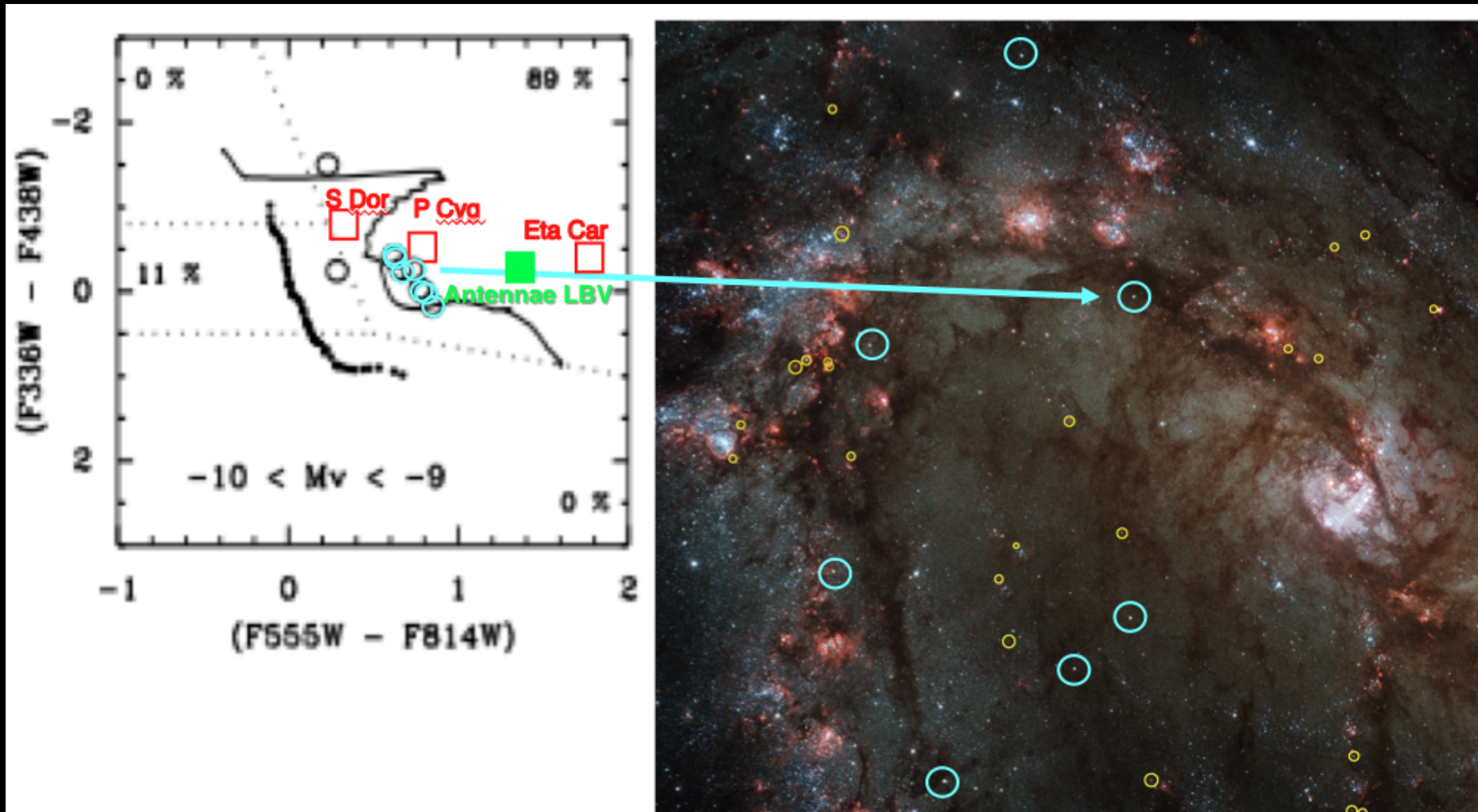
The Use Cases **guide the development of a set of** tools and **requirements** for the project

LBV example – Data quality **requirements**

- **Accuracy of photometry** better than 0.1 mag in UBVI
- **< 5 % artifacts** (mismatched objects, false detections, ...)
- **Measurement of a concentration index** (sufficient to separate stars from clusters at $m-M = 30$ with **< 5 % artifacts**)
- **Include WFPC2, ACS, WFC3**
- **Estimation of completeness** as function of magnitude / filter

In Chandar et al. 2010, we found that Luminous Blue Variable (LBV) candidates could be identified from a color-color diagram in M83.

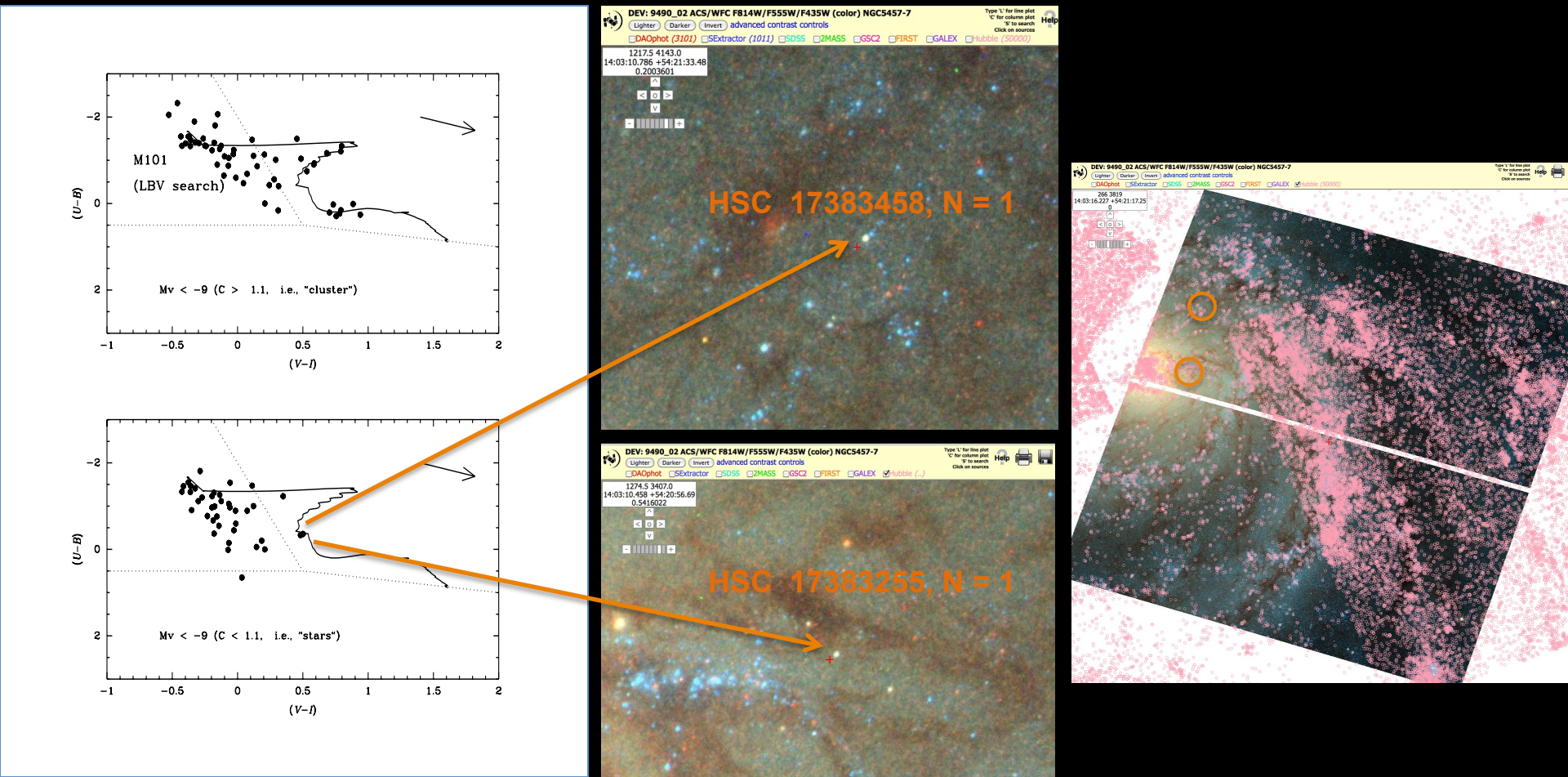
- We then used this technique to identify a LBV in the Antennae.
- What makes this a challenging archival task is the cross-matching of several instruments over long periods of time.



As part of this use case investigation, **2 excellent LBV candidates were found in M101** (using IRAF rather than the HSC since F336W is not currently in the list of included HSC filters).

Unfortunately, the two are not in the region with extensive coverage, in both time and filters, hence it is not possible to determine if they are actually variable using the HSC.

Will need to wait for version 1 of the HSC before this is easy to do (e.g, when a better ACS SExtractor catalog becomes available for image 9490_02).



Community Involvement

Four working groups will be formed to help guide the development of the HSC and insure that people using a variety of different analysis tools can use the catalog and tools effectively.

Point and Nearly-point source photometry (e.g., color-magnitude diagrams in globular clusters and in the field, compact clusters in external galaxies) [Potential members: Tom Brown, Jay Anderson, Peter Stetson, Abi Saha, Julianne Dalcanton, Brad Whitmore, Nate Bastian, Soeren Larsen, ...]

Extended object photometry and morphology (e.g., faint galaxies in survey fields, photometric redshifts) [Potential members: Harry Ferguson, Anton Koekemoer, Mark Dickinson, ...]

Multi wavelength (and spectroscopic) cross matches (e.g., ULIRGs, gamma-ray bursts, x-ray counterparts) [Potential members: Tamas Budavai, Lee Armus, Bob Hanisch, Knox Long, ...]

Time-resolved phenomena (e.g., Cepheids, supernova, variable stars, ...) [Potential members: Stefano Casertano, Adam Riess, Nathan Smith, ...]

Timing ? Start 1 or 2 groups now, others later ?

Early MAST Survey Results

11. We have recently introduced a preliminary beta version of the Hubble Source Catalog (HSC), with the goal of cross-matching the tens of thousands of individual visit-based source lists that are currently in the Hubble Legacy Archive into a single master catalog.

- How frequently do you think you would use the HSC for your research? (250 respondents)

Extensively	27
Moderately	121
Rarely	82
None	20

- What modes might you want to use in interacting with this catalog?

Download full catalog	80
Cone search (including web services callable from programs)	163
Casjobs/SQL	69
Other	12

- Would you be interested in serving on a HSC Working Group to help guide the development of this new initiative?

- **31 yes** (out of 232 respondents)

- Any other feedback/suggestions on how this catalog might be useful for your research?

“ This is a great first stab at this (and a long time coming). There are a lot of limitations in the catalog at present (e.g., no flags for suspected cosmic rays, only wide-band filters, assessment of variability/light curve) that need to be sorted out and expanded upon to make the catalog more useful to a broader audience/community.”

“ Include cross matching of catalog sources with source catalogs from other missions/surveys, and conditional matching (e.g., 'find me all HSC sources detected in X-ray and mid-IR and that were observed with HST in at least two bands/filters or grism').”

“ I think the catalog will be more useful with the highest possible precision in the absolute coordinates of sources.”

16. How important are the following current or planned features ?

Cross-correlations with other missions (e.g., SDSS, 2MASS,...) (235 respondents)

5 not useful

21 unsure

209 very useful

This was the highest “useful” rating out of 13 options.

Others included:

- footprints from other missions 199
- single sign-on user accounts 192
- integration with publications 187

Current HSC Web Form

Barbara A. MIKULSKI ARCHIVE OF SPACE TELESCOPES

MAST STScI Tools Mission Search Tutorial Site Search

HST Home About HST Getting Started Registration Archive Status HST Search HSTonline Search

Archive Status Hubble Source Catalog Summary (Help) Field Descriptions

Standard Form File Upload Form

Search Reset Clear Form

Target Name Resolver Radius (arcmin)

Right Ascension Declination Equinox

NumImages

User-specified field 1 Field Descriptions User-specified field 2 Field Descriptions

Output Columns up down

Sort By: ang_sep (") Reverse Match ID Reverse null Reverse

Output Coords: Sexagesimal Degrees Hours

Output Format HTML Table

Show Query Make Rows Distinct

Maximum Records: 1001

Records per Page: 100

Search Reset Clear Form

HLA_CAT_SUM Search Results

Display numeric columns graphically using VOPLOT

Object name m83 resolved by NED (via SANTA cache) to MESSIER 083 (G) RA: 13 37 09.95 Dec: -29 51 55.51 (J2000)

number of rows returned = 83

Match ID	NumFilters	NumVisits	NumImages	Target Name	Match RA (J2000)	Match DEC (J2000)	Obs. Start Time	Obs. Stop Time	Mag_F814W	Mag_F606W	Mag_F555W	Mag_F336W
3623982	4	1	4	M83	13 37 00.868	-29 51 55.89	2000-05-02 12:35:14	2000-05-02 19:32:14				
3629990	3	1	3	M83	13 37 00.839	-29 51 55.29	2000-04-25 16:59:14	2000-04-25 21:50:04	14.440			
3629999	3	1	3	M83	13 37 01.120	-29 51 56.62	2000-04-25 16:59:14	2000-04-25 21:50:04	15.432			
3623993	4	1	4	M83	13 37 00.731	-29 51 56.92	2000-05-02 12:35:14	2000-05-02 19:32:14				
3629971	3	1	3	M83	13 37 00.684	-29 51 56.45	2000-04-25 16:59:14	2000-04-25 21:50:04	15.132			
3623984	4	1	4	M83	13 37 01.165	-29 51 58.24	2000-05-02 12:35:14	2000-05-02 19:32:14				
3623994	2	1	2	M83	13 37 01.702	-29 51 50.25	2000-05-02 15:46:14	2000-05-02 19:32:14				
3623979	3	1	3	M83	13 37 00.477	-29 51 54.89	2000-05-02 12:35:14	2000-05-02 19:32:14				
3629975	3	1	3	M83	13 37 00.430	-29 51 54.76	2000-04-25 16:59:14	2000-04-25 21:50:04	14.635			
3629948	3	1	3	M83	13 37 00.950	-29 52 02.40	2000-04-25 16:59:14	2000-04-25 21:50:04	15.013			
3629940	3	1	3	M83	13 37 00.990	-29 52 02.82	2000-05-02 12:35:14	2000-05-02 19:32:14				
3629941	3	1	3	M83	13 37 00.792	-29 52 02.62	2000-04-25 16:59:14	2000-04-25 21:50:04	15.482			
3629960	3	1	3	M83	13 37 00.493	-29 52 00.80	2000-04-25 16:59:14	2000-04-25 21:50:04	14.955			
3623990	4	1	4	M83	13 37 00.543	-29 52 01.45	2000-05-02 12:35:14	2000-05-02 19:32:14				
3629991	4	1	4	M83	13 37 00.402	-29 51 59.69	2000-05-02 12:35:14	2000-05-02 19:32:14				
3629969	3	1	3	M83	13 37 00.368	-29 51 59.22	2000-04-25 16:59:14	2000-04-25 21:50:04	14.879			
3623997	4	1	4	M83	13 37 00.328	-29 51 58.18	2000-05-02 12:35:14	2000-05-02 19:32:14				
3629987	3	1	3	M83	13 37 00.298	-29 51 57.76	2000-04-25 16:59:14	2000-04-25 21:50:04	15.273			

archive.stsci.edu/hst/hla_cat/

Barbara A. MIKULSKI ARCHIVE OF SPACE TELESCOPES

MAST STScI Tools Mission Search Tutorial Site Search

HST Home About HST Getting Started Registration Archive Status HST Search HSTonline Search

HST Target Search

HST Abstract Search

FAQ

Search & Retrieval

MAST Services

Daily Data Reports

About HST Data

High-Level Science Products

Data Reduction & Analysis

Catalogs

Proposal Support

Proprietary Rights

Documentation

Gallery

Related Sites

Data Use Policy

Beta Version 1.0 of the Hubble Source Catalog

Search with Summary Form now (one row per match)
Search with Detailed Form now (one row per source)

The Hubble Source Catalog contains information about sources detected in any of the ACS and WFCP2 images. It is based on the positional crossmatching of source lists in the [Hubble Legacy Archive](#) (HLA) across HST visits (pointings) and/or detectors.

The catalog contains all members of the ACS/WFC and WFCP2 Source Extractor source lists in the HLA version DR6 that are considered to be valid detections because they have flag values less than 5. The catalog describes which sources crossmatch together. The crossmatching process involves adjusting the relative astrometry of overlapping images so as to minimize positional offsets between closely aligned sources in different images. After correction, the astrometric residuals of crossmatched sources are significantly reduced to typically less than 10 mas. In addition, the catalog describes source nondetections. The crossmatching algorithms and the properties of the catalog are described in [Budavari & Lubow \(2012\)](#).

Interfaces

There are two forms interfaces for querying the catalog. They follow the standards of MAST and provide a convenient way to browse the catalog. The [Detailed Search Form](#) provides a view of the source catalog that displays an entry for each source. The [Summary Search Form](#) displays an entry for each set of matching sources. In addition, we will be providing a more general SQL query interface for more complex and longer running searches through a [CatsJobs](#) system.

HLA Source Lists

The catalog is based on HLA Source Extractor [source lists](#). To build source lists, the HLA first constructs a white light detection image for each detector within each visit. This filter-combined drizzled image provides added depth. Source Extractor is run on the white light detection image to identify each source and its position. Next, the combined drizzled image for each filter used in the detection image is searched for a source at each position indicated by the detection image. If a valid source is detected, then its properties are entered into an HLA source list for the visit, detector, and filter with a flag value less than 5. Sources that are found in the white light detection image, but not in some filter with value flag less than 5 are regarded as nondetections or color dropouts for that filter. Some of the visits involve a single filter. In that case, the white light image is the same as the filter (color) image and no color dropouts are present.

Limitations

The catalog is preliminary and some of its limitations are listed below. We are working to make improvements on these issues.

Overlay only available in hla/dev at present



http://archtest.stsci.edu/hst/hla_cat/

Barbara A. MIKULSKI ARCHIVE OF SPACE TELESCOPES

MAST STScI Tools Mission Search Tutorial Site Search

HST Home About HST Getting Started Registration Archive Status HST Search HSTonline Search

Archive Status Hubble Source Catalog Beta Version (Help) Field Descriptions General Info

Standard Form File Upload Form

Search Reset Clear Form

Target Name Resolver Radius (arcmin)

Right Ascension Declination Equinox

NumImages

User-specified field 1 Field Descriptions User-specified field 2 Field Descriptions

Output Columns up down

Sort By: ang_sep (") Reverse Match ID Reverse null Reverse

Output Coords: Sexagesimal Degrees Hours

Output Format HTML Table

Show Query Make Rows Distinct

Maximum Records: 1001

Records per Page: 100

Search Reset Clear Form

HLA_CAT Search Results

Display numeric columns graphically using VOPLOT

Object name m83 resolved by NED (via SANTA cache) to MESSIER 083 (G) RA: 13 37 09.95 Dec: -29 51 55.51 (J2000)

number of rows returned = 207

Previous 1 2 3 4 5 6 Next

Page 1 of 6

Match ID	Mem ID	Source ID	Det	Target Name	Image Name	Match RA (J2000)	Match DEC (J2000)	D	D Sigma	Mode	Aperture	Exp Time	Obs. Start Time
19953474	1	2121018001	Y	ANY	HST_10579_at_ACS_WFC_F606N	13 37 00.878	-29 51 55.20	0.000		ACCUM	WFC	1996	2006-02-25 19:08:45
3623982	1	4220778942	Y	M83	hst_08234_02_wfc2_487e_wf	13 37 00.868	-29 51 55.89	0.000		IMAGE	PCI-FIX	3300	2000-05-02 12:35:14
3623982	2	4220778542	Y	M83	hst_08234_02_wfc2_7502e_wf	13 37 00.868	-29 51 55.89	0.000		IMAGE	PCI-FIX	2400	2000-05-02 14:07:14
3623982	3	4220778542	Y	M83	hst_08234_02_wfc2_856e_wf	13 37 00.868	-29 51 55.89	0.000		IMAGE	PCI-FIX	1200	2000-05-02 19:09:14
3623982	4	4220778542	Y	M83	hst_08234_02_wfc2_967e_wf	13 37 00.868	-29 51 55.89	0.000		IMAGE	PCI-FIX	2400	2000-05-02 15:46:14
3626990	1	4539036698	Y	M83	hst_08234_01_wfc2_490e_wf	13 37 00.839	-29 51 55.29	0.000		IMAGE	PCI-FIX	2100	2000-04-25 16:59:14
3626990	2	4539036698	Y	M83	hst_08234_01_wfc2_547e_wf	13 37 00.839	-29 51 55.29	0.000		IMAGE	PCI-FIX	930	2000-04-25 19:53:14
3626990	3	4539036698	Y	M83	hst_08234_01_wfc2_814e_wf	13 37 00.839	-29 51 55.29	0.000		IMAGE	PCI-FIX	710	2000-04-25 21:34:14
3626899	1	4539035457	Y	M83	hst_08234_01_wfc2_490e_wf	13 37 01.120	-29 51 56.62	0.000		IMAGE	PCI-FIX	2100	2000-04-25 16:59:14
3626899	2	4539035457	Y	M83	hst_08234_01_wfc2_547e_wf	13 37 01.120	-29 51 56.62	0.000		IMAGE	PCI-FIX	930	2000-04-25 19:53:14
3626899	3	4539035457	Y	M83	hst_08234_01_wfc2_814e_wf	13 37 01.120	-29 51 56.62	0.000		IMAGE	PCI-FIX	710	2000-04-25 21:34:14
19953564	1	2121015114	Y	ANY	HST_10579_at_ACS_WFC_F606N	13 37 00.726	-29 51 54.54	0.000		ACCUM	WFC	1996	2006-02-25 19:08:45

One row per observation

HSC Timeline

June 2012 – Beta-1 version released (Lubow + Budavari paper posted)

- Includes ACS and WFPC2 based on HLA SExtractor source lists
- Only detailed / time-resolved version of form available

October 2012 – Questionnaire sent. Early results encouraging

January 2013 – Beta-2 version of catalog

- Includes ACS and WFPC2 based on HLA SExtractor source lists
- Include “summary” form (early version made available in version 1.5)
- Improved source filtering and matching

May 2013 – Version 1 of HSC released

- Includes ACS, WFPC2, WFC3
- Preliminary cross-matching (e.g., SDSS)
- Improved tools (over plots, correlation plotting, ...)

Future – SQL CASJobs, value-added catalogs, workshops, PASP ...

Summary

An early **Beta version** of the HSC has been online since June.

We are using it, in conjunction with several use cases, to test the quality of the HSC and to set detailed requirements.

An aggressive timeline has been established, including **strong involvement with the community**. How would the STUC like to be involved ?

The HSC will add a valuable new dimension to the information available to Hubble users for **decades** to come, and provides a **bridge to future missions** both in space (e.g., JWST), and on the ground (e.g., LSST).

