

# *Advanced TOPCAT-STILTS*

Miriam Cortés



Astronomy ESFRI & Research Infrastructure Cluster  
ASTERICS - 653477



# TOPCAT & STILTS

- Both do basically the same things but
  - TOPCAT
    - Easier to learn.
    - Good for interactive use, especially exploring data to get a feel for what's there.
  - STILTS
    - Better for reproducible work (it can be scripted).
    - Steeper learning curve.

# TOPCAT & STILTS

- Which is the best format?

- [4.1.1.1 FITS](#)
- [4.1.1.2 Column-oriented FITS](#)
- [4.1.1.3 VOTable](#)
- [4.1.1.4 CDF](#)
- [4.1.1.5 ASCII](#)
- [4.1.1.6 IPAC](#)
- [4.1.1.7 Comma-Separated Values](#)
- [4.1.1.8 GBIN](#)
- [4.1.1.9 Tab-Separated Table](#)
- [4.1.1.10 SQL Database Queries](#)
- [4.1.1.11 World Data Center](#)

- Small table (<1000 rows): **doesn't matter.**

- Medium-sized (rows\*cols) < 20million): **FITS.**

- Big (millions of rows, especially with lots of columns):  
**colfits.**

- If the input file is not in this format you can convert it using STILTS:

- *stilts tpipe in=xxx.csv ifmt=csv out=xxx.fits*

# TOPCAT & STILTS

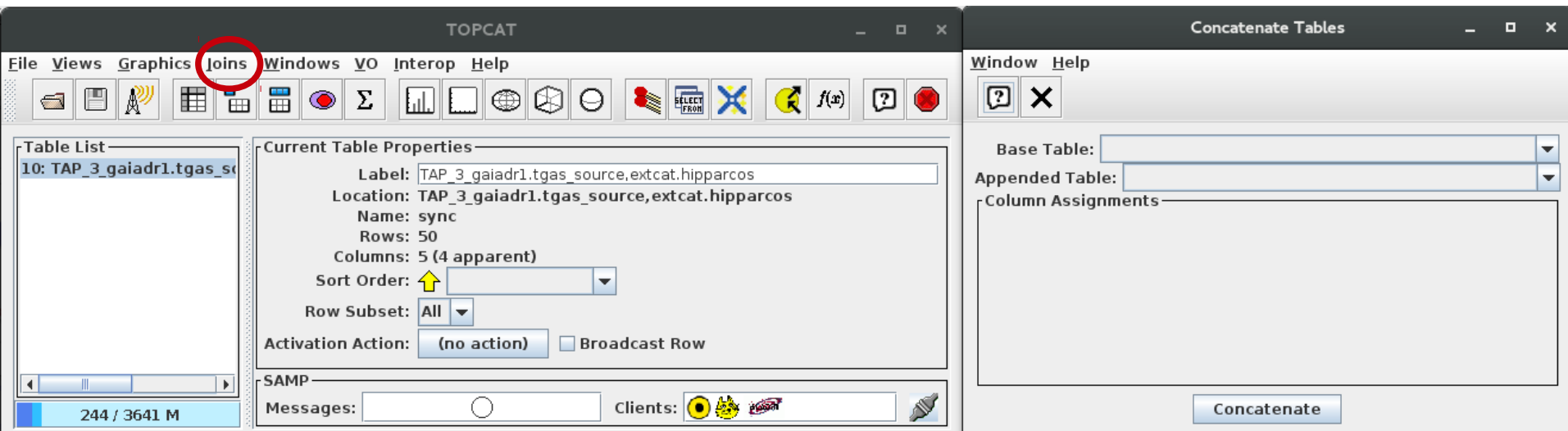
- Output in Latex

The image shows two overlapping windows. The background window is TOPCAT, displaying 'Current Table Properties' for a table named 'TAP\_3\_gaiadr1.tgas\_source,extcat.hipparcos'. The properties include: Label: TAP\_3\_gaiadr1.tgas\_source,extcat.hipparcos; Location: TAP\_3\_gaiadr1.tgas\_source,extcat.hipparcos; Name: sync; Rows: 50; Columns: 5; Sort Order: ↑; Row Subset: All. The foreground window is Emacs, editing a file named 'tabla\_latex.tex'. The LaTeX code defines a table with 5 columns and 6 rows of data.

```
documentclass{article}
\begin{document}
\begin{table}
\begin{tabular}{|r|r|r|r|r|}
\hline
\multicolumn{1}{|c|}{hip} &
\multicolumn{1}{|c|}{g_mag_abs_gaia} &
\multicolumn{1}{|c|}{g_mag_abs_hip} &
\multicolumn{1}{|c|}{b_v} &
\hline
95905 & 2.90110612385656 & 3.08139684809066 & 0.394 & \\
95838 & 3.36666243484313 & 3.60007543840966 & 0.707 & \\
95662 & 4.21575480915181 & 4.96691320323364 & 0.683 & \\
96089 & 3.67412200337596 & 3.99711049720092 & 0.609 & \\
97946 & 3.95220466256254 & 3.93122863291356 & 0.495 & \\
98189 & 4.08580555128650 & 3.90903495748743 & 0.639 & \\
\end{tabular}
\end{table}
\end{document}
```

# TOPCAT & STILTS

- Concatenating tables in TOPCAT



- Only two tables at a time.

# TOPCAT & STILTS

- Concatenating multiple tables in STILTS

## B.24.2 Examples

Here are some examples of `tcat`:

```
stilts tcat ifmt=ascii in=t1.txt in=t2.txt in=t3.txt out=table.txt
```

Concatenates the three named ASCII format tables to produce an output table. All three must have compatible numbers and types of columns.

```
stilts tcat ifmt=ascii in="t1.txt t2.txt t3.txt" out=table.txt
```

Has exactly the same effect as the previous example.

```
stilts tcat ifmt=ascii in=@inlist.lis out=table.txt
```

This will have the same effect as the previous two examples if a file name "inlist.lis" in the current directory contains three lines, "t1.txt", "t2.txt" and "t3.txt".

- Same input format → `tcatn`
- Similar columns (in number and class).

# TOPCAT & STILTS

- Concatenating multiple tables in STILTS

```
stilts tcatn nin=2 in1=survey.vot.gz ifmt2=csv in2=more_data.csv
      icmd1='addskycoords fk5 galactic RA2000 DEC2000 GLON GLAT' \
      icmd1='keepcols "OBJ_ID GLON GLAT"' \
      icmd2='keepcols "ident gal_long gal_lat"' \
      loccol=FILENAME
      omode=topcat
```

In this case we are trying to concatenate results from two tables which are quite dissimilar to each other. In the first place, one is a VOTable (no `ifmt1` parameter is required since VOTables can be detected automatically), and the other is a comma-separated-values file (for which the `ifmt2=csv` parameter must be given). In the second place, the column structure of the two tables may be quite different. By pre-processing the two tables using the `icmd1` & `icmd2` parameters, we produce in each case an input table which consists of three columns of compatible types and meanings: an integer identifier and floating point galactic longitude and latitude coordinates. The second table contains such columns to start with, but the first table requires an initial step to convert FK5 J2000.0 coordinates to galactic ones. `tcatn` joins the two doctored tables together, to produce a table which contains only these three columns, with all the rows from both input tables, and sends the result directly to a new or running instance of TOPCAT. An additional column named FILENAME is appended to the table before sending it; this contains "survey.vot.gz" for all the columns from the first table and "more\_data.csv" for all the columns from the second one.

# TOPCAT & STILTS

- Functions in TOPCAT

The screenshot displays two windows from the TOPCAT software. The left window, titled 'Define Synthetic Column', contains a form with fields for Name, Expression, Units, Description, UCD (set to 'no UCD'), and Index (set to 47). A red circle highlights the  $f(x)$  icon in the top-left corner of this window. The right window, titled 'Available Functions', shows a tree view of function categories. The 'Times' category is expanded, listing various date and time conversion functions. The function `julianToMjd( julianEpoch )` is selected and highlighted. The right pane of this window provides details for the selected function, including its description, parameters, return value, and signature.

**Function `julianToMjd( julianEpoch )`**

**Description:**  
Converts a Julian Epoch to Modified Julian Date. For approximate purposes, the argument of this routine consists of an integral part which gives the year AD and a fractional part which represents the distance through that year, so that for instance 2000.5 is approximately 1 July 2000.

**Parameters:**  
`julianEpoch` (floating point)  
julian epoch

**Return Value (floating point):**  
modified julian date

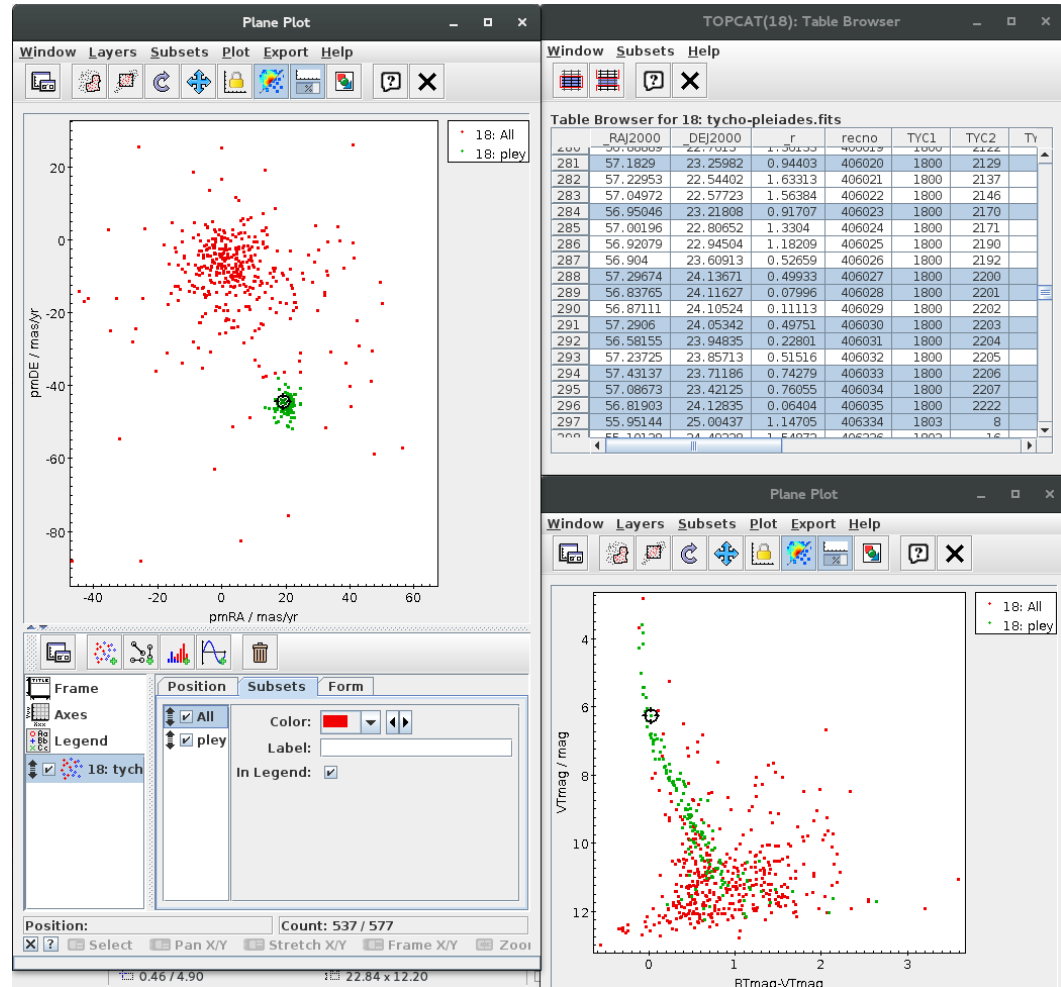
**Example:**  
`julianToMjd(2000.0) = 51544.5`

**Signature:**  
`double julianToMjd(double)`



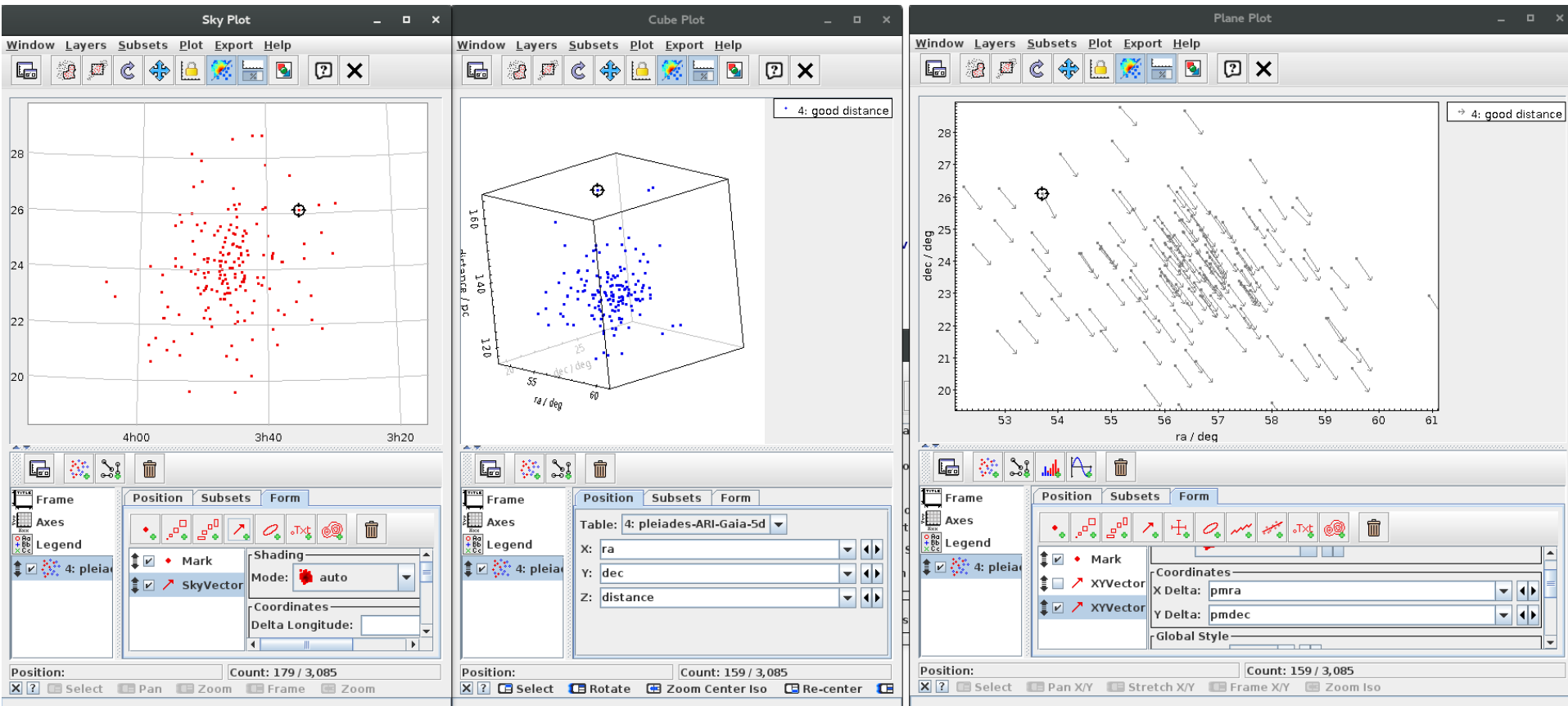
# TOPCAT & STILTS

- Linked views



# TOPCAT & STILTS

- Linked views



# TOPCAT & STILTS

- Crossmatching



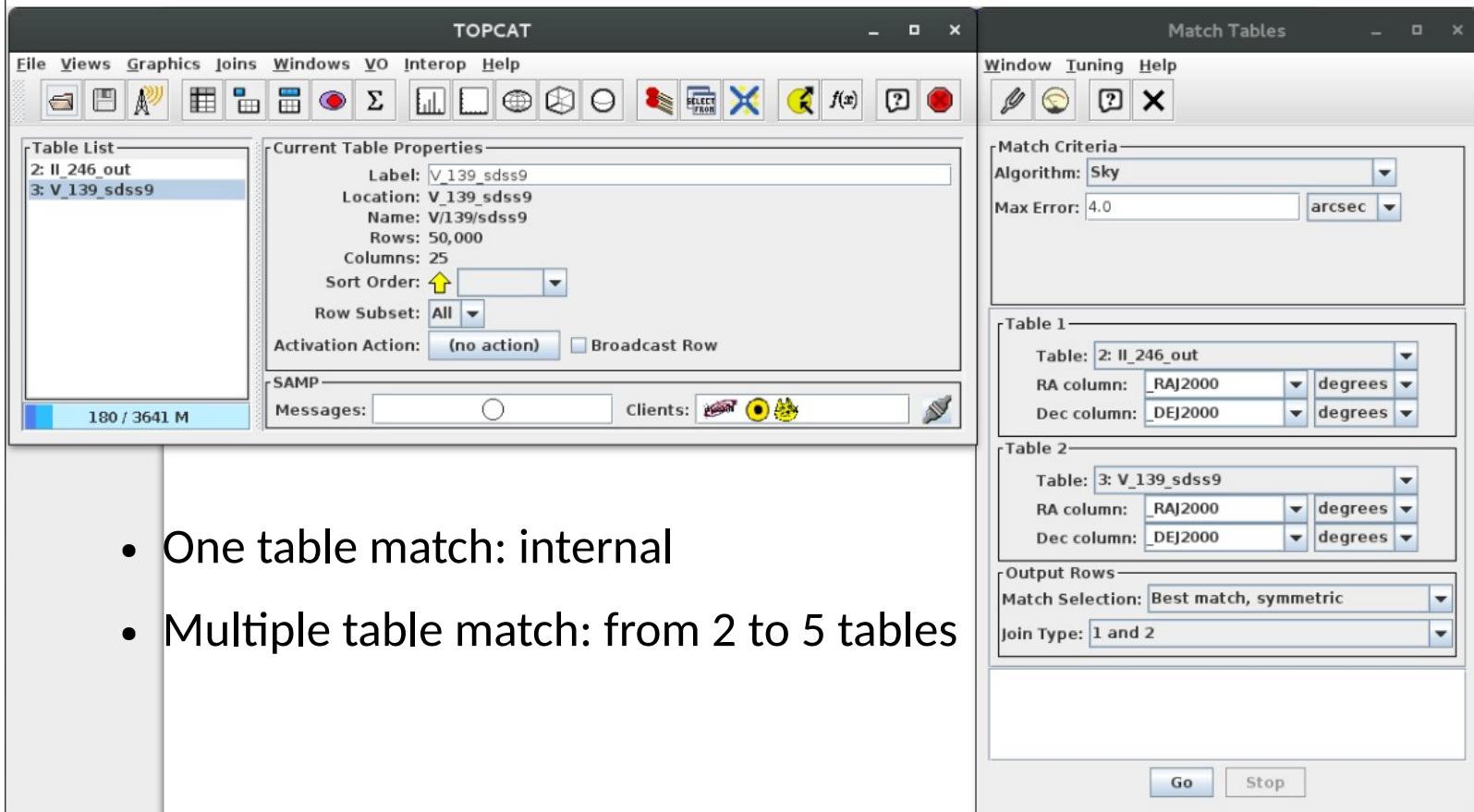
```
stilts tskymatch2 \  
  in1=tycho-pleiades.fits ra1=_RAJ2000 dec1=_DEJ2000 \  
  in2=2mass-pleiades.fits ra2=_RAJ2000 dec2=_DEJ2000 \  
  join=1and2 find=best error=1 \  
  out=tycho-2mass.fits \  
  \
```

- There are lots of different match types (Algorithm selector), not just Sky.
- Think about the output options. Especially in crowded fields, the default Best Match, Symmetric can give surprising results.
- For large tables (> million rows) , the crossmatch can run out of memory.
  - Tip: Increase heap memory (run with `java -jar -Xmx2048M topcat-full.jar`) or use the `java -disk` option.

# TOPCAT & STILTS

- **Crossmatching**  - How to x-match two **medium-size** catalogues?

TOPCAT → Joins / Pair match



The screenshot displays the TOPCAT software interface. The main window is titled 'TOPCAT' and features a menu bar (File, Views, Graphics, Joins, Windows, VO, Interop, Help) and a toolbar with various icons. On the left, a 'Table List' shows two tables: '2: II\_246\_out' and '3: V\_139\_sdss9'. The 'Current Table Properties' panel for 'V\_139\_sdss9' shows details like Label, Location, Name, Rows (50,000), Columns (25), Sort Order, Row Subset (All), and Activation Action. The 'Match Tables' window on the right is titled 'Match Tables' and contains settings for 'Match Criteria' (Algorithm: Sky, Max Error: 4.0 arcsec), 'Table 1' (Table: II\_246\_out, RA column: RAJ2000, Dec column: DEJ2000), and 'Table 2' (Table: V\_139\_sdss9, RA column: RAJ2000, Dec column: DEJ2000). The 'Output Rows' section is set to 'Match Selection: Best match, symmetric' and 'Join Type: 1 and 2'. 'Go' and 'Stop' buttons are at the bottom.

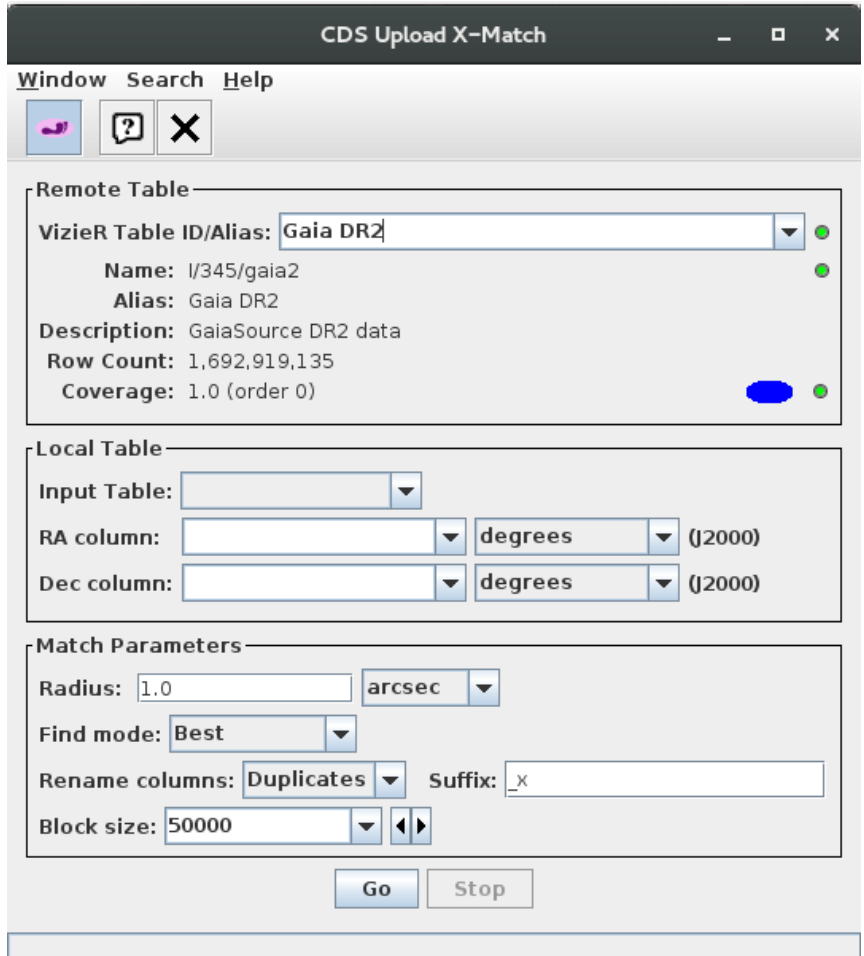
- One table match: internal
- Multiple table match: from 2 to 5 tables

# TOPCAT & STILTS

- **Crossmatching**  - How to x-match my catalogue with a **large catalogue** (in CDS) ?

TOPCAT → Joins → CDS Upload X-Match

- **Advantages:** Efficiency
- **Disadvantages:**
  - Only CDS catalogues
  - Only default columns



**CDS Upload X-Match**

Window Search Help

Remote Table

VizieR Table ID/Alias: Gaia DR2

Name: I/345/gaia2

Alias: Gaia DR2

Description: GaiaSource DR2 data

Row Count: 1,692,919,135

Coverage: 1.0 (order 0)

Local Table

Input Table:

RA column: degrees (12000)

Dec column: degrees (12000)

Match Parameters

Radius: 1.0 arcsec

Find mode: Best



Rename columns: Duplicates Suffix: \_x

Block size: 50000

Go Stop

# TOPCAT & STILTS

- **Crossmatching**  - How to x-match my catalogue with a **large catalogue** (in CDS) ?

Radmm  X Y  Hot Stuff for One Year (HSOY) (Altmann+, 2017) [2017A&A...600L...4A](#) [ReadMe+ftp](#) [Similar Catalogs](#)

I/339 [Post annotation](#)

1.I/339/hsoy The HSOY catalogue (583001653 sources) (original column names in green) (583001653 rows)

**Simple Constraint** **List Of Constraints**

Query by [Constraints](#) ? applied on Columns (Output Order:  +  -)

Standard  Original

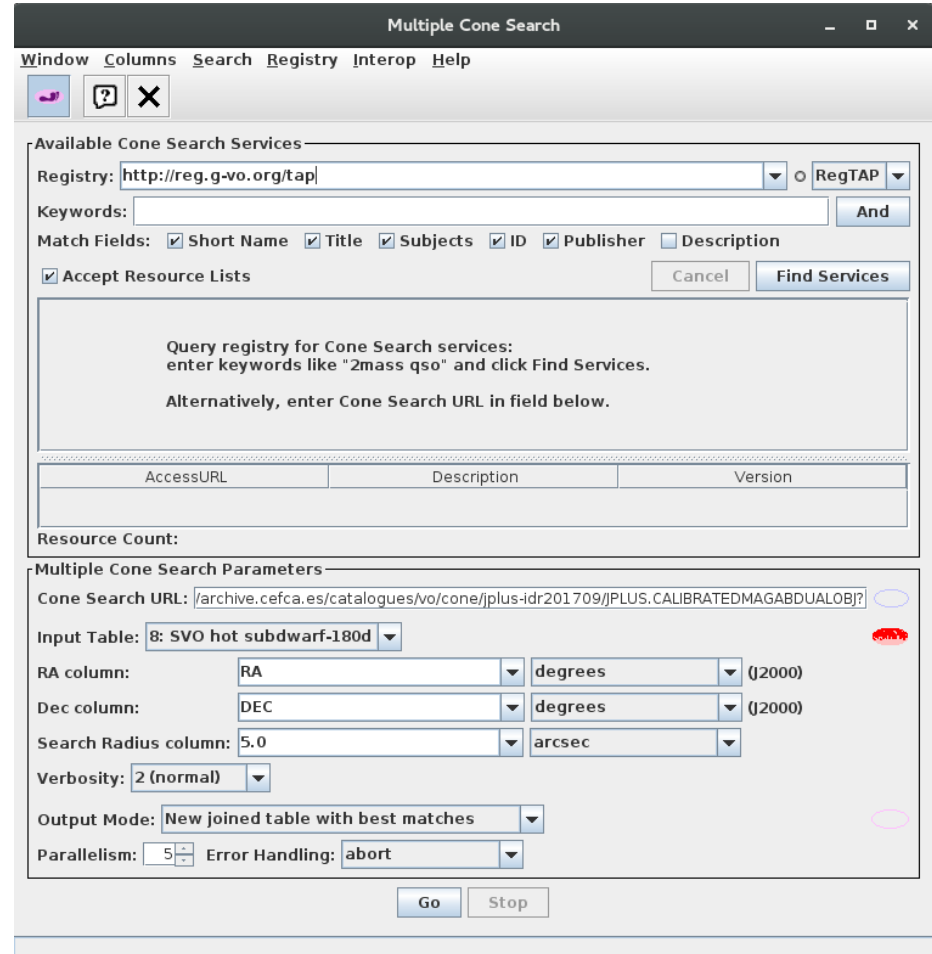
Show	Sort	Column	Clear	Constraint	Explain (UCD)
<input checked="" type="checkbox"/>	<input type="radio"/>	RAJ2000	<input type="text"/>	<a href="#">deg</a>	(i) Right ascension, J2000.0, at epoch 2000 (raj2000) ( <a href="#">pos.eq.ra;meta.main</a> )
<input checked="" type="checkbox"/>	<input type="radio"/>	DEJ2000	<input type="text"/>	<a href="#">deg</a>	(i) Declination, J2000.0, at epoch 2000 (dej2000) ( <a href="#">pos.eq.dec;meta.main</a> )
<input checked="" type="checkbox"/>	<input type="radio"/>	ipix	<input type="text"/>		(n)(i) PPMXL object identifier (ipix) ( <a href="#">Note 1</a> ) ( <a href="#">meta.id;meta.main</a> )
<input checked="" type="checkbox"/>	<input type="radio"/>	comp	<input type="text"/>		[0/4] Disambiguation counter (where multiple DR1 objects match one PPMXL object) (comp) ( <a href="#">Note 1</a> ) ( <a href="#">meta.code.multip</a> )
<input type="checkbox"/>	<input type="radio"/>	e_RAJ2000	<input type="text"/>	<a href="#">mas</a>	Mean error: RA*cos(DE) at mean epoch EpRA (e_ra) ( <a href="#">stat.error;pos.eq.ra</a> )
<input type="checkbox"/>	<input type="radio"/>	e_DEJ2000	<input type="text"/>	<a href="#">mas</a>	Mean error: DE at mean epoch EpDE (e_de) ( <a href="#">stat.error;pos.eq.dec</a> )
<input checked="" type="checkbox"/>	<input type="radio"/>	pmRA	<input type="text"/>	<a href="#">mas/yr</a>	Proper motion in RA, pmRA*cos(DE) (pmra) ( <a href="#">pos.pm;pos.eq.ra</a> )
<input checked="" type="checkbox"/>	<input type="radio"/>	pmDE	<input type="text"/>	<a href="#">mas/yr</a>	Proper motion in DE (pmde) ( <a href="#">pos.pm;pos.eq.dec</a> )
<input type="checkbox"/>	<input type="radio"/>	e_pmRA	<input type="text"/>	<a href="#">mas/yr</a>	Mean error in pmRA (e_pmra) ( <a href="#">stat.error;pos.pm;pos.eq.ra</a> )
<input type="checkbox"/>	<input type="radio"/>	e_pmDE	<input type="text"/>	<a href="#">mas/yr</a>	Mean error in pmDE (e_pmde) ( <a href="#">stat.error;pos.pm;pos.eq.dec</a> )

# TOPCAT & STILTS

- Crossmatching  - How to x-match my catalogue with a **large catalogue** (in CDS) ?

TOPCAT → VO → Multicone

- Disadvantages:
  - slow



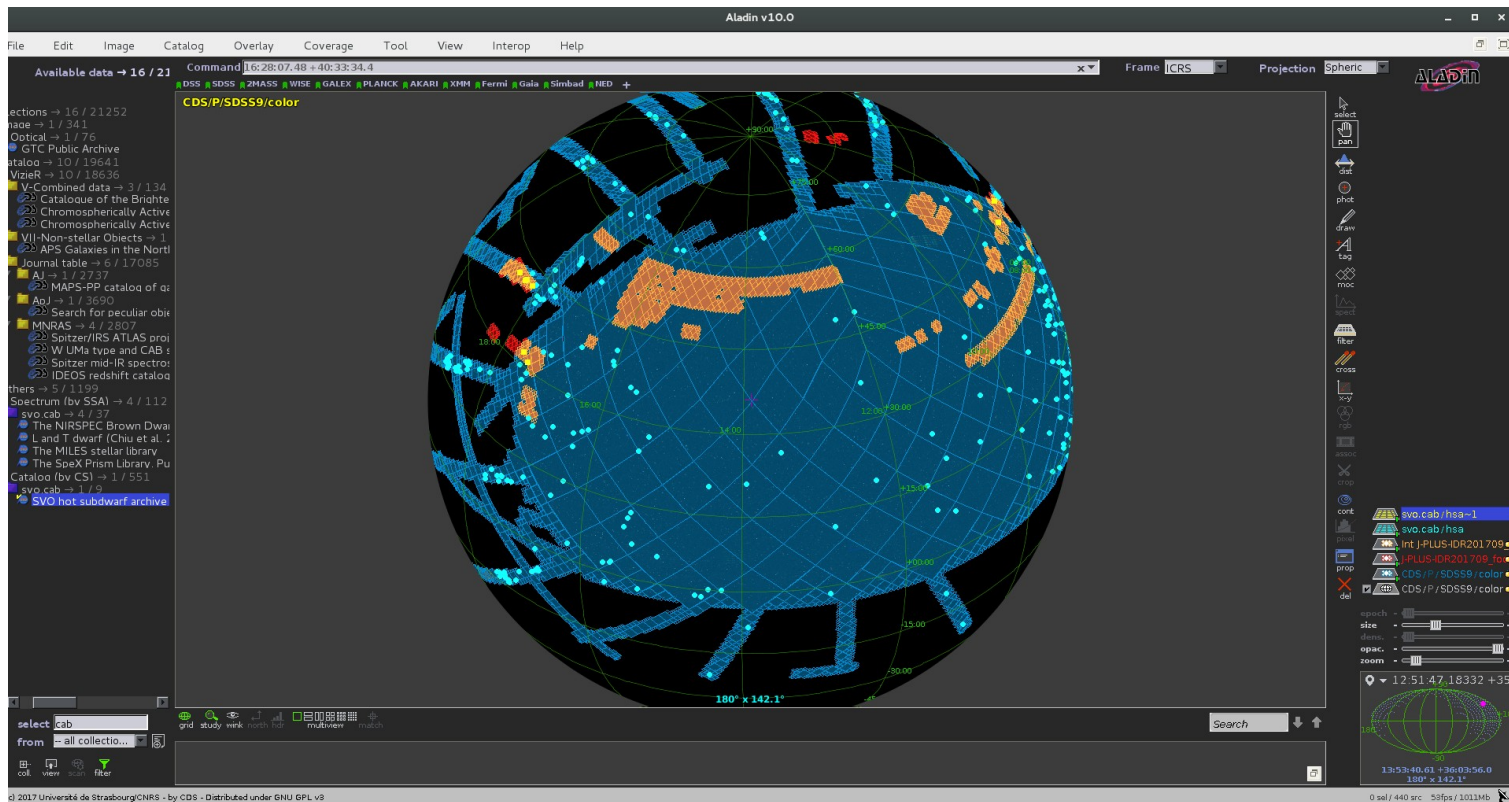
The screenshot shows the 'Multiple Cone Search' window. The 'Available Cone Search Services' section has 'Registry' set to 'http://reg.g-vo.org/tap' and 'RegTAP' selected. The 'Match Fields' section has checkboxes for 'Short Name', 'Title', 'Subjects', 'ID', 'Publisher', and 'Description', with 'Short Name', 'Title', 'Subjects', and 'ID' checked. The 'Find Services' button is visible. Below this is a text box with instructions: 'Query registry for Cone Search services: enter keywords like "2mass qso" and click Find Services. Alternatively, enter Cone Search URL in field below.' A table with columns 'AccessURL', 'Description', and 'Version' is shown below the text box. The 'Resource Count:' section is empty. The 'Multiple Cone Search Parameters' section has 'Cone Search URL' set to '/archive.cefca.es/catalogues/vo/cone/jplus-idr201709/JPLUS.CALIBRATEDMAGABDUALOBJ?'. The 'Input Table' is '8: SVO hot subdwarf-180d'. The 'RA column' is 'RA' in 'degrees' with '(J2000)'. The 'Dec column' is 'DEC' in 'degrees' with '(J2000)'. The 'Search Radius column' is '5.0' in 'arcsec'. The 'Verbosity' is '2 (normal)'. The 'Output Mode' is 'New joined table with best matches'. The 'Parallelism' is '5' and 'Error Handling' is 'abort'. 'Go' and 'Stop' buttons are at the bottom.

# TOPCAT & STILTS

- **Crossmatching**  - How to x-match my catalogue with a **large catalogue** (in CDS) ?

Alternative (for non all-sky surveys)

- Filter a table by MOC → X-match the filtered table





# TOPCAT & STILTS

- Crossmatching  - How to x-match two large catalogues (in CDS) ?

- Disadvantages:
  - No filtering
  - Large outputs

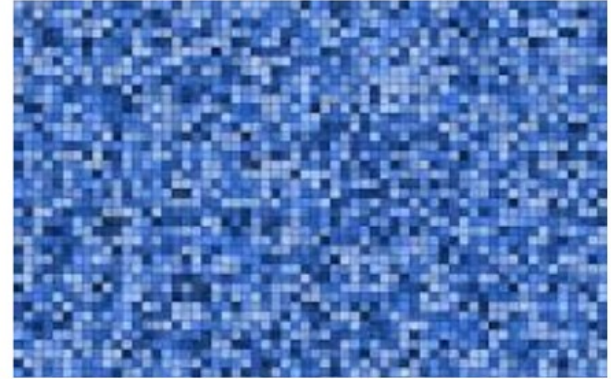


The screenshot shows the CDS X-Match Service interface. At the top, there is a navigation bar with links for Portal, Simbad, VizieR, Aladin, X-Match, Other, and Help. Below this, the main heading is "CDS X-Match Service" with sub-links for X-match, Tables management, and Documentation. The main content area is titled "Choose tables to cross-match". It features two input fields: "Gala DR2" and "PanSTARRS DR1". Each field has a dropdown menu with options "VizieR", "SIMBAD", and "My store". Below the input fields, there are two table cards. The first card is for "Gaia DR2 (Gaia Collaboration, 2018)" with 1,692,919,135 rows and a thumbnail image of a galaxy. The second card is for "The Pan-STARRS release 1 (PS1) Survey - DR1 (Chambers+, 2016)" with 1,919,106,885 rows and a thumbnail image of a galaxy. A "Show options" button is located below the table cards. A "Begin the X-Match" button is positioned below the "Show options" button. At the bottom of the interface, there is a section titled "Visualize and manage your cross-match jobs" which contains a table with columns for Table 1, Table 2, Options, Begin, Status, and Actions. The table currently shows "No job in list". A "Delete" button is located at the bottom right of the interface.

# TOPCAT & STILTS

- Crossmatching  • - How to x-match two large catalogues (in CDS) ?  
(Alternative)

- STILTS



- Cross-match

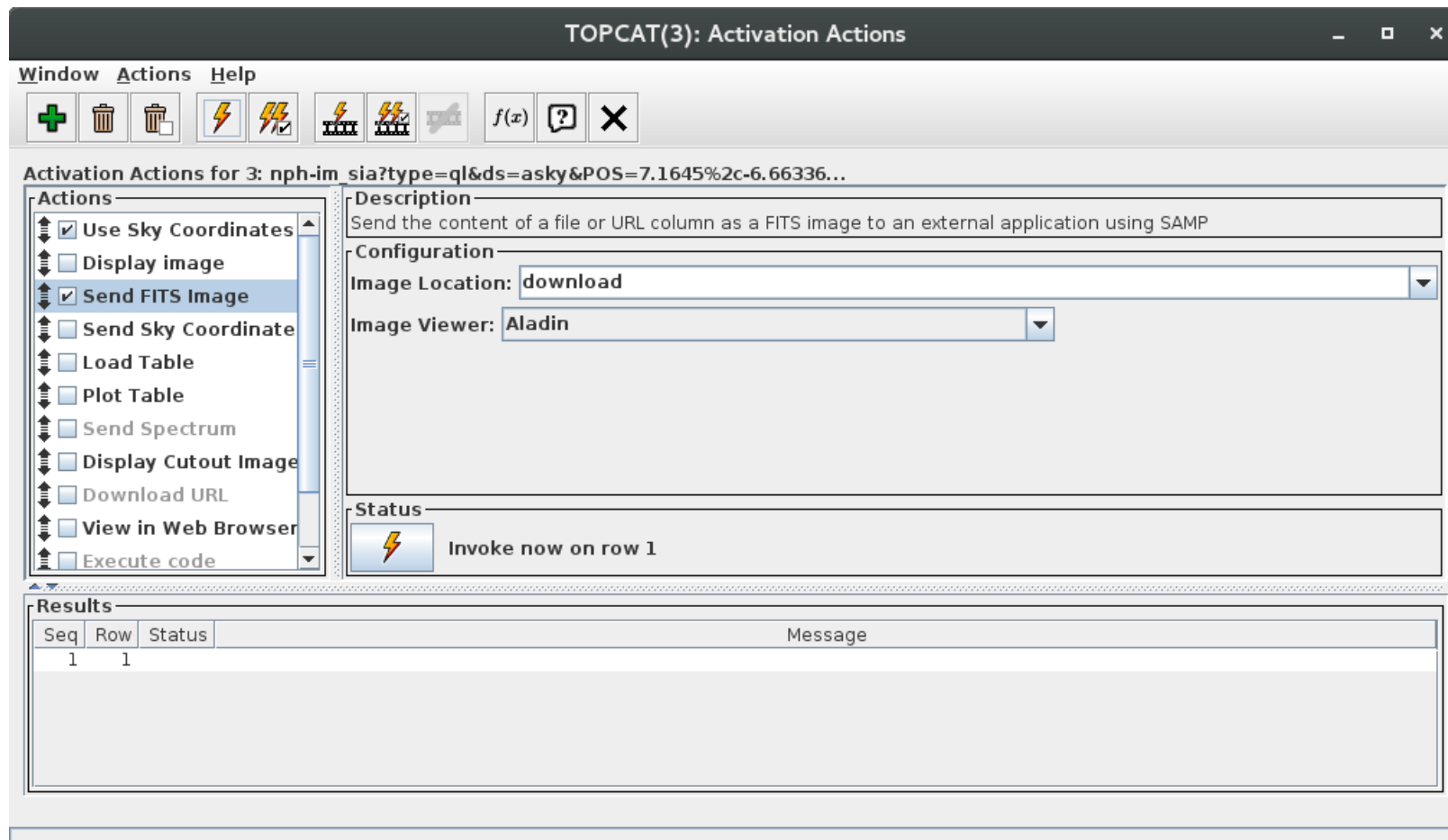
```
java -jar stilts.jar tskymatch2 ifmt1=votable in1=2mass.xml ifmt2=votable  
in2=sdss.xml ra1="RAJ2000" dec1="DEJ2000" ra2="RAJ2000" dec2="DEJ2000"  
error=10 find=all out=cross.xml ofmt=votable'
```

- Filtering

```
java -jar stilts.jar tpipe ifmt=votable in=cross.xml cmd="select  
zmag>12&&zmag<19.5&&rmag-kmag>(zmag+0.5)/2.5&&(rmag-  
kmag)<(zmag+10.5)/2.5&&e_Kmag>0" out=rmkz.xml ofmt=votable
```

# TOPCAT & STILTS

- Activation actions and activation window



# TOPCAT & STILTS

- Activation actions and activation window

Starlink SPLAT-VO: A Spectral Analysis Tool

Global list of spectra: /tmp/SPLAT294236732607

Properties of current spectra:

Short name: /tmp/SPLAT294236732607326303.fits  
Full name: http://sdc.cab.inta-csic.es:80/cqi-ines/SingleDownload?filename=LWP12752LL.FIT  
Format: TABLE

Columns: WAVELENGTH FLUX SIGMA

Colour: Save Reset

Composite: 100%  
Line type: polyline  
Line width: 1 Style: line

TOPCAT

Table List  
29: ll\_246\_out  
31: ssas(29)

Current Table Properties  
Label: ssas(29)  
Location: ssas(29)  
Name:  
Rows: 34  
Columns: 50  
Sort Order:  
Row Subset: All  
Activation Action: spectrum(Spectrum)  Broadcast Row

SAMP  
Messages: Clients:

Starlink SPLAT-VO: <plot0>

Displaying: /tmp/SPLAT294236732607326303.fit Remove Y limits (%): automatic  :V-hair

Wavelength: 3239.469  log Data count: 1.446742E-16  log  Track free

X scale: 1.0 Y scale: 1.0

2-d compound coordinate system

Data count (erg/cm<sup>2</sup>/s/A)

Wavelength (Angstrom)

TOPCAT(31): Table Browser

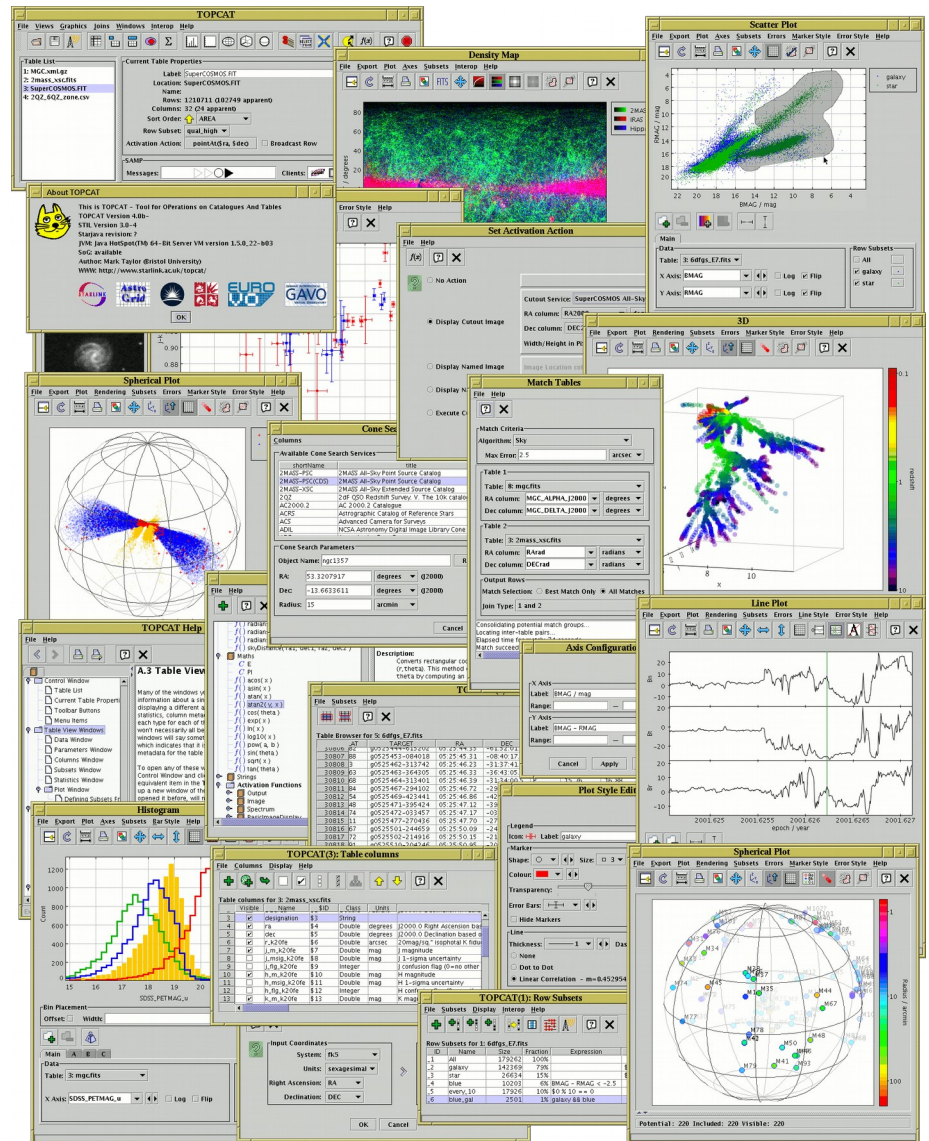
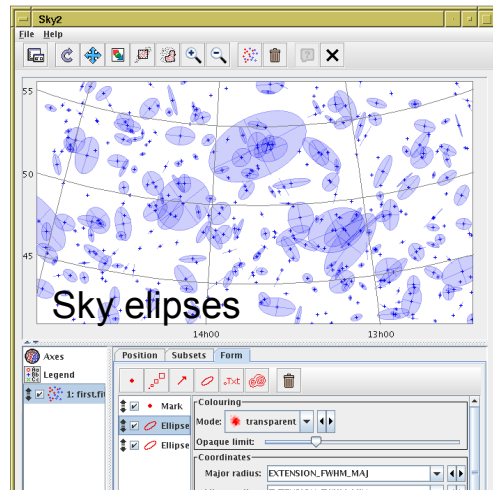
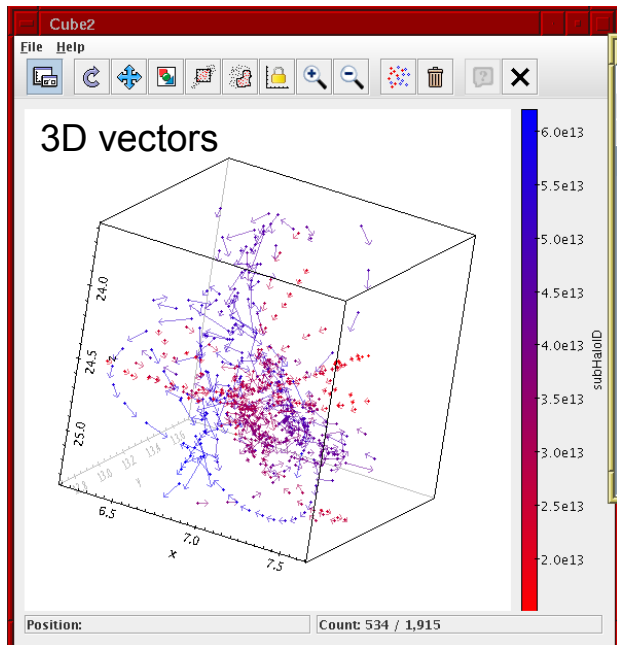
Window Subsets Help

Table Browser for 31: ssas(29)

	AXES	UNITS	DIMEQ
1	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
2	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
3	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
4	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
5	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
6	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
7	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
8	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
9	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
10	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
11	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
12	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
13	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
14	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
15	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
16	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
17	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3
18	WAVELENGTH FLUX SIGMA QUALITY	ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a	L ML-1T-3 ML-1T-3

Go Stop

# TOPCAT: Visualization



# TOPCAT & STILTS

- More at:

- TOPCAT v 4.6-1

<http://www.star.bris.ac.uk/~mbt/topcat/sun253/sun253.html>

- STILTS v 3.1-4

<http://www.star.bris.ac.uk/~mbt/stilts/sun256/sun256.html>

- TOPCAT/STILTS advanced tutorial

<http://andromeda.star.bris.ac.uk/topcat/tutorial-asterics1/>