

# *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 a composite screenshot of the TOPCAT and Emacs interfaces. The TOPCAT window displays the '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: (ascending arrow icon); Row Subset: All. The Emacs window, titled 'tabla\_latex.tex - emacs@esm.laptop', shows the LaTeX code generated for this table. The code includes document class, begin, and table/tabular commands, with data rows separated by \hline commands.

**TOPCAT Current Table Properties:**

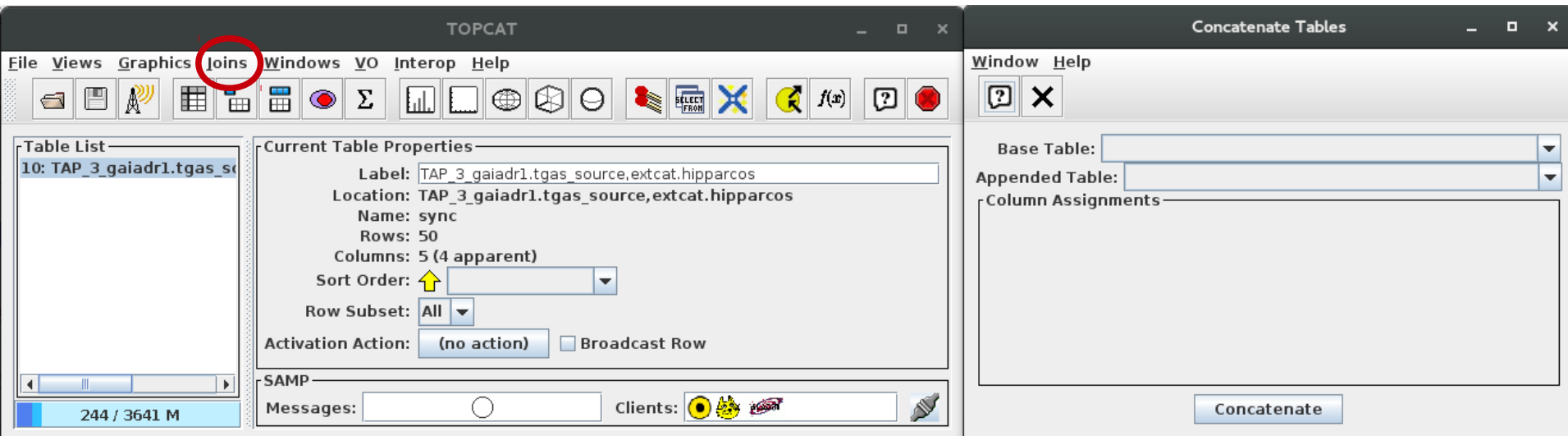
- Label: TAP\_3\_gaiadr1.tgas\_source,extcat.hipparcos
- Location: TAP\_3\_gaiadr1.tgas\_source,extcat.hipparcos
- Name: sync
- Rows: 50
- Columns: 5
- Sort Order: (ascending arrow icon)
- Row Subset: All

**Emacs LaTeX Output:**

```
\documentclass{article}
\begin{document}
\begin{table}
\begin{tabular}{|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

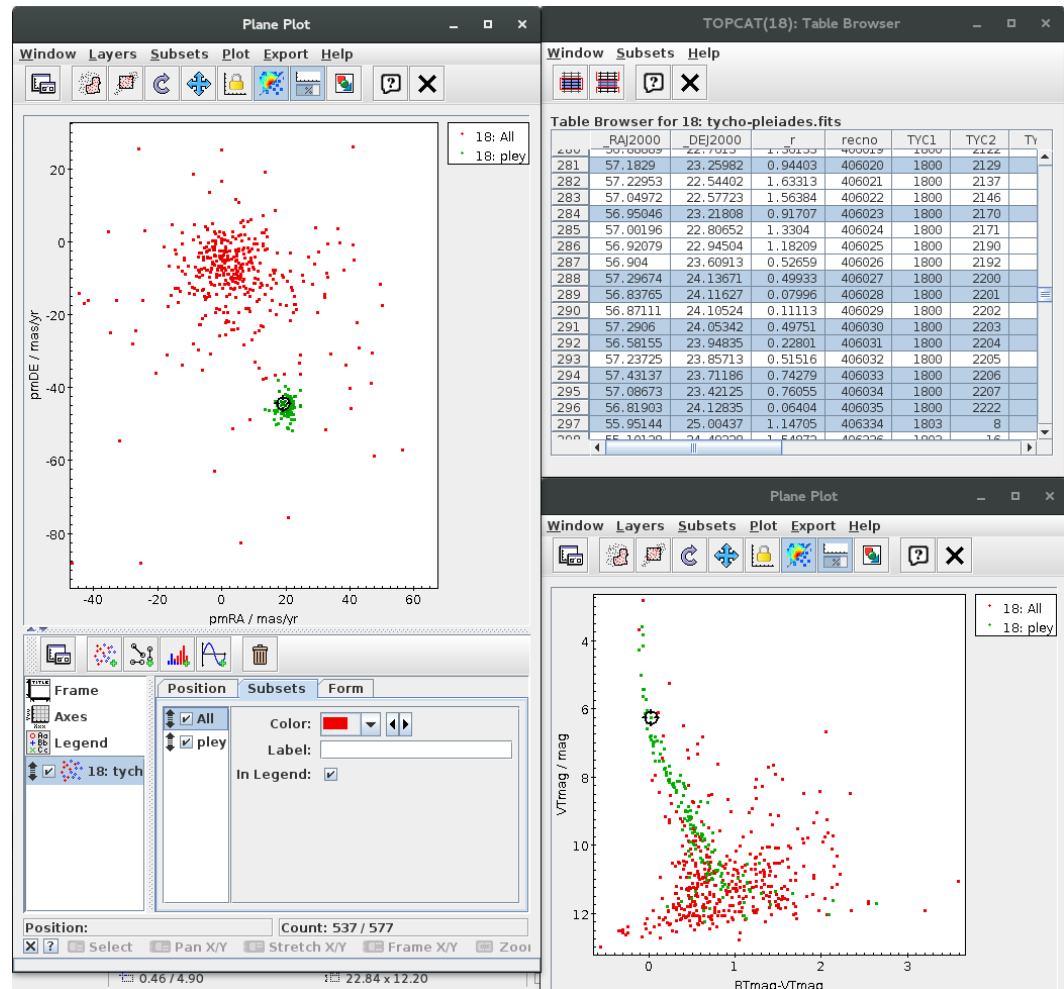
- Functions in TOPCAT

The screenshot displays two windows from the TOPCAT software interface. The left window, titled 'Define Synthetic Column', contains fields for 'Name', 'Expression', 'Units', 'Description', 'UCD' (with a dropdown menu showing '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: Arithmetic, Arrays, Conversions, CoordsDegrees, CoordsRadians, Coverage, Distances, Fluxes, Formats, KCorrections, Maths, Strings, Tilings, and Times. The 'Times' category is expanded, listing various date and time conversion functions. The function 'julianToMjd( julianEpoch )' is highlighted in blue. To the right of the function list, a detailed description is provided: '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', and 'Signature: double julianToMjd(double)'.



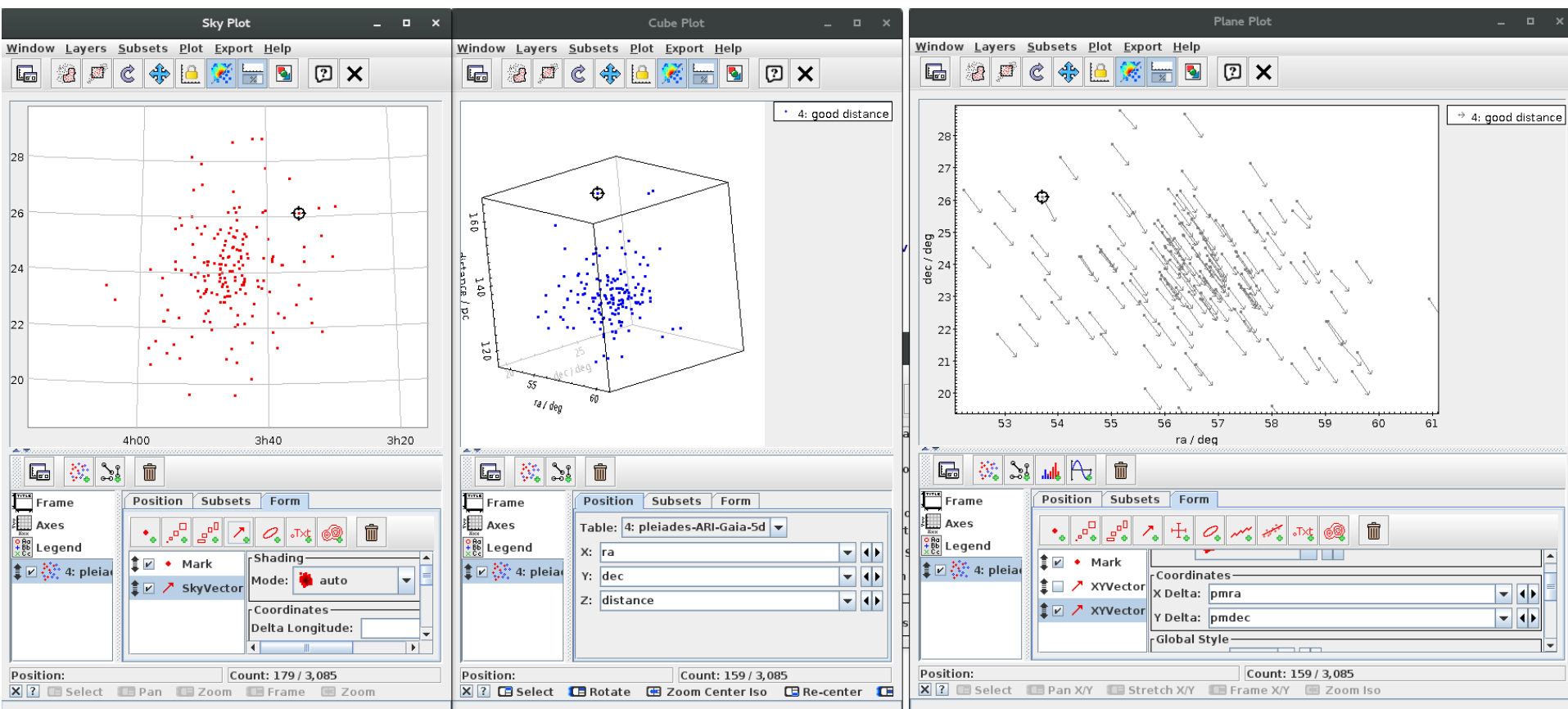
# TOPCAT & STILTS

- 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 \  

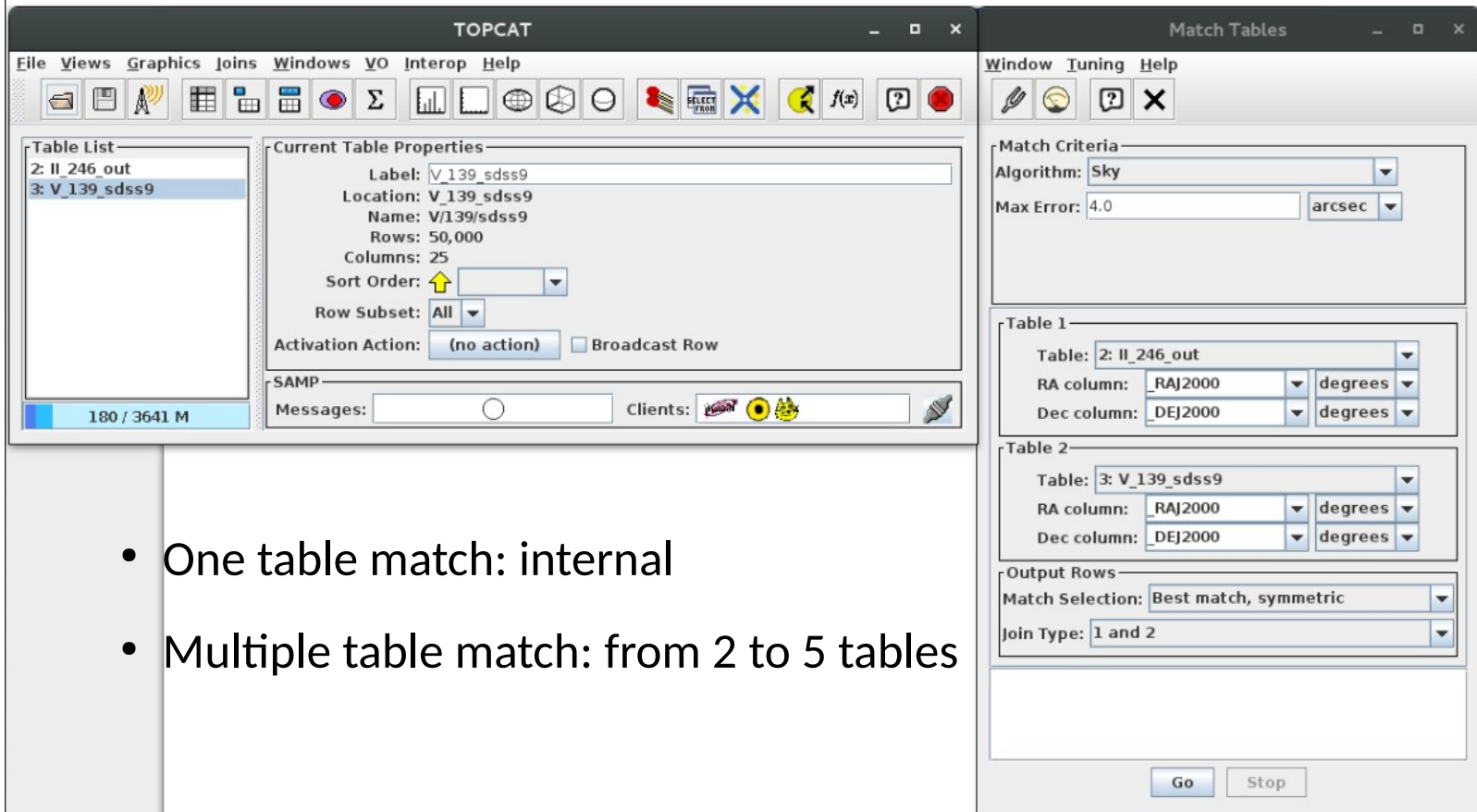
```

- 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 image shows two side-by-side windows from the TOPCAT software. The left window is the main TOPCAT interface, titled 'TOPCAT'. It has a menu bar (File, Views, Graphics, Joins, Windows, VO, Interop, Help) and a toolbar. On the left is a 'Table List' showing two tables: '2: II\_246\_out' and '3: V\_139\_sdss9'. The 'Current Table Properties' panel on the right shows details for 'V\_139\_sdss9': Label, Location, Name, Rows (50,000), Columns (25), Sort Order (up arrow), Row Subset (All), and Activation Action (no action). The bottom status bar shows '180 / 3641 M'. The right window is titled 'Match Tables' and contains settings for a table match. It has a 'Match Criteria' section with 'Algorithm' set to 'Sky' and 'Max Error' set to '4.0' in 'arcsec'. Below are sections for 'Table 1' and 'Table 2'. 'Table 1' is '2: II\_246\_out' with RA column 'RAJ2000' and Dec column 'DEJ2000', both in 'degrees'. 'Table 2' is '3: V\_139\_sdss9' with RA column 'RAJ2000' and Dec column 'DEJ2000', both in 'degrees'. The 'Output Rows' section has 'Match Selection' set to 'Best match, symmetric' and 'Join Type' set to '1 and 2'. At the bottom are 'Go' and 'Stop' buttons.

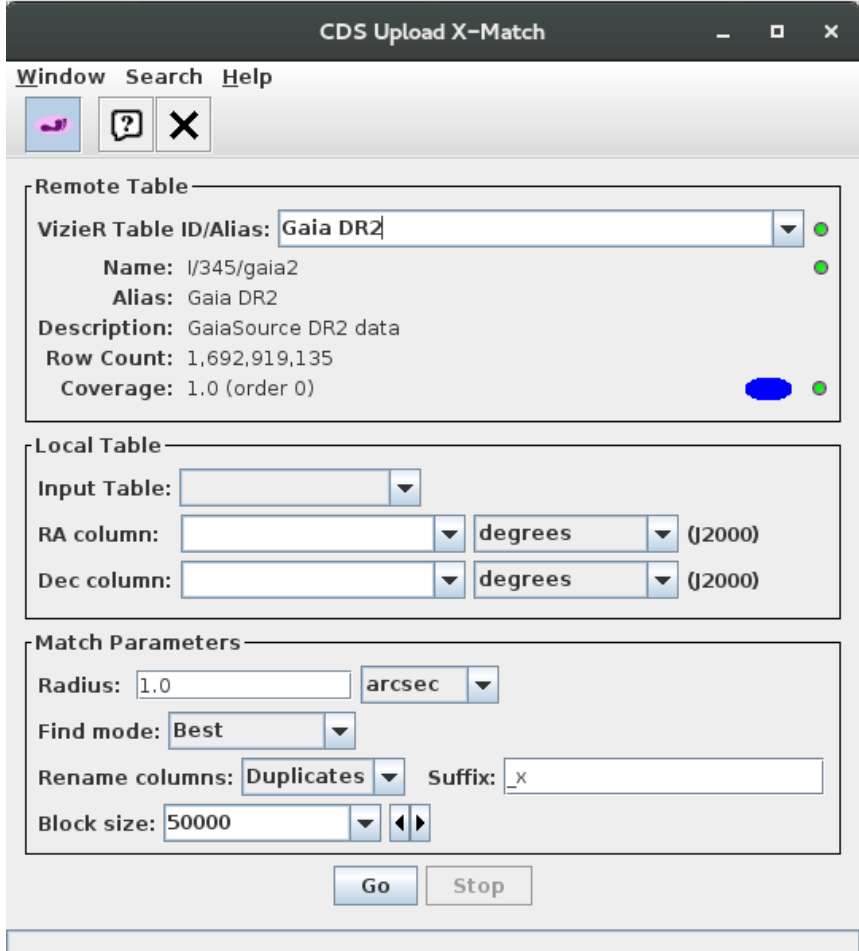
- One table match: internal
- Multiple table match: from 2 to 5 tables

# 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 (J2000)

Dec column: degrees (J2000)

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) ?

Radmind  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** Submit Reset All

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

Standard ☒ Original ☐

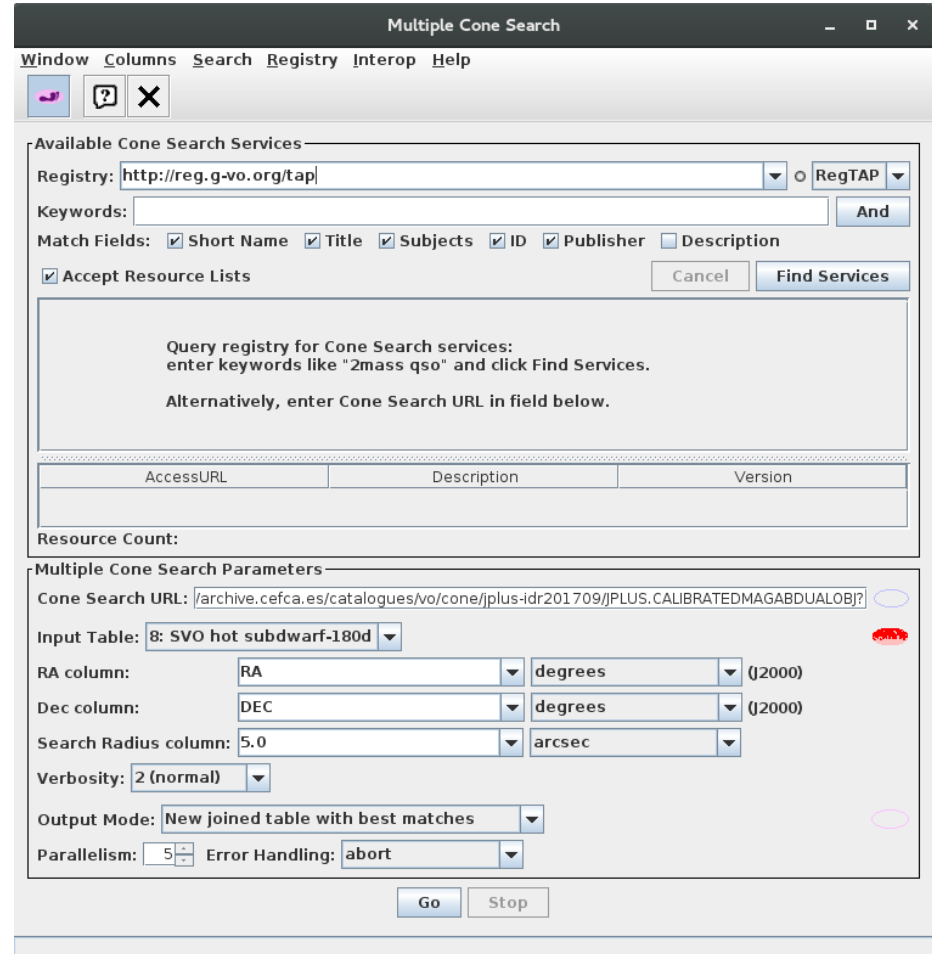
Show	Sort	Column	Clear	Constraint	Explain (UCD)
<input checked="" type="checkbox"/>	<input type="radio"/>	RAJ2000		<a href="#">deg</a> (i) Right ascension, J2000.0, at epoch 2000 ( <a href="#">raj2000</a> ) ( <a href="#">pos.eq.ra;meta.main</a> )	
<input checked="" type="checkbox"/>	<input type="radio"/>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



Multiple Cone Search

Window Columns Search Registry Interop Help

Available Cone Search Services

Registry:

Keywords:

Match Fields: ☒ Short Name ☒ Title ☒ Subjects ☒ ID ☒ Publisher ☐ Description

☒ Accept Resource Lists

Query registry for Cone Search services:  
enter keywords like "2mass qso" and click Find Services.  
Alternatively, enter Cone Search URL in field below.

AccessURL	Description	Version
-----------	-------------	---------

Resource Count:

Multiple Cone Search Parameters

Cone Search URL:

Input Table:

RA column:   (J2000)

Dec column:   (J2000)

Search Radius column:

Verbosity:

Output Mode:

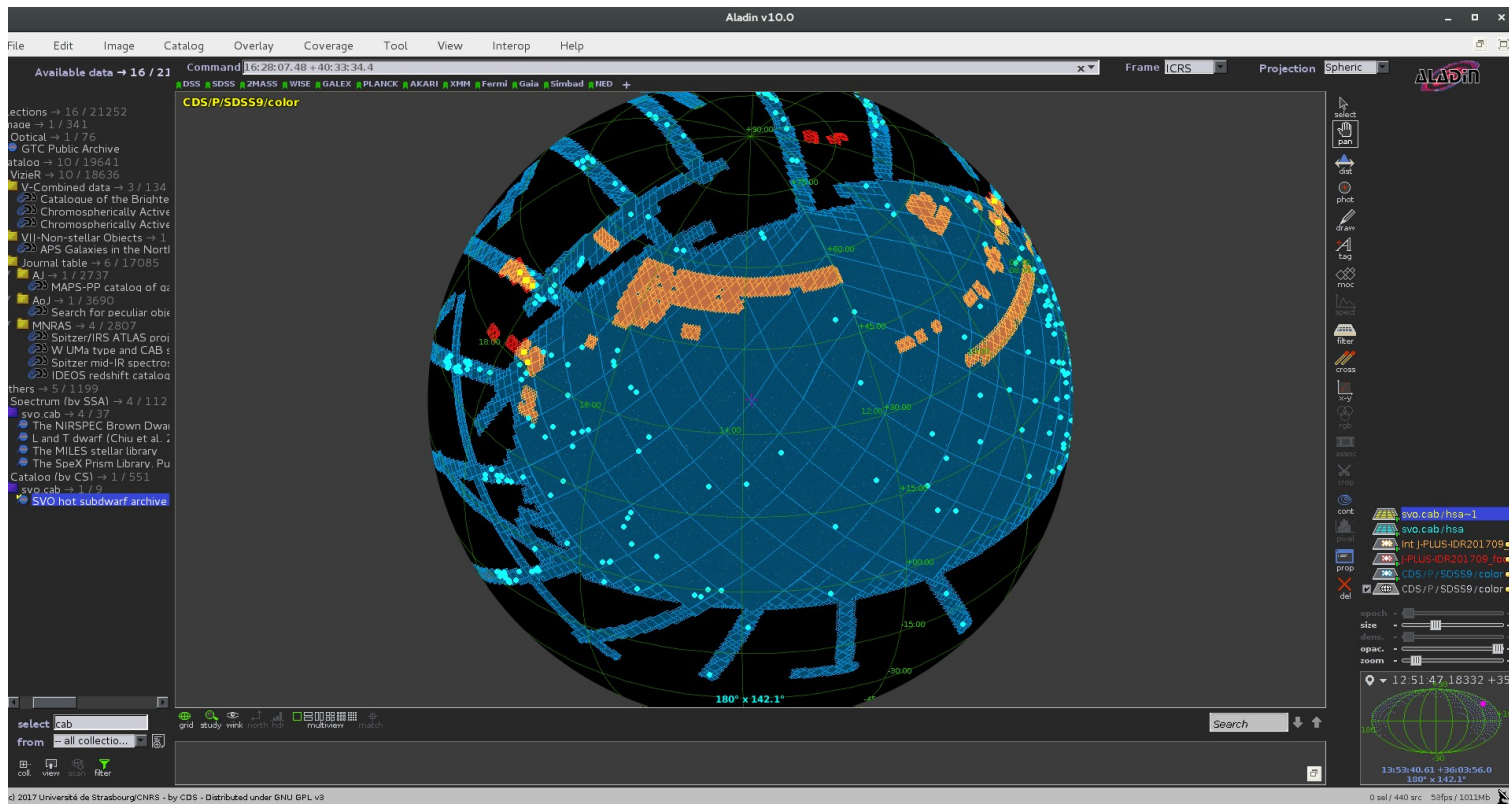
Parallelism:  Error Handling:

# 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



CDS X-Match Service

Choose tables to cross-match

Gala DR2 PanSTARRS DR1

VizieR SIMBAD My store VizieR SIMBAD My store

Gaia DR2 (Gaia Collaboration, 2018)  
1,692,919,135 rows

The Pan-STARRS release 1 (PS1) Survey - DR1 (Chambers+, 2016)  
1,919,106,885 rows

Show options

Begin the X-Match

Visualize and manage your cross-match jobs

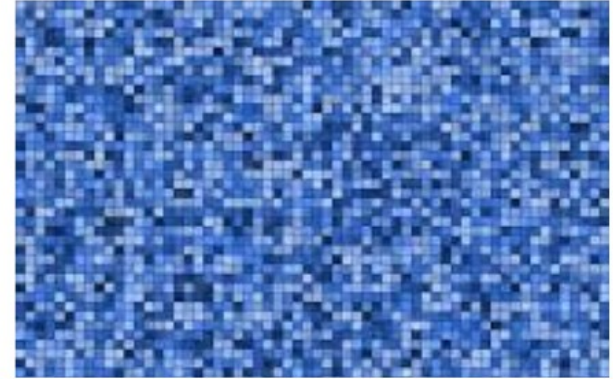
Table 1	Table 2	Options	Begin	Status	Actions
No job in list					

For the selected job(s): Delete

# TOPCAT & STILTS

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

- STILTS



- Cross-match

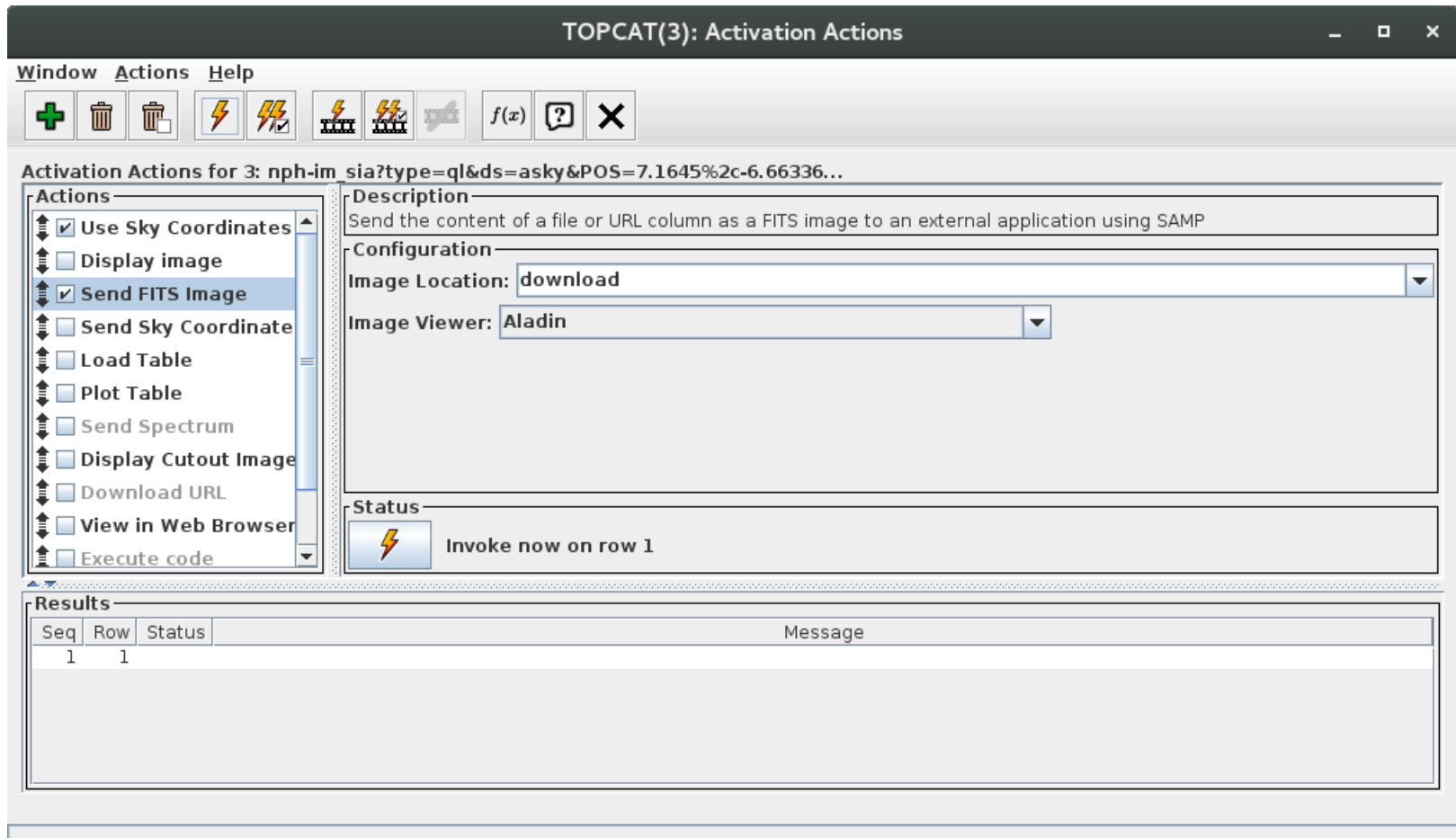
```
java -jar stilts.jar tskymatch2 ifmt1=votable in1=2mass.xml ifmt2=votable  
in2=sdss.xml ra1="RAJ2000" dec1="DEJ2000" ra2="RAJ2000" dec2="DEJ2000"  
error=10 find=all out=cross.xml ofmt=votable'
```

- 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
```

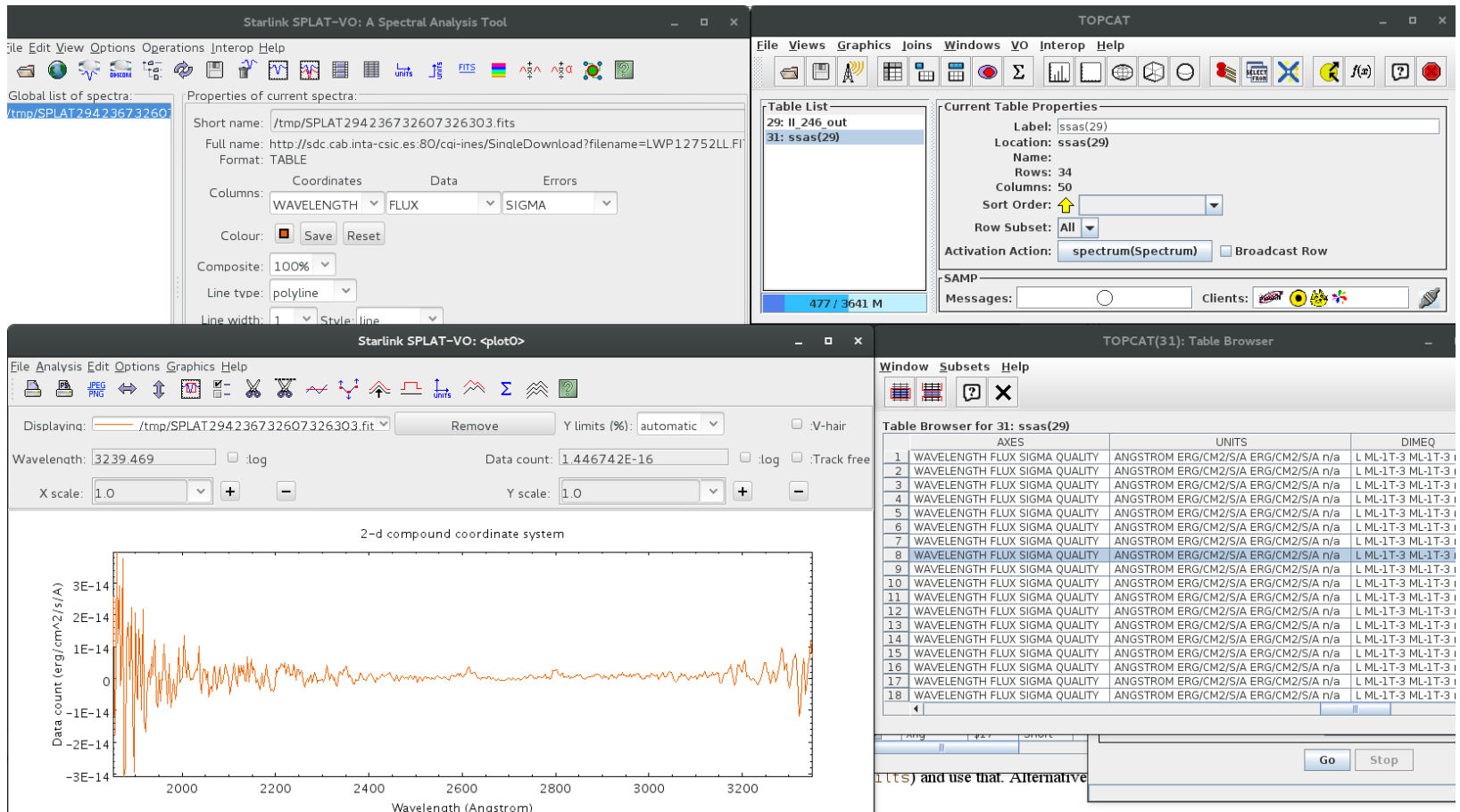
# TOPCAT & STILTS

- Activation actions and activation window

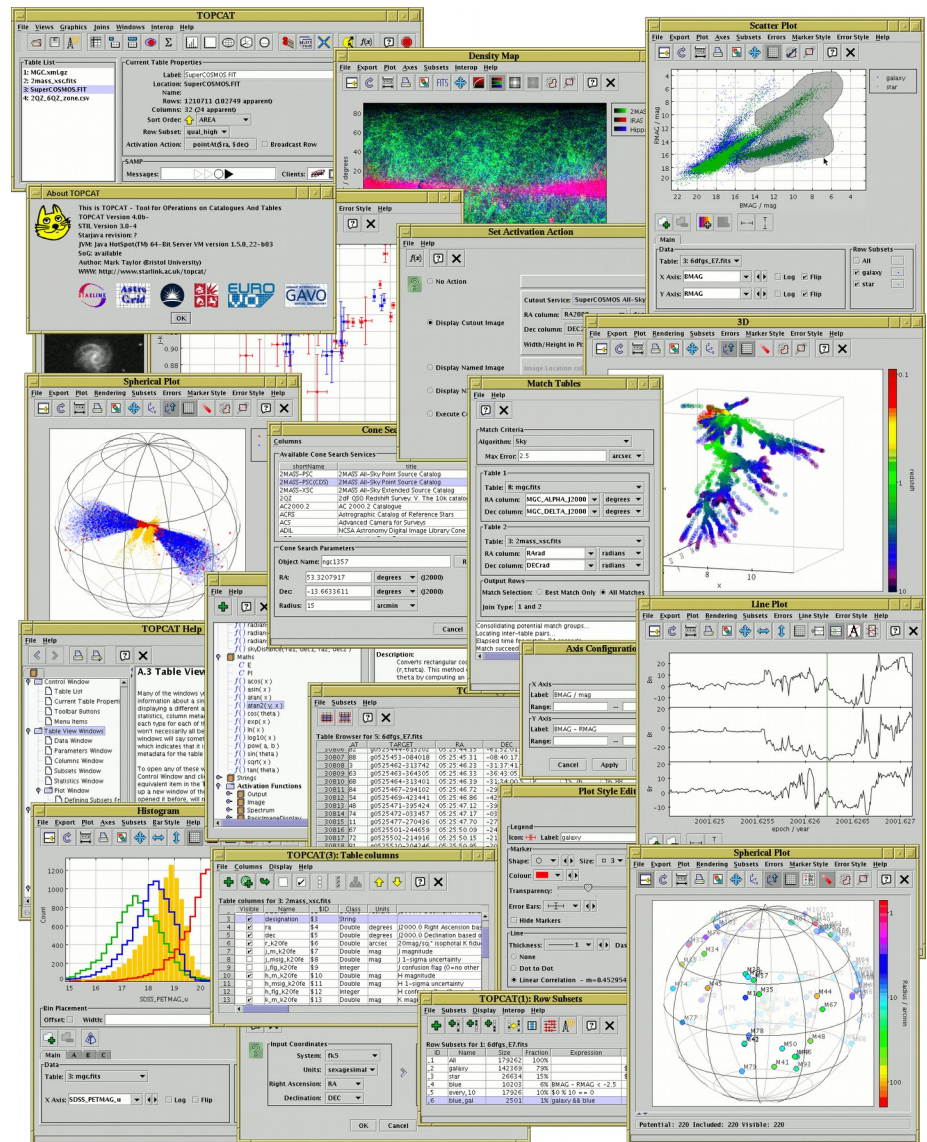


# TOPCAT & STILTS

- Activation actions and activation window



The screenshot shows the Cube2 software interface. At the top is a menu bar with 'File' and 'Help'. Below it is a toolbar with various icons for file operations, navigation, and visualization. The main window displays a 3D plot titled '3D vectors'. The plot shows a dense cluster of points, with red points forming the core and purple points forming the outer shell. Blue arrows represent vectors originating from the points. The axes are labeled x, y, and z. A color bar on the right indicates the 'sub-HaloID' values, ranging from 2.0e13 to 6.0e13. The status bar at the bottom shows 'Position:' and 'Count: 534 / 1,915'.



# 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/>