# 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

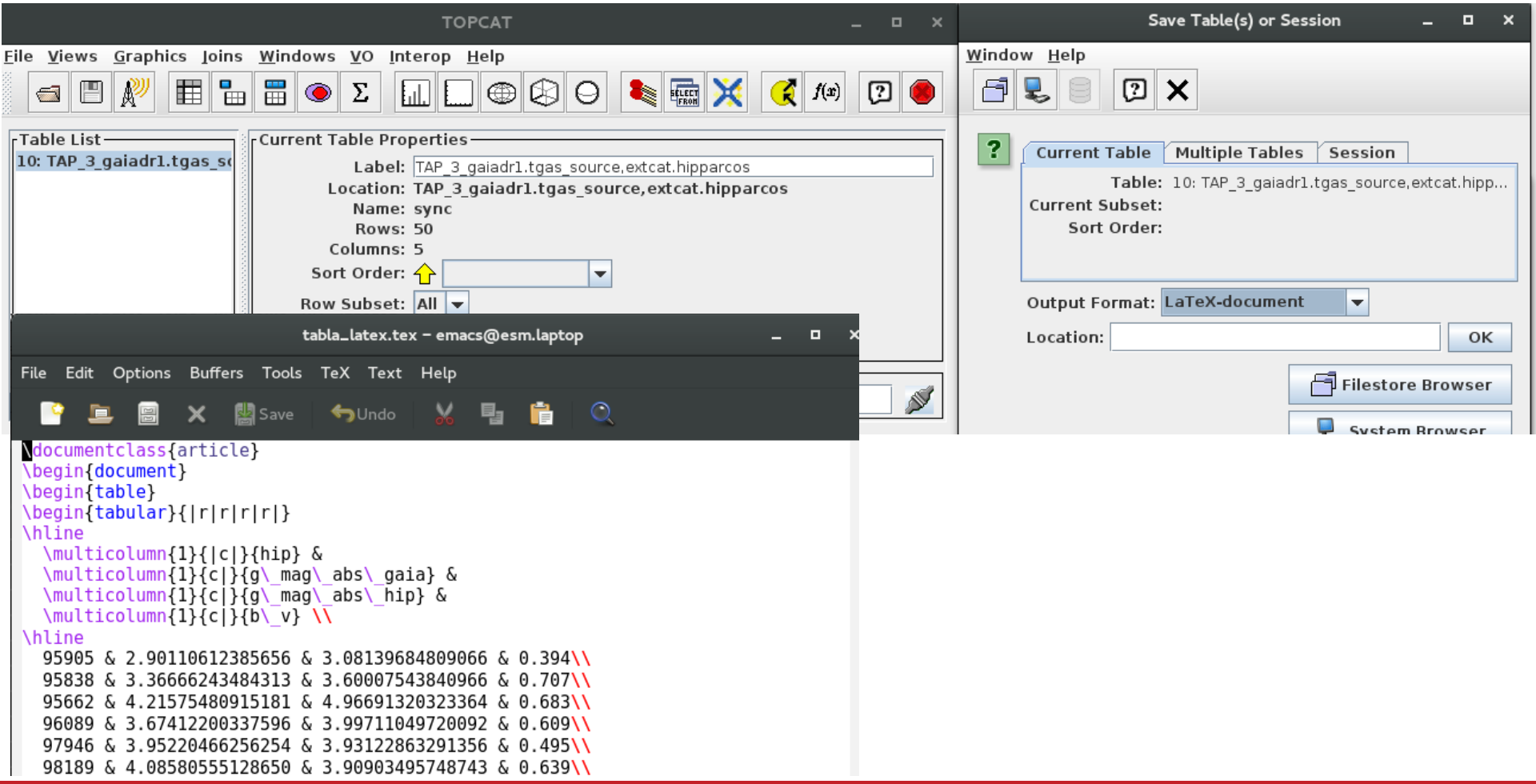
- 4.1.1.1 FITS
- 4.1.1.2 Column-oriented FITS
- 4.1.1.3 VOTable
- 4.1.1.4 CDF
- 4.1.1.5 ASCII
- 4.1.1.6 IPAC
- 4.1.1.7 Comma-Separated Values
- 4.1.1.8 GBIN
- 4.1.1.9 Tab-Separated Table
- 4.1.1.10 SQL Database Queries
- 4.1.1.11 World Data Center

- Which is the best format?
  - 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*

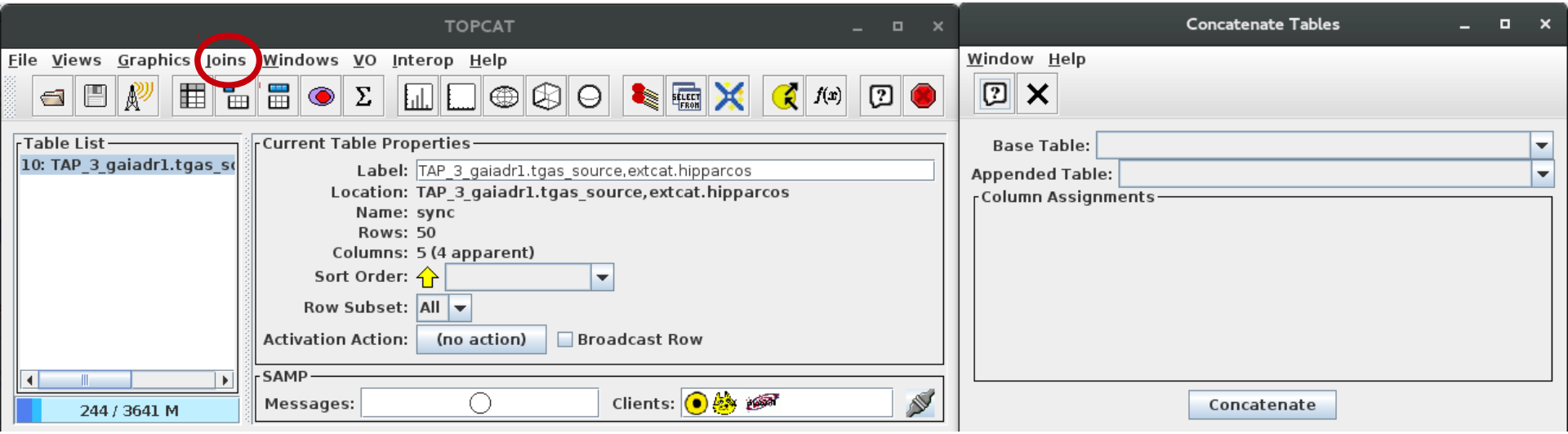
# TOPCAT & STILTS

- Output in Latex



# 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

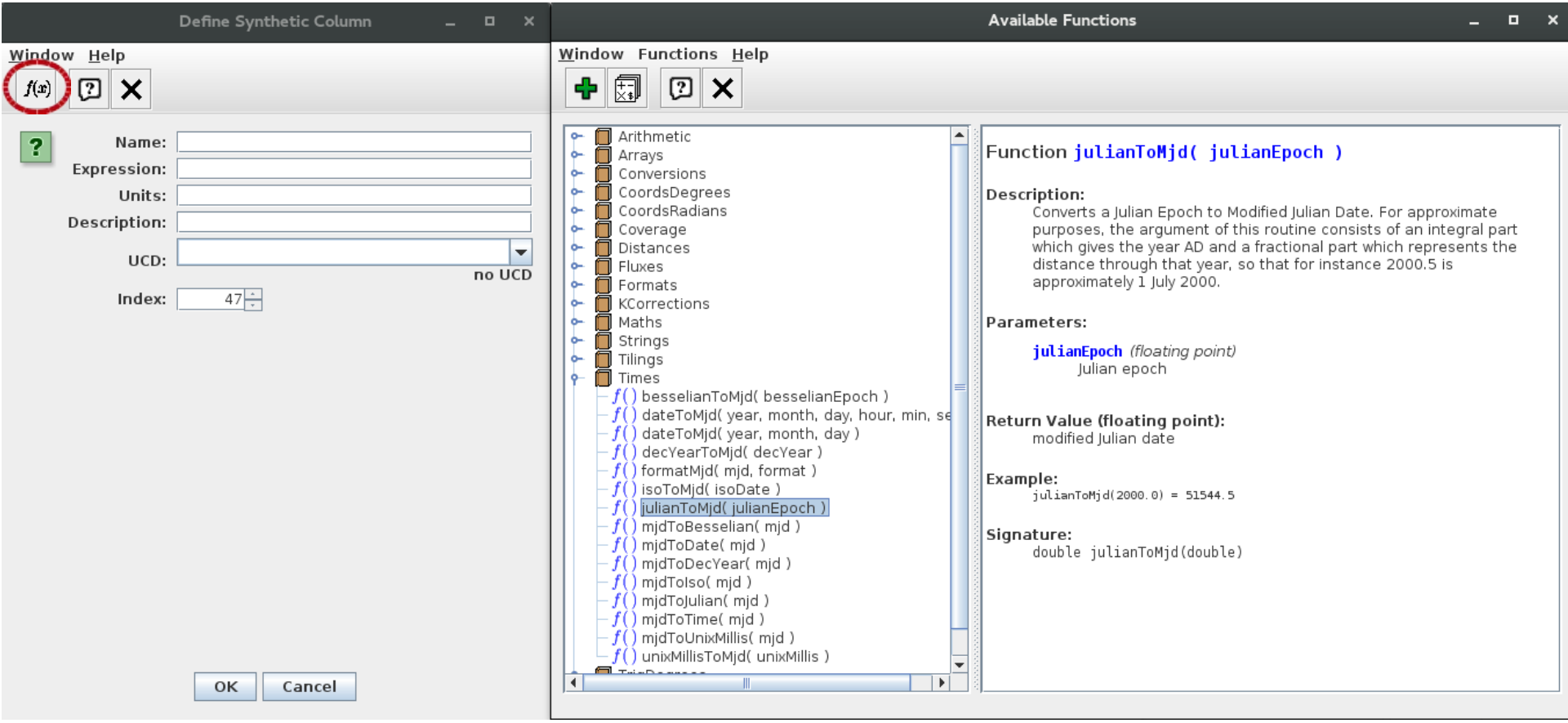
- Concatenating multiple tables in STILTS

```
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.

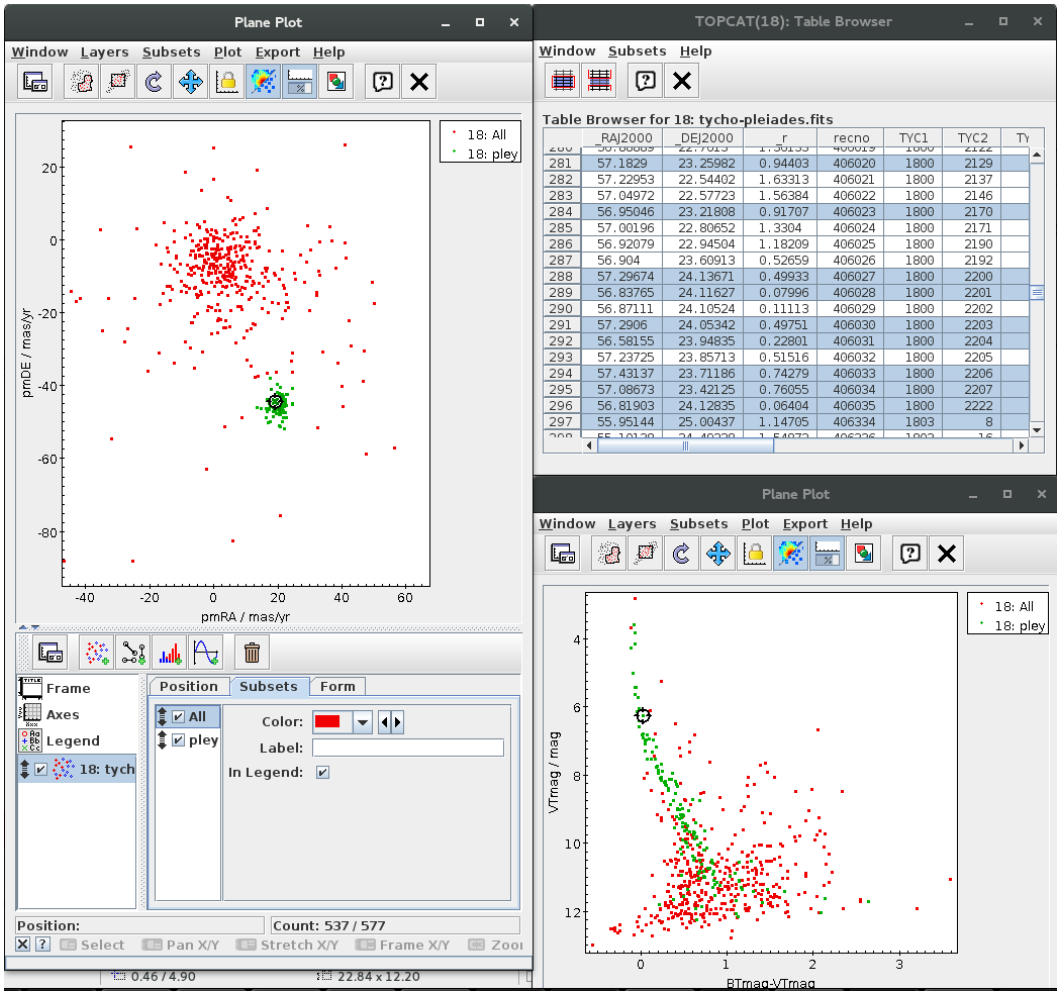
# TOPCAT & STILTS

- Functions in TOPCAT



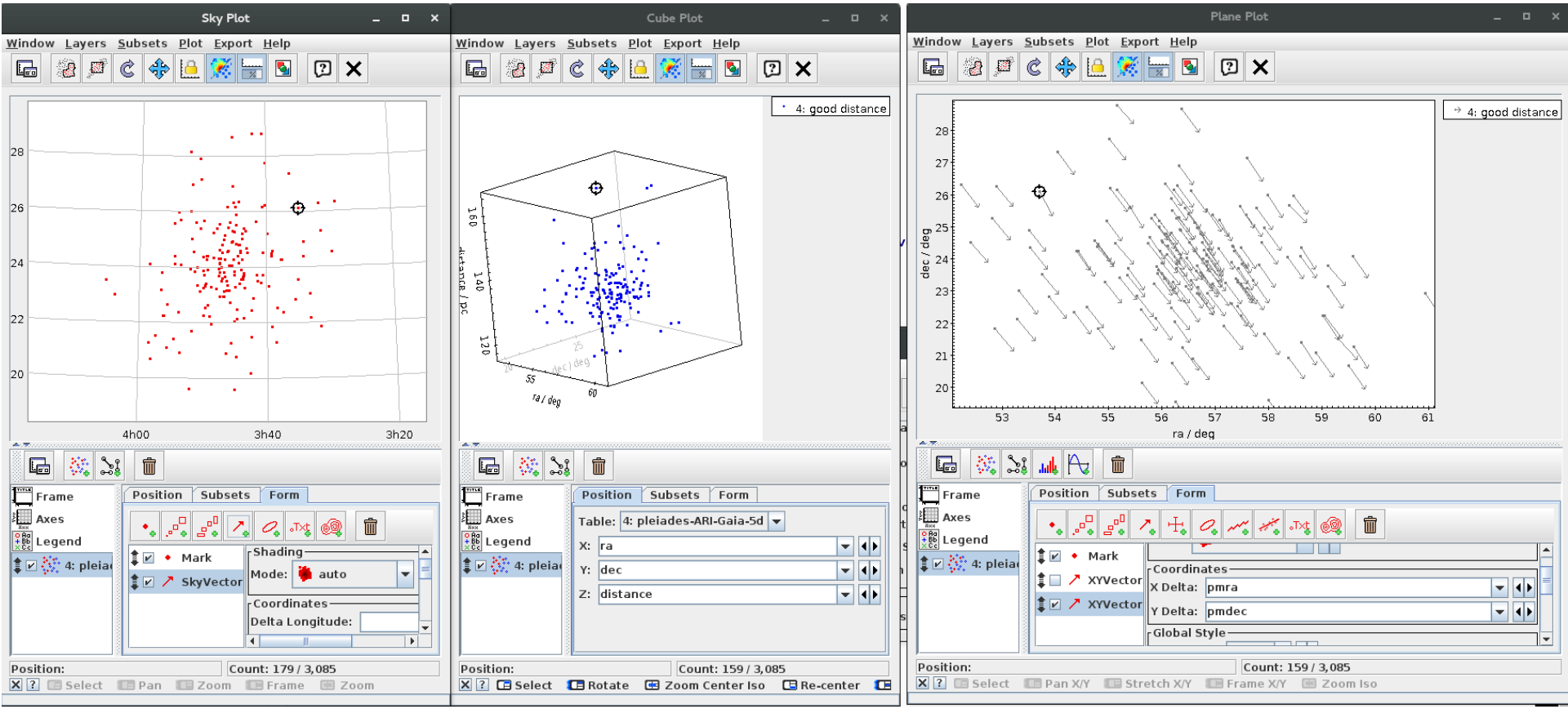
# TOPCAT & STILTS

- Linked views



# TOPCAT & STILTS

- Linked views



# TOPCAT & STILTS

- Crossmatching 

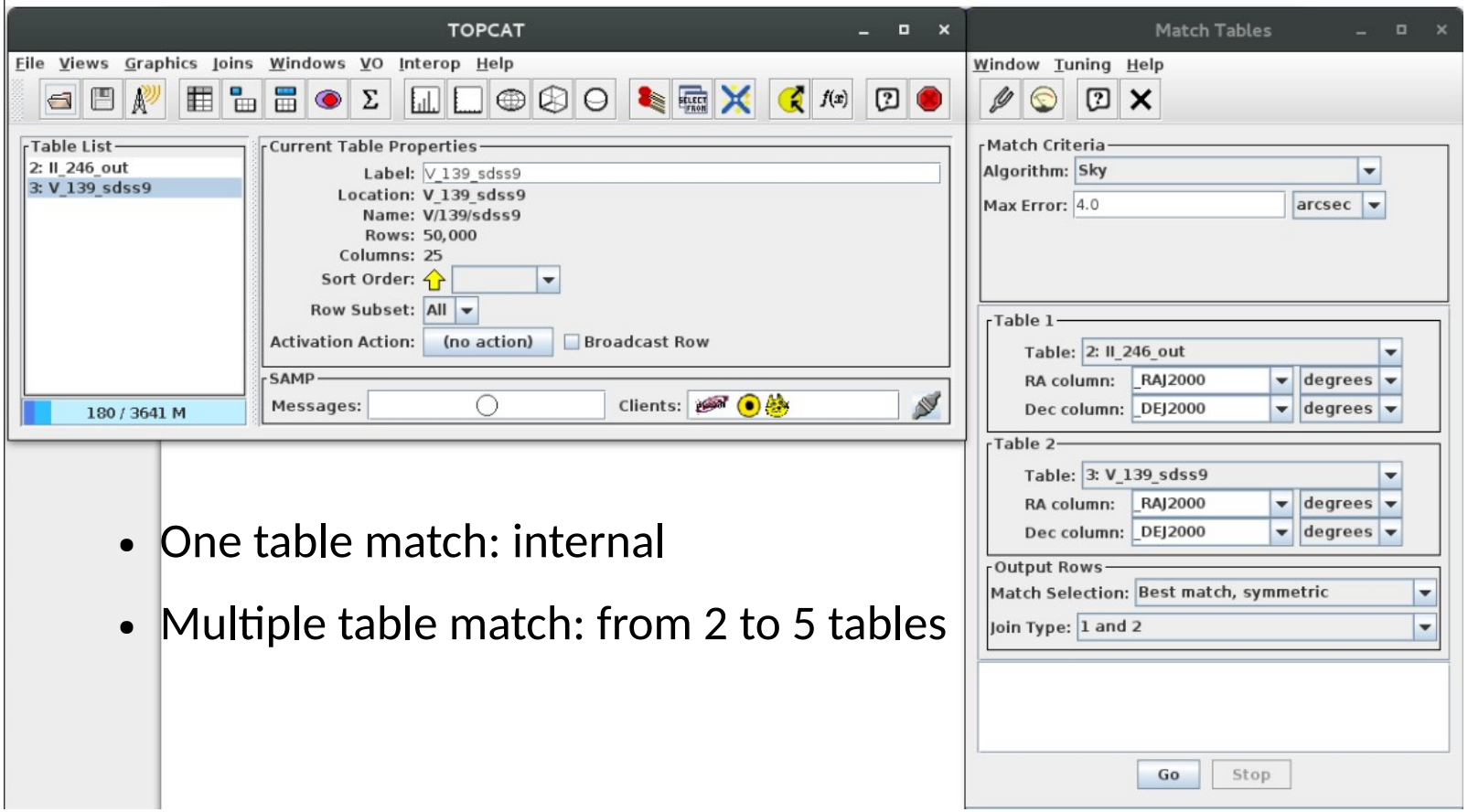
```
stilts tskymatch2 \
       in1=tycho-pleiades.fits ra1=_RAJ2000 dec1=_DEJ2000 \
       in2=2mass-pleiades.fits ra2=_RAJ2000 dec2=_DEJ2000 \
       join=1and2 find=best error=1 \
       out=tycho-2mass.fits \
```

  - 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



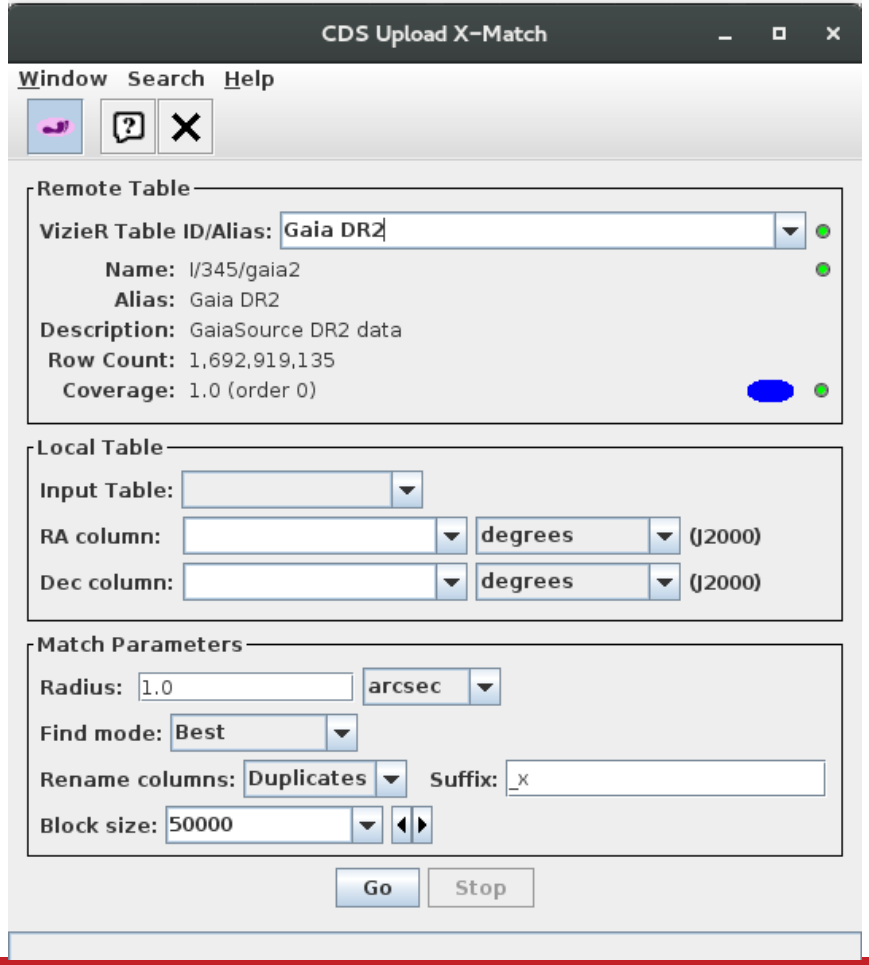

- One table match: internal
- Multiple table match: from 2 to 5 tables

# 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



# TOPCAT & STILTS

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

I/339 Hot Stuff for One Year (HSOY) (Altmann+, 2017) 2017A&A...600L...4A ReadMe+ftp Similar Catalogs

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

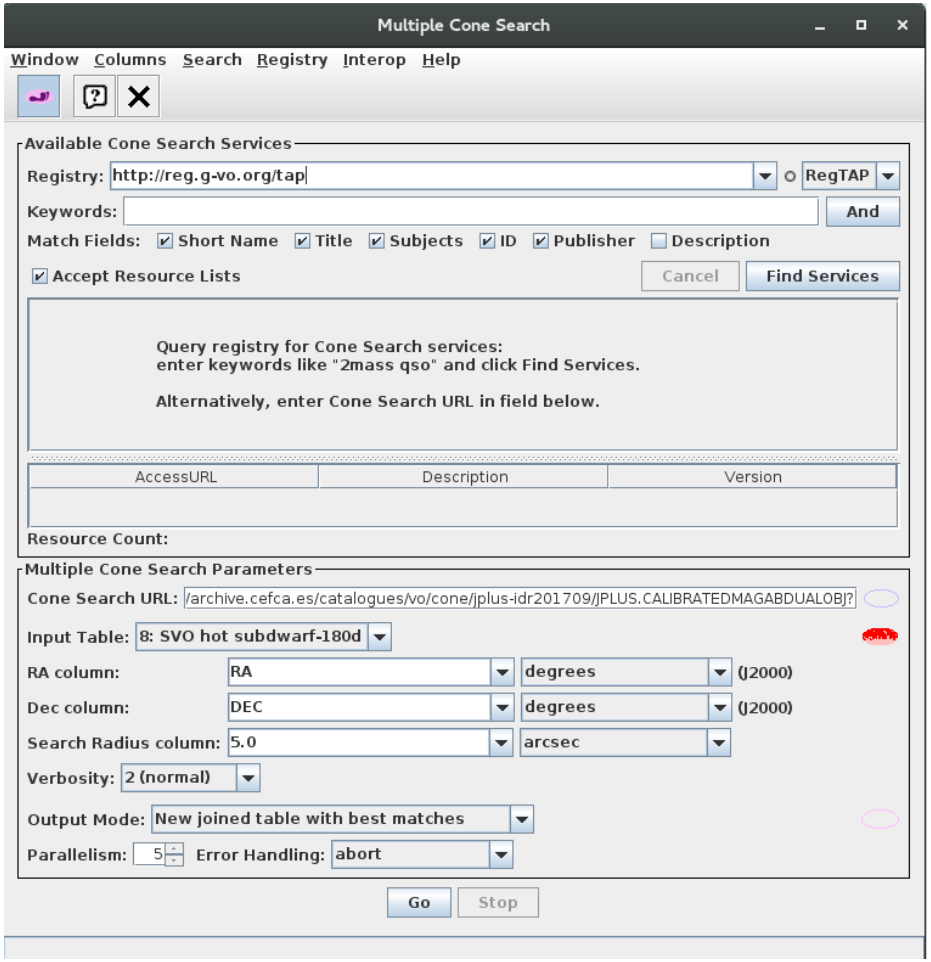
| Show | Sort | Column | Clear | Constraint | | Explain (UCD) |
|---|---|---|---|---|---|---|
| ☑ | ○ | RAJ2000 | | | deg | (i) Right ascension, J2000.0, at epoch 2000 (raj2000) (pos.eq.ra;meta.main) |
| ☑ | ○ | DEJ2000 | | | deg | (i) Declination, J2000.0, at epoch 2000 (dej2000) (pos.eq.dec;meta.main)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

# TOPCAT & STILTS

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

*TOPCAT → VO → Multicone*

- Disadvantages:
  - slow

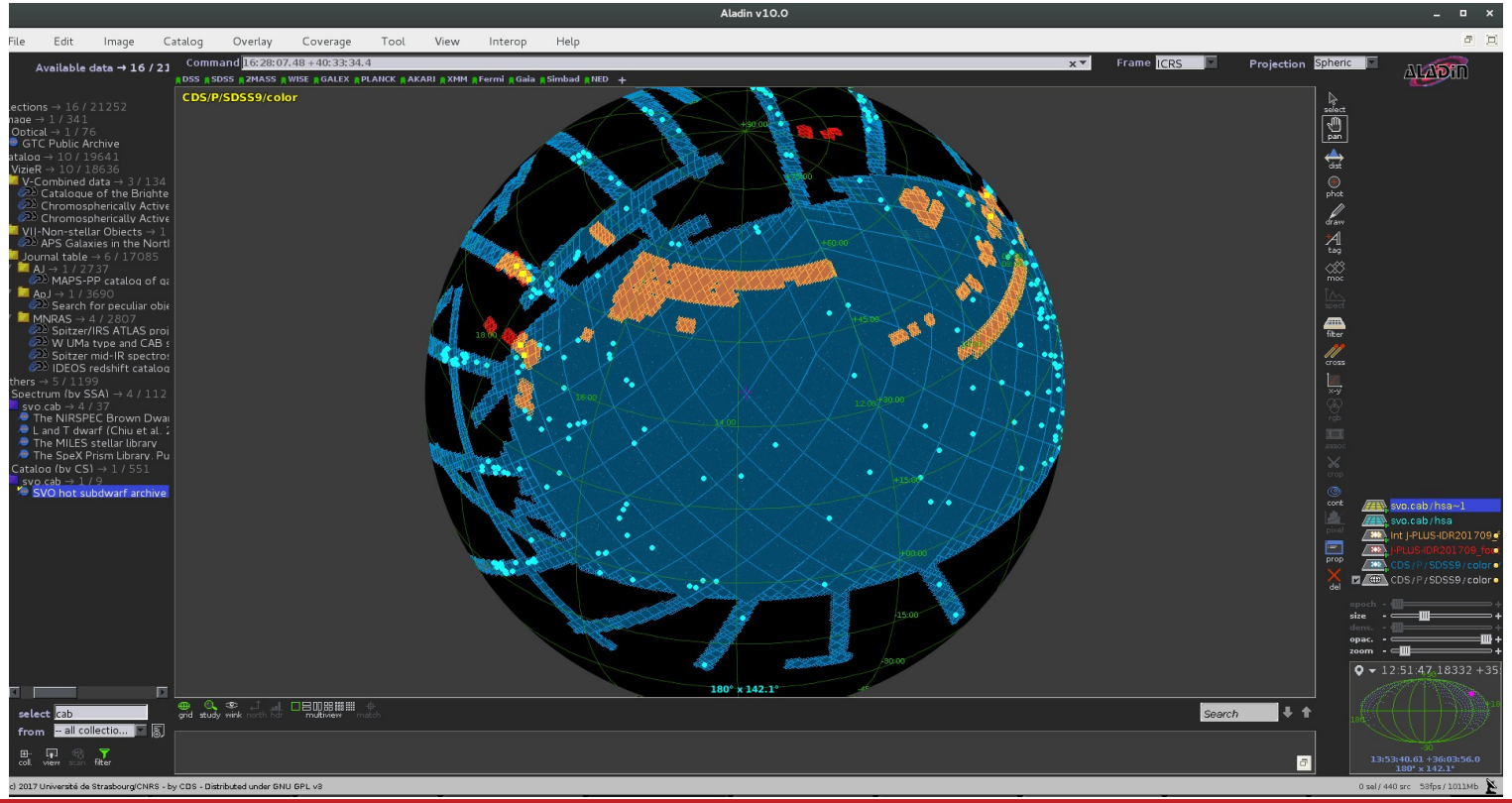


# TOPCAT & STILTS

- 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

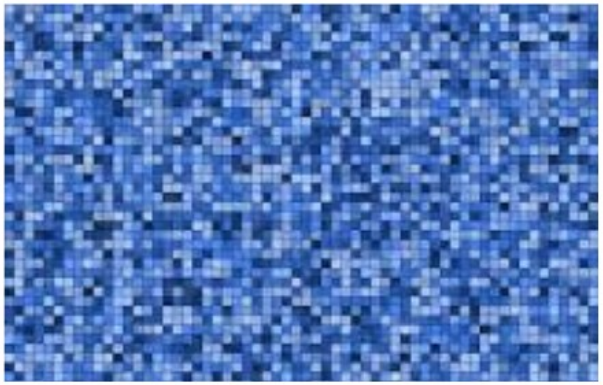
- Crossmatching - How to x-match **two large catalogues** (in CDS) ?

  - Disadvantages:
    - No filtering
      → Large outputs

# TOPCAT & STILTS

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

  - STILTS



- Cross-match

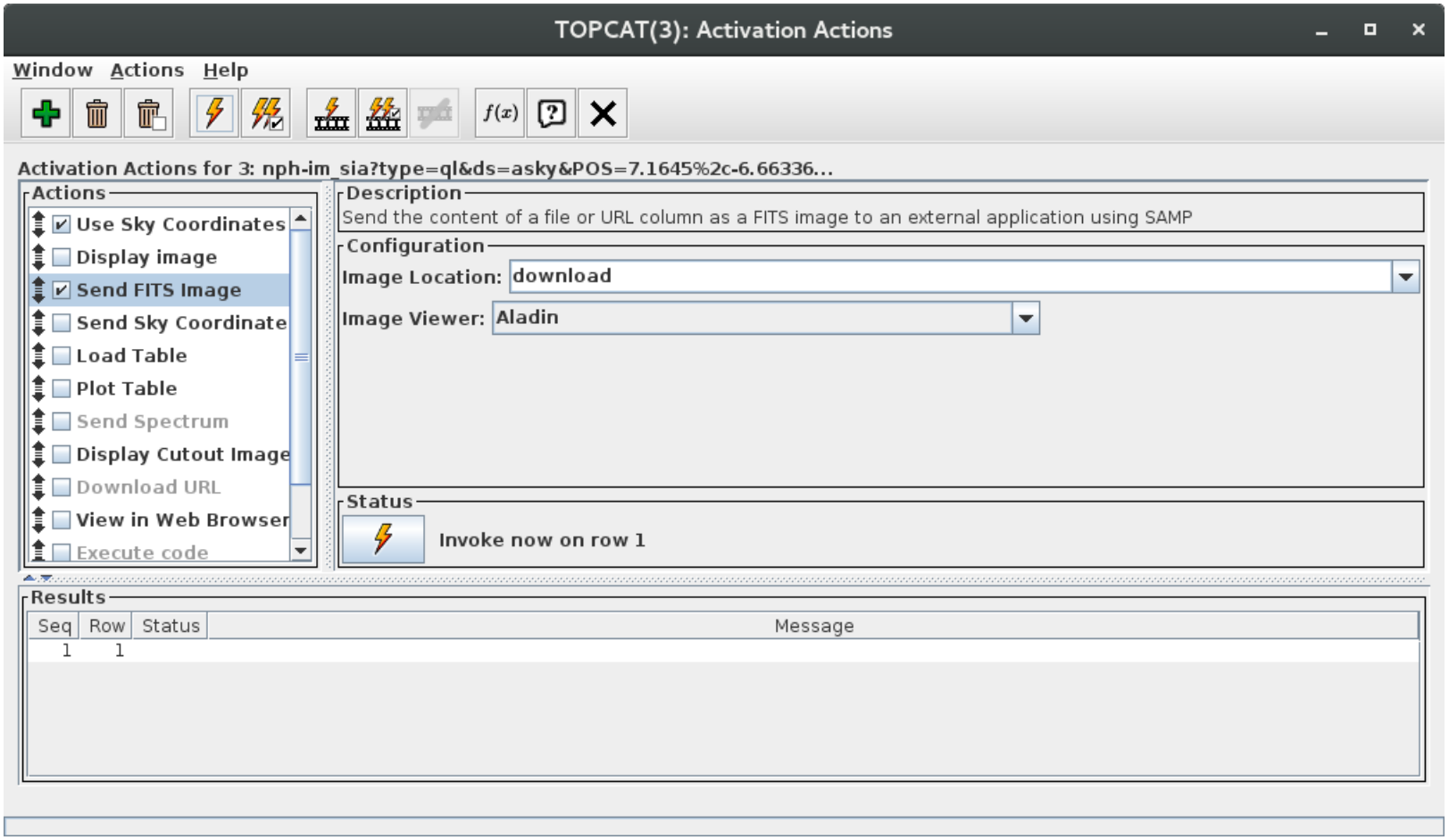
```
java -jar stilts.jar tskymatch2 ifmt1=votable in1=2mass.xml ifmt2=votable
in2=sdss.xml ra1="RAJ2000" dec1="DEJ2000" ra2="RAJ2000" dec2="DEJ2000"
error=10 find=all out=cross.xml ofmt=votable'
```

- 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
```

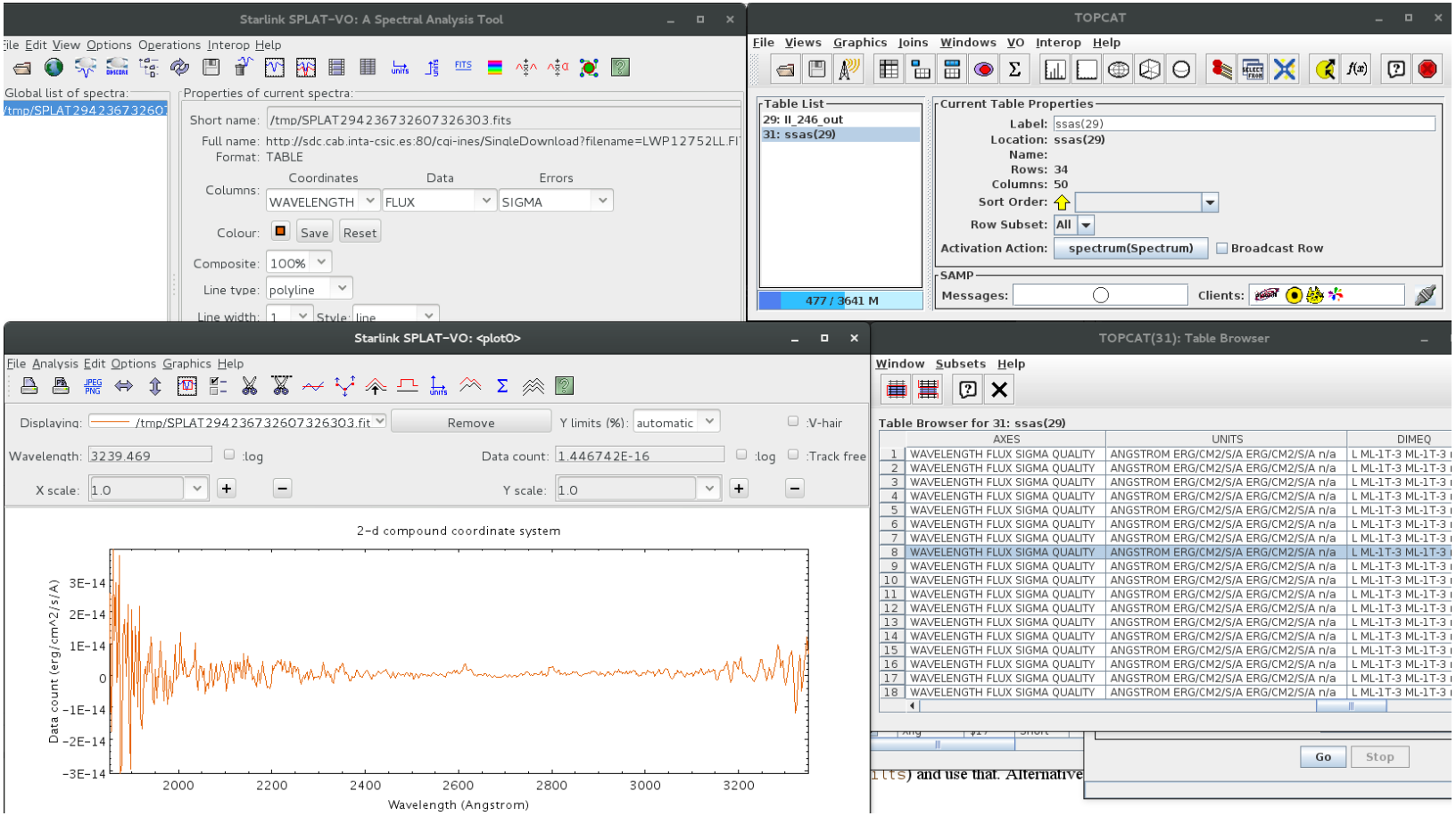
# TOPCAT & STILTS

- Activation actions and activation window

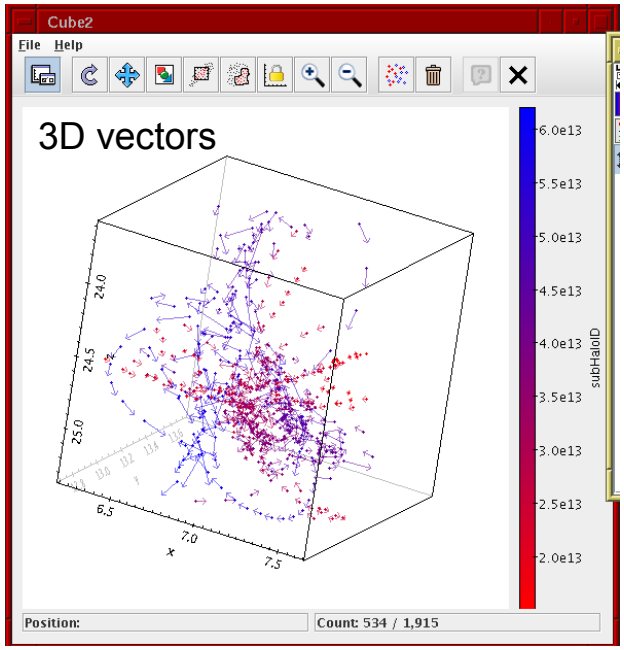


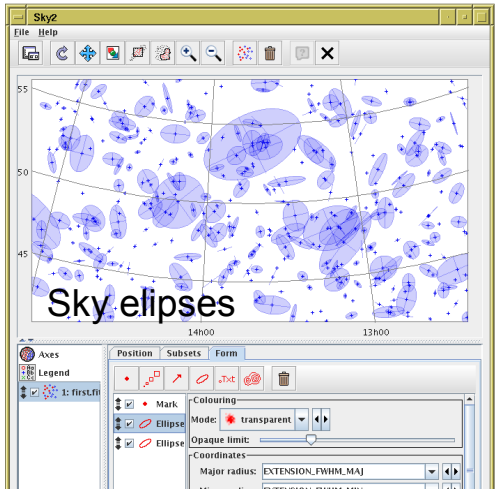
# TOPCAT & STILTS

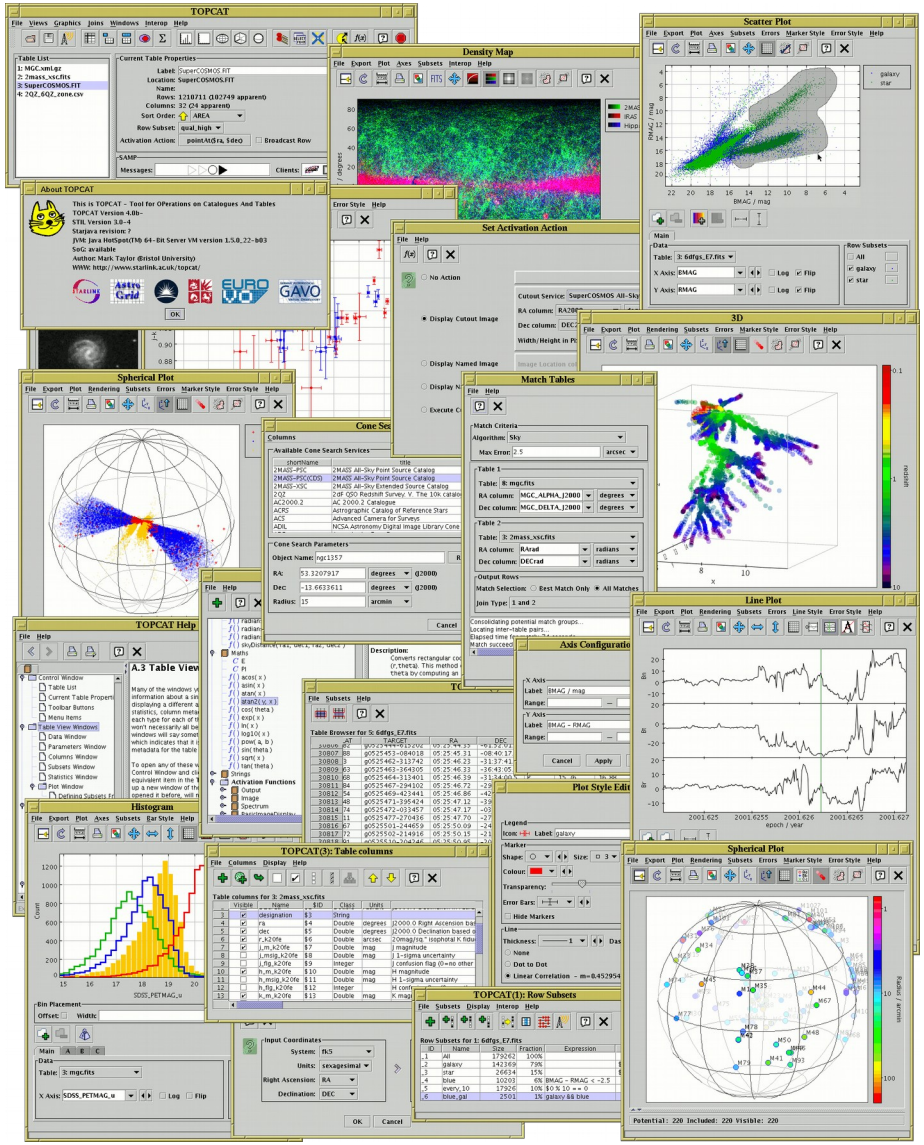
- Activation actions and activation window



# 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/