

# *Advanced TOPCAT-STILTS*

## Enrique Solano



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?

- Small table (<1000 rows): **doesn't matter.**
- Medium-sized (rows\*cols) < 20million): **FITS.**
- Big (millions of rows, especially with lots of columns): **colfits.**

- [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](#)

- 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 Emacs window shows LaTeX code for a table with 5 columns and 5 rows, with columns labeled 'hip', 'gaia', 'hip', and 'v'. The table data is as follows:

hip	gaia	hip	v
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

# TOPCAT & STILTS

- Concatenating tables in TOPCAT

The image shows two windows from the TOPCAT software. The main window, titled 'TOPCAT', has a menu bar with 'File', 'Views', 'Graphics', 'Joins', 'Windows', 'VO', 'Interop', and 'Help'. The 'Joins' menu is circled in red. Below the menu bar is a toolbar with various icons. The 'Table List' on the left shows a table named '10: TAP\_3\_gaiadr1.tgas\_s...'. The 'Current Table Properties' panel shows details for a table: Label: TAP\_3\_gaiadr1.tgas\_source,extcat.hipparcos; Location: TAP\_3\_gaiadr1.tgas\_source,extcat.hipparcos; Name: sync; Rows: 50; Columns: 5 (4 apparent); Sort Order: (arrow icon); Row Subset: All; Activation Action: (no action) [Broadcast Row]. The 'SAMP' section shows 'Messages: [ ]' and 'Clients: [ ]'. The 'Concatenate Tables' dialog box is open on the right, with fields for 'Base Table:' and 'Appended Table:', and a 'Concatenate' button at the bottom.

- 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.
- Similar columns.

# 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

- Eliminating duplicated rows in TOPCAT

The screenshot displays the TOPCAT software interface with three main windows:

- TOPCAT (Table Browser):** Shows a list of tables (15: V\_139\_sdss9, 16: II\_246\_out, 17: match(16,15)) and current table properties for 'match(16,15)', including 63 rows and 46 columns.
- Match Tables:** Shows match criteria (Algorithm: Sky, Max Error: 5.0 arcsec) and table selection (Table 1: 16: II\_246\_out, Table 2: 15: V\_139\_sdss9). The 'Output Rows' section shows 'Match Selection: All matches' circled in red.
- Plane Plot:** A scatter plot of DEJ2000\_2 / deg vs RAJ2000\_2 / deg showing three data points connected by lines.

The Table Browser window displays the following data table:

RAJ2000_1	DEJ2000_1	r_1	RAJ2000_1	DEJ2000_1	2MASS	Jmag	e_Jmag	Hm...	
46	10.01233	9.9538	0.04777	10.01233	9.9538	00400295+0957136	16.446	0.122	15.61
47	10.01233	9.9538	0.04777	10.01233	9.9538	00400295+0957136	16.446	0.122	15.61
48	10.05632	9.95983	0.06848	10.05632	9.95983	00401351+0957353	15.831	0.072	15.35
49	10.01043	9.95304	0.04807	10.01043	9.95304	00400250+0957109	16.591	0.149	16.11
50	10.08967	9.99552	0.08842	10.08967	9.99552	00402152+0959438	16.011	0.09	15.19
51	10.08967	9.99552	0.08842	10.08967	9.99552	00402152+0959438	16.011	0.09	15.19
52	10.03678	10.06028	0.07032	10.03678	10.06028	00400882+1003370	15.354	0.046	14.88
53	10.00555	10.0184	0.0192	10.00555	10.0184	00400133+1001062	16.173	0.098	15.51
54	10.00555	10.0184	0.0192	10.00555	10.0184	00400133+1001062	16.173	0.098	15.51
55	10.00555	10.0184	0.0192	10.00555	10.0184	00400133+1001062	16.173	0.098	15.51
56	10.04468	10.0234	0.04984	10.04468	10.0234	00401072+1001242	15.546	0.062	15.17
57	10.01584	10.03806	0.04113	10.01584	10.03806	00400380+1002170	14.662	0.029	14.11
58	10.03417	10.01356	0.03628	10.03417	10.01356	00400820+1000488	14.333	0.029	13.64
59	10.0802	10.00655	0.07925	10.0802	10.00655	00401924+1000235	16.351	0.117	15.40
60	10.0832	10.04873	0.09533	10.0832	10.04873	00401996+1002554	16.103	0.083	15.48
61	10.09697	10.02244	0.0981	10.09697	10.02244	00402327+1001207	13.249	0.026	12.73
62	10.01117	10.0939	0.09455	10.01117	10.0939	00400268+1005380	16.373	0.127	15.99
63	10.03102	10.06308	0.07009	10.03102	10.06308	00400744+1003470	16.412	0.117	15.60



# TOPCAT & STILTS

- Eliminating duplicated rows in TOPCAT

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, the 'Table List' panel shows three tables: '15: V\_139\_sdss9', '16: ll\_246\_out', and '17: match(16,15)'. The 'Current Table Properties' panel for the selected table shows: Label: match(16,15), Location: match(16,15), Name: Joined, Rows: 63, Columns: 46, Sort Order: (up arrow icon), Row Subset: All, and Activation Action: (no action) with a 'Broadcast Row' checkbox. The status bar at the bottom indicates '141 / 3641 M' and 'SAMP'.

A secondary window titled 'Match Tables' is open on the right. It has a 'Window Tuning Help' menu and a toolbar. The 'Match Criteria' section shows 'Algorithm: Exact Value'. The 'Table' section shows 'Table: 17: match(16,15)' and 'Matched Value column: 2MASS'. The 'Action' section has three radio button options: 'Mark Groups of Rows', 'Eliminate All Grouped Rows' (which is selected and circled in red), and 'Eliminate All But First of Each Group'. Below this is a 'New Table With Groups of Size' field set to '2'. At the bottom of the window, a status bar shows 'Identifying isolated links...', 'Walking links...', 'Elapsed time for match: 0 seconds', and 'Match succeeded', with 'Go' and 'Stop' buttons.

# TOPCAT & STILTS

- Functions in TOPCAT

The screenshot displays two windows from the TOPCAT software. The left window, titled 'Define Synthetic Column', contains a form for defining a new column. The 'Name' field is empty, and the 'Index' field is set to 47. A red circle highlights the 'f(x)' icon in the top-left corner of the window. The right window, titled 'Available Functions', shows a tree view of function categories. The 'Arithmetic' category is expanded, and the 'julianToMjd( julianEpoch )' function is selected and highlighted. The right pane of this window provides details for the selected function, including its description, parameters, return value, and signature.

**Define Synthetic Column**

Window Help

f(x) ? X

? Name:

Expression:

Units:

Description:

UCD:  no UCD

Index:

OK Cancel

**Available Functions**

Window Functions Help

+ ? X

- Arithmetic
- Arrays
- Conversions
- CoordsDegrees
- CoordsRadians
- Coverage
- Distances
- Fluxes
- Formats
- KCorrections
- Maths
- Strings
- Tilings
- Times

f() besselianToMjd( besselianEpoch )

f() dateToMjd( year, month, day, hour, min, se

f() dateToMjd( year, month, day )

f() decYearToMjd( decYear )

f() formatMjd( mjd, format )

f() isoToMjd( isoDate )

f() **julianToMjd( julianEpoch )**

f() mjdToBesselian( mjd )

f() mjdToDate( mjd )

f() mjdToDecYear( mjd )

f() mjdToIso( mjd )

f() mjdToJulian( mjd )

f() mjdToTime( mjd )

f() mjdToUnixMillis( mjd )

f() unixMillisToMjd( unixMillis )

**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

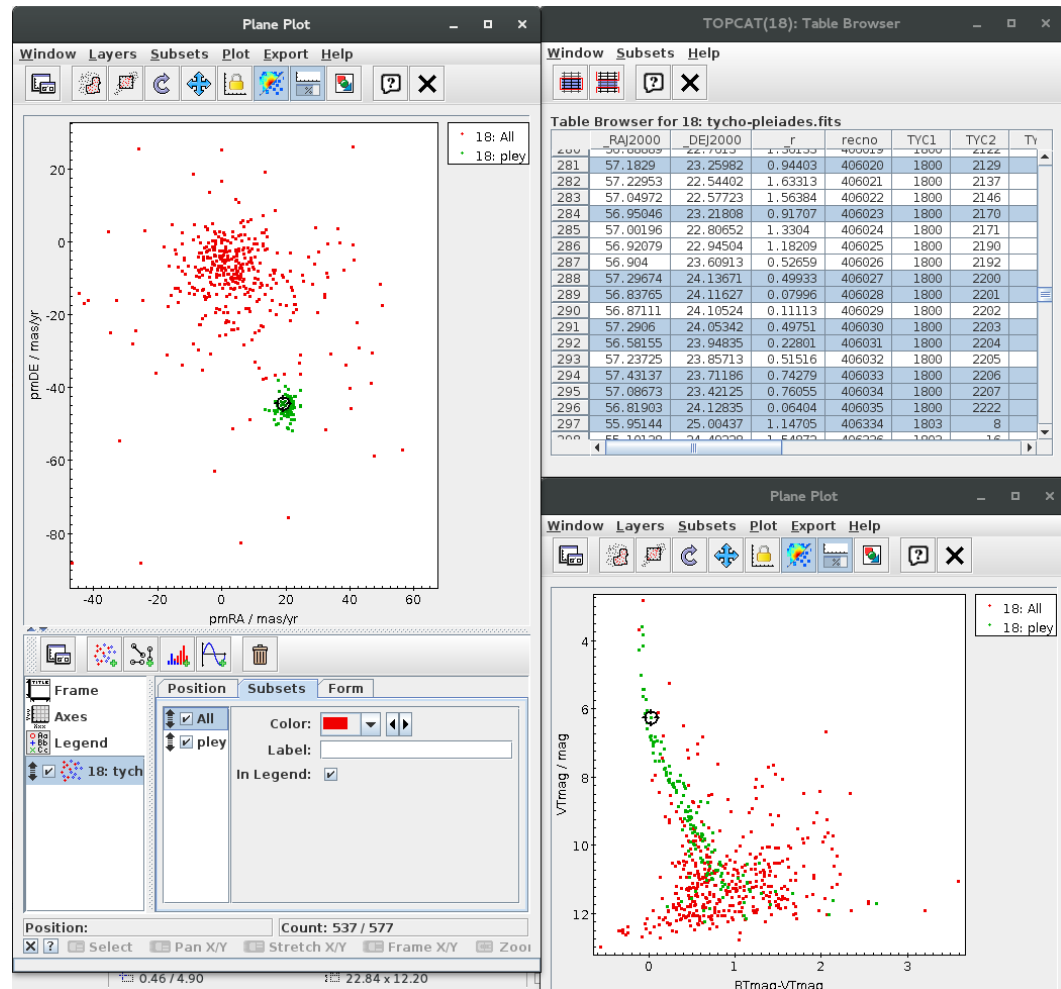
**Return Value (floating point):**  
modified julian date

**Example:**  
julianToMjd( 2000. 0 ) = 51544. 5

**Signature:**  
double julianToMjd(double)

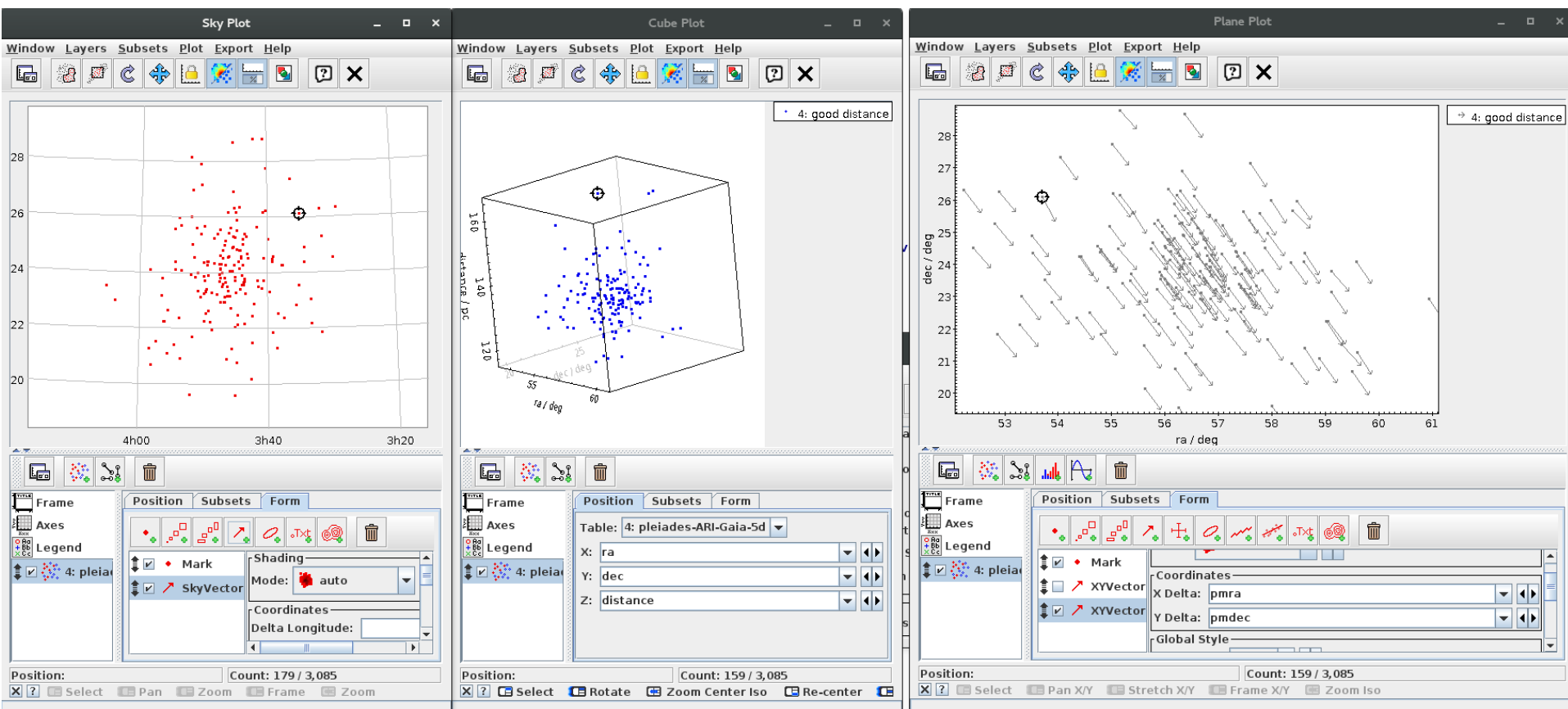
# TOPCAT & STILTS

- 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

- CDS X-match 
  - Very fast, even for large tables.
  - Only a restricted set of columns is available (not all from VizieR).
  - You can only specify simple positional criteria.

```
stilts cdsskymatch cdstable=2MASS \  
    in=tycho-pleiades.fits \  
    ra=_RAJ2000 dec=_DEJ2000 radius=1 \  
    find=best out=tycho-2mass.fits
```

# TOPCAT & STILTS

- Multicone 

- Slower but necessary when CDS X-match cannot be used.
- Multi-SIA, multi-SSA similar (but for images or spectra).

```
stilts coneskymatch \  
    serviceurl='http://vizier.u-strasbg.fr/viz-bin/votable/-A?-  
out.all&-source=II%2F246%2Fout&' \  
    in=tycho-pleiades.fits icmd=progress \  
    sr=0.0002777 find=best \  
    out=tycho-2mass.fits
```

# TOPCAT & STILTS

- Activation windows

The image displays four windows from the Starlink SPLAT-VO and TOPCAT software suite:

- Starlink SPLAT-VO: A Spectral Analysis Tool**: Shows the 'Properties of current spectra' for a file named `/tmp/SPLAT294236732607326303.fits`. The 'Columns' section is set to 'WAVELENGTH', 'FLUX', and 'SIGMA'. The 'Line type' is 'polyline'.
- TOPCAT**: Shows the 'Table List' with two tables: '29: ll\_246\_out' and '31: ssas(29)'. The 'Current Table Properties' for '31: ssas(29)' are displayed, including 'Label: ssas(29)', 'Location: ssas(29)', 'Name: ssas(29)', 'Rows: 34', and 'Columns: 50'. The 'Activation Action' is set to 'spectrum(Spectrum)'.
- Starlink SPLAT-VO: <plotO>**: Shows a 2-d compound coordinate system plot. The y-axis is 'Data count (erg/cm2/s/A)' ranging from  $-3E-14$  to  $3E-14$ . The x-axis is 'Wavelength (Angstrom)' ranging from 2000 to 3200. The plot shows a noisy spectrum with a peak around 2100 Angstroms.
- TOPCAT(31): Table Browser**: Shows a table with 18 rows and 6 columns. The columns are 'AXES', 'UNITS', and 'DIMEQ'. The rows contain data for 'WAVELENGTH FLUX SIGMA QUALITY' and 'ANGSTROM ERG/CM2/S/A ERG/CM2/S/A n/a'.





# TOPCAT & STILTS

- More at:

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

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

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