Data mining with TOPCAT and ADQL

Creating a target list

Harry Dawson
Research workshop on evolved stars
30.08.2022

Overview

- Topcat
  - Basic overview
  - Table visualisation/manipulation
  - Visualisation tools
  - Crossmatching
- ADQL
  - Basic commands & hands-on exercise
- Creating our target list for spectroscopy (groups 1 & 2)
  - Searching for RV variability in the catalogue of BHBs
  - Observational constraints
- Creating our target list for photometry (groups 3 & 4)
  - Group 3 -> Hunt for HW Vir systems in ATLAS and ZTF
  - Group 4 -> ADQL query the Gaia variable database
  - Observational constraints!

TOPCAT

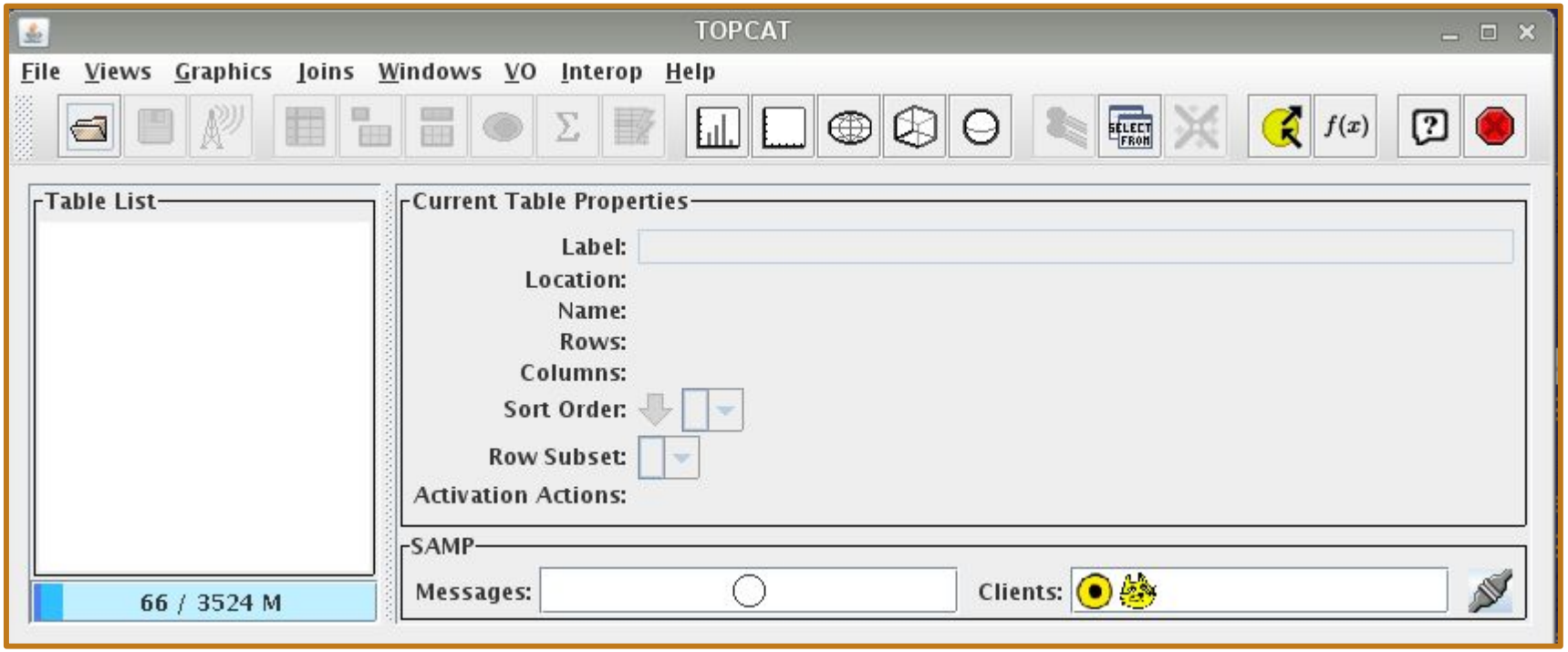
Tool for OPerations on Catalogues And Tables

Does what you want with tables

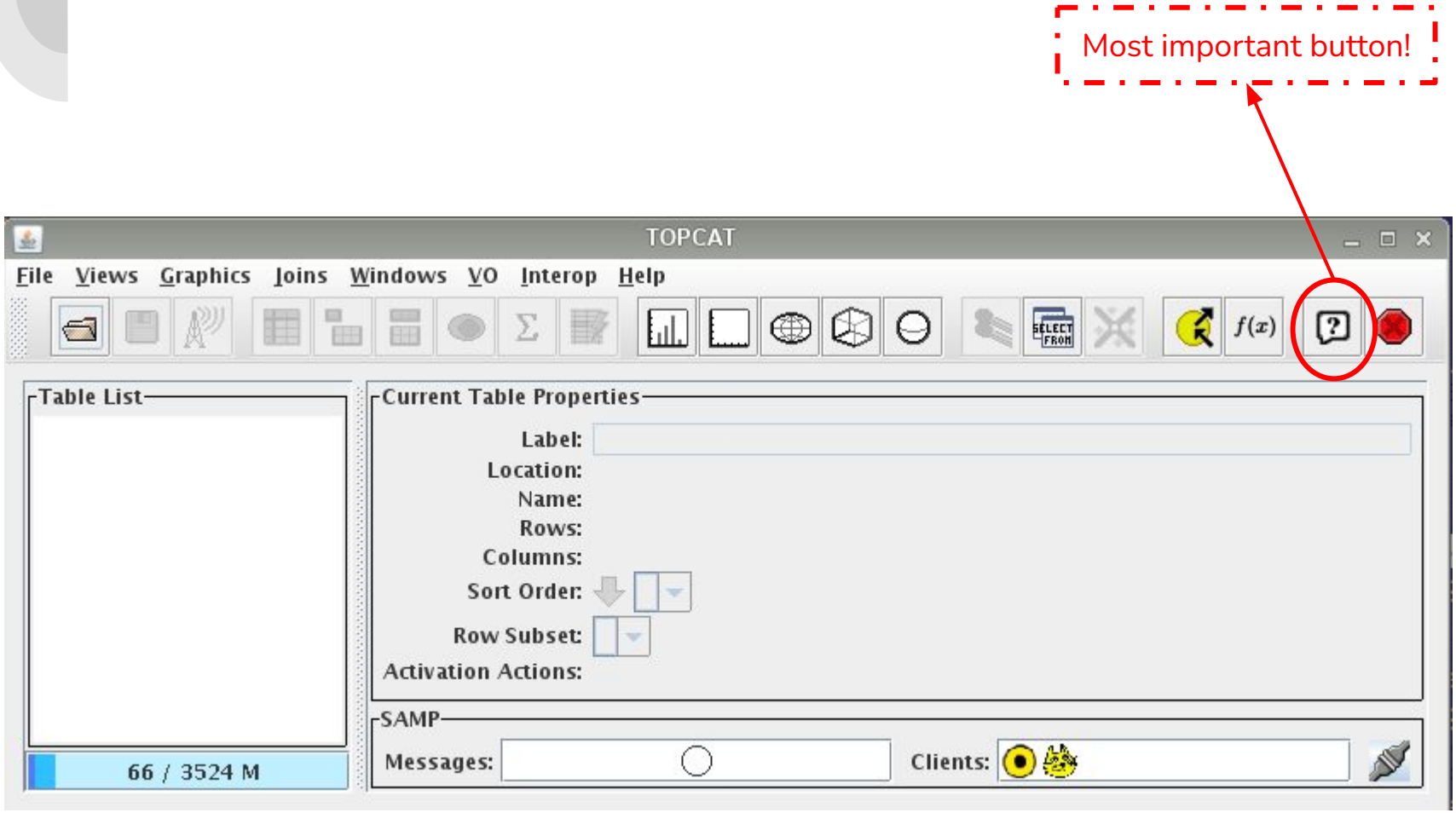
- Website: http://www.star.bristol.ac.uk/~mbt/topcat/
- Manual: http://www.starlink.ac.uk/topcat/sun253/
- Why TOPCAT?
  - Easy to use
  - Easy to learn
  - Easy to investigate data — good for exploratory analysis
  - Simple things obvious, complicated things documented
  - Easy to install and run
  - Fast
  - Copes with large data sets

- What can we do with TOPCAT?
  - Read/write tables in multiple formats
  - View/edit data
  - View/edit metadata
  - Plot data
  - Crossmatch — efficient and very flexible
  - (Simple) Calculations
  - Access Virtual Observatory (VO) services - hands-on exercise later
- What can’t we do?
  - Images, spectra (tables only!)
  - Cannot script it, you have to point and click!
  - Not ALL file formats
  - Do astronomy for you

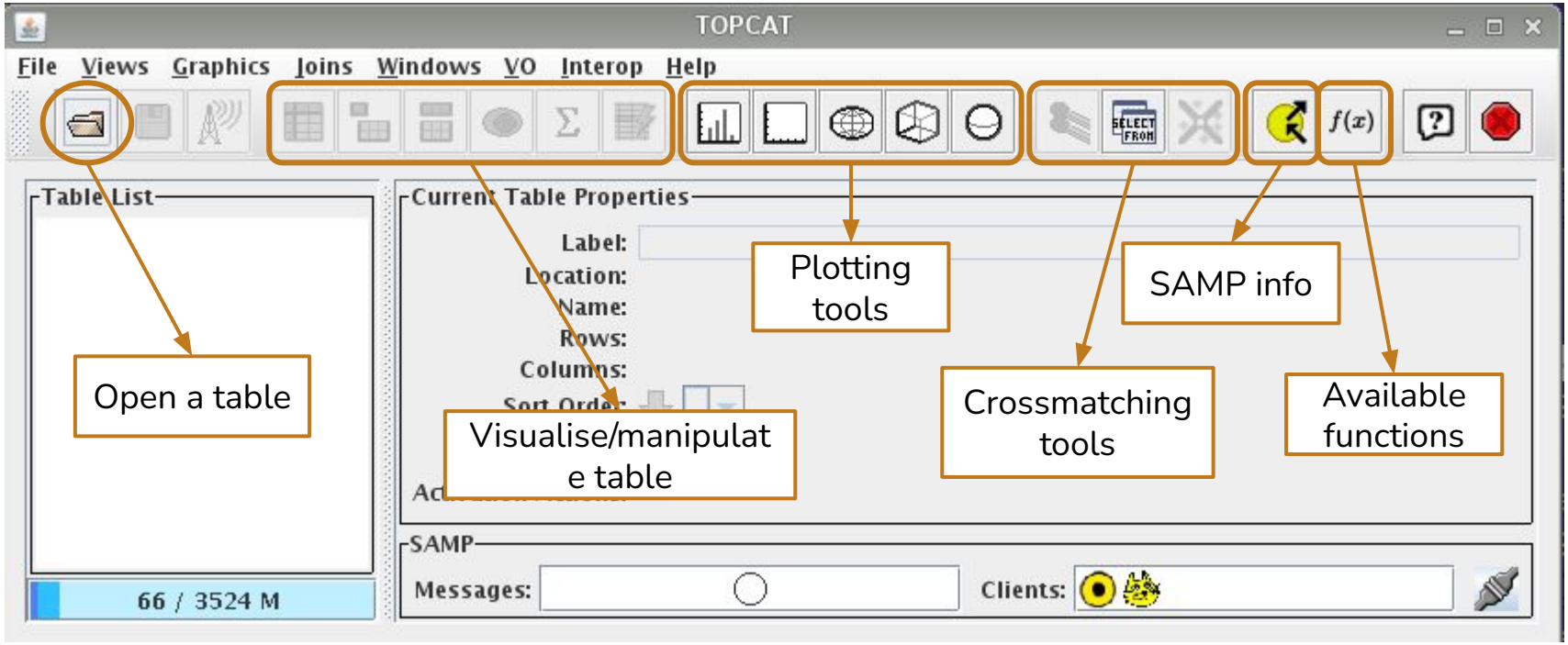
TOPCAT – start window



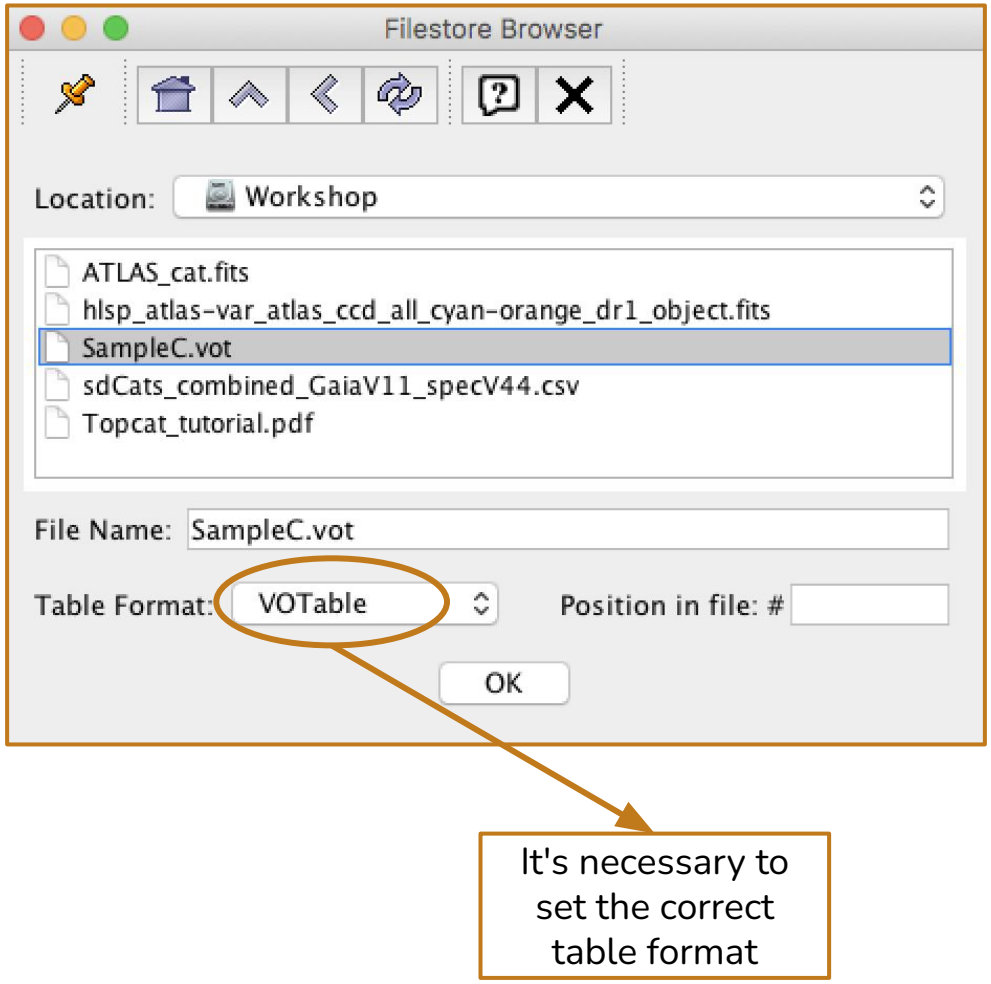
TOPCAT – start window



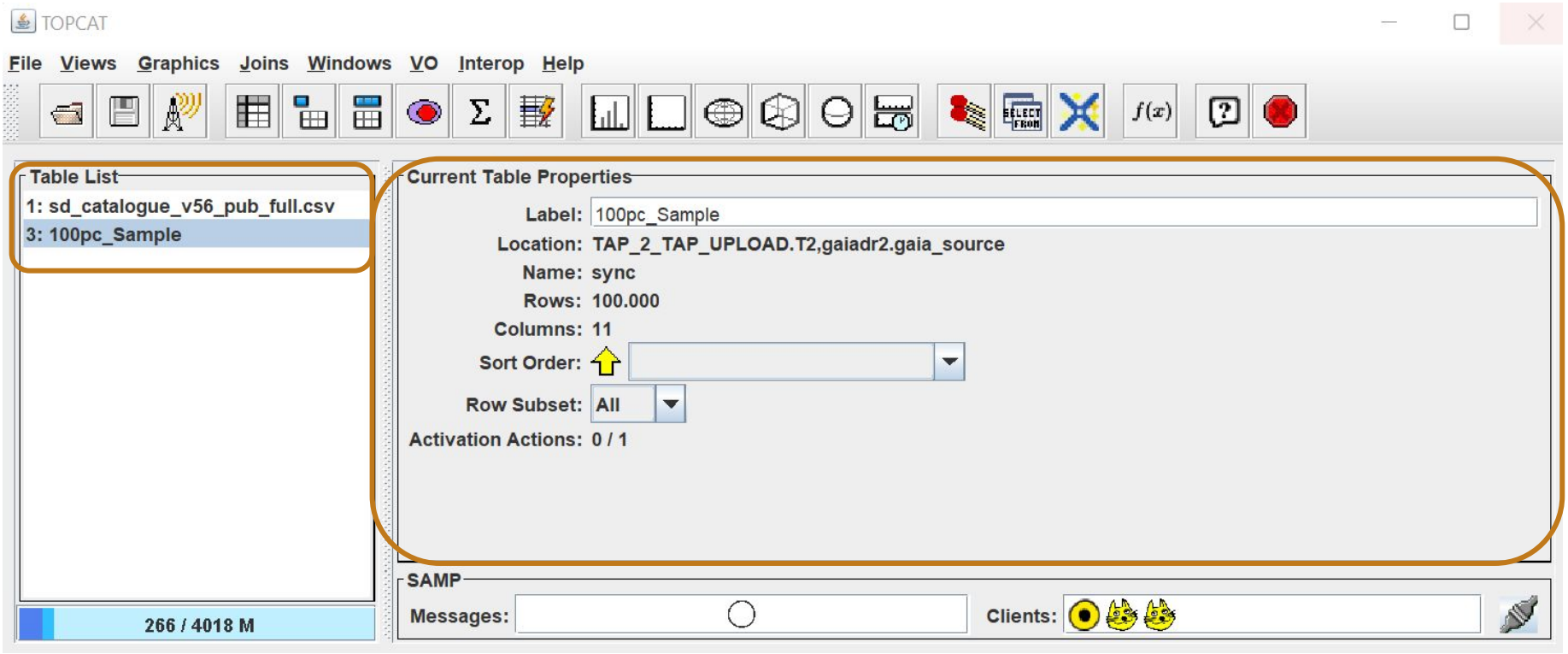
TOPCAT – start window



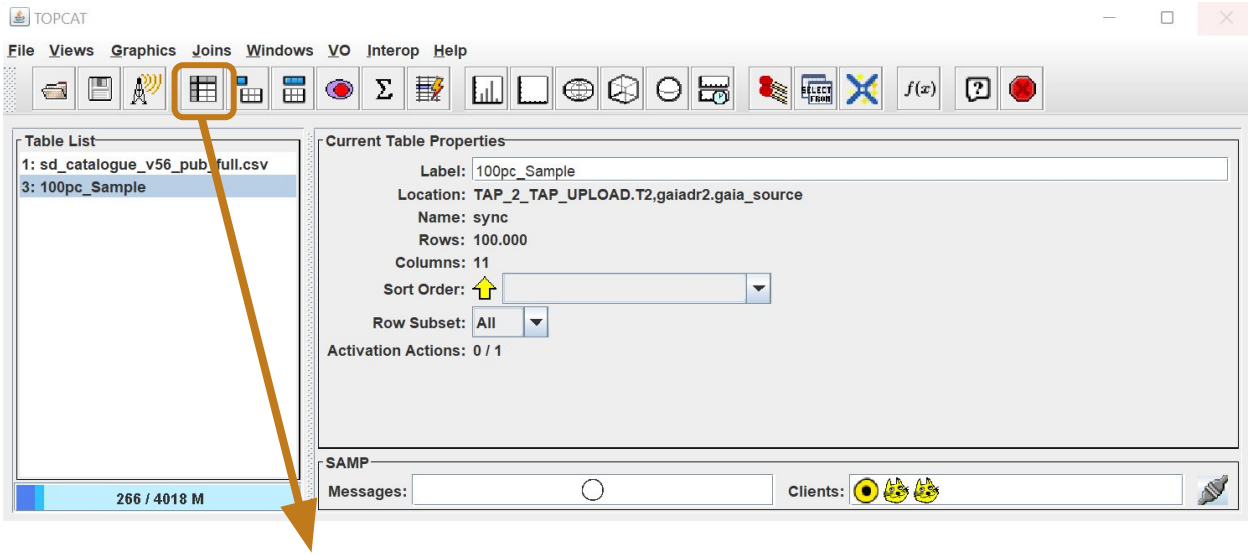
TOPCAT – open a table



TOPCAT – tables



TOPCAT – browse a table



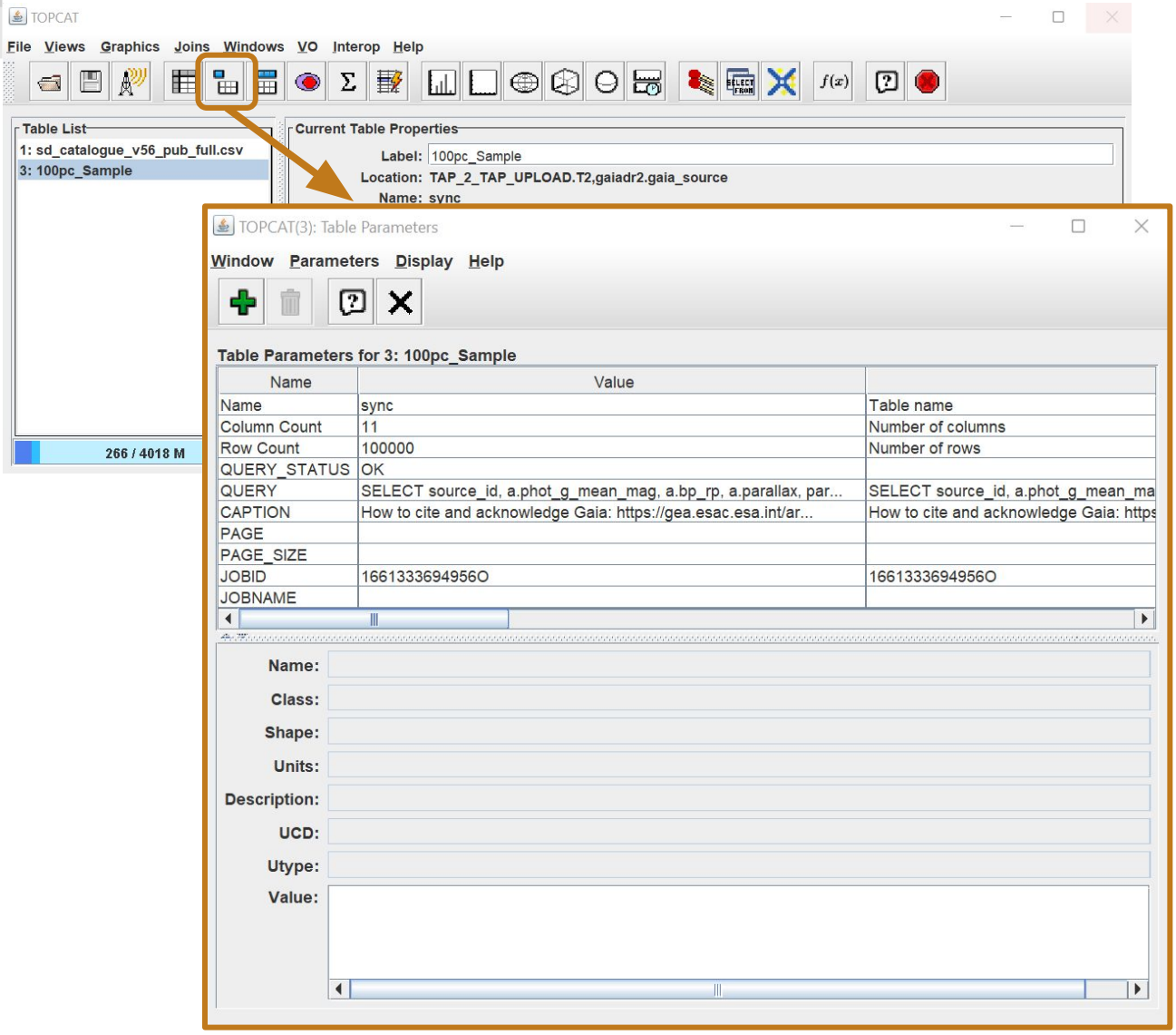
TOPCAT(2): Table Browser

Window Subsets Help

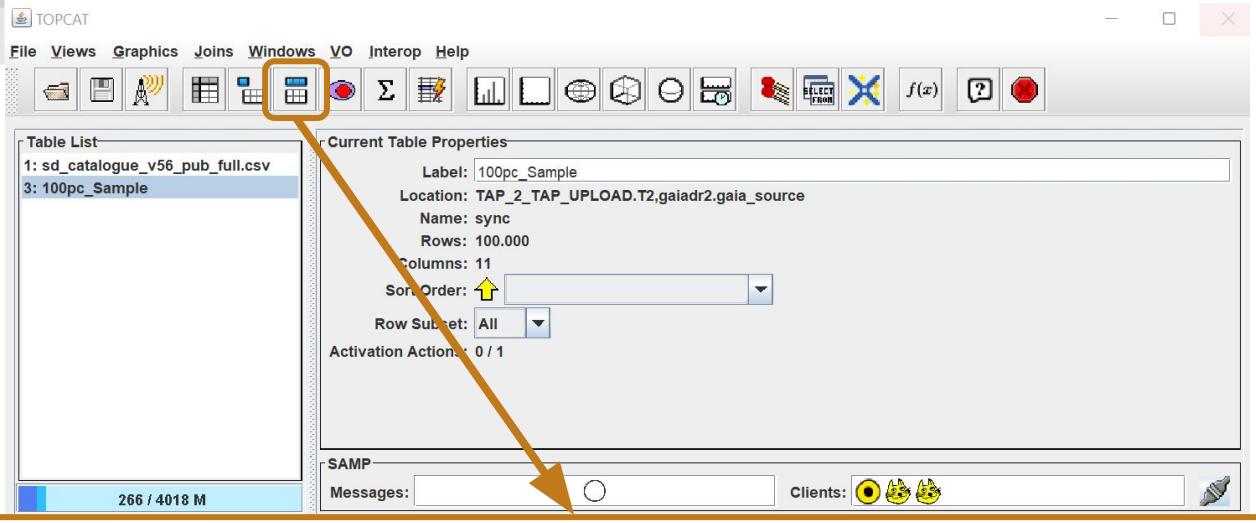
Table Browser for 2: SampleC.vot

	source_id	ra	dec	parallax	pmra	pmdec	phot_g_me...	phot_bp_m...	phot_rp_m...	bp_rp	teff_val	radius_val	radial_velocity
1	5256215443991096192	147.86761	-61.24324	14.45812	12.03787	-69.37827	15.9087	17.5931	14.6429	2.9502	4061.37		
2	5256330686560451584	151.56722	-60.97767	11.94937	-22.95639	71.97418	16.0123	17.8669	14.7033	3.16366	3719.83		
3	5256385455986316288	151.27972	-60.70641	12.54169	31.90794	80.67874	8.88798	9.19604	8.46277	0.733274	5956.	1.07332	-7.42609
4	5253416396637155072	153.5164	-61.03644	12.63063	-105.39727	-45.1931	15.137	16.5149	13.956	2.55884	3806.61		
5	5253387156502079744	152.8841	-61.23938	10.00575	-104.01756	50.83406	7.99488	8.28995	7.58982	0.700138	6150.75	1.89712	78.49139
6	5256366489408398336	150.23835	-60.96456	13.98831	-94.56353	119.33368	15.208	16.8438	13.9581	2.88572	3942.28		
7	5251098523021221376	144.83717	-61.32796	15.20927	-42.29215	19.4506	4.43662	4.46872	4.54535	-0.076632	9450.		
8	5257162462774509440	145.37644	-60.51155	19.26591	-186.61478	102.95347	11.6907	12.6408	10.7458	1.89492	4121.07	0.501863	15.92912
9	5258941648688757888	153.40699	-57.19364	13.69926	-19.40082	84.64139	15.0713	16.5366	13.8802	2.6564	3866.73		
10	5258898488554176384	151.62451	-57.25991	32.36492	48.46716	-62.36505	12.7897	14.1319	11.6591	2.47282	3764.82		
11	5259661897522690688	151.14651	-57.02871	11.71382	-114.0676	60.93288	14.3115	15.8704	13.0872	2.78321	3683.46		
12	5258429379357599232	152.07547	-58.19864	14.42705	-4.00739	-13.83841	6.47879	6.80914	6.04574	0.763399	6011.5	2.77469	-10.38242
13	5255092876977182976	153.85899	-59.60026	15.49247	-59.49346	11.34791	16.459	18.0891	15.1948	2.89426	4120.11		

TOPCAT – table metadata



TOPCAT – column metadata



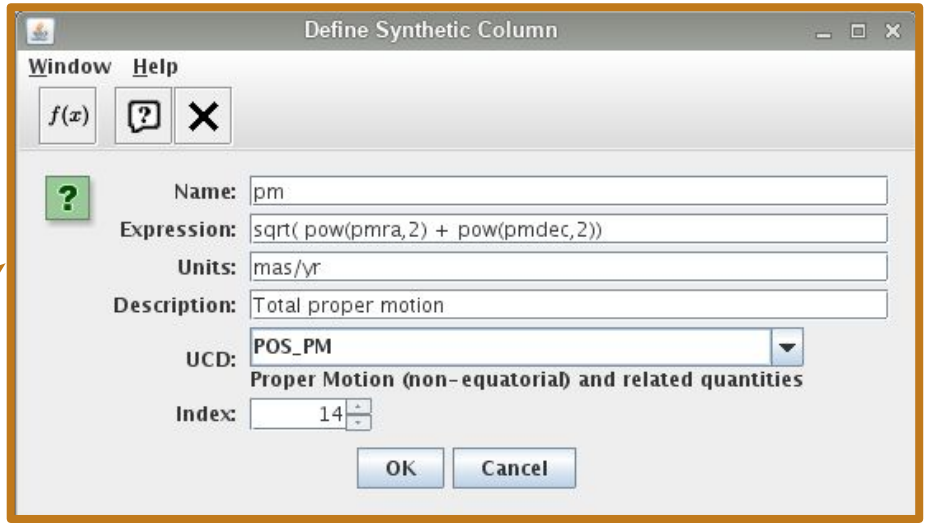
TOPCAT(3): Table Columns

Window Columns Display Help

Table Columns for 3: 100pc_Sample

△	Index	Visible	Name	$ID	Class	Units	Description	UCD	Datatype
0		☐	Index	$0	Long		Table row index		
1	1	☑	source_id	$1	Long				long
2	2	☑	phot_g_mean_mag	$2	Float	mag	G-band mean magnitude	phot.mag;stat.mean;em.opt	float
3	3	☑	bp_rp	$3	Float	mag	BP - RP colour	phot.color	float
4	4	☑	parallax	$4	Double	mas	Parallax	pos.parallax	double
5	5	☑	parallax_over_error	$5	Float		Parallax divided by its error	arith.ratio	float
6	6	☑	phot_bp_mean_flux_over_error	$6	Float		Integrated BP mean flux divided by its error	arith.ratio	float
7	7	☑	phot_rp_mean_flux_over_error	$7	Float		Integrated RP mean flux divided by its error	arith.ratio	float
8	8	☑	phot_bp_rp_excess_factor	$8	Float		BP/RP excess factor		float
9	9	☑	astrometric_chi2_al	$9	Float		AL chi-square value	stat.fit.chi2	float
10	10	☑	astrometric_n_good_obs_al	$10	Integer		Number of good observations AL	meta.number	int
11	11	☑	astrometric_excess_noise	$11	Double	mas	Excess noise of the source	stat.value	double

TOPCAT – create new column



Column names become variables

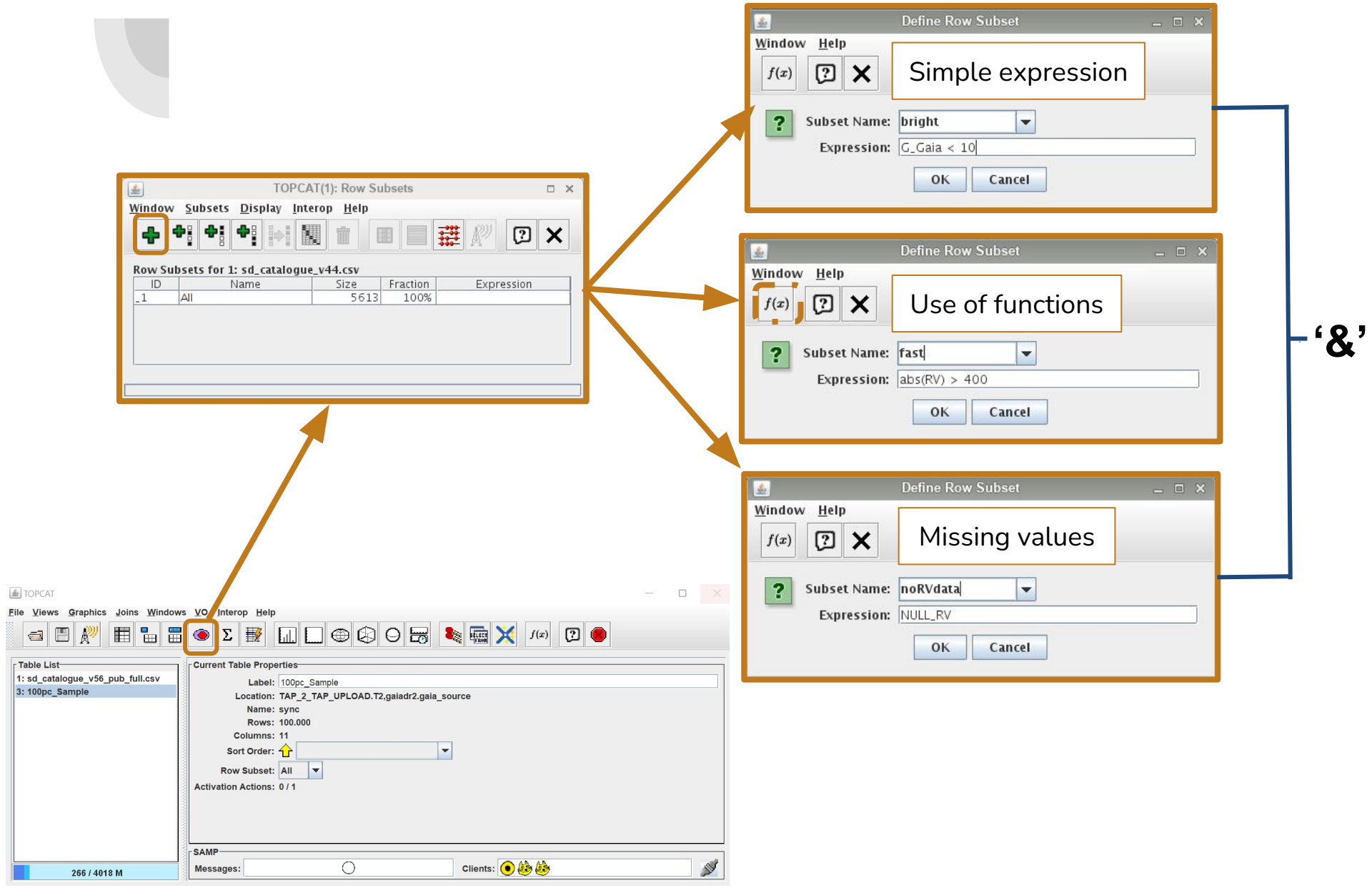
TOPCAT(2): Table Columns

Window Columns Display Help

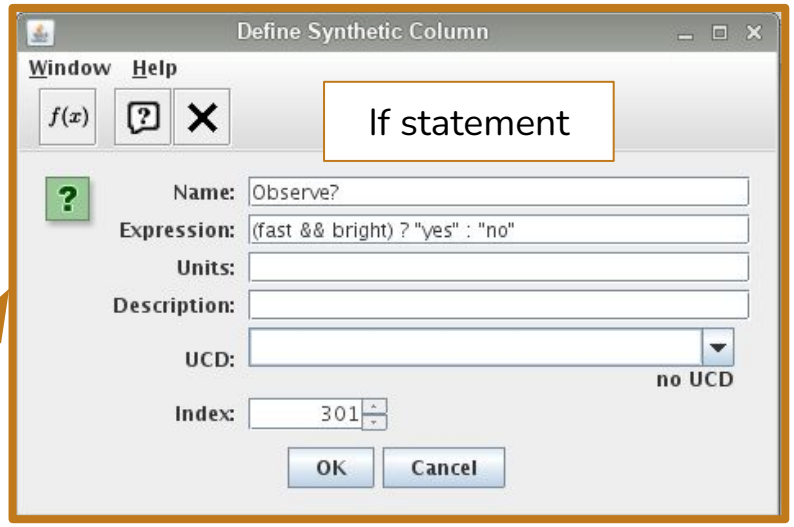
Table Columns for 2: SampleC.vot

Δ	Index	Visible	Name	$ID	Class	Units	Description	UCD	Datatype	VOTable ID
0		☐	Index	$0	Long		Table row index			
1	1	☑	source_id	$1	Long				long	col_0
2	2	☑	ra	$2	Double	deg	Right ascension	pos.eq.ra;meta.main	double	col_1
3	3	☑	dec	$3	Double	deg	Declination	pos.eq.dec;meta.main	double	col_2
4	4	☑	parallax	$4	Double	mas	Parallax	pos.parallax	double	col_3
5	5	☑	pmra	$5	Double	mas/yr	Proper motion in right ascension direction	pos.pm;pos.eq.ra	double	col_4
6	6	☑	pmdec	$6	Double	mas/yr	Proper motion in declination direction	pos.pm;pos.eq.dec	double	col_5
7	7	☑	phot_g_mean_mag	$7	Float	mag	G-band mean magnitude	phot.mag;stat.mean;em.opt	float	col_6
8	8	☑	phot_bp_mean_mag	$8	Float	mag	Integrated BP mean magnitude	phot.mag;stat.mean	float	col_7
9	9	☑	phot_rp_mean_mag	$9	Float	mag	Integrated RP mean magnitude	phot.mag;stat.mean	float	col_8
10	10	☑	bp_rp	$10	Float	mag	BP – RP colour	phot.color	float	col_9
11	11	☑	teff_val	$11	Float	K	Stellar effective temperature	phys.temperature.effective	float	col_10
12	12	☑	radius_val	$12	Float	solRad	Stellar radius	phys.size.radius	float	col_11
13	13	☑	radial_velocity	$13	Double	km/s	Radial velocity	spect.dopplerVeloc.opt	double	col_12

TOPCAT – create & combine subsets



TOPCAT – create column based on subset



Expression language:

- Standard arithmetic/functions: (+, -, /, *), (abs, max, round, sin, cos, pow...)
- Conditional: (**a?b:c**)

Examples:

➔ BP - RP

➔ skyDistanceDegrees(ra, dec, 25, 52.1) < 5 [degrees]

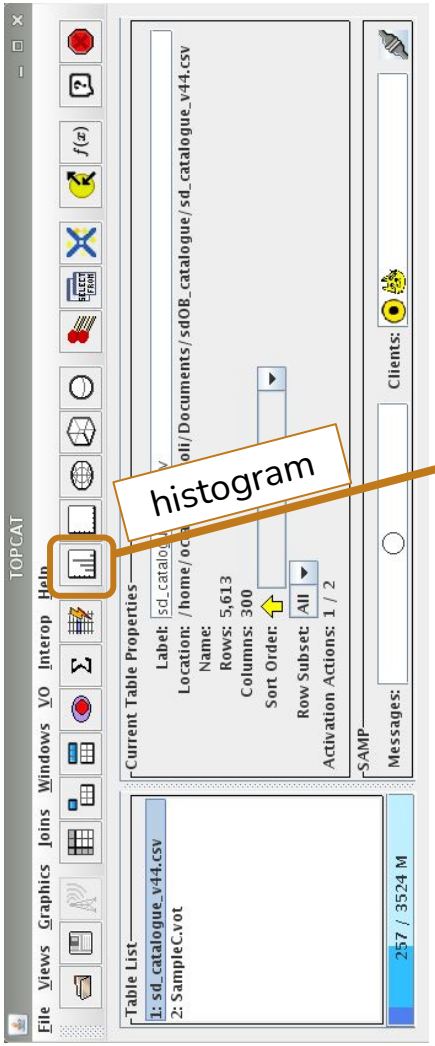
TOPCAT(2): Table Columns

Window Columns Display Help

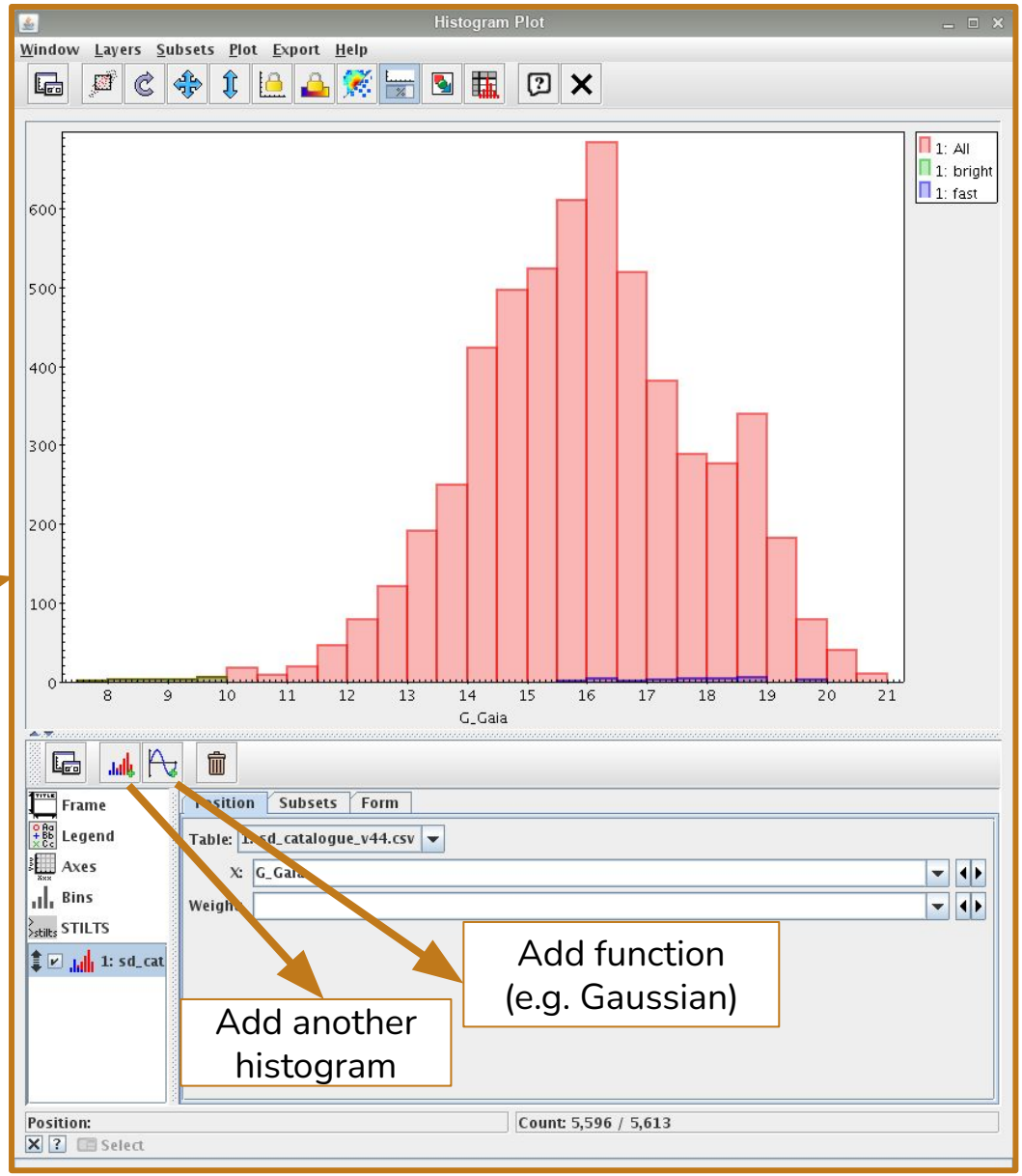
Table Columns for 2: SampleC.vot

Δ	Index	Visible	Name	$ID	Class	Units	Description	UCD	Datatype	VOTable ID
0		☐	Index	$0	Long		Table row index			
1	1	☑	source_id	$1	Long				long	col_0
2	2	☑	ra	$2	Double	deg	Right ascension	pos.eq.ra;meta.main	double	col_1
3	3	☑	dec	$3	Double	deg	Declination	pos.eq.dec;meta.main	double	col_2
4	4	☑	parallax	$4	Double	mas	Parallax	pos.parallax	double	col_3
5	5	☑	pmra	$5	Double	mas/yr	Proper motion in right ascension direction	pos.pm;pos.eq.ra	double	col_4
6	6	☑	pmdec	$6	Double	mas/yr	Proper motion in declination direction	pos.pm;pos.eq.dec	double	col_5
7	7	☑	phot_g_mean_mag	$7	Float	mag	G-band mean magnitude	phot.mag;stat.mean;em.opt	float	col_6
8	8	☑	phot_bp_mean_mag	$8	Float	mag	Integrated BP mean magnitude	phot.mag;stat.mean	float	col_7
9	9	☑	phot_rp_mean_mag	$9	Float	mag	Integrated RP mean magnitude	phot.mag;stat.mean	float	col_8
10	10	☑	bp_rp	$10	Float	mag	BP – RP colour	phot.color	float	col_9
11	11	☑	teff_val	$11	Float	K	Stellar effective temperature	phys.temperature.effective	float	col_10
12	12	☑	radius_val	$12	Float	solRad	Stellar radius	phys.size.radius	float	col_11
13	13	☑	radial_velocity	$13	Double	km/s	Radial velocity	spect.dopplerVeloc.opt	double	col_12

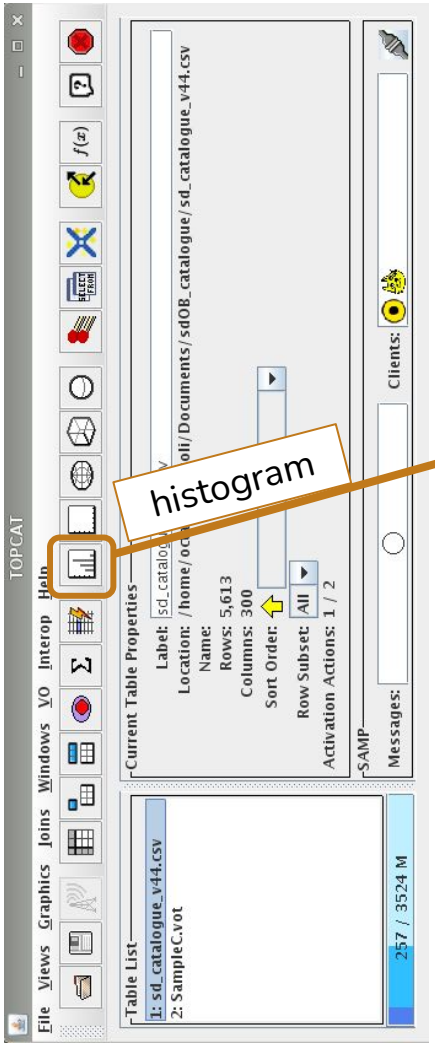
TOPCAT – Visualisation tools



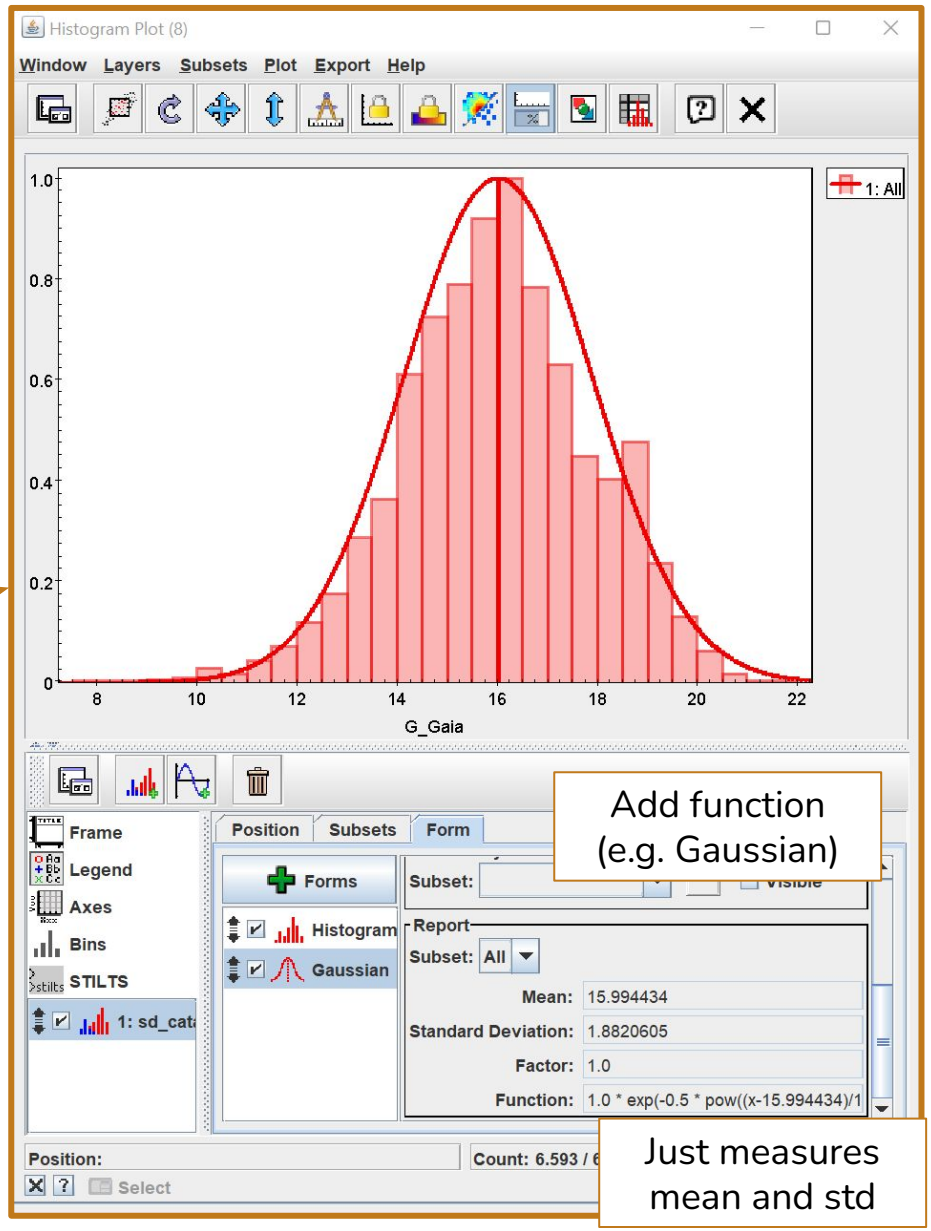
histogram



TOPCAT – Visualisation tools



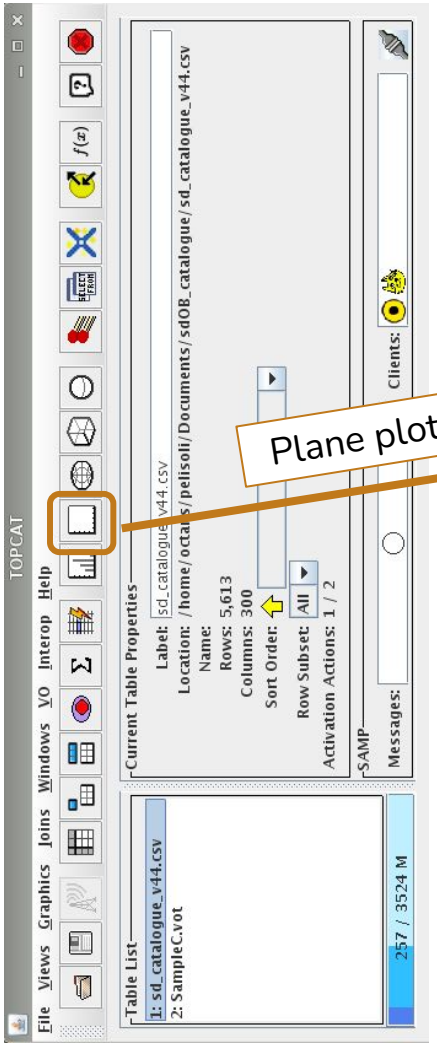
histogram



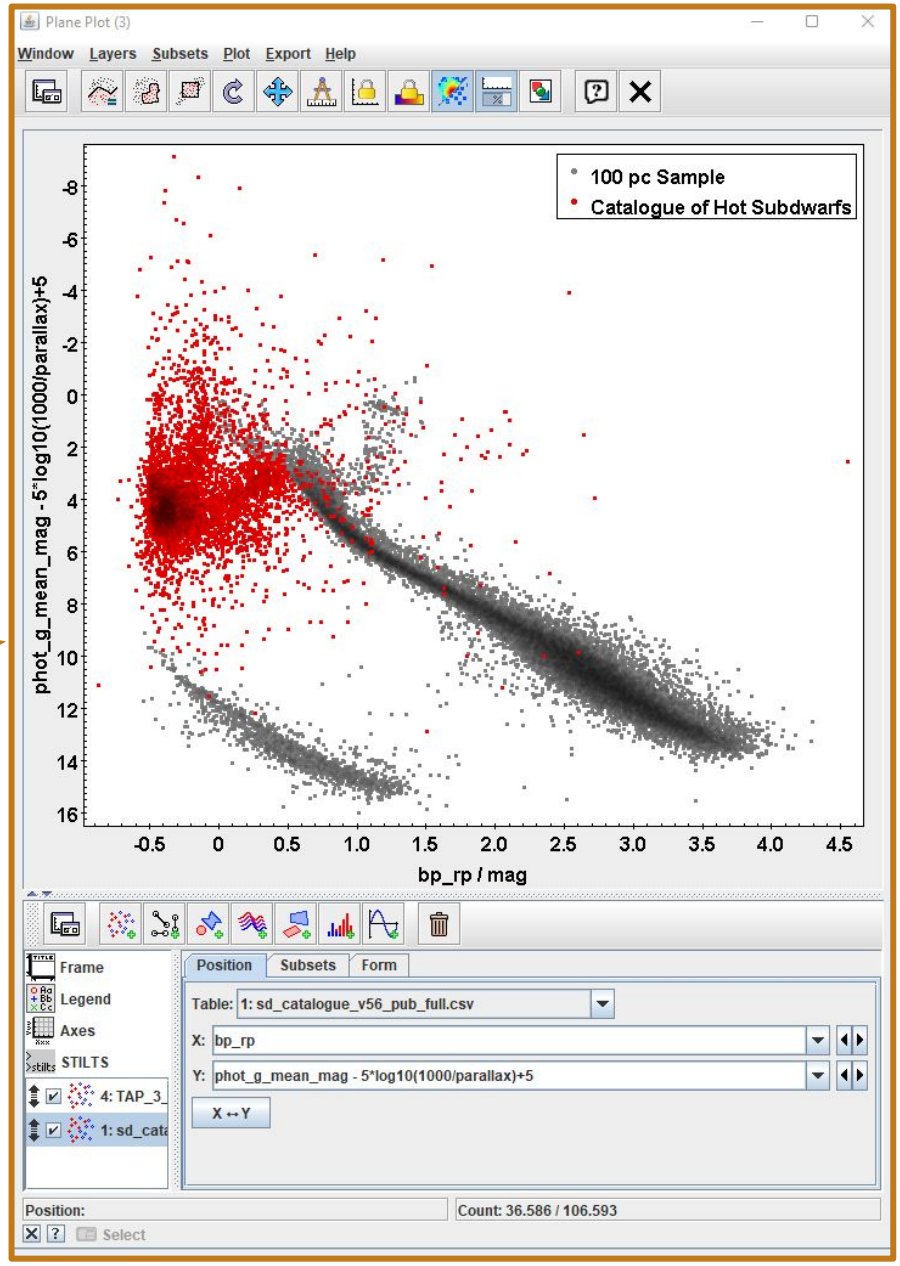
Add function (e.g. Gaussian)

Just measures mean and std

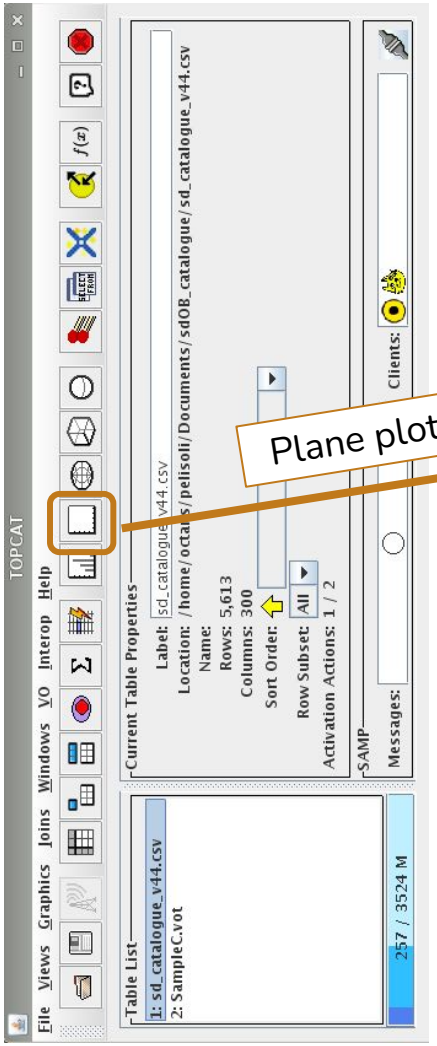
TOPCAT – Visualisation tools



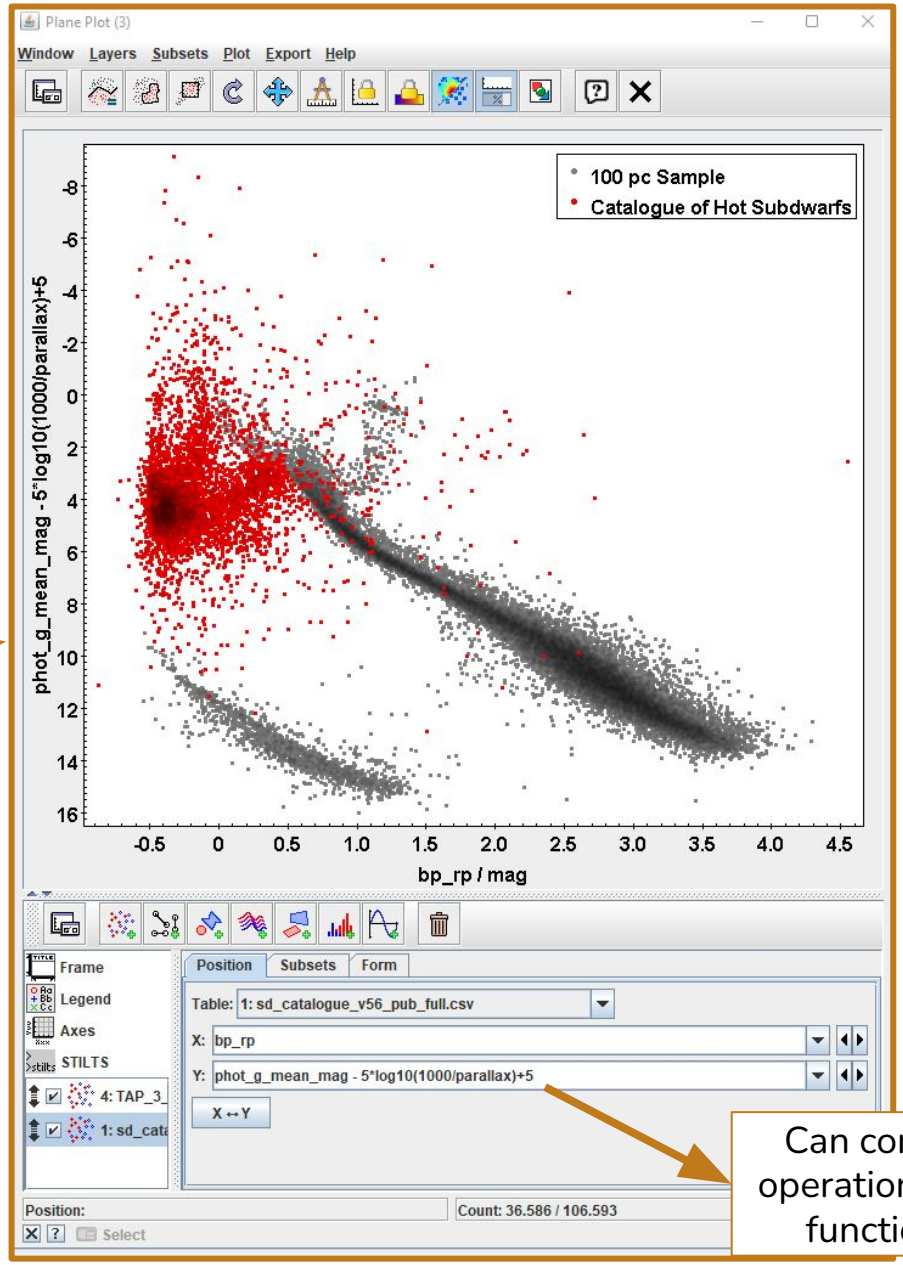
Plane plot



TOPCAT – Visualisation tools

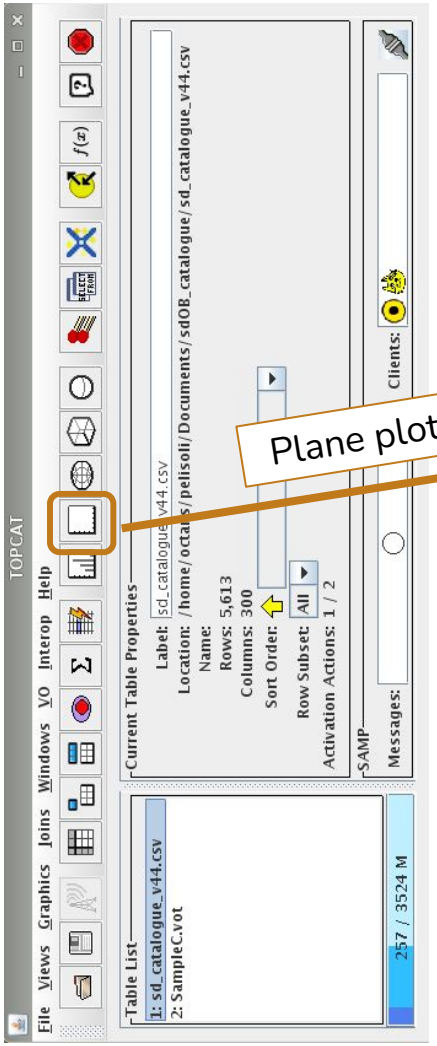


Plane plot

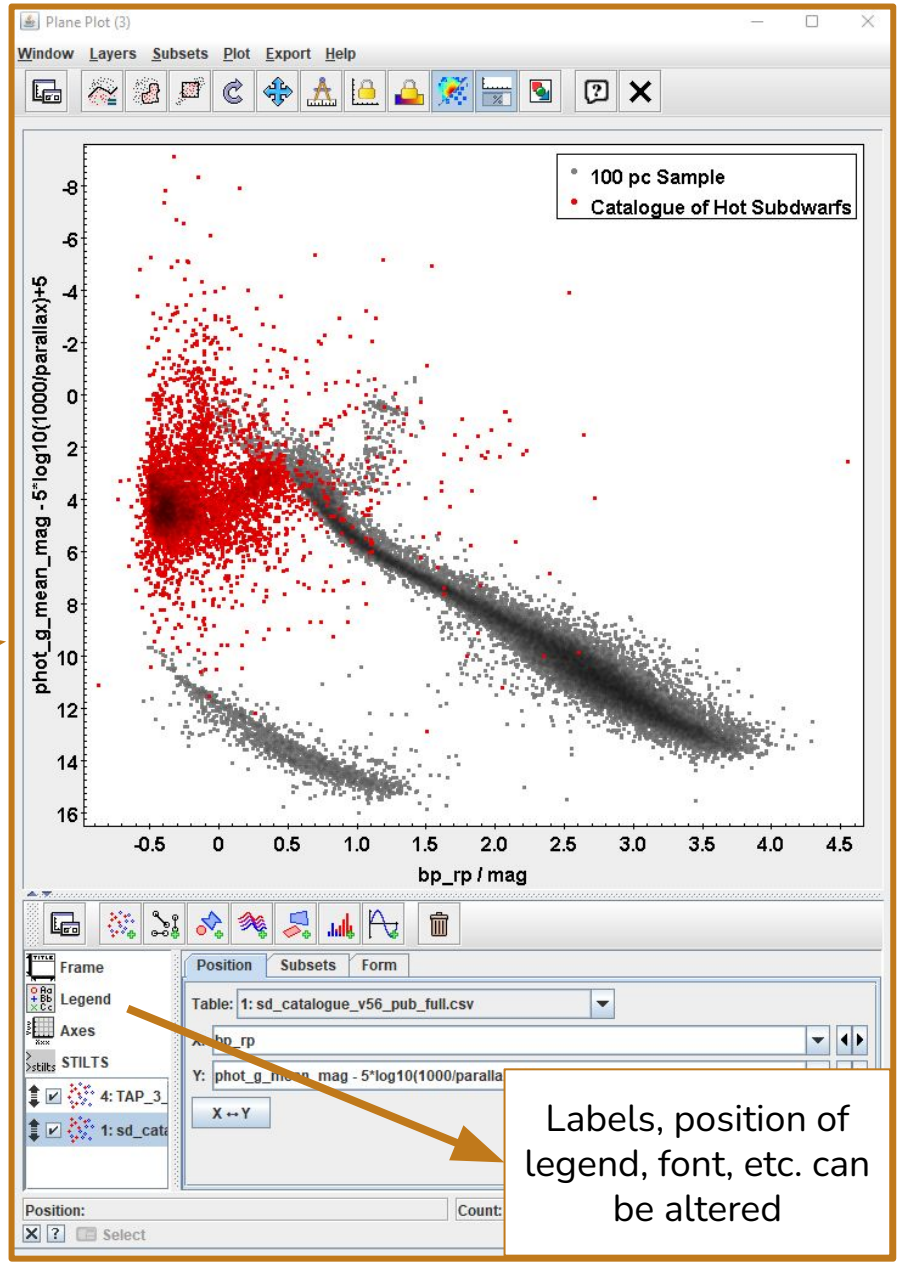


Can contain operations and functions

TOPCAT – Visualisation tools

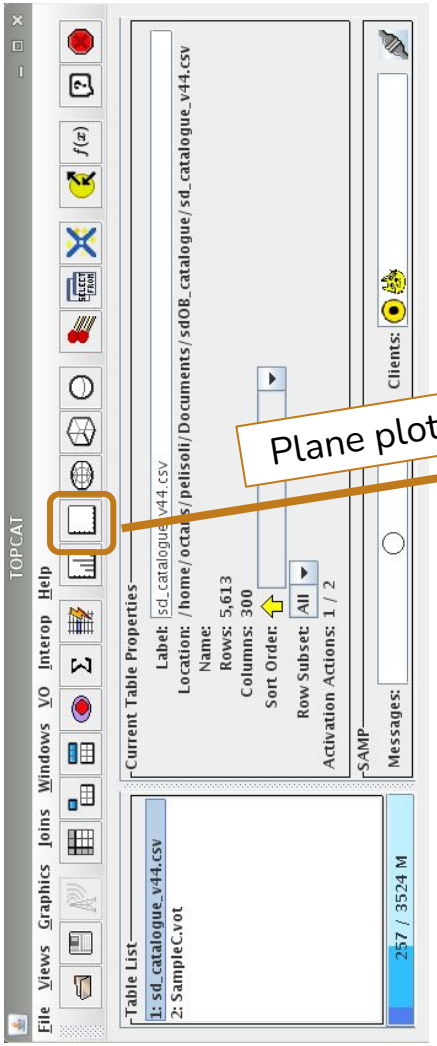


Plane plot

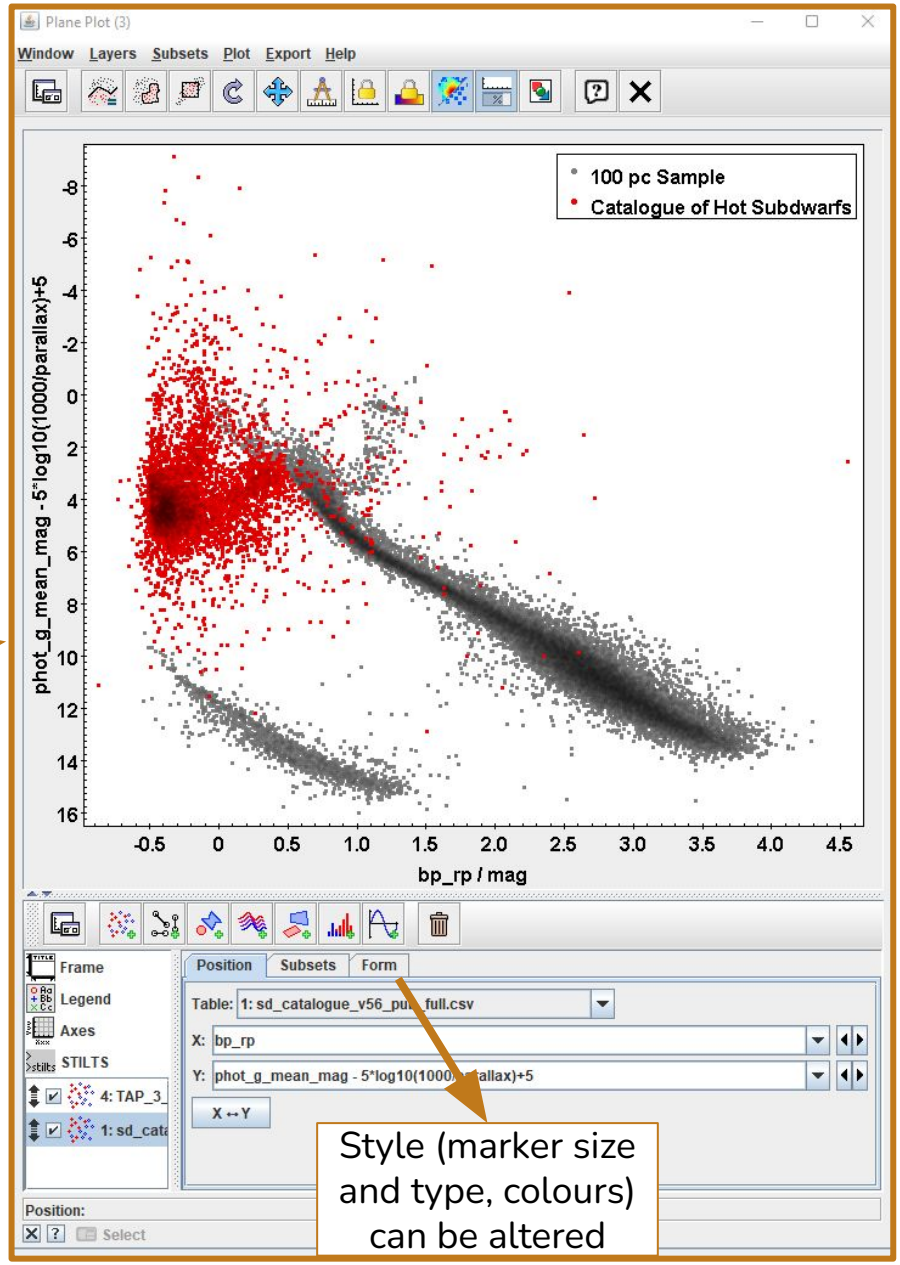


Labels, position of legend, font, etc. can be altered

TOPCAT – Visualisation tools

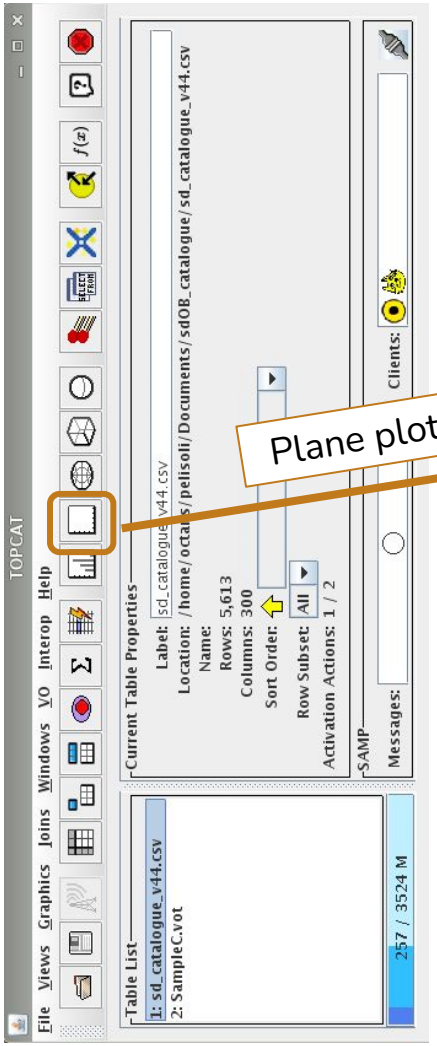


Plane plot

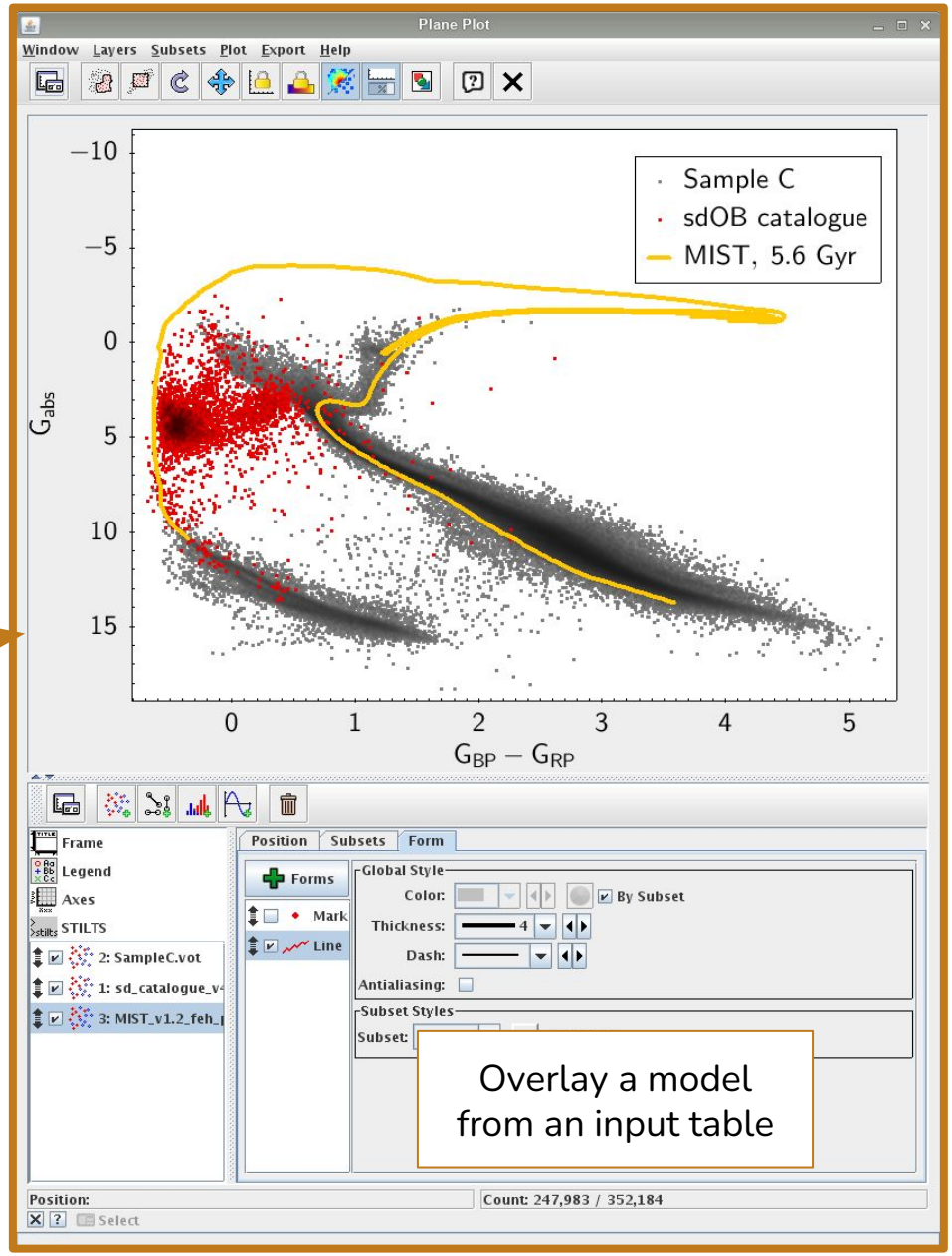


Style (marker size and type, colours) can be altered

TOPCAT – Visualisation tools

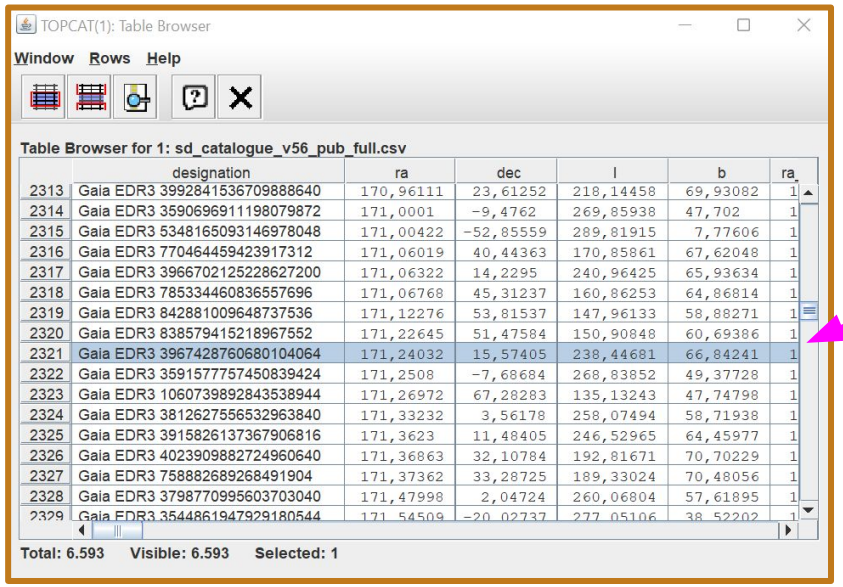


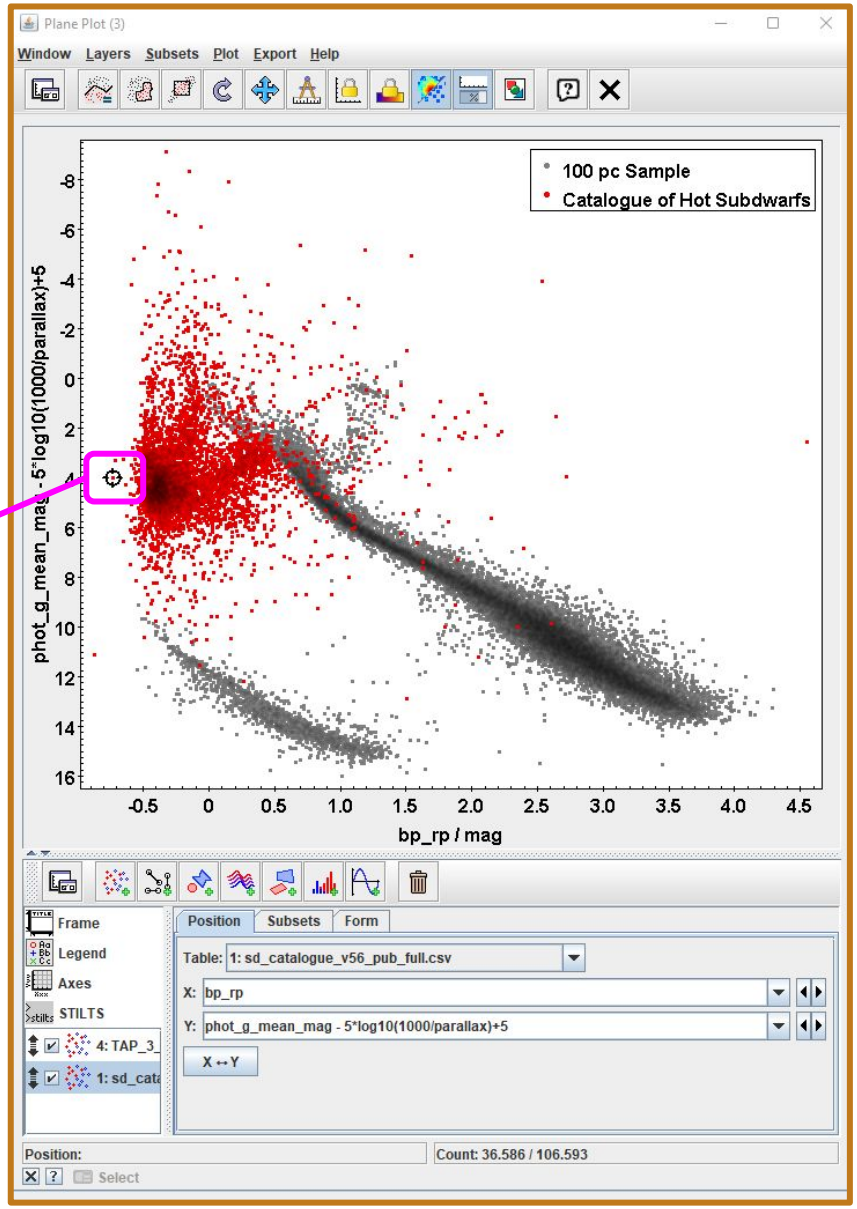
Plane plot



Overlay a model from an input table

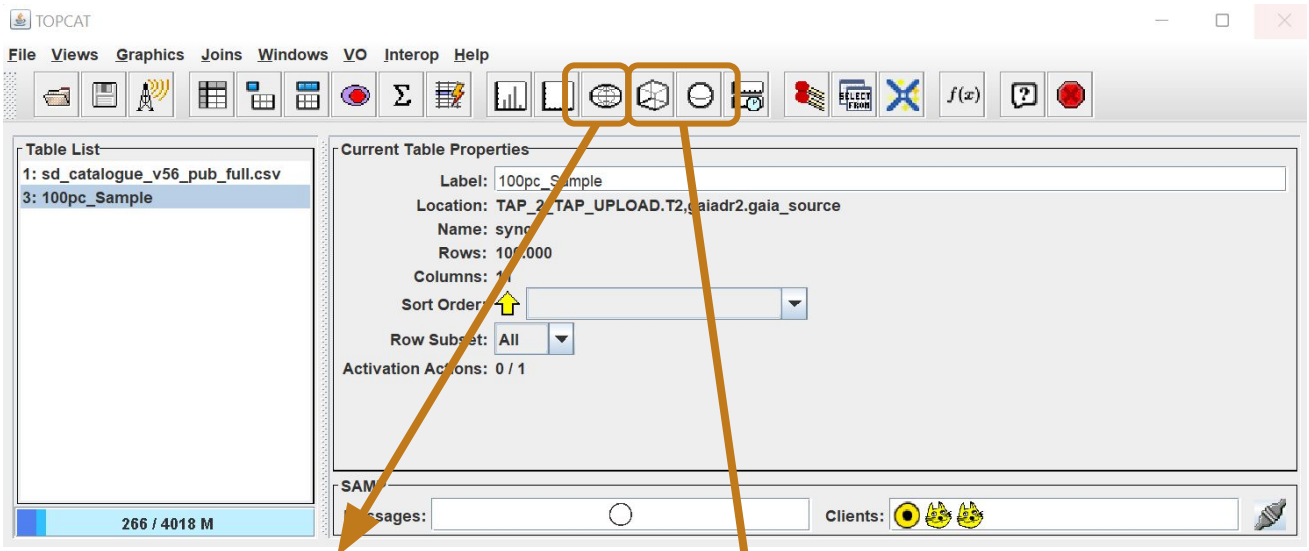
TOPCAT – Visualisation tools



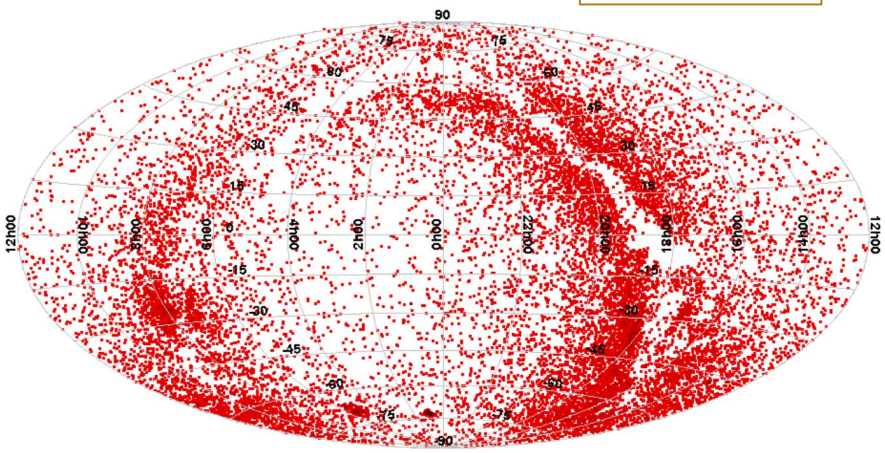


Identify single points
-> Highlights those points in all tables and plots!

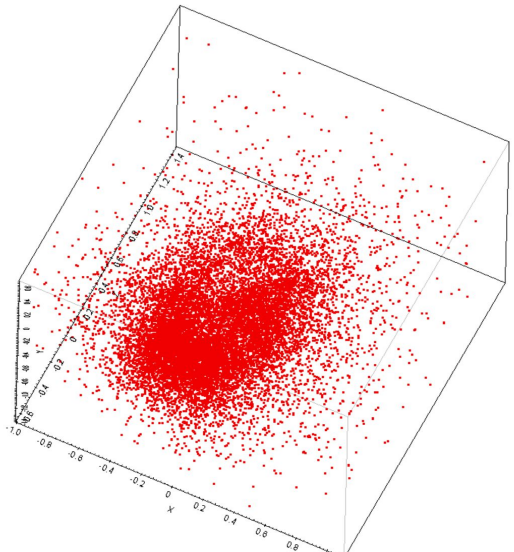
TOPCAT – Visualisation tools



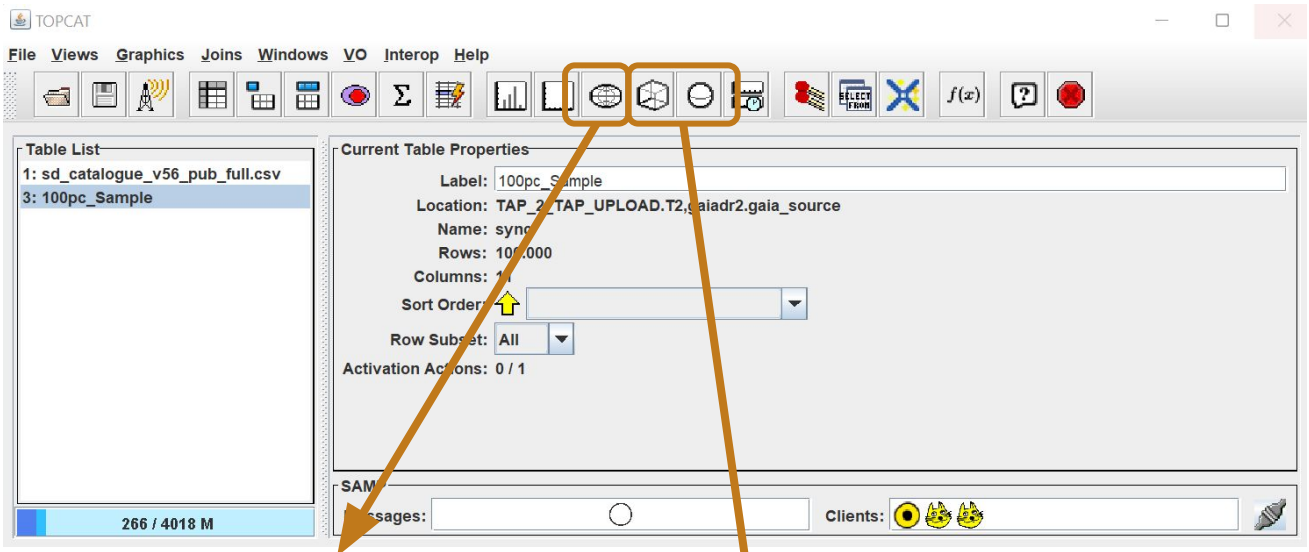
Sky plotting



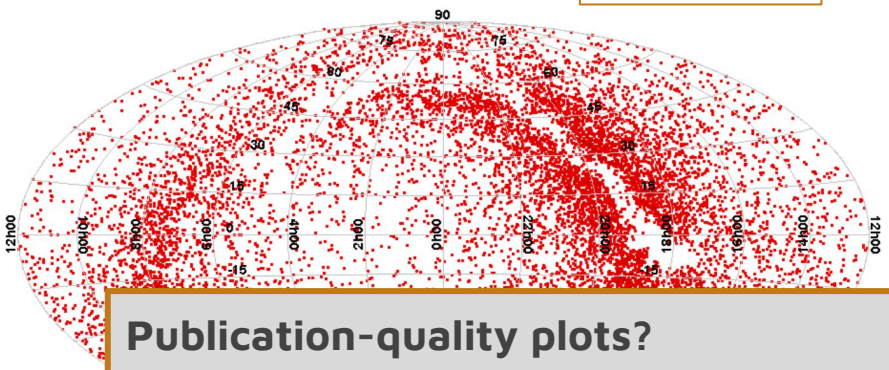
3D plotting



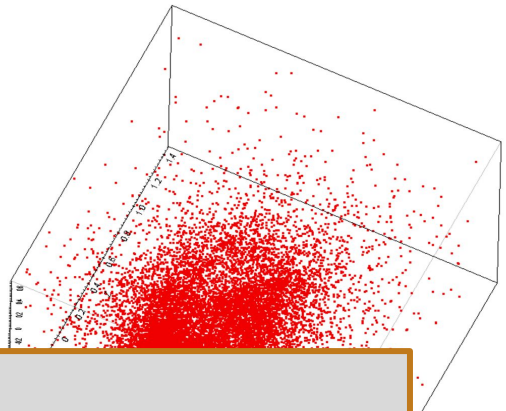
TOPCAT – Visualisation tools



Sky plotting

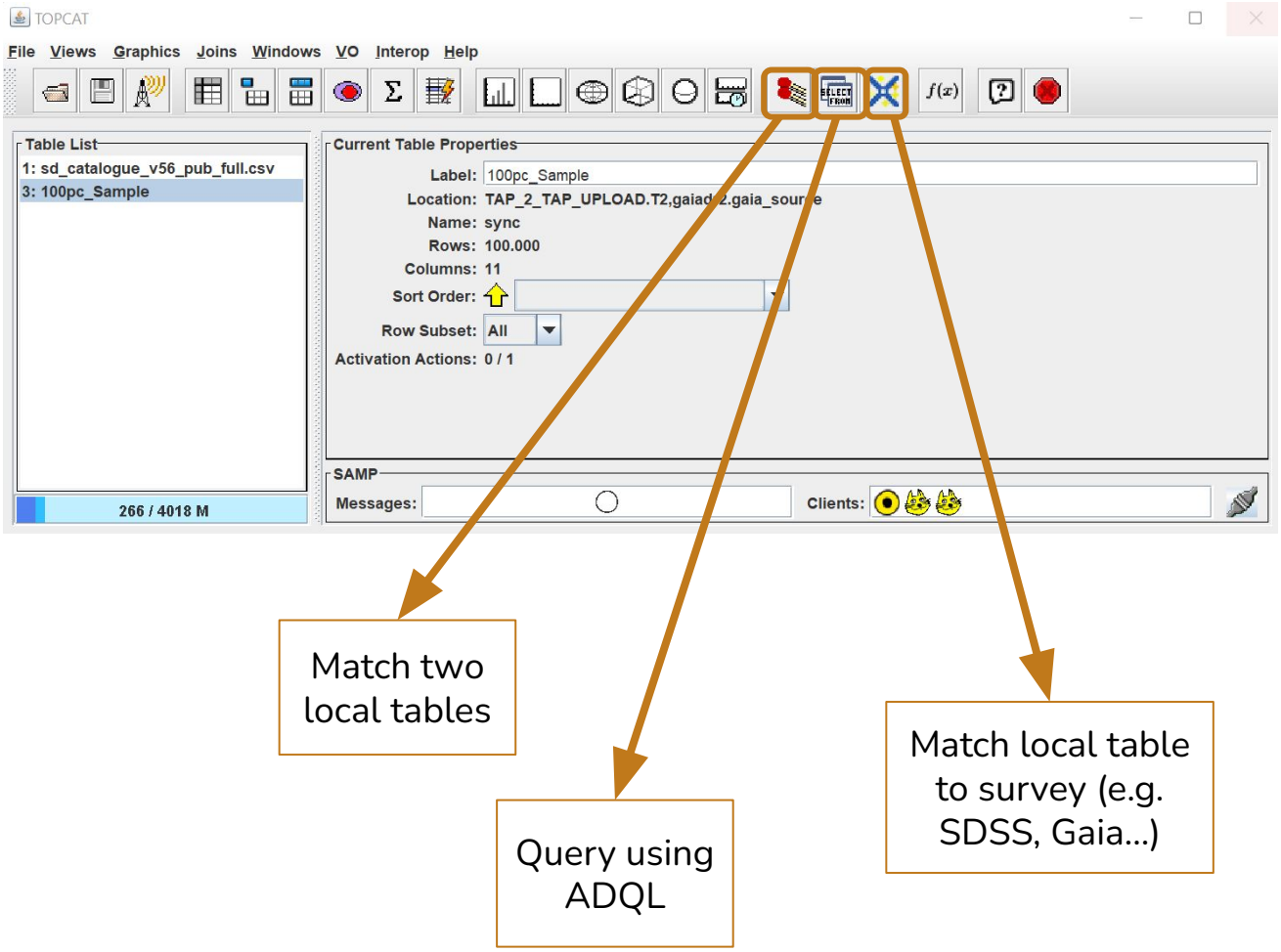


3D plotting

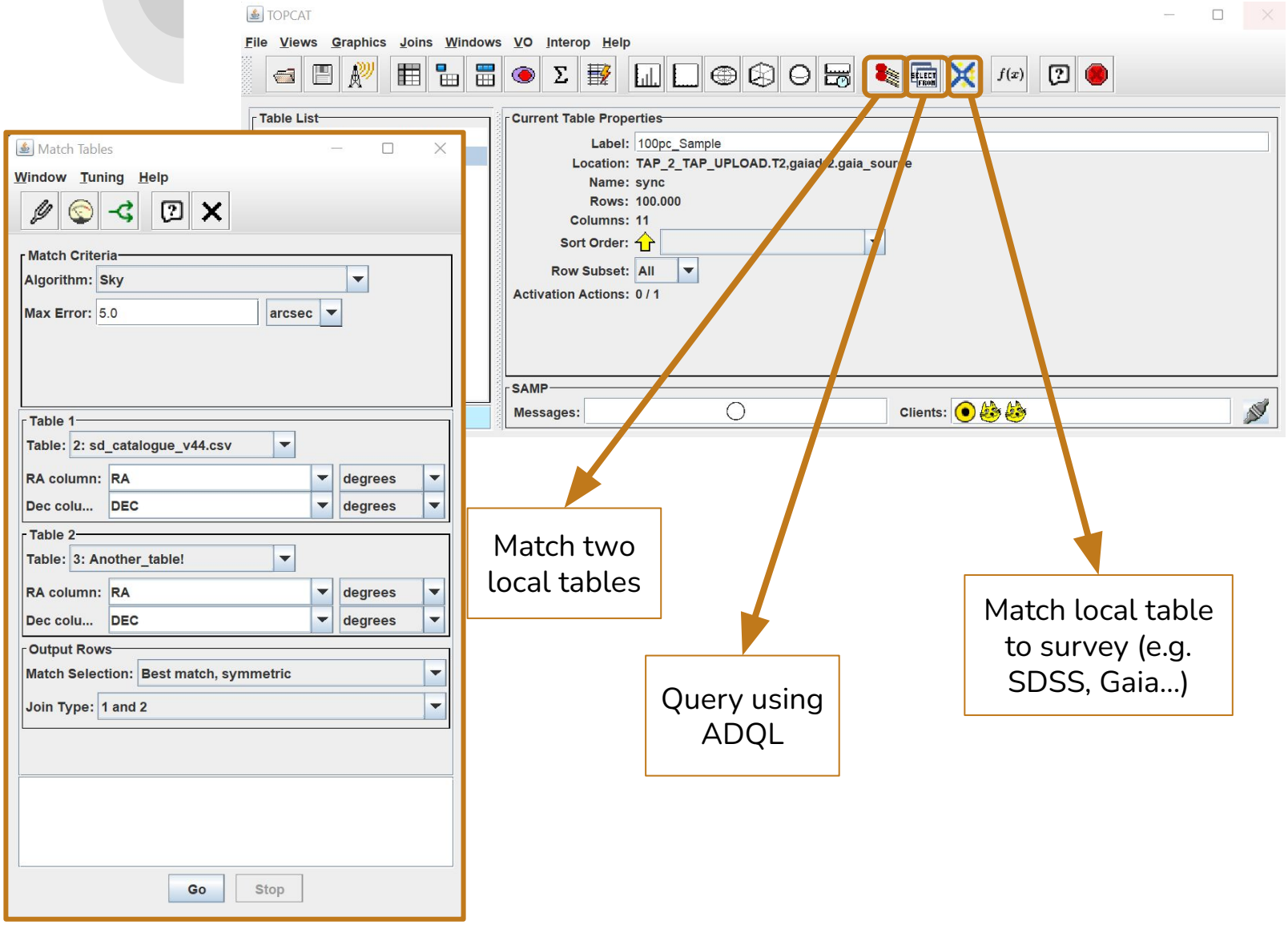


Publication-quality plots?
Sure! -> Export to PDF, EPS, PNG, GIF, SVG…
But not quite as good as Matplotlib etc. But best for millions of data points!

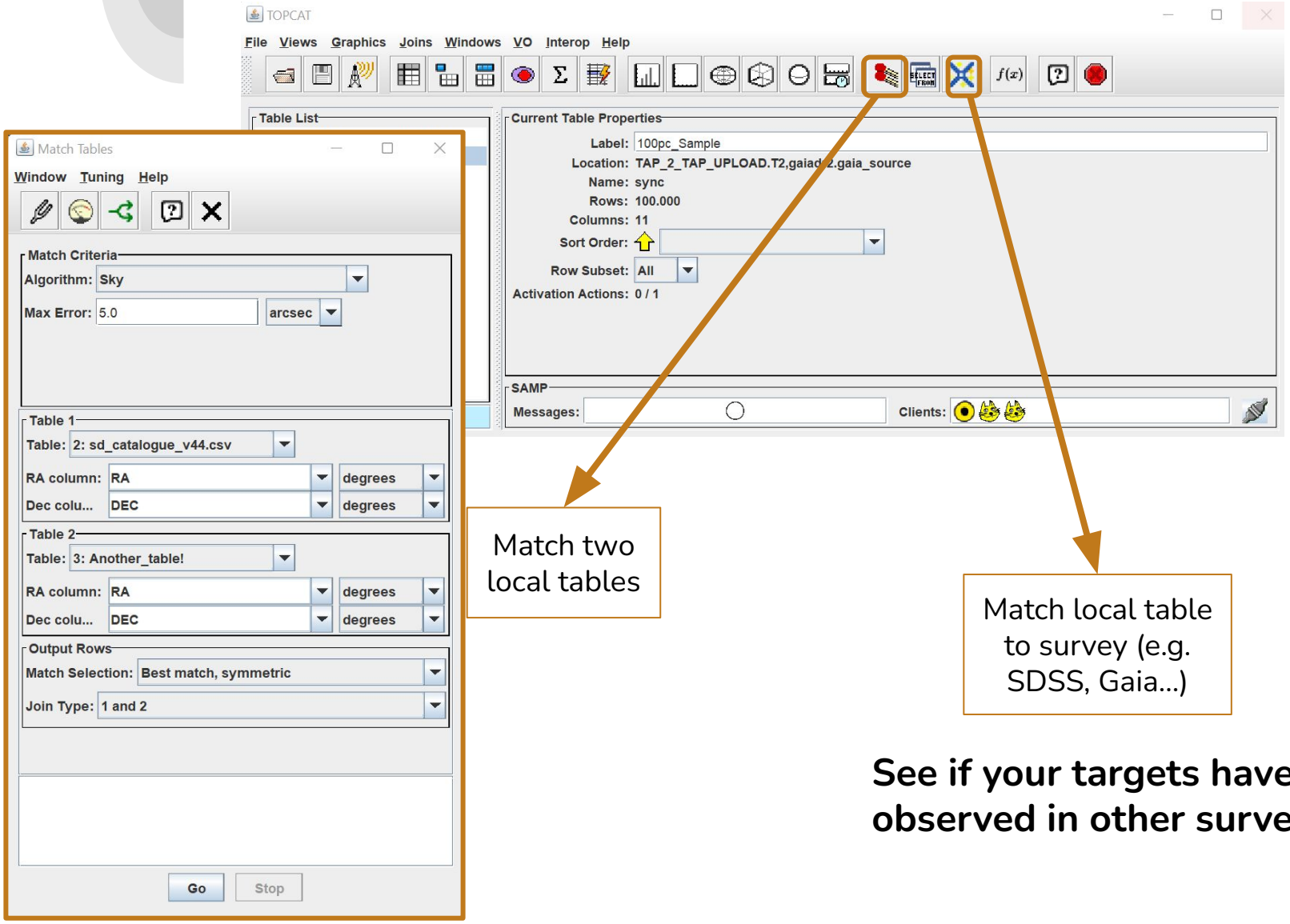
TOPCAT – Crossmatching



TOPCAT – Crossmatching



TOPCAT – Crossmatching

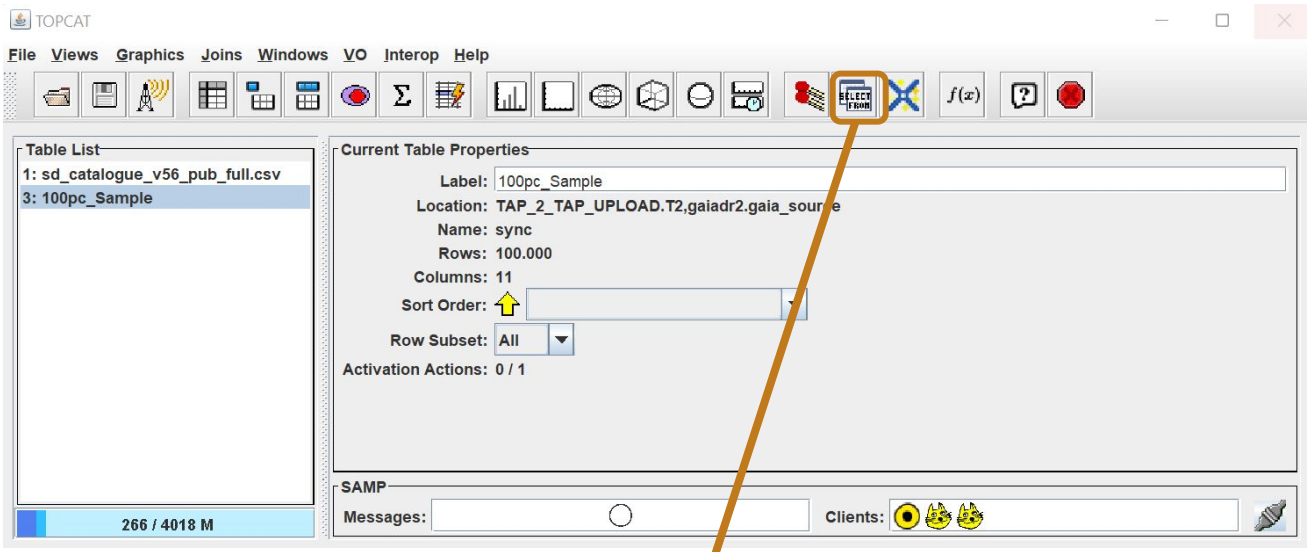


Match two local tables

Match local table to survey (e.g. SDSS, Gaia...)

See if your targets have been observed in other surveys!

TOPCAT – ADQL



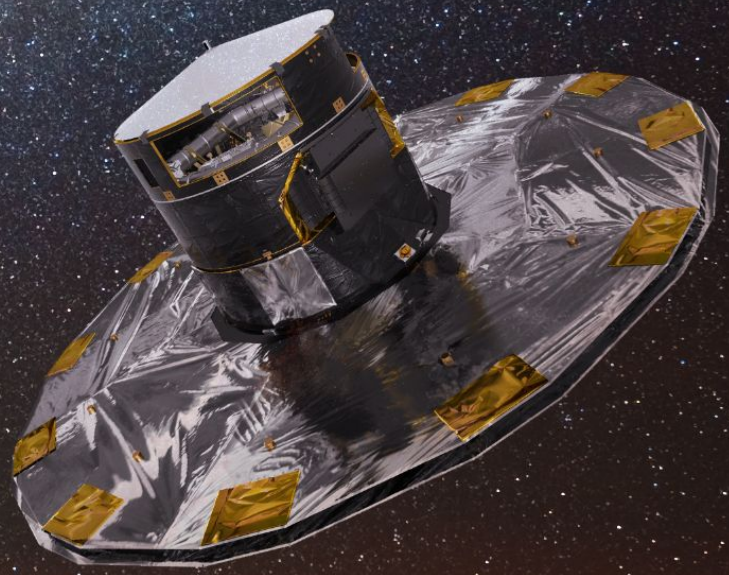
Query for new targets in existing public catalogues

Query using ADQL

- ➔ ADQL = Astronomical Data Query Language
- ➔ For examples see the ‘EXTRAS’ at the end

The Dawn of Gaia

- An ambitious mission to chart the three-dimensional map of our Galaxy
- Full astrometric solution (parallax, proper motion, position) for 1.46 billion sources
- Magnitudes (G, BP, RP) and colours (BP-RP) etc are provided for 1.8 billion sources (~1% of the Galaxy)



Exercise – ADQL queries in TOPCAT

Over to you!

Aims:

- Perform two Table Access Protocol (TAP) queries using an SQL-like language to see how you can obtain ***real*** data from TOPCATS Virtual Observatory (VO)
- Get familiar with some of Gaia Quality Criteria (QC) and other parameters
- Put what you now know about creating subsets into action and implement some QC to clean your data - this is very similar to what RIck did when constructing his catalogue!

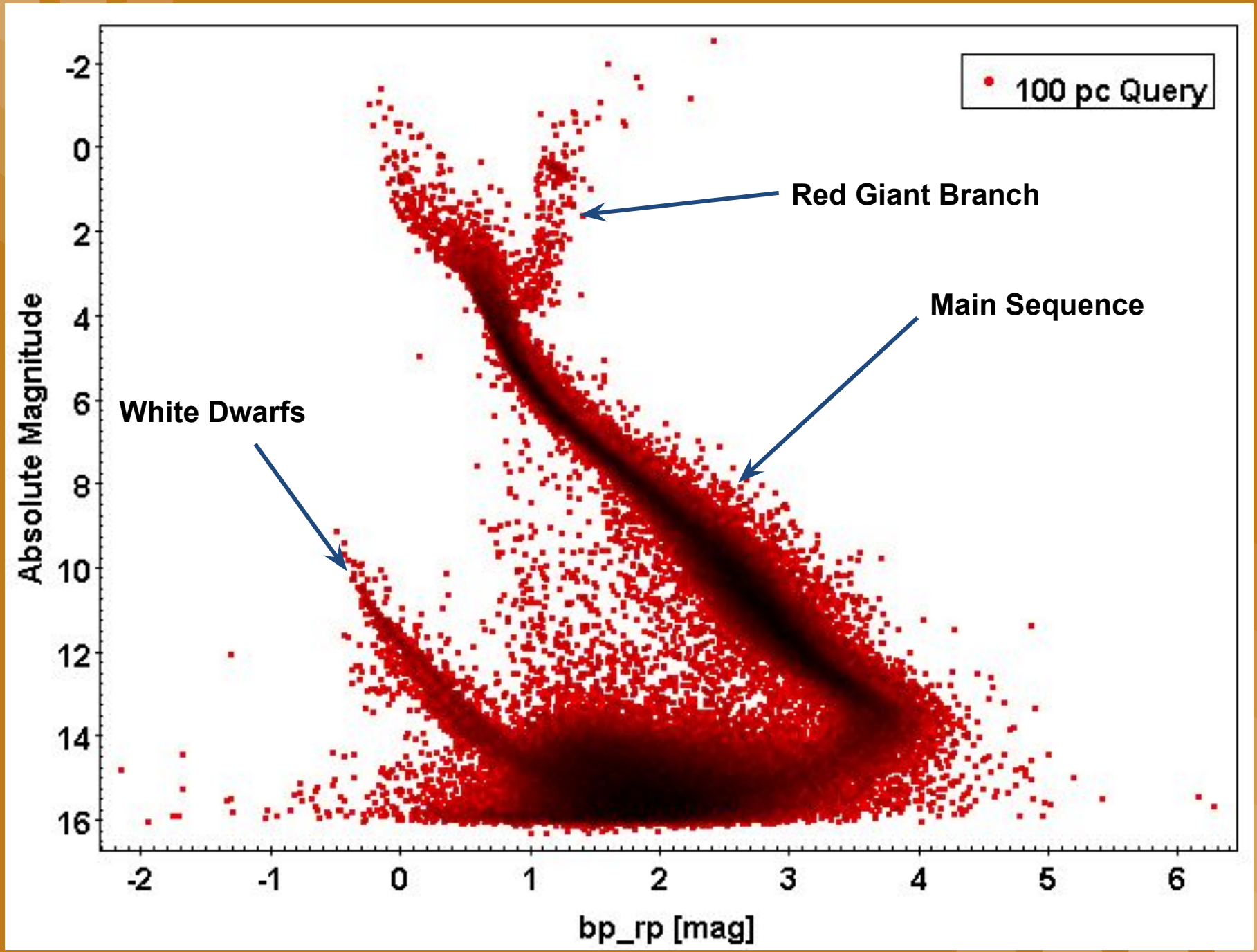
Exercise – ADQL queries in TOPCAT

- We want to query the Gaia database and select 100,000 stars closer than 100 pc. Execute the following:

```
SELECT TOP 100000 source_id, phot_g_mean_mag, bp_rp, parallax,ra,dec
FROM gaiadr2.gaia_source
WHERE parallax > 10
```

- Plot the data in a CMD: absolute magnitude vs colour (bp-rp)
- As gaia only gives apparent magnitude (phot_g_mean_mag), you need to convert it using:
  abs_mag = phot_g_mean_mag - 5*log10(1000/parallax)+5
- Inspect this diagram. Is there something wrong with it? Why?

For more details and examples on these ADQL query conditions, see the `EXTRA' slides at the end


100 pc Query
Red Giant Branch
Main Sequence
White Dwarfs
Absolute Magnitude
bp_rp [mag]
-2
0
2
4
6
8
10
12
14
16
-2
-1
0
1
2
3
4
5
6

Exercise – ADQL queries in TOPCAT

- Gaia also offers quality control (QC) parameters to help identify spurious sources
- Let's retrieve some of these to improve our selection using the 'JOIN USING' query command (appends to our existing table). Execute the following

```
SELECT source_id, a.phot_g_mean_mag, a.bp_rp, a.parallax, a.ra, a.dec,
parallax_over_error, phot_bp_mean_flux_over_error,
phot_rp_mean_flux_over_error, phot_bp_rp_excess_factor, astrometric_chi2_al,
astrometric_n_good_obs_al, astrometric_excess_noise

FROM TAP_UPLOAD.T1 AS a

JOIN gaiadr2.gaia_source AS b USING(source_id)
```

Your table index in 'table list'

- Now create a subset with parallax_over_error < 5 and see how the CMD changes.

- **Now you know how important quality control parameters are!**
- **Use the following conditions to further improve your HR-diagram: (use the '&' symbol when creating a subset to combine all QCs)**

astrometric_excess_noise < 1.0

phot_bp_mean_flux_over_error > 10

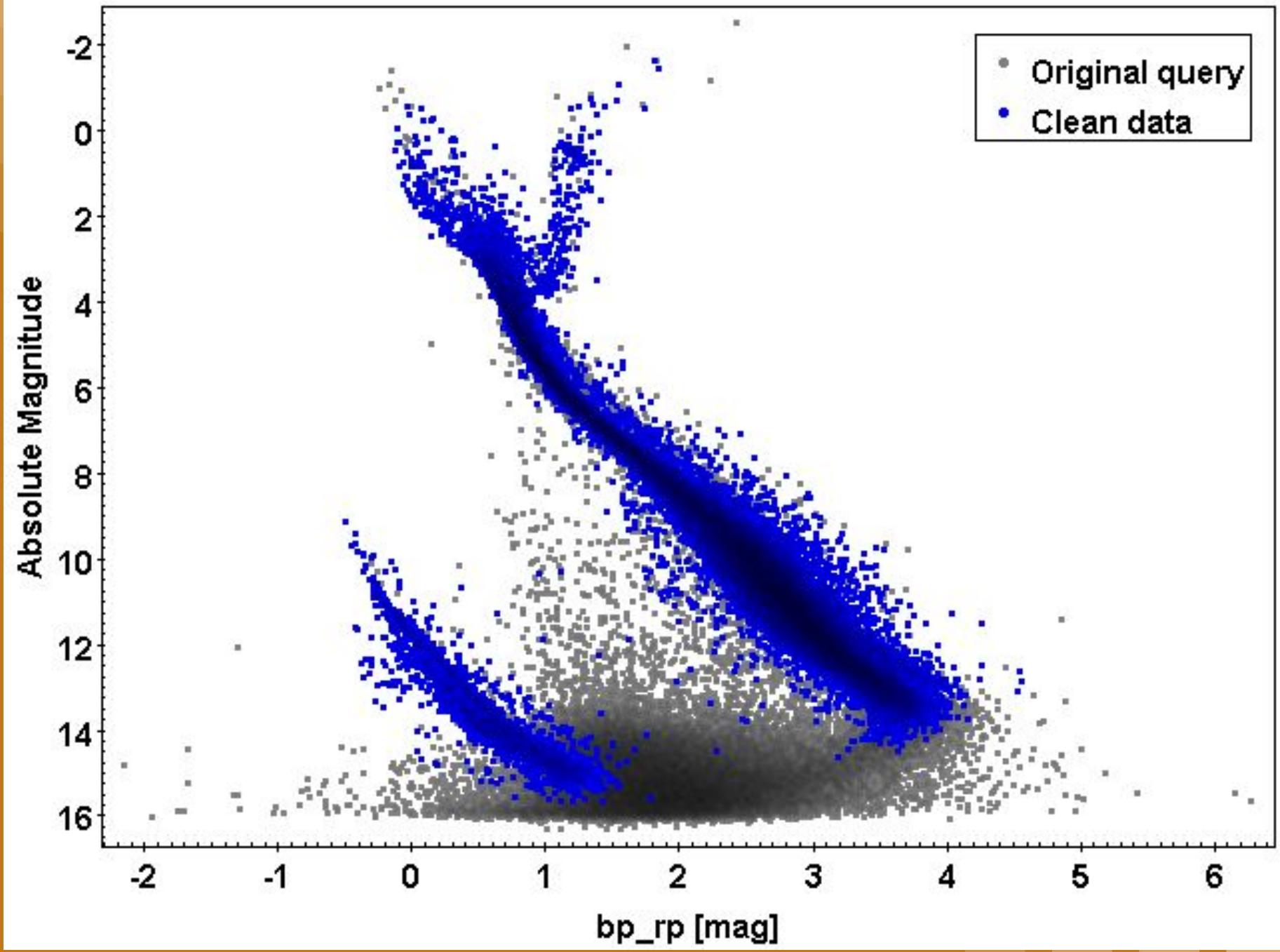
phot_rp_mean_flux_over_error > 10

phot_bp_rp_excess_factor < 1.3+0.06*pow(bp_rp,2)

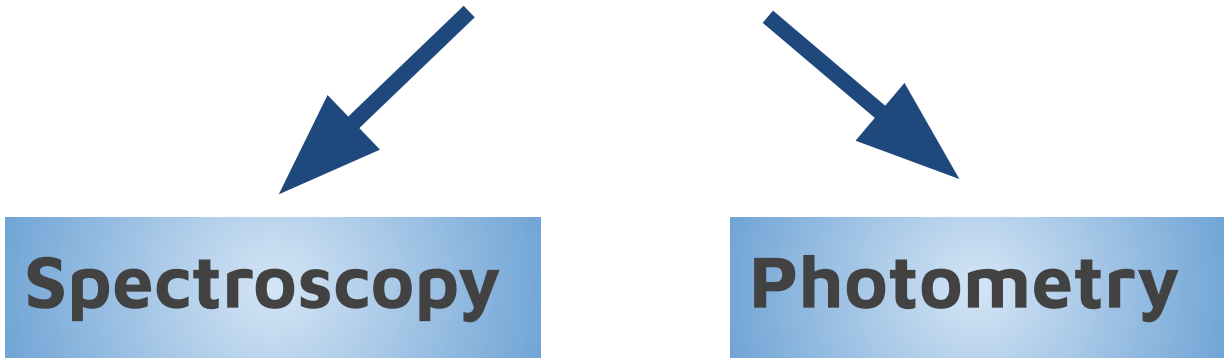
phot_bp_rp_excess_factor > 1.0+0.015*pow(bp_rp,2)

astrometric_chi2_al/(astrometric_n_good_obs_al-5)
< 1.44*max(1, exp(-0.4*(phot_g_mean_mag-19.5)))

Check out this paper: https://arxiv.org/abs/1804.09366 if you want to understand more about where all of these parameters come from.


Original query
Clean data
Absolute Magnitude
-2
0
2
4
6
8
10
12
14
16
-2
-1
0
1
2
3
4
5
6
bp_rp [mag]

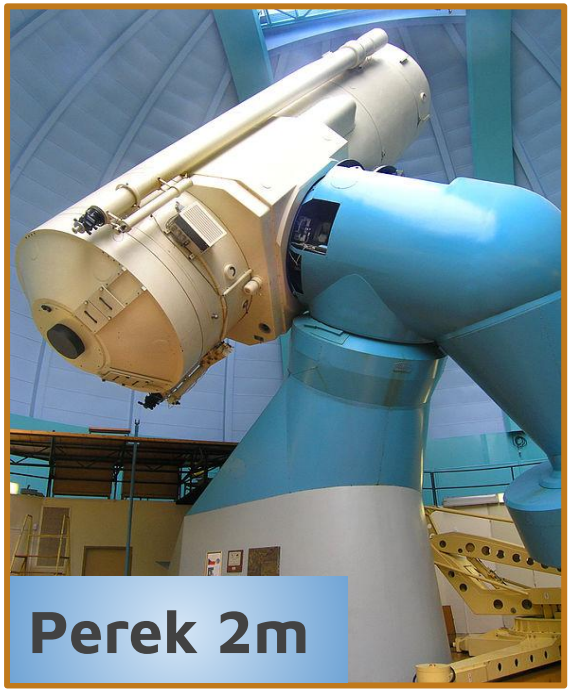
Creating our target lists



Creating our target lists

Spectroscopy

Groups A & B



Perek 2m

Photometry

Groups C & D

Mayer 65cm

Creating our target lists

Observational constraints

Use TOPCAT subsets to select observable stars

- Target declination (dec) restricted by our **latitude**
  (Ondřejov location: 14.78°E **49.9°N** 500m, UT-offset +2)
- Target right ascension (ra) restricted by the **time of year**
  Check the suns position!
- Also consider 'circumpolar' stars. (Dec > ?? always visible!)
  - Check all targets with staralt once you have your list:
    http://catserver.ing.iac.es/staralt/index.php

- Brightness constraints of our telescopes (under perfect conditions!):
  - Perek (spectroscopy): < ~**10.5 Gmag**
  - Mayer (photometry): < ~**16.5 Gmag**

➡ **See Max's presentation slides on Target visibility**

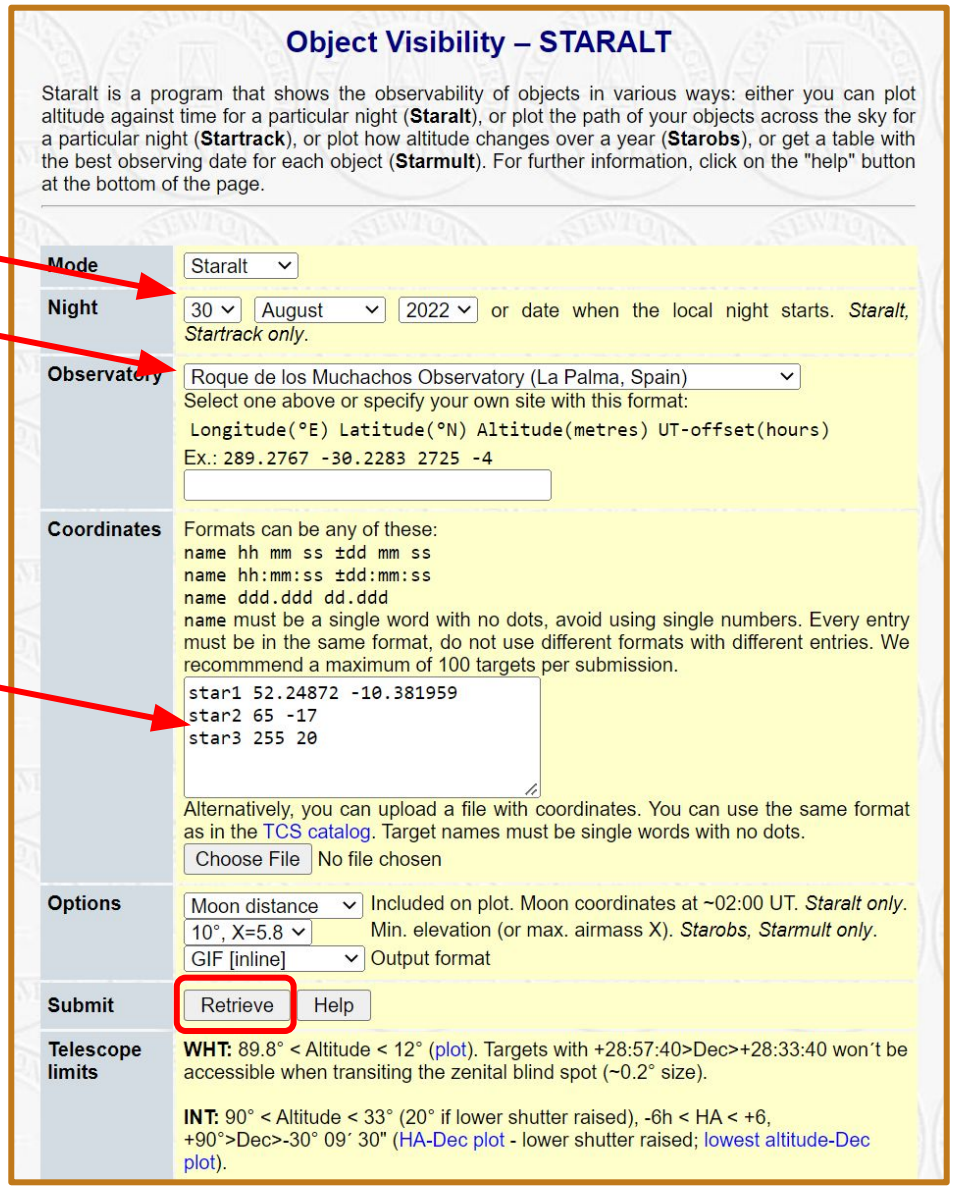
Creating our target lists

Observational constraints

Date

Location

Targets:
name ra dec



Object Visibility – STARALT

Staralt is a program that shows the observability of objects in various ways: either you can plot altitude against time for a particular night (**Staralt**), or plot the path of your objects across the sky for a particular night (**Startrack**), or plot how altitude changes over a year (**Starobs**), or get a table with the best observing date for each object (**Starmult**). For further information, click on the "help" button at the bottom of the page.

Mode Staralt

Night 30 August 2022 or date when the local night starts. *Staralt, Startrack only.*

Observatory Roque de los Muchachos Observatory (La Palma, Spain)
Select one above or specify your own site with this format:
`Longitude(°E) Latitude(°N) Altitude(metres) UT-offset(hours)`
Ex.: `289.2767 -30.2283 2725 -4`

Coordinates Formats can be any of these:
`name hh mm ss ±dd mm ss`
`name hh:mm:ss ±dd:mm:ss`
`name ddd.ddd dd.ddd`
`name` must be a single word with no dots, avoid using single numbers. Every entry must be in the same format, do not use different formats with different entries. We recommmend a maximum of 100 targets per submission.

```
star1 52.24872 -10.381959
star2 65 -17
star3 255 20
```

Alternatively, you can upload a file with coordinates. You can use the same format as in the TCS catalog. Target names must be single words with no dots.
Choose File No file chosen

Options Moon distance — Included on plot. Moon coordinates at ~02:00 UT. *Staralt only.*
10°, X=5.8 — Min. elevation (or max. airmass X). *Starobs, Starmult only.*
GIF [inline] — Output format

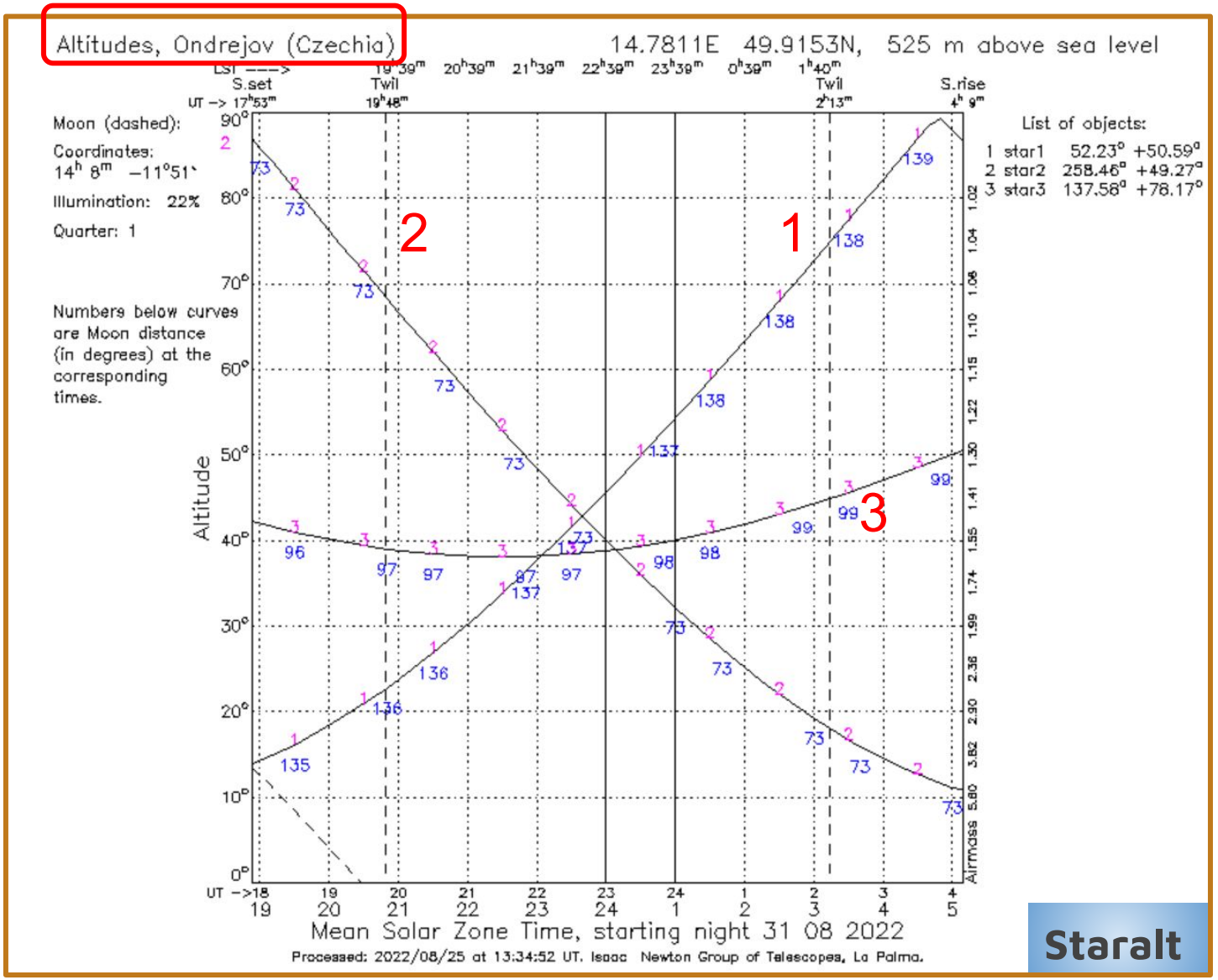
Submit Retrieve Help

Telescope limits **WHT:** 89.8° < Altitude < 12° (plot). Targets with +28:57:40>Dec>+28:33:40 won´t be accessible when transiting the zenital blind spot (~0.2° size).

INT: 90° < Altitude < 33° (20° if lower shutter raised), -6h < HA < +6, +90°>Dec>-30° 09´ 30" (HA-Dec plot - lower shutter raised; lowest altitude-Dec plot).

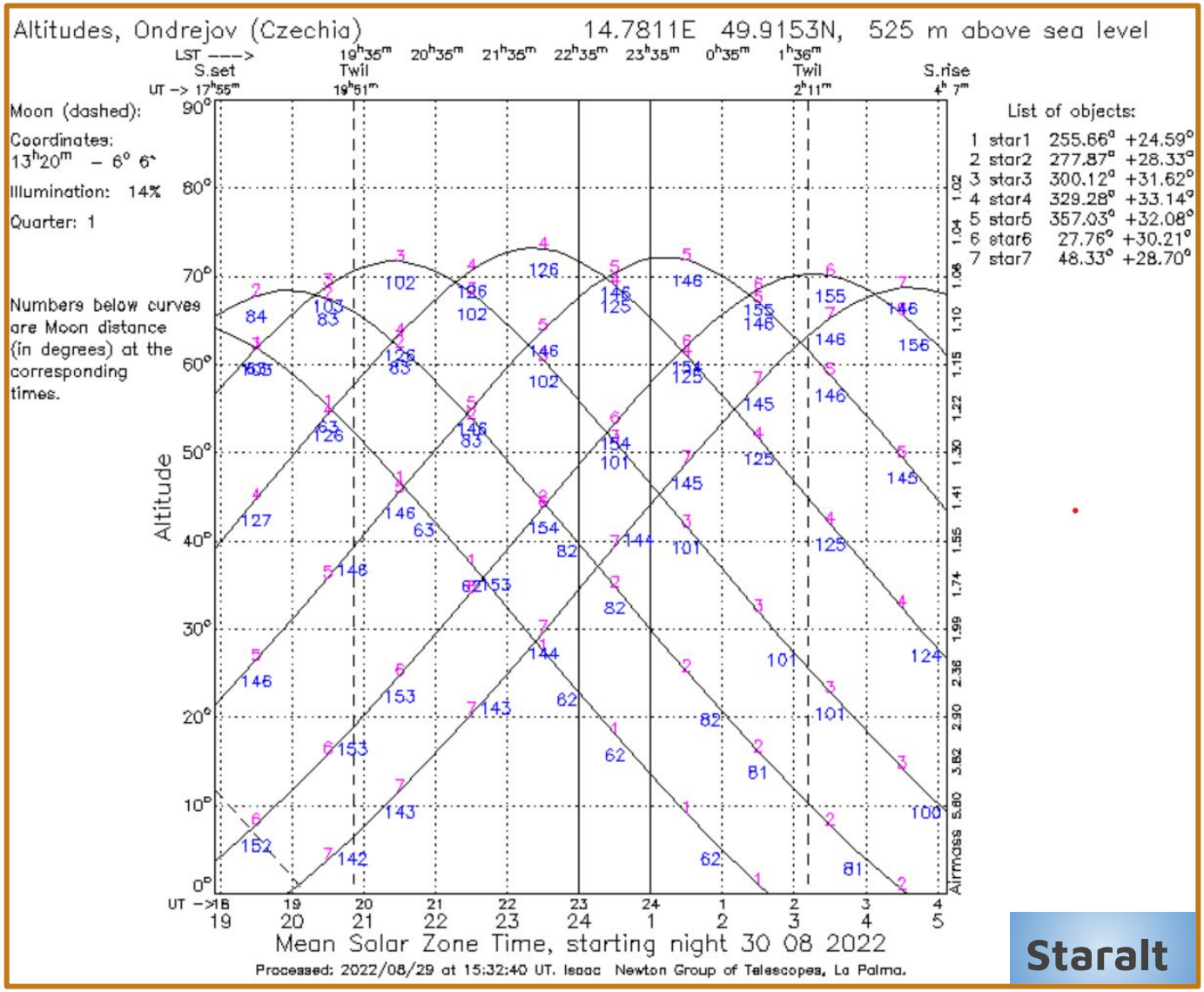
Creating our target lists

Observational constraints



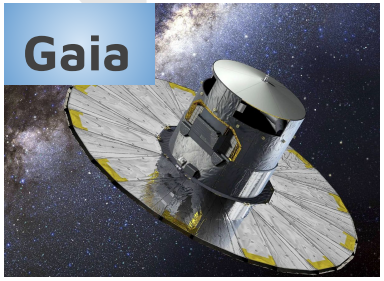
Creating our target lists

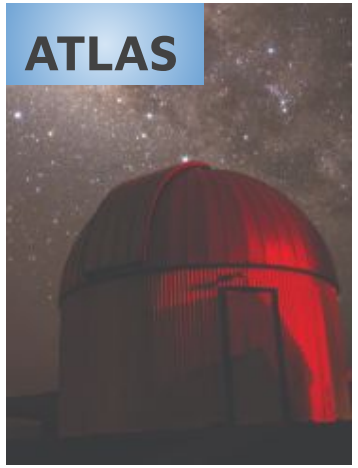
Observational constraints

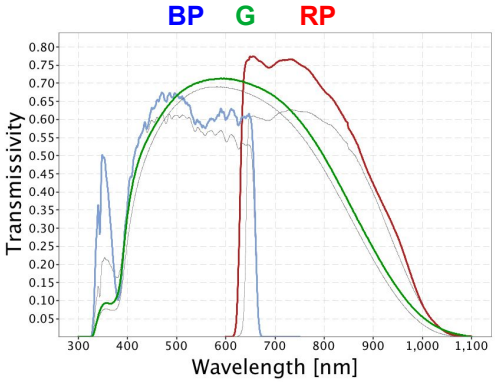


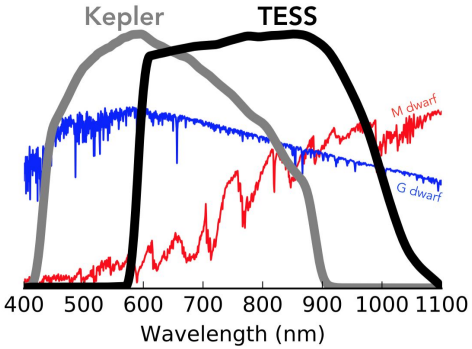
Creating our target lists

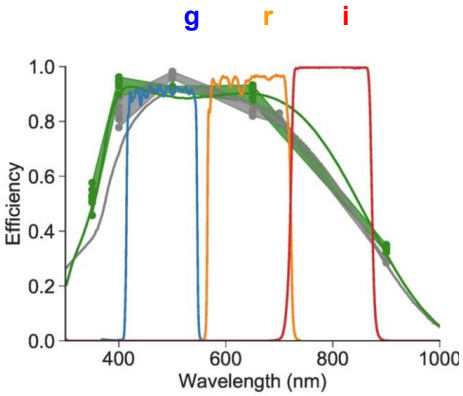
Photometric data

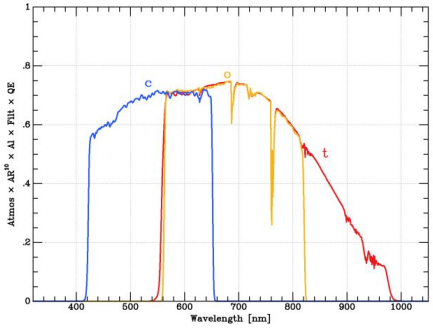






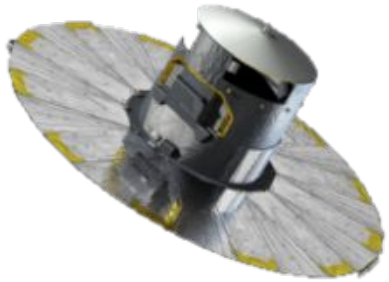






Plus parallaxes!

Creating our spectroscopic target list (groups A & B)



- We want to perform high-res, multi-epoch observations of **blue horizontal branch** stars to systematically probe for binarity. Not been done before!
- We will be using an already published catalogue of BHBs based on Gaias EDR3. Note that these **are not confirmed BHBs** so we first need to confirm their nature! Check SIMBAD if they are already known or have plenty of spectra already.

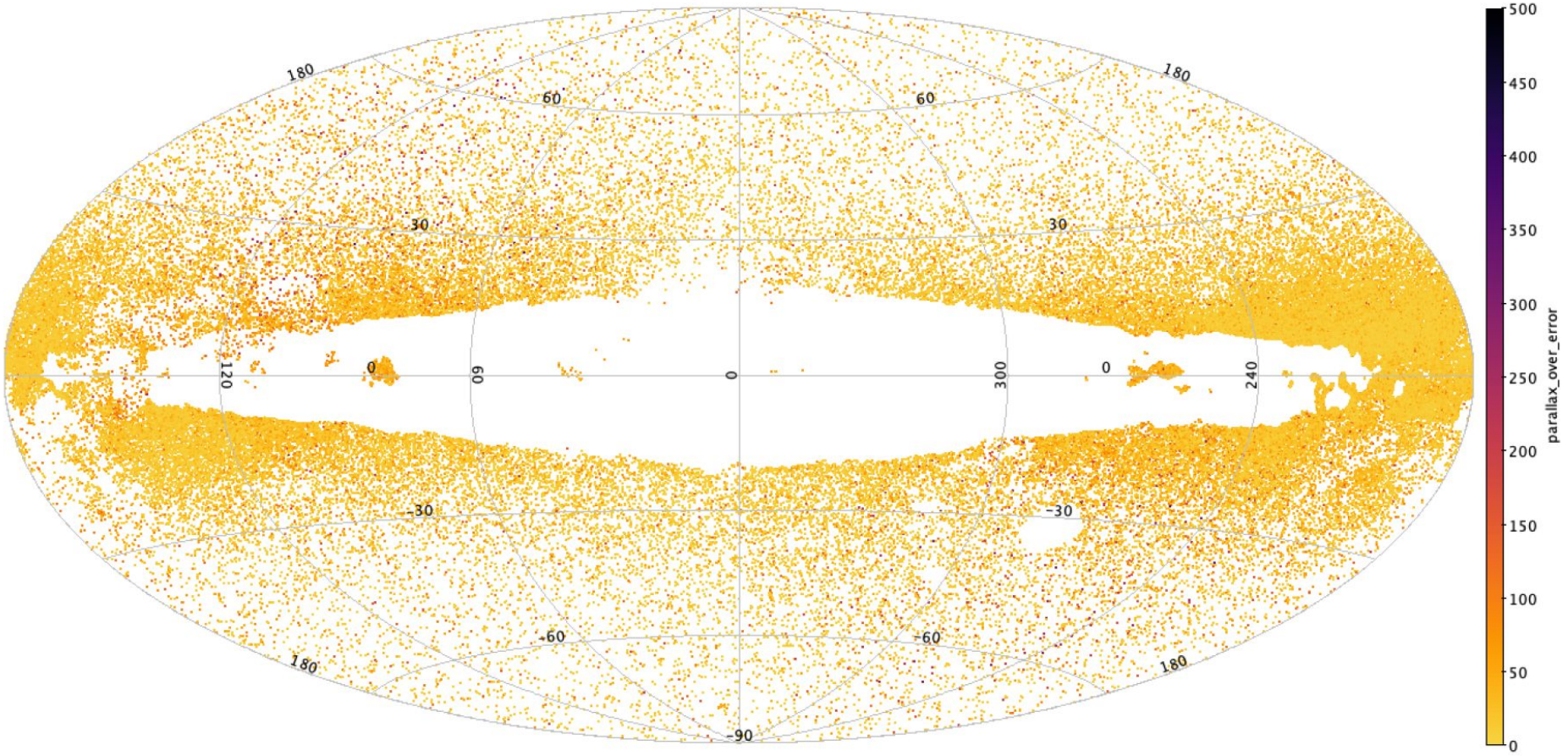
➡ See Rick's paper (and presentation earlier today) for more details on how it was constructed!
2021A&A...654A.107C

- The catalogue of BHBs should be in your shared virtualbox folder as a 'FITS-PLUS' table format called: ...

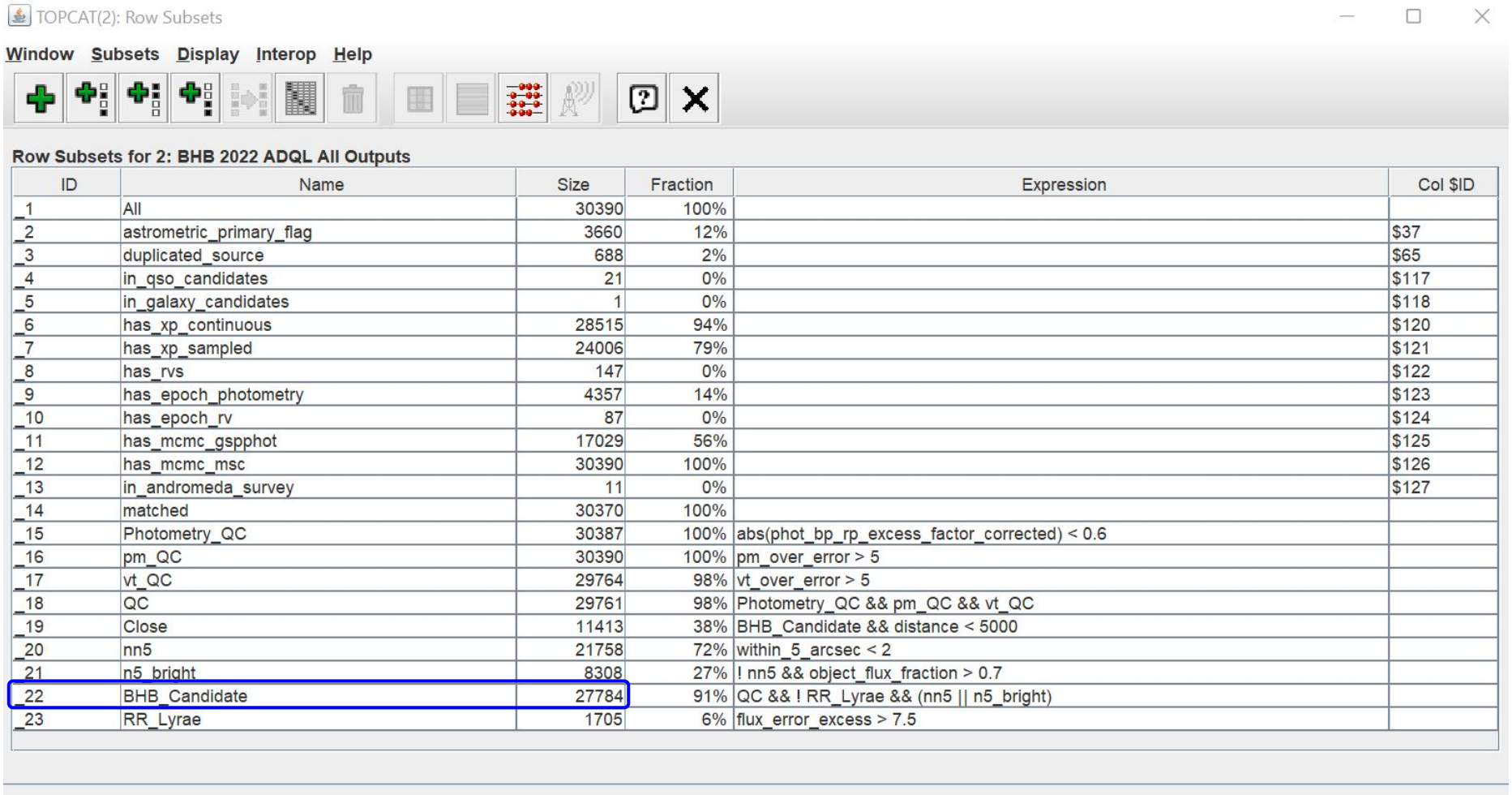
  Open it in TOPCAT and inspect the data. Several subsets have already been made for you.

Creating our spectroscopic target list (groups A & B)

Use TOPCAT to plot a ‘skyplot’ of the data



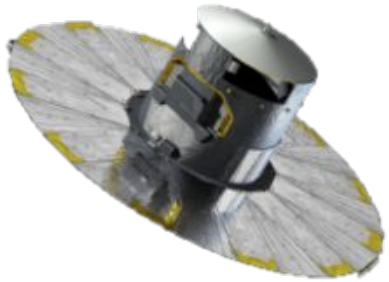
Creating our spectroscopic target list (groups A & B)



Row Subsets for 2: BHB 2022 ADQL All Outputs

ID	Name	Size	Fraction	Expression	Col $ID
_1	All	30390	100%		
_2	astrometric_primary_flag	3660	12%		$37
_3	duplicated_source	688	2%		$65
_4	in_qso_candidates	21	0%		$117
_5	in_galaxy_candidates	1	0%		$118
_6	has_xp_continuous	28515	94%		$120
_7	has_xp_sampled	24006	79%		$121
_8	has_rvs	147	0%		$122
_9	has_epoch_photometry	4357	14%		$123
_10	has_epoch_rv	87	0%		$124
_11	has_mcmc_gspphot	17029	56%		$125
_12	has_mcmc_msc	30390	100%		$126
_13	in_andromeda_survey	11	0%		$127
_14	matched	30370	100%		
_15	Photometry_QC	30387	100%	abs(phot_bp_rp_excess_factor_corrected) < 0.6	
_16	pm_QC	30390	100%	pm_over_error > 5	
_17	vt_QC	29764	98%	vt_over_error > 5	
_18	QC	29761	98%	Photometry_QC && pm_QC && vt_QC	
_19	Close	11413	38%	BHB_Candidate && distance < 5000	
_20	nn5	21758	72%	within_5_arcsec < 2	
_21	n5_bright	8308	27%	! nn5 && object_flux_fraction > 0.7	
_22	BHB_Candidate	27784	91%	QC && ! RR_Lyrae && (nn5 || n5_bright)	
_23	RR_Lyrae	1705	6%	flux_error_excess > 7.5	

Creating our spectroscopic target list (groups A & B)



- Now we need to create an observing plan - **half the job**!

  A rough “road map” for for this is given below:

  Step 1: See which targets are observable **right now** by constraining their right ascensions and declinations (in TOPCAT using subsets), **please don’t point the telescope at the ground!**

  Also consider ‘circumpolar’ targets. You can use staralt.

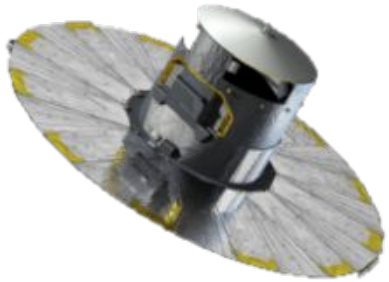
  Step 2: What is the limiting magnitude of your telescope?

  Step 3: Create your observing schedule for the **next three nights** considering:

  - ➔ Exposure and overhead times. How many targets can we observe per night?
  - ➔ Position of the moon as well as sunrise/sunset times. ~30 minutes before and after twilight is usually fine.
  - ➔ We’re observing each target **at least two times!** Plan accordingly.
  - ➔ Which targets should be observed first?
  - ➔ Check which BHBs were **observed last year** for follow-up

  Step 4: Create a spreadsheet to share!

Creating our spectroscopic target list (groups A & B)



- Now we need to create an observing plan - **half the job**!

  A rough “road map” for for this is given below:

  Step 1: See which targets are observable **right now** by constraining their right ascensions and declinations (in TOPCAT using subsets), **please don’t point the telescope at the ground!**

  Also consider ‘circumpolar’ targets. You can use staralt.

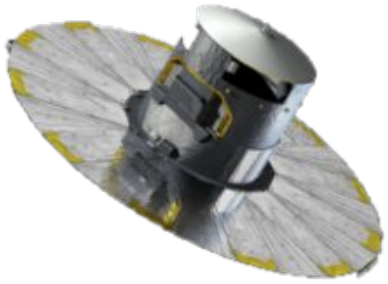
  Step 2: What is the limiting magnitude of your telescope?

  Step 3: Create your observing schedule for the **next three nights** considering:

  - ➔ Exposure and overhead times. How many targets can we observe per night?
  - ➔ Position of the moon as well as sunrise/sunset times. ~30 minutes before and after twilight is usually fine.
  - ➔ We’re observing each target **at least two times!** Plan accordingly.

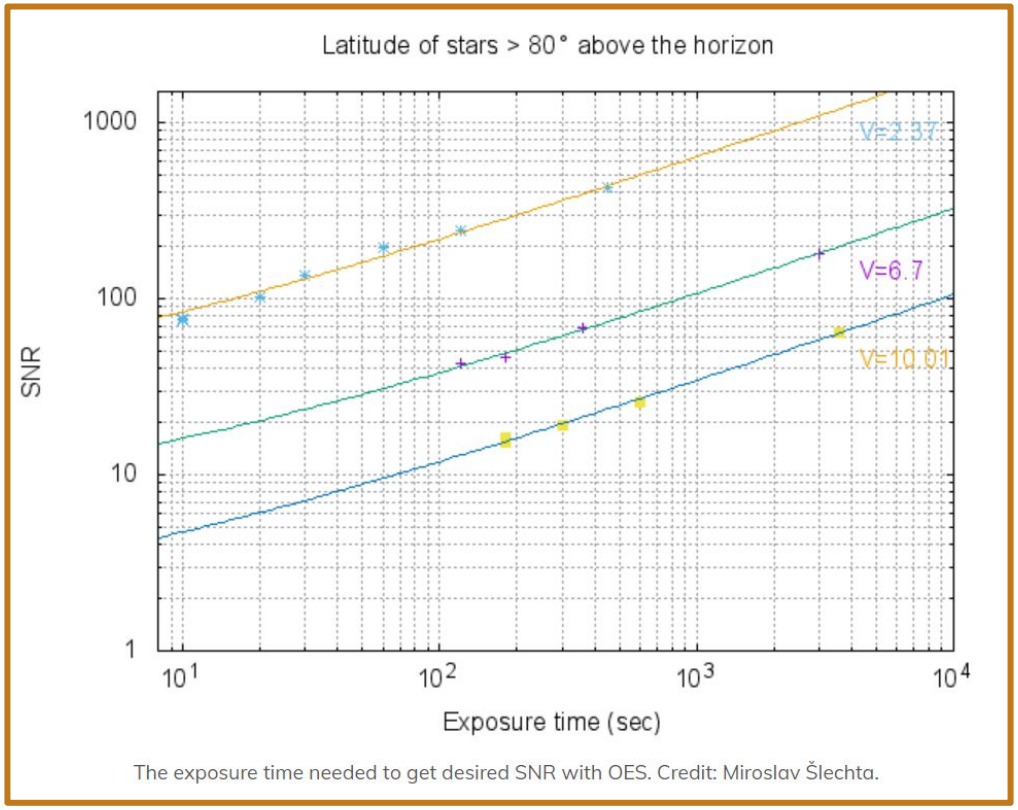
Check the observation log from last years workshop:
https://docs.google.com/spreadsheets/d/17yzRLfii18cOfXg3jlWMX2errtwyF03DSi5Gp2NGHzM/edit?usp=sharing

Creating our spectroscopic target list (groups A & B)



- Exposure time estimators for the Perek telescope

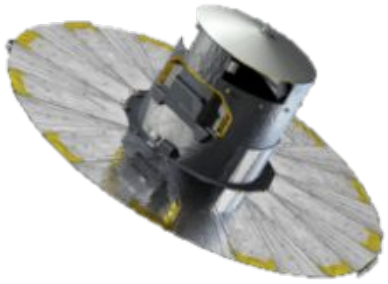
Magnitude (G)	8	9	10	11
Exposure (s)	1800	2700	3600	5200 !



The exposure time needed to get desired SNR with OES. Credit: Miroslav Šlechta.

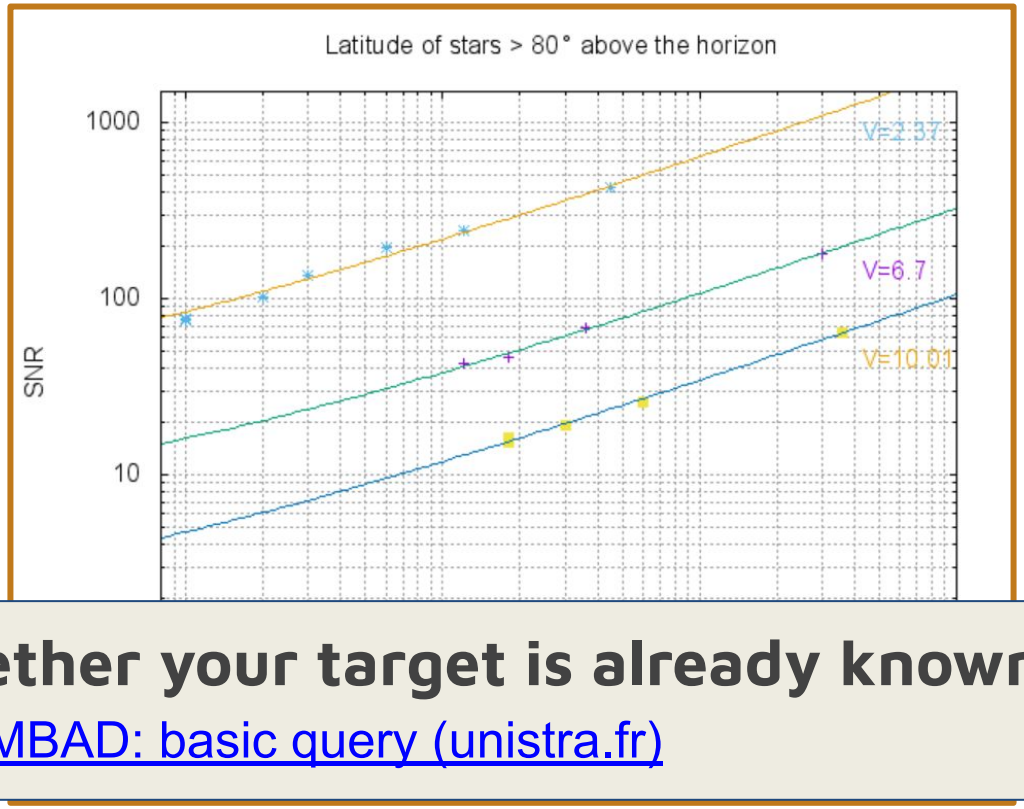
Echelle spectrograph (cas.cz)

Creating our spectroscopic target list (groups A & B)



- Exposure time estimators for the Perek telescope

Magnitude (G)	8	9	10	11
Exposure (s)	1800	2700	3600	5200 !



Check whether your target is already known in SIMBAD!

➡ SIMBAD: basic query (unistra.fr)

Creating our photometric target list (groups C & D)

- We want to observe **hot subdwarf stars with suspected variability**.
- Each group will be working on a **different project**:

  Group 3 -> Follow-up of HW Vir systems discovered in ATLAS and ZTF

  Group 4 -> Classifying hot subdwarf candidates from the Gaia variability catalogue
- We are going to use a table containing 60,000+ hot subdwarf candidates:

  -> Find this catalogue within the shared virtualbox folder and open it in TOPCAT

Creating our photometric target list (groups C & D)

- Faintest magnitude we can observe with the 65 cm Mayer telescope: **16.5 mag (exposure time: 2 min)**
- When planning your run, remember :
  - we want to observe a complete orbital period in BVR or I filter, for Group 3 we want light curves of the same target in different filters, for Group 4 for classifying only one light curve in the filter with the highest S/N
  - in case of bad weather we will need **bright back-up targets**, so prepare targets of different brightnesses from 12/13 mag to the limiting magnitude
  - we need **at least ~100 data points per orbit**, so fainter short-period targets might still not be observable

Creating our photometric target list
Group C (project 1)

- To identify candidate variables, we will use the ATLAS catalogue:

  https://archive.stsci.edu/prepds/atlas-var/

  (The “Object Table”)

  * This table is 7GB in size! Instead, here we will use the compressed version:

  -> Find instead the file called ‘ATLAS_cat.fits’ (300MB) with fewer columns or download at:
  http://www.astro.physik.uni-potsdam.de/~pelisoli/AstroWorkshop/ATLAS_cat.fits

- We will also check **ZTF** (different passbands)

  -> There is a given CSV file of targets called you can upload into TOPCAT ‘table_paper_publication_simplified.csv’

Creating our photometric target list
Group C (project 1)

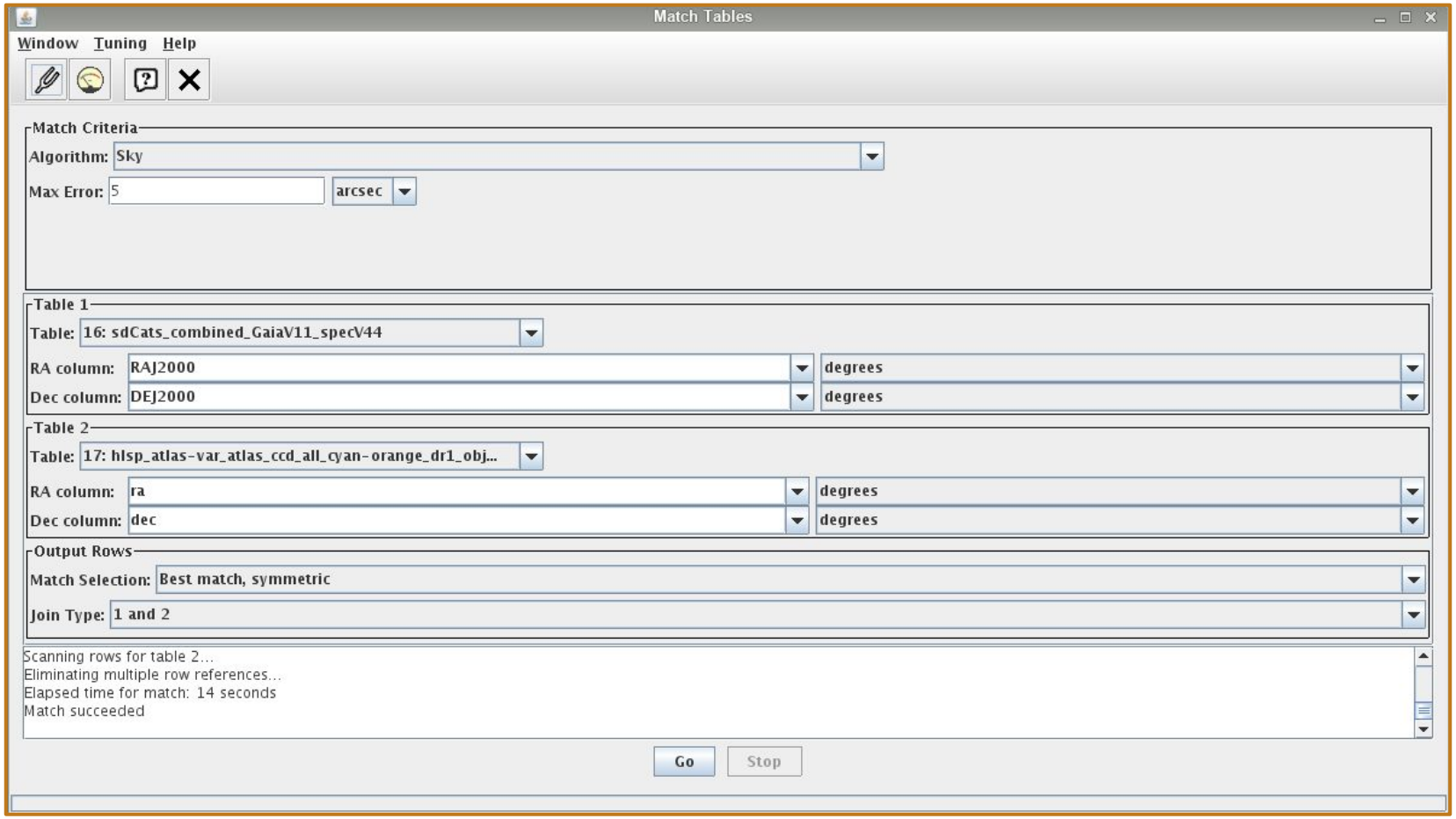
- **Step 1:** import all three tables to TOPCAT.
- **Step 2:** select only relevant columns from the ATLAS table.

  - There are 197 (!) columns in the full table – they describe many parameters in the variability search algorithm run by ATLAS.
  - Using the column metadata shortcut, all columns but the following are deselected:
    - ATO_ID
    - ra and dec (we need those to do a crossmatch)
    - fp_period
    - fp_fitrms
    - fp_fitchi
    - CLASS (this is the type of variation ATLAS identified)

We are interested in short period binaries. These parameters describe the fitted period, root-mean-square, and chi-square of the short-period algorithm in ATLAS. However periods given by ATLAS are not always correct!!!.

Creating our photometric target list
Group C (project 1)

- **Step 3:** cross-match ATLAS with the catalogue of hot subdwarf candidates

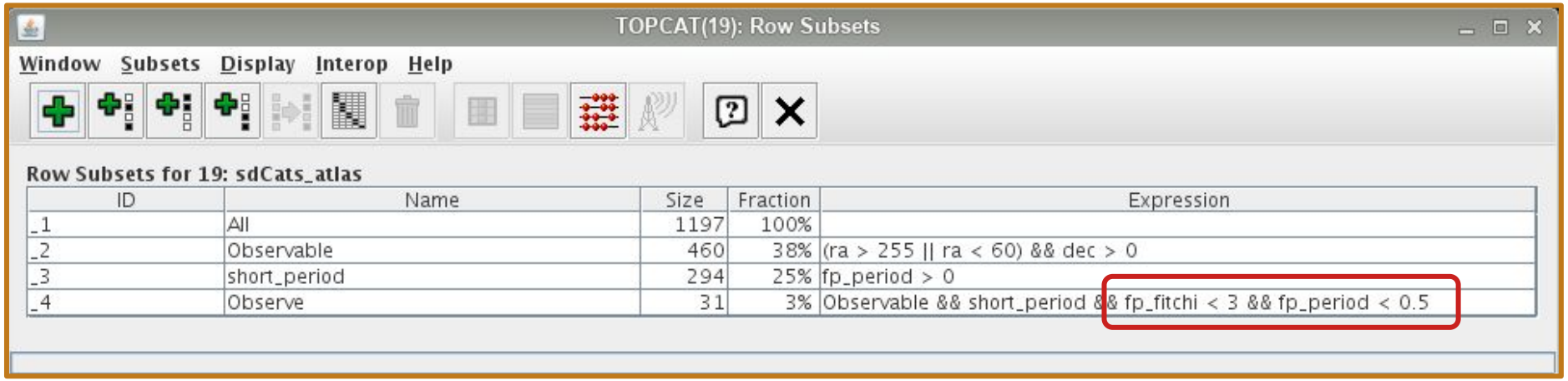


Creating our photometric target list

Group C (project 1)

- **Step 4:** Once again create a subset of observable objects for both ATLAS and ZTF targets given our location.
  - You can again use staralt: http://catserver.ing.iac.es/staralt/index.php

    (Ondřejov location: 14.46°E 49.54°N 32m, UT-offset +2)
  - The **moon is like a giant light bulb**, make sure it's not too close to your targets! This is particularly important for photometry.



ID	Name	Size	Fraction	Expression
_1	All	1197	100%	
_2	Observable	460	38%	(ra > 255 || ra < 60) && dec > 0
_3	short_period	294	25%	fp_period > 0
_4	Observe	31	3%	Observable && short_period && fp_fitchi < 3 && fp_period < 0.5

Creating our photometric target list

Group C (project 1)

- To determine the best targets, you need to inspect the light curves and perform a period search. Using the shape of the phase-folded light curve you will be able to identify HW Vir candidates
  - **Step 5:** For your observable targets, we want to have a look at the photometric data already available to assess it's quality.
    - Use the provided 'ATLAS_ZTF_LombScargle' python script to download the data, perform a Lomb-Scargle periodogram and phase-fold the data.
    - What is the period? We preferably want targets whose orbit we can cover in a single night (**period < 1 d**)
    - Check to see if the target is already published in SIMBAD, our goal is to analyse new systems!
    - **Important**: always also have bright back-up targets in case of bad weather, so have enough targets of different brightnesses in your target list

Check whether your target is already known in SIMBAD!

➡ SIMBAD: basic query (unistra.fr)

Creating our photometric target list
Group D (project 2)

- Gaias most recent data release (DR3) released a catalogue of over **10 million variable objects**. See arXiv.2206.06416 for details.
- We will perform an ADQL cross-match to see how many of our confirmed hot subdwarfs are variable in gaia:

```
SELECT *
FROM TAP_UPLOAD.T1 AS t
JOIN gaiadr3.vari_summary AS g
USING (source_id)
```

➔ This should return 831 variable abjects. How many are observable at Ondrejov? **Create your target list (step 1).**

➔ Exposure time is more complicated for photometry and depends on the orbital period.

Creating our photometric target list
Group D (project 2)

- We will now use the provided script to download and inspect the Gaia light curves, as well as from TESS and ZTF as a cross-match. Use 'GAIA_TESS_ZTF_download_LombScargle' -> this needs to be done one at a time.
- Is there enough data points in the G, BP and RP bands? Has the Lomb-scargle periodogram returned a clear peak? What is the period?
- Once again check the data from ZTF and TESS to see if the target is also variable there. If there is no data, we should definitely observe it! Otherwise you can use the TESS and ZTF data already to classify the target.
- **Important**: always also have bright back-up targets in case of bad weather, so have enough targets of different brightnesses in your target list

Check whether your target is already known in SIMBAD!

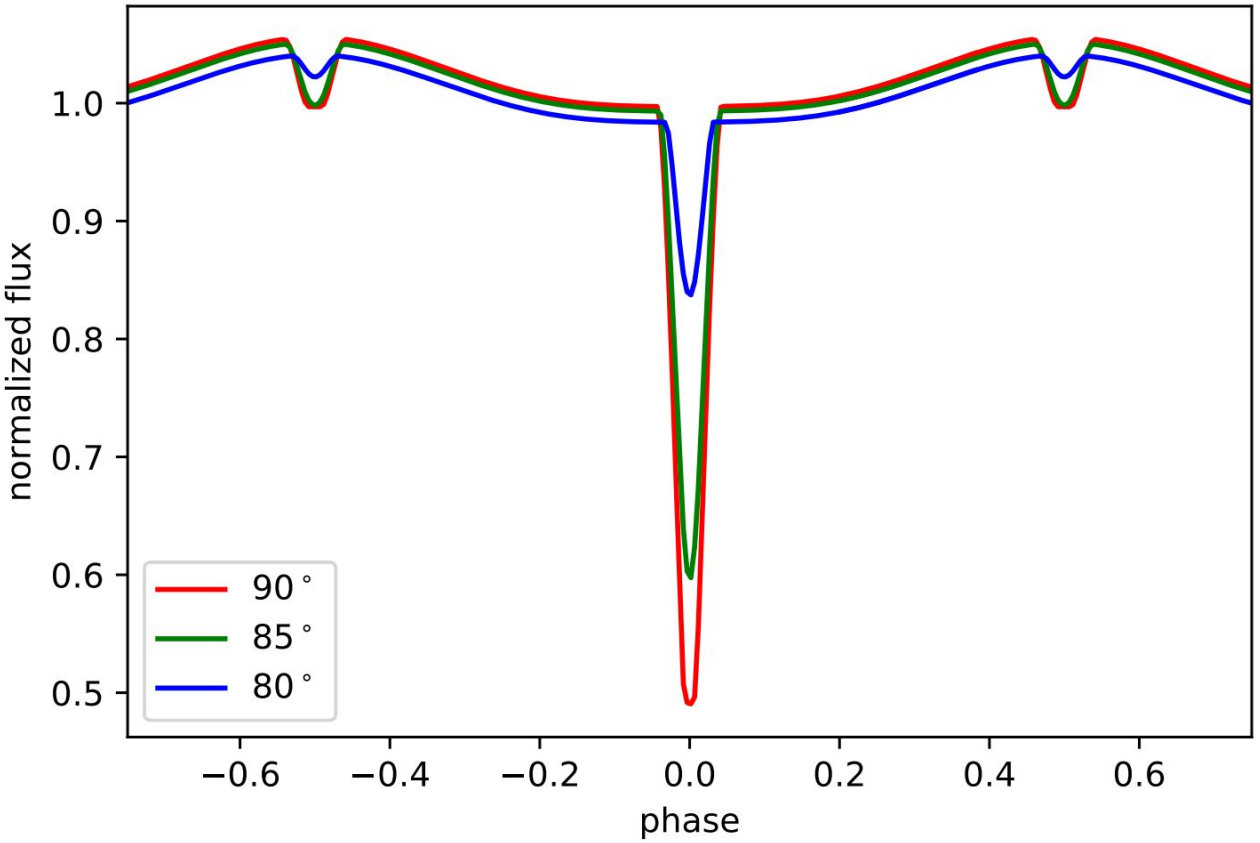
➡ SIMBAD: basic query (unistra.fr)

Bonus!

Light curve classification

Eclipsing Reflection effect systems (HW Vir candidates) P = 0.05 - 1.26d

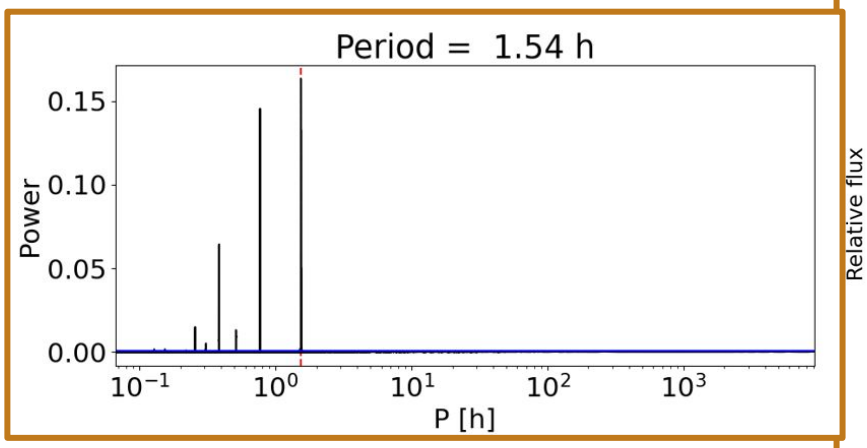
Our target!

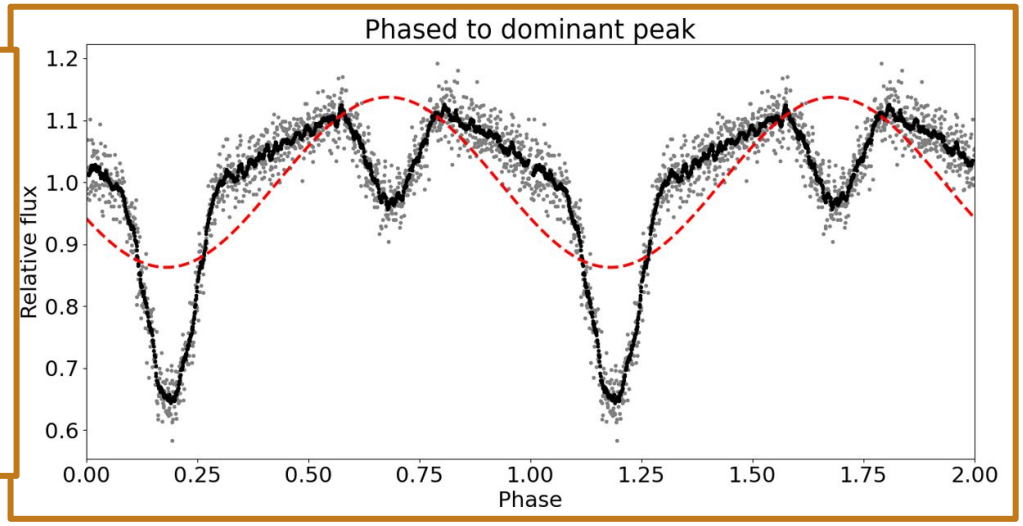


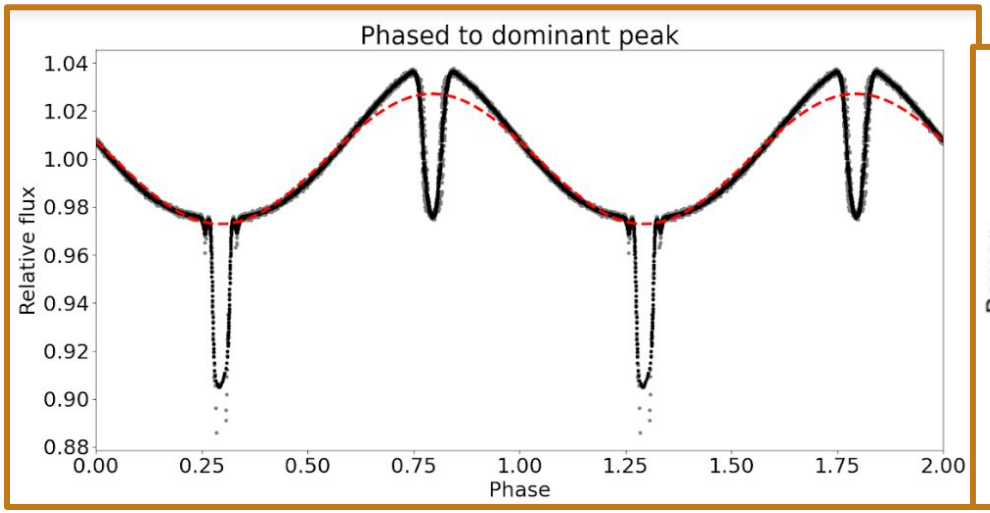
➡ See Veronikas slides on Photometry

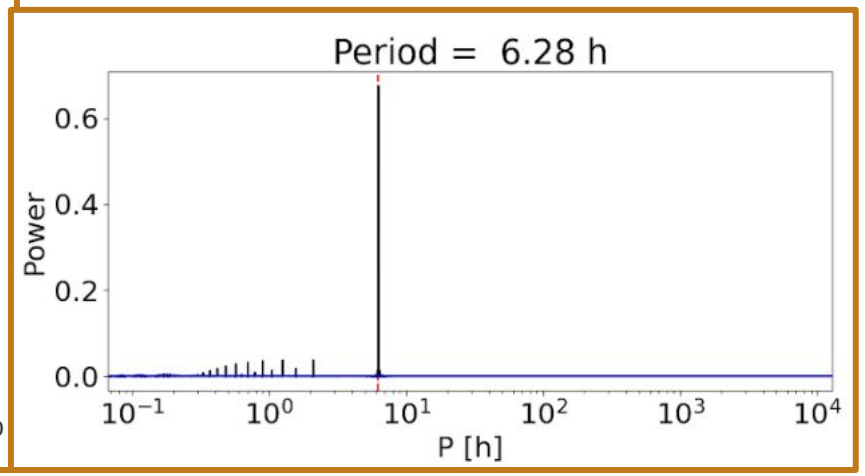
Light curve classification

Eclipsing Reflection effect systems (HW Vir candidates) P = 0.05 - 1.26d



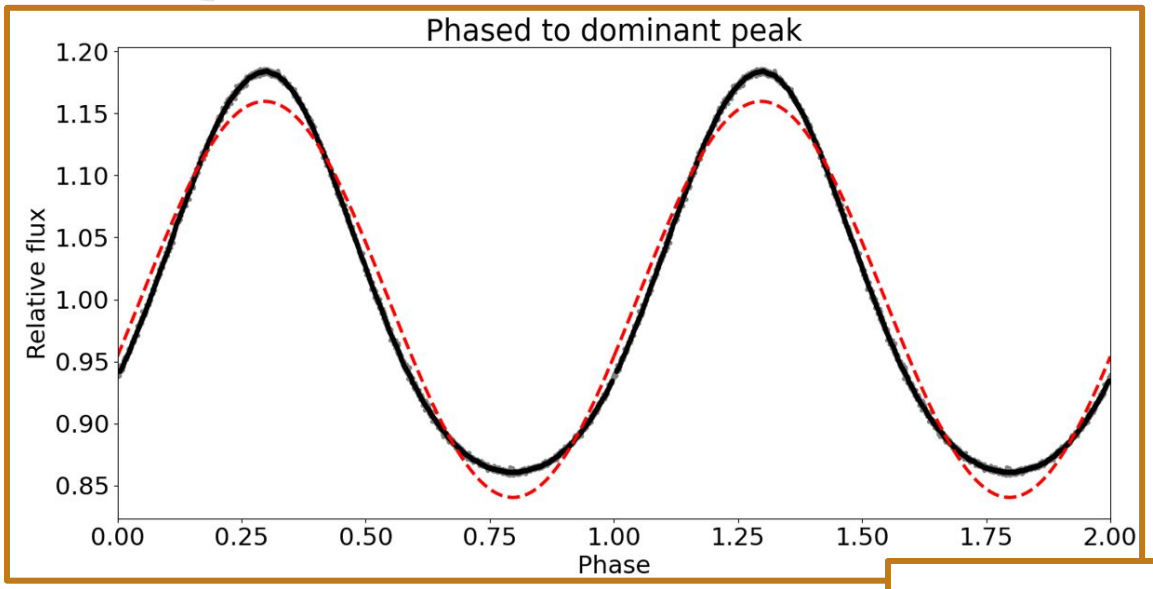


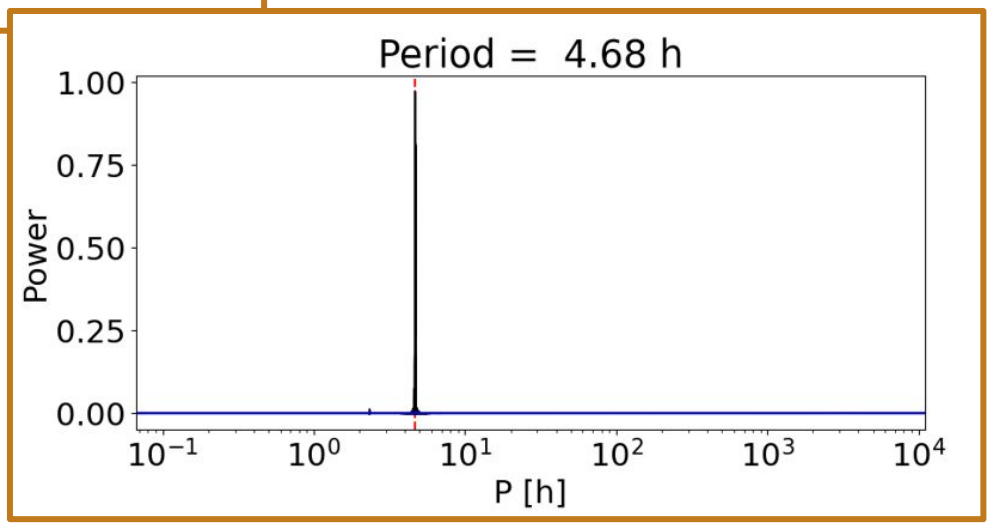




Light curve classification

Reflection/irradiation effect (without eclipses) - Not sinusoidal!

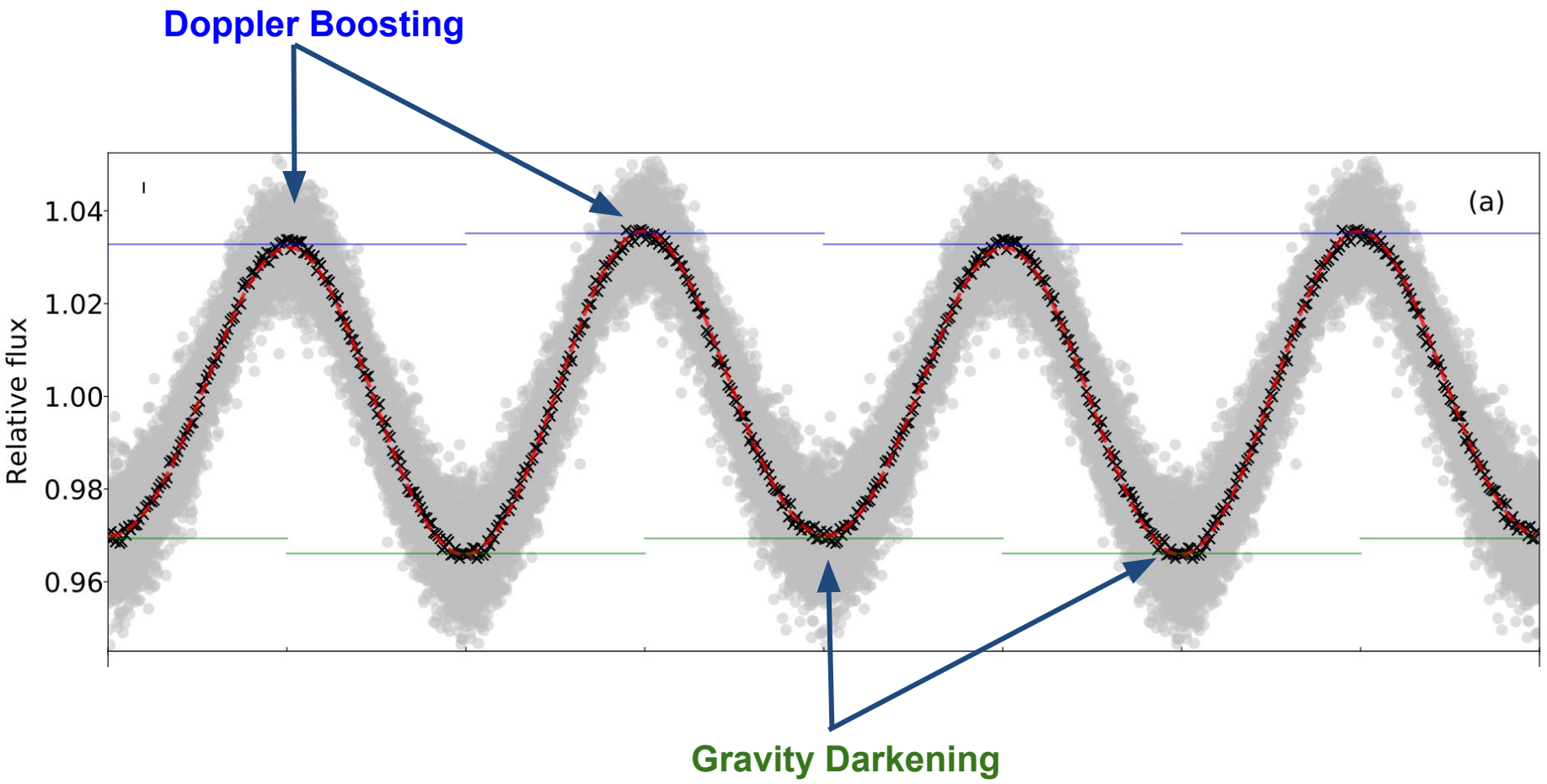




Light curve classification

What else could we expect?

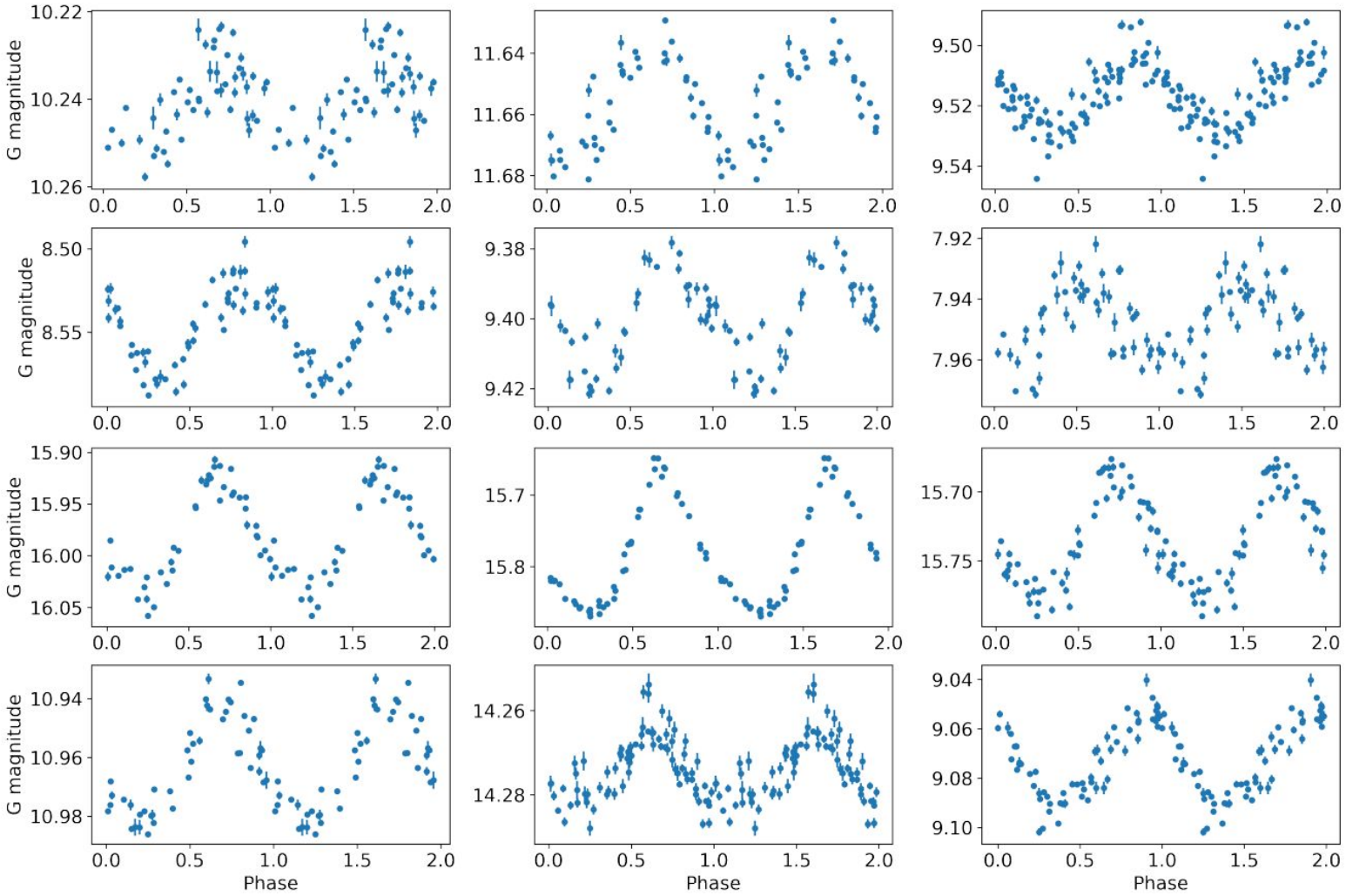
Ellipsoidal Deformation - Sinusoidal



Light curve classification

What else could we expect?

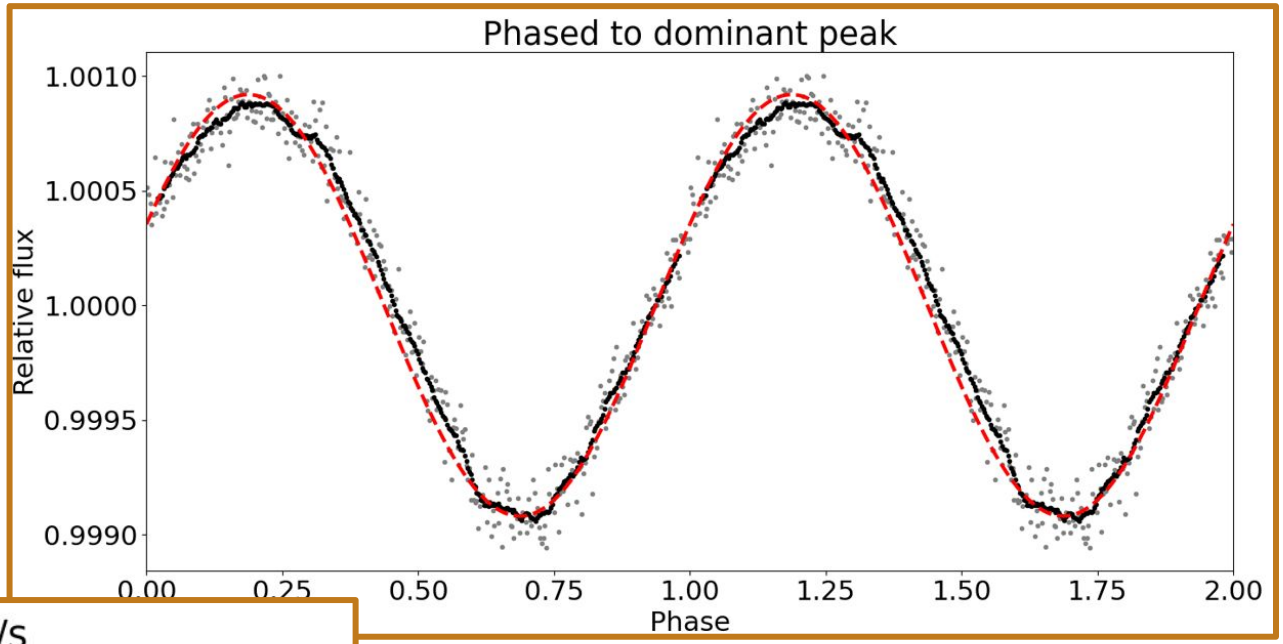
Pulsations (OBAF-Type Main Sequence) - Usually small amplitudes which are not constant

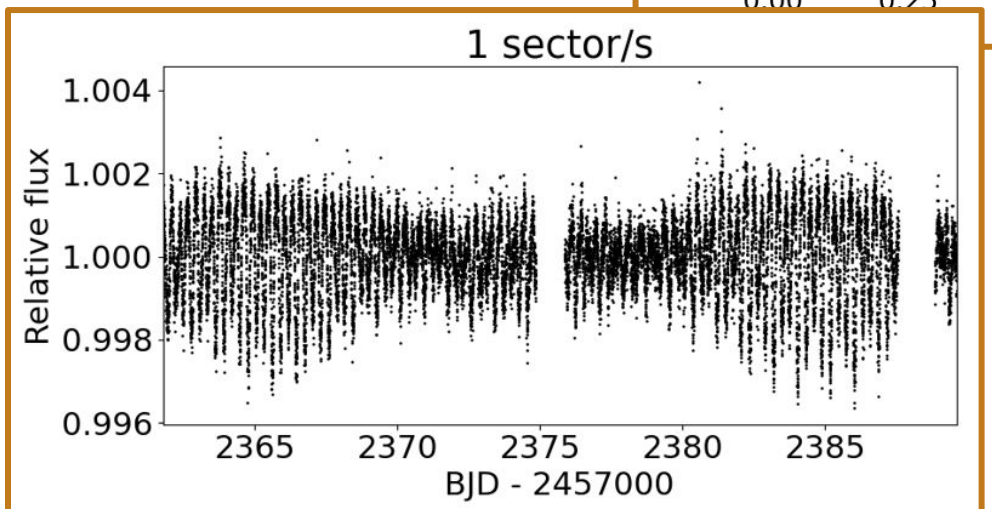


Light curve classification

What else could we expect?

Pulsations (OBAF-Type Main Sequence) - Usually small amplitudes which are not constant



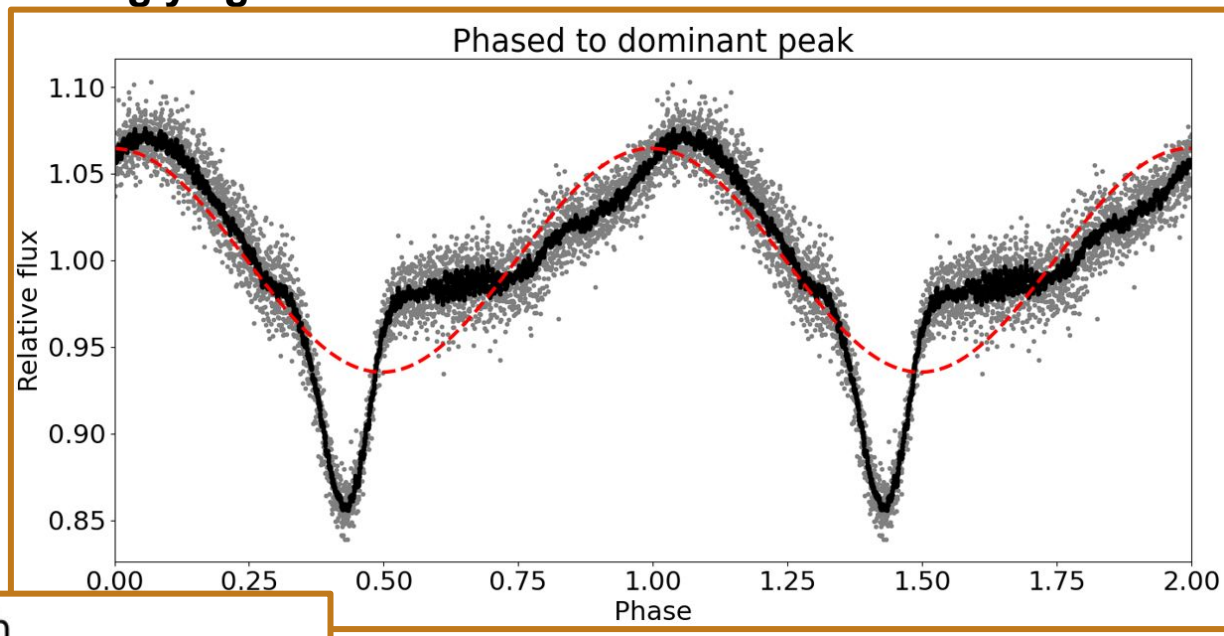


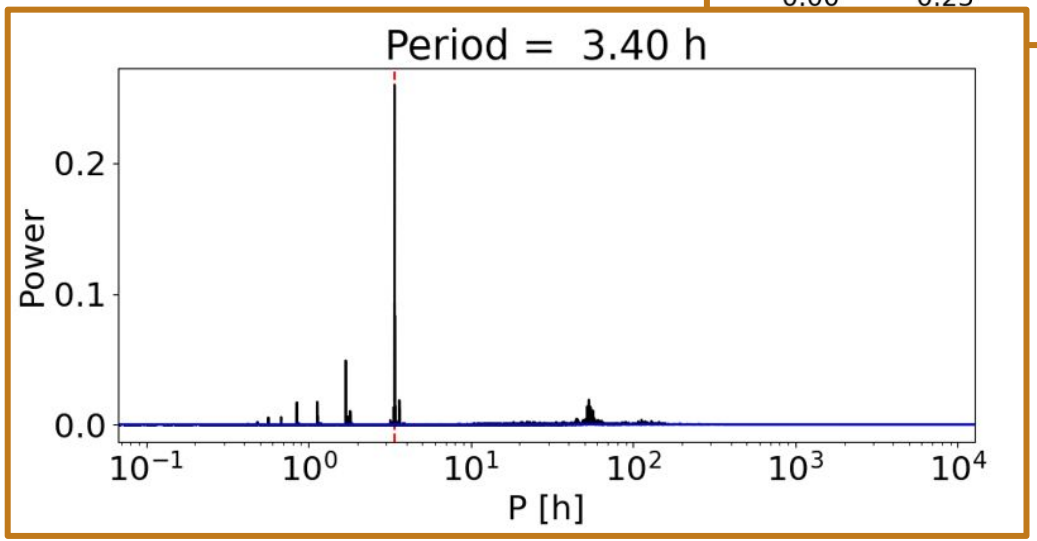
➡ **Beats might be seen in the raw data**

Light curve classification

What else could we expect?

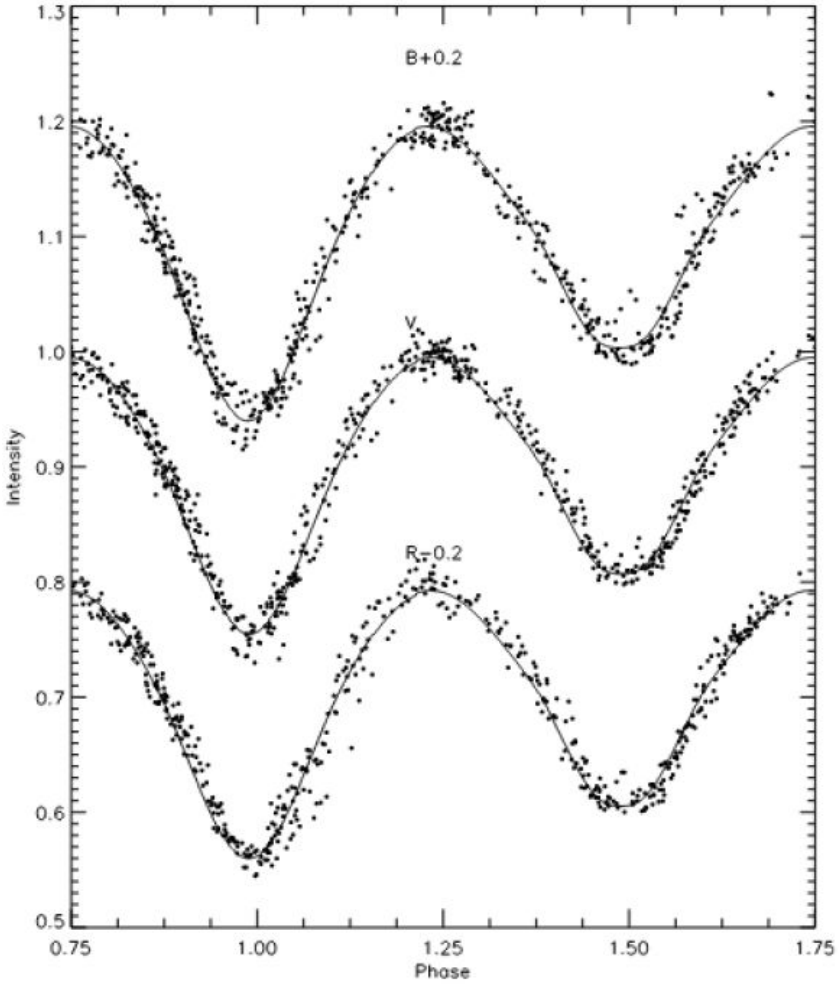
Cataclysmic Variables - Ugly light curves!

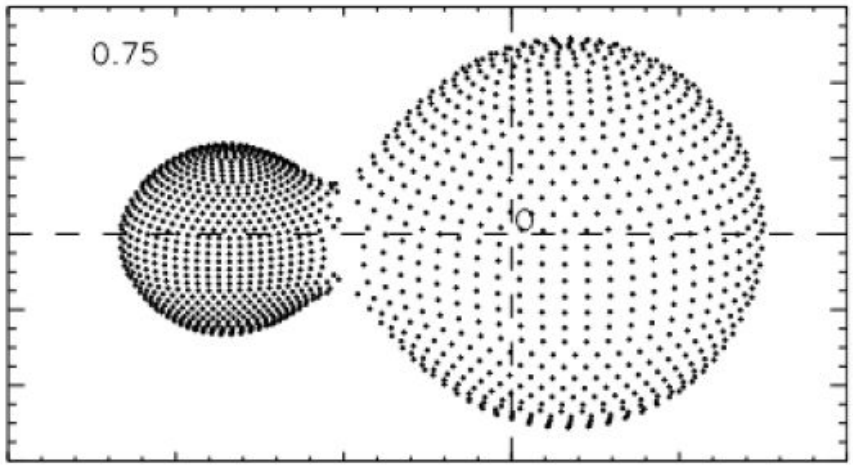




Light curve classification
What else could we expect?

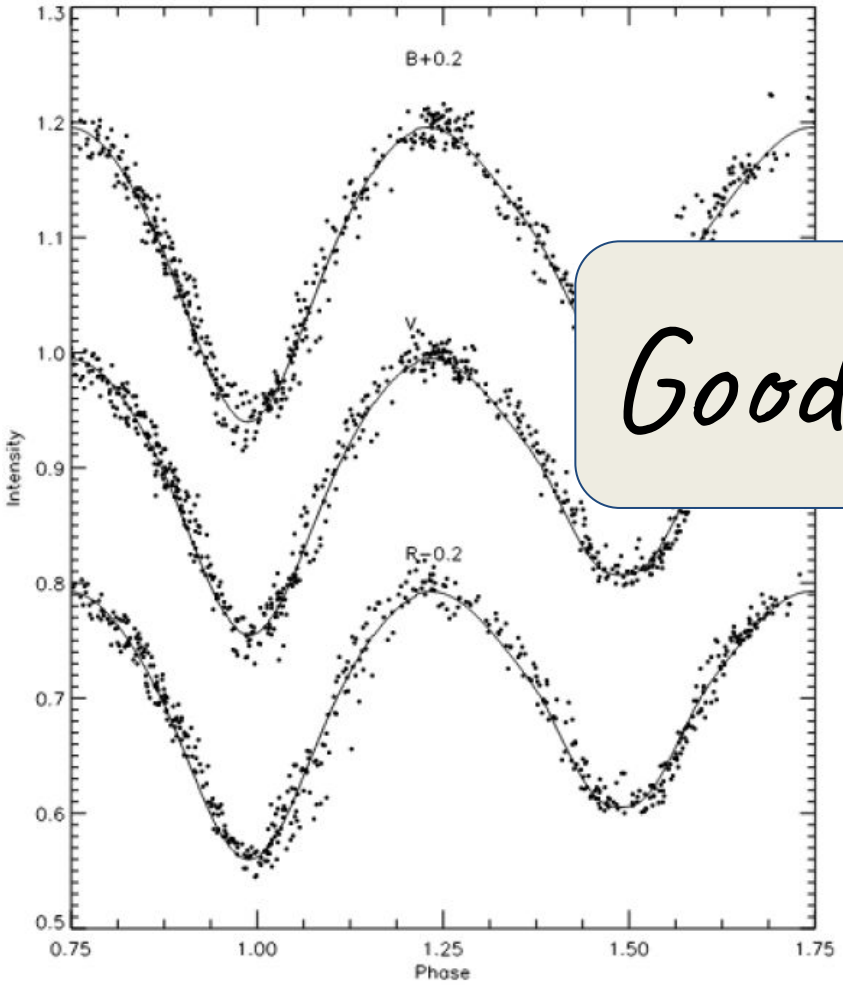
Contact systems: V-shaped LCs



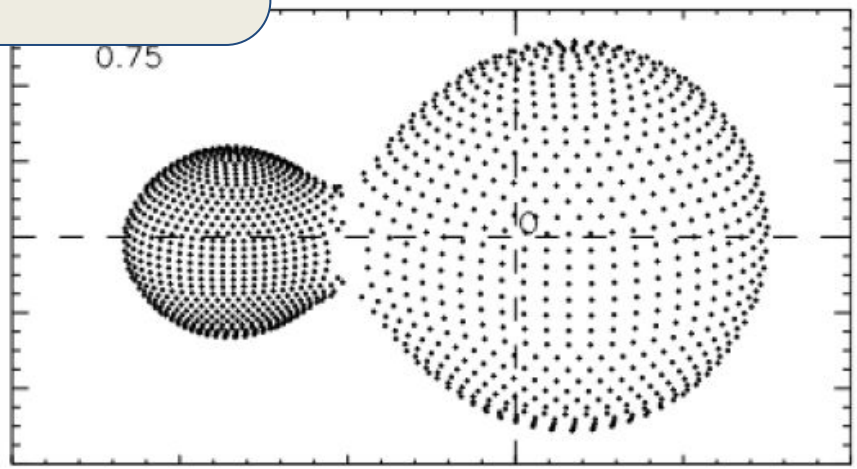


Light curve classification
What else could we expect?

Contact systems: V-shaped LCs



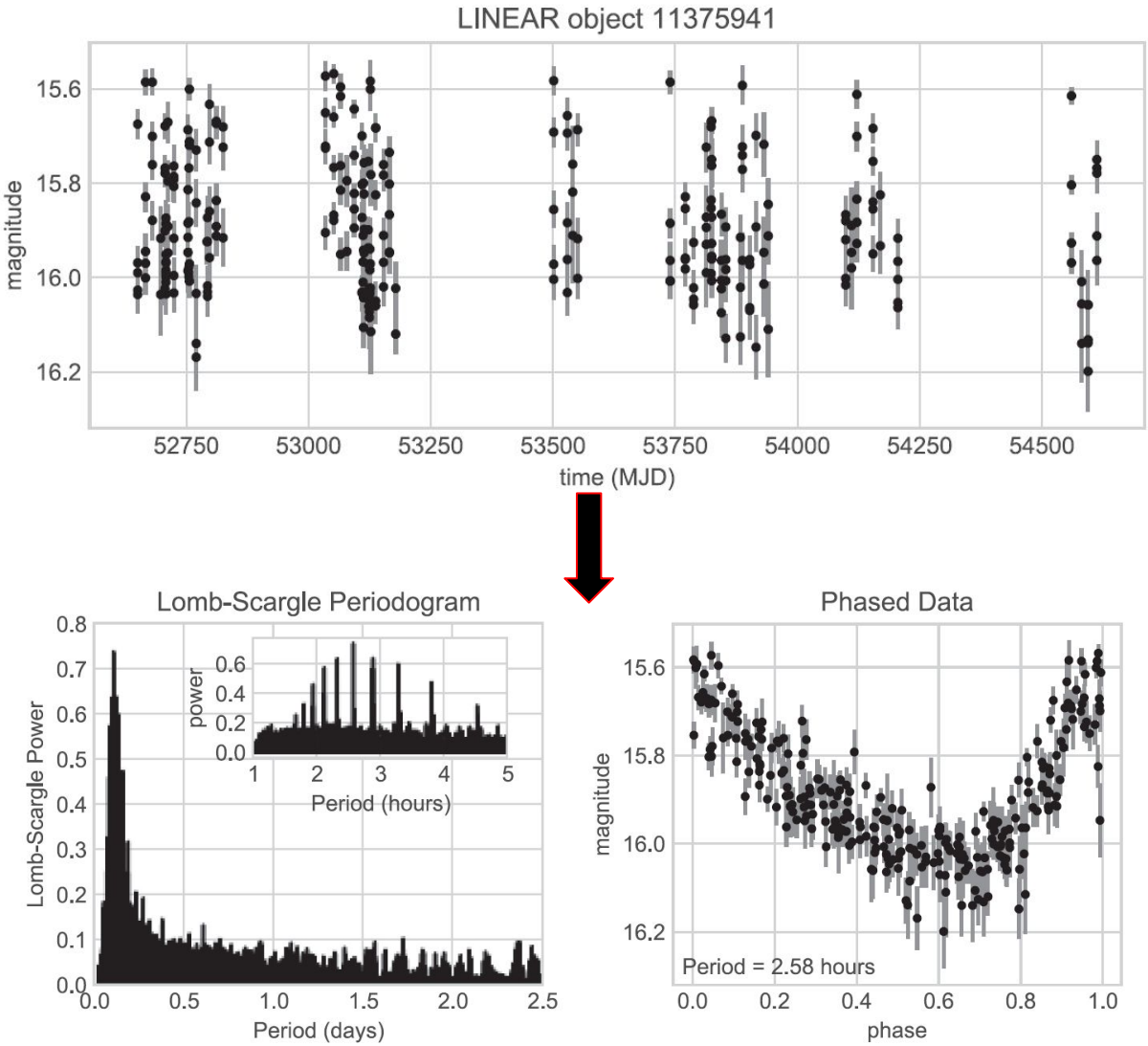
Good luck!



EXTRAS

Light curve classification

Lomb-Scargle Periodogram



https://doi.org/10.3847/1538-4365/aab766

ADQL queries

- ADQL = Astronomical Data Query Language
- Useful tutorial http://docs.g-vo.org/adql-gaia/html/
- A dialect of SQL

Very basic summary of a query:

```
SELECT [TOP (number of rows)] [source table index].(variables you need)
FROM (table you're querying) [AS (table index)]
[WHERE (condition 1) AND (condition 2) OR (condition 3)]
[ORDER BY (variable)]
```

ADQL queries – SELECT: ORDER BY

- Useful to select brightest, fastest, etc. from a table
- E.g.: 50 brightest stars in Gaia DR2

- E.g.: 20 highest proper motion stars in Tycho

ADQL queries – SELECT: ORDER BY

- Useful to select brightest, fastest, etc. from a table
- E.g.: 50 brightest stars in Gaia DR2

```
SELECT TOP 50 source_id, phot_g_mean_mag, parallax, bp_rp
FROM gaiadr2.gaia_source
ORDER BY phot_g_mean_mag
```

- E.g.: 20 highest proper motion stars in Tycho

```
SELECT TOP 20 source_id, parallax, phot_g_mean_mag,
              SQRT(POWER(pmra,2)+POWER(pmdec,2)) AS pm
FROM gaiadr1.tgas_source
ORDER BY pm DESC
```

ADQL queries – SELECT: WHERE clause

- WHERE introduces a logical expression, in a similar way to other languages, with operators AND and OR.
- E.g.: stars brighter than 12, closer than 50 pc.

ADQL queries – SELECT: WHERE clause

- WHERE introduces a logical expression, in a similar way to other languages, with operators AND and OR.
- E.g.: stars brighter than 12, closer than 50 pc.

```
SELECT source_id, phot_g_mean_mag, parallax, bp_rp
FROM gaiadr2.gaia_source
WHERE phot_g_mean_mag < 12.0 AND parallax > 20.0
```

ADQL queries – SELECT: JOIN USING

- For joining two tables with a same column
- E.g.: get Gaia DR2 proper motions for stars with known `source_id`

ADQL queries – SELECT: JOIN USING

- For joining two tables with a same column
- E.g.: get Gaia DR2 proper motions for stars with known `source_id`

```
SELECT source_id, a.phot_g_mean_mag, a.parallax,
       a.bp_rp, b.pmra, b.pmdec
FROM TAP_UPLOAD.t6 AS a
JOIN gaiadr2.gaia_source AS b USING(source_id)
```

ADQL queries – Geometries

- Useful for searching a radius around given coordinates
- E.g.: get Gaia DR2 proper motions for stars with *unknown* `source_id` (3" search)

ADQL queries – Geometries

- Useful for searching a radius around given coordinates
- E.g.: get Gaia DR2 proper motions for stars with *unknown* `source_id` (3" search)

```
SELECT b.source_id, a.NAME_SDCAT, b.pmra, b.pmdec
FROM TAP_UPLOAD.t10 AS a
JOIN gaiadr2.gaia_source AS b ON 1=CONTAINS (
     POINT('ICRS', a.RAJ2000, a.DEJ2000),
     CIRCLE('ICRS', b.ra, b.dec, 3./3600.))
```

- Note: same thing could be done with a TOPCAT crossmatch, but that is not always the case (e.g. if a table is not listed for crossmatching).

Exercise: Pleiades

(Credit: Niall Deacon, Hawaii)

From the tutorial at
http://andromeda.star.bris.ac.uk/topcat/tutorial/

- Open the VizieR load dialog () **(click on "VO" at the top bar menu)**
- Search for all the objects within 3 degrees of the Pleiades in the Tycho-2 catalogue:
  - Check **Cone Selection** button
  - Object name `Pleiades`, **Resolve**
  - Radius 3 degrees
  - Catalogue Selection **Surveys** tab
  - Click on row Tycho-2 (Name column is ordered alphabetically)
  - Click **OK**
  - Loads 2 tables (2 tables in VizieR under that heading) - pick the one with most rows
- Visualise proper motions:
  - Open a scatter plot window 
  - X = pmRA, Y = pmDE
  - Zoom in to find a cluster with non-zero motion
  - Draw a blob round it to create a new subset (click  ; drag out the cluster region, click  again)
- Draw colour-magnitude diagram:
  - Open a different scatter plot window 
  - X = VTmag - BTmag, Y = VTmag, flip Y
  - See where the new cluster subset you identified sit in colour-magnitude space (main sequence?).
- Save the cluster identification:
  - Go to the Subsets window 
  - Select the row corresponding to the cluster subset
  - Create a new boolean table corresponding to this subset by clicking the **To Column**  toolbar button
  - Save the table.

Creating our target list - STEP 3

➡ TAP service to query

- Write an ADQL query in *Gaia* recovering stars within your cut, also using quality control parameters:

Selects all targets that fulfill the criteria in Gaia

```
select *, (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) as abs_g_mag, (4.74/parallax*pm) as vt
from gaiadr3.gaia_source
where parallax_over_error > 5
and parallax > 0
and bp_rp < 0.53
and bp_rp > -0.1
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) < 2.5 - 2.77*bp_rp
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) > -1
and astrometric_sigma5d_max < 1.5 and 4.74/parallax*pm > 145
```

If TOPCAT ‘times out’, go directly to the Gaia archive: https://gea.esac.esa.int/archive/

Creating our target list - STEP 3

➡ ADQL commands

- Write an ADQL query in *Gaia* recovering stars within your cut, also using quality control parameters:

```
select *, (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) as abs_g_mag, (4.74/parallax*pm) as vt
from gaiadr3.gaia_source
where parallax_over_error > 5
and parallax > 0
and bp_rp < 0.53
and bp_rp > -0.1
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) < 2.5 - 2.77*bp_rp
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) > -1
and astrometric_sigma5d_max < 1.5 and 4.74/parallax*pm > 145
```

If TOPCAT ‘times out’, go directly to the Gaia archive: https://gea.esac.esa.int/archive/

Creating our target list - STEP 3

Gaia parameters. See the documentation for details:
13.1.1 gaia_source‣ 13.1 Main tables ‣ Chapter 13 Datamodel description ‣ Part IV Gaia archive ‣ Gaia Early Data Release 3 Documentation release 1.1 (esa.int)

- Write an ADQL query in *Gaia* recovering stars within your cut, also using quality control parameters:

```
select *, (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) as abs_g_mag, (4.74/parallax*pm) as vt
from gaiadr3.gaia_source
where parallax_over_error > 5
and parallax > 0
and bp_rp < 0.53
and bp_rp > -0.1
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) < 2.5 - 2.77*bp_rp
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) > -1
and astrometric_sigma5d_max < 1.5 and 4.74/parallax*pm > 145
```

If TOPCAT ‘times out’, go directly to the Gaia archive: https://gea.esac.esa.int/archive/

Creating our target list - STEP 3

➡ Equations can be used and returned as a variable of your choice

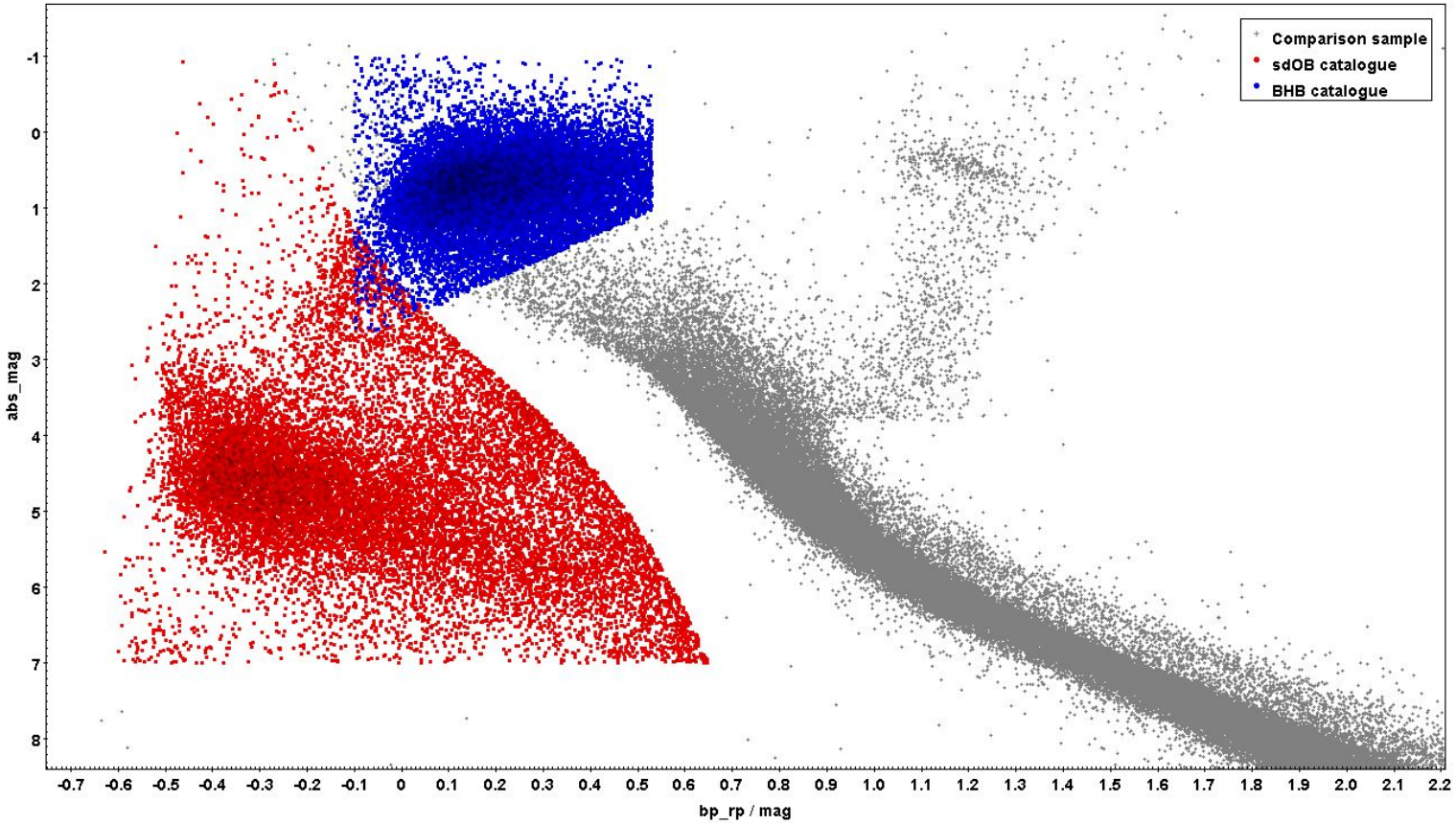
- Write an ADQL query in *Gaia* recovering stars within your cut, also using quality control parameters:

```
select *, (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) as abs_g_mag, (4.74/parallax*pm) as vt
from gaiadr3.gaia_source
where parallax_over_error > 5
and parallax > 0
and bp_rp < 0.53
and bp_rp > -0.1
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) < 2.5 - 2.77*bp_rp
and (phot_g_mean_mag + 5 + 5*log10(parallax/1000)) > -1
and astrometric_sigma5d_max < 1.5 and 4.74/parallax*pm > 145
```

If TOPCAT ‘times out’, go directly to the Gaia archive: https://gea.esac.esa.int/archive/

Current samples of BHBs and sdBs

- Cuts are a bit more elaborate!



- Download the combined catalogues here: http://www.astro.physik.uni-potsdam.de/~hdawson/AstroWorkshop/sdOB_BHB_catalogue_edr3_combined_RUWE.fits