- **Title: Identification and characterization of stellar cluster members using Clusterix and VOSA.**
- **Author: Enrique Solano. Spanish Virtual Observatory.**
- **Last update: 2020 Mar 1st.**

- **VO-Tools:** Clusterix, TOPCAT, VOSA

- **Scientific background:** Open clusters (OCs) are coeval groups of stars formed from the same molecular cloud and, thus, having the same age and initial chemical composition. This makes them ideal targets to study the formation and evolution of stellar objects.

  The determination of the mean properties of open clusters requires prior knowledge of their members to optimise the costly process of obtaining and reducing high resolution spectroscopic data on a large scale. Hence, a precise identification of the stars that compose a cluster is critical to accurately determine the kinematic and fundamental parameters of the clusters (age, total mass, etc.), which are essential for studies of Galactic dynamics.

  In this tutorial we will work with M67, an old cluster (~3.6 Gyr) at about 900 pc with a near solar metallicity and low reddening. M67 is one of the best studied open clusters, considered a cornerstone of stellar astrophysics and used as a calibrator in many surveys. However there is no study covering a large area in spite of some studies on its corona showing that it is an extended cluster. Gaia Collaboration et al. (2018b) studied an area of 1 deg with G <20 and found 1520 members.

  Detailed information on Clusterix is given in Balaguer-Núñez et al. (2020MNRAS.492.5811B )

- **Using Clusterix to estimate cluster membership probabilities.**

  - Go to Clusterix web server → http://clusterix.cab.inta-csic.es/clusterix/

  - Fill in the fields as shown in Figure 1. Click *"Search".* **DO NOT forget to apply the filter in magnitudes (G < 16 )!!**

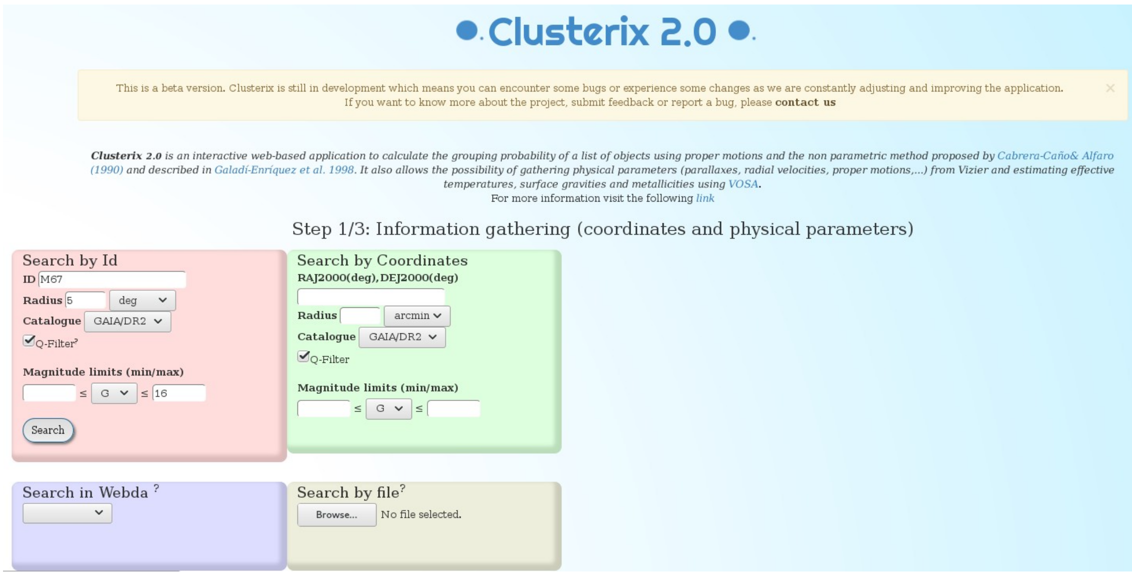


Figure 1

- The result of the query is visualized using Aladin, where the objects obtained from the query are plotted as red diamonds. The individual representation of objects is limited to 40 000. If the result of a query exceeds this amount, then just a yellow circle enclosing the search region is drawn. (Figure 2)

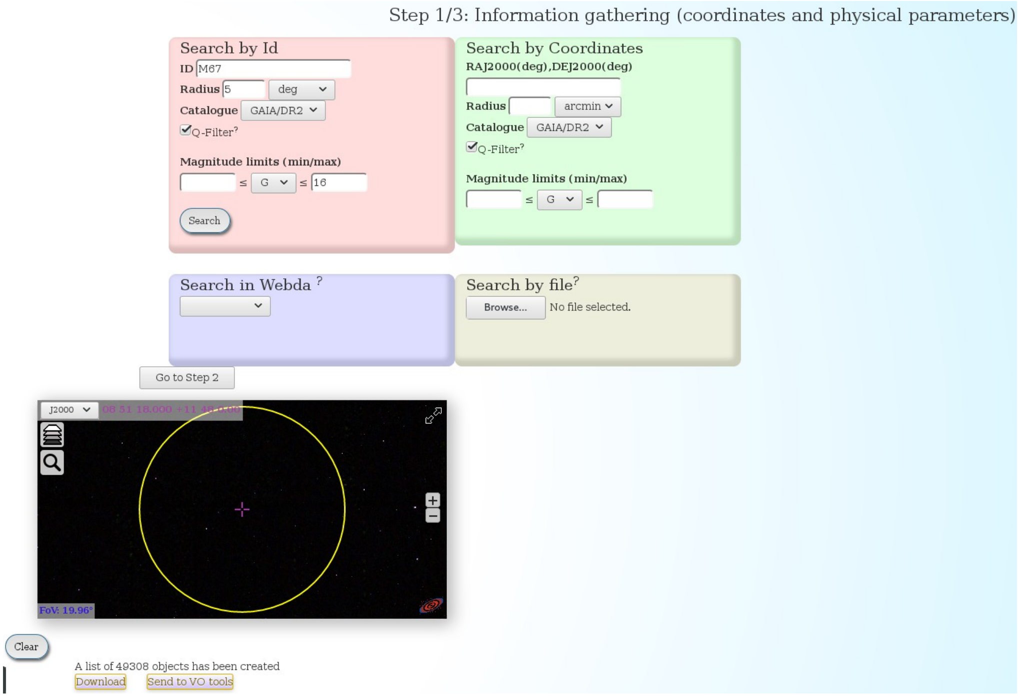



Figure 2

- Click on "Go to Step 2 / Membership from proper motions" to proceed to **Step2.**

  - The second step is to select the *cluster+field* (c+f) and *only field* (f) regions. The definition of these areas is one of the most critical decisions to take by the user, and Clusterix offers several ways to interactively shape and reshape these areas. The simplest option relies on mouse clicks to draw circles that define the *c+f* and *f* regions. The system also includes an easy way to set up a "clean" area around the c+f region to avoid a region that could still have a significant number of cluster members.

    Alternatively, the user can specify the circular areas directly writing their parameters (in decimal degrees) in the corresponding boxes (format: "ra,dec,radius;"). For this tutorial, fill in the fields with the values given in Figure 3.

    Also, in this second step, the user can customize the following parameters:

    - Proper motion limits. Maximum value of the total proper motion (to discard objects that clearly cannot belong to the expected cluster population), and maximum value of the total proper motion error (to remove data of dubious quality).

    - Magnitude range to further limit the selection done in Step 1.

    - Smoothing parameter. Clusterix 2.0 proposes a default value for the smoothing parameter *h.* It represents the radius of the kernel windows used to compute the frequency functions. A large value would blur out the details of the frequency functions, while a small value would yield noisy results. See Balaguer-Núñez's paper for more information on the h parameter.

    - Fine tuning values. To avoid meaningless probability values, Clusterix 2.0 restricts the probability calculations to stars with densities $\gamma$ times above the noise.

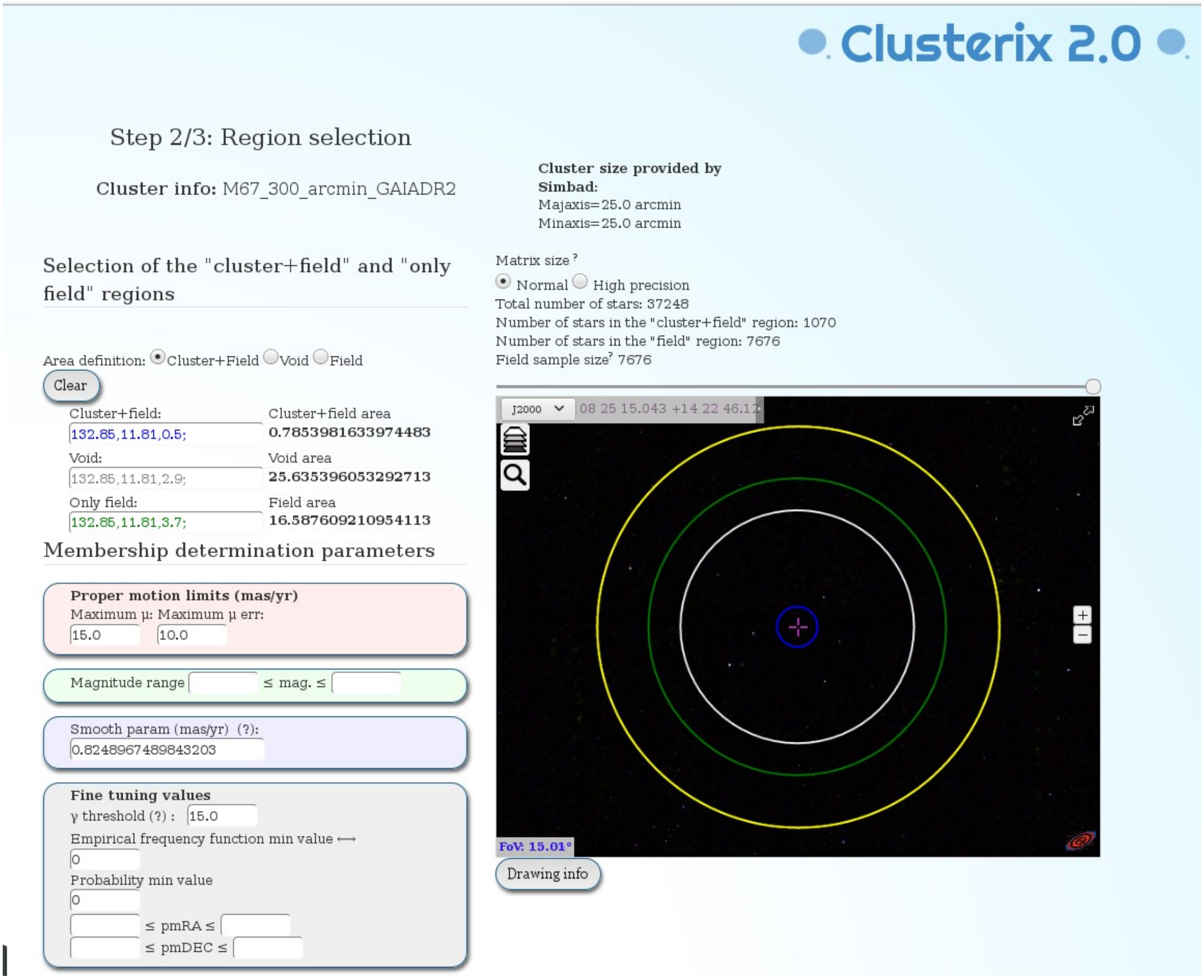



Figure 3.

- The frequency functions computed using the parameters given in Figure 3 are shown in Figure 4.

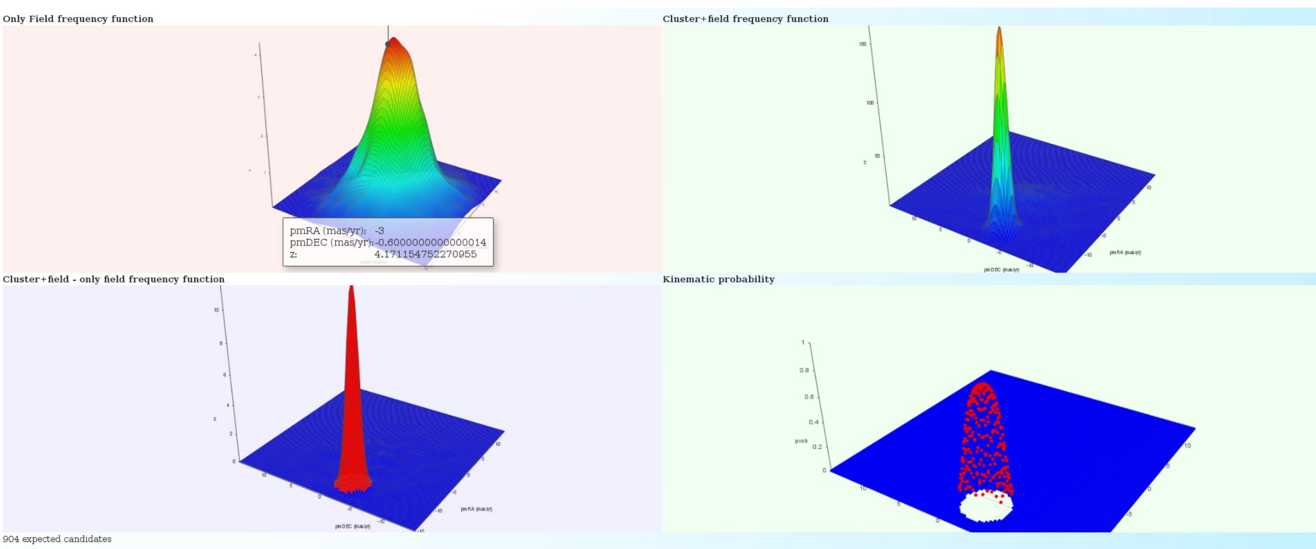
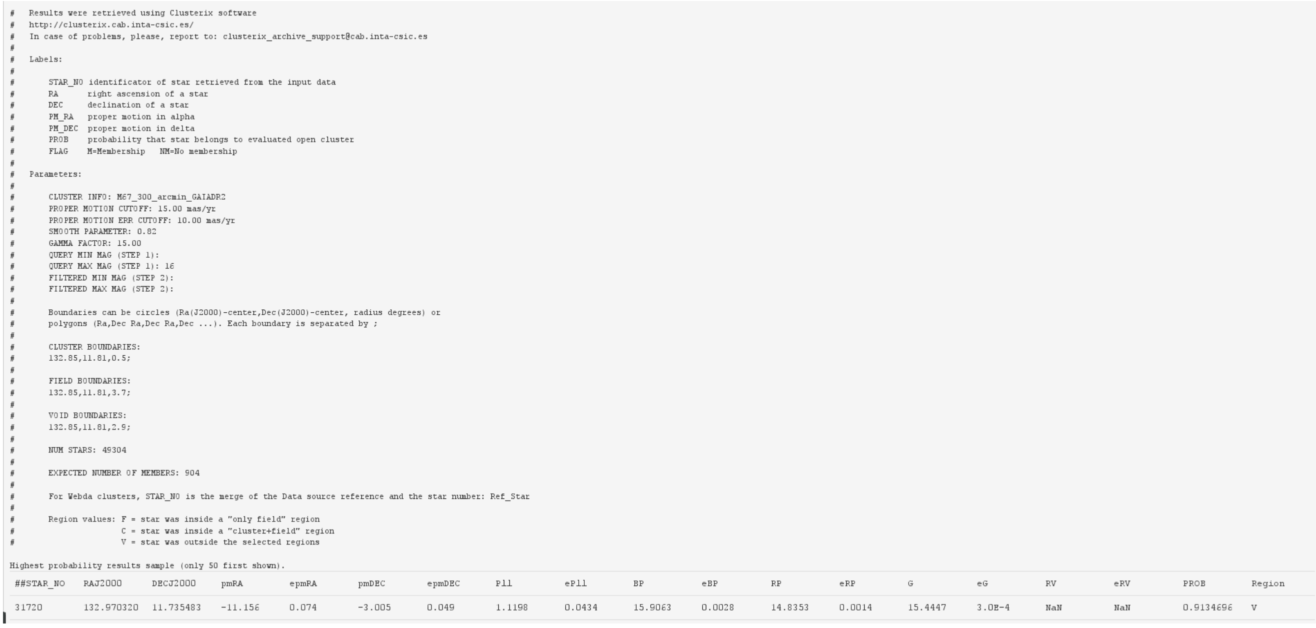



Figure 4

- Once the frequency functions have been computed you can go to Step 3 by clicking "Go to Step 3". Here you can find a summary of the main parameters used in the query as well as the list of cluster members.Those with the highest probabilities are ranked first. (Figure 5).



```
#   Results were retrieved using Clusterix software
#   http://clusterix.cab.inta-csic.es/
#   In case of problems, please, report to: clusterix_archive_support@cab.inta-csic.es
#
#   Labels:
#
#       STAR_NO identificator of star retrieved from the input data
#       RA      right ascension of a star
#       DEC     declination of a star
#       PM_RA   proper motion in alpha
#       PM_DEC  proper motion in delta
#       PROB    probability that star belongs to evaluated open cluster
#       FLAG    M=Membership   NM=No membership
#
#   Parameters:
#
#       CLUSTER INFO: M67_300_arcmin_GAIADR2
#       PROPER MOTION CUTOFF: 15.00 mas/yr
#       PROPER MOTION ERR CUTOFF: 10.00 mas/yr
#       SMOOTH PARAMETER: 0.82
#       GAMMA FACTOR: 15.00
#       QUERY MIN MAG (STEP 1):
#       QUERY MAX MAG (STEP 1): 16
#       FILTERED MIN MAG (STEP 2):
#       FILTERED MAX MAG (STEP 2):
#
#       Boundaries can be circles (Ra(J2000)-center,Dec(J2000)-center, radius degrees) or
#       polygons (Ra,Dec Ra,Dec Ra,Dec ...). Each boundary is separated by ;
#
#       CLUSTER BOUNDARIES:
#       132.85,11.81,0.5;
#
#       FIELD BOUNDARIES:
#       132.85,11.81,3.7;
#
#       VOID BOUNDARIES:
#       132.85,11.81,2.9;
#
#       NUM STARS: 49304
#
#       EXPECTED NUMBER OF MEMBERS: 904
#
#       For Webda clusters, STAR_NO is the merge of the Data source reference and the star number: Ref_Star
#
#       Region values: F = star was inside a "only field" region
#                      C = star was inside a "cluster+field" region
#                      V = star was outside the selected regions

Highest probability results sample (only 50 first shown).
```

| ##STAR_NO | RAJ2000 | DECJ2000 | pmRA | epmRA | pmDEC | epmDEC | Pll | ePll | BP | eBP | RP | eRP | G | eG | RV | eRV | PROB | Region |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 31720 | 132.970320 | 11.735483 | -11.156 | 0.074 | -3.005 | 0.049 | 1.1198 | 0.0434 | 15.9063 | 0.0028 | 14.8353 | 0.0014 | 15.4447 | 3.0E-4 | NaN | NaN | 0.9134696 | V |

Figure 5.

- At the top left of this page there is a button "Send to VO tools". Click here to send the table with the membership probabilities to TOPCAT (NOTE: TOPCAT must be previously open. SAMP broadcasting requires user authorization. A new window ("SAMP Hub Security") may pop up asking for authorization. If so, click "Yes".). After this, a new table with 49 304 rows will be created.

- **Using TOPCAT to visualize the results**
  - Once the table has been uploaded, we can visualize the results in TOPCAT by doing this:
    - In the TOPCAT main window → *Graphics / Plane Plot.* A new window *"Plane Plot"* will pop up.
    - Select the columns to be plotted (X: RA_PM; Y: DEC_PM). Use the mouse to center / zoom in / zoom out the graphic. You will clearly see the M67 overdensity in the proper motion parameter space (Figure 6).

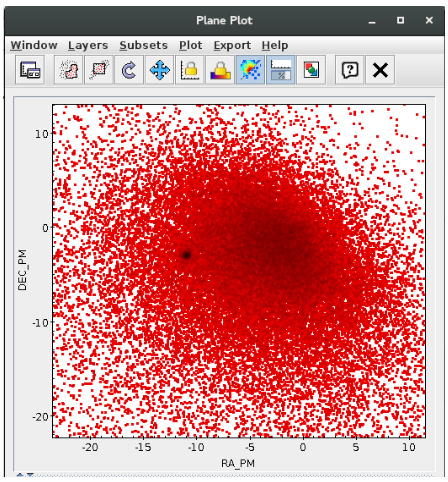



Figure 6.

- Let's now select the M67 members according the probabilities calculated by Clusterix.

  - In the TOPCAT main window → *Views / Row Subsets*. A new window *“Row Subsets”* appears.

  - In the Row Subsets window → *Subsets / New subset*. A new window *“Define Row Subset”* pops up.

  - In the Define Row Subset window → *Subset name: filt; Expression:prob>0.8*. Click “OK”. (Figure 7). 1080 sources will be selected.

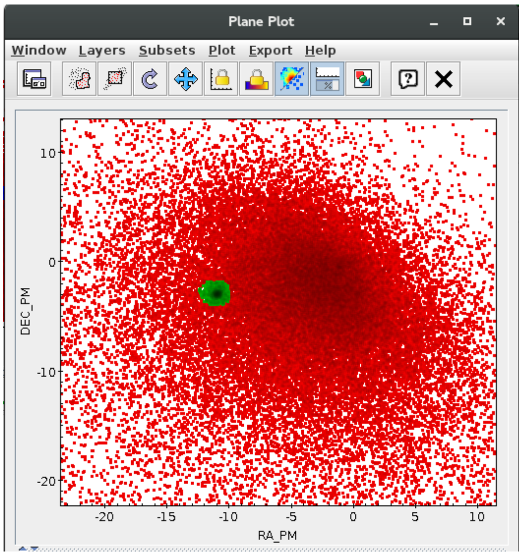



Figure 7

  - Alternatively, you can build an histogram with the probability values (TOPCAT main window → *Graphics / Histogram Plot.* A new window *“Histogram Plot*” appears → Select the column to be plotted (X: Prob) and use it to define your probability cut (Figure 8).

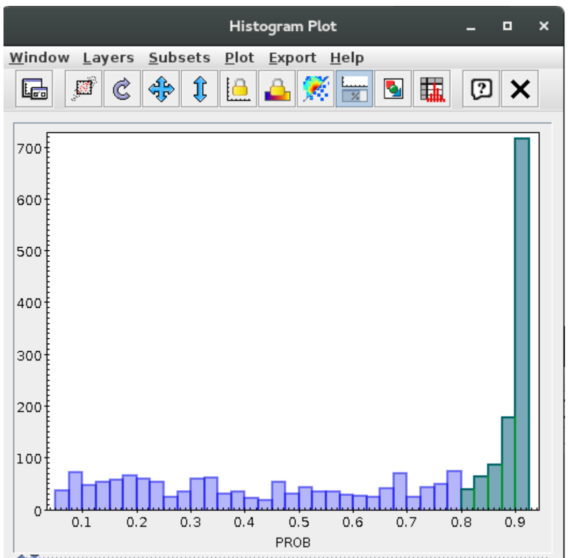



Figure 8

  - Let's build now a colour-magnitude diagram with all the sources in the selected field.

- In the TOPCAT main window → *Graphics / Plane Plot.* A new window *"Plane Plot"* will pop up.

•

- Select the columns to be plotted (X: BP-RP; Y: Gmag).
- Click Axes → Tick "Y Flip". (Figure 9). If the selected objects (in blue) do not appear, click on *Subset tab* (it is between "Position" and "Form", see Figure 9) and tick *"filt".*

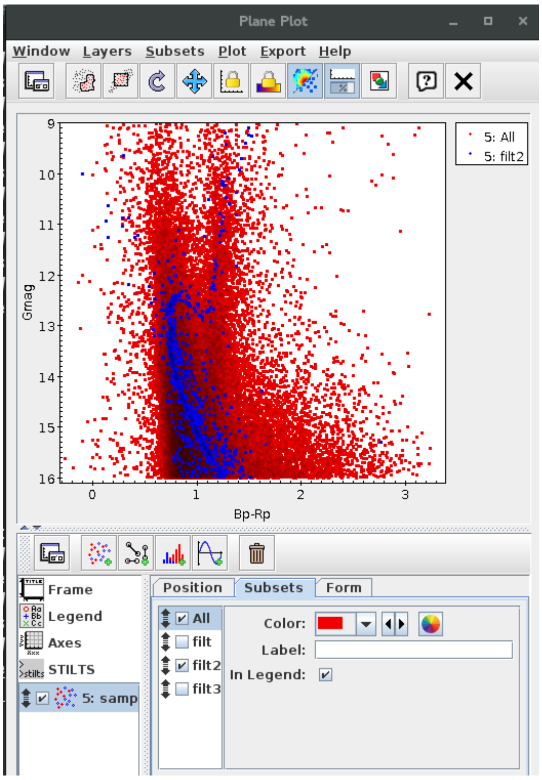



Figure 9.

- The membership selection made by Clusterix is based on proper motions only. Additionally, you can use the information on parallaxes to refine the selection.
  - In the TOPCAT main window → Graphics / Histogram Plot (X: PLL). In the "Histogram Plot" window, tag "Subset", deselect "All" and tick "filt" (Figure 10).
  - Click "Axes" → "Range" → Minimum X: 0.0 / Maximum X: 2.0 to get a zoomed view of the histogram shown in Figure 10.

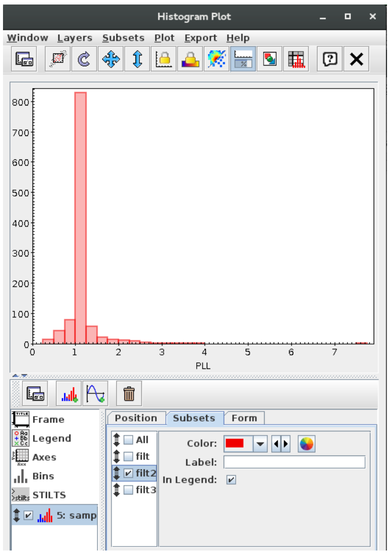



Figure 10

- To perform a sigma clipping in the filtered data (blue subset) we can do the following.
  - In the TOPCAT main window, hit Σ
  - In the *"Row Statistics"* window → Subset for calculations: **filt.** The value for the mean value and standard deviation for the parallax should be something like 1.16 and 0.36, respectively
  - In the TOPCAT main window → *View / Row Subsets*. A new window *"Row Subsets"* appears.
    - In the Row Subsets window → *Subsets / New subset*. A new window *"Define Row Subset"* pops up.
    - In the Define Row Subset window → *Subset name: filt2; Expression:filt&&pll>1.16-0.36&&pll<1.16+0.36* Click "OK". 952 sources will be selected. The cluster sequence is now refined in the colour-magnitude diagram (Figure 11).

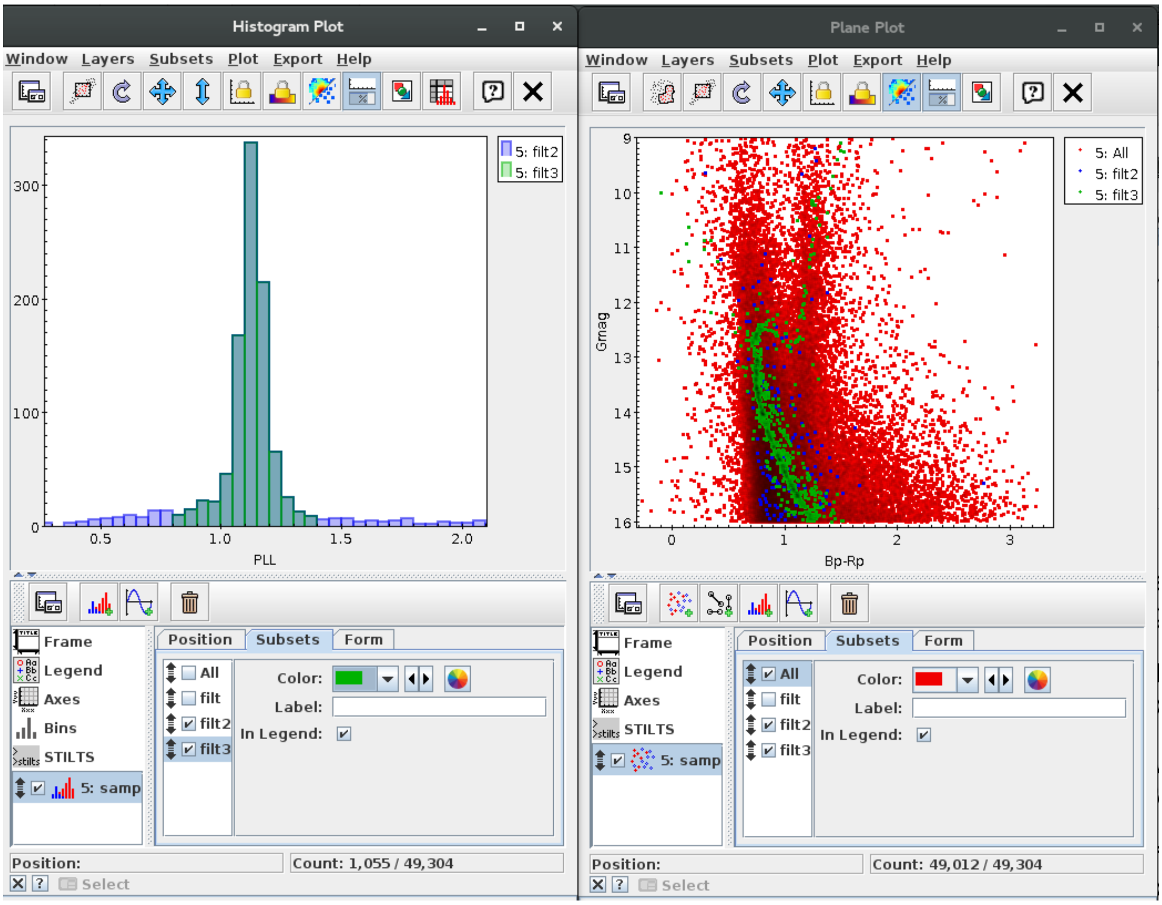



Figure 11

Now we have identified the cluster members, we can do useful thing with them.

- **Distance to the cluster**
  - Distances in parsecs are the reciprocal of parallaxes in arcsec. However, inverting parallaxes to get distances is problematic if parallax errors are large, say > 20%
  - To select sources with parallax error < 20%
    - In the TOPCAT main window → *View / Row Subsets*. A new window *"Row Subsets"* appears.
      - In the Row Subsets window → *Subsets / New subset*. A new window *"Define Row Subset"* pops up.
      - In the Define Row Subset window → *Subset name: filt3; Expression:filt2&&epll/pll<0.2* Click "OK". 952 sources will be selected (in this case, all sources selected in the previous step have relative errors in parallax < 20%.
    - Create a new column with the distances derived from parallaxes.
      - In the TOPCAT main window → **Views / Column Info.** A new window *"Table Columns"* appears
      - In the Table Columns window → **Columns / New Synthetic column.** A new window "Define Synthetic Column" appear.
        - **Name:** *distance*
        - **Expression:** *1000/pll*
        - **Units:** *pc*
      - Σ Using the Statistics button, will get a mean value for the distance (Subset for calculations: filt3) of 892 pc, quite close to that given in the Introduction (900 pc).

- **Visualizing proper motions**
  - Let's visualise proper motions of the *filt3* subset (Figure 12).
    - In the TOPCAT main window: **Graphics / Sky Plot**. A new window *“Sky plot”* will appear.
    - In the “Sky Plot” window, → Form → Hit 

    - *“Add sky vector”*
      - **Delta Longitude:** RA_PM
      - **Delta Latitude: DEC_PM**

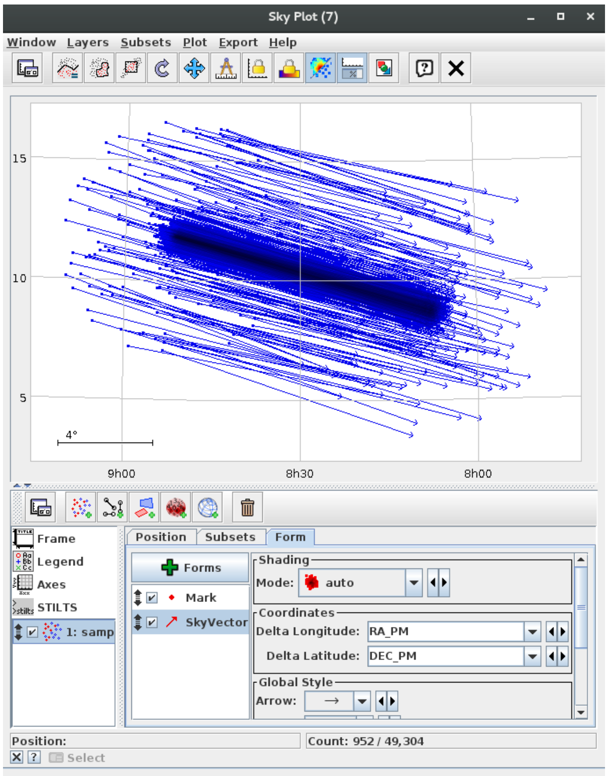



Figure 12

- **Visualizing the cluster members on the sky**
  - Open Aladin (java -jar Aladin.jar)
  - In Aladin, write *M67* in the **Command** box. A M67 DSS coloured image will be loaded.
  - In the TOPCAT man window
    - Row subset: **filt3**
    - Interop / Send table to… Aladin → a new plane in Aladin with the 952 sources will appear.
    - We can see how, thanks to the excellent quality of Gaia DR2 data and VO tools, we have been able to identify cluster members really far away from the cluster center. (Figure 13)

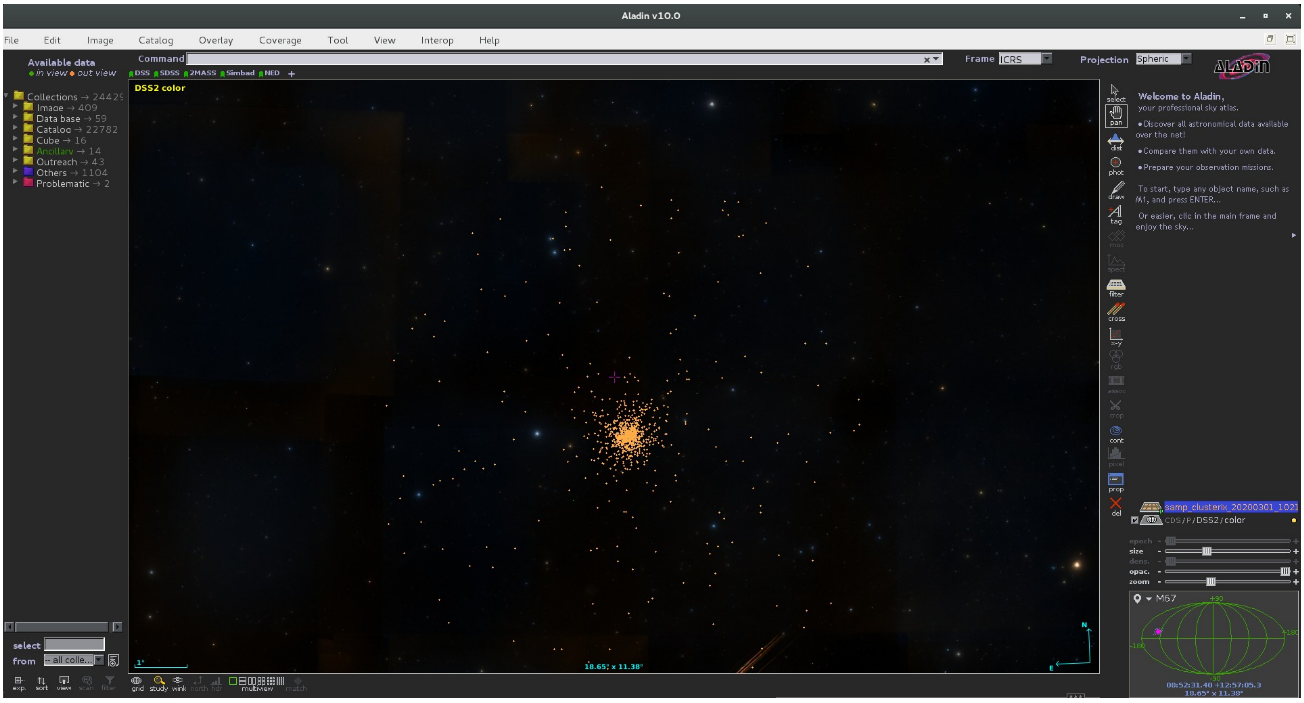



Figure 13

- **Estimating physical parameters**

  - The next step is to use VOSA to estimate effective temperatures, luminosities and radii for a subset of these objects.

- Step 1.- Go to [http://svo2.cab.inta-csic.es/theory/vosa/](http://svo2.cab.inta-csic.es/theory/vosa/)

- Step 2.- To use VOSA you need to be registered. Click on “Register” and fill in the fields (email, name and passwd).

- Tag "Files"

  - Step 3.- Upload the vosa_granada2020*.txt* file (available from the web page of the school) in VOSA (“File to upload”). This file is a subset (2 objects, 3 columns – Id, RA, Dec – ) of the list of objects obtained in Clusterix Step 3. We will work with this subset in order to speed up the workflow.

    Give a description (free text). And then, click “Upload” (do not bother about the File type). The message “*your- file-name has been succesfully uploaded!”* will appear. Click “Continue”. If the message does not appear, go to “your files” section and click “Select”.

- Tag "Objects"

  - Step 4.- Place the cursor on the*“Objects”* tag and then click *“Distances”*. Click *“Search for Obj. Distances”.* To make the Gaia DR2 distances the “final” distances, do the following: Go to the *“Actions for all the objects in the file”* panel, tick *“Select values by ranking”* and choose Gaia2 in the first place. Click *“Make all changes”*. The Gaia DR2 coordinates will appear in bold in the “Final” column. (Figure 14)

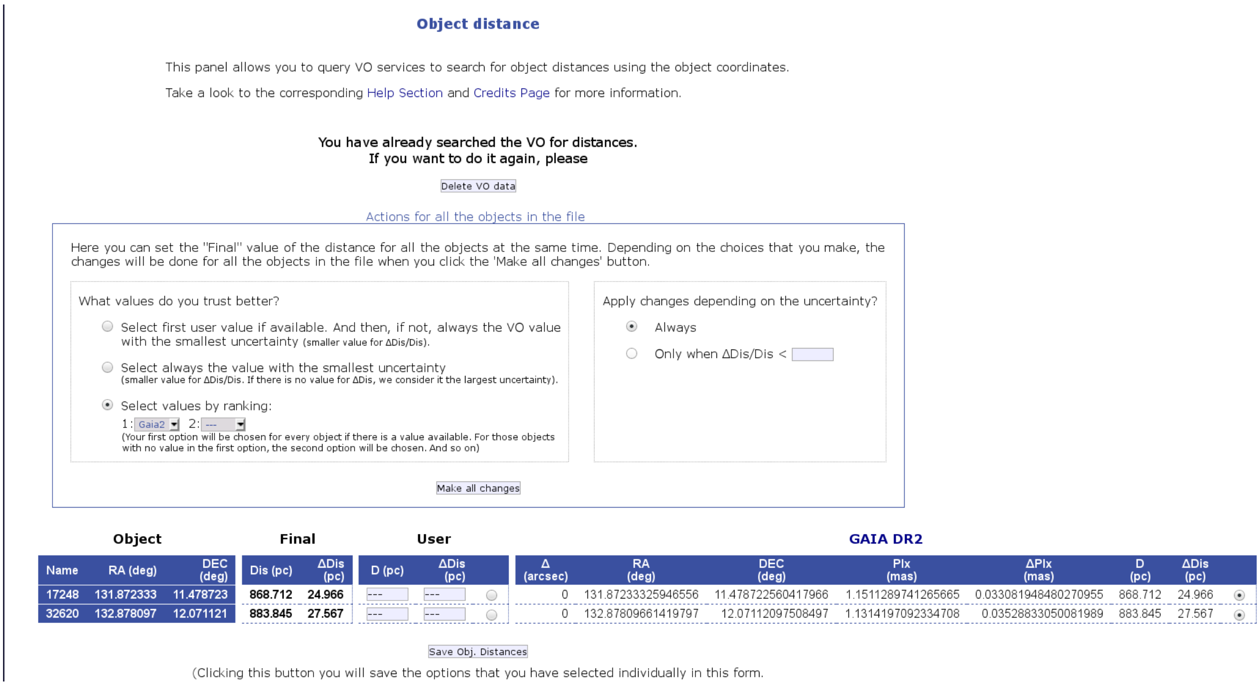



Figure 14.

- Tag "Build SEDs"

  - Step 5.- Place the cursor on the *"Build SEDs" tag* and then click *"VO photometry".* Here we will be able to look for phometric information of our objects in different VO catalogues. In order not to slow down too much the tutorial, click *"unmark All"* and select only 2MASS, WISE and APASS9. Then, click *"Query selected services"* at the bottom of the page. Once this is done, a summary table with the VO photometry (in flux units) will appear (Figure 15).

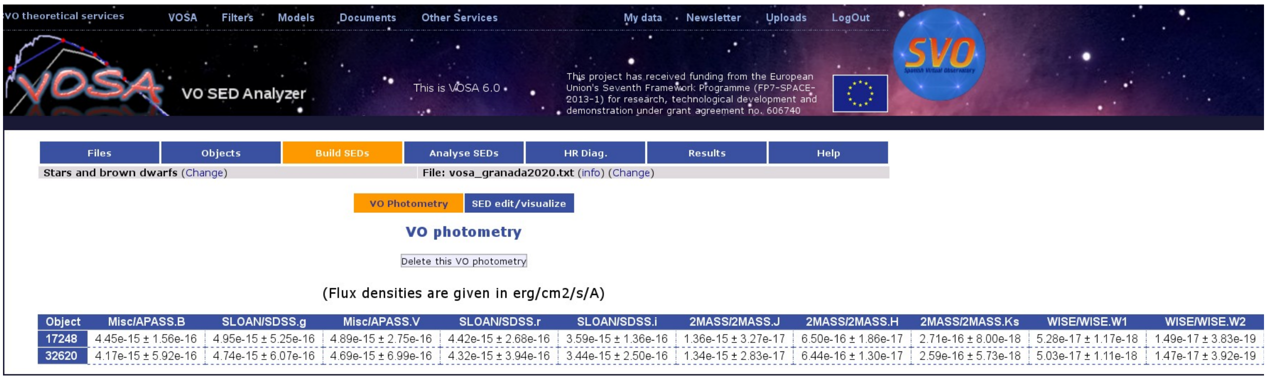
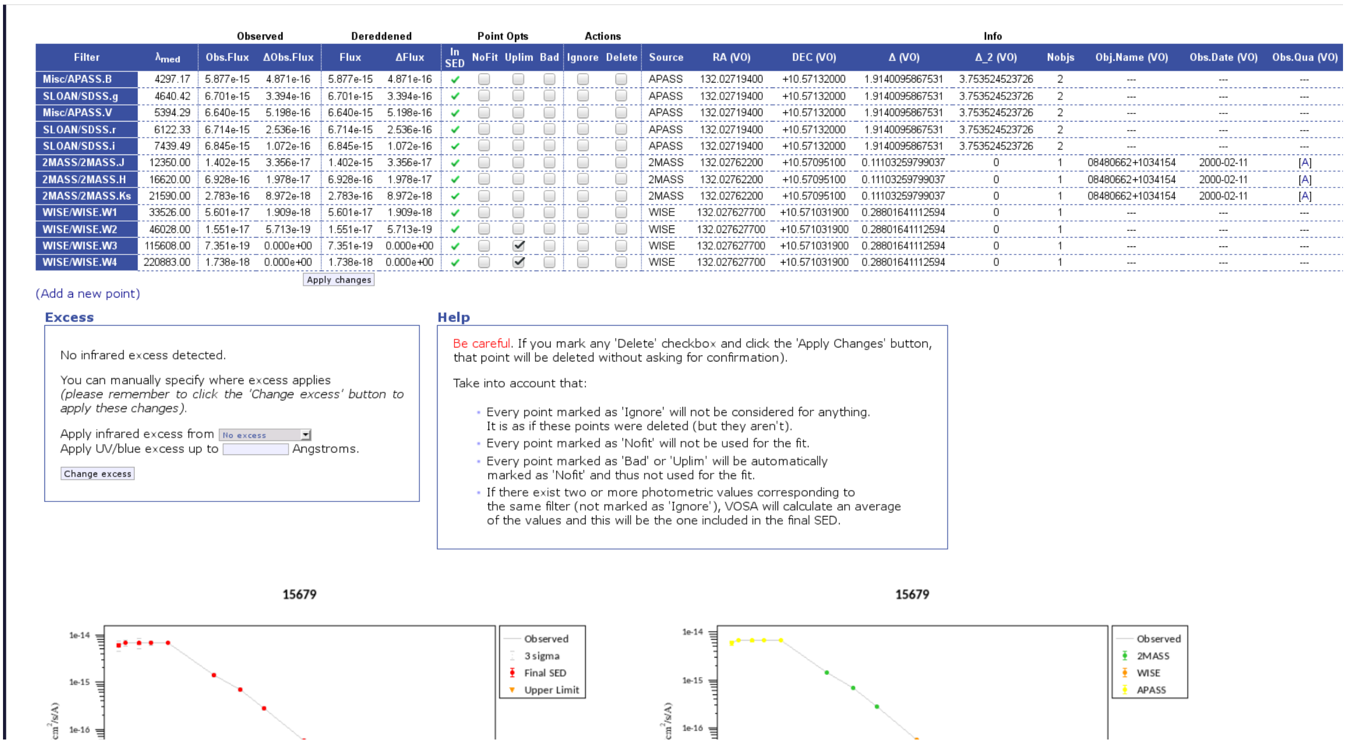



Figure15

  - Step 6.- Place the cursor on the *"Build SEDs" tag* and then click *"SED edit/visualize".* This tag gives us the possibility of visualising/modifying the SED before the model fitting. VOSA gathers from VO services not only the photometric information but also different metadata of interest (Object name, observing date and information on quality). In particular, VOSA uses the information on quality to automatically identify bad photometric points and remove them from the fitting. Upper limits are treated in a similar way (see, for instance, WISE W3 and W4 for object "17248"). The user can manually override this selection of photometric points by ticking/unticking the appropriate boxes. (Figure 16)

    For this use case, do not make any change in the SED edit/visualize section.



| Filter | λmed | Obs.Flux | ΔObs.Flux | Flux | ΔFlux | In SED | NoFit | Uplim | Bad | Ignore | Delete | Source | RA (VO) | DEC (VO) | Δ (VO) | Δ_2 (VO) | Nobjs | Obj.Name (VO) | Obs.Date (VO) | Obs.Qua (VO) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Misc/APASS.B | 4297.17 | 5.877e-15 | 4.871e-16 | 5.877e-15 | 4.871e-16 | ✔ | | | | | | APASS | 132.02719400 | +10.57132000 | 1.9140095867531 | 3.753524523726 | 2 | --- | --- | --- |
| SLOAN/SDSS.g | 4640.42 | 6.701e-15 | 3.394e-16 | 6.701e-15 | 3.394e-16 | ✔ | | | | | | APASS | 132.02719400 | +10.57132000 | 1.9140095867531 | 3.753524523726 | 2 | --- | --- | --- |
| Misc/APASS.V | 5394.29 | 6.640e-15 | 5.198e-16 | 6.640e-15 | 5.198e-16 | ✔ | | | | | | APASS | 132.02719400 | +10.57132000 | 1.9140095867531 | 3.753524523726 | 2 | --- | --- | --- |
| SLOAN/SDSS.r | 6122.33 | 6.714e-15 | 2.536e-16 | 6.714e-15 | 2.536e-16 | ✔ | | | | | | APASS | 132.02719400 | +10.57132000 | 1.9140095867531 | 3.753524523726 | 2 | --- | --- | --- |
| SLOAN/SDSS.i | 7439.49 | 6.845e-15 | 1.072e-16 | 6.845e-15 | 1.072e-16 | ✔ | | | | | | APASS | 132.02719400 | +10.57132000 | 1.9140095867531 | 3.753524523726 | 2 | --- | --- | --- |
| 2MASS/2MASS.J | 12350.00 | 1.402e-15 | 3.356e-17 | 1.402e-15 | 3.356e-17 | ✔ | | | | | | 2MASS | 132.02762200 | +10.57095100 | 0.11103259799037 | 0 | 1 | 08480662+1034154 | 2000-02-11 | [A] |
| 2MASS/2MASS.H | 16620.00 | 6.928e-16 | 1.978e-17 | 6.928e-16 | 1.978e-17 | ✔ | | | | | | 2MASS | 132.02762200 | +10.57095100 | 0.11103259799037 | 0 | 1 | 08480662+1034154 | 2000-02-11 | [A] |
| 2MASS/2MASS.Ks | 21590.00 | 2.783e-16 | 8.972e-18 | 2.783e-16 | 8.972e-18 | ✔ | | | | | | 2MASS | 132.02762200 | +10.57095100 | 0.11103259799037 | 0 | 1 | 08480662+1034154 | 2000-02-11 | [A] |
| WISE/WISE.W1 | 33526.00 | 5.601e-17 | 1.909e-18 | 5.601e-17 | 1.909e-18 | ✔ | | | | | | WISE | 132.027627700 | +10.571031900 | 0.28801641112594 | 0 | 1 | --- | --- | --- |
| WISE/WISE.W2 | 46028.00 | 1.551e-17 | 5.713e-19 | 1.551e-17 | 5.713e-19 | ✔ | | | | | | WISE | 132.027627700 | +10.571031900 | 0.28801641112594 | 0 | 1 | --- | --- | --- |
| WISE/WISE.W3 | 115608.00 | 7.351e-19 | 0.000e+00 | 7.351e-19 | 0.000e+00 | ✔ | | ✔ | | | | WISE | 132.027627700 | +10.571031900 | 0.28801641112594 | 0 | 1 | --- | --- | --- |
| WISE/WISE.W4 | 220883.00 | 1.738e-18 | 0.000e+00 | 1.738e-18 | 0.000e+00 | ✔ | | ✔ | | | | WISE | 132.027627700 | +10.571031900 | 0.28801641112594 | 0 | 1 | --- | --- | --- |



Figure 16

.

○ Tag "Analyse SEDs"

- Step 7: Place the cursor on the *"Analyse SEDs" tag* and then click *"Chi-square fit".* Different grids of theoretical models covering different ranges of physical parameters are displayed. For this tutorial select only the "Kurucz ODFNEW /NOVER models". Click *"Next: Select mode params".* (Figure 17).

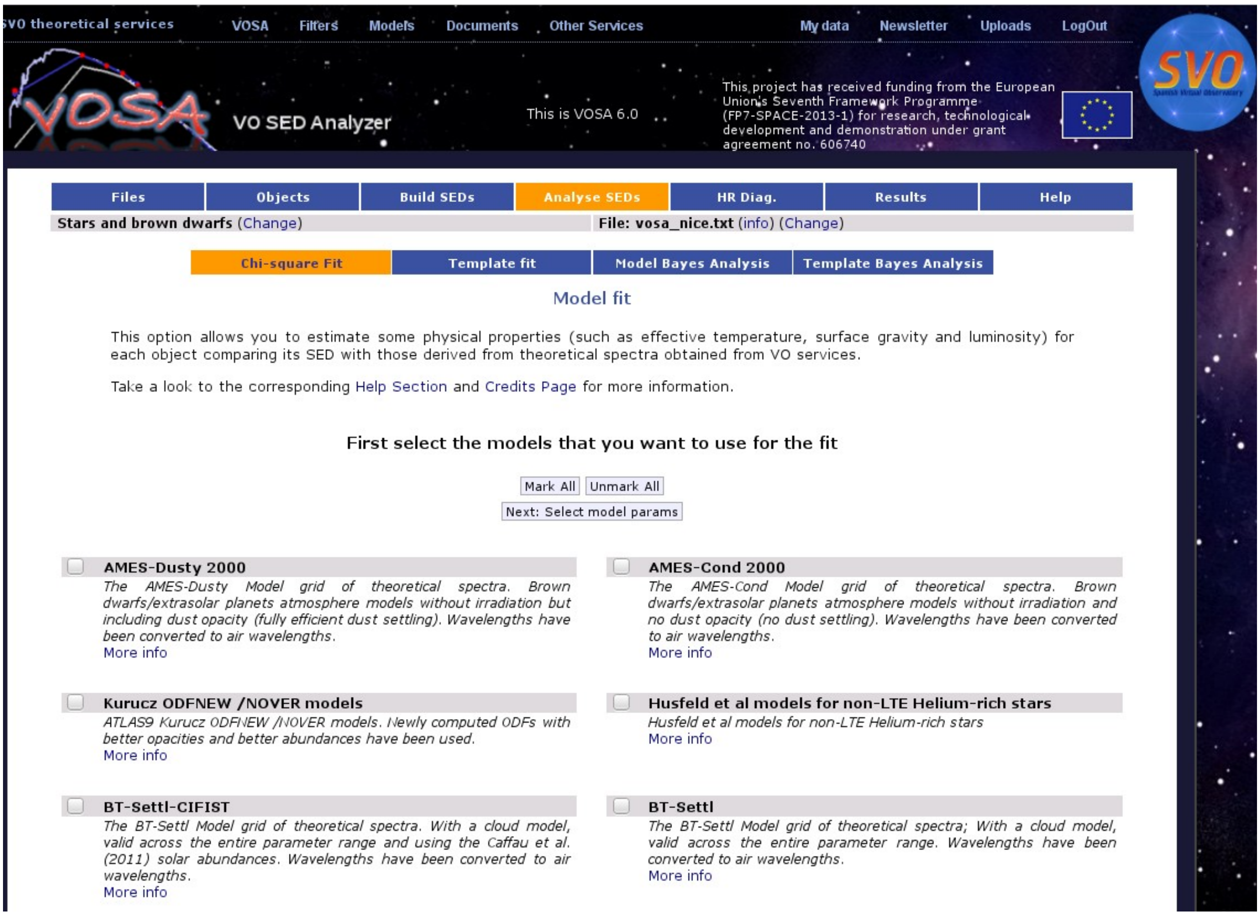
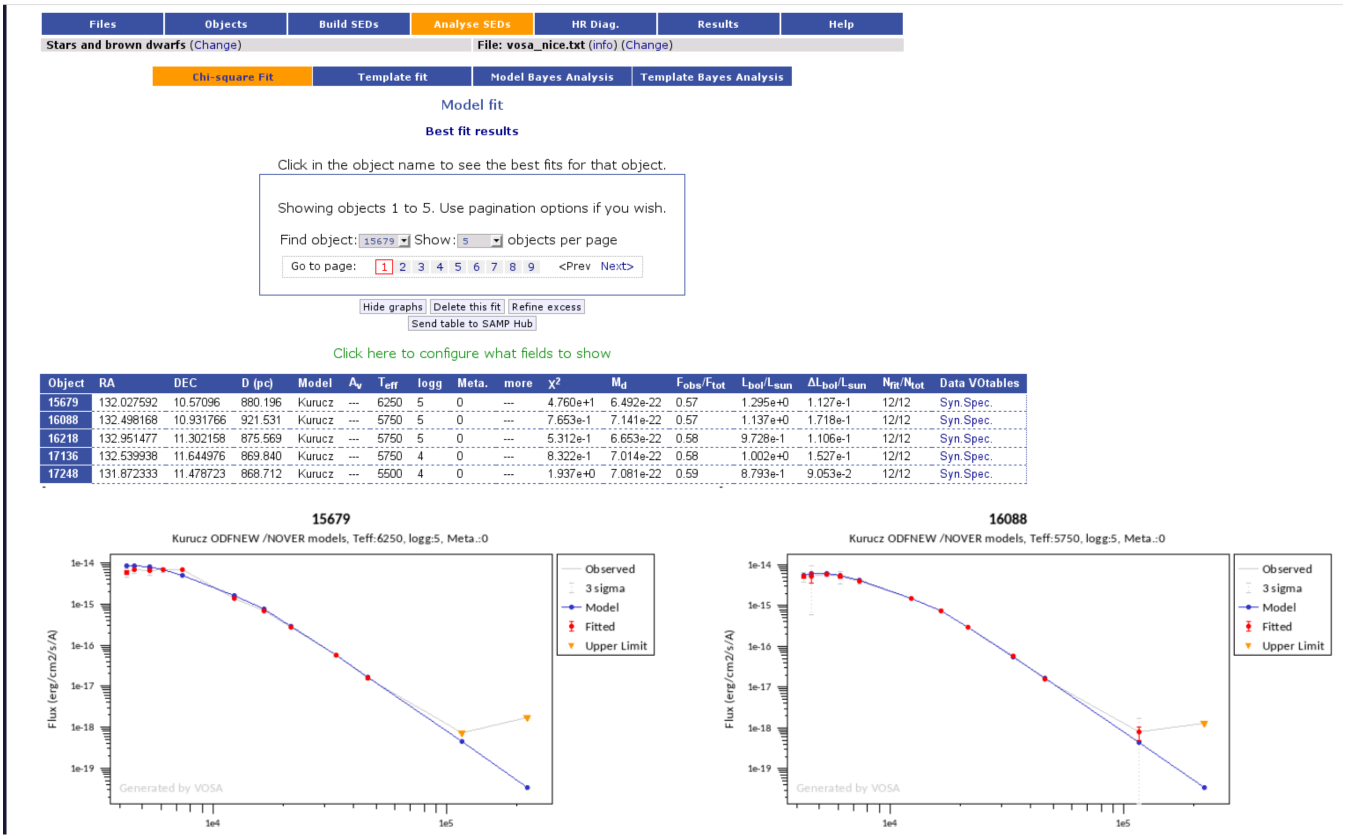


Figure 17

- Step 8.- In this window, we can limit the range of physical parameters that will be used for the fit. To save time we will make the following assumptions:
  - Teff: 4000-8000K
  - logg: 4.0-5.0 dex.
  - [M/H]=0.0

  Then, click *"Make the fit"*

- Step 9.- We will see now a summary table with the best fit results. Click on "Show graphs" to have a look at the graphics. If needed, the table can be sent to TOPCAT using the "Send Table to SAMP hub" button. (Figure 18).




| Object | RA | DEC | D (pc) | Model | $A_v$ | $T_{eff}$ | logg | Meta. | more | $\chi^2$ | $M_d$ | $F_{obs}/F_{tot}$ | $L_{bol}/L_{sun}$ | $\Delta L_{bol}/L_{sun}$ | $N_{fit}/N_{tot}$ | Data VOtables |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15679 | 132.027592 | 10.57096 | 880.196 | Kurucz | --- | 6250 | 5 | 0 | --- | 4.760e+1 | 6.492e-22 | 0.57 | 1.295e+0 | 1.127e-1 | 12/12 | Syn.Spec. |
| 16088 | 132.498168 | 10.931766 | 921.531 | Kurucz | --- | 5750 | 5 | 0 | --- | 7.653e-1 | 7.141e-22 | 0.57 | 1.137e+0 | 1.718e-1 | 12/12 | Syn.Spec. |
| 16218 | 132.951477 | 11.302158 | 875.569 | Kurucz | --- | 5750 | 5 | 0 | --- | 5.312e-1 | 6.653e-22 | 0.58 | 9.728e-1 | 1.106e-1 | 12/12 | Syn.Spec. |
| 17136 | 132.539938 | 11.644976 | 869.840 | Kurucz | --- | 5750 | 4 | 0 | --- | 8.322e-1 | 7.014e-22 | 0.58 | 1.002e+0 | 1.527e-1 | 12/12 | Syn.Spec. |
| 17248 | 131.872333 | 11.478723 | 868.712 | Kurucz | --- | 5500 | 4 | 0 | --- | 1.937e+0 | 7.081e-22 | 0.59 | 8.793e-1 | 9.053e-2 | 12/12 | Syn.Spec. |



Figure 18

To get information on the radii derived from VOSA, click *"Click here to configure what fields to show"* and tick R1 and R2. Click *"Save config."* Radii obtained using two different approaches will now appear in the summary table. More information on how VOSA calculates radii can be found at:
http://svo2.cab.inta-csic.es/theory/vosa/helpw4.phpotype=star&action=help&what=fit#fit:radius

## Tag "HR diagram"

- Step 10.- Let's estimate now masses and ages from isochrones and evolutionary tracks. Click on **"HR diagram".** The Siess collection of isochrones and evolutionary tracks will be shown. Select ages between 1-6 Gy and masses between 0.5-2 solar masses. Click **Make the HR diagram.** A table including the luminosities, masses and ages, as well as the HR diagram will be shown. Click on **Data range** below the HR diagram for a more clear view of the objects (Figure 19). The calculated ages are in agreement with the value given in the Introduction (3.6 Gyr).

## Tag "Results"

- Step 11.- Place the cursor on the **Results** tag and then click **Activity Log**. You will find here a summary of all steps executed to carry out the workflow.

- Step 12.- Place the cursor on the **Results** tag and then click **Download Results**. You can save different type of results (plots, VO photometry, HR diagram,…) in different formats.

- Step 13.- Place the cursor on the **Results** tag and click **Download Results**. Click **Retrieve** (do not tick anythnig). A new window with the message "Tar file successfully created" will appear. By clicking **Download your results** in the tar file (folder info), among other files, you will get a couple of files (refs.dat, refs.bibtex.bib) including the bibliographic references of all the resources you have made use of.

Files | Objects | Build SEDs | Analyse SEDs | HR Diag. | Results | Help

**Stars and brown dwarfs** (Change) **File: vosa_granada** (info) (Change)

HR Diagram

Delete this HR Diagram

Send results table to SAMP Hub

**Objects**

| Object | Model | $T_{eff}$ | | LogL | | Age | | Mass | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 17248 | siess | 5500 | (5375,5625) | -0.0559 | (-0.1031,-0.0133) | 3.2035 | ..(---,3.2035) | 1.0023 | (---,1.0989) | [2] |
| 32620 | siess | 5500 | (5375,5625) | -0.0583 | (-0.1237,-0.0016) | 3.0114 | ..(---,4.9913) | 1.0026 | (---,1.0998) | [2] |

[2] The distance to both the closest curves has been estimated as the one to the closest point in each curve

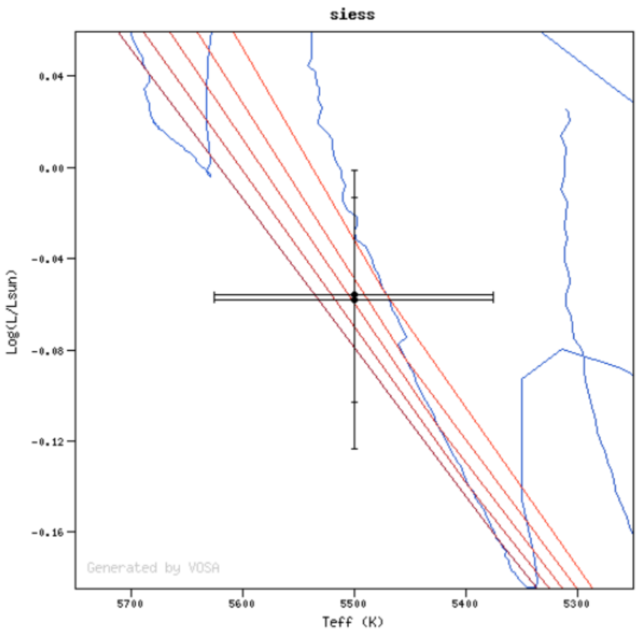



Figure 19