

CARACTERIZACIÓN DE CÚMULOS ESTELARES ABIERTOS CON HERRAMIENTAS DEL OBSERVATORIO VIRTUAL

Astronomía en el Aula con datos
reales y herramientas profesionales



Joaquín Álvaro Contreras



CARACTERIZACIÓN DE CÚMULOS ESTELARES ABIERTOS CON HERRAMIENTAS DEL OBSERVATORIO VIRTUAL

**Astronomía en el Aula con datos
reales y herramientas profesionales**

Joaquín Álvaro Contreras

Publicaciones de ApEA

Asociación para la Enseñanza de la Astronomía

Nº 44 (Secundaria)- Abril 2024

Publicaciones de ApEA nº 44
Asociación para la Enseñanza de la Astronomía

© Texto: Joaquín Álvaro Contreras
© Dibujos e imágenes: Joaquín Álvaro Contreras, NASA.
Dibujo de Portada: Tanguy
Foto de Portada: Las Pléyades. Marta Dueñas Mínguez.

Comité de Redacción:
Ederlinda Viñuales, Antonio Arribas y Ricardo Moreno

Dirección:
Ricardo Moreno (Vocal Editor de Publicaciones de ApEA)
rmluquero@gmail.com

ISBN: 978-84-942304-6-2



Esta obra está licenciada bajo la Licencia Creative Commons Atribución 4.0 Internacional. Para ver una copia de esta licencia, visita <http://creativecommons.org/licenses/by/4.0/>.

EL AUTOR

Joaquín Álvaro Contreras



Es Licenciado en Ciencias Físicas, especialidad de Astrofísica, por la Universidad Complutense de Madrid (UCM). Hizo el programa de Doctorado en Física Teórica y Física Automática y Robótica. Ha sido investigador sobre radiación cósmica en el Instituto Nacional de Técnica Aeroespacial, Grupos de la Comisión Nacional de Investigación del Espacio (INTA-CONIE) en la década de los 80, con Juan Sequeiros y José Medina. Durante 14 años ha sido docente en la UNED, (1985-1999).

Su actividad profesional se ha movido en el campo de Informática, (grandes sistemas). Además, ha mantenido una actividad dilatada como divulgador de la Astronomía y de la Física: cursos para docentes, conferencias, etc. Es socio fundador y presidente actual de AstroCuenca. También tiene experiencia en diseño museístico del área de Astronomía en el Museo de las Ciencias de Castilla La Mancha. Es miembro de la *Sociedad Española de Astronomía* (SEA), de la *European Astronomical Society* (EAS) y de la *Europlanet Society* (EPS).

Actualmente es Presidente de la Federación de Asociaciones Astronómicas de España (FAAE), miembro de la Comisión Pro-Am de la SEA y Miembro del *Regional Hub for Spain and Portugal of the Europlanet Society*.

Es autor de la Publicación nº 43 de ApEA titulada "Exoplanetas minería con datos reales de aplicación en el aula".

ÍNDICE

El autor.....	3
Presentación.....	7
1ª Parte: Obtención de los datos y primeros ejercicios con Excel.....	9
Introducción.....	9
Primeros pasos.....	11
Primeros ejercicios.....	14
Conclusión.....	18
2ª Parte: Caracterización con TOPCAT y ALADIN	19
Primer ejercicio	20
Segundo ejercicio.....	30
3ª Parte: CLUSTERIX.....	31
I Parte	32
II Parte	37

PRESENTACIÓN

Este trabajo tiene el propósito de hacer posible un desarrollo temático en el aula con alumnos de secundaria. Al igual que en la Publicación anterior sobre exoplanetas, se trata de facilitar el acceso a los grandes conjuntos de datos disponibles en Astronomía, al tiempo que se introduce al alumnado en los conocimientos astronómicos del tema tratado y, también, en los rudimentos del método científico de la mano de las herramientas que usan habitualmente los astrónomos profesionales (*).

El objetivo es acercarnos a las tareas de caracterización de Cúmulos abiertos (OCs), es decir, determinar qué estrellas de todas las que aparecen en el campo visual de observación forman parte de un cúmulo y, a partir de ahí, definir sus parámetros generales: masa, edad, distancia, propiedades dinámicas, etc.

En la primera parte de esta Publicación, se explicará cómo obtener de internet los datos reales y se propondrán ejercicios sencillos para trabajar con ellos con Excel. En la segunda parte se dará un paso más, al utilizar las herramientas del Observatorio Virtual, TOPCAT y Aladin. En la tercera y última parte, usaremos la herramienta Clusterix.

(*) Nuestro agradecimiento al Spanish Virtual Observatory (SVO), responsable de las herramientas que hacen posible el desarrollo de estos trabajos.

1ª PARTE

OBTENCIÓN DE LOS DATOS Y PRIMEROS ANÁLISIS CON EXCEL



Messier 25. Créditos & Copyright: Jean-Charles Cuillyre (CFHT) & Giovanni Anselmi (Coelum Astronomia), Hawaiian Starlight.



Messier 67 (NGC 2682). Créditos: Sloan Digital Sky Survey (SDSS).

Introducción

Los Cúmulos estelares abiertos, (en adelante 'OCs' –Open Clusters–), son grupos de estrellas unidas gravitacionalmente que se formaron a partir de una misma nube de gas molecular, por lo que comparten la misma composición química y edad, y tienen además posiciones relativas muy próximas y movimientos propios similares. Los OCs son objetos fundamentales para comprender la estructura espiral, la dinámica y la evolución química de la galaxia, al tiempo que también ayudan a comprender los procesos de formación estelar.

Aunque la mayoría de las estrellas de la Vía Láctea se presentan aisladas, se considera que en su mayor parte, tal vez todas, se forman en entornos agrupados y pasan un periodo de tiempo unidas gravitacionalmente entre sí incrustadas en la nube progenitora, (Clarke *et al.* 2000; Portegies Zwart *et al.* 2010). La evolución de estos sistemas tiende a disgregarlos en pocos millones de años debido a interacciones gravitatorias con otros sistemas, a fuerzas de marea galácticas y/o a mecanismos que involucran la pérdida de gas por retroalimentación estelar, (Brinkmann *et al.* 2017). Una fracción de estos sistemas sobrevive, no obstante, en la fase incrustada y permanece ligada a escalas de tiempo más largas.

Los OCs jóvenes permiten investigar las regiones de formación de estrellas y comprender los mecanismos que dan lugar a éstas. Por su parte los cúmulos intermedios y antiguos informan sobre los procesos estelares y la evolución del disco galáctico. Algunas órbitas muy perturbadas podrían también proporcionar evidencia de sucesos recientes de fusión y rastros de acreción pasada desde fuera de la galaxia, (Cantat-Gaudin *et al.* 2016).

El estudio de los OCs ha tomado nuevo impulso a partir del inmenso y preciso conjunto de datos de la misión GAIA (ESA), tanto para la revisión de los ya conocidos como para la determinación de otros nuevos, habiéndose aumentado el catálogo anterior de cúmulos abiertos hasta prácticamente doblarlo con unos 3.800 OCs.

Los OCs pueden contener desde sólo unas decenas de componentes a unos pocos miles de estrellas y suelen encontrarse casi exclusivamente en el plano de la galaxia.

No hay que confundir los **cúmulos abiertos** con los **Cúmulos Globulares**. Estos últimos, también son conjuntos de estrellas muy unidas entre sí por la gravedad, pero en este caso compuestos de varios centenares de miles, lo que les da una forma esférica y densidades estelares muy altas en sus centros. Se encuentran en el halo galáctico con órbitas que cruzan el plano de la galaxia. De estos sólo se conocen alrededor de unos 150 en nuestra galaxia, son mucho más antiguos y su origen y evolución está poco claro todavía, aunque en general se asume que pueden ser núcleos de galaxias fallidas en las cercanías de la Vía Láctea y que terminaron en una fase muy temprana formando parte de ésta.



Fig. 1 - Las Pleiades, un OC compuesto aproximadamente por un centenar de estrellas, situadas a una distancia de 400 años-luz (M45 – NGC 1432/35).



Fig. 2 - NGC 2516 – con más de 500 estrellas, situado a 1.300 años-luz. Es un cúmulo muy joven, apenas tiene 150 millones de años.

Primeros pasos

Antes de entrar en faena conviene hacer un breve recorrido por los cúmulos más representativos con imágenes incluidas. Esta fase, además de ayudar como elemento motivador, debe servir para tener una idea más completa de los OCs. Para ello es suficiente la Wiki:

https://commons.wikimedia.org/wiki/Category:Open_clusters

Esta etapa puede extenderse todo lo que se considere necesario y es susceptible de ser complementada con observaciones, con telescopios de campo, de cúmulos abiertos. Otros recursos que pueden resultar de interés son:

https://www.youtube.com/watch?v=iPUFkoM_SDM

<https://www.youtube.com/watch?v=Wqqj2T6ox70>

<https://www.youtube.com/watch?v=nnGVL9LEVDE>

https://www.cosmos.esa.int/web/gaia/iow_20201015

Ya es momento de acceder al fichero que contiene el catálogo de los OCs que vamos a utilizar ^[1]:

<http://vizier.cds.unistra.fr/viz-bin/VizieR-3?-source=B/ocl/clusters>

En esta página nos encontraremos con lo necesario para hacer la consulta al catálogo de OCs y algunas herramientas adicionales, (no es necesario registrarse).

The screenshot shows the VizieR web interface. At the top, there are search criteria for 'B/ocl' and 'clusters'. Below that, there is a 'Simple Constraint' section with a table of columns to query. The table has columns for 'Show', 'Sort', 'Column', and 'Explain (UCD)'. The 'Column' column contains various parameters like 'ra', 'dec', 'class', 'diam', 'dist', 'mag', 'age', 'pmra', 'pmra_err', 'pmdec', 'pmdec_err', 'nc', 'k14', 'rv', 'rv_err', 'o_rv', 'r_rv', 'feh', 'feh_err', 'trtype', 'webda', and 'lynga'. The 'Explain (UCD)' column provides detailed descriptions for each parameter. The 'Query by Constraints' section is also visible, showing the applied constraints and the resulting table of columns.

Fig. 3 – Acceso al catálogo CDS-VizieR.

[1] CDS-VizieR Catalog – Créditos: “This research has made use of the VizieR catalog access tool, CDS, Strasbourg, France, <https://ui.adsabs.harvard.edu/abs/2002A%26A...389..871D/abstract>”. Este fichero tiene 2.167 registros. Hay otros catálogos con un mayor número de OCs, ([https://vizier.cds.unistra.fr/viz-bin/VizieR?-source=J/A+A/652/A102, 3.794 rows](https://vizier.cds.unistra.fr/viz-bin/VizieR?-source=J/A+A/652/A102,3.794%20rows)), pero para los propósitos de este trabajo es más adecuado el seleccionado.

El cuerpo principal de la página presenta las diferentes columnas que integran el fichero, con indicación de las unidades en que están expresados los datos y una breve explicación de su contenido, (seleccionaremos únicamente las marcadas en la fig. 3. Ver también la fig. 4).

También hay una ventana para configurar la consulta (fig. 4) y, a la derecha, un botón 'submit', repetido tres veces, para ejecutarla y obtener los datos.

La configuración de la consulta es bastante sencilla y sólo hay que tener en cuenta unos pocos detalles (fig. 4):

1. En 'preferencias', el parámetro '*max*' se refiere al número máximo de registros que devolverá la consulta. Como sabemos que el número total de OCs inventariados en este catálogo es de 2.167, (ver nota ^[1]), si queremos recuperar todos, habrá que poner aquí un número superior a este total.

2. El segundo detalle a tener en cuenta es el formato de los datos recuperados. Por defecto se ofrece la opción de una tabla HTML. En este caso los datos se presentarán en el navegador en forma de tabla. (Antes de descargar los datos es recomendable hacer la consulta HTML con objeto de ver que es correcta, fig. 5)

Para descargar el catálogo completo y guardarlo en el ordenador para trabajar después con ellos es conveniente seleccionar aquí la opción '*XML + CSV (Astrores)*', fig. 4. El fichero descargado de esta manera tendrá una cabecera y un final '*.xml*', que es fácil de eliminar con un editor de texto sencillo, y dejar así el contenido '*.txt limpio*' para su importación a una tabla Excel.

También interesa solicitar las coordenadas en formato decimal. Esto nos ahorrará algunas manipulaciones después.

Hecho esto ya se puede ejecutar la consulta (***submit***).

Fig. 4 – Configuración de la consulta para descargar los datos deseados.

Una vez obtenidos los datos relativos a los OCs catalogados, no son necesarias demasiadas explicaciones. El contenido de los valores recuperados es fácilmente interpretable. No obstante, podemos aprovechar para aclarar algunos conceptos:

Full	RAJ2000 deg	DEJ2000 deg	Cluster	RAJ2000 "h:m:s"	DEJ2000 "d:m:s"	Class	Diam arcmin	Dist pc	E(B-V) mag	Age [yr]	pmRA mas/yr	pmDE mas/yr	Nc	RV km/s	o	[Fe/H] Sun	TrType
1	000.05000	+60.96667	Berkeley 58	00 00 12	+60 58 00		11.0	2700	0.720	8.470	0.56	1.56	88	-87.00	1		
2	000.08750	+50.74167	NGC 7801	00 00 21	+50 44 30	r	8.0	1275	0.170	9.230	-3.56	-2.60	85				4-3-p-*
3	000.16250	+59.23889	FSR 0459	00 00 39	+59 14 20	IR	2.0	3800	1.103	8.950	-1.66	-0.01	8				
4	000.40417	+64.62500	Stock 18	00 01 37	+64 37 30		6.0	1242	0.710	8.100	-2.78	-0.15	109				4-2-p-*
5	000.55833	+67.41667	Berkeley 59	00 02 14	+67 25 00		10.0	1000	1.241	6.100	-2.11	-1.20	2	-5.17	3		3-3-m-n
6	000.87500	+63.58333	Berkeley 104	00 03 30	+63 35 00		3.0	4365	0.450	8.890	-2.35	-0.30	76			0.070	4-2-p-*
7	001.02917	-29.83333	Blanco 1	00 04 07	-29 50 00		70.0	269	0.010	7.796	20.17	3.00	27	5.53	49	0.040	
8	001.17083	+56.08333	Stock 19	00 04 41	+56 05 00		4.0				1.43	-2.04	42				2-2-p-*
9	001.32083	-20.69167	NGC 7826	00 05 17	-20 41 30	r	20.0	620	0.030	9.340	6.48	-6.20	39	-4.62	1		4-2-p-*
10	001.84167	+64.97278	Majaess 1	00 07 22	+64 58 22	IR	5.0				4.63	1.72	20				
11	001.90833	+61.47500	Czernik 1	00 07 38	+61 28 30		2.5	2530	1.230	6.700	-3.83	1.66	5				4-2-m-*
12	002.08333	+51.72083	SAI 1	00 08 20	+51 43 15		4.0	2170	0.340	9.100	0.31	-0.88	35				
13	002.40000	+60.47500	Berkeley 1	00 09 36	+60 28 30		5.0	2420	0.780	8.600	-1.83	-0.75	135				3-1-p-*
14	002.40000	+62.68000	ASCC 1	00 09 36	+62 40 48		24.0	4000	0.160	8.250	-3.27	1.19	1189	-74.10	3		
15	002.52500	+61.16667	King 13	00 10 06	+61 10 00		5.0	3100	0.820	8.500	-4.15	-2.46	127				2-2-m-*
16	002.64167	+58.76000	Alessi 20	00 10 34	+58 45 36	e	40.0	600	0.220	6.480	8.73	-3.11	17	-11.50	1		
17	002.72083	+59.74722	FSR 0474	00 10 53	+59 44 50	IR	4.6	4638	1.145	8.900	-0.41	-0.81	53				
18	003.72917	+61.47306	FSR 0480	00 14 55	+61 28 23	IR	1.5	5492	0.937	9.300	-1.66	-1.81	24				
19	004.08333	+59.96194	Juchert Saloran 1	00 16 20	+59 57 43		5.0	3859	0.830	9.050	-0.59	0.58	60				
20	004.17917	+64.50583	Majaess 8	00 16 43	+64 30 21	IR	3.0				-2.38	2.06	24				
21	004.42500	+60.93333	Berkeley 60	00 17 42	+60 56 00		3.0	4365	0.860	8.200	-3.64	0.20	40			0.070	3-1-p-*

Fig. 5 – Muestra del catálogo (formato HTML).

Las coordenadas celestes, (*RAJ2000*, *DEJ2000*), son la *ascensión recta* y *declinación* (coordenadas ecuatoriales) de cada uno de los OCs. Como los cúmulos abiertos son objetos extensos, (no puntuales, como una estrella), se está dando una posición media considerada como ‘centro del cúmulo’.

El sufijo ‘J2000’ indica que los valores están referidos a la ‘Época’ de las coordenadas [2], (en este caso, 1 de enero de 2000, a las 12 TT). Esto es algo que debe tenerse en cuenta si se van a cruzar los objetos de este catálogo con otros catálogos, (por ejemplo con el de Gaia).

Como al configurar la consulta se ha añadido la opción ‘decimal’ para las posiciones, se añaden al comienzo de la tabla las coordenadas (*AR*, *DEC*) en grados en formato decimal.

Otros parámetros que ofrece el catálogo son los referidos a los *movimientos propios* de los OCs, tanto en *AR*, (*pmRA*), como en *DEC*, (*pmDE*), (en *velocidad radial*, *RV*, no en todos los casos se tiene dato empírico), así como la distancia media al cúmulo (en parsecs), la edad estimada (su logaritmo decimal en años), un valor del diámetro o ‘anchura’ del cúmulo, (en minutos de arco), el ‘exceso de color’ (*B-V*, en magnitud fotométrica). Estos parámetros serán necesarios para los ejercicios que haremos a lo largo de este trabajo. También se indica una estimación del número de estrellas que componen cada OC y, en muchos de ellos, su tipología *Trumpler*. Ver:

http://www.astrophoton.com/trumpler_class.htm.

[2] Según la IAU, las coordenadas celestes se ‘refrescarán’ cada 50 años, (J1950, J2000, J2050, ...). Estas correcciones se deben fundamentalmente a los movimientos de precesión y nutación del eje terrestre. La recomendación de la IAU es que las coordenadas utilizadas estén referidas a la ‘Época’ con ± 25 años, esto es, a partir de 2025 las que deberían utilizarse son las J2050, así hasta el año 2075.

Primeros ejercicios

Con la tabla ya descargada, la importación a Excel no requiere demasiados cuidados. Una vez eliminadas las líneas que no contienen datos, como la cabecera y final 'xml', el fichero se importa sin mayor complicación teniendo en cuenta el separador ';' y que los valores decimales en los datos van con un punto '.', que tal vez convenga cambiar a coma ',' antes de hacer el *import*. La primera línea será la cabecera con los nombres de las columnas y, como puede verse, la segunda línea son las unidades en que vienen expresados los datos. Conviene agregar esta información a cada columna en la primera línea y eliminar la segunda línea, (también la tercera, que son meros guiones para remarcar la cabecera). A medida que vayamos haciendo uso de esta tabla se irán concretando detalles de las columnas que aún no han sido comentadas.

Por ahora lo primero que haremos, (después de comprobar la integridad de los datos y ajustar los formatos para trabajar con comodidad), es intentar un mapa de distribución de OCs en el cielo. Si lo hacemos en coordenadas ecuatoriales (con los datos de las columnas `_RAJ2000`, `_DECJ2000`), es casi inmediato. Si no hubiéramos descargado las coordenadas de posición en formato decimal, deberíamos haberlas calculado a partir de las coordenadas en sexagesimal (h:m:s, °, m, s):

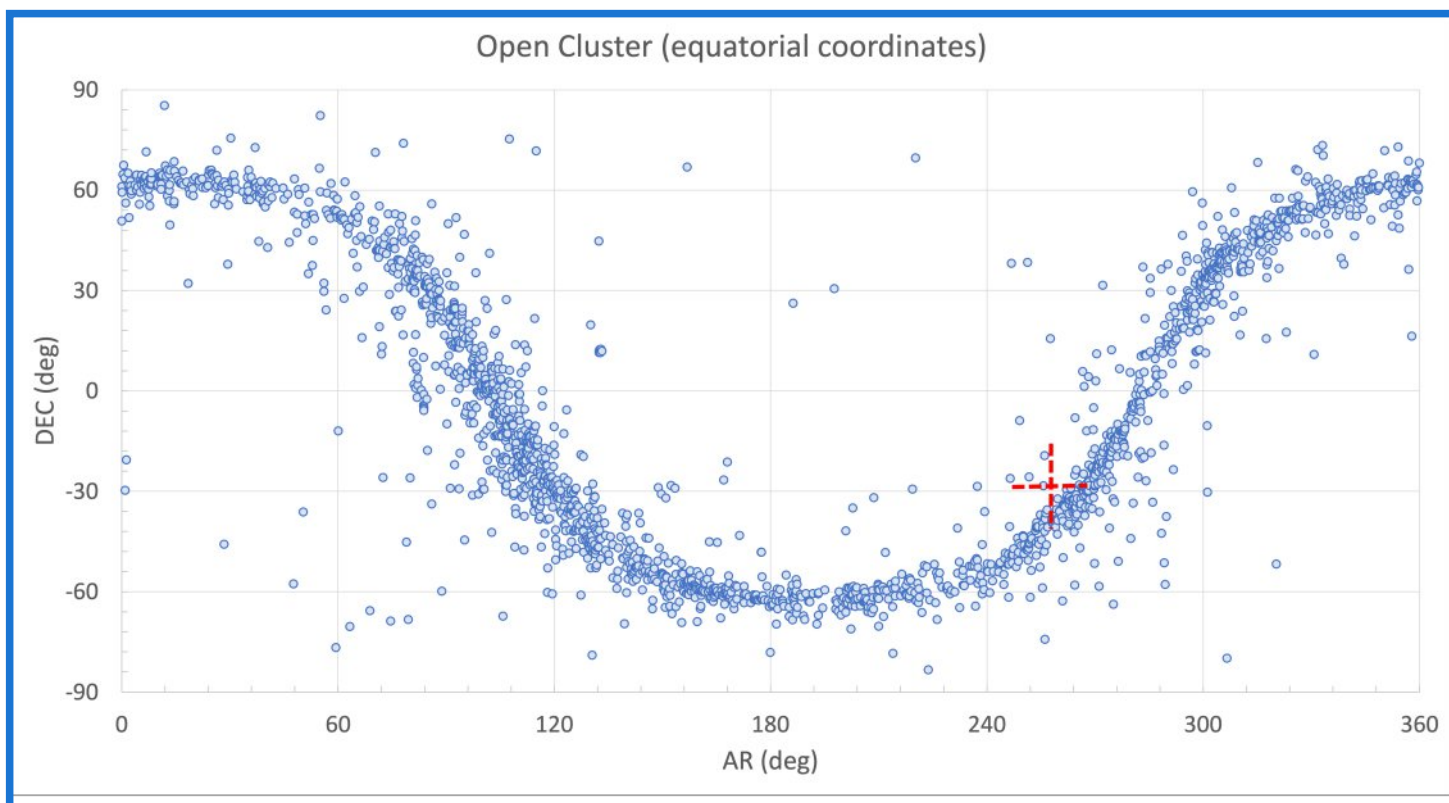


Fig. 6 – Distribución de OCs en coordenadas ecuatoriales.

Se puede intentar añadir a la gráfica el Centro Galáctico (marcado con guiones en rojo en la fig. 6), que para las coordenadas de la Época J2000 está situado en 17h 45m y 40.04s de ascensión recta y $-29^{\circ} 00'$ y $28.1''$ en declinación. Como resulta evidente se requiere aplicar las transformaciones de unidades correspondientes (h, m, s) a grados.

Otra representación, más significativa, de estos mismos datos puede hacerse sobre las coordenadas galácticas (fig. 7). Pero ahora sí, debemos añadir un par de columnas a nuestra tabla de referencia donde insertar los valores calculados de las coordenadas en *longitud* y *latitud* galácticas. Ver nota al pie ^[3]. Como resulta evidente, este cálculo no es tan sencillo pero puede resultar un buen ejercicio utilizando un poco de programación, o incluso en el propio Excel haciendo uso de algunas columnas adicionales para pasos intermedios.

En la gráfica de la fig. 7 se hace patente que todos, (o casi), los OCs están situados en el plano galáctico.

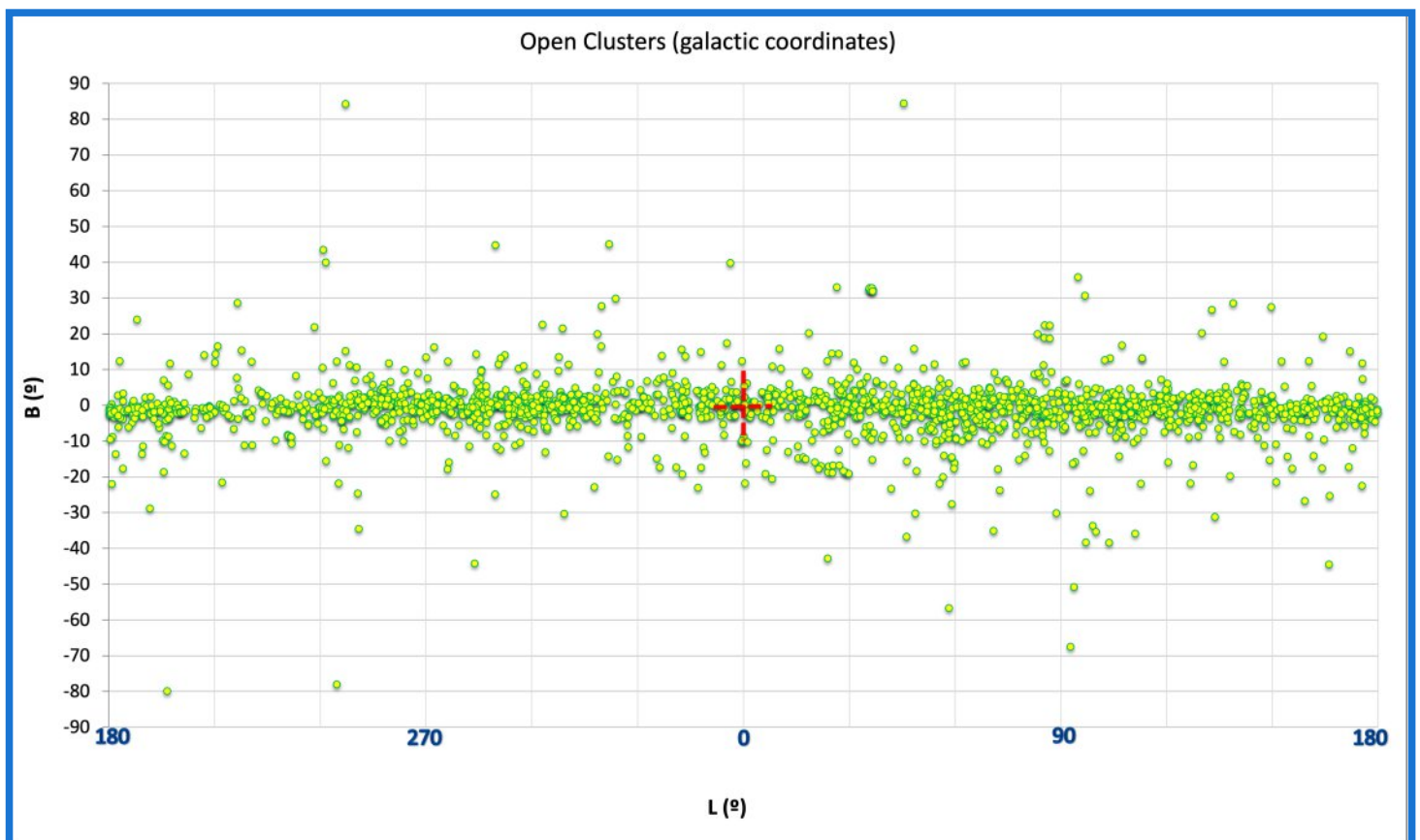


Fig. 7 – Distribución de cúmulos abiertos en coordenadas galácticas. Nótese que para situar el centro galáctico en la posición central de la gráfica se ha desplazado el '0' 180 grados en la coordenada 'longitud'.

Esta imagen puede resultar engañosa haciendo pensar que realmente tenemos una representación de la galaxia completa a través de los cúmulos que la pueblan. En realidad nuestro alcance visual de la galaxia es limitado. El Sol, (y con él, la Tierra), se encuentra en el llamado Puente de Orión, entre los brazos de Perseo y Sagitario, de manera que el entorno al que los telescopios, tanto terrestres como espaciales, tienen acceso en el plano galáctico se reduce considerablemente debido en su mayor parte a las nubes de gas presentes en estos brazos, (ver fig. 8).

[3] https://es.wikipedia.org/wiki/Conversión_de_coordenadas_ecuatoriales_a_coordenadas_galácticas

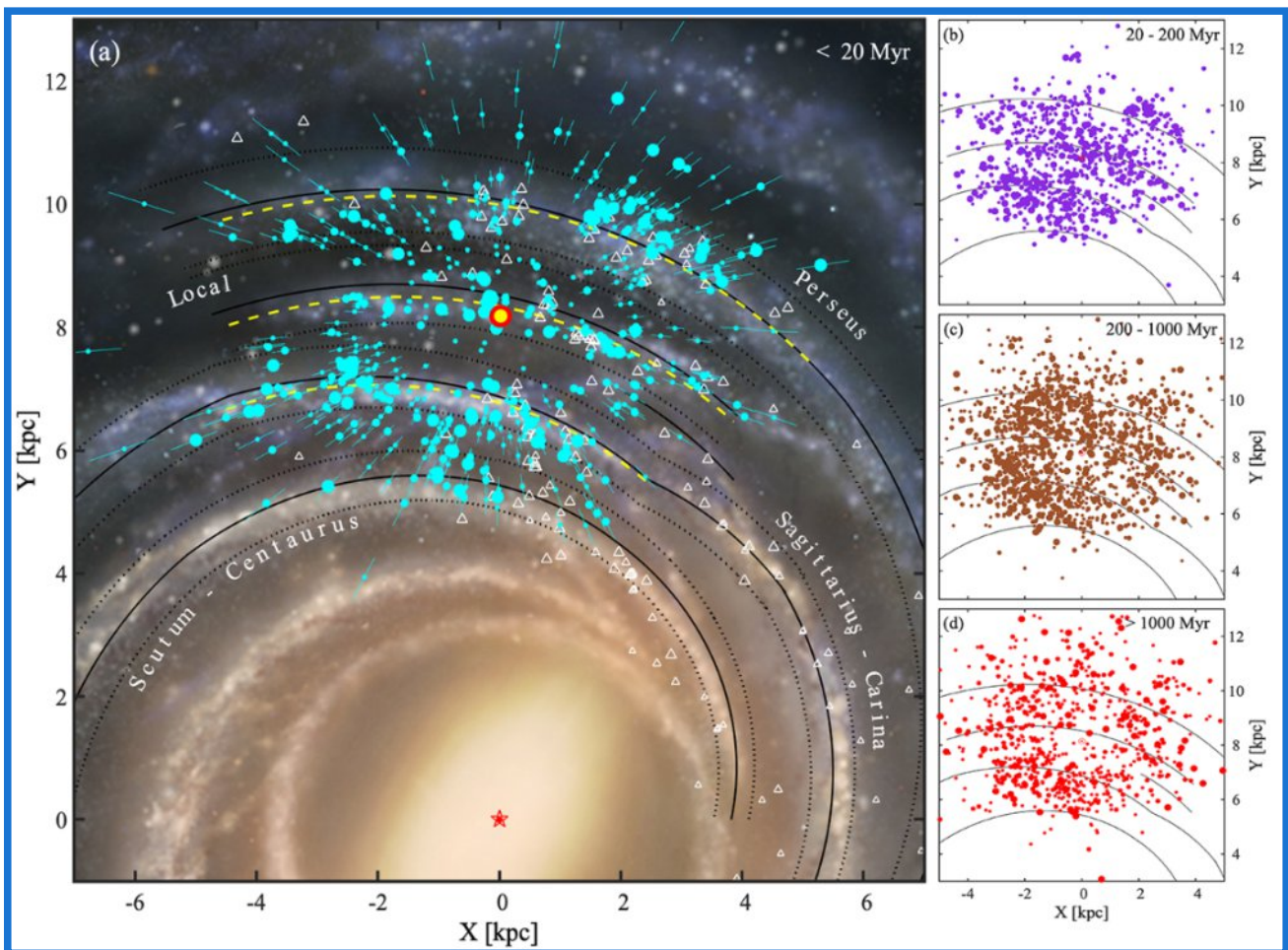


Fig. 8 – Vista cenital de la galaxia con la posición de los cúmulos abiertos conocidos. En primer plano los más jóvenes, con edades inferiores a 20 millones de años. A la derecha, sucesivas representaciones para cúmulos más antiguos. El Sol está representado en la imagen central por un círculo amarillo. (Créditos imagen: “[Evolution of the local spiral structure of the Milky Way revealed by open clusters](#)” – C. J. Hao et al. - Astronomy and Astrophysics 652, A102 [2021].

Nos centraremos ahora en el parámetro ‘Distancia’. También en este caso el valor que aparece para cada OC en el catálogo es un valor medio, (recuérdese que un OC es un objeto extenso compuesto por algunos centenares de estrellas). Las unidades de este campo en el fichero son *parsecs* ^[4], pero el procedimiento empírico por el que se estima este valor es el de la *paralaje* ^[5], por lo que es interesante añadir una columna a nuestra tabla con el valor calculado de la ‘distancia’ en forma de paralaje (en unidades de mili-segundo de arco, *mas*), que nos servirá más adelante. La relación *parsec-paralaje* es sencilla:

$$distancia (pc) = 1 / plx (segundos de arco) \rightarrow plx (mas) = 1000 / dist (pc)$$

Por otra parte una representación bidimensional de la distancia frente a cualquier otro parámetro, (RA, DEC por ejemplo), es poco significativo. Sí es interesante un mapa 3D, para lo que necesitamos convertir las coordenadas ecuatoriales (esféricas) a cartesianas, teniendo en cuenta que ‘distancia’ es el módulo del vector que buscamos, (no olvidar expresar los ángulos en radianes):

$$\begin{aligned} x &= distancia \cdot \coseno(DEC) \cdot \coseno(AR) \\ y &= distancia \cdot \coseno(DEC) \cdot \seno(AR) \\ z &= distancia \cdot \seno(DEC) \end{aligned}$$

Con un poco de programación Python, (ver nota ^[6]) se pueden obtener estas representaciones fácilmente. En la fig. 10 se ofrecen otras perspectivas. El origen de coordenadas (0,0,0) corresponde a la posición de la Tierra.

[4] Pársec. Ver <https://es.wikipedia.org/wiki/Pársec>.

[5] Paralaje estelar: https://es.wikipedia.org/wiki/Paralaje_estelar.

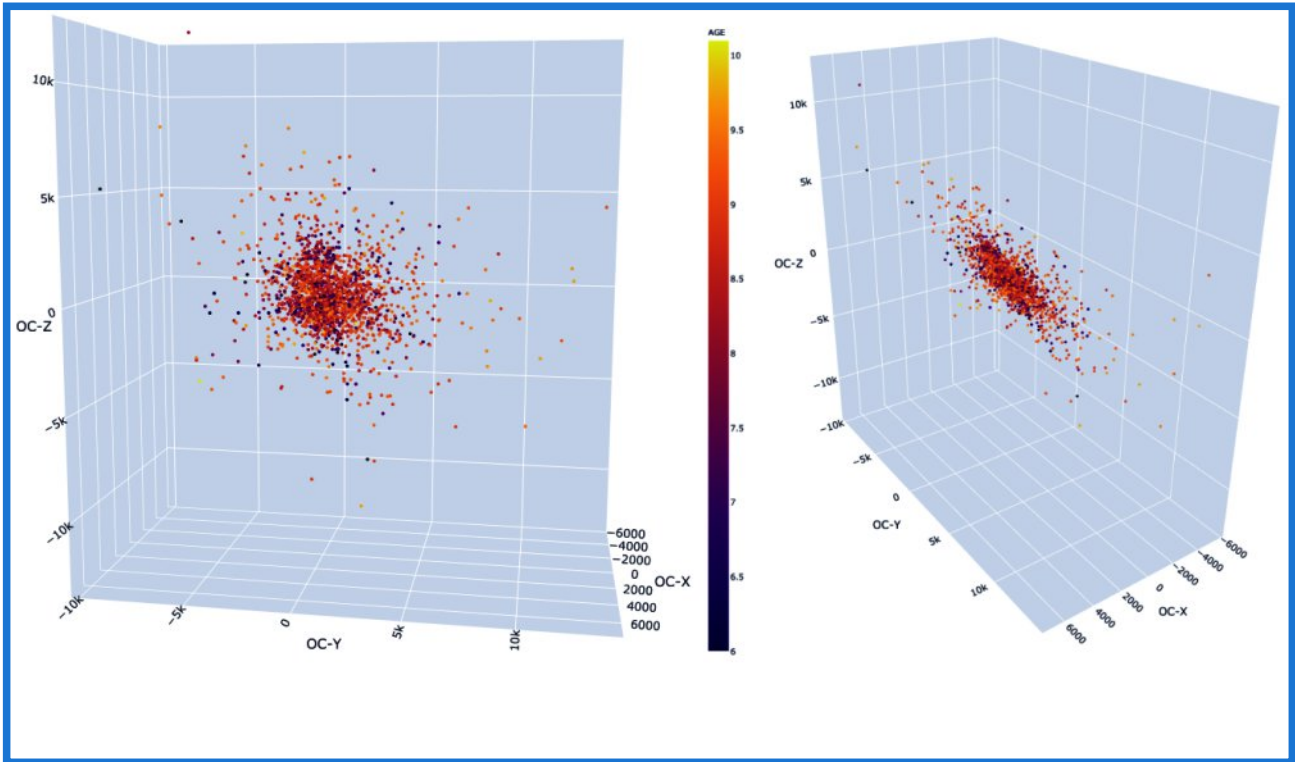


Fig-9 – Mapa 3D de la distribución de los cúmulos abiertos desde distintas perspectivas al rotar los ejes.

En la fig. 10 (izquierda) se aprecia cómo los cúmulos representados se distribuyen en el plano de la galaxia. Debe tenerse en cuenta lo comentado en la fig. 8 acerca de las limitaciones para observar en el plano galáctico. A pesar de las altas capacidades instrumentales de observación disponibles, en este entorno no dejamos de tener una visión bastante local.

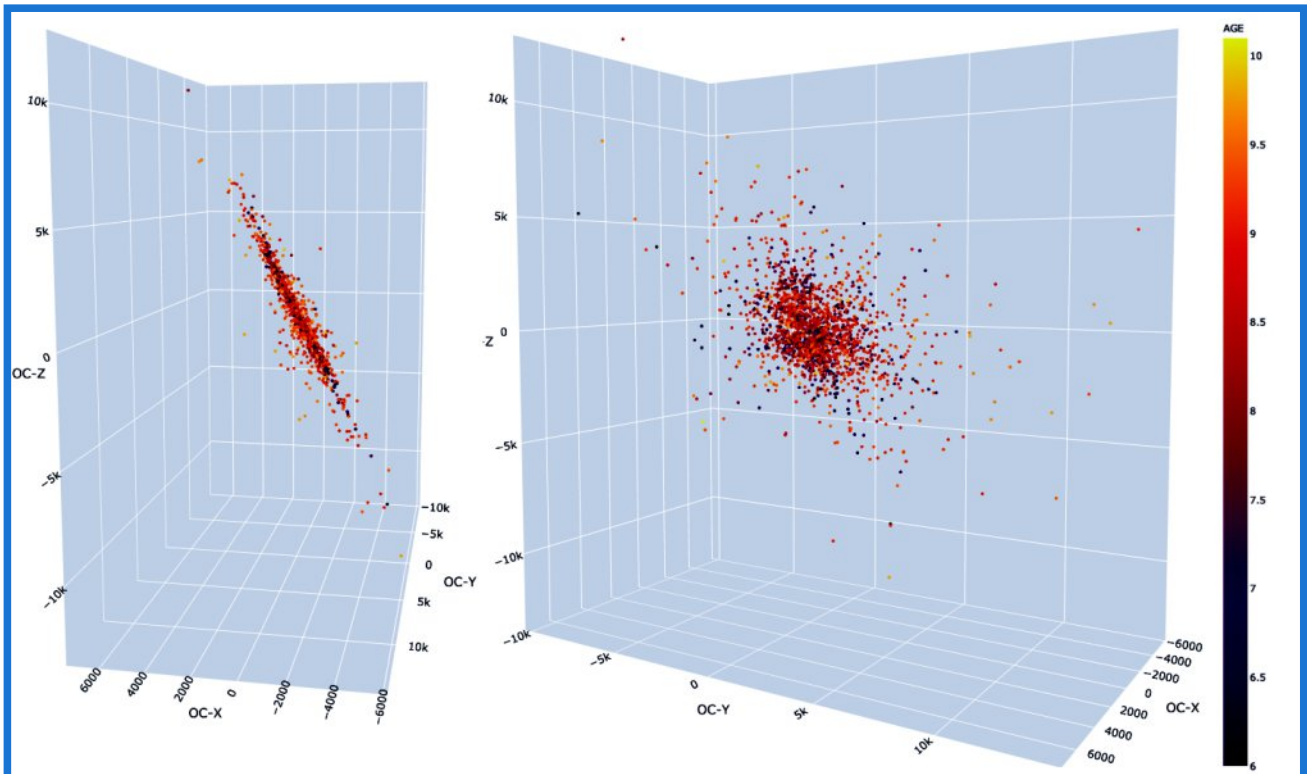


Fig. 10 – Otras perspectivas en la distribución de OCs. Los colores representan el campo 'edad'^[6].

Conclusión

En este artículo se ha pretendido acercarnos al conocimiento de los cúmulos estelares abiertos y situarnos en su contexto. Debe quedar claro que los datos utilizados, y en general todo el catálogo presentado, no son el comienzo del estudio de OCs, sino justamente su resultado. El problema real es la *caracterización* del cúmulo, esto es, allí donde hay una sobre-densidad de estrellas, y por tanto un potencial OC, el trabajo consistirá en determinar qué estrellas de entre todas las presentes en el campo visual observado forman parte del posible cúmulo y, a partir de éstas, determinar las propiedades del mismo. En unos casos es sencillo y el cúmulo se hace fácilmente evidente, pero en otros no es una tarea trivial. Una vez caracterizado el cúmulo se le da entrada en el catálogo.

Como ya se ha comentado, la ingente cantidad de datos, de exquisita calidad astrométrica y fotométrica, recopilados por la misión Gaia ha permitido casi duplicar en los últimos años el número de OCs conocidos utilizando herramientas de inteligencia artificial [7]. Nuestro propósito en los próximos artículos es hacer uso de herramientas más clásicas del Observatorio Virtual, (TOPCAT, Aladin y Clusterix), para introducirnos en la caracterización de OCs, utilizando para ello algunos de los objetos ya catalogados lo que permitirá contrastar los resultados obtenidos.

[6] Código Python utilizado para las representaciones 3D. Genera un .html (mapa_3D.html) que se ejecuta directamente en el navegador con varias funcionalidades, (zoom, rotación de ejes, etc.). A partir de un fichero.csv con los campos de las coordenadas x, y, z, (se ha incluido también el campo edad (AGE) para representarlo a través del color), se obtiene 'mapa_3D.html':

```
import numpy as np
import pandas as pd

df = pd.read_csv('fichero.csv', sep=';', names=['Cluster', 'OC-X', 'OC-Y', 'OC-Z', 'AGE'], header=0)

import plotly
import plotly.express as px

fig = px.scatter_3d(df, x='OC-X', y='OC-Y', z='OC-Z', color='AGE')
fig.update_traces(marker=dict(size=2))
plotly.offline.plot(fig, filename='mapa_3D.html')
```

[7] Cantat-Gaudin et al. – “A Gaia DR2 view of the Open Cluster population in the Milky Way” - <https://arxiv.org/abs/1805.08726> .
Castro-Ginard et al. – “Hunting for open clusters in Gaia: 582 new OCs in the Galactic disc” - <https://arxiv.org/pdf/2001.07122.pdf>.

2ª PARTE

CARACTERIZACIÓN CON TOPCAT Y ALADIN



NGC 2002 - Cúmulo abierto observado por el Telescopio Espacial Hubble en la Gran Nube de Magallanes, galaxia satélite de la Vía Láctea a 160.000 años luz. El cúmulo es muy joven, sólo tiene unos 18 millones de años, y está formado por unas 1.100 estrellas. Como se desprende de esta imagen, los OCs (Open Clusters) no son exclusivos de nuestra galaxia. El proceso de formación de estrellas es similar en todo el universo.

En la figura 1 se aprecian miles de puntitos azules, formando una esfera en el centro de la imagen. Son todas las estrellas (hasta magnitud 21) presentes en el campo visual con un determinado radio desde el centro. Sólo las marcadas como puntos amarillos en ese conjunto se consideran integrantes de M67.

¿Cómo se llega a discriminar este subconjunto, esto es, el cúmulo? Ésta es la tarea que nos proponemos a partir de aquí. El proceso de análisis puede articularse en torno a los tres criterios principales que han de cumplir las estrellas de un cúmulo estelar.



Fig. 1 – M67, (imagen obtenida con Aladin).

Como por definición un OC es un conjunto de estrellas nacidas de una misma nube de gas molecular en una ventana temporal relativamente corta, las estrellas del cúmulo van a compartir al menos tres propiedades:

1. La dinámica de la nube de gas progenitora. Esto significa que estas estrellas presentarán movimientos propios similares, ($pmRA$, $pmDEC$).
2. Serán estrellas próximas entre sí. Es decir, su distancia a un punto de referencia como puede ser la Tierra debe ser la misma, dentro de un rango más o menos reducido debido a la propia extensión del cúmulo.
3. Además, estas estrellas tendrán una misma composición química, (definida por la nube original), y una misma edad, también dentro de un rango temporal bastante estrecho.

Antes de empezar necesitamos ahora descargar e instalar en nuestro ordenador dos herramientas del Observatorio Virtual:

a) **TOPCAT** (*Tool for OPerations on Catalogues And Tables*)

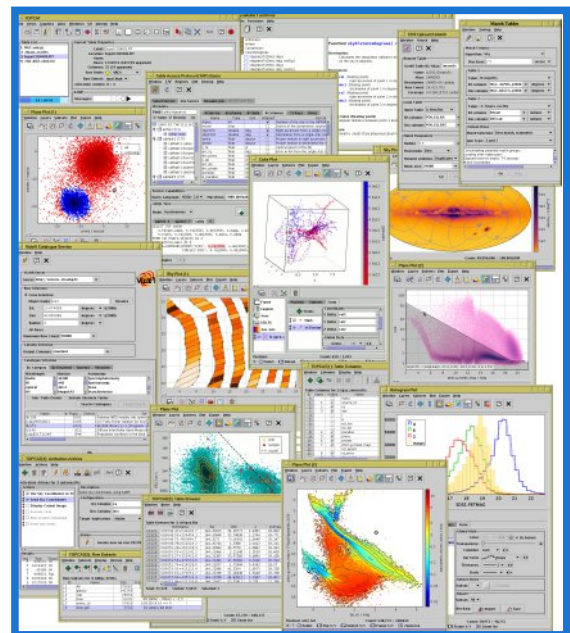
<https://www.star.bris.ac.uk/~mbt/topcat/>



En la página del enlace anterior se puede descargar TOPCAT. Está disponible para múltiples plataformas (Windows, Mac, Unix), y ahí se dan instrucciones completas acerca de su instalación, tutoriales, etc.

Al ejecutar TOPCAT se abre una simple ventana, pero a medida que se trabaja con esta herramienta van apareciendo ventanas nuevas, de manera que al final la pantalla puede llegar a tener un aspecto similar al de la figura adjunta.

Iremos avanzando paso a paso, de manera que no es necesario por ahora dedicar tiempo a los tutoriales o a una exploración exhaustiva de TOPCAT^[1].



b) **Aladin**



<http://aladin.cds.unistra.fr/java/nph-aladin.pl?frame=downloading>

Aladin Desktop^[2] es un atlas profesional con múltiples funcionalidades y acceso a numerosos catálogos de objetos del cielo. Disponible para todas las plataformas y fácil de instalar. Hay una versión más ligera, **Aladin Lite**, que se ejecuta en el navegador, pero utilizaremos la versión Desktop que nos permitirá enlazar con TOPCAT directamente.

Primer ejercicio.- Tomaremos el cúmulo **NGC 2516** como objeto de nuestro primer trabajo. Por ahora del catálogo que ya se descargó anteriormente y que nos ocupó en el primer artículo de esta serie, sólo anotaremos un dato: los 30 *arm* que figuran en dicha tabla como diámetro del cúmulo. (Hacemos esto para evitar en este primer ejercicio tener que ensayar un poco a ciegas diferentes valores, como veremos después).

[1] Micro-curso sobre TOPCAT del SVO (en español): <https://www.youtube.com/watch?v=IJ3qeliOOGY>

[2] Micro-curso sobre Aladin del SVO (en español): <https://www.youtube.com/watch?v=Ajro94X7kUc&t=2498s>

1. Abrimos Aladin y en la línea de comando de la parte superior escribimos NGC 2516, (*enter*). Aladin devolverá de inmediato la imagen del cielo ocupada por el objeto que le hemos pedido resolver:

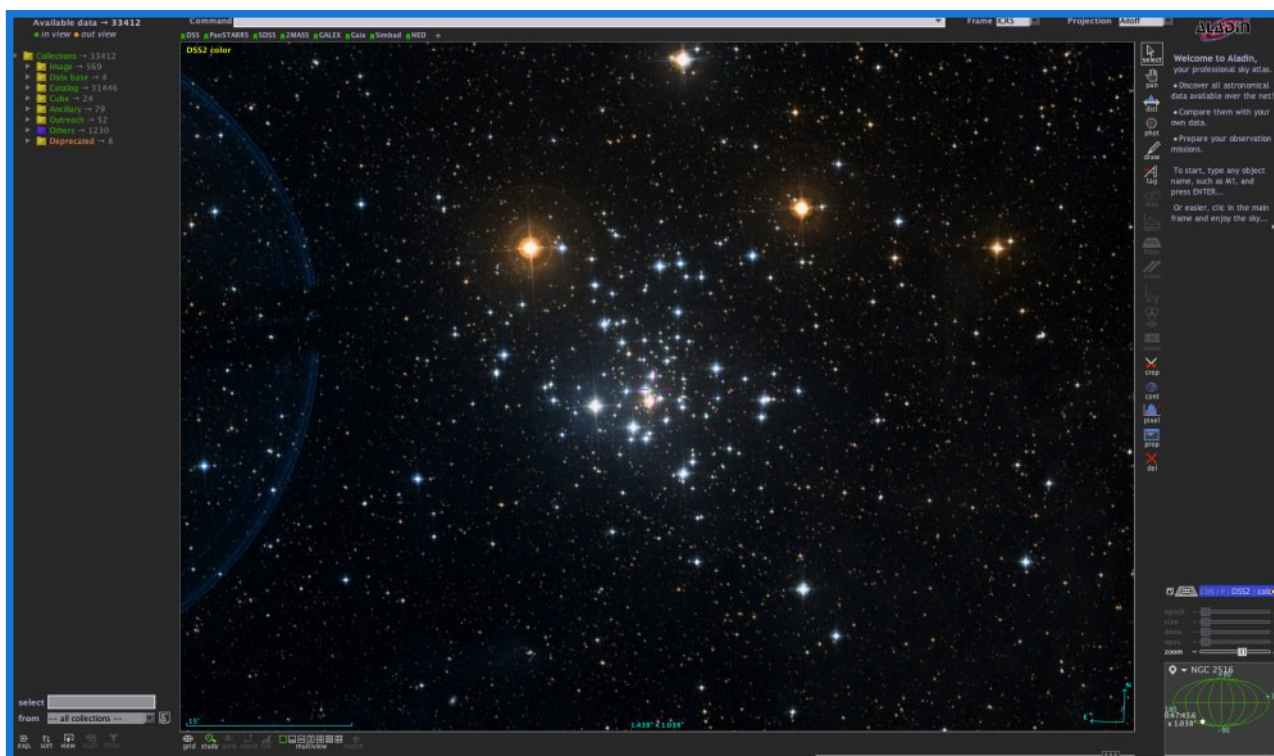


Fig. 2 – NGC 2516 – imagen de DSS2 ofrecida por Aladin.

2. Dejamos Aladin abierto y ahora ejecutamos TOPCAT. Nuestro propósito aquí es obtener de Gaia/EDR3 todas las estrellas presentes en el campo visual de las coordenadas del cúmulo NGC 2516, dentro de un radio en el que se considera que está contenido este objeto.

Como se verá al ejecutar TOPCAT, se abre una ventana bastante poco significativa a pesar de que los botones de la parte superior son contextuales e indican su funcionalidad.

Lo que pretendemos aquí es *‘abrir una tabla nueva’*, botón marcado con un círculo rojo. Esto abrirá una nueva ventana en la que debemos actuar sobre el botón *‘obtener fuentes de un catálogo usando un cono para la búsqueda’*, (fig. 4, también marcado).

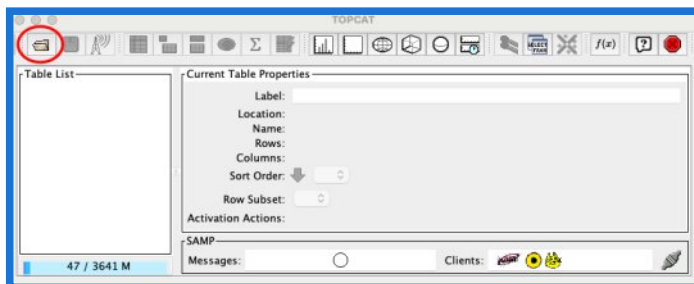


Fig. 3 – Ventana inicial de TOPCAT.

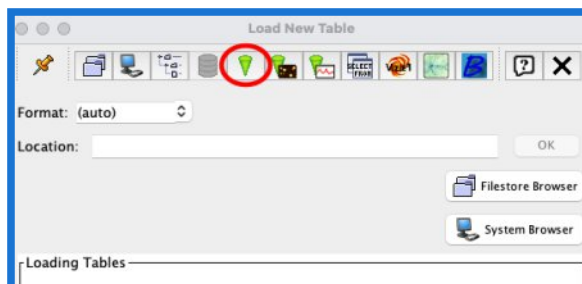


Fig. 4.

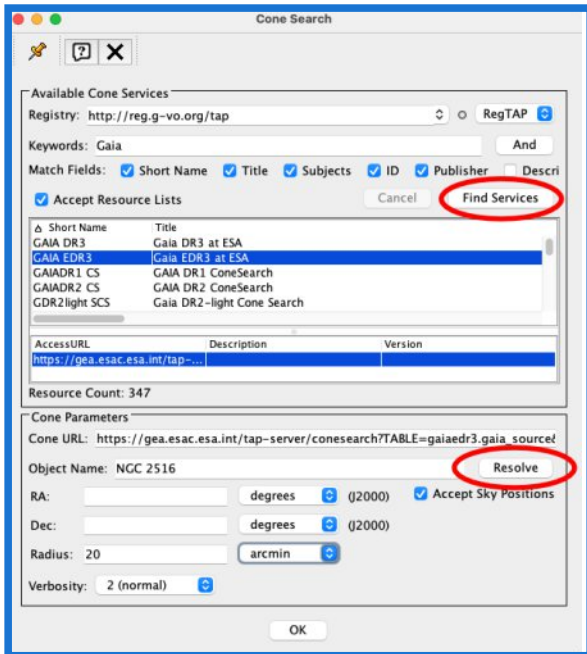


Fig. 5.

objeto, (aparecen sus coordenadas RA, Dec - J2000), y se inicia la descarga de datos. Concluida ésta, se cierran solas las dos ventanas anteriores y en la ventana de inicio nos da información del número de filas descargadas desde Gaia y también del número de columnas, (10.504 y 100 respectivamente).

Estas 10.504 filas corresponden cada una a una estrella de las que Gaia tiene información en ese campo del cielo con radio de 20 *arcmin*.

3. En este momento ya todos los botones de la ventana inicial, (fig. 3), están activos. El segundo por la izquierda nos permite guardar estos datos recuperados de Gaia en diferentes formatos. El tercero, (con aspecto de antena), envía estos datos a otras aplicaciones. Como ya tenemos Aladin abierto, (por eso se pidió antes dejarlo en este estado), la tabla con los 10.504 registros es automáticamente representada en este atlas del cielo, (fig. 6). Probémoslo.

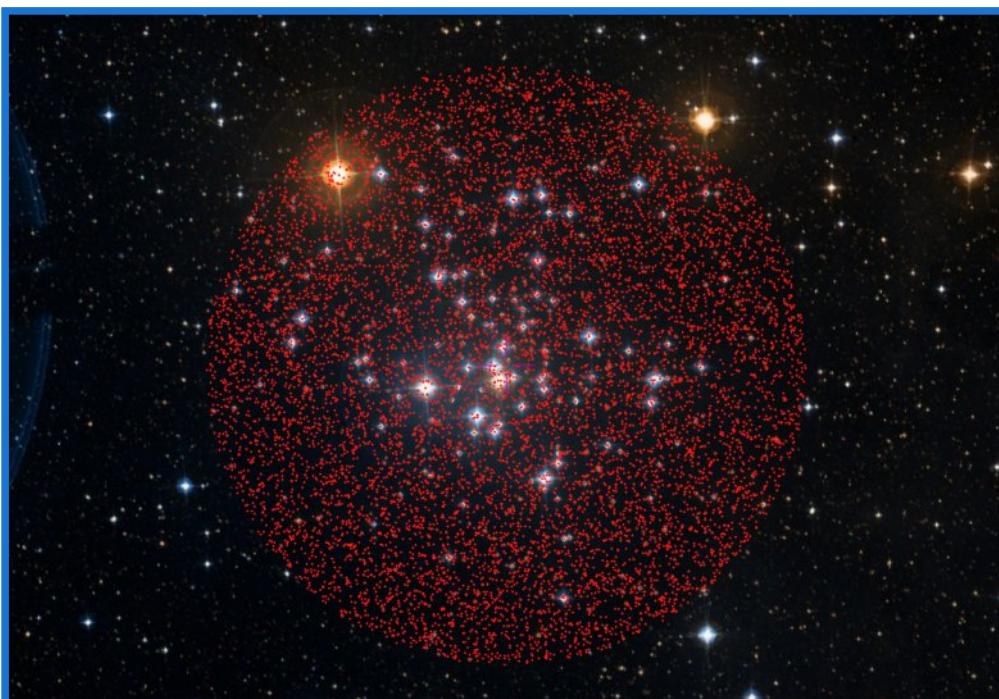


Fig. 6 – Representación de posiciones en Aladin de todas las estrellas, (puntos rojos), recuperadas de Gaia en la consulta hecha a través de TOPCAT.

Se abrirá una nueva ventana (fig.5) en la que ya se debe concretar qué buscamos y dónde:

- En *Keywords* escribiremos Gaia y a continuación el botón '*encontrar servicios*'. Se abrirá un combo debajo en el que seleccionaremos '*GAIA EDR3*', (última *release* disponible de los datos de Gaia).

- En la parte inferior, '*parámetros del cono*', pondremos el objeto que buscamos, NGC 2516 en '*nombre del objeto*', y también rellenamos el tamaño de cono: '*Radius*' donde, como sabemos que este cúmulo está inventariado con un diámetro estimado de 30 minutos de arco, pondremos radio: 20 *arcmin*, (le damos un poco más de holgura).

- Y le pedimos que "*resuelva*"...

Lo que ocurre a continuación es que se localiza el

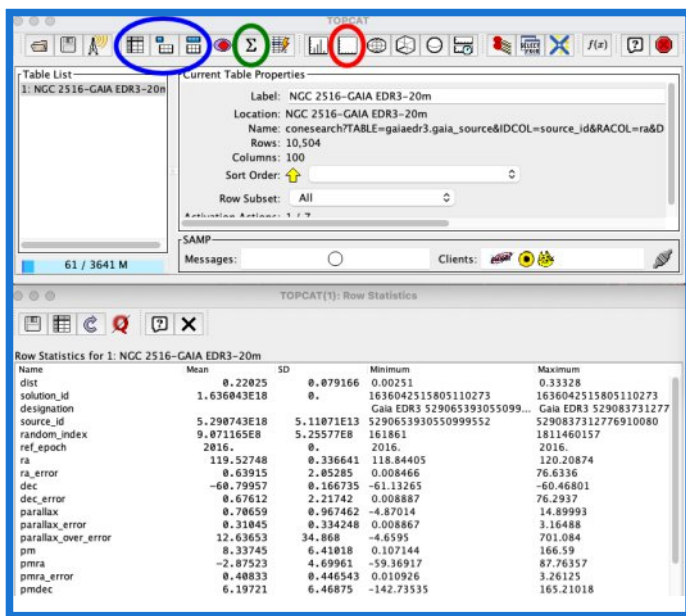


Fig. 7 – Detalle sobre datos estadísticos de la consulta.

4. El siguiente grupo de botones presentan información de los datos obtenidos, tanto de los datos en sí como del metadato asociado o el contenido y unidades de cada columna en formato tabla. Es interesante ahora explorar un poco por encima la información estadística sobre este conjunto de estrellas, (botón remarcado en verde, en la fig. 7).

Como se podrá comprobar aquí, para cada campo se da un valor medio, la desviación estándar, un valor mínimo y un máximo y también el número de estrellas, de las 10.504, para las que el dato presente en cada columna se considera bueno, (o no nulo).

Por ahora tampoco hay que perder demasiado tiempo con esto, pero como detalle nos fijaremos en los valores (*mín*, *máx*) del *parallax*, (-4.8, 14.89 en milisegundos de arco).

Sabemos que la paralaje está relacionada con la distancia, (ver artículo anterior de la serie), y el amplio rango entre estos dos valores nos dan idea de la distinta profundidad en el cielo de las estrellas que estamos tratando.

Otro dato interesante aquí puede ser el rango de magnitudes fotométricas en el que se extiende este conjunto, lo que también nos informa de sus diferentes brillos. Buscaremos en esta tabla moviendo la barra de scroll el campo '*phot_g_mean_mag*' y veremos que este rango está entre 4.29885 y 21.766 para las más débiles.

5. El siguiente bloque de botones en la parte superior de la ventana inicial sirve para desarrollar gráficos. Utilizaremos por ahora el marcado en rojo en la fig.7.

Se abre una nueva ventana, en la que debemos seleccionar qué parámetros queremos representar en cada eje.

Para obtener esta gráfica (fig. 8) se han elegido las coordenadas (*ra*, *dec*) en los ejes (X, Y).

Como puede verse esta es la información enviada a Aladin en el punto 3 (fig. 6).

En esta ventana hay otras funcionalidades para los gráficos que exploraremos según vaya siendo necesario. Por ahora interesa avanzar en nuestro objetivo: *caracterizar el*

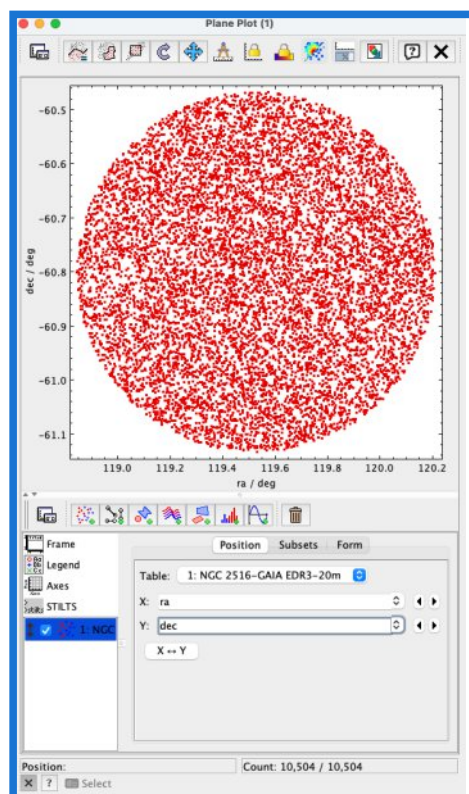


Fig. 8 – Representación gráfica de posiciones.

cúmulo NGC 2516. Y en este sentido el primer paso, como se indicó al principio, es centrar nuestra atención en los movimientos propios, (*pmra* y *pmdec*).

Cambiamos los valores representados en los ejes por estos últimos y obtendremos una nueva representación (fig. 9), en este caso del espacio de configuración de velocidades en milisegundos de arco por año (*mas/year*).

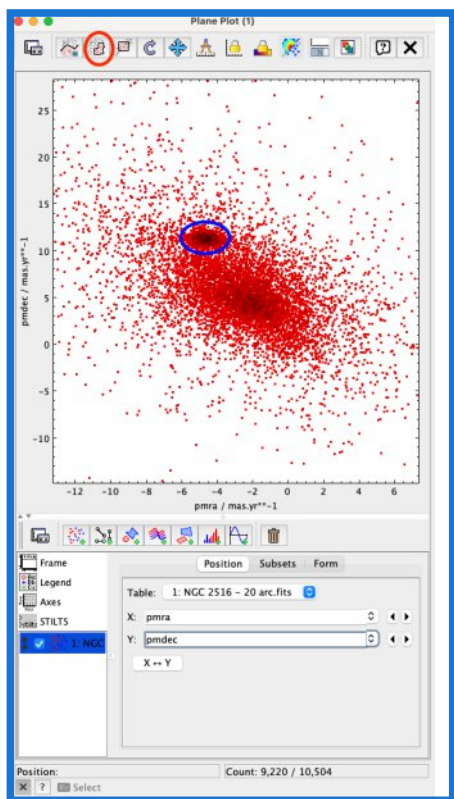


Fig. 9 – Representación de movimientos propios, (*pmra*, *pmdec*)

da opción a generar una nueva tabla con las estrellas seleccionadas en ese área. Esta selección es considerada como un subconjunto al que hay que dar un nombre, (como esto será algo que vamos a repetir alguna vez más, podemos ir numerando las sucesivas versiones).

Llegados aquí, lo que tenemos es el conjunto total de estrellas inicial (10.504) y un primer subconjunto al que hemos etiquetado, por ejemplo, con NGC2516v1, (ver fig. 10).

Ahora tenemos representados dos colectivos de estrellas: a) el total original en rojo, y b) el subconjunto (v1) en azul.

En la ventana inicial de TOPCAT también podemos elegir sobre cuál de estos conjuntos actuamos:

Si en 'row subset' se selecciona el subconjunto -v1- podemos ver el número de componentes, (1.164, en este ejemplo).

6. Se obtiene así una nueva imagen, (se puede modificar el zoom con el scroll del ratón en el área del gráfico). En un campo visual cualquiera del cielo se debería esperar una distribución de velocidades repartidas con carácter básicamente aleatorio en torno a un centro más poblado. Aquí, (ver figura 9), se aprecia una sobre-densidad (remarcada con un círculo azul en la figura) y que es significativa de un posible cúmulo.

Ahora un poco de artesanía. Pinchando en el botón remarcado en rojo, podremos dibujar con el puntero del ratón el área que encierra la sobre-densidad observada, (más o menos la misma que la marcada en azul).

Hecho esto se vuelve a pinchar el mismo botón. Se abrirá un cuadro de diálogo en el que se nos

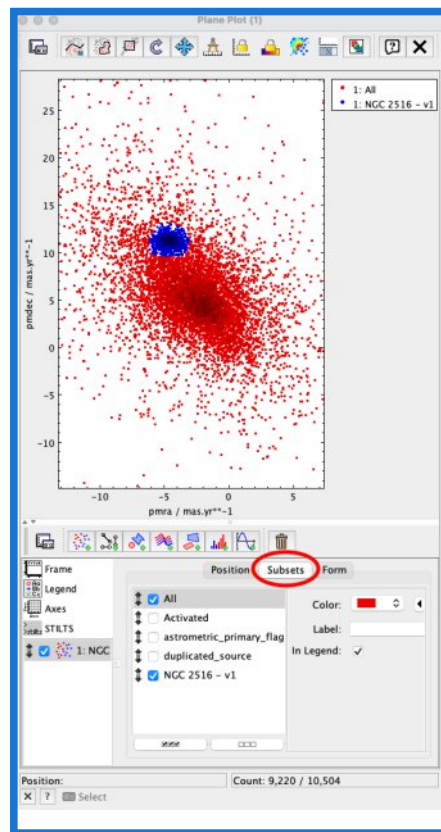
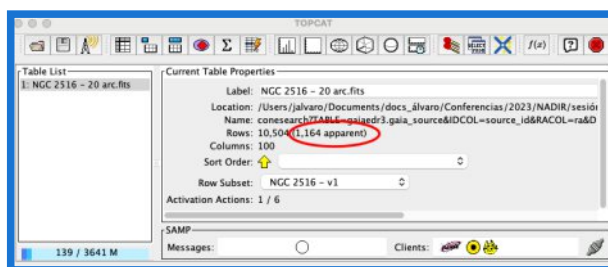


Fig. 10 – Selección del área con sobre-densidad.



También este grupo se puede enviar a Aladin con el tercer botón superior por la izquierda de esta ventana.

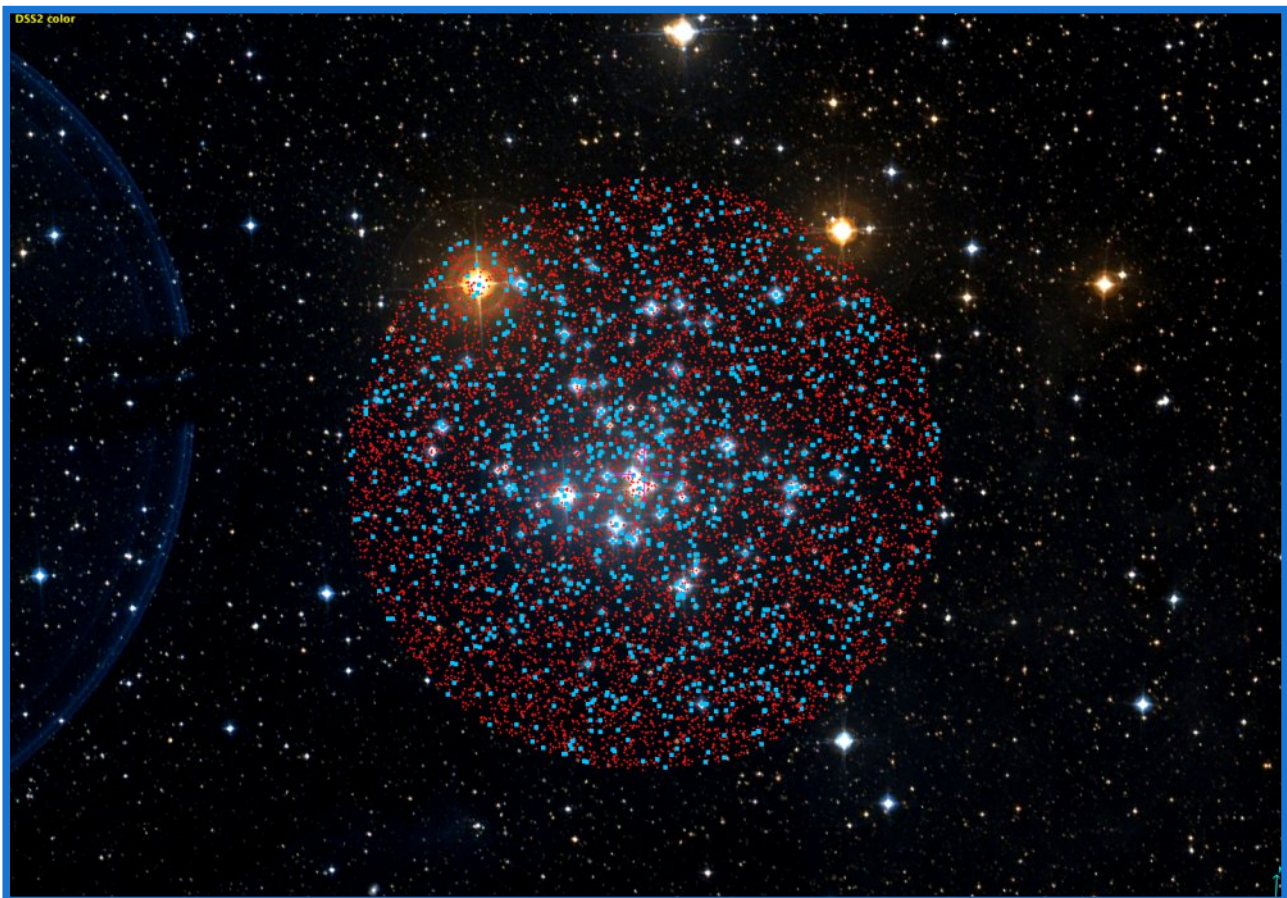
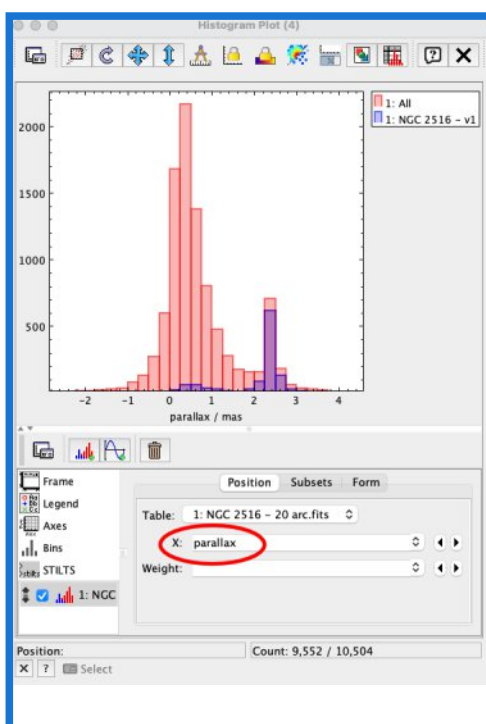


Fig. 11 – Representación en Aladin del subconjunto -v1- (en azul) superpuesto al conjunto inicial total.



7. Todavía puede haber estrellas en el campo visual del posible cúmulo que no formen parte del mismo. Es el momento de aplicar el segundo requisito planteado al principio: la distancia. Se asume que las estrellas de un OC están próximas entre sí y por tanto deben tener un valor de paralaje muy definido y con una dispersión pequeña en torno a un valor medio.

Nos centramos ahora sobre el campo del paralaje. En la ventana inicial de TOPCAT pedimos un histograma, (primero de los botones del grupo de gráficos), y seleccionamos el campo '*parallax*' en el eje X para la representación:

Si nos aseguramos de tener seleccionados ambos conjuntos (pestaña 'subset' en esta ventana) y ajustamos el zoom del área del gráfico con el ratón, deberíamos obtener algo similar a lo representado en esta figura.

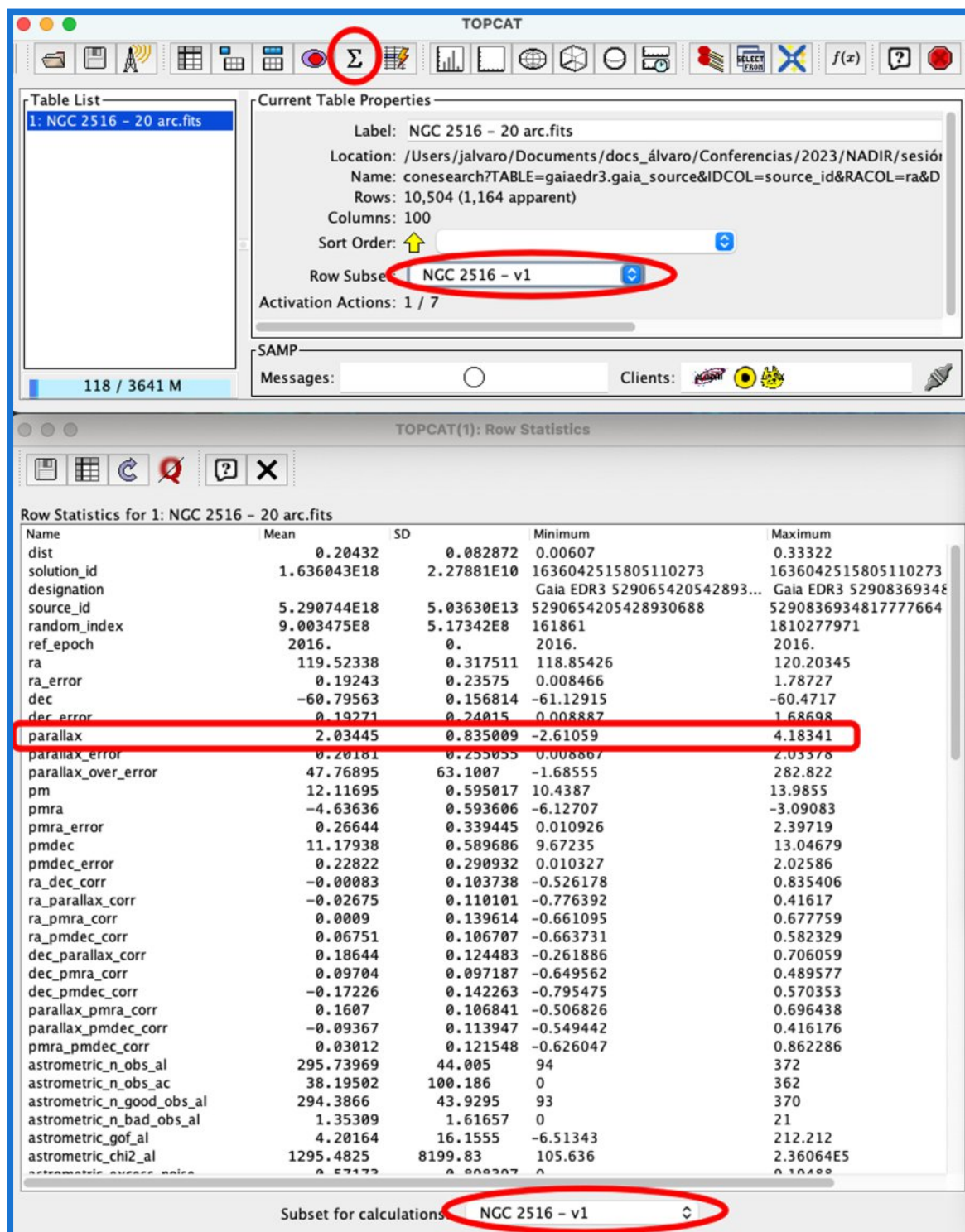
Fig. 12 – Distribución de paralajes del conjunto original (en rojo) y del subconjunto -v1- (en azul).

Como se puede ver hay una anomalía significativa respecto a lo que debería ser una distribución más o menos aleatoria respecto a un centro. Esta sobre-abundancia, también presente aquí y en torno a un valor de 2.4 mas, debe estar señalando el cúmulo que buscamos. Debemos estrechar de alguna manera su dispersión y, de paso, eliminar la fracción de estrellas que ‘se nos han colado’ en la selección anterior que aquí aparecen con un paralaje entre 0 y 1 mas.

Para ello volvemos de nuevo a la ventana inicial de TOPCAT, solicitamos estadística sobre el subconjunto -v1- y nos fijamos en los valores del campo ‘parallax’, (fig. 13).

Resulta evidente que no está centrado en los aproximados 2.4 mas y que su dispersión es grande.

Fig. 13 – Valores estadísticos en el subconjunto -v1-



Pasos que vamos a seguir:

a) En la ventana inicial de TOPCAT nos aseguramos de tener seleccionado el subconjunto -v1- y pinchamos sobre el botón 'display row subsets', (parece un 'ojo'). Se abrirá una nueva ventana, (fig. 14), sobre la que pinchamos en el primer botón '+', (definir nuevo subconjunto), que abrirá a su vez otra ventana:

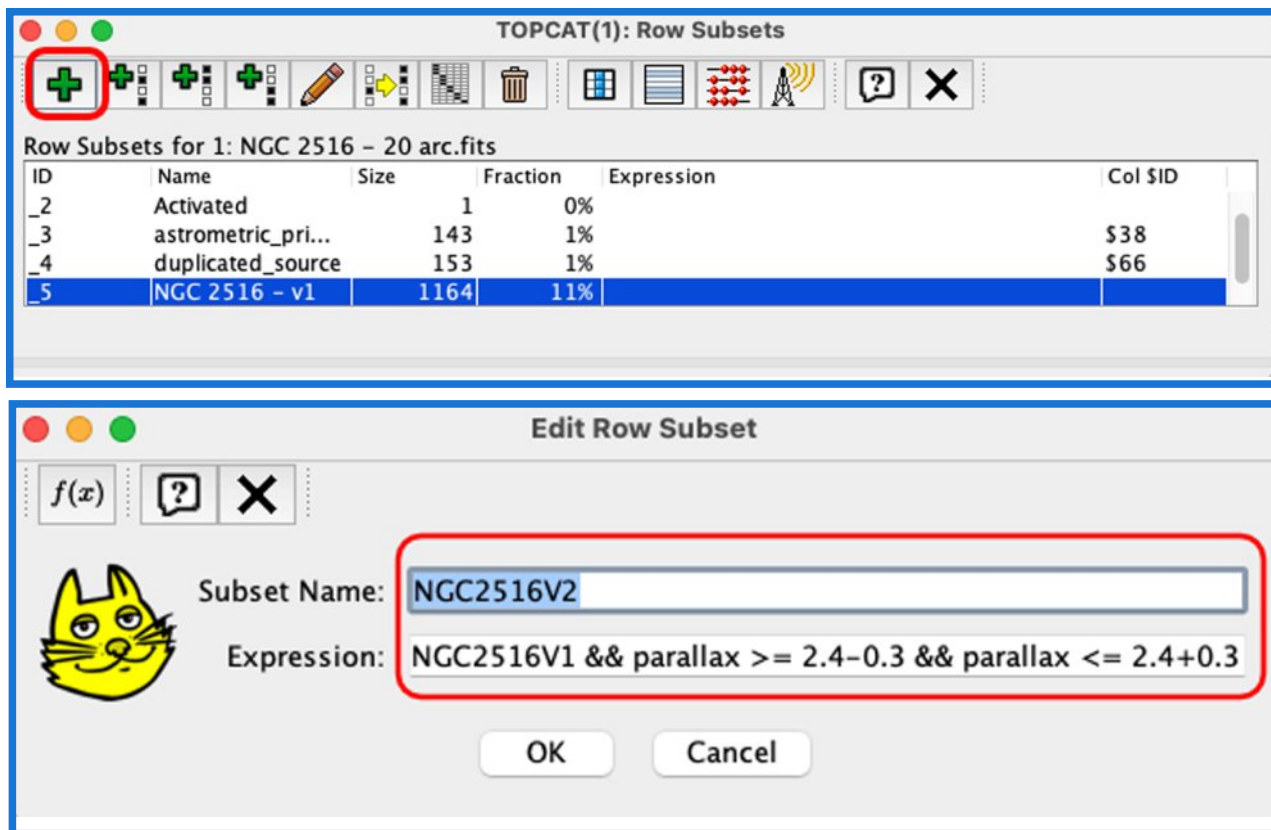


Fig. 14 – Ventanas TOPCAT para generar un nuevo subconjunto a partir de -v1-

Aquí daremos nombre al nuevo subconjunto y escribiremos el código con las condiciones que definirán este *subset*.

Lo que queremos es seleccionar, a partir del *subset -v1-*, sólo estrellas con un paralaje centrado en 2.4 mas y una dispersión máxima que podemos estimar en torno al 10 % de este valor. Se pueden ensayar otros límites, pero ésta es generalmente una buena aproximación. El código por tanto sería éste:

`NGC2516v1 && parallax >= 2.4-0.3 && parallax <= 2.4+0.3`

Tras el OK, tendremos el *subset -v2-* con 798 estrellas, (en este ejemplo), y un histograma de paralajes como el de la fig. 15.

Lo damos por bueno ..., es sólo un ejercicio.

Ya sólo nos queda un último paso. Analizar la edad de las estrellas de nuestra última selección.

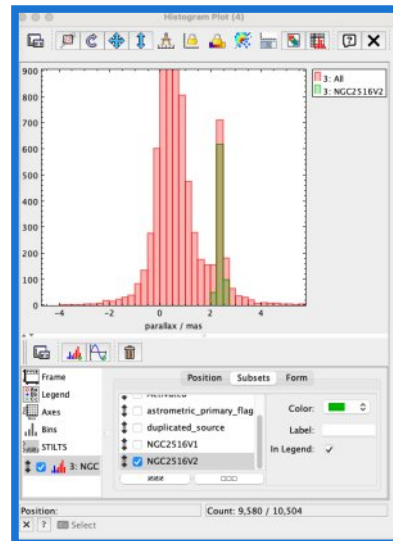


Fig. 15 – Histograma de paralajes del subconjunto -v2- frente al conjunto original.

8. En realidad lo que vamos a hacer es representar un diagrama H-R del subconjunto -v2-, utilizando el parámetro ‘exceso de color’, (diferencia entre las magnitudes fotométricas del azul y el infrarrojo), en sustitución de las temperaturas efectivas de las estrellas, frente a la magnitud fotométrica de Gaia en el visible (verde). Para ello podemos cerrar las ventanas de TOPCAT que molesten salvo la inicial y procederemos a solicitar un nuevo gráfico, (segundo botón del grupo de gráficos). Los ejes de este nuevo gráfico serían:

$$X = bp-rp, Y = phot_g_mean_mag$$

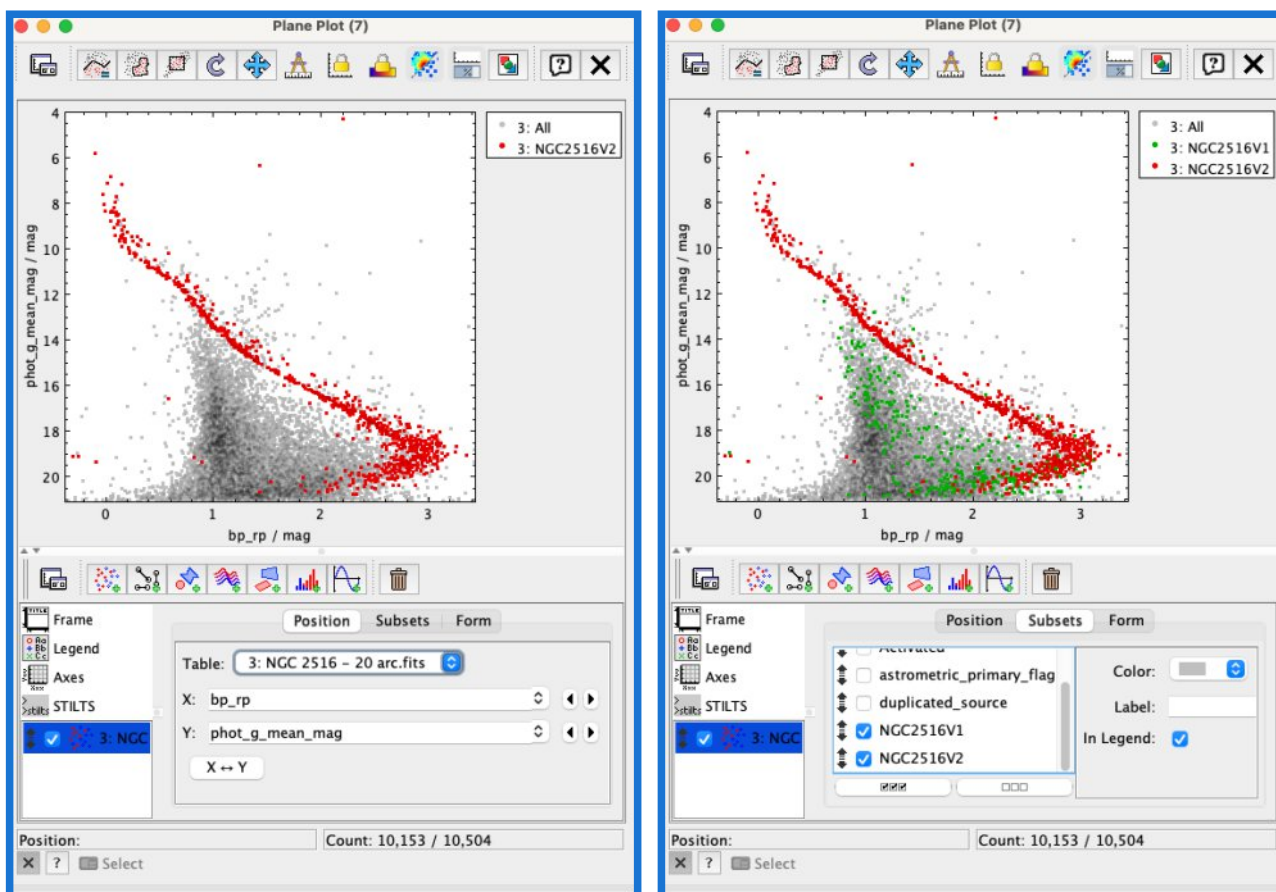


Fig. 16 – Diagramas ‘exceso de color’ frente a magnitud fotométrica. A la izquierda el subconjunto -v2- y a la derecha, ambos, -v1- y -v2-. De fondo el conjunto total inicial. El eje Y se ha invertido, con las magnitudes en sentido decreciente, más débiles abajo.

Lo que se observa en estos diagramas es un perfil bastante nítido de las estrellas finalmente seleccionadas, (candidatas a ser las componentes del cúmulo). Podría decirse que todas ellas están en la secuencia principal del diagrama H-R, apreciándose en la parte superior un incipiente punto de inflexión en el que empieza la derivación a la fase de gigantes rojas. Determinar la edad del cúmulo a partir de aquí, asumiendo que esta aproximación es razonablemente buena, es algo más complejo de lo que aquí podemos intentar. Habría que comparar estos datos con otros perfiles derivados de modelos matemáticos ad hoc y tener en cuenta a su vez otros parámetros como metalicidad, y otros indicadores de la composición química. Pero ya sabemos por la información de este cúmulo en el catálogo que tenemos como referencia que NGC 2516 es un cúmulo bastante joven, con poco más de 100 millones de años: $10^{8.052}$.

En la parte inferior del gráfico (fig. 16) hay un buen número de estrellas que parecen necesitar explicación por el hecho de ‘emborronar’ de alguna manera ese perfil nítido comentado. Hay que advertir aquí que hay literatura científica suficiente sobre Gaia, (de la propia ESA), advirtiendo del

empobrecimiento en la calidad de los datos fotométricos por encima de la magnitud 19. Podríamos haber filtrado también este conjunto eliminándolo directamente del conjunto original. Igualmente al análisis con movimientos propios se podría haber añadido un paso teniendo en cuenta las velocidades radiales, siempre y cuando este dato no presente demasiados registros con valores nulos.

En la fig. 16, derecha, se ven (en verde) las estrellas excluidas en el paso del subconjunto -v-1 al -v2-

Si comparamos los resultados obtenidos con los datos relativos a NGC 2516 en el catálogo de OCs que manejamos en el primer artículo, podemos sacar algunas conclusiones:

a) Por el número de componentes: hemos obtenido 798, frente a los estimados 770 del catálogo. Es evidente que sólo hemos hecho un ejercicio de aproximación, que podríamos haber refinado aún un poco y del que también se podrían haber eliminado de partida las estrellas con magnitud superior a la 19. El resultado es muy aceptable.

b) Si atendemos a la distancia media al cúmulo, la estadística de nuestro -v2- indica un paralaje medio de $2.41 \text{ mas} \pm 0.10 \text{ mas}$. En el catálogo se indica 2.44 mas .

c) En relación a los movimientos propios, los datos catalogados son:

pmRA: -5.43 mas/yr

pmDEC: 8.41 mas/yr

nuestros resultados aquí difieren un poco más:

pmRA: -4.64 mas/yr ,

pmDEC: 11.25 mas/yr

Nuevamente aquí hay que ver la conveniencia de rehacer el análisis eliminando las estrellas más débiles, ($\text{mag} > 19$), aunque en este caso los datos obtenidos en nuestro trabajo deben considerarse más fiables que los del catálogo, ya que hemos trabajado con los datos astrométricos de GAIA/EDR3, mucho más precisos que los existentes a la fecha de publicación del catálogo de referencia.

d) Y ya como imagen final, podemos llevar a Aladin las estrellas del *subset -v2-* que quedan representadas en amarillo en la fig. 17.

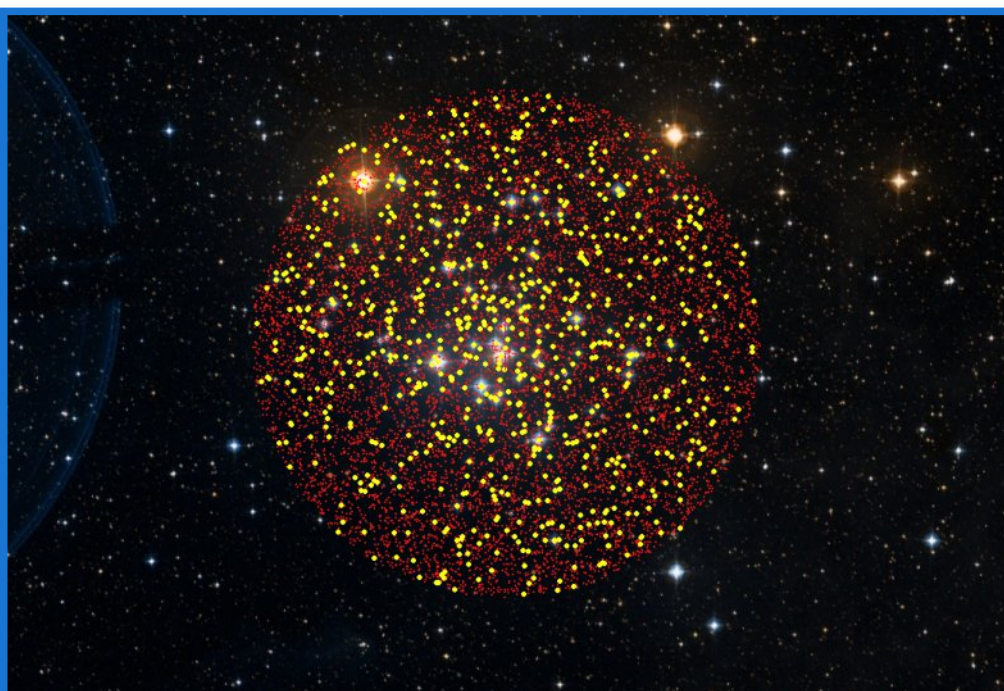


Fig. 17 – Estrellas consideradas finalmente como componentes del cúmulo NGC 2516, (en amarillo), a partir de Aladin. Esta imagen sugiere que debería ampliarse todavía un poco el radio del campo de observación, (20 minutos de arco en nuestro ejercicio), ya que parece haber demasiadas estrellas del cúmulo en el borde del área seleccionada.

9. Antes de dar por terminado el proceso y cerrar TOPCAT, conviene guardar todas las tablas de la sesión, (original, v1 y v2), por si fuera necesario rehacer el análisis en algún momento. (Menú principal: *File/Save table(s)/Session*, hay varias opciones en cuanto al formato. Después pueden abrirse de nuevo en TOPCAT sin necesidad de volver a descargar el conjunto de Gaia y/o repetir pasos ya dados).

Segundo ejercicio.- La propuesta para este ejercicio es repetir todo lo hecho con NGC 2516 en el primer ejercicio, pero en este caso sobre **NGC 2682 (M67)**. Aquí sólo se ofrecen unos posibles resultados finales como guía.

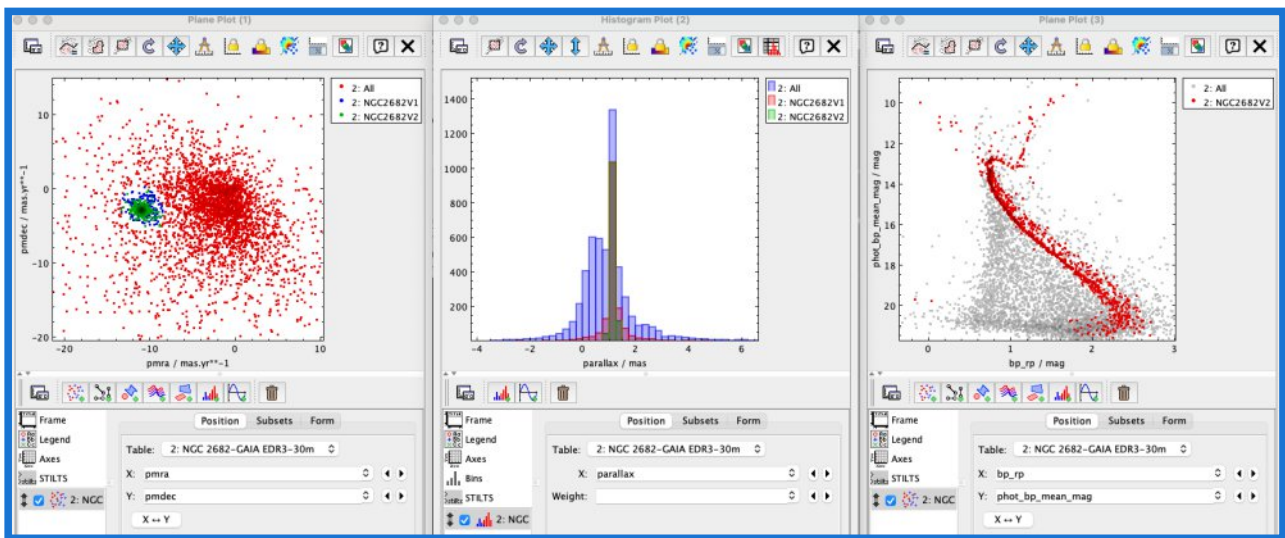


Fig. 18 – Diferentes fases del proceso de caracterización de NGC 2682 (M67) en TOPCAT. El perfil en el diagrama de magnitudes presenta el punto de giro hacia la rama de las gigantes rojas, donde ya hay componentes, más evidente que en el cúmulo anterior. Este cúmulo tiene una edad de unos 2.800 millones de años, (véase el catálogo: 109.45 años).

El número de estrellas resultantes es ahora de 1.192 con un valor medio de paralaje de $1.15 \text{ mas} \pm 0.076$ y en movimientos propios, $pmra = -10.95$ y $pmdec = -2.91 \text{ mas/yr}$.

Algunas conclusiones.- Es seguro que este proceso que hemos seguido puede haber parecido laborioso, e incluso bastante farragoso en ciertos momentos, pero eso ocurre por desconocimiento o falta de dominio de la herramienta utilizada. A medida que se trabaja con TOPCAT y los conceptos en sí mismos empiezan a resultar familiares, este análisis puede hacerse en pocos minutos y puede emplearse algo más de tiempo en ser más exigente y riguroso en la tarea.

Lo que sí se hace evidente es una cierta dosis de ‘*artesanía*’, que depende en buena medida de la experiencia del investigador, como cuando se hace la selección del área de sobre-densidad en movimientos propios o se van aplicando criterios en las diferentes etapas de filtrado. Un proceso así se define como **supervisado** y **parametrizado**, que significa que la actuación sobre los parámetros a tratar es necesaria y necesita la vigilancia y decisión del analista.

En el próximo y último artículo de la serie veremos una nueva herramienta, **Clusterix**, también directamente conectada con TOPCAT, y que se autodefine como no supervisada y no parametrizada (¿?). Veremos entonces en qué medida se simplifica el trabajo.

3ª PARTE

CLUSTERIX



NGC 2682 – M67 – Cúmulo estelar abierto situado a 2.900 a-l, compuesto por unas 1.000 estrellas con una edad estimada de 3.600 millones de años. Imagen de Aladin.

El procedimiento seguido en el apartado anterior para caracterizar cúmulos estelares abiertos es un proceso supervisado y parametrizado, con todos los condicionantes que esto conlleva y que, como veíamos entonces, implica una cierta dosis de artesanía mezclada con alguna experiencia previa y/o conocimiento del OC.

En esta tercera parte haremos uso de una nueva herramienta que ayudará bastante en las etapas iniciales del análisis. Se trata de Clusterix 2.0 : “A Virtual Observatory tool to estimate cluster membership probability”. Se define como una herramienta interactiva, (vía navegador web), y no parametrizada, para la determinación de las probabilidades de ser miembros (estrellas) pertenecientes a un cúmulo estelar, basada en los datos sobre movimientos propios.

I Parte.-

Acceso: <http://clusterix.cab.inta-csic.es/clusterix/>

Clusterix 2.0

Clusterix 2.0 is an interactive web-based application to calculate the grouping probability of a list of objects using proper motions and the non parametric method in the approach described in Balaguer-Nuñez et al (2020). It also allows the possibility of gathering physical parameters (parallaxes, radial velocities, proper motions,...) from Vizier and estimating effective temperatures, surface gravities and metallicities using VOSA.

An step-by-step example on how to use Clusterix can be found [here](#)

NOTE: When working with Clusterix, it is recommendable to clear cache to avoid caching problems. Clusterix has been tested in
in
Chrome (109.0.5414.74) and Firefox (109.0)

Step 1/3: Information gathering (coordinates and physical parameters)

Search by Id

ID

Radius deg

Catalogue GAIA/EDR3

RUWE ? ≤

Magnitude limits (min/max)

≤ G ≤

Search

Search by Coordinates

RAJ2000(deg),DEJ2000(deg)

Radius arcmin

Catalogue GAIA/EDR3

RUWE ? ≤

Magnitude limits (min/max)

≤ G ≤

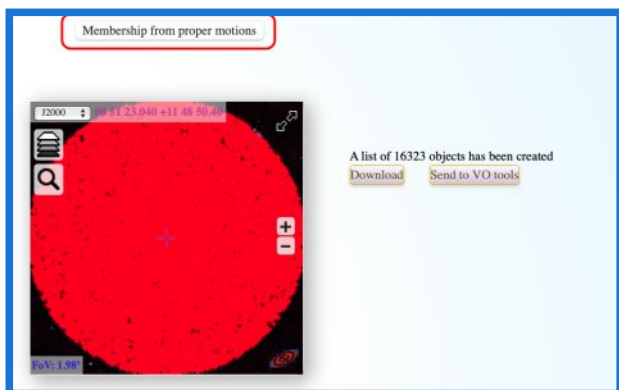


Fig. 1 – Izquierda, pantalla inicial de búsqueda. Arriba, resultado de la búsqueda.

La pantalla inicial (fig. 1) nos pide la región del cielo que se quiere explorar con dos opciones: búsqueda por coordenadas, o búsqueda por el nombre del objeto (entiéndase OC). En ambos casos hay que especificar el radio o anchura del cono angular centrado en las coordenadas del objeto de interés. También se puede seleccionar el catálogo al que se quiere acceder, (utilizaremos aquí GAIA/EDR3) y, por defecto, se asume el parámetro RUWE 1.4, para descartar objetos extensos –sólo nos interesan estrellas–. Igualmente se pueden acotar la búsqueda poniendo límites (min/máx) en los valores de la magnitud fotométrica.

En el ejemplo que vamos a desarrollar nos centraremos en el cúmulo NGC 2682 (M 67), imagen de portada. En el formulario de búsqueda pondremos ID: M67 y el radio de apertura: 1º (60 arcmin) . Una vez activado el botón de búsqueda, Clusterix nos devuelve el número de estrellas encontradas en este campo centrado en las coordenadas del mismo, (fig. 1 derecha). En este caso son 16.323.

Ahora podemos pasar al paso 2 pinchando en el botón ‘miembros por movimientos propios’, (recuadrado en rojo en la imagen). Y aquí, a pesar de que Clusterix se considera un procedimiento no parametrizado, vemos que aún debemos establecer algunos criterios de trabajo dando ciertos valores a una serie de campos, (fig. 2).

El bloque remarcado por ‘1’ en la figura 2 es clave para iniciar el proceso de análisis de frecuencias sobre movimientos propios en los objetos catalogados. En este momento hay que definir tres áreas y escribir en los tres campos a rellenar tres parámetros separados por ‘comas’ y finalizando en cada caso con ‘punto y coma’. Estos parámetros son: 1) ascensión recta (en grados), 2) declinación (también en grados y 3) radio de apertura (igualmente en grados).

En los tres casos, la ascensión recta y declinación tienen los mismos valores y corresponden al centro del campo de visión a analizar. Esto significa que debemos conocer cuáles son estas coordenadas previamente, (en este caso para M67). Por esta razón, ya que finalmente es necesario, es siempre preferible optar por hacer la selección en el ‘paso 1’ a partir de la ventana de búsqueda ‘por coordenadas’ recurriendo bien al fichero de Vizier o al atlas Aladin, por ejemplo.

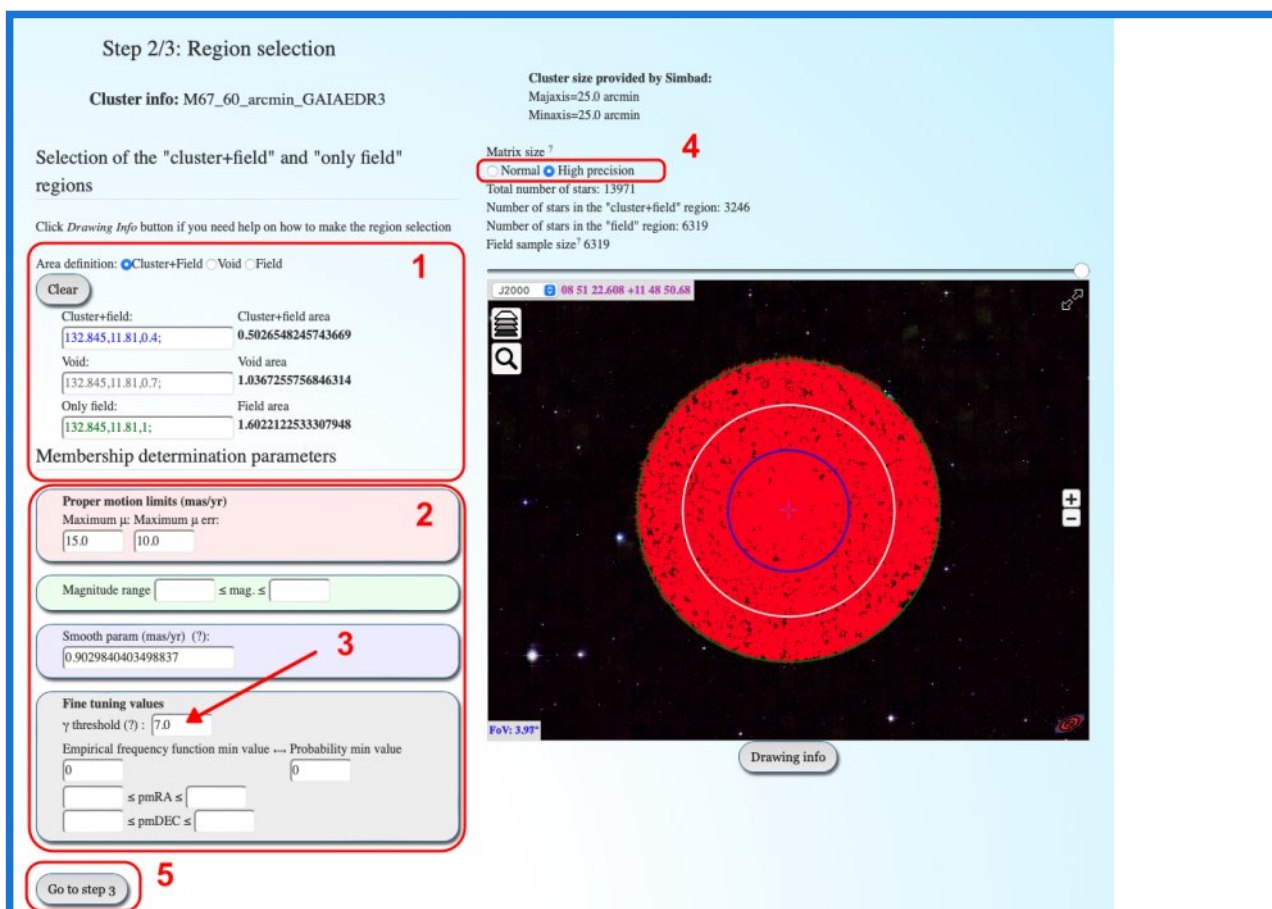


Fig. 2 – Clusterix 2.0 - paso 2/3.

[1] Micro-curso sobre TOPCAT del SVO (en español): <https://www.youtube.com/watch?v=Ij3qeliOOGY>

[2] Micro-curso sobre Aladin del SVO (en español): <https://www.youtube.com/watch?v=Ajro94X7kUc&t=2498s>

El tercero de los valores (radio de apertura) en cada caso es diferente. Estos radios definen tres áreas cuyo significado, por orden, es: a) área donde se presume que está contenido el cúmulo, b) área vacía que se excluye del análisis de frecuencias para evitar contaminación si todavía contuviera estrellas marginales del cúmulo, y c) área del campo del cielo donde ya no se esperan estrellas del cúmulo y que servirá de referencia a Clusterix para el cálculo de frecuencias. Como se puede ver en la parte derecha de la figura 2, cada uno de estos radios da lugar al círculo correspondiente a las respectivas áreas, (azul, blanco y verde), que se van dibujando a medida que se rellenan los valores del bloque 1.

Esta definición de áreas es crítica. Si no se hace correctamente, los resultados pueden no ser efectivos y habría que hacer sucesivos intentos probando con valores de radios distintos, especialmente el que marca la diferencia entre el primer círculo y el segundo. El tercero, (verde), normalmente será el límite del campo con objetos determinado en el paso 1, en nuestro ejemplo 1º.

En el bloque 2 (fig. 2) se puede ajustar aún el valor de algunos parámetros, restringiendo algunos límites como los relativos a los movimientos propios y magnitudes. Por defecto el campo threshold se sitúa en 3.0, pero puede cambiarse por otros como 5, 6, 7 Las ayudas (?) son contextuales y ofrecen alguna información adicional.

Finalmente se puede también decidir sobre el tamaño de la matriz de trabajo ('4' en la fig. 2), eligiendo entre una resolución estándar (Normal) o de alta precisión (High precision).

Completados estos requerimientos Clusterix inicia por su cuenta el análisis y en breves segundos presenta resultados:

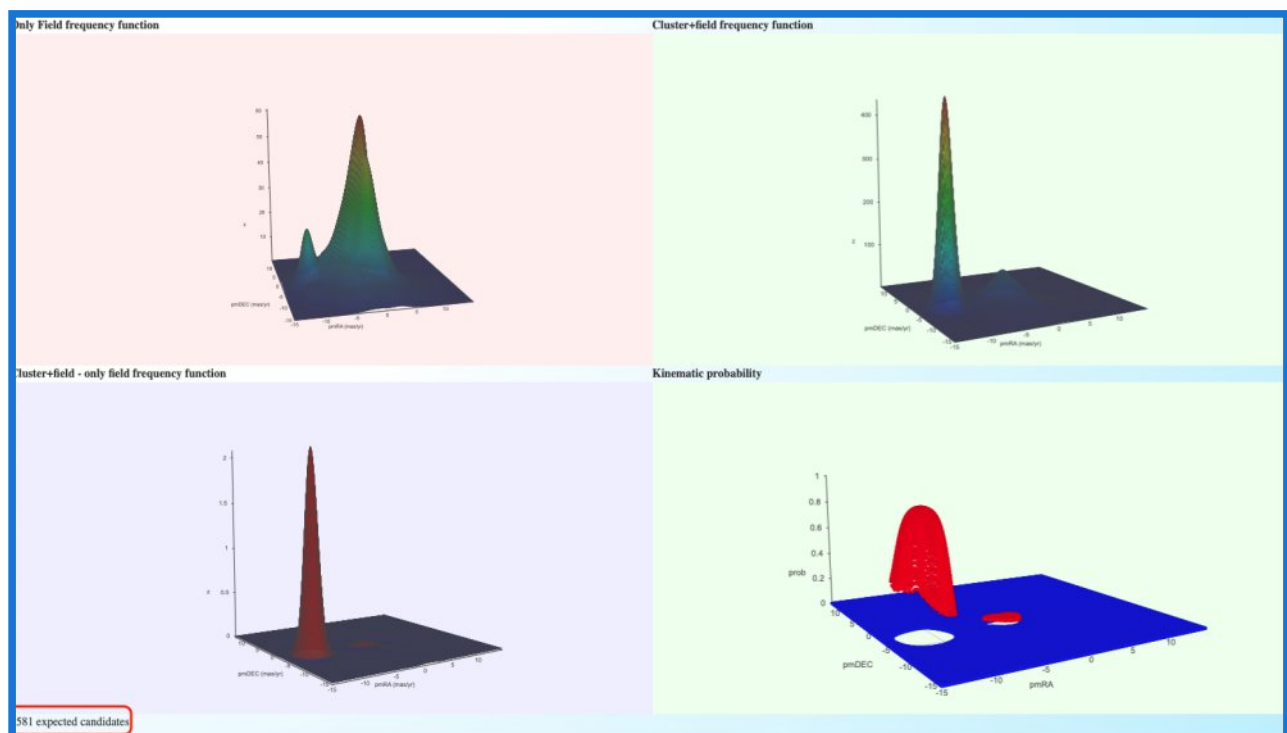


Fig. 3 – gráficos de resultados del análisis de frecuencias.

Estas cuatro ventanas son independientes y, en todas ellas, puede ser cambiada la perspectiva y/o tamaño con el movimiento del ratón presionado sobre ellas. La que más nos interesa es la de abajo/derecha, que presenta un campo de probabilidad asociado a cada una de las estrellas del conjunto inicial. Esto significa que se han encontrado correlaciones sobre movimientos propios, que es lo que se buscaba. Si este campo de probabilidad no aparece, o resultara poco significativo a simple vista, habría que volver a redefinir las áreas de análisis y, tal vez, algunos otros parámetros.

En la esquina inferior izquierda, Clusterix nos indica un número de miembros estimado para el cúmulo después del análisis. En este caso, 1.581 estrellas.

Si damos por buena la aproximación atendiendo a los resultados nítidos de un campo de probabilidad bien localizado, pasaríamos al 'paso 3' pinchando sobre el botón correspondiente: marcado con '5' en la figura 2. Esto nos lleva a una nueva ventana donde se presenta un resumen del flujo de trabajo seguido y los 50 primeros objetos, en formato tabla, ordenados por el valor del campo PROB (probabilidad). Ver figura 4.

```

Step 3/3. Determination of membership probabilities
Send to VO tools Save as text

# 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      identifier of star retrieved from the input data
# RAJ2000     right ascension of a star
# DECJ2000    declination of a star
# pmRA        proper motion in right ascension
# epmRA       error in proper motion in right ascension
# pmDEC       proper motion in declination
# epmDEC      error in proper motion in declination
# P11         Parallax
# eP11        error in parallax
# BP          BP magnitude
# eBP         error in BP magnitude
# RP          RP magnitude
# eRP         error in RP magnitude
# G           G magnitude
# eG          error in G magnitude
# RV         Radial velocity
# eRV        error in radial velocity
# Source_ID  Unique Source Identifier
# R11ME      re-normalised unit weight error
# PROB       probability that star belongs to evaluated open cluster
# FLAG       W=Membership  N=No membership
#
# 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

```

Fig. 4 – recorte de la pantalla del paso 3/3.

En este punto podemos guardar esta información y la tabla de objetos en formato texto, pero es más interesante enviar los datos a las herramientas ya conocidas, TOPCAT y Aladin, mediante el botón remarcado en la imagen.

Como ya sabemos, antes de hacerlo se deben abrir ambas aplicaciones y ya después se hace el envío desde Clusterix. Todo lo que sigue lo haremos sobre TOPCAT y Aladin, igual que se describió en el segundo artículo de la serie.

Si todo ha funcionado correctamente, nos encontramos de nuevo con una tabla que se trabajará con TOPCAT, (fig. 5), y en la que, de entrada, podemos ver que está compuesta por los 16.323 objetos iniciales.

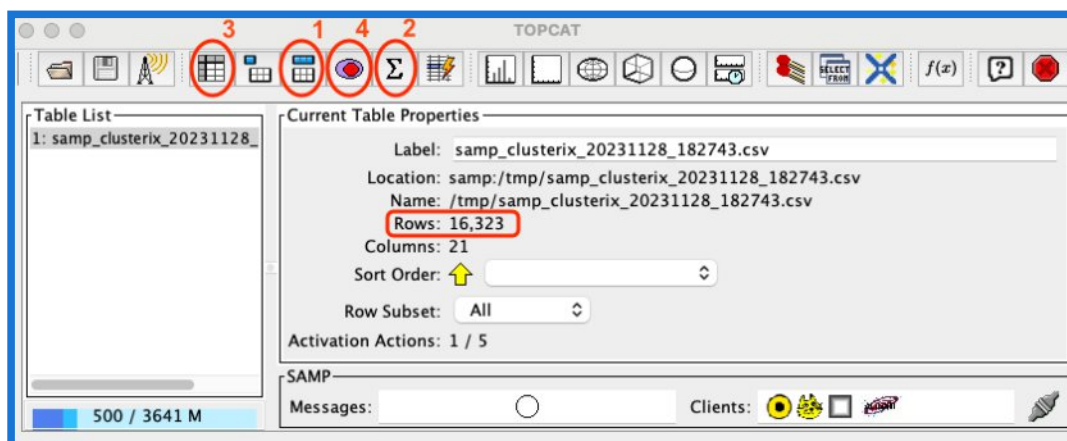


Fig. 5 – tabla incorporada a TOPCAT desde Clusterix

Bueno ..., y ¿dónde están los 1.581 objetos propuestos por Clusterix como potenciales miembros del cúmulo?

Si analizamos el contenido de la tabla transferida, (botón 1 – fig. 5), podemos ver que a los datos recuperados de GAIA se ha añadido un nuevo campo: PROB, que es el que ha calculado Clusterix para cada objeto (fig. 6a). Igualmente, con el botón 2 (fig. 5), se puede explorar la estadística de los diferentes campos de la tabla. Para el campo ‘PROB’ hay un valor máximo de 0.85494 y un mínimo de 0., con una media de 0.11527, (fig. 6b). Podemos explorar el contenido real de la tabla a través del botón 3 (fig. 5). Estos datos están ordenados por el valor de PROB de mayor a menor, y se puede constatar que los 1.581 objetos seleccionados por Clusterix tienen todos un valor PROB > 0.843, (fig. 7)^[3].

Index	Visible	Name	SID	Class	Description	Datatype	NULL VALUE
0		Index	\$0	Long	Table row index		
1	<input checked="" type="checkbox"/>	STAR_NO	\$1	Short		short	-32768
2	<input checked="" type="checkbox"/>	RAJ2000	\$2	Double		double	
3	<input checked="" type="checkbox"/>	DECJ2000	\$3	Double		double	
4	<input checked="" type="checkbox"/>	RA_PM	\$4	Float		float	
5	<input checked="" type="checkbox"/>	eRA_PM	\$5	Float		float	
6	<input checked="" type="checkbox"/>	DEC_PM	\$6	Float		float	
7	<input checked="" type="checkbox"/>	eDEC_PM	\$7	Float		float	
8	<input checked="" type="checkbox"/>	PLL	\$8	Float		float	
9	<input checked="" type="checkbox"/>	ePLL	\$9	Float		float	
10	<input checked="" type="checkbox"/>	BP	\$10	Float		float	
11	<input checked="" type="checkbox"/>	eBP	\$11	Double		double	
12	<input checked="" type="checkbox"/>	RP	\$12	Float		float	
13	<input checked="" type="checkbox"/>	eRP	\$13	Double		double	
14	<input checked="" type="checkbox"/>	Gmag	\$14	Float		float	
15	<input checked="" type="checkbox"/>	eGmag	\$15	Double		double	
16	<input checked="" type="checkbox"/>	RV	\$16	Float		float	
17	<input checked="" type="checkbox"/>	eRV	\$17	Float		float	
18	<input checked="" type="checkbox"/>	GAIADR2_SOURCE_ID	\$18	Long		long	-9223372036854775807
19	<input checked="" type="checkbox"/>	PROB	\$19	Double		double	
20	<input checked="" type="checkbox"/>	FLAG	\$20	String		char	
21	<input checked="" type="checkbox"/>	Region	\$21	Character		char	

Fig. 6a – contenido

Name	Mean	SD	Minimum	Maximum
STAR_NO	8162.	4712.04	1	16323
RAJ2000	132.82695	0.488785	131.82461	133.86379
DECJ2000	11.80825	0.482003	10.81515	12.813
RA_PM	-4.17171	9.5476	-229.438	163.074
eRA_PM	0.45499	0.538091	0.013	3.28
DEC_PM	-4.84892	9.96686	-387.849	111.493
eDEC_PM	0.32137	0.37015	0.01	2.726
PLL	0.92472	1.24862	-6.1096	25.461
ePLL	0.41349	0.462695	0.0121	3.5925
BP	18.98415	2.28812	6.92644	22.7605
eBP	0.06465	0.063721	0.0028	1.51276
RP	17.48861	2.04352	4.49649	21.0746
eRP	0.03838	0.036145	0.00378	1.23897
Gmag	18.31524	2.16497	5.63165	21.0078
eGmag	0.00433	0.002116	0.00276	0.0284
RV	24.47854	28.5746	-47.62	153.59
eRV	1.34876	1.95143	0.14	16.99
GAIADR2_SOURCE_ID	6.833799E17	3.14182E15	598613540781594752	6081442836
PROB	0.11527	0.279764	0.	0.85494
FLAG			M	NM

Fig. 6b – resumen estadístico

eGmag	RV	eRV	GAIADR2_SOURCE_ID	PROB	FLAG	Region
1569	0.00293		604968133874491392	0.84422	M	C
1570	0.00538		605026030034504576	0.84422	M	F
1571	0.00276		604918522708195584	0.84397	M	C
1572	0.00276		604933400474853248	0.84347	M	V
1573	0.00511		604910924910708224	0.84347	M	C
1574	0.0029		604898246166933504	0.84336	M	C
1575	0.00281		604971466769552128	0.84329	M	C
1576	0.00534		598887770148211584	0.84318	M	V
1577	0.00424		604731807594113408	0.84306	M	V
1578	0.00538		604928036059785088	0.84306	M	V
1579	0.00431		601970388436281984	0.84301	M	F
1580	0.00345		604906531159506304	0.84301	M	C
1581	0.00417		604695012608940672	0.84301	M	V
1582	0.00498		604897902569851264	0.84292	M	C
1583	0.00415		604687453466527744	0.84292	M	F
1584	0.00946		604963941987041280	0.84292	M	C
1585	0.00311		604702571751893760	0.8429	M	C
1586	0.00517		608016254984805504	0.8429	M	F
1587	0.00323		604682643103730432	0.8429	M	F
1588	0.00478		598674220084691200	0.84285	M	F

Fig. 7 – datos de la tabla

[3] En muchos otros casos la proyección de probabilidades de Clusterix no es tan clara. En esa situación una buena aproximación puede ser tomar los objetos cuya probabilidad esté por encima de la media.

II Parte.-

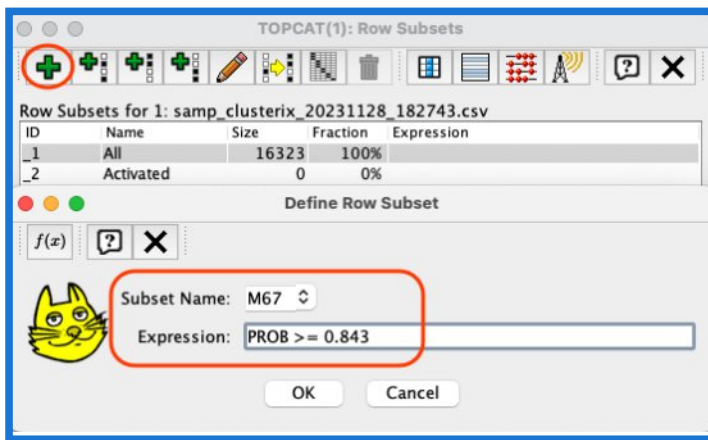


Fig. 8 – se crea el subconjunto M67.

A partir de aquí trabajaremos el resultado final con TOPCAT siguiendo los pasos ya vistos en el artículo 2. Lo primero que debemos hacer es definir el subconjunto con los 1.581 objetos propuestos por Clusterix. Para ello creamos un nuevo dataset con el botón 4, (fig. 5).

Hecho esto, lo representamos en el espacio de movimientos propios. (Figura 9a). Y después vemos el histograma de paralajes, (figura 9b).

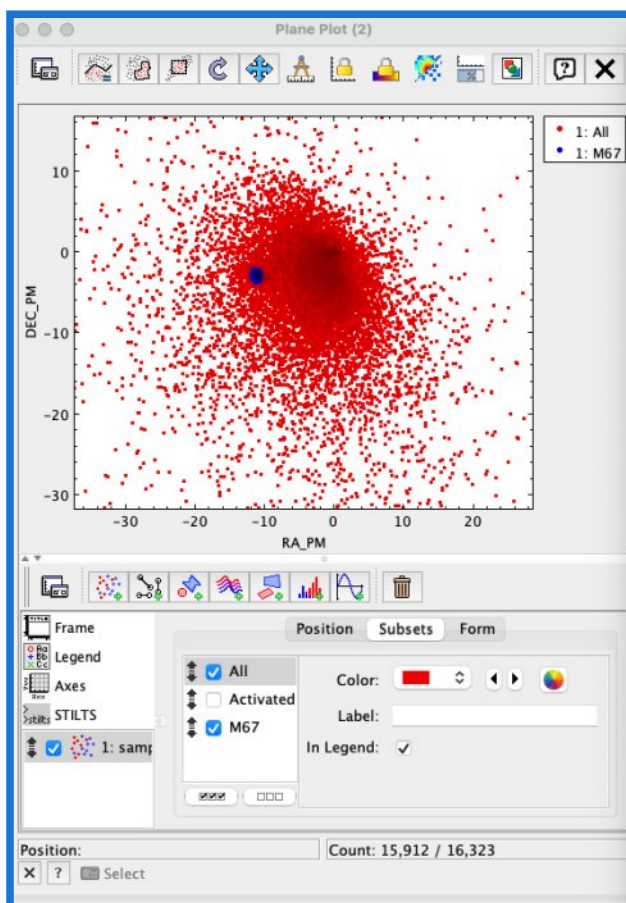


Fig. 9a – movimientos propios

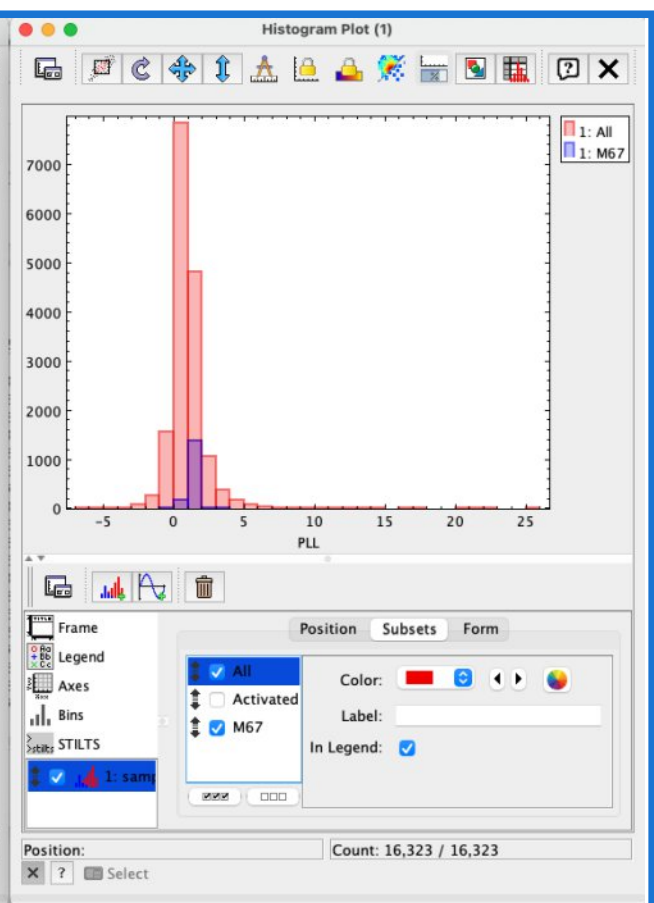
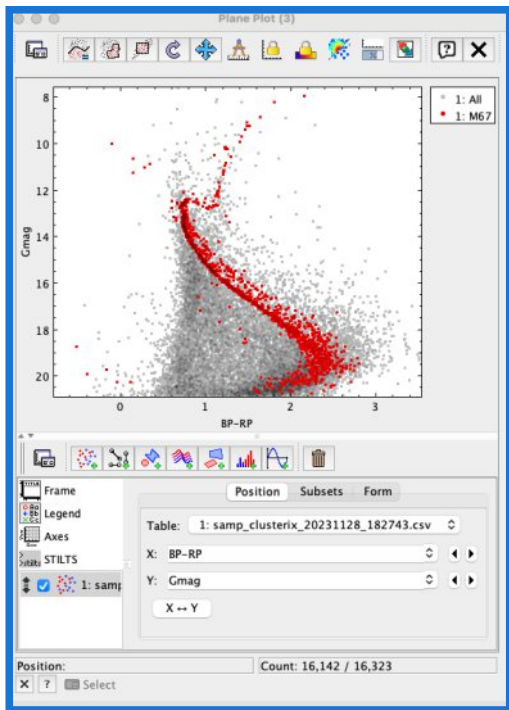


Fig. 9b – histograma de paralajes

En la figura 9a vemos localizada la sobre-densidad atribuida al cúmulo M67, que sin la ayuda de Clusterix hubiéramos tenido que ‘recortar a mano’. En la imagen de la derecha (fig. 9b) se aprecia también la contribución de las estrellas del cúmulo en la forma del histograma de ‘paralajes’. El paralaje está directamente relacionado con la distancia de los objetos analizados: $d=1000/\text{parallax}$ (recuérdese que el paralaje viene expresado en milisegundos de año por año, lo que nos da distancias en parsecs). No tienen sentido por tanto los valores negativos que aparecen en el histograma. Se deben a datos erróneos que habría que descartar.



Ya podemos representar el diagrama ‘exceso de color’ frente a ‘magnitud fotométrica’ y con esto habríamos terminado la caracterización del cúmulo, (fig. 10).

Fig. 10 – diagrama color/magnitud.

En este diagrama se aprecia claramente el perfil del cúmulo con la mayoría de las estrellas en la secuencia principal y un punto de inflexión nítido con estrellas que derivan hacia la región de gigantes rojas. Sin embargo también resulta evidente una cierta cantidad de ruido que hace aconsejable refinar un poco las condiciones de la caracterización de M67.

Para ello crearemos un nuevo filtro restringiendo las estrellas con paralaje a sólo una distancia de 1 sigma a uno y otro lado del valor medio del subconjunto M67, (fig. 11).

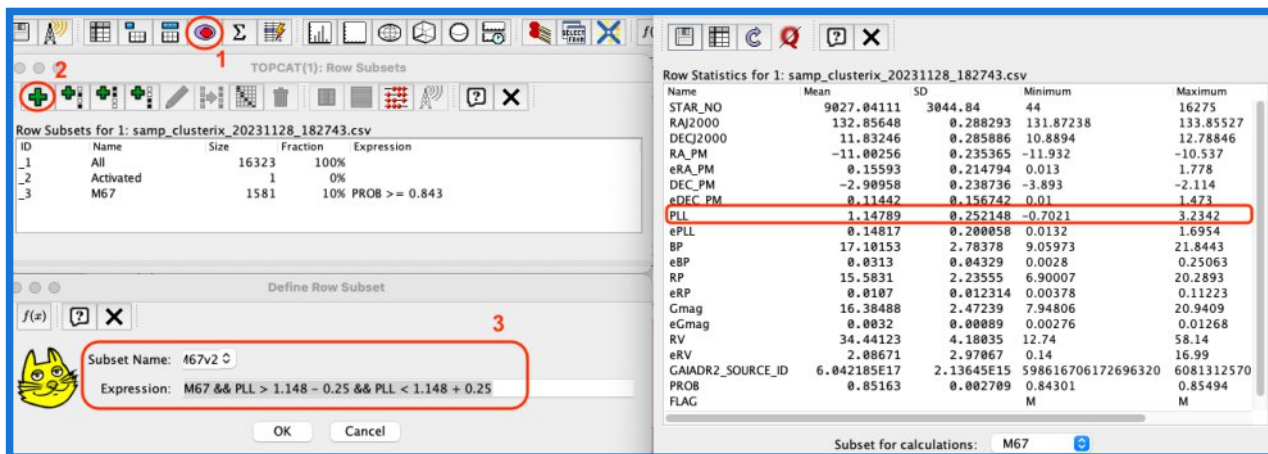


Fig. 11 – se crea un nuevo dataset (M67v2) con la ayuda del sumario estadístico para M67 restringiendo los valores del campo PLL (paralaje) a ± 1 del valor medio.

Pero ya puestos, también podemos ser un poco más estrictos y eliminar del último subconjunto los objetos con un error relativo en los valores de movimientos propios superior al 10%.

ID	Name	Size	Fraction	Expression
1	All	16323	100%	
_2	Activated	1	0%	
_3	M67	1581	10%	PROB >= 0.843
4	M67v2	1356	8%	M67 && P11 > 1.148 - 0.25 && P11 < 1.148 + 0.25
_6	M67v3	1288	8%	M67v2 && RA_PM/eRA_PM < -10 && DEC_PM/eDEC_PM < -10

Fig. 12 – filtro para un nuevo dataset (M67v3) donde se restringe la pertenencia al cúmulo a sólo los objetos con un error relativo en movimientos propios menor del 10%. (El signo '-' se debe a que los valores de RA_PM y DEC_PM son negativos).

Así, finalmente, se tendría para M67:

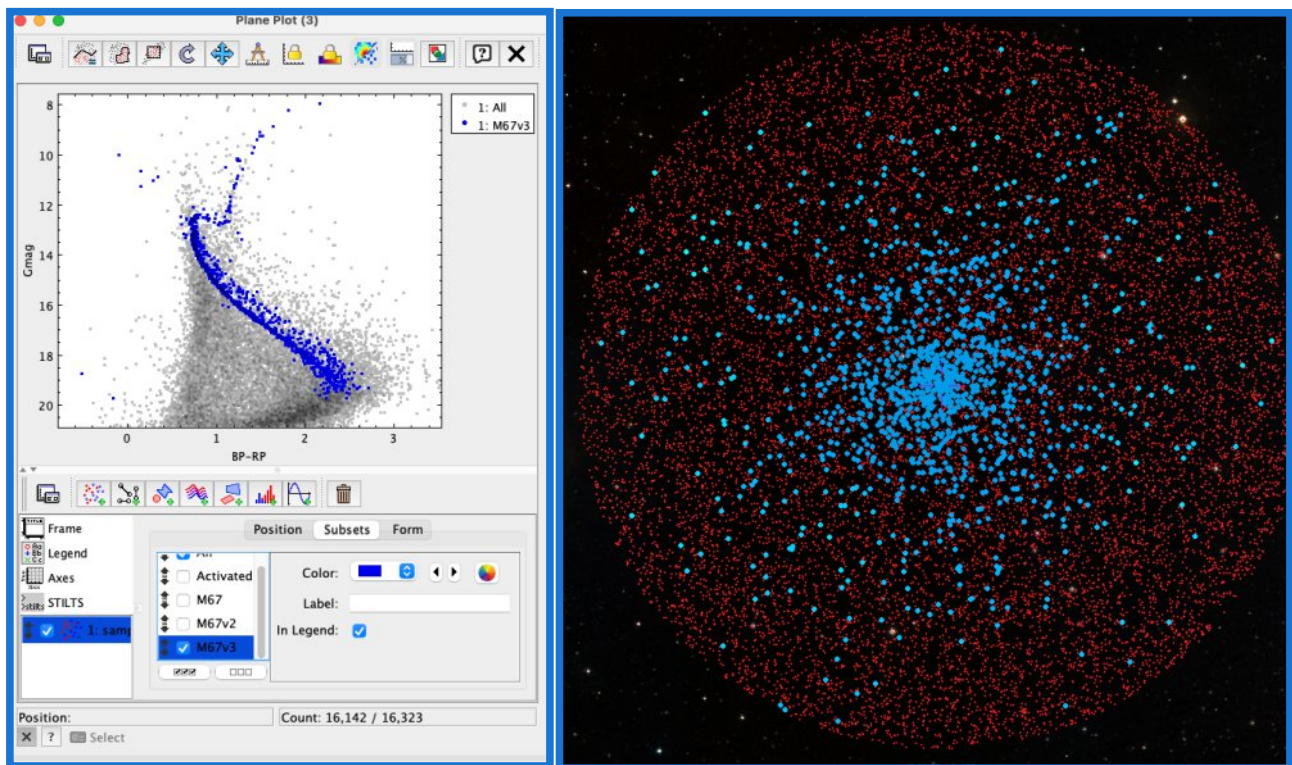


Fig. 13. M67 versión 3. Diagrama color/magnitud y representación del cúmulo en Aladin (azul) frente a las estrellas de fondo (rojo).

Para terminar se propone realizar un nuevo ejercicio con el cúmulo NGC 2516, de coordenadas AR = 119.52° y DEC = -60.75°, y un radio de 90 arcmin.

Se recomienda utilizar los radios de apertura 0.7°, 1.2° y 1.5° respectivamente en Clusterix.



Fig. 14 – NGC 2516 – [Aladin].

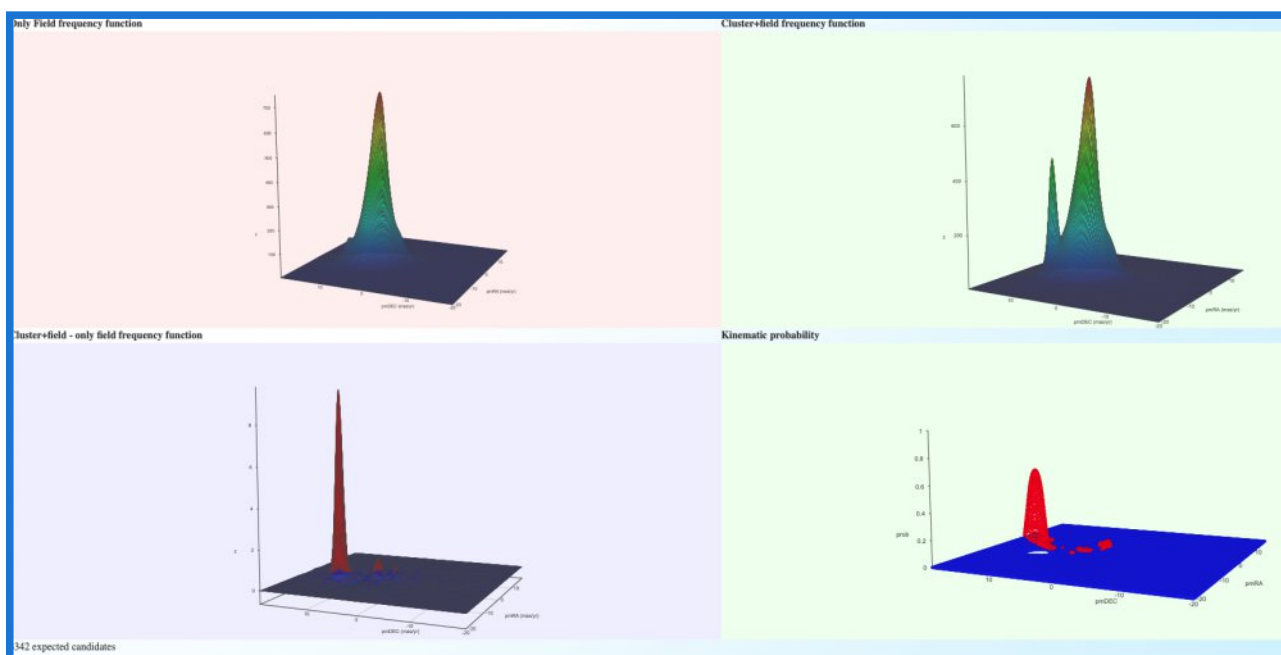


Fig. 15 – NGC 2516 – Análisis de frecuencias en Clusterix.

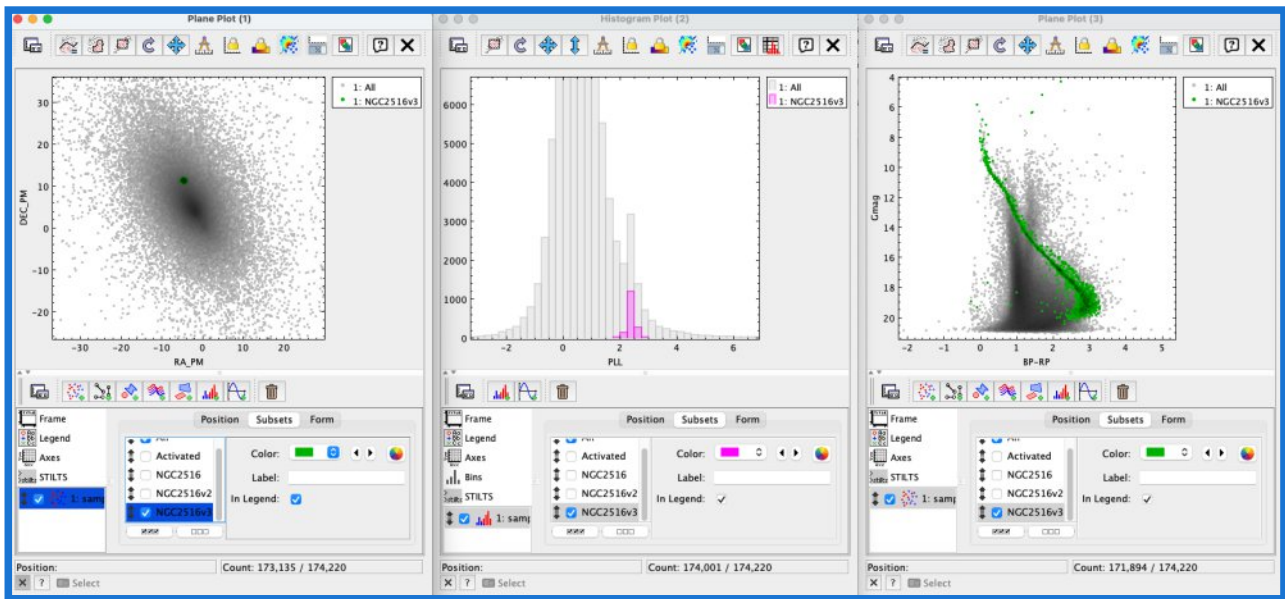


Fig. 16 – NGC 2516 – filtrado y optimizado en TOPCAT.

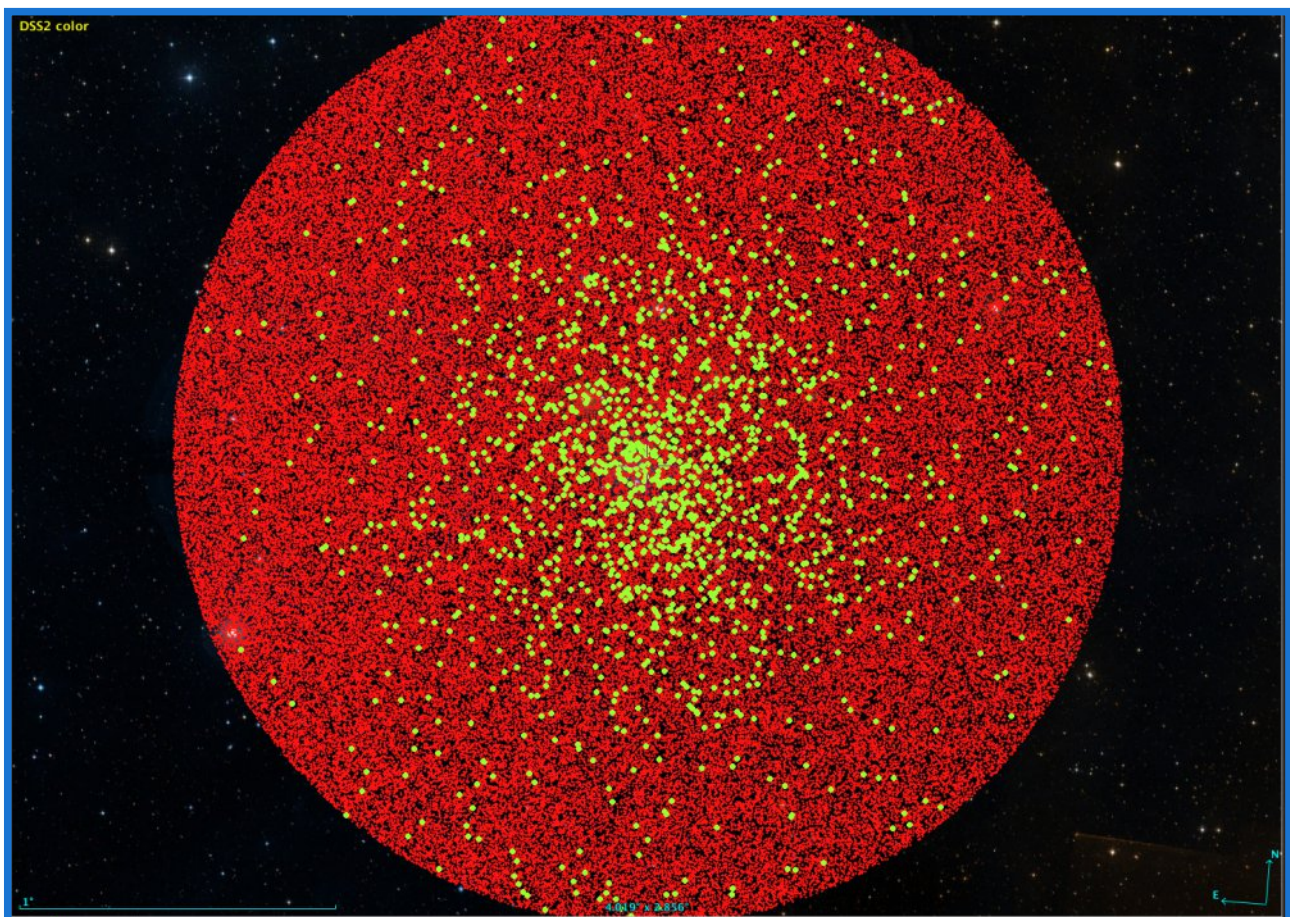


Fig. 17 – NGC 2516 – estrellas del cúmulo en Aladin.



ApEA, la Asociación para la Enseñanza de la Astronomía, nació en 1995 para acoger a todas las personas que se dedican a la enseñanza de la Astronomía en centros educativos, planetarios, museos de la ciencia, agrupaciones de aficionados y clubes de estudiantes.

ApEA engloba a todos los interesados en la enseñanza de todos los niveles educativos reglados -desde la enseñanza primaria hasta la universitaria- así como los no reglados. También organiza reuniones de formación para sus socios y publica materiales de interés didáctico, como la presente colección. Más información en www.apea.es

