

## SECTION: HERRAMIENTAS DEL OBSERVATORIO VIRTUAL

# Caracterización de cúmulos estelares abiertos

Joaquín Álvaro Contreras<sup>1</sup><sup>1</sup>FAAE, Madrid, Spain. E-mail: [jalvaro@citelan.es](mailto:jalvaro@citelan.es).

**Keywords:** cúmulos estelares, herramientas VO, TOPCAT, Aladin, observatorio virtual, Gaia  
© Este artículo está protegido bajo una licencia Creative Commons Attribution 4.0 License

## Resumen

El problema de discriminar qué estrellas forman parte de un cúmulo estelar abierto, en adelante **OC** (*Open Cluster*), descartando aquellas que no son parte del cúmulo, es fundamental para determinar la existencia de un OC, y en definitiva caracterizarlo, atribuyéndole una estimación precisa de sus componentes y a partir de estos, la masa del cúmulo, edad, distancia, movimiento propio relativo en la galaxia, composición química, etc. El proceso de caracterizar un OC no es fácil y, aunque ya se utilizan técnicas de inteligencia artificial eficaces, conocer los fundamentos y tareas para resolver este problema es conveniente si se desea trabajar en este campo, ya sea para refinar datos de OCs conocidos como para descubrir otros que aún no lo son. Afortunadamente hay algunas herramientas del Observatorio Virtual (VO) que facilitan la tarea. En este artículo, y los siguientes de esta serie, haremos un repaso amplio sobre algunos casos prácticos utilizando herramientas VO como TOPCAT [1], Aladin [2], Clusterix 2.0 [3] y VOSA [4], e intentaremos finalmente un desarrollo con técnicas de inteligencia artificial con este mismo propósito.

## Abstract

The problem of identifying which stars belong to an open star cluster (hereinafter OC, Open Cluster), while discarding those that do not, is fundamental to determining the existence of an OC and, ultimately, characterizing it. This involves providing an accurate estimate of its members and, based on them, calculating properties such as the cluster's mass, age, distance, relative proper motion within the galaxy, chemical composition, and more. The process of characterizing an OC is not straightforward, and although effective artificial intelligence techniques are already in use, understanding the fundamentals and tasks required to address this problem is essential for those who wish to work in this field, whether to refine data for known OCs or to discover previously unidentified ones. Fortunately, there are tools from the Virtual Observatory (VO) that make this task easier. In this article, and in subsequent ones in this series, we will provide a comprehensive review of practical cases using VO tools such as TOPCAT [1], Aladin [2], Clusterix 2.0 [3] and VOSA [4], and we will conclude by attempting a development using artificial intelligence techniques for the same purpose.

## 1. Introducción

Los cúmulos estelares abiertos 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* [5]; *Portegies Zwart et al. 2010* [6]). 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. 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. 2018 [7]*).

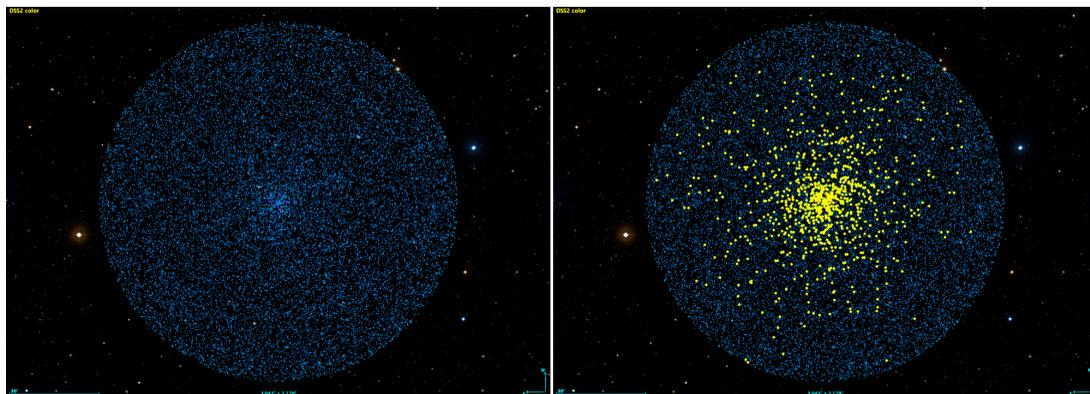
Los OCs pueden contener desde sólo unas decenas de componentes a unos pocos miles de estrellas y suelen encontrarse casi exclusivamente en el plano de la galaxia.

El estudio de los OCs ha tomado nuevo impulso a partir del inmenso y preciso conjunto de datos de la misión GAIA (ESA) [8], tanto para la revisión de los ya conocidos como para la determinación de otros nuevos, lo que ha aumentado el catálogo de OCs conocido hasta fechas recientes a prácticamente doblarlo con unos 3.800 en el momento actual [9].

## 2. Caracterización de cúmulos estelares abiertos

En este primer artículo analizaremos, como caso práctico, el cúmulo **NGC 2682 (Messier 67)** [10], bien conocido, lo que nos permitirá un test sencillo de los resultados obtenidos. Para éste, y otros ejercicios similares, puede tomarse como referencia de datos ya catalogados el *CDS VizieR Catalogue of Open Clusters* [11], del que sólo haremos uso para tomar una estimación inicial del campo de cielo a estudiar y, al final, para comparar resultados con los datos aquí recogidos.

Antes de empezar podemos ver las dos imágenes presentadas en la Fig. 1. A la izquierda se muestran, (*como puntos azules*) todas las estrellas presentes en el campo de visión seleccionado<sup>1</sup>, donde se presume que está localizado NGC 2682. A la derecha, (*como puntos amarillos*), se resaltan las estrellas que forman parte del cúmulo. Esto nos da idea de cuál es el objetivo perseguido, la caracterización del cúmulo, que empieza por determinar las estrellas que forman parte del mismo.



**Figura 1.** Izq: 20.284 estrellas a partir de Gaia/EDR3, (ver nota al pie). Dcha: sobrepuestas y resaltadas en amarillo 920 estrellas estimadas como componentes de NGC 2682 a partir de un análisis rápido con TOPCAT. Ambas imágenes son representaciones en *Aladin Sky Atlas* con los datos transferidos desde TOPCAT.

<sup>1</sup>Obtenidas de Gaia/EDR3 con un campo de cielo de 1° para la posición estimada del centro del cúmulo en coordenadas ecuatoriales. En total aparecen 20.284 estrellas hasta la magnitud visual 22,26

Para el propósito que nos ocupa en este artículo se utilizará **TOPCAT** y, opcionalmente **Aladin** como atlas del cielo, con el único objetivo de tener una representación visual ilustrativa. Si se desea reproducir el hilo de procesos que vamos a seguir (*deseable*), es necesario instalar ambas herramientas VO, (ver referencias [1] y [2]; en ambos casos hay instrucciones para la instalación y tutoriales de las dos aplicaciones, si se desea profundizar en ellas. También el *Observatorio Virtual Español (SVO)* tiene publicadas varias escuelas y talleres al respecto) [12], (ver caso práctico del SVO, por Francisco Jiménez-Esteban y Belén López Martí [13]).

El análisis de un campo estelar que presumiblemente contiene un OC, con objeto de caracterizarlo, inicialmente puede hacerse a partir de algunas características esbozadas en la definición de 'cúmulo estelar abierto': *conjunto de estrellas nacidas de una misma nube de gas molecular en una ventana temporal relativamente corta*. Esto significa que las estrellas del cúmulo van a compartir una serie de propiedades. A saber:

1. La dinámica de la nube de gas progenitora. Esto indica que estas estrellas presentarán movimientos propios similares, (tanto en ascensión recta como en declinación), frente al conjunto de estrellas del campo de visión que, en general, tendrán una distribución aleatoria de velocidades.
2. Además 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. Por último, estas estrellas tendrán una misma composición química, (heredada de la nube original), y una misma edad, también dentro de un rango temporal bastante estrecho.

### 3. Primeros pasos - Obtención de datos y análisis sobre velocidades

Empezaremos por cargar en TOPCAT, desde Gaia/EDR3, datos estelares del campo de cielo deseado, para iniciar seguidamente un primer filtrado atendiendo a los movimientos propios, según el punto 1.

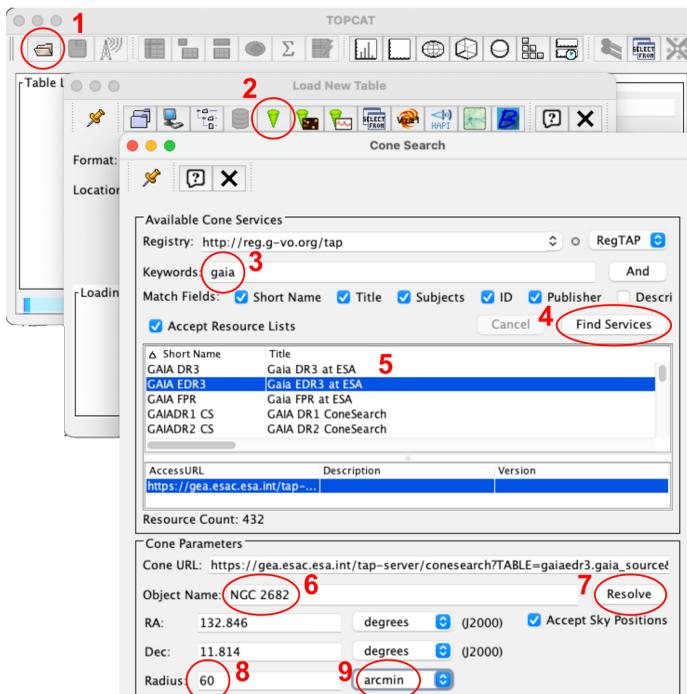
#### 3.1. Acceso a datos en Gaia desde TOPCAT

Se ejecuta la aplicación TOPCAT que se iniciará presentando una ventana simple sin datos ni opciones de menú activas ya que aún no se ha cargado ninguna tabla. La única opción posible es la de 'cargar una nueva tabla'. Seguiremos la secuencia de pasos indicados en la Fig. 2 y que se detallan a continuación.

1. cargar nueva tabla - Se abre una segunda ventana donde seleccionaremos el icono marcado con el número 2 'obtener catálogos de fuentes usando los servicios web de búsqueda por conos'. (Los menús son contextuales)
2. al seleccionar el ítem marcado con el número 2 se abre una tercera ventana.
3. en *Keywords* escribimos 'gaia' y,
4. pulsamos en 'encontrar servicios'
5. en la lista de recursos que se despliega, seleccionamos 'GAIA DR3' o 'GAIA EDR3'
6. ahora indicamos el nombre del objeto que nos interesa. En este caso 'NGC 2682' y,
7. pedimos que lo resuelva. En este paso se rellenarán automáticamente las coordenadas (RA, Dec) del objeto buscado. (También es posible escribir directamente estas coordenadas, si se conocen, en los campos correspondientes).
8. ya sólo falta indicar el radio del campo de cielo a considerar. En este caso elegimos 1º (60 minutos de arco)<sup>2</sup>
9. una vez completados estos pasos, confirmamos la búsqueda pulsando 'OK'

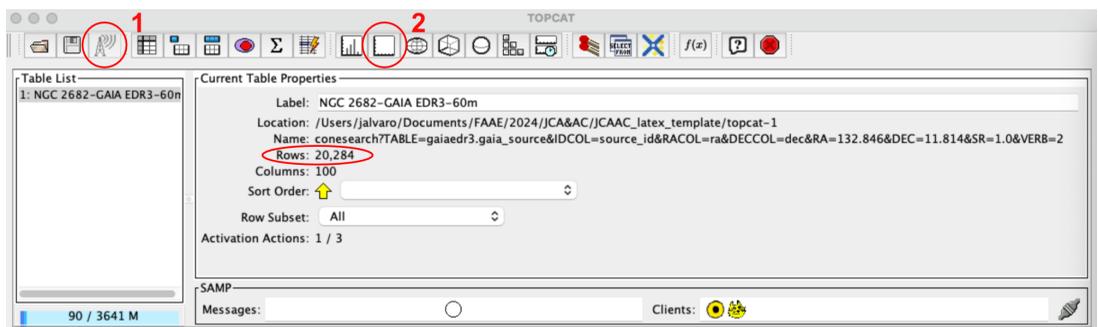
---

<sup>2</sup>De los datos conocidos de este cúmulo -ver referencia [7]- sabemos que su diámetro estimado es de 25 arcmin. Aquí seremos algo más generosos con el tamaño del campo a analizar.



**Figura 2.** Captura de pantalla de tres ventanas consecutivas en TOPCAT. (Las ventanas se van desplegando a medida que se pulsán las acciones 1, 2, ...) Los números correlativos indican los pasos sucesivos a seguir para obtener los datos de Gaia para NGC 2682. Los detalles en cada caso se explican en el texto del artículo.

En pocos segundos se habrá cargado la tabla correspondiente con los datos extraídos de Gaia y la ventana principal de TOPCAT nos dará información de su contenido (Fig. 3).



**Figura 3.** 20.284 registros recuperados de Gaia con los datos estelares disponibles para esta región del cielo.

Aquí podemos explorar esta información, tanto los datos brutos de la tabla como los metadatos de la misma o las columnas que la componen. Las opciones disponibles están accesibles desde los iconos contextuales en la parte superior de esta ventana.

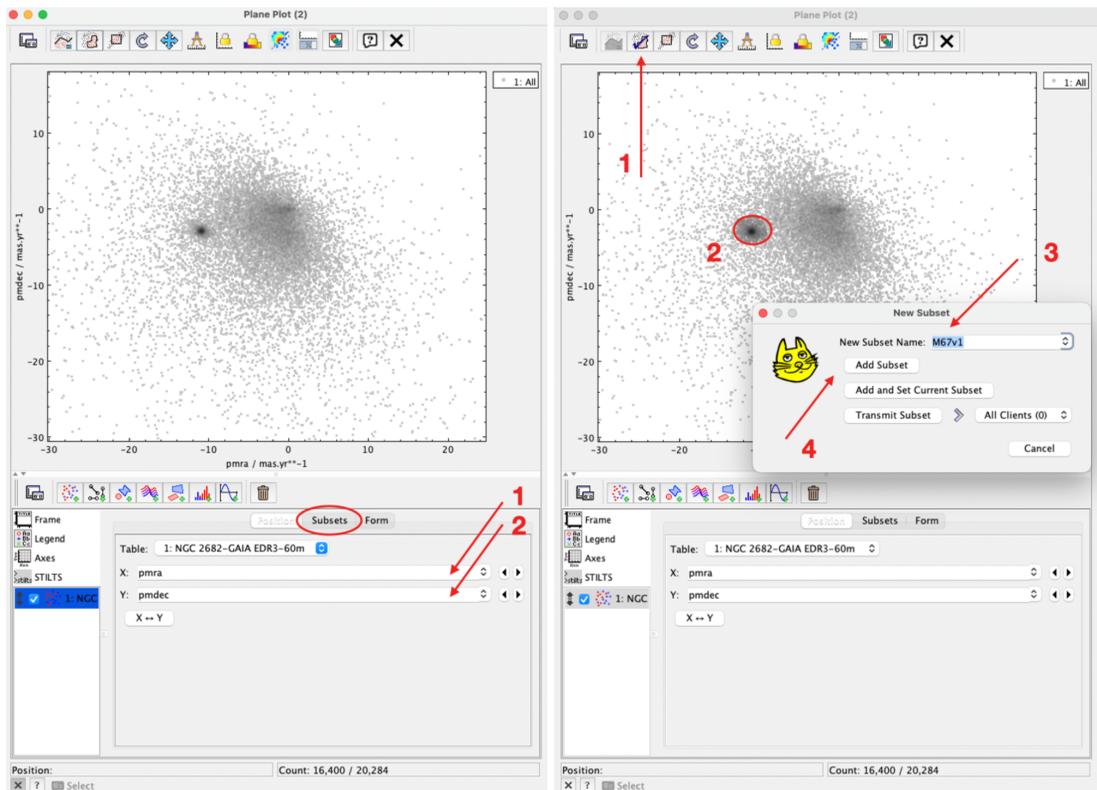
Por ahora nos limitaremos a dos de estas funcionalidades. La primera, opcional, será exportar estos datos a *Aladin*. Para ello es preciso abrir esta aplicación (*Aladin*) y escribir en la línea superior de comando: NGC 2682 (+INTRO), con objeto de situar en el atlas del cielo una posición y escala aproximada. Hecho esto podemos pulsar el icono correspondiente en TOPCAT, (marcado con '1' en la Fig. 3). El resultado será similar a lo presentado en la Fig. 1 (izquierda).

Pero lo que nos interesa para continuar en nuestro objetivo de caracterizar el cúmulo es ver, sin más

preámbulos, la información disponible en el espacio de configuración de velocidades, (movimientos propios en AR y Dec).

### 3.2. Análisis de movimientos propios en las estrellas del campo seleccionado

Empezaremos por representar el mapa de movimientos propios del *dataset* descargado. Para ello pulsamos en el icono superior 'ventana de trazado de mapas', (marcado con '2' en la Fig. 3). Se abrirá una nueva ventana en la que debemos seleccionar los parámetros que queremos representar, en este caso movimiento propio en ascensión recta y movimiento propio en declinación, (*pmra* y *pmdec*, del conjunto de datos). Ver Fig.4 - izquierda.



**Figura 4.** Izquierda: mapa de velocidades del conjunto de estrellas descargado de Gaia. Derecha: selección de candidatos a formar parte del cúmulo. En la parte inferior de ambas imágenes se puede ver que del total de 20.284 estrellas extraídas de Gaia, sólo 16400 han sido representadas. El resto corresponden a estrellas que no tienen informada alguna de las dos velocidades o ninguna de ellas.

Para obtener el gráfico de movimientos propios, (Fig.4 - izquierda), se debe seleccionar el parámetro 'pmra' en el eje X y el parámetro 'pmdec' en el eje Y, en las posiciones marcadas con '1' y '2' en la figura. La representación será inmediata. Se puede ajustar la escala de ambos ejes en el mapa usando el ratón en la manera habitual para ajustar el 'zoom'. También se puede elegir el color con el que se presentan los puntos de este mapa a partir de la pestaña 'subsets' marcada en la imagen.

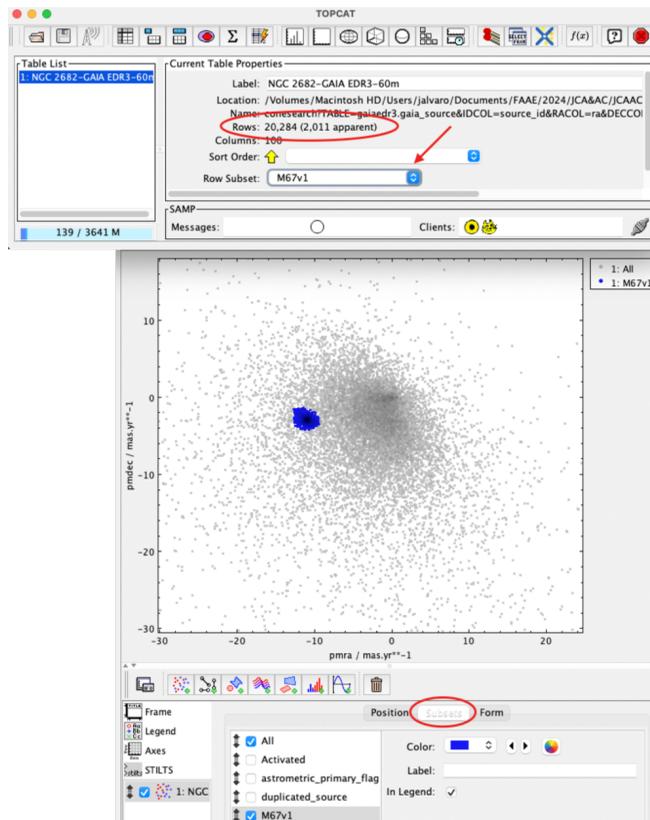
Lo que podemos ver en este gráfico es que las velocidades de las estrellas se distribuyen en general con cierto carácter aleatorio presentando una mayor abundancia en torno a los '0' en ambos ejes. Pero se

puede observar también una *sobre-densidad*, muy evidente y concentrada, en torno a  $pmra -10$  y  $pmdec -2$ . Esto es justamente lo que buscábamos: una concentración de estrellas con movimientos propios muy similares y definidos, como ésta, revela con toda seguridad la existencia de un OC.

El siguiente paso es seleccionar el conjunto de puntos de la sobre-densidad para crear un *subset*. Asumimos que con ello estamos creando un filtro que separará los posibles componentes del cúmulo del resto de estrellas de fondo. Para ello veamos ahora la parte derecha de la Fig. 4:

Empezaremos por seleccionar el icono correspondiente a '*dibujar una región a mano alzada para definir un nuevo subset*', (paso número '1' de la Fig. 4 -derecha-). Utilizaremos el ratón para encerrar la región señalada con un bucle en forma de lazo, (punto '2' Fig. 4 -derecha-). Una vez cerrado el lazo se debe volver a *pinchar* el mismo icono de antes ('1') y en la ventana emergente daremos nombre a este subset (punto '3'). Conviene añadir un número a estos nombres porque serán varios los filtros sucesivos que aplicaremos. Hecho esto, pulsaremos el botón '*add subset*' (paso '4').

Después de estas primeras acciones tenemos un subconjunto del total al que hemos llamado '*M67v1*' que contiene 2.011 estrellas (ver remarque en la Fig. 5) y que es resultado de la selección hecha casi *artesanalmente* a mano alzada sobre el mapa inicial.

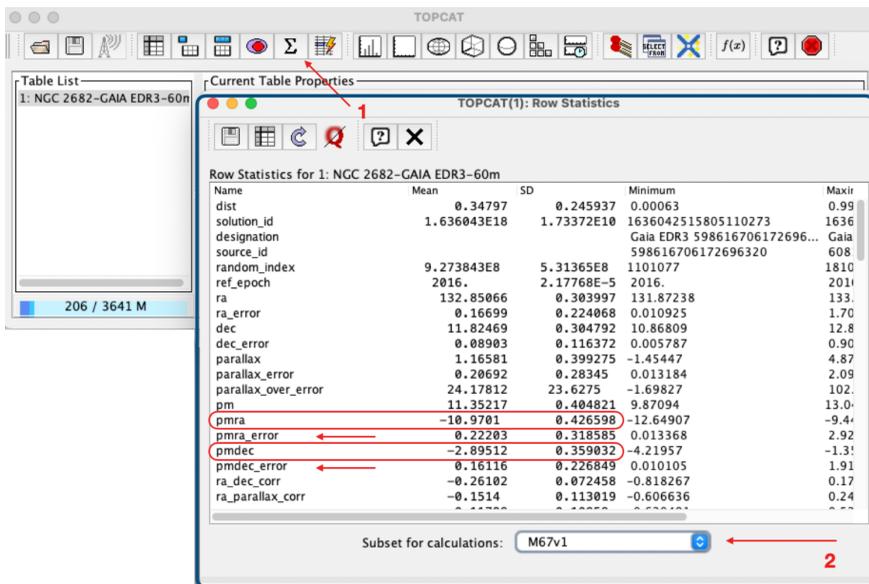


**Figura 5.** Representación del subset resultado del primer filtrado (en azul), sobre el fondo del dataset total inicial. En la ventana superior se puede elegir entre los datos del dataset original 'all' y cualquiera de los subsets que se irán creando -combo marcado con la flecha-. En la ventana inferior, la pestaña 'subsets' permite activar/desactivar los diferentes conjuntos/susconjuntos representados, así como otras propiedades gráficas de los mismos, como el color.

El hecho de ir paso a paso enumerando con detalle cada una de las acciones a realizar puede distraernos de alguna manera del propósito real, que es ir progresando en el proceso de discriminar los componentes del OC. Se ha optado por ser minucioso en lo posible en esta primera fase con objeto de facilitar la tarea de seguir este caso práctico a las personas que afrontan por primera vez el trabajo con TOPCAT. A medida que vayamos avanzando no será necesario ser tan prolijos en detalles y el hilo conceptual irá ganando peso.

Lo que debe quedar claro después de esta primera selección es que, a todas luces, es insuficiente para dar por resuelto el problema. Se trata de una primera aproximación sobre la que se debe seguir trabajando. Lo siguiente que haremos es aplicar un nuevo filtro sobre el subset 'M67v1'. Para ello acotaremos algo la posible dispersión de las velocidades en 'M67v1' que, como hemos visto, se ha seleccionado un poco 'a ojo' y en el que, por precaución, se ha sido algo generoso (por exceso) a la hora de *dibujar* la región de interés en el mapa del espacio de configuración de velocidades. También aprovecharemos el nuevo filtro para descartar aquellas estrellas en las que el error en los datos de movimientos propios sea mayor de lo razonable.

Hacemos uso ahora de otra funcionalidad de TOPCAT que nos permite tener información resumizada y estadística de los datos que manejamos.



**Figura 6.** Detalle de valores estadísticos para cada columna de la tabla. Datos referidos al subset 'M67v1'. Remarcados en la figura los campos significativos en esta etapa del trabajo. Las indicaciones correspondientes se detallan en el texto del artículo.

En la ventana inicial de TOPCAT seleccionamos ahora la función de '*mostrar estadísticas*', (paso '1' en la Fig. 6). Esto nos abre una nueva ventana, (remarcada en azul en la figura anterior), que presenta diversos parámetros (media, desviación estándar, mínimos, máximos, ..) para cada columna de la tabla. Seguimos interesados en los campos relativos a los movimientos propios: pmra, pmdec y sus respectivos errores empíricos. Nótese que esta información está referida al subset 'M67v1' (indicado por el '2' en la imagen).

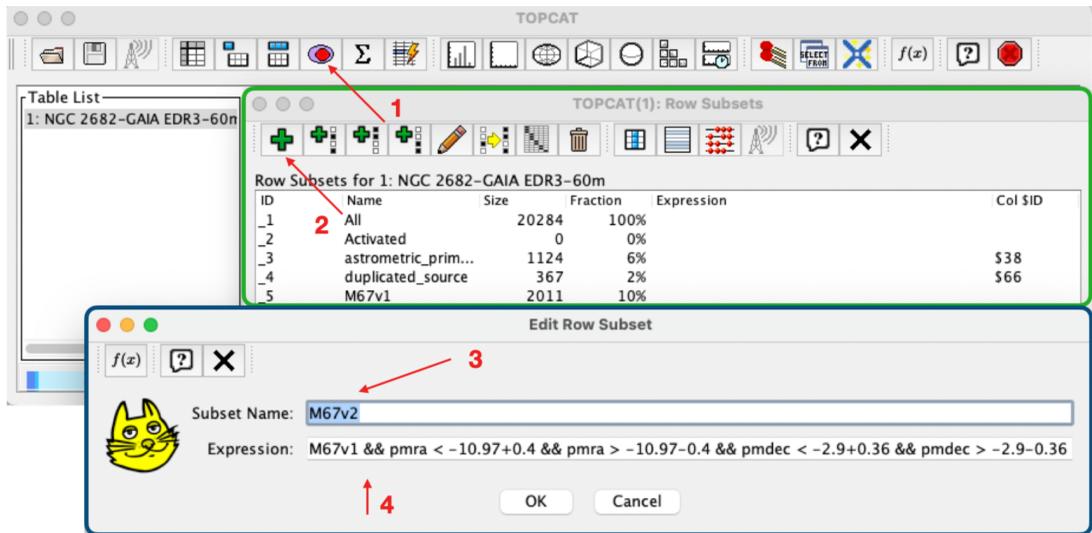
Nuestro objetivo aquí es añadir un nuevo filtro a los objetos seleccionados en la etapa anterior con la finalidad de acotar en lo posible y de manera razonable los candidatos a ser componentes del cúmulo.

Para ello vamos a tener en cuenta el valor medio de  $pmra$  y  $pmdec$ , así como la dispersión de estos respecto a los valores medios tomando como referencia la *desviación estándar*. También tendremos en cuenta la 'calidad de los datos empíricos' recuperados de Gaia, por lo que los campos 'pmra\_error' y 'pmdec\_error' serán significativos, pero en este caso no los que presenta la información estadística, sino los correspondientes a cada objeto de 'M67v1'. Así pues construiremos nuestro nuevo filtro en base a las siguientes condiciones:

1. descartar aquellas estrellas que presenten valores en  $pmra$  y  $pmdec$  que estén fuera del intervalo ' $media(pmra) \pm desviación\ estándar(pmra)$ '. Lo mismo haremos con  $pmdec$ .
2. descartar las estrellas cuyo error relativo en  $pmra$  y  $pmdec$  sea superior al 10%. Esto es:  
 $pmra\_error / pmra > 0.1$  y  $pmdec\_error / pmdec > 0.1$ .

Evidentemente estos criterios pueden ser modificados según se estime oportuno. En el caso práctico que aquí estamos desarrollando serán los utilizados a modo de ejemplo.

Ya estamos listos para construir nuestro nuevo filtro. Ver indicaciones a partir de la Fig. 7.



**Figura 7.** Detalle de un nuevo filtro para 'M67v1'. Remarcados en la figura los sucesivos pasos para hacerlo efectivo. Las indicaciones se detallan en el texto.

Para definir un filtro y generar un nuevo subset a partir de 'M67v1' utilizaremos la función '*display row subsets*', marcada con '1' en la Fig. 7. Se abrirá una ventana, (remarcada en verde en la figura). Ahí, paso '2', seleccionamos '*definir nuevo subset usando una expresión algebraica*' que nos abrirá la ventana remarcada en azul. Damos nombre al nuevo subconjunto, (paso '3', seguimos con el criterio sencillo de añadir un número secuencial a nuestro juego de subsets), y finalmente en la 'caja' destinada a la 'expresión' escribimos nuestro conjunto de reglas sin olvidar que queremos aplicarlas al último subset válido que ya tenemos, (en este caso 'M67v1'), y que para no olvidarlo lo escribiremos en primer lugar:

*Expression:*

M67v1 &&

$$\begin{aligned} pmra < -10.97 + 0.4 \quad \&\& \quad pmra > -10.97 - 0.4 \quad \&\& \\ pmdec < -2.9 + 0.36 \quad \&\& \quad pmdec > -2.9 - 0.36 \quad \&\& \\ pmra\_error/pmra < 0.1 \quad \&\& \quad pmdec\_error/pmdec < 0.1 \end{aligned}$$

Las cuatro líneas de la expresión anterior deben escribirse seguidas en una misma línea continua dentro de la caja 'expression'. El doble '&' (&&) es la forma de indicar que todas las condiciones deben cumplirse, (operador lógico 'Y').

También se pueden aplicar por separado de manera sucesiva, pero en este caso debe ponerse especial atención en el *subset* sobre el que se van ejecutando, cuyo nombre deberá estar siempre presente en cada una de las expresiones. Los valores (-10.97, 0.4, -2.9, 0.36) insertados en la expresión del filtro anterior corresponden al valor medio y desviación estándar de cada variable (ver Fig. 6).

Ejecutado el filtro, el nuevo subconjunto 'M67v2' tiene ahora 1.130 estrellas y ya estamos preparados para tener en cuenta la segunda de las propiedades enumeradas en el apartado 2. sobre los OCs: sus componentes estarán a una misma distancia respecto al origen en el sistema de referencia desde el que se observan, dentro de un intervalo razonable propio del carácter extenso de un cúmulo abierto.

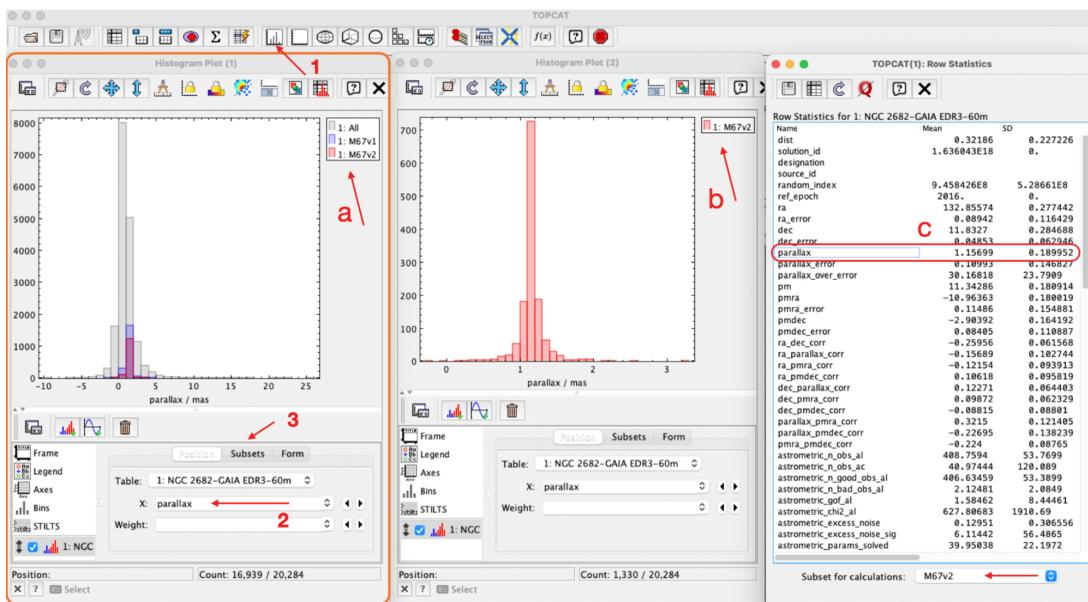
### 3.3. Análisis de distancias. El paralaje

En la información que tenemos de NGC 2682, recuperada de Gaia, la distancia está disponible en el campo 'parallax', esto es en forma de paralaje y unidades 'mas' (milisegundos de arco). Es fácil trasladar este dato a *parsec* (pc), o bien a *años-luz*, (' $distancia(pc) = 1000/parallax(mas)$ ' y ' $distancia(años - luz) = distancia(pc) \cdot 3.26$ '). Pero no es necesario, seguiremos trabajando con el valor del *paralaje*.

Vamos a TOPCAT de nuevo (ver Fig. 8). Ahora seleccionamos la función de '*representar un histograma*' para el valor de paralajes, (en la ventana principal de TOPCAT, paso '1' de la figura). Se abre una nueva ventana, remarcada en la imagen. Sólo debemos seleccionar el campo 'parallax' para el eje X, (paso '2'). En la pestaña '*subsets*' se pueden activar los subconjuntos representados, (por ahora es interesante marcar 'all', 'M67v1' y 'M67v2'), y tendremos la representación mostrada en la parte izquierda de la figura (véase leyenda, marcada con 'a'). Se puede ajustar la escala de la representación pulsando la 'cruz' en la parte superior (cuarto icono por la izquierda).

Podemos jugar un poco con la escala de la representación en ambos ejes para tener una visión más detallada de esta distribución (ver Fig. 9), donde se aprecia mejor cómo impacta el subconjunto 'M67v2' sobre el fondo de estrellas. En otros casos, (OCs más cercanos, por ejemplo), el histograma mostraría un 'pico' sensiblemente desplazado sobre la distribución del total. Evidentemente, no son esperables paralajes negativos. Los que aparecen aquí responden a datos empíricos de mala calidad. No nos afectan y los ignoramos.

En la parte central de la Fig. 8 se muestra sólo la distribución para 'M67v2'. Como se hizo en el apartado anterior con las velocidades, vamos a intentar estrechar la dispersión en distancias acotando valores y filtraremos igualmente estrellas con errores relativos mayores a un 10 % en el valor del paralaje. Para ello recurrimos de nuevo a los datos estadísticos de 'M67v2' y tomamos nota de la 'media' y la 'desviación estándar' del campo 'parallax'. Seguiremos el mismo criterio que con los movimientos propios.

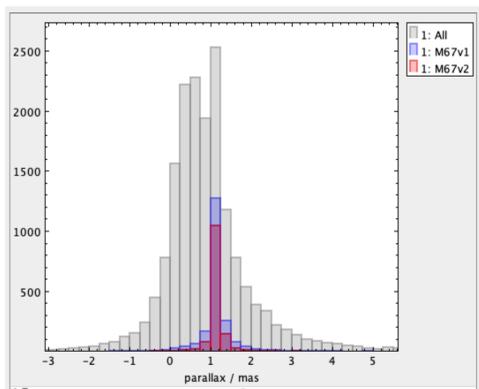


**Figura 8.** *histograma que representa la distribución del campo parallax. En la figura se presentan tres ventanas diferentes (izquierda, centro y derecha) cuyos detalles se comentan en el texto del artículo.*

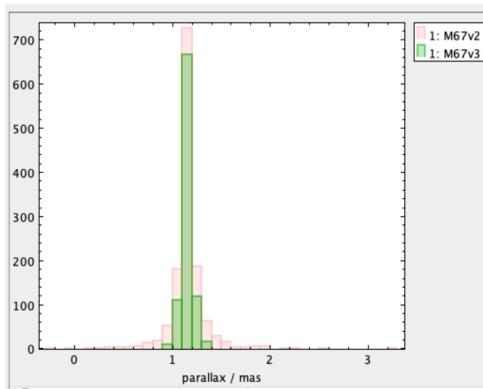
La expresión para el nuevo filtro sería como se muestra a continuación y lo aplicamos para crear un nuevo *subset* 'M67v3' de igual forma que se hizo antes con el resultado que se muestra en la Fig. 10.

M67v2 &&

$$parallax < 1.16 + 0.19 \quad \&\& \quad parallax > 1.16 - 0.19 \quad \&\& \\ parallax\_error / parallax < 0.1$$



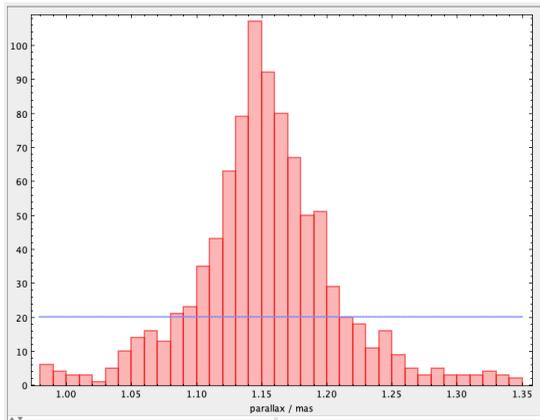
**Figura 9.** *Detalle de la Fig. 8 (izquierda) con escalas en los ejes ajustadas para una mayor resolución de la región de interés.*



**Figura 10.** *Histograma de parallax para 'M67v3' sobre el fondo de 'M67v2'. En la Fig. 11 se presenta con mayor detalle.*

El resultado de los dos filtros aplicados sobre la selección inicial, hecha a *'mano alzada'* a partir del mapa de velocidades del conjunto total, nos deja un *dataset* con **920 estrellas** que damos por bueno en el ejemplo desarrollado aquí. Si comparamos el número de componentes obtenido para NGC 2682 después de este simple análisis con el ofrecido en la *WiKi* [10] veremos que resulta bastante aproximado. Comparado con los datos del catálogo CDS-VizieR [11], el resultado excede sensiblemente a los registrados allí, pero hay que tener en cuenta que hemos ampliado de forma notoria el radio del campo de cielo y también, y esto es más importante, que los datos de Gaia son más completos que los recogidos en *surveys* anteriores.

Si analizamos el resumen estadístico de los datos de nuestra última tabla, *'M67v3'*, y en concreto los relativos a la distancia, *'parallax'*, algo que queda bien representado en el histograma de paralaje (Fig. 11), tendremos una estimación de su distancia relativa a nuestra posición.



**Figura 11.** *histograma ampliado de parallax para 'M67v3'. La línea de corte sobre el valor '20' en el número de estrellas se corresponde muy aproximadamente con una nueva estimación del tipo  $media(parallax) \pm sd$  (desviación estándar), que en términos numéricos es  $1.153 \pm 0.055$ .*

Asumiendo una distancia media al cúmulo de  $1.153$  (*mas*) con un intervalo de confianza del  $68\% = 0.055$  (*mas*)<sup>3</sup>, tendremos una distancia de  $867.3(pc) \pm 42(pc)$ , (recordar que  $distancia(pc) = (1/parallax(mas)) \cdot 1000$ ), o bien en *años-luz*  $= 2.827 \pm 137$  que sería un rango razonable de amplitud en su profundidad de distancia.

Todavía no podemos estimar datos como la masa total del cúmulo, luminosidad, composición química, edad, etc. Para terminar de caracterizarlo, una vez que asumimos cuáles son las estrellas componentes, deberíamos recabar datos de cada una de ellas y aplicar algunos modelos *ad hoc* para este cometido. Intentaremos alguna aproximación en la tercera entrega de esta serie.

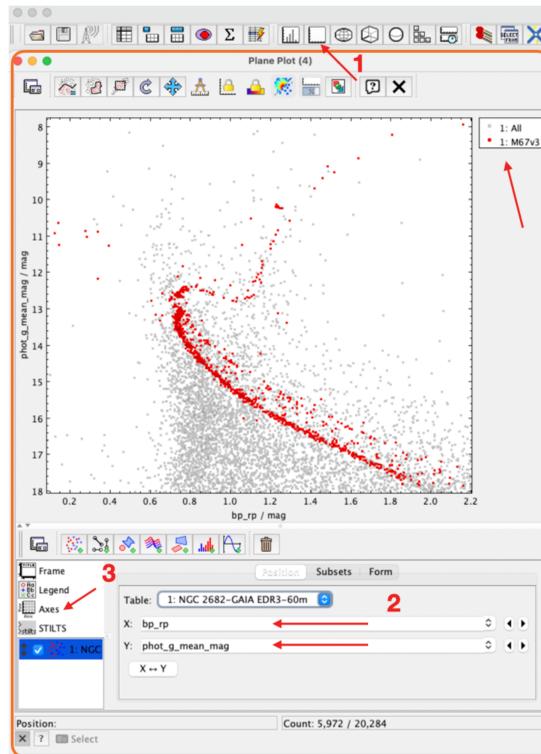
Lo que si podemos hacer es transferir *'M67v3'* a Aladin con objeto de reproducir la Fig. 1, tal como ya se comentó en la parte inicial de este artículo. Pero antes de terminar, aún podemos hacer algo más ...

#### 4. El diagrama color-magnitud

Sin abandonar TOPCAT vamos a representar el diagrama color-magnitud de *'M67v3'*, similar a un diagrama *Hertzsprung-Russell*, pero en vez de flujos y temperaturas utilizaremos el *índice de color* (*azul-rojo*) y la *magnitud visual* (*gaia*), aprovechando que en la información descargada de Gaia tenemos también estos datos fotométricos debidamente calibrados. Para ello volvemos a representar un mapa con estos dos parámetros en los respectivos ejes X e Y. Ver Fig. 12.

De nuevo seleccionamos la función de *'display plane window'*, (paso '1' en la Fig. 12), que nos abrirá una nueva ventana en la que asignaremos el campo *'bp-rp'* al eje X y *'phot-g-mean-mag'* en el eje Y,

<sup>3</sup>En una distribución normal el rango *'media ± σ'* incluye el 68% de la muestra.



**Figura 12.** Diagrama color-magnitud para 'M67v3' con el fondo del conjunto inicial descargado para el área de cielo de 1º centrada en las coordenadas de NGC 2682. Los detalles en el texto del artículo.

(paso '2'), (esto es 'magnitud en el azul - magnitud en el rojo = índice de color (b-r) y magnitud visual respectivamente).

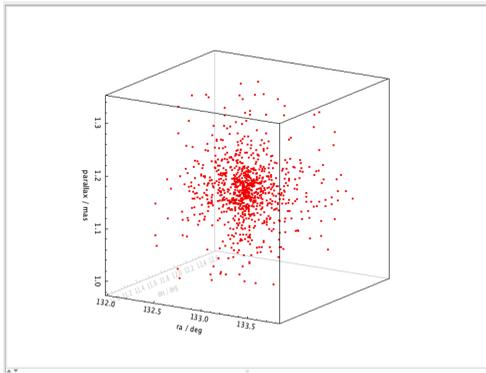
Será conveniente invertir el eje Y en esta representación para ver el diagrama en su forma más familiar. Para ello, en la parte inferior izquierda, se selecciona la opción 'ejes' y se procede sobre el eje Y, (paso '3'). Como ya sabemos, en la pestaña 'subsets' podemos activar los subconjuntos a visualizar y asignarles los colores deseados. En la Fig. 12 se presenta el subset 'M67v3', esto es, la última versión de este breve análisis después de los filtros aplicados y, de fondo, el conjunto total inicial. Si nos fijamos en el pie del gráfico se muestra un contador que indica *Count: 5,972 / 20,284* y que significa que, de los 20284 objetos recuperados de Gaia, sólo 5972 se han representado; el resto carecen de alguno de los datos dibujados o de ambos. De las 920 estrellas de 'M67v3', 919 están incluidas en el gráfico.

Como puede apreciarse en el gráfico el conjunto de estrellas (en rojo), que finalmente hemos atribuido a NGC 2682 como componentes del mismo, muestran un perfil bastante nítido en el que se aprecia como la mayoría de las estrellas están en la *secuencia principal* y también se advierte el punto de inflexión en el que ya algunas de ellas inician su recorrido hacia la región de gigantes y super-gigantes rojas. Un número muy pequeño se observa en la zona de las estrellas más brillantes (azules y/o blancas). De aquí puede deducirse que éste, además de ser un cúmulo bastante poblado, no es un OC especialmente joven. Aunque para intentar la aproximación a una determinación más precisa de su edad y otras características del mismo tendremos que esperar a la tercera entrega de esta serie de artículos, en la que haremos uso de la herramienta VO 'VOSA' [4].

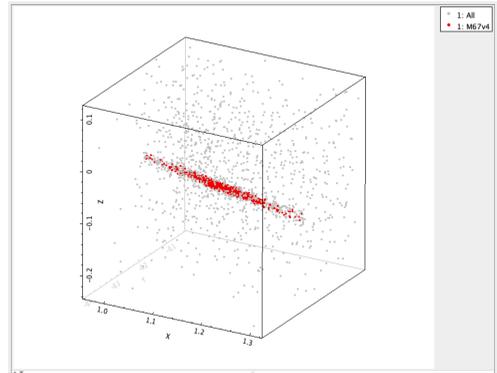
## 5. Otros aspectos de interés

Antes de terminar con TOPCAT hay que advertir que una de las funcionalidades de esta herramienta, (segundo icono por la izquierda en la ventana principal), permite guardar localmente bien una o todas las tablas utilizadas y, así mismo, la sesión completa, con los filtros y procesos desarrollados. Hay varios formatos disponibles para salvar estos objetos.

Igualmente podemos tener una representación 3D del cúmulo tanto del mapa de posiciones espaciales (ejes:  $X = RA$ ,  $Y = Dec$  y  $Z = parallax$  en coordenadas cartesianas, Fig. 13), así como una imagen del *chorro* dinámico que presenta frente a su entorno más próximo a partir de un mapa 3D sobre los movimientos propios en coordenadas cilíndricas, Fig. 14, (ejes:  $X = pmra$ ,  $Y = pmdec$  y  $Z = parallax$ ).



**Figura 13.** Mapa 3D de NGC 2682 en coordenadas cartesianas, ( $RA$ ,  $Dec$ ), en el eje  $Z$  se representa la distancia. Esta funcionalidad en TOPCAT permite realizar movimientos de rotación en los tres ejes para una visualización más completa.



**Figura 14.** Mapa 3D en coordenadas cilíndricas del espacio de configuración de velocidades, (en el eje  $Z$  se representa también la distancia). Esta representación da idea de la dinámica propia del cúmulo, en forma de 'chorro', en relación a su entorno de fondo de cielo.

Puesto que tenemos una estimación del movimiento propio del cúmulo podríamos aventurar también una evolución temporal de su posición relativa en la galaxia, tanto hacia el futuro como al pasado, por ejemplo en millones de años, ( $10^6$ ,  $10^7$ , ...). Para esto sería conveniente utilizar coordenadas galácticas y hacer las correspondientes transformaciones de los valores con los que hemos trabajado ( $ra$ ,  $dec$ ,  $pmra$  y  $pmdec$ ) al nuevo sistema de referencia. (TOPCAT permite añadir campos nuevos a cualquiera de las tablas en curso definiendo al tiempo su contenido, de esta manera sin demasiadas complicaciones se pueden realizar simulaciones como las comentadas aquí). Un ejemplo al respecto podemos verlo en *Cameren Swiggum, et al.* [14], artículo en el que se detallan 3 superestructuras, o *familias* de cúmulos *cercanos*, generadas a partir de *macro-nubes* de gas en una simulación hacia atrás en el tiempo [ $30 \cdot 10^6$  años] y circunscrita a un espacio de 6500 años-luz en torno al Sol [15].

En el próximo artículo haremos uso de *Clusterix 2.0* [3], otra herramienta del Observatorio Virtual que facilita sensiblemente el trabajo que hemos hecho hasta aquí, al menos en su primera parte. Como hemos comentado en algún momento, hay cierto grado de *artesanía* en todo esto. En realidad un procesado de este tipo debemos definirlo como *supervisado* y *parametrizado* en tanto que exige en todo momento la actuación del analista definiendo criterios y siguiendo paso a paso su desarrollo. *Clusterix* se define a sí mismo como herramienta *no supervisada* y *no parametrizada*, aunque ya veremos que tiene otros condicionantes . . .

## Agradecimientos

### TOPCAT -

This research has made use of "TOPCAT VO"

Author: Mark Taylor, Astrophysics Group, Physics Department, University of Bristol

Email: m.b.taylor@bristol.ac.uk

TOPCAT WWW page: <http://www.starlink.ac.uk/topcat/>

<https://ui.adsabs.harvard.edu/abs/2005ASPC..347...29T/abstract>

### Aladin -

This research has made use of "Aladin sky atlas" developed at CDS, Strasbourg Observatory, France

<https://ui.adsabs.harvard.edu/abs/2000A%26AS..143...33B/abstract>

<https://ui.adsabs.harvard.edu/abs/2014ASPC..485..277B/abstract>

<https://ui.adsabs.harvard.edu/abs/2022ASPC..532....7B/abstract>

### Spanish Virtual Observatory-

This research has made use of the Spanish Virtual Observatory <https://svo.cab.inta-csic.es> project funded by MCIN/AEI/10.13039/501100011033/ through grant PID2020-112949GB-I00.

## References

- [1] TOPCAT: Tool for OPerations on Catalogues And Tables - <https://www.star.bris.ac.uk/~mbt/topcat/>
- [2] Aladin Sky Atlas - <https://aladin.cds.unistra.fr>
- [3] Clusterix 2.0 - <http://clusterix.cab.inta-csic.es/clusterix/>
- [4] VO SED Analyzer - <http://svo2.cab.inta-csic.es/theory/vosa/index.php>
- [5] Clarke, C.J. et al - "The Formation of Stellar Clusters" - <https://ui.adsabs.harvard.edu/abs/2000prpl.conf..151C/abstract>
- [6] Portegies Zwart et al. - "Young Massive Star Clusters" - <https://ui.adsabs.harvard.edu/abs/2010ARA%26A..48..431P/abstract>
- [7] Cantat-Gaudin, T. et al. - "A Gaia DR2 view of the open cluster population in the Milky Way" - [https://www.aanda.org/articles/aa/full\\_html/2018/10/aa33476-18/aa33476-18.html](https://www.aanda.org/articles/aa/full_html/2018/10/aa33476-18/aa33476-18.html)
- [8] Misión Gaia (ESA) - <https://sci.esa.int/web/gaia/>
- [9] Alfred Castro-Ginard et al. - "Hunting for open clusters in gaia dr2: 582 new open clusters in the galactic disc" - Astronomy & Astrophysics, 635:A45, 2020.
- [10] Wikipedia - [https://es.wikipedia.org/wiki/Cúmulo\\_abierto\\_M67](https://es.wikipedia.org/wiki/Cúmulo_abierto_M67)
- [11] Open Clusters - CDS VizieR Catalogue - <http://vizier.cds.unistra.fr/viz-bin/VizieR-3?-source=B/ocl/clusters>
- [12] Spanish Virtual Observatory <https://svo.cab.inta-csic.es>. Escuelas y tutoriales: <https://svo.cab.inta-csic.es/docs/index.php?pagename=Meetings>
- [13] Francisco Jiménez-Esteban & Belén López (SVO) - "Determination of the distance to Pleiades using TOPCAT" - <https://zenodo.org/records/10720670>
- [14] Cameren Swiggum et a. - "Most nearby young star clusters formed in three massive complexes" - <https://arxiv.org/abs/2406.06510>
- [15] YouTube @Cameren Swiggum Astro - <https://www.youtube.com/watch?v=4iBy5hz6eNE>