

Rendimiento de métodos semi-automáticos de búsqueda de cúmulos estelares

M. Chiarpotti^{1,2} & A.E. Piatti^{1,2}

¹ Instituto Interdisciplinario de Ciencias Básicas, CONICET-UNCuyo, Argentina

² Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Received: 09 February 2024 / Accepted: 29 May 2024

©The Authors 2024

Resumen / En este trabajo evaluamos el rendimiento de métodos computacionales de detección de cúmulos estelares utilizando la base de datos infrarroja VVV. La evaluación de dicha capacidad de detección la llevamos a cabo comparando los resultados obtenidos con métodos computacionales que desarrollamos con aquellos basados en la inspección visual de las imágenes VVV. De la comparación entre los resultados obtenidos a partir del método computacional desarrollado y los resultados que provienen de la detección de candidatos a cúmulo estelar por inspección visual, encontramos que el método computacional confirmó un 60 % de los candidatos seleccionados. Estos candidatos resultaron ser sobredensidades estelares que destacan, por lo menos, cinco veces por sobre la densidad media del fondo del cielo, lo cual interpretamos como un umbral de sensibilidad del ojo humano.

Abstract / In this work we evaluate the performance of computational methods for detecting star clusters using the VVV infrared database. We carried out the assessment of detecting star cluster candidates by comparing the results obtained from computational methods and those from visual inspection of the VVV images. We found that the computational methods recovered 60% of visual detected candidates. These candidates turned out to be stellar overdensities that stand out, at least, five times above the average density of the background of the sky, which we interpret as a threshold for the sensitivity of the human eye.

Keywords / open clusters and associations: general — stars: fundamental parameters — techniques: photometric

1. Introducción

El estudio de cúmulos estelares ha sido fundamental desde tiempos inmemoriales para mejorar nuestro conocimiento acerca de la formación, estructura, evolución dinámica y química de las galaxias. Particularmente, el análisis de los cúmulos abiertos es importante para conocer el pasado y presente de las propiedades de la Vía Láctea. Para la detección y posterior análisis de nuevos candidatos a cúmulos estelares a partir de relevamientos en el rango visible del espectro electromagnético, se utilizan actualmente programas de búsquedas automáticas o semiautomáticas con resultados satisfactorios (Ferreira et al. 2020). Sin embargo, estos métodos no han sido aún evaluados extensamente para detectar nuevos candidatos a cúmulos estelares en la región infrarroja del espectro. La región infrarroja es favorable para estudiar zonas de la Vía Láctea marcadamente afectadas por extinción interestelar, en la cual también abundan un gran número de estrellas, como por ejemplo las regiones centrales de la Galaxia. Estas características imponen nuevos desafíos a la búsqueda de candidatos a cúmulos estelares.

Borissova et al. (2011) identificaron, con un método de detección visual, 96 candidatos a cúmulos estelares y grupos estelares en el infrarrojo, utilizando los datos recolectados por *Visible and Infrared Survey Telescope for Astronomy, Variables in the Vía Láctea* (VVV:Fig. 1). El VVV es un relevamiento público que tiene como

objetivo observar homogéneamente la región central y el disco interno de la Vía Láctea, utilizando el telescopio VISTA de 4m ubicado en el Observatorio Europeo Austral (ESO, Chile). Cada región aproximadamente cuadrada (2x2 grados) cubierta por el mosaico de detectores empleados por el telescopio VISTA se denomina *tile*. Los datos que proporciona para cientos de millones de estrellas observadas en esta región de la galaxia consisten en la posición, movimientos propios y magnitudes en 5 filtros diferentes (Z, Y, J, H, Ks). Borissova et al. (2011) utilizaron solamente estrellas con $K_s < 13.5$ mag.

Piatti et al. (2016) desarrollaron un método de búsqueda de cúmulos semiautomático con el cual identificaron 38 candidatos a cúmulos estelares en el infrarrojo utilizando los datos proporcionados por VISTA para el relevamiento *Vista Magellanic Clouds* (VMC) en una zona altamente enrojecida de la Nube Mayor de Magallanes. Para su desarrollo, utilizaron un programa de búsqueda de sobredensidades basado en una librería de *Machinelearning* llamado KERNEL DENSITY ESTIMATOR (KDE), el cual utiliza como variables principales el ancho de banda o *bandwidth*, h , y la función de distribución de densidad estelar (gaussiana). Para obtener los valores de h apropiados confeccionaron inicialmente una lista de 68 cúmulos estelares conocidos en la región estudiada, utilizando el catálogo de Bica et al. (2008), de los cuales extrajeron los valores de los radios y densidades medias. En particular, utilizaron 3 valores de h diferentes, correspondientes a los valores mínimo, inter-

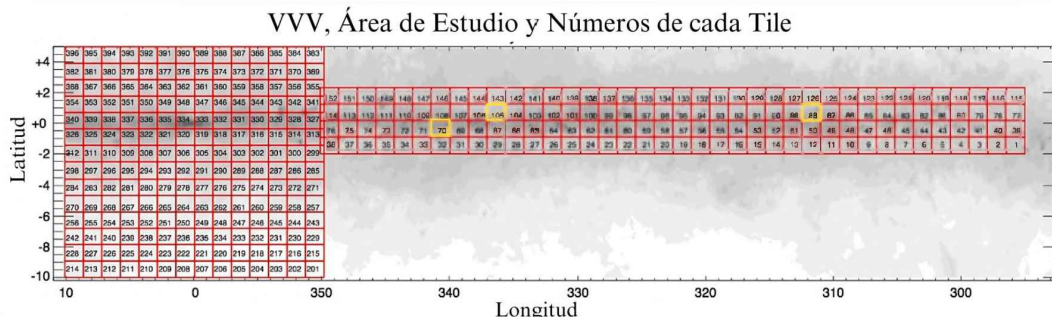


Fig. 1. Esquema de las regiones de la Galaxia observada por VVV representadas con grillas rojas. Dichas regiones fueron relevadas uniformemente con campos de 2 grados de lado aproximadamente llamados *tiles*. En amarillo se representan los *tiles* seleccionados en este trabajo (d070, d105, d088 en orden de izquierda a derecha). Imagen obtenida de la página oficial de VVV survey (<https://vvvsurvey.org/about/survey-area/>).

medio y máximo de los radios de los cúmulos previamente conocidos. De todas las sobredensidades detectadas por KDE en las imágenes VMC, Piatti et al. (2016) seleccionaron aquellas cuya densidad y extensión (radio) se encontraran dentro de los rangos respectivos de los cúmulos estelares conocidos en la región analizada. De la búsqueda de nuevos candidatos en la misma región del cielo, pero a partir de datos en el infrarrojo, Piatti et al. (2016) detectaron 143 sobredensidades estelares nuevas. Posteriormente, realizaron una limpieza de posibles estrellas de campo, obteniendo así 38 objetos candidatos a cúmulo estelar genuinos.

En el presente trabajo desarrollaremos un método de detección y análisis automático de cúmulos estelares y lo aplicamos a los mismos datos utilizados por Borissova et al. (2011). El método propuesto se basará principalmente en la implementación de KDE. A partir de su utilización y de sus resultados comparados con aquéllos de Borissova et al. (2011) evaluamos su rendimiento en la detección de cúmulos estelares en la región infrarroja del espectro electromagnético.

2. Método

La detección de candidatos a cúmulos estelares la llevamos a cabo utilizando como base de datos los *tiles* d070, d088, d105 del relevamiento VISTA VVV, mostrados con amarillo en la Fig 1. Elegimos estos *tiles* porque contienen un número significativo de candidatos a cúmulos estelares identificados visualmente por Borissova et al. (2011), lo cual permite evaluar mejor estadísticamente los métodos semiautomáticos de búsqueda de cúmulos. Cada *tile* contiene 5 candidatos detectados visualmente por Borissova et al. (2011).

El análisis de los 15 candidatos a cúmulos estelares lo llevamos a cabo con un programa de detección semiautomático desarrollado ad-hoc, que tiene en cuenta sobredensidades estelares mayores a un cierto umbral, y la correspondiente distribución de las estrellas que componen una sobredensidad en el diagrama color-magnitud. Luego, comparamos nuestros resultados con los de Borissova et al. (2011).

A continuación describimos la metodología, paso a paso.

2.1. Comprobación de candidatos

En primer lugar seleccionamos una región de 0.11 grados de radio alrededor de las coordenadas centrales de los 15 candidatos a cúmulos estelares seleccionados. Dichas coordenadas centrales las tomamos de Borissova et al. (2011). El tamaño del campo lo elegimos de modo que en ellos se incluyera no solamente el campo del candidato a cúmulo estelar, sino también una buena parte de su campo circundante, por razones que resultarán claras más adelante. Luego, graficamos las posiciones (R.A. y Dec., por sus siglas en inglés) de todas las estrellas observadas en el campo de cada candidato a cúmulo estelar en un área cuadrada de 0.11 grados de lado. Sobre estos campos seleccionados aplicamos KDE para reconocer dichas sobredensidades estelares. Además, realizamos un conteo estelar para cada candidato para verificar los datos extraídos de Borissova et al. (2011).

2.2. Mapas de densidad estelar

KDE provee de información muy valiosa para detectar la existencia de sobredensidades reales de entre cientos de miles de estrellas examinadas en la base de datos VVV. Como resultado de diferentes corridas de KDE, obtuvimos las posiciones centrales, en ascensión recta y declinación, (R.A. y Dec.) de las sobredensidades estelares detectadas. Para filtrar sobredensidades estelares espúreas, por ejemplo, aquellas debidas a fluctuaciones de la densidad estelar del campo, sustrajimos a las densidades obtenidas por KDE el valor medio de la densidad real de cada campo en los alrededores del candidato a cúmulo estelar, y dividimos dicho valor por la dispersión de la densidad de las estrellas del campo, que representamos con la variable sigma (σ). La Fig. 2 muestra la distribución de densidad " σ " en función de la posición espacial de las estrellas.

El modo concreto de utilizar KDE lo detallamos a continuación: Inicialmente generamos una grilla de celdas cuadradas distribuidas en todo el campo a explorar; en cada una de ellas KDE aplica una función de distribución de densidades definida previamente. El tamaño de las celdas cuadradas elegido fue ($\Delta Dec.$, $\Delta R.A.$) = (826pixel, 826pixel) (0.479 segundos de lado), y el mismo resultado suficientemente grande como para evitar detec-

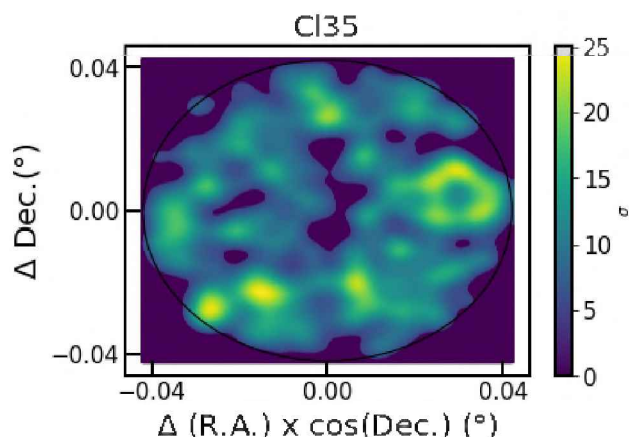


Fig. 2. Mapa de densidad estelar generado por KDE para el candidato a cúmulo estelar Cl35. El círculo negro está centrado en el objeto y su radio es el estimado por Borissova et al. (2011). El Norte está hacia arriba y el Este hacia la izquierda.

tar fluctuaciones espúreas y suficientemente chico como para detectar sobredensidades reales pequeñas. De entre todos los parámetros que tuvimos que variar dentro de KDE, éste resultó el más sensible en cuanto al tiempos de ejecución que demandó cada corrida de KDE.

Para utilizar un valor de ancho de banda apropiado decidimos adoptar un valor igual a 4 veces el radio más pequeño de los candidatos a cúmulo estelar, expresado en grados, de modo que cada sobredensidad estelar asociada a un candidato a cúmulo estelar pueda ser mapeada por KDE por una cantidad razonable (número de celdas) de funciones de distribución. El candidato con menor radio en nuestra muestra es Cl88 con un valor de 0.0033grados. En base a esto, adoptamos $h = 0.0012$ grados.

2.3. Comparación con diagramas color-magnitud

Como última condición necesaria para corroborar si efectivamente las sobredensidades estelares reportadas en Borissova et al. (2011) son candidatos a cúmulos estelares, analizamos sus diagramas color-magnitud. Para construir este diagrama, utilizamos las magnitudes en los filtros Ks y J.

Para confeccionar el diagrama color-magnitud de cada candidato a cúmulo estelar tomamos las estrellas contenidas en un círculo con centro correspondiente a las coordenadas R.A. y Dec. y el radio reportado en Borissova et al. (2011). El diagrama color-magnitud resultante lo comparamos con aquel de un campo estelar adyacente, para un círculo de igual área. Si el objeto de interés es un candidato a cúmulo estelar, su diagrama color-magnitud contiene secuencias de estrellas que se distinguen del diagrama color magnitud de las estrellas del campo adyacentes al objeto.

3. Análisis y resultados

Luego de realizadas las corridas de los métodos semi-automáticos mencionados obtuvimos los siguientes resultados. Las Figs. 3 y 4 muestran los resultados obtenidos a partir de la aplicación del método semi-automático para Cl36. Se realizó un análisis de completitud de KDE de forma individual en los 3 *tiles* seleccionados, resultando 100 %, 100 %, y 75 %, respectivamente.

3.1. Sobredensidades detectadas con KDE

A partir del conteo estelar observamos que el número de estrellas contabilizadas computacionalmente dentro del radio de cada candidato a cúmulo estelar es aproximadamente un 100 % mayor que el número total de estrellas miembros estimado por Borissova et al. (2011). Esta diferencia puede ser debida a las limitaciones del método visual para reconocer estrellas detrás de otras, aunque ellos no mencionan si el número de estrellas miembros obtenido incluye la substracción estadística de estrellas del campo.

Los mapas de densidad estelar para los candidatos a cúmulo estelar Cl35, Cl36, Cl37, Cl73, Cl74, Cl86, Cl87, Cl88 muestran picos de sobredensidad centrados en las coordenadas de los candidatos identificados por Borissova et al. (2011), o poseen múltiples picos dentro del radio respectivo. En cambio, los candidatos Cl33 y Cl34 presentan sobredensidades con picos que están desplazados ligeramente de las posiciones centrales de los mismos, aunque en todos los casos dentro del radio de cada candidato a cúmulo estelar. Por otro lado, Cl76 y Cl90 no poseen ningún pico de densidad dentro del radio respectivo, por lo que podemos concluir que estos candidatos a cúmulo estelar no presentan ninguna sobredensidad.

3.2. Diagramas color-magnitud

Realizamos un análisis de los diagramas color-magnitud construidos y concluimos que los candidatos a cúmulo estelar Cl36, Cl37, Cl74, Cl87 y Cl88 poseen secuencias de estrellas que insinúan la presencia de un cúmulo estelar. Mientras que para los candidatos Cl33, Cl34, Cl35, Cl73, Cl74, Cl76, Cl88 y Cl90 no detectamos secuencias de estrellas que puedan pertenecer a un cúmulo estelar.

4. Discusión y conclusiones

Luego de inspeccionados y analizados los gráficos de distribución espacial y distribución de densidades decidimos descartar los candidatos a cúmulo estelar Cl76 y Cl90, debido a que ninguno cumple con la condición de conformar una sobredensidad estelar; no fueron detectados por KDE por lo que no los tuvimos en cuenta en adelante. También descartamos los candidatos a cúmulo estelar Cl77, Cl78 y Cl91 debido a que las estrellas de la base de datos VVV que componen dichos candidatos no está completa y presenta errores. Por otro lado, los candidatos a cúmulo estelar Cl34, Cl35, Cl73, y Cl88 no presentan una secuencia de estrellas propias de un cúmulo estelar en su diagrama color-magnitud, y por lo tanto también los descartamos como candidatos

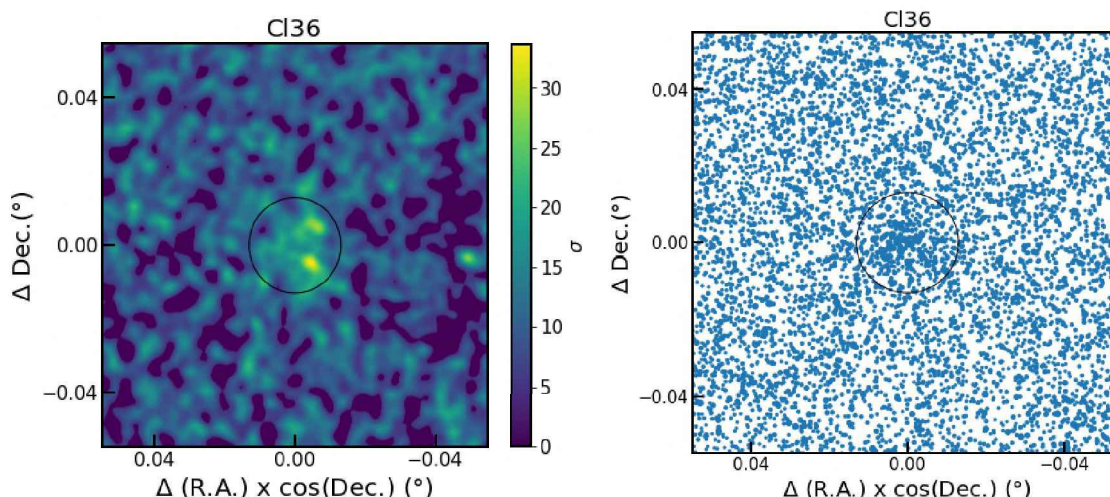


Fig. 3. *Panel izquierdo:* Diagrama de distribución de sobredensidades, *Panel derecho:* Diagrama de distribución espacial de las estrellas pertenecientes al candidato a cúmulo Cl36 y estrellas del campo cercano. El círculo negro representa el radio reportado por Borissova et al. (2011).

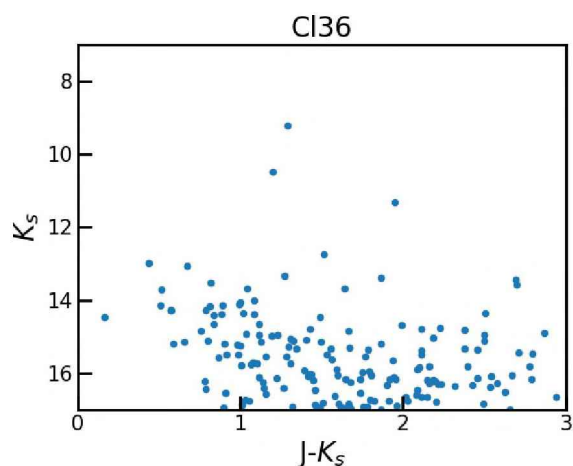


Fig. 4. Diagrama color-magnitud, sin descontaminar, del candidato a cúmulo estelar Cl36.

a cúmulo. Finalmente, los 6 candidatos a cúmulo estelar Cl133, Cl136, Cl137, Cl174, Cl187 y Cl166 presentan sobredensidades estelares y diagramas color-magnitud compatibles con los de cúmulos estelares, y por lo tanto confirmamos los mismos como candidatos a cúmulo estelar. Concluimos que de 10 candidatos a cúmulos estelares detectados en Borissova et al. (2011), solamente 6 fueron detectados por KDE. Esto implica que el rendimiento o capacidad de KDE para detectar los candidatos a cúmulo estelar de Borissova et al. (2011) es de: 6/10, lo cual representa un 60%. Este porcentaje hace referencia a candidatos a cúmulo estelar donde KDE detecta sobredensidades bien definidas y sus respectivos diagramas color-magnitud también muestran secuencias com-

patibles con las de cúmulos estelares. De los candidatos descartados, donde la sobredensidad no es clara o los diagramas color-magnitud de sobredensidades no muestran secuencias claras de un cúmulo estelar no podemos concluir definitivamente que no sean cúmulos estelares. Especulamos que esos candidatos a cúmulo estelar reconocidos por inspección visual pueden estar constituidos por pocas estrellas o embebidos en nebulosas, las cuales fueron identificadas por Borissova et al. (2011) como candidatos a cúmulo estelar.

Una de las ventajas que hemos identificado en los métodos computacionales es que las coordenadas centrales de las sobredensidades son estimadas con mayor certeza, al igual que la estimación de sus extensiones (radios). Por ejemplo, las sobredensidades detectadas para los candidatos a cúmulo estelar Cl133 y Cl134 están desplazadas ligeramente de las coordenadas provistas por Borissova et al. (2011). Por otro lado, hemos detectado que los candidatos a cúmulo estelar estudiados y finalmente confirmados se distinguen de la densidad del fondo del cielo en $\sigma > 5$. De aquí concluimos que la sensibilidad del ojo humano permitiría distinguir sobredensidades si ellas sobresalen claramente de la distribución de las estrellas del campo.

Referencias

- Bica E., et al., 2008, Monthly Notices of the Royal Astronomical Society, 389, 678
 Borissova J., et al., 2011, Astronomy & Astrophysics, 532, A131
 Ferreira F.A., et al., 2020, Monthly Notices of the Royal Astronomical Society, 496, 2021
 Piatti A.E., et al., 2016, Monthly Notices of the Royal Astronomical Society, 460, 383