

Identificación de cúmulos estelares en Gaia DR3 para correlacionarlos con los brazos espirales de la Vía Láctea

Tesis presentada para cumplir con los requisitos del título de Astrónomo

SANTIAGO LASTRA RAMIREZ

FACULTAD DE CIENCIAS EXACTAS Y NATURALES

UNIVERSIDAD DE ANTIOQUIA. INSTITUTO DE FÍSICA
Medellín, Colombia 2024
www.udea.edu.co

TRABAJO DE GRADO 2024

**En el camino de crear un gran catálogo
de cúmulos estelares**

SANTIAGO LASTRA RAMIREZ



UNIVERSIDAD DE ANTIOQUIA

1 8 0 3

Instituto de Física
UNIVERSIDAD DE ANTIOQUIA
Medellín, Colombia 2024

En el camino de crear un gran catálogo de cúmulos estelares
SANTIAGO LASTRA RAMIREZ

© SANTIAGO LASTRA RAMIREZ, 2024.

Supervisor: Esteban Silva, FACOM (Física y Astrofísica Computacional)

Trabajo de Grado 2024
Instituto de Física
Universidad de Antioquia
Calle 67 No. 53 - 108, Medellín
Colombia
Teléfono [+57-604] 219 8332

Escrito en L^AT_EX
Medellín, Colombia 2024

En el camino de crear un gran catálogo de cúmulos estelares

SANTIAGO LASTRA RAMIREZ

Instituto de Física

Universidad de Antioquia

Resumen

El estudio de los cúmulos estelares es esencial para comprender la evolución estelar, la dinámica gravitacional de las estrellas y la estructura evolutiva de la galaxia. En este trabajo, se ha generado un catálogo de 534 cúmulos mediante métodos de agrupamiento como DBSCAN y pyUPMASK, permitiendo la estimación de parámetros clave como la edad, metalicidad y extinción a través del ajuste de la isocrona al CMD del cúmulo. Este catálogo, centrado en el catálogo de Hunt and Reffert (2023), muestra una significativa correlación con otros catálogos, como los de Cantat-Gaudin et al. (2020) y Bossini et al. (2019). Además, se concluye que los cúmulos más jóvenes o con una mayor metalicidad tienden a ubicarse más cerca del plano y a los brazos de la galaxia.

Palabras clave: Cúmulos estelares, brazos de la galaxia.

Lista de Acrónimos

KPC	Kilo Pársec
GC	Cúmulo Globular
OC	Cúmulo Abierto
CMD	Diagrama Color Magnitud
HR	Hertzsprung-Russell
MS	Secuencia Principal (<i>Main Sequence</i>)
RGB	Rama de Gigantes Rojas (<i>Red-Giant Branch</i>)
HB	Rama Horizontal (<i>Horizontal Branch</i>)
AGB	Rama de Gigantes Asintóticas (<i>Asymptotic Giant Branch</i>)
HR23	Catálogo de cúmulos identificados por Hunt & Refert 2023
KMS	Método K-means utilizado en el algoritmo UPMASK
RFR	Método de rechazo utilizado en el algoritmo UPMASK
KDE	Estimación de densidad de Kernel (<i>Kernel Density Estimation</i>)
HR23	Catálogo de Hunt & Refert 2023

Nomenclatura

Índices

i, j	Indicar elementos individuales
α	Indicar ascensión recta
δ	Indicar declinación

Variables

μ_α o $pmRA$	Movimiento propio en ascensión recta
μ_δ o $pmDEC$	Movimiento propio en declinación
l	Latitud galáctica
b	Longitud galáctica
ϖ	Paralaje
ϵ	Parámetro de DBSCAN

Contenido

Lista de Acrónimos	vii
Nomenclatura	ix
Lista de Figuras	xiii
1. Introducción	1
1.1. GAIA	1
1.2. Cúmulos Estelares	2
1.2.1. Propiedades de los cúmulos estelares	3
1.2.2. Tipos de cúmulos estelares	4
1.2.2.1. Cúmulos Abiertos	4
1.2.2.2. Cúmulos Globulares	4
1.2.3. Formación de Cúmulos Estelares	4
1.2.3.1. Colapso de la nube interestelar	4
1.2.3.2. Fragmentación de la Nube	5
1.2.3.3. Formación del Núcleo Protoestelar	5
1.2.3.4. Acreción	5
1.2.3.5. Disociación e Ionización	5
1.2.3.6. Fase de estrella de presecuencia principal	6
1.2.4. Formación de cúmulos	6
1.2.5. Importancia de los cúmulos estelares	6
1.3. Distribución de Cúmulos en la Vía Láctea	7
1.3.1. Cúmulos en el Disco	8
1.3.1.1. Características del disco de la Vía Láctea	8
1.3.1.2. Distribución de los cúmulos abiertos en el disco de la Vía Láctea	9
1.3.2. Cúmulos en el Halo y en el Bulbo	10
1.4. Diagramas HR	12
1.4.1. Regiones de los diagramas HR	13
1.4.1.1. Secuencia Principal	13
1.4.1.2. Subgigantes (SGB)	14
1.4.1.3. Rama de Gigantes Rojas (RGB)	14
1.4.1.4. Rama Horizontal (HB)	14
1.4.1.5. Rama de Gigantes Asintóticas (AGB)	14
1.4.1.6. Enanas Blancas	15
1.4.2. Movimiento de una estrella a través del diagrama	15
1.4.3. Importancia de los diagramas HR en los cúmulos estelares	16

2. Objetivos	19
2.1. Objetivo General	19
2.2. Objetivos Específicos	19
3. Metodología	21
3.1. Compilación de Catálogos	21
3.2. Cúmulos con menos de 100 estrellas	21
3.3. Importar estrellas de GAIA	23
3.4. DBSCAN	24
3.4.1. DBSCAN comparado con cúmulos de la bibliografía	26
3.5. pyUPMASK	28
3.5.1. Probabilidad de Membresía	29
3.6. fitCMD	31
3.6.1. Estudio fitCMD	32
3.6.1.1. Comparación de cúmulos según la edad	32
3.6.1.2. Comparación cúmulos según la distancia	34
3.7. Obtención parámetros astrométricos	35
4. Resultados	39
4.1. Catálogo	39
4.1.1. Edad dinámica del catálogo	40
4.1.2. Comparación de valores con HR23	42
4.1.3. Cúmulos detectados en relación con otros catálogos	45
4.1.4. Coeficiente de Spearman	45
4.1.5. Correlación de la metalicidad con otros catálogos	47
4.1.6. Cúmulos no detectados	48
4.1.7. Perspectivas Futuras	50
4.2. Estructura de la galaxia	51
4.2.1. Brazos de la galaxia	51
4.2.1.1. Metalicidad	53
4.2.2. Radiogalactocéntrico y coordenada Z	54
5. Conclusiones	57
Bibliography	59

Listas de Figuras

1.1.	Distribución de las distintas secciones de la galaxia	8
1.2.	Diagrama de la galaxia vista de cara	10
1.3.	Distribución de cúmulos abiertos en el disco de la galaxia	11
1.4.	Distribución bimodal de la metalicidad del bulbo	12
1.5.	Diagrama HR	13
1.6.	Ramas en el diagrama HR	15
1.7.	HR para estrellas de masas distintas	16
1.8.	Isocronas de NGC752 e IC4651	17
3.1.	Edad Dinámica para HR23	22
3.2.	Error nominal en el paralaje para distintas magnitudes	24
3.3.	Ejemplo de DBSCAN	25
3.4.	DBSCAN: NGC 6253	26
3.5.	DBSCAN: NGC 2682 comparación con el catálogo de Cantat Gaudin . .	27
3.7.	pyUPMASK: NGC 1039 PM % mayor y menor que el 0.5	30
3.6.	pyUPMASK: probabilidad de membresía de NGC 1039	30
3.8.	Clasificación cúmulos de HR23 con fitCMD	32
3.9.	Edad cúmulos GOOD y MEDIUM de HR23 con fitCMD	33
3.10.	Distancia cúmulos GOOD y MEDIUM de HR23 con fitCMD	34
3.11.	Uso de KDE para obtención de la moda	36
4.1.	Cálculo del radio donde se encuentra la mitad de la luminosidad . .	40
4.2.	Edad dinámica para el catálogo de este trabajo	41
4.3.	Múltiple CMD de OC	42
4.4.	Diferencia de distancia en el plano galáctico	43
4.5.	Diferencia de distancia en la coordenada Z	44
4.6.	Histogramas de Spearman	46
4.7.	Cúmulos no incluidos por mal ajuste	48
4.8.	Extinción en cúmulos detectados y no detectados	49
4.9.	Distribución espacial de cúmulos no detectados	49
4.10.	Mal ajuste de Alessi 33 con FitCMD	50
4.11.	Cúmulos distribuidos en el plano de la galaxia	51
4.12.	4 intervalos de edad distribuidos en el plano de la galaxia	52
4.13.	Extinción de los cúmulos distribuidos en la galaxia	53
4.14.	Relación metalicidad con Reid y Poggio	54
4.15.	Cúmulos distribuidos en el eje Z galáctico	55

Capítulo 1

Introducción

Los cúmulos estelares han permitido estudiar tanto la dinámica colectiva de estrellas sujetas a una gravedad mutua como la composición, estructura y dinámica de una galaxia. Por ejemplo, los cúmulos globulares en la Vía Láctea están principalmente en el Halo y en el Bulbo, mientras que los cúmulos abiertos están sobre el disco de la galaxia.

Actualmente múltiples trabajos han estudiado la relación que tienen los cúmulos abiertos jóvenes con los brazos de la galaxia, algunos de estos son: [Cantat-Gaudin et al. \(2018\)](#), [Castro-Ginard et al. \(2020\)](#), [Bossini et al. \(2019\)](#), [He et al. \(2021a\)](#) y [Hunt and Reffert \(2023\)](#). Estos estudios emplean el catálogo de GAIA y utilizan algoritmos como DBSCAN y pyUPMASK para identificar cúmulos en todo el cielo. Además mediante el ajuste de isocronas determinan la edad, metalicidad y masa de los cúmulos.

En este trabajo se hará un procedimiento similar al realizado por los anteriores autores estudiando los parámetros astrofísicos de cúmulos que tengan más de 100 estrellas. Y así observar, mediante un catálogo propio, cuál es la relación que hay entre los cúmulos jóvenes con el modelo de [Reid et al. \(2014\)](#) para los brazos de la galaxia.

Antes de continuar con la metodología y con una explicación detallada a las consideraciones tomadas, es necesario explicar los conceptos utilizados en todo el trabajo:

1.1 | GAIA

La misión Gaia nace con el objetivo de crear un mapa tridimensional de la Vía Láctea, estudiando el movimiento, posición, temperatura y edad de las estrellas. También busca estudiar el pasado y el futuro de nuestra galaxia utilizando un telescopio capaz de medir datos astrométricos de las estrellas, planetas, asteroides y otros objetos celestes ([Gaia Collaboration and Prusti \(2016\)](#)).

Esta misión nace gracias a la “European Space Agency” (ESA) en busca de crear una de las misiones más ambiciosas en la historia de la astronomía. La construcción de la nave espacial fue aceptada en el año 2006 y el 19 de diciembre del año 2013 fue lanzada al espacio, 7 meses después (el 19 de Julio de 2014) la nave ya se encontraba en el segundo punto de Lagrange para el sistema Sol-Tierra-Luna y lista para la

toma de datos. Desde ese momento hasta el día de hoy se han obtenido más de 1.8 mil millones de objetos celestes ([Gaia Collaboration and Brown \(2021\)](#)).

Dos años después de tener a la nave en órbita (septiembre de 2016) sale el primer data release con información de posición y brillo de 1142 millones de fuentes con una cobertura del 90 % del cielo. Aunque, únicamente se obtuvieron la distancia y movimiento de 2 millones de estas con una precisión astrométrica de alrededor de 10 micro arco segundos para cuerpos brillantes y de 200 micro arco segundos para cuerpos débiles; también, una precisión fotométrica de alrededor del 1 % para cuerpos brillantes y del 20 % para cuerpos débiles ([Gaia Collaboration and Brown. \(2016\)](#)).

El segundo data release, que fue publicado en abril del 2018, trajo consigo posiciones y brillos de 600 millones de fuentes más que el primero y 1300 millones de fuentes con distancias y movimientos propios. Además, la precisión astrométrica y fotométrica aumentaron el doble, y se crearon nuevos elementos: Información de color en las bandas GBP y GRP, velocidades radiales para 7 millones de fuentes; estimaciones de la temperatura efectiva, extinción, enrojecimiento, radio y luminosidad estelar para entre 77 y 161 millones de fuentes, y para 14 000 planetas menores se presenta la astrometría y fotometría de época ([Gaia Collaboration and Brown \(2018\)](#)).

Finalmente con su último data release publicado en su totalidad el 13 de junio de 2022 se mejoró la precisión de los datos astrométricos y fotométricos, y se aumentó la cantidad de fuentes observadas hasta un total de 1800 millones con 1600 millones con distancias y movimientos propios. Además, se analizaron objetos celestes que no se habían analizado en los pasados data release, como por ejemplo los sistemas multi estelares ([Gaia Collaboration and Brown \(2021\)](#)).

Todos estos datos fueron obtenidos utilizando varios instrumentos con los que está dotado GAIA, los más importantes son: el instrumento fotométrico que captura longitudes de onda entre los 330 nm y los 1050 nm, el espectrómetro de velocidad radial y el instrumento astrométrico el cual consta de dos telescopios donde el espejo primario de cada telescopio es de 1.45m de diámetro y con una longitud focal de 35m ([Gaia Collaboration and Prusti \(2016\)](#)).

Gracias a GAIA se han obtenido características de cuerpos estelares que no se habían obtenido antes y con una alta precisión. Se espera que para el finales del 2025 con la publicación del DR4 se tenga un catálogo completo de todos los cuerpos celestes observados con datos astrométricos, fotométricos y de velocidad radial.

1.2 | Cúmulos Estelares

Los cúmulos estelares son colecciones de estrellas que pueden tener una estructura tanto esférica (cúmulos globulares) o irregular (cúmulos abiertos). Estos objetos celestes son importantes para la astronomía ya que permiten estudiar la historia, formación y estructura tanto de la Vía Láctea como de las galaxias en general. Esto es debido a que las estrellas en los cúmulos nacieron prácticamente al mismo tiempo

y de la misma nube molecular o nebulosa interestelar, dando paso a poder estudiar la edad y la metalicidad de una región del cielo ([Carroll and Ostlie \(2017\)](#)).

1.2.1. Propiedades de los cúmulos estelares

Los cúmulos estelares pueden tener distintas formas dependiendo del tipo de cúmulo que sea y de la ubicación del cúmulo en la galaxia, ya que si están por fuera del disco es más probable que tengan una simetría esférica, mientras que si están dentro de este pueden tener simetrías irregulares; esto debido a la cercanía de los cuerpos y por lo tanto su atracción gravitacional.

También es importante mencionar que hay cúmulos estelares viejos y cúmulos estelares jóvenes, donde su edad es determinada gracias a sus diagramas color-magnitud (CMD). Los cúmulos estelares jóvenes están compuestos de estrellas de población 1 y tienden a tener mucho más gas en su interior que los cúmulos viejos, dando paso a que sigan formando estrellas. A diferencia de estos, los cúmulos viejos están compuestos de estrellas de población 2 ([LeBlanc \(2010\)](#)).

Cabe recalcar que se asume que las estrellas de ambos tipos de cúmulos fueron creadas prácticamente al mismo tiempo y provenientes de la misma nube; por lo tanto las estrellas están formadas con, esencialmente, metalicidades similares. Aunque esto no quiere decir que todas las estrellas tienen el mismo tiempo de vida, las estrellas grandes tienen menos tiempo de vida que las estrellas chicas ([Pols \(2011\)](#)).

Además, es correcto pensar que en el cúmulo no siempre se tendrá la misma cantidad de estrellas, esto debido a que hay muchas estrellas que pueden desprenderse gracias a que se desligan gravitacionalmente por masas vecinas al cúmulo. Una de las masas más importantes a tener en cuenta es la masa de la galaxia, la cual le ejerce al cúmulo una fuerza gravitacional constante o variable dependiendo si este se mantiene a una misma distancia o si se acerca a los brazos de la galaxia.

Si el cúmulo se mantiene aproximadamente a una misma distancia de la galaxia, la fuerza gravitacional ejercida es constante tanto en magnitud como en dirección. Una fuerza de marea constante es importante en la tasa de evaporación producida por los campos internos fluctuantes dentro del cúmulo, disminuyendo así la energía requerida para escapar de este y produciendo un aumento en la tasa de abandono.

Por otro lado, si el cúmulo cruza el plano galáctico la fuerza gravitacional infringida por la galaxia hará que este se comprima en dirección perpendicular del plano galáctico y por lo tanto caliente las partes más externas del cúmulo.

Por último, otro efecto que puede causar el desprendimiento de estrellas (pero en una menor medida), es cuando un cúmulo pasa a través de una nube o cuando cruza cerca a otro cúmulo, estos dos mecanismos también calientan las partes externas pero no causan un gran cambio dentro del mismo ([Spitzer Jr \(2014\)](#)).

1.2.2. Tipos de cúmulos estelares

En general hay dos tipos de cúmulos: los globulares y los abiertos, pero cada uno de estos se puede dividir en más subgrupos dependiendo de la cantidad de estrellas, de su edad, de su lejanía a la galaxia.

1.2.2.1. Cúmulos Abiertos

Son sistemas estelares irregulares que poseen entre 10^2 y 10^4 estrellas y su tamaño es de aproximadamente 10 pc. Son cuerpos relativamente jóvenes compuestos de estrellas de población 1 y continuamente se están formando muchos de estos en el disco galáctico. La edad de la mayoría de estos cúmulos es menor a 1 Gyr ya que cúmulos que se hayan formado antes lo más probable es que hayan perdido a sus estrellas debido a la atracción gravitacional de la galaxia. Hoy en día se cree que la mayoría de las estrellas en el disco galáctico fueron formadas en cúmulos abiertos que se disolvieron.

1.2.2.2. Cúmulos Globulares

Son mucho más masivos que los cúmulos abiertos, poseen entre 10^4 y 10^6 estrellas, tienen simetría elíptica o esférica. Estos cuerpos carecen de gas, polvo o estrellas jóvenes ya que son cuerpos viejos y gran parte del material que se utilizaba para formar estrellas ya fue consumido o expulsado a través de fuertes vientos solares o explosión de supernovas. Además las estrellas que están dentro del cúmulo ya han completado muchas órbitas ($\sim 10^4$) desde que se formó el cúmulo; por lo tanto, ya superaron su tiempo de relajación olvidando cualquier vestigio de su velocidad inicial.

Estos cuerpos son interesantes ya que tienen una alta densidad estelar, típicamente en el centro de los cúmulos globulares este valor es de $10^4 M_{\odot} \text{ pc}^{-3}$ comparado con un $0.05 M_{\odot} \text{ pc}^{-3}$ en el vecindario solar.

1.2.3. Formación de Cúmulos Estelares

Para explicar la formación de cúmulos estelares es necesario primero explicar la formación de estrellas en sí. Las observaciones han indicado que las estrellas se forman típicamente en las nubes moleculares con masas del orden de $10^5 M_{\odot}$, con dimensiones de ~ 10 parsec, temperaturas entre 10-100 K y densidades de 10-300 moléculas/cm³. Estas nubes están en equilibrio hidrostático y se pueden identificar 6 etapas en el proceso de formación de la estrella.

1.2.3.1. Colapso de la nube interestelar

La formación estelar empieza cuando la nube sufre una perturbación, ya sea por una explosión cercana de una supernova o por la colisión con otra nube, causando una fractura en el equilibrio hidrostático y posteriormente haciendo que la nube colapse por causa de la gravedad. El criterio para que una nube colapse fue determinado a principios del siglo XX por el matemático James Jean, el cual determina que si

una nube tiene una masa mayor que la masa de Jeans (fórmula 1.1) implica una contracción en la nube.

$$M_J = 4 \times 10^4 M_{\odot} \left(\frac{T}{100 \text{ K}} \right)^{3/2} \left(\frac{n}{\text{cm}^{-3}} \right)^{-1/2} \quad (1.1)$$

Donde n es la densidad numérica. Tomando valores típicos de T y de n en nubes moleculares se obtiene que $M_J \sim 10^3 - 10^4 M_{\odot}$. Además, el hecho de que el proceso es dinámico y que la densidad es pequeña, el tiempo que se demora en colapsar la nube es del orden de millones de años.

1.2.3.2. Fragmentación de la Nube

A medida que la densidad en la nube incrementa debido al colapso, la masa de Jeans decrece (fórmula 1.1). Como resultado, el criterio de estabilidad en la nube se rompe y ésta se divide en pequeñas partes, las cuales también se fragmentan. Este proceso de fragmentación puede continuar hasta que la masa de las partes resultantes sea inferior a $0.1 M_{\odot}$ ([Pols \(2011\)](#)).

1.2.3.3. Formación del Núcleo Protoestelar

Cuando la densidad de una nube colapsante aumenta, el gas se vuelve opaco a los fotones infrarrojos, lo que provoca que la radiación quede atrapada en la parte central de la nube. Esto genera un aumento en la presión y el calentamiento del gas, haciendo que el núcleo de la nube entre en equilibrio hidrostático y que el colapso dinámico se ralentice hasta una contracción cuasiestática. En esta etapa, se puede hablar de una protoestrella.

1.2.3.4. Acreción

Debido a que el gas sigue cayendo hacia el núcleo protoestelar y dado que las nubes en contracción contienen una cantidad sustancial de momento angular, el gas que cae forma un disco de acreción alrededor de la protoestrella. Estos discos son característicos del proceso de formación estelar y se observan en muchas estrellas jóvenes.

La acreción de gas genera energía gravitacional, esta energía se divide en dos partes: una parte es radiada y expulsada fuera del disco, mientras que la otra parte se dirige hacia el centro, calentando el núcleo protoestelar.

1.2.3.5. Disociación e Ionización

Cuando la temperatura del núcleo alcanza una temperatura de $\sim 2000 \text{ K}$, el hidrógeno molecular (H_2) empieza a disociarse, causando un aumento en el calor específico. Esto da paso a que el equilibrio hidrostático ya no sea posible y por lo tanto empieza otro colapso dinámico durante el cual la liberación de energía gravitacional es absorbida por las moléculas disociadas sin un aumento significativo de la temperatura. Cuando H_2 se disocia por completo en hidrógeno atómico el equilibrio

hidrostático se recupera y la temperatura empieza a aumentar.

Después de cierto tiempo, cuando la temperatura alcanza alrededor de 10^4 K, comienza otro proceso de colapso dinámico a medida que el hidrógeno y el helio se ionizan. Este colapso continuará hasta que la ionización de la estrella esté completa y se restablezca un equilibrio hidrostático. Como resultado, la estrella experimentará una reducción significativa en su tamaño (Pols (2011)).

1.2.3.6. Fase de estrella de presecuencia principal

Como su nombre lo indica, son estrellas que no han alcanzado la secuencia principal (se explica en la sección 1.4.1.1). Esta fase se caracteriza por ser un proceso en el cual la protoestrella no ha comenzado a tener fusión nuclear, lo que lleva a que la estrella siga contrayéndose para compensar la pérdida de energía por radiación, causando un aumento en la temperatura al interior de la estrella y por lo tanto una quema de hidrógeno en sus núcleos. Esto da paso a que la estrella pase finalmente a la secuencia principal. Larson (2003)

1.2.4. Formación de cúmulos

Como ya se comentó en la sección 1.2.3.2, la nube colapsante se fragmenta en varias secciones y muchas de estas secciones pueden llevar a formar estrellas, estrellas que nacieron prácticamente al mismo tiempo, en el mismo lugar del espacio y con la misma metalicidad; esto da paso a que se formen los cúmulos estelares.

Además, la dependencia de la cantidad de miembros del cúmulo estelar se debe al tamaño de la nube estelar. Ya que si es una nube “pequeña” es imposible que esta forme cúmulos globulares, generalmente estos cúmulos se forman por el colapso de dos o más nubes moleculares. Por otro lado, puede ocurrir incluso que de una sola nube molecular nazcan dos o más cúmulos abiertos, como es el caso de Hodge 301 y R136 que se encuentran en la Gran Nube de Magallanes (Evans et al. (2015)).

1.2.5. Importancia de los cúmulos estelares

Los cúmulos estelares son importantes para la astrofísica por dos motivos principales: uno de ellos es porque permite estudiar tanto la dinámica colectiva de las estrellas sujetas a una gravedad mutua como la dinámica de las estrellas en un ambiente cercano y el segundo motivo es porque permiten dar un trazado a la estructura, composición y dinámica de una galaxia (este último será mejor explicado en la sección 1.3.1).

Para explicar el primero, es necesario decir que es incorrecto aplicar los conceptos de dinámica de gases para los sistemas estelares ya que incluso la estrella promedio puede tener una distancia libre media mayor que el tamaño del sistema (contrario a los gases, el camino libre medio de las partículas de estos suele ser bastante pequeña en comparación con el tamaño del sistema); por lo tanto, conceptos familiares como la presión no pueden aplicarse a los sistemas estelares. Dando paso a decir que para estos sistemas sería inapropiado aplicar las ecuaciones usuales de la dinámica de

fluidos. (Tomado de [Michie \(1964\)](#))

Una forma de determinar el movimiento de las estrellas en los cúmulos es haciendo uso de la ecuación de Boltzmann (ecuación 1.2), la cual describe los procesos dinámicos acercándose más a la física de partículas libres que a la física de fluidos, esto debido a los grandes caminos libres medios de las estrellas. Otra forma de determinar el movimiento de las estrellas es haciendo uso del problema de N-cuerpos.

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = Q(f, f) \quad (1.2)$$

Donde f es una función que depende tanto de la posición como de la velocidad ($f(\mathbf{x}, \mathbf{v})$) y $Q(f, f)$ es conocido como *integral de colisión*, donde se define como una integral en el espacio de momentos sobre el producto de funciones de distribución de una partícula. Tomado de [Gyu Eun Lee \(2019\)](#)

Por otro lado, es importante aclarar que el movimiento de las estrellas en el cúmulo no depende únicamente de la atracción gravitacional entre ellas, también hay factores externos que modifican la trayectoria de una estrella. Como por ejemplo es el caso de la fuerza de marea en un cúmulo que surge del gradiente de la fuerza galáctica a través del cúmulo, esta fuerza puede distorsionar la forma del cúmulo y arrancar estrellas de su órbita ([Kinman \(1962\)](#)). El radio límite de marea está definido como $R(M_c/M_g)^{1/3}$, donde R es la distancia del centro de la galaxia al cúmulo, M_c es la masa del cúmulo y M_g es la masa de la galaxia (esta relación se cumple únicamente si el cúmulo está lo suficientemente lejos del centro de la galaxia).

Para regiones más cercanas al centro galáctico, como lo dice [Michie \(1964\)](#), la fuerza de marea es independiente de la masa del cúmulo y de la distancia al centro galáctico. Por lo tanto, la fuerza de marea no altera significativamente la estructura de los cúmulos en regiones cercanas o intermedias al centro de la galaxia.

1.3 | Distribución de Cúmulos en la Vía Láctea

La Vía Láctea es una galaxia espiral Sbc con 100 – 400 miles de millones de estrellas distribuidas en el Bulbo (tiene una simetría esférica con un radio de 4 kpc), en el disco (tanto delgado como grueso tienen un radio de 25 kpc) y un Halo (que puede ser tanto estelar como de materia oscura. Tiene simetría esférica con un radio $\gtrsim 100$ kpc y $\gtrsim 230$ kpc, respectivamente). Todas estas secciones tienen estrellas distribuidas en su radio, aunque, no todas estas estrellas tienen la misma metalicidad ni edad. Dependiendo de la sección es más probable encontrar ciertas estrellas, por ejemplo, en el Bulbo y en el Halo están principalmente los cúmulos globulares mientras que en el disco están los cúmulos abiertos.

Se puede observar en el gráfico 1.1 cómo es la forma y tamaño de las distintas partes de la Vía Láctea, cabe aclarar que el halo no está a escala ya que el halo estelar es aproximadamente 4 veces el tamaño del disco y el halo de materia oscura es aproximadamente 20 veces el tamaño del disco. Además, se plasma cuál es la distancia

del centro de la galaxia al Sol, pero es importante decir que el Sol tampoco está a escala, si este estuviera a escala en la imagen su tamaño sería más pequeño que un píxel (Carroll and Ostlie (2017)).

Se procede entonces a estudiar las distintas regiones de la galaxia a través del estudio de los cúmulos estelares que se encuentran en cada una de ellas.

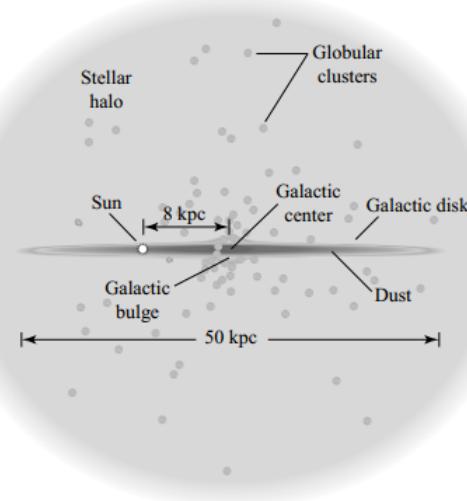


Figura 1.1: Distribución de las distintas secciones de la galaxia, tomado de Carroll and Ostlie (2017).

1.3.1. Cúmulos en el Disco

1.3.1.1. Características del disco de la Vía Láctea

El disco de la Vía Láctea está compuesto por el disco delgado y el disco grueso: El disco delgado, el cual está conformado por estrellas jóvenes, polvo y gas, y tiene una escala vertical de 350 pc. Dentro del disco delgado, encontramos una subregión conocida como el disco delgado joven, que es el plano central de distribución de gas y polvo en la galaxia, tiene una altura de alrededor de 90 pc y es donde están las estrellas más jóvenes de toda la galaxia. Por otro lado, está el disco grueso, tiene una escala vertical de 1000 pc y se caracteriza por tener estrellas en menor cantidad y más viejas que en el disco delgado (Carroll and Ostlie (2017)).

Otro factor importante es la metalicidad en ambas regiones. En el disco delgado las estrellas tienen metalicidades entre $-0.5 < [Fe/H] < 0.3$ mientras que en el disco grueso $-0.6 < [Fe/H] < -0.4$ dando a entender que el disco delgado tiene estrellas con una mayor metalicidad que el disco grueso. Esto ocurre, en cierta parte, debido a la explosión de supernovas, ya que estas envían hierro al medio interestelar que próximamente será adquirido por nuevas estrellas en el momento de su formación. Así que, la metalicidad se correlaciona con la edad, siendo las estrellas más jóvenes y recientemente formadas las que tienen la mayor abundancia de hierro. Y esto es correcto según el estudio realizado por Bensby et al. (2014), el cual dice que la edad

media de las estrellas del disco delgado está entre 7 – 8 Gyr, mientras que para las estrellas del disco grueso la edad media está entre 9 – 10 Gyr.

1.3.1.2. Distribución de los cúmulos abiertos en el disco de la Vía Láctea

El disco de una galaxia espiral (como la Vía Láctea) tiene mucho más polvo y gas que una galaxia elíptica; por lo tanto, la formación estelar está activa. Según la NASA ([Dunbar \(2008\)](#)) día a día se crean 7 estrellas en el disco de la Vía Láctea, estas estrellas jóvenes que tienen una alta metalicidad debido al enriquecimiento en metales del medio intergaláctico, se les llama estrellas de Población 1.

Actualmente en el disco delgado de la Vía Láctea se tienen estrellas de Población 1, mientras que en el disco grueso hay estrellas tanto de Población 1 como de Población 2. Las estrellas de Población 1 están caracterizadas por ser estrellas jóvenes, ricas en metales y fueron formadas en cúmulos abiertos ([Guidry \(2019\)](#)). Las estrellas de Población 2 son estrellas con menos metalicidad y más viejas que las estrellas de población 1, se encuentran en mayor medida en el bulbo de la galaxia o en los cúmulos globulares del halo galáctico. Por otro lado, están las hipotéticas estrellas de Población 3, las cuales son estrellas que nacieron en el universo primitivo y que poseen una metalicidad casi nula; son hipotéticas ya que hasta el momento no se ha observado una de estas.

Las estrellas de Población 1 son las que se encuentran principalmente en el disco de la galaxia y se distribuyen alrededor de este según las sobredensidades de HII, de gas molecular, de polvo y de estrellas ([Van den Bosch et al. \(2010\)](#)). Estas sobredensidades son las que representan los brazos de la galaxia y se pueden observar en la figura 1.2. Además debido a que las estrellas de esta población fueron formadas en cúmulos abiertos, por lo tanto es recíproco decir que la gran mayoría de estos cúmulos se encuentra en los brazos de la galaxia.

Según [Xu et al. \(2018\)](#) “...las estrellas OB y las asociaciones estelares jóvenes son trazadores principales de los brazos espirales, especialmente sus regiones HII asociadas, que son brillantes en longitudes de onda de radio y casi inmunes a la extinción del polvo interestelar, y se pueden detectar ampliamente en todo el plano galáctico...”. Esto se puede notar en los trabajos realizados por [Crampton and Georgelin \(1975\)](#), [Fich and Blitz \(1984\)](#), [Avedisova \(1985\)](#); los cuales realizaron mapas de regiones de HII repartidas por la galaxia y lograron trazar “The Perseus Arm”, “The Sagittarius-Carina Arm”, “The Scutum-Crux-Centaurus Arm” y “The Norma Arm”.

Por otro lado, el trabajo realizado por [Castro-Ginard et al. \(2021\)](#) cataloga cúmulos estelares en la galaxia y determina la edad de estos haciendo uso del diagrama color magnitud (el cual será explicado en la sección 1.4). Con estos cúmulos logran hacer un mapa de la galaxia como se muestra en la Figura 1.3, donde los segmentos de los brazos espirales en el vecindario solar (región sombreada modelada por [Reid et al. \(2014\)](#)) se ven claramente en las sobredensidades de OC para el intervalo de edad más joven explorado (0 – 30 Myr), además, muestran cómo estas sobredensidades sufren una dispersión creciente con el tiempo y por lo tanto una dilución lenta.

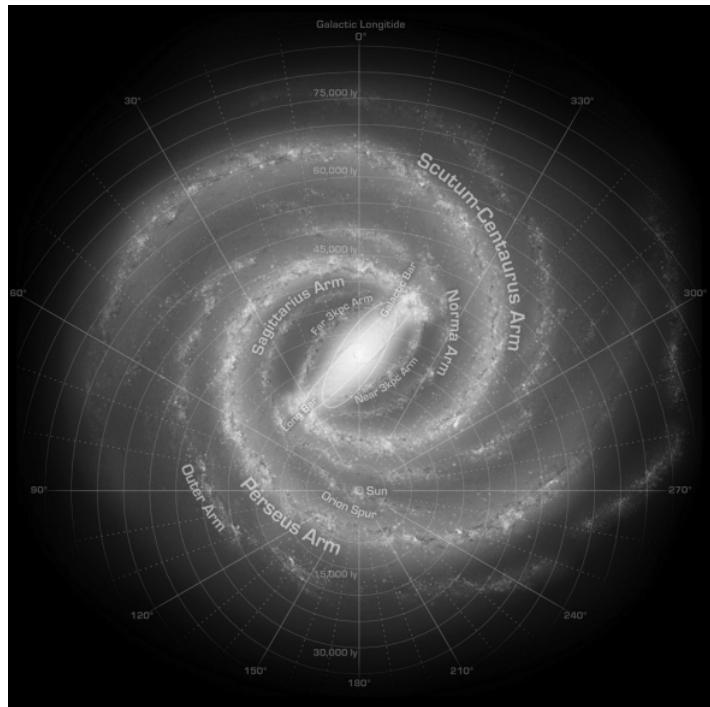


Figura 1.2: Diagrama de cómo es probable que aparezca la Galaxia vista de cara. Este diagrama fue realizado Robert Hurt y la imagen fue extraída de [Churchwell et al. \(2009\)](#) y se basa en datos obtenidos de la literatura en longitudes de onda de radio, infrarrojo y visible.

Algunos cúmulos abiertos conocidos de la Vía Láctea son las Pléyades, las Híades y el Pesebre los cuales son cúmulos jóvenes con edades de $\sim 110 - 160$ Myr, ~ 680 Myr y ~ 590 Myr respectivamente ([Gossage et al. \(2018\)](#)). Estos cúmulos son los más cercanos a la vía láctea y tanto las Pléyades como las Híades se encuentran sobre el brazo de Orión, el mismo brazo de la galaxia que se encuentra el sistema solar.

1.3.2. Cúmulos en el Halo y en el Bulbo

El bulbo de la galaxia está compuesto principalmente por estrellas viejas de población II, pero la metalicidad de estas estrellas no necesariamente es baja. Como se explica en [Hill et al. \(2011\)](#) el bulbo de la galaxia tiene una distribución bimodal en las metalicidades, donde se tienen estrellas tanto de baja metalicidad ($[Fe/H] < 0$) como de alta metalicidad ($[Fe/H] > 0$), así como se muestra en la Figura 1.4. Estos dos tipos de estrellas se distribuyen en el bulbo de la siguiente forma: las estrellas más ricas en metales están más concentradas hacia el plano mientras que las estrellas con menos metalicidad están en las zonas más lejanas del plano, incluso en los estudios de [Hill et al. \(2011\)](#) y [Rojas-Arriagada et al. \(2017\)](#) observaron que el bulbo bajo en metales está relacionado químicamente con el disco grueso.

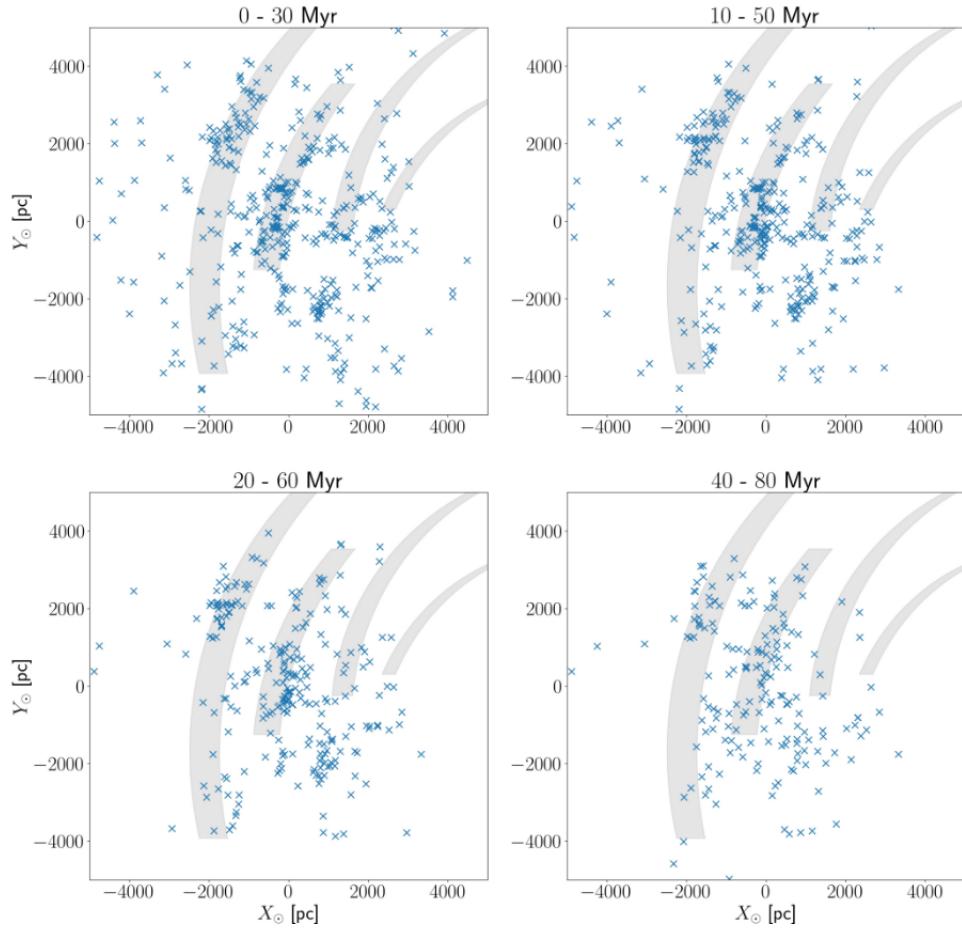


Figura 1.3: Gráfico tomado de Castro-Ginard et al. (2021) en el cual se plasma la distribución de cúmulos abiertos en el disco de la galaxia, los OC son las cruces azules con su edad. De izquierda a derecha y de arriba a abajo corresponde a: edades menores 30 Myr, de 10 a 50 Myr, de 20 a 60 y de 40 a 80 Myr. Las regiones sombreadas son los brazos de la galaxia modelados por Reid et al. (2014)

En el bulbo prevalecen principalmente cúmulos globulares (Harris (2010) identificó 157 GC pero el reciente estudio de Bajkova and Bobylev (2022) haciendo uso de GAIA EDR3 descartó 5 de estos cúmulos), incluso cerca a este es que se han observado los GC con mayor metalicidad, siendo NGC 6528 el primero en la lista, con una metalicidad de $[Fe/H] = -0.1$ (Muñoz et al. (2018)). Es interesante que siendo el cúmulo con mayor metalicidad pertenezca al componente de bulbo pobre en metales.

De manera similar, en el Halo de la galaxia habitan dos tipos de cúmulos globulares (hasta el momento según Carroll and Ostlie (2017) han encontrado alrededor de 150 GC en el Halo), los primeros tienen una metalicidad de $[Fe/H] < -0.8$ que son los más alejados del disco y se distribuyen de manera esférica por el halo; los segundos tienen una metalicidad de $[Fe/H] > -0.8$ y tienen estrellas más jóvenes, además tienen una distribución mucho más plana e incluso pueden estar asociados al disco grueso. Además, la edad de los cúmulos más jóvenes es alrededor de 11 Gyr, mientras que para los más viejos es un poco menos de 13 Gyr (Datos tomados de Krauss and Chaboyer (2003)).

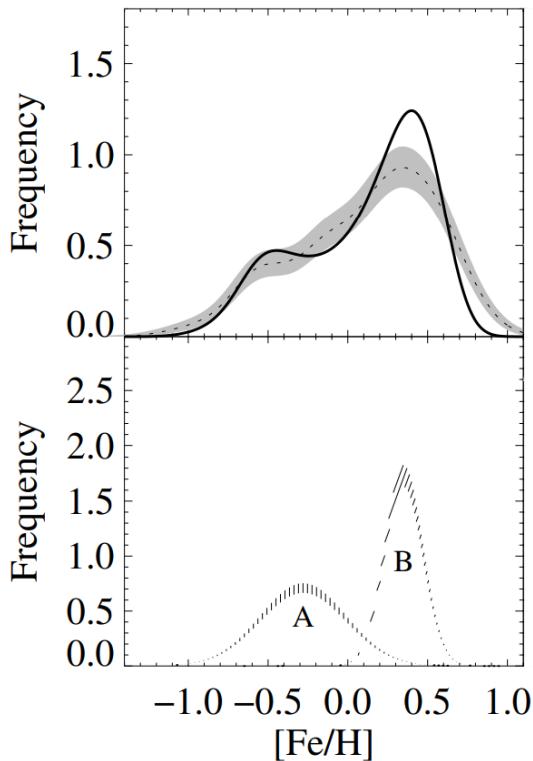


Figura 1.4: Gráfico tomado de Hill et al. (2011) donde se muestra la distribución de las metalicidades del bulbo de la galaxia, notando 2 picos: uno de ellos son estrellas con baja metalicidad (Grupo A) y otro es de estrellas con alta metalicidad (Grupo B).

Para este trabajo se hará uso principalmente de los cúmulos abiertos en el disco de la galaxia y no tanto de los cúmulos globulares en el Halo o en el Bulbo.

1.4 | Diagramas HR

Los Diagramas Color-Magnitud (CMD) o también conocidos como diagramas Hertzsprung - Russell (HR) son diagramas de los colores y la magnitud de las estrellas que permiten determinar diversos estados evolutivos de las estrellas o conjuntos de estrellas. El primer predecesor de estos diagramas se publicó en 1913 por Henry Norris-Russell el cual analizó la magnitud absoluta y el tipopectral de 200 estrellas y lo plasmó en un gráfico, notando una correlación entre medidas. De la misma forma Ejnar Hertzsprung, en la misma época, notó esta correlación.

Luego, se siguió estudiando la correlación entre estas medidas hasta obtener el conocido diagrama CMD, donde se estudia la relación entre la luminosidad y el color de las estrellas para así determinar características importantes de estas, como por ejemplo su temperatura, su masa, su edad y su metalicidad. Un ejemplo de estos gráficos se puede observar en la Figura 1.5.

En los diagramas HR, las estrellas se distribuyen según su etapa de evolución, su

masa y su color. Como resultado, estos diagramas presentan diversas regiones que han sido objeto de un extenso estudio (Chiosi et al. (1992), Langer and Kudritzki (2014), Larsson-Leander (1972)).

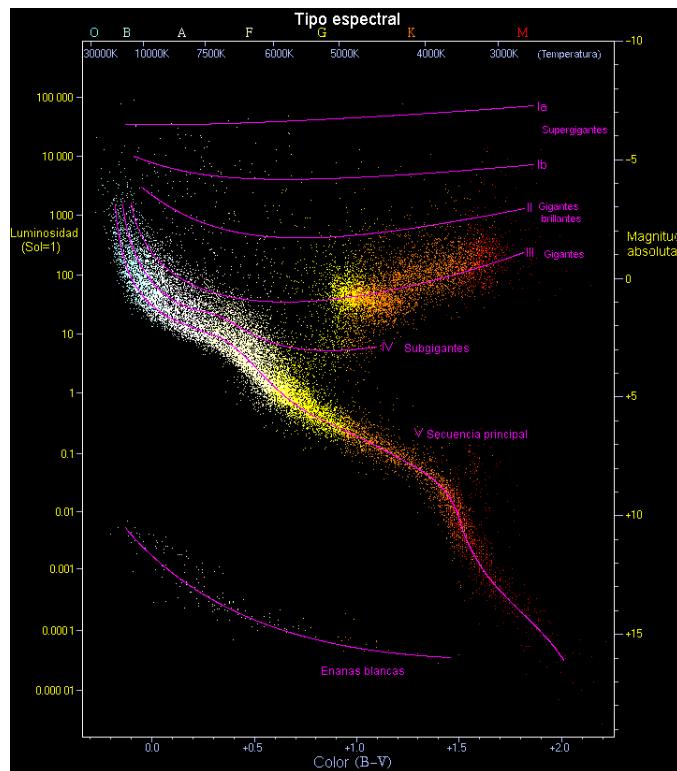


Figura 1.5: Diagrama HR con 22,000 estrellas del catálogo de Hipparcos y 1,000 del catálogo de Gliese. El Sol se encuentra en la secuencia principal exactamente a una luminosidad de 1 y un color B-V de 0.66.

1.4.1. Regiones de los diagramas HR

1.4.1.1. Secuencia Principal

Luego de que la estrella sale de fase de estrella de presecuencia principal (explicado en la sección 1.2.3.6) y la fusión nuclear del hidrógeno se convierte en el proceso de producción de energía dominante en la estrella y el exceso de energía obtenido de la contracción gravitacional se ha perdido, esta se encuentra en una de las curvas del diagrama denominada **Secuencia Principal**. La estrella permanece gran parte de su vida en esta curva hasta consumir una gran cantidad de hidrógeno en el núcleo.

En general, cuánto más masiva es una estrella, más corta es su vida útil en la secuencia principal. Ya que, una vez que la estrella ha agotado su combustible de hidrógeno en el núcleo, comienza a evolucionar fuera de la secuencia principal en el diagrama HR. Dependiendo de su masa, la estrella puede evolucionar hacia una supergigante, una gigante roja o puede pasar directamente a convertirse en una enana blanca (Adams and Laughlin (1997)).

1.4.1.2. Subgigantes (SGB)

Es una etapa de la evolución estelar únicamente para estrellas de masa baja a intermedia ($0.5 M_{\odot}$ a $8 M_{\odot}$), y ocurre cuando el núcleo empieza a colapsar por su propio peso ya que no hay suficiente energía brindada por la fusión nuclear que genera presión en la estrella. Esto lleva a que aumente la temperatura cerca al núcleo y el hidrógeno se comience a fusionar en una capa más externa, dando paso a que la estrella se expanda y se enfrié.

Cuando la estrella alcanza la temperatura alrededor de los 5000 K, la estrella pasa a la fase de **Gigante Roja**. Esta transición desde estrella de secuencia principal a gigante roja es conocida como la rama de las **Subgigantes**.

1.4.1.3. Rama de Gigantes Rojas (RGB)

La rama de gigantes rojas se produce cuando la estrella no posee la energía necesaria para fusionar helio en su núcleo. Esto provoca que las capas externas al núcleo comiencen a fusionar hidrógeno, generando más helio y, como consecuencia, aumentando la masa y temperatura del núcleo estelar.

Luego de cierto tiempo (alrededor de 1 Gyr para estrellas similares al Sol, 2.5 Gyr para estrellas de $0.6 M_{\odot}$ y de 25 Myr para estrellas de $2 M_{\odot}$, [Pols et al. \(1998\)](#)) el núcleo de la estrella obtiene la energía suficiente para comenzar la fusión llevando esto a que la luminosidad descienda pero la temperatura aumente, migrando entonces a la **Rama Horizontal**.

1.4.1.4. Rama Horizontal (HB)

En esta fase el núcleo de la estrella tiene la energía suficiente para quemar el Helio en su interior, dando paso a que la temperatura aumente y que la luminosidad disminuya. Este proceso sigue por millones de años hasta que se agote el núcleo de Helio y se tenga ahora un núcleo de elementos más pesados, lo que lleva entonces a terminar el proceso de fusión en este. Como resultado, cesa la fusión en las capas más internas, aunque en las capas externas aún se produce la quema de helio.

Cuando la estrella alcanza este punto, deja de pertenecer a la rama horizontal y pasa a estar en la **Rama Gigante Asintótica**.

1.4.1.5. Rama de Gigantes Asintóticas (AGB)

Una vez que la estrella ha completado la fusión del helio en su núcleo, esta comienza a enfriarse y a aumentar su luminosidad. Su trayectoria se asemeja en gran medida a la de la fase de gigante roja, por lo que se le conoce como la rama asintótica gigante. Sin embargo, durante esta etapa, la estrella se vuelve aún más luminosa que en la rama gigante roja.

Tras este proceso la estrella se aproxima a sus últimos años de vida.

1.4.1.6. Enanas Blancas

La estrella con un núcleo de metal e incapaz de generar la temperatura para la fusión de Carbono, procede a acumular una masa de Carbono y Oxígeno debido a la fusión de las capas externas de Helio. Dando paso a que la estrella se desprenda de sus capas más externas y deje atrás un núcleo expuesto de elementos pesados, este núcleo es más conocido como enana blanca ([Pols \(2011\)](#)).

El movimiento de la estrella por el diagrama HR a través de las distintas ramas mencionadas se observa en la Figura 1.6.

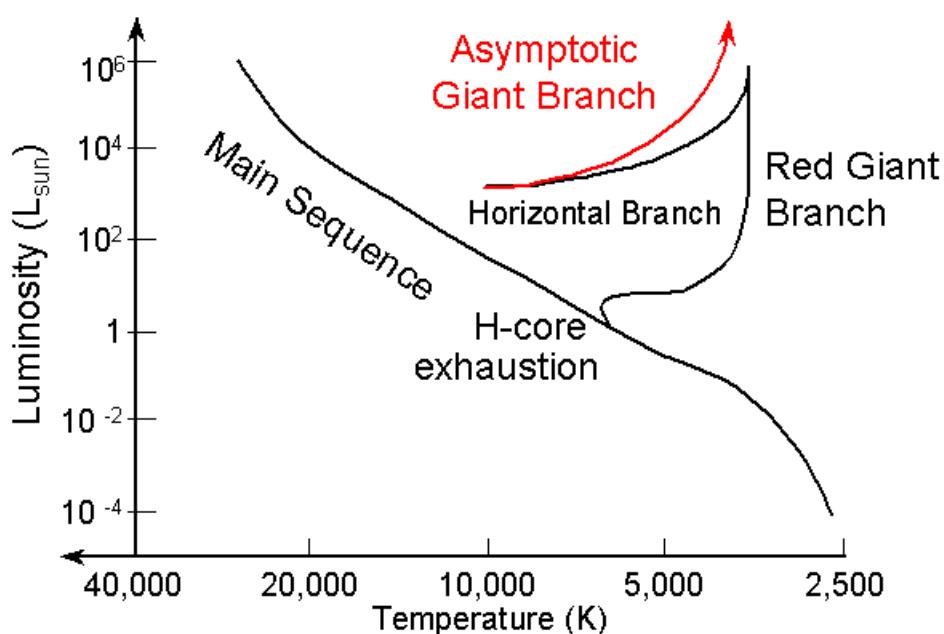


Figura 1.6: Se muestran las distintas ramas en el diagrama HR, donde el eje “*x*” representa la temperatura y el eje “*y*” representa la luminosidad. Imagen extraída de [Pogge \(2006\)](#).

1.4.2. Movimiento de una estrella a través del diagrama

En términos generales, la ubicación y trayectoria en el diagrama HR difieren considerablemente entre estrellas de baja masa y estrellas de alta masa. Esto se puede apreciar en la Figura 1.7, la cual ilustra la transición de las estrellas de diferentes masas desde una etapa previa a la secuencia principal hasta finalmente alcanzarla.

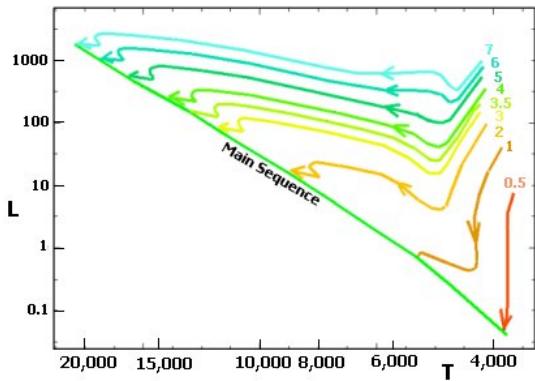


Figura 1.7: Diagrama HR de la transición entre estrella de presecuencia principal a estrella de secuencia principal. La diferencia en las curvas es debido a la masa de la estrella, por ejemplo, 7 se refiere a una masa de $7 M_{\odot}$. Imagen tomada de [Northen Iowa](#).

Como se habló en la sección 1.2.3.6 las estrellas en su última etapa de formación pasan a ser estrellas de presecuencia principal, estas se encuentran al lado derecho del diagrama color magnitud plasmado en la Figura 1.5, pasa el tiempo hasta que la estrella empieza a quemar hidrógeno en su núcleo y se convierte en una estrella de la secuencia principal. Luego de quemar gran parte del hidrógeno en su núcleo la estrella pasa a ser una subgigante y poco tiempo después una gigante roja, finalmente cuando la estrella ya terminó de quemar tanto el hidrógeno como el helio en su núcleo se aproxima la muerte de la estrella.

1.4.3. Importancia de los diagramas HR en los cúmulos estelares

Los diagramas HR en los cúmulos estelares sirven para determinar la edad de estos, ya que analizando los tipos de estrellas se puede determinar qué tan viejos son. Por ejemplo, la presencia de estrellas en la AGB es indicativa de una población estelar antigua; un número considerable de estrellas en la RGB está asociado con una actividad de formación estelar de edad intermedia a antigua.

Pero un método que permite determinar la edad del cúmulo es ajustar una isocrona a las estrellas pertenecientes a este. Una isocrona es, como lo dice Moyano (2019) “...es el conjunto de puntos correspondiente a modelos estelares de igual edad, pero distinta masa, en un diagrama HR”. En general las isócronas son curvas que se ajustan a las estrellas cuando estas pertenecen a la misma población estelar, un ejemplo de estas se puede observar en la Figura 1.8.

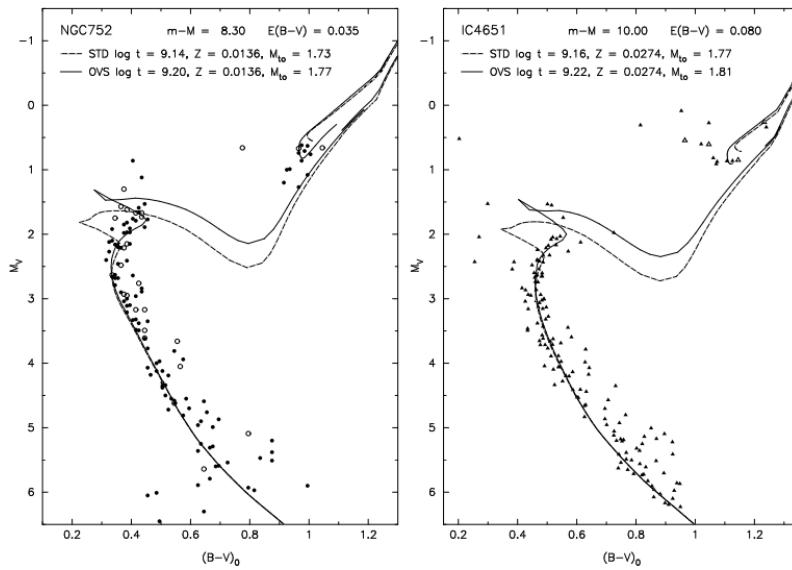


Figura 1.8: Imagen tomada de [Pols \(2011\)](#) donde se ilustra dos ejemplos de ajuste de isocronas para dos cúmulos abiertos, NGC752 (gráfico izquierdo) e IC4651 (gráfico derecho). El primer ajuste es usando los modelos estándar de evolución estelar (std) y el segundo es modelos con sobrepasso convectivo (ovs). Se puede observar que este último es capaz de reproducir de manera más precisa la extensión superior de la banda de la secuencia principal en ambos casos.

Estas curvas se han caracterizado a través de los años permitiendo conocer cuál es la edad del cúmulo y se han creado librerías y programas que permiten la obtención de las isocronas con un alto grado de confiabilidad; un par de estos programas son **AStECA** ([Perren et al. \(2015\)](#)) y **fitCMD** ([Charles Bonatto \(2019\)](#)).

Es importante decir que el hecho de que las estrellas pertenezcan a un cúmulo estelar, da paso a que hayan nacido prácticamente al mismo tiempo y lo que las diferencia en el diagrama HR es la masa de estas, así como se explicó en la Figura 1.7.

Capítulo 2

Objetivos

2.1 | Objetivo General

Estudiar la correlación de los brazos de la Vía Láctea con un catálogo personalizado de cúmulos estelares.

2.2 | Objetivos Específicos

- Evaluar la efectividad de los programas de clustering y de ajuste de isocronas aplicados, mediante la comparación con cúmulos previamente estudiados.
- Obtener una base de datos de cúmulos estelares que contenga información detallada acerca de sus parámetros astrofísicos: edad, metalicidad y distancia, utilizando los datos recopilados en el Gaia Data Release 3.
- Examinar la ubicación espacial de los cúmulos pertenecientes al catálogo y determinar su asociación o relación con los distintos brazos de nuestra Vía Láctea.
- Establecer la correlación entre la composición química predominante de los cúmulos estelares y su posición relativa con respecto a los brazos de nuestra Vía Láctea.

Capítulo 3

Metodología

3.1 | Compilación de Catálogos

Inicialmente, se recopilaron distintos catálogos provenientes de fuentes como Cantat-Gaudin et al. (2018), Cantat-Gaudin et al. (2020), Castro-Ginard et al. (2020), Castro-Ginard et al. (2021), Castro-Ginard et al. (2022) y Bossini et al. (2019). Estos 6 catálogos, de acceso abierto y disponibles en VizieR, se juntaron en una única base de datos con el propósito de establecer un catálogo completo. En este catálogo se eliminaron aquellos cúmulos repetidos. Este corte se hizo utilizando 2 requerimientos: el primero es si los centros de ambos cúmulos están presentes en un conesearch de 0.5 grados en las coordenadas galácticas y el segundo es si sus paralajes y movimientos propios son compatibles dentro de 2σ . Así, se obtuvo un catálogo total de 2499 cúmulos.

Luego de pasados unos meses de estar trabajando con este catálogo total, Hunt and Reffert (2023) publicaron su catálogo con 7167 cúmulos, incluyendo a los 2495 cúmulos de la unión de los otros siete catálogos e identificando 4668 nuevos cúmulos. Se optó entonces por trabajar con este catálogo ya que poseía una mayor cantidad de cúmulos y por lo tanto una cantidad más significativa en la muestra.

3.2 | Cúmulos con menos de 100 estrellas

Se llevó a cabo un estudio sobre los cúmulos de Hunt and Reffert (2023) (en adelante denominado “HR23”) que tienen entre 50 y 100 estrellas, así como aquellos con más de 100 estrellas. Este estudio fue propuesto por Gieles and Portegies Zwart (2011), quienes hacen una distinción entre cúmulos y asociaciones estelares a través del cálculo de la edad dinámica, definida como:

$$\Pi = \frac{\text{Age}}{T_{\text{cr}}} \quad \text{donde} \quad T_{\text{cr}} = \left(\frac{R_{\text{eff}}^3}{GM} \right)^{1/2}$$

“Age” se refiere a la edad del cúmulo y “ T_{cr} ” es el tiempo de cruce. “ R_{eff} ” es el radio donde se encuentra la mitad de la luminosidad, “M” la masa del cúmulo y “G” la constante de gravitación universal.

Lo que Gieles and Portegies Zwart (2011) plantean es que la frontera de $\Pi = 1$ permite distinguir entre una asociación estelar y un cúmulo estelar. Las asociaciones

estelares tienen un $\Pi < 1$, ya que el radio aumenta aproximadamente de manera proporcional a la edad y, por lo tanto, T_{cr} también aumenta. En cambio, para los cúmulos estelares, $\Pi > 1$, ya que se observa que, en primera instancia, R_{eff} y el tiempo de cruce permanecen aproximadamente constantes con el tiempo.

Se hizo entonces el cálculo de la edad dinámica para los cúmulos de HR23, tanto para los 2004 cúmulos con más de 100 estrellas como para los 1853 cúmulos que tienen entre 50 y 100 estrellas. En la Figura 3.1 se presenta un gráfico de la edad dinámica versus la edad del cúmulo, donde los puntos azules representan los cúmulos con más de 100 estrellas y los puntos naranjas representan los cúmulos con entre 50 y 100 estrellas. La línea punteada indica $\Pi = 1$.

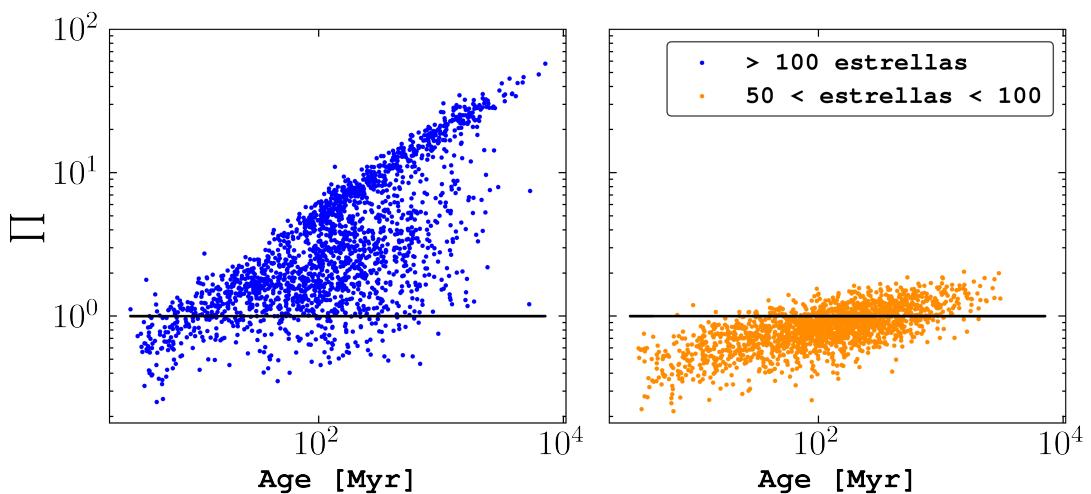


Figura 3.1: Edad dinámica en los cúmulos del catálogo de HR23 en relación con la edad de cada cúmulo. Los cúmulos se clasifican en dos grupos: los naranjas, que contienen menos de 100 estrellas, y los azules, que albergan más de 100 estrellas.

Para los cúmulos con más de 100 estrellas, se observa que los cúmulos que tienen una edad dinámica menor a 1 son generalmente jóvenes, y a medida que aumenta la edad, es menos probable encontrar un cúmulo con $\Pi < 1$. En total, el 17.42 % de cúmulos de HR23 con más de 100 estrellas tienen un $\Pi < 1$. En contraste, el 71.65 % de los cúmulos con menos de 100 estrellas (puntos naranjas en el gráfico) tienen un $\Pi < 1$, donde a medida que observamos cúmulos más jóvenes es más probable que estos estén por debajo de $\Pi = 1$ en comparación con los cúmulos viejos. Otro valor importante a tener en cuenta es que el 96.35 % de los 3310 cúmulos que tienen menos de 50 estrellas tienen un $\Pi < 1$.

Algunos ejemplos de cúmulos de HR23 con entre 50 y 100 estrellas con una edad superior a los 1000 Myr y con un $\Pi > 1$ son: Ruprecht 147 y Berkeley 75 los cuales son cúmulos que han perdido muchas de sus estrellas más externas con el paso del tiempo quedando actualmente el núcleo de un previo cúmulo estelar más grande (e.g. [Curtis et al. \(2013\)](#) y [Im et al. \(2023\)](#)).

Se decidió continuar trabajando exclusivamente con los cúmulos de HR23 que poseen más de 100 estrellas, principalmente porque un corte mayor eliminaría un gran porcentaje de cúmulos (el 40 % de los 2004 cúmulos tiene entre 100 y 150 estrellas). Además, estudios como el de [Payne-Gaposchkin \(1979\)](#) sugieren que los cúmulos abiertos contienen entre 10^2 y 10^4 estrellas. Tampoco se eliminaron aquellos cúmulos con más de 100 estrellas que tuvieran un $\Pi < 1$. Esto se debe a que el porcentaje de tales cúmulos es bajo y resulta interesante identificar cuáles de estos cúmulos son detectados en este trabajo.

3.3 | Importar estrellas de GAIA

Luego de conocer la posición de los 2004 cúmulos se realiza una búsqueda de objetos celestes dentro de un cono tridimensional, definido por la posición central de cada cúmulo y un radio angular de 1° . Se filtraron aquellas estrellas que tuvieran paralajes negativos para evitar datos con un error elevado.

También, se seleccionaron únicamente las estrellas con un $\text{ruwe} < 1.4$, garantizando la fiabilidad de los datos provenientes de Gaia (valor recomendado por [Lindegren et al. \(2021\)](#)). Este parámetro indica la fiabilidad de los valores astrométricos medidos por Gaia donde un valor menor a 1.4 indica que las observaciones se ajustan bien a las esperadas, lo que sugiere mediciones de alta calidad y confiabilidad. Por otro lado, valores significativamente mayores a 1.4 pueden indicar problemas en los datos, como la presencia de compañeros estelares cercanos que afectan las mediciones, errores sistemáticos en los datos o variaciones en el brillo de la estrella que no se tienen en cuenta.

Además, se restringió la selección a estrellas con una magnitud aparente menor a 20 con el objetivo de excluir aquellas que presentan una incertidumbre promedio en el paralaje de 0.6 mas. Esto se debe a que, al aumentar el límite de magnitud, por ejemplo a 21, el error en el paralaje se incrementa a 2 mas. Estos datos se basan en la Figura 3.2, obtenida de [Cantat-Gaudin et al. \(2018\)](#), donde se muestra la relación entre el error nominal del paralaje y la magnitud aparente de 10,000 estrellas. Se observa que para magnitudes superiores a 20, el error en el paralaje aumenta de manera más pronunciada que para magnitudes inferiores.

Este corte en la magnitud es importante para este trabajo, ya que la selección de estrellas que pertenecen al cúmulo se hará únicamente con los parámetros astrométricos. Sería interesante identificar miembros del cúmulo entre las estrellas con grandes incertidumbres astrométricas si se emplean otros criterios como selecciones fotométricas, aunque muchas de las fuentes sin una solución astrométrica completa también carecen de fotometría G_{BP} y G_{RP} que es necesaria en este trabajo para el ajuste de la isocrona que se explicará más adelante.

En términos generales, no se restringió excesivamente la búsqueda en GAIA, sino que se priorizó hacer una rigurosa selección de datos que no presentaran grandes errores en sus mediciones astrométricas. Esto permitió recopilar una gran cantidad

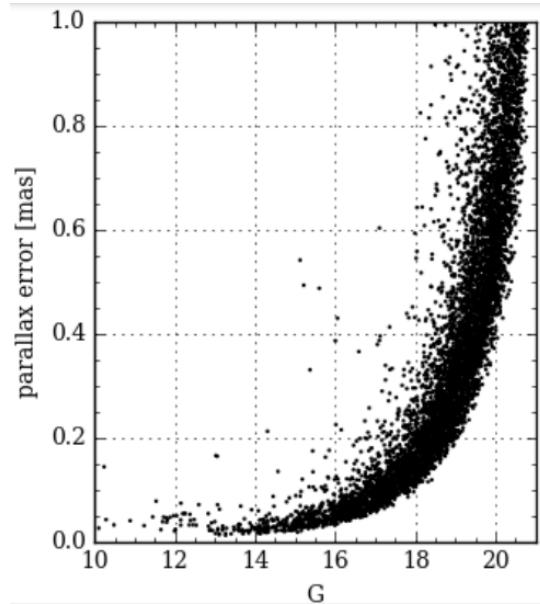


Figura 3.2: Error nominal en el paralaje contra la magnitud G para una muestra aleatoria de 10,000 estrellas de Gaia, gráfico tomado de [Cantat-Gaudin et al. \(2018\)](#).

de estrellas con información confiable para realizar un análisis detallado en la posterior búsqueda de cúmulos.

Luego de hacer la búsqueda en la base de datos de GAIA para los 2004 cúmulos se pasan los datos a DBSCAN.

3.4 | DBSCAN

DBSCAN es un algoritmo no supervisado de clustering que permite segmentar un conjunto de datos en 1 o múltiples cúmulos. Este algoritmo trabaja con 2 parámetros: ϵ que hace referencia al radio en el cuál hay una cantidad mínima de puntos, y **minPts**, que establece el mínimo de puntos requeridos para formar un grupo, ambos parámetros se ingresan manualmente. Para más información, leer los artículos de Ester et al. (1996) y Castro-Ginard et al. (2018).

Para comprender mejor el funcionamiento, es esencial entender cómo DBSCAN categoriza los datos en tres tipos diferentes: datos núcleo, datos alcanzables y datos ruido. Los datos núcleo son aquellos que, alrededor de ellos, forman una esfera de radio ϵ y contienen al menos una cantidad mínima de puntos (**minPts**). Si no se cumple esta condición, se consideran datos ruido. Por otro lado, si dentro de la esfera no hay la cantidad mínima de puntos pero al menos un punto es un dato núcleo, se clasifican como datos alcanzables. En la Figura 3.3 se puede ver un ejemplo de un sistema que tiene un **minPts**=2 y un radio ϵ donde los puntos verdes son los datos núcleo, el punto azul es el dato alcanzable ya que encierra a un dato núcleo y el punto rojo es el dato ruido.

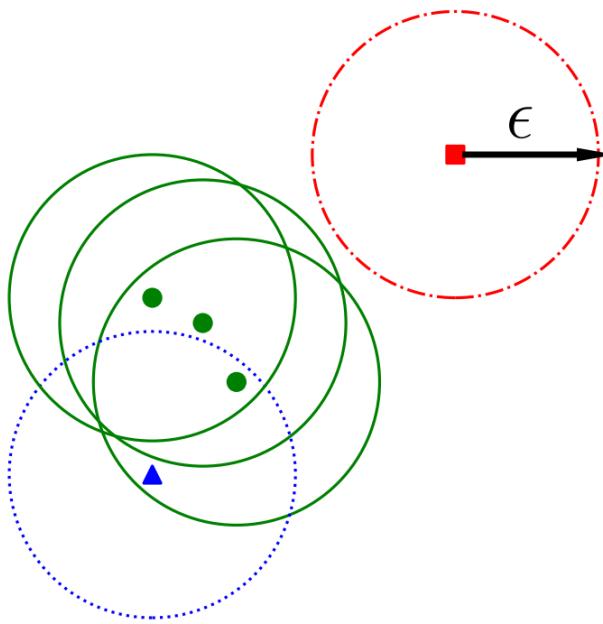


Figura 3.3: Un ejemplo de DBSCAN tomado de Castro-Ginard et al. (2018) para un sistema que se asignó un minPts=2 y un radio ϵ , los verdes son los datos núcleo, el azul el dato alcanzable y el rojo el dato ruido.

Teniendo esto en consideración, se procedió a proporcionar a DBSCAN cinco parámetros: l (latitud galáctica), b (longitud galáctica), ϖ (paralaje), μ_α (movimiento propio en ascension recta) y μ_δ (movimiento propio en declinación). Utilizando estos parámetros DBSCAN calculó la distancia de cada estrella a las otras en un hiperespacio de 5 dimensiones y determinó cuáles eran los puntos núcleo, alcanzables y ruido. La distancia entre dos estrellas (i y j) se calculó haciendo uso de la fórmula 3.1 (tomada de Castro-Ginard et al. (2018)).

$$d(i, j) = \sqrt{(l_i - l_j)^2 + (b_i - b_j)^2 + (\varpi_i - \varpi_j)^2 + (\mu_{\alpha,i} - \mu_{\alpha,j})^2 + (\mu_{\delta,i} - \mu_{\delta,j})^2} \quad (3.1)$$

Para este proyecto en particular, se emplearon los valores de minPTS y ϵ de la siguiente manera: se mantuvo un valor constante de minPTS = 8, basado en un estudio realizado por He et al. (2021b). Respecto al valor de ϵ , este se determinó evaluando múltiples valores dentro del rango de 0.03 a 0.4, que son un valor más pequeño y el otro más grande a los típicamente utilizados por Castro-Ginard et al. (2018) y He et al. (2021a). El valor tomado de ϵ variaba dependiendo del cúmulo, donde para regiones con muchas estrellas el valor era cercano a 0.03, mientras que para regiones con menos cantidad de estrellas el valor era cercano a 0.4 (nunca llegó a los valores extremos ya que se trabajó con un intervalo más amplio al normal). El objetivo era asegurar que en la región explorada solo se identificara un único cúmulo para evitar pequeños cúmulos que pudieran encontrarse en la misma región pero que no fueran el cúmulo identificado por HR23 con más de 100 estrellas. Además, de todos los valores de ϵ que identificaran un único cúmulo, se tomaba el valor más grande, asegurando así incluir a la gran mayoría de estrellas que posteriormente pasarían a otro filtro.

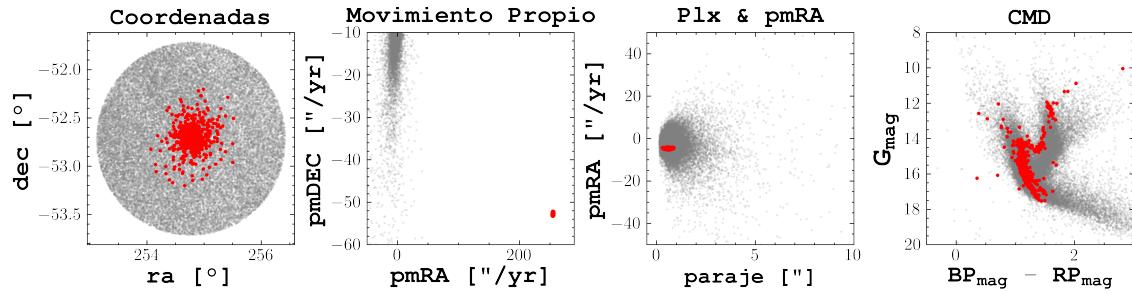


Figura 3.4: Gráfico del cúmulo NGC 6253 en comparación con las estrellas de fondo utilizando DBSCAN. *Primera imagen:* Ascension recta y declinación tanto de las estrellas del cúmulo (puntos rojos) como las estrellas de GAIA DR3. *Segunda imagen:* Movimientos propios de las estrellas. *Tercera imagen:* Paralaje y movimiento propio en la ascension recta. *Cuarta imagen:* CMD tanto del cúmulo como de las estrellas de fondo.

Luego de hacer la búsqueda en las 2004 regiones, DBSCAN logró identificar 1776 cúmulos utilizando el procedimiento anteriormente mencionado. Cúmulos que posteriormente se harán pasar por el código creado por Pera et al. (2021): pyUPMASK, para estudiar la probabilidad de membresía de las estrellas y descartar aquellas con una probabilidad baja.

3.4.1. DBSCAN comparado con cúmulos de la bibliografía

Para validar este procedimiento, se llevó a cabo una comparación del porcentaje de estrellas obtenidas mediante este método con las obtenidas por Cantat Gaudin (2018), (2020) y Castro Ginard (2018), (2020), (2021), (2022). Para ello, se recomilaron los catálogos publicados desde 2018 hasta la fecha actual, de los autores, y se realizó una depuración de cúmulos duplicados mediante una búsqueda en un radio de 1 arcmin en coordenadas ecuatoriales alrededor de cada cúmulo. En caso de encontrar varios cúmulos, se seleccionó aquel que presentara los datos más actualizados, resultando en un total de 2495 cúmulos para la combinación de los catálogos.

Posteriormente, se procedió a identificar los cúmulos presentes tanto en el catálogo final de los autores mencionados como en el catálogo obtenido en este trabajo. Para ello, se realizó otra búsqueda de 1 arcmin y se obtuvo un total de 710 cúmulos.

Finalmente, el porcentaje de estrellas totales obtenidas en relación con las obtenidas por Cantat Gaudin y Castro Ginard fue del 66.04 %.

NGC 6253

El cúmulo abierto NGC 6253, localizado en la constelación de Ara y en proximidad al centro galáctico ($l=335.46$ deg, $b=-6.25$ deg), contiene una considerable cantidad de estrellas de fondo, lo que complica la determinación precisa del cúmulo. No obstante, el empleo del algoritmo DBSCAN en un hiperespacio de 5 coordenadas mejora significativamente la fiabilidad en la identificación del cúmulo.

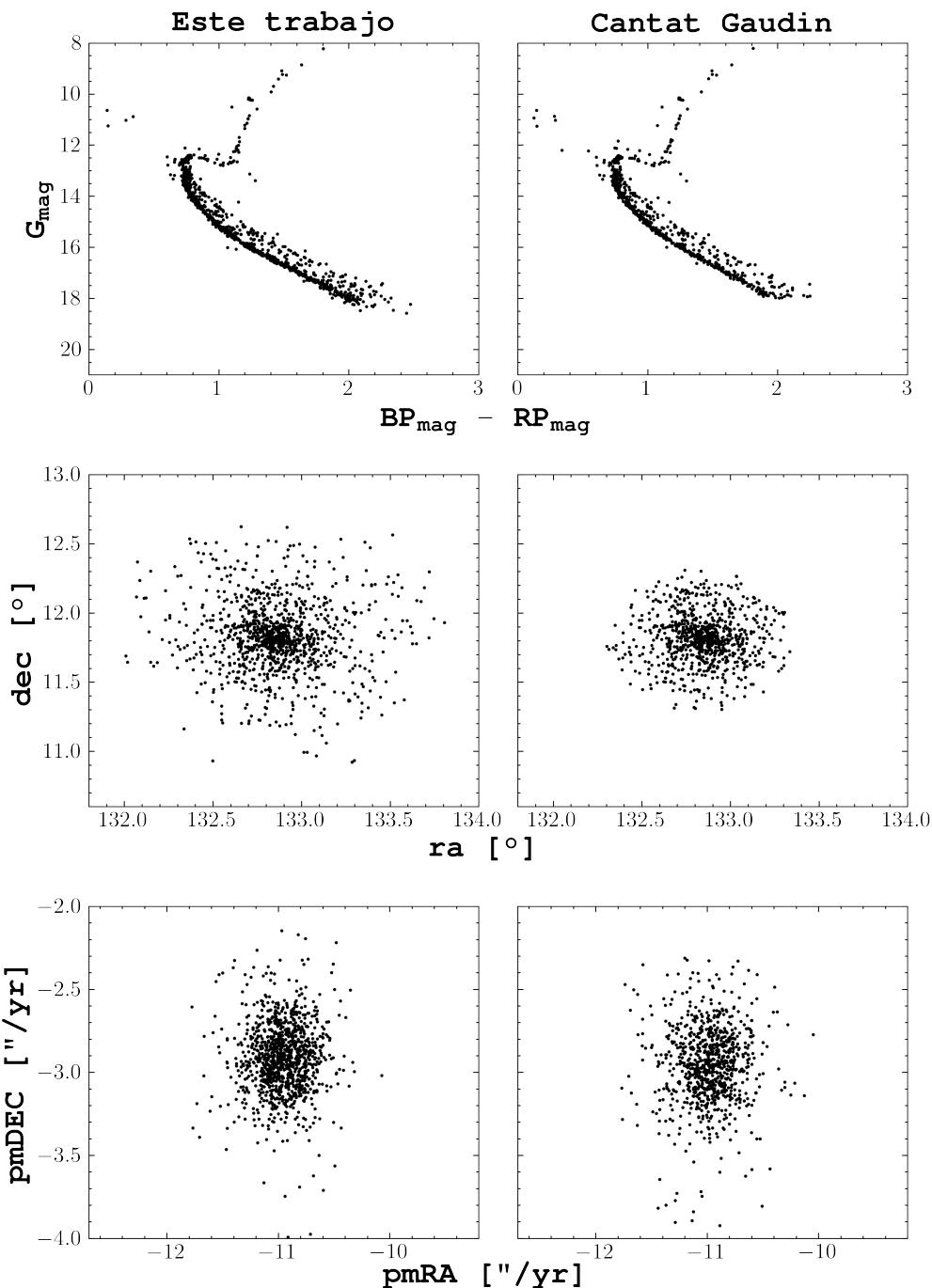


Figura 3.5: Gráficos de NGC 2682 donde se plasma al lado izquierdo las 1103 estrellas obtenidas por este trabajo y al lado derecho las 691 estrellas obtenidas por Cantat-Gaudin et al. (2018). El *gráfico superior* representa el CMD del cúmulo, el *gráfico del medio* es de las coordenadas ecuatoriales y el *gráfico inferior* el de los movimientos propios.

La Figura 3.4 muestra que los parámetros que influyen notablemente en la capacidad de DBSCAN para identificar el cúmulo NGC 6253 son los relacionados con el

movimiento propio (segundo gráfico). Se evidencia que las coordenadas ecuatoriales y el paralaje no presentan una diferenciación significativa en la identificación de este cúmulo.

NGC 2682

Se estudia el caso de NGC 2682, también conocido como M67, un cúmulo abierto ubicado en la constelación de Cáncer. La cantidad de estrellas pertenecientes a este cúmulo e identificadas por [Cantat-Gaudin et al. \(2018\)](#) fue de 691. Sin embargo, con el método aplicado en este trabajo, se identificaron un total de 1103 estrellas.

El gráfico correspondiente a este cúmulo se encuentra representado en la Figura 3.5, donde se realiza una comparación entre las posiciones, movimientos propios y el CMD presentes en ambos trabajos. Se observa una notable similitud entre los gráficos, aunque en las coordenadas ecuatoriales, el gráfico generado con los datos de este proyecto muestra algunas estrellas más distantes del centro en comparación con el trabajo de Cantat Gaudin. En consecuencia, se opta por la utilización de pyUPMASK para asignar a cada estrella una probabilidad de pertenencia al cúmulo, lo que permite eliminar aquellas que podrían ser ruido y no formar parte necesariamente del cúmulo. Este procedimiento se aplica a todos los cúmulos analizados.

3.5 | pyUPMASK

pyUPMASK es un código actualizado y mejorado de la versión original “UPMASK” que tiene como objetivo identificar y separar grupos de estrellas de las estrellas de campo utilizando el método de agrupación K-MEANS y aplicando un proceso de rechazo basado en la relación de los grupos con una distribución uniforme aleatoria.

Más específicamente, el algoritmo UPMASK se compone de dos bucles principales: bucle interno y bucle externo. El interno se encarga de determinar las estrellas que pertenecen al cúmulo y el externo analiza las incertidumbres generadas por el bucle interno asignando un valor de probabilidad de membresía a cada estrella.

El bucle interno contiene dos partes principales que conforman el núcleo del algoritmo: el método de agrupación (método K-means o KMS) y el método de rechazo (RFR). El método de agrupación divide el número total de estrellas del cúmulo en N grupos, el valor cambia a medida que se hacen las múltiples iteraciones en el bucle.

Luego de obtener los N grupos, el método RFR calcula la densidad de núcleo (KDE, en inglés “*Kernel Density Estimation*”) para cada uno de los grupos y compara el KDE de estos con el KDE de una distribución uniforme. Si no son similares, el grupo se descarta y a todas las estrellas se les asigna un valor de 0, que equivale a ser estrellas de campo. Los grupos que sobreviven al proceso de RFR se conservan para una iteración posterior del bucle interno y dependiendo de la cantidad de veces que la misma estrella aparezca en diversos grupos se le calcula una incertidumbre. Este proceso se repite hasta que no se rechazan más grupos y el bucle finalice, todas las

estrellas dentro de los grupos sobrevivientes se les asigna un valor de 1 y pasan al bucle externo.

En el bucle externo cada estrella se reemplazará por una distribución Gaussiana donde la media es el valor original y la desviación estándar es la incertidumbre calculada en el bucle interno. Luego se toma un valor cualquiera de la distribución Gaussiana y se ejecuta de nuevo el bucle interno para las estrellas con errores simulados.

Este proceso se repite la cantidad de veces que el usuario quiera (se utilizó el valor de 50 que es el recomendado por Krone-Martins and Moitinho (2014)) donde cada ejecución utiliza las estrellas con diferentes errores simulados. Una vez completadas todas las iteraciones, se recopilan los resultados y se calcula el valor de probabilidad de membresía, este es la fracción de ejecuciones en las que se le asignó la condición de miembro del cúmulo estelar.

Para este proyecto, a pyUPMASK se le enviaron los mismos 5 parámetros que a DBSCAN pero únicamente de las estrellas que este último catalogó como pertenecientes al cúmulo. Luego que pyUPMASK terminara de correr se obtuvieron las probabilidades de membresía de cada una de las estrellas de los 2004 cúmulos.

3.5.1. Probabilidad de Membresía

Para este estudio, se utilizó un valor de probabilidad de membresía del 0.5 con el propósito de eliminar aquellas estrellas más distantes del centro de la hipersfera, permitiendo obtener un conjunto más depurado de estrellas dentro del cúmulo. El valor de 0.5 se seleccionó basándose en la investigación llevada a cabo por Krone-Martins and Moitinho (2014), los fundadores de UPMASK. En su estudio, crearon regiones sintéticas de más de 10.000 estrellas donde no más de 500 estrellas eran las pertenecientes al cúmulo y el resto estrellas de campo. Luego, utilizando únicamente UPMASK determinan que si no se quiere perder información del cúmulo lo óptimo es usar el 0.5.

NGC 1039

NGC 1039, más conocido como M34, es un cúmulo abierto localizado en la constelación de Perseo y contiene 1095 estrellas según HR23. Mediante el uso del método DBSCAN, se logró identificar 630 estrellas distribuidas de la siguiente manera: 243 estrellas con una probabilidad de membresía mayor al 0.9, 308 estrellas con una probabilidad entre el 0.7 y el 0.9, 43 con probabilidad entre el 0.5 y el 0.7 y 36 con una probabilidad menor al 0.5 (si se desea observar la diferencia de valores entre los cúmulos del catálogo de HR23 con los de este trabajo puede ingresar a <https://github.com/slastrar/Thesis-Catalogue> el archivo de nombre “Comparison with Hunt.csv”).

La representación gráfica de esta distribución se muestra en la Figura 3.6. En el CMD (gráfico derecho), se evidencia que la mayoría de las estrellas identificadas por

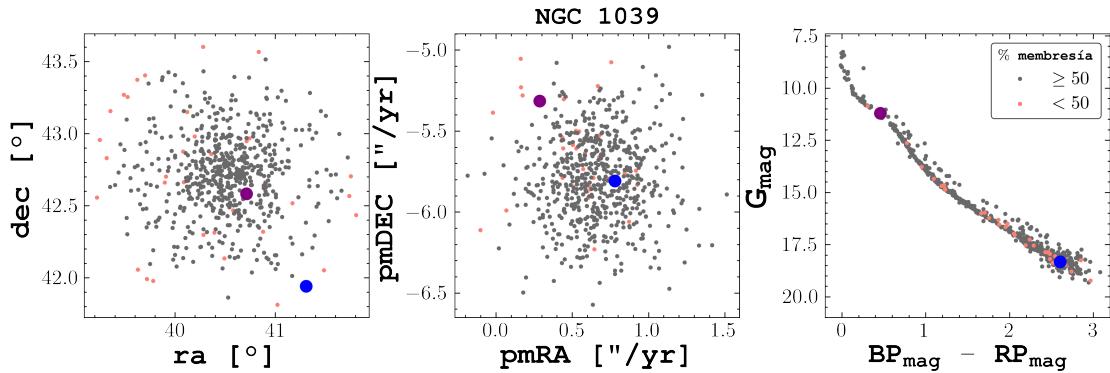


Figura 3.7: NGC 1039 donde los puntos gris son estrellas con probabilidad de membresía mayor al 0.5, mientras que los puntos naranjas, azules y morados indican estrellas con una PM % menor al 0.5.

DBSCAN tienen la misma edad. Sin embargo, se destaca que algunas estrellas tienen una probabilidad de membresía más alta que otras debido a su mayor proximidad a las estrellas vecinas (como se observa en los otros 2 gráficos).

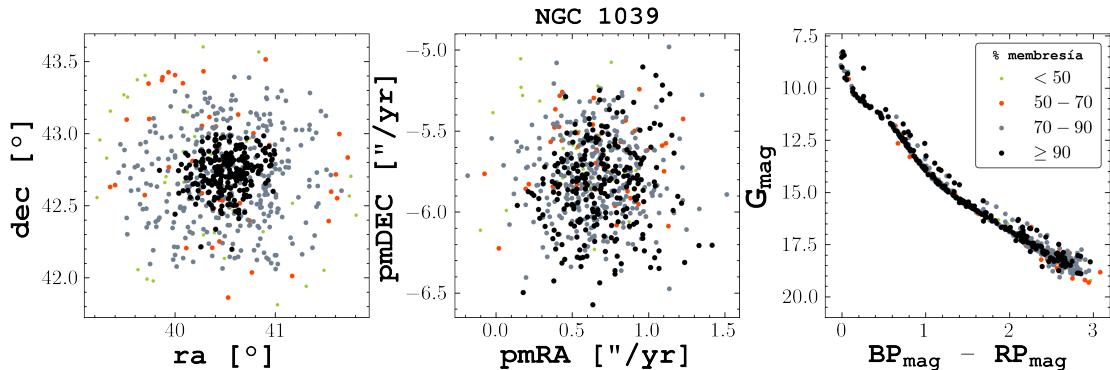


Figura 3.6: Cúmulo NGC 1039 donde el color negro representa las estrellas que tienen una probabilidad de membresía mayor al 0.9, el gris entre el 0.7 y el 0.9, el rojo entre 0.5 y 0.7, y el verde menor al 0.5. El *gráfico izquierdo* representa las coordenadas ecuatoriales, el *gráfico del centro* el de movimientos propios y el *gráfico derecho* el CMD.

En la Figura 3.7, se presentan únicamente las estrellas que tienen una probabilidad de membresía superior al 0.5 (puntos grises) y aquellas con una probabilidad inferior al 0.5 (puntos naranjas). Se han seleccionado específicamente dos estrellas con una probabilidad menor al 0.5 (representadas en morado y azul) para ilustrar cómo, aunque una de ellas pueda estar ubicada en el centro de la distribución en algunos valores, estará fuera de ese centro en otros. Este ejemplo ilustra el funcionamiento de pyUPMASK al mostrar cómo se comportan estas estrellas en diferentes puntos de la distribución.

Finalmente, con los cúmulos identificados y limpiados se guardan los datos y se procede a correr `fitCMD` para calcular la isócrona que mejor se ajusta a los datos.

3.6 | fitCMD

fitCMD es un código que, a través del análisis del CMD de un cúmulo estelar, facilita la estimación de parámetros astrofísicos como edad, distancia, masa y metalicidad, al generar un CMD artificial que mejor se ajuste al CMD original del cúmulo.

Para asegurar el correcto funcionamiento de **fitCMD** con todo tipo de cúmulo, es esencial definir un rango amplio de valores de los siguientes parámetros:

- M_{cl} (Masa del cúmulo): Se seleccionó un intervalo de $[1, 1 \times 10^6]$ M_{\odot} , ya que se explora el rango más grande permitido por el programa, valor tomado de [Marks and Kroupa \(2010\)](#).
- t_{age} (Edad del cúmulo): Se abarca tanto cúmulos de 1Myr de edad hasta cúmulos con edad cercana a la edad del universo.
- Z (Metalicidad): Metalicidad de hasta 100 veces menor que la del Sol (la metalicidad del Sol definida por fitCMD es de $Z = 0.0152$) hasta 2 veces mayor, ya que estos valores están dentro de los límites permitidos por el programa.
- XC (Exceso de color): Debe estar en el rango de 0 a 1.
- DM (Módulo de la distancia): Entre 10 pc y 16 kpc, tomando así un amplio rango en la distancia.

Una vez establecido este rango, el programa analiza las isócronas que se encuentren dentro de estos valores y genera diagramas CMD artificiales para cada una de ellas. Posteriormente, el programa divide el CMD original en celdas y calcula la densidad relativa de estrellas en cada celda, lo que permite la creación del diagrama de Hess correspondiente, denominado “ $H_M = H_M(M_{cl}, t_{age}, Z)$ ”; este mismo proceso se hace para los CMD artificiales.

Una vez obtenidos los diagramas de Hess tanto para el CMD original como para los CMD artificiales, se calcula el valor mínimo absoluto de $R_H = (H_{ori} - H_{art})^2$, donde R_H se denomina como la hiper-superficie residual, H_{ori} representa el diagrama de Hess del CMD original y H_{art} el de los CMD artificiales. El propósito de este cálculo es identificar cuál de los CMD artificiales (y, por lo tanto, la isócrona correspondiente) se asemeja más al CMD original, lo que permite determinar los parámetros astrofísicos asociados a este último.

Este programa se alimenta con un archivo por cada cúmulo, el cual contiene información de paralaje, coordenadas ecuatoriales, movimientos propios, magnitudes en las bandas G, G_{BP}, G_{RP} así como sus respectivos errores de cada una de las estrellas que forman parte del cúmulo. Dado que se estudian cúmulos de diferentes poblaciones, se considera necesario extender el rango de valores para los parámetros M_{cl} , t_{age} , Z , XC y DM al máximo posible. Esto asegura que se abarque un amplio espectro que engloba las 3978 isocronas disponibles en **fitCMD**, lo cual es fundamental para evitar posibles errores en el proceso de ajuste.

Finalmente, **fitCMD** devuelve un archivo con la edad, metalicidad, distancia, extinción, masa del cúmulo y módulo de la distancia. También un archivo PDF con la respectiva isocrona ajustada al diagrama Hess del CMD original.

3.6.1. Estudio fitCMD

Para validar los resultados obtenidos con fitCMD, se seleccionaron específicamente los miembros de los 2004 cúmulos del catálogo HR23 y se procesaron a través del programa. El objetivo de este análisis fue verificar la consistencia de las edades y distancias en comparación con los datos presentes en el catálogo HR23. Estos datos fueron revisados uno por uno, permitiendo la clasificación individual de cada cúmulo de la siguiente forma:

- GOOD: Un buen ajuste a un CMD de un cúmulo.
- MEDIUM: Un ajuste medio a un CMD no tan claro que sea un cúmulo.
- BAD: Un CMD de estrellas que no pertenece a un cúmulo.
- BADFIT: Un mal ajuste al CMD del cúmulo abierto.
- NODATA: FitCMD no reconoció el conjunto de estrellas como un cúmulo.

Se muestran en la Figura 3.8 los cúmulos BADFIT, MEDIUM, BAD y GOOD, y en la tabla 3.1 la cantidad por cada clasificación.

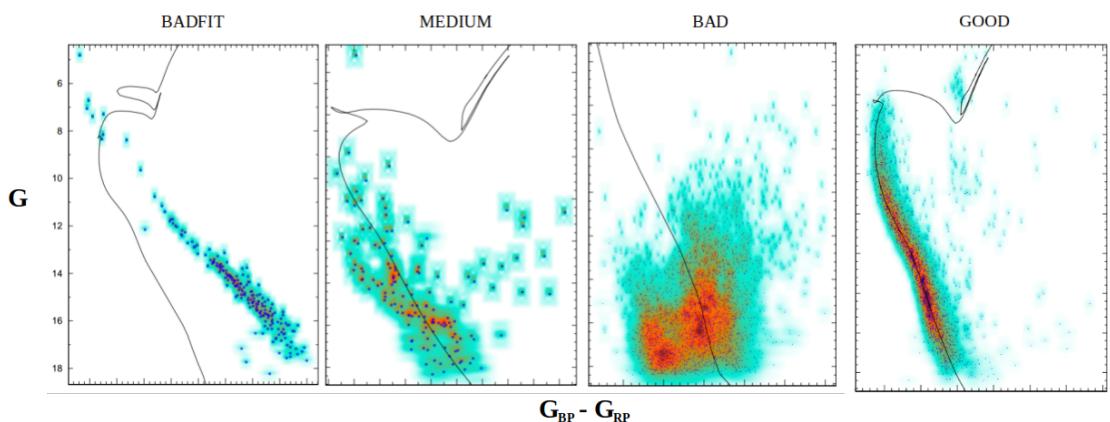


Figura 3.8: Representación gráfica de la clasificación de los cúmulos para los BADFIT, MEDIUM, BAD y GOOD.

Tipo	Cantidad	Porcentaje
GOOD	1051	59.2
BAD	162	9.1
MEDIUM	154	8.7
BADFIT	407	22.9
NODATA	2	0.1

Tabla 3.1: Cantidad de cúmulos de HR23 divididos por clasificación.

3.6.1.1. Comparación de cúmulos según la edad

En la Figura 3.9 se puede observar los cúmulos que fueron categorizados como “GOOD” o “MEDIUM” después de procesarlos a través de fitCMD. Los puntos negros representan cúmulos para los cuales la edad calculada con fitCMD difiere en más de 0.5 Gyr en comparación con la edad calculada por HR23 en su catálogo.

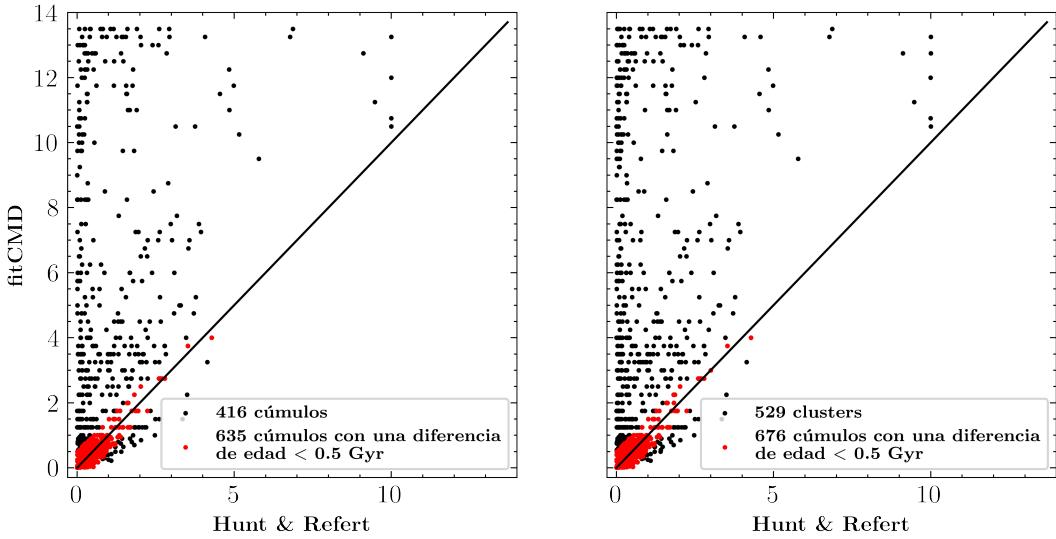


Figura 3.9: Las edades de los cúmulos se muestran a la izquierda para los cúmulos “GOOD” y a la derecha para los cúmulos “GOOD” y “MEDIUM”. En este gráfico, los puntos negros representan todos los cúmulos, mientras que los puntos rojos representan aquellos cuya diferencia de edad entre el catálogo de HR23 y el resultado de fitCMD es < 0.5 Gyr.

“GOOD”		
Dif. edad [Gyr]	Cant. cúmulos	Porcentaje
< 0.7	684	65.1
< 0.5	635	60.4
< 0.3	526	50.0
< 0.2	438	41.7
< 0.1	292	27.8
< 0.05	179	17.0

Tabla 3.2: Cantidad de cúmulos para distintas diferencias de edad entre el cúmulo de HR23 y el de este trabajo para los cúmulos clasificados como “GOOD”.

“GOOD” y “MEDIUM”		
Dif. edad [Gyr]	Cant. cúmulos	Porcentaje
< 0.7	731	60.7
< 0.5	676	56.1
< 0.3	558	46.3
< 0.2	463	38.4
< 0.1	306	25.4
< 0.05	188	15.6

Tabla 3.3: Cantidad de cúmulos para distintas diferencias de edad entre el cúmulo de HR23 y el de este trabajo para los cúmulos clasificados como “GOOD” y “MEDIUM”.

Las tablas 3.2 y 3.3 muestran la cantidad de cúmulos con una diferencia de edad entre `fitCMD` y el catálogo de HR23, agrupados por intervalos de tiempo: < 0.7 Gyr, < 0.5 Gyr, < 0.3 Gyr, < 0.2 Gyr, < 0.1 Gyr y < 0.05 Gyr. Se muestran tanto para los cúmulos clasificados como “GOOD” como para aquellos clasificados como “MEDIUM”.

3.6.1.2. Comparación cúmulos según la distancia

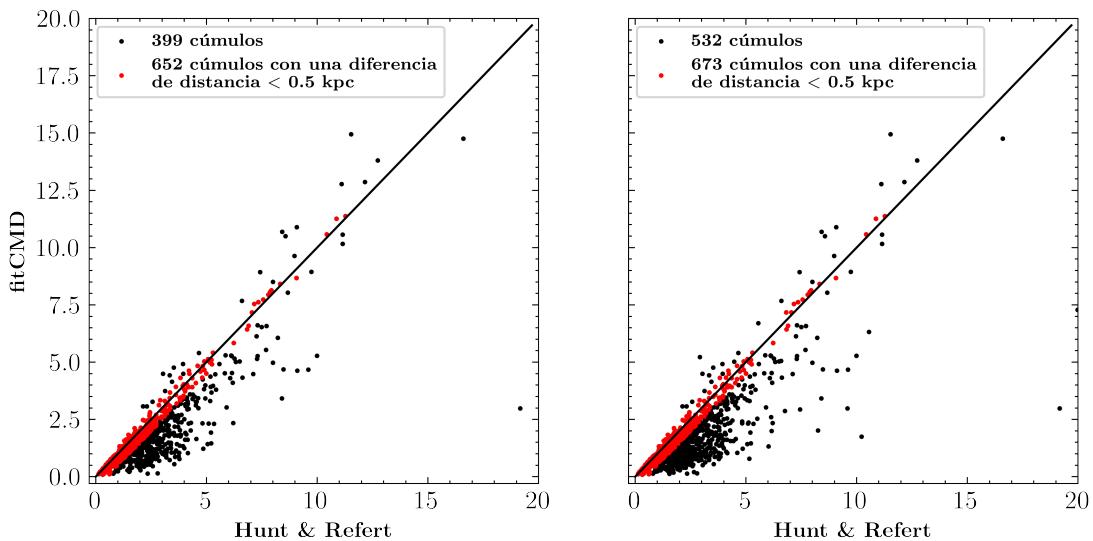


Figura 3.10: Las distancias de los cúmulos se muestran a la izquierda para los cúmulos “GOOD” y a la derecha para los cúmulos “GOOD” y “MEDIUM”. En este gráfico, los puntos negros representan todos los cúmulos, mientras que los puntos rojos representan aquellos cuya diferencia de distancia entre el catálogo de HR23 y el resultado de `fitCMD` es < 0.5 kpc.

En la Figura 3.10 y las Tablas 3.4 y 3.5, se lleva a cabo el mismo procedimiento que en la sección de 3.6.1.1, pero esta vez se realiza utilizando la distancia en kpc.

“GOOD”		
Dif. dist [kpc]	Cant. cúmulos	Porcentaje
< 0.7	745	70.1
< 0.5	652	62.0
< 0.3	539	51.3
< 0.2	451	42.9
< 0.1	336	32.0
< 0.05	229	21.8

Tabla 3.4: Cantidad de cúmulos para distintas diferencias de distancia entre el cúmulo de HR23 y el de este trabajo para los cúmulos clasificados como “GOOD”.

“GOOD” y “MEDIUM”		
Dif. dist [kpc]	Cant. cúmulos	Porcentaje
< 0.7	780	64.7
< 0.5	673	55.9
< 0.3	551	45.7
< 0.2	459	38.1
< 0.1	342	28.4
< 0.05	232	19.2

Tabla 3.5: Cantidad de cúmulos para distintas diferencias de distancia entre el cúmulo de HR23 y el de este trabajo para los cúmulos clasificados como “GOOD” y “MEDIUM”.

Con esto se puede decir que el programa `fitCMD` obtiene un ajuste más cercano al catálogo de HR2023 en el cálculo de la distancia, ya que el 64.7 % de los cúmulos GOOD o MEDIUM tienen una diferencia de distancia entre `fitCMD` y el catálogo menor a 0.7 kpc. Mientras que, para la edad, el porcentaje para la diferencia menor a 0.7 Gyr es del 60.5 %. Pero, en general, se puede observar en las figuras 3.9 y 3.10 que el gráfico de las edades de los cúmulos no se ajustan tanto a una distribución lineal en comparación con el gráfico de las distancias.

3.7 | Obtención parámetros astrométricos

Para determinar los valores de los parámetros astrométricos de los cúmulos (movimientos propios, paralaje, coordenadas ecuatoriales y galácticas), se tomó la moda de las estrellas del cúmulo usando el método del Kernel Density Estimation (KDE) de `scikit-learn`. Esta técnica estadística sirve para determinar la presencia de múltiples modas (picos) o la presencia de valores atípicos a través de la representación suave y continua de la distribución de datos, lo que facilita la comprensión de la forma y la estructura de los datos. Entonces, para evitar algunos problemas que tienen los histogramas, como por ejemplo, el proceso de agrupación (la cantidad de bins) y la anchura de un pico específico, se optó por utilizar el KDE para visualizar correctamente la distribución de los valores de los parámetros astrométricos.

La principal variable que afecta la forma final de un KDE es su ancho de banda, que para variables aproximadamente gaussianas, puede tomarse como $N^{-1/5}$ (Silverman (1986), Guidoum (2015)), donde N es el número de valores en la muestra, o en este caso, el número de estrellas que pertenecen al cúmulo. Después de calcular el KDE gaussiano con el ancho de banda mencionado, se determinó la moda de la distribución, es decir, el punto en el que la densidad de probabilidad tiene un máximo.

Se optó por utilizar la moda en lugar de la media debido a que la moda es menos influenciada por los valores extremos. Mientras que la moda se basa únicamente en la observación más frecuente o común en el conjunto de datos, la media considera toda la distribución, lo que la hace más susceptible a ser afectada por valores extremos. Esta diferencia se puede apreciar en la Figura 3.11, donde el histograma

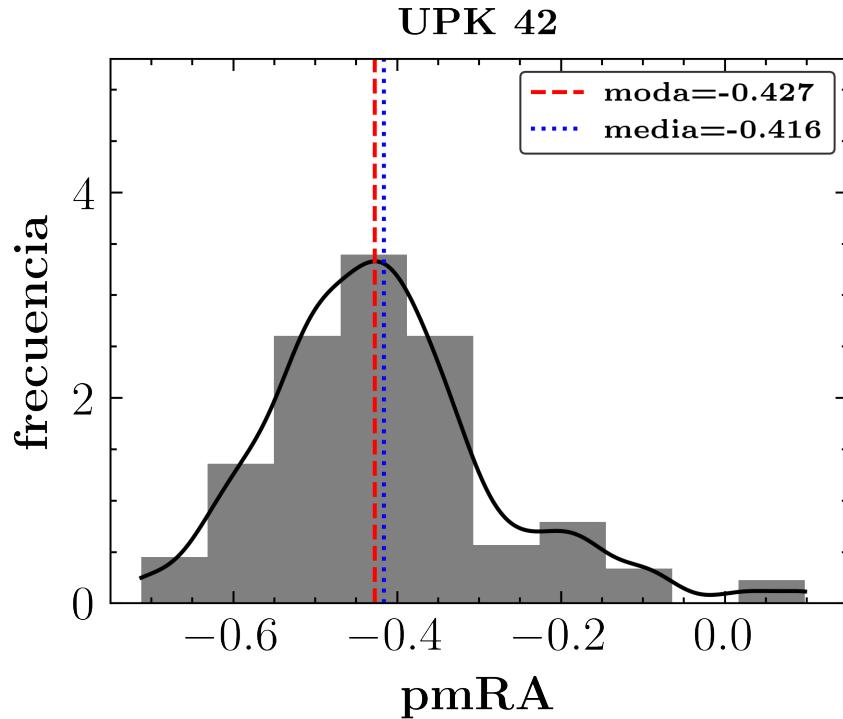


Figura 3.11: Histograma del movimiento propio en ascensión recta para el cúmulo UPK 42, la línea negra representa el KDE del histograma, la línea roja la moda de los datos y la azul la media.

normalizado que se muestra es el del movimiento propio en ascensión recta del cúmulo UPK 42, que contiene 109 estrellas identificadas en esta investigación; la línea negra representa el KDE del histograma. Se observa que el valor de la media está desplazado hacia la derecha respecto a la moda, debido a la influencia de los datos con valores positivos en el movimiento propio. En cambio, la moda refleja el valor más recurrente en los datos, ignorando estos valores extremos.

Se eligió este método en lugar de los utilizados en los trabajos de Cantat-Gaudin et al. (2018), Castro-Ginard et al. (2020) y Castro-Ginard et al. (2022), que recurren a la mediana o la media. Esta decisión se tomó debido a las observaciones realizadas en la sección de DBSCAN (sección 3.4.1), donde se notó que, muchas de las estrellas estaban cercanas entre sí en alguno de los parámetros, pero lejanas en otros, como se observa en la Figura 3.4. Esto implica que, al calcular la media del parámetro astrométrico del cúmulo que presente valores similares o más distantes del centro que los mostrados en la Figura 3.11, dicha media se verá sesgada por estos valores extremos.

Además, algo que se observó es que para cúmulos con aproximadamente más de 500 estrellas, la media era diferente de la moda después de la 5 cifra significativa. A medida que se consideraban cúmulos con una mayor cantidad de estrellas, esta diferencia tendía a disminuir, lo que sugiere una convergencia entre la media y la moda con el incremento del número de estrellas.

Error en la medida

Se determinó que el intervalo definido por el percentil 16 y el percentil 84 iban a hacer los errores de los parámetros astrométricos. Esta elección se basó en la distribución normal de los datos, donde estos dos percentiles incluyen aproximadamente al 68 % de los datos alrededor del valor central. Al seleccionar estos percentiles se asegura que el error represente una desviación estándar de la media en cada dirección, lo que es un estándar en estadística para describir la variabilidad en datos gaussianos.

Capítulo 4

Resultados

4.1 | Catálogo

Tras correr fitCMD para los 2004 cúmulos identificados por el método empleado en este trabajo se hizo el mismo procedimiento realizado en la sección 3.6.1 de observar el CMD de los cúmulos, obteniendo las siguientes cantidades en su clasificación:

Tipo	Cantidad	Porcentaje
GOOD	247	12.3
MEDIUM	323	16.2
BAD	1112	55.5
BADFIT	92	4.6
NODATA	230	11.4

Tabla 4.1: Cantidad de cúmulos de este trabajo divididos por clasificación.

Con estos datos se determinó que el catálogo de cúmulos identificados por el método acá empleado es de 570 cúmulos (los clasificados como GOOD y MEDIUM) que equivale al 28.5 % de los 2004 cúmulos identificados de HR23 para el estudio acá utilizado. En general combinaciones de diversos factores pueden dificultar la detección de un cúmulo, como la densidad de fuentes de fondo, la extinción interestelar, la densidad poblacional de un cúmulo o diferencias en sus movimientos propios respecto a las estrellas del campo.

Igualmente, este porcentaje de cúmulos aceptados no es bajo, ya que [Cantat-Gaudin et al. \(2018\)](#) empleó un método similar de tomar cúmulos identificados por la literatura y posicionar los datos de GAIA DR2 alrededor de estos cúmulos y, tras buscar cúmulos en estas áreas, finalmente aceptaron el 50 %.

Además, el método empleado por HR23 difiere del utilizado en este trabajo. En su proceso, ellos no emplean inicialmente un corte en la magnitud, sino que utilizan una red neuronal que clasifica una solución astrométrica como confiable o no confiable. Esta red neuronal fue presentada por [Rybicki et al. \(2022\)](#) y no está disponible como código abierto. Tampoco emplean DBSCAN para la obtención de los cúmulos, sino que usan HDBSCAN, una versión más nueva de DBSCAN que funciona con otros parámetros distintos a minPts y ϵ . El problema de este algoritmo es su tendencia a generar muchos cúmulos falsos positivos, ya que identifica fluctuaciones densas y

aleatorias en un conjunto de datos como cúmulos. HR23 emplearon este algoritmo e identificaron 7.1 millones de cúmulos en primera instancia, luego tras muchas observaciones que hicieron a los cúmulos obtuvieron su catálogo final de 7167 cúmulos.

Por lo tanto, se eligió aplicar DBSCAN ya que trabajar en zonas previamente identificadas en la literatura como lugares con presencia de cúmulos, minimiza la probabilidad de obtener falsos positivos. Además, la extensa y sólida utilización de DBSCAN en la literatura proporciona un entendimiento más sólido y extenso sobre su funcionamiento.

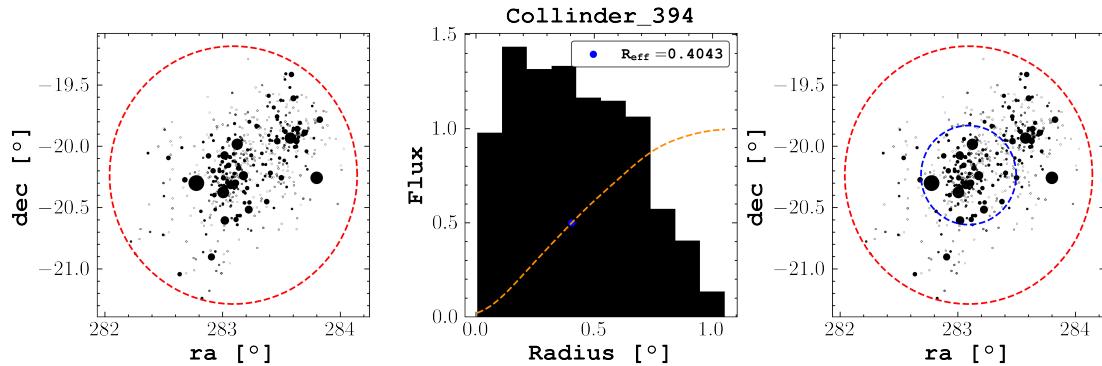


Figura 4.1: Se presenta el cúmulo “Collinder 394” como un ejemplo del cálculo del R_{eff} . En el primer gráfico se muestran las coordenadas ecuatoriales de las estrellas pertenecientes al cúmulo, donde el tamaño de cada punto representa el flujo de la estrella. En el segundo gráfico, se visualiza la suma del flujo de las estrellas en función de la distancia al centro del cúmulo donde la línea naranja representa la función de distribución acumulada (CDF), y el punto azul indica el valor de R_{eff} . El último gráfico es similar al primero, pero incluyendo también la representación de R_{eff} .

4.1.1. Edad dinámica del catálogo

Ahora, es necesario calcular la edad dinámica de los 570 cúmulos identificados para obtener el catálogo final de cúmulos. Para lograr esto, primero se calcula el radio donde se encuentra la mitad de la luminosidad (R_{eff}). Para ello, se toma el centro del cúmulo en ascensión recta y declinación, luego se calcula la distancia del centro a cada una de las estrellas, posteriormente el valor de la distancia máxima se divide en intervalos de igual tamaño (la cantidad de intervalos es del mismo tamaño que el tamaño del cúmulo) y se calcula la suma del flujo en la banda G para las estrellas que están en el mismo intervalo. Estos valores permiten entender cómo se distribuye el flujo de las estrellas a través del radio, valores que posteriormente serán estudiados por la función `rv.continuos` de la librería `scipy.stats` para obtener una distribución continua de los datos. A esta distribución se le calcula la función de distribución acumulada, o `cdf` (por sus siglas en inglés “cumulative distribution function”), la cual permite determinar dónde se encuentra el 50 % de la luminosidad del cúmulo.

Los pasos anteriormente mencionados se plasman en la Figura 4.1. En el gráfico izquierdo se plasma el cúmulo “Collinder 394” en coordenadas ecuatoriales junto con la distancia del centro a la estrella más externa (círculo rojo). El gráfico del centro muestra un gráfico de barras para la suma del flujo de las estrellas en función de la distancia al centro, junto con la curva de la cdf calculada por `rv.continuos`. El gráfico de la derecha es la representación del R_{eff} del cúmulo. Es importante destacar que el valor del R_{eff} calculado por HR23 para este cúmulo es 0.4017° mientras que el calculado por este trabajo es de 0.4043° .

Una vez calculado este parámetro, se estima el valor de T_{cr} , junto con el valor de Π correspondiente. Este proceso se ilustra en la Figura 4.2, donde se presenta la relación entre Π y la edad del cúmulo para el catálogo generado en este trabajo. Se identificaron 36 cúmulos con un valor de $\Pi < 1$, lo que equivale al 6.32 % de los 570 cúmulos iniciales. Por lo tanto, estos 36 cúmulos se eliminan del catálogo, obteniendo así un catálogo final de 534 cúmulos.

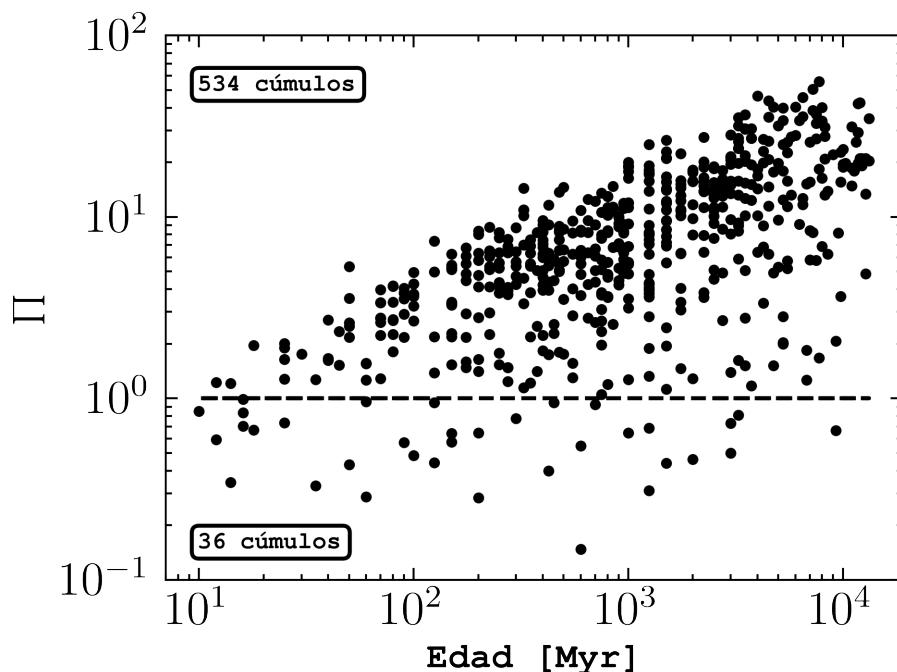


Figura 4.2: Se calcula la edad dinámica para los 570 cúmulos y se obtiene que 36 están por debajo del corte de $\Pi = 1$

Finalmente, en la Figura 4.3 se plasma el CMD de 12 cúmulos escogidos aleatoriamente del catálogo con su respectiva edad y ajuste de la isocrona.

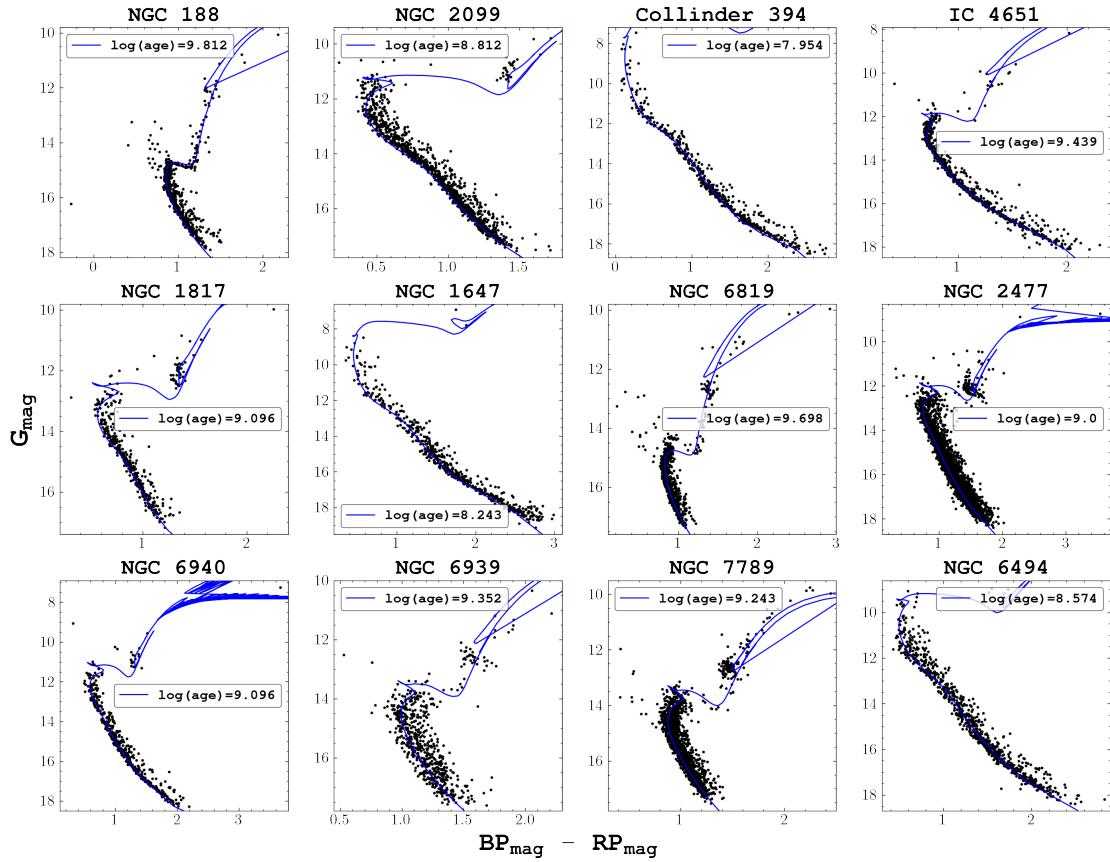


Figura 4.3: Se presenta el CMD de 12 cúmulos escogidos aleatoriamente del catálogo con la edad y la isocrona obtenidas por fitCMD.

4.1.2. Comparación de valores con HR23

Valores astrométricos

Para comparar los valores astrométricos de los cúmulos entre el catálogo de HR23 y el catálogo obtenido en este trabajo, se tomó principalmente las posiciones X, Y y Z de los cúmulos para el plano galáctico ya que así se pueden comparar las coordenadas l , b y la distancia al cúmulo. También se compararon los movimientos propios de los cúmulos.

En primera instancia, se estudió primero la relación en las coordenadas X y Y que son las que más varían y se plasma en la Figura 4.4 donde se halla que el 66.10 % de los datos tienen una diferencia en la distancia menor a 350 pc para la coordenada X y el 57.30 % de los datos para la coordenada Y. Para la diferencia de distancia menor a 200 pc se halló que el 48.13 % es para la coordenada X y el 46.63 % para la coordenada Y. Finalmente, para la diferencia de distancia menor a 100 pc se halló que el 34.64 % es para la coordenada X y 32.58 % para la coordenada Y.

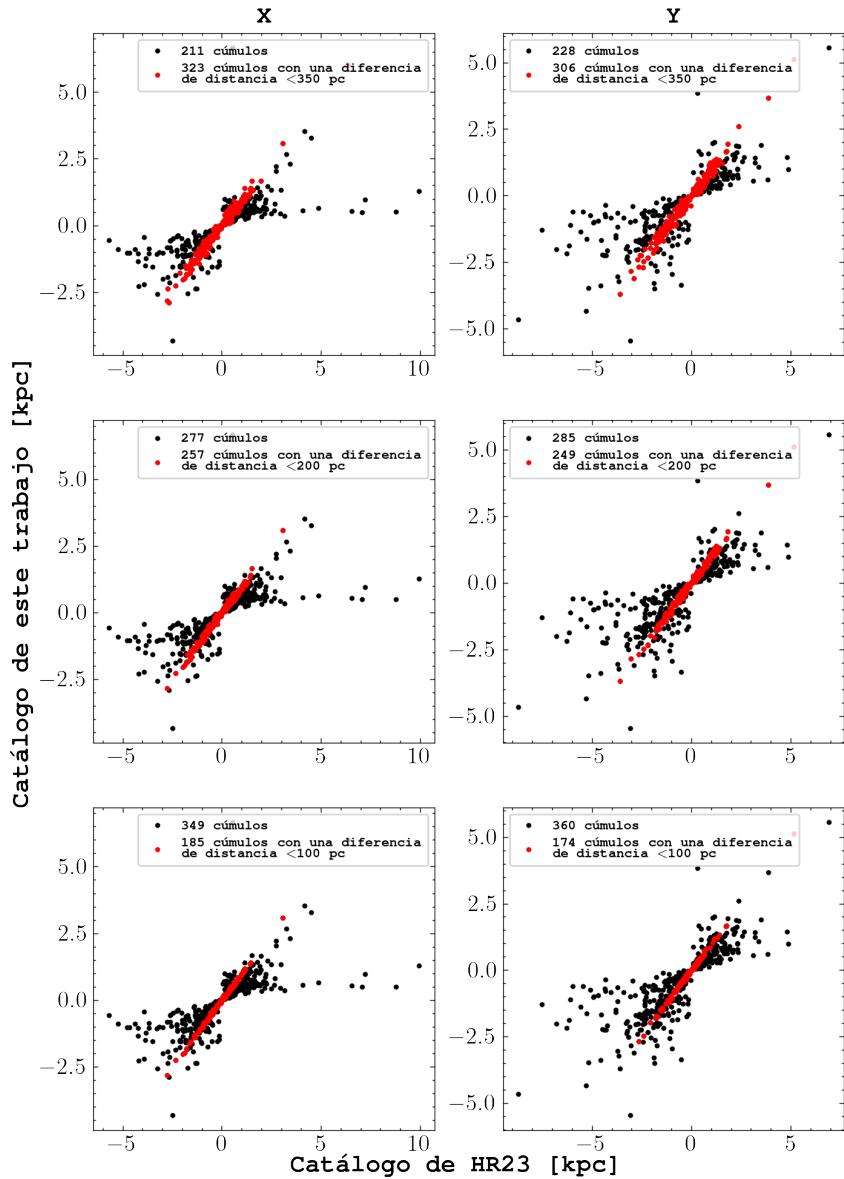


Figura 4.4: Diagrama que muestra la diferencia de distancias entre el catálogo de HR23 y el generado en esta investigación para las coordenadas X e Y en el plano galáctico. El *gráfico superior* representa los cúmulos con una discrepancia de distancia en las coordenadas menor a 350 pc. El *gráfico intermedio* para una discrepancia de 200 pc. Y el *gráfico inferior* para una discrepancia de 100 pc.

Por otro lado, para la coordenada Z se utilizaron los valores de diferencia de distancia de 350, 100, 50 y 10 pc; y se obtuvieron porcentajes del 97.19 %, 88.39 %, 71.35 % y 26.59 %, respectivamente. Se observa que 15 cúmulos tienen una diferencia en la coordenada Z mayor a la distancia del disco delgado (0.30 ± 0.05 kpc, tomado de McMillan (2011)). Algunos de estos cúmulos son NGC 362, NGC 6341 y NGC 6254, cúmulos globulares localizados fuera del disco de la galaxia.

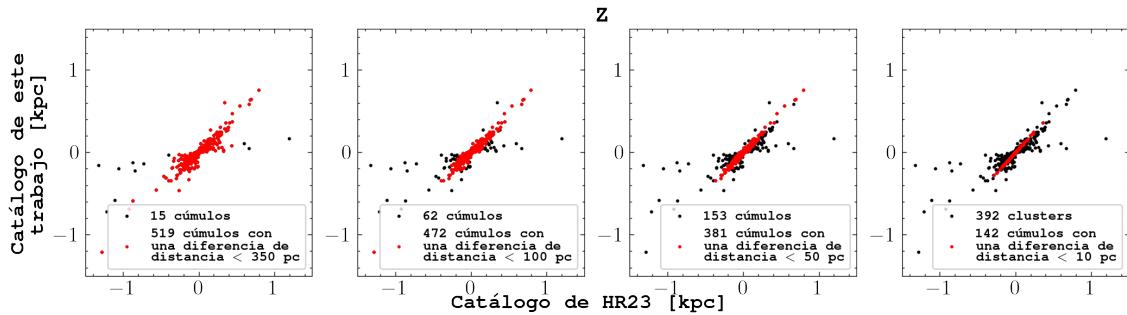


Figura 4.5: Diagrama que muestra la diferencia de distancias entre el catálogo de HR23 y el generado en esta investigación para las coordenadas Z del plano galáctico. En orden de izquierda a derecha se plasman los cúmulos con una discrepancia de distancia de 350, 100, 50, 10 pc.

Algo que se analizó, es que estos cúmulos que tienen una diferencia en distancia mayor a 300 pc son principalmente cúmulos globulares, específicamente 13 de los 15 cúmulos son cúmulos globulares que están a una distancia mayor a la del disco grueso (0.9 ± 0.15 kpc, tomado de [McMillan \(2011\)](#)). Los restantes dos cúmulos se sitúan dentro del disco grueso, pero no se obtuvo una medida precisa de su distancia mediante fitCMD.

Posteriormente se estudió también la diferencia de los movimientos propios de la misma forma que la de las coordenadas y los valores obtenidos fueron: el 74.16 % para una diferencia de 0.5 arcseg/año, el 65.92 % para una diferencia de 0.1 arcseg/año, el 59.74 % para una diferencia de 0.05 arcseg/año y el 26.40 % para 0.01 arcseg/año.

Edad, extinción y metalicidad

De la misma forma se calculó la diferencia para la edad de los cúmulos y se obtuvo que: el 56.90 % tienen una diferencia menor a 1 Gyr, el 40.29 % una diferencia menor a 500 Myr, el 12.49 % una diferencia menor a 100 Myr y el 6.16 % (que equivale a 33 cúmulos) una diferencia menor a 50 Myr.

Mientras que para la extinción total en la banda visible ($A(V)$) la diferencia entre los valores tienen porcentajes de: el 66.98 % para una diferencia menor a 1, el 39.74 % una diferencia menor a 0.5, el 7.86 % para una diferencia menor a 0.1 y el 4.11 % para una diferencia menor a 0.05.

Por otro lado, el catálogo de HR23 no ofrece información acerca de la metalicidad del cúmulo, mientras que el catálogo de este trabajo sí posee esta información. Esto permite la observación y análisis de cómo se distribuyen los cúmulos con metalicidades altas en el plano de la galaxia. Así, se puede comparar este trabajo con el de [Poggio et al. \(2022\)](#), quienes estudiaron estrellas gigantes brillantes en el disco de la galaxia, identificando regiones con una mayor presencia de estas estrellas de alta metalicidad y trazando los brazos de la galaxia.

4.1.3. Cúmulos detectados en relación con otros catálogos

Se realizó un crossmatch entre el catálogo de este estudio y catálogos reconocidos de cúmulos, tales como Cantat-Gaudin (se unen los catálogos del (2018) y del (2020)), Zhihong He et al. (2023), Castro-Ginard (se unen los catálogos del (2018), (2020), (2021) y (2022)) y Bossini et al. (2019) , que contienen 1779, 1656, 1102 y 263 cúmulos, respectivamente. Para confirmar la presencia de un cúmulo en ambos catálogos, se utilizaron dos requerimientos: el primero es si los centros de ambos están presentes en un conesearch de 0.5 grados en las coordenadas galácticas y el segundo es si sus paralajes y movimientos propios son compatibles dentro de 2σ , donde σ es la suma cuadrática de las incertidumbres citadas en ambos catálogos para cada cantidad.

Se obtuvo que 521 cúmulos de los 1779 (25.97 %) de Cantat-Gaudin están presentes en el catálogo acá obtenido, así como 206 de los 1102 (17.97 %) cúmulos de Castro-Ginard, 186 de los 1656 (10.99 %) de Zhihong y 162 de los 263 (61.60 %) de Bossini. Se destaca que, en términos porcentuales, el catálogo con el menor número de cúmulos en común es el de Zhihong. Esto se debe, en gran parte, a que ellos hicieron la búsqueda de cúmulos con HDBSCAN a diferencia de los otros que utilizaron DBSCAN directamente.

4.1.4. Coeficiente de Spearman

El coeficiente de correlación de Spearman, desarrollado por el psicólogo británico Charles Spearman (1904), es una medida estadística que sirve para evaluar la correlación entre dos variables que están clasificadas ordinalmente. Este coeficiente no implica causalidad entre las variables; más bien, cuantifica la fuerza y dirección de la asociación entre las variables, independientemente de la escala de medida utilizada. Es particularmente útil cuando se trata con datos que no se distribuyen de manera normal.

El rango de valores de Spearman está entre -1 a 1, indicando los valores extremos correlaciones exactas, donde -1 indica una correlación perfecta decreciente y 1 una correlación perfecta creciente; 0 por el contrario indica que no hay correlación entre las variables. Según Schober et al. (2018) las correlaciones se clasifican de la siguiente forma:

- 0.00–0.10: Correlación nula.
- 0.10–0.39: Correlación débil.
- 0.40–0.69: Correlación moderada.
- 0.70–0.89: Correlación fuerte.
- 0.90–1.00: Correlación muy fuerte.

Hay otro factor que también está asociado con el coeficiente de correlación de Spearman, el p-value. Este valor ayuda a determinar si esa correlación es estadísticamente significativa, es decir, si es probable que no se deba al azar. Para valores del p-value mayor a 0.05 indica que no hay suficiente certeza para afirmar correlación entre los valores, mientras que menor a 0.05 indica que las dos variables están significativamente correlacionadas.

Para este trabajo en particular el coeficiente de Spearman permite calcular cuál es la relación entre los valores astrométricos y astrofísicos ya que estos no necesariamente se distribuyen de manera normal. Por ejemplo, la edad de los cúmulos se plasma en la Figura 4.6, donde se aprecia que hay en mayor medida cúmulos jóvenes tanto para el catálogo de HR23 (histograma rojo) como para el catálogo de este trabajo (histograma gris).

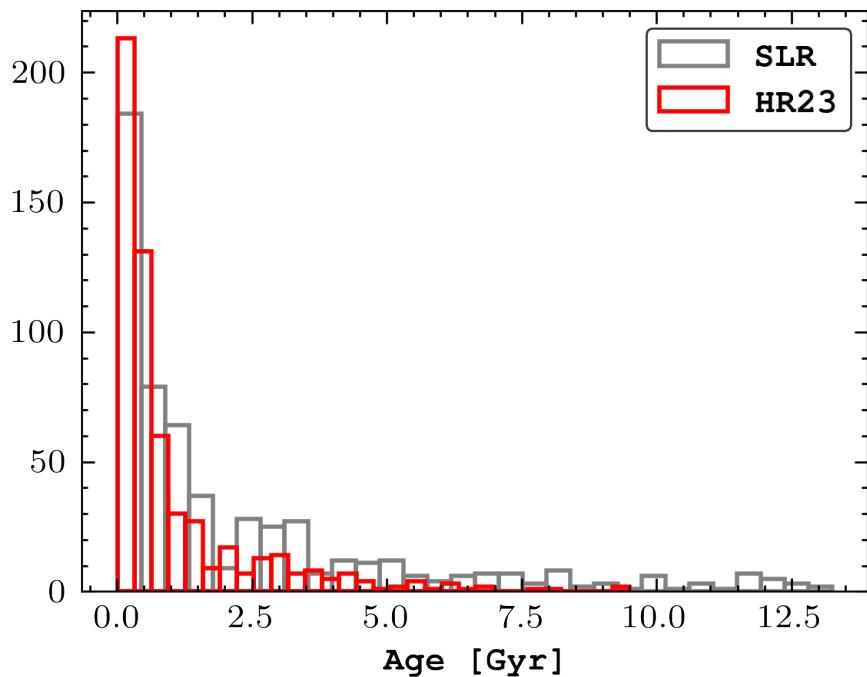


Figura 4.6: Histogramas de la edad de los cúmulos, el histograma rojo es del catálogo de HR23 y el negro es el catálogo de este trabajo.

HR23 con el catálogo de este trabajo

Calculando el coeficiente de Spearman para la distancia de los cúmulos entre el catálogo de HR23 y el de este trabajo se obtuvo un valor de 0.5572 (correlación moderada), de la misma forma el coeficiente para los movimientos propios entre el valor del catálogo de HR23 y el de este trabajo es de 0.8207 (correlación fuerte). Mientras que el valor para la edad es de 0.1572 y para la extinción de 0.2314, ambos correlación débil.

Por otro lado, el p-value del coeficiente para la distancia y el movimiento propio es de 8.62×10^{-48} y de 3.39×10^{-140} , respectivamente. Mientras que el de la edad es de 0.25 y el de la extinción de 0.17.

Como se puede evidenciar, el catálogo de HR23 y el de este trabajo está más correlacionado en los valores astrométricos (correlación media) que en los astrofísicos (correlación débil).

Otros catálogos con el catálogo de este trabajo

A diferencia de HR23, si se calcula el coeficiente de Spearman para la edad y la extinción de los cúmulos que están tanto en el catálogo de este estudio como los catálogos de Cantat-Gaudin y de Bossini (no se utilizaron los otros dos catálogos debido a que poco más de 40 de los 206 cúmulos de Castro-Ginard tienen valores en la edad. Y el catálogo de Zhihong no posee información de la extinción) se muestran los valores en la Tabla 4.2.

Catálogo	distancia	m. propio	edad	extinción
C-Gaudin	0.7649	0.9351	0.4426	0.5280
Bossini	0.8271	0.8945	0.5612	0.6108

Tabla 4.2: Coeficiente de Spearman entre el catálogo de HR23 y el de este trabajo para 4 parámetros distintos.

El coeficiente de Spearman en estos catálogos experimentó un notable aumento, especialmente en relación con la edad y la extinción, pasando de una correlación débil a una moderada. Además, al calcular el p-value para ambas medidas, se obtiene un valor inferior a 0.05, lo que sugiere una significación estadística. Por otro lado, los parámetros astrométricos muestran coeficientes elevados, clasificados como correlación fuerte o incluso muy fuerte.

4.1.5. Correlación de la metalicidad con otros catálogos

Se comparó la correlación de la metalicidad de los cúmulos identificados en este trabajo con otros dos catálogos conocidos Kharchenko et al. (2013) y Dias et al. (2002). Donde el catálogo de Kharchenko, que cuenta con 3006 cúmulos, solo 387 presentan valores de metalicidad, mientras que los 1747 cúmulos de Dias sí poseen estos valores.

Para determinar cuántos cúmulos están en este trabajo y en los otros dos, es necesario realizar un conesearch de 0.5 grados en las coordenadas galácticas de los centros de los cúmulos y también analizar aquellos cúmulos cuyos paralajes y movimientos propios estuvieran dentro de 2σ , donde σ es la suma cuadrática de las incertidumbres citadas en ambos catálogos para cada cantidad. Se encontró que 97 cúmulos están presentes tanto en este trabajo como en el de Kharchenko, mientras que 407 están en común entre este trabajo y el de Dias.

Luego, para calcular el coeficiente de Spearman se necesitó la fórmula 4.1 en el catálogo obtenido en este trabajo para calcular el exceso de metalicidad $[Fe/H]$ que es como lo presentan en Kharchenko y Dias. El parámetro Z es como se obtiene la metalicida del cúmulo haciendo uso de fitCMD.

$$[Fe/H] = \log(Z/Z_{\odot}) \quad (4.1)$$

Calculando entonces el coeficiente de Spearman entre Kharchenko y el catálogo de este trabajo, así como entre Dias y el catálogo de este trabajo, obteniéndose valores de 0.9576 y 0.8221 respectivamente, ambos tienen un p-value menor a 0.05. Esto

indica que el programa fitCMD proporciona valores de metalicidad similares a los reportados previamente por otros catálogos y, lo que brinda mayor confianza al estudiar la relación de la metalicidad con los brazos de la galaxia.

4.1.6. Cúmulos no detectados

La cantidad de cúmulos que se lograron identificar es inferior al 30 % de la lista inicial de candidatos. Esto se debe a diversos factores, entre ellos el ajuste de la isocrona realizado por fitCMD, ya que muchos de los cúmulos que podrían haber sido incluidos en el catálogo fueron excluidos porque la isocrona no se adaptaba de manera adecuada a los datos. Un total de 92 cúmulos fueron descartados debido a este problema de mal ajuste, entre ellos están los cúmulos NGC 6144, NGC 2420 y UPK 166, para los cuales el método presentado en este trabajo identificó 1815, 256 y 138 estrellas respectivamente. El CMD de estos se representa en la Figura 4.7 donde se puede notar que el ajuste no es preciso.

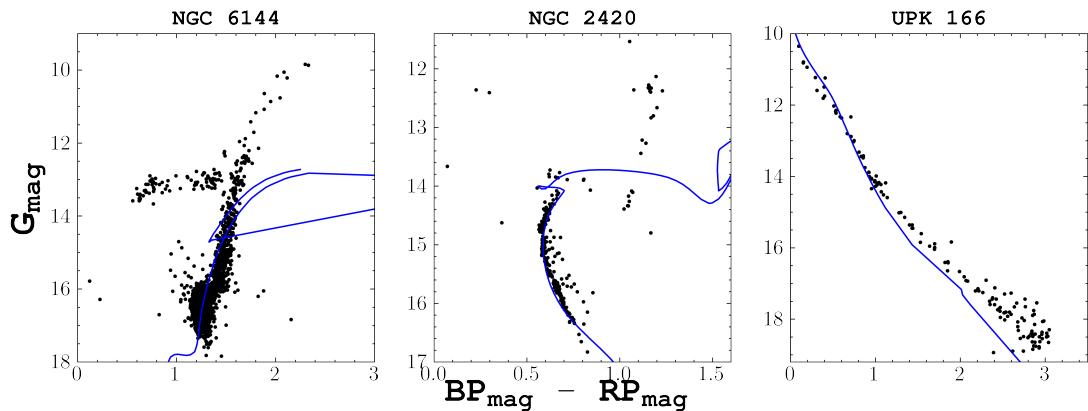


Figura 4.7: Gráfico de 3 cúmulos no incluidos en el catálogo debido a que el ajuste realizado por FitCMD no es preciso.

Por otro lado, la Figura 4.8 ilustra la distribución de distancias y extinciones E(B-V) según aparecen en el catálogo de HR23, tanto para la muestra de OC detectados como para aquellos que no lo fueron. Ambos muestran similitudes para valores de extinción y de distancia bajos, pero la cantidad de cúmulos detectados disminuye en regiones de alta extinción, específicamente para valores de E(B-V) mayores a 3. Además, cúmulos con una distancia mayor a 10.1 kpc no se incluyeron en el catálogo debido a que DBSCAN detectó cúmulos con muy pocas estrellas en las regiones de búsqueda.

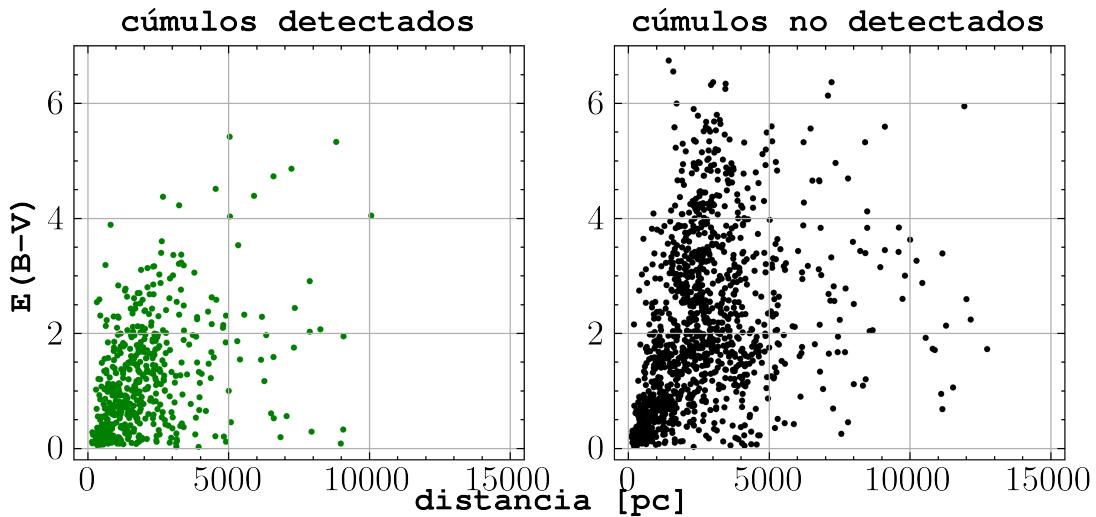


Figura 4.8: Extinción contra distancia (valores tomados de HR23) tanto para cúmulos detectados (gráfica izquierda) como para cúmulos no detectados (derecha).

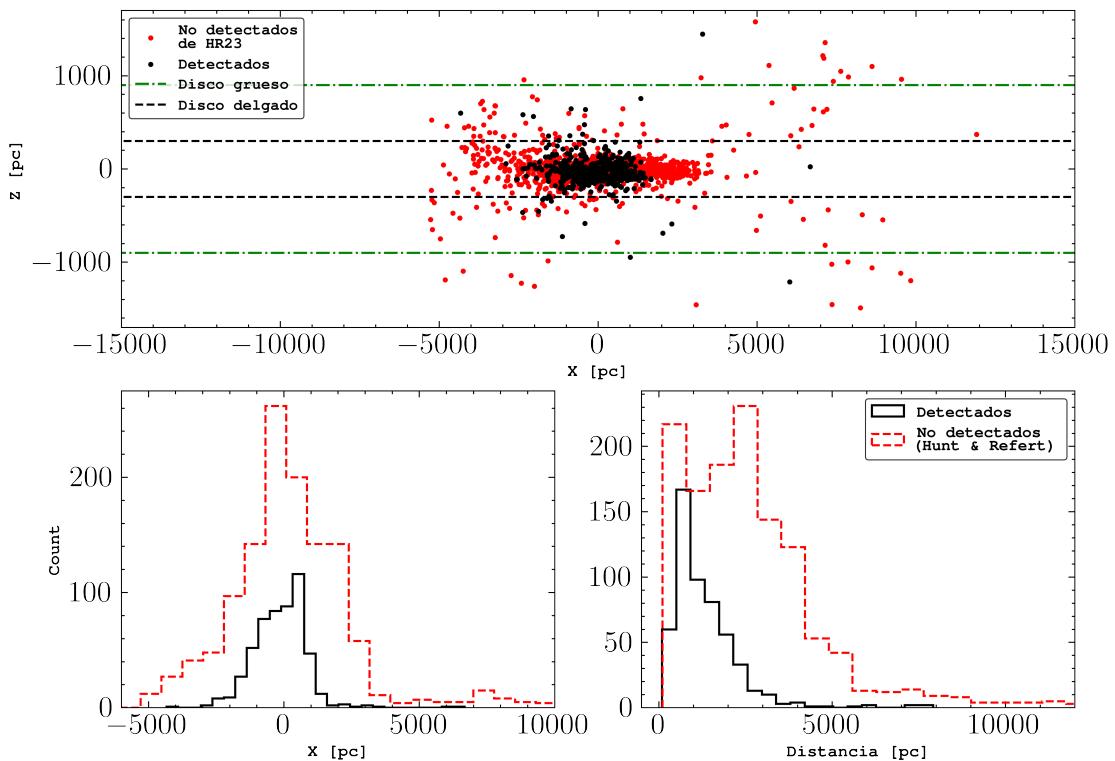


Figura 4.9: *Gráfico superior:* Cúmulos no detectados del catálogo de HR23. Los valores del grosor del disco delgado y del disco grueso fueron tomados de McMillan (2011). *Gráfico inferior izquierdo:* Distribución en distancia Galactocéntrica de los cúmulos abiertos identificados (negro) y aquellos que no se lograron identificar (rojo), incluyendo únicamente los cúmulos de HR23. *Gráfico inferior derecho:* Igual que el panel anterior, para la distribución en distancia.

Finalmente, en el gráfico superior e inferior izquierdo de la Figura 4.9, se estudia la relación entre los cúmulos no detectados y su distribución posicional. Se observa que la mayoría de estos se encuentran en el disco delgado. Además, en el gráfico inferior derecho de la misma Figura se evidencia que la gran mayoría de cúmulos no detectados están a una distancia de 3000 pc. Estos dos desafíos complican la tarea de DBSCAN en la identificación de cúmulos, ya que, más allá de la simple densidad de fuentes a lo largo de la línea de visión, el contraste entre las estrellas del cúmulo y las estrellas del campo en el espacio astrométrico es la variable relevante.

4.1.7. Perspectivas Futuras

A pesar de que investigaciones como las de Jordi et al. (2010) han demostrado que variar la ley de extinción dentro del rango señalado por Fitzpatrick and Massa (2007) tiene un impacto insignificante en la fotometría de Gaia, sería prudente evitar depender únicamente de una única ley de enrojecimiento, como la utilizada por fitCMD. Esto se debe a que las áreas de formación estelar, donde se encuentran los cúmulos más jóvenes, a veces muestran leyes de enrojecimiento inusuales que difieren del entorno interestelar general. Este fenómeno se ha evidenciado en estudios como los de Hur et al. (2012) y Kumar et al. (2014).

Además, es crucial aguardar la conclusión del programa FitCMD para obtener datos más precisos. Aunque este programa facilita la obtención de la edad, extinción y metalicidad de los cúmulos, a veces varía la isocrona que mejor se ajusta a los datos al ejecutar el programa varias veces para el mismo conjunto de estrellas, como se ilustra en la Figura 4.10. Este aspecto resalta la importancia de esperar a la finalización del proceso para asegurar resultados más coherentes y confiables.

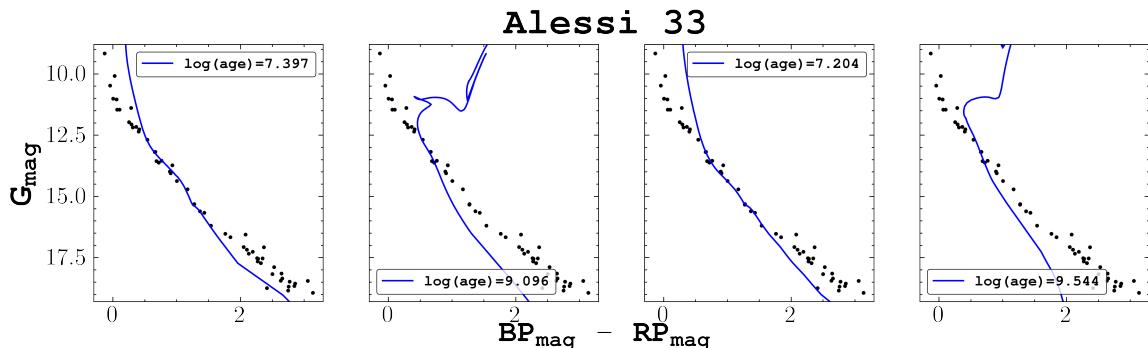


Figura 4.10: Se plasma la discrepancia en el ajuste de la isocrona al realizar cuatro ejecuciones de FitCMD en el cúmulo catalogado como BADFIT, Alessi 33, utilizando el mismo conjunto de datos.

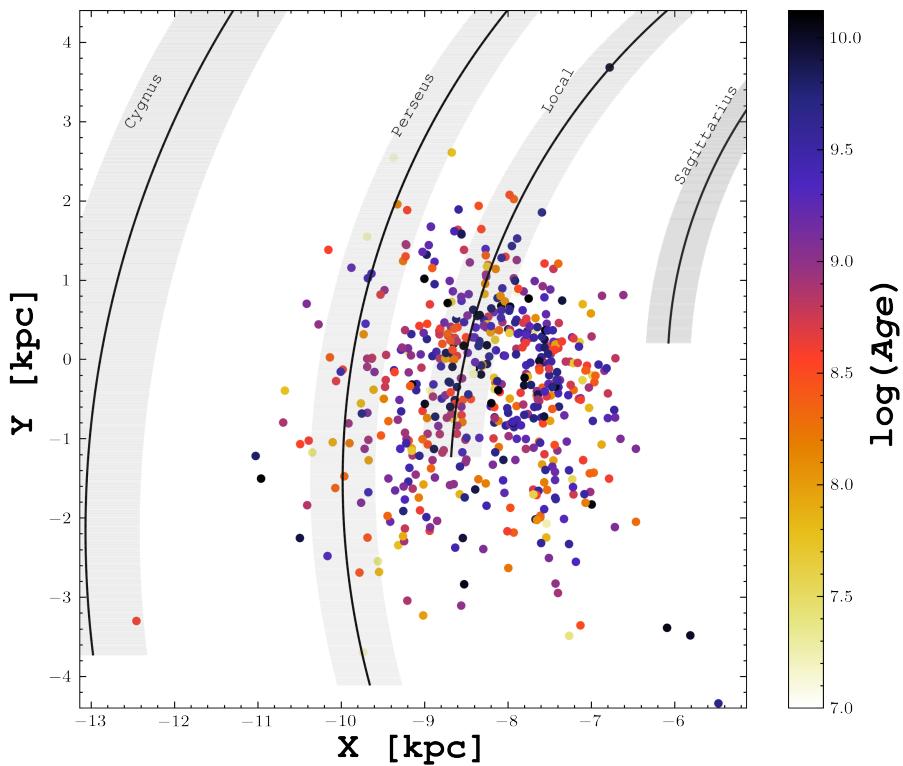


Figura 4.11: Proyección de los cúmulos en el plano galáctico. El color depende de la edad del cúmulo, donde los puntos amarillos son cúmulos jóvenes y los azules cúmulos antiguos.

4.2 | Estructura de la galaxia

Con el valor de la distancia obtenido por fitCMD, se calculan las coordenadas cartesianas galácticas (X,Y,Z) y se plasma en la Figura 4.11 la proyección de la distribución de cúmulos en el plano de la galaxia. Cabe aclarar que el valor tomado como radio galactocéntrico es de 8.15 kpc, que es el valor del modelo de brazo espiral de Reid et al. (2019).

4.2.1. Brazos de la galaxia

En varios estudios, como los realizados por Morgan et al. (1953), Becker and Fenkart (1970), Dias and Lépine (2005), Castro-Ginard et al. (2021) y Castro-Ginard et al. (2022), se ha analizado la relación entre los brazos de la galaxia y los cúmulos jóvenes. Estos estudios sugieren que los grupos de cúmulos jóvenes suelen estar ubicados en los brazos galácticos.

Esta correlación se muestra en la Figura 4.12, donde la línea sombreada en los gráficos representa los brazos de la galaxia, cuyo modelo se basa en los datos recopilados por Reid et al. (2019). Estos investigadores se centraron en las regiones de formación estelar de alta masa en nuestra galaxia y concluyeron que la Vía Láctea presenta una estructura espiral con cuatro brazos principales, con algunos segmentos adicionales

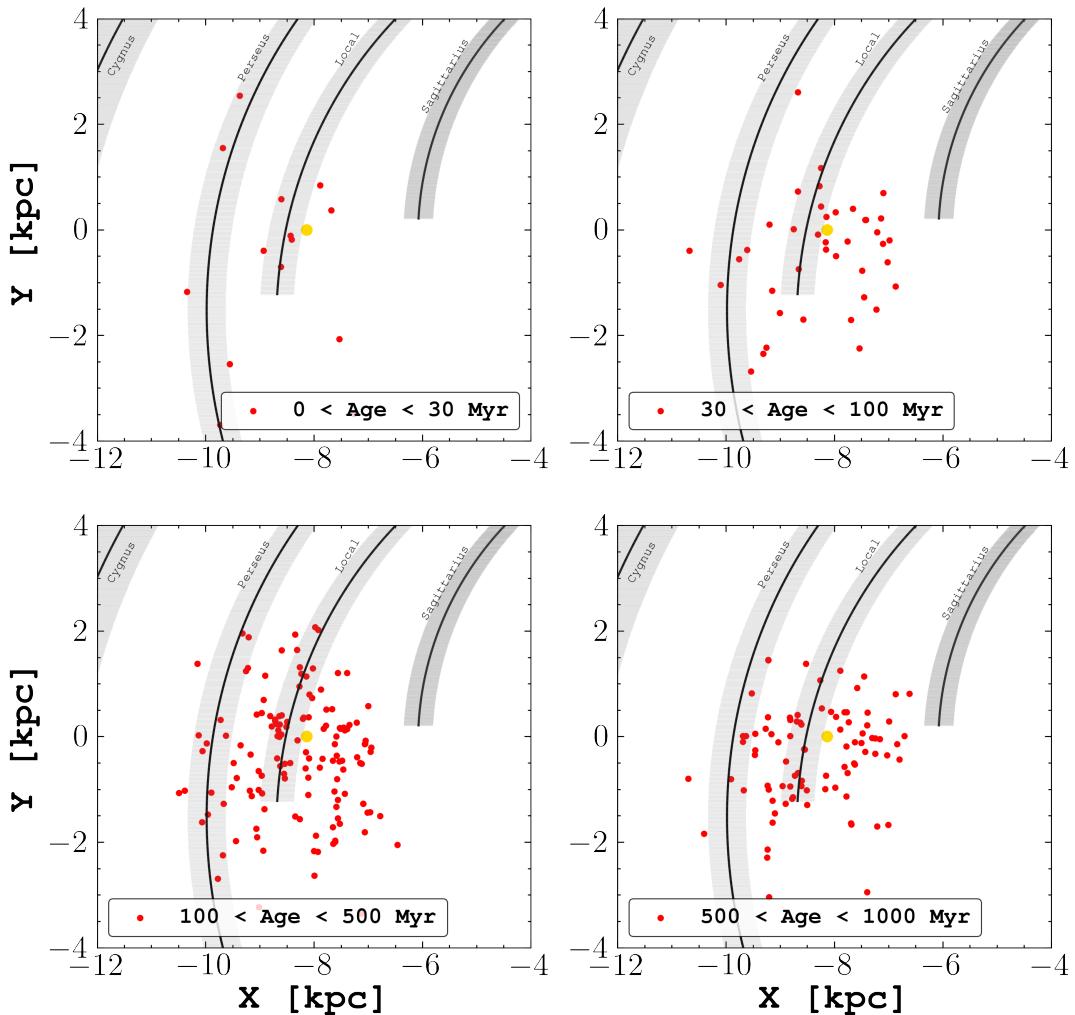


Figura 4.12: Distribución de los Cúmulos Abiertos en coordenadas galactocéntricas XY en diferentes rangos de edad. El centro galáctico se encuentra hacia valores positivos de X, y la dirección de la rotación galáctica es hacia valores positivos de Y. De izquierda a derecha y de arriba abajo corresponden a cúmulos con: edades inferiores a 30 millones de años, de 30 a 100 millones de años, de 100 a 500 millones de años, y de 500 a 1000 millones de años. La región sombreada son los brazos espirales definidos por Reid et al. (2019).

de brazo. Además, también se grafica la posición de los cúmulos estelares según su edad, siendo el gráfico superior izquierdo el de los cúmulos con edades menores a 30 Myr y el inferior derecho muestra los cúmulos con edades entre 500 y 1000 Myr. Se observa que los cúmulos más jóvenes están cerca a los brazos de la galaxia mientras que los más viejos tienen una distribución más uniforme alrededor del plano de esta. Para los rangos utilizados en las cuatro gráficas de la Figura, se observa que los porcentajes de cúmulos dentro de las regiones sombreadas comparado con la cantidad de cúmulos dentro de ese rango de edades son del 66.6 %, 37.9 %, 35.3 % y 29.4 % en orden del intervalo más joven al más antiguo, sin considerar la coordenada Z.

Además, si se distribuyen los cúmulos en los 3 intervalos usados por [Cantat-Gaudin et al. \(2020\)](#): $t < 158$ Myr (cúmulos jóvenes), $158 < t < 1000$ Myr (cúmulos intermedios), $t > 1000$ Myr (cúmulos viejos); el porcentaje de cúmulos en las regiones sombreadas es de 52.5 %, 36.0 % y 21.3 %, respectivamente. Mostrando así que, en porcentaje, es más probable encontrar cúmulos jóvenes cerca a los brazos de la galaxia.

Por otro lado, en cuanto a la Figura 4.11, se aprecia que los cúmulos estelares no se encuentran muy distantes del Sol, con coordenadas de 8.15 kpc en la dirección X y 0 kpc en la dirección Y (valores tomados del modelo de Reid). En su mayoría, estos cúmulos presentan coordenadas X inferiores a 11 kpc, siendo la excepción el poco estudiado cúmulo “King 23” ([Sariya et al. \(2021\)](#)), que muestra una coordenada X mayor. Además, se nota la falta de cúmulos en las regiones demarcadas como los brazos “Cygnus” y “Sagittarius”. Esto se debe al método de cálculo de distancias utilizado por fitCMD, que, como se analiza en la sección 3.6.1, se ve influenciado principalmente por el modelo de extinción empleado por el programa.

Tras examinar la extinción calculada por el ajuste CMD de los cúmulos y su representación en la Figura 4.13, se observa que los cúmulos más distantes muestran una extinción más pronunciada en comparación con los más cercanos. Además, se observa que los cúmulos orientados hacia el centro de la galaxia sufren un aumento más significativo en la tasa de extinción en comparación con aquellos ubicados en las regiones externas de la galaxia. Asimismo, en el gráfico de la derecha de la misma Figura se observa que los cúmulos con mayor extinción se encuentran cerca del plano galáctico.

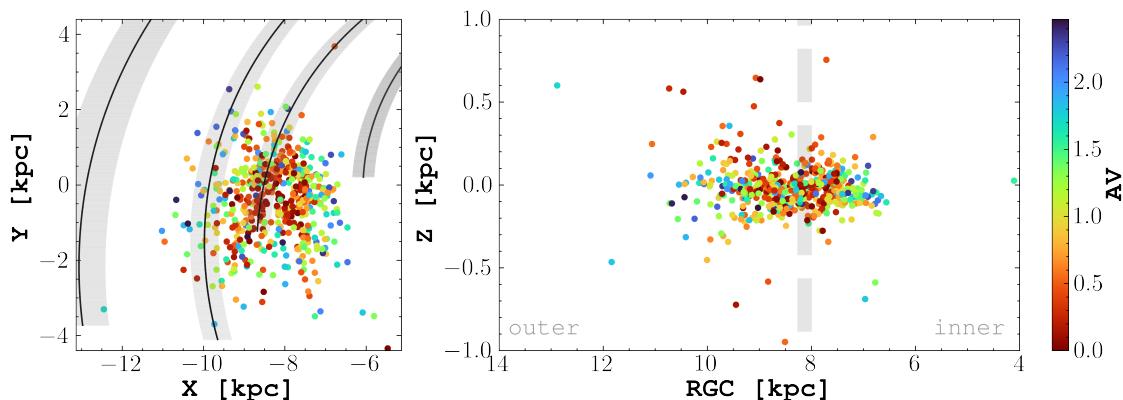


Figura 4.13: Distribución de cúmulos en las coordenadas galácticas XY (izquierda) y altitud contra radio galactocéntrico (derecha), el color depende de la extinción AV.

4.2.1.1. Metalicidad

Se estudió la relación entre la metalicidad de los cúmulos y los brazos de la galaxia utilizando 2 modelos diferentes: el de [Reid et al. \(2019\)](#) y el de [Poggio et al. \(2022\)](#). Este último se basa en el estudio de estrellas gigantes brillantes en el disco de la galaxia, identificando regiones con una mayor presencia de estas estrellas con alta metalicidad, las cuales están asociadas a los brazos de la galaxia.

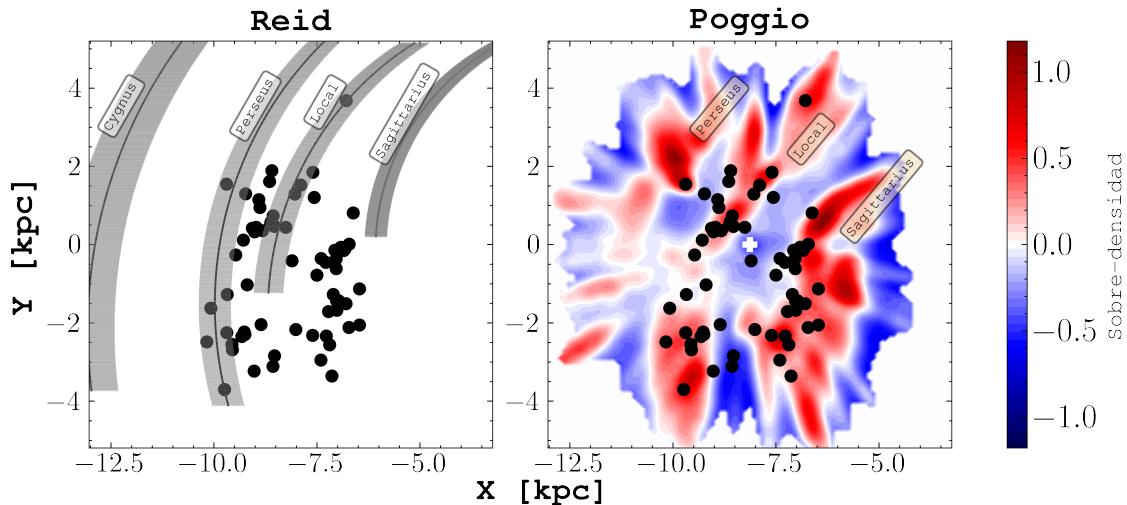


Figura 4.14: Correlación entre los cúmulos con una metalicidad superior a 1.5 veces la del Sol y los modelos para los brazos de la galaxia de Reid et al. (2019) (gráfico izquierdo) y Poggio et al. (2022) (gráfico derecho).

En la Figura 4.14 se plasman las posiciones galácticas de los cúmulos del catálogo acá presentado que tienen una metalicidad mayor a 1.5 veces la metalicidad solar. En el gráfico izquierdo, se muestra el modelo de Reid, donde la región sombreada representa los brazos de la galaxia. En el gráfico derecho, se presenta el modelo de Poggio, donde las regiones con valores positivos (zona de color rojo) indican las áreas más ricas en metales. Se observa que el porcentaje de cúmulos en las regiones sombreadas es del 27.69 % para Reid mientras que para Poggio es del 84.62 %, esto se debe a que los brazos de Reid son curvas con un ancho establecido mientras que para Poggio son formas más irregulares y más grandes.

También se destaca que la estructura y la posición de los brazos de la galaxia varían según los parámetros que se están investigando. Reid se centró principalmente en los “molecular masers” asociados a estrellas muy jóvenes de alta masa, mientras que Poggio se enfocó en la metalicidad de las estrellas gigantes brillantes. Aunque ambos lograron trazar los brazos, lo hicieron de manera diferente. Esto sugiere que no existe un modelo único y completo para la estructura correcta y precisa de los brazos de la galaxia, lo que destaca la importancia de continuar estudiando y mapeando su forma y ubicación.

4.2.2. Radiogalactocéntrico y coordenada Z

Al analizar la Figura 4.15, que estudia la relación entre la coordenada Z y el radiogalactocéntrico (RGC) de los cúmulos, se observa que se destacan dos picos en la distribución de los cúmulos jóvenes. Estos picos corresponden al brazo local y al brazo de Perseo, localizados aproximadamente a $RGC = 8$ kpc y $RGC = 10$ kpc, respectivamente.

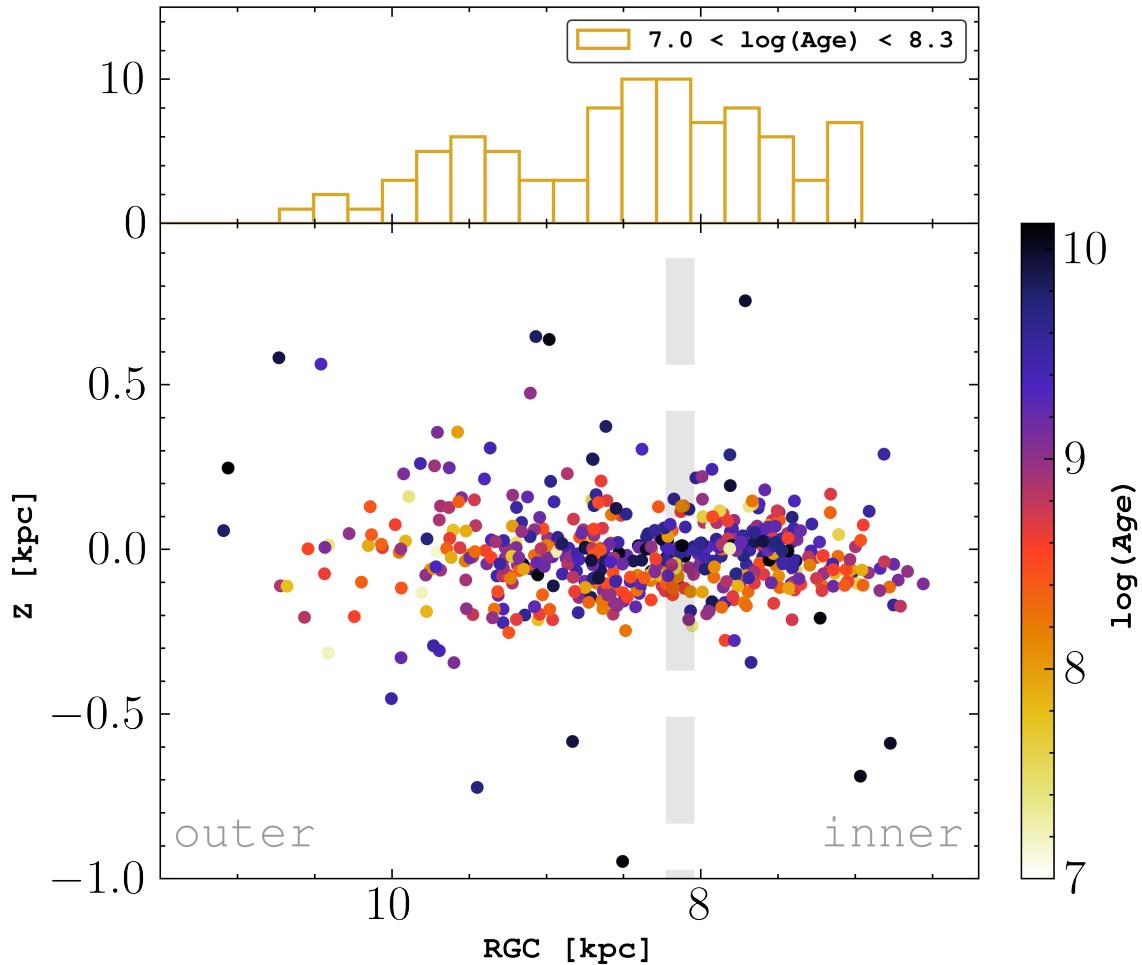


Figura 4.15: *Panel superior:* Histograma de cantidad de cúmulos con una edad menor a 200 Myr. *Panel inferior:* Distribución de todos los cúmulos en el eje Z y el color depende de la edad del cúmulo. La línea vertical grisácea representa la posición del Sol $R_\odot = 8.15$ kpc

Además, se aprecia que los cúmulos más jóvenes (puntos amarillos) están principalmente cerca al plano ($Z=0$), mientras que los cúmulos más antiguos (puntos azules y negros) están tanto cerca al plano como alejados de este. Asimismo, al alejarnos del centro de la galaxia, se evidencia un incremento en la cantidad de cúmulos estelares ubicados fuera del plano galáctico. Por el contrario, en las regiones más internas, estos cúmulos tienden a encontrarse más cerca del plano.

Analizando ahora cómo se distribuye la metalicidad de los cúmulos en la coordenada Z de la galaxia, lo que se hace es dividir el rango de las metalicidades (se utiliza la columna de Z/Z_{sun} que para estos cúmulos tienen valores entre 0 y 1.97) en 3 subrangos de igual tamaño y tomar los cúmulos que pertenecen al rango. Luego, para cada rango se calcula cuál es la distancia al plano en la que está el 95 % de los cúmulos del rango. Obteniendo así que para los cúmulos que tienen un Z/Z_{sun} entre $[0, 2/3]$ el 95 % de los cúmulos están en una distancia aproximada al plano de 378 pc, para los cúmulos con un Z/Z_{sun} entre $[2/3, 4/3]$ el 95 % de los cúmulos están a 285

pc y para los cúmulos con Z/Z_{sun} entre $[4/3, 2]$ el 95 % de los cúmulos están a 259 pc.

También se estudió la diferencia entre los cúmulos con metalicidades altas ($Z/Z_{\text{sun}}=[1/10, 3]$) y cúmulos con metalicidades bajas ($Z/Z_{\text{sun}} = [1/1000, 1/10]$), y se obtuvo que la media de la distancia al plano de la galaxia para los cúmulos con metalicidad alta es de 118.7 pc y para los cúmulos de metalicidad baja es de 235.6 pc. Mostrando así que los cúmulos con mayor metalicidad están más cerca del plano de la galaxia que aquellos con una metalicidad más baja.

Capítulo 5

Conclusiones

La metodología empleada para determinar cúmulos estelares es una tarea de la minería de datos que utiliza un algoritmo de agrupamiento no supervisado, DBSCAN, para encontrar grupos de estrellas que comparten coordenadas comunes (l , b , parallax, μ_α , μ_δ) y analiza si hay o no un cúmulo. También, haciendo uso de pyUPMASK se estudia la probabilidad de membresía de cada estrella y se eliminan aquellas con una probabilidad menor al 50 %. Además, el diseño completo de la metodología necesita una intervención manual mínima, indicando que es factible aplicarla a un conjunto de datos extenso como el Gaia DR3.

La aplicación de la metodología descrita en este trabajo condujo a obtener un catálogo de 570 cúmulos con información acerca distancia, posición en la galaxia, movimientos propios en declinación y ascensión recta, edad, extinción, masa y metalicidad. Muchos de estos valores fueron tomados al ajustar una isocrona al CMD de los datos a través del programa fitCMD. Este catálogo de cúmulos junto con los miembros de estos se encuentra en <https://github.com/slastrar/Thesis-Catalogue>.

Se hizo crossmatch con otros catálogos importantes aparte de HR23 y se obtuvo que hasta 521 de los 534 cúmulos están en esos catálogos. Obteniendo valores del coeficiente de correlación de Spearman de hasta el 93.5 % para parámetros astrométricos y 56.1 % para parámetros astrofísicos. A diferencia de estos catálogos, es que acá se ofrece un catálogo completo con propiedades astrofísicas que muchos no agregan. Por ejemplo, [Bossini et al. \(2019\)](#) no brinda información acerca de la masa de los cúmulos y HR23 de la extinción.

Es importante mencionar que la metodología utilizada es similar a la de otros trabajos y que, aunque estos puedan descubrir más cúmulos no identificados por este escrito, a su vez podrían pasar por alto algunos otros que acá sí se identificaron. En resumen, se concluye que la búsqueda utilizando un único modelo de detección no logra identificar todas las estructuras estelares existentes, y que diferentes algoritmos de aprendizaje automático con este fin se complementan mutuamente.

Se concluye que los cúmulos más jóvenes están más cerca a los brazos y al plano de la galaxia que aquellos más viejos. También se hace este análisis para la metalicidad de los cúmulos, donde se observa que los de mayor metalicidad están más cerca al plano y a los brazos de la galaxia.

La próxima década de investigación de cúmulos estelares promete mantener su carácter emocionante y dinámico. Esto se debe a que la calidad de los datos continuará en aumento gracias al próximo lanzamiento del Gaia DR4, el cual tardará casi el doble de meses de observación que el Gaia DR3. Además, el esperado lanzamiento final de Gaia, el DR5, se basará en una recopilación de datos de aproximadamente diez años, lo que supondrá mejoras significativas en la precisión de los datos astrométricos. Este avance inevitablemente conducirá a la identificación de nuevos cúmulos estelares y a mejoras en la relación señal/ruido.

Bibliografía

Emily L. Hunt and Sabine Reffert. Improving the open cluster census. II. An all-sky cluster catalogue with Gaia DR3. *A&A*, 673:A114, May 2023. doi: 10.1051/0004-6361/202346285.

T. Cantat-Gaudin, F. Anders, A. Castro-Ginard, C. Jordi, M. Romero-Gómez, C. Soubiran, L. Casamiquela, Y. Tarricq, A. Moitinho, A. Vallenari, A. Bragaglia, A. Krone-Martins, and M. Kounkel. Painting a portrait of the Galactic disc with its stellar clusters. *A&A*, 640:A1, August 2020. doi: 10.1051/0004-6361/202038192.

D. Bossini, A. Vallenari, A. Bragaglia, T. Cantat-Gaudin, R. Sordo, L. Balaguer-Nuñez, C. Jordi, A. Moitinho, C. Soubiran, L. Casamiquela, R. Carrera, and U. Heiter. Age determination for 269 Gaia DR2 open clusters. *A&A*, 623:A108, March 2019. doi: 10.1051/0004-6361/201834693.

T. Cantat-Gaudin, C. Jordi, A. Vallenari, A. Bragaglia, L. Balaguer-Nuñez, C. Soubiran, D. Bossini, A. Moitinho, A. Castro-Ginard, A. Krone-Martins, L. Casamiquela, R. Sordo, and R. Carrera. A Gaia DR2 view of the open cluster population in the Milky Way. *A&A*, 618:A93, October 2018. doi: 10.1051/0004-6361/201833476.

A. Castro-Ginard, C. Jordi, X. Luri, J. Álvarez Cid-Fuentes, L. Casamiquela, F. Anders, T. Cantat-Gaudin, M. Monguió, L. Balaguer-Nuñez, S. Solà, and R.M. Badia. Hunting for open clusters in Gaia DR2: 582 new open clusters in the Galactic disc. *A&A*, 635:A45, March 2020. doi: 10.1051/0004-6361/201937386.

Zhi-Hong He, Ye Xu, and Li-Gang Hou. Search for age pattern across spiral arms of the Milky Way. *Research in Astronomy and Astrophysics*, 21(1):009, January 2021a. doi: 10.1088/1674-4527/21/1/9.

M.J. Reid, K.M. Menten, A. Brunthaler, X.W. Zheng, T.M. Dame, Y. Xu, Y. Wu, B. Zhang, A. Sanna, M. Sato, K. Hachisuka, Y.K. Choi, K. Immer, L. Moscadelli, K.L.J. Rygl, and A. Bartkiewicz. Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way. *Astrophysical Journal*, 783(2):130, March 2014. doi: 10.1088/0004-637X/783/2/130.

Gaia Collaboration and Prusti . The Gaia mission. *A&A*, 595:A1, November 2016. doi: 10.1051/0004-6361/201629272.

Gaia Collaboration and Brown . Gaia Early Data Release 3. Summary of the contents and survey properties. *A&A*, 649:A1, May 2021. doi: 10.1051/0004-6361/202039657.

- Gaia Collaboration and Brown. Gaia Data Release 1. Summary of the astrometric, photometric, and survey properties. *A&A*, 595:A2, November 2016. doi: 10.1051/0004-6361/201629512.
- Gaia Collaboration and Brown. Gaia Data Release 2. Summary of the contents and survey properties. *A&A*, 616:A1, August 2018. doi: 10.1051/0004-6361/201833051.
- Bradley W Carroll and Dale A Ostlie. *An introduction to modern astrophysics*. Cambridge University Press, 2017.
- Francis LeBlanc. *An Introduction to Stellar Astrophysics*. Wiley, 2010.
- O.R. Pols. *Stellar Structure and Evolution*. Astronomical Institute Utrecht, 2011.
- Lyman S Spitzer Jr. *Dynamical evolution of globular clusters*, volume 25. Princeton University Press, 2014.
- Richard B. Larson. The physics of star formation. *Reports on Progress in Physics*, 66(10):1651–1697, October 2003. doi: 10.1088/0034-4885/66/10/R03.
- C.J. Evans, M.B. Kennedy, P.L. Dufton, I.D. Howarth, N.R. Walborn, N. Markova, J.S. Clark, S.E. de Mink, A. de Koter, P.R. Dunstall, V. Hénault-Brunet, J. Maíz Apellániz, C.M. McEvoy, H. Sana, S. Simón-Díaz, W.D. Taylor, and J.S. Vink. The VLT-FLAMES Tarantula Survey. XVIII. Classifications and radial velocities of the B-type stars. *A&A*, 574:A13, February 2015. doi: 10.1051/0004-6361/201424414.
- Richard W. Michie. The Dynamics of Star Clusters. *Annual Review of Astron and Astrophysics*, 2:49, January 1964. doi: 10.1146/annurev.aa.02.090164.000405.
- Gyu Eun Lee. The Boltzmann equation. December 2019.
- T. D. Kinman. Notes on Faint Star Clusters. *Annual Review of Astron and Astrophysics*, 1962.
- T. Bensby, S. Feltzing, and M.S. Oey. Exploring the Milky Way stellar disk. A detailed elemental abundance study of 714 F and G dwarf stars in the solar neighbourhood. *A&A*, 562:A71, February 2014. doi: 10.1051/0004-6361/201322631.
- Brian Dunbar. Milky way churns out seven new stars per year, scientists say, Feb 2008. URL https://www.nasa.gov/centers/goddard/news/topstory/2006/milkyway_seven.html.
- M.W. Guidry. *Stars and stellar processes*. 2019.
- Van den Bosch, Frank C. Houjun, Mo, and Simon White. *Galaxy Formation and Evolution*. 2010.
- Ye Xu, Li-Gang Hou, and Yuan-Wei Wu. The spiral structure of the Milky Way. *Research in Astronomy and Astrophysics*, 18(12):146, December 2018. doi: 10.1088/1674-4527/18/12/146.

- D. Crampton and Y.M. Georgelin. The distribution of optical H II regions in our Galaxy. *A&A*, 40:317–321, May 1975.
- M. Fich and L. Blitz. Optical H II regions in the outer Galaxy. *Astrophysical Journal*, 279:125–135, April 1984. doi: 10.1086/161872.
- V.S. Avedisova. Galactic Spiral Structure Outlined by Diffuse Nebulae in the Solar Neighborhood. *Soviet Astronomy Letters*, 11:185–188, June 1985.
- Ed Churchwell, Brian L. Babler, Marilyn R. Meade, Barbara A. Whitney, Robert Benjamin, Remy Indebetouw, Claudia Cyganowski, Thomas P. Robitaille, Matthew Povich, Christer Watson, and Steve Bracker. The spitzer/glimpse surveys: A new view of the milky way. *Publications of the Astronomical Society of the Pacific*, 121(877):213, mar 2009. doi: 10.1086/597811. URL <https://dx.doi.org/10.1086/597811>.
- A. Castro-Ginard, P.J. McMillan, X. Luri, C. Jordi, M. Romero-Gómez, T. Cantat-Gaudin, L. Casamiquela, Y. Tarricq, C. Soubiran, and F. Anders. Milky Way spiral arms from open clusters in Gaia EDR3. *A&A*, 652:A162, August 2021. doi: 10.1051/0004-6361/202039751.
- Seth Gossage, Charlie Conroy, Aaron Dotter, Jieun Choi, Philip Rosenfield, Philip Cargile, and Andrew Dolphin. Age Determinations of the Hyades, Praesepe, and Pleiades via MESA Models with Rotation. *Astrophysical Journal*, 863(1):67, August 2018. doi: 10.3847/1538-4357/aad0a0.
- V. Hill, A. Lecureur, A. Gómez, M. Zoccali, M. Schultheis, C. Babusiaux, F. Royer, B. Barbuy, F. Arenou, D. Minniti, and S. Ortolani. The metallicity distribution of bulge clump giants in Baade’s window. *A&A*, 534:A80, October 2011. doi: 10.1051/0004-6361/200913757.
- A. Rojas-Arriagada, A. Recio-Blanco, P. de Laverny, Š. Mikolaitis, F. Matteucci, E. Spitoni, M. Schultheis, M. Hayden, V. Hill, M. Zoccali, D. Minniti, O.A. Gonzalez, G. Gilmore, S. Randich, S. Feltzing, E.J. Alfaro, C. Babusiaux, T. Bensby, A. Bragaglia, E. Flaccomio, S.E. Koposov, E. Pancino, A. Bayo, G. Carraro, A.R. Casey, M.T. Costado, F. Damiani, P. Donati, E. Franciosini, A. Hourihane, P. Jofré, C. Lardo, J. Lewis, K. Lind, L. Magrini, L. Morbidelli, G.G. Sacco, C.C. Worley, and S. Zaggia. The Gaia-ESO Survey: Exploring the complex nature and origins of the Galactic bulge populations. *A&A*, 601:A140, May 2017. doi: 10.1051/0004-6361/201629160.
- William E. Harris. A New Catalog of Globular Clusters in the Milky Way. *arXiv e-prints*, art. arXiv:1012.3224, December 2010. doi: 10.48550/arXiv.1012.3224.
- A.T. Bajkova and V.V. Bobylev. A New Catalog of orbits of 152 Globular Clusters from Gaia EDR3. *arXiv e-prints*, art. arXiv:2212.00739, November 2022. doi: 10.48550/arXiv.2212.00739.

- C. Muñoz, D. Geisler, S. Villanova, I. Saviane, C.C. Cortés, B. Dias, R.E. Cohen, F. Mauro, and C. Moni Bidin. Chemical analysis of NGC 6528: one of the most metal-rich bulge globular clusters. *A&A*, 620:A96, December 2018. doi: 10.1051/0004-6361/201833373.
- Lawrence M. Krauss and Brian Chaboyer. Age Estimates of Globular Clusters in the Milky Way: Constraints on Cosmology. *Science*, 299(5603):65–70, January 2003. doi: 10.1126/science.1075631.
- Cesare Chiosi, Gianpaolo Bertelli, and Alessandro Bressan. New developments in understanding the HR diagram. *Annual Review of Astron and Astrophys*, 30: 235–285, January 1992. doi: 10.1146/annurev.aa.30.090192.001315.
- N. Langer and R.P. Kudritzki. The spectroscopic Hertzsprung-Russell diagram. *A&A*, 564:A52, April 2014. doi: 10.1051/0004-6361/201423374.
- G. Larsson-Leander. The Hertzsprung-Russel Diagram and Stellar Ages (review Paper). page 2, January 1972.
- Fred C. Adams and Gregory Laughlin. A dying universe: the long-term fate and evolutionof astrophysical objects. *Reviews of Modern Physics*, 69(2):337–372, April 1997. doi: 10.1103/RevModPhys.69.337.
- Onno R. Pols, Klaus-Peter Schröder, Jarrod R. Hurley, Christopher A. Tout, and Peter P. Eggleton. Stellar evolution models for $Z = 0.0001$ to 0.03. *Monthly Notices of the Royal Astronomical Society*, 298(2):525–536, 08 1998. ISSN 0035-8711. doi: 10.1046/j.1365-8711.1998.01658.x. URL <https://doi.org/10.1046/j.1365-8711.1998.01658.x>.
- Richard Pogge. Lecture 16: The evolution of low-mass stars, Jan 2006. URL <https://www.astronomy.ohio-state.edu/pogge.1/Ast162/Unit2/lowmass.html>.
- Universe of Northen Iowa. Stellar birth and middle age. URL <https://sites.uni.edu/morgans/astro/course/Notes/section2/new7.html>.
- Facundo David Moyano. Tesis de Licenciatura en Astronomía. *Universidad Nacional de La Plata*, 03 2019. URL http://sedici.unlp.edu.ar/bitstream/handle/10915/77876/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y.
- G.I. Perren, R.A. Vázquez, and A.E. Piatti. AStCA: Automated Stellar Cluster Analysis. *A&A*, 576:A6, April 2015. doi: 10.1051/0004-6361/201424946.
- Charles Bonatto. An efficient approach to extract parameters from star cluster CMDs: fitCMD. *Monthly Notices of the RAS*, 483(2):2758–2765, February 2019. doi: 10.1093/mnras/sty3291.
- A. Castro-Ginard, C. Jordi, X. Luri, T. Cantat-Gaudin, J.M. Carrasco, L. Casamiquela, F. Anders, L. Balaguer-Nuñez, and R.M. Badia. Hunting for open clusters in Gaia EDR3: 628 new open clusters found with OCfinder. *A&A*, 661:A118, May 2022. doi: 10.1051/0004-6361/202142568.

Mark Gieles and Simon F. Portegies Zwart. The distinction between star clusters and associations. *Monthly Notices of the Royal Astronomical Society*, 410(1):L6–L7, January 2011. doi: 10.1111/j.1745-3933.2010.00967.x.

Jason L. Curtis, Angie Wolfgang, Jason T. Wright, John M. Brewer, and John Asher Johnson. Ruprecht 147: The Oldest Nearby Open Cluster as a New Benchmark for Stellar Astrophysics. *Astronomical Journal*, 145(5):134, May 2013. doi: 10.1088/0004-6256/145/5/134.

Hyobin Im, Sang Chul Kim, Jaemann Kyeong, Hong Soo Park, and Joon Hyeop Lee. Photometry of the Four Anti-Galactocentric Old Open Clusters: Czernik 30, Berkeley 34, Berkeley 75, and Berkeley 76. *The American Astronomical Society*, 165(2):53, February 2023. doi: 10.3847/1538-3881/aca7fb.

C. Payne-Gaposchkin. *Stars and clusters*. 1979.

L. Lindegren, S.A. Klioner, J. Hernández, A. Bombrun, M. Ramos-Lerate, H. Steidelmüller, U. Bastian, M. Biermann, A. de Torres, E. Gerlach, R. Geyer, T. Hilger, D. Hobbs, U. Lammers, P.J. McMillan, C.A. Stephenson, J. Castañeda, M. Davidson, C. Fabricius, G. Gracia-Abril, J. Portell, N. Rowell, D. Teyssier, F. Torra, S. Bartolomé, M. Clotet, N. Garralda, J.J. González-Vidal, J. Torra, U. Abbas, M. Altmann, E. Anglada Varela, L. Balaguer-Nuñez, Z. Balog, C. Barache, U. Becciani, M. Bernet, S. Bertone, L. Bianchi, S. Bouquillon, A.G.A. Brown, B. Bucciarelli, D. Busonero, A.G. Butkevich, R. Buzzi, R. Cancelliere, T. Carlucci, P. Charlot, M. R.L. Cioni, M. Crosta, C. Crowley, E.F. del Peloso, E. del Pozo, R. Drimmel, P. Esquej, A. Fienga, E. Fraile, M. Gai, M. Garcia-Reinaldos, R. Guerra, N.C. Hambly, M. Hauser, K. Janssen, S. Jordan, Z. Kostrzewska-Rutkowska, M.G. Lattanzi, S. Liao, E. Licata, T.A. Lister, W. Löffler, J.M. Marchant, A. Masip, F. Mignard, A. Mints, D. Molina, A. Mora, R. Morbidelli, C.P. Murphy, C. Pagani, P. Panuzzo, X. Peñalosa Esteller, E. Poggio, P. Re Fiorentin, A. Riva, A. Sagristà Sellés, V. Sanchez Gimenez, M. Sarasso, E. Sciacca, H.I. Siddiqui, R.L. Smart, D. Souami, A. Spagna, I.A. Steele, F. Taris, E. Utrilla, W. van Reeven, and A. Vecchiato. Gaia Early Data Release 3. The astrometric solution. *A&A*, 649:A2, May 2021. doi: 10.1051/0004-6361/202039709.

Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. pages 226–331, January 1996.

A. Castro-Ginard, C. Jordi, X. Luri, F. Julbe, M. Morvan, L. Balaguer-Nuñez, and T. Cantat-Gaudin. A new method for unveiling open clusters in Gaia. New nearby open clusters confirmed by DR2. *A&A*, 618:A59, October 2018. doi: 10.1051/0004-6361/201833390.

Zhi-Hong He, Ye Xu, Chao-Jie Hao, Zhen-Yu Wu, and Jing-Jing Li. A catalogue of 74 new open clusters found in Gaia Data-Release 2. *Research in Astronomy and Astrophysics*, 21(4):093, May 2021b. doi: 10.1088/1674-4527/21/4/93.

- M.S. Pera, G.I. Perren, A. Moitinho, H.D. Navone, and R.A. Vazquez. pyUPMASK: an improved unsupervised clustering algorithm. *A&A*, 650:A109, June 2021. doi: 10.1051/0004-6361/202040252.
- A. Krone-Martins and A. Moitinho. UPMASK: unsupervised photometric membership assignment in stellar clusters. *A&A*, 561:A57, January 2014. doi: 10.1051/0004-6361/201321143.
- Michael Marks and Pavel Kroupa. Initial conditions for globular clusters and assembly of the old globular cluster population of the Milky Way. *Monthly Notices of the RAS*, 406(3):2000–2012, August 2010. doi: 10.1111/j.1365-2966.2010.16813.x.
- B. W. Silverman. Density estimation for statistics and data analysis. 1986.
- Arsalane Chouaib Guidoum. Kernel estimator and bandwidth selection for density and its derivatives. *Department of Probabilities and Statistics, University of Science and Technology, Houari Boumediene, Algeria*, 2015. doi: 10.13140/RG.2.34597.93925.
- Jan Rybizki, Gregory M. Green, Hans-Walter Rix, Kareem El-Badry, Markus Demleitner, Eleonora Zari, Andrzej Udalski, Richard L. Smart, and Andrew Gould. A classifier for spurious astrometric solutions in Gaia eDR3. *Monthly Notices of the RAS*, 510(2):2597–2616, February 2022. doi: 10.1093/mnras/stab3588.
- Paul J. McMillan. Mass models of the Milky Way. *Monthly Notices of the RAS*, 414(3):2446–2457, July 2011. doi: 10.1111/j.1365-2966.2011.18564.x.
- E. Poggio, A. Recio-Blanco, P. A. Palicio, P. Re Fiorentin, P. de Laverny, R. Drimmel, G. Kordopatis, M. G. Lattanzi, M. Schultheis, A. Spagna, and E. Spitoni. The chemical signature of the Galactic spiral arms revealed by Gaia DR3. *Astronomy and Astrophysics*, 666:L4, October 2022. doi: 10.1051/0004-6361/202244361.
- Zhihong He, Xiaochen Liu, Yangping Luo, Kun Wang, and Qingquan Jiang. Unveiling Hidden Stellar Aggregates in the Milky Way: 1656 New Star Clusters Found in Gaia EDR3. *Astrophysical Journal, Supplement*, 264(1):8, January 2023. doi: 10.3847/1538-4365/ac9af8.
- Charles Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904. ISSN 00029556. URL <http://www.jstor.org/stable/1412159>.
- Patrick Schober, Christa Boer, and Lothar A Schwarte. Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia*, 126(5):1763–1768, 2018.
- N. V. Kharchenko, A. E. Piskunov, E. Schilbach, S. Röser, and R. D. Scholz. Global survey of star clusters in the Milky Way. II. The catalogue of basic parameters. *Astronomy and Astrophysics*, 558:A53, October 2013. doi: 10.1051/0004-6361/201322302.

- W. S. Dias, B. S. Alessi, A. Moitinho, and J. R. D. Lépine. New catalogue of optically visible open clusters and candidates. *Astronomy and Astrophysics*, 389: 871–873, July 2002. doi: 10.1051/0004-6361:20020668.
- C. Jordi, M. Gebran, J.M. Carrasco, J. de Bruijne, H. Voss, C. Fabricius, J. Knude, A. Vallenari, R. Kohley, and A. Mora. Gaia broad band photometry. *A&A*, 523: A48, November 2010. doi: 10.1051/0004-6361/201015441.
- E.L. Fitzpatrick and D. Massa. An Analysis of the Shapes of Interstellar Extinction Curves. V. The IR-through-UV Curve Morphology. *Astrophysical Journal*, 663 (1):320–341, July 2007. doi: 10.1086/518158.
- Hyeonoh Hur, Hwankyung Sung, and Michael S. Bessell. Distance and the Initial Mass Function of Young Open Clusters in the η Carina Nebula: Tr 14 and Tr 16. *Astronomical Journal*, 143(2):41, February 2012. doi: 10.1088/0004-6256/143/2/41.
- Brajesh Kumar, Saurabh Sharma, Jean Manfroid, Eric Gosset, Gregor Rauw, Yaël Nazé, and Ram Kesh Yadav. Investigation of the stellar content in the western part of the Carina nebula. *A&A*, 567:A109, July 2014. doi: 10.1051/0004-6361/201323027.
- M. J. Reid, K. M. Menten, A. Brunthaler, X. W. Zheng, T. M. Dame, Y. Xu, J. Li, N. Sakai, Y. Wu, K. Immer, B. Zhang, A. Sanna, L. Moscadelli, K. L. J. Rygl, A. Bartkiewicz, B. Hu, L. H. Quiroga-Nuñez, and H. J. van Langevelde. Trigonometric Parallaxes of High-mass Star-forming Regions: Our View of the Milky Way. *Astrophysical Journal*, 885(2):131, November 2019. doi: 10.3847/1538-4357/ab4a11.
- W. W. Morgan, A. E. Whitford, and A. D. Code. Studies in Galactic Structure. I. a Preliminary Determination of the Space Distribution of the Blue Giants. *Astrophysical Journal*, 118:318, September 1953. doi: 10.1086/145754.
- W. Becker and R. B. Fenkart. Galactic Clusters and H II Regions. 38:205, January 1970.
- Wilton S. Dias and J. R. D. Lépine. Direct Determination of the Spiral Pattern Rotation Speed of the Galaxy. *Astrophysical Journal*, 629(2):825–831, August 2005. doi: 10.1086/431456.
- Devesh P. Sariya, Ing-Guey Jiang, D. Bisht, R. K. S. Yadav, and G. Rangwal. Astrometric and Photometric Investigation of Three Old Age Open Clusters in the Gaia Era: Berkeley 32, Berkeley 98, and King 23. *Astronomical Journal*, 161 (3):102, March 2021. doi: 10.3847/1538-3881/abd31f.

INSTITUTO DE FISICA
UNIVERSIDAD DE ANTIOQUIA
Medellín, Colombia
www.udea.edu.co



UNIVERSIDAD DE ANTIOQUIA

1 8 0 3