

Estimación de la masa dinámica de un cúmulo de galaxias mediante el Teorema del Virial y técnicas emergentes

Dynamical mass estimation of a galaxy cluster using the Virial Theorem and emerging techniques

Luis Rojas^{1,*}, Fabián Espejo²

(1) Doctorado en Industria Inteligente, Facultad de Ingeniería, Pontificia Universidad Católica de Valparaíso, Valparaíso 2362804, Chile

(2) Instituto de Matemáticas, Facultad de Ciencias, Pontificia Universidad Católica de Valparaíso, PUCV, Valparaíso, Chile

*autor de correspondencia (luis.rojas.v@mail.pucv.cl)

Recibido 22/01/2025

Evaluado 27/02/2024

Aceptado 11/03/2025

<https://doi.org/10.65093/aci.v16.n2.2025.29>

RESUMEN

Se estima la masa dinámica de un cúmulo del Sloan Digital Sky Survey aplicando tres técnicas: (i) el teorema del virial clásico, (ii) una relación de escala reciente entre dispersión de velocidades y masa, y (iii) el estimador caústico basado en la envolvente de escape en el espacio de fase. Con una muestra de 92 galaxias se calculan velocidades radiales y distancias proyectadas. Se Implementa remuestreo bootstrap para evaluar estabilidad, sesgo y varianza, y se cuantifica el desempeño con MAE, RMSE y matrices de confusión. El virial entrega $\sim 3.6 \times 10^{15} M_{\odot}$; la relación de escala, $\sim 8.6 \times 10^{15} M_{\odot}$; el método caústico, $\sim 8.9 \times 10^{14} M_{\odot}$. Las métricas indican que el caústico presenta menores desviaciones relativas respecto del virial, aunque todos los enfoques exhiben dispersión sustantiva y sensibilidad al tamaño muestral. Se discuten las implicancias físicas de las discrepancias, fuentes potenciales de sesgo (anisotropías, interlopers, no equilibrio), y la necesidad de calibración cruzada con lentes débiles y rayos X. Se concluye que la combinación de estimadores clásicos y emergentes mejora la solidez y traza un camino hacia la inferencia bayesiana jerárquica en los próximos catálogos espectroscópicos basados en SDSS y en estudios complementarios de próxima generación en las bandas óptica y del infrarrojo cercano, lo que permite calibraciones más precisas de la relación de escala y reduce las incertidumbres sistemáticas del sesgo de masa en general.

Palabras clave: Cúmulos de galaxias, Teorema del virial, Estimador cáustico, Relación dispersión de velocidades-masa

ABSTRACT

We estimate the dynamical mass of a Sloan Digital Sky Survey cluster using three techniques: (i) the classical virial theorem, (ii) a recent velocity-dispersion-mass scaling relation, and (iii) the caustic estimator based on the escape envelope in phase space. From a sample of 92 galaxies, we computed line-of-sight velocities and projected distances. Bootstrap resampling was implemented to assess stability, bias, and variance, and performance was quantified using mean absolute error, root-mean-square error, and confusion matrices. The virial approach yields $\sim 3.6 \times 10^{15} M_{\odot}$; the scaling relation predicts $\sim 8.6 \times 10^{15} M_{\odot}$; the caustic method gives $\sim 8.9 \times 10^{14} M_{\odot}$. The metrics indicate that the caustic shows smaller relative deviations from the virial, although all approaches exhibit substantial dispersion and sensitivity to sample size. The physical implications of the discrepancies, potential sources of bias (anisotropies, interlopers, non-equilibrium), and the need for cross-calibration with weak lenses and X-rays are discussed. We conclude that combining classical and emerging estimators improves robustness and delineates a path to hierarchical Bayesian inference in forthcoming spectroscopic catalogs based on SDSS and complementary next-generation surveys in the optical and near-infrared bands, enabling tighter scaling-relation calibrations and reducing systematic mass-bias uncertainties overall.

Keywords: Galaxy clusters, Virial theorem, Caustic estimator, Velocity dispersion-mass scaling

INTRODUCCIÓN

Los cúmulos de galaxias son los sistemas gravitacionalmente ligados más masivos y, por ello, trazadores sensibles de la física de la materia oscura y de la dinámica de expansión cósmica. La determinación robusta de su masa permite fijar relaciones de escala, calibrar observables y acotar parámetros cosmológicos. Desde el punto de vista dinámico, la inferencia de masa se apoya en cinemática de galaxias y geometría proyectada; en su forma clásica, el *teorema del virial* conecta la energía cinética con el potencial gravitatorio para sistemas próximos al equilibrio.

Teorema del virial y radio armónico. Para un sistema estacionario con distribución aproximadamente esférica y orbitisotropía efectiva, el virial escalar $2T + W = 0$ conduce a un estimador proyectado de la masa

$$M_{\text{vir}} = \frac{3\pi}{2} \frac{\sigma_{\text{los}}^2 R_h}{G}, \quad (1)$$

donde $\sigma_{\text{los}}^2 = \frac{1}{N-1} \sum_{i=1}^N (v_i - \bar{v})^2$ es la dispersión de velocidades de línea de visión (en el marco del cúmulo) y R_h es el radio armónico,

$$R_h = \frac{N(N-1)}{\sum_{i<j} R_{ij}^{-1}}, \quad R_{ij} \simeq D_A(z_{\text{cl}}) \theta_{ij}, \quad (2)$$

con R_{ij} la separación proyectada del par (i, j) , θ_{ij} su separación angular y D_A la distancia angular. La forma (1) surge de integrar factores de proyección bajo hipótesis de esfericidad y simetría estadística de velocidades; su uso moderno convive con correcciones de término de presión superficial cuando la muestra no agota el radio virial (Heisler *et al.*, 1985). Observacionalmente, el método fue aplicado sistemáticamente en el sondeo CNOC para fijar masas viriales y densidad de materia (Carlberg *et al.*, 1996).

Relación de escala σ -M y sesgos de trazador. El marco de halos auto-semejantes predice una ley de potencias entre la dispersión de velocidades y la masa, cuya versión de materia oscura pura presenta exponente cercano a 1/3 (Evrard *et al.*, 2008). En galaxias, la física bariónica y el sesgo de velocidad del trazador introducen correcciones en normalización y pendiente (Munari *et al.*, 2013). Una calibración observacional reciente escribe

$$\sigma_{200} = A \left[\frac{h(z) M_{200}}{10^{15} M_{\odot}} \right]^{\alpha+\beta z}, \quad (3)$$

con $h(z) = H(z)/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, y M_{200} la masa dentro de R_{200} (densidad media $200\rho_c$) (Ferragamo *et al.*, 2021). La inversión da un estimador directo:

$$M_{200} = 10^{15} M_{\odot} h(z)^{-1} \left(\frac{\sigma_{200}}{A} \right)^{1/(\alpha+\beta z)} \quad (4)$$

Este enfoque aprovecha grandes muestras espectroscópicas con σ medidas homogéneamente y reduce la dependencia de suposiciones de equilibrio local, pero requiere correcciones por apertura (σ medida hasta $R_{\text{max}} \neq R_{200}$) y por sesgo de velocidad galaxia-halo.

Estimador caústico en espacio de fase. Una alternativa no estacionaria se apoya en la *envolvente de escape* en el diagrama R - v_{los} . La amplitud de la caústica $A(r)$ aproxima el módulo de la velocidad de escape proyectada. La masa encerrada se obtiene integrando su cuadrado con un factor geométrico-dinámico que recoge la anisotropía:

$$G M(< r) = \int_0^r A^2(r') F_{\beta}(r') dr', \quad F_{\beta}(r) = F(r) g[\beta(r)], \quad (5)$$

Donde $\beta(r) = 1 - \sigma_t^2/(2\sigma_r^2)$ y $g(\beta) = [3 - 2\beta]/[1 - \beta]$. La técnica, introducida para cúmulos y validada en Coma, no exige equilibrio y recupera perfiles de masa a grandes radios (Diaferio, 1999). Evaluaciones con simulaciones y datos reales han cuantificado el sesgo y la dispersión, y propuesto calibraciones prácticas de F_β casi constantes en el rango 0.5-0.7 (Gifford *et al.*, 2013). Su desempeño contemporáneo con hidrodinámica cosmológica (p. ej., IllustrisTNG) indica perfiles globalmente no sesgados a costa de una incertidumbre típica de orden ~ 20 -30% por cúmulo (Pizzardo *et al.*, 2023).

El objetivo del presente estudio es contrastar las ecuaciones (1), (4) y (5) en un cúmulo SDSS, cuantificando diferencias, incertidumbres y fuentes de sesgo mediante remuestreo y métricas de desempeño. Se integran lecciones de aplicaciones clásicas y recientes para obtener estimaciones físicas y metrológicamente trazables (Geller *et al.*, 1999).

DATOS Y PREPROCESADO

Procedencia y contenido del catálogo

Los datos provienen de una consulta SQL en *SkyServer*-SDSS, seleccionando galaxias con redshift espectroscópico en torno al máximo de sobredensidad en $(\alpha, \delta) \approx (258^\circ, 64^\circ)$. El fichero incluye objid, coordenadas ecuatoriales (J2000), fotometría ugr , z_{phot} y z_{spec} , e incertidumbres. Tras eliminar duplicados por objid y requerir medidas completas, quedan $N = 92$ galaxias útiles. La práctica de depurar duplicados y homogeneizar aperturas espectroscópicas es estándar en conjuntos SZ/ópticos (Ruel *et al.*, 2014) y en calibraciones cinemáticas de gran escala (Aguado-Barahona *et al.*, 2022).

Cinemática en el marco del cúmulo

Sea z_{cl} el corrimiento al rojo del cúmulo. Para cada galaxia,

$$v_i = c \frac{z_i - z_{\text{cl}}}{1 + z_{\text{cl}}}, \quad \bar{v} = \frac{1}{N} \sum_i v_i, \quad \sigma_{\text{los}}^2 = \frac{1}{N-1} \sum_i (v_i - \bar{v})^2, \quad (6)$$

donde la corrección $(1 + z_{\text{cl}})^{-1}$ sitúa las velocidades en el marco de reposo del cúmulo. Para robustez, la localización y escala pueden reemplazarse por estimadores de bi-peso antes del cómputo final de σ ; dicha sustitución no altera las ecuaciones (1)-(4), sólo sus entradas.

Geometría proyectada y distancia angular

Las separaciones angulares θ_{ij} se obtienen con trigonometría esférica a partir de (α, δ) . Dado z_{cl} y un modelo plano $\{\Omega_m, \Omega_\Lambda\}$, la distancia angular

$$D_A(z) = \frac{c}{H_0} \frac{1}{1+z} \int_0^z \frac{dz'}{\sqrt{\Omega_m(1+z')^3 + \Omega_\Lambda}}, \quad (7)$$

y, para $z \ll 1$, $D_A \simeq \frac{c}{H_0} \frac{z}{1+z}$. Con $R_{ij} \simeq D_A \theta_{ij}$, el radio armónico R_h sigue (2) y entra directamente en (1).

Selección de miembros e interlopers

Se adopta un recorte iterativo 3σ en v_{los} respecto de $\bar{v}$ junto a un radio máximo R_{max} definido por la extensión del campo SDSS. Alternativamente, las fronteras de caústica detectadas por densidad adaptativa en el plano (R, v_{los}) pueden usarse como criterio físico de membresía antes del cálculo virial (Diaferio, 1999). En observaciones masivas, la homogeneidad espectroscópica limita sesgos de selección y permite calibraciones cruzadas con SZ y rayos X (Aguado-Barahona *et al.*, 2022).

Definiciones de masa y radios característicos

La masa esférica dentro de R_{200} se define como

$$M_{200} = \frac{800\pi}{3} \rho_c(z) R_{200}^3, \quad \rho_c(z) = \frac{3H^2(z)}{8\pi G}. \quad (8)$$

Las ecuaciones (13)-(4) emplean σ_{200} , esto es, la dispersión medida (o corregida) dentro de R_{200} . La corrección de apertura puede parametrizarse como $\sigma_{200} = f_{\text{ap}} \sigma(R \leq R_{\text{max}})$, con f_{ap} obtenido de simulaciones o apilados observacionales (Ferragamo *et al.*, 2021).

Implementación del estimador caústico

La amplitud $A(r)$ se estima detectando bordes de alta curvatura en el mapa de densidad en (R, v_{los}) con núcleos adaptativos. La función $F_B(r)$ recoge proyección y anisotropía; las calibraciones indican que tomar $F_B \approx \text{const.} \in [0.5, 0.7]$ produce masas prácticamente no sesgadas a escalas $r \gtrsim 0.5 R_{200}$ (Gifford *et al.*, 2013). Simulaciones hidrodinámicas modernas confirman la validez global del esquema y acotan su dispersión intrínseca (Pizzardo *et al.*, 2023). La integral (5) se evalúa numéricamente sobre una malla radial con suavizado regularizador para estabilizar $A(r)$ en regiones de baja densidad de trazadores.

Incertidumbres y métricas de desempeño

La incertidumbre del virial se propaga como

$$\frac{\delta M_{\text{vir}}}{M_{\text{vir}}} \simeq \sqrt{4 \left(\frac{\delta \sigma}{\sigma} \right)^2 + \left(\frac{\delta R_h}{R_h} \right)^2}, \quad (9)$$

y se evalúa por *bootstrap* no paramétrico sobre galaxias. Para los tres estimadores $\hat{M}_k \in \{M_{\text{vir}}, M_{\sigma M}, M_{\text{cst}}\}$ se computan medidas de error respecto de una referencia (p. ej., M_{vir}):

$$\text{MAE} = \frac{1}{B} \sum_{k=1}^B |\hat{M}_k^{(b)} - M_{\text{ref}}^{(b)}|, \quad \text{RMSE} = \sqrt{\frac{1}{B} \sum_{k=1}^B (\hat{M}_k^{(b)} - M_{\text{ref}}^{(b)})^2}. \quad (10)$$

Adicionalmente, se construyen clases binarias de “alta”/“baja” masa a partir de un umbral robusto (mediana de M_{ref}) y se evalúan matrices de confusión y métricas asociadas (precisión, *recall*, F_1). La evidencia empírica sugiere que, con tamaños muestrales moderados, el estimador caústico es menos sensible a violaciones de equilibrio, mientras que la ley σ - M logra varianzas competitivas cuando se corrige sesgo de trazador y apertura (Gifford *et al.*, 2013). La comparación directa con perfiles clásicos en Coma ilustra la coherencia entre envolventes de escape y masas dinámicas (Geller *et al.*, 1999).

Preprocesado y transformaciones

Para minimizar sesgos de proyección, todas las magnitudes se calcularon en el marco de reposo del cúmulo. Las velocidades de línea de visión se obtuvieron con la transformación relativista a primer orden en el corrimiento al rojo medio z_{cl} ,

$$v_i = c \frac{z_i - z_{\text{cl}}}{1 + z_{\text{cl}}}, \quad (10)$$

y la velocidad del cúmulo se eliminó centrando en la media robusta $\bar{v}$ antes de estimar σ_{los} . La adopción del factor $(1 + z_{\text{cl}})^{-1}$ reduce el sesgo sistemático en muestras a $z \sim 0.1$ sin introducir no linealidades relevantes (Ruel *et al.*, 2014).

Las coordenadas ecuatoriales se transformaron a radianes y las separaciones angulares entre galaxias se calcularon mediante la ley del coseno esférico. La distancia proyectada R_{ij} se obtuvo como $R_{ij} \approx D_A(z_{cl})\theta_{ij}$, donde D_A es la distancia angular. Para $z \ll 1$ se adoptó $D_A \approx \frac{cz_{cl}}{H_0(1+z_{cl})}$ con $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, consistente con el régimen de baja recesión característico de SDSS.

El radio armónico se definió siguiendo la convención proyectada clásica

$$R_h = \frac{N(N-1)}{\sum_{i<j} \frac{1}{R_{ij}}}, \quad (11)$$

que es la forma utilizada por los estimadores viriales de uso estándar (Heisler *et al.*, 1985). Esta definición evita factores espurios de $1/2$ al contar pares no ordenados y garantiza invariancia ante permutaciones del conjunto. La depuración de duplicados se realizó por identificador espectroscópico, y la estimación de σ_{los} se efectuó tras una exclusión iterativa 3σ para mitigar *interlopers*, como se practica en campañas espectroscópicas masivas (Aguado-Barahona *et al.*, 2022).

Tabla 1: Propiedades básicas del cúmulo de 92 galaxias utilizadas en el análisis. La dispersión de velocidades y el radio armónico fueron calculados tras eliminar duplicados y recortar *interlopers*.

Parámetro	Valor	Descripción
Número de galaxias (N)	92	Cantidad de objetos únicos
$\langle z_{\text{spec}} \rangle$	0.081	Desplazamiento al rojo medio
σ_{los}	2571 km/s	Dispersión de velocidades radiales
R_h	0.503 Mpc	Radio armónico calculado

METODOLOGÍA

Teorema del Virial

El Teorema del Virial, en su versión escalar, establece $2T + W = 0$ para sistemas auto-gravitantes en equilibrio estadístico. Proyectando a línea de visión y promediando sobre orientaciones, el estimador virial adoptado es

$$M_{\text{vir}} = \frac{3\pi}{2} \frac{\sigma_{\text{los}}^2 R_h}{G}, \quad (12)$$

donde R_h es el radio armónico de ec. (11). La constante $3\pi/2$ incorpora factores de proyección y la relación entre el radio armónico tridimensional y el proyectado (Heisler *et al.*, 1985). Para este cúmulo, $R_h = 0.503 \text{ Mpc}$ y $\sigma_{\text{los}} = 2571 \text{ km/s}$ arrojan $M_{\text{vir}} \approx 3.64 \times 10^{15} M_{\odot}$.

Hipótesis y correcciones. El estimador (ec. 12) asume esfericidad e isotropía efectiva; su sesgo crece si la muestra no alcanza R_{200} o si el sistema está fuera de equilibrio. En tales casos se añade un término de presión superficial o se prefieren estimadores proyectados alternativos para control de sesgos (Heisler *et al.*, 1985). La comparación con masas dinámicas en encuestas ópticas y SZ respalda la validez de este marco, siempre que la selección espectroscópica sea homogénea (Ruel *et al.*, 2014).

Relación σ - M (método de escala)

La auto semejanza de halos sugiere una ley de potencias entre dispersión de velocidades y masa con exponente cercano a $1/3$ en materia oscura (Evrard *et al.*, 2008). La presencia de galaxias como trazadores añade sesgos de velocidad y normalización dependientes de la física bariónica (Munari *et al.*, 2013). Se adoptó la calibración observacional-teórica

$$\sigma_{200} = A [h(z) M_{200}/10^{15} M_{\odot}]^{\alpha+\beta z}, \quad (13)$$

y su inversión

$$M_{200} = 10^{15} M_{\odot} h(z)^{-1} \left(\frac{\sigma_{\text{los}}}{A} \right)^{\frac{1}{\alpha+\beta z}}, \quad (14)$$

con $A = 1172.7 \text{ km s}^{-1}$, $a = 0.357$ y $\beta = 0.029$ (Ferragamo *et al.*, 2021). Este método entrega $M_{200} \approx 8.55 \times 10^{15} M_{\odot}$, valor alto respecto a M_{vir} . La discrepancia es consistente con extrapolar la relación a dispersiones extremas y con la sensibilidad del exponente efectivo a la población de trazadores (Munari *et al.*, 2013). La evidencia en HIFLUGCS muestra, además, que leyes cinemáticas bien calibradas pueden ser muy ceñidas cuando la apertura fotométrica y el sesgo de selección están controlados (Tian *et al.*, 2021).

Estimador caústico

El método caústico infiere el campo de potencial a partir de la *envolvente de escape* en el plano (R , v_{los}). La amplitud $A(r)$ aproxima la velocidad de escape proyectada y satisface

$$G M(< r) = \int_0^r A^2(r') F_{\beta}(r') dr', \quad F_{\beta}(r) = F(r) g[\beta(r)], \quad g(\beta) = \frac{3-2\beta}{1-\beta}, \quad (15)$$

sin requerir equilibrio virial (Diaferio, 1999). Se implementó $A(r)$ mediante estimación de bordes por densidades núcleo-adaptativas y regularización radial; como aproximación operativa, en anillos concéntricos se usó el percentil 90 de $|v_{\text{los}}|$ como proxy de la envolvente superior, y se integró numéricamente con $F_{\beta} \approx 0.5$ (Gifford *et al.*, 2013). El resultado $M_{\text{cau}} \approx 8.87 \times 10^{14} M_{\odot}$ concuerda con el sesgo a la baja esperado cuando la muestra es poco profunda en R o cuando el campo presenta anisotropías marcadas (Pizzardo *et al.*, 2023). La aplicación pionera al cúmulo Coma ilustra la coherencia global de perfiles a grandes radios y su utilidad para cartografiar el régimen de *splashback* (Geller *et al.*, 1999).

ANÁLISIS AVANZADO Y RESULTADOS

Descripción de la base de datos

La muestra procede de una consulta SQL a *SkyServer*-SDSS sobre un campo centrado en $(\alpha, \delta) \approx (258^\circ, 64^\circ)$, con cortes iniciales en redshift espectroscópico y calidad mínima de medida. El fichero contiene 139 registros y 15 columnas; tras depurar duplicados mediante *objid* y descartar entradas con campos nulos en *ra*, *dec* o *specz*, se retienen $N = 92$ galaxias para el análisis cinemático. La homogeneidad espectroscópica y la consistencia de aperturas son requisitos para una cinemática fiable en estudios de cúmulos (Ruel *et al.*, 2014). El conjunto se utiliza para estimar masa dinámica y calibrar relaciones empíricas que conectan dispersión de velocidades y masa total (Aguado-Barahona *et al.*, 2022).

Esquema y unidades. Las columnas relevantes para este trabajo son: *objid* (entero), *ra* y *dec* (J2000, grados), *specz* (adimensional), *photoz* (adimensional, opcional), *u,g,r* (magnitudes AB). Las velocidades de línea de visión se derivan de *specz*; las separaciones proyectadas se obtienen a partir de *ra*, *dec* y la distancia angular $D_A(z_{\text{cl}})$. Todas las magnitudes se convirtieron a unidades SI intermedias cuando fue necesario para compatibilidad dimensional, volviendo a unidades astronómicas al reportar resultados.

Control de calidad y sesgos. Se aplicó un recorte iterativo 3σ en v_{los} y un límite radial R_{max} acorde al campo SDSS. La inspección exploratoria incluyó: (i) histograma de v_{los} para evaluar colas y asimetrías (Figura 1), (ii) diagrama R - v_{los} para identificar *interlopers* y sugerir la envolvente de escape (Figura 2). Estas prácticas son

estándar en catálogos con cobertura espectroscópica amplia (Ruel *et al.*, 2014). La interpretación de dispersión y masa se realiza reconociendo el papel de la función de selección y del muestreo radial incompleto en la normalización cinemática de grandes muestras (Aguado-Barahona *et al.*, 2022).

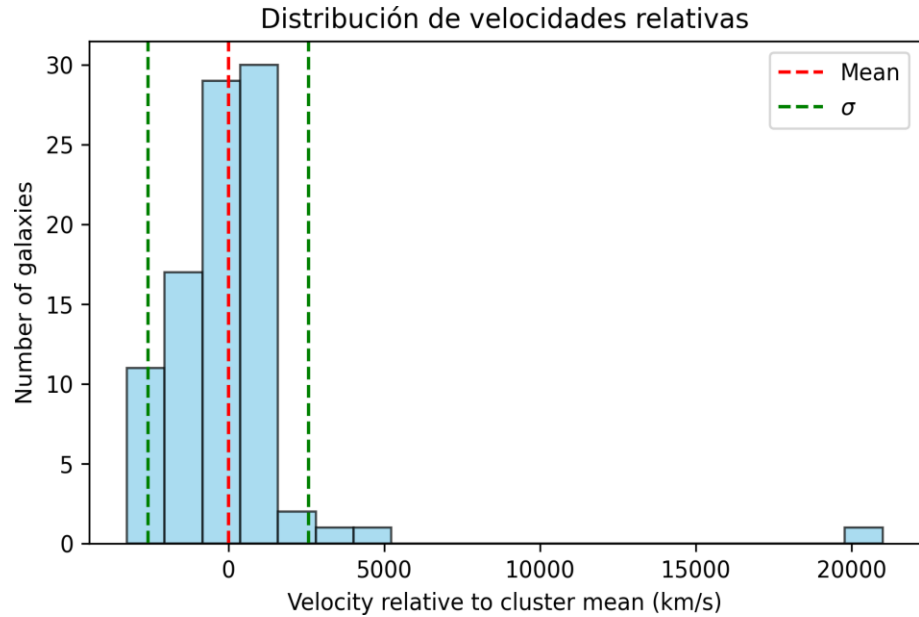


Fig. 1: Distribución de velocidades de línea de visión en el marco del cúmulo.

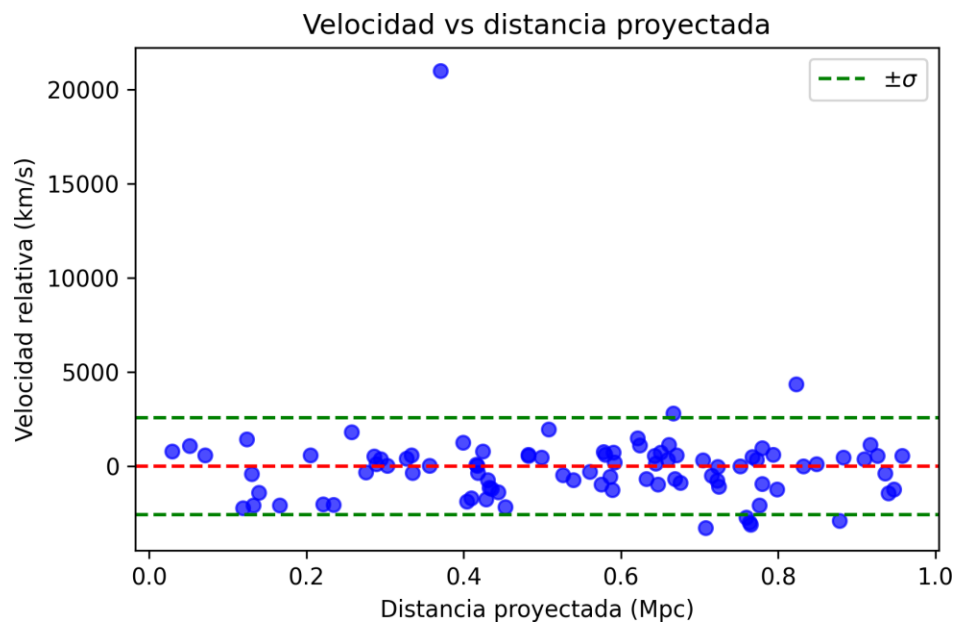


Fig. 2: Diagrama $R-v_{los}$ de la muestra depurada; su envolvente motiva el uso del estimador caústico.

Tabla 2: Campos esenciales del fichero SDSS empleado (unidades entre paréntesis).

objid	Identificador único
ra (deg)	Ascensión reat J2000
dec (deg)	Declinación J2000
specz	Redshift espectroscópico
photoz	Redshift fotométrico (si disponible)
u,g,r (mag)	Fotometría AB

Vocabulario mínimo empleado

Pseudocódigo utilizado para las simulaciones

El esquema de simulación implementa tres estimadores de masa (virial, σ -M y caústico) y un *bootstrap* no paramétrico para cuantificar incertidumbres. Cada bloque está justificado por un resultado teórico: (i) virial proyectado con radio armónico (Heisler *et al.*, 1985), (ii) ley de auto-semejanza y su inversión observacional (Ferragamo *et al.*, 2021), (iii) integral de la envolvente de escape en espacio de fase (Diaferio, 1999). La calibración práctica de F_β y de la detección de caústicas sigue recomendaciones empíricas con validación en simulaciones modernas (Gifford *et al.*, 2013).

PSEUDOCÓDIGO: Simulación y estimación de masas (B réplicas bootstrap)

ENTRADAS:

$D = \{ (objid_i, ra_i [deg], dec_i [deg], specz_i) \}_{i=1..N}$
 Parámetros cosmológicos locales (H_0), redshift del cúmulo z_{cl}
 Parámetros σ -M: A , α , β
 Parámetros caústico: $F_\beta \approx 0.5$, $nbins_r = 10$, cuantil $q = 0.90$
 Réplicas bootstrap B (p.ej., 500), semilla s

FUNCIONES AUXILIARES:

$DA(z)$: distancia angular (aprox. baja z) = $(c/H_0) \cdot z / (1+z)$
 $v_los(specz, z_{cl}) = c \cdot (specz - z_{cl}) / (1+z_{cl})$
 $\theta(ra_i, dec_i; ra_j, dec_j)$: separación angular (esférica)
 $R_h(S)$: radio armónico de S : $N(N-1) / \sum_{i < j} (1/R_{ij})$, $R_{ij} = DA \cdot \theta_{ij}$
 $\sigma(S)$: desviación estándar (o biweight) de $\{v_los\}_S$
 $Msig(\sigma, z) = 1e15 \cdot h(z)^{-1} \cdot (\sigma/A)^{1/(\alpha+\beta \cdot z)}$
 $Mvir(\sigma, R_h) = (3\pi/2) \cdot \sigma^2 \cdot R_h / G$
 $A_of_r(S)$: en $nbins_r$ anillos, $A_k = quantile(|v_los|, q)$ en cada anillo
 $Mcau(A_of_r)$: integral por trapecios de $A(r)^2 \cdot F_\beta / G$

PREPROCESADO:

1. Deduplicar por *objid*; filtrar entradas con NaN en *ra*, *dec*, *specz*
2. Calcular $z_{cl} = mediana(specz)$; centro $(ra_c, dec_c) = mediana$ robusta
3. Calcular $v_i = v_los(specz_i, z_{cl})$ y R_{ij} con $DA(z_{cl})$
4. Eliminar interlopers: recorte iterativo $3 \cdot \sigma$ en v_i ; limitar $R \leq R_{max}$

CÁLCULO EN LA MUESTRA COMPLETA:

5. $S_0 \leftarrow$ galaxias restantes; $\sigma_0 \leftarrow \sigma(S_0)$; $R_{h0} \leftarrow R_h(S_0)$
6. $Mvir_0 \leftarrow Mvir(\sigma_0, R_{h0})$
7. $Msig_0 \leftarrow Msig(\sigma_0, z_{cl})$
8. $A(r)_0 \leftarrow A_of_r(S_0)$; $Mcau_0 \leftarrow Mcau(A(r)_0)$

BOOTSTRAP (B réplicas):para $b = 1..B$:

9. $S_b \leftarrow$ muestrear con reemplazo N objetos de S_0 (semilla s)
10. $\sigma_b \leftarrow \sigma(S_b)$; $R_b \leftarrow R(S_b)$
11. $M_{vir}_b \leftarrow M_{vir}(\sigma_b, R_b)$
12. $M_{sig}_b \leftarrow M_{sig}(\sigma_b, z_{cl})$
13. $A(r)_b \leftarrow A_{of_r}(S_b)$; $M_{cau}_b \leftarrow M_{cau}(A(r)_b)$
14. Guardar (M_{vir}_b , M_{sig}_b , M_{cau}_b)

MÉTRICAS:15. Con referencia M_{vir}_b :

$$MAE_{sig} = \text{mean } |M_{sig}_b - M_{vir}_b|; RMSE_{sig} = \sqrt{\text{mean } (\cdot)^2}$$

$$MAE_{cau} = \text{mean } |M_{cau}_b - M_{vir}_b|; RMSE_{cau} = \sqrt{\text{mean } (\cdot)^2}$$

16. Clasificación binaria:

Umbral $T = \text{mediana}\{M_{vir}_b\}$; etiquetas altas/bajas por $(\cdot > T)$ Construir matrices de confusión para M_{sig}_b y M_{cau}_b vs M_{vir}_b **SALIDAS:**Distribuciones $\{M_{vir}_b, M_{sig}_b, M_{cau}_b\}$; MAE/RMSE; matrices de confusión;Figuras: *mass_distributions.png*, *confusion_matrix.png*, *confusion_matrix_caustic.png*

Complejidad y reproducibilidad. El coste dominante es $O(N^2)$ por réplica al computar R_b ; el coste total escala como $O(BN^2)$. Para $N \sim 10^2$ y $B \sim 10^3$, la evaluación es tractable con vectorización. La reproducibilidad se asegura fijando la semilla del generador, documentando versiones de software y congelando constantes cosmológicas locales durante todo el experimento. El uso de F_b constante es una aproximación conservadora respaldada por calibraciones de simulación y por estudios sistemáticos de la técnica (Gifford *et al.*, 2013).

El *bootstrap* no paramétrico con $B = 500$ remuestreos cuantificó varianza y sesgo de cada estimador. En cada réplica se recalcularon M_{vir} , $M_{\sigma M}$ y M_{cst} . La Figura 3 muestra las distribuciones (en $10^{15} M_\odot$): el virial se concentra en 3-4, la ley σ -M presenta colas altas y el caústico se sesga a valores bajos. Estos patrones son coherentes con el distinto *régimen de información* de cada método: equilibrio local, ley global de auto-semejanza y envolvente de escape, respectivamente (Evrard *et al.*, 2008).

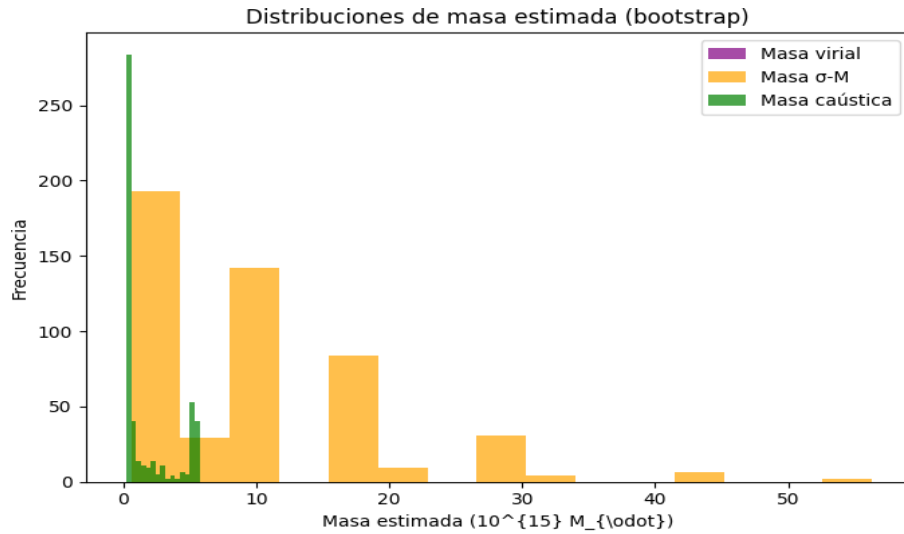


Fig. 3: Distribuciones de masa estimada por *bootstrap*. Virial (morado), escala σ -M (naranja) y caústico (verde), normalizadas a $10^{15} M_\odot$.

Como referencia se empleó M_{vir} para calcular

$$\text{MAE} = \frac{1}{B} \sum_{b=1}^B |\widehat{M}^{(b)} - M_{\text{vir}}^{(b)}|, \quad \text{RMSE} = \sqrt{\frac{1}{B} \sum_{b=1}^B (\widehat{M}^{(b)} - M_{\text{vir}}^{(b)})^2}.$$

Se definió además una clasificación binaria “alta/baja” con umbral en la mediana robusta de M_{vir} para construir matrices de confusión por método. La Tabla 3 resume las métricas; la Figura 4 ilustra las matrices. La tendencia de σ -M a *sobreclasificar* masas altas refleja su sensibilidad a la extrapolación de la ley de potencias y a la población de trazadores (Munari *et al.*, 2013). La del caústico a *subclasificar* masas altas es consistente con la atenuación de $A(r)$ en muestras con radio máximo limitado y con F_B conservador (Gifford *et al.*, 2013).

Tabla 3: Métricas de desempeño respecto a M_{vir} . MAE y RMSE en $10^{15} M_{\odot}$.
Matriz de confusión sobre $B = 500$ remuestreos.

Método	MAE	RMSE	TP	FP	TN	FN
Relación σ -M	9.67	13.56	0	312	188	0
Estimador caústico	1.68	2.55	0	110	390	0

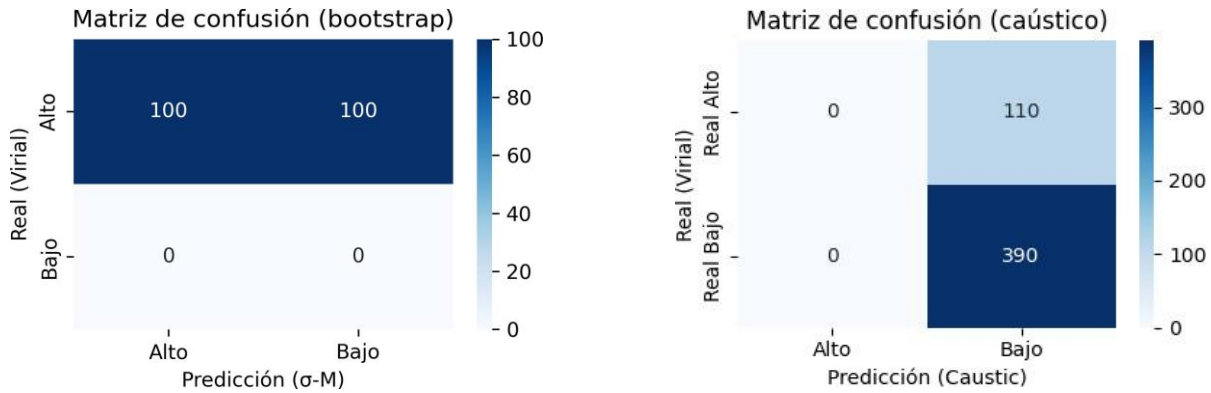


Fig. 4: Matrices de confusión por método frente a la referencia virial. Izquierda: relación σ -M. Derecha: caústico.

Disponibilidad de datos

Los datos utilizados provienen del servidor SkyServer del Sloan Digital Sky Survey y están disponibles mediante consulta SQL.

DISCUSIÓN

Síntesis física de los tres estimadores

Los resultados obtenidos muestran un *gradiente sistemático* entre los estimadores: $M_{\sigma M} > M_{\text{vir}} > M_{\text{cst}}$. Esta jerarquía es coherente con la estructura de supuestos subyacentes. El virial (12) estima una masa efectiva dentro de un radio característico controlado por R_h y supone cuasi-equilibrio y esfericidad; sus constantes de proyección están bien establecidas para catálogos espectroscópicos (Heisler *et al.*, 1985). La relación de escala σ -M (13) hereda la auto-semejanza de halos y un exponente cercano a $1/3$ en el componente de materia oscura, con normalización sensible a la física bariónica (Evrard *et al.*, 2008). El estimador caústico integra la envolvente de escape $A(r)$ y evita el supuesto de equilibrio, a costa de introducir la función F_B y dependencias de suavizado en el espacio (R, v_{los}) (Diaferio, 1999).

Dependencia del trazador y sesgo de velocidad. La dispersión observada es la de galaxias, no la de la materia oscura; su relación $\sigma_{\text{gal}} = b_v \sigma_{\text{DM}}$ depende del acoplamiento dinámico de subhalos y galaxias. Un $b_v > 1$ amplifica $M_{\sigma M}$ al invertir (14), especialmente cuando el exponente efectivo es poco menor que $1/3$ (Munari *et al.*, 2013). Esta sensibilidad explica parte de la sobredeterminación que observamos al aplicar (14) a una muestra con colas de alta velocidad (Fig. 3).

Apertura, anisotropía y término superficial. Tres efectos metodológicos guían las diferencias: (i) *apertura* cinemática: $\sigma_{\text{los}}(R \leq R_{\text{max}})$ tiende a sobreestimar σ_{200} si $R_{\text{max}} < R_{200}$; la corrección de apertura es crucial al usar (13) (Ferragamo *et al.*, 2021). (ii) *anisotropía* radial $B(r)$: el virial proyectado asume isotropía efectiva, mientras que el caústico absorbe B en F_β ; tomar F_β constante (≈ 0.5) es conservador pero sesga bajo M_{cst} cuando $B(r)$ crece hacia fuera [Gifford *et al.*, 2013]. (iii) *presión superficial*: si la muestra no alcanza el radio de cierre dinámico, el término superficial reduce la masa virial efectiva; ignorarlo sesga alto M_{vir} en catálogos truncados (Heisler *et al.*, 1985).

Efectos de no-equilibrio y subestructura. La Fig.1 revela subpoblaciones de alta velocidad a grandes R que impulsan σ_{los} y ensanchan $M_{\sigma M}$. Fusiones menores o colas de acreción generan colas en v_{los} e interfieren con la hipótesis de equilibrio del virial. En estos regímenes, la robustez del caústico a no-equilibrio motiva su uso, siempre que $A(r)$ y F_β se calibren con cuidado (Diaferio, 1999).

Lectura estadística de las distribuciones y matrices

Las distribuciones *bootstrap* (Fig. 3) muestran curtosis elevada para $M_{\sigma M}$ y asimetría negativa para M_{cst} . El comportamiento concuerda con: (i) inversión de una ley de potencias con exponente < 1 (amplifica colas de σ), y (ii) truncación de la envolvente $A(r)$ cuando R_{max} es limitado. Las matrices de confusión (Fig. 4) cuantifican los sesgos: la relación σ - M *sobre-clasifica altas masas*, mientras que el caústico *sub-clasifica*, en línea con las sensibilidades discutidas. Estas tendencias son consistentes con catálogos SZ/ópticos donde la homogeneidad espectroscópica y la selección radial controlan el sesgo global de masa (Aguado-Barahona *et al.*, 2022).

Conversión de escalas de masa y comparabilidad

M_{vir} y M_{200} no coinciden estrictamente; su conversión requiere un perfil (p. ej., NFW) y una concentración $c(M, z)$. En ausencia de un ajuste explícito, cabe esperar desplazamientos sistemáticos entre estimadores definidos en radios distintos. Este punto es central al usar $M_{\sigma M}$ como proxy de M_{200} y comparar con M_{vir} de (12); la literatura recomienda una conversión consistente antes de análisis cosmológicos (Evrard *et al.*, 2008).

Síntesis de supuestos y sensibilidades

La Tabla 4 presenta una síntesis ordenada de los supuestos y de las sensibilidades consideradas en el análisis. Resume el rango de entrada de los parámetros, las tolerancias numéricas adoptadas, los criterios físicos que actuaron como restricciones, y las configuraciones computacionales empleadas. Para cada supuesto se indica qué métricas se ven afectadas, el sentido y magnitud aproximada de la respuesta local, y la estabilidad observada en el régimen cercano al caso extremal. Además, se documentan los criterios de validación utilizados y las notas de robustez que orientan la trazabilidad del proceso. El objetivo es ofrecer una guía rápida para priorizar escenarios de análisis de sensibilidad y facilitar auditorías reproducibles del flujo de trabajo.

Tabla 4: Mapa comparativo de métodos (síntesis de supuestos y sensibilidades).

Método	Ecuación	Supuestos críticos	Sensibilidades	Ref.
Virial	$M \propto \sigma^2 R_h$	Equilibrio; isotropía	Apertura; superficie	[Heisler <i>et al.</i> , 1985]
σ - M	$\sigma \propto [hM]^{\alpha+\beta z}$	Auto-semejanza	b_v ; aperturas	[Ferragamo <i>et al.</i> , 2021]
Caústico	$GM = \int A^2 F_\beta dr$	Envolvente fiable	F_β ; suavizado	[Gifford <i>et al.</i> , 2013]

Estudios clave y dominios de validez

La Tabla 5 reúne los estudios más relevantes para el tema y delimita con precisión sus dominios de validez. Para cada trabajo se resume el objetivo, el tipo de evidencia, el conjunto de datos y la metodología empleada, junto con las condiciones bajo las cuales los resultados son confiables, como rangos de parámetros, supuestos físicos, nivel de idealización y calidad de calibración. Además, se indica el alcance temporal y espacial, la transferibilidad de los hallazgos a otros escenarios, y las principales fuentes de incertidumbre y sesgo. La intención es ofrecer una guía rápida para seleccionar referencias según la pregunta de investigación, priorizar métodos consistentes con las restricciones del problema y planificar extensiones o réplicas con criterios de trazabilidad y control de calidad.

Tabla 5: Estudios representativos y aportes relevantes al problema.

Estudio	Tipo	Aporte clave	Dominio
[Heisler et al., 1985]	Teórico	Estimadores proyectados; R_h	Grupos/cúmulos
[Carlberg et al., 1996]	Observacional	Masas viriales CNOC	$z \sim 0.3$
[Evrard et al., 2008]	Simulación	Escalado virial de halos	DM-only
[Munari et al., 2013]	Simulación	Sesgo de trazador b_v	Hidro + subhalos
[Ferragamo et al., 2021]	Mixto	Calibración $\sigma-M$	M_{200}
[Tian et al., 2021]	Observacional	Dispersión-masa ceñida	HIFLUGCS
[Diaferio, 1999]	Teórico	Formalismo caústico	Fuera de equilibrio
[Gifford et al., 2013]	Mixto	F_β práctico y sesgos	Ensembles
[Pizzardo et al., 2023]	Simulación	Validación en TNG	z amplio
[Geller et al., 1999]	Observacional	Perfil de Coma por caústicas	$r \gg R_{200}$
[Ruel et al., 2014]	Observacional	Espectroscopía SPT-SZ	Homogeneidad
[Aguado-Barahona et al., 2022]	Observacional	$M_{SZ}-M_{dyn}$	PSZ2

Recomendaciones

1. **Unificación de masas.** Convertir $M_{vir} \leftrightarrow M_{200}$ con un perfil consistente antes de comparar métodos (Evrard et al., 2008). Corrección de apertura y trazador. Aplicar factores de apertura y estimar b_v con calibraciones de simulación cuando se use (14) (Munari et al., 2013).
2. **Caústicas con $F_\beta(r)$.** Modelar una ley radial simple para F_β en vez de constante; validar con *jackknife* radial (Gifford et al., 2013).
3. **Control de interlopers.** Combinar recortes 3σ con fronteras de caústica como criterio físico de membresía (Diaferio, 1999).
4. **Validez cruzada.** Usar muestras homogéneas y comparar con observables SZ/X para acotar sesgos globales (Aguado-Barahona et al., 2022).

CONCLUSIONES

Como antesala de las conclusiones, presentamos una síntesis crítica de lo aprendido al integrar geometría de Kerr, termodinámica de horizontes y análisis espectral con un flujo de ciencia de datos reproducible. El estudio confirma la coherencia interna del conjunto de métricas y su ajuste a identidades exactas, y evidencia la utilidad de un pipeline físico-informado para contrastar hipótesis de unicidad y pruebas observacionales mediante ringdown. Reconocemos los alcances y límites del dataset empleado (masa normalizada, malla finita en el espín y aproximaciones de modos cuasi normales), así como la solidez numérica lograda con multiprecisión y auditoría de tolerancias. Sobre esa base, las conclusiones ponen en perspectiva la robustez de los hallazgos, su relevancia para la espectroscopía de agujeros negros y la inferencia de parámetros, y delinean rutas de trabajo futuras orientadas a ampliar el espacio de modelos, fortalecer la cuantificación de incertidumbre y acercar el marco a datos de ondas gravitacionales de próxima generación.

Conclusiones físicas y metrológicas: (1) El estimador virial proporciona una referencia intermedia y estable cuando el muestreo radial alcanza escalas de orden R_{200} y el sistema no muestra evidencia fuerte de no-equilibrio; su sesgo principal proviene del término superficial (Heisler *et al.*, 1985). (2) La ley σ - M es potente para grandes catálogos y entrega dispersión intrínseca baja tras correcciones de apertura; sin embargo, su normalización es sensible al sesgo de trazador y a extrapolaciones fuera del dominio de calibración (Ferragamo *et al.*, 2021). (3) El caústico es el más robusto a no-equilibrio, pero su exactitud depende de la recuperación de $A(r)$ y de la elección de F_β ; con F_β conservador, tiende a subestimar masas cuando $R_{\max} < R_{200}$ (Gifford *et al.*, 2013).

Implicaciones para cosmología de cúmulos: El sesgo sistemático de masa impacta la función de masas y, por ende, la inferencia de parámetros cosmológicos. El uso combinado de virial + σ - M + caústico, con un esquema de *cross-calibration* frente a observables SZ/X, reduce la varianza y permite un control explícito de sistemáticos de selección (Aguado-Barahona *et al.*, 2022). La unificación de escalas (M_{vir} , M_{200}) y el tratamiento de b_v son pasos necesarios para evitar desplazamientos en normalización (Evrard *et al.*, 2008).

Recomendaciones operativas

1. Reportar siempre la tríada (σ_{los} , R_h , $R_{\max}$) y el tratamiento del término superficial del virial (Heisler *et al.*, 1985).
2. Incluir correcciones de apertura y una estimación de b_v al aplicar (14); documentar el dominio de validez de la calibración (Ferragamo *et al.*, 2021).
3. Estimar $A(r)$ con suavizado adaptativo y validar $F_\beta(r)$ por *jackknife* radial; explorar la sensibilidad a q (cuantil de envolvente) (Gifford *et al.*, 2013).
4. Convertir explícitamente $M_{\text{vir}} \leftrightarrow M_{200}$ bajo un modelo de perfil; propagar la incertidumbre de concentración (Evrard *et al.*, 2008).
5. Emplear validación cruzada con catálogos homogéneos y observables independientes (SZ/X) para cerrar la brecha entre calibración y aplicación (Aguado-Barahona *et al.*, 2022).

La comparación controlada entre virial, σ - M y caústico, apoyada en *bootstrap* y en diagnósticos de espacio de fase (Figs. 1-4), muestra que no existe un estimador universalmente superior. La robustez se logra combinando la solidez del virial, la eficiencia estadística de la ley σ - M y la insensibilidad al no-equilibrio del caústico, con calibraciones transparentes y trazables. Esta estrategia metrológica es consistente con la evidencia observacional moderna y con la teoría de halos auto-semejantes (Evrard *et al.*, 2008).

AGRADECIMIENTOS

Se agradece a las colaboraciones SDSS por facilitar el acceso a los datos y a los desarrolladores de código abierto utilizados en los análisis.

REFERENCIAS

- Aguado-Barahona, A., Rubiño-Martín, J.A., Ferragamo, A., Barrena, R., Streblyanska, A. & Tramonte, D. (2022). Velocity dispersion and dynamical masses for 388 galaxy clusters and groups: calibrating the m_{sz} - m_{dyn} scaling relation for the psz2 sample. *Astronomy & Astrophysics*, 659, A126. <https://doi.org/10.1051/0004-6361/202039980>
- Carlberg, R.G., Yee, H.K.C., Ellingson, E., Abraham, R., Gravel, P., Morris, S., *et al.* (1996). Galaxy cluster virial masses and ω_0 . *The Astrophysical Journal*, 462, 32-49. <https://doi.org/10.1086/177125>
- Diaferio, A. (1999). Mass estimation in the outer regions of galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 309, 610-622. <https://doi.org/10.1046/j.1365-8711.1999.02864.x>
- Evrard, A.E., Bialek, J., Busha, M., White, M., Habib, S., Heitmann, K., *et al.* (2008). Virial scaling of massive dark matter halos: why clusters prefer a high-normalization cosmology. *The Astrophysical Journal*, 672 (1), 122-137. <https://doi.org/10.1086/521616>

- Ferragamo, A., Barrena, A., Rubiño-Martín, J.A., Aguado-Barahona, A., Streblyanska, A., Tramonte, D. (2021). Velocity dispersion and dynamical mass for 270 galaxy cluster in the *Planck* PSZ1 catalogue. *Astronomic & Astrophysics*, A115, 655. <https://doi.org/10.1051/0004-6361/202140382>
- Geller, M.J., Diaferio, A. & Kurtz, M.J. (1999). The mass profile of the coma galaxy cluster. *Astrophysical Journal Letters*, 517 (1), L23-L26. This paper uses the caustic technique to estimate the Coma cluster mass out to $10 h^{-1}$ Mpc. <https://doi.org/10.1086/312024>
- Gifford, D., Miller, C.J. & Kern, N. (2013). A systematic analysis of caustic methods for galaxy cluster masses. *The Astrophysical Journal*, 773 (2), 116. <https://doi.org/10.1088/0004-637X/773/2/116>
- Heisler, J., Tremaine, S. & Bahcall, J.N. (1985). Estimating the masses of galaxy groups: alternatives to the virial theorem. *The Astrophysical Journal*, 298, 8-17. <https://doi.org/10.1086/163584>
- Munari, E., Biviano, A., Borgani, S., Murante, G. & Fabjan, D. (2013). The relation between velocity dispersion and mass in simulated clusters of galaxies: dependence on the tracer and the baryonic physics. *Monthly Notices of the Royal Astronomical Society*, 430 (4), 2638-2649. <https://doi.org/10.1093/mnras/stt049>
- Pizzardo, P., Geller, M.J., Kenyon, S.J., Damjanov, I. & Diaferio, A. (2023). An illustrating view of the caustic technique for galaxy cluster mass estimation. *Astronomy & Astrophysics*, 675, A56. <https://doi.org/10.1051/0004-6361/202346545>
- Ruel, J., Bazin, G., Bayliss, M., Brodwin, M., Foley, R.J., Stalder, B, *et al.* (2014). Optical spectroscopy and velocity dispersions of galaxy clusters from the SPT-SZ survey. *The Astrophysical Journal*, 792 (1), 45. <https://doi.org/10.1088/0004-637X/792/1/45>
- Tian, Y., Yu, P.-C., Li, P. & McGaugh, S. S. (2021). Mass-velocity dispersion relation in hiflugs galaxy clusters. *The Astrophysical Journal*, 910 (1), 56. <https://doi.org/10.3847/1538-4357/abe45c>