

# ROGER:Reconstructing Orbits of Galaxies in Extreme Regions using machine learning techniques (2010.11959)

Martín de los Ríos, Héctor J. Martínez, Valeria Coenda, Hernán Muriel, Andrés N. Ruiz, Cristian A. Vega-Martínez & Sofía A. Cora



# 1. Introducción

- a. Motivación.
- b. Técnicas de *machine learning*.
- c. Multidark simulations + SAG.
- d. Tipos de galaxias según sus órbitas.

# 2. Resultados de ROGER

- a. Estimación de probabilidades según la posición en el espacio de fases proyectado.
- b. Clasificación a partir de las probabilidades.

# 3. Tutorial

# 4. Conclusiones

# 1. Introducción:

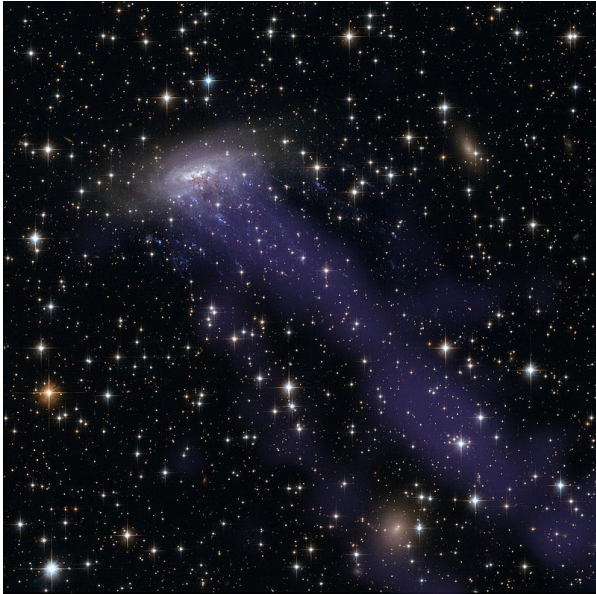
## a. Motivación y trabajos previos

La formación y evolución de las galaxias es un proceso complicado que implica la acción de diferentes mecanismos físicos que actúan a *diferentes escalas temporales y espaciales*.



En cuanto a su origen, estos procesos pueden deberse a factores **internos (masa)** o **externos (ambiente)**.

# MECANISMOS EXTERNOS



Jellyfish galaxy

- Ram pressure stripping (e.g. Gunn & Gott 1972, Abadi et al. 1999, Rasmussen et al. 2006, Jaffé et al. 2012, Hess & Wilcots 2013)
- Starvation/Strangulation (e.g. Larson et al. 1980, Bekki 2009, McCarthy et al. 2008, Bahé et al. 2013, Vijayaraghavan & Ricker 2015)

# MECANISMOS EXTERNOS



- Tidal stripping (e.g. Gnedin 2003b, Villalobos et al. 2014) and thermal evaporation (Cowie & Songaila 1977)
- Galaxy-galaxy interactions (e.g. Moore et al. 1996, Moore et al. 1999, Gnedin 2003a)
- Fusiones de Galaxias, principalmente en grupos de galaxias.

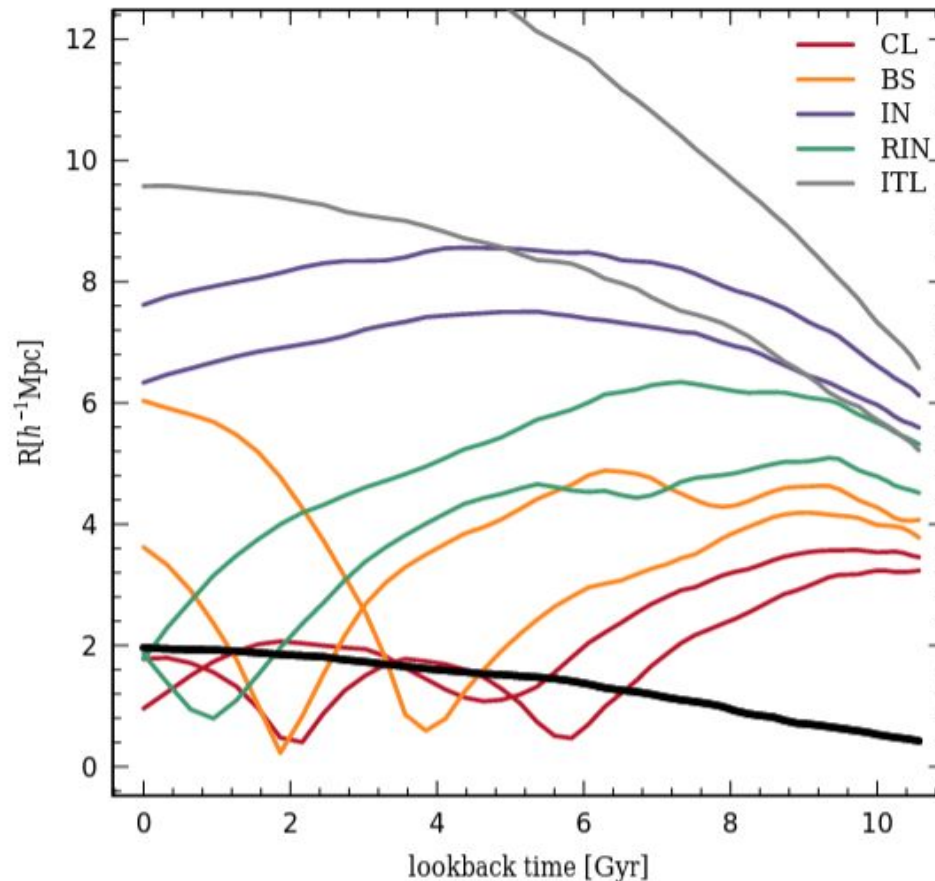
# 1. Introducción:

## b. Multidark simulations + SAG.

- Simulación cosmológica Multidark N-body con  $3840^3$  partículas en un volumen de  $1\text{Gpc}^3$ , evolucionadas desde  $z=120$  a  $z=0$  con una cosmología Planck.
- Modelo semi-analítico SAG (Cora et al. 2018).
- Halos más masivos  $M_{\{200\}} > 10^{15} M_{\{\text{sun}\}}/h$  (34 Halos)
- Criterio de aislamiento.
- Estudiamos órbitas de galaxias con una masa estelar mínima de  $10^{8.5} M_{\{\text{sub}\}}/h$  ( $\sim 30000$  galaxias).

# 1. Introducción:

## c. Tipos de galaxias según sus órbitas.



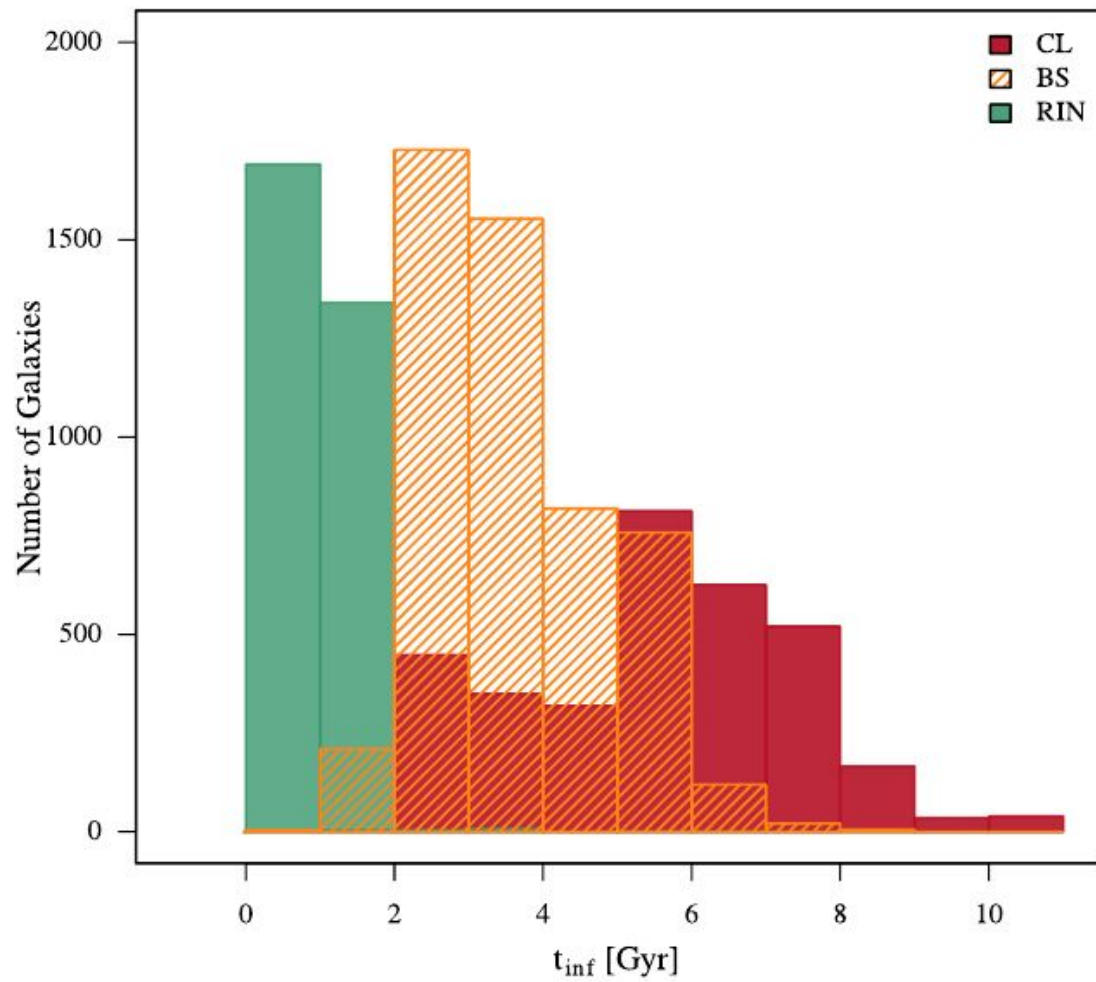
**Galaxias de Cúmulos:** Galaxias que están orbitando alrededor del centro del cúmulo.

**Galaxias en Infall:** Galaxias que están cayendo hacia el cúmulo.  $d(t) > R_{200}$

**Galaxias “Recent infallers”:** Galaxias que ingresaron recientemente al cúmulo.  $d(t_0) < R_{200}$ .

**Galaxias Backsplash:** Galaxias que cruzaron  $R_{200}$  dos veces, una hacia adentro y otra hacia afuera.  $d(t_0) > R_{200}$

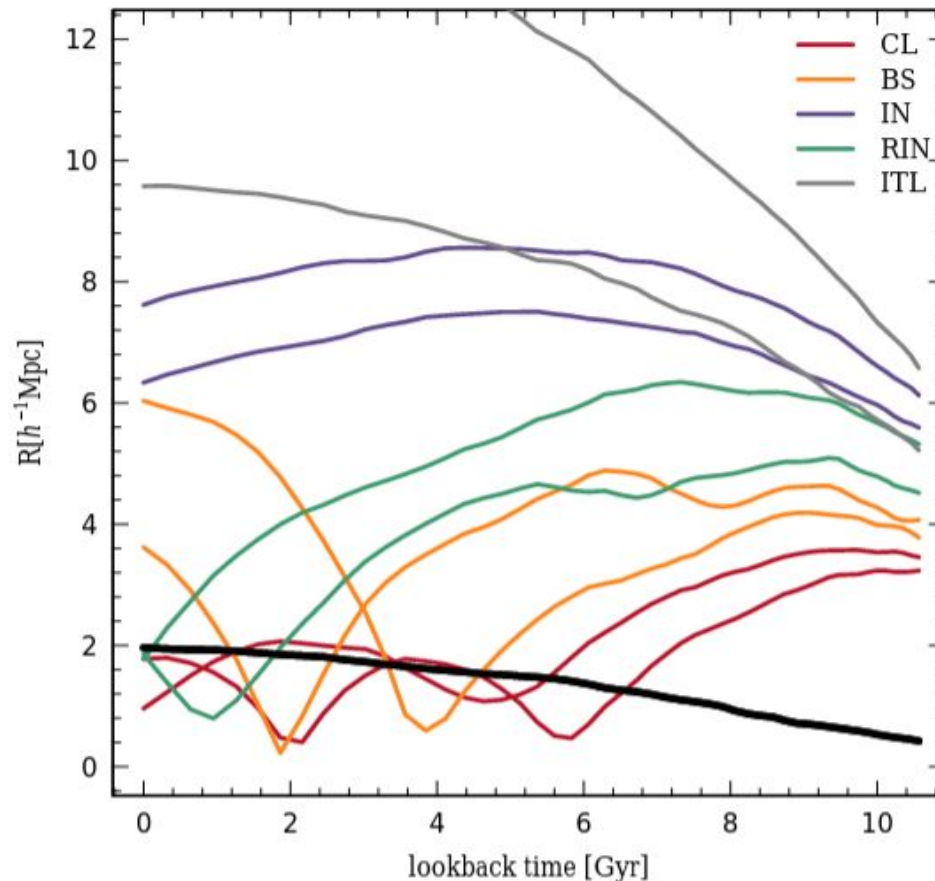
**Galaxias Interlopers:** Galaxias que por proyección están en el espacio de fases proyectado.





# 1. Introducción:

## c. Tipos de galaxias según sus órbitas.



**Galaxias de Cúmulos:** Galaxias que están orbitando alrededor del centro del cúmulo.

**Galaxias en Infall:** Galaxias que están cayendo hacia el cúmulo.  $d(t) > R_{200}$

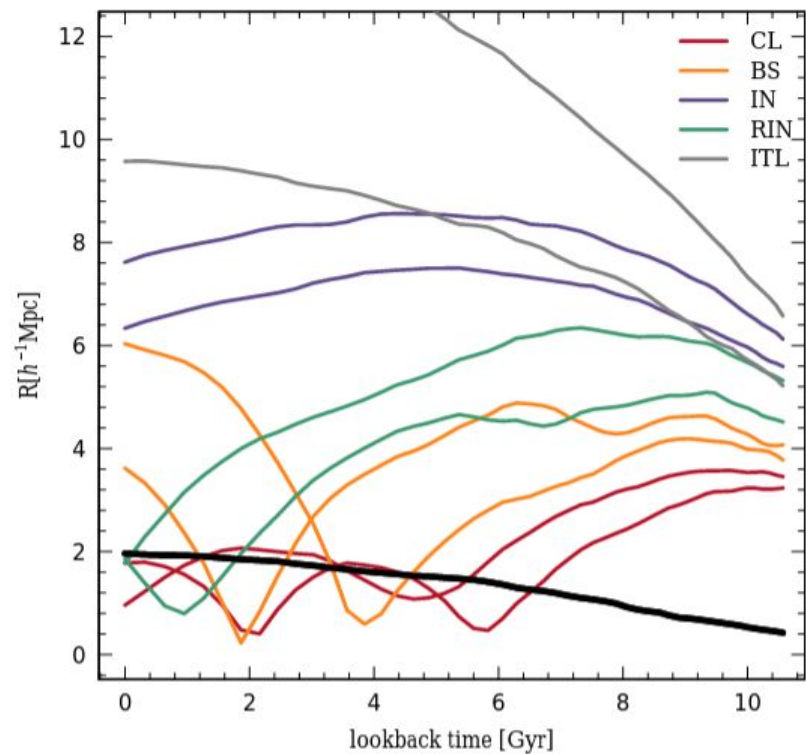
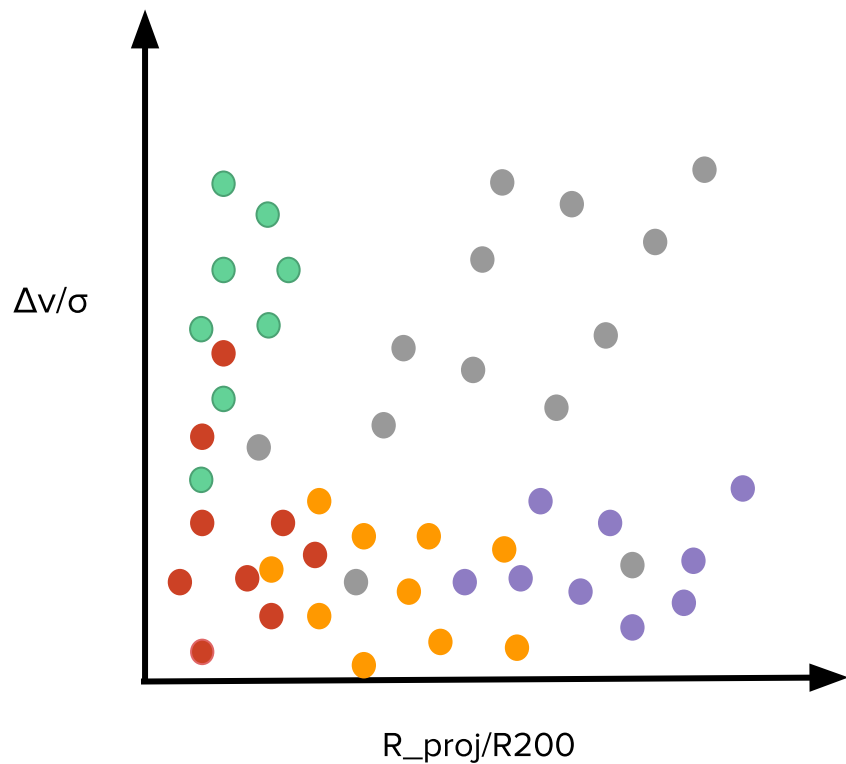
**Galaxias “Recent infallers”:** Galaxias que ingresaron recientemente al cúmulo.  $d(t_0) < R_{200}$ .

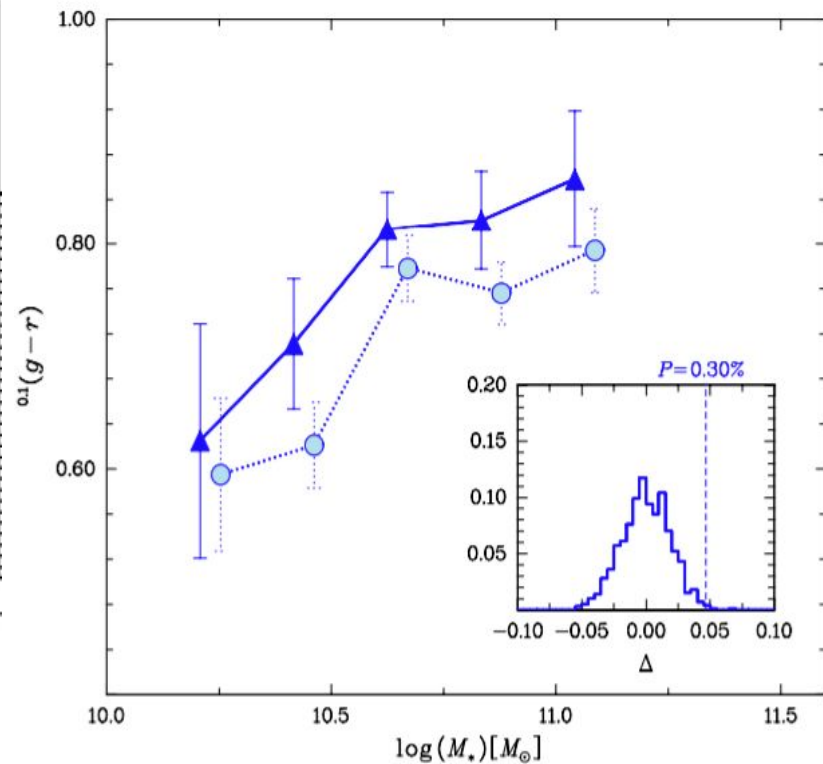
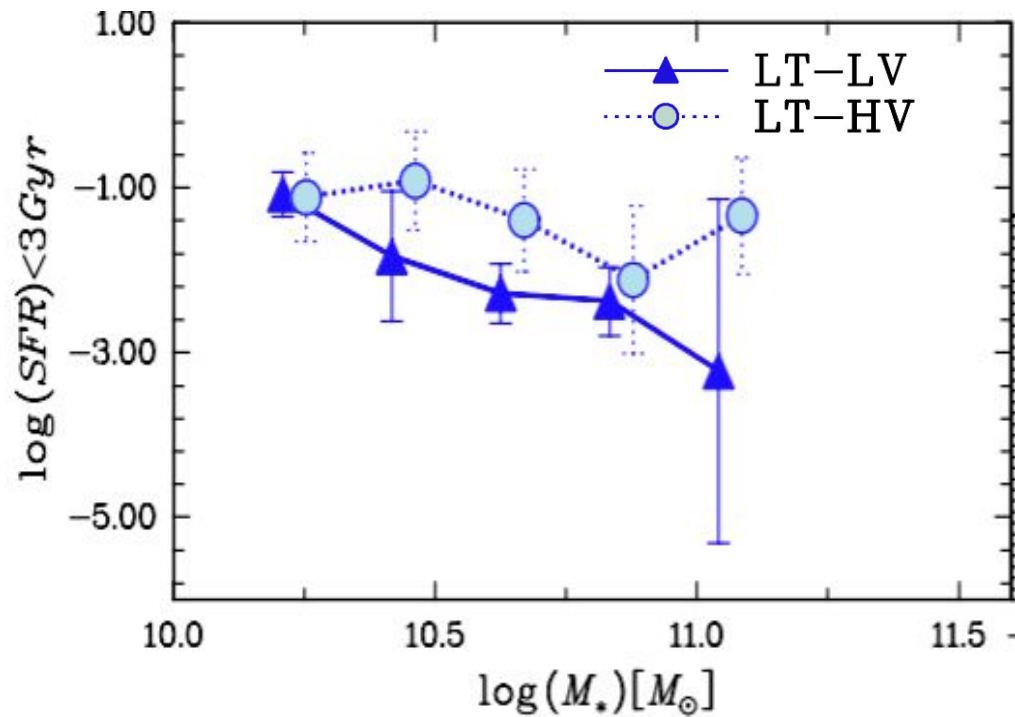
**Galaxias Backsplash:** Galaxias que cruzaron  $R_{200}$  dos veces, una hacia adentro y otra hacia afuera.  $d(t_0) > R_{200}$

**Galaxias Interlopers:** Galaxias que por proyección están en el espacio de fases proyectado.

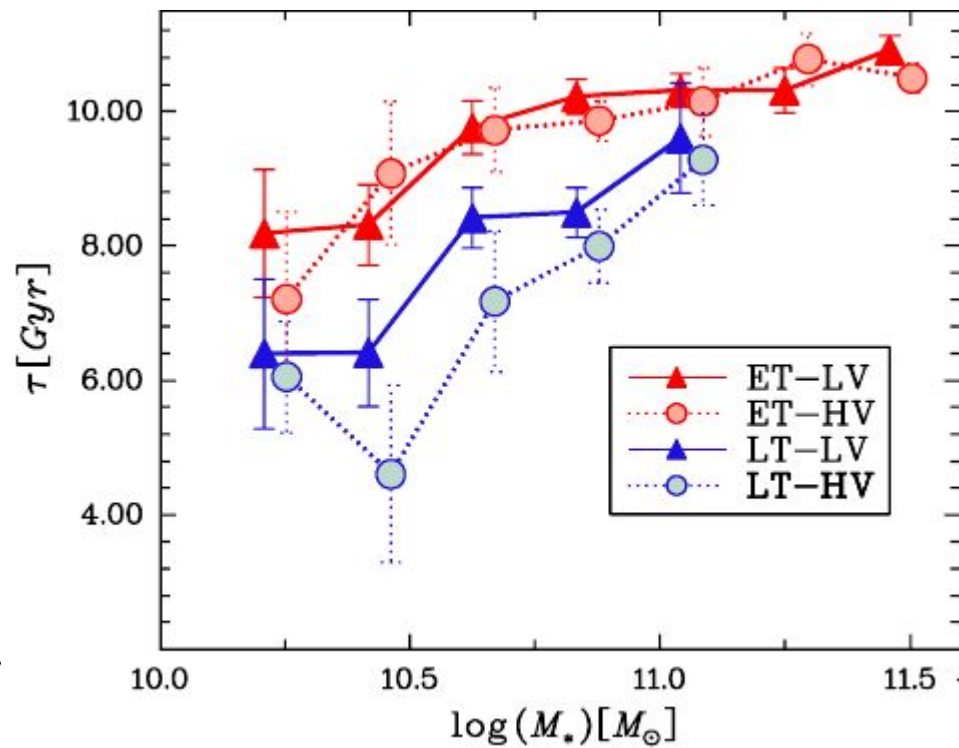
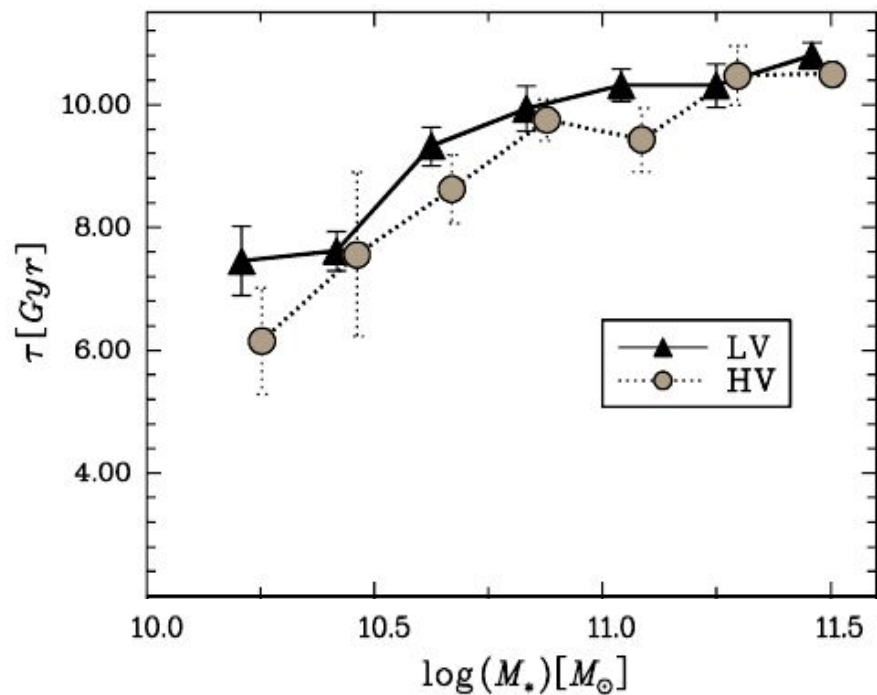
Información  
Observacional

Información 3D

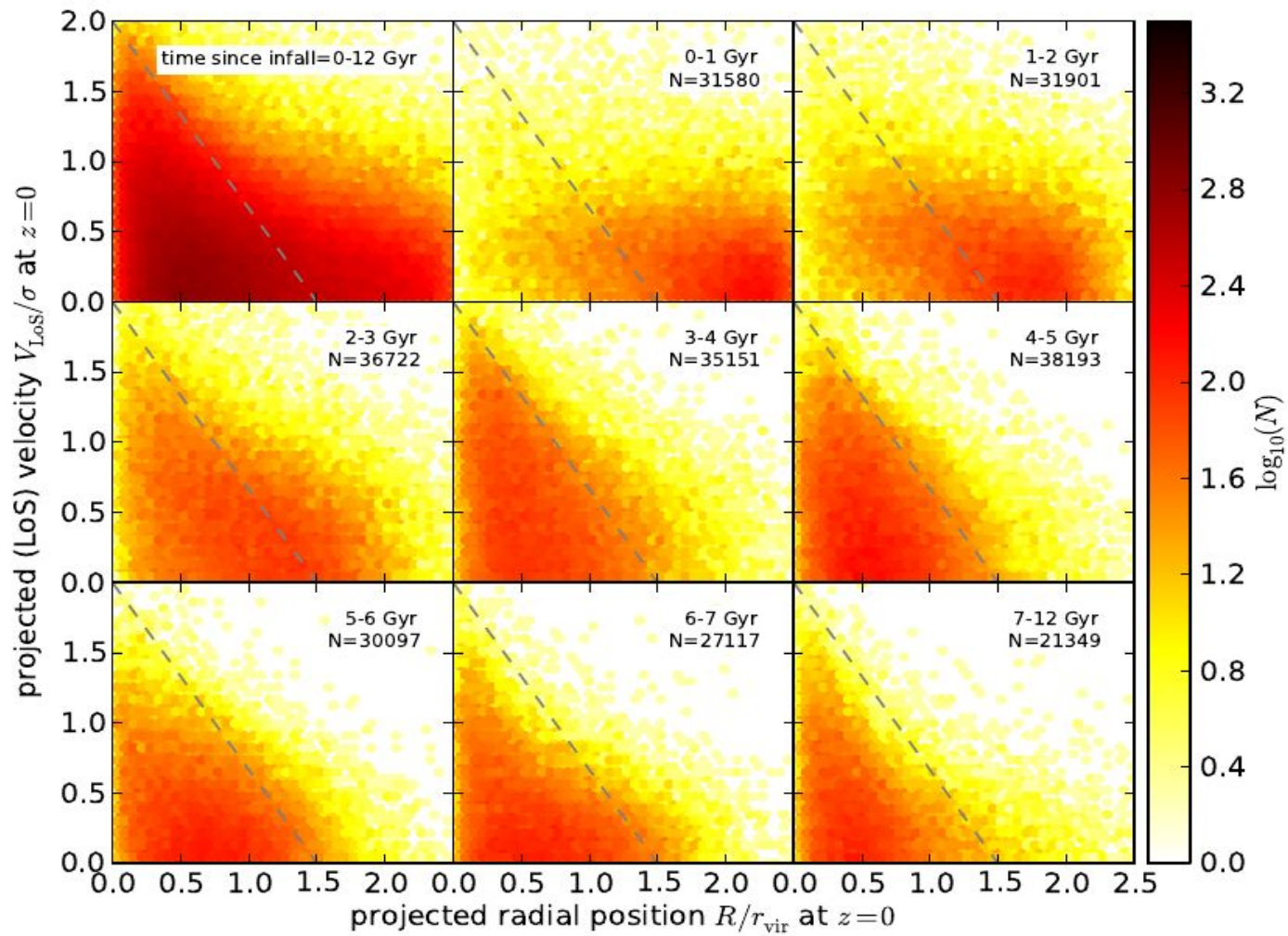




Muriel & Coenda 2014 (1402.3594)

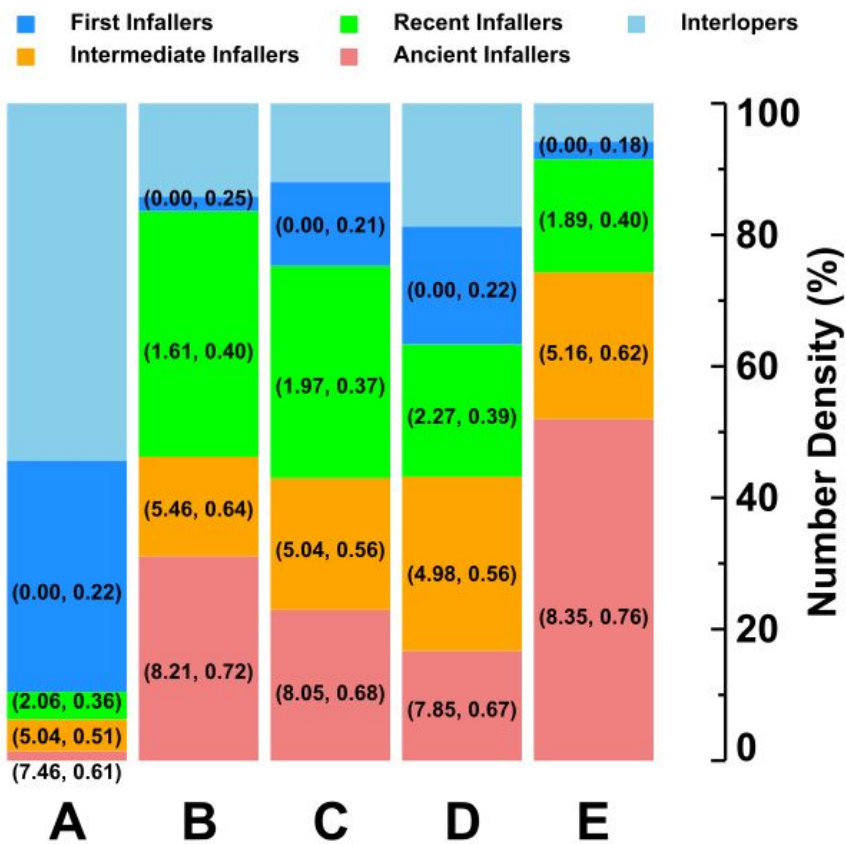
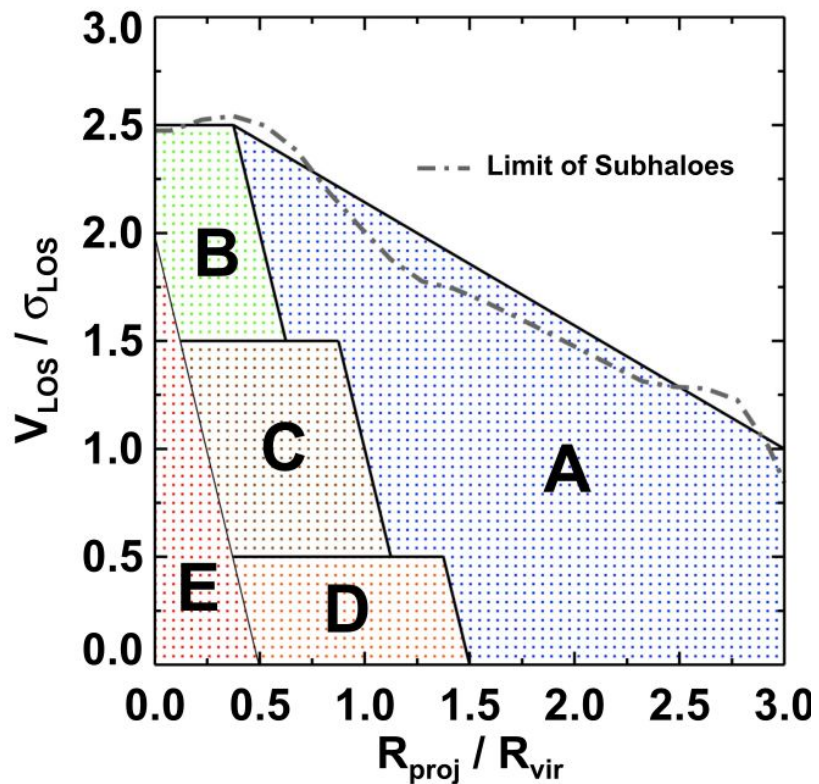


Muriel & Coenda 2014 (1402.3594)

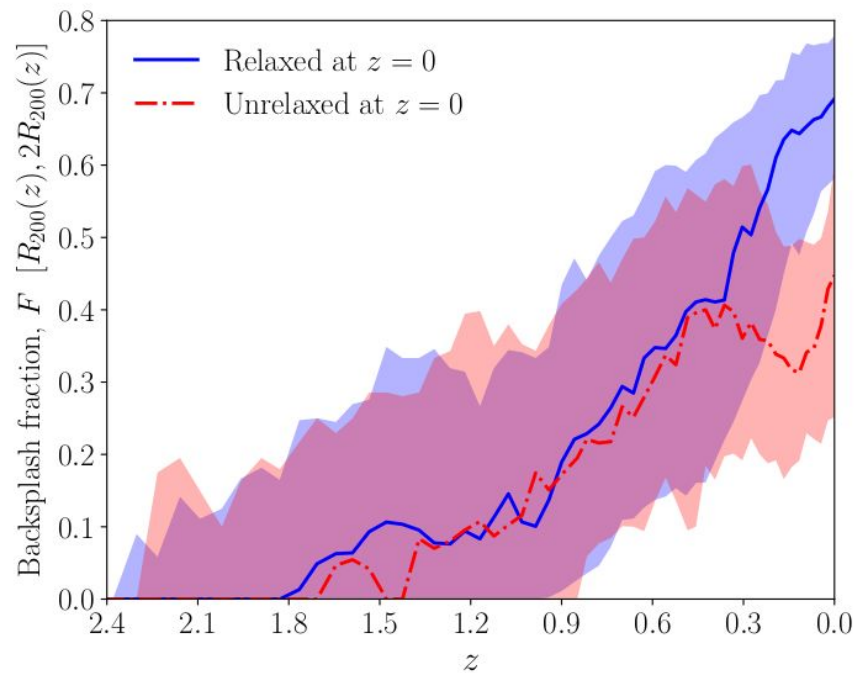
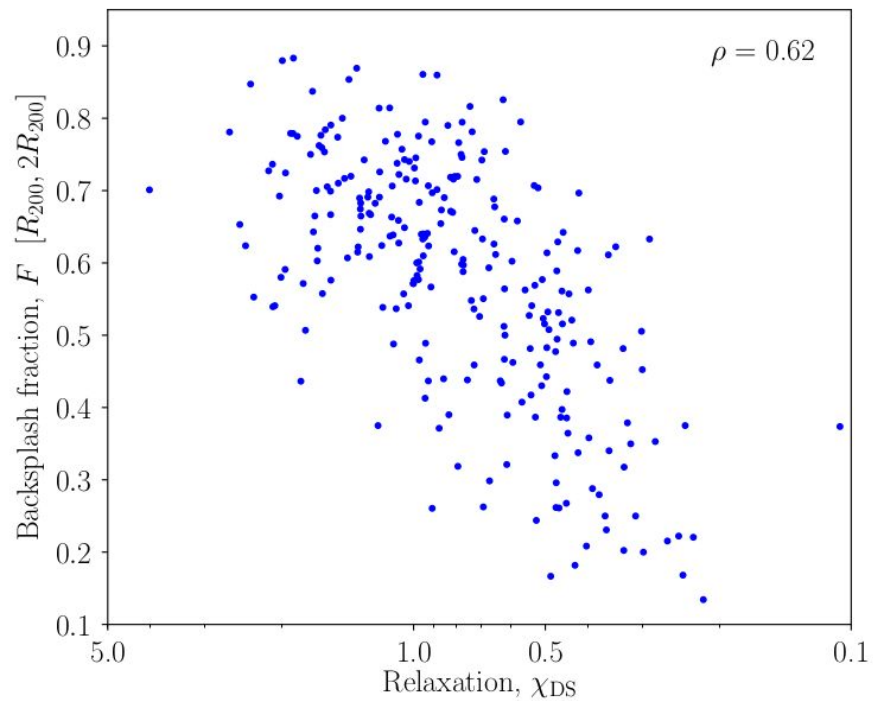


Oman et al. 2013  
(1301.6757)





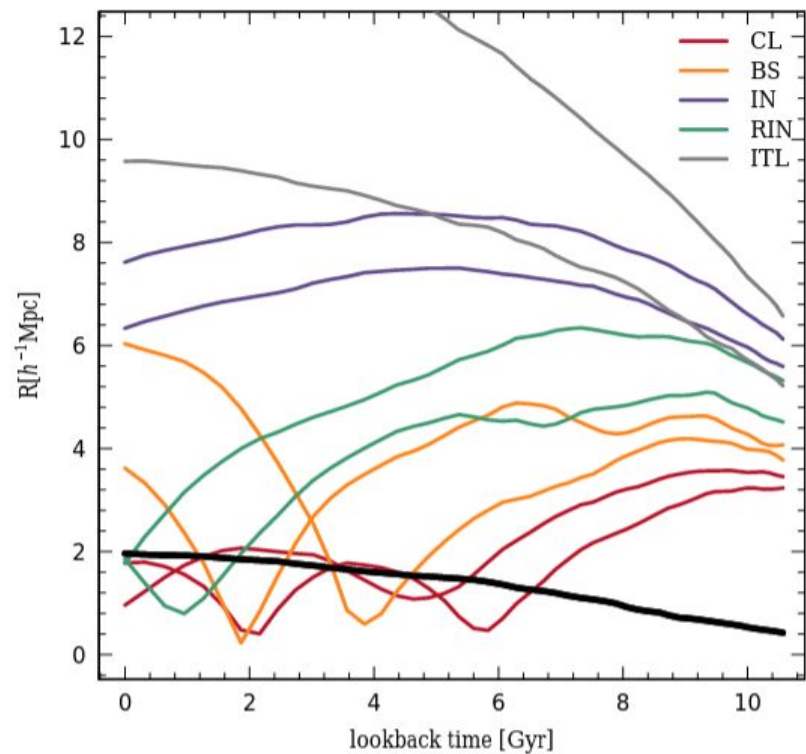
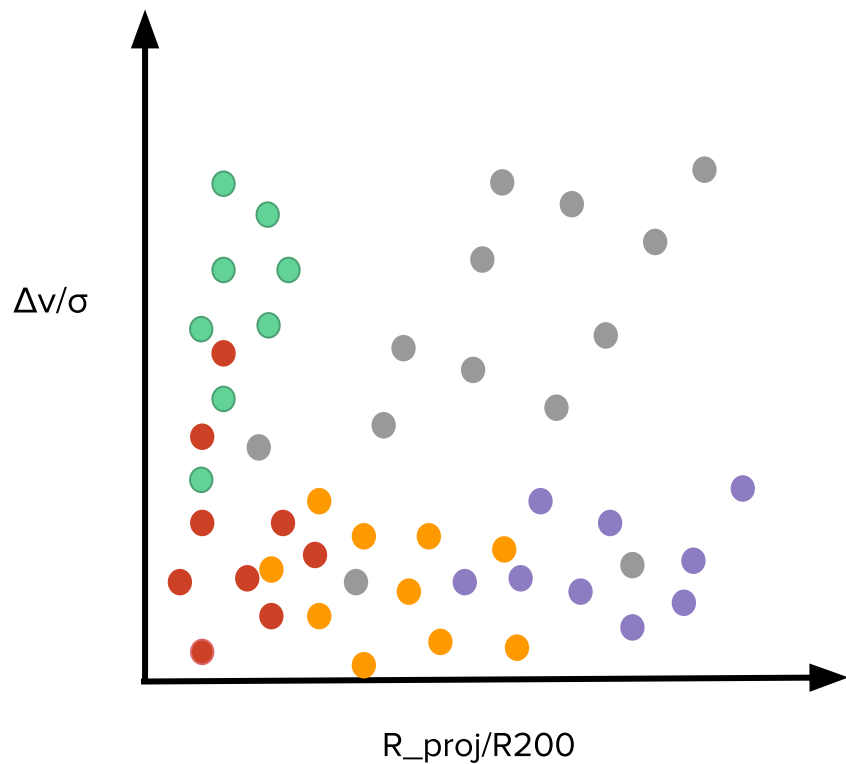
Rhee et al. 2017 (1704.04243)



Haggar et al. 2020 (2001.11518)

Información  
Observacional

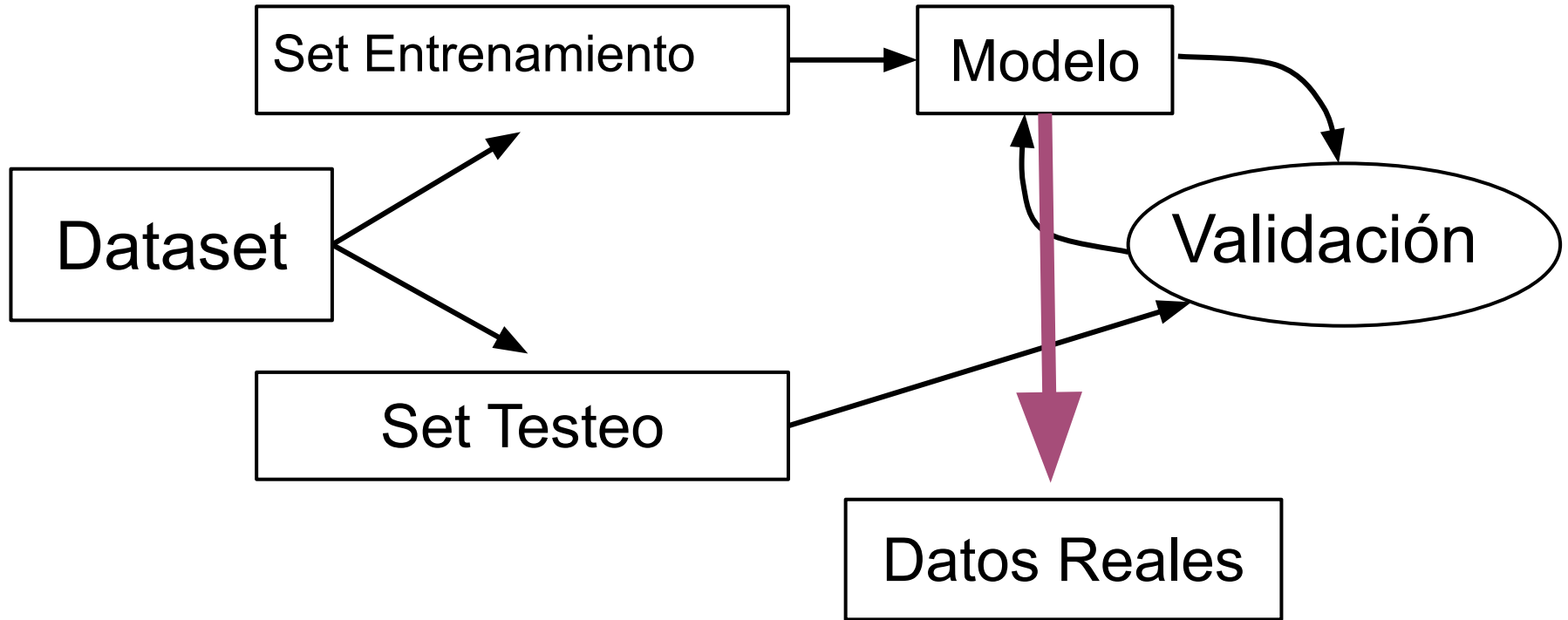
Información 3D

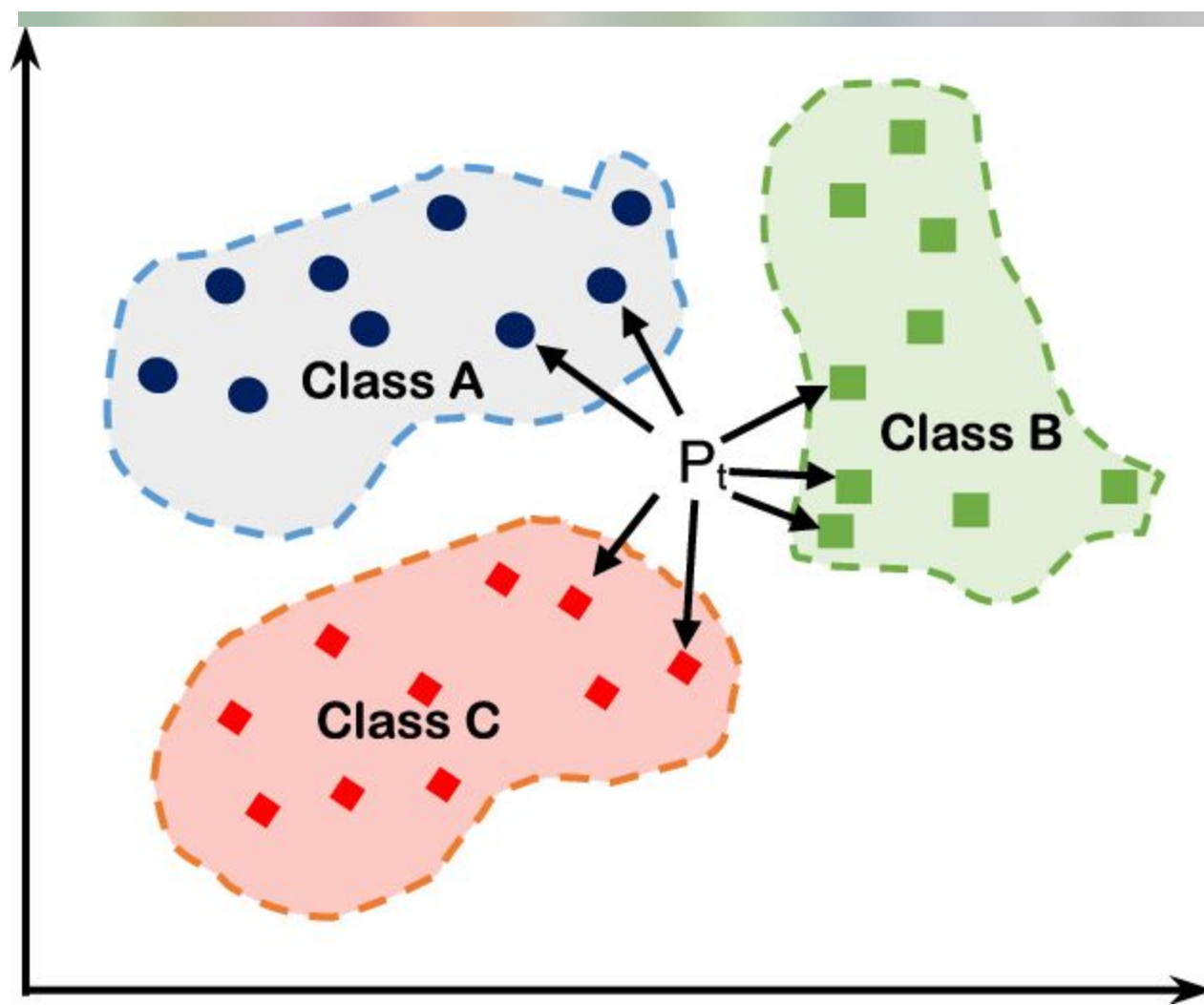




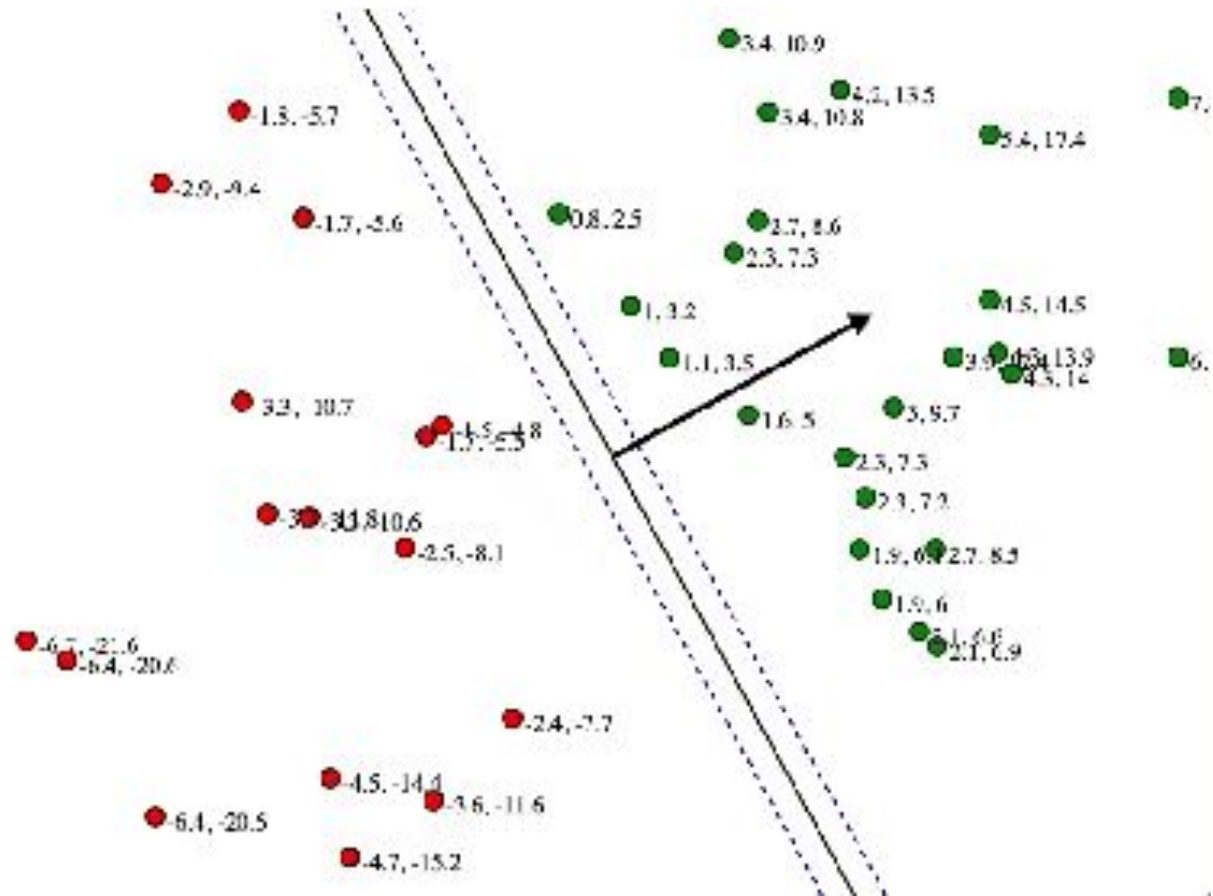
# 1. Introducción:

d. Técnicas de Machine Learning

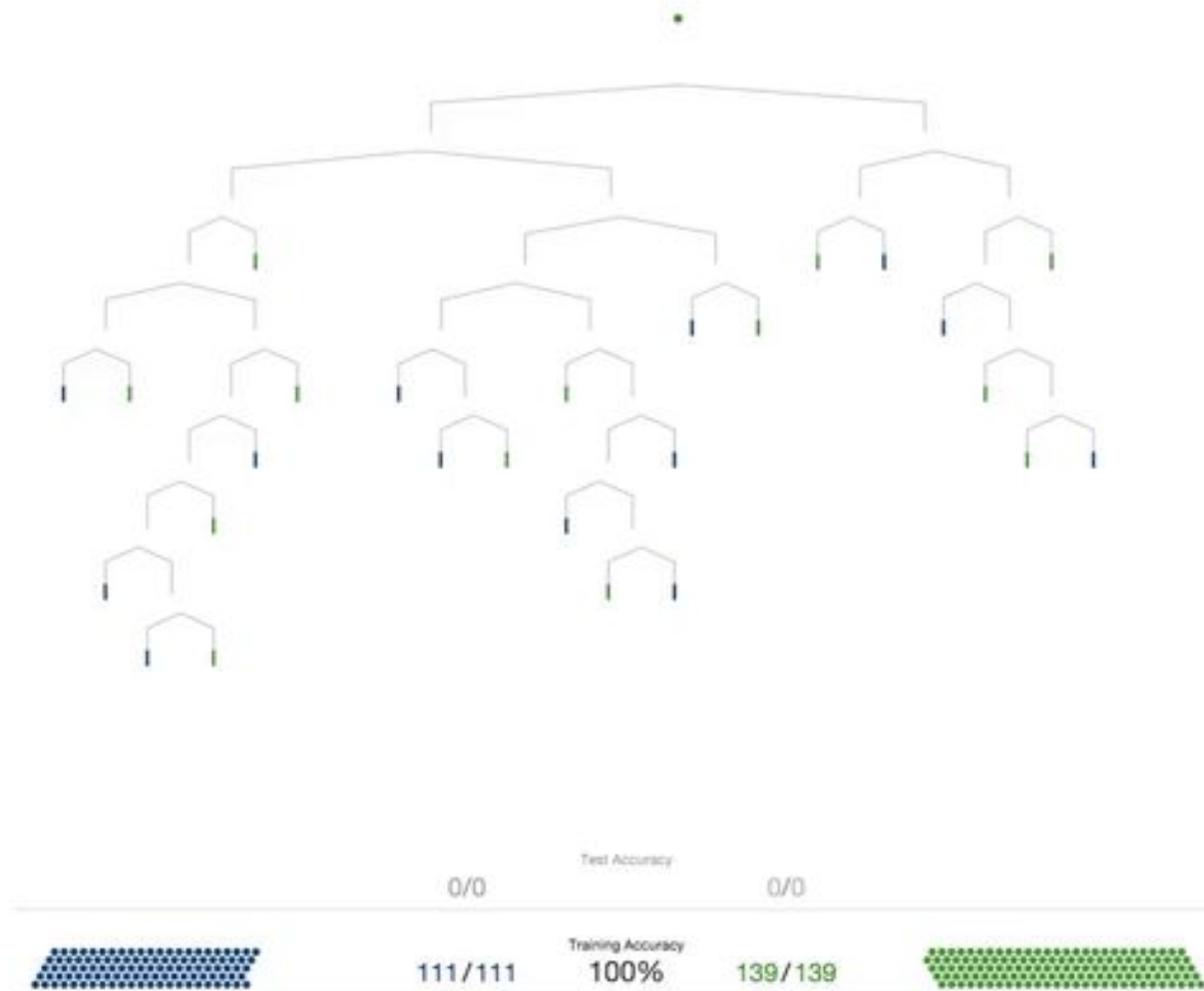




# Support Vector Machines



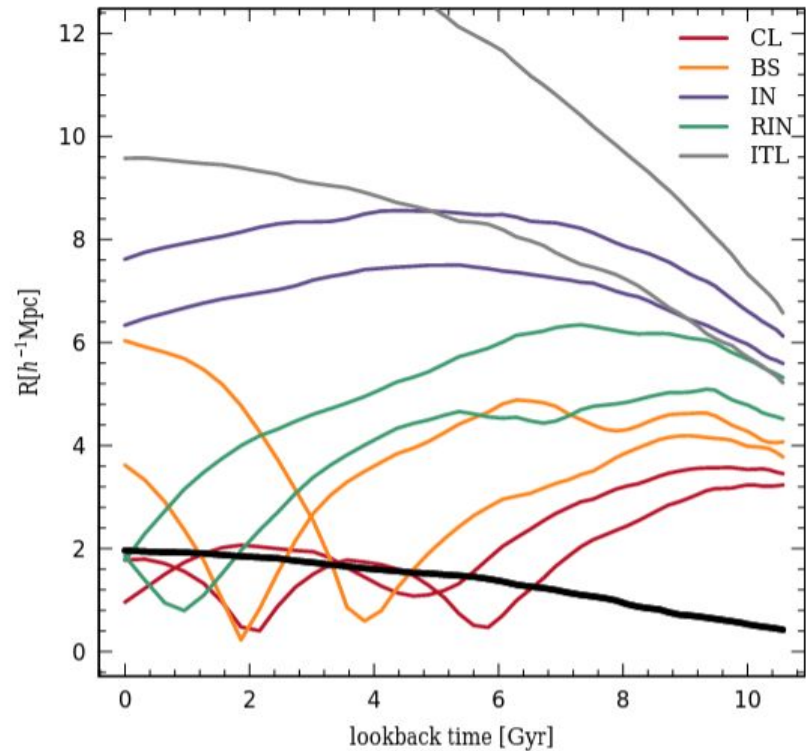
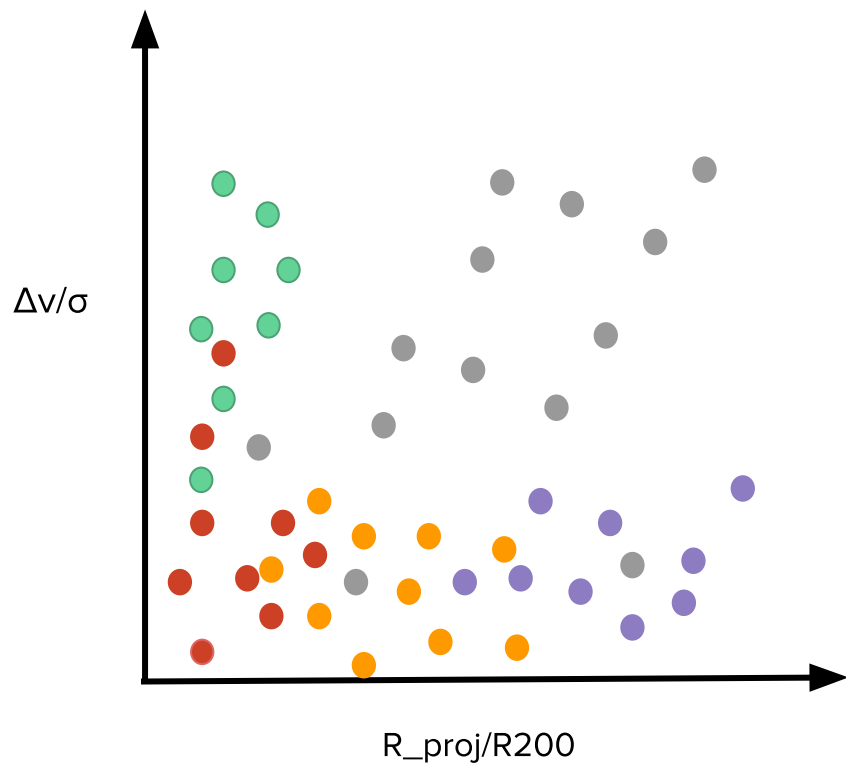
# Random Forest



Información  
Observacional

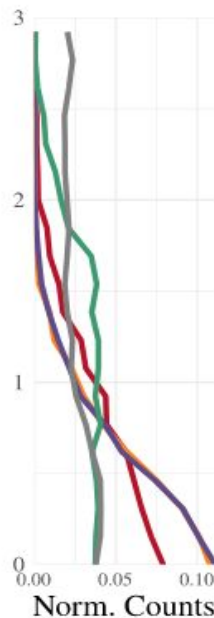
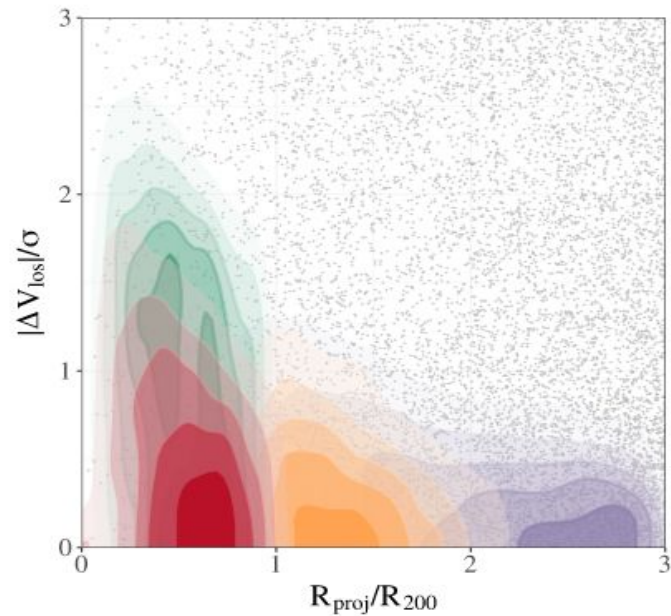
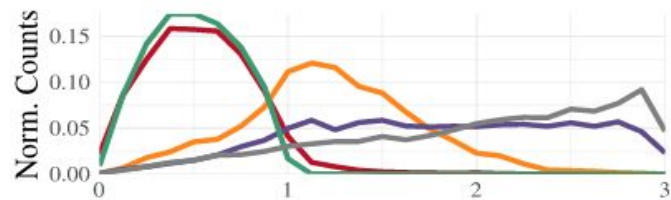
**ROGER**

Información 3D

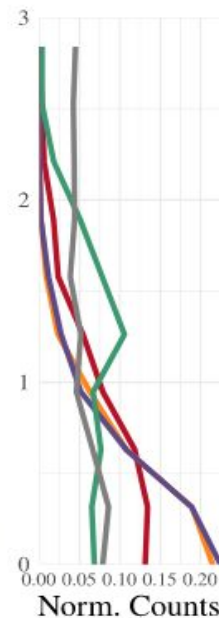
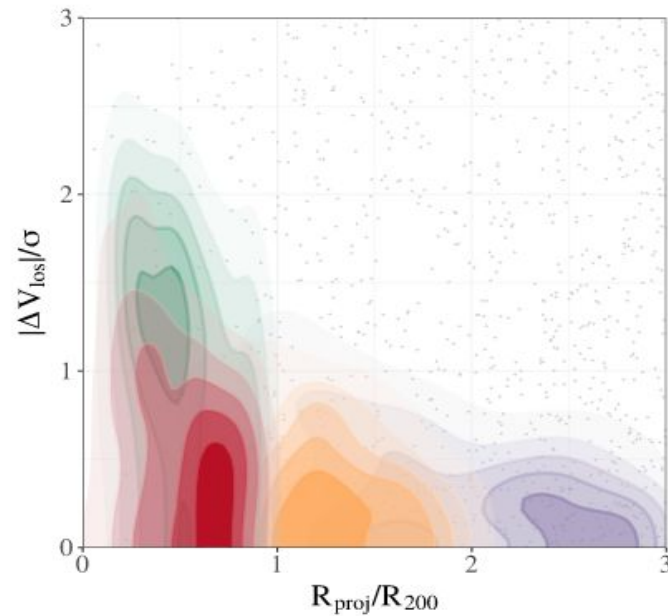
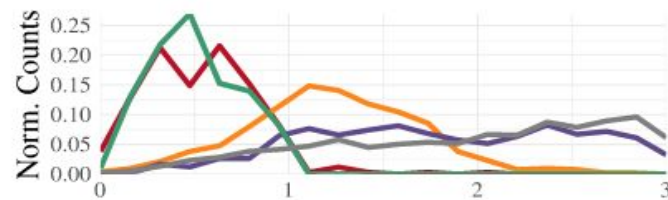


## 2. Resultados

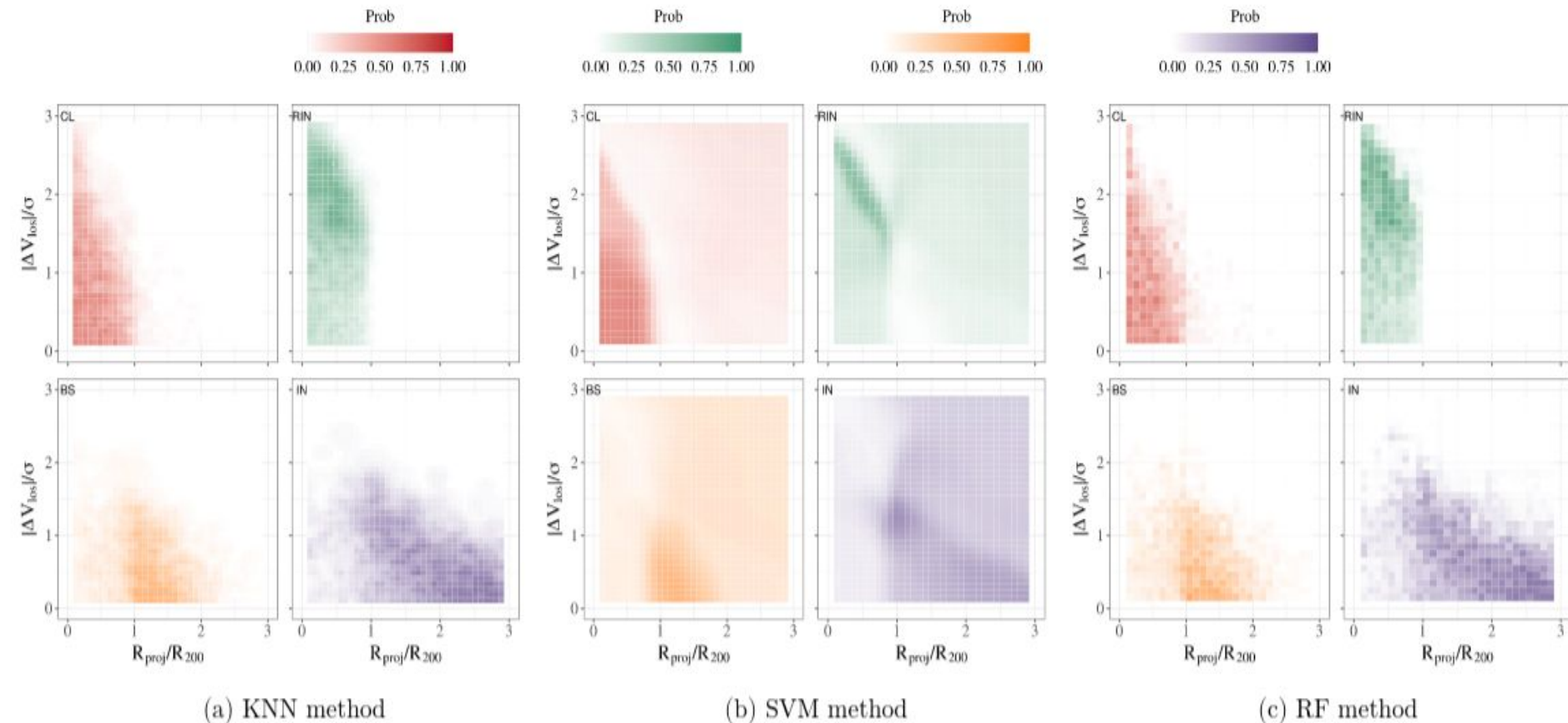
Set Entrenamiento



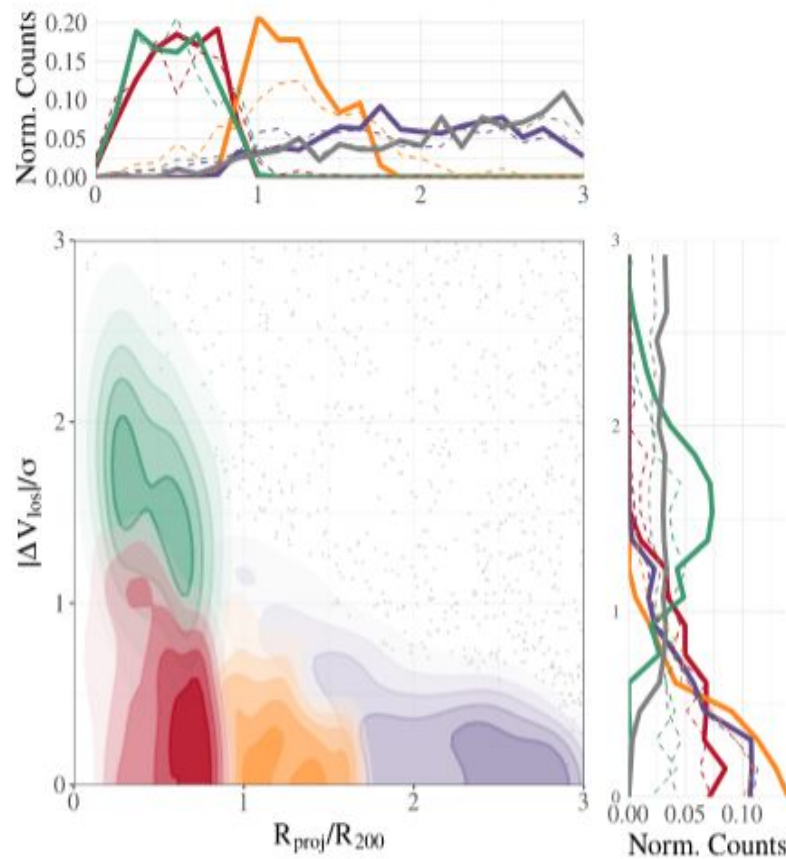
Set Testeo



# Probabilidades según la posición en el espacio de fases



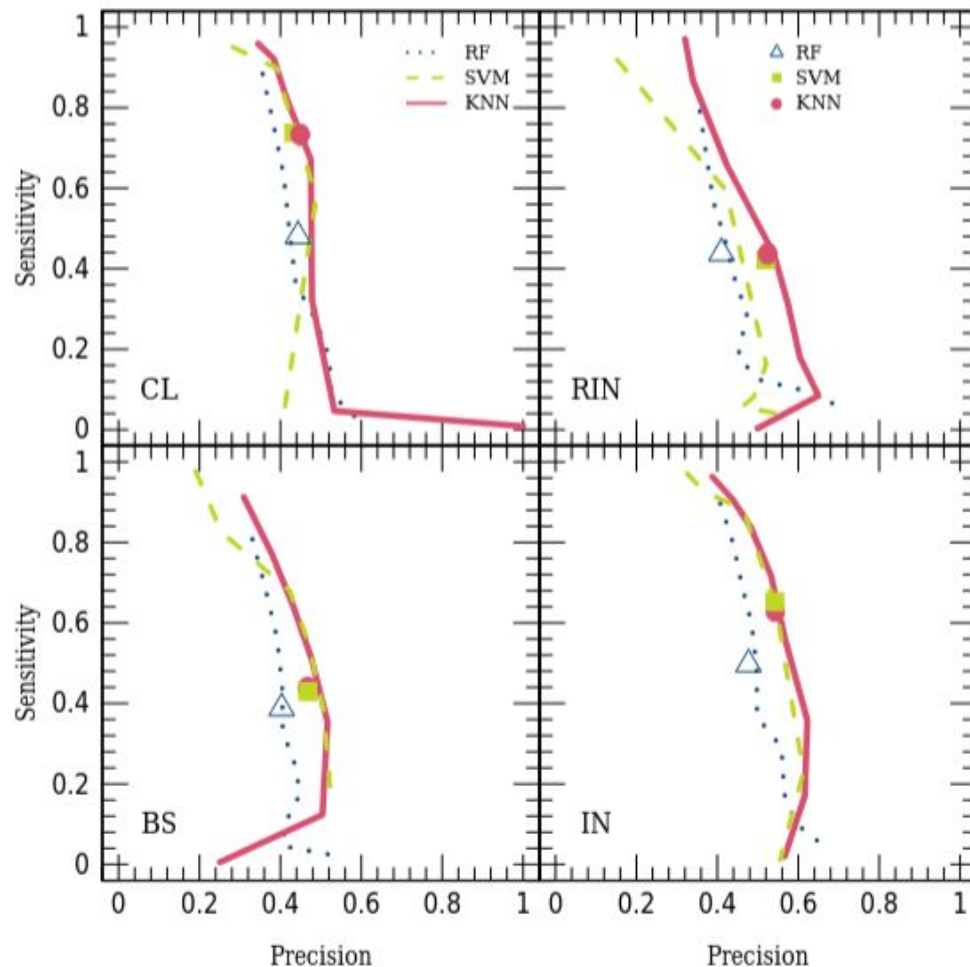
Clasificamos teniendo en cuenta la clase más probable.



(a) KNN method

Predicted	CL	RIN	BS	IN	ITL
	74% n=254	45% n=139	14% n=74	6% n=51	5% n=46
	15% n=50	45% n=137	5% n=27	2% n=19	3% n=29
	8% n=27	6% n=20	45% n=232	20% n=166	6% n=57
	3% n=11	3% n=8	34% n=180	63% n=526	27% n=245
ITL	0% n=1	1% n=4	2% n=13	9% n=75	59% n=539
	CL	RIN	BS	IN	ITL
Real					

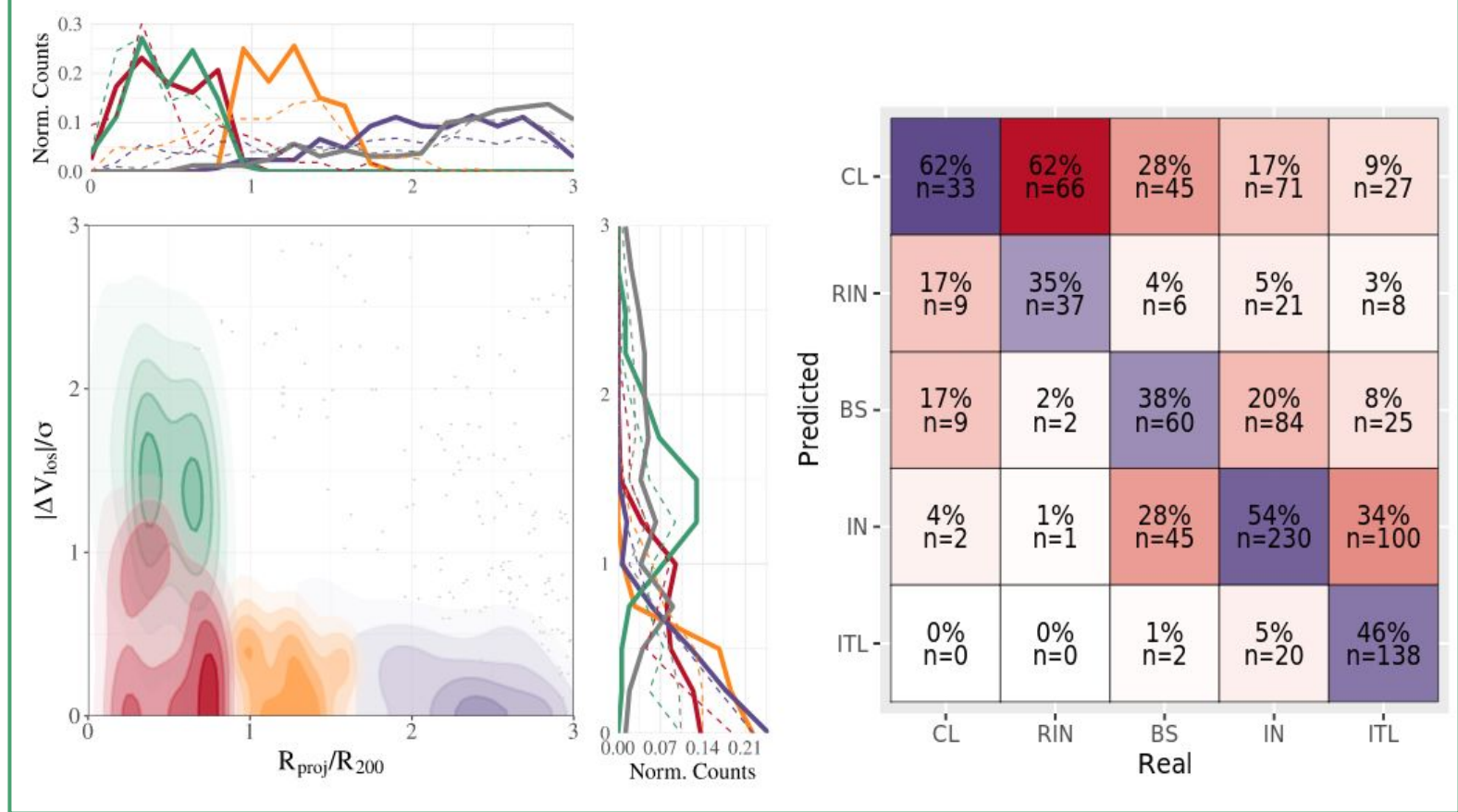




Clasificamos una galaxia como perteneciente a una dada clase, si la probabilidad de que esa galaxia pertenezca a dicha clase es mayor a un valor crítico.

**Precisión:** Número de galaxias bien clasificadas de una dada clase, dividido el número total de galaxias clasificadas como de dicha clase.

**Sensitividad:** Número de galaxias bien clasificadas de una dada clase, dividido el número total de galaxias de dicha clase



Realizamos una clasificación random para cada galaxia teniendo en cuenta las probabilidades predichas.

# Tutorial

- <https://github.com/Martindelosrios/ROGER>
- [https://mdelosrios.shinyapps.io/roger\\_shiny/](https://mdelosrios.shinyapps.io/roger_shiny/)

# Trabajos Futuros

- Estudiar las propiedades de los diferentes tipos de galaxias.
- Analizar la evolución temporal de los diferentes tipos de galaxias.
- Estimar la correlación entre la fracción de cada clase de galaxia y el estado dinámico de los cúmulos (MeSsl).
- Analizar las posibles dependencias angulares en el ingreso y salida de las galaxias backsplash.

# Conclusiones

- Desarrollamos un algoritmo de aprendizaje automático para clasificar dinámicamente a las galaxias según su posición en el espacio de fases proyectado.
- El algoritmo estima la probabilidad de cada galaxia de pertenecer a cada clase orbital.
- El algoritmo con mejor performance es el KNN, alcanzando una precisión de 66%, 52%, 75% y 38% al clasificar galaxias de cúmulos, backsplash, infallers and recent infallers respectivamente.
- El código puede ser utilizado libremente como un paquete de R (<https://github.com/martindelosrios/ROGER>) o a través de una interfaz web ([https://mdelosrios.shinyapps.io/roger\\_shiny/](https://mdelosrios.shinyapps.io/roger_shiny/)).



**GRACIAS!**