

ROGER:Reconstructing Orbits of Galaxies in Extreme Regions using machine learning techniques

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



1. Introducción

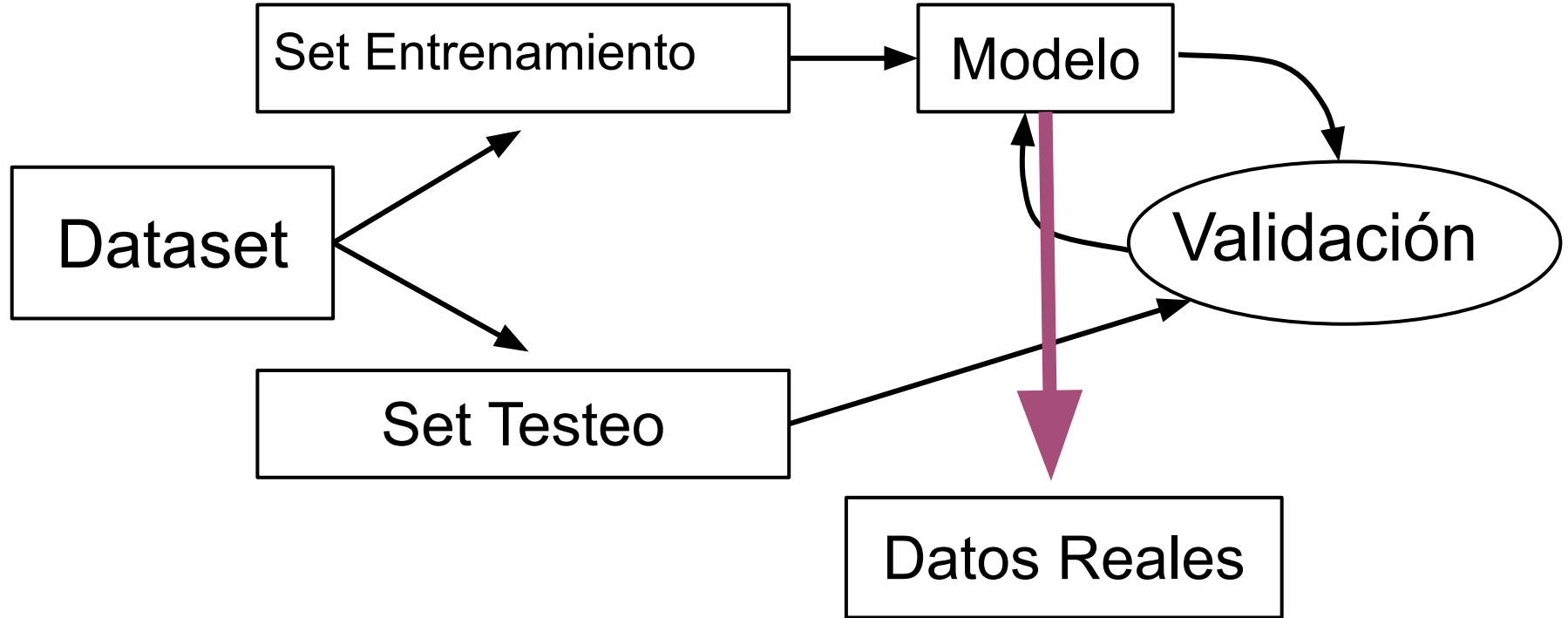
- a. Técnicas de *machine learning*.
- b. Multidark simulations + SAG.
- c. Tipos de galaxias según sus órbitas.

2. Resultados de ROGER

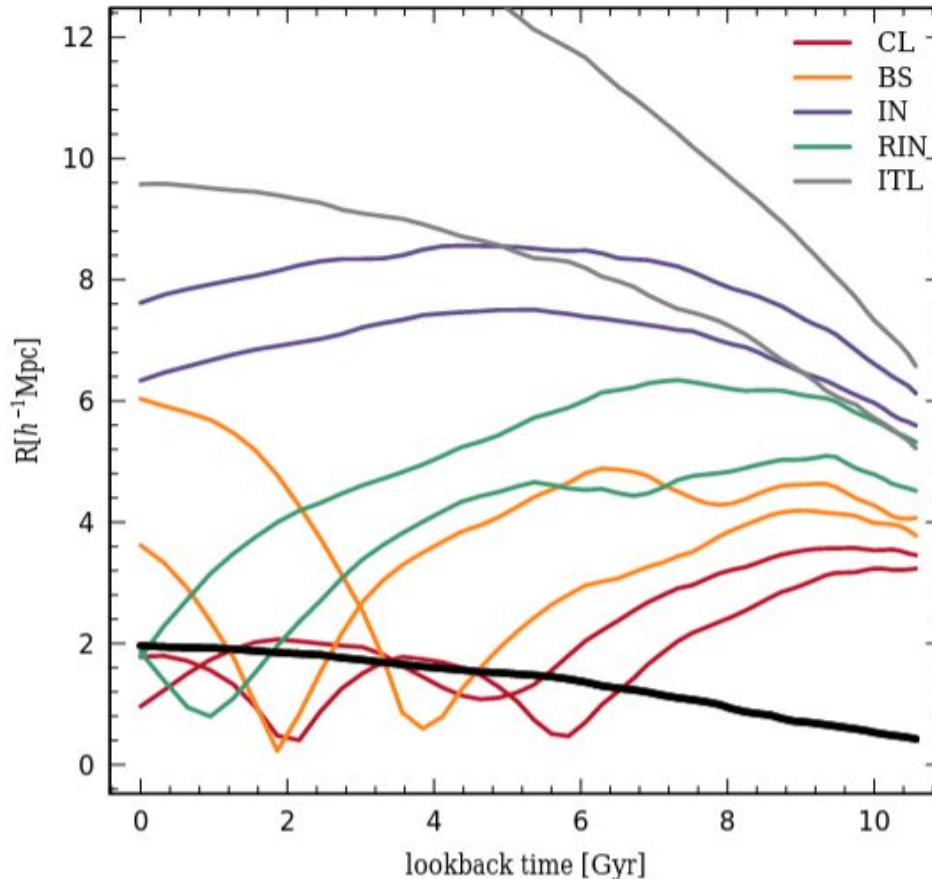
- a. Estimación de probabilidades según la posición en el espacio de fases proyectado.
- b. Clasificación a partir de las probabilidades.

3. Conclusiones

1. Introducción



- Simulación cosmológica Multidark N-body con 3840^3 partículas en un volumen de 1Gpc^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 $10^{8.5} M_{\{\text{sub}\}}/h$ (~30000 galaxias).



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. $d(t_0) < R_{200}$.

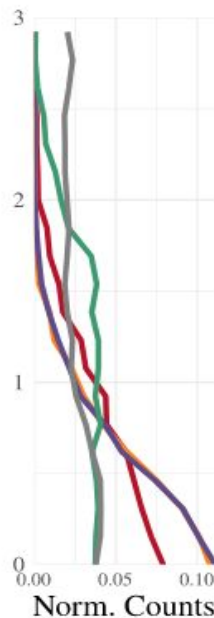
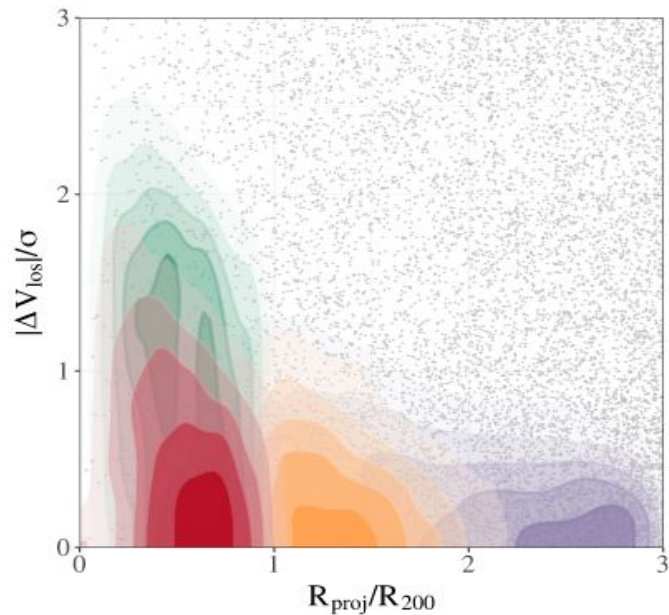
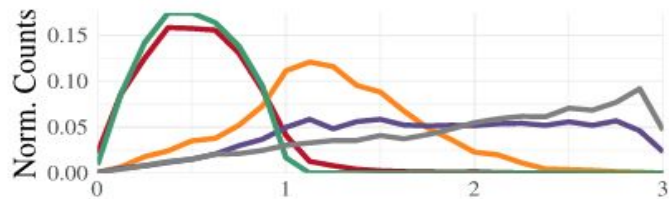
Galaxias Backsplash: Galaxias que cruzaron R_{200} dos veces, una hacia adentro y otra hacia afuera. $d(t_0) > R_{200}$

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

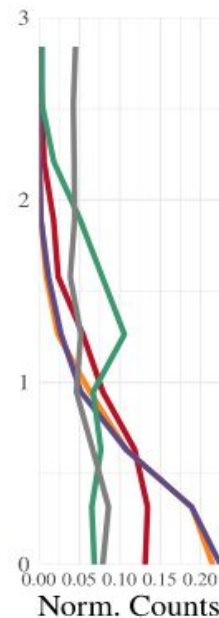
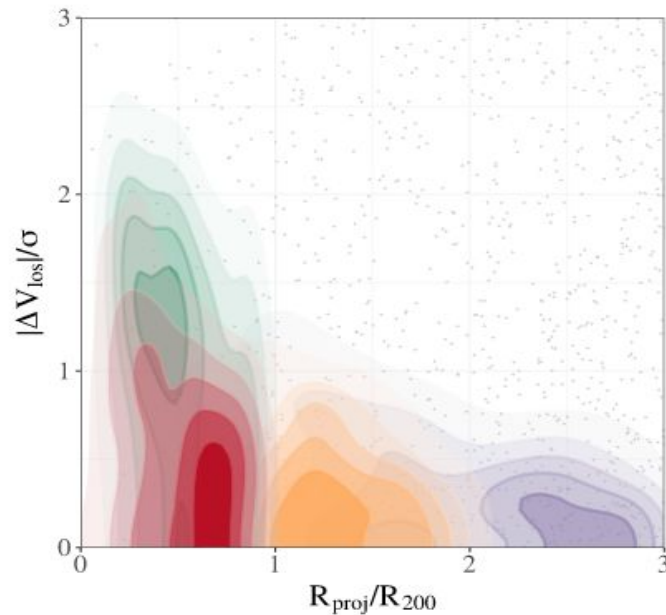
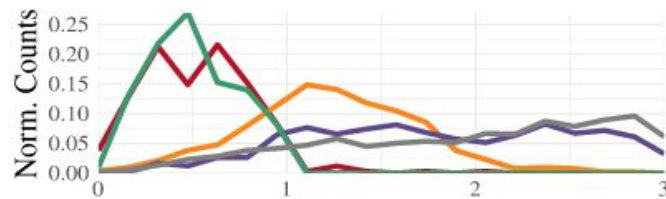
- Se espera que las galaxias backsplash sean estadísticamente más rojas que las “infalling”.
- Se espera que las galaxias “recent infalling” sean más azules que las galaxias del cúmulo.
- Cúmulos relajados tienen una mayor fracción de galaxias backsplash. (Haggar et al. 2020)

2. Resultados

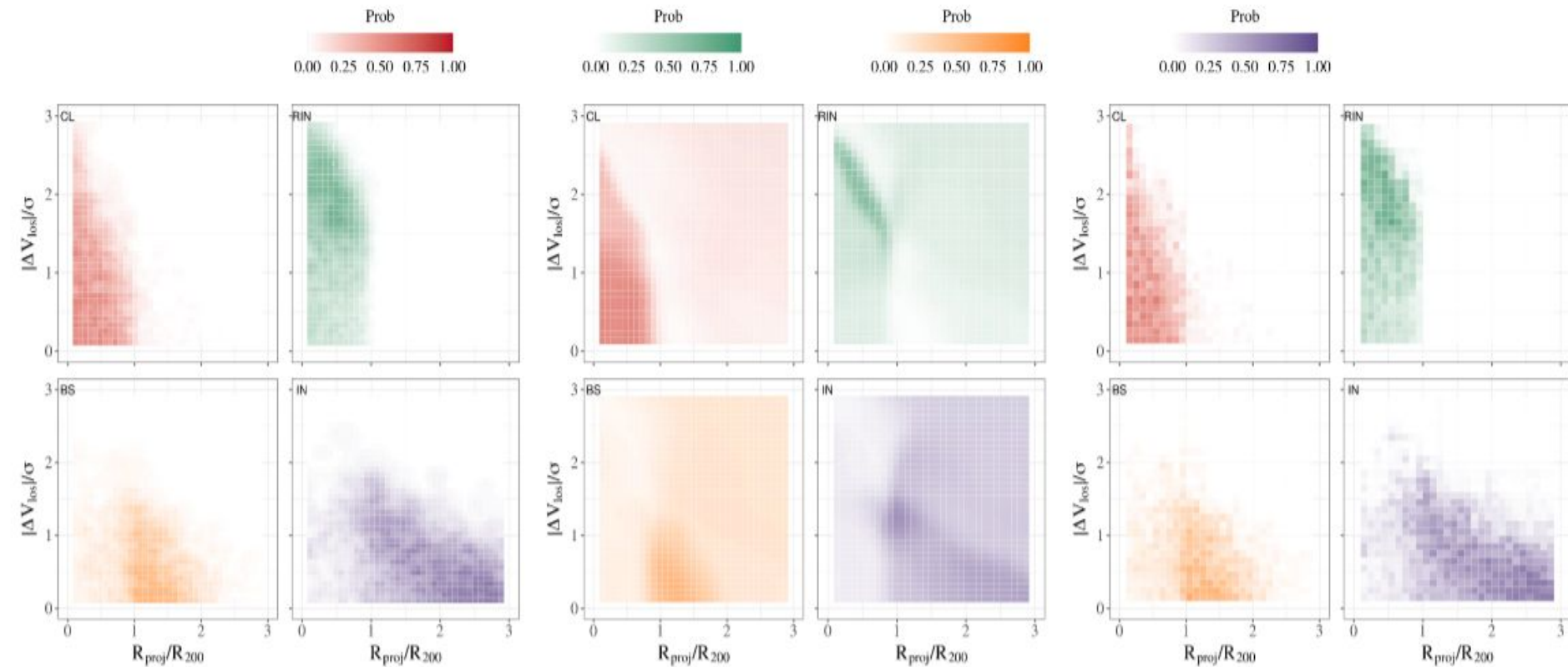
Set Entrenamiento



Set Testeo



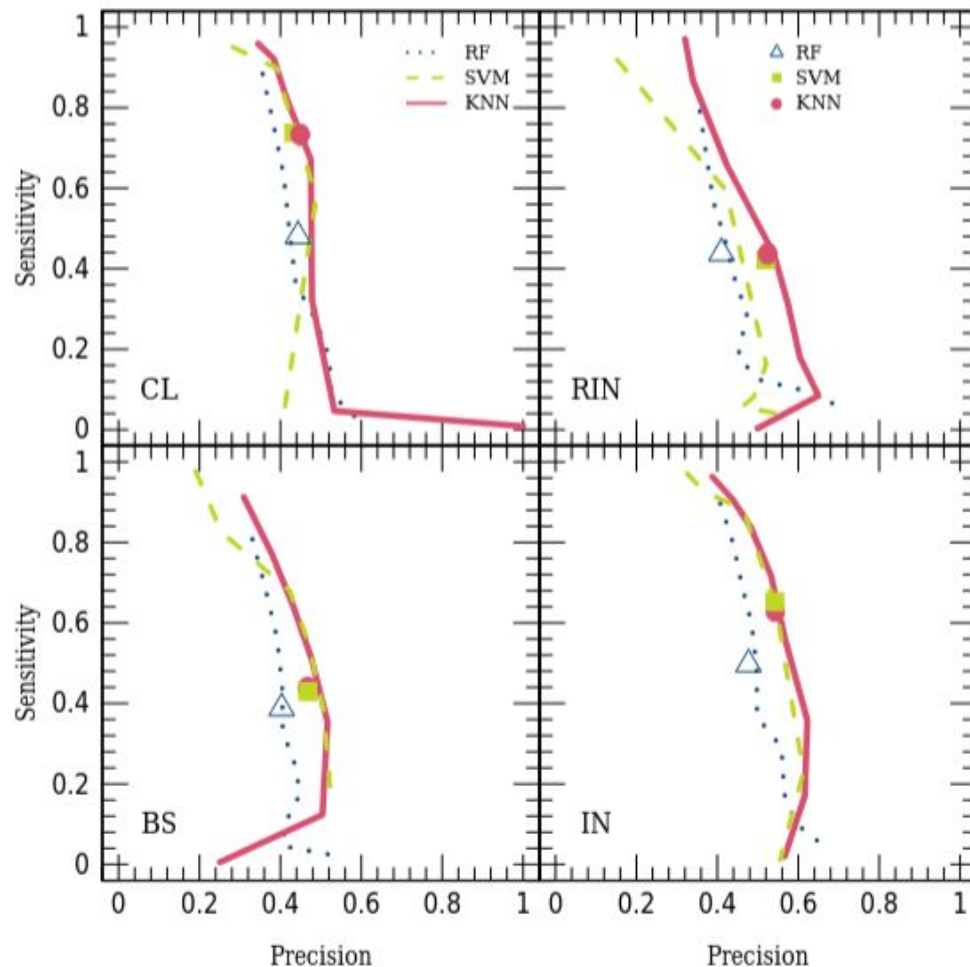
Probabilidades según la posición en el espacio de fases



(a) KNN method

(b) SVM method

(c) RF method

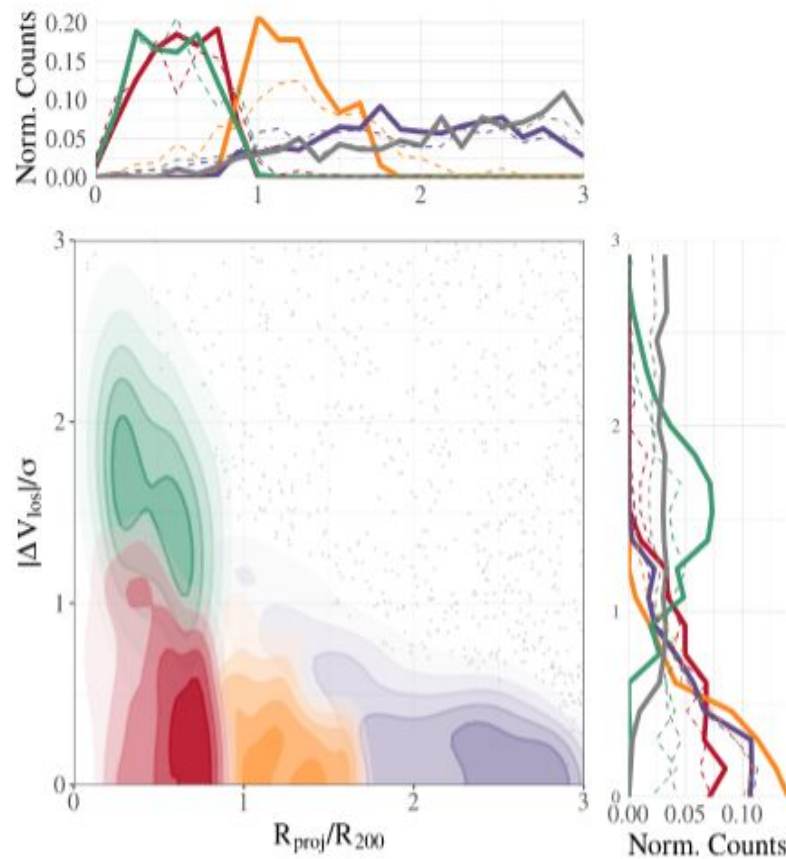


Clasificamos una galaxia como perteneciente a una dada clase, si la probabilidad de que esa galaxia pertenezca a dicha clase es mayor a un valor crítico.

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

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
	0% n=1	1% n=4	2% n=13	9% n=75	59% n=539
	CL	RIN	BS	IN	ITL
	Real				

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, backplash, 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/).



GRACIAS!