

# Estadísticas Globales de usos agrícolas en Europa con Sentinel e IA en Eurostat





## Cultivos



# Fuente de Estadísticas Agrícolas



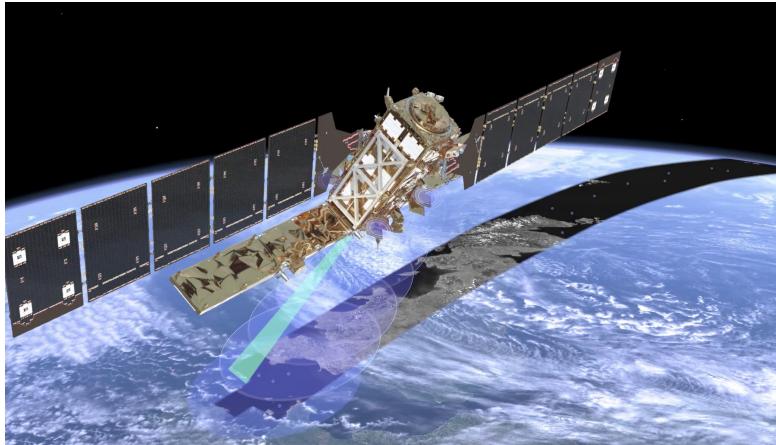


jornadas **siglibre**  
Geotech/spatial data science

# Transición hacia nuevas fuentes

*La revolución de los datos satélites y la IA*

**Sentinel 2-A**



**LPIS**



**LUCAS Survey**



## Modelos de clasificación



JOINT RESEARCH CENTRE



**Random Forest**



**European  
Environment  
Agency**



**Transformer-  
based Feature  
Extractor**

**CNN**



# ¿Son estas estadísticas fiables?



## El problema del Sesgo

### Sesgo datos de entrenamiento

- Subrepresentación ciertos cultivos minoritarios
- Distribución geográfica desigual
- Etiquetas incorrectas

### Sesgos temporales

- Diferencias en las fechas de adquisición de imágenes y puntos de LUCAS con etiquetas
- Falta de representación de variaciones climáticas interanuales.
- Sesgo por cobertura nubosa en primavera/verano

### Sesgos espaciales

- Diferencias de tamaño medio de parcela entre regiones
- Falta de cobertura homogénea en regiones montañosas o de difícil acceso.

...

*IA ≠ Realidad Perfecta*



## ¿Por qué son necesarias unas estadísticas agrarias de calidad?



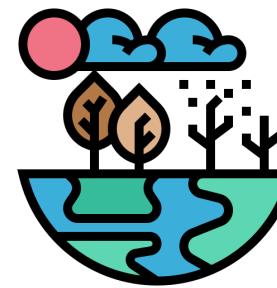
Planificación y  
política agrícola



Monitoreo de  
productividad y  
rendimientos



Gestión ambiental  
y sostenible



Cambio climático y  
adaptación

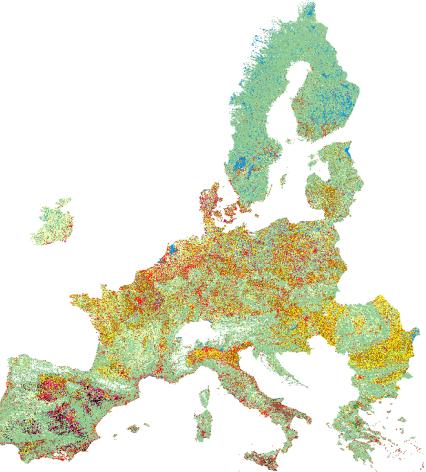


Mercado y  
economía



# ¿Cómo corregir el sesgo?

Mapa Clasificado



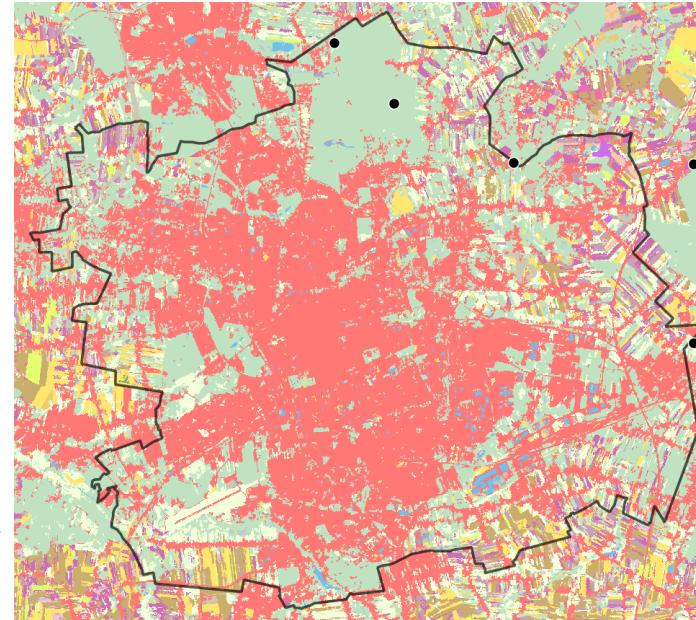
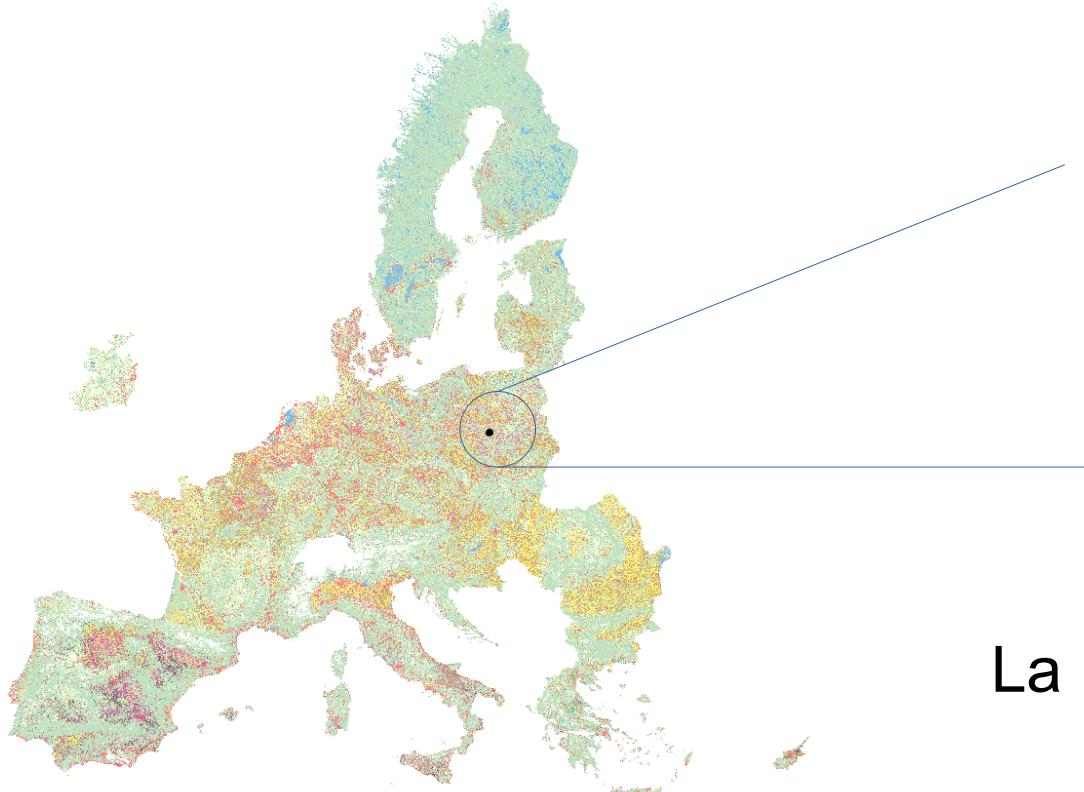
LUCAS Survey

Matriz de Confusión

		Confusion Matrix for Europe																													
		Reference																													
		300 - Woodland and shrubland type of vegetation																													
		600 - Bare land and lichens/moss																													
Predicted		100 - Artificial land	101 - Common wheat	102 - Durum wheat	103 - Barley	104 - Rye	105 - Oats	106 - Maize	107 - Rice	108 - Triticale	109 - Other cereals	110 - Potatoes	111 - Sugar beet	112 - Other non permanent-industrial crops	113 - Other roots crops	114 - Sunflower	115 - Rape and turnip rape	116 - Dry pulses, vegetables and flowers	117 - Soya	118 - Other fodder crops	119 - Bare arable land	120 - Grassland	121 - Wetlands	122 - Water	Total						
		1649	4	9	7	0	1	14	0	1	1	8	1	1	4	4	4	2	24	9	25	449	567	51	9	1	2847				
		99	1717	81	306	96	71	36	0	97	12	6	5	10	61	7	123	4	30	73	188	75	356	8	7	4	3472				
		38	25	87	28	0	14	1	0	3	1	0	0	0	2	3	1	0	9	20	38	27	72	1	2	1	373				
		12	64	20	466	16	31	3	0	24	4	0	2	3	10	4	39	1	7	20	84	18	78	2	0	0	908				
		8	17	0	17	83	6	6	0	36	2	0	0	3	1	0	2	0	3	5	9	5	39	2	0	0	244				
		7	35	5	49	7	99	3	0	12	3	0	0	0	2	0	0	0	3	12	15	12	73	0	0	0	337				
		118	33	1	18	6	9	1188	5	7	13	10	5	9	20	24	7	37	40	23	22	71	223	5	1	2	1897				
		0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1			
		0	2	0	7	15	2	1	0	28	0	0	0	0	0	1	0	1	0	1	3	0	3	0	0	0	0	63			
		0	0	0	0	0	0	1	0	0	1	8	0	0	0	1	0	0	1	1	3	2	1	6	0	0	0	25			
		7	7	0	0	1	0	1	0	0	1	118	8	1	3	10	2	1	6	3	8	2	13	0	0	0	192				
		13	5	0	0	0	1	1	0	0	1	3	149	7	3	5	2	0	8	4	3	7	10	0	0	0	222				
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8			
		11	1	0	1	0	0	0	0	0	0	0	0	1	39	2	0	0	6	3	1	4	8	0	0	0	0	77			
		20	4	2	2	0	0	9	0	0	0	9	2	4	5	264	3	1	18	6	13	17	29	1	0	0	409				
		18	29	1	3	3	0	1	0	4	2	1	1	4	2	5	427	2	7	4	24	9	41	1	0	0	589				
		4	2	0	0	0	0	4	0	0	1	1	0	0	1	2	0	52	1	2	2	2	8	1	0	0	83				
		3	7	1	0	0	0	2	1	0	0	2	6	2	6	7	0	8	0	67	4	14	3	21	0	0	0	154			
		20	7	5	6	1	4	1	0	1	0	1	0	1	2	0	0	0	3	175	12	8	101	1	0	0	349				
		24	2	3	11	0	10	0	0	0	0	1	1	0	5	0	0	9	10	172	43	56	9	0	0	0	357				
		973	20	5	25	7	14	22	0	8	3	5	0	3	2	5	5	1	17	25	49	9909	1426	62	18	75	12679				
		338	34	8	16	14	14	21	0	8	1	4	1	1	4	4	4	0	13	240	24	604	5735	15	7	26	7136				
		18	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	5	10	13	19	1	1	68				
		11	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	12	21	11	20	3	81			
		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	55	83			
		3392	2015	228	962	250	281	1313	5	230	55	175	178	61	169	345	628	102	272	643	713	11296	8917	189	66	169	32659				

Estimador  
insesgado

## Pero hay una limitación... ¿Y en áreas pequeñas?



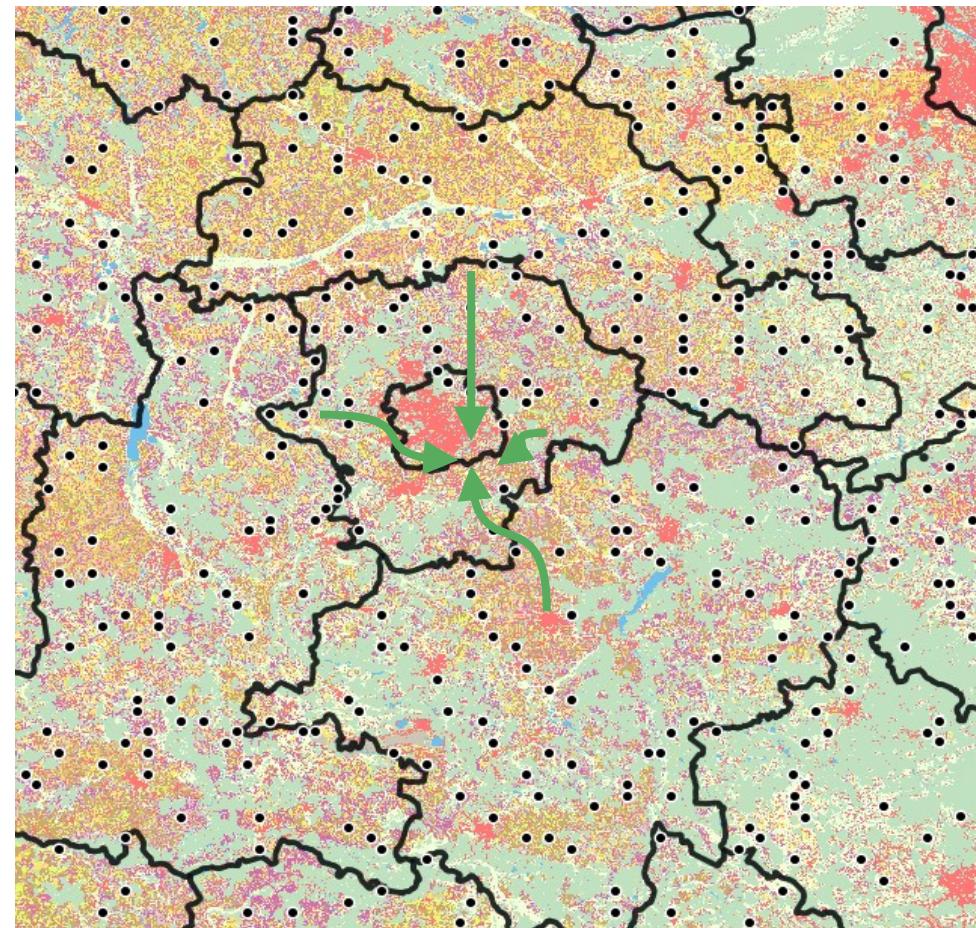
La muestra de puntos es demasiado pequeña

↓  
Estimaciones inestables



## Small Area Estimator (SAE)

*Prestando fuerza de las áreas vecinas*



## EBLUP en acción

- ✓ Aprovecha datos de referencia de áreas cercanas
- ✓ Mejora la estabilidad en áreas pequeñas
  
- ✗ Solo se basa en la clasificación más probable
- ✗ No mantiene proporciones



## Un paso más allá: Modelos Multinomiales

*Unit-Level Logistic Multinomial Mixed Model*



**CLAVE:** Considera la variable explicativa como una variable categórica

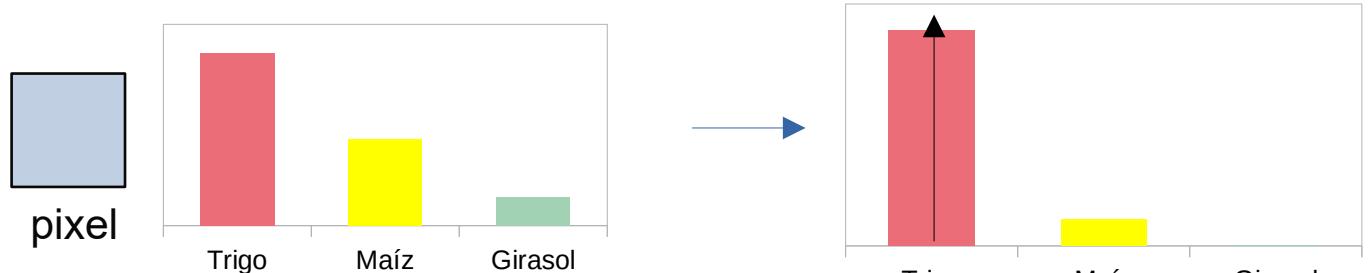


Las variables categóricas no toman valores numéricos continuos, sino que representan categorías o clases.



Cada observación (pixel) pertenece a una y solo una categoría

## ¿Cómo funciona?

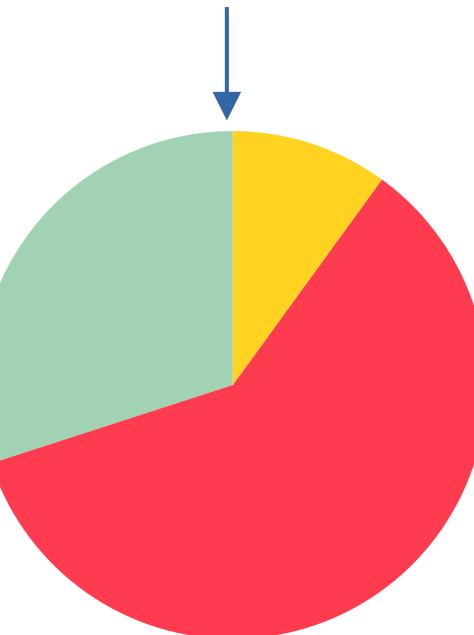


**Referencia = Trigo**



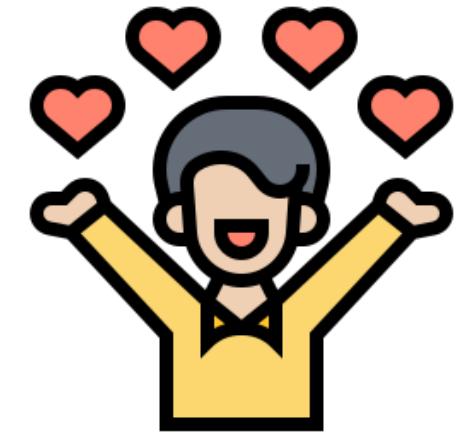
**Referencia = Maíz**

Al corregir pixel por pixel,  
conseguimos proporciones  
más insesgadas de la zona



## Ventajas de los Modelos Multinomiales

- ✓ Utiliza las probabilidades de clasificación
- ✓ Todas las categorías suman el 100% del área total
- ✓ Las proporciones son coherentes globales y localmente
- ✓ El dato se da a nivel de pixel
- ✓ Proporciona un cálculo de error y, por tanto, de exactitud de la cifra





jornadassiglibre  
Geotech/spatial data science

Universitat de Girona  
Servei de Sistemes d'Informació  
Geogràfica i Teledetecció

# Muchas gracias

