



Comparación de árboles de decisión CART y bajo criterio experto en la clasificación de una zona de riego extensa

**Autores: Mónica Guillén, Marcos Rodrigues,
Miquel Febrer, M^a Auxiliadora Casterad**

¿POR QUÉ?

Conocer la **superficie de cultivos** es clave para en la **gestión de las zonas regables**. Los gestores de las comunidades de regantes demandan estrategias que permitan optimizar el uso del agua.

Los **nuevos modelos y estrategias de gestión** que están adoptando Comunidades de Regantes, Confederaciones Hidrográficas, Organismo públicos, ... tienen como **protagonistas las imágenes de satélite**. Se están invirtiendo grandes esfuerzos en la elaboración de **cartografía de cultivos precisa y ajustada a la realidad**.

Los nuevos métodos de clasificación basados en algoritmos **“Machine Learning”** pueden contribuir a mejorar y agilizar la clasificación a la vista de los resultados obtenidos en diferentes zonas agrícolas (Rodríguez-Galiano y Chica-Rivas, 2012; Toro y col., 2015; Larrañaga y Álvarez-Mozos, 2016; Gomariz-Castillo y col., 2017; Siontstantinou y col., 2018).



CITA
CENTRO DE INVESTIGACIÓN Y TECNOLOGÍA
AGROALIMENTARIA Y AGROPECUARIA



Universitat de Lleida
Departament d'Enginyeria
Agroforestal



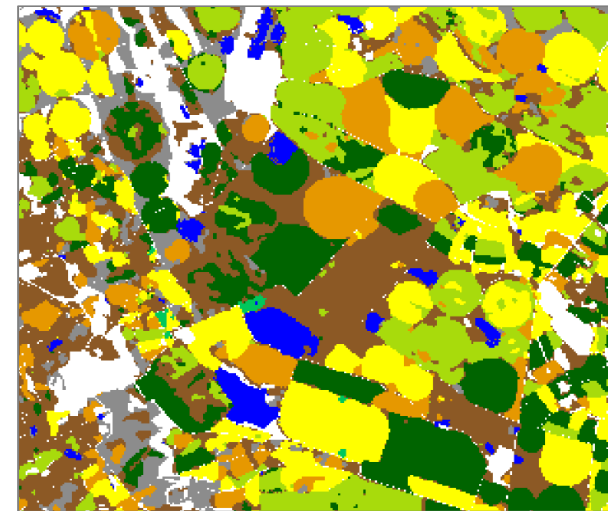
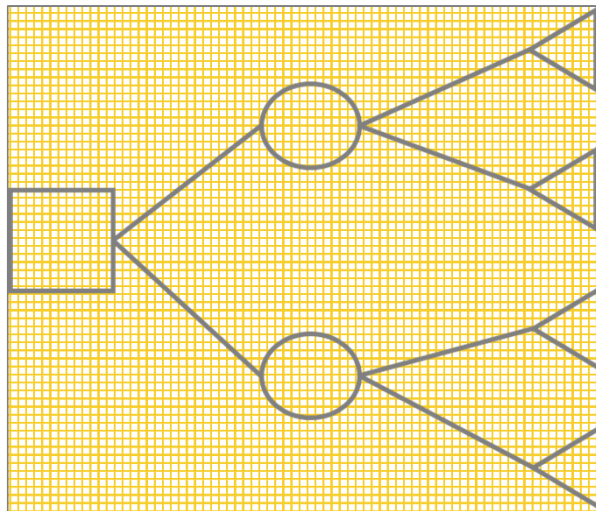
RemOT
Technologies



GOBIERNO
DE ARAGON

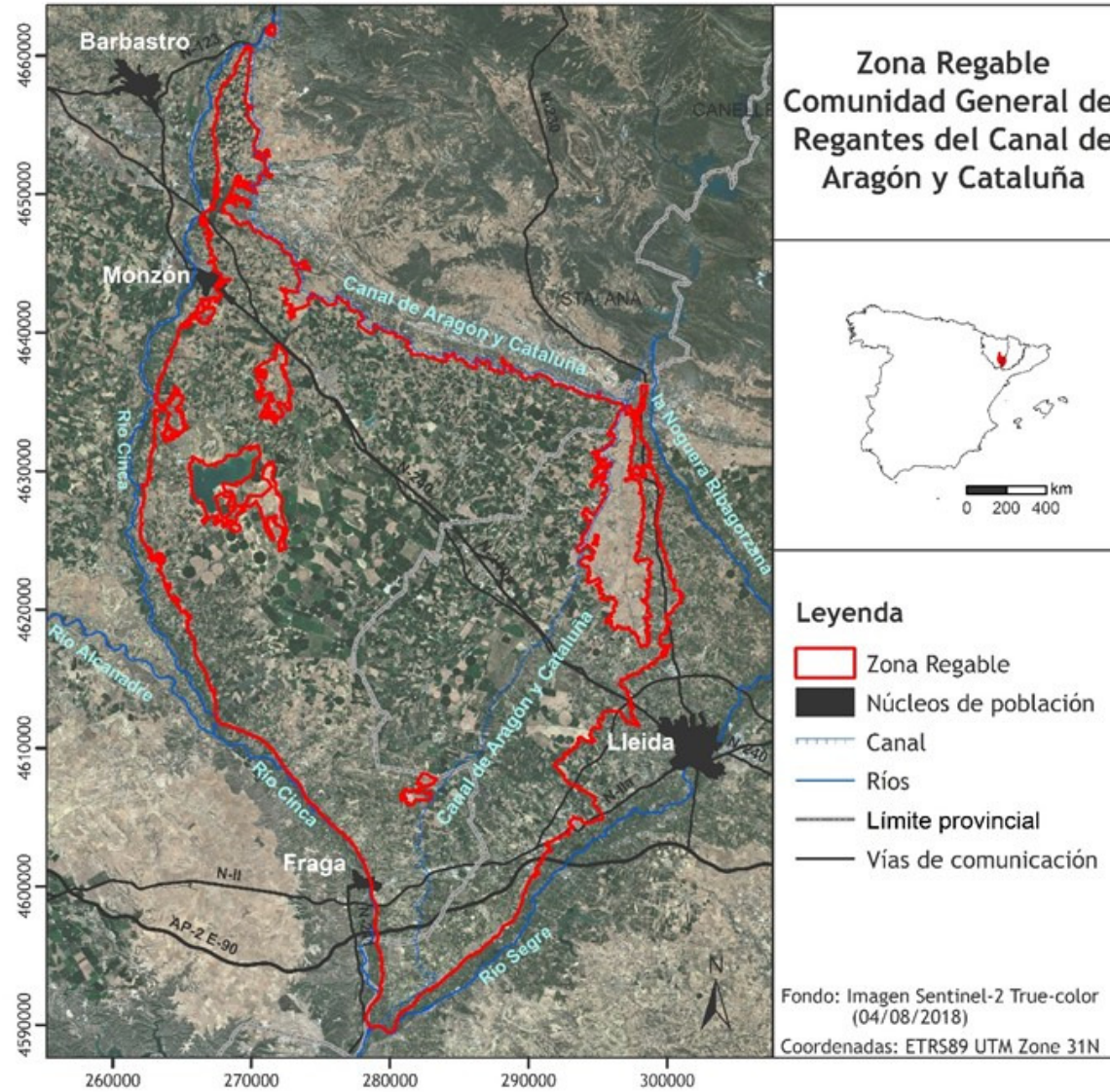
Departamento de Ciencia,
Universidad,
y Sociedad del Conocimiento

El **objetivo** del trabajo es **comparar y evaluar** desde un punto de vista **metodológico y predictivo** la aplicación de árboles de decisión **CART** y árboles de decisión **bajo criterio experto** en la clasificación de tipologías de **cultivos** en la Zona Regable de la Comunidad General de Regantes del Canal de Aragón y Cataluña (CAyC) .



¿DÓNDE?

Zona Regable del CGRCAYC



CITA
CENTRO DE INVESTIGACIÓN Y TECNOLOGÍA
AGROALIMENTARIA DE ARAGÓN



Universitat de Lleida
Departament d'Enginyeria
Agroforestal

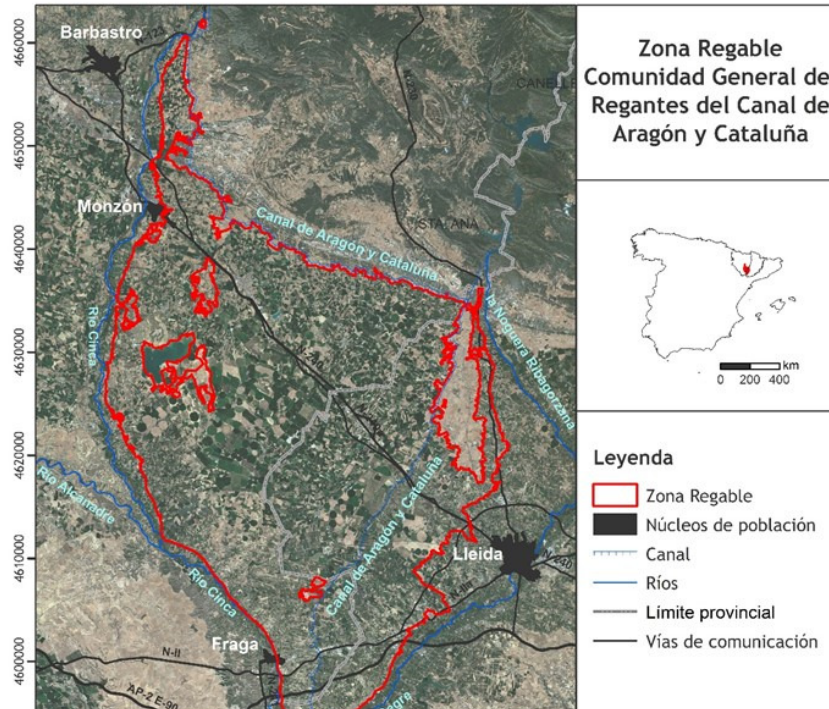


RemOT
Technologies



GOBIERNO
DE ARAGON
Departamento de Ciencia, Universidad,
y Sociedad del Conocimiento

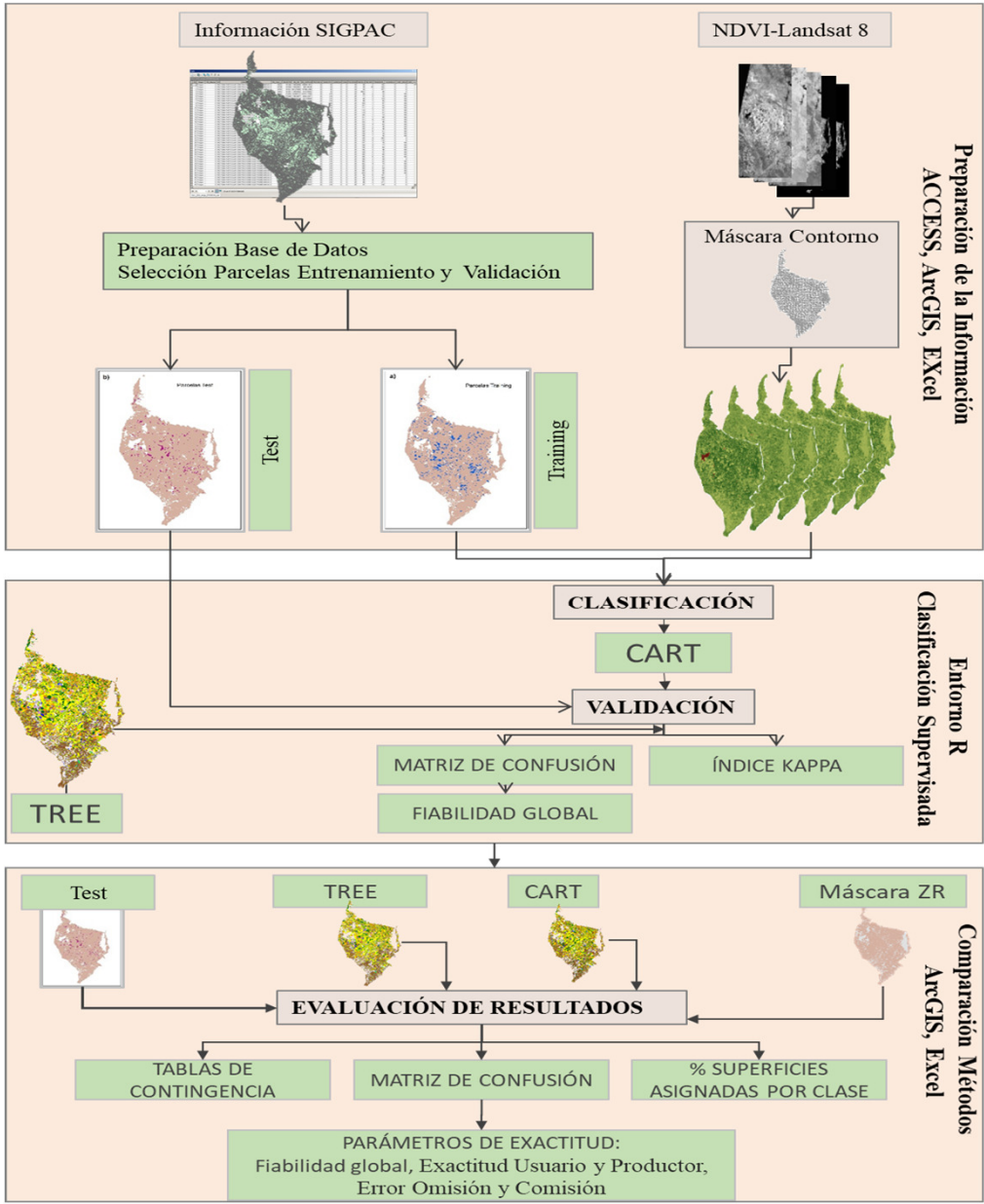
Zona Regable del CGRCAYC



- Su gran extensión:
 - ✓ 105.000 hectáreas
 - 60 % en Aragón
 - 40 % en Cataluña
- Diversidad de sistemas de riegos
 - ✓ 50 % riego por aspersión
 - ✓ 27 % riego por gravedad
 - ✓ 23 % riego por goteo
- Diversidad de cultivos

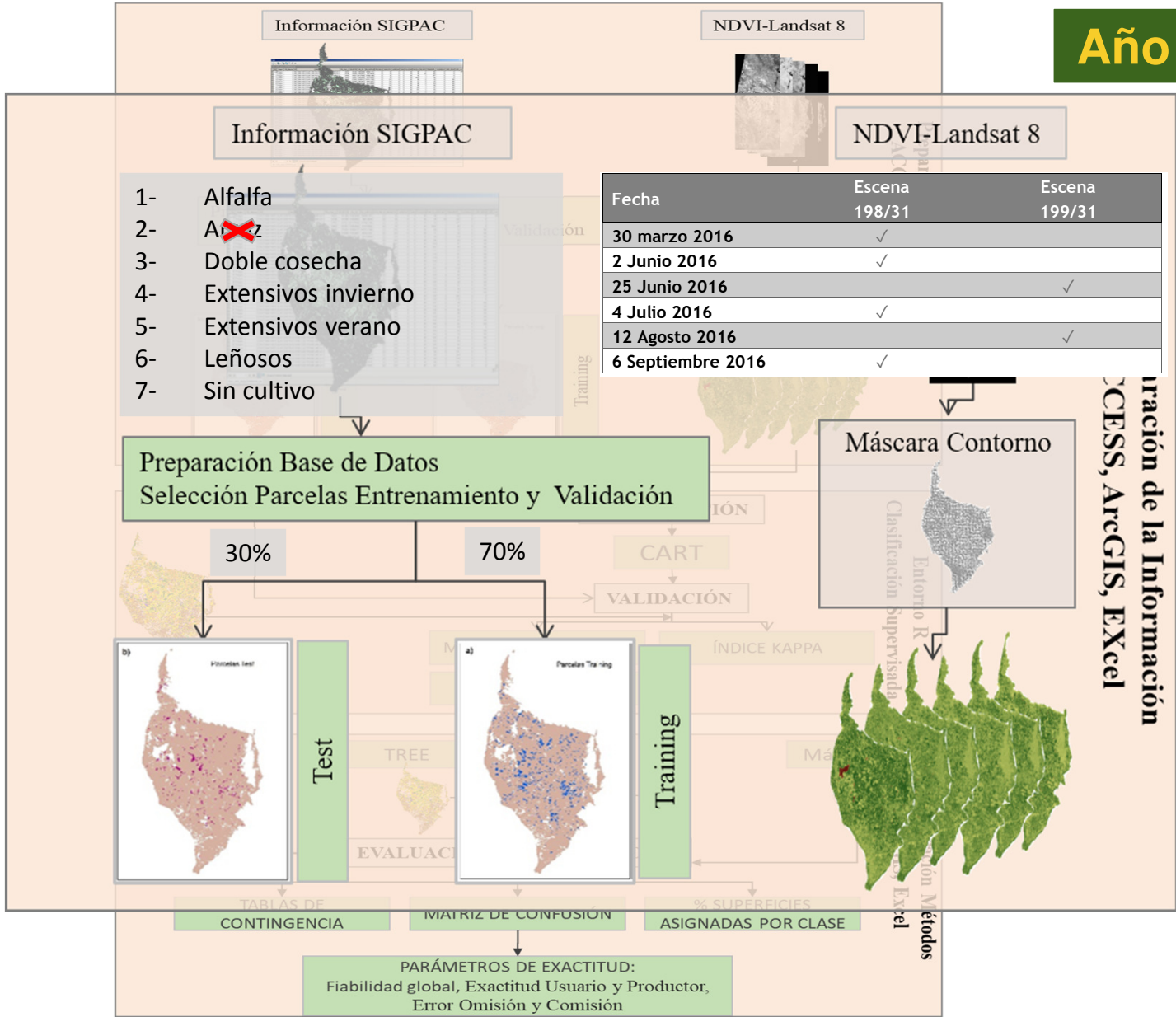
- Experiencias previas de identificación de cultivos
 - **clasificación supervisada** a través de **árboles de decisión bajo criterio experto**
 - basado en la evolución del índice de vegetación **NDVI**.
- Interés en mejorar y agilizar la **monitorización** con teledetección de su **superficie cultivada**.

Flujo de trabajo



Año 2016

Flujo de trabajo

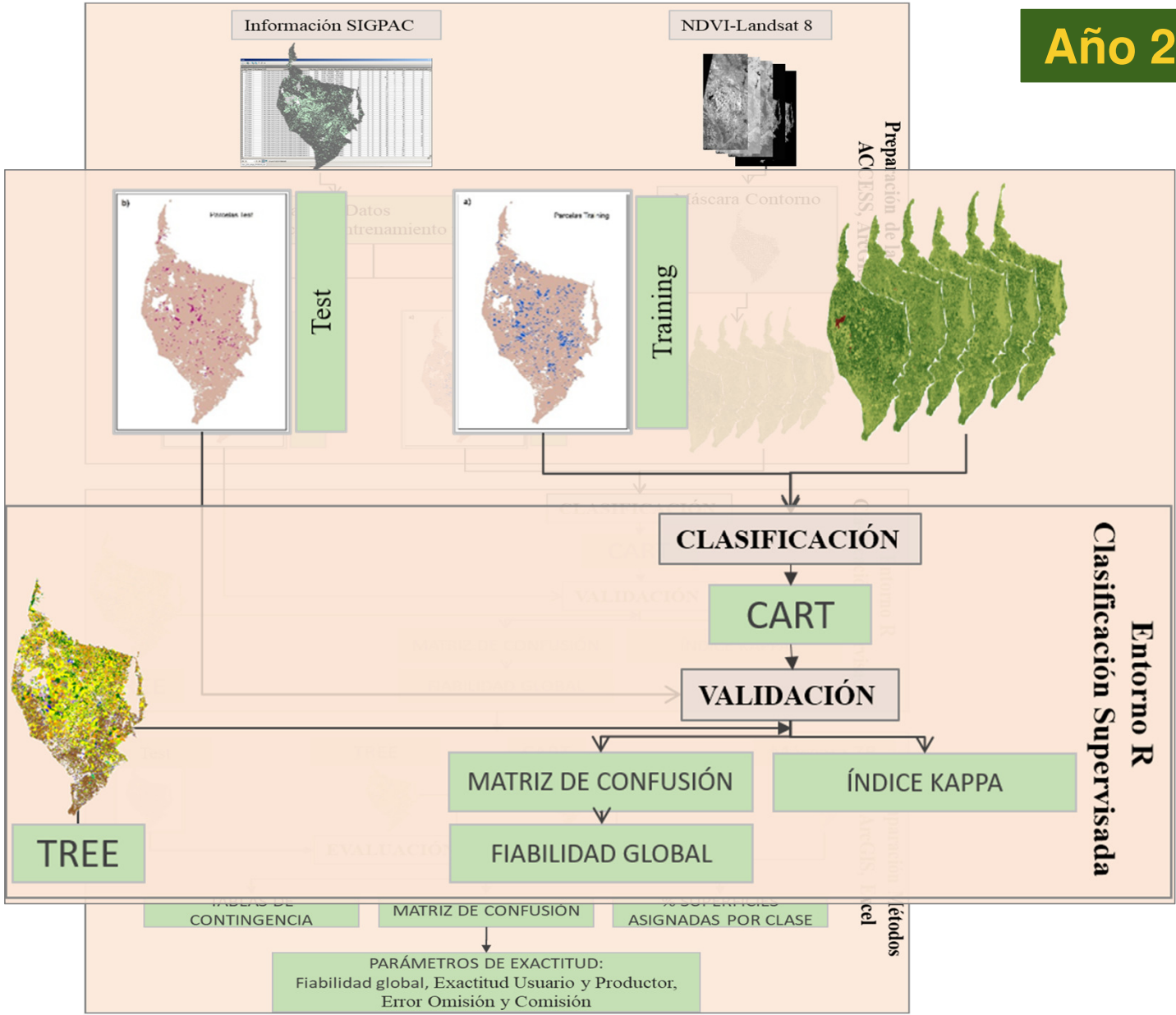


¿CÓMO?



Año 2016

Flujo de trabajo

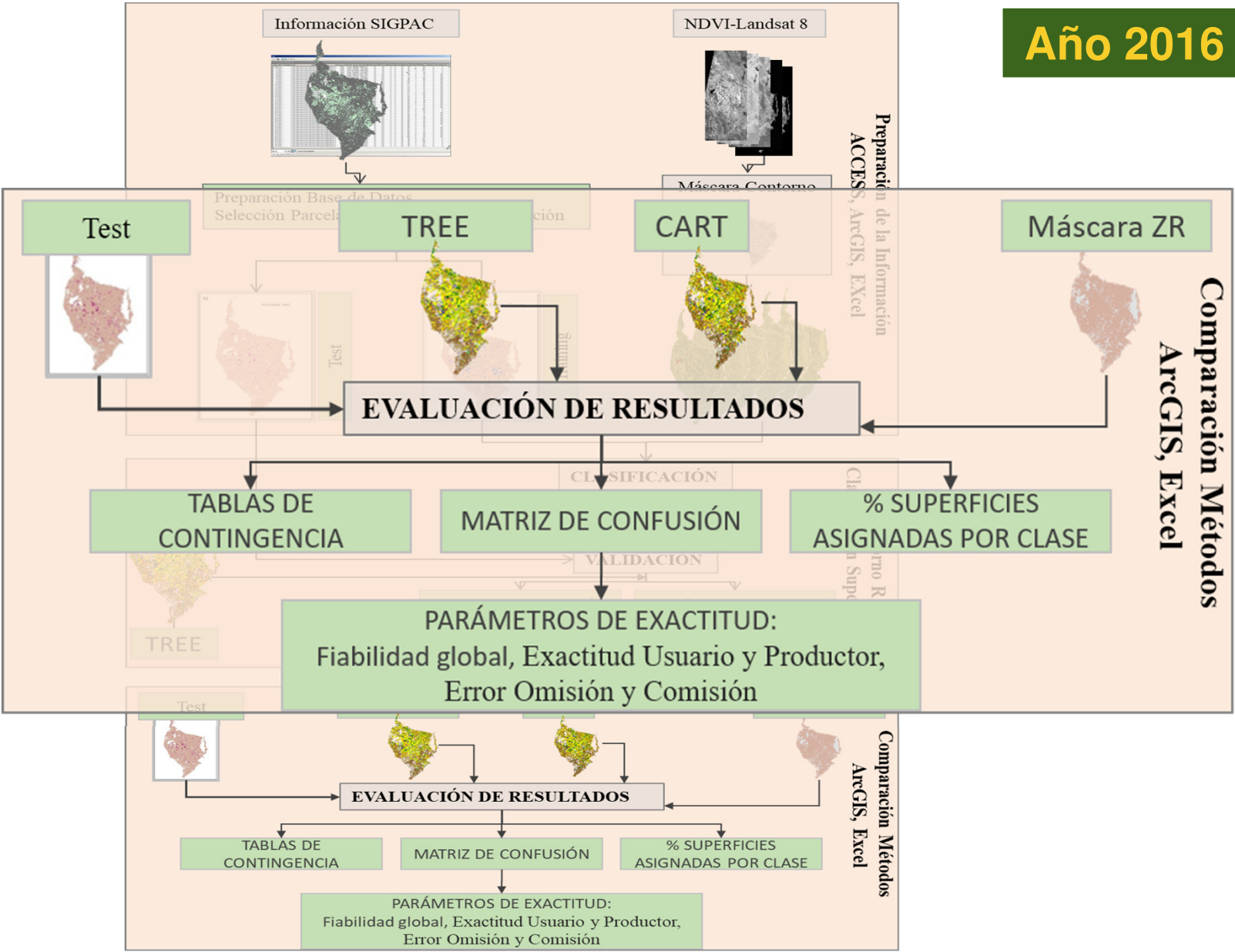


¿CÓMO?

Logos of participating institutions: cita (Centro de Investigación y Tecnología Agroforestal, Instituto de Investigación y Tecnología Agraria y Alimentaria), Universitat de Lleida Departament d'Enginyeria Agroforestal, Geoforest, RemOT Technologies, and GOBIERNO DE ARAGON (Departamento de Ciencia, Universidad, y Sociedad del Conocimiento).

Año 2016

Flujo de trabajo



¿CÓMO?



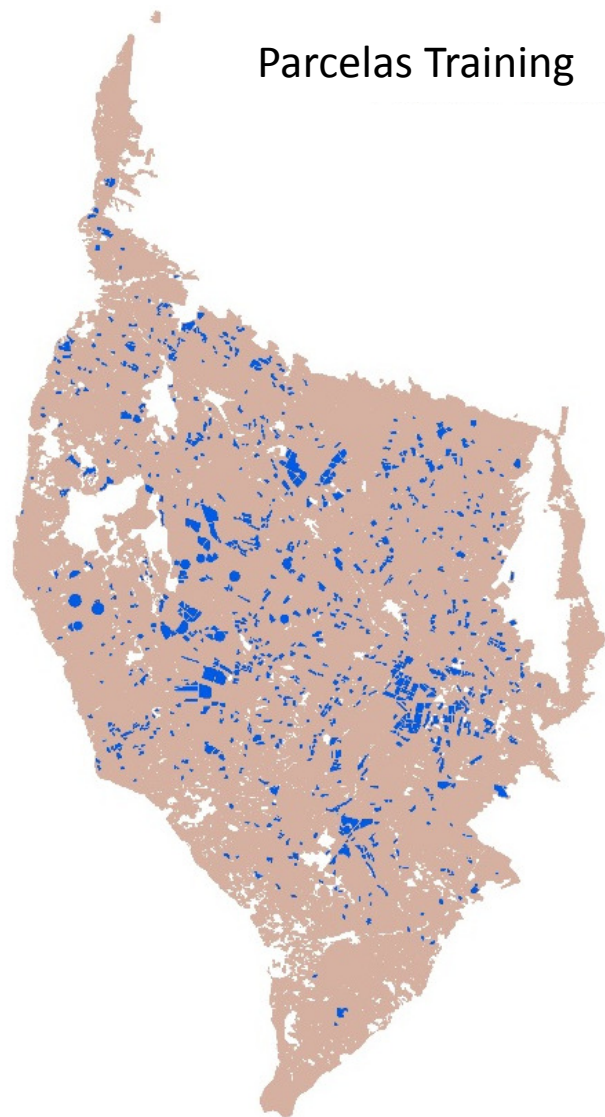
Universitat de Lleida
Departament d'Enginyeria
Agroforestal



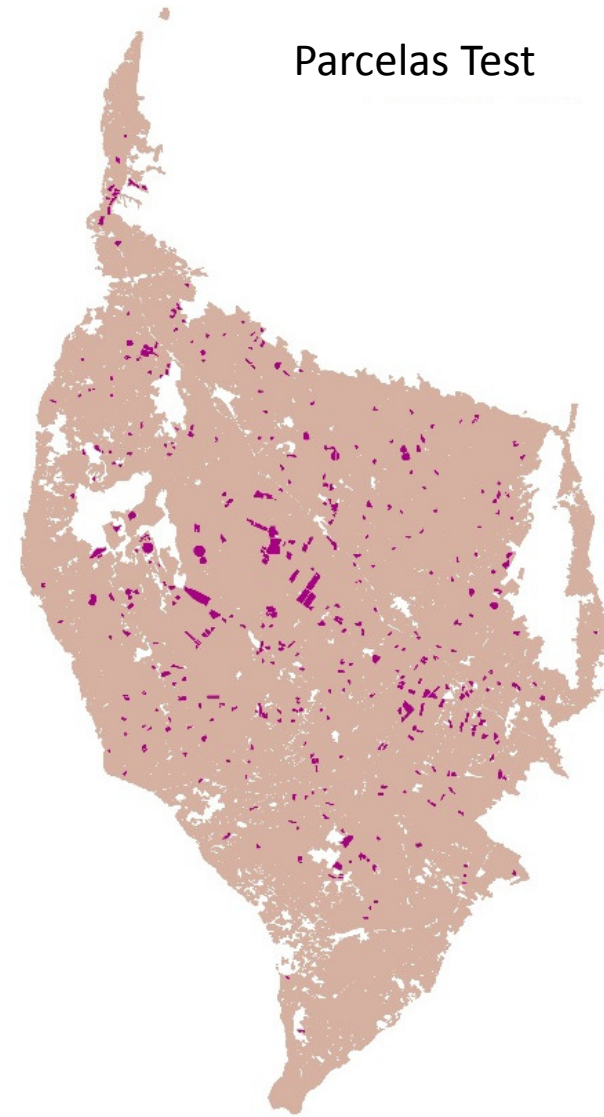
Departamento de Ciencia, Universidad,
y Sociedad del Conocimiento

VT seleccionada

Parcelas Training



Parcelas Test



¿CÓMO?



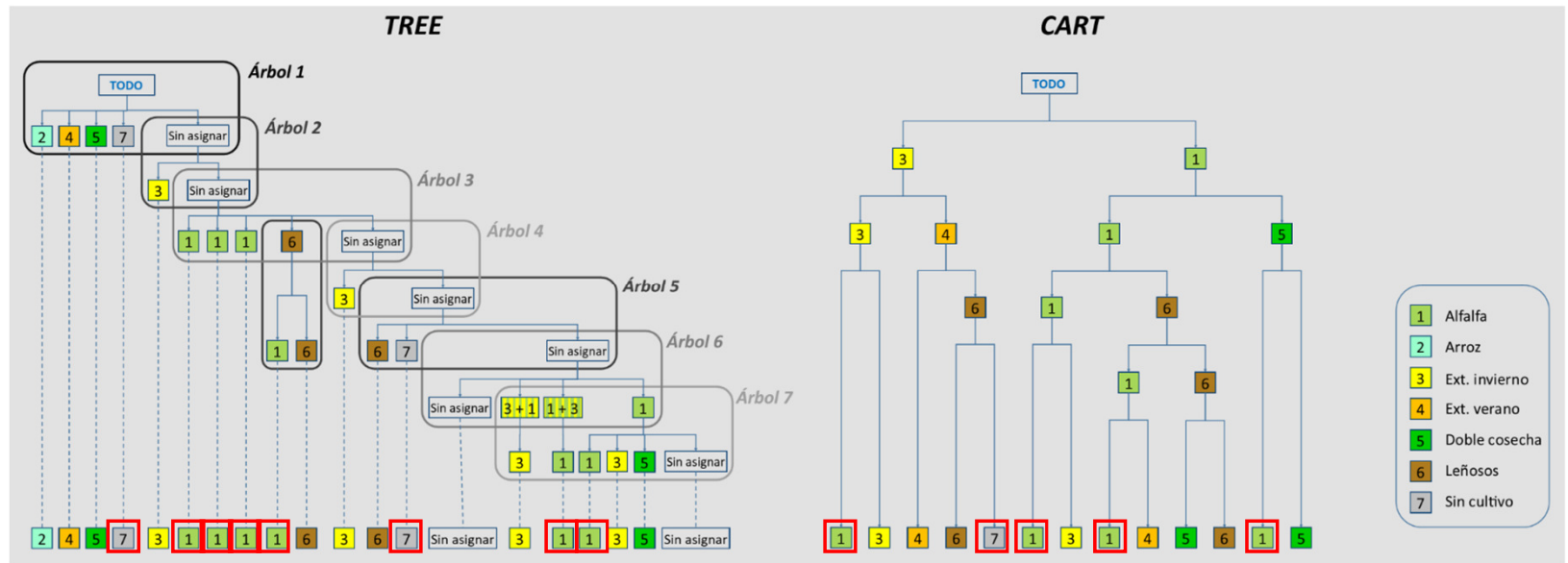
Universitat de Lleida
Departament d'Enginyeria
Agroforestal



GOBIERNO DE ARAGON
Departamento de Ciencia, Universidad,
y Sociedad del Conocimiento

¿QUÉ SE OBTIENE?

Árboles de Decisión



Importancia de las variables

CART

fecha_6r	fecha_5r	fecha_1r	fecha_4r	fecha_3r	fecha_2r
6/09/2016	12/08/2016	30/03/2016	4/07/2016	25/06/2016	2/06/2016
21	19	18	16	15	11

Verificación de resultados

	CART	TREE
\hat{F}	0.911	0.835
κ	0.889	--



¿QUÉ SE OBTIENE?

TREE

		VERDAD TERRENO							Exactitud Usuario (%)	Error Comisión (%)
NºParcelas		ALF	DC	EI	EV	LEÑ	SC	TOTAL		
TREE	Alfalfa	65	1	0	7	3	1	77	84.4%	15.6%
	Doble Cosecha	1	80	0	0	0	1	82	97.6%	2.4%
	Ext. Invierno	0	1	61	0	1	1	64	95.3%	4.7%
	Ext. Verano	1	0	0	48	0	0	49	98.0%	2.0%
	Leñoso	1	0	0	0	61	0	62	98.4%	1.6%
	Sin Cultivo	0	0	0	0	6	4	10	40.0%	60.0%
	Sin Identificar	23	3	2	0	9	1	38		
	TOTAL	91	85	63	55	80	8	382		

$\hat{F} = 84\%$

Exactitud Productor(%)	71.4%	94.1%	96.8%	87.3%	76.3%	50.0%
Error Omisión (%)	28.6%	5.9%	3.2%	12.7%	23.8%	50.0%

CART

		VERDAD TERRENO							Exactitud Usuario (%)	Error Comisión (%)
NºParcelas		ALF	DC	EI	EV	LEÑ	SC	TOTAL		
CART	Alfalfa	82	5	0	0	4	0	91	90.1%	9.9%
	Doble Cosecha	2	79	0	2	0	0	83	95.2%	4.8%
	Ext. Invierno	1	1	62	0	0	1	65	95.4%	4.6%
	Ext. Verano	2	0	0	52	5	1	60	86.7%	13.3%
	Leñoso	4	0	1	1	68	1	75	90.7%	9.3%
	Sin Cultivo	0	0	0	0	3	5	8	62.5%	37.5%
	TOTAL	91	85	63	55	80	8	382		

$\hat{F} = 91\%$

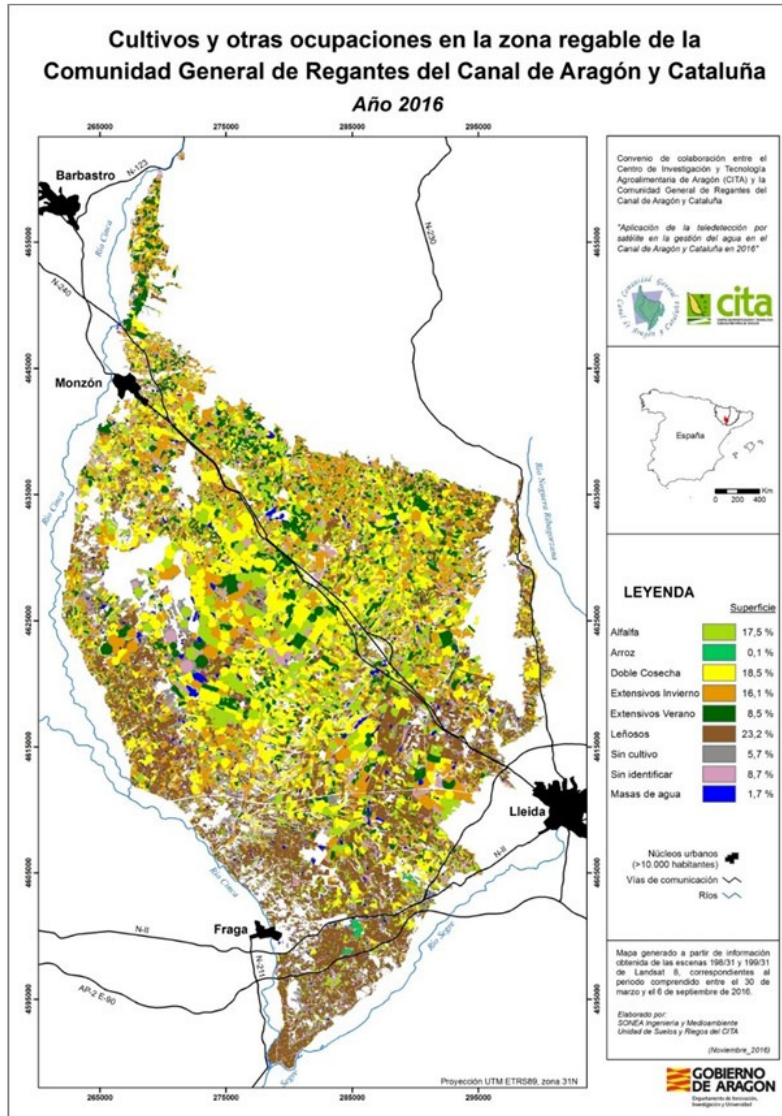
Exactitud Productor(%)	90.1%	92.9%	98.4%	94.5%	85.0%	62.5%
Error Omisión (%)	9.9%	7.1%	1.6%	5.5%	15.0%	37.5%



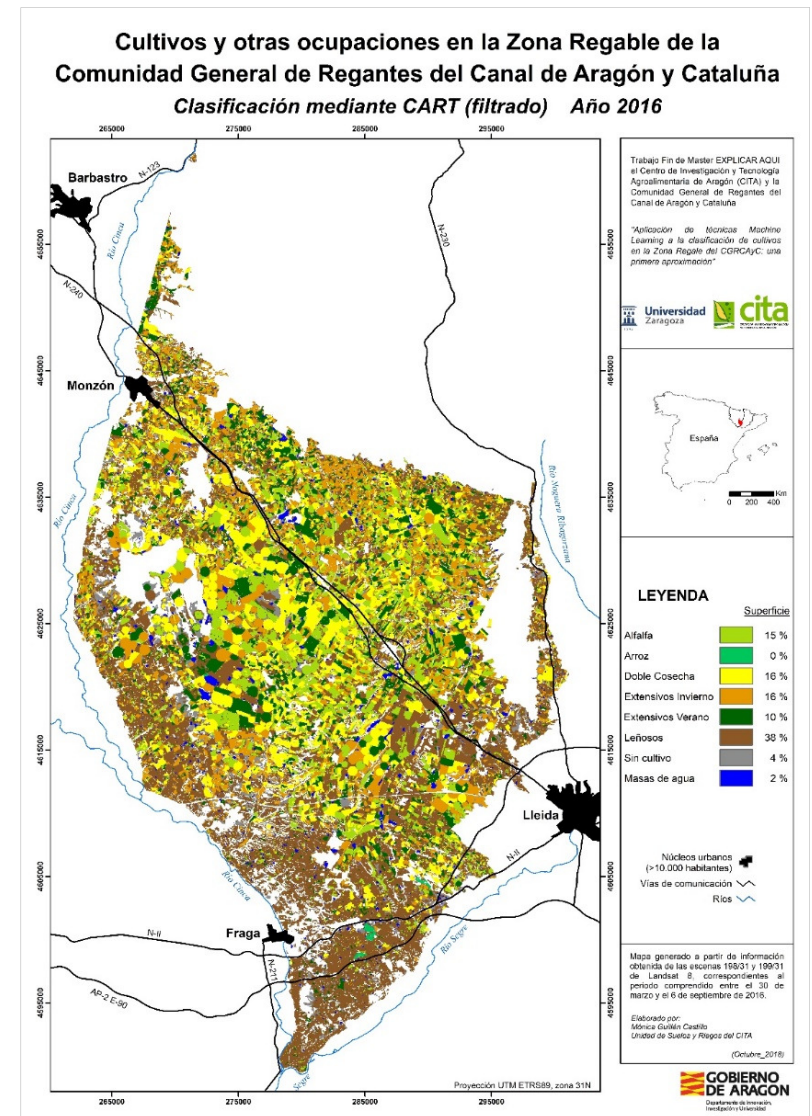
¿QUÉ SE OBTIENE?

Clasificaciones generadas

TREE



CART



¿QUÉ SE OBTIENE?

Clasificaciones generadas

TREE

CART

Cultivos y otras ocupaciones en la zona regable de la Comunidad General de Regantes del Canal de Aragón y Cataluña

Porcentaje de la superficie total clasificada de cada cultivo y ocupación

Cultivos y otras ocupaciones en la Zona Regable de la Canal de Aragón y Cataluña (filtrado) Año 2016

	TREE	CART
Alfalfa	17.8%	14.6%
Arroz	0.1%	0.1%
Doble Cosecha	18.8%	16.2%
Ext. Invierno	16.1%	15.8%
Ext. Verano	8.5%	9.9%
Leñosos	23.7%	37.6%
Sin Cultivo	5.7%	4.0%
Masas de Agua	1.7%	1.7%
Sin Indentificar	7.6%	--

Trabajo Fin de Master EXPLICAR AQUÍ el Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA) y la Comunidad General de Regantes del Canal de Aragón y Cataluña

Aplicación de Algoritmo Machine Learning a la clasificación de cultivos en la Zona Regable del CGRCAyC: una primera aproximación

Universidad Zaragoza 



LEYENDA

Superficie	Porcentaje
Alfalfa	15 %
Arroz	0 %
Doble Cosecha	16 %
Extensivos invierno	16 %
Extensivos Verano	10 %
Leñosos	38 %
Sin cultivo	4 %
Masas de agua	2 %

Mapa generado a partir de información obtenida de las escenas 19831 y 19931 de Landsat 5, correspondientes al período comprendido entre el 30 de marzo y el 6 de septiembre de 2016.

Elaborado por: Mónica Guzmán Casallo
Unidad de Suelo y Riego del CITA
(Octubre 2016)

 GOBIERNO DE ARAGON
Departamento de Ingeniería Agroforestal



Universitat de Lleida
Departament d'Enginyeria
Agroforestal



 GOBIERNO DE ARAGON
Departamento de Ciencia, Universidad,
y Sociedad del Conocimiento

- ❑ Los **dos modelos** comparados en este trabajo proporcionan una **muy buena clasificación**, si bien **CART** ($\hat{F} = 0,91$) **mejora** la fiabilidad global de TREE ($\hat{F} = 0,84$).
- ❑ **Todas las variables** elegidas, 6 fechas diferentes de valor NDVI obtenidos de imágenes Landsat 8, **son importantes** para los modelos, lo que entra dentro de lo **esperado** ya que estas fechas corresponden a una **selección previa** hecha de la experiencia.
- ❑ Los dos clasificadores coinciden en que la categoría **Alfalfa** es la que a través de **más nodos terminales se discrimina**, mostrando la complejidad de su clasificación. La categoría **Sin Cultivo** es la que con **menos nodos terminales** se discrimina y peores fiabilidades consigue.



CITA
CENTRO DE INVESTIGACIÓN Y TECNOLOGÍA
AGROALIMENTARIA DE ARAGÓN



Universitat de Lleida
Departament d'Enginyeria
Agroforestal



RemOT
Technologies



GOBIERNO
DE ARAGON
Departamento de Ciencia, Universidad,
y Sociedad del Conocimiento

- ❑ El **modelo CART** es **más sencillo de interpretar** que TREE y permite **identificar** de forma más **rápida y eficiente** las **variables** más importantes. Además, toma las **reglas de decisión** de manera **objetiva**, no así en TREE que depende del criterio experto.
- ❑ La metodología **TREE** permite **clasificar** aquellas **categorías con poca** representación de **verdad terreno**. No necesita de verdad terreno, ya que permite diferenciar clases o categorías espectrales similares a partir de la comparación visual de las distintas firmas.
- ❑ A diferencia de **CART**, que clasifica todos los píxeles que se introducen en el modelo, **TREE deja píxeles sin identificar** si no llega a cumplir todas las condiciones impuestas.

CITA
CENTRO DE INVESTIGACIÓN Y TECNOLOGÍA
AGROALIMENTARIA DE ARAGÓNUniversitat de Lleida
Departament d'Enginyeria
AgroforestalRemOT
TechnologiesGOBIERNO
DE ARAGONDepartamento de Ciencia, Universidad,
y Sociedad del Conocimiento

A
D
E
M
Á
S

Comparación de Clasificadores

	CART	RF	TREE-2016
Identificar variables importantes de forma rápida y eficaz	○ ○ ○	○ ○ ○	○
Reglas de decisión objetivas	○ ○ ○	○ ○ ○	○
Sencillez del modelo	○ ○ ○	○	○ ○
Rapidez en ejecución	○ ○ ○	○ ○ ○	○
Fiabilidad del método (orden)	○ ○	○ ○ ○	○
Permitir dejar pixeles sin identificar	—	—	○ ○ ○
Imprescindible parcelas de entrenamiento	○ ○ ○	○ ○ ○	○
Imprescindibles parcelas de validación	○ ○ ○	○	○ ○ ○
Influencia de poca verdad terreno	○ ○ ○	○ ○ ○	○
Influencia de poca disponibilidad de imágenes	○ ○ ○	○ ○ ○	○





Comparación de árboles de decisión CART y bajo criterio experto en la clasificación de una zona de riego extensa

**Autores: Mónica Guillén, Marcos Rodrigues,
Miquel Febrer, M^a Auxiliadora Casterad**

mguillenc@aragon.es



Agradecimiento: A la Comunidad General de Regantes del Canal de Aragón y Cataluña por apoyar estos trabajos con dos contratos de colaboración con el CITA (2016 y PDR-2016-2018).