

Cartografía del abandono de parcelas en regadíos tradicionales de ribera con datos de Sentinel 2 y la PAC

Mónica Guillén⁽¹⁾, Rosa Gómez-Báguena⁽¹⁾, Miquel Febrer⁽³⁾, Lucía Martínez-Cebrián⁽⁴⁾, Raquel Salvador^(1,2), M^a Auxiliadora Casterad⁽¹⁾

⁽¹⁾ Dpto. de Sistemas Agrícolas, Forestales y Medio Ambiente (Unidad asociada a EEAD-CSIC Suelos y Riegos), Centro de Investigación y Tecnologías Agroalimentaria de Aragón, Avda. Montañana 930, 50059-Zaragoza, España. mguillenc@cita-aragon.es, rgomezba@cita-aragon.es, rsalvadore@cita-aragon.es, acasterad@cita-aragon.es

⁽²⁾ Instituto Agroalimentario de Aragón-IA2 (CITA-Universidad de Zaragoza), Zaragoza, España.

⁽³⁾ Geoinnova, C/ Luis Morondo Urra nº 8 Bajo A31.006 Pamplona, España. miquel.febrer@geoinnova.es

⁽⁴⁾ Investigador independiente. lucia.martinezcebrian@gmail.com

Resumen: El abandono de la actividad agraria es un problema importante en Europa en general y en España en particular. Los regadíos de ribera sufren este proceso de abandono en mayor medida, por sus peculiares características. No existen estudios cuantitativos sobre el porcentaje de superficie y número de parcelas abandonadas en las zonas de regadío de ribera de Aragón. Por ello, el objetivo de este trabajo es desarrollar una metodología para localizar y cuantificar estas parcelas abandonadas mediante un modelo de clasificación aplicado a tres años consecutivos, y la definición de reglas de decisión basadas en la clasificación de cada año. El modelo se obtuvo con el algoritmo *Support Vector Machine Radial* y diferentes variables basadas en los índices NDVI, NDMI y NDTI y características texturales, todas ellas derivadas de imágenes de satélite Sentinel 2. Parcelas abandonadas, no abandonadas y parcelas sin determinar fueron las categorías identificadas en las dos áreas de estudio, las riberas de los ríos Martín y Guadalupe de Teruel, con una fiabilidad global en su clasificación del 85%. La aplicación del procedimiento al conjunto de regadíos de ribera de Teruel revela el predominio de las parcelas abandonadas, 52%, frente a las no abandonadas, 38%, quedando menos del 10% sin asignar.

Palabras clave: parcelas abandonadas, clasificación, imagen satélite, *machine learning*.

Using Sentinel 2 and CAP data to map land abandonment in traditional riparian irrigation lands

Abstract: The abandonment of agricultural activity is a major problem in Europe in general and in Spain in particular. Riverside irrigated areas are more affected by this abandonment process due to their specific characteristics. There are no quantitative studies on the percentage of area and number of abandoned plots in the irrigated riverside lands of Aragón. Therefore, the aim of this work is to develop a methodology to locate and quantify these abandoned plots by means of a classification model applied to three consecutive years and the definition of decision rules based on the classification of each year. The model was obtained using the *Support Vector Machine Radial* algorithm and different variables based on NDVI, NDMI and NDTI indices and textural characteristics, all derived from Sentinel-2 satellite images. Abandoned, non-abandoned and undetermined plots were the categories identified in the two study areas, the riverbanks of the Martín and Guadalupe rivers in Teruel, with an overall reliability in their classification of 85%. The application of the procedure to all the irrigated riverside lands of Teruel shows the predominance of abandoned plots, 52%, compared to non-abandoned plots, 38%, with less than 10% remaining unassigned.

Keywords: abandoned plots, classification, satellite image, *machine learning*.

1. INTRODUCCIÓN

En regadíos tradicionales de ribera que sufren abandono de la actividad agrícola, identificar las parcelas abandonadas es esencial para su recuperación. La gran cantidad de parcelas abandonadas a considerar, a menudo de pequeño tamaño, unida en muchos casos a la dificultad de acceso a ellas complica su identificación *in-situ*. Por ello, se hace necesario disponer de una metodología que permita agilizar y automatizar el proceso de detección y cuantificación, adaptándolo a las peculiaridades de estos regadíos ribereños. El uso de imágenes de satélite se muestra como una posible opción para la identificación de superficies abandonadas como ponen de manifiesto los trabajos de Portalés-Juliá, *et al.* (2021), Morell-Monzó, *et al.* (2021),

Vidal-Macua, *et al.* (2018), en los que utilizan en algunos casos imágenes Sentinel 2 y en otros Landsat obteniendo buenos resultados. En estos trabajos, tanto las áreas de estudio como los cultivos o usos del suelo utilizados son variados y presentan semejanzas con las condiciones en las que se encuentran las áreas objeto de estudio en este trabajo, aunque en cada uno de los trabajos existe un concepto de abandono diferente.

El objetivo del trabajo es disponer de una metodología para localizar y cuantificar con imágenes de satélite la superficie ocupada por parcelas abandonadas en regadíos tradicionales de ribera.

2. MATERIAL Y MÉTODOS

2.1. Ámbito de estudio

Los regadíos de ribera de la provincia de Teruel fueron seleccionados como área de estudio debido a que no existía información clara y cartografiada sobre el proceso de abandono que están sufriendo las zonas de ribera dentro de la provincia. Diversos entes locales y agricultores habían mostrado su preocupación sobre el abandono que están sufriendo estas tierras de cultivo. En ellas, generalmente, existen suelos de buena calidad y se encuentran cerca del casco urbano, pero se hallan fuera del modelo productivo actual, perdiéndose así un patrimonio medioambiental, productivo y paisajístico de gran interés local. Así, se seleccionaron los regadíos de las riberas de nueve ríos principales (Tabla 1) con curso de agua permanente, representando una superficie de algo más de 252 000 ha y un tamaño medio de parcela de 0.22 ha.

El tamaño de parcelas y la información de partida disponible sobre ellas y sus usos condicionó la estrategia a seguir y las imágenes a utilizar. Se optó por imágenes Sentinel 2, dada su resolución espacial y acceso libre, y por técnicas de clasificación *machine learning*. Las riberas de los ríos Martín y Guadalope fueron las escogidas como pilotos para la puesta a punto de la metodología de clasificación que posteriormente se aplicó al resto de riberas.

Tabla 1. Distribución de parcelas, número, superficie total y superficie media (ha), en subcuencas de ríos de Teruel.

Cuenca	Subcuenca.	Recintos SIG-PAC		
		Nº	Sup. (ha)	Sup. media (ha)
Ebro	Río Aguasvivas	7854	759	0.14
	Río Guadalope	24894	4705	0.28
	Río Huerva	1382	381	0.37
	Río Jiloca	25028	7535	0.37
	Río Martín	30495	3688	0.16
	Río Matarraña	8472	1338	0.26
Júcar	Río Mijares	13732	1714	0.16
	Río Turia	37267	4977	0.17
Tajo	Río Gallo	449	70	0.24
	Total	71789	7035	28

2.2. Modelo de clasificación

Se probaron los algoritmos de clasificación *Random Forest* (RF), *Support Vector Machine Polinomial* (SVM_POLY) y *Support Vector Machine Radial* (SVM_RADIAL) para diferenciar parcelas con y sin manejo productivo. Las parcelas de referencia para entrenar y validar se obtuvieron del SIGPAC y las declaraciones PAC de 2018. Se eligieron aleatoriamente el 5% de los recintos agroforestales que contenían más de 5 píxeles de 10 m, previa eliminación con un índice de forma de los recintos con “efecto pasillo”. Las parcelas elegidas que no presentan declaración PAC en 2018 son posibles candidatas a parcelas abandonadas. Mediante fotointerpretación de ortofotos del PNOA 2018 e imágenes Sentinel 2 se constató cuáles de éstas no presentan realmente un manejo productivo, eliminando aquellas dudosas. Quedó una muestra de 1423 parcelas, de

las cuales 679 no presentaban manejo productivo. El 70% de las parcelas de referencia se utilizaron para el entrenamiento y el 30% para la validación.

Como variables explicativas para la clasificación se consideraron el valor máximo mensual, y los valores máximos, mínimo, medio y desviación anual, de los índices de vegetación NDVI, de humedad NDMI y de laboreo NDTI obtenidos de las imágenes Sentinel 2 de 2018. También se consideró la media y desviación anual de las siguientes variables de medida de textura a través del método de matriz de co-ocurrencia en niveles de grises (Haralick *et al.*, 1973): segundo momento angular, contraste, correlación, momento diferencial inverso (homogeneidad), suma promedio y varianza. Para depurar las variables a utilizar se aplicó el test estadístico Kruskal-Wallis (prueba no paramétrica para distribuciones no normales de datos como es el caso) que revela si existen diferencias significativas entre las parcelas de referencia con y sin manejo productivo, y se analizó la correlación existente entre variables

Para la calibración de los modelos se utilizó el paquete Caret (Kuhn, 2008), manteniendo el porcentaje de cada categoría en cada una de las muestras para no sesgar los resultados del modelo hacia una categoría u otra. La calibración del modelo se realizó con el método de validación cruzada ($n=10$). Además, para cada uno de los tres modelos de clasificación se analizó la importancia de las variables explicativas finalmente escogidas. Se descartaron las variables anuales NDVI y NDMI mínimo, NDMI medio, NDTI máximo y textura-desviación estándar, al no existir diferencias significativas entre las parcelas de referencia con y sin manejo productivo. Aunque existía una correlación elevada entre algunas de las variables restantes, se optó por considerar todas para la clasificación ya que, al eliminarlas, el resultado final del modelo no mejoraba. La capacidad discriminatoria de los modelos y la fiabilidad de la clasificación se evaluó con los estadísticos exactitud y Kappa. Para ello, se utilizaron 10 remuestros distintos para evitar posibles sesgos en la elección de los datos de calibración y validación

2.3. Categorización y cartografía

Ya que un solo año no es determinante para considerar si una parcela está o no abandonada, se procedió a aplicar el modelo de clasificación escogido a tres años consecutivos. Los años elegidos fueron 2017, 2018 y 2019, siendo 2017 el año más seco de la década y 2018 el más húmedo.

Las parcelas de referencia se dividieron en dos grupos según el valor medio de la predicción realizada en 2018 con el modelo de clasificación seleccionado: $P>0.5$, (posibilidad de que no haya manejo productivo y, por tanto, de abandono), o $P<0.5$ (mayor probabilidad de manejo productivo). En la Tabla 2 se presenta el criterio utilizado para poder considerar una parcela como abandonada o no. Lo umbrales se obtienen para cada año a partir de las parcelas de referencia.

En este estudio, una parcela se considera o no abandonada en función de la clasificación que tiene en los tres años (2017, 2018 y 2019), salvo las parcelas con declaración PAC en el año de referencia (2018), que se consideran directamente no abandonadas. El resto de las parcelas se

clasifican según las reglas de la Tabla 3, tras probar y evaluar tres opciones diferentes.

Tabla 2. Criterios de categorización de parcelas de referencia

Parcelas Referencia	Umbral	Criterio	Asignación
Grupo 1 Sin manejo	Q3 (75%)	Probabilidad > umbral Q3	Abandonada
		Probabilidad < umbral Q3	Sin determinar
Grupo 2 Con manejo	Q1 (25%)	Probabilidad < umbral Q1	No abandonada
		Probabilidad > umbral Q1	Sin determinar

Tabla 3. Reglas de decisión para identificar si una parcela pertenece o no a una categoría y su probabilidad de ocurrencia, en función de las posibles respuestas de la parcela en el periodo de estudio (2017-2019).

Categoría y Probabilidad ocurrencia	Criterio
ABANDONADO	
Muy alta	3 años abandonado
Alta	2 años abandonado y 1 año sin determinar
Media	2 años abandonado y 1 año no abandonado 1 año abandonado y 2 años sin determinar
Baja	3 años sin determinar con $P > 0.5$ los 3 años Cada año una categoría con $P > 0.5$ los 3 años
NO ABANDONADO	
Muy alta	Parcelas con declaración PAC en 2018
Muy alta	3 años No abandonado
Alta	2 años No abandonado y 1 año sin determinar
Media	2 años no abandonado y 1 año abandonado 1 año no abandonado y 2 años sin determinar
Baja	3 años sin determinar con $P < 0.5$ los 3 años Cada año una categoría con $P < 0.5$ los 3 años
SIN DETERMINAR	
	3 años sin determinar si los 3 años no tiene $P > 0.5$ o $P < 0.5$ Cada año una categoría si los 3 años no tiene $P > 0.5$ o $P < 0.5$

Con objeto de conocer fiabilidad en la identificación de parcelas abandonadas y no abandonadas se comparó lo clasificado como Abandonado, No abandonado y Sin determinar con la asignación de Grupo 1-Abandonadas y Grupo 2-No abandonadas en las parcelas de validación.

3. RESULTADOS Y DISCUSIÓN

3.1. Clasificación

Para cada uno de los modelos RF, SVM_POLY y SVM_RADIAL, se analizó la importancia de las variables explicativas. Se realizaron varias pruebas de eliminación de variables de escasa importancia, no detectándose mejoras en ningún modelo, por lo que se asumió que estas variables aportaban cierta mejora al modelo a pesar de su importancia relativa menor. Además, se advirtió que incluso las variables altamente correlacionadas podrían tener información no redundante y la eliminación de cualquiera

de ellas reduciría la capacidad de discriminación del modelo.

En la Tabla 4, se representan los valores Kappa y la exactitud de los modelos para los datos de entrenamiento y validación. Los tres modelos muestran valores medios muy similares, siendo ligeramente superiores para SVM_Radial, modelo finalmente elegido. El modelo RF presenta una mayor variabilidad. Los valores de acierto son superiores a 0.85, confirmando la fiabilidad de los modelos con los datos de calibración.

Podemos observar cómo ambos estadísticos muestran unos valores medios muy similares entre los datos de validación y los de entrenamiento. Los valores mínimos de los datos de validación son más elevados, demostrando una mayor estabilidad.

Tabla 4. Bondad de ajuste (Kappa y Exactitud en %) de los modelos RF, SVM radial y SVM polinomial para los datos de calibración (Modelo_ENT) y validación (Modelo_VAL).

Modelo_ENT	Kappa			Exactitud		
	Min.	Mean.	Max.	Min.	Mean.	Max.
RF	0.597	0.702	0.798	0.800	0.852	0.899
SVM_RADIAL	0.658	0.740	0.798	0.830	0.871	0.899
SVM_POLY	0.695	0.734	0.797	0.848	0.868	0.899
Modelo_VAL						
RF	0.712	0.730	0.749	0.856	0.865	0.875
SVM_RADIAL	0.702	0.759	0.811	0.852	0.880	0.906
SVM_POLY	0.713	0.760	0.802	0.802	0.880	0.901

La predicción del modelo elegido sobre las parcelas de referencia se muestra gráficamente de la Figura 1.

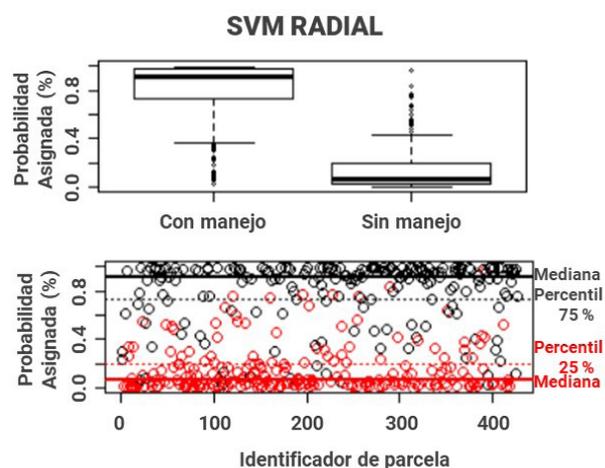


Figura 1. Predicción del modelo elegido sobre las parcelas de referencia. Boxplot (gráfica superior) y valores de predicción individuales, parcelas con manejo en rojo y sin manejo en negro (gráfica inferior).

En la Figura 2 se muestra uno de los resultados cartográficos que se consiguen con la información obtenida de la clasificación.

En las cuencas de los ríos Martín y Guadalupe, de las 426 parcela de referencia 155 se clasificaron como abandonadas, 244 como no abandonadas y 27 se quedaron sin determinar. La fiabilidad global de la clasificación fue 85%, con un error de omisión del 15.7% y de comisión del 2.6% para la categoría abandonadas y del 1.8 % y 13.1%

respectivamente para las categorías no abandonadas. A la vista de estos resultados se presume que la incorporación de más años en la clasificación puede ser una opción para mejorar la predicción, así como disminuir las parcelas en las que no se puede determinar si están o no abandonadas.

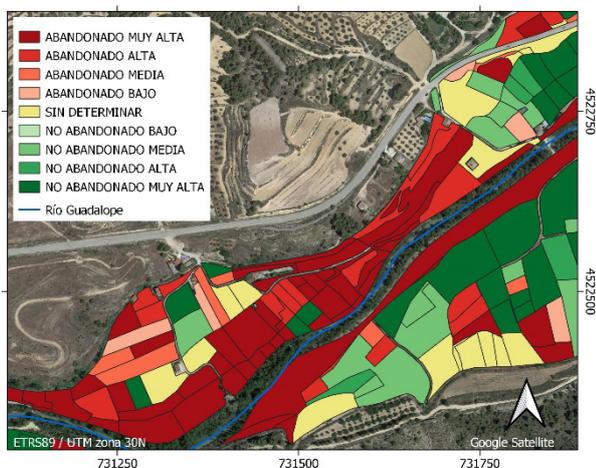


Figura 2. Detalle de distribución de clasificación de parcelas.

3.2. Abandono en los regadíos de ribera de Teruel

La aplicación del procedimiento de clasificación descrito al conjunto de regadíos de ribera de Teruel (Tabla 5) clasifica 149.573 parcelas. Más del 38% de las parcelas fueron identificadas como abandonadas, mientras que el 52% aparecen como no abandonadas y menos del 10% quedaron sin determinar. En las clasificadas como abandonadas predominan los usos SIGPAC de Tierra Arable (un poco más el 18%), pastos y pastizales (un 10%) y forestal (6.3%).

Tabla 5. Distribución (número, superficie en ha y %) de parcelas abandonadas en los regadíos de ribera de Teruel por subcuenca.

Cuenca	Subcuenca	Recintos SIG-PAC		
		Nº	ha	%
Ebro	Río Aguasvivas	4944	301	40
	Río Guadalope	9973	1309	28
	Río Huerva	325	21	5
	Río Jiloca	5389	1229	14
	Río Martín	17 508	1518	41
Júcar	Río Matarraña	3908	452	34
	Río Mijares	10242	932	54
Tajo	Río Turia	19390	1465	29
	Río Gallo	110	8	12
Total		71 789	7035	28

La distribución del abandono por subcuencas (Tabla 5), es muy variable. La subcuenca con mayor porcentaje de superficie abandonada es Mijares seguida de la del río Martín y Aguavivas. Estas subcuencas corresponden a las que el tamaño medio de las parcelas es más pequeño (Tabla 1).

4. CONCLUSIONES

Los tres modelos de clasificación, *Random Forest*, *Support Vector Machine Polinomial* y *Support Vector Machine Radial*, comparados en este trabajo para diferenciar parcelas con y sin manejo productivo utilizando datos de Sentinel 2, presentan valores de exactitud muy similares, siendo ligeramente superiores para *Support Vector Machine Radial*.

El procedimiento propuesto, clasificación con *Support Vector Machine Radial* y reglas de decisión basadas en la clasificación obtenida en los tres años estudiados (2017-2019), ha permitido identificar si una parcela está abandonada, no abandonada o se queda sin determinar con una fiabilidad global del 85%.

Los resultados obtenidos, si bien presentan margen de mejora, son optimistas en la identificación de parcelas abandonadas con teledetección en zonas con características similares a las de los regadíos de ribera.

5. AGRADECIMIENTOS

Los proyectos FITE, RegATeA y RegATeA-Pro donde se desarrolló el trabajo han sido financiados por el Gobierno de Aragón, a través del fondo de inversiones de Teruel (años 2019-2020 y 2022-2024), con participación del Gobierno de España (Ministerio de Hacienda).

6. BIBLIOGRAFÍA

Haralick, R.M., Shanmugam, K., Dinstein, I.H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, 6, 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>

Kuhn, M. (2008). Caret package. *Journal of Statistical Software*, 28(5), 1-26. <https://doi.org/10.18637/jss.v028.i05>

Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.T. (2020). Comparison of Sentinel-2 and high-resolution imagery for mapping land abandonment in fragmented areas. *Remote Sensing*, 12, 2062. <https://doi.org/10.3390/rs12122062>

Portalés-Julià, E., Campos-Taberner, M., García-Haro, F.J., Gilabert, M.A. (2021). Assessing the sentinel-2 capabilities to identify abandoned crops using deep learning. *Agronomy*, 11(4), 654. <https://doi.org/10.3390/agronomy11040654>

Vidal-Macua, J.J., Ninyerola, M., Zabala, A., Domingo-Marimon, C., Gonzalez-Guerrero, O., Pons, X. (2018). Environmental and socioeconomic factors of abandonment of rained and irrigated crops in northeast Spain. *Applied Geography*, 90, 155-174. <https://doi.org/10.1016/j.apgeog.2017.12.005>