

Clasificación de imágenes Landsat 8 en la Demarcación Hidrográfica del Segura

Rodríguez-Valero, M.I. ^{1*}, Alonso-Sarria, F. ²

¹ Universidad de Murcia, Avda. Teniente Flomesta, 5, 30003, Murcia, España.

² Instituto Universitario de Agua y Medio Ambiente, Universidad de Murcia. Edificio D, Campus de Espinardo, s/n, 30100, Murcia. España.

Resumen: En el presente trabajo se obtiene, a través de un proceso de clasificación digital de imágenes Landsat 8 de 2017, una cartografía de usos del suelo en la Demarcación Hidrográfica del Segura; la clasificación se ha llevado a cabo empleando dos clasificadores: Máxima Verosimilitud (MV) y *Random Forest* (RF). Puesto que las áreas de entrenamiento se obtienen de imágenes históricas de alta resolución hasta 2016, previamente a la clasificación, se hace un análisis de validación cruzada de las áreas de entrenamiento para determinar cuáles de ellas pueden haber sufrido un cambio de uso entre 2016 y 2017. Los resultados obtenidos con MV y con RF, tanto con el conjunto original de áreas como con el obtenido eliminando la problemática, se comparan para determinar cuál es la mejor opción. En el caso de MV mejoran notablemente los resultados cuando se realiza la depuración de las áreas de entrenamiento, pasando de un porcentaje de acierto del 81% al 84,9%; sin embargo, con RF esta mejora no es tan importante ya que se pasa del 84,1% al 85,1%. Por tanto, se puede concluir que, con ambos métodos la clasificación es más exacta cuando se emplean las áreas de entrenamiento depuradas y, aunque los resultados obtenidos son bastante aceptables tanto para MV como para RF, este último realiza una clasificación más exacta.

Palabras clave: Random Forest, Máxima Verosimilitud, usos del suelo, teledetección, Landsat 8.

Classification of Landsat 8 images in the Segura Hydrographic Demarcation

Abstract: This work presents a cartography of land uses in the Segura Hydrographic Demarcation obtained by classifying 2017 Landsat 8 images. The classification was carried out using two classifiers: Maximum Likelihood (ML) and Random Forest (RF). Training areas were obtained from historical high resolution imagery until 2016. Prior to classification, a cross validation analysis of the training areas was carried out to determine which of them may have undergone a change of use between 2016 and 2017. The results obtained with ML and RF, both with the original set of training areas and with the one obtained eliminating the problem, are compared to determine the best option. In the case of ML, the results improve after eliminating the changing training areas, from 77.7% to 81.4%; however, with RF this improvement is not so important, going from 84.1% to 85.1%. Therefore, it can be concluded that, with both methods, the classification is more exact when the modified training areas are used and, although the results obtained are quite acceptable for both ML and RF, the latter performs a more accurate classification.

Key words: Random Forest, Maximum Likelihood, land use, remote sensing, Landsat 8.

To cite this article: Rodríguez-Valero, M.I., Alonso-Sarria, F. 2019. Classification of Landsat 8 images in the Segura Hydrographic Demarcation. *Revista de Teledetección*, 53, 33-44. <https://doi.org/10.4995/raet.2019.11016>

* Corresponding author: mariaisabel.rodriguez1@um.es

1. Introducción

Uno de los objetivos de la clasificación digital de imágenes de satélite es dividir la imagen original en un número finito de clases disjuntas, de tal forma que cada píxel que compone la imagen es asignado a una clase en particular (Ayala y Menenti, 2001), obteniendo como resultado una cartografía de las categorías objeto de estudio (Chuvieco, 1996). En este trabajo se emplea la clasificación supervisada, que requiere de un cierto conocimiento del área de estudio ya que, antes de efectuar la clasificación, se definen las categorías o clases que se pretenden discriminar (fase de entrenamiento).

Diversas estrategias se han propuesto para añadir nueva información relevante a la reflectividad de las diferentes bandas de una imagen de satélite. El uso de información textural destaca la variación espacial del brillo de una imagen, favoreciendo así la separabilidad de las clases (Berberoglu *et al.*, 2007, Gomariz *et al.*, 2017). Wang y Tenhunen (2004), Ezzine *et al.* (2014) o Gomariz *et al.* (2017) muestran cómo, debido a la estacionalidad de usos y coberturas, el uso de una imagen por estación aumenta la exactitud de la clasificación de usos del suelo.

La liberación de las imágenes del satélite Landsat en 2008 permite clasificar (sin excesivo coste) imágenes históricas para obtener series de mapas de cobertura y usos del suelo. El problema que es necesario afrontar, en ese caso, es la falta de áreas de entrenamiento para todas las fechas. Una posible solución es buscar áreas que no hayan cambiado de uso a lo largo del período de estudio, utilizando para ello ortofotografía de alta resolución y mapas de usos de suelo de diferentes fechas. En todo caso, sería necesario un procedimiento para evaluar qué áreas pueden haber sufrido un cambio de uso entre la fecha de referencia de las áreas de entrenamiento y la de la imagen a clasificar.

Se han propuesto diversos algoritmos para la clasificación de imágenes de satélite. Entre los más utilizados están Máxima Verosimilitud (MV) que asume que la función densidad de las reflectividades en cada clase es una normal multivariante. Recientemente se han propuesto diversos métodos basados en técnicas de aprendizaje automático que no están condicionados por supuestos estadísticos previos. Entre ellos destacan los basados en

árboles de decisión, como *Random Forest* (RF) o *boosting*, los basados en redes neuronales, o los basados en funciones *kernel* como *support vector machines* (Maxwell *et al.*, 2018).

El objetivo de este trabajo es clasificar los usos del suelo en 2017 en la Demarcación Hidrográfica del Segura utilizando MV y RF a partir de tres imágenes correspondientes a primavera, verano y otoño. Se parte de una colección de áreas de entrenamiento, obtenidas a partir de diversas fuentes entre 1984 y 2016, cuyo uso puede considerarse constante durante este período. Previamente a la clasificación, se hace un análisis de validación cruzada de las áreas para determinar cuáles de ellas pueden haber sufrido un cambio de uso entre 2016 y 2017. Los resultados obtenidos con MV y con RF, tanto con el conjunto original de áreas como con el obtenido eliminando las áreas que presentan cambios, se comparan para determinar cuál es la mejor opción.

2. Área de estudio

La Demarcación Hidrográfica del Segura (DHS) se localiza en el sureste de España, y cuenta con una superficie aproximada de 18.870 km². Afecta a cuatro comunidades autónomas, englobando casi en su totalidad a la Región de Murcia y parcialmente a Andalucía (Jaén, Granada y Almería), Castilla la Mancha (Albacete) y Comunidad Valenciana (Alicante). Presenta una gran variedad orográfica y climática, destacando el contraste entre períodos de fuertes sequías y las frecuentes inundaciones debidas a las lluvias torrenciales. A pesar de ello, hay un intensivo uso agrícola con una gran variedad de cultivos y coberturas.

Para este estudio se han diferenciado un total de 9 usos del suelo presentes en la zona: 1. Bosque: engloba a aquellas superficies forestales que albergan vegetación de porte arbóreo; 2. Matorral: se corresponde con las zonas de típico matorral mediterráneo y es fácilmente localizable en las zonas de solana; 3. Arbóreo poco denso: alberga cultivos de viñedos, olivos o almendros, entre otros; 4. Arbóreo denso: en este uso se localizan cultivos de cítricos y frutales de hueso; 5. Herbáceo de secano: es el suelo dedicado principalmente al cultivo de cereales y algunas leguminosas; 6. Herbáceo de regadío: Comprende áreas dedicadas al cultivo hortícola, además de los arrozales; 7. Superficies

impermeables: suelo que ha sido sometido a una fuerte transformación antrópica, provocando que la escorrentía en este sea máxima y no haya infiltración de agua. 8. Láminas de agua: zonas del suelo que están cubiertas de agua. 9. Invernaderos: suelos dedicados al cultivo en invernaderos.

3. Metodología

3.1. Áreas de entrenamiento

En una primera fase del proyecto se identificaron los usos del suelo de 1181 áreas, de las que un 30% serán utilizadas como áreas de validación. Las parcelas de entrenamiento y validación se han obtenido a partir de análisis multitemporales de fotografías aéreas y mapas de usos del suelo, comparando diferentes fuentes, que abarcan un periodo temporal de 30 años. Las imágenes empleadas para ello han sido: Ortofoto Digital Pancromática de Andalucía 1984-1985, Ortofoto Digital Pancromática del SIG Oleícola de Andalucía 1997-98, Ortofoto 1997 SIG Oleícola (Murcia), Ortofoto Digital de Andalucía coloreada 2001-02, SIGPAC 97-03 (ortoimágenes del vuelo SIGPAC realizadas entre 1997 y 2003 sobre la totalidad del territorio español), OLISTAT 97-98 (ortoimágenes del vuelo OLISTAT Oleícola realizado entre 1997 y 1998 sobre territorio español) y Ortoimagen PNOA 2016 (ortofoto de máxima actualidad del proyecto PNOA). Debido a que los servidores no siempre están activos, se ha procurado utilizar 3 imágenes como mínimo; el acceso a las ortofotos se ha llevado a cabo en QGIS, mediante una conexión WMS a servidores oficiales. Si el uso del suelo de una determinada zona no

ha sufrido cambios en ese espacio de tiempo, se creará un área de entrenamiento (Tabla 1). Es importante que estas sean lo más homogéneas posible.

Tabla 1. Usos del suelo identificados en la DHS.

Código identificador	Usos
1	Bosque
2	Matorral
3	Arbóreo poco denso
4	Arbóreo denso
5	Herbáceo de secano
6	Herbáceo de regadío
7	Superficies impermeables
8	Láminas de agua
9	Invernaderos

3.2. Obtención de imágenes

Para llevar a cabo la clasificación digital, se utilizará un conjunto de 12 imágenes Landsat 8 del año 2017, 4 por cada estación del año, excepto invierno, ya que la cuenca del Segura se reparte entre 4 imágenes de Landsat 8. Las fechas son 2/5 y 12/6 para primavera, 21-30/7 para verano y 10-19/11 para otoño. Las correspondientes al periodo invernal no se han descargado por tener una nubosidad demasiado alta; además, ha quedado demostrado en investigaciones anteriores que la imagen de invierno es, al menos en el entorno del área de estudio, la que menos información aporta de las cuatro estaciones (Gomariz *et al.*, 2017). Las imágenes Landsat están identificadas por sus *paths/rows* del *Worldwide Reference System of Landsat (WRS)*; para la DHS son: 199/33, 199/34, 200/33 y 200/34. Las imágenes para llevar a cabo este estudio (Tabla 2) se han descargado a través

Tabla 2. Imágenes descargadas a través de “Semiautomatic Classification Plugin”.

Id. imagen	Fecha	% Nubes
LC08_L1TP_199033_20170612_20170628_01_T1	12/06/17	0
LC08_L1TP_199034_20170612_20170628_01_T1	12/06/17	0
LC08_L1TP_200033_20170502_20170515_01_T1	02/05/17	0
LC08_L1TP_200034_20170502_20170515_01_T1	02/05/17	0
LC08_L1TP_199033_20170730_20170811_01_T1	30/07/17	0
LC08_L1TP_199034_20170730_20170811_01_T1	30/07/17	0
LC08_L1TP_200033_20170721_20170728_01_T1	21/07/17	2
LC08_L1TP_200034_20170721_20170728_01_T1	21/07/17	2
LC08_L1TP_199033_20171119_20171205_01_T1	19/11/17	0
LC08_L1TP_199034_20171119_20171205_01_T1	19/11/17	0
LC08_L1TP_200033_20171110_20171121_01_T1	10/11/17	1
LC08_L1TP_200034_20171110_20171121_01_T1	10/11/17	0

del complemento “Semiautomatic Classification Plugin” (SCP); es un *plugin* de código abierto y libre acceso para QGIS (Congedo, 2016) que permite descargar imágenes de forma semiautomática indicando las coordenadas geográficas del área de estudio, las fechas deseadas y el máximo porcentaje de nubes que podemos aceptar para realizar la clasificación, que en este trabajo fue de un 20%. Este *plugin* permite el preprocesado de imágenes; en este trabajo se obtuvo la reflectividad en la superficie terrestre utilizando el algoritmo de Moran *et al.* (1992) y se corrigió el efecto de la atmósfera utilizando el método de Chávez (1996). A continuación, se importan las imágenes en GRASS y se renombran, con el fin de que su manejo resulte más rápido y sencillo, y posteriormente se procede a la unión de las cuatro imágenes que abarcan el área de estudio, por estación y para todas las bandas, obteniendo un total de 11 capas para cada estación. Para verificar que la superposición de imágenes captadas en fechas ligeramente diferentes no genera cortes en el espacio, se hicieron composiciones de color RGB; así, además se puede comprobar que no existe nubosidad que haya sido indetectable para el algoritmo de la NASA.

3.3. Cálculo de índices

Una vez comprobado que la calidad de las imágenes es buena, se crean una serie de índices normalizados, para cada periodo estacional, a partir de las bandas de Landsat 8, cuyo cálculo se basa en la diferencia entre dos reflectividades divididas entre la suma de esos dos mismos valores. Estos índices establecen filtros selectivos para aprovechar la mejor respuesta en las distintas bandas de lectura de acuerdo a las reflectividades de los componentes de la superficie (Chiarito y Chiarito, 2015), y son los siguientes:

El Índice Normalizado de Vegetación (NDVI): es uno de los más utilizados. Se genera empleando las bandas 4 y 5 de Landsat 8 (Ecuación 1):

$$NDVI = \frac{\rho_5 - \rho_4}{\rho_5 + \rho_4} \quad (1)$$

donde ρ_5 es la reflectividad de la banda 5 y ρ_4 es la reflectividad de la banda 4. Este índice se ha correlacionado con el contenido de clorofila en la hoja, el flujo neto de CO_2 , la radiación fotosintéticamente activa, la productividad neta de la

vegetación, el índice de área foliar o la dinámica fenológica entre otros (Chuvieco, 2010).

El Índice Normalizado de Edificaciones (NDBI): permite estimar zonas con superficies construidas o edificadas (Chen *et al.*, 2006) junto a las superficies desnudas (Santana *et al.*, 2010); mide la proporción de edificios en cada píxel. En zonas semiáridas en las que los suelos tienen poca vegetación, es difícil que este índice permita acotar las áreas edificadas, ya que la reflectividad de ambas cubiertas es muy similar en estas regiones del espectro electromagnético (García *et al.*, 2014). Se genera a partir de las bandas 5 y 6 de Landsat 8 (Ecuación 2):

$$NDBI = \frac{\rho_6 - \rho_5}{\rho_6 + \rho_5} \quad (2)$$

donde ρ_5 es la reflectividad de la banda 5 y ρ_6 es la reflectividad de la banda 6.

El Índice Normalizado de Desnudez de Suelo (NDBaI): identifica zonas sin vegetación. Se genera a partir de las bandas 5 y 10 de Landsat 8 (Ecuación 3):

$$NDBaI = \frac{ND5 - ND10}{ND5 + ND10} \quad (3)$$

donde ND5 es el nivel digital de la banda 5 y ND10 es el nivel digital de la banda 10. Puesto que la banda 10 es el térmico, se utilizan valores de nivel digital (ND) en lugar de reflectividades.

El Índice Normalizado Modificado de Agua (MNDWI): mide la cantidad de agua presente en el píxel. Este método fue propuesto por Xu (2006), y utiliza las bandas 3 y 6 de Landsat 8 (Ecuación 4):

$$MNDWI = \frac{\rho_3 - \rho_6}{\rho_3 + \rho_6} \quad (4)$$

donde ρ_3 es la reflectividad de la banda 3 y ρ_6 es la reflectividad de la banda 6.

3.4. Algoritmos de clasificación

El siguiente paso es la clasificación; en este trabajo se ha empleado el clasificador de Máxima Verosimilitud (MV) y el clasificador *Random Forest* (RF). El primero, asume que las variables utilizadas siguen una distribución normal multivariante en cada clase, de tal forma que cada categoría tiene asignada una función gaussiana, que se obtiene a partir de su vector de medias y matriz

de varianza-covarianza, permitiendo así calcular la probabilidad de que un píxel pertenezca a una determinada clase, asignando dicho píxel a aquella clase que maximice la función de probabilidad (Chuvieco, 2010).

Por otra parte, RF es un clasificador por *ensembles* propuesto por Breiman (2001). Los clasificadores de *ensembles* construyen un conjunto de clasificadores “débiles” y combinan sus resultados obteniéndose un clasificador más exacto (Zhou, 2012). En el caso de RF, se combinan varios árboles de decisión, que se caracterizan por ser clasificadores de bajo sesgo y alta varianza, para obtener una predicción con similar sesgo, pero mucha menos varianza (Zhou, 2012). Una de las ventajas adicionales de RF es que aporta una estimación interna de exactitud mediante una forma de validación cruzada (Cánovas *et al.*, 2016). RF entrena un gran número de árboles de decisión (por defecto 500) y cada uno de ellos, a su vez, se entrena con un subconjunto de los datos de entrenamiento obtenido mediante *bootstrapping*. Esta técnica es la más conocida de las técnicas de remuestreo; consiste en extraer con reemplazamiento, a partir de la muestra original, una nueva muestra del mismo tamaño y que tendrá las mismas propiedades estadísticas (Berk, 2016). Además, en cada nodo de los árboles de decisión, la variable utilizada para dividir el nodo se obtiene a partir de un subconjunto aleatorio de las variables empleadas. El hecho de que RF entrene cada árbol con un conjunto de datos y variables diferentes, hace que disminuya la correlación entre árboles, además del error de generalización, obteniéndose así mejores resultados que con otros algoritmos (Cánovas *et al.*, 2016).

En ambos métodos se utilizaron como predictores tanto los valores de reflectividad como los índices de cada una de las 3 fechas. En total 45 predictores.

3.5. Selección de áreas de entrenamiento

Para determinar qué áreas de entrenamiento podrían haber cambiado de uso entre 2016 y 2017 se llevó a cabo un análisis de validación cruzada dejando fuera polígonos completos y se contabilizó el porcentaje de celdas correctamente clasificadas en cada polígono. Se considera que los polígonos con ninguna celdilla correctamente clasificada

han cambiado de uso y se eliminan del conjunto de áreas de entrenamiento. Este procedimiento se realiza tanto con MV como con RF de manera que tras la eliminación se tienen dos nuevos conjuntos de áreas de entrenamiento, uno para cada método.

3.6. Clasificación y validación

Se han entrenado 4 modelos: MV con las áreas de entrenamiento originales, MV con las modificadas, RF con las áreas originales y RF con las modificadas. Estos polígonos se calibran con el 70% de las áreas correspondientes y se validan con el 30% restante empleando el método *hold-out* (Pérez *et al.*, 2015).

Las áreas de validación se utilizarán para calcular la matriz de confusión y el estadístico Kappa; a través del análisis de la matriz de confusión “no solo se conoce la fiabilidad global de la clasificación, sino también la exactitud conseguida para cada una de las clases, además de los principales conflictos entre ellas” (Chuvieco, 2010:501). En esta matriz, cada fila corresponde a los píxeles de una clase en las áreas de validación e indica cuántos de ellos han sido asignados a cada una de las clases por el proceso de clasificación utilizado, en este caso MV y RF; por otra parte, cada columna corresponde a una clase en el mapa clasificado e indica a qué clases pertenecían en realidad estos píxeles. Mediante el cálculo de la matriz de confusión se conocerá el error de omisión y de comisión del modelo utilizado; el error de omisión es el porcentaje de píxeles que no han sido clasificados en esa clase y el de comisión el porcentaje de píxeles clasificados en una clase pero que realmente no pertenecen a ella.

Por otro lado, el estadístico Kappa mide la diferencia entre el acuerdo mapa-realidad observado y el que cabría esperar simplemente por azar; intenta delimitar el grado de ajuste debido solo a la exactitud de la clasificación, prescindiendo del causado por factores aleatorios (Chuvieco, 2010). Se calcula con la Ecuación 5:

$$K = \frac{N \cdot \sum_{i=1}^n x_{i,i} - \sum_{i=1}^n s_{i,i}}{N^2 - \sum_{i=1}^n s_{i,i}} \quad (5)$$

donde:

$X_{i,j}$ = elemento situado en la fila i columna j de la matriz de confusión.

$s_{.i}$ = suma de los elementos de la columna i (clase i).

s_i = suma de los elementos de la fila i (clase i).

N = número de píxeles; n = número de clases.

El valor del índice Kappa oscila entre -1 y 1 , de tal forma que a medida que este es más pequeño, menor acuerdo existe. Para su interpretación, se sigue la valoración propuesta por Landis y Koch (1977), presentada en la Tabla 3, que indica el grado de concordancia del modelo empleado.

Tabla 3. Valoración del coeficiente Kappa (Landis y Koch, 1977).

Kappa	Grado de acuerdo
<0,00	Sin acuerdo
0,00-0,20	Insignificante
0,21-0,40	Discreto
0,41-0,60	Moderado
0,61-0,80	Sustancial
0,81-1,00	Casi perfecto

Debido al carácter aleatorio de RF, diferentes ejecuciones producirán diferentes resultados, además y con intención de poder calcular los intervalos de confianza de los valores de kappa de modo no paramétrico, los modelos se validan a partir de 50 remuestreos obtenidos mediante *bootstrapping* a partir de los polígonos de validación, tal como se recomienda, por ejemplo, en Berk (2016) o Efron y Hastie (2018).

3.7. Recursos informáticos

Este estudio se ha desarrollado en un entorno Windows utilizando, como software para el almacenamiento y gestión de imágenes, QGIS y GRASS (*Geographic Resources Analysis Support System*). Ambos son Sistemas de Información Geográfica (SIG) de código abierto que se distribuyen bajo la Licencia Pública General (GPL).

Concretamente, se ha utilizado la versión QGIS 2.12.3 con GRASS 6.4.3. La clasificación con RF se ha llevado a cabo utilizando “R”, que es un software libre que permite realizar análisis estadísticos; es el más usado en la comunidad científica. Se puede descargar en <http://www.rproject.org>.

4. Resultados y discusión

4.1. Selección de polígonos

Al hacer la validación cruzada uno a uno con RF se encontró que, de los 1181 polígonos de entrenamiento disponibles, en 374 (31,7%) se clasificaban correctamente todos sus píxeles y en 72 (6,1%) no se clasificaba correctamente ningún píxel. Con MV los resultados fueron 331 (28%) y 118 (10%) respectivamente. En 58 (5%) polígonos ninguno de los métodos logró clasificar correctamente ningún píxel. En consecuencia, los polígonos con un 100% de error para cada algoritmo se eliminaron de los correspondientes conjuntos de entrenamiento.

4.2. Modelos

Una vez aplicados los algoritmos de clasificación, se presentan los resultados, para cada uno de los métodos de clasificación (Figuras 1 y 2) (disponibles para descarga en <https://umubox.um.es/index.php/s/LAnJwWLHAFqdhk0>), las matrices de confusión para MV empleando todas las áreas de entrenamiento (Tabla 4), MV empleando las áreas de entrenamiento depuradas (Tabla 5), RF empleando todas las áreas de entrenamiento (Tabla 6) y RF empleando las áreas de entrenamiento depuradas (Tabla 7), los porcentajes de errores de omisión y comisión para ambos

Tabla 4. Matriz de confusión del modelo MV con el conjunto de áreas de entrenamiento original.

Clase	Bosq.	Mator.	Arb. Poc. D.	Arb. Denso	Herb. Sec.	Herb. Reg.	Sup. Imp.	L. Agua	Inv.	Total
Bosque	2.531	110	70	0	28	0	6	23	0	2.768
Matorral	367	3.861	186	116	191	79	17	81	0	4.898
Arb. P. denso	104	50	641	66	173	84	32	16	0	1.166
Arb. Denso	1	0	19	1.280	58	394	36	4	0	1.792
Herb. Sec.	1	2	73	37	744	28	15	8	0	908
Herb. Reg.	28	55	54	214	162	188	17	20	4	742
Sup. Imp.	7	7	28	22	87	98	2.188	29	96	2.562
Lám. Agua	0	0	0	2	0	0	0	5.778	0	5.780
Invern.	0	0	0	0	0	0	34	3	127	164
Total	3.039	4.085	1.071	1.737	1.443	871	2.345	5.962	227	

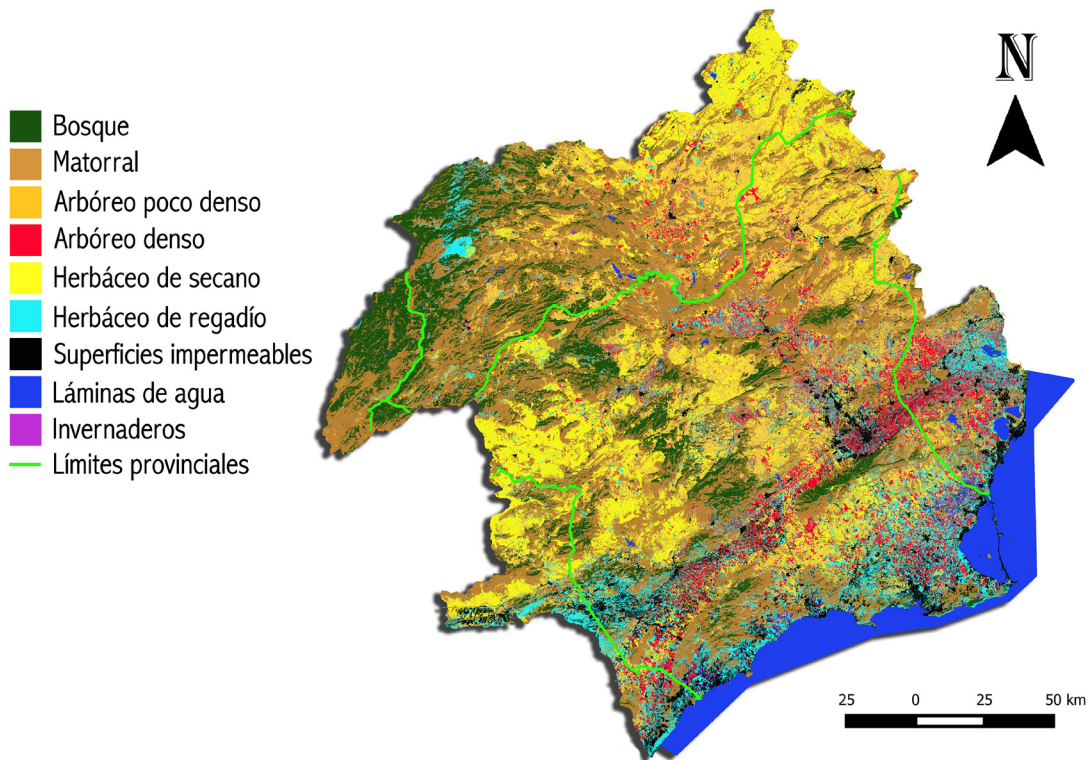


Figura 1. Mapa de predicción correspondiente al algoritmo de MV con las áreas de entrenamiento depuradas.

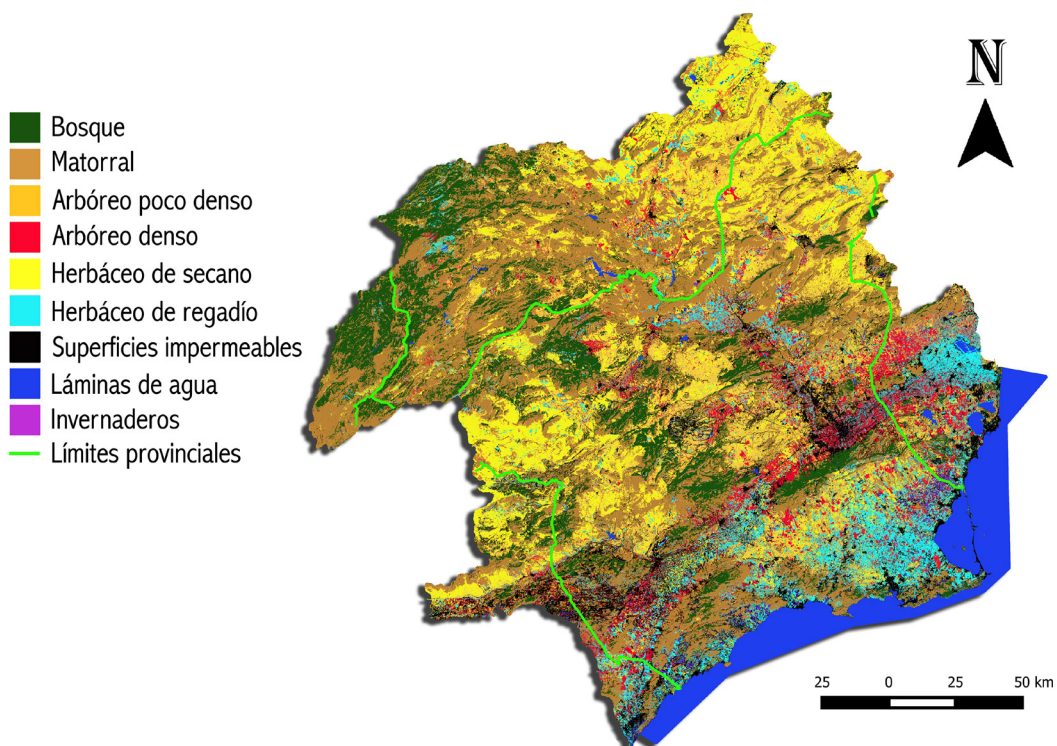


Figura 2. Mapa de predicción correspondiente al algoritmo RF con las áreas de entrenamiento depuradas.

Tabla 5. Matriz de confusión del modelo MV con el conjunto de áreas de entrenamiento modificado.

Clase	Bosq.	Mator.	Arb. Poc. D.	Arb. Denso	Herb. Sec.	Herb. Reg.	Sup. Imp.	L. Agua	Inv.	Total
Bosque	2.600	117	11	8	19	0	5	6	0	2.766
Matorral	183	3.115	39	181	223	15	38	5	0	3.799
Arb. P. denso	30	63	662	37	259	8	39	1	1	1.100
Arb. Denso	0	0	21	1.303	30	249	34	0	14	1.651
Herb. Sec.	47	2	53	17	2.021	20	2	0	1	2.163
Herb. Reg.	25	69	39	104	150	303	186	5	18	899
Sup. Imp.	5	6	20	25	54	53	2.604	35	25	2.827
Lám. Agua	0	0	0	1	0	0	0	4.799	0	4.800
Invern.	0	0	0	0	0	0	10	0	129	139
Total	2.890	3.372	845	1.676	2.756	648	2.918	4.851	188	

métodos utilizando los dos conjuntos de áreas de entrenamiento (Tabla 8) y, por último, los estadísticos de la exactitud obtenida en los diferentes remuestreos (Tabla 9).

Teniendo en cuenta los resultados de los errores de comisión y omisión para ambos métodos (Tabla 8), la clasificación que MV realiza de las celdas de vegetación natural (bosque y matorral) se puede considerar aceptable, tanto con las áreas de entrenamiento originales como con las modificadas. Sin embargo, RF para estas mismas clases, realiza una clasificación más exacta cuando se aplica a las áreas de entrenamiento originales. En ambos métodos existe un cierto grado de confusión entre bosque y matorral debido al carácter difuso de la frontera entre ambas clases.

Respecto al arbóreo poco denso, presenta porcentajes altos de errores de omisión y de comisión tanto en MV como en RF, disminuyendo estos errores en los casos en los que se emplean las áreas de entrenamiento modificadas. El arbóreo poco denso tiene tendencia a confundirse con aquellas clases que contienen vegetación, aunque con el herbáceo de regadío parece distinguirse mejor en todos los casos. Cuando se emplea MV con las áreas de entrenamiento originales se obtienen los

peores resultados, ya que un 45% de los píxeles clasificados como arbóreo poco denso pertenecen en realidad a otros usos del suelo, y un 40% de los píxeles que sí pertenecen a esta clase se han asignado a otras; en general, la clasificación obtenida para esta clase no puede considerarse buena.

En relación al arbóreo denso, llama la atención que tanto MV como RF arrojan un alto error de omisión con las áreas de entrenamiento modificadas (32,6% y 41,7%, respectivamente); en la clasificación llevada a cabo con RF, los píxeles que pertenecen a arbóreo denso han sido asignados, mayoritariamente, a las clases de arbóreo de secano, matorral y herbáceo de regadío, mientras que MV los asigna a las dos últimas, fundamentalmente. El error de omisión es menor en MV, sin embargo, el de comisión es mayor que en RF, habiendo clasificado un 22,1% de los píxeles como arbóreo denso, cuando en realidad estos pertenecen a otras clases, principalmente a herbáceo de regadío.

El herbáceo de secano presenta un error de omisión elevado (48,4%) cuando se emplea MV con las áreas de entrenamiento originales; este error disminuye significativamente cuando se utilizan las áreas modificadas, tanto para MV como para

Tabla 6. Matriz de confusión del modelo RF con el conjunto de áreas de entrenamiento original.

Clase	Bosq.	Mator.	Arb. Poc. D.	Arb. Denso	Herb. Sec.	Herb.R.	Sup. Imp.	L. Agua	Inv.	Total
Bosque	3.018	180	19	23	68	0	13	0	0	3.321
Matorral	82	3.753	85	47	177	4	25	0	0	4.173
Arb. P. denso	11	94	747	116	187	25	51	5	2	1.238
Arb. Denso	8	12	91	741	43	72	37	4	1	1.009
Herb. Sec.	0	33	228	15	1.162	311	19	10	1	1.779
Herb. Reg.	9	51	11	96	100	599	93	8	16	983
Sup. Imp.	11	40	63	39	48	37	2.060	19	108	2.425
Lám. Agua	0	0	0	3	0	12	4	5.042	0	5.061
Invern.	0	0	0	0	0	0	18	0	93	111
Total	3.139	4.163	1.244	1.080	1.785	1.060	2.320	5.088	221	

Tabla 7. Matriz de confusión del modelo RF con el conjunto de áreas de entrenamiento modificado.

Clase	Bosq.	Mator.	Arb. Poc. D.	Arb. Denso	Herb. Sec.	Herb. Reg.	Sup. Imp.	L. Agua	Inv.	Total
Bosque	2.372	219	0	1	38	10	9	47	0	2.696
Matorral	130	4.017	156	284	119	24	16	23	0	4.769
Arb. P. denso	12	54	702	240	99	11	27	0	0	1.145
Arb. Denso	0	12	42	1.160	12	110	48	7	2	1.393
Herb. Sec.	1	18	197	78	1.678	323	17	0	0	2.312
Herb. Reg.	5	53	10	177	103	1.269	32	12	12	1.673
Sup. Imp.	6	41	53	40	27	199	2.074	14	36	2.490
Lám. Agua	0	0	0	10	0	22	0	4.330	1	4.363
Invern.	0	0	0	0	0	0	1	0	85	86
Total	2.526	4.414	1.160	1.990	2.076	1.968	2.224	4.433	136	

RF. Teniendo en cuenta solo el porcentaje de error de comisión, se obtiene un resultado más acertado empleando el algoritmo de MV, cuando se emplean las áreas de entrenamiento modificadas, ya que solo un 7,4% de los píxeles se han clasificado como herbáceo de seco no perteneciendo a esta clase, mientras que con RF un 27,4% de estas celdas se asignan erróneamente a dicha clase. Sin embargo, en lo que respecta al error de omisión en el herbáceo de seco y siguiendo con las áreas de entrenamiento modificadas, es RF el que presenta un error más bajo (19,2%) frente a un 27,5% de celdas que, siendo herbáceo de seco, han sido asignadas a otros usos por MV, mayoritariamente a matorral, arbóreo de seco y herbáceo de regadío.

La clase correspondiente a herbáceo de regadío cuenta con unos porcentajes muy altos de errores con el algoritmo de MV, tanto si se emplean todas las áreas de entrenamiento como si se utilizan solo las modificadas, presentando problemas de clasificación con el arbóreo denso, el herbáceo de seco y las superficies impermeables; sin embargo, con RF la clasificación del herbáceo de regadío mejora considerablemente, sobre todo en

el caso en el que se han empleado las áreas de entrenamiento depuradas, aunque siguen existiendo problemas de confusión con las clases anteriormente mencionadas.

La distinción entre seco y regadío depende mucho de la disponibilidad de agua de regadío, en periodos secos en áreas semiáridas es frecuente encontrar “regadíos no regados” cuya signatura espectral es difícil de distinguir de la de las áreas de seco.

En general, las superficies impermeables y las láminas de agua se han clasificado de forma muy acertada en todos los casos, siendo esta última clase la que presenta los porcentajes de errores más bajos.

Por último, la clase correspondiente a los invernaderos es una de las pocas que no presenta problemas con respecto a aquellas clases que albergan vegetación natural o cultivos, sin embargo, sí que tiende a confundirse con las superficies impermeables debido a las similitudes entre las signaturas espectrales de ambas clases. Llama la atención la diferencia entre el error de omisión y comisión que existe cuando se emplea RF con

Tabla 8. Porcentaje de errores de comisión y omisión para MV y RF.

Clase	MV originales		MV modificados		RF originales		RF modificados	
	Comisión	Omisión	Comisión	Omisión	Comisión	Omisión	Comisión	Omisión
Bosque	8,6	16,7	11,0	12,7	9,1	3,9	12,0	6,1
Matorral	21,2	5,5	16,6	6,5	10,1	9,8	15,8	9,0
Arb. poco denso	45,0	40,1	41,3	38,0	39,7	40,0	38,7	39,5
Arb. denso	28,6	26,3	22,1	32,6	26,6	31,4	16,7	41,7
Herb. seco	18,1	48,4	7,4	27,5	34,7	34,9	27,4	19,2
Herb. regadío	74,7	78,4	84,4	72,3	39,1	43,5	24,1	35,5
Sup. impermeable	14,6	6,7	4,0	12,5	15,1	11,2	16,7	6,7
Lám. agua	0,0	3,1	0,2	0,8	0,4	0,9	0,8	2,3
Invernaderos	22,6	44,1	56,8	26,6	16,2	57,9	1,2	37,5

Tabla 9. Primer cuartil, mediana y tercer cuartil de la fiabilidad global (f) y kappa (k) a partir de los remuestreos de validación con ambos grupos de áreas de entrenamiento.

Modelo	f q25	f q50	f q75	k q25	k q50	k q75
MV (datos originales)	0,777	0,810	0,829	0,740	0,774	0,792
MV (datos modificados)	0,814	0,849	0,861	0,784	0,816	0,832
RF (datos originales)	0,811	0,841	0,856	0,780	0,807	0,828
RF (datos modificados)	0,839	0,851	0,867	0,810	0,824	0,842

áreas de entrenamiento modificadas; apenas un 1% de los píxeles que se clasifican como invernaderos no pertenecen a esta clase, frente a un 37,5% de celdas que son invernaderos y se han clasificado en otros usos de suelo, fundamentalmente como superficies impermeables.

Para concluir, analizando la mediana de los estadísticos obtenidos para ambos métodos (Tabla 9) y siguiendo el criterio de Landis y Koch (1977) para su interpretación (Tabla 3), se observa una mejora significativa en la clasificación llevada a cabo por MV cuando se emplean las áreas de entrenamiento depuradas, pasando de un grado de acuerdo calificado como “sustancial” ($k=0,774$) según Landis y Koch (1977), a un grado de acuerdo “casi perfecto” (0,816). Respecto a los resultados obtenidos para RF, mejoran también cuando se emplean las áreas de entrenamiento modificadas, sin embargo, esta mejoría no es tan importante, puesto que se pasa de un $k=0,807$ a un 0,824 obteniendo en ambos casos una clasificación “casi perfecta”.

5. Conclusiones

Teniendo en cuenta todos los apartados anteriores, se llegan a las siguientes conclusiones:

- Los usos del suelo agrícolas son los más complicados de clasificar, tanto para MV como para RF, siendo estos los que mayor confusión introducen. Todos los tipos de cultivos, en general, dan problemas de clasificación, especialmente el herbáceo de regadío, que presenta problemas de clasificación con el arbóreo denso, el herbáceo de secano y las superficies impermeables. Con RF la clasificación del herbáceo de regadío mejora considerablemente, sobre todo cuando se emplean las áreas de entrenamiento depuradas.
- El arbóreo poco denso, presenta porcentajes altos de errores de omisión y de comisión tanto en MV como en RF, disminuyendo estos errores en los casos en los que se emplean las áreas de

entrenamiento modificadas. Tiene tendencia a confundirse con aquellas clases que contienen vegetación, aunque con el herbáceo de regadío parece distinguirse mejor en todos los casos. La falta de agua para regadío dificulta la distinción entre coberturas de secano y de regadío.

- Las clases correspondientes a usos del suelo con vegetación natural (bosque y matorral) son clasificadas satisfactoriamente por ambos algoritmos, sin embargo, presenta un menor porcentaje de error RF cuando se aplica a las áreas de entrenamiento originales.
- Los suelos formados en su mayor parte por materiales metamórficos se confunden con superficies impermeables, cometiendo el mismo error tanto RF como MV.
- La clasificación realizada por MV y RF respecto de las superficies impermeables es bastante acertada en todos los casos, no existiendo errores significativos.
- Ambos algoritmos clasifican con bastante exactitud las láminas de agua.
- Los errores de omisión para la clase invernaderos son elevados para ambos métodos en todos los casos, presentando el porcentaje más bajo cuando se emplea MV en las áreas de entrenamiento depuradas. Los invernaderos son diferenciados mejor por RF cuando se depuran las áreas de entrenamiento, aunque un porcentaje considerable de celdas que son invernaderos se han clasificado en otros usos de suelo, fundamentalmente como superficies impermeables. La problemática que presenta esta clase se debe a que muchos invernaderos son temporales, y los plásticos que utilizan pueden modificar la reflectividad del terreno.
- Abarcando la superficie total del área de estudio, se consigue una clasificación más exacta con RF, pero si esta clasificación se analiza por zonas de área reducida, dependiendo de las

características de estas, será más conveniente emplear un algoritmo u otro.

Como conclusión final cabe decir que, en el caso de MV mejoran notablemente los resultados cuando se realiza la depuración de las áreas de entrenamiento, obteniendo un porcentaje de acierto del 81,7%; sin embargo, con RF empleando las áreas depuradas, esta mejora no es tan significativa, aunque sí que se obtiene un mejor resultado con respecto al obtenido realizando la clasificación con las áreas de entrenamiento originales, siendo el porcentaje de acierto de 82,3%. Por tanto, se puede concluir que, con ambos métodos la clasificación es más exacta cuando se emplean las áreas de entrenamiento depuradas y, aunque los resultados obtenidos son bastante aceptables tanto para MV como para RF, este último realiza una clasificación más acertada.

Referencias

- Ayala, R., Menenti, M. 2001. Metodología para la búsqueda del mejor clasificador de imágenes de satélite. *Teledetección, Medio Ambiente y Cambio Global*, 469-472.
- Berberoglu, S., Curran, P.J., Lloyd, C.D., Atkinson, P.M. 2007. Texture classification of Mediterranean land cover. *International Journal of Applied Earth Observation and Geoinformation*, 9(3), 322-334. <https://doi.org/10.1016/j.jag.2006.11.004>
- Berk, R.A. 2016. *Statistical learning from a regression perspective*. Springer. <https://doi.org/10.1007/978-3-319-44048-4>
- Breiman, L. 2001. Random Forests *Machine Learning* 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Cánovas, F., Alonso, F., Gomariz, F. 2016. Modificación del algoritmo Random Forest para su empleo en clasificación de imágenes de Teledetección. *Aplicaciones de las Tecnologías de la Información Geográfica (TIG) para el desarrollo económico sostenible*, 359-368.
- Chávez, P.S. 1996. Image-Based Atmospheric Corrections - Revisited and Improved. *Photogrammetric Engineering and Remote Sensing*, 62, 1025-1036.
- Chen, X., Zhao, H., Li, P., Yin, Z. 2006: Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote Sensing of Environment*, 104, 133-146. <https://doi.org/10.1016/j.rse.2005.11.016>
- Chiarito, G., Chiarito, E. 2015. Evaluación de temperaturas urbanas de acuerdo al uso del suelo: Rosario. *Avances en Energías Renovables y Medio Ambiente*, 19, 93-102.
- Chuvieco, E. 1996. *Fundamentos de Teledetección Espacial*. Madrid: Rialp. Tercera Edición Revisada.
- Chuvieco, E. 2010. *Teledetección Ambiental. La observación de la Tierra desde el espacio*. Barcelona: Ariel Ciencia.
- Congedo, L. 2016. Semi-Automatic Classification Plugin Documentation. Último acceso: 14 de mayo, 2018, de <https://doi.org/10.13140/RG.2.2.29474.02242/1>
- Efron, B., Hastie, T. 2018. *Computer Age Statistical Inference* Cambridge University Press.
- Ezzine, H., Bouziane, A., Ouazar, D. 2014. Seasonal comparisons of meteorological and agricultural drought indices in Morocco using open short time-series data. *International Journal of Applied Earth Observation and Geoinformation*, 26(1), 36-48. <https://doi.org/10.1016/j.jag.2013.05.005>
- García, P., Pérez, M.E., García, J.M., Redondo, M.M., Sanz, J.J., Navarro, A. 2014. *Sellado de suelos a partir de teledetección y SIG: estudio en el Tajo medio-alto*. Dpto. de Análisis Geográfico Regional y Geografía Física. Universidad Complutense de Madrid.
- Gomariz, F., Alonso, F., Cánovas, F. 2017. Improving Classification Accuracy of Multi-Temporal Landsat Images by Assessing the Use of Different Algorithms, Textural and Ancillary Information for a Mediterranean Semiarid Area from 2000 to 2015. *Remote Sensing*, 9(10), 1058. <https://doi.org/10.3390/rs9101058>
- Landis, J.R., Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- Maxwell, A.E., Warner T.A., Fang, F. 2018. Implementation of machine-learning classification in remote sensing: an applied review, *International Journal of Remote Sensing*, 39(9), 2784-2817. <https://doi.org/10.1080/01431161.2018.1433343>
- Moran, M.S., Jackson, R.D., Slater, P.N., Teillet, P.M. 1992. Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output. *Remote Sensing of Environment*, 41(2-3), 169-184. [https://doi.org/10.1016/0034-4257\(92\)90076-V](https://doi.org/10.1016/0034-4257(92)90076-V)
- Pérez, L.I., Delegido, J., Rivera, J.P., Verrelst, J. 2015. Análisis de métodos de validación cruzada para la obtención robusta de parámetros biofísicos. *Revista de Teledetección*, 44, 55-65. <https://doi.org/10.4995/raet.2015.4153>

- Santana, L.M., Escobar, L., Capote, P. 2010. Estimación de un índice de calidad ambiental urbano, a partir de imágenes de satélite. *Revista de Geografía Norte Grande*, 45, 77-95. <https://doi.org/10.4067/S0718-34022010000100006>
- Wang, Q.N., Tenhunen, J.D. 2004. Vegetation mapping with multitemporal NDVI in North Eastern China Transect (NECT). *International Journal of Applied Earth Observations and Geoinformation*, 6(1), 17-31. <https://doi.org/10.1016/j.jag.2004.07.002>
- Xu, H. 2006. Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, 27, 3025-3033. <https://doi.org/10.1080/01431160600589179>
- Zhou, Z.H. 2012. *Ensemble methods. Foundations and algorithms* CRCPrss. <https://doi.org/10.1201/b12207>