

OBTENCIÓN DE CARTOGRAFÍAS DE USOS Y COBERTURAS DEL SUELO DE LA DEMARCACIÓN HIDROGRÁFICA DEL SEGURA PARA EL PERIODO 1986-2019, EMPLEANDO TELEDETECCIÓN Y CLASIFICACIÓN DIGITAL DE IMÁGENES

María Isabel Rodríguez Valero¹, Francisco Alonso Sarria¹
Universidad de Murcia

RESUMEN

Los cambios en los usos y coberturas del suelo conllevan consecuencias ambientales de diverso tipo. La clasificación digital de imágenes obtenidas mediante teledetección es una potente herramienta para evaluar el grado de transformación ambiental debido a factores antropogénicos.

El objetivo de este trabajo es desarrollar un esquema de trabajo basado en técnicas de teledetección y clasificación digital de imágenes para obtener cartografías de usos y coberturas del suelo en la Demarcación Hidrográfica del Segura, para el periodo comprendido entre 1986 y 2019. Para ello se ha utilizado el algoritmo de clasificación supervisada Random Forest y como variables predictoras las bandas espectrales de imágenes de Landsat 5, 7 y 8, y cuatro índices normalizados derivados de estas. Aunque la exactitud global conseguida indica que existe un margen de mejora en el ajuste del modelo, las clasificaciones obtenidas se consideran fiables. Respecto a los usos y coberturas del suelo obtenidos tras el proceso de clasificación, se observa una disminución del uso forestal y un aumento de los usos agrícolas, la superficie cubierta por matorral, los suelos desnudos y los invernaderos.

Palabras clave: Random Forest; Landsat; cobertura terrestre; SIG; clasificación digital.

OBTAINING CARTOGRAPHIES OF LAND USE AND LAND COVER OF THE SEGURA RIVER BASIN DISTRICT FOR THE PERIOD 1986-2019, USING REMOTE SENSING AND DIGITAL IMAGE CLASSIFICATION

ABSTRACT

Changes in land use and land cover lead to environmental consequences of various kinds. Digital classification of remotely sensed images is a powerful tool to assess the degree of environmental transformation due to anthropogenic factors.

The aim of this work is to develop a working scheme based on remote sensing techniques and digital image classification to obtain cartographies of land use and land cover in the Segura Hydrographic Demarcation, for the period between 1986 and 2019. For this purpose, the Random Forest supervised classification algorithm has been used and the spectral bands of Landsat 5, 7 and 8 images, and four normalised indices derived from these, have been used as predictor variables. Although the overall accuracy achieved indicates that there is room for improvement in the model fit, the classifications obtained are considered reliable. With regard to the land uses and land covers obtained after the classification process, a decrease in forest use and an increase in agricultural uses, area covered by scrubland, bare soil and greenhouses can be observed.

Key words: Random Forest; Landsat; land cover; GIS; digital classification.

¹ Instituto Universitario del Agua y del Medio Ambiente, Universidad de Murcia. E-mail: mariaisabel.rodriguez1@um.es, alonsarp@um.es

1. INTRODUCCIÓN

A nivel mundial, regional y local, un gran número de factores ambientales, demográficos, económicos y socioculturales tienen influencia en el cambio de uso de suelo (LÓPEZ *et al.*, 2014). Uno de los procedimientos más fiables para evaluar el grado de transformación ambiental debido a factores antropogénicos, es el análisis del cambio de uso o coberturas del suelo (TURNER *et al.*, 1994). El estudio de la evolución de los cambios en el uso del suelo ha suscitado especial interés en los últimos años, debido a sus implicaciones en relación con la pérdida de hábitat, biodiversidad o la capacidad productiva de los ecosistemas (FERNÁNDEZ *et al.*, 2017). A todos estos efectos provocados por los cambios en la cobertura terrestre hay que sumar el aumento del riesgo de inundaciones y la desertificación, que afecta especialmente a la zona mediterránea de la Península Ibérica. El suelo es un recurso limitado y el continuo crecimiento poblacional obliga a disponer de este para cubrir las necesidades humanas. En la actualidad, se suelen utilizar terrenos forestales para usos agrícolas o urbanos, lo que implica la eliminación de la vegetación natural de la superficie además de la roturación de los suelos, en el caso de la agricultura. Además, la reducción o eliminación de la cubierta vegetal provoca un aumento de la escorrentía y eleva la probabilidad de sufrir deslizamientos de tierra, lo que supone un grave peligro para la población. Además, el abandono de los terrenos agrícolas, fenómeno común en la cuenca mediterránea (REY *et al.*, 2007), genera, especialmente en zonas semiáridas, áreas de suelo desnudo erosionables con importantes riesgos de desertificación (KOSMAS *et al.*, 2014). La erosión en la cuenca del Río Segura se ha convertido en un problema relevante, ya que repercute en la calidad de las aguas superficiales además de suponer la pérdida de suelos. A todos estos efectos se suma la pérdida de biodiversidad, que ambientalmente supone uno de los grandes problemas generados por el cambio en el uso del suelo. Ello es debido a que, con la eliminación de la vegetación natural se reduce el número de especies, tanto de flora (desbroce y desmonte) como de fauna (destrucción de hábitats), provocando importantes cambios ecosistémicos además de una homogenización en los paisajes. No obstante, desde el punto de vista socioeconómico, estas modificaciones pueden resultar beneficiosas, ya que podrían suponer el acceso a nuevos mercados y una activación económica, además de la creación de puestos de trabajo y la dinamización de las áreas rurales.

La continua evolución de las tecnologías y la creciente disponibilidad de imágenes de satélite permiten analizar los usos y coberturas del suelo y sus cambios a lo largo de los años. Para ello, se emplea la clasificación digital de imágenes obtenidas mediante teledetección. Una de las grandes ventajas de la aplicación de estas técnicas es su bajo coste, además de un ahorro de tiempo considerable. La clasificación es una forma de aprendizaje con la que se agrupan conjuntos de objetos en clases mediante técnicas de aprendizaje automático o estadísticas (ARANGO *et al.*, 2005). La clasificación digital de imágenes tiene como propósito la categorización automática de todos los píxeles que forman una imagen, de tal forma que cada píxel es asignado a una determinada clase, obteniendo como resultado una cartografía de las categorías objeto de estudio (CHUVIECO, 2010). Dicha clasificación, se puede llevar a cabo empleando métodos supervisados o no supervisados; en este trabajo se ha utilizado la clasificación supervisada que, al contrario de lo que ocurre con la no supervisada, requiere un conocimiento previo del área de estudio y la definición a priori de las clases que se desean discriminar. En clasificación supervisada se utilizan técnicas que provienen del aprendizaje automático y de la estadística; se trata de algoritmos que clasifican el píxel en función de los valores de reflectividad de dicho píxel (CHUVIECO, 2010) y en ocasiones otras variables auxiliares, poniéndolos en relación con los valores registrados para las diferentes clases a partir de las áreas de entrenamiento. De esta forma, es posible obtener un mapa con la clasificación de usos y coberturas del suelo. Entre los métodos predictivos basados en técnicas de aprendizaje automático más utilizados se encuentran Random Forest (RF), las redes neuronales (*Natural Networks*-NN) o las máquinas de vectores de soporte (*Support Vector Machine* – SVM) (CÁNOVAS-GARCÍA, 2012). Estos algoritmos se apoyan sobre un modelo explícito o implícito establecido que posibilita categorizar los patrones analizados, a partir de los datos de entrenamiento (RIVERO, 2014). En este trabajo se ha empleado RF como algoritmo clasificador. En GOMARIZ-CASTILLO *et al.* (2017) se demuestra un rendimiento significativamente mejor de RF sobre SVM y la máxima verosimilitud en este tipo de ambientes, caracterizados por una elevada fragmentación espacial derivada de la fuerte

presión antrópica sobre el territorio. Resultados similares en este tipo de ambientes se obtienen en trabajos como el realizado por SLUITER *et al.* (2010).

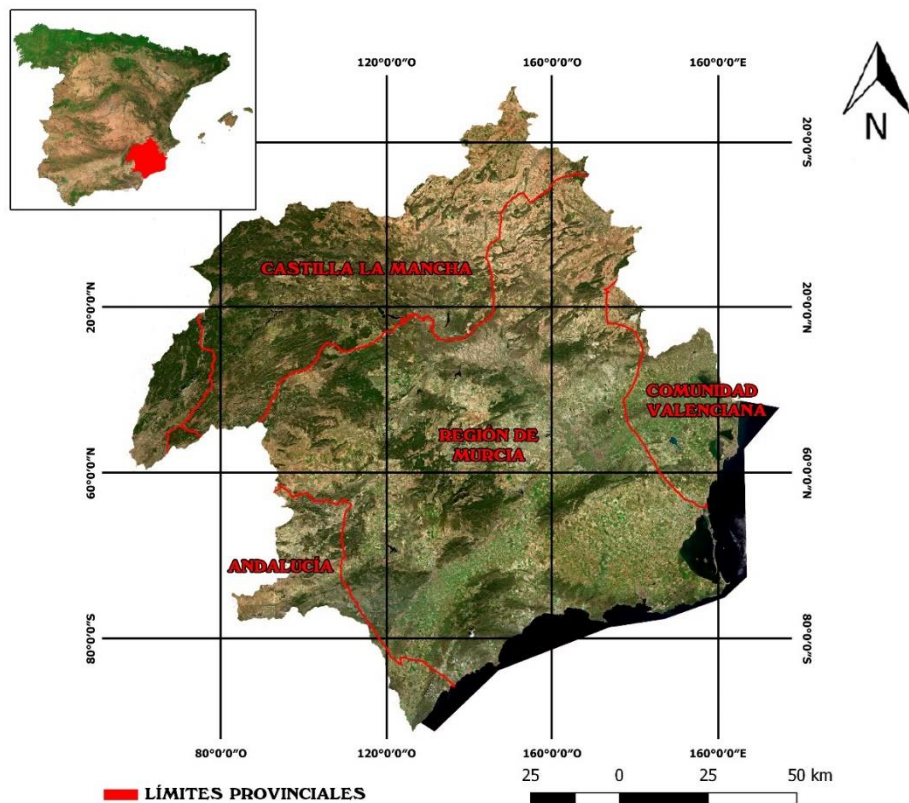
MARTÍNEZ-FERNÁNDEZ *et al.* (2011) o ALONSO-SARRIA *et al.* (2010) han desarrollado sus investigaciones en la Demarcación Hidrográfica del Segura, ya que esta zona resulta de gran interés por ser una de las regiones semiáridas que más modificaciones en los usos y coberturas del suelo ha sufrido.

El objetivo de este trabajo es emplear teledetección y clasificación digital de imágenes para obtener cartografías de usos del suelo de la Demarcación Hidrográfica del Segura, para el periodo comprendido entre 1986 y 2019. De esta forma, es posible obtener una colección histórica de cartografía que permita conocer e interpretar los cambios más significativos experimentados en el área de estudio. Además, en este estudio se propone un esquema de trabajo basado en herramientas de código abierto que pretende ser reproducible a otras zonas de estudio y con fuentes de información adicionales a las utilizadas en el presente trabajo.

2. ÁREA DE ESTUDIO

La Demarcación Hidrográfica del Segura (DHS) se localiza en el sureste de España, y cuenta con una superficie total aproximada de 20.234 km² (Fig. 1). Si se tiene en cuenta solo la parte continental, su extensión es de 19.025 km². Afecta a cuatro comunidades autónomas, englobando a la Región de Murcia casi en su totalidad 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.

FIGURA 1
Localización de la Demarcación Hidrográfica del Segura.

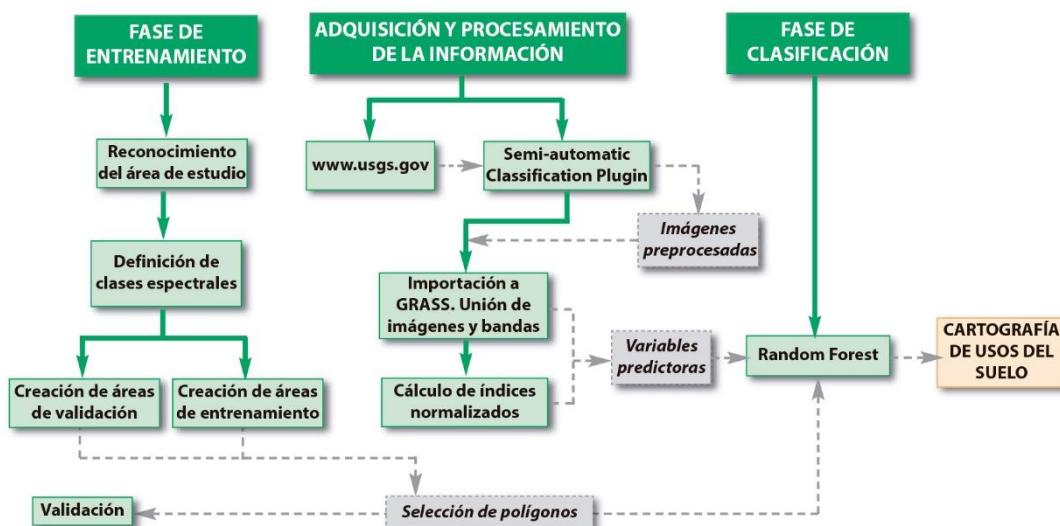


Según los últimos datos del Sistema de Información sobre Ocupación del Suelo de España (SIOSE), recogidos en 2014 (SIOSE, 2015), los principales usos del suelo que se localizan en la DHS son: el forestal, matorral y pastos con un 49% de ocupación total del suelo, seguido del uso agrícola que ocupa un 45%; el suelo artificial supone un 5% del total y las superficies de agua (incluyendo también las zonas húmedas) representan el 1% del suelo restante. Comparando los datos del SIOSE para el año 2014 con los obtenidos por el Sistema de Información Geográfico de Parcelas Agrícolas (SIGPAC) en el año 2010, se observa una clara tendencia al aumento de los suelos agrícolas, pasando de una ocupación del 40,6% en 2010 a un 45% en 2014. Dicho cambio se debe entre otros factores a que la agricultura es un importante motor económico en la zona, lo que en ocasiones ha justificado la transformación de suelos, principalmente forestales, en agrícolas.

3. METODOLOGÍA

La metodología que se ha seguido en este estudio se puede dividir en tres etapas (Fig. 2): a) fase de entrenamiento, b) fase de adquisición y procesamiento de la información y c) fase de clasificación.

FIGURA 2
Esquema de trabajo seguido en este estudio.



Este estudio se ha desarrollado en un entorno Linux utilizando, como software informático para el almacenamiento y gestión de la información espacial los programas SIG QGIS 3.2.2 y GRASS 7.4.1. Son programas de código abierto que se distribuyen bajo la Licencia Pública General (GPL).

Las clasificaciones se han realizado utilizando el programa de análisis de datos R a través de GRASS. R 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>.

3.1 FASE DE ENTRENAMIENTO

3.1.1 Definición de las clases espectrales y creación de las áreas de entrenamiento y validación

La clasificación digital supervisada requiere un conocimiento previo del área de estudio para decidir cuáles serán las clases que se desean discriminar (clases informacionales). En este caso, se han propuesto un total de diez usos del suelo presentes en la DHS (Tabla 1). Cada uno de ellos lleva asignado un código identificador. A continuación, se hace una descripción detallada de cada uno:

1. Bosque: superficies forestales que albergan vegetación de porte arbóreo; los pinares son el grupo predominante en la zona. 2. Matorral: áreas de matorral mediterráneo; es fácilmente localizable en las zonas de solana. 3. Arbóreo de secano: también denominado de secano. En este uso se localizan cultivos como viñedos, olivos o almendros. 4. Arbóreo de regadío: este uso engloba áreas de cultivo de cítricos (limoneros y naranjos principalmente) y frutales de hueso. 5. Herbáceo de secano: es el uso dedicado al cultivo de cereales como el maíz, la cebada o el trigo y algunas leguminosas. 6. Herbáceo de regadío: áreas dedicadas al cultivo hortícola, además de los arrozales. 7. Superficies impermeables: suelos fuertemente antropizados en los que la escorrentía es máxima y no hay infiltración (como las zonas construidas, las infraestructuras de transporte o zonas comerciales e industriales). 8. Láminas de agua: áreas cubiertas de agua (mares, lagos, embalses o balsas de riego). 9. Suelo desnudo: suelos sin cobertura vegetal (como zonas de roquedo o zonas alteradas derivadas de la minería). 10. Invernaderos: suelos dedicados al cultivo en invernaderos; suelen ser estructuras metálicas recubiertas de plástico, aunque también los hay de madera.

TABLA 1
Usos del suelo identificados en la Demarcación Hidrográfica del Segura.

CÓDIGO IDENTIFICADOR	USO
1	Bosque
2	Matorral
3	Arbóreo secano
4	Arbóreo regadío
5	Herbáceo secano
6	Herbáceo regadío
7	Superficies impermeables
8	Láminas de agua
9	Suelo desnudo
10	Invernaderos

Una vez identificadas las clases a utilizar, se crean las áreas de entrenamiento y validación. Estas se obtienen a partir de fotointerpretación multitemporal de fotografías aéreas y mapas de usos del suelo. Las imágenes empleadas para ello han sido: Ortofotografía Digital Pancromática de Andalucía 1984-1985, Ortofotografía Digital Pancromática del SIG Oleícola de Andalucía 1997-98, Ortofotografía de 1997 del SIG Oleícola (Murcia), Ortofotografía Digital de Andalucía coloreada 2001-02, SIGPAC 1997-2003 (ortoimágenes del vuelo SIGPAC realizadas entre 1997 y 2003 sobre la totalidad del territorio español), OLISTAT 1997-1998 (ortoimágenes del vuelo OLISTAT Oleícola realizado entre 1997 y 1998 sobre territorio español) y Ortoimagen PNOA 2016 (ortofotografía de máxima actualidad del proyecto PNOA). El acceso a las ortofotos se ha llevado a cabo a través de QGIS empleando los diferentes servicios *Web Map Service* (WMS) disponibles en el Centro Nacional de Información Geográfica (CNIG).

El procedimiento para la creación de las áreas de entrenamiento y validación puede resumirse en los siguientes pasos: a) se han comparado 3 ortofotografías de las mencionadas anteriormente abarcando un periodo temporal de 20 años; b) si en un área determinada no ha variado el uso del suelo, se procede a la creación del polígono y se le asigna el identificador correspondiente (Tabla 1). Es importante que las áreas de entrenamiento y validación sean lo más representativas posible, ya que esto ayudará a que se consiga una clasificación más aproximada a la realidad. De esta forma, se han creado un total de 1.493 polígonos en toda la DHS.

3.1.2 Selección de polígonos

Para conseguir una clasificación más representativa, se ha llevado a cabo una depuración de las áreas de entrenamiento y validación. Este proceso, utilizado con éxito en esta misma área en trabajo realizado por GOMARIZ-CASTILLO *et al.* (2017), consiste en un análisis de validación cruzada que deja fuera polígonos completos y contabiliza el porcentaje de celdillas que se clasifican correctamente en cada polígono. Si este porcentaje es del 0% significa que el uso del suelo comprendido en el área de entrenamiento ha variado, por tanto, el polígono no es válido y debe ser eliminado del conjunto de áreas de entrenamiento y validación. Podría rebajarse el criterio del 0% y eliminar también aquellos polígonos cuyo porcentaje de acierto fuese inferior a

un umbral dado, pero de esta forma se reduciría significativamente el número de áreas de entrenamiento y la variabilidad de los valores de reflectividad de estas, obteniendo una clasificación poco fiable. Este análisis se ha realizado para cada uno de los años analizados, obteniendo para la clasificación las áreas de entrenamiento depuradas (Tabla 2).

TABLA 2
Número de áreas de entrenamiento utilizadas para cada año tras la depuración.

AÑO	Áreas de entrenamiento y validación
1986	1.106
1990	1.142
1996	1.214
2001	1.239
2010	1.134
2015	1.278
2019	1.282

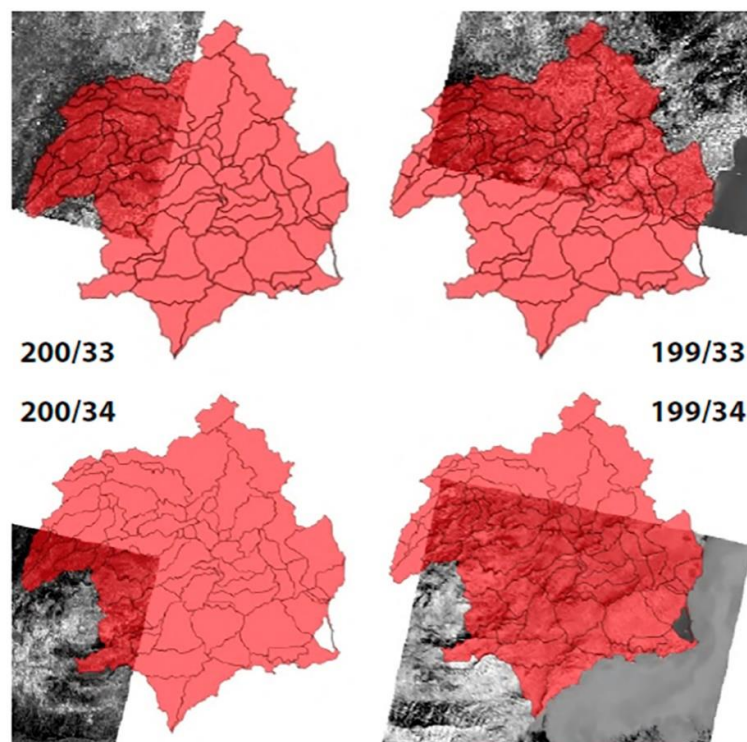
3.2 ADQUISICIÓN Y PROCESAMIENTO DE LA INFORMACIÓN

La DHS se reparte en cuatro escenas de Landsat (Fig. 3), a pesar de que los sensores de este satélite abarcan una superficie superior que la del área de estudio. Las imágenes Landsat utilizadas en este estudio, identificadas por los códigos de paths/rows del Worldwide Reference System of Landsat (WRS) han sido 199/33, 199/34, 200/33 y 200/34.

Para cada uno de los años analizados se han descargado un total de 12 imágenes: cuatro escenas para cada estación del año, excepto para invierno, ya que estas imágenes son las que menos información aportan de las cuatro estaciones (GOMARIZ-CASTILLO *et al.*, 2017) y suelen ser las que presentan un mayor porcentaje de nubosidad. Las imágenes se han obtenido del servidor web del Servicio Geológico de los Estados Unidos (USGS) mediante el plugin de código abierto y libre acceso acceso Semi-Automatic Classification Plugin (SCP) del programa SIG QGIS. Para su adquisición, en este trabajo se han seleccionado las escenas con un rango de nubosidad entre el 0 y el 20%.

FIGURA 3

Ejemplo de la distribución de escenas Landsat necesarias para cubrir a la superficie de la DHS.



A las imágenes digitales se les ha aplicado una corrección atmosférica, que elimina el efecto de los aerosoles y la radiancia intrínseca que se introduce en el sensor y se ve reflejado en la imagen, consecuencia de la interacción del sensor con la atmósfera (AGUILAR *et al.*, 2014). Como algoritmo de corrección atmosférica se ha utilizado el propuesto por CHÁVEZ en 1996; se trata del método de sustracción de objeto oscuro (DOS1) que asume la existencia de una reflectancia de superficie nula o pequeña en determinados píxeles de la imagen, que se ve incrementada por la bruma atmosférica. El valor representativo de esa diferencia se sustrae en cada banda, en todos los píxeles de la escena (BRIZUELA *et al.*, 2007). Dependiendo del satélite Landsat del que provengan las imágenes, estas contarán con más o menos bandas: para cada escena se dispone de, 7 bandas para Landsat 5, 8 bandas para Landsat 7 y 11 en el caso de Landsat 8. En este trabajo se han utilizado, tanto como covariables de entrada en la clasificación como para obtener los índices normalizados, las bandas disponibles en las tres plataformas del espectro visible, infrarrojo cercano e infrarrojo medio (6 en total, con una resolución espacial de 30 m). De forma adicional, se ha utilizado la banda del infrarrojo térmico (con una resolución espacial de 120 m en Landsat 5, 60 m en Landsat 7 y 100 m en Landsat 8) para la obtención del Índice Normalizado de Desnudez del Suelo.

3.2.2 Obtención de los índices espectrales

Descargadas las cuatro escenas Landsat que abarcan la DHS para cada estación del año, se ha llevado a cabo en GRASS la unión de estas por estación y para cada una de las bandas. Como se aprecia en la Figura 3 las imágenes se solapan, por lo que el orden seguido en la unión ha sido: 199/34, 199/33, 200/34 y 200/33. Para verificar que las imágenes no contenían cortes se hicieron composiciones de color RGB, pudiendo además así comprobar que las escenas no contenían nubosidad que hubiese pasado desapercibida para el algoritmo de la NASA.

Con la finalidad de tener un mayor número de variables predictoras que aporten más exactitud al modelo, se han calculado cuatro índices normalizados para cada una de las estaciones del año. Los índices se crean a partir álgebra de mapas y las reflectividades de cada píxel, usando dos o más bandas de una misma imagen; su cálculo se basa en la diferencia entre dos reflectividades divididas entre la suma de esos dos mismos valores y establecen filtros selectivos para aprovechar la mejor respuesta en las distintas bandas, de acuerdo a las reflectividades de los componentes de la superficie (CHIARITO *et al.*, 2015). Los índices normalizados que se han calculado son:

A) Índice Normalizado Diferencial de Vegetación (NDVI)

Es uno de los índices más conocidos y utilizados; se basa en la reflectividad de longitudes de onda pertenecientes a la parte óptica del espectro electromagnético (ANAYA *et al.*, 2013). Se creó con el objetivo de separar la vegetación del brillo que produce el suelo (ROUSE *et al.*, 1974) y se ha correlacionado con el contenido de clorofila en la hoja, el flujo neto de CO₂, la radiación fotosintéticamente activa, la productividad neta de la vegetación, el índice de área foliar o la dinámica fenológica entre otros (CHUVIECO, 2010). Se calcula mediante la diferencia de la reflectancia del infrarrojo cercano y el visible rojo dividido entre la suma de estas dos bandas. La ecuación es la siguiente:

$$NDVI = \frac{NIR - VISr}{NIR + VISr}$$

Donde NIR es la banda espectral del infrarrojo cercano y VISr la banda espectral del visible rojo.

B) Índice Normalizado de Edificaciones (NDBI)

Este índice permite estimar zonas con superficies edificadas (CHEN *et al.*, 2006) junto a las superficies desnudas (SANTANA *et al.*, 2010) midiendo la proporción de edificios en cada píxel. Presenta ciertas limitaciones ya que, en zonas semiáridas en las que los suelos tienen poca vegetación, es difícil que este índice permita acotar las áreas edificadas debido a que la

reflectividad de ambas cubiertas es muy similar en estas regiones del espectro electromagnético (GARCÍA *et al.*, 2014). Se calcula como la diferencia entre la reflectancia del infrarrojo de onda corta y el infrarrojo cercano dividido entre la suma de ambas bandas; su ecuación es la siguiente:

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR}$$

Donde SWIR es la banda espectral del infrarrojo de onda corta y NIR la banda espectral del infrarrojo cercano.

C) Índice Normalizado de Desnudez del Suelo (NDBaI)

Identifica suelos que apenas tienen vegetación; en este caso los cálculos se realizan con valores de niveles digitales (ND) en lugar de reflectividades, ya que para el cálculo de este índice es necesario emplear la banda del infrarrojo térmico. Su ecuación es la siguiente:

$$NDBaI = \frac{ND_{NIR} - ND_{TIRS}}{ND_{NIR} + ND_{TIRS}}$$

Donde ND_{NIR} es la banda espectral (niveles digitales) del infrarrojo cercano y ND_{TIRS} la banda espectral (niveles digitales) del infrarrojo térmico.

D) Índice Normalizado Modificado del Agua (MNDWI)

Este índice normalizado propuesto por XU en 2006, es una modificación del índice de diferencia de agua normalizado (NDWI); en lugar de utilizar la banda del infrarrojo cercano emplea la del visible verde. A través de este índice es posible identificar masas de agua y zonas de elevada saturación, ya que mide la cantidad de agua presente en un píxel. Su ecuación es la siguiente:

$$MNDWI = \frac{VIS_V - SWIR}{VIS_V + SWIR}$$

Donde VIS_V es la banda espectral del visible verde y SWIR la banda espectral del infrarrojo de onda corta.

3.3 FASE DE CLASIFICACIÓN

En esta fase se aplica el algoritmo clasificador que dará como resultado las cartografías para cada uno de los años seleccionados. En este trabajo se ha utilizado el algoritmo Random Forest (RF) (BREIMAN, 2001), implementado en el programa de análisis de datos R en el paquete randomForest (LIAW *et al.*, 2002). Se trata de un algoritmo de aprendizaje automático basado en un clasificador por ensembles, que combina varios clasificadores individuales (árboles de decisión); cada uno de estos árboles clasifica de forma individual un píxel y este será asignado finalmente a la clase con mayor número de votos.

Un árbol de decisión es un conjunto de particiones organizadas jerárquicamente que se aplican sucesivamente desde la raíz, que incluye todos los casos de la muestra que se quiere clasificar, hasta llegar a un conjunto de nodos terminal del árbol (BREIMAN *et al.*, 1984; QUINLAN, 1993). Todos los píxeles en un mismo nodo terminal se asignan a la clase más frecuente entre ellos. RF utiliza un conjunto de árboles de decisión, por defecto 500, como clasificadores base (RODRÍGUEZ *et al.*, 2012), que se construyen automáticamente empleando subconjuntos de datos diferentes obtenidos mediante bootstrapping (BREIMAN, 1996), y se dejan crecer hasta su máxima extensión posible. El bootstrapping consiste en hacer un remuestreo aleatorio con reemplazamiento en el conjunto de datos originales, obteniendo una nueva muestra del mismo tamaño que tiene las mismas propiedades estadísticas (BERK, 2016). Al existir reemplazamiento, algunos datos pueden ser usados más de una vez en el entrenamiento de los clasificadores individuales, lo que hace que RF sea menos sensible a variaciones de los datos de entrada,

incrementando al mismo tiempo la exactitud de las clasificaciones (BREIMAN, 2001; RODRÍGUEZ *et al.*, 2012). Del subconjunto de variables predictivas seleccionadas al azar, para realizar las divisiones de los nodos de cada árbol se utiliza la que minimice el índice de Gini (G) de las particiones resultantes. El índice de Gini mide la heterogeneidad en las diversas particiones y puede también utilizarse para medir la homogeneidad del árbol de decisión en su conjunto. Su valor oscila entre 0 y 1; un valor bajo indica que el nodo contiene predominantemente observaciones de una sola clase, por lo tanto, cuanto más bajo sea el índice de Gini, más homogéneo será el árbol.

Una de las ventajas de RF es el aporte de una estimación interna de exactitud mediante una forma de validación cruzada (CÁNOVAS-GARCÍA *et al.*, 2016); cada árbol de decisión se entrena con un subconjunto de los datos de entrenamiento obtenido mediante bootstrapping, que deja fuera aproximadamente 1/3 de los casos, estas se utilizarán para estimar el error en cada uno de los árboles. Sin embargo, CÁNOVAS-GARCÍA *et al.*, (2017) demuestran que esta estimación puede sobreestimar la exactitud cuándo se aplica a datos espaciales. Otra ventaja de RF es que permite obtener un ranking de importancia de las variables, pudiendo así conocer cuál de ellas es la que más ha influido en todo el proceso de decisión (DEL TORO *et al.*, 2015).

Los principales parámetros de RF son *mtry* y *ntree*; el primero hace referencia al número de variables que se seleccionarán en cada partición de cada árbol, mientras que *ntree* es el número de árboles que conforman el modelo. En este trabajo se han utilizado los valores que RF emplea por defecto ya que en anteriores trabajos como GOMARIZ-CASTILLO *et al.* (2017) se demuestra que no existen diferencias significativas con RF optimizado.

RF ha comenzado a ser uno de los algoritmos más utilizados en teledetección en la última década y se ha aplicado con éxito en varios problemas de regresión y clasificación de los datos de observación de la Tierra, logrando buenos resultados en general (GISLASON *et al.*, 2006).

En este trabajo se han utilizado como variables predictoras en cada año los valores de reflectividad y los índices normalizados de cada una de las 3 fechas. En total se han empleado 30 variables predictoras para cada clasificación.

3.4 VALIDACIÓN DE LAS CLASIFICACIONES

Para evaluar la capacidad predictiva del modelo se ha llevado a cabo un proceso de validación que estima el error promedio que comete el modelo al intentar predecir nuevas observaciones que no estaban incluidas entre las que se emplearon para generarlo (AMAT, 2016). En este trabajo se ha optado por la validación cruzada.

Dentro de estas técnicas de validación, destaca por su sencillez y por ser uno de los más usados, el método de validación simple o *hold-out*, que consiste en dividir el conjunto de datos en dos subconjuntos, utilizando uno de ellos para entrenar el modelo y el otro para validarlo; con el modelo creado se clasifican los datos de validación. Los estadísticos de error de esta clasificación indican la fiabilidad del método empleado (PÉREZ *et al.*, 2015).

Otra opción es la validación cruzada usando *k* grupos, más conocido como *k-fold*. Esta técnica divide el conjunto de los datos en *k* subconjuntos y aplica la técnica *hold-out* durante *k* iteraciones, utilizando cada vez un subconjunto distinto para validar el modelo entrenado con los restantes (JUNG *et al.*, 2015). El error se calcula como la media aritmética de los errores de cada iteración para obtener un único resultado (PÉREZ *et al.*, 2015).

Con la validación del modelo se crea la matriz de confusión, además de una serie de estadísticos que permiten conocer su exactitud y fiabilidad. En la diagonal de esta matriz aparecen los píxeles que han sido asignados correctamente a su clase, por tanto, todos los elementos que quedan fuera de esta diagonal indican errores en la clasificación. 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, p. 501) permitiendo conocer el error de omisión y comisión del modelo. “El primero se refiere a una no asignación a una categoría, mientras el segundo a una delimitación excesivamente amplia” (CHUVIECO, 2010, p. 504), es decir, el error de omisión es el porcentaje de píxeles que perteneciendo a una clase no han sido clasificados como tal y el de comisión el

porcentaje de píxeles clasificados en una clase pero que realmente no pertenecen a ella.

La evaluación de los modelos se determina atendiendo a dos estadísticos: la exactitud y el índice kappa. La exactitud es el número de píxeles clasificados correctamente divididos entre el total de píxeles que componen la imagen; el índice kappa es un estadístico de concordancia que estima la exactitud descontando los aciertos debidos al azar (CHUVIECO, 2010). El valor del índice kappa oscila entre 0 y 1; valores bajos indican un menor grado de acuerdo. Se estima mediante la siguiente ecuación:

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

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.

Para la interpretación del índice kappa, se ha seguido la valoración propuesta por LANDIS y KOCH (1977) (Tabla 3), que indica el grado de concordancia del modelo empleado:

TABLA 3
Valoración del coeficiente kappa (Landis *et al.*, 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

4. RESULTADOS Y DISCUSIÓN

4.1 VALIDACIÓN GLOBAL Y POR CLASE DE LAS CLASIFICACIONES OBTENIDAS

Una vez hechas las clasificaciones, se presentan los mapas de uso del suelo obtenidos para cada uno de los años (Fig. 4, 5, 6, 7, 8, 9 y 10) y los estadísticos de cada clasificación (Tabla 4); las matrices de confusión de las clasificaciones figuran en el Anexo I.

TABLA 4
Porcentaje de acierto, kappa y grado de acuerdo para cada clasificación.

Año	Kappa	Grado de acuerdo	% acierto
1986	0,605	Sustancial	66
1990	0,627	Sustancial	68
1996	0,596	Moderado	65
2001	0,647	Sustancial	70
2010	0,605	Sustancial	66
2015	0,689	Sustancial	73
2019	0,695	Sustancial	74

FIGURA 4
 Mapa de usos y coberturas del suelo en la DHS para el año 1986.

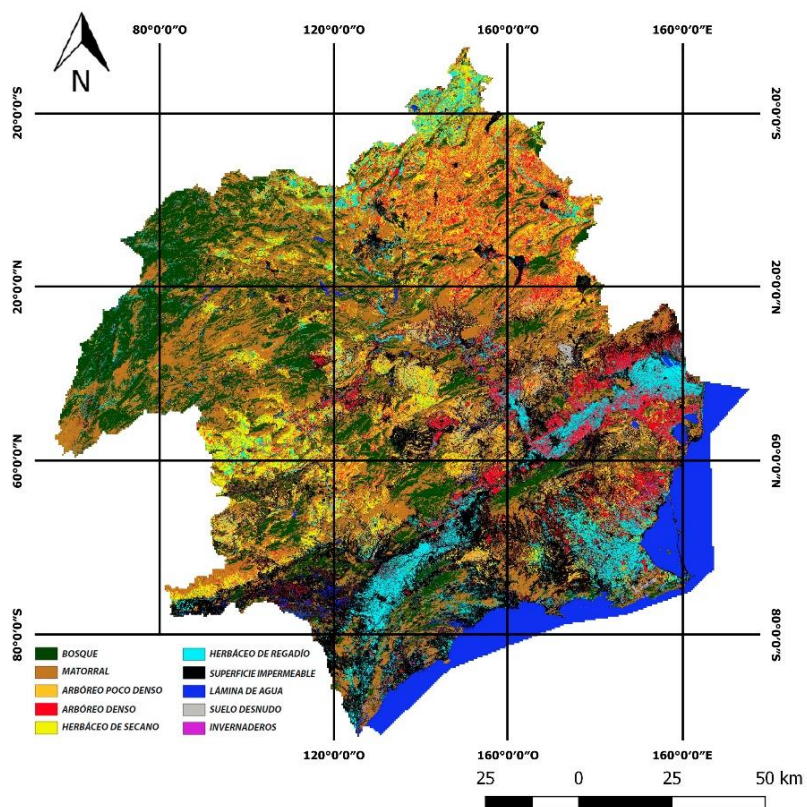


FIGURA 5
 Mapa de usos y coberturas del suelo en la DHS para el año 1990.

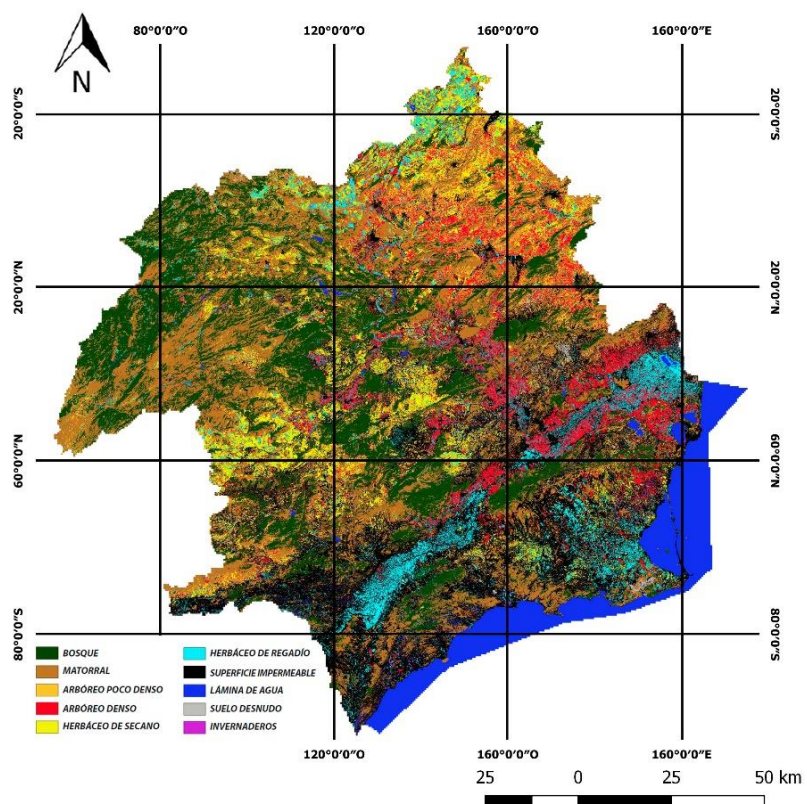


FIGURA 6

Mapa de usos y coberturas del suelo en la DHS para el año 1996.

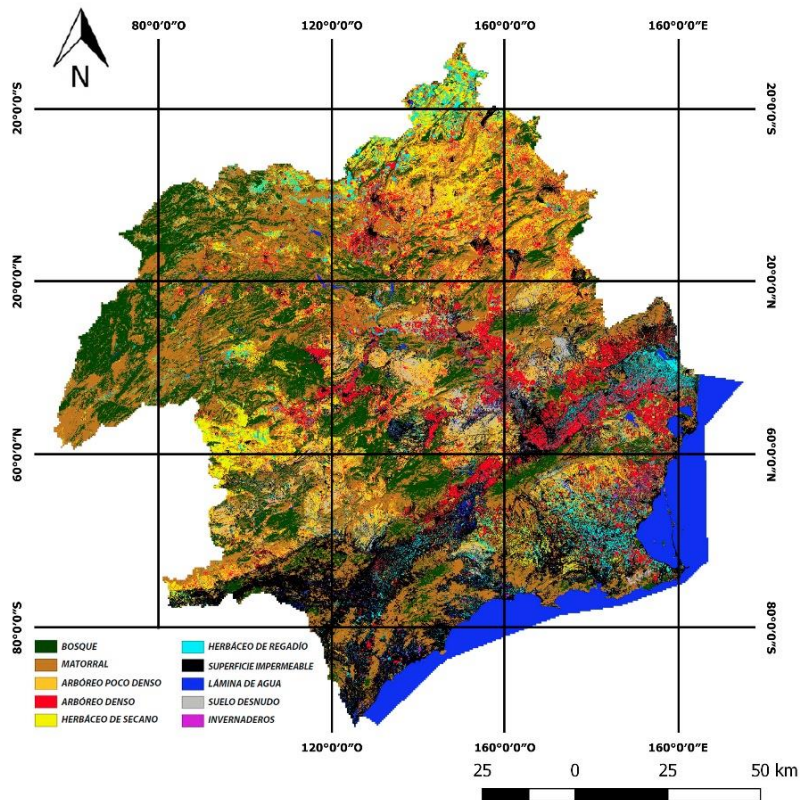


FIGURA 7

Mapa de usos y coberturas del suelo en la DHS para el año 2001.

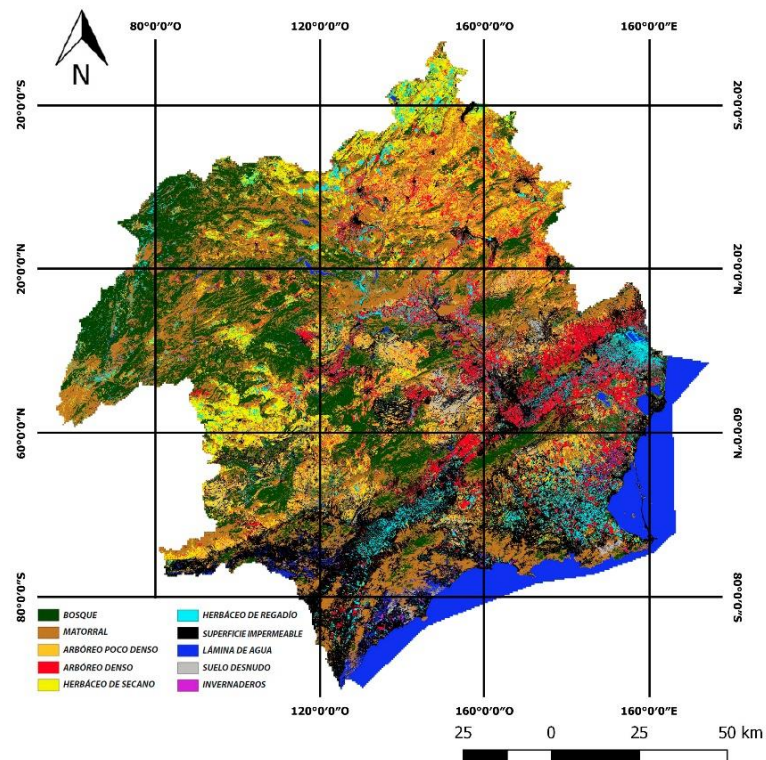


FIGURA 8
 Mapa de usos y coberturas del suelo en la DHS para el año 2010.

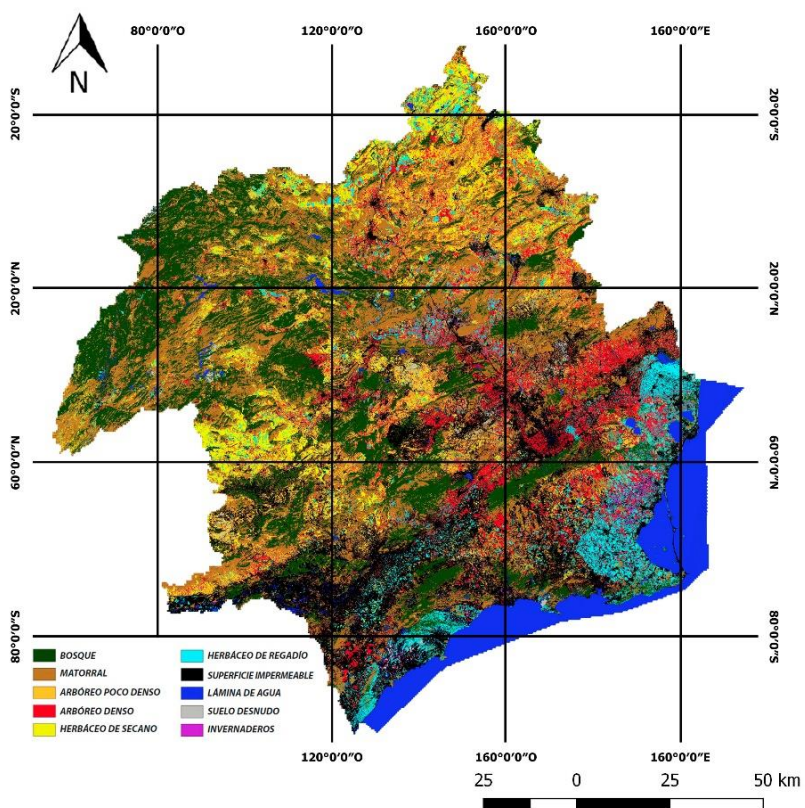


FIGURA 9
 Mapa de usos y coberturas del suelo en la DHS para el año 2015.

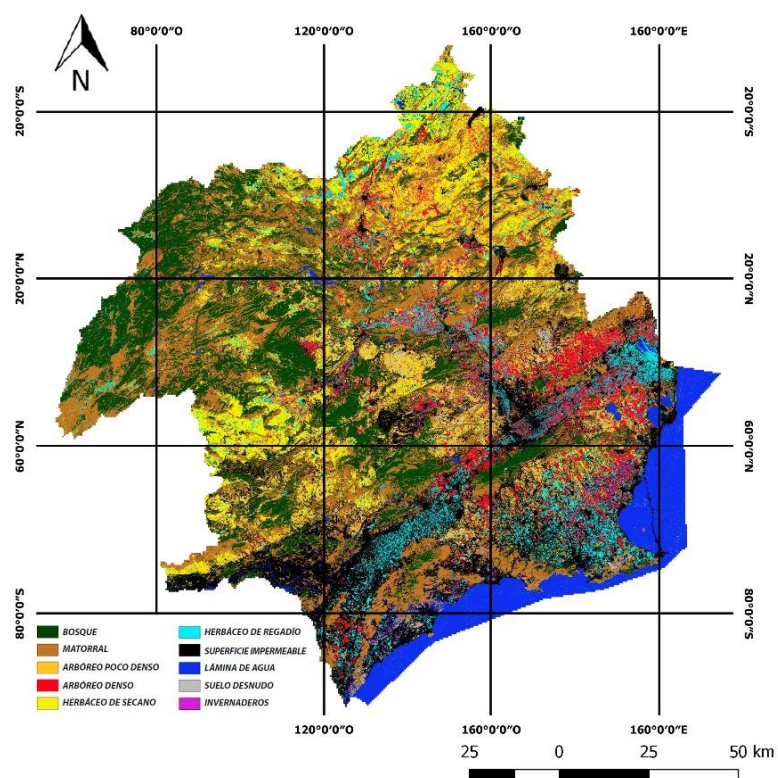
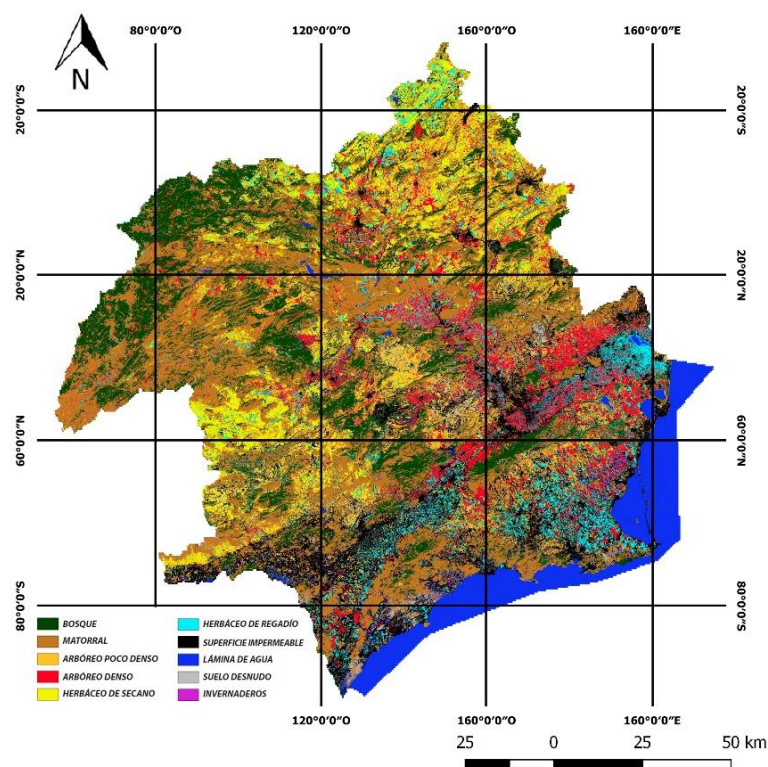


FIGURA 10
Mapa de usos y coberturas del suelo en la DHS para el año 2019.



Con el análisis visual de las cartografías se pueden señalar algunos aspectos relevantes. Las zonas boscosas de las sierras de la DHS se distinguen claramente, y en general, se pone de manifiesto el protagonismo del matorral en el área de estudio, siendo la clase predominante. Respecto a los suelos agrícolas, el herbáceo de regadío se localiza fundamentalmente en el campo de Cartagena, las zonas de Lorca, Totana y en el noreste de la DHS. Se aprecia como el arbóreo de regadío se concentra en la zona de la huerta de Murcia, Caravaca, Moratalla, Cieza y en la zona suroeste de la Comunidad Valenciana. Respecto al secano, el herbáceo predomina en el norte de la DHS, principalmente en la zona de Albacete, y en la zona oeste del área de estudio. Los núcleos urbanos más poblados se distinguen claramente, sin embargo, en los mapas correspondientes a los años 1996, 2001 y 2010 se observa como las superficies impermeables han sido sobreestimadas notablemente, asignándose esta clase a zonas ocupadas por herbáceo de regadío (zona Totana y Lorca, fundamentalmente). Es habitual la utilización de plásticos protectores para maximizar el rendimiento agrícola, lo que hace posible que esta confusión entre signaturas espectrales sea debido a esto, ya que los valores de reflectividad correspondientes al regadío se asemejan a los de las superficies impermeables.

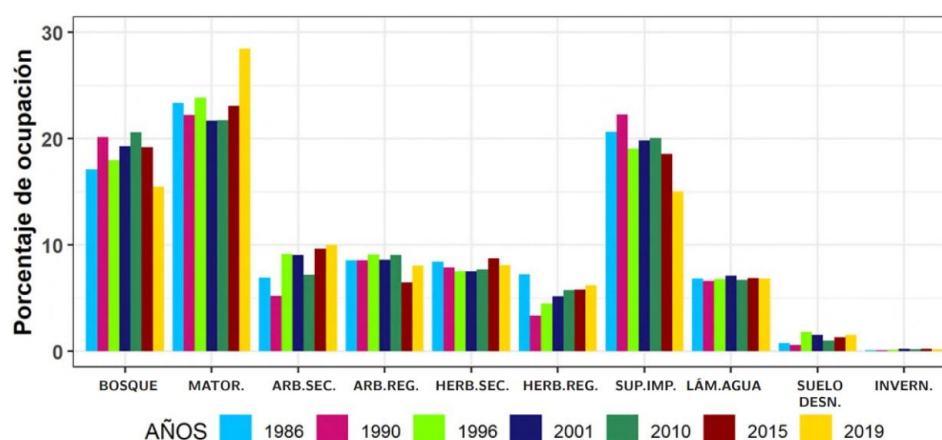
Aunque los estadísticos obtenidos en las clasificaciones y las cartografías de usos del suelo indican que los resultados son aceptables, se esperaba obtener una fiabilidad mayor y un mejor ajuste del modelo ya que hay una sobreestimación en la clase que representa las superficies impermeables. Una opción para lograr una mejora en los resultados puede ser la creación de nuevas áreas de entrenamiento que tras un análisis de validación cruzada su porcentaje de acierto sea del 100%; el inconveniente es que conseguir una cantidad de polígonos representativa de cada una de las clases que cumpla estas condiciones requiere una inversión de tiempo considerable. También el uso de variables auxiliares como mapas de altitud, orientación, pendiente o bandas de textura, podría incrementar la fiabilidad, pero para ello se debe repetir la fase de entrenamiento teniendo en cuenta estas variables a la hora crear los nuevos polígonos.

4.2 EVALUACIÓN DE LOS CAMBIOS DE LAS COBERTURAS DE USO DEL SUELO

En la Figura 11 se resume el porcentaje de ocupación por año de las coberturas del uso del

suelo en la DHS durante el periodo analizado, elaborado a partir de las cartografías obtenidas:

FIGURA 11
Evolución de los usos y coberturas del suelo en la Demarcación Hidrográfica del Segura entre 1986 y 2019.



En los años 80, la cubierta forestal de la DHS era de un 17%, llegando a ocupar en el 2010 más de un 20% de esta; sin embargo, se observa como en 2019 este uso ha sufrido una significativa disminución, suponiendo un 15% aproximadamente de la superficie total de la DHS. Esta reducción se produce fundamentalmente en la zona noroeste de la DHS, concretamente en la Sierra de Cazorla y en el Calar del Mundo; también la Sierra de la Almenara, Sierra de Moratalla y los alrededores del Embalse del Cenajo acusan esta disminución de la cubierta forestal. En su lugar, se observan amplias zonas de matorral, que es el uso con mayor presencia en todos los años analizados. Aunque en 1986 ya contaba con un alto porcentaje de ocupación (más del 23%) y no ha seguido una clara tendencia de crecimiento, en 2019 experimenta un incremento importante situándose por encima del 28%. Las principales causas de la deforestación suelen ser el aprovechamiento de los terrenos para actividades agrícolas o ganaderas y la expansión de zonas urbanas, además de los incendios. El abandono de terrenos agrícolas deja a los suelos sometidos a la dinámica del matorral; este desarrollo a menudo es considerado como una consecuencia paisajística negativa, sin embargo, es posible que el contexto sociocultural haga variar esta percepción (ROBLEDANO *et al.*, 2016).

En 1986 la clase del arbóreo poco denso ocupaba casi un 7% de la superficie de la DHS, sufriendo una disminución en 1990 debido, probablemente, a la fuerte sequía sufrida en ese año que se extendió hasta 1995; este comportamiento puede deberse a que los cultivos de secano se suelen regar fundamentalmente con el agua de lluvia. Se observa como en 1996 este uso vuelve a aumentar su extensión por encima del 9%, al igual que en 2001, para descender en 2010 de nuevo y situarse en un 7%. Este descenso pudo deberse a otra situación de escasez hídrica sufrida durante los años 2004 y 2005. Entre 2015 y 2019 la superficie ocupada por el arbóreo de secano aumenta, llegando a ocupar este último año un total del 10% de la DHS. Si bien el arbóreo de secano es un tipo de cultivo permanente, dado que los resultados obtenidos mediante técnicas de teledetección dependen de forma directa de la respuesta espectral de dicho uso, es probable que los periodos de sequía hayan influido significativamente en los resultados finales, al afectar en el comportamiento vegetativo de los cultivos. Por el contrario, la extensión del arbóreo de regadío en la DHS presenta cierta estabilidad; en 1986 su superficie era de un 8,5%, manteniéndose alrededor de este porcentaje hasta el año 2015, que se redujo a un 6,4%, para en 2019 volver a situarse en torno al 8%.

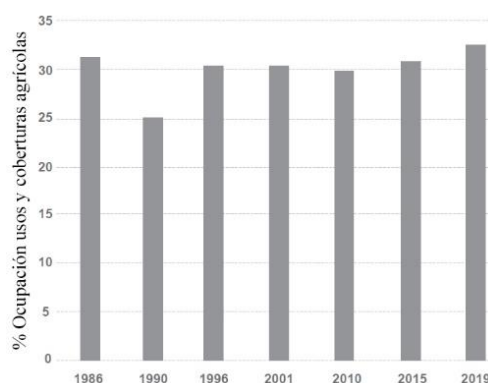
La clase correspondiente al herbáceo de secano no sufre variaciones significativas en su extensión, oscilando esta entre el 7,5% y el 8,7% de la superficie total de la DHS. Por el contrario, el herbáceo de regadío sufre una disminución drástica en 1990 respecto del 86, reduciéndose en casi un 4% la superficie ocupada por este uso; concretamente pasa de un 7,3% en 1986 a un 3,4% en 1990. La DHS es una región semiárida que experimenta importantes periodos de sequía. Entre 1981 y 1986 se produjo uno de estos ciclos secos (RUIZ-ÁLVAREZ *et al.*, 2021) que

probablemente influyó en la reducción de la extensión del herbáceo de regadío. En los siguientes años analizados se observa como este uso va aumentando su extensión, hasta llegar en 2019 a ocupar un 6,2% de la demarcación.

Los usos del suelo agrícolas son los más complicados de clasificar y todos los tipos de cultivos, en general, dan problemas de clasificación confundiendo unas clases con otras. El elevado fraccionamiento parcelario y la heterogeneidad de los cultivos en la DHS dificulta la clasificación de las cubiertas agrícolas. Por ello, se han agrupado las clases correspondientes a usos agrícolas (Fig. 12) para analizar su evolución de forma global.

FIGURA 12

Gráfico que muestra la evolución de los usos del suelo agrícolas en la Demarcación Hidrográfica del Segura entre 1986 y 2019.



Como se observa en el gráfico de la Figura 12, los suelos dedicados a la agricultura ocupan una gran superficie de la DHS, suponiendo más de un 30% de esta, excepto en los años 90, donde el suelo agrícola se reduce ligeramente, llegando a ocupar el 25% de la DHS.

La clase de las láminas de agua se mantiene estable durante el periodo analizado, sin sufrir variaciones significativas. Los suelos desnudos junto con los invernaderos ocupan una mínima superficie de la DHS, y aunque existen pequeñas variaciones en la extensión de estas clases en los años analizados, ambas presentan una tendencia al aumento.

Respecto a las superficies impermeables se detectó una elevada sobreestimación de este uso en las cartografías obtenidas, por lo que se considera que los porcentajes mostrados en el gráfico de la Fig. 11 para este uso no se ajustan a la realidad. Uno de los errores más significativos de esta sobreestimación se localiza en la zona suroeste de la DHS, donde el clasificador confunde una zona de suelos de naturaleza metamórfica con superficies impermeables; este área debería haberse clasificado como suelo desnudo. Probablemente este error se deba a la similitud de los valores de reflectividad de los píxeles pertenecientes a estas clases.

5. CONCLUSIONES

Los estadísticos obtenidos en las distintas clasificaciones indican que existe un margen de error considerable y se esperaba obtener una exactitud superior, por lo que en futuras investigaciones se plantea utilizar en cada caso aquellas bandas e índices que resulten más útiles en función del objetivo fijado, además de información textural y contextual. PADIAL *et al.* (2019) logra obtener una buena discriminación en las cubiertas naturales y especialmente en las agrícolas con la aplicación de filtros multicriterio a las áreas de entrenamiento basados en el NDVI, que permiten la exclusión de píxeles ajenos a las categorías de interés; con las coberturas de urbano y suelo desnudo no se consigue una buena discriminación. GUAN *et al.* (2012) propone obtener de forma automática una selección de características de las áreas urbanas con RF utilizando LIDAR. Por otra parte, RIAÑO *et al.* (2010) plantea recurrir a datos LIDAR para identificar las diferentes alturas de la vegetación, así podría solventarse la confusión entre clases de porte arbóreo y

herbáceo. No solo la selección de información mejora las clasificaciones ya que en otros trabajos (ALONSO-SARRIA *et al.*, 2019) incrementan la exactitud combinando distintas técnicas de aprendizaje automático. No obstante, teniendo en cuenta que el número de variables predictoras del modelo no es elevado (30 por cada clasificación) y observando las cartografías, se considera que las clasificaciones obtenidas son fiables y cumplen con el objetivo del estudio. Además, la metodología propuesta permite obtener una serie temporal y espacial de mapas de usos y coberturas del suelo para conocer los cambios más significativos que ha experimentado la superficie terrestre del área de estudio.

Los usos agrícolas tienen un alto porcentaje de ocupación en la DHS, superando el 30% de la superficie casi todos los años analizados, alcanzando su ocupación máxima en 2019. Esta tendencia es de gran importancia, dado que la agricultura es una actividad que juega un importante papel en la DHS y modifica significativamente la cobertura natural de la superficie de la zona. Los invernaderos y el suelo desnudo son las únicas clases que presentan una tendencia al aumento, mientras que los porcentajes de ocupación del resto de clases oscilan en los años analizados, excepto las láminas de agua que no muestran cambios significativos. En cuanto a la cubierta forestal, su extensión se ve considerablemente reducida en 2019 respecto a 1986, al contrario de lo que ocurre con el matorral; su ocupación aproximada en 2019 es de casi el 30% de la DHS. Respecto a los usos del suelo, teniendo en cuenta la dinámica socioeconómica de la demarcación, cabe esperar una disminución de los suelos forestales y un aumento de los usos agrícolas, el matorral, los suelos impermeables y los suelos desnudos.

A la vista de los resultados globales obtenidos se puede concluir que hay una disminución generalizada de la cobertura forestal y un aumento de los usos agrícolas, la superficie cubierta por matorral y los suelos desnudos. Respecto a las superficies impermeables no ha sido posible analizar su evolución, puesto que el modelo sobreestima de forma significativa este uso.

6. BIBLIOGRAFÍA

- AGUILAR, H., MORA, R. y VARGAS, C. (2014): *Revista Geográfica de América Central*. “Metodología para la corrección atmosférica de imágenes Aster, Rapideye, Spot 2 y Landsat 8 con el módulo Flaash del software Envi, 53, pp. 39-59.
- ALONSO-SARRIA, F., VALDIVIESO-ROS, C., RODRÍGUEZ-VALERO, M. y GOMARIZ-CASTILLO, F. (2019): *XVIII Congreso de Teledetección. Revista Teledetección*. “Combinación de random forest y redes neuronales para clasificar usos del suelo”, 337-340.
- ALONSO-SARRIA, F., GOMARIZ-CASTILLO, F. y CÁNOVAS-GARCÍA (2010): *Cuaternario y geomorfología: Revista de la Sociedad Española de Geomorfología y Asociación Española para el Estudio del Cuaternario*. “Análisis temporal de los cambios de usos del suelo en la cuenca del Segura mediante teledetección. Implicaciones sobre la degradación”, 24(3), pp. 71-86.
- AMAT, J. (2016): *Technical report*. “Validación de modelos de regresión: Cross-validation, oneleaveout, bootstrap”. Recuperado de: https://rpubs.com/Joaquin_AR/238251
- ANAYA, J.A. y VALENCIA, G.M. (2013): *Geofocus*. “Fenología de ambientes tropicales en el marco de la Teledetección”, 13-2, pp. 195-211.
- ARANGO, M., BRANCH, J.W. y BOTERO, V. (2005): *Revista Facultad Nacional Agronomía Medellín*. “Clasificación no supervisada de coberturas vegetales sobre imágenes digitales de sensores remotos: Landsat -ETM+”, Vol. 58, N°1, pp. 2611-2634.
- BERK, R.A., (2016): *Springer*. “Statistical learning from a regression perspective”. En Springer International Publishing (Ed), 347 pp.
- BREIMAN, L., (1996): *Machine Learning*. “Bagging predictors”, 24 (2), pp.123-140.
- BREIMAN, L., (2001): *Machine Learning*. “Random forest”, 45(1), pp. 5-32.
- BREIMAN, L., FRIEDMAN, J., OLSHEN, R. y STONE, C. (1984): *Wadsworth International Group*. “Classification and Regression Trees”.
- BRIZUELA, A., AGUIRRE, C. y VELASCO, I. (2007): *Teledetección: hacia un mejor entendimiento de la dinámica global y regional*. “Aplicación de métodos de corrección

- atmosférica de datos Landsat 5 para análisis multitemporal”. En Martin (Ed). Madrid, pp. 207-214.
- CÁNOVAS-GARCÍA, F. (2012): *Tesis doctoral, Universidad de Murcia*. “Análisis de imágenes basado en objetos (OBIA) y aprendizaje automático para la obtención de mapas de coberturas del suelo a partir de imágenes de muy alta resolución espacial: aplicación en la Unidad de Demanda Agraria nº 28, cabecera del Argos”.
- CÁNOVAS-GARCÍA, F., ALONSO-SARRIA, F. y GOMARIZ-CASTILLO, F. (2016): *Aplicaciones de las Tecnologías de la Información Geográfica (TIG) para el desarrollo económico sostenible*. “Modificación del algoritmo Random Forest para su empleo en clasificación de imágenes de Teledetección”, pp. 359-368.
- CHÁVEZ, P.S. (1988): *Remote Sensing of Environment*. “An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data”, 24:459-479.
- CHÁVEZ, P.S. (1996): *Photogrammetric Engineering and Remote Sensing*. “Image-Based Atmospheric Corrections - Revisited and Improved”, 62, pp. 1025-1036.
- CHEN, X., ZHAO, H., LI, P. y YIN, Z. (2006): *Remote Sensing of Environment*. “Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes”, 104, pp. 133-146.
- CHIARITO, G. y CHIARITO, E. (2015): *Facultad de Ciencias Exactas, Ingeniería y Agrimensura de Rosario*. “Evaluación de temperaturas urbanas de acuerdo al uso del suelo: Rosario”.
- CHUVIECO, E (2010): “Teledetección Ambiental. La observación de la Tierra desde el espacio”. Ed. Ariel Ciencia, Barcelona.
- DEL TORO, N., GOMARIZ-CASTILLO, F., CÁNOVAS-GARCÍA, F. y ALONSO-SARRIA, F. (2015): *Boletín de la Asociación de Geógrafos Españoles*. “Comparación de clasificación de imágenes de satélite en la cuenca del río Argos (Región de Murcia)”, 67, pp. 327-347.
- FERNÁNDEZ, D. y CORBELL, E. (2017): *Biblio3W, Revista Bibliográfica de Geografía y Ciencias Sociales*. “Cambios en los usos de suelo en la Península Ibérica: un meta-análisis para el período 1985-2015”, Vol. XXII (1215).
- GARCÍA, P., PÉREZ, M.E., GARCÍA, J.M., REDONDO, M.M., SANZ, J.J. y NAVARRO, A. (2014): *Universidad Complutense de Madrid*. “Sellado de suelos a partir de teledetección y SIG: estudio en el Tajo medio-alto”.
- GISLASON, P.O., BENEDIKTSSON, J.A. y SVEINSSON, J.R. (2006): *Pattern Recognition Letters*. “Random Forests for land cover classification”, 27(4), pp. 294-300.
- GOMARIZ-CASTILLO, F., ALONSO-SARRIA, F. y CÁNOVAS-GARCÍA, F. (2017): *Remote Sensing*. “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”, 9(10), p. 1058.
- GUAN, H., YU, J., LIA, J. y LUO, L. (2012): *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. “Random Forest-Based feature selection for land-use classification using LIDAR data and orthoimagery”.
- JUNG, Y. y HU, J. (2015): *Journal of Nonparametric Statistics*. “A k-fold averaging crossvalidation procedure”, 27(2), pp. 167-179.
- KOSMAS, C. et al. (2014): *Catena*. “An exploratory analysis of land abandonment drivers in areas prone to desertification”, 128, pp. 252-261.
- LANDIS, J.R. y KOCH, G.G. (1977): *Biometrics*. “The measurement of observer agreement for categorical data”, 33, pp. 159-174.
- LIAW, A. y WIENER, M. (2002): *R news*. “Classification and Regression by Random Forest”, 2(3), pp. 18-22.
- LÓPEZ, V.H., BALDERAS, M.A., CHÁVEZ, M.C., JUAN, J.I. y GUTIÉRREZ, J.G. (2014): *CIENCIA ergo sum*. “Cambio de uso de suelo e implicaciones socioeconómicas en un área mazahua del altiplano mexicano”, 22-2, pp. 136-144.
- MARTÍNEZ-FERNÁNDEZ, J., ESTEVE-SELMA, M.A., GUAITA, N. y FITZ, C. (2011): Biodiversidad en España. Una visión desde la sostenibilidad ante el cambio global. “Cambios de ocupación del suelo en la cuenca del Segura y sus implicaciones para la biodiversidad”, capítulo 3, pp. 120-174. En OSE-Ediciones Mundi-Prensa.

- MORAN, M.S., JACKSON, R.D., SLATER, P.N. y TEILLET, P.M. (1992): *Remote Sensing of Environment*. "Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output", 41(2:3), pp. 169-184.
- PADIAL, M., VIDAL, J.J.; SERRA, P.; NINYEROLA, M. y PONS, X. (2019): *XVIII Congreso de Teledetección. Revista Teledetección*. "Aplicación de filtros multicriterio basados en NDVI para la extracción de áreas de entrenamiento desde la base de datos SIOSE", 311-314.
- PÉREZ, LI., DELEGIDO, J., RIVERA, J.P. y VERRELST, J. (2015): *Revista de Teledetección, Asociación Española de Teledetección*. "Análisis de métodos de validación cruzada para la obtención robusta de parámetros biofísicos". 44, pp. 55-65.
- QUINLAN, J.R. (1993): *C4.5: Programs for Machine Learning*. En Morgan Kaurmann (Ed), San Mateo.
- REY, J.M., MARTINS, A., NICOLAU, J.M. y SCHULZ, J. (2007): *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*. "Abandonment of agricultural land: an overview of drivers and consequences", nº 57.
- RIANO, D., SALAS, J. y CHUVIECO, E. (2010): *Tecnologías Geográficas para el Desarrollo Sostenible*. "Corrección Atmosférica y Topográfica, Información Multitemporal y Auxiliarias Aplicadas a la Cartografía de Modelos de Combustibles con Imágenes Landsat-TM", Ed: Universidad de Alcalá: Asociación Española de Geografía, pp. 222-239.
- RIVERO, J. (2014): *Revista Cubana de Ciencias Informáticas*. "Técnicas de aprendizaje automático para la detección de intrusos en redes de computadoras", 8 (4), pp. 52-73.
- ROBLEDANO, F., ROMERO, A., BELMONTE, F., FERNÁNDEZ, M.Á., MARTÍNEZ, C., SÁNCHEZ, M. y ZAPATA, V.M. (2016): *Abandono de cultivos en la Región de Murcia: consecuencias ecogeomorfológicas*. "Consecuencias del abandono de cultivos en la Región de Murcia y recomendaciones de gestión", Universidad de Murcia, pp. 227-263.
- RODRÍGUEZ, V.F., GHIMIRE, B., ROGAN, J., CHICA, M. y RIGOL, J.P. (2012): *ISPRS Journal of Photogrammetry and Remote Sensing*. "An assessment of the effectiveness of a random forest classifier for land-cover classification", 67, pp. 93-104.
- ROUSE, J.W., HAAS, R. H., DEERING, D.W. y SEHELL, J.A. (1974): *Final Rep. RSC 1978-4, Remote Sensing Center*. "Monitoring the vernal advancement and retrogradation (Green wave effect) of natural vegetation", Texas A&M Univ., College Station.
- RUIZ-ÁLVAREZ, V., PORTELA, M.M., ESPINOSA, L.Á., GARCÍA, R., BELMONTE, F. y RUIZ, M. (2021): *Boletín de la Asociación de Geógrafos Españoles*. "Regionalization of droughts from the SPI in the Segura hydrographic basin and Commonwealth of Taibilla Channels (Southeast of Spain)", 88. Recuperado de: <https://doi.org/10.21138/bage.3045>
- SANTANA, L.M., ESCOBAR, L. y CAPOTE, P. (2010): *Revista de Geografía Norte Grande*. "Estimación de un índice de calidad ambiental urbano, a partir de imágenes de satélite", 45, pp. 77-95.
- SIOSE (2015): *Instituto Geográfico Nacional*. "Documento técnico SIOSE 2011, Versión 1.2". Ministerio de Fomento. Gobierno de España., Madrid, España.
- SLUITER, R. y PEBESMA, E.J. (2010): *Int. J. Remote Sensing*. "Comparing techniques for vegetation classification using multi- and hyperspectral images and ancillary environmental data", 31, 6143-6161.
- TURNER, B., MEYER, W.B. y SKOLE, D.L. (1994): *Ambio*. "Global land-use/land-cover change: towards an integrated study", 23(1), pp. 91-95.
- XU, H. (2006): *International Journal of Remote Sensing*. "Modification of normalised difference water index (NDWI) to enhanced open water features in remotely sensed imagery", 27(14), pp. 3025-3033.

Agradecimientos:

Este trabajo fue en parte financiado por el Ministerio de Economía, Industria y Competitividad/Agencia Estatal de Investigación/FEDER (Fondo Europeo de Desarrollo Regional) número CGL2017-84625-C2-2-R.

ANEXO I: MATRICES DE CONFUSIÓN DE LAS CLASIFICACIONES

TABLA 5

Matriz de confusión correspondiente a la clasificación del año 1986.

CLASES	Bosque	Matorr.	Arb. Sec.	Arb. Reg.	Herb. Sec.	Herb. Reg.	Sup. Imp.	Lám. Agua	Suelo desn.	Invern.	Total
Bosque	21.069	4.288	501	165	370	461	377	31	1	0	27.263
Matorral	4.255	21.137	651	1.006	840	597	442	52	42	0	29.022
Arb. Sec.	166	514	3.138	1.897	1.368	442	387	15	123	0	8.050
Arb. Reg.	121	624	2.344	4.359	3.004	3.153	640	21	49	17	14.332
Herb. Sec.	342	963	1.813	3.659	6.125	2.702	604	4	58	5	16.275
Herb. Reg.	197	235	999	1.905	1.149	5.652	434	126	4	19	10.720
Sup. Imp.	406	725	1.119	1.760	932	1.486	18.292	394	541	143	25.798
Lám. Agua	1	9	32	20	0	17	13	20.550	112	0	20.754
Suelo desn.	0	4	38	5	11	1	44	24	325	0	452
Invern.	0	0	0	6	0	0	21	0	0	102	129
Total	26.557	28.499	10.635	14.782	13.799	14.511	21.254	21.217	1.255	286	

TABLA 6

Matriz de confusión correspondiente a la clasificación del año 1990.

CLASES	Bosque	Matorr.	Arb. Sec.	Arb. Reg.	Herb. Sec.	Herb. Reg.	Sup. Imp.	Lám. Agua	Suelo desn.	Invern.	Total
Bosque	21.503	3.788	398	157	400	358	207	79	0	0	26.890
Matorral	3.931	21.443	697	789	1.099	593	445	71	123	0	29.191
Arb. Sec.	134	473	3.545	2.399	1.572	367	359	8	102	1	8.960
Arb. Reg.	177	366	2.022	6.095	2.052	1.817	682	6	232	16	13.465
Herb. Sec.	428	1.255	1.945	2.282	6.428	4.467	579	4	49	0	17.437
Herb. Reg.	84	428	687	2.293	1.341	5.166	499	86	45	2	10.631
Sup. Imp.	242	739	1.270	1.017	832	1.376	18.403	109	819	130	24.937
Lám. Agua	1	11	0	23	1	37	6	20.891	66	5	21.041
Suelo desn.	0	9	19	30	1	7	71	18	138	0	293
Invern.	0	0	0	15	0	0	12	4	0	200	231
Total	3.931	26.500	28.512	10.583	15.100	13.726	14.188	21.263	1.574	354	

TABLA 7

Matriz de confusión correspondiente a la clasificación del año 1996.

CLASES	Bosque	Matorr.	Arb. Sec.	Arb. Reg.	Herb. Sec.	Herb. Reg.	Sup. Imp.	Lám. Agua	Suelo desn.	Invern.	Total
Bosque	20.532	4.810	402	486	439	570	285	32	7	0	27.593
Matorral	4.673	20.242	740	1.036	852	672	421	39	148	0	28.823
Arb. Sec.	189	939	4.002	2.084	2.104	603	487	20	293	5	10.726
Arb. Reg.	298	711	1.523	4.679	2.168	3.077	828	42	80	5	13.411
Herb. Sec.	347	727	1.943	3.098	6.507	2.365	321	1	186	15	15.510
Herb. Reg.	131	339	1.141	1.846	1.110	4.577	300	36	56	0	9.536
Sup. Imp.	352	889	1.444	1.357	676	2.051	18.432	455	676	69	26.401
Lám. Agua	2	9	31	15	0	75	70	20.948	56	15	21.221
Suelo desn.	3	27	172	30	50	47	67	75	270	1	712
Invern.	0	0	0	1	0	0	72	0	0	244	317
Total	26.557	28.693	11.398	14.632	13.906	14.037	21.283	21.648	1.742	354	

TABLA 8

Matriz de confusión correspondiente a la clasificación del año 2001.

CLASES	Bosque	Matorr.	Arb. Sec.	Arb. Reg.	Herb. Sec.	Herb. Reg.	Sup. Imp.	Lám. Agua	Suelo desn.	Invern.	Total
Bosque	21.661	4.423	199	393	365	282	156	22	19	0	27.520
Matorral	4.019	21.362	408	569	841	754	525	94	102	0	28.674
Arb. Sec.	172	485	4.220	3.371	1.376	179	378	13	228	3	10.425
Arb. Reg.	159	399	2.286	5.590	1.539	2.654	411	150	75	0	13.263
Herb. Sec.	237	909	2.235	2.000	7.640	2.930	357	28	134	5	14.477

Herb. Reg.	126	300	215	2.140	1.169	5.983	415	116	26	0	10.490
Sup. Imp.	158	663	1.069	712	817	1.689	18.802	237	442	205	24.821
Lám. Agua	0	12	16	39	1	61	23	21.270	61	2	21.485
Suelo desn.	3	29	180	44	39	8	56	11	609	3	982
Invern.	0	0	0	0	0	0	169	8	0	335	512
Total	26.535	28.582	10.855	12.860	73.787	14.540	21.292	21.949	1.696	553	

TABLA 9

Matriz de confusión correspondiente a la clasificación del año 2010.

CLASES	Bosque	Matorr.	Arb. Sec.	Arb. Reg.	Herb. Sec.	Herb. Reg.	Sup. Imp.	Lám. Agua	Suelo desn.	Invern.	Total
Bosque	20.325	4.012	439	548	292	284	176	100	1	0	26.177
Matorral	4.243	19.754	1.186	804	541	1.033	469	78	145	0	28.253
Arb. Sec.	307	872	4.103	2.733	1.411	456	347	6	152	1	10.388
Arb. Reg.	292	526	2.035	5.376	1.134	2.429	469	9	105	14	12.389
Herb. Sec.	142	381	1.579	1.949	7.474	3.863	300	1	406	0	16.095
Herb. Reg.	113	509	419	2.054	1.730	5.127	208	33	93	3	10.289
Sup. Imp.	212	708	936	917	779	1.409	17.920	198	228	110	23.417
Lám. Agua	2	16	1	42	0	74	20	11.451	3	12	11.621
Suelo desn.	0	74	54	40	56	121	31	2	321	0	699
Invern.	0	0	0	0	0	0	7	1	4	218	230
Total	25.636	26.852	10.752	14.463	13.417	14.796	19.947	11.879	1.458	358	

TABLA 10

Matriz de confusión correspondiente a la clasificación del año 2015.

CLASES	Bosque	Matorr.	Arb. Sec.	Arb. Reg.	Herb. Sec.	Herb. Reg.	Sup. Imp.	Lám. Agua	Suelo desn.	Invern.	Total
Bosque	20.302	4.437	381	960	637	103	322	70	5	0	27.217
Matorral	5.083	21.445	832	443	868	572	529	89	253	0	30.114
Arb. Sec.	429	833	5.607	2.082	1.376	564	501	16	193	0	11.601
Arb. Reg.	196	289	1.213	9.358	498	1.316	417	38	13	1	13.339
Herb. Sec.	196	908	2.052	551	8.448	2.680	195	11	275	0	15.316
Herb. Reg.	41	109	232	768	1.586	8.023	406	37	14	0	11.216
Sup. Imp.	285	638	890	740	659	1.563	18.736	319	732	145	24.707
Lám. Agua	1	17	25	52	0	117	31	21.666	37	3	21.949
Suelo desn.	2	82	108	7	42	13	49	11	311	0	625
Invern.	0	0	0	0	0	0	106	2	0	364	625
Total	26.535	58.758	11.340	14.961	14.114	14.951	21.292	22.259	1.833	513	

TABLA 11

Matriz de confusión correspondiente a la clasificación del año 2019.

CLASES	Bosque	Matorr.	Arb. Sec.	Arb. Reg.	Herb. Sec.	Herb. Reg.	Sup. Imp.	Lám. Agua	Suelo desn.	Invern.	Total
Bosque	20.887	4.325	268	177	296	140	207	84	0	0	26.384
Matorral	4.628	21.443	997	462	958	562	425	72	214	0	29.761
Arb. Sec.	355	870	5.875	1.783	1.161	536	430	10	339	2	11.361
Arb. Reg.	137	509	1.056	10.181	667	1.348	418	26	12	0	14.354
Herb. Sec.	111	893	1.826	796	7.360	4.017	301	2	167	0	15.473
Herb. Reg.	110	201	354	736	2.276	6.971	188	81	26	8	10.951
Sup. Imp.	243	489	780	477	792	1.196	19.108	217	366	108	23.776
Lám. Agua	0	18	2	57	0	102	26	21.713	4	6	21.927
Suelo desn.	5	49	342	10	50	12	42	9	752	0	1.271
Invern.	0	0	0	0	0	0	147	14	10	362	533
Total	26.476	28.796	11.500	14.679	13.560	14.884	21.292	22.228	1.890	486	