



ESCUELA DE INGENIERÍA AGRARIA Y FORESTAL

TRABAJO FIN DE MÁSTER

**MÁSTER UNIVERSITARIO EN GEOINFORMÁTICA PARA LA GESTIÓN DE LOS RECURSOS
NATURALES**

**IDENTIFICACIÓN Y CLASIFICACIÓN DE HÁBITATS EUNIS MEDIANTE IMÁGENES
MULTIESPECTRALES EN LA RÍA DE ORTIGUEIRA (GALICIA)**

**EUNIS HABITAT IDENTIFICATION AND CLASSIFICATION BY MULTISPECTRAL IMAGERY IN
RÍA DE ORTIGUEIRA (GALICIA)**

Raquel Rodríguez Cereijo

Tutor: Eduardo Manuel González Ferreiro

Cotutor: Juan Guerra Hernández

Ponferrada, septiembre de 2022

RESUMEN

Contar con una cartografía de hábitats actualizada y precisa es un aspecto clave a la hora de analizar y gestionar espacios naturales, por ello, la teledetección tiene un papel fundamental. En este caso el análisis se ha centrado en una parte de la Ría de Ortigueira y Ladrado (A Coruña), incluida en diversos instrumentos de protección debido a su relevancia ecológica y a la fragilidad de este tipo de ecosistemas.

En este trabajo se plantea como objetivo principal realizar clasificaciones supervisadas empleando análisis orientado a objetos a partir de imágenes multiespectrales de dos sensores con diferentes características e información complementaria, en base a las categorías de hábitats establecidas en el sistema EUNIS (*European Nature Information System*). Posteriormente se analizarán los resultados obtenidos y las variables más representativas en cada caso.

Los pasos seguidos han sido los siguientes: descarga y procesado de las imágenes e información complementaria, segmentación de la ortofotografía del PNOA (Plan Nacional de Ortofotografía Aérea) para utilizarla como punto de partida en ambos casos, preparación de la base de datos formada por las diferentes variables y sus métricas, cálculo de estadísticas zonales para cada objeto, realización de clasificaciones supervisadas mediante algoritmos de aprendizaje automático y validación de los resultados empleando diferentes estadísticos. Estos procesos se han realizado empleando una combinación de diferentes softwares, tanto comerciales como libres.

Los resultados muestran mayor precisión global en la clasificación generada en base a las imágenes de WorldView 2 en comparación con la realizada con Sentinel 2. Además, en ambos casos se observa la importancia que presentan los índices de vegetación a la hora de caracterizar los diferentes hábitats.

Palabras clave: humedal, Sentinel 2, WorldView 2, OBIA, Random Forest, EUNIS

ÍNDICE

| | |
|--|----|
| Resumen..... | I |
| 1. Introducción | 1 |
| 2. Objetivos..... | 3 |
| 3. Material y métodos | 3 |
| 3. 1. Material | 4 |
| 3.1.1. Área de estudio | 4 |
| 3.1.2. Información espectral | 6 |
| 3.1.3. Datos LiDAR | 8 |
| 3.1.4. Áreas de entrenamiento y validación..... | 8 |
| 3.1.5. Software | 10 |
| 3.2. Métodos | 11 |
| 3.2.1. Definición de clases | 11 |
| 3.2.2. Procesado de imágenes..... | 12 |
| 3.2.3. Segmentación..... | 15 |
| 3.2.4. Cálculo de estadísticos zonales | 16 |
| 3.2.5. Clasificación..... | 16 |
| 3.2.6. Validación | 17 |
| 4. Resultados | 19 |
| 4.1. Clasificación Sentinel 2 | 19 |
| 4.2. Clasificación con imagen WorldView 2 | 21 |
| 5. Discusión..... | 23 |
| 6. Conclusiones y líneas de trabajo futuras..... | 28 |
| 7. Bibliografía..... | 29 |
| Anexo I. Descripción de los hábitats de la zona de estudio | 38 |
| Anexo II. Script de clasificación | 39 |
| Anexo III. RESULTADOS CLASIFICACIONES | 43 |
| A. Clasificación con Sentinel 2 | 43 |
| B. Clasificación con WorldView 2 | 44 |

Índice de tablas

| | |
|---|----|
| Tabla 1. Características de las bandas de Sentinel 2. Fuente | 6 |
| Tabla 2. Resumen de las características de las imágenes Sentinel 2 MSI utilizadas..... | 7 |
| Tabla 3. Características de las bandas de WorldView 2..... | 7 |
| Tabla 4. Resumen de las características de las imágenes WorldView 2 utilizadas. | 8 |
| Tabla 5. Áreas de entrenamiento y validación por cada tipo de hábitat. | 9 |
| Tabla 6. Hábitats de la zona de estudio. | 12 |
| Tabla 7. Índices espectrales utilizados. | 14 |
| Tabla 8. Códigos de las variables empleadas. | 16 |
| Tabla 9. Matriz de confusión para la clasificación con Sentinel 2 con todas las variables. | 20 |
| Tabla 10. Matriz de confusión para la clasificación con WorldView 2..... | 22 |

Índice de figuras

| | |
|---|----|
| Figura 1. Flujo de trabajo | 4 |
| Figura 2. Área de estudio | 5 |
| Figura 3. Distribución de las áreas de entrenamiento y puntos de validación..... | 9 |
| Figura 4. Resultado del proceso de segmentación de la ortofotografía..... | 15 |
| Figura 5. Importancia de las variables en el modelo de clasificación de Sentinel 2 en base a la reducción media en la precisión y la reducción media del índice de Gini. | 19 |
| Figura 6. Importancia de las variables en el modelo de clasificación de WorldView 2 en base a la reducción media en precisión y en la reducción media de Gini..... | 21 |

1. INTRODUCCIÓN

A lo largo de las últimas décadas se ha puesto de manifiesto la necesidad de protección que presentan los humedales, debido tanto a sus características intrínsecas como a la presión antrópica que soportan. Este hecho se refleja en el aumento de instrumentos y normativa a diferentes escalas para favorecer su conservación, como el Convenio Ramsar, nombre habitual que se da a la Convención Relativa a los humedales de Importancia Internacional especialmente como Hábitat de Aves Acuáticas (UNESCO, 1971), la Directiva Marco del Agua (Parlamento Europeo y Consejo de la Unión Europea, 2000) o la Ley de Patrimonio Natural y de la Biodiversidad (España, 2007), entre otras. A pesar de estas medidas, la tendencia actual es desfavorable, ya que únicamente el 12% de los sitios Ramsar españoles se encuentran bien conservados (MITECO, 2021). Entre los principales factores que afectan a estos espacios destacan las desecaciones, contaminación o la presión urbanística.

Por otro lado, contar con una nomenclatura estandarizada es un aspecto básico en la gestión de hábitats, especialmente a la hora de aplicar instrumentos internacionales en los diferentes estados. Por ello, en los años 90 el Centro Temático Europeo sobre Protección de la Naturaleza y Biodiversidad (ETC/NPB) creó para la Agencia Europea de Medio Ambiente la clasificación de hábitats EUNIS (Davies y Moss, 1999; Davies *et al.*, 2004). Está caracterizada por la organización de hábitats de forma jerárquica, abierta y por criterios objetivos, asignando a cada uno de ellos un código identificativo (Ramil-Rego *et al.*, 2008). Estos aspectos, junto con su compatibilidad con las unidades empleadas en la Red Natura 2000 (Martínez *et al.*, 2010), hace que haya sido propuesta por la Directiva INSPIRE (Parlamento Europeo y Consejo de la Unión Europea, 2007) como la clasificación de referencia para hábitats (Ichter *et al.*, 2014).

La cartografía de hábitats en la Unión Europea se ha visto favorecida, entre otros factores, por la Directiva Hábitats (Consejo de las Comunidades Europeas, 1992), ya que en ella se indica que en las zonas incluidas en la Red Natura 2000 se debe analizar su estado de conservación. Además, la Estrategia de Biodiversidad 2020 (Comisión Europea, 2011), incluye la cartografía de los ecosistemas y sus servicios como una de las medidas para cumplir con el objetivo de “Mantenimiento y restauración de los ecosistemas y sus servicios” (Ichter *et al.*, 2014). Es por ello que la teledetección juega un papel clave en la conservación de los hábitats, ya que permite conocer su estado actual, cambios sufridos y hacer estimaciones sobre su posible evolución; elementos fundamentales para definir las medidas de gestión y en su caso, evaluar la eficacia de las decisiones tomadas (Varela *et al.*, 2007; Álvarez-Martínez *et al.*, 2018). Otras

características por las que resulta útil su utilización son la capacidad de monitorizar grandes extensiones, contar con información actualizada de diversas fuentes y características o la posibilidad de realizar análisis históricos (Agrillo *et al.*, 2021). El uso de imágenes de satélite en cartografía de hábitats se ha popularizado, entre otras cosas, por la disponibilidad de datos de libre acceso, como los productos de la misión Sentinel 2 del programa Copernicus de la Unión Europea (Spoto *et al.*, 2012). Además, existen numerosos satélites comerciales, entre los que se encuentran los WorldView (Tarantino *et al.*, 2012), que permiten contar con imágenes de muy alta resolución espacial.

Entre las técnicas que se pueden realizar a partir de imágenes de satélite, se encuentra el análisis orientado a objetos u *Object Based Image Analysis* (OBIA) (Blaschke, 2010). Se trata de una metodología basada en objetos formados por grupos de píxeles homogéneos cuyas características se definen en el proceso de segmentación. Este paso consiste en dividir los datos de entrada en conjuntos de píxeles, dependiendo del enfoque que se utilice, se realizará uniendo píxeles en base a sus características (*bottom-up*) o fragmentando la imagen original en objetos de menor tamaño (*top-down*) (Mauro *et al.*, 2013).

Las clasificaciones se pueden realizar empleando diferentes algoritmos, actualmente están en auge los basados en aprendizaje automático. De entre ellos, *Random Forest* (RF) descrito por Breiman (2001), es uno de los más utilizados, por ejemplo, en clasificaciones de cobertura del suelo empleando imágenes de satélite multi e hiperespectrales (Pal, 2005; Lawrence *et al.*, 2006). Entre sus ventajas destaca la posibilidad de utilizar todas las muestras del terreno, especialmente importante en zonas de difícil acceso, permite cuantificar la importancia relativa de cada variable y además se trata de un método no paramétrico que permite el uso de grandes cantidades de datos y variables sin necesidad de que se ajusten a la normalidad (Rodríguez-Galiano *et al.*, 2012; Feng *et al.*, 2017).

La combinación de OBIA con imágenes de satélite se ha utilizado en varios estudios para analizar hábitats, a continuación se van a mencionar algunos ejemplos en los que se han empleado imágenes Sentinel 2: i) Agrillo *et al.* (2021) clasificaron hábitats forestales en Italia a distintos niveles EUNIS empleando una combinación de imágenes Sentinel 2, variables ambientales y RF, alcanzando unas precisiones globales de entre el 76,14 % y 87 %, ii) Le Dez *et al.* (2021) evaluaron la capacidad de las imágenes Sentinel 2 para elaborar cartografía de hábitats en el estuario del Loira, iii) Calleja *et al.* (2019) analizaron la evolución de la especie invasora *Baccharis halimifolia* L. en la costa de Cantabria empleando imágenes Sentinel 2A y

Landsat-8, iv) Whyte *et al.* (2018) combinaron imágenes Sentinel 1 y 2, técnicas OBIA y algoritmos RF y SVM (*Support Vector Machine*) para el mapeo de humedales en Sudáfrica. En el caso del sensor WorldView 2 también existen estudios: i) Tarantino *et al.* (2012) analizaron el comportamiento de las imágenes para caracterizar coberturas de uso del suelo como agua, suelo, vegetación y artificial, desagregadas en diferentes subcategorías, ii) Jackson y Adam (2021) analizaron la capacidad de las imágenes WorldView 2 para diferenciar especies de árboles en bosques tropicales y iii) Immitzer *et al.* (2012), en cuyo estudio evaluaron el potencial de estas imágenes para clasificar 10 especies de arbolado en Austria empleando RF.

Sin embargo, hasta la fecha no existen estudios específicos para el mapeado automático de hábitat en la Península Ibérica con el sistema de clasificación EUNIS usando imágenes de WorldView 2 y Sentinel 2. Además, no se han encontrado referencias previas sobre la influencia de la resolución de ambos sensores multispectrales en la precisión de las clasificaciones de imágenes mediante OBIA y la importancia de usar variables LiDAR para mejorar los resultados.

2. OBJETIVOS

El objetivo principal de este trabajo consiste en desarrollar una metodología para clasificar hábitats en humedales mediante OBIA con imágenes multispectrales de satélite, Sentinel MSI 2A/B y WorldView 2, en combinación con datos LiDAR-PNOA de Galicia, siguiendo el sistema de clasificación EUNIS.

Como objetivo secundario se pretende analizar si los sensores del programa Copernicus tienen capacidad para poder ser usados en clasificaciones en zonas de humedal con un nivel de detalle elevado.

3. MATERIAL Y MÉTODOS

Este apartado describe la metodología seguida para realizar las clasificaciones supervisadas en base a OBIA empleando imágenes de los sensores Sentinel 2 MSI y WorldView 2 WV110. En el flujo de procesos (Figura 1) se muestran los pasos realizados, detallando su orden, la información utilizada y los programas empleados.

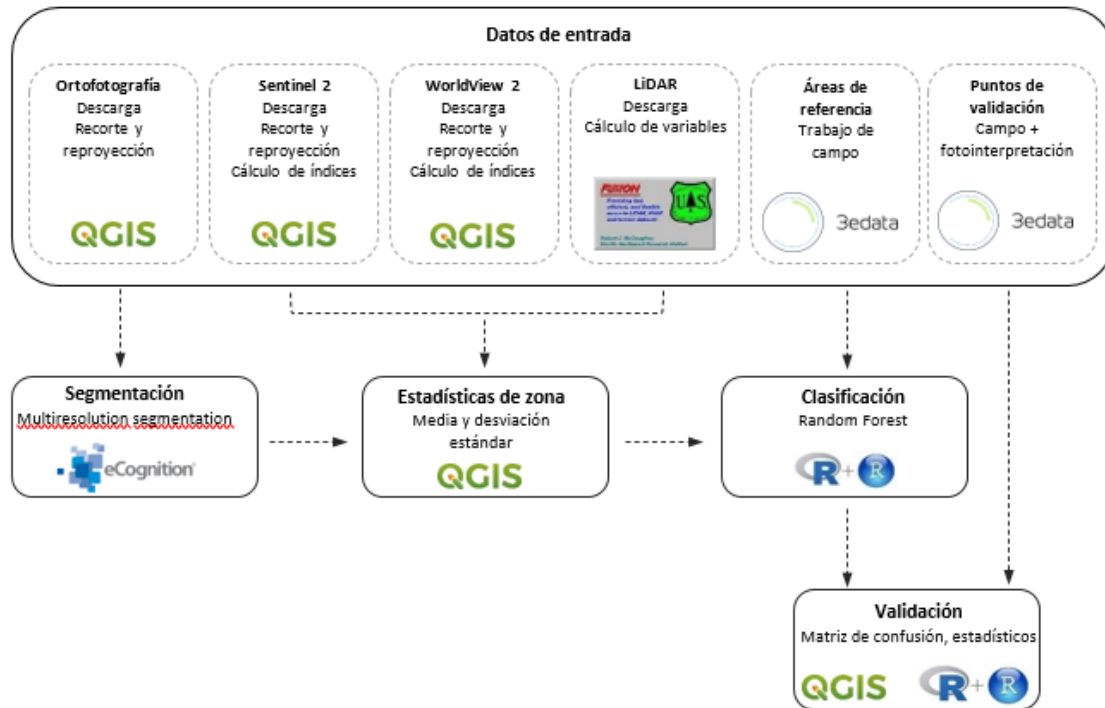


Figura 1. Flujo de trabajo

3. 1. Material

A continuación, se van a describir los datos empleados en el proceso y los programas informáticos con los que se ha realizado cada paso.

3.1.1. Área de estudio

El área analizada en este trabajo se trata de la Enseada de Mera (**¡Error! No se encuentra el origen de la referencia.**), ubicada en la zona suroeste de la Ría de Ortigueira y Ladrado, al norte de la provincia de A Coruña (Galicia). Cuenta con una superficie de aproximadamente 365 ha en las que se pueden encontrar una amplia variedad de ambientes que van desde masas arboladas hasta áreas permanentemente cubiertas por agua, entre ellos destacan las formaciones de *Zostera noltii* Hornemann y *Z. marina* L. ya que suponen prácticamente la mitad de la superficie vegetada, seguidas por las comunidades dominadas por *Spartina maritima* (Curtis) Fernald y los juncales, mayoritariamente halófilos (Izco y Sánchez, 1996).

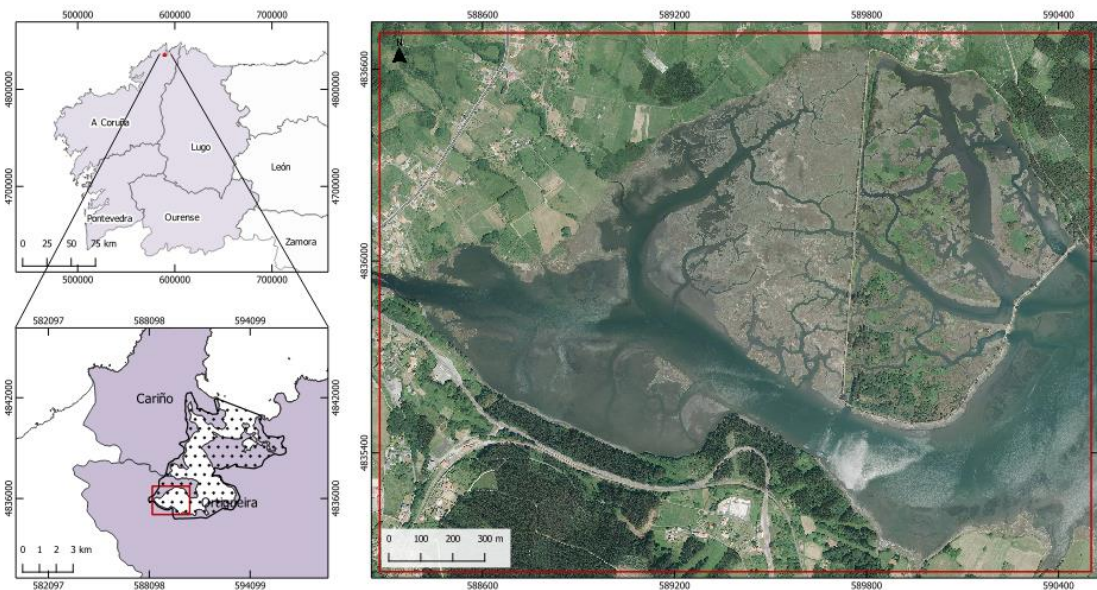


Figura 2. Área de estudio

Se trata de un área con relieves suaves formados por rocas ultrabásicas y esquistos en el margen izquierdo, debido al Complejo de Cabo Ortegal, y rocas metamórficas en el lado derecho, por el Anticlinal Ollo de Sapo (Ferreiro *et al.*, 2001). La caracterización bioclimática de la zona, según la clasificación de Rivas-Martínez (2004), indica que le corresponde el Macrobioclima templado, bioclima hiperoceánico subhiperoceánico, termotipo termotemplado superior y ombrotipo húmedo inferior (Rodríguez y Ramil-Rego, 2007). Esto implica que no haya sequía estival ni heladas en invierno y que las temperaturas medias mensuales tengan variaciones poco marcadas.

La elección de esta zona se debe a su importancia a nivel ecológico y la necesidad de su conservación, siendo protegida por numerosos instrumentos de diverso rango, entre los que destaca su incorporación en lista Ramsar de humedales de importancia internacional (UNESCO, 1971) en 1989, su declaración en 1990 como Zona de Especial Importancia para las Aves (ZEPA) y su nombramiento como Zona de Especial Conservación (ZEC) en 2014. Entre los factores que han llevado a su protección destacan la abundante presencia de aves limícolas, especialmente en invierno, además de la fragilidad intrínseca de los humedales. Otro aspecto que se ha tenido en cuenta a la hora de seleccionar el área de estudio ha sido la disponibilidad de información de campo, proporcionada por la empresa 3edata Ingeniería Ambiental S.L. en el ámbito del proyecto TERRA3, desarrollado en el marco de la *Civil UAVs Initiative* y financiado por la Agencia Gallega de Innovación de la Xunta de Galicia, de forma que ha permitido contar con datos de partida de elevada calidad.

3.1.2. Información espectral

La selección de la información espectral de partida se ha realizado en base a la disponibilidad de imágenes y a los objetivos del estudio. De este modo, se ha utilizado una ortofotografía de elevada resolución espacial del PNOA e imágenes satelitales procedentes de dos sensores, Sentinel 2 MSI y WorldView 2 WV110, con diferentes resoluciones espectrales y espaciales. Además, a partir de sus bandas se han calculado varios índices espectrales con la intención de facilitar la caracterización de los hábitats. En este caso, al tratarse de un entorno acuático con presencia de hábitats intermareales, se han utilizado imágenes en condiciones de marea alta y marea baja para facilitar su identificación. Además, se ha buscado que las imágenes coincidan temporalmente con el trabajo de campo realizado por la empresa 3edata Ingeniería Ambiental S.L.

3.1.2.1. Imágenes Sentinel 2

El proyecto Sentinel 2 está formado por dos satélites idénticos lanzados por la Agencia Espacial Europea (ESA) en los años 2015 y 2017. Sus imágenes están formadas por 13 bandas en el espectro visible e infrarrojo con una resolución espacial que varía entre los 10 y 60 m. En la **¡Error! No se encuentra el origen de la referencia.** se muestran sus características.

Tabla 1. Características de las bandas de Sentinel 2. Fuente:
<https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument>.

| Banda | Satélite 2A | | Satélite 2B | | Resolución espacial (m) | |
|-------|-----------------------------------|---------------------|-------------------------------|---------------------|-------------------------|----|
| | Longitud de onda central (nm) | Ancho de banda (nm) | Longitud de onda central (nm) | Ancho de banda (nm) | | |
| 1 | Azul costero | 442,7 | 21 | 442,3 | 21 | 60 |
| 2 | Azul | 492,4 | 66 | 492,1 | 66 | 10 |
| 3 | Verde | 559,8 | 36 | 559,0 | 36 | 10 |
| 4 | Rojo | 664,6 | 31 | 665,0 | 31 | 10 |
| 5 | Borde del rojo (Red Edge) | 704,1 | 15 | 703,8 | 16 | 20 |
| 6 | Red Edge | 740,5 | 15 | 739,1 | 15 | 20 |
| 7 | Red Edge | 782,8 | 20 | 779,7 | 20 | 20 |
| 8 | Infrarrojo Cercano (NIR) I | 832,8 | 106 | 833,0 | 106 | 10 |
| 8A | NIR II | 864,7 | 21 | 864,0 | 22 | 20 |
| 9 | Vapor de agua | 945,1 | 20 | 943,2 | 21 | 60 |
| 10 | Cirros | 1373,5 | 31 | 1376,9 | 30 | 60 |
| 11 | Infrarrojo de onda corta (SWIR) I | 1613,7 | 91 | 1610,4 | 94 | 20 |
| 12 | SWIR II | 2202,4 | 175 | 2185,7 | 185 | 20 |

La descarga de imágenes se realizó desde la plataforma *Copernicus Open Access Hub* (<https://scihub.copernicus.eu/>) filtrando por área, fecha y porcentaje de nubosidad, en este caso, inferior al 10 %. Las imágenes seleccionadas (Tabla 2) tienen un nivel de procesado 2A, por lo que se encuentran corregidas atmosféricamente y presentan valores de reflectancia en la zona inferior de la atmósfera o *Bottom Of Atmosphere* (BOA).

Tabla 2. Resumen de las características de las imágenes Sentinel 2 MSI utilizadas.

| Nombre | Marea | Fecha |
|--|-------|------------|
| S2A_MSIL2A_20190326T112111_N0211_R037_T29TN | Baja | 26/03/2019 |
| S2B_MSIL2A_20190712T113329_N0213_R080_T29TNJ | Alta | 12/07/2019 |

3.1.2.2. Imágenes WorldView 2

El satélite WorldView 2 fue puesto en órbita en el año 2009 por DigitalGlobe con la intención de proporcionar imágenes de muy alta resolución tanto espacial como temporal. Entre sus características destaca la incorporación de las bandas Azul Costero, Amarillo, Borde del rojo y NIR II, especialmente interesantes para ambientes acuáticos según las conclusiones del trabajo de Tarantino *et al.* (2012). En la Tabla 3 se detallan las propiedades de cada banda.

Tabla 3. Características de las bandas de WorldView 2. Fuente: Shamsoddini *et al.* (2013).

| Banda | Longitud de onda central (nm) | Ancho de banda (nm) | Resolución espacial (m) |
|-----------------------|-------------------------------|---------------------|-------------------------|
| Pancromática | 632 | | 0,50 |
| 1 Azul Costero | 427 | 47,3 | 1,84 |
| 2 Azul | 478 | 54,3 | 1,84 |
| 3 Verde | 546 | 63,3 | 1,84 |
| 4 Amarillo | 608 | 37,4 | 1,84 |
| 5 Rojo | 659 | 57,4 | 1,84 |
| 6 Red Edge | 724 | 39,3 | 1,84 |
| 7 NIR | 831 | 98,9 | 1,84 |
| 8 NIR II | 908 | 99,6 | 1,84 |

Las imágenes WorldView 2 se descargaron desde la plataforma web de DigitalGlobe (<https://discover.digitalglobe.com/>), donde, al igual que en el caso anterior, se seleccionaron imágenes con bajo porcentaje de nubosidad temporalmente similares al trabajo de campo (Tabla 4). Los productos fueron proporcionados por la empresa 3edata Ingeniería Ambiental S.L. y se recibieron preprocesados con corrección radiométrica y un refinado pancromático o *pansharpening* a 0,5 m de resolución espacial en base a la banda pancromática.

Tabla 4. Resumen de las características de las imágenes WorldView 2 utilizadas.

| Nombre | Marea | Fecha |
|-------------------------|-------|------------|
| 103001008F8B4700 | Baja | 22/03/2019 |
| 103005008EF94700 | Alta | 12/07/2019 |

3.1.2.3. Ortofotografía

Se ha descargado una ortofotografía del Plan Nacional de Ortofotografía Aérea (PNOA), accesible desde: <http://centrodedescargas.cnig.es/>, con la intención de contar con una imagen de elevada resolución espacial como base en el proceso de segmentación y así obtener un punto de partida común para las clasificaciones posteriores realizadas a partir de las imágenes de los dos sensores. De entre las opciones disponibles, se ha seleccionado una ortofotografía con resolución espacial de 25 cm del año 2017, ya que es la más cercana en el tiempo a los trabajos de campo.

3.1.3. Datos LiDAR

Los datos LiDAR (*Light Detection and Ranging*) aéreos se han empleado con el objetivo principal de ayudar a clasificar los hábitats en base a sus características estructurales. Para ello se ha partido de los archivos en formato laz del vuelo PNOA-LiDAR del año 2015-2017 (segunda cobertura), disponibles en: <https://centrodedescargas.cnig.es/>, cuya densidad es de 0,5 primeros retornos por metro cuadrado. Estos puntos se encuentran clasificados automáticamente y coloreados en RGB en base a ortofotografías.

3.1.4. Áreas de entrenamiento y validación

Las áreas de entrenamiento han sido proporcionadas por la empresa 3edata Ingeniería Ambiental S.L. (Tabla 5), que realizó la campaña de trabajo de campo durante el verano de 2019. Como resultado definieron 173 áreas de entrenamiento en las que indicaron la ubicación y el tipo de hábitat correspondiente. Por otro lado, con la intención de obtener una muestra independiente para validar la clasificación, generaron 387 puntos distribuidos por la zona de estudio mediante muestreo aleatorio. Estos puntos fueron analizados por especialistas en hábitats mediante fotointerpretación y visitas de campo, definiendo su localización y categoría correspondiente. En la Figura 3 se representa su distribución por la zona de estudio.

Tabla 5. Áreas de entrenamiento y validación por cada tipo de hábitat.

| Código trabajo | Descripción | Objetos para entrenamiento | Puntos de validación |
|----------------|---|----------------------------|----------------------|
| 102400 | Sedimentos litorales mixtos (llanuras intermareales no vegetadas) | 18 | 165 |
| 102527 | Matorrales salinos atlánticos | 26 | 40 |
| 102531 | Comunidades atlánticas del supraestero | 33 | 77 |
| 102554 | Praderas de <i>Spartina maritima</i> | 18 | 21 |
| 102600 | Sedimentos litorales dominados por angiospermas acuáticas | 24 | 36 |
| 105420 | Sedimentos sublitorales mixtos con salinidad variable (estuarios) | 6 | 20 |
| 105310 | Helechales subatlánticos | 20 | 6 |
| 161214 | Alisedas del norte de la Península Ibérica | 19 | 14 |
| 162810 | Plantaciones de eucaliptos (<i>Eucalyptus</i> spp.) | 9 | 8 |
| | Total | 173 | 387 |



Figura 3. Distribución de las áreas de entrenamiento y puntos de validación.

3.1.5. Software

3.1.5.1. eCognition

El software eCognition Developer 9.0 (Trimble Inc., 2014) está principalmente enfocado al análisis de imágenes orientado a objetos. En este caso se ha empleado como herramienta para realizar la segmentación de la ortofotografía, mediante el algoritmo *Multiresolution Segmentation* propuesto por Baatz y Schäpe (2000). Este proceso busca minimizar la heterogeneidad media de los objetos (Trimble Inc., 2014) basándose en los parámetros de escala, forma y compacidad. La escala determina la variabilidad máxima permitida en los objetos, condicionando su tamaño; por otro lado, la forma y compacidad constituyen el criterio de homogeneidad: en el primer caso afectará al peso que tendrá el componente espacial respecto al espectral, y en el segundo, a lo compactos que serán los objetos resultantes.

3.1.5.2. QGIS

QGIS versión 3.16 (QGIS, 2021) es un sistema de información geográfica (SIG) libre y de código abierto. Este programa se ha empleado para el preprocesado de la información de entrada, que incluye tanto el recorte de las bandas en base a la zona de estudio como su reproyección para homogeneizar el sistema de referencia de coordenadas. Además, se ha utilizado para calcular las estadísticas zonales de los objetos, realizar el análisis visual de las áreas de entrenamiento y validación, asociar los resultados de las clasificaciones a los puntos de referencia y, por último, para generar la cartografía mostrada en el documento.

3.1.5.3. R/Rstudio

RStudio se trata de un entorno de desarrollo integrado para el lenguaje de programación R, que entre otras cosas, permite realizar análisis estadísticos y de datos espaciales. En este caso se ha utilizado la versión 4.0.3 de R (R Core Team, 2020) y la 1.1.456 de RStudio (RStudio Team, 2018). Se ha empleado para realizar la clasificación supervisada de hábitats mediante el algoritmo RF, disponible en el paquete *randomForest* (Liaw y Wiener, 2002). Tras ello se ha realizado la validación de los resultados mediante funciones del paquete *caret* (Khun, 2020), con las que se pueden obtener matrices de confusión y los estadísticos para cada clasificación.

3.1.5.4. FUSION

FUSION es un software gratuito especializado en el procesado de datos LIDAR, en este caso se ha utilizado la versión 4.40 (USDA, 2022). La información se puede visualizar y procesar desde su interfaz o a partir de líneas de comando, indicando el algoritmo a aplicar y sus parámetros. En este caso se ha optado por la segunda opción, utilizando los ejecutables *GroundFilter*, *GridSurfaceCreate* y *CanopyModel* para el procesado de los datos LIDAR del PNOA.

3.2. Métodos

En este apartado se describe la secuencia seguida para realizar las clasificaciones con las imágenes de ambos sensores en base a lo mostrado en la Figura 1. De forma general, la metodología propuesta consiste en cuatro puntos principales: procesado de las imágenes, segmentación, clasificación y validación. A pesar de que se han utilizado imágenes de distintos sensores, los pasos seguidos han sido los mismos para ambos casos.

3.2.1. Definición de clases

Uno de los elementos clave a la hora de realizar una clasificación es definir las categorías que se pretenden diferenciar, en este caso, la bibliografía disponible sobre la Ría de Ortigueira y Ladrado, como las obras de Izco y Sánchez (1996) o Ramil-Rego *et al.* (2008), junto con la información proporcionada en el trabajo de campo, han llevado a considerar que en la zona de estudio se pueden encontrar nueve hábitats, descritos y organizados a continuación siguiendo el sistema EUNIS (Tabla 6).

Tabla 6. Hábitats de la zona de estudio.

| Nivel | Código EUNIS | Código trabajo | Hábitat | Descripción |
|--|--------------|----------------|---|---|
| Nivel A Hábitats marinos | A2.4 | 102400 | Sedimentos litorales mixtos | Llanuras intermareales no vegetadas |
| | A2.527 | 102527 | Matorrales salinos atlánticos | Formaciones de arbustos y plantas perennes crasas que se desarrollan sobre suelos húmedos salinos |
| | A2.531 | 102531 | Comunidades atlánticas del supraestero | Formaciones herbáceas no pioneras dominadas por los juncales |
| | A2.554 | 102554 | Praderas de <i>Spartina maritima</i> | Praderas dominadas por gramíneas del género <i>Spartina</i> que se encuentran en zonas intermareales con acumulación de sedimentos y aguas tranquilas |
| | A2.6 | 102600 | Sedimentos litorales dominados por angiospermas acuáticas | Llanuras intermareales en las que abundan formaciones de <i>Zostera marina</i> y <i>Zostera noltii</i> |
| | A5.42 | 105420 | Sedimentos sublitorales mixtos con salinidad variable | Hábitats con sustrato no rocoso desde la bajamar hasta 200 m de profundidad |
| Nivel E Pastizales y zonas dominadas por herbáceas, musgos o líquenes | E5.31 | 105310 | Helechales subatlánticos | Helechales de <i>Pteridium aquilinum</i> (L.) Kuhn. |
| Nivel G Bosques | G1.214 | 161214 | Alisedas del norte de la Península Ibérica | Bosques de ribera formados por alisos (<i>Alnus glutinosa</i> (L.) Gaertn.), fresnos (<i>Fraxinus</i> spp.) y otras especies de galería |
| | G2.81 | | Plantaciones de eucaliptos (<i>Eucalyptus</i> spp.) | Masas de eucaliptos |

En el Anexo I se incluye una descripción más completa de cada tipo de hábitat y su equivalencia entre el sistema EUNIS, la Lista Patrón Española de los Hábitats Terrestres, LPETH (España, 2020) y el Anexo I de la Directiva 92/43/CEE o Directiva Hábitats (Consejo de las Comunidades Europeas, 1992). Cabe destacar la clase G1.214 “Alisedas del norte de la Península Ibérica” ya que tiene correspondencia con el hábitat 91E0 “Bosques aluviales de *Alnus glutinosa* y *Fraxinus excelsior* (*Alno-Padion*, *Alnion incanae*, *Salicion albae*)” catalogado como prioritario en el Anexo I de la Directiva 92/43/CEE.

3.2.2. Procesado de imágenes

Una vez descargada la información de partida, se ha unificado el sistema de referencia de coordenadas a ETRS89 UTM Zona 29N (EPSG 25829) mediante la herramienta “Combar (reproyectar)” para los archivos en formato ráster y “Reproyectar capa” para los vectoriales (QGIS, 2021). Tras ello y con el objetivo de reducir el tiempo de procesado, se han recortado las capas para ajustarlas a la zona de estudio utilizando “Cortar ráster por capa de máscara” (QGIS,

2021). Por último, para homogeneizar el tamaño de píxel de las bandas de Sentinel 2 a 10 m, se ha utilizado la herramienta “Resample” de QGIS sobre las bandas cuya resolución espacial es a 20 m, en este caso son: banda 5, banda 6, banda 7, banda 11 y banda 12. El método escogido ha sido vecino más cercano o *Nearest Neighbour*.

3.2.2.1. Índices espectrales

A partir de las imágenes de ambos sensores, se han calculado cinco índices espectrales. El primero de ellos es el NDVI (*Normalized Difference Vegetation Index*), se trata de un índice ampliamente empleado para estimar el vigor de la vegetación en base a la intensidad de la radiación en las bandas del rojo y NIR (Escribano *et al.*, 2014). Su rango de valores se encuentra entre el -1 y 1, correspondiendo los valores positivos a zonas con vegetación y los cercanos a cero a las áreas desprovistas de ella. Hay que tener en cuenta que es sensible a la reflectividad del suelo, por lo que en casos en los que hay baja densidad de vegetación podrían existir problemas a la hora de detectarla (Muñoz, 2013).

$$NDVI = \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R}$$

El GNDVI (*Green Normalized Difference Vegetation Index*) es una versión modificada del NDVI que permite detectar con más detalle la actividad fotosintética de la vegetación al emplear la banda del verde.

$$GNDVI = \frac{\rho_{NIR} - \rho_G}{\rho_{NIR} + \rho_G}$$

El índice RENDVI (*Red Edge Normalized Difference Vegetation Index*) es otra modificación del NDVI que, en este caso, aprovecha la sensibilidad de la vegetación en el Red Edge.

$$RENDVI = \frac{\rho_{NIR} - \rho_{RedEdge}}{\rho_{NIR} + \rho_{RedEdge}}$$

El NDWI (*Normalized Difference Water Index*) evalúa la cantidad de agua de la vegetación o saturación del suelo. En este caso se ha optado por el método de McFeeters (1996) para resaltar las masas de agua.

$$NDWI = \frac{\rho_G - \rho_{NIR}}{\rho_G + \rho_{NIR}}$$

Por último, el índice OSAVI (*Optimized Soil Adjusted Vegetation Index*) tiene en cuenta la reflectancia del cultivo y del suelo, reduciendo su efecto con el parámetro Y , importante cuando la cobertura de vegetación no es total (Redel y Zubillaga, 2014).

$$OSAVI = 1 + Y \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R + Y}$$

Las bandas de cada sensor utilizadas para el cálculo de los índices espectrales se pueden consultar en la Tabla 7.

Tabla 7. Índices espectrales utilizados.

| Índice | Fórmula para Sentinel 2 | Fórmula para WorldView 2 | Referencia |
|---------------|---|---|-------------------------------|
| NDVI | $NDVI = \frac{B8 - B4}{B8 + B4}$ | $NDVI = \frac{B7 - B5}{B7 + B5}$ | Rouse <i>et al.</i> (1974) |
| NDWI | $NDWI = \frac{B3 - B8}{B3 + B8}$ | $NDWI = \frac{B3 - B7}{B3 + B7}$ | McFeeters (1996) |
| GNDVI | $GNDVI = \frac{B8 - B3}{B8 + B3}$ | $GNDVI = \frac{B7 - B3}{B7 + B3}$ | Gitelson <i>et al.</i> (1996) |
| RENDVI | $RENDVI = \frac{B8 - B6}{B8 + B6}$ | $RENDVI = \frac{B7 - B6}{B7 + B6}$ | Gitelson y Merzlyak (1994) |
| OSAVI | $OSAVI = 1 + 0,16 \frac{B8 - B4}{B8 + B4 + 0,16}$ | $OSAVI = 1 + 0,16 \frac{B7 - B5}{B7 + B5 + 0,16}$ | Rondeaux <i>et al.</i> (1996) |

El cálculo de los índices espectrales se ha realizado con la herramienta “Calculadora Ráster” de QGIS (QGIS, 2021), en la que se han seleccionado las bandas necesarias en base a la fórmula de cada índice.

3.2.2.2 Datos LIDAR

Los datos LIDAR del PNOA se han procesado para generar varios productos derivados a una resolución espacial de 2 m. Las variables calculadas han sido el Modelo Digital de Elevaciones o *Digital Elevation Model* (DEM), el Modelo Digital de Superficie o *Digital Surface Model* (DSM) y el Modelo Digital de Copas con alturas normalizadas respecto al terreno o *Canopy Height Model* (CHM). Para generar el DEM se han seleccionado los puntos del terreno de la nube de puntos de entrada mediante el ejecutable *GroundFilter* y la opción o *switch surface*, que permite interpolar los puntos clasificados como suelo. El resultado obtenido se ha utilizado como entrada en el algoritmo *GridSurfaceCreate*, cuyo objetivo es generar un modelo continuo. Como opciones se ha incluido el *switch minimum* para que lo realice a partir de los puntos con

valor más bajo. Tras ello, para generar el DSM se ha utilizado el algoritmo *CanopyModel*, en este caso interpola los retornos más altos para representar la superficie de copas. Por último, para calcular el CHM se ha empleado de nuevo *CanopyModel*, pero en este caso se ha incluido el DEM generado en el primer paso para normalizar las alturas.

3.2.3. Segmentación

El proceso de segmentación se ha realizado empleando el algoritmo *Multiresolution Segmentation* disponible en el software eCognition (Trimble Inc., 2014). Se trata de una segmentación *bottom-up*, es decir, comienza con píxeles semilla que va uniendo a otros en base a su homogeneidad. El objetivo es formar objetos uniformes y lo más diferenciados de su entorno posible (Trimble Inc., 2014).

A la hora de realizar la segmentación, se han probado diferentes combinaciones de los parámetros de escala, forma, compacidad y peso de las bandas de la imagen, con la intención de encontrar la que mejor se adapta tanto a la zona como a los objetivos del estudio. Tras realizar un análisis visual de los resultados de cada combinación, se ha optado por asignar el mismo peso a las bandas rojo, verde y azul de la ortofotografía del PNOA y dar unos valores de 100 para la escala, 0,3 para la forma y 0,5 para la compacidad, obteniendo como resultado una imagen formada por 9.492 objetos (Figura 4).

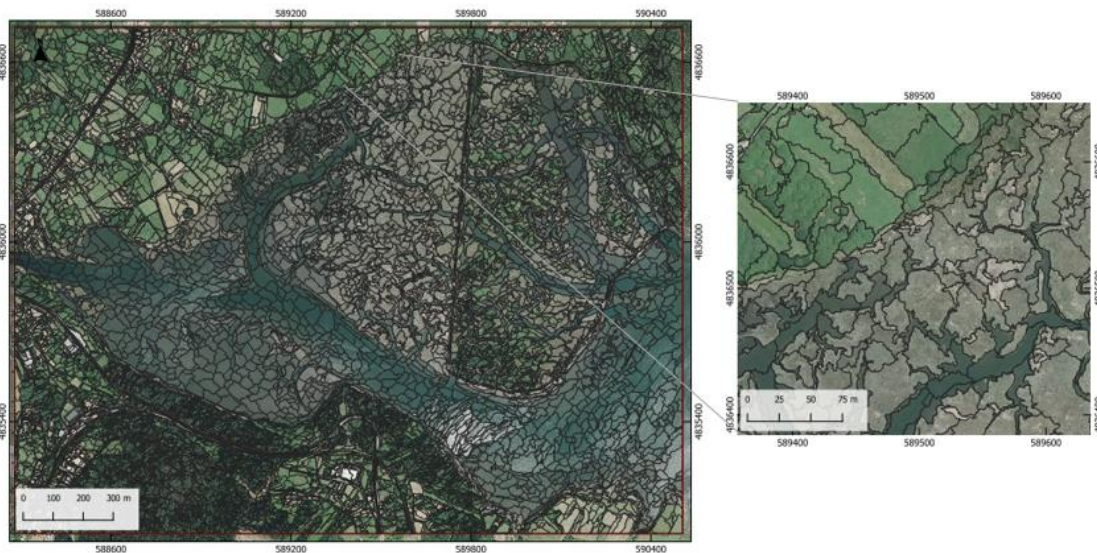


Figura 4. Resultado del proceso de segmentación de la ortofotografía.

3.2.4. Cálculo de estadísticas zonales

Tomando como referencia los objetos generados en la segmentación, se ha realizado el cálculo de la media y desviación estándar de las variables mediante la herramienta “Estadísticas de zona” disponible en QGIS (QGIS, 2021). La finalidad de este proceso es que cada objeto tenga asignado el valor que le corresponda de las distintas variables y así poder utilizarlos en la clasificación. Como resultado se ha generado una base de datos formada por 114 variables que permitirán caracterizar cada objeto y categoría. En la Tabla 8 se muestra cómo se ha codificado cada variable para facilitar su identificación en pasos posteriores.

Tabla 8. Códigos de las variables empleadas.

| Variable | Código |
|----------------------------|--|
| Bandas | S2b2a_m, S2b2a_d, S2b2b_m, S2b2b_d, S2b3a_m, S2b3a_d, S2b3b_m, S2b3b_d, S2b4a_m, S2b4a_d, S2b4b_m, S2b4b_d, S2b5a_m, S2b5a_d, S2b5b_m, S2b5b_d, S2b6a_m, S2b6a_d, S2b6b_m, S2b6b_d, S2b7a_m, S2b7a_d, S2b7b_m, S2b7b_d, S2b8a_m, S2b8a_d, S2b8b_m, S2b8b_d, S2b11a_m, S2b11a_d, S2b11b_m, S2b11b_d, S2b12a_m, S2b12a_d, S2b12b_m, S2b12b_d, WWb1a_m, WVb1a_d, WWb1b_m, WVb1b_d, WWb2a_m, WVb2a_d, WWb2b_m, WVb2b_d, WWb3a_m, WVb3a_d, WWb3b_m, WVb3b_d, WWb4a_m, WVb4a_d, WWb4b_m, WVb4b_d, WWb5a_m, WVb5a_d, WWb5b_m, WVb5b_d, WWb6a_m, WVb6a_d, WWb6b_m, WVb6b_d, WWb7a_m, WVb7a_d, WWb7b_m, WVb7b_d, WWb8a_m, WVb8a_d, WWb8b_m, WVb8b_d |
| Índices espectrales | S2NDVla_m, S2NDVla_d, S2NDVlb_m, S2NDVlb_d, S2NDWla_m, S2NDWla_d, S2NDWlb_m, S2NDWlb_d, S2GDVla_m, S2GDVla_d, S2GDVlb_m, S2GDVlb_d, S2OSAVla_m, S2OSAVla_d, S2OSAVlb_m, S2OSAVlb_d, S2RENDVla_m, S2RENDVla_d, S2RENDVlb_m, S2RENDVlb_d, WVNDVla_m, WVNDVla_d, WVNDVlb_m, WVNDVlb_d, WVNDWla_m, WVNDWla_d, WVNDWlb_m, WVNDWlb_d, WWGDVla_m, WWGDVla_d, WWGDVlb_m, WWGDVlb_d, WVOSAVla_m, WVOSAVla_d, WVOSAVlb_m, WVOSAVlb_d, WVRENDVla_m, WVRENDVla_d, WVRENDVlb_m, WVRENDVlb_d |
| LiDAR | chm_m, chm_d, dem_m, dem_d, dsm_m, dsm_d |

Abreviaturas utilizadas: S2 (imágenes de Sentinel 2), WV (imágenes de WorldView 2), b2, b3, b4, b5, b6, b7, b8, b11 y b12 (banda), a (marea alta), b (marea baja), m (media) y d (desviación estándar).

3.2.5. Clasificación

El primer paso para realizar la clasificación supervisada ha sido la preparación del conjunto de datos de entrenamiento, para ello se han unido las áreas de entrenamiento con los objetos generados en la segmentación en base a su localización. El resultado obtenido se ha revisado para asegurar que cada área de referencia se encontraba en un único objeto y que este fuera lo más homogéneo posible. De esta forma, se ha obtenido un conjunto de datos formado por 173 objetos caracterizados con la categoría de verdad terreno correspondiente y las variables calculadas en el apartado anterior. Los datos se han exportado en formato csv para

continuar con el proceso de clasificación en RStudio (RStudio Team, 2018). Una vez cargados los datos de referencia, se han generado dos subconjuntos para trabajar por separado con las variables de Sentinel 2 y las de WorldView 2, incluyendo las métricas LiDAR en ambos casos. La clasificación se ha realizado empleando distintas funciones de los paquetes *randomForest* (Liaw y Wiener, 2002) y *caret* (Khun, 2020). En primer lugar, se ha generado un modelo a partir de los datos de entrenamiento mediante la función *randomForest*, indicando 500 para el número de árboles o *n**tree* y dejando el valor por defecto, raíz cuadrada del número de variables, para los predictores empleados en cada nodo o *m**try*. Tras ello, se analizaron los resultados obtenidos y la importancia de las variables, realizando varias combinaciones con el objetivo de optimizar las clasificaciones, de esta forma se han realizado modelos que incluyen todas las variables y sólo con las variables en base al decrecimiento de Gini. Por último, los modelos generados se aplicaron al conjunto de la población y se exportaron para realizar el proceso de validación. El script completo se puede ver en el Anexo II.

3.2.6. Validación

Para evaluar la calidad de los resultados obtenidos, en primer lugar, se utilizó la herramienta “Unir atributos por localización” (QGIS, 2021). Como capa base se empleó la clasificación generada en el paso anterior y como capa de unión, los puntos de validación; de esta forma cada punto cuenta con la categoría de referencia y la generada en la clasificación. Este proceso se repitió con cada clasificación generada. Los resultados se exportaron en formato csv para continuar el proceso en RStudio (RStudio Team, 2018). Una vez cargados el fichero con los puntos de validación y los resultados de las clasificaciones en RStudio (RStudio Team, 2018), se ha generado una matriz de confusión mediante la función *confusionMatrix* disponible en el paquete *caret* (Khun, 2020). En ella se comparan los resultados obtenidos o predichos por el modelo con los esperados o de referencia. Además de la matriz se calculan diferentes estadísticos, el primero es la exactitud global, mediante la que se relaciona el número de muestras correctamente clasificadas respecto al total, expresada en este caso con un intervalo de confianza del 95% según el método Exacto o de Clopper-Pearson.

Otra métrica que se obtiene con la función *confusionMatrix* es el coeficiente Kappa (Cohen, 1960). Este valor indica la diferencia que existe entre la clasificación actual y los resultados que se obtendrían al realizarla por azar (Cohen, 1960; Rosenfield y Fitzpatrick-Lins, 1986). Su fórmula es:

$$K = \frac{n \sum_{i=1}^k X_{ii} - \sum_{i=1}^k X_{i+} X_{+i}}{n^2 - \sum_{i=1}^k X_{i+} X_{+i}}$$

Donde n es el tamaño de la muestra, X_{ii} el acuerdo observado, X_{i+} la suma de valores de la fila, X_{+i} la suma de valores de la columna y $X_{i+}X_{+i}$ el producto de los marginales de cada categoría.

A partir de los valores obtenidos en la matriz de confusión se pueden calcular otras variables como la exactitud del usuario (EU), con la que para cada clase se relaciona el número de elementos clasificados correctamente respecto del total. Se complementa con el error de comisión (EC) que indica la probabilidad de incluir una clase en una que no le corresponda. Las fórmulas son:

$$EU = \frac{X_{ij}}{n_{+i}} 100 \quad EC = 100 - EU$$

Donde X_{ij} es el valor de muestras correctamente clasificadas para una clase y n_{+i} el número total de elementos de una clase en la clasificación.

También se puede calcular la exactitud del productor (EP), con la que se refleja la relación de las muestras de referencia correctamente identificadas frente al total de muestras de referencia de esa clase. La métrica complementaria en este caso es el error de omisión (EO), mostrando la probabilidad de no incluir un objeto en la clase que le correspondería.

$$EP = \frac{X_{ij}}{n_{+j}} 100 \quad EO = 100 - EP$$

Donde X_{ij} es el valor de muestras correctamente identificadas para la una clase y n_{+j} el número total de elementos en la clase de referencia.

Con estas cuatro métricas es posible identificar las categorías que se han definido correctamente y aquellas en las que existen problemas de caracterización. Por último, se han utilizado dos métricas para evaluar la importancia de las variables incluidas en el modelo. La primera es la disminución media de la precisión o *Mean Decrease Accuracy* (MDA), mediante la que se refleja la reducción en la precisión del modelo al no tener en cuenta cada una de las variables, por ello, a mayor MDA, más importancia tendrá la variable. La segunda métrica es la disminución media del índice de Gini o *Mean Decrease Gini* (MDG), calculada en base al índice de Gini (Breiman *et al.*, 1984). El MDG indica la contribución de una variable en la formación de los nodos del árbol de clasificación, de forma que aquellas con valor más elevado afectarán más al modelo.

4. RESULTADOS

4.1. Clasificación Sentinel 2

La clasificación obtenida empleando las 30 variables procedentes de las imágenes Sentinel 2 y LIDAR con mayor valor en el índice de decrecimiento medio de Gini, ha tenido una exactitud global del 55,56 %, con un intervalo de confianza del 95 % de entre el 50,45 y 60,58 %. El coeficiente Kappa es de 0,4402 reflejando que la exactitud de la clasificación es moderada. En la Tabla 9 se muestra la matriz de confusión con la predicción sobre los puntos de validación.

Las clases en las que se han detectado mayores errores de comisión son la clase estuario (105420) y matorral salino (102527), ambas con un error de comisión superior al 60 %. En cuanto al error de omisión, las clases que muestran mayores problemas son la intermareal vegetado (102600) y los eucaliptales (162810), en el segundo caso con un error del 100%.

En la Figura 5 se puede ver la importancia de las variables en función de su valor respecto al MDA y MDGI, en este caso los valores más altos los presentan la media de los índices de vegetación, destacando el GNDVI en condiciones de marea alta. Otras variables con elevada relevancia son la media de las bandas 4, 11 y 12 con ambas mareas, seguidas por los índices en marea baja, las variables LIDAR y la media del resto de bandas.

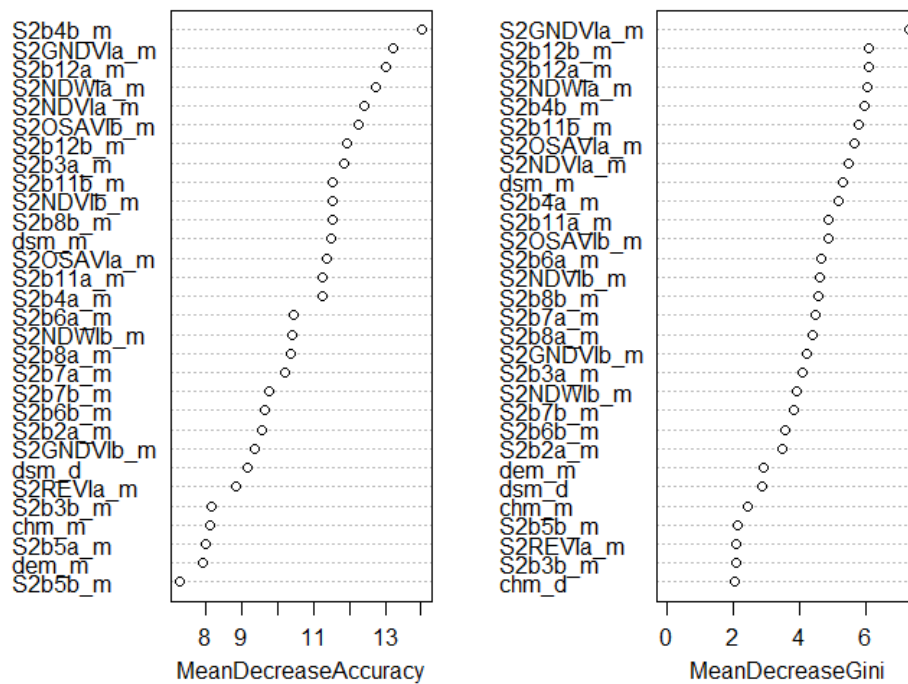


Figura 5. Importancia de las variables en el modelo de clasificación de Sentinel 2 en base a la reducción media en la precisión y la reducción media del índice de Gini.

Tabla 9. Matriz de confusión para la clasificación con Sentinel 2 con todas las variables.

| | Referencia | | | | | | | | | Total | EU (%) | EC (%) | |
|---------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|-------|--------|------------------------|------------------------|
| | 102400 | 102527 | 102531 | 102554 | 102600 | 105420 | 145310 | 161214 | 161810 | | | | |
| Clasificación | 102400 | 89 | 2 | 0 | 4 | 21 | 6 | 0 | 0 | 0 | 122 | 72,95 (64,43-80,07) | 27,05 (19,93-35,57) |
| | 102527 | 22 | 22 | 25 | 1 | 1 | 0 | 0 | 0 | 2 | 73 | 30,14 (20,78-41,48) | 69,86 (58,52-79,22) |
| | 102531 | 3 | 16 | 45 | 0 | 0 | 0 | 0 | 1 | 2 | 67 | 67,16 (55,22-77,25) | 32,84 (22,75-44,78) |
| | 102554 | 14 | 0 | 0 | 15 | 1 | 0 | 0 | 0 | 0 | 30 | 50,00 (32,06-67,94) | 50,00 (32,06-67,94) |
| | 102600 | 8 | 0 | 0 | 1 | 12 | 0 | 0 | 0 | 0 | 21 | 57,14 (36,52-75,56) | 42,86 (24,44-63,48) |
| | 105420 | 29 | 0 | 0 | 0 | 1 | 14 | 0 | 0 | 0 | 44 | 31,82 (19,93-46,63) | 68,18 (53,37-80,07) |
| | 145310 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 5 | 100,00 (59,90-100) | 0,00 (0,00-40,10) |
| | 161214 | 0 | 0 | 7 | 0 | 0 | 0 | 1 | 13 | 4 | 25 | 52,00 (36,79-74,39) | 48,00 (25,61-63,21) |
| | 161810 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - |
| | Total | 165 | 40 | 77 | 21 | 36 | 20 | 6 | 14 | 8 | 387 | | |
| EP (%) | 53,94 (46,33-61,37) | 55,00 (39,82-69,30) | 58,44 (47,28-68,80) | 71,43 (49,79-86,44) | 33,33 (20,14-49,75) | 70,00 (47,87-85,68) | 83,33 (41,78-98,86) | 92,86 (66,46-99,99) | 0,00 (0,00-29,34) | | | | |
| EO (%) | 46,06 (38,63-53,67) | 45,00 (30,70-60,18) | 41,56 (31,20-52,72) | 28,57 (13,56-50,21) | 66,67 (50,25-79,86) | 30,00 (14,32-52,13) | 16,67 (1,14-58,22) | 7,14 (<0,01-33,54) | 100,0 (70,66-100,0) | | | | |

Códigos de hábitats: 102400 Intermareal no vegetado, 102527 Matorral salino, 102531 Junciales, 102554 Spartina, 102600, Intermareal vegetado, 105420 Estuario, 145310 Helechales, 161214 Alisedas y 162810 Eucaliptales. EU: exactitud del usuario con intervalo de confianza de Wald ajustado al 95 % de probabilidad, EC: error de comisión con intervalo de confianza de Wald ajustado al 95 % de probabilidad, EP: exactitud del productor con intervalo de confianza de Wald ajustado al 95 % de probabilidad, EO: error de omisión con intervalo de confianza de Wald ajustado al 95 % de probabilidad.

4.2. Clasificación con imagen WorldView 2

En la clasificación en base a las 30 variables de WV2 y LIDAR con mayor valor en el índice de decrecimiento medio de Gini se ha obtenido una precisión global del 65,37 % con un intervalo de confianza del 95 % entre el 60,40 y 70,11 %. El índice Kappa para este caso es de 0,5588 por lo que al igual que en el caso anterior, sigue siendo moderado.

Las clases con mayor error de comisión son el matorral salino (102527) y la categoría Spartina (102554), siendo en ambos casos mayor del 50%, el resto de las clases presentan una exactitud de usuario mayor del 70%. Por otro lado, la clase intermareal vegetado (102600) tiene un error de omisión del del 58,33 %, siendo el más alto de toda la clasificación. En la Tabla 10 se muestran los resultados obtenidos en la matriz de confusión.

Del mismo modo que en la clasificación con Sentinel 2, las variables con mayor importancia para el modelo según el MDA y MDG son las medias de índices de vegetación en condiciones de marea alta, destacando en este caso el NDVI (Figura 6). Otras variables con elevado peso en la clasificación son la media de las bandas 7 y 8 con marea alta, aunque seguidas de cerca por la media del resto de bandas con diferente marea y las variables LIDAR.

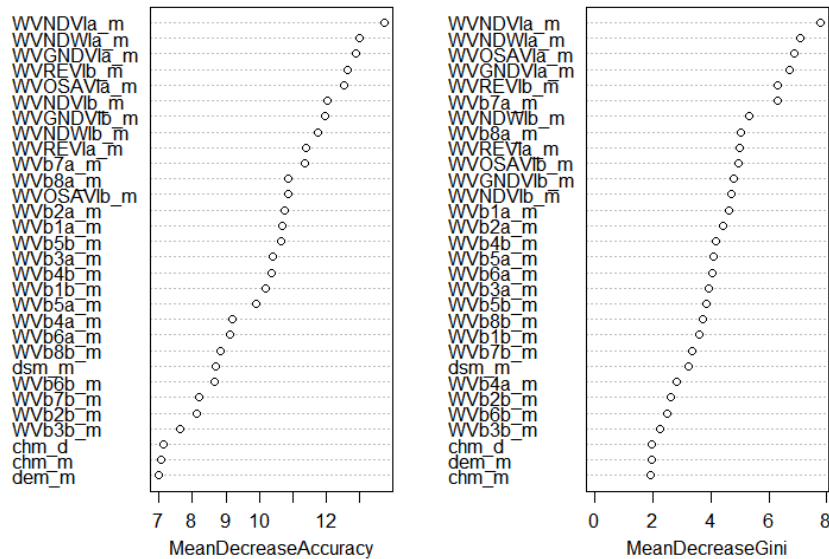


Figura 6. Importancia de las variables en el modelo de clasificación de WorldView 2 en base a la reducción media en precisión y en la reducción media de Gini.

En el Anexo III se pueden ver los resultados de ambas clasificaciones sobre la zona de estudio.

Tabla 10. Matriz de confusión para la clasificación con WorldView 2.

| | Referencia | | | | | | | | | Total | EU (%) | EC (%) | |
|---------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------------------|-------------------------|------------------------|-------|--------|-------------------------|-------------------------|
| | 102400 | 102527 | 102531 | 102554 | 102600 | 105420 | 145310 | 161214 | 161810 | | | | |
| Clasificación | 102400 | 115 | 1 | 0 | 2 | 12 | 6 | 0 | 0 | 0 | 136 | 84,56 (77,47-89,75) | 15,44 (10,25- 22,53) |
| | 102527 | 7 | 26 | 29 | 1 | 0 | 0 | 0 | 0 | 0 | 63 | 41,27 (29,95-53,59) | 58,73 (46,41-70,05) |
| | 102531 | 3 | 10 | 43 | 0 | 0 | 0 | 0 | 1 | 3 | 60 | 71,67 (59,16-81,56) | 28,33 (18,44-40,84) |
| | 102554 | 36 | 3 | 2 | 17 | 9 | 0 | 0 | 0 | 0 | 67 | 25,37 (16,41-37,01) | 74,63 (62,99-83,59) |
| | 102600 | 1 | 0 | 0 | 1 | 15 | 0 | 0 | 0 | 0 | 17 | 88,24 (64,41-97,97) | 11,76 (2,03-35,59) |
| | 105420 | 3 | 0 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 17 | 82,35 (58,16-94,62) | 17,65 (5,38-41,84) |
| | 145310 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 1 | 0 | 7 | 85,71 (46,65- 99,47) | 14,29 (0,53-53,35) |
| | 161214 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 12 | 0 | 14 | 85,71 (58,81- 97,24) | 14,29 (2,76-41,19) |
| | 161810 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 5 | 6 | 83,33 (41,78- 98,86) | 16,67 (1,14-58,22) |
| | Total | 165 | 40 | 77 | 21 | 36 | 20 | 6 | 14 | 8 | 387 | | |
| EP (%) | 69,70 (62,29- 76,21) | 65,00 (49,45- 77,92) | 55,84 (44,73- 66,40) | 80,95 (59,41- 92,92) | 41,67 (27,12- 57,82) | 70,00 (47,87- 85,68) | 100,00 (64,26- 100,0) | 85,71 (58,81- 97,24) | 62,50 (30,38-86,51) | | | | |
| EO (%) | 30,30 (23,79-37,71) | 35,00 (22,08-50,55) | 44,16 (33,60-55,27) | 19,05 (9,05-47,77) | 58,33 (42,18-72,88) | 30,00 (14,32-52,13) | 0,00 (0,00-35,74) | 14,29 (2,76-41,19) | 37,50 (13,49-69,62) | | | | |

Códigos de hábitats: 102400 Intermareal no vegetado, 102527 Matorral salino, 102531 Juncales, 102554 Spartina, 102600, Intermareal vegetado, 105420 Estuario, 145310 Helechales, 161214 Alisedas y 162810 Eucaliptales. EU: exactitud del usuario con intervalo de confianza de Wald ajustado al 95 % de probabilidad, EC: error de comisión con intervalo de confianza de Wald ajustado al 95 % de probabilidad, EP: exactitud del productor con intervalo de confianza de Wald ajustado al 95 % de probabilidad, EO: error de omisión con intervalo de confianza de Wald ajustado al 95 % de probabilidad.

5. DISCUSIÓN

En base al objetivo principal de este trabajo, se han realizado clasificaciones supervisadas sobre una zona de humedal, siguiendo una metodología fácilmente reproducible y adaptable a otras áreas de estudio, siempre teniendo en cuenta sus características y la información disponible.

Los resultados obtenidos, una exactitud global del 55,56 % con Sentinel 2 y un 65,37 % con WorldView 2, se encuentran en la línea de otros trabajos en humedales con objetivos similares y en base a las mismas fuentes de datos de satélite. En el caso de Sentinel 2, por ejemplo, en el estudio de Calleja *et al.* (2019) sobre clasificación de hábitats en una zona de humedal de Cantabria, se logró una precisión global del 49 ± 16 % con un índice Kappa del 0.42 ± 0.13 en una clasificación en base a OBIA. En cambio, existen otros ejemplos en los que las precisiones son superiores; Whyte *et al.* (2018) superaron el 70% al clasificar usos del suelo en una zona de humedal, Agrillo *et al.* (2021) obtuvieron precisiones superiores al 75%, aunque en ese caso se trataba de diferentes tipos de masas forestales. Le Dez *et al.* (2021) también superaron el 75% de precisión global, destacando en este caso la gran cantidad de datos de campo con los que contaron, además de un enfoque multitemporal al incluir imágenes Sentinel 2 de diferentes periodos del año.

En el caso de WorldView 2, los valores obtenidos son inferiores a los observados en varios estudios, como Jackson y Adam (2021), en el que lograron una precisión global del $73,2 \pm 2,5$ % en la diferenciación de varias especies forestales, Tarantino *et al.* (2012) alcanzaron precisiones superiores al 80% en diferentes clases de usos del suelo al utilizar las 8 bandas de WV2 o Immitzer *et al.* (2012) donde superaron el 80% diferenciando especies de arbolado.

Analizando los resultados por categorías, se han observado varias clases que presentan problemas de diferenciación en las dos clasificaciones. La primera es matorral salino (102527), con errores de comisión en torno al 55 % en ambos casos, ya que en esta categoría se incluyen objetos correspondientes principalmente a las clases intermareal no vegetado (102400) y juncales (102531). En el caso de los juncales, puede deberse a que tienen una estructura relativamente similar por lo que las variables LIDAR no han podido ser determinantes, pero la confusión con hábitats no vegetados indica que en este caso las variables espectrales no son demasiado representativas, probablemente debido a la elevada influencia del agua.

La segunda categoría en la que se han observado dificultades de diferenciación en ambas clasificaciones es en intermareal vegetado (102600), en este caso con errores de omisión superiores al 55 % dado que gran parte de los objetos de esta categoría se incluyen en la clase intermareal no vegetado (102400). Estos resultados pueden estar influenciados por la mayor importancia que presentan las variables en condiciones de marea alta respecto a las de marea baja en los modelos de clasificación.

La categoría formada por plantaciones de eucaliptos (162810), es una de las clases que ha obtenido peores resultados en la clasificación con Sentinel 2, ya que ninguna de las muestras de validación ha sido clasificadas como tal. En cambio, en la clasificación con WV2 el error de omisión se reduce al 37,50 %, este caso podría reflejar la importancia de la diferente resolución espacial de las imágenes de ambas clasificaciones.

Otra categoría que presenta problemas de diferenciación con las imágenes de Sentinel 2 es la de Estuario (105420), debido a que la clasificación incluye hábitats intermareales no vegetados (102400). Se trata de zonas con elevada influencia del agua y espectralmente muy similares, en las que la diferencia entre las clases se basa principalmente en criterios ecológicos, por lo que para distinguirlas podría ser útil incorporar información complementaria.

En la clasificación con WorldView 2, la clase *Spartina* (102554) presenta elevados errores de comisión (74,63%) debido a que se incluyen muestras de hábitat intermareal vegetado (102600) y no vegetado (102400). Resulta llamativa la confusión con zonas no vegetadas dado que apenas presentar respuesta en los índices de vegetación, pudiendo deberse a la influencia del agua.

En ambos casos, las categorías que muestran mejores resultados son intermareal no vegetado (102400), juncales (102531), helechales (145310) y alisedas (161214). En WorldView 2 también destacan los estuarios (105420) y los eucaliptales (162810).

La calidad de la clasificación depende de varios factores, como las características del área de estudio, el método seleccionado, los datos de entrenamiento y validación empleados, las clases objetivo y las habilidades del analista (Lu *et al.*, 2014; Khatami *et al.*, 2016). De hecho, uno de los aspectos clave a la hora de realizar clasificaciones es la definición de las clases, sobre ello, Ruiz *et al.* (2005) afirman que, para obtener resultados coherentes en clasificaciones con

imágenes multiespectrales, las clases tienen que ser unidades homogéneas a nivel estructural y espectral, debiendo tener una respuesta suficientemente específica. Hay que tener en cuenta que en este caso la definición de clases se ha realizado en base a criterios ecológicos, por lo que la información complementaria juega un papel fundamental al aportar información sobre rasgos no espectrales que pueden ser clave para ciertas categorías.

Por tanto, aunque de forma habitual se emplean datos espectrales como las bandas o índices de vegetación para la clasificación de coberturas del suelo (Lu y Weng, 2007), existen diversos estudios como los de Han *et al.* (2012), Lu *et al.* (2014), Fois *et al.* (2018) y Li *et al.* (2018), en los que aseguran que la combinación de rasgos espectrales y espaciales como la estructura, topografía o condiciones locales del clima o el suelo, incrementa la precisión de las clasificaciones. Además, los trabajos de Lu *et al.* (2014) y Feng *et al.* (2017) destacan la importancia de los rasgos espaciales en clasificación de coberturas en base a imágenes de alta resolución espacial. Otro ejemplo es el estudio de Xie *et al.* (2019), en el que se refleja cómo el uso de datos procedentes de varias fuentes como las bandas espectrales, índices de vegetación, texturas y factores topográficos permiten aumentar la precisión en clasificaciones de coberturas y de bosque, comparado sólo con el uso de bandas espectrales.

Por ello, en este estudio se ha optado por combinar diversas fuentes de datos, además de analizar las bandas espectrales, índices y variables LiDAR más relevantes para cada modelo. En ambos casos la media de los índices NDVI, NDWI, GNDVI y OSAVI en condiciones de marea alta han mostrado ser las variables más importantes, esto puede deberse a que al combinar las características de varias bandas consiguen optimizar la información que aportan. Coincide con las conclusiones de Tarantino *et al.* (2012) en las que las bandas Red Edge y NIR 2 de WorldView 2 son las más relevantes a la hora de discriminar categorías en las que existe una combinación de agua y vegetación. En el caso de Sentinel 2, en el estudio de Agrillo *et al.* (2021) las variables espectrales que mostraron mayor importancia para la clasificación de hábitats forestales fueron las relacionadas con el verdor, como el EVI (*Enhanced Vegetation Index*), junto con las bandas 3 y 12. La importancia de las variables también está en línea con las conclusiones de Dolz (2020) donde considera que el índice NDWI es el más adecuado para detectar masas de agua en humedales.

Otro aspecto clave en las clasificaciones supervisadas son las áreas de entrenamiento, según Sabat-Tomala *et al.* (2020), su calidad es uno de los factores que más influye en la precisión de la clasificación. El traslado de información de las áreas de referencia a los objetos

muestra la importancia de los parámetros utilizados en la segmentación, porque de esta forma se evitará tener áreas de referencia en objetos que engloben varios tipos de hábitats. En este caso hay que tener presente que ninguna de las categorías tiene un número de muestras de verdad terreno superior a 50, que según Congalton (1991) es el número mínimo a tener en cuenta cuando las clases a definir son menos de 12 y el área de estudio no supera los 4.000 km², por lo que un aspecto a mejorar sería el incremento de la muestra de verdad terreno para lograr una mejor representación de las clases.

Por otro lado, el sistema de clasificación con el que se obtendrán los mejores resultados no está claro ya que depende de cada zona de estudio, los objetivos e información disponible, de hecho, existen varias comparativas entre el enfoque en base a píxel y a objetos en zonas urbanas (Whiteside et al., 2014) y marismas (Ouyang et al., 2011) obteniendo diferentes resultados. En el caso del OBIA, las unidades básicas de procesado son los objetos generados mediante el proceso de segmentación (Lebourgeais et al., 2017), por lo que la clasificación posterior depende en gran medida de ellos (Shackelford y Davis, 2003; Su y Zheng, 2017). Seleccionar los valores óptimos de los parámetros de la segmentación resulta complicado debido a que implica criterios subjetivos y depende de la complejidad de las coberturas, especialmente de la composición y tamaño de los parches (Xie et al., 2019). Según varios estudios como los de Clinton et al. (2010) o Drăguț et al. (2014), el método recomendado para evaluar la calidad del resultado de la segmentación es el análisis visual de los objetos generados, ya que hasta el momento no existe un protocolo estandarizado para definir los parámetros a utilizar (Deur et al., 2021).

En este estudio, los parámetros se seleccionaron tras realizar múltiples combinaciones y analizar los resultados obtenidos con ellas. Finalmente se ha optado por una escala de 100, para evitar la sobresegmentación que se había detectado a escalas superiores y reducir los tiempos de procesado. Una vez fijada la escala, se han hecho pruebas con diferentes valores para los parámetros de forma y compacidad. Los mejores resultados se han obtenido con la combinación de 0,3 para forma y 0,5 para compacidad, ya que los objetos generados se ajustan a las formas visibles en la ortofotografía. Además, se han realizado diferentes asignaciones de pesos a las bandas R, G y B de la ortofotografía, obteniendo mejor resultado al dar el mismo peso a las 3.

Una vez escogido el enfoque, OBIA o clasificación a nivel de píxel, es necesario seleccionar el algoritmo de clasificación, que dependerá de la zona de estudio, el periodo analizado y el sistema de clasificación de coberturas (Li *et al.*, 2018). Actualmente están teniendo interés creciente las técnicas de *machine learning* o aprendizaje automático basadas en conjuntos de modelos. Rodríguez-Galiano *et al.* (2012) afirman que estos clasificadores generan resultados más robustos que uno de forma individual, aunque han desarrollado múltiples estudios (Lu *et al.*, 2004; Li *et al.*, 2012; Lu *et al.*, 2014; Wessel *et al.*, 2018) comparando diferentes clasificadores sin que se hayan obtenido conclusiones generales claras, parece existir cierto acuerdo en que cuando se trata de datos de varias fuentes, *Random Forest* y *Support Vector Machine* (SVM) han mostrado mejores resultados que *Maximum Likelihood Classifier* (MLC) (Xie *et al.*, 2019).

En las clasificaciones de cobertura del suelo, RF ha demostrado ser estable y relativamente eficiente debido a que incluye pocos parámetros a definir por el usuario (Pal, 2005; Lawrence *et al.*, 2006; Chan y Paelinckx, 2008). Esta técnica permite identificar el grado en el que cada variable es importante para una clasificación mediante el cálculo del descenso medio en precisión (MDA). Cuanto más disminuya la precisión del modelo, reflejado mediante valores elevados de MDA, debido a la exclusión o cambio de una variable, más importancia tendrá esa variable (Cutler *et al.*, 2007). Se trata de una característica importante y a tener en cuenta para evitar incluir variables que no aporten información relevante para la diferenciación de las clases y que incluso que puedan generar ruido sobre las determinantes.

En las primeras clasificaciones generadas se han obtenido precisiones globales medias, con índices Kappa moderados, según la interpretación del índice de Landis y Koch (1977). Entre las opciones planteadas para mejorar los resultados se encuentran el aumento y revisión de las clases de entrenamiento, realizar pruebas para intentar mejorar los parámetros utilizados en el algoritmo clasificador y la selección de variables o la agrupación de las clases en categorías fácilmente distinguibles como arbolado, pastizales y hábitats marinos, para realizar posteriormente una clasificación intragrupos, generando de esta forma una clasificación por niveles.

6. CONCLUSIONES Y LÍNEAS DE TRABAJO FUTURAS

- Los resultados obtenidos con imágenes WorldView 2 tienen una precisión global más elevada que con Sentinel 2, aunque resultan bajas en ambos casos. Esto hace necesario replantear la estrategia a seguir, valorando la posibilidad de modificar el enfoque de la clasificación, ajustar las clases a definir o realizar pruebas con diferentes algoritmos para intentar lograr resultados más precisos.

- Se ha descrito una metodología fácilmente reproducible y adaptable a otras zonas de estudio que permitiría generar cartografía actualizada de una forma relativamente rápida y sencilla.

- Debido a la importancia ecológica de los humedales, es necesario realizar una investigación más profunda que permita mejorar los resultados y que puedan ser empleados como fuente de información para planes de gestión.

7. BIBLIOGRAFÍA

Agrillo, E., Filippini, F., Pezzarossa, A., Casella, L., Smiraglia, D., Orasi, A., Attorre, F. y Taramelli, A. (2021) "Earth Observation and Biodiversity Big Data for Forest Habitats Types Classification and Mapping", *Remote Sensing*, 13, 1231.

Álvarez-Martínez, J.M., Jiménez-Alfaro, B., Barquín, J., Ondiviela, B., Recio, M., Silió-Calzada, A. y Juanes, J.A. (2018) "Modelling the area of occupancy of habitat types with remote sensing", *Methods in Ecology and Evolution*, 9, pp. 580-593.

Baatz, M. y Schäpe, A. (2000) "Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation" en Strobl, J., Blaschke, T. y Griesebner, G. (eds.), *Angewandte Geographische Informations-Verarbeitung XII*, Wichmann Verlag, Karlsruhe, Germany, pp. 12-23.

Blaschke, T. (2010) "Object based image analysis for remote sensing", *International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing*, 65(1), pp. 2-16.

Breiman, L. (2001) "Random Forests", *Machine Learning*, 45, pp. 5-32.

Breiman, L., Friedman, J., Stone, C., y Olshen, R. (1984) "Classification and regression trees", California, Estados Unidos: Wadsworth, Inc.

Calleja, F., Ondiviela, B., Galván, C., Recio, M., Juanes, J.A. (2019) "Mapping estuarine vegetation using satellite imagery: The case of the invasive species *Baccharis halimifolia* at a Natura 2000 site", *Continental Shelf Research*, 174, pp. 35-47.

Chan, J.C.W. y Paelinckx, D. (2008) "Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery", *Remote Sensing of Environment*, 112, pp. 2999-3011.

Cohen, J. (1960) "A coefficient of agreement for nominal scales", *Educational and Psychological Measurement*, 20 (1), pp. 37-46.

Congalton, R.G. (1991) "A review of assessing the accuracy of classifications of remotely sensed data", *Remote Sensing of Environment*, 37, pp. 35-46.

Clinton, N., Holt, A., Scarborough, J., Yan, L. y Gong, P. (2010) "Accuracy Assessment Measures for Object-based Image Segmentation Goodness", *Photogrammetric Engineering & Remote Sensing*, 76, pp. 289-299.

Comisión Europea (2011) "Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, Estrategia de la UE sobre la biodiversidad hasta 2020: nuestro seguro de vida y capital natural", 4 de mayo de 2011, pp. 1-19.

Consejo de las Comunidades Europeas (1992) "Directiva 92/43/CEE del Consejo, de 21 de mayo de 1992, relativa a la conservación de los hábitats naturales y de la fauna y flora silvestres" *Diario Oficial de las Comunidades Europeas*, 22 de julio de 1992 (206), pp. 7-50.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. y Lawler, J.J. (2007) "Random forests for classification in ecology" *Ecology*, 88 (11), pp. 2783-2792.

Davies, C.E. y Moss, D. (1999) "EUNIS Habitat Classification", *Final Report to the European Topic Centre on Nature Conservation*, European Environment Agency.

Davies, C.E., Moss, D. y Hill, M.O. (2004) "EUNIS Habitat Classification Revised 2004", *Report to the European Topic Centre on Nature Protection and Biodiversity*. European Environment Agency.

Deur, M., Gašparović, M. y Balenović, I. (2021) "An Evaluation of Pixel and Object-Based Tree Species Classification in Mixed Deciduous Forests Using Pansharpened Very High Spatial Resolution Satellite Imagery", *Remote Sensing*, 13, 1868.

Dolz, S. (2020) *Cartografía de masas de aguas utilizando imágenes WorldView-3 en el Marjal de Gandía*- Trabajo Fin de Grado. Universitat Politècnica de València, Escola Politècnica Superior de Gandía.

Drăguț, L., Csillik, O., Eisank, C. y Tiede, D. (2014) "Automated parameterisation for multi-scale image segmentation on multiple layers", *International Society for Photogrammetry and Remote Sensing*, 88, pp. 119-127.

Escribano, J.A., Hernández, C.G. y Tarquís, A.M. (2014) "Selección de índices de vegetación para la estimación de la producción herbácea en dehesas", *Pastos*, 44 (2), pp. 6-18.

España (2007) "Ley 42/2007, de 13 de diciembre, del Patrimonio Natural y de la Biodiversidad", *Boletín Oficial del Estado*, 14 de diciembre de 2007 (299), pp. 51275-51327.

España (2020) "Resolución de 3 de diciembre de 2020, de la Secretaría de Estado de Medio Ambiente, por la que se revisa y amplía la lista patrón de las especies silvestres presentes en España" *Boletín Oficial del Estado*, 18 de diciembre de 2020 (329), pp. 116743-116744.

Feng, Y., Lu, D., Moran, E., Dutra, L., Calvi, M. y De Oliveira, M. (2017) "Examining spatial distribution and dynamic change of urban land covers in the Brazilian Amazon using multitemporal multisensor high spatial resolution satellite imagery", *Remote Sensing*, 9, 381.

Ferreiro, R., Brea, C., Cerqueira, F., Constela, D., Pereira, L., Tabeayo, F. y Tellado, I. (2001) "Avaliación do estado de conservación dos hábitats e das poboacións de anfibios nas áreas propostas como LICs na Provincia da Coruña", Universidade da Coruña, Grupo de Investigación en Biología Evolutiva.

Fois, M., Cuenca-Lombraña, A., Fenu, G. y Bacchetta, G. (2018) "Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues and future directions", *Ecological Modelling*, 385, pp. 124-132.

Gitelson, A. y Merzlyak, M.N. (1994) "Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves", *Journal of Photochemistry and Photobiology*, 22, pp. 247-252.

Gitelson, A., Kaufman, Y. y Merzylak, M. (1996) "Use of a green channel in remote sensing of global vegetation from EOS-MODIS", *Remote Sensing of Environment*, 58, pp. 289-298.

Han, N., Wang, K., Yu, L. y Zhang, X. (2012) "Integration of texture and landscape features into object-based classification for delineating *Torreya* using IKONOS imagery", *International Journal of Remote Sensing*, 33, pp. 2003-2033.

Ichter, J., Evans, D., Richard, D., Poncet, L., Spyropoulou, R. y Martins, I. (2014) "Terrestrial habitat mapping in Europe: an overview", *European Environment Agency, Luxembourg*.

Immitzer, M., Atzberger, C. y Koukal, T. (2012) "Tree species classification with Random forest using very high spatial resolution 8-band 545 WorldView-2 satellite data", *Remote Sensing*, 4, pp. 2661-2693.

Izco, J. y Sánchez, J. M. (1996) "Los medios halófilos de la ría de Ortigueira (A Coruña, España): vegetación de dunas y marismas", *Thalassas: an international journal of marine sciences*, 12, pp. 63-100.

Jackson, C.M., y Adam, E. (2021) "Machine Learning Classification of Endangered Tree Species in a Tropical Submontane Forest Using WorldView-2 Multispectral Satellite Imagery and Imbalanced Dataset", *Remote Sensing*, 13(24), 4970.

Khatami, R., Mountrakis, G. y Stehman, S.V. (2016) "A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: general guidelines for practitioners and future research", *Remote Sensing of Environment*, 177, pp. 89-100.

Khun, M. (2020). "caret: Classification and Regression Training. R package version 6.0-86". Disponible en: <https://CRAN.R-project.org/package=caret>

Landis, J. R., y Koch, G. G. (1977) "The Measurement of Observer Agreement for Categorical Data", *Biometrics*, 33(1), pp. 159-174.

Lawrence, R.L., Wood, S.D. y Sheley, R.L. (2006) "Mapping invasive plants using hyperspectral imagery and breiman cutler classifications (randomforest)", *Remote Sensing of Environment*, 100, pp. 356-362.

Le Dez, M., Robin, M. y Launeau, P. (2021) "Contribution of Sentinel-2 satellite images for habitat mapping of the Natura 2000 site 'Estuaire de la Loire' (France)", *Remote Sensing Applications: Society and Environment*, 24, 100637.

Lebourgeois, V., Dupuy, S., Vintrou, É., Ameline, M., Butler, S. y Bégué, A. (2017) "A Combined Random Forest and OBIA Classification Scheme for Mapping Smallholder Agriculture at Different Nomenclature Levels Using Multisource Data (Simulated Sentinel-2 Time Series, VHRS and DEM)", *Remote Sensing*, 9 (3), 259.

Li, G., Lu, D., Moran, E.; y Sant'Anna, S.J.S. (2012) "A comparative analysis of classification algorithms and multiple sensor data for land use/land cover classification in the Brazilian Amazon", *Journal of Applied Remote Sensing*, 6 (1), pp. 061706-061717.

Li, N., Lu, D., Wu, M., Zhang, Y. y Lu, L. (2018) "Coastal wetland classification with multiseasonal high-spatial resolution satellite imagery", *International Journal of Remote Sensing*, 39 (23), pp. 8963-8983.

Liaw, A. y Wiener, M. (2002) "Classification and Regression by randomForest", *R News*, 2(3), pp. 18-22.

Lu, D. y Weng, Q. (2007) "A survey of image classification methods and techniques for improving classification performance", *International Journal of Remote Sensing*, 28, pp. 823-870.

Lu, D., Li, G., Moran, E. y Kuang, W. (2014) "A comparative analysis of approaches for successional vegetation classification in the Brazilian Amazon", *GIScience & Remote Sensing*, 51, pp. 695-709.

Lu, D., Mausel, P., Batistella, M. y Moran, E. (2004) "Comparison of land-cover classification methods in the Brazilian Amazon Basin", *Photogrammetric Engineering and Remote Sensing*, 70, pp. 723-732.

Martínez, S., Ramil, P. y Chuvieco, E. (2010) "Monitoring loss of biodiversity in cultural landscapes. New methodology based on satellite data", *Landscape and Urban Planning*, 94, pp. 127-140.

Mauro, F., Hernando, A., García, A., Núñez, M., Martín-Hernández, S. y Rosario, R. (2013) "Delineación de rodales para la ordenación forestal a partir de información LIDAR", *6º Congreso Forestal Español, Sociedad Española de Ciencias Forestales, Vitoria-Gasteiz*.

McFeeters, S. K. (1996) "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features", *International Journal of Remote Sensing*, 17(7), pp. 1425-1432.

MITECO (Ministerio para la Transición Ecológica y el Reto Demográfico) (2021) "Situación actual y perspectivas de conservación y restauración de los humedales españoles hasta 2030. Informe con motivo del Día Internacional de los Humedales y 50 aniversario del Convenio Ramsar".

Muñoz, P. (2013) "Apuntes de Teledetección: Índices de vegetación", *Centro de Información de Recursos Naturales, Ministerio de Agricultura, Chile*.

Ouyang, Z.T., Zhang, M.Q., Xie, X., Shen, Q., Guo, H.Q. y Zhao, B. (2011) "A comparison of pixel-based and object-oriented approaches to VHR imagery for mapping saltmarsh plants", *Ecological Informatics*, 6, pp. 136-146.

Pal, M. (2005) "Random forest classifier for remote sensing classification", *International Journal of Remote Sensing*, 26, pp. 217-222.

Parlamento Europeo y Consejo de la Unión Europea (2000) "Directiva 2000/60/CE del Parlamento Europeo y el Consejo de 23 de octubre de 2000 por la que se establece un marco comunitario de actuación en el ámbito de la política de aguas" *Diario Oficial de las Comunidades Europeas*, 22 de diciembre de 2000 (327), pp. 1-73.

Parlamento Europeo y Consejo de la Unión Europea (2007) "Directiva 2007/2/CE del Parlamento Europeo y del Consejo de 14 de marzo de 2007 por la que se establece una infraestructura de información espacial en la Comunidad Europea (Inspire)" *Diario Oficial de la Unión Europea*, 25 de abril de 2007 (108), pp. 1-14.

QGIS (2021). *QGIS Project*. Disponible en: <https://www.qgis.org/>.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Disponible en: <https://www.R-project.org/>.

Rivas-Martínez, S. (2004). *Global bioclimatics (Clasificación Bioclimática de la Tierra)*. Disponible en: https://webs.ucm.es/info/cif/book/bioc/global_bioclimatics_0.html (Accedido: 1 mayo 2022).

RStudio Team (2018). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA URL Disponible en: <http://www.rstudio.com/>.

Ramil-Rego, P., Rodríguez Guitián, M. A., Ferreiro da Costa, J., Rubinos Román, M., Gómez-Orellana, L., de Nóvoa Fernández, B., Hinojo Sánchez, B. A., Martínez Sánchez, S., Cillero Castro, C., Díaz Varela, R. A., Rodríguez González, P. M. y Muñoz Sobrino, C. (2008) Os Hábitats de Interese Comunitario en Galicia, Fichas descriptivas, Monografías do Ibader. Universidade de Santiago de Compostela. Lugo.

Redel, H.M. y Zubillaga, M.M. (2014) "Aproximaciones para el manejo sitio-específico de la fertilización nitrogenada en maíz a través de índices espectrales", *Informaciones Agronómicas de Hispanoamérica, International Plant Nutrition Institute*, 15, pp. 15-22.

Rodríguez, M. y Ramil-Rego, P. (2007) "Clasificaciones climáticas aplicadas a Galicia: revisión desde una perspectiva biogeográfica", *Recursos Rurais* (1), 3, pp. 31-53.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M. y Rigol-Sanchez, J.P. (2012) "An assessment of the effectiveness of a random forest classifier for land-cover classification", *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, pp. 93-104.

Rosenfield, G.H. y Fitzpatrick-Lins, K. (1986) "A coefficient of agreement as a measure of thematic classification accuracy", *Photogrammetric Engineering and Remote Sensing*, 52, pp. 223-227.

Rondeaux, G., Steven, M. y Baret, F. (1996) "Optimization of soil-adjusted vegetation indices", *Remote Sensing of Environment*, 55(2), pp- 55-107.

Rouse J.W., Haas, R.H., Schell, J.A., Deering, D.W. (1974) "Monitoring vegetation systems in the Great Plains with ERTS", *Proceedings of the Third Earth Resources Technology Satellite - 1 Symposium*, pp. 301-317.

Ruiz, L.A., Recio, J.A. y Fernández, A. (2005) "Clasificación de entornos forestales mediterráneos mediante técnicas de análisis de texturas", *Cuadernos de la Sociedad Española de Ciencias Forestales*, 19, pp. 187-192.

Sabat-Tomala, A., Raczko, E. y Zagajewski, B. (2020) "Comparison of Support Vector Machine and Random Forest Algorithms for Invasive and Expansive Species Classification Using Airborne Hyperspectral Data", *Remote Sensing*, 12 (3), 516.

Shamsoddini, A., Trinder, J.C. y Turner, R. (2013) "Pine plantation structure mapping using WorldView-2 multispectral image", *International Journal of Remote Sensing*, 34 (11), pp. 3986-4007.

Shackelford, A.K. y Davis, C.H. (2003) "A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas", *IEEE Transactions on Geoscience and Remote Sensing*, 41, pp. 2354-2364.

Spoto, F., Sy, O., Laberinti, P., Martimort, P., Fernandez, V., Colin, O., Hoersch, B. y Meygret, A. (2012) "Overview Of Sentinel-2", *IEEE International Geoscience and Remote Sensing Symposium*, pp. 1707-1710.

Su, T. y Zheng, S. (2017) "Local and global evaluation for remote sensing image segmentation", *International Society for Photogrammetry and Remote Sensing*, 130, pp. 256-276.

Tarantino, C., Adamo, M., Pasquariello, G., Lovergine, F., Blonda, P. y Tomaselli, V. (2012) "8-Band image data processing of the WorldView-2 satellite in a wide area of applications", *National Council of Researches (CNR), Italy*.

Trimble (2014) "eCognition Developer Reference Book", *Germany GmbH, Arnulfstrasse 126, D-80636 Munich, Germany*.

UNESCO (1971) “Convención relativa a los Humedales de Importancia Internacional especialmente como hábitat de aves acuáticas”, *Oficina de Normas Internacionales y Asuntos Legales*, 3 de febrero de 1971, pp. 1-3.

USDA, United States Department of Agriculture (2022) *FUSION/LDV: Software for LIDAR Data Analysis and Visualization, version 4.40*. Pacific Northwest Research Station, Bob McGaughey.

Varela, R.A., Ramil-Rego, P., Calvo, S. y Muñoz, S. (2007) “Automatic habitat classification methods based on satellite images: A practical assessment in the NW Iberia coastal mountains”, *Environmental Monitoring and Assessment*, 144(1-3), pp. 229-250.

Wessel, M., Brandmeier, M. y Tiede, D. (2018) “Evaluation of different machine learning algorithms for scalable classification of tree types and tree species based on Sentinel-2 data”, *Remote Sensing*, 10, 1419.

Whiteside, T.G., Maier, S.W. y Boggs, G.S. (2014) “Area-based and location-based validation of classified image objects”, *International Journal of Applied Earth Observation and Geoinformation*, 28, pp. 117-130.

Whyte, A., Ferentinos, K. y Petropoulos, G. (2018) “A new synergistic approach for monitoring wetlands using Sentinels-1 and 2 data with object-based machine learning algorithms”, *Environmental Modelling & Software*, 104, pp. 40-54.

Xie, Z., Chen, Y., Lu, D., Li, G. y Chen, E. (2019) “Classification of land cover, forest and tree species classes with ZiYuan-3 multispectral and stereo data”, *Remote sensing*, 11, 164.

ANEXO I. DESCRIPCIÓN DE LOS HÁBITATS DE LA ZONA DE ESTUDIO

| Código trabajo | EUNIS | | | LPETH | | Anexo I Directiva 92/43/CEE | | Prioritario |
|----------------|---------------|---------------|---|--------|--|-----------------------------|--|-------------|
| | Código (2012) | Código (2017) | Nombre | Código | Nombre | Código | Nombre | |
| 1024 | A2.4 | - | Littoral mixed sediments | 1 | Comunidades costeras y halofíticas | 1140 | Llanos fangosos o arenosos que no están cubiertos de agua cuando hay marea baja | |
| 102527 | A2.527 | - | Atlantic salt scrubs | 15.62 | Matorrales salinos atlánticos | 1420 | Matorrales halófilos mediterráneos y termoatlánticos (<i>Sarcocornetea fruticosi</i>) | |
| 102531 | A2.531 | - | Atlantic upper shore communities | 15.33 | Comunidades atlánticas del supraestero | 1330 | Pastizales salinos atlánticos (<i>Glauco-Puccinellietalia maritimae</i>) | |
| 102554 | A2.554 | - | Flat-leaved [Spartina] swards | 15.21 | Praderas de Spartina marítima | 1320 | Pastizales de Spartina (<i>Spartinion maritimae</i>) | |
| 1026 | A2.6 | - | Littoral sediments dominated by aquatic angiosperms | 1 | Comunidades costeras y halofíticas | 1140 | Llanos fangosos o arenosos que no están cubiertos de agua cuando hay marea baja | |
| 10542 | A5.42 | | Sublittoral mixed sediment in variable salinity (estuaries) | 1 | Comunidades costeras y halofíticas | 1130 | Estuarios | |
| 14531 | E5.31 | - | Sub-Atlantic bracken fields | 31.861 | Helechales de <i>Pteridium aquilinum</i> subatlánticos | - | - | |
| 161214 | G1.214 | T1-214 | Northern Iberian [Alnus] galleries | 44.34 | Alisedas del norte de la Península Ibérica | 91E0 | Bosques aluviales de <i>Alnus glutinosa</i> y <i>Fraxinus excelsior</i> (<i>Alno-Padion</i> , <i>Alnion incanae</i> , <i>Salicion albae</i>) | si |
| 16281 | G2.81 | T2-A1 | Eucalyptus plantations | 83.322 | Plantaciones de eucaliptos (<i>Eucalyptus</i> spp.) | - | - | |

ANEXO II. SCRIPT DE CLASIFICACIÓN

```
getwd()
setwd("C:/TFM/_2_PROCESOS/_3_CLASIFICACION")
library(randomForest)
library(caret)

### Modelo con variables de WorldView 2 y LIDAR
#1. Datos de entrenamiento y población
Train_WV<-read.csv("C:/TFM/_2_PROCESOS/_3_CLASIFICACION/ZS_VT_WV.csv")
View(Train_WV)
head(Train_WV)

Pop_WV<-read.csv("C:/TFM/_2_PROCESOS/_3_CLASIFICACION/ZS_POP_WV.csv")
View(Pop_WV)
head(Pop_WV)

#2. RF con todas las variables
set.seed(1000)
Modelo_WV<-randomForest(as.factor(Train_WV$VT)~
chm_m+ chm_d+ dsm_m+ dsm_d+ dem_m+dem_d+WVb1a_m+ WVb1a_d+WVb2a_m+WVb2a
_d+WVb3a_m+WVb3a_d+WVb4a_m+ WVb4a_d+WVb5a_m+WVb5a_d+WVb6a_m+WVb6a_d
+ WVb7a_m+ WVb7a_d+WVb8a_m+ WVb8a_d+ WVGNDVla_m+WVGNDVla_d+ WVNDVla_m+
WVNDVla_d+WVNDWla_m+ WVNDWla_d+ WVOSAVla_m+ WVOSAVla_d+ WVREVla_m+WVR
EVla_d+ WVb1b_m+WVb1b_d+ WVb2b_m+WVb2b_d+ WVb3b_m+ WVb3b_d+WVb4b_m+WV
b4b_d+WVb5b_m+WVb5b_d+WVb6b_m+WVb6b_d+WVb7b_m+ WVb7b_d+ WVb8b_m+WVb
8b_d+WVGNDVlb_m+WVGNDVlb_d+WVNDVlb_m+WVNDVlb_d+WVNDWlb_m+WVNDWlb_d
+ WVOSAVlb_m+ WVOSAVlb_d+WVREVlb_m+WVREVlb_d, ntree=500, data=Train_WV,
importance=TRUE, na.action=na.exclude)

Modelo_WV
Modelo_WV$confusion
Modelo_WV$importance
Modelo_WV$serr.rate
```

```
varImpPlot(Modelo_WV)
importance(Modelo_WV)
```

#3. RF con las variables importantes según valor de MDG

```
set.seed(1000)
Modelo_RF_Gini<-randomForest(as.factor(Train_WV$VT)~ WVNDVla_m+ WVNDWla_m+
WVOSAVla_m+ WVGNDVla_m+ WVREVIb_m+ WVb7a_m+ WVNDWib_m+ WVb8a_m+
WVREVIa_m+ WVOSAVib_m+ WVGNDVib_m+ WVNDVib_m+ WVb1a_m+ WVb2a_m+
WVb4b_m+ WVb5a_m+ WVb6a_m+ WVb3a_m+ WVb5b_m+ WVb8b_m+ WVb1b_m+
WVb7b_m+ dsm_m+ WVb4a_m+ WVb2b_m+ WVb6b_m+ WVb3b_m+ chm_d+ dem_m+
chm_m, ntree=500, data=Train_WV, importance=TRUE, na.action=na.exclude)
```

```
Modelo_RF_Gini
```

#4. Aplicación del modelo sobre la población

```
Pred_RF_Gini <-predict(Modelo_RF_Gini, Pop_WV, type ="response")
Prob_RF_Gini <-predict(Modelo_RF_Gini, Pop_WV, type ="prob" )
```

```
Pop_pred_RF_Gini<-data.frame(Pop_WV, Pred_RF_Gini, Prob_RF_Gini)
View(Pop_pred_RF_Gini)
```

```
write.csv(Pop_pred_RF_Gini[,c(1,60:69)],file="C:/TFM/_2_PROCESOS/_3_CLASIFICACION/Mod
elo_RF_WV_Gini.csv")
```

#5. Validación, carga del csv que contiene las predicciones del modelo y las clases de referencia

```
Valid_WV<-
read.csv("C:/TFM/_2_PROCESOS/_4_VALIDACION/Valid_Modelo_RF_WV_Gini.csv")
View(Valid_WV)
```

```
CM_WV<-confusionMatrix(as.factor(Valid_WV$Pred_RF_Gini),as.factor(Valid_WV$Valid))
CM_WV
```

```
### Modelo con variables de Sentinel 2 y LIDAR
```

```
#1. Datos de entrenamiento y población
```

```
Train_S2<-read.csv("C:/TFM/_2_PROCESOS/_3_CLASIFICACION/ZS_VT_S2.csv")
```

```
View(Train_S2)
```

```
head(Train_S2)
```

```
Pop_S2<-read.csv("C:/TFM/_2_PROCESOS/_3_CLASIFICACION/ZS_POP_S2.csv")
```

```
View(Pop_S2)
```

```
head(Pop_S2)
```

```
#2. RF con todas las variables
```

```
set.seed(1000)
```

```
Modelo_S2<-randomForest(as.factor(Train_S2$VT)~
```

```
chm_m+ chm_d+ dsm_m+ dsm_d+ dem_m+dem_d+S2b2a_m+ S2b2a_d+ S2b3a_m+ S2b3a_d+
```

```
S2b4a_m+ S2b4a_d+ S2b5a_m+ S2b5a_d+ S2b6a_m+ S2b6a_d+ S2b7a_m+ S2b7a_d+
```

```
S2b8a_m+ S2b8a_d+ S2NDVla_m+ S2NDVla_d+ S2GNDVla_m+ S2GNDVla_d+ S2NDWla_m+
```

```
S2NDWla_d+ S2OSAVla_m+ S2OSAVla_d+ S2REVIa_m+ S2REVIa_d+ S2b2b_m+ S2b2b_d+
```

```
S2b3b_m+ S2b3b_d+ S2b4b_m+ S2b4b_d+ S2b5b_m+ S2b5b_d+ S2b6b_m+ S2b6b_d+
```

```
S2b7b_m+ S2b7b_d+ S2b8b_m+ S2b8b_d+ S2GNDVlb_m+ S2GNDVlb_d+ S2NDVlb_m+
```

```
S2NDVlb_d+ S2NDWlb_m+ S2NDWlb_d+ S2OSAVlb_m+ S2OSAVlb_d+ S2REVIb_m+ S2REVIb_d+
```

```
S2b11a_m+ S2b11a_d+ S2b12a_m+ S2b12a_d+ S2b11b_m+ S2b11b_d+ S2b12b_m+ S2b12b_d,
```

```
ntree=500, data=Train_S2, importance=TRUE, na.action=na.exclude)
```

```
Modelo_S2
```

```
Modelo_S2$confusion
```

```
Modelo_S2$importance
```

```
Modelo_S2$serr.rate
```

```
varImpPlot(Modelo_S2)
```

```
importance(Modelo_S2)
```

```
#3. RF con las variables importantes según valor de MDG
```

```
set.seed(1000)
```

```
Modelo_RF_Gini_S2<-randomForest(as.factor(Train_S2$VT)~
```

```
S2GNDVla_m+S2b12b_m+ S2b12a_m+ S2NDWla_m+ S2b4b_m+ S2b11b_m+ S2OSAVla_m+
```

```
S2NDVla_m+ dsm_m+ S2b4a_m+ S2b11a_m+ S2OSAVlb_m+ S2b6a_m+ S2NDVlb_m+
```

```
S2b8b_m+ S2b7a_m+ S2b8a_m+ S2GNDV1b_m+ S2b3a_m+ S2NDW1b_m+ S2b7b_m+ S2b6b_m+  
S2b2a_m+ dem_m+ dsm_d+ chm_m+ S2b5b_m+ S2REV1a_m+ S2b3b_m+ chm_d, ntree=500,  
data=Train_S2, importance=TRUE, na.action=na.exclude)
```

```
Modelo_RF_Gini_S2
```

#4. Aplicación del modelo sobre la población

```
Pred_RF_Gini_S2 <-predict(Modelo_RF_Gini_S2, Pop_S2, type ="response")
```

```
Prob_RF_Gini_S2 <-predict(Modelo_RF_Gini_S2, Pop_S2, type ="prob" )
```

```
Pop_pred_RF_Gini_S2<-data.frame(Pop_S2, Pred_RF_Gini_S2, Prob_RF_Gini_S2)
```

```
View(Pop_pred_RF_Gini_S2)
```

```
write.csv(Pop_pred_RF_Gini_S2[,c(1,64:73)],file="C:/TFM/_2_PROCESOS/_3_CLASIFICACION/  
Modelo_RF_S2_Gini.csv")
```

#5. Validación, carga del csv que contiene las predicciones del modelo y las clases de referencia

```
Valid_S2<-read.csv("C:/TFM/_2_PROCESOS/_4_VALIDACION/Valid_Modelo_RF_S2_Gini.csv")
```

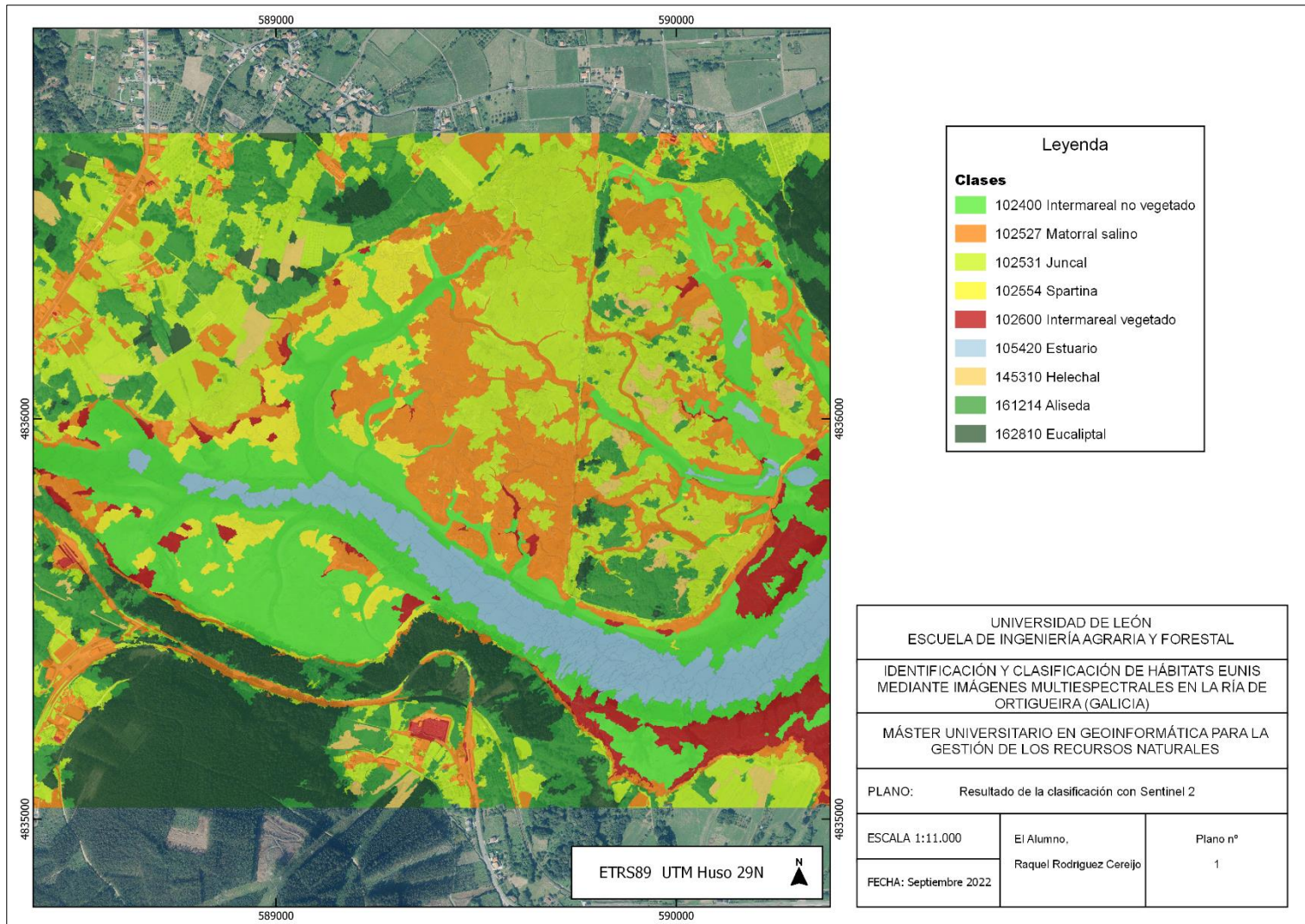
```
View(Valid_S2)
```

```
CM_S2<-confusionMatrix(as.factor(Valid_S2$Pred_RF_Gini),as.factor(Valid_S2$Valid))
```

```
CM_S2
```

ANEXO III. RESULTADOS CLASIFICACIONES

A. Clasificación con Sentinel 2



B. Clasificación con WorldView 2

