UNIVERSITAT POLITÈCNICA DE VALÈNCIA

DEPARTAMENTO DE INGENIERÍA HIDRÁULICA Y MEDIO AMBIENTE

Desarrollo de procedimientos para la detección del abandono de cultivos de cítricos utilizando técnicas de teledetección

TESIS DOCTORAL

Autor:

D. Sergio Morell Monzó

Directores:

Dra. Dña. María Teresa Sebastiá Frasquet

Dr. D. Javier Estornell Cremades

Valencia, marzo 2023

Agradecimientos

En primer lugar, me gustaría dar las gracias a mis directores, la Dra. María Teresa Sebastiá Frasquet y el Dr. Javier Estornell Cremades, por darme la oportunidad de realizar esta tesis, por guiarme en el arte de la investigación y por toda su ayuda en lo profesional y personal. En especial, quiero dar las gracias a mis padres, por hacer siempre por mi lo posible e imposible, a Pilar, por acompañarme en mis aventuras y por ponérmelo todo tan fácil, y al resto de mi familia y amigos. También me gustaría dar las gracias al Dr. Enrique Moltó García por acogerme durante mi estancia en el Instituto Valenciano de Investigaciones Agrarias. Quiero dar las gracias a mis compañeros, profesores e investigadores, del Campus de Gandia de la UPV con los que he compartido el día a día, por acogerme tan bien, por ayudarme en los momentos de crisis y por la buena convivencia durante estos años. Finamente quiero agradecer a todas las personas que se han interesado por este trabajo de una u otra manera y que me han ayudado durante el camino.

RESUMEN

El abandono de tierras agrícolas es un problema creciente en muchas regiones del planeta con importantes consecuencias a nivel socioeconómico y medioambiental. La Comunitat Valenciana (España) es la principal región productora de cítricos de Europa, sin embargo, actualmente se está produciendo un abandono significativo de estos cultivos. La presente tesis pretende desarrollar procedimientos para la identificar el abandono de los cultivos de cítricos a través de técnicas de teledetección y aprendizaje automático. Los métodos basados en teledetección se consideran especialmente convenientes debido a su alto grado de automatización y a su eficiencia en tiempo y coste, comparado con otras técnicas cartográficas habituales basadas en fotointerpretación y visitas de campo. La tesis desarrolla procedimientos para la identificación del estado de los cultivos utilizando diferentes tipos de datos de observación de la Tierra, como son: imágenes multiespectrales de los satélites Sentinel-2 y WorldVIew-3, imágenes aéreas, y nubes de puntos altimétricas derivadas de LiDAR aerotransportado y procesos fotogramétricos a partir de imágenes tomadas con vehículos aéreos no tripulados. Se desarrollaron diferentes metodologías para explotar dichas fuentes de datos en un paisaje agrícola altamente fragmentado con un tamaño promedio de las parcelas entre 0,25 ha y 0,50 ha. El análisis tomó como referencia la base de datos catastral de parcelas agrícolas y se identificaron tres estados de los cultivos en campo: productivo, no productivo y abandonado. Los resultados más precisos se obtuvieron al utilizar las imágenes aéreas o imágenes de satélite WorldView-3, incluyendo el cálculo de descriptores de textura, con una exactitud global entre 90-95%. La importancia de los descriptores de textura extraídos a partir de imágenes de resolución submétrica radica en su capacidad para identificar los patrones de plantación de los cultivos. Este enfoque permitió obtener una alta exactitud a partir de una única imagen. No obstante, también son destacables los resultados obtenidos al utilizar series temporales de imágenes Sentinel-2 con exactitudes globales en torno al 90%. Las diferencias existentes en las series temporales de índices de vegetación y humedad de los tres estados del cultivo permitieron mejorar el rendimiento de clasificación con respecto al enfoque basado en una única fecha. Diferentes casos de estudio demostraron la capacidad de identificar el abandono de tierras a través de teledetección. La información cuantitativa y espacialmente explícita generada sobre el estado de los cultivos puede ser una herramienta útil para la gestión agrícola, para la implementación de medidas de gestión del paisaje y para mejorar las estimaciones anuales de rendimiento de los cítricos.

RESUM

L'abandonament de terres agrícoles és un problema creixent en moltes regions del planeta amb importants conseqüències en l'àmbit socioeconòmic i mediambiental. La Comunitat Valenciana (España) és la principal regió productora de cítrics d'Europa, no obstant això, actualment està produint-se'n un abandonament significatiu d'estos cultius. La present tesi pretén desenvolupar procediments per identificar l'abandonament dels cultius de cítrics per mitjà de tècniques de teledetecció i aprenentatge automàtic. Els mètodes basats en teledetecció es consideren especialment convenients a causa del seu alt grau d'automatització i a la seua eficiència en temps i cost, comparat amb altres tècniques cartogràfiques habituals basades en fotointerpretació i visites de camp. La tesi desenvolupa procediments per a la identificació de l'estat dels cultius utilitzant diferents tipus de dades d'observació de la Terra, com són: imatges multiespectrals dels satèl·lits Sentinel-2 i WorldVIew-3, imatges aèries i núvols de punts altimètrics derivats de LiDAR aerotransportat i processos fotogramètrics a partir d'imatges preses amb vehicles aeris no tripulats. Es van desenvolupar diferents metodologies per explotar estes fonts de dades en un paisatge agrícola altament fragmentat amb una grandària mitjana de les parcel·les entre 0,25 ha i 0,50 ha. L'anàlisi va prendre com a referència la base de dades cadastral de parcel·les agrícoles i es van identificar tres estats dels cultius al camp: productiu, no productiu i abandonat. Els resultats més precisos es van obtenir en utilitzar imatges aèries o imatges de satèl·lit WorldVIew-3, incloent-ne el càlcul de descriptors de textura, amb una exactitud global entre 90-95%. La importància dels descriptors de textura extrets a partir d'imatges de resolució submètrica radica en la seua capacitat per a identificar els patrons de plantació dels cultius. Aquest enfocament va permetre obtenir una alta exactitud a partir d'una única imatge. No obstant això, també són destacables els resultats obtinguts en utilitzar sèries temporals d'imatges Sentinel-2, amb exactituds globals entorn del 90%. Les diferències existents en les sèries temporals d'índex de vegetació i humitat dels tres estats del cultiu van permetre millorar el rendiment de classificació respecte a l'enfocament basat en una única data. Diferents casos d'estudi han demostrat la capacitat d'identificar l'abandonament de terres per mitjà de teledetecció. La informació quantitativa i espacialment explícita generada sobre l'estat dels cultius pot ser un ferrament útil per a la gestió agrícola, per a la implementació de mesures de gestió del paisatge i per millorar les estimacions anuals de rendiment dels cítrics.

ABSTRACT

Agricultural land abandonment is a growing issue in many regions of the planet with important socioeconomic and environmental consequences. The Comunitat Valenciana region (Spain) is the main citrus-producing in Europe, however, there is a currently significant abandonment of these crops. This thesis aims to develop procedures to identify citrus crop abandonment through remote sensing and machine learning techniques. Remote sensing-based methods are considered particularly convenient due to their high degree of automation and their efficiency in time and cost, compared to other usual cartographic techniques based on photo-interpretation and field visits. The thesis develops procedures for crop status identification using different types of Earth observation data, such as: multispectral images from Sentinel-2 and WorldView-3 satellites, aerial images, and altimetric point clouds derived from airborne LiDAR and photogrammetric processes from images taken by unmanned aerial vehicles. Different methodologies were developed to apply these data sources in a highly fragmented agricultural landscape with an average parcel size between 0.25 ha and 0.50 ha. The analysis was based on the cadastral database of agricultural parcels and three crop statuses were identified at field: productive, non-productive and abandoned. The most accurate results were obtained using aerial images or WorldView-3 satellite images, including the computation of texture descriptors, with an overall accuracy between 90-95%. The importance of texture descriptors extracted from sub-metric resolution images lies in their ability to identify crop planting patterns. This approach made it possible to obtain high accuracy from a single image. However, the results obtained when using time series of Sentinel-2 images are also remarkable, with an overall accuracy around 90%. The differences in the time series of vegetation and moisture indices of the three crop statuses improved the classification accuracy compared to the single date approach. Different case studies demonstrate the ability to identify agricultural land abandonment through remote sensing. The quantitative and spatially explicit information generated can be a useful tool for agricultural management, for implementing landscape management actions and for improving annual citrus yield estimations.

ÍNDICE

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

CAPÍTULO 1

Introducción. El problema del abandono de los cultivos de cítricos en la Comunitat Valenciana

1.1. INTRODUCCIÓN

1.1.1. El abandono de tierras agrícolas: concepto y problemática

El abandono de tierras agrícolas se considera una problemática global con importantes implicaciones medioambientales y socioeconómicas (Prishchepov, 2020). Sin embargo, el concepto de abandono de tierras o tierra abandonada es complejo y no existe definición única en todas las fuentes científicas y legales. El abandono de la actividad agrícola no siempre genera una situación de ausencia de uso de la tierra. En ocasiones el abandono de la actividad agrícola se produce por un cambio en el uso del suelo (procesos de urbanización, plantaciones forestales, etc.). Por tanto, en primer lugar, es necesario diferenciar entre el concepto más amplio de abandono de la actividad agrícola y el concepto más específico de abandono de tierras agrícolas, que se reserva para aquellos casos sin uso de la tierra ni actividad económica (Corbelle-Rico y Cresente-Maceda, 2008).

Keenleyside y Tucker (2010) definen la tierra abandonada como aquella en la que se ha detenido la gestión agrícola y ha comenzado la sucesión de otras especies vegetales. Gałecka-Drozda y Zachariasz (2017) llevan este concepto más allá y enfatizan que para que un campo se considere abandonado la sucesión de especies y los cambios físicos sobre los cultivos deben avanzar y ser claramente visibles. La Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO) considera una tierra abandonada cuando no ha habido actividad agrícola durante un periodo mayor o igual a cinco años (FAOSTAT, 2020). Esta definición es útil desde el punto de vista administrativo, pero al mismo tiempo es difícil de utilizar en un contexto científico, ya que es difícil afirmar, a partir de imágenes o incluso visitas *in situ*, el tiempo que un campo lleva abandonado. Por tanto, en este trabajo nos basaremos en los cambios físicamente visibles del abandono de tierra, los cuales se pueden determinar en el campo. De manera general los principales cambios que se producen en un campo abandonado son la expansión de la vegetación silvestre, hierbas, arbustos y bosques, y la desaparición de los cultivos y de los patrones tradicionales del uso de la tierra (Cocca et al., 2012 y Czesak et al., 2021). Como señala Pointereau et al. (2008), esta gran variedad de definiciones en la literatura científica prueba la complejidad y variabilidad del abandono de tierras agrícolas, que se presenta como un problema complejo.

El abandono de tierras agrícolas (en adelante ATA) es un problema creciente alrededor del mundo que influye directamente sobre aspectos clave del medio ambiente (biodiversidad, fijación de CO_2, generación de nuevos ecosistemas, incendios forestales, suelo y recursos hídricos) y la socioeconomía (producción alimentaria, economía, valores culturales y paisaje) (Prishchepov, 2020; Subedi et al., 2021). Aunque el ATA ha recibido atención desde principios de la década de 1990, alrededor del 95% de los artículos se ha publicado en los últimos 15 años (Subedi et al. 2021), lo que indica una creciente preocupación por parte de la comunidad científica. La revisión bibliográfica realizada por Subedi et al. 2021 muestra el surgimiento de la problemática a nivel mundial y el cada vez mayor interés por parte de la comunidad científica. En los últimos años se han publicado multitud de estudios acerca de los factores impulsores del abandono y sus consecuencias. La mayoría de casos de estudio se ubican en Asia y Europa, siendo China y España los países con más estudios publicados (Subedi et al. 2021). No obstante, el ATA es un problema de alcance global que está presente en

multitud de áreas del planeta. El ATA ocurre en contextos sociales, ambientales y económicos variables (Kosmas et al. 2015). Sin embargo, es particularmente frecuente en zonas montañosas y en áreas con estructuras parcelarias altamente fragmentadas (Czesak et al., 2021).

Los factores impulsores del ATA son diversos y se agrupan en las siguientes categorías: factores demográficos y sociales, características de los campos de cultivo, factores biofísicos (ambientales), factores económicos, legislativos y sociopolíticos (Subedi et al. 2021). Con respecto a los efectos del ATA, diversos estudios han reportado efectos negativos derivados del ATA como la degradación de los paisajes rurales y agrarios, el aumento en la frecuencia de incendios forestales, la erosión del suelo, impactos negativos sobre la economía y la seguridad alimentaria, reducción de la biodiversidad, degradación de aspectos sociales y culturales ligados a la agricultura, disminución de algunas poblaciones de aves y aumento de la escorrentía y erosión del suelo. Sin embargo, otros estudios han reportado efectos positivos derivados del ATA como el aumento de la capacidad de fijación de CO_2, el aumento de la biodiversidad, la mejora de las características del suelo, disminución de la escorrentía y la erosión del suelo y mayor disponibilidad de hábitat y aumento de poblaciones de aves (Subedi et al. 2021). Dependiendo del tipo de cultivo y el contexto social y ambiental, el ATA puede percibirse como un problema o como una oportunidad. Estas características hacen del abandono de tierras un fenómeno complejo que debe ser estudiado a diferentes escalas (global, regional y local) (Strijker, 2005).

La agricultura representa más de la mitad del territorio de la Unión Europea, garantiza la producción alimentaria, gestiona importantes recursos naturales y apoya el desarrollo socioeconómico de las zonas rurales (Terres et al. 2015). Además, se estima que el 50% de todas las especies de plantas y animales (incluidas algunas de las que figuran en la Directiva Hábitat de la Unión Europea) dependen de la gestión agrícola (Terres et al. 2015). La actividad agrícola es esencial para garantizar estas funciones y, por tanto, evitar el ATA debe ser un objetivo importante de la Política Agraria Común Europea (PAC). Sin embargo, en la Unión Europea en torno al 11% de la superficie agrícola se encuentra en alto riesgo de abandono para el periodo 2015-2030 (Perpiña-Castillo et al., 2018). El ATA es una importante barrera para el desarrollo sostenible de muchas regiones europeas (Leal Filho et al. 2017) y según las proyecciones, es probable que haya niveles significativos de ATA en Europa durante las próximas décadas (Terres et al. 2013). Por tanto, es necesario desarrollar métodos eficientes en tiempo y coste para identificar el ATA y proporcionar información sobre los cambios en el uso del suelo agrícola y mejorar continuamente la PAC. La información temporal y espacialmente explícita sobre el ATA puede ayudar a mejorar los instrumentos de política para contrarrestar o revertir el proceso y también puede ayudar a implementar medidas adecuadas de seguimiento y gestión del paisaje (Volpi et al., 2023). En este sentido, Perpiña-Castillo et al. (2021) desarrollaron un modelo para estimar el riesgo de ATA en la Unión Europea y Reino Unido a partir de factores biofísicos, socioeconómicos agrícolas y factores regionales. Según sus estimaciones se esperan niveles crecientes de abandono de tierras agrícolas en los próximos años y es probable que España y Polonia representen un tercio del abandono total de tierras de la Unión Europea, mientras que Francia, Alemania e Italia complementan el grupo líder (responsables en conjunto de más del 70 % del abandono total) (Perpiña-Castillo et al., 2021).

1.1.2. El abandono de los cultivos de cítricos en la Comunitat Valenciana

Durante la mayor parte del siglo XX, el motor económico que permitió la acumulación de capital para el desarrollo industrial y la modernización social de muchas zonas valencianas fue el sector citrícola (Noguera, 2010). La buena situación de los precios del mercado internacional y la ausencia de competidores generó importantes beneficios y capacidad financiera a la amplia base social formada por agricultores (Tomás-Carpí, 1977; Salom y Albertos, 2001), y rápidamente se produjo una polarización de la agricultura valenciana hacia los cítricos. De esta manera los cítricos se convirtieron en el cultivo mayoritario y en uno de los símbolos culturales de la Comunitat Valenciana.

En la actualidad, España es el mayor productor de cítricos de Europa, el sexto del mundo (FAO, 2021) y el primer exportador de cítricos para consumo en fresco. En la Comunitat Valenciana los cítricos son el cultivo mayoritario que ocupa en torno al 30% de la superficie agrícola (MAPA, 2021). La Comunitat Valenciana produce más del 50% de los cítricos españoles. Se cultivan principalmente variedades de naranjas y mandarinas que suponen una producción anual de más de 3 millones de toneladas. Según la Encuesta sobre Superficies y Rendimientos de Cultivos (ESYRCE), en 2021 la superficie dedicada al cultivo de cítricos en la Comunitat Valenciana se estimó en 160.088 ha (MAPA, 2022). La figura 1.1 muestra la distribución espacial de los cítricos en la Comunitat Valenciana. Sin embargo, en los últimos años se ha producido un abandono significativo de las explotaciones citrícolas. Desde el año 2000 hasta el 2020 se ha producido una disminución en torno al 20% de la superficie dedicada a los cítricos en la Comunitat Valenciana (MAPA, 2021). En los próximos años se espera un aumento de las zonas abandonadas debido a los cambios socioeconómicos que se están produciendo en la Unión Europea (Rounsevell et al., 2006, Terres et al. 2013).
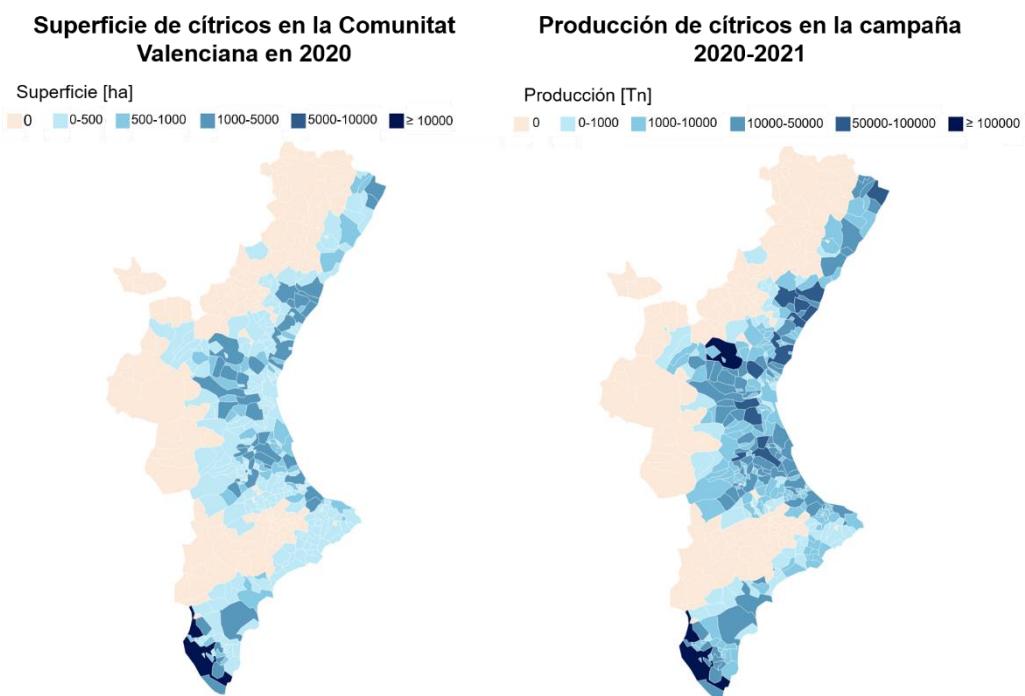
Figura 1.1. Distribución espacial de los cultivos de cítricos en la Comunitat Valenciana. Fuente: Generalitat Valenciana. Conselleria d'Agricultura, Desenvolupament Rural, Emergència Climàtica i Transició Ecològica (modificado).

1.2. ESTADO DEL ARTE

1.2.1. Antecedentes sobre la detección del abandono de tierras agrícolas utilizando datos de observación de la tierra

En este contexto, la teledetección se presenta como una potente herramienta para abordar el problema del ATA que puede reemplazar la actualización manual de mapas a través de fotointerpretación de imágenes aéreas y campañas de campo. En 2018 la Comisión Europea aprobó el uso de metodologías alternativas a la verificación in situ para el seguimiento de cultivos basadas en datos de observación de la Tierra y técnicas de aprendizaje automático (regulación 63 No. 746/2018). De esta manera se motiva a los estados miembros a utilizar datos del programa Copernicus (Sentinel-1 y Sentinel-2) para verificar las declaraciones de los agricultores a la hora de acceder a subsidios de la PAC.

La identificación de tipos de cultivos utilizando datos de observación de la tierra es una cuestión ampliamente estudiada con múltiples trabajos a lo largo de los años y algunas publicaciones recientes relacionadas con el seguimiento de las subvenciones de la PAC (ej. Sitokonstantinou et al., 2018; Campos-Taberner et al., 2019; Navarro et al., 2021; Dimov, 2022). En la Comunitat Valenciana, Amorós López et al. (2011) desarrollaron una metodología multietapa para identificar los cultivos de cítricos y actualizar el Sistema de Información Geográfica de cítricos de la Comunitat Valenciana. Sin embargo, la identificación del ATA ha sido menos estudiado desde el campo de la teledetección, ya que es un fenómeno complejo, multiescalar y con una alta variabilidad (Yin et al., 2018).

La mayoría de estudios existentes sobre la detección del ATA se centran en el uso de series temporales de imágenes Landsat (ej. Yin, H., 2018, Grădinaru et al., 2019, Prishchepov et al., 2012) y Moderate Resolution Imaging Spectroradiometer (MODIS) (ej. Löw et al., 2018, Alcántara, et al., 2012, Estel et al., 2012). Estos estudios utilizaron series temporales de imágenes de resolución media y baja para detectar el abandono de cultivos estacionales en grandes áreas que no requieren de una resolución espacial elevada. Los primeros estudios datan del año 2012 en el cual Alcantara et al. (2012) estudiaron por primera vez el potencial de los datos MODIS para identificar el abandono de tierras de cultivo. En este mismo año Prishchepov et al. (2012) estudiaron el potencial de los datos Landsat y también la influencia en la fecha de adquisición de las imágenes para identificar el ATA. Estel et al. (2015) ampliaron el estudio a otras zonas de Europa para la detectar tierras de cultivo abandonadas y recultivadas utilizando series temporales largas de MODIS. Este estudio es el primero que realiza una segmentación espacial (¿dónde?) y temporal (¿en qué periodo?) del ATA. Posteriormente Yin et al. (2018) realizaron una segmentación espacial y temporal del ATA utilizando datos Landsat. En 2018 y 2019 estas técnicas se aplicaron en otras áreas (Löw et al. 2018) y otros contextos (Grădinaru et al., 2019). Todos estos estudios se basaron en el uso de algoritmos de aprendizaje automático (árboles de clasificación y regresión, bosques aleatorios y máquinas de vectores de soporte). Las imágenes Sentinel-2 han permitido abordar esta problemática a una resolución espacial mayor. Sin embargo, los estudios sobre el ATA utilizando datos de Sentinel-2 siguen siendo escasos. Szostak et al. (2017) se centraron en la detección de sucesiones forestales de tierras previamente ocupadas por cultivos estacionales en Baja Silesia (Polonia). Ruiz et

al. (2020) abordaron el problema del abandono de diferentes cultivos en la Comunitat Valenciana sin incluir los cítricos. Por otro lado, Volpi et al. (2022) analizaron esta problemática para el cultivo de los olivos. Finalmente, Portalés-Julià et al. (2021) aplicaron series temporales para estudiar el abandono en cítricos, utilizando únicamente datos Sentinel-2 a diferencia de esta tesis que explora el uso de otras fuentes de datos. También existen experiencias previas sobre el uso de datos LiDAR para la detección del ATA. Los trabajos de Kolecka et al. (2015) y Czesak et al. (2021) utilizaron datos LiDAR para detectar estados avanzados de abandono y sucesiones forestales secundarias en paisajes altamente fragmentados que antiguamente fueron ocupados por cultivos.

De manera general los estudios sobre la identificación del ATA se han centrado en cultivos estacionales (ej. trigo, maíz, arroz, cebada) que presentan ciclos productivos anuales / bianuales (labrado, siembra, germinación, crecimiento y cosecha). La revisión realizada por Goga et al. (2019) muestra que la identificación del ATA asociado a cultivos estacionales, tierras arables y pastizales es un problema ampliamente cubierto, a pesar de la carencia de trabajo de campo de referencia. Por el contrario, existen pocos estudios centrados en la detección del abandono de cultivos permanentes de hoja perenne, como los cítricos, que pueden ser más desafiantes a la hora de identificar el ATA puesto que no presentan estos ciclos que pueden facilitar la identificación de los campos abandonados.

1.2.2. Consideraciones para la detección del estado de los cultivos de cítricos en la Comunitat Valenciana

Actualmente, en la Comunitat Valenciana, no existe una base de datos cartográfica con el detalle temático y espacial suficiente para abordar el problema del ATA. El Sistema de Información Geográfica de Parcelas Agrícolas (SIGPAC) permite identificar las parcelas de cítricos, pero no está dedicado a aportar información sobre el estado de los cultivos. Las parcelas abandonadas se registran en SIGPAC como una incidencia (código 117). Sin embargo, en la práctica se registran muy pocas parcelas abandonadas (Tomás et al., 2021). Después de un largo tiempo desde el abandono de una parcela y cuando el abandono es muy evidente en las ortofotos, algunas parcelas abandonadas se clasifican como pastizal. Sin embargo, en muchas ocasiones las parcelas abandonadas siguen clasificadas como cítricos en la base de datos. Actualmente se están introduciendo técnicas para la clasificación automática de parcelas, sin embargo, la actualización de SIGPAC se sigue realizando mayoritariamente de forma manual a través de fotointerpretación y visitas *in situ*. Además, este sistema se actualiza progresivamente de forma manual, lo que puede provocar errores y cambios no registrados hasta que se vuelve a actualizar la cartografía de la región.

Al abordar la identificación del estado de los cultivos de cítricos en la Comunitat Valenciana, cabe considerar ciertas características de la estructura agraria y del propio cultivo. La Comunitat Valenciana destaca por el pequeño tamaño de las explotaciones agrícolas, en especial el sector citrícola. El paisaje agrario se caracteriza por una alta fragmentación con un tamaño promedio de las parcelas entre 0,2 y 0,5 ha. Este aspecto es un factor limitante en relación con el tipo de datos a utilizar. Por este motivo descartamos el uso imágenes MODIS y Landsat con resoluciones espaciales de 250-500 m y 30 m, respectivamente. Por otra parte, a diferencia de los estudios previos donde se estudió el ATA asociado a cultivos estacionales, los cítricos son cultivos permanentes de hoja perenne. A diferencia de los cultivos estacionales, el cultivo de cítricos no está

sujeto a ciclos anuales/bianuales de arado, siembra, crecimiento y cosecha. El crecimiento de los cítricos es muy lento y puede durar en torno a 15 años hasta el máximo volumen de los árboles. Además, el cultivo se mantiene verde durante todo el año. Por este motivo, al ser el abandono un proceso lento cuyos cambios no se detectan de manera inmediata, la firma espectral de estos campos cambia poco y gradualmente. Estas características pueden ser un desafío de cara a la identificación del abandono de los cultivos de cítricos en la Comunitat Valenciana.

1.3. OBJETIVOS Y ALCANCE DE LA TESIS

El presente trabajo pretende dar respuesta al problema del seguimiento del abandono de cultivos de cítricos. El objetivo general de esta tesis es desarrollar métodos semiautomáticos para la identificación del estado de los cultivos utilizando diferentes fuentes de datos de observación de la Tierra. El trabajo explota el uso de imágenes ópticas Sentinel-2, WorldView-3, imágenes aéreas de muy alta resolución y datos altimétricos derivados de LiDAR y fotogrametría SfM.

Los objetivos específicos de esta investigación son los siguientes:

• Proponer un sistema de clases para el estudio del abandono de los cultivos de cítricos que permita diferenciar los principales estados del cultivo y que sea adecuado para complementar la cartografía de SIGPAC.

• Obtener datos de referencia sobre el abandono de los cultivos de cítricos, a través de campañas de campo y fotointerpretación, que permitan calibrar y validar los modelos de clasificación.

• Desarrollar procedimientos semiautomáticos para identificar el estado de los cultivos utilizando diferentes fuentes de datos de teledetección libres y comerciales.

• Analizar la capacidad de los datos multiespectrales Sentinel-2 para identificar el abandono de los cultivos de cítricos a través de una única imagen y compararlos con el uso de imágenes de mayor resolución espacial.

• Evaluar el rendimiento de la clasificación al incluir información espacial derivada de imágenes de muy alta resolución en forma de características de textura de la imagen.

• Evaluar el uso combinado de información altimétrica, espectral y textural procedente de varias fuentes de datos en el rendimiento de la clasificación.

• Analizar la capacidad de las series temporales de imágenes Sentinel-2 para identificar el abandono de los cultivos de cítricos y evaluar la transferibilidad espacial y temporal de las firmas espectrales-temporales.

   Un Plan de Seguimiento Agrícola completo está constituido por varias etapas como la extracción del área agrícola, delimitación de parcelas, clasificación de tipos de cultivos, etc. (figura 2). El alcance de este trabajo se limita a los cultivos de cítricos. A nivel funcional, este trabajo se centra en la clasificación del estado de las parcelas de cítricos previamente identificadas. Los métodos propuestos no conforman un Plan de Seguimiento Agrícola completo, sino que deben concebirse como una etapa dentro del mismo. Los pasos previos a la identificación del estado de los cultivos están disponibles a través de la base de datos de SIGPAC.

Plan de Seguimiento Agrícola

general

Extracción de la superficie agrícola

Delimitación de parcelas

Identificación / clasificación
de tipos de cultivos

**Estado de los
cultivos**

Salud y plagas

Fenología

Abandono

Estimación del
rendimiento

específico

Figura 1.2. Esquema de las etapas que constituyen un Plan de Seguimiento Agrícola completo.

1.4. ESTRUCTURA DE LA TESIS

Esta tesis se presenta como un compendio de artículos. La memoria de tesis está estructurada en 8 capítulos. El presente capítulo (capítulo 1) introduce la problemática del abandono de los cultivos de cítricos en la Comunitat Valenciana, revisa el estado del arte sobre la detección del abandono de tierras a través de datos de observación de la Tierra y define los objetivos, alcance y estructura de la tesis. Del capítulo 2 al 6 se presentan cinco estudios dedicados a la identificación del estado de los cultivos de cítricos a través de diferentes metodologías y tipos de datos. Estos capítulos contienen la versión de autor de cinco artículos científicos, los cuales han sido enviados a diferentes revistas de reconocido prestigio y sometidos a revisión por pares. La tabla 1.1 detalla cada uno de estos artículos y el estado en el que se encuentran en el momento de finalizar la presente memoria de tesis. En el capítulo 7 se incluye una discusión general de los resultados obtenidos durante estas investigaciones. Finalmente, en el capítulo 8 se analizan los avances producidos durante esta tesis, las conclusiones obtenidas y las futuras líneas de investigación propuestas en base a la experiencia adquirida.

Tabla 1.1. Relación de capítulos de la tesis y su respectivo artículo científico.

Capítulo	Título del artículo	Autores	Revista	Indexación	Estado
Cap. 2	Comparison of Sentinel-2 and High-Resolution Imagery for Mapping Land Abandonment in Fragmented Areas	Morell-Monzó, S.; Estornell, J.; Sebastiá-Frasquet, M.T.	Remote Sensing ISSN: 2072-4292	Q1 [JCR LE-Geosciences, Multidisciplinary] Impact factor: 5,349	Publicado en: (2020) Vol. 12(12) 2062.
Cap. 3	Land Use Classification of VHR Images for Mapping Small-Sized Abandoned Citrus Plots by Using Spectral and Textural Information	Morell-Monzó, S.; Sebastiá-Frasquet, M.T. Estornell, J.	Remote Sensing ISSN: 2072-4292	Q1 [JCR LE-Geosciences, Multidisciplinary] Impact factor: 5,349	Publicado en: (2021) Vol. 13(4) 681.
Cap. 4	Cartografía del abandono de cultivos de cítricos mediante el uso de datos altimétricos: LiDAR y fotogrametría SfM	Morell-Monzó, S.; Sebastiá-Frasquet, M.T.; Estornell, J.	Revista de Teledetección ISSN: 1133-0953	Q3 [SJR Geography, Planning and Development] CiteScore: 1,3	Publicado en: (2022) Vol. 59, 49-60.
Cap. 5	Assessing the capabilities of high-resolution spectral, altimetric, and textural descriptors for mapping the status of citrus parcels	Morell-Monzó, S; Estornell, J.; Sebastiá-Frasquet, M.T.	Computer and Electronics in Agriculture ISSN: 0168-1699	Q1 [JCR AH-Agriculture, Multidisciplinary] Impact factor: 6,757	Publicado en: (2023) Vol. 204, 107504
Cap. 6	Detecting abandoned citrus crops using Sentinel-2 time series. A case study in the Comunitat Valenciana region (Spain)	Morell-Monzó, S.; Sebastiá-Frasquet, M.T.; Estornell, J.; Moltó, E.	ISPRS Journal of Photogrammetry and Remote Sensing ISSN: 0924-2716	Q1 [JCR LE-Geosciences, Multidisciplinary] Impact factor: 11,774	En revisión

1.5. REFERENCIAS

Alcántara, C., Kuemmerle, T., Prishchepov, A.V., Radeloff, V.C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sensing of Environment, 2012, 124, 334–347. https://doi.org/10.1016/j.rse.2012.05.019

Amorós López, Izquierdo Verdiguier, E., Gómez Chova, L., Muñoz Marí, J., Rodríguez Barreiro, J.Z., Camps Valls, G., Calpe Maravilla, J. Land cover classification of VHR airborne images for citrus grove identification. ISPRS Journal of Photogrammetry and Remote Sensing, 2011, 66, 115-123. https://doi.org/10.1016/j.isprsjprs.2010.09.008

Campos-Taberner, M., García-Haro, F.J., Martínez, B., Sánchez-Ruíz, S., Gilabert, M.A. A copernicus Sentinel-1 and Sentinel-2 classification framework for the 2020+ European common agricultural policy: A case study in València (Spain). Agronomy, 2019, 9(9), 556. https://doi.org/10.3390/agronomy9090556

Cerdá, A., Ackerman, O., Terol, E., Rodrigo Comino, J. Impact of Farmland Abandonment on Water Resources and Soil Conservation in Citrus Plantations in Eastern Europe. Water, 2019, 11(4), 824. https://doi.org/10.3390/w11040824

Cocca, G., Sturaro, E., Gallo, L., Ramanzin, M. Is the abandonment of traditional livestock farming systems the main driver of mountain landscape change in Alpine areas? Land Use Policy, 2012, 29, 878–886. https://doi.org/10.1016/j.landusepol.2012.01.005

Corbelle-Rico, E.; Crecente-Maseda, R. Estudio da evolución da superficie agrícola na comarca da Terra Chá a partir de fotografía aérea histórica e mapas de usos, 1956-2004. Recursos Rurais 2018, 4, 57–65. https://doi.org/10.15304/rr.id5312

Czesak, B., Różycka-Czas, R., Salata, T., Dixon-Gough, R., Hernik, J. Determining the Intangible: Detecting Land Abandonment at Local Scale. Remote Sensing, 2021, 13(6), 548 1166. https://doi.org/10.3390/rs13061166

Dimov, D. Classification of remote sensing time series and similarity metrics for crop type verification. Journal of Applied Remote Sensing, 2022, 16(2). https://doi.org/10.1117/1.JRS.16.024519

Estel, S., Kuemmerle, T., Alcántara, C., Levers, C., Prishchepov, A.V., Hostert, P. Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. Remote Sensing of Environment, 2015, 163, 312–325. https://doi.org/10.1016/j.rse.2015.03.028

FAO. Citrus Fruit Fresh and Processed Statistical Bulletin 2020. Citrus Fruit Statistical Compendium 2020; FAO: Rome, Italy, 2021. http://www.fao.org/3/cb6492en/cb6492en.pdf

FAO-FAOSTAT. Methods & Standards. Disponible en: https://www.fao.org/statistics/methods-and-standards/en/ (mayo 2021).

Goga, T., Feranec, J., Bucha, T., Rusnák, M., Sačkov, I., Barka, I., Kopecká, M., Papčo, J., Oťaheľ, J., Szatmári, D., Pazúr, R., Sedliak, M., Pajtík, J., & Vladovič, J. A Review of the Application of Remote Sensing Data for Abandoned Agricultural Land Identification with Focus on Central and Eastern Europe. Remote Sensing, 2019, 11(23), 2759. https://doi.org/10.3390/rs11232759

Grădinaru, S.R., Kienast, F., Psomas, A. Using multi-seasonal Landsat imagery for rapid identification of abandoned land in areas affected by urban sprawl. Ecological Indicartors, 2019, 96, 79–86. https://doi.org/10.1016/j.ecolind.2017.06.022

Keenleyside, C., Tucker, G. M. Farmland Abandonment in the EU: an Assessment of Trends and Prospects. Report prepared for WWF. Institute for European Environmental Policy, London, United Kingdom, 2010.

Kolecka, N., Kozak, J., Kaim, D., Dobosz, M., Ginzler, C., Psomas, A. Mapping Secondary Forest Succession on Abandoned Agricultural Land with LiDAR Point Clouds and Terrestrial Photography. Remote Sensing, 2015, 7(7), 8300-8322. https://doi.org/10.3390/rs70708300

Kosmas, C., Kairis, O., Karavitis, C., Acikalin, S., Alcalá, M., Alfama, P., Atlhopheng, J., Barrera, J., Belgacem, A., Solé-Benet, A., et al. An exploratory analysis of land abandonment drivers in areas prone to desertification. CATENA 2015, 128, 252–261. https://doi.org/10.1016/j.catena.2014.02.006

Leal Filho, W., Mandel, M, Al-Amir, A.Q., Feher, A., Chiappetta Jabbour, C.J. An assessment of the causes and consequences of agricultural land abandonment in Europe. International Journal of Sustainable Development & World Ecology, 2017, 24, 554-560. https://doi.org/10.1080/13504509.2016.1240113

Löw F., Prishchepov, A.V., Waldner, F., Dubovyk, O., Akramkhanov, A., Biradar, C., Lamers, J.P.A. Mapping Cropland Abandonment in the Aral Sea Basin with MODIS Time Series. Remote Sensing, 2018, 10(2), 159. https://doi.org/10.3390/rs10020159

Ministerio de Agricultura, Pesca y Alimentación. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2020; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2021. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Ministerio de Agricultura, Pesca y Alimentación. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2021; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2022. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Navarro, A., Silva, I., Catalao, J., Falcao, J. An operational Sentinel-2 based monitoring system for the management and control of direct aids to the farmers in the context of the Common Agricultural Policy (CAP): A case study in mainland Portugal. International Journal of Applied Earth Observation and Geoinformation, 2021, 103, 102469. https://doi.org/10.1016/j.jag.2021.102469

Noguera, J. Viabilidad y competitividad del sistema citrícola valenciano. Boletín de la Asociación de Geógrafos Españoles, 2010, 52, 81-99. ISSN: 0212-9426.

Perpiña-Castillo, C., Kavalov, B., Diogo, V. Agricultural Land Abandonment in the EU within 2015–2030. Technical Report JRC113718. European Commission Joint Research Centre, Brussels, Belgium, 2018.

Perpiña-Castillo C., Jacobs-Crisioni, C., Diogo, V., Lavalle, C. Modelling agricultural land abandonment in a fine spatial resolution multi-level land-use model: An application for the EU. Environmental Modelling & Software, 2021, 136, 104946. https://doi.org/10.1016/j.envsoft.2020.104946

Pointereau, P., Coulon, F., Girard, P., Lambotte, M., Stuczynski, T., Sánchez Ortega, V., del Rio, A. Analysis of Farmland Abandonment and the Extent and Location of Agricultural Areas that are Actually Abandoned or are in Risk to be Abandoned. Technical Report JRC46185. European Commission Joint Research Centre: Brussels, Belgium, 2008.

Portalés-Julià E., Campos-Taberner, M., García-Haro, F.J., Gilabert, M.A. Assessing the Sentinel-2 Capabilities to Identify Abandoned Crops Using Deep Learning. Agronomy, 2021, 11(4), 654. https://doi.org/10.3390/agronomy11040654

Prishchepov, A.V. Agricultural Land Abandonment. Oxford Bibliographies. Environmental Sciences. Oxford University Press 2020. https://doi.org/10.1093/obo/9780199363445- 0129

Prishchepov, A.V., Radeloff, V.C., Dubinin, M., Alcantara, C. The effect of Landsat ETM/ETM image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Remote Sensing of Environment, 2012, 126, 195–209. https://doi.org/10.1016/j.rse.2012.08.017

Rounsevell, M.D.A., Reginster, I., Araújo, M.B., Carter, T.R., Dendoncker, N., Ewert, F., House, J.I., Kankaanpää, S., Leemans, R., Metzger, M.J., Schmit, C., Smith, P., Tuck, G. A coherent set of future land use change scenarios for Europe. Agriculture, Ecosystems, and Environment, 2006, 114, 57–68. https://doi.org/10.1016/j.agee.2005.11.027

Ruiz, L.A., Almonacid-Cabaliler, J., Crespo-Peremarch, P., Recio, J.A., Pardo-Pascual, J.E., Sánchez-García, E. Automated classification of crop types and condition in a Mediterranean area using a fine-tuned convolutional neural network. The International Archcives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B3, 2020, 1061–1068. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1061-2020

Salom, J., Albertos, J.M. El modelo de desarrollo de la Comunitat Valenciana. En La Comunidad Valenciana en la Europa de las Regiones, 1ª ed. Romero, J.S.J., Vera F., Eds. Ariel Geografía: Valencia, España, 2001.

Sitokonstantinou, V., Papoutsis, I., Kontoes, C., Lafarga Arnal, A., Armesto Andrés, A.P., Garraza Zurbano, J.A. Scalable parcel-based crop identification scheme using Sentinel-2 data time-series for the monitoring of the common agricultural policy. Remote Sensing, 2018, 10, 911. https://doi.org/10.3390/rs10060911

Strijker, D. Marginal lands in Europe - causes of decline. Basic and Applied Ecology 2005, 6, 99–106. https://doi.org/10.1016/j.baae.2005.01.00

Subedi, Y.R., Kristiansen, P., Cacho, O. Drivers and consequences of agricultural land abandonment and its reutilization pathways: A systematic review. Environmental Development, 2021, 42, 100681. https://doi.org/10.1016/j.envdev.2021.100681

Szostak, M., Hawryło, P., Piela, D. Unsing of Sentinel-2 images for automation of forest succession detection. European Journal of Remote Sensing, 2018, 51, 142-149. https://doi.org/10.1080/22797254.2017.1412272

Terres, J.M, Nisini, L., Anguiano, E. Assessing the risk of farmland abandonment in the EU. JRC Scientific and Policy Reports. Publications Office of the European Union. https://doi.org/doi:10.2788/81337

Terres, J.-M., Scacchiafichi, L. N., Wania, A., Ambar, M., Anguiano, E., Buckwell, A., Coppola, A., Gocht, A., Källström, H. N., Pointereau, P., Strijker, D., Visek, L., Vranken, L., Zobena, A. Farmland abandonment in Europe: Identification of drivers and indicators, and development of a composite indicator of risk. Land Use Policy, 2015, 49, 20–34. https://doi.org/10.1016/j.landusepol.2015.06.009

Tomás-Carpí, J.A. La economía valenciana: Modelos de interpretación. En Contribución Invisible de las Mujeres en la Economía: El Caso Específico del Mundo Rural, 1ª ed. Vera A., Ed.; Instituto De La Mujer. Madrid, España, 1977.

Tomás, J.V., Calafat, C., Gallego, Á., Martínez, V. Estimación de la superficie agrícola abandonada en cultivos permanentes: registros y teleobservación. XIII Congreso de Economía Agroalimentaria. Universidad Politécnica de Cartagena. https://doi.org/10.31428/10317/10648

Volpi, I., Marchi, S., Petacchi, R., Hoxha, K., Guidotti, D. Detecting olive groove abandonment with Sentinel-2 and machine learning: The development of a web-based tool for land management. Smart Agricultural Technology, 2023, 3, 100068. https://doi.org/10.1016/j.atech.2022.100068

Yin, H., Prishchepov, A.V., Kuemmerle, T., Bleyhl, B., Buchner, J., Radeloff, V.C. Mapping agricultural land abandonment from spatial and temporal segmentation of Landsat time series. Remote Sensing of Environment, 2018, 210, 12–24. https://doi.org/10.1016/j.rse.2018.02.050

CAPÍTULO 2

Comparison of Sentinel-2 and High-Resolution Imagery for Mapping Land Abandonment in Fragmented Areas

2.1. INTRODUCTION

Agricultural land abandonment is considered a global issue with relevant environmental and socio-economic consequences (MacDonald et al., 2000; Kosmas et al., 2015). The concept of land abandonment (or abandoned land) is complex and not unanimous across all scientific and legal sources. The abandonment of agricultural activity does not always generate a situation of use absence. Sometimes, there is a change in land use (for example, urbanization processes or forestry plantations). Thus, it is necessary to differentiate between the broadest concept of abandonment of agricultural activity and the more specific concept of land abandonment, which is reserved for those cases without any land use or economic activity (Corbelle-Rico & Cresente-Maseda, 2018; Baudry, 1991; Pinto-Correia, 1991). In this study, we use the term land abandonment as the equivalent of agricultural land abandonment or farmland abandonment.

In some European countries, land abandonment is documented to have occurred at other times in history, in the Middle Ages, during the nineteenth century due to industrialization, and after World War II (Gellrich et al., 2007). For the 2015–2030 period, it is estimated that around 11% (more than 20 million hectares) of agricultural land in the European Union (EU) is at risk of abandonment (Perpiña-Castillo et al., 2018). In the European Mediterranean region, land abandonment is a common issue (Rey-Benayas, 2007), with variable social, environmental and economic contexts (Romero-Díaz & Martínez-Hernández, 2014). Mediterranean coastal areas have been one of the most significantly changing areas of Europe due to rapid urbanization, industrialization and tertiarization process (Árgyelán, 2015). A good example of abandonment due to these processes is the Valencian citrus industry. Spain is the most important citrus producer in the EU and the fifth highest producer in the world (IVIA, 2020), and the Valencian Region (East Spain) is its most important producer. Citruses are the main crop in this region, with an area of 165,000 ha, about 60% of the national area of this crop (MAPA, 2017). It produces more than 3 million Tn of mandarins and oranges annually (50% and 45% of the Spain's citrus industry, respectively) (IVIA, 2020). Other important crops in the area are olives (91,843 ha), almonds (91,291 ha) and vineyards (64,407 ha) with minor extension (Generalitat Valenciana, 2019). However, in the areas with freshwater availability and favourable weather conditions, citrus is generally the only crop because it has higher profitability than other crops.

For most of the twentieth century, the economic engine that allowed the accumulation of capital for the industrial development and social modernization of many Valencian areas was orange cultivation (Noguera, 2010). The good situation of international market prices and the absence of competitors brought significant benefits and financial capacity to the extensive social base formed by farmers (Tomás-Carpi, 1997; Piqueras, 1999; Salom & Albertos, 2001), and there was a quick polarization of Valencian agriculture towards citrus. This situation made citrus become the majority crop and the symbol of the Valencian Region. However, in recent years, there has been a massive abandonment of these agricultural holdings due to internal and external causes. There was a significant decrease of 14% in the citrus-cultivated area from 2008 to 2017 (from 190,000 ha to 165,000 ha) (MAPA, 2017). The main internal causes are the emergence of more profitable land uses, such as urban use, and the strong competitiveness of foreign producers (outside the EU), so many land owners stop farming their holdings and abandon their plots. In the coming years, an increase in

15

abandoned areas is expected due to the socio-economic changes that are taking place in the EU (Rounsevell et al., 2018). Therefore, it is necessary to obtain detailed, descriptive and quantitative information to understand the abandonment consequences in terms of the landscape, the environment and socioeconomics (Verbug et al., 2006). Land abandonment can generate negative and positive impacts on the environment, depending on the type of crop, the weather, the characteristics of agricultural practices, etc. Among the negative environmental impacts, we can highlight the development of successional vegetation that is highly flammable and increases the risk of forest fires (Dubinin et al., 2010). This vegetation also increases the pressure of weeds, pests and pathogens in nearby agricultural fields (Smelansky et al., 2003). In addition, land abandonment may promote soil erosion (Ruiz-Flano et al., 1992) and the loss of landscapes of high cultural value and agrobiodiversity (Fischer et al., 2012), and promote soil salinization (Penov, 2004). On the positive side, agricultural land abandonment can provide environmental services such as carbon sequestration and the mitigation of anthropogenic CO_2 emissions due to the vegetation recovery and increase in soil organic matter (Novara et al., 2017), which also helps in reducing soil erosion (Cerdà et al., 2019). This is especially true in citrus land abandonment and with other crops with flood irrigation and organic matter removal. In these cases, land abandonment can help reduce soil erosion (Rey-Benayas & Bullock, 2012).

The effects of land abandonment on the environment and society are complex and need to be studied at different scales. We face current challenges that require the efficient management of agricultural and wilderness areas (Rey-Benayas & Bullock, 2012). This involves food provisioning at efficient scales and maintaining a network of rewild areas that offer trade-offs between food provisioning, ecosystem services and biodiversity conservation (Navarro & Pereira, 2016). In this sense, remote sensing is a powerful tool that allows us to manage these problems (land use and land cover change) from a spatial and temporal point of view at different scales.

Land abandonment and vegetation changes have been studied using remote sensing techniques, but to our knowledge, the abandonment of citrus lands has not been studied. Previous studies have tried to develop methods to estimate citrus production using Landsat images. Although they have found positive correlations between the Landsat results and manual censuses, there are serious discrepancies due to the inability to detect abandoned and unproductive plots using Landsat images (Shrivastava & Gebelein, 2007). Some previous studies about land abandonment have focused on the use of MODIS (Moderate Resolution Imaging Spectroradiometer) imagery to map land abandonment in large areas, which does not require a high spatial resolution (Löw et al., 2018; Alcántara et al., 2012; Estel et al., 2015). However, most studies have focused on Landsat imagery (Dara et al., 2018; Müller et al., 2013; Yin et al., 2018; Kuemmerle et al., 2008; Grădinaru et al., 2019; Prishchepov et al., 2012; Baumann et al., 2011). The spatial resolution of Landsat images also has serious limitations in very heterogeneous areas with high spatial fragmentation. In this sense, Sentinel-2 (with spatial resolutions of 10 m and 20 m) and high-resolution imagery can provide a solution to these limitations. Sentinel-2 use for mapping land abandonment is poorly exploited. Previous studies only differentiate between active crops and highly advanced forest successions (Szostak et al., 2017; Kanjir et al., 2018; Proisy et al., 2018).

In general, previous remote sensing studies about land abandonment have been based on two types of approaches. Firstly, for crops with clearly differentiated phenological

states throughout the year, and crops with fallow seasons, the multi-temporal-based approach allows the detection of anomalies in phenological dynamics (Alcántara et al., 2012; Estel et al., 2015; Yin et al., 2018; Prishchepov et al., 2012; Kanijir et al., 2018). However, this is not characteristic of citrus, which have slow phenological dynamics. Secondly, the mono-temporal approach is based on the detection of succession vegetation that invades abandoned crops (Szostak et al., 2017). In this study, the mono-temporal approach was chosen, mainly because the characteristics of citrus, but it has other advantages, i.e., it only requires one field campaign for validation, and processing is more efficient because it uses only one image. The success of this strategy depends on the capacity of the sensor and the classification algorithm in capturing differences in characteristics between wild vegetation and active crops.

There are different algorithms for classifying images from remote sensors. These methods range from unsupervised algorithms (i.e., ISODATA or K-means) to parametric supervised algorithms (i.e., maximum likelihood) and machine learning algorithms such as Artificial Neural Networks (ANN), k-Nearest Neighbors (k-NN), Decision Trees (DT), Support Vector Machines (SVM) and Random Forests (RF) (Noi & Kappas, 2017). Machine learning algorithms offer the potential for effective and efficient classification, as they are capable of handling high dimensional data and modelling complex nonlinear relationships (Maxwell et al., 2018). Numerous studies have been conducted to find the best classification algorithm for land use and cover studies. However, their conclusions are quite different for different contexts. Despite this, SVM and RF are usually the most accurate algorithms and are insensitive to noise and overtraining (Noi & Kappas, 2017).

The aim of this work is to compare the performance of Sentinel-2 images and high-resolution airborne images for the mapping of abandoned agricultural plots in a highly heterogenous and fragmented area. We applied a supervised classification based on Random Forest algorithms, and three classification models were compared: Model 1 uses the 10 m Sentinel-2 bands; Model 2 uses the 10 m and 20 m Sentinel-2 bands; and Model 3 uses the Red, Green, Blue and near infrared bands of airborne images. We mapped the recent land abandonment to differentiate between two types of land cover with very similar spectral responses.

2.2. DATA AND METHODS

2.2.1. Study area

The study area is Oliva, a coastal city in the Valencian Region (Spain) (Figure 2.1). In this area, various land uses—such as urban, agricultural and forestry—coexist in a reduced geographic area. The study area is 4,616 ha, and the agricultural land use is approximately 40%. The relief of the area is flat, with an average slope of 2%. There is a gradient from the coast, which is characterized by young and little-evolved soils associated with old sandbanks that closed primitive lagoons, to the inland alluvial zone, which is characterized by fluvial materials. Alluvial soils are highly permeable and have higher agronomic capacity (Viñals, 1995; Viñals, 1996).
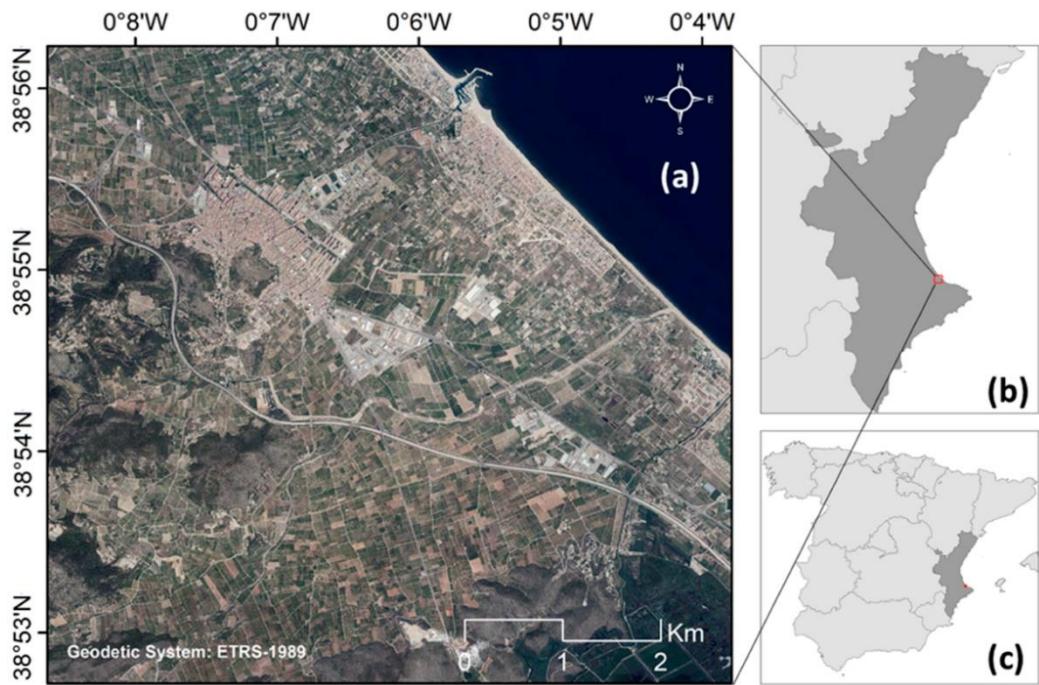
Figure 2.1. Location of the study area (a) in the Valencian Region (b), Spain (c).

The main crop in this area is citrus, which occupies more than 95% of the cultivated agricultural area. The average size of the agricultural plots is 0.38 ha ± 0.27, and it is representative of many municipalities in the Valencian Region, which had maximum production in the late 1970s and are currently experiencing massive land abandonment.

2.2.2. Land abandonment process in citrus

Three types of plots were classified according to the type of coverage found in the study area (Figure 2.2): (a) Not in production—these are plots where bare soil occupies most of the surface of the plot. These plots are in replacement periods. In this class, trees may be planted but are generally not yet productive. (b) In production—these are plots with citrus productive cultivation. In this class, the citrus crop occupies most of the surface of the plot. (c) Abandoned—these are abandoned plots dominated by wild vegetation.
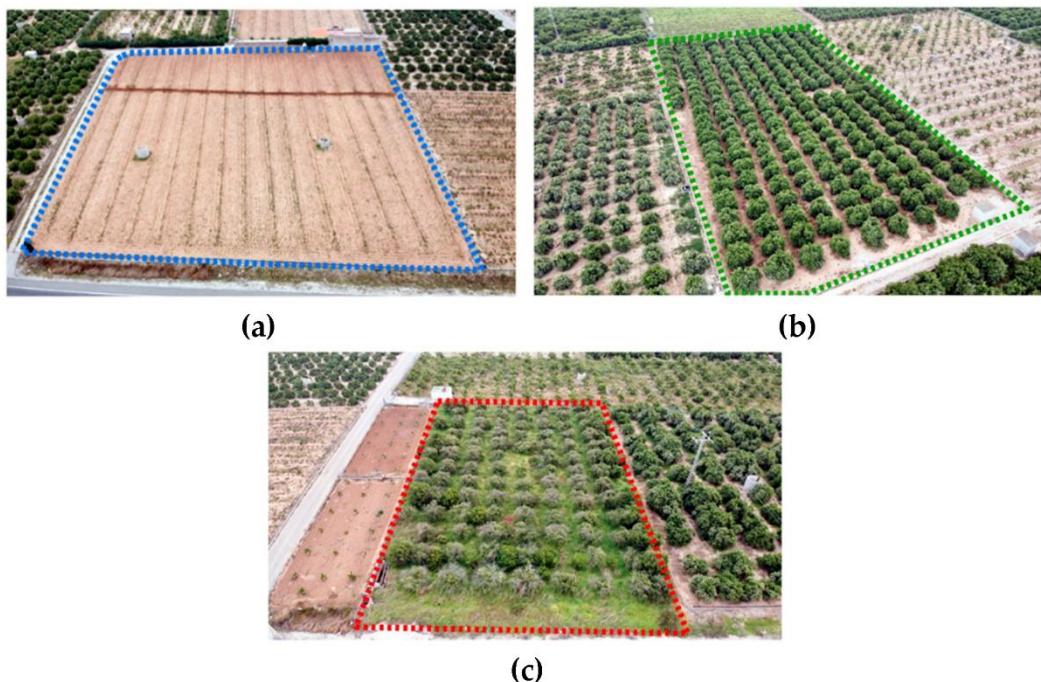
Figure 2.2. Types of plots classified according to the type of coverage: (a) not in production plots, (b) in productions plots and (c) abandoned plots.

In the study area, active crops coexist with several herbaceous species that grow under the tree layer formed by orange trees, although these herbs are usually controlled by farmers. Differences in environmental and soil moisture generate different compositions of vegetation in winter and summer. In winter, species such as *Oxalis pes-caprae* and *Hordeum murinum* predominate, while in summer, species like *Portulaca oleracea* and *Setaria verticillata* appear.

The ecological succession of an abandoned plot consists of the progressive replacement of the crop by wild vegetation of the area. In the first stage of abandonment, there is a growth of this herbaceous vegetation that extends progressively in the plot. This process occurs from the first to the fourth year after abandonment and is visible by around the third year. The type of the new vegetation and its growth depend on the characteristics of the soil, although species such as *Asparagus acutifolius* and *Imperata cylindrica* are common. In the second stage of abandonment, from the fourth or fifth year, the crop loses vigour (there is a decrease in leaf density and greenness) due to the lack of irrigation and the wild vegetation height and cover increase. Finally, the death of the trees can occur, and the wild vegetation occupies the entire plot, even colonizing the trees' remains. This process is completed by the 10th year, approximately. The height of the successive vegetation depends on the type of species that appear, although generally, by the 10th year of abandonment, the vegetation measures around 1 m in height.

In this last stage, shrub species appear, such as *Rubus ulmifolius* and *Pistacia lentiscus*. *Olea europea* and *Chamaerops humilis* may appear occasionally, and exotic species such as *Arundo donax* may appear in areas with water.

2.2.3. Data and processing

The image geoprocessing was programmed in an R environment (R Core Team, 2019) using the raster (Hijmans, 2020) and rgdal (Bivard et al., 2020) packages.

A Sentinel-2 image acquired on 11 May 2019 with processing level L1C was downloaded from the Copernicus Open Access Hub: https://scihub.copernicus.eu/dhus/#/home. The Sentinel-2 image was processed as follows (Figure 2.3). First, it was atmospheric corrected using the sen2cor algorithm (Richter et al., 2012). Then, bands of 20 m resolution were resampled to 10 m. Finally, four spectral indices were calculated from these images: the Enhanced Vegetation Index (EVI) (Huete, 1994), Thiam's Transformed Vegetation Index (TTVI) (Thiam, 1998), and Normalized Difference Moisture Indices, which use Sentinel-2 short-wave infrared or SWIR (Bands 11 and 12 respectively) (NDMI1 and NDMI2) (Wilson & Shader, 2002):

$$EVI = \frac{(NIR - R)}{(NIR + 6R - 7.5B + 1)} \tag{2.1}$$

$$TTVI = \sqrt{|\frac{(NIR - R)}{(NIR + R)}| + 0.5} \tag{2.2}$$

$$NDMI = \frac{(NIR - SWIR)}{(NIR + SWIR)} \tag{2.3}$$

where NIR is the near infrared band, R is the red band, B is the blue band and $SWIR$ is the short-wave infrared band. The bands of each image are detailed in Table 2.1.
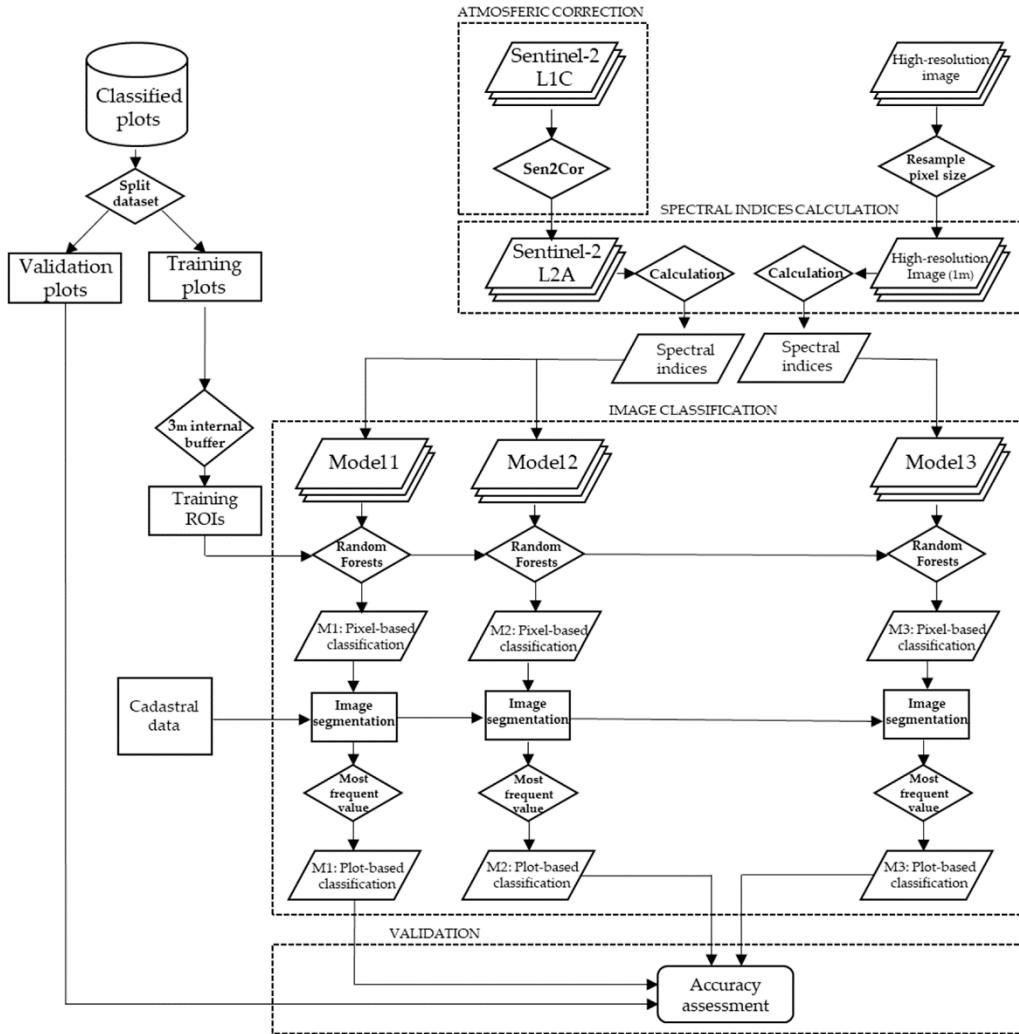
Figure 2.3. Chart describing the methodology of the experiment.

Table 2.1. Comparison between the MultiSpectral Instrument bands of Sentinel-2 and Qioptic Vexcel HR Digaron bands of UltraCam Eagle.

Chanel	Sentinel-2 Imagery (MultiSpectral Instrument)			High-resolution Airborne Imagery (UltraCam Eagle)		
	Band	Central Wavelength (nm)	Spatial Resolution (m)	Band	Central Wavelength (nm)	Spatial Resolution (m)
Blue	2	490	10	3	430	0.25
Green	3	560	10	2	530	0.25
Red	4	665	10	1	620	0.25
Red edge 1	5	705	20			
Red edge 2	6	740	20	4	720	0.25
Red edge 3	7	783	20			
NIR	8	842	10			
Red edge 4	8A	865	20			
SWIR 1	11	1610	20			
SWIR 2	12	2190	20			

These three spectral indices were selected to capture differences in characteristics between not in production, in production and abandoned plots. The TTVI is a transformation of the Normalized Difference Vegetation Index (NDVI) that avoids

21

negative values but has no technical differences from the NDVI (Silleos et al., 2006). This index is sensitive to the vegetation greenness and the biomass amount. However, dense vegetation can saturate this spectral index. The EVI allows the correction of this saturation effect (Huete, 1999). The NDMI is sensitive to soil and vegetation moisture (Wilson & Shader, 2002); it was added to capture the moisture differences that may exist between irrigated and abandoned plots (non-irrigated). The possibility of adding other spectral indices such as the Soil-Adjusted Vegetation Index (SAVI) (Huete, 1988) and Green Normalized Difference Vegetation Index (GNDVI) (Gitelson et al., 1996) was studied, but this option was rejected due to the high correlation of these indices with the TTVI. Adding these indices could destabilize the classification model. The Spearman's correlations between the indices used were as follows: ρTTVI-EVI: 0.11, ρEVI-NDMI: 0.14 and ρTTVI-NDMI: 0.58.

High-resolution airborne images were downloaded from the Valencian Spatial Data Infrastructure (IDEV): http://www.idev.gva.es/es. These images were taken on May 14, 2019 with an UltraCam Eagle UC-E-1-50016095-f80 camera from the company Vexcel Imaging GmbH © with a Qioptic Vexcel HR Digaron sensor. The Valencian Cartographic Institute (ICV) provides an annual image of the entire Valencian Region with a spatial resolution of 0.25 m and four spectral bands (Table 2.1). These images were geometrically corrected and radiometrically calibrated by Vexcel Imaging GmbH ©. According to the company report, the radiometric calibration was based on a series of 50 flat field images for each aperture size and sensor. The flat field was illuminated by eight normal light lamps with known spectral illumination curves. These images were used to calculate the specific sensitivity of each pixel to compensate for local as well as global variations in sensitivity. Sensitivity tables are calculated for each sensor and aperture setting, and applied during the radiometric calibration process for the images. The results of the bundle adjustment showed an RMSE value lower than 4 cm for the X, Y and Z coordinates using a set of 22 points according to the calibration report of the company.

Seven images were downloaded and merged to create a single mosaic covering the study area. The mosaic image was resampled to a resolution of 1 m to reduce the computational requirements and the pixels' spectral variability within each parcel. Additionally, two spectral indices were calculated: the EVI (Huete, 1994) (Equation 2.1) and TTVI (Thiam, 1998) (Equation 2.2). The spectral indices calculated from the high-resolution image cannot be compared directly with the indices of the Sentinel-2 images due to the differences in the wavelengths of their spectral bands. The most important difference between both sensors is in the NIR band, which, in fact, is more similar to the Vegetation Red Edge of Sentinel-2 (Table 2.1).

The ground truth data were collected during a field campaign conducted between 11 and 14 July 2019. A set of 240 plots was selected (80 not in production, 80 in production and 80 abandoned). Stratified random sampling with proportional distribution was performed. The original dataset was randomly segmented using 144 training plots and 96 validation plots. The same number of plots for each class was selected for training (48 plots of each class) and validation (32 plots of each class) (Figure 2.4).
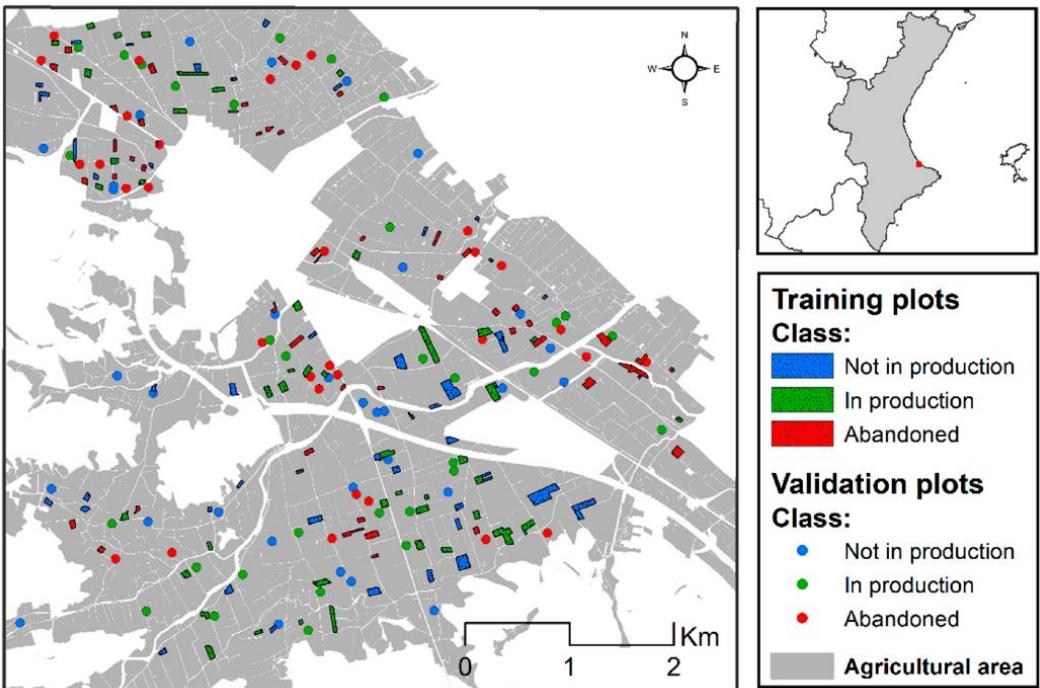
Figure 2.4. Location of parcels surveyed during ground measurements in July 2019. These data were used to create the training dataset and assess classifier performance.

The regions of interest (ROIs) were created by applying an internal buffer of 3 m on the training plots to avoid the vegetation that occupied the plot boundaries. Each pixel value within these plots was extracted from Sentinel-2 and airborne bands and spectral index to create the set of pixels to train the Random Forests classifiers. Three models were analysed using the coverage type as the dependent variable. The independent variables of each model were grouped according to the spatial and spectral resolution of the bands of each sensor. Model 1: The pixel values of Bands 2, 3, 4, 8, EVI and TTVI vegetation indices retrieved from the Sentinel-2 images were used as independent variables. These variables have a spatial resolution of 10 m. Model 2: The pixel values of Bands 2, 3, 4, 5, 6, 7, 8, 8A, 11 and 12; the vegetation indices EVI and TTVI; and the moisture indices NDMI1 (Band 11) and NDMI2 (Band 12) were used as independent variables. These variables have a minimum spatial resolution of 20 m. Model 3: The Red, Green, Blue and Near-Infrared bands and the vegetation indices EVI and TTVI retrieved from airborne image were used as independent variables. These variables have a spatial resolution of 1 m.

2.2.4. Classification algorithm

Pixel-based classification was done using Random Forests, which is a non-parametric algorithm commonly used in remote sensing for supervised classifications. This algorithm was proposed by Breiman (1999).

The Random Forests classifier consists of an ensemble of tree classifiers $\{h(x, \Theta k), k = 1, ..., \}$ where the $\{\Theta k\}$ are independent, identically distributed random vectors and X is an input pattern (Breiman, 2001). In training, the algorithm creates multiple CARTS-like trees (Breiman et al., 1984), each one trained on a bootstrapped sample of the original training data, and searches only across a randomly selected subset

23

of the input variables to determinate a split (for each node) (Gislason et al., 2006). Bootstrap Aggregation (or Bagging) is the method of selecting a training dataset for each tree by randomly drawing with N examples, where N is the size of the original training set (Breiman, 1996).

The design of a decision tree required the choice of an attribute selection measure and pruning method. The Random Forests classifier uses the Gini Index as an attribute selection measure, which measures the impurity of an attribute with respect to the classes (Pal, 2005). For a given training set T, selecting one case (training pixel) at random and saying that it belongs to some class C_i, the Gini Index can be written as:

$$\sum \sum_{j \neq i} (\int (C_i, T) \, / \, |T|)(\int (C_j, T) / \, |T|)$$

(2.4)

where $\int (C_j, T) / \, |T|$ is the probability that the selected case belongs to class Ci.

The Mean Decrease in Gini (MDG) was used as a measure of the importance of the variables of each model. The Gini importance measures the average gain of purity by splits of a given variable. Therefore, if the variable is useful, it tends to split mixed labeled nodes into pure single class nodes. This means that a variable is important if it is capable of differentiating a class by itself or, in other words generating a single class node. This measure of importance should not be confused with the Mean Decrease Accuracy, which corresponds to the decrease in the average accuracy of a model when a variable is permuted.

The number of variables in Bagging and the number of trees grown are user-defined parameters. As the number of trees increases, the generalization error always converges, and overfitting is not a problem because of the Strong Law of Large Numbers. Despite this, the number of trees grown must be limited so as not to excessively increase computing time: $cTv(MN \log(N))$, where c is a constant, T is the number of trees in the ensemble, M is the number of variables and N is the number of samples in the training dataset (Gislason et al., 2006).

In this work, the Random Forest algorithm was applied using the `randomForest` function (Liaw & Wiener, 2002) from the `caret` package (Kuhn et al., 2019) in the R environment (R Core Team, 2019). Model 1 has six input variables and was trained with pixels covered by 144 training ROIs. Model 2 has 14 input variables and was trained with pixels covered by 144 training ROIs. Model 3 has six input variables and was trained with pixels covered by 144 training ROIs. The classifiers were optimized with 100 decision trees and were trained with the same number of pixels in each category. The generalization errors of the models converged around 100 decision trees (Figure 2.5). Therefore, a more complex model would require more computing time without improving the classification. The details of each model are shown in Table 2.2.
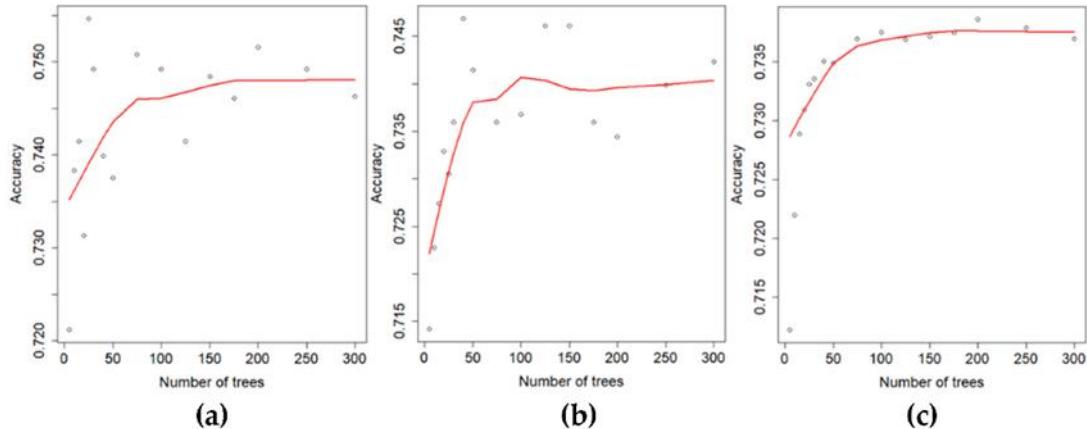
Figure 2.5. Convergence of the generalization error of the models. The models were trained with an increasing number of decision trees, and the accuracy of classification was evaluated from the pixel-based classification. The convergence of the three models occurred at around 100 decision trees. Images (a), (b) and (c) show the convergence of Models 1, 2 and 3, respectively.

Table 2.2. Summary of the parameters used in each model.

	Input variables	N input variables	Training ROIs (parcels)	Samples by variable (pixels)	N trees	Selected variables in each bagging
Model 1	B2, B3, B4, B8, EVI, TTVI (Sentinel-2)	6	144	2847	100	2
Model 2	B2, B3, B4, B6, B7, B8, B8A, B11, B12, EVI, TTVI, NDMI1, NDMI2 (Sentinel-2)	14	144	2847	100	8
Model 3	B1, B2, B3, B4, EVI, TTVI (High-resolution image)	6	144	283,329	100	2

The image classified by pixels was masked to remove other land uses. The cadastral data were used to frame urban areas. Later, the rustic areas where agricultural use is not allowed were masked through the cartography of the Forest Territorial Action Plan of the Valencian Region (PATFOR). The resulting image was segmented using a cadastral layer that defines agricultural plots. Finally, each plot was classified as an object following the most frequent pixel value criteria inside the plot.

2.2.5. Accuracy assessment and validation

A key issue in the classification of remote sensing images is the evaluation of the results' precision (Breiman, 1996). Originating from traditional validation methods for pixel-based classifications, many object-based image analysis (OBIA) results have been assessed by point-based sampling, although object-based strategies are highly recommended (Congalton & Green, 2019; Radoux et al., 2010; Hernando et al., 2012). A polygon that represents a meaningful and homogenous patch of land is more suitable for accuracy assessments of maps that are produced with OBIA (Congalton & Green, 2019; Olofson et al., 2014; Whiteside et al., 2014). In this case, the agricultural plots are clearly defined units that present a homogeneous and independent behavior of the neighboring plots. For this reason, an evaluation approach based on plots was used.

For the accuracy assessment, 96 plots were randomly selected. The same number of plots of each class were selected to calculate the overall accuracy, user accuracy and

producer accuracy. Finally, the area of the misclassified and correctly classified plots was computed, and basic statistics were calculated to determine possible relationships between the classification accuracy and the size of the plots.

2.2. RESULTS

According to the classification of agricultural plots in the cadastral database, the agricultural area was segmented into 11,305 plots that occupy around 1900 Ha. Table 2.3 compares the area and the number of parcels classified as not in production, in production and abandoned by each model. The surface classified by each model was similar, but Model 3 classified more abandoned plots and not in production plots than Models 1 and 2. The spatial distribution of abandoned plots was similar in the three models, which indicates that there are more abandoned parcels in the coastal zone (Figure 2.6).
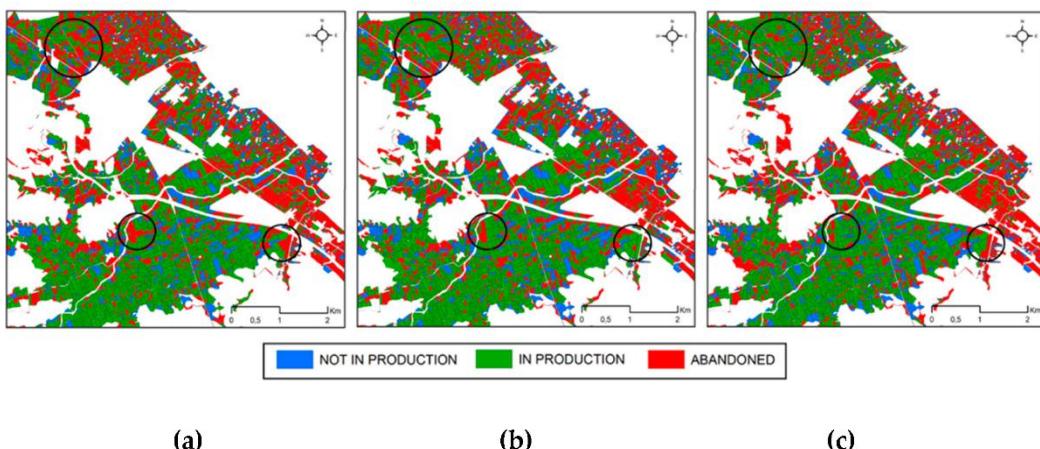
Figure 2.6. Land abandonment map generated by each model. Model 1 (a), Model 2 (b) and Model 3 (c). The circles show areas where the models predicted differently.

Table 2.3. Classification results of each model used.

	Not in production plots	Not in production surface (%)	In production parcels	In production surface (%)	Abandoned plots	Abandoned surface (%)
Model 1	2080	14.3	5069	51.0	4156	34.6
Model 2	2308	16.0	4959	51.9	4038	32.1
Model 3	2415	15.4	4587	52.2	4303	32.2

Table 2.4 shows the overall accuracy for each model. The most accurate map was obtained with Model 3 (Figure 2.7), which uses high-resolution airborne images. This model improved, by more than 10%, the accuracy of Models 1 and 2, which use Sentinel-2 images. Adding the Sentinel-2 20 m bands did not improve the classification accuracy.
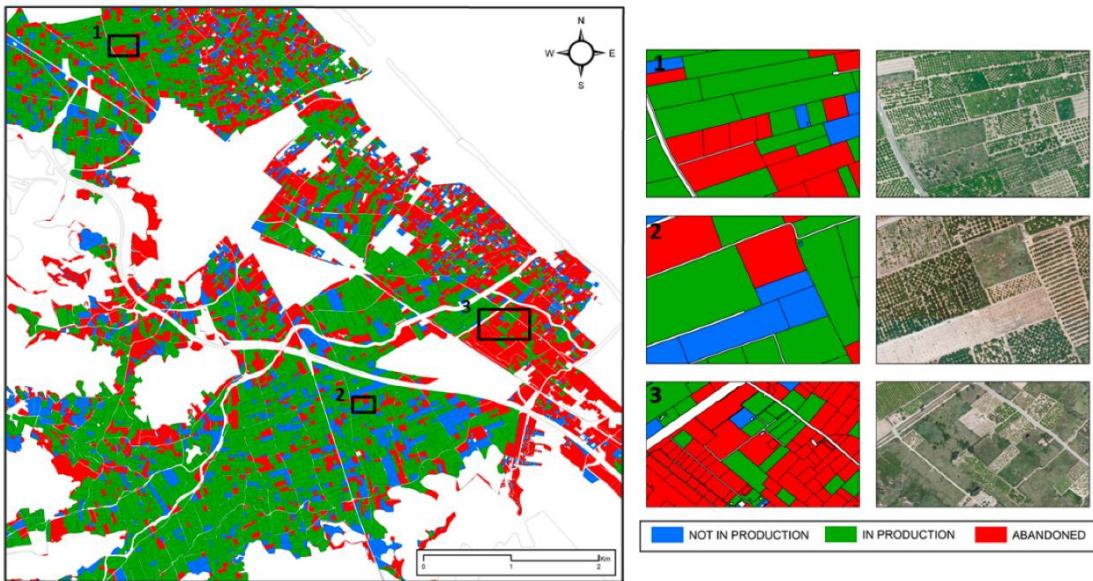
Figure 2.7. Model 3 map.

Table 2.4. Summary of accuracies obtained with each model.

	Overall accuracy
Model 1	77.1 %
Model 2	76.0 %
Model 3	88.5 %

The not in production class was easily detectable by all the models, with producer accuracies close to 100%. The distinguishing between in-production and abandoned classes was more inaccurate due to a similar spectral response. All the models obtained a classification accuracy for the abandoned class of 78%; however, Model 3 improved the classification accuracy for the in-production class compared to the other models. Model 3 obtained the best results for the differentiation of these two classes, with a producer's accuracy of 87.5% for the in-production class and 78.1% for the abandoned class. The confusion matrices for each model are shown in Table 2.5.

Table 2.5. Model confusion matrix.

Model 1					
	Reference				
Classified data	**Non-productive**	**Productive**	**Abandoned**	**Total map**	**User's accuracy**
Non-productive	31	4	2	37	83.8%
Productive	0	18	5	23	78.3%
Abandoned	1	10	25	36	69.4%
Total field	32	32	32		
Producer's accuracy	96.9%	56.3%	78.1%		
Model 2					
	Reference				
Classified data	**Non-productive**	**Productive**	**Abandoned**	**Total map**	**User's accuracy**
Non-productive	31	3	0	34	91.2%
Productive	0	17	7	24	70.8%
Abandoned	1	12	25	38	65.8%
Total field	32	32	32		
Producer's accuracy	96.9%	53.1%	78.1%		
Model 3					
	Reference				
Classified data	**Non-productive**	**Productive**	**Abandoned**	**Total map**	**User's accuracy**
Non-productive	32	2	0	34	94.1%
Productive	0	28	7	35	80.0%
Abandoned	0	2	25	27	92.6%
Total field	32	32	32		
Producer's accuracy	100%	87.5%	78.1%		

The analysis of the plot area shows that the Sentinel-2 images tend to fail with small plots (Figure 2.8). The average size of the 22 plots misclassified by Model 1 was 3416 m2 ± 2392, while the 74 correctly classified plots have an average area of 3986 m2 ± 2845. This difference was larger for Model 2, which uses lower resolution bands (20 m). The average size of the 23 plots misclassified by this model was 2798 m2 ± 1838, while the average size of the 73 plots correctly classified was 4162 m2 ± 2895. By contrast, the accuracy of Model 3 is not affected by the size of the plots. The average area of the 11 plots misclassified by Model 3 was 3956 m2 ± 3035, while the average area of the 85 plots classified correctly was 3834 m2 ± 2717.
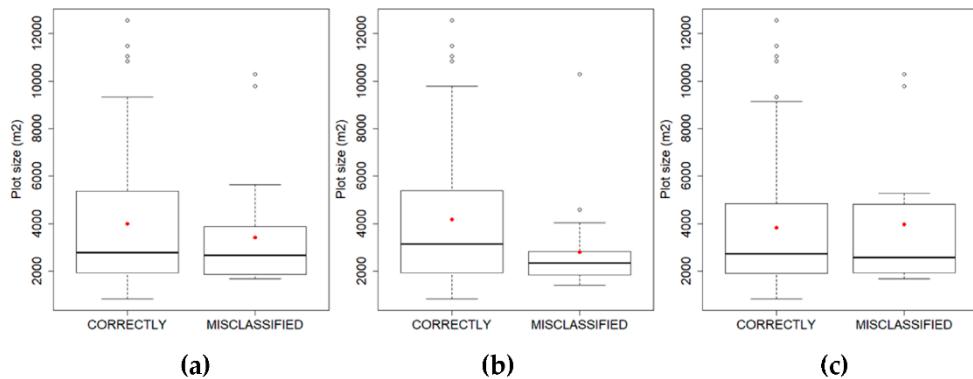
Figure 2.8. Boxplots of the plots correctly classified and misclassified by Model 1 (a), Model 2 (b) and Model 3 (c).

According to the Gini results (Figure 2.9), the TTVI was the most important variable in general. This means that the TTVI was able to separate categories by itself. TTVI was the most important variable in Models 1 and 2 using Sentinel-2 bands. In Model 3, the TTVI was the second most important variable. This last result may be related to the wavelength offset of the NIR band of the high-resolution image towards the red edge. The red band was also very important in Models 1 and 3. In Model 2, the NDMI1 was the second most important variable for pixel classification. However, adding these variables did not improve the classification accuracy of the plots. According to our interpretation, this means that the NDMI1 may differentiate between abandoned and in production plots, but the pixel size is not enough in our study area. Model 2 did not improve compared to Model 1 because most of the plots correctly classified by the NDMI1 could also be correctly classified by the TTVI.
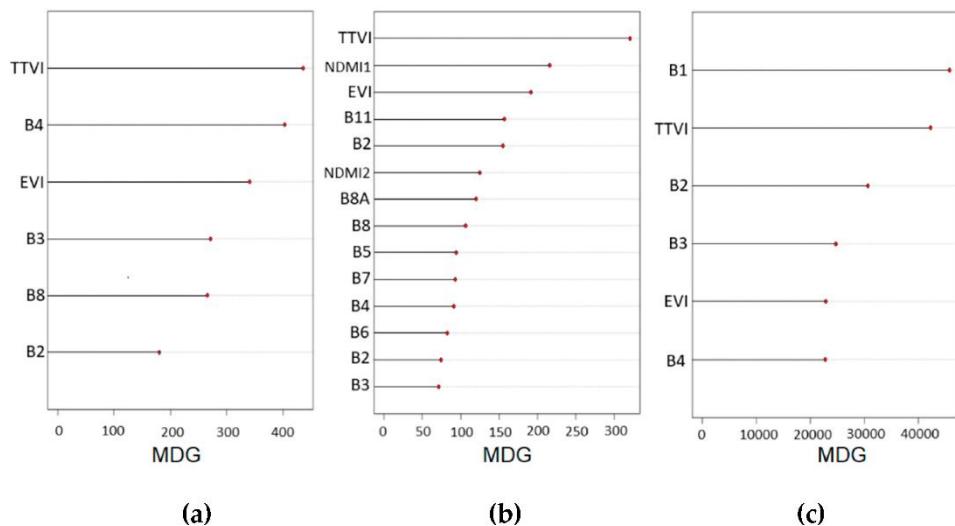
Figure 2.9. Importance of the variables according to Mean Decrease in Gini. Importance of Model 1 variables (a), Model 2 (b) and Model 3 (c).

2.3. DISCUSSION

Our results showed that high-resolution airborne images with four spectral bands generate more accurate results than Sentinel-2 images for mapping abandoned citrus plots. High-resolution images were the best alternative for the study of abandonment at

29

the cadastral plot level. However, processing these images is more costly from a computational point of view than processing Sentinel-2 images. Adding 20 m Sentinel-2 bands did not improve the accuracy of the classification, despite having more spectral information. The NDMI calculated from the 20 m bands of Sentinel-2 was not able to improve the classification accuracy. This is because the spatial resolution of these bands is not sufficient for the small plots in our study area. This fact does not exclude the use of the NDMI to detect abandonment in other less-fragmented areas, since it was an important variable according to the Gini Index. Although the map accuracy at the plot level was higher for high-resolution images, the abandoned areas obtained with both sources of information were similar. Our results show that Sentinel-2 images tend to fail with small plots. These results show that the resolution of Sentinel-2 images has limitations for the detection of abandonment in highly fragmented areas. However, the ability of Sentinel-2 images to correctly classify large plots may mask better apparent performance. This behaviour can explain the similar surface area obtained using both sources of information. Despite these limitations, Sentinel-2 probably offers better results than Landsat in fragmented areas due to its higher spatial resolution, but this should be studied.

The spatial distribution of the abandoned plots is in accordance with our previous experience (Morell-Monzó & Membrado-Tena, 2020). There is a high concentration of abandoned plots in the coastal zone due to the lower agronomic capacity of the soils and the prospects for urban growth. In the coastal area, many owners keep their plots without agricultural interest, and they expect future urban processes. There is also a high concentration of abandoned plots adjacent to urban areas where future urban expansion is expected. The active and replacement (productive and non-productive) plots are located in the inland area, where the soils are more productive and the crops are less exposed to the cold winds from the east. In this area, the citrus is harvested sooner and acquires greater market value.

The proposed methodology was adequate for studying the abandonment of citrus at the cadastral plot level. Our results recommend the use of high-resolution images for studies that require high precision at the plot level. However, performing the detection of changes based on multiple dates with different sensors requires the radiometric normalization of the images. Some radiometric standardization techniques have already been developed for the detection of citrus changes (Yang & Mueller, 2007). This is one of the main drawbacks of using high-resolution airborne images compared to Sentinel-2.

The availability of high-resolution images with four spectral bands for the entire Valencian territory since 2017 allows a retrospective investigation of the issue. Mapping the citrus land abandonment in the Valencian Region during recent years could help to better understanding of this problem, the detection of hot spots and the spatial location of actions to improve the landscape. The use of these images could help to automate the process of estimating annual citrus production in the Valencian Region. This would allow us to predict the economic impact of citrus based on the analysis of airborne spectral data to determine final crop harvests. Our study provides an overview of the capabilities of airborne images for authorities who wish to implement a monitoring system based on airborne images. This methodology could be applied in other areas of the planet with small plots of citrus or other types of crops when land abandonment is associated with the growth of wild vegetation. In other areas and for

other types of crops, the success of the classification will depend on the ability of the sensor to discriminate between wild vegetation and crops. It would be interesting to test this methodology for other types of crops, since these may result in different classification accuracy. Other types of airborne or satellite images could also be processed following this methodology. This would reduce the error in the estimation of citrus production reported by other authors in other areas such as Florida (USA) using images of moderate resolution (Shrivastava & Gebelein, 2007).

Although the use of high-resolution images is appropriate, we do not exclude the use of Sentinel-2 images for studies looking at general land-abandonment trends, and the plot level is not as important. The use of Sentinel-2 images is the best alternative when other types of images that are more expensive and computationally demanding are not available. Sentinel-2 can be useful for detecting land abandonment in areas where, probably, the Landsat resolution is not high enough. Furthermore, these images could improve the accuracy of the maps created with Landsat in previous studies (Grădinaru et al., 2019). High-resolution imagery currently has little applicability for large-area studies due to these images not being open access and the high computational requirements. However, Sentinel-2 images are free and have global spatial coverage and a temporal resolution of 5 days. In addition, new cloud-based processing technologies allow these images to be used on a large area efficiently. More research is needed to test Sentinel-2 images using methodologies for mapping abandoned land proposed by other authors (Estel et al., 2015; Yin et al., 2018; Baumann et al., 2011).

Future studies could exploit other approaches using these images. High-resolution images allow other approaches such as object-based classification by adding spatial and textural variables to classification models. This approach could increase object-level classification accuracy, differentiate more specific subcategories, or even provide information about crop age. On the other hand, the 5-day temporal resolution of the Sentinel-2 images allows a multi-temporal approach to mapping land abandonment. It would be interesting to evaluate the performance of Sentinel-2 images with these methodologies that have already been tested with other types of lower resolution images (Low et al., 2018; Alcántara et al., 2012; Yin et al., 2018). Although, a priori, this strategy is not the most appropriate for mapping the abandonment of citrus lands, this approach should be studied. A multi-time approach can also be useful for discriminating other types of crops, in mixed areas.

Remote sensing technology can be used to obtain spatial and temporal information about different ecosystem services and trade-offs. In this study, we have been able to map land abandonment in a highly fragmented crop, and in a study area with different types of soils at specific times. These results can be applied for monitoring agricultural ecosystems, providing useful information for their management. Future investigation with a multi-temporal approach will allow the study of the impact of land use change on soil properties. The abandonment of such an intensive activity as citrus farming can produce a recovery in soil properties. In this detrital plain, soils have very important environmental value and should also be considered in management plans (Smith et al., 2016).

2.4. CONCLUSIONS

Citrus land abandonment was studied in a highly fragmented area using a mono-temporal approach and pixel-based classification with the Random Forests algorithm.

The methodology used allowed the successful mapping of the abandonment of citrus lands using the Sentinel-2 and high-resolution airborne images. The plot-based map accuracy assessment confirmed the reliability of our approach. High-resolution airborne images showed better accuracies at the plot level (overall accuracy, 88.5%). However, at the surface level, both sources of information produced similar results. This is because Sentinel-2 images misclassify many small parcels but are able to correctly classify large parcels. Our results show that Sentinel-2 images tend to fail with small plots. The use of 20 m Sentinel-2 bands did not improve the classification accuracy. Furthermore, adding 20 m bands generated the misclassification of more small plots. Our results suggest the use of high-resolution images for studies that require a level of detail of the cadastral plot. Meanwhile, Sentinel-2 images can be used for studies that do not need as much detail. Ultimately, our methodology for mapping land abandonment can lead to a deeper understanding of land use and land cover change and allows its application with other images, in other areas and for other types of crops.

2.6. REFERENCES

Alcántara, C.; Kuemmerle, T.; Prishchepov, A.V.; Radeloff, V.C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sensing of Environment 2012, 124, 334–347. https://doi.org/10.1016/j.rse.2012.05.019

Árgyelán, T. Abandonment phenomenon in Europe. Acta Univ. Sapientiae Agric. Environ. 2015, 7, 89–97. https://doi.org/10.1515/ausae-2015-0008

Baudry, J. Ecological consequences of grazing extensification and land abandonment: Role of interactions between environment, society and techniques. In Land Abandonment and Its Role in Conservation, 1st ed.; Baudry, J., Bunce, R.G.H., Eds.; Options Méditerranéennes: Série A; Séminaires Méditerranéens: Zaragoza, Spain, 1991; Volume 15, pp. 13–19.

Baumann, M.; Kuemmerle, T.; Elbakidze, M.; Ozdogan, M.; Radeloff, V.C.; Keuler, N.S.; Prishchepov, A.V.; Kruhlov, I.; Hostert, P. Patterns and drivers of post-socialist farmland abandonment in Western Ukraine. Land Use Policy 2011, 28, 552–562. https://doi.org/10.1016/j.landusepol.2010.11.003

Bivard, R.; Keitt, T.; Rowlingson, B. rgdal: Bindings for the 'Geospatial' Data Abstraction Library, R package version 1.5-8; 2020. Available online: https://cran.r-project.org/web/packages/rgdal/index.html (accessed on 20 June 2020).

Breiman, L. Bagging predictors. Machine Learning 1996, 24, 123–140. https://doi.org/10.1007/BF00058655

Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Classification and Regression Trees, 1st ed.; Wadsworth: Monterey, CA, USA, 1984.

Breiman, L. Random Forests—Random Features; Technical Report 567; Statistics Department, University of California: Berkeley, CA, USA, 1999.

Breiman, L. Random Forests. Machine Learning 2001, 45, 5–32. https://doi.org/10.1023/A:1010933404324

Cerdà, A.; Ackermann, O.; Terol, E.; Rodrigo-Comino, J. Impact of Farmland Abandonment on Water Resources and Soil Conservation in Citrus Plantations in Eastern Spain. Water 2019, 11, 824. https://doi.org/10.3390/w11040824

Cerdà, A.; Brevik, C.E. The impact of abandonment of traditional flood irrigated citrus orchards on soil infiltration and organic matter. In Geoecología, Cambio Ambiental y

Paisaje: Homenaje al profesor José María García-Ruiz; Instituto Pirenaico de Ecología: Zaragoza, Spain, 2014; pp. 267–276.

Congalton, R.G.; Green, K. Assessing the Accuracy of Remotely Sensed Data, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2019. https://doi.org/10.1201/9780429052729

Conselleria de Agricultura, Desarrollo Rural, Emergencia Climática y Transición Ecológica. In Informe del Sector Agrario Valenciano 2019; Generalitat Valenciana: Valencia, Spain, 2020.

Corbelle-Rico, E.; Crecente-Maseda, R. Estudio da evolución da superficie agrícola na comarca da Terra Chá a partir de fotografía aérea histórica e mapas de usos, 1956-2004. Recursos Rurais 2018, 4, 57–65. https://doi.org/10.15304/rr.id5312

Dara, A.; Baumann, M.; Kuemmerle, T.; Pflugmacher, D.; Rabe, A.; Griffiths, P.; Hölzel, N.; Kamp, J.; Freitag, M.; Hostert, P. Mapping the timing of cropland abandonment and recultivation in northern Kazakhstan using annual Landsat time series. Remote Sens. Environ. 2018, 213, 49–60. https://doi.org/10.1016/j.rse.2018.05.005

Dubinin, M.; Potapov, P.; Lushchekina, A.; Radeloff, V.C. Reconstructing long time series of burned areas in arid grasslands of southern Russia by satellite remote sensing. Remote Sensing of Environment 2010, 114, 1638–1648. https://doi.org/10.1016/j.rse.2010.02.010

Estel, S.; Kuemmerle, T.; Alcántara, C.; Levers, C.; Prishchepov, A.V.; Hostert, P. Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. Remote Sensing of Environment 2015, 163, 312–325. https://doi.org/10.1016/j.rse.2015.03.028

Fischer, J.; Hartel, T.; Kuemmerle, T. Conservation policy in traditional farming landscapes. Conservation Letters 2012, 5, 167–175. https://doi.org/10.1111/j.1755-263X.2012.00227.x

Gellrich, M.; Baur, P.; Koch, N.; Zimmermann, E. Agricultural land abandonment and natural forest re-growth in the Swiss mountains: A spatially explicit economic analysis. Agriculture, Ecosystems and Environment. 2007, 118, 93–108. https://doi.org/10.1016/j.agee.2006.05.001

Gislason, P.O.; Benediktsson, J.A.; Sveinsson, R.J. Random Forests for land cover classification. Pattern Recognition Letters 2006, 27, 294–300. https://doi.org/10.1016/j.patrec.2005.08.011

Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment. 1996, 58, 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Grădinaru, S.R.; Kienast, F.; Psomas, A. Using multi-seasonal Landsat imagery for rapid identification of abandoned land in areas affected by urban sprawl. Ecological Indicators 2019, 96, 79–86. https://doi.org/10.1016/j.ecolind.2017.06.022

Hernando, A.; Tiede, D.; Albrecht, F.T.; Lang, S. Novel parameters for evaluating the spatial and thematic accuracy of land cover maps, International Conference on Geographic Object-Based Image Analysis. In Proceedings of the International Conference on Geographic Object-Based Image Analysis, 4. (GEOBIA), Rio de Janeiro, Brazil, 7–9 May 2012.

Hijmans, R.J. raster: Geographic Data Analysis and Modeling, R package version 3.1-5; 2020. Available online: https://rdrr.io/cran/raster/ (accessed on 20 June 2020).

Huete, A. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 1988, 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Huete, A.; Justice, C.; Liu, H. Development of vegetation and soil indices for MODIS-EOS. Remote Sensing of Environment 1994, 49, 224–234. https://doi.org/10.1016/0034-4257(94)90018-3

Huete, A.; Justice, C.; Van Leeuwen, W. MODIS vegetation index (MOD13). Algorithm Theor. Basis Doc. 1999, 3, 1–129.

Instituto Valenciano de Investigaciones Agrarias (IVIA). Citricultura Valenciana: Gestión Integrada de Plagas y Enfermedades en Cítricos. Available online: http://gipcitricos.ivia.es/citricultura-valenciana (accessed on 11 May 2020).

Kanjir, U.; Đurić, N.; Veljanovski, T. Sentinel-2 Based Temporal Detection of Agricultural Land Use Anomalies in Support of Common Agricultural Policy Monitoring. ISPRS International Journal of Geo-Information 2018, 7, 405. https://doi.org/10.3390/ijgi7100405ç

Kosmas, C.; Kairis, O.; Karavitis, C.; Acikalin, S.; Alcalá, M.; Alfama, P.; Atlhopheng, J.; Barrera, J.; Belgacem, A.; Solé-Benet, A.; et al. An exploratory analysis of land abandonment drivers in areas prone to desertification. CATENA 2015, 128, 252–261. https://doi.org/10.1016/j.catena.2014.02.006

Kuemmerle, T.; Hostert, P.; Radeloff, V.C.; Linden, S.; Perzanowski, K.; Kruhlov, I. Cross-border Comparison of Post-socialist Farmland Abandonment in the Carpathians. Ecosystems 2008, 11, 614–628. https://doi.org/10.1007/s10021-008-9146-z

Kuhn, M.; Cont. Caret: Classification and Regression Training, R package version 6.0-84; 2019; Available online: https://cran.r-project.org/web/packages/caret/caret.pdf (accessed on 20 June 2020).

Liaw, A.; Wiener, M. Classification and regression by randomForest. R News 2002, 2, 18–22.

Löw, F.; Prishchepov, F.; Waldner, F.; Dubovyk, O.; Akramkhanov, A.; Biradar, C.; Lamers, J. Mapping Cropland Abandonment in the Aral Sea Basin with MODIS Time Series. Remote Sensing 2018, 10, 159. https://doi.org/10.3390/rs10020159

MacDonald, D.; Crabtree, J.R.; Wiesinger, T.; Dax, T.; Stamou, N.; Fleury, P.; Lazpita, J.G.; Gibon, A. Agricultural abandonment in mountain areas or Europe: Environmental consequences and policy response. Journal of Environmental Management. 2000, 59, 47–69. https://doi.org/10.1006/jema.1999.0335

Maxwell, A.E.; Warner, T.A.; Fand, F. Implementation of machine-learning classification in remote sensing: An applied review. International Journal of Remote Sensing 2018, 39, 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. ESYRCE: Encuesta Sobre Superficies y Rendimientos; (N.I.P.O.: 013-17-120-0.); Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2017. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Morell-Monzo, S.; Membrado-Tena, J.C. Causas y consecuencias del crecimiento urbanístico del litoral valenciano a través de la evolución de los usos del suelo. El caso de Oliva. Cuadernos de Turismo 2020, 44, 303–326. https://doi.org/10.6018/turismo.44.404861

Müller, D.; Leitão, P.J.; Sikor, T. Comparing the determinants of cropland abandonment in Albania and Romania using boosted regression trees. Agricultural Systems 2013, 117, 66–77. https://doi.org/10.1016/j.agsy.2012.12.010

Navarro, L.M.; Pereira, H.M. Rewilding Abandoned Landscapes in Europe. In Rewilding European Landscapes, 21st ed.; Pereira, H.M., Navarro, L.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; Volume 26, pp. 3–25.

Noguera, J. Viabilidad y competitividad del sistema citrícola valenciano. Boletín de la Asociación de Geógrafos Españoles 2010, 52, 81–99.

Noi, P.T.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors 2017, 18, 18. https://doi.org/10.3390/s18010018

Novara, A.; Gristina, L.; Sala, G.; Galati, A.; Crescimanno, M.; Cerdà, A.; Badalamenti, E.; La Mantia, T. Agricultural land abandonment in Mediterranean environment provides ecosystem services via soil carbon sequestration. Science of The Total Environment 2017, 576, 420–429. https://doi.org/10.1016/j.scitotenv.2016.10.123

Olofson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment 2014, 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Pal, M. Random Forest classifier for remote sensing classification. International Journal of Remote Sensing 2005, 26, 217–222. https://doi.org/10.1080/01431160412331269698

Penov, I. The use of irrigation water in Bulgaria's Plovdiv Region during transition. Environmental Management 2004, 34, 304–313. https://doi.org/10.1007/s00267-004-0019-8

Perpiña-Castillo, C.; Kavalov, B.; Diogo, V. Agricultural Land Abandonment in the EU within 2015–2030; (Technical Report JRC113718); European Commission: Brussels, Belgium, 2018. https://EconPapers.repec.org/RePEc:ipt:iptwpa:jrc113718

Pinto-Correia, T. Land abandonment: Changes in the land use patterns around the Mediterranean basin. Etat de l'Agriculture en Méditerranée. Les sols dans la région méditerranéenne: Utilisation, gestion et perspectives d'évolution. Options Méditerranéennes Série A Séminaires Méditerranéens 1991, 1, 13–19. http://om.ciheam.org/article.php?IDPDF=95605230

Piqueras, J. El Espacio Valenciano. Una Síntesis Geográfica, 1st ed.; Gules: Valencia, Spain, 1999.

Prishchepov, A.V.; Radeloff, V.C.; Dubinin, M.; Alcantara, C. The effect of Landsat ETM/ETM image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Remote Sensing of Environment 2012, 126, 195–209. https://doi.org/10.1016/j.rse.2012.08.017

Proisy, C.; Viennois, G.; Sidik, F.; Andayani, A.; Enright, J.A.; Guitet, S.; Gusmawati, N.; Lemonnier, H.; Muthusankar, G.; Olagoke, A.; et al. Monitoring mangrove forests after aquaculture abandonment using time series of very high spatial resolution satellite images: A case study from the Perancak estuary, Bali, Indonesia. Marine Pollution Bulletin 2018, 131, 61–71. https://doi.org/10.1016/j.marpolbul.2017.05.056

Radoux, J.; Bogaert, P.; Defourny, P. Overall accuracy estimation for geographic object-based image classification. In Proceedings of the Ninth International

Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, UK, 20–23 July 2010.

R Core Team. R: A Language and Environment for Statistical Computing; R Core Team: Viena, Austria, 2019.

Rey-Benayas, J. Abandonment of agricultural land: An overview of drivers and consequences. CAB Reviews: Perspectives in Agriculture, Veterinary Sciences, Nutrition and Natural Resources. 2007, 2, 57.

Rey-Benayas, J.M.; Bullock, J.M. Restoration of Biodiversity and Ecosystem Services on Agricultural Land. Ecosystems 2012, 15, 883–899. https://doi.org/10.1007/s10021-012-9552-0

Richter, R.; Louis, J.; Müller-Wilm, U. Sentinel-2 MSI—Level 2A products algorithm theoretical basis document. Eur. Space Agency 2012, 49, 1–72.

Romero-Díaz, A.; Martínez-Hernández, C. Usos del suelo y abandono de tierras de cultivo en el altiplano Jumilla-Yecla (Región de Murcia). In Geoecología, Cambio Ambiental y Paisaje: Homenaje al Profesor José María García-Ruiz; Instituto Pirenaico de Ecología: Zaragoza, Spain, 2014; pp. 461–470.

Rounsevell, M.D.A.; Reginster, I.; Araújo, M.B.; Carter, T.R.; Dendoncker, N.; Ewert, F.; House, J.I.; Kankaanpää, S.; Leemans, R.; Metzger, M.J.; et al. A coherent set of future land use change scenarios for Europe. Agriculture, Ecosystems and Environment 2006, 114, 57–68. https://doi.org/10.1016/j.agee.2005.11.027

Ruiz-Flano, P.; Garcia-Ruiz, J.M.; Ortigosa, L. Geomorphological evolution of abandoned fields. A case study in the Central Pyrenees. CATENA 1992, 19, 301–308. https://doi.org/10.1016/0341-8162(92)90004-U

Salom, J.; Albertos, J.M. El modelo de desarrollo de la Comunidad Valenciana. In La Comunidad Valenciana en la Europa de las Regiones, 1st ed.; Romero, J.S.J., Vera, F., Eds.; Ariel Geografía: Valencia, Spain, 2001.

Silleos, N.G.; Alexandridis, T.K.; Gitas, I.Z.; Perakis, K. Vegetation Indices: Advances Made in Biomass Estimation and Vegetation Monitoring in the Last 30 Years. Geocarto International 2006, 21, 21–28. https://doi.org/10.1080/10106040608542399

Shrivastava, R.J.; Gebelein, J.L. Land cover classification and economic assessment of citrus groves using remote sensing. ISPRS J. Photogrammetry and Remote Sensing 2007, 61, 341–353. https://doi.org/10.1016/j.isprsjprs.2006.10.003

Smelansky, I.E. Biodiversity of Agricultural Lands in Russia: Current State and Trends. In UICN–The World Conservation Union; Ladonina, N.N., Gorelova, Y.V., Chernyakhovsky, D.A., Eds.; IUCN Representative Office for Russia and CIS: Moscow, Russia, 2003; p. 52.

Smith, P.; House, J.I.; Bustamante, M.; Sobocka, J.; Harper, R.; Pan, G.; West, P.C.; Clark, J.M.; Adhya, T.; Rumpel, C.; et al. Global change pressures on soils from land use and land management. Global Change Biology. 2016, 22, 1008–1028. https://doi.org/10.1111/gcb.13068

Szostak, M.; Hawryło, P.; Piela, D. Using of Sentinel-2 images for automation of the forest succession detection. European Journal of Remote Sensing 2017, 51, 142–149. https://doi.org/10.1080/22797254.2017.1412272

Thiam, A.K. Geographic Information Systems and Remote Sensing Methods for Assessing and Monitoring Land Degradation in the Sahel Region: The Case of Southern Mauritania. ProQuest Dissertations and Theses. Ph.D. Thesis, Clark University, Worcester, MA, USA, 1998.

Tomás-Carpi, J.A. La economía valenciana: Modelos de interpretación. In Contribución Invisible de las Mujeres en la Economía: El Caso Específico del Mundo Rural, 1st ed.; Vera, A., Ed.; Instituto De La Mujer: Madrid, Spain, 1977.

Verbug, P.H.; Schulp, C.J.E.; Witte, N.; Veldkamp, A. Downscaling of land use change scenarios to assess the dynamics of European landscapes. Agriculture, Ecosystems and Environment 2006, 114, 39–56. https://doi.org/10.1016/j.agee.2005.11.024

Viñals, M.J. El Marjal de Oliva-Pego: Geomorfología y Evolución de un Humedal Costero Mediterráneo, 1st ed.; Conselleria de Agricultura y Medio Ambiente, Generalitat Valenciana: Valencia, Spain, 1996.

Viñals, M.J. Secuencias Estratigráficas y Evolución Morfológica del Extremo Meridional del Golfo de Valencia (Cullera-Dénia). El Cuaternario del País Valenciano, 1st ed.; Universitat de València-AEQUA: Valencia, Spain, 1995.

Whiteside, T.G.; Maier, S.W.; Boggs, G.S. Area-based and location-based validation of classified image objects. International Journal of Applied Earth Observation and Geoinformation 2014, 5, 117–130. https://doi.org/10.1016/j.jag.2013.11.009

Wilson, E.H.; Shader, S.A. Detection of forest harvest type using multiple dates of Landsat TM imagery. Remote Sensing of Environment 2002, 80, 385–396. https://doi.org/10.1016/S0034-4257(01)00318-2

Yang, Z.; Mueller, R. Heterogeneously sensed imagery radiometric response normalization for citrus grove change detection. In Optics for Natural Resources, Agriculture, and Foods; SPIE: Washington, DC, USA, 2007. https://doi.org/10.1117/12.735345

Yin, H.; Prishchepov, A.V.; Kuemmerle, T.; Bleyhl, B.; Buchner, J.; Radeloff, V.C. Mapping agricultural land abandonment from spatial and temporal segmentation of Landsat time series. Remote Sensing of Environment 2018, 210, 12–24. https://doi.org/10.1016/j.rse.2018.02.050

CAPÍTULO 3

Land Use Classification of VHR Images for Mapping Small-Sized Abandoned Citrus Plots by Using Spectral and Textural Information

Morell-Monzó, S.; Sebastiá-Frasquet, M.-T.; Estornell, J. Land Use Classification of VHR Images for Mapping Small-Sized Abandoned Citrus Plots by Using Spectral and Textural Information. Remote Sensing, 2021, 13, 681. https://doi.org/10.3390/rs13040681

3.1. INTRODUCTION

The Comunitat Valenciana (CV) region is located in eastern Spain. Its weather conditions, soil characteristics and water availability have made this region one of the richest agricultural areas in the Mediterranean basin (Amorós et al., 2011; IVIA, 2020). The CV region is the most important citrus producer in Spain with an area of 160,912 ha, about 60% of the national area of this crop according to the Survey on Crop Areas and Yields of 2019 (MAPA, 2019). Spain is the most important citrus producer in the European Union (EU) and the fifth highest producer in the world (Noguera, 2010). However, in recent years, there has been a massive abandonment of these agricultural holdings. In the CV region there has been a decrease of 15% in citrus areas from 2008 to 2018 (from 188,650 ha to 161,944 ha) (MAPA, 2008; MAPA, 2018) and an increase in abandoned areas is expected due to the socio-economic changes that are taking place in the EU (Rounsevell et al., 2006).

Monitoring this issue is an important point for land and landscape management, but also for creating a citrus inventory and making annual yield estimations. Currently there are no public cartographic sources of land use to map land abandonment at parcel resolution with appropriate thematic detail. The Spanish Land Occupation Information System (SIOSE) is an object-based and hierarchical database of land use/land cover in Spain. SIOSE allows the identification of abandoned areas, but its minimum area of representation is 2 ha. On the other hand, the Geographic Information System for Agricultural Parcels (SIGPAC) is available in the member states of the EU. This database has cadastral plot resolution and identifies as abandoned plots next land uses: unproductive land, grassland and arable land. However, there are other plots classified as crops (e.g., citrus) that are actually abandoned. Therefore, it is possible to estimate the abandoned area through SIGPAC, but its insufficient thematic detail inevitably causes an underestimation of the abandoned area. In addition, this system is periodically updated manually, which can cause errors and unregistered changes during the update period.

Traditionally, land use maps such as SIOSE or SIGPAC have been generated through either photo-interpretation or visiting agricultural plots directly, because of the high level of accuracy required for further administrative actions (Amorós et al., 2011). Nowadays, remote sensing is a powerful tool for obtaining maps that can replace the creation of manual land use maps with a more efficient personal resources use. Land abandonment has been previously approached by remote sensing by other authors (Löw et al., 2018; Alcántara et al., 2012; Estel et al., 2015; Yin et al., 2018; Grădinaru et al., 2019). These previous studies have been conducted in areas that do not require high spatial resolution, with MODIS and Landsat products. However, detecting abandoned citrus plots in the CV region by remote sensing is challenging due to the high spatial fragmentation and small size of the agricultural plots which makes necessary the use of high-resolution products. Furthermore, previous studies have focused on land abandonment of crops with fast phenological dynamics. However, land abandonment in woody evergreen crops (e.g., citrus) has been little studied. Previous studies on citrus already showed serious discrepancies between agriculture census data and satellite-derived cropland area using medium-resolution satellite imagery as Landsat (Schrivastava & Gebelein, 2007). Sentinel-2 images overcome some of these limitations and showed its potential in agricultural applications (Ramírez et al., 2020; Campos-

Taberner et al., 2020), however in areas with high spatial fragmentation they may be not enough (Vajsová et al., 2020; Vajsová et al., 2019). The Sentinel-2 satellite images did not show enough accuracy to identify abandoned plots in our study area due to resolution limitations and the small size of the plots, for this reason, the use of higher resolution images is recommended (Morell-Monzó et al., 2020).

Highly fragmented areas force the use of Very High-Resolution (VHR) images. However, the use of VHR image time series is costly from an economic and operational point of view, which forces us to focus on procedures that use a single image (mono-temporal approaches). In addition, VHR images usually have limitations derived from their low spectral resolution, which makes mapping vegetation difficult especially for very similar covers. The limitations of low spectral resolution can be reduced using robust machine learning algorithms and textural image analysis (Feng et al., 2015).

There is a wide set of techniques for the extraction of texture features: statistical methods such as the Gray Level Co-occurrence Matrix (GLCM) (Haralick et al., 1973), Local Binary Patterns (Wang & He, 1990), filtering techniques such as energy filters (Laws, 1980) or Gabor filters (Jain & Farrokhnia, 1991), methods based on wavelet decomposition (Mallat, 1989) and methods derived from geostatistical functions (Balaguer et al., 2010). Texture measures quantitatively describe relationships of Digital Numbers (DN) values of neighbor pixels (Hall-Beyer, 2017). Texture features provide contextual information that allows us to differentiate patterns in the spatial distribution of objects in the image and increases classification accuracy.

As far as we know there are no previous studies on land abandonment which use VHR images and texture features. However, the GLCM features have shown to improve classification accuracy (Lan & Liu, 2018), especially in high resolution images (Feng et al., 2015; Wang et al., 2015; Jia et al., 2012; Stumpf & Kerle, 2011). There are previous studies which use GLCM texture features and VHR images to map vegetation and determine the health status of trees (Feng et al., 2015; Wang et al., 2015). These studies share characteristics with the detection of abandonment in citrus since an abandoned plot shows a growth of wild vegetation and trees in poor health conditions. In these cases, GLCM features and the Random Forests algorithm were used. There are also previous studies on the identification of citrus plots using VHR images but these studies did not focus on identifying abandoned plots. These studies are generally based on two types of approaches. The first group is based on the extraction of spectral and textural features of the plot (Gil-Yepes et al., 2016; Ruiz et al., 2011), while the second group is based on the detection of trees (Amorós et al., 2011; Recio et al., 2013). The second approach has several limitations, the most important the fine spatial resolution images needed to detect a single citrus tree. In addition, the development of wild vegetation during land abandonment can mask citrus trees and makes more difficult their identification. Although single tree approach is widely applied in precision agriculture to analyze intra-plot variations, information of land abandonment at level tree is not so relevant. The objective of land abandonment studies is to classify the whole plots into a unique category what can be done considering their spectral and textural information. The use of textural information for a single tree approach will require more advanced algorithms to extract and classify them as abandoned.

Since 2017 a VHR airborne image is available annually for the entire CV region through the Valencian Cartographic Institute (ICV). These images are currently used to update the SIGPAC of the CV region. In 2011, Amorós et al. (2011) created a

methodology to identify citrus plots automatically. The main goal of this methodology was to detect plots containing citrus at the cadastral plot level using a multi-stage object detection approach to identify the citrus trees (Amorós-López et al., 2008). These advances allowed automating the classification of citrus plots to update the SIGPAC annually. However, this methodology does not allow to differentiate between productive citrus plots or abandoned. In the current context of land abandonment in the CV region, it is especially important to develop a procedure to map abandoned citrus plots to avoid overestimating the productive area.

In this work we present a multistage methodological approach to detect abandoned plots in areas with a predominance of citrus. Our approach is based on three general steps: First, a textural feature extraction based on the GLCM was performed from the Normalized Difference Vegetation Index (NDVI) of the high-resolution airborne images. Then, a supervised pixel-based classification was performed using the Random Forests algorithm. Finally, a refinement of the classification based on majority voting within each plot was carried out to classify each cadastral plot with a unique value.

3.2. DATA AND METHODS

This section describes the proposed methodology to detect abandoned plots in areas with a predominance of citrus. A classification based on three types of plots was proposed (Figure 3.1): (a) Not in production—these are plots where bare soil occupies most of the plot surface. These plots are in replacement periods. In this class, trees may be planted but are not yet productive. (b) In production—these are parcels with citrus productive cultivation. In this class, the citrus crop occupies most of the plot surface. (c) Abandoned—these are abandoned plots dominated by wild vegetation that show obvious signs of abandonment.
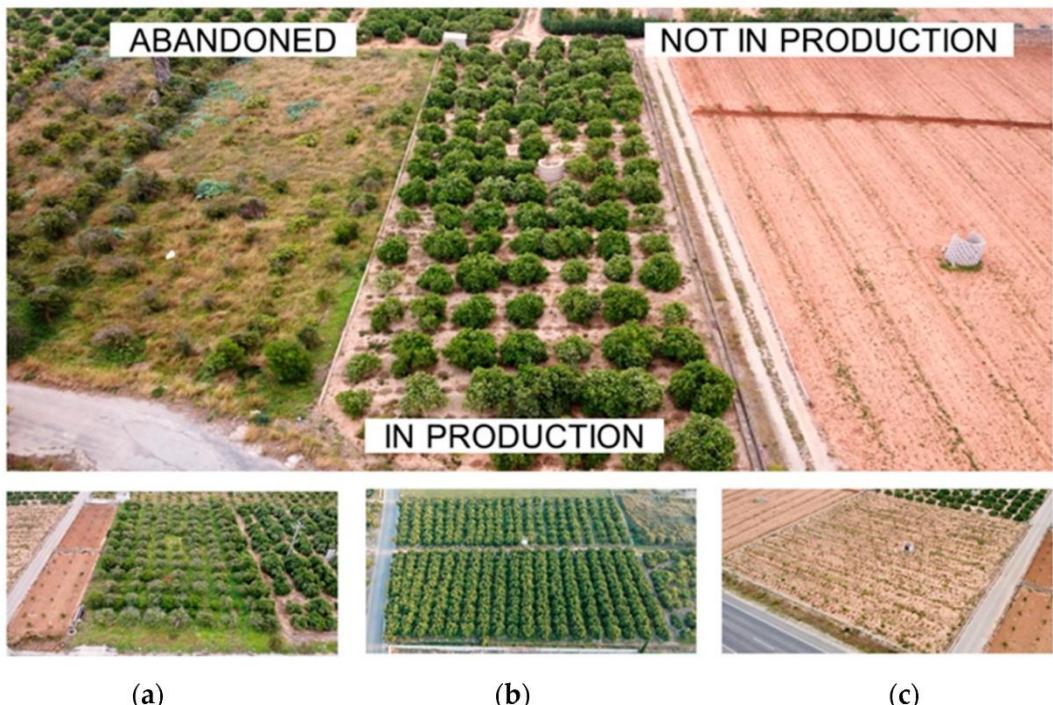
Figure 3.1. Aerial view of the three types of parcels classified: (a) abandoned, (b) in production and (c) non-productive.

41

The detection of abandoned parcels is based on the spectral and textural differences caused by the growth of wild vegetation, the loss of vigour of the trees (decreased foliar density and greenness) and the loss of regular spatial patterns of the crop. A more detailed description of the ecological succession of an abandoned parcel can be found in Morell et al. (2020).

3.2.1. Study area and data

The study area is divided into 4 locations of the CV region with a predominance of citrus. The main study area is in the municipality of Oliva (Figure 3.2), a coastal city in the CV region that covers an area of 59,600 ha, of which 50% is available for agriculture. Most of this area is located on a coastal plain with high agrological capacity (Viñals, 1995; Viñals, 1996). However, urban expansion and tourism compete with agricultural land uses and are promoting land abandonment. The main crop in this area is citrus, which represents more than 95% of the crops in the area (Generalitat Valenciana, 2020). In this area, the citrus plots have an average extension of 0.32 ha according to SIGPAC.
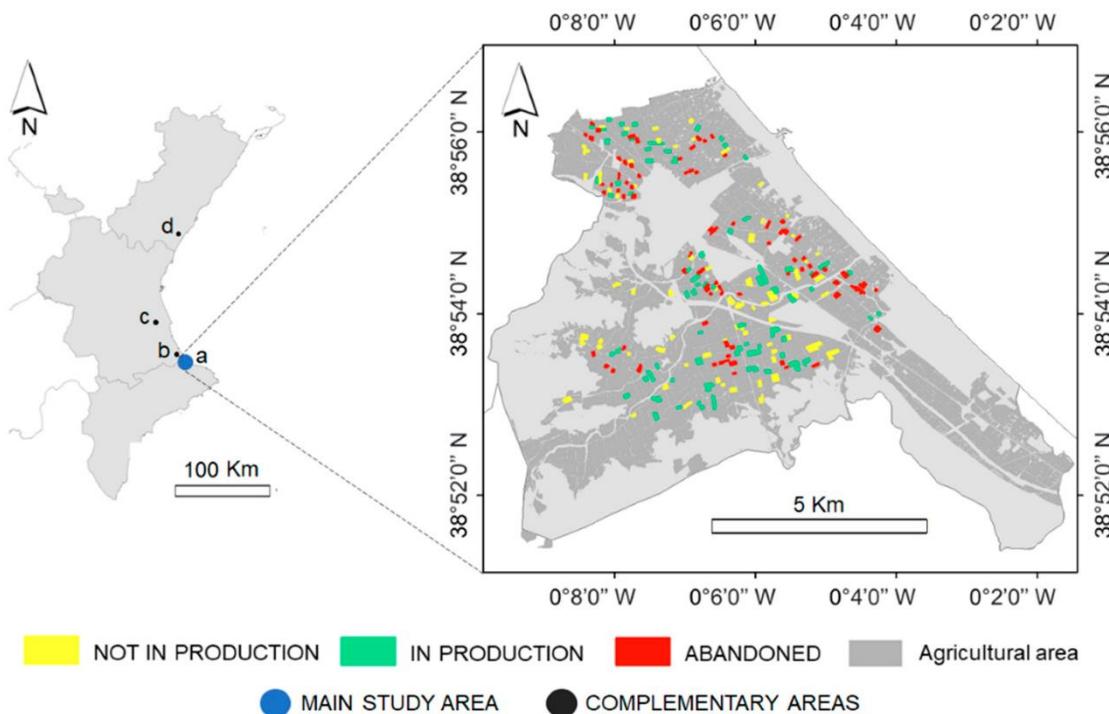
Figure 3.2. Study area. On the left the Comunitat Valenciana (CV) region and the study areas: (a) Oliva, (b) Bellreguard-Almoines, (c) Benicull-Polinyà del Xúquer and (d) Nules. On the right the main study area in Oliva municipality and the plots classified during the field survey.

Additionally, 3 complementary areas of 435 ha with a predominance of citrus were chosen in the municipalities of: Bellreguard-Almoines, Benicull-Polinyà del Xúquer and Nules (Figure 3.2).

The VHR images used were provided by the Valencian Cartographic Institute (ICV). The ICV provides an annual image (VHR) of the entire Valencian Region with a spatial

resolution of 0.25 m and four spectral bands: R (430 nm), G (530 nm), B (620 nm) and Red-edge (720 nm). These images were taken using an UltraCam Eagle UC-E-1-50016095-f80 camera from the company Vexcel Imaging GmbH © with a Qioptic Vexcel HR Digaron sensor. The images were taken on different dates in each study area (Table 3.1).

Table 3.1. Reference data and image acquisition date in each study area.

Study area	Number of parcels	Reference data acquisition	Image acquisition date
Oliva (a)	240	Field survey[1]	14, 15 May 2019
Bellreguard-Almoines (b)	90	Photointerpretation	14 May 2019
Benicull-Polinyà del Xuquer (c)	90	Photointerpretation	15 June 2019
Nules (d)	90	Photointerpretation	17 June 2019

[1] The field survey was conducted between 11 and 14 July 2019.

The ground truth dataset was collected during a field survey conducted between 11 and 14 July 2019 in the main study area (Oliva). Stratified random sampling with proportional distribution was performed. A set of 240 plots in an area of 98 ha was selected (80 not in production, 80 in production and 80 abandoned). These plots were used to train and validate the model. In addition, 270 plots were digitalized and classified in the complementary study areas (Table 3.1). To avoid errors during the photointerpretation process of the plots, the 2018 and 2020 images of the ICV were also used. These plots were used to test the model outside the study area.

3.2.1. Image processing and classification

Image processing is described as follows (Figure 3.3). The images were geometrically corrected and radiometrically calibrated by Vexcel Imaging GmbH©. Then, the images were resampled at 1 m resolution to reduce computing time. A mosaic was computed with the tiles that cover each study area. Then, the images were pre-processed to extract the GLCM-based texture features using the `GLCMTextures` package (Ilich, 2020) using an R 4.0.3 environment (R Core Team, 2020). The standard procedure for the automatic extraction of GLCM texture features is based on a grey level image considering the panchromatic band or calculating the first principal component of multispectral images (Huang et al., 2014). However, our imagery does not have a panchromatic band and performing a Principal Component Analysis would mean a loss of interpretability of our information. For that reason, the NDVI transformation was used to extract the GLCM textural features. This approach was used by other authors successfully (Akar & Güngör, 2015; Jin & Liu, 2018).
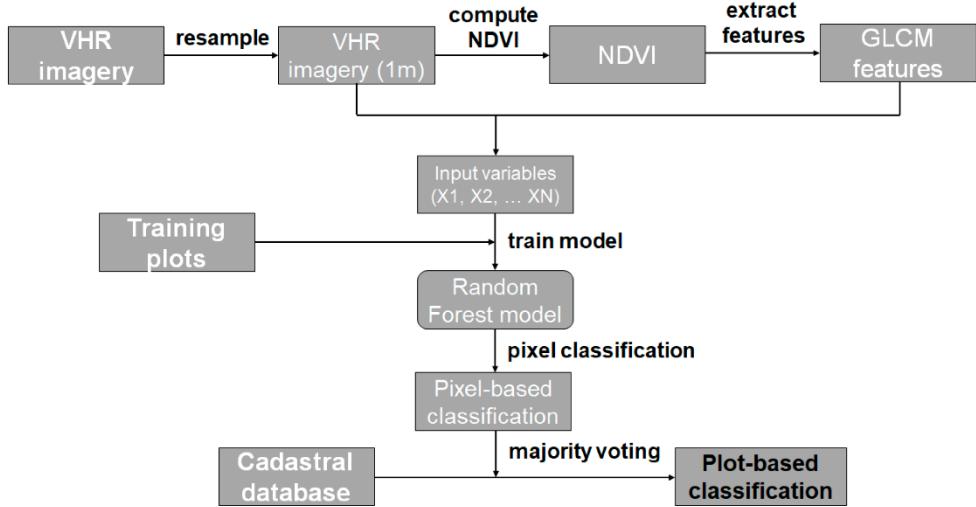
Figure 3.3. Workflow describing the image processing and classification.

The NDVI band was discretized in 32 gray levels to compute the GLCM measures. Then, 4 kernels with different window sizes were created (3×3, 5×5, 7×7, and 9×9) and 7 texture measures were computed. The texture measures were calculated in an invariant spatial direction that is the average of the four directions: 0°, 45°, 90°, and 135°. This omnidirectional approach was chosen as the crops do not present a specific direction.

Only 7 texture measures were computed out of the 14 originally proposed by (Haralick et al., 1973). These texture measures are the most widely used in remote sensing and image analysis (Hall-Beyer, 2017): mean, variance, contrast, dissimilarity, homogeneity, angular second moment (ASM), and entropy.

$$GLCM\ Mean\ (\mu) = \sum_{i,j=0}^{N-1} i(P_{i,j}) \tag{3.1}$$

$$GLCM\ Variance\ (\sigma^2) = \sum_{i,j=0}^{N-1} P_{i,j}(i - \mu)^2 \tag{3.2}$$

$$Contrast = \sum_{i,j=0}^{N-1} (i - j)^2 \tag{3.3}$$

$$Dissimilarity = \sum_{i,j=0}^{N-1} P_{i,j}\ |i - j| \tag{3.4}$$

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i + j)^2} \tag{3.5}$$

$$Angular\ Second\ Moment\ (ASM) = \sum_{i,j=0}^{N-1} P_{i,j}^2 \qquad (3.6)$$

$$Entropy = \sum_{i,j=0}^{N-1} P_{i,j}(-\text{In}\ P_{i,j}) \qquad (3.7)$$

where N is the number of rows or columns in the GLCM (equal to the number of gray levels), i are the row indices of the GLCM matrix (equal to gray level of reference cell), j are the column indices of the GLCM matrix (equal to gray level of neighboring cell) and $P_{i,j}$ is the Probability (relative frequency) of neighboring cells having gray levels i i and i.

Pixel-based classification was done using the RandomForests (RF) algorithm (Breiman, 1999), which is a non-parametric algorithm commonly used in remote sensing for supervised classifications (Gislason et al., 2006; Pal, 2005). The RF classifier consists of a set of classification trees such that each one is trained with an independent random vector and with the same distribution for each tree. Therefore, classifier consists of an ensemble of classifiers $\{h(x, \Theta k), k = 1, ..., \}$ where the $\{\Theta k\}$ are independent, identically distributed random vectors, and X is an input pattern (Breiman, 2001).

In training, the algorithm creates multiple classification trees, each one trained on a bootstrapped sample of the original training data, and searches only across a randomly selected subset of the input variables to determinate a split for each node. Therefore, trees require a measure of variable selection and a measure of the division purity to construct the nodes. The Gini Diversity Index was used as a measure of impurity (Equation 4).

In RF different classification trees are combined in the form of weak predictors, this means that the prediction of the model is determined by the majority vote of the trees. For this reason, RF is actually an assembly model that uses the bagging technique to combine the predictions of various decision trees. Bagging (or bootstrap aggregating) consists of training each decision tree with a bootstrapped sample (random selection with replacement) from the original training data set. At each bagging, 2/3 of the training data set is randomly selected to train each individual tree. Trees grow to their full extent without pruning. Even so, overfitting does not occur due to the use of bootstrapped samples to train individual trees and the fact that the model consists of a multitude of independently trained trees with random samples.

Our RF model was created using the `randomForest` package (Liaw & Wiener, 2002) from the `caret` API (Kunh, 2020) using an R 4.0.3 environment (R Core Team, 2020). Five classification models were trained using 100 decision trees. These models were trained with pixels extracted from 210 training plots out of the total 240 plots from the main study area. The first model took as independent variables (input variables) the pixel values of the Red (430 nm), Green (530 nm), Blue (620 nm) and Red-edge (720 nm) bands of the image. The remaining 4 models were trained with the pixel value of the spectral bands and the 7 texture features, each one at a different window size (3×3, 5×5, 7×7, and 9×9). The models took as the dependent variable (output variable) the

type of coverage according to our classification: not in production, in production and abandoned.

The image resulting from the pixel-based classification was segmented using the cadastral database. Then, these segments were used to enhance the classification results of the pixel based Random Forests algorithm using the majority voting within each segment (Morell-Monzó et al., 2020). This procedure creates an object-based map where each object is a cadastral plot (Ghorbanian et al., 2020) and improves the accuracy of pixel-based classification by removing misclassified pixels, within homogeneous segments, known as salt-and-pepper noise (Wulder et al., 2018). Finally, urban and forest plots were masked using the cadastral database and the Forestry Territorial Action Plan of the CV region.

3.2.3. Validation and accuracy assessment

An important issue in classifying remote sensing images is evaluating the results in terms of accuracy. Our approach was validated from two different points of view. First the pixel-based accuracy was evaluated (that is, the accuracy before the majority voting refinement). Second, the plot-based accuracy was evaluated (that is, the accuracy after the majority voting refinement). This double information allows us to know the performance improvements generated by the GLCM measures and the improvements generated by the majority voting.

The accuracy assessment was carried out through an 8-fold cross validation. In each step a balanced set of 30 plots was used for validation. The average accuracy and standard deviation of the pixel-based classification were obtained using a balanced set of pixels extracted from the 30 validation plots at each step. Then, average accuracy, standard deviation and average confusion matrix of the plot-based classification was obtained using the same 30 validation parcels through the 8-folds cross-validation. Additionally, the most accurate model was tested in the complementary areas and the plot-based accuracy was obtained. The model was applied in the complementary areas without any additional training data to test if the model can generalize outside of the training area. To this end a set of validation plots were identified by photointerpretation for each complementary area (Table 3.1).

3.2.4. Feature importance and model pruning

The Boruta feature selection algorithm (Kursa & Rudncki, 2010) was used to select the relevant features for classification. Boruta is a wrapping method around Random Forest classification algorithm. This algorithm is based on a backward removal of features that are not relevant, but it implements some improvements over the traditional Recursive Feature Elimination (Degenhardt et al., 2017; Kumar & Shaikh, 2017).

The measure of importance used by Boruta is the Z score. The importance measure of an attribute is obtained as the loss of accuracy of classification caused by the random permutation of attribute values between objects. It is computed separately for all trees which use a given attribute for classification. Then the mean decrease in accuracy is divided by its standard deviation (σ) is obtained for computing the Z score:

$$Z\ score = \frac{MDA}{\sigma} \tag{3.7}$$

46

Since we cannot use Z score directly to decide if importance of any given attribute is significant (that is, whether it is discernible from importance which may arise from random fluctuations), Boruta compares the importance of real variables with the importance of random variables by design, called shadow features. Then it removes the features that are less important than the best shadow variable and repeats the process iteratively for a few steps (Kursa et al., 2010).

This approach was used to decide which variables are significantly important for our classification. Further, a ranking of importance of variables was carried out using the Z score, and then a pruning of the model was carried out, removing the less relevant variables recursively.

3.3. RESULTS

3.3.1. Model tunning

The five models were trained with an increasing number of trees up to 500 and the out-of-bag error (OOB error) was used as a measure of accuracy. All models showed convergence before 100 decision trees. After the 100 trees, the models did not generate an out-of-bag error reduction (Figure 3.4), therefore the models were trained with 100 trees. The training history also showed a better performance of the models that include spectral and GLCM textural features regarding the spectral model (Figure 3.4).
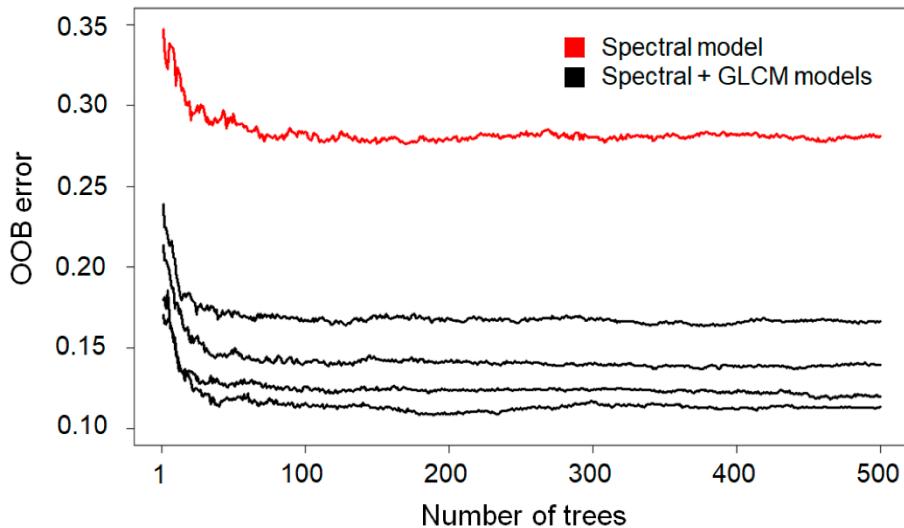
Figure 3.4. Random Forest models training history. The plot shows the evolution of the out-of-bag error as the number of trees increases. All models converge before training 100 decision trees.

3.3.2. Classification and accuracy assessment

A summary of the average accuracy and its standard deviation for each model obtained through cross-validation can be observed in Table 3.2. The average accuracy of the pixel-based and plot-based classification is shown separately. The least accurate model was the one that only included spectral variables. This model showed a pixel-based classification accuracy of 71% and a plot-based classification accuracy of 91%. The most accurate model was the one that included spectral and GLCM textural features at a 9×9 window size with a pixel-based classification accuracy of 87% and a plot-based

accuracy of 95%. However, models that included spectral and GLCM textural features with 5×5 and 7×7 window sizes also showed a similar accuracy. Table 3.3 shows the performance of the model in terms of precision, recall, and F1-score for the plot-based classification.

Table 3.2. Average accuracy of each model in pixel-based classification and parcel-based classification through an 8-folds cross validation.

Model	Average Pixel Accuracy	Standard Deviation Pixel Accuracy	Average Parcel Accuracy	Standard Deviation Parcel Accuracy
Spectral features	0.71	0.04	0.91	0.07
Spectral + 3x3 GLCM features	0.84	0.03	0.95	0.05
Spectral + 5x5 GLCM features	0.86	0.03	0.95	0.05
Spectral + 7x7 GLCM features	0.86	0.03	0.95	0.06
Spectral + 9x9 GLCM features	0.87	0.03	0.95	0.04

Table 3.3. Precision, recall, and F1-score averages of each model through 8-folds cross-validation.

Model	Precision			Recall			F1-score		
	Non-productive	Productive	Abandoned	Non-productive	Productive	Abandoned	Non-productive	Productive	Abandoned
Spectral features	0.94	0.90	0.92	0.99	0.86	0.88	0.96	0.87	0.89
Spectral + 3x3 GLCM features	0.97	0.93	0.95	0.96	0.94	0.94	0.96	0.93	0.94
Spectral + 5x5 GLCM features	0.97	0.93	0.95	0.96	0.94	0.94	0.96	0.93	0.94
Spectral + 7x7 GLCM features	0.97	0.95	0.93	0.98	0.91	0.95	0.97	0.93	0.94
Spectral + 9x9 GLCM features	0.96	0.96	0.94	0.98	0.91	0.96	0.96	0.92	0.95

Adding GLCM textural features resulted in an accuracy improvement of up to 16% for pixel-based classification. As the window size was increased, there was a decrease in noise in the pixel-based classification (Figure 3.5). The noise reduction in the 3×3 window size model regarding the spectral model led to an improvement in object-based classification. However, in the models with larger window sizes the noise reduction did not lead to an improvement in the object-based classification compared with the 3×3 window size model. This is because as the noise from the pixel-based classification is reduced, the improvement capacity of the majority voting refinement is smaller. On the other hand, the use of majority voting refinement led to an improvement in accuracy in

all cases. The most notable accuracy improvement was in the model that only includes spectral features where the accuracy improved by 20%.

| | NOT IN PRODUCTION | | IN PRODUCTION | | ABANDONED |
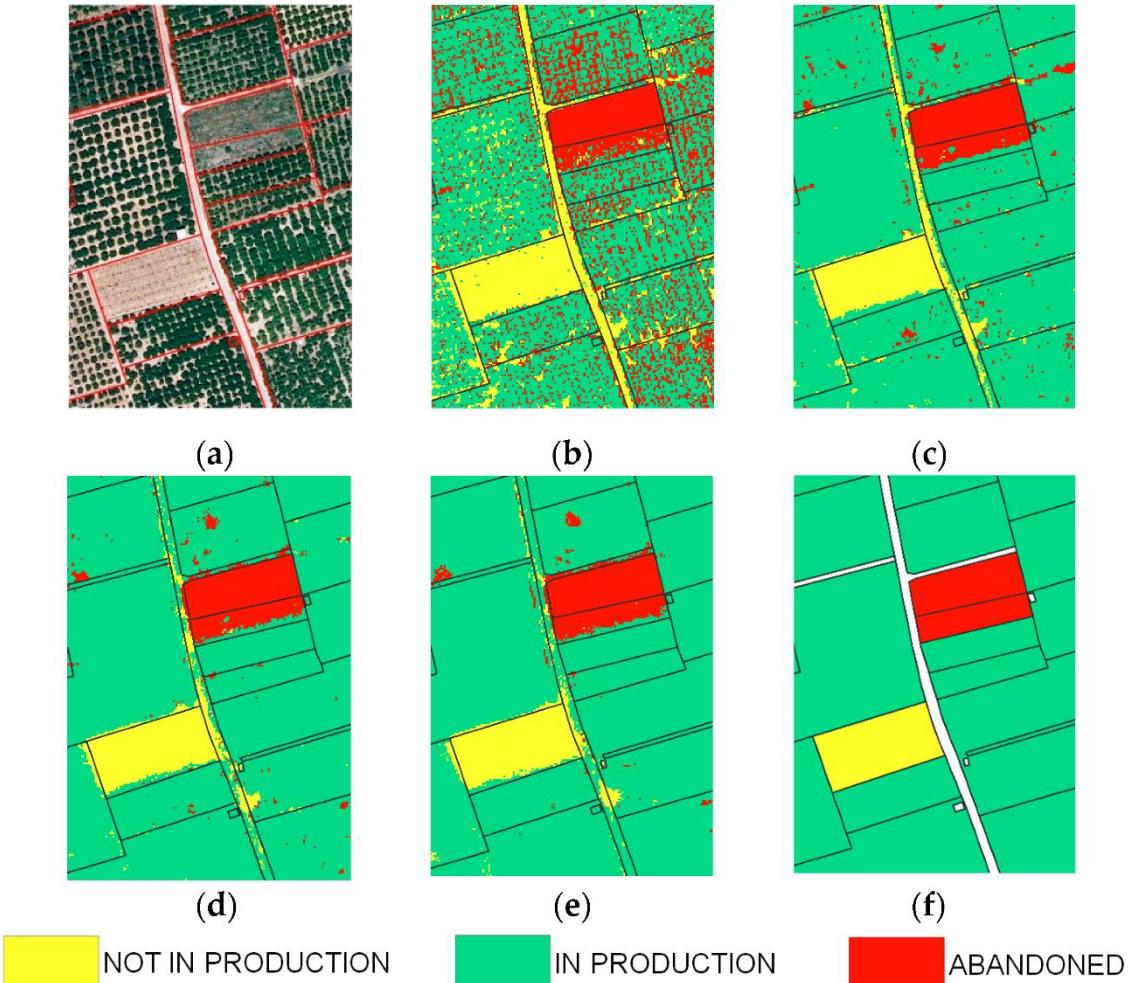
Figure 3.5. Example of reducing salt and pepper noise by adding Gray Level Co-Occurrence Matrix (GLCM) textural features. The noise decreases as the window size increases and this causes an increase in the accuracy of the pixel-based classification. (a) true colour image, (b) spectral model classification, (c) 3×3 GLCM model classification, (d) 5×5 GLCM model classification, (e) 9×9 GLCM model classification, (f) parcel-based classification.

For the most accurate model (9×9 GLCM model classification), the highest producer's accuracy (98%) was obtained for the category non-productive (Table 3.4). This category is more easily detectable due to the spectral differences compared to the two other categories. The in production and abandoned categories obtained a producer's accuracy of 91% and 96%, respectively. User's accuracy for categories not in production, in production and abandoned was 95%, 96%, and 94%, respectively.

Table 3.4. Average confusion matrix obtained through 8-folds cross-validation.

		Reference			Average User's Accuracy
		Non-productive	Productive	Abandoned	
Predicted	**Non-productive**	0.98	0.04	0.01	0.95
	Productive	0.01	0.91	0.03	0.96
	Abandoned	0.01	0.05	0.96	0.94
Average Producer's Accuracy		0.98	0.91	0.96	

According to the cadastral database and our masking process, the main study area contains 14,953 agricultural plots occupying an area of 2949 ha. Our most accurate model detected 2373 not in production plots (345 ha), 6064 in production plots (1513 ha), and 6516 abandoned plots (1091 ha) (Figure 3.6). There is a high concentration of abandoned plots in the coastal zone due to the lower agronomic capacity of the soils and the prospects for urban growth. In the coastal area, many owners keep their plots without agricultural interest, and they expect future urban development processes. There is also a high concentration of abandoned plots adjacent to urban areas where future urban expansion is expected. The in production and not in production plots are located in the inner area, where the soils are more productive, and the crops are less exposed to the cold winds from the east (Morell-Monzó et al., 2020). According to our results, 31% of the plots classified as citrus by SIGPAC are abandoned.

The most accurate model in the main study area was tested in the three complementary areas (Figure 3.6). According to the photo-interpreted plots, the model has a high overall accuracy (93%) (Table 3.5). The high accuracy obtained indicates that the model can be applied outside the study area, although the accuracy was lower than in the training area. The most accurate class was not in production, while the most inaccurate class was the abandoned class. This is due to the spectral similarity of the abandoned and in production plots.

Table 3.5. Overall accuracy the in complementary areas.

	Overall accuracy	Producer's Accuracy			User's Accuracy		
		Non-productive	Productive	Abandoned	Non-productive	Productive	Abandoned
Bellreguard – Almoines (b)	0.98	1.00	1.00	0.93	1.00	0.94	0.91
Benicull – Polinyà del Xuquer (c)	0.88	1.00	0.93	0.70	0.86	0.85	0.95
Nules (d)	0.92	1.00	0.97	0.83	0.91	0.91	0.96
Average	0.93	1.00	0.97	0.82	0.92	0.91	0.94
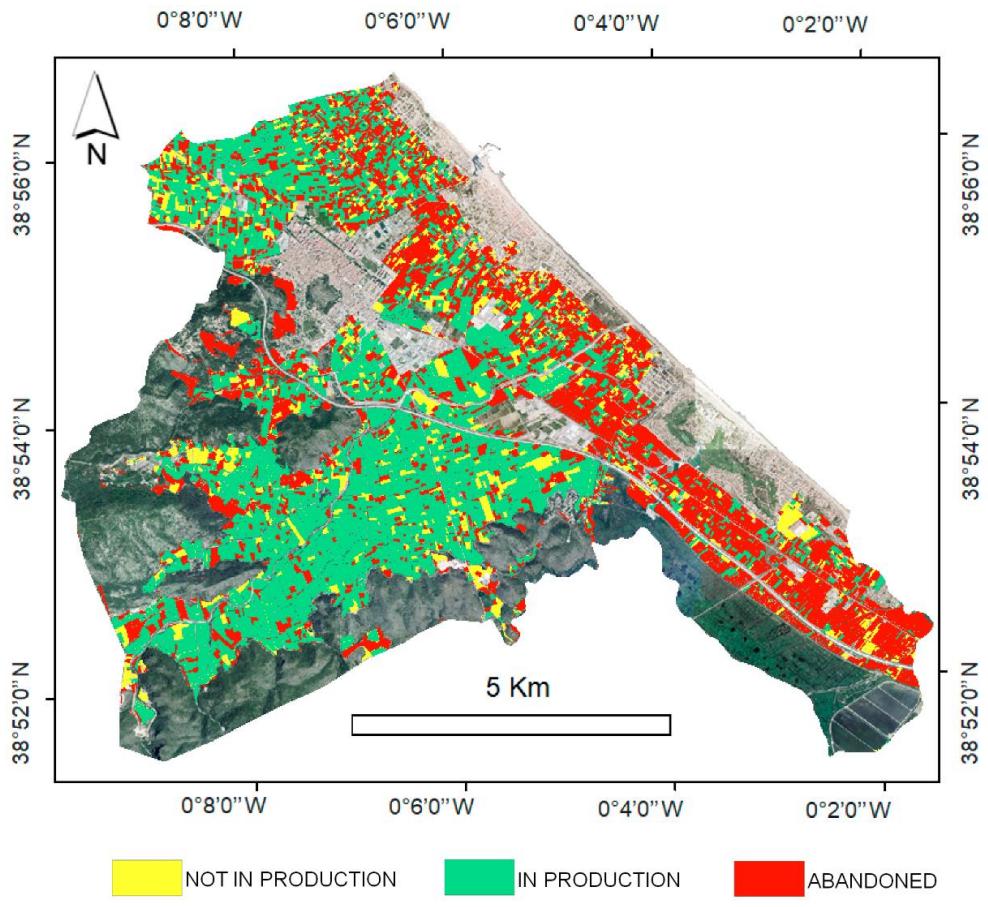
Figure 3.6. Land abandonment map of the Oliva municipality in 2019.
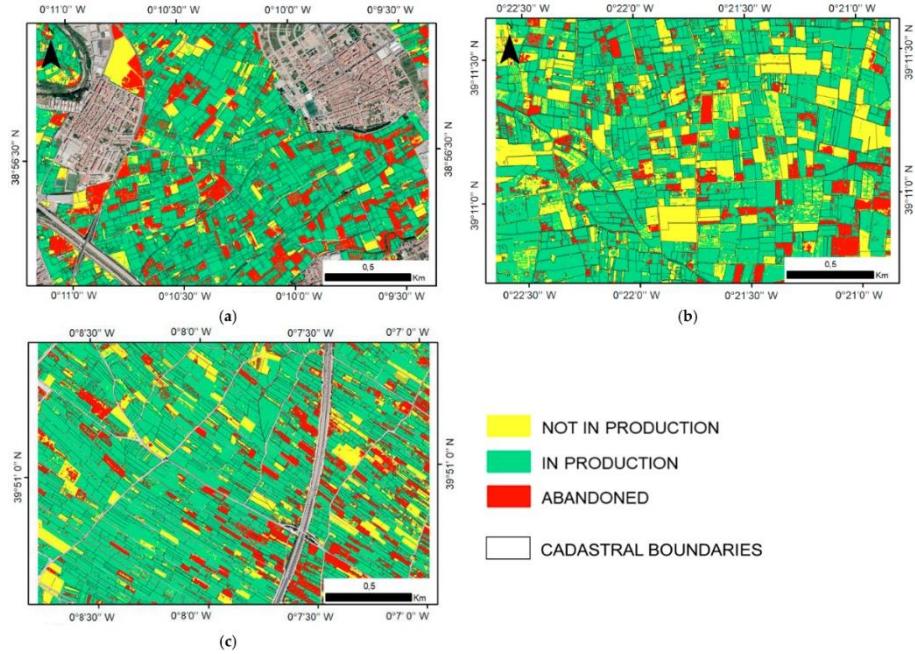
Figure 3.7. Land abandonment map obtained in the complementary areas: (a) Bellreguard / Almoines, (b) Benicull / Polinyà del Xuquer, (c) Nules.

52

3.3.3. Variables importance and model pruning

Figure 3.8 shows the variables importance ranking for each model. According to the Boruta test, all variables showed significant importance for classification. Regarding the spectral variables, all the models showed that the most important band is the NIR followed by the blue, red, and green bands. Furthermore, the NIR was always among the three most important features. Regarding the GLCM texture measures, the most important feature was the contrast for all models. The second most important GLCM texture feature was dissimilarity for the largest window sizes (5×5, 7×7 and 9×9) and mean for the smallest window model (3×3). The ASM texture feature was the least important, although it also passed the Boruta significance test.
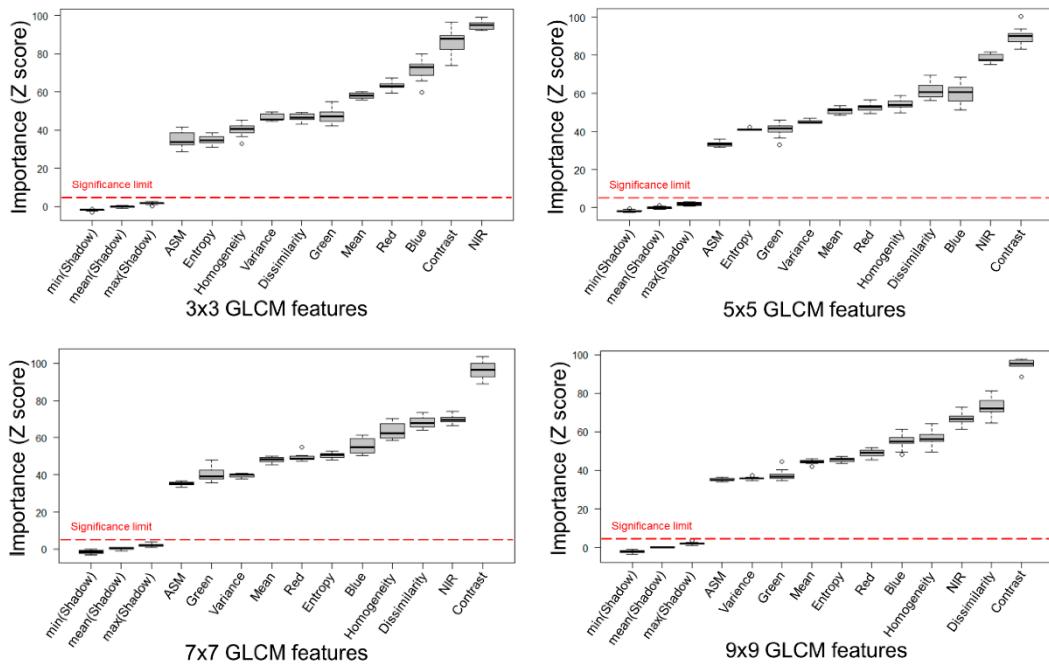
Figure 3.8. Feature importance ranking of each model. The red line indicates the limit of significance in the Boruta test.

Recursive pruning of the 9×9 window size model showed a gradual decrease in accuracy (Figure 3.9). This behaviour matches with the Boruta results which indicated that all the variables are significant for the classification. The greatest loss of accuracy occurred in the last three steps of pruning where the most important features, which are contrast and dissimilarity, are removed. The elimination of the ASM did not cause a decrease in the accuracy of the model and it could be proposed for elimination.
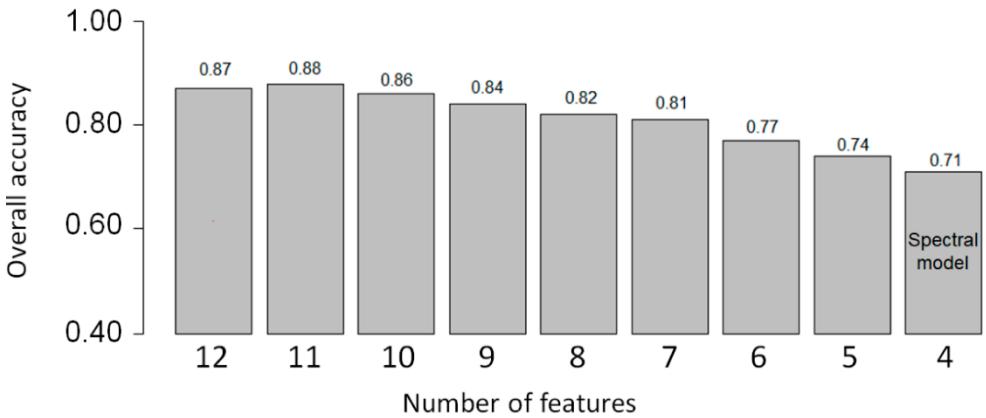
Figure 3.9. Recursive pruning of the best model (Spectral + 9×9 GLCM features). The complete model contains 11 features. These features were recursively eliminated according to the Boruta ranking without eliminating spectral features. The last pruning contains only the 4 spectral bands.

3.4. DISCUSSION

The proposed methodology was appropriate for mapping land abandonment in citrus crops at cadastral plot resolution. The combination of spectral information and texture features allowed high classification accuracies. The results confirmed that adding GLCM texture features to spectral information improves classification accuracy for detecting land abandonment. The overall accuracy of the pixel-based classification was 87% and the overall accuracy of the plot-based classification was 95%. These results represent important improvements compared to the use of Sentinel 2 images (Morell-Monzó et al., 2020). Our accuracy levels are similar to previous studies using GLCM features and the RF algorithm in VHR images in comparable contexts, such as vegetation mapping and health status, which obtained overall accuracies between 88–97% (Feng et al., 2015; Wang et al., 2015).

The best results were obtained when generating the texture features from a 9x9 window size. The model accuracy increased as the window size increased due to the reduction of salt and pepper noise. This is because as the window size increases, the texture features better capture regular citrus planting patterns. However, increasing the window size too much can decrease accuracy due to the inclusion of pixels from adjacent plots. This is an important point to consider in areas with small sized plots such as our study area where the average plot size is 0.32 ha.

All features are statistically significant for classification. However, the most important features were contrast and dissimilarity together with the NIR and blue bands. NIR and contrast were among the most important in all models. Dissimilarity increased while blue band decreased in importance as window size increased. On the other hand, the ASM feature was the least important in all the models.

The proposed methodology is an accurate and efficient alternative to generate land abandonment maps since it allows obtaining maps with high accuracy from a reduced set of training data. In addition, the limitations of Sentinel-2 for areas with high spatial fragmentation was overcome thanks to the use of VHR images (Vajsová et al., 2020; Vajsová et al., 2019; Morell-Monzó et al., 2020). We created a complete map of abandoned citrus plots in the Oliva municipality for the year 2019, that detected a 31% of erroneously classified plots in the current methodology used by the public

administration. The trained model was also tested outside the training area. However, its performance was about 3% lower than in the main study area. This difference can be explained considering that photointerpretation can offer some limitations to define the classes of the test plots, requiring additional field data for validating the complementary study areas. In addition, we recommend adjusting the window size to the needs of each area. The window size can be adjusted by evaluating the performance of the classifier in each study area or by using automatic processes to define the window size (Franklin et al., 1996). These results indicate that the proposed methodology is scalable to other citrus areas of the CV region.

The availability of an annual VHR image of the entire CV region enables to generate land abandonment maps in key areas from manageable field data sets. Furthermore, these land abandonment maps are perfectly integrated with the information from SIGPAC and other methodologies for citrus detection (Amorós et al., 2011). Combining these data sources could generate valuable information for managing the land abandonment problem and improve the citrus inventory in the CV region. This new information could correct the overestimation of citrus crop yield through SIGPAC. However, this approach is not limited only to the CV region. This approach could be implemented in other areas of the planet with citrus crops. VHR images are becoming more common. In addition to airborne images, the increasing availability of satellite images with fine spectral and spatial resolutions encourage to develop further research on the land abandonment issue. This new information would help us to improve the accuracy of the citrus inventory and annual yield, especially in areas with small plots or that have land abandonment problems. Finally, this approach could be adapted to other types of crops with regular planting pattern where abandonment is associated with the growth of wild vegetation and the loss of vigour of evergreen crops.

3.5. CONCLUSIONS

In this paper, a methodology for mapping abandoned citrus plots in highly fragmented areas from VHR images with four spectral bands (430 nm, 530 nm, 620 nm, and 720 nm) is presented. The proposed methodology combines spectral information and GLCM-based texture features from the NDVI to generate a pixel-based classification using Random Forests. Then, a unique classification value is assigned to each plot by majority voting. The best results were obtained when generating the texture features from a 9x9 window size. The proposed methodology allowed to obtain satisfactory results with an overall accuracy of 95%. Model uses 7 texture features: mean, variance, contrast, dissimilarity, homogeneity, angular second moment, and entropy. All features showed statistically significant importance for classification. However, the most important were contrast, dissimilarity, and homogeneity together with the spectral bands. This methodology allowed creating a complete map of abandoned plots in the municipality of Oliva (Comunitat Valenciana region) and it could be applied in other areas and other type of crops. This new methodological approach could correct the citrus yield overestimation generated by the current data sources available by public administrations. However, evaluating the performance of the model in areas far from the training area requires further study. This research provides new information to map small agricultural plots, especially to detect land abandonment in woody evergreen crops that have been little studied until now.

3.6. REFERENCES

Akar, Ö.; Güngör, O. Integrating multiple texture methods and NDVI to the Random Forest classification algorithm to detect tea and hazelnut plantation areas in northeast Turkey. International Journal of Remote Sensing 2015, 36, 442–464. https://doi.org/10.1080/01431161.2014.995276

Alcántara, C.; Kuemmerle, T.; Prishchepov, A.V.; Radeloff, V.C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sensing of Environment 2012, 124, 334–347. https://doi.org/10.1016/j.rse.2012.05.019

Amorós López, J.; Izquierdo Verdiguier, E.; Gómez Chova, L.; Muñoz Marí, J.; Rodríguez Barreiro, J.Z.; Camps Valls, G.; Calpe Maravilla, J. Land cover classification of VHR airborne images for citrus grove identification. ISPRS Journal of Photogrammetry and Remote Sensing 2011, 66, 115–123. https://doi.org/10.1016/j.isprsjprs.2010.09.008

Amorós-López, J.; Izquierdo-Verdiguier, E.; Gómez-Chova, L.; Muñoz-Marí, J.; Rodríguez-Barreiro, J.Z.; Camps-Valls, G.; Calpe-Maravilla, J. Multi-stage robust scheme for citrus identification from high resolution airborne images. In Image and Spatial Processing for Remote Sensing XIV, 1st ed.; Bruzzone, L., Notarnicola, C., Posa, F., Eds.; SPIE: Cardiff, UK, 2008; Volume 7109, p. 710908.

Balaguer, A.; Ruiz, L.A.; Hermosilla, T.; Recio, J.A. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. Computers & Geosciences 2010, 36, 231–240. https://doi.org/10.1016/j.cageo.2009.05.003

Breiman, L. Random Forests. Machine Learning 2001, 45, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L. Random Forests—Random Features; Technical Report 567; Statistics Department, University of California: Berkeley, CA, USA, 1999.

Campos-Taberner, M.; García-Haro, F.J.; Martínez, B.; Gilabert, M.A. Deep learning for agricultural land use classification from Sentinel-2. Revista de Teledetección 2020, 56, 35–48. https://doi.org/10.4995/raet.2020.13337

Degenhardt, F.; Seifert, S.; Szymczak, S. Evaluation of variable selection methods for random forests and omics data sets. Briefings in Bioinformatics 2017, 20, 492–503. https://doi.org/10.1093/bib/bbx124

Estel, S.; Kuemmerle, T.; Alcántara, C.; Levers, C.; Prishchepov, A.V.; Hostert, P. Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. Remote Sensing of Environment 2015, 163, 312–325. https://doi.org/10.1016/j.rse.2015.03.028

Feng, Q.; Liu, J.; Gong, J. UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. Remote Sensing 2015, 7, 1074–1094. https://doi.org/10.3390/rs70101074

Franklin, S.E.; Wulder, M.A.; Lavigne, M.B. Automated derivation of geographic window sizes for use in remote sensing digital image texture. Computers & Geosciences 1996, 22, 665–673. https://doi.org/10.1016/0098-3004(96)00009-X

Geleralitat Valenciana. Portal Estadístico de la Generalitat Valenciana. Fichas Municipales 2020. Available online: http://www.pegv.gva.es/es/fichas (accessed on 25 December 2020).

Ghorbanian, A.; Kakooei, M.; Amani, M.; Mahdavi, S.; Mohammadzadeh, A.; Hasanlou, M. Improved land cover map of Iran using Sentinel imagery withing Google Earth Engine and novel automatic workflow for land cover classification using migrated training samples. ISPRS Journal of Photogrammetry and Remote Sensing 2020, 167, 276–288. https://doi.org/10.1016/j.isprsjprs.2020.07.013

Gil-Yepes, J.L.; Ruiz, L.A.; Recio, J.A.; Balaguer-Beser, A.; Hermosilla, T. Description and validation of a new set of object-based temporal geostatistical features for land-use/land-cover change detection. ISPRS Journal of Photogrammetry and Remote Sensing 2016, 121, 77–91. https://doi.org/10.1016/j.isprsjprs.2016.08.010

Gislason, P.O.; Benediktsson, J.A.; Sveinsson, R.J. Random Forests for land cover classification. Pattern Recognition Letters 2006, 27, 294–300. https://doi.org/10.1016/j.patrec.2005.08.011

Grădinaru, S.R.; Kienast, F.; Psomas, A. Using multi-seasonal Landsat imagery for rapid identification of abandoned land in areas affected by urban sprawl. Ecological Indicators 2019, 96, 79–86. https://doi.org/10.1016/j.ecolind.2017.06.022

Hall-Beyer, M. GLCM Texture: A Tutorial v. 3.0 March 2017. PRISM University of Calgary Digital Repository 2017. http://dx.doi.org/10.11575/PRISM/33280

Hall-Beyer, M. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. International Journal of Remote Sensing 2017, 38, 1312–1338. https://doi.org/10.1080/01431161.2016.1278314

Haralick, M.R.; Shanmugan, K.; Distein, I. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics 1973, 3, 610–620. https://doi.org/10.1109/TSMC.1973.4309314

Huang, X.; Liu, X.; Zhang, L. A Multichannel Gray Level Co-Ocurrence Matrix for Multi/Hyperspectral Image Texture Representation. Remote Sensing 2014, 6, 8424–8445. https://doi.org/10.3390/rs6098424

Ilich, A. GLCMTextures. Available online: http://doi.org/10.5281/zenodo.4310187 (accessed on 25 December 2020).

Instituto Valenciano de Investigaciones Agrarias (IVIA). Citricultura Valenciana: Gestión Integrada de Plagas y Enfermedades en Cítricos. Available online: http://gipcitricos.ivia.es/citricultura-valenciana (accessed on 25 December 2020).

Jain, A.K.; Farrokhnia, F. Unsupervised texture segmentation using Gabor filters. Pattern Recognition 1991, 24, 1167–1186. https://doi.org/10.1016/0031-3203(91)90143-S

Jia, L.; Zhou, Z.; Li, B. Study of SAR Image Texture Feature Extraction Based on GLCM in Guizhou Karst Mountainous Region. In Proceedings of the 2012 2nd International Conference on Remote Sensing, Environment and Transportation Engineering, Nanjing, China, 1–3 June 2012.

Jin, Y.; Liu, X. Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: A case study of central Shandong. International Journal of Remote Sensing 2018, 39, 8703–8723. https://doi.org/10.1080/01431161.2018.1490976

Kumar, S.S.; Shaikh, T. Empirical Evaluation of the Performance of Feature Selection Approaches on Random Forests. In Proceedings of the 2017 International Conference on Computer and Applications (ICCA), Doha, United Arab Emirates, 6–7 September 2017.

Kunh, M. Caret: Classification and Regression Training, R package version 4.0.3; 2020. Available online: https://cran.r-project.org/web/packages/caret/caret.pdf (accessed on 25 December 2020).

Kursa, M.B.; Jankowski, A.; Rudnicki, W.R. Boruta—A system for Feature Selection. Fundamenta Informaricae 2010, 101, 271–285.

Kursa, M.B.; Rudncki, A. Feature Selection with Boruta Package. J. Stat. Softw. 2010. Available online: https://www.jstatsoft.org/article/view/v036i11 (accessed on 25 December 2020).

Lan, Z.; Liu, Y. Study on Multi-Scale Window Determination for GLCM Texture Description in High-Resolution Remote Sensing Image Geo-Analysis Supported by GIS and Domain Knowledge. ISPRS International Journal of Geo-Information 2018, 7, 175. https://doi.org/10.3390/ijgi7050175

Laws, K.I. Texture Image Segmentation. Ph.D. Dissertation, University Southern California, Los Angeles, CA, USA, January 1980.

Liaw, A.; Wiener, M. Classification and regression by randomForest. R News 2002, 2, 18–22.

Löw, F.; Prishchepov, F.; Waldner, F.; Dubovyk, O.; Akramkhanov, A.; Biradar, C.; Lamers, J. Mapping Cropland Abandonment in the Aral Sea Basin with MODIS Time Series. Remote Sensing 2018, 10, 159. https://doi.org/10.3390/rs10020159

Mallat, S.G. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 1989, 11, 674–693. https://doi.org/10.1109/34.192463

Ministerio de Agricultura, Pesca y Alimentación. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2019; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2019. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Ministerio de Agricultura, Pesca y Alimentación. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2008; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2008. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Ministerio de Agricultura, Pesca y Alimentación. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2018; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2018. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Morell-Monzó, S.; Estornell, J.; Sebastiá-Frasquet, M.T. Comparison of Sentinel-2 and High-Resolution Imagery for Mapping Land Abandonment in Fragmented Areas. Remote Sensing 2020, 12, 2062. https://doi.org/10.3390/rs12122062

Noguera, J. Viabilidad y competitividad del sistema citrícola valenciano. Boletín de la Asociación de Geógrafos Españoles 2010, 52, 81–99.

Pal, M. Random Forest classifier for remote sensing classification. International Journal of Remote Sensing 2005, 26, 217–222. https://doi.org/10.1080/01431160412331269698

Ramírez, M.; Martínez, L.; Montilla, M.; Sarmiento, O.; Lasso, J.; Diaz, S. Obtaining agricultural land cover in Sentinel-2 satellite images with drone image injection

using Random Forest in Google Earth Engine. Revista de Teledetección 2020, 56, 35–48. https://doi.org/10.4995/raet.2020.14102

R Core Team. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Viena, Austria, 2020; Available online: https://www.R-project.org/ (accessed on 25 December 2020).

Recio, J.A.; Hermosilla, T.; Ruiz, L.A.; Palomar, J. Automated extraction of tree and plot-based parameters in citrus orchards from aerial images. Computers and Electronics in Agriculture 2013, 90, 24–34. https://doi.org/10.1016/j.compag.2012.10.005

Rounsevell, M.D.A.; Reginster, I.; Araújo, M.B.; Carter, T.R.; Dendoncker, N.; Ewert, F.; House, J.I.; Kankaanpää, R.; Leemans, R.; Metzger, M.J.; et al. A coherent set of future land use change scenarios for Europe. Agriculture, Ecosystems and Environment 2006, 114, 57–68. https://doi.org/10.1016/j.agee.2005.11.027

Ruiz, L.A.; Recio, J.A.; Fernández-Sarría, A.; Hermosilla, T. A feature extraction software tool for agricultural object-based image analysis. Computers and Electronics in Agriculture 2011, 76, 284–296. https://doi.org/10.1016/j.compag.2011.02.007

Schrivastava, R.J.; Gebelein, J.L. Land cover classification and economic assessment of citrus groves using remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 2007, 61, 341–353. https://doi.org/10.1016/j.isprsjprs.2006.10.003

Stumpf, A.; Kerle, N. Object-oriented mapping of landslides using Random Forests. Remote Sensing of Environment 2011, 115, 2564–2577. https://doi.org/10.1016/j.rse.2011.05.013

Vajsová, B.; Fasbender, D.; Wirnhardt, C.; Lemajic, S.; Devos, W. Assessing Spatial Limits of Sentinel-2 Data on Arable Crops in the Context of Checks by Monitoring. Remote Sensing 2020, 12, 2195. https://doi.org/10.3390/rs12142195

Vajsová, B.; Fasbender, D.; Wirnhardt, C.; Lemajic, S.; Sima, A.; Astrand, P. Applicability Limits of Sentine-2 Data Compared to Higher Resolution Imagery for CAP Checks by Monitoring; JRC Technical Report: JRC115564; Publications Office of the European Union: Luxembourg, 2019; ISBN 978-92-76-01935-0.

Viñals, M.J. El Marjal de Oliva-Pego: Geomorfología y Evolución de un Humedal Costero Mediterráneo, 1st ed.; Conselleria de Agricultura y Medio Ambiente, Generalitat Valenciana: Valencia, Spain, 1996.

Viñals, M.J. Secuencias Estratigráficas y Evolución Morfológica del Extremo Meridional del Golfo de Valencia (Cullera-Dénia). In El Cuaternario del País Valenciano, 1st ed.; Universitat de València-AEQUA: Valencia, Spain, 1995.

Wang, H.; Zhao, Y.; Pu, R.; Zhang, Z. Mapping Robinia Pseudoacacia Forest Health Conditions by Using Combined Spectral, Spatial, and Textural Information Extracted from IKONOS Imagery and Random Forest Classifier. Remote Sensing 2015, 7, 9020–9044. https://doi.org/10.3390/rs70709020

Wang, L.; He, D.-C. Texture classification using texture spectrum. Pattern Recognition 1990, 23, 905–910. https://doi.org/10.1016/0031-3203(90)90135-8

Wulder, M.A.; Coops, N.C.; Roy, D.P.; White, J.C.; Hermosilla, T. Land cover 2.0. International Journal of Remote Sensing 2018, 39, 4254–4284. https://doi.org/10.1080/01431161.2018.1452075

Yin, H.; Prishchepov, A.V.; Kuemmerle, T.; Bleyhl, B.; Buchner, J.; Radeloff, V.C. Mapping agricultural land abandonment from spatial and temporal segmentation of

Landsat time series. Remote Sensing of Environment 2018, 210, 12–24. https://doi.org/10.1016/j.rse.2018.02.050

CAPÍTULO 4

Cartografía del abandono de cultivos de cítricos mediante el uso de datos altimétricos: LiDAR y fotogrametría SfM

4.1. INTRODUCCIÓN

El abandono de tierras agrícolas es un fenómeno creciente a nivel mundial con relevantes implicaciones medioambientales (ej. biodiversidad, captación de CO2, incendios forestales, recursos hídricos, etc.) y socio-económicas (ej. Producción alimentaria, sustento y paisaje) (Prishchepov, 2020). El abandono de tierras agrícolas ocurre frecuentemente en áreas montañosas y en zonas con estructuras parcelarias fragmentadas (Czesak et al., 2021). Además, según el tipo de cultivo, el contexto social y ambiental, el abandono de tierras se puede percibir como un problema o como una oportunidad (Czesak et al., 2021). Algunos de los impactos negativos que se pueden deriver del abandono de tierras están relacionados con la seguridad alimentaria, la degradación del paisaje, la proliferación de especies invasoras o la pérdida de aspectos sociales y culturales (Subedi et al., 2021). Por otra parte, pueden producirse impactos positivos como el incremento en la fijación de CO2, la mejora de la calidad del suelo, la creación de nuevos hábitats y el aumento de la biodiversidad (Subedi et al., 2021). Sin embargo, algunos de estos impactos pueden ser tanto positivos como negativos dependiendo del contexto donde se produce el abandono de tierras (Subedi et al., 2021). Estas características hacen del abandono de tierras un fenómeno complejo que debe ser estudiado a distintas escalas (global, regional/nacional y local).

España es el mayor productor de cítricos de Europa, y la Comunitat Valenciana alberga más del 50% de la producción nacional (MAPA, 2020). Sin embargo, la Comunitat Valenciana es una de las regiones donde se está produciendo un mayor abandono de tierras en las últimas décadas. Desde el año 2000 hasta el 2020 se ha producido una disminución en torno al 20% de la superficie cultivada (MAPA, 2020). En 2020, la superficie dedicada al cultivo de cítricos se estimaba en 159.140 ha (MAPA, 2020). Las causas que han propiciado el abandono de tierras son complejas, sin embargo, se podrían resumir en: a) competencia con productores extranjeros fuera de la Unión Europea, b) aparición de usos del suelo más rentables (usos urbanos), c) políticas de importación europeas desfavorables, d) ineficiencias propias del sector citrícola valenciano (Compés et al., 2019). En los próximos años se espera un aumento de las zonas abandonadas dentro y fuera de la Comunidad Valenciana debido a los cambios socio-económicos que se están produciendo en la Unión Europea (Rounsevell et al., 2006; Subedi et al., 2021).

La teledetección es una potente herramienta que nos permite observar los cambios en el uso de la tierra relacionados con el abandono de tierras agrícolas, desde un punto de vista espacial y temporal, a diferentes escalas. Los esfuerzos previos para cartografiar el abandono de tierras se han centrado en la monitorización de cultivos estacionales en grandes áreas que no requieren una resolución espacial fina. Con este fin se han utilizado series temporales MODIS y Landsat (ej. Alcantara et al., 2012; Estel et al., 2015; Löw et al., 2018; Yin et al., 2018; Grădinaru et al., 2019; Prishchepov et al., 2012; Dara et al., 2018). Las imágenes Sentinel-2 han permitido abordar el problema a una mayor resolución espacial (ej. Szostak et al., 2017), sin embargo, la monitorización de cultivos a nivel de parcela en paisajes fragmentados sigue siendo un desafío utilizando este tipo de imágenes (Vajsová et al., 2019 y 2020; Morell-Monzó et al.,

2020). Además, la monitorización de cultivos permanentes, especialmente los de hoja perenne como los cítricos, es aún más compleja debido a la baja separabilidad espectral entre las parcelas abandonadas y activas. Aunque en una parcela de cítricos abandonada generalmente se produce una pérdida de vigor en los árboles y un crecimiento de la vegetación silvestre, estas diferencias son difíciles de captar utilizando imágenes de resolución moderada. Estas circunstancias han motivado el uso de imágenes de alta resolución (sub-métricas) que permiten la extracción de texturas (ZuritaMilla et al., 2017; Neigh et al., 2018; Zhang et al., 2020) y el análisis de patrones espaciales (GilYepes, et al., 2016; Recio et al., 2013). Teniendo en cuenta estos datos, Amorós-López et al. (2011) y Morell-Monzó et al. (2021) consiguieron mejorar los resultados utilizando imágenes de alta resolución para la monitorización de parcelas de cítricos. Sin embargo, hasta donde sabemos, no se ha explorado el uso de datos altimétricos para la monitorización de cultivos perennes. Sí existe una experiencia previa, los trabajos de Kolecka et al. (2015) y Czesak et al. (2021) que utilizaron datos LiDAR para cartografiar el abandono de cultivos estacionales y sucesiones forestales secundarias en zonas con paisajes altamente fragmentados.

En este trabajo se realiza una primera exploración del uso de datos altimétricos para la detección de cultivos de cítricos abandonados. El objetivo del trabajo es explorar el uso de los datos LiDAR del Plan Nacional de Ortofotografía Aérea (PNOA) y datos altimétricos utilizando un vehículo aéreo no tripulado (UAV) y generados mediante procesos fotogramétricos Structure from Motion (SfM) para cartografiar el estado (en producción, abandonado, no en producción) de parcelas de cítricos.

Se utilizó el algoritmo Random Forests (Breiman, 2001) como método de clasificación. Sin embargo, el enfoque utilizado para la extracción de características fue adaptado a la naturaleza y características de cada fuente de información. Aunque existen muchos algoritmos de aprendizaje automático que han sido usados en el campo de la teledetección, Random Forests se ha convertido en uno de los más utilizados (Sheykhmousa et al., 2020). Esto se debe a sus características en lo que respecta a exactitud de clasificación, facilidad de uso (por su reducido número de hiperparámetros), robustez frente al sobreajuste (por su aprendizaje no secuencial) y buena escalabilidad frente a conjuntos de datos grandes (por su baja complejidad de tiempo de entrenamiento y la capacidad de ser paralelizado (Belgiu y Drăguţ 2016; Pal, 2005). Estas características lo han convertido en uno de los algoritmos estándar para clasificaciones supervisadas en problemas de teledetección que no requieren soluciones basadas en aprendizaje profundo o modelos de ensamble más complejos. El estudio se llevó a cabo en una zona de la comarca de La Safor (Comunitat Valenciana, España), donde se está produciendo un acelerado abandono de los cultivos de cítricos.

4.2. MATERIAL Y MÉTODOS

4.2.1. El área de estudio

El área de estudio se enmarca en la comarca de La Safor, una zona costera de la Comunidad Valenciana (España). En esta zona coexisten diversos usos del suelo, como el uso urbano, forestal y agrícola, en el cual se centra este estudio. El cultivo mayoritario de esta zona son los cítricos, que ocupan más del 95% de la superficie agrícola (GVA, 2020). La superficie agrícola se ubica en la llanura costera la cual se

caracteriza por su relieve plano. Esta zona destaca por suelos con una alta capacidad agronómica y una alta disponibilidad de agua que han convertido a la comarca en una zona históricamente agrícola. Sin embargo, en los últimos años se han abandonado un gran porcentaje de explotaciones agrícolas. La estructura agraria se caracteriza por el pequeño tamaño de las parcelas (promedio 0,25-0,50 ha) y la alta fragmentación del paisaje. El estudio fue realizado en una zona costera entre los municipios de Gandia y Oliva. Se tomaron datos en 12 zonas que ocupan una superficie aproximada de 1300 ha (figura 4.1). Se muestrearon 12 localizaciones con el objetivo de cubrir la variabilidad ambiental de la zona.
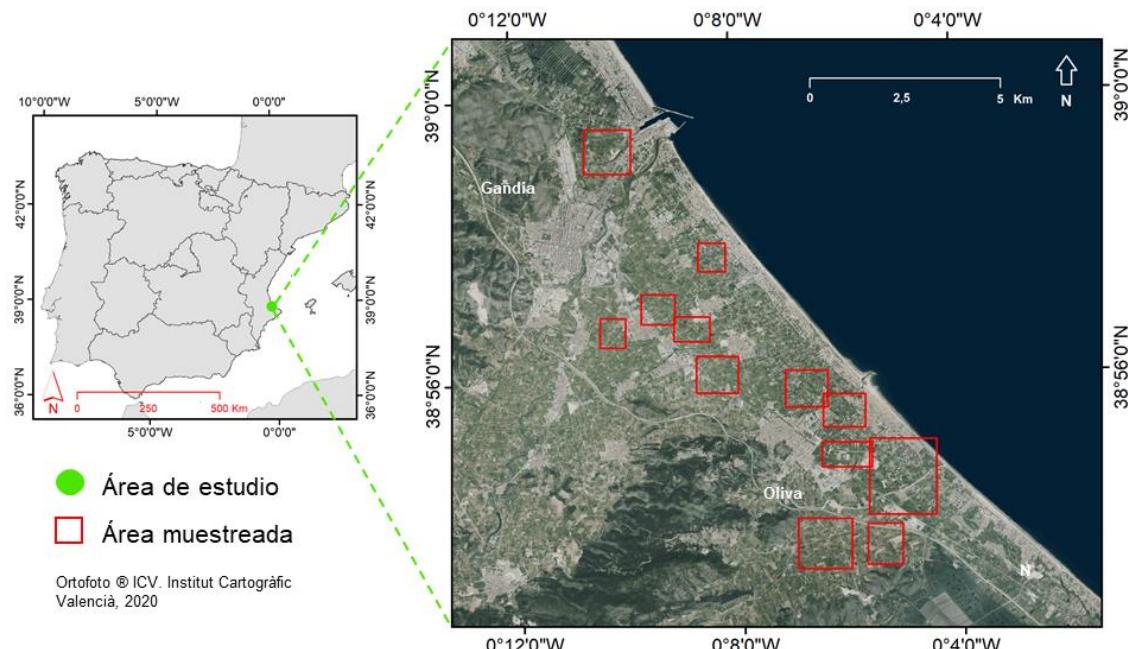
Figura 4.1. Área de estudio en la comarca de La Safor (Comunitat Valenciana, España).

4.2.2. Datos

Con el objetivo de generar información útil para la detección del abandono de cultivos de cítricos y su gestión se definió una clasificación basada en tres tipos de parcelas: no productiva (NP), productive (PR) y abandonada (AB) como se muestra en la figura 4.2. La identificación de la categoría AB tiene interés para el seguimiento del abandono de los cultivos. Además, la diferenciación entre las categorías NP y PR permite realizar predicciones de producción de cítricos más precisas. Las parcelas de cítricos pueden ser identificadas a partir de la base de datos del Sistema de Información Geográfica de Parcelas Agrícolas (SIGPAC). La clasificación propuesta en este trabajo se diseñó para complementar la información de SIGPAC y ser aplicada sobre las parcelas de cítricos previamente identificadas.

Los datos verdad terreno se obtuvieron mediante campañas de campo, fotografías de drones y fotointerpretación de ortofotos. Se clasificaron un total de 400 parcelas citrícolas según la base de datos de SIGPAC; 120 de ellas fueron usadas para entrenar y validar el modelo de clasificación basado en datos LiDAR del PNOA y 280 fueron utilizadas para entrenar y validar el modelo de clasificación basado en datos altimétricos derivados de fotogrametría SfM. El conjunto de 280 parcelas fue generado mediante visitas de campo entre febrero y marzo de 2021 y revisión de fotografías tomadas por el

UAV. Por otro lado, el conjunto de 120 parcelas fue generado mediante fotointerpretación sobre ortofotos de octubre de 2015. El objetivo inicial era fotointerpretar la clase para el año 2015 de las mismas parcelas que fueron visitadas en el campo. Sin embargo, la clasificación de las parcelas PR y AB mediante fotointerpretación no permitió conocer la clase real de todas las parcelas. Además, un conjunto significativo de parcelas había cambiado de categoría durante el periodo 2015-2021, por lo que finalmente se seleccionó un conjunto más pequeño de 120 parcelas cuya clase si se pudo asegurar mediante fotointerpretación. Los límites parcelarios se obtuvieron a partir de SIGPAC.

Figura 4.2. Clasificación utilizada: no productiva–NP (izquierda), productiva–PR (centro) y abandonada–AB (derecha).

En relación a la nube de puntos LiDAR, se utilizaron datos del proyecto PNOA (PNOA 2015 CC BY 4.0 www.scne.es). En el marco de este proyecto se genera una cobertura LiDAR de todo el territorio español utilizando un sistema LiDAR aéreo con una frecuencia aproximada de 6 años desde el año 2009. En este estudio se utilizaron los datos de la 2ª cobertura correspondientes al año 2015. Para nuestra zona de estudio los datos LiDAR fueron tomados en octubre de 2015. Estos datos se distribuyen de forma gratuita en áreas de 2×2 km. La nube de puntos de la 2ª cobertura tiene una densidad de puntos de 0,5- 4 puntos/m2 y una precisión altimétrica RMSEz < 20 cm. Además del valor posicional (x, y, z), los archivos incluyen los atributos de intensidad, número de retorno, cantidad de retornos, valores de clasificación, número de pasada, color RGB, tiempo GPS, ángulo y dirección de escaneo para cada punto.

La nube de puntos fotogramétrica se obtuvo a través de 12 vuelos con un UAV realizados en febrero de 2021. El vehículo no tripulado fue un senseFly eBee X que incorpora una cámara S.O.D.A. RGB 20,1MP y un sistema de posicionamiento RTK/PPK (error 3-8 cm). Los vuelos fueron realizados a una altura de 123,7 m/AED en tomas nadirales. Se tomaron un total de 5510 fotografías con un solape lateral de 60% y un solape en el sentido del vuelo de 80%. Las fotografías se procesaron utilizando la técnica SfM en Agisoft PhotoScan Professional Edition software (Agisoft LLC) version 1.2.5 build 2735. No se utilizaron puntos de control adicionales debido al sistema de posicionamiento del RTK/PPK equipado en el UAV. Se obtuvo un error de alineación <6 cm. Después se generó una nube de puntos densa (325 puntos/m^2) con un total aproximado de 4723 millones de puntos ocupando un espacio de memoria de 91 Gb.

4.2.3. Procesamiento

La nube de puntos LiDAR con una densidad promedio 0,8 puntos/m2 fue reclasificada para extraer los puntos correspondientes al suelo. Se utilizó el método Progressive

Morphological Filter (PMF) (Zhang et al., 2003) para la identificación de los puntos correspondientes al suelo. Se probó también el método Cloth Simulation Filter (Zhang et al., 2016). Este método es computacionalmente más eficiente, pero proporcionó peores resultados según la interpretación visual de los resultados, por lo que se optó por el método PMF. Los puntos de suelo fueron interpolados para crear un modelo digital del terreno rasterizado con 1 m de resolución siguiendo el enfoque de k vecinos más cercanos con una ponderación de distancia inversa (k=2). Este método de interpolación es uno de los métodos estándar en muchas librerías de software. Después, la altura de la nube de puntos fue normalizada al suelo. Finalmente, se realizó una extracción de estadísticas descriptivas de la altura utilizando la nube de puntos y los límites parcelarios para realizar una predicción wall-to-wall de cada parcela. Siguiendo este enfoque basado en objetos, se extrajeron un total de 13 características de cada parcela. Las características extraídas fueron: promedio, desviación estándar, mínimo, máximo y rango de Z; asimetría, curtosis y entropía de Z; número de primeros retornos dividido entre el número de retornos por encima de 2 m, porcentaje de retornos sobre la Z promedio, número de primeros retornos dividido entre el número total de retornos y Leaf Área Index (LAD) (Bouvier et al., 2015), donde Z es altura. Las características extraídas fueron utilizadas para entrenar un modelo Random Forests con 200 árboles de decisión y 1/3 de las muestras como muestras out-of-bag. La figura 4.3 muestra el flujo de trabajo para el procesamiento de los datos LiDAR.

La nube de puntos fotogramétrica con una densidad de 325 puntos/m2 fue homogenizada y se redujo su densidad para reducir el tiempo de procesamiento. La homogenización se realizó a una densidad objetivo de 50 puntos/m^2. Después se realizó una clasificación de los puntos del suelo utilizando el mismo método aplicado para los datos LiDAR. Los puntos de suelo fueron interpolados para crear un modelo digital del terreno rasterizado con 1 m de resolución siguiendo el enfoque de k vecinos más cercanos con una ponderación de distancia inversa (k=20). Después se calculó el modelo digital de superficie utilizando el método pit-free (Khosravipour et al., 2014). Finalmente se calculó el Canopy Heigh Model (CHM) y 9 características a partir de la nube de puntos normalizada: promedio, desviación estándar, máximo, mínimo y rango de Z, varianza, asimetría, curtosis y entropía de Z. Adicionalmente, el CHM fue utilizado para calcular 7 características de textura basadas en la Matriz de Co-Ocurrencia de Nivel de Gris (GLCM), adaptando el enfoque propuesto en Morell-Monzó et al. (2021). Las características extraídas fueron: promedio, varianza, homogeneidad, contraste, disimilitud, entropía y segundo momento angular, que fueron calculadas a 12 tamaños de ventana distintos (3×3, 5×5, 7×7, 9×9, 11×11, 13×13, 15×15, 17×17, 19×19, 21×21, 23×23 y 25×25). Las características de textura fueron calculadas en una dirección espacial invariante, que es el promedio de las cuatro direcciones 0°, 45°, 90° y 135° a partir de una discretización del CHM a 64 niveles de gris. La discretización del CHM se realizó en intervalos uniformes entre el mínimo y el máximo. La extracción de características de textura a partir de CHM ha sido utilizada por otros autores de forma exitosa (ej. Ozdemir y Donoghue, 2013; Niemi y Vauhkonen, 2016).

Las 9 características extraídas de la nube de puntos, junto con el CHM y las 7 características de textura fueron apiladas en un fichero ráster de 1 m de resolución. Estas bandas fueron utilizadas para entrenar un modelo Random Forests con 200 árboles de decisión y 1/3 de las muestras como muestras out-of-bag. El modelo Random

Forests permitió realizar una clasificación de los píxeles de la imagen generando una imagen segmentada semánticamente. Finalmente, los límites de las parcelas se utilizaron para mejorar la clasificación basada en píxeles utilizando el voto mayoritario de los píxeles de cada parcela. Este procedimiento crea un mapa basado en objetos donde cada objeto es una parcela (Ghorbanian et al., 2020) y mejora la precisión basada en píxeles eliminando píxeles mal clasificados, dentro de segmentos homogéneos, conocido como ruido de sal y pimienta (Wulder et al., 2018). Esta estrategia post-clasificación produjo mejoras de rendimiento en nuestros trabajos anteriores (Morell-Monzó et al., 2020; Morell-Monzó et al., 2021). La figura 4.4 muestra el flujo de trabajo para el procesamiento de los datos fotogramétricos.
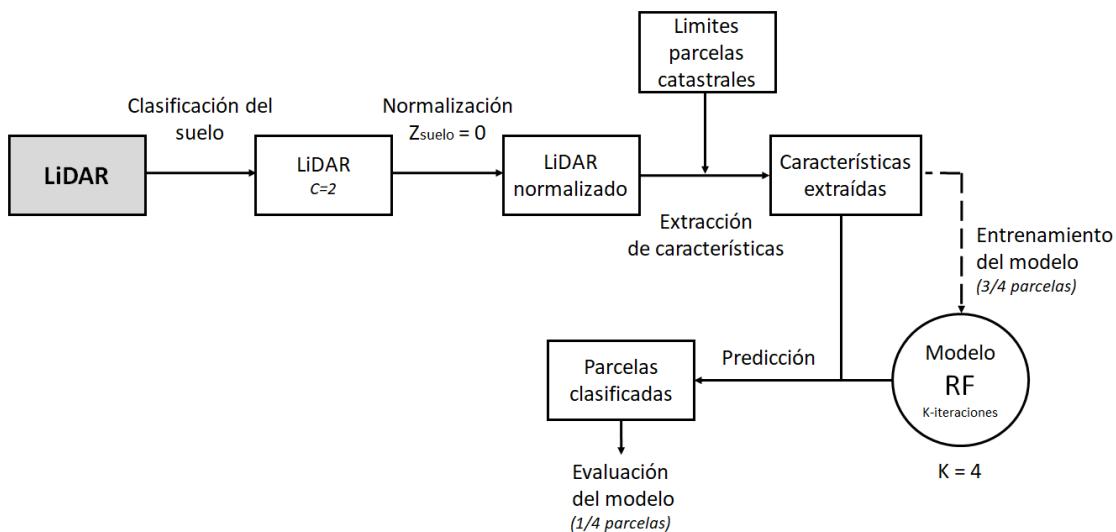
Figura 4.3. Flujo de trabajo para el procesamiento de los datos LiDAR.
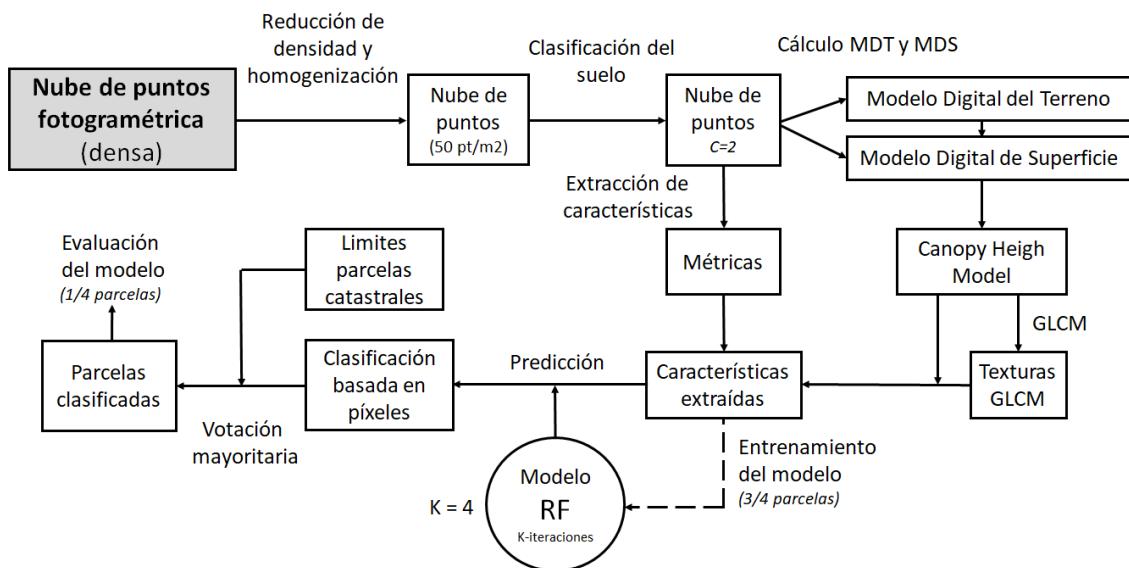
Figura 4.4. Flujo de trabajo para el procesamiento de los datos fotogramétricos.

La lectura y procesamiento de las nubes de puntos se realizó utilizando el paquete `lidR` (Roussel y Auty, 2021). La lectura y manipulación de archivos raster y vectoriales se realizó utilizando los paquetes `raster` (Hijmans, 2021) y `rgdal` (Bivand et al.,

2021) respectivamente, y la extracción de características de textura se realizó utilizando el paquete `glcm` (Zvoleff, 2020) en un entorno R 3.6.1 (R Core Team, 2019).

4.2.4. Evaluación

Se evaluó el rendimiento de clasificación para cada fuente de información utilizada (LiDAR y fotogrametría SfM). La evaluación de la precisión se realizó a través de una validación cruzada de 4 iteraciones. En cada iteración se utilizaron ¾ de las parcelas para entrenamiento y las parcelas restantes (¼), que nunca fueron vistas por el modelo, fueron utilizadas como conjunto de test. Para la validación cruzada de los datos LiDAR, en cada iteración, se utilizaron 90 parcelas para entrenamiento y 30 para test. Para los datos derivados de fotogrametría SfM, en cada iteración, se utilizaron 210 parcelas para entrenamiento y 70 para test. Se calculó la exactitud global, índice Kappa de Cohen y métricas derivadas de la matriz de confusión (precisión y exhaustividad) utilizando el promedio de las 4 iteraciones de la validación cruzada. La exactitud global es la medida más simple y representa el número de muestras correctamente clasificadas entre el total de muestras de test. Debido a que la clasificación propuesta consta de pocas categorías, la probabilidad de acierto utilizando un modelo aleatorio es alta. Por este motivo se incluyó el índice Kappa que compara la exactitud observada con la exactitud esperada (probabilidad aleatoria). Se incluyeron también la precisión y la exhaustividad de cada clase, que corresponden a la exactitud del productor y del usuario respectivamente.

4.3. RESULTADOS

Todos los modelos Random Forests creados convergieron antes de los 200 árboles de decisión. El hiperparámetro `mtry`, que define el número de características en cada embolsamiento, fue ajustado a 6 para el conjunto de características derivadas del LiDAR y a 5 para el conjunto de características derivadas de la nube de puntos fotogramétrica.

Los datos LiDAR obtuvieron una exactitud global de clasificación del 67,9%, con un índice Kappa de 0,52. La precisión para las parcelas NP y AB fue del 65% y del 67% respectivamente, mientras que la clase PR obtuvo una precisión del 73%. Los datos LiDAR tuvieron mayor capacidad para clasificar las parcelas PR que las parcelas NP y AB. Esto se debe a la semejanza altimétrica que hay entre las parcelas no productivas y las parcelas abandonadas con vegetación muy baja, que son frecuentes en nuestro conjunto de datos. Los valores de precisión y exhaustividad también muestran esta tendencia (tabla 4.1).

Tabla 4.1. Rendimiento de clasificación de los datos LiDAR mediante una validación cruzada de 4 iteraciones.

LiDAR	Validación cruzada (k=4)		
	No productivo	Productivo	Abandonado
Precisión	0,6507	0,7326	0,6713
Exhaustividad	0,6750	0,7000	0,6625
Exactitud global		0,6792	

Kappa de Cohen	0,5188

Los datos derivados de la nube de puntos fotogramétrica mostraron un rendimiento de clasificación mayor que los datos LiDAR. Se obtuvo una exactitud global máxima de 83,56%. No obstante, la exactitud global de clasificación varió dependiendo del tamaño de ventana utilizado para calcular las características de textura entre el 79,29-83,57% (tabla 4.2) con un promedio de 80,89%. El refinamiento de la clasificación utilizando la votación mayoritaria dentro de cada parcela produjo una mejora del rendimiento a todos los tamaños de ventana con una mejora promedio del 1,76% (tabla 4.2).

Tabla 4.2. Exactitud global de clasificación basada en píxeles (antes de aplicar el voto mayoritario por parcela) y basada en parcelas (después de aplicar el voto mayoritario por parcela).

GLCM kernel	Exactitud global (pixel)	Exactitud global (parcela)
3x3	0,7588	0,7929
5x5	0,7674	0,8000
7x7	0,7757	0,7929
9x9	0,7831	0,8072
11x11	0,7883	0,8286
13x13	0,7966	0,8357
15x15	0,7984	0,8143
17x17	0,8044	0,8072
19x19	0,8055	0,8072
21x21	0,8056	0,8072
23x23	0,8058	0,8072
25x25	0,8060	0,8072
max	0,8060	0,8357
min	0,7588	0,7929
promedio	0,7913	0,8089

El mejor rendimiento se obtuvo al extraer las características de textura a un tamaño de ventana de 13×13. A este tamaño de ventana, la exactitud global de la clasificación basada en píxeles fue del 79,65% (Kappa 0,66) y el refinamiento mediante votación mayoritaria aumentó el rendimiento hasta una exactitud global de 83,57% (Kappa 0,74), con una precisión de 83,96 para la clase NP, 92,88 para la clase PR y 76,58 para la clase AB. La categoría más fácilmente detectable por esta fuente de información fue también PR. Sin embargo, las parcelas NP fueron más fácilmente detectables que las AB (tabla 4.3).

Tabla 4.3. Rendimiento de clasificación de los datos fotogramétricos mediante una validación cruzada de 4 iteraciones.

SfM	Validación cruzada (k=4)		
	No productivo	Productivo	Abandonado
Precisión	0,8396	0,9288	0,7658
Exhaustividad	0,5000	0,9658	0,8800
Exactitud global		0,8357	
Kappa de Cohen		0,7379	

A continuación, la figura 4.5 muestra la predicción de los datos LiDAR para el año 2015 frente a la predicción de los datos fotogramétricos SfM para 2021. Se pueden observar diferencias entre ambas predicciones debido tanto a los cambios en los usos del suelo producidos entre 2015 y 2021 como por las diferencias de exactitud de clasificación de ambas fuentes de datos utilizadas.
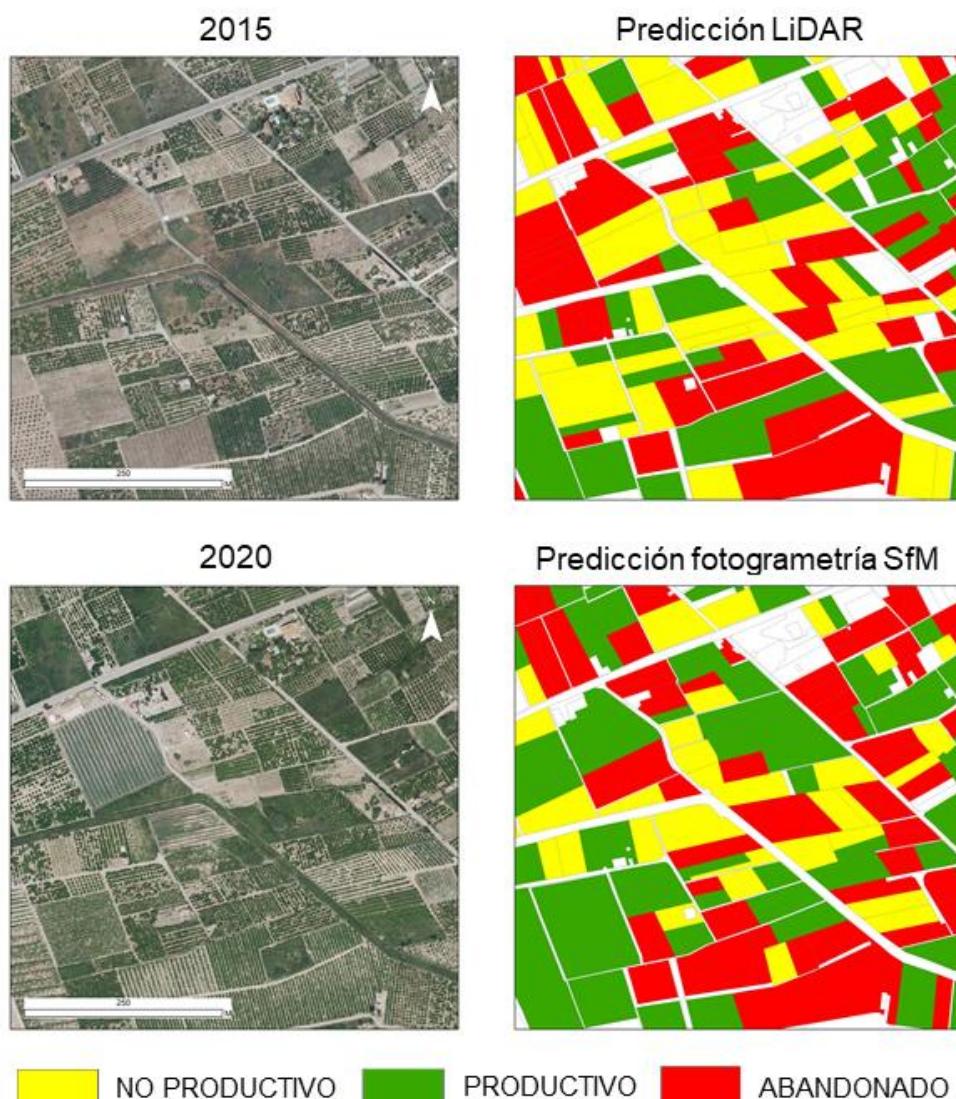
Figura 4.5. Detalle de predicción obtenida a partir de los datos LiDAR (2015) frente a la obtenida por los datos fotogramétricos SfM (2021).

4.4. DISCUSIÓN

El estudio evaluó el uso de datos altimétricos para la detección de parcelas de cítricos abandonadas. Se evaluó el rendimiento de clasificación de tres tipos de parcelas útiles para el estudio del abandono de tierras y para la estimación de rendimientos agrícolas. Los resultados obtenidos mostraron un rendimiento de clasificación bajo para los datos LiDAR (68%), con una resolución de 0,8 puntos/m^2. Sin embargo, los datos altimétricos derivados de un proceso fotogramétrico SfM obtuvieron una precisión de clasificación del 83,56%. Este mejor rendimiento podría ser explicado teniendo en cuenta que, debido a la alta densidad de la nube de puntos fotogramétrica, se pudieron extraer características de textura a partir del CHM. Nuestras experiencias previas (Morell-Monzó et al., 2021) mostraron que las características de textura mejoran el rendimiento de clasificación para monitorizar parcelas de cítricos. Los datos LiDAR, con una resolución de 0,8 puntos/m^2, no permitieron la extracción de características de textura.

Si comparamos los resultados obtenidos con estudios anteriores, observamos las siguientes diferencias en cuanto a rendimiento. Para la detección de parcelas abandonadas de cítricos, las imágenes de alta resolución (sub-métricas) con 4 bandas espectrales (R, G, B, NIR) ofrecieron un mejor rendimiento (exactitud global: 93-95%) (Morell-Monzó et al., 2021) que ambas fuentes de información altimétrica. Sin embargo, los datos altimétricos obtenidos mediante fotogrametría mostraron un rendimiento similar a un enfoque basado en una única imagen Sentinel-2 (exactitud global: 77%) (Morell-Monzó et al., 2020). En relación con los estudios de Kolecka et al. (2015) y Czesak et al. (2021) que utilizan datos LiDAR de una resolución mayor (4-12 puntos/m^2), la exactitud global para la detección del abandono de cultivos estacionales fue del 82% (Czesak et al., 2021) mientras que para la monitorización de sucesiones forestales secundarias fue del 95% (Kolecka et al., 2015). Estos mejores rendimientos pueden explicarse por la mayor densidad de puntos LiDAR utilizados en estos trabajos en comparación a los del PNOA aplicados en nuestro estudio (0,8 puntos/ m^2) y teniendo en cuenta que la variación que se produce en los cultivos estacionales abandonados facilita su detección frente a una cubierta de árboles frutales perennes (cítricos). Los mejores resultados obtenidos a partir de la nube de puntos derivada por procesos fotogramétricos basados en SfM y en otros estudios en los que utilizan datos LiDAR de mayor densidad sugieren que este parámetro podría ser de interés para aportar resultados más precisos a los obtenidos con los datos LiDAR utilizados en nuestro trabajo.

4.5. CONCLUSIONES

Este trabajo proporciona una de las primeras experiencias en el uso de datos altimétricos para cartografiar el abandono de tierras agrícolas de cultivos perennes. El estudio calcula el rendimiento de datos LiDAR (0,8 puntos/m^2) y de datos altimétricos derivados de un proceso de fotogrametría SfM para el caso específico del abandono de cultivos de cítricos. El potencial de los datos altimétricos es especialmente notorio en el caso de las nubes de puntos fotogramétricas Structure from Motion de alta densidad. Los resultados mostraron una exactitud global del 67,9% para los datos LiDAR y de un

83,6 % de los datos fotogramétricos. La extracción de características de textura a partir del Canopy Heigh Model fue clave para identificar parcelas no productivas, productivas y abandonadas. Investigaciones futuras deberían explorar la fusión de datos espectrales, texturales y altimétricos para el estudio de parcelas de cítricos y otros cultivos perennes y analizar el rendimiento de las futuras coberturas de datos LiDAR con una mayor densidad de puntos.

4.6. REFERENCIAS

Alcantara, C., Kuemmerle, T, Prishchepov, A. V., Radeloff, V. C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sensing of Environment 2012, 124, 334-347. https://doi.org/10.1016/j.rse.2012.05.019

Amorós-Lopez, J., Izquierdo-Verdiguier, E., GómezChova, L., Muñoz-Marí, J., Rodríguez-Barreiro J. Z., Camps-Valls, G., Calpe-Maravilla, J. Land cover classification of VHR airborne images for citrus grove identification. ISPRS Journal of Photogrammetry and Remote Sensing 2011, 66, 115-123. https://doi.org/10.1016/j.isprsjprs.2010.09.008

Belgiu, M., Drăguţ L. Random Forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 2016, 114, 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bivand, R., Keitt, T., Rowlingson, B. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.5-23, 2021. https://CRAN.R-project.org/package=rgdal

Bouvier, M., Durrieu, S., Fournier, R. A., Renaud, J. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne las data. Remote Sensing of Environment 2015, 156, 322- 334. http://doi.org/10.1016/j.rse.2014.10.004

Breiman, L. Random Forests. Machine Learning 2001, 45, 5-32. https://doi.org/10.1023/A:1010933404324

Compés, R., García, J. M., Martínez, V. La crisis citrícola en la Comunidad Valenciana y el Acuerdo de Asociación Económica con el sur de África. Comunicación. Universitat Politècnica de València, 2019.

Czesak, B., Różycka-Czas, R., Salata, T., DixonGough, R., Hernik, J. Determining the Intangible: Detecting Land Abandonment at Local Scale. Remote Sensing 2021, 13, 1166. https://doi.org/10.3390/rs13061166

Dara, A., Baumann, M., Kuemmerle, T., Pflugmacher, D., Rabe, A., Griffiths, P., Hölzel, N., Kamp, J., Freitag, M., Hostert, P. Mapping the timing of cropland abandonment and recultivation in northern Kazakhstan using annual Landsat time series. Remote Sensing of Environment 2018, 213, 49-60. https://doi.org/10.1016/j.rse.2018.05.005

Estel, S., Kuemmerle, T., Alcántara, C., Levers, C., Prishchepov, A. V., Hostert, P. Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. Remote Sensing of Environment 2015, 163, 312-325. https://doi.org/10.1016/j.rse.2015.03.028

Generalitat Valenciana. 2020. Estadísticas agrícolas. Superficies y producción de la Comunitat Valenciana (Principales cultivos). https://agroambient.gva.es/es/estadistiques-agricoles

Gil-Yepes, J. L., Ruiz, L. A., Recio, J. A., BalaguerBeser, A., Hermosilla, T. Description and validation of a new set of object-based temporal geostatistical features for land-use/ land-cover change detection. ISPRS Journal of Photogrammetry and Remote Sensing 2016, 121, 77-91. https://doi.org/10.1016/j.isprsjprs.2016.08.010

Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., Hasanlou, M. Improved land cover map of Iran using Sentinel imagery withing Google Earth Engine and novel automatic workflow for land cover classification using migrated training samples. ISPRS Journal of Photogrammetry and Remote Sensing 2020, 167, 276-288. https://doi.org/10.1016/j.isprsjprs.2020.07.013

Grădinaru, S. R., Kienast, F., Psomas, A. Using multi-seasonal Landsat imagery for rapid identification of abandoned land in areas affected by urban sprawl. Ecological Indicators 2019, 96(2), 79-86. https://doi.org/10.1016/j.ecolind.2017.06.022

Hijmans, R. H. raster: Geographic Data Analysis and Modeling. R package version 3.4-10, 2021. https://CRAN.R-project.org/package=raster

Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., Hussin, Y. A. Generating pitfree canopy height models from airborne Lidar. Photogrammetric Engineering & Remote Sensing 2014, 80(9), 863-872. https://doi.org/10.14358/PERS.80.9.863

Kolecka, N., Kozak, J., Kaim, D., Dobosz, M., Ginzler, C., Psomas, A. Mapping Secondary Forest Succession on Abandoned Agricultural Land with LiDAR Point Clouds and Terrestrial Photography. Remote Sensing 2015, 7(7), 8300-8322. https://doi.org/10.3390/rs70708300

Löw, F., Prishchepov, F., Waldner, F., Dubovyk, O., Akramkhanov, A., Biradar, C., Lamers, J. Mapping Cropland Abandonment in the Aral Sea Basin with MODIS Time Series. Remote Sensing 2018, 10(2), 159. https://doi.org/10.3390/rs10020159

Ministerio de Agricultura, Pesca y Alimentación ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2019; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2020. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/ agricultura/esyrce/

Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.-T. Comparison of Sentinel-2 and HighResolution Imagery for Mapping Land Abandonment in Fragmented Areas. Remote Sensing 2020, 12(12), 2062. https://doi.org/10.3390/rs12122062

Morell-Monzó, S., Sebastiá-Frasquet, M.T., Estornell, J. Land Use Classification of VHR Images for Mapping Small-Sized Abandoned Citrus Plots by Using Spectral and Textural Information. Remote Sensing 2021, 13(4), 681. https://doi.org/10.3390/rs13040681

Neigh, C. S. R., Carroll, M. L., Wooten, M. R., McCarty, J. L., Powell, B. F., Husak, G. J., Enenkel, M., Hain C. R. Smallholder crop area mapped with wall-to-wall WorldView sub-meter panchromatic image texture: A test case for Tigray, Ethiopia. Remote Sensing of Environment 2018, 212, 8-20. https://doi.org/10.1016/j.rse.2018.04.025

Niemi, M. T., Vauhkonen, J. Extracting Canopy Surface Texture from Airborne Laser Scanning Data for the Supervised and Unsupervised Prediction of Area-Based Forest Characteristics. Remote Sensing 2016, 8(7), 582. https://doi.org/10.3390/rs8070582

Ozdemir, I., Donoghue D. N. M. Modelling tree size diversity from airborne laser scanning using canopy height models with image texture measures. Forest, Ecology and Management 2013, 295, 28-37. https://doi.org/10.1016/j.foreco.2012.12.044

Pal, M. Random Forest classifier for remote sensing classification. International Journal of Remote Sensing 2005, 26(1), 217-222. https://doi.org/10.1080/01431160412331269698

Prishchepov, A.V. Agricultural Land Abandonment. Oxford Bibliographies Environmental Science. Oxford University Press 2020, https://doi.org/10.1093/obo/9780199363445-0129

Prishchepov, A.V., Radeloff, V.C., Dubinin, M., Alcantara, C. The effect of Landsat ETM/ ETM image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Remote Sensing of Environment 2012, 126, 195-209. https://doi.org/10.1016/j.rse.2012.08.017

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.Rproject.org/.

Recio, J. A., Hermosilla, T., Ruiz, L. A., Palomar, J. Automated extraction of tree and plot-based parameters in citrus orchards from aerial images. Computer and Electronics in Agriculture 2013, 90, 24-34. https://doi.org/10.1016/j.compag.2012.10.005

Rounsevell, M. D. A., Reginster, I., Araújo, M. B., Carter, T. R., Dendoncker, N., Ewert, F., House, J. I., Kankaanpää, S., Leemans, R., Metzger, M. J. A coherent set of future land use change scenarios for Europe. Agriculture Ecosystems and Environment 2006, 114, 57-68. https://doi.org/10.1016/j. agee.2005.11.027

Roussel, J. R., Auty D. Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. R package version 3.1.2, 2021. https://cran.rproject.org/package=lidR

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., Homayouni, S. Support Vector Machine versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2020, 13, 6308-6325. https://doi.org/10.1109/JSTARS.2020.3026724

Subedi, Y. R., Kristiansen, P., Cacho, O. Drivers and consecuences of agricultural land abandonment and its reutilisation pathways: A systematic review. Environmental Development 2021, (in press). https://doi.org/10.1016/j.envdev.2021.100681

Szostak, M., Hawryło, P., Piela, D. Using of Sentinel-2 images for automation of the forest succession detection. European Journal of Remote Sensing 2017, 51, 142-149. https://doi.org/10.1080/22797 254.2017.1412272

CAPÍTULO 5

Assessing the capabilities of high-resolution spectral, altimetric, and textural descriptors for mapping the status of citrus parcels

Morell-Monzó, S., Estornell, J., & Sebastiá-Frasquet, M.-T. (2023). Assessing the capabilities of high-resolution spectral, altimetric, and textural descriptors for mapping the status of citrus parcels. Computers and Electronics in Agriculture, 204, 107504. https://doi.org/10.1016/j.compag.2022.107504

5.1. INTRODUCTION

Agricultural land abandonment (ALA) is an increasing issue around the world with relevant environmental (e.g., biodiversity, carbon sequestration, novel ecosystems, wildfires, and water resources) and socio-economic (e.g., food production, livelihood, and landscape) implications (Prishchepov, 2020). This issue occurs in variable social, environmental, and economic contexts (Kosmas et al., 2015). However, it is particularly prevalent in mountainous zones and in areas with highly fragmented parcel structures (Czesak et al., 2021). Depending on the type of crop, the social, and environmental context, ALA can be perceived as a problem or as an opportunity since it can have negative and positive effects on the environment and society (Subedi et al., 2021). These characteristics make ALA a complex phenomenon that must be studied at different scales (global, regional, and local) (Strijker, 2005). In the European Union about 11 % of agricultural land is at high risk of abandonment (Perpiña-Castillo et al., 2018). Therefore, time and cost-effective methods are urgently needed to identify ALA to provide land managers and policymakers with information on land use changes and to continuously update the European Common Agricultural Policy.

The Comunitat Valenciana (CV) region (Eastern Spain) is the most important citrus producer in Europe. It produces more than 3 million Tn of mandarins and oranges annually (50 % and 45 % of Spain's citrus industry, respectively) (IVIA, 2022). In 2021, the area dedicated to citrus was estimated at 160,088 ha (MAPA, 2021). However, from 2000 to 2020 there has been a decrease of around 20 % of the area dedicated to citrus crops (MAPA, 2022). Identifying the status of citrus crops at the parcel level with high accuracy is essential for the public administration to make decisions on agricultural policy and to supervise the payment of subsidies to farmers. This temporal and spatial-explicit information on ALA may help customize policy instruments for counteracting or reversing this process and to implement tailored monitoring and management measures of the landscape (Volpi et al., 2023).

In 2018, the European Commission approved the use of alternative methodologies to in situ verification for crop monitoring based on Earth Observation data (regulation No. 746/2018). First efforts on detecting ALA via Earth Observation data focused on the use of Landsat (e.g., Yin et al., 2018, Grădinaru et al., 2019, Prishchepov et al., 2012) and MODIS (e.g., Löw et al., 2018, Alcántara et al., 2012, Estel et al., 2015) data. These studies focused on the use of time series of low and moderate-resolution imagery to detect ALA in seasonal crops and large areas that do not require high spatial resolution. More recent studies show that the higher spatial resolution of Sentinel-2 imagery made possible to address the ALA issue (e.g., Szostak et al., 2017, Volpi et al., 2023). These new capabilities motivated some authors to use these images to identify ALA at patch-level and pixel-level (Ruiz et al., 2020, Portalès-Julià et al., 2021). There are also experiences on the use of 3D data for the detection of ALA. The works of Kolecka et al., 2015, Czesak et al., 2021 used LiDAR data to map the abandonment of seasonal crops. LiDAR data allowed to detect of advanced stages of ALA and secondary forest successions in areas with highly fragmented landscapes that were once occupied by crops.

Despite the advances, the operational implementation of citrus crop monitoring is still challenging in highly fragmented landscapes. These are common in the CV, where the size of the parcels can be a limiting factor to apply Sentinel-2 imagery. Fragmented

landscapes are also common in many European areas and some authors have studied the applicability limits of Sentinel-2 but for monitoring seasonal crops with higher spectral-temporal separability (Vajsová et al., 2020). Also, in citrus crops parcels in production and abandoned have a very similar spectral response. In addition, unlike seasonal crops, citrus is a perennial crop that does not has a cycle throughout the year (e.g., wrought, planting, germination, growth, and harvest) which could help to distinguish them from abandoned fields. The VHSR data can produce better results because its sub-metric resolution can capture textures caused by planting patterns (also called planting frames) characteristic of fruit crops, which is not possible with Sentinel-2 10 m resolution. Furthermore, 3D point clouds can provide new information that can improve the identification of citrus crops' status. Several studies have combined aerial and satellite imagery with 3D point clouds to improve crop type and other land cover classifications (Prins & Van Niekerk, 2021).

The objective of this paper is to evaluate the use of three VHSR data sources to identify the status of citrus crops: WorldView-3 (WV-3) imagery including visible and near-infrared (VNIR) and short-wave infrared (SWIR) bands, Very High-Resolution Airborne (VHRA) imagery, and 3D point clouds generated applying Structure from Motion (SfM) on images captured by a Remotely Piloted Aircraft System – RPAS. The classification performance including texture features for each data type is also analysed. Additionally, the study evaluates the fusion of multi-spectral data and 3D point clouds to identify citrus crops' status. Two algorithms widely used in land cover classification problems are compared: Random Forest (RF) and Support Vector Machines (SVM). Finally, the relevance of the descriptors used for the classification and the class separability is statistically analysed.

5.2. MATERIAL AND METHODS

5.2.1. Study area

The study area is located in the CV region (Eastern Spain) (Figure 5.1), between the coastal municipalities of Gandia and Oliva (38°56′14.14″N, 0° 8′42.03″W), which occupies around 10,000 ha. In this area various land uses coexist such as urban, forestry, and agricultural use. The majority crop in this zone is citrus which occupies more than 95 % of the agricultural area (Generalitat Valenciana, 2021). The agricultural area is located on the coastal plain which is characterized by its flat relief, soils with high agronomic capacity, and availability of water. These characteristics have made this region a historically agricultural area. In recent years there has been a general abandonment of citrus. The agricultural structure is characterized by the small size of parcels (average 0.20–0.40 ha) and the high fragmentation of the landscape.
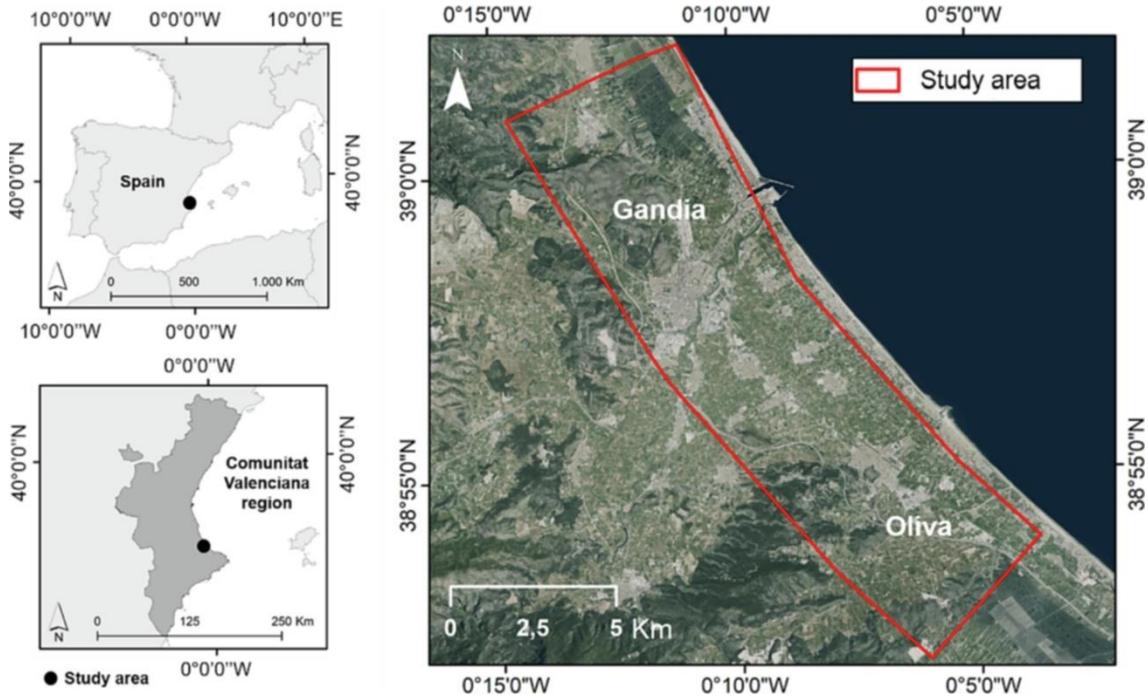
Figure 5.1. Study area between the municipalities of Gandia and Oliva in the Comunitat Valenciana region (eastern Spain). Centre of the image: 38°56′14.14″N, 0° 8′42.03″W.

5.2.2. Spatial data and feature extraction

WorldView-3 images with values of top-of-atmosphere radiance were acquired on November 11, 2020. The acquired WV-3 image contains 8 VNIR bands (Coastal blue 425 nm, Blue 480 nm, Green 545 nm, Yellow 605 nm, Red 660 nm, Red- edge 725 nm, NIR1 835 nm, and NIR2 950 nm) of 11-bits and 2 m resolution, one pan-chromatic band (450–800 nm) of 11-bits and 0.5 m resolution, and 8 SWIR bands (SWIR1 1210 nm, SWIR2 1570 nm, SWIR3 1660 nm, SWIR4 1730 nm, SWIR5 2165 nm, SWIR6 2205 nm, SWIR7 2260 nm, SWIR8 2330 nm) of 14-bits and 3.7 m resolution (Digital Globe, 2014). The WV-3 product was orthorectified and atmospherically corrected using the ATCOR2 algorithm (Richter & Schläpfer, 2019, Richter & Schläpfer, July 2021) to obtain surface reflectance images. To combine images with different spatial resolutions resampling processes were applied. The panchromatic band with a resolution of 0.5 m was resampled at a resolution of 1 m. The average was used as the aggregating criteria of the pixels. Then the VNIR (2 m) and SWIR (3.7 m) bands were resampled to 1 m using the nearest neighbor and de panchromatic band as reference.

To produce new descriptors for the classifier, 7 texture features were computed from the panchromatic band. The texture is an important characteristic of images that informs about the spatial distribution of the different pixel intensity levels. This spatial information is not necessarily correlated with spectral data, so including texture descriptors could improve classification accuracy (Hall-Beyer, 2017). In this study, the Gray Level Co-Occurrence Matrix (GLCM) method (Haralick et al., 1973) was used to compute texture features. These were calculated from 64 gray levels in an invariant direction that is the average of the four directions: 0°, 45°, 90°, and 135°. This omnidirectional approach was chosen due to planting frames do not present a specific direction. Two first-order histogram texture features were calculated: mean, variance,

and five GLCM texture features: entropy, contrast, dissimilarity, angular second moment (ASM), and homogeneity (Morell-Monzó et al., 2021).

Very High-Resolution Airborne (VHRA) Imagery was also used in this study. This data was provided by the Valencian Cartographic Institute (ICV) (Source: https://idev.gva.es/va Orthophoto 2020 CC BY 4.0 © Institut Cartogràfic Valencià, Generalitat Valenciana). These images were taken on a photogrammetric flight on May 6, 2020 with an UltraCam Eagle UC-E-1–50016095-f80 camera from the company Vexcel Imaging GmbH © with a Qioptic Vexcel HR Digaron sensor. The ICV provides an annual VHRA image of the entire Valencian Region in tiles of 15160×10160 pixels. These images have a spatial resolution of 0.25 m and 4 spectral bands (Blue 430 nm, Green 530 nm, Red 620 nm, Red-edge 720 nm). These images were radiometrically calibrated and geometrically corrected (RMSE < 4 cm) by Vexcel Imaging GmbH ©. 11 tiles were processed and resampled to 1 m to reduce the computational requirements and the pixels' spectral variability within each parcel. The value of the resampled pixels was based on the average value of the aggregated pixels. Finally, the same 7 texture features were computed from the Normalized Difference Vegetation Index (NDVI). The NDVI was computed by the normalized difference of the Red and Red-edge bands.

A photogrammetric point cloud was obtained through 12 RPAS flights carried out in February 2021. The RPAS was a senseFly eBee X which incorporates a S.O.D.A. RGB 20.1 MP and an RTK/PPK positioning system (3–8 cm positioning error). The flight height was around 123.7 m in nadir shots. A total of 5510 images were taken with 60 % lateral overlap and 80 % flight direction overlap. The images were processed using the SfM technique to obtain a 3D point cloud. No additional control points were required thanks to the RTK/PPK positioning system. An alignment error < 6 cm was obtained. Then a dense point cloud (325 points/m2) with approximately 4723 million points was generated. The dense point cloud was homogenized and its density was reduced to decrease processing time. Homogenization was performed at a target density of 50 points/m2. The Progressive Morphological Filter (PMF) method (Zhang et al., 2003) was used to identify the ground points. It is well known that correctly identifying ground points from SfM point clouds can be a challenging task, especially in steeply sloping areas and dense vegetation. However, citrus crops in our study area have some characteristics that make this task easier. Citrus, which are irrigated crops, must necessarily be located in parcels with low slopes (<3%) to apply irrigation properly. In addition, agricultural practices require spaces between tree rows and at the parcel boundaries. These spaces make that a significant set of ground points are retrieved correctly. On the other hand, farmers need to keep their parcels free of vegetation to facilitate agricultural practices, which helps to reduce DTM overestimation. In this context, the PMF algorithm, which is one of the standard methods, provided sufficiently accurate results. The Digital Terrain Model (DTM) with a cell size of 1 m × 1 m was computed. Then the Digital Surface Model (DSM) was computed using the pit-free method (Khosravipour et al., 2014). Finally, the Canopy Height Model (CHM) and 9 statistics were extracted from the normalized point cloud: mean H, standard deviation H, maximum H, minimum H, range of H, variance H, skewness H, kurtosis H, and entropy of H, where H is the height above the ground. These features were stored in a 1 m raster grid. Additionally, the same 7 texture features computed from the CHM were calculated and added to the raster grid to produce new descriptors for the classifier.

In order to generate useful information to detect, quantify and manage citrus crops abandonment, a classification was defined based on three types of parcels: non-productive (NP), productive (PR), and abandoned (AB) as shown in Figure 5.2. Identify the AB parcels is useful for monitoring land abandonment. Furthermore, identify the NP and PR parcels allows making more accurate crop yield estimations. Citrus parcels were previously identified from the Land Parcel Identification System (LPIS) database of Spain known as Sistema de Información Geográfica de Parcelas Agrícolas (SIGPAC).

Figure 5.2. Example of the three parcel types classified: non-productive (left), productive (center) and abandoned (right).

Ground truth data were obtained through field campaigns conducted in February and March 2021. A systematic sampling of the area covered with the three data sources (WV-3, VHRA, and RPAS-SfM) was carried out. Verification was done through photointerpretation of RPAS images and the ICV orthophotos due to the different data acquisition of ground truth data and remote sensing. A total of 280 parcels without changes during the period May 2020 - February 2021 were selected. This dataset contains 60 NP, 120 PR, and 100 AB parcels. This unbalanced dataset keeps the proportions of each category observed in the field.

5.2.3. Classification

The classification approach used in this study is based on that proposed in Morell-Monzó et al. (2021) adapted to each data source. The general workflow consists of the following steps: 1) preprocessing raw data; 2) feature extraction (altimetric features, texture features, and pixel values for each band); 3) model training (RF or SVM) using pixels as training samples; 4) semantic segmentation of the image to create a pixel-based classification; 5) apply majority voting to obtain a single classification value per parcel.

In this work two commonly employed classification algorithms were used, Random Forest (Breiman, 2001) and Support Vector Machines (Cortes & Vapnik, 1995). Although there is a growing trend towards the use of deep learning algorithms, RF and SVM are still benchmark algorithms for many remote sensing problems (Sheykhmousa et al., 2020). This is due to their good performance with small amounts of training data and ease of use for non-expert users as they require few hyperparameters to be adjusted, they use non-sequential training, they are quite robust against overfitting they are computationally less demanding and they have higher interpretability compared with deep learning) (Sheykhmousa et al., 2020). In land cover classifications, RF and SVM are two of the most popular methods (Saini & Ghosh, 2018).

The RF algorithm was applied using the `randomForest` R package (Liaw and Wiener, 2002). RF requires adjusting two main hyperparameters: the number of trees (`ntree`) that make up the forest, and the number of variables randomly selected in each node split (`mtry`). In this study, the `ntree` was adjusted to 250 and `mtry` was adjusted to $\sqrt{M}$, where M is the number of input variables. The impurity is used to search the optimal threshold at each node of the tree. RF used the Gini Index as a measure of impurity.

The SVM algorithm was applied using the e1071 R package (Meyer et al., 2021). A SVM with a radial basin function kernel was used. It has two hyperparameters to adjust: the regularization parameter (C) and kernel bandwidth (γ). In this study the constant of regularization C was fixed to 1 and γ was adjusted to $1/M$, where M is the number of input variables (data dimensionality). The tolerance of termination criterion was set by default in 0.001.

5.2.4. Accuracy assessment and validation

First, the optimal kernel size for extracting texture features was obtained. The following kernel sizes were evaluated: 3x3, 5x5, 7x7, 9x9, 11x11, 13x13, 15x15, 17x17, 19x19, 21x21, 23x23 and 25x25. The optimal kernel size was the one that maximized the overall accuracy (OA) based on the RF classifier.

Once the optimal kernel size was obtained, the performance of each data bundle and its respective combinations was evaluated using both RF and SVM. Each model was evaluated using a 4-folds cross-validation approach using random splits without replacement. In each iteration, 3/4 of the parcels were used for training and 1/4 for validation, ensuring a balanced dataset with the same number of parcels of each class by iteration. The following performance metrics were calculated: OA, Cohen's Kappa, precision (producer's accuracy by class), and recall (user's accuracy by class).

$$Overall\ Accuracy = \frac{CM_{ii}}{N} \tag{5.1}$$

Cohen's Kappa can be generalized to the m classes,

$$Cohen's\ Kappa = \frac{N \sum_{i=i}^{m} CM_{ii} - \sum_{i=1}^{m} Ci_{corr} Ci_{pred}}{N^2 - \sum_{i=1}^{m} Ci_{corr} Ci_{pred}} \tag{5.2}$$

where N is the total number of samples, m is the number of classes, CM_{ii} the diagonal elements of the confusion matrix, and Ci_{corr} and Ci_{pred} are the correct and predicted labels of the class i respectively.

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \tag{5.3}$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \tag{5.4}$$

where TP_i are the true positives for the i class, FP_i are the false positives for the i class, and FN_i are the false negative for the i class.

A second validation step was performed. In this case a spatial cross-validation with 4 clusters was performed. This validation technique was applied to avoid an underestimation of the model error due to spatial correlation of the samples which is inherent to remote sensed data (Karasiak et al., 2021). Random cross- validation can lead to error underestimation when data are strongly aggregated (Wadoux et al., 2021, Stock, 2022), as in our case. In this validation step, only the performance of the most accurate data bundle from each data source was evaluated. Figure 5.3 shows the different splits used in random and spatial cross-validation. Spatial clusters were generated by k-Means clustering from the XY coordinates of the parcels. This procedure allows to generate spatial groups of parcels with the greatest distance between groups. However, no balanced data splits are generated.
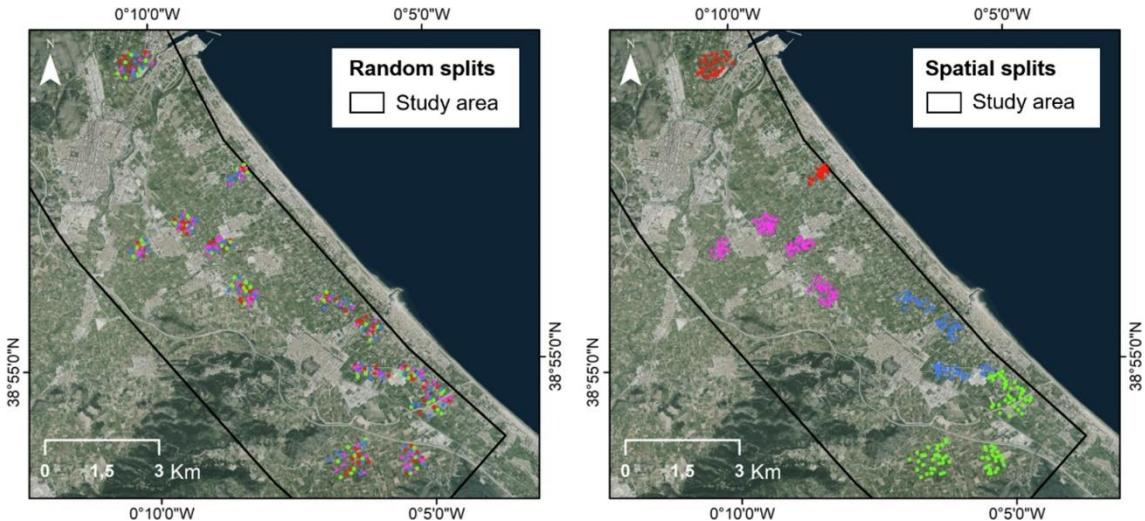
Figure 5.3. Data splits used in random cross-validation (left) and spatial cross-validation (right).

5.2.5. Descriptors' relevance and class separability

To evaluate the relevance of each descriptor, the Jeffries-Matusita (JM) distance was used. JM distance measures the separability between a pair of probability distributions. This method does not inform about the influence of a descriptor on the model performance but it collects the intrinsic structure of the variables. Unlike methods based on permutation or impurity (such as RF feature importance measures), this is a statistical method not biased against collinearity and cardinality. In addition, JM distance allows knowing the separability between pairs of classes, which favours the interpretation of the problem. The JM distance takes the range [0, √2] and is defined as:

$$J_{xy} = \sqrt{2(1 - e^{Bxy})} \qquad (5.5)$$

where B_{xy} is the multivariate Bhattacharyya distance,

$$B_{xy} = \frac{1}{8}(x-y)^t \left(\frac{\sum x + \sum y}{2}\right)^{-1}(x-y) + \frac{1}{2}\ln\left(\frac{\left|\frac{\sum x + \sum y}{2}\right|}{\left|\sum x\right|^{\frac{1}{2}} + \left|\sum y\right|^{\frac{1}{2}}}\right) \tag{5.6}$$

where x is the first spectral signature vector, y is the second spectral signature vector, $\sum x$ is the covariance matrix of x, and $\sum y$ is the covariance matrix of y.

5.3. RESULTS

5.3.1. Kernel size analysis

The kernel size used to extract texture features influences the performance of the model. The optimal kernel size depends on the size and shape of the parcels as well as the spatial resolution of the image. The theoretical behaviour is that as the kernel size increases, a higher OA of the pixel-based classification is obtained due to the reduction of salt-and-pepper noise. For this reason, the larger kernel, produce the smaller improvements in OA of the parcel-based classification when majority voting is applied. However, if the window size increases too much, a decrease in parcel-based OA can occur. This can be explained considering that a too large window can include pixels from other categories and adjacent parcels. So, the pixels at the edges of the parcels can be misclassified.

The analysis of the kernel size used to extract the texture features is shown below (Figure 5.4). The optimal window size to compute texture features from the panchromatic band of WV-3 image was 7x7. For the NDVI derived from the VHRA image was 11x11. However, in this case the OA of the parcel-based classification remained stable for all window sizes (0.975–0.986). For the CHM derived from the SfM the optimal window size was 11x11. For the SfM data, the OA of the parcel-based classification was even lower than the OA of the pixel-based classification (Figure 29c) using a 25x25 kernel. On the other hand, the OOB error of the RF model decreased as the window size increased, so this measure is not recommended to adjust the kernel size as this will result in an underestimation of the error. This occurs because the OOB samples and the training samples are included in the same window. As the window size increase the OOB samples and the training samples are more correlated as they share more common neighbours.
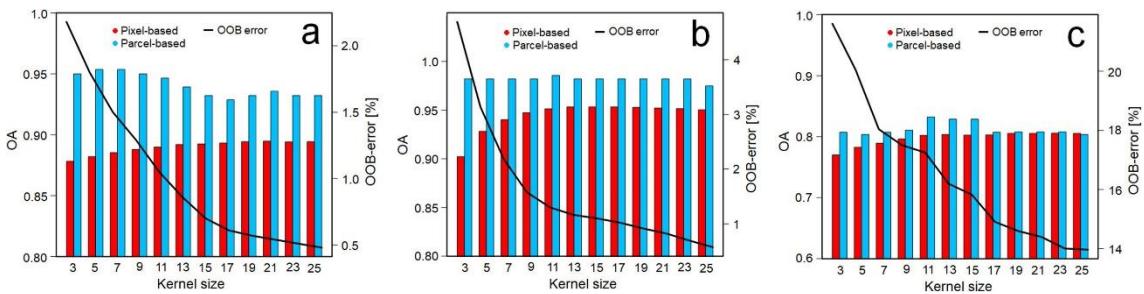
Figure 5.4. Overall accuracy (OA) of the pixel-based classification (before majority voting step) and parcel-based (after majority voting step) and Random Forest's Out-Of-Bag error at each window size from 3x3 to 25x25. a) corresponds to the WorldView-3 data, b) corresponds to the Very High-Resolution Airborne Imagery, and c) corresponds to the Structure from Motion data.

5.3.2. Classification accuracy

The results of the random cross-validation for each data bundle are shown below (Figure 5.5). The data source that produced the best results was the VHRA image (OA max = 0.986) followed by the WV-3 image (OA max = 0.954) and the SfM point cloud (OA max = 0.836). RF and SVM showed similar performance in VHRA image. But RF was slightly more accurate in the WV-3 images and SfM point clouds. Furthermore, RF is more convenient to larger datasets due to its lower time complexity. SVM has a time complexity $O(dn^2)$ or $O(dn)^3$ while RF has a time complexity of $O(n\log(n)dk)$.
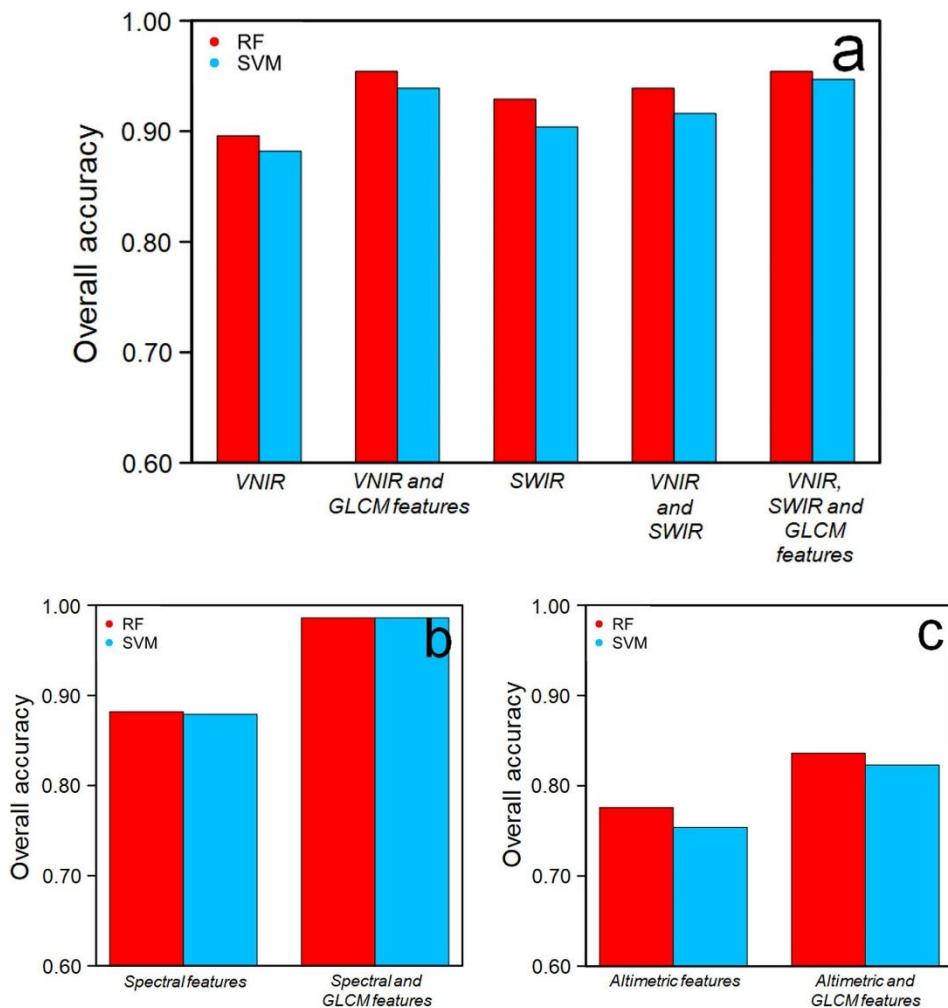
Figure 5.5. Overall accuracy of the classification based on Random Forest (RF) and Support Vector Machines (SVM) for each data bundle: a) WorldView-3 subsets, b) Very High-Resolution Airborne Image subsets, and c) SfM point cloud subsets.

All the WV-3-based data bundles produced an OA between 0.882 and 0.954 (Table 13). The best performance was obtained using the VNIR bands, SWIR bands and GLCM texture features. This combination produced an OA of 0.954 using RF and 0.947 using SVM. Similar results were obtained by the VNIR bands and GLCM texture features with an OA of 0.954 using RF and 0.939 using SVM. Finally, the data bundle with the lowest performance was obtained when the VNIR bands were used only with an OA of 0.896 using RF and 0.882 using SVM.

Table 5.1. Accuracy metrics of WorldView-3 based data bundles using random cross-validation. NP, PR and AB correspond to the non-productive, productive and abandoned parcels respectively.

	OA	Kappa	NP Precision	PR Precision	AB Precision	NP Recall	PR Recall	AB Recall
	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)
VNIR bands	0.896	0.836	0.960	0.909	0.856	0.733	0.983	0.890
	0.883	0.827	0.951	0.900	0.846	0.742	0.973	0.893
VNIR bands + GLCM features	0.954	0.927	0.948	0.959	0.931	0.900	0.975	0.940
	0.939	0.915	0.951	0.949	0.908	0.867	0.974	0.965
SWIR bands	0.929	0.905	0.864	0.949	0.909	0.853	0.978	0.940
	0.904	0.846	0.970	0.909	0.856	0.775	0.983	0.890
VNIR bands + SWIR bands	0.939	0.905	0.964	0.959	0.909	0.833	0.975	0.960
	0.916	0.833	0.960	0.931	0.858	0.775	0.973	0.890
VNIR bands, SWIR bands + GLCM features	0.954	0.927	0.968	0.973	0.924	0.965	0.973	0.927
	0.947	0.917	0.960	0.961	0.912	0.962	0.973	0.924

Table 5.1 shows the performance by class of each WV-3 based data bundle. Adding GLCM texture features from the panchromatic band to VNIR multispectral data improved OA around 5.5 %. Adding texture features improved recall of NP class primarily. The texture features also improved the precision of the PR and AB classes and the recall of the AB class. Combinations which do not include texture features showed a larger commission error of the NP class due to the erroneous classification between PR and AB classes. Furthermore, the combinations incorporating VNIR bands without texture features generated a lower recall but higher precision for the NP class. In this case, the high commission error of the NP class was caused by the erroneous classification of the AB class and, to a lesser extent, of the PR class. Adding SWIR bands to the VNIR bands produced improvements in AB class detection (improving both precision and recall). Finally, the model which includes all the data bundles (VNIR, SWIR, and GLCM features) obtained the best performance and showed a good balance between user's and producer's accuracy. According to this model, the most difficult class to detect was AB.

The VHRA image-based data bundles produced an OA between 0.882 and 0.986 (Table 5.2). The best performance was obtained using the VNIR bands and GLCM texture features derived from NDVI. This combination produced an OA of 0.986 using both RF and SVM. The data bundle that uses only the VNIR bands obtained a lower recall for the NP class but a high precision. It produces a higher commission error for the NP class mainly due to the erroneous classification to PR class and AB class to a lesser extent. Adding texture features allowed to identify all the NP parcels correctly. Adding texture features also improved the precision and recall of the PR and AB classes, producing a good balance between precision and recall for all classes Table 5.2.

Table 5.2. Accuracy metrics of Very High-Resolution Airborne Imagery based data bundles using random cross-validation. NP, PR and AB correspond to the non-productive, productive and abandoned parcels respectively.

	OA	Kappa	NP Precision	PR Precision	AB Precision	NP Recall	PR Recall	AB Recall
	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)
VNIR bands	0.882	0.814	1.000	0.806	0.944	0.867	0.958	0.800
	0.879	0.806	1.000	0.803	0.924	0.841	0.956	0.807
VNIR bands + GLCM	0.986	0.978	1.000	0.992	0.971	1.000	0.975	0.990
features	0.986	0.973	1.000	0.989	0.974	1.000	0.986	0.981

The data bundles derived from SfM point cloud produced an OA between 0.754 and 0.836. The best performance was obtained using the 3D point cloud features and the GLCM texture features derived from the CHM. This combination produced an OA of 0.8336 using RF and 0.823 using SVM. 3D data showed lower performance than the WV-3 and VHRAI. But these results show the potential of 3D point clouds data to monitor the status of citrus crops, specially including texture features derived from the CHM. Texture features improved the OA around 6.0 %.

The 3D point clouds and texture features allows to identify the NP class with high accuracy (0.929 of precision and 0.966 of recall using RF). However, these data showed low ability to separate the PR and AB parcels.

Table 5.3. Accuracy metrics of SfM point clouds-based data bundles using random cross-validation. NP, PR and AB correspond to the non-productive, productive and abandoned parcels respectively.

	OA	Kappa	NP Precision	PR Precision	AB Precision	NP Recall	PR Recall	AB Recall
	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)
Altimetric features	0.776	0.568	0.827	0.758	0.692	0.627	0.851	0.840
	0.754	0.557	0.822	0.717	0.692	0.631	0.843	0.835
Altimetric + GLCM	0.836	0.738	0.929	0.840	0.786	0.966	0.504	0.892
features	0.823	0.738	0.909	0.831	0.788	0.960	0.528	0.863

The results of the spatial cross-validation for the most accurate data bundle are shown below (Table 5.4). Spatial cross-validation show an average OA around 1.5 % lower than using the random cross-validation. These results suggest a small underestimation of the model error due to the spatial correlation of the parcels when using random cross-validation. Both RF and SVM were affected by this spatial correlation. In this case the most accurate data source was also the VHRA image (OA between 0.967–0.962) followed by the WV-3 image (OA between 0.936–0.935) and the SfM point cloud (OA = 0.822).

Table 5.4. Accuracy metrics for the best model of each data bundle. NP, PR and AB correspond to the non-productive, productive and abandoned parcels respectively using spatial cross-validation.

	OA	Kappa	NP Precision	PR Precision	AB Precision	NP Recall	PR Recall	AB Recall
	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)	(RF/ SVM)
WV-3 (VNIR bands + SWIR bands + GLCM features)	0.936 0.935	0.907 0.900	0.917 0.926	0.960 0.957	0.914 0.921	0.961 0.945	0.958 0.961	0.911 0.913
VHRAI (VNIR bands + GLCM features)	0.967 0.962	0.951 0.948	1.000 1.000	0.958 0.967	0.949 0.930	1.000 1.000	0.942 0.939	0.964 0.945
SfM point cloud + GLCM features	0.822 0.822	0.714 0.718	0.914 0.918	0.838 0.835	0.776 0.781	0.957 0.954	0.618 0.615	0.878 0.865

Figure 5.6 shows the map of abandonment of citrus parcels in the study area. The map was generated by using the VHRA image and the Random Forest classifier. It can be seen that abandonment is concentrated mainly in coastal areas. In addition, there is a particularly high concentration of abandoned parcels near the protected wetlands of Marjal de la Safor (north of the map) and Marjal Pego-Oliva (south-east of the map). These results agree with our field work in the study area and our previous experiences (Morell-Monzó et al., 2021).
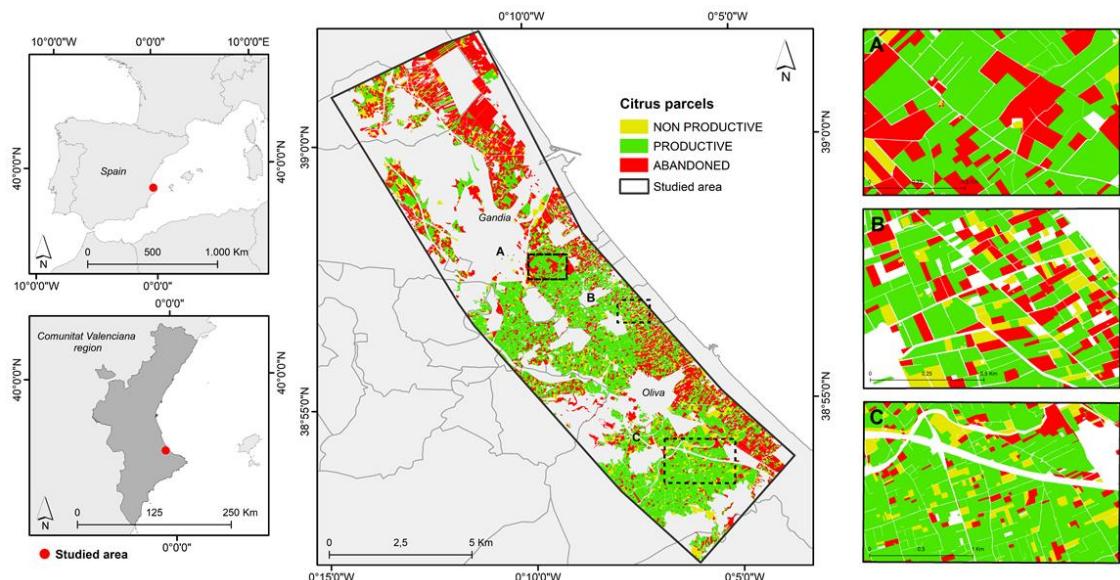
Figure 5.6. Map of abandonment of citrus parcels between the municipalities of Gandia and Oliva in the Comunitat Valenciana region (Spain). The map was obtained using the Very High-Resolution Imagery of the Valencian Cartographic Institute of 2020 and the Random Forest classifier, which proved to be the most accurate results.

5.3.3. Spectral, altimetric, and textural combination

The performance obtained by combining spectral and 3D data is shown below (Table 5.5). The results indicate an improvement of 1.85 % in the OA when 3D point cloud features are added to the WV-3 images and an improvement of 0.55 % in the OA when these features are added to the VHRA image. The improvements produced by the 3D information were limited due to the good performance of the WV-3 images and the VHRA image, which have small margins of improvement.

Table 5.5. Random Forest-based accuracy metrics obtained by combining spectral, 3D point cloud and textural descriptors. R-CV are the results obtained through random cross-validation and S-CV are the results obtained through spatial cross-validation. S + A is the accuracy metric obtained by combining spectral features, its texture features, 3D point cloud features, and its CHM texture features and Δ is the percentage of improvement produced by adding altimetric features to the spectral information. NP, PR and AB correspond to the non-productive, productive and abandoned parcels respectively.

	OA	Kappa	NP Precision	PR Precision	AB Precision	NP Recall	PR Recall	AB Recall
	S+A	S+A	S+A	S+A	S+A	S+A	S+A	S+A
	(Δ%)	(Δ%)	(Δ%)	(Δ%)	(Δ%)	(Δ%)	(Δ%)	(Δ%)
WV-3 + Altimetric features (R-CV)	0.971 +1.70	0.955 +2.80	1.000 +3.2	1.000 +2.70	0.931 +0.70	0.895 -7.00	1.000 +2.70	1.000 +7.30
WV-3 + Altimetric features (S-CV)	0.956 +2.00	0.927 +2.00	0.981 +6.4	0.968 +0.80	0.943 +2.90	0.972 +1.20	0.971 +1.30	0.902 -0.90
VHRAI + Altimetric features (R-CV)	0.993 +0.70	0.989 +1.10	1.000 +0.00	0.992 +0.00	0.990 +1.90	0.983 -1.70	1.000 +2.50	0.994 +0.40
VHRAI + Altimetric features (S-CV)	0.971 +0.40	0.961 +1.09	1.000 +0.00	0.968 +0.80	0.946 -0.30	1.000 +0.00	0.969 +2.90	0.950 -1.40

5.3.4. Descriptors' relevance and class separability

Regarding the WV-3 descriptors, the bands in the visible spectrum (coastal blue, blue, green, yellow, red) produced a greater separability between the NP-AB classes, followed by NP-PR and PR-AB (Figure 5.6a). In the near-infrared spectrum (red edge, NIR1, NIR2) the separability is minimized for all class combinations. The first order texture features, mean and variance, produced a similar separability to the visible bands due to its high correlation to the panchromatic band. However, the second order texture features improved the separability between the PR-AB classes, which are more difficult to separate. This explains the performance improvement produced by the texture features.
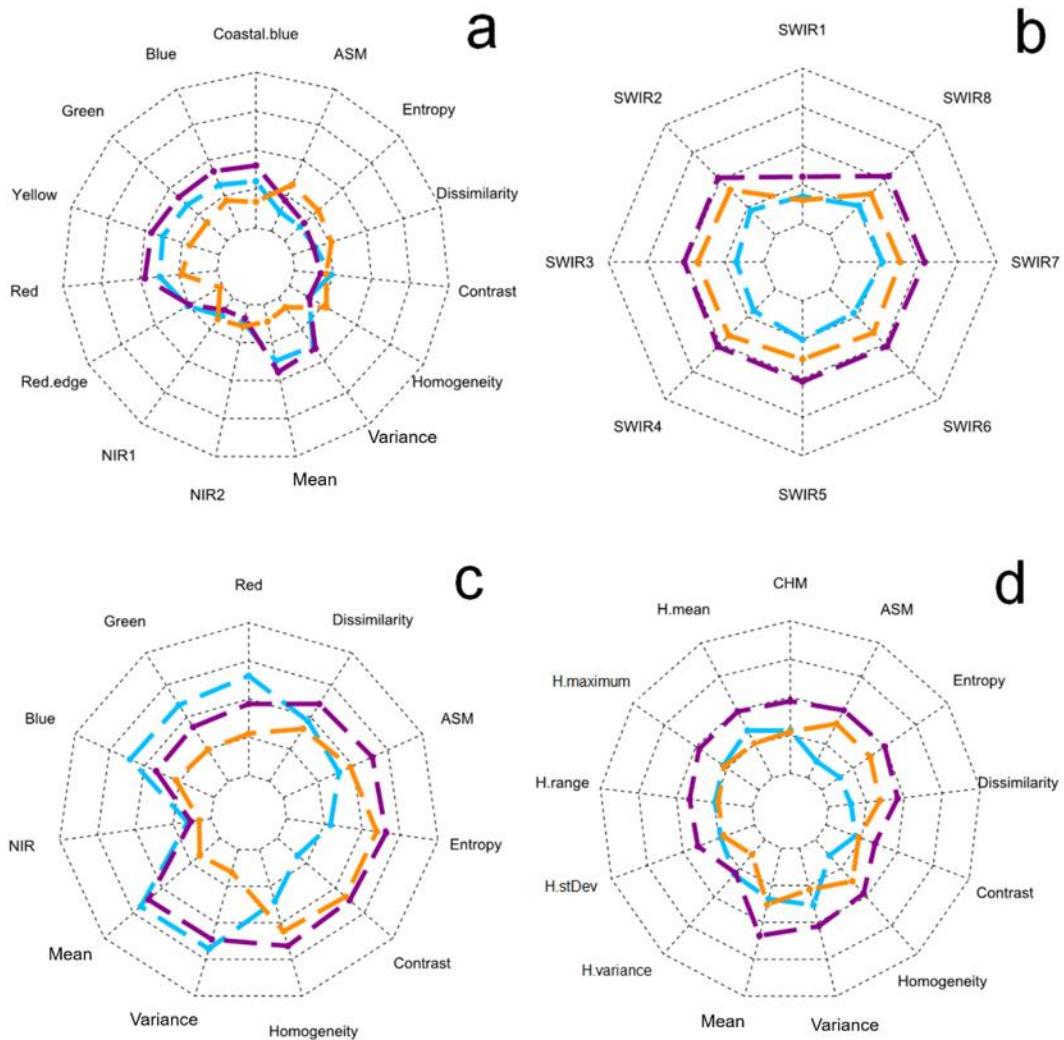
Figure 5.6. Class separability produced by each descriptor according to the Jeffries-Matusita distance. a) WorldView-3 VNIR bands and GLCM features, b) WorldView-3 SWIR bands, c) Very High-Resolution Airborne Image bands and GLCM features. d) SfM point cloud altimetric and GLCM features. Jeffries-Matusita distance varies between 0 and √2, where 0 is null separability and √2 is maximum separability.

On the other hand, the SWIR bands produce a higher separability between the NP-AB classes, followed by PR-AB and NP-PR (Figure 5.6b). This greater separability between the PR-AB classes explains the improvement in precision and recall of the AB class produced by the SWIR bands. There is a maximum separability of the PR-AB classes between 1570 nm and 2330 nm, which correspond to the SWIR2 and SWIR8 bands. In addition, there is a low separability in the SWIR1 band, which is the closest to the near-infrared spectrum.

Regarding the VHRA image-derived descriptors, the bands of the visible spectrum produced a greater separability between the NP-PR classes, followed by NP-AB and PR-AB (Figure 5.6c). In this wavelength region, the separability between the NP-AB and PR-AB classes was similar to that of WV-3. However, the separability between the NP-PR classes was higher. The VHRA image also show a minimum of separability for all classes in the NIR - Red-edge wavelengths. On the other hand, texture features extracted from NDVI improved the separability between NP-AB and PR-AB classes. The results seem to indicate that extracting texture features from NDVI produce better

results than from the WV-3 panchromatic band. This can be explained because NDVI produces a higher contrast between trees and bare soil.

Regarding the SfM point cloud descriptors, the highest separability occurred between the NP-AB classes (Figure 5.6d). The combinations of NP-PR and PR-AB classes produced lower separability. Finally, texture features improved the separability of the PR-AB classes. These results show a certain potential of the texture features extracted from the CHM.

5.4. DISCUSSION

This study evaluated the capabilities of high spatial resolution descriptors for mapping the status of citrus parcels in a highly fragmented agricultural system. The results show the potential of WV-3 imagery (including VNIR and SWIR bands), VHRA imagery with four VNIR bands, and SfM-derived point clouds to identify citrus crop status. Both WV-3 and VHRA imagery showed a high potential (OA ≥ 0.90) to identify the crop status using a single image.

The WV-3 image achieved an OA between 0.954 and 0.935 using the VNIR band, SWIR bands, and GLCM texture features extracted from the panchromatic band. This image was taken in February where rainfall and higher soil moisture occur. Although farmers often control scrub, these conditions produce vegetation growth not associated with land abandonment that can affect imagery performance. The WV-3 SWIR bands showed a high potential specially to detect the abandoned parcels. The separability analysis showed that SWIR bands produce a higher separability between PR and AB crops, which are the most spectrally similar classes. These results can be explained as SWIR bands are related to vegetation and soil moisture. The abandoned parcels do not receive irrigation while in productive parcels the farmers irrigate their crops avoiding a water stress situation. This fact suggests that the separability between PR and AB parcels could be greater in the dry season where there is hardly any rainfall. These facts, encourage the study of ALA in evergreen crops through time-series of vegetation and moisture indices.

The VHRA image achieved an OA between 0.986 and 0.962 using the VNIR bands (Red, Green, Blue, Red-edge) and the GLCM texture features extracted from the NDVI. The VHRA image was the most accurate data source. However, these images have lower radiometric quality than WV-3. These images do not have surface reflectance products. In addition, the images are taken during photogrammetric flights carried out on different dates and therefore with different atmospheric conditions and different solar angles, even using different sensors every year. These characteristics make it difficult to create a spatially and temporally transferable model. However, the excellent results obtained allow a monitoring plan for citrus crops in the entire Valencian territory by using several locally trained models. Alternatively, the radiometric quality of the VHRA imagery could be enhanced by applying image fusion techniques (e.g., Wang & Atkinson, 2018, Houborg & McCabe, 2018, Zhao & Liu, 2022) that allows fusion with Sentinel-2.

SfM point clouds achieved a lower performance than optical spectral data. SfM point clouds achieved an OA between 0.836 and 0.822 using altimetric features and GLCM texture features extracted from the CHM. Combining multispectral and altimetric data produced a slight improvement in model performance. However, generating 3D point

clouds using SfM is computationally and operationally demanding, so it is not suitable for large areas. This fact may not justify the use of 3D point clouds generated by SfM for this problem. An alternative for large-scale studies would be to use Airborne Laser Scanning (ALS) data to extract texture features from the CHM. The classification performance obtained is promising and highlights the importance of texture features extracted from CHM for identifying spatial patterns of crops. This information could complement moderate-resolution imagery such as Sentinel-2, of free access, large coverage, and high temporal resolution when spatial planting patterns are a key feature. Texture features extracted from the CHM allowed to model citrus planting frames. However, we assume that there may be a CHM underestimation under certain circumstances. The first is the DTM overestimation in the presence of dense vegetation. The second is the smoothing effect of the tree canopies produced by the SfM algorithm. Although the characteristics of citrus crops make it easy to obtain a consistent CHM, in other scenarios this can be a limiting factor. The use of ALS systems with higher penetration (even more for Full-Waveform LiDAR systems), could also be an alternative to compute a more accurate DTM. Another alternative is the use of the algorithm proposed by Tan et al. 2018 to remove misclassified ground points on vegetation using the spectral information of the images captured by the RPAS.

Including texture features produced improvements in all tested models, especially when they were computed from the NDVI. The descriptors relevance showed that texture features complement the spectral data and covers gaps where spectral information does not produce enough separability between classes. These results agree with our previous experiences (Morell-Monzó et al., 2021) and confirm the importance of texture features extracted from very high spatial resolution data to identify the status of evergreen crops (e.g., citrus crops). The importance of textures lies on their ability to model spatial planting patterns (planting frames) typical of fruit crops.

This work provides new research to identify the status of citrus crops and other evergreen crops at parcel level using remote sensing data. In particular, the work provides advances towards the operational implementation of checks by monitoring in the context of Common Agricultural Policy and the quantification of agricultural land abandonment in the Valencian Community region. One of the biggest challenges of remote sensing applications is to create accurate models that require little training data, reduce recalibration efforts, and that are spatially and temporally transferable. In the first steps of application of these products for monitoring the status of agricultural parcels, it is necessary to validate the models with more data and in other areas, since a lower performance is expected outside the study area and other dates. Future efforts should focus on quantifying the spatial–temporal transferability of the models. Furthermore, future research should clarify the optimal acquisition date of the images according to the rainfall and temperature regime.

5.5. CONCLUSIONS

This work assessed the potential of high-resolution spectral, SfM point clouds and textural descriptors to identify the status of citrus parcels in the Comunitat Valenciana region (Spain). Different analyses involving combinations of WorldView-3, Very High-Resolution Airborne Imagery and RPAS-SfM point clouds data were carried out to assess the impact of combining the different data sources on classification accuracy. It

was shown the high potential of WV-3 and VHRA imagery to identify citrus crops' status (greater than 0.900) from a single image. The SfM data showed a lower potential ($\approx$ 0.825) by itself. Combining multispectral and 3D point clouds produced little improvement in the classification accuracy. The WV-3 SWIR bands showed a high potential to detect abandoned parcels as they produce more spectral separability over the productive parcels in the 1570 nm – 2330 nm spectrum. The results also show the importance of GLCM texture features extracted from sub-metric images due to their ability to model spatial planting patterns typical of fruit crops.

This work provides new research for the operational implementation of checks by monitoring and quantification of agricultural land abandonment in the CV region. The experience of this research can help the implementation of tools that enable monitoring the citrus crop status in other areas and other similar perennial crops. The research provided useful information on which sources of remotely sensed data are most effective for citrus crop status classification, either on its own or in combination. It also showed the performance of two machine learning algorithms: Random Forest and Support Vector Machines. This provides valuable information for selecting the appropriate combination of remotely sensed data sources and machine learning algorithms.

5.6. REFERENCES

Alcántara, C., Kuemmerle, T., Prishchepov, A.V., Radeloff, V.C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sensing of Environment 2012, 124, 334–347. https://doi.org/10.1016/j.rse.2012.05.019.

Breiman, L. Random forests. Machine Learning 2001, 45, 5–32. https://doi.org/10.1023/A:1010933404324

Cortes, C., Vapnik, V. Support-vector networks. Machine Learning 1995, 20, 273–297. https://doi.org/10.1007/BF00994018

Czesak, B., Różycka-Czas, R., Salata, T., Dixon-Gough, R., Hernik, J. Determining the intangible: detecting land abandonment at local scale. Remote Sensing 2021, 13, 1166. https://doi.org/10.3390/rs13061166

Estel, S., Kuemmerle, T., Alcantara, ´ C., Levers, C., Prishchepov, A.V., Hostert, P. Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. Remote Sensing of Environment 2015, 163, 312–325. https://doi.org/10.1016/j. rse.2015.03.028

Geleralitat Valenciana, 2021. Portal Estadístico de la Generalitat Valenciana. Fichas Municipales. Available online: http://www.pegv.gva.es/es/fichas (accessed on 05 December 2021).

Gradinaru, S.R., Kienast, F., Psomas, A. Using multi-seasonal Landsat imagery for rapid identification of abandoned land in areas affected by urban sprawl. Ecological Indicators 2019, 96, 79–86. https://doi.org/10.1016/j.ecolind.2017.06.022

Hall-Beyer, M. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. International Journal of Remote Sensing 2017, 38 (5), 1312–1338. https://doi.org/10.1080/01431161.2016.1278314

Haralick, M.R., Shanmugam, K., Distein, I. Textural features for image classification. IEEE Transactions on Systems Man Cybernetics 1973, 3, 610–620. https://doi.org/10.1109/TSMC.1973.4309314

Houborg, R., McCabe, M.F. A cubesat enabled spatio-temporal enhancement method (CESTEM) utilizing planet, landsat and MODIS data. Remote Sensing of Environment 2018, 209, 211–226. https://doi.org/10.1016/j.rse.2018.02.067

Institut Valencià d'Investigacions Agràries – IVIA. Citricultura valenciana. http://gipcitr icos.ivia.es/citricultura-valenciana (accessed on 20 October 2022).

Karasiak, N., Dejoux, J.F., Monteil, C., Sheeren, D. Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. Machine Learning 2022, 111, 2715-2740. https://doi.org/10.1007/s10994-021-05972-1

Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., & Hussin, Y. A. Generating pitfree canopy height models from airborne lidar. Photogrammetric Engineering & Remote Sensing 2014, 80(9), 863-872. https://doi.org/10.14358/PERS.80.9.863

Kolecka, N.; Kozak, J.; Kaim, D.; Dobosz, M.; Ginzler, C.; Psomas, A. Mapping Secondary Forest Succession on Abandoned Agricultural Land with LiDAR Point Clouds and Terrestrial Photography. Remote Sensing 2015, 7(7), 00-8322. https://doi.org/10.3390/rs70708300

Kosmas, C.; Kairis, O.; Karavitis, C.; Acikalin, S.; Alcalá, M.; Alfama, P.; Atlhopheng, J.; Barrera, J.; Belgacem, A.; Solé-Benet, A.; et al. An exploratory analysis of land abandonment drivers in areas prone to desertification. CATENA 2015, 128, 252–261. https://doi.org/10.1016/j.catena.2014.02.006

Liaw, A.; Wiener, M. (2002). Classification and Regression by randomForest. R News 2(3), 18--22. https://cran.r-project.org/web/packages/randomForest/index.html

Löw F.; Prishchepov, A.V.; Waldner, F.; Dubovyk, O.; Akramkhanov, A.; Biradar, C.; Lamers, J.P.A. Mapping Cropland Abandonment in the Aral Sea Basin with MODIS Time Series. Remote Sensing 2018, 10(2), 159. https://doi.org/10.3390/rs10020159

Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A. Leisch, F. (2021). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-9. https://CRAN.R-project.org/package=e1071

Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2020; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2021. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2021; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2022.

Morell-Monzó, S.; Sebastiá-Frasquet, M.T.; Estornell, J. Land Use Classification of VHR Images for Mapping Small-Sized Abandoned Citrus Plots by Using Spectral and Textural Information. Remote Sensing 2021, 13(4), 681. https://doi.org/10.3390/rs13040681

Perpiña-Castillo, C.; Kavalov, B.; Diogo, V. Agricultural Land Abandonment in the EU within 2015–2030; (Technical Report JRC113718); European Commission: Brussels, Belgium, 2018.

Portalès-Julià, E.; Campos-Taberner, M.; García-Haro, F.J.; Gilabert, M.A. Assessing the Sentinel-2 Capabilities to Identify Abandoned Crops using Deep Learning. Agronomy 2021, 11(4), 654. https://doi.org/10.3390/agronomy11040654

Prins, A.J.; Van Niekerk, A. Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms. Geo-spatial Information Science 2021, 24. https://doi.org/10.1080/10095020.2020.1782776

Prishchepov, A. V. Agricultural Land Abandonment. Oxford Bibliographies. Environmental Science. Oxford University Press. 2020. https://doi.org/10.1093/obo/9780199363445- 0129

Prishchepov, A.V.; Radeloff, V.C.; Dubinin, M.; Alcantara, C. The effect of Landsat ETM/ETM image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Remote Sensing of Environment 2012, 126, 195–209. https://doi.org/10.1016/j.rse.2012.08.017

Richards, J. A. and Jia, X.Remote Sensing Digital Image Analysis: An Introduction. Berlin, Germany: Springer 2006, 349-351. https://link.springer.com/book/10.1007/3-540-29711-1

Richter, R. Schläpfer, D., 2019. Atmospheric and Topographic Correction (ATCOR Theoretical Background Document), DLR report:DLR-IB 564-03/2019, pp 142. https://www.rese-apps.com/pdf/atcor_atbd.pdf

Richter, R. Schläpfer, D., Atmospheric / Topographic Correction for Satellite Imagery, DLR report DLR-IB 565-01/2019, Wessling, Germany, pp 210. ATCOR-2/3 User Guide, Version 9.4.0, July 2021: https://www.rese-apps.com/pdf/atcor3_manual.pdf

Ruiz, L.A.; Almonacid-Caballer, J.; Crespo-Peremarch, P.; Recio, J.A.; Pardo-Pascual, J.E.; Sánchez-García, E. Automated classification of crop types and condition in a Mediterranean area using fine-tunned convolutional neural network. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B3-2020, 2020.

Saini, R., Ghosh, S.K. Crop classification on single date Sentinel-2 imagery using Random Forest and Support Vector Machine. The International Archives of Photogrammetry, Remote Sensing and Spatial Information, Volume XLII-5, 2018.

Sheykhmousa, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P. and Homayouni, S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2020, 13, 6308-6325. https://doi.org/10.1109/JSTARS.2020.3026724.

Stock, A. Spatiotemporal distribution of labeled data can bias the validation and selection of supervised machine learning algorithms: A marine remote sensing example. ISPRS Journal of Photogrammetry and Remote Sensing 2022, 187, 46-60. https://doi.org/10.1016/j.isprsjprs.2022.02.023

Strijker, D. Marginal lands in Europe - causes of decline. Basic and Applied Ecology 2005, 6, 99–106. https://doi.org/10.1016/j.baae.2005.01.001

Subedi, R.Y.; Kristiansen, P.; Cacho, O. Drivers and consequences of agricultural land abandonment and its reutilization pathways: A systematic review. Environmental Development, in press. https://doi.org/10.1016/j.envdev.2021.100681

Szostak, M.; Hawryło, P.; Piela, D. Using of Sentinel-2 images for automation of the forest succession detection. European Journal of Remote Sensing 2017, 51, 142–149. https://doi.org/10.1080/22797254.2017.1412272

Tan, Y.; Wang, S.; Xu, B.; Zhang, J. An improved progressive morphological filter for UAV-based photogrammetric point clouds in river bank monitoring. ISPRS Journal of Photogrammetry and Remote Sensing 2018, 146, 421-429. https://doi.org/10.1016/j.isprsjprs.2018.10.013

Vajsová, B.; Fasbender, D.; Wirnhardt, C.; Lemajic, S.; Devos, W. Assessing Spatial Limits of Sentinel-2 Data on Arable Crops in the Context of Checks by Monitoring. Remote Sensing 2020, 12, 2195. https://doi.org/10.3390/rs12142195

Volpi, I.; Marchi, S.; Petacchi, R.; Hoxha, K.; Guidotti, D. Detecting olive grove abandonment with Sentinel-2 and machine learning: The development of a web-based tool for land management. Smart Agricultural Technology 2023, 3, 100068. https://doi.org/10.1016/j.atech.2022.100068

Wang, Q.; Atkinson, P.M. Spatio-temporal fusion for daily Sentinel-2 images. Remote Sensing of Environment 2018, 204, 31-42. https://doi.org/10.1016/j.rse.2017.10.046

Wadoux, A.M.J.-C, Heuvelink, G.B.M., de Bruin, S., Brus, D.J. Spatial cross-validation is not the right way to evaluate map accuracy. Ecological Modelling 2021, 457. https://doi.org/10.1016/j.ecolmodel.2021.109692

Yin, H.; Prishchepov, A.V.; Kuemmerle, T.; Bleyhl, B.; Buchner, J.; Radeloff, V.C. Mapping agricultural land abandonment from spatial and temporal segmentation of Landsat time series. Remote Sensing of Environment 2018, 210, 12–24. https://doi.org/10.1016/j.rse.2018.02.050

Zhang, K., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J., & Zhang, C. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE Transactions on Geoscience and Remote Sensing 2003, 41, 4, 872-882. https://doi.org/10.1109/TGRS.2003.810682

Zhao, Y.; Liu, D. A robust and adaptive spatial-spectral fusion model for PlanetScope and Sentinel-2 imagery. GIScience & Remote Sensing 2022, 59, 520-546. https://doi.org/10.1080/15481603.2022.2036054

CAPÍTULO 6

Detecting abandoned citrus crops using Sentinel-2 time series. A case study in the Comunitat Valenciana region (Spain)

6.1. INTRODUCTION

Agricultural land abandonment (ALA) is a growing phenomenon around the world that has received attention since the early 1990s, however, around 95% of the papers were published in the last 15 years (Subedi et al., 2021). This indicates a growing concern in the scientific community. In the European Union (EU), agriculture represents more than half of the territory. Agriculture guarantees food production, manages important natural resources, and supports the socio-economic development of rural areas (Terres et al., 2015). However, around 11% of the agricultural area is at high risk of abandonment for the period 2015-2030 (Perpiña-Castillo et al., 2018). ALA is a major obstacle to the sustainable development of many European regions (Leal Filho et al., 2017) and according to projections will increase in Europe over the next decades (Terres et al., 2013). In this context, it is necessary to develop time and cost-efficient methods to identify ALA to update inventories of land use. Temporally and spatially explicit information about ALA can help to support policy instruments to counteract or reverse the process and can also help implement adequate landscape management measures (Volpi et al., 2023).

The development of the Copernicus program through the European Space Agency (ESA) and the European Commission has meant a great advance in remote sensing systems that have significantly improved agricultural monitoring techniques (Phiri et al., 2020). ALA is particularly frequent in mountainous areas and in areas with highly fragmented landscapes (Czesak et al., 2021). In these areas, the spatial resolution of other commonly used sensors, such as MODIS (Moderate Resolution Imaging Spectroradiometer) (500 m) or Landsat (30 m) is not enough to precisely monitor the small-sized parcels. The higher spatial resolution of Sentinel-2 (10-20m) can improve these capabilities. Morell-Monzó et al. (2020) used the Sentinel-2 spectral data in the first experience to identify the abandonment status of citrus crops obtaining an overall accuracy of around 77% using a single scene. The agricultural landscape of this crop in the Valencian region (Spain) is characterized by small-sized parcels, between 0.25 and 0.50 ha, especially in coastal areas.

The information in the spectral domain of the Sentinel-2 data provides great opportunities to discriminate between types of crops with different spectral signatures (Asgarian et al., 2016). However, information in the temporal domain is also important because it allows tracking the evolution of a crop in time. The higher temporal resolution of Sentinel-2 has been an opportunity to improve the identification and mapping of crop types and crop rotations (Vuolo et al., 2018). The inherently dynamic nature of crops and their spectral-temporal signatures make classification approaches based on a single date limited (Vuolo et al., 2018). In many experiments, the use of multi-temporal data has shown better results than approaches based on a single date (Gómez et al., 2016). However, the benefits of multi-temporal data for detecting the status of permanent evergreen crops, such as citrus, may be not so obvious. Unlike seasonal crops, the growth of citrus to productive age lasts about 5 years (reaching its maximum growth lasts 15 years). So, during the half-life of the crop usually there is only one plantation process, or if repeated is after more than fifteen years. The growth of trees is relatively slow until it reaches a mature state when the growth is controlled by pruning tasks. Harvesting does not produce important reflectance changes as in annual or bi-annual crops. Citrus cultivated in Europe are irrigated crop. When a

productive parcel is abandoned, its evolution will be more conditioned by the temperature and rainfall regime since it lacks irrigation. Being abandonment a long process in time, orchard spectral signature changes little and gradually, making it difficult to detect abandoned parcels using a single scene. These characteristics make detecting the status of citrus crops especially challenging.

Spain is the largest citrus producer in Europe, the fourth in the world, and the first exporter of citrus for fresh consumption. Comunitat Valenciana region accounts for more than 50% of the Spanish citrus production. A massive abandonment of citrus orchards has been experienced from 2000 to 2020, representing a 20% decrease of the cultivated surface (MAPA, 2021). In this region the current estimated area dedicated to citrus 160,088 ha (MAPA, 2022). In this context, it is of priority interest to develop time and cost-efficient methods to identify and quantify citrus crop abandonment.

The main objective of this work is to propose a method to automatically identify the status of citrus crops using Sentinel-2 data. For this purpose, a machine learning-based classification framework was developed using time series of spectral indices derived from Sentinel-2 to identify the crop status at parcel level. The impact of multi-temporal Sentinel-2 data was assessed, and the results were compared with our previous experiences (Morell-Monzó et al., 2020). In addition, motivated by the promising results of other authors (Vuolo et al., 2018), we assess the transferability of spectral-temporal signatures across regions and years inside the Comunitat Valenciana region (Spain).

6.2. DATA AND METHODS

6.2.1. Study area

Our study area is in La Safor, a coastal region of 43,000 ha, approximately centered at (34º56'14''N, 0º8'42''E) in Comunitat Valenciana. It was selected due to its large tradition of citrus production. Its landscape is characterized by two clearly differentiated structural elements. On the west, a mountainous forest area. On the east a plain alluvial zone and the coastline, which is highly populated and where most of the irrigated agricultural production is located. This plain was generated by successive sedimentary systems of Quaternary origin. It stands out for its flat relief, soils with high agronomic potential and high availability of water that have made it a historically agricultural area. The major crops in this region are citrus representing 95% of cultivated area (Generalitat Valenciana, 2021). However, in recent years a large part of the farms has been abandoned. The agrarian structure is characterized by the small size of the farms (between 0.25 and 0.5 ha) resulting in a highly fragmented landscape.

This region has a typical Mediterranean climate with two distinct seasons, wet and dry (Gasith & Resh, 1999; González-Hidalgo et al., 2005). The characteristic rainfall is scarce and torrential and has an extremely high spatial and temporal variability being the raining seasons autumn and spring. In contrast, summer (June to August) is the dry period in this Mediterranean zone.

6.2.2. Map classes and reference data

This work proposes a classification based on three orchard status: non-productive (NP), productive (PR), and abandoned (AB) (Figure 6.1). NP class includes bare soil orchards

before tree planting and recently planted trees that have not reached productive age and are predominantly occupied by bare soil. PR class includes orchards at productive age. Finally, AB class includes visually evident abandoned orchards.

Major changes that occur in an abandoned orchard are the growth of spontaneous vegetation, loss of vigor of the trees, and, therefore, progressive loss of the typical planting patterns of fruit crops. However, growers may carry out management practices that generate different evolutions on abandoned orchards. In some cases, farmers cut the tree canopies, then sprouts and small branches may appear. In other cases, farmers carry out periodic pruning or weeding, limiting the height and density of spontaneous vegetation. In general, the lack of water leads to the progressive death of the trees at different rates.

Figure 6.1. Examples of the three types of abandonment status: non-productive (left), productive (center), and abandoned (right) parcels. The grey grid corresponds to the 10-meter Sentinel-2 pixels.

Identifying AB parcels is essential for monitoring citrus land abandonment and differentiating PR from NP parcels is interesting to improve citrus yield estimations. Declared citrus parcels are georeferenced in the Land Parcel Identification System of Spain (SIGPAC). However, this database has no information on crop status. The proposed classification framework was designed to complement the SIGPAC information and be applied to previously identified citrus parcels.

Seven ground truth datasets were obtained through: a) field inspection campaigns, b) from images captured by unmanned aerial vehicles (UAV), and c) from very-high resolution orthophoto photointerpretation. Field campaigns and UAV data were obtained in previous published studies (Morell-Monzó et al. 2020; Morell-Monzó et al., 2021), and for this research we enlarged the dataset with parcels classified by photointerpretation of very-high resolution orthophotos. Table 6.1 summarizes these ground truth sources. The first dataset (OLV-19) contains 240 parcels classified during a field campaign carried out between July 11 and 14, 2019 in the municipality of Oliva (Figure 6.2b). Parcels have an average size of 0.41 ha, summing up 100.5 ha. They were selected through stratified random sampling with equal proportion for each category. This dataset contains 80 NP, 80 PR, and 80 AB parcels. The second dataset (LSF-20) consists of 280 parcels classified by visual interpretation of UAV photogrammetric flights carried out in February 2021 (Figure 6.2c). Parcels were selected through a systematic and intensive sampling of the area covered by the UAV. The data set 60 NP (21.4%), 120 PR (42.9%), and 100 AB parcels (35.7%). The parcels have an average size of 0.48 ha, summing up 136.7 ha. The third dataset (LSF-21) consist of 836 parcels classified during field campaigns carried out from March 9 to May 28, 2021 (Figure

6.2d). This dataset contains 188 NP (22.5%), 368 PR (44%), and 280 AB parcels (33.5%). The parcels have an average size of 0.44 ha, summing up 368.2 ha. All three datasets were in the coastal area of La Safor, which is about 20 km long and 6 km wide, hereafter named the training area. To avoid possible errors caused by changes in land use between field campaigns and the acquisition date of the images, abandonment status was verified using high-resolution airborne images from the Valencian Cartographic Institute (ICV) https://idev.gva.es/va/, taken annually in May.

Table 6.1. Summary and characteristics of the ground truth data used.

Dataset	Number of parcels	Sampling method	Acquisition date (verification)	Sampled area (Ha)	Average parcel size (Ha)	NP parcels ratio	PR parcels ratio	AB parcels ratio
OLV-19	240	Field campaigns	July 11-14, 2019 (May 14, 2019)	100.5	0.41	33.3%	33.3%	33.3%
LSF-20	280	UAV photointerpretation	February 1-29, 2020 (May 6, 2020)	136.7	0.48	21.4%	42.9%	35.7%
LSF-21	836	Field campaigns	March 9, 2021 (May 16, 2021)	368.2	0.44	22.5%	44.0%	33.5%
PV-21	80	Photointerpretation	May 16, 2021	34.3	0.36	21.3%	45.0%	33.7%
TV-21	80	Photointerpretation	May 15, 2021	36.6	0.46	12.6%	63.7%	23.7%
BP-21	80	Photointerpretation	May 13, 2021	65.9	0.86	16.3%	55.7%	25.0%
NL-21	80	Photointerpretation	July 22, 2021	39.5	0.49		55.1%	24.9%

Additionally, 4 datasets were generated outside the training area (Figure 6.2a). These datasets are located in the municipalities of Potries and Villalonga (PV-21), Tavernes de la Valldigna (TV-21), Benicull and Polinyà del Xuquer (BP-21), and Nules (NL-21). Datasets PV-1 and TV-21 are located in La Safor region, close to the training area, while datasets BP-21 and NL-21 are located far from the study area. Each of these datasets contains 80 parcels randomly samped and classified by photointerpretation of the 2021 orthophoto offered by ICV. Orthophotos from previous years and Street View images from Google Maps were also used for verification.
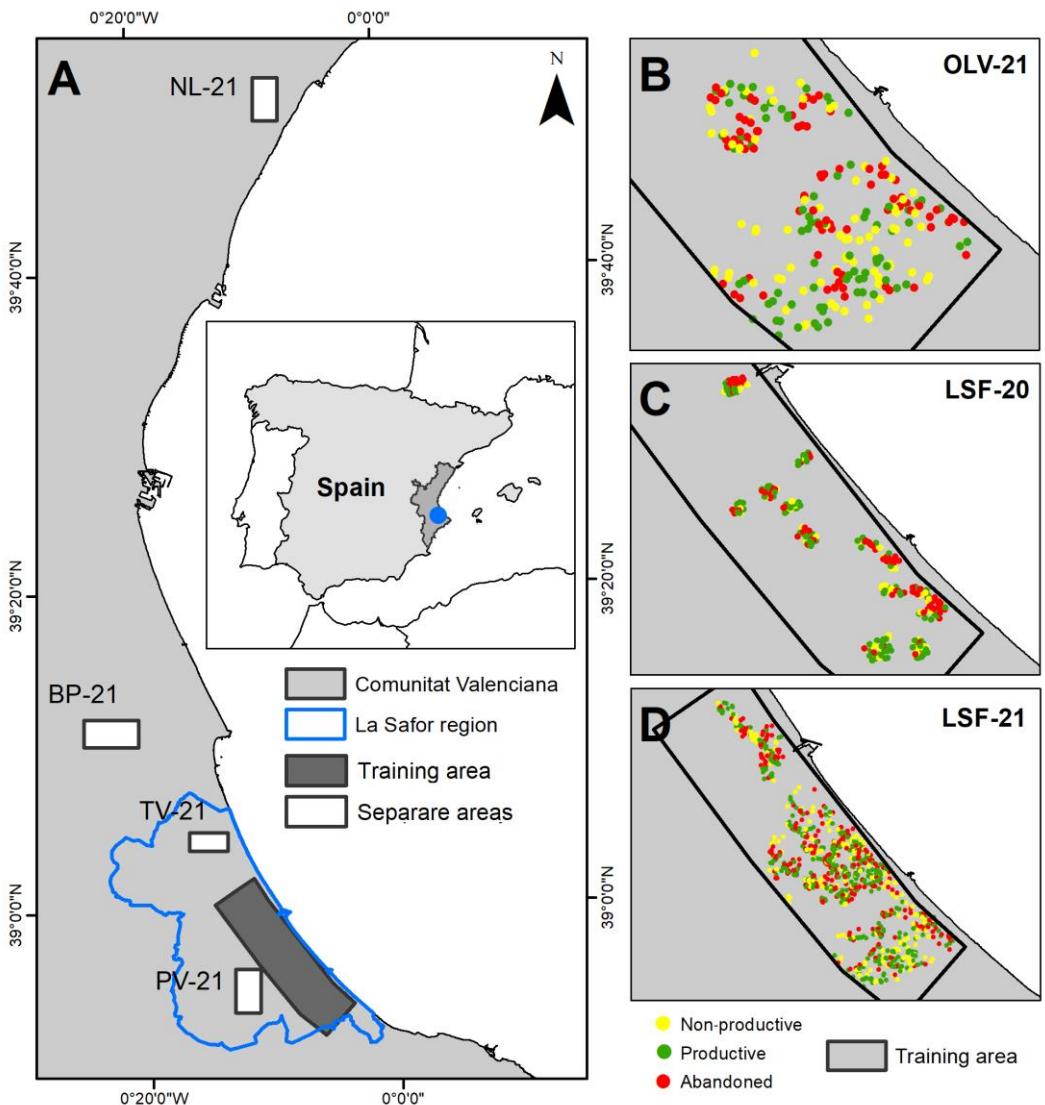
Figure 6.2. Study area and ground truth datasets. (A) Study area in La Safor, Comunitat Valenciana region (Spain) and ground truth datasets. (B) OLV-19 dataset. (C) LSF-20 dataset. (D) LSF-21 dataset.

6.2.3. Sentinel-2 time series processing and feature extraction

Three annual collections of Sentinel-2A/B L2A bottom-of-atmosphere (BOA) reflectance images were acquired from the Google Earth Engine platform (Gorelick et al., 2017) generating annual data time series: a) September 2018 - September 2019, representing the 2019 season, b) September 2019 - September 2020, representing the 2020 season and c) September 2020 - September 2021, representing the 2021 season. The time series consist of 65, 57 and 54 images respectively. Images were projected in UTM/WGS84 30N and they were filtered to include only those with less than 75% cloud percentage. The L2A product provides a scene classification band (SC). Pixels classified as no data, saturated or defective, cast shadows, cloud shadows, unclassified, cloud medium and high probability, and thin cirrus in the SC band were masked in all reflectance bands. These bands were resampled to 10 m resolution by pixel disaggregation of the lower resolution bands. Finally, two spectral indices were calculated using the bands B4, B8, and B11: the Optimized Soil Adjusted Vegetation Index (OSAVI) (Rondeaux, 1996) and the Normalized Difference Moisture Index

(NDMI) (Gao et al., 1996). These spectral indices were selected based on our previous work (Morell-Monzó et al., 2022) where they obtained the best discriminating results among a total of 15 indices.

For Sentinel-2 images, OSAVI is calculated from the near-infrared band (B8 ~0.83 µm) and red band (B4 ~0.66 µm). The OSAVI is an optimization of the Soil Adjusted Vegetation Index (SAVI) proposed by Huete (1988). This index uses a canopy bottom adjustment value (L), which is a function of vegetation density. OSAVI uses a value L=0.16 proposed by Rondeaux (1996). This index is suitable when the vegetation is sparse or when the soil is visible through the canopy, which is common in fruit crops. NDMI is calculated from the near-infrared band (B8 ~0.83 µm) and short-wave infrared band (B11 ~1.61 µm). The near-infrared spectral channel collects the reflectance of the internal structure of the leaf and the dry matter content of the leaf. At this wavelength the absorption produced by liquid water in the soil and vegetation is insignificant. In contrast, the short-wave infrared spectral channel is sensitive to soil and vegetation water content and to the internal mesophyll structure of the leaf. As a result, the NDMI is an indicator of the moisture content of soil and vegetation. Furthermore, the NDMI is less sensitive to atmospheric effects than other spectral indices such as the Normalized Difference Vegetation Index (NDVI) (Gao et al., 1996).

Both OSAVI and NDMI time series contained temporal gaps due to the removal of images (those having more than 75% coverage) and spatial gaps due to masked pixels. In order to create continuous time series with non-empty values every 5 days, those gaps were filled by linear interpolation expressed in equation 6.1.

$$Y_i = Y_{i-1} + (Y_{i+1} - Y_{i-1}) * \left(\frac{t_i - t_{i-1}}{t_{i+1} - t_{i-1}} \right) \tag{6.1}$$

where Y_i is the interpolated pixel value, Y_{i-1} is the pixel value of the previous image, Y_{i+1} is the pixel value of the next image, t_i is the time value of the image containing the pixel to interpolate, and t_{i-1} and t_{i+1} are the time values of the previous and next images.

To reduce noise in the time series introduced by the undetected clouds, partial cloud shading or unfavorable atmospheric conditions, the original OSAVI and NDMI signals were smoothed using the Savitzky-Golay filter (Savitzky & Golay, 1964; Chen et al., 2004). This is a moving window filter in the time domain that removes outliers and smoothes data in the time series while preserving the trends. It is based on a simplified least square fitting convolution to smooth and compute derivatives of a set of consecutive values. The convolution can be understood as a weighted moving average filter, which uses weights given by a polynomial of a user-defined degree. When these weights are applied to the signal, they perform a least squares polynomial fit within the filter window (Chen et al., 2004). In this work, a moving window of 50 days and a third-order polynomial were used.

Features were extracted from the smoothed signals to be used as inputs to train a supervised classifier. They can be categorized as: a) central tendency statistics: mean (12 per year); b) dispersion statistics: range (2 per year) and standard deviation (2 per year); c) distribution shape statistics: skewness (1 per year) and kurtosis (1 per year) of each spectral index, providing a total of 36 features per pixel of the smoothed time series. The length of the intervals used for feature extraction was optimized through a search grid with values of 12, 6, 4, 3, 2, and 1 month. Finally, monthly intervals were

used for the central tendency statistics, semi-annual intervals for the dispersion statistics, and annual intervals for the distribution shape statistics (Figure 6.3).
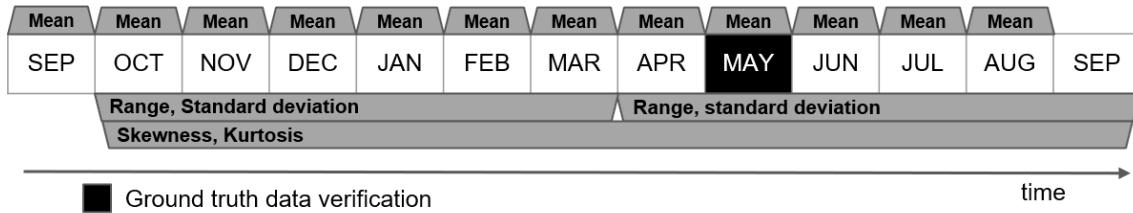
Figure 6.3. Feature extraction by time intervals from Sentinel-2 OSAVI and NDMI smoothed time series.

6.2.4. Parcel-based classification

In this work, an ad hoc parcel-based classification method was developed. A parcel-based classification allows spatial indexing with agricultural databases (e.g., SIGPAC). The proposed method is described below:

1. Using the extracted features, a supervised classification model is trained using pixels as training samples. Only pixels completely covered by the parcels are used for training.

2. Then the fitted model is used to predict the probability of each pixel being in each class (Figure 36). Therefore, for each pixel of the image, a probability vector with length C is predicted, being C the number of classes. In this work the number of classes is 3.

3. Finally, parcel-based classification is performed. For each parcel, the average probability of each class is calculated using the full pixels inside the parcel boundaries (Figure 36). Full pixels are those that have a coverage threshold of 100%, however, the method iteratively reduces this threshold to 75% and 50% in case there are no full pixels inside the parcel boundaries. Finally, each parcel is assigned the class with the maximum probability.

It should be outlined that in majority voting (Figure 36, left), a classification value is predicted for each pixel and then parcel is classified by applying the mode operator. In the proposed method (Figure 36, right), the class probability is predicted for each pixel. Then, the probability values for each class are averaged and finally, each parcel is classified according to the class with the maximum probability, thus applying the argument of the maxima (argmax). In practice, the outcome of the proposed method is similar to the majority voting method. However, the use of pure pixels is especially interesting in highly fragmented agricultural structures since it allows for avoiding edge effects. Additionally, this method returns the probability of each parcel being in each category. This data, combined with the number of complete pixels in the parcel, can be employed to automatically filter parcels for manual class re-assignment, removal, or further specific processing.
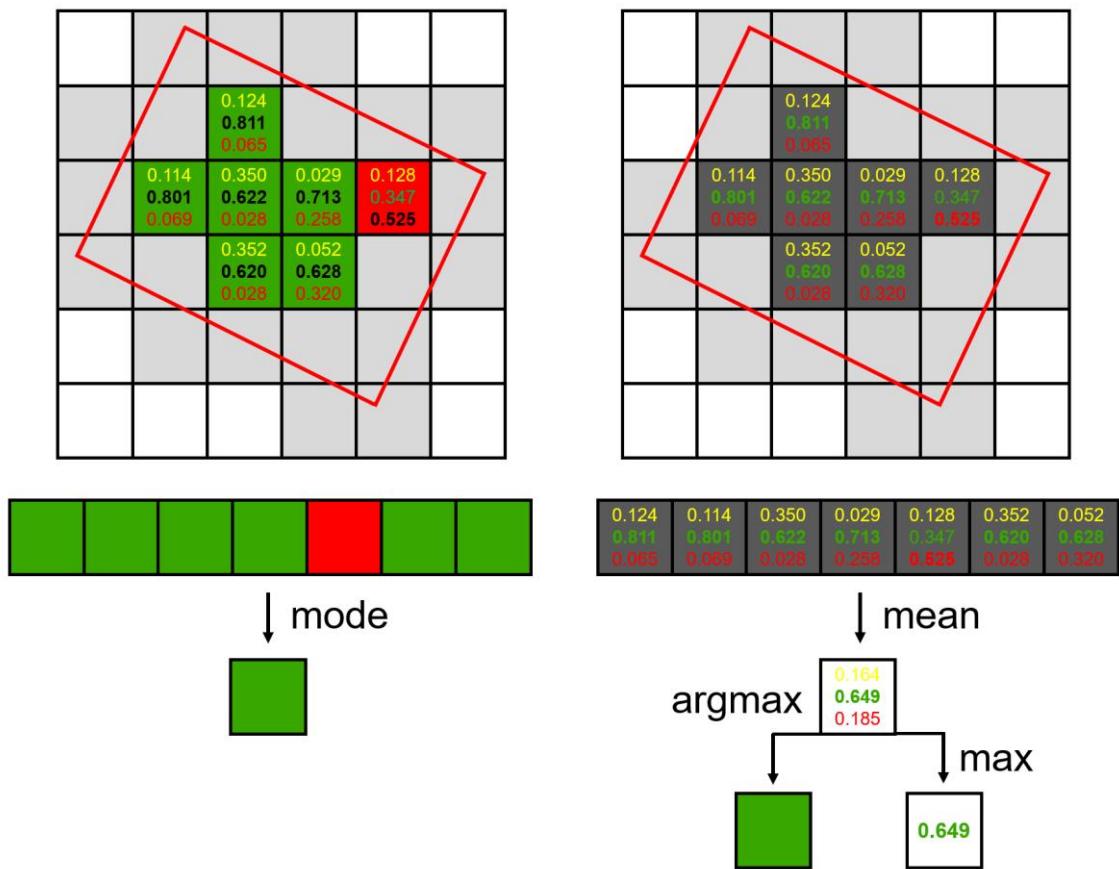
Figure 6.4. Differences between classification based on majority voting (left) and the proposed method based on maximum averaged probability (right).

The Random Forest (RF) method (Breiman, 2001) was used to build the pixel classifier. Although there is a growing trend toward the use of deep learning algorithms, RF remains as a benchmark algorithm for many remote sensing problems (Sheykhmousa et al., 2020). This is due to its good performance with small amounts of training data and its robustness against overfitting. RF requires tuning two key hyperparameters: the number of trees (`ntree`) that make up the forest and the number of randomly selected variables at each node split (`mtry`). RF models was trained with an increasing `ntree` up to 150 trees and the hyperparameter `mtry` was set to 6, corresponding to the square root of the number of variables. The OOB samples were used to measure model convergence. In this work, the RF algorithm was applied from the `randomForest` R package (Liaw & Wienner, 2002).

6.2.5. Accuracy assessment

One of the greatest challenges in many machine learning applications is to build accurate and robust models that reduce recalibration efforts and ensure model transferability to never seen data distributions (Dimov, 2022). Three models for citrus crop status identification were trained using the OLV-19, LSF-20, and LSF-21 datasets. These models were validated using two cross-validation strategies. First, 4-fold cross-validation was performed using random splits without replacement. In each iteration, 3/4 of the parcels were used for training and 1/4 of the parcels were used for validation,

ensuring balanced data sets with the same number of parcels in each class. Second, 4-fold spatial cross-validation was performed. This strategy was used to avoid an underestimation of the model error due to spatial autocorrelation of the samples, which is inherent in remote sensing data (Karasiak et. al., 2021).

Random cross-validation is suitable when the data is randomly distributed. However, when the data is strongly aggregated, this strategy may underestimate the model error (Wadoux et. al., 2021; Stock et al., 2022). In our case, the LSF-20 dataset is clustered due to the sampling method through UAV flights. The other datasets were obtained through random sampling. However, the agricultural areas are not randomly distributed in the area, which causes a given degree of samples aggregation. These circumstances motivated the use of spatial cross-validation. Spatial partitions were generated using the k-Means method on the XY coordinates of the parcels. This procedure allows for generating data partitions with the maximum distance between groups, although in this case it cannot be ensured that the data is balanced.

To assess the model performance the following accuracy metrics were calculated: overall accuracy (OA), Cohen's kappa index, precision (producer's accuracy by class), and recall (user's accuracy by class). Due to the small size of the agricultural parcels, relationships between model accuracy and parcel size were explored using the Kruskall-Wallis non-parametric test. The effect of dataset training size on model performance was also studied.

One of the biggest challenges for the operational implementation of a remote sensing-based monitoring system is to obtain robust spatial and temporal transferability. For this reason, the spatial and temporal transferability of the trained models was additionally evaluated with the aim of knowing the generalization capacity outside the training area and in years different from the training one. To assess spatial transferability, data from separate areas (PV-21, TV-21, BP-21, NL-21) were used. The model trained in the study area with LSF-21 data was used to evaluate its performance outside the study area without using additional data to calibrate the model. To evaluate the temporal transferability of the models, cross-validation of the models trained in the years 2019, 2020, and 2021 was carried out. Each model was evaluated in the two years other than the training year without using additional training samples.

6.2.6. Influence of features on class separability

The Jeffries-Matusita (JM) distance measures the separability between a pair of probability distributions. This method does not inform about the influence of a descriptor on the model performance, but rather collects the intrinsic structure of the variables. Unlike methods based on permutation or impurity (such as measures of the importance of RF features), this statistical method is not biased by collinearity or cardinality. The JM distance is widely used as a separability criterion for evaluation of classification results. In this work the JM distance defined in Richards & Jia (2006) was used, so it takes the range [0,2]. Class separability of features based on JM distances can be illustrated in radar charts.

6.3. RESULTS

6.3.1. Spectral-temporal signature

This section shows the temporal behaviour of the spectral indices for the three types of parcels studied. Figure 6.5 shows the average time series of OSAVI and NDMI for each crop status. Regarding OSAVI, in average, NP parcels had lower values than PR and AB parcels along the three years of study, being easily separated. A periodical behaviour is observed in the NP profile with a maximum between December – February and a minimum between June – August. This cyclical behaviour corresponds to the growth of spontaneous vegetation in the cold and wet winter and its wilt in the hot and dry summer. The effect of spontaneous vegetation cycles is more evident in the NP parcels than in the others, since planted tree canopies represent only a small part of the surface. In contrast, PR and AB parcels had a higher OSAVI values along the whole period of study and showed much lower separability. The OSAVI profile of the PR and AB classes also showed periodic but less defined behaviour. Mayor differences between PR and AB were found in the three summers (June – August). In these seasons, AB parcels reduced OSAVI levels more than PR parcels.

Similar patterns are observed from NDMI signals, with maximum values in winter and lower values in summer. NP parcels have a had lower NDMI values than PR and AB parcels, showing a higher separability with to PR and AB classes. NDMI differences between PR and AB parcels are greater than OSAVI. Based on these results, the major challenge is to separate the PR and AB classes. In this sense, NDMI produces greater differences between these classes. However, NDMI differences are severely reduced in May – June.
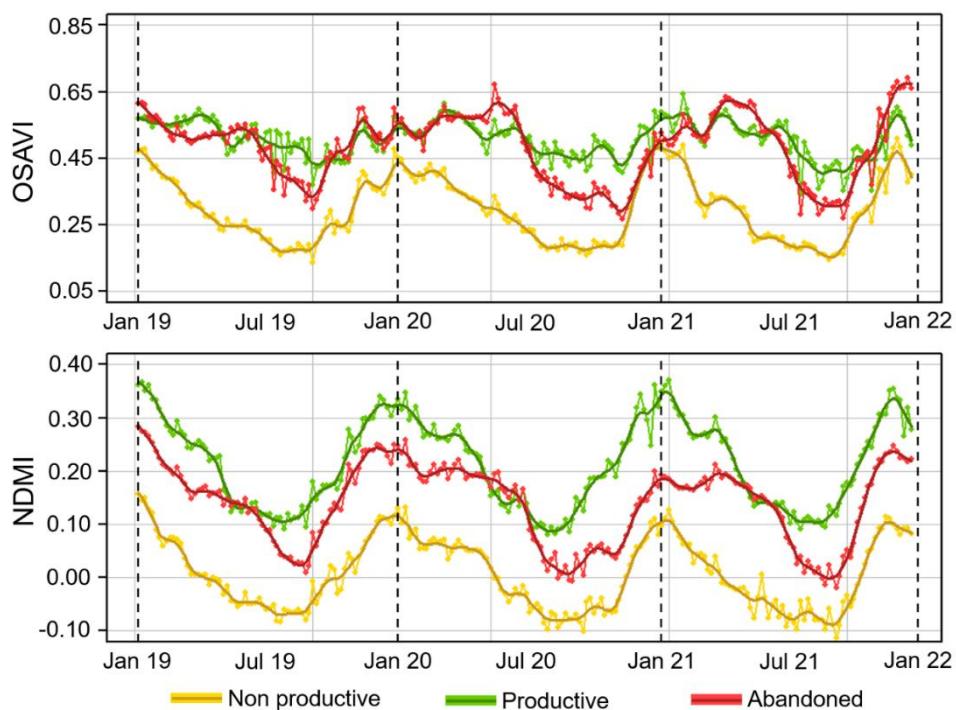
Figure 6.5. Average temporal profile of OSAVI and NDMI for the three classes: non-productive, productive and abandoned. Dots joined by light color lines represent original values, dark color lines are the smoothed time series after applying the Savitzky-Golay filter.

Figure 6.6 shows OSAVI evolution of three sample parcels from 2018 to 2022. PR and AB parcels have higher OSAVI values (between 0.65 and 0.30). In contrast, the NP parcel, having a limited tree canopy and large areas of bare soil during most of the year, had lower values (between 0.25 and 0.05). The PR parcel has higher OSAVI values than the AB parcel. In addition, the PR parcel is more stable over time. Moreover, the effect of changing the status from AB to NP can be observed at Figure 6.6c. Figure also shows the effect of the Savitzky-Golay filter. The filter smooths the original signal and reduces noise (sudden lower/higher OSAVI values), mainly caused by cloudy pixels not detected by the SC band of the Sentinel-2 L2A product.

ç
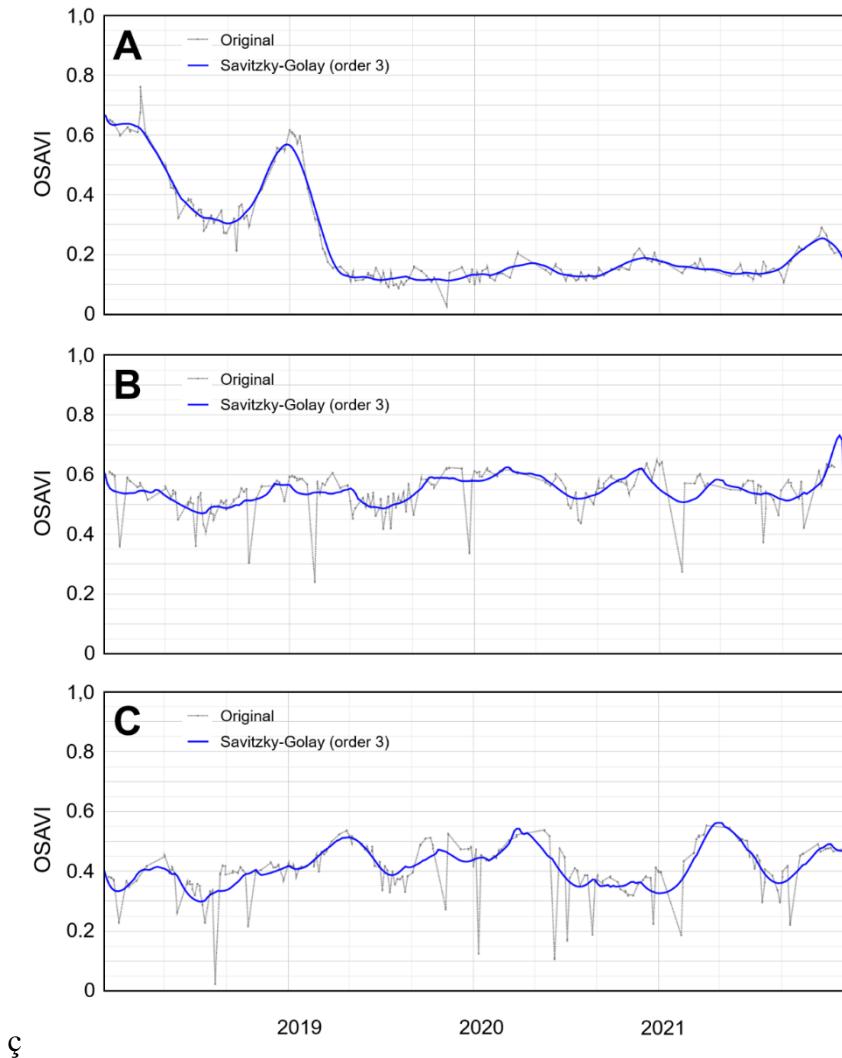
Figure 6.6. OSAVI from 2018 to 2022 of three sample parcels: (A) productive parcel, (B) abandoned parcel, (C) productive parcel replanted at the beginning of 2019.

6.3.2. Random Forest model

All RF models built during cross-validation converged before 150 decision trees (between 50 and 100 trees). A larger number of trees did not result in a reduction of OOB-error. The average OOB-error was reduced to 2.0%, 2.1% and 3.7% for the OLV-19, LSF-20 and LSF-21 datasets, respectively. Therefore, a higher OOB-error was

obtained as the number of parcels in the dataset increased. In all three datasets, the NP class produced the lowest error (0.8-3.2 %), followed by the PR class (1.4-2.3 %), and the AB class (2.9-7.0 %).

6.3.3. Accuracy assessment

This section evaluates the performance of the classification models. Figure 6.7 shows a summary of the OA obtained with each dataset and each validation strategy. Considering all the experiments, the average OA ranged from 0.89 to 0.96, with an average of 0.92. For the OLV-19 dataset, the OA obtained through random cross-validation was 0.92 +/- 0.03 while through spatial cross-validation was 0.89 +/- 0.04. For the LSF-2020 dataset, the OA obtained through random cross-validation was 0.96 +/- 0.02 while through spatial cross-validation was 0.94 +/- 0.03. Finally, for the LSF-21 dataset, the OA obtained through random cross-validation was 0.90 +/- 0.02 while through spatial cross-validation was 0.89 +/- 0.02. All the above errors expressing a 95% confidence interval.

OA was higher for the LSF-20 dataset. This fact agrees with our previous experiences (Morell-Monzó et al. 2023) and suggest that this dataset is less demanding than the rest. It can be attributed to the fact orchards were selected systematically, thus being types of abandonment more evident types of abandonment or a more advanced process of abandonment that made them easily detectable. Consequently, we propose the more conservative OA values (0.89 to 0.92). Additionally, small (up to 3%) OA differences between random cross-validation and spatial cross-validation, were found. These differences were smaller in the larger dataset (LSF-21).
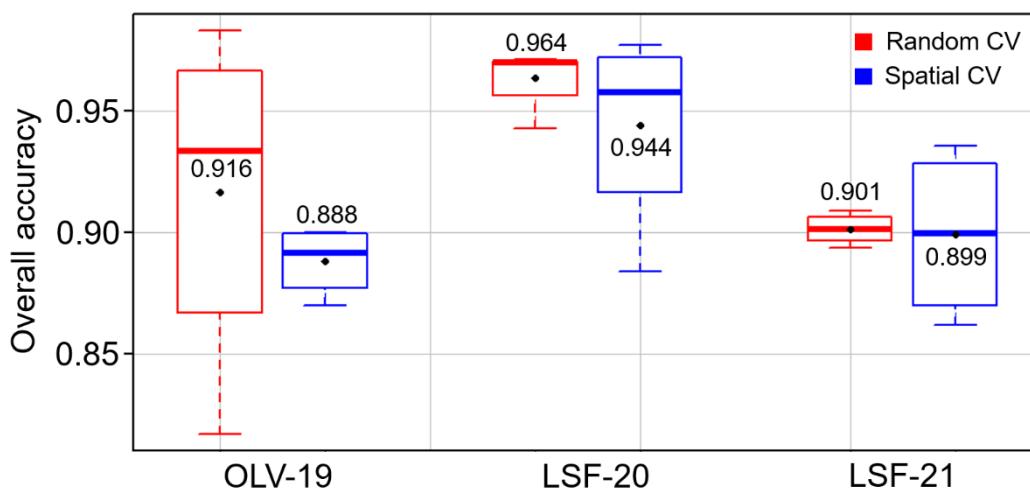
Figure 6.7. Overall accuracy of the models obtained using 4-fold cross-validation for the OLV-19, LSF-20 and LSF-21 datasets. In red are shown the results obtained using the random cross-validation strategy and in blue those obtained using the spatial cross-validation strategy. Dots show the average overall accuracy.

Below, the model performance by classes is studied. Table 6.2 shows precision and recall by classes of random and spatial cross-validation on the 3 datasets. The following average class accuracy was obtained: NP = 0.935 > PR = 0.921 > AB = 0.899. NP class was the class detected with more accuracy, while AB was the most difficult to detect. On the other hand, recall values were in general higher than 0.900, except for the AB class both in the OLV-19 and LSF-21 datasets. This last dataset produced an omission

error of around 18% for the AB class, which was mostly misclassified as PR. These results suggest a limitation of the model to detect some AB parcels.

Table 6.2. Cross-validation results of the Random Forest model trained on the OLV-19, LSF-20, and LSF-21 datasets.

Dataset	Performance	Random cross-validation	Spatial cross-validation
OLV-19	Overall accuracy (95% C.I.)	0.916 +/- 0.03	0.888 +/- 0.04
	Cohen's Kappa	0.875	0.827
	NP precision	0.942	0.932
	PR precision	0.878	0.857
	AB precision	0.932	0.883
	NP recall	0.962	0.942
	PR recall	0.912	0.916
	AB recall	0.875	0.805
LSF-20	Overall accuracy (95% C.I.)	0.964 +/- 0.02	0.944 +/- 0.03
	Cohen's Kappa	0.944	0.909
	NP precision	0.983	0.985
	PR precision	0.945	0.918
	AB precision	0.980	0.956
	NP recall	0.900	0.897
	PR recall	0.991	0.972
	AB recall	0.970	0.926
LSF-21	Overall accuracy (95% C.I.)	0.901 +/- 0.02	0.899 +/- 0.02
	Cohen's Kappa	0.846	0.841
	NP precision	0.915	0.915
	PR precision	0.883	0.890
	AB precision	0.919	0.915
	NP recall	0.919	0.934
	PR recall	0.953	0.941
	AB recall	0.820	0.818

Regarding the effect of parcel size in the results, Figure 6.8 shows the size of correctly classified and misclassified parcels. In all datasets, the model showed higher accuracy on larger parcels. However, Kruskal-Wallis's analysis of variance showed no significant differences between the size of correctly classified and misclassified parcels.
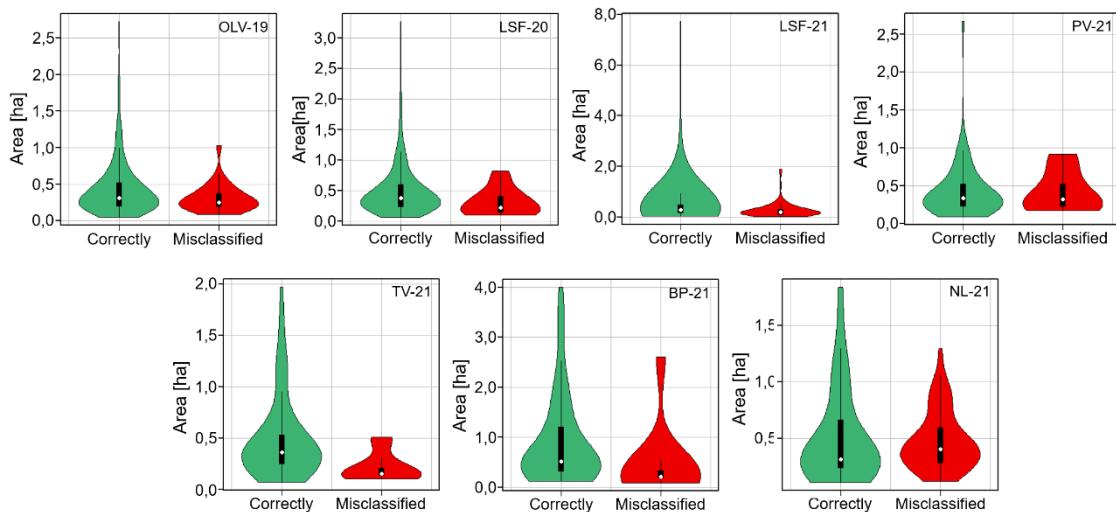
Figure 6.8. Size of correctly classified and misclassified parcels in each data set.

Regarding the influence of the size of the training dataset on model performance, Figure 6.9 shows OA of the models using increasing proportions of the training data. As expected, a decrease in model performance is observed as the percentage of training

samples is reduced. Reducing the training samples by 90% decreased the model's accuracy between 5.4% and 1.6%, considering all experiments. This reduction in accuracy is reasonable and indicates a good ability to downscale the model to smaller datasets. This feature is particularly interesting for model recalibration, as obtaining ground truth data can be expensive. These results confirm the good performance of RF with small amounts of training data, which can outperform deep learning-based algorithms (Portalés-Julià et al. 2021).
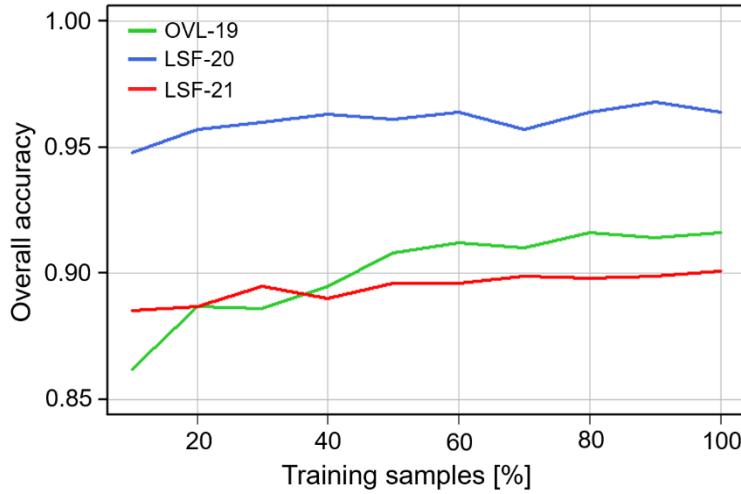
Figure 6.9. Effect of training dataset size on model performance.

6.3.4. Spatial and temporal transferability

To assess spatial transferability, the model trained with the LSF-21 dataset was used to classify citrus crop status in parcels of the four separate areas (PV-21, TV-21, BP-21, and NL-21) without using additional training. Table 6.3 shows the classification results. OA ranged from 0.87 to 0.92. Highest accuracies were obtained in the NL-21 (0.92 +/- 0.06) and PV-21 (0.92 +/- 0.06) datasets, while TV-21 and BP-21 had the lowest (0.88 +/- 0.07 and 0.87 +/- 0.07 respectively). Average OA outside the study area was 0.898, which is only 1% below the classification within the training area. These results show a good transferability of the model within the studied territory. However, the recall of the AB class was highly compromised when transferring the model out of the study area (0.700 and 0.579 for BP-21 and TV-2,1 respectively). This indicates that the model is not able to recall a large part of the abandoned parcels in these areas, suggesting some inability of the model to detect new abandonment patterns that are not well represented in the training dataset. We attribute these errors to the photointerpretation process, since arable lands are frequent in this area, they may have been erroneously labelled as NP.

Dataset	Overall accuracy (95% C.I.)	Cohen's Kappa	NP precision	PR precision	AB precision	NP recall	NP recall	NP recall
PV-21	0.924 +/- 7.4%	0.879	0.933	0.900	0.958	0.875	1.000	0.852
TV-21	0.875 +/- 5.9%	0.742	1.000	0.862	0.846	0.900	0.980	0.579
BP-21	0.870 +/- 5.8%	0.773	0.733	0.913	0.875	0.846	0.955	0.700
NL-21	0.922 +/- 7.2%	0.858	1.000	0.968	0.700	0.833	0.968	0.875
Average	0.899	0.813	0.916	0.910	0.845	0.863	0.975	0.752

Temporal transferability of the trained models was cross-evaluated on models trained with the OLV-19, LSF-20, and LSF-21. The performance of each model was tested with data from years other than the training year. Table 6.4 shows the results of time transferability.

Table 6.4. Temporal transferability. Cross-evaluation of models trained with the OLV-19, LSF20 and LSF-21 datasets.

Training dataset (year)	Tested with year:	Overall accuracy (95% C.I.)	Cohen's Kappa	NP precision	PR precision	AB precision	NP recall	PR recall	AB recall
OLV-19 (2019)	2020	0.889 +/- 3.7%	0.829	0.983	0.979	0.776	0.983	0.775	0.970
	2021	0.851 +/- 2.1%	0.769	0.922	0.914	0.751	0.819	0.838	0.888
LSF-20 (2020)	2019	0.883 +/- 4.0%	0.825	0.962	0.772	0.966	0.950	0.975	0.725
	2021	0.883 +/- 2.2%	0.826	0.950	0.975	0.725	0.981	0.856	0.987
LSF-21 (2021)	2019	0.850 +/- 4.5%	0.775	0.917	0.733	0.980	0.926	0.962	0.625
	2020	0.946 +/- 2.6%	0.916	0.936	0.923	0.988	0.983	1.000	0.860

In the different cross-experiments OA ranged from 0.85 to 0.95, with an average of 0.88. This represents about 3.5% reduction in performance when using the model out of the training year, which can be considered as acceptable. However, we should highlight that a reduction of the precision and recall of the PR and AB classes, leading to more confusion between the two categories can be observed. Figure 6.10 shows a summary of the OA obtained by classifying each dataset. It compares the performance obtained using models trained in years other than the prediction year versus the cross-validation performance in the same year.
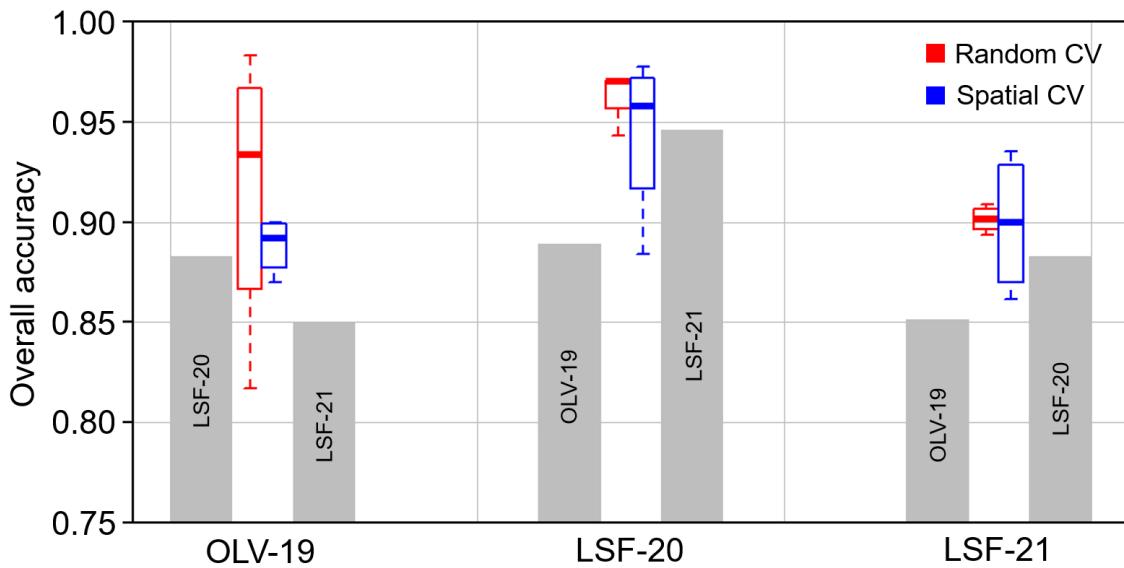
Figure 6.10. OA obtained in each dataset as a function of the model used. The gray bars show the OA obtained using models trained on data sets from other years and the boxplots show the accuracy obtained by cross-validation.

6.3.5. Feature importance on class separability

Figure 6.11 shows 4 radar charts of JM distances, which are related to feature separability assessment. The upper part of the figure displays calculated pairwise JM distances of the 3 classes generated by the features previously categorized as central tendency statistics (12 average monthly values of OSAVI and NDMI). The lower part refers to JM values of the 6 features previously categorized as dispersion-related (range October-March, range April-September, standard deviation October-March, standard deviation April-September, annual kurtosis annual skewness of OSAVI an NDMI) for the 3 classes of abandonment status.
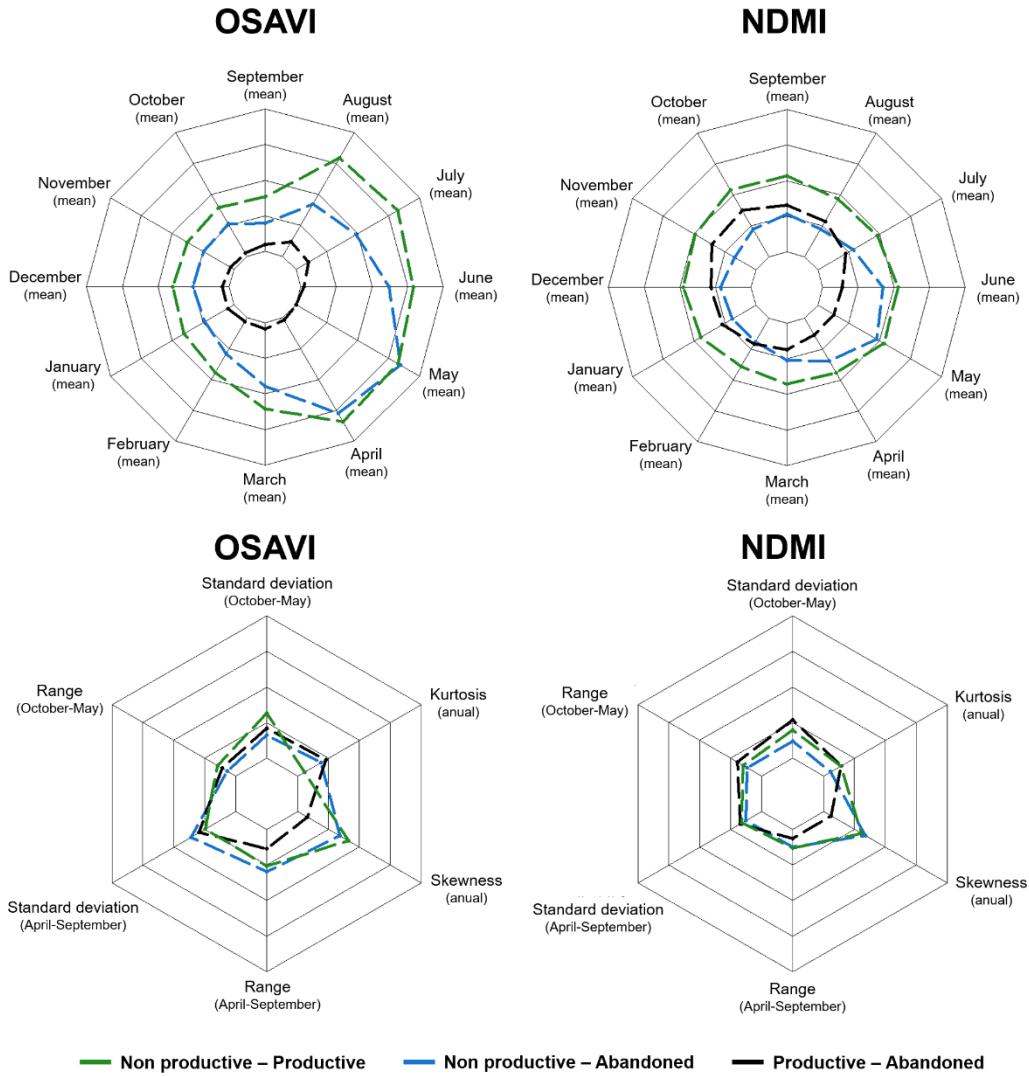
Figure 6.11. Jeffries-Matusita separability for each variable studied and each pair of categories.

All JM values were in the range 0.1 – 1.6. Despite this apparently low separability, the RF classifier produced high classification accuracy, which reveals both the importance of using a multivariate approach and the potential of RF classifiers. Monthly averages of both OSAVI and NDMI showed the highest JM values. In fact, removing the rest of variables resulted only in a ~2% decrease in model performance (data not shown).

Comparison NP-PR produced the highest JM distances for both OSAVI and NDMI features, and PR-AB the lowest. OSAVI derived features induced greater NP-PR and NP-AB separability. However, NDMI derived features produced greater PR-AB separability, being these the most difficult to separate, as observed in Figure 6.11.

OSAVI derived features showed higher separability for NP-PR and NP-AB classes in April - August when spontaneous vegetation grows. PR-AB presented higher separability in July-August, which are the warmest and driest months, when spontaneous vegetation wilts. This is the typical lifecycle of *Oxallis pes-caprae*, one of the most extended weeds in the study area. The pattern particularly affects NP orchards, where tree canopies occupy a very small proportion of land. In any case, JM values remained lower than 0.5, which are commonly considered as very low.

112

Regarding NDMI derived features, separation of NP-PR classes showed a relatively stable pattern throughout the year, with JM values around 1. Separability of NP-AB classes increased in May-June and decreased in December-October. PR-AB separability was higher in September-November. This may be attributed to differences in humidity due to the absence of irrigation in the abandoned orchards. However, JM values of PR-AB remained low (less than 0.7).

6.4. DISCUSSION

This work studied the use of Sentinel-2 time series to identify the status of citrus crops in a highly fragmented agricultural land. The work develops a classification framework for processing OSAVI and NDMI time series extracted from Sentinel-2 L2A data to identify three types of parcels (non-productive, productive, and abandoned). The proposed approach involves a supervised classification using the Random Forest algorithm and a strategy for classification at parcel level. Results on three different temporal and spatial datasets data provided OA ranging from 89 to 92%, which represents significant improvements of about 13% over previous classification methods based on single date (Morell-Monzó et al., 2020). Overall, the work demonstrates the potential of Sentinel-2 dense time series for monitoring abandonment of citrus parcels and highlights the importance of the time domain information to separate parcels with similar spectral responses.

Several research studies highlight the importance of multitemporal data for monitoring seasonal crops (Vuolo et al., 2018; Roumenina et al., 2015; Palchowdhuri et al., 2018), reporting on increasing accuracy when including additional multi-temporal information to RF models. Our work confirms that this importance also extends to permanent, perennial crops. To our knowledge there are few studies on the identification of permanent crop abandonment. Although a direct comparison of accuracy estimates is not possible, our results support those obtained in detecting abandonment in other perennial crops, like olives (Volpi et al., 2022) or in forest perennial tree species identification (Hemmerling et al., 2021).

Despite the low spectral separability between the studied parcel types, Random Forest algorithm allowed to obtain the following average class accuracy: NP = 0.935 > PR = 0.921 > AB = 0.899. However, the results also showed limitations in recall of the AB parcels. This fact suggests that some types of abandonment are unidentifiable through our approach. Statistical analysis could not confirm significant differences between the size of correctly classified and misclassified parcels. However, Vajsová, et. al., (2020) pointed out to other characteristics of the parcels not linked to size but shape (for instance, the ratio between the number of full pixels and the number of incomplete pixels within the parcel). More research is needed to associate the characteristics of the parcels and the types of abandonment that can be identified due to the particular resolution of Sentinel-2 images.

The model showed acceptable spatial and temporal transferability. When transferred the model outside the training area, an average OA of 0.89 was obtained, which represents around a 1% reduction. However, transferring the model to other areas notably reduced the capacity to recall the AB parcels. Omission errors for the AB class raised to 42.1% and 30.0% in some cases. These results highlight the importance of having a database with a well-represented AB class. Abandonment is a process, not a

status, that depends on the time elapsed since its beginning, management practices and environmental conditions that affect soil and vegetation. These conditions affect the crop remaining in the orchard and the vegetative patterns of spontaneous species. This may cause that intra-class variability is not well represented in the training dataset. Volpi et al. 2023 studied the abandonment of olive groves in Tuscany (Italy) and obtained a poor recall of AB class. When transferring the model out of the training year, an average OA of 0.88 was achieved, which represents a 3.5% reduction. A decrease in recall metric of the AB class was observed. When the model was transferred out of the training year the confusion between AB and PR classes increases. The phenological profile of the AB parcels, which do not receive irrigation, is more sensitive to the climatic regime of each particular year. The high variability of rainfall between years makes it difficult to obtain a good recall of the AB class when a model trained with data from a single year is transferred to other years. Such performance reductions may be attributed to a lack of correlation between the index values along the different years, which depend on agrometeorological conditions like soil fertility, temperature, and precipitation.

Overall, this study provides promising results for monitoring the abandonment of citrus orchards in the Comunitat Valenciana region. Accuracies obtained allow the possibility of identifying general trends of land abandonment at medium to large scales, which may help to support policies for adequate landscape management.

Further research is required to improve accuracy and the recall of AB class. We suggest adding texture features capable to detect planting patterns (also called planting frames). However, this would require higher spatial resolution images, as we showed in past work (Morell-Monzó et al., 2021). Texture information can be extracted from Very High-Resolution (VHR) airborne and satellite images. But airborne images may show limitations to transfer the model spatially and temporally due to differences in acquisition conditions (sensor, date, sun angle, atmospheric conditions, etc.). Image fusion techniques can optimally merge information from image sources of different time and spatial resolutions (Moltó, 2022). Future research should focus on developing methods to combine Sentinel-2 time series and VHR images in a transferable way. Furthermore, it is possible to add new information from synthetic aperture radar (i.e. form Sentinel-1 mission). Moreover, using longer time series or algorithms enable to model sequential information (e.g. Long Short-Therm Memory or 1D Convolutional Neural Networks) could also improve model performance.

6.5. CONCLUSIONS

This work studied the use of Sentinel-2 time series to identify the status of citrus crops in a highly fragmented agricultural landscape in the Comunitat Valenciana region (Spain). The work developed a classification framework to identify three types of parcels (non-productive, productive, and abandoned). Results show the potential of Setinel-2 time series to identify citrus crop status. Results on three different temporal and spatial datasets provided overall accuracies ranging from 0.89 to 0.92, outperforming the classification based on a single date around 13%. The model showed acceptable spatial and temporal transferability. However, transferring the model to other areas notably reduced the capacity to recall the abandoned parcels. When transferring the model out of the training year there was more confusion between the productive and

abandoned parcels. Overall, this study provides promising results for monitoring the abandonment of citrus orchards in the Comunitat Valenciana region. More research is needed to improve identification, generate more transferable models, and understand the abandonment patterns or parcel types that cannot be identified.

6.6. REFERENCES

Asgarian, A., Soffianian, A., Pourmanafi, S. Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal Landsat 8 imagery. Computers and Electronics in Agriculture, 2016, 127, 531–540. https://doi.org/10.1016/j.compag.2016.07.019

Bhattacharyya, A. On a measure of divergence between two statistical populations defined by their probability distributions. Bulletin of the Calcutta Mathematical Society 1943, 35, 99–109. Available on-line: https://archive.org/details/dli.calcutta.11603/page/98/mode/2up

Breiman, L. Random Forest. Machine Learning 2001, 45(1), 5–32. https://doi.org/10.1023/a:1010933404324

Chen, J., Jönsson, Per., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote Sensing of Environment 2004, 91(3–4), 332–344. https://doi.org/10.1016/j.rse.2004.03.014

Czesak, B., Różycka-Czas, R., Salata, T., Dixon-Gough, R., Hernik, J. Determining the Intangible: Detecting Land Abandonment at Local Scale. Remote Sensing 2021, 13(6), 1166. https://doi.org/10.3390/rs13061166

Dimov, D. Classification of remote sensing time series and similarity metrics for crop type verification. Journal of Applied Remote Sensing 2022, 16(02). https://doi.org/10.1117/1.jrs.16.024519

Gao, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 1996, 58 3, 257–266. https://doi.org/10.1016/s0034-4257(96)00067-3

Gasith, A., Resh, V.H. Streams in the Mediterranean climate regions: abiotic influences and biotic responses to predictable seasonal events. Annual Review of Ecology and Systematics 1999, 30, 51-81. http://www.jstor.org/stable/221679.

Geleralitat Valenciana. Portal Estadístico de la Generalitat Valenciana. Fichas Municipales 2020. Available online: http://www.pegv.gva.es/es/fichas (accessed on 15 November 2022).

Gómez, C., White, J. C., Wulder, M. A. Optical remotely sensed time series data for land cover classification: A review. ISPRS Journal of Photogrammetry and Remote Sensing 2016, 116, 55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

González-Hidalgo, J.C., de Luís Arrillaga, M., Peña Monné, J.L. Extreme rainfall events, climate variability and soil erosion. Some comments related to climate change in Mediterranean environments. Cuaternario y Geomorfologia. 2005, 19(1–2), 49-62.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment 2017, 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Hemmerling, J., Pflugmacher, D., Hostert, P. Mapping temperate forest tree species using dense Sentinel-2 time series. Remote Sensing of Environment 2021, 267, 112743. https://doi.org/10.1016/j.rse.2021.112743

Huete, A. R. A soil-adjusted vegetation index (SAVI). In Remote Sensing of Environment 1988, 25, 3, 295–309. https://doi.org/10.1016/0034-4257(88)90106-x

Nisini, L., Anguiano, E., Terres, J., Assessing the risk of farmland abandonment in the EU, Publications Office, 2013, https://data.europa.eu/doi/10.2788/81337

Karasiak, N., Dejoux, J.-F., Monteil, C., Sheeren, D. Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. Machine Learning 2021, 111, 7, 2715–2740. https://doi.org/10.1007/s10994-021-05972-1

Leal Filho, W., Mandel, M., Al-Amin, A. Q., Feher, A., Chiappetta Jabbour, C. J. An assessment of the causes and consequences of agricultural land abandonment in Europe. International Journal of Sustainable Development & World Ecology 2016, 24, 6, 554–560. https://doi.org/10.1080/13504509.2016.1240113

Liaw, A., Wiener, M. Classification and Regression by randomForest. R News 2002, 2(3), 18--22. https://cran.r-project.org/web/packages/randomForest/

Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2020; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2021. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2021; Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente: Madrid, Spain, 2022. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/

Moltó, E. Fusion of Different Image Sources for Improved Monitoring of Agricultural Plots. Sensors 2022, 22, 17, 6642. https://doi.org/10.3390/s22176642

Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.T. 2022. Clasificación del estado de parcelas de cítricos utilizando datosmultitemporales Sentinel-2. Teledetección para una Agricultura Sostenible en la era del Big Data, Actas del XIX Congreso de la Asociación Española de Teledetección 2022, 35-38. http://www.aet.org.es/?q=congresos

Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.-T. Comparison of Sentinel-2 and High-Resolution Imagery for Mapping Land Abandonment in Fragmented Areas. Remote Sensing 2020, 12(12), 2062. https://doi.org/10.3390/rs12122062

Morell-Monzó, S., Sebastiá-Frasquet, M.-T., Estornell, J. Land Use Classification of VHR Images for Mapping Small-Sized Abandoned Citrus Plots by Using Spectral and Textural Information. Remote Sensing 2021, 13, 4, 681. https://doi.org/10.3390/rs13040681

Palchowdhuri, Y., Valcarce-Diñeiro, R., King, P., Sanabria-Soto, M. Classification of multi-temporal spectral indices for crop type mapping: a case study in Coalville, UK. The Journal of Agricultural Science 2018, 156(1), 24–36. https://doi.org/10.1017/s0021859617000879

Perpiña Castillo C., Kavalov B., Diogo V., Jacobs-Crisioni C., Batista e Silva F., Lavalle C. Agricultural Land Abandonment in the EU within 2015-2030. JRC113718, European Commission 2018.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V., Murayama, Y., Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. Remote Sensing 2020, 12, 14, 2291. https://doi.org/10.3390/rs12142291

Portalés-Julià, E., Campos-Taberner, M., García-Haro, F. J., Gilabert, M. A. Assessing the Sentinel-2 Capabilities to Identify Abandoned Crops Using Deep Learning. In Agronomy 2021, 11(4), 654. https://doi.org/10.3390/agronomy11040654

Richards, J. A. and Jia, X., 2006. Remote Sensing Digital Image Analysis: An Introduction. Berlin, Germany: Springer.

Rondeaux, G., Steven, M., Baret, F. Optimization of soil-adjusted vegetation indices. Remote Sensing of Environment 1996, 55(2), 95–107. https://doi.org/10.1016/0034-4257(95)00186-7

Roumenina, E., Atzberger, C., Vassilev, V., Dimitrov, P., Kamenova, I., Banov, M., Filchev, L., Jelev, G. Single- and Multi-Date Crop Identification Using PROBA-V 100 and 300 m S1 Products on Zlatia Test Site, Bulgaria. Remote Sensing 2015, 7, 10, 13843–13862. https://doi.org/10.3390/rs71013843

Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry, 1964, 36, 1627-1639.

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., Homayouni, S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2020, 13, 6308–6325. https://doi.org/10.1109/jstars.2020.3026724

Stock, A., Subramaniam, A. Iterative spatial leave-one-out cross-validation and gap-filling based data augmentation for supervised learning applications in marine remote sensing. In GIScience & Remote Sensing 2022, 59(1), 1281–1300. https://doi.org/10.1080/15481603.2022.2107113

Subedi, Y. R., Kristiansen, P., Cacho, O. Drivers and consequences of agricultural land abandonment and its reutilisation pathways: A systematic review. Environmental Development 20222, 42, 100681. https://doi.org/10.1016/j.envdev.2021.100681

Terres, J.-M., Scacchiafichi, L. N., Wania, A., Ambar, M., Anguiano, E., Buckwell, A., Coppola, A., Gocht, A., Källström, H. N., Pointereau, P., Strijker, D., Visek, L., Vranken, L., Zobena, A. Farmland abandonment in Europe: Identification of drivers and indicators, and development of a composite indicator of risk. Land Use Policy 2015, 49, 20–34. https://doi.org/10.1016/j.landusepol.2015.06.009

Vajsová, B., Fasbender, D., Wirnhardt, C., Lemajic, S., Devos, W. Assessing Spatial Limits of Sentinel-2 Data on Arable Crops in the Context of Checks by Monitoring. Remote Sensing 2020, 12(14), 2195. https://doi.org/10.3390/rs12142195

Volpi, I., Marchi, S., Petacchi, R., Hoxha, K., Guidotti, D. Detecting olive grove abandonment with Sentinel-2 and machine learning: The development of a web-based tool for land management. Smart Agricultural Technology 2023, 3, 100068. https://doi.org/10.1016/j.atech.2022.100068

Vuolo, F., Neuwirth, M., Immitzer, M., Atzberger, C., Ng, W.-T. How much does multi-temporal Sentinel-2 data improve crop type classification? In International

Journal of Applied Earth Observation and Geoinformation 2018, 72, 122–130. https://doi.org/10.1016/j.jag.2018.06.007

Wadoux, A. M. J.-C., Heuvelink, G. B. M., de Bruin, S., Brus, D. J. Spatial cross-validation is not the right way to evaluate map accuracy. Ecological Modelling 2021, 457, 109692. https://doi.org/10.1016/j.ecolmodel.2021.109692

CAPÍTULO 7

Discusión de resultados

7.1. IDENTIFICACIÓN DEL ESTADO DE LOS CULTIVOS DE CÍTRICOS UTILIZANDO DIFERENTES TÉCNICAS DE TELEDETECCIÓN

Esta tesis ha abordado el problema de la identificación del estado de los cultivos de cítricos tratando de dar respuesta al seguimiento del abandono de tierras agrícolas en la Comunitat Valenciana. Se han analizado diferentes casos de estudio en la comarca de La Safor y a modo complementario en otras zonas de la Comunitat Valenciana. La tesis ha comparado diferentes metodologías y fuentes de datos de observación de la tierra como Sentinel-2, World-View-3, imágenes aéreas de muy alta resolución y datos altimétricos derivados de LiDAR y fotogrametría SfM. Se han desarrollado diferentes enfoques de clasificación para identificar tres tipos de parcelas citrícolas: no productivas, productivas y abandonadas.

El primer caso de estudio (Capítulo 2) se llevó a cabo en el municipio de Oliva (Valencia). En este caso se compararon imágenes Sentinel-2 e imágenes aéreas de muy alta resolución, utilizando el algoritmo Random Forest como método de clasificación. La comparación mostró una mayor capacidad de las imágenes aéreas para identificar el estado de los cultivos a partir de una única imagen. La exactitud global para una única imagen Sentinel-2 fue de aproximadamente el 77%, mientras que para una única imagen aérea fue del 88%. Los valores de exactitud obtenidos a través de una única imagen Sentinel-2 fueron más bajos que en otros estudios donde se utilizó Random Forest para tipos de cultivos con más categorías (Immitzer et al. 2016; El Hachimi et al., 2021), lo que pone de manifiesto la dificultad para identificar las parcelas abandonadas. En nuestro enfoque utilizamos la base de datos catastral de parcelas agrícolas para realizar una clasificación a nivel de parcela, aplicando la votación mayoritaria a partir de la clasificación basada en píxeles. Esta estrategia fue utilizada previamente por otros autores de manera exitosa (Ghorbanian et al., 2020) y permitió mejorar la clasificación debido a la reducción del ruido sal y pimienta (Wulder et al., 2018). Las mejoras producidas por el refinamiento basado en la votación mayoritaria fueron mayores en las imágenes de muy alta resolución debido a la relación entre el tamaño del píxel y el tamaño del objeto a clasificar. Las ventajas de la alta resolución permiten desplegar el alto potencial del análisis basado en objetos (OBIA) (Blaschke, 2010). Por otra parte, añadir las bandas 5, 6, 7, 11 y 12 de Sentinel-2, correspondientes al borde rojo e infrarrojo de onda corta (705 nm – 2190 nm) no mejoró la detección a partir de una única imagen. Atribuimos este resultado a la menor resolución (20 m) de estas bandas y al pequeño tamaño de las parcelas estudiadas. La mayor confusión se produjo entre las parcelas productivas y abandonadas debido a la similitud de su firma espectral. Utilizando las imágenes aéreas se obtuvo una precisión del 80% con un *recall* del 87,5% para la categoría productivo y una precisión del 92,6% con un *recall* del 78,1% para la categoría abandonado.

Estos resultados nos llevaron a explotar el uso de las imágenes aéreas de muy alta resolución con mayor profundidad. En el segundo caso de estudio (Capítulo 3) desarrollamos un marco de clasificación que combina información espectral e información espacial en forma de características de textura, adoptando un enfoque similar al de Feng et al. (2015). Utilizamos la Matriz de Co-Ocurrencia de Nivel de Gris para extraer siete características de textura a partir del Índice de Vegetación de Diferencia Normalizada (NDVI): promedio, varianza, contraste, disimilitud, homogeneidad, segundo momento angular y entropía. Se utilizó un enfoque de ventana

móvil para extraer un valor de textura para cada píxel de la escena a partir de un vecindario definido por el usuario. A mayor tamaño de ventana se capturan mejor los patrones espaciales del cultivo y se reduce el ruido sal y pimienta en la clasificación. Sin embargo, aumentar demasiado el tamaño de ventana puede producir efectos de borde en las zonas de transición entre dos categorías (Ruiz et al., 2001), es decir, en los límites de las parcelas. Para nuestro caso de estudio el tamaño de ventana óptimo para la extracción de características de textura fue 9x9, teniendo en cuenta que la resolución de la imagen utilizada es de 1 m y el tamaño promedio de las parcelas está entre 0,25 y 0,50 ha. Este enfoque permitió elevar la exactitud global de clasificación basada en píxeles hasta un 87% debido a la reducción del efecto sal y pimienta. Después de aplicar la votación mayoritaria se obtuvo una exactitud global de clasificación en torno al 95% a nivel de parcela. Estos resultados suponen una mejora en torno al 7% con respecto al método de clasificación basado únicamente en información espectral. En este experimento mejoramos el método de validación utilizando una validación cruzada en lugar de un método *hold-out*. Se eligió la validación cruzada ya que los métodos de *hold-out* pueden dar lugar a grandes variaciones en las estimaciones de precisión (Lyons et al., 2018). Utilizamos el método Boruta para identificar la importancia de las variables utilizadas. Este es un método de envoltura en torno al clasificador Random Forest que permite identificar las variables con una importancia estadísticamente significativa. Los resultados mostraron que todas las variables utilizadas por el modelo son significativamente importantes. La importancia de las características de textura radica en su capacidad para modelar los patrones espaciales típicos de los cultivos frutales, también llamados marcos de plantación. Los patrones espaciales son un rasgo identificativo de los cultivos de cítricos que se pierde al abandonarse una parcela. La resolución espacial de 10 m de Sentinel-2 es insuficiente para capturar los marcos de plantación de los cítricos (Moltó, 2022). Sin embargo, incluir características de textura derivadas de imágenes con resolución menor a 1 m permitió mejorar significativamente la clasificación (de un 71% a un 87% a nivel de píxel). La adición de características de textura repercutió mayoritariamente en una mejora en la separación de las categorías productivo y abandonado. A nivel de parcela se obtuvo una precisión del 96% con un *recall* del 91% para la categoría productivo y una precisión del 94% con un *recall* de 96% para la categoría abandonado. Nuestro estudio utilizó la Matriz de Co-Ocurrencia de Nivel de Gris descrito en Haralick et al. (1973) que es un método ampliamente implementado en muchos paquetes de software. Sin embargo, existen otros métodos para la extracción de características de textura que podrían explorarse como: patrones binarios locales (Wang y He, 1990), técnicas de filtrado como los filtros de energía (Laws, 1980), los filtros de intensidad de bordes (Sutton y Hall, 1972), o los bancos de filtros Gabor (Jain y Farrokhnia, 1991), métodos basados en la descomposición *wavelet* (Mallat, 1989) o métodos basados en funciones geoestadísticas (Balaguer et al., 2010). El enfoque propuesto fue testeado en otras zonas fuera del área de estudio y se obtuvo una exactitud global promedio mayor del 90%. Sin embargo, la mayor limitación de esta metodología está relacionada con la calidad radiométrica de las imágenes. A diferencia de las imágenes Sentinel-2, las imágenes aéreas ofrecidas por el Institut Catogràfic Valencià no disponen de productos de reflectancia de la superficie. Este hecho limita la transferibilidad del modelo fuera del área y año de entrenamiento debido a las diferentes condiciones de adquisición de las imágenes, como son: el tipo de sensor utilizado la fecha de adquisición, diferentes condiciones atmosféricas, diferente ángulo

de incidencia solar, altura del vuelo, etc. (Feng et al., 2015). Sin embargo, la alta exactitud obtenida permitiría utilizar estas imágenes para entrenar diversos modelos que fueran capaces de generar inferencias regionalmente (a escala local o comarcal) y así, cubrir gran parte del territorio valenciano. Hermosilla et al. (2022) propusieron una estrategia para entrenar e implementar diferentes modelos de clasificación a nivel regional con el objetivo de cartografiar grandes áreas. El enfoque propuesto se podría extender a regiones que no disponen de imágenes aéreas anuales utilizando imágenes satelitales de muy alta resolución de constelaciones comerciales. Actualmente existen constelaciones con resolución submétrica como WorldView, SkySat, BlackSky, Pleaides Neo, NewSat, Vision-1 o Jilin-1 (ESA, 2022), entre otras.

El siguiente caso de estudio (Capítulo 4) se llevó a cabo en la franja litoral de comarca de La Safor (Valencia), entre los municipios de Gandía y Oliva. Se realizó un muestreo en 12 pequeñas áreas sobrevoladas por un vehículo aéreo no tripulado (UAV, por sus siglas en inglés). En este experimento evaluamos el potencial de los datos altimétricos para identificar el estado de los cultivos de cítricos. Se evaluó el rendimiento de las nubes de puntos LiDAR del PNOA 2015 (2ª cobertura) con una densidad de puntos 0,8 puntos/m^2 y nubes de puntos altimétricas extraídas a través de procesos fotogramétricos SfM con una densidad de puntos > 50 puntos/m^2. Las nubes de puntos LiDAR permitieron clasificar el estado de los cultivos con una exactitud global baja, en torno al 67,9%. Sin embargo, las nubes de puntos SfM permitieron obtener una exactitud global en torno al 83,6%, superando incluso a una única imagen Sentinel-2 (Morell-Monzó et al., 2020). Esto se debe a la alta densidad de puntos de la nube de puntos SfM que permitió extraer características de textura a partir del cálculo del Canopy Height Model (CHM). Estas texturas extraídas a partir de la información altimétrica de la parcela permitieron capturar los patrones de plantación de los cultivos. A pesar de que las nubes de puntos derivadas de fotogrametría no tienen penetración sobre la vegetación, los espacios entre árboles y caminos permitieron obtener un CHM consistente para extraer texturas. No obstante, la subestimación del CHM debido a una sobreelevación del MDT puede ser un factor limitante en zonas con alta pendiente o zonas con vegetación muy densa (Tan et al., 2018). Sin embargo, los cítricos, que son cultivos de regadío, requieren poca pendiente (habitualmente < 3%) y a menudo los agricultores eliminan la vegetación de los campos para facilitar las prácticas agrícolas. Estas características junto con el espacio entre las filas de árboles facilitan la generación de un MDT coherente. Los resultados obtenidos en nuestro trabajo a través de los datos LiDAR no alcanzaron la exactitud obtenida por Kolecka et al. (2015) y Czesak et al. (2021) que utilizaron datos LiDAR de mayor densidad (4-12 puntos/m^2). En estos trabajos la exactitud global para la detección del abandono de cultivos estacionales fue del 82% (Czesak et al., 2021) mientras que para la monitorización de sucesiones forestales secundarias fue del 95% (Kolecka et al., 2015). Estos mejores resultados pueden explicarse por la mayor densidad de puntos y el tipo de vegetación. Además, en estos casos, las variaciones altimétricas en los cultivos estacionales abandonados que dieron paso a sucesiones forestales fueron mayores que las encontradas en el abandono de cítricos. Será de interés analizar las próximas coberturas LiDAR con mayor densidad de puntos lo que permitirá la extracción de características de textura y su aplicación en la detección del abandono de cítricos.

En el siguiente caso de estudio (Capítulo 5) extendimos este enfoque de clasificación a imágenes del satélite WorldView-3 y analizamos las mejoras producidas por la fusión

de datos multiespectrales y altimétricos. En este conjunto de datos, las imágenes WorldView-3 permitieron obtener una exactitud global de 95,4%. Otros autores (McCarty et al., 2017; Zurita-Milla et al., 2017; Neigh et al., 2018) también reportaron las mejoras obtenidas al usar imágenes satelitales de muy alta resolución con respecto a las imágenes de resolución moderada para la monitorización de parcelas de reducido tamaño. La información altimétrica por sí misma mostró un menor potencial que la información espectral. Combinar información espectral y altimétrica produjo mejoras entre el 0,4-2%. No obstante, debido a los resultados obtenidos en términos de precisión y el alto coste de obtención y procesamiento de las nubes de puntos SfM, no recomendamos esta técnica para la identificación del estado de los cultivos de cítricos. Sin embargo, en base al análisis de separabilidad Jeffries-Matusita, las características de textura extraídas a partir del CHM mostraron un cierto potencial para identificar los patrones de plantación de los cítricos. Nuestros resultados, tanto en imágenes multiespectrales, como pancromáticas y de elevación, demuestran la importancia de las características de textura extraídas a partir de imágenes con resolución < 1 m (submétricas) para identificar los patrones de plantación típicos de los cultivos de cítricos, que a su vez ayudan a detectar los cultivos abandonados. En este sentido, la identificación de patrones espaciales presentes en diferentes paisajes antrópicos, como los patrones de plantación de cultivos y árboles en parcelas agrícolas, pueden ser particularmente útiles en diferentes problemas agrícolas como reportaron varios autores (Ruiz et al., 2007; Trías-Sanz et al., 2006, Helmholz et al., 2007, Recio, 2009). Las características de textura extraídas a partir de la banda pancromática de WorldView-3 produjeron una menor separabilidad entre las clases productivo y abandonado que las características de textura extraídas a partir del NDVI. La mayor separabilidad al utilizar el NDVI se debe a que produce mayores diferencias entre los niveles de gris del suelo y el dosel del árbol, lo que genera un mayor contraste. Por el contrario, la banda pancromática, con un mayor ancho de banda (450 – 800 nm), produce un menor contraste que repercute en la textura de la imagen. La menor separabilidad producida por las texturas extraídas a partir de la banda pancromática también puede atribuirse a la fecha de adquisición de la imagen. La imagen WorldView-3 fue adquirida en el mes de noviembre. En esta época del año las precipitaciones, la temperatura y la humedad del suelo promueven el crecimiento de *Oxalis pes-caprae* (comúnmente llamado agrios o *agret,* en valenciano). Esta vegetación también puede afectar al contraste entre el suelo y el dosel de árbol. Por otra parte, el análisis de separabilidad también muestra la importancia de las bandas en el infrarrojo de onda corta de WorldView-3 puesto que existe un máximo de separabilidad entre los 1570 nm – 2330 nm entre las parcelas productivas y abandonadas. Esta mayor separabilidad se produce porque estas bandas son sensibles al contenido de humedad del suelo y la vegetación. Las parcelas abandonadas no reciben riego y por tanto el aporte de agua a la parcela solamente depende del régimen de precipitaciones. En cambio, las parcelas productivas reciben riego periódicamente para evitar que el cultivo entre en una situación de estrés hídrico.

En el último caso de estudio (Capítulo 6) estudiamos las capacidades de los datos multitemporales Sentinel-2 para identificar el estado de los cultivos de cítricos. Para ello, ampliamos el conjunto de datos de referencia con datos de tres años consecutivos. Adicionalmente generamos cuatro conjuntos de datos de referencia fuera del área de estudio. Desarrollamos un marco de clasificación que utiliza series temporales del Índice de Vegetación Ajustado al Suelo Optimizado (OSAVI) y del Índice De Humedad

De Diferencia Normalizada (NDMI) derivado de imágenes Sentinel-2. También mejoramos el proceso de refinamiento de clasificación a nivel de parcela utilizando la máxima probabilidad promedio, en lugar de la votación mayoritaria. La naturaleza inherentemente dinámica de los cultivos y su firma espectral provoca que los enfoques de clasificación basados en una única fecha sean limitados. La alta resolución temporal de Sentinel-2 (tiempo de revisita de 5 días) permite monitorizar la evolución del cultivo y reconstruir series temporales continuas a pesar de eventos meteorológicos adversos prolongados (Vuolo et al., 2018). Diversos autores reportaron las mejoras obtenidas al utilizar datos multitemporales en clasificaciones de tipos de vegetación de fenología marcada (Gómez et al., 2016), como los cultivos y pastos estacionales (Vuolo et al., 2018; Valero et al., 2016). Sin embargo, los beneficios de los datos multitemporales para identificar el estado de los cultivos permanentes, como los cítricos, han sido poco estudiados. Nuestros resultados muestran las mejoras producidas por los datos multitemporales Sentinel-2. Obtuvimos exactitudes globales que variaron entre 89-92% para los tres años estudiados. Esto supone una mejora significativa de en torno al 13% con respecto al enfoque basado en una única fecha. Estos resultados nos permiten extender la importancia de los datos multitemporales también al seguimiento de otros cultivos permanentes de hoja perenne. La inspección de los perfiles espectrales-temporales de las parcelas no productivas, productivas y abandonadas permitió observar diferencias fenológicas entre las diferentes categorías. Las parcelas no productivas tienen valores de OSAVI más bajos durante todo el año. Se observa un comportamiento cíclico del OSAVI que corresponde al crecimiento de la vegetación en la época húmeda. Por el contrario, las parcelas productivas y abandonadas tienen niveles de OSAVI más altos debido a la mayor densidad vegetal. Estos dos tipos de parcelas maximizan su separabilidad en el OSAVI en verano, ya que la vegetación silvestre, predominante en las abandonadas, pierde vigor, mientras que las parcelas productivas mantienen los niveles de OSAVI más altos. Con respecto al NDMI, que está relacionado con el contenido de humedad del suelo y la vegetación, también se observa un comportamiento cíclico que concuerda con las estaciones seca y húmeda. El NDMI produce una mayor separabilidad entre las parcelas productivas y abandonadas, lo que concuerda con los resultados obtenidos en Morell-Monzó et al. (2023) donde las bandas del infrarrojo de onda corta también mostraron una mayor separabilidad entre ambas clases. Sin embargo, la resolución espacial de las bandas del infrarrojo de onda corta de Sentinel-2 (20 m) puede ser un factor limitante en un paisaje altamente fragmentado. No pudimos confirmar diferencias significativas entre el tamaño de la parcela y la exactitud de clasificación, sin embargo, las limitaciones espaciales de Sentinel-2 pueden estar relacionadas con otras características de la parcela, como la forma, que determinan el número de píxeles completos dentro de la misma, como muestra el estudio de Vajsová et al. 2020. Evaluamos la transferibilidad espacial y temporal de los modelos generados. Al transferir el modelo fuera del área de estudio obtuvimos una buena capacidad de generalización, sin embargo, en algunas zonas la capacidad de recuperación (*recall*) de las parcelas abandonadas se redujo notablemente. En algunas zonas los errores por omisión de esta categoría fueron superiores al 40%. Este hecho concuerda con la experiencia de Volpi et al. (2023) quienes estudiaron el abandono de cultivos de olivar en La Toscana (Italia) y también obtuvieron un *recall* bajo para la categoría abandonado. Estos resultados ponen de manifiesto la dificultad para detectar el abandono debido a su alta variabilidad y la importancia de obtener un conjunto de datos

de abandono bien representado. Transferir el modelo fuera del año de entrenamiento también supuso una mayor confusión entre las categorías productivo y abandonado. Atribuimos esta reducción de rendimiento a la falta de correlación entre las series temporales de las parcelas abandonadas entre los diferentes años. Las parcelas abandonadas que no reciben riego son más sensibles al régimen de precipitaciones y temperatura de cada año en particular, que puede tener una gran variabilidad de un año a otro. Pensamos que una mayor representatividad de las condiciones agroclimáticas de diferentes años podría mejorar la transferibilidad del modelo. Se requiere más investigación para tratar de conocer las características de tamaño y forma, así como las casuísticas de abandono que no pueden detectarse a través de Sentinel-2.

7.2. CONSIDERACIONES METODOLÓGICAS SOBRE LA CLASIFICACIÓN DEL ESTADO DE LOS CULTIVOS DE CÍTRICOS

Hasta donde sabemos esta es la primera tesis que estudia la identificación del abandono de tierras agrícolas en la Comunitat Valenciana utilizando técnicas de teledetección. Sin embargo, durante el desarrollo de este trabajo se publicaron interesantes estudios sobre la detección del abandono de tierras agrícolas en la Comunitat Valenciana (Ruiz et al., 2020 y Portalés-Julià et al., 2021). No obstante, ninguno de estos trabajos tuvo como objetivo realizar una clasificación del estado de los cultivos a nivel de parcela. El trabajo de Ruiz et al. (2020) se centró en la identificación del abandono de tierras agrícolas a nivel de parche utilizando Redes Neuronales Convolucionales (CNN), pero no incluyó cultivos de cítricos. Por otra parte, el trabajo de Portalés-Julià et al. (2021) se centró en la identificación del ATA a nivel de píxel utilizando una arquitectura Long-Short-Term Memory (LSTM) con datos Sentinel-2. Por el contrario, nuestro trabajo realiza una exploración y validación de diferentes datos de teledetección para detectar el estado de los cultivos de cítricos a nivel de parcela.

Nuestra investigación requirió definir una clasificación propia del estado de los cultivos de cítricos. El enfoque propuesto parte de la utilización de la base de datos de SIGPAC. Utilizamos una clasificación basada en tres tipos de parcelas: no productivas, productivas y abandonadas. Además, de las categorías productiva y no productiva, creamos una categoría más correspondiente a las parcelas no productivas, debido a sus grandes diferencias con respecto a las dos anteriores. Es necesario remarcar que las parcelas de la clase no productiva son aquellas donde predomina el suelo desnudo. Nuestra investigación se aseguró de que estas parcelas cumplieran esta condición en el momento de la adquisición de los datos. Sin embargo, en algunas ocasiones las parcelas no productivas pueden estar cubiertas de vegetación, lo que las hace indistinguibles de las parcelas abandonadas. La aparición de la vegetación depende de las prácticas de manejo agrícola, la época del año y las características ambientales de la parcela. Esta puede ser una limitación metodológica para los enfoques de clasificación basados en una única fecha.

La investigación utilizó un enfoque de clasificación basado en *machine learning*. Aunque existe una tendencia creciente al uso de algoritmos de aprendizaje profundo por parte de la comunidad científica, debido a la cantidad limitada de datos disponibles, elegimos un enfoque basado en algoritmos de *machine learning* tradicional, que habitualmente requieren una menor cantidad de datos. En esta tesis se utilizaron dos algoritmos ampliamente utilizados en clasificación de uso y cobertura del suelo:

Random Forest y *Support Vector Machines*. Estos algoritmos destacan por su buen desempeño con pequeñas cantidades de datos. Nuestros resultados publicados en Morell-Monzó et al. (2023) muestran un rendimiento similar de ambos clasificadores. Sin embargo, *Random Forest* presenta ventajas desde el punto de vista computacional y de escalabilidad con respecto a Support *Vector Machines*. *Support Vector Machines* tiene una complejidad de tiempo entre $O(dn^2)$ y $O(dn^3)$, mientras que *Random Forest* tiene una complejidad de tiempo de $O(n\log(n)dk)$. Por este motivo, se utilizó *Random Forest* en la mayoría de casos de estudio de esta tesis. No obstante, sería interesante hacer una comparación más exhaustiva de ambos clasificadores en términos de exactitud de clasificación, transferibilidad espacial y temporal, sobreajuste, sensibilidad al número de muestras de entrenamiento, tiempo de ejecución y escalabilidad, ya que estos son aspectos importantes para la implementación operativa de estas metodologías.

7.3. GENERALIZACIÓN DE RESULTADOS

Esta investigación presenta metodologías eficientes en tiempo y coste para el seguimiento del abandono de cítricos en la Comunitat Valenciana a través de técnicas de teledetección y aprendizaje automático. Estas metodologías proporcionan información cuantitativa y espacialmente explícita sobre el abandono de tierras que puede ser de gran ayuda a los responsables de políticas agrarias y permitir implementar medidas de gestión del paisaje y generar indicadores de la efectividad de estas políticas. La investigación fue desarrollada en la comarca de La Safor (Valencia). Adicionalmente se utilizaron otros conjuntos de datos en Nules (comarca de la Plana Baixa, Castellón) y Benicull y Polinyà del Xúquer (comarca de la Ribera Baixa, Valencia). Sin embargo, los resultados obtenidos pueden ser generalizados a muchas otras áreas del territorio valenciano con un alto predominio de cultivos citrícolas. Los resultados también pueden generalizarse a otras regiones productoras de cítricos, sin embargo, las prácticas de manejo, clima y condiciones ambientales pueden afectar a los resultados.

Los resultados obtenidos en esta investigación son prometedores. Los niveles de exactitud obtenidos a través de los datos Sentinel-2 son más adecuados para monitorizar dinámicas generales del abandono de tierras agrícolas en el territorio. Sin embargo, estos niveles de exactitud pueden ser insuficientes para otros propósitos como la supervisión del pago de subsidios a los agricultores adscritos a la PAC. Por otro lado, las exactitudes obtenidas a través de imágenes aéreas de alta resolución son más cercanas a las necesidades de la administración para este propósito. Sin embargo, utilizar estas imágenes a grandes escalas aún presenta retos en cuanto a la transferibilidad espacial y temporal de los modelos. La implementación operativa de controles por monitoreo para el seguimiento del abandono de tierras debería aprovechar los beneficios de los datos Sentinel-2 y las imágenes aéreas de alta resolución de manera combinada.

Finalmente, el conocimiento generado por esta investigación también es útil para el estudio del abandono de otros cultivos, especialmente los cultivos de hoja perenne. En esta tesis se han desarrollado procedimientos que combinan técnicas de teledetección, análisis de imagen, series temporales y aprendizaje automático que pueden utilizarse en otros contextos. Las experiencias obtenidas en esta investigación ayudan a comprender el comportamiento espacial, espectral y temporal de los cultivos abandonados y a definir estrategias para su identificación, en particular de los cultivos de cítricos.

7.4. REFERENCIAS

Balaguer, A., Ruiz, L. A., Hermosilla, T., & Recio, J. A. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. Computers & Geosciences, 2010, 36(2), 231–240. https://doi.org/10.1016/j.cageo.2009.05.003

Blaschke, T. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 2010, 65(1), 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004

Czesak, B., Różycka-Czas, R., Salata, T., Dixon-Gough, R., & Hernik, J. Determining the Intangible: Detecting Land Abandonment at Local Scale. Remote Sensing, 2021, 13(6), 1166. https://doi.org/10.3390/rs13061166

El Hachimi, J., El Harti, A., Ouzemou, J.-E., Lhissou, R., Chakouri, M., Jellouli, A. Assessment of the benefit of a single sentinel-2 satellite image to small crop parcels mapping. Geocarto International, 2021. https://doi.org/10.1080/10106049.2021.1974955

European Spatial Agency – ESA. VHR, HR and MR Optical Missions. https://earth.esa.int/eogateway/activities/edap/vhr-hr-mr-optical-missions (accessed on 26 december, 2022)

Feng, Q., Liu, J., Gong, J. UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. Remote Sensing, 2015, 7(1), 1074–1094. https://doi.org/10.3390/rs70101074

Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., Hasanlou, M. (2020). Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 167, 276–288. https://doi.org/10.1016/j.isprsjprs.2020.07.013

Gómez, C., White, J. C., Wulder, M. A. Optical remotely sensed time series data for land cover classification: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 116, 55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Haralick, R. M., Shanmugam, K., Dinstein, I. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics, 1973, 6, 610–621. https://doi.org/10.1109/tsmc.1973.4309314

Helmholz, P., Gerke, M., Heipke, C. (2007). Automatic discrimination of farmland types using Ikonos imagery. In Silla, U. et al. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2007, 36. https://ris.utwente.nl/ws/portalfiles/portal/176674603/81_pia07.pdf

Hermosilla, T., Wulder, M., White, J.C., Coops, N.C. Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes. Remote Sensing of Environment 2022, 268, 112780. https://doi.org/10.1016/j.rse.2021.112780

Immitzer, M., Vuolo, F., Atzberger, C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. Remote Sensing, 2016, 8(3), 166. https://doi.org/10.3390/rs8030166

Jain, A. K., Farrokhnia, F. Unsupervised texture segmentation using Gabor filters. Pattern Recognition, 1991, 24(12), 1167–1186. https://doi.org/10.1016/0031-3203(91)90143-s

Kolecka, N., Kozak, J., Kaim, D., Dobosz, M., Ginzler, C., Psomas, A. Mapping Secondary Forest Succession on Abandoned Agricultural Land with LiDAR Point Clouds and Terrestrial Photography. Remote Sensing, 2015, 7(7), 8300–8322. https://doi.org/10.3390/rs7070830

Laws, K.I. Texture Image Segmentation. Ph.D. Dissertation, University Southern California, Los Angeles, CA, USA, 1980. https://apps.dtic.mil/sti/citations/ADA083283

Lyons, M. B., Keith, D. A., Phinn, S. R., Mason, T. J., Elith, J. A comparison of resampling methods for remote sensing classification and accuracy assessment. Remote Sensing of Environment, 2018, 208, 145–153. https://doi.org/10.1016/j.rse.2018.02.026

Mallat, S. G. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(7), 674–693. https://doi.org/10.1109/34.192463

McCarty, J. L., Neigh, C. S. R., Carroll, M. L., Wooten, M. R. Extracting smallholder cropped area in Tigray, Ethiopia with wall-to-wall sub-meter WorldView and moderate resolution Landsat 8 imagery. Remote Sensing of Environment, 2017, 202, 142–151. https://doi.org/10.1016/j.rse.2017.06.040

Moltó, E. Fusion of Different Image Sources for Improved Monitoring of Agricultural Plots. Sensors, 2022, 22(17), 6642. https://doi.org/10.3390/s22176642

Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.-T. Assessing the capabilities of high-resolution spectral, altimetric, and textural descriptors for mapping the status of citrus parcels. In Computers and Electronics in Agriculture, 2023, 204, 107504. https://doi.org/10.1016/j.compag.2022.107504

Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.-T. Comparison of Sentinel-2 and High-Resolution Imagery for Mapping Land Abandonment in Fragmented Areas. Remote Sensing, 2020, 12(12), 2062. https://doi.org/10.3390/rs12122062

Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.T. (2022). Clasificación del estado de parcelas de cítricos utilizando datos multitemporales Sentinel-2. En Teledetección para una Agricultura Sostenible en la era del Big Data, Actas del XIX Congreso de la Asociación Española de Teledetección, 2022, 35-38, Pamplona. http://www.aet.org.es/?q=congresos

Neigh, C. S. R., Carroll, M. L., Wooten, M. R., McCarty, J. L., Powell, B. F., Husak, G. J., Enenkel, M., Hain, C. R. Smallholder crop area mapped with wall-to-wall WorldView sub-meter panchromatic image texture: A test case for Tigray, Ethiopia. Remote Sensing of Environment, 2018, 212, 8–20. https://doi.org/10.1016/j.rse.2018.04.025

Portalés-Julià, E., Campos-Taberner, M., García-Haro, F. J., Gilabert, M. A. Assessing the Sentinel-2 Capabilities to Identify Abandoned Crops Using Deep Learning. Agronomy, 2021, 11(4), 654. https://doi.org/10.3390/agronomy11040654

Recio, J. Técnicas de extracción de características y clasificación de imágenes orientada a objetos aplicadas a la actualización de bases de datos de ocupación del suelo. Ph.D Dissertation Universitat Politècnica de València, València, Spain, 2009. https://doi.org/10.4995/Thesis/10251/6848

Ruiz, L. A., Almonacid-Caballer, J., Crespo-Peremarch, P., Recio, J. A., Pardo-Pascual, J. E., Sánchez-García, E. (2020). Automated classification of crop types and condition in a Mediterranean área using fine-tuned convolutional neural network. In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences: XLIII-B3-2020, 1061–1068. https://doi.org/10.5194/isprs-archives-xliii-b3-2020-1061-2020

Ruiz, L.A., Fernandez-Sarría, A., Recio, J. Evaluation of textura análisis techniques to characterize vegetation. https://cgat.webs.upv.es/BigFiles/texture_vegetation.pdf

Ruiz, L.A., Recio, J.A., Hermosilla, T. (2007). Methods for automatic extraction of regularity patterns and its application to object-oriented image classification. In Silla, U. et al. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2007, 36. http://www.pf.bgu.tum.de/isprs/pia07/puba/PIA07_Ruiz_et_al.pdf

Sutton, R. N., Hall, E. L. Texture Measures for Automatic Classification of Pulmonary Disease. IEEE Transactions on Computers, 1972, C–21 7, 667–676. https://doi.org/10.1109/t-c.1972.223572

Tan, Y., Wang, S., Xu, B., Zhang, J. An improved progressive morphological filter for UAV-based photogrammetric point clouds in river bank monitoring. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 146, 421–429. https://doi.org/10.1016/j.isprsjprs.2018.10.013

Trias-Sanz, R. Texture Orientation and Period Estimator for Discriminating Between Forests, Orchards, Vineyards, and Tilled Fields. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(10), 2755–2760. https://doi.org/10.1109/tgrs.2006.875784

Vajsová, B., Fasbender, D., Wirnhardt, C., Lemajic, S., Devos, W. Assessing Spatial Limits of Sentinel-2 Data on Arable Crops in the Context of Checks by Monitoring. Remote Sensing, 2020, 12(14), 2195. https://doi.org/10.3390/rs12142195

Valero, S., Morin, D., Inglada, J., Sepulcre, G., Arias, M., Hagolle, O., Dedieu, G., Bontemps, S., Defourny, P., Koetz, B. Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions. Remote Sensing, 2016, 8(1), 55. https://doi.org/10.3390/rs8010055

Volpi, I., Marchi, S., Petacchi, R., Hoxha, K., Guidotti, D. Detecting olive grove abandonment with Sentinel-2 and machine learning: The development of a web-based tool for land management. Smart Agricultural Technology, 2023, 3, 100068. https://doi.org/10.1016/j.atech.2022.100068

Vuolo, F., Neuwirth, M., Immitzer, M., Atzberger, C., Ng, W.-T. How much does multi-temporal Sentinel-2 data improve crop type classification? International Journal of Applied Earth Observation and Geoinformation, 2018, 72, 122–130. https://doi.org/10.1016/j.jag.2018.06.0

Wang, L., He, D.-C. Texture classification using texture spectrum. Pattern Recognition, 1990, 23, 8, 905–910. https://doi.org/10.1016/0031-3203(90)90135-8

Wulder, M. A., Coops, N. C., Roy, D. P., White, J. C., Hermosilla, T. Land cover 2.0. International Journal of Remote Sensing, 2018, 39(12), 4254–4284. https://doi.org/10.1080/01431161.2018.1452075

Zurita-Milla, R., Izquierdo-Verdiguier, E., de By, R. A. (2017). Identifying crops in smallholder farms using time series of WorldView-2 images. In 9th International

Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 2017. https://doi.org/10.1109/multi-temp.2017.8035246

CAPÍTULO 8

Conclusiones y futura investigación

8.1. CONCLUSIONES

Las principales conclusiones que se derivan de la tesis se exponen a continuación:

- La utilización de imágenes aéreas de alta resolución genera resultados muy satisfactorios para el análisis del abandono de cítricos. De entre todos los tipos de datos utilizados, la combinación de características espectrales y de textura extraídas a partir del índice NDVI, a partir de estas imágenes, proporcionó los resultados más precisos de esta tesis. El uso de características de textura extraídas a partir del NDVI con resolución submétrica fue especialmente importante ya que permite modelizar los marcos de plantación típicos de los cultivos de cítricos productivos. No obstante, el uso de estas imágenes muestra alguna limitación en su aplicación debido a su calidad radiométrica y a la complejidad para obtener imágenes reflectancia a nivel de superficie lo que limita su alcance espacial para su aplicación.

- Las imágenes de satélites de alta resolución espacial y espectral, como las imágenes Worldview-3, también generaron unos resultados similares a los obtenidos al utilizar las imágenes aéreas. En este caso se combinaron bandas del espectro visible, infrarrojo cercano, infrarrojo de onda corta y del infrarrojo y características de textura a partir de la banda pancromática. A pesar de esta mayor resolución espectral (16 bandas), en comparación a las imágenes aéreas, no se constató una mejora en los resultados en términos de precisión. Sin embargo, este tipo de imágenes, con una mayor calidad espectral y radiométrica, permiten obtener valores de reflectancia a nivel de suelo lo que puede favorecer su aplicación a zonas de mayor extensión. Un inconveniente de este tipo de productos es que no son de acceso abierto lo que implica un coste económico que dificulta la adquisición de estos datos.

- Las imágenes Sentinel-2 mostraron un potencial de interés significativo para el estudio del abandono de cítricos. Destacan las mejoras obtenidas mediante el uso de series temporales en contraposición al uso de una única imagen. La alta resolución temporal de la constelación Sentinel-2 (5 días) permite registrar diferencias en la evolución temporal de las parcelas productivas, no productivas y abandonadas. Esta información en el dominio temporal permite distinguir de manera precisa estos tres tipos de parcelas. El uso de las imágenes Sentinel-2, con una alta resolución temporal, espectral, radiométrica y de acceso libre muestran un atractivo significativo para el seguimiento del fenómeno estudiado en esta tesis. Además, los datos Sentinel-2 son uno de los productos estándar para el seguimiento de los cultivos en el contexto de la Política Agraria Común Europea. Como inconvenientes, cabe destacar que la resolución espacial de estas imágenes puede ser un factor limitante en un paisaje agrícola altamente fragmentado, como el valenciano.

- En cuanto a la altimetría, cabe destacar que se pueden obtener resultados relevantes para la detección de parcelas abandonadas a partir de puntos 3D derivados de procesos fotogramétricos. La elevada densidad de puntos que se obtienen a partir de esta técnica permite extraer características de textura a partir del CHM lo que contribuye a mejorar los resultados de las clasificaciones realizadas. En contraste, los resultados menos precisos de esta tesis se

132

obtuvieron al utilizar datos LiDAR con una densidad de puntos reducida. Cabe esperar que las nuevas coberturas de datos LiDAR del PNOA mejoren este aspecto y por tanto sea de interés científico evaluar su rendimiento en investigaciones futuras. Las ventajas de utilizar nubes de puntos derivadas de procesos fotogramétricos a partir de imágenes registradas con drones varias. Esta tecnología presenta una mayor flexibilidad y la adquisición de datos es menos compleja comparado a otras técnicas y permite obtener mayor resolución de las imágenes y mayor densidad de puntos. Sin embargo, no es una técnica idónea para monitorizar extensiones de superficie elevadas debido al coste computacional y operativo de la información que se registra. Por tanto, su uso puede tener más interés para estudios de carácter local. Finalmente, la disponibilidad de fuentes de datos 3D puede también ser utilizada en combinación con información espectral lo que puede mejorar los resultados para el estudio del abandono de cítricos como se ha demostrado en esta tesis.

8.2. FUTURA INVESTIGACIÓN

La futura investigación debería explorar el uso de series temporales Sentinel-2 e imágenes de muy alta resolución de manera combinada. La alta resolución en el dominio temporal de los datos Sentinel-2 junto con la alta resolución en el dominio espacial de las imágenes aéreas podría mejorar los resultados obtenidos en investigación. En vista de los resultados obtenidos, pensamos que la combinación de ambas fuentes de datos podría mejorar la exactitud de la clasificación. Sin embargo, el uso de imágenes aéreas de manera transferible requiere avances importantes. Se necesita más investigación para mejorar la calidad radiométrica de las imágenes aéreas, con objeto de mejorar la transferibilidad de los modelos. Futura investigación debería centrarse en desarrollar técnicas para la fusión de dichas imágenes con el objetivo de aprovechar la alta resolución de las imágenes aéreas junto con la gran calidad radiométrica de las imágenes Sentinel-2.

Por otra parte, nuevos estudios deberían analizar los límites espaciales de Sentinel-2, especialmente en paisajes agrícolas altamente fragmentados. En este sentido, sería interesante conocer otras características espaciales de las parcelas distintas a su tamaño que no pueden monitorizarse a través de Sentinel-2, como por ejemplo la morfología. Además, sería de interés caracterizar los cambios en las parcelas desde el inicio del abandono, la velocidad de los cambios, tipos de vegetación, etc. que pueden influir en la detección. Esta información podría ayudar a plantear estrategias para mejorar la precisión en la detección de las parcelas abandonadas.

En esta tesis se utilizaron dos algoritmos de aprendizaje automático ampliamente utilizados en teledetección, como son Random Forest y Support Vector Machines. Sin embargo, en la mayoría de casos de estudio se utilizó Random Forest debido a sus ventajas relacionadas con el tiempo de computación. Sería interesante probar otros algoritmos de clasificación como Extreme Gradient Boosting, redes neuronales artificiales o Support Vector Machines para compararlos en términos de exactitud, transferibilidad espacial y temporal, tiempo de computación y escalabilidad. También sería interesante explorar otros métodos de extracción de características espaciales para extraer características a partir de las imágenes de muy alta resolución y otros enfoques para la extracción de características a partir de las series temporales. El uso de series

temporales más largas también podría ser interesante para este propósito. Una mayor cantidad de datos etiquetados permitiría el uso de métodos de aprendizaje profundo. En este sentido, las series temporales podrían beneficiarse de arquitecturas Long short-term memory o Redes Neuronales Convolucionales en 1D, mientras que las imágenes de muy alta resolución podrían beneficiarse de las Redes Neuronales Convolucionales.

Finalmente, futuras investigaciones deberían centrarse en aplicar y validar las metodologías desarrolladas en esta tesis a otros tipos de cultivos y a otros escenarios de clasificación útiles para la gestión agrícola.