



Universidad Politécnica
de Madrid



**Escuela Técnica Superior de
Ingenieros Informáticos**

Grado en Ingeniería Informática

Trabajo Fin de Grado

**Diseño de un Algoritmo Automático de
Cartografía de Usos del Suelo y sus
Cambios a partir de Imágenes de
Satélite**

Autor: Diego Madruga Ramos
Tutor(a): Consuelo Gonzalo Martín
Cotutor(a): Alicia Palacios Orueta

Madrid, enero 2024

Este Trabajo Fin de Grado se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

Trabajo Fin de Grado

Grado en ingeniería informática

Título: DISEÑO de un ALGORITMO AUTOMÁTICO de CARTOGRAFÍA de USOS del SUELO y sus CAMBIOS a partir de IMÁGENES de SATÉLITE

Enero 2024

Autor: Diego Madruga Ramos

Tutor:

Consuelo Gonzalo Martín

Departamento de Arquitectura y Tecnología de Sistemas Informáticos

ETSI Informáticos

Universidad Politécnica de Madrid

Cotutor:

Alicia Palacios Orueta

Departamento de Ingeniería Agroforestal

ETSI Agronómica, Alimentaria y de Biosistemas

Universidad Politécnica de Madrid

Agradecimientos

Este trabajo de fin de grado ha sido posible gracias a la oportunidades y soporte que me han brindado las siguientes personas e instituciones.

Gracias a mi tutora Consuelo Gonzalo Martín, por su labor docente en asignaturas que han sido de gran apoyo e inspiración para este TFG y por su ayuda en la elaboración del mismo.

A mi cotutora Alicia Palacios Orueta por su mentoría y brindarme la oportunidad de ser parte de un equipo de investigación.

Gracias a todos los miembros del equipo de investigación GEO-Qubidy, por su incesante apoyo y ayuda en la elaboración de este trabajo.

A mis compañeros de buhardilla, Cesar, Tomás, Alfonso y Sergio por su ánimo y paciencia.

Finalmente, gracias a mi familia por su apoyo e inspiración en la consecución de mis metas.

Resumen

Como consecuencia del continuo crecimiento de los avances en el campo de la informática, ciertos sectores pueden crear y automatizar distintos procesos con resultados significativos. Uno de los sectores que más se ve beneficiado por estos avances es el sector ambiental. La teledetección como técnica de observación de la Tierra mediante satélites, aporta información del estado de la superficie terrestre, produciendo grandes beneficios para el monitoreo del medio ambiente, por su aportación de información tanto en tiempo real como de forma histórica.

El objetivo de este estudio es el diseño, desarrollo e implementación de una metodología y herramienta de análisis de la superficie vegetal, mediante el uso de la teledetección junto con el empleo de series de tiempo, para el análisis de patrones en la vegetación.

El desarrollo de esta herramienta comienza con la descarga de las imágenes brutas de teledetección desde el portal oficial. Las misiones seleccionadas para la obtención de las imágenes han sido TERRA y AQUA de la NASA, las cuales llevan en su interior el sensor MODIS que lleva proporcionando imágenes más de 20 años. Una vez obtenidas estas imágenes se calculará el índice espectral NDVI para cada una de las imágenes, con las bandas espectrales necesarias. Una vez calculado el índice espectral se agruparán todas las imágenes para generar la serie de tiempo de la zona de estudio. Durante el proceso de obtención de imágenes puede haber zonas con datos erróneos originados por diversas causas, como pueden ser nubes o reflejos, esto hace necesario que la serie de tiempo sea filtrada y suavizada.

La otra parte de la herramienta consiste en los análisis estadísticos. Como técnica novedosa en el sector, se ha desarrollado un algoritmo que utiliza como fuente principal de información, los datos generados por la función de autocorrelación (ACF) a nivel de píxel. El diseño del algoritmo se basa en el estudio y análisis de pendientes, porcentaje de negativos y frecuencias participantes en la ACF, para su posterior clasificación. La unificación de estas características permite detectar hasta 5 categorías distintas de píxeles determinando ciclo, tendencia o cambio estructural.

Los resultados obtenidos después de la clasificación generan una cartografía que permite observar la distribución espacial de las distintas clases. Estas clases sirven para identificar distintos patrones de comportamiento de la vegetación, determinando las zonas que se mantienen inmutables y que zonas han sufrido cambios a lo largo del rango de estudio. Siendo útil para identificar zonas agrícolas, incendios e incluso servir como indicador de en sistemas de alerta temprana para detectar zonas con aumento de la biomasa y ser potencialmente combustibles.

Abstract

As a consequence of the continuous progress in the field of computer science, certain sectors are able to create and automate different processes with significative results. One of the sectors that benefits the most from these improvements is the environmental sector. Remote sensing as an Earth observation technique using satellites provides information on the state of the Earth's surface, producing important benefits for environmental monitoring by providing information in real time and historically.

The objective of this research is the design, development and implementation of a methodology and tool for the analysis of the vegetation surface, using remote sensing and time series, for the analysis of vegetation patterns.

The development of this tool starts with the download of the raw remotely sensed images from the official portal. The missions selected to obtain the images were NASA's TERRA and AQUA, which contain the MODIS sensor that has been providing images for more than 20 years. Once these images have been obtained, the NDVI spectral index will be calculated for each of the images, with the necessary spectral bands. Once the spectral index has been calculated, all the images will be grouped together to generate the time series of the study area. During the imaging process there may be areas with erroneous data due to various causes, such as clouds or reflections, which makes it necessary for the time series to be filtered and smoothed.

The other part of the tool consists of statistical analyses. As a novel technique in the sector, an algorithm has been developed that uses the data generated by the autocorrelation function (ACF) at pixel level as the main source of information. The design of the algorithm is based on the study and analysis of slopes, percentage of negatives and frequencies participating in the ACF, for their subsequent classification. The unification of these characteristics allows the detection of up to 5 different categories of pixels, determining cycle, trend or structural change.

The results obtained after classification generate a cartography that allows to observe the spatial distribution of the different classes. These classes are used to recognize different patterns of vegetation behaviour, identifying areas that remain unchanged and areas that have changed over the study range. It is useful to identify agricultural areas, fires and even serve as an indicator in early warning systems to detect areas with an increase in biomass and potentially being combustible.

Índice de contenidos

1	Introducción	7
2	Desarrollo	10
2.1	Materiales	10
2.1.1	Área de Estudio.....	10
2.1.2	Índices de Vegetación.....	11
2.1.3	Software Utilizado	12
2.2	Metodología.....	14
2.2.1	Series de Tiempo.....	15
2.2.1.1	Filtros en series de tiempo	15
2.2.1.2	Función de Autocorrelación (ACF)	16
2.2.2	Descarga y Procesado de Imágenes	17
2.2.2.1	Búsqueda y Descarga de Productos MODIS Disponibles	17
2.2.2.2	Procesamiento de Imágenes	21
2.3	Análisis y Diseño	21
2.3.1	Análisis de Pendientes	22
2.3.2	Porcentaje de Negativos.....	23
2.3.3	Transformada de Fourier	23
2.3.4	Autocorrelaciones y pendientes	24
2.3.5	Aplicación de la Transformada de Fourier	29
2.3.6	Clasificación de Píxeles	30
3	Resultados y conclusiones	32
3.1	Resultados	35
3.2	Conclusiones.....	37
4	Análisis de Impacto	40
5	Bibliografía	42
6	Anexos	44
6.1	Costes de la investigación.....	44
6.2	Imagen Clasificada	46

Índice de Ilustraciones

Ilustración 1: Mapa Área de Estudio	10
Ilustración 2: NDVI del Área de Estudio (elaboración propia)	12
Ilustración 3: Logos del software utilizado en el desarrollo.	13
Ilustración 4: Diagrama de flujo de la metodología propuesta	14
Ilustración 5: Serie de tiempo NDVI	15
Ilustración 6: Comparación entre una serie sin filtrar y dos series filtradas con el método de Savitzky-Golay [17].....	16
Ilustración 7: Filtros para la búsqueda de productos en el catálogo.	17
Ilustración 8: Introduciendo fechas en el visor de Earth Data.	18
Ilustración 9: Ventanas que muestran: A la izquierda las herramientas para crear el área de búsqueda, a la derecha el polígono creado	18
Ilustración 10: Ventana que muestra el menú de búsqueda avanzado de imágenes.....	19
Ilustración 11: Lista de productos disponibles en sus distintas versiones para su descarga.....	19
Ilustración 12: Ventana de elección de imágenes para descargar.	20
Ilustración 13: Pagina final de descarga de imágenes.....	20
Ilustración 14: Calculo de la pendiente en el código.....	22
Ilustración 15: Calculo del porcentaje de negativos en el código.....	23
Ilustración 16: Normalización de FFT realizada en el código.....	24
Ilustración 17: Series de tiempo de píxeles de no cambio, 1 ciclo (a, clase 0) y 2 ciclos (b, clase 2).....	25
Ilustración 18: Funciones de autocorrelación de píxeles de no cambio (azul), 1 ciclo (a, clase 0) y 2 ciclos (b, clase 2) y pendiente (rojo).	25
Ilustración 19: Serie de tiempo de píxeles con tendencias, 1 ciclo (c, clase 1) y 2 ciclos (d, clase 3).....	26
Ilustración 20: Funciones de autocorrelación con tendencia (azul), 1 ciclo (c, clase 1) y 2 ciclos (d, clase 3) y pendiente (rojo).	27
Ilustración 21: Serie de tiempo de pixeles con cambios estructurales (clase 4).	28
Ilustración 22: Funciones de autocorrelación con cambios (azul, clase 4) y pendiente (rojo).	28
Ilustración 23: FFT sobre autocorrelaciones.....	30
Ilustración 24: Árbol de decisión CCT.	30
Ilustración 25: Logo Satchange	32
Ilustración 26: Diagrama de clases Satchange.	33
Ilustración 27: Menú principal de la aplicación.	34
Ilustración 28: Mapa de clasificación de píxeles CCT.....	35
Ilustración 29: Zoom del sur de Portugal con ejemplos de las ACF respecto a la clasificación del CCT.	36
Ilustración 30: Incendio de Riba de Saelices.....	37
Ilustración 31: Dehesa Extremeña.	38
Ilustración 32: Zonas de cultivo de Castilla y León.....	38
Ilustración 33: Diagrama de Gantt.	44

Índice de Tablas

Tabla 1: Clasificación de píxeles en base al análisis de la autocorrelación.....	22
Tabla 2: Tabla de análisis de características.	29
Tabla 3: Resumen de la clasificación CCT de la Península Ibérica.....	35
Tabla 4: Distribución de tareas respecto a su duración.....	44
Tabla 5: Lista de componentes del equipo informático utilizado.	45

1 Introducción

La cartografía de usos del suelo y sus cambios desempeña un papel esencial en la gestión del territorio, particularmente en el contexto del cambio global. Históricamente, esta tarea se realizaba mediante visitas de campo, lo que requería un consumo significativo de tiempo y recursos económicos. Sin embargo, los usos del suelo y sus cubiertas están en constante evolución, lo que subraya la necesidad de disponer de metodologías operativas y fiables que permitan un seguimiento frecuente y dinámico. Esta necesidad se ha visto satisfecha en gran medida gracias a la sinergia entre la teledetección y la informática.

La informática juega un papel fundamental en este escenario, permitiendo la adquisición, procesamiento, análisis y visualización de grandes conjuntos de datos recopilados desde satélites, drones y otras fuentes remotas. Esta unión entre teledetección e ingeniería informática ha impulsado avances significativos al ofrecer la capacidad de gestionar eficazmente datos multidimensionales a gran escala. La disponibilidad de series temporales de imágenes de satélite, combinada con la capacidad de cómputo y herramientas geoespaciales, ha dado lugar al desarrollo de herramientas que permiten analizar vastas áreas a lo largo del tiempo[1].

Los avances en la teledetección a través de la informática y sensores satelitales han transformado nuestra capacidad para comprender y monitorear la superficie terrestre a nivel global. Estos sensores, mediante la detección de la radiación reflejada o emitida por la Tierra en diversas longitudes de onda, permiten una identificación detallada de distintas coberturas, como vegetación, suelo desnudo, agua o nieve. Además, la información multiespectral capturada por estos sensores se puede procesar utilizando índices espectrales, como el Índice de Vegetación de Diferencia Normalizada (NDVI)[2], el cual está estrechamente relacionado con la biomasa y la capacidad fotosintética de la vegetación. Esta capacidad para obtener datos multiespectrales y calcular índices como el NDVI es fundamental para la teledetección y la monitorización ambiental.

Además, cuando se integran estos índices en un marco temporal, se abre la puerta a un análisis exhaustivo de la dinámica de la vegetación a diferentes escalas temporales, lo que abarca desde los ciclos naturales de la vegetación hasta los cambios en el uso del suelo. Esto ha permitido un seguimiento detallado de la vegetación y otros aspectos ambientales a nivel global y a lo largo del tiempo, lo que ha sido valioso para numerosos estudios y aplicaciones.

El sensor MODIS (Moderate Resolution Imaging Spectroradiometer) (<https://modis.gsfc.nasa.gov>), transportado en los satélites TERRA y AQUA de la NASA, ha desempeñado un papel crucial en esta área desde el año 2000. MODIS proporciona datos multiespectrales con una alta resolución temporal, ofreciendo información diaria sobre la superficie terrestre. Actualmente hay sensores a bordo de satélites con unas capacidades muy potentes en cuestión de resolución espacial y temporal como Landsat (<https://landsat.gsfc.nasa.gov>) y Sentinel (<https://sentinels.copernicus.eu/web/sentinel/home>); sin embargo, los productos de MODIS ofrecen un equilibrio excelente entre resolución espacial, frecuencia temporal y longitud del periodo de observación lo que permite evaluar la superficie de la tierra a escala regional y durante un largo periodo de tiempo.

La presencia de tendencias es muy frecuente en el medio natural, pudiendo ser debidas a distintos procesos ambientales, como aumento progresivo de biomasa o procesos de desertificación, entre otros. La teledetección multitemporal permite la observación de tendencias en grandes superficies por lo que puede generar una gran cantidad de información muy relevante para la gestión de la naturaleza. En este contexto, el desarrollo de algoritmos capaces de identificar tendencias a partir de datos de satélite se convierte en un área de gran interés. Una de las metodologías capaces de identificar tendencias es el análisis de series de tiempo (AST) [3].

El AST ha sido aplicado principalmente en sectores como el de la economía, la física y la medicina entre otros [4]. Uno de sus principales objetivos es la predicción, para lo cual es común utilizar la metodología Box-Jenkins mediante la cual se identifica el modelo de predicción más apropiado [5]. La función de autocorrelación (ACF) es una de las herramientas utilizadas en esta metodología para la identificación de periodicidad y estacionalidad de las series de tiempo [6].

En el AST el estudio de las tendencias se apoya generalmente en dos tipos de análisis, el test de Mann-Kendall, que evalúa el nivel de parentesco entre dos conjuntos de datos comparándolos entre si a lo largo del tiempo, determinando si los datos muestran un patrón claro de cambio o si son aleatorios [7], y el test de raíces unitarias que evalúa una dependencia temporal con una hipótesis nula que determina si la raíz unitaria es estacionaria o no [8]. En este mismo ámbito, se ha observado que la función de autocorrelación de una serie de tiempo con tendencia presenta unos patrones específicos, lo cual permite la identificación de tendencias más sutiles y de otros tipos de patrones.

En este contexto, en este trabajo de fin de grado se propone desarrollar un algoritmo innovador para identificar cambios de uso del suelo mediante el análisis estadístico de series temporales provenientes de imágenes de satélite. El producto de teledetección seleccionado para este proyecto es el MOD09Q1V61 de MODIS ([link del producto](#)), que proporciona compuestos de 8 días con una resolución espacial de 250 metros. El periodo de estudio abarca desde 2000 hasta finales de 2022, y la zona de estudio abarca toda la Península Ibérica.

El diseño de este algoritmo se enmarca en el proceso de desarrollo de una herramienta que ofrecerá un servicio automatizado de cartografía de usos del suelo y sus cambios, aprovechando imágenes de teledetección, específicamente satelitales. Este enfoque permitirá la generación operativa, dinámica y frecuente de mapas de cambios de uso y cobertura del suelo, reduciendo significativamente la necesidad de tareas de campo para la adquisición de información. Esta herramienta de análisis y seguimiento ambiental que se encuentra en desarrollo en el Grupo de Investigación en UPM GEO-Qubidy (<https://blogs.upm.es/dynamicland/>).

El algoritmo a desarrollar comprenderá dos partes esenciales: (1) el cálculo de la función de autocorrelación a nivel de píxel, y (2) el diseño de un procedimiento para identificar y clasificar los cambios de uso del suelo basados en la dependencia temporal, derivada de las características de la función de autocorrelación a nivel de píxel.

El desarrollo de este algoritmo implica una serie de actividades fundamentales, que incluyen la comprensión de los productos de imágenes de satélite (MODIS), la familiarización con las metodologías de estudio de series temporales, el análisis de las series de tiempo y sus autocorrelaciones, el diseño y desarrollo del algoritmo de detección de cambios en la cubierta del suelo, la evaluación de los resultados obtenidos y la propuesta de mejoras y futuras direcciones en este ámbito.

2 Desarrollo

2.1 Materiales

2.1.1 Área de Estudio

La Península Ibérica, ubicada en el suroeste de Europa, emerge como un entorno de estudio excepcional en el ámbito de la teledetección y la monitorización ambiental. Esta vasta región, que abarca tanto España como Portugal, se caracteriza por su impresionante diversidad de cubiertas terrestres, que incluyen bosques frondosos, extensiones de matorrales, cultivos agrícolas, áreas urbanas y una variada costa litoral. La singular topografía de la Península Ibérica, que abarca desde las altas cumbres de los Pirineos hasta las extensas llanuras de la Meseta Central, crea una amplia gama de condiciones ambientales que la convierten en un laboratorio natural ideal para la investigación en teledetección [9].

Esta región presenta desafíos únicos y oportunidades para el desarrollo de algoritmos de detección y seguimiento de cambios en el uso del suelo, así como para la evaluación de la dinámica de la vegetación a lo largo del tiempo. Debido a su variabilidad climatológica, que abarca desde climas mediterráneos cálidos y secos en el sur hasta climas oceánicos más frescos y húmedos en el norte, la Península Ibérica experimenta una amplia gama de temperaturas y niveles de humedad a lo largo del año. Esta variabilidad climática influye directamente en la distribución de las cubiertas terrestres y en la respuesta de la vegetación, lo que la convierte en un lugar estratégico para el desarrollo de proyectos e investigaciones relacionadas con la monitorización ambiental y la gestión sostenible de los recursos naturales.



Ilustración 1: Mapa Área de Estudio

2.1.2 Índices de Vegetación

Los índices de Vegetación son un subconjunto de lo que se denominan índices espectrales. Los índices espectrales son combinaciones lineales de bandas que proporcionan un beneficio para la visualización/detección de un determinado producto [10].

En este caso nos vamos a centrar en los índices de vegetación, que como su propio nombre señala, se centra en resaltar cubiertas vegetales. Hay diversos índices de vegetación como el EVI, el SAVI y el NDVI que es el usaremos en el proyecto.

El NDVI (Índice de Vegetación de Diferencia Normalizada) es un indicador utilizado para analizar y evaluar la vegetación mediante la medición de la radiación infrarroja reflejada o emitida por la superficie terrestre. Esto se logra mediante sensores remotos instalados en satélites. El NDVI se calcula comparando la intensidad de la radiación en la banda del infrarrojo cercano (NIR) y la banda roja (RED) del espectro electromagnético utilizando Ecuación 1 [2]:

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

Ecuación 1: Formula NDVI

El NDVI proporciona información valiosa sobre la salud y el estado de la vegetación. Puede identificar áreas con vegetación en crecimiento, vegetación saludable o vegetación dañada, por ejemplo, por incendios forestales. Aunque los valores del NDVI siempre oscilan entre -1 y 1, varios factores pueden influir en estos valores, como las condiciones atmosféricas, la exposición del suelo y la fase fenológica de la vegetación [2].

En la Ilustración 2 se muestra el índice NDVI calculado para una imagen de la zona de estudio, procedente del producto MOD09Q1 de MODIS, donde se puede observar cómo varía los valores en función del tipo y la cantidad de vegetación de la superficie. Al trabajar con series de tiempo el volumen de información que se almacena depende directamente del formato numérico utilizado, para reducir el tamaño en disco de las imágenes, se multiplica el NDVI por 10000 pudiendo pasar del formato Float32 a Int16.

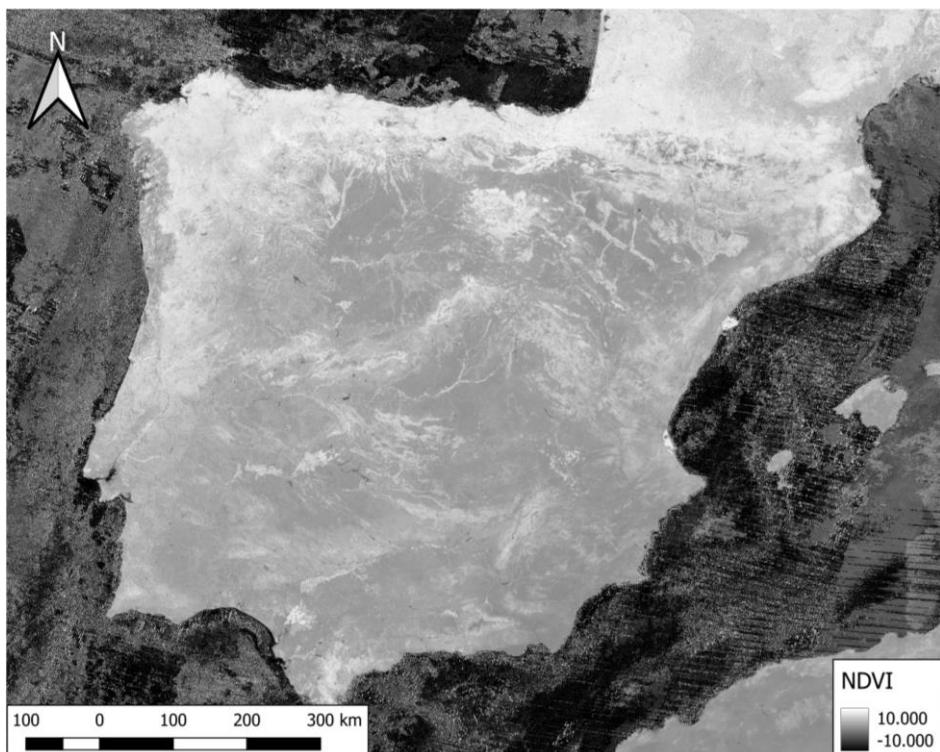


Ilustración 2: NDVI del Área de Estudio (elaboración propia)

2.1.3 Software Utilizado

En este punto comentaremos y analizaremos las herramientas utilizadas en el desarrollo del proyecto.

Se han empleado una variedad de herramientas de software clave que desempeñaron un papel fundamental en la obtención de los objetivos de investigación y análisis. Estos programas incluyeron Visual Studio Code (VSCode), Python [11], Jupyter Notebook [12], QGIS (Quantum Geographic Information System) [13] y GDAL (Geospatial Data Abstraction Library) [14].

- **Visual Studio Code (VSCode):** Se ha utilizado VSCode como entorno de desarrollo integrado (IDE) principal. Este versátil editor de código abierto de Microsoft que proporciona un entorno de programación cómodo y altamente personalizable para escribir y depurar código en Python, el lenguaje que se eligió para el desarrollo. VSCode permite organizar y gestionar eficazmente los scripts y módulos, lo que ha resultado crucial para el desarrollo y la mantenibilidad del código.
- **Python:** Python fue el lenguaje de programación central en el desarrollo. Ha sido elegido debido a su amplio ecosistema de bibliotecas y su popularidad en la comunidad de análisis de datos y ciencia de datos. Con Python, se ha podido implementar algoritmos y técnicas de análisis de datos de manera eficiente y realizar cálculos estadísticos y geoespaciales. Además, la biblioteca NumPy[15] desempeñó un papel esencial en el manejo y procesamiento eficiente de datos numéricos, enriqueciendo significativamente la capacidad de análisis.

- **Jupyter Notebook:** Para realizar pruebas del diseño y desarrollo del algoritmo a escala pequeña, se ha utilizado Jupyter Notebook. Esta aplicación web me permitió crear documentos interactivos que combinaban código Python, visualizaciones y explicaciones detalladas. Esto resultó muy útil para mostrar el proceso de análisis, los resultados y facilitar la reproducibilidad de la investigación.
- **QGIS (Quantum Geographic Information System):** QGIS fue una herramienta importante, especialmente para el manejo de datos geoespaciales. Se ha utilizado para visualizar y analizar datos espaciales, lo que fue esencial para revisar resultados de tipo ráster.
- **GDAL (Geospatial Data Abstraction Library):** Se ha utilizado GDAL para tareas avanzadas de procesamiento y análisis de datos geoespaciales. Esta biblioteca de código abierto ha permitido llevar a cabo conversiones de formatos, transformaciones de proyecciones y manipulación de datos geoespaciales de manera eficiente, garantizando la integridad y la calidad de los datos.

En resumen, la unión de todas estas herramientas ha proporcionado una base sólida, permitiendo realizar análisis de datos geoespaciales con precisión y rigor, y comunicar los avances de manera efectiva. Estas herramientas desempeñaron un papel crucial en la adquisición de los objetivos del proyecto y en la presentación de los resultados de manera profesional.



Ilustración 3: Logos del software utilizado en el desarrollo.

2.2 Metodología

En la Ilustración 4 se muestra el diagrama de flujo para obtener la imagen clasificada. Para generar las series de tiempo; será necesario descargar todas las imágenes comprendidas en el periodo de estudio, se calculará el índice espectral seleccionado para cada imagen y se unirán todas las imágenes generando un “stack” o cubo de imágenes que contiene todas las series de tiempo a nivel de píxel. A estas series de tiempo se les aplicarán varios métodos estadísticos, se ha utilizado el filtro Savitzky-Golay para suavizar la serie y eliminar datos aberrantes, junto a la función de autocorrelación para observar la estabilidad temporal de cada píxel.

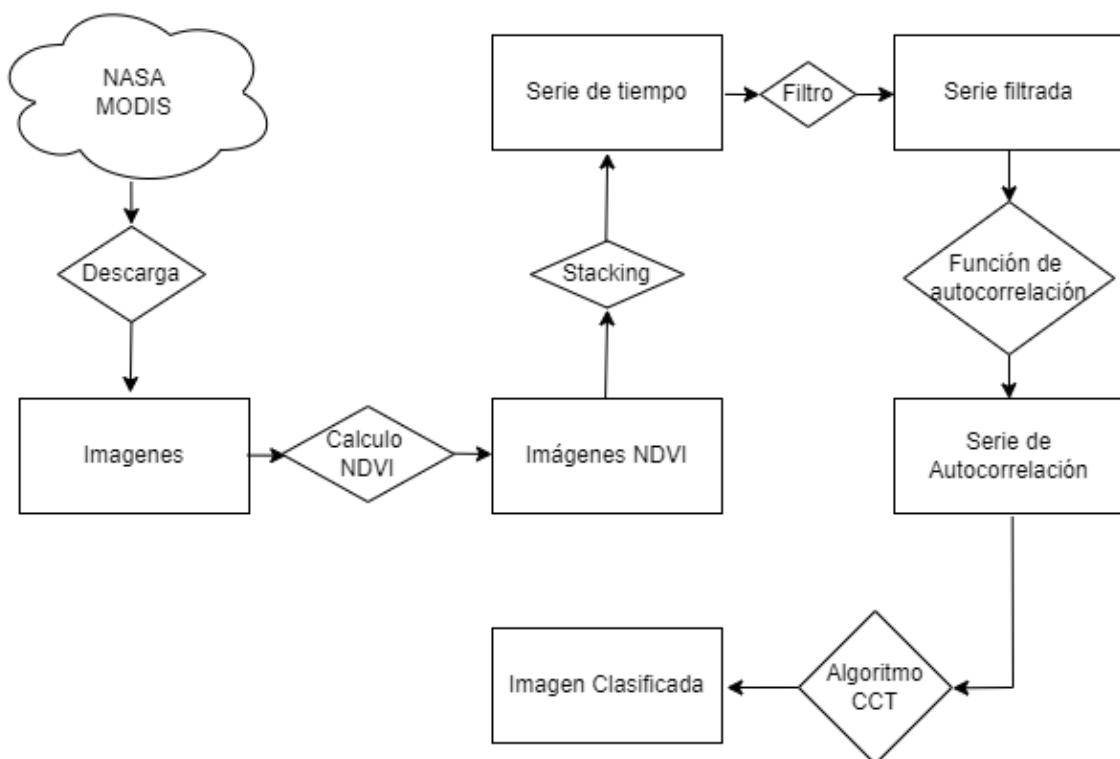


Ilustración 4: Diagrama de flujo de la metodología propuesta

2.2.1 Series de Tiempo

Las series de tiempo son combinaciones de datos que conforman en su unión una temporalidad. Estas series temporales nos indican como ha variado o no un dato a lo largo del rango de estudio.

En nuestro caso generaremos series de tiempo con las imágenes NDVI descargadas y calculadas anteriormente (2.2.2), lo que nos permiten ver y analizar cómo ha variado el verdor de un píxel MODIS (250m) a lo largo de los años, como se puede observar en la Ilustración 5.

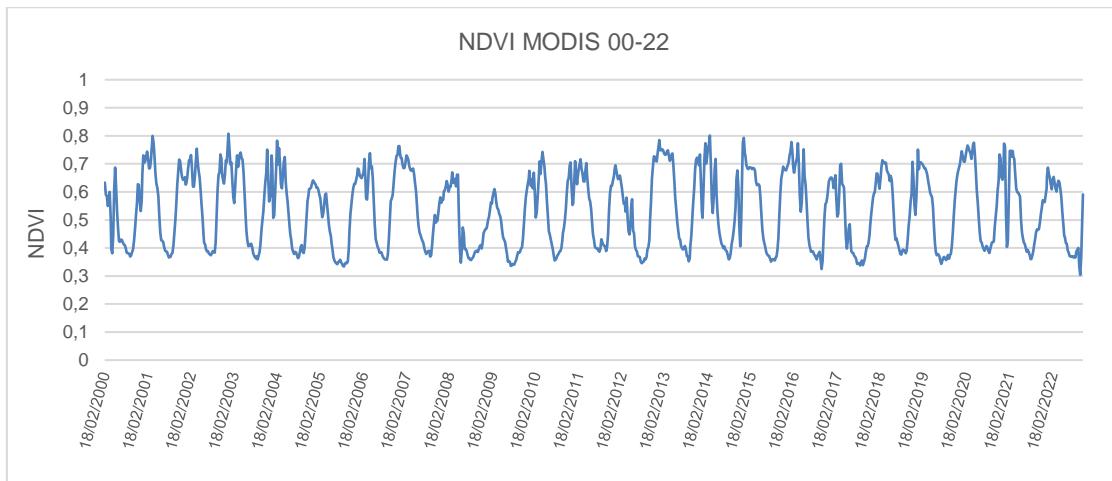


Ilustración 5: Serie de tiempo NDVI

El análisis estadístico de series de tiempo permite extraer información sobre la dinámica de las variables en estudio en términos de su variabilidad estadística, presencia de tendencias, ciclos y periodicidades, entre otras, así como modelizar y hacer predicciones. Aplicado a la teledetección ambiental, este tipo de análisis permite evaluar el funcionamiento de los ecosistemas a partir de su dinámica general [16].

2.2.1.1 Filtros en series de tiempo

El filtrado es una herramienta que utilizamos sobre la serie de tiempo para suavizar la curva y eliminar el ruido producido por diversas razones.

En este caso se ha usado el filtro de Savitzky-Golay[17] para suavizar la curva de la función. Este filtro se basa en el cálculo de una regresión polinomial local (de grado k), con al menos $k+1$ puntos equiespaciados, para determinar el nuevo valor de cada punto. La función que resultará será similar a la función de entrada, pero suavizada. La ventaja de este filtro es que preserva características de la serie inicial como la distribución de los máximos y mínimos relativos y el ancho de los picos. En la Ilustración 6, podemos observar las diferencias entre una serie original (puntos azules), y dos series ya filtradas (negro y rojo).

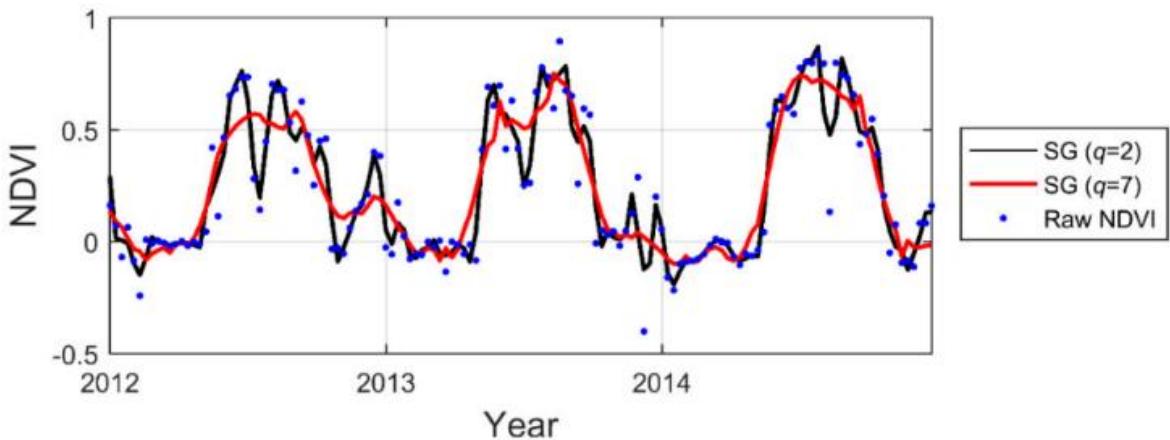


Ilustración 6: Comparación entre una serie sin filtrar y dos series filtradas con el método de Savitzky-Golay [17]

2.2.1.2 Función de Autocorrelación (ACF)

La autocorrelación es una serie temporal en la que el valor actual depende de valores pasados se denomina una serie temporal autocorrelacionada. Si una serie es estacionaria, la dependencia en la serie entre dos puntos separados por k unidades de tiempo (es decir su retardo) puede describirse mediante el coeficiente de autocorrelación r_k en el retardo k . Este mide el grado en que un valor de la serie por encima o por debajo de la media en el tiempo t tiende a ser seguido por un valor de la serie por encima o por debajo de la media k unidades de tiempo posteriormente. El gráfico de los coeficientes de autocorrelación r_k respecto de k para $k = 1, 2, \dots$ se denomina función de autocorrelación de la serie. Esta describe cómo la correlación en la serie desaparece a medida que aumenta la separación o el retardo k entre dos puntos en el tiempo. Diferentes modelos dan lugar a patrones característicos de autocorrelación. Por lo tanto, pueden usarse para reconocer qué modelo ajustar inicialmente a una serie temporal [18].

La expresión para el cálculo de la función de autocorrelación es:

$$\hat{r}_k = \frac{\sum_{t=k+1}^N (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^N (y_t - \bar{y})^2}$$

Ecuación 2: Función de autocorrelación.

donde:

- y : es la serie temporal de la variable en estudio.
- \bar{y} : es la media de la serie temporal de la variable en estudio.
- t : indica el instante de tiempo.
- K : indica el número de retardos.

2.2.2 Descarga y Procesado de Imágenes

2.2.2.1 Búsqueda y Descarga de Productos MODIS Disponibles

Si se quiere saber qué producto o productos MODIS se adaptan mejor al objetivo del trabajo, se deberá acudir al siguiente enlace, donde se encuentran todos los productos disponibles con sus características técnicas:

https://lpdaac.usgs.gov/product_search/?view=cards&sort=title

Cada producto tiene distintas características en cuanto a (1) nivel de procesado, (2) resolución temporal, (3) resolución espacial, etc. En la parte superior de la página, (Ilustración 7), tendremos distintos filtros para poder buscar el producto con la resolución espacial y temporal que mejor se adapte a nuestro objetivo y área estudiada.

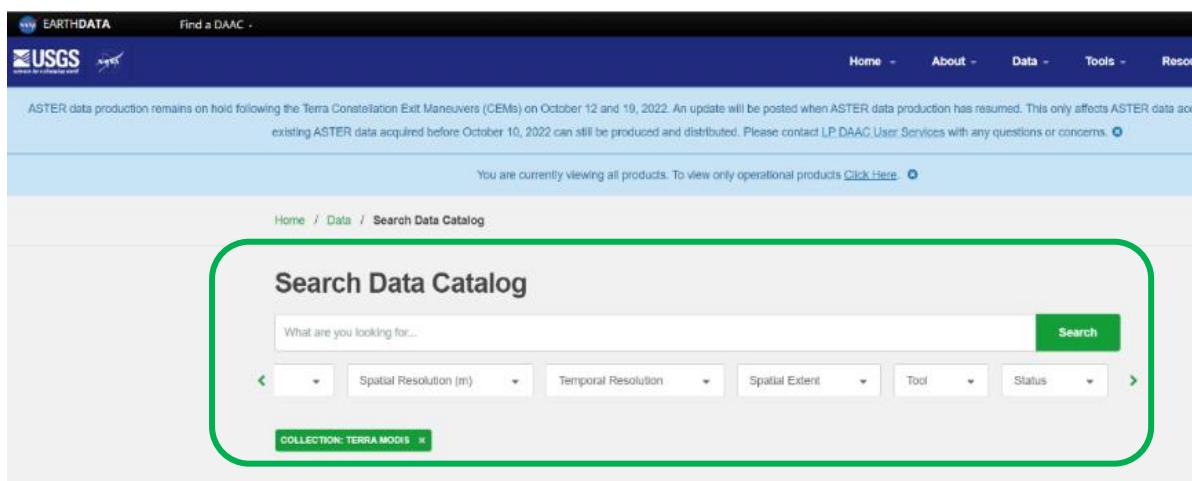


Ilustración 7: Filtros para la búsqueda de productos en el catálogo.

Para continuar con el proceso de descarga de imágenes, en este ejemplo se usará el producto MOD09Q1. Este producto, tiene dos versiones. En la fecha en la que se obtuvieron las imágenes, la versión más reciente era la v061. Este producto proporciona imágenes cada 8 días (resolución temporal) con un tamaño de píxel de 250m (resolución espacial).

2.2.2.1.1 Búsqueda de imágenes

Cuando busquemos las imágenes, lo primero será escoger las fechas de inicio y fin del periodo de estudio (Ilustración 8). Las fechas de inicio y fin se introducen en el botón que se encuentra en el recuadro verde.

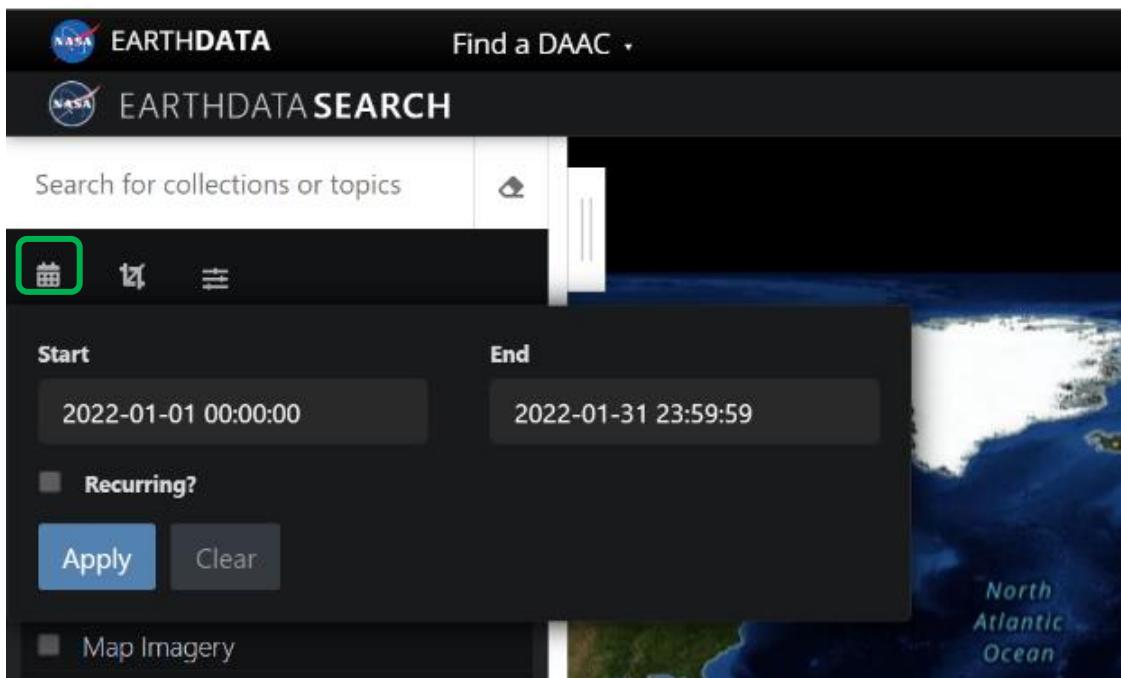


Ilustración 8: Introduciendo fechas en el visor de Earth Data.

El siguiente paso será crear un polígono que comprenda el área de estudio. Eso se puede hacer con el botón del recuadro naranja, de la Ilustración 9. Se puede elegir entre varias formas de crear el recinto que delimita el área de estudio. Aquí se usará el polígono como ejemplo.

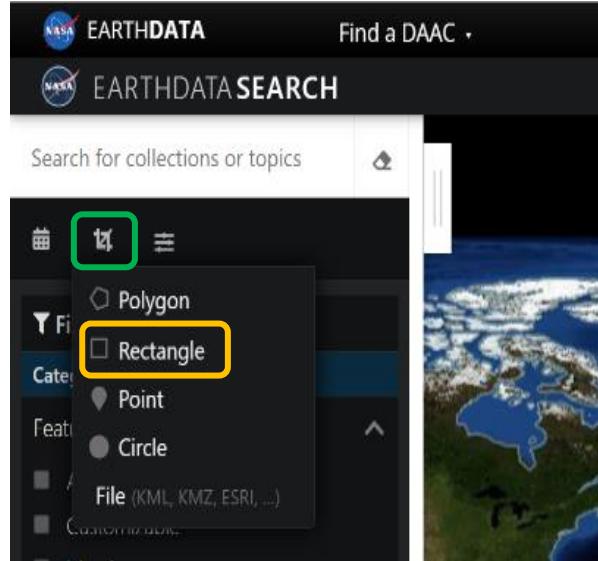


Ilustración 9: Ventanas que muestran: A la izquierda las herramientas para crear el área de búsqueda, a la derecha el polígono creado

Ahora en el menú de la izquierda (Ilustración 10: Ventana que muestra el menú de búsqueda avanzado de imágenes.), se buscará el submenú “Instruments” y de entre todos los instrumentos disponibles se escogerá MODIS.

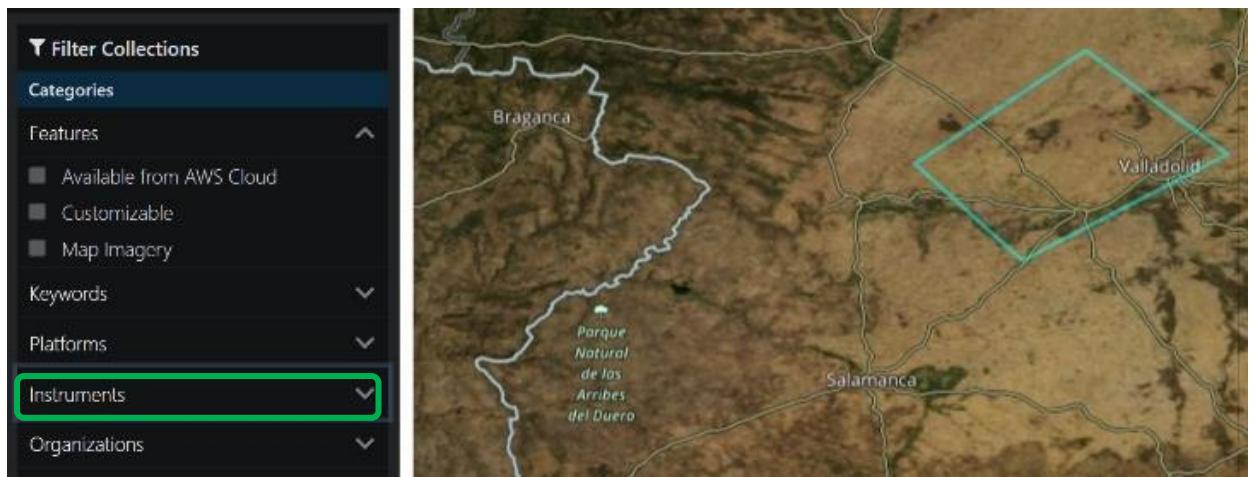


Ilustración 10: Ventana que muestra el menú de búsqueda avanzado de imágenes.

2.2.2.1.2 Descarga de imágenes

Una vez escogido el producto, se podrá continuar con el proceso de descarga de imágenes. Primero se introducirá el nombre del producto en el buscador y después se presionará “Enter” para iniciar la búsqueda de imágenes.

Para iniciar el proceso de descarga, se hará clic sobre el producto escogido, (Ilustración 11). Como comentamos anteriormente, se escogerá la versión más actual disponible (V061).

MOD09Q1

3 Matching Collections

Showing 3 of 3 matching collections

MODIS/Terra Surface Reflectance 8-Day L3 Global 250m SIN Grid V006
5 Granules 2000-02-24 ongoing

The MOD09Q1 Version 6 product provides an estimate of the surface spectral reflectance of Terra Moderate Resolution Imaging Spectroradiometer (MODIS)...
GEOSS • MOD09Q1 v006 - LP DAAC

MODIS/Terra Surface Reflectance 8-Day L3 Global 250m SIN Grid V061
5 Granules 2000-02-18 ongoing

The MOD09Q1 Version 6.1 product provides an estimate of the surface spectral reflectance of Terra Moderate Resolution Imaging Spectroradiometer (MODIS)...
GEOSS • MOD09Q1 v061 - LP DAAC

Ilustración 11: Lista de productos disponibles en sus distintas versiones para su descarga.

Se abrirá una nueva página (Ilustración 12) con las distintas imágenes que hay en el periodo temporal escogido. Se podrán bajar de una en una haciendo clic en el recuadro rojo, o descargar todas haciendo clic en el recuadro “Download All”

Ilustración 12: Ventana de elección de imágenes para descargar.

En la nueva página, se clicará sobre el botón verde “Download Data” para continuar con el proceso de descarga donde entraremos en una nueva página (Ilustración 13), que nos proporcionará varias opciones de descarga.

Download Status

This page will automatically update as your orders are processed. The Download Status page can be accessed later by visiting <https://search.earthdata.nasa.gov/downloads/5687405453> or the [Download Status and History page](#).

Ilustración 13: Pagina final de descarga de imágenes.

Si se dispone de Linux, la forma más sencilla es usando el Script de descarga o “*Download Script*” que proporciona la página. En caso de disponer de Windows, la forma más sencilla es, o bien haciendo clic en cada enlace para iniciar la descarga, o mediante un gestor de links. En la *Store* de *Google Chrome* hay varias extensiones para gestionar links, una de ellas se llama “*Simple Mass Downloader*” que nos permitirá bajar múltiples enlaces o archivos de un sitio web.

2.2.2.2 Procesamiento de Imágenes

Una vez descargadas las imágenes, se procede a calcular el índice espectral seleccionado para el análisis. En este caso se realizará el cálculo del NDVI, al haber seleccionado MODIS a 250m es el índice de vegetación más sencillo de calcular con las bandas que proporciona cada imagen. Cabe mencionar que MODIS proporciona otro producto (MOD13Q1) que contiene el NDVI calculado, pero la resolución temporal de este se reduce de 8 a 16 días, y al realizar un análisis de series de tiempo es preferible tener una mayor densidad de imágenes. Para calcular este índice se han seguido los siguientes pasos:

1. Extracción de las bandas del rojo (RED) y del infrarrojo cercano (NIR).
2. Utilizar la Ecuación 1 para calcular el NDVI, con las dos bandas extraídas.
3. Guardar el resultado como una imagen nueva manteniendo la proyección y transformación geográfica de la original.

2.3 Análisis y Diseño

El algoritmo se basa en el estudio de la función de autocorrelación calculada sobre la serie temporal de NDVI, ayudando a abstraerse de las variabilidades climáticas y de cubierta comentadas en el área de estudio 2.1.1, debido a que se compara el píxel consigo mismo, permitiendo su aplicabilidad en todas las posibles coberturas de la Península Ibérica y otras regiones del globo. Al superar los desafíos que plantea la variabilidad climática y de cubierta, el algoritmo aspira a proporcionar una herramienta efectiva y precisa para la monitorización ambiental y la gestión de recursos naturales en esta rica y dinámica área geográfica. De esta manera, se contribuye al avance de la teledetección y al entendimiento de los procesos terrestres en un contexto tan variado como la Península Ibérica.

Después de analizar las formas de las autocorrelaciones, se observó que, dependiendo del cambio sufrido en la serie de tiempo del píxel, la serie de autocorrelación tiene una forma distinta, por lo que se generaron cinco clases de píxeles:

Tabla 1: Clasificación de píxeles en base al análisis de la autocorrelación.

Clase	Descripción
0	<u>1 ciclo:</u> Son píxeles que no han sufrido un cambio drástico en el rango de fechas de estudio. Su forma se define por una onda sin grandes cambios de amplitud y con una frecuencia constante.
1	<u>1 ciclo con tendencia:</u> Son píxeles que han sufrido un cambio en su historial, pero han recuperado su nivel de verdor (NDVI). Su forma se define como una onda con una pendiente negativa, sin grandes cambios de amplitud y con una frecuencia constante.
2	<u>2 ciclos:</u> Son píxeles que no han sufrido un cambio drástico en el rango de fechas de estudio. Su forma se define por una onda sin grandes cambios de amplitud y con una frecuencia constante dominante y una frecuencia secundaria
3	<u>2 ciclos con tendencia:</u> Son píxeles que han sufrido un cambio en su historial, pero han recuperado su nivel de verdor (NDVI). Su forma se define como una onda con una pendiente negativa, sin grandes cambios de amplitud y con una frecuencia dominante y una frecuencia secundaria.
4	<u>Cambio Estructural:</u> son píxeles que han sufrido uno o varios cambios a lo largo de su temporalidad, y no han recuperado sus niveles de NDVI pudiendo haber sido producido por un suceso que cause una degradación de la cubierta vegetal. Su forma no está definida siendo un cumulo de ruido de valores bajos.

2.3.1 Análisis de Pendientes

Tal y como se explica en la Tabla 1, una de las fuentes de información para clasificar a los distintos píxeles es la pendiente de la onda.

Para calcular la pendiente de la onda, la dividiremos en dos mitades a las cuales se calcularán la media de sus valores. Los valores resultantes de las dos mitades se restarán para obtener así la pendiente que se calcula siguiendo la Ecuación 3.

$$slope = \bar{p2} - \bar{p1}$$

Ecuación 3: Calculo de la pendiente.

donde:

- $\bar{p1}$: Media de la primera parte de la onda.
- $\bar{p2}$: Media de la segunda parte de la onda.

```
1 p1, p2 = np.split(serie, 2)
2 self.slope = (np.mean(p2) - np.mean(p1)).round(2)
```

Ilustración 14: Calculo de la pendiente en el código.

2.3.2 Porcentaje de Negativos

Otro factor importante fue la relación entre los valores positivos y negativos dentro de la onda. Teniendo en cuenta que para una onda estable el baremo de la cantidad de números positivos y negativos debe ser cercano al 50% y en un cambio hay más valores negativos, el diseño del algoritmo iba a estar afectado por el porcentaje de valores negativos que tiene la onda. Siendo una característica para tener en cuenta dentro de la clasificación.

$$\%_{negativos} = \frac{\text{Número de valores negativos de la serie}}{\text{Longitud total de la serie}} \times 100$$

Ecuación 4: Calculo de porcentaje de negativos

```
1 self.negvalues = round(len(np.where(serie < 0)[0])*100/len(serie), 2)
```

Ilustración 15: Calculo del porcentaje de negativos en el código.

2.3.3 Transformada de Fourier

La Transformada de Fourier es una herramienta matemática utilizada en una amplia gama de disciplinas, desde la ingeniería hasta la física y la música. Su propósito principal es descomponer una señal, que puede ser cualquier función, en sus componentes de frecuencia individuales. En otras palabras, nos permite analizar una señal y determinar las frecuencias que la componen [19].

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-j2\pi ft} dt$$

Ecuación 5: Transformada de Fourier

donde:

- $X(f)$ es la representación en el dominio de la frecuencia de la señal normalizada.
- $x(t)$ es la señal original en el dominio del tiempo.
- f es la frecuencia.
- j es la unidad imaginaria ($j^2 = -1$).
- La integral se toma desde menos infinito hasta más infinito.

Se utilizó la Transformada Rápida de Fourier (Fast Fourier Transform)[20] debido a que realiza de forma más eficiente los cálculos, que se define siguiendo la Ecuación 6.

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j(2\pi/N)kn}$$

Ecuación 6: Transformada Rápida de Fourier

donde:

- $X(k)$ es el coeficiente de frecuencia en la posición k .
- $x(n)$ es el valor de la señal en el tiempo n .
- N es la longitud total de la señal.
- j es la unidad imaginaria.

Para investigar las frecuencias presentes en la onda de la función de autocorrelación, normalizando la serie de Fourier resultante elevando al cuadrado sus valores, para obtener la representación de la densidad espectral de potencia normalizada.

```
1 fft = np.fft.fft(serie)
2 fftfreq = np.fft.fftfreq(len(serie))
3 fftnorm = abs(fft[:int(len(fft**2)/2)])**2
4 fftfreqnorm = fftfreq[:int(len(fftfreq)/2)]
```

Ilustración 16: Normalización de FFT realizada en el código.

Esta técnica me permitió identificar las frecuencias dominantes en la señal, lo que resultó fundamental para la identificación de dobles ciclos y tendencias.

2.3.4 Autocorrelaciones y pendientes

La autocorrelación permite observar las frecuencias que participan dentro de la serie de tiempo, pudiendo detectar estacionalidades [6]. A continuación, veremos diferentes tipos series de tiempo, sus funciones de autocorrelación y sus pendientes en función de la clase de píxel (Tabla 1).

En la Ilustración 17 se observa las series de ciclos sin tendencias, correspondientes a las clases 0 y 2. Tal y como se han descrito estas series tienen una estabilidad en su temporalidad. En la Ilustración 18, se detectan fácilmente que ondas están formadas por una y dos frecuencias predominantes respectivamente y a su vez las pendientes (línea roja) son prácticamente paralelas al eje x, es decir, cercanas a 0.

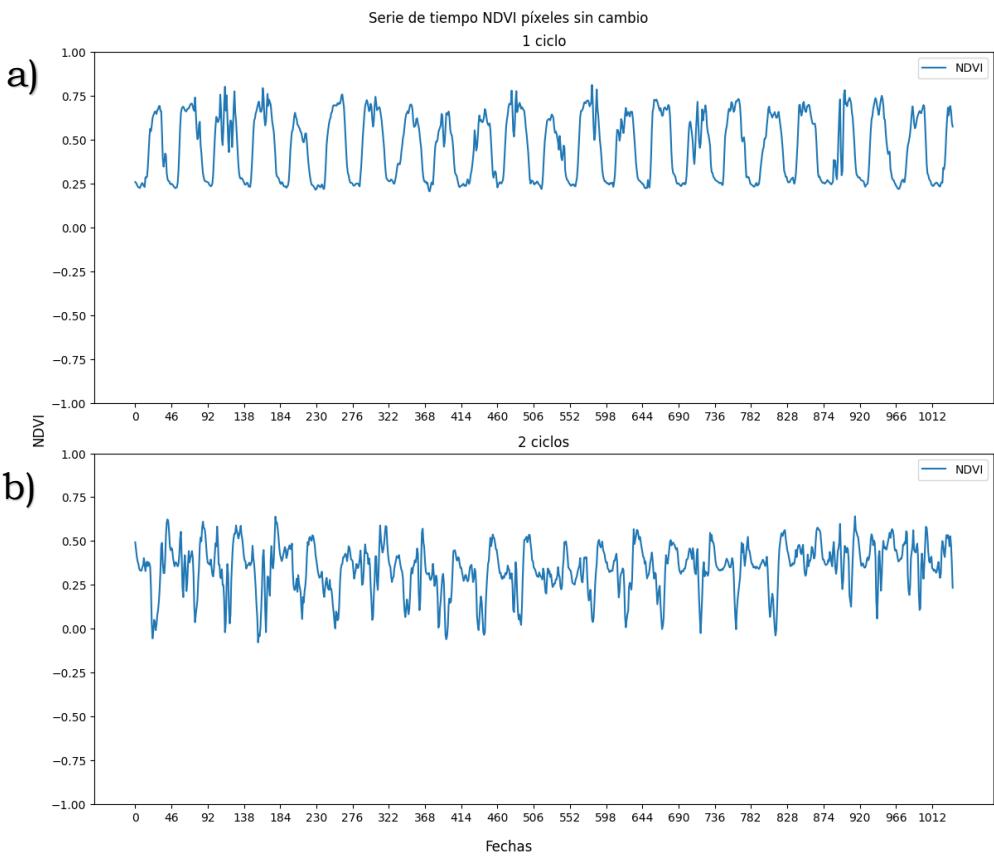


Ilustración 17: Series de tiempo de píxeles de no cambio, 1 ciclo (a, clase 0) y 2 ciclos (b, clase 2).

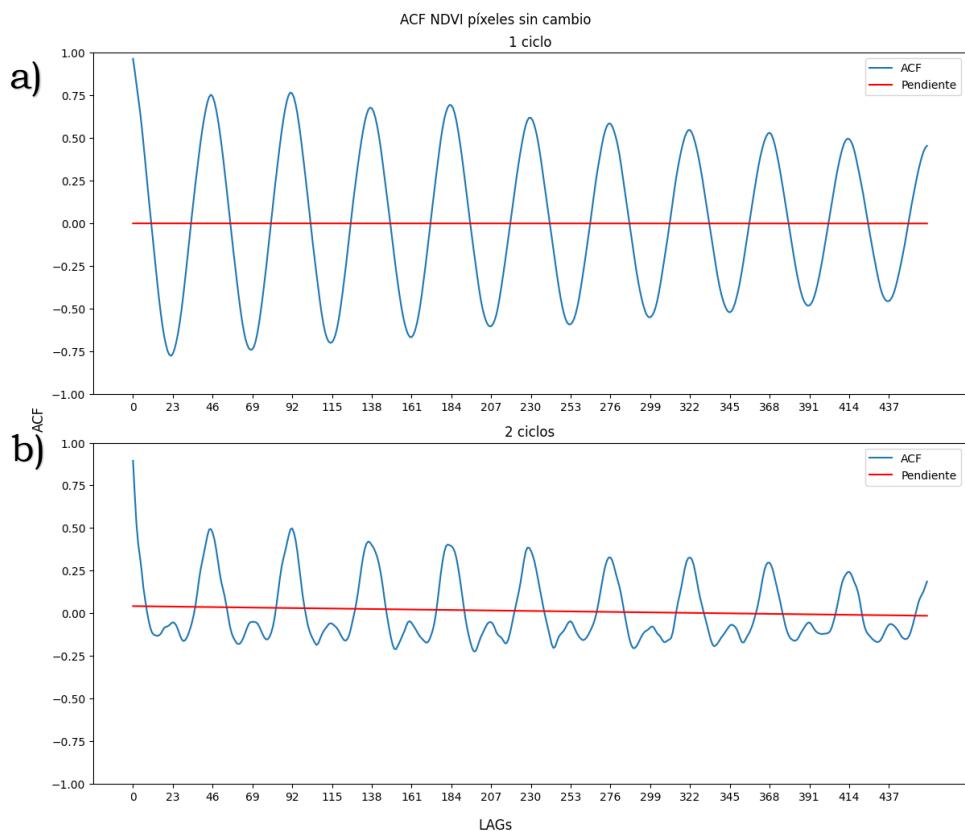


Ilustración 18: Funciones de autocorrelación de píxeles de no cambio (azul), 1 ciclo (a, clase 0) y 2 ciclos (b, clase 2) y pendiente (rojo).

En la Ilustración 19, se muestra series de tiempo con tendencias, correspondientes a las clases 1 y 3, que suelen ser píxeles que han sufrido un cambio, pero han conseguido recuperar sus niveles o zonas que aumente la cantidad de masa vegetal generando un aumento del NDVI a lo largo de los años. En la Ilustración 20, las ondas tienen las mismas frecuencias de las clases 0 y 2, pero añadiendo una frecuencia que hace que la serie obtenga una pendiente descendiente como se observa con la línea roja.

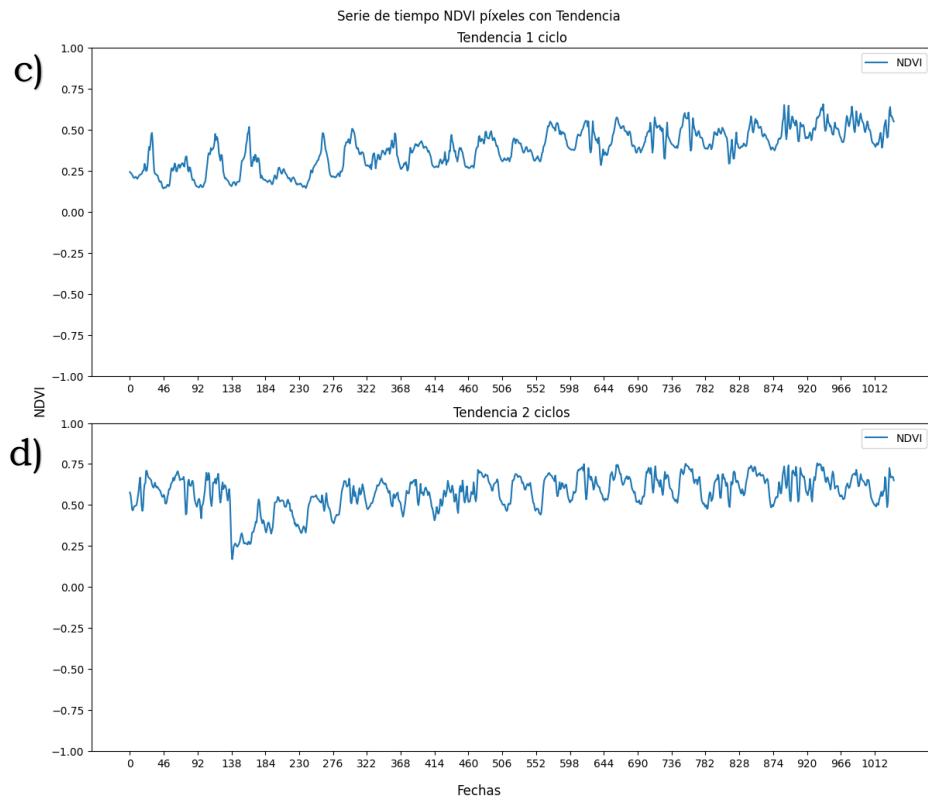


Ilustración 19: Serie de tiempo de píxeles con tendencias, 1 ciclo (c, clase 1) y 2 ciclos (d, clase 3).

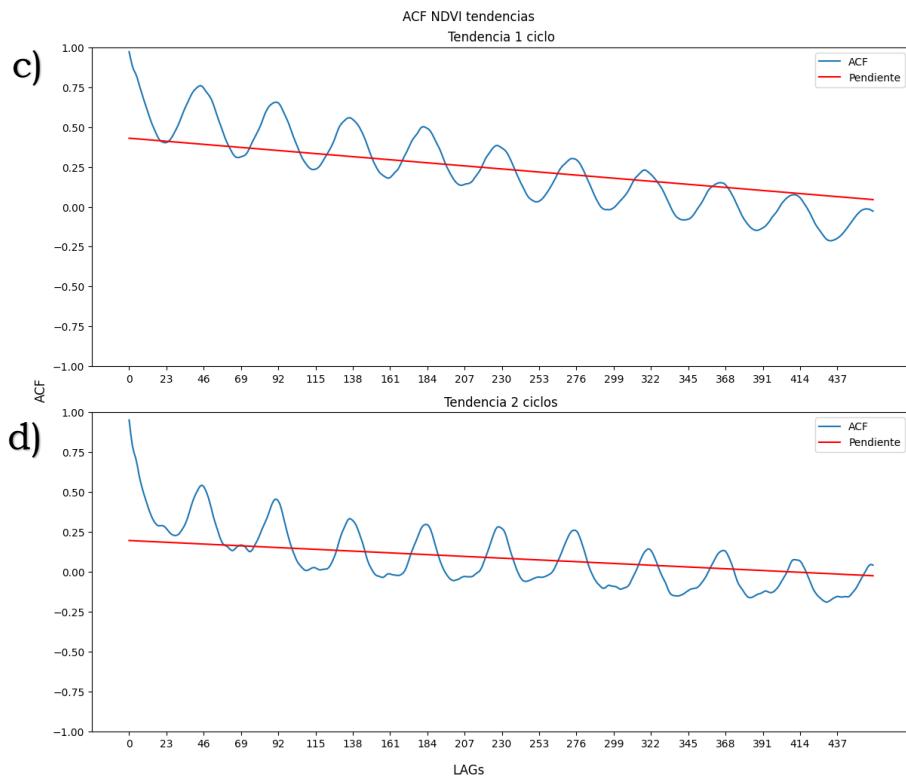


Ilustración 20: Funciones de autocorrelación con tendencia (azul), 1 ciclo (c, clase 1) y 2 ciclos (d, clase 3) y pendiente (rojo).

Las series de tiempo de pixeles de cambio se representan en la Ilustración 21. Se caracterizan por tener varios cambios, un cambio al final del periodo de estudio o no poder recuperar sus niveles y estabilidad después de una ruptura, marcadas en rojo en la serie. Las ondas representadas en la Ilustración 22 son funciones de autocorrelación con poca estabilidad, no pudiendo identificar las frecuencias dominantes asociadas al periodo de muestras del satélite y teniendo una pendiente es muy parecida a los pixeles de no cambio (clases 0 y 2).

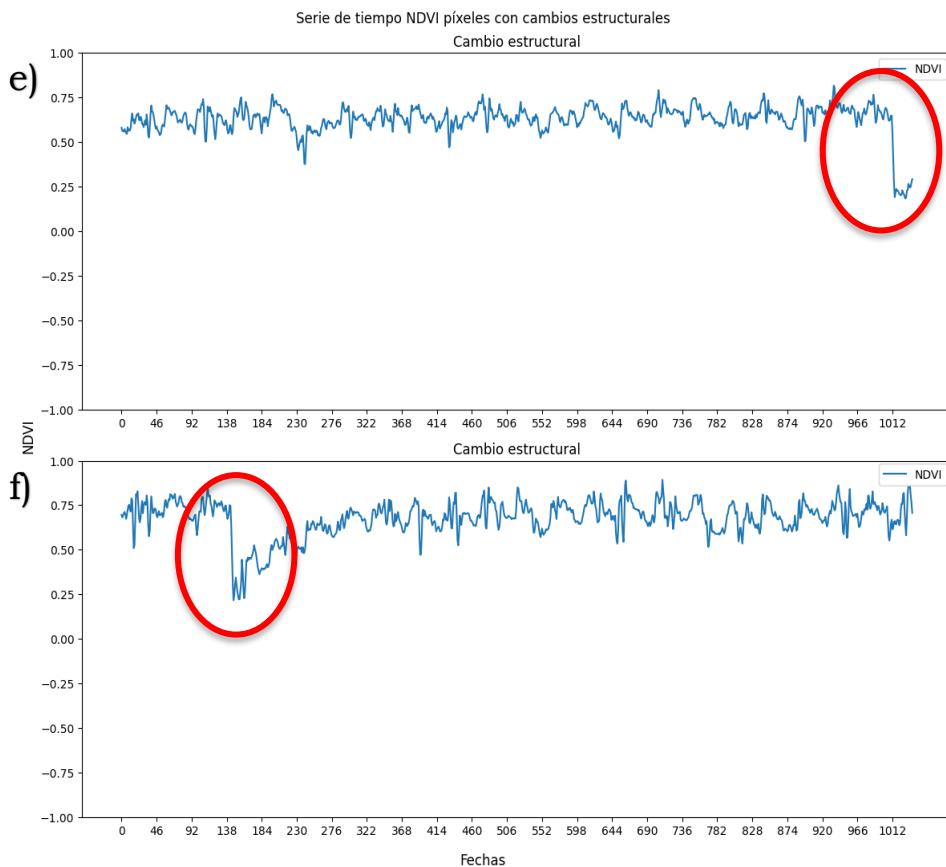


Ilustración 21: Serie de tiempo de píxeles con cambios estructurales (clase 4).

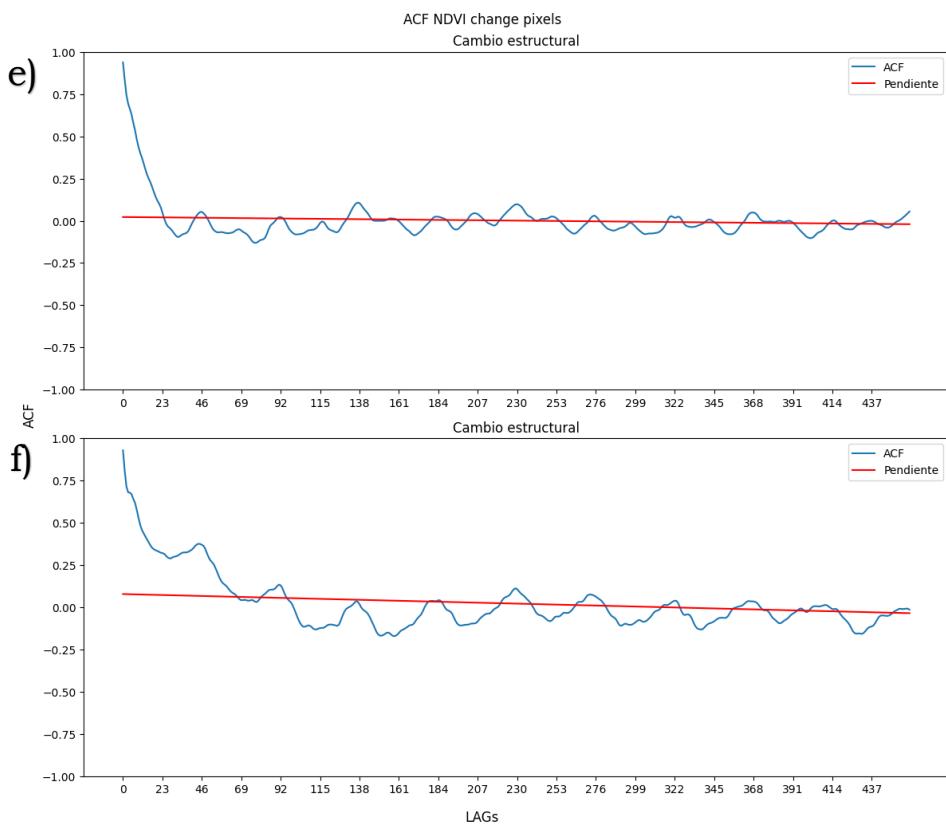


Ilustración 22: Funciones de autocorrelación con cambios (azul, clase 4) y pendiente (rojo).

Reuniendo y analizando los datos de las series de autocorrelación, se hizo un primer diseño en el que se usaban solo los datos que aparecen en la Tabla 2, donde podemos observar el tipo de píxel que es, la predicción, la pendiente y los porcentajes de valores positivos y negativos.

Tabla 2: Tabla de análisis de características.

Tipo	Tipo predicho	Pendiente	% positivos	% negativos
Sin cambio	Sin cambio, 1 ciclo	-0.59	48.91	50.87
Sin cambio	Sin cambio, 2 ciclos	-56.48	36.52	63.48
Tendencia	Tendencia 1 ciclo	-384.9	79.35	20.22
Tendencia	Tendencia 2 ciclos	-220.87	58.7	40.87
Cambio	Cambio	-42.45	34.35	62.61
Cambio	Cambio	-113.13	41.52	57.39

2.3.5 Aplicación de la Transformada de Fourier

La primera versión del diseño del algoritmo clasificó de forma correcta la mayoría de los píxeles, pero el resultado generaba ruido produciendo una sensación de sal y pimienta. Este ruido era producto de la fina línea en las condiciones de la primera versión de identificación cambios y no cambios, ya que en ciertos casos sus características se parecían en gran medida.

Para mitigar este problema, se optó por usar Fourier. Esta decisión no solo mejoró la forma de identificar cambios, si no que gracias a las frecuencias que se detectaban permitió la identificación de dobles ciclos y de tendencias de una forma más sencilla.

En la Ilustración 23, se puede observar el resultado de realizar la transformada de Fourier sobre la autocorrelación y normalizarla. A la izquierda de la imagen podemos ver las series de 1 ciclo detectándose el pico en el periodo 46, y a la derecha se detecta un segundo pico en el periodo 23 identificando el doble ciclo. Adicionalmente como se ha comentado las series con tendencias (c y d) tienen su máximo en la frecuencia 0, identificando la pendiente negativa de la serie, y finalmente las series de cambio (e y f) tal y como se hipotetizaban son conjuntos de frecuencias, sin detectar los períodos clave.

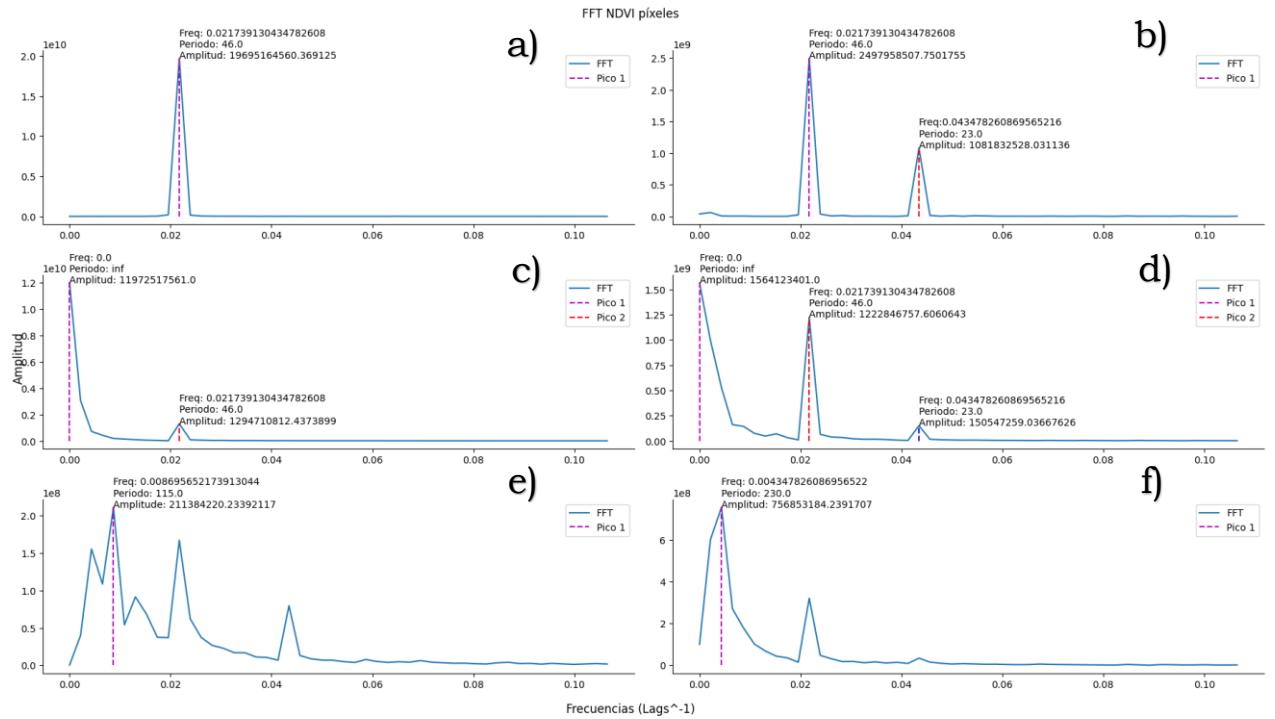


Ilustración 23: FFT sobre autocorrelaciones.

2.3.6 Clasificación de Píxeles

Para la evaluación de los píxeles, tras saber cómo se comportan las ondas en cada caso, y unificando las tres técnicas para analizar las ondas se diseña y plantea la estructura del algoritmo que se muestra en Ilustración 24.

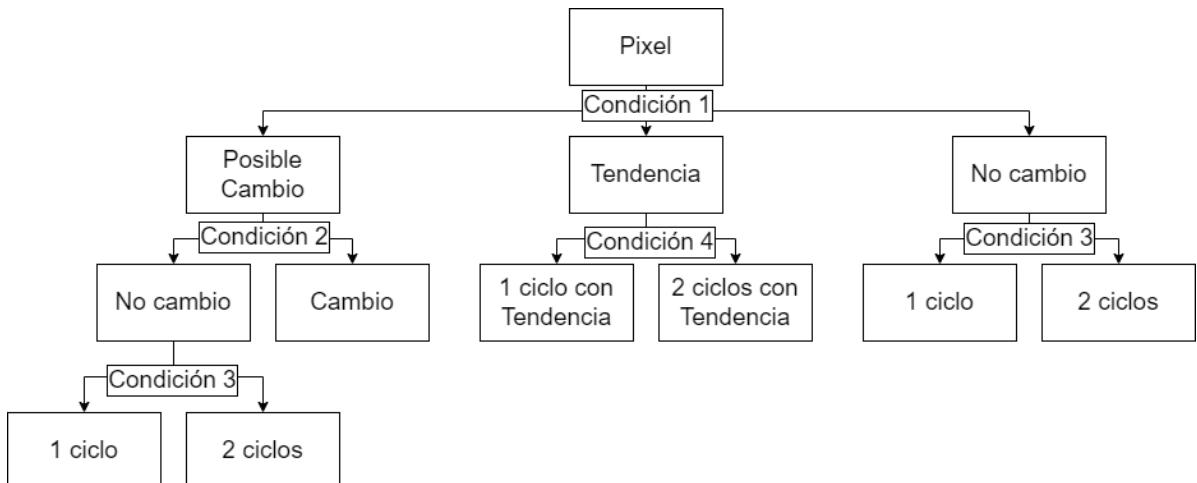


Ilustración 24: Árbol de decisión CCT.

A continuación, se explicará en qué consisten cada una de las condiciones:

Condición 1. Esta condición separa en 3 grupos los píxeles, posibles cambios, tendencias y no cambios. Para ello se usará la pendiente, el porcentaje de negativos y la transformada de Fourier. Si la pendiente es menor de -42 y tiene más de un 55% de valores negativos, se cataloga como posible cambio. Si el máximo dentro del espacio de frecuencias de Fourier se localiza en la posición 0 o la pendiente es inferior a -105 y tiene menos de un 55% de valores negativos se categoriza como tendencia. Si el píxel no ha entrado en ninguna de las dos condiciones se considera que es un píxel de no cambio.

Condición 2. En este caso, dentro de este posible cambio entran muchos píxeles que son de no cambio, para ello se realizará otro filtro para determinar si es un cambio o no. Este filtro se hará observando el espacio de frecuencias de Fourier, para ello se localiza el máximo de las frecuencias y se calcula el periodo (1/frecuencia), si este periodo coincide con el periodo de la serie de tiempo, en nuestro caso 46 al ser MODIS, podemos estimar que ese píxel no es un cambio.

Condición 3. Para identificar los píxeles de no cambio con un ciclo o con dos, volveremos a observar el espacio de frecuencias de Fourier. Si tiene un doble ciclo la serie, se podrá observar una segunda frecuencia dentro de la serie de Fourier, que su periodo coincide con la mitad del periodo anual, es decir, 23 y que este segundo pico de frecuencias supere un mínimo (800) para obtener los dobles ciclos más puros.

Condición 4. Al igual que los píxeles de no cambio las tendencias también pueden tener dobles ciclos, teniendo la misma condición de periodo que en los no cambios, pero en este caso el pico de frecuencia secundaria es más bajo al ser una tendencia por lo que el mínimo umbral se coloca a 150.

3 Resultados y conclusiones

El algoritmo CCT se incluye dentro de una aplicación (Satchange) que permite aplicar de forma cómoda y eficiente la metodología que se ha aplicado en esta investigación. En el marco de la Unión Europea financiado por el Fondo Europeo de Desarrollo Regional (FEDER) se llevó a cabo un proyecto en colaboración con la empresa Prosecar, cuyo objetivo principal era desarrollar una aplicación para cartografiar los usos del suelo, donde participé como único desarrollador.



Ilustración 25: Logo Satchange

En la Ilustración 26 se muestra el diagrama de clases de la aplicación pudiendo observar los distintos módulos que la componen. La leyenda de los elementos de la ilustración es la siguiente:

- Caja simple: archivos de una única funcionalidad.
- Caja de clase: contiene el nombre del archivo o de la clase que participa seguido de sus atributos y finalmente sus métodos junto con sus atributos si aplica.
- Caja discontinua: Directorio dentro del proyecto.
- Flecha: indica el uso entre clases/archivos.

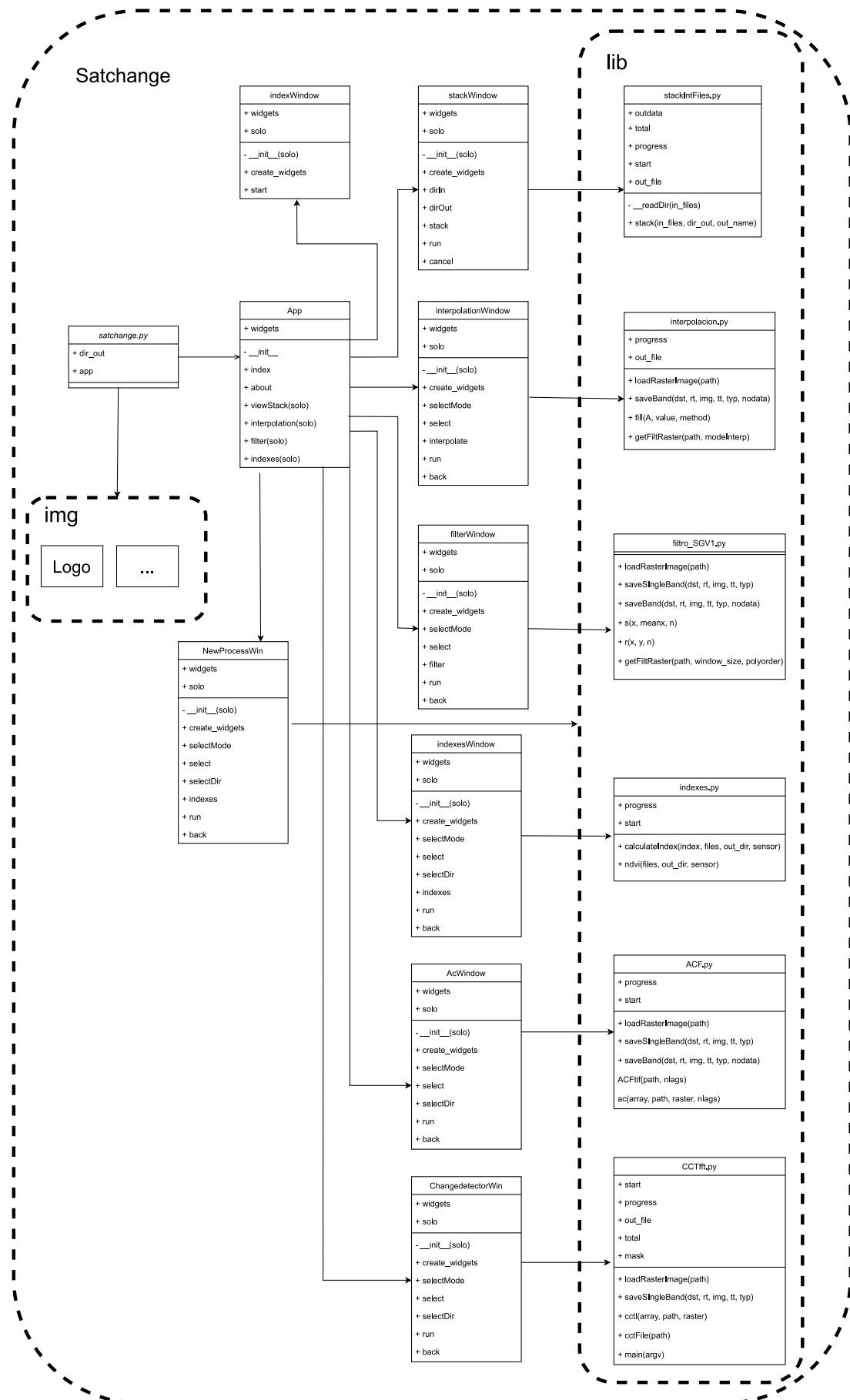


Ilustración 26: Diagrama de clases Satchange.

A continuación, en la Ilustración 27 se muestra el menú de la herramienta donde se pueden ver las distintas funcionalidades relacionadas con la metodología explicada. A continuación, se explica la función de cada botón en la interfaz:

- Stack: muestra la ventana de generación de series de tiempo.
- Interpolation: muestra la ventana para interpolar series de tiempo con datos sin valores (Sentinel-2), en este proyecto al trabajar con MODIS no se aplica.
- Autocorrelation: muestra la ventana para calcular la autocorrelación a una serie de tiempo.
- Filter: muestra la ventana de filtrado de series de tiempo donde se encuentra implementado entre otros el Savitzky-Golay.
- Indexes: muestra la ventana para calcular índices espectrales como el NDVI.
- Change detection: muestra la ventana para aplicar el algoritmo CCT a una serie de autocorrelación.

Cada una de las iteraciones de pueden hacer por separado, también pudiendo iniciar un proceso completo donde se unifican todas las iteraciones individuales en una única ejecución.

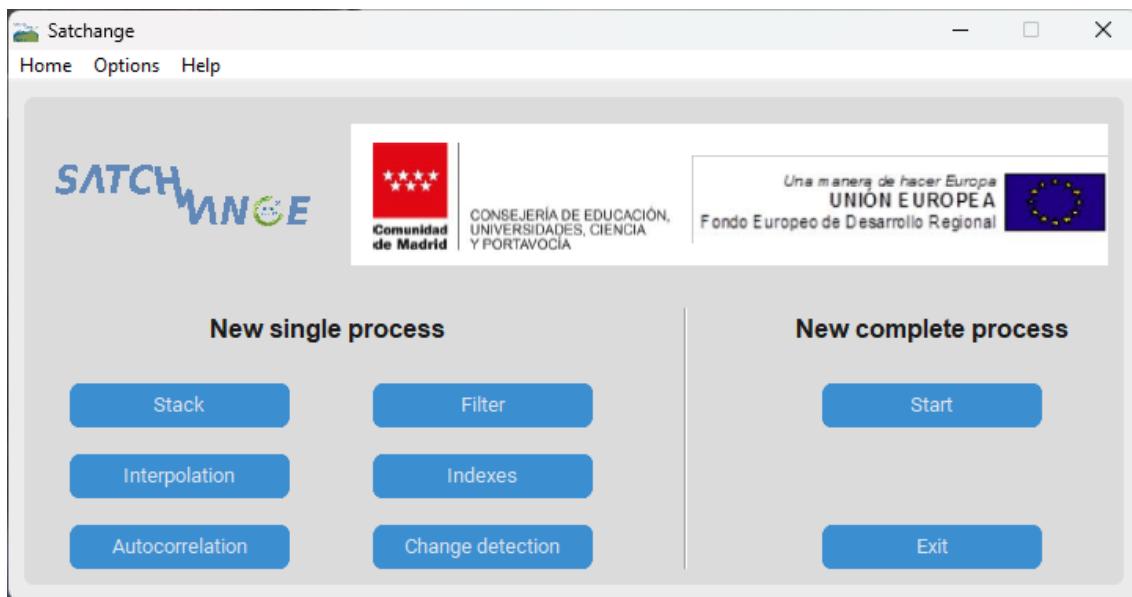


Ilustración 27: Menú principal de la aplicación.

3.1 Resultados

Después de depurar y completar el diseño final del algoritmo se integrando todas las propiedades exploradas en el documento. En la Ilustración 28 se ha conseguido caracterizar la superficie vegetal de toda la península Ibérica consiguiendo identificar el número de ciclos, tendencia o cambio estructural de cada píxel de la imagen, pudiéndose ver en mayor detalle en el anexo Imagen Clasificada.

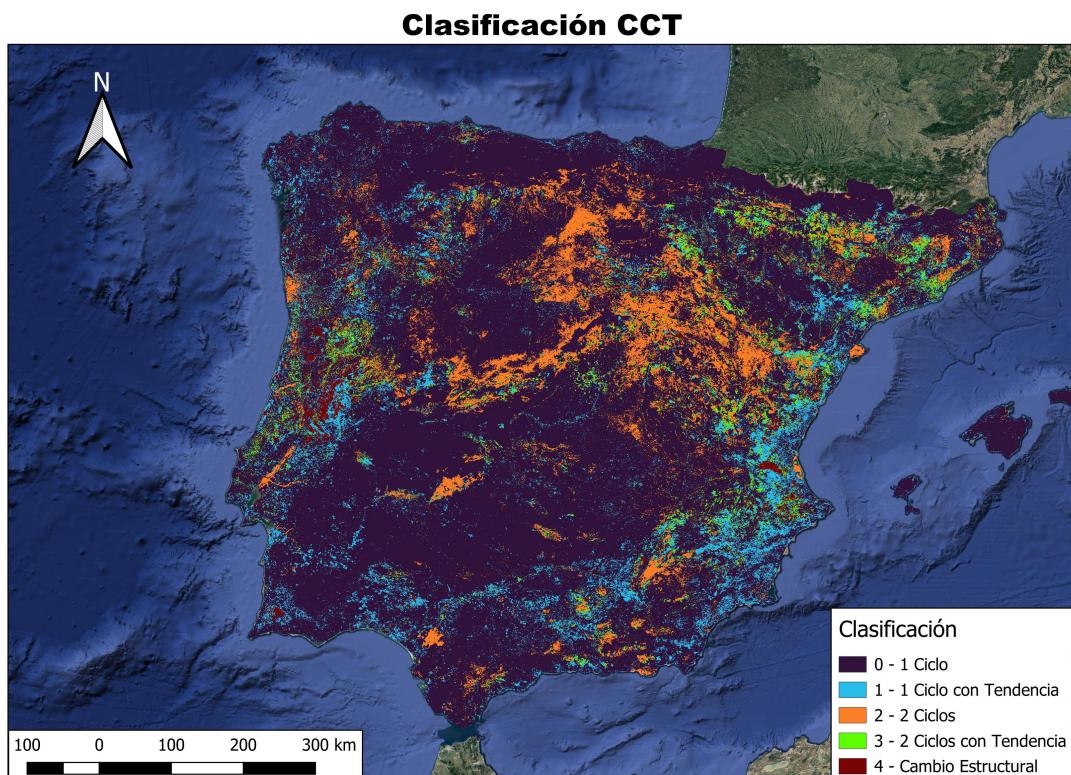


Ilustración 28: Mapa de clasificación de píxeles CCT.

En la Tabla 3 se ha hecho un resumen de la clasificación donde se observa que el 78% de los píxeles son de 1 ciclo, 9% de 2 ciclos, 12% son píxeles con tendencias y solo un 2% han sufrido cambios estructurales en la península Ibérica.

Tabla 3: Resumen de la clasificación CCT de la Península Ibérica

Clase	Valor píxel	N.º Píxeles	%
1 ciclo	0	8493232	78
1 ciclo con tendencia	1	935790	9
2 ciclos	2	938944	9
2 ciclos con tendencia	3	360664	3
Cambio estructural	4	224004	2

La Ilustración 29 muestra el sur de Portugal, se ha seleccionado esta zona debido a la gran variedad de clases que se observan, pudiendo mostrar de forma sencilla ejemplos de las series de los píxeles clasificados en función de su tipo.

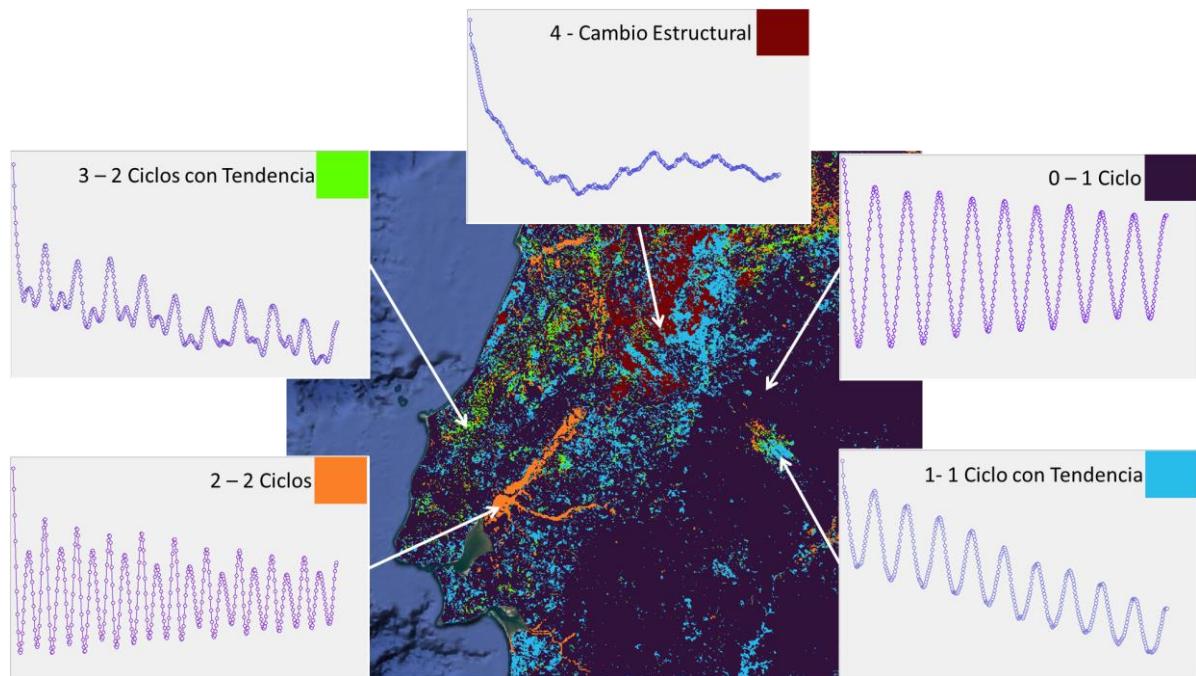


Ilustración 29: Zoom del sur de Portugal con ejemplos de las ACF respecto a la clasificación del CCT.

Tras observar el tiempo consumido a la hora de aplicar la metodología de análisis de series de tiempo (filtrado y cálculo de autocorrelación), llegando a tardar horas en ciertos procesos debido a la envergadura de los datos. El algoritmo consigue procesar casi 27 millones de píxeles multiplicado por el número de bandas (460) que contiene la autocorrelación correspondiente a los LAGs en aproximadamente 10 minutos.

3.2 Conclusiones

El objetivo principal del proyecto ha sido desarrollar un algoritmo que permita cartografiar en función de la dinámica de la vegetación. Este comportamiento ha hecho posible desarrollar el algoritmo gracias al análisis de series de tiempo junto a la función de autocorrelación, que ha permitido clasificar distintos tipos de píxeles.

Podemos observar en las siguientes imágenes como este algoritmo es capaz de ajustarse en gran medida a la clasificación propuesta. Por ejemplo, en la Ilustración 30 observamos el cambio estructural provocado por el incendio de Riba de Saelices, en 2005 Guadalajara.

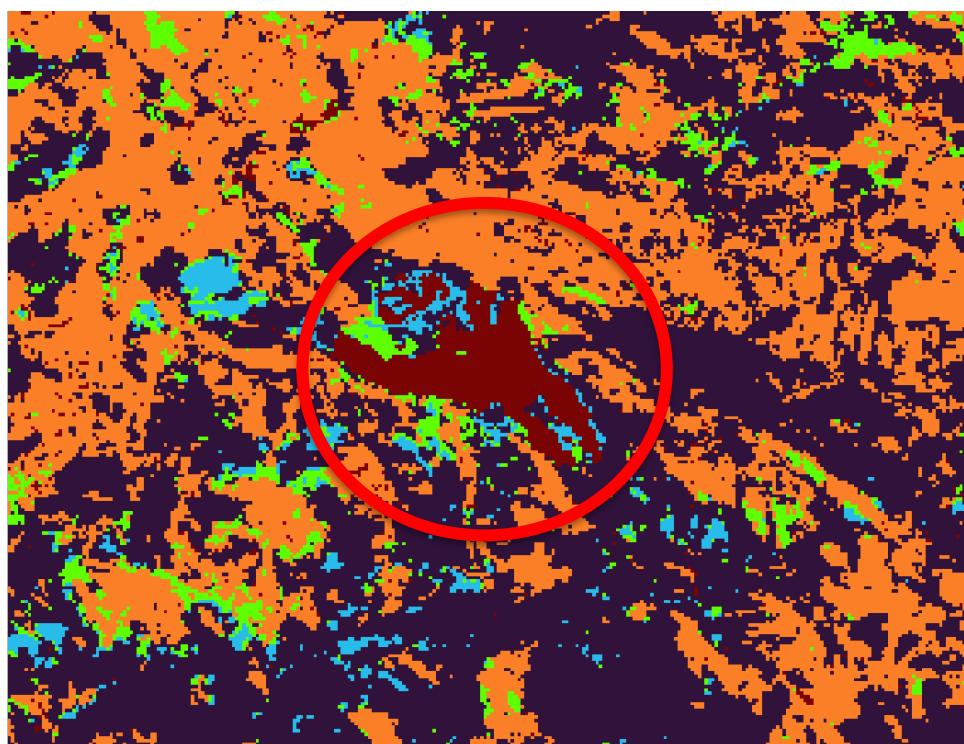


Ilustración 30: Incendio de Riba de Saelices.

Otros ejemplos que podemos encontrar para comprobar el ajuste de este algoritmo son la Dehesa Extremeña de la Ilustración 31 que presentan un ciclo sin tendencia, y en la Ilustración 32 que muestran los dobles ciclos ubicados en las zonas de cultivo en Castilla y León procedentes de las provincias de Burgos, Valladolid y Palencia.

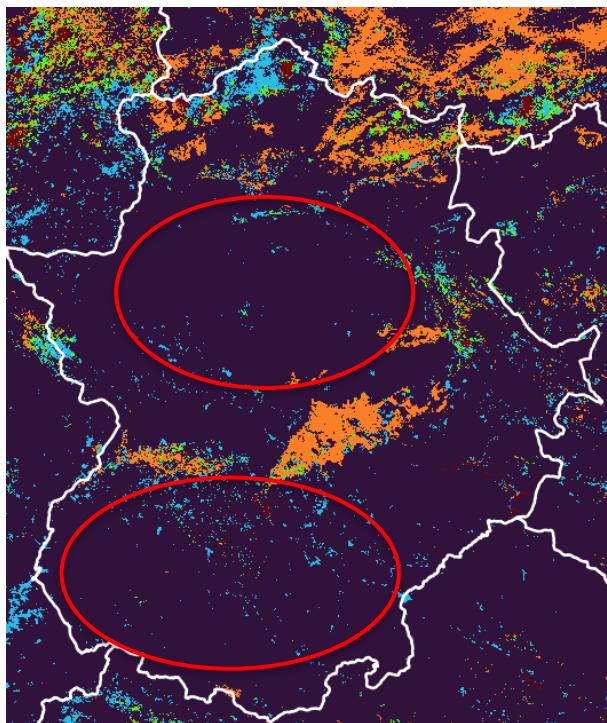


Ilustración 31: Dehesa Extremeña.

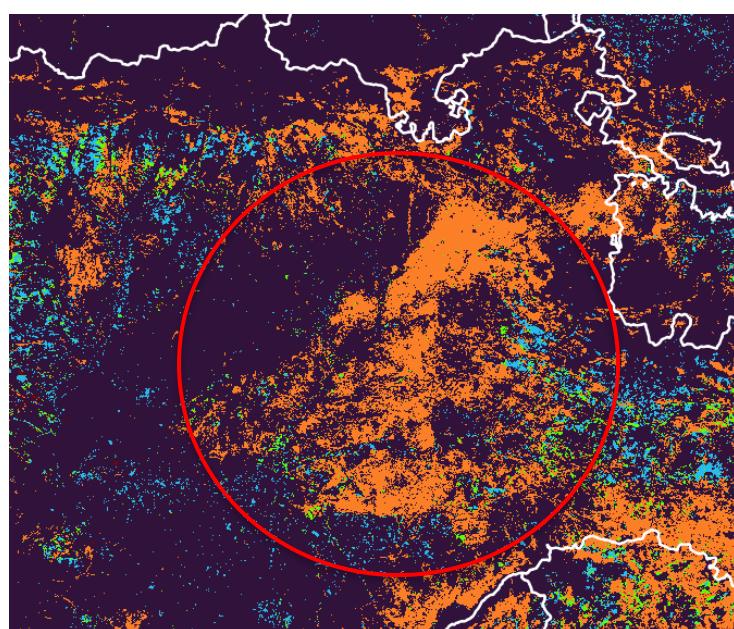


Ilustración 32: Zonas de cultivo de Castilla y León.

El algoritmo es útil para identificar distintas zonas de cultivo en función de sus ciclos, incendios e incluso servir como indicador en sistemas de alerta temprana para detectar zonas potencialmente combustibles debido aumento de la biomasa en zonas con tendencias.

En líneas futuras para la validación del algoritmo, habría que seleccionar una muestra aleatoria de píxeles y contrastar los resultados obtenidos por el algoritmo CCT, con algoritmos ya validados y comúnmente usados como el test de Man-Kendall o raíces unitarias. Debido a la duración del proyecto y complejidad de estos dos algoritmos no se ha realizado este proceso.

Adicionalmente se podrían implementar mejoras o ajustes en el algoritmo para mejorar la precisión.

4 Análisis de Impacto

Tras el desarrollo y analizando los resultados obtenidos, el análisis de impacto se puede enmarcar en las siguientes categorías, relacionándolo con los Objetivos de Desarrollo Sostenible:

- **Personal:**

A lo largo de los años ha sido necesaria la monitorización con personal en campo para analizar y cartografiar la superficie vegetal. Sin embargo, este sistema permite una actualización y análisis rápido de zonas de estudio, reduciendo la necesidad de salidas de campo a la hora de hacer un estudio de vegetación.

Relación con los ODS:

- **ODS 4: Educación de Calidad:** La reducción de salidas de campo permite asignar recursos a la mejora de otros aspectos educativos.
- **ODS 9: Industria, Innovación e Infraestructura:** La implementación de tecnologías de teledetección reduce la dependencia de métodos tradicionales, fomentando la innovación.

- **Empresarial:**

El uso de la herramienta no solo permite el monitoreo de la vegetación, también puede ser útil a la hora de detectar el crecimiento de zonas urbanizadas, desforestaciones o detección de incendios, siendo útil para grupos y empresas de observación de la Tierra.

Relación con los ODS:

- **ODS 9: Industria, Innovación e Infraestructura:** Utilización de herramientas tecnológicas para la detección de cambios en la vegetación y áreas urbanas.
- **ODS 11: Ciudades y Comunidades Sostenibles:** Contribución al monitoreo y gestión sostenible del crecimiento urbano.

- **Económico:**

Al reducir la cantidad de visitas al campo, se reducen el número de viajes necesarios ya sean mediante vehículos convencionales o mediante transporte aéreo, lo que produce un ahorro económico al prescindir de estos gastos y poder invertirlo en otros puntos de desarrollo de las empresas.

Relación con los ODS:

- **ODS 9: Industria, Innovación e Infraestructura:** Uso eficiente de herramientas tecnológicas para impulsar la innovación y la eficiencia económica.
- **ODS 12: Producción y Consumo Responsable:** La reducción de costos de viajes y transporte puede contribuir a un consumo responsable.

- **Medioambiental:**

Otra consecuencia de la reducción de visitas es el impacto medioambiental, ya que el monitoreo mediante teledetección es menos invasivo. Aunque se tenga que llevar el procesamiento en ordenadores que consumen mucha energía, el impacto de una visita a la zona de estudio conlleva un mayor impacto ambiental ya que hay que contar la contaminación emitida en el transporte y la alteración del medio al personarse físicamente.

Relación con los ODS:

- **ODS 13: Acción por el Clima:** Reducción de la huella de carbono al minimizar las visitas de campo.
- **ODS 15: Vida de Ecosistemas Terrestres:** Contribución a la conservación y monitoreo menos invasivo de la vegetación.

5 Bibliografía

- [1] J. A. Sobrino, *Teledetección*. Universitat de Valencia, 2001.
- [2] C. J. Tucker, «Red and photographic infrared linear combinations for monitoring vegetation», *Remote Sens Environ*, vol. 8, n.º 2, pp. 127-150, 1979, doi: [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0).
- [3] M. Mudelsee, «Trend analysis of climate time series: A review of methods», *Earth Sci Rev*, vol. 190, pp. 310-322, 2019, doi: <https://doi.org/10.1016/j.earscirev.2018.12.005>.
- [4] M. Jaramillo Ayerbe, D. E. González Gómez, M. E. Núñez Cabrera, G. E. Portilla, y J. H. Lucio García, «Análisis de series de tiempo univariante aplicando metodología de Box-Jenkins para la predicción de ozono en la ciudad de Cali, Colombia», *Revista Facultad de Ingeniería Universidad de Antioquia*, n.º 39, pp. 79-88, 2007.
- [5] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, y G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [6] L. Recuero *et al.*, «Fallowing temporal patterns assessment in rainfed agricultural areas based on NDVI time series autocorrelation values», *International Journal of Applied Earth Observation and Geoinformation*, vol. 82, p. 101890, 2019.
- [7] M. Guo, J. Li, H. He, J. Xu, y Y. Jin, «Detecting Global Vegetation Changes Using Mann-Kendal (MK) Trend Test for 1982–2015 Time Period», *Chin Geogr Sci*, vol. 28, n.º 6, pp. 907-919, dic. 2018, doi: 10.1007/s11769-018-1002-2.
- [8] J. Lee, «Testing for a unit root in time series with trend breaks», *J Macroecon*, vol. 18, n.º 3, pp. 503-519, 1996, doi: [https://doi.org/10.1016/S0164-0704\(96\)80034-6](https://doi.org/10.1016/S0164-0704(96)80034-6).
- [9] B. L. é Ibiza, *Regiones botánicas de la Península Ibérica*. Fortanet, 1895.
- [10] P. Muñoz Aguayo, «Apuntes de teledetección: índices de vegetación.», 2013.
- [11] G. Van Rossum y F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.
- [12] T. Kluyver *et al.*, «Jupyter Notebooks – a publishing format for reproducible computational workflows», en *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, F. Loizides y B. Schmidt, Eds., 2016, pp. 87-90.
- [13] QGIS Development Team, «QGIS Geographic Information System». 2023. [En línea]. Disponible en: <https://www.qgis.org>
- [14] GDAL/OGR contributors, «GDAL/OGR Geospatial Data Abstraction software Library». 2020. [En línea]. Disponible en: <https://gdal.org>
- [15] C. R. Harris *et al.*, «Array programming with NumPy», *Nature*, vol. 585, pp. 357-362, 2020, doi: 10.1038/s41586-020-2649-2.
- [16] L. Recuero, J. Litago, J. E. Pinzón, M. Huesca, M. C. Moyano, y A. Palacios-Orueta, «Mapping Periodic Patterns of Global Vegetation Based on Spectral Analysis of NDVI Time Series», *Remote Sens (Basel)*, vol. 11, n.º 21, 2019, doi: 10.3390/rs11212497.

- [17] Z. Cai, P. Jönsson, H. Jin, y L. Eklundh, «Performance of smoothing methods for reconstructing NDVI time-series and estimating vegetation phenology from MODIS data», *Remote Sens (Basel)*, vol. 9, n.º 12, p. 1271, 2017.
- [18] G. M. Jenkins, «Practical experiences with modelling and forecasting time series», *(No Title)*, 1979.
- [19] P. Bloomfield, *Fourier analysis of time series: an introduction*. John Wiley & Sons, 2004.
- [20] H. J. Nussbaumer y H. J. Nussbaumer, *The fast Fourier transform*. Springer, 1982.

6 Anexos

6.1 Costes de la investigación

A la hora de evaluar el coste del proyecto, lo dividiremos en dos secciones, coste de personal y de equipos informáticos necesarios.

- **Personal:**

Para saber el presupuesto del salario esperado por el desarrollo realizado, se tendrá en cuenta el tiempo consumido para completar el desarrollo. Para ello se analizará el diagrama de Gantt (Ilustración 33) actualizado para obtener el tiempo consumido en el desarrollo.

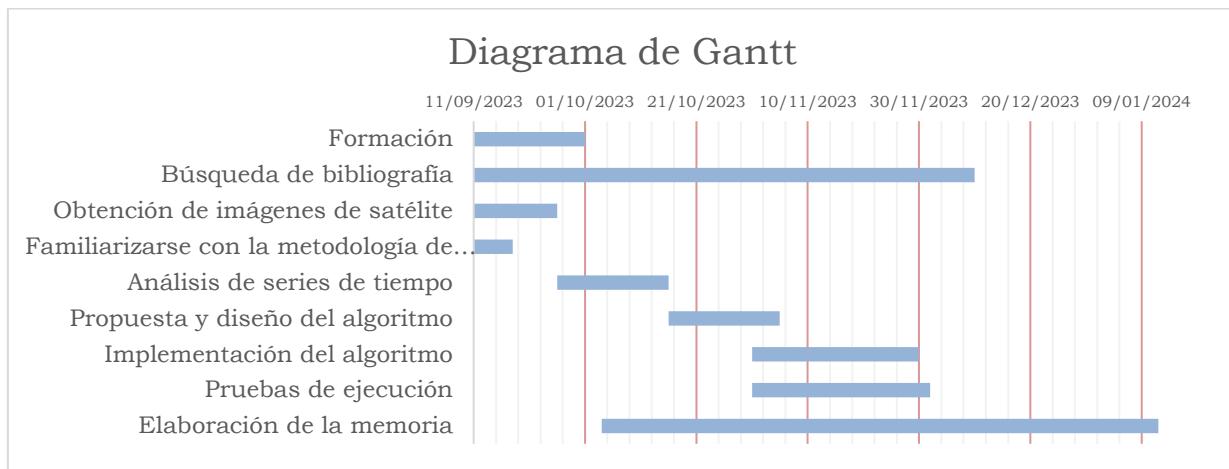


Ilustración 33: Diagrama de Gantt.

En la Tabla 4 se muestran las tareas del diagrama en formato tabular para obtener el número de días totales de trabajo, en este caso han salido 123 días.

Tabla 4: Distribución de tareas respecto a su duración.

Actividad	Fecha inicio	Duración (días)	Fecha fin
Formación	11/09/2023	20	01/10/2023
Búsqueda de bibliografía	11/09/2023	90	10/12/2023
Obtención de imágenes de satélite	11/09/2023	15	26/09/2023
Familiarizarse con la metodología de análisis de series de tiempo	11/09/2023	7	18/09/2023
Análisis de series de tiempo	26/09/2023	20	16/10/2023
Propuesta y diseño del algoritmo	16/10/2023	20	31/10/2023
Implementación del algoritmo	31/10/2023	30	15/11/2023
Pruebas de ejecución	31/10/2023	32	07/11/2023
Elaboración de la memoria	04/10/2023	100	12/01/2024
Total de días			123

Según la página web Talent.com el sueldo de un ingeniero informático junior (<https://es.talent.com/salary?job=ingeniero+informático+junior>), es de 12.67€ a la hora. El número de horas en un puesto de trabajo de jornada completa son 8 horas. Teniendo en cuenta todos los parámetros que se han comentado, el cálculo del salario por la realización de este proyecto sería; $12.67 * 8 = 101.36\text{€}$, correspondiente al día laboral y $101.36 * 123 = 12467.28\text{€}$ como salario total.

- **Equipos informáticos**

El análisis de series de tiempo a la escala realizada, no se puede realizar en un equipo con características comunes. Al trabajar con datos de decenas de gigabytes, se necesita un equipo con la capacidad de cálculo y de memoria suficientes para no ralentizar o imposibilitar el proceso. Por esta razón se montó un ordenador con componentes a medida que fuese capaz de dar las mejores prestaciones en el procesamiento.

Tabla 5: Lista de componentes del equipo informático utilizado.

Parte del equipo	Modelo	Precio (€)
Procesador	Intel core i9 13900KF	639.00
Refrigeración	MSI MAG CoreLiquid C240 ARGB Kit Refrigeración Líquida	139.98
Placa base	Asus prime Z790-P	300.00
Memoria RAM	Corsair VENGACE DDR5 5200MHz 64GB 2x32GB	275.99
Tarjeta gráfica	MSI RTX 3060 VENTUS 2X OC 12GB	379.89
Disco principal	Samsung 980 SSD 1TB PCIe 3.0 NVMe M.2	98.40
Disco secundario	Samsung 870 EVO SSD 2.5" 2TB SATA3 Negro	175.99
Fuente de alimentación	EVGA SuperNOVA 750 G6 750W 80 Plus Gold Full Modular	134.99
Caja	Corsair 110Q USB 3.0 Negra	99.90
Total		2243.14

La combinación de las dos partes que influyen en los costes, obtenemos los costes totales que hacen la suma de 14710.42€.

6.2 Imagen Clasificada

Clasificación CCT

