



UNIVERSITAT JAUME I
Escola Superior de Tecnologia i
Ciències Experimentals

MOSAICADO MULTI-TEMPORAL PARA
PRODUCTOS L2 DE SENTINEL 3/FLEX

TFM Máster Sistemas Inteligentes

Autor: Damián Ibáñez Fernández

Tutor: Filiberto Pla Bañón

Cotutor: Rubén Fernández Beltrán

Fecha: Castelló de la Plana, septiembre 2020

Índice

Introducción.....	1
Remote Sensing	2
2.1. ¿Qué es la Remote Sensing?	2
2.2. Misión Sentinel-3	4
2.3. Misión FLEX.....	6
Mosaicado multitemporal de imágenes.....	7
3.1. WELD	7
3.2. Median NIR composite	8
3.3. Medoide	8
3.4. Sen2Three.....	9
3.5. S2GM.....	10
Requerimientos para el software.....	11
Descripción del sistema	12
5.1. Datos de entrada	13
5.2. Metodología	15
5.2.1. Adquisición operacional de datos.....	15
5.2.2. Mosaicado espacial para productos OLCI L2	16
5.2.3. Composición temporal de productos OLCI L2.....	17
5.3. Adaptación para FLEX.....	23
Experimentación	25

6.1. Conjunto de datos.....	25
6.2. Proceso experimental.....	27
6.3. Resultados.....	27
6.3.1. Experimento 1	27
6.3.2. Experimento 2.....	32
Planificación	36
7.1. Planificación inicial	36
7.2. Planificación final	37
Conclusión.....	39
Anexos	40
9.1. Referencias.....	40

Capítulo 1

Introducción

El aumento de datos en el campo de Remote Sensing en los últimos años está abriendo el paso a retos de estudios y análisis globales que antes éramos incapaces de afrontar debido a la falta de información disponible. Debido a esto, las técnicas existentes de mosaicado multitemporal se limitaban a producir composiciones de imágenes espectrales sin considerar características biofísicas de alto nivel como los que se obtienen a través de misiones como Sentinel-3 (S3) o la futura FLuorence EXplorer (FLEX).

Este trabajo tiene como objetivo desarrollar un algoritmo de mosaicado multitemporal para productos derivados de S3, y estudiar el futuro uso de este para la misión FLEX. Concretamente se pretende diseñar una nueva metodología operacional para producir mosaicos multitemporales de productos derivados de forma automática, facilitando así el procesado de productos biofísicos de alto nivel para un día concreto, de forma semanal, mensual, estacional o anual. Es decir, automatizar todo el proceso desde la adquisición de datos hasta la obtención de los mosaicos multitemporales y el cálculo de confianza de estos. Esta tarea se ha realizado como parte del proyecto RTI2018-098651-B-C54 financiado por el Ministerio de Ciencia, Innovación y Universidades, de productos avanzados L3 y L4 para la misión FLEX-S3, y se ha publicado un artículo científico sobre ello [51].

Este algoritmo ha sido diseñado principalmente en Matlab, aunque para el proceso completo también se han utilizado otras herramientas como Copernicus Open Access Hub [42] o SNAP [40]. Para conseguir los mosaicados multitemporales de productos derivados S3 ha sido necesario integrar técnicas como registro de imagen, modificación y proyección de coordenadas, filtrado de datos, mosaicado espacial y composición temporal junto con el correcto uso y entendimiento de la información obtenida de estos productos.

Además, se ha realizado un estudio experimental con diferentes colecciones de productos S3 para mostrar los resultados obtenidos y comparar las ventajas que este nuevo método presenta con respecto a otras alternativas de mosaicado tradicionales.

Capítulo 2

Remote Sensing

2.1. ¿Qué es la Remote Sensing?

Remote Sensing o teledetección es el campo que abarca todo lo referente a la obtención de información de un objeto o un fenómeno sin realizar contacto físico directo con este. Remote Sensing por lo tanto es un campo amplio y sus técnicas se utilizan en diferentes áreas como la geografía, agrimensura, meteorología, ecología y el resto de las ciencias de la tierra, aunque no se limita a ellas ya que también tiene usos en aplicaciones militares, comerciales, o económicas.

Dependiendo del tipo de información que se desea capturar se utilizan diferentes medios. Para imágenes de muy alta resolución, cubriendo áreas pequeñas se utilizan UAVs (vehículos aéreos no tripulados) y drones. Cuando se necesitan imágenes de alta resolución, pero con una mayor cobertura se utilizan aviones o helicópteros. Por último, cuando se necesitan cubrir áreas muy amplias se usan satélites, de baja, media o alta órbita terrestre en los que encontramos resoluciones de todo tipo.

No obstante, cuando hablamos de Remote Sensing generalmente nos referimos a la información que se obtiene a través de los diferentes sensores que utilizan estos satélites para obtener datos de la tierra, desde la atmosfera y los océanos a ciudades o zonas terrestres ya que ofrecen una mayor información espacial. Mientras tanto los UAVs, helicópteros y demás se utilizan para resolver problemas locales.

Podemos dividir los sensores que se utilizan en Remote Sensing para obtener la información remota en dos grupos: los activos que emiten algún tipo de señal y la información que se obtiene proviene de la detección de esta señal reflejada como sensores RADAR o LiDAR que utilizan ondas de radio o luz como se muestra en la [Figura 1](#); y los pasivos que obtienen la información de radiación

emitida o reflejada naturalmente por el objeto, siendo la fuente más común la luz solar, como en la [Figura 2](#).

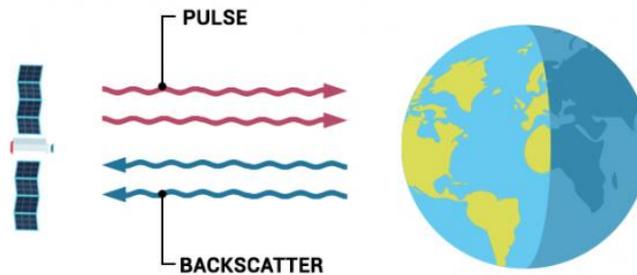


Figura 1. Sensores de Remote Sensing activos.

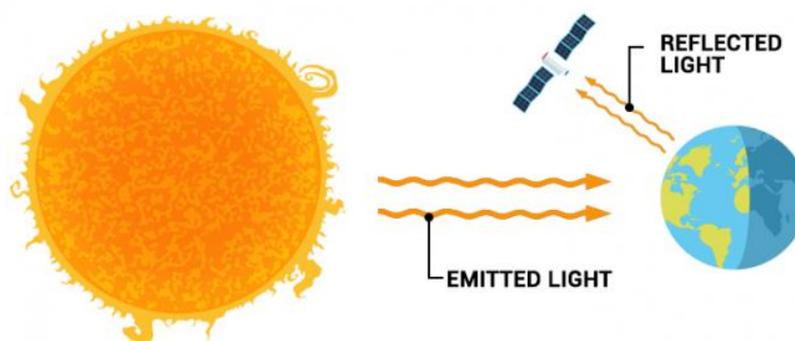


Figura 2. Sensores de Remote Sensing pasivos.

Pese a que en Remote Sensing se utilizan muchos tipos de datos, los más comunes son imágenes. Sin embargo, estas imágenes contienen mucha más información que las que acostumbramos a obtener con cámaras normales, tanto en resolución espacial como espectral. Al aumentar el área que se cubre se suele perder resolución espacial, pero hay satélites capaces de obtener resoluciones de 50 cm por píxel. En cuanto a la resolución espectral muchos de estos datos son imágenes multiespectrales, es decir que contiene varias bandas que contienen información que el sensor es capaz de captar en un conjunto de longitudes de onda; o hiperespectrales, que contiene bandas con muchas menos longitudes de onda, pero puede contener cientos o miles de bandas.

Los otros tipos de datos comunes que se utilizan son los obtenidos a través de sensores como LiDAR o RADAR para construir modelos topográficos. Los LiDAR se utilizan comúnmente para tierra, bajo agua se utilizan sonares u otro tipo de radiómetros o espectrómetros porque los LiDAR utilizan láseres de luz, la cual se absorbe o difracta en medios acuáticos complicando las mediciones.

2.2. Misión Sentinel-3

La misión Sentinel-3 forma parte del Programa Copernicus de la Agencia Espacial Europea para lanzar varios satélites de observación de la tierra. Actualmente la misión S3 consta de dos satélites Sentinel-3A y Sentinel-3B, aunque otros satélites están en futuro desarrollo. El programa Copernicus tiene el objetivo de obtener a través de estas misiones datos e información precisa para ayudar a gestionar, comprender y reducir los efectos del cambio climático.

Concretamente, la misión S3 tiene como objetivo medir la topografía, temperatura y color de la superficie del mar y la tierra con precisión, monitorización y modelado de clima, mapeado de superficies de bosques y detección de fuegos entre otros. Los datos que proporcionan estos satélites son prácticamente en tiempo real, con un ciclo orbital de 27 días. Lo que significa que tarda ese tiempo en volver a pasar sobre un mismo punto geográfico en la tierra, sin embargo, durante esta orbita los satélites pasan varias veces por puntos cercanos por lo que se pueden obtener productos de una zona prácticamente cada día.

Los productos obtenidos a través de S3 son públicos y se puede acceder a ellos a través de sitios como Copernicus Open Access Hub [43] que proporciona datos de este y otros satélites como otras misiones Sentinel para poder analizarlas de forma organizada.

Estos datos, por lo tanto, vienen proporcionados por una serie de sensores, que se encargan de cumplir los propósitos mencionados anteriormente. Podemos observar su disposición también en la [Figura 3](#). Los instrumentos principales son los siguientes:

- **SLSTR** (Sea and Land Surface Temperature Radiometer). Su objetivo principal es medir la temperatura de la superficie del mar a través de nueve canales espectrales y dos bandas adicionales para la monitorización de fuegos. Estos canales cubren el espectro visible e infrarrojo cercano, así como el infrarrojo de onda corta.

- **OLCI** (Ocean and Land Colour Instrument). Este sensor mide la radiación solar reflejada por la tierra, con una resolución de 300 metros por píxel, tiene 21 bandas espectrales desde los 400 nm hasta los 1020 nm. Su objetivo principal es obtener información de la superficie terrestre y oceánica relacionada con la biología con información como la actividad clorofílica de las plantas o la absorción de vapor de agua.
- **SRAL** (Synthetic Aperture Radar Altimeter). Este instrumento es el que se encarga de las mediciones topográficas precisas en conjunto con DORIS, MWR y LRR.
- **DORIS** (Doppler Orbitography and Radiopositioning Integrated by Satellite) es un receptor del posicionamiento en la órbita.
- **MWR** (Microwave Radiometer). Este sensor tiene como función principal medir la cantidad de vapor de agua, nubes y la radiación térmica emitida por la tierra.
- **LRR** (Laser Retroreflector). Este láser se encarga de localizar el satélite en la órbita.
- **GNSS** (Global Navigation Satellite System) realiza la determinación de la órbita y permite el seguimiento de varios satélites a la vez.



Figura 3. Sensores de Sentinel-3.

2.3. Misión FLEX

La misión FLEX (FLuorecence EXplorer) es el octavo satélite de exploración de la superficie terrestre de la Agencia Espacial Europea del programa Copernicus, que se encuentra actualmente en desarrollo. El objetivo principal de esta misión es obtener información de la actividad fotosintética de la superficie terrestre, medida a través de la fluorescencia que produce la clorofila durante la fotosíntesis. Está planificado su lanzamiento para el año 2024.

La fluorescencia inducida por el sol (SIF) es uno de los mecanismos de los cloroplastos para disipar el exceso de radiación debido a la fotosíntesis. En este proceso se emite radiación en unos espectros muy concretos, con unos picos en 690 y 740 nm por lo que es detectable y medible. Esta información tiene utilidad para valorar el estado de la vegetación, analizar el ciclo del carbono y su absorción, y el estudio y monitorización de bosques y campos de cultivo. Para cumplir este objetivo FLEX llevará el sensor FLORIS (Fluorescence Imaging Spectrometer) que obtendrá datos del rango espectral de 500 a 780 nm, con un muestreo de tan solo 0.1 nm y una resolución espacial de 300m al igual que los datos de S3.

FLEX se relaciona estrechamente con S3, ya que cuando se lance este orbitará en tándem delante de uno de los dos satélites de S3 para tomar imágenes del mismo punto geográfico con una diferencia temporal de tan solo unos pocos segundos. Esto permitirá que ambos satélites complementen sus mediciones e información.

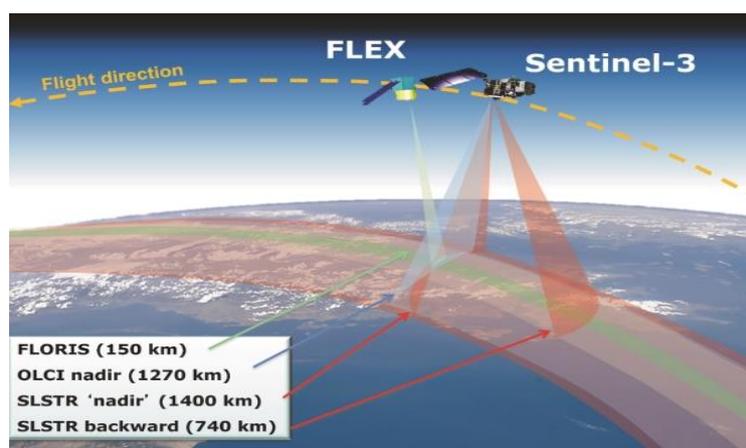


Figura 4. FLEX y S3 en tándem, mostrando la superposición de los datos detectados por sus sensores.

Capítulo 3

Mosaicado multitemporal de imágenes

El mosaicado espacial de imágenes es el proceso por el cual se construye una sola imagen a través de varias imágenes que se superponen en algún punto. Con esto se consigue una imagen de mayor tamaño que contiene más información de la escena que una sola imagen. De una forma similar, el mosaicado multitemporal es un proceso que pretende obtener imágenes más completas o precisas de una escena utilizando diferentes muestras a lo largo de un periodo de tiempo. En el área de Remote Sensing son dos técnicas de gran utilidad, permitiéndonos ver zonas más amplias de la tierra o datos más robustos.

Este trabajo se enfoca en generar mosaicos multitemporales, sin embargo, se han desarrollado diversos algoritmos de este tipo para otros satélites y tipos de datos en Remote Sensing. A continuación, se describirán brevemente los más relevantes.

3.1. WELD

Para las imágenes del satélite Landsat, el algoritmo WELD [29] se diseñó con el objetivo de seleccionar los píxeles terrestres con la menor cantidad de nubes, nieve u otros tipos de contaminación atmosférica. WELD utiliza conjuntos de imágenes de reflectancia mensuales, estacionales y anuales para generar mosaicos geofísicos y biofísicos consistentes para realizar valoraciones regionales de las dinámicas de cobertura del terreno.

Para estas composiciones se utiliza un árbol de decisiones que considera: el NDVI (Índice de Vegetación de Diferencia Normalizada), la temperatura del brillo, la temperatura superficial máxima aparente, el ángulo mínimo de escaneo, la máxima diferencia entre reflectancia roja y NIR (Infrarrojo cercano) y combinaciones de estos para comparar varios píxeles y obtener el mejor representante.

En la [Ecuación 1](#) se muestra la expresión correspondiente al NDVI, donde ρ_{nir} representa la reflectancia de la banda NIR y ρ_{red} la de la banda roja. El mayor problema de este método es que no es muy robusto si el espectro no satisface el control de calidad en la clasificación de datos previa al árbol de decisiones.

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$$

Ecuación 1. Índice NDVI.

3.2. Median NIR composite

Existen otros métodos más estadísticos, como el Median NIR composite que se ha utilizado para diferentes estudios regionales utilizando datos de Landsat en Rusia [30] y en la República Democrática del Congo [35]. En detalle, este método es una evolución del método presentado por Hansen [36], que utiliza datos filtrados con una clasificación entre agua, nubes, sombras y otras clases similares de un conjunto de imágenes entre las que se elige la mejor muestra para realizar la composición. Se considera como la mejor muestra el valor de la mediana NIR. Este método ha mostrado ventajas para series temporales de datos con no muy buena resolución espacial comparado con la técnica del máximo NDVI, además, probó producir resultados con muy poco ruido en áreas boscosas.

3.3. Medioide

Otro enfoque distinto es el propuesto por Flood en [37] utilizando la función del medioide. El medioide es el objeto representativo de un conjunto de datos del cual la disimilitud media a todos los objetos del conjunto de datos es mínima [38]. El medioide aplicado a reflectancia se estima calculando la distancia entre todos los píxeles de reflectancia en cada una de las bandas y minimizando su suma. El medioide ha mostrado ser mejor que la media en producir imágenes que sean representativas de un periodo de tiempo.

Este método se aplicó también a datos de Landsat TM/ETM+ para crear imágenes estacionales de reflectancia. Un análisis de los valores de reflectancia estacionales sugiere que esta composición es más representativa que la hecha mediante el valor máximo de NDVI. La función del mediodide se muestra en la [Ecuación 2](#), donde se define como el punto x_i en el conjunto x que minimiza la suma de las distancias entre este punto y todos los otros puntos x_j en el conjunto. composición.

$$\text{Medioide}(x) = \underset{x_i \in X}{\text{arg min}} \sum_{x_j \in X} ||x_j - x_i||$$

Ecuación 2. Expresión del medioide.

3.4. Sen2Three

En cuanto a las misiones de Sentinel, en el caso de Sentinel-2 (S2) el algoritmo Sen2Three [32] genera productos sintéticos espacio-temporales corregidos atmosféricamente S2 L2 por el algoritmo de corrección atmosférica Sen2Cor [39]. Este método cambia los píxeles que se consideran contaminados por otros a lo largo de una serie temporal siguiendo tres criterios: las adquisiciones más nuevas tienen más valor, una categorización según la suma de buenos píxeles, y otra según la media de la densidad de aerosol óptico o la media del ángulo solar cenital. En la [Figura 5](#) podemos observar el funcionamiento de este algoritmo.

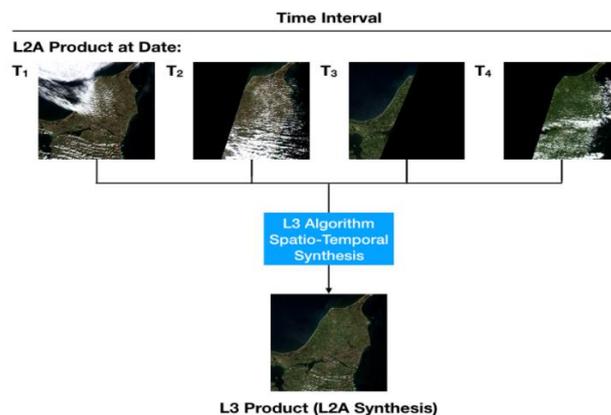


Figura 5. Ejemplo de funcionamiento de Sen2Trhee.

3.5. S2GM

Otro método de composición temporal es el propuesto para el proyecto S2 Global Mosaics (S2GM) [33]. Este fue diseñado para generar mosaicos de productos L2 de reflectancia de la superficie utilizando Sentinel-2. Este algoritmo toma una serie temporal y píxel por píxel dependiendo del número de muestras válidas:

- Utiliza el algoritmo del mediodo.
- Utiliza una variación de WELD adaptada para productos S2 con una clasificación llamada Short Term Composite para S2 (STC-S2).

En más detalle, S2GM realiza un submuestreo espacial con una interpolación del vecino más próximo para observaciones de diferente resolución y realiza el mosaicado espacial utilizando herramientas de Sentinel Application Platform (SNAP) [40].

Capítulo 4

Requerimientos para el software

El objetivo final de este proyecto es crear una metodología mediante la cual obtener mosaicos multitemporales de productos L2 de Sentinel-3. Estos son productos filtrados y derivados de los productos L1, que a su vez se obtienen de los datos en crudo de los sensores de Sentinel-3. Los productos L2 son interesantes ya que contienen información de mayor nivel biofísica. Para cumplir este objetivo principal el algoritmo debe ser capaz de realizar una serie de procesados de datos y cumplir unos requerimientos para que el resultado final sea el deseado. Las funcionalidades que se buscan conseguir con el diseño de este algoritmo son las siguientes:

- Obtención de los datos de forma automática.
- Realizar un mosaicado espacial de los productos L2 S3.
- Realizar un mosaicado o composición temporal de los productos L2 S3.
- Obtener una medida de confianza para los resultados.

Otros requerimientos clave fueron que la exigencia de procesamiento del programa fuese la menor posible, pero con un modelo robusto, o que el código fuese fácilmente interpretable e intuitivo para futuros cambios y adaptaciones, por ejemplo, a datos de FLEX.

Además del diseño de un algoritmo que cumpliera estas tareas se realizó una experimentación con conjuntos de datos de diferentes espacios de tiempo, así como una comparación con otros métodos utilizados para este propósito.

Capítulo 5

Descripción del sistema

En esta sección se presenta la parte metodológica del trabajo que ha sido diseñada para generar mosaicos multitemporales de OLCI. Primero se definirá el tipo de datos de entrada que utilizará el sistema, y a continuación se describirá la metodología. Finalmente mencionaremos los cambios necesarios para utilizar este algoritmo con FLEX. Esta se puede dividir en cuatro módulos secuenciales como podemos observar en [Figura 6](#) :

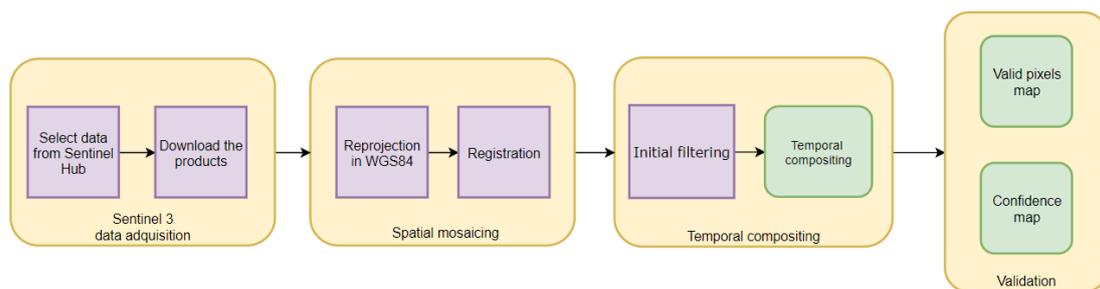


Figura 6. Metodología propuesta.

1. La adquisición operacional de los datos, dónde los datos originales se obtienen y descargan a través de Copernicus Open Hub.
2. El mosaicado espacial y la reorganización de los productos descargados, realizando una reproyección a coordenadas World Geodetic System 1984 (WGS84) y un proceso de registrado espacial.
3. La composición temporal tras filtrar los datos.
4. Una validación final de los mosaicos de productos multitemporales mostrando el número de píxeles y su correspondiente nivel de confianza.

5.1. Datos de entrada

Este algoritmo se ha implementado para mosaicar temporal y espacialmente productos de Sentinel-3. Concretamente, productos terrestres L2, aunque podría ser adaptado a productos marinos o de aguas continentales modificando los datos de entrada y la toma de decisiones explicada más adelante.

Los productos biofísicos L2 elegidos para realizar el mosaicado han sido los del índice OTCI. El índice OTCI es un índice de la clorofila terrestre de las plantas obtenido por el sensor OLCI de reflectancia, específicamente de las bandas centradas en 754 nm, 709 nm y 681 nm. Este índice es clave para modelar la productividad terrestre, intercambios de gas y la salud de la vegetación en general, además de estar relacionado con los productos que ofrecerá FLEX. Sin embargo, este algoritmo debería dar buenos resultados con los otros productos terrestres ofrecidos por S3 mostrados en la [Tabla 1](#).

Product Name	Parameter Name
Marine and inland waters products	
OLCI water-leaving reflectance (16 bands)	Rxxx
OLCI water inherent optical properties	
-CDM absorption coefficient	ADG443
OLCI Ocean Colour products	
-Algal pigment concentration	CHL
-Total Suspended Matter concentration	TSM
-Diffuse attenuation coefficient	KD490
OLCI Atmosphere by-products	
-Photosynthetically Active Radiation	PAR
-Aerosol optical Depth	T865
-Aerosol Angstrom exponent	A865
-Integrated water vapour column	IWV
Land products	
Surface products	
-OLCI Global Vegetation Index and rectified reflectances	OGVI, RC681, RC865
-OLCI Terrestrial Chlorophyll index	OTCI
Atmosphere by-products	
-OLCI Integrated Water vapour column	IWV

Tabla 1. Productos L2 ofrecidos por Sentinel-3.

Flag Name	Flag Description
Classification and Quality Flags	
INVALID	Invalid flag_ instrument data missing or invalid
WATER	Clear sky water
LAND	Clear sky land
CLOUD	Cloudy pixel
CLOUD_AMBOGUOUS	Potentially cloudy pixels
CLOUD_MARGIN	A margin around CLOUD and CLOUD_AMBIGUOUS of 2 pixels in RR and 4 pixels in FR products
SNOW_ICE	Possible sea-ice or snow contamination
INLAND_WATER	Fresh inland waters flag: based on Level-1 land_water flag
TIDAL	Pixel is in shallow water based on Level-1 land_water flag
COSMETIC	Cosmetic flag (from Level-1B): missing data filled in by interpolation
SUSPECT	Suspect flag (from Level-1B): transmission errors means measurements may be Unreliable
HISOLZEN	High solar zenith: >70°
SATURATED	Saturation flag: saturated within any band from 400 to 754 nm or in bands 779, 865 885 and 1020 nm
WVFAIL	Suspect values derived for the water vapour over land: see ATBD SD-03-C02 for details. Set when the following internal flags are raised: ORINPWV_F or OROUTWV_F or L_WV_FAIL
OGVI_FAIL	Suspect values derived for the OGVI (FAPAR): see ATBD SD-03-C13 for details - set when the following internal flags are raised: ORINP1_F or OROUT1_F
OTCI_FAIL	Suspect values derived for the OTCI_ see ATBD for SD-03-C14 details - set when the following internal flags are raised: ORINP2_F (OTCI input out of range) or OROUT2_F (OTCI output out of range) or LRAYFAIL_F (problems deriving Rayleigh reflectance)
Science Flags	
LRAYFAIL	Problems deriving the Rayleigh reflectance over the land: see ATBD SD-03-C15 for Details
OGVI_CLASS_BAD	Flag bad data from OGVI spectral tests
OGVI_CLASS_WS	Flag water or deep shadow from OGVI spectral tests
OGVI_CLASS_CSI	Flag Cloud, snow or ice from OGVI spectral tests
OGVI_CLASS_BRIGHT	Flag bright from OGVI spectral tests
OGVI_CLASS_INVALID_REC	Flag invalid rectification
OTCI_BAD_IN	Input data bad quality: (at least one of B12, B11, B10 is not Valid) or (B12-B11) < Threshold1 or (B11-B10) < Threshold2
OTCI_CLASS_ANG	View angle flag: OTCI_CLASS_IN OK AND view angle > Threshold (TBD) AND sun angle > Threshold (TBD)
OTCI_CLASS_CLSN	Cloud and snow flag: input data quality flag OK but cloud shadow or partial snow

Tabla 2. Información en los productos LQSF L2 S3.

Además de esta información para realizar la clasificación de los datos se ha utilizado otro producto ofrecido por Sentinel-3, y es el de Land Quality and Science Flag (LQSF). Este proporciona una clasificación de cada píxel incluyendo agua, tierra, nubes y nieve entre otros. Esta clasificación se ha realizado teniendo en cuenta índices, ratios en las bandas, los valores de reflectancia y TOA (Top of Atmosphere), el Índice Global de Vegetación de OLCI (OGVI) y el Índice Terrestre de Clorofila (OTCI). Hay 25 clases en total, incluyendo valores para la calidad del píxel, posibles errores en las lecturas y saturaciones. Existe otro producto L2 de clasificación para datos marinos, pero como la experimentación se ha centrado en productos terrestres esta no ha sido utilizada. Podemos ver la clasificación en la [Tabla 2](#).

5.2. Metodología

5.2.1. Adquisición operacional de datos

El primer paso de este proceso tiene como objetivo coleccionar los productos de OLCI L2 de interés para generar el mosaico multitemporal final. Es importante remarcar que la tarea de generar composiciones globales suele demandar trabajar con enormes cantidades de datos operacionales de Remote Sensing [41], lo cual lógicamente aumenta la cantidad de desafíos técnicos (en términos de adquisición de datos, almacenamiento y automatización). Específicamente se hace uso de los datos disponibles que Copernicus provee en sus servicios de la siguiente forma. Primero se define una Región de Interés (ROI) R basada en una forma cerrada poligonal de geo-coordenadas utilizando el formato Geo-JavaScript Notation (GeoJSON). Además, se considera un intervalo temporal T del mosaico multitemporal como un dato de entrada para el proceso de adquisición de datos. Después, teniendo una ROI R y un intervalo temporal T , se demandan la lista de productos OLCI L2 que cumplen estas características al Copernicus Open Acces Hub [42]. Una vez la lista de identificadores de productos relevantes se han obtenido, se itera (en un modo de procesamiento por lotes) para automáticamente descargar los productos operacionales correspondientes (junto con su información de metadatos) en un formato comprimido.

5.2.2. Mosaicado espacial para productos OLCI L2

El segundo paso consiste en mosaicar todos los productos descargados en un cubo global espacio-temporal, teniendo en cuenta las posibles desviaciones de geo-localización entre los productos. Siendo X_r^t una imagen producto derivado L2 que cubre la región r en diferentes periodos de tiempo t , se proyectan sus píxeles a coordenadas WGS84 utilizando la herramienta SNAP [40]. Con este sistema de coordenadas, los píxeles transformados son todos del mismo tamaño, haciendo mucho más sencillo encajar y superponerlos en la estructura de mosaicado. Ya que este proceso puede introducir pequeños errores de localización, un proceso de submuestreo es necesario.

En consecuencia, un cubo 3D espacio-temporal se construye para un número de imágenes producto L2 $G_r(lt, lg, t)$ representando la región r y los instantes de tiempo necesarios para producir el mosaico deseado. Los ejes de este cubo se definen en un rango de latitud lt (altura), longitud lg (anchura) y un intervalo de tiempo t (profundidad) que permite que todas las imágenes X_r^t se registren espacialmente en este cubo, dónde la resolución espacial es de 300 m como en el sensor OLCI S3 del que se obtuvieron los datos. Ya que las coordenadas de este cubo no concuerdan generalmente con las calculadas para cada producto X_r^t , se realiza un proceso de interpolación en el que un kernel estima los valores esperados en las coordenadas del cubo como se muestra en la Figura 7. Los datos de los productos transformado espacialmente de X_r^t se añaden en las capas temporales t del cubo 3D, es decir $G_r(lt, lg, t) = L(X_r^t)$ dónde L es la transformación latitud-longitud desde las coordenadas UTM.

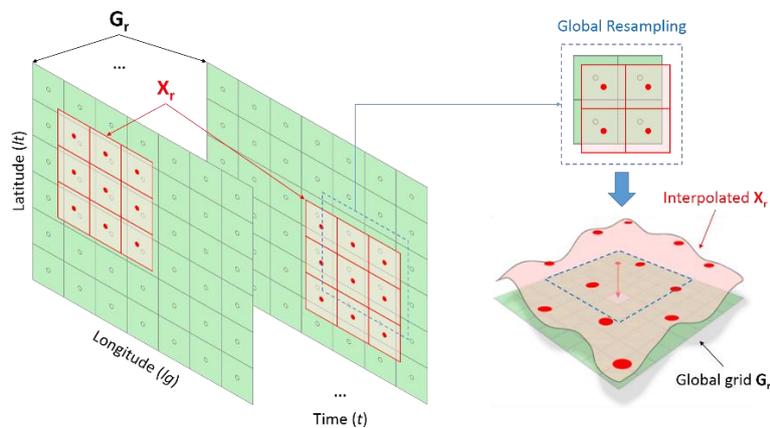


Figura 7. Interpolación del producto para generar el cubo 3D espacio-temporal $G_r(lt, lg, t)$.

5.2.3. Composición temporal de productos OLCI L2

Una vez todos los productos de interés están representados en un cubo uniforme espacio-temporal, el tercer paso del método propuesto es realizar la composición temporal. Específicamente, la composición temporal de datos se realiza eligiendo el mejor valor de pixel del conjunto de muestras temporales dependiendo de diversos factores: número de muestras temporales válidas, clasificación de píxeles Land Use/Land Cover (LULC), valor nominal del producto y otros aspectos importantes como el ángulo solar cenital durante la adquisición o la media estadística del valor producto.

Como se vio en la sección Mosaicado multitemporal de imágenes, trabajos previos han realizado el mosaicado de bandas de reflectancia L2 de productos de S2 o de Landsat. Sin embargo, ninguno de estos métodos puede ser utilizado directamente para crear mosaicos de productos L2 derivados. Estos trabajos previos mostraron que el medoide y la mediana son los algoritmos más adecuados para integrar valores de series temporales, pero para obtener resultados satisfactorios se necesita un número mínimo de píxeles válidos. Como el objetivo de este trabajo es mosaicar productos biofísicos y otros derivados a través de la reflectancia espectral de las bandas de las imágenes, la falta de observaciones temporales puede ser crítica. Dado que los datos de entrada corresponden a cantidades invariantes de parámetros biofísicos y geofísicos de alto nivel, la robustez del algoritmo de medoide contra valores atípicos es la misma que la de la mediana [37].

Teniendo esto en cuenta, se ha utilizado la mediana de las observaciones cuando hay un número mínimo de muestras válidas para cada pixel. En caso de que no haya suficientes muestras para un pixel, tal y como los métodos S2GM [33] y WELD [29], se utiliza un árbol de decisiones. Este árbol se ha diseñado para considerar la clasificación LQSF L2 de OLCI, así como el índice del valor objetivo y consiste en una variación del algoritmo STC-S2 utilizado para productos L2 de S2.

Una vez el cubo espaciotemporal $G_r(lt, lg, t)$ es creado mediante el proceso de registrado geo-espacial, el algoritmo de mosaicado propuesto para productos OLCI S3 se puede dividir en dos módulos principales: un filtrado previo de los datos de $G_r(lt, lg, t)$ y la composición temporal basada en la mediana y el algoritmo definido como Short-Term Composite para S3, STC-S3.

5.2.3.1. Filtrado

Dado el cubo 3D temporal $G_r(lt, lg, t)$, se necesita un filtrado preliminar para considerar exclusivamente los píxeles válidos en el conjunto temporal. El producto LQSF provee una clasificación de cada píxel como se introdujo anteriormente. Sin embargo, cabe mencionar que este producto no contiene clases para vegetación o suelo. Estas se encuentran incluidas en la clase de tierra ya que el sensor OLCI inicialmente estaba orientado a caracterización y monitorización de aguas.

Sin embargo, otros indicadores en este producto pueden proveer información adicional sobre los índices de vegetación. Este es el caso del indicador OGVI, que significa Índice Global de Vegetación OLCI de S3. Este índice se calcula a través de la estimación de la Fracción de la Radiación Fotosintética Activa (FAPAR) utilizando el algoritmo del producto LQSF.

Para ilustrar como funciona este algoritmo, se ha elegido el índice OTCI de S3 como producto L2 biofísico para realizar mosaicos. El índice OTCI estima la clorofila y puede definirse como se muestra en la [Ecuación 3](#), donde Oa_{12} , Oa_{11} y Oa_{10} son las bandas de reflectancia centradas en 753, 709 y 681 nm respectivamente del sensor OLCI.

$$OTCI = \frac{Oa_{12} - Oa_{10}}{Oa_{11} - Oa_{10}}$$

Ecuación 3. Definición de OTCI.

Prioridad	Criterio
1	El píxel se clasifica como LAND, WATER o SNOW-ICE
2	El valor del píxel es un número finito válido
3	El ángulo solar cenital es inferior a 70°

Tabla 3. Criterios para el filtrado de datos.

Un valor de píxel se considera válido cuando las tres condiciones expresadas en la [Tabla 3](#) se cumplen. Los valores con un cenit solar alto (mayores que 70°) no se toman en cuenta porque estos datos pueden ser físicamente irreales. Esto se concluyó en [44] para la precisión de MODIS, y se puede aplicar a S3 también. Además, los valores del producto que no son válidos o finitos también son filtrados. Los píxeles que no son clasificados como LAND, WATER o SNOW-ICE también se deben filtrar, eliminando los clasificados como INVALID o CLOUD entre otros. Después del paso del filtrado las coordenadas espaciales (lt, lg) agrupan un número de valores temporales válidos $\{G_r(lt, lg, t); t \in V_t(lt, lg)\}$, donde $V_t(lt, lg)$ es la marca de tiempo establecida vinculada a (lt, lg) . Los datos resultantes pasan a la siguiente fase.

5.2.3.2. Composición temporal

La composición y mosaicado temporal integra los valores de los píxeles anteriormente filtrados para obtener un representante $C_r(lt, lg)$ para cada píxel (lt, lg) del correspondiente conjunto temporal de valores de productos válidos $\{G_r(lt, lg, t); t \in V_t(lt, lg)\}$. Después del filtrado, generalmente el resultado son diferentes cantidades de muestras válidas por cada píxel. Dependiendo del número de observaciones válidas se utiliza una estrategia u otra. Siendo $N_t(lt, lg) = |V_t(lt, lg)|$ el número de muestras temporales válidas en el cubo.

Cuando hay disponibles suficientes observaciones válidas, el valor representante $C_r(lt, lg)$ se obtiene mediante la mediana del conjunto $\{G_r(lt, lg, t); t \in V_t(lt, lg)\}$. La mediana es un estimador robusto cuando se trata de una pequeña proporción de valores atípicos en suficientes muestras de datos. Sin embargo, cuando no hay suficientes muestras la mediana puede proporcionar una estimación sesgada. Es por esto que la mediana solo se utiliza cuando hay suficientes muestras temporales, que en este caso se ha definido como superior a un threshold th_m .

Si el número de muestras temporales es superior a este valor, se utiliza la mediana; en caso contrario se utiliza la adaptación de STC para S3, $STC_S3(lt, lg)$. En este trabajo el threshold que se ha utilizado ha sido 4, con tal de conseguir que la mediana trabaje correctamente con un número de muestras mínimo (cómo en [33]), reduciendo así la probabilidad de que los valores atípicos afecten a la estimación final. En cuanto al método STC, incluso con pocas muestras es capaz de obtener información útil a través de la identificación auxiliar de los píxeles. Estos métodos se consideran lo suficientemente simples y apropiados para realizar

composiciones de productos L2 operacionales de S2 de forma automatizada. En resumen, el algoritmo de mosaicado temporal se puede formular con la expresión de la [Ecuación 4](#).

$$C_r(lt, lg) = \begin{cases} \text{si } N_t(lt, lg) > th_m & \text{entonces } M_r(lt, lg) \\ \text{en otro caso} & STC_S3(lt, lg) \end{cases}$$

Ecuación 4. Algoritmo de mosaicado temporal.

El algoritmo propuesto de STC para S3 está basado en el STC-S2 del algoritmo de mosaicado global S2 [33], el algoritmo WELD [29] y el Land Cover Climate Change Initiative (LC-CCI) WELD for S2 [45]. Estos algoritmos se han propuesto para reducir la contaminación de aerosoles y nubes residuales en series temporales. En el algoritmo adaptado para productos terrestres L2 de S3, el mejor píxel de un pequeño conjunto de muestras temporales se elige basado en un árbol de decisiones utilizando los productos derivados, la clasificación de píxeles y otros indicadores.

El algoritmo $STC_S3(lt, lg)$ va tomando por cada píxel (lt, lg) dos de las muestras temporales, $G_r(lt, lg, t_1)$ y $G_r(lt, lg, t_2)$ donde $t_1, t_2 \in V_t(lt, lg)$ para compararlos. La lógica del algoritmo de comparación entre píxeles se resume en la [Tabla 4](#). Como en WELD y en STC-S2 cada fila de la tabla es una comparación entre dos muestras temporales, y si esta comparación no se cumple se comprueba la siguiente condición. Este proceso se repite hasta que se cumple una condición o se alcanza la última regla.

Las reglas del árbol de decisiones se basan en los siguientes criterios, dependiendo de las propiedades de las muestras que se están comparando:

- Los píxeles clasificados como LAND tienen la mayor prioridad dado que entre las posibles clasificaciones son los que tienen más probabilidades de ser vegetación.
- Si los píxeles no están clasificados como LAND entonces las clases SNOW_ICE y WATER tienen el segundo y tercer nivel de prioridad respectivamente. SNOW_ICE tiene el segundo nivel ya que es probable que sean píxeles LAND erróneamente clasificados, por lo que WATER es la última opción.

- Una comprobación más exhaustiva se realiza utilizando OGVI para incrementar la probabilidad de una caracterización correcta del píxel como vegetación, ya que los píxeles de vegetación también se pueden clasificar con este índice como SNOW_ICE, CLOUD o WATETR.
- Cuando dos muestras temporales pertenecen a la misma clase, la muestra con mayor valor de OTCI se elige, es decir, el mayor valor de clorofila se convierte en la estimación del píxel considerado en ese periodo de tiempo.
- Los valores de OTCI pueden variar para las muestras de un mismo píxel en un corto periodo de tiempo. El motivo de este comportamiento se debe a las diferencias en fenología. Para minimizar este efecto se trata de seleccionar las muestras temporales en el mismo estado fenológico, por ejemplo, cuando las hojas están fotosintéticamente activas. Un criterio aproximado y simplificado para mantener la mayor actividad fotosintética en la composición es seleccionar el mayor valor de OTCI.

Prioridad	Criterio
1	Si Gr(t1) y Gr(t2) son LAND (por LQSF) y ninguno es CLOUD o SNOW_ICE (por OGVI): Elige el que tenga mayor valor de OTCI
2	Si Gr(t1) y Gr(t2) son LAND (por LQSF) y uno es CLOUD o SNOW_ICE (por OGVI): Elige el valor de OTCI no clasificado como CLOUD o SNOW_ICE
3	Si Gr(t1) y Gr(t2) son LAND (por LQSF) y ambos son CLOUD o SNOW_ICE (por OGVI): Elige el que tenga mayor valor de OTCI
4	Si Gr(t1) o Gr(t2) es LAND (por LQSF) y el otro es CLOUD o SNOW_ICE (por OGVI): Elige el valor de OTCI clasificado como LAND
5	Si Gr(t1) y Gr(t2) son SNOW_ICE (por LQSF): Elige el que tenga mayor valor de OTCI
6	Si Gr(t1) o Gr(t2) es SNOW_ICE (por LQSF) y el otro es WATER (por LQSF): Elige el valor de OTCI clasificado como SNOW_ICE
7	Si Gr(t1) y Gr(t2) son WATER (por LQSF) y ninguno es WATER (por OGVI): Elige el que tenga mayor valor de OTCI
8	Si Gr(t1) y Gr(t2) son WATER (por LQSF) y uno es WATER (por OGVI): Elige el valor de OTCI clasificado como WATER por OGVI
9	Si Gr(t1) y Gr(t2) son WATER (por LQSF) y ambos son WATER (por OGVI): Elige el que tenga mayor valor de OTCI

Tabla 4. Criterio de selección entre muestras temporales.

Estos criterios se diseñaron para clasificar y seleccionar valores de píxeles con el objetivo de obtener mosaicos temporales de productos OTCI L2 S3. En caso de querer mosaicar otros productos, se deberían definir otros criterios para el algoritmo STC, por ejemplo, para productos marítimos utilizar Algal Pigment Concentration. Sin embargo, exceptuando los criterios para diferentes productos el resto de la composición y mosaicado temporal serían idénticos.

5.2.3.2. Validación de los mosaicos

Con tal de validar los mosaicos multitemporales generados, al no existir ningún dato real y válido del terreno con qué compararlo, se han propuesto dos medidas de confianza para juzgar la calidad de los mosaicos generados. A través de un mapa que muestra el número de observaciones válidas se puede obtener de un vistazo información sobre los niveles de confianza del mosaico.

Otra estadística podría ser la desviación estándar de los productos temporales válidos por cada píxel, no obstante, teniendo en cuenta que el número de observaciones válidas suele ser pequeño y puede incluir valores atípicos, la desviación estándar no se considera una medida confiable.

En lugar de utilizar la desviación estándar de la media $S(\bar{G}_r(lt, lg))$ como medida de incertidumbre, si hay al menos dos valores, la desviación estándar de las observaciones válidas de cada píxel se pesan con el valor crítico de la distribución t de Student [46] correspondiente al número de observaciones válidas y a un nivel de significancia $\alpha = 0.05$. El índice de confianza propuesto se define como:

$$U_A(lt, lg) = e^{-t_{N_t-1; 95\%} * S(\bar{G}_r(lt, lg))}$$

Ecuación 5. Índice de confianza propuesto.

Dónde la desviación estándar $S(\bar{G}_r(lt, lg))$ se define como:

$$S(\bar{G}_r(lt, lg)) = \frac{\sqrt{\frac{1}{N_t - 1} * \sum_{t=1}^{N_t} (\bar{G}_r(lt, lg) - (G_r(lt, lg, t))^2)}}{\sqrt{N_t}}$$

Y la media $\bar{G}_r(lt, lg)$ se define como:

$$\bar{G}_r(lt, lg) = \frac{\sum_{t=1}^{N_t} G_r(lt, lg, t)}{N_t}$$

En resumen, el índice de confianza $U_A(lt, lg)$ de un píxel se define como la exponencial inversa de la distribución $t_{n-1;95\%}$ por la desviación estándar $S(\bar{G}_r(lt, lg))$ para que este índice se encuentre en el intervalo $[0,1]$. El uso de $t_{n-1;95\%}$ aumenta la confianza proporcionalmente con el número de observaciones válidas, debido al pequeño número de valores críticos asociados a la distribución t. Por lo tanto, el índice de confianza propuesto genera una medida de confianza normalizada que tiene en cuenta tanto la desviación estándar como el número de muestras válidas.

5.3. Adaptación para FLEX

El algoritmo propuesto se ha diseñado para S3, para poder utilizarlo en un futuro con productos derivados L2 de FLEX se necesita realizar algunos cambios. Como S3 y FLEX volarán en tándem siguiendo la misma órbita con algunos segundos de retraso y FLEX proporcionará productos con la misma resolución espacial que S3, 300x300 m por píxel, se necesitará realizar un registrado geo-espacial entre S3 y FLEX. Este registrado espacial podría realizarse como se ha descrito en la sección 5.2.2. Mosaicado espacial para productos OLCI L2 construyendo un cubo 3D temporal para S3 y FLEX simultáneamente.

En el caso en que FLEX proporcione una clasificación por píxeles para productos L2 entonces el árbol de decisiones STC-S3 tendrá que ser adaptado a las clases de FLEX. En cualquier caso, el mismo producto LQSF L2 S3 puede ser utilizado tras registrar las imágenes de FLEX y S3. Consiguiendo así asignar la clasificación de S3 al correspondiente píxel de FLEX que ya está pensado para productos de tierra y de vegetación. El mayor cambio sería sustituir el producto de OTCI por el producto de FLEX elegido, como fluorescencia, que es el producto principal ofrecido por el sensor FLORIS de FLEX.

Sin embargo, esta modificación también tendría implicaciones adicionales debido a la diferencia entre la naturaleza del fenómeno de la fluorescencia y los niveles de clorofila. Mientras que la concentración de clorofila tiende a ser estable en cortos periodos de tiempo, la fluorescencia es un parámetro biofísico dinámico que requiere un modelo suplementario para caracterizar las variaciones existentes. Es más, la órbita de FLEX/S3 afectará a los niveles de fluorescencia capturados por FLORIS y por lo tanto pueden generarse algunas desviaciones artificiales entre diferentes productos de fluorescencia. Para solucionar esta cuestión, se necesita un modelo empírico (basado en medidas reales de FLEX) para caracterizar como afecta la el punto de adquisición de datos en la órbita a la fluorescencia. Para solucionarlo, cada producto individual podría ser corregido según las condiciones en las que se adquirieron los datos antes de realizar el mosaicado espacial. Esto podría aplicarse también para otros productos disponibles de FLEX.

Capítulo 6

Experimentación

En este capítulo se muestra la parte experimental del trabajo, incluyendo detalles de los conjuntos de datos utilizados, el proceso experimental y los resultados obtenidos, así como su validación.

6.1. Conjunto de datos

Para la experimentación se han utilizado diferentes conjuntos de productos L2 OLCI multitemporales para generar sus correspondientes mosaicos. Específicamente, se ha elegido la región entre París (Francia) y Trento (Italia) como ROI. Por lo tanto, solo se han considerado los productos que incluyen alguna zona del área rectangular entre las coordenadas latitud-longitud (48.864716, 2.349014) y (46.06787, 11.12108). Para esta ROI, se han definido los siguientes conjuntos de datos:

- **Invierno:** Desde el 15 hasta el 31 de enero de 2019 (27 productos).
- **Primavera:** Desde el 15 hasta el 21 de abril de 2019 (15 productos).
- **Verano:** Desde el 15 hasta el 21 de julio de 2019 (15 productos).
- **Otoño:** Desde el 14 hasta el 20 de octubre de 2019 (15 productos).
- **Abril-1:** Desde el 1 al 7 de abril de 2019 (16 productos).
- **Abril-2:** Desde el 8 al 14 de abril de 2019 (15 productos).
- **Abril-3:** Desde el 15 hasta el 21 de abril de 2019 (15 productos).

- **Abril-4:** Desde el 22 hasta el 30 de abril de 2019 (17 productos).

Como ejemplo, en la [Figura 8](#) se muestra en el área azul los productos OLCI L2 considerados para el conjunto *Otoño* y la ROI en naranja. Todos los conjuntos son de una semana excepto el conjunto *Invierno* que consta de dos semanas debido a la gran cantidad de nubes presentes en Europa a lo largo de la estación de invierno. Estos datos se han elegido para analizar la evolución estacional de la vegetación en el centro de Europa a través de mosaicos. Ya que los valores de los índices de vegetación cambian mucho a lo largo del tiempo, también se han considerado conjuntos de datos semanales para abril, dónde la tercera semana (*Abril-3*) corresponde al conjunto *Primavera*.

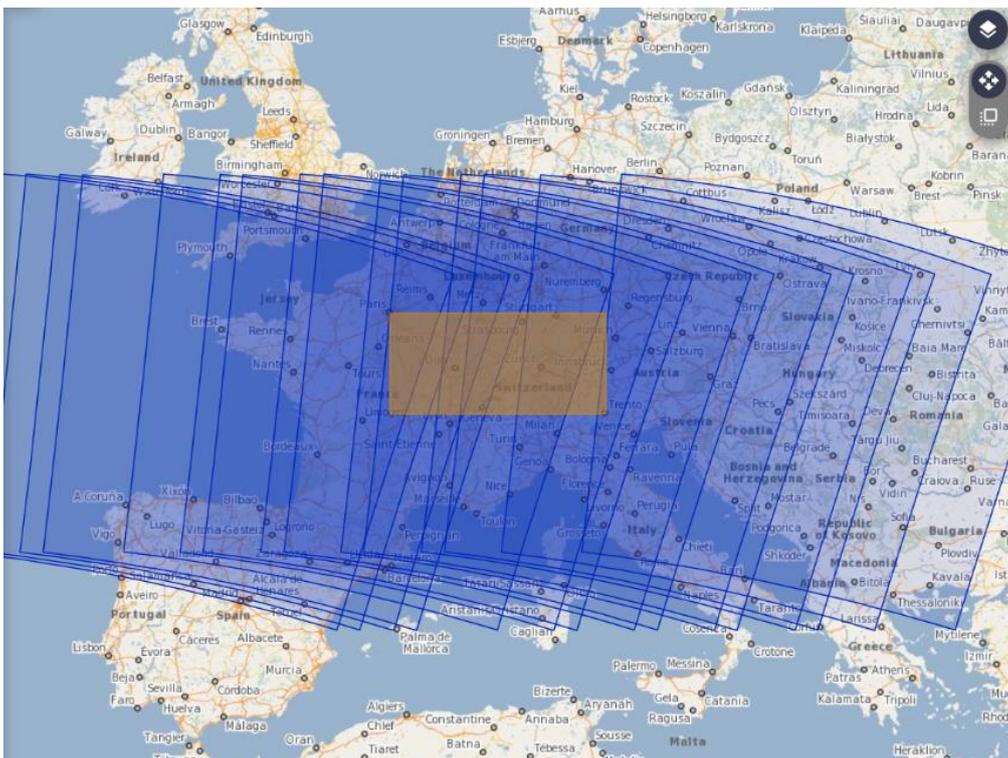


Figura 8. Productos considerados para el conjunto *Otoño*.

6.2. Proceso experimental

Para cada uno de los conjuntos de datos, se ha utilizado la metodología propuesta para descargar los productos, generar sus correspondientes mosaicos de OTCI y validar los resultados según las medidas de confianza previamente explicadas. Aunque los productos OLCI L2 contienen muchos datos [23], es importante recalcar que únicamente la información de los productos OTCI y LQSF se ha utilizado para generar los mosaicos. Para organizar mejor los resultados, se han dividido los experimentos en dos:

- **Experimento 1:** En este experimento se han utilizado los conjuntos de datos *Invierno*, *Primavera*, *Verano* y *Otoño* para realizar un análisis estacional del año 2019.
- **Experimento 2:** Este experimento incluye los conjuntos de datos de las diferentes semanas de abril para estudiar la evolución de los mosaicos de OTCI a lo largo de este mes.

En cuanto al software y hardware, todos los experimentos se han realizado mediante Python 3.5.2 y MATLAB 2019^a, y han sido ejecutados en un servidor Ubuntu 16.04 x64 con 24 Intel® Xeon® ES-2640 procesadores con 189 Gb de RAM.

6.3. Resultados

6.3.1. Experimento 1

En la [Figura 9](#) se muestran los mosaicos OTCI generados para (a) *Invierno*, (b) *Primavera*, (c) *Verano* y (d) *Otoño*. Como se puede observar, los valores de OTCI bajos se representan en colores azules, mientras que los valores de OTCI alto, indicando un alto contenido en clorofila, se muestran en verde. Adicionalmente, los píxeles invariables o inválidos se muestran en gris. Esta escala de colores es la proporcionada oficialmente para ser utilizada en los servicios de Sentinel-Hub, donde los colores rojizos representan bajas concentraciones de clorofila y los verdosos expresan altos valores de clorofila relacionados con vegetación más densa y saludable.

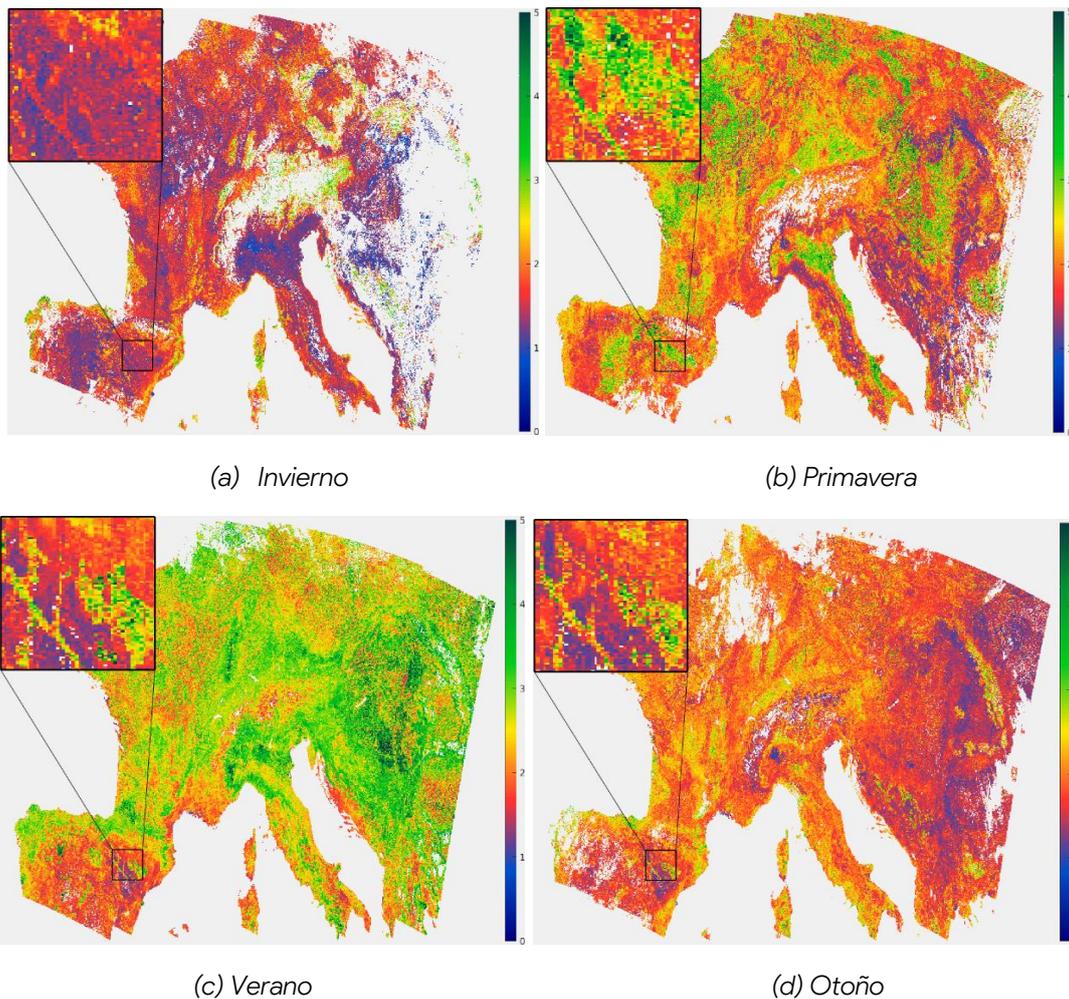


Figura 9. Mosaicos generados en el experimento 1.

De los mosaicos obtenidos se pueden hacer algunas observaciones importantes. En general, el análisis estacional muestra que la mayor cantidad de vegetación verde se alcanza en verano, seguido por primavera, otoño y por último invierno, tal y como se esperaría en una región europea [47]. Sin embargo, algunas áreas de interés específico muestran otras tendencias, como es el caso del detalle mostrado en la Figura 9, donde se incluye la parte del norte de la comunidad autónoma de Aragón. Específicamente, esta región contiene parques naturales como son la Sierra y Cañones de Guara [49], que principalmente contienen coníferas como *Pinus nigra* y *Pinus sylvestris*, que incrementan su actividad fotosintética durante la primavera a través de un pigmento foliar superior [50]. Precisamente este incremento se puede observar en la composición multitemporal de la Figura 9 (b), mostrando la consistencia del sistema generando mosaicos de productos derivados OLCI en distintas fechas.

En cuanto a la disponibilidad de datos, el número de observaciones válidas para cada mosaico es otro aspecto que debe ser analizado. En la [Figura 10](#) se muestran los correspondientes mapas de píxeles válidos para cada estación. En este caso, el azul a amarillo muestra el número de muestras válidas de OTCI para cada conjunto de datos, y el color gris expresa falta de información. Cada mapa tiene un rango ligeramente distinto ya que el número de muestras válidas por píxel es lógicamente diferente en cada periodo.

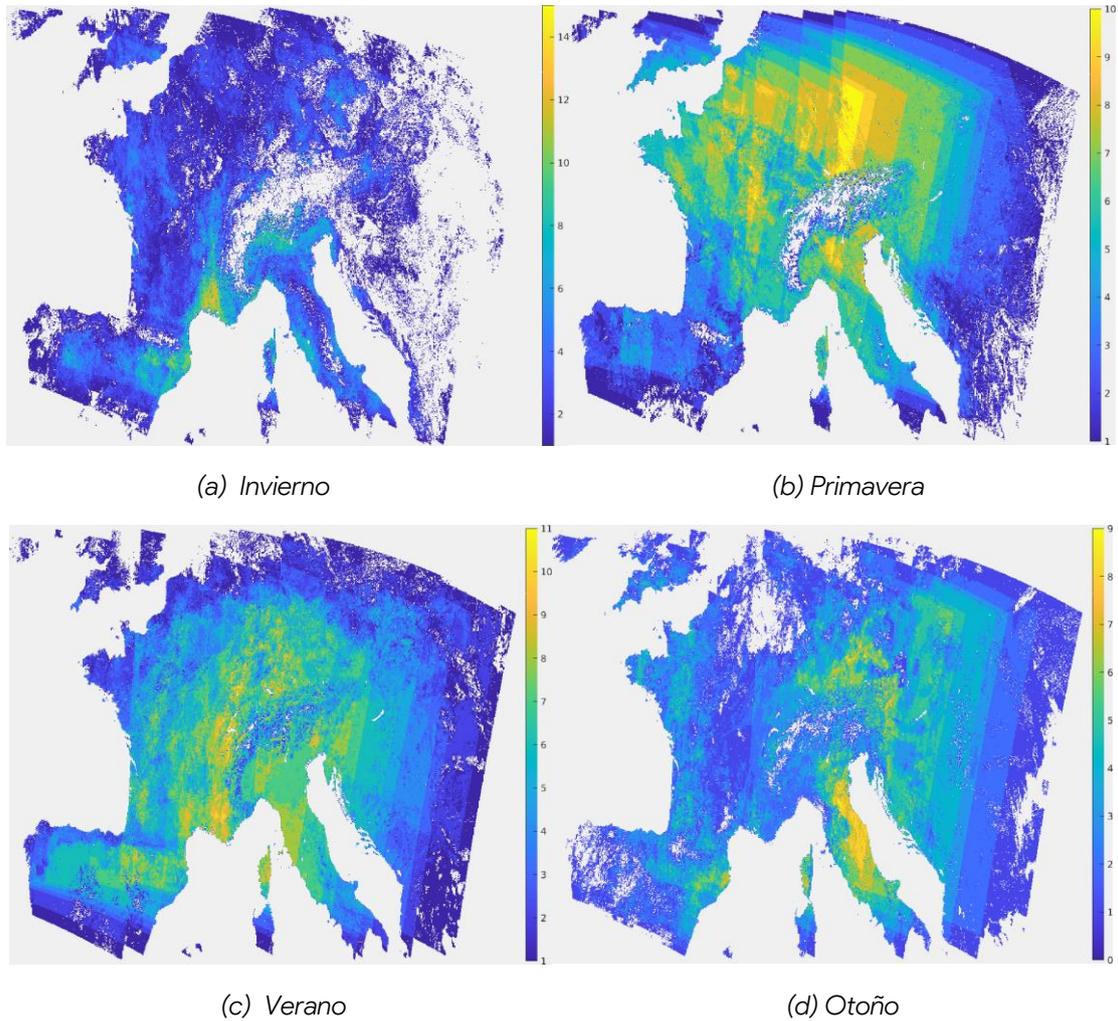


Figura 10. Mapas de píxeles válidos para los mosaicos generados en el Experimento 1.

Por ejemplo, en [Figura 10](#) (a) vemos como no hay muchos píxeles válidos debido a la gran cantidad de nubes en invierno. Aunque intentamos reducir este efecto utilizando productos a lo largo de dos semanas, la cantidad de datos para

la estación de invierno es considerablemente menos que en otros conjuntos de datos, siendo el de *Primavera* y *Verano* los conjuntos con más información disponible. Ya que el número total de observaciones válidas tiende a ser mucho menor que el número de productos considerados en cada conjunto de datos, hace más útil el diseño del algoritmo de mosaicado multitemporal.

También es necesario analizar las diferencias que aporta el uso del algoritmo propuesto comparado con diferentes alternativas de mosaicado multitemporal. Para ello, en la [Figura 11](#) se muestran mosaicos generados mediante el método propuesto ([Figura 11 \(a\)](#)), la media ([Figura 11 \(b\)](#)) y la mediana ([Figura 11 \(c\)](#)) usando el conjunto de Primavera. Además, se muestran los detalles del parque natural Sierra y Cañones de Guara.

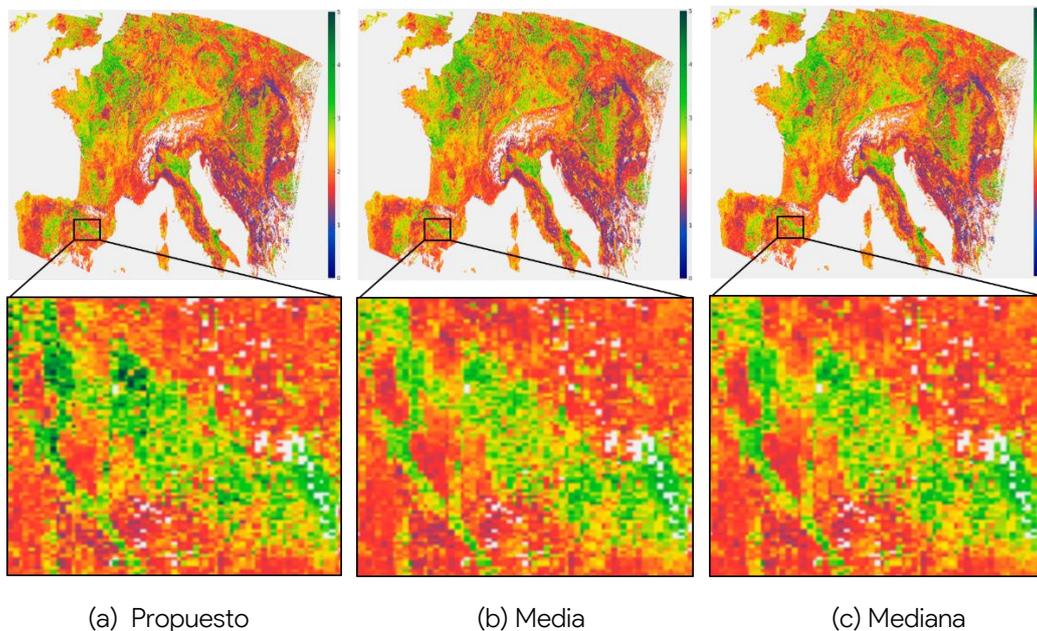


Figura 11. Comparación cualitativa entre diferentes mosaicos OTCI en el conjunto Abril.

En general en esta área el número de observaciones es inferior al threshold propuesto ($th_m = 4$) que ayuda a identificar mejor las diferencias entre los algoritmos. Se puede observar como el algoritmo propuesto es capaz de mantener valores mayores de OTCI aun cuando hay pocas observaciones disponibles. Mientras que la media y la mediana tienen a reducir la cantidad de clorofila detectada en el mosaico, el algoritmo propuesto es capaz de seleccionar valores de OTCI que representan niveles de pigmentos mayores durante el periodo temporal.

Por lo tanto, las reglas de decisión del árbol son capaces de discernir picos mayores durante la estación de primavera cuando la actividad fotosintética se espera que alcance los máximos en este parque natural.

Estos resultados cualitativos revelan que el método propuesto tiene ventajas respecto a los métodos de la media y la mediana ya que la combinación de las observaciones de OTCI se ven menos afectadas por desviaciones en las muestras. En el caso de la media, es uno de los más afectados ya que reduce los picos. En el caso de la mediana, podemos ver que lógicamente tiene un comportamiento similar al método propuesto cuando el número de observaciones es suficiente. Sin embargo, el método propuesto muestra mejores capacidades combinando valores multitemporales de OTCI a la vez que mostrando la clorofila fielmente con muestras limitadas. Tanto el efecto de la contaminación atmosférica como las limitaciones temporales son un factor clave en entornos dónde el método propuesto puede dar ventajas significantes con datos limitados.

Por último, en la [Figura 12](#) se muestran los mapas de confianza de los mosaicos de OTCI correspondientes según el índice de confianza propuesto. En esta figura, el color azul hasta el rojo muestra los valores de confianza y el color gris muestra información no disponible (gris claro) o zonas donde no se puede definir la métrica propuesta (gris oscuro). Como se puede observar, niveles de confianza bajos no tienen por qué corresponder con un alto número de píxeles válidos en la [Figura 10](#). Se pueden observar variaciones en los niveles de confianza dependiendo de las localizaciones geográficas y de las observaciones disponibles. Por ejemplo, en el caso de la fig.(d), el norte de Francia tiene niveles más bajos que el centro de Italia donde hay más muestras válidas disponibles. En general, los valores de confianza de los mosaicos generados son altos exceptuando zonas con nubes, bordes y otras excepciones como las áreas con muy pocas observaciones.

A pesar de que las medidas de OTCI son correctas en las áreas con muy pocas observaciones, el nivel de confianza que presentan es muy bajo ya que los valores de estos productos no se pueden contrastar durante el conjunto temporal del mosaico. Es importante mencionar que los niveles de confianza están entre 0 y 1, así que los mapas se han escalado acorde a este intervalo.

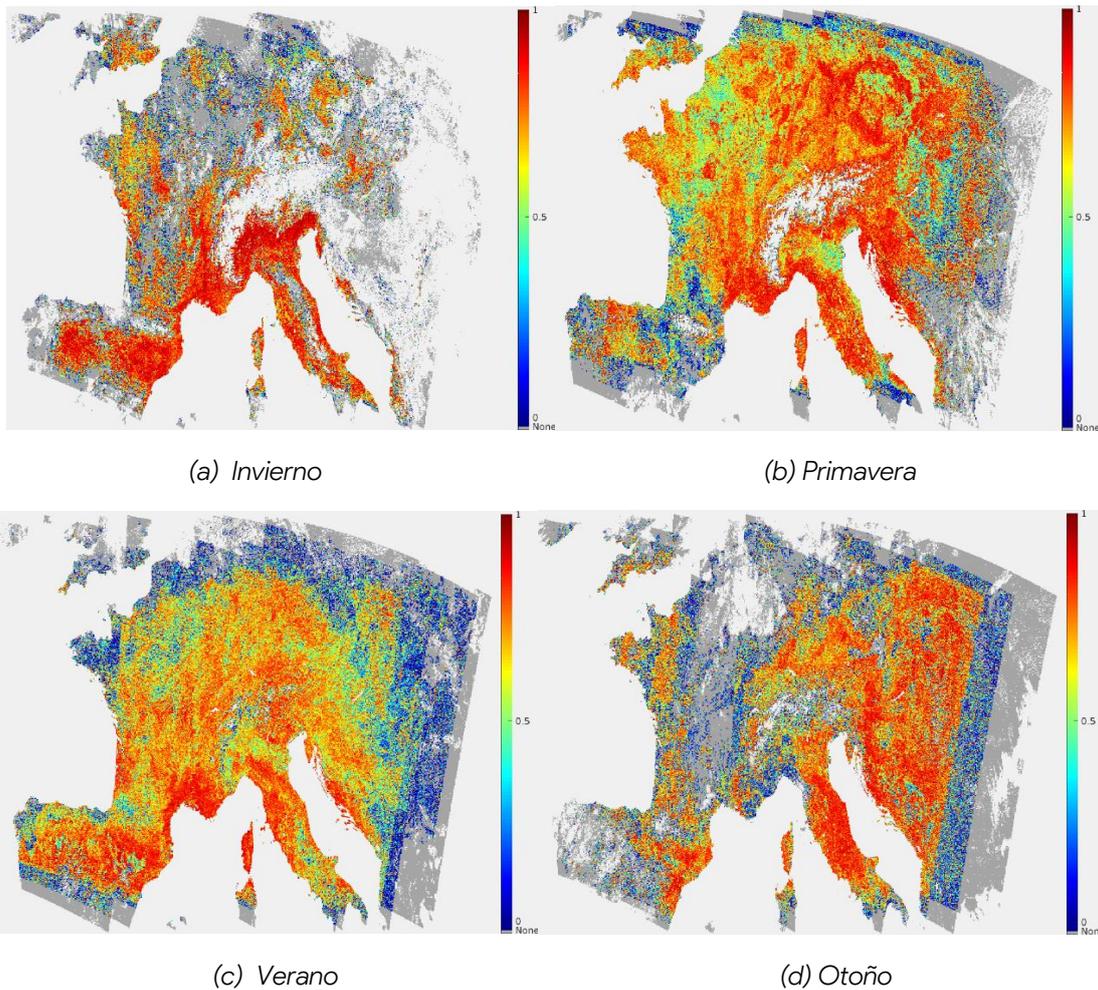


Figura 12. Mapas de confianza para los mosaicos de OTCI del experimento 1.

6.3.2. Experimento 2

En cuanto al segundo experimento, en la [Figura 13](#) se muestran los mosaicos de OTCI generados para los conjuntos de datos de las diferentes semanas de abril. Como en el experimento anterior, los valores de OTCI bajos se representan en azul y en verde los valores altos. Teniendo en cuenta los resultados mostrados en la [Figura 13](#) es posible ver que estos mosaicos de abril de 2019 tienen similitudes importantes. Sin embargo, se pueden encontrar puntos relevantes en cuanto al crecimiento esperado en la vegetación de la zona de interés. Cuando se analizan los resultados desde una perspectiva global, los mosaicos multitemporales revelan que la zona central de Europa tiene a incrementar los niveles de clorofila de la vegetación a lo largo de abril, siendo

consistente con el previo análisis dónde la mayor cantidad de vegetación se encuentra en verano. En el caso de los detalles que se muestran de la Sierra y Cañones de Guara, podemos observar cómo hay algunos cambios relevantes. Concretamente, los mayores niveles de OTCI se encuentran en el conjunto 3-Abril seguido por 4-Abril, 2-Abril y 1-Abril en ese orden. Este es el pico de la estación de primavera en este parque natural, que podemos observar en la [Figura 13 \(c\)](#) indicando que el método propuesto es capaz de generar mosaicos multitemporales de forma consistente reteniendo las peculiaridades de las diferentes regiones.

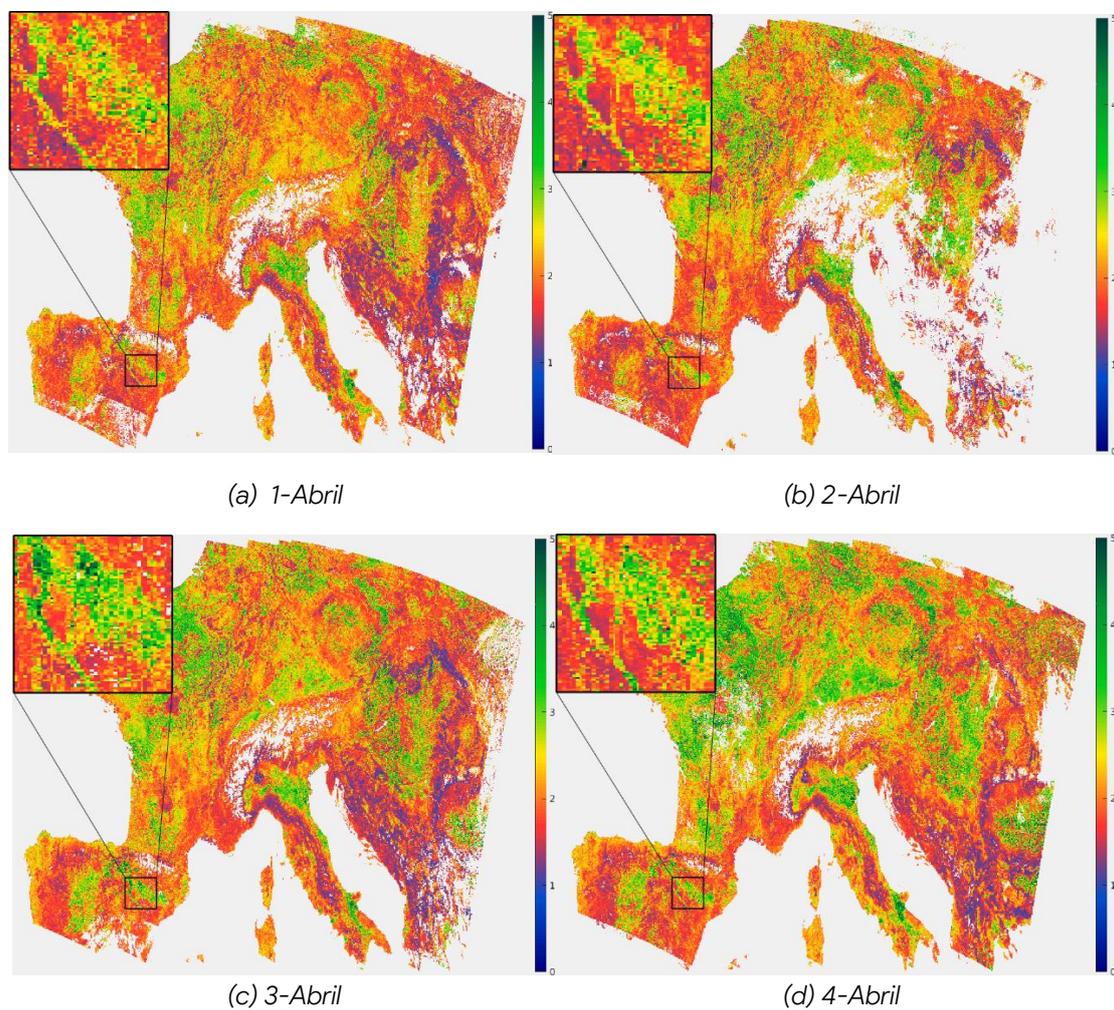
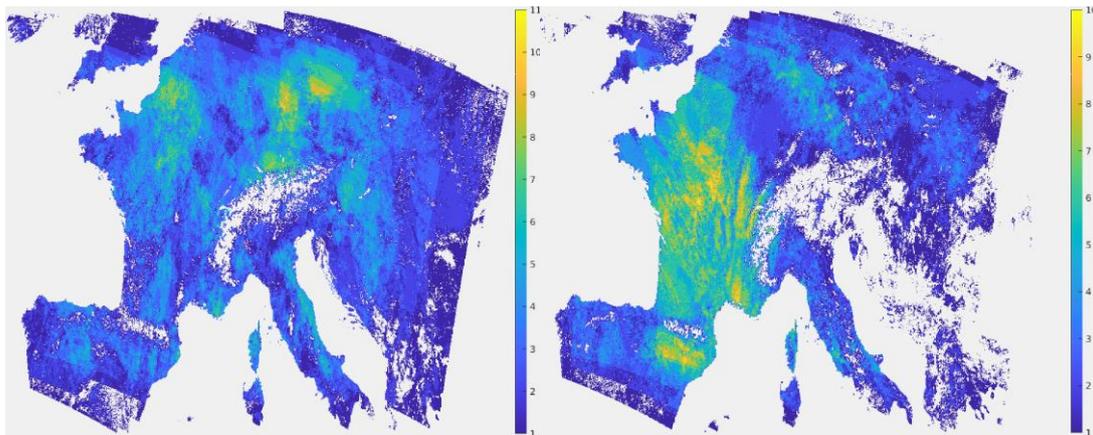


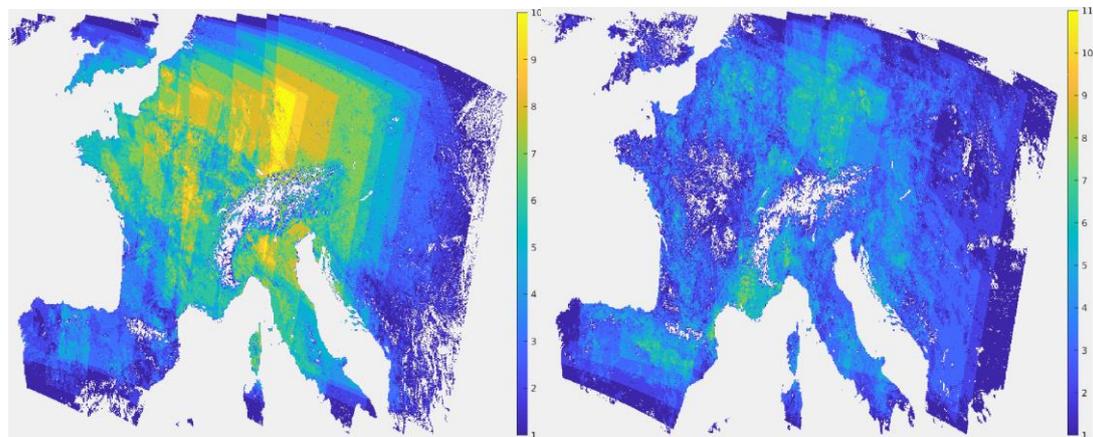
Figura 13. Mosaicos de OTCI generados para el experimento 2.

La [Figura 14](#) muestra el mapa de observaciones disponibles en el experimento 2. Como se puede observar hay diferencias importantes entre los mosaicos semanales ya que la cantidad de nubes es un gran problema en Europa central durante abril [44]. El número de muestras válidas por cada conjunto es significativamente inferior al número total de productos disponibles, siendo la excepción 3-Abril.



(a) 1-Abril

(b) 2-Abril



(c) 3-Abril

(d) 4-Abril

Figura 14. Mapas de píxeles válidos del experimento 2.

Por último, en la [Figura 15](#) se muestran los mapas de confianza de los mosaicos semanales. Como se puede observar, los valores de confianza son mayores en áreas con más observaciones válidas y con un comportamiento más estable respecto a la cantidad de vegetación detectada por el OTCI. La coloración en todos los mapas ha sido la misma que en el experimento 1.

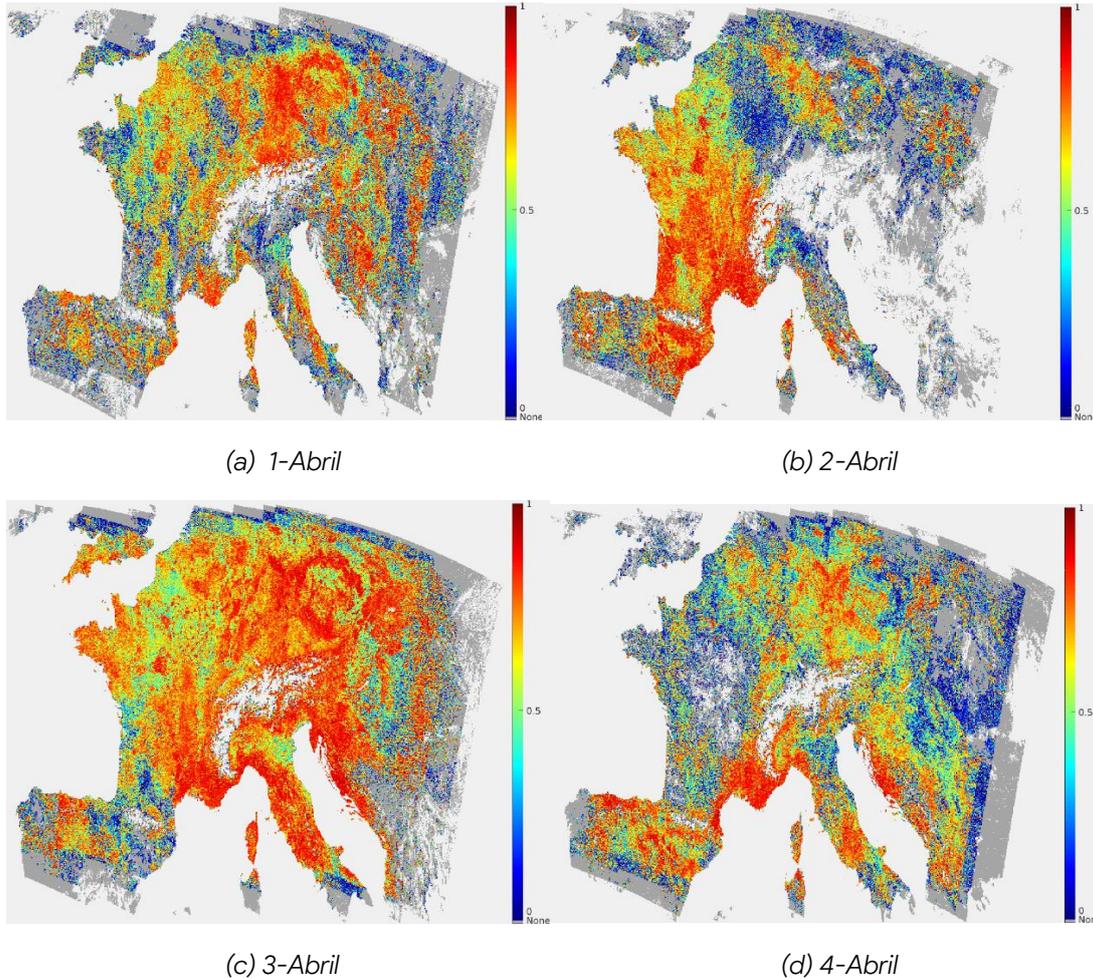


Figura 15. Mapas de confianza del experimento 2.

Capítulo 7

Planificación

Previamente a la realización de este proyecto se realizó una planificación inicial del tiempo que se iba a emplear para completar este proyecto. En este caso la planificación final ha cumplido prácticamente los tiempos deseados.

7.1. Planificación inicial

- Definición del proyecto y reuniones. Se realizarán varias reuniones con el profesor tutor, cotutor y otros participantes en el proyecto para decidir la metodología a implementar, revisar el progreso y comentar las modificaciones necesarias.
Tiempo estimado: 20 horas.
- Investigación. Realizar una investigación sobre el campo de Remote Sensing, así como de los satélites Sentinel-3 y FLEX, los sensores que utilizan, datos que proporcionan etc. También sobre las diferentes técnicas necesarias para el algoritmo y como programarlas.
Tiempo estimado: 80 horas.
- Diseño de la metodología. Diseñar el algoritmo y el resto de procesos necesarios tras haber realizado la investigación y conocer las funciones que debe realizar.
Tiempo estimado: 40 horas.
- Desarrollo de la metodología. Implementar el diseño definido en el paso anterior mediante código y conseguir que funcione correctamente.
Tiempo estimado: 40 horas.

- Experimentación. Experimentar con diferentes conjuntos de datos para obtener resultados que muestren si el diseño implementado es el deseado.
Tiempo estimado: 30 horas.
- Revisión y mejoras. Revisar y mejorar el algoritmo tras los primeros resultados experimentales obtenidos.
Tiempo estimado: 20 horas.
- Documentación. Tiempo dedicado a redactar la memoria, realizar la planificación, etc.
Tiempo estimado: 60 horas.
- Presentación. Diseñar la presentación del TFM propuesto y su explicación al jurado.
Tiempo estimado: 10 horas.

La suma total de las horas planificadas corresponde a 300 horas.

7.2. Planificación final

En esta sección se describen las horas utilizadas para los apartados anteriores realmente a lo largo del proyecto.

- Definición del proyecto y reuniones. Dado el escaso conocimiento del campo al inicio del proyecto fueron necesarias más reuniones.
Tiempo estimado: 20 horas. Tiempo empleado: 25 horas.
- Investigación. Esta parte del proyecto también ocupó más tiempo del esperado, por motivos similares al anterior.
Tiempo estimado: 80 horas. Tiempo empleado: 90 horas.

- Diseño de la metodología. Esta parte, sin embargo, se realizó en menor tiempo del esperado, ya que se encontró un buen diseño inicialmente.
Tiempo estimado: 40 horas. Tiempo empleado: 30 horas.
- Desarrollo de la metodología. Para el desarrollo se tardó el tiempo esperado, el mayor problema fue tratar con los datos, debido a su estructuración y los cambios de coordenadas mayormente.
Tiempo estimado: 40 horas. Tiempo empleado: 40 horas.
- Experimentación. Para la experimentación fue necesario menos del tiempo esperado.
Tiempo estimado: 30 horas. Tiempo empleado: 20 horas.
- Revisión y mejoras. Para las revisiones también se necesitó más tiempo debido a que procesar este tipo de datos toma una gran cantidad de tiempo, y para cada cambio que se realizaba se necesitaba un tiempo hasta volver a tener los resultados
Tiempo estimado: 20 horas. Tiempo empleado: 30 horas.
- Documentación. En cuanto a la documentación fueron necesarias algunas horas más debido a la redacción del artículo científico y reportes.
Tiempo estimado: 60 horas. Tiempo empleado 65 horas.
- Presentación. El diseño y preparación de la presentación duró el tiempo planificado.
Tiempo estimado: 10 horas. Tiempo empleado 10 horas.

El tiempo empleado finalmente ha sido de 310 horas, estando cerca del tiempo planificado inicialmente.

Capítulo 8

Conclusión

En este proyecto se ha presentado un algoritmo de mosaicado multitemporal novedoso diseñado especialmente para productos biofísicos operacionales S3 o de otro tipo para la misión S3/FLEX. Además, se ha publicado un artículo científico sobre el trabajo desarrollado [51]. El método propuesto se ha definido siguiendo cuatro procesos secuenciales (adquisición operacional de datos, mosaicado espacial, composición temporal y estimación de niveles de confianza) para generar mosaicos consistentes de productos derivados S3. Los experimentos realizados utilizando diferentes conjuntos de datos S3 han mostrado que el método propuesto es más conveniente.

Una de las primeras conclusiones que se obtienen de este trabajo es la importancia de los productos derivados biofísicos en la misión S3/FLEX y como generar estos mosaicos multitemporales puede ayudar en estas misiones. Por otro lado, el uso de datos multitemporales permite reducir la influencia de la oclusión producida por nubes y otros tipos de contaminación en los datos. Además, permite el uso de amplias áreas de interés para realizar análisis biofísicos globales.

Otra cuestión importante es la relacionada con la disponibilidad de observaciones válidas y su inconsistencia. El proceso de seleccionar un píxel representativo para un periodo de tiempo es generalmente muy sensible al número de muestras válidas. Al contrario que las funciones convencionales de composiciones temporales, el método propuesto ha sido diseñado para tener en cuenta los casos que aportan menos información aparentemente considerando la naturaleza de estos datos, que en el marco global resultan producir resultados más relevantes.

En un futuro se planea avanzar en las siguientes direcciones: extender el método propuesto para usarse en datos de Remote Sensing multimodales, desarrollar modelos empíricos para la fluorescencia de FLEX globalmente y estudiar el uso de diferentes productos, así como utilizar diferentes técnicas para los datos que faltan.

Capítulo 9

Anexos

9.1. Referencias

1. N. Pettorelli, H. Schulte to Buhne, A. Tulloch, G. Dubois, C. Macinnis- " Ng, A. M. Queiros, D. A. Keith, M. Wegmann, F. Schrodt, M. Stellmes ´ et al., "Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward," *Remote Sens. Ecol. Conserv.*, vol. 4, no. 2, pp. 71–93, 2018.
2. R. Fernandez-Beltran, J. M. Haut, M. E. Paoletti, J. Plaza, A. Plaza, and F. Pla, "Multimodal probabilistic latent semantic analysis for sentinel-1 and sentinel-2 image fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1347–1351, 2018.
3. R. Fernandez-Beltran, P. Latorre-Carmona, and F. Pla, "Single-frame super-resolution in remote sensing: a practical overview," *International Journal of Remote Sensing*, vol. 38, no. 1, pp. 314–354, 2017.
4. J. Haas and Y. Ban, "Urban land cover and ecosystem service changes based on sentinel-2a msi and landsat tm data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 2, pp. 485–497, 2018.
5. R. Fernandez-Beltran, J. M. Haut, M. E. Paoletti, J. Plaza, A. Plaza, and F. Pla, "Remote sensing image fusion using hierarchical multimodal probabilistic latent semantic analysis," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 12, pp. 4982–4993, 2018.
6. R. Fernandez-Beltran, F. Pla, and A. Plaza, "Sentinel-2 and sentinel-3 intersensor vegetation estimation via constrained topic modeling," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 10, pp. 1531–1535, 2019.
7. D. Poursanidis and N. Chrysoulakis, "Remote sensing, natural hazards and the contribution of esa sentinels missions," *Remote Sensing Applications: Society and Environment*, vol. 6, pp. 25–38, 2017.

8. L. Zhuo, Q. Dai, D. Han, N. Chen, B. Zhao, and M. Berti, "Evaluation of remotely sensed soil moisture for landslide hazard assessment," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 162–173, 2019.
9. R. Fernandez-Beltran, F. Pla, and A. Plaza, "Intersensor remote sensing image registration using multispectral semantic embeddings," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 10, pp. 1545–1549, 2019.
10. R. Fernandez-Beltran, A. Plaza, J. Plaza, and F. Pla, "Hyperspectral unmixing based on dual-depth sparse probabilistic latent semantic analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6344–6360, 2018.
11. R. Fernandez-Beltran, F. Pla, and A. Plaza, "Endmember extraction from hyperspectral imagery based on probabilistic tensor moments," *IEEE Geoscience and Remote Sensing Letters*, 2020, dOI: 10.1109/LGRS.2019.2963114.
12. J. Aschbacher and M. P. Milagro-Perez, "The European Earth monitoring (GMES) programme: Status and perspectives," *Remote Sens. Environ.*, vol. 120, pp. 3–8, 2012.
13. M. Berger, J. Moreno, J. A. Johannessen, P. F. Levelt, and R. F. Hanssen, "Esa's sentinel missions in support of earth system science," *Remote Sens. Environ.*, vol. 120, pp. 84–90, 2012.
14. C. Donlon, B. Berruti, A. Buongiorno, M.-H. Ferreira, P. Fernandes, J. Frerick, P. Goryl, U. Klein, H. Laur, C. Mavrocordatos et al., "The global monitoring for environment and security (gmes) sentinel-3 mission," *Remote Sens. Environ.*, vol. 120, pp. 37–57, 2012.
15. J. Vicent, N. Sabater, C. Tenjo, J. R. Acarreta, M. Manzano, J. P. Rivera, P. Jurado, R. Franco, L. Alonso, J. Verrelst et al., "Flex end-to-end mission performance simulator," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4215–4223, 2016.
16. D. Arnas, P. Jurado, I. Barat, B. Duesmann, and R. Bock, "Flex: A parametric study of its tandem formation with sentinel-3," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 7, pp. 2447–2452, 2019.
17. J. Kravitz, M. Matthews, S. Bernard, and D. Griffith, "Application of sentinel 3 olci for chl-a retrieval over small inland water targets: Successes and challenges," *Remote Sens. Environ.*, vol. 237, p. 111562, 2020.
18. D. Kyryliuk and S. Kratzer, "Evaluation of sentinel-3a olci products derived using the case-2 regional coastcolour processor over the baltic sea," *Sensors*, vol. 19, no. 16, p. 3609, 2019.
19. L. A. Brown, J. Dash, A. L. Lidon, E. Lopez-Baeza, and S. Dransfeld, "Synergetic exploitation of the sentinel-2 missions for validating the sentinel-3 ocean and land color instrument terrestrial chlorophyll index over

- a vineyard dominated mediterranean environment,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 7, pp. 2244–2251, 2019.
20. A. Kokhanovsky, M. Lamare, O. Danne, C. Brockmann, M. Dumont, G. Picard, L. Arnaud, V. Favier, B. Jourdain, E. Le Meur et al., “Retrieval of snow properties from the sentinel-3 ocean and land colour instrument,” *Remote Sensing*, vol. 11, no. 19, p. 2280, 2019.
 21. Z. Zhang, L. Zhao, and A. Lin, “Evaluating the performance of sentinel3a olci land products for gross primary productivity estimation using ameriflux data,” *Remote Sensing*, vol. 12, no. 12, p. 1927, 2020.
 22. H. Lyu, Z. Yang, L. Shi, Y. Li, H. Guo, S. Zhong, S. Miao, S. Bi, and Y. Li, “A novel algorithm to estimate phytoplankton carbon concentration in inland lakes using sentinel-3 olci images,” *IEEE Trans. Geosci. Remote Sens.*, pp. 1–12, 2020.
 23. P. Team, “GMES Space Component Sentinel-3 Payload Data Ground Segment Products Definition Document,” European Space Agency (ESA), Tech. Rep., 11 2016.
 24. Z. Xiao, “Compositing, smoothing, and gap-filling techniques,” *Advanced remote sensing terrestrial information extraction and applications*, pp. 75–90, 2012.
 25. Y. Guo, F. Li, P. A. Caccetta, and D. Devereux, “Multiple temporal mosaicing for landsat satellite images,” *J. Appl. Remote Sens.*, vol. 11, no. 1, p. 015021, 2017.
 26. B. N. Holben, “Characteristics of maximum-value composite images from temporal avhrr data,” *Int. J. Remote Sens.*, vol. 7, no. 11, pp. 1417–1434, 1986.
 27. J. Cihlar, D. Manak, and M. D’lorio, “Evaluation of compositing algorithms for avhrr data over land,” *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 2, pp. 427–437, 1994.
 28. A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, “Overview of the radiometric and biophysical performance of the modis vegetation indices,” *Remote Sens. Environ.*, vol. 83, no. 1-2, pp. 195–213, 2002.
 29. D. P. Roy, J. Ju, K. Kline, P. L. Scaramuzza, V. Kovalsky, M. Hansen, T. R. Loveland, E. Vermote, and C. Zhang, “Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States,” *Remote Sens. Environ.*, vol. 114, no. 1, pp. 35–49, 2010.
 30. P. Potapov, S. Turubanova, and M. C. Hansen, “Regional-scale boreal forest cover and change mapping using landsat data composites for european russia,” *Remote Sens. Environ.*, vol. 115, no. 2, pp. 548–561, 2011.
 31. V. Kovalsky and D. P. Roy, “The global availability of landsat 5 tm and landsat 7 etm+ land surface observations and implications for global 30 m

- landsat data product generation,” *Remote Sens. Environ.*, vol. 130, pp. 280–293, 2013.
32. “SEN2THREE: SENTINEL-2 LEVEL 3 SPATIO-TEMPORAL SYNTHESIS PROCESSOR, VERSION 1.0.0. DOCUMENTATION.”
<https://step.esa.int/thirdparties/sen2three/1.0.0/>, 2017, Accessed: 2020-04-01.
 33. “The Sentinel-2 Global Mosaic service.” <https://land.copernicus.eu/imagery-in-situ/global-image-mosaics/>, 2018, Accessed: 2020-04-01.
 34. G. Schmidt, C. B. Jenkerson, J. Masek, E. Vermote, and F. Gao, “Landsat ecosystem disturbance adaptive processing system (ledaps) algorithm description,” US Geological Survey, Tech. Rep., 2013.
 35. P. V. Potapov, S. A. Turubanova, M. C. Hansen, B. Adusei, M. Broich, A. Altstatt, L. Mane, and C. O. Justice, “Quantifying forest cover loss in democratic republic of the congo, 2000–2010, with landsat etm+ data,” *Remote Sens. Environ.*, vol. 122, pp. 106–116, 2012.
 36. M. C. Hansen, D. P. Roy, E. Lindquist, B. Adusei, C. O. Justice, and A. Altstatt, “A method for integrating modis and landsat data for systematic monitoring of forest cover and change in the congo basin,” *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2495–2513, 2008.
 37. N. Flood, “Seasonal composite landsat tm/etm+ images using the medoid (a multi-dimensional median),” *Remote Sensing*, vol. 5, no. 12, pp. 6481–6500, 2013.
 38. A. Struyf, M. Hubert, P. Rousseeuw et al., “Clustering in an objectoriented environment,” *Journal of Statistical Software*, vol. 1, no. 4, pp. 1–30, 1997.
 39. J. Louis, V. Debaecker, B. Pflug, M. Main-Knorn, J. Bieniarz, U. Mueller-Wilm, E. Cadau, and F. Gascon, “Sentinel-2 sen2cor: L2a processor for users,” in *Proceedings Living Planet Symposium 2016*. Spacebooks Online, 2016, pp. 1–8.
 40. “Sentinel Application Platform (SNAP),” <https://step.esa.int/main/toolboxes/snap/>, 2017, Accessed: 2020-04-21.
 41. Y. Cheng, K. Zhou, J. Wang, and J. Yan, “Big earth observation data integration in remote sensing based on a distributed spatial framework,” *Remote Sensing*, vol. 12, no. 6, p. 972, 2020.
 42. “Copernicus Open Access Hub,” <https://scihub.copernicus.eu/>, 2017, Accessed: 2020-04-21.
 43. “Ocean and Land Colour Instrument (OLCI) Algorithm Theoretical Basis Document),” https://sentinels.copernicus.eu/documents/247904/349589/OLCI_L2_ATBD_FAPAR.pdf, 2010, Accessed: 2020-05-19.

44. X. Wang and C. S. Zender, "Modis snow albedo bias at high solar zenith angles relative to theory and to in situ observations in greenland," *Remote Sens. Environ.*, vol. 114, no. 3, pp. 563–575, 2010.
45. "High-Resolution Landcover CCI project," <http://cci.esa.int/hlandcover/>, 2018, Accessed: 2020-04-12.
46. Student, "The probable error of a mean," *Biometrika*, pp. 1–25, 1908.
47. R. Stockli and P. L. Vidale, "European plant phenology and climate as seen in a 20-year avhrr land-surface parameter dataset," *Int. J. Remote Sens.*, vol. 25, no. 17, pp. 3303–3330, 2004.
48. M. Palahí, M. Tome, T. Pukkala, A. Trasobares, and G. Montero, "Site index model for *pinus sylvestris* in north-east spain," *Forest Ecology and Management*, vol. 187, no. 1, pp. 35–47, 2004.
49. J. Riedel, A. Bernues, and I. Casaus, "Livestock grazing impacts on herbage and shrub dynamics in a mediterranean natural park," *Rangeland Ecology & Management*, vol. 66, no. 2, pp. 224–233, 2013.
50. A. Trasobares, T. Pukkala, and J. Miina, "Growth and yield model for uneven-aged mixtures of *pinus sylvestris* l. and *pinus nigra* arn. in catalonia, north-east spain," *Annals of forest science*, vol. 61, no. 1, pp. 9–24, 2004
51. Ibañez, D.; Fernandez-Beltran, R.; Sotoca, J.M.; Mollineda, R.A; Moreno, J. and Pla, F.; "Multi-temporal Mosaicing for Sentinel-3/FLEX Derived Level-2 Product Composites", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 13, pp. 5439 – 5454, 2020.