GEOQUBIDY

geoqubidy@upm.es

MANUAL DE USUARIO FEGAPP

Interfaz de usuario gráfica, Texto

Descripción generada automáticamenteInterfaz de usuario gráfica

Descripción generada automáticamente con confianza baja

****Contenido****

[1 Antecedentes 2](#_Toc193983757)

[2 Introducción 2](#_Toc193983758)

[3 Instalación y estructura 4](#_Toc193983759)

[3.1 Instalación 4](#_Toc193983760)

[3.1.1 Python 4](#_Toc193983761)

[3.1.2 Dependencias 4](#_Toc193983762)

[3.1.3 Variables de entorno 5](#_Toc193983763)

[3.2 Estructura del proyecto 6](#_Toc193983764)

[4 Productos 8](#_Toc193983765)

[4.1 Sigpac crono 8](#_Toc193983766)

[4.2 Imágenes satelitales de Sentinel-2 8](#_Toc193983767)

[4.3 marcadores de estabilidad 8](#_Toc193983768)

[5 Funcionalidades 8](#_Toc193983769)

[5.1 Obtención de sigpac crono 8](#_Toc193983770)

[5.2 Descarga de imágenes 9](#_Toc193983771)

[5.3 Generación de marcadores 10](#_Toc193983772)

[6 Justificación científica 11](#_Toc193983773)

[6.1 Manejo de la base de datos 11](#_Toc193983774)

[6.1.1 Gestión de la base de datos SIGPAC 11](#_Toc193983775)

[6.1.2 Tratado de los datos SIGPAC 11](#_Toc193983776)

[6.1.3 Generación de capas temporales 12](#_Toc193983777)

[6.2 Series de tiempo 13](#_Toc193983778)

[6.2.1 Índices espectrales 13](#_Toc193983779)

[6.2.2 Suavizado de series de tiempo 14](#_Toc193983780)

[6.3 Cálculo de marcadores 14](#_Toc193983781)

[6.3.1 Comparación de perfiles temporales 14](#_Toc193983782)

[6.3.2 Detección de ciclos y productividad 14](#_Toc193983783)

[6.3.3 Parametrización de umbrales 15](#_Toc193983784)

# Antecedentes

El convenio entre el FEGA y la UPM busca optimizar los controles de la Política Agrícola Común (PAC) mediante el uso de imágenes satelitales de la constelación Sentinel-2. El objetivo principal es desarrollar modelos informáticos capaces de detectar el abandono de parcelas agrícolas (viñedo, frutales, olivar y pastos permanentes) de forma automatizada y precisa.

Este acuerdo responde a la necesidad del FEGA de cumplir con la normativa europea, que exige controles rigurosos para evitar fraudes y penalizaciones financieras. La tecnología Sentinel-2 permite analizar la evolución de la vegetación y detectar cambios que indiquen abandono de cultivos.

La UPM se encargará de diseñar y ejecutar un proyecto de investigación para desarrollar algoritmos y metodologías de detección, aportando su experiencia técnica y recursos informáticos. El FEGA, por su parte, proporcionará información, bases de datos y financiación, además de validar e implementar los resultados en sus controles oficiales.

El convenio tiene una duración de cuatro años, prorrogables a ocho, y cuenta con una financiación máxima de 250.750,60 euros por parte del FEGA. Se espera que este acuerdo permita al FEGA mejorar la eficiencia de sus controles, reducir el número de inspecciones manuales y minimizar el riesgo de sanciones por incumplimiento de la normativa de la PAC.

# Introducción

FEGAPP es una aplicación diseñada para obtener y analizar la evolución temporal de las parcelas a lo largo de los años. Para ello, se emplea la base de datos del SIGPAC, a partir de la cual se genera un producto que recopila información temporal de las parcelas en estudio. Para extraer conclusiones más precisas, es fundamental analizar el comportamiento del suelo. En este sentido, se utilizan técnicas de teledetección basadas en imágenes del sensor Sentinel-2, que ofrece una resolución espacial máxima de 10 metros. La combinación de ambas fuentes de datos permite analizar dinámicas y establecer correlaciones entre las declaraciones registradas y la "verdad terreno".

A partir del estudio conjunto de la base de datos y de las series temporales, se han diseñado unos marcadores que analizan las dinámicas de las parcelas para inferir su estabilidad temporal. Estos marcadores pueden ser útiles en la identificación de parcelas abandonadas.

La Figura 1 muestra un resumen de las funcionalidades de la aplicación junto con los productos generados. Actualmente, la interfaz permite la generación de la base de datos SIGPAC\_CRONO y la descarga de series temporales de Sentinel-2. Se está trabajando en la integración de los marcadores en la aplicación gráfica, pero, mientras tanto, se proporciona un tester para su generación, aunque sigue en fase de desarrollo.

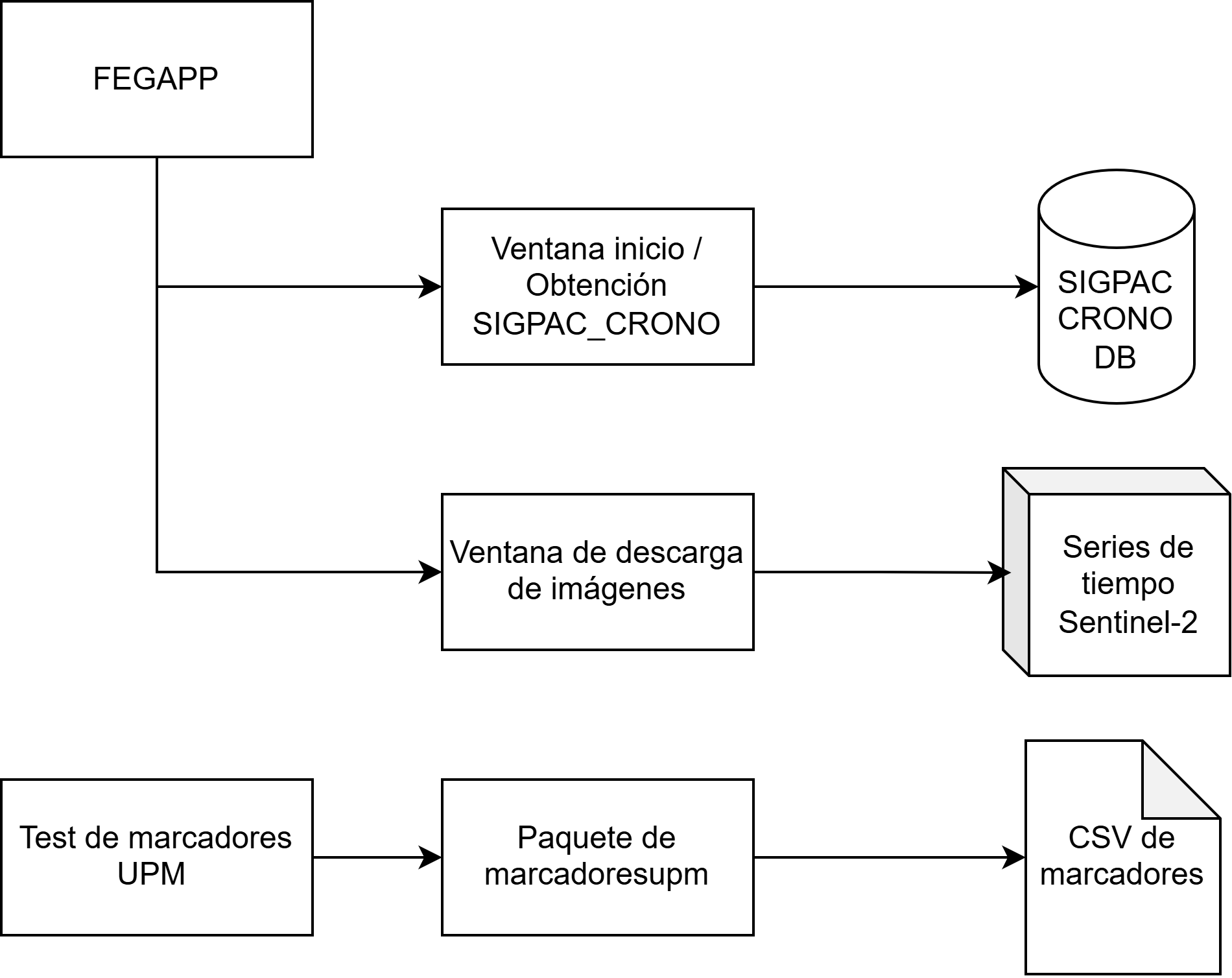


Figura 1: Resumen de las funcionalidades ofertadas en la aplicación y sus productos.

# Instalación y estructura

## Instalación

### Python

La aplicación esta implementada y ejecutada en el lenguaje de programación Python, en concreto en su versión 3.

Se ha planteado un modelo de actualización y mejoras de la aplicación por lo que creemos que es mejor no compilar la aplicación en un ejecutable binario por dos motivos:

* Es necesario generar un ejecutable nuevo cada vez que se añadan mejoras o parches.
* Al estar escrito en Python, se permite la ejecución en distintos sistemas operativo, mientras que si se generará un ejecutable sería necesario generar binarios para cada sistema operativo que se de soporte.

Es decir, para ejecutar la aplicación será necesario tener instalado del software de Python en la maquina local.

### Dependencias

Si ya se tiene instalado Python en la maquina local, se procede a explicar cómo instalar las dependencias de uso para su correcto funcionamiento de la aplicación.

Primero es necesario obtener el código fuente de la aplicación, esto se puede conseguir mediante la descarga desde el repositorio en formato comprimido, o clonando el repositorio de GitHub a través de la herramienta de git.

git clone <https://github.com/TomPug/FEGA_DETECT.git>

Una vez con la carpeta del proyecto en local, y situándonos dentro del repositorio, se encuentra un pequeño fichero de texto llamado ‘requirements.txt’, este archivo contiene todas las dependencias de uso de la aplicación para su ejecución a través de Python. Es recomendable crear un entorno virtual de Python e instalar los requisititos ahí:

pip install -r requirements.txt

Adicionalmente, es necesario instalar de forma manual una librería que no es posible desde el pip de Python. Este paquete es GDAL muy útil para el manejo de datos geoespaciales y se puede instalar usando un Wheel de instalación obtenible desde el repositorio de GitHub de <https://github.com/cgohlke/geospatial-wheels/releases> . En este repositorio hay muchos paquetes y versiones de estos, se debe descargar e instalar el Wheel que coincida con la versión deseada de GDAL (la más reciente sirve), junto con la versión de Python3 instalada en la máquina. Por ejemplo, para Python3.10 y una máquina de 64bits este sería el archivo por descargar e instalar:

pip install GDAL-3.10.1-cp310-cp310-win\_amd64.whl

Finalmente, cuando se termine la instalación de los requisitos ya dispondremos de todas las características necesarias para ejecutar la aplicación, para ello introduciremos el siguiente comando:

python3 main.py

Cabe destacar que este comando se puede introducir en comando de bash o PowerShell autoejecutables y que la aplicación se abra como si fuera un icono en el escritorio.

### Variables de entorno

Para conseguir que la aplicación se conecte con la base de datos seleccionada es necesario generar dos ficheros:

* .env 🡪 Archivo de tipo ‘entorno’ que sirve para insertar variables de entorno en la máquina, permitiendo no almacenar en el código de la aplicación información sensible. Este archivo debe contener, ‘MYUSER’ que hace referencia al usuario de la base de datos, ‘MYPASSWORD’ siendo la contraseña del usuario, ‘MYHOST’ haciendo referencia a la IP de la maquina con el servicio y finalmente ‘MYPORT’ que es el puerto abierto para la conexión con el servicio de la base de datos en el host. Hay que destacar que, si no se detecta en la aplicación este archivo de configuración, se pedirá por pantalla los datos y se generará dicho fichero al inicio de esta.
* CSV\_CONFIG.csv 🡪 alojado dentro de la carpeta de ‘src/config’. Debe contener que, para cada comunidad autónoma, las fechas de inicio de cada campaña de sigpac para cada año que esté en la base de datos del sigpac, aparte de ello es necesario añadir el nombre que se le ha dado a la base de datos de cada comunidad autónoma y añadirla en una columna llamada nombre base de datos.

Interfaz de usuario gráfica, Aplicación, Tabla, Excel

Descripción generada automáticamente

Figura 2: Ejemplo del documento CSV\_CONFIG.csv

## Estructura del proyecto

Aquí se muestra la estructura de carpetas y archivos en los que se compone el repositorio de GitHub con todas las funcionalidades implementadas hasta el momento, junto con una breve descripción del archivo:

**📂 FEGA\_DETECT**

**📁 docs**

Carpeta donde se contiene el manual de uso extenso.

* 📄 manual\_uso.docx

**📁 img**

Carpeta donde se encuentran las imágenes asociadas con el README.

* 📄 **descarga.png**
* 📄 **FEGAPP.png**

**📁 src**

Contiene la información necesaria para el correcto funcionamiento del **SIGPAC Crono**:

* **📁 config**
  + **CSV\_CONFIG.csv**: Archivo CSV con la información de comunidades autónomas, años de campaña, fechas de inicio y nombre de la base de datos instalada.
* **📁 data**
  + **📁 TILES**

Carpeta con los *tiles* de Sentinel-2 necesarios para obtener la geometría correcta para la descarga de imágenes, en formato **ShapeFile**.

* + **📁 umbrals**

Carpeta con los umbrales calculados de los marcadores tras la función parametrización, en formato **CSV**.

* **📁 img**

Carpeta con los archivos de imágenes necesarias:

* + 📄 **composición.png**: Composición del FEGA y el grupo de investigación.
  + 📄 **IconoFegaApp.ico**: Imagen de GEOQUBIDY.
* **📁 lib**

Contiene los códigos fuente de las funcionalidades de la aplicación:

* + 📄\_\_init\_\_.py: Marca el comienzo de un paquete.
  + **📁 descarga**
    - 📄\_\_init\_\_.py: Marca el comienzo de un paquete.
    - 📄DESCARGA\_GUI.py: Código de la interfaz gráfica para la descarga de imágenes.
    - 📄descarga\_planet.py: Código para la descarga de imágenes.
    - 📄Stack\_stacks.py: Algoritmo para el apilado de todas las imágenes descargadas.
  + **📁 procesamiento**
    - 📄\_\_init\_\_.py: Marca el comienzo de un paquete.
    - **📁 sigpac**
      * 📄\_\_init\_\_.py: Marca el comienzo de un paquete.
      * 📄FEGA\_REC\_APP.py: Algoritmos de descarga de la base de datos del SIGPAC y generación del SIGPAC Crono.
    - **📁 upm**
      * 📄\_\_init\_\_.py: Marca el comienzo de un paquete.
      * 📄marcadoresupm.py: Algoritmos de cálculo de marcadores estadísticos de estabilidad de parcelas, analizando dinámicas temporales.
* **📁 test**
  + 📄\_\_init\_\_.py: Marca el comienzo de un paquete.
  + 📄conext.py: Importación de paquetes para la carpeta test
  + 📄marcadores\_trial.py: Prueba del cálculo de marcadores.
* 📄\_\_init\_\_.py: Marca el comienzo de un paquete.
* 📄 **Fega.py**: Código principal que integra todos los algoritmos dentro de una interfaz gráfica para ejecutar la aplicación.

📄 **LICENSE**: Licencias de uso de la aplicación.

📄 **main.py**: Código para ejecutar la aplicación.

📄 **README.md**: Texto plano del README.

📄 **requirements.txt**: Requisitos de instalación de paquetes de Python.

# Productos

En este punto se describirán los productos que se pueden obtener mediante el uso de la aplicación utilizando las diferentes funcionalidades.

## Sigpac crono

El producto SIGPAC\_CRONO es una base de datos espacial calculada respecto a los datos almacenados en el SIGPAC de recintos y líneas de declaración. Estas capas se combinan para obtener unas geometrías que contengan los atributos de ambas capas, que hemos denominado ‘recintos declarados’. Estos ‘recintos declarados’ se calculan anualmente, generando tantas capas como años en el intervalo de estudio se haya seleccionado. El producto final es una base datos espacial con las capas de los ‘recintos declarados’ anuales y la capa crono que es una interpretación temporal de cada geometría, conteniendo datos importantes para analizar su estabilidad y usos.

## Imágenes satelitales de Sentinel-2

Las imágenes satelitales que se obtienen a través de esta herramienta son cubos de datos que contienen las componentes espaciales junto la temporalidad de cada píxel, es decir series de tiempo. Esto es útil para analizar las dinámicas de las coberturas del suelo, dando información de si han sufrido cambios o no, junto con la resolución espacial y temporal que proporciona la misión de Sentinel-2.

## marcadores de estabilidad

Los marcadores de estabilidad proporcionan información sobre la dinámica de la una parcela haciendo uso de la técnica de teledetección junto con el análisis estadístico de series de tiempo. Se proporcionan varios marcadores:

* TAM (Temporal Angle Mapper): que infiere si en la vida de la parcela se ha sufrido un cambio brusco de dinámica.
* Anual cycles (ciclos anuales): Describe la forma en el tiempo de la parcela, introduciendo el valor de ‘1’ si ese año ha tenido un ciclo estable y ‘0’ en caso contrario.

# Funcionalidades

A continuación, se mostrarán las interfaces de la aplicación para obtener los productos descritos en el punto anterior.

## Obtención de sigpac crono

Para generar el producto del SIGPAC\_CRONO se usará la ventana principal de la aplicación, donde se deben introducir toda la información de configuración, tal y como se muestra en la Figura 2. El resultado se generará dentro de la carpeta de salida seleccionada incluyendo los datos de los años de estudio seleccionados de la zona y usos elegidos.

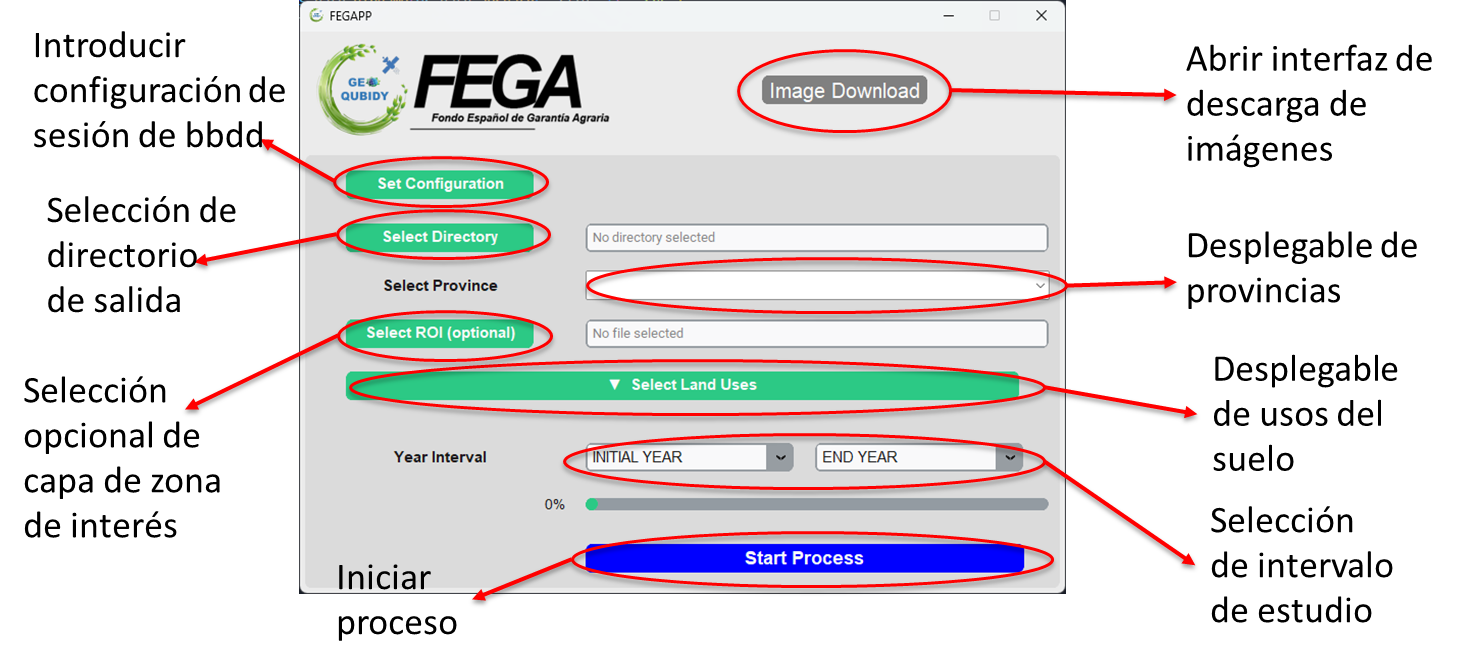


Figura 3: Ventana de inicio y de obtención del SIGPAC\_CRONO.

## Descarga de imágenes

Dentro de la ventana principal hay un botón que nos conducirá a la ventana de descarga de imágenes, la cual nos pedirá toda la configuración mostrada en la Figura 3. Las series de tiempo descargadas se encontrarán dentro de la carpeta de salida, y habrán sido calculadas en función del índice espectral y resolución seleccionados.

Adicionalmente, para definir la zona de estudio hay dos posibilidades, si se escribe dentro del cuadro de texto habilitado el nombre del TILE de Sentinel-2, se descargará toda la imagen de Sentinel-2 asociada al TILE. Pero si se marca la casilla de ‘Shapefile’, pedirá seleccionar un fichero en formato ‘Shapefile’, este se analizará donde se encuentra respecto a los TILEs de Sentinel-2 y solo se descargará la fracción de superficie asociada a la geometría introducida.

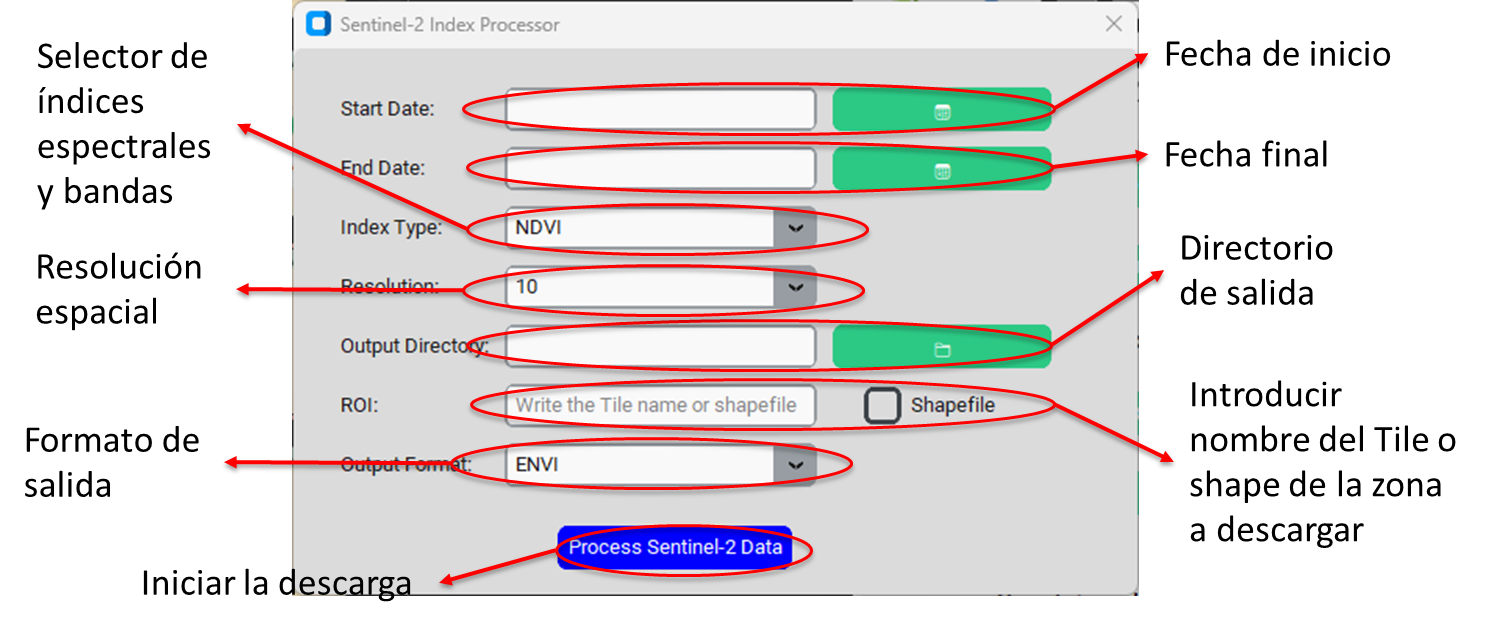


Figura 4: Ventana de descarga de imágenes de Sentinel-2.

## Generación de marcadores

Actualmente, no se tiene integrado en la interfaz gráfica la funcionalidad de los marcadores, pero se proporciona dentro de la carpeta de ‘src/test’ un pequeño código para ejecutar la funcionalidad para calcular los marcadores en terminal. La Figura 4 muestra las opciones que están disponibles dentro de la prueba, que son las funcionalidades que se pueden realizar con la librería de ‘marcadoresupm.py’.

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

Figura 5: Menú de la prueba de marcadores.

Para obtener los marcadores es necesario:

1. Calcular los índices espectrales de las parcelas.
2. Descargar los CSVs de datos.
3. Parametrizar umbrales (opcional).
4. Calcular marcadores.

El resultado es un CSV que contiene los identificadores de las parcelas y los marcadores calculados.

# Justificación científica

## Manejo de la base de datos

### Gestión de la base de datos SIGPAC

Para la gestión de la base de datos del SIGPAC, se ha empleado una caché local actualizada hasta el año 2024. La extracción de los datos relevantes se ha realizado a partir de cuatro tablas clave:

1. **t$recinto\_ex:** Esta tabla resulta de la unión de `t$recinto` y `t$atrib\_recinto`, consolidando la información geométrica y de atributos de los recintos en una única estructura. De esta tabla se extraerán los polígonos de cada recinto junto con sus atributos de interés.
2. **t$linea\_declaracion:** Esta tabla, por su parte, contiene las declaraciones realizadas por los agricultores sobre las actividades agrícolas llevadas a cabo en cada recinto. Esta información es esencial para contrastar la clasificación del SIGPAC con la realidad productiva del terreno.

En resumen, la primera tabla (`t$recinto\_ex`) proporcionan la clasificación oficial del SIGPAC, mientras que la tabla `t$linea\_declaracion` ofrece la perspectiva de los agricultores sobre el uso real de los recintos. La combinación de estas fuentes de datos permite realizar un análisis exhaustivo y preciso de la actividad agrícola en el área de estudio.

### Tratado de los datos SIGPAC

La gestión de la base de datos se lleva a cabo utilizando PostgreSQL como sistema de gestión de bases de datos relacionales. Para interactuar con la base de datos y extraer los datos necesarios, se emplea el lenguaje de programación Python en conjunto con el paquete SQLAlchemy, que proporciona una interfaz de alto nivel para trabajar con bases de datos SQL.

El proceso de obtención de los recintos se realiza de manera iterativa sobre un rango de años definido. Para cada año, se ejecuta una consulta SQL a la tabla t$recinto\_ex para extraer información relevante sobre los recintos agrícolas. La consulta selecciona campos específicos como el identificador único (dn\_oid), la geometría espacial (dn\_geom), las fechas de inicio y fin del recinto (dn\_initialdate, dn\_enddate), el coeficiente de admisibilidad prevalente (cap\_prevalente), el uso SIGPAC (uso\_sigpac), las incidencias registradas (incidencias) y el área del recinto (area\_m2).

Se aplican filtros a la consulta para seleccionar únicamente los recintos que cumplen con ciertos criterios, como encontrarse dentro del rango de fechas del año en cuestión, pertenecer a una provincia específica y tener un área superior 500m2. Además, se excluyen ciertos usos del suelo SIGPAC como 'CA' (Caminos), 'AG' (Cuerpos de Agua), 'ZU' (Zonas Urbanas), 'ED' (Edificaciones), 'ZC'(Zona Concentrada), 'ZV' (Zona censurada) e 'IV' (Invernaderos).

Los datos resultantes de cada consulta anual se almacenan en un objeto GeoDataFrame de la biblioteca GeoPandas y se agregan a una lista llamada rc\_df. Esta lista contendrá un GeoDataFrame por cada año, lo que facilita el posterior análisis y procesamiento de los datos de recintos a lo largo del tiempo.

De manera similar, se obtiene la información de las líneas de declaración contenida en la tabla t$linea\_declaracion. Se realiza un bucle para extraer año por año los datos de las declaraciones y sus polígonos asociados.

Una vez obtenidas las capas geométricas de recintos y líneas de declaración, se realiza una superposición espacial de ambas capas. Esta operación permite asociar a cada polígono de recinto la información de la declaración correspondiente realizada por el agricultor. El resultado de esta superposición es una nueva capa que denominamos "Recintos Declarados", la cual contendrá información tanto de la clasificación SIGPAC como de la declaración del agricultor para cada recinto. Esta capa combinada es fundamental para el posterior análisis y clasificación de los estados de abandono y otras singularidades presentes en los recintos agrícolas.

Imagen que contiene Forma

El contenido generado por IA puede ser incorrecto.

Figura 6: Esquema de unión entre Recintos y Líneas de declaración.

### Generación de capas temporales

La creación de la cronología de estados de abandono se basa en el análisis temporal de los Recintos Declarados a lo largo de los años seleccionados. Para ello, se implementan las siguientes etapas:

* **Clipado y Conversión:**
  + Se recorta la capa de Recintos Declarados a la zona de interés definida por el usuario.
* **Overlay Temporal:**
  + Se realiza un *overlay* espacial de tipo "union" entre las capas de Recintos Declarados de cada año consecutivo.
  + Este proceso genera una nueva capa que combina la información de los recintos a lo largo del tiempo, manteniendo la geometría y los atributos relevantes.
* **Cálculo de Campos de Cronología:**
  + Se crean nuevos campos en la capa resultante del *overlay* para representar la evolución temporal de los estados de abandono (a\_crono), los productos declarados (p\_crono) y los usos del suelo (u\_crono).
  + Estos campos se codifican utilizando una secuencia de dígitos, donde cada dígito representa el estado o valor correspondiente en un año determinado.
  + Finalmente, se exportan las capas resultantes (recintos declarados anuales, capa crono por usos, y capa crono conjunta) en formato GPKG para su análisis y visualización.

La cronología generada permite identificar patrones de cambio en el uso del suelo agrícola a lo largo del tiempo, facilitando la detección de áreas con tendencias de abandono, cambios en los cultivos predominantes y otras dinámicas relevantes para la gestión y planificación agrícola.

## 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.

### Índices espectrales

Los índices espectrales son combinaciones de diferentes bandas en los que se obtiene un beneficio en la observación. Hay diferentes tipos de índices espectrales, este caso de estudio se usa el NDVI (1), siendo un índice normalizado, y por otro lado tenemos índices de forma espectral como el AR (2) y AS1 (3).

Gráfico

El contenido generado por IA puede ser incorrecto.

Figura 7: Firma espectral con los ángulos AR y AS1

|  |  |
| --- | --- |
|  | (1) |
|  | (2) |
|  | (3) |

### Suavizado de series de tiempo

Las series temporales resultantes de la aplicación de una máscara de calidad no contienen valores de datos en los casos en que el píxel está oculto por una capa de nubes, o en los casos en que falta una fecha. Para resolver este problema, se emplea una técnica de interpolación de regresión lineal, por la que los valores que faltan se sustituyen por valores lineales separados por los puntos de datos válidos más cercanos en los extremos del intervalo de datos que faltan.

Además, ocasionalmente hay valores fuera de la máscara de calidad que no representan una observación válida. Esos valores fuera de la máscara se eliminan de la serie mediante una técnica de suavizado; existen numerosos filtros, pero, en este caso, se aplicó el filtro Savitzky-Golay a la serie temporal puliendo la forma de la señal.

Gráfico, Histograma

Descripción generada automáticamente

Figura 8: Comparación entre una serie bruta y suavizada de NDVI.

## Cálculo de marcadores

En este apartado se explicará de forma resumida las características que se han usado para implementar los marcadores.

### Comparación de perfiles temporales

La comparación de perfiles espectrales se ha usado mucho para el estudio de diferentes usos del suelo, pero al trabajar con series de tiempo, el estudio de perfiles permite observar si la serie ha sufrido algún cambio brusco en su dinámica. Este marcador es el llamado TAM (Temporal Angle Mapper) que calculado respecto a los años obtenidos en la serie de tiempo de NDVI.

### Detección de ciclos y productividad

Para detectar si la parcela a generado un ciclo de productividad vegetal se usan los índices del AR y AS1, que al combinarse es habilitan la posibilidad de discernir entre los periodos de vegetación, vegetación seca y suelo. Esto es útil para determinar durante un año si se ha observado un ciclo vegetación-suelo, infiriendo que existe un aprovechamiento, si se detecta un ciclo dentro de unos límites de días y si la diferencia de los índices supera un umbral.

### Parametrización de umbrales

Para obtener los valores umbrales de los marcadores se ha implementado un sistema de filtros de histograma. Un filtro Gaussiano se implementa iterando sobre el histograma eliminando los datos fuera del intervalo de confianza hasta que no se retiran elementos del conjunto. Dependiendo de la naturaleza del marcador se usan diferentes componentes del intervalo de confianza final:

* TAM: en este caso se utiliza el límite superior del intervalo de confianza ya que se estima que valores superiores a esos las parcelas sufren cambios de dinámica interanual.
* Número de días mínimo del ciclo: tras el filtro se usa como umbral mínimo de días la media de la muestra del número de días.
* Productividad de ciclo: al igual que en el número de días, el umbral de productividad se estima utilizando el valor de la media de la muestra final tras la aplicación del filtro.

Esta parametrización se puede hacer a nivel global para un estudio general, o se pueden realizar parametrizaciones para cada zona de estudio permitiendo un ajuste mayor del análisis.