

# Herramienta de soporte para el análisis de nitrogenado de cultivos

Jerónimo Fernández Álvarez

**Resumen**—La agricultura actualmente está tecnificándose con el objetivo de aumentar el rendimiento. Hoy en día, los sistemas de detección remota se utilizan para conocer distintos indicadores del cultivo (vigor, estrés hídrico, estimación de cosecha) y se obtienen imágenes aéreas multiespectrales que, debidamente procesadas, caracterizan los cultivos. Este proyecto se centra en el desarrollo de una herramienta para estudiar la validez de los índices espectrales para determinar diferencias de fertilización en las diferentes regiones de un cultivo. A partir del estudio del estado del arte y de conversaciones con expertos en el área, se ha diseñado e implementado una herramienta para evaluar regiones de un cultivo a través de índices espectrales. La herramienta permite, de manera sencilla y autónoma, detectar diferencias entre regiones de un cultivo a través de la aplicación de distintos índices espectrales.

**Palabras clave**— Agricultura de precisión, índices de vegetación, índices de nitrogenación, Soil Line, imágenes multiespectrales, capa ráster, capa vectorial, histograma, superposición, discriminación de cultivos, MATLAB, interfaz gráfica, sistema de información geográfica.

**Abstract**—Nowadays, the agriculture is tecnifying with the aim of raising its performance. At the moment, the remote sensing systems are being used so as to survey different crop indicators (vigour, hidric stress, harvest prediction), and multispectral aerial images are obtained, if they are properly processed, to characterize the crops. This project focuses in the design and implementation of a spectral indices validation tool in order to determine fertilization difference between regions of the same crop. Starting off the study of the state of art y discussing with experts of the subject, a spectral index validation tool for regions has being developed. This tool allows, in an easy and autonomous way, to detect differences between crop regions through the applicattion of spectral indices.

**Keywords**— precision agriculture, vegetation indices, nitrogrenation indices, Soil Line, multispectral images, ráster layer, vector layer, histogram, overlap, crop discrimination, matlab, graphical interface, geographic information system.

## 1 INTRODUCCIÓN

Uno de los objetivos de la agricultura, es obtener el máximo rendimiento con el mínimo gasto en recursos. Desde su descubrimiento, la agricultura ha estado estrechamente relacionada a la productividad de alimentos y hambrunas, adquiriendo así un alto grado de importancia para la prosperidad de la humanidad. Con la llegada de nuevas tecnologías, sistemas de detección remotos así como sistemas de posicionamiento global (GPS) pueden ser utilizados para evaluar la variabilidad espacial en los campos de cultivo y obtener diferentes indicadores de rendimiento. Las técnicas habituales, basadas en recoger in-

formación directamente del propio cultivo, acostumbran a ser costosas, requerir mucho tiempo y no son aplicables a todas las regiones. Una alternativa a estas técnicas es la teledetección o “*remote sensing*”.

En los últimos años, se ha aumentado el uso de imágenes hiperespectrales tomadas desde vehículos aéreos tripulados y no tripulados, para calcular tanto el estado del cultivo como para la predicción de su posterior cosecha. Utilizando diversas bandas del espectro electromagnético se puede monitorizar parámetros tan relevantes como la salud del cultivo, la humedad de la tierra que lo aloja, la cantidad de nitrógeno de este e incluso hacer estimaciones de una futura cosecha. Por ejemplo, diversos estudios han demostrado que los cultivos más saludables se caracterizan por una alta absorción de energía del espectro rojo así como una alta reflexión de energía en el espectro cercano al infrarrojo. Así, una imagen con sensibilidad a estos espectros en bandas separadas

- E-mail de contacto: jeronimo.fernandez@e-campus.uab.cat
- Mención realizada: Ingeniería de Computación
- Trabajo tutorizado por: Daniel Ponsa Mussarra
- Curso 2015/16

permite estimar la salud del cultivo.

Este trabajo se centra en el desarrollo de una herramienta para el estudio y análisis de imágenes hiperespectrales de campos de cultivo. La motivación reside en dar soporte a un trabajo realizado en “*Departament d’Agricultura, Ramaderia, Pesca i Alimentació*” de la Generalitat de Catalunya. Su departamento de investigación y desarrollo trata de definir un método para poder diferenciar cultivos que han recibido tratamientos de nitrogenación diferentes. Como punto de partida de este desarrollo, se aplican distintos tratamientos de nitrogenación sobre distintas regiones de un campo de cultivo en un centro de investigación agrónoma. Posteriormente, mediante el uso de vehículos aéreos y cámaras hiperespectrales, se toman imágenes aéreas de estos centros de investigación en diferentes anchos de banda del espectro. Ahora se podrán evaluar las imágenes y correlacionar estos resultados con los datos conocidos del tratamiento aplicado.

Se utilizan índices espectrales, tanto de nitrogenación como de vegetación, para obtener valores relacionados al estado de los cultivos. Un índice espectral no es más que una combinación de la señal recogida en distintos anchos de banda del espectro electromagnético, retornando un único valor que permite caracterizar un determinado estado del cultivo. Hay docenas de índices espectrales relacionados con la vegetación, cada uno de ellos arroja resultados diferentes dependiendo del tipo de cultivo en el que se aplique. Por tanto, es difícil determinar qué índice es bueno para un cultivo, obligando a realizar un análisis amplio de todos ellos para poder extraer alguna conclusión.

La herramienta desarrollada en este proyecto facilita a los usuarios la discriminación entre cultivos a través de la aplicación de índices de vegetación y nitrogenación en una fuente de datos, compuesta por imágenes hiperespectrales.

## 1.1. Objetivos

El objetivo principal de este trabajo es ofrecer una herramienta escalable, automatizada y con una mínima interacción por parte del usuario, para la discriminación del nitrogenado de cultivos mediante la aplicación de índices espectrales. El sistema debe ser capaz de obtener todas las combinaciones posibles de ejecución de los índices, ofreciendo así una colección de resultados que puedan facilitar un posterior análisis.

Para desarrollar la herramienta, es necesario un amplio estudio de los índices espectrales, así como de las imágenes espectrales, fuente elemental de datos. Debido a la envergadura del proyecto y para su correcto desarrollo, se divide en las siguientes fases o subobjetivos:

1. **Estudio del estado del arte:** Análisis de los índices espectrales, tanto de nitrogenación como de vegetación, los más utilizados y los que mejores resultados ofrecen. También se analizaron las posibles alteraciones que puede sufrir su aplicación, debidos a factores externos a la propia vegetación y parámetros para reducir el impacto de estos.

2. **Preparación de la fuente de datos o dataset:** Definición de la estructura de archivos necesaria para poder llevar a cabo un análisis automatizado de las diferencias de fertilización de cultivos. Esta estructura debe ser capaz de soportar diversos tipos de datos. Se realiza mediante el encapsulado en clases.
3. **Implementación de los índices espectrales:** Implementación y validación de los índices de nitrogenación y vegetación. Construcción de un flujo de ejecución capaz de ejecutarlos con diversas combinaciones de bandas afines.
4. **Implementación de métricas:** Implementación de métricas de comparación entre los cultivos evaluados en base a un índice espectral. Definición de un criterio entendible que permita saber cuánta es la diferencia entre ellos.

A continuación se va a explicar con detalle el trabajo realizado. El artículo está organizado en los siguientes apartados:

- **Estado del arte:** Donde se describen los usos de las imágenes multiespectrales en la agricultura.
- **Metodología:** Explicación detallada de la metodología utilizada en el desarrollo de este proyecto.
- **Implementación:** Detalle de la implementación del proyecto, así como su flujo de ejecución y detalle de sus características.
- **Resultados:** Detalle de los resultados obtenidos en el proyecto.
- **Conclusiones:** Exposición de las conclusiones obtenidas a través del desarrollo y los resultados del proyecto.

## 2 ESTADO DEL ARTE

Cada día es más frecuente el uso de imágenes multiespectrales en el campo de la agricultura con el objetivo de obtener datos fiables del estado del cultivo, predecir la cantidad de cosecha y determinar la ubicación de incidencias. En este momento se están realizando múltiples estudios relacionados a recursos agrícolas con diversos objetivos, desde predecir una cosecha, a la detección de problemas en el cultivo, con el objetivo de ayudar al agricultor en la asignación de recursos y a tomar mejores decisiones de mercado.

Con el creciente mercado de UAV, ya no es necesario fletar un vehículo aéreo tripulado con el fin de tomar fotografías aéreas del área de cultivo. Esto reduce costes operacionales, facilitando el acceso a estos métodos a más empresas agrícolas y autónomos. Esto es posible dado que, las últimas mejoras tecnológicas en el campo de la imagen, proporcionan cámaras multiespectrales potentes con un peso asombrosamente bajo, permitiendo que sean montadas en UAV. Otra ventaja de estas imágenes, es que acostumbran a georeferenciarse posteriormente, posibilitando la localización y el cálculo del tamaño de áreas.

Combinando imágenes de distinto ancho de banda del espectro electromagnético, es posible monitorizar y obtener

medidas tanto del estado de la vegetación como del suelo. En el caso del suelo, se puede determinar el contenido de materia orgánica del suelo, ya que un suelo rico en ella acostumbra a ser más oscuro que un suelo arenoso, con bajo contenido de materia orgánica.

En el caso de la vegetación, las imágenes pueden proporcionar datos relevantes al cultivo tales como su salud, detección de cambios de crecimiento, análisis de secciones del suelo donde se ubica, patrones de regadío aplicados, entre otros muchos parámetros. Por ejemplo, una hoja sana, refleja gran parte de la luz en la banda cercana al infrarrojo así como parte de la banda verde. Por el contrario, una hoja estresada se comporta igual que una sana, a excepción de que absorbe más luz en la banda cercana al infrarrojo. De otro modo, una hoja enferma, absorbe gran parte de la luz en la banda cercana al infrarrojo, así como gran parte de la luz en la banda verde [Fig 1].

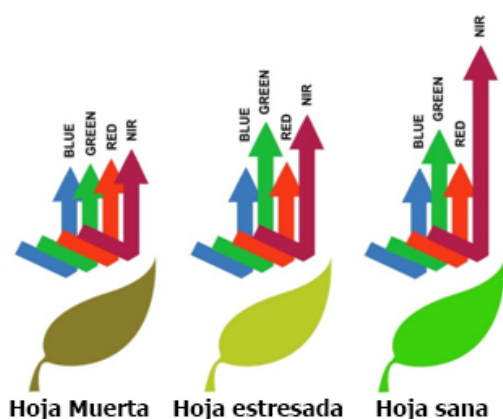


Fig. 1: Reflexión de bandas de espectro, en relación a la salud de una hoja

Así pues, mediante combinaciones de dichas bandas, se puede aproximar el estado de un cultivo. Este es el principio de funcionamiento de los índices espectrales de nitrogenación y vegetación. Los primeros se centran en la concentración de clorofila en la hoja [1] [2] [3], medida a través de la banda verde, obteniendo los valores de esta a través de la combinación de bandas. Los índices de vegetación sin embargo, acostumbran a combinar las bandas más cercanas al rojo y al infrarrojo cercano, con el objetivo de medir la vigorosidad de las hojas de la planta [4]. Hay docenas de índices espectrales, cada uno formulado de manera concreta, en función de las bandas con las que se realiza el cálculo. Sin embargo, hay factores externos que al afectar a la luz, afectan irremediablemente al resultado del índice espectral. Entre estos factores, está la forma de la planta y de su hoja, así como su suelo y su contenido de humedad. Además, la reflexión de la propia atmósfera terrestre es un problema si estas imágenes se toman desde un satélite artificial.

Para reducir estas alteraciones, algunos índices cuentan con parámetros de corrección. Algunos son simples valores numéricos, que se ajustan en función del tipo de planta o vienen definidos por el propio índice. Otros parámetros, dependen del tipo de suelo al desnudo donde se encuentra el cultivo alojado, que se estima con un parámetro conocido como recta Soil Line o línea de suelo [5]. Se llama así porque se caracteriza a partir de estimar la ecuación de una recta que, en un gráfico de correlación entre el espectro rojo

y el cercano al infrarrojo, separa los valores del suelo de los de la vegetación [6].

Esta se puede calcular en base a una recta, que delimita la distribución de un gráfico de puntos del espectro rojo contra el espectro cercano al infrarrojo [Fig 2].

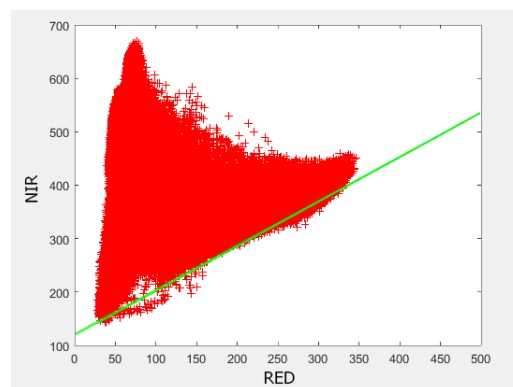


Fig. 2: Gráfica de correlación entre bandas de rojo y cercano al infrarrojo y recta Soil Line.

Por ejemplo, uno de los índices de vegetación más extendidos, es el NDVI (Normalized Difference Vegetation Index) [Fig 3]. Este se basa en la diferencia de absorción entre el espectro rojo visible y del cercano al infrarrojo, sin ningún parámetro de corrección. Por su definición, los valores que retorna se encuentra entre el -1.0 y +1.0. Generalmente, los valores cercanos a -1 corresponden a agua, así como los cercanos a 0 representan suelo rocoso o arenoso. Los valores positivos hasta 0.5 corresponderían a arbustos no muy densos y los más cercanos al 1 se darían en bosques tropicales o vegetación densa. Sin embargo, NDVI se ve afectado por factores atmosféricos, debido a la composición de la propia atmósfera, por la humedad del suelo donde se encuentra el cultivo y el tipo de reflexión causada por la forma de la planta.

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

Fig. 3: Índice de vegetación NDVI (Normalized Difference Vegetation Index).

Por otro lado, el índice de vegetación TSAVI (Transformed Soil Adjusted Vegetation Index) también se basa en las bandas rojo del espectro visible y la banda cercana al infrarrojo, como NDVI. Sin embargo, este incluye dos parámetros de corrección para minimizar el impacto de la influencia del suelo. El primero de ambos, el parámetro 'x', nominal y válido para todos los casos ya que su valor 0.08, viene dado por la propia definición del índice. El segundo, se trata en realidad de dos parámetros que provienen de la misma fuente, la ecuación de la recta Soil Line, donde 'b' es el intercepto y 'a' la pendiente. Esta ecuación, dependerá directamente del suelo del cultivo que se esté evaluando con el índice. Nótese que, obviando el parámetro x (o asignándole el valor 0), con a=1 y b=0, el índice de vegetación TSAVI pasa a ser el NDVI. Los valores 0 retornados por TSAVI corresponden al suelo desnudo, aumentando el valor conforme la densidad de la vegetación es mayor.

Aún con todas las variedades de índices espectrales y los

$$TSAVI = \frac{[\alpha(\rho_{nir} - \alpha \times \rho_{nir} - b)]}{[(\rho_{red} + \alpha \times \rho_{nir} - \alpha \times b + x(1 + \alpha^2))]}$$

Fig. 4: Índice de vegetación TSAVI (Transformed Soil Adjustment Vegetation Index).

parámetros de corrección, no hay una norma general que determine qué índice es mejor para que tipo de cultivo. Para obtener conclusiones y/o mediciones precisas, es necesario evaluar todos ellos sobre distintos cultivos o áreas, interpretar los resultados y determinar qué índice es el más apropiado para el fin que se desea evaluar.

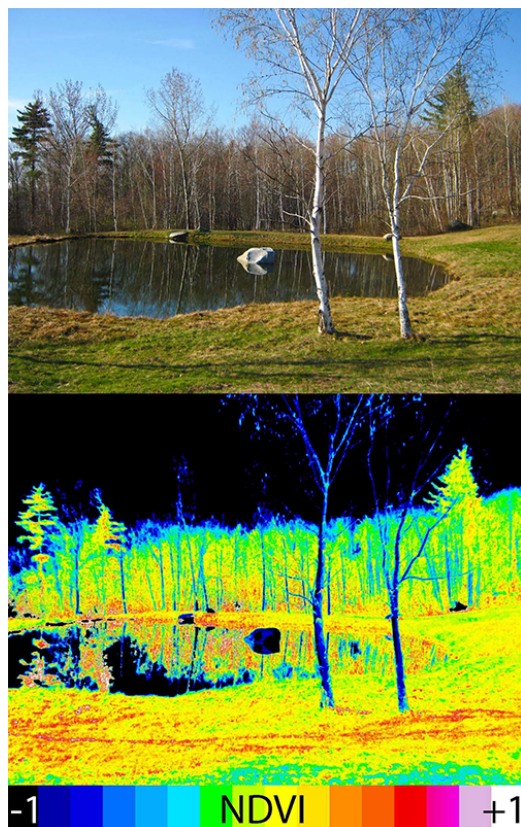


Fig. 5: Imágen RGB (arriba) e imágen NDVI (Normalized Difference Vegetation Index). Notesé la diferencia entre elementos no fotosintéticos (troncos, rocas, hierba seca, ...) y la vegetación sana, con valores entre 0,1 y 0,9.

### 3 METODOLOGÍA

Se utilizó un desarrollo de software basado en iteraciones e incremental. Con este método, se pretende ir incorporando nuevas funcionalidades al software de una forma escalonada. Todos los resultados fueron analizados y validados posteriormente a su obtención. Además se realizaron reuniones semanales siguiendo el método Scrum, para determinar los objetivos de cada semana.

El entorno de desarrollo escogido ha sido MATLAB en su versión 2015a. La decisión fue tomada debido a las herramientas de procesamiento de imágenes georreferenciadas que dispone [7], facilitando el tratamiento de este tipo de datos. Además, debido a que el grueso de proyecto se basa

en operaciones de matrices, este entorno está especialmente optimizado para trabajar con este tipo de datos.

Para la implementación se utiliza la filosofía de programación orientada a objetos. Esto facilita el mantenimiento, ampliación y comprensión del código fuente, habilitando la opción de una futura expansión o línea de continuación. También, por la propia naturaleza de los índices espectrales, dado que aparecen nuevas versiones, versiones combinadas o variaciones de estos, este método permite acoplar un nuevo índice de manera rápida sin comprometer al resto del sistema.

La fuente de datos que evaluará el software, está compuesto por diversos componentes de un sistema de información geográfica o GIS [8]. Como pieza clave, tenemos las imágenes multispectrales. Estas pueden ser un archivo separado en capas, donde cada una contiene una banda, o diversos archivos, cada uno con una capa y una banda del espectro diferente. Estas imágenes geotiff, denominadas capas raster [9] en un entorno GIS, están geolocalizadas de manera que, se puede relacionar información conocida del terreno a través de capas vectoriales, que permiten delimitar distintas áreas de cultivo dentro de la misma imagen. La capa vectorial está almacenada en un archivo Shape [10]. Mediante la combinación del archivo de capas vectoriales con las imágenes [Fig 6], se pueden obtener las distintas áreas de la imagen a evaluar en imágenes independientes, lo cual permitirá comparar los resultados.

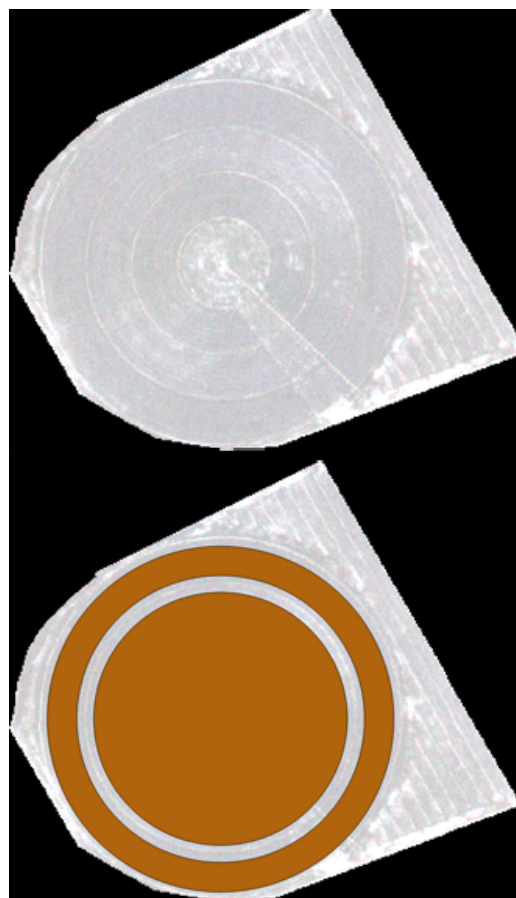


Fig. 6: Arriba: Imágen Ráster. Abajo: Imágen Ráster con capa vectorial superpuesta

El sistema está preparado para soportar múltiples fuentes



de datos que mantengan esos requisitos. Sin embargo, una pequeña configuración por parte del usuario es necesaria para poder ejecutar la evaluación con éxito.

El usuario debe, dentro de un fichero de configuración, especificar por cada dataset las bandas disponibles de este. Esto lo hará colocando el ancho de banda de cada imagen en el orden que se encuentra en la fuente de datos. Además, es necesario poner una etiqueta a las diferentes capas de la capa vectorial, en orden de aparición. Esta etiqueta puede relacionar áreas, bien ya sea porque tienen el mismo tratamiento, o porque el usuario desea relacionarlas por otra razón. La razón de esta etiqueta es para poder ofrecer información de relación entre áreas una vez realizados los cálculos (Información más detallada de este proceso se puede consultar en el Apéndice A-1).

## 4 IMPLEMENTACIÓN

Con el objetivo de tener un código escalable y de fácil mantenimiento, la implementación del software se basa en la filosofía de Programación Orientada a Objetos. Recientemente MATLAB, tras verse abrumado por sus competidores, ha mejorado su sintaxis en cuanto a esta filosofía de programación, haciéndolo más potente y ofreciendo ayudas a los desarrolladores. El proyecto al completo se compone de diversas clases, algunas no relacionadas directamente con el tema principal de este. Estas se utilizan como soporte para realizar operaciones de mejora del rendimiento y/o cálculos de soporte, como por ejemplo el cálculo de MD5 [11], poco óptimo en MATLAB, o la gestión de la memoria cache, independiente del tipo de programa que se desea desarrollar.

El grueso principal del programa se compone de 4 clases:

- **Dataset:** Clase raíz de la fuente de datos que utilizará el programa. En ella se almacena toda la información relacionada a las fuentes de datos, desde los datos de las propias imágenes hasta las rutas de los archivos de capa vectorial.

A través de un constructor recibe los archivos de la capa ráster, la capa vectorial de regiones y la de Soil Line, opcionalmente. A continuación, lee automáticamente el archivo de configuración del dataset, generando las estructuras de datos en función del contenido de este. Puede generar los datos ráster a través de un solo archivo multicapa o de múltiples archivos de una sola capa.

Dentro de sus estructuras de datos, también contiene las estructuras de control que determinarán que operaciones se realizarán sobre este durante la ejecución del programa.

- **FixedSizeDatasetSampler:** Trabajar con imágenes puede ser un proceso lento, puesto que dependiendo de la resolución de estas, la cantidad de datos a procesar puede ser enorme. Para poder llevar a cabo ejecuciones más rápidas con muestreos más pequeños,

se creó esta clase.

Todas sus operaciones se basan en el objeto Dataset que contenga. Este se recibe por parámetro en el constructor, junto con el número de píxeles que se desea evaluar posteriormente. En el caso de que este valor sea un 0, se evaluará la imagen en su totalidad.

En esta clase es donde se obtiene la estructura final de datos que se evaluará en el índice espectral. En su formato original, la imagen se trata de una matriz tridimensional, donde la tercera dimensión contiene las diferentes bandas. Por motivos de rendimiento y de estandarización dentro de la aplicación, este objeto transforma cada imagen en un vector columna, generando una matriz donde cada columna será una imagen del archivo ráster [Fig 7].

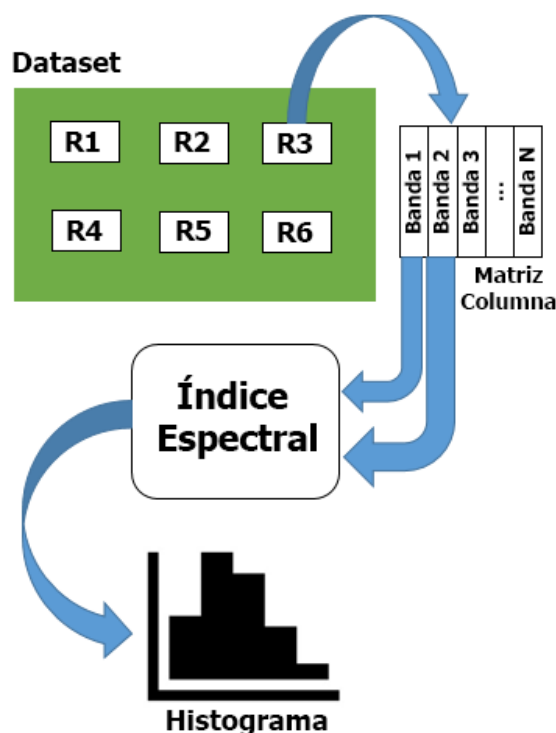


Fig. 7: Recorrido de los datos desde el Dataset hasta el histograma

Cabe decir que este muestreo se realizará por cada región. Es en esta clase donde se procesan las regiones de la imagen, para posteriormente aplicar el muestreo, que consiste en coger un número N de píxeles al azar.

Finalmente, esta clase se encarga también del proceso de cálculo de la recta Soil Line. A pesar de que este, es un parámetro de índices de vegetación, sus valores están intrínsecamente ligados al dataset. Usando una capa vectorial para este propósito, se aplica una búsqueda de mínimos por intervalos en un gráfico de correlación bandas. Posteriormente, a través de una regresión lineal, se obtiene la ecuación de la recta. Además, esto se hace para todas las combinaciones posibles dentro de la definición de Soil Line. Este procedimiento también se puede llevar a cabo realizando una regresión cuantílica, dejando

el 99 % de la muestra por encima de la recta. Sin embargo, a pesar de que la aplicación también lo tiene implementado, se optó por el método de mínimos por intervalo debido a su mayor velocidad de cómputo [12].

- **SpectralIndex:** Clase abstracta que define la mínima estructura necesaria para un índice espectral, que se pretende ejecutar posteriormente en el programa. Todos los índices espectrales, ya sean de vegetación o nitrogenación, son derivados de esta clase.

Una vez derivada, cada índice contendrá la fórmula o cálculo necesario para su ejecución. Será necesario especificar por cada índice espectral el rango de bandas aceptadas en formato de matriz columna. Esta información se proporcionará luego al controlador en el momento de comprobar si es posible su ejecución con el dataset evaluado.

Opcionalmente y dependiendo del índice, se especificará si este requiere algún parámetro de corrección o si se trata sencillamente de un coeficiente. En caso de necesitar un parámetro dependiente del dataset como la Soil Line, se especificará en esta estructura de control también.

- **SpectralIndexEvaluator:** Clase que combina las anteriores para realizar el cálculo final. El controlador ejecutará su método de evaluación con un objeto `FixedSizeDatasetSampler`, un objeto `SpectralIndex`, las bandas afines entre ellos y los parámetros de corrección, de ser necesarios. El índice espectral determina que bandas de `FixedSizeDatasetSampler` le son afines, y luego calcula el valor del índice para las distintas combinaciones válidas.

Durante su ejecución, obtendrá los bandas afines de las regiones del Dataset, ejecutando el índice espectral por cada una de ellas. En caso de que el parámetro de corrección de Soil Line sea necesario, se le solicitará al Dataset .

Para caracterizar cada región, se generarán los histogramas normalizados de la imagen del índice calculado, mostrando la variabilidad de sus valores y la distribución de estos. Al mismo tiempo, se construye una matriz identidad de  $N \times N$ , donde  $N$  es el número de regiones procesadas.

Para cada pareja de regiones, se valora la similitud de sus histogramas aplicando el coeficiente de Bhattacharyya [13]. Este coeficiente trata de medir la cantidad de solapamiento entre dos conjuntos, con lo cual, puede ser utilizado para determinar la proximidad relativa entre ellos. Para calcular el coeficiente de Bhattacharyya, se deben especificar los intervalos de valores escogidos en la muestra y el número de miembros que los componen [Fig 8], que en este caso son el número de barras del histograma y los valores de estas, respectivamente. Si el resultado de aplicar el coeficiente es 0, no existe superposición alguna entre los dos conjuntos. En cambio, si el resultado es 1, los

conjuntos tienen una superposición total.

$$BC(p,q) = \sum_{i=1}^n \sqrt{p_i q_i}$$

Fig. 8: Fórmula del coeficiente de Bhattacharyya.

Con la cantidad de superposición de cada pareja de histogramas calculada, se rellena la diagonal superior de la matriz identidad. No se rellena la diagonal inferior, ya que sería idéntica a esta última. Esta matriz será donde un experto centrará su análisis, buscando el mínimo valor de superposición entre las regiones. Un valor bajo de superposición, indicará que hay mayor diferencia entre esas regiones y que por lo tanto, el índice aplicado es más sensible a ese tipo de cultivo, permitiendo de esa manera, detectar diferencias con más eficacia.

#### 4.1. Secuencia de ejecución

A continuación se muestra el pseudocódigo del hilo de ejecución del programa:

```

1  Seleccionar datasets a procesar
2  Seleccionar índices espectrales a evaluar

4  Para cada dataset
5      Muestrear datos en regiones
        delimitadas

7  Para cada índice espectral
8      Seleccionar bandas afines del dataset
        para el índice espectral
9      Obtener todas las combinaciones de
        bandas afines
10     Calcular índice espectral para cada
        región delimitada
11     Estimar el número de barras del
        histograma
12     Generar un histograma por cada región
        delimitada
13     Calcular superposición por de los
        histogramas por cada pareja de
        regiones
14     Registrar los datos en una matriz
15     Guardar datos en fichero caché y
        generar entrada en fichero índice

```

#### 4.2. Interfaz gráfica de Análisis

La ejecución del programa no está pensada para retornar los resultados de manera que un usuario pueda identificar rápidamente lo que visualiza, puesto que no se encuentra dentro de su especificación. La idea es procesar una determinada selección de datasets y índices, registrando los resultados obtenidos en disco.

Sin embargo, el objetivo del proyecto es facilitar a este una serie de cálculos con el fin de poder ser analizados, para poder generar conclusiones o detectar comportamientos anormales en las regiones de los cultivos evaluados. Por este motivo, como complemento a la herramienta de cómputo principal, se ha implementado una interfaz gráfica de

análisis. En ella, el usuario no solo podrá ver el coeficiente de superposición de los histogramas de las regiones, sino la propia distribución de estos y su superposición en un diagrama de barras.

La interfaz gráfica, carga datos de disco recuperando cualquier cálculo realizado anteriormente guardado en los ficheros caché. No se podrán ejecutar nuevos cálculos a través de esta herramienta.

Ofreciendo al usuario 3 inputs del tipo desplegable, seleccionando un dataset, el sistema completará el segundo desplegable con los índices que han sido calculados sobre este. Una vez seleccionado un índice, se obtendrán todas las combinaciones de bandas que fueron calculadas sobre este [Fig 9].

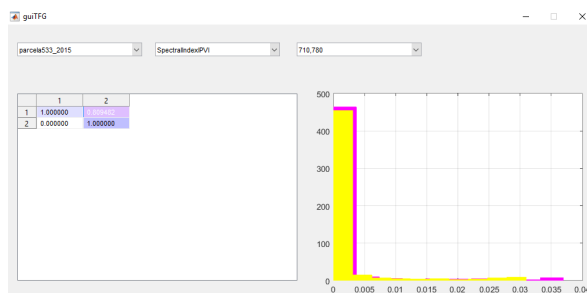


Fig. 9: Pantalla principal de la interfaz gráfica

Finalmente, seleccionando una combinación de bandas, se mostrará la matriz con el coeficiente de superposición por cada par de ellas. En este punto, el usuario ya obtiene la información básica que calcula el programa, pero tiene la opción de ir más allá. Haciendo clic sobre una casilla con valor, la interfaz mostrará en su parte derecha los histogramas de las regiones relacionadas a dicha casilla. Diferenciados por colores, el usuario podrá visualizar la distribución de sus histogramas, detectando picos, desplazamientos, etc [Fig 10].

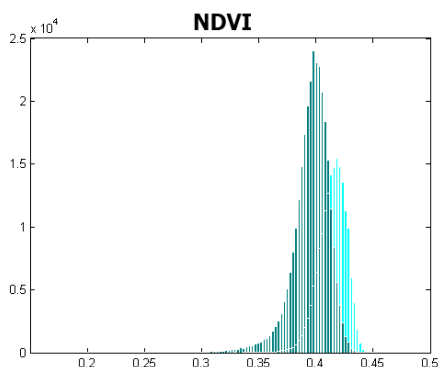


Fig. 10: Comparación de los histogramas de las imágenes de dos regiones procesadas con NDVI.

## 5 RESULTADOS

Como resultado principal del proyecto, se ha diseñado e implementado un software escalable, mantenible y basado en objetos, para evaluar imágenes multispectrales de cultivos con índices de vegetación y nitrogenación. Cuenta con un programa de proceso, y de una herramienta de visualización interactiva.

Se ha definido un protocolo para estandarizar el registro de datos en el sistema. Mediante este protocolo, se puede adaptar cualquier fuente de datos, basada en capas ráster y vectoriales, para su ejecución en el sistema.

La aplicación permite a un usuario final analizar la comparación la cantidad de superposición de histogramas entre regiones. Además, mediante el apoyo de una interfaz gráfica, el usuario puede recuperar datos ya procesados y ver las distribuciones y su superposición mediante un gráfico de barras.

### 5.1. Ejemplo de análisis

A continuación se procede a analizar un caso real de discriminación de nitrogenado de regiones de un cultivo.

Para llevar a cabo el análisis se utilizará el dataset de un campo de cultivo de maíz con dos regiones, llamado "Parcela 533". Este tiene una región circular interior rodeada por otra región anular, la cual tiene un tratamiento de nitrogenado o fertilización más elevado. Como se puede apreciar en la Figura 6, ya disponemos de una capa vectorial que nos diferencia las regiones dentro de la zona de cultivo.

En este caso, su capa ráster nos ofrece seis anchos de banda del espectro diferentes:

- 550nm, correspondiente al verde visible
- 670nm, correspondiente al naranja visible
- 700nm, correspondiente al rojo o el borde rojo visible
- 710nm, correspondiente al rojo o el borde rojo visible
- 730nm, correspondiente al cercano al infrarrojo
- 780nm, correspondiente al cercano al infrarrojo

Para evaluar este dataset, vamos a utilizar los siguientes índices espectrales (una lista de todos los implementados en el programa se puede consultar en el Apéndice A-2) con todas las combinaciones de bandas afines posibles:

- Normalized Difference Vegetation Index (NDVI)
- Transformed Soil Adjustment Vegetation Index (TSAVI)
- Modified Chlorophyll Absorption in Reflectance Index (MCARI)

En este análisis, tanto para el NDVI como para el TSAVI, el único valor de cercano al infrarrojo aceptado será el ancho

de banda de 780nm. Como rojo se aceptará cualquier ancho de banda entre los 670nm y los 730nm. Para el MCARI, único índice de nitrogenación propuesto para este análisis, se utilizará como verde el ancho de banda de 550nm, 670nm como borde del rojo o “red edge” y 700nm como rojo visible. No existe la posibilidad de hacer más de una combinación de bandas para este último índice de nitrogenación.

Como último apunte antes del análisis, NDVI y MCARI no precisan de parámetros externos, a diferencia de TSAVI que sí lo hace. A su parámetro de corrección ‘x’ le asignaremos el valor recomendado en su definición de 0,08. Para el parámetro Soil Line, se calculará la ecuación de la recta para cada pareja de bandas afines.

Una vez analizados estos índices espectrales sobre el dataset propuesto, se procede a analizar los resultados. Se ha generado una salida de programa, basada en la cantidad de superposición de los histogramas de regiones, tras aplicar un índice espectral determinado, un ejemplo resumido del resultado que retorna el programa se muestra en la siguiente Tabla 1.

Índice espectral y ancho de banda	Solapamiento
NDVI $\rho_{670nm}/\rho_{780nm}$	0.9596
NDVI $\rho_{700nm}/\rho_{780nm}$	0.8870
NDVI $\rho_{710nm}/\rho_{780nm}$	0.8094
NDVI $\rho_{730nm}/\rho_{780nm}$	0.7740
TSAVI $\rho_{670nm}/\rho_{780nm}$	0.9250
TSAVI $\rho_{700nm}/\rho_{780nm}$	0.8912
TSAVI $\rho_{710nm}/\rho_{780nm}$	0.8316
TSAVI $\rho_{730nm}/\rho_{780nm}$	0.7661
MCARI $\rho_{550nm}/\rho_{670nm}/\rho_{700nm}$	0.9122

Tabla 1: Resultados retornados por el programa para el análisis propuesto.

A partir de los resultados de la tabla se puede concluir que, para este cultivo, el índice de nitrogenación MCARI no es muy útil, ya que solo puede calcularse para una combinación de bandas y esta a su vez, retorna un valor de solapamiento muy alto entre los histogramas de sus regiones.

En cambio, en los índices de vegetación NDVI y TSAVI, se observa un patrón similar en cuanto a la tendencia de los valores en función de las bandas evaluadas. Conforme la banda del rojo se acerca a valores de la banda cercana al infrarrojo, la superposición entre los histogramas de las regiones disminuye, indicando una mayor diferencia entre ellos. El valor más bajo de la tabla de resultados, lo retorna

TSAVI, utilizando como rojo el ancho de banda de 730nm y como cercano al infrarrojo el ancho de banda de 780nm. La diferencia entre los valores mínimos de NDVI y TSAVI es de un 0,0079, lo cual nos indica que los factores de corrección de TSAVI no han sido muy eficientes. Una causa de esto, podría deberse a que no hay ninguna sección de suelo desnudo en el dataset, lo cual ha obligado a calcular el parámetro Soil Line sobre la propia vegetación, disminuyendo o anulando su eficacia. De ser así, los valores de la recta de Soil Line han hecho que los valores de TSAVI tiendan a ser los del NDVI, ya que para  $a=1$  y  $b=0$ ,  $TSAVI = NDVI$ .

A partir del análisis realizado, en base a los resultados anteriores, se puede concluir que:

- De los índices utilizados, NDVI y TSAVI son los mejores para detectar diferencias entre regiones en este cultivo.
- Las diferencias se incrementan cuando el rojo escogido tiene un ancho de banda más cercano al ancho de banda cercano al infrarrojo.
- Se podría tratar de mejorar los resultados del índice de vegetación TSAVI calculando la recta Soil Line a partir una región de suelo desnudo o sin vegetación.

## 6 CONCLUSIONES

A lo largo del desarrollo de este proyecto, se ha construido una aplicación que permite a un usuario analizar la discriminación del nitrogenado de los cultivos, mediante imágenes multiespectrales. Se han implementado diversos índices espectrales, tras un amplio estudio del estado de arte, seleccionando los más utilizados del momento y que mejores resultados proporcionan a la comunidad de investigación. Además, se ha diseñado el software de manera que sea escalable y de fácil mantenimiento, ya que cada día se proponen nuevos índices espectrales y nuevas bandas en las cámaras multiespectrales.

Se ha estudiado el estado del arte del uso de imágenes en la agricultura, conociendo así el amplio abanico de posibilidades que este entorno brinda, analizando el uso de las imágenes espectrales en él y indagando sobre nuevos conceptos que pueden ayudar a mejorar el cálculo de variables asociadas a cultivos.

Se han cumplido la gran parte de los objetivos establecidos al inicio del proyecto, y a pesar de los cambios de planificación necesarios debido a la ampliación del ámbito de este, se ha logrado obtener la herramienta final que se propuso al inicio. Esta herramienta permite a expertos del sector de la agricultura hacer análisis y sacar conclusiones de sus campos de investigación, determinando así que índices espectrales funcionan mejor en las diferentes especies de plantas, que bandas ofrecen mejores resultados, etc. Incluso pueden ver las distribuciones de los histogramas de las regiones, en superposición con otras regiones, mediante una interfaz gráfica de usuario, detectando así desplazamientos,



similitudes o diferencias.

Sin embargo, algunos puntos no se han podido realizar, debido a la falta de tiempo. Como punto pendiente, se propuso crear un índice de vegetación propio, tratando de separar muestras de píxeles mediante SVM [14] o LDA [15]. Este se descartó, invirtiendo el tiempo en la implementación del software y en sus mejoras.

## 6.1. Líneas de continuación

A pesar de que la aplicación es funcional, hay una serie de mejoras e incorporaciones que pueden ayudar al usuario final.

Primeramente y para una mejor experiencia de usuario, sería interesante hacer una interfaz de usuario avanzada. De este modo, se evitaría al usuario la interacción con el código de manera definitiva, siempre y cuando no desee añadir un nuevo índice espectral o hacer alguna modificación mayor. Además, se podrían incluir más elementos visuales a banda de la superposición de histogramas que ahora se muestra.

Por otro lado, se podría implementar una memoria caché con una granularidad más fina. En este momento, si se añade un índice a una lista de índices precalculada para un dataset, el sistema volverá a realizar todos los cálculos, sin aprovechar los ya realizados.

Como añadido, se podría dotar al usuario de un método interactivo de entrada usando el clic del ratón, para seleccionar una zona de suelo vacío de la imagen. Esto nos servirá para calcular el parámetro corrector de Soil Line, en caso de falta del archivo vectorial asociado.

## AGRACEDICIMIENTOS

Me gustaría agradecer principalmente a mi tutor, Daniel Ponsa Mussarra, las ayudas y consejos que me ha dado durante la realización de este proyecto. Gracias a ellos, se ha logrado cumplir los objetivos básicos propuestos y el resultado ha sido mejor de lo esperado.

Este trabajo se ha realizado con el soporte de los proyectos FireDMMI (TIN2014-56919-C3-2-R) y LIFE+FUTUR AGRARI (LIFE 12-ENV/ES/000647).

## REFERENCIAS

- [1] L. Dextrazec D. Haboudanea; J. R. Millera; N. Tremblay; P. J. Zarco-Tejadad. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, (81):416–426, 2009.
- [2] C. Wu; Z. Niu; Q. Tang; W. Huang. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and forest meteorology*, (7):1230–1241, 2008.
- [3] E. R. Hunt Jr.; P. C. Doraiswamy; J. E. McMurtreya; C. S.T. Daughtry; E. M. Perry; B. Akhmedova. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Remote Sensing of Environment*, pages 3468–3478, 2011.
- [4] A. Viña; A. A. Gitelson; A. L. Nguy-Tobertsom; Y. Peng. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment*, pages 3468–3478, 2011.
- [5] F. Bareta; S. Jacquemouda; J. F. Hanocqa. The soil line concept in remote sensing. *Remote Sensing Reviews*, (7):65–82, 2009.
- [6] H. Yoshioka; T. Miura; J. A. M. Demattê; K. Batchily; A. R. Huete. Soil line influences on two-band vegetation indices and vegetation isolines: A numerical study. *Remote Sensing of Environment*, (2):545–561, 2010.
- [7] Mathworks. Image processing toolbox. <http://es.mathworks.com/products/image/>, Fecha de último acceso: 20/02/2016.
- [8] Wikipedia. (geographic information system). [https://en.wikipedia.org/wiki/Geographic\\_information\\_system](https://en.wikipedia.org/wiki/Geographic_information_system), Fecha de último acceso: 18/03/2016.
- [9] U.S. government work. (raster product format). <http://www.digitalpreservation.gov/formats/fdd/fdd000298.shtml>, Fecha de último acceso: 18/03/2016.
- [10] Environmental Systems Research Institute. Esri shapefile technical description. *An ESRI White Paper*, 1998.
- [11] Z. Yong-Xia; Z. Ge. Md5 research. *Multimedia and Information Technology (MMIT), 2010 Second International Conference*, (15):271–273, 2010.
- [12] G. A. Fox; G. J. Sabbagh; S. W. Searcy; C. Yang. An automated soil line identification. *Routine for Remotely Sensed Images. Soil Science Society of America*, 2007.
- [13] N. N. Soe. Image matching scheme by using bhattacharyya coefficient algorithm. *International Journal of Innovative Research in Computer and Communication Engineering*, (3):232–243, 2015.
- [14] N. Cristianini; J. Shawe-Taylor. An introduction to support vector machines and other kernel-based learning methods. *Cambridge University Press*, 2000.
- [15] T. Li; S. Zhu; M. Ogihara. Using discriminant analysis for multi-class classification: An experimental investigation. *Knowledge and Information Systems*, (10):453–472, 2006.

## APÉNDICE

### A.1. Instalación de un dataset

A continuación se especifican los pasos para la construcción de un dataset aceptado por la aplicación:

1. Dentro de un directorio, colocar el archivo Shape de capa vectorial.
2. En el mismo directorio, colocar el archivo ráster. Puede ser un solo archivo, o múltiples archivos. En este último caso, crear un nuevo directorio dentro del anterior.
3. En el mismo directorio del archivo de capa vectorial (Shape File). Se crea un documento de texto TXT, con el nombre "settings.txt" como muestra la figura 3.. Ahora se procede a editar su contenido:

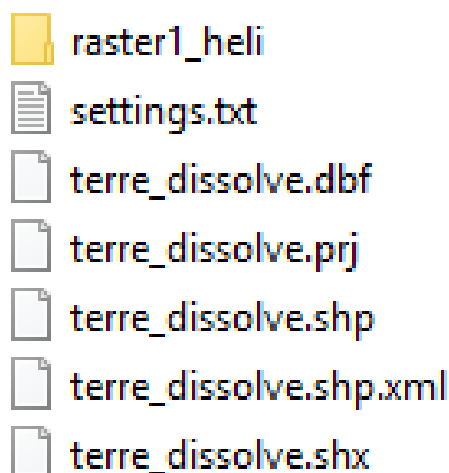


Fig. 11: Ejemplo de contenido de una carpeta de Dataset

- 3.1. En la primera línea, se introducirán por orden de aparición y con un espacio de separación, el ancho de banda de todas las bandas disponibles en el archivo ráster. En caso de un archivo, el orden lo determina la profundidad de la matriz, siendo primera la banda menos profunda y última la que tenga un índice de acceso mayor en esa profundidad. En caso de que sea un directorio con múltiples archivos, estas se introducirán en función del orden alfabético de los archivos allí alojados.
- 3.2. En la segunda línea, se introduce un identificador por cada región o capa del archivo vectorial. Deben de ser identificadores numéricos [3.2.]. Por ejemplo, si dentro de 4 regiones, las pares han sido tratadas diferentes a las impares, pero iguales entre sí, colocaremos: 1 2 1 2.

```
780 700 550 670 710 730
1 2 3 4 3 1 4 2 2 4 1 3
```

Fig. 12: Ejemplo de contenido del archivo settings.txt

4. Accedemos al controlador de la aplicación.
5. Ahora por cada dataset, procederemos a editar los atributos rasterFileName, shapeFileName y en caso de ser necesario soilShapeFileName. Recordemos que la ubicación del fichero settings.txt, ha de ser la misma que la de shapeFileName.
6. Llegados a este punto, estamos listos para evaluar el dataset recién instalado.

## A.2. Listado de índices espectrales de vegetación y nitrogenación implementados

Índice Espectral	Fórmula
Infrared Percentage Vegetation Index	$IPVI = \frac{\rho_{nir}}{\rho_{nir} + \rho_{red}}$
Perpendicular Vegetation Index	$PVI = \frac{1}{\sqrt{\alpha^2 + 1}} ((\rho_{nir} - \alpha \rho_{red} - \beta))$
Normalized Difference Vegetation Index	$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$
Normalized Difference Red Edge Index	$NDRE = \frac{\rho_{nir} - \rho_{red\ edge}}{\rho_{nir} + \rho_{red\ edge}}$
Optimized Soil Adjusted Vegetation Index	$OSAVI = (1,5) \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red} + 0,16}$
Soil Adjusted Vegetation Index	$SAVI = (1 + L) \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red} + L}$
Transformed Soil Adjusted Vegetation Index	$TSAVI = \frac{\alpha(\rho_{nir} - \alpha \rho_{red} - \beta)}{\alpha \rho_{nir} + \rho_{red} - \alpha \beta + \chi(1 + \alpha^2)}$
Modified Soil Adjusted Vegetation Index 2	$MSAVI2 = \frac{(2\rho_{nir} + 1 - \sqrt{(2\rho_{nir} + 1)^2 - 8(\rho_{nir} - \rho_{red})})}{2}$
Transformed Chlorophyll Absorption in Reflectance Index	$TCARI = 3((\rho_{red\ edge} - \rho_{red}) - 0,2(\rho_{red\ edge} - \rho_{green}))(\frac{\rho_{red\ edge}}{\rho_{red}})$
Modified Chlorophyll Absorption Ratio Index	$MCARI = (\rho_{red\ edge} - \rho_{red}) - 0,2(\rho_{red\ edge} - \rho_{green})(\frac{\rho_{red\ edge}}{\rho_{red}})$

Tabla 2: Listado de índices de vegetación y de nitrogenación implementados en el programa por defecto.