Instrucciones para la clasificación orientada a objetos usando código abierto

Un tutorial sobre cómo usar el software de código abierto, la aplicación ORFEO en QGIS y datos de satélite de percepción remota para realizar clasificaciones orientadas a objetos.

Por: Miguel A. Pavon, MAP

Para

Desert Landscape Conservation Cooperative

Para ayuda con clasificaciones, OTB o QGIS, visite las referencias en línea:

- https://www.orfeo-toolbox.org/documentation/
- https://www.qgis.org/en/site/forusers/index.html

Ayuda adicional puede ser proporcionada contactando a Miguel Pavon en pavonma@hotmail.com o (512) 466-3936.

Cita sugerida: Pavon, Miguel A., Instrucciones para la clasificación orientada a objetos usando código abierto, 2017.

Contenido.

1. Descargue e instale el software QGIS para su plataforma 2
   1.1 Descargar Software QGIS 2
   1.2 Instalar el software QGIS 3
   1.3 Verificar la instalación de su software QGIS 10
2. Selecciona y descarga una imagen en Earth Explorer para usarla en tu Clasificación Orientada a Objetos 13
   2.1 Landsat 8 13
   2.2 Crea un login en Earth Explorer 13
   2.3 Busque y seleccione los datos que necesita 14
   2.4 Descargue sus datos 18
3. Clasificación orientada a objetos con QGIS y ORFEO Toolbox 22
   3.1 Polígonos de referencia y pruebas de campo 21
   3.2 Crear la pila de imágenes 24
   3.3 Segmentación vectorial y exploración de datos 28
   3.4 Segmentación ráster y clasificación orientada a objetos 35
      3.4.1 Flujo de trabajo de segmentación de cambio medio 36
      3.4.2 Preparación de los datos de referencia 41
      3.4.3 Clasificación basada en objetos utilizando el algoritmo SVM 42
      3.4.4 Evaluación de la precisión 49

Instrucciones para la clasificación orientada a objetos usando código abierto

Este conjunto de instrucciones fue compilado por Miguel A. Pavón, MAP, trabajando con Bird Conservancy of the Rockies para la Cooperativa de Conservación del Paisaje del Desierto (DLCC). La justificación detrás de esto es poner a disposición un tutorial para la clasificación de la cubierta de tierra que se basa en datos y herramientas de fácil acceso que todo el mundo puede usar para mejorar las prácticas de conservación, planificación y manejo de la tierra. La clasificación de imágenes orientada a objetos implica la identificación de objetos de imagen, o segmentos, que están compuestos de píxeles vecinos (espacialmente contiguos) de contenido espectral similar (color). El caso en la mano utiliza software QGIS de código abierto, la aplicación ORFEO en QGIS y los datos de satélite de Landsat puestos a disposición por USGS a través de Earth Explorer.

Este tutorial se divide en tres secciones, la primera es la instalación del software QGIS, la segunda es la descarga de datos (Landsat) de Earth Explorer y, por último, instrucciones sobre cómo utilizar ORFEO dentro de QGIS para realizar una clasificación orientada a objetos en la imagen Landsat.

QGIS es un Sistema de Información Geográfica que se ejecuta en múltiples plataformas, por lo que es ideal para los nuevos usuarios y partes interesadas de la DLCC para utilizar en la región. Es fácil de usar y bien documentado, aunque tiene una gran variedad de herramientas que pueden hacer que se vea un reto, pero cuanto más tiempo pasas usando, más intuitivo se hace.

Earth Explorer es un sistema basado en sitios web para consultar y ordenar imágenes satelitales, fotografías aéreas y otros productos geográficos y cartográficos a través de la USGS. El usuario puede especificar un criterio de búsqueda, seleccionar conjuntos de datos y explorar los resultados antes de descargar las imágenes y los productos necesarios para su investigación.

ORFEO significa Optical and Radar Federated Earth Observation y es comúnmente descargado como una caja de herramientas que se puede usar dentro de QGIS. Una de las ventajas es que permite análisis y clasificación de imágenes basadas en objetos.

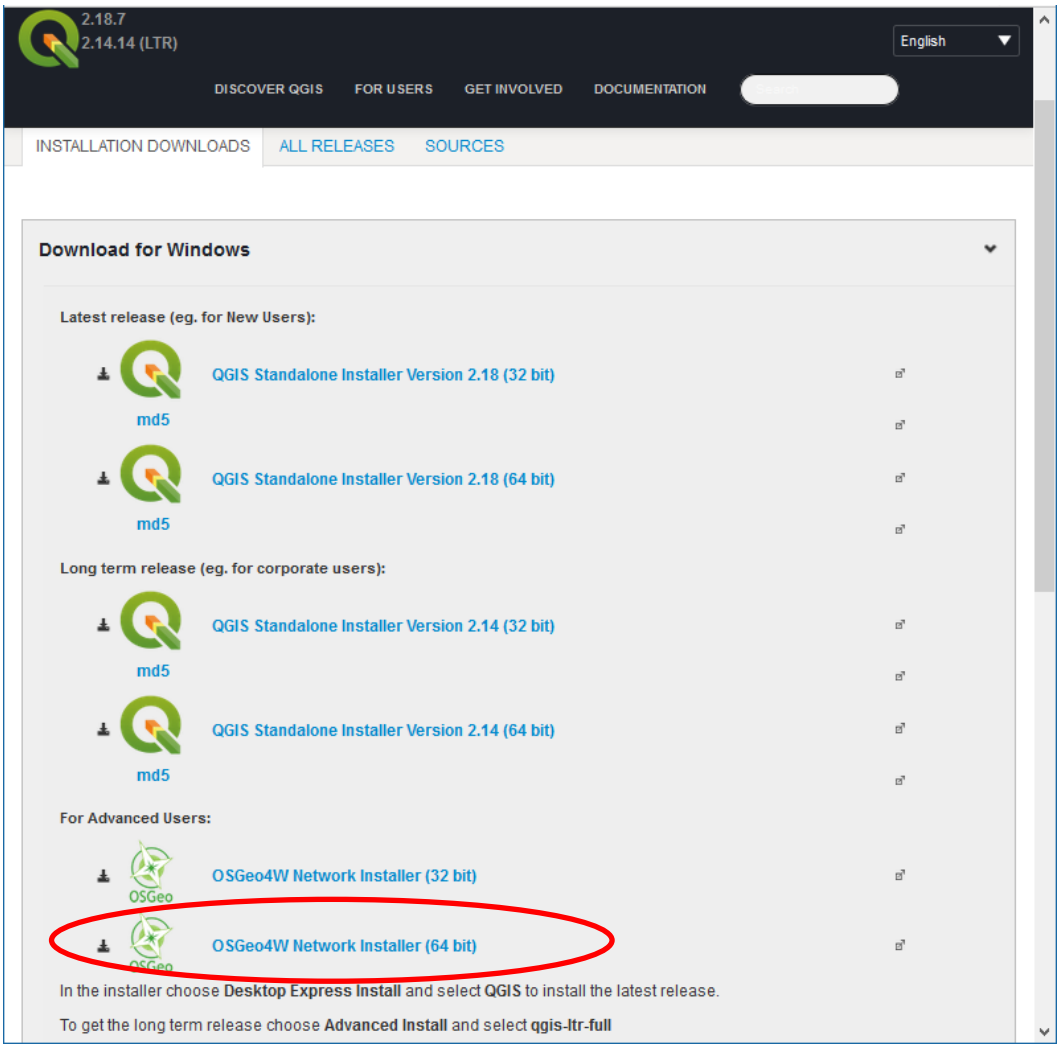
Agradecemos al DLCC por hacer accesibles los fondos para este proyecto y facilitar seminarios web, Bird Conservancy of the Rockies para la gestión del proyecto, QGIS y ORFEO equipos por su colaboración para hacer software de código abierto integrado y disponible y USGS para apoyar la recolección de datos y la difusión en el Interfaz de usuario de Earth Explorer.

1. Descargar e instalar el software de QGIS para su plataforma

1.1 Descargar el software QGIS

Abra su navegador y vaya a: https://www.qgis.org/en/site/forusers/download.html o https://www.qgis.org/es/site/forusers/download.html (en Español)

Usted notará la versión actual, fecha de lanzamiento, y que QGIS está disponible en Windows, MacOS X, Linux y Android en varios idiomas. En este ejemplo vamos a trabajar con una instalación de Windows 64 bits, pero usted puede elegir lo que es más apropiado para sus necesidades. Hay dos opciones: la última versión (herramientas más nuevas, pero también la "oportunidad" para descubrir errores) o la versión a largo plazo, que es más estable. En lugar de tomar el camino más conservador con la liberación a largo plazo de 64 bits, vamos a seleccionar a los usuarios la opción avanzada: **OSGEO4W del instalador de red (64 bits)** cerca de la parte inferior de las descargas para Windows que incluye la Caja de Herramientas Orfeo (OTB).

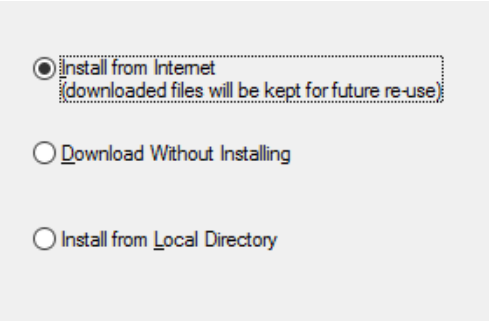


Descargue el software e inicie la instalación.

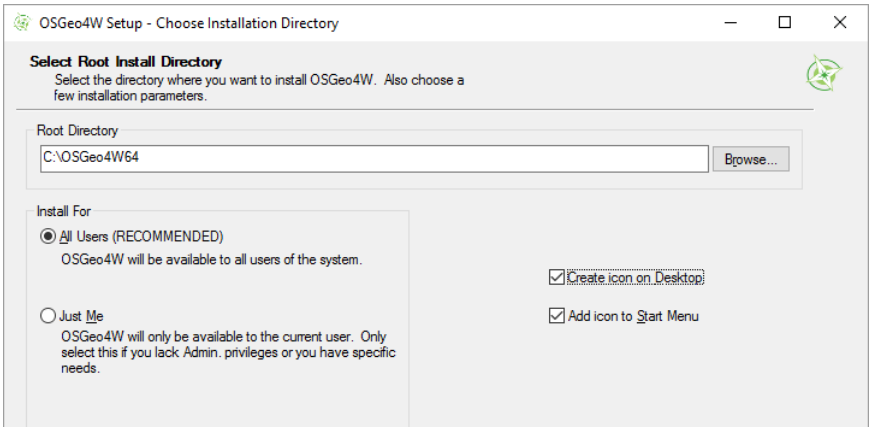
1.2 Instalar el software QGIS

Haga clic en **Instalación Avanzada** (“**Advanced Install”)** y luego **Siguiente** (**“Next”)** para continuar.

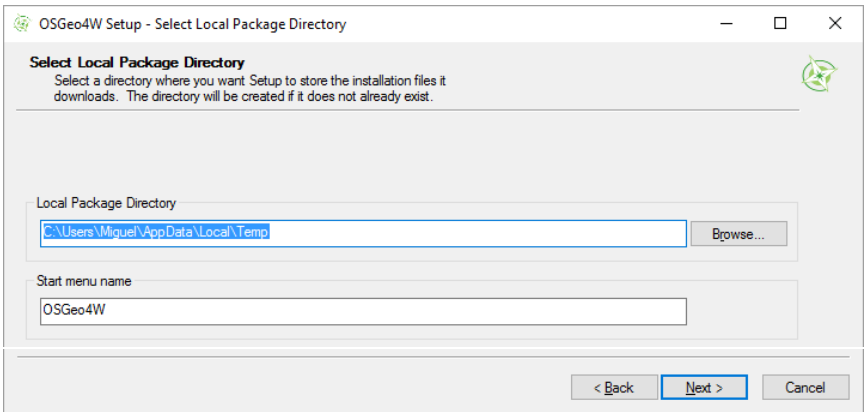
Elija instalar desde Internet (cada componente se elegirá y luego se descargará individualmente del repositorio). Haga clic en **Siguiente** (“**Next”)**.



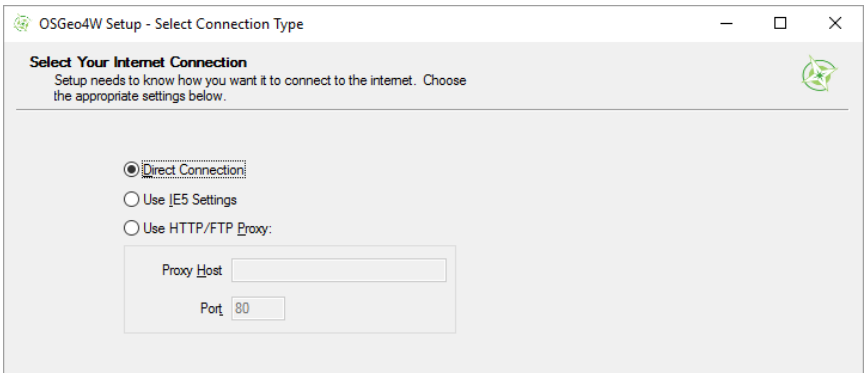
Seleccione el directorio raíz para la instalación y elija la opción para instalar el software con disponibilidad para todos los usuarios. Si desea crear un icono en el escritorio y agregarlo al menú Inicio, seleccione esta opción, así, haga clic en **Siguiente** (“**Next”)**.



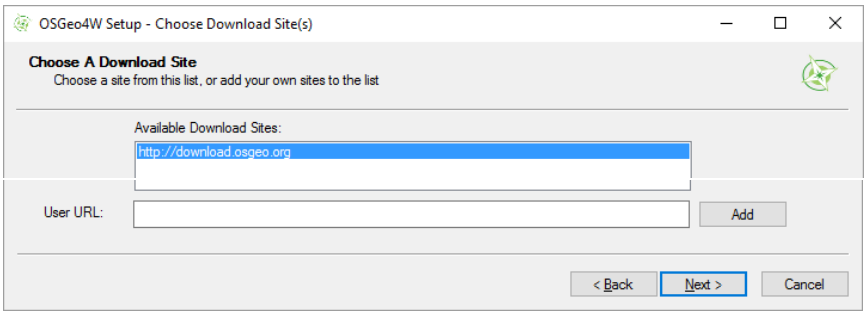
Si es necesario navegar al **directorio del paquete local (Local Package Directory)**, a continuación, haga clic en **Siguiente (Next).**



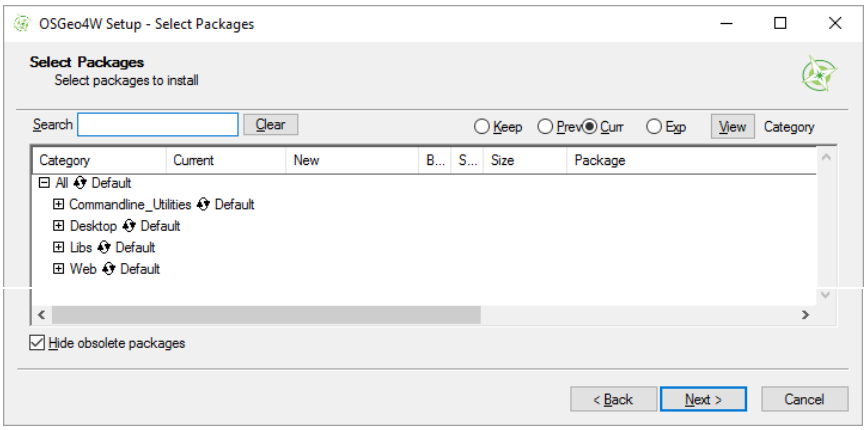
Seleccione la **Conexión directa (Direct Connection),** haga clic en **Siguiente** (**Next).**



Haga clic en el **sitio de descarga disponibles (available download site),** haga clic en **Siguiente (Next).**



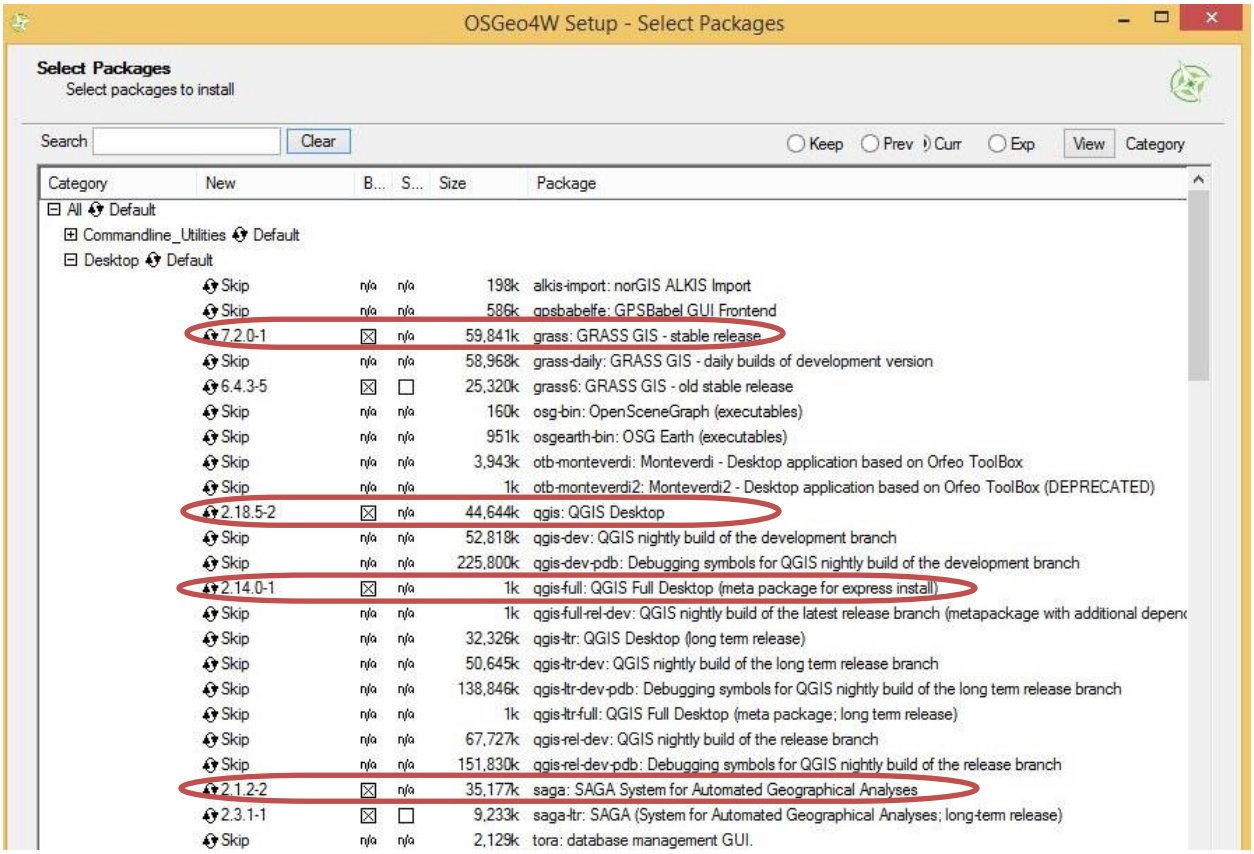
Seleccione paquetes.



Desde la página de selección de paquetes, expanda la sección de Escritorio y haga clic una vez en "Omitir" para al menos los siguientes paquetes de modo que muestre el último número de versión donde solía decir "Omitir":

- Grass: GRASS GIS - liberación estable
- Qgis: QGIS Desktop
- Qgis-full: QGIS Full Desktop (paquete de meta para la instalación rápida)
- Saga: Sistema SAGA de Análisis Geográfico Automatizado

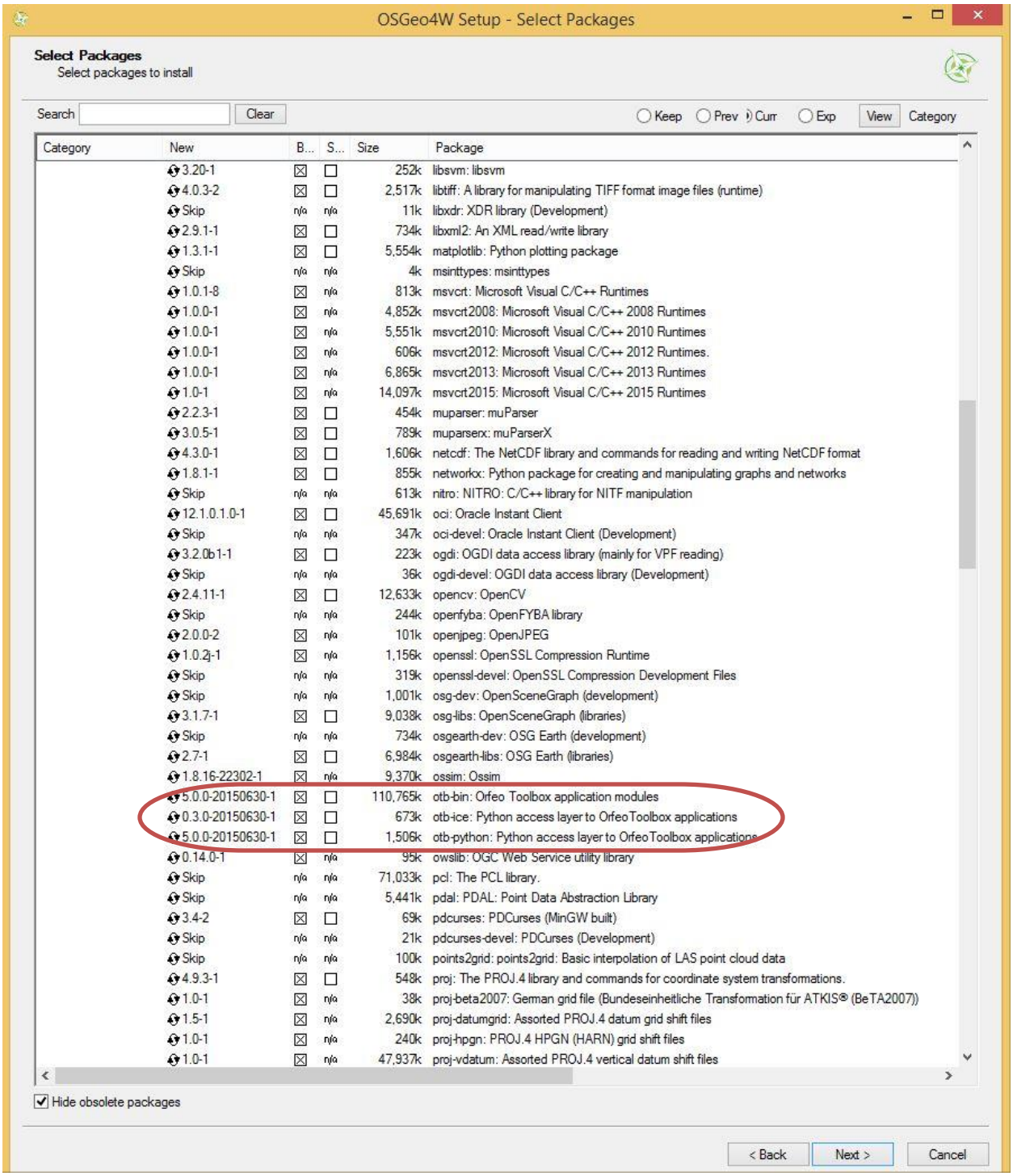
Otros paquetes con las versiones mostradas en la captura de pantalla, pero no marcadas en rojo, se agregarán automáticamente al llegar al paso posterior en el que se selecciona **"Instalar estos paquetes para satisfacer las dependencias"** (**Install these packages to meet dependencies)**.

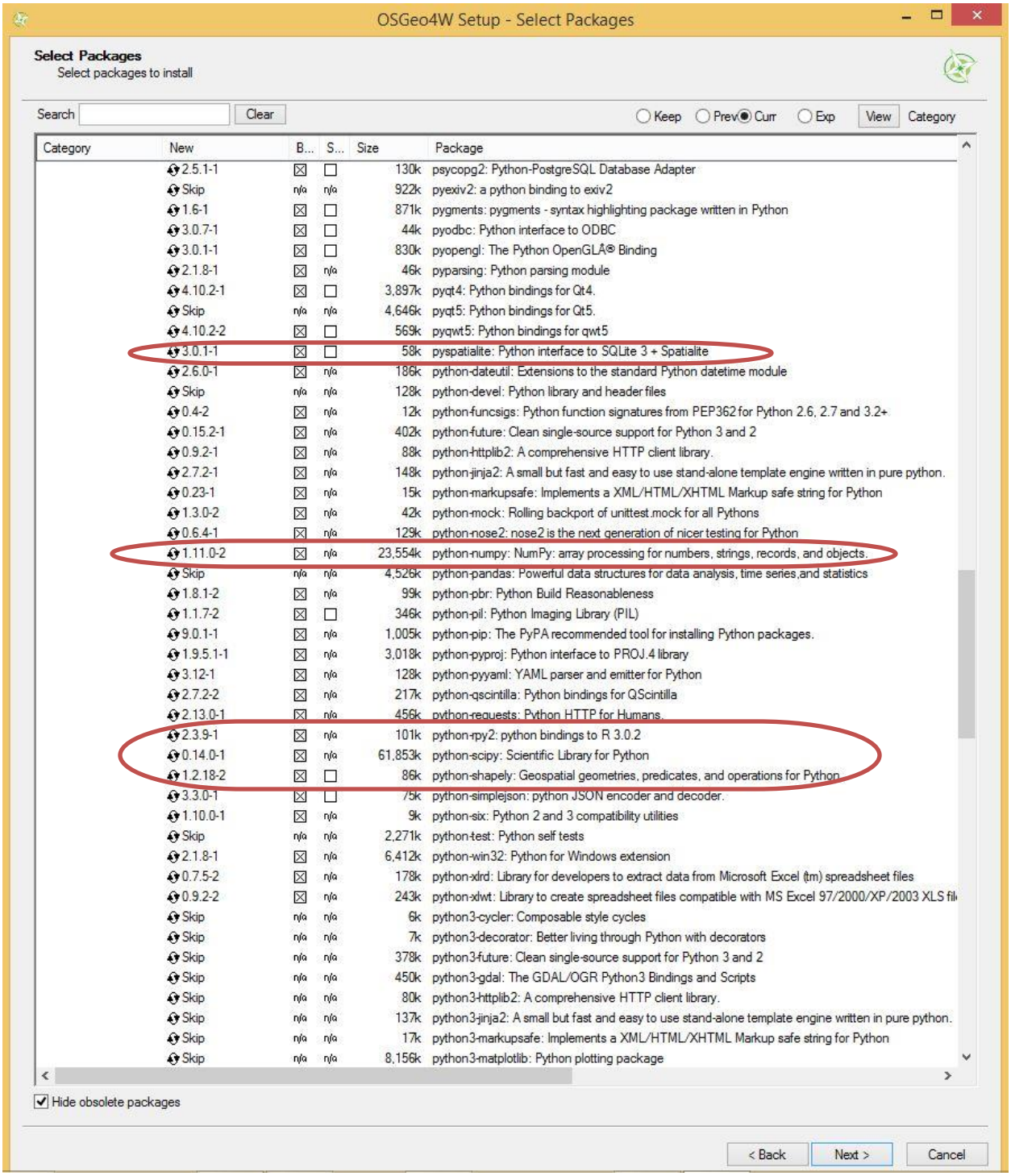


A continuación, expanda la sección Libs para verificar que hay un cheque para instalar (Keep) Orfeo Tool Box (OTB) al menos:

- Otb-bin: Módulos de aplicación Orfeo Toolbox
- Otb-ice: capa de acceso de Python a las aplicaciones de Orfeo Toolbox
- Otb-python: capa de acceso de Python a las aplicaciones de Orfeo Toolbox
- Pyspatialite: Interfaz de Python para SQLite 3 + Spatialite
- Python-numpy: Num Py: Procesamiento de matrices para números, cadenas, registros y objetos
- Python-rpy2: enlaces de python a R 3.0.2
- Python-scipi: Biblioteca científica para Python
- Python-shapely: Geoespaciales, predicados y operaciones para Python.

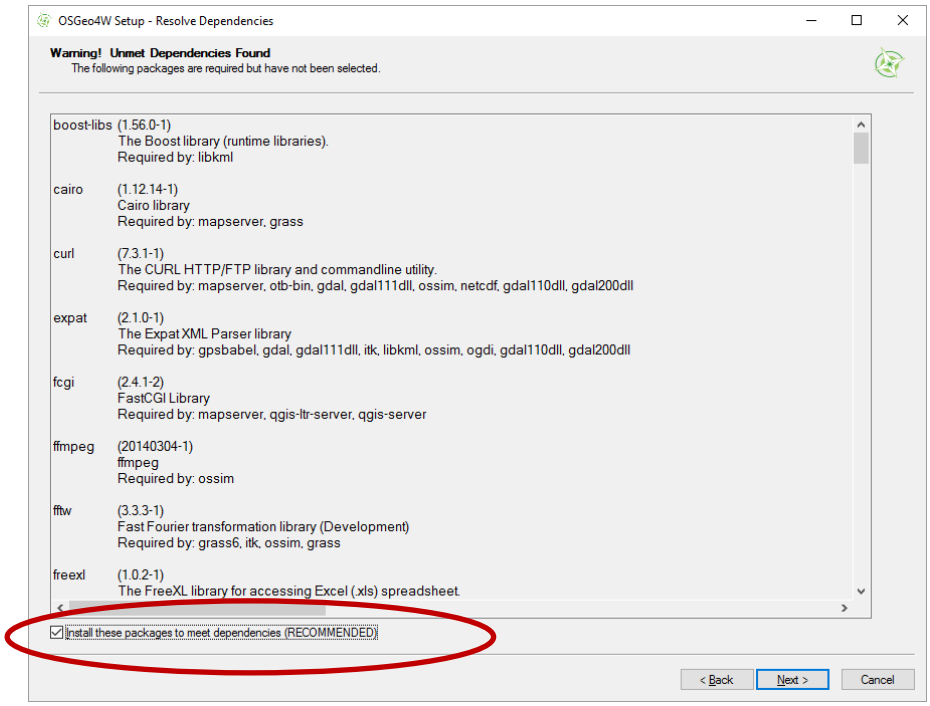
Si el espacio no es un factor limitante, puede considerar una instalación completa. La opción **Instalar todo** también tardará más tiempo en descargar e instalar. Haga clic en la pantalla de arriba en **"Categoría" (Category)** inmediatamente a la derecha de **"todo" (All)** en las flechas que giran, y cicle a través de las iteraciones hasta que la opción **"Instalar"(Install)** está al lado de **"todo" (All)**. Esto instalará todas las opciones que garantizan que tiene la Caja de Herramientas ORFEO y todas las otras herramientas disponibles en el momento del lanzamiento, puede saltar al final de este capítulo si decide ir con una instalación completa, de lo contrario, continúe con las opciones.


OSGeo4W Setup - Select Packages
Select Packages
Select packages to install
Search
Clear
Keep
Prev
Curr
Exp
View
Category
Category
New
B...
S...
Size
Package
3.20-1 252k libsvm: libsvm
4.0.3-2 2,517k libtiff: A library for manipulating TIFF format image files (runtime)
Skip 11k libxdr: XDR library (Development)
2.9.1-1 734k libxml2: An XML read/write library
1.3.1-1 5,554k matplotlib: Python plotting package
Skip 4k msinttypes: msinttypes
1.0.1-8 813k msvcrt: Microsoft Visual C/C++ Runtimes
1.0.0-1 4,852k msvcrt2008: Microsoft Visual C/C++ 2008 Runtimes
1.0.0-1 5,551k msvcrt2010: Microsoft Visual C/C++ 2010 Runtimes
1.0.0-1 606k msvcrt2012: Microsoft Visual C/C++ 2012 Runtimes.
1.0.0-1 6,865k msvcrt2013: Microsoft Visual C/C++ 2013 Runtimes
1.0-1 14,097k msvcrt2015: Microsoft Visual C/C++ 2015 Runtimes
2.2.3-1 454k muparser: muParser
3.0.5-1 789k muparserx: muParserX
4.3.0-1 1,606k netcdf: The NetCDF library and commands for reading and writing NetCDF format
1.8.1-1 855k networkx: Python package for creating and manipulating graphs and networks
Skip 613k nitro: NITRO: C/C++ library for NITF manipulation
12.1.0.1.0-1 45,691k oci: Oracle Instant Client
Skip 347k oci-devel: Oracle Instant Client (Development)
3.2.0b1-1 223k ogdi: OGDI data access library (mainly for VPF reading)
Skip 36k ogdi-devel: OGDI data access library (Development)
2.4.11-1 12,633k opencv: OpenCV
Skip 244k openfyba: OpenFYBA library
2.0.0-2 101k openjpeg: OpenJPEG
1.0.2j-1 1,156k openssl: OpenSSL Compression Runtime
Skip 319k openssl-devel: OpenSSL Compression Development Files
Skip 1,001k osg-dev: OpenSceneGraph (development)
3.1.7-1 9,038k osg-libs: OpenSceneGraph (libraries)
Skip 734k osgearth-dev: OSG Earth (development)
2.7-1 6,984k osgearth-libs: OSG Earth (libraries)
1.8.16-22302-1 9,370k ossim: Ossim
5.0.0-20150630-1 110,765k otb-bin: Orfeo Toolbox application modules
0.3.0-20150630-1 673k otb-ice: Python access layer to OrfeoToolbox applications
5.0.0-20150630-1 1,506k otb-python: Python access layer to OrfeoToolbox applications
0.14.0-1 95k owslib: OGC Web Service utility library
Skip 71,033k pcl: The PCL library.
Skip 5,441k pdal: PDAL: Point Data Abstraction Library
3.4-2 69k pdcurses: PDCurses (MinGW built)
Skip 21k pdcurses-devel: PDCurses (Development)
Skip 100k points2grid: points2grid: Basic interpolation of LAS point cloud data
4.9.3-1 548k proj: The PROJ.4 library and commands for coordinate system transformations.
1.0-1 38k proj-beta2007: German grid file (Bundeseinheitliche Transformation für ATKIS® (BeTA2007))
1.5-1 2,690k proj-datumgrid: Assorted PROJ.4 datum grid shift files
1.0-1 240k proj-hpgn: PROJ.4 HPGN (HARN) grid shift files
1.0-1 47,937k proj-vdatum: Assorted PROJ.4 vertical datum shift files
Hide obsolete packages
< Back
Next >
Cancel

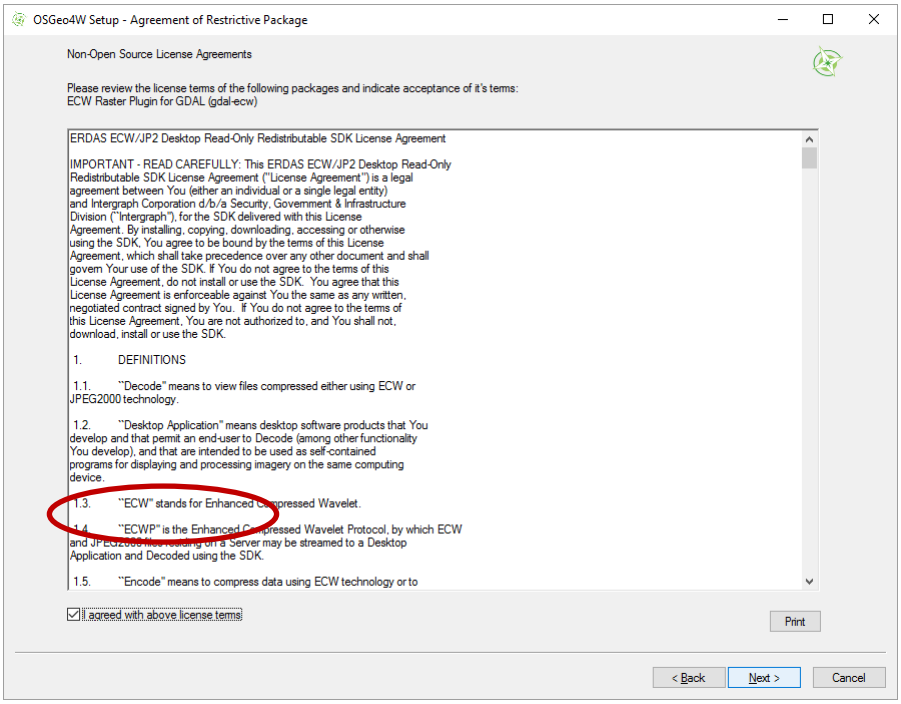


Haga clic en **Siguiente (Next)**.

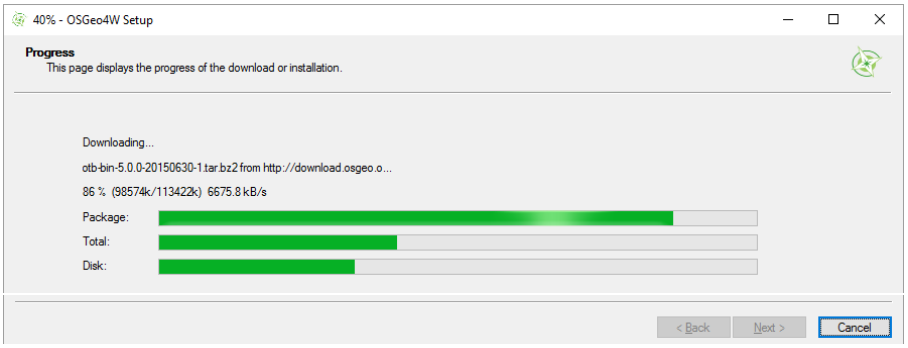
En la nueva pantalla, asegúrese de que la marca de verificación está activada para: Instalar estos paquetes para cumplir con las dependencias.



Haga clic en **Siguiente** (**Next)** y esté de acuerdo con los términos de licencia de software y las opciones del plugin seleccionado.



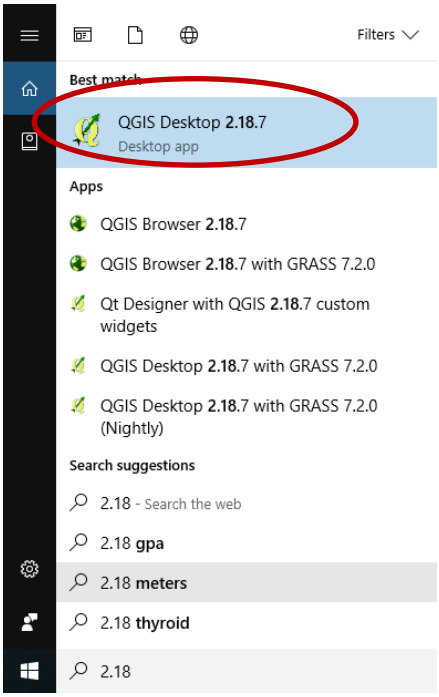
El software comenzará a descargar todas las opciones ya instalarlas.



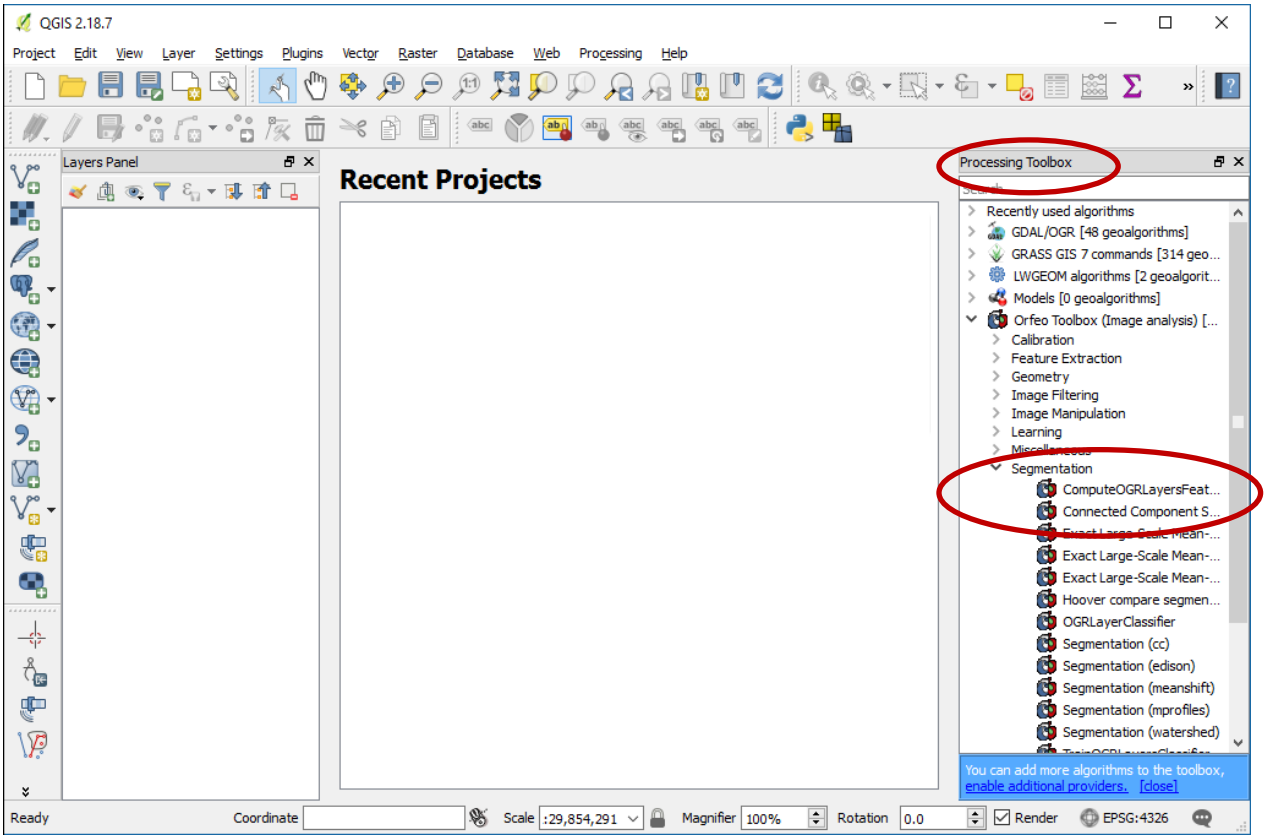
Después de la instalación con éxito, haga clic en **Finalizar (Finish).**

1.3 Verificar la instalación de su software QGIS

En aplicaciones, en su plataforma, **abra (open)** la versión recién instalada de escritorio QGIS.

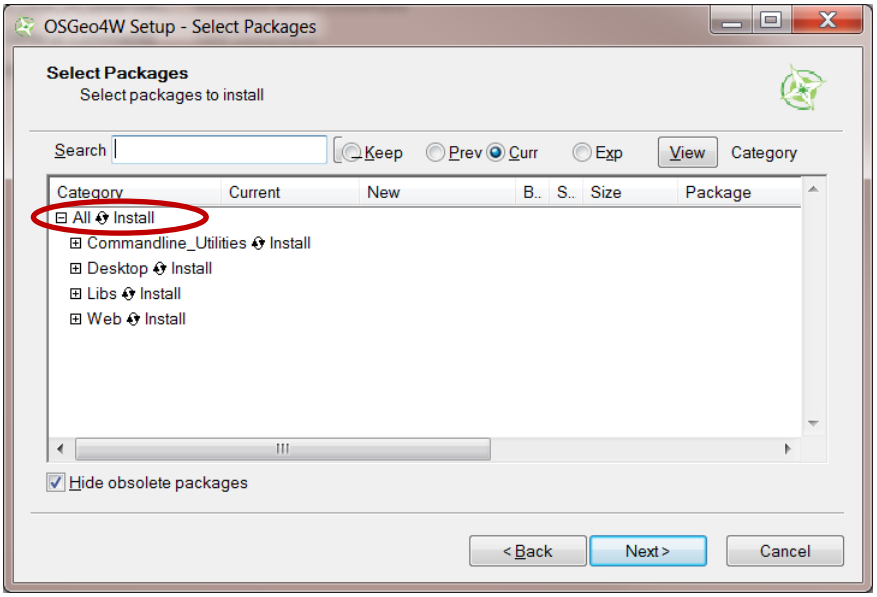


Abra la **caja de herramientas de procesamiento (Processing toolbox)**, bajo **Orfeo Caja de herramientas> Segmentación (Orfeo Toolbox** > **Segmentation)**, se encuentran varias herramientas de segmentación, que se pueden utilizar para la clasificación orientada a objetos, de imágenes de percepción remota.



Si la opcion **Orfeo Caja de herramientas> Segmentación (Orfeo Toolbox** > **Segmentation)** no está allí, o si está vacío, puede ser necesario volver a ejecutar el **instalador OSGEO4W red (64 bits) (OSGEO4W Network Installer (64 bit))** para revisar sus opciones. A medida que cambian las versiones, estas opciones también cambian; En caso de duda, una forma segura de obtener el Orfeo Toolbox es hacer una instalación completa.

Ir a las descargas y ejecutar **OSGEO4W del instalador de red (64 bit) (OSGEO4W Network Installer (64 bit))** de nuevo. Cerca del final de la instalación, en el paso donde debe seleccionar Paquetes, preste atención a las opciones. En la columna Categoría, haga clic en siguiente, justo a la derecha de "**todos" (All)** y desplazarse a través de las iteraciones hasta que la opción de **"instalar" (install)** está al lado de **todos (All),** como a continuación:



La instalación de todos los paquetes en una instalación completa tomará más espacio en su sistema, pero también asegurará que tiene todas las herramientas, incluyendo las herramientas de segmentación y clasificación necesarias para realizar una clasificación orientada a objetos con OTB en QGIS.

2 Seleccionar y descargar una imagen en Earth Explorer para usar en su clasificación orientada a objetos

Earth Explorer es una interface de usuario para consultar y ordenar imágenes de satélite y otros productos de la USGS. El usuario puede especificar un criterio de búsqueda, seleccionar conjuntos de datos y explorar los resultados antes de descargar las imágenes y los productos necesarios para su proyecto. Esta sección muestra cómo identificar y descargar imágenes de detección remota disponibles para usar en su Clasificación Orientada a Objetos.

2.1 Landsat 8.

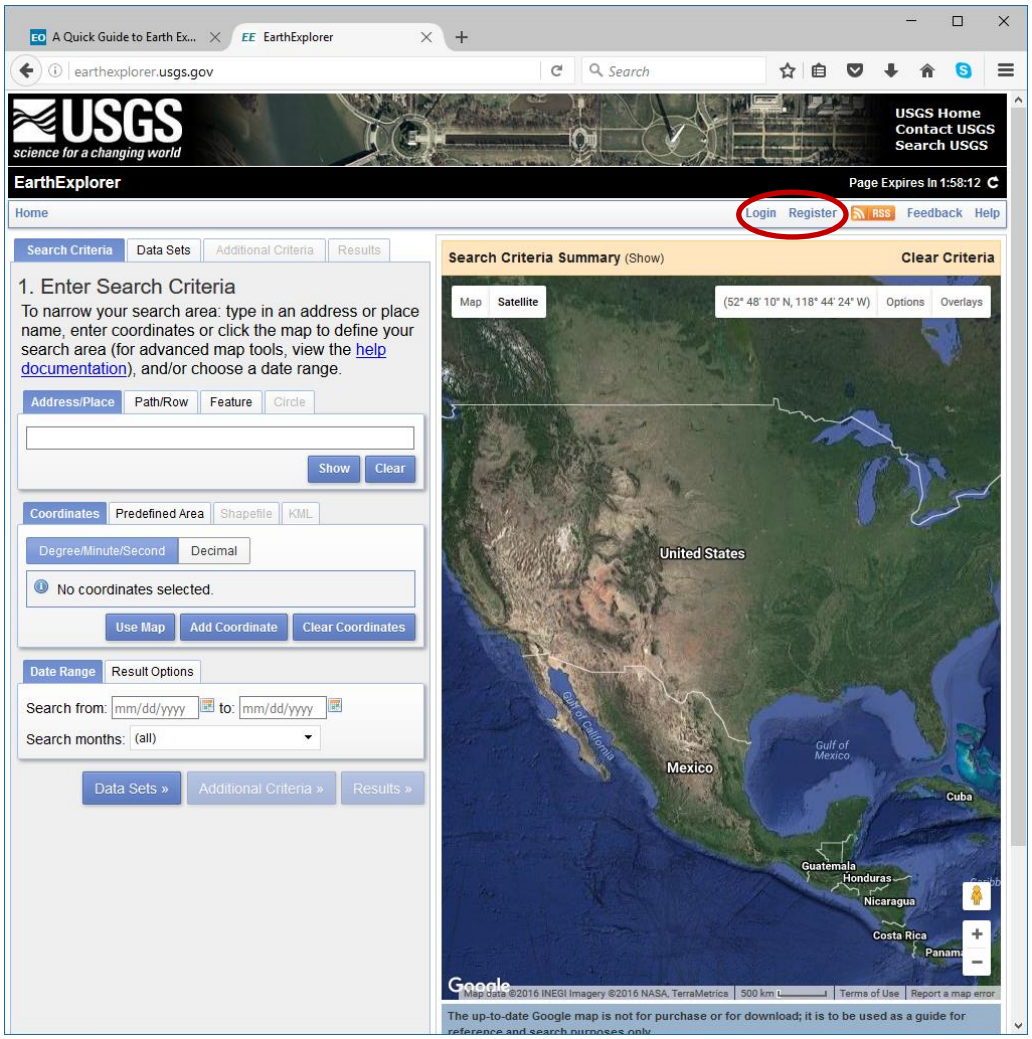
Landsat 8 es el último satélite de la NASA en una serie que ha producido un registro multiespectral ininterrumpido de la superficie terrestre de la Tierra desde principios de los años setenta. El período de retorno (re-visita) del Landsat 8 es de 16 días. Hay dos instrumentos a bordo, el Operational Land Imager (OLI, 8 bandas a 30 m de resolución espacial más una banda pancromática a 15 m) y el sensor infrarrojo térmico (TIRS, 2 bandas a 100 m, re-muestreados y co-registrados A 30 m). Los datos recogidos por estos sensores están disponibles para su descarga gratuita a través de EarthExplorer. Además, hay otras interfaces, por ejemplo, GloVis y LandsatLook Viewer. Los datos Landsat normalmente están disponibles dentro de las 24 horas de la recepción. Preguntas sobre LLandsat pueden ser contestadas en la lista de preguntas frecuentes del USGS'(https://Landsat.usgs.gov/frequently-asked), el Landsat 8 Manual (https://Landsat.usgs.gov/Landsat-8-l8-data -Los usuarios-manual), a través de correo electrónico (custserv@usgs.gov) o por teléfono (800-252-4574).

Para este ejemplo, descargaremos una imagen de Landsat 8 en la Cooperativa de Conservación del Paisaje del Desierto (DLCC) cerca de Cd. Juárez - El Paso en la frontera entre Estados Unidos y México. Las imágenes calibradas están disponibles a través de Earth Explorer.

2.2 Crea un login en EarthExplorer

El sitio web de Earth Explorer de USGS permite al usuario personalizar los parámetros de búsqueda para los datos de Landsat. Después de entrar en el sitio web, http://earthexplorer.usgs.gov (nuevos usuarios deben crear un inicio de sesión, el acceso al sitio y los datos son todavía libre), puede buscar datos mediante la selección de una región en un mapa, introduciendo Coordenadas o introduciendo un nombre de lugar. El usuario también puede seleccionar parámetros específicos tales como porcentajes de cobertura de nubes aceptables y el intervalo de fechas de adquisición deseadas.

En un navegador, vaya al sitio http://earthexplorer.usgs.gov , y entrar o registrar una nueva cuenta.



2.3 Busque y seleccione los datos que necesita

En el panel izquierdo, la primera pestaña es Criterios de búsqueda que contiene tres conjuntos de opciones de búsqueda.

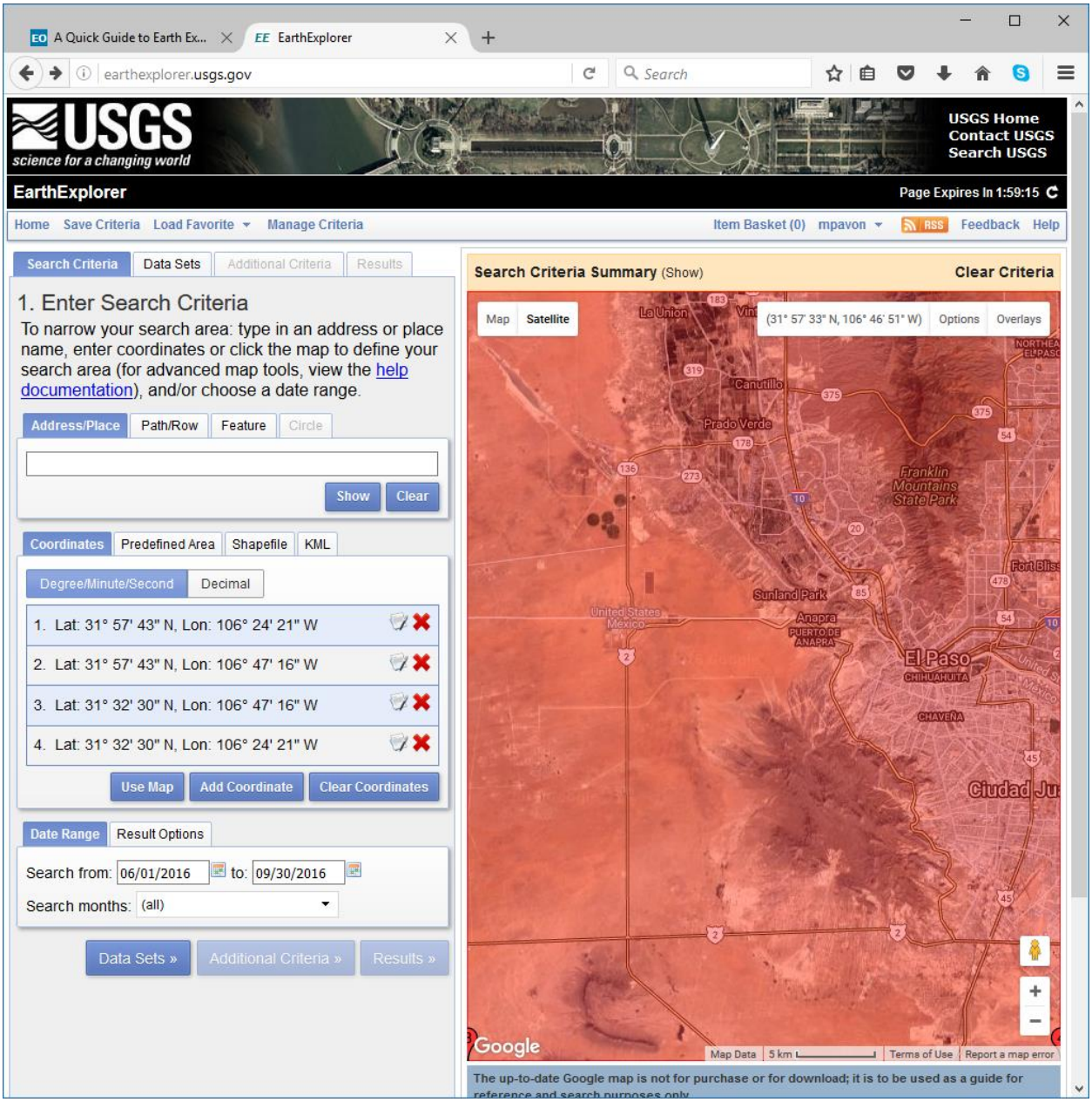
El primer conjunto de opciones permite al usuario identificar una dirección particular o un nombre de lugar, una combinación de fila de rutas o el nombre de una función en su área de interés (AOI), que puede definir aquí.

El segundo conjunto de opciones permite escribir coordenadas o utilizar un shapefile o KML con un área predefinida como entrada. Además, el usuario puede utilizar la opción "Mapa" simplemente haciendo clic en el mapa y rellenar un conjunto de coordenadas.

El tercer conjunto de opciones durante los criterios de búsqueda permite intervalos de fechas. Si el usuario tiene un año o un mes en particular en el que se realizó la recolección de datos de campo, un

año que no estaba demasiado seco o demasiado húmedo u otras fechas, en particular, se puede especificar aquí un intervalo.

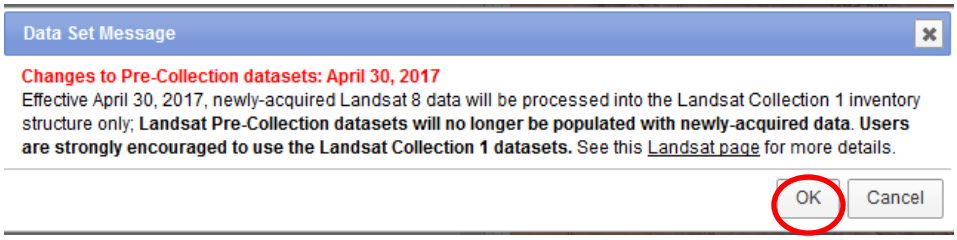
Zoom en una zona al oeste de Cd. Juarez - El Paso, en la frontera internacional y haga clic en "Usar mapa" que tomará las coordenadas de su vista como parte de los criterios de búsqueda. Para fechas, utilice el verano de 2016, del 06/01/2016 al 30/09/2016 para este tutorial. Observe que en la sección Registro de inicio de sesión ahora se define un usuario.

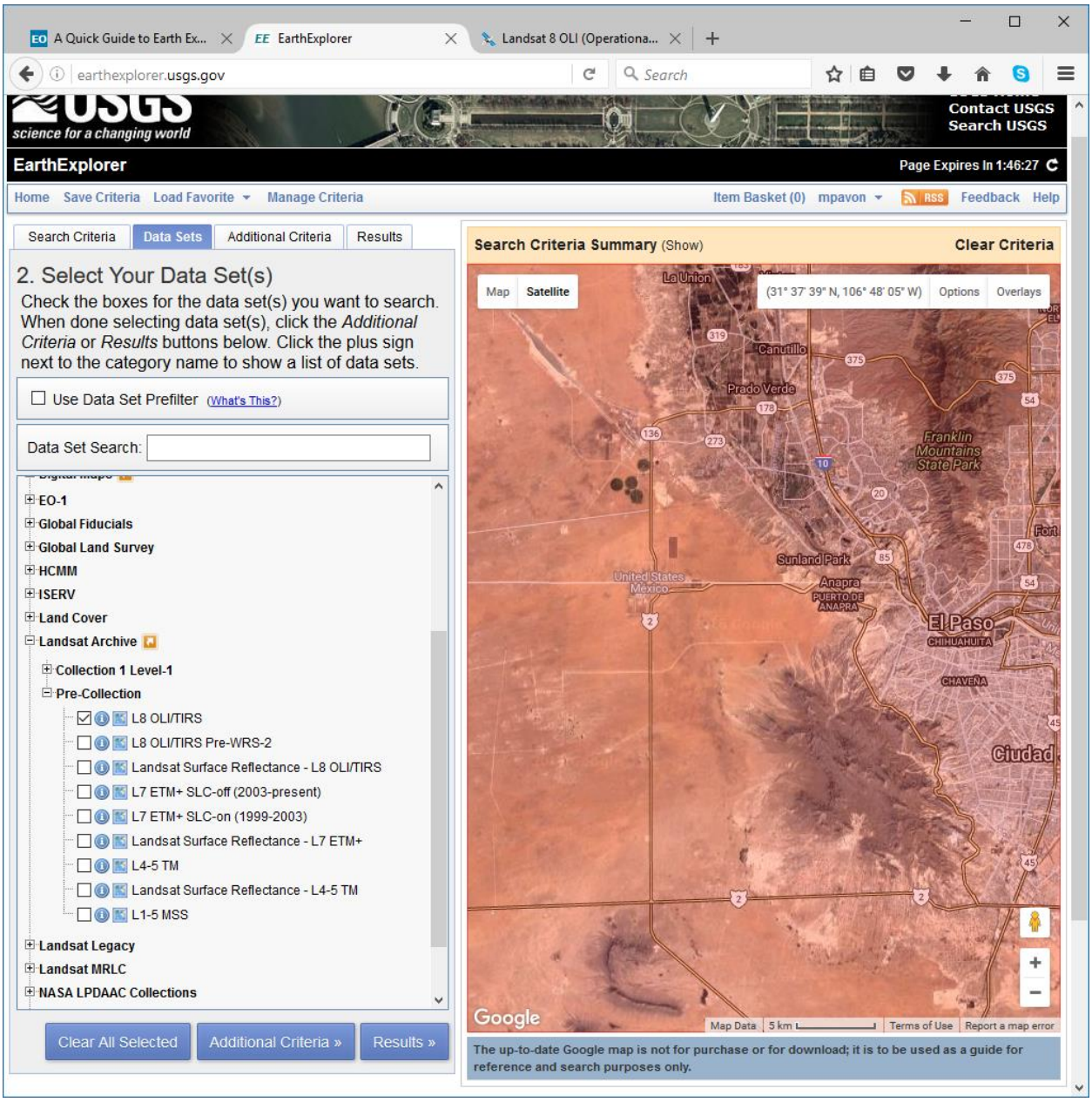


The second tab is the “Data Sets” page. Many datasets are available, from “Aerial Imagery” to “Lidar Elevation” to “AVHRR Phenology” that may be useful for DLCC. In this case, we will look for a Landsat 8 image.

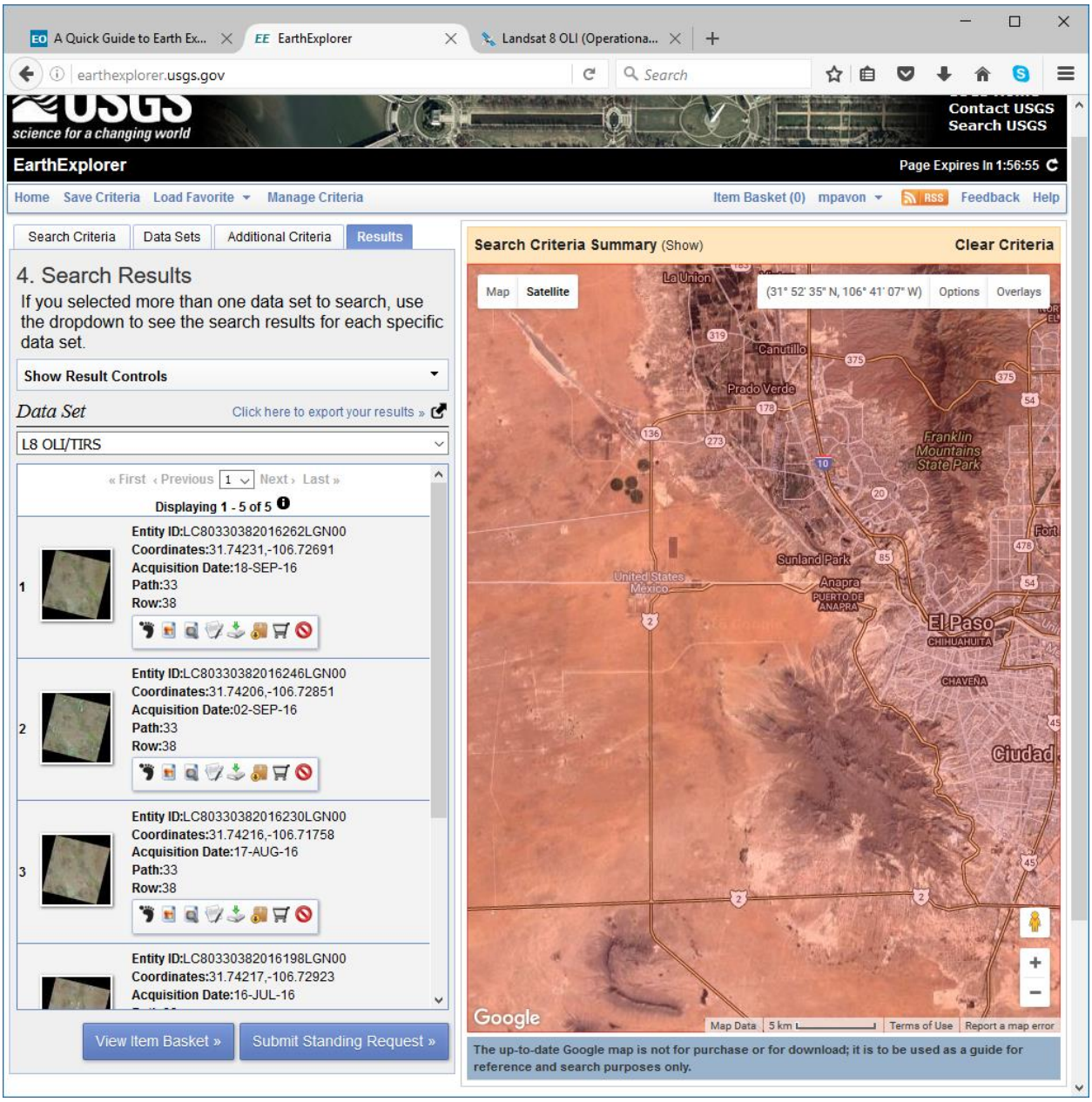
La segunda pestaña es la página "Conjuntos de datos". Muchos conjuntos de datos están disponibles, desde "Aerial Imagery" hasta "Lidar Elevation" hasta "AVHRR Phenology" que puede ser útil para DLCC. En este caso, buscaremos una imagen Landsat 8.

Vaya a la segunda pestaña, "Conjuntos de datos". Ampliar "Landsat Archive", bajo "Pre-Colección" poner una marca en "L8 OLI / TIRS". Estos son los sensores de imágenes para Landsat 8: OLI (Operational Land Imager) y TIRS (Thermal Infrared Sensor). Para este tutorial, trabajaremos solo con los datos OLI. Usted puede obtener una advertencia para los conjuntos de datos Landsat nuevos después del 30 de abril de, 2017 redirigir a que “Colección 1 conjuntos de datos”, reconocen el mensaje haciendo clic **en Aceptar (OK)** y continúe.





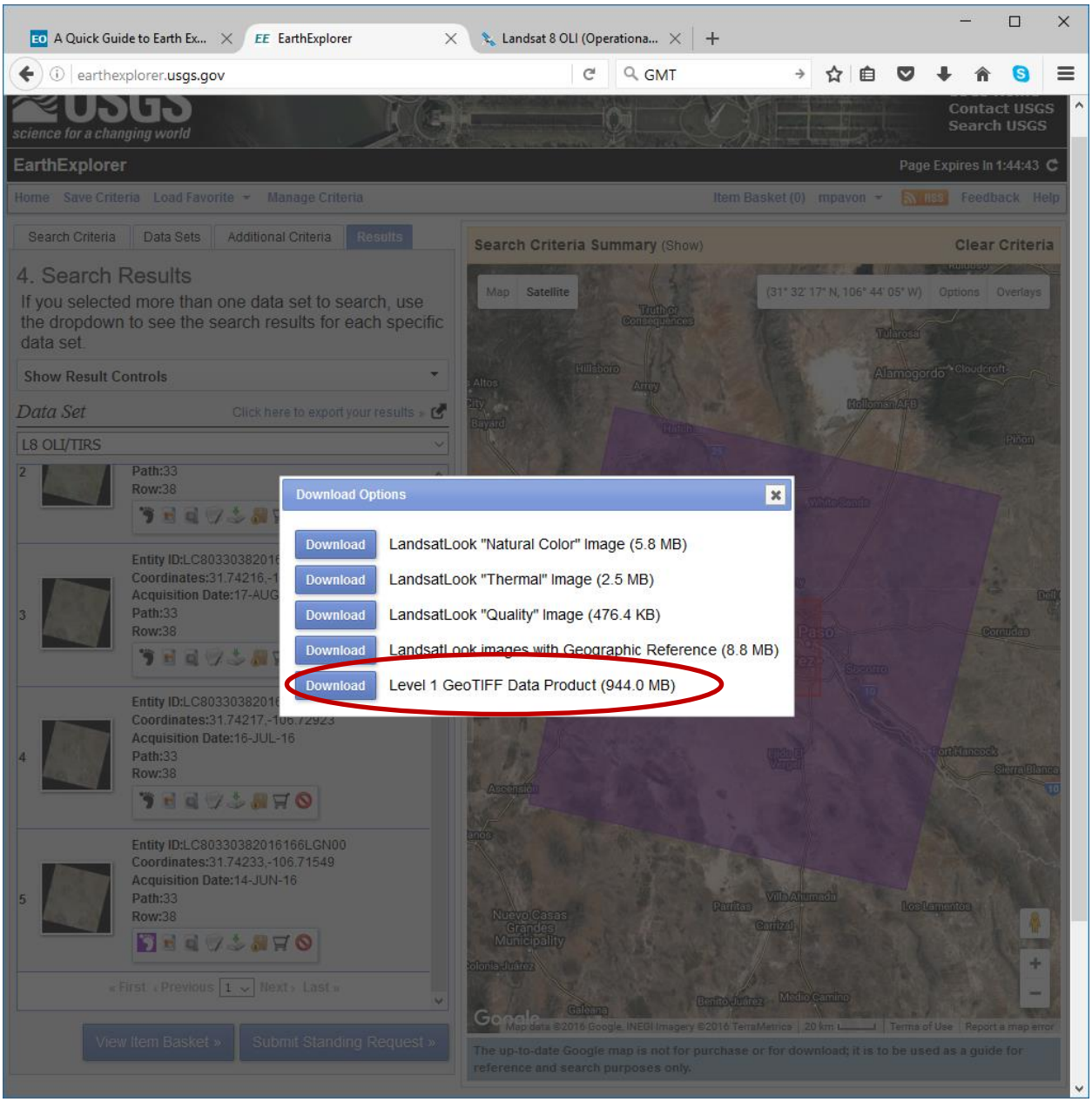
Haga clic en la ficha Criterios adicionales. Desplácese hasta "Cubierta de nubes" y seleccione "Menos del 10%". Eso debería limitar sus resultados a aquellos que tienen una mínima cobertura de nubes. Haga clic en la pestaña Resultados. El usuario puede obtener una lista de imágenes disponibles que se ajusten a los criterios con miniaturas y enlaces a más información. Las miniaturas no son lo suficientemente grandes para mostrar mucho, así que haga clic en uno para ver una imagen ligeramente más grande.



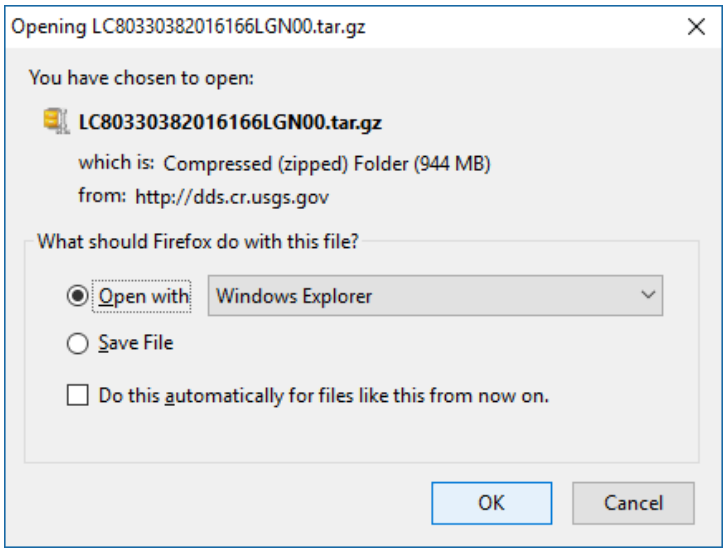
Al hacer clic en una imagen se abre una nueva ventana que permite al usuario ver mejor la imagen y algunos metadatos. Pasar por todos ellos hasta que se familiarice con los resultados. Parece que el gránulo para 14/06/2016 tiene una nubosidad mínima y se tomó alrededor de las 11 AM hora local (El Paso es GMT -6).

2.4 Descargue sus datos

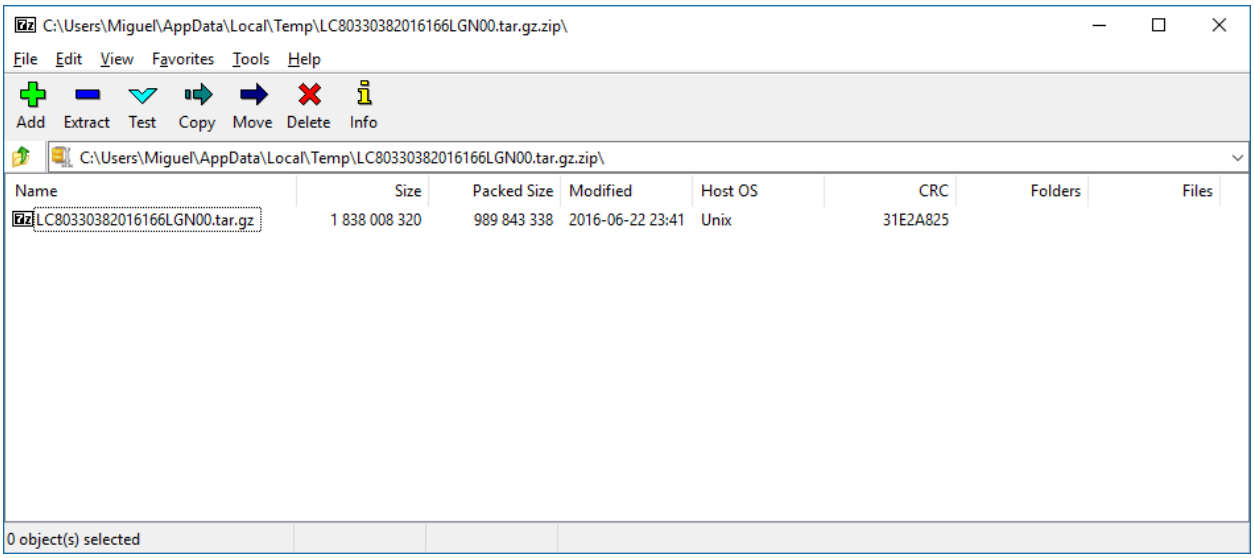
Pasar el ratón sobre los diferentes iconos de cada imagen muestra un pequeño texto de lo que hacen; En orden están: mostrar huella (extensión), mostrar extensión, comparar imagenes, metadatos de la escena, las opciones de la transferencia directa, agrega a la descarga en bloque, ordenar la escenta y excluye escena de resultados. Explora los iconos haciendo clic en ellos. Haga clic en las opciones de descarga para Entity ID: LC80330382016166LGN00.



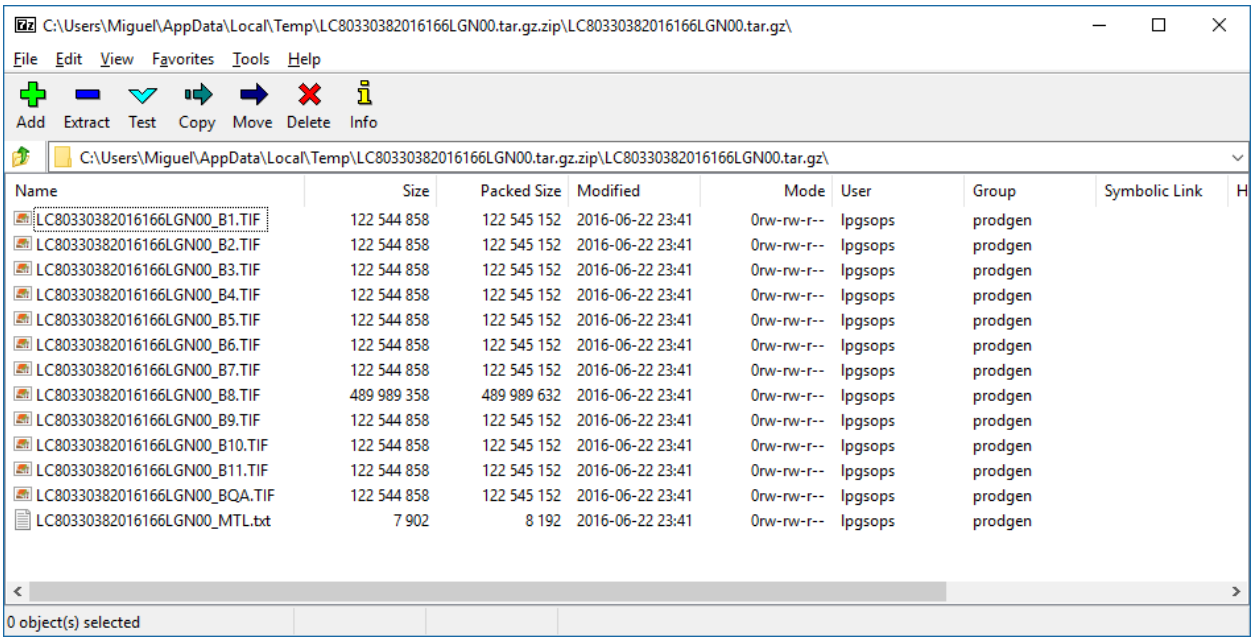
La opción más grande, "Producto de Datos GeoTIFF de Nivel 1" es la que proporciona un producto sin comprimir de resolución completa. El producto de nivel 1 es datos corregidos por terreno, georeferenciados y calibrados, un paquete de 16 bits y GeoTIFFs de un solo canal. Los otros son solo vistas previas JPEG de resolución completa para comprobar la calidad de la imagen. Haga clic en el botón "Descargar" (Download) junto a ella. Un archivo comprimido iniciará su descarga.



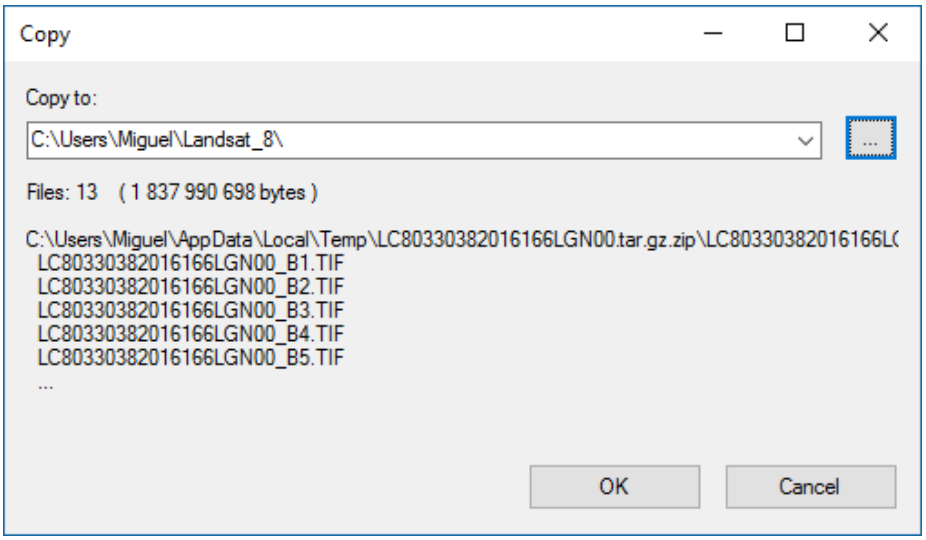
Elegir un programa con el que se puede abrir un fichero comprimido (WinZip (http://www.winzip.com), 7Zip (http://www.7-zip.org), etc.) Dependiendo de su conexión a Internet, este Puede tomar de uno a varios minutos. Se utilizó 7-Zip como una manera de abrirla (7-Zip es software de código abierto disponible en http://www.7-zip.org/ , y hay aplicaciones y utilidades también está disponible para los usuarios de Mac: línea de comandos, RAR extractor (http://www.unrarx.com), Unarchiver (http://unarchiver.c3.cx/unarchiver), etc.). Tar.gz es un archivo comprimido dentro de un archivo comprimido.



Haga doble clic en el archivo gz para descomprimir el archivo gz en sus componentes.



Haga clic en el botón de extracción y descargue las bandas individuales 1-11, la imagen de evaluación de calidad (BQA) y los metadatos (txt) en una carpeta local.



Ahora los datos de Landsat 8 están en su disco duro local. En el próximo tutorial vamos a trabajar en convertirlo en una pila de imágenes y utilizar muestras de campo para realizar la clasificación orientada a objetos.

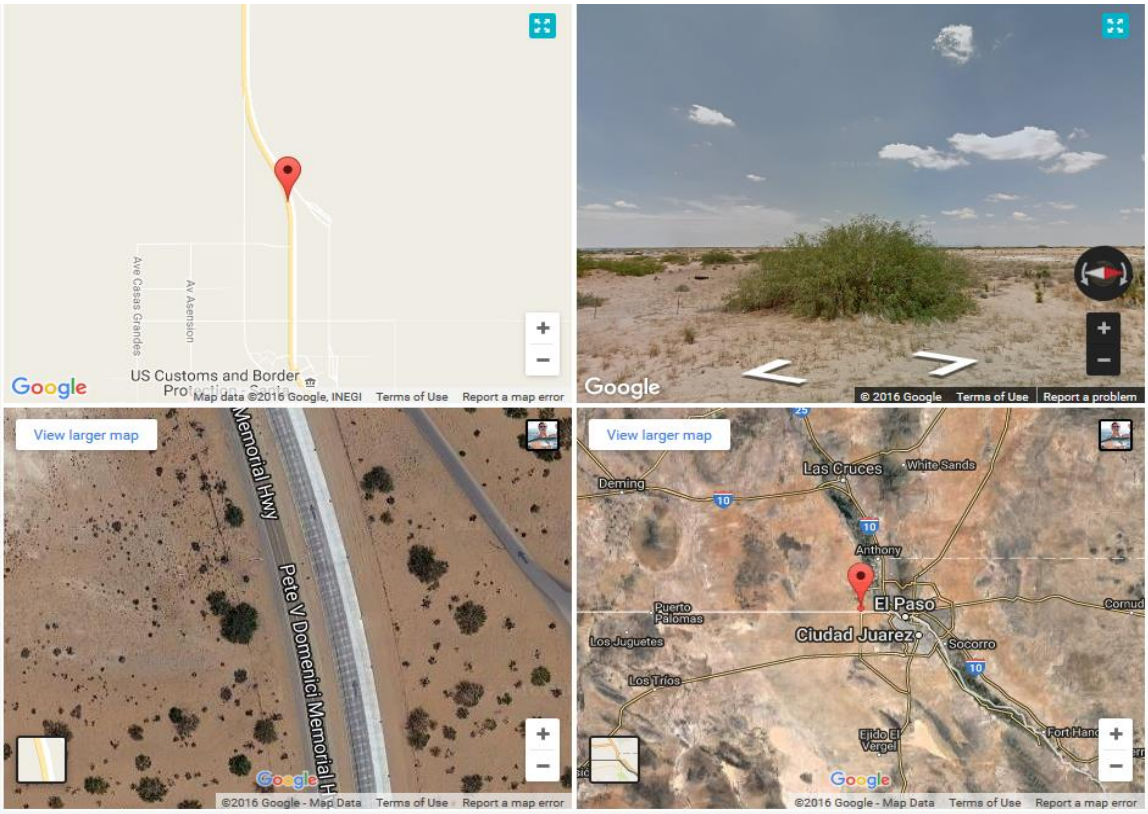
3 Clasificación orientada a objetos con QGIS y ORFEO Toolbox.

La Herramienta Orfeo, Orfeo Toolbox (u OTB) es una biblioteca para el procesamiento de imágenes de sensores remotos, la Clasificación Orientada a Objetos (OOC) está incluida en ella. El OTB se puede describir como un complemento para QGIS para realizar análisis de imágenes.

3.1 Polígonos de referencia y comprobación de campo

Para OOC necesitamos tener polígonos de formación que definan las clases de cobertura de tierra que esperamos crear, que se llaman las clases objetivas. Usted crea cada polígono de entrenamiento alrededor de una clase de cubierta de tierra conocida para "enseñar" al algoritmo de clasificación cómo se ve esa cubierta terrestre. Un polígono o punto de entrenamiento es una referencia a un lugar en el suelo donde usted es positivo de la cubierta de tierra en el momento de la recolección, registrando los atributos de cada punto de verificación debe ser la cubierta de tierra, hora / fecha de recolección y coordenadas X, Y (En un sistema de coordenadas definido). Además, necesitamos crear la pila de imágenes de todas las bandas Landsat 8. Una pila de imágenes es la combinación de 11 rangos diferentes de frecuencia de luz disponibles para una imagen, por ejemplo, azul, verde, rojo, infrarrojo cercano, etc., superpuestos en el mapa para crear una imagen multicapa) y el suelo Verificando puntos de verificación para evaluar la exactitud de la clasificación.

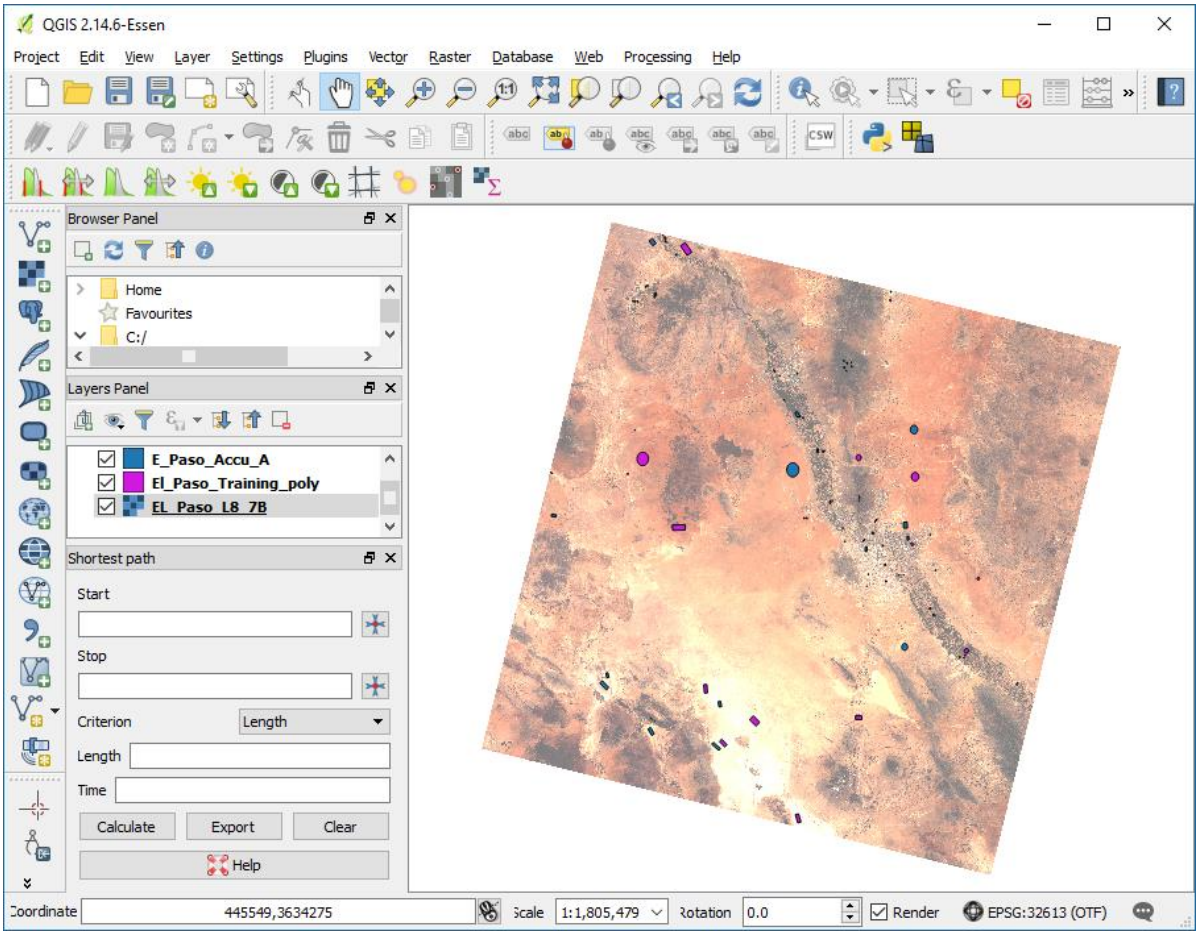
Idealmente tanto los polígonos de entrenamiento como los puntos de verificación son verificados en el campo y documentados con fotografías. La recopilación en campo de puntos de comprobación de terreno (o polígonos) también debería estar temporalmente cerca de la fecha de recopilación de las imágenes de detección remota (usaremos Landsat 8 en este ejercicio). Esta información nos da la tierra de verdad para lo que realmente estaba en el suelo en el momento de la colección de imágenes. Los polígonos de formación y los puntos de control se distribuirán en toda la extensión de la pila de imágenes y deberán incluir un mínimo recomendado de 10 polígonos y puntos por clase de cobertura terrestre, para proporcionar suficiente diversidad estadística al clasificador. En proyectos reales, el trabajo de campo es necesario para capturar la información de comprobación del terreno, es decir, determinar cuál es la clase correcta de uso del suelo basada en visitar la ubicación del polígono. En nuestro caso, como las fotografías de campo no fueron recolectadas o fácilmente disponibles, utilizamos como una herramienta de instrucción imágenes de google maps vista de la calle, imágenes aéreas de alta resolución y fotografías de usuario. Un ejemplo de tales fotografías y localizaciones cerca de la frontera de Tejas / de México es visto abajo:



En la vista de la calle en Google Maps, se cita la fecha de la imagen - es mejor si la fecha de la fotografía es cercana a la de la pila de imágenes de detección remota. Esta no es una condición que podemos influenciar mucho, una razón más para incluir en su plan de OOC para llevar a cabo la comprobación del trabajo de campo durante el mismo período de tiempo que la (s) fecha (s) de la pila de imágenes LANDSAT que planea descargar.

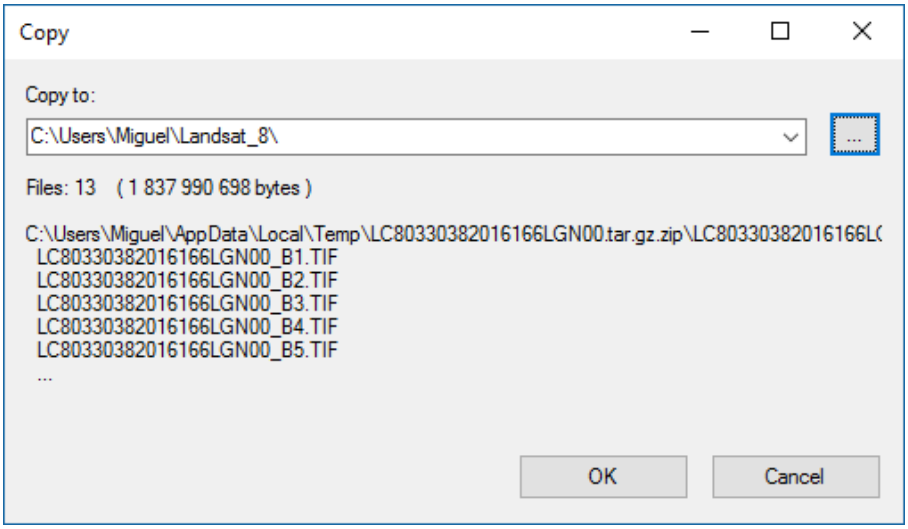
Para la creación de polígonos de entrenamiento, una vez que tengamos nuestro conjunto de imágenes: El Paso Landsat 8, 7 bandas, en QGIS, necesitamos agregar nuestra comprobación de terreno, recogida con GPS, como puntos. Creamos una nueva capa de polígonos y digitalizamos polígonos para cada punto basados en las clases, ubicación y extensión en el campo, documentados en nuestras notas y fotografías. Para este ejercicio, utilizamos la mitad de nuestros puntos para crear polígonos de entrenamiento y la otra mitad para crear polígonos que más tarde serán utilizados para la evaluación de exactitud de clasificación como polígonos de validación.

Para este ejercicio, se crearon cerca de 90 polígonos para seis clases: agua, desarrollado (urbano), hierba y agricultura, matorrales, tierras áridas y bosques. Los polígonos de entrenamiento se utilizarán para la formación de muestras para su uso en la clasificación, mientras que los polígonos de validación se utilizarán para la evaluación de la exactitud de la clasificación resultante. Los polígonos para las muestras de entrenamiento (en un shapefile llamado "El_Paso_Training_poly.shp") se pueden ver en magenta encima de la imagen Landsat 8 en color natural (Combinación de colores naturales, (4,3,2) se refiere a las bandas Landsat mostradas en rojo, verde, azul (RGB) respectivamente). Los polígonos reservados para la evaluación de la exactitud se ven encima de la imagen en azul (reservados en un shapefile llamado: "El_Paso_accu_A.shp").

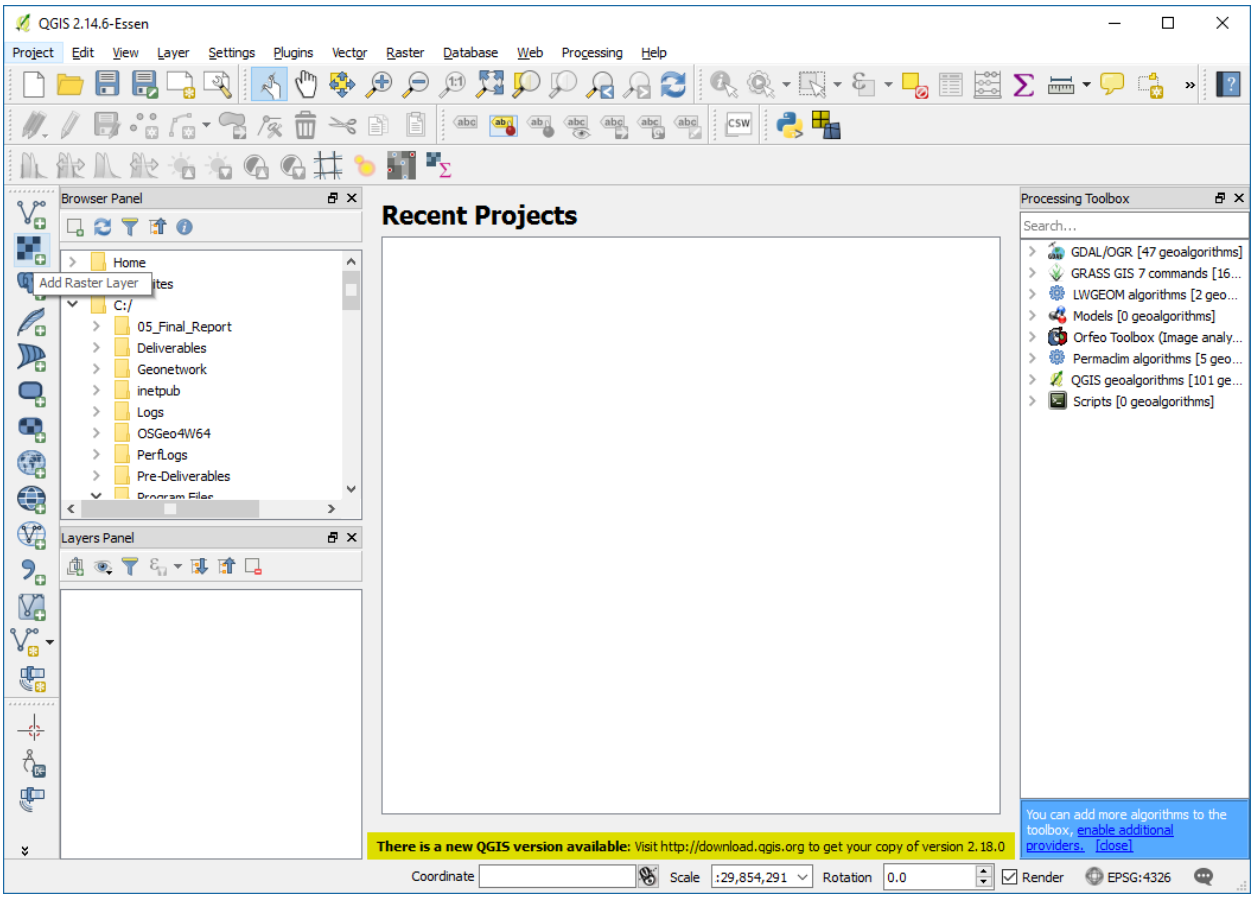


3.2 Crear la pila de imágenes

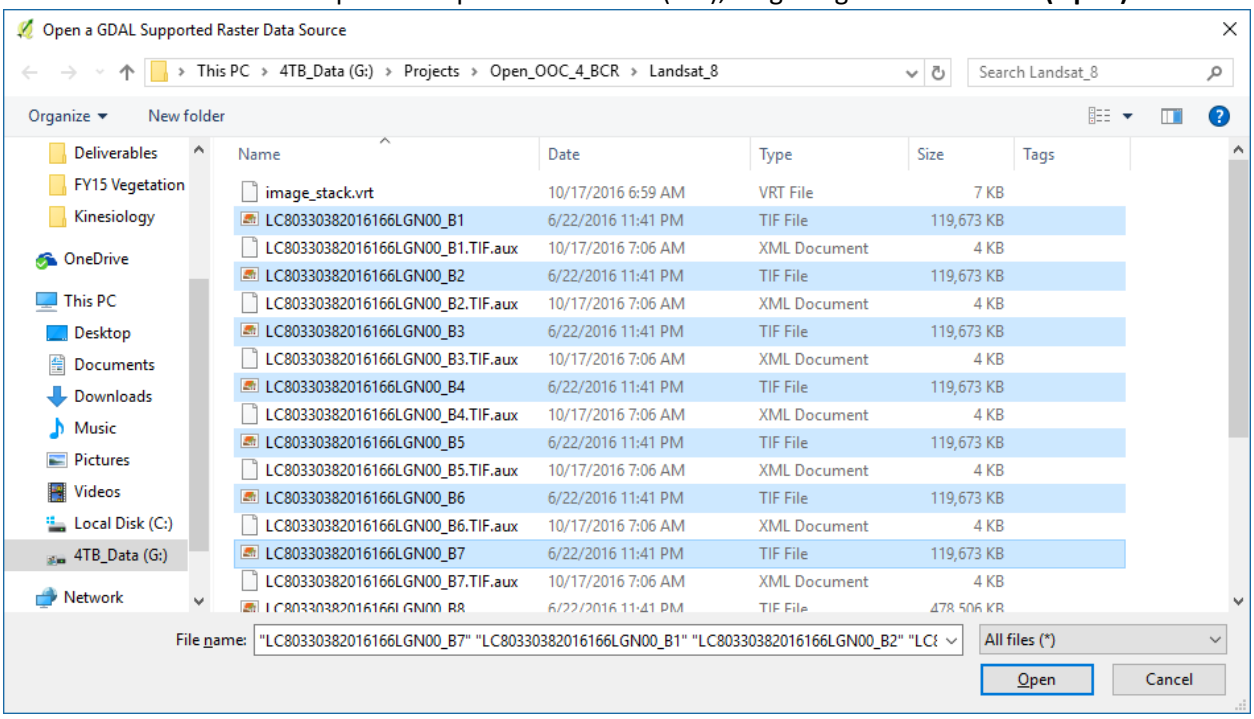
Vaya al directorio donde guardó las bandas individuales de imágenes de Landsat 8 descargadas de Earth Explorer.



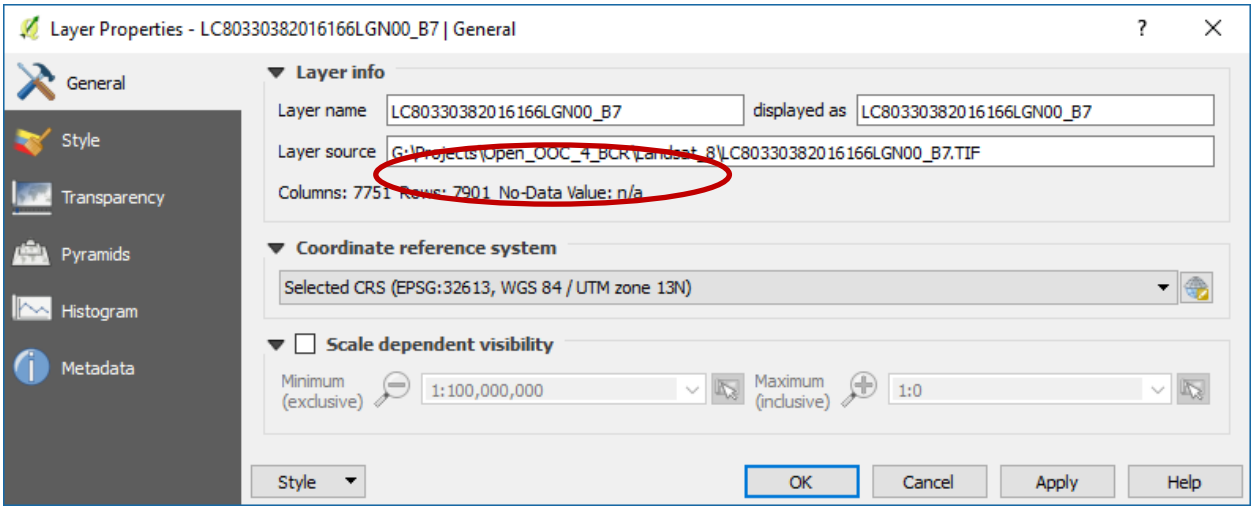
Abra QGIS y añada las capas con el botón **"Añadir capa de Raster**" (**Add Raster Layer**). Abra las primeras 7 bandas de la imagen de El Paso Landsat 8. Estas corresponden a las 7 primeras bandas del sensor OLI sin utilizar la pancromática.



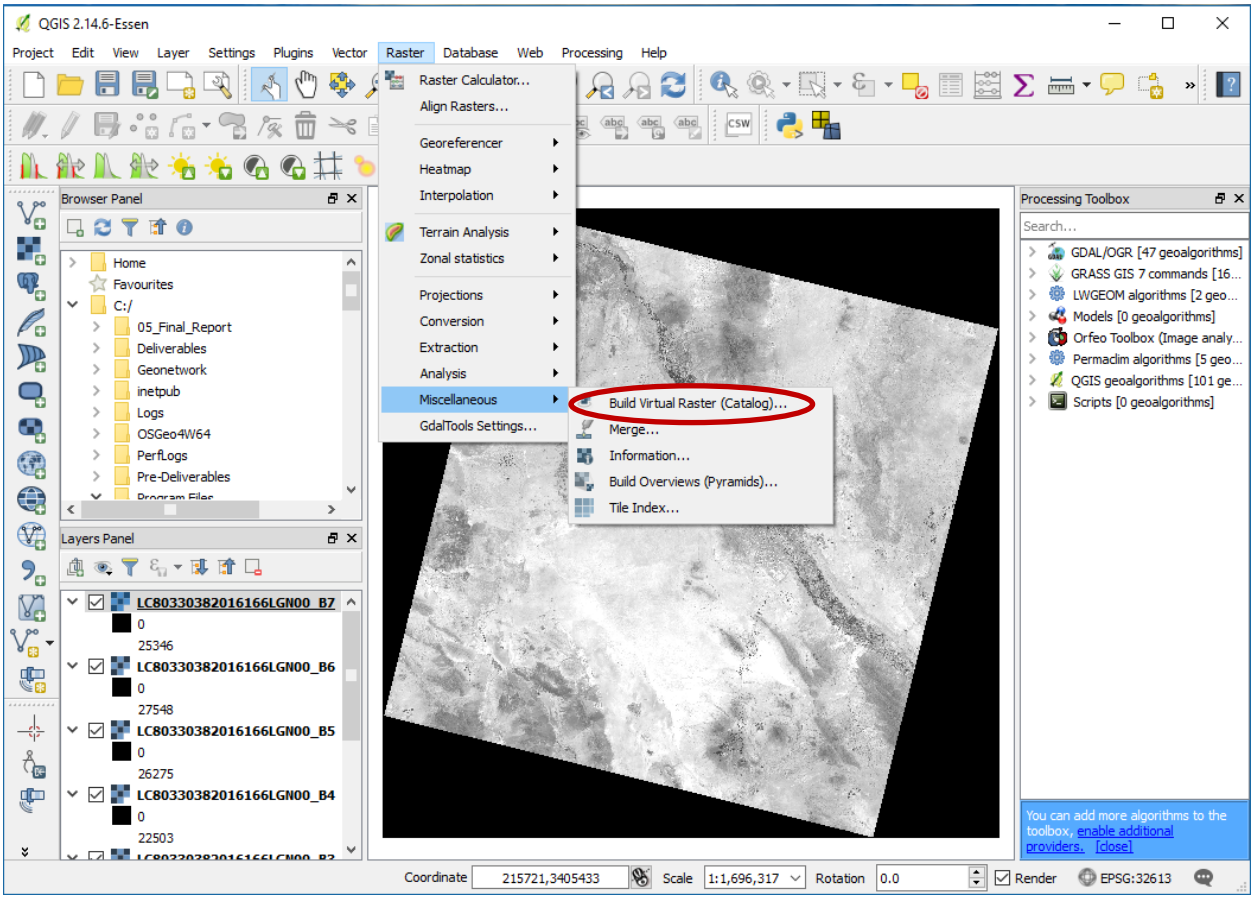
Seleccione los 7 archivos tif para las 7 primeras bandas (1-7), luego haga clic en **"Abrir" (Open)**.



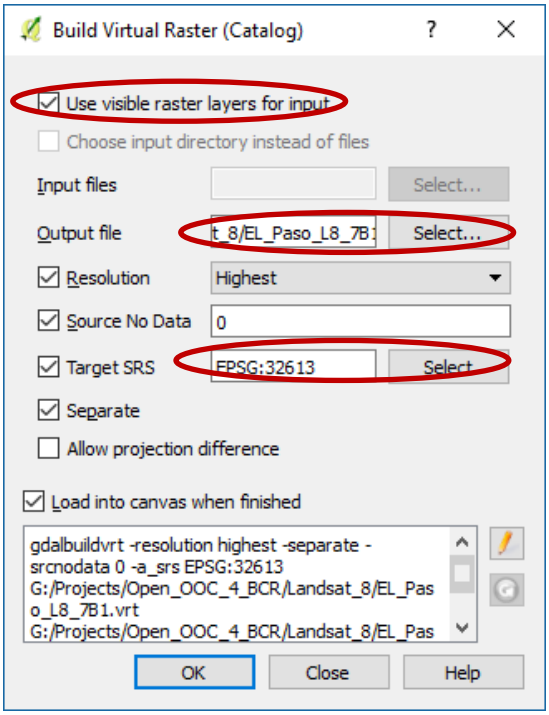
Una vez que se hayan añadido las capas, haga doble clic en cualquiera de ellas para acceder a la información general de cada imagen. Observe la información de referencia espacial: EPSG: 32613, WGS 84 / UTM zona 13N).

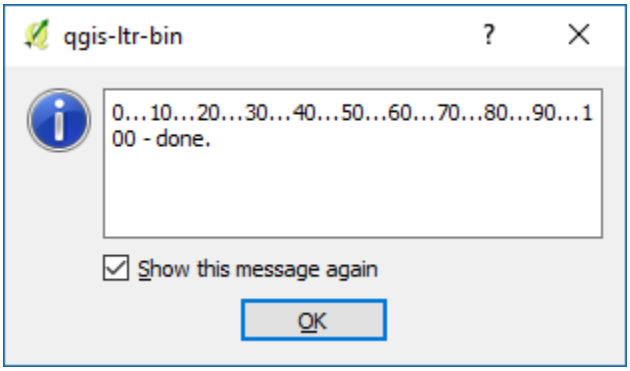


En el menú superior de QGIS, seleccione **Raster> Otros> Construir trama virtual (Catálogo) (Raster > Miscellaneous > Build Virtual Raster (Catalog))**.



Bajo las opciones de **generación virtual de mapa de bits (Catálogo),** utilizar los raster visibles para la entrada, escriba un nombre para el archivo de salida (El_Paso_L8_7B), compruebe resolución a la más alta, la fuente de ningún dato como cero y se debe cargar en el lienzo cuando haya terminado por Haciendo clic en las marcas de verificación. Para seleccionar el Sistema de Referencia Espacial (SRS) de destino, haga clic en el botón "Seleccionar" y, bajo filtro, escriba el número EPSG (32613), luego seleccione "EPSG: 32613, WGS 84 / UTM zona 13N" como su SRS. Haga clic **en OK** para la selección SRS y se separan también, haga clic en **OK** para construir su catálogo.





Haga doble clic en la pila de capas (El_Paso_L8_7B, nueva imagen con las 7 bandas). Usted debe ser capaz de cambiar la reproducción de la banda con un clic derecho e ir a las propiedades (para ver las tres bandas que se pueden ver en un momento dado). El "tipo de representación" debe ser de color Multibanda. Para una aproximación de color natural, elija rojo: banda 4, verde: banda 3, azul: banda 2. Bajo Realce de contraste, puede seleccionar en el menú desplegable "Estirar a MinMax". Haga clic en Aceptar. Una descripción de las 7 bandas de su pila está abajo, todas las cuales se adquieren a una resolución de tierra de 30 m.

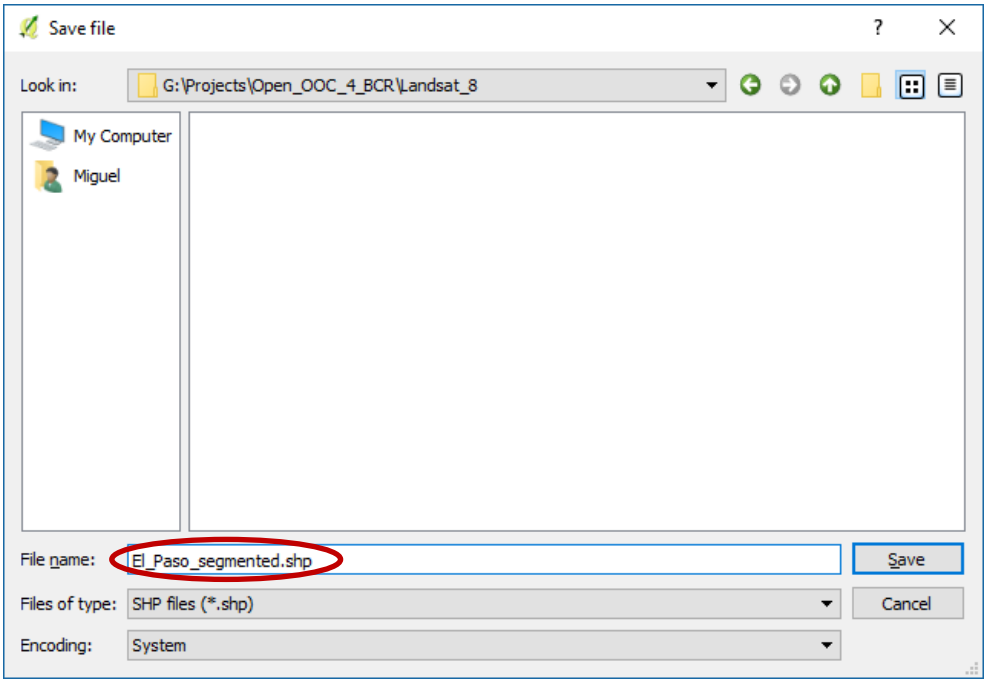
Landsat 8 Nombre de la Banda (Band Name)	**Ancho de Banda (µm) (Bandwith)**
Band 1 Costa	0.43 – 0.45
Band 2 Azul	0.45 – 0.51
Band 3 Verde	0.53 – 0.59
Band 4 Rojo	0.64 – 0.67
Band 5 Infrarojo Cercano	0.85 – 0.88
Band 6 Infrarojo de Onda Corta (SWIR) 1	1.57 – 1.65
Band 7 Infrarojo de Onda Corta (SWIR) 2	2.11 – 2.29

3.3 Segmentación vectorial y exploración de datos

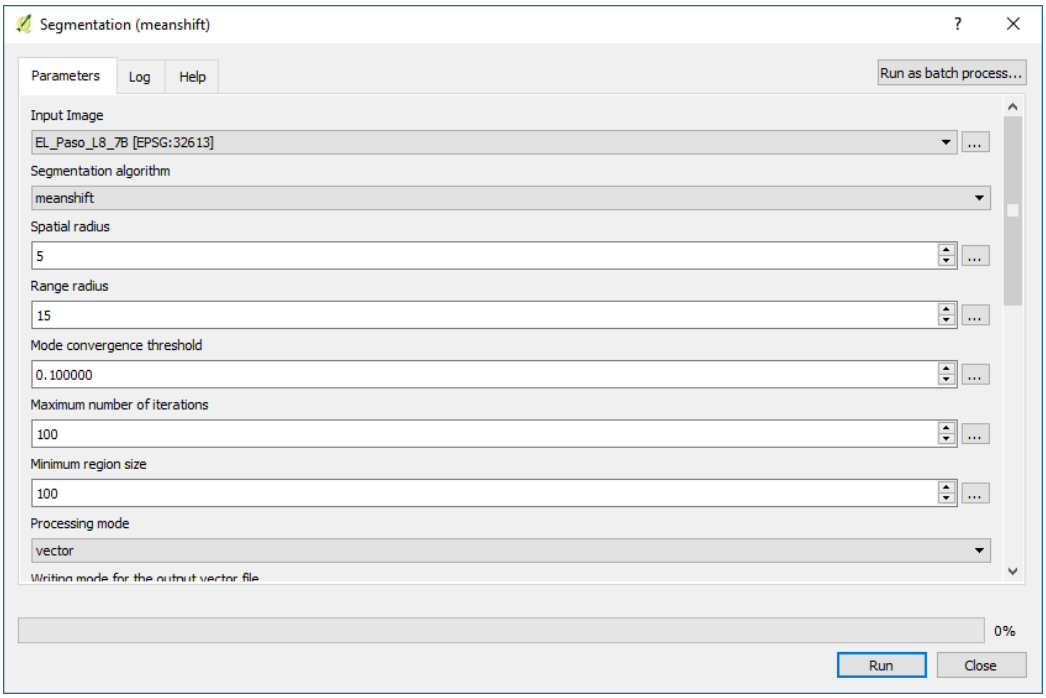
En el menú superior de QGIS, seleccione **Producción> Caja de herramientas (Processing > Toolbox)**, para obtener acceso a los diferentes módulos de la caja de herramientas de procesamiento, uno de ellos debe ser **Orfeo Caja de herramientas (análisis de imágenes)** (**Orfeo Toolbox (Image analysis))**.

Expanda el cuadro de herramientas de Orfeo y el conjunto de herramientas Segmentación. Hay varias opciones para segmentar una imagen, el proceso de agrupar píxeles contiguos con un color similar (valores espectrales) juntos en segmentos; para este ejercicio vamos a tratar la opción de **Segmentación (meanshift) (Segmentation (meanshift))** haciendo doble clic en ella.

La imagen de entrada es "El_Paso_L8_7B [EPSG: 32613]". Mantendremos los valores predeterminados tal y como son. Vamos a definir un archivo de forma que el archivo vectorial de salida, algo significativo como: **"El_Paso_segmented.shp".** Haga clic en los tres periodos a la derecha del campo Archivo de vector de salida, a continuación, guardar en el archivo, vaya a su directorio, y el tipo de **"El_Paso_segmented.shp".**



La herramienta de segmentación debe ser algo como esto:

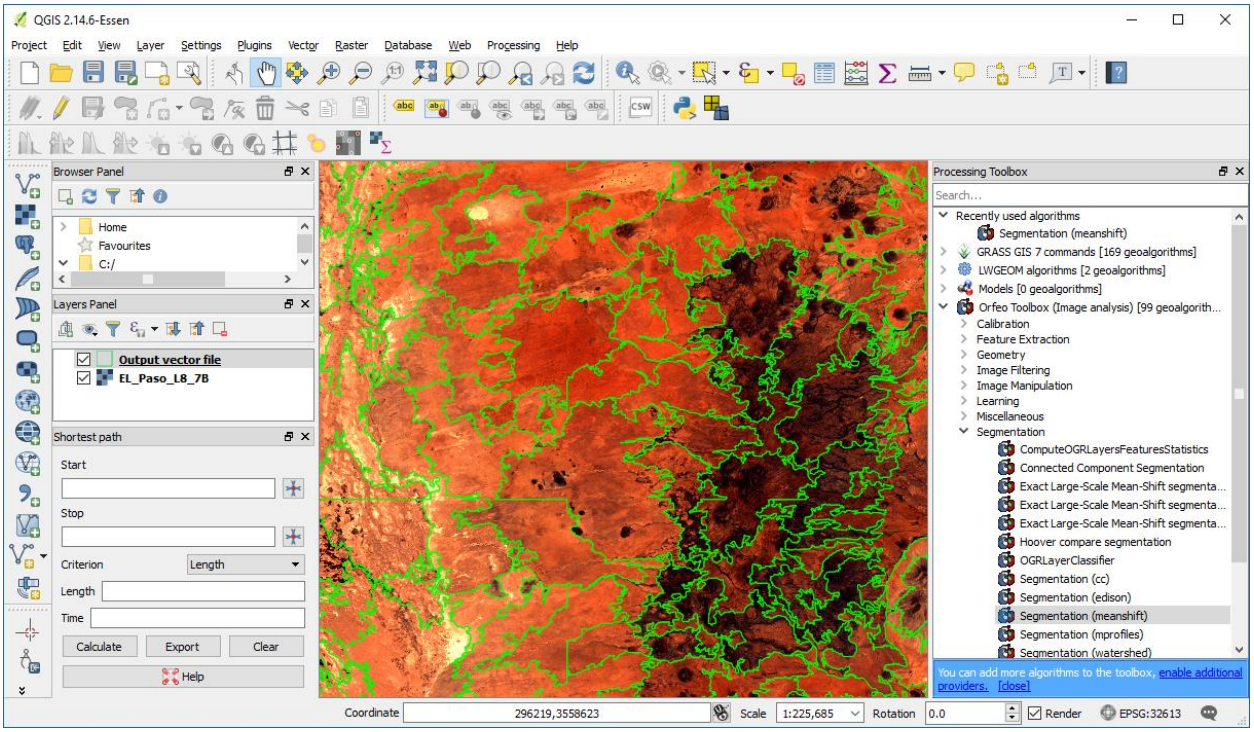


Click the **Run** button and let it process the image, it should take several minutes.

Haga clic en el botón **Ejecutar** y se deja procesar la imagen, se debe llevar varios minutos.

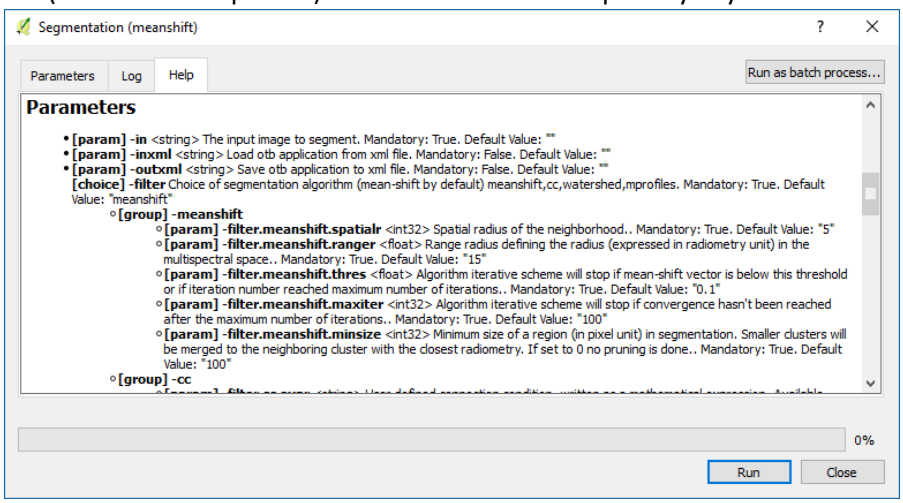
Dependiendo del nivel de resolución y el número de clases que desee alcanzar, puede jugar con una combinación diferente en todos los parámetros de la segmentación de desplazamiento medio hasta obtener resultados satisfactorios. A continuación se muestra el resultado de una ejecución anterior con

un **Radio de Rango (Range radius)** movido desde 15 a 30 y de **Regiones de tamaño mínimo** (**Minimum regions size)** de 100 a 9000 después de ampliar y haga doble clic sobre la capa de salida, al hacer clic en Estilo, al hacer clic en **Relleno (Fill)**, y seleccionando **"Relleno transparente"(Transparent fill)**.



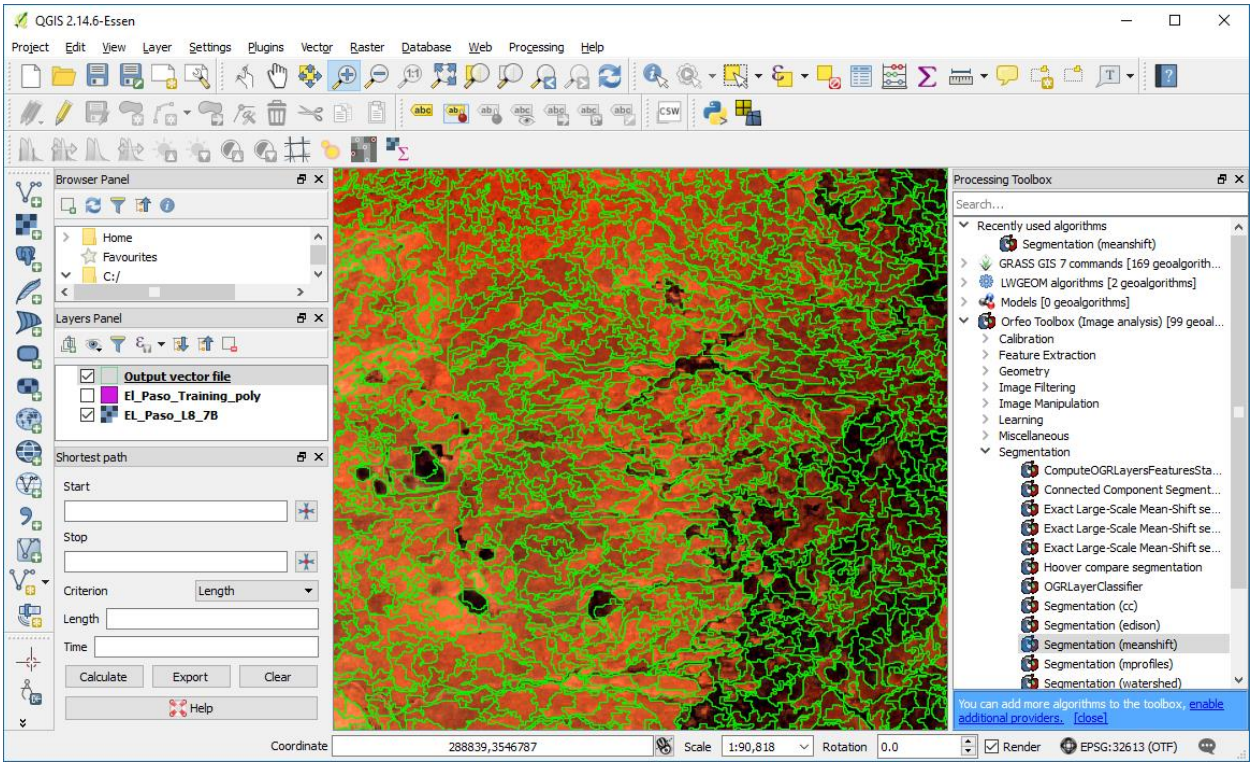
En el ejemplo anterior se puede ver que todavía hay puntos oscuros agregados con el fondo de color más claro. Esto significa que el radio debe ser reducido y también el tamaño de regiones mínimo debe ser menor, pero ahora hemos explorado lo que sucede si se proporcionan grandes números para esos parámetros, se obtienen segmentos que son demasiado grandes y generales.

Abra la herramienta de segmentación (meanshift) de nuevo y volver el **Radio de rango (Range radious)** a 15 (valor por defecto) y tamaño de la región a 100 (también valor por defecto). OTB busca píxeles vecinos cuya posición espacial (en número de píxeles) cae dentro del radio espacial y cuya distancia euclídea (expresada en unidades radiométricas) entre los valores espectrales cae dentro del radio de alcance. Se puede acceder a la explicación de lo que hace la herramienta y de lo que significa este parámetro haciendo clic en la pestaña "Ayuda" proporcionada con la

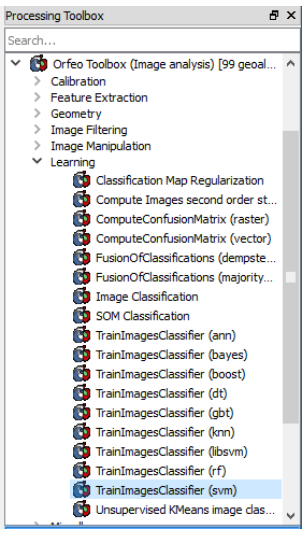


herramienta. Acercar cerca del centro de la captura de pantalla anterior para ver la mayoría de los puntos más oscuros ahora pertenecen a sólo un segmento como se desea.

A continuación se muestra el resultado de los valores por defecto con Radio de rango en 15 (valor predeterminado) y tamaño de región en 100 (también valor predeterminado).

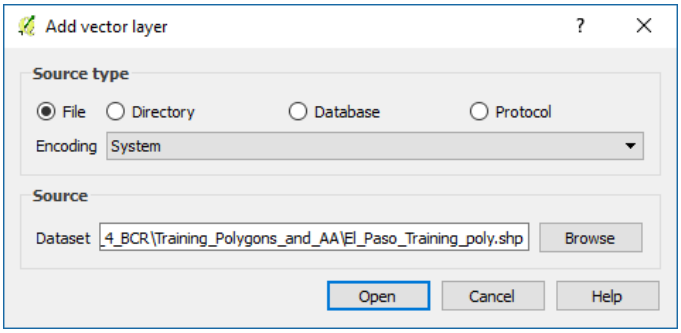


Once we have a segmentation that fits our needs, we need to train a classifier, and algorithm that does the classification work assigning segments to classes, it will use information from the Landsat image (“El_Paso_L8_7B.vrt”) and the training polygons vector data (“El_Paso_Training_poly.shp”).

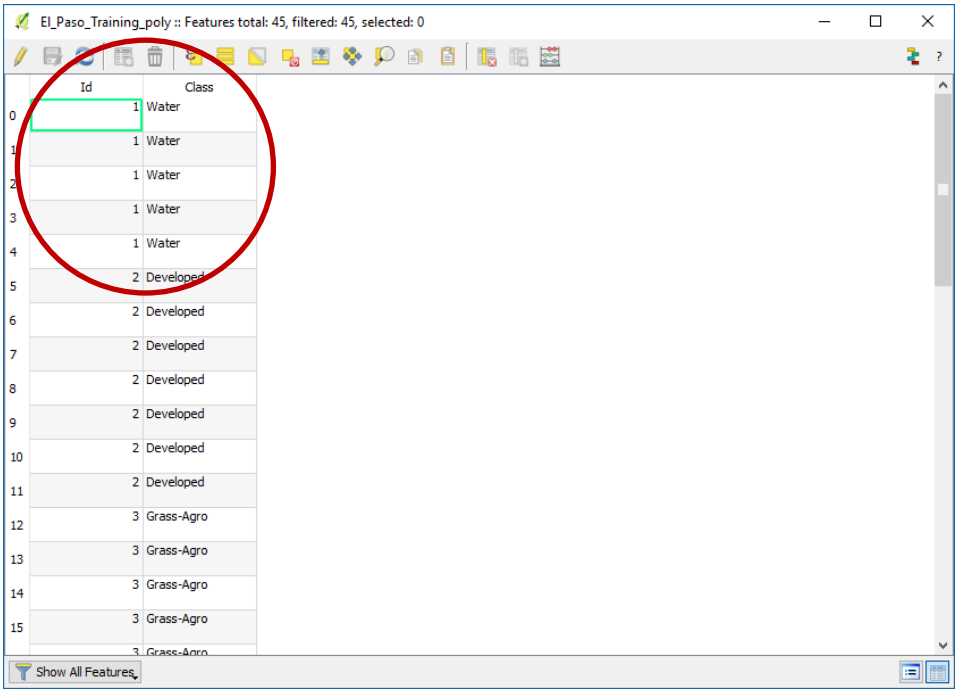


Una vez que tengamos una segmentación que se ajuste a nuestras necesidades, necesitamos entrenar un clasificador y un algoritmo que haga el trabajo de clasificación asignando segmentos a clases, utilizará información de la imagen de Landsat ("El_Paso_L8_7B.vrt") y los datos del vector de polígonos de entrenamiento ("El_Paso_Training_poly.shp").

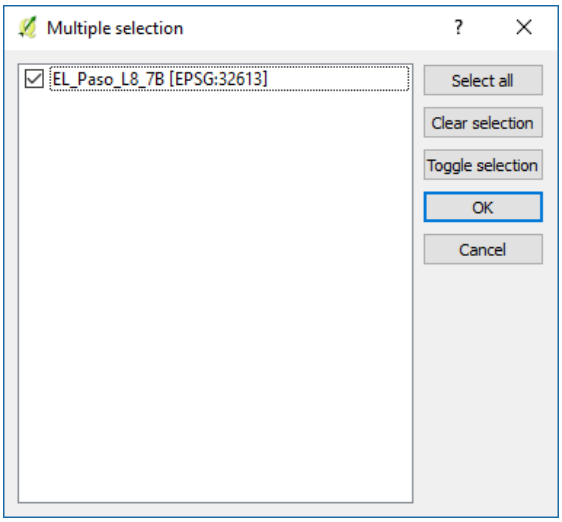
En el lado izquierdo de su ventana QGIS, puede ver varios iconos con un signo más. Estos se refieren a conjuntos de datos que puede agregar a su sesión. Pasa el cursor sobre el primero Hasta que vea una etiqueta llamada **"Agregar capa vectorial" (Add Vector Dataset**), haga clic en ella. Se abrirá un cuadro de diálogo para añadir una capa vectorial, seleccionará el tipo de fuente como un "archivo" y, bajo fuente, navegar hasta donde tenga los polígonos de entrenamiento, en este ejemplo utilizamos: "El_Paso_Training_poly.shp". Haga clic en “**Abrir” (Open).**



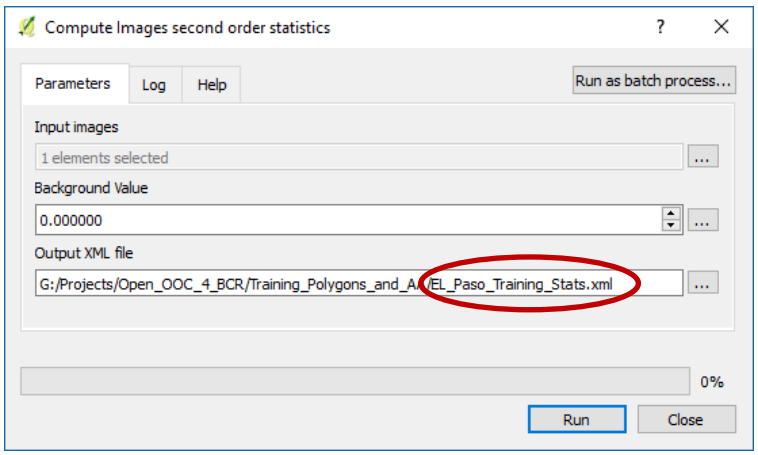
Los datos del vector de entrenamiento deben contener polígonos con un campo entero positivo que representa la etiqueta de la clase. Puede hacer clic con el botón derecho del ratón en el panel Capas: "El Paso Training poly" y hacer clic en **"Abrir tabla de atributo" (Open atribute table**). Se abre la tabla. En este ejemplo se utilizaron 1 para agua (Water), 2 urbano (developed), 3 para pasto y agricultura (Grass-Agro), 4 para arbustos (Brushland), 5 para tierras desnudas (Bareland) y 6 para bosque (Forest). El primer valor está bajo un campo llamado "Id" y el segundo está bajo un campo llamado "class". Éstos se pueden ver en la tabla de "El_Paso_Training_poly.shp" a continuación.



Para tener una idea de nuestros contenidos de datos de imagen, necesitamos calcular la media global y la desviación estándar para cada banda de la imagen Landsat ("El_Paso_L8_7B.vrt") y guardar los resultados en un archivo XML.



En el panel de la caja de herramientas de procesamiento de ir a **Orfeo Caja de herramientas> Aprendizaje> Imágenes de calcular las estadísticas de segundo orden (Orfeo Toolbox > Learning > Compute Images second order statistics)** and double click on it. y haga doble clic en él. Como entrada utilizamos "El_Paso_L8_7B.vrt" (haga clic en los tres puntos para seleccionarlo). El valor de fondo es cero y definimos un archivo xml de salida (haga clic en los tres puntos y guarde en archivo): "El_Paso_Training_Stats.xml". **"Ejecutar" (Run).**

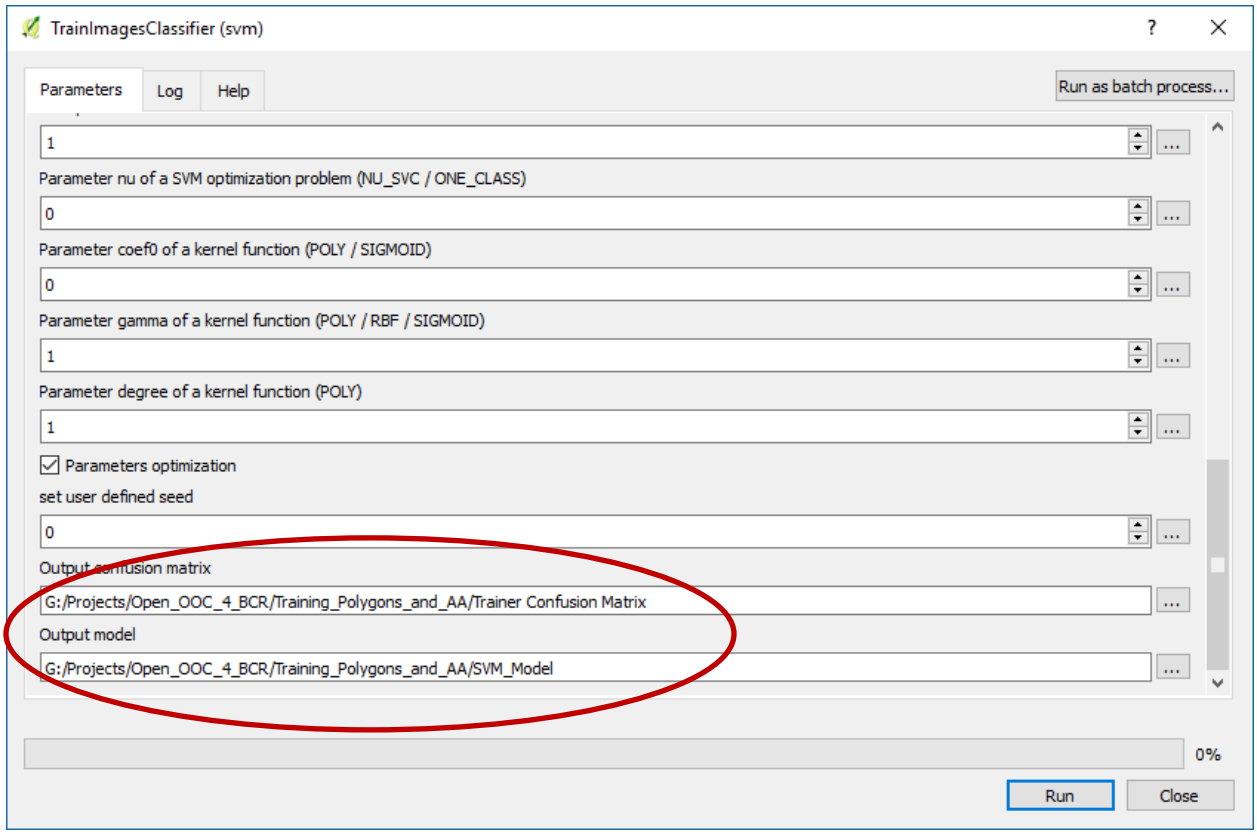


Eso calcula la media global y la desviación estándar de nuestra imagen y la almacena en el archivo "El_Paso_Training_Stats.xml". Este archivo puede ser utilizado opcionalmente por el algoritmo TrainerImagesClassifier (svm), o puede generar las estadísticas en sí mismo.

En el panel de la caja de herramientas de procesamiento de ir a **Orfeo Caja de herramientas> Aprendizaje> TrainImagesClassifier (SVM)** (**Orfeo Toolbox > Learning > TrainImagesClassifier (svm))** y haga doble clic en él. Esta herramienta realiza un entrenamiento clasificador a partir de una imagen y datos vector de entrenamiento. Los datos del vector de entrenamiento deben contener polígonos con un campo entero positivo que representa la etiqueta de la clase (que es el campo "Id" en nuestro caso).

Compruebe la imagen y los archivos de polígonos de entrenamiento que serán utilizados, estamos utilizando la imagen Landsat ("El_Paso_L8_7B.vrt") como los datos de **Imagen de entrada (Input Image)** y polígonos de entrenamiento de vectores ("El_Paso_Training_poly.shp") como la **Entrada de datos vectoriales** (**Input Vector Data)**. Vamos a utilizar la mayor parte de los valores por defecto en la herramienta **TrainImagesClassifier (SVM),** a excepción del **Nombre del campo discriminante (Name of**

the Discriminating field) (un campo entero) que va a escribir **"Id".** Al desplazarse hasta cerca del final del diálogo, encontrará un lugar para especificar la matriz de confusión de salida y el modelo de salida. El primero nos dará una idea de cómo son similares y discernibles las clases entre ellos. Cuantos más valores aterricen en la diagonal, la menor confusión (llame al archivo de salida algo significativo como "Trainer Confusion Matrix"). El segundo nos dará el modelo necesario para realizar la clasificación (llámelo "SVM_Model") en el siguiente paso. Un modelo de máquina de soporte vectorial (SVM) es un modelo de aprendizaje por máquina supervisado con algoritmos de aprendizaje asociados en QGIS que analizan los datos utilizados para la clasificación orientada a objetos como es éste caso.



Haga clic en **Ejecutar (Run)** para iniciar la herramienta TrainImageClassifier.

La herramienta: TrainImageClassifier produce una matriz de confusión en formato de texto, se puede añadir una extensión ".txt" para que sea reconocible por editores de texto como el notepad (u otro editor de texto). Una vez que pueda ser abierto, nos dará una idea de cómo las clases son potencialmente clasificadas con los polígonos de entrenamiento utilizados. También se puede importar a una hoja de cálculo y volver a formatear para incluir los nombres de las clases. Los niveles de confianza de todas las etiquetas de clase se estiman a partir de una comparación de los mapas de clasificación con las clases verificadas en campo, lo que da como resultado una matriz de confusión. Las etiquetas de referencia para cada clase son filas, una por clase definida en los polígonos de entrenamiento. Las etiquetas Producidas son columnas, una por clase en la clasificación resultante en la clasificación. Una matriz de confusión aceptable tendrá la mayoría de los valores más altos en la diagonal, lo que significa que la clasificación se parece a lo que se ha observado en el campo, por ejemplo, 128 píxeles

referenciados como Agua durante la comprobación de tierra fueron correctamente clasificados como Agua en la siguiente matriz.

Los valores que no están en la diagonal están confundidos. Por ejemplo, hay 5 píxeles referenciados como Urbanos en la siguiente matriz que fueron clasificados erróneamente como Agua. Estos pueden ser asfalto u otra cubierta terrestre que el clasificador confunde con agua, por lo tanto, el nombre dado; matriz de confusión. Para cada mapa de clasificación, estas matrices de confusión se utilizan para estimar la "confiabilidad" de cada etiqueta de clase. Cuantos más valores aterricen en la diagonal, más creencia en que el clasificador está haciendo un buen trabajo para una clase dada. Si la discernibilidad es pobre, eso muestra muchos píxeles no en la diagonal; Puede ser necesario realizar más trabajo de campo o re-selección de los polígonos de entrenamiento usados. Este es un ejemplo de lo que se puede esperar como la matriz Confusión, que da una idea de la discernibilidad de las clases:

#Produced labels (columns)	**1**	**2**	**3**	**4**	**5**	**6**
#Reference labels (rows)						
1 Water	128	3	4	5	0	2
2 Developed	5	128	0	5	0	0
3 Grass-Agro	6	0	144	2	0	1
4 Brushland	1	2	0	151	0	0
5 Barren land	0	0	0	6	159	0
6 Forest	0	1	1	0	0	138

Más adelante en el proceso se utilizará una matriz de confusión cuando se realice la evaluación de la exactitud comparando el mapa creado por la Clasificación Orientada a Objetos con los recursos de validación en campo disponibles.

3.4 Segmentación Raster y clasificación orientada a objetos

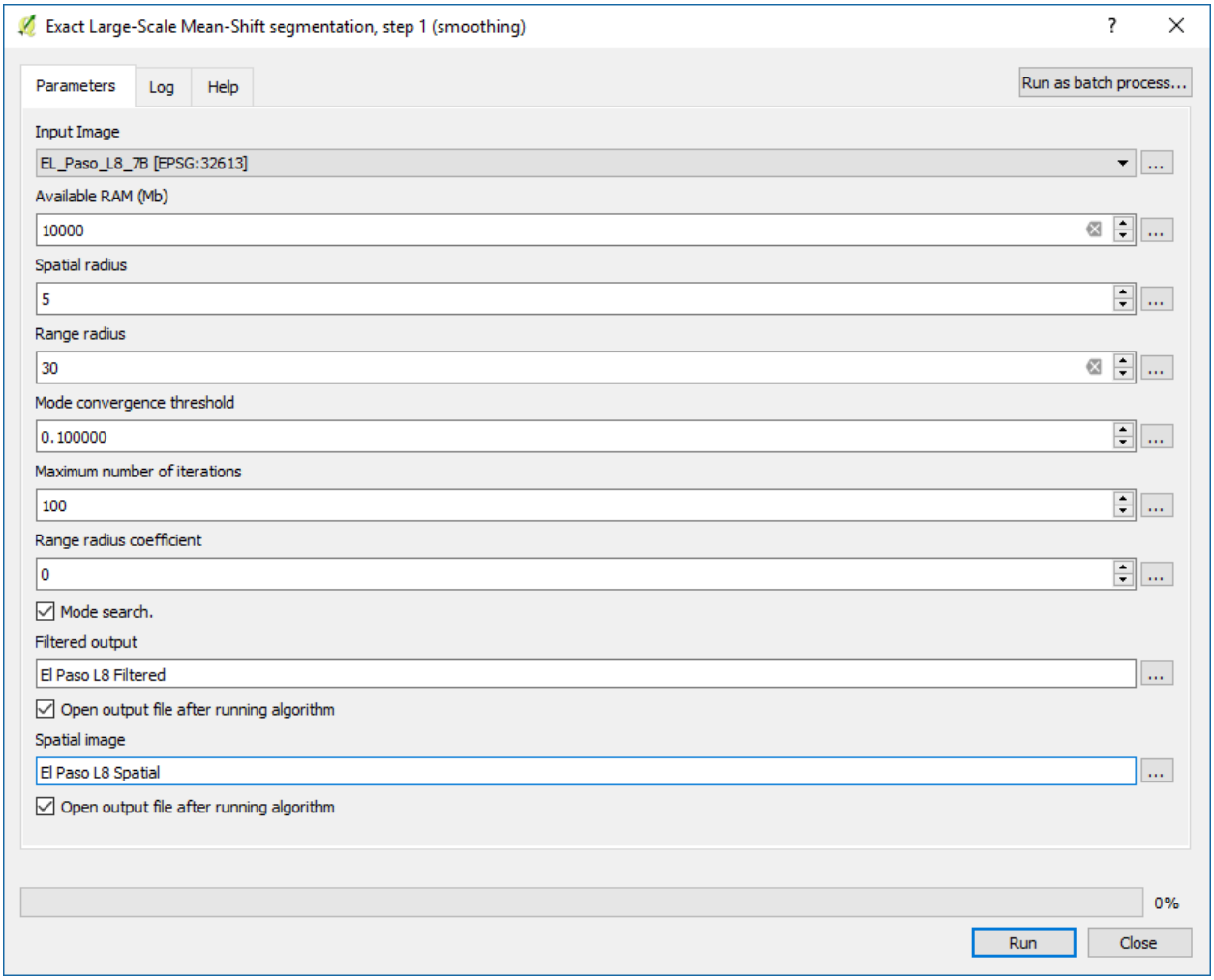
La segmentación ráster es el proceso de partición de una imagen digital en múltiples segmentos, conjuntos de píxeles u objetos. El objetivo de la segmentación es pre-procesar o simplificar la representación de una imagen en objetos que sean más significativos y más fáciles de analizar. Los píxeles espacialmente cercanos y espectralmente similares se agrupan en objetos. Con las herramientas de segmentación y clasificación de OTB, los usuarios pueden preparar rasters segmentados como un primer paso para crear conjuntos de datos raster clasificados orientados a objetos.

Ahora que tenemos una mejor comprensión de la imagen de Landsat 8, los polígonos de entrenamiento, y sus clases, una segmentación raster es necesaria para la clasificación orientada a objetos. Cada paso de procesamiento puede tomar varios minutos dependiendo de la cantidad de datos a procesar y la velocidad de la computadora. El flujo de trabajo en QGIS viene en tres procedimientos principales:

- Secuencia raster de desplazamiento medio (mean-shift raster segmentation) de imágenes grandes (4 pasos).
- Preparación de datos de referencia (polígonos de formación).
- Clasificación Orientada a Objetos utilizando el algoritmo SVM.

3.4.1 Flujo de trabajo de segmentación de cambio medio (Mean-shift segmentation workflow)

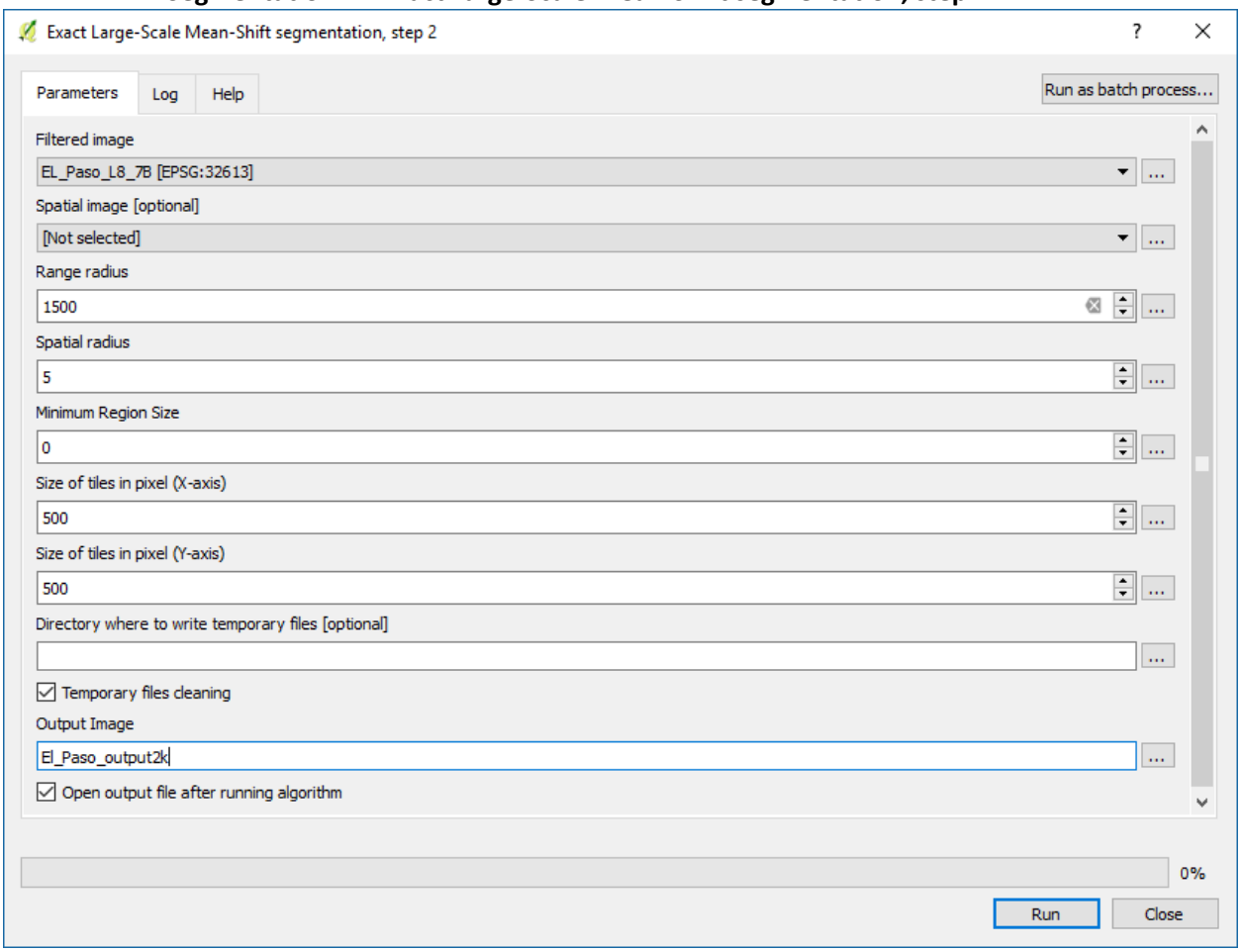
a. Realice un filtro de desplazamiento medio de "El_Paso_L8_7B.vrt". Esto ayudará a extraer objetos homogéneos de una imagen filtrada. En el panel de la caja de herramientas de procesamiento de ir a **Orfeo Toolbox > Image Filtering > Exact Large-Scale Mean-Shift segmentation, step 1 (Smoothing).**



i. Use “El_Paso_L8_7B.vrt” como la **Imagen de entrada (Input image)**.
ii. Establecer la **memoria RAM disponible (Mb)** (**Available RAM)** a lo que se quiere poner a disposición de los SIG para procesar. Cuanta más RAM tenga su computadora disponible, más rápido se ejecutará el proceso.
iii. Establecer el **Radio de rango (Range Radius)** a 30.

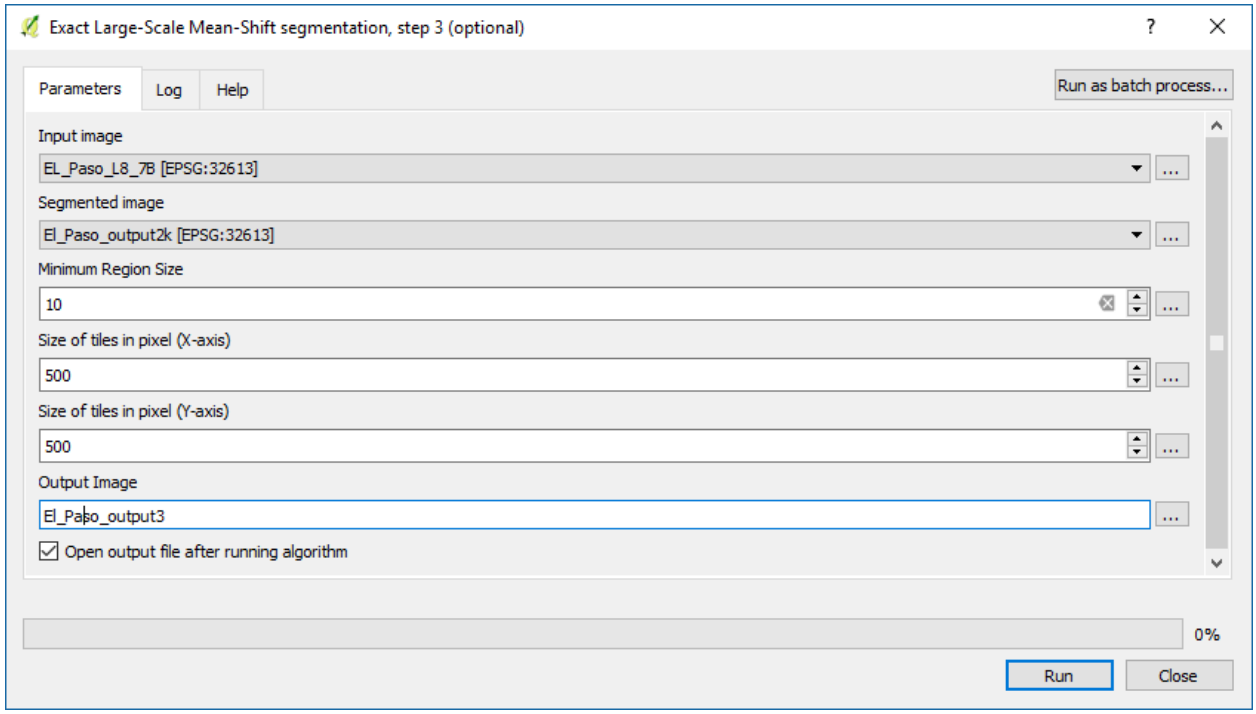
iv. Haga clic en los tres puntos y seleccionar “guardar en el archivo” para definir una trayectoria de **“salida filtrada” (filtered output”)** (nombre el archivo: “El_Paso_L8_filtered.tif”) y la **imagen espacial (spatial image)** (nombre el archivo: “El_Paso_L8_spatial.tif”).
v. Hacer el parámetro de **Modo de búsqueda (Mode Search)** deshabilitado.
vi. Deje todos los otros parámetros con los valores por defecto y haga clic en **Ejecución (Run).**
vii. Tenga en cuenta que "EL_Paso_L8_filtered.tif" tiene menos variación que la imagen original, parece más suave. Algunas de las cabezas y las colas en el original se han filtrado dando por resultado una imagen más lisa. Esto se puede verificar haciendo clic con el botón derecho en las propiedades de la imagen, haciendo clic en Histograma y calculando el histograma para cada capa. Puede capturar la pantalla de ambos histogramas como referencia, la imagen filtrada tiene más píxeles cerca del centro de las distribuciones y menos hacia los extremos (cabezas y colas de la curva de distribución).

b. Convierte la imagen filtrada en un objeto de una banda. Esta etapa producirá una imagen etiquetada en la que los píxeles vecinos cuya distancia de alcance esté por debajo del radio de alcance (y opcionalmente la distancia espacial por debajo del radio espacial) se agruparán en el mismo grupo. Ir al **Processing-Toolbox --> Orfeo Toolbox (image analysis) --> Segmentation --> Exact Large-Scale Mean-Shift segmentation, step 2.**



i. Use “EL_Paso_L8_filtered.tif” como **Imagen filtrada (Filtered Image)** (salida del paso anterior).
ii. Deja **Spatial image (optional)** vacía (no seleccionado).
iii. Aumentar el “**Rango de radio” (Range Radius)** a “1500”.
iv. Haga clic en los tres puntos y seleccionar “guardar en el archivo” para guardar la **Imagen de salida (Output Image)** como un nombre de archivo significativo para usted; En este ejemplo usamos "El_Paso_L8_output2k.tif".
v. Deje todos los otros parámetros con los valores por defecto y haga clic en **Ejecución (Run).**
vi. Observe las agrupaciones de píxeles en objetos en la imagen de salida ("El_Paso_L8_output2k"). Puede modificar el “**Rango de radio” (Range Radius)** a otros valores. Cuando las características distintas del suelo (es decir, las manchas oscuras mencionadas anteriormente), están en segmentos por sí mismos, usted sabe que tiene buenos parámetros, rasgos continuos en el suelo partidos excesivamente en pequeños segmentos o grandes segmentos que cubren varias características o clases, necesitan ser refinados. Repita como sea necesario para satisfacer sus necesidades de clasificación si los resultados iniciales no son satisfactorios.

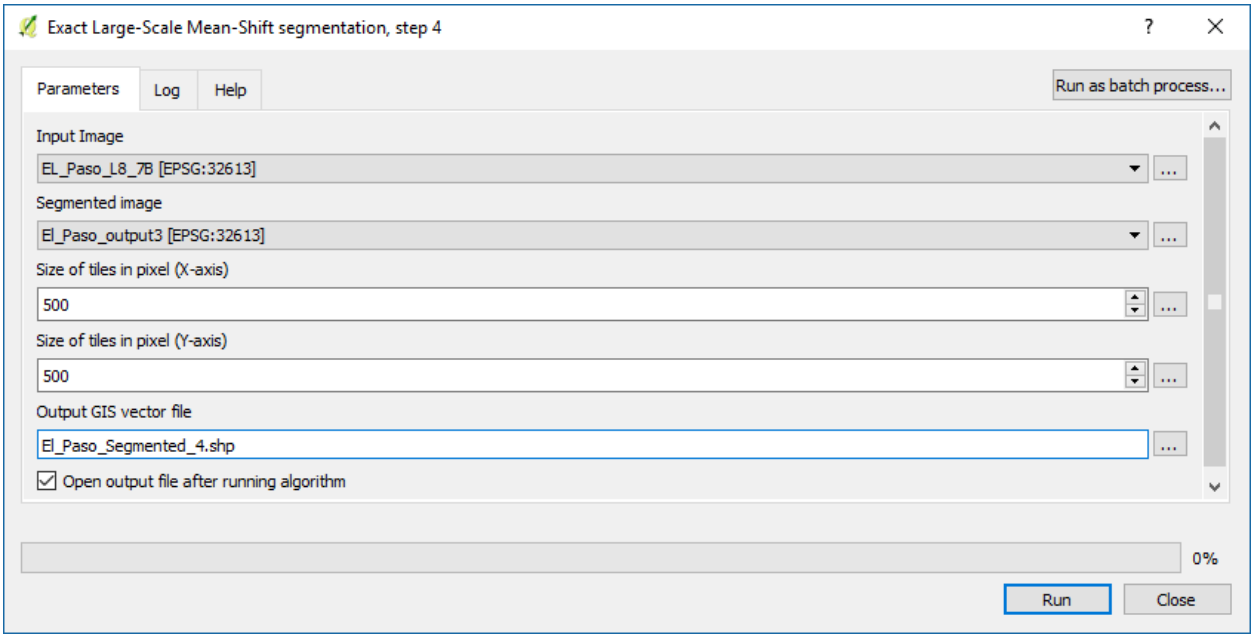
c. Combinar pequeños objetos cuyo tamaño en píxeles es menor que **minsize** parámetro con la región adyacente que tiene distancia radiométrica cerca euclidiana y tamaño aceptable. Ir a **Processing-Orfeo Toolbox (image analysis) --> Segmentation --> Exact Large-Scale Mean-Shift segmentation, step 3 (optional).**



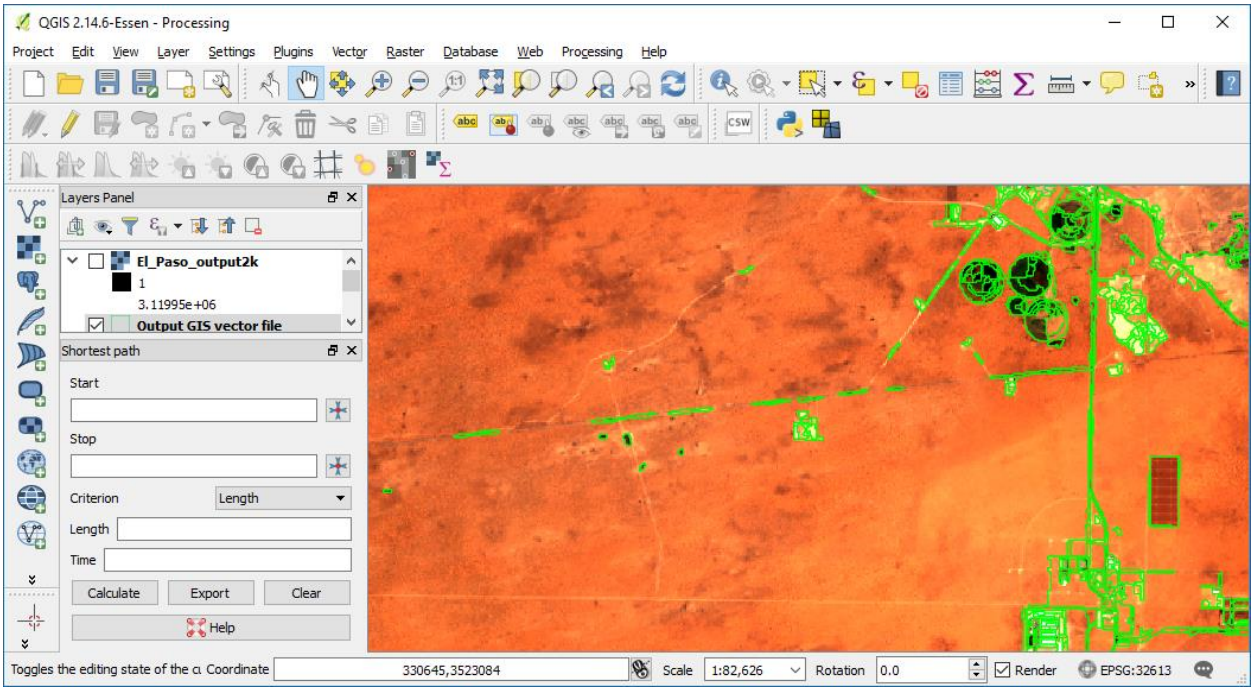
i. La **Imagen de entrada (Input image)** es la pila de imágenes original (“EL_Paso_L8_7B”).
ii. La **Imagen segmentada (Segmented image)** es el resultado de la segmentación raster anterior (paso 1b).
iii. The **Minimum Region size** is defined as 10.

iv. Haga clic en los tres puntos y seleccionar “guardar en el archivo” para guardar la **Imagen de Salida (Output Image)** un nombre de archivo significativo para usted. En este ejemplo usamos "El_Paso_output3.tif".
v. Deje todos los otros parámetros con los valores por defecto y haga clic en **Ejecución (Run).**
vi. Puede notar que los objetos pequeños han sido absorbidos por objetos más grandes y contiguos que están cerca radiométricamente (una distancia euclidiana en valores digitales que está dentro de una tolerancia predefinida).

d. Transforme los objetos de imagen ráster en polígonos (objetos vectoriales). Cada polígono contiene campos adicionales: media y varianza de cada canal de la imagen de entrada (parámetro), etiqueta de imagen de segmentación y número de píxeles en el polígono. Ir a **Processing-Orfeo Toolbox (image analysis) --> Segmentation-Exact Large-Scale Mean-Shift segmentation, step 4.**



i. Utilice "El_Paso_L8_7B.vrt". como la **Imagen de entrada (Input image)**.
ii. Use “El_Paso_output3.tif” como la **Imagen segmentada** (**Segmented Image).**
iii. Haga clic en los tres puntos, selecciona “Guardar archivo” y el nombre el **archivo vectorial de salida SIG (Output GIS vector file).** Este ejemplo utiliza “El_Paso_Segmented_4.shp”.
iv. Deje todos los otros parámetros con los valores por defecto y haga clic en **Ejecución (Run).**



i. Observe que ahora tenemos objetos vectoriales que tienen estadística de media y varianza para cada una de las 7 bandas en las imágenes de entrada (0-6). Clic derecho en el **Archivo vectorial de salida SIG (Output GIS vector file)** en su sesión de QGIS ("El_Paso_Segmented_4.shp" en la realidad) y haga clic en **Abrir tabla de atributos** (**Open Attribute Table)**. Compruebe que los valores de etiqueta, nbPixels, media y varianza se han poblado para cada banda (columna).

Output GIS vector file :: Features total: 192678, filtered: 192678, selected: 0

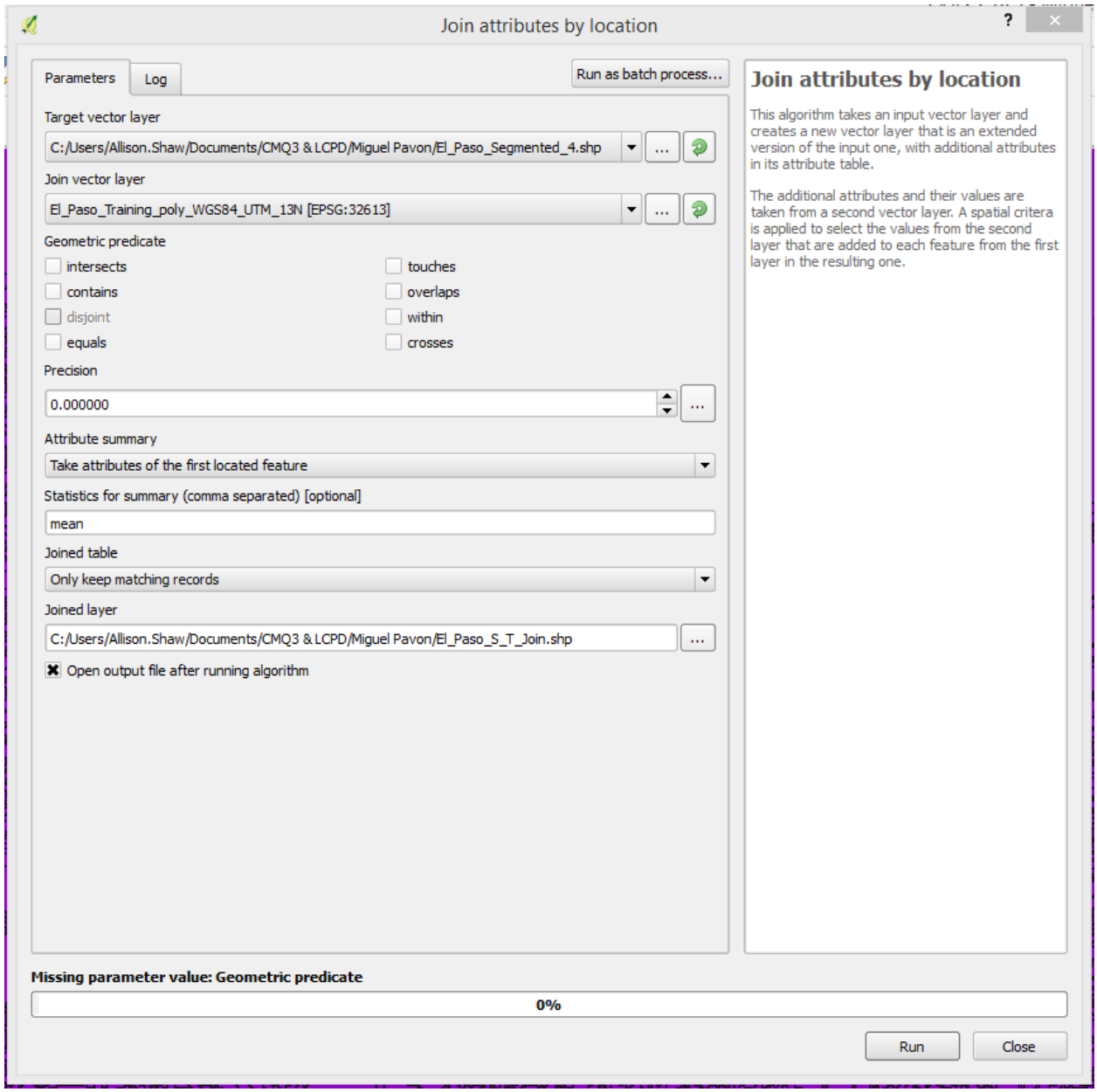
	label	nbPixels	meanB0	meanB1	meanB2	meanB3	meanB4	meanB5
0	1	19718168	0.094608888030...	0.275439739227...	0.330502420663...	0.238587483763...	0.551012992858...	1.149394750
1	4	4103039	12602.39160156...	12572.40234375...	13067.26269531...	15758.64257812...	19901.46679687...	22304.12695
2	93	24	13466.79199218...	13455.62500000...	13941.29199218...	16065.54199218...	19210.54101562...	23482.75000
3	85	20	13397.75000000...	13365.75000000...	13734.15039062...	15657.54980468...	18769.05078125...	23712.80078
4	103	15	13501.79980468...	13512.86621093...	14064.73339843...	16265.46679687...	19482.06640625...	23667.19921
5	83	52	13616.05761718...	13660.50000000...	14247.50000000...	16521.90429687...	19873.71093750...	24201.30859
6	90	36	14015.52734375...	14145.27734375...	14941.55566406...	17488.97265625...	21064.86132812...	25007.13867
7	331	14	14348.28613281...	14562.85742187...	15530.64257812...	18408.28515625...	22183.57226562...	24282.57226
8	495	18	13097.22265625...	13005.77734375...	13243.72265625...	15007.83300781...	18390.38867187...	22302.33398
9	732	25	13546.75976562...	13596.59960937...	14216.04003906...	16417.11914062...	20588.88085937...	23189.35937
10	852	12	12210.08300781...	11949.83300781...	12195.58300781...	13315.33300781...	20328.50000000...	19670.58398
11	1129	13	12309.46191406...	11958.46191406...	11835.61523437...	12605.92285156...	17257.46093750...	17260.46093
12	1260	12	13577.58300781...	13529.16699218...	13865.66699218...	15815.83300781...	19057.58398437...	20733.41601
13	1365	17	14553.76464843...	14677.52929687...	15295.35253906...	17505.64648437...	19836.64648437...	20095.47070
14	1261	193	13952.73046875...	14034.22265625...	14732.34179687...	17060.87500000...	20674.81250000...	22896.20703
15	1400	31	14671.54882812...	14775.25781250...	15427.48339843...	17806.19335937...	21565.54882812...	21854.87109

Show All Features

3.4.2 ***Preparación de los datos de referencia***

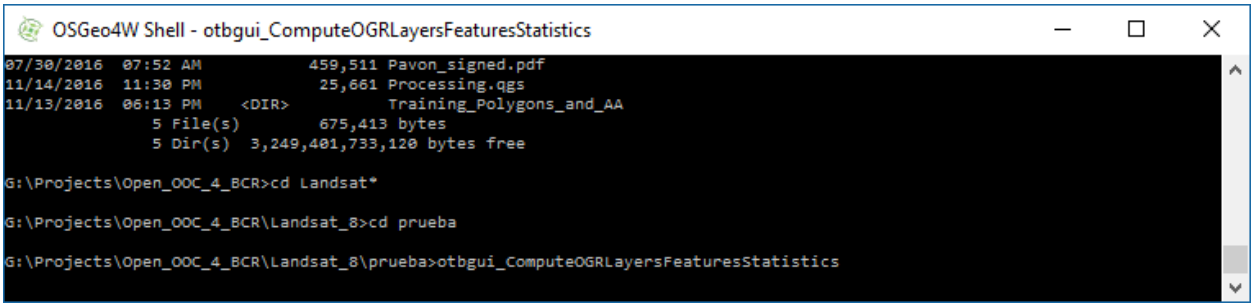
Para este paso usaremos datos de vector de polígonos de entrenamiento ("El_Paso_Training_poly.shp") creados como polígonos de muestra y utilizados antes durante la exploración de datos.

a. Úniremos el atributo de clases en los polígonos de entrenamiento con los segmentos de imagen. Añadir conjunto de datos vectorial “El_Paso_Training_poly.shp” a su Sesión QGIS (Icono **Add Vector Layer** en la barra de la izquierda).
b. Ir al menú principal en QGIS (barra superior) **Vector --> Data Management Tools --> Join Attributes by Location.**
    i. El **Archivo Vector objetivo** (**Target vector file)** es la salida del paso 1d (“El_Paso_Segmented_4.shp”).
    ii. Establecer la **capa de vectores de Ingreso (Join vector layer) como** “El_Paso_Training_poly.shp”.
    iii. Haga clic en los tres puntos para el **Archivo Unido (Joined Layer)**, seleccione “Guardar en archivo” y guardar Como "El_Paso_S_T_Join.shp"
    iv. Deje el resto de los parámetros con los valores por defecto y haga clic en **Ejecutar (Run).**

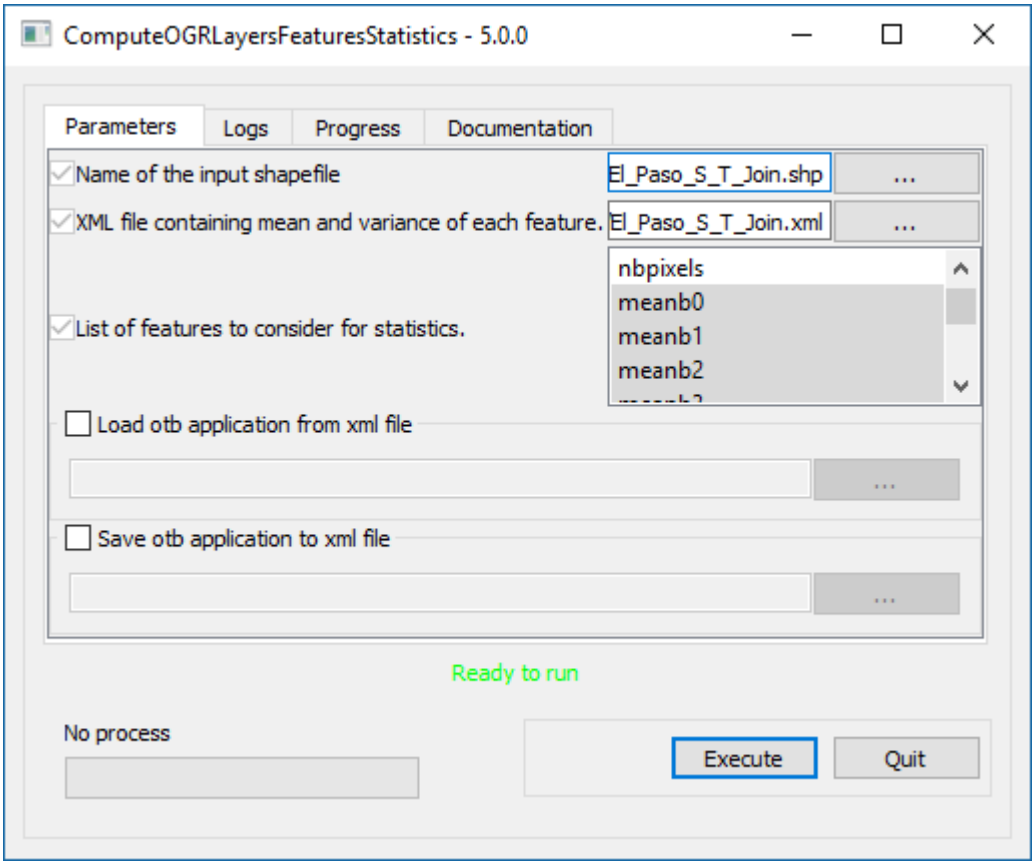


3.4.3 Clasificación basada en objetos utilizando el algoritmo SVM

a. Abra la Interfaz gráfica de la caja de herramientas Orfeo buscando la OSGeo4W Shell. En Windows se puede encontrar aquí: **Start --> All Programs --> OSGeo4W --> OSGeo4W Shell.**



b. En la consola de comandos, escriba: **otbgui_ComputeOGRLayersFeaturesStatistics,** para abrir una nueva interfaz gráfica de usuario (GUI).



i. Haga clic en los tres puntos por **Nombre del archivo de entrada (Name of the input file)** y vaya a "El_Paso_S_T_Join.shp".

ii. Haga clic en los tres puntos por el **archivo XML que contiene la media y la varianza de cada función (XML file containing mean and variance of each feature)** y teclee: "El_Paso_S_T_Join.xml".

iii. Seleccione de la **lista de características a tener en cuenta para las estadísticas (List of features to consider for statistics):** meanb0, meanb1, meanb2, meanb3, meanb4 meanb5 and meanb6.

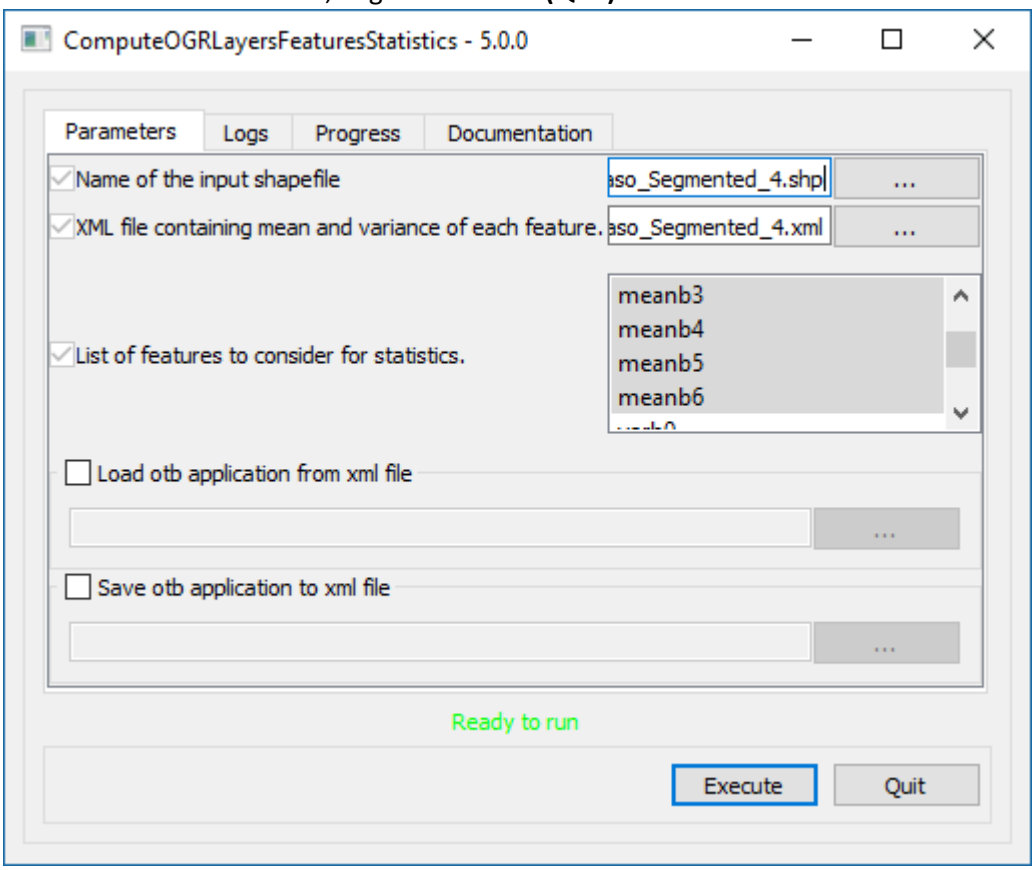
iv. Haga clic en: **Ejecutar (Execute).**

v. Observe que se creó un nuevo archivo xml ("El_Paso_S_T_Join.xml") con un contenido que define la media y la desviación estándar de cada columna considerada en el paso 3.b.iii:

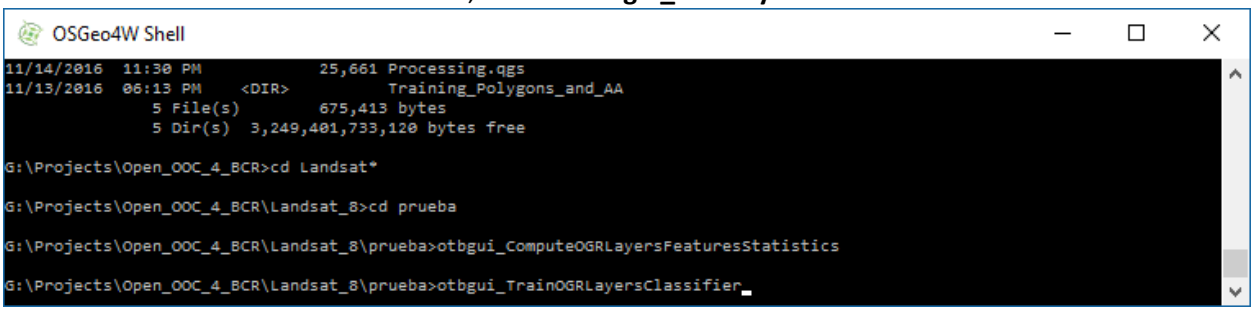

```
<?xml version="1.0"?>

<FeatureStatistics><Statistic name="mean"><StatisticVector
value="12233.5"/><StatisticVector value="11915.7"/><StatisticVector
value="12262"/><StatisticVector value="13163.8"/><StatisticVector
value="20137.1"/><StatisticVector value="18113"/><StatisticVector
value="15200.3"/></Statistic><Statistic name="stddev"><StatisticVector
value="1676.7"/><StatisticVector value="1970.47"/><StatisticVector
```

value="**2305.63**"/><StatisticVector value="**3251.66**"/><StatisticVector value="**4175.49**"/><StatisticVector value="**3956.02**"/><StatisticVector value="**4072.67**"/></Statistic></FeatureStatistics>

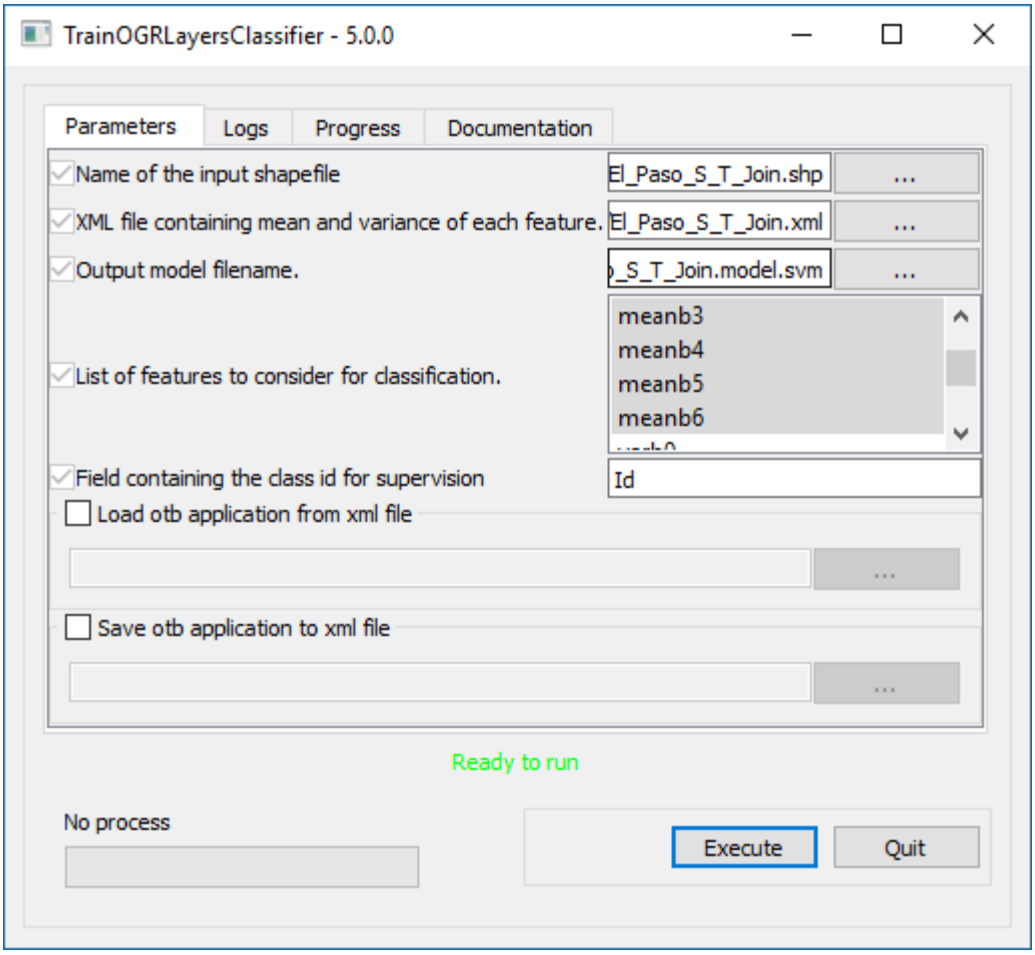
vi. Repita los pasos i.-iv. Pero cambiando "Nombre del shapefile de entrada" a "El_Paso_Segmented_4.shp" y "Archivo XML que contiene la media y la varianza de cada característica a" El_Paso_Segmented_4.xml "para calcular estadísticas xml utilizando los mismos parámetros que antes. Cuando el texto verde "listo para funcionar" (Ready to Run) vuelve a aparecer, haga clic en **"Ejecutar"** (**Execute**), cuando este hecho, haga clic en **Salir (Quit).**



c. En la consola de comandos, escriba: **otbgui_OGRLayerClassifier.** Se abre una nueva GUI.



i. Haga clic en los tres puntos por **Nombre de archivo de forma de entrada (Name of input shapefile)** y seleccione “El_Paso_S_T_Join.shp”.
ii. Haga clic en los tres puntos de **archivo XML que contienen media y la varianza de cada función (XML file containing mean and variance of each feature)** y seleccione “El_Paso_S_T_Join.xml”.
iii. Haga clic en los tres puntos de **salida de la señal nombre del modelo (Ouput model name**) y teclee: “El_Paso_S_T_Join.model.svm”.
iv. Seleccione de la **lista de características a tener en cuenta para la clasificación (List of features to consider for classification):** meanb0, meanb1, meanb2, meanb3, meanb4 meanb5 and meanb6.
v. En el **campo que contiene el ID de clase de la supervisión (Field containing the class id for supervisión),** escriba el campo con el valor numérico: “Id”.



vi. Haga clic en: **Ejecutar (Execute).** Después de terminado, haga clic en **Salir (Quit).**
vii. Observe que hay notificaciones en el shell de comandos y también hay un nuevo modelo en su directorio llamado "El_Paso_S_T_Join.model.svm" que se puede abrir con Notepad u otro editor de texto y debe ser similar a:

El_Paso_S_T_Join.model.svm - Notepad

File Edit Format View Help

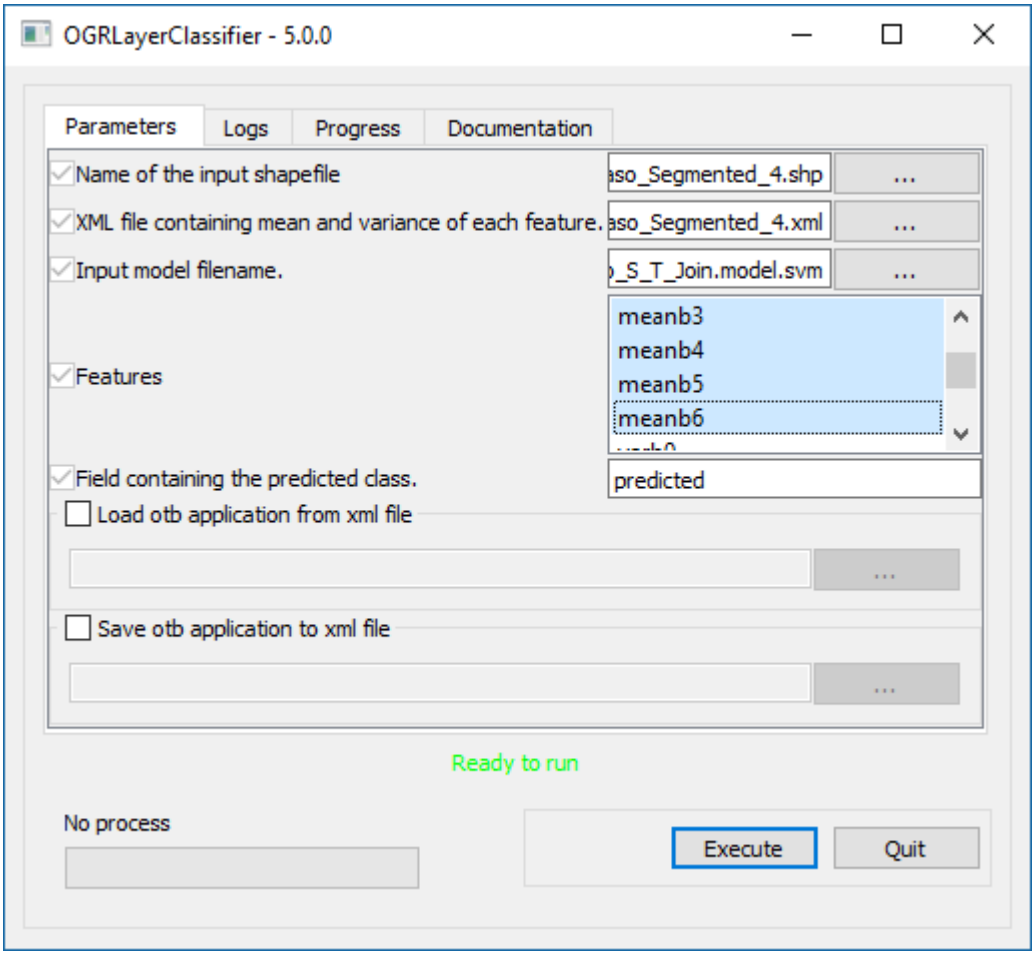
```
svm_type c_svc
kernel_type linear
nr_class 6
total_sv 118
rho -1.91128 -3.09791 -2.9023 -1.0501 -0.574891 -0.44408 -3.56953 0.631438 -0.154755 -
1.06383 1.19069 0.513531 2.82539 0.496643 0.436799
label 3 4 6 5 1 2
nr_sv 30 27 10 6 16 29
SV
0.8705505632961241 0 0 0.8705505632961241 0.8705505632961241 1:-0.26961966 2:-0.25678449
3:-0.25279747 4:-0.067684239 5:-1.1318672 6:-0.74745039 7:-0.44360516
0 0.2881804030260093 0 0.8705505632961241 0.8705505632961241 1:-0.88775571 2:-0.9465078
3:-1.1070905 4:-0.91377522 5:-1.8883441 6:-2.0532819 7:-1.7007676
0.8705505632961241 0 0.8705505632961241 0.5155751139340982 0.8705505632961241 1:1.1308742
2:1.1746296 3:1.2501019 4:1.2692312 5:0.63154068 6:1.2275882 7:1.1586257
0 0 0 0 0.07549942109361806 1:-0.6438434 2:-0.41068277 3:-0.16256826 4:0.0057890503
5:0.9906125 6:-0.78761268 7:-1.0938821
0.8705505632961241 0 0.4872095764884074 0 0 1:0.38519266 2:0.76212147 3:1.2737028
4:1.7371412 5:1.2517907 6:2.271461 7:1.4416767
0.8705505632961241 0 0 0 0 1:0.076779001 2:0.16535613 3:0.23558713 4:0.35698317 5:-
```

d. En la consola de comandos, escriba: **otbgui_OGRLayerClassifier.** Se abre una nueva GUI.

OSGeo4W Shell

```
optimization finished, #iter = 16
nu = 0.030075
obj = -2.401287, rho = 0.513531
nSV = 5, nBSV = 3
*
optimization finished, #iter = 7
nu = 0.120910
obj = -1.878415, rho = 2.825394
nSV = 4, nBSV = 1
.*.*
optimization finished, #iter = 296
nu = 0.048472
obj = -3.925506, rho = 0.496643
nSV = 9, nBSV = 4
.*
optimization finished, #iter = 224
nu = 0.176169
obj = -19.569696, rho = 0.436799
nSV = 27, nBSV = 22
Total nSV = 118
G:\Projects\Open_OOC_4_BCR\Landsat_8\prueba>otbgui_OGRLayerClassifier
```

Esta aplicación aplicará un modelo de aprendizaje de máquina entrenado en las características seleccionadas para obtener una clasificación de cada polígono contenido en una capa. La lista de características debe coincidir con la lista utilizada para la formación. La clase de cobertura de terreno predicha se escribe para cada geometría en un campo de atributo definido por el usuario que hemos denominado "Predicted”.



i. Haga clic en los tres puntos para el **nombre de archivo de forma de entrada (Name of input shapefile)** y utilizar la salida de la segmentación: "El_Paso_Segmented_4.shp".
ii. Haga clic en los tres puntos de **archivo XML que contienen media y la varianza de cada función (XML file containing mean and variance of each feature)** y seleccione "El_Paso_Segmented_4.xml".
iii. Haga clic en los tres puntos por **Nombre de modelo de entrada** (**Input model name)** y use "El_Paso_S_T_Join.model.svm".
iv. Seleccione de la lista de **Características (Features):** meanb0, meanb1, meanb2, meanb3, meanb4 meanb5 y meanb6.
v. Use "predicted" para el **campo que contiene la clase predicha** (**Field containing the predicted class).**
vi. Haga clic en: **Ejecutar (Execute).** Se debe tomar un par de minutos, después de terminado, haga clic en **Salir (Quit).**

e. Ahora puede abrir el archivo vectorial resultante en QGIS
   i. Abra una nueva sesión QGIS.
   ii. Añada conjunto de datos raster "EL_Paso_L8_7B" a su sesión de QGIS (**Add Raster Layer** Icono en la barra izquierda).

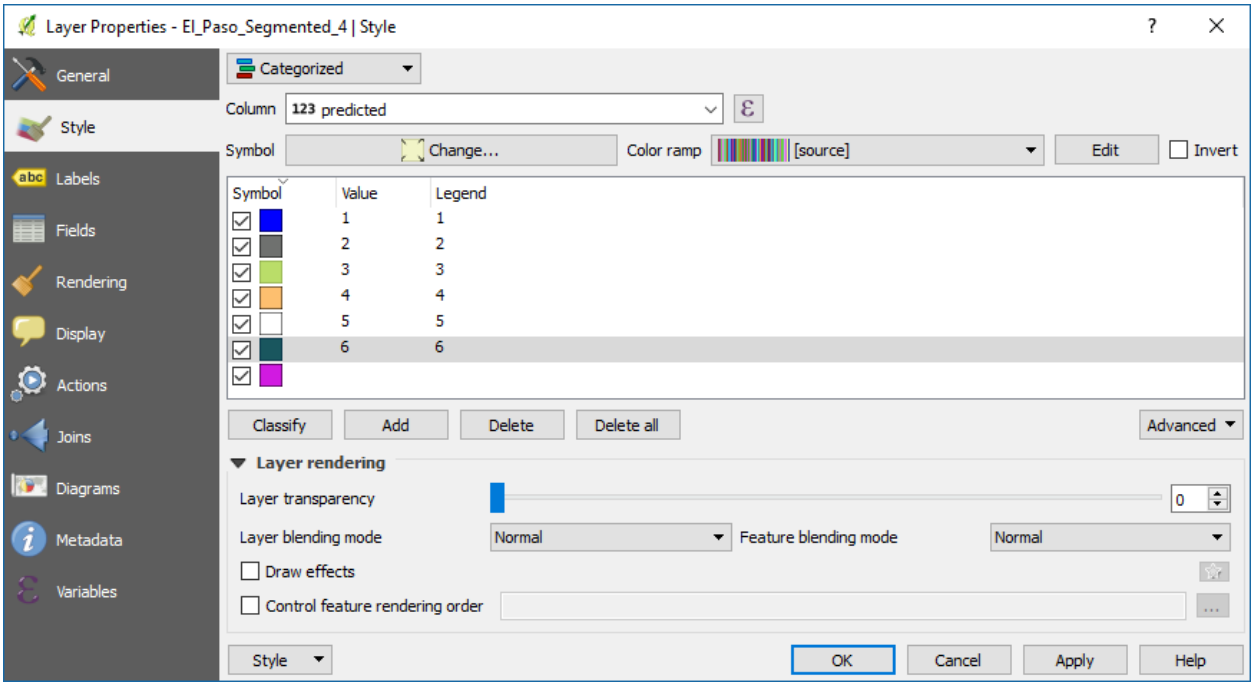
iii. Add vector dataset “El_Paso_Segmented_4.shp” to your QGIS Session (**Add Vector Layer** Icon in the left bar).

iv. Haga clic en “El_Paso_Segmented_4” y seleccione **Propiedades (Properties).** Haga clic en el Menú de Estilo (Style menu) en la parte superior izquierda. Cerca de la parte superior, cambie de "Símbolo individual" (Single symbol) a "Categorizado" (Categorized). Para Columna (Column), desplácese hacia abajo para seleccionar ”predicted”.

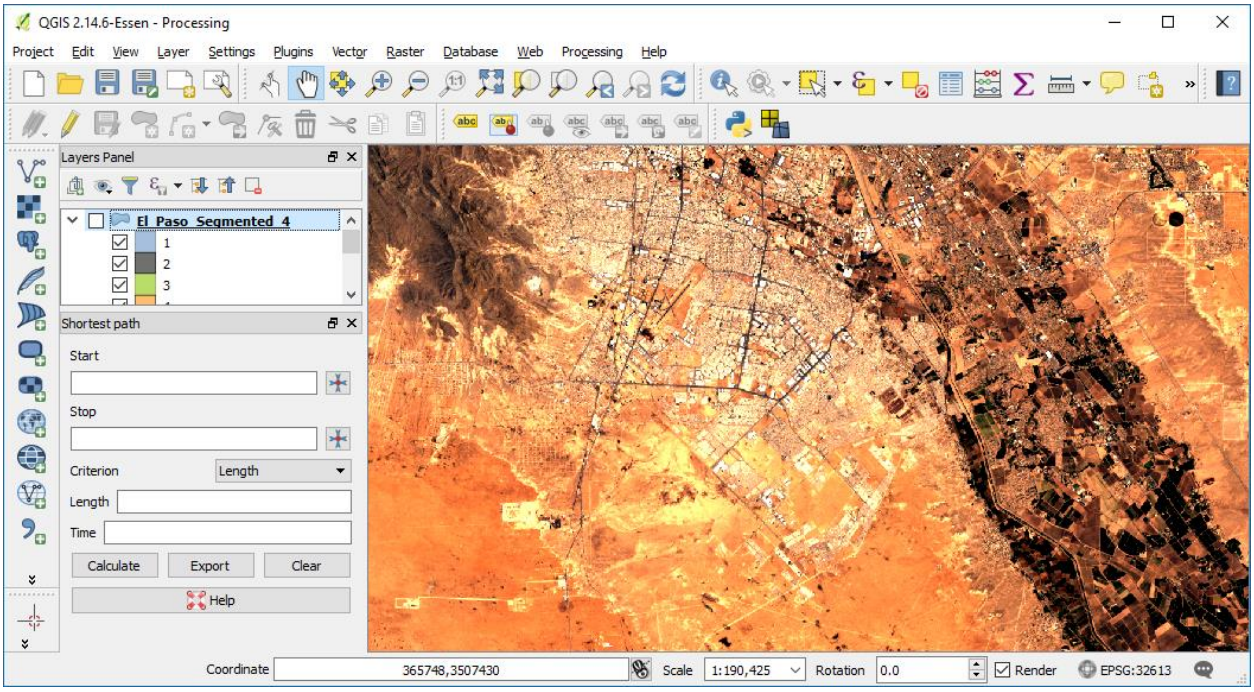
v. Haga clic en **Clasificar (Classify),** los valores por debajo del símbolo del 1-6 de las clases definidas en nuestros polígonos de formación (“El_Paso_Training_poly.shp”). Las identificaciones de las clases son las siguientes: 1 para el agua, 2 para el desarrollo, 3 para el césped y la agricultura, 4 para el pincel, 5 para el terreno estéril y 6 para el bosque. Hay un séptimo símbolo con otros polígonos (sin clasificación). En la tabla, se puede agregar un nuevo campo para nombrar explícitamente las clases informativas.

vi. Mueva el control deslizante Transparencia de capa a la derecha y haga clic en Aplicar con prueba y error hasta que sea lo suficientemente transparente para ver las imágenes debajo, pero no demasiado transparentes para indicar qué color y, por lo tanto, la clase de cobertura de suelo que es.

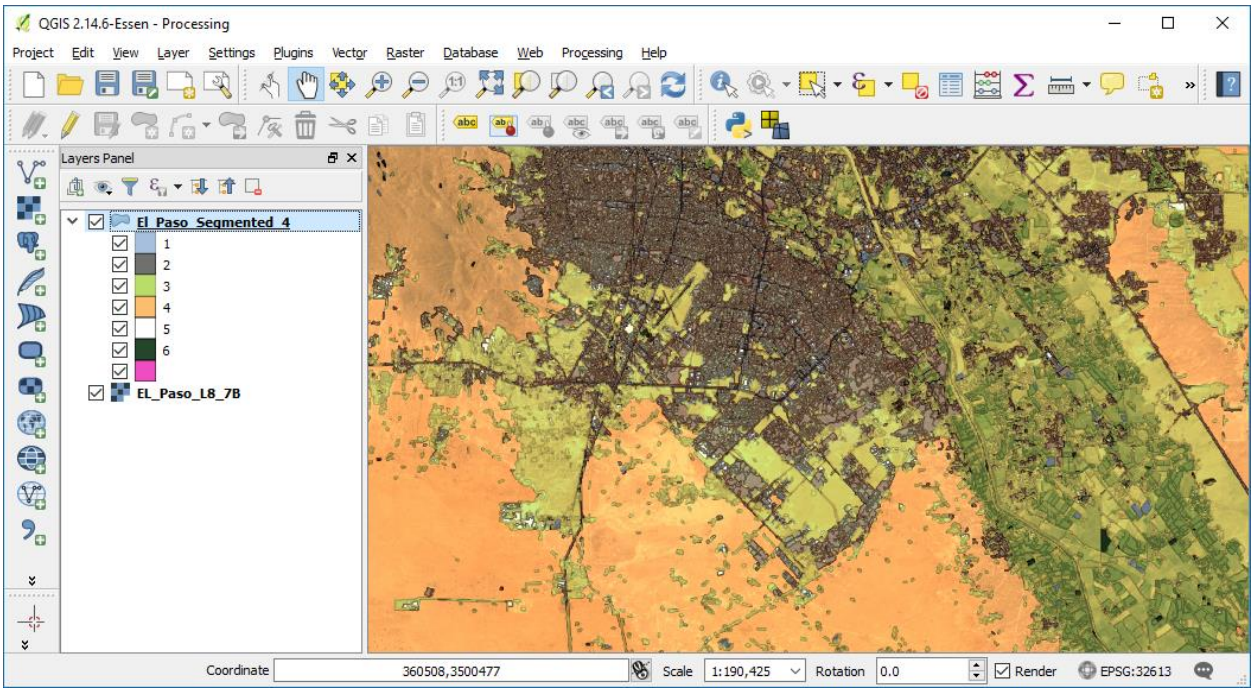
vii. Haga clic en Aceptar (OK).



Ahora podemos echar un vistazo a las imágenes.



Y podemos ver la clasificación encima de las imágenes.



3.4.4 Evaluación de la precisión

Se puede ver que nuestro primer paso en una Clasificación Orientada a Objetos no es perfecto - algunas parcelas agrícolas fueron clasificadas como bosque. Esto puede ser ayudado con la comprobación de

campo en el terreno y polígonos de entrenamiento mejor definidos. La mayoría de las clases fueron discernibles y ahora, después de este tutorial, usted se sentirá experto en la realización de la clasificación orientada a objetos en QGIS. Una vez que esté visualmente satisfecho con los resultados de clasificación, puede seguir un par de pasos: disolver polígonos contiguos de la misma clase y realizar una evaluación de precisión de su clasificación con el segundo conjunto de polígonos reservados para eso ("El_Paso_accu_A.shp"). Información y documentación acerca de cómo realizar una serie de procesos auxiliares como la forma de digitalizar polígonos, seleccionar una parte, el proyecto de un archivo de formas, etc., se pueden encontrar bajo la etiqueta de documentación de la página web qgis: http://qgis.org/en/site / .

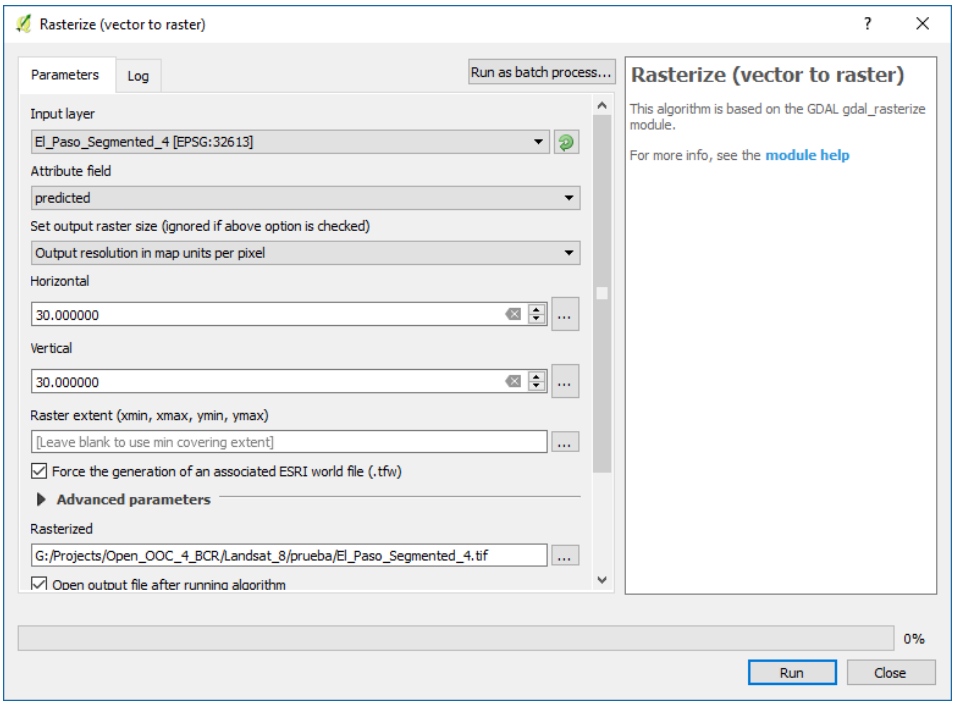
Para reducir el número de polígonos, es posible que desee disolver los polígonos que son contiguos pero no se han combinado antes, esto creará un shapefile más pequeño que será más fácil de compartir.

a. Usted puede encontrar la herramienta disolver en la barra de menú superior: **Vector --> Geoprocessing Tools --> Dissolve.** Desmarque la casilla: "Disolver todo (no use campos)".

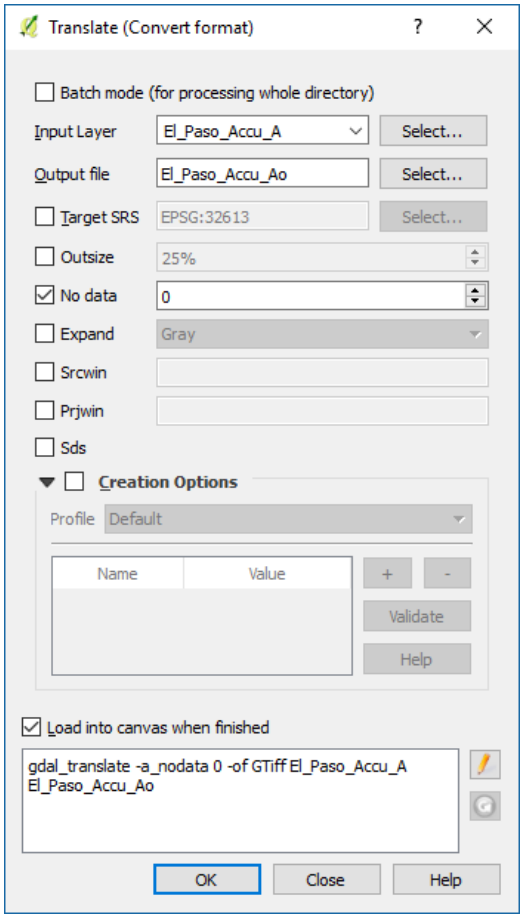
   Seleccione el atributo “predicted” como su campo de disolución y haga clic en > para transferirlo a la casilla Seleccionada (Selected box). Haga clic en los tres puntos de "Disuelto", seleccione "Guardar en archivo", y el nombre de su salida "El_Paso_Segmented_4_dissolved.shp".

Para calcular la matriz de precisión (también llamada de confusión), tanto los polígonos referenciados ("El_Paso_accu_A.shp", utilizando el campo "Id"), como los polígonos resultantes de la clasificación ("El_Paso_Segmented_4_dissolved.shp" o "El_Paso_Segmented_4_dissolved.shp" ") Deben convertirse en imágenes (rasters). La evaluación de precisión utilizando productos completamente vectoriales no está disponible en este momento..

b. En la parte superior la barra de búsqueda del menu encuentre: **Raster --> Conversion --> Rasterize (Vector to Raster)**
c. Rasterize: “El_Paso_accu_A.shp”, usando el campo "Id" para crear: "El_Paso_accu_A.tif", seleccione tamaños de píxeles horizontales y verticales sobre el tamaño de los píxeles landsat utilizados para la entrada de imagen, unos 30 m. Haga clic en la marca de verificación para forzar la generación de un archivo ESRI mundial asociado (.tfw) y haga clic en "ejecutar" (run).



d. Partiendo de que "El_Paso_accu_A.tif" tiene ceros como rellenos, y no tenemos una clase cero, necesitamos convertir los ceros a "No_data". En la parte superior barra de menú encuentre: **Raster --> Conversion --> Translate (Convert format).** Seleccione "El_Paso_Accu_A.tif" como su capa de entrada, el nombre de su capa de salida como "El_Paso_Accu_Ao", seleccione la marca de "Sin datos" y asignarle un "0" como valor para No-datos y haga clic en Aceptar.



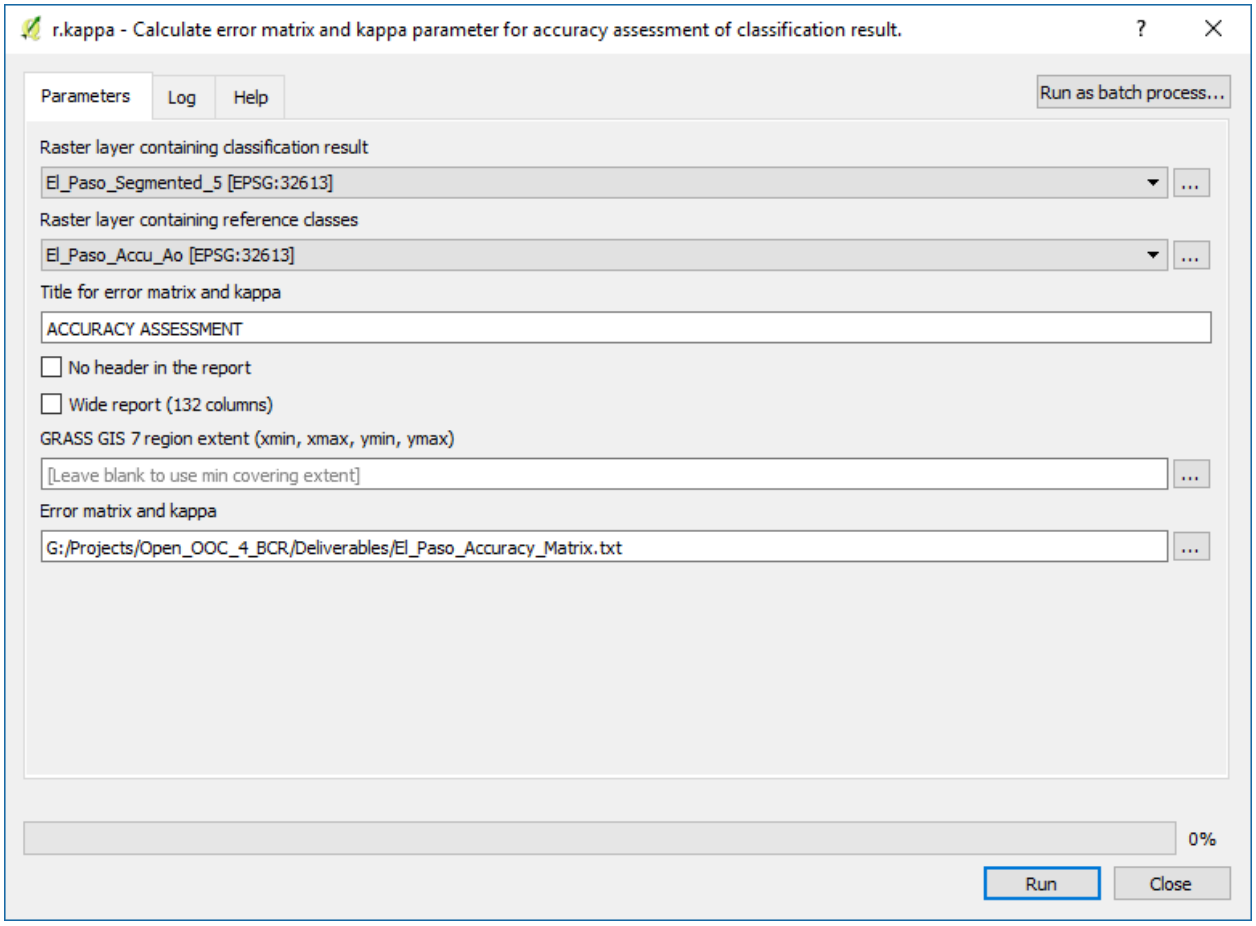
e. Rasterize: "El_Paso_Segmented_4.shp" utilizando el campo "predicho" en "El_Paso_Segmented_5.tif", seleccione tamaños de píxeles horizontales y verticales sobre el tamaño de los píxeles Landsat utilizados para la entrada de imagen, unos 30 m. Haga clic en la marca de verificación para forzar la generación de un archivo ESRI mundial asociado (.tfw) y haga clic en "ejecutar". Este paso puede tomar un tiempo más largo, ya que hay más procesamiento.

f. Encuentre: **Processing Toolbox --> GRASS GIS & commands --> Raster -->r.kappa**

g. Seleccione la capa que contiene el resultado de la clasificación "El_Paso_Segmented_4.tif" y la capa ráster de los polígonos de evaluación de exactitud "El_Paso_accu_A.tif".

h. Dar un nombre a la matriz de error, como "**El_Paso_Error_Matrix.txt"**

i. A continuación, ejecute **r.kappa** . Esto calcula la matriz de error y el parámetro kappa para la evaluación de la exactitud del resultado de la clasificación.



Los valores de Kappa para el total y cada clase se calculan junto con sus varianzas. También se tabulan el porcentaje de errores de comisión y omisión, el total de resultados clasificados correctos por número de píxeles, el área total en píxeles y el porcentaje de píxeles clasificados correctamente. ¿Qué significa todo esto? La estadística Kappa (o valor) es una métrica que compara una precisión observada con una precisión esperada (probabilidad aleatoria). Según Landis y Koch: valores kappa de 0-0.20 pueden considerarse como ligeros, 0.21-0.40 como justos, 0.41-0.60 como moderados, 0.61-0.80 como sustanciales, y 0.81-1 como casi perfectos, mientras que otros autores como Fleiss consideran Kappas valores> 0,75 como excelente, 0,40-0,75 como justo a bueno, y <0,40 como pobre.

El error de clasificación ocurre cuando un objeto (o píxel en el caso de clasificaciones supervisadas o no supervisadas) perteneciente a una clase (es decir, cobertura de la tierra) se asigna a otra clase. Los errores de omisión ocurren cuando un objeto se deja fuera de la clase que se está evaluando; Errores de comisión se producen cuando un objeto se incluye incorrectamente en la clase que se está evaluando. Un error de omisión en una categoría se contará como un error en la comisión en otra clase.

En general, las clasificaciones de precisión de más del 70% se consideran aceptables, pero mucho depende del propósito de la misma, si desea comparar la cobertura de la tierra con el tiempo para ver cómo están funcionando las prácticas de conservación de la tierra, debemos minimizar el error, en este caso un valor Por encima del 85% de precisión en la clasificación se prefiere.

El informe de evaluación se escribirá en un archivo de salida que está en formato de texto sin formato y nombrado por el usuario al momento de ejecutar el programa. Este informe se puede trasladar a Excel o un procesador de textos como Open Office para su formateo si es necesario.

Ejemplo de archivo de texto de la Evaluación de Precisión: El_Paso_Accuracy_Matrix.txt:

```
                 ACCURACY ASSESSMENT

LOCATION: temp_location                    Mon May 08 22:08:05 2017

MASK: none

MAPS: MAP1 = (untitled) (tmp1494299281396 in PERMANENT)

      MAP2 = (untitled) (tmp1494299281395 in PERMANENT)


Error Matrix
Panel #1 of 2
                  MAP1
     cat#  1     2     3     4     5
 M     1   144   147   50    0     0
 A     2   531   2709  0     0     31
 P     3   7     245   6136  5060  44
 2     4   0     0     0     35306 15700
       5   0     2     0     0     244
       6   0     0     175   0     0
Col Sum    682   3103  6361  40366 16019

Panel #2 of 2
                  MAP1
     cat#  6     Row Sum
 M     1   0     341
 A     2   0     3271
 P     3   57    11549
 2     4   476   51482
       5   0     246
       6   16    191
Col Sum    549   67080


Cats  % Commission     % Ommission      Estimated Kappa
1     57.771261  78.885630  0.416353
2     17.181290  12.697390  0.819854
3     46.869859  3.537180   0.482200
4     31.420691  12.535302  0.211013
5     0.813008   98.476809  0.989319
6     91.623037  97.085610  0.076209
```

```
Kappa      Kappa Variance
0.352544   0.000013

Obs Correct     Total Obs  % Observed Correct
     44555      67080      66.420692

MAP1 Category Description
1:    Water
2:    Developed
3:    Grass-Agro
4:    Brushland
5:    Barren land
6:    Forest

MAP2 Category Description
1:    Water
2:    Developed
3:    Grass-Agro
4:    Brushland
5:    Barren land
6:    Forest
```

Como se puede ver con los datos sintéticos utilizados para este tutorial, polígonos de entrenamiento y polígonos de evaluación de exactitud, el valor Kappa calculado en general es .35 y la precisión de clasificación es del 66%. Ambos valores están un poco debajo de lo que nos gustaría como un mínimo. Con los datos verificados en el campo alrededor de la fecha de la colección de imágenes, los valores de kappa y la precisión de clasificación serán mejores que los que se muestran en este ejercicio. Esperemos que esta experiencia nos permita señalar la importancia de la comprobación de campo verificado in situ.

Si los resultados de su clasificación no son lo que desea, puede omitir polígonos de entrenamiento y polígonos de precisión con certidumbre baja o con clases mixtas. Medidas adicionales pueden ser ajustar parámetros como el radio espacial (Spatial Radious) (píxeles) y radio de alcance (range radious) (radiométrico) a valores más pequeños para obtener segmentos (objetos) más pequeños, minimizando la confusión y facilitando su clasificación.

Para obtener ayuda adicional o información sobre clasificaciones, OTB o QGIS, por favor, mantenga la ayuda en línea como referencia:

- Https://www.orfeo-toolbox.org/documentation/
- Https://www.qgis.org/en/site/forusers/index.html

Ayuda adicional puede ser proporcionada contactando a Miguel Pavon en pavonma@hotmail.com o (512) 466-3936.