

## Taller de clasificación Supervisada de imágenes satelitales en R

### Resumen metodológico

El presente documento, resume e interpreta brevemente los pasos metodológicos llevados a cabo en R por el código: “Script\_Clasificación\_Supervisada.R”, y los materiales anexos incluidos con este.

En la carpeta entregada para este taller (“Workshop\_Clasificaciones”), podrá encontrar los siguientes archivos que servirán para la realización del taller:

#### Materiales:

- Carpeta “raster” con una escena satelital de **reflectividad de superficie Landsat 8** y el **índice de vegetación NDVI** en formato GeoTiff (~raster/2016/L8\_ref-ndvi\_2016.tif). Corresponde a un archivo multibanda (varias imágenes apiladas en una). En orden posee las siguientes bandas:
  - Azul (B)
  - Verde (G)
  - Rojo (R)
  - Infrarrojo cercano (NIR)
  - Infrarrojo de onda corta 1 (SWIR1)
  - Infrarrojo de onda corta 2 (SWIR2)
  - Índice de vegetación de diferencia normalizada (NDVI)
- **Puntos de entrenamiento-validación** creados previamente (mediante Google Earth) para **cinco clases de cobertura de suelo** conocidas y presentes en la escena, y en formato KML. Cada set de puntos representa una de las siguientes clases
  - **Urbano** (n=165) → Cubiertas creadas/construidas por el humano.
  - **Vegetación manejada** (n=77) → Cultivos agrícolas, plantaciones y parques urbanos.
  - **Matorral y suelo desnudo** (n=120) → Matorrales esclerófilos/espinosos y suelos desnudos
  - **Bosque** (n=43) → Bosque esclerófilo asociado principalmente a quebradas.
  - **Cuerpos y cursos de agua** (n=54) → Lagunas, tranques y ríos presentes.

#### Pasos metodológicos:

- Carga de **paquetes de R + definición de directorios** de entrada y salida dentro de cada ordenador.
- Carga de dos **funciones de R**, que permiten visualizar firmas espectrales en pasos posteriores (sacspect)
- Lectura de **archivo ráster multibanda** (función “stack”). Deja disponibles las imágenes satelitales en un archivo multibanda propio de R (*RasterStack*). Cada una de las bandas de este archivo

- Lectura de **puntos de entrenamiento-validación** (función “readOGR”). Deja disponibles estos archivos vectoriales de puntos en el formato propio de R (*SpatialPointsDataFrame*).
- **Exploración visual de la información espacial y satelital:**
  - **Apertura de archivos KML** utilizando software Google Earth Pro (fuera de R): Cada archivo KML fue creado previamente utilizando las herramientas de Google Earth, visualizando la imagen de alta resolución para el año en cuestión, agregando marcas de posición dependiendo de cada una de las cinco clases consideradas, agrupando estas marcas en una misma carpeta, y exportando esta última a formato KML.
  - **Visualización interactiva de firmas espectrales** de un píxel determinado (“sacspect”). Esto permite verificar la variabilidad de las firmas espectrales presentes en la escena.
- **Generación de insumos previos para clasificación supervisada** (paquete “rasclass”). Este paquete requiere conectarse a una carpeta que posea a cada una de las variables espaciales (raster) a considerar en la clasificación supervisada, así como los puntos de entrenamiento-validación del algoritmo clasificador (raster):
  - **Exportación de cada banda por separado** (x6 bandas de reflectividad de superficie + índice de vegetación NDVI), en formato de texto ASCII (.asc).
  - **Transformación a ráster** (con los mismos parámetros de las bandas de reflectividad) de cada *SpatialPointsDataFrame* representativo de cada una de las cinco clases a clasificar (cinco ráster con píxeles con un solo valor donde interceptan los puntos, y el resto con valores “no dato”)
  - **Sumatoria de los cinco ráster** de los puntos, para formar uno que incluya a las cinco clases.
  - Exportación del **ráster resumen de los puntos de entrenamiento-validación** a formato ASCII dentro de la misma carpeta anterior.
- **Creación del objeto “rasclass” para almacenar clasificación supervisada** mediante la lectura de la carpeta con los archivos en formato ASCII (“readRasterFolder”).
- **Evaluación de separabilidad entre las clases** utilizando la **Medida de separabilidad de divergencia transformada** (toma valores entre 0 y 2). Esto permite establecer si las clases representadas por los distintos puntos de entrenamiento son o no grupos homogéneos de información que se separan del resto de las clases en cuestión.
  - Si es que presentan una media-baja separabilidad ( $\sim < 1.7$ ), debiera editarse la capa de puntos de entrenamiento en Google Earth, para poder acercarse lo más posible al valor 2 de separabilidad.
- **Ejecución de la clasificación supervisada utilizando SVM** (“classifyRasclass”), y considerando una partición para validación del 30% (*splitfraction*=0.3). La clasificación realizada (mapa) así como sus medidas de precisión, se almacenan dentro del mismo objeto “rasclass”.
- **Exportación de estadísticas de acuracidad de la clasificación y mapa final** para visualizarlas y almacenarlas fuera de la sesión de R, y considerando distintos formatos para su manejo. Estas salidas incluyen:
  - **Archivo formato CSV** con la **Precisión Global (overall accuracy)**, **índice Kappa** y la **matriz de confusión**.

- Mapa con colores representativos de la clasificación final en formato PNG.
- Archivo ráster en formato GeoTiff de la clasificación final.
- **Postclasificación, Estadísticas y gráficos a partir de las clasificaciones efectuadas**, lo que incluye:
  - **Cálculo de superficies por cada clase de cobertura de suelo.**
  - Exportación de **tabla resumen de estas superficies** calculadas por cada clase de cobertura de suelo (CSV).
  - Generación de un **gráfico de torta resumen** a partir de estas superficies calculadas por cada clase de cobertura de suelo (PNG).