

EL INDICE NDVI Y LA CLASIFICACIÓN DE ÁREAS SEMBRADAS APRENDIZAJE AUTOMÁTICO NO SUPERVISADO “K-MEANS”¹

Girimonte¹, Patricia y García Fronti², Javier

¹ Universidad de Buenos Aires, Facultad de Ciencias Económicas, Av. Córdoba 2122 – 1120AAQ, Ciudad Autónoma de Buenos Aires, República Argentina

² Universidad de Buenos Aires, Facultad de Ciencias Económicas, CIMBAGE IADCOM Argentina

Resumen

Recibido: 04-07-2020

Aceptado: 15-08-2020

Palabras clave

NDVI, Teledetección,
Gestión agropecuaria.

La teledetección satelital es utilizada en distintas partes del mundo como una estrategia para la gestión agropecuaria. La información de los satélites permite calcular diversos índices para la gestión, uno de los más utilizados es el de vegetación de diferencia normalizada (NDVI), el cual está asociado a la fracción de la radiación solar absorbida por las plantas.

El objetivo de este trabajo es describir un procedimiento que permite clasificar áreas sembradas a partir de imágenes satelitales utilizando el índice NDVI. Para realizarlo se clasificarán áreas entrenando un algoritmo de aprendizaje automático de clasificación no supervisado “K-means”, el cual se implementa en lenguaje R.

Copyright: Facultad de Ciencias Económicas, Universidad de Buenos Aires.

ISSN (En línea) 2362 3225

¹ Los autores agradecen los comentarios recibidos de Pablo M. Herrera, Verónica García Fronti y de los participantes del 1er Seminario de Riesgo Agropecuario 2020.

EL INDICE NDVI Y LA CLASIFICACIÓN DE ÁREAS SEMBRADAS APRENDIZAJE AUTOMÁTICO NO SUPERVISADO “K-MEANS”

Abstract

KEYWORDS

NDVI, Remote sensing, agricultural management

Satellite remote sensing is used in different parts of the world as a strategy for agricultural management. The information from the satellites makes it possible to calculate various indices, one of the most used is the normalized difference vegetation (NDVI), which is associated with the fraction of solar radiation absorbed by plants.

The objective of this work is to describe a procedure that allows to classify planted areas based on NDVI. To do this, areas will be classified by training an unsupervised classification machine learning algorithm "K-means", which is implemented in R language.

Copyright: Facultad de Ciencias Económicas, Universidad de Buenos Aires.

ISSN (En línea) 2362 3225

INTRODUCCIÓN

La incorporación de la información que brinda la teledetección satelital en las ciencias agronómicas tiene su origen en la década del 70, donde también se definieron los primeros índices relacionados con la productividad. Uno de los más utilizados es el Índice de Vegetación de Diferencia Normalizada (NDVI) también llamado Índice Verde (Rouse, et al, 1973).

La política de datos abiertos iniciadas en el año 2013 por la NASA y en el año 2017 por la ESA, ha posibilitado a los distintos países poner a disposición de sus ciudadanos información georreferenciada relacionada con distintos indicadores económicos. Las imágenes satelitales más usadas por los organismos gubernamentales son las del Landsat 8 de la NASA, Sentinel 2 (Satélite) de la ESA, y complementariamente el SPOT (Satélite Para la Observación de la Tierra) desarrollado por el CNES (Centro Nacional de Estudios Espaciales francés) en colaboración con Bélgica y Suecia.

Dentro del ámbito académico la disponibilidad de esta información permite plantear diferentes líneas de investigación, como las relacionadas con la productividad en las actividades agropecuarias, la tributación agropecuaria en base a la productividad, la sustentabilidad del suelo y la contaminación ambiental entre otras. También en el ámbito privado existen diferentes propuestas que le permiten al productor gestionar su campo de manera remota a través de imágenes satelitales provistas por la NASA, ESA, CONAE y en algunas iniciativas más costosas por drones. Algunas de estas propuestas digitalizan los diferentes puntos del campo, identifican y realizan un mapeo de la variabilidad del terreno, para así poder elegir la mejor estrategia para maximizar el rinde, reducir los costos y el impacto ambiental.

El objetivo de este trabajo es presentar un modelo de aprendizaje automático no supervisado que clasifique áreas sembradas a partir de imágenes satelitales utilizando el índice NDVI, en una zona con siembra de soja cercana a Adela María en el departamento de Río Cuarto. Para lograr lo propuesto este trabajo se divide en dos secciones. La primera presenta la definición de NDVI utilizada y explica cómo se ha realizado el cálculo para la región elegida. En la segunda se implementa el modelo propuesto y se clasifican las áreas.

1. CÁLCULO DEL NDVI

La utilización de datos espectrales para evaluar parámetros de vegetación se basa en la reflectancia diferencial de los tejidos fotosintéticos dentro de la porción rojo e infrarrojo del espectro electromagnético (Rouse et .al., 1973). La fenología de la vegetación se define como los ciclos recurrentes de las actividades biológicas estacionales y su relación con condiciones ambientales y meteorológicas como temperatura, luz, humedad y tipo de suelo. Los distintos tipos de vegetación responden de manera diferente a dichas condiciones. Debido a la absorción de la clorofila, las hojas verdes reflejan muy poca luz correspondiente al rojo, mientras que muestran una alta reflectancia en la zona del infrarrojo cercano.

El NDVI (Normalized Difference Vegetation Index) es un índice que deriva del cociente entre la reflectancia del rojo y el infrarrojo cercano: $(I_r - R) / (I_r + R)$ (1) donde I_r es la reflectancia correspondiente al infrarrojo cercano y R la reflectancia correspondiente al rojo del espectro electromagnético (Rouse et .al., 1973). Los valores del índice oscilan entre -1 y 1. Los valores negativos están relacionados con cuerpos de agua y superficies degradadas por acción del fuego, mientras que valores positivos más bajos (cerca de 0) corresponden a vegetación senescente o de baja cobertura. Los valores positivos altos (cerca de +1) representan alto contenido de biomasa fotosintética. Se han encontrado fuertes relaciones entre el NDVI y algunas características funcionales y estructurales de la vegetación como biomasa, índice de área foliar, cobertura y productividad primaria neta, razón por la que es uno de los índices más utilizados como indicador de la productividad. (Deering 1978).

1.1. Determinación de la zona y período de análisis

La soja es uno de los principales cultivos de Argentina no sólo por la producción, sino por la superficie ocupada. Nuestro país es el tercer productor de soja después de Estados Unidos y Brasil, siendo las provincias de Buenos Aires, Córdoba y Santa Fe las de mayor producción interna. La producción de soja en la provincia de Córdoba en la campaña 2018/19 alcanzó las 14.970.100 toneladas, representando el 27% del total del país, que de acuerdo con el Ministerio de Agricultura, Ganadería y Pesca se produjeron más de 55 millones de toneladas. El rendimiento promedio ponderado fue de 37,6 qq/ha, siendo de 5 quintales superior a la media nacional (Bolsa de Cereales de Córdoba, 2019). La provincia de Córdoba provee información detallada de la cobertura de su suelo (IDECOR, 2019) con mapas de coberturas obtenidos mediante trabajo en campo y teledetección satelital.

A partir de esta información de la cobertura de suelo y la disponible en la Bolsa de Cereales, correspondientes a la provincia de Córdoba, se eligió una zona de análisis con siembra de soja cercana a Adela María en el departamento de Río Cuarto. (Bolsa de Cereales de Córdoba, 2019). Para determinar el período se tuvo en cuenta que en Argentina el cultivo de la soja en secano

comienza entre los meses de septiembre y diciembre (siembra tardía, generalmente para la soja de segunda, luego del trigo), siendo entonces el período febrero-marzo el de mayor vigorosidad de la planta.

1.2. Cálculo del NDVI por píxel

Existen distintos satélites que permiten obtener imágenes terrestres, entre los más utilizados en la actualidad para los análisis de cobertura digital se encuentra el Landsat 8, con dos sensores, uno de ellos, Oli (Operational Land Imager) permite obtener mediciones en el infrarrojo cercano y el TIERS (Thermal Infrared Sensor).

Dentro de las opciones disponibles están las imágenes “LandsatLook”, estos son archivos optimizados para la interpretación visual y no se recomienda su utilización para procesamiento; mientras que “Level 1 GeoTiff data Product” contiene imágenes georreferenciadas y corregidas en el terreno, éstas son las adecuadas para un posterior procesamiento y análisis (CONAE, 2018). Antes de descargar la imagen, que comprimida en formato “tar” tiene un tamaño aproximado de 1.65 GB, puede visualizarse para corroborar el porcentaje de nubes y que corresponda a la región buscada,

Una vez descargada la imagen al descomprimirla se obtienen 11 archivos, 9 de formato TIFF (Tagged Image File Format), y 2 correspondientes a los metadatos que contiene información de fecha, hora, latitud y longitud de la zona. El nombre del archivo comprimido y luego la primera parte del nombre de cada una de las bandas corresponde a la plataforma, el sensor, path-row, y el año y día juliano, de adquisición. Descomprimidas, cada una de las bandas tienen un tamaño aproximado de 113 MB salvo la banda 8 que tiene un tamaño de 453 MB. El formato Tiff permite el procesamiento de las imágenes satelitales en diferentes softwares específicos y otros no específicos como Python y R.

Una vez obtenidas las 11 bandas de la imagen se realizó un primer procesamiento en SOPI (Software de procesamiento de imágenes) desarrollado por la CONAE. Este primer análisis descriptivo permitió observar la región de interés, y decidir el apilado de bandas adecuado para visualizar la imagen de acuerdo con nuestro interés de clasificar la cobertura del suelo.

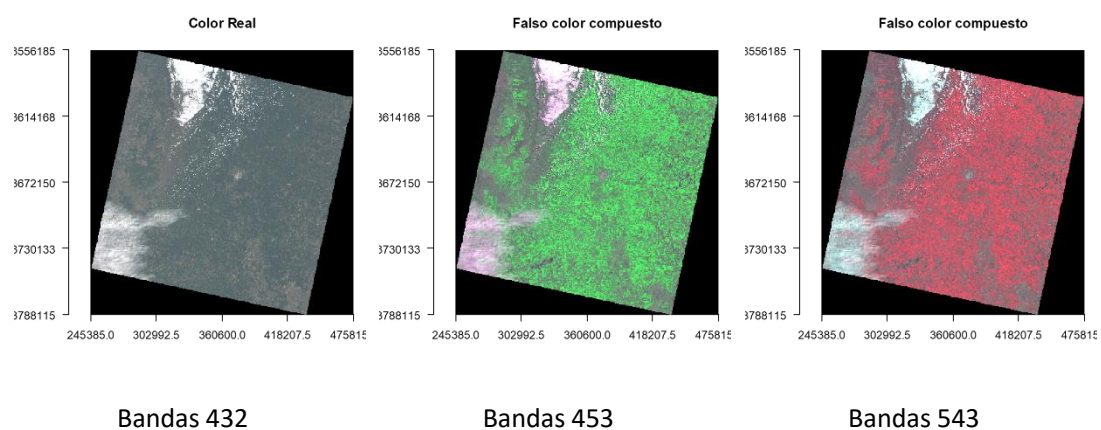
La interpretación visual es una parte fundamental de la teledetección, dado que nos permite observar características biofísicas y fenológicas de los diversos usos y coberturas del suelo. La reflectancia como se mencionó anteriormente es la energía que refleja una cobertura del total de energía que recibe del sol, toma valores entre 0 y 1, pero suele expresarse también como porcentaje, por ejemplo, una reflectancia del 20% para un cultivo, significa que refleja el 20% de la energía solar incidente. La reflectancia espectral es la medida para una longitud de onda

determinada, nos permite definir la respuesta espectral de una cobertura. En una imagen, cuanto más alto (e intenso) es el valor del píxel de una determinada cobertura, mayor reflectancia presentará en determinada zona del espectro (CONAE, 2018).

A partir de la interpretación de las firmas espectrales y de la disponibilidad de bandas de las misiones satelitales, es posible observar la respuesta espectral de los usos y coberturas y relacionarlas con variables biofísicas. Existen diferentes combinaciones de bandas y su utilización dependerá del fenómeno en estudio. Por ejemplo, para visualizar vegetación se utiliza una combinación de bandas que incluyan la del rojo e infrarrojo cercano.

El posterior procesamiento de las imágenes se realizó en R 3.6.1 con las librerías raster (Hijmans, 2019), dplyr (Wickham et.al. ,2019), sp (Pebesma, E.J., R.S. Bivand, 2005) y rgdal (Bivand, et.al, 2019), con un script realizado para tal fin (Hijmans R., 2019) (Ver Apéndice). Para visualizar las imágenes se realizaron 3 apilados de bandas distintos (**Figura 1**).

Figura 1: Apilados de bandas



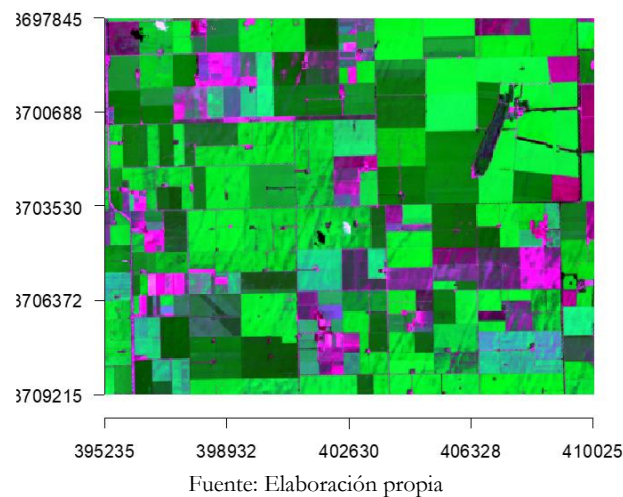
Fuente: Elaboración Propia

Como nuestro objeto es clasificar la cobertura del suelo a partir de NDVI, se continuó el análisis con el apilado de las bandas 4 (infrarrojo cercano),5 (infrarrojo de onda corta) y 3 (rojo), dado que permite tener una apreciación visual de las áreas cultivadas.

Se definió un polígono aproximadamente 18 km al norte de Adela María en el departamento de Río Cuarto, determinando su posición de acuerdo con la longitud y latitud en Google Maps. Para poder extraer de la imagen satelital este polígono se lo definió como un “objeto espacial” con las mismas coordenadas geográficas que la imagen obtenida del Landsat 8 (Apéndice, Figuras 1 y 2).

La **Figura 2** corresponde a la región seleccionada. La imagen pixelada obtenida permite observar distintas coberturas del suelo. Las zonas con verde más claro deberían corresponderse con las zonas donde los cultivos presentan mayor vigorosidad.

Figura 2: Región de interés



Luego para la región seleccionada se obtuvo para cada píxel el valor de NDVI operando entre las bandas 5 y 4 de acuerdo con la fórmula (1).

2. CLASIFICACIÓN Y ANÁLISIS DE LA COBERTURA DEL SUELO A PARTIR DEL NDVI

Clasificar a partir del índice NDVI y analizar los valores obtenidos es de importancia para evaluar la cobertura del suelo y detectar, si las hubiera, zonas donde el crecimiento del cultivo podría presentar problemas, dado que se obtendría un valor de NDVI inferior al esperado.

Para el área seleccionada, de acuerdo con la resolución del Landsat 8 el número de píxeles es de 186847; como nos interesa un mapeo del suelo, a partir del NDVI, se trabajó con el total de píxeles.

La **Tabla 1**, muestra las medidas resumen para el NDVI calculado.

Tabla 1: Medidas descriptivas de NDVI

Mínimo	Máximo	Media	Mediana	Rango	Desvío	MAD
0.096	0.650	0.624	0.539	0.554	0.102	0.093

Las Figuras 3 y 4 corresponden, respectivamente, al histograma y boxplot del NDVI. A partir del análisis descriptivo realizado puede observarse una distribución asimétrica a izquierda, con presencia de valores atípicos.

Figura 3

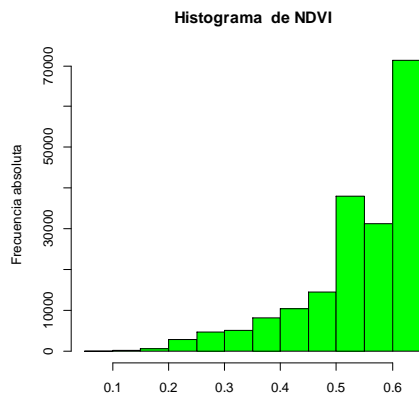
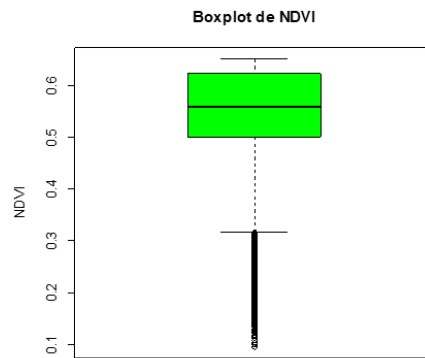


Figura 4



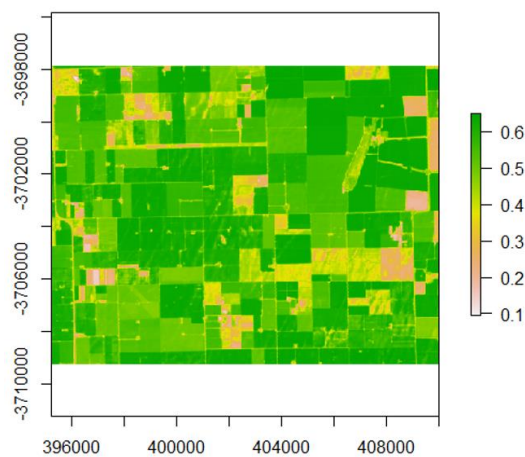
Los valores de NDVI se encuentran en un rango entre 0.1 y 0.65, siendo en su mayoría mayores a 0.5 indicando un nivel alto de vigorosidad, esto es de esperar dado que corresponde a una zona con cultivo principalmente de soja, y al período elegido de mayor vigorosidad de la planta.

De acuerdo con la clasificación de NDVI los valores más bajos corresponderían a zonas con vegetación senescente o de baja cobertura (Tabla 1, Anexo).

La **Figura 5** corresponde al gráfico por default de R, que realiza un mapeo de la zona a partir del valor de NDVI de cada píxel.

Figura 5

Región-NDVI-Landsat8



Fuente: Elaboración Propia

Para realizar este mapeo se combina la información de la capa raster con los valores calculados de NDVI, con la capa vectorial con los datos georreferenciados del área de estudio. Las coordenadas de la capa vectorial son las que permiten identificar la región.

En este gráfico puede observarse, como ya se mencionó, que el NDVI toma valores entre 0.10 y 0.65. El color preponderante en la región es el verde en distintas tonalidades, cuanto más intenso, es mayor el valor del índice.

2.1. Clasificación del NDVI mediante el algoritmo de K- means

Como hemos mencionado, diversas investigaciones relacionadas con la productividad de los cultivos se basan principalmente en el NDVI como un indicador de la productividad. El aprendizaje automático es el subcampo de las ciencias de la computación y una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas que permitan que las computadoras “aprendan”. En muchas ocasiones el campo de actuación del aprendizaje automático se solapa con el de la estadística inferencial, ya que las dos disciplinas se basan en el análisis de datos. Sin embargo, el aprendizaje automático incorpora las preocupaciones de la complejidad computacional de los problemas. De acuerdo con el abordaje de los problemas se lo clasifica en aprendizaje supervisado y no supervisado (Hastie et. al, 2001). En el aprendizaje supervisado el objetivo es predecir el valor del output en función de distintas variables medidas; mientras que en el aprendizaje no supervisado no existe un output y el objetivo es describir las asociaciones y patrones entre un conjunto de medidas de entrada. Dentro de los métodos de aprendizaje no supervisado que permiten resolver problemas de clasificación y que se aplicará en este trabajo, se encuentra el de k-means (Hastie et. al, 2001).

El análisis de *cluster* se usa para encontrar estructuras en los datos sin proveer una explicación ni interpretación, solo “descubre” estructura en los datos, es luego el analista quien debe interpretar los posibles motivos de esa estructura.

Para poder realizar esta segmentación o agrupamiento es necesario establecer una medida de semejanza entre dos objetos (o de desemejanza). El algoritmo de las K-means considera como medida de semejanza la distancia euclídea y es usado cuando todas las variables son del tipo cuantitativo.

En este trabajo para cada píxel se obtuvo el valor de NDVI, por lo cual los datos son unidimensionales.

2.2. Implementación del algoritmo de K-means en R

Para clasificar la zona seleccionada mediante el algoritmo de K-means se trabajó con el total de los píxeles y se utilizó la función `kmeans` de R (R Core Team, 2019) considerando un máximo de 500 interacciones y el algoritmo de Lloyd (Lloyd, 1982). Esta función informa como una “medida de la calidad” de la clasificación obtenida el valor porcentual del cociente entre la suma de cuadrados entre los *cluster* (SSBC) y la suma de cuadrados total (SSTC), este valor podría considerarse con el porcentaje de variabilidad explicada por la clasificación.

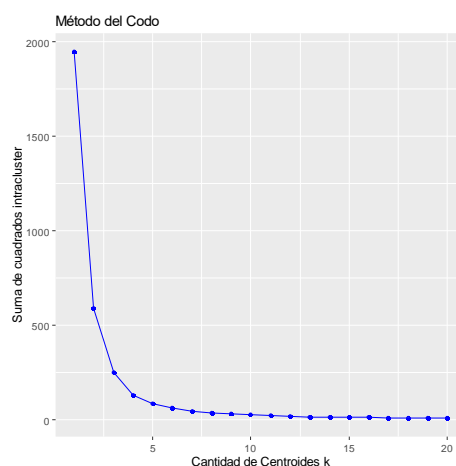
El número de *clusters* se determinó con el método heurístico del “codo”, para esto se utilizó la función `kmeans` considerando de 1 a 20 clusters. Para cada una de las clasificaciones se consideró la suma de cuadrados “*intra-clusters*” (SSWC) y se determinó gráficamente cuando comienza a mantenerse constante.

De acuerdo con el resultado se analizaron las clasificaciones obtenidas mediante el método de K-means considerando de 5 a 20 *clusters*. Para cada una de estas clasificaciones se realizó además del mapeo de la región clasificada, el gráfico de los boxplots en paralelo del NDVI en cada uno de los *clusters* obtenidos. Asimismo, se complementaron los resultados con un análisis descriptivo de NDVI dentro de cada *cluster*.

2.3. Resultados obtenidos a partir del procesamiento y análisis de los datos.

Con los resultados obtenidos mediante el método de K-means considerando de 1 a 15 *clusters*, se construyó la **Figura 6**, que muestra la suma de cuadrados intra-cluster, en función del número de clusters. Puede observarse que a partir de 5 *clusters* (donde se forma el “codo”) no hay una alta variación en los resultados de la suma de cuadrados intra-cluster

Figura 6: Intra-cluster



Fuente: Elaboración propia

A partir de los resultados obtenidos y teniendo en cuenta la clasificación de NDVI (Anexo, Tabla 1), se compararon los resultados obtenidos considerando 5, 7, y 9 clusters. Los gráficos de las figuras 6, 7 y 8 muestran los resultados obtenidos correspondientes al mapeo de la región considerada y los boxplots en paralelo.

En los boxplots de la **Figura 7** se observan 3 *clusters* con valores de NDVI mayor a 0.5, un cluster con valores menores que 0.3, y dos cluster con valores entre 0.3 y 0.5, que de acuerdo a la clasificación de NDVI, sería conveniente separar. Para esta clasificación el valor porcentual del cociente entre la suma de cuadrados entre los *cluster* (SSBC) y la suma de cuadrados total (SSTC) es 95.80%.

Puede observarse en los boxplots de la **Figura 8**, 3 *clusters* con valores de NDVI mayores a 0.5, un cluster con valores menores a 0.3, un cluster con valores entre 0.3 y 0.4 y un cluster con algunos

valores menores que 0.4 y otros mayores que, de acuerdo a la clasificación de NDVI, sería conveniente separar. Para esta clasificación el valor porcentual del cociente entre la suma de cuadrados entre los *cluster* (SSBC) y la suma de cuadrados total (SSTC) es 97.68%.

Figura 7: Clasificación con 5 clusters

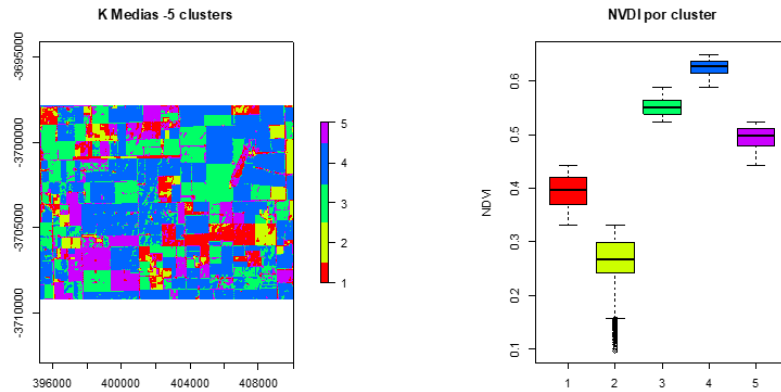
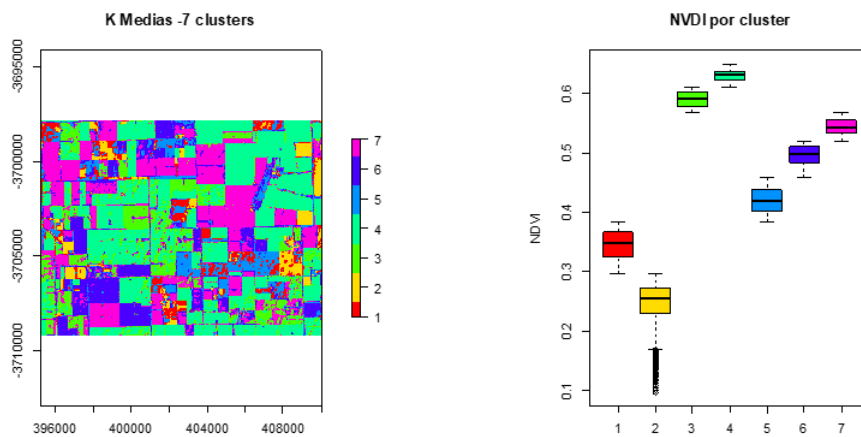


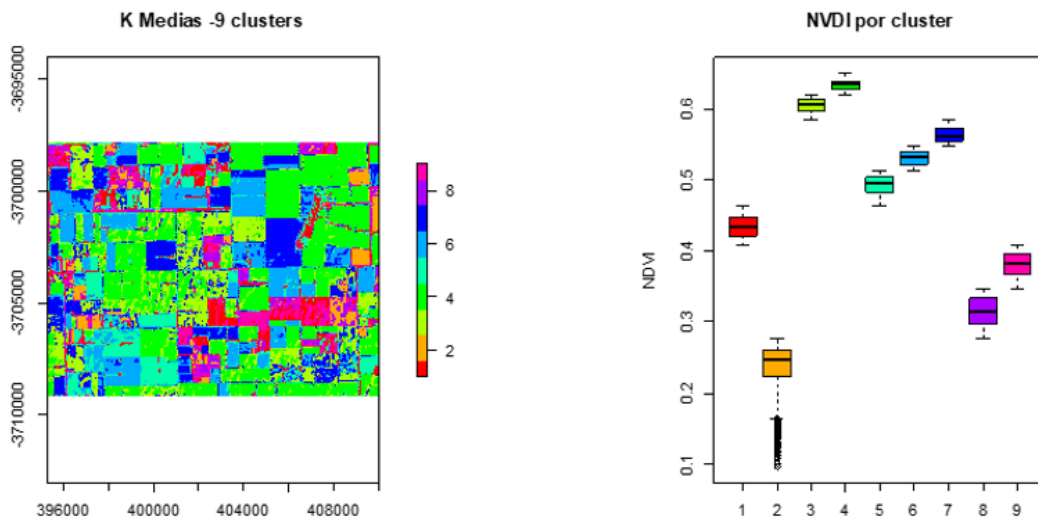
Figura 8: Clasificación con 7 clusters



Puede observarse en los boxplots del Gráfico 8 que no hay ningún cluster con valores menores y mayores de 0.4. Para esta clasificación el valor porcentual del cociente entre la suma de cuadrados entre los *cluster* (SSBC) y la suma de cuadrados total (SSTC) es 98.53%.

En las tres clasificaciones realizadas no se logra separar un *cluster* con valores menores y mayores a 0.2. Para lograrlo hay que considerar 20 *clusters*, pero se pierde la separación lograda con 9 clusters con punto de corte 0.4. Esta dificultad se debe a la presencia de valores atípicos en donde hay algunos valores muy bajos respecto del resto.

Figura 9: Clasificación con 9 clusters



Teniendo en cuenta que la zona analizada se conoce previamente se considera adecuado considerar 9 *cluster*.

CONCLUSIÓN

En este trabajo se describió en detalle el procedimiento seguido para clasificar un área sembrada a partir de imágenes satelitales. La clasificación de las imágenes se realizó a partir del índice NDVI que es utilizado ampliamente en el sector agropecuario para determinar la productividad del área sembrada.

A partir de este desarrollo se proponen futuras líneas de investigación que involucren desde lo metodológico la adquisición de una serie de imágenes de una misma región con el objeto de plantear modelos para estimar la productividad y detectar anomalías. Por otro lado, como se ha mencionado, la actividad agropecuaria es una de las que mayor porcentaje aporta a la tributación, por lo cual, en relación con la política tributaria, otra línea sería analizar un escenario posible que permita establecer un valor de “renta presunta” a la actividad agrícola proporcional a la productividad, estimada mediante métodos de machine learning a partir del NDVI.

REFERENCIAS BIBLIOGRÁFICAS

- Bivand R., Keitt T., Rowlingson R. (2019). rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.4-4. <https://CRAN.R-project.org/package=rgdal>
- Boken, V. K. y Shaykewich, C. F. (2002). Improving an operational wheat yield model using phenological phase-based Normalized Difference Vegetation Index. *International Journal of Remote Sensing*, 23, 4155–4168.
- Bolsa de Cereales de Córdoba. (2019). Obtenido de <http://www.bccba.com.ar>
- CONAE, (2016). Índices Espectrales derivados de imágenes satelitales Landsat 8 Sensor OLI. Guía de Usuario.
- CONAE, (2018). Material del Curso SoPI: Introducción a la Teledetección
- Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Deering, D. W. (1978). Rangeland reflectance characteristics measured by aircraft and spacecraft sensors. Ph.D. Dissertation, Texas A & M University, College Station, TX, 338 pp
- Doraiswamy, P. C. y Cook, P. W. (1995). Spring wheat yield assessment using NOAA AVHRR data. *Canadian Journal of Remote Sensing*, 21, 43–51.
- FADA. (Septiembre de 2019). Obtenido de <http://fundacionfada.org/informes/indice-fada-septiembre-2019-564/>
- Hastie, T., Tibshirani, R., Friedman, J. H. (2001). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer.
- Hijmans, R (2019). raster: Geographic Data Analysis and Modeling. R package version 3.0-2. <https://CRAN.R-project.org/package=raster>
- Hijmans, R (2019). <https://rspatial.org/>
- IDECOR ,2019. Obtenido de: <https://idecor.cba.gov.ar>
- INTA, 2017. (s.f.). Obtenido de https://inta.gob.ar/sites/default/files/inta_informe_estadistico_del_mercado_de_soja.pdf

- Lloyd, S. P. (1957, 1982). Least squares quantization in PCM. Technical Note, Bell Laboratories. Published in 1982 in *IEEE Transactions on Information Theory*, 28, 128–137.
- Long, L.S., Solana, C., Canters, F., Chen, L., & Kervyn, M. (2017). Testing random forest classification for identification 1 and aging of lava flows from a single Landsat 8 image 2.
- Ministerio de Agricultura, Ganadería y Pesca. Presidencia de la Nación. (2019). Obtenido de <https://geoadmin.agroindustria.gob.ar/geonetwork/srv/spa/catalog.search#/home>
- Open Data Charter. (s.f.). Obtenido de <https://opendatacharter.net/principles-es/>
- Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2), <https://cran.r-project.org/doc/Rnews/>.
- Presidencia de la Nación-Ministerio de Modernización. (s.f.). Obtenido de Paquete de Apertura de Datos de la República Argentina: <https://datosgobar.github.io/paquete-apertura-datos/guia-subnacionales/#1-que-son-los-datos-abiertos>
- Roger S. Bivand, Edzer Pebesma, Virgilio Gomez-Rubio, 2013. Applied spatial data analysis with R, Second edition. Springer, NY. <http://www.asdar-book.org/>
- Rouse, J.W., Jr., R.H. Haas, J.A. Schell, and D.W. Deering. (1973). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Prog. Rep. RSC 1978-1, Remote Sensing Center, Texas A&M Univ., College Station, 93p. (NTIS No. E73-106393).
- Schomwandt, David. (2015). Teledetección aplicada a las Ciencias Agronómicas y recursos naturales. http://www.siaa.gob.ar/joomla_files/images/mapas/ManualSensores.pdf
- Sobrino, J. A., Raissouni, N., Kerr, Y., Oliosio, A., López-García, M. J., Belaid, A., El Kharraz, M. H., Cuenca, J., Dempere, L., (2000). Teledetección. Sobrino, J. A. (Ed.), Servicio de Publicaciones, Universidad de Valencia (ISBN 84-370-4220-8), Valencia (España).
- Tucker, C. J., Holben, B. N., Elgin, J. H. y McMurtrey, J. E. (1980). Relationships of spectral data to grain yield variation. *Photogrammetric Engineering and Remote Sensing*, 46, 657–666.
- U.S. Department of the Interior, U.S. Geological Survey. (2015). LANDSAT 8 (L8) Data users handbook (pdf). p. 106.
- USGS. (s.f.). Obtenido de <https://earthexplorer.usgs.gov/>

Wickham H., François R., Henry L., Müller K. (2019). dplyr: A Grammar of Data Manipulation. R
package version 0.8.3. <https://CRAN.R-project.org/package=dplyr>