

EVALUACIÓN DE MODELOS DE APRENDIZAJE ESTADÍSTICO APLICADOS EN  
ESTIMACIÓN DE VARIABLES BIOFÍSICAS EN CULTIVOS DE CAÑA DE AZÚCAR

Proyecto de Maestría



CAMILO ALBERTO HERRERA ROZO

UNIVERSIDAD DE BUENOS AIRES  
FACULTAD DE AGRONOMIA  
MAESTRÍA EN BIOMETRÍA Y MEJORAMIENTO  
Buenos Aires - Argentina, Junio  
2014

## Introducción

En la actualidad existe una gran cantidad de métodos estadísticos donde cada uno de estos tiene una infinidad de variaciones y posibles aplicaciones. El uso de la inteligencia artificial (visión computacional, sistemas expertos, sistemas de ayuda de decisión, etc.) y otras técnicas prometedoras de esta (redes neuronales, lógica difusa y bioinformática) pueden proporcionar soluciones a los problemas en sistemas agrícolas complejos de manera eficaz (Bustos M, 2005). En este contexto hemos seleccionado tres métodos de interés en los cuales deseamos evaluar su funcionamiento, adaptación, eficiencia y precisión, ya que estos métodos son herramientas con un importante potencial de aplicación en casos de aprendizaje y clasificación estadística (Mitchell, 1997)(Samuel, 1959), los métodos seleccionados son: PLS o Regresión por mínimos cuadrados parciales (De Jong, 1990), Random Forest (Breiman, 2001) y el Aprendizaje por cuantificación vectorial (LVQ) (Kohonen, 1997). Estos tres métodos serán aplicados sobre imágenes multiespectrales (Hough, 1991) y datos obtenidos mediante procesos de simulación del comportamiento de cultivos de caña de azúcar.

La Agricultura de precisión y manejo sitio-específico (AEPS), se define como el arte de realizar las prácticas agronómicas requeridas por una especie vegetal, de acuerdo con las condiciones espaciales y temporales del sitio donde se cultiva, para obtener de ellas su rendimiento potencial (Sandoval et al., 2012). La medición de variables biofísicas como lo son Área foliar (AF) o Índices de área foliar (IAF) en cultivos de caña de azúcar se realiza en la actualidad por medio de procesos destructivos. También realizar el conteo de cantidad de tallos (NT) en cultivos de caña de azúcar es un proceso de estimación empírico y poco preciso, estas 2 características plantean un panorama interesante de investigación con la finalidad de lograr generar nuevos métodos de estimación indirecta y precisa, que se espera lograr por medio de estos métodos estadísticos aplicados sobre imágenes aéreas y procesos de simulación.

Este proyecto pretende hacer uso de herramientas novedosas, integrando métodos estadísticos y sistemas de visión artificial o visión por computador (Gonzalez y Woods, 1996), para mejorar la comprensión de procesos naturales que son reflejados por los cultivos y captados por los sensores remotos, que facilitan la toma de información para lograr el reconocimiento de patrones espaciales presentes en los cultivos (Gutiérrez, 2006). Adicionalmente se desea cubrir la necesidad cada vez mayor de automatizar y mejorar procesos costosos o de características destructivas relacionados a la toma de información, permitiendo hacer un mejor manejo agronómico y apuntando a sistemas de agricultura de precisión y manejo sitio-específico en cultivos de caña de azúcar.

## **Objetivos**

### **Objetivo general:**

Integrar métodos estadísticos y sistemas de visión artificial o visión por computador, para mejorar la comprensión de procesos naturales, reflejados por cultivos de caña de azúcar y captados por sensores remotos.

### **Objetivos específicos:**

1. Evaluar el funcionamiento, adaptación, eficiencia y precisión de tres algoritmos de inteligencia artificial [Aprendizaje por cuantificación vectorial (LVQ)], [La Regresión por mínimos cuadrados Parciales] y [Random Forest] aplicados en imágenes multiespectrales de cultivos de caña de azúcar.
2. Estimar el área foliar (AF) y el índice de área foliar (IAF) en cultivos de caña de azúcar y la cantidad de tallos (NT) en cultivos de caña de azúcar.
3. Ajustar modelos estadísticos y de inteligencia artificial que relacionen correctamente la información de las imágenes multiespectrales a las variables de cultivo (AF), (IAF) y (NT).

## Bibliografía

- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.
- Bustos M, J. R. (2005). Inteligencia artificial en el sector agropecuario. *Seminario Investigativo*, Universidad Nacional de Colombia. Bogotá, Colombia,.
- De Jong, S. (1990). Multivariate calibration, h. martens and t. naes, wiley, new york, 1989. isbn 0 471 90979 3 no. of pages: 504. *Journal of Chemometrics*, 4(6):441–441.
- Gonzalez, R. and Woods, R. (1996). *Tratamiento digital de imágenes*. Editorial Díaz de Santos, S.A.
- Gutiérrez, C.P. y Nieto, L. (2006). *Teledetección: nociones y aplicaciones*. Carlos Pérez Gutiérrez, Ángel Luis Muñoz Nieto.
- Hough, H. (1991). *Satellite Surveillance*. Loompanics Unlimited.
- Kohonen, T., editor (1997). *Self-organizing Maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3):210–229.
- Sandoval, P., González, J., de Investigación de la Caña de Azúcar de Colombia, C., and Colombia. Departamento Administrativo de Ciencia, T. e. I. (2012). *Principios y aplicaciones de la percepción remota en el cultivo de la caña de azúcar en Colombia*. Centro de Investigación de la Caña de Azúcar de Colombia.