

# Predicción de la sequía agrícola en Chile: regresión lineal vs deep learning

30 de abril de 2019

## Resumen

La seguridad alimentaria mundial se ve afectada negativamente por la sequía. Las proyecciones climáticas predicen que la frecuencia y la intensidad de la sequía pueden aumentar en diferentes partes del mundo, lo que es particularmente peligroso para los países en desarrollo. Los pronósticos tempranos de la temporada sobre la ocurrencia y severidad de la sequía podrían ayudar a mitigar mejor sus consecuencias negativas. El objetivo de este estudio fue evaluar si la sequía agrícola se puede predecir con precisión a partir de fuentes de datos casi en tiempo real disponibles de forma gratuita. Como variable de respuesta, utilizamos la puntuación estándar del NDVI acumulado estacional (zcNDVI) en base a los datos de 2000-2017 de MODIS como proxy de las anomalías de la productividad primaria estacional. Las predicciones se realizaron con tiempos de entrega pronosticados entre uno y seis meses antes del final de la temporada, que variaban entre las unidades censales. Las variables predictoras incluyeron el zcNDVI obtenido al acumular NDVI desde el inicio de la temporada hasta el tiempo de predicción; Índices de precipitación estandarizados para escalas de tiempo de uno, tres, seis, 12 y 24 meses derivados de estimaciones de lluvia satelital; dos índices de oscilación climática que incluyen la Oscilación Decenal del Pacífico y el índice ENOS Multivariado; la duración de la temporada de crecimiento; y latitud y longitud. Para cada una de las 758 unidades censales consideradas, las series temporales de la respuesta y las variables predictoras se promediaron para las áreas agrícolas, lo que dio como resultado una serie temporal de 17 estaciones por unidad para cada variable. Los enfoques de predicción utilizados fueron: (i) regresión lineal óptima (OLR), por lo que para cada unidad de censo se seleccionó el único predictor que explicaba mejor la variabilidad interanual de zcNDVI, y (ii) una arquitectura de red neuronal multicapa de alimentación directa a menudo llamada aprendizaje profundo (DL) donde todos los predictores para todas las unidades se combinaron en un solo modelo espacio-temporal. Ambos enfoques se evaluaron con un procedimiento de validación cruzada de un año de ausencia. Ambos métodos mostraron una buena precisión de predicción para tiempos de entrega pequeños y valores similares para todos los tiempos de entrega. Los valores medios de  $R_{cv}^2$  para OLR fueron 0.95, 0.83, 0.68, 0.56, 0.46 y 0.37, contra 0.96, 0.84, 0.65, 0.54, 0.46 y 0.38 para DL, para uno, dos, tres, cuatro, cinco y seis meses de tiempo de entrega, respectivamente. Dada la amplia gama de climas y tipos de vegetación cubiertos dentro del área de estudio, esperamos que los modelos presentados puedan contribuir a un sistema mejorado de alerta temprana para la sequía agrícola en todo el mundo.

En este estudio el procesamiento de los datos espaciales se realizó con los paquetes ‘raster’ [1], ‘sp’ [2] y ‘maptools’ [3]. Para el resumen y manejo de datos se utilizaron la serie de paquetes de ‘tidyverse’ [4]. La implementación de las redes neuronales multi-capas (deep learning) se realizó con el framework de java ‘h2o’ mediante el paquete ‘wrapper’ de R ‘h2o’ [5]. Los análisis de regresión lineal con los paquetes ‘base’ de R. Todos los gráficos se realizaron con el paquete ‘ggplot2’ [6].

Este estudio se publicó en Diciembre del año 2018 [7] en la revista ‘Remote Sensing of Environment’, el artículo original se puede descargar del siguiente link <https://www.sciencedirect.com/science/article/pii/S0034425718304541>.

## References

1. R. J. Hijmans, *Raster: Geographic data analysis and modeling* (2019).
2. E. J. Pebesma & R. S. Bivand, Classes and methods for spatial data in R. *R News*, **5** (2005) 9–13.

3. R. Bivand & N. Lewin-Koh, *Maptools: Tools for handling spatial objects* (2019).
4. H. Wickham, *Tidyverse: Easily install and load the 'tidyverse'* (2017).
5. E. LeDell, N. Gill, S. Aiello, A. Fu, A. Candel, C. Click, T. Kraljevic, T. Nykodym, P. Aboyoun, M. Kurka, & M. Malohlava, *H2o: R interface for 'h2o'* (2019).
6. H. Wickham, *Ggplot2: Elegant graphics for data analysis* (Springer-Verlag New York, 2016).
7. F. Zambrano, A. Vrieling, A. Nelson, M. Meroni, & T. Tadesse, Prediction of drought-induced reduction of agricultural productivity in Chile from MODIS, rainfall estimates, and climate oscillation indices. *Remote Sensing of Environment*, **219** (2018) 15–30. <https://doi.org/10.1016/J.RSE.2018.10.006>.