

# Prediction of agricultural drought in Chile from multiple spatio-temporal data sources

Francisco Zambrano<sup>1</sup>, Anton Vrieling<sup>2</sup>, Andy Nelson<sup>2</sup>, Michele Meroni<sup>3</sup>, and Tsegaye Tadesse<sup>4</sup>

<sup>1</sup>Héméra Centro de Observación de la Tierra

<sup>2</sup>International Institute for Geo-Information Science and Earth Observation

<sup>3</sup>Institute for Environment and Sustainability JRC

<sup>4</sup>National Drought Mitigation Center, University of Nebraska-Lincoln

November 26, 2022

## Abstract

Global food security is negatively affected by drought. Climate projections show that drought frequency and intensity may increase in different parts of the globe. Early season forecasts on drought occurrence and severity could help to better mitigate the negative consequences of drought. The objective of this study was to assess if interannual variability in agricultural productivity in Chile can be accurately predicted from freely-available, near real-time data sources. As the response variable, we used the standard score of seasonal cumulative NDVI (zcNDVI), based on 2000-2017 data from Moderate Resolution Imaging Spectroradiometer (MODIS), as a proxy for anomalies of seasonal primary productivity. The predictions were performed with forecast lead-times from one- to six-month before the end of the growing season, which varied between census units in Chile. Predictor variables included the zcNDVI obtained by cumulating NDVI from season start up to prediction time; standardised precipitation indices, derived from satellite rainfall estimates, for time-scales of 1, 3, 6, 12 and 24 months; the Pacific Decadal Oscillation and the Multivariate ENSO oscillation indices; the length of the growing season, and latitude and longitude. We used two prediction approaches: (i) optimal linear regression (OLR) whereby for each census unit the single predictor was selected that best explained the interannual zcNDVI variability, and (ii) a multi-layer feedforward neural network architecture, often called deep learning (DL), where all predictors for all units were combined in a single spatio-temporal model. Both approaches were evaluated with a leave-one-year-out cross-validation procedure. Both methods showed good prediction accuracies for small lead times and similar values for all lead times. The mean R<sup>2</sup>cv values for OLR were 0.95, 0.83, 0.68, 0.56, 0.46 and 0.37, against 0.96, 0.84, 0.65, 0.54, 0.46 and 0.38 for DL, for one, two, three, four, five, and six months lead time, respectively. Given the wide range of climates and vegetation types covered within the study area, we expect that the presented models can contribute to an improved early warning system for agricultural drought in different geographical settings around the globe.

# Prediction of agricultural drought in Chile from multiple spatio-temporal data sources

Francisco Zambrano<sup>1</sup>, Anton Vrieling<sup>2</sup>, Andy Nelson<sup>3</sup>, Michele Meroni<sup>3</sup>, Tsegaye Tadesse<sup>4</sup>

**Motivation:** Although there is debate whether drought frequency has increased in recent years (Dai, 2012;), climate change is expected to exacerbate the phenomenon and lead to more frequent and intense drought periods, which may even occur in regions where overall precipitation increases are expected (IPCC, 2013). Planning for effective adaptation strategies is thus crucial to mitigate future impacts (Roco et al., 2014). In addition, the ability to anticipate the impact of drought early in the season and take in-season mitigation measures such as more targeted irrigation, or reducing stand density (Bodner et al., 2015) could help to reduce crop losses (Pulwarty and Sivakumar, 2014).

**Goal:** The main goal of this study is to assess if interannual variability in crop biomass productivity can be accurately predicted using freely-available, near real-time data sources.

## Methods

Study area: main agricultural land of Chile

