Remote Sensing of Environment 219 (2018) 15–30

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Prediction of drought-induced reduction of agricultural productivity in Chile from MODIS, rainfall estimates, and climate oscillation indices

Francisco Zambrano[a,b,*], Anton Vrieling[b], Andy Nelson[b], Michele Meroni[c], Tsegaye Tadesse[d]

[a] Hémera Centro de Observación de la Tierra, Escuela de Agronomía, Facultad de Ciencias, Universidad Mayor, La Piramide 5750, Huechuraba, Santiago, Chile
[b] University of Twente, Faculty of Geo-Information Science and Earth Observation, P.O. Box 217, 7500 AE Enschede, the Netherlands
[c] Directorate D - Sustainable Resources, Food Security Unit, European Commission, Joint Research Centre (JRC), Ispra, VA, Italy
[d] National Drought Mitigation Center, University of Nebraska-, Lincoln, NE 68583-0988, USA

ARTICLE INFO

Keywords:
Early-warning
Drought
Agriculture
Productivity
Prediction
Chile
MODIS
zcNDVI
Phenology

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. These increases are particularly hazardous for developing countries. 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. For each of the 758 census units considered, the time series of the response and the predictor variables were averaged for agricultural areas resulting in a 17-season time series per unit for each variable. 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_{cv}^2 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.

1. Introduction

Droughts cause major agricultural production losses worldwide (Campbell et al., 2016). Although there is debate whether drought frequency has increased in recent years (Dai, 2012; Sheffield et al., 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; McVicar et al., 2012; Trenberth et al., 2014; Wild, 2009). Although increased levels of carbon dioxide in the atmosphere may increase the water use efficiency of crops (Donohue et al., 2013; Lu et al., 2016; Zhu et al., 2016), the combined effects of global mean temperature (Zhao et al., 2017) and drought occurrence (Dai, 2012) are expected to cause an overall reduction of global crop yields (Ray et al., 2015; Zhao et al., 2017) and may have a negative impact on cropping frequency and sown area (Cohn et al., 2016). 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 (e.g., by applying a regulated deficit irrigation), or

* Corresponding author.
E-mail address: francisco.zambrano@umayor.cl (F. Zambrano).

https://doi.org/10.1016/j.rse.2018.10.006
Received 10 October 2017; Received in revised form 26 September 2018; Accepted 4 October 2018

reducing stand density (Bodner et al., 2015) could help to reduce crop losses (Pulwarty and Sivakumar, 2014; Wilhite et al., 2000, 2014).

Satellite image time series have been widely used for monitoring agricultural drought (AghaKouchak et al., 2015). Commonly-derived variables from such time series include vegetation indices and rainfall estimates, which can be translated into anomalies by comparing the values in the current year against historic distributions (Ashouri et al., 2015; Funk et al., 2015; Huffman et al., 2007). The most commonly-used vegetation index for this purpose is the NDVI (Normalised Difference Vegetation Index; Rouse et al., 1974) from which multiple anomaly measures have been derived (Jiao et al., 2016; Kogan, 1990; Peters et al., 2002; Sandholt et al., 2002) and applied for monitoring agricultural drought (Cunha et al., 2015; McVicar and Jupp, 1998; Rojas et al., 2011; Zambrano et al., 2016; Zhang and Jia, 2013). Besides NDVI, other vegetation indices have been developed with the aim to further increase the sensitivity to vegetation activity, including the enhanced vegetation index (EVI; Huete et al., 2002) and the Vegetation Index based on the Universal Pattern Decomposition (VIUPD; Zhang et al., 2007). An often-used approach to translate rainfall information into a meteorological drought measure is through the calculation of the Standardised Precipitation Index (SPI; McKee et al., 1993), an anomaly measure that is closely related to soil moisture availability when computed for short time scales (i.e. < 9 months) (Quiring and Ganesh, 2010). While the SPI can be calculated from weather station data, in countries with a low station density and short historical records, satellite-derived rainfall estimates (RFE) can be a good alternative source for SPI calculation (Tapiador et al., 2012; Zambrano et al., 2017). Other satellite-derived products that have relevance for drought monitoring include estimates of soil moisture and evapotranspiration (Hao and AghaKouchak, 2013; Mu et al., 2013; Sheffield et al., 2004; Tsakiris et al., 2007). Drought indices are also constructed by combining multiple parameters. For example, the SPEI (Standardised Precipitation Evapotranspiration Index) considers both precipitation and evapotranspiration to account for the effects of temperature variability on drought assessment (Vicente-Serrano et al., 2010) and has been used in various studies for monitoring agricultural drought (Moorhead et al., 2015; Potopová et al., 2015; Vicente-Serrano et al., 2012). Climate-based drought indicators and satellite-derived vegetation metrics with other biophysical information are used by the United States National Drought Mitigation Center to calculate the Vegetation Drought Response Index (VegDRI, J.F. Brown et al., 2008; Brown et al., 2008) by analysing historical input data with a regression tree approach that produces a map of drought conditions. While we can accurately assess and monitor agricultural drought as it occurs with a variety of indices, early prediction is more complex.

The prediction of vegetation conditions is challenging for three reasons: 1) the underlying uncertainties in weather and climate prediction (Morss et al., 2008); 2) changes in precipitation patterns (Dore, 2005), such that precipitation amount and distribution may substantially deviate from *normal* regional patterns; and 3) the effect of both of these on crop production which in turn depends on land management decisions and the sensitivity of different crop stages to the intensity and duration of water shortage or excess (Knapp et al., 2008; Sykes, 2001), amongst other factors. Despite this challenge, various efforts have been made to predict vegetation conditions. Table 1 provides a non-exhaustive overview of relevant studies that aim to predict agricultural productivity, or a related proxy, from remote sensing-derived predictors.

Several studies have used a single predictor to explain the interannual variability in seasonal vegetation productivity. In those cases, early prediction was achieved by using lagged relationships whereby the predictor was available before the end of the season (Meroni et al., 2014; Vrieling et al., 2016). Alternatively, rainfall has been used as a predictor of seasonal vegetation productivity. For example, Meroni et al. (2017) found for the Sahel that on average about 40% of the variability in seasonal vegetation productivity could be explained by selecting the optimal time-scale and timing of SPI per grid cell. In addition, climatic oscillation indices, such as the Pacific Decadal Oscillation (PDO) and the Multivariate ENSO Index (MEI), have been shown to affect weather across the globe and can explain variability in agricultural productivity (Boisier et al., 2016; Garreaud and Battisti, 1999; Hansen et al., 1998; Marj and Meijerink, 2011; Montecinos and Aceituno, 2003; Reilly et al., 2003).

Brown et al. (2010) demonstrated that the timing of the growing season cumulative NDVI depends significantly on the PDO and the MEI across multiple locations in Africa. Van Leeuwen et al. (2013) showed that the MEI and Antarctic Oscillation (AAO) index could explain part of the interannual variability in annual NDVI-based productivity and phenology for South America. While these studies assessed the explanatory power of climatic oscillation indices on vegetation variability, they did not specifically address the prediction of vegetation productivity shortfalls before they occur. Although studies focusing on single parameters offer interesting directions for the prediction of vegetation productivity, combining multiple predictors could increase the prediction skills.

The use of multiple predictors to estimate vegetation response has been evaluated by applying different techniques such as multiple linear regression models and regression trees. Commonly-used predictors include lagged NDVI, precipitation-derived indices such as SPI, soil moisture, and oscillation indices (Table 1). Optimal prediction skills of NDVI variability range approximately between 90% (one month before) and 50% (three months before), as achieved for example by Tadesse et al. (2014). More recently, machine learning methods have been used for predicting daily and monthly rainfall (Abbot and Marohasy, 2012, 2014; Deo and Şahin, 2015; Nastos et al., 2014), mostly because they can accommodate a large number of input variables and can automatically combine these into complex functions that describe multiple, non-linear relationships between the independent and explanatory variables (LeCun et al., 2015). Given the mentioned advantages of machine-learning methods and the lack of current drought prediction tools, there is scope to evaluate if machine learning methods could provide more accurate early prediction of drought than linear regression models.

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. The sources included NDVI time series, anomalies of cumulative rainfall at different monthly time-steps obtained from satellite-derived RFEs, and climatic oscillation indices. To achieve our goal we tackle the following objectives: (i) derivation of a proxy for seasonal crop biomass productivity for the growing season; (ii) development of two prediction models for that proxy, one based on optimal linear regression model (OLR) per spatial unit, and the other combining information from all units in a feed-forward multi-layer neural network, known as deep learning (DL); and (iii) evaluation of the OLR and DL model predictions for the period 2000–2017. The Methods, Results and Discussion sections are structured according to these three objectives.

2. Study area

We selected Chile as a case study area for this analysis, because it contains a large diversity in climate and crops, and frequently suffers from drought-induced crop losses (Zambrano et al., 2016). Future crop losses can be expected because a decrease in precipitation is predicted for the Central and South part by the global climate models (IPCC, 2013) where about 84% of the agricultural activities are concentrated (INE, 2007). As a result, wheat and maize yields for Chile are expected to decrease by about 15% to 20% by 2050 (IPCC, 2014; Meza and Silva, 2009) based on crop growth simulation models with future climate scenario data. During the past decade, an unusual long period of dry conditions persisted over Central Chile and has been termed a *mega drought* (Garreaud et al., 2017). The focus of this study is the main