Artificial Intelligence and machine learning model for spatial and temporal prediction of drought in the Colombia Caribbean region.

Daissy Herrera [1] dmherrerap@unal.edu.co – Edier Aristizábal [1] evaristizabalg@unal.edu.co
[1] Universidad Nacional de Colombia. Sede Medellin. Facultad de Minas. Departamentos de Geociencias y Medio Ambiente.

1. Introduction

Drought is one of the most critics hydrometeorological phenomena in terms of impacts on society [1]. Drought has produced several impacts in Colombia with important affectations on the population, such as forest fires, water shortages [2] and agriculture and fish farming losses [3]. However, a very small number of studies have been carried out to predict and evaluate drought events. In this research, we propose a data-based model using Random Forest (RF) and Bagging Decision tree (DTC) in the northern Colombia zone.

2. Study Area.

Study Area: Magdalena, is a Department in the north region of Colombia (Figure 1) with an area of 23,188 km2 [4] and a population of 1'263,788 [5].

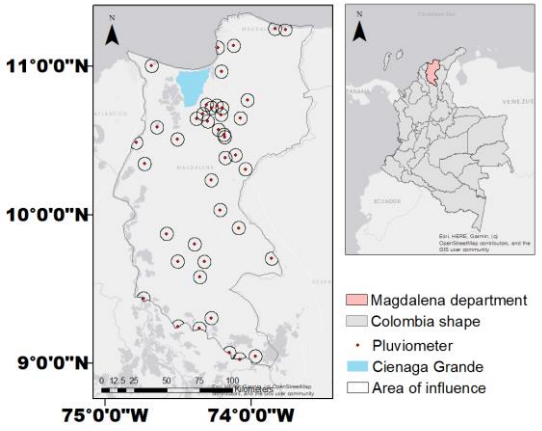

Figure 1. Study area location.

2. Data.

Remote sensing data and macro climatic variables was collected as independent variables (Figure 2) for the model in a period of time from 2010 to 2019 on a monthly time scale

As a response variable to identify agricultural drought events, the Standardized Precipitation Index (SPI) was used, with a 3-month time scale. It was calculated using pluviometry information derivate from IDEAM, that is the institute that measure environmental variables in Colombia.

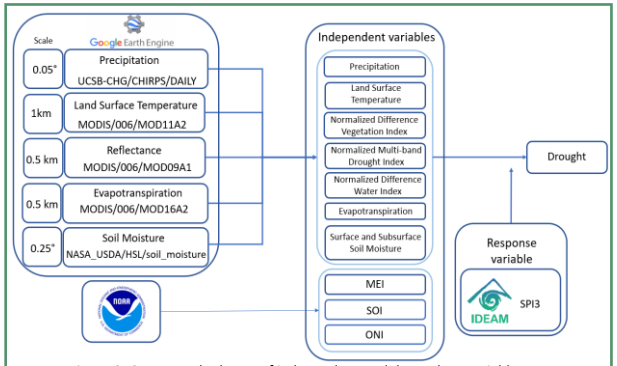

Figure 2. Conceptual scheme of independent and dependent variables.

3. Drought Forecasting Model.

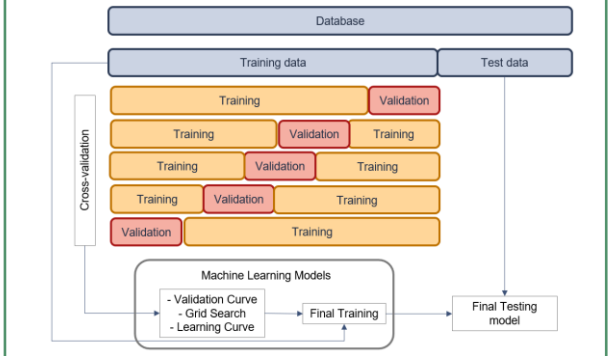

Figure 3. Conceptual scheme of the application of the RF and DTC models.

Bagging Decision tree:
The results of the final testing showed a precision of 0.33 and a recall of 0.8 for predicted drought conditions.

TN: 66.1%	FP: 21%	FN: 2.5%	TP: 10.3%

Random forest:
The results of the final testing showed a precision of 0.93 and a recall of 0.84 for predicted drought conditions.

TN: 79.6%	FP: 7.6%	FN: 2%	TP: 10.7%

3. Drought Forecasting Model.

The outcomes showed that Random Forest provides the best outcomes. Both models made fewer mistakes predicted real drought events, but DTC model had more errors forecasting real normal or humid conditions.

4. Distributed Model.

Taking all feature data over the Department in August of 2014 and forecasts SPI3, using RF, and comparing with real data, in this case, municipalities who reported impacts or affectation because of drought in August of 2014. It is obtained that:

- Over the entire reported municipalities, the model forecasted both conditions drought and normal or humid.
- Exist zones where the model predicted drought conditions but are not reported by governmental entities with drought in August of 2014.

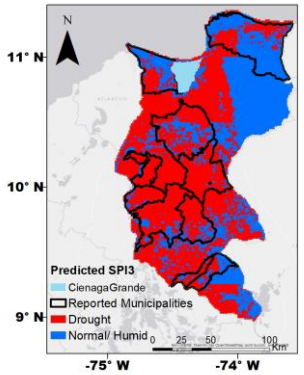

Figure 4. Predicted SPI3 and reported municipalities in August of 2014.

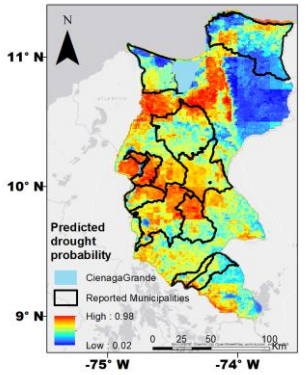

Figure 5. Predicted drought probability and reported municipalities in August of 2014

5. Conclusions.

- RF and DTC predicted well drought condition, but Random Forest provides the best outcomes because is the model that had fewer mistakes forecasting real drought events and normal or humid conditions in comparison with DTC.
- The implementation of the developed model can allow governmental entities assessment and monitor agricultural drought over time
- In the way that the governmental entities measure and report with exactitude when and where drought events occur, we can compare how close SPI3 describes real drought conditions in Magdalena Department and have comprehensive evaluations about the development model.

References: [1] FAO. Sequía. 2020. [Online]. Available: http://www.fao.org/emergencies/tipos-de-peligros-y-de-emergencias/sequia/en/. [2] Revista SEMANA. Las graves secuelas económicas de la sequía. 2014. [Online]. Available: https://www.semana.com/nacion/articulo/las-graves-secuelas-economicas-de-la-sequia/398750-3. [3] EL HERALDO. (2015) Magdalena, el más azotado por la temporada de sequía. [Online]. Available: https://www.elheraldo.co/magdalena/magdalena-el-mas-azotado-por-la-temporada-de-sequia-219590 [4] Gobernación del Magdalena. Nuestro departamento. 2020. [Online]. Descargado de http://www.magdalena.gov.co/departamento/nuestro-departamento
DANE. Resultados censo nacional de población y vivienda 2018. [Online]. Available: https://www.dane.gov.co/files/censo2018/informacion-tecnica/presentaciones-territorio/191004-CNPV-presentacion-Magdalena.pdf [5] DANE. Resultados censo nacional de población y vivienda 2018. [Online]. Available: https://www.dane.gov.co/files/censo2018/informacion-tecnica/presentaciones-territorio/191004-CNPV-presentacion-Magdalena.pdf