EGU2020-8173
https://doi.org/10.5194/egusphere-egu2020-8173
EGU General Assembly 2020

Deep Learning for Drought and Vegetation Health Modelling: Demonstrating the utility of an Entity-Aware LSTM

Thomas Lees[1], Gabriel Tseng[2], Steven Reece[3], and Simon Dadson[1]
[1]Oxford University, School of Geography and Environment, Oxford, United Kingdom of Great Britain and Northern Ireland (thomas.lees@chch.ox.ac.uk)
[2]Okra Solar, Phnom Penh, Cambodia
[3]University of Oxford, Pattern Recognition Group, Oxford, UK

Tools from the field of deep learning are being used more widely in hydrological science. The potential of these methods lies in the ability to generate interpretable and physically realistic forecasts directly from data, by utilising specific neural network architectures.

This approach offers two advantages which complement physically-based models. First, the interpretations can be checked against our physical understanding to ensure that where deep learning models produce accurate forecasts they do so for physically-defensible reasons. Second, in domains where our physical understanding is limited, data-driven methods offer an opportunity to direct attention towards physical explanations that are consistent with data. Both are important in demonstrating the utility of deep learning as a tool in hydrological science.

This work uses an Entity Aware LSTM (EALSTM; cf. Kratzert et al., 2019) to predict a satellite-derived vegetation health metric, the Vegetation Condition Index (VCI). We use a variety of data sources including reanalysis data (ERA-5), satellite products (NOAA Vegetation Condition Index) and blended products (CHIRPS precipitation). The fundamental approach is to determine how well we can forecast vegetation health from hydro-meteorological variables.

In order to demonstrate the value of this method we undertook a series of experiments using observed data from Kenya to evaluate model performance. Kenya has experienced a number of devastating droughts in recent decades. Since the 1970s there have been more than 10 drought events in Kenya, including droughts in 2010-2011 and 2016 (Haile et al 2019). The National Drought Monitoring Authority (NDMA) use satellite-derived vegetation health to determine the drought status of regions in Kenya.

First, we compared our results to other statistical methods and a persistence-based baseline. Using RMSE and R-squared we demonstrate that the EALSTM is able to predict vegetation health with an improved accuracy compared with other approaches. We have also assessed the ability of the EALSTM to predict poor vegetation health conditions. While better than the persistence baseline the performance on the tails of the distribution requires further attention.

Second, we test the ability of our model to generalise results. We do this by training only with subsets of the data. This tests our model's ability to make accurate forecasts when the model has not seen examples of the conditions we are predicting. Finally, we explore how we can use the EALSTM to better understand the physical realism of relations between hydro-climatic variables embedded within the trained neural network.

References:

Gebremeskel, G., Tang, Q., Sun, S., Huang, Z., Zhang, X., & Liu, X. (2019, June 1). Droughts in East Africa: Causes, impacts and resilience. Earth-Science Reviews. Elsevier B.V. https://doi.org/10.1016/j.earscirev.2019.04.015

Klisch, A., & Atzberger, C. (2016). Operational drought monitoring in Kenya using MODIS NDVI time series. Remote Sensing, 8(4). https://doi.org/10.3390/rs8040267

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., & Nearing, G. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrology and Earth System Sciences, 23(12), 5089–5110. https://doi.org/10.5194/hess-23-5089-2019

Github Repository: https://github.com/esowc/ml_drought