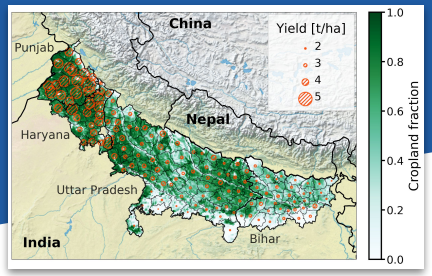
Explainable deep learning to predict and understand crop yield estimates

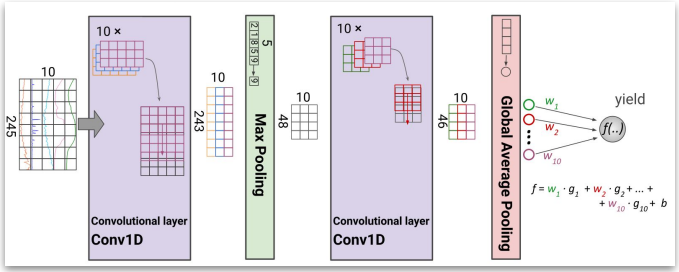
Aleksandra Wolanin[1,*], Gonzalo Mateo-García[2], Gustau Camps-Valls[2], Luis Gómez-Chova[2], Michele Meroni[3], Gregory Duveiller[3], You Liangzhi[4], and Luis Guanter[5]

[1] GFZ, Potsdam, Germany
[2] Universitat de València, València, Spain,
[3] European Commission Joint Research Centre (JRC), Ispra, Italy
[4] The International Food Policy Research Institute (IFPRI), Washington, D.C., United States of America
[5] Centro de Tecnologías Físicas, Universitat Politècnica de València, València, Spain

* ola@gfz-potsdam.de



Wolanin et al., 2020, *Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt*, Environmental Research Letters 2020



GFZ Helmholtz Centre Potsdam

HELMHOLTZ

EGU 2020

Short summary

- Environmental conditions influence crops and final grain yield in a complex non-linear manner, for which machine learning (ML) techniques can account
- ML typically lacks transparency and interpretability, though understanding which are the underlying factors behind both a predicted loss or gain is important
- **Aim:** maintaining the ability to **interpret** how the models achieve their results while benefiting from the **increased predictive performance of DL**
- Deep neural network (Fig. 1a) was applied to multivariate time series of vegetation and meteorological data to estimate the wheat yield in the Indian Wheat Belt
- Features and yield drivers learned by the model were analyzed and visualized and with the use of regression activation maps (Fig. 1b and 2)
- DL model outperformed other models and facilitated the interpretation of variables and processes leading to yield variability
- Learned features were mostly related to the length of the growing season, temperature, and light conditions during the growing season
- For example,high yields in 2012 were associated with low temperatures accompanied by sunny conditions during the growing period (Fig. 3)
- The proposed methodology can be used for other crops and regions in order to facilitate application of DL models in agriculture

Figure 1

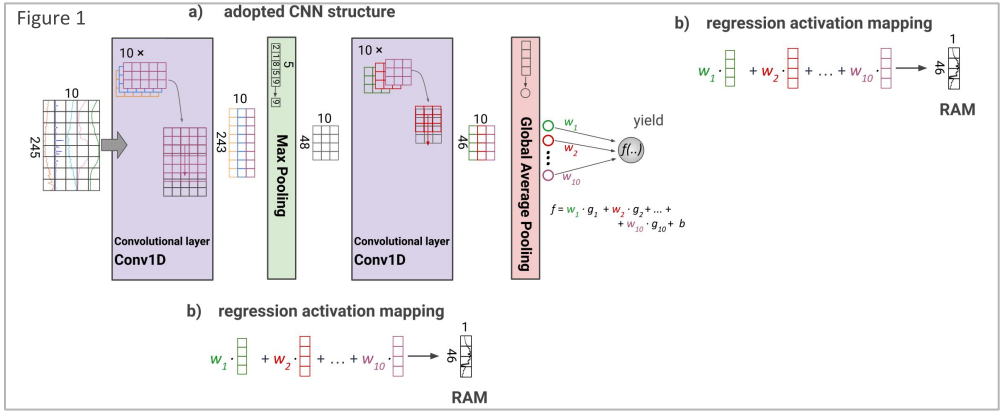


Figure 2 Regression activation maps for Patiala (2006)

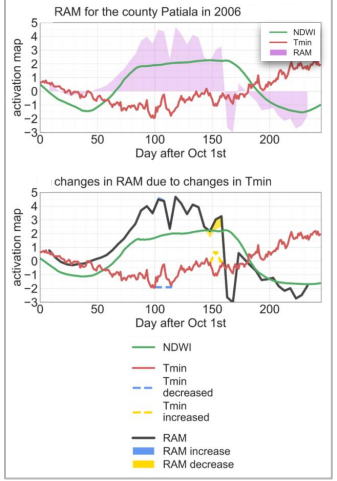
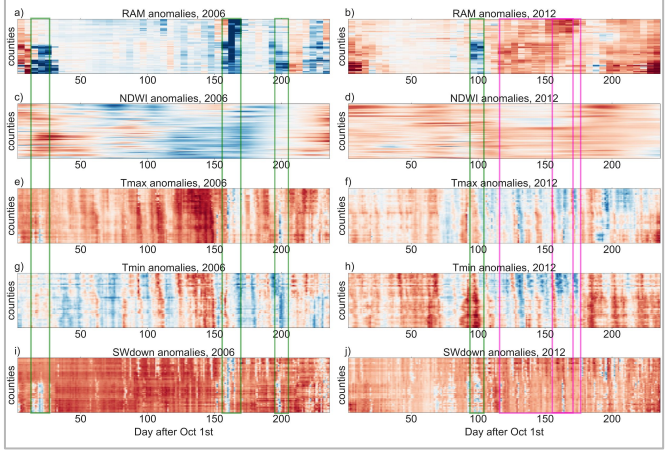


Figure 3 Anomalies of RAMs and selected input variables for all districts for the years 2006 and 2012. Time series for each district are shown as rows, and columns correspond to time steps.



More details in: Wolanin A., Mateo-García G., Camps-Valls G., Gómez-Chova L. ,Meroni, M., Duveiller, G., You, L., Guanter L. (2020) Estimating and Understanding Crop Yields with Explainable Deep Learning in the Indian Wheat Belt. Environmental Research Letters, 15, 02401