

Scalable Country-Level Crop Yield Modeling for Food Security and Risk Mitigation

Tackling Climate Change with Machine Learning

NeurIPS, Dec 2025

Andrew Hobbs

Benson Adomako

Addi Joof

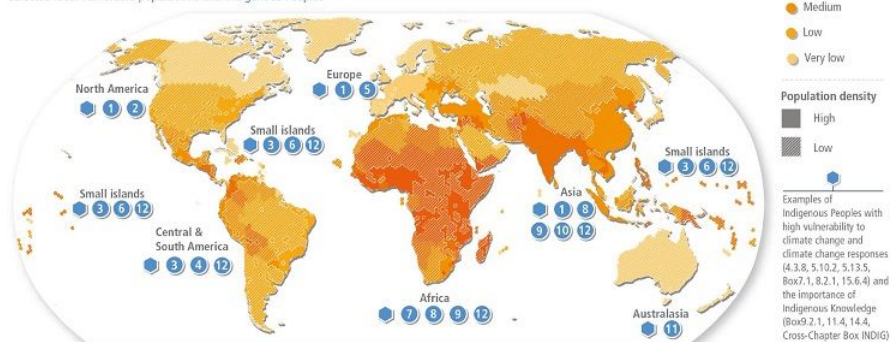
Jesse Anttila-Hughes

University of San Francisco, Department of Economics

Many of the world's most climate vulnerable are farmers

Observed human vulnerability differs between and within countries and strongly determines how climate hazards impact people and society

(a) Map of observed human vulnerability based on two comprehensive global indicator-systems using national data, plus examples of selected local vulnerable populations and Indigenous Peoples

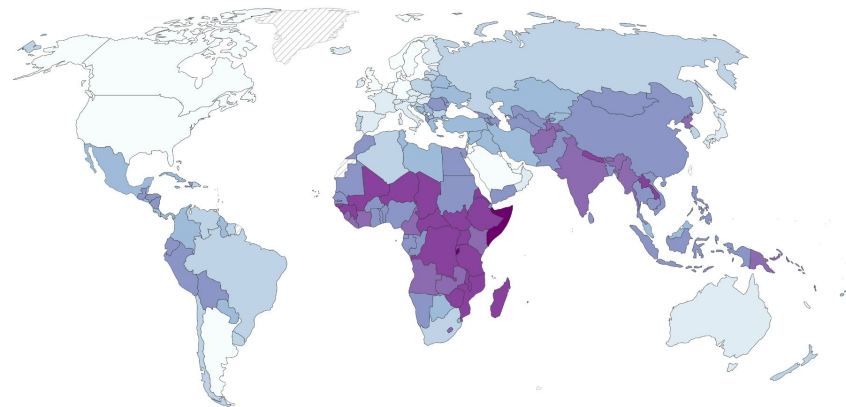


Examples of local vulnerable populations | Examples of some aspects of vulnerability | Chapter references

- 1 Indigenous Peoples of the Arctic | health inequality, limited access to subsistence resources and culture | CCP 6.2.3, CCP 6.3.1
- 2 Urban ethnic minorities | structural inequality, marginalisation, exclusion from planning processes | 14.5.9, 14.5.5, 6.3.6
- 3 Smallholder coffee producers | limited market access & stability, single crop dependency, limited institutional support | 5.4.2
- 4 Indigenous Peoples in the Amazon | land degradation, deforestation, poverty, lack of support | 8.2.1, Box 8.6
- 5 Older people, especially those poor & socially isolated | health issues, disability, limited access to support | 8.2.1, 13.7.1, 6.2.3, 7.1.7
- 6 Island communities | limited land, population growth and coastal ecosystem degradation | 15.3.2
- 7 Children in rural low-income communities | food insecurity, sensitivity to undernutrition and disease | 5.12.3
- 8 People uprooted by conflict in the Near East and Sahel | prolonged temporary status, limited mobility | Box 8.1, Box 8.4
- 9 Women & non-binary | limited access to & control over resources, e.g. water, land, credit | Box 9.1, CCB-GENDER, 4.8.3, 5.4.2, 10.3.3
- 10 Migrants | informal status, limited access to health services & shelter, exclusion from decision-making processes | 6.3.6, Box 10.2
- 11 Aboriginal and Torres Strait Islander Peoples | poverty, food & housing insecurity, dislocation from community | 11.4.1
- 12 People living in informal settlements | poverty, limited basic services & often located in areas with high exposure to climate hazards | 6.2.3, Box 9.1, 9.8, 10.4.6, 12.2.2, 12.3.5, 15.3.4

Share of the labor force employed in agriculture, 2019

Agriculture includes the cultivation of crops and livestock production, as well as forestry, hunting, and fishing. Employment includes anyone engaged in any activity to produce goods or services for pay or profit.



Source: Our World in Data based on International Labor Organization (via the World Bank) and historical sources
OurWorldInData.org/employment-in-agriculture • CC BY

Problem

- When crop yields are poor, farmers must rely on costly coping strategies, including reduced food consumption, removing children from school, or selling productive assets.
- Accurate, timely data on regional crop yields is essential to effective agricultural insurance, famine early warning, and government programs.
- Developing accurate crop yield measures requires costly ground-truth data
 - Precisely georeferenced farm yield data are scarce in developing countries
 - Existing measures are often not tested for accuracy.

Our Approach

- We combine reported yields existing household survey data from the World Bank's Living Standards Measurement Survey (LSMS) with remotely sensed satellite data on climate and environmental features from six African countries
- Apply machine learning techniques to develop a predictive index of maize production within each country and across countries.
- Develop a promising new approach for generating accurate regional yield estimates for smallholder farmers.

Data

Crop Yields (Maize): World Bank LSMS-ISA
Data from Nigeria, Uganda, Malawi, Tanzania,
Mali, & Ethiopia

Croplands: NASA Global Food Security
Analysis Support Data (GFSAD)

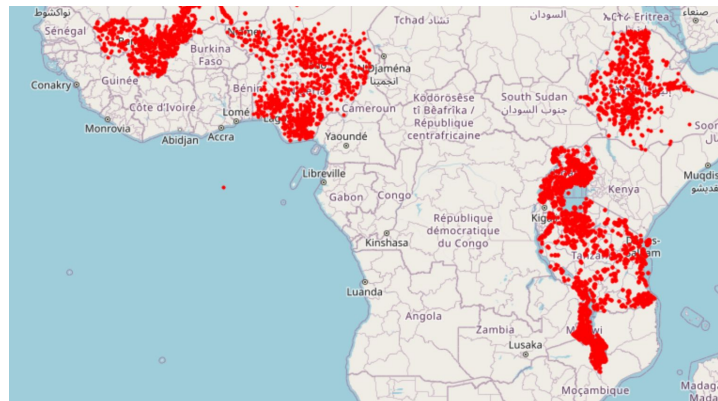
Vegetation Indices: NDVI, GCVI, and EVI
(LANDSAT) monthly maximum.

Growing/ Killing Degree Days, Temperature

+ Soil Moisture: ECMWF ERA5

Precipitation: CHIRPS

Total of ~30,000 observations in ~7,500
enumeration areas.

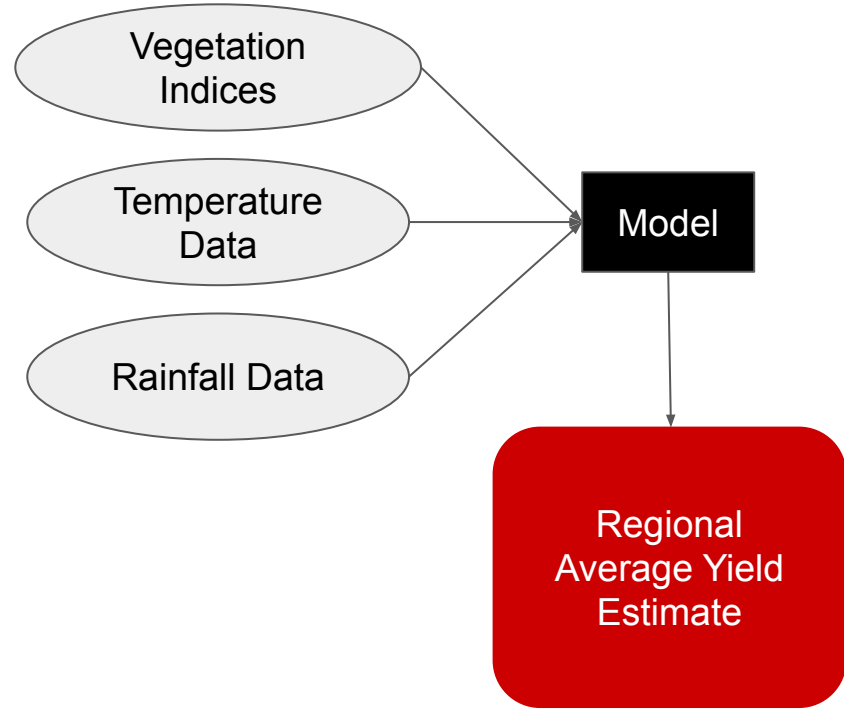


Why machine learning?

Multiple types of data tell us something about crop yields

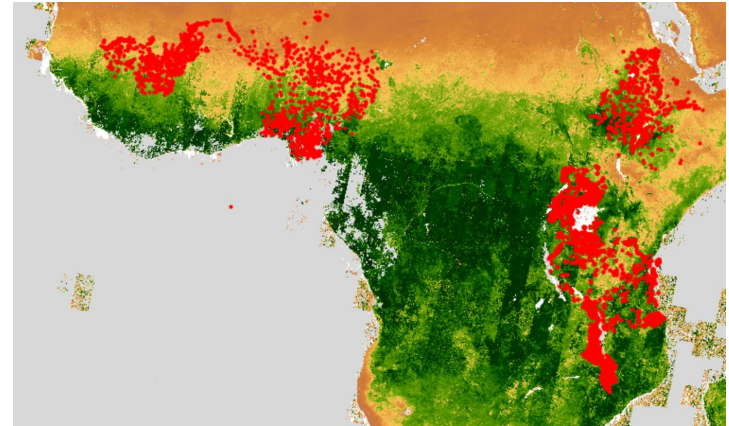
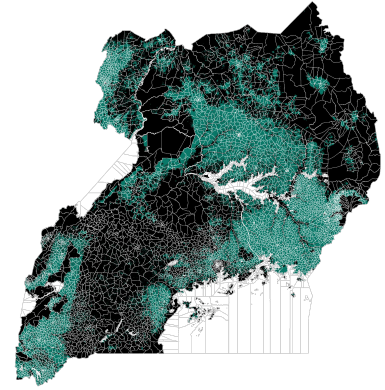
- Vegetation indices
- Precipitation
- Temperature
- Root/ soil moisture

Combining these data into a single index of average crop yields in an area is complex and machine learning is well suited to this sort of task.

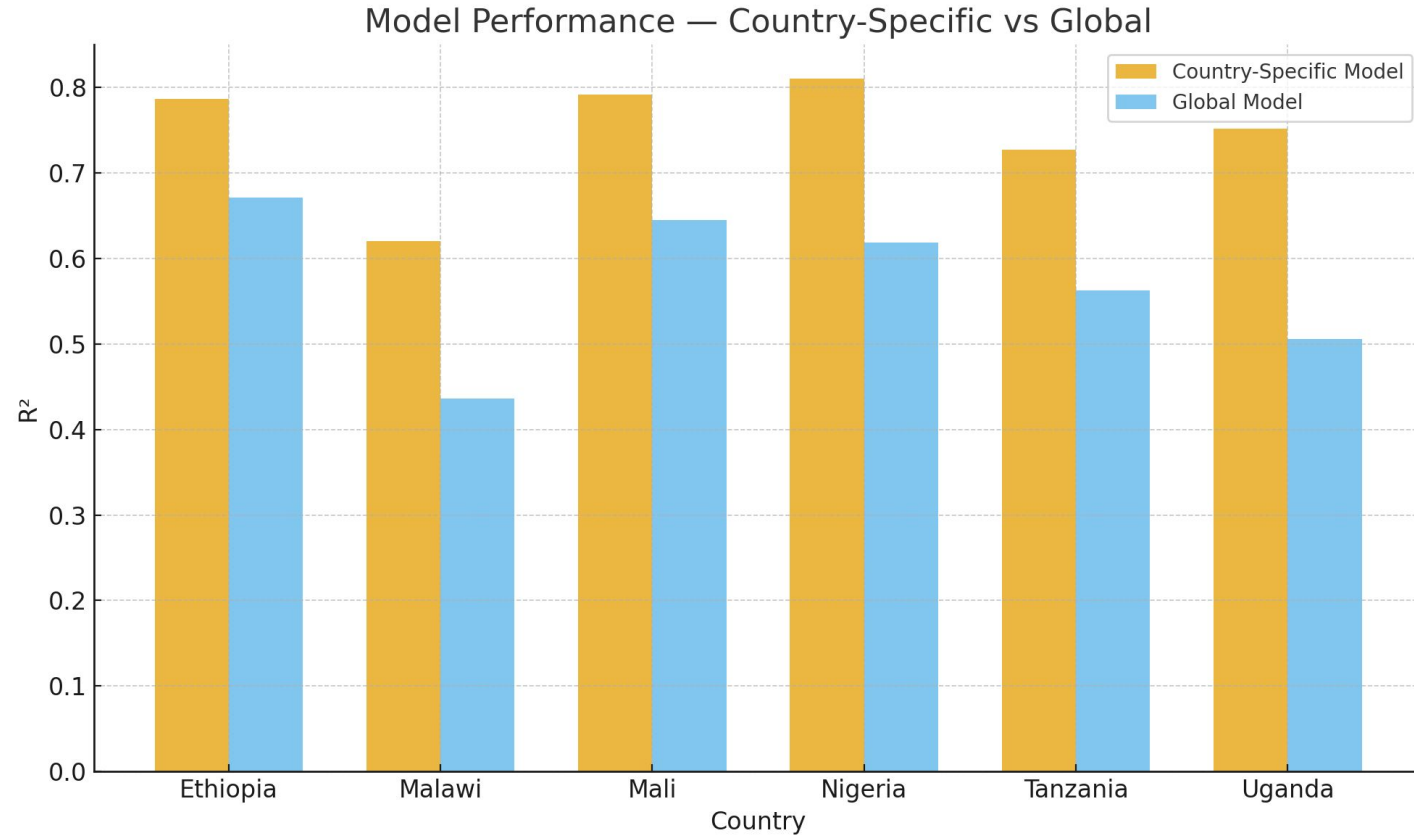


Linking satellite and survey data

1. Points indicate approximate center of 'enumeration area.'
2. For each enumeration area, collect all data within a 10km circle on cropland pixels.
3. Calculate spatial mean.
4. Calculate monthly mean or max for 12 months leading to harvest.
5. Link those data with LSMS survey data on crop yields from the same location.



Results: Country-specific and Global models



Conclusions

- Early results suggest reasonably good accuracy in all countries.
- Global (or continent-wide) model does nearly as well as country-specific model in many cases.
- Next steps:
 - Risk averse hyperparameter tuning (emphasize correctness in ‘bad’ years).
 - Test models on smaller, higher accuracy datasets with field-specific data.
 - Link predicted regional yields to health and economic outcomes from other surveys.
 - Try CNNs, giving the model pixel-level + foundation model data rather than spatially aggregated data.