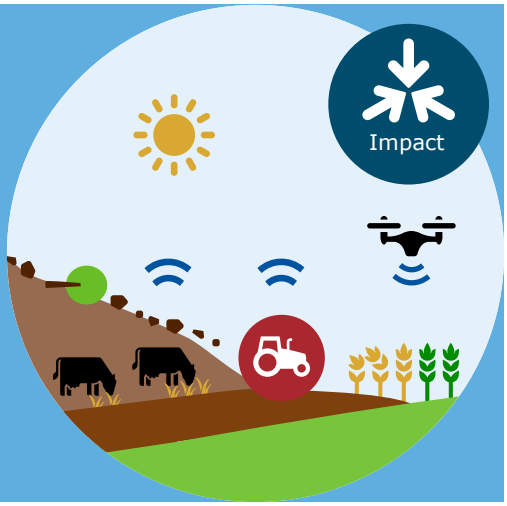
Yield loss information from machine learning for insurance payouts

A case study in Ethiopia for maize

Emerging DS/AI methods



WAGENINGEN
UNIVERSITY & RESEARCH

Data Driven Discoveries in a changing climate (D3C2)

Objective: The development of high-accuracy, high-resolution, near-real-time yield loss products through machine learning models for crop insurance pay-outs.

Activities

Our study delved into estimating maize yield in Ethiopia using machine learning models and satellite data. We leveraged MODIS Leaf Area Index (LAI) and MODIS Fraction of Photosynthetically Active Radiation (FPAR) at a 500m spatial resolution from 2006 to 2016. Recognising the importance of technological advancements, we included the 'year' variable in our analysis, capturing the discernible upward trend in yield since 2006. Zonal-level observations from Ethiopia's Central Statistical Agency (CSA) served as predictor outcomes for training our machine learning models, reporting maize yields in metric tons per hectare. We resampled predictors and outcomes to an intermediate spatial resolution of 5 km to facilitate accurate yield estimation. This strategic choice enabled the development of machine learning models while maintaining computational efficiency.

Using a crop mask from 2021 at a 10m resolution, we extracted LAI and FPAR values within each 5km x 5km grid. Our focus was exclusively on maize cultivation areas. Despite limitations in data availability, we opted for the 2021 crop mask for pragmatic reasons, ensuring data continuity and meaningful analysis within the constraints of available resources. Our machine learning modelling framework employed diverse algorithms, including gradient-boosted regression, XG Boost, and Random Forest. Hyperparameter tuning involved rigorous testing and validation on a dataset partitioned into training and test sets (2006-2014). K-fold cross-validation (k=3) and temporal out-of-sample validation for 2015 and 2016 enhanced model robustness and generalisation. Post-hyperparameter tuning, we applied the optimised models to predict high-resolution yields (500m) across representative years, ensuring spatial alignment with a crop mask derived from a 10m resolution dataset.

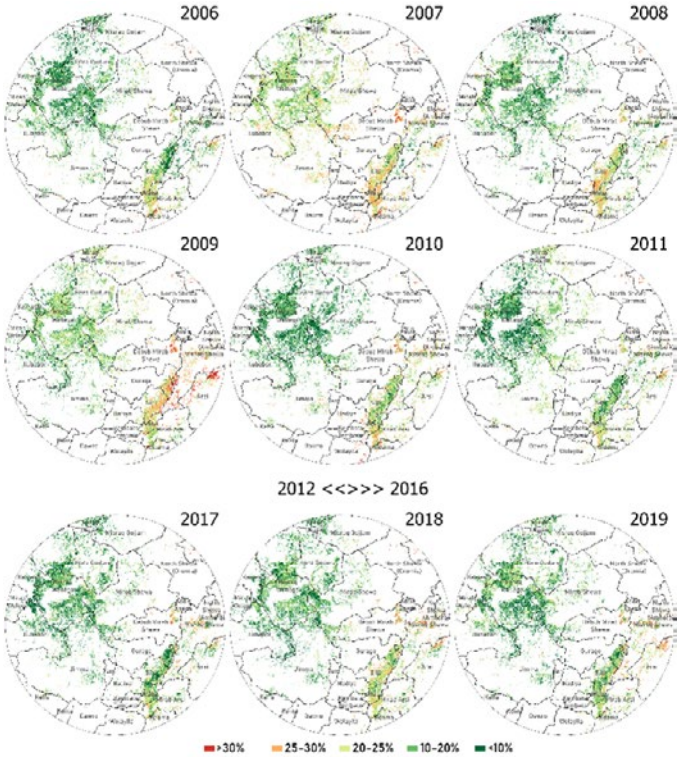


Figure Yield loss information at high resolution (~500m) from ensembled machine learning models for insurance payouts.

Achievement

I am pleased to report the successful implementation of three cutting-edge machine learning models to revolutionise high-resolution, high-accuracy, and near-real-time yield loss estimation. This pioneering approach is a superior alternative to existing insurance products, particularly valuable in regions with limited observation density, notably in the global south. This achievement was made possible through the unwavering support of the D3-C2 investment team, whose funding played a pivotal role in realising the project's objectives.

The final analysis assessed the relative importance of predictors (LAI and FPAR) in the overall model architecture. Our study resulted in a comprehensive dataset of 68,642 observations for 2006-2016 within delineated zones in Ethiopia, providing valuable insights into maize yield dynamics.

Outlook

Our recent engagement with UNDP's Insurance and Risk Financing facility has initiated a promising collaboration in sustainable finance and agricultural resilience. Key representatives like Edgar Aguilar and Encok Sing'oei discussed a project across Ethiopia, Uganda, Tanzania, Bangladesh, and India, focusing on living income and value chain resilience. Our ongoing project in Ethiopia demonstrated our commitment, attracting UNDP's interest. Simultaneously, another team member, Yuca, shared her encounter with Edgar Aguilar at an FAO workshop, where his eagerness to explore collaboration was evident. With detailed project descriptions expected in February, the groundwork has been laid for a collaborative exploration, focusing on design and pricing strategies. Yuca has volunteered to facilitate connections, showcasing our dedication to effective collaboration. This marks a significant step in our journey towards sustainable practices and resilient agriculture, and we eagerly anticipate the positive impact this collaboration could bring to the targeted countries.

In our continuous efforts to forge meaningful partnerships, we've initiated contact with the UN Economic Commission for Africa (UN-ECA) based in Ethiopia. Sonia Essobmadje, Chief Innovative Finance and Capital Markets Section at UN-ECA, expressed her interest in index insurance for de-risking and climate adaptation among smallholder farmers. We consider this as a potential opportunity for a Weather and Climate Resilience project.

Deliverables

- Research Article Outcome Report: 'Enhanced Remote Sensing and Machine Learning-based Crop Insurance Index Product: A Case Study in Ethiopia for Maize', submitted in March 2024. We are aiming to submit in the Remote Sensing of Environment Journal (https://www.sciencedirect.com/journal/remote-sensing-of-environment)

Lessons learned

In a landscape where prototype research funding is often scarce, the backing from the D3-C2 investment team has been instrumental. The collaborative efforts of the project team have been commendable, with active involvement contributing significantly to the project's success. This initiative not only advanced technological capabilities but fostered a cross-disciplinary learning environment. As a data scientist, I gained valuable insights into the economic intricacies of insurance products while non-technical team members delved into the technical nuances of the project. This project has not only delivered on its technical objectives but has become a platform for knowledge exchange within the team. The diverse perspectives have enriched our understanding, creating a model for collaborative innovation. I thank the D3-C2 investment team for their foresight and the project team for their dedication and expertise. This success marks a significant stride towards enhancing agricultural risk assessment methodologies and underscores the potential impact of innovative machine learning applications in addressing global challenges.

Contact

Dr. Ponraj Arumugam
Researcher Climate Resilience,
Wageningen University and Research
ponraj.arumugam@wur.nl

Wageningen University & Research
P.O. Box 47
6700 AB Wageningen
The netherlands
T +31 317 48 07 00
www.wur.eu