Measuring agricultural adaptation to climate change in Zambia
using satellite imagery and machine learning
(Capstone Continuation)

Proposal Authors

- Tamma Carleton
  Assistant Professor
  Bren School, University of California, Santa Barbara
  tcarleton@ucsb.edu
  415-816-7870
- Carlo Broderick
  MEDS Student
  Bren School, University of California, Santa Barbara
  209-404-4123
- Andrew Bartnik
  MEDS Student
  Bren School, University of California, Santa Barbara
  213-801-2540

Client

- Tamma Carleton
  Assistant Professor
  Bren School, University of California, Santa Barbara
  tcarleton@ucsb.edu
  415-816-7870

a. Objectives.

Food security in sub-Saharan Africa depends heavily on local agricultural productivity and is increasingly under threat from global climate change (Porter et al., 2014). For example, a combination of severe drought followed by floods in southern Zambia in 2017 and 2018 pushed 2.3 million people into food insecurity (OCHA, 2019). Farmers across Africa generally operate small-scale farms, live in poverty, and have limited access to the adaptation technologies, such as irrigation or drought-resistant seed varieties, that are being used in other regions to mitigate crop yield losses caused by extreme climate conditions. Despite the importance of agriculture and climate change adaptation for wellbeing across Africa, data limitations imply that little is known about how farmers are adapting to climate change in these resource-poor contexts. Agricultural statistics in most of sub-Saharan Africa are available only at the national level (FAOSTAT, 2022), are measured with substantial error (Lobell et al., 2019), and rarely provide information on adaptation strategies. Lack of information on farmer adaptation in the most agriculturally vulnerable region of the world substantially limits local and global policy efforts to ensure food security under accelerating climate change. This capstone project will build on our client's capstone project from 2021, CropMOSAIKS, to generate the first assessment of farmer adaptation to drought across Zambia. In 2021, the CropMOSAIKS team demonstrated that satellite imagery and machine learning could be used to predict spatial and temporal variation in maize yields across Zambia (e.g., see SM Figure 1). This was a critical first step toward understanding farmer adaptation to climate change. Here, we propose to dramatically scale up this initial proof of concept, using hundreds of thousands of newly accessed household-level training data observations with a richer set of ground truth variables that will allow us to directly quantitatively assess farmer adaptation.

b. Significance.

There are three target audiences for this work.

1. **The climate change impacts research community.** Very little is known about whether and how farmers in low-income contexts can respond to drought and other aspects of climate change (Auffhammer, 2018). Current efforts to generate empirically-based estimates of climate change impacts on agricultural production generally fail to incorporate farmer crop switching or other forms of adaptation. This research gap is particularly stark in low-income communities, where data limitations make measurement of climate change impacts and adaptation exceedingly difficult. The first output from this project – satellite imagery-based high-resolution maps of farmer adaptation activity over time – will enable future research on climate adaptation, unlocking research possibilities that were previously out of reach due to data limitations. The second output from this project – analysis and data visualizations of the relationship between climate adaptation actions and drought events will fill a research gap directly by quantifying which, if any, measurable adaptation strategies are Zambian farmers willing and able to adopt when severe drought conditions occur.
2. **Policymakers working to develop climate-smart agricultural practices in Zambia.** Adaptation planning is becoming a critical component of climate policy, both in developed and low-income countries. The outputs from this project will provide critical information and data inputs to policymakers seeking to support effective climate-smart agriculture in Zambia and beyond. For example, the U.N. FAO has a large project on climate-smart agriculture,

but no case study examples from Zambia.

3. **Scholars and practitioners in developing countries looking to leverage satellite imagery to solve social and environmental problems.** This project sits within a broader research agenda aimed at democratizing access to social and environmental monitoring using satellite imagery. The clients are jointly pursuing this agenda as part of a team that developed a new unsupervised machine learning approach to featurize satellite imagery and generate predictions of a wide range of on-the-ground conditions. This project serves as a valuable demonstration of this method, as applied in a developing country context.

c. Background.

This capstone will allow students to leverage new household-level survey data and combine it with satellite imagery and machine learning to provide a much more comprehensive picture of climate change adaptation than would be possible with ground truth data alone, given the costs of traditional data collection methods in countries like Zambia. Zambia is an ideal place to study questions at the intersection of climate and agriculture, as over 30% of its land area is dedicated to agricultural production, 55% of the population works in agriculture, and it faces accelerating drought intensity under climate change. Building off the CropMOSAIKS pipeline, students this year will have access to 30,000 household surveys per year measuring a range of variables, from crop switching to irrigation. This is in contrast to aggregate district-level statistics on a single variable (maize yield) used in CropMOSAIKS. Students will leverage the featurization process developed last year, saving time and allowing them to focus on the machine learning prediction component of the pipeline, as well as the analysis of drought effects. This year we will produce higher resolution predictions of more variables, and anticipate far higher accuracy than last year's capstone project, due to the infrastructure already in place and the dramatically improved training data obtained. Additionally, the focus this year will be on directly measuring adaptation decisions (e.g., crop switching), instead of measuring maize yield only.

d. Equity.

Low-income regions are consistently under-represented in both the production and consumption of remote sensing products. This occurs despite the fact that these regions are likely to benefit the most from such products, given the paucity of other systematically collected data (Haack and Ryerson, 2016). These inequities in access to the rich information contained within satellite imagery arise due to large barriers to entry into the remote sensing field, driven by high computational, data storage, expertise, and financial resource costs. As documented in Rolf et al., the MOSAIKS method separates users from raw imagery, lowering the computational cost by many orders of magnitude, while being simple and easy to implement. This project will both make final output data available covering a region of the globe in which data are widely known to be scarce and of low quality *and* will release featurized images that will enable and empower people in data-poor regions to make predictions themselves for new tasks that are not studied here. Through these avenues, we hope to facilitate equitable access to environmental monitoring processes and outputs.

e. Data.

This project will combine publicly available satellite imagery from Sentinel and Landsat 7, as accessed via Microsoft Planetary Computer, with farm-level survey data on crop planting, irrigation, and other adaptation decisions, provided by the clients. Drought data will come in the form of Palmer Drought Severity Index data from the Global Drought Crops Monitoring dataset. These data are obtained by our Zambian collaborator Protensia Hadunka and his PhD advisor and UCSB faculty member Kathy Baylis. They are not yet publicly available, but have been shared with our client via this Box folder.

f. Computational tools & needs.
Imagery data will be accessed and featurized (turning images into tabular data via MOSAIKS) via Microsoft Planetary Computer (PC), following the 2021 Capstone. Then ridge regression (using Python) linking adaptation actions observed in household survey data to imagery features will be performed on the tabular data on Taylor. Correlations and data visualizations using drought data will also be done on Taylor. The majority of computational burden will be carried by the Microsoft Planetary Computer (PC) which can be used by MEDS students during this project free of charge. Access to these free computational services can sometimes vary. The client has committed to providing the capstone team with additional computational services from Microsoft's paid cloud platform Azure if access to PC becomes a problem.

g. Possible approaches.
The implementation steps are as follows:

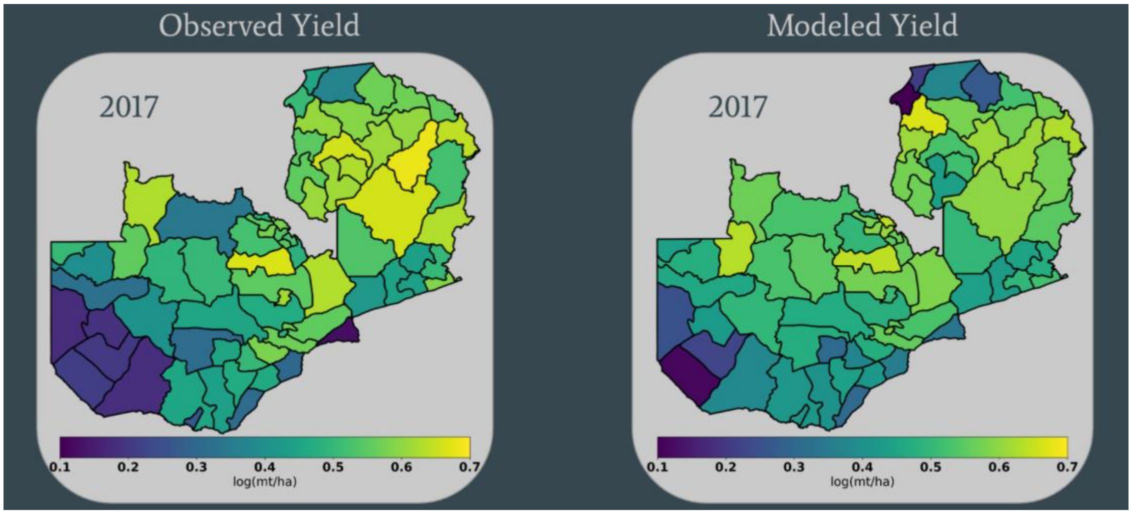
1. Use the "random convolutional features" approach outlined in and coded up by Rolf et al. to featurize 8-day Landsat 7 satellite imagery over the 30,000 farms surveyed each year in ground truth Zambian data.
2. Merge geo-located imagery features to administrative records of cropped area by crop, multi-cropping, and irrigation. Collapse to include only growing season measurements.
3. Run cross-validated ridge regressions to predict adaptation actions using the MOSAIKS features, pooling across all years.
4. Use the regression models from Step 3 alongside features for all grid cells in Zambia to produce a set of time-varying, high-resolution maps of adaptation across the country.
5. Correlate adaptation predictions to measures of drought severity and generate data visualizations of these results.
6. Work with the clients to integrate this output, and the intermediate features, into an existing public-facing API (API maintenance is not MEDS students' responsibility).

h. Deliverables.
Deliverables include:
1. A database of high-resolution (1km x 1km) maps over 15 years (2007-2022) showing predicted farmer adaptations, including crop switching, multi-cropping, and irrigation.
2. Estimates and data visualizations of the effects of drought on these adaptation actions.

This new data product and its open-source codebase will be disseminated through an existing API. Results for the impacts of drought on farmer adaptation will be released publicly as a report and will be handed off to our Zambian collaborator, Protensia Hadunka.

i. Audience.
Please see detailed information on the target audiences above in the "Significance" section.

SUPPORTING MATERIALS (not counted toward 3-page limit):



SM Figure 1: Comparison of observed yields (left panel) to predicted yields (right panel), where predictions are estimated using MOSAIKS features computed from Sentinel and Landsat imagery and the CropMOSAIKS pipeline. Observed yield is only available at the district level, so all estimates in CropMOSAIKS are performed at the district scale.

a. Citations.

Auffhammer, Maximilian. "Quantifying economic damages from climate change." *Journal of Economic Perspectives* 32, no. 4 (2018): 33-52.

Burke, Marshall, et al. “Using Satellite Imagery to Understand and Promote Sustainable Development.” Science (New York, N.Y.), vol. 371, no. 6535, Mar. **2021**, p. eabe8628. PubMed, https://doi.org/10.1126/science.abe8628.

Burke, Marshall, and David B. Lobell. “Satellite-Based Assessment of Yield Variation and Its Determinants in Smallholder African Systems.” Proceedings of the National Academy of Sciences, vol. 114, no. 9, Feb. **2017**, pp. 2189–94. pnas.org (Atypon), https://doi.org/10.1073/pnas.1616919114.

“ERA5.” ECMWF, 3 Nov. **2017**, https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5.

“FAOSTAT.” Food and Agriculture Organization of the United Nations Data Repository, https://www.fao.org/faostat/en/#data. Accessed 13 Oct. **2022**.

Haack, Barry, and Robert Ryerson. “Improving Remote Sensing Research and Education in Developing Countries: Approaches and Recommendations.” International Journal of Applied Earth

Observation and Geoinformation, vol. 45, Mar. **2016**, pp. 77–83. NASA ADS, https://doi.org/10.1016/j.jag.2015.11.003.

“Harvested Area and Yield for 175 Crops.” EarthStat, http://www.earthstat.org/harvested-area-yield-175-crops/. Accessed 13 Oct. **2022**.

Hultgren Andrew, Tamma Carleton, Michael Delgado, Diana Gergel, Michael Greenstone, Trevor Houser, Solomon Hsiang, et al. “The Impacts of Climate Change on Global Grain Production Accounting for Adaptation.” *In prep.*, (**2021**).

IPCC, “Food Security and Food Production Systems.” Climate Change 2014 – Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects: Working Group II Contribution to the IPCC Fifth Assessment Report: Volume 1: Global and Sectoral Aspects, vol. 1, Cambridge University Press, 2014, pp. 485–534. Cambridge University Press, https://doi.org/10.1017/CBO9781107415379.012.

Lobell, David B., et al. “Eyes in the Sky, Boots on the Ground: Assessing Satellite- and Ground-Based Approaches to Crop Yield Measurement and Analysis.” American Journal of Agricultural Economics, vol. 102, no. 1, **2020**, pp. 202–19. Wiley Online Library, https://doi.org/10.1093/ajae/aaz051.

“MOSAIKS.” Multi-Task Observation Using Satellite Imagery & Kitchen Sinks, https://siml.berkeley.edu/home/index/. Accessed 13 Oct. **2022**.

OCHA (United Nations Office for the Coordination of Humanitarian Affairs) (2019). Zambia: Prolonged drought increases food insecurity. https://www.unocha.org/story/zambia-prolonged-drought-increases-food-insecurity

Porter, J.R., L. Xie, A.J. Challinor, K. Cochrane, S.M. Howden, M.M. Iqbal, D.B. Lobell, and M.I. Travasso, 2014: Food security and food production systems. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L.White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 485-533.

Proctor, Jonathan. “Atmospheric Opacity Has a Nonlinear Effect on Global Crop Yields.” Nature Food, vol. 2, no. 3, 3, Mar. **2021**, pp. 166–73. www.nature.com, https://doi.org/10.1038/s43016-021-00240-w.

QUARMBY, N. A., et al. “The Use of Multi-Temporal NDVI Measurements from AVHRR Data for Crop Yield Estimation and Prediction.” International Journal of Remote Sensing, vol. 14, no. 2, Jan. **1993**, pp. 199–210. Taylor and Francis+NEJM, https://doi.org/10.1080/01431169308904332.

Rising, James, and Naresh Devineni. “Crop Switching Reduces Agricultural Losses from Climate Change in the United States by Half under RCP 8.5.” Nature Communications, vol. 11, no. 1, 1, Oct. **2020**, p. 4991. www.nature.com, https://doi.org/10.1038/s41467-020-18725-w.

Rolf, Esther, Jonathan Proctor, Tamma Carleton, Ian Bolliger, Vaishaal Shankar, Miyabi Ishihara, Benjamin Recht, and Solomon Hsiang. "A generalizable and accessible approach to machine learning with global satellite imagery." *Nature communications* 12, no. 1 (**2021**): 1-11.

“USGS.” Landsat 7 | U.S. Geological Survey, https://www.usgs.gov/landsat-missions/landsat-7?qt-science_support_page_related_con=0#qt-science_support_page_related_con. Accessed 13 Oct. **2022**.

b. Budget and justification.

This project does not require any additional funding. The only possible cost would be computational, as the students need to transform imagery into random convolutional features, which is a somewhat costly computational step (see Rolf et al., 2021 for a detailed breakdown of computation costs). All other computational steps in the analysis (i.e., geospatial merging, linear regression and prediction) do not require large computational resources (as documented in Rolf et al., 2021). This featurization step, however, can be conducted using Microsoft’s Planetary Computer, where Landsat 7 and Sentinel images are already made available.

The majority of computational burden will be carried by the Microsoft Planetary Computer (PC) which can be used by MEDS students during this project free of charge. Access to these free computational services can sometimes vary. The client has committed to providing the capstone team with additional computational services from Microsoft’s paid cloud platform Azure if access to PC becomes a problem.

c. Client letter of support.

The client in this case is also an author of this proposal. The client commits to supporting MEDS students throughout the course of the capstone project by providing: input data (as described above); a well-documented existing codebase that can be used to directly build from for this new context; expertise in both agricultural systems and the computational infrastructure of MOSAIKS, an architecture the clients co-developed; and regular troubleshooting guidance and general project support.

c-i. Funding:

If additional computational resources are required the client will provide access to the paid cloud computing platform Microsoft Azure.

c-ii. Data:

This project will combine publicly available satellite imagery from Sentinel and Landsat 7, as accessed via Microsoft Planetary Computer, with farm-level survey data on crop planting, irrigation, and other adaptation decisions, provided by the clients. Drought data will come in the form of Palmer Drought Severity Index data from the Global Drought Crops Monitoring dataset. These data are obtained by our Zambian collaborator Protensia Hadunka and his PhD advisor and UCSB faculty member Kathy Baylis. They are not yet publicly available, but have been shared with our client via this Box folder.