ENVIRONMENTAL RESEARCH
LETTERS

PERSPECTIVE • OPEN ACCESS

Considerations for AI-EO for agriculture in Sub-Saharan Africa

To cite this article: Catherine Nakalembe and Hannah Kerner 2023 *Environ. Res. Lett.* **18** 041002

ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
14 December 2022

REVISED
1 March 2023

ACCEPTED FOR PUBLICATION
15 March 2023

PUBLISHED
24 March 2023

PERSPECTIVE

Considerations for AI-EO for agriculture in Sub-Saharan Africa

Catherine Nakalembe[1,*] and Hannah Kerner[2]

[1] Department of Geographical Sciences, University of Maryland, College Park, MD 20740, United States of America
[2] School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85282, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: cnakalem@umd.edu

Keywords: Artificial intelligence, satellite Earth observations, Africa, agriculture

1. Introduction

Adapting to and mitigating climate change while addressing food insecurity are top priorities in Sub-Saharan Africa that require technologies to improve rural livelihoods with minimal environmental costs [1]. Artificial intelligence (AI) offers great promise for climate-smart solutions that improve food security outcomes. While precision agriculture is often the foremost use case for AI in agriculture (e.g. automation of farm equipment or nutrient application), precision agriculture is out of reach for most African farmers due to the required capital and infrastructure.

AI solutions using satellite Earth observations (EOs), which we call AI-EO, are more accessible in the near term. EO enables agricultural analyses and insights at global scales, and many datasets are freely available, making EO-based solutions affordable [2]. AI-EO-derived products such as crop type maps and yield estimates are necessary to forecast food production surpluses or deficits, inform trade, and aid decisions. These products can support policies that accelerate the design and adoption of climate-smart agriculture and impact farmer livelihoods by increasing access to actionable early warning, risk financing or insurance [3], farm inputs, markets, and cost-reducing interventions [2, 4].

Despite their promise, AI-EO solutions for agriculture in Africa are still limited. Most techniques are not generalizable across heterogeneous landscapes. In this paper, we describe the principal sub-fields of research in AI-EO for agriculture in Africa and discuss examples and limitations of existing work. We also propose ten considerations for future work to help increase the impact of AI-EO research in Africa.

2. Key AI-EO applications

In this section, we discuss AI-EO applications for agriculture and current limitations that need to be addressed in future work in Africa.

2.1. Cropland and crop type mapping

Cropland maps indicate where crops are growing spatially (figure 2(a)), while crop type maps indicate the specific crop is growing in each spatial unit (e.g. maize). In EO-based yield or conditions modeling, these maps are required to restrict the analysis to pixels that include cropland or a specific crop type. Crop type maps should be updated seasonally because farmers may change crops grown in a particular field [2].

Cropland mapping involves classifying spatial units (e.g. pixels with a specific spatial resolution) as containing cropland. Crop type mapping is usually framed as a multi-class classification but can be framed a binary classification where the positive class is the crop type of interest, and the negative class includes all other crop types and non-crop classes. The resulting map's spatial resolution depends on the satellite data inputs (e.g. 10 m/pixel if using Sentinel-2 [5]). Most ML models used for cropland and crop type classification are tree-based classifiers such as decision trees and random forests. Deep learning approaches, especially recurrent neural networks that learn important crop-specific growth patterns in time series data, have gained popularity in recent years and are the current state of the art [6].

While the accuracy of public cropland maps is generally high in developed regions, prior studies have shown significantly lower accuracy in Africa. The errors reported for crop area estimates in [8] were lowest in Africa, and [9] found that user accuracy was as low as 17% in eastern Africa (e.g. figure 1).

This is partly because farms in Sub-Saharan Africa are predominantly smallholder farms which can be difficult to detect accurately without very high-resolution satellite datasets (which are not freely available). Another limiting factor is the lack of publicly available labeled African agriculture datasets.

While some recent work has proposed approaches for improving crop type classification results with limited labeled data using transfer learning and

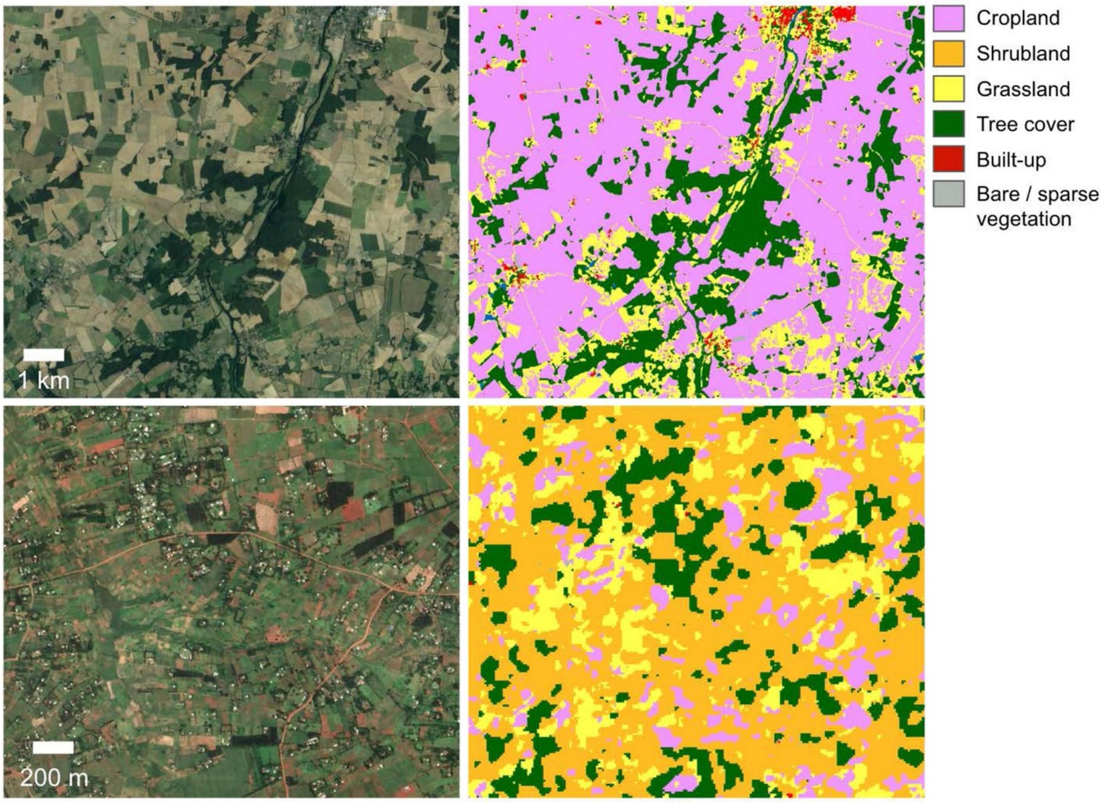

Figure 1. Example region in ESA WorldCover land cover map [7] in western France (top) and western Kenya (bottom). Both regions show dense crop fields, but most fields are confused for shrubland in Kenya but accurately mapped as cropland in France. Reproduced from [7].

few-shot learning methods like meta-learning [6], most prior work relies on large labeled datasets that are not currently available for Africa. Even if a model can be trained efficiently using a small number of labeled samples, robust evaluation of the resulting crop type map (which involves dense inference over an entire region) still requires many labels. While land cover labels (including cropland) can usually be annotated using photo-interpretation of high-resolution satellite images, crop type must be determined through ground-truth observation. Intercropping is a common practice in Africa that presents an additional challenge for AI-EO methods. While it is possible to physically identify multiple crops growing in the field from in-situ observation, it is difficult to disaggregate them at the same pixel location in satellite data. Current methods typically treat intercropped fields as a single aggregate class or use the label for the crop assumed to be the dominant crop.

2.2. Yield estimation

Models that can support strategies for increasing yields are essential for climate-smart agriculture in Africa, where increases in crop production have largely been driven by increases in the cultivated area, not yields. Yield estimation is a regression task in which models estimate the crop harvested per unit area (e.g. kg per hectare). Most yield estimation work is done at the regional or national scale, with less work at the field scale (figure 2(b)). Yield estimation methods may also be evaluated by how early the end-of-season yield can be accurately forecast in the growing season. Most ML approaches for yield estimation use tree-based methods such as random forests and XGBoost. Recent work has presented various deep learning solutions, including Deep Gaussian Processes, graph neural networks, and recurrent neural networks. Other approaches leverage crop simulation and statistical regression models. However, few studies have focused on yield estimation for Africa [10, 11].

2.3. Field boundary delineation

Field boundaries guide sampling and area estimation methods to provide statistically sound estimates of cropped area and are helpful for sub-field assessments of inputs, crop performance, and production (figure 2(c)) [12]. Some studies have used instance segmentation methods from computer vision like Mask R-CNN [13], but most recent work has used semantic segmentation methods like U-Nets followed by post-processing to isolate individual field instances [14].

A major roadblock to field boundary delineation in Africa is the insufficient spatial resolution of publicly available satellite datasets. Smallholder fields are often smaller than 1 ha (100 m × 100 m), thus,

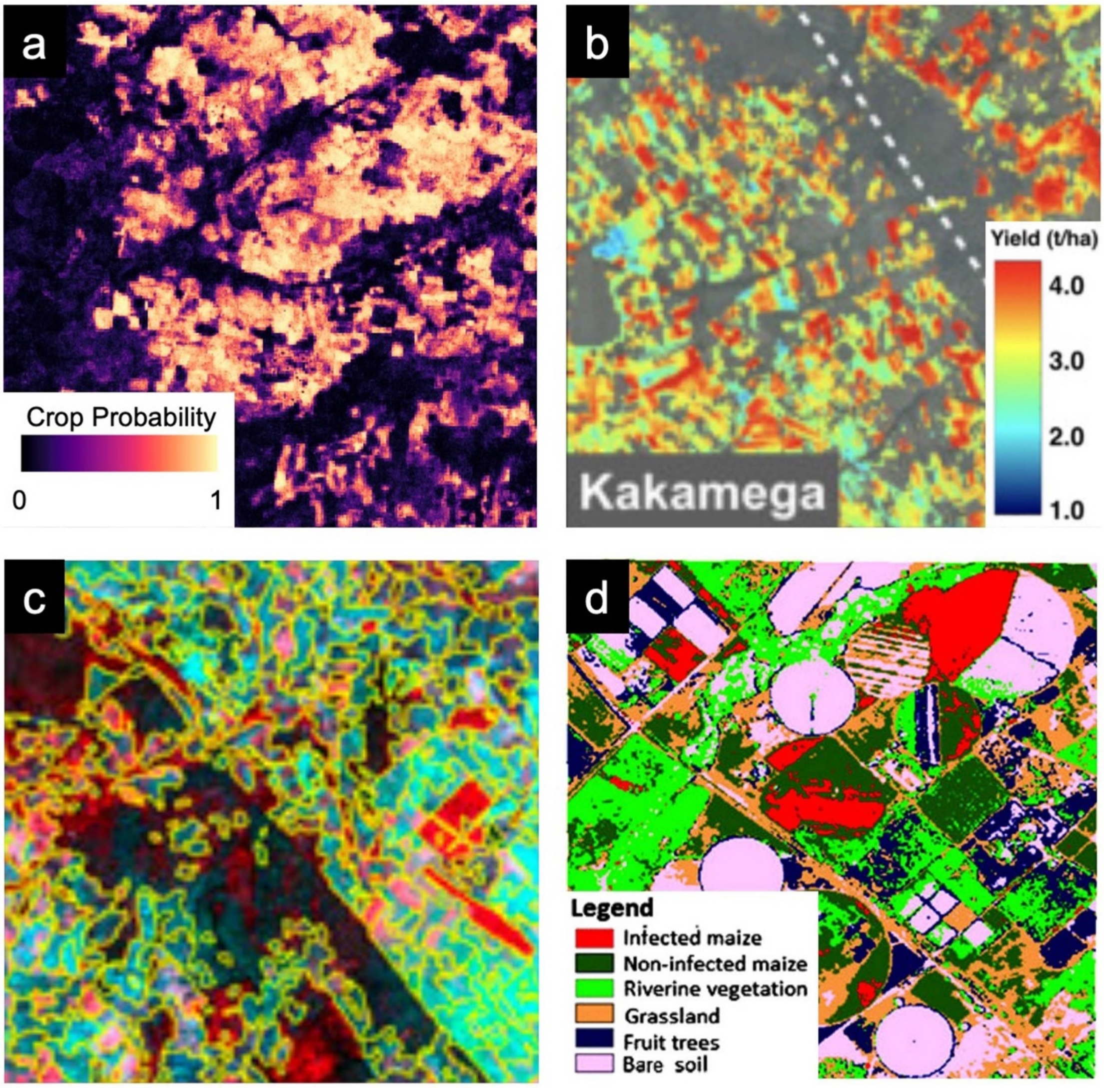

Figure 2. Example outputs from AI-EO models for (a) cropland classification in Togo. Reproduced from [5]. CC BY 4.0. (b) yield estimation in Kenya. Reprinted from [11], Copyright 2019, with permission from Elsevier. (c) field boundary delineation in Ghana. Reproduced from [15]. CC BY 4.0. and (d) pest detection Reproduced with permission from [16].

delineation requires very high-resolution datasets currently only available commercially [14]. Another challenge is that few public datasets provide field boundary labels in Africa. Prior work has proposed solutions using active learning [15] and combining transfer learning with weak supervision [14].

2.4. Pest, disease, and anomaly detection

There has been some research on pest/disease detection using EO data, for example, but most studies use high-resolution datasets that are not publicly or globally available and do not leverage modern AI techniques. Techniques leveraging AI and EO can be useful for detecting pest and disease impacts over large areas during the growing season to minimize crop damage (figure 2(c)) [16]. These approaches could also detect other in-field anomalies, such as nutrient deficiencies or weeds. However, few studies have used AI-EO to detect anomalies over large areas, partly due to limited access to high-resolution satellite datasets and ground-truth labels needed to train and evaluate AI models [17]. Current AI techniques for crop disease surveillance mostly focus on in-situ plant disease diagnosis using cell phone images or ground-based robots [18].

3. Considerations

In this section, we propose ten considerations that should be incorporated into future work to increase the positive impact of AI-EO research in Africa.

(i) **Interdisciplinary teams are a requirement.** Developing robust, practical, and contextually relevant AI-EO methods for agriculture in Africa requires interdisciplinary teams, including experts in AI, agriculture/agronomy, remote sensing, climate science, soil science, and local and regional practices.

(ii) **Consider the resource context.** The intended stakeholders' resource context (e.g. cost, energy, or internet bandwidth) must be considered at all stages of the research process, including algorithm selection/design and dissemination of results. For example, internet costs may be prohibitive for running specific AI systems (1 GB of data costs an average of 4.3% of monthly net income in Africa). These constraints are rarely considered in AI research due to a widespread belief that AI innovations will eventually become available to individuals in resource-poor settings rather than welcoming these requirements as drivers for AI innovation [19].

(iii) **Prioritize methods for limited labeled data.** Prioritizing research on AI-EO methods that optimize performance with limited labeled data is necessary for addressing the labeled data bottleneck for agriculture in Africa. The community should also prioritize more community-wide efforts to create publicly available labeled datasets for agriculture in Africa.

(iv) **Methods should be transparent and reproducible.** There is a need for greater transparency and accountability in AI-EO, including proper documentation and open methods with sufficient details that enable researchers across fields to trace and replicate prior work and compare methods to directly track progress in the field.

(v) **Work with stakeholders from the beginning.** Stakeholders and end-users should be consulted early in the research and development to ensure researchers work on problems that solve real end-user needs. Stakeholders can help inform the choice of metrics for evaluating AI-EO methods representative of real-world performance rather than defaulting to commonly used metrics in AI.

(vi) **Decolonize research methods and practices.** Most discussions of equity and ethics in AI and EO have traditionally focused on the diversity and inclusion of under-represented groups. Decolonizing current practices in research should be an added focus of such efforts. Researchers need to recognize the harms of colonialism, avoid parachute science, and conduct research in ways that do not perpetuate harm or reinforce negative power structures [20].

(vii) **Form meaningful partnerships with local institutions.** The majority of AI research today is being conducted in the US, Europe, and China, making international partnerships between institutions in these regions and Africa a critical component for the success of agriculture applications in Africa. International partnerships could work toward creating an enabling environment for AI in Africa that provides resources for capacity building, recruiting talent, infrastructure for data and computing, and inclusive innovations. These partnerships should go beyond data collection to ensure the benefits of AI-EO innovations are realized, sustained, and extended by local stakeholders.

(viii) **Institutionalized investments.** For the investments above to be sustainable, there is a need for investments in research to be institutionalized by enduring organizations focused on the public good such as universities or government. Core government agencies should build, invest in, and embrace AI-EO solutions beyond human resources to include infrastructure for data collection, analysis, scientific research and collaboration, and communication with broad stakeholders [4].

(ix) **Open access to high-resolution imagery.** Access to high-resolution satellite datasets is a significant roadblock to AI-EO application development for smallholder agriculture. Community and cross-institution efforts to purchase data from commercial providers like Planet or Maxar for agriculture-related research in Africa are needed. For example, the Norway International Climate and Forests Initiative Satellite Data Program made high-resolution Planet basemaps freely available for projects related to forests in the world's tropics. A similar initiative should be implemented for agriculture.

(x) **Limitations of AI-EO solutions should be assessed and communicated.** While there is vast potential and excitement for AI-EO technologies for agriculture in Africa, researchers need to assess and communicate the limitations of these solutions. Over-promising and under-delivering on the capabilities of AI risks disillusionment and loss of interest for stakeholders and funders of AI-EO solutions. While published research shows possible benefits, practical demonstrations that realize the promise of these solutions for stakeholders are still limited. Furthermore, it is critical to perform rigorous evaluations to thoroughly assess and communicate the strengths and limitations of AI-EO models (indeed, any models) used to inform decisions or policies that affect people's livelihoods and outcomes.

4. Opportunities for future work

There are common topics for future research that would substantially impact many applications. To address the lack of labeled data, the research could focus on learning efficiently from limited labeled data and sparse, partial, or noisy labels. More efficient and scalable approaches to data collection are needed to create datasets that represent real-world conditions and progress for agriculture in Africa. Future work could develop approaches for intercropping in crop type classification and yield prediction. Techniques for multi-fidelity data fusion are also required to combine satellite data sources with variable temporal/spatial resolution and quality. Finally, there is a need for more equitable programs that fund researchers and organizations based in Africa to develop the datasets, models, and other capabilities required to advance AI-EO for agriculture in Africa.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

Catherine Nakalembe https://orcid.org/0000-0002-2213-593X
Hannah Kerner https://orcid.org/0000-0002-3259-7759

References

[1] Sasson A 2012 Food security for Africa: an urgent global challenge *Agric. Food Secur.* **1** 1–16
[2] Nakalembe C *et al* 2021 A review of satellite-based global agricultural monitoring systems available for Africa *Glob. Food Secur.* **29** 100543
[3] Benami E, Jin Z, Carter M R, Ghosh A, Hijmans R J, Hobbs A, Kenduiywo B and Lobell D B 2021 Uniting remote sensing, crop modelling and economics for agricultural risk management *Nat. Rev. Earth Environ.* **2** 140–59
[4] Nakalembe C 2020 Urgent and critical need for Sub-Saharan African countries to invest in earth observation-based agricultural early warning and monitoring systems *Environ. Res. Lett.* **15** 121002
[5] Kerner H, Tseng G, Becker-Reshef I, Nakalembe C, Barker B, Munshell B, Paliyam M and Hosseini M 2020 Rapid response crop maps in data sparse regions *ACM SIGKDD Conf. on Data Mining and Knowledge Discovery Workshops*
[6] Tseng G, Kerner H, Nakalembe C and Becker-Reshef I 2021 Learning to predict crop type from heterogeneous sparse labels using meta-learning *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition*
[7] Zanaga D *et al* 2021 ESA WorldCover 10 m 2020 v100 (version v100)[data set] (Zenodo)
[8] Potapov P, Turubanova S, Hansen M C, Tyukavina A, Zalles V, Khan A, Song X-P, Pickens A, Shen Q and Cortez J 2022 Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century *Nat. Food* **3** 19–28
[9] Leroux L, Jolivot A, Bégué A, Seen D L and Zoungrana B 2014 How reliable is the modis land cover product for crop mapping sub-saharan agricultural landscapes? *Remote Sens.* **6** 8541–64
[10] Van Klompenburg T, Kassahun A and Catal C 2020 Crop yield prediction using machine learning: a systematic literature review *Comput. Electron. Agric.* **177** 105709
[11] Jin Z, Azzari G, You C, Di Tommaso S, Aston S, Burke M and Lobell D B 2019 Smallholder maize area and yield mapping at national scales with Google Earth Engine *Remote Sens. Environ.* **228** 115–28
[12] Masoud K M, Persello C and Tolpekin V A 2019 Delineation of agricultural field boundaries from Sentinel-2 images using a novel super-resolution contour detector based on fully convolutional networks *Remote Sens.* **12** 59
[13] Meyer L, Lemarchand F and Sidiropoulos P 2020 A deep learning architecture for batch-mode fully automated field boundary detection *The Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* vol 43 pp 1009–16
[14] Wang S, Waldner F and Lobell D B 2022 Unlocking large-scale crop field delineation in smallholder farming systems with transfer learning and weak supervision *Remote Sens.* **14** 5738
[15] Estes L D *et al* 2022 High resolution, annual maps of field boundaries for smallholder-dominated croplands at national scales *Front. Artif. Intell.* **4** 744863
[16] El-Ghany N M, Abd El-Aziz S E and Marei S S 2020 A review: application of remote sensing as a promising strategy for insect pests and diseases management *Environ. Sci. Pollut. Res.* **27** 33503–15
[17] Gold K M 2021 Plant disease sensing: studying plant-pathogen interactions at scale *Msystems* **6** e01228–21
[18] Mrisho L M, Mbilinyi N A, Ndalahwa M, Ramcharan A M, Kehs A K, McCloskey P C, Murithi H, Hughes D P and Legg J P 2020 Accuracy of a smartphone-based object detection model, PlantVillage Nuru, in identifying the foliar symptoms of the viral diseases of cassava-CMD and CBSD *Front. Plant Sci.* **11** 590889
[19] De-Arteaga M, Herlands W, Neill D B and Dubrawski A 2018 Machine learning for the developing world *ACM Trans. Manage. Inf. Syst.* **9** 1–14
[20] Roldan-Hernandez L, Boehm A B and Mihelcic J R 2020 Parachute environmental science and engineering *Environ. Sci. Technol.* **54** 14773–4