A Landscape Analysis Of Remote Sensing Providers Offering Solutions Tailored To Smallholder Farmers In Sub-Saharan Africa

Executive Summary

With the rapid advancement of low cost earth observation (EO) technologies, the remote sensing sector is actively building new solutions that could support smallholder-facing organizations to optimize their climate change adaptation and mitigation strategies - including precision agriculture, crop insurance, soil health improvements, carbon verification and crop health monitoring.

This report, published by One Acre Fund and the Syngenta Foundation, explores the remote sensing ecosystem to learn more about the various providers, and evaluate their feasibility for use within the smallholder farmer context in sub-Saharan Africa. Specifically, the research focuses on the use of remote sensing technologies across the following key verticals: **1)** Agroforestry and aboveground biomass carbon monitoring; **2)** Flood detection; **3)** Soil properties, including soil organic carbon (SOC); **4)** Yield measurement; **5)** Field Delineation; **6)** Crop health.

For each of these applications, the report summarizes the technical approaches that are used, with a focus on what is currently considered 'state of the art'. The report also highlights some of the top providers in each vertical. Ultimately, this report is intended to serve as a learning resource for smallholder farmer-facing organizations, which might inform their future collaborations with the fast-developing remote sensing ecosystem. .

- *<u>Report Author:</u> George Azzari, Innovation Advisor, Atlas AI*

- *<u>Special Thanks</u> to Syngenta Foundation for providing financial support for this project, along with their expertise and guidance throughout the research process.*

- ***For inquiries about this report,** please contact:*
    - *John Mundy, One Acre Fund Senior Partnerships Manager, Carbon and Technology (john.mundy@oneacrefund.org)*

1 Landscape Overview

1.1 Types of Earth Observation Stakeholders

Fundamentally, the landscape of Earth Observation (EO) stakeholders can be divided into four main categories:

- **Satellite Imagery Providers:** in addition to the main public providers of imagery (NASA and ESA), there are only a handful of commercial entities - although the number of startups focusing on new types of satellites has been growing at a fast pace. The leading players in this field are Maxar, Airbus, Planet, and Satellogic. Smaller, venture-backed companies focusing on SAR data[1] are also surfacing in the market (e.g., Capella and Ursa).

- **Earth Observation (EO) Platform Providers:** warehousing, preprocessing, and running analysis on satellite imagery are complex and resource-intensive tasks. A few providers have been focusing on making things easier for users of all levels of expertise by building UI and API solutions (Google Earth Engine, Descartes Labs, and ESRI, among others).

- **Analytics and Recommendations Providers:** Often composed of highly interdisciplinary teams, these providers combine domain knowledge, remote sensing science, and machine learning to extract higher-level information from EOs (e.g., estimation of crop yield or segmentation of agricultural fields). A few of these organizations go a step further by building recommendation systems on top of their analytics (e.g. irrigation and inputs regimes). A key aspect of this class of providers is that they strive for automated, Software as a Service (SaaS-type) solutions as opposed to more bespoke solutions offered by Consulting Providers. This is likely the class of EO stakeholders that has shown the fastest growth, especially in the agricultural sector.

- **Consulting Providers:** Providers in this class have similar technical capabilities to Analytics Providers; however, they mainly focus on end-to-end bespoke solutions for their customers rather than automation and software. Their services may also include project management and strategy expertise. Examples of providers in this category are Lobelia Earth and Vito. To some extent, Academia and other research organizations could be considered similar to consulting providers; their work can play a crucial role in

[1] Synthetic Aperture Radar, or SAR, is a type of active data collection where a sensor produces its own energy and then records the amount of that energy reflected back after interacting with the Earth. [Source: NASA]

exploring new approaches but they may not be able to provide custom solutions or end-to-end products.

Often, organizations position themselves at the intersection of two or more of these categories, with one primary function complemented by secondary ones. For example, almost all satellite imagery providers offer analytics capabilities and products, and it is not unusual for platform providers or analytics providers to offer technical consulting services.

1.2 Summary Table

Below is a summary of the applications offered and the stakeholders' suitability. The color scale here represents the potential suitability of each organization’s product/offering for the six target applications - not the final recommendation.

✓	Offered
✗	Not offered
✓/✗	Offered/Not offered, requires additional follow up with vendor to clarify capabilities

Organization	ABG Carbon	Soil Carbon	Floods	Field Delin.	Yield	Pest / Disease
6th Grain	✗	✓	✗	✓	✓	✓
Agribora	✗	✓	✗	✗	✓	✗
Agripredict	✗	✗	✗	✗	✗	✓
Agritask	✗	✗	✗	✗	✗	✓
Airbus	✓	✓	✗	✓	✓	✓
Assimila	✗	✗	✗	✗	✓	✓
Boomitra	✗	✓	✗	✗	✗	✗
Cloud to Street	✗	✗	✓	✗	✗	✗
CropIn	✗	✗	✗	✗	✓	✓
CropNuts	✗	✓	✗	✗	✗	✗
Descartes Labs	✗	✗	✗	✗	✗	✗

Organization	ABG Carbon	Soil Carbon	Floods	Field Delin.	Yield	Pest / Disease
Digifarm	✓	✗	✗	✓	✗	✗
EarthDaily	✗	✗	✗	✗	✓	✗
GeoGecko	?	?	?	✗	✓	?
Habiterre	✗	✓	✗	✗	✓	✗
Ignitia	✗	✗	✓	✗	✗	✗
Land and Carbon Lab (WRI)	✓	✗	✗	✗	✗	✗
Lobelia Earth	✓	✗	✓	✗	✗	✓
Measure.io	✗	✗	✗	✗	✗	✓
NASA HARVEST	✗	✗	✗	✗	✓	✗
NCX	✓	✗	✗	✗	✗	✗
OneSoil	✗	✗	✗	✓	✓	✗
Planet	✓	✗	✗	✓	✗	✗
QED	✗	✗	✗	✓	✗	✗
Radiant Earth Foundation	✗	✗	✗	✗	✗	✗
Regrow	✓	✓	✗	✓	✓	✓
Resilience Constellation	✓	✗	✗	✗	✗	✗
Satellogic	✗	✗	✗	✗	✗	✗
SatSure Sparta	✗	✗	✓	✗	✓	✗
Tomorrow.io	✗	✗	✓	✗	✗	✓
University of Copenhagen/NASA Goddard	✓	✗	✗	✗	✗	✗
Vito	✗	✗	✗	✗	✗	✗

Note: Information in this table is based on interviews (where possible) and/or information that was publicly available online. The research was conducted from April - May 2022, and is subject to change.

1.3 Automatic Field Delineation

1.3.1 Technical Background and State of the Art

The ability to automatically recognize and extract the footprint of individual objects (such as buildings or roads) or other semantic classes (such as land-use or land-cover types) is known as image segmentation. This is likely the most sought-after application of EO across the board; from defense and humanitarian missions to agricultural and socio-economic studies, the EO community has been trying to detect the boundaries of all sorts of things for decades[2]. In the last five years, a lot of progress has been made toward this goal thanks to the convergence of a couple of factors. First, natural image segmentation became a key component in many consumer applications (e.g. self-driving cars, medical image diagnostic, and facial recognition security); this triggered a downpour of capital into research and development and fueled significant technical advances. Convolutional Neural Networks (CNNs) and other deep neural network architectures emerged as the dominant tool for image segmentation applications and became widespread in many sectors[3]. This includes the EO world, where older edge-detection and classification algorithms have increasingly been replaced by CNNs[4,5]. The impact of these techniques in EO-based applications has been significant; even though CNNs have not always increased the accuracy of the outputs compared to more traditional approaches, they have improved the generalizability of the models and have removed the need to pre-process images with specialized filters to extract predictive features. Second, the spatial resolution and revisit time of earth-observing satellites have increased significantly, both from public and commercial providers. Smaller objects and finer-scale spatiotemporal variability have become more detectable, opening up more applications with real-life implications. These advances have also benefited the automatic delineation of agricultural field boundaries. More frequent imagery allows for capturing the high temporal variability typical of agricultural systems and accounting for the occurrence of clouds in coincidence with the growing season in rainfed fields (the vast majority of them). Higher resolutions enabled models to resolve smaller fields with higher accuracy. Unfortunately, publicly available imagery tops at a resolution of 10 m; this is still not enough to segment smallholder landscapes, and expensive, sub-meter imagery is necessary for accurate results[23]. Deep neural architectures provide better generalization and context awareness that help with the high variability in the shape and conditions of fields in these

[2] E.g. Hoeser, T., Bachofer, F., & Kuenzer, C. (2020). Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications. *Remote Sensing*

[3] E.g. Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J., (2021). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Transactions on Neural Networks and Learning Systems.

[4] E.g. Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing

[5] E.g. Hoeser, T., & Kuenzer, C. (2020). Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends. *Remote Sensing*.

landscapes[6]. Several studies have segmented agricultural fields with accuracies above 80%[7,8], enough to push a few stakeholders to build commercial automated SaaS platforms for this.

1.3.2 Top Stakeholders

Similarly to the case of ABGC, the best providers of accurate field delineation will have excellent computer vision (CV) skills combined with access to high-resolution imagery (1 m and below). Because field boundaries can change from season to season and be confused by intercropping and other factors, access to ground validation data plays an important role in this case. Finally, stakeholders capable of incorporating crop phenology and crop classification into the stack will likely have a competitive advantage. We reviewed 10 providers in this domain. **Airbus** stood out among all because of the quality of the output, likely due to the combination of good CV technique and access to imagery from the Pleiades constellation (50 cm resolution). Interestingly, they also provide a full suite of field-level analytics that include yield estimates. A downside is that this was only the result of a small pilot in Kenya, although they are currently looking for partners to extend it (see here for their deck on Project Kilimo). Given the high price of Airbus imagery, questions remain about the scalability of the finished product within reasonable budgets. **DigiFarm** *(Norwegian ag-tech startup)* and **OneSoil** also showed impressive results, both packaged as SaaS platforms with APIs and extended geographical variability (here is a live demo for DigiFarm). Neither has produced datasets for SSA smallholders yet. OneSoil relies completely on Sentinel imagery (10 m resolution) and, at this resolution, their product is not likely to perform well on small fields. On the other hand, Digifarm leverages super-resolved Sentinel imagery enhanced to 1 m, which gives them a good chance to resolve fields smaller than 1 acre. This approach consists in training a Convolutional Neural Network to learn the relationship between a fine-resolution class of imagery (let's call it A) and an overlapping class of coarser-resolution imagery (let's call it B)[9,10]. The outcome is a model capable of generating an accurate A-like interpretation of an image from B. This approach is particularly useful in EOs, as high-resolution imagery is often hard to come by for large regions and with high frequency; a relatively small number of Airbus images (0.5 m) acquired over sparse locations and times, could be used to enhance the resolution of Sentinel images (10 m) acquired across a

---

[6] Persello, C., Tolpekin, V. A., Bergado, J. R., & de By, R. A. (2019). Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. *Remote Sensing of Environment*, *231*, 111253.

[7] E.g. Waldner, F., & Diakogiannis, F. I. (2020). Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sensing of Environment*, *245*, 111741.

[8] Wang, S., Waldner, F., & Lobell, D. B. (2022). *Unlocking large-scale crop field delineation in smallholder farming systems with transfer learning and weak supervision*.

[9] E.g. "PlanetScope Radiometric Normalization and Sentinel-2 Super-Resolution (2.5 m): A Straightforward Spectral-Spatial Fusion of Multi-Satellite Multi-Sensor Images Using ResidualConvolutional Neural Networks". Accessed 22 Apr. 2022.

[10] E.g. "Super-Resolution of Sentinel-2 Images UsingConvolutional Neural Networks and Real GroundTruth Data". Accessed 22 Apr. 2022.

whole country several times a year. This solution may provide enough accuracy for smallholder fields while also reducing the cost of the inputs. Although the performance of Digifarm's models still has to be verified, they seem to have positioned themselves in an ideal accuracy-vs-cost spot. **Planet**'s recent acquisition of Vandersat signals a renewed interest in the agricultural domain; as for Airbus, the unlimited access to high-resolution imagery and the internal computer vision capabilities make this provider a good candidate for a partnership in this domain. Unfortunately, Planet has not produced a field boundary dataset yet, although they are planning to release a public access one in the near future.

1.4 Above-ground Biomass Carbon Estimation

1.4.1 Technical Background and State of the Art

The core idea behind monitoring AGBC from EO is to measure Above Ground Biomass (AGB) via known proxies and then convert the total biomass into carbon mass[11]. Some approaches directly measure tree morphological parameters (e.g., crown projected area, tree height, etc.), which can then be converted into AGB via empirical allometric equations[12]. This class of methods has to rely on either LIDAR measurements (for height)[13] or instance segmentation of crowns projected area (CPA) based on high-resolution imagery[14] (see also Section 1.8.1 for more information on segmentation). Other approaches use spectral signatures (from optical imagery) or backscatter patterns (from SAR imagery) to estimate canopy density variables (e.g., Leaf Area Index or Fraction of Vegetation Cover) that are empirically related to AGB[15]. In other cases, optical and SAR imagery is used to classify trees against the surrounding environment and produce a pixel-level tree/no-tree binary output and roughly convert it to biomass. While methods in the second and third categories may work well for large-scale applications within dense ecosystems (e.g., dense tropical forests), methods relying on segmentation of individual tree CPA lead to more accurate carbon estimates, especially in sparse landscapes such as savannas[16]. Estimated CPA can be converted into Diameter at Breast Height (DBH) via empirical calibrations and used with

---

[11] Spawn, S.A., Sullivan, C.C., Lark, T.J. et al. Harmonized global maps of above and belowground biomass carbon density in the year 2010. Sci Data 7, 112 (2020)

[12] Ahmad, A., Gilani, H., & Ahmad, S. R. (2021). Forest Aboveground Biomass Estimation and Mapping through High-Resolution Optical Satellite Imagery—A Literature Review. *Forests*, *12*(7), 914.

[13] E.g. Patenaude et al. Quantifying forest above ground carbon content using LiDAR remote sensing, Remote Sensing of Environment, 2004.

[14] E.g. Brandt, M., Tucker, C.J., Kariryaa, A. et al. An unexpectedly large count of trees in the West African Sahara and Sahel. Nature 587, 78–82 (2020)

[15] E.g. Forkur et al. Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets - A case study, Remote Sensing of Environment, 2020.

[16] "The Contribution of Trees Outside of Forests to Landscape ... - MDPI." 28 Nov. 2021, https://www.mdpi.com/1999-4907/12/12/1652/pdf. Accessed 6 Apr. 2022.

standard allometric equations, avoiding the need for new biomass measurements[9], which are invasive and time-consuming. One challenge with this approach is that such relationships are specific to tree species and growth stages. While tree species could be identified via classification[17], the growth stage may be more difficult to estimate. Furthermore, tree crown pruning and other management practices may create noise in the CPA/DBH conversion and ultimately introduce errors in the AGB estimates. The errors related to allometric differences across species, sizes, and management practices can be estimated and ultimately evaluated against the benefit of scalability and cost reduction that remote measurements bring when compared to field work.

1.4.2 Top Stakeholders

The most scalable and accurate approach to remote ABG estimates in these conditions seems to be via the conversion to biomass of CPA measurements from high-resolution imagery segmentation. This process faces two challenges involving different expertise: one is the accurate segmentation of individual tree crowns, and the second is the conversion of CPA values into ABG via allometric equations while accounting for variability in tree species, sizes, management, etc. The ideal partner would have both excellent computer vision skills and the ability to design and execute ground experiments to account for allometric variability issues. Because of One Acre Fund's existing experience with ground-based forestry operations, we will assume that the second aspect could be addressed internally, and we will also consider partners focusing exclusively on computer vision. Moreover, we will consider having access to high-resolution imagery (1m or below) - or the ability to deploy it as needed - a significant advantage as it would vastly improve the segmentation of newly planted trees. Eight organizations were reviewed in this domain: the University of Copenhagen, Lobelia Earth, Airbus, Planet, Satellogic, Descartes Lab, NCX, Resilience Constellation, the, and WRI's Land and Carbon Lab. **The University of Copenhagen** stood out for its project on the accurate segmentation of tree crowns at 0.5m resolution across all SSA (see Figure 2). The resulting product is freely available along with the source code[18,19]; however, it wouldn't be useful operationally in this context because the authors of the study combined together images spanning 15 years and therefore the model output is not necessarily representative of today's conditions. Nevertheless, this study presents valid proof of concept of what would be achievable with current images acquired on a regular basis over One Acre Fund's AOIs. In fact, **Lobelia Earth** started collaborating with the University of Copenhagen to actively monitor tree carbon and health in Niger; their stack includes CPAs segmented from Airbus' high-res imagery, a ground campaign for allometric measurements, and soil moisture and greenness estimates for tree health. Interestingly, carbon sequestration is tracked with blockchain, and tokens can be exchanged for money via a mobile

---

[17] Wan, H., Tang, Y., Jing, L., Li, H., Qiu, F., & Wu, W. (2021). Tree Species Classification of Forest Stands Using Multisource Remote Sensing Data. *Remote Sensing*, *13*(1), 144.
[18] Data product available here: https://doi.org/10.3334/ORNLDAAC/1832
[19] Source code available here: https://doi.org/10.5281/zenodo.3978185

phone app. Both **Airbus** and **Planet** have significant in-house computer vision capabilities in addition to the key advantage of being providers of sub-meter satellite imagery. Although neither of them provides an out-of-the-box CPA product, this is not necessarily a deal-breaker; segmentation models based on annotated images are transferable to different tasks (access to high-resolution imagery, on the other hand, is a major obstacle for most). Furthermore, Planet has recently acquired Vandersat as part of their agricultural team. Vandersat has developed a Sentinel 1 and Sentinel 2 global biomass product; while the product's 10m resolution won't cut One Acre Fund's requirements, Planet has plans to integrate its imagery to produce a high-resolution version. **WRI** has developed a global dataset of tree density that captures tree presence even in sparse landscapes. The dataset is based on a solid ML approach but it still remains too coarse for small trees (10 m resolution) and does not have a direct way to convert to carbon. The rest of the organizations did not come close to satisfying the requirements for this application.

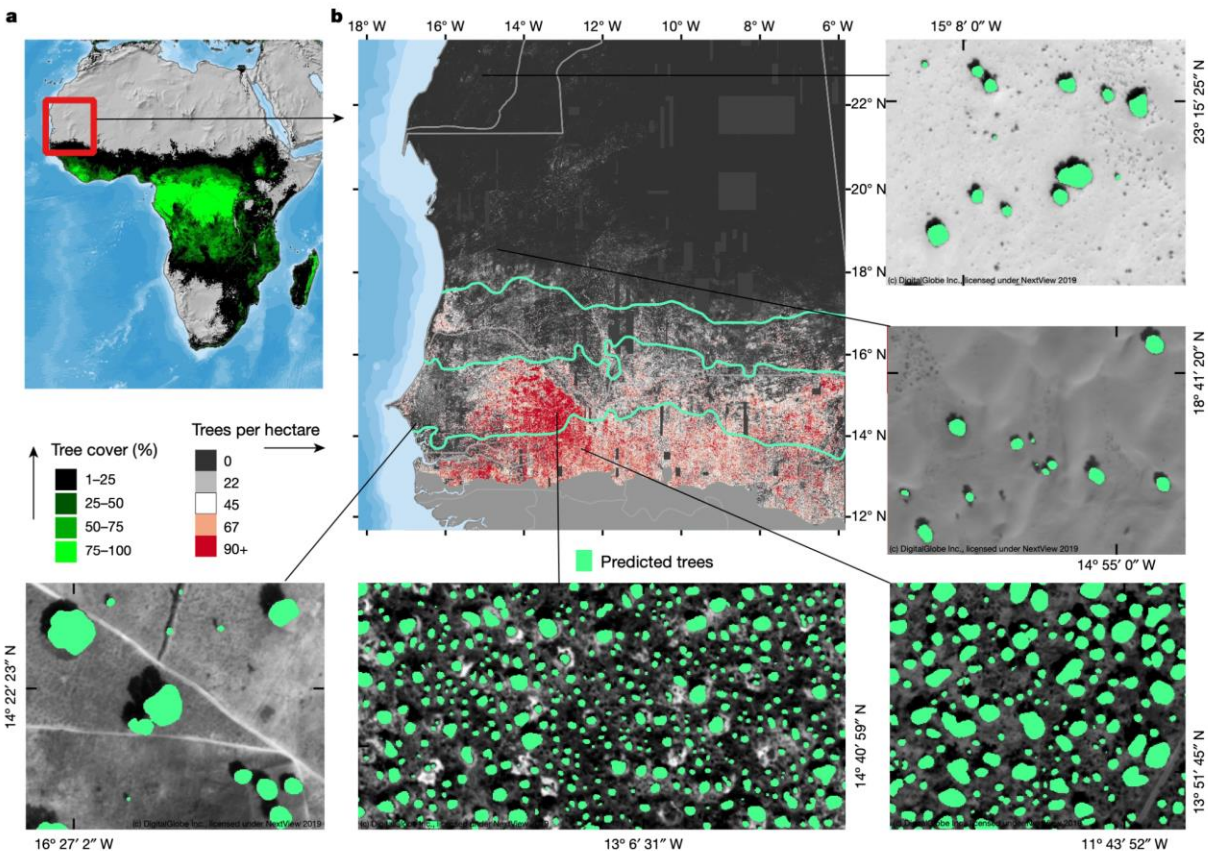

Figure 2: Individual tree segmentation for all SAA, from Brandt et al, 2020.

1.5 Flood Detection

1.5.1 Technical Background and State of the Art

For the last 20 years, Synthetic Aperture Radar (SAR) has been the main tool for the detection and characterization of floods from airborne and satellite measurements. SAR is a type of sensor that actively emits microwaves[20] toward the ground target and measures the variation in intensity and polarization (i.e. the orientation) of the returning signal. SAR signals undergo complex interactions on their path to the ground and back, making the resulting imagery very susceptible to noise[21]. However, with the right expertise in pre-processing techniques, these images can become powerful predictive tools. Two characteristics make SAR particularly useful for flood detection 1) it's capable of collecting images in all weather conditions, and 2) its images capture surface water and dry land as very clearly distinguished, highly contrasted surfaces[12]. A variety of accurate methods have been developed to leverage SAR's water/land contrast and automatically detect changes in surface water over time, from simple thresholding approaches to more sophisticated machine learning ones[22]. However, for many real-life applications of flood analytics, that is only the first step: historical assessments of floods extent and frequency, timely information on the impact and extent of ongoing events, and accurate forecasts of upcoming ones are other key components with important operational implications[23]:

- Historical flood patterns can be helpful predictors of the risk level of an area; the main challenges associated with this task are the availability of geolocated historical data for validation and the availability of archive SAR imagery (only in 2015, with Sentinel-1, SAR imagery started to be collected regularly and frequently at the global level). A solution to this problem is to rely on optical imagery as well, which is available in longer historical records[24].
- Timely monitoring of ongoing floods and their impact is particularly important to mobilize aid and estimate insurance payouts. The main challenge here is that satellite revisit times may not be frequent enough to capture the immediate aftermath (especially after the recent failure of the second Sentinel-1 satellite, which brought its revisit time down to 12

---

[20] Most SAR systems use microwaves with wavelengths ranging from 3 mm to 30 cm; for reference, optical systems (i.e. Landsat, Sentinel-2, Planet, etc.) measure wavelengths between 0.1um and 100 um.

[21] Cian, F., Marconcini, M., &; Ceccato, P. (2018). Normalized Difference Flood Index for rapid flood mapping: Taking advantage of EO big data. Remote Sensing of Environment.

[22] Dasgupta, A., Grimaldi, S., Ramsankaran, R. A. A. J., Pauwels, V. R. N., & Walker, J. P. (2018). Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. *Remote Sensing of Environment*, *215*, 313–329.

[23] de Leeuw, J., Vrieling, A., Shee, A., Atzberger, C., Hadgu, K., Biradar, C., Keah, H., & Turvey, C. (2014). The Potential and Uptake of Remote Sensing in Insurance: A Review. *Remote Sensing*, *6*(11), 10888–10912.

[24] E.g. Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., & Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *80 | Nature |*, *596*.

days). In this case, the integration of different data sources, including new commercial SAR satellites and even publicly available optical data (e.g. Sentinel 2) may provide enough frequent observations.

- Finally, forecasting of events is key for mitigation efforts; accurate weather forecasts, which are hard to come by, and reliable hydrodynamic models are the main current challenges to this task[25].

1.5.2 Top Stakeholders

From the state of the art presented above, in my opinion, the key requirements for the ideal partner in flood detection are 1) mastery of SAR image processing, 2) scalable and reliable models that do not rely too extensively on local fine-tuning, 3) historical risk estimates, 4) forecasting of risk, and 5) automatic and continuous monitoring over time with an alert system. Three stakeholders were reviewed in this domain: Cloud To Street, SarMap, and Lobelia Earth. Of the three, we interviewed only Cloud To Street and SarMap, while Lobelia Earth was reviewed based on online materials. **Cloud To Street** stood out for the maturity of its product and the soundness of its methods (see here for their demo deck). Its core team has worked on EO-based flood detection academically for some time before spinning off the company, and they published peer-reviewed studies on their methodology and applications. Their analytics include flood risk estimates based on long historical records of Landsat data (optical, 30 m resolution), and weekly assessments of flooded areas (the standard resolution is 10 m, but they're able to integrate commercial SAR imagery to generate higher resolution estimates). They do not offer flood risk forecasting yet, although they claim it is in their short-term roadmap. Cloud to Street seems also to be the most mature of the three organizations in terms of product commercialization and adoption; their analytics are already used by several organizations and are nicely packaged into a SaaS platform. This platform continuously monitors customers' AOIs and allows easy visualization of the data and intersection with other data sources (e.g. estimates of the affected population, acres of cropland, etc). They also provide alert systems. **SarMap** presented very reliable methodological approaches (here are a few slides). Its team has over twenty years of experience in SAR image processing - likely the strongest of the three organizations. On the modeling side, they seem to prefer simpler but highly curated techniques; this choice may increase accuracy (they have an 85% minimum requirement for their estimates) but it could have a cost on scalability. SarMap's flood analytics are very similar to Cloud To Street's in terms of what they measure and at what resolution; however, they also have a broader scope of applications that include cropland mapping and yield estimates. Overall, they have a less mature product and they seem to operate mostly as bespoke solution consultants. **Lobelia Earth** has built a flood forecasting and early warning system in West Africa[26] in collaboration with the

[25] Giustarini, L., Chini, M., Hostache, R., Pappenberger, F., & Matgen, P. (2015). Flood Hazard Mapping Combining Hydrodynamic Modeling and Multi Annual Remote Sensing data. *Remote Sensing*.

[26] https://www.lobelia.earth/case-studies/fanfar, last accessed April 26th, 2022.

Swedish Meteorological and Hydrological Institute (SMHI). The core idea is to leverage EOs to measure water level for the Niger River basin and couple this information with a hydrologic model that predicts river discharge. The final product is a 14-day forecasting system that automatically sends early warning signals via SMS, email, or API. It is noteworthy that they have introduced the forecasting element that is missing for both Cloud To Street and SarMap. Lobelia has also done some work at a larger scale and coarser resolution to estimate climatological flood risk based on historical data and river discharge models.

1.6 Soil Carbon Estimation

1.6.1 Technical Background and State of the Art

Classic wet-lab methods that measure Soil Organic Carbon (SOC) concentration and other soil properties are reasonably accurate, but they are time-consuming and expensive. Ground-based hyperspectral measurements, a significantly cheaper and faster solution, have shown significant predictive power for SOC, especially with modern machine learning techniques[27]. Like all field-based sampling approaches, these methods do not scale well to large areas and are not capable of capturing SOC's intrinsically high spatial and temporal variability. The predictive power of ground spectral measurements has triggered several attempts to leverage EO for surface SOC estimates. A few studies deployed multispectral and hyperspectral EO - either simulated or observed - to directly estimate SOC via statistical models. Wang et al (2022)[28] reported promising results; for example, estimates based on the new generation of forthcoming hyperspectral satellites were capable of explaining about 80% of SOC variability (see Figure 1). Unfortunately, not only these satellites do not yet exist (the analysis was performed on simulated data), but they will also have long revisit times (about 30 days) or coarser resolutions (30 m). These characteristics make them likely more suitable for longer-term, large-scale estimates than for smallholder applications. A fusion approach based on existing multispectral satellites was able to explain about 64% of the surface SOC variability with high revisit time and about 10m resolution. Whether this performance level is enough will depend on the use case and its operational requirements. These estimates are more likely to be reliable at detecting change at larger spatial and temporal scales than at the individual field level. For example, this approach may be enough to assess the impact of conservation practices in a small region or very large fields, but it is probably unsuitable for monitoring carbon offset credits from individual farmers. Addressing the suitability of these methods for One Acre Fund's applications will ultimately require running some experiments. Other approaches aim to measure all other components of

[27] E.g. Machine learning and soil sciences: a review aided by machine learning tools - SOIL. 2020, https://soil.copernicus.org/articles/6/35/2020/. Accessed 27 Mar. 2022.
[28] E.g. Wang et al Using soil library hyperspectral reflectance and machine learning to predict soil organic carbon: Assessing potential of airborne and spaceborne optical soil sensing. Remote Sensing of Environment, 2022. Accessed 27 Mar. 2022.

croplands' carbon mass balance (i.e., GPP, respiration, harvest, etc.) and resolve SOC. Direct measures of other components can also be difficult, although they can be approximated using process-based models or EO[29].

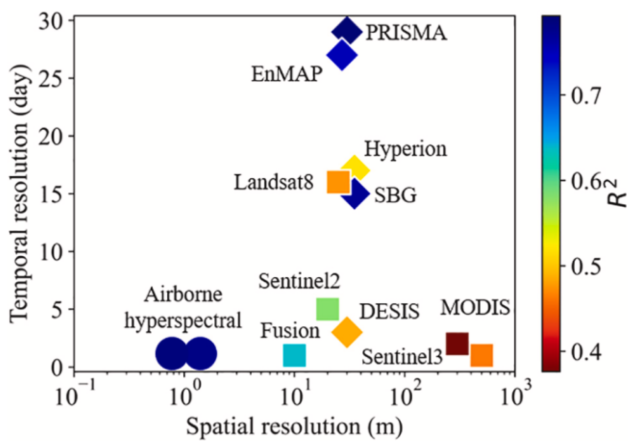

Figure 1: R2, Temporal Resolution, and Spatial Resolution of SOC estimates from simulated satellite and airborne observations. From Wang et al, 2022.

1.6.2 Top Stakeholders

Because of the early development stage of this field, any partnership for SOC estimation will likely involve collaboratively building a research-based prototype rather than acquiring a reliable and scalable product. For this reason, the right partners in SOC monitoring should be capable researchers: transparent about their methodology, rigorous with their validation approaches, and accustomed to navigating experimental uncertainty. Six organizations were reviewed in this domain: Agribora, Airbus, Boomitra, Cropnuts, Habiterre, and 6th Grain. **Habiterre** stood out for its research capability and methodological transparency (its approach is based on years of peer-reviewed academic work). **Boomitra** promises strong results directly from EO by leveraging deep learning; although the performance of their models would need to be verified, potentially through small scale pilots (in ideal conditions, the best models that were reviewed in the literature based on hyperspectral EO explained only about half of SOC variability). **Cropnuts** seem to have a significant capacity in sample data collection and analysis; although not enough to be considered an EO partner, they can complement other partners by helping calibrate and validate models based on ground data.

[29] E.g. Quantifying carbon budget, crop yields and their responses to environmental variability using the ecosys model for U.S. Midwestern agroecosystems. Agricultural and Forest Meteorology, 2021. Accessed 27 Mar. 2022.

1.7 Yield Estimation

1.7.1 Technical Background and State of the Art

Measuring field-level crop yield is a crucial aspect of several applications including optimization of management practices and precision agriculture, risk or loss assessment for crop insurance, and decision-making on interventions or policies by governments and multilaterals. Methods to estimate yield from EOs vary significantly, but they all fundamentally rely on leveraging empirical or modeled relationships between crop physiology and EOs throughout the growing season[30]. The simpler approaches rely on a single EO variable recorded at specific stages of the growing season and empirically relate it to ground-measured post-harvest yield. The regression coefficients can be used to make predictions of yield at the grid level of the EO source. Some studies have used proxies of wetness, such as precipitation or soil moisture. Others leveraged EO Vegetation Indices (VIs), which are combinations of different spectral bands of optical satellites that are well known to correlate with vegetation greenness and productivity. Each of these proxies may be more predictive than the others depending on growing conditions, limiting variables, soil characteristics, management practices, and other factors; for example, in Kenya, precipitation seems to be more predictive in drier, less productive regions, while VIs seem to perform better in areas of more consistent, medium productivity that are more climatically limited. In some cases, the most predicting observable may change interannually or even throughout the season; for example, precipitation can be the better predictor during the emergence phase, while VIs capture more for the rest of the season.[31] This variability makes it difficult to build single-source models that are scalable and reliable across larger regions. Combining multiple sources of data (e.g. both weather and VIs) may increase the accuracy and flexibility of these models, especially with more sophisticated ML techniques capable of capturing complex, non-linear relationships. However, the higher the complexity, the larger the amount of representative ground data[32] - possibly spanning several seasons and large regions - that will be necessary to produce reliable models. To avoid this problem, researchers have developed new methods that rely on large sets of crop simulations constrained with a range of growing conditions representative of the AOI across time and space. The resulting growing curves are used to estimate multivariate relationships between yield and EOs as a function of the growing season stage. Regression coefficients are then used to infer yield from available EOs[33],

---

[30] Benami, E., Jin, Z., Carter, M. R., Ghosh, A., Hijmans, R. J., Hobbs, A., Kenduiywo, B., & Lobell, D. B. (2021). Uniting remote sensing, crop modelling and economics for agricultural risk management. *Nature Reviews Earth & Environment*, *2*(2), 140–159.

[31] Davenport, F. M., Harrison, L., Shukla, S., Husak, G., Funk, C., & McNally, A. (2019). Using out-of-sample yield forecast experiments to evaluate which earth observation products best indicate end of season maize yields. *Environmental Research Letters*, *14*(12).

[32] Crop cuts or full harvest, self-reported data are not reliable enough.

[33] Lobell, D. B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment*, *164*, 324–333.

and ground data is used at the end of the process only to validate and possibly calibrate the final output, requiring much fewer data points than other approaches.

1.7.2 Top Stakeholders

The ideal provider of yield estimates is able to take into account the complexity of smallholder landscapes with minimal dependency on ground data. The spatiotemporal variability in management practices and growing conditions can have a significant impact on the accuracy and scalability of remote yield estimates. The current state of the art shows that approaches relying on calibrated models from single sources of data (e.g. weather or EOs) do not capture sufficient variability. Therefore we look for providers capable of exploiting multiple sources of information (field data, EOs, crop simulations, and weather) via classic or machine learning models. Of the eleven stakeholders reviewed for this analysis, only NASA Harvest, Airbus, and Assimila matched these criteria, while most of the others rely on simple calibrated models on single data sources. **Airbus** has developed an accurate Leaf Area Index (LAI)[34] model based on EO that they have deployed in several agricultural applications. In collaboration with ITK, they used their EO-derived LAI to constrain a biophysical crop model known as Cropwin[35] and estimated yield for maize in Kenya at the field and district levels, both in season and post-harvest (an effort named project Kilimo[36]). The accuracy of their models has not been reported to us, but we will continue to explore their performance further. When evaluating Airbus' yield estimates, one factor to consider is that their imagery has a resolution of 50 cm. Such resolution may not be necessary for insurance-related risk assessments or other regional-level applications, but it would make a significant difference in field-level recommendations on best practices. **NASA Harvest** is a consortium of research-oriented organizations from both the private and public sectors. Taken together, this group likely carries the highest level of knowledge and technical competence on EO-based yield estimates. Individual projects in the consortium span a range of food security applications and many of them focus on SSA smallholder systems. At the smallholder field level, David Lobell's work seems to be still among the best performing and most reliable ones, with uncalibrated yield estimates capturing about half of the total variability of full plot harvest measurements in western Kenya[37]. Besides its technical depth, NASA Harvest remains an academic-type organization and therefore it may be more suitable for exploratory projects than for building an end-to-end product.

[34] LAI is a biophysical parameter measuring the fraction of unit area covered by leaf area. It is usually well correlated with yield in maize and wheat.
[35] https://www.itk.fr/en/cropwin/
[36] https://kilimoproject.itk.fr/en/, Slides here.
[37] Lobell, D. B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., & Murray, S. (2020). 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*, *102*(1), 202–219.

1.8 Detection of Crop Health Disruption

1.8.1 Technical Background and State of the Art

Quoting Terentev et al (2022)[38]: "*Plant stress is a state of the plant in which it is influenced by unfavorable abiotic (light, heat, air, humidity, soil composition, and relief conditions) and biotic factors (phytogenic, zoogenic, microbogenic, and mycogenic). Plant responses to both abiotic and biotic stress are usually complex and include both nonspecific (common for different stressors) and specific components. In a state of stress, plants stop their growth, sharply reduce the activity of their root systems, and reduce the intensity of photosynthesis and protein synthesis"*. As long as the reduction of normal physiological activity is correlated with a change in spectral proxies, detecting ongoing disruptions in crop health directly from EOs is possible. Abrupt declines in the spectral proxies correlated to photosynthetic activity, such as NDVI or fluorescence [39], are usually good indicators of such events. However, establishing the cause of the detected changes (e.g. pests vs mismanagement vs drought) can be difficult without knowledge of existing ground conditions or without post-harvest benchmarks. On the other hand, ongoing crop stress can go undetected if only spectral EO is used, depending on the timing, cause, and other factors. For example, heat and drought can affect certain crops during the grain-filling phase, compromising the productivity of the plant but not the greenness of the canopy[40]. For pests and diseases, timing and growing conditions can be particularly confusing factors. Some pests and diseases may take weeks from the onset before they affect the greenness of their host plants and become detectable from EO. Also, favorable conditions for pest emergence often favor high yield as well, possibly compensating for each other's effect on EO-detectable greenness. This could complicate not only the ongoing monitoring of crop health but also remote yield estimates at the end of the season. A different, perhaps more reliable and effective approach to pest and disease detection, is to monitor and forecast environmental and crop conditions known as favorable to pest emergence. This can be done via weather data, measurements of soil moisture, and crop greenness. If favorable conditions are forecasted or detected, on the ground enumerators or even farmers can be dispatched via text message to verify the presence of any pest or disease. This approach can be streamlined with smartphone applications based on visual recognition.

---

[38] Terentev, A., Dolzhenko, V., Fedotov, A., & Eremenko, D. (2022). *Current State of Hyperspectral Remote Sensing for Early Plant Disease Detection: A Review*.

[39] Guan, K., Wu, J., Kimball, J. S., Anderson, M. C., Frolking, S., Li, B., Hain, C. R., & Lobell, D. B. (2017). The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields. *Remote Sensing of Environment*.

[40] Benami, E., Jin, Z., Carter, M. R., Ghosh, A., Hijmans, R. J., Hobbs, A., Kenduiywo, B., & Lobell, D. B. (2021). Uniting remote sensing, crop modelling and economics for agricultural risk management. *Nature Reviews Earth & Environment*, *2*(2), 140–159.

1.8.2 Top Stakeholders

Given the operational importance of detecting crops pests and diseases in a timely manner, we will be looking for organizations capable of integrating several sources of information and different techniques to evaluate risk, monitor current conditions, and forecast possible outbreaks. Of the 11 providers interviewed, several showed promising approaches although all in the form of bespoke solutions or pilot projects; no ready-to-use product was found in this domain. **TomorrowNow,** the non-profit arm of Tomorrow.io, has run a pilot project in collaboration with CGIAR to predict locust infestation in Kenya[41]. Their approach consisted in using Tomorrow.io weather data forecasts, in combination with swarm movement and hatching models to predict outbreak locations. Crop conditions were monitored in real-time directly from farmers through a crowdsourcing system. In their case, no EO was used, except for the weather variables. **Assimila** and **6th Grain** followed a similar approach. Assimila added Landsat surface temperature to the data stack to predict Fall ArmyWorm outbreaks in China and Laos. 6th Grain included an EO-based crop-monitoring component to track fungal diseases in Egypt. Also **Lobelia**, in collaboration with **FAO**, followed a similar approach; in their case environmental conditions were measured via soil moisture and then included in a pest and disease model. Soil moisture was measured at a 1 km grid by integrating SMAP and MODIS measurements; SMAP - a longer wavelength SAR satellite that is sensitive to water content in the first few inches of soil - offers quite reliable measurements of soil moisture but at a maximum resolution of 9 km. Lobelia's approach uses MODIS surface temperature observations to disaggregate SMAP measurements into a 1 km grid. These are the organizations that stood for their holistic approach and methodological robustness. Other providers offered interesting solutions for single components of the stack (e.g. weather or picture-based recognition of pests) but not end-to-end solutions (see the table).

2 Recommendations

- **AGBC and Automatic Field Delineation:** From a technical standpoint, these verticals have some common requirements. They both need high-resolution imagery (1 m or finer), training and validation labels, and well-designed computer vision models for object segmentation. The modeling aspect is challenging, but it belongs to a class of problems that has been explored widely - there are a few practitioners that can tackle that. Models can be trained on fairly inexpensive annotated images and require minimal ground data for validation. So the hottest commodity is the imagery, and a general recommendation is to **focus on partnering directly with one of the high-resolution imagery providers**. In addition to access to the most important resource, they usually have pretty strong

[41] Farmers Can Get Ahead of Locust Swarms | ICARDA

computer vision teams that can support these efforts. On this front, potential partner would be **Airbus:** they have high-quality imagery at 50 cm (the next-gen, which was recently launched, is 30 cm), a good computer vision team that has already worked on segmenting smallholder farmers' fields, and they're looking for new partners to extend their project Kilimo in Kenya. Even though they have not worked on single-tree segmentation, this could be feasible, especially given they could leverage the open models from the University of Copenhagen. Costs may be prohibitive at scale, but this needs to be verified. AGBC has a further requirement: the conversion of CPA values into ABG while capturing allometric variability. Providers such as **Lobelia Earth** could support this type of work given their ongoing effort in end-to-end tracking of the AGBC market in Niger. Lobelia's work in collaboration with the University of Copenhagen relies on Airbus imagery for CPA segmentation, which reinforces the benefits of a direct partnership with imagery providers.

- **Pest and disease:** The research did not identify a mature, ready-to-use product in this field. Tackling this problem will likely need some degree of experimentation and active collaboration. The most reliable approaches aim at leveraging EOs to monitor and forecast ideal pest conditions. The ongoing efforts of **TomorrowNow with CGIAR** seem promising because they bring together a great combination of agronomic expertise, swarms modeling, and weather forecasting. TomorrowNow and CGIAR are looking for new partners to extend this project in Kenya, and this could be a good opportunity for Africa-based organizations to test these models on the ground. Partnering with TomorrowNow also has a key strategic value for other verticals: the parent organization, Tomorrow.io, is a good provider of global weather data and forecasting (with new dedicated satellites launching soon). Access to this data, perhaps at a reduced cost, could open the doors for farmer-facing organizations to potentially develop their own pest and disease monitoring systems and further improve the advice provided to client farmers. **Lobelia Earth** has done some excellent work producing soil moisture estimates - an important driver for some pests - although they did not own the end-to-end pipeline, which was modeled by FAO.

- **Flood:** Although Sarmap and Cloud To Street are methodologically comparable, ultimately the product maturity of **Cloud To Street** will likely better serve real-life applications that require constant, automatic monitoring and warning. One aspect Cloud to Street is missing is the shorter-term risk forecasting associated with weather events. Although their team says they are working on this feature, it will be important to get a realistic timeline for their roadmap. **Lobelia Earth**, thanks to their collaboration with SMHI to build an early-warning platform in West Africa, may prove useful in this area. In addition, I would flag here that good weather forecast data is a key piece for proper flood forecasting; **Tomorrow.io** weather data may be beneficial here as well.

- **Soil Carbon:** Of all verticals, this one is likely the most challenging to tackle from EOs, and it will likely require a targeted research collaboration. **Habiterre** is a promising partner for this; their experience in this domain, research skills, and methodological transparency stood out from the other stakeholders. Their work will need to be supported by ground data collection, which could be supported by **CropNuts**. Their approach, which is based on collecting hyperspectral soil samples, is lower cost and more scalable than other alternatives (such as wet-lab measurements). In addition, hyperspectral soil measurements will be a good investment for the near future, when new hyperspectral satellites will become available and possibly open the door to more accurate remote SOC measurements.

- **Yield:** The most promising provider in this domain is **Airbus**. While this report did not have an opportunity to evaluate Airbus' yield predictions in Kenya, the methodological approach they follow falls in the category of the most reliable ones (EO for greenness and weather + crop simulations). In addition, the fact that they provide very high-resolution estimates opens the doors to a wider range of applications in smallholder landscapes. Another promising partner to highlight is **NASA Harvest**. They would be helpful in validating and improving methodologies and results, making sure that everything is done according to the state of the art.

So to summarize, below are a few highlighted partnership opportunities that smallholder facing organizations can consider:

Vendor	*Recommend for:*
Lobelia Earth	lead ABGC pipeline, support pest/disease pipeline, and support flood risk monitoring.
Airbus	access to high-resolution imagery, lead field delineation, support ABGC, and lead yield estimates.
TomorrowNow/Tomorrow.io	access to weather data and forecast, lead pest and disease pipeline with CGIAR, support flood pipeline
Cloud to Street	lead flood pipeline (risk, monitoring, and forecasting).
Habiterre	lead SOC pipeline
CropNuts	support SOC pipeline for ground support
NASA Harvest	support yield estimates.