Monitoring Agricultural Investments in Ethiopia: A Remote Sensing Based Approach

Matthias Hack , Fabian Löw , Guido Lemoine , Oliver Schönweger , Mulugeta Tadesse , Felix Rembold , Dimo Dimov

Introduction

The Government of Ethiopia envisages fostering food security and increase agricultural production by extending cultivated area for example through leasing out large areas of agricultural land to foreign and domestic investors. Experience so far show, that these large scale land acquisitions (LSLA) have not delivered the envisaged benefits so far. In order to track successful development and implementation of such leases there is a need for monitoring the implementation of these agricultural projects. To overcome this shortcoming, the *Ethiopian Horticulture and Agricultural Investment Authority* (EHAIA) with the support of GIZ is currently working on the establishment of a monitoring tool based on satellite remote sensing data. Monitoring results would allow to describe the state of land use for every LSLA in Ethiopia on a regular basis and thereby assess the performance and compliance of the agricultural investors. Free satellite images from the Copernicus earth observation program and visual interpretation of these data could constitute a basic version of the tool that would enable basic monitoring functions at very low costs and little technical requirements. The tool could be further elaborated to generate detailed information e.g. on the spatial distribution and vital status of specific crops by incorporating additional remote sensing data and sophisticated algorithms for analysis. Thereby, the tool becomes flexible and could be used for other spatial planning context e.g. in the forest sector.

Problem Statement

Ethiopia has made substantial progress in boosting food production and reducing hunger and malnutrition over the last two decades. However, Ethiopia has remained food insecure, with food deficit each year. The Government of Ethiopia five-year strategy, “Growth and Transformation Plan (GTP)”, envisages that Ethiopia becomes a food-secure, middle-income country by 2025 and increases output of major crops from 19 to 27 million tons during the period of the plan [1].

To increase private investment in the agriculture sector, areas suitable for large-scale agricultural investments (LSAI) have been identified to be allotted to investors. As a result, 3.6 million hectares of land were identified out of which about 2.4 million hectares of land transferred to investors for agricultural investments by the federal and regional government agencies over the period up to 2014. On the national level, EHAIA is mandated to guide and administer the areas suitable for large-scale commercial farming. It is responsible for facilitating agricultural investments as well as land administration and transfer processes. Large-scale agricultural investments are expected to bring foreign currency as well as technology transfer to the country, while the local communities would benefit from employment, technology and infrastructures related to these investments according to a trickle-down effect. Proper design and implementation of these projects would allow the achievements of these Government policy objectives.

Progress towards achieving these objectives has been rather limited so far. Out of the 2.4 million ha of land transferred to about 6,000 private investors, only one-third has been developed by these investors until 2014. Same plots have been given out to different investors, and many investors have not used provided land productively, even years after the contractual start. Besides the remoteness of many investments sites and corresponding infrastructural problems, partly agro-ecological unfavourable conditions or overlaying land use claims, administrative capacity constraints hamper the Government's efforts to promote sustainable large-scale commercial agricultural development. One major shortcoming in administrating LSLA is poor geo-information, which is needed throughout the whole investment process:

- Before the investment: The choice of location has to be based on an intensive investigation of prevailing land uses and the natural endowment. The replacement of existing agricultural or housing activities should be prevented and unused land without important ecology value, but with favourable agro-ecological conditions should be prioritized for the allocation of LSAI. Such information-based planning is crucial for a conflict-free and productive investment.

- During the investment: Spatial monitoring of the investment is needed to control the investment spatially and timely evolves as fixed in the contract. Such spatial monitoring enhances compliance of contracts and guiding of investments.

EHAIA has little capacities to collect needed spatial data and to process this data to relevant information. This is especially true for in-situ data collection at the (potential) investment sites.

Solution

A promising solution to alleviate problems related to spatial information is utilizing earth observation (remote sensing, RS) for monitoring agricultural activities within LSAI areas. It provides repetitive, timely, and objective information about the surface of the earth from a distant platform, usually a satellite or airborne sensor, or unmanned aerial vehicle (UAV). This platform collects reflected electromagnetic radiation over large regions and processes the data into a digital image and finally into digital maps. RS data as the ones which are freely available with Sentinel-1 (S1) and 2 (S2) from the recently launched Copernicus earth observation program[1] of the EU constitute a major asset for this kind of application [2,3] and provide an unprecedented opportunity to monitor also fragmented agricultural landscapes with small-scale farming such as in Ethiopia. Time series[2] of Sentinel data reach expectations for agricultural monitoring in many regards: frequent (S1 every 6 days, S2 every 5 days), systematic and full coverage with free, high-resolution data (S1 up to 5x5m pixel size, S2 up to 10x10m pixel size) support agricultural monitoring even at the field level. S1's payload is a Synthetic Aperture Radar (SAR) operating in the radar's C-Band (microwave). It is hence not affected by atmospheric conditions (clouds, aerosols). The complementarity of the sensor attributes (SAR, optical) enhances the accurate recognition of crop types, crop acreage estimation, and the separation of land cover categories [4,5]. The Copernicus program operates four satellites of S1 and S2 each, to guarantee data availability at least until 2027 (with "next generation" versions beyond). This together with the "free and open" policy for Sentinel data and with the rapidly decreasing costs of local and cloud based IT infrastructure, provides a suitable basis for initiating a long-term monitoring tool in Ethiopia [6]. Existing projects like JECAM[3]

---

[1] http://www.copernicus.eu

[2] A "time series" refers to a set of consecutive images acquired from one region along a given time frame, e.g. along the growing season within one year.

[3] http://www.jecam.org

convincingly demonstrate the utility of RS for agricultural monitoring in various agricultural landscapes in Africa.

Open and free geographical information systems (GIS[4]) enable the storage, management and analysis of large quantities of such spatially distributed data, which are associated with their respective geographic features (e.g. agricultural fields). These two applications merge, when the remotely sensed data used for mapping and spatial analysis is overlaid with other spatial GIS data of the same geographic site. It is a scientifically sound, proven and potentially cost-effective approach [7], that can be maximized for accurate cropland mapping in smallholder-dominated Savannah landscapes [8,9], burnt area mapping [10], land cover mapping in Ethiopia. [11,12]. Therefore, it is suitable to assess agricultural activities in LSAI areas in Ethiopia [13]. [13] first explored and demonstrated the potential of using Sentinel data for land conversion monitoring in Gambella. Likewise, previous studies point to the potential cost-effectiveness of RS for agricultural statistics [7].

Thus, EHAIA with support of GIZ and the EU is currently developing an operational, agricultural monitoring system, based on the aforementioned user needs and requirements. It concentrates on defining an integrated, low-cost tool primarily based on multi-sensor RS data, which shall satisfy the needs of the national level users, especially for knowledge about LSAI areas (i.e., cropped area and crop acreage, evolvement monitoring). It is supposed to (i) regularly deliver this information for each LSAI and (ii) to bring added value and savings to existing ground surveying techniques or where ground surveying is restricted. The tool can support monitoring single investment sites and supports assessing whether and to what degree an investor of a LSAI is compliant with the contractual agreements.

The development of the tool is technically supported by MapTailor Geospatial Consulting GbR and incorporates advise from the Joint Research Center of the European Commission.

Concepts of Operation

As a first step of conceptualization, it was examined which prerequisites (technical, system, software, personnel) and processing routines (e.g. algorithms, workflows, tools) meet different requirements best. Such a tool and the information that it delivers to the end users can be of various degrees of sophistication. The functionalities of a monitoring tool, which satisfies the user´s information demand, can be regarded as "modules" and added on demand. An increasing information demand thus goes along with an increasing number of functionalities, a higher demand of different input data (for example, reference data) and higher prerequisites in terms of the skills of the operators of such a tool, or processing power of the IT infrastructure. Setting up a monitoring tool is not a "one-of" process. Whilst a basic system could primarily focus on cultivated land delineation and supporting surveys, a more sophisticated version could be progressively built as an extension of the basic system. Thus, the implementation of such a tool can be considered as a phased approach from basic to more sophisticated. The following section exemplarily presents a simple (basic) and enhanced-basic (basic +) manifestations of this modular concept. These basic versions have little technical requirements and depict a good starting point to establish the monitoring tool. As stated before, the tool could be further enhanced with more sophisticated ways of analysis which would generate detailed information and enable further applications. However, more sophisticated manifestations of the tool are not

[4] A GIS is a computer based technology consisting of hardware, software, data and applications. It allows capturing, editing, storing, reorganising, modelling and analysing all forms of geographically referenced information. With the help of GIS, data can be interpreted and visualised in many ways that reveal relationships, patterns and trends in the form of maps, reports and charts.

explained in this paper, as the next steps of implementation will initially focus on the basic versions.

Basic version and basic version +

The rationale behind the basic option is to ensure maximum utility for providing basic spatial information for mapping and tracking the spatial evolvement of commercially cultivated areas (CCA) within the investment sites of a contracted LSAI. At the same time, it shall help to reach a compromise between limited budget and personnel capacities available for ground surveys, sampling correctness, and quality of information provided. The basic monitoring tool uses freely available satellite images from the Copernicus earth observation program to produce georeferenced products (digital maps) at the end of the season annually. Such maps are characterised by a classification that entails several land use/land cover (LULC) classes. The basic monitoring tool helps to monitor single investment sites and assessing if an investor is compliant with contractual agreements.

The workflow for the basic version can be summarized as followed:

(i) Downloading optical and SAR satellite images from the Sentinel missions.

(ii) pre-processing these data sets, e.g. clipping data on areas of interest, atmospheric corrections, etc.;
in case of option “basic +”: gathering reference data from e.g. field surveys or existing high resolution images.

(iii) Transforming the data into information - here: digital maps of LULC by visual interpretation. For the visual interpretation of satellite images, additional steps could be undertaken to enhance the visibility and accurate detection of spatial structures in satellite images that indicate the presence of commercially cultivated areas (CCA), for example the establishment of (canals, road) infrastructure and land clearance (detection of burnt area). As a proof of concept, next to the pre-processed images, false colour composites were displayed in the open and free GIS “QGIS”, and a principal component analysis (PCA) and unsupervised image classification using k-means algorithm have been performed to support the interpretation;
in case of option “basic +”: preparation from training and validation data derived from reference data, supervised image classification

(iv) Visualizing the results as digital maps, reports, and statistics.

Following chart summarizes the general workflow for the basic options:

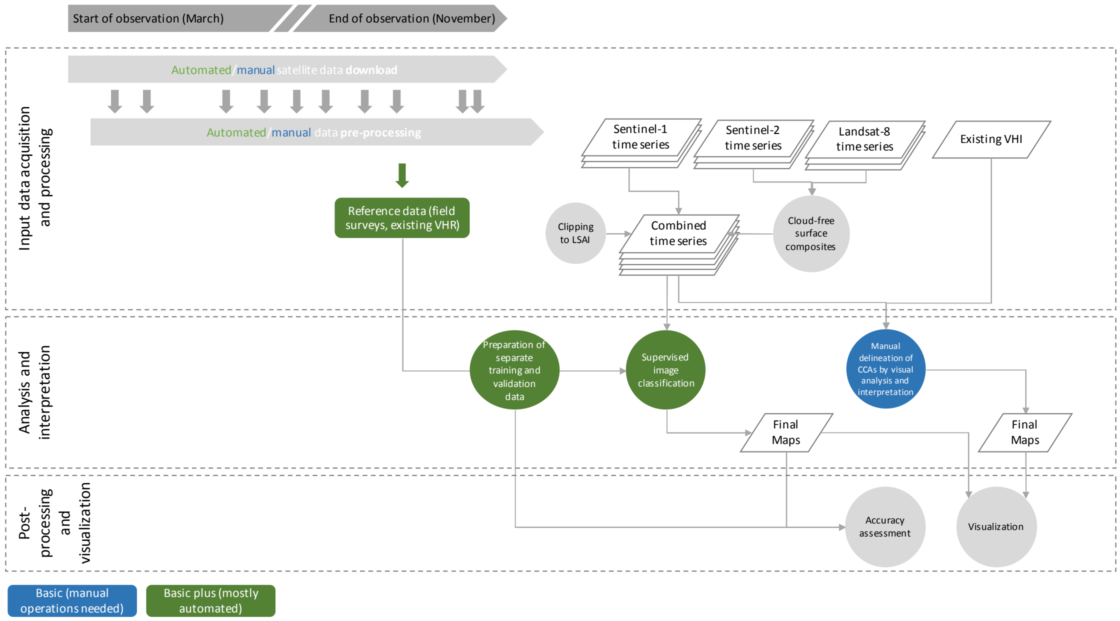

Figure 1: General workflow of the basic module

Setting up the tool requires a high performance IT infrastructure, 1-2 staff and optionally field assistants. The processing chain can either be done manually using free and open software packages, which is relatively straightforward to implement, or be largely automated by coding in programming languages, allowing significant time and thus cost savings after the initial setup phase. The initial setup phase is estimated 1–3 years and should be framed by user/stakeholder workshops and trainings.

As foreseen for the basic approach, unlabelled map products can be delivered annually in GeoTIFF format and are dedicated for manual interpretation through the system operators. The maps in Figure 2 show these deliverables, which enable separating CCA from non-CCA derived from visual interpretation. The unsupervised classification even revealed various cropping intensities / stages of development in CCA and supported the visual interpretation.

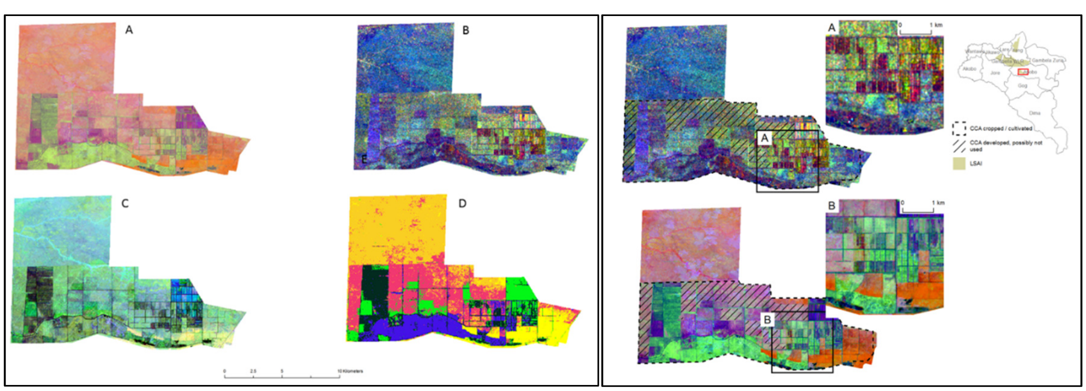

Figure 2: Left: Examples of maps used to assist the visual image interpretation and identification of CCA. A: S2 PCA, B:S1 PCA, C: NDVI 3-monthly composites, and D: unsupervised classification. Right: CCA actually cultivated and other CCA prepared /developed for cultivation, delineated by on-screen digitization in QGIS.

The basic + option foresees a supervised image classification using a machine learning algorithm, for example Random Forest (RF) [14,15] or Support Vector Machines (SVM) [15–17]. To do so, training data (i.e., pixel coordinates and associated class labels) are used to train a classifier algorithm and to classify the pre-processed times series (S1 and S2). This can completed on demand by an a-posteriori filtering for smoothing out minor errors in the resulting maps and to create more accurate and visually appealing maps (see figure 3), as was done in the PoC and other studies in Ethiopia [11].

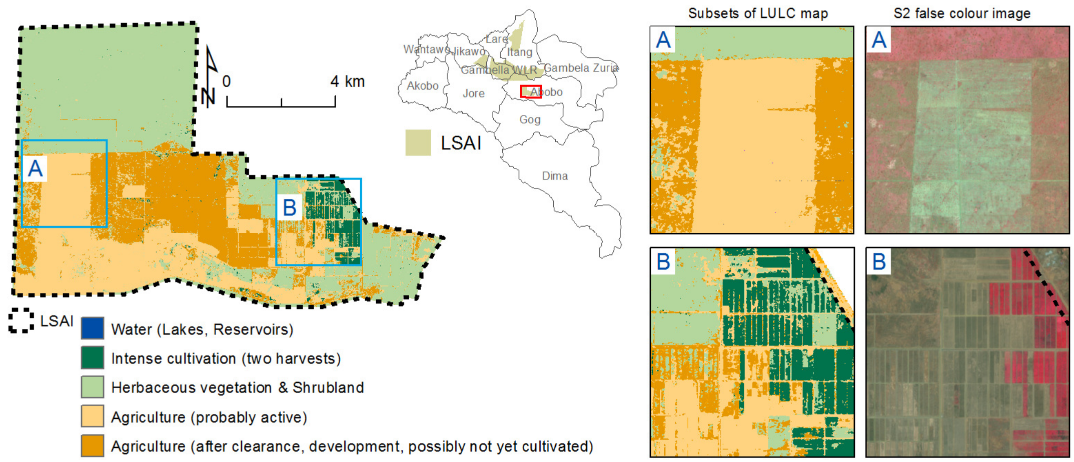

Figure 3: Example of a map created by supervised image classification including post-processing filtering in Gambella. Non-active CCA includes land prepared and developed (e.g. vegetation clearance, infrastructure like roads and canals) for CCA, see image subset A. Sentinel-2 image false color composite (vegetation in red color) in the small image subsets support the visual interpretation. The area is a mix of cleared areas, some of which are already cultivated (image subset B), and areas that are not yet cleared.

Second to the last step of the processing chain, the quality of the maps is quantitatively assessed, based on independent (external) validation data for creating detailed accuracy assessments [18,19,20]. This serves the purpose of providing users with validated algorithms and information about the quality of the maps. Validation data is used to assess the map quality through the calculation of confusion matrices[18]. Such confusion matrices compare, on a class-by-class basis, the relationship between the validation data and the corresponding results of the classification. The predefined quality aim of the final product is an overall accuracy of 85%, preferably with higher per-class accuracy for CCA.

Conclusion

A regular, accurate and cost effective spatial monitoring of every LSAI in Ethiopia with satellite data is feasible. It provides a solution for spatially consistent, timely and objective information retrieval about spatial evolution of commercially cultivated areas. Using satellite data enables monitoring land use state and land conversion where local data is not sufficient or comprehensive ground surveys are limited. A detailed prove of concept revealed how open and free software, in conjunction with the free Sentinel satellite data from the Copernicus earth observation programme reaches expectations for an accurate and low-cost agricultural monitoring even at the parcel level. Accuracies of maps can be higher than 85% (for certain land use / land cover categories even higher than 90%).

The actual implementation of the monitoring tool would require some initial investments in an appropriate IT infrastructure and personnel and should be preceded by measures of user requirement evaluation, to tailor the specifics of the tool and its actual implementation to the requirements and capacities of the mandated agency.

References

1. Keeley, J.; Seide, W. M.; Eid, A.; Kidewa, A. L. *Large-scale land deals in Ethiopia : Scale, trends, features and outcomes to date*; London:, 2014.

2. Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; Koetz, B. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sens.* **2015**, *7*, 12356–12379.

3. Atzberger, C. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs. *Remote Sens.* **2013**, *5*, 949–981.

4. Inglada, J.; Vincent, A.; Arias, M.; Marais-Sicre, C. Improved early crop type identification by joint use of high temporal resolution sar and optical image time series. *Remote Sens.* **2016**, *8*.

5. Navarro, A.; Rolim, J.; Miguel, I.; Catalao, J.; Silva, J.; Painho, M.; Vekerdy, Z. Crop monitoring based on SPOT-5 Take-5 and sentinel-1A data for the estimation of crop water requirements. *Remote Sens.* **2016**, *8*.

6. JRC-MARS *Concept note -Towards Future Copernicus Service Components in support to Agriculture ?*; 2016.

7. FAO *Cost-Effectiveness of Remote Sensing for Agricultural Statistics in Developing and Emerging Economies - Technical Report Series GO-09-2015*; 2015.

8. Sweeney, S.; Ruseva, T.; Estes, L.; Evans, T. Mapping cropland in smallholder-dominated savannas: Integrating remote sensing techniques and probabilistic modeling. *Remote Sens.* **2015**, *7*, 15295–15317.

9. Zoungrana, B. J. B.; Conrad, C.; Amekudzi, L. K.; Thiel, M.; Da, E. D.; Forkuor, G.; L??w, F. Multi-temporal landsat images and ancillary data for land use/cover change (LULCC) detection in the Southwest of Burkina Faso, West Africa. *Remote Sens.* **2015**, *7*, 12076–12102.

10. Verhegghen, A.; Eva, H.; Ceccherini, G.; Achard, F.; Gond, V.; Gourlet-Fleury, S.; Cerutti, P. The Potential of Sentinel Satellites for Burnt Area Mapping and Monitoring in the Congo Basin Forests. *Remote Sens.* **2016**, *8*, 986.

11. Eggen, M.; Ozdogan, M.; Zaitchik, B.; Simane, B. Land Cover Classification in Complex and Fragmented Agricultural Landscapes of the Ethiopian Highlands. *Remote Sens.* **2016**, *8*, 1020.

12. Lemmen, C.; Zevenbergen, J.; Lengoiboni, M.; Deininger, K.; States, U.; Burns, T. First experiences with High Resolution Imagery Based Adjudication Approach for Social Tenure Domain Model in Ethiopia. In *FIG – World Bank Conference, Land Governance in Support of the Millennium Development Goals – Responding to New Challenges*; 2009; pp. 1–22.

13. Lemoine, G.; Rembold, F. *Ethiopia land conversion monitoring: Gambella state - JRC Technical Reports*; Ispra, 2016.

14. Rodriguez-Galiano, V. F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J. P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104.

15. Vapnik, V. *Statistical learning theory*; 1st ed.; John Wiley & Sons: New York, 1998.

16. Mathur, A.; Foody, G. M. Crop classification by support vector machine with intelligently selected training data for an operational application. *Int. J. Remote Sens.* **2008**, *29*, 2227–2240.

17. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259.

18. Congalton, R. G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46.

19. Olofsson, P.; Foody, G. M.; Herold, M.; Stehman, S. V.; Woodcock, C. E.; Wulder, M. A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57.

20. Olofsson, P.; Foody, G. M.; Stehman, S. V.; Woodcock, C. E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131.