Monitoring Agricultural Investments in Ethiopia: A Remote Sensing Based Approach

Matthias Hack[1], Fabian Löw[2], Guido Lemoine[3], Oliver Schönweger[4], Mulugeta Tadesse[5], Felix Rembold[6], Dimo Dimov[7]

Abstract:

Between 2005 and 2016, the Ethiopian Government leased about 2.4 million hectares of land for commercial agricultural investments to private domestic and foreign investors as a means of economic growth, food security and job creation. In order to steer these vast amounts of large-scale agricultural investments towards the envisaged benefits, it is crucial to monitor the investments' implementation progress frequently. But so far, the mandated agencies neither have the resources nor the capacities to check the implementation status of every contracted investment at certain regular intervals.

Germany and the European Union are supporting the Government of Ethiopia in improving policies and management related to large scale agricultural investments towards sustainable economically and socially responsible outcomes. Therefore, the Project Support to Responsible Agricultural Investments (S2RAI) - implemented by Deutsche Geselllschaft für Internationale Zusammenarbeit (GIZ) GmbH - provides technical support to the official Ethiopian partner institutions and other involved stakeholders like investors, local population and civil society organizations. One field of action is the establishment of a functional remote sensing monitoring system (RSMS), that will deliver regular information on the status of each large-scale agricultural investment. The project collaborates in this regard i.a. with the Joint Research Center of the European Commission (JRC).

The proposed paper describes how the envisaged RSMS will make use of high-resolution satellite imagery from the Copernicus Earth Observation Program of the European Union, namely Sentinel-1 and Sentinel-2 satellites. The data as well as the software to process and download the result are provided by Google through Earth Engine cloud infrastructure for free for non-commercial use. The workflow of accessing satellite images, pre-processing these data sets, and transforming the satellite data into information related to the actual land use and land use change within large scale agricultural investments is designed in a way that it is easy to implement and to replicate.

Key Words: Ethiopia, large-scale agricultural investments, monitoring, remote sensing, sentinel satellite

[1] *Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, matthias.hack@giz.de*
[2] *MapTailor Geospatial Consulting GbR, fabian.loew@maptailor.net*
[3] *Joint Research Center (JRC) European Commission, guido.lemoine@jrc.ec.europa.eu*
[4] *GIZ, oliver.schoenweger@giz.de*
[4] *GIZ, mulugeta.tadesse@giz.de*
[5] *JRC, felix.rembold@ec.europa.eu*
[5] *MapTailor Geospatial Consulting GbR, dimo.dimov@maptailor.net*
[6] *JRC, felix.rembold@ec.europa.eu*
[7] *MapTailor Geospatial Consulting GbR, dimo.dimov@maptailor.net*

Background to Large-Scale Land Investments In Ethiopia

Between 2002 and 2012, the Government of Ethiopia (GoE) leased about 2.4 million hectares of land for commercial agricultural investments to private domestic and foreign investors, while more than 3 million of hectares have been classified as suitable for agricultural investments. Large-scale agricultural investments (LSAI) are expected to bring foreign currency as well as technology transfer to the country while the local communities are supposed to benefit from employment, technology transfer and infrastructural improvements. In the *Growth and Transformation Plan for 2010-2015* it is stated:

> "Fundamentals of the strategy include the shift to produce high value crops, a special focus on high-potential areas, facilitating the commercialization of agriculture, supporting the development of large-scale commercial agriculture where it is feasible. The commercialization of smallholder farming will continue to be the major source of agricultural growth." (Federal Democratic Republic of Ethiopia, 2010)

Most of the LSAI are located in the western lowlands of Ethiopia in the Regional States "Benishangul-Gumuz" (BSG) and "Gambella" (GB). These areas are identified as suitable for LSAI in particular, mainly for their sparse population density, the availability of fertile land and the supposedly underdeveloped agricultural structures:

> "The Gambella as well as the Benishangul Gumuz Regional States are geographically located at the lowlands of the western part of the country. In the two regional states the sparsely settled inhabitants eke out a living from shifting cultivation and animal husbandry. Both regional states famous known for bounty and plenty are blessed with a vast amount of arable lands. Ironically though, the fecund lands used to lie fallow. Sad as it may sound their farming system too has been archaic. In light of these facts there is a call for breathing life in to their way of living. Rendering people of the two regional states beneficiaries of socio-economic services necessitates clustering dots of huts scattered across the respective regional states and introducing modern agriculture." (Ministry of Agriculture and Natural Resources, 2016)

So far, the strategy of commercializing the agricultural sector through LSAI did not pay of as envisaged. Many investment sites did not develop the land as fixed in the contracts. According to discussions with officials from GB and BSG, only a fraction of the land that was supposed to be cultivated until now is actually under cultivation in GB. In BSG a bigger share of leased land is cultivated, but still by far not the amount as it was fixed in the contracts. Besides the underperformance of the investment projects, other problems affiliated to LSAI are reported by various sources – such as forced evictions, not permitted under-leasing or clearing and planting outside the investment boundaries (Keeley, Seide, Eid, & Kidewa, 2014). Reasons for the unfavorable development of LSAI are manifold, e.g.

- General framing conditions like fluctuating food prices
- Strategic policy decisions like very low lease prices
- Administrative shortcomings e.g. in the identification of suitable land, screening of investors, contract design and monitoring due to capacity constraints
- Miscalculations or even speculative ambitions of the investors

Supported by federal institution, former *Ethiopian Horticulture and Agricultural Investment* Agency (EHAIA), the regional land administration is responsible to identify suitable areas, to lease out sites to the investors and to control the investments during implementation phase.

Germany and the European Union are supporting the GoE in improving policies and management related to LSAI towards sustainable and economically and socially responsible outcomes. Therefore, the Project *Support to Responsible Agricultural Investments* (S2RAI) - implemented by *Deutsche Geselllschaft für Internationale Zusammenarbeit (GIZ) GmbH* - provides technical support to the official Ethiopian partner institutions and other involved stakeholders like investors, local population and civil society organizations. One field of action is the establishment of a functional monitoring system that will deliver regular information on the status of each LSAI. The project collaborates in this regard i.e. with the *Joint Research Center of the European Commission* (JRC).

Monitoring Large-Scale Land Investments

Frequent and comprehensive monitoring of LSAI is a crucial part in steering the process of establishing LSAI as mean of sustainable development in Ethiopia. Regular monitoring information on the status of implementation of the projects is important to hold the investors accountable based on commitments entered into the contractual agreement and business plans. It helps to understand, prevent and solve land-related conflicts among investors or investors and the local population. Furthermore, aggregated information on the performance and impact of LSAI is important for respective policy and decision makers.

However, due to capacity limitations of the mandated agencies, monitoring could only be carried out in a piecemeal approach and only few investments were checked by field inspections. Nonetheless, in both GB and BSG the authorities managed to carry out monitoring most of LSAI projects between 2017 and 2018. However, this was intended as a onetime exercise, not a repeating measure of constant monitoring.

S2RAI works on the realization of a monitoring system that allows regional and federal authorities to generate comprehensive information on each LSAI on a regular basis.

According to the concept of this monitoring system, the scope of the aspects of field level monitoring are broad: general data, compliance with legal aspects, compliance with the contract (i.e. compliance with the environmental management plan and the business plan) and performance. To each of these themes a set of parameters are assigned that clearly measures the status of the various aspects of LSAI. Major data sources identified include file keeping and sharing which also includes self-declarations by the investors as well as file sharing among various line departments, and field inspections.

Remote sensing analysis (RSA): By analyzing satellite imagery, all LSAI sites can be monitored throughout a year. The information generated by RSA will e.g. indicate the extent and pace of cultivation. Occasional field inspections might be triggered by such analysis.

The focus of this paper is the conceptualization and implementation of RSA as one important component of monitoring LSAI.

Obviously, many outcomes and impacts of LSAI have a physical manifestation that can be observed with remote sensing (RS). If the approach is well designed, such a remote sensing based monitoring tool (RSM) can deliver valid, objective and replicable information on the spatial dimension of LSAI at low costs and efforts. In respect to the limited capacities of the mandated agencies, RS becomes a crucial pillar of the monitoring system.

Major monitoring questions that are addressed by the RSM:

- How much area land within an investment site is cultivated?
- At which pace was the land cultivated?
- Which crops are cultivated?
- Do implementation activities like clearing, planting and constructing cross investment boundaries?
- Does other land use occur within the investment site?

The following subsequent sections describe which RS data and approaches of analysis are chosen to meet the monitoring needs and the prevailing capacities of the mandated agencies.

Eco-Zonal and Agricultural Characteristics of the Area of Interest

Profound knowledge of the eco-zonal characteristics as well as of prevailing agricultural practices are important to specify the right RS data and analysis methods for the RSM. Therefore, investigation on the prevailing climate conditions and the phenology of natural vegetation and cropland was undertaken.

According to the KÖPPEN-GEIGER climate classification, the western lowlands of Ethiopia belong to the class "Tropical wet and dry or savanna climates (Aw)" (Peel, Finlayson, & McMahon, 2007). The average temperature of every month is above + 18 degree Celsius, with highest temperatures occurring just before the rain session (April to October). The month from December to March depict dry season. Precipitation mainly reaches maximum in the month of August.

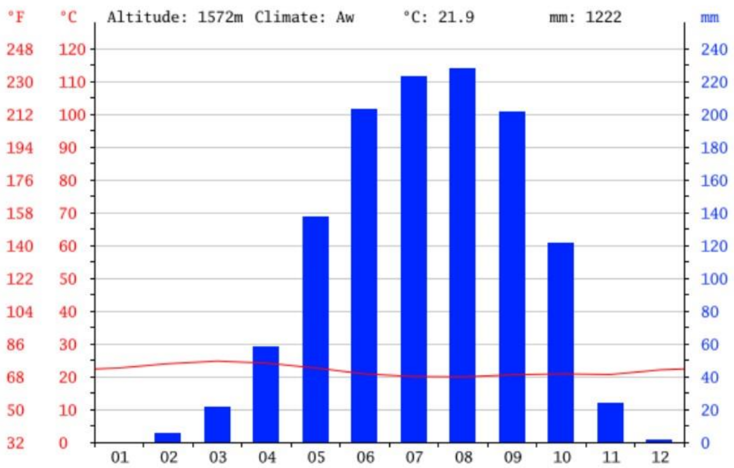


Figure 1: Climate graph of Assosa (regional capital of BSG). (Merkel, 2016)

Referring to the eco-zonal classification scheme of SCHULZ (2002) the prevailing natural vegetation characterized as "arid savannah" in the east and "humid savannah" in the west, respectively and "short grass -" and "high grass savannah". Density of scrub- and tree cover differs according to the water availability and land use intensity. The variability of precipitation during the year heavily determines the phenological appearance of the natural vegetation. In the dry session parts of the herbaceous plants aboveground dry out, followed by leaf fall of the alternate green wood species (Schultz, 2002).

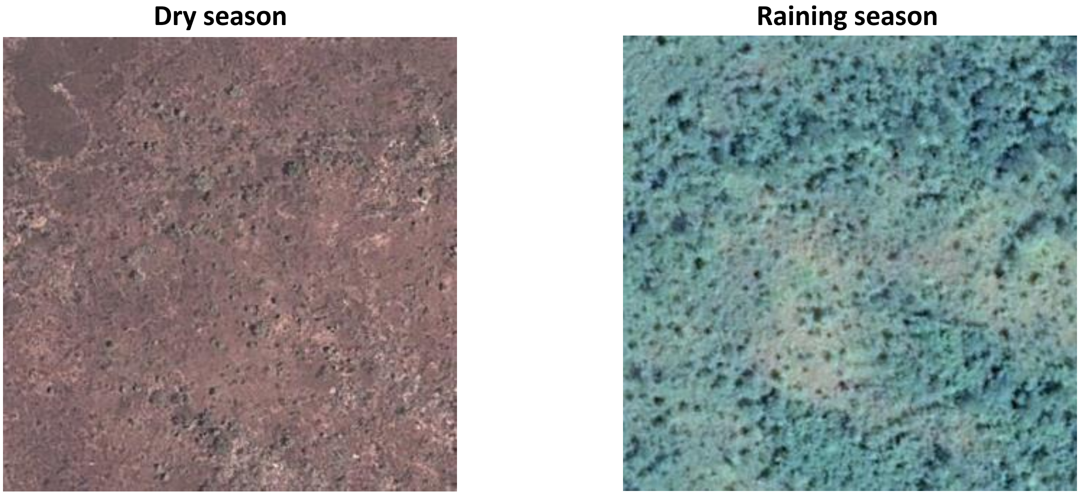


Figure 2: Phenological difference of natural vegetation between dry and raining season. Very high-resolution images from Google Earth: left: acquisition date: 09.05.2012, right: acquisition date: 11.10.2016. (Google Inc.)

Fire events are a common phenomenon in the area. Natural fires peak right at the beginning of the October and then decrease continuously until the end of May. Man-made fires as means to clearance peak later in November (Molinario, Davies, Schroeder, & Justice, 2013).

The local population in the western lowlands of Ethiopia is heavily involved in subsistence agriculture. For them, the availability of agricultural inputs like fertilizer, machinery or irrigation infrastructure is very limited. Half of the small-scale farmers cultivate plots of less than 1.5 ha. Most small-scale farmers practice shifting cultivation – after clearance, they cultivate the area for a period of four to eight years, after which it becomes fallow land and natural succession takes place. Most planted crops are sorghum, millet and maize, followed by rice and teff. Additionally, oil seeds like sesame, Niger-seeds and sunflower are planted, as well as mango, groundnut and cotton. Mainly due to the precipitation patterns, there is only one crop cycle per year with harvest in the second half of the year (Dorosh & Rashid, 2012).

Figure 3: Left: Plots of small-scale farming in BSG. (photograph taken by Matthias on the 20.05.2016); Right: Crop Calendar for Ethiopia. (FAO, 2016)

Agricultural practices related to LSAI contrast very much with small-scale farming. Here, larger sites are cultivated and inputs like fertilizer, machinery and irrigation schemes are used. The mechanization often leads to regular shapes of the cropland. The major crops cultivated by the investors are sesame, cotton, maize and sorghum.

Concept of the Remote Sensing Monitoring Tool

Many sensors installed on distant carrying platforms - usually a satellite, plane or unmanned aerial vehicle (drone), collect data about the surface of the earth. These sensors collect reflected electromagnetic radiation over large regions. Using specific workflows, these data can then be further processed into a digital image and finally into digital maps. The term “Remote Sensing” entails the whole system from data acquisition over data processing up to data analysis and interpretation.

The workflows applied and the information derived from RS can be of various degrees of sophistication. Thus, the complexity of a RS application can be managed. Different functionalities of a RSM, which satisfy different information demands of the user, can be regarded as “modules” and added on demand. An increasing information demand thus goes along with an increasing number of functionality, 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. Consequently, setting up a RSM does not need to be a “one-off” 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.

Accordingly, it is foreseen to establish a RSM at federal and regional levels incrementally, starting with a basic version, that has little technical requirements and that delivers basic monitoring results.

In the course of increasing capacities and technical knowledge, the RSM can be extended towards a more sophisticated option that has higher technical requirements (more input data, automated classifications), but delivers more enhanced monitoring results such as crop identification or crop condition estimations.

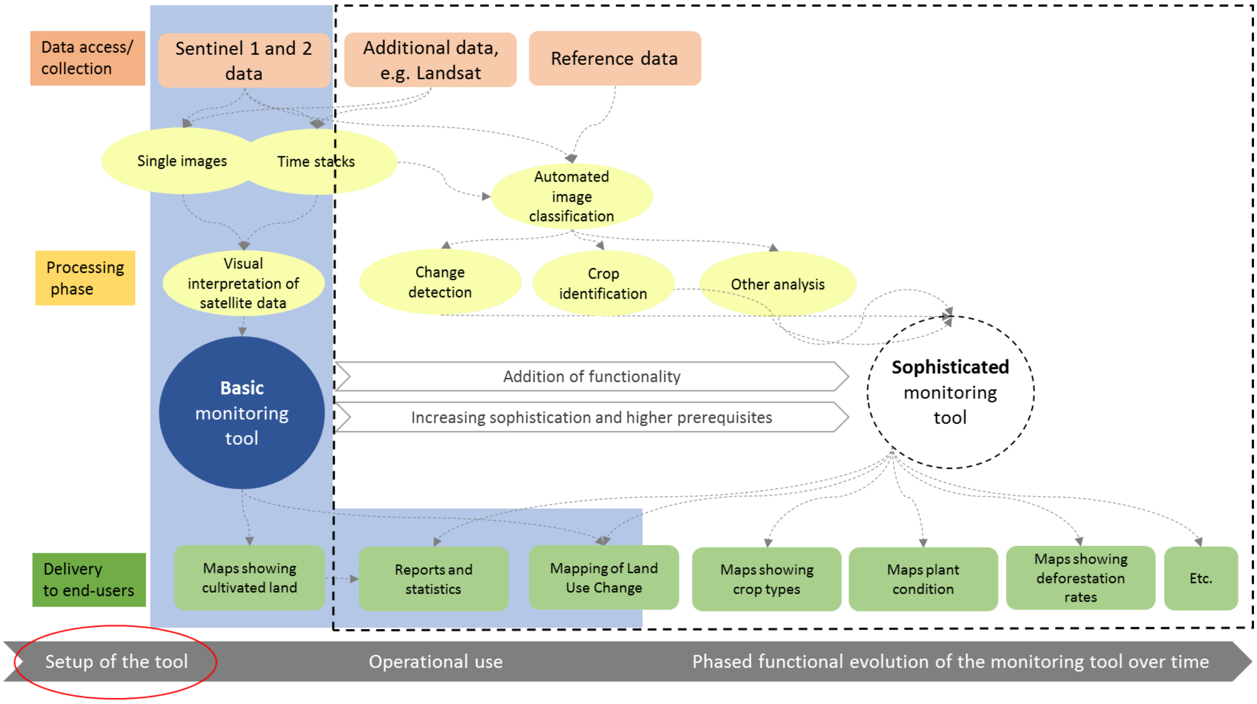


Figure 4: General workflows of the basic and more sophisticated versions of the remote sensing-based monitoring tool and their incremental implementation.

The texts below describe the workflow of the basic version of the RTM and first deliverables from testing this workflow.

Sentinel-1 and Sentinel-2 Data

Data from the Sentinel-1 - (S1) and Sentinel-2 (S2) - satellites of the *Copernicus Earth Observation Program* of the European Union[8] constitute a major asset for this kind of application and provide an unprecedented opportunity to monitor even fragmented agricultural landscapes with small-scale farming such as in Ethiopia. Time series[9] of Sentinel data reach expectations for agricultural monitoring in many regards:

- The S1 and S2 data are free of charge and publicly accessible.
- Frequent (for BSG and GB: S1 every 6 days, S2 every 5 days), systematic and full coverage with free, high-resolution data (for regarding application: S1: 20mx20m, S2: 10x10m pixel size) support agricultural monitoring even at the field level.
- S1's carries a Synthetic Aperture Radar (SAR) operating in the microwave C-Band, and two polarizations (vertical and horizontal). It is hence not affected by atmospheric

[8] *http://www.copernicus.eu*
[9] *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.*

conditions (clouds, aerosols) and can acquire imagery independent from sun illumination (descending orbits in the morning, ascending orbits at night).

- S2´s carries a Multi-spectral sensor, which collects electromagnetic reflectance of sunlight within 12 narrow spectral bands in the visual (blue, green, red), near infrared (red-edge, NIR) and shortwave infrared regions of the spectrum. This selection of spectral bands is very useful for discrimination of bare soil, vegetation growth and crop specific phenomena and events that affect the surface (e.g. burn scars, inundation, land conversion).
- 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.
- The Copernicus program foresees the simultaneous operation of two 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 provides a suitable basis for initiating a long-term monitoring tool in Ethiopia. Existing projects like JECAM [10] convincingly demonstrate the utility of RS for agricultural monitoring in various agricultural landscapes in Africa.
- The European space agency provides data that already went through basic steps of correction: S2 data is geometrically corrected to the “Top of atmosphere” radiance.

  For atmospheric correction of S2 and geocoding of S1, ESA provides software toolkits, to create GIS-ready imagery. This eases the workflow of the data user.

The whole data set of S2, which steadily increased by repeating image acquisition, is subdivided into so-called “granules”. A granule, also called tile, is the minimum indivisible partition of a product (containing all possible spectral bands). For Level-1C and Level-2A, the granules are 10,000 km^2 ortho-images in UTM/WGS84 projection (ESA, 2018). S1 is provided as frames of 185x185 km^2, cut along the orbit path.

A total of 20 S1-frames and 34 S2-granules cover the area of BSG and GB.

[10] *http://www.jecam.org*

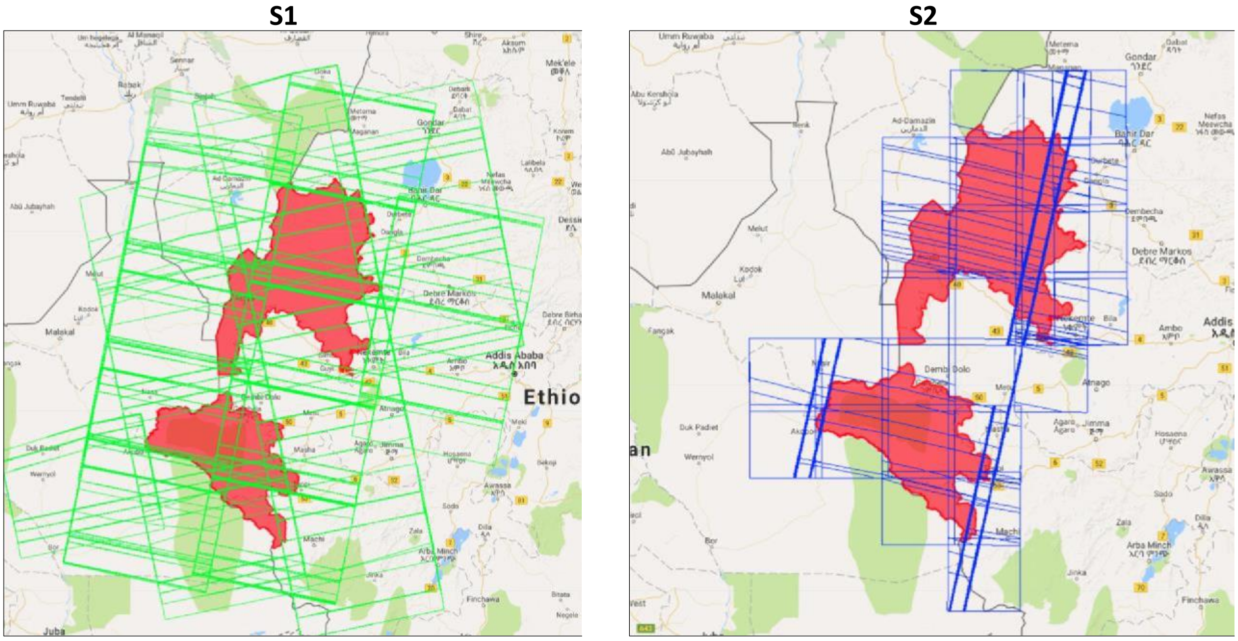


Figure 5: Coverage of S1 and S2 granules for BSG and GB. (ESA, 2018)

Data Access

S1 and S2 data are free of charge and publicly accessible. The data can either be downloaded or accessed and pre-processed in a cloud-computing environment, without downloading the input data. Both options have their advantages and disadvantages.

Downloading the data

There are different download hubs for S1 and S2 data. It is recommended to use the *Copernicus Open Access Hub*[11]. After registration, the image archive can be accessed and images can be searched according to certain criteria. Most important search criteria are 1) data type (S1 or S2) 2) time range for the acquisition date of the data 3) in case of S2, maximum cloud coverage. From the result list, single or a set of images can be selected and downloaded. Instead of manually selecting images for the download, this process can also be semi-automated by using the application-programming interface of the *Copernicus Open Access Hub*.

One single S1 granule is around 1.4 GB, one single S2 Granule around 0.7 GB. The amount of data that needs to be downloaded to cover both BSG and GB several times per year can easily increase, especially if time stacks are produced as input data for the analysis. Currently, internet connectivity both at federal and regional levels is rather poor, download rates are low and connection is unstable. Under these conditions, probably will not be able to download the needed data themselves. Thereby, downloading the S1 and S2 becomes a serious bottleneck in the workflow. As an intermediate solution, S2RAI could support the workflow by using the good internet connection in the German head office of GIZ to download the data on a hard drive, which then will be send to Addis. Same could be done by other cooperating institutions with better internet connection. However, this should only be considered as a fallback solution, more sustainable solutions that support the institutions independence from external support

[11] https://scihub.copernicus.eu/

should be envisaged. This could be realized by increasing internet capacities of governmental departments, enhancing collaboration amongst governmental departments or using cloud-computing environments.

Accessing data in cloud computing environment

Instead of downloading the data, there are also options to view, process and analyse the data in a cloud-computing environment. In addition, images, intermediate or final results of analysis can be downloaded from there. The major advantage of cloud computing for the regarded application case is that the user is independent from good download rates and own computing capacities, because the storage of data and its processing and analysis is done on the server of the cloud-computing provider. Considering the low bandwidth availability in this regard, cloud computing seems to be a valid solution. A bottleneck for this solution could be that it might be necessary to upload confidential data like the shape-files of the investment outlines onto the server to carry out certain analysis. The GoE might not agree on uploading such data on a server of an independent provider. Furthermore, programming skills are often required to make use of cloud-computing.

Data Pre-Processing

As all input and output data of the analysis is referring to spatial information, special geoinformation software or coding will need to be applied for further data processing and analysis. It is highly recommendable to create a spatially enabled database, e.g. using the open source Postgresql/Postgis, for spatial data storage. All necessary steps in the processing chain can either be done in a fully automated way by coding in programming languages (R/R-Studio, Python, Java) (often needed in a cloud computing environment like the *Google Earth Engine)*, or using open and free software like QGIS, SNAP Toolbox and other open source components to successively perform the processing steps manually on the workstation computers. For the starting phase of implementation, primarily open source software will be used that is easier to handle than applications that require programming skills.

Before the data is ready for visual interpretation, the images will be further processed for better visual discrimination of the various land use/land cover types.

Pre-processing of S1

Correction of S1 data

Generally, SAR data (S1) pre-processing consists of steps like terrain correction, orbit correction, thermal noise and border noise removal to ensure seamless inter-scene processing (in case a LSAI overlaps two or more S1 scenes), calibration to sigma naught by conversion of ground range detection (GRD) to radiometric calibrated sigma naught backscatter, and speckle filtering. Furthermore, the correction of the horizontal accuracy of S1 scenes through the precise orbit pass files, provided by the ESA S1 mission planning, is available at earliest 6 weeks after the last S1 image acquisition. With the open source software *SNAP S1-Toolbox* all these steps can be carried out either manually for each single image, or a workflow can be defined using a so-called visual programming tool, which will allow an automated processing of all selected images.

Display single S1 images

For agricultural analyses we are interested in backscattering intensity, which depends on the surface type and conditions, and the "coherence" between two successive S1 images, which will depend on how consistent the backscatter signal is over the time period between those images.

One S1-product actually contains two bands as separated raster-files, one for each combination of emitted and received polarization. Over land, the S1-SAR emits only vertical polarized microwave pulses and receives those in vertical and horizontal polarization, thus creating a VV and VH channel.

On-screen visualization of the SAR bands is given through RGB composition, e.g. VV in red, VH in green and a difference (or ratio) image of VV and VH bands in the blue channel.

In cases where only one polarization band is available, the VV image is shown as greyscale.

Time-stacks of S1 data

Principals of time composites:

Time composites (or time stacks) are a product of remote sensing data processing. Thereby, images from different point of times covering the same area are stacked to one image – for most applications in a chronological order. Each pixel then has several values, which capture the phenological change over time of the land cover type that the pixel covers. Discrimination of land cover types can then be based on richer data inputs and thereby becomes more precise and differentiating than classifying single images from only one point of time. Time stacks can be composed of radar as well as of optical RS data

Time stacks can be composed of radar as well as of optical RS data. Sentinel-1 time stacks are useful to highlight changes over time. This is particular useful if Sentinel-2 imagery is not available, because Sentinel-1 is insensitive to cloud cover. Thus, it is much easier to build consistent time stacks from Sentinel-1. Transitions from bare to vegetated state (or vice versa) are clearly visible in Sentinel-1 time composites as well as crop specific preparation and growth cycles (e.g. for rice, cotton, maize). Due to speckle, small parcels are less easy to delineate in Sentinel-1, but are typically detectable if they occur in clusters (e.g. an area with smallholder cultivation).

In QGIS, time stacks can be built via the raster menu option "Build Virtual Raster". For visual analysis, one will typically select 3 images for 3 successive dates, to display in the RGB channels of a computer video screen. Changes over time are then visible in the combination of the prime additive colors red, green and blue. Deeper time stacks can be built with more than three channels, for instance for selection of temporal signatures for specific locations or parcel averages.

The image below is a time composite of S1 images from 23rd of October 2017 (in the red channel), 10th of December 2017 (in the green channel) and 15th of January 2018 (in the blue channel) showing agricultural activities in the Afar Region. The bands for the VV-polarization are used at 10-meter resolution. The bright yellow to red colors highlight a cropping pattern for which harvesting has lowered the VV backscattering in the last period (the blue channel). The bright blue parcel in the lower dark cultivated area is a crop that emerges in the December-January period, in an apparent paddy rice area (the dark colors are due to low backscattering of flat, possibly inundated fields. The reddish area to the eastern side of the image is natural vegetation that dries out with the onset of the dry season after October.

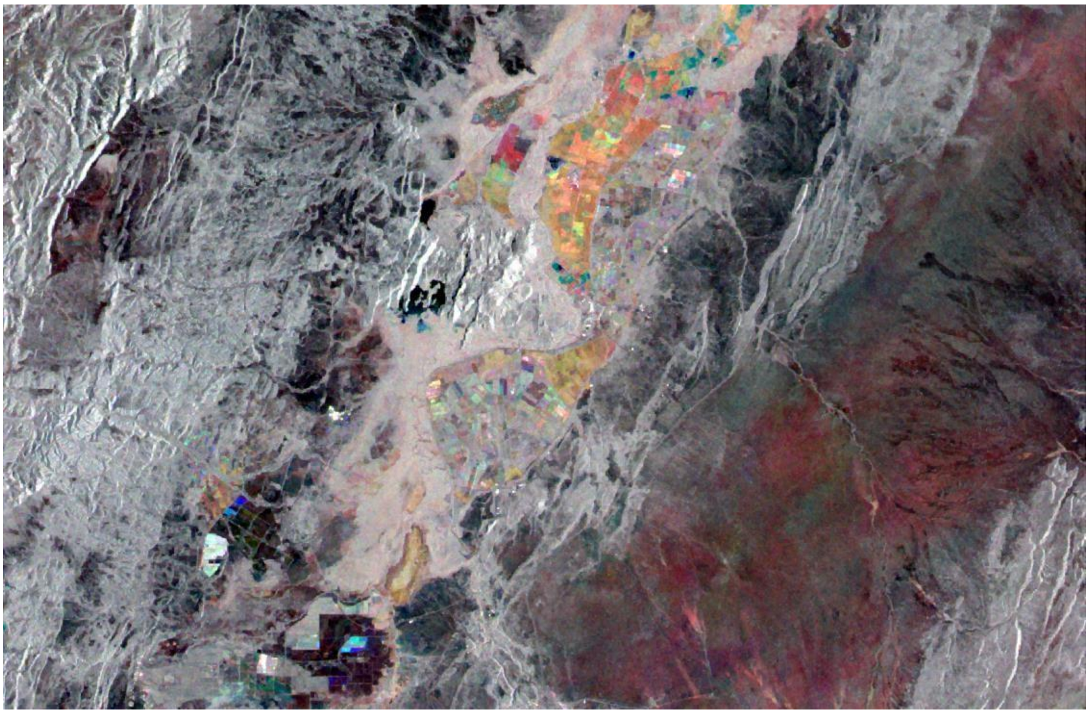


Figure 6: S1 time composite showing agricultural activities in Afar Region.

Preprocessing of S2

Correction of S2 data

For optical data, information about the earth surface in raw satellite images is encoded as so-called digital numbers (DN), which need to be converted to top-of-atmosphere (TOA). The data provided by the *Copernicus Programme* already have this level of pre-processing. Sentinel-2 imagery is already very well georeferenced and is always projected in the local UTM coordinate reference system (unfortunately, Sentinel-2B is shifted one pixel south compared to Sentinel-2A, which will be corrected in the future). For the analysis of data under the basic version, no further correction work needs to be done on the data[12].

[12] For more sophisticated analysis it might be necessary to further process the data to top-of-canopy (TOC) reflectance values via atmospheric correction algorithms (thereby reducing the effect of dust and aerosol scattering) before further processing and analyzing.)

Display single false color composites of S2

One S2 image actually consists of 13 separated image-files – one image-file for each of the 12 bands that cover different parts of the wavelength plus 1 band for a quick look in true colors.

Different band combinations loaded into the R, G and B channels will reveal different aspects of land cover. Vegetation and its' different growth status have significant reflectance values especially in the red and near infrared spectrum. Therefore, S2 bands 4,5,6,7 and 8 (so called red edge bands) are most suitable for visual interpretation of cropland. But also band 12, 11 or 3 might be of interested. A set of combinations of bands displayed in the RGB-channels should be applied for visual interpretation, amongst them: R: 4 – G: 3, B: 2 (natural color), R: 8 - G: 12 - B: 4 (false color), R: 8 - G: 4 - B: 3 (alternative false color) and R:7 – G:6 – B:5 (red edge).

Calculating indices

Further processing steps for the S2 data include computing the vegetation index "normalized difference vegetation index" (NDVI)[13], using a standard formulation applied to the S2 red (B4, 665 nm) and vegetation near-infrared (B8, 842 nm) bands. NDVI has proven to be useful in discriminating difference land cover and crop types.

Time-stacks of S2 data

As a subsequent step, multi-temporal image compositions out of a set of NDVI-images should be created.

Therefore, all S2 images with less than 20% cloud coverage available for a three-month period are selected. For each image, the NDVI is calculated. Then all NDVI images are stacked and the mean value for each pixel is calculated. The output is an image file (raster), that describes the average vegetation density and activity for each pixel for one quarter of the year [14]. The same should be done for subsequent quarters of the year.

Data Analysis

Once the input data went through the described pre-processing, it is rather easy to visually interpret the image. Nevertheless, background knowledge on prevailing vegetation and crop types and their typical vegetation periods and crop/cultivation cycle are needed to correctly interpret the data. In addition, knowledge on the weather conditions, especially the precipitation patterns, of the analyzed year are valuable asset for interpretation. Basic background knowledge for interpretation is provided under section "Eco-zonal and agricultural characteristics of the area of interested". More precise background knowledge will enable deeper interpretations.

---

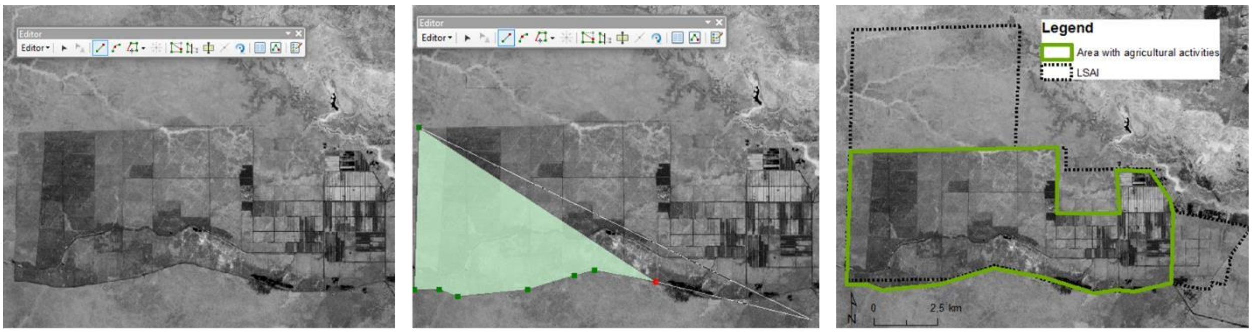
[13] *NDVI is a dimensionless index that is indicative for vegetation density and is calculated by comparing the visible and near-infrared sunlight reflected by the surface (reflectance).*

[14] *The compositing period can be adjusted, depending on the coverage, to enhance the availability of composites (e.g. 2 monthly or monthly).*

The concept of deriving monitoring results from visual interpretation is the same for the different input data, rather single S1 or S2 images or time composites are used (see figure 7).

1. The interpreter visually identifies the cropland according to specific characteristics like regular shapes, specific colors, color intensities[15] or contrasts. The interpreter might also identify other land uses/cover like small scale farming, housing, burned areas, etc. or even distinguish crop types (in case of interpreting time composites). Obviously, as an important reference, the outlines of the investment sites should always displayed as a layer above the image.
2. The interpreter draws a polygon (vector data) around the identified cropland or other land uses/covers of interest. Such a polygon already depicts a quantitative measurement of cropland. As the outlines of the investment sites are laid over the images, the interpreter can visually estimate if and how much cropland or other land uses/cover can be found within or around a certain investment site.
3. Simple calculations from the derived polygon and the polygon of the investment site can be carried out to derive quantitative information on the share of cropland in the whole investment site.

Interpreting single S1 or S2 images is less challenging. The pre-processing phase is shorter and requires less expertise (see figure 7). The interpretation is more intuitive, as the interpreter does not need to comprehend the logic of time composite images when identifying cropland. However, single images contain less data; differences between cropland and natural vegetation are less visible than with time composites. Mix-up of the different land cover types and thereby misinterpretations are more likely, especially if the agricultural area is not characterized by straight boundaries and regular shape (see figure 7).



Visual interpretation: Identification of structures indicative of agricultural activities (**green arrows**), e.g. clearance, fields, canal, rods, etc.

On-screen digitization of the identified structure in a GIS (**green polygon**)

Final product: map containing information about agricultural activities (green polygon) within LSAI. Uncultivated LSAI clearly visible (**red arrow**)

Figure 7: Visual identification and delineation of cropland within an investment site in GB based on a single NDVI image.

Each land use/cover type has its characteristic phenological cycle throughout a certain time. Time composites describe these phenological changes of the different land use/cover types - in the case of S1 data by measuring changing roughness in surface and surface cover (vegetation), in case of S2 data by measuring changing electromagnetic reflectance. Consequently, with time composites of S1 or S2 images, land use/cover types can better be

[15] *Depending on the input data, the colors or color intensities describe different characteristics.*

distinguished (e.g. cropland from natural vegetation) then with single S1 or S2 images. Even different crop types can be distinguished with time composites if the interpreter has profound knowledge on different crop cycles. However, pre-processing of time composites and its interpretation is more challenging and demands more background and technical knowledge.

For testing the RSM process, an S2 and NDVI composite from S2 images from 29th of July, 27th of Oct and 26th of Nov was composed (see figure 8, left side). It shows cultivated crop fields as green colors, i.e. the crops are at full cover in October and being harvested in November. Significant parts of the investment is not yet cultivated (greyish areas) and most cultivated fields do not follow the outline of internal parcels included in the polygon outline. Some cultivated fields "spill over" into areas outside the investment outline (e.g. Northeastern edge).

For the sake of comparing, a time composite from S1 VH bands for the same periods has been composed – but not three single images were used, instead several images for each of the month (July, October and November) were stacked and averaged (see figure 8, right side). Parcel outlines are somewhat less distinct than in the S2 time composite, due to speckle, but also because this area has significant relief, which is not as apparent in the NDVI time composite. Still, several land use classes can be delineated and related to detection of agricultural activity. More intensive pre-processing of the S1 time composite might have increased the visibility of differences between various land uses.

NDVI time composite from S2 data VH time composite from S1 data

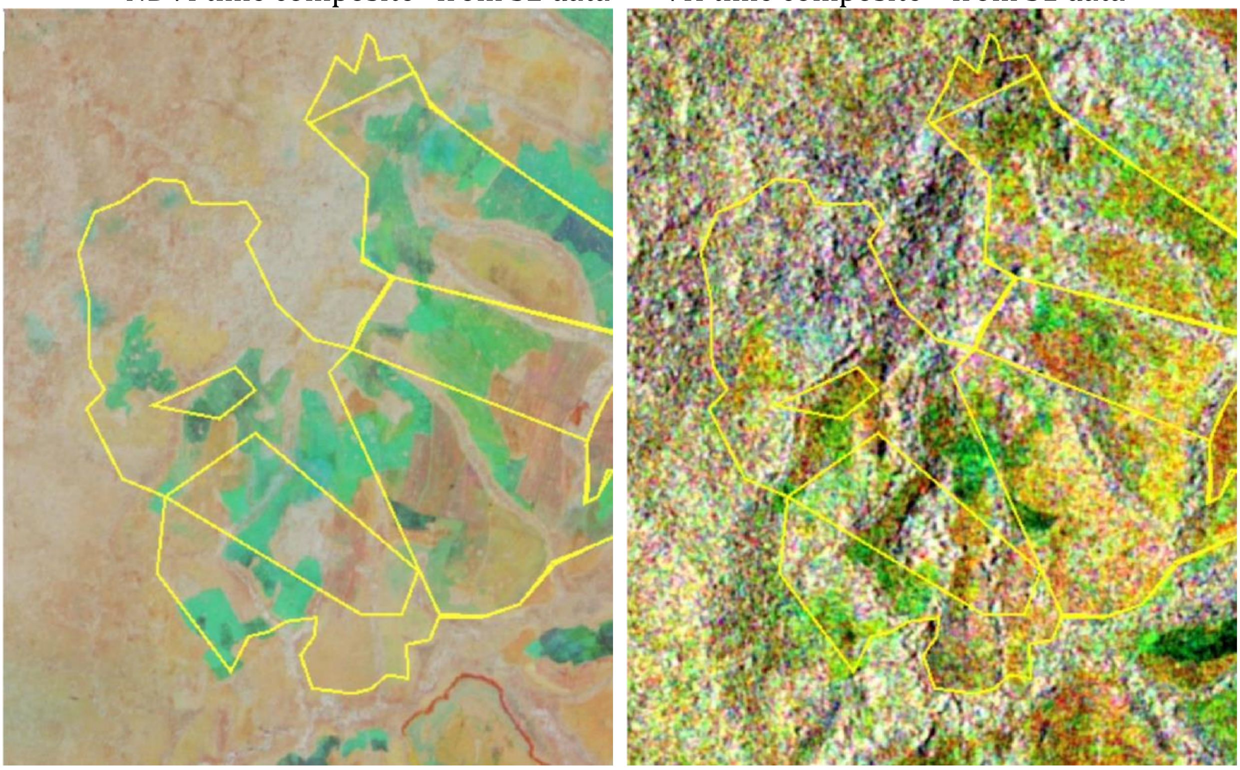


Figure 8: Comparison of time composites from S2 and S1 data. (Matthias Hack)

Implementation Status and Way Forward

The initial steps of implementing the RSM at BSG have already been undertaken. The users' needs and existing capacities have been identified. Accordingly, the concept for the RSM has been developed as described in the previous chapters. Operationalization started by providing a first training to technical staffs from federal and BSG institutions in February 2018.

Figure 9: Training of EHAIA and LAUD staff on the basic version of the RSM, 19-22.02.2018, Addis Abeba. (Photo by Matthias Hack)

As next steps of operationalization, responsible federal institution and BSG staff has examined a considerable number of LSAI by applying the workflow of the basic version of the RSM. This will help reinforce the gained knowledge, to test the basic versions and to gain experience that will be used in further development of the RSM.

The first basic version of the system has been developed using Google Earth Engine during the first year and is operational. A total of 2–3 years has been scheduled for the complete implementation of the tool. However, it should be noted that the implementation of the monitoring tool should not be regarded as a one-off process – it should be seen in the context of opportunities for on-going updating (e.g., adaption of new methods for classification, general software upgrades), sporadic upgrading (e.g. incorporation of new reference data for adjusting the calibration of methods and improving the accuracy) and incremental improvement whenever relevant or necessary for fulfilling the objectives of the tool.

After the initial setup-phase of the monitoring tool, resource savings are possible because (i) a well calibrated method theoretically works without further calibration, provided that temporarily regular satellite time series can be provided (which will be the case for the Sentinels), and thus the need for field surveys could be minimized. Furthermore, (ii) the processing chain can be largely automated using coding. Still, the validation of the map products and the usability of the monitoring tool is an iterative process and should remain so in the longer term parallel to the development of the tool and consider integrating the availability of new classification methods (e.g., decision fusion methods or image segmentation), RS sensors (e.g., new Sentinel missions) and modelling techniques. Specifically, it is recommended that at least one training per year after the initial should be accomplished in order to keep pace with these technical developments and to further improve the tool (e.g.

improving field surveys, if envisaged, and the product dissemination according to end-users´ needs). The figure below summarizes recommended some basic steps to be undertaken especially during the setup phase in the initial 1-3 years.

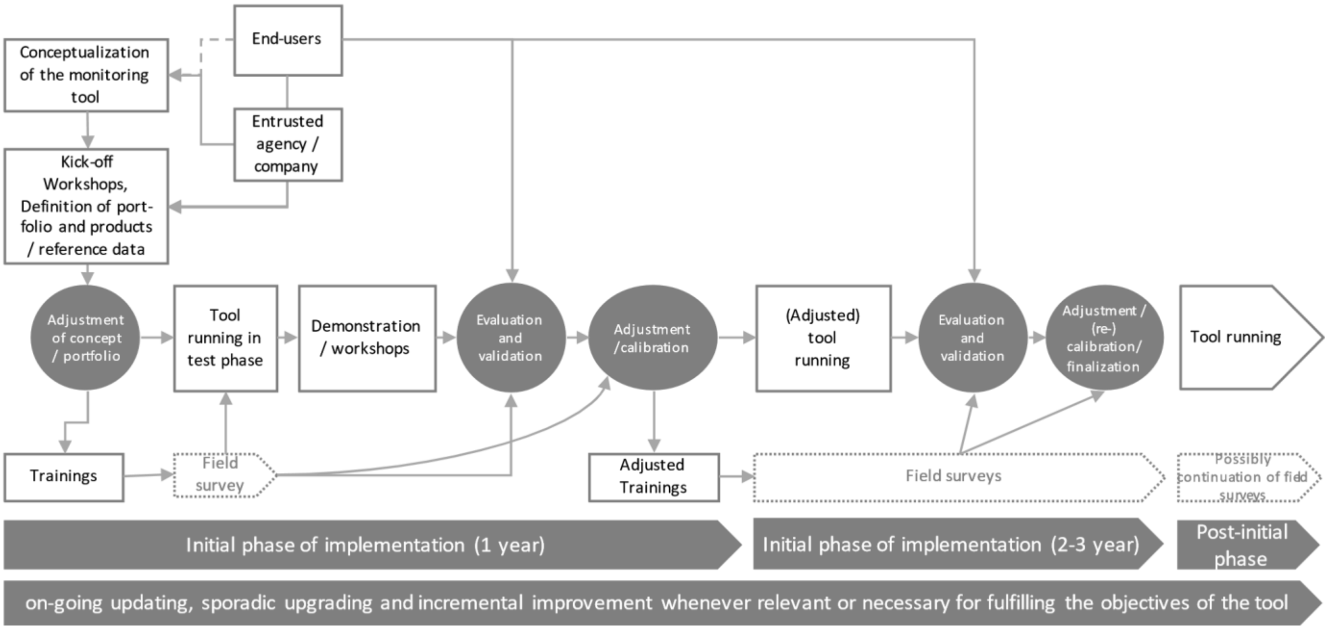


Figure 10: General workflow of tool conceptualization, validation and adjustment (Adjusted from GMFS[16]).

Conclusion

S2 and S1 data are valuable data source for monitoring LSAI in Ethiopia: Spatial and temporal resolution of the data is high enough to detect agricultural land, furthermore other major land cover and land use can be distinguished.

Another advantage is that there are no procurement costs for the data. Thereby, important monitoring questions concerning the spatial evolvement of LSAI could be answered with little financial and technical effort.

Computing capacities and technical knowledge required for the workflow of accessing, processing, analyzing and interpreting the data are an obstacle for federal and regional states of Ethiopia. Starting with a basic version of such monitoring tool that requires less technical knowledge, would incrementally introduce the use of Sentinel data and thereby match the prevailing capacities. Using single NDVI images from relevant points in time seems to be the best starting point. Such images already entail valid information and are easy to produce and to interpret. As a second step, time composites of several NDVIs could be generated and analyzed. This requires more knowledge, but also delivers enhanced results. As a third stage of sophistication and improvement, S1 time composites could be generated and used for visual interpretation.

[16] www.gmfs.info

References

Dorosh, P., & Rashid, S. (2012). *Food and Agriculture in Ethiopia.Progress and Policy Challenges.* Philadelphia: University of Pennsylvania Press.

ESA. (20. 01 2018). *Data Product*. Von Sentinel Online: https://sentinel.esa.int/web/sentinel/missions abgerufen

FAO. (14. 09 2016). *GIEWS - Global Information and Early Warning System*. Abgerufen am 27. 10 2016 von Country Brief: http://www.fao.org/giews/countrybrief/country.jsp?code=ET

Federal Democratic Republic of Ethiopia. (2010). *Growth and Transformation Plan (GTP) 2010/112014/15.* Addis Abeba: Ministry of Finance and Economic.

Google Inc. (satellite imagery from different acquisition dates). Google Earth Pro.

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

Merkel, A. (06. 08 2016). *Climate-Data.org*. Abgerufen am 27. 10 2016 von Climate: Asosa: http://en.climate-data.org/location/32103/

Ministry of Agriculture and Natural Ressources. (12. 12 2016). *As the Benefit of Agricultural Investment is Multipronged Give Not Ear to Doomsayer Bloggers!* Von Agriculture investment Agency: http://www.moa.gov.et/de/agricultural-investment-directorate abgerufen

Molinario, G., Davies, D. K., Schroeder, W., & Justice, C. O. (2013). Chararterizing the spati-temporal fire regime in Ethiopia using the MODIS-active fire product: a replicable methodology for countrylevel fire reporting. *African Geographical Review*.

Peel, M. C., Finlayson, B. L., & McMahon. (2007). *Updated world map of the Köppen-Geiger climate classification*. Von Spatial data access tool: http://webmap.ornl.gov/ogcdown/dataset.jsp?ds_id=10012 abgerufen

Schultz, J. (2002). *Die Ökozonen der Erde.* Stuttgart: Eugen Ulmer GmbH & Co.

Acronyms

BSG	Benishangul-Gumuz
EHAIA	Ethiopian Horticulture and Agricultural Investment Agency
GB	Gambela
GIZ	Deutsche Geselllschaft für Internationale Zusammenarbeit GmbH
GoE	Government of Ethiopia
JRC	Joint Reseach Center of the European Commission
LAUD	Land Administration and Use Directorate
LSAI	Large-scale agricultural investments
RS	Remote sensing
RSA	Remote sensing analysis
RSM	Remote sensing-based monitoring tool
S2RAI	Support to Responsible Agricultural Investments