### <u>Annex 5</u>

## **Excerpt of Remote Sensing Chapter from:**

## HARITA-IRI Report to Oxfam America, June 2010 IRI Technical Report 10-8

Full report available at: <u>http://iri.columbia.edu/publications/id=1002</u>

Note:

The proposed ILO project would perform the activities suggested in section 4.4, which are currently unfunded.

# tech report **10-08**

HARITA IRI Report to Oxfam America June 2010





**Contact** Daniel Osgood, deo@iri.columbia.edu

# HARITA IRI Report to Oxfam America

#### Final Report for IRI MIEL Planning & Technical Support for HARITA Micro-Insurance Pilot USA 536 /09

June 2010

Project PI: Daniel Osgood (deo@iri.columbia.edu)

Contributors: Jiehua Chen, Tufa Dinku, Eric Holthaus, Bradfield Lyon, Malgosia Madajewicz, Connor Mullaly, Michael Norton, Daniel Osgood, Angelica Pascualini, Saurabh Pant, Nicole Peterson, Ricardo Saavedra, Jessica Sharoff, Kenneth E. Shirley, Christopher Small, Sarah Stiffel, Tse-Ling Teh

This report was funded by Oxfam America, the Rockefeller Foundation and the US National Oceanic and Atmospheric Administration (NOAA), which provided its support under cooperative agreement NA050AR431104. Additional funding sources include: US National Science Foundation's cooperative agreement NSF-SES 0345840 with CRED, the Earth Institute, the Applied Statistics Center at Columbia University and U.S. Dept. of Agriculture National Institute of Food and Agriculture CSREES 2006 35615-1671. Some report elements are adapted from student activities in EAEE E4010.001/Engineering 252 "Remote Sensing and Environmental Change" at Columbia University, taught by Michael Bell and Pietro Ceccato and utilize the IRI data library.

> IRI Technical Report 10-08 http://iri.columbia.edu/publications/id=1002

### 4 Remote Sensing

### 4.1 Introduction

In creating robust and accurate index insurance contracts for the HARITA project, it is critical to have a reliable way to estimate crop growth/loss, in order to design contracts and identify triggers. While farmer feedback is important in determining poor growing years, it also remains crucial to build a system that relies on verifiable data. On-the-ground sensing instruments, such as rain gauges, are valuable scientific tools, however these too present some complications. These technologies can be expensive to implement and limited geographic spacing can create difficulties in contract design. The use of remote sensing (satellite) technologies to leverage the limited amount of ground-based information may make the up-scaling of weather index insurance feasible, as long as robust, responsible, and cost effective techniques can be employed for calibration and validation.

Currently, the HARITA project utilizes ARC satellite rainfall estimates to trigger payouts. It is important to identify the areas where remote sensing of rainfall is not an accurate mechanism to build contracts. Sometimes ARC data has difficulty reflecting actual rainfall and/or crop loss. For example, geographic boundaries, such as mountains, where rainfall measurements vary greatly due to localized weather patterns, can limit reliability. It is essential to identify locations where these problems may occur. To address this problem we are investigating various other remote sensing tools that estimate greenness, the vegetative vigor of plants. These are often referred to as vegetative indices, and are described further below. Our eventual goal is to develop processes through which remote sensing of vegetation can leverage the minimal ground level information available to accurately perform this role, and perhaps to be utilized as part of the index product.

The remote sensing activities in this report can be organized into three categories

- 1. Satellite estimates of rainfall for index calculations.
- 2. Exploratory validation of satellite rainfall estimates using off the shelf satellite vegetation sensing products.
  - (a) Sites: Hade Halga, Geneti, Awat Bikalsi, Adi Ha, and Hadush Adi
  - (b) Years: 2000-2009
  - (c) Season: Late summer (phase 2 of contracts, flowering and filling phases for crops)
- 3. Determination of what the vegetative satellite imagery actually represents so that it can be appropriately utilized and improved for validating and mapping problem areas in satellite estimates of rainfall. This activity includes:
  - (a) Ground verification activities comparing the dry and wet seasons in Adi Ha

(b) Exploratory use of rare but high resolution satellite images to determine if they can be used to reduce the amount of necessary ground verification information

### 4.2 Remote sensing of rainfall

The 2009 contract used remotely sensed estimates of rainfall to determine payouts, and we have been using the same products to develop all of the indexes for 2010. The product used is NOAA CPC ARC, available directly from NOAA as well as through the IRI data library. The draft indexes presented in the 2010 contracts are all based on remote sensing of rainfall. In addition, IRI has been working with the Ethiopian Met service and other partners to develop more accurate remotely sensed estimates of rainfall, and to expand the length of the of the data historical series from 15 years to 30 years. In this activity, the Ethiopian Met service will utilize their archive of historical data to develop improved and extended satellite rainfall estimates.

Much of the analysis for remote sensing of rainfall is the development of systematic statistical methods to combine and compare and validate the rainfall data against vegetative sensing, ground measurements of rainfall, and other data sources. See the section on statistical methods for more description of these activities.

Ground measured rainfall is the most critical measure to compare satellite rainfall estimates. Figure 10 shows the location of available ground based met stations in relation to the sites for 2010. Many differences between the dataset are not problematic for index insurance so long as the index can be strategically designed to avoid having these differences lead to missing or inappropriate payouts. For example, differences in exact daily rainfall levels are not critical if the total over the contract window reflects droughts. In addition, if the rainfall estimates are systematically higher or lower than ground observations, that is not a problem if the years with low precipitation estimates in the satellite data sets are the same as those in the ground measurements. The figure includes the three new automated met stations installed as part of the HARITA project.

### 4.3 Off the shelf vegetative sensing satellite products

One emerging methodology for overcoming sparse field data is to measure levels of precipitation and/or vegetation from a remote sensing platform. Currently, the HARITA utilizes ARC data for index payouts, a remote-sensing product used to estimate rainfall amounts. In previous years, these values have been compared to station gauge data at Maychew, Adishewshu and Alamata to test for accuracy. However, remote sensing tools that measure levels of plant growth, or vegetation, may also represent what is happening on the ground and may provide another means of verification, as well as a possible future source for deriving contracts. In order to test this idea, we compared five different off the shelf satellite-based



Figure 10: Geographic placement of new villages and existing rain gauges.

indices, of which four are derived from the presence of chlorophyll (or the greenness in vegetation) and one from the presence of water in the vegetation; each of these products varies spatially and temporally in resolution.

NDVI is the most widely used satellite vegetative index. It is available from the early 1980s to the present for most of the world, with several images per month. NDVI stands for Normalized Difference Vegetation Index and senses the greenness in a given area by measuring the wavelengths of light that are reflected off the surface. Plants have evolved to absorb some wavelengths of light for photosynthesis as well as reflect those wavelengths of light which are harmful to them. By measuring the relative ratio of the two wavelengths, we can passively monitor vegetation on the earth's surface from a satellite at various points in time.

There have been some index insurance projects that use indexes like NDVI for the basis of their contracts, such as the project by Chris Barrett of Cornell in the rangelands of Kenya. At IRI we have designed NDVI triggered contracts for village level index insurance contracts for nearly a dozen countries in Africa for the MVP project. These contracts have been transacted in Kenya and Ethiopia. NDVI is very different from rainfall indexes, especially in a place like Ethiopia which has much mixed vegetation such as bushes and trees which could confuse the measurement. Also, factors such as soil reflectance, solar angle, cloud cover, satellite angle, and differences between satellites must also be taken into account.

We are conducting these investigations not only to gain more insight of conditions on the ground, but also with an eye to potential enhancements in the contract. Depending on the results of the research, it might be advantageous to include NDVI data in our contracts so

the the final product is a mixture of rainfall and NDVI data. Currently, we are utilizing NDVI as an independent source of information to validate the performance of rainfall based indexes, including those based on remotely sensed rainfall estimates. We are also very interested in using vegetative sensing to map the area for which a given rainfall based index can accurately protect. NDVI data is easily accessed through the IRI's data library most of the globe, which allows us to study the data for our various sites in Ethiopia, and provides data delivery avenues convenient for going to scale.

These indices include:

- 1. Normalized Difference Vegetation Index (NDVI-AVHRR) from the National Oceanic and Atmospheric Administration (NOAA), which is measured every ten days at an 8km spatial resolution.
- 2. NDVI-Spot Vegetation Index from the Centre Nationale d'Etudes Spatiales (CNES), which is measured every ten days at an 1km spatial resolution.
- 3. NDVI/MODIS Vegetation Index from National Aeronautical and Space Administrations (NASA), which is a 16-day composite index measured at a 250m spatial resolution.
- 4. Enhanced Vegetation Index (EVI) derived from the same sources as NDVI/MODIS, but has a different quantitative relationship that is more sensitive to canopy structure and type.
- 5. Normalized Difference Wetness Index (NDWI) also derived from the same sources as NDVI/MODIS, but reflects changes in the vegetations water content. The MODIS-derived indices (number 3-5) are only available starting from the year 2000, while the NDVI-Spot vegetation index is available from 1998; NDVI-AVHRR is used starting from 1995. Thus, when comparing these sources of information we start with the year 2000. It is very important to have a long enough time series in order to capture the frequency of drought. Unfortunately, the highest quality products are new and data collection has not occurred for as long as desired. Therefore, it is important to try to figure out the reliability of older products and a robust combination of both tools. The comparison of these different vegetation indices against the rainfall gauges and the ARC satellite-derived rainfall estimates can help understand the applicability of vegetative remote sensing tools to incorporate into the design of index insurance products.

These vegetative index satellite products have substantial limitations for their use. Many products are distorted by satellite view angle, solar illumination angle, atmospheric dust and humidity, bare earth, clouds, changes in satellite sensors, compositionally mixed pixels and other factors. Since these products measure the vegetative vigor (or for some products, the soil moisture) over a particular region, in order to understand what they are indicating, it is important to understand the land cover in the region. Typically, the agricultural crop being insured is not the majority of the vegetation. Instead, the products see mostly grasslands,

trees, or bare earth. Understanding what is being observed is necessary to use the products. In addition, many crops (such as maize) can have vigorous, green foliage without producing much grain. The satellite vegetation indices are therefore not used as direct measurements of the health of the crop. Instead they are used strategically to proxy the response of the landscape to rainfall. For example, the index may be used to determine if catastrophically low levels of rainfall occurred during a key month for the crops being insured, by observing if the surrounding grasslands lost their vegetative vigor in the month following key rainfall.

For the initial diagnostic of these products, we focus on the second phase of the indexes, and the vegetative sensing in the month following the rainfall. Typically, the vegetative response to rainfall lags the actual rainfall by about a month. We compare

- Capped satellite estimated rainfall for the second phase of the contract. (ARC)
- Capped rainfall from nearby raingauges. Note that these raingauges are expected to experience different rainfall patterns from the index site, either due to distance or features such as mountains. The raingauges were not considered sufficiently related to the index site to be workable as triggers for payments.
- Initial farmer design team reporting of "bad" years. Note that these years include problems related to sowing, and in some cases are not specific to the crop being presented. In follow up design meetings "bad" years were identified by the farmers in much more precise terms. We hope to repeat the presentation utilizing the updated information.
- WRSI, a useful, but limited agronomic water stress calculation used in the design software, based on the ARC rainfall estimates.
- The remote sensing indices listed above for one month after the rainfall window.

It is important to note that this exercise was performed in parallel with the index design and used in discussions to improve the indexes. Therefore the index windows used for rainfall ranking in this analysis are not necessarily the finalized contracts. Resources permitting, we intend to repeat the analysis with the finalized information.

As would be expected, results from using the off the shelf vegetative sensing products and nearby raingauges are mixed. The datasets reflect many of the major events as bad years. However, many major events are not reflected in the indexes and the indexes have a substantial amount of disagreement between them. Table 13 is an example of one of these comparisons, presenting the information for Teff in Hade Alga (the technical annex presents these comparison tables for more indexes and sites). We compare the rankings to see if the worst years are reflected in the other datasets. This past year was the worst or second worst year across all of the available datasets, and was noted as "bad" by the farmers. Widely recognized as being a drought year, 2004 is the worst in the ARC dataset, the eighth worst year in two of the raingauge datasets, the fourth worst year for the NDVIg-rg (the vegetative

Yr	ARC	Maychew	Adisheshu	Alamata	Farmer	WRSI	NDVIg-rg	SPOT-Veg	MODIS	NDWI	EVI
2000	8	1	6	1	-	6	1	1	1	1	1
2001	3	3	1	8	-	7	2	1	4	7	3
2002	6	2	9	3	-	5	4	3	2	2	2
2003	1	6	7	9	-	2	5	4	3	8	4
2004	10	8	8	5	Bad	9	7	6	5	9	6
2005	5	9	4	6	-	4	6	8	8	6	8
2006	7	5	3	4	Bad	8	3	7	6	5	7
2007	4	7	2	7	Bad	3	8	5	7	3	5
2008	2	4	5	2	Bad	1	9	9	10	4	10
2009	9	#N/A	#N/A	#N/A	Bad	10	10	10	9	10	9

Table 13: Hade Alga Teff Ranking.csv

	Adi Ha Maize	AdiHa Teff	Hade Alga Teff	Geneti Teff	Hadush Adi Wheat
NDVI	2	2	2	1	1
SPOT Veg	3	2	2	2	1
MODIS	1	0	1	1	2
NDWI	2	2	2	1	1
EVI	2	1	2	2	2

Table 14: 3 in 5 Worst Years.csv

index with the longest history but the least advanced), and the second worst year for the NDWI, a wetness index. It is important to note that 2004 is in the worst half of the other indices, but is not indicated as a particularly bad year.

Interestingly, 2008, is an extremely bad year for all of the vegetative indices (except for the wetness index) while all of the rainfall measures indicate either average or very high levels of rainfall. Since the farmers reported 2008 as a bad year, it may be that the vegetative indices are picking up a meaningful signal that is for reasons somewhat different than simple measures of rainfall. It may be, that with more research, vegetative indexes may play a valuable complimentary role to water measurements in creating a more effective index in the future. In the Millennium Villages Project, for a site with more data for design than is available in the HARITA project, we developed an index that was a blend of remote sensing of vegetation and ground based rainfall measurements. This index outperformed both of the component indexes.

Table 14 presents a summary of the vegetative indices across the sites. It presents a sum of how many of the three worst ARC years are represented in the worst half of the different indices. One can observe that in most cases about two of the worst ARC years fall in the worst half of the years of any vegetative index, although there are many cases for which there are fewer years. It is not clear that any particular vegetative index greatly out performs another, although SPOT Veg and EVI tend to have the highest level of agreement with the ARC rankings.

This initial analysis of off the shelf vegetative products serves to motivate our research into

the improvement of vegetative sensing. It suggests that in spite of their known limitations there may be real promise in the use of vegetative sensing to validate satellite rainfall indexes if their behavior can be better understood. The vegetative indices may also at some point be valuable in improving a precipitation based index. However, at this point it is difficult to get robust or consistent performance from any of the off the shelf products. this work also shows that it will be critical to continue this work to improve upon and better understand the performance of existing products in order to provide robust, responsible, and reliable metrics.

#### 4.4 Project research activities on remote sensing of vegetation

While simple vegetation indices can indicate the presence, absence and in some cases change in vegetation cover, these indices are vulnerable to bias from exposed soil and cannot generally distinguish between indigenous vegetation and agriculture. The difference is important because agriculture and indigenous vegetation generally have different phonology (green-up and senescence periods) and response to rainfall. The wide field sensors used to monitor temporal changes in vegetation indices are also limited in their spatial resolution to pixel dimensions of 250 x 250 meters or more typically 1000 x 1000 meters. Hence, the vegetation imaged in most pixels will be some unknown spatial mixture of agriculture and indigenous vegetation. However, both the spatial heterogeneity of mixed pixels and the distinction between agriculture and indigenous vegetation can be addressed with the use of additional imagery with higher spatial resolution and more sophisticated methods for estimation of vegetation type and abundance.

We are developing a strategy of multi-resolution spatial-temporal analysis with telescoping validation. This approach addresses the issues described above by using a combination of high temporal resolution (daily to 8 day composite), low spatial resolution (250-1000 m) MODIS imagery with moderate spatial resolution (30 m), low temporal resolution (16 day) Landsat imagery and high spatial resolution (2.4 m) Quickbird imagery collected on specific dates. The MODIS imagery resolves spatial variations in the temporal phenology but cannot distinguish the relative abundance of agriculture and indigenous vegetation. The Landsat imagery can distinguish different types and abundances of vegetation at plot scales but only as snapshots on specific dates. The Quickbird imagery can resolve individual trees, bushes, herbaceous understory and agriculture but only on the date the image is acquired. We are able to combine the information provided by each sensor into a single spatial-temporal map of vegetation type, abundance and phenology through the use of spectral mixture analysis (SMA).

SMA provides a robust, scaleable, and verifiable method to identify spectrally pure endmembers (e.g. illuminated foliage, soil, water, rock, shadow) and estimate the spatial abundance of each within spectrally mixed pixels (Adams et al, 1986; Smith et al, 1990). SMA can be applied to both Landsat and Quickbird imagery allowing the field-verifiable estimates from the Quickbird imagery to be used for quantitative validation of the Landsat-derived

estimates of the same endmember abundances at 30 m resolution. Accuracy of vegetation abundance estimates from Landsat have been quantified at 94% in previous studies (Small and Lu, 2006). The telescoping validation approach uses field observations to verify vegetation type in Quickbird imagery then uses Quickbird-derived mixture fractions to validate Landsat-derived mixture fractions to distinguish different vegetation types and abundances at regional scales.

We expect that validation of Landsat fraction estimates in both wet and dry seasons should allow for distinction of indigenous vegetation and agriculture at plot scales. This will allow us to combine seasonal maps of vegetation type and abundance from Landsat with spatial phenology maps derived from MODIS to distinguish the aggregate phenology of different mixtures of agriculture and indigenous vegetation at scales of 250-1000 meters. By comparing the phenology of a given year with the phenology from previous years, it should be possible to quantify anomalous rates of green up and senescence in different areas and to identify the relative abundance of agriculture and indigenous vegetation contributing to the aggregate phenology observed by MODIS.

We conducted two visits to Tigray in 2009 to better inform satellite vegetation measurements by providing contrast between dry season vegetation (in March), such as bushes, trees and crops, and that just before harvest (in October), in the period of their greatest extent. These visits provide the ability to analyze the appropriateness of utilizing remote sensing data and how accurately remote sensing techniques reflects on-the-ground observations. The May visit was conducted by Dr. Christopher Small and the November visit was by Michael Norton. The field validation in March was conducted using a January 2006 Quickbird image for comparison.

Because the presence of indigenous vegetation in non-agricultural areas does not generally change from year to year the primary differences are expected to be related to interannual variability of precipitation (for drought sensitive herbaceous vegetation) and grazing by goats (for smaller shrubs and herbaceous vegetation). In the course of this investigation, approximately 20 landscape panoramas and 50 GPS waypoints were recorded during each visit. To see the difference between the non-agricultural season and when the crops are being grown, see Figures 12 and 13 which are the same subsection of one of the landscape panoramas (figure 14). Figure 12 was taken in the March groundtruthing trip and clearly illustrates the lack of greenness before the cropping season has begun. Figure 13 illustrates the same view in October, and is much greener. By comparing the two time periods, we can better understand how well remote sensing may be able to proxy crop vigor.

In spite of potential differences, a very strong correspondence was found between the 2006 Quickbird image and the 2009 vegetation cover. Using the GPS geotagged field photos in conjunction with the Quickbird imagery we were able to associate different dry season vegetation fraction abundances with different amounts and types of herbaceous cover identified in the field.

We are currently developing spectral mixture models for both the Landsat and Quickbird

imagery to conduct vicarious validation of the Landsat-derived vegetation fraction estimates at 30 m pixel scale. We are comparing the Jan. 14, 2006 Quickbird image with the two closest Landsat acquisitions from Dec. 2006 (julian day 344) and February 2009 (j.d. 43) to quantify both the temporal change and agreement with the higher resolution validation data. These vegetation maps will then be compared with that derived from the Landsat image acquired Feb. 2009 (j.d. 51) to further quantify the multi-temporal uncertainty in the dry season estimates. The dry and wet season vegetation differences will be mapped using Landsat fractions from Feb. 2009 and Oct. 2009 (j.d. 309). These fraction estimates will be used to calibrate the 2000-2010 MODIS EVI. By using high resolution imagery to understand what off the shelf products are representing, our hope is that the off the shelf products can be corrected in order to perform well in validating the performance of rainfall estimates. If the performance of the off the shelf products can be sufficiently improved, they may also play a role in the future as part of the index itself.



Figure 11: Teff crop in Adi Ha, October 2009.



Figure 12: Adi Ha Landscape photograph from groundtruthing exercise– May.



Figure 13: Adi Ha Landscape photograph from groundtruthing exercise–October



Figure 14: Panorama containing view in previous two figures.



Figure X Wet and dry season comparison of orchard and fields west of Adi Ha village. Some slight differences in vegetation phenology are apparent for shrubs (A), and trees (B, C, F). More pronounced differences are seen in agricultural fields - particularly at boundaries (D, E). The far hill slopes (G) are darker because of both tree foliage and resulting shadow.

Figure 15: Wet and dry season comparison west of Adi Ha



Figure Y Wet and dry season comparison of agriculture and indigenous vegetation north of Adi Ha. As in previous example, the indigenous trees are a bit more leafed out and casting deeper shadow in October. The most pronounced difference is in the increased greenness of the agriculture. Note extensive intermingling of trees and agriculture.

Figure 16: Wet and dry season comparison north of Adi Ha