# Using Demographic Data to Estimate Cropped Area in Highly Undernourished Countries

## Introduction

The populations of many developing countries face chronic undernourishment[1] due to food production and import shortages or food accessibility issues. Approximately one sixteenth, of the world's undernourished live in Central America and the Caribbean (Food and Agricultural Organization (FAO) 2006). While the proportion of the undernourished is much smaller than Africa, which contains one fourth of the population of the undernourished, the Central America and Caribbean region contains two countries with some of the highest concentration of undernourishment on the planet, Haiti and Guatemala. In Haiti, almost 50% of the population is undernourished, while in Guatemala a third of the population is undernourished and this proportion is increasing (FAO 2006). War, economic change and political factors have been correlated to food insecurity and food shortages and are very likely components of the high rates of undernourishment in Guatemala and Haiti – two countries with violent and war-torn pasts (FAO 2006, Husak et al. 2008). These countries also have some of the highest fertility rates and the lowest rates of contraception in the Western Hemisphere (PRB 2008). The combination of limited food and rapidly growing populations portends a dire future for both countries.

An important component of undernourishment is the measure of a country's cropped area. As food production is the product of cropped area and yield, having a better estimate of the cropped area gives a more accurate projection of food production and ultimately food insecurity (Husak et al. 2008). Determining the amount of land in cropped area and understanding the relationship between demographic factors and cropped area provides information important for predicting the amount of food to import and helps to identify areas at risk for famines.

Unfortunately measuring the amount of cropped area is not exact. This is due to many factors, a few of which are limited accessibility, poor communication between farmers and government agencies, and political or economic motivations in crop area reporting. While intensive surveys have been undertaken and remote sensing tools and techniques developed in the 1970s and 1980s have helped to improve cropped area estimates large variability in country-level estimates persist (Hammond 1975, Hixson et al. 1981, Husak 2008). Additionally, current pixel counting and classification techniques rely either are very sensitive to variation in pixels and cannot always distinguish between subsistence farms and other growing areas – leading to misclassification of pixels and a incorrect estimate of cropped area Even today, many state-of-the-practice remote sensing techniques exist to estimate cropped area, and each has its limitations (Tian 2002, Husak 2008). Recently, the area frame sampling approach, a technique relying on high and medium resolution data used by organizations within Europe and the United States, has

[1] According to the FAO, undernourishment "refers to the condition of people whose dietary energy consumption is continuously below a minimum dietary energy requirement for maintaining a healthy life…"

been applied to Ethiopia. In developing a model for Ethiopia, the researchers relied on different types of images, including IKONOS and Landsat, and landscape information to determine the amount of cropped area within Ethiopia (Husak 2008). This approach uses information from several different data sources to model the relationship between the proportion of a landscape that is cropped known attributes, like elevation. Applying these techniques to Ethiopian data researchers have improved the cropped area estimates and international food aid agencies currently have better knowledge of the food situation in Ethiopia (Husak et al. 2008, Husak 2009).

In terms of demographic and sociological research, previous studies have linked, at the household level, land use and household demographic characteristics (Cain 1984, Binswanger and McIntire 1987, Clay and Johnson 1992, Grace & Carr, in submission, etc.). However, demographic characteristics have not been used to predict the amount of cropped land and based on the results of related research, it is not unlikely that the average household's wealth, the average number of people in a house or the proportion women headed households, for example, may be related to the amount of land a household has cropped. The purpose of this study is to build on the area frame sampling models estimating Ethiopia's cropped area through the incorporation of household and community level demographic data. The results of this study will determine if demographic data can improve cropped area estimates and will provide an alternative perspective on the relationship between demographic factors and land use.

## Data

Three primary types of data will be used in this study: high resolution (~1m) remote sensing data, geophysical data gathered from a variety of sources to capture landscape characteristics and demographic data. The high resolution imagery are provided through existing relationships with the Famine Early Warning System Network (FEWSNET). Geophysical data are collected from a variety of sources including the United States Geological Survey (USGS) , the National Oceanic and Atmospheric Association (NOAA) and potentially from in-country partnerships. The high resolution images will be selected for use and processed by the FEWSNET project. FEWSNETresearchers select images with limited cloud cover taken during the growing season when crops are easiest to distinguish from other vegetation. Because of their interest in exploring the effectiveness of incorporating demographic data into the determination of cropped area, the FEWSNET researchers have granted us access to the geophysical data.

In addition to the geophysical data, demographic data for each of the two countries will be used. Demographic and Health Survey (DHS) data is available for Haiti (2005/06) and Reproductive Health Survey (RHS) data is available for Guatemala (2010 (or 2002 if 2010 data is not released)). These data sets are high quality, country representative data sets containing detailed health and education information at the household and individual level.

## Methods

Manual interpretation of the high resolution satellite imagery serves as the "truth" for this project, and has been shown to be a reliable proxy for field samples. Logistic regression analysis[2] where blocks of interpreted points serve as the dependent observation will be used. The proportion of each block that is cropped, as determined by the interpretation of a regular grid of points within the imagery, is then related to the independent measures such as slope and elevation as well as selected demographic variables will be included. This relationship, developed over areas covered by high resolution imagery, is then expanded to the wider region of interest. The use of aerial images and statistical analysis provides an improved estimate of cropped area over traditional agricultural survey based methods which generally involve aggregating cropped area information from only a subset of sampled locations. In the Ethiopia case, using the remotely sensed images of the entire country h provided a more accurate estimate of the amount of cropped area in Ethiopia. The addition of demographic variables is expected to further improve the accuracy of the pixel classification.

Models will be evaluated for fit using standard fit techniques (log likelihood values and Akaike Information Criterion (AIC), (McCullagh and Nelder 1999). The regression results and the resulting cropped area estimates will be compared to the estimates gathered by governmental organizations within Haiti and Guatemala.

**Significance and Contribution**

Currently satellite imagery and remotely sensed data exist for many developing countries. Because of advances in technology this data can relatively quickly, easily and inexpensively be processed to facilitate the estimation of a country's food production. This research proposes to use the available remotely sensed data with the widely available DHS/RHS demographic data to improve estimates of food resources in developing countries. Understanding how much food a country or region produces and will improve understanding of who has access to food and which population groups may be at risk for famine. Ultimately the results of this study can be used to reduce the risk of widespread famine in food-instable countries.

**References**

Binswanger, H. and J. McIntire (1987), Behavioral and material determinants of production relations in land-abundant tropical agriculture, *Economic Development and Cultural Change,* **25**(4).

Cain, M. (1984), *On the relationship between landholding and fertility,* New York: New York Center for Policy Study.

[2] In previous research a smoothing splines were incorporated to improve the fit of the model with only the slope and elevation dependent variables, this strategy will also be explored in the event that the relationship between the independent and dependent variables does not appear to be linear.

Clay, D. and N.E. Johnson (1992), Size of the farm or size of the family: Which comes first?, *Population Studies,* **42**(3), 491-505.

FAO (2006), *The State of Food Insecurity in the World,* 44pp., Food and Agriculture Organization of the United Nations, Rome.

Grace, K. and Carr, D. (In Submission), The impact of the individual, the household and the community on fertility behavior in Peten, Guatemala.

Hammond, A.L. (1975), Crop Forecasting from space: Toward a global food watch, *Science,* **188***,* 433-436.

Hixson M., et al. (1981), Evaluation of a segment-based landsat full-frame approach to crop area estimation, paper presented at *Proceedings of the 7th International Symposium on Machine Processing of Remotely Sensed Data,* NASA, West Lafayette, Ind., 23-26 June.

Husak , G., et al. (2008), Crop area estimation using high and medium resolution satellite imagery in areas with complex topography, *Journal of Geophysical Research,* **113**.

Husak, G. (2009), Information given through personal communication

McCullagh, P., and J. A. Nelder (1989), *Generalized Linear Models,* 2nd ed., 511 pp., CRC Press, Boca Raton, Fla.

PRB (2008), World population data sheet. Technical report, Population Reference Bureau, Washington DC.

Tian, Y. et al. (2002), Multiscale analysis and validation of the MODIS LAI product II. Sampling strategy, *Remote Sensing of Environment,* **83**:431-441