A High Resolution Satellite Interpretation Technique for Crop Area Monitoring in Developing Countries

Marshall, M.[a], G.J. Husak[b], J. Michaelsen[c], C. Funk[d], D. Pedreros[e], and A. Adoum[f]

[a] Department of Geography, University of California at Santa Barbara, Santa Barbara, CA 93106-4060, USA, Tel: 001-8058938322, Fax: 001-8058933146, *marshall@geog.ucsb.edu*.

[b] Department of Geography, University of California at Santa Barbara, Santa Barbara, CA 93106-4060, USA, *husak@geog.ucsb.edu*.

[c] Department of Geography, University of California at Santa Barbara, Santa Barbara, CA 93106-4060, USA, *joel@geog.ucsb.edu*.

[d] Department of Geography, University of California at Santa Barbara, Santa Barbara, CA 93106-4060, USA, *chris@geog.ucsb.edu*.

[e] Department of Geography, University of California at Santa Barbara, Santa Barbara, CA 93106-4060, USA, *pedreros@usgs.gov*.

[f] AGRHYMET Regional Center, No 0425, rue 001 Boulevard de l'Université, Quartier Université, Commune V, NIGER, *a.Alkhalil@agrhymet.ne*.

Abstract
District-level crop area (CA) is a highly uncertain term in food production equations, which are used to allocate food aid and implement appropriate food security initiatives. Remote sensing studies typically overestimate CA and production, as subsistence plots are exaggerated at coarser resolution, which leads to over optimistic food reports. In this study, medium resolution Landsat ETM+ images were manually classified for Niger and corrected using CA estimates derived from high resolution sample image, topographic, and socioeconomic data. A logistic model with smoothing splines was used to compute the block-average (0.1 degree) probability of an area being cropped. Livelihood zones and elevation explained 75% of the deviance in cropped area, while medium resolution did not add explanatory power. The model overestimates crop area when compared to the national inventory, perhaps due to temporal changes in intercropping, and the exclusion of some staple crops in the national inventory.

Keywords: Remote Sensing, GIS, Agriculture, Classification, Decision Support

1. Introduction

The target of halving undernourished people in developing countries made at the 1996 World Food Summit in Rome is far from being achieved, as the number of undernourished people has decreased from 823 million to 820 million people since the beginning of the 1990's [*FAO*, 1999]. In sub-Saharan Africa, the trend in undernourished has continued to rise for the past 30 years. In Niger, where poverty and malnutrition are a staggering 63% and 40% respectively, the 2004 drought led to a 12% drop in crop production from the five year average. With insufficient and sporadic rains at the end of the 2007 growing season, access to food in 2008 is impaired [*USAID* 2007]. In Niger, and many countries in Africa, the limited purchasing power of people, combined with food shortages, requires preventative market based strategies such as import planning to combat food insecurity on the continent.

Population growth in poverty stricken regions, often leads to farming intensification in climate-sensitive areas (e.g. agro-pastoral), as farmers with low purchasing power are forced to increase production to meet growing demand [*Bjerknes, et al.*, 1969]. Crop production estimates at the district and national levels are therefore becoming important in long term planning of food security initiatives, natural resources management, and desertification prevention. For a given crop, production is the product of yield and crop area (CA). The latter is typically derived from ground truth and demographic surveys. The CA term is highly uncertain: field samples are typically too few to extrapolate to the district or national level, and the majority of plots contain two or more crops, both of which are assigned the total CA of the plot, so that when total CA is calculated as the aggregate of all plots on all farms, overestimates occur. In addition, field samples are costly and time consuming, particularly in remote areas of the world. Satellite image interpretation provides an efficient and cost-effective alternative to traditional CA estimation, and as such has been employed with several broad-scale surveys in Canada, Europe, and the United States: Statistics Canada [*Ryerson, et al.*, 1985], Monitoring Agriculture with Remote Sensing [*Gallego, et al.*, 1997] and the Italian AGRIT Project [*Giovacchini and Brunetti*, 1992], and the Large Area Crop Inventory Experiment [*MacDonald, et al.*, 1975] respectively.

Crop area estimation from satellite imagery is typically calculated using the product of the resolution of an image and the area of an agricultural feature delineated with a spectral classifier [*Gallego, et al.*, 1997] or by direct pixel count [*Bauer, et al.*, 1978; *Shao, et al.*, 2001; *Sridhar, et al.*, 1994]. Studies in west Africa primarily used corase resolution AVHRR to establish linear [*Maselli, et al.*, 1992] or polynomial [*Groten*, 1993] relationships between NDVI and crop yield. To avoid problems of extrpolating these functions through time, a time average of NDVI at the end of the growing season when agricultural landcover has much higher chlorophyll content than background vegetation has also been attempted [*Rasmussen*, 1992]. These methods often lead to serious overestimation of crop area in subsistence-based agricultural systems, as the spectral characteristics of uncultivated land is obscured at relatively coarse resolution [*Ozdogan and Woodcock*, 2006]. Co-location inaccuracies and considerable overlap between spectral categories can induce further error [*Gallego, et al.*, 1997]. Regression estimators developed from the proportion of cropped area in ground surveys with pre-determined segments (area frames) in satellite imagery can produce highly accurate CA estimates [*Pradhan*, 2001]. The cost and difficulty of taking field measurements with this method is still a major limitation. The use of high resolution (HR) satellite imagery in the absence of ground data to determine the bias estimator is a relatively unexplored area, however a study involving Landsat and IKONOS imagery for a high production zone in Ethiopia shows promise, as the unbiased estimates explained 77% of the deviance in agricultural surveys [*Husak, et al.*, 2008].

This study produces unbiased estimates of CA for the primary crop producing season in Niger for 2005. The objectives of this study were: 1) to determine crop area from medium resolution (MR) satellite imagery and area frames defined by the shape and extent of HR satellite imagery; 2) to derive a bias estimator from statistical comparison of area frames with MR satellite imagery; and 3) to compare unbiased CA estimates with estimates produced by the national survey. Unlike the Ethiopian study, which focused on a small crop producing region, estimates in this study were made nation-wide. In addition, this study investigates parameters for the statistical model that better represent the spatiotemporal distribution of crop area in Niger, including: maps of major livelihood zones, latitude, precipitation, the Normalized Difference Vegetation Index (NDVI), slope, and elevation were analyzed together with the HR bias estimator.

2. Material and Methods

2.1 Study Area and Data

Niger is one of the Sahel nations (0°09'E – 15°59'E longitude and 11°59'N – 23°31'N latitude) located in West Africa **Figure 1**. The topography of Niger is subtle: the highest peak is atop the Aïr Massif volcano (1944m) in the north-central portion of the country, which quickly descends to lowlands in the extreme north and highlands in the south [*Gu and Adler*, 2004]. The cycle of rainfall in Niger is typically unimodal, with a peak in (August), due to the penetration of moist Atlantic Monsoon air into the dry trade winds. Rainfall totals south of 15°N can be as high as 870mm/yr and food production is characterized by rainfed subsistence and small-scale irrigated agriculture. More than 85% of the country lies north of 15°N latitude with rainfall totals less than 250mm/yr. These areas are semi-arid to arid and consist primarily of pastoral land use and sparse subsistence agriculture. Crops throughout the country are sown at the beginning of the rainy season and harvested at the end of the rainy season. The high frequency of intense rain, combined with poor infiltration and high evaporation puts considerable stress on crops. As such, subsistence farmers typically cultivate staple foods (millet and sorghum), mixed with cowpeas or groundnuts, in order to reduce the risk of crop failure.

The study utilized seven data sources to integrate information on farming practices, plant biomass, rainfall, latitude, relief, HR classification and MR classification of "crop" and "not crop" for the 2005 growing season, including: polygons delineating primary livelihood zones in Niger, 0.05° resolution seasonal (September-November) average and cumulative grids of MODIS NDVI from 2000-2008 and FEWSNET Africa rainfall estimates from 1998-2008 respectively, the SRTM 90m resolution digital elevation model (DEM) for northern Africa, 0.61cm/1m resolution 2005 September-November Quickbird/IKONOS panchromatic images, and 30m resolution 2005 September-October Landsat ETM+ composites. The map delineating livelihood zones was developed from preliminary interviews and workshops with key informants at the national and regional level and later refined through a series of local meeting and visits [*FEWSNET*, 2005]. The livelihood zones include: agro-pastoral, cultivated Aïr Mountains, Bilma Oasis, desert, cash crops of Lake Chad and the Komadougou River, irrigated rice, pastoral, rainfed agriculture, irrigated cash crops, and high work out-migration. The MODIS product is available for download on the EOS data gateway. The DEM was processed by FAO-SDRN (Environmental and Natural Resources Service). Slope and elevation were determined using standard flow routing algorithms available in ArcGIS®.

The satellite imagery and the Rapid Landcover Mapping (RLCM) tool used to perform the classification were freely provided by the United States Geological Survey (USGS) EROS Data Center in Sioux Falls, SD. The RLCM tool was developed in ArcGIS for the manual classification of landcover area using satellite imagery and a stratified sampling regime. The grid is a set of vector points that can be overlaid with satellite imagery. Each point or sample can be selected and classified while visualizing corresponding satellite images. The grid of samples and corresponding attribute information can then be exported into various formats for analysis. Forty-two HR images stratified over 12 MR images were used to develop the bias estimator. An additional 16 MR images were classified in order to determine nation-wide CA. The area covered by a single HR image ranged between 400km^2 in the more productive agricultural areas to 25km^2 in the less productive agricultural areas, while the Landsat images had an area close to 30,000km^2. The HR imagery covered nearly 6000km^2 of the area covered by the MR imagery. Landsat ETM+ and Quickbird/IKONOS images were cloud- free (<1% cloud cover) and registered in a Lambert Equal Area Azimuthal projection (datum WGS-84) for comparative purposes with the aforementioned data products. The MR and HR images were collected towards the end of the primary growing season. It was expected that the greatest contrast in texture and shape of planted/harvested fields and natural photosynthetic vegetation occurs at this time. Sown fields would be effective for classification as well, as emergent crops tend to have relatively lower photosynthetic rates than other vegetation, however cloud-free images were not available at the beginning of the rainy season.

District and sub-district estimates of CA in Niger have been made by the Enquête pour la Production et l'Estimation des Rendements (EPER) since the mid-seventies. The EPER uses a two step process to determine CA: villages are randomly selected in proportion to village size from a sub-district first, and within each selected village, farms are randomly selected with uniform probability. The area of the selected farms provides an average, which is multiplied by the total number of farms in a sub-district from the national census to calculate the district total. The district totals were not used in this study, because of severe problems with double counting. To account for double counting at the national level, EPER uses adjustment factors derived in Adamou (2001) to estimate crop area for each staple crop. These fractions allow EPER to estimate the total planted area, accounting for the fact that the same field shares two crops and is therefore counted twice in food production estimates. The fractions are determined for cowpea and groundnuts, as these crops are most commonly mixed with other staple crops. In Adamou (2001), 89% of the cropped area in cowpea is represented in the total estimated area of millet grown. To determine the total area of millet grown, EPER simply subtracts 89% of the area reported as cow pea from the millet total. Similarly, the total area of sorghum without cowpea can be estimating using the cowpea to sorghum proportion of 41%. These two figures are then added to the remaining cowpea area. Forty percent of groundnuts are counted in other crops, so 40% of the area is removed from the groundnut total and the remainder is added to the group total. Other crops are a relatively lower proportion of cropped area and are therefore ignored. The fractions are derived for national estimates of cropped area, so estimates at the sub-district or district level are not possible at this time.

2.2 Classification of HR and medium resolution imagery

Interpretation of HR and MR imagery consisted of the manual classification of samples with the RLCM tool. A grid of unclassified samples were overlaid with the satellite imagery and classified individually or in clusters as “crop” or “not crop.” The assigned samples were given a unique identifier that was later merged with information on NDVI, precipitation, latitude, slope, elevation, and livelihood zone. More than one sample can be classified at a time using the RLCM tool, so the scale at which samples were visualized depended on the size and shape of features being classified, complexity of the landscape, and the resolution of the imagery being interpreted. The HR imagery was interpreted as a single grayscale band, while Landsat images were interpreted from three band composites that enhanced photosynthetic features (green, red, and near infrared or mid-infrared). The HR image frames and corresponding Landsat subsets were sampled at 500m intervals, while the full Landsat scenes used to determine unbiased CA estimates for the entire country were sampled at 2km intervals. For both high and low frequency samples, remaining indicators of CA were extracted from their respective data sources and assigned to the corresponding interpreted samples. It was assumed that minor co-location errors in the various datasets are accounted for in the block averaging technique described below. After omitting samples that were identified as cloud, cloud shadow, or “no data” due to Landsat SLC-off, more than 18,500 high frequency and 144,800 low frequency samples were used to derive the bias estimator and unbiased estimates respectively.

A visit of Niger was made to determine the general feasibility of using HR imagery in lieu of actual field visits. The scope of the survey was limited by budget and time constraints, so it was not considered in the validation phase. The field survey was performed in early October of 2007 to evaluate six IKONOS images taken in the previous month. The sampling regime consisted of 125 stratified roadside pairs. The latitude and longitude of each sample was determined using a Garmin ETREX 2 (spatial error < 15m) and recorded in ArcGIS, along with a “crop” or “not crop” classifier. The interpretation of the HR imagery was not performed in a Geographic Information Systems environment. Instead, HR interpretation was made independent of field survey results using snapshots of the sample locations overlaid with corresponding HR imagery provided by the field personnel. This prevented registration errors between interpretations made in the field and remote sensing imagery.

2.3. Statistical Analysis

Logistic regression generalized linear modeling [*McCullagh and Nelder*, 1989] was used to represent the binary response (crop/non-crop estimate for HR satellite imagery) to various potential predictors. In a simple multivariate logistic regression model, the probability of an area being crop **p** is defined as the logistic transformation of a linear function. The logistic coefficient is defined as the log of the odds ratio (probability of an event/probability of no event):

$$\log(p_j / (1 - p_j)) = \beta_0 + \Sigma\, \beta_{ij}\, x_{ij}$$

Where β_0 and β_{ij} are that intercept and slope coefficients of the odds ratio respectively, and **x** is the independent variable **i** (livelihood zone, MR crop or not crop, NDVI, elevation, slope, and latitude) for block **j**. The possibility for non-linear relationships was investigated using smoothing spline models in a generalized additive model framework [*Hastie et al.*, 2001]. Spline functions are piecewise polynomial fits. The model becomes:

$$\log(p_j / (1 - p_j)) = \beta_0 + \Sigma f_{ij}(x_{ij})$$

Where $f_{ij}(x_{ij})$ is the spline function for independent variable **i** and block **j**. The basic modeling approach used here was a somewhat modified version of the model presented in Husak et al. (2008). There are two main differences. First, the independent variables with the exception of MR were averaged over 0.1 degree blocks in order to reduce the impact of possible mis-registration errors and to facilitate interpretation of the results. The block size was a tradeoff between overpopulating the sample size using smaller blocks and obscuring details using larger blocks. The number of 500m crop (non-crop) samples in the HR blocks divided by the total number of samples in each block was used to calculate probability in the logistic model. Continuous predictors like slope and elevation were averaged. The aggregation produced 155 blocks with 1 to 484 interpreted samples per block. The median was 89 samples per block, and approximately three quarters of the blocks had at least 40 samples. The number of interpreted samples varied, because the sampling grid was chosen independent of the distribution of HR imagery, and the blocks were adjusted, so that blocks did not overlap two or more livelihood zones.

The second difference, involved a change in the way the livelihood zone (land use indicator) was used in the model. In the Ethiopian analysis, the data were partitioned based on land use and separate models were fit with the remaining predictors for each land cover class. In that case, there were strong relationships between elevation, slope and probability of crop that varied between different classes. In this study the livelihood zone predictor was used first as a mask to remove all desert pixels from the analysis and then as a nominal predictor in a single model fit over all non-desert livelihood zones.

3. Results

The high frequency samples were distributed among the primary crop producing livelihood zones: 30% agro-pastoral, 14% irrigated cash crops, 13% high work out-migration, 5% pastoral, 24% rainfed agriculture, and 14% irrigated rice. The absence of agricultural landcover in the pastoral zone reflects the low representation of sampling there. Desert was not sampled, as agricultural landcover was assumed to be negligible. The remaining livelihood zones were aggregated into the aforementioned categories that best represented their climatic, eco-physiological, textural, and topographic characteristics. Overall, 23% of the HR samples were classified as cropland. The HR evaluation with ground data is not robust enough to consider for validation, but was a convenient and inexpensive collection of samples used for proof of concept. Of the 249 samples classified from the field visit and HR image interpretation, 38.1% were crop and there was agreement for 223 of the samples, giving an overall accuracy of 89.6%. The errors of omission and commission were very similar, resulting in a modeled estimate of crop at the samples that was within 1.5% of the field survey.

A total of 13,786 Landsat samples (crop = 720 and not crop = 13,066) were correctly classified when compared to the HR on a sample by sample basis, while the remaining 4,768 samples were classified incorrectly, yielding an overall accuracy of 74.3%. The errors of commission and omission were substantially higher for crop than non-crop: 82.9% and 63.7% respectively. A summary of the commission errors, omission errors, and accuracies for each livelihood zone are presented in **Table 1**. The overall accuracy was highest for the pastoral livelihood zone, followed by agro-pastoral, high work out-migration, rainfed agriculture, irrigated cash crops, and irrigated rice. The omission errors were highest for irrigated cash crops and rainfed agriculture. The commission and omission error for each livelihood zone, with the exception of pastoral, was 100% and 0% respectively. Agro-pastoral had a high omission error, which was countered by a low commission error for non-crop.

The final logistic model used livelihood zone, MR interpretation, NDVI, and elevation as predictors. The model explained 84% of the binomial distribution deviance. The deviance explained is comparable to the variance explained in a linear model. **Figure 2** shows the actual block-level crop pixel frequencies and predicted probabilities. The R^2, which is used to illustrate the relationship between the two, is 0.73. The residuals are not homogenous, because the distribution is binomial. The blocks at the (0,0) intercept are smaller than the other blocks and contain no cropped area. These blocks were classified from imagery in the northeast of the country, which is dominated by pastoral landuse. Livelihood was the most important predictor by a large margin, explaining 62% of the deviance. This is evident in **Figure 3**, which shows box-plots of the range of crop frequencies for each livelihood zone. The pastoral zone has almost no cropped pixels, and the resulting model produces a very small probability of roughly 0.001. Also unlikely to be cropped are the agro-pastoral (fitted probability, p=0.07) and Out Migration (p=0.12) zones. Most of the cropped pixels are concentrated in the rice (p=0.32), rainfed agriculture (p=0.33), and irrigated cash (p=0.47) zones.

The second most important predictor was elevation, which captures an additional 13% of the deviance. The elevation range is small (~200-500m) and the differences appear to be primarily localized, rather than large scale. The relationship between elevation and predicted crop probability in **Figure 4** is non-linear and step-like with log odds ratio values of around -0.4 (decreasing the odds ratio by a factor of about 0.60, or 3:5 against crop) below 300m, and then a rapid increase to a peak log odds ratio of 0.90 (increasing the odds ratio by a factor of about 2.5, or 5:2 in favor of crop) around 400m. The curve declines at the highest elevations, above 400m, but this is essentially an edge effect, as there are only a few observations above 450m.

The next most important predictor is NDVI, which adds roughly 5% to the predicted deviance. The relationship in **Figure 5** shows a monotonic increase in predicted crop probability, starting from a very low log odds ratio of around -3 (decreasing odds ratio by a factor of about 0.04, or 25:1 against) at NDVI values around 0.10, indicating almost no chance of crop, climbing rapidly up to a peak of about 0.40 (increasing odds ratio by a factor of about 1.5, or 3:2 in favor) at an NDVI value of around 0.25, and then basically becoming flat. It appears that NDVI can indicate areas where there will not be any cropped area, but is less successful at distinguishing between crop and not crop in areas with NDVI values above the threshold of around 0.25. Medium resolution satellite interpretation explained the remaining deviance in the final model. As with NDVI, the

MR relationship had a threshold effect, with a rapid rise in predicted crop probabilities around MR frequencies of about 0.4.

The model was applied to the low frequency nationwide sample set to get estimates of total cropped area. The 2km data were again aggregated into 0.1 degree blocks and then probability estimates were produced based on the slope and intercept of the HR analysis. **Figure 6** shows the estimated frequencies. There is a clear latitudinal stratification in spite of the fact that latitude was not included explicitly. This mainly reflects the fact that latitude is implicit in the livelihood zones, which show strong latitudinal stratification as well. The total predicted cropped area was 9,008,500 ha, or 16.7% of the total non-desert land area, with a standard error of ±171,162 ha. The national estimate provided by EPER for 2005 was 7,454,571 ha, which is outside the limits of model error.

4. Discussion

The primary objectives of this study were to make unbiased estimates of crop area using MR satellite imagery and HR sub-sets and to compare these estimates to national inventories. It is difficult to make direct comparisons between results of this paper and the existing national estimate, as the proportions defined in Adamou (2001) were determined from 2001 scale statistics, which may not account for temporal shifts in double cropping. For example, 2005 was a good harvest year, so optimistic farmers may have shifted their strategy and inter-cropped less, which would lead to lower 2005 proportions and a higher cropped area. This would suggest that the national survey underestimates CA using the 2001 proportions. In addition, other staples with lower cropped area, such as rice and maize were not included in the EPER total. The USGS produced two land use maps for Niger using Landsat imagery from 1975 and 2000. Agricultural land use for Niger in 2000 was estimated to be 9,292,400 ha, which is nearly one and a half times more than the 2001 adjusted national estimate. The USGS estimate is somewhat higher than the estimate made in this study, which is expected, because no bias corrections were made. The relatively high commission and omission errors and lower accuracy for agriculture in MR classification, reveals interpreter bias towards choosing non-agriculture at coarser resolutions. This appears to be in contradiction to the Ethiopia study, which showed that commission and omission errors were typically higher for non-agriculture, leading to an over estimation of cropped area. The confusion matrix should be viewed with caution, due to the disproportionate sample sizes of crop and non-crop. The confusion matrix was standardized (not shown here) to account for disproportionate sample sizes. The results were similar to those shown in **Table 1**. Standardization is difficult, because of the high (100%) commission errors for most livelihood zones: non-agriculture in HR is never classified as agriculture in MR. Co-location errors between HR and MR samples may further inhibit interpretation of the confusion matrix as well. The bias is lower for livelihoods with relatively less vegetation (pastoral and agro-pastoral), as reflected in the individual commission and omission errors. This could reflect the difficulty of interpreting agriculture and natural vegetation at coarser resolution when pixels are more mixed. One advantage of the logistic model is that it assumes a binomial distribution, so the assumptions of normality, as in the confusion matrix are irrelevant. It appears from model development that the high-medium resolution hybrid model produces CA estimates that more closely resemble ground estimates than MR alone. The model also reveals that the inclusion of MR interpretation in the model shows no real cost-benefit over HR interpretation alone. The use of HR imagery for bias estimation and extrapolation of unbiased CA estimates using data on livelihood, elevation, and NDVI appears to be the most cost-effective technique for estimating CA.

The use of this technique is critical to food security monitoring and response to famine. Field surveys in developing countries are often too sparse, because of poor infrastructure, leading to extra expenses, unanticipated delays and a biased sample. Four interpreters took approximately five days to classify the 18,500 high resolution samples. This equates to approximately one day for an individual to interpret an image frame with 1000 samples. If the low frequency classification is included in the time budget, approximately two additional weeks with two image interpreters were need to complete the interpretation. The low frequency interpretation takes less time, because features were less distinct in the coarser Landsat imagery and downloading/uploading is reduced. The rate at which low and high frequency samples were classified is far superior to the number of samples that could be visited on the ground in less than one month. With access to the United States government image library, the opportunity to analyze imagery at no cost makes this method even more viable.

The livelihood zones, which capture physical gradients such as rainfall and temperature and the associated change in life style, are the single most important predictor of CA in Niger. The low probability of CA north of 15°N latitude is a clear demarcation of monsoonal and semi-arid climates, as well as agricultural and pastoral livelihoods in the country. The model shows that farmers generally farm in the southern highlands (>300m). These areas are near Zinder and are the most populated in the country. Given the high correlation between elevation and cropped area and the subtleness of relief in the country, it appears that elevation is an indicator population density. Livelihood zone, which is the most likely predictor variable to explain population density, is not significantly correlated with elevation. Slope shows a relatively weak relationship with CA and this most likely reflects the subtleness of relief in Niger. The inability of NDVI to improve crop area estimates above a value of 0.25 is due to the relatively coarse resolution of the NDVI product and the presence of mixed pixels. Images (including NDVI) were chosen towards the end of the growing season, so that agriculture, which typically has a stronger signal than natural vegetation at this time, would maximize delineation between the two. In moderate to densely vegetated areas (NDVI > 0.25) the ability of NDVI to delineate cropped area is less, because natural vegetation acts to lower the NDVI signal. Niger is a sparsely vegetated country, so NDVI gives some added explanatory power. It is expected that in countries with more dense vegetation, NDVI would add less explanatory power to the model. Precipitation was also considered as a potential predictor, but was excluded, due to its high correlation with NDVI.

Uncertainties the MR interpretation can be attributed due in part to misclassification and co-location errors. Misclassification is a problem that can be improved with better training and tools to aid the classification of MR imagery in countries where MR imagery is a more significant predictor of CA than in Niger. Additionally, attempts to automate the classification of HR and MR imagery could reduce the human time investment involved and yield results consistent with the manual interpretation. Co-location errors are inherent and difficult to accommodate without investing considerable time to co-registering the various datasets that were used to develop the model.

The simple test of the HR interpretation, in lieu of ground truth confirmed a high-level of correspondence between the field assessment of crop and the image interpretation. It should be noted that time and cost constraints limited the significance and robustness of this test, as the sample size was too small and limited to roadside samples. One additional step that proved beneficial to the 2007 HR and field comparison was access to notes and photos from the field trip. This aided the image interpreter, by revealing general patterns in texture, shape, and brightness of vegetated areas in the region, without biasing actual data samples. This sort of training does not require much of an investment, but proved extremely valuable in helping landscape interpretation and ultimately the determination of whether a data sample was crop.

Technicians working at the Regional Center for Mapping of Resources in Nairobi, Kenya are now using the methods outlined in this paper to estimate CA for many food insecure countries in Africa using high-resolution or high-medium resolution satellite imagery provided freely by the USGS. With the technique developed outline in this paper, they have been able to determine bias estimators for given growing seasons with modest accuracy. The initial bias estimator developed for each country is used as an educated guess for rapid response, while the latest HR images are used to reevaluate the bias estimator to account for temporal changes over growing seasons.

5. Conclusions and Implications

The study highlights the benefit of maximizing the use of digital spatial data and products, to provide cost-effective cropped area estimates for countries in Africa. In addition, the study demonstrates the ability of HR imagery to substitute for ground truth in regions where such techniques are currently not feasible, and the added benefit of using a minimal set of training imagery to guide interpretation of HR imagery. With this development, it is possible to design extensive sampling regimes that can be interpreted using HR imagery and provide “ground truth” over large areas. The relationships developed can then be extrapolated over entire countries using data such as livelihood zones, elevation, and NDVI. The added benefit of a HR-MR hybrid model is limited in Niger and should be evaluated on a country by country basis. For early warning networks, the methodology discussed here enhances the use of an area-frame sampling approach where image interpretations provide wide area crop percentages which can be confidently used to build statistical models relating physical and sociological information to cropped area.

The introduction and application of this method in government sponsored offices in Niger will undoubtedly save time and money when considering the limitations of field sampling. The bias estimator can be used to provide initial estimates of CA for a given growing season, but more importantly, the methodology presented here can be used by regional offices provided with HR imagery and existing vector products throughout Africa. Automated classification approaches, including segmentation of moderate resolution imagery for CA estimation, remains unexplored. Its application in countries with mixed agriculture and pastoral landcover like Niger could show a clear advantage over spectral classifiers, because structurally these landcover types are quite dissimilar, while spectrally they are similar. The time saved using segmentation would undoubtedly further reduce resource expenditures.

Acknowledgements
This work was supported by the US Agency for International Development through a US Geological Survey cooperative agreement (04HQAG0001).

List of References

Adamou, M. (2001), Etude de la Diversité Génétique des Niébés (Vigna unguiculata (L.) Walp.) du Niger: Caractères Agromorphologiques, Thesis, 90 pp, Université Abdou Moumouni: Niamey, Niger.

Bauer, M. E., et al. (1978), Area estimation of crops by digital analysis of Landsat data *Photogrammetric Engineering and Remote Sensing*, 44, 1033-1043.

Bjerknes, J., et al. (1969), Satellite mapping of the Pacific tropical cloudiness, *Bulletin of the American Meteorological Society*, 50, 313-322.

FAO (1999), The state of food insecurity in the world - 1999, 32 pp, Food and Agriculture Organization.

FEWSNET (2005), Niger Livelihood Profiles, 18pp, Famine Early Warning System Network, Washington D.C.

Gallego, F.J., et al. (1997), Regional crop inventories in Europe assisted by remote sensing: 1988-1993, 60pp, Joint Research Centre, Ispra, Italy.

Giovacchini, A. and A. Brunetti (1992), Agricultural Statistics by Remote Sensing in Italy: an Ultimate Cost Analysis, paper presented at the Conference on the Application of Remote Sensing to Agricultural Statistics, Office for Publications of the E.C., Luxembourg City, Luxembourg.

Groten, S. M. E. (1993), NDVI—crop monitoring and early yield assessment of Burkina Faso, *International Journal of Remote Sensing*, *14*, 1495 - 1515.

Gu, G. and R.F. Adler (2004), Seasonal Evolution and Variability Associated with the West African Monsoon System, *Journal of Climate*, 17, 3364-3377.

Hastie, T., et al. (2001), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 1st ed., 552 pp, Spring: New York, United States.

Husak, G. J., et al. (2008), Crop area estimation using high and medium resolution satellite imagery in areas with complex topography, (in press).

MacDonald, R. B., et al. (1975), The use of LANDSAT data in a Large Area Crop Inventory Experiment /LACIE, paper presented at the 2nd International Symposium On Machine Processing Of Remotely Sensed Data, National Aeronautics and Space Administration, West Lafayette, IN, June 3-5.

Marsh, S. E., et al. (1992), Comparison of multi-temporal NOAA-AVHRR and SPOT-XS satellite data for mapping land-cover dynamics in the west African Sahel, *International Journal of Remote Sensing*, *13*, 2997 - 3016.

Maselli, F., et al. (1992), Use of NOAA-AVHRR NDVI data for environmental monitoring and crop forecasting in the Sahel. Preliminary results, *International Journal of Remote Sensing*, *13*, 2743 - 2749.

McCullagh, P. and J.A. Nelder (1989) Generalized Linear Models, 2nd ed., 511 pp, Chapman and Hall: New York.

Ozdogan, M., and C. E. Woodcock (2006), Resolution dependent errors in remote sensing of cultivated areas, *Remote Sensing of Environment*, 103, 203-217.

Pradhan, S. (2001), Crop Area Estimation Using GIS, Remote Sensing and Area Frame Sampling, *International Journal of Applied Earth Observation and Geoinformation*, 3, 86-92.

Rasmussen, M. S. (1992), Assessment of millet yields and production in northern Burkina Faso using integrated NDVI from the AVHRR, *International Journal of Remote Sensing*, *13*, 3431 - 3442.

Ryerson, R. A., et al. (1985), Timely crop area estimates from Landsat, *Photogrammetric Engineering and Remote Sensing*, 51, 1735-1743.

Shao, Y., et al. (2001), Rice monitoring and production estimation using multitemporal RADARSAT, *Remote Sensing of Environment*, 76, 310-325.

Sridhar, V. N., et al. (1994), Wheat production forecasting for a predominately unirrigated region in Madhya Pradesh (India), *International Journal of Remote Sensing*, 15, 1307-1316.

USAID (2007), Niger Food Security Update, 6 pp, United States Agency for International Development, Niamey, Niger.

Wentling, M.G. (2008), Niger-Annual Food Security Report: Current Situation and Future Prospects, 7 pp, United States Agency for International Development, Niamey, Niger.

Table 1: Total number of HR and MR points classified, commission error (Co), omission error (Om), and overall accuracy for crop (Ag) and not crop (NonAg) in each livelihood zone.

Livelihood Zone	HR Ag	HR NonAg	MR Ag	MR NonAg	CoAg (%)	CoNonAg (%)	OmAg (%)	OmNonAg (%)	Accuracy (%)
Agro-pastoral	391	5243	170	5464	100.0	4.0	43.5	0.0	96.1
High work out-migration	298	2078	64	2312	100.0	10.1	21.5	0.0	90.2
Irrigated cash crops	1202	1351	665	1888	100.0	28.4	55.3	0.0	79.0
Irrigated Rice	819	1740	261	2298	100.0	24.3	31.9	0.0	78.2
Pastoral	1	958	23	936	4.3	0.0	100.0	2.3	97.7
Rainfed Agriculture	1493	2959	800	3652	100.0	19.0	53.6	0.0	84.4

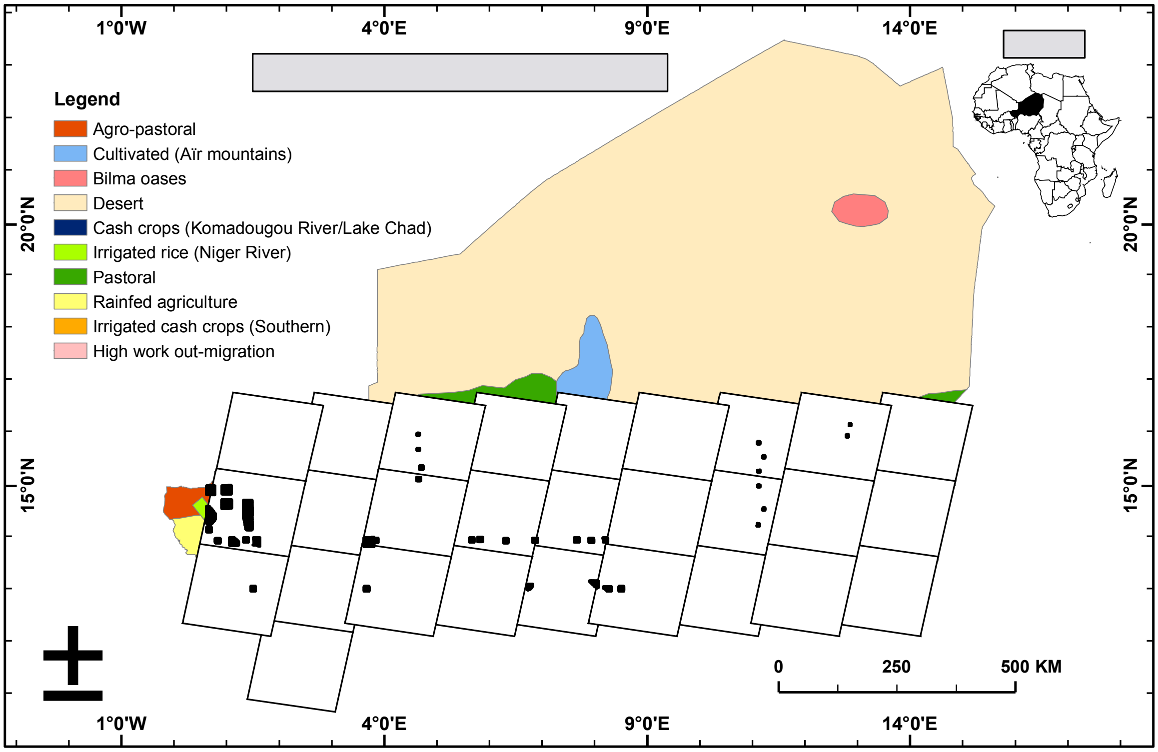

Figure 1: The study area included 28 moderate resolution Landsat scenes (-) and 42 HR IKONOS/Quickbird scenes (■) used in determining the bias estimator. The livelihood zone map was provided by FEWSNET®.

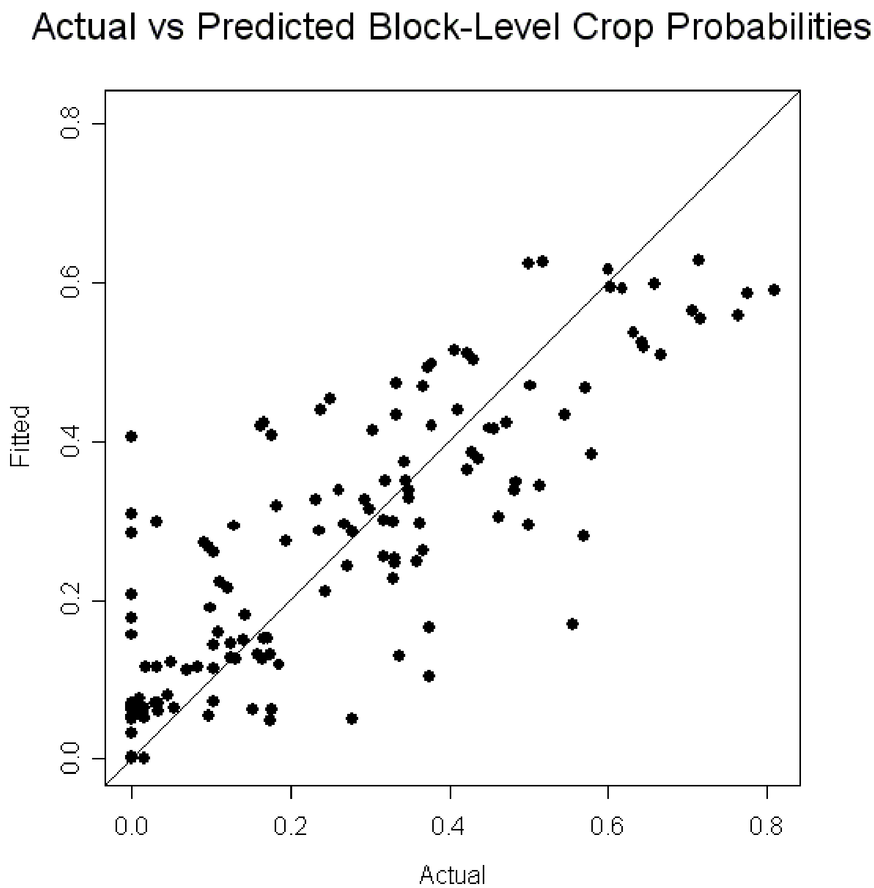

Figure 2: Actual versus predicted block-level crop probabilities and generalized additive model fit.

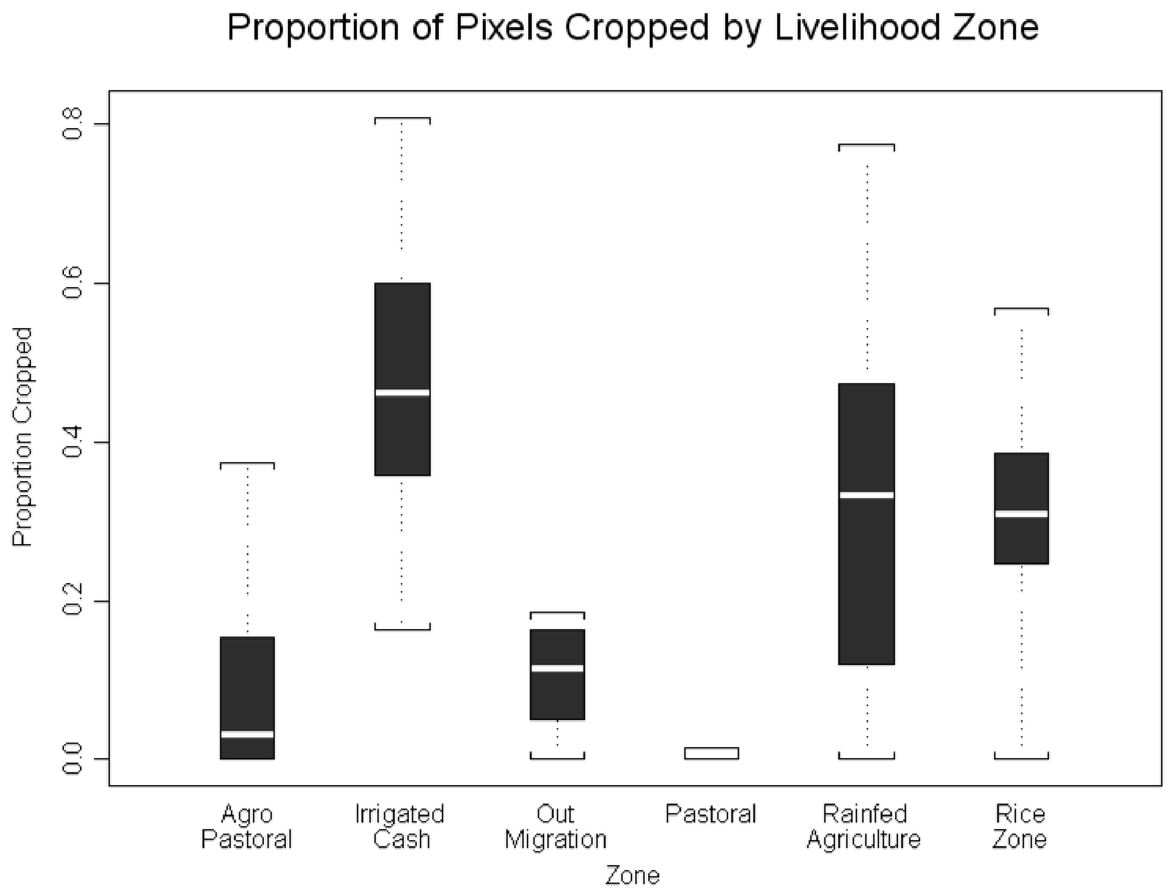

Figure 3: Box-plots of the proportion of cropped pixels based for each livelihood zone.

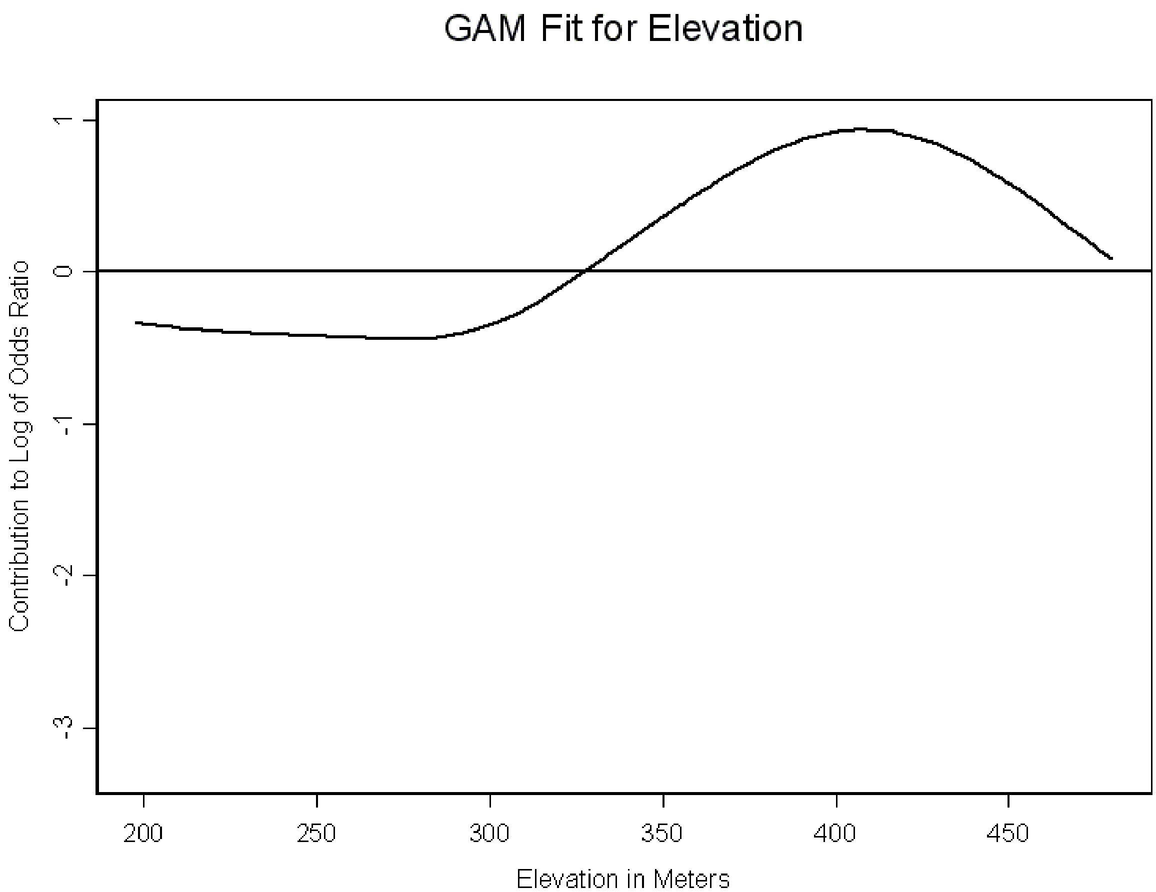

Figure 4: Relationship between elevation and predicted crop area probabilities expressed as log of odds ratio.

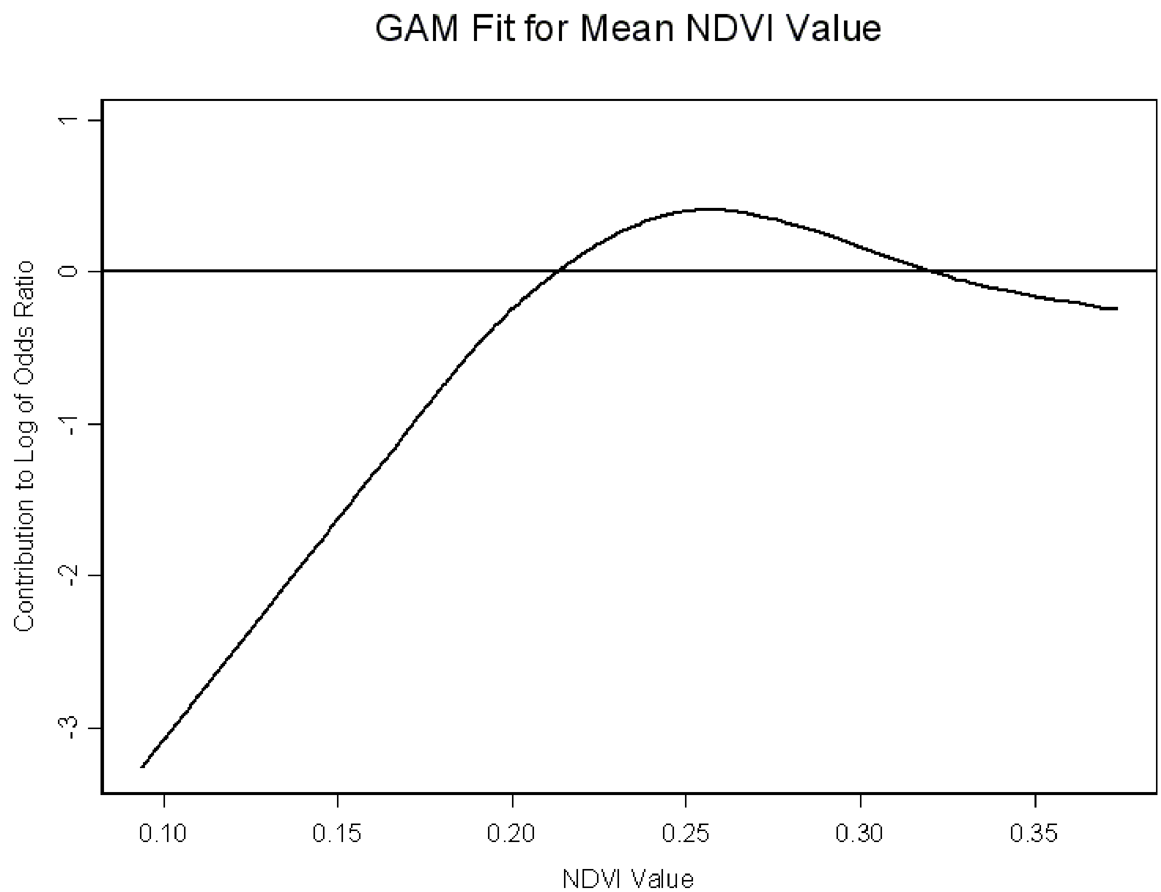

Figure 5: Relationship between NDVI value and predicted crop area probabilities expressed as log of odds ratio.

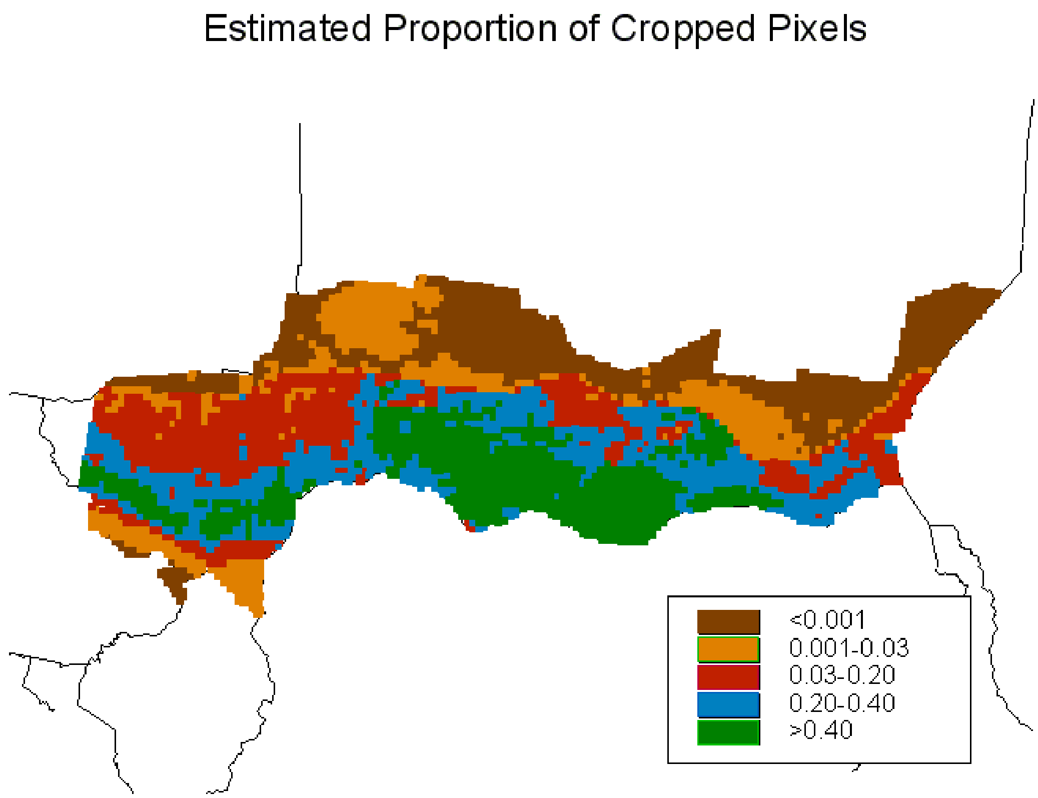

Figure 6: Modeled probability of cropped pixels.