Applied Geography 55 (2014) 257–265

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Estimating agricultural production in marginal and food insecure areas in Kenya using very high resolution remotely sensed imagery

Kathryn Grace [a, *], Greg Husak [b], Seth Bogle [a]

[a] Department of Geography, University of Utah, USA
[b] Climate Hazards Group, University of California, Santa Barbara, USA

ARTICLE INFO

Article history:
Available online 17 October 2014

Keywords:
Kenya
Food insecurity
Land use
High resolution imagery

ABSTRACT

Regularly monitoring the amount of food produced in food insecure, isolated, subsistence farming areas can be used to help identify households or communities who may be in need of additional food resources. Measuring seasonal food production in developing countries, particularly at a sub-national level, is complicated by lack of data. In this study we use high resolution remotely sensed data to calculate cultivated area in two different growing areas, during two different seasons in Kenya. The results of the research support the usefulness of this approach for agricultural monitoring in the developing world and suggest that monitoring cultivated area requires attention to the specific growing characteristics of an area.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

Agricultural production in developing countries, while generally relatively consistent year to year at the country-level, can mask large variation in sub-national production. In fact, it is a common occurrence to observe notable food production failures in one area of a country while another area experiences average or even above-average yields. Because developing countries are characterized by limited food storage and transportation infrastructure, moving food from one high producing part of the country to a less productive part of the country with food needs can be difficult. Regularly monitoring the amount of food produced in food insecure, isolated, subsistence farming areas can be used to help identify households or communities who may be in need of additional food resources.[1] Measuring seasonal food production in developing countries, particularly at a sub-national level, is complicated by lack of data. It is difficult and costly to access all of the farming households engaged in subsistence farming. However, recent research has focused on the use of remotely sensed data to aid in the estimation of area under cultivation. Because food production is the measure of yield (production per hectare) multiplied by area (number of hectares), we can use the area measure to reduce uncertainty in small-scale food production estimates.

One strategy for estimating cultivated area relies on manual interpretation of very high resolution data. With sufficient very high resolution data it is possible to construct estimates of cultivated area, even in communities where fields are small, as is commonly the case in poor, subsistence communities. While this strategy has been used effectively to estimate cultivated area (Grace, Husak, Harrison, Pedreros, & Michaelsen, 2012; Husak et al., 2008; Marshall et al., 2011), questions remain about the spatial and temporal generalizability of this approach. Specifically the ability of this approach to approximate cultivated area in both marginal and highly productive growing areas is unknown.

The purpose of this paper is to examine the ability of the technique to estimate cultivated area in two very different agricultural areas of Kenya, a highly food insecure country in East Africa, during two different agricultural seasons. The results of this research are twofold. First, this research will provide insight into the ability of this estimation strategy to estimate cultivated area for different years, in one high producing and one marginally productive area, based on a suite of readily accessible input variables. Second, the results will provide a remotely-sensed based estimate of cultivated area for two sub-regions of Kenya for two different time periods. As no comparable sub-national data on cultivated area for Kenya exists, this information will be helpful in establishing a baseline of

* Corresponding author.
E-mail addresses: grace@geog.utah.edu, katgrace@gmail.com (K. Grace), husak@geog.ucsb.edu (G. Husak), bogle@geog.utah.edu (S. Bogle).

[1] According to the established definition of food insecurity adopted by the FAO, an individual is food secure when food is available, accessible, able to be utilized and is stable (World Food Summit, 1996). This research, focused on small-scale agricultural yield and production relates to the availability component of food insecurity.

http://dx.doi.org/10.1016/j.apgeog.2014.08.014
0143-6228/© 2014 Elsevier Ltd. All rights reserved.

food production for two different growing regions in a food insecure country.

Using remotely sensed data to estimate agricultural production

Observed rainfall and remotely sensed based estimates of vegetation are often used as measures or serve as proxies for the amount of cultivation or vegetation in a particular area (Cracknell, 2001; Tucker, 1979). The satellite-based estimate of vegetation, Normalized Difference Vegetation Index (NDVI) is frequently used to estimate small-scale food production (De Beurs & Henebry, 2004; Funk & Budde, 2007; Grace, Brown, & McNally, 2014; Townshend & Justice, 1986; Tucker, 1979). Rainfall can be used as a proxy measure of yield and food production on its own (Grace et al., 2012) but can be used in combination with NDVI to help identify growing areas where, for example, some type of irrigation is employed (Omuto, 2011).

Research has found, however, that estimates of the amount of land area under cultivation may be improved through the use of other types of supporting data beyond simply rainfall or vegetation measures. The presence of cultivation is likely dependent on many factors including topographical and landscape features in addition to rainfall and vegetation estimates. In conjunction with NDVI, researchers have used various land cover classifications developed from high resolution imagery and pixel counting methods to estimate the amount of cultivated area (Bauer, Hixson, & Davis, 1978; Fang, 1998; Sridhar et al., 1994). Issues of mixed pixels[2] (Genovese, Vignolles, Nègre, & Passera, 2001; Omuto, 2011; Rojas, 2007), coregistration errors and spatially coarse (rainfall or other remotely sensed) data can make any estimate based on a single data source problematic. These issues may be particularly acute in developing countries where subsistence farms, small in size by their very nature, abut areas of natural vegetation (Ozdogan & Woodcock, 2006). Using multiple types of data available at different scales may improve the understanding of what is actually happening on the ground and mask out areas where agricultural production is not relevant (Husak et al., 2008; Rojas, 2007).

Here we use existing remotely sensed data and build on the area frame sampling approach.[3] This approach involves selecting sampling units at various scales and estimating statistical relationships of the outcome of interest, in this case cultivated area, and the supporting independent data (in this case this is topographic information). We use very high resolution imagery gathered during the local growing season (this data is only available for a subset of the area of interest and serves as the finest sampling unit) and combine the imagery with landscape and geophysical information to construct a model of cropped area at a sub-national to national scale. As Husak et al. (2008), Marshall et al. (2011) and Grace et al. (2012) have shown, this approach is well suited to monitor land cover in areas of complex topography, and where farm sizes are small. No existing applications of this methodology, however, have explored the usefulness of this technique for predicting future or past cultivation. Additionally the usefulness of these models has not been compared in the context of different types of growing areas, specifically areas where the agricultural intensity varies dramatically. In this study, we compare the models used to predict cropped area for two distinct growing areas as well as compare the models' usefulness in prediction. This research will expand scientific understanding of the usefulness of high resolution imagery for agricultural monitoring in a country characterized by diverse topography.

Kenya context

Kenya is heavily dependent on agriculture for food and income. Nearly 75% of the labor force is involved in agriculture (Kenya National Bureau of Statistics, 2006). In most cases irrigation is inefficient (Mati, Mutie, Gadain, Home, & Mtalo, 2008) and fertilizer use is relatively low (Duflo, Kremer, & Robinson, 2008) leaving the majority of farmers heavily dependent on rainfall for crop production. While cropping strategies and growing seasons vary across the country, maize represents one of the largest crops grown and it is grown for both personal consumption and for sale. Periods of inconsistent rainfall, drought and the resultant reductions in agricultural production followed by significant food insecurity are not uncommon in Kenya (Funk et al., 2008). Kenya's agricultural production varies dramatically across the country. This variation is related to differences in social/development factors (Jayne & Muyanga, 2012), rainfall onset, duration and consistency where drought and dry-spells during the rainy season impact the eastern areas of the country more than any other area (Funk et al., 2008; Ngetich et al., 2014). Alternatively, near the middle of the country around Lake Victoria and just west of Nairobi, these highlands represent some of the most productive zones in Kenya where droughts and dry-spells are much less common (Ngetich et al., 2014). The dependence on rainfed agriculture and this area's consistent rainfall patterns result in extensive planting to generate income in some way.

Moving east of the capital towards the south there exists some of the most marginal and inconsistent production areas in the country. More marginal areas have historically reported less area under cultivation because of inconsistent rainfall (Ngetich et al., 2014). What is particularly noteworthy is that even though this area of the country receives more rainfall than the eastern arid area because rainfall is less reliable the people who live here may be at greater risk of food insecurity than the people who live in, and are habituated to, the drier northern regions where drought/dry-spells occur more often and with greater severity. What may be occurring is that farming households depend on the marginal areas for agricultural production and if and when the rains are delayed, sporadic or limited, the resulting reduction in agricultural production leaves these households in a precarious food situation (Osman-Elasha, 2007). Because there are significant spatial variations in agricultural productivity (Funk et al., 2008; Ngetich et al., 2014), regularly examining agricultural production at the sub-national-level is pivotal to identifying communities at risk for food production challenges (Jayne & Muyanga, 2012; Osman-Elasha, 2007).

Because of the dependence on rainfed agriculture and the low level of development, Kenya is one of 18 countries that the Famine Early Warning System Network (FEWS NET - a division of research climatologists and earth scientists supported by US Agency for International Development) regularly monitors for indications of climate/weather issues that might impact food security. A component of the FEWS NET monitoring strategy is the development of livelihood maps. These maps reflect the dominant livelihood (wage and food producing) strategies of an area and are constructed using climate, topographic, economic and local farmer/expert knowledge (http://www.fews.net/east-africa/kenya/livelihood-description/thu-2011-09-29). In this study, we focus on the rich, highly fertile area in the highlands of Kenya near Lake Victoria – this is known as the "Western High Potential Zone" (we will refer to this as HPZ in the remainder of the text) in the FEWS NET livelihood descriptions. We also focus on the marginal area to the south of Nairobi, where

[2] One pixel may represent more than one land cover type.

[3] A report produced by the Food and Agricultural Organization (FAO) in 1998 explores different applications of the area frame sampling approach.

rainfall is inconsistent and where population growth is increasing as a result of poor farmers leaving the crowded fertile growing areas in search of land. This area is referred to as the Southeastern Marginal Mixed Farming Zone (from this point we will refer to this as the MMZ) in the FEWS NET livelihood classification. Fig. 1 highlights the HPZ and MMZ.

The HPZ has a continuous rainy season from March to September with annual rainfall averaging from 1200 to 2200 mm. Temperatures peak around 29 °C during the dry season of December to February. For the remainder of the year, temperatures average from 13 °C to 29 °C. Because of the continuous rain and favorable temperatures this zone is able to maintain agricultural productivity nearly continuously throughout the year. This equates to three vegetable harvests per year. This high productivity allows many farmers to source a third of the food consumed from self-production. Maize is a staple food in the region and constitutes a large portion of the food crop grown along with beans, bananas, and other vegetables. Tea, coffee, and wheat are also produced and sold as a cash crop. In terms of maize yields, the Kenyan Ministry of Agriculture reports that this zone yielded the greatest amount of maize as compared to all other zones in the country.

The MMZ has two rainy seasons each year. The short rainy season extends from October to December. During this period rainfall averages a fairly reliable 500 mm. The second, long rainy season extends from March until May and averages about 800 mm of precipitation. The long rainy season is relied on for most of the food consumed in the area. However, this rainy season is often less reliable than the short rainy season. Temperatures in the zone range from 12 °C to 35 °C with an annual average of 25 °C. Maize is the most relied upon crop in this zone. It is grown for both household consumption and sale. The producing household often consumes over half of the maize they grow. The remainder is sold at market along with other cash and food crops such as cowpeas, beans, sorghum and millet. Here maize yields are about 7 times less than those of the HPZ and are among the lowest yields in the country (excluding the dryer pastoral zones). No data on cultivation for either zone currently exists.

The rainfall difference between these two zones drives much of the difference in cropped area and resulting food production. Because of the dependence on rainfall and the differences between the zones, there is generally more cultivated area in the west and likely with less annual variation. These differences are due to a number of physical factors. Specifically, the HPZ is characterized by a long growing season, abundant and consistent rains, as well as social and agricultural factors (i.e. sending excess food to market, advanced technological investment, non-food crops), which combine to create an environment that supports widespread cultivation year after year. The MMZ, on the other hand, with sporadic and occasionally (and increasingly) failed rains (Funk et al., 2008), experiences reduced and variable cropped areas as a result of limited availability of seeds and agricultural inputs, economic obstacles, and less-developed technology (Jayne & Muyanga, 2012), all of which combine to limit and vary the cropped area in any given year.

Climate models suggest that this marginal area of the country will continue to face irregular rainfall potentially increasing the likelihood of failed crops (Williams & Funk, 2010, 2011). However, little quantitative examination of the link between landscape characteristics, rainfall and crop production exists at the sub-national scale in Kenya leaving researchers and policy makers with no clear models of future food production in this area. This research seeks to employ existing remotely sensed imagery to model cultivated area in these two different growing areas during two different growing seasons.

Fig. 1 shows the locations of the very high resolution imagery within the two livelihood zones studied.

Data and methods

Data

We use very high resolution (~1m) remotely sensed panchromatic data from the WorldView satellites, gathered during the June–September period in order to capture cultivation and

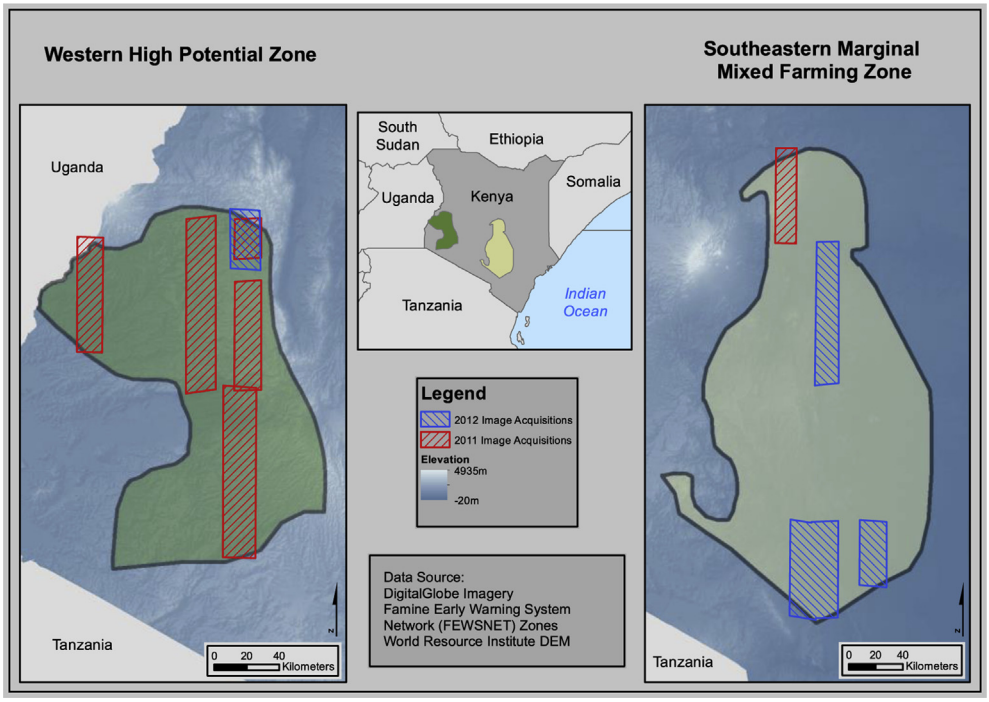

Fig. 1. Two study areas in Kenya.

Table 1
Imagery coverage and geophysical and landscape data according to year and study area.

	High Potential Zone			Sample 2012			Sample 2011		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
CTI	27.5	91.2	56.4	25.0	78.0	51.2	27.3	95.0	55.7
Rainfall (2011)	566	1206	893	590	791	693	585	1100	858
Rainfall (2012)	645	1347	973	694	860	764	641	1169	941
Elevation	1122	2775	1872	1925	2509	2090	1215	2499	1879
Slope	4	71.8	13.6	22	89.2	19.5	6	74.7	15.1
NDVI (2011)	0.67	0.89	0.8	0.75	0.86	0.81	0.68	0.88	0.81
NDVI (2012)	0.66	0.88	0.78	0.75	0.83	0.79	0.63	0.88	0.79
N (% coverage)	844			26 (3%)			320 (38%)		
	Marginal Mixed Zone			Sample 2012			Sample 2011		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
CTI	23.7	113.8	59.0	34.8	100.1	61.9	28.0	79.6	46.3
Rainfall (2011)	18	429	142	71	247	143	54	351	200
Rainfall (2012)	61	566	205	68	335	170	133	349	265
Elevation	294	2112	811	377	970	703	709	2124	1248
Slope	0	84.2	12.0	3	46.7	96	61	66.7	31.5
NDVI (2011)	0.21	0.86	0.47	0.25	0.79	0.48	0.17	0.86	0.59
NDVI (2012)	0.25	0.84	0.56	0.37	0.75	0.52	0.34	0.83	0.63
N (% coverage)	1489			230 (15%)			53 (4%)		

Note: NDVI indicates Normalized Difference Vegetation Index and CTI indicates Compound Topographical Index. Rainfall is measured in mm, elevation is measured in, slope is measured in degrees.

harvesting activity in both the MMZ and the HPZ. Only a small portion of each area of interest has corresponding high resolution imagery. This data is then interpreted using head's-up interpretation (more details on the processing of the data is provided in the upcoming Methods Section). In addition to the high resolution imagery we also rely on geophysical and landscape data from the US Geological Survey's HYDRO1k.[4] Specifically, HYDRO1k provides slope, compound topographical index[5] (CTI) and elevation information at a resolution of 1 km. We also include seasonal (February–July) rainfall totals for both 2011 and 2012 and the maximum seasonal normalized difference vegetation index (NDVI). The rainfall estimates are produced by the National Oceanic and Atmospheric Administration (NOAA) satellite rainfall estimate (RFE2), produced at 0.1-degree resolution (Xie & Arkin, 1996). The MODerate resolution Imaging Spectroradiometer (MODIS) based NDVI has a resolution of 250 m Table 1 provides summary information of the data used in the study. It is important to note that sample sizes were limited – 2011 MMZ high resolution imagery and 2012 HPZ imagery was scant.

While the input data is of varying resolution, these products are used to define landscape characteristics, and not to develop a spatially explicit map of cropped areas. Individual points are attributed with the data from these input datasets and the points are later aggregated to capture broad landscape features, differentiating highlands from lowlands, green from brown, and wet from dry. Although, there is uncertainty in these datasets at the individual pixel level, these data have been chosen because of their ability to capture the characteristics over the landscape.

Methods

We build on methodologies used in related research where high resolution imagery is linked to independent variables through the use of generalized additive models (GAMs) (Grace et al., 2012; Husak et al., 2008; Marshall et al., 2011). The approach begins with a 1 km grid overlayed on the high resolution imagery. Each point in the grid is then classified, by hand, as a location with cultivated area or a location where there is no cultivated area. These points are then attributed with the HYDRO1k, rainfall and NDVI data.

Fig. 2 shows an agricultural area as seen in the very high resolution imagery. Each point overlaid on the imagery was classified as cropped, not cropped, or unclassifiable (due to cloud obstruction).

We then aggregate to blocks of 5×5 km and calculate the percent crop (based on the number of observed crop points in a block) and aggregate the independent variables to the spatial scale (average slope of the block, for example). The percent crop in each block then serves as the dependent variable and the aggregated (typically aggregated to averages) external information serves as the independent or the predictor variables. In other words, we divide each livelihood zone into blocks of 5×5 km and calculate the percent crop in each block (as measured from the very high resolution data). We then explore how the variation in percent crop relates to variation in our independent variables (also at the 5×5 km blocks).

The GAM provides a flexible regression framework for this strategy. It allows relationships between independent and dependent variables to be non-linear (Hastie & Tibshirani, 1990; Wood, 2006), which is often more appropriate for analysis of agriculture. This means that we are not restricted to a linear relationship between rainfall and percent crop, but rather allow for the possibility that high and low rainfall areas, for example, might both be characterized by lower crop, whereas areas that receive moderate rainfall are characterized by high levels of crop. The models take the following form:

$$Y \sim \text{Binomial}(\text{percent crop at block level})$$

$$\text{Logit}(\text{percent crop at block level}) = s(\text{elev}) + s(\text{NDVI}) + s(\text{rainfall}) + s(\text{CTI}) + s(\text{slope}) + s(\text{latitude}),$$

where $s()$ are smooth functions. We also include a spatial smooth of each block's latitude, representing a spatial random effect.

After a suitable model is estimated – all, or a subset of, the independent variables are included in the model based on their significance – the model is then used to estimate cultivated area for

[4] Data available from the U.S. Geological Survey.

[5] CTI is referred to as the "wetness index" and is calculated as a function of the slope and the upstream contributing area (Moore, Grayson, & Ladson, 1991).

Fig. 2. Very high resolution remotely sensed imagery with 1 km grid overlay (Kenya).

those blocks where high resolution imagery was not available using the independent variables observed at the blocks. We then aggregate the block estimates to estimate the percent of cultivated area per year for both the HPZ and the MMZ. Since we have data for both 2011 and 2012 for both zones, we estimate a total of four GAMS – one per year for each zone.

Box 1 is used to summarize the methodological approach – a multi-step process involving a novel use of existing high resolution imagery.

Results

The results are presented in Figs. 2–5. Because GAMs allow more flexibility, and capture the dynamic relationship between the predictors and crop activities, the results are curves presented in graphical form. In all the plots, the *y*-axis captures the probability, or proportion, of a given area being cropped, as a logit value. The line on the plot, then captures the relationship between the predictor variable (% crop), and the proportion of an area with that predictor value being cropped. Fig. 3, for example, demonstrates how NDVI is significantly related to the percent of area cultivated. For the lowest observed values of NDVI the percent crop is expected to be very low, keep in mind the *y*-axis is on the logit scale. As NDVI increases to around 0.5, the likelihood of crop dramatically increases as well and then plateaus for the highest observed NDVI values. The 2012 results (Fig. 4), show a somewhat similar relationship between NDVI and crop for the observed values (the 2012 NDVI range of values is significant smaller than the 2011 range). In terms of elevation, we observe the highest values of percent crop, associated with an elevation of around 800 m for both 2011 and 2012. In areas of higher or lower elevation, we expect a lower amount of cultivation. For 2012, we observe the significance of the relationship between rainfall and cultivated area – in places where rainfall is between 100 and 150 mm we observe the highest percentage of cultivation with a reduction in the likelihood of cultivation as rainfall increases. Places with large slopes are also expected to be less cultivated.

Using the same interpretation approach for Figs. 5 and 6 – as graphical representations of the relationship between percent cropped area and the independent variables we now turn to the HPZ results. Fig. 6 presents the 2012 data and highlights the challenge of fitting these models to small data sets. With only 26 blocks, our analysis is limited. Interpreting the results that we do have, mid-high CTI and high slope values are associated with the highest percent of cultivated area. Because of a much larger and more representative sample, the 2011 results (Fig. 5) provide greater insight into cultivation in the HPZ. Places of high elevation or with high rainfall or high NDVI values are the least likely to be cultivated.

The resulting GAMs and the relationship between cultivation and the independent variables highlight the context specific nature of agricultural production. In other words, significant determinants of crop production are not the same in different areas of the country. Therefore, we can apply this methodology in other contexts as long as there is adequate high resolution imagery to inform

Box 1
Methodological Approach

1. Obtain high resolution imagery over area of interest during time period of interest (ideal coverage is at least 10% of the area of study during harvest period)
2. Overlay a regular grid (1 km) of points over the imagery and interpret points – crop or not-crop
3. Attribute points with external independent variables that are available for the entire region of interest - geophysical, socioeconomic, or landscape characteristics
4. Aggregate independent and dependent variables to blocks of 5×5 km
5. Construct regression models (in this case we use GAMS) predicting percept
6. Input external independent variables into the model to estimate percent crop for blocks **without** high resolution imagery

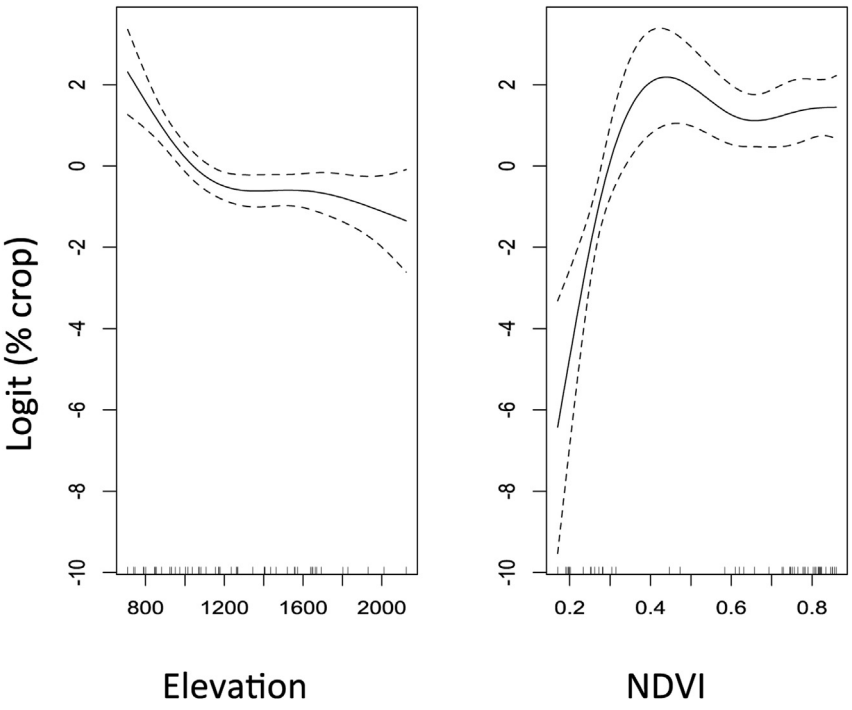

Fig. 3. Marginal Mixed Zone, 2011.

the context-specific model of cultivated area. We summarize the models in Table 2 by indicating the impact of the variable on the explained deviance (the ability of the variable to explain variation in the dependent variable).

NDVI and elevation are significant for all but one case. The explanation for the lack of significance of NDVI and elevation in the 2012 High Productive zone model is due to the small sample size and lack of observed variation in those variables in our sample. For

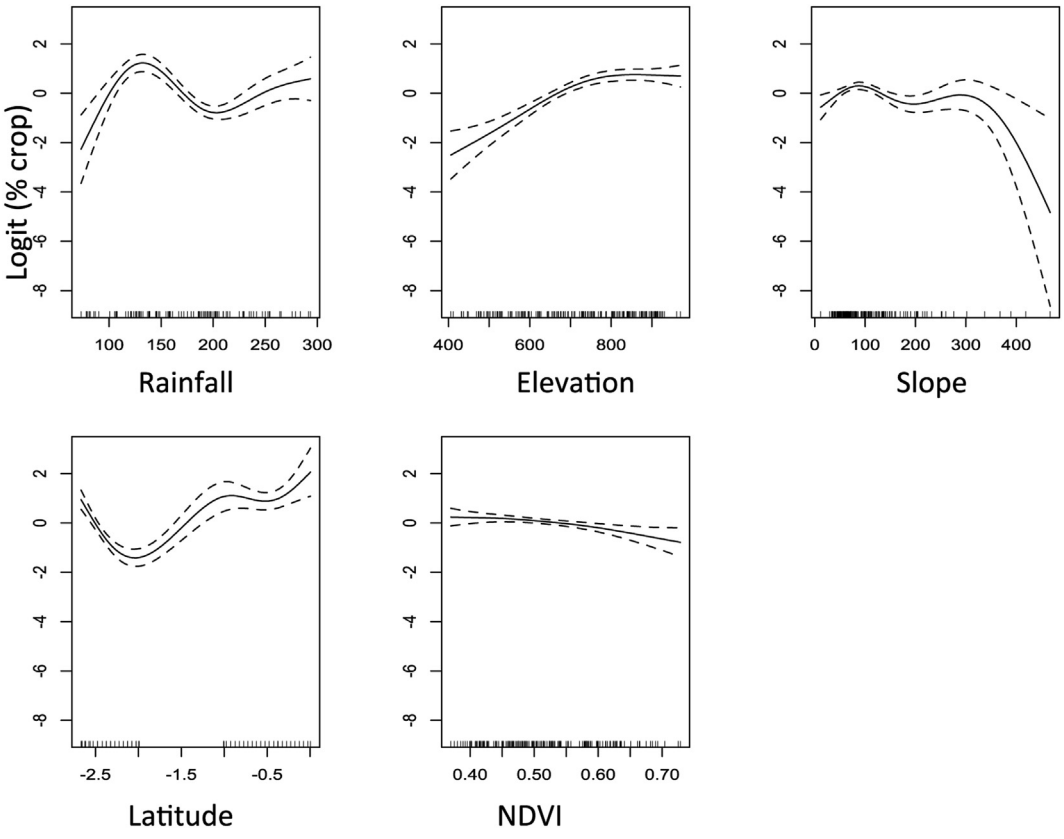

Fig. 4. Marginal Mixed Zone, 2012.

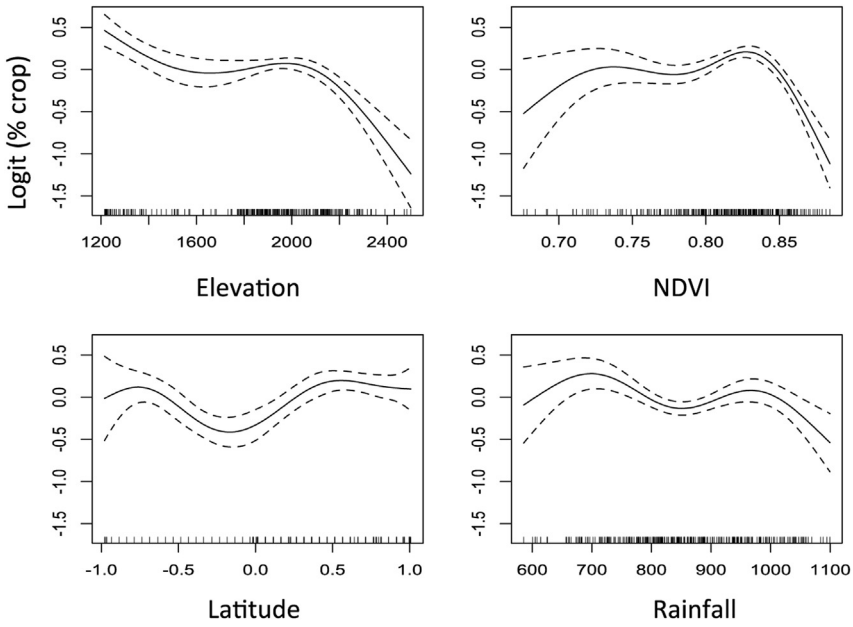

Fig. 5. High Potential Zone, 2011.

the two zones/time periods with the largest sample sizes, rainfall, elevation and NDVI are significant.

Estimates of cultivated area

One of the goals of the analysis was to determine annual cultivation based on the statistical analysis and to explore the use of the models to hind/forecast.

Because of the limited imagery for these zones during these periods, the estimates resulting from these models should be interpreted with caution because the data does not capture the variation in the landscape. In fact, we include the estimates based on these sample sizes to highlight the importance of adequate coverage when using this approach. In applying the approach to other contexts, a user must be careful to verify that the diversity of the landscape under study is captured adequately by the high resolution data. In application, the estimates from the years with scant data should not be used to estimate the cropped area of the entire zone but only of areas within the zone with the same characteristics. Additionally, models based on years with scant data should not be used to estimate past or future cultivation. We present all estimates here for comparison purposes. These estimates are presented in Table 3.

In 2011 in the MMZ, we estimate that approximately 55% of the area was cultivated. In 2012, the period where there was significantly more data used for model construction, we estimated the

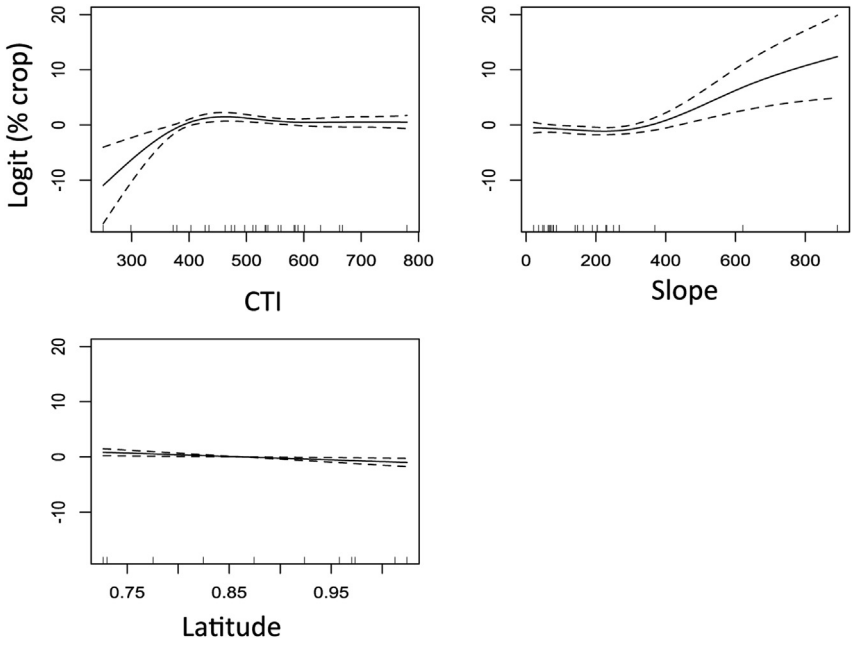

Fig. 6. High Potential Zone, 2012.

Table 2
Summary of GAM results according to year and study area.

Independent variables	Marginal Mixed Zone		High Productive Zone	
	2011	2012	2011	2012
CTI				8.9
Rainfall		26.9	12.5	
Elevation	6.8	12.7	8	
NDVI	41.5	1.3	14.5	
Slope		9.6		6.6
Total Deviance	66.3	56	41.9	81.4
Simple Spatial Model		12.7	18.8	48

area to be 16%. Using the 2012 model to estimate the 2011 cultivation, we estimate 18%. The limited sampling of 2011 overestimated the role of NDVI and built a relationship that exaggerated cropped area because of the influence created by limited, especially green areas included in the sample. In a cross-validation exercise, where a sample of the interpreted data (we assume the high resolution interpretations are "known" values) was withheld and then compared to the results produced from the model, the 2012 model performed much better than the 2011 model, providing further support to the quality of the 2012 model. In the HPZ, the 2011 model, outperforms the 2012 model, again because of more data during the 2011 period. Our estimate for the area under cultivation in 2011 is 42% which is almost 50% larger than the 2012 estimate. The 2012 estimate using the more robust 2011 model produces an estimate of 38%. Given the similarity of results and the performance of the 2012 model (for the MMZ) and the 2011 model (for the HPZ) under cross validation, the ability of these models to forecast/hindecast data is promising. The models did a poor job of estimating cultivated area in the other zone, however, suggesting that models of cultivation must be constructed with attention to small scale contextual information, in this case we rely on livelihood zone as the spatial context.

Discussion and conclusion

The purpose of this paper is to examine the ability of the technique to estimate cultivated area in two very different agricultural areas of Kenya, a highly food insecure country in East Africa, during two different agricultural seasons. The results suggest that this estimation strategy is adaptable enough to estimate cultivation in two different areas of Kenya but the estimation is limited by the amount of high resolution data. The results also provided insight into the determinants of cultivation in both areas which may be useful in forecasting food production given future rainfall and growing predictions.

This study focused on the estimation of cultivated area in two relatively small zones in Kenya. Cultivated area is a key component of land use for understanding food insecurity in developing countries but annually varying cultivation data rarely exists. Using remotely sensed data to estimate and model agricultural land can provide information about potential food needs for poor communities. While the scope of the study is small, the characteristics of these areas are common in many developing countries and provide insight into the usability of this cultivated area process in different types of agricultural settings generally reflective of the developing world. Agriculturalists produce food on fairly small farms in areas of complex topography and rely heavily on rainfall for their crop production. The results of this research indicate that even though these regions share some traits – small producers dependent on rainfall – the relationship between the landscape/geophysical characteristics and cultivated area is dependent on livelihood zone.

Table 3
Estimates of cultivated area by year and study area.

Marginal Mixed Zone	
2011 Estimate	55%
2012 Estimate	16%
2011 Estimate from 2012 Model	18%
High Potential Zone	
2011 Estimate	42%
2012 Estimate	29%
2012 Estimate from 2011 Model	38%

The reason for these differences is a combination of different factors. In different livelihood zones, producers may grow different products – maize is common to both zones but in the HPZ zone, cash crops and vegetables are also grown. The diversity of crops may facilitate cultivating more areas as rainfall, elevation, soil, etc. needs vary. In the MMZ maize is dominant while there is some production of other cereals and some small cash crops. The dominance of maize in this area, may reflect the difficulty in producing any other crop or it may reflect local culture and knowledge which favor maize production over other cultivars. Livestock is also dominant in the MMZ suggesting that some land will be reserved of animal feed/grazing.

Using livelihood designations rather than political or geophysical boundaries can help create a micro-scale context within which to monitor food production. Our results build on the relevance of the livelihood zone as a spatial unit for the analysis of food production. The use of livelihood zones to aid in understanding the variation in rainfall (or related factors) in other countries should be explored as well.

Currently, high resolution imagery is regularly collected and, as this research shows, can be used to estimate land use related to food insecurity and development in poor countries. The implementation of this approach relies on pre-existing data and after the initial model is constructed, can be used with easily accessible updated rainfall and NDVI variables to update estimates of on-the-ground cultivation. Understanding small-scale (sub-national- and community-level) issues of agricultural production is necessary to developing effective strategies to address food insecurity in the developing world. Monitoring production is complex and survey-based methods can be out of reach to many poor countries, this approach provides an alternative strategy that may facilitate improved agricultural monitoring.

In practice, a researcher of community- and household-level food insecurity could use this approach, given adequate high resolution data, to estimate food production at a micro-level for a given community (Grace & Nagle, 2013). The estimate of food production could then be used to explore health and socio-political outcomes related to food availability. Any food production model could also be used to create a time series of food production and evaluate various population outcomes over time as they relate to the greater context of cultivated area.

Because of various climate change scenarios, food production and cultivation are of prime concern for countries that have faced chronic food insecurity (Brown & Funk, 2008; Schmidhuber & Tubiello, 2007). Understanding the human impacts of climate change through thorough studies of food insecurity requires empirical research that evaluates a variety of spatial scales of food production and vulnerability. One of the primary constraints to studies of micro-level food insecurity is the lack of relevant data on food availability. The technique proposed here offers one way of using available data that could be useful for improving scientific understanding of climate, food insecurity and population outcomes.

Acknowledgments

We are very thankful to the support of Gary Eilerts of USAID and FEWS NET, Chris Funk of USGS and FEWS NET, James Verdin and Frank Davenport of FEWS NET and James Rowland of USGS and FEWS NET. This work was supported by USAID/USGS grants, USGS award #G09AC00001 and #G14AC00042.

References

Bauer, M. E., Hixson, M. M., & Davis, B. J. (1978). Area estimation of crops by digital analysis of landsat data. *Photogrammetric Engineering and Remote Sensing, 44*, 1033–1043.

Brown, M. E., & Funk, C. C. (2008). *Food security under climate change*.

Cracknell, A. P. (2001). The exciting and totally unanticipated success of the AVHRR in applications for which it was never intended. *Advances in Space Research, 28*(1), 233–240.

De Beurs, K. M., & Henebry, G. M. (2004). Land surface phenology, climatic variation, and institutional change: analyzing agricultural land cover change in Kazakhstan. *Remote Sensing of Environment, 89*(4), 497–509.

Duflo, E., Kremer, M., & Robinson, J. (2008). How high are rates of return to fertilizer? Evidence from field experiments in Kenya. *The American economic review*, 482–488.

Fang, H. (1998). Rice crop area estimation of an administrative division in China using remote sensing data. *International Journal of Remote Sensing, 19*(17), 3411–3419.

Funk, C., & Budde, M. (2007). *National MODIS NDVI-based production anomaly estimates for Zimbabwe*. Santa Barbara: University of California, 12.

Funk, C., Dettinger, M. D., Michaelsen, J. C., Verdin, J. P., Brown, M. E., Barlow, M., et al. (2008). Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development. *Proceedings of the National Academy of Sciences, 105*(32), 11081–11086.

Genovese, G., Vignolles, C., Nègre, T., & Passera, G. (2001). A methodology for a combined use of normalised difference vegetation index and CORINE land cover data for crop yield monitoring and forecasting. A case study on Spain. *Agronomie, 21*(1), 91–111.

Grace, K., Brown, M., & McNally, A. (2014). Examining the link between food prices and food insecurity: a multi-level analysis of maize price and birthweight in Kenya. *Food Policy, 46*, 56–65.

Grace, K., Husak, G. J., Harrison, L., Pedreros, D., & Michaelsen, J. (2012). Using high resolution satellite imagery to estimate cropped area in Guatemala and Haiti. *Applied Geography, 32*(2), 433–440.

Grace, K., & Nagle, N. (2013, 11–13 April). *Using high resolution remotely sensed data to re-examine the relationship between agriculture and fertility in a pre-transitional setting*. New Orleans, LA, USA: Presented at the Population Association of America (PAA) annual meeting.

Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models* (Vol. 43). CRC Press.

Husak, G. J., Marshall, M. T., Michaelsen, J., Pedreros, D., Funk, C., & Galu, G. (2008). Crop area estimation using high and medium resolution satellite imagery in areas with complex topography. *Journal of Geophysical Research: Atmospheres (1984–2012), 113*(D14).

Jayne, T. S., & Muyanga, M. (2012). Land constraints in Kenya's densely populated rural areas: implications for food policy and institutional reform. *Food Security, 4*(3), 399–421.

Kenya National Bureau of Statistics. (2006). *Economic survey 2006 [Data file]*. Available from http://www.knbs.or.ke.

Marshall, M. T., Husak, G. J., Michaelsen, J., Funk, C., Pedreros, D., & Adoum, A. (2011). Testing a high-resolution satellite interpretation technique for crop area monitoring in developing countries. *International Journal of Remote Sensing, 32*(23), 7997–8012.

Mati, B. M., Mutie, S., Gadain, H., Home, P., & Mtalo, F. (2008). Impacts of land-use/cover changes on the hydrology of the transboundary Mara River, Kenya/Tanzania. *Lakes & Reservoirs: Research & Management, 13*(2), 169–177.

Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological processes, 5*(1), 3–30.

Ngetich, K. F., Mucheru-Muna, M., Mugwe, J. N., Shisanya, C. A., Diels, J., & Mugendi, D. N. (2014). Length of growing season, rainfall temporal distribution, onset and cessation dates in the Kenyan highlands. *Agricultural and Forest Meteorology, 188*, 24–32.

Omuto, C. T. (2011). A new approach for using time-series remote-sensing images to detect changes in vegetation cover and composition in drylands: a case study of eastern Kenya. *International Journal of Remote Sensing, 32*(21), 6025–6045.

Osman-Elasha, B. (2007). Africa's vulnerability to environmental stress and climate change. *Tiempo, 63*, 3–9. Retrieved from http://www.tiempocyberclimate.org/portal/archive/pdf/tiempo63high.pdf.

Ozdogan, M., & Woodcock, C. E. (2006). Resolution dependent errors in remote sensing of cultivated areas. *Remote Sensing of Environment, 103*(2), 203–217.

Rojas, O. (2007). Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya. *International Journal of Remote Sensing, 28*(17), 3775–3793.

Schmidhuber, J., & Tubiello, F. N. (2007). Global food security under climate change. *Proceedings of the National Academy of Sciences, 104*(50), 19703–19708.

Sridhar, V. N., Dadhwal, V. K., Chaudhari, K. N., Sharma, R., Bairagi, G. D., & Sharma, A. K. (1994). Wheat production forecasting for a predominantly unirrigated region in Madhya Pradesh India. *Remote Sensing, 15*(6), 1307–1316.

Townshend, J. R., & Justice, C. O. (1986). Analysis of the dynamics of African vegetation using the normalized difference vegetation index. *International Journal of Remote Sensing, 7*(11), 1435–1445.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment, 8*(2), 127–150.

Williams, A. P., & Funk, C. (2010). *A westward extension of the warm pool intensifies the walker circulation, drying Eastern Africa*. US Geological Survey Open-File Report, 1199(11).

Williams, A. P., & Funk, C. (2011). A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Climate Dynamics, 37*(11–12), 2417–2435.

Wood, S. (2006). *Generalized additive models: an introduction with R*. CRC press.

World Food Summit. (1996, Nov 13). *Rome declaration on world food security*. Available from http://www.fao.org/wfs/index_en.htm.

Xie, P., & Arkin, P. A. (1996). Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *Journal of Climate, 9*(4), 840–858.