An approach to improving early warning systems: Using spatially and temporally rich data to predict food insecurity crises in Malawi

Submission to the Agricultural and Applied Economics Association Annual Meeting
May 2017 draft – please do not cite without permission

Erin Lentz, Hope Michelson, and Kathy Baylis

Abstract

While the causes of famine are complex and include institutional failure and conflict, policymakers and practitioners often lack the necessary information to identify food insecure populations and, therefore, to effectively allocate scarce resources to mitigate hunger. As a result, humanitarian responses tend to trail the onset of food security crises. Our paper aims to enhances the early warning for food insecurity crises. The last decade has seen a dramatic increase in the available quantity and quality of data related to food security, rainfall, and prices. The full potential of these data has not yet been exploited, and they are often evaluated in isolation. We connect disparate datasets and use variation over space and time to build an evidence-based approach to estimating food security outcomes. We utilize price and precipitation data collected in near real-time to predict sub-national food insecurity, as assessed by the Integrated Food Security Phase Classification (IPC). By overlaying mapped measures by season and analysing changes over time, our model has the potential to predict future sub-national food insecurity at relevant spatial and temporal scales. We find that contemporaneous price and measures of the prior year's precipitation consistently predict the IPC in Malawi. We then compare the IPC food security evaluations to household measures of food insecurity in Malawi to ask whether our simple price and weather model improves the prediction of food security outcomes. We find that the IPC is strongly associated with the reduced Coping Strategies Index, a household measure of food insecurity, indicating that the IPC method does successfully capture food security. This model is a crucial step toward improving governments' and NGOs' abilities to target potential food crises at an early stage.

Introduction

The world faced one of its worst humanitarian crisis since World War II in the Spring of 2017, with famine or near-famine conditions in South Sudan, Somalia, northeast Nigeria and Yemen (Gettleman, 2017). These events followed on the heels of Somalia's 2011 famine, in which an estimated 258,000 people died (Maxwell et al. 2016). While the causes of famine are complex and include institutional failure and conflict (Devereux 2009), policymakers and practitioners often lack the necessary information to identify food insecure populations and, therefore, to effectively allocate scarce resources to mitigate hunger (Barrett and Headey 2014). As a result, humanitarian responses tend to trail the onset of food security crises, hampering the efficacy of these initiatives. Our paper aims to enhances the early warning for food insecurity crises. By leveraging real-time, readily available spatially and temporally granular data, we develop a first-cut at a parsimonious, replicable early warning model for food security crises. Once validated in other countries, such a model could assist policymakers in identifying how to best target resources in response to a food security crisis. Further, a model that can predict where problems are likely to occur could save resources if early intervention lessens the severity of the crisis. Finally, our model could also help prioritize where and when to send resources in the case of multiple simultaneous crises.

The last decade has seen a dramatic increase in the available quantity and quality of data related to food security, rainfall, and prices. The full potential of these data has not yet been exploited, and they are often evaluated in isolation, with few analyses linking disaggregated weather and price data. Thus, while these data are available, we are unaware of any food security early warning and monitoring systems that incorporate them into a single predictive model.

Further, most food security measures are at either the national or the household level – leaving a critical gap for programming and research at the meso sub-national level: regions, districts, or sub-districts. The sub-national level is particularly relevant to identifying and understanding food and nutrition insecurity. Food insecurity can be a highly localized phenomenon, with some zones within a country experiencing acute – and regular – insecurity while others do not (Sen 1980; Lentz and Barrett 2013; see also Brown et al. 2014 for a review). Further, a region's food security status can change multiple times over the course of a year and it is not clear that current methods adequately capture these changes. Seasonal hunger appears to be more common than annual surveys of food security would suggest (Devereux et al. 2008). In Malawi, for example, Gelli et al. (2017) find that while food expenditures increase during the lean season relative to the post-harvest season, reflecting the higher cost of buying maize, food insecurity and nutritional status worsen, including a 26 percent decrease in dietary diversity scores for children and a ten-point increase in the proportion of the population consuming less than 1800 kilocalories per day. Given that the range of available food products increases during the lean, rainy season, a decrease in dietary diversity may well underestimate the true fall in food insecurity.

Our work takes advantage of two underexploited resources: a wealth of data available for integration, analysis, and prediction and a relatively new set of sub-national food security indicators. We connect disparate datasets and use variation over space and time to build an evidence-based approach to estimating food security outcomes. We first utilize existing data collected in near-real time (e.g., price and precipitation data) to predict sub-national food insecurity, as assessed by the Integrated Food Security Phase Classification (IPC). The IPC is led by 12 agencies, including the United Nations Food and Agriculture Organization, the United Nations World Food Programme, and the US Agency for International Development's (USAID's) Famine Early Warning System Network (FEWS NET). The IPC is a widely adopted set of tools and protocols to rank food insecurity within and across countries in a consistent manner (IPC 2012). By overlaying mapped measures by season and analysing changes over time, our model has the potential to predict future sub-national food insecurity at relevant spatial and temporal scales. We then compare the IPC food security evaluations to household measures of food insecurity to ask whether our simple price and weather model improves the prediction of food security outcomes. This model is a crucial step toward improving governments' and NGOs' abilities to target potential food crises at an early stage.

While the IPC has become the standard metric for early warning for food insecurity, its production and use are characterized by six important challenges. First, the IPC assessment is generated relatively infrequently for each country, only two to four times each year. Alerts can come too late to mobilize resources to avert human suffering. Second, because IPC assessments are made at the district and/or livelihood zone level, their scope is limited to predicting variation of food security for an entire IPC zone, but not within zones or communities. Third, the IPC assessment is expensive and time intensive because it requires a group of experts to convene to make assessments. Fourth, the assessments are also prone to accusations of political bias by the media and outside watchers (The Economist, 2017). Fifth, the assessments are data intensive in places where data can be dangerous or difficult to collect, with famine classifications requiring specific death rate thresholds that can be extremely hard to measure. Last, since experts incorporate a series of different data sources into their analysis, the analytical process cannot predict food insecurity "out of sample" (i.e., in areas without food security outcome data). Yet, the IPC remains the best available data on early warning that is reliably collected at high frequency and more spatially disaggregation than national measures. Inasmuch as early warning contributes to funding decisions for aid, there is value to understanding whether readily available data can approximate the IPC. Identifying the extent to which real-time price and precipitation data approximate, and can potentially enhance, the IPC can help ameliorate these challenges.

This paper makes three primary contributions to the analysis and prediction of food insecurity. First, ours is among the first analyses to integrate weather and price data together at a sub-national level. We describe protocols for doing so in a way that reflects variable population densities; these protocols can be adopted by researchers interested in complementary questions. Second, this is the first attempt to analyse subnational food insecurity using high-frequency price and weather data. Since we often do not observe yield or local production, particularly in time for early warning measures, we use maize prices as a measure of local food access and precipitation as a measure of food availability. We find that these weather and price data are predictive of sub-national food security and that our model contributes to the explanation of household level food insecurity beyond the IPC. Finally, a primary contribution of this work is validation of the IPC. Food security early warning systems require high frequency and spatially disaggregated data on food security; the IPC is the best high-frequency sub-national indicator of food security currently available. We establish that the IPC is well correlated the reduced Coping Strategies Index, a household-level measure of food insecurity, for the four quarterly assessments that overlap with Malawian household survey data. Thus, we find that the IPC assessments reflect and anticipate household food insecurity. As far as we know, this is the first paper to validate the IPC against household level food security data. By assessing the factors associated with IPC predictions and validating the IPC against a year of high frequency, household-level food security measures, we aim to improve food insecurity prediction and to supplement existing IPC assessments.

In the next section, we review the literature. We then describe our data and the data matching process. In the third section, we present our empirical strategy and discuss our findings. We first ask whether our high-frequency and spatially disaggregated rainfall and price measures are predictive of the IPC. We examine the relationship between our rainfall measures and local maize prices, to determine whether the relationship between prices and precipitation is sensible, and therefore, is predicting the IPC in a consistent manner. We also estimate how well the IPC predicts household food security, measured as the reduced Coping Strategies Index (rCSI) for 2010, and, once we control for the IPC, whether our rainfall and prices, measured at the IPC level, have any additional predictive power. Finally, we present limitations and robustness checks before concluding.

Literature and overview of early warning for food security

Beginning with Amartya Sen's work in 1980, researchers have identified the drivers of acute food insecurity crises as not merely lack of food, but also lack of access (or entitlements) and lack of political will (see Sen (1980) and Devereux (2009)). At the same time, early warning systems have gained greater attention as a way to identify and track impending food security crises and famines. These early warning systems monitor a variety of information, including remote sensing data, prices, and changes to livelihood patterns. The USAID established FEWS NET in 1985, after the catastrophic famines in Ethiopia and Sudan in 1984, when warnings came too late. For several chronically food-insecure countries, FEWS NET produces maps of livelihoods zones and remote sensing data (e.g., greenness maps from normalized difference vegetation indices (NDVI) data). FEWS also assembles and graphs prices. Vegetation maps and price graphs are produced separately; that is, FEWS does not integrate and process the data together, although FEWS does make recommendations that aim to synthesize the body of evidence. FEWS analyses and recommendations are currently used to determine responses by the US government, including food aid allocations. Early warning data on prices and remote sensing for food security have not yet been combined into a single predictive model, meaning FEWS analysts, prior to the IPC, had to assign relative weights to the indicators (e.g, when is a decrease in greenness a cause for alarm?).

A consortium of agencies developed the IPC, at least in part, as (1) a response and complement to the ad hoc nature of the first order analysis and (2) to address concerns over lack of comparability and consistency in food security assessments across countries, time, and crises. Established in 2004, the IPC is a set of tools and procedures that produces qualitative scales and maps assessing the severity of food insecurity within and across countries over time. The IPC, a sub-national classification of food insecurity, has become the standard metric for early warning of food insecurity. Assessments are issued biannually or quarterly depending on the country and assessments are made at the geographic scale of "IPC zone" – often below a country's district level. By 2015, twenty countries adopted the IPC, and by 2018, the number is anticipated to increase up to 50 (IPC 2015). IPC assessments assist governments and humanitarian actors deciding where and when to target limited resources, standardizing scales of food security crisis magnitude and severity across time, contexts, and situations.[1] Within each IPC zone, IPC assessments assign classifications ranging from 1 = none or minimal to 5 = humanitarian catastrophe / famine. The assessments also include the number of people affected and relevant social characteristics (e.g., pastoralists), causes, and possible responses. See Figure 1 for an example in Malawi.

Figure 1 here

IPC assessments follow a standard set of protocols to facilitate comparability across countries. The IPC uses a convergence of evidence approach (IPC 2012) and the IPC does not collect data, instead relying on meta-analysis of existing information. The IPC manual (2012) explains "Since the IPC approach is not based on a mathematical model, it requires critical thinking on the part of food security analysts. While the IPC is designed to structure the analysis process as systematically as possible, it does require the analysts to have a strong understanding of the concepts and technical details of conducting food security, nutrition and livelihood analysis" (p. 14). The IPC also relies on consensus of the team of trained IPC food security analysts, drawn from civil society and government, to help avoid analyst or institutional biases.

Analysts compare specific indicators that are related to food consumption and livelihood changes, nutritional status, and mortality against standardized IPC reference tables to identify the appropriate food security classification for an IPC zone. The area classification and household group classifications rely on evidence with universal thresholds that are comparable across situations (see the 2012 IPC manual Tables 4 & 5 on pp. 32-33 for specific cut-offs and data requirements). These specific findings need to be buttressed with additional evidence. For example, an IPC area classification "3" or "crisis" is consistent with 20 to 40 percent of the population within an IPC zone has a body mass index (BMI) of less than 18.5 percent (a measure of undernutrition). If at least 40 percent of the population's BMI is less than the benchmark of 18.5, this range is

[1] For more information, see the IPC website: http://www.ipcinfo.org/ipcinfo-about/what-is-the-added-value-of-ipc/en/ For more information on the Malawi IPC zone designations and Household Economy Approach, see the Malawi National Vulnerability Assessment Committee 2005 baseline report.

consistent with the IPC area classification of at least "4", or "emergency." However, before the IPC zone classification is determined, analysts examine additional evidence to ensure these food security measures are trustworthy. The analysts reach consensus on an IPC classification for a given IPC zone based on the total body of evidence for that zone. The analysts also indicate their confidence in the amount and type of reliable data available with a ranking ranging from one to three.

The IPC, by aiming for cross-national and intra-nationally consistent rankings, has added rigor to early warning systems. Nonetheless, the IPC is often limited by the lack of high-frequency and spatially disaggregated data. For example, 2014 IPC assessments in South Sudan indicate areas without assessments, because no data were available (FEWs NET 2014). One proposed solution to the lack of reliable, frequent data is to establish sentinel sites (Barrett 2010). Barrett and Headey (2014) argue that strategically locating sentinel sites in the worlds' most vulnerable locations could provide information for targeting, an ability to monitor vulnerability and early warning as situations deteriorate. They argue for inclusion of vulnerable countries that face natural disasters, high levels of emergency aid, and high levels of child undernutrition. At a sub-national level, the authors illustrate a method of selecting select sentinel sites by combining agro-ecological data with nutritional outcomes, arguing that agro-ecological factors are associated with food security (Barrett and Headey 2014). Yet, within food insecure countries, which sub-national factors are predictive of food insecurity are still not well understood.

Several areas of research outside of food security early warning systems are relevant both methodologically and have the potential to provide additional predictors of food insecurity. For example, considerable ongoing research is working to predict crop production in developed and developing countries using available data on temperature, rainfall, agro-ecology, and satellite imagery. The Global Agricultural Monitoring initiative GeoGLAM tracks crop conditions in countries with poor food security by integrating and analyzing remote sensing data, on the ground reports, and information from national and regional experts. Nonetheless, availability alone is not a sufficient condition to solve food insecurity nor is it the only predictor (Sen 1980).

Earlier approaches, such as poverty mapping, combine multiple sources of data to predict spatially-disaggregated outcomes that are related to food insecurity. Popular in the 1990s and early 2000s, poverty mapping uses small area estimation techniques to combine household-level survey data with census data to estimate poverty over small areas. The technique involves numerous methodological challenges, including considerable data requirements, data non-normality, heteroskedasticity of errors and likely spatial autocorrelation (Davis, 2003). The small area estimation approaches to generating poverty maps require that the data used for the maps are collected during the same timeframe and that the samples are representative of the larger population, which limits the ability of poverty maps to be dynamic and time-responsive (Davis 2003; Hyman et al. 2005; Bedi et al. 2007).

While most poverty maps present income or expenditure information, expenditure and income data have been found to be imprecisely correlated to food security and nutritional outcomes in many parts of the world (Kadiyala et al. 2014; Brown et al. 2017) limiting their usefulness for researchers and policymakers interested in food and nutrition security. Regularly updated poverty maps that include non-income or expenditure measures of food and nutrition security, such as anthropomorphic measures, are far less common. One valuable alternative to small area estimation is to generate maps with community or district level data, which have the benefit of not requiring the same degree of disaggregated data (Kristjanson et al. 2005). These maps also miss the temporality, unless there is access to high frequency community and district level data. In addition, poverty maps are generally descriptive, useful for targeting, but not generally used for analysis or prediction of specific events.

Current spatial analyses of food security or agriculture rarely join spatial data that are representative at different spatial scales. More broadly, techniques to combine remote sensing data with household surveys and price data are in their infancy. Johnson et al. (2013), for example, detail the technical challenges of matching satellite data on Malawian forest cover with Demographic Health Survey (DHS) data whose global positioning system (GPS) coordinates are randomly displaced to ensure confidentiality (Johnson et al. 2013). A second study, discussing approaches to linking environmental data and household surveys, concludes that researchers have not yet converged on how to incorporate food price and risk data (Brown et al. 2014, p. 21). Below we discuss our protocols for combining sub-national IPC, price, and rainfall data.

One challenge with mapping food security is that it is not directly observable (Barrett and Lentz 2014). Food security is often operationalized as adequate physical availability of food, the economic and physical access to food, and the effective utilization of food. These attributes are inherently nested: without availability there can be no access, and without access, there can be no utilization. A fourth component, stability, reflects the temporal dimension of these three aspects. If stability is included in the assessment, individuals who have food at one point in the year but not at a later point, perhaps due to seasonal hunger, ought to be considered food insecure (FAO n.d.).

Analysts interested in (rural) food security at a sub-national level often rely on proxy measures: precipitation or greenness to capture availability and prices to capture access. Utilization is generally measured within households, and therefore rarely explicitly included in early warning systems. Stability, arguably important for early warning, poses serious measurement challenges. As a result, stability is both under-studied and under-theorized. Rapidly available nationally representative surveys capturing seasonal trends in undernutrition are rare (Barrett and Headey, 2014). Few surveys collect household-level food security measures that are both high frequency and spatially representative at the sub-national level. Yet, many households and individuals experience food insecurity for part of year (see Devereux et al. 2008) – and depending on livelihood strategy, incomes, and locations, the duration and timing of household food insecurity will vary. Annual food security surveys likely under-report the population who experiences food insecurity in a given year. Thus, due to data limitations, few long-term, nationally representative studies have evaluated the temporal variation of sub-national food security.

Just as flooding and drought can disrupt food security, limited market access and transactions costs can contribute to food insecurity (Barrett et al. 2009; Timmer 2014). Market access and transactions costs vary not only spatially but temporally, and, as described in the research on famines and severe food insecurity, are worse during times of acute food insecurity (see Sen 1980 and Mukherjee 2015). Given the limited spatial market integration in many developing countries (Mallory and Baylis 2012) and high transaction costs in rural markets (Montgomery et al. 2017), we expect that rural agrarian households who experience acute food security are harmed by a lack of stability in local measures of climate, prices, and other factors. Analysis of spatial and temporal variation within a country can provide insight into the relationship between availability-based (i.e., precipitation) and access-based measures (i.e., prices and stock-outs) of food insecurity and the IPC. Thus, mapping food security, prices, and availability measures with a focus on prediction has clear value.

Data sources and descriptive statistics

Malawi's ongoing history of food insecurity, recent high-profile agricultural interventions targeting small farmers (e.g., Targeted Input Program, Farm Input Subsidy Program), and the breadth of data available make it an ideal candidate for this study. In 2013, Malawi was ranked the 18th poorest country in the world (UNDP, n.d.). In 2010, approximately 70 percent of the population lived below $1.90 per day (World Bank, n.d.). About 85 percent of the population resides in rural areas (UNDP, n.d.). Malawi faces both acute and chronic food insecurity. In the past fifteen years, Malawi experienced several severe food insecurity crises, including a famine in 2002, a food security crisis in 2005, a food security crisis in 2012-2013, and major flooding in southern Malawi in 2015 with resulting widespread food insecurity. In 2013, the percent of children under age five who are moderately or severely stunted was between 37 and 47.8 percent (UNICEF, n.d.; UN WFP, n.d.). The stunting rate, a measure of long-term food insecurity, is the proportion of children below two standard deviations from the median height-for-age reference population.

Malawi's rainy season starts in October and runs through April. While the lean season varies for each family (and across years for families), the FEWs seasonal calendar indicates that the lean season generally runs between November and March. The main crops grown by rural Malawians include maize, sorghum, cassava, and rice, and, as a cash crop, tobacco. While dated, a 2008 report indicates that 90 percent of Malawian agriculture is rainfed (Semu Banda 2008). About 95 percent of farmers in Malawi grow maize and it provides the primary source of calories for Malawians (Jones et al. 2014; MVAC 2015). The Malawi Vulnerability Assessment Committee (MVAC) (2015) argues that Malawians' over-reliance on maize contributes to their food insecurity because maize yields are more susceptible to both droughts and dry-spells, than other staples, such as cassava. In 2005, the Malawi National Vulnerability Assessment Committee (a precursor to MVAC) estimated that about one-third of the rural population did not grow enough food to meet their needs (2005). These households relied on day labor (ganyu) or other activities to cover shortfalls that ranged between two and six months' worth of food.

Because the majority of rural Malawians rely on maize production for food and maize production is heavily dependent on rainfall, estimating food insecurity as a function of maize prices and rainfall can provide us insight into how these factors influence food security. The importance of rain to small farmers and the dominance of maize in the diet means both of these factors may be useful sentinel indicators to understanding temporal and spatial variation in food security at the IPC zone level. Further, food insecure households have little influence or control of these factors, allowing us to gauge the influence of external factors on sub-national food security.

IPC

In 2005, using a Household Economy Mapping Approach, the Malawi National Vulnerability Assessment Committee, which included individuals from the Ministry of Economic Planning and Development, the Ministry of Agriculture, Irrigation and Food Security, the Department of Local Government, the National Statistics Office, the Ministry of Health and Population, WFP, FEWS NET, Save the Children, and World Vision International identified 18 agro-ecological livelihood zones in Malawi. To increase resolution and allow for variation within each livelihood zone, livelihood zones were intersected with sub-district administrative zones, for a total of 60 IPC zones. National parks, which are not contiguous, are given their own zone and not assessed; there are 59 regularly assessed IPC-zones. We use shape files provided by FEWS NET to identify each IPC zone.[2]

We received IPC levels for Malawi from July of 2009 to April 2016 from FEWS NET. Assessments are generally conducted quarterly, although in some quarters, assessments were updated (e.g., if a situation rapidly deteriorates). There are 2360 possible assessments (the number of IPC zones times the number of assessment periods during this period). In 29 out of a total of 2360 assessments, there was insufficient information to complete an assessment. The roughly quarterly IPC assessments in Malawi focus on identifying current acute food insecurity, which is often "on top of" chronic food insecurity.[3] The IPC assessments in Malawi are "IPC compatible", which means that the Malawi assessments are done with a smaller group of IPC-trained technical analysts, primarily housed within the government. The analysts reach "working consensus" rather than complete consensus (i.e., sign off by all stakeholders). IPC compatible assessments are often undertaken due to time constraints but remain directly comparable for the broader IPC community (IPC 2012).

Nearly 77 percent of Malawi's IPC assessments between 2009 and 2016 were phase 1, or "minimal." Slightly over one-fifth (21 percent) were assessed at phase 2, "stressed," and only 3 percent were phase 3, "crisis."

[2] The IPC zones and Household Economy Analyses were updated in 2015, and, from mid-2016 onward, the IPC assessments utilize the 2015 zoning. We use the 2005 IPC zones because they were the base for the IPC assessments prior to mid-2016.

[3] Other IPC protocols focus on identifying chronic food insecurity and or future projections. See IPC 2012.

The periods of greatest phase 3 food insecurity occurred in 2012 (11 percent of assessments for that year), and 2015 (6 percent of assessments for that year). Figure 2 shows IPC and reduced Coping Strategies Index (rCSI) values for two periods of 2010. The IPC rankings are indicated by the borders on the IPC zones, with thicker borders indicating greater insecurity. First, in April 2010, 4 of 58 IPC zones were assessed as phase 2 - severe. The others were phase 1. By July 2010, nine IPC zones were assessed as phase 2 - severe and three zones were assessed phase 3 – crisis. The lean season usually starts in October or November, so the fact that people were already experiencing food insecurity suggests a deteriorating situation.

Figure 2 here

CHIRPS

The Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data set, collects daily data on rainfall at a high spatial resolution (Funk et al. 2015). We utilize the CHIRPS data to compute a series of rainfall variables, described above. USAID funded the development of CHIRPS to support FEWS Net's early warning system, with a focus on drought monitoring. CHIRPS uses a "smart interpolation" approach and has been validated against independent weather station and Global Precipitation Climatology Centre (GPCC) data. CHIRPS data are highly disaggregated, at a resolution of 0.05 degrees, or a little over 5km^2. A dataset with comparable historical reach, the Climate Prediction Center Merged Analysis of Precipitation (CMAP) is collected at a coarser resolution of 2.5 degrees. Boyle et al. (2014) note that the level of spatial disaggregation is critical to understanding land use, as highly aggregated spatial data can miss pockets of important variation, particularly for rainfall.

We use the CHIRPS data to generate several agronomically-relevant measures of precipitation. We calculate a measure of the total rainfall for the prior year, on the assumption that current food security will be largely affected by last year's harvest. Second, we develop a measure of the beginning of the rains for the prior agricultural season, to capture late-onset monsoons, which have been shown to affect agricultural yield (Guan et al, 2015). Third, we develop a measure for length of dry spells during the past rainy season.

Prices

Existing studies on spatial integration find that major markets in Malawi are relatively well-integrated in the long run. Nyogo (2014) finds for six markets in Malawi during 2000 and 2008, long run integration is relatively good, but short run integration is not. Myers (2013) examines maize prices for ten markets in Malawi, during 2001-2008, finding that markets are relatively well integrated and that in general, when price shocks occur, shocks are arbitraged away. However, both of these studies use primary markets where spatial integration may be more likely. No analyses have characterized spatial price integration in Malawi for secondary or tertiary markets where we expect local production to affect local prices, determining food access.

We have high frequency price data for 72 markets distributed across Malawi between 2000 and 2017. Our price data include weekly maize prices from a variety of markets, including remote, rural markets and urban markets. These data were collected by the Malawi Statistical Division of the Ministry of Agriculture on a weekly basis. The Ministry of Agriculture purposively sampled these markets as those that are representative of markets across Malawi (Ministry of Agriculture personal communication). See Figure 3 for locations (some monitored urban markets are in close proximity and appear on the map to share a location).

Figure 3 here

When data are missing for a certain market-week, we interpolate the average market price. We first use weekly prices to generate an average price for each month. When prices for some weeks are missing during a month, we generate the monthly average based on the average for the reported weeks in each month. When no data are collected in an entire month, we linearly interpolate the average monthly price per market using data for that same market in the nearby months. We replace missing monthly prices in the price series with interpolated prices. We have 7416 points of data across 72 markets; 2037 are interpolated values.

Figure 3 plots the coefficient of variation for the retail maize price over two different recent five-year periods.[4] The maps exhibit considerable spatial and temporal variation confirming the value of attending to both spatial and temporal variation in prices. Given the variation in Figure 3 and the limited spatial market integration in many developing countries (Mallory and Baylis 2012; Sekhar 2012), we expect that the effect of agriculture and other factors may be largest at the local level.

rCSI

We utilize data from the 2010-2011 Malawi Third Integrated Household Survey (IHS3) to benchmark the IPC. The IHS collect the reduced Coping Strategies Index (rCSI), a household food security measure, in a nationally representative sample during March 2010 through March 2011. The rCSI is a continuous measure based on responses to a five-question survey module that asks households about the frequency of various coping strategies used. A higher rCSI value indicates greater use of coping strategies, and greater food insecurity. While there are numerous ways to measure food insecurity, the rCSI is a relatively "universal" indicator and seems to best capture lack of quantity of food within a household (Maxwell et al. 2014). However, relative to the Coping Strategies Index, which includes the rCSI questions as well as additional question, the rCSI likely captures less severe coping mechanisms (Maxwell et al. 2013). Thus, the rCSI is well-suited for early warning because it can identify when food security is deteriorating, though it might not capture the full extent of food insecurity once a crisis hits.

GPS coordinates were also collected as part of the IHS3 survey. To maintain respondent anonymity GPS coordinates are randomly offset between one and five kilometres. We use these offset coordinates to identify each household's IPC zone. While it is possible that a handful of households would be categorized incorrectly into a neighboring IPC zone, due to the offset, this should random and, any resulting error should tend toward zero in the aggregate. Offset households were kept within the same district; thus for at least those IPC zone boundaries that are contiguous with district boundaries, the zone assignment will be correct. Each month, the IHS was fielded in several different locations across Malawi. This spatial variation means that, in each month, households were interviewed from most IPC zones. The monthly data collection from a variety of locations allows us to consider how rCSI varies across households both spatially and temporally.

Figure 3 shows the variation of the IPC in 2010. Due to random sampling used by IHS, we do not have statistically representative rCSI values for rural residents for all IPC zones. For the zones in which we observe rCSI values, we see that rCSI values tend to be better for many IPC zones in July, relative to values in April. The lower Shire valley, Malawi's most southern IPC zone, was classified as in crisis in July by the IPC assessment but has a worse average rCSI score in April relative to July. This opens an important question as to whether the rCSI is, relative to the IPC, a bell-weather or leading indicator of increasingly deteriorating conditions. The figure also suggests the value of predicting food insecurity for months when the IPC is not available. Conditions worsened rapidly over the spring of 2010, due to the failure of crops resulting from dry spells during the rainy season (FEWS NET 2010). Being able to identify these patterns as early as possible can help with response efforts.

Data matching process

[4] Of the 63 markets, 8 markets did not have latitudes and longitudes assigned by the Ministry of Agriculture. We found locations for Hewe, Kasiya, Chimbiya, Sharpe Valley, and Embangweni from Geonames, a website that has latitude and longitude coordinates for cities across the globe. Chataloma, Bemebeke Turnoff, and Mayaka were too small to be included in GeoNames. Therefore, we identified latitude and longitude for these markets based on Facebook links with local businesses in each of these markets.

To utilize both price and precipitation data, we aggregated market data from point formats and precipitation and population data from raster formats to the IPC zones, which are polygon locations and our unit of analysis. Combining time-varying spatial data appropriately requires care. We assign market prices to each IPC zone using the following approach. First, we assume that people visit the market closest to them by the straightest path or Euclidean distance. Thiessen polygon boundaries are drawn based on the midpoint between market locations. Everyone within a Thiessen polygon boundary is closest to the market within that polygon boundary. This approach creates a marketshed for individuals, based on market proximity. We then overlay the IPC zones onto the Thiessen polygons and clip each Thiessen polygon by the IPC zone that they fall into. The final market price for each IPC zone is the average of markets prices weighted by populations within the intersection of the IPC zone and marketshed. From a spatial perspective, each IPC zone may include some, one, or none of our 63 sampled maize markets within it. By our approach, people within that IPC zone experience market prices closest to them. An implication of this approach is that residents within each IPC zone could go to a market outside of an IPC zone if it is closer than the nearest market within an IPC zone.[5] Further, using raster-level population statistics from the 2011 Landscan files, we weight marketshed prices by the population of each intersection of the Thiessen polygon and IPC zone to ensure that the resulting prices within the IPC zone polygons are representative of the prices accessible to the population. See Annex 1.

Similar to the market price data, the CHIRPS rainfall data are collected at a finer scale than the IPC zones. Therefore, we aggregate the rainfall data to the IPC zone by taking a population weighted average of the rainfall values within each IPC zone. We again apply 2011 Landscan population values.

Model

Our primary model estimates the IPC classifications as a function of prices and precipitation. Our goal is to produce a parsimonious, first-cut or "IPC-light" model that could supplement the IPC at more granular spatial and temporal scales. We intentionally present the simplest models possible, to assess the ability of readily available high-frequency and spatially disaggregated data to explain the IPC. Our motivation for this parsimony is that a simple model might be more broadly applicable: a simple model relying on widely-available data could be more easily be replicated in other locations facing periodic food insecurity.

Our model utilizes high frequency price data from primary, secondary, and tertiary markets and satellite data on precipitation for Malawi going back to 2009 to estimate historical IPC levels for Malawi at the IPC-zone level. By overlaying mapped measures by season and assessing changes over time, our model has the potential to predict future sub-national food insecurity at relevant spatial and temporal scales. Subsequent models will assess whether adding less available data, such as distribution of fertilizer subsidies across Malawi, substantially improve the overall fit and prediction of our model. Using spatial panel estimation methods (see Baylis et al. 2011), our approach to estimating the IPC is:

$$IPC_{jt} = Lnprices_{jt} + TotalRainfall_{js-1} + FirstRain_{js-1} + MaxNoRain_{js-1} + FloodMax_{jt} + y + m + e$$

where j = the IPC zones; s = rainy season (defined as October to April for Malawi); and t = quarter. The IPC can take values between one to five, although in Malawi, IPC assessments varied between one and three. Lnprices is the natural log of average maize prices within each quarter, *t*, for IPC zone, *j*. TotalRainfall captures the total rainfall during last year's rainy season (between October and April) by IPC zone and year (FEWS NET, n.d.). Incorporating a lag of last year's precipitation allows us to capture the effects of the last

[5] An alternative approach would be to assign people to markets based on road networks (e.g., it may be faster to go to a market that is 5 km away on a paved road than a market that is 3km away on a dirt path). This would be a valuable alternative approach, but Malawi's road network information does not allow us to make this refinement.

growing season on current food security. FirstRain captures the timing of the beginning of the rains for the prior year. It measures the number of days following the first of October when the rains began. The beginning of the rains is defined as when it rains for at least three of the past five days for a total accumulation of at least 10 millimeters. Dry spells during the rainy season are a particular concern for crop production, especially in Malawi (FEWS NET 2010). MaxNoRain is the number of days without rain during last year's rainy season (October – April), which also captures inconsistency of rain potentially harming current food insecurity. FloodMax is the maximum daily precipitation that month in regions that are susceptible to floods. While other variables are intended to capture possible adverse effects of droughts, this variable picks up too much rain, which can cause flooding, such as in the lower Shire valley in 2015. We also estimate a model including year, *y*, and month, *m*, fixed effects to capture any additional unobservable effects. In some specifications, we include IPC zone fixed effects to capture other time-invariant influnences on food insecurity.

Following the estimation of the IPC, we assess the relationship between prices and precipitation in our data. If current price is a good prediction of local food access, and if markets have sufficiently high transaction costs such that they are affected by local production, then our measures of the past year's rainfall should help predict local price. We also control for current flooding, since it may increase market transaction costs. By identifying whether the relationship between prices and precipitation is sensible, we gain insight into how these factors predict the IPC.

$$\text{Lnprices}_{jt} = \text{TotalRainfall}_{js-1} + \text{FirstRain}_{js-1} + \text{MaxNoRain}_{js-1} + \text{FloodMax}_{jt} + y + m + e \qquad (1)$$

where j = the IPC zones; s = rainy season (defined as October to April for Malawi); and t = month. The precipitation variables are the same as defined above, but the natural log of maize prices are now monthly, rather than quarterly, averages.

Both the IPC and our model mean to capture household food security. As noted above, a constraint of food security estimations is that high-frequency food security measures are rare. To assess whether the IPC captures household-reported outcomes of food insecurity, we then estimate a pooled cross-section of 13 months of a food security measure, the rCSI, as a function of the IPC and other variables. The rCSI is a continuous variable. Data on rCSI were collected from households across Malawi between March 2010 and March 2011 as part of the 2010-2011 Malawi Third Integrated Household Survey (IHS3). We leverage the temporal variation in IHS3 data collection during to understand how food security values vary within each IPC zone, by month of data collection. We first estimate the rCSI as a function of our price and weather model described above, to determine if our model directly predicts food security.

$$\text{rCSI}_{jt} = \text{Lnprices}_{jt} + \text{TotalRainfall}_{js-1} + \text{FirstRain}_{js-1} + \text{MaxNoRain}_{js-1} + \text{FloodMax}_{jt} + e_{jt} \qquad (2)$$

where j = the IPC zones; s = rainy season (defined as October to April for Malawi); and t = month. We do not include year fixed effects because we only have thirteen months of data. Instead, we include an indicator variable to indicate when the year is 2011. The IPC is a three category ordinal variable. The omitted variable is IPC = 1, indicating minimal food insecurity.

Second, we estimate the rCSI as a function of the IPC, to establish the relationship between the IPC – a sub-national measure of food insecurity – and the rCSI – a household level measure, aggregated to the IPC region. In one specification, we include IPC zone fixed effects. Next, we extract the residuals from this regression to determine whether our weather and price model used above predicts further variation in the regional aggregated measure of rCSI, over and above that predicted by the IPC. Thus, we estimate the following two equations:

$$\text{rCSI}_{jt} = \text{IPC}_{jq} + m_t + e_{jt} \qquad (3)$$

$$e_{jt} = Lnprices_{jt} + TotalRainfall_{js-1} + FirstRain_{js-1} + MaxNoRain_{js-1} + FloodMax_{jt} + u_{jt} \quad (4)$$

This decomposition allows us to determine whether our precipitation and price variables have value as predictors of household food insecurity, beyond what is captured by the IPC.

Findings

Table 1 presents descriptive statistics for price and precipitation variables by IPC classification. We find that, when moving from IPC = 1 (minimal) to IPC = 2 (stressed), the natural log of maize prices increases and rainfall during the last rainy season falls. Thus, moving from an assessment of 1 to 2 (indicating deteriorating food insecurity) coincides with rising prices and lower rainfall during the last planting season. The shift from an IPC assessment of IPC = 2 (stressed) to IPC = 3 (crisis), does not coincide with higher prices or decreased rainy season rainfall. However, the amount of precipitation in flood zones increases, suggesting that IPC = 3 often captures food insecurity associated with periods of severe flooding.

In Table 2, we present the relationship between the IPC and several household level food security measures drawn from the 2010-2011 IHS. As dietary diversity and food consumption score decrease, indicating the deterioration of dietary quality (Maxwell et al. 2014), the IPC increases. Similarly, as the rCSI increases, indicating the use of more coping strategies, the IPC also increases. A large jump in the number of coping strategies used – from roughly 6 to 10 – coincides with an increase in IPC assessment of stressed (=2) to crisis (=3).

In Table 3, we assess how well monthly measures of price and precipitation across 58 IPC zones in Malawi can predict IPC values. Results from model 1 suggest that last year's rainy season precipitation and days of dry spells and contemporaneous flooding are significantly associated with current IPC rankings in the expected ways. The inclusion of contemporaneous log prices in model 2 strongly predicts the IPC; the rainfall measures' coefficients decrease but are still significant and informative, indicating that the effect of precipitation on food security is not completely captured by local price, possibly indicating that poor rains affect food security both through local incomes as well as local prices. The results also indicate the potential value of tracking the number of days of dry spells to identify possible food security crises in the upcoming year. Each additional rainless day within the longest dry spell during the rainy season is associated with a .0027 to .0034 increase in IPC rankings. This metric is a potentially valuable leading indicator of a region's future food insecurity. The R-square suggests moderate prediction. Our model explains about 17.5 percent of the variation in the IPC assessments.

Turning to unpacking our findings from Table 3, we first confirm that last year's rainfall is strongly predictive of maize prices, which are often considered informative of food insecurity. Table 4 shows a highly statistically significant relationship between prices and total rainfall during the last year and between prices and last year's duration of dry spells. The coefficients' signs are as expected. More rainfall is associated with a lower log price, while longer dry spells are associated with higher prices. Across models 1 – 5, which incorporate alternative specifications of precipitation and/or IPC zone fixed effects, these results remain highly statistically significant. Including IPC zone fixed effects in models 2 and 5 results in qualitatively consistent estimates, although the magnitudes of the coefficients on rainfall decrease. This effect may result from some areas being dryer and higher priced, on average. We cannot distinguish whether our fixed effects are capturing a long-term causal relation between precipitation and location, or whether both of these factors are correlated with a different time-invariant unobservable such as higher transaction costs. Given the relatively small number of seasons that we observe in our data, it is not surprising that some of the explanatory power of annual rainfall is captured by the location fixed effect. In contrast, the coefficient on the number of dry days is larger with IPC zone fixed effects. We also see that contemporaneously high rainfall in the flood prone areas also raise prices. The R-square values are relatively high in these models because of the inclusion of annual fixed effects, which soak up the year-on-year variation.

Ideally, early warning systems would be based on household or individual-level experiences with food insecurity. However, these data are rarely available at a high enough frequency. We utilize the rCSI, collected across 13 months in 2010-2011, to uncover how well the IPC does at capturing household level food insecurity. In Table 5, we turn to estimating the value of household reported rCSI as a function of the IPC and month fixed effects. The month fixed effects are consistent with the lean and post-harvest seasons. The coefficients show that coping strategies tend to increase toward the end of the lean season (i.e., in March and April), and then begin to reduce in July and August, after harvest (Gelli et al. 2017).[6] We also find that increases in the IPC are strongly associated with increasing reports of household food insecurity. When the situation is severe (e.g., IPC ranking is 3), its large coefficient indicates that many more rCSI strategies are being deployed by households. Model 2 introduces IPC zone fixed effects, which absorb much of the explanatory power of the IPC rankings. This suggests that some zones are regularly or chronically more food.

We are also interested in whether our selected price and precipitation measures can explain variation in the IPC. In Table 6, we re-run the same regressions as in Table 4, but estimate the rCSI instead of log price and now include price on the right-hand side. We find similar effects for last season's total rainfall on household food security status as we observed with local price. An increasing amount of rain last year is associated with decreasing rCSI (or better food security). As expected, increasing prices increases household rCSI, indicating that when maize prices increase, households rely on more coping strategies. Results associated with precipitation variables are more mixed. As with Table 1, including IPC zone fixed effects dampens the coefficients on total rainfall last year and current prices, as would be expected since we only observe the food security outcome of a single production season. The fixed effects are highly significant, suggesting that some parts of the country regularly face worse food insecurity, as measured by the rCSI, even when controlling for precipitation and prices.

We also explore whether our price and precipitation variations add any explanatory power to the estimation of rCSI, once we account for the IPC. We decompose the estimation of the rCSI in two steps. In Table 7, we estimate the rCSI as a function of the IPC as model 1. We then compute and store the residuals and estimate these as a function of log prices, precipitation and month fixed effects, shown in Table 7, model 2. Both current log prices, last year's rainfall during the rainy season, and the first day of rain of last year's rainy season are all strongly associated with the residuals, that is the portion of the rCSI unexplained by the IPC. This decomposition shows that while the IPC does predict the rCSI, incorporating price and precipitation data can improve the overall fit and increase the explanatory power.

Robustness checks (forthcoming)

- Evaluate whether results are sensitive to weighing Theissen polygons by population
- We will examine the relationship between the IPC and the FCS. The food consumption score (FCS) measures the lack of adequate quality or dietary diversity of food which captures a different form of food insecurity than the rCSI (Maxwell et al. 2014).
- We will include additional variables into future estimations, including:
    - lagged values for IPC
    - lagged values for prices
    - measure of market thinness and market stock-outs
    - fertilizer inputs from two fertilizer subsidy programs: the Targeted Input Program, and the Farm Input Subsidy Program.
- Estimate ordinal logits for IPC outcomes.

Limitations

---

[6] Available upon request.

Food security data are often representative at different spatial scales, from country-level assessments using FAO production data, to individual-level measures based on child anthropometry Z-scores (Dangour et al. 2012). Our analysis and results are at the sub-district zone-level. The zones used in our analysis are intended to facilitate emergency response, and are relatively coarse (Chris Hillbruner, personal communication). We cannot estimate variations within communities or within households, which remain important drivers of food insecurity and undernutrition. Brown et al. (2017), for example, recently found that nearly 75 percent of food insecure individuals reside outside of the poorest 20 percent of households. New technologies that can target particular households or communities, such as SMS and voice response, offer some possibilities to complement and enrich these findings (Bauer et al. 2015). As of this writing, findings from these technologies have not been adequately validated against more traditional measures of food insecurity, and this area remains an important avenue for future research. Nonetheless, once such tools are better understood and it is established that they do effectively capture food insecurity, incorporating a more localized collection of rCSI data could be a valuable complement to current early warning systems. Pairing IPC assessments with voice-response surveys and or other innovative ways of capturing useful supplemental data could improve the ability to drill down beyond IPC zones to household level food insecurity.

A further limitation is that the IPC in Malawi does not have a great deal of variation. Malawi is also a fairly simple case: there was no conflict, no mass migration, and no famine. Nonetheless, in places where conflicts are emergent or ongoing, our approach can provide an important source of information. Our future research will evaluate the IPC in other locations. Finally, there are other important sources of variation that we hope to incorporate into our future analyses. Measures such as the NDVI and soil moisture could provide valuable insights. Similarly, incorporating crops models and household level factors could potentially improve the fit of our models. We aim to replicate these models elsewhere to improve the models, as well as incorporate and evaluate other forms of data to assess the trade-offs between increasing complexity and parsimony.

Discussion and next steps

Our work takes advantage of underexploited resources: a wealth of time-varying spatial data available for integration, analysis, and prediction of food crises. We find that contemporaneous price and measures of the prior year's precipitation consistently predict the IPC. These results represent a first step in developing a parsimonious early warning model that utilizes real-time data. Incorporating precipitation and market data into modelling of the IPC can provide an important first step for policymakers aiming to determine where to focus data collection efforts, to monitor changes more frequently than the quarterly or semi-annual IPC assessments, and to target assistance efforts.

Our research advances the frontiers of food security early warning and monitoring in three crucial ways. First, in the case of Malawi, we use a small set of readily available data to model food security crises at a fine time scale, both to allow rapid responses and to identify factors that affect regional resilience to rainfall shocks and price increases. Second, our model demonstrates the possibility of making predictions in remote areas of countries, where food-security data are often lacking. Because such areas are more difficult, expensive and time-intensive to reach, improving predictions in these places has a special importance. Following further validation of our model in other contexts, our future work will aim to generate improved predictions out of sample. Third, this is the first quantitative validation that we are aware of the IPC against household level food security data. We find that the IPC is strongly associated with the rCSI, indicating that the IPC method does successfully capture food security, even though the IPC relies on a meta-analysis of available data to reach a working group consensus on IPC rankings. However, tracking precipitation and prices can also improve predictions (e.g., model 2 in Table 3 versus model 2 in Table 7).

The IPC has become the standard for early warning systems for over twenty countries. The objective of our model is to complement the IPC's current approach by developing a first-cut or "IPC-light" model that could supplement the IPC at more granular spatial and temporal scales for more countries. A predictive model, such as the one presented here, is a crucial step toward governments' and NGOs' abilities to target potential food crises in early stages. In other words, our model does not replace an IPC analysis, but rather is a first cut that can help IPC users identify which areas to focus assessments efforts on, or to support analysis and reduce uncertainty of IPC findings in areas where the IPC lacks other sources of data. Potentially, after future demonstrated proofs of concept, our model could be extended to out of sample predictions, which could be particularly important in areas of conflict or areas where situations are rapidly changing. After replicating this work elsewhere, we intend to assess whether maching learning learning algorithms can improve prediction. We also hope to produce analysis of multiple geospatial time series data to identify possible sentinel sites for monitoring food and nutrition security (Barrett 2010). On possible outcome could be the automation of data collection and analysis to support decision-makers.

This work is critical for both programming and research. Practitioners and policymakers are often inundated with many different streams of data that are poorly integrated with one another, if at all. Thus, an ongoing challenge is to identify predictive relationships between readily observable, real-time data, such as weather and price data, and food insecurity, particularly at the sub-national level. A next step is producing maps that incorporate multiple sources of information in a clear, predictive manner that can aid practitioners and policymakers aiming to better target interventions and identify candidate sentinel sites. Mapping and analyzing time-series information on these relationships will aid identifying possible mechanisms by which to improve the targeting and timeliness of interventions and policies. Ultimately, utilizing readily available precipitation and price data to predict these real-time nutritional needs will enable practitioners and policymakers to engage in real-time analysis.

References

Barrett, C. B., Bell, R., Lentz, E. C., and Maxwell, D. 2009. "Market Information and Food Insecurity Response Analysis." *Food Security* 1 (June): 151-68.

Barrett, C. B., and Headey, D. 2014. "Measuring Resilience in a Volatile World: A proposal for a multicountry system of sentinel sites." 2020 Conference paper. May.

Baylis, K., N. Paulson and G. Piras. 2011. "Spatial Approaches to Panel Data in Agricultural Economics," *Journal of Agriculture and Applied Economics* 43(3):325–338.

Bedi, T., Coudouel, A., and Simler, K., ed. 2007. *More than a pretty picture: Using poverty maps to design better policies and interventions.* . World Bank, Washington, DC.

Boyle, S. Christina M. Kennedy, Julio Torres, Karen Colman, Pastor E. Pérez-Estigarribia, and Noé U. de la Sancha. 2014. "High-resolution satellite imagery is an important yet underutilized resource in conservation biology." *PLOS One.* 9(1). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900690/

Brown, C., Ravallion, M., van de Walle, D. 2017. "Are Poor Individuals Mainly Found in Poor Households?" Policy Research Working Paper; No. 8001. World Bank, Washington, DC. https://openknowledge.worldbank.org/handle/10986/26342

Brown, M., Grace, K., Shively, G., Johnson, K., and Carroll, M. 2014. "Using satellite remote sensing and household survey data to assess human health and nutrition response to environmental change." *Population and Environment.* 36: 48-72.

Davis, B. 2003. "Choosing a method for poverty mapping." Food and Agriculture Organization of the United Nations: Rome.

Devereux, S. 2009. "Why does famine persist in Africa?" *Food Security. 1:25.*

Devereux, S., Vaitla, B. and Hauenstein Swan, S. 2008. *Seasons of Hunger: Fighting cycles of quiet starvation among the world's rural poor.* London: Pluto Press.

The Economist. 2017. "The Economist explains: How is famine declared? The UN relies on technocratic calculations but politics still play a part." April 4.

FEWS NET, n.d. "Seasonal Calendar of Malawi." Accessed May 10, 2017. http://www.fews.net/file/87824

FEWS NET, 2010. "Malawi Food Security Outlook Update." May. Accessed May 10, 2017. http://www.fews.net/sites/default/files/documents/reports/Malawi_FSOU_2010%20_05_final.pdf

FEWS NET, 2014. "South Sudan PC Technical Release – Jul/Aug 2014." August. Accessed May 10, 2017. http://www.fews.net/east-africa/south-sudan/food-security-outlook-update/august-2014

Food and Agricultural Organization. 2015. *The State of Food Insecurity in the World: Meeting the 2015 international hunger targets: taking stock of uneven progress.* Accessed January 9, 2017. http://www.fao.org/3/a4ef2d16-70a7-460a-a9ac-2a65a533269a/i4646e.pdf

Food and Agricultural Organization. N.d. "An introduction to the basic concepts of food security." Accessed May 10, 2017. http://www.fao.org/docrep/013/al936e/al936e00.pdf

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell & Joel Michaelsen. 2015. "The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes". *Scientific Data* 2, 150066. doi:10.1038/sdata.2015.66 2015

Gelli, A. et al. 2017. "Lean-Season Food Transfers Affect Children's Diets and Household Food Insecurity: Evidence form a Quasi-Experiment in Malawi." *The Journal of Nutrition: Nutritional Epidemiology.* April 5.

Gettleman, J. 2017. "Drought and Famine Heighten Threat of Not Just 1 Famine, but 4." *New York Times.* March 27. Accessed May 10, 2017. https://www.nytimes.com/2017/03/27/world/africa/famine-somalia-nigeria-south-sudan-yemen-water.html?_r=0

Hyman, G., Larrea, C., and Farrow, A. 2005. "Methods, results, and policy implications of poverty and food security mapping assessments." *Food Policy.* (30): 453-460.

IPC 2012. "Technical Manual Version 2.0: Evidence and Standards for Better Food Security Decisions." The Food and Agricultural Organization of the United Nations. Rome.

IPC 2015. "The IPC Global Strategic Programme (2014-2018): Vision and Action Plan." Accessed May 10, 2017. http://www.ipcinfo.org/fileadmin/templates/ipcinfo/Docs/GSP_Borchure_2015.pdf

Johnson, K., Jacob, A., and Brown, M. 2013. "Forest cover associated with improved child health and nutrition: Evidence from the Malawi Demographic and Health Survey and satellite data." *Global Health Science and Practice.* 1(2): 237-248.

Jones, A., Shrinivas, A., and Bezner-Kerr, R. 2014. "Farm production diversity is associated with greater household dietary diversity in Malawi: findings from nationally representative data." *Food Policy.* 46:1-12.

Kadiyala, S. Harris, J., Headey, D., Yosef, S., and Gillespie, S. 2014. "Agriculture and nutrition in India: mapping evidence to pathways." *Annals of the New York Academy of Sciences.* 1-14.

Kristjanson, P., Radeny, M., Baltenweck, I., Ogutu, J., and Notenbaert, A. 2005. "Livelihood mapping and poverty correlates at a meso-level in Kenya." *Food Policy* (30): 568-583.

Lentz, E. C., and Barrett, C. B. (2013) "The economics and nutritional impacts of food assistance policies and programs." *Food Policy.* 42: 151-163

Malawi, National Statistics Office. 2008. *2008 Malawi Population and Housing Census.*

Malawi National Vulnerability Assessment Committee. "Malawi Livelihoods Baseline Profile. Version 1." September. Accessed May 10, 2017. http://pdf.usaid.gov/pdf_docs/Pnadg538.pdf

MVAC (Malawi Vulnerability Assessment Committee (MVAC). 2015. "Livelihoods Baseline: National Overview Report: 2015." Accessed May 10, 2017. http://www.fews.net/sites/default/files/documents/reports/Malawi-livelihood-baseline-profiles.pdf

Mallory, M., and Baylis, K. (2012). The Food Corporation of India and the Public Distribution System: Impacts on Market Integration in Wheat, Rice, and Pearl Millet. *Agribusiness* 30(2): 225-250.

Maxwell, D. Majid, N., Adan, G., Abdirahman, and Kim, J. J. 2016. "Facing famine: Somali experiences in the famine of 2011." *Food Policy.* 65:63-73.

Maxwell, D., Vaitla, B., and Coates, J. 2014. "How do indicators of household food insecurity measure up? An empirical comparison from Ethiopia." *Food Policy.*

McMillen, D. and Soppelsa, M. (2012). A conditionally parametric probit model of micro-data land use in Chicago. Working paper. Univeristy of Illinois at Urbana-Champaign.

Mukherjee, J. 2015. *Hungry Bengal: War, Famine and the End of Empire.* Oxford University Press: New York, NY.

Myers, R. 2013. "Evaluating the effectiveness of inter-regional trade and storage in Malawi's private sector maize markets." *Food Policy.* August. 41: 75-84. http://www.sciencedirect.com/science/article/pii/S0306919213000468

Nyogo, L. 2014. "Maize price differences and evidence of spatial integration in Malawi: the case of selected markets." *Working Paper.* http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/127993

Semu Banda, P. for Inter Press Service News Agency 2008. "Development – Malawi: Irrigation promises to increase food security." http://www.ipsnews.net/2008/09/development-malawi-irrigation-promises-to-increase-food-security/

Sen, A. 1980. "Famines." *World Development*. 8: 613-621.

Sen, A. 1981. *Poverty and Famines: an essay on entitlement and deprivation*. Clarendon Press: Oxford, UK.

Timmer, P. 2014. *Food Security and Scarcity: Why ending hunger is so hard.* Philadelphia: University of Pennsylvania Press.

UNICEF. N.d. "UNICEF data: monitoring the situation of children and women: Malawi." Accessed May 10, 2017. https://data.unicef.org/country/mwi/

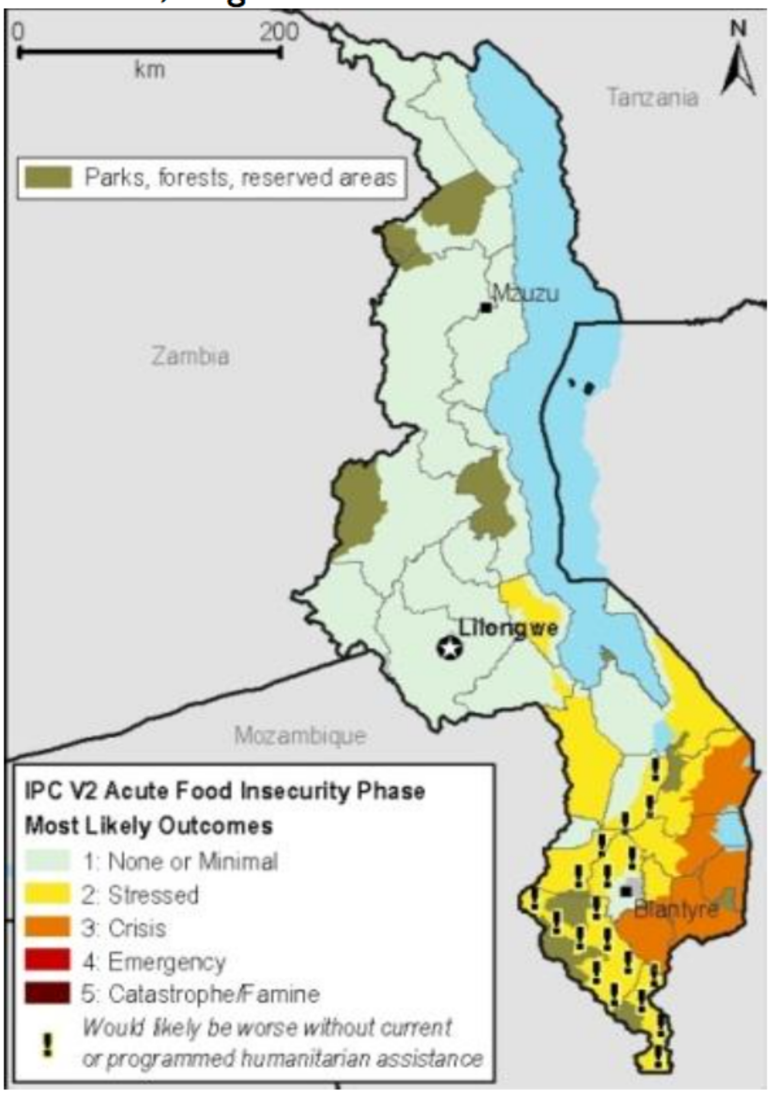
UNDP. N.d. "About Malawi." Accessed May 10, 2017. http://www.mw.undp.org/content/malawi/en/home/countryinfo.html

UN WFP. N.d. "Where we work: Malawi." Accessed May 10, 2017. http://www1.wfp.org/countries/malawi

World Bank. N.d. "Poverty and Equity: Malawi." Accessed May 10, 2017. http://povertydata.worldbank.org/poverty/country/MWI

Figures and Tables

Figure 1. Current estimated food security outcomes, August 2012



Source: FEWS NET

Figure 1: IPC zones for 2012. Source: FEWS NET

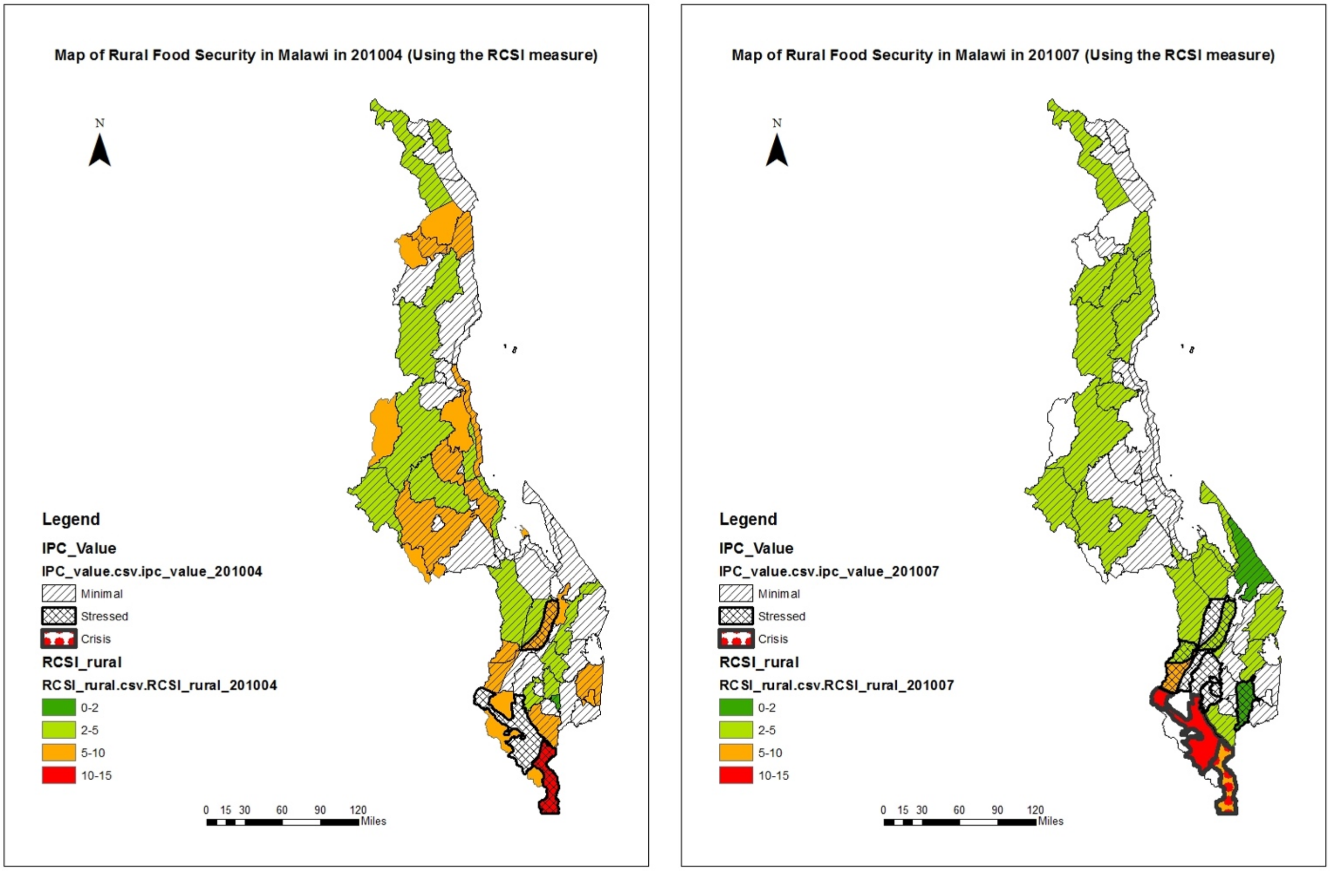


Figure 2: IPC and rCSI values in April and July 2010.

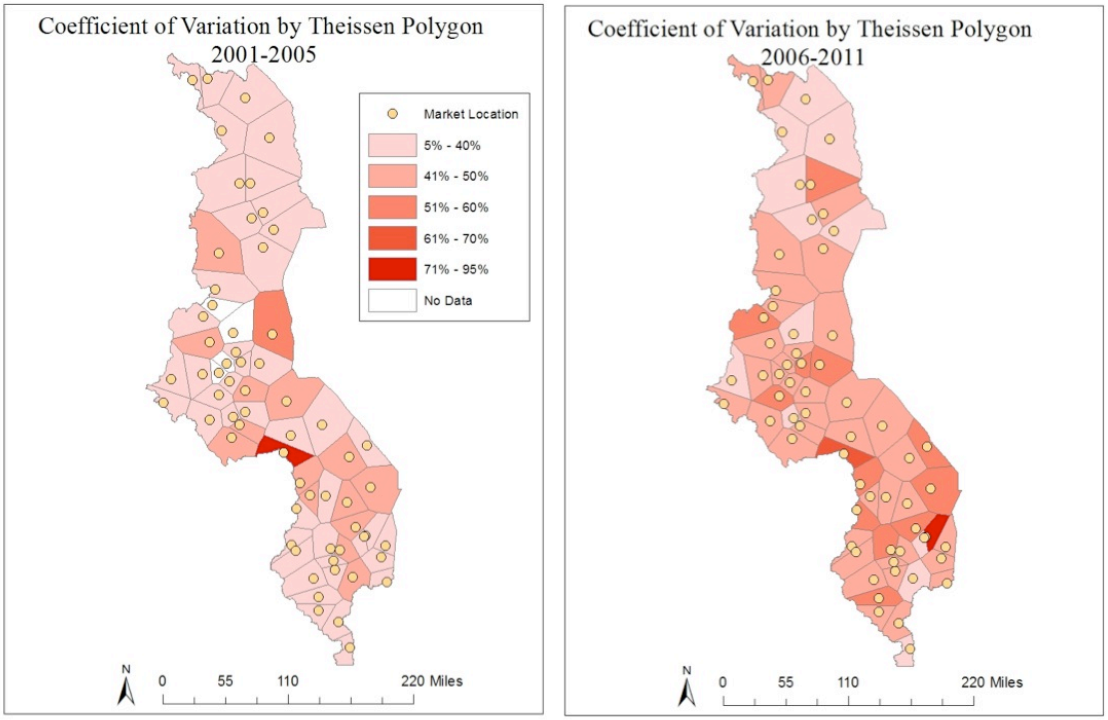


Figure 3: Coefficient of variation (CV), maize retail price, Malawi, 2001-2005 and 2006-2011

Table 1: Mean and standard deviation of price and precipitation data by IPC classification

	All values	All values with IPC values not missing	IPC = 1 (minimal)	IPC = 2 (severe)	IPC = 3 (crisis)
Mean Ln Price	3.93	4.00	3.91	4.36	4.26
SD Ln Price	0.68	0.71	0.72	0.59	0.57
Mean L12 Seasonal Rainfall	962.79	964.45	977.86	911.52	930.59
SD L12 Seasonal Rainfall	198.27	201.00	207.51	162.22	182.11
Mean Max Precipitation * Flood Zone for only flood zones	93.93	93.36	83.15	102.02	107.04
SD Max Precipitation * Flood Zone for only flood zones	109.62	105.01	103.97	107.67	94.87
Mean Max Precipitation * Flood Zone including zeros	4.70	4.77	2.58	12.46	17.26
SD Max Precipitation * Flood Zone including zeros	31.91	31.37	23.25	50.21	54.34
Mean L12 Day of 1st rain	42.04	42.35	42.47	41.87	42.13
SD L12 Day of 1st Rain	19.35	19.26	19.34	18.67	21.05
Mean L12 Max days of dryness	18.66	18.60	18.31	19.84	18.58
SD L12 Max days of dryness	7.82	8.04	8.14	7.79	5.22
n for LN price	5749	4489	3567	828	94
n for precipitation	5760	4933	3900	909	124
n for flood precipitation without zero-values	324	252	121	111	20

Table 2: Household food security measures presented by contemporaneous IPC classification

	IPC=1	IPC=2	IPC=3
RCSI	3.11 (6.08)	5.66 (7.71)	9.71 (8.94)
Food consumption score	49.47 (18.34)	46.16 (16.65)	39.76 (12.96)
Household dietary diversity	5.23 (1.27)	4.99 (1.23)	4.42 (1.30)
n	9,778	1,770	144

Table 3: Relationship between ipc and climate variables, price

VARIABLES	(1) ipc ranking	(2) ipc ranking
lnprice		0.261***
		(0.0435)
total rainfall Oct-Apr L12	-0.000387***	-0.000316***
	(5.73e-05)	(5.80e-05)
no days after Oct 1w five-day rain > 10 and rained at least 3/5 days L12	0.000180	-0.000732
	(0.000617)	(0.000649)
longest dry spell during the rainy season (Oct-Mar) L12	0.00340**	0.00265*
	(0.00141)	(0.00142)
Max daily rain in flood-prone region	0.00142***	0.00145***
	(0.000310)	(0.000317)
Year FE	Y	Y
month FE	Y	Y
Observations	1,749	1,617
R-squared	0.175	0.177

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Relationship between price and climate variables

VARIABLES	(1) maize price (ln)	(2) maize price (ln)	(3) maize price (ln)	(4) maize price (ln)	(5) maize price (ln)
total rainfall Oct-Apr L12	-0.000225***	-0.000160***	-0.000192***	-0.000206***	-0.000178***
	(1.94e-05)	(4.02e-05)	(2.28e-05)	(2.49e-05)	(4.56e-05)
Max daily rain in flood-prone region	0.000119	0.000202	0.000132	0.000123	0.000199
	(0.000118)	(0.000147)	(0.000118)	(0.000118)	(0.000146)
num days after Oct 1 w five-day rain > 10 and rained at lst 3/5 dys	-0.000336	0.000224	-0.000390*	-4.06e-05	0.000378
	(0.000208)	(0.000222)	(0.000209)	(0.000287)	(0.000319)
longest dry spell during (Oct-Mar) L12	0.00301***	0.00536***	0.00307***	0.00150**	0.00285***
	(0.000483)	(0.000601)	(0.000483)	(0.000648)	(0.000781)
positive rain dev 12m lag			-8.04e-07***		
			(2.94e-07)		
negative rain dev 12m lag			2.54e-07		
			(4.38e-07)		
dry zone * tot rainfall				-4.27e-05*	4.02e-05
				(2.28e-05)	(7.19e-05)
dry zone * maxdays L12				0.00347***	0.00587***
				(0.000943)	(0.00119)
dry zone * first rain L12				-0.000461	-0.000126
				(0.000363)	(0.000440)
IPC zone FE		Y			Y
month FE	Y	Y	Y	Y	Y
year FE	Y	Y	Y	Y	Y
Observations	5,166	5,166	5,166	5,166	5,166
R-squared	0.858	0.876	0.858	0.859	0.877

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Relationship between IPC and household food security (RCSI) from the IHS3, 2010-11

	(1)	(2)
VARIABLES	RCSI	RCSI
IPC = 2	2.851***	0.0401
	(0.167)	(0.272)
IPC = 3	7.484***	1.656**
	(0.540)	(0.695)
Month FE	Y	Y
IPC zone FE		Y
Observations	11,692	11,692
R-squared	0.051	0.093

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Relationship between household RCSI and climate variables, price (Similar to Table 3)

VARIABLES	(1) RCSI	(2) RCSI	(3) RCSI	(4) RCSI	(5) RCSI
price (log)	2.043***	0.00143	2.157***	1.784***	0.0207
	(0.250)	(0.326)	(0.252)	(0.254)	(0.327)
total rainfall Oct-Apr L12	-0.00440***	0.00186	-0.00512***	-0.00398***	-0.00295
	(0.000367)	(0.00165)	(0.000442)	(0.000496)	(0.00202)
Max daily rain in flood-prone region	0.00667***	-0.00545**	0.00654***	0.00637***	-0.00531**
	(0.00183)	(0.00259)	(0.00184)	(0.00183)	(0.00259)
num days after Oct 1 w five-day rain > 10 and rained at lst 3/5 dys	0.0264***	0.01000	0.0299***	0.0225***	-0.0239*
	(0.00486)	(0.00917)	(0.00492)	(0.00692)	(0.0135)
longest dry spell during (Oct-Mar) L12	0.0185	0.0689**	0.0146	-0.0542***	0.0713*
	(0.0127)	(0.0309)	(0.0128)	(0.0193)	(0.0428)
positive rain dev 12m lag			1.04e-05***		
			(2.66e-06)		
neg rain dev 12m lag			2.61e-05		
			(1.70e-05)		
dry zone * tot rainfall				-0.00287***	0.00949***
				(0.000412)	(0.00221)
dry zone * maxdays L12				0.113***	-0.0337
				(0.0241)	(0.0624)
dry zone * first rain L12				0.00943	0.0576***
				(0.00763)	(0.0170)
year FE	Y	Y	Y	Y	Y
month FE	Y	Y	Y	Y	Y
IPC zone FE		Y			Y
Observations	12,271	12,271	12,271	12,271	12,271
R-squared	0.057	0.115	0.058	0.061	0.117

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Column (1) regresses the ipc on the RCSI; Column (2) regresses the residuals of this regression on “our model” as it currently stands: price and rainfall information

VARIABLES	(1) RCSI	(2) residuals
IPC = 2	2.552***	
	(0.165)	
IPC = 3	6.602***	
	(0.537)	
price (log)		1.175***
		(0.247)
total rainfall Oct-Apr L12		-0.00123***
		(0.000373)
flood_maxprecip		-2.33e-05
		(0.00202)
num days after Oct 1 w five-day rain > 10 and rained at lst 3/5 dys		0.0126***
		(0.00476)
longest dry spell during (Oct-Mar) L12		-0.00312
		(0.0131)
year=2011		-5.056***
		(0.371)
Month FE		Y
Observations	11,692	11,692
R-squared	0.031	0.049

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Annex 1: Matching marketsheds to IPC zones

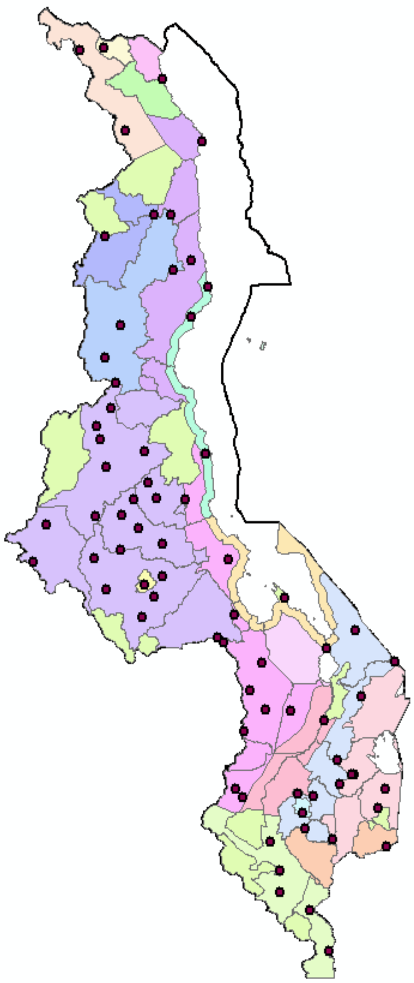


Figure A1. Map of IPC zones and districts, with 63 markets. Some markets are close to one another (e.g., urban markets in Blantyre or Lilongwe) and therefore do not appear as distinct dots.

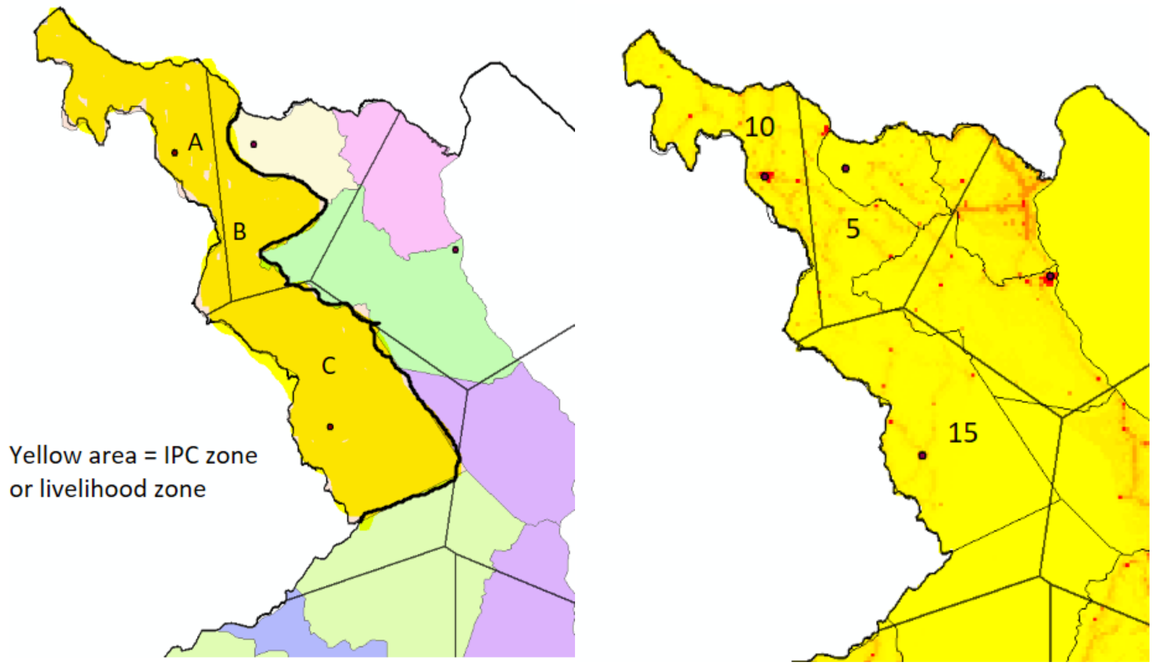


Figure A2. Weighting the Thiessen polygons that fall within a livelihood zone by population where the dot within each polygon is the closest market

Zone A - population of 10 people
Zone B - population of 5 people
Zone C - population of 15 people
Total population for the livelihood zone = Zone A (10) + Zone B (5) + Zone C (15) = 30 people
To compute the weighted price, take the weighted average for each zone's price data by population. Note that the residents in portion B are most proximate to a market in a different IPC zone.