October, 2019

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

Erin Lentz, *University of Texas at Austin*
Hope Michelson, *University of Illinois at Urbana-Champaign*
Kathy Baylis
Yujun Zhou, *University of Illinois at Urbana-Champaign*

World Development 122 (2019) 399–409

Contents lists available at ScienceDirect

World Development

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

A data-driven approach improves food insecurity crisis prediction

E.C. Lentz [a,*], H. Michelson [b,*], K. Baylis [b,*], Y. Zhou [b,*]

[a] Lyndon B. Johnson School of Public Affairs, University of Texas at Austin, Austin, TX 78713-8925, United States
[b] Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

ARTICLE INFO

Article history:
Accepted 10 June 2019
Available online 20 June 2019

Keywords:
Food insecurity
Crisis
Prediction
Early warning
Sub-Saharan Africa
Famine

ABSTRACT

Globally, over 800 million people are food insecure. Current methods for identifying food insecurity crises are not based on statistical models and fail to systematically incorporate readily available data on prices, weather, and demographics. As a result, policymakers cannot rapidly identify food insecure populations. These problems delay responses to mitigate hunger. We develop a replicable, near real-time model incorporating spatially and temporally granular market data, remotely-sensed rainfall and geographic data, and demographic characteristics. We train the model on 2010–2011 data from Malawi and forecast 2013 food security. Our model correctly identifies the food security status of 83 to 99% of the most food insecure village clusters in 2013, depending on the food security measure, while the prevailing approach correctly identifies between 0 and 10%. Our results show the power of modeling food insecurity to provide early warning and suggest model-driven approaches could dramatically improve food insecurity crisis response.

1. Introduction

The global food insecure population is greater than 800 million and rising (FAO, 2017). Currently, international food crisis assessments are not based on statistical models. Instead, assessments use a convergence-of-evidence methodology, in which local stakeholders make infrequent projections based on available data. We envision and propose the following: a set of country-specific models combining readily available data to produce high-frequency, sub-national predictions of food security status. These predictions would complement the existing tools used to assess global food crises. Model predictions would inform status quo methods to support a more effective deployment of assessment resources by identifying particular regions in need of more careful scrutiny or by documenting locations where circumstances are deteriorating quickly. The research we present here is a first step towards achieving this larger objective of developing standardized methods that analysts can use to supplement the current approach, the Integrated Food Security Phase Classification System (IPC), to identify food crises. A lack of near-term, sub-national predictions delays effective response, and governments, nongovernmental organizations, and other donors face considerable challenges allocating scarce resources to mitigate hunger (Headey & Barrett, 2015). Thus, improving forecasts can quickly and directly improve the circumstances of affected populations.

Our model uses only readily accessible secondary data to predict sub-national food security, facilitating its application to other countries and contexts and avoiding primary data collection, a major limitation faced by other proposed estimation approaches (Headey & Barrett, 2015). Using Malawi as a test case, a country with frequent food insecurity challenges, we estimate and forecast three measures of food security as functions of remotely-sensed measures of weather shocks, market prices, and demographic data related to food access. We propose our model as a transparent, replicable, and intuitive contribution to a redesigned global early warning system. Incorporating model-driven predictions into such a system would enhance and hasten humanitarian response and reduce the potential for political manipulation (De Waal, 2018).

The most widely accepted definition of food security includes a hierarchy of components: availability, access, and utilization; and one cross-cutting dynamic component, stability (Webb, Coates, Frongillo, Rogers, & Swindale, 2006). Food must be available for people to access it; target populations must have the logistical wherewithal to access and consume it. Current efforts to anticipate food security crises have focused on using detailed, remotely sensed data to predict local crop production quantities, a measure of availability (e.g., Lobell et al., 2008). While our research builds on these production-focused methods, to our knowledge, ours is

* Corresponding authors at: Lyndon B. Johnson School of Public Affairs, UT Austin, PO BOX Y, Austin, TX 78713-8925, United States (E.C. Lentz). Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, 1301 W. Gregory Drive, Urbana, IL 61801, United States (H. Michelson and K. Baylis).

E-mail addresses: erinclentz@utexas.edu (E.C. Lentz), hopecm@illinois.edu (H. Michelson), baylis@illinois.edu (K. Baylis), zhou100@illinois.edu (Y. Zhou).

https://doi.org/10.1016/j.worlddev.2019.06.008
0305-750X/

the first to predict subnational food insecurity using readily-available data to capture availability, access, and stability in an integrated model. Because efforts focused exclusively on crop production also by definition focus exclusively on availability (Niles & Brown, 2017; Shively, 2017; Hidrobo, Hoddinott, Kumar, & Olivier, 2018), they commonly disregard critical allocation dynamics: how harvested food moves through heterogeneous geographies, infrastructures, and market systems so that households can access it. As economist Amartya Sen famously noted (Sen, 1982. Page 1):

> "Starvation is the characteristic of some people not *having* enough food to eat. It is not the characteristic of there *being* not enough food to eat. While the latter can be a cause of the former, it is but one of many *possible* causes."

The global current standard for early warning to guide emergency aid allocation and response, the IPC, uses information related to both availability and access, but is characterized by its own set of limitations and challenges. IPC classifications are made quarterly or semi-annually, which are infrequent relative to the rate at which food security can worsen on the ground. Moreover, because the IPC uses a Delphic method relying on a convergence of evidence approach rather than a formal model (IPC, 2012), it has faced criticism that it is too complex, requires an excess of detailed information that has uneven availability, produces assessments that are difficult to replicate and confirm, and shows vulnerability to political influence (The Economist, 2017; De Waal, 2018). Finally, the IPC fails to make use of a full scope of available data, particularly readily available secondary data, in a replicable and transparent manner (IPC, 2012).

In short, although the last decade has seen a dramatic increase in the available quantity and quality of data related to food availability and access, including high-resolution measures of soil quality, rainfall, and prices, no current food security early warning and monitoring system systematically incorporates these data into a predictive model.

We use linear and log-linear models based on 2010 data from Malawi to estimate three food security outcome measures, and then use those models to forecast food security outcomes three years later. Our best forecasts accurately predict the out-of-sample food security status for between 65 and 88% of village clusters, depending on the food security measure. Critically important for the effective targeting of resources, while our analysis indicates that the IPC fails to predict any of the most food insecure village clusters in 2013, our model correctly predicts 83–99% of the most food insecure village clusters in the sample. Because our model uses readily available monthly data, our approach generates food security predictions in near real time, providing assessments at least two months earlier than standard IPC quarterly evaluations. We are focused on model cost, replicability and scalability and we identify and evaluate a series of model trade-offs across (1) spatial granularity, (2) availability of a range of predictors, and (3) errors of inclusion and exclusion.

2. Materials and methods

We use readily available data to model the 2010 food security status of village clusters in Malawi, for three, commonly-used food security measures. We model each of these measures as functions of weather and price fluctuations, accounting for geographic variation and intra-annual seasonal trends. We then evaluate the model's accuracy using out-of-sample prediction of 2013 data.

The guiding principle of our modeling approach is to employ data likely to be widely available at high spatial granularity and, when possible, high frequency, allowing rapid, real-time assessment of sub-national food security. In response to concerns about the "black box" nature of the IPC (The Economist, 2017; De Waal, 2018), our method is reproducible, transparent, and parsimonious.

We study model performance across different combinations of spatial scales and data classes to explore the relative improvement in prediction associated with increasing spatial granularity and adding more (and potentially more costly to access) data to the model. Increasing granularity and incorporating additional variables are likely to improve prediction and targeting but both also increase the cost of data processing and access. Therefore, we ask, what is the right spatial level for prediction? And which data seem to contribute the most? We run model specifications at three increasingly granular spatial scales and progressively include three classes of variables as predictors. These data classes range along a continuum: from widely available, remotely sensed data collected daily to information about household assets and demographics derived from infrequently fielded surveys.

Governments, donors and national and international agencies face tradeoffs regarding modeling choices: valuable funds and time spent gathering additional information to refine targeting could, instead, be spent on providing rapid assistance to more people (Basu, 1996). Our interest is to understand the amount of additional precision gained from estimations made at increasingly spatially granular scales and from estimations incorporating variables requiring more resources to access and process. In our results, we highlight the impacts of increasing granularity and adding predictors on accuracy, and assess tradeoffs.

We begin by describing our food security measures and the three classes of data that we use in the analysis. Further details on measuring food security, matching data across spatial scales, and data sources are available in the Supplementary Materials (SM), Sections S1, S2, and S3, respectively. No institutional review board approval was required; we used only anonymized, secondary data. The collecting agencies received informed consent for personal data.

We predict three measures of food security used by international humanitarian organizations including the US Agency for International Development (USAID) and the United Nations World Food Programme (WFP): the reduced coping strategies index (rCSI), the household dietary diversity score (HDDS) and the food consumption score (FCS). We use the 2010–11 Living Standards Measurement Survey (LSMS) for Malawi as our source for our dependent variables because the survey's sampling strategy generates household-level food security outcomes with good temporal variation and spatial granularity. We take the average at the village-level for each measure. Further details on the LSMS sample are in the SM, along with Table S1, which summarizes data sources and how the data were processed.

The three dependent variable measures capture different dimensions of food insecurity; rCSI measures household coping strategies and is associated with food quantity, while FCS and HDDS reflect dietary quality (Maxwell, Vaitla, & Coates, 2014; Vaitla, Coates, Glaeser, Hillbruner, & Biswal, 2017). The HDDS is a count of the number of food categories that a household consumes in a week. The FCS weights this count of food groups according to their nutrient density. Higher values of both the FCS and the HDDS indicate higher food security. The rCSI reflects the number of coping strategies a household uses from a set of five universal strategies to address possible food shortages by frequency and severity; a higher rCSI indicates lower food security. Most often, governments and international agencies apply cut-offs to categorize food security status rather than use the continuous measures of food security (Vaitla et al., 2017).

To predict the three measures of food security: rCSI, FCS, and HDDS, we employ what we term Class 1, 2, and 3 data. These are three nested classes of variables with increasing processing requirements and decreasing availability. We compare predictions of the food security measures using these three data classes to a naïve prediction of cluster (or village-level) average food security using the IPC food security assessment as the only right-hand side variable. We refer to the IPC assessment as our ***Class 0 data***.

The IPC is an appropriate benchmark for our model. It is used to assess food security quarterly or semi-annually in 26 countries; in addition to these countries, a committee of 17 West African countries uses IPC protocols with minor differences to assess food security. In Malawi, an IPC assessment exists for each of Malawi's 60 IPC livelihood zones. The IPC livelihood zones - combinations of livelihood zones and administrative districts - are the highest spatial level of aggregation in our analysis. The IPC assessment data and livelihood zone boundaries are available from the Famine Early Warning System Network (FEWS Net).

IPC-trained teams answer a series of questions and sift through available evidence including food prices, anthropometric measures, and mortality rates, to reach convergence on assigning a food insecurity classification to each IPC zone on a scale of one through five (IPC, 2012). The IPC aims for consistency in assigning food security classifications across the countries where it is regularly used. A challenge is that availability and quality of the evidence used to forecast food insecurity varies across and within countries and over time. As a result, the IPC assessing teams inevitably rely on some interpretation of the evidence.

Class 1 data are readily available data including precipitation, market prices, soil quality and geographic variables combined with the Class 0 IPC food security assessments. Class 1 data are generally collected remotely and are widely available, making them especially useful for difficult-to-reach areas at risk of food insecurity. Agricultural product prices are often readily available and collected in person by government agents, although these can also be collected using cellular phone technologies. We generate measures of agriculturally-relevant precipitation generated from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) dataset (Funk et al., 2015). We use the total amount of precipitation that fell during the October–April rainy season. For the same season, we define the length of the longest dry spell as the number of continuous days with no rain. To measure the beginning of the rainy season, we calculate the number of days after October 1st in which rainfall greater than 10 mm fell 3 days out of 5 (Guan, Sultan, Biasutti, & Lobell, 2015). These three variables are taken from the prior agricultural season to predict food availability for the June/July maize harvest. We include the maximum amount of rainfall in the current month to control for possible flooding for flood-prone regions, which can affect transportation and local economic outcomes. We also include quarter fixed effects to capture seasonality.

We include a monthly average of local market price data of maize (the primary staple cereal), collected weekly from the Malawi Statistical Division (MSD) of the Ministry of Agriculture. In addition to the average price of maize in the market by month, we generate a measure of the number of weeks per month in which the maize price is missing in the nearest market reporting data as a measure of potential local availability and market thinness (Mallory & Baylis, 2016). Along with these high-frequency weather and market variables, we use measures of elevation, distance to market, distance to road, amount of agricultural land in a 1 km radius and an indicator for limited soil nutrient retention as a measure of soil quality. These data come from the 2010–11 LSMS survey but would all be available from secondary sources in other contexts.

Class 2 data are data likely to be available but also likely to require additional work to be accessed and processed, such as data on household roof type (a coarse predictor of poverty) and cell phone ownership. For the Class 2 data, we draw from the LSMS but note that these could come from other sources. We include the percent of households who own a cell phone and the roof type (metal versus thatch), both of which could be collected from sources other than recent surveys, such as cellular companies and remote sensing.

Class 3 data consist of infrequently gathered but publicly available household-level data including demographics and assets. These data are often available in a census or large household survey such as the Demographic Health Surveys or LSMS. Data for our Class 3 variables also come from the Malawi LSMS surveys: demographic data, including the gender and age of the household head, number of household members, and asset holdings.

We emphasize that we extract only a minimal set of variables from the LSMS: food security, basic demographics, and assets. Information on basic demographics and roofing could also be available from a census or rapid assessment tool. We stress this point in order to emphasize that the LSMS is not fundamental to our strategy. Surveys collecting representative, geo-referenced food security data could also be used to train and validate the initial model. For, example, parsimonious, rapid assessments such as SMART surveys or monitoring tools using a community reporter model (e.g., Knippenberg, Jensen, & Constas, 2019) would suffice. Ideally, an initial model would be trained and tested using existing secondary data (such as the LSMS) and then the model's real-time predictions would be validated against periodic high-frequency, household food security data from rapid assessments. These findings could be further incorporated into the model to update and improve the accuracy of forecasts over time.

We conduct the estimation at three spatial scales for two periods: in 2010, our sample is drawn from (a) IPC zone (n = 55, which excludes 5 urban zones), (b) traditional authority (TA) level (n = 281), and (c) cluster of villages, or cluster level (n = 736); in 2013, our sample is (a) IPC zone (n = 46); TA level (n = 153) and village cluster level (n = 204). Class 1, 2, and 3 data are aggregated to the appropriate spatial scale for each spatial scale of estimation. For example, the IPC-district level estimations use the share of households in the IPC zone that own a cellular phone, the zone's average total precipitation during the prior agricultural season, etc. Similarly, the cluster level estimations use the share of households in the cluster that own a cellular phone, or the cluster's average total precipitation during the prior agricultural season.

For the Class 1 data, we develop a protocol to combine monthly data for 72 geocoded markets with seasonal rainfall data to cluster-geocoded annual demographic and asset data. We match markets with gridded rainfall data from the most recent agricultural season. For Class 2 and 3 data, we combine this information with household-level cluster-geocoded information to capture demographic and assets data. See the SM (Section S2) for details and Fig. S1 for a graphical representation of the matching protocol.

We use these data in a linear model to predict average cluster-level HDDS and rCSI, and a log-linear model to predict log FCS. We log FCS in the analysis to place more weight on the lower end of the distribution. We use these regressions in our analysis to enable the replicability of our work in other locations and to evaluate how our model relates to existing food security research; in particular we are interested in understanding the potential magnitude and source of bias in the results when applied out-of-sample.

For each cluster-level spatial scale, we estimate:

$$FS_{it} = \mathrm{b}_{0t} + \mathrm{b}_1 Class1_{\mathrm{jt}} + \mathrm{b}_2 Class2_{\mathrm{jt}} + \mathrm{b}_3 Class3_{\mathrm{jt}} + e_{\mathrm{it}}$$

where *FS* is the cluster average food security measure (rCSI, HDDS, logFCS) for cluster *i* and month *t*. *Class 1–3* data, described above, are included additively across specifications and are averages at spatial level *j*, where *j* is IPC zone, TA, or cluster. At the IPC zone and traditional authority level, we compute the cluster-level food security as predicted by variables calculated at the IPC zone and

TA level, respectively. The *Class 0* estimations only include lagged IPC assessments, which are computed only at the IPC zone.

We also apply a machine learning technique, lasso (least absolute shrinkage and selection operator) regressions, to our data. We chose the lasso technique for a few reasons. First, the lasso yields easily interpretable coefficients, which is less true for more complex techniques. We believe that interpretable coefficients are critical for policymakers interested in understanding and responding to the drivers of food security. Second, lasso techniques are sparse prediction functions, meaning if a particular variable does not contribute to the prediction, its coefficient will reduce to zero. This helps us understand whether the variables we have included in our main linear and log-linear regressions add value to our predictions.

3. Results

We begin by presenting descriptive statistics. We then present our main results: predictions of food insecurity in the 2013 out-of-sample data. We evaluate our model's responsiveness to sample variation by splitting the data by region and apply a lasso regression to our data, comparing findings between our linear and log-linear regression results and the machine-learning model. We then consider errors of inclusion and exclusion associated with our main predictions. We validate our main 2010 model. We examine the role of prices in our out-of-sample predictions and compare results against the IPC.

3.1. Descriptive statistics

Summary statistics of our food security measures and explanatory variables are presented in Table 1. The 2010 measures are from Malawi's 2010–11 LSMS, which surveyed 12,271 households over 12 months. The 2010–11 sample is representative at a national level for each month and, for the crop year, for each district. The 2013 measures used for the out-of-sample prediction are from the 2013 Malawi LSMS, which surveyed 3,999 households. Approximately 60% of our sample reports an rCSI of zero, indicating that the majority of households did not use any coping strategies associated with food insecurity during the survey period.

Food security varies substantially over time and space. Mean rCSI peaks in April in the Malawi 2010–11 data, and FCS and HDDS scores are at their lowest levels in February and March, Malawi's lean season (Fig. 1a). Food security measures also vary systematically over space, with low levels of food security as measured by rCSI in the Shire valley, the southern tip of Malawi (Fig. 1b). Following the poverty prediction literature, we use the cluster-level average of household food security measures as outcomes in our model to reduce within-cluster heterogeneity (Hyman, Larrea, & Farrow, 2005; Jean et al., 2016). A cluster is similar to a village or group of small villages. In the SM (Fig. S2), we present the degree of variation of our 2010 household food security measures explained only by month and/or geographic variation.

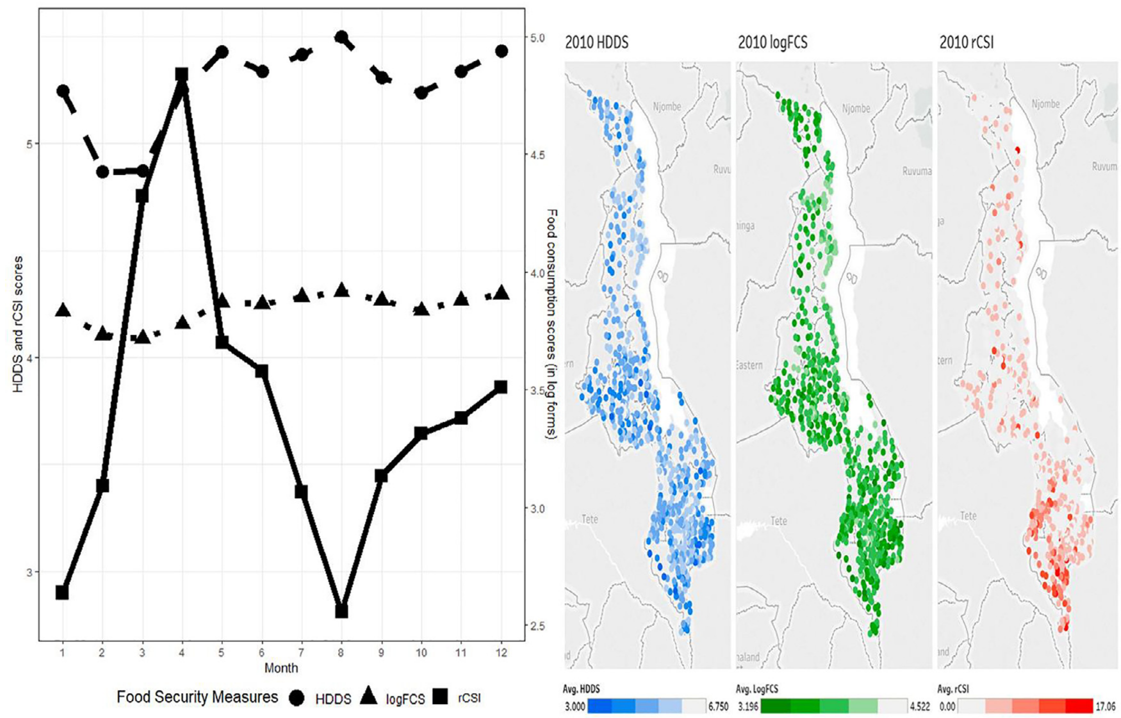
3.2. Out-of-sample predictions: Spatial variation and data availability classes

To generate out-of-sample predictions, we first fit our model to our 2010 Class 0, 1, 2 and 3 data. We retain the coefficients from the fitted model and apply them to the 2013 data. Results show how much of the actual variation in food security status in 2013 correctly predicted by the model.

We first present the share of the variance of the three 2013 food security outcomes explained by our estimation models, as measured by R-squared (Fig. 2). We plot the R-squared of the predicted versus actual food security on the vertical axis for models at different levels of spatial aggregation, listed across the horizontal axis. For example, the ''IPC zone'' level on the x-axis uses data processed at the highest level of spatial aggregation (agro-ecological zones within districts). Moving to the right, the same data are processed

Table 1
Summary statistics of 2010 and 2013 Malawi food security measures and predictors for both household and cluster levels. Ranges of possible values are in parentheses in column 1. Higher rCSI values indicates lower food security; lower HDDS and FCS values indicate lower food security. The 2010–11 LSMS sample is representative at a national level for each month and, for the crop year, for each district. The 2013 food security measures are used for the out-of-sample predictions. We present the cluster average of the household-level variables, which includes food security measures and asset measures.

	Year 2010 (12270 households)					Year 2013 (3999 households)				
Variable	Mean	Median	Std. Dev.	Min	Max	Mean	Median	Std. Dev.	Min	Max
Household logFCS (0–4.72)	3.82	3.83	0.37	1.1	4.72	3.91	3.91	0.36	2.2	4.72
Household rCSI (0–42)	3.68	0	6.51	0	42	4.23	0	6.95	0	42
Household HDDS (0–12)	5.18	5	1.27	1	7	5.56	6	1.15	1	7
	Year 2010 (768 clusters)					Year 2013 (204 clusters)				
Cluster mean logFCS	3.87	3.86	0.21	3.23	4.53	3.96	3.95	0.18	3.45	4.49
Cluster mean rCSI	3.68	2.88	3.05	0.00	17.25	4.26	3.70	2.66	0.00	16.28
Cluster mean HDDS	5.18	5.19	0.70	3.00	6.75	5.55	5.55	0.57	4.10	6.86
Total rainfall (meters)	1.00	0.99	0.18	0.57	1.58	0.93	0.90	0.17	0.59	1.58
First day of rain	45.02	41.00	10.25	3.00	72.00	53.36	51.00	13.26	30.00	72.00
Max days without rain	23.76	21.00	8.14	10.00	52.00	24.30	24.00	6.75	12.00	43.00
Rainfall in flood prone regions (meters)	2.04	0.00	9.66	0.00	75.97	1.11	0.00	7.34	0.00	69.15
Number of cellphones owned	0.60	0.44	0.60	0.00	4.13	0.94	0.70	0.71	0.05	4.40
Maize price (log form)	3.34	3.41	0.39	2.40	5.19	4.46	4.50	0.26	2.77	5.05
Market thinness	0.46	0.49	0.29	0.00	1.00	0.27	0.25	0.16	0.00	1.00
Percent of non-natural roof	0.36	0.49	0.25	0.00	1.00	0.46	0.40	0.27	0.00	1.00
Household size	4.56	4.56	0.68	2.31	7.06	5.00	4.89	0.72	3.58	7.53
Household age	42.15	42.06	4.65	30.81	56.38	42.56	42.37	4.06	33.63	56.19
Household gender (1 for male, 2 for female)	1.24	1.25	0.12	1.00	1.69	1.23	1.22	0.10	1.00	1.50
Asset Index	0.00	−0.02	0.35	−0.83	1.05	−0.02	−0.30	0.54	−0.30	2.94
Distance to road (km)	8.37	4.36	10.15	0.07	56.19	7.68	4.18	8.40	0.06	44.68
Distance to Admarc market (km)	8.06	6.56	5.77	0.38	37.32	7.81	6.24	5.04	1.20	32.89
Percentage of agricultural land	0.36	0.49	0.25	0.00	1.00	0.34	0.41	0.22	0.00	0.75
Nutrition retention constrained	0.28	0.00	0.44	0.00	1.00	0.26	0.05	0.38	0.00	1.00
Elevation (km)	0.87	0.90	0.35	0.04	1.73	0.94	1.02	0.29	0.12	1.55
IPC Value (1 month lag)	1.17	1.00	0.40	1.00	3.00	1.08	1.00	0.22	1.00	3.00



A B

Fig. 1. Food security measures vary across space and time. (A) Average food security measures by month for 2010 vary by season. January = 1. Read rCSI (reduced coping strategies index) and HDDS (Household Dietary Diversity Score) against the left axis and FCS (Food Consumption Score) against the right axis. A high rCSI indicates the use of more coping strategies, thus low food security, whereas a high HDDS or FCS indicates high food security. (B) Darker shades of cluster-level average food security measures in Malawi 2010–11 LSMS data indicate lower regional food security. Specifically, for the household dietary diversity score (HDDS, left panel) and log food consumption score (logFCS, central panel), darker shades indicate lower dietary quality. For the reduced coping strategies index (rCSI, right panel), a darker shade indicates that on average households in the cluster are employing more coping strategies.

at more granular levels of aggregation (the traditional authority or TA, and the cluster level, a cluster of villages). We use a different color to represent each food security measure and a different shape to represent the data classes employed in each iteration of the model.

Our initial estimation uses only the IPC measure as a predictor (Class 0 – represented by circles in Fig. 2B). Because this metric is only compiled at the IPC zone level, higher levels of spatial variation do not increase the amount of variation and it explains very little: from 0.00 for HDDS, 0.005 for logFCS and 0.016 for rCSI. In contrast to the "base" IPC only model, when we incorporate other relevant factors that explain food security and increase the degree of the spatial disaggregation, moving from the IPC zone to the TA and then village cluster, the share of variation explained increases. This is particularly true for the HDDS and logFCS, and particularly for specifications that include some measure of assets (Class 2 and Class 3 data). The specification using only Class 1 data (geographic and price variables) does substantially worse than those using Class 2 data (which add measures of three assets).

Our best predictions employ the full complement of available data (Class 1, 2, and 3) and are conducted at the most spatially disaggregated level, the cluster. Our predictions for the two dietary diversity measures HDDS and FCS, are relatively strong: explaining 64 and 65% of the variation respectively (measured by R-squared). The rCSI prediction fares worse, only predicting 12% of the variation. In comparison to the out-of-sample IPC-only (Class 0) results, our predictions' explanatory power are substantial improvements: the variation in each food security measure explain by our model relative to the IPC: is 64% vs. 0.0% for the HDDS; 65% vs. 0.5% for the log FCS; and 12% vs 1.6% for the rCSI. Across geospatial levels, we identify substantial gains in prediction accuracy for the Malawi model when we move from Class 1 to Class 2 data. Our predictions continue to improve, albeit by smaller amounts, moving from Class 2 to Class 3 data.

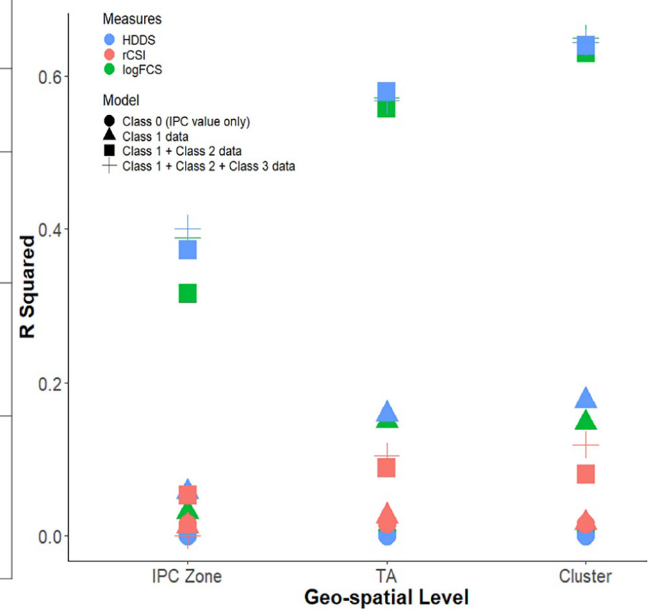
While our model explains more variation in food security as we increase the spatial granularity of the predictor variables (moving to the right along the x-axis), the marginal benefits of granularity are largest when moving from IPC-level to TA-level. The additional explanatory power gained when moving from TA to cluster is smaller. Further, spatial granularity matters more when we include Class 2 or Class 3 data than Class 1 data. Thus, if the predictive model relies solely on measures like rainfall, price and soil quality, a larger geographic aggregation may suffice. Both increasing spatial granularity and incorporating relevant variables that are less easily collected improves prediction and therefore targeting, but will increase the cost and complexity of data collection, processing, and access.

Comparing our out-of-sample results to our 2010 results, we find the amount of explained variation is similar for HDDS and logFCS; while the 2013 models explain 64% and 65% of the variation, the 2010 models explain 91% and 62% respectively (See Table 2). Our out-of-sample results for the rCSI explain less than the in-sample results; in 2010 we explained 34% compared to only 12% in 2013.

In general, our models explain less variation in rCSI, likely reflecting the non-linearity and lower variation in rCSI scores. In both years, the rCSI has a median value of 0, indicating the median respondent did not engage in any coping strategies. In 2010–11, almost two-thirds of the rCSI observations are 0, which limits the observed variation available to train our model. The model is more

	IPC Zone	TA	Cluster Village
Data Class 0	IPC Prediction **+ Quarter Fixed Effect**		
Data Class 1 *Available*	+ Precipitation, Soil Quality, Market Access		
Data Class 2 *Obtainable*	+ Asset Measures		
Data Class 3 *Prior Census*	+ Demographics		

A



B

Fig. 2. The share of variation in out-of-sample cluster-level food security predicted by our models improves with greater spatial granularity and richer data. (A) We predict food security outcomes using three levels of spatial granularity and 4 classes of models. Class 0 data include the IPC early warning value only. Class 1 data contains: past IPC values, precipitation, market prices, market access measures, and soil quality. Class 2 data contains: share of households owing cellular phone and share of dwellings with metal versus thatch roof. Class 3 data contains: household demographics and assets. (B) The explanatory power of the models, measured as R-squared, increases with the Classes of data and the spatial granularity of data used. Our best model explains 64% of the variation in cluster-averages of household dietary diversity (HDDS) and 65% of logged Food Consumption Scores (logFCS). However, at the most spatially disaggregated level, the cluster level, additional household and demographic variables add little additional information.

effective at explaining the variation in food security for the more normally distributed measures of dietary diversity such as HDDS and FCS, than for the truncated rCSI measure.

A second approach to out-of-sample predictions is training our data on some regions and estimating out-of-sample predictions for the remaining regions. This approach allows us to understand how well our model applies across geography, while using both years of data for training. Splitting the sample by geography could be valuable in situations where data is collected in one part of the country but not others. Using data from the southern and central region to predict the north, we find that, while informative, the model performs worse when applied to a different geographies than when used to forecast future years (See Table S2 in the SM for out-of-sample regional results). This result might be explained by the fact that the north has a slightly different agro-ecology and market structure, and thus applying coefficients estimated from the south and central regions does not fit the northern data as well. We do not recommend this approach for extremely remote, conflict-ridden, or cut-off regions, as they may face substantially greater risk of food insecurity compared to regions where data collection is easier. In what follows, we focus on the out-of-sample predictions that vary by time, given our primary interest in improving early warning forecasting.

We also estimate the model for each of the food security measures using a lasso regression. We include all of the variables to allow the lasso to identify which have the most explanatory power. Table 3 shows that the lasso performs reasonably well, particularly for logFCS and HDDS; the lasso predicts 64% of the variation in logFCS (compared to 65% in our main model), 61% in HDDS (compared to 64%), and 6% in rCSI (compared to 11%). The lasso largely selects the variables that we include in our Class 3 models and signs are similar on the estimated coefficients (See Table S3 in the SM for the model coefficients). In fact, for two of the food security measures, the lasso retains all possible variables, confirming that these variables add explanatory power to our base model. The lasso's prediction accuracy is slightly worse for forecasting 2013 than for the base model, suggesting the lasso model may be overfitting. In other words, the lasso may be embedding more information from 2010 than is useful for predicting food security in later years. In contrast, the lasso does better than the standard ordinary least squares (OLS) regression when predicting over region, suggesting the more flexible functional form may perform better when being used to estimate across different climactic and market settings.

3.3. Out-of-sample predictions: Errors of inclusion and exclusion

In what follows, we focus on the results of the specification employing the richest set of explanatory variables conducted at the most disaggregated spatial scale, estimating the model at the cluster level aggregation and incorporating Class 1, 2, and 3 data. These are the results presented furthest to the right on the x-axis in Fig. 2 and represented by the plus signs. Fig. 3 shows a scatterplot of predicted values against actual values for each of the three food security measures. The 45-degree, diagonal lines in each figure indicate where predictions equal measured values.

Our initial results, described above, focus on the amount of variation we explain with our model for continuous outcomes (i.e., the R-squared). However, practitioners arguably care more about how well we classify clusters by food security category (i.e., above or below a cut-off). Such cut-offs are often used by humanitarian agencies to target aid to village clusters suffering from food insecurity as measured by rCSI, log FCS, or HDDS, and therefore predicting cut-offs rather than continuous values may be more useful for

Table 2
Regression results for each food security measure using 2010 LSMS data for Malawi confirms that food security measures are associated with common drivers. The results are estimated at the cluster-level and include predictors from all Class 1 + Class 2 + Class 3. Standard errors are presented in parentheses and asterisks indicate level of statistical significance of coefficients where three asterisks indicate 1%; two indicate 5% and one indicates 10%.

	logFCS	HDDS	rCSI
IPC Value (Previous month)	−0.023	−0.081*	2.223***
	(0.015)	(0.047)	(0.274)
Total rainfall during rainy season	0.104***	0.101	−1.302**
	(0.032)	(0.103)	(0.598)
First day of rain	−0.001	−0.002	0.017*
	(0.001)	(0.002)	(0.010)
Maximum days without rain	−0.0005	−0.005**	−0.047***
	(0.001)	(0.002)	(0.013)
Maize price	−0.001**	−0.002*	0.025***
	(0.0004)	(0.001)	(0.007)
Maize market thinness	−0.010	−0.085	−0.272
	(0.017)	(0.055)	(0.317)
Percent of agricultural land	0.011	0.034	−0.117
	(0.020)	(0.066)	(0.382)
Elevation	−0.060***	−0.048	−1.216***
	(0.018)	(0.060)	(0.346)
Soil nutrient retention constrained	−0.014	−0.079**	0.509**
	(0.012)	(0.040)	(0.229)
Distance to nearest road	−0.002***	−0.010***	−0.018*
	(0.001)	(0.002)	(0.010)
Distance to nearest Admarc market	0.003***	0.005*	−0.007
	(0.001)	(0.003)	(0.017)
Roof type	0.111***	0.572***	0.158
	(0.029)	(0.093)	(0.537)
Number of cellphones owned	0.143***	0.332***	−1.084***
	(0.014)	(0.046)	(0.266)
Household size	−0.015*	−0.041	0.631***
	(0.008)	(0.025)	(0.146)
Household age	−0.003**	−0.015***	−0.016
	(0.001)	(0.004)	(0.021)
Household gender	−0.040	0.064	0.445
	(0.043)	(0.138)	(0.799)
Asset index	0.123***	0.462***	−0.704**
	(0.017)	(0.055)	(0.317)
Quarter 1	−0.059***	−0.193***	−0.169
	(0.015)	(0.048)	(0.276)
Quarter 2	−0.012	−0.048	−0.166
	(0.016)	(0.053)	(0.304)
Quarter 3	0.007	−0.057	−1.121***
	(0.013)	(0.043)	(0.248)
Observations	760	760	760
R^2	0.605	0.623	0.337

these agencies. To assess the accuracy of our predictions, we compute Type I errors of inclusion (i.e., targeting those who do not need assistance) and Type II errors of exclusion (i.e., not targeting those who are insecure) using commonly used cut-offs for food security status. Details on the selection of food security cut-offs are included in the SM.

We quantify errors of inclusion and exclusion by calculating the percent of correct predictions, i.e., those household clusters correctly identified as food secure or insecure. We then calculate the percent of the clusters the model predicts to be in a *better* food security state than they are (an exclusion or Type II error) and the percent we predict to be in a *worse* food security state than they are (an inclusion or Type I error). We compare these numbers against random allocation and the allocation one would obtain if one assumed all clusters were in the largest category. See SM Fig. S4 for the results of analyses conducted at other spatial scales.

We present the cut-offs of food security measures in the scatterplots (Fig. 3), using vertical lines to represent the actual category cut-offs and horizontal lines for the predicted. We shade and label areas of the graph to identify where food security is over-predicted (areas I and II) and under-predicted (areas III and IV). Clusters in area I, where we predict food insecure clusters as food secure, are errors of exclusion. Clusters in area IV, where we predict food secure clusters as food insecure (area IV), are errors of inclusion. Areas II and III indicate areas where the continuous food security measure is over or under-predicted, but the predicted category is the same as the actual.

Given the priority to correctly categorize households while minimizing the number of food-insecure households that we miss, our model performs quite well. For the HDDS in Fig. 3A vertical dotted lines at 3 and 6 indicate food security category cutoffs of low and medium dietary diversity (proxies for low and medium food security, respectively). Most cluster averages lie between these two categories. While our model tends to under-predict food security in the middle-range of the distribution (we predict 79% have lower than actual HDDS values and 21% have higher values than their true level), we correctly classify the food security category of 88% of the sample, outperforming the simple assumption of all households being in the largest category (medium dietary diversity), which would correctly categorize 76%. Similarly, we observe more type I than type II errors (11% versus than less than 1%); thus, our model misses very few clusters that are truly food insecure.

For logFCS, shown in Fig. 3B, the model slightly under-predicts food security for the majority of the sample (we predict 82% have lower FCS values than their true level and 18% have higher). The model performs especially well for households in the lowest food security category in the 2013 data ("borderline"), correctly classifying 83% of the most food insecure clusters. Overall, we correctly predict 76% of the cluster categories, under-estimating the food security category of 23% of the sample (area IV), and over-estimating in less than 1% (area I). This result is much better than random, which would correctly categorize only one third of households. On the other hand, due to the fact that the cut-off levels used by humanitarian agencies are designed to capture relatively rare events, our prediction performs slightly worse than if we simply

Table 3
The comparison of model with different specifications. The results are out-of-sample predictions in 2013. Models are estimated at the cluster-level, using only the IPC value and include predictors from all Class 1 + Class 2 + Class 3, respectively. The percentage of food insecure clusters that are predicted to be in the insecure category is also known as the true positive rate or "sensitivity" of prediction. The categorical accuracy is the percentage of clusters that are correctly predicted to be the same as the actual food security category.

	R squared			Sensitivity of Food Insecure Category			Categorical accuracy		
	logFCS	HDDS	rCSI	logFCS	HDDS	rCSI	logFCS	HDDS	rCSI
Main model	0.649	0.643	0.119	0.833	0.994	0.86	0.759	0.882	0.650
Main model without cluster price variables	0.638	0.661	0.119	0.222	0.968	0.258	0.916	0.882	0.591
Main Model with region fixed effect and without cluster price variables	0.557	0.554	0.073	0.278	0.955	0.290	0.916	0.862	0.616
Main Model with GIEWS price and without cluster price variables	0.634	0.661	0.036	0.722	0.994	0.978	0.793	0.892	0.468
LASSO	0.640	0.611	0.061	0.833	1	1	0.665	0.857	0.463

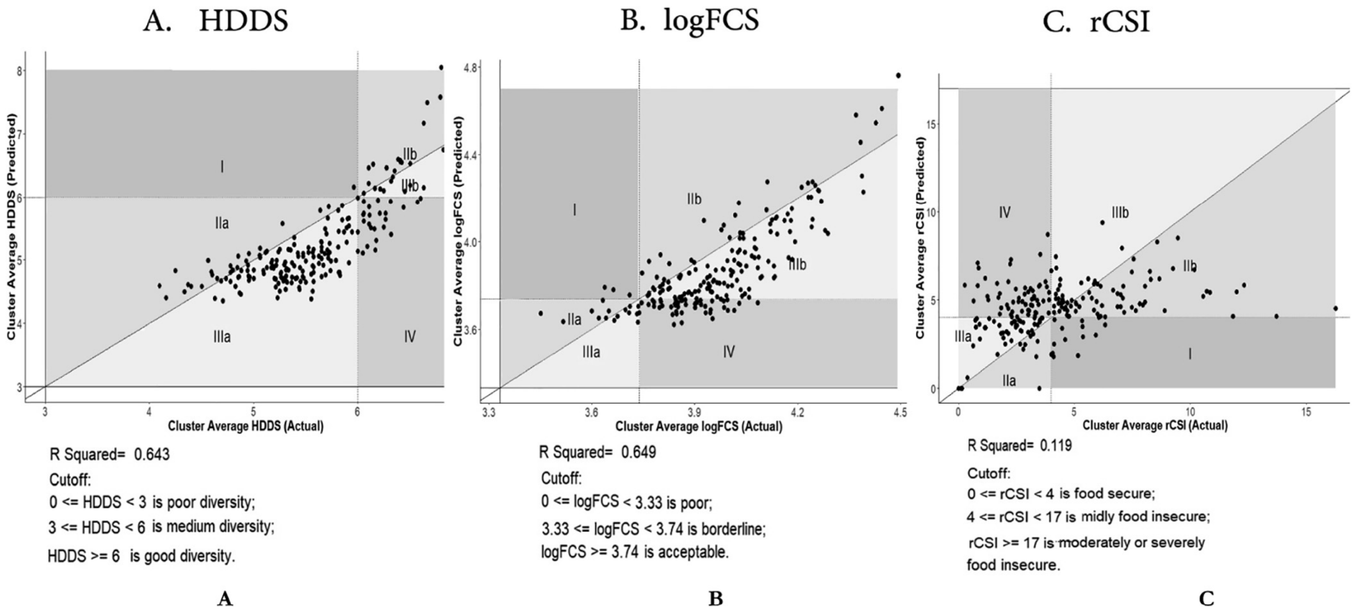


Fig. 3. Actual versus predicted out-of-sample cluster food security estimates, with errors of exclusion and inclusion, where errors of exclusion may be of more concern to relief agencies. Across all three panels, Area I shows areas in which the model under predicts food insecurity status (errors of exclusion) and Area IV shows areas in which the model over predicts food insecurity (errors of inclusion). Areas IIa, IIb, IIIa and IIIb are more nuanced. Areas IIa and IIIa show: clusters where the model correctly predicted that households were food secure but under or over predicted the continuous measure, respectively. Area IIb and area IIIb show: clusters where categorical status was correctly predicted to be food insecure but continuous measures were under and over predicted, respectively. (A) Our model correctly classifies 88% of the household dietary diversity scores. (B) Our model correctly predicts 76% of the log food consumption score categories. (C) Our model correctly predicts 65% of reduced coping strategies index categories.

assumed that each cluster was food secure, on average, which would correctly categorize 91% of the clusters.

For rCSI, shown in Fig. 3C, our best model obtains 65% accuracy, under-predicting the food security status of 61% of households, and over-predicting food security for 39%. As with the HDDS, our best model does better than the naïve prediction that all clusters are food secure on average, which would correctly classify 54%, and considerably better than random allocation, which would only accurately predict the food security category of one quarter of the clusters.

In the SM we include density plots for the predicted and actual values (Fig. S3) and figures presenting the prediction accuracy for the full sample and the tail-only samples (Fig. S4). Relative to predictions using the full 2010 sample, predictions using only the food insecure population decreases errors of exclusion while increasing errors of inclusion, as one might expect. A humanitarian agency aiming to reach most of the food insecure population may prefer fewer errors of exclusion relative to errors of inclusion.

3.4. Validation of 2010 model

The 2013 data allow us to examine the accuracy of our predictive model, given the coefficients on the predictors generated from the 2010 data. Here we validate our model estimated using the 2010 data by asking whether our predictive variables affect food security as expected. We find that across all three food security measures, the signs and significance of included variables are consistent with theory (Barrett, 2010) and other empirical analyses (Jones, Shrinivas, & Bezner-Kerr, 2014; Shively, 2017). Cluster-level coefficient estimates are presented in Table 2 (Table S4 presents these estimates for the clusters in the tails). Holding other variables constant, greater precipitation, earlier start of the rainy season and shorter dry spells during the last agricultural season (for HDDS) and better soil quality improves current cluster average food security (likely through increased production, and thus greater food availability); measures of improved food access such as lower food prices, proximity to road and markets, cell-phone ownership and assets are all associated with greater food security.

In-sample, our most spatially granular estimations with Class 1, 2 and 3 data predict 62% of the variation in HDDS, 61% of the variation in logFCS and 34% of rCSI. The explanatory power is comparable to that obtained in cross-sectional studies that include fixed effects as well as highly detailed information on production and farm size (Jones et al., 2014). Yet, we note that models used in such studies are data intensive and not well-suited for prediction.

3.5. The role of price data

While the data available for early warning has expanded, key data remain missing, incomplete or difficult to access. In particular, high-frequency (e.g., monthly) price data collected from tertiary (rural, remote) markets are rare and there is little research to date on whether primary and secondary market prices are adequately integrated with more remote, rural market prices. In Table 3, we show results using our high-frequency, high geo-spatial resolution prices, compared to (1) having no price data and (2) using the Global Information and Early Warning System (GIEWS) price data for Malawi, which includes monthly maize prices for six cities. In each model, we retain quarterly fixed effects, which help capture seasonal effects, such as pre-harvest hunger.

When relying on the GIEWS data, we find similar results for logFCS and HDDS, but the rCSI predictions are worse, reflecting that the rCSI is more sensitive to prices. This finding is even more pronounced when using only the quarterly fixed effects and no price data; we lose further accuracy in our prediction of rCSI, but we retain good predictive power for logFCS and HDDS (See Table S2 in the SM for results using an alternative specification with month fixed effects). These results suggest that food prices are particularly valuable for understanding rCSI in our Malawi model. Additional research would help us to understand whether focusing early warning data collection efforts on prices from a sample of tertiary markets could yield substantial improvements in early warning, and whether these improvements matter for CSI, for other food security measures, or whether relying on price data from primary markets is adequate.

3.6. Comparison to the IPC, the current approach to food insecurity early warning

A final and primary contribution of this work is validation of the IPC. For the four quarterly assessments that overlap with Malawian household survey data in 2010–11, we find that the IPC assessment (i.e., Class 0 model) is correlated with household-level measures of food insecurity. See Table 4. Among the three food security measures, the IPC is most closely related to measures of rCSI. The rCSI is believed to capture inadequate quantities of food consumed, which is consistent with acute food insecurity, the focus of the IPC (De Waal, 2018). Overall, the IPC assessments anticipate the LSMS household food insecurity measures. That said, a significant constraint of the IPC is that while it may work relatively well on average, it does not capture the distribution of food security outcomes, and therefore may miss the food insecure clusters and households it is meant to identify.

We compare the sensitivity of predictions from our model against those of the IPC. Our sensitivity measure is the percentage of food insecure clusters correctly predicted to be food insecure (also known as a true positive rate) for our 2013 out-of-sample data. High sensitivity indicates greater accuracy in predicting food insecurity hotspots. In Table 5, we present this sensitivity measure for both our model and for a model using only IPC classification values. For FCS, the IPC value only model fails to identify any of the "borderline" food insecure clusters, compared to our model which identifies more than three quarters, resulting in a sensitivity of 83% (no clusters had average FCS scores that were in the poor category). Our model captures four times more of the clusters classified as food insecure based on the rCSI (86% versus 10%). Because the measure is relatively more discrete, and the majority of households are in the medium dietary diversity category, the HDDS results are similar for the two models (100% versus 99%).

Table 4
In-sample IPC value regression results indicate that the IPC is significantly associated with food insecurity but the explanatory power (R-squared) is quite low. Standard errors are presented in parentheses and asterisks indicate level of statistical significance of coefficients where three asterisks indicate 1%; two indicate 5% and one indicates 10%.

	(1) IPC Value	(2) IPC Value	(3) IPC Value
rCSI	0.055***		
	(0.004)		
logFCS		−0.282***	
		(0.069)	
HDDS			−0.087***
			(0.021)
Constant	0.963***	2.255***	1.616***
	(0.021)	(0.267)	(0.109)
Observations	760	760	760
R^2	0.175	0.022	0.023

Table 5
The percentage of food insecure clusters correctly predicted to be food insecure. The results are out-of-sample predictions at the food insecure category in 2013. They are estimated at the cluster-level, first using only the IPC value and then including predictors from all data Classes ("Full Model"). The percentage of food insecure clusters that are predicted to be in the insecure category is also known as the true positive rate or "sensitivity" of prediction.

Model	HDDS	logFCS	rCSI
IPC only	1.000	0.000	0.097
Full model	0.994	0.833	0.860

4. Discussion

Our approach makes several contributions, improving upon the current best practice in early warning predictions and food insecurity analysis. The model decreases the lag associated with current early warning, helping to identify food insecure populations at least two months earlier than the IPC; faster responses save lives (Nikulkov, Barrett, Mude, & Wein, 2016; Gelli et al., 2017). It also targets food insecure communities at a more spatially granular level, better allocating scarce resources. By developing a model that relies on widely available, pre-existing, spatially disaggregated, and temporally frequent data, we ensure that our model can be applied to most countries with sizable food insecure populations and, depending on the method of price data collection, generate predictions in near-real time. We prioritize parsimony and modeling transparency with the objective of reproducibility, including to contexts where other data may be difficult to obtain, such as conflict zones, and the model could be expanded to incorporate data gathered quickly via remote-sensing or cell-phone technology. Thus, the method could be especially useful in areas where on-the-ground information is limited or entirely unavailable. By utilizing only easily accessible, secondary data to construct a transparent, replicable and intuitive early warning system, we not only can enhance humanitarian response but also address current concerns of politicization of early warning resulting, in part, from perceptions that current approaches are black boxes (The Economist, 2017).

4.1. Leveraging readily available data improves early warning

Our results demonstrate that a relatively small set of readily available, secondary data can effectively predict food security, and improve upon current best practice in early warning, the IPC. Our use of readily available data means that agencies, donors, and governments aiming to better predict food security could adopt our approach without additional on-the-ground data collection efforts. Note that while nearly all governments regularly collect the market prices of basic foodstuffs as data inputs into central bank inflation calculations in addition to other applications, these data are not always as widely collected and made available as they are in Malawi. Our research demonstrates the importance of such data for humanitarian applications.

By utilizing relatively simple regression techniques, we have shown that our model is consistent both across food security measures and at different spatial scales. Further, the variables are associated with food security outcomes as expected based on existing theory and research. More complex machine-learning based models may improve predictions, and remain an important avenue for future research. In related work, we apply additional machine-learning methods to this same prediction problem (Authors, 2018). In prediction, variable selection involves a trade-off between bias and variance. We could take an approach where we add a large set of variables, higher order effects and their interactions, which might reduce the prediction variance and increase the prediction bias. Yet, to adopt models without validating that their results are consistent with expectations runs the real risk of biased predictions, potentially harming the very people our models are intended to serve. For these reasons, the simple parametric model proposed here serves as a robust and interpretable alternative.

4.2. Frequency matters

This model can produce high-frequency predictions of sub-national food insecurity. People fall in and out of food security over the course of a year, but currently, standard data collection techniques largely miss such dynamics. Many food security surveys are outdated (for prediction) by the time the data are available. The IPC is assessed quarterly at most, sometimes semi-annually. The data we use in our model are available in near-real time, and the dynamic variables are higher frequency than the IPC; precipitation data are available daily and price data are released weekly or monthly. Even this one-month lag in our model provides information earlier and more frequently than currently available approaches.

Researchers and analysts desiring a model that more closely approximates real-time could use other techniques to collect prices, such as cell-phone based data collection, particularly in regions where government price data collection does not exist. While cellular data is a promising way to increase data frequency, real methodological concerns remain, especially for the collection of household-level data. Potential sample biases in cellular phone survey data are not yet well understood (Blumenstock, Cadamuro, & On, 2015).

4.3. Spatial granularity matters

Our model predicts local food security status at a significantly more granular spatial level than what is currently available. Moreover, spatial granularity in the data can further improve prediction. In particular, when spatially detailed asset and demographic information are available, conducting the analysis at a more spatially granular scale substantially improves predicting power. If such data are not available and the predictive model is limited to spatially correlated Class 1 data measures like precipitation and market data, conducting the analysis at a larger geographic aggregation may suffice.

4.4. Targeting tradeoffs

Our model highlights the tradeoffs faced by food security analysts and policymakers balancing errors of inclusion and exclusion. Our results show that if the objective is to identify those clusters most at risk of food insecurity, analysts would do well to choose a measure that is normally distributed and has categories defined such that most of the measure is not in a single category, particularly if that category is defined as being food secure. As a second best, weighting observations by their food insecurity status to place more emphasis on the lower tail of the distribution can help. The cost is that this approach increases the number of potentially truly food secure households predicted to be food insecure (errors of inclusion). The severity of this drawback depends on the goal of and funding available for programming.

4.5. Policy implications and future work

Our approach is designed to be replicable in other locations by relying on readily available data and utilizing consistent protocols to incorporate data at different spatial levels within a site. Across sites, each type of data is similarly scaled, enabling cross-site comparisons (Sachs et al., 2010). Our approach can be applied to locations that lack recent census or household survey information, such as conflict areas or those in the midst of another crisis, such as a natural disaster.

A limitation of our data is that Malawi did not experience any major shock during this time period. For this reason, we believe that our current model may best be used as a complement to the more Delphic approach used by the IPC that incorporates difficult-to-predict shocks such as conflict or catastrophic events. Predictions of these events remain an open and unresolved area of research (Bazzi et al., 2018). In applications of this model to places with long-term conflict, it may be sensible to include an indicator variable in the model for such regions of instability. Methodologically, such a spatial fixed effects approach to incorporating conflict presupposes that conflict, and its effect on food security, is time invariant. We also note that prices and market assessments provide some information about infrastructure disruptions and shortages likely to occur during conflict, natural disasters, mass migration, or other difficult-to-predict crises. Thus, our model could be expanded to incorporate emergent efforts to predict these events, such as tracking population movements through cell-phone tower usage, collecting data via cell-phones, or scraping reporting in newspapers (Blumenstock et al., 2015).

A second limitation is the degree to which Malawi can serve as a generalizable test case for extending this work into other countries. Two points are relevant with respect to the external validity of our work. First, we argue that the method we propose to integrate and analyze readily available data can be meaningfully applied in other countries. First, we do not intend for the coefficients from our Malawi model to be applied to other countries; the estimated coefficients are unlikely to be applicable in other settings, with different weather patterns and market infrastructure. Malawi is a small country with chronic food security problems punctuated by crises and may provide limited insight into other areas facing different challenges. Instead, we intend the *approach* to be applicable in other contexts. A second concern is that Malawi may be distinct in its available relative abundance of secondary data. In particular, Malawi is distinguished by good quality price data collected reliably and weekly with high spatial granularity (72 markets in a relatively small area) and multiple rounds of LSMS data with more community observations per unit area than found in other LSMS surveys in the region. Application in other countries will require solving country-specific challenges related to identifying appropriate food security outcome data and price data to run and validate the model (e.g., using SMS or other innovative technologies). In Table S5 in the SM, we summarize data sources that could support extending this work elsewhere.

For these reasons, future work validating our models in other contexts is warranted. Our models are more effective at predicting HDDS and logFCS relative to rCSI in Malawi during 2010–2011 and 2013. This may reflect the experience of food insecurity specific to Malawi or it may reflect the underlying distribution of each food security measure (Maxwell et al., 2014). Extending our model to other locations and during periods of more severe food insecurity present exciting possibilities to better understand which outcome measure might be best suited to prediction. Our model improves on the status quo in food security early warning. Building a better, transparent, data-driven early warning system that is replicable and intuitive can save lives and resources, and encourage policymakers to pay more attention to early warnings of hunger and famine.

5. Data and materials availability

LSMS data: http://microdata.worldbank.org/index.php/catalog/lsms.

Weather data: CHIRPS http://chg.geog.ucsb.edu/data/chirps/.

Elevation: NASA SRTM 90 m http://opentopo.sdsc.edu/raster?opentopoID=OTSRTM.042013.4326.1.

Soil: FAO Harmonized World Soil Database http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/.

Agricultural land cover:GlobCover v 2.3 http://due.esrin.esa.int/page_globcover.php.

Global Information and Early Warning System (GIEWS): Food Price Monitoring Tool http://www.fao.org/giews/food-prices/tool/public/#/dataset/domestic.

Additional data related to this paper (i.e., IPC and price data) may be requested from the authors.

Acknowledgments

Hina Acharya, Richard Barad, Gary Eilerts, Christopher Hillbruner, Alison McGuignan, Angela Hamann, Nour Nourey, Caleb Rudow, Anne Speca, and Guyu Ye helped identify and compile data. Chris Barrett, Molly Brown, Peter Christensen, Jason Cons, Christopher Hillbruner, Seppe Kuehn, Daniel Maxwell, Sonja Perakis, Silke Pietzsch, Joanna Upton, two anonymous reviewers, and participants at the University of Illinois at Urbana-Champaign International Policy and Development seminar and Agricultural and Applied Economics Association 2017 and 2018 annual meetings provided valuable feedback.

Funding

This work was supported by the Policy Research Institute at the Lyndon B Johnson School of Public Affairs, University of Texas at Austin; the National Science Foundation Award# 1520683 (HAZARD Sees); and National Institute of Food and Agriculture Hatch Project# 1014548.

Author contributions

EL and HM and KB designed and performed research; KB and YZ analyzed data; all wrote the paper. Senior authorship is shared among EL, HM and KB.

Declaration of Competing Interest

EL, HM, KB, and YZ declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.worlddev.2019.06.008.

References

Barrett, C. B. (2010). Measuring food insecurity. *Science, 327*(5967), 825–828.
Basu, K. (1996). Relief programs: When it may be better to give food rather than cash. *World Development, 24*(1), 91–96.
Bazzi, S., Blair, R. A., Blattman, C., Dube, O., Gudgeon, M., & Peck, R. (2018). The Promise and Pitfalls of Conflict Prediction: Evidence from Columbia and Indonesia. January. Working Paper.
Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science, 350*, 1073–1076.
De Waal, A. (2018). *Mass Starvation: The history and future of famine*. Cambridge: UK: Wiley Press.
The Economist (2017). The Economist explains: How is famine declared? The UN relies on technocratic calculations but politics still play a part. *The Economist*.
Food and Agricultural Organization (2017). The State of Food Security and Nutrition in the World: Building resilience for peace and food security. Accessed January 2, 2018. http://www.fao.org/3/a-I7695e.pdf.
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... Michaelson, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data, 2*. https://doi.org/10.1038/sdata.2015.66 2015 150066.
Gelli, A., Aberman, N. L., Margolies, A., Santacroce, M., Baulch, B., & Chirwa, E. (2017). Lean-season food transfers affect children's diets and household food insecurity: Evidence from a quasi-experiment in Malawi. *Journal of Nutrition, 147*(5), 869–878.
Guan, K., Sultan, B. M., Biasutti, M., & Lobell, D. B. (2015). What aspects of future rainfall changes matter for crop yields in West Africa. *Geophysical Research Letters, 42*(19).
Headey, D., & Barrett, C. B. (2015). Opinion: Measuring development resilience in the world's poorest countries. *Proceedings of the National Academy of Sciences, 112*(37), 11423–11425.
Hidrobo, M., Hoddinott, J., Kumar, N., & Olivier, E. (2018). Social protection, food security and asset formation. *World Development, 101*, 88–103.
Hyman, G., Larrea, C., & 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*. Food and Agricultural Organization: Rome.
Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science, 353*(6301), 790–794. https://doi.org/10.1126/science:aaf7894.
Jones, A., Shrinivas, A., & 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.
Knippenberg, E., Jensen, N., & Constas, M. (2019). Quantifying household resilience with high frequency data: Temporal dynamics and methodological options. *World Development, 121*, 1–15.
Lobell, D., Burke, M. B., Tebaldi, C., Mastrandra, M., Falcon, W. P., & Naylor, R. L. (2008). Prioritizing climate change adaptation needs for food security in 2030. *Science, 319*(5863), 607–610.
Mallory, M., & Baylis, K. (2016). 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., Vaitla, B., & Coates, J. (2014). How do indicators of household food insecurity measure up? An empirical comparison from Ethiopia. *Food Policy, 47*, 107–116.
Nikulkov, A., Barrett, C. B., Mude, A. G., & Wein, L. M. (2016). Assessing the impact of U. S. food assistance delivery policies on child mortality in northern Kenya. *PLoS ONE, 11*(12). doi:10.1371/journal.pone.0168432 e0168432.
Niles, M., & Brown, M. (2017). A multi-country assessment of factors related to smallholder food security in varying rainfall conditions. *Scientific Reports, 7*(16277), 1–11.
Sachs, J., Remans, R., Smukler, S., Andelman, S. J., Cassman, K. G., Castle, D., ... Sanchez, P. A. (2010). Monitoring the world's agriculture. *Nature, 466*, 558–560.
Sen, A. (1982). *Poverty and famines: An essay on entitlement and deprivation*. Oxford University Press.
Shively, G. (2017). Infrastructure mitigates the sensitivity of child growth to local agriculture and rainfall in Nepal and Uganda. *Proceedings of the National academy of Sciences of the United States of America, 114*(5), 903–908. https://doi.org/10.1073/pnas.1524482114.
Vaitla, B., Coates, J., Glaeser, L., Hillbruner, C., Biswal, P., & Maxwell, D. (2017). The measurement of household food security: Correlation and latent variable analysis of alternative indicators in a large multi-country dataset. *Food Policy, 68*, 193–205.
Webb, P. J., Coates, E. A., Frongillo, B. L., Rogers, A., & Swindale, P. Bilinksy (2006). Measuring household food insecurity: Why it's so important and yet so difficult to do. *Journal of Nutrition, 136*, 1404S–1408S.