Received: 4 February 2021 | Accepted: 13 April 2021

DOI: 10.1002/aepp.13175

FEATURED ARTICLE

Predicting poverty and malnutrition for targeting, mapping, monitoring, and early warning

Linden McBride[1] | Christopher B. Barrett[2] | Christopher Browne[3] | Leiqiu Hu[4] | Yanyan Liu[5] | David S. Matteson[6] | Ying Sun[7] | Jiaming Wen[7]

[1]Department of Economics, St. Mary's College of Maryland, St. Mary's City, Maryland, USA

[2]Dyson School of Applied Economics and Management, Cornell University, Ithaca, New York, USA

[3]Center for Applied Mathematics, Cornell University, Ithaca, New York, USA

[4]Department of Atmospheric and Earth Science, University of Alabama – Huntsville, Huntsville, Alabama, USA

[5]Markets, Trade and Institutions Division, International Food Policy Research Institute, Washington, District of Columbia, USA

[6]Department of Statistics and Data Science, Cornell University, Ithaca, New York, USA

[7]School of Integrative Plant Science Soil and Crop Sciences, Cornell University, Ithaca, New York, USA

Correspondence
Linden McBride, Department of Economics, St. Mary's College of Maryland, Kent Hall 203, 47645 College Drive, St. Mary's City, MD 20686, USA.
Email: lemcbride@smcm.edu

Funding information
United States Agency for International

Abstract

Increasingly plentiful data and powerful predictive algorithms heighten the promise of data science for humanitarian and development programming. We advocate for embrace of, and investment in, machine learning methods for poverty and malnutrition targeting, mapping, monitoring, and early warning while also cautioning that distinct objectives require distinct data and methods. In particular, we highlight the differences between poverty and malnutrition targeting and mapping, the differences between structural and stochastic deprivation, and the modeling and data challenges of early warning system development. Overall, we urge careful consideration of the purpose and use cases of machine learning informed models.

KEYWORDS

big data, humanitarian assistance, machine learning, poverty mapping, poverty prediction

JEL CLASSIFICATION

D31; I32; O15

Development, Grant/Award Number: 7200AA18CA00014

Editor in charge: Daniel Petrolia

Recent extreme weather events, food price shocks, the COVID-19 pandemic, and outbreaks of violent conflict have all vividly demonstrated that food emergencies arise quickly and effect their greatest devastation on already vulnerable and disadvantaged communities. Significant information gaps have long impeded effective humanitarian response to food emergencies. This was true during the "four famines" declared by the United Nations in 2017 and remains true amid the current pandemic response around the world today. Agencies responding to such crises or working on longer term development to reduce vulnerability to food emergencies need tools to help manage the coupled challenges of relief and development programming.

The tools agencies need vary according to the task at hand. Accurate and up-to-date household targeting and deprivation mapping can help ensure that scarce resources get directed to the subpopulations that most need and benefit from them, minimizing waste and accelerating relief. Rigorous monitoring and evaluation (M&E) systems can accelerate learning how best to deploy limited resources for maximum impact. Early warning systems can help with resource mobilization and accelerate timely delivery of contextually appropriate interventions. The appropriate intervention may differ, however, by type of deprivation being targeted, mapped, monitored, or forecasted.

The rapid digitization of even low-income regions makes unprecedented volumes of data available in near-real time, potentially enabling rapid and significant advances in development agencies' toolkits. Call detail records (CDRs), mobile phone apps, remote sensing, and social media abound with data. We have more available information about human movement, behavior, and interactions, and on the natural and cultivated environments, than ever before. Simultaneous improvements in data science methods, which we broadly lump under the heading of "machine learning," equip analysts to process ever larger volumes of data faster and extract from those "big data" highly predictive feature sets for a given response variable as well as develop models with highly flexible functional forms.

Machine learning methods are improving the accuracy of household-level poverty and malnutrition targeting tools (Aiken et al., 2020; Blumenstock, 2021; Knippenberg et al., 2019; Kshirsagar et al., 2017; McBride & Nichols, 2018). Recent advances in both data and methods have enabled production of high-quality subnational maps with tolerably accurate estimates of current and probable poverty and malnutrition conditions (Jean et al., 2016; Yeh et al., 2020). Likewise, poverty and malnutrition forecasting for the purpose of early warning is making incremental gains (Browne et al., 2021; Tang et al., 2021; Yeh et al., 2020). While development and humanitarian programming are enjoying a machine learning and big data revolution, there exist real risks that these new data series and methods underdeliver on the promise of improving targeting, mapping, monitoring, and early warning. Careful evaluation and validation work are as important as the efforts to advance the data and methodological frontiers of these tools.

While the appeal of machine learning applications to poverty and malnutrition targeting, mapping, M&E, and forecasting during an era of increasingly big data derives from the methods' ability to produce flexible, data-driven, predictive models, analysts must still carefully consider the purpose and possible use cases of different models. Consider, for example, the trade-offs between a household-level proxy means test (PMT) targeting tool, with household consumption expenditures as the dependent variable, that has as its features (i.e., explanatory variables) a set of household asset holdings that jointly explain permanent income versus a tool that simply includes the feature set that exhibits the greatest predictive power in the available data. The former, asset-based, model might include such assets as materials of housing construction, livestock holdings, and the education of heads of household—all features that well-accepted theory suggests reflect long-term welfare status—while the latter model, built on the most predictive feature set available, might include any strong correlates of current consumption, such as the amount of meat consumed in the last 7 days and whether a household member is currently employed or not. The latter model may better identify the currently poor, but may also lead to greater errors of inclusion and exclusion if the objective is to target the structurally poor. Moreover, opportunistic models may prove more vulnerable to degradation in predictive accuracy over time than might models founded in empirically validated theory.

Trade-offs also exist between highly predictive models built with hundreds of features and smaller, simpler models that predict less accurately but are transparent, easy to interpret, implement, and update. The former may prove more useful for an agency, replete with technical expertise and adequate budget to harvest data continuously, that seeks a poverty map, but may be less useful for poverty targeting, program evaluation, monitoring, and the like, for a smaller operational agency if the more onerous data collection task that supports the model is infeasible to undertake on either an urgent or regular basis. Overall, effectively fitting tools to tasks requires attention to an operational agency's objective function (Zhou et al., 2021).

As Carter and Barrett (2006) note, most of the work we do in identifying the poor is "almost unavoidably backward looking" due to the nature of data; this observation also applies to the task of identifying the malnourished and food insecure.[1] However, the question most operational agencies seek to answer is: who will be poor or malnourished in the future? Of course, no method can reveal the future with certainty; the combination of machine learning and big data does not provide us with a crystal ball. We can, however, leverage big data and machine learning methods to make effective, efficient targeting tools by paying careful attention to the trade-offs to various approaches to modeling and data inputs. Valuable progress has been made on all these fronts. We review and discuss this progress while also highlighting the differences between targeting and mapping, the differences between identification of the stochastically poor and the structurally poor, and discussing the modeling challenges of developing early warning systems as well as data collection efforts to support all these efforts going forward.

TARGETING VERSUS MAPPING

Emergencies disproportionately impact places that suffer high levels of chronic poverty. Therefore, understanding spatial patterns of deprivation is one of the primary analytical tasks confronting an operational agency. Both targeting and mapping tools assist with these endeavors as well as related tasks such as M&E. The basic difference between targeting and mapping is that targeting identifies households or individuals while mapping identifies geographic areas. However, the two can and often are productively combined. For example, one might use geographic targeting based on poverty maps as a coarse-level tool to identify poor regions or regions with high welfare heterogeneity and then develop PMTs to target individuals/households within the targeted geographic areas (see Blumenstock, 2021 for a great application of such two stage targeting). Below, we make some distinctions between targeting and mapping, note recent advancements, and identify research frontiers in both areas.

At its most basic level, a targeting tool such as a PMT assigns weights to a set of easily identified household characteristics so as to approximate either households' welfare or households' probability of falling below a welfare threshold (Coady et al., 2004; Grosh & Baker, 1995; Schreiner, 2007). In most cases, stock variables, such as household asset holdings, are on the right-hand side of a targeting model, and flow variables, such as consumption, food security status, or a binary indicator variable capturing consumption or food security status below a threshold, are on the left-hand side. Many development and social protection programs identify eligible participants based on PMTs. They also use such tools for monitoring, evaluation, and impact evaluation purposes, where appropriate.

Recent innovations in poverty and malnutrition targeting have built upon the "scorecard" approach to tool development—a scorecard includes a short list of observable household characteristics with corresponding estimated weights (Schreiner, 2007)—using machine learning algorithms and techniques for dimension reduction and out-of-sample validation. Building on the basic function of a PMT, McBride and Nichols (2018) demonstrate the targeting gains of out-of-sample testing for model selection, whether using standard approaches, such as logit and quantile regression, or machine learning approaches, such as random forests. Recognizing that agencies need simple scorecards and often have to build them using limited (as opposed to "big") data, Kshirsagar et al. (2017) demonstrate that highly predictive models using no more than 10 variables can be built on limited data using cross-validation and parameter regularization for feature selection.

Kshirsagar et al. (2017) also highlight the important role of bootstrap aggregation—parameterizing a model by averaging over many subsamples from the data—within model development to achieve a model that performs consistently well across subpopulations within a country. Baez et al. (2019) emphasize the value and importance of using child malnutrition indicators, as opposed to consumption or expenditure measures, as the outcome variable in a targeting model when one is interested in child well-being among agricultural households. In producing a drought-contingent targeting model using Demographic and Health Survey (DHS) data combined with Normalized Difference Vegetation Index (NDVI), Baez et al. (2019) find that simple and easily interpreted methods such as logistic regression and classification trees do just as well as black box methods such as random forest and gradient boosting in the prediction of child-level stunting, suggesting that the trade-offs between interpretability and strong prediction may be modest. Baez et al. (2019) also make clear the value of augmenting relatively sparse data, such as DHS household-level surveys, with big data such as NDVI satellite data for targeting purposes.

One challenge of targeting tools is that model parameters may shift over time and space, meaning the model needs to be reparameterized to remain accurate when applied to a new (including future) population. This is a challenge because nationally representative household surveys are few and far between. Another challenge is that targeting tools require further data collection for application to a given subpopulation for the purpose of targeting: to be used for targeting or monitoring purposes, the model weights must be applied to current household assets and characteristics. Therefore, important innovations in targeting tool development include those that look to new data sources for parameterization and/or application of the model.

Such innovations include Altındağ et al. (2021), Blumenstock (2021), and Aiken et al. (2020). Altındağ et al. (2021) demonstrate that a PMT tool developed to assist the Lebanese government in disbursing aid to Syrian refugees could be applied to administrative data already held by the government. This innovation accelerates disbursement and reduces costs by obviating the need for a costly household survey. Blumenstock (2021) uses a representative mobile phone-based survey, in areas identified through a poverty mapping exercise, to build and deploy a PMT targeting tool based on mobile phone use during the COVID-19 pandemic. With the noted limitation that those who do not own a mobile phone SIM card will not have access to the targeted transfers (and the research team does a great deal to limit the population that will fall into this category), this approach resolves many of the challenges of conventional targeting tool development and deployment and will become more effective only as mobile phone ownership becomes more ubiquitous. Likewise, Aiken et al. (2020) find that CDRs perform as well as a survey-based PMT in identifying the ultra poor in the Balkh province of Afghanistan for the subset of households that own phones. They also find, however, that CDRs are not strong predictors of household wealth in general in this setting, likely due to sample homogeneity. Use of satellite imagery to capture household-level asset holdings, such as housing roof top material, may be another way forward in streamlining household-level targeting.

Poverty mapping is a related but distinct task from household- or individual-level targeting. In contrast to targeting tools, mapping focuses not on households but on geographic aggregates showing the spatial distribution of welfare (Coudouel & Bedi, 2007; Elbers et al., 2003; Ghosh & Rao, 1994). Historically, mapping has been conducted at a relatively coarse scale, such as first or second administrative levels within or across countries. With increased spatial precision from various Earth observation products, it is now feasible to produce welfare estimates at far more precise—village, or sub-village, prospectively even plot—scales, which means the lines between geographic mapping and household-level targeting are blurring. Further blurring the lines is the fact that the geographic aggregation of household-level data may be used to communicate welfare distributions even though the disaggregate, household-level data are readily available. The distinction between mapping and targeting then may come down to whether it is the objective of the analysis to identify variation in welfare across geographic space, or across households as identified by their observable characteristics (possibly including their geographic location).

Recent advances in poverty mapping and small-area estimation include the use of mobile phone records, nightlights data, daytime satellite imagery, and other remote sensing data combined with various machine learning models (or models more recently grouped into machine learning, such as elastic nets) for increasingly accurate projections of the spatial distribution of poverty, food (in)security, welfare, or measures of human development (Blumenstock et al., 2015; Browne et al., 2021; Engstrom et al., 2017; Head et al., 2017; Hersh et al., 2020; Jean et al., 2016; Masaki et al., 2020; Noor et al., 2008; Pokhriyal & Jacques, 2017; Steele et al., 2017; Yeh et al., 2020).

In particular, Blumenstock et al. (2015) introduce CDRs as a predictor of the distribution of asset wealth in regions of the world for which household-level data are otherwise sparse. Jean et al. (2016) demonstrate that convolutional neural nets (CNN) trained to predict nighttime light intensity from daytime satellite imagery can capture 55%–75% of the variation in asset wealth (as reflected in the widely used and internationally standardized DHS wealth index) within a given country and 41%–56% of the variation in consumption expenditures. While Jean et al. (2016) find success in predicting asset wealth, they find less success predicting consumption expenditures. This finding is echoed by findings in Barriga Cabanillas et al. (2021) and Tang et al. (2021); overall, stock measures appear easier to predict than flow measures.

While the trend in mapping is to combine disparate data sources, it should be noted that, as the data input demands grow, so does the burden and expense of updating the model/map. A number of papers directly address these and other data challenges. Hersh et al. (2020)

demonstrate that reasonably accurate poverty maps can be developed using only open source data, reducing the cost burden of map development and maintenance for national statistics offices. Yeh et al. (2020) demonstrate that CNNs trained on only nightlights or only multispectral daytime imagery do about as well as those trained on both nightlights and daytime imagery in the prediction of asset poverty across space, as the inputs capture the same variation. Such observations help reduce the input demands for high-resolution spatial mapping.

Yeh et al. (2020) also find that nightlights data performed poorly relative to daytime imagery in predicting asset wealth over time, in large part because the nightlights data did not vary sufficiently over time in the sample regions. This finding highlights another key challenge of mapping with big data: often the data that are most abundant have limited variation among the poorest of the poor (Blumenstock, 2016, 2020). For example, Barriga Cabanillas et al. (2021) demonstrate the limitation of CDRs in the prediction of food security outcomes in Haiti; they find that these records do not capture variation at the lower end of the welfare distribution. Given the limitations of nightlights and CDRs, some recent mapping models focus on use of NDVI, solar-induced chlorophyll fluorescence (SIF), and other remotely sensed data that may better correlate with the welfare of households reliant on local agricultural systems for their livelihoods or for food supplies (Browne et al., 2021; Tang et al., 2021).

In addition to poverty mapping, recent advancements have been made in the mapping of other measures of deprivation. For example, Njuguna and McSharry (2017) combine data on mobile phone ownership and the number of calls per phone with nightlights and population density data to map the multidimensional poverty index for Rwanda. Pokhriyal and Jacques (2017) also map the multidimensional poverty index within Senegal at the commune level by using a combination of data sources including mobile phone records, climate and vegetation data, data on soils and crop production, and other remotely sensed and geospatial data.

Big data and machine learning-informed mapping methods have not performed as well in predicting indicators of malnutrition as they have in predicting asset wealth and poverty. For example, Head et al. (2017) replicate the success of Jean et al. (2016) in predicting asset wealth, explaining approximately 70% of the variation in the DHS wealth index across space using satellite imagery. They also find moderate success predicting electricity, education, and mobile phone ownership, with r^2 values from 24% to 64%. However, they find very poor performance in predicting child anthropometric outcomes.

A significant challenge in predicting malnutrition outcomes is that they tend to be measured with noise, and some workhorse measures, such as severe wasting, are, thankfully, less prevalent than poverty (Baez et al., 2019; Head et al., 2017). To address this challenge, Zhou et al. (2021) use oversampling combined with random forest and gradient boosting models; they demonstrate that oversampling rare and noisy food insecurity outcomes in model development can significantly improve predictive accuracy in the targeting and forecasting of food insecurity at the cluster level. Noting that malnutrition outcomes are strongly correlated with wealth, which Jean et al. (2016), Head et al. (2017), and many others predict with accuracy, Browne et al. (2021) propose joint prediction of malnutrition and wealth measures, who find modest improvement in predicting malnutrition indicators over time using multivariate random forests. The objective is to parse signal from noise by harnessing multivariate analysis; this line of research merits further exploration.

Mapping and targeting differ fundamentally in terms of objectives and (geographic) scale of output, though they can be, and often are, fruitfully used together. Machine learning and big data have enabled exciting advancements in the efficiency and accuracy of both targeting and mapping. However, we join Head et al. (2017) in urging caution when generalizing the results of successful approaches such as Jean et al. (2016) to other contexts and outcomes. Sufficient variation in the input variables and sufficient correlation with the outcomes of interest are minimum requirements for a targeting or mapping model to accurately identify the poor or malnourished. In comparison to mapping, targeting has not enjoyed as much innovation in recent years. This may be due to the fact that geography is a strong determinant of poverty and malnutrition, and therefore household-level targeting is not as great a priority as is geographic targeting for operational agencies. It is also in part due to the fact that big data are not yet abundant at the household level. Targeting advances by Blumenstock and co-authors based on CDRs prove the exception to this; however, as noted above, in many places these data still offer limited information about the population of interest, the poorest of the poor (Barriga Cabanillas et al., 2021; Blumenstock, 2020).

STRUCTURAL VERSUS STOCHASTIC POVERTY AND FOOD INSECURITY

The poor include those who are always poor as well as those who move in and out of poverty, with some evidence that the latter group makes up a substantial proportion of the poor at any given point in time. For example, Baulch and Hoddinott (2000) find that 20%–65% of households across 13 panel studies are classified as "sometimes poor," a category more numerous than the "always poor" classification. They also find considerable heterogeneity in the duration of poverty within the transitory poor.[2] Likewise, Knippenberg et al. (2020) find that most food-insecure households in Malawi transition in and out of food insecurity, with significant heterogeneity in food insecurity spell length. Distinguishing between structural and stochastic deprivation[3] is important for well-targeted interventions and requires the targeting, mapping, monitoring, or early warning model to account for the structural determinants of deprivation (Carter & Barrett, 2006).

How can one account for the structural determinants of deprivation in such a model? Because permanent (i.e., expected) income and thus consumption expenditures are endogenous to the stock of productive assets one controls, identifying the structural determinants of poverty entails considering the dynamics of asset accumulation, where "asset is understood to broadly include conventional, privately held productive and financial wealth, as well as social, geographic and market access positions that confer economic advantage" (Carter & Barrett, 2006). Much empirical work supports this asset-based theory of welfare dynamics (Balboni et al., 2020; Barrett et al., 2006; Barrett et al., 2019; Carter & Lybbert, 2012).

An asset-based understanding of welfare dynamics can assist one in identifying the structural poor in a targeting or mapping setting. One useful, data-driven, distinction is offered by Carter and May (1999): households that are currently poor (according to their current level of consumption or expenditures) but for whom a conditional expectation function of consumption/expenditures given current asset holdings predicts a non-poor standard of living are considered stochastically poor; those that are currently poor and for whom a conditional expectation function of consumption/expenditures given current asset holdings predicts a poor standard of living are considered structurally poor. This distinction reinforces the value of including asset holdings as features when developing targeting models where the primary group of interest is the structural poor. Asset holdings have the additional advantage that they are often verifiable, whereas other predictors of welfare, such as how much meat was consumed in the past week, may be more challenging to verify. However, there is tension between the asset-based theory of welfare dynamics, wherein asset holdings are important in determining long-term deprivation, and standard big data/machine learning approaches to model development, which are often based on the most predictive feature set for a given response variable. Big data rarely include household-level asset holdings, and models with the highest out of sample r^2 may not do a great job in identifying the structurally poor and food insecure, as these are generally a subset of the overall currently poor.

Where the distinction between stochastic and structural poverty or food insecurity is needed for well-targeted interventions, more household-level survey data may be needed. Where such data are not available, mapping for the purpose of geographic targeting may best capture the persistent spatial distribution of poverty and malnutrition. Geography typically reflects the confluence of multiple market and state failures and may directly influence structural poverty and malnutrition. Moreover, geographic targeting is relatively quick and cost effective, especially as compared to household- or individual-level targeting. Indeed, there is growing evidence that much persistent poverty is place-based and that escape from such poverty often involves migration (Beegle et al., 2011; Jalan & Ravallion, 2002; Pritchett & Hani, 2020; Ravallion and Wodon, 1999; de Weerdt, 2010). Yeh et al. (2020) suggest, and take a few steps in the direction of, using poverty mapping as an avenue to learn more about the determinants of geographically distributed structural poverty; this is an important direction for further research.

STATIC VERSUS DYNAMIC MODELS

Can advances in machine learning methods and data availability improve early warning systems, such as would allow us to predict the mass movement into (or deepening of) poverty or malnutrition in an upcoming period? For the identification of individuals, households, and/or regions that will experience poverty in the next period, and the anticipation of the consequences of shocks on vulnerable populations, dynamic rather than static models may be necessary. By dynamic models we mean models that take as inputs and produce as outputs changes over time. With few exceptions, poverty mapping and poverty targeting efforts tend to produce static models, defined as fixed-point or interval values (for stocks and flows, respectively) rather than as intertemporal change in those values. This is in large part due to data limitations. Although similar in nature, the tasks—early warning versus identifying the presently poor or malnourished—require different tools and inputs. Most critically, dynamic models require panel data, that is, repeated observations over time of the units of observation.

Several works have made progress in these directions. Mude et al. (2009) generate a famine early warning and emergency needs assessment model using lagged high-frequency data from northern Kenya. Lentz et al. (2019) demonstrate that simple linear regression with high-frequency data inputs outperforms the prevailing food insecurity early warning model, the Integrated Food Security Phase Classification (IPC) System, in Malawi. Using such inputs as remotely sensed climate data, data on food prices, and demographic data, Lentz et al. (2019) make near-real-time food security predictions that would allow agencies to monitor food security dynamics over time and at a local (survey cluster) level. Cooper et al. (2019) produce a map that predicts where drought is likely to have the most severe impacts on child stunting by accounting for factors such as arid environments, poor governance, and political instability. And as mentioned above, Zhou et al. (2021) use oversampling to improve the forecasting of food insecurity at the cluster level, while Browne et al. (2021) jointly predict next period asset and malnutrition outcomes at the cluster level using multivariate random forests. Andree et al. (2020) predict food crises in 21 countries between 2009 and 2020. With particular attention to the trade-offs between false positive and false negative rates in prediction outcomes, Andree et al. (2020) demonstrate that model developers can adjust the loss function of their random forest model to give more or less conservative (in terms of the extent to which they avoid false positives) predictions. Like that of Zhou et al. (2021), this approach pays careful attention to agency needs and common decision-making parameters. Using this strategy, Andree et al. (2020) are able to predict future food crises with lower false negative rates than can the currently used IPC System.

In addition to the above strategies, each of which predicts levels of food security or malnutrition for early warning and monitoring purposes, several early warning efforts have focused on predicting changes in outcomes based on changes in inputs over time. Tang et al. (2021) demonstrate that a CNN trained on changes in NDVI over time can predict future changes in poverty in Uganda. Yeh et al. (2020) also show that mapping models can predict changes over time with modest accuracy. The Yeh et al. (2020) deep-learning model, using multispectral daytime imagery, captures 17% of the out-of-sample variation in cluster level in changes in welfare. Using simulations, Yeh et al. (2020) demonstrate that an r^2 of 17% is about as well as a model could do predicting out-of-sample changes in the wealth index, given the variation in the index over time.

An additional promising area of research for early warning purposes relates to development resilience (Barrett & Constas, 2014). Recent theoretical (Barrett & Constas, 2014) and empirical developments (Cissé & Barrett, 2018; Knippenberg et al., 2019) in the study of development resilience suggest that, with high-frequency longitudinal data, targeting expected resilience may be possible. Conceptualizing resilience as a measure of future well-being allows one to estimate the probability of staying above a welfare threshold as a function of assets and exposure to shocks (Cissé & Barrett, 2018; Knippenberg et al., 2019). Along these lines, Baez et al. (2019) propose a drought-contingent early warning system to identify regions where children are at risk of stunting due to drought. Knippenberg et al. (2019) estimate a food security early warning system, with the food security coping strategies index as the response variable, using LASSO and random forests for feature selection.

An exciting data development for potentially enhancing early warning models is the increasing availability and accessibility of satellite measurements of SIF (Frankenberg et al., 2011; Mohammed et al., 2019; Sun et al., 2017; Wen et al., 2020). Because SIF is the only remotely detectable optical signal that occurs during photosynthetic machinery, it possess exciting potential to reveal physiological stress earlier than can the more commonly used NDVI measure of vegetation growth/health (Daumard et al., 2010; Song et al., 2018). In addition, whereas the conventional use of NDVI to capture vegetation growth usually involves recalibration for different cropping systems, SIF, mechanistically linked to photosynthesis itself, has the potential to be more scalable for yield estimation across crop types (Guanter et al., 2014; Sun et al., 2017). However, at present, SIF data do not have as high a resolution as NDVI (Wen et al., 2020). This remains a central limitation to their use for early warning.

In fact, many of these early warning models depend on, or would perform better with, access to higher frequency and higher resolution data from surveys or satellites. While high-frequency survey monitoring data can increasingly be supplemented by remote sensing data—for example, NDVI, SIF, weather, food prices, daytime or nightlights imagery—to significantly improve poverty and malnutrition M&E, mapping, and targeting, data preprocessing and maintenance is time consuming and expensive. Multilateral collaboration to produce and maintain cloud-based platforms with curated open access data would facilitate low-cost data reuse and augmentation, along with model development and improvement. Such a platform would allow researchers to pick up where others have left off. Google Earth Engine makes some raster data available, but the data list is still not comprehensive. Efforts such as the data gateway of the Food Security Portal, hosted by IFPRI (https://api.foodsecurityportal.org/), are a useful step in this direction. And recent developments such as the Rolf et al. (2020) Multi-task Observation using Satellite Imagery and Kitchen Sinks, a procedure for the extraction and labeling of "task-agnostic" features from satellite imagery, promise to significantly increase the accessibility of remotely sensed data.

For those data that cannot be observed via remote sensing, high-frequency data collection using sentinel sites, which allow long-term monitoring and measurement of welfare and well-being, would go a long way toward supporting development resilience and early warning (Barrett, 2010; Headey & Barrett, 2015). Taking advantage of widespread access to mobile phones, such sentinel sites could be set up affordably for the purpose of collecting a small set of key features not observed through remote sensing, while costlier "thick" surveys, which include welfare outcomes of interest for ground truthing of early warning models, could be conducted less frequently (Headey & Barrett, 2015).

As mentioned in the previous sections, combining different types of data (temporal, spatial) and tools (mapping, monitoring, targeting) may also allow us to overcome some current data limitations. For example, where household-level outcomes are of interest but only course spatial data are available, mapping may allow the researcher to identify regions in which more intensive on-the-ground data gathering is needed, such as vulnerable regions, regions with high heterogeneity, or regions where our models perform poorly. We might also interact data that varies temporally with data that varies spatially to better predict changes over time. Machine learning methods such as transfer learning may also allow the modeler to overcome some data limitations such as sample size constraints, and neural nets may allow the modeler to resolve the problem of combining data at multiple temporal and spatial scales.

Early warning models also suffer from endogeneity of intervention. If prior indicators of food crises caused intervention/response, then the outcomes needed to train such an early warning model may not be available in the data (V. Kshirsagar, personal communication, 2021). Analysts need data not only on food insecurity and its structural correlates but also on intervention and mitigation efforts in order to design and maintain effective early warning systems. Such data would also require coordination among operational agencies and could be vastly facilitated by interagency/intergovernmental platforms.

However, even as we clamor for more data, we also urge caution in data collection, handling, and use. Informed consent and anonymity must be scrupulously maintained. While randomized offsets in geotagged household-level survey data can frustrate targeting, mapping, and M&E efforts, they also serve to protect the populations under study. Similar strategies must be used to protect the geolocations of individuals who share data via CDRs and mobile phone-based surveys. As data continue to grow, especially geotagged data and data from mobile phones and apps, researchers will need to continue to protect the privacy of vulnerable communities in tandem with improving the rapid deployment of services and transfers that can meet their needs.

Finally, one can only predict with accuracy states and processes that have been previously observed in data. This presents a challenge to the development of early warning systems in an era of climate change. Nonstationarity processes can result in a shift into a previously unobserved state. For example, with sea level rise, past data-generating processes will likely perform poorly in predicting future outcomes for coastal communities increasingly faced with regular or permanent flooding. We will struggle to anticipate outcomes in such regions with accuracy no matter how abundant our data and powerful our algorithms. Such data limitations suggest that simulation may be an additional important tool for developers of early warning systems in the future.

CONCLUSION

Any machine learning-based model, map, or tool will only be as good as the data used to train and test it. Consequently, data availability has enormous implications for the questions that can be asked and answered with machine learning applications. Progress is limited by a serious undersupply of the global public good of data collection, standardization, updating, and open access curation of key variables that are typically useful in poverty and malnutrition targeting, mapping, M&E, and forecasting.

The COVID pandemic, the rising challenges of climate change, and related food crises combined with advanced data collection techniques and powerful algorithms all increase the value of rapid assessment while simultaneously increasing the complexity of and care we must take in data use and model building. While we embrace the power and possibility of big data and machine learning, we simultaneously urge thoughtful consideration of the purposes and uses of these models. Most importantly, no data or targeting, mapping, or early warning model will be effective without the political will and financial support to take action and intervene to reduce unnecessary human suffering.

ACKNOWLEDGMENTS

We thank Nathan Jensen, Varun Kshirsagar, and participants and presenters at the AAEA ASSA 2021 Big Data and Near-Real-Time Monitoring of Food Emergencies session for feedback and insights on earlier drafts of this work. Any remaining errors are our own. We thank the United States Agency for International Development for financial support under cooperative agreement # 7200AA18CA00014, "Innovations in Feed the Future Monitoring and Evaluation—Harnessing Big Data and Machine Learning to Feed the Future." The contents are solely the authors' responsibility and do not necessarily reflect the views of USAID or the United States Government.

ENDNOTES

[1] We are abstracting from important definitional challenges throughout this paper: determining a poverty line, local and global nutrition standards, and food security threshold (even measuring food security) are all non-trivial tasks. See Barrett (2010) for an introduction to the challenges of measuring food security, UNHCR (2005) for an introduction to the challenges of defining nutrition standards, and Ravallion (2020) for an introduction to the challenges of measuring poverty.

[2] An important caveat to these observations is that transience in welfare status is typically overestimated in available data due to measurement error and short panel intervals (Naschold & Barrett, 2011). In addition, while the transitory poor and food insecure make up a majority of the poor and food insecure on headcount measures, they are not necessarily the majority in more distributionally sensitive metrics, as the depth of poverty and food insecurity is typically greatest among the most chronically deprived.

[3] One might consider seasonal food insecurity along these same lines: some seasonal food insecurity is structural, some stochastic (V. Kshirsagar, personal communication, 2021). Where parsing structural seasonal food insecurity from stochastic seasonal food security is important for a well-targeted intervention, the concerns we identify here apply.

REFERENCES

Aiken, E., G. Bedoya, A. Coville, and J. E. Blumenstock. 2020. *Targeting Development Aid with Machine Learning and Mobile Phone Data*. Working Paper, University of California, Berkeley.

Altındağ, O., S.D. O'Connell, A. Şaşmaz, Z. Balcıoğlu, P. Cadoni, M. Jerneck, and A. Kunze Foong. 2021. "Targeting Humanitarian Aid Using Administrative Data: Model Design and Validation." *Journal of Development Economics* 148: 102564.

Andree, B. P. J., A. Chamorro, A. Kraay, P. Spencer, and D. Wang. 2020. *Predicting Food Crises*. World Bank Policy Research Working Paper 9412, World Bank.

Baez, J. E., V. Kshirsagar, and E. Skoufias. 2019. *Adaptive Safety Nets for Rural Africa: Drought-Sensitive Targeting with Sparse Data*. World Bank Policy Research Working Paper 9071, World Bank.

Balboni, C., O. Bandiera, R. Burgess, M. Ghatak, and A. Heil. 2020. *Why Do People Stay Poor? Evidence on Poverty Traps from Rural Bangladesh*. LSE Working Paper, London School of Economics.

Barrett, C.B. 2010. "Measuring Food Insecurity." *Science* 327(5967): 825–8.

Barrett, C.B., and M.A. Constas. 2014. "Toward a Theory of Resilience for International Development Applications." *Proceedings of the National Academy of Sciences* 111(40): 14625–30.

Barrett, C.B., P.P. Marenya, J. McPeak, B. Minten, F. Murithi, W. Oluoch-Kosura, F. Place, J.C. Randrianarisoa, J. Rasambainarivo, and J. Wangila. 2006. "Welfare Dynamics in Rural Kenya and Madagascar." *Journal of Development Studies* 42(2): 248–77.

Barrett, C.B., T. Garg, and L. McBride. 2016. "Well-Being Dynamics and Poverty Traps." *Annual Review of Resource Economics* 8: 303–27.

Barrett, C.B., M.R. Carter, and J.P. Chavas. 2019. *The Economics of Poverty Traps*. Chicago, IL: University of Chicago Press.

Barriga Cabanillas, Oscar, Joshua Blumenstock, Daniel Putman, and Travis J. Lybbert. 2021. "Digital Breadcrumbs and Dietary Diversity: Can Mobile Phone Metadata Predict Food Security Outcomes Among Rural Haitian Households?" In *AAEA ASSA 2021 Big Data and Near-Real-Time Monitoring of Food Emergencies session*, Virtual Meeting.

Baulch, B., and J. Hoddinott. 2000. "Economic Mobility and Poverty Dynamics in Developing Countries." *Journal of Development Studies* 36(6): 1–24.

Beegle, K., J. de Weerdt, and S. Dercon. 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics* 93(3): 1010–33.

Blumenstock, J. 2016. "Fighting Poverty with Data." *Science* 353(6301): 753–4.

Blumenstock, J. 2020. "Machine Learning Can Help Get COVID-19 Aid To Those Who Need It Most." *Nature* 581(7807).

Blumenstock, Joshua. 2021. "Using Mobile Phone and Satellite Data to Target Emergency Cash Transfers." *CEGA Blog Post*, accessed January 12, 2021, https://medium.com/center-for-effective-global-action/using-mobile-phone-and-satellite-data-to-target-emergency-cash-transfers-f0651b2c1f3f.

Blumenstock, J., G. Cadamuro, and R. On. 2015. "Predicting Poverty and Wealth from Mobile Phone Metadata." *Science* 350(6264): 1073–6.

Browne, C., D. Matteson, L. McBride, L. Hu, Y. Liu, Y. Sun, J. Wen, and C. B. Barrett. 2021. *Multivariate Random Forest Prediction of Poverty and Malnutrition Prevalence*. Working paper, Cornell University.

Carter, M.R., and C.B. Barrett. 2006. "The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach." *Journal of Development Studies* 42(2): 178–99.

Carter, M.R., and T.J. Lybbert. 2012. "Consumption Versus Asset Smoothing: Testing the Implications of Poverty Trap Theory in Burkina Faso." *Journal of Development Economics* 99(2): 255–64.

Carter, M.R., and J. May. 1999. "Poverty, Livelihood and Class in Rural South Africa." *World Development* 27(1): 1–20.

Cissé, J.D., and C.B. Barrett. 2018. "Estimating Development Resilience: A Conditional Moments-Based Approach." *Journal of Development Economics* 135(C): 272–84.

Coady, D., M. Grosh, and J. Hoddinott. 2004. *Targeting of Transfers in Developing Countries: Review of Lessons and Experience*. Washington, DC: The World Bank.

Cooper, M.W., M.E. Brown, S. Hochrainer-Stigler, G. Pflug, I. McCallum, S. Fritz, J. Silva, and A. Zvoleff. 2019. "Mapping the Effects of Drought on Child Stunting." *Proceedings of the National Academy of Sciences* 116 (35): 17219–24.

Coudouel, A., and T. Bedi, eds. 2007. *More Than a Pretty Picture: Using Poverty Maps to Design Better Policies and Interventions*. Washington, DC: The World Bank.

Daumard, F., S. Champagne, A. Fournier, Y. Goulas, A. Ounis, J.F. Hanocq, and I. Moya. 2010. "A Field Platform for Continuous Measurement of Canopy Fluorescence." *IEEE Transactions on Geoscience and Remote Sensing* 48(9): 3358–68.

Elbers, C., J.O. Lanjouw, and P. Lanjouw. 2003. "Micro-Level Estimation of Poverty and Inequality." *Econometrica* 71(1): 355–64.

Engstrom, R., J. Hersh, and D. Newhouse. 2017. *Poverty from Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being*. Working Paper, World Bank.

Frankenberg, C., J. Fisher, J. Worden, G. Badgley, S. Saatchi, J. Lee, Geoffrey C. Toon, et al. 2011. "New Global Observations of the Terrestrial Carbon Cycle from GOSAT: Patterns of Plant Fluorescence with Gross Primary Productivity." *Geophysical Research Letters* 38(17): 1–6.

Ghosh, M., and J.N.K. Rao. 1994. "Small Area Estimation: An Appraisal." *Statistical Science* 9(1): 55–76.

Grosh, M.E., and J.L. Baker. 1995. *Proxy Means Tests for Targeting Social Programs: Simulations and Speculation*. Washington, DC: World Bank.

Guanter, L., Y. Zhang, M. Jung, J. Joiner, M. Voigt, J.A. Berry, C. Frankenberg, et al. 2014. "Global and Time-Resolved Monitoring of Crop Photosynthesis with Chlorophyll Fluorescence." *Proceedings of the National Academy of Sciences* 111(14): E1327–33.

Head, Andrew, Mélanie Manguin, Nhat Tran, and Joshua Blumenstock. 2017. "Can Human Development Be Measured with Satellite Imagery?" In *Proceedings of ICTD '17, Lahore, Pakistan*, November 16–19, 2017 8: 1–11.

Headey, D., and C.B. Barrett. 2015. "Opinion: Measuring Development Resilience in the World's Poorest Countries." *Proceedings of the National Academy of Sciences* 112(37): 11423–5.

Hersh, J., R. Engstrom, and M. Mann. 2020. "Open Data for Algorithms: Mapping Poverty in Belize Using Open Satellite Derived Features and Machine Learning." *Information Technology for Development* 27(2): 263–92.

Jalan, J., and M. Ravallion. 2002. Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China. *Journal of Applied Econometrics* 17(4): 329–46.

Jean, N., M. Burke, M. Xie, W.M. Davis, D.B. Lobell, and S. Ermon. 2016. "Combining Satellite Imagery and Machine Learning to Predict Poverty." *Science* 353(6301): 790–4.

Knippenberg, E., N. Jensen, and M. Constas. 2019. "Quantifying Household Resilience with High Frequency Data: Temporal Dynamics and Methodological Options." *World Development* 121: 1–15.

Knippenberg, E., H. Michelson, E. Lentz, and T. Lallemant. 2020. *Early Warning; Better Data, Better Algorithms*. Working Paper.

Kshirsagar, V., J. Wieczorek, S. Ramanathan, and R. Wells. 2017. *Household Poverty Classification in Data-Scarce Environments: A Machine Learning Approach*. Working Paper, arXiv:1711.06813 arXiv preprint.

Lentz, E.C., H. Michelson, K. Baylis, and Y. Zhou. 2019. "A Data-Driven Approach Improves Food Insecurity Crisis Prediction." *World Development* 122: 399–409.

Masaki, T., D. Newhouse, A. R. Silwal, A. Bedada, and R. Engstrom. 2020. *Small Area Estimation of Non-Monetary Poverty with Geospatial Data, Background Paper for World Development 2021 Report*. World Bank Policy Research Working Paper No. 9383, World Bank.

McBride, L., and A. Nichols. 2018. "Retooling Poverty Targeting Using Out-of-Sample Validation and Machine Learning." *World Bank Economic Review* 32(3): 531–50.

Mohammed, G.H., R. Colombo, E.M. Middleton, U. Rascher, C. van der Tol, L. Nedbal, Y. Goulas, et al. 2019. "Remote Sensing of Solar-Induced Chlorophyll Fluorescence (SIF) in Vegetation: 50 Years of Progress." *Remote Sensing of Environment* 231: 111177.

Mude, A. G., C. B. Barrett, J. G. McPeak, R. Kaitho, and P. Kristjanson. 2009. Empirical Forecasting of Slow-Onset Disasters for Improved Emergency Response: An Application to Kenya's Arid North. *Food Policy* 34(4): 329–39.

Naschold, F., and C.B. Barrett. 2011. "Do Short-Term Observed Income Changes Overstate Structural Economic Mobility?" *Oxford Bulletin of Economics and Statistics* 73(5): 705–17.

Njuguna, C., and P. McSharry. 2017. "Constructing Spatiotemporal Poverty Indices from Big Data." *Journal of Business Research* 70: 318–27.

Noor, A.M., V.A. Alegana, P.W. Gething, A.J. Tatem, and R.W. Snow. 2008. "Using Remotely Sensed Night-time Light as a Proxy for Poverty in Africa." *Population Health Metrics* 6(1): 5.

Pokhriyal, N., and D.C. Jacques. 2017. "Combining Disparate Data Sources for Improved Poverty Prediction and Mapping." *Proceedings of the National Academy of Sciences* 114(46): E9783–92.

Pritchett, L., and F. Hani. 2020. "The Economics of International Wage Differentials and Migration." In *Oxford Research Encyclopedia of Economics and Finance*. Oxford: OUP.

Ravallion, M., and Q. Wodon. 1999. "Poor Areas, Or Only Poor People?" *Journal of Regional Science* 39(4): 689–711.

Ravallion, M. 2020. "On Measuring Global Poverty." *Annual Review of Economics* 12: 167–88.

Rolf, Esther, Jonathan Proctor, Tamma Carleton, Ian Bolliger, Vaishaal Shankar, Miyabi Ishihara, Benjamin Recht, and Solomon Hsiang. 2020. *A Generalizable and Accessible Approach to Machine Learning with Global Satellite Imagery*. NBER Working Paper 28045.

Schreiner, M. 2007. "A Simple Poverty Scorecard for the Philippines." *Philippine Journal of Development* 34(2): 43–70.

Song, L., L. Guanter, K. Guan, L. You, A. Huete, W. Ju, and Y. Zhang. 2018. "Satellite Sun-Induced Chlorophyll Fluorescence Detects Early Response of Winter Wheat to Heat Stress in the Indian Indo-Gangetic Plains." *Global Change Biology* 24(9): 4023–37.

Steele, J.E., P.R. Sundsøy, C. Pezzulo, V.A. Alegana, T.J. Bird, J. Blumenstock, J. Bjelland, et al. 2017. "Mapping Poverty Using Mobile Phone and Satellite Data." *Journal of the Royal Society Interface* 14(127): 20160690.

Sun, Y., C. Frankenberg, J.D. Wood, D.S. Schimel, M. Jung, L. Guanter, D.T. Drewry, et al. 2017. "OCO-2 Advances Photosynthesis Observation from Space via Solar-Induced Chlorophyll Fluorescence." *Science* 358 (6360): eaam5747.

Tang, Binh, Yanyan Liu, and David Matteson. 2021. "Dynamic Poverty Prediction with Vegetation Index." In *AAEA ASSA 2021 Big Data and Near-Real-Time Monitoring of Food Emergencies Session*, Virtual Meeting.

UNHCR. 2005. *A Manual: Measuring and Interpreting Malnutrition and Mortality*. UNHCR. Accessed January 30, 2021. www.unhcr.org/45f6abc92.pdf.

de Weerdt, J. 2010. "Moving Out of Poverty in Tanzania: Evidence from Kagera." *Journal of Development Studies* 46(2): 331–49.

Wen, J., P. Köhler, G. Duveiller, N.C. Parazoo, T.S. Magney, G. Hooker, L. Yu, C.Y. Chang, and Y. Sun. 2020. "A Framework for Harmonizing Multiple Satellite Instruments to Generate a Long-Term Global High Spatial-Resolution Solar-Induced Chlorophyll Fluorescence (SIF)." *Remote Sensing of Environment* 239: 111644.

Yeh, C., A. Perez, A. Driscoll, G. Azzari, Z. Tang, D. Lobell, S. Ermon, and M. Burke. 2020. "Using Publicly Available Satellite Imagery and Deep Learning to Understand Economic Well-Being in Africa." *Nature Communications* 11(1): 1–11.

Zhou, Y., K. Baylis, E. Lentz, and H. Michelson. 2021. "Predicting Food Security in Sub-Saharan Africa with Machine Learning." In *AAEA ASSA 2021 Big Data and Near-Real-Time Monitoring of Food Emergencies Session*, Virtual Meeting.

How to cite this article: McBride, Linden, Christopher B. Barrett, Christopher Browne, Leiqiu Hu, Yanyan Liu, David S. Matteson, Ying Sun, Jiaming Wen. 2021. "Predicting poverty and malnutrition for targeting, mapping, monitoring, and early warning." *Applied Economic Perspectives and Policy* 1–14. https://doi.org/10.1002/aepp.13175