Open-Source Models for Famine Prediction

Methodology and Technical Notes

Andrew J. Peterson, Ph.D.

May 15, 2025

We provide a basic codebase for creating early warnings of famine from data on food availability and prices. In contrast to models for predicting agricultural yield based on satellite imagery, these tools are meant to work in the context of rapidly-evolving situations, such as war affected regions where parties to the conflict may cut off access to food aid. We provide methods for resource-scarcity based models and price-models, along with utilities for conducting sensitivity analyses and diagnostic plots. We also provide Jupyter notebooks to illustrate model usage, alongside detailed code comments and technical documentation. Having found no similar repositories on GitHub, we aim for this resource to inform and enhance public discussions of famine prediction and response – an endeavor that, as current crises in Sudan, Gaza, and elsewhere unfortunately illustrate, remains urgently needed.

Code available on Github here, and archived at Zenodo (DOI: 10.5281/zenodo.15366708).

Assistant Professor, University of Poitiers
andrew.peterson (at) univ-poitiers (dot) fr

1. Project Overview

This project aims to bring together diverse methods for data-driven famine early warning analysis in a coherent, documented Python package. While there are a number of sophisticated models for analyzing satellite and crop data, the focus here is instead basic analysis of price and food availability calculations, with a focus on tools that can be used for prediction in low-information and rapidly-evolving contexts.

While crop yield estimates are obviously essential, we believe it is also important to develop methods for understanding the evolution of famine conditions in a situation in which violence and the efforts of warring parties to close off humanitarian aid corridors, for example, can quickly generate a crisis for affected populations.

The ongoing famine in Sudan, and in particular the reports by the Clingendael Institute (Hoffmann 2024; Gaasbeek 2024), were a primary inspiration for this work. Their predictions of thousands or perhaps even millions of deaths from famine requires greater public attention and transparency about the assumptions underlying such models. Note that while the resource scarcity model tries to follow their methods as close as possible as described in the reports, at times we introduce what we believe to be reasonable approaches in order to create a programmatic approach. For example, we introduce simple functional forms for describing the distribution of grain across the population in the place of expert-derived estimates, which allows for simulation analysis (See Section 4.4).

Currently these situations are analyzed by specialist experts who have knowledge of local conditions (e.g. IPC Partners (2018) and Famine Early Warning Systems Network (FEWS NET) (1985).) Given the assumptions necessary for the models below, that is a natural approach. However, we believe that open-source models, so long as their potential shortcomings or limitations are acknowledged, could provide an important complement to hand-curated expert analyses. This may be especially true in situations in which in-country experts are prevented from working for political or resource reasons.[1] For a certain segment of the population, open data and modeling may provide greater understanding in a way reliance on experts cannot. The ideal is likely a hybrid approach, in which fully-open source models are developed and improved iteratively in connection with in-country experts and curation of more public data.

[1] For example, the U.S. cut the budget for FEWS NET and in Sudan the government is no longer cooperating with the IPC (Spitzer 2025).

With these thoughts in mind, we focus on two types of models:

a. Price-based models: Track abrupt changes in food prices as an early indicator of scarcity and access issues.

b. Resource scarcity models: Estimate the total available (and potentially accessible) food and compare it to population requirements over a given period.

Again, thinking of Sudan, and the importance of the already-grave and potentially catastrophic impact of this famine, we believe it is essential to draw more awareness and public discussion of these models, and hope that a transparent repository may contribute to anyone who wants to understand what underlies such models. In addition to this document, we provide the code, some sample data for Sudan, and Jupyter notebooks that walk through the basic calculations, so users can also explore alternative data or parameters, or modify the code themselves (See Github repository here). Finally, in addition to the above models, we also provide code for sensitivity analyses and visualizations of model outputs.

Disclaimer: This work is intended as an initial contribution to ongoing discussions rather than as a definitive resource. While I have discussed these matters with others, to whom I'm grateful for their help, I am not an expert in the field, and remain solely responsible for any errors. Please feel free to reach out with questions, comments or concerns, and the repository and documentation will be updated accordingly.

2. Related Work

Famines have increasingly been understood as socio-economic disasters rather than purely natural shortages. A foundational theory is Amartya Sen's entitlement approach, which posits that famines result from failures in people's *exchange entitlements* – their ability to command food through work, trade, or assets – rather than from an absolute lack of food (Sen 1981). In other words, a person or group can starve amid plenty if they lose access to food due to personal, structural or economic problems. Sen's analysis of the Bengal famine of 1943, for example, showed that food production was relatively adequate, but inflation and unequal purchasing power meant the poorest could not afford grain, leading to mass starvation, while others emphasize political failures.

Economic and political factors such as poverty, market collapse, or conflict are often central to famine dynamics. In Darfur, Sudan, the mid-1980s famine followed drought but was fundamentally

driven by conflict, displacement, and the collapse of local livelihoods rather than nationwide food unavailability (de Waal 2005). Political instability and war disrupted markets and eroded communities' entitlements to food. Such evidence shifted famine analysis away from simple food supply explanations towards a political economy perspective: famines are "man-made" in the sense that human institutions and policies determine *who* can access food during crises.

Famine prediction has long been a challenge for policymakers and researchers. Early studies in the 1980s began to identify key indicators and methodologies for anticipating famines. For example, analysts noted that market conditions and food prices often provide early warning signals of impending food crises Seaman and Holt (1980); Cutler (1984). There were also early calls for proactive famine prevention through predictive modeling Mellor (1986). Despite these warning signs, devastating famines continued, and vulnerabilities and response failures continue to occur (Devereux 2000), underscoring the need for more effective models and institutions.

2.1. Institutional Early Warning Systems

In response to recurring food crises, institutionalized early warning systems were developed to improve famine prediction and promote timely interventions. One prominent example is the Famine Early Warning Systems Network (FEWS NET) (1985) established in 1985, which integrates diverse data sources (e.g., climate indicators, crop yields, market prices, and nutrition surveys) to project potential food security crises (Ross et al. 2009), which has however recently suffered from a resource shortfall. Over time, analytical frameworks and data inputs for such systems have expanded. Advances in technology have enabled the incorporation of satellite remote sensing data (rainfall, vegetation indices, etc.) to detect early signs of drought and crop failure (Brown 2010). Analytical tools have also evolved, such as the 'Food Economy Approach', that applies exchange entitlements theory to assess risk levels within different livelihood systems through risk-mapping Boudreau (1998); Seaman (2000). These developments have made modern early warning systems more data-driven and comprehensive.

Despite these improvements, accurately predicting famines remains complex. Food security crises often result from a confluence of factors – including weather shocks, conflict, political and economic disruptions – that challenge simple modeling approaches Misselhorn (2005). And of course early warning must be coupled with effective response. The 2011 Somalia famine tragically illustrated that even when warnings are available, a slow or insufficient humanitarian response can still lead to catastrophe Maxwell and Fitzpatrick (2012). In that case, early warning analysts had flagged the deteriorating conditions, but aid was mobilized too late, resulting in tens of thousands

of deaths Hillbruner and Moloney (2012). This experience underscored that predictive systems are only as valuable as the actions they trigger: timely, coordinated interventions are essential to translate warnings into famine prevention.

2.2. Advances in Modeling Techniques

Recent years have seen growing use of advanced statistical and machine learning techniques to enhance famine prediction models. Researchers are leveraging richer datasets and more powerful algorithms to improve the accuracy and lead time of famine forecasts. One line of work applies supervised machine learning to historical food security data, such as the work of Okori and Obua (2011) using household survey data in Uganda. Mwebaze, Okori, and Quinn (2010) explore a causal modeling approach, learning probabilistic relationships among famine risk factors (rainfall, crop yields, market prices, etc.) to improve early warning capabilities.

Variables ranging from remote-sensed drought indices to conflict intensity and food price trends may be combined within a single predictive framework, and use machine learning methods that allow for non-linear relationships in high dimensional spaces. Significant challenges remain, however. True famines are relatively rare events, which means limited data for model training and an imbalanced prediction problem. Furthermore, the interpretability of model outputs is crucial – stakeholders need to trust and understand predictions in order to act on them. Ongoing research and initiatives (for example, the World Bank's Famine Action Mechanism) aim to integrate these advanced predictive techniques into operational use. With this in mind, we hope to include more approaches to modeling in future versions of our work, such as work on incorporating information on conflict, resource networks, and mobility (Maxwell 2019). and adapting to specific circumstances relevant to individual cases (Howe 2018).

3. Price Anomaly Detection Models

Food prices are an intuitive early-warning indicator for famine conditions, and are often available even from simple anecdotal evidence of individuals who observe the change of prices in their community. For populations with relatively fixed budgets for whom food makes up a sizeable part of their expenses, a rapid increase in prices can make reducing food consumption the only option. Finally, in contrast to estimates of crop productivity, these prices can act as signals that include many difficult-to-observe conditions about the local availability of stocks or alternatives.[2]

[2]Conversely, if individuals are able to buy up significant parts of available reserves, they could create artificial shortages to raise prices.

Thus, with a few caveats, these prices can be useful to help detect stress before hunger becomes widespread. A common approach is to identify price anomalies – unusual surges in food prices relative to expected seasonal patterns. For example, the FAO's Global Information and Early Warning System (GIEWS) is based on an indicator of price anomalies that relies on a weighted compound growth rate of prices to identify abnormal increases (Baquedano 2015). By comparing current monthly price growth against historical trends (accounting for seasonal fluctuations and inflation), this indicator signals when prices are rising quickly. Such compound growth models are relatively simple, yet effective in diverse markets, serving as an early warning tool even when other data such as harvest estimates are unavailable.[3]

3.1. Baquedano Price Anomaly Model

We provide code to implement the price anomaly detection model based on an approach initially proposed by Araujo, Araujo-Bonjean, and Brunelin (2012) and improved by Baquedano (2015). The model accounts for volatility, non-stationarity of prices, and seasonal effects.

Compound Growth Rate (CGR)

Following Baquedano (2015), we characterize monthly price movements using the compound growth rate (CGR) over a rolling window of length w months. Let P_t denote the real (inflation-adjusted) price at month t, and P_{t+w-1} the price at the end of that w-month window. Then

$$\mathrm{CGR}_t = \left(\frac{P_{t+w-1}}{P_t}\right)^{\frac{1}{(w-1)}} - 1.$$

Because there are $(w-1)$ intervals between t and $(t+w-1)$ in monthly data, the exponent is $1/(w-1)$, consistent with typical compound growth definitions. In the source code, this is computed by the function `compute_cgr(series, window)`. This helps detect sustained price changes while mitigating short-term fluctuations.

Volatility Adjustment

Price time series in developing regions frequently exhibit high volatility. To lessen false positives from 'normal' short-term swings, we define volatility as the standard deviation of log differences

[3] Recent research has further enhanced price-based warning models such as through Bayesian spatio-temporal models, to incorporate price information across multiple markets and over time, treating neighboring markets' prices as correlated. These approaches help distinguish broad inflationary trends from localized spikes (Wang et al. 2020).

within the same effective window. Specifically, let:

$$\sigma_{\Delta \log P}(w) \;=\; \mathrm{std}\Big(\log P_{t+1} - \log P_t\Big) \quad \text{over } (w-1) \text{ monthly steps.}$$

In code, the function `compute_volatility` calculates this rolling standard deviation of log price differences over $(w-1)$ intervals.

Then for CGR, to reduce the impact of highly volatile price swings, Baquedano (2015) suggests deflating the CGR by $\big[1 - \sigma_{\Delta \log P}(w)\big]$. Hence the volatility-adjusted CGR (vCGR) is:

$$\nu\mathrm{CGR}_t \;=\; \mathrm{CGR}_t \;\times\; \Big[1 - \sigma_{\Delta \log P}(w)\Big].$$

In our code, this is handled by `adjust_cgr_for_volatility(cgr_series, vol_series, clip=False)`, where `clip=True` optionally caps $\sigma_{\Delta \log P}$ at 1.0 so that $(1-\sigma)$ does not become negative.

Weighted Estimation of Mean and Standard Deviation

Once the volatility-adjusted CGR is calculated, we can compare each month's νCGR against historically typical values for that same month. Because older data might be less relevant, Baquedano (2015) incorporates weights. Let there be γ total observations (years) for a given month, indexed by y. Denote $\nu\mathrm{CGR}_y$ the volatility-adjusted CGR in year y for that month, and w_y the corresponding weight. Equation (4) from Baquedano (2015) gives a weighted mean:

$$\overline{\nu\mathrm{CGR}}_W \;=\; \frac{\sum_{y=1}^{\gamma} w_y \, \nu\mathrm{CGR}_y}{\sum_{y=1}^{\gamma} w_y}.$$

Equation (5) specifies the weighted standard deviation:

$$\hat{\sigma}_{\nu\mathrm{CGR},W} \;=\; \sqrt{\frac{\sum_{y=1}^{\gamma} w_y \left(\nu\mathrm{CGR}_y - \overline{\nu\mathrm{CGR}}_W\right)^2}{\sum_{y=1}^{\gamma} w_y \;\times\; \frac{(\gamma-1)}{\gamma}}}.$$

In our code, these are computed via `weighted_mean(values, weights)` and `weighted_std(values, weights)`.

Monthly Anomaly Score (IPA_score)

Having a historical reference for each calendar month, Baquedano (2015) defines an anomaly score for a current month t:

$$\text{IPA\_score}_t = \frac{\nu\text{CGR}_t - \overline{\nu\text{CGR}}_W}{\hat{\sigma}_{\nu\text{CGR},W}}.$$

If $\hat{\sigma}_{\nu\text{CGR},W} = 0$, or if any relevant data are missing, the score is undefined (`NaN`). Our code uses `compute_anomaly_score(vcgr_value, weighted_mean, weighted_std)`, and the function `classify_anomaly` assigns thresholds:

- **Price Alert**: if IPA_score ≥ 1.
- **Price Watch**: if $0.5 \leq$ IPA_score < 1.
- **Normal**: otherwise.

Quarterly vs. Annual Signals and Combination

Because price anomalies may occur over short horizons (within a season) or reflect persistent long-run shifts, we typically compute two forms of vCGR: a quarterly window (3 months) and an annual window (12 months), with corresponding anomaly scores, $\text{IPA\_score}_{3m,t}$ and $\text{IPA\_score}_{12m,t}$. Baquedano (2015) then combines them via a convex combination with weight γ:

$$\text{IPA}_t = \gamma \times \text{IPA\_score}_{3m,t} + (1-\gamma) \times \text{IPA\_score}_{12m,t}.$$

We determine γ by principal component analysis (PCA) on the covariance matrix of $\left(\nu\text{CQR}_t,\ \nu\text{CAG}_t\right)$, where νCQR_t is the 3-month vCGR series and νCAG_t is the 12-month vCGR series. Specifically, γ is set to the fraction of total variance explained by the first principal component. If insufficient data exist (e.g. no variability), the code defaults to $\gamma = 0.5$.

Dealing with Non-stationary Price Series

Whereas some prior studies (Araujo, Araujo-Bonjean, and Brunelin 2012) relied on fitting a trend to P_t directly, the CGR-based approach can be more robust when the price level itself is non-stationary. Our code includes optional stationarity tests (Augmented Dickey–Fuller, KPSS, etc.), but following Baquedano (2015), we do not require strict stationarity in levels, since we focus on growth rates.

Implementation notes and handling missing data

Before computing CGRs, we typically adjust for missing data via the function: `handle_missing_data(series, method='interpolate', ...)`, which allows for linear

interpolation, forward fill, backward fill, or dropping missing dates. For high-inflation environments, we also recommend adjusting P_t by an external CPI or inflation index.

Once the data are prepared, the key steps for anomaly detection are:

a. **Data preprocessing**: Fill or drop missing values; adjust for inflation if needed.
b. **Compute CGRs**: Obtain rolling 3-month (quarterly) and 12-month (annual) CGRs.
c. **Calculate volatility**: Using rolling std of log differences over $(w-1)$ steps.
d. **Volatility-adjust the CGR**: νCGR = CGR $\times$ $[1 - \sigma_{\Delta \log P}]$ (with optional clipping).
e. **Monthly historical statistics**: For each month of the year, compute weighted mean and weighted std dev of νCGR.
f. **Anomaly scores**: Standardize the current νCGR_t using the historical mean and std dev for that month.
g. **Combine signals**: Merge the quarterly and annual signals with weight γ by PCA.
h. **Classify anomalies**: Label each month as Normal, Watch, or Alert.

Discussion. For a concrete example on Sudan wheat-price data, see the companion notebook `price_model_walkthrough.ipynb`, which demonstrates these steps in detail.

This CGR-based model aims to serve as a broad screening tool for abnormal price movements. Improvements can likely be achieved in specific contexts by tuning the weighting scheme, the interpolation or data-cleaning strategies, and the alert thresholds to local market conditions.

4. Resource Scarcity Model

4.1. Overview

The resource scarcity model is based on an intuitive if stark calculation. For a given country or region, it attempts to calculate the calories available from different sources of food and compare this to caloric needs of the population. This can help identify where and when shortages become critical, particularly during conflict or other disruptions.

While the model depends on a number of parameters which are not fully observable in some contexts, such as the availability of reserve stocks, alternative food sources, migration rates, etc., the estimates can be refined over time and provide a tool for analyzing situations that are uncommon in many societies, such as the disruption of food transportation due to conflict or aid embargoes.

The current implementation centers on caloric availability and basic anthropometric (BMI) impacts. However, real famine conditions often intersect with broader health crises (e.g., epidemics)

and social barriers (e.g., unequal distribution, conflict-driven displacement). While the model estimates how calorie deficits evolve and affect mortality over time, it does not fully address related factors such as healthcare infrastructure or socio-political constraints. Despite such limitations, the model offers a flexible basis for rapid assessments, scenario testing, and integration with other modules (e.g., pricing or disease models). Users should therefore regard it as a first-pass framework, to be refined with situational data and expert judgment.

4.2. Estimating total available calories

We begin by providing a step-by-step overview of how to estimate the total calories available in a given region over a specified period of time. In particular, we focus on grains (e.g., wheat, sorghum, and millet) and then scale to account for non-grain foodstuffs.

We then use this to estimate whether enough food exists to meet the caloric needs of a population. The basic steps are:

a. **Compile production data.** Identify the locally produced amounts of each relevant cereal (e.g., wheat, sorghum, millet) in the specified time interval, typically measured in metric tons (MT).
b. **Incorporate additional grain sources.** Add estimates for other major inputs, such as imports, food aid, wild-harvested grains, or leftover household stocks, to the production totals. In practice, each added source may be disaggregated across major categories such as wheat, rice, sorghum, etc.
c. **Adjust for unconsumed remainders.** Because not all food stocks are necessarily depleted by the end of the modeling period, we subtract a small fraction to represent carry-over stocks remaining after consumption, seed-grains, losses due to spoilage, etc.
d. **Convert grain quantities to caloric values.**
    i. *Moisture content assumption.* Convert wet weight to dry matter by applying a moisture content factor (often 10–15%, depending on local conditions (Balkrishna and Visvanathan 2019; Abdollahpour, Kosari-Moghaddam, and Bannayan 2020; Sadaka et al. 2015).
    ii. *Caloric density.* Multiply the dry matter by the kilocalories per kilogram of each grain (e.g., approximately 3000–3400 kcal/kg for wheat dry matter, see Jocelyne, Béhiblo, and Ernest (2020)). This yields total grain-derived kilocalories.
e. **Estimate non-grain sources.** Many populations obtain a substantial share of total calories from cereals, but not 100%. If estimates of these non-grain sources are not available, they may be estimated by assuming a fixed proportion of grain sources.

Finally, in situations where the data may be particularly unreliable, it may be worth reviewing the analysis with a 'sanity check', such as by assessing how many days of food the computed total would cover for the relevant population, given an assumed daily caloric intake of 2000–2200 kcal

per person.

Implementation notes

An `est_kcal_walkthrough.ipynb` Jupyter notebook is included in the repository, illustrating how to implement this method in Python. This assumes access to the following data as input:

- **Cereal Production.** Quantities of harvested wheat, sorghum, millet, or other staple grains (MT).
- **Additional Stocks.** Imports, food aid, household stocks, and informal trade volumes.
- **Moisture Content.** An estimate of post-harvest moisture content for each grain.
- **Caloric Densities.** Average kilocalories per kg (dry) of each grain type.
- **Dietary Composition.** Fraction of total calories assumed to come from cereals versus other foods (often reported by agencies such as FEWS NET).
- **Population.** A best estimate of the total population, alongside average daily per-capita caloric needs.

  The final output of this method is:

$$\text{Total Kcal} = \left(\sum_{i \in \{\text{grains}\}} (\text{MT}_i \times 1000 \times (1 - m_i) \times \text{kcal/kg dry}_i) \right) \times \frac{1}{\text{grain share}},$$

where:

- MT_i are the metric tons of each grain i,
- m_i is the fractional moisture content of grain i,
- kcal/kg dry_i is the kilocalorie density of the dry matter for grain i,
- grain share is the fraction of total caloric intake assumed to come from cereals (e.g., 0.70).

4.3. Energy Requirements and Intake

The model estimates month-on-month changes in food consumption and caloric requirements across percentiles of the population. The scope here is limited to capturing the physiological effects of caloric deficits using body-mass index (BMI) as a central state variable. We reference and adapt elements from (Gaasbeek 2024; Henry 2001). For each subgroup (or percentile) of the population, the model tracks:

- The fraction of energy requirements met (via total dietary intake).
- The resulting change in BMI due to sustained deficits or surpluses.
- The excess mortality associated with low BMI, applied to the population of that subgroup.

This process is then iterated over many months to project the potential evolution of malnutrition

and mortality across a population.

The model first distinguishes among three BMI ranges, each associated with a different baseline energy requirement per day:

- BMI $< 15\,\mathrm{kg/m^2}$: 1900 kcal/day
- $15 \leq$ BMI $< 18.5\,\mathrm{kg/m^2}$: 2100 kcal/day
- BMI $\geq 18.5\,\mathrm{kg/m^2}$: 2200 kcal/day

Although these thresholds and values are borrowed from prior famine assessment guidelines, they should be interpreted cautiously, since different activity levels, cultural norms, and regional biological factors could change overall requirements by large margins (Henry 2001; Joint FAO/WHO/UNU Expert Consultation on Human Energy Requirements 2004).

In each monthly step, the individual's actual energy intake (in kcal/day) is compared to the relevant requirement for their current BMI category. The fraction by which intake falls short of requirement defines the *energy deficit*:

$$\text{Deficit} = \frac{\text{Required\_kcal} - \text{Actual\_kcal}}{\text{Required\_kcal}}, \tag{1}$$

Because BMI reflects body mass relative to height (BMI $= m/\ell^2$), monthly weight change in response to energy deficits can be projected, following Gaasbeek (2024) as:

$$\Delta\text{BMI} = \alpha \cdot (\text{Deficit}) \; - \; \beta \left(22 - \text{BMI}_{\text{prev}}\right), \tag{2}$$

where α and β are constants fitted to match the general shape of BMI decline or recovery patterns discussed in (Henry 2001). In the default implementation, α can be around 3.41 for deficit scenarios and up to 3.52 when modeling BMI gains after a surplus (the *recovery* scenario). The term $22 - \text{BMI}_{\text{prev}}$ loosely represents a homeostatic effect (i.e., lower BMI leads the body to conserve energy, reducing the net loss). The model subtracts the total ΔBMI from the individual's previous BMI:

$$\text{BMI}_{\text{new}} = \text{BMI}_{\text{prev}} - \Delta\text{BMI}.$$

In practice, both α and β can be fine-tuned or replaced by more physiologically detailed rules (Henry 2001).

Persistent low BMI is associated with substantial increases in mortality risk (Gaasbeek 2024; Aune et al. 2016). The model applies an exponential function for monthly excess mortality, conditioned on the previous month's BMI:

$$\text{Mortality}(\text{BMI}_{t-1}) = 0, \quad \text{if } \text{BMI}_{t-1} \geq 18.5, \tag{3}$$
$$\text{Mortality}(\text{BMI}_{t-1}) = \delta \ \exp\Big((18.5 - \text{BMI}_{t-1})^{\gamma}\Big), \quad \text{if } \text{BMI}_{t-1} < 18.5.$$

Typical defaults in the code are $\delta \approx 0.00028$ and $\gamma \approx 1.33$. This approach, while simplistic, can reflect an individual-level probability of dying from complications related to severe malnourishment (e.g., increased susceptibility to infectious disease). However, it should be supplemented or refined if disease prevalence or other factors are known to be significant.

BMI is assumed homogeneous within each population percentile or subgroup; real-world differences by age, sex, health status, or occupation are neglected. Moreover, the baseline energy requirement used here (e.g., 2100–2200 kcal/day) may underestimate the needs of individuals who are sick or performing strenuous physical labor (Gaasbeek 2024).

4.4. The full model

The above elements are combined into the the month-by-month simulation of how each population subgroup (percentile) is assigned a particular calorie intake, and the resulting change in health status) under varying levels of overall food availability, thereby providing a projected trajectory of famine severity.

We divide the population into 100 equally sized percentiles. Each percentile begins the simulation with a specified body-mass index (BMI). The code allows for multiple ways of specifying this initial BMI distribution—for instance, a linear gradient from BMI 30 down to BMI 18, or a logarithmic profile that produces a heavier concentration of low-BMI individuals in the poorest percentiles. The first step each month is then to assign a share of the total available calories to each percentile.

It is crucial to model not just average caloric intake but also how consumption is spread across socioeconomic classes. Rather than hand-specifying specific assumptions based on local knowledge, we introduce two simple distributions:

- **Linear distribution:** A simple linear function of the percentile index, with parameter *beta1* controlling the slope of increasing caloric intake toward wealthier percentiles.
- **Piecewise-linear distribution:** A linear increase in intake for lower percentiles, transitioning to a constant (flat) caloric level for higher percentiles (e.g., those wealthy enough to meet their basic needs). This method is controlled by *beta1* (the slope) and a cutoff percentile *c* above which the intake plateaus.

Which part of these distributions is the most crucial depends on the specific context. In many situations, high inequality leads a minority to have inadequate access to food, in which case the part that must be most carefully modeled is the lower percentiles. Regardless of the method, the distribution function is constrained to ensure that the total calories assigned across all percentiles equals the fixed monthly budget of calories. The model includes configurable lower and upper calorie bounds (`kcal_min` and `kcal_max`, respectively) for each percentile.

The primary logic is represented in the code in the `ResourceScarcityModel` class. It loops over each month, performing a sequence of updates that reflect how food availability, BMI, and mortality co-evolve:

a. Update grain stock: Each month, new inputs (e.g., from harvests or aid) are added to the existing grain stock. This total stock is then the maximum that can be consumed over that month.
b. Determine monthly caloric supply: A user-defined monthly demand or consumption schedule is used to decide how many metric tons of grain will be consumed in that period. That consumption is converted to calories via a factor (e.g., 3,500,000 kcal per metric ton). In reality, anticipated future harvests or price signals may guide consumption; here the code allows the user to supply any monthly time series.
c. Distribute calories across population percentiles Once the total caloric consumption is established, the model applies one of the aforementioned distribution functions. The result is an array of per-person daily intake for each percentile group.
d. Update BMI and compute excess mortality Each percentile's new caloric intake is compared against the energy requirement (conditional on current BMI). The difference translates into an estimated monthly energy deficit (or surplus). In turn, this deficit modifies BMI using one of two update formulas, depending on whether a surplus-based *recovery* scenario is enabled. The model then calculates the expected mortality for each percentile based on the aforementioned formula.
e. Update population figures. Besides famine-related mortality, each month can feature additional changes: births, deaths from non-famine causes, and net migration. In the model, these effects are read from a data file and spread evenly across all "alive" percentiles. This ensures that no single group (other than those already at critical BMI) is disproportionately affected by routine demographic changes.

As the simulation proceeds, intermediate and final values are recorded. Two primary outputs are generated:

- *Monthly Aggregates.* The quantity of available grain, consumption, updated stock, number of deaths (both from immediate BMI cutoffs and from the excess mortality curve), and the residual total population are stored in a `DataFrame`.
- *Per-Percentile Series.* For diagnostic and analytic purposes, the model also stores at each timestep

the per-percentile BMI, assigned calories, deaths, and remaining population. This data can be used to visualize how low-BMI groups experience faster depletion or elevated mortality as the simulation advances.

4.5. Visualization

Clear presentation of simulation outputs is an essential step for analyzing results and diagnosing the predictions of the model. We provide simple plotting utilities to generate heatmaps from per-percentile data for each time period (e.g., month). By default, the script uses `matplotlib` and `seaborn` to create color-coded tables that reveal spatial (percentile) and temporal (month) patterns in key model variables such as BMI, calorie intake, or mortality:

- **BMI:** Reds indicate severely low BMIs, blending toward orange or yellow for moderate values, then green for healthy BMI ranges (`cmap_bmi`). See for example Figure 1
- **Calorie:** Displays red below approximately 600–700 kcal/day and turns green once values exceed typical minimum intake goals (`cmap_kcal`).
- **Mortality:** Shows low mortality rates in green, moderate in yellow, and high rates in red (`cmap_mortality`).

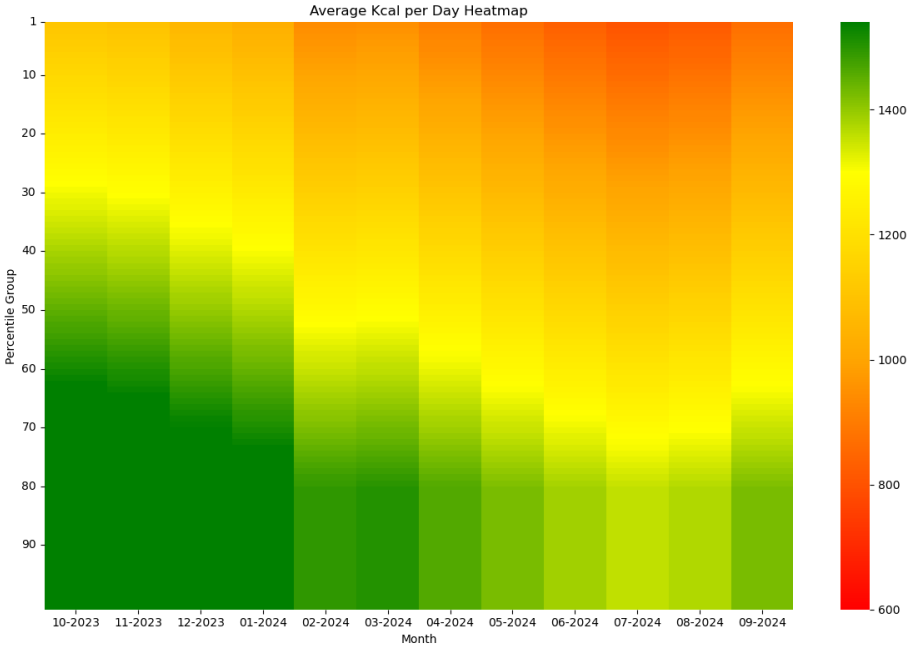

FIGURE 1. Heatmap of caloric values across population percentiles over time

4.6. Discussion and model scope

This approach is intended as a transparent, modular, tool to be calibrated to specific contexts. Users are advised to combine these estimates with local survey results and expert judgment and to account for economic, social, and political factors not captured. Although the model focuses on nutrition and weight loss, famine conditions often worsen morbidity and mortality via outbreaks of communicable diseases, displacement of communities, breakdown of healthcare systems, etc. which can have interactive effects with famine. Furthermore, conflict can constrain people's ability to search for alternative food sources (Gaasbeek 2024).

The number of inputs and assumptions needed to get the model off the ground can seem daunting. That said, sensitivity analysis (see Section 5) can help identify how uncertainty in the inputs or parameters affect the model's predictions. This may in some cases may reassure users that the basic predictions of the model are correct even if the model is incorrectly specified on certain inputs. Alternatively, high dependence on a specific input would provide reason for a user to expend additional effort to achieve greater certainty and precision on that factor.

We review a few caveats in particular:

- **Distribution and access.** Even if total calories appear sufficient, local famines may still occur due to inequality, breakdowns in distribution, or conflict. We specify simple functions for distribution, but more nuanced functions might be required to capture real-world complexities in disease environment, social safety nets, and humanitarian interventions.
- **Population heterogeneity and physiological dynamics.** Different subpopulations have varied energy needs (e.g., children, pregnant women, labor-intensive occupations). A single daily per-capita caloric requirement is a rough average, and the parameters are based in some cases on studies of individuals in situations very different from those of famine populations.
- **Food source uncertainties** Our model focuses on grain and assumes a fixed fraction of non-grain sources. The fraction of total calories from cereals may vary across different regions or time periods. Furthermore, populations may have access to hard-to-observe food stocks not captured in broadly available statistics, or travel between regions, etc.

Future work could refine the monthly distribution function to account for nontrivial market-based rationing, inter-annual carryovers of stocks, or the effect of conflict on food logistics. Perhaps the most useful would be a framework for updating the analysis iteratively based on additional data points, which might range from survey data (such as those conducted for example in Sudan by (Sudan Nutrition Sector 2024)) or a systematic integration of news and NGO reports about blockades, passage of food-aid trucks, available stocks, access or denial of access to agricultural land, etc.

For more information

- Clingendael Institute Reports: (Gaasbeek 2024; Hoffmann 2024)
- FEWS NET. Please note that the vital FEWS NET website (`https://fews.net`) is currently down following President Trump's suspension of foreign aid (Spitzer 2025). Multiple sources discuss their approach: Brown (2008); Famine Early Warning Systems Network (FEWS NET) (2018, 2021)
- FAO (e.g. Joint FAO/WHO/UNU Expert Consultation on Human Energy Requirements (2004); Food and Agriculture Organization of the United Nations (2024)).
- IPC Technical Manual (IPC Partners 2018) and note on Indicators Integrated Food Security Phase Classification (IPC) (2021)
- On physiological dynamics: Arsenault and McDonald (2015); Thomas et al. (2009); Preedy and Patel (2019)

5. Sensitivity Analysis

In order to assess how uncertainties in key parameters influence the output of the resource scarcity model, we provide code to implement a global sensitivity analysis following the variance-based method Saltelli (2008). This type of analysis enables researchers to systematically quantify which input factors exert the greatest influence on a model's variance in its main outcome of interest.

We use the Sobol method, which decomposes the total variance of the model output into contributions associated with individual parameters (first-order indices) and their interactions (higher-order indices). From these, one can also derive a total-order index to capture the overall effect of each parameter across all interaction terms. Because famine outcomes often depend on multiple interconnected factors, it is crucial to evaluate interactions rather than focusing solely on single-parameter changes.

5.1. Example Implementation: resource scarcity model

We provide an example of sensitivity analysis for the resource scarcity model in the `sensitivity.ipynb` Jupyter notebook. Five parameters were selected to reflect both the direct drivers of food availability and the more technical parameters governing how caloric deficits affect the population:

- *grain_multiplier*, capturing the possibility that total grain availability may differ from baseline estimates (e.g., due to international aid or improved harvests).
- *factor_deficit* and *factor_adj*, which modify how an individual's body mass index (BMI) responds to a sustained caloric deficit.

- *beta1*, controlling the degree of inequality in how any available food is distributed across different population percentiles.
- *distrib_kcal_min*, a minimum level of daily kilocalories allocated to the most food-insecure group in the population.

Each parameter was varied within a plausible range that captured the uncertainty or variability that might be encountered in real-world scenarios.

For each set of parameter values sampled using the Sobol sequence, the model was run to completion (e.g., simulating up to three years). The final number of excess deaths was recorded, though additional interim outputs (e.g., excess deaths after one or two years) were also tracked to examine how parameter effects evolved over time. This procedure was repeated for all sampled parameter combinations, producing a distribution of model outputs that spanned the chosen parameter space.

We compute the first-order (main) effects, total-order effects (and, optionally, higher-order indices). See for an example Figure 2.

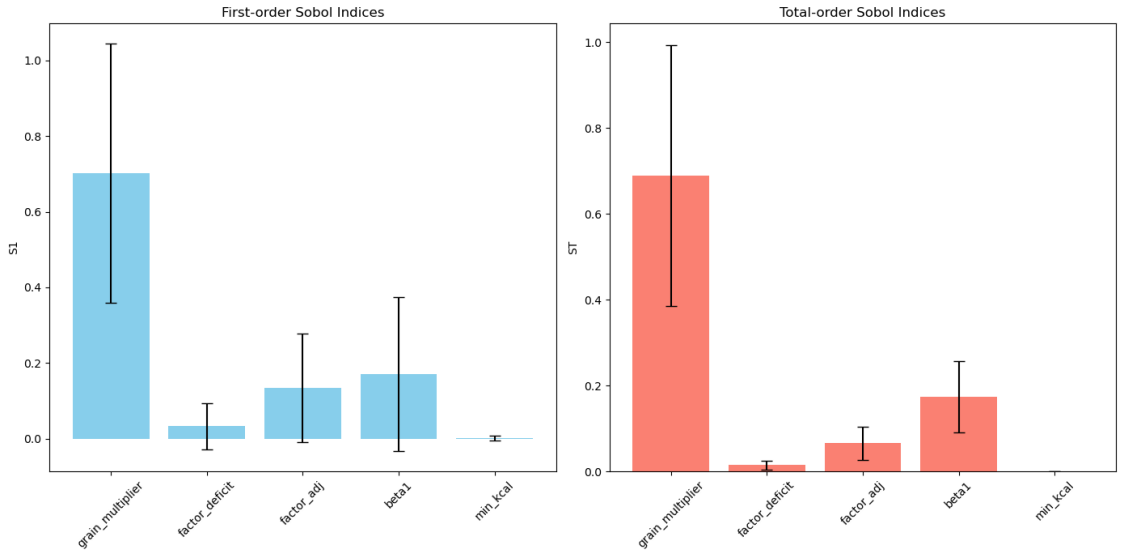

FIGURE 2. Example Sobol Indices from sensitivity analysis of the resource scarcity model

- **First-order indices (S1)** estimate the proportion of output variance that can be ascribed to each parameter by itself. A high first-order index implies that the parameter directly contributes a large share of the variability in excess mortality.
- **Total-order indices (ST)** measure the overall contribution of each parameter, including all interactions with other parameters.

Variance-based sensitivity methods are particularly suitable in this context for two reasons. First,

famine progression depends on both supply-side changes (total food availability) and population-level dynamics (how deficits accumulate in individuals, leading to health impacts). Hence, interactions among these parameters can have non-linear effects on excess mortality, making a global approach to sensitivity analysis preferable to simpler one-at-a-time or local analyses.

Second, global sensitivity analysis helps distinguish between parameters that are inconsequential within realistic ranges and those that demand more accurate empirical estimates. This prioritization is invaluable for guiding data collection efforts or for tailoring policy interventions. For instance, discovering that total food availability has a dominating effect on model outcomes might focus policy discussions on ensuring sufficient imports or aid distributions, whereas discovering that inequality in distribution patterns heavily shapes mortality outcomes might suggest a different, more targeted, set of interventions.

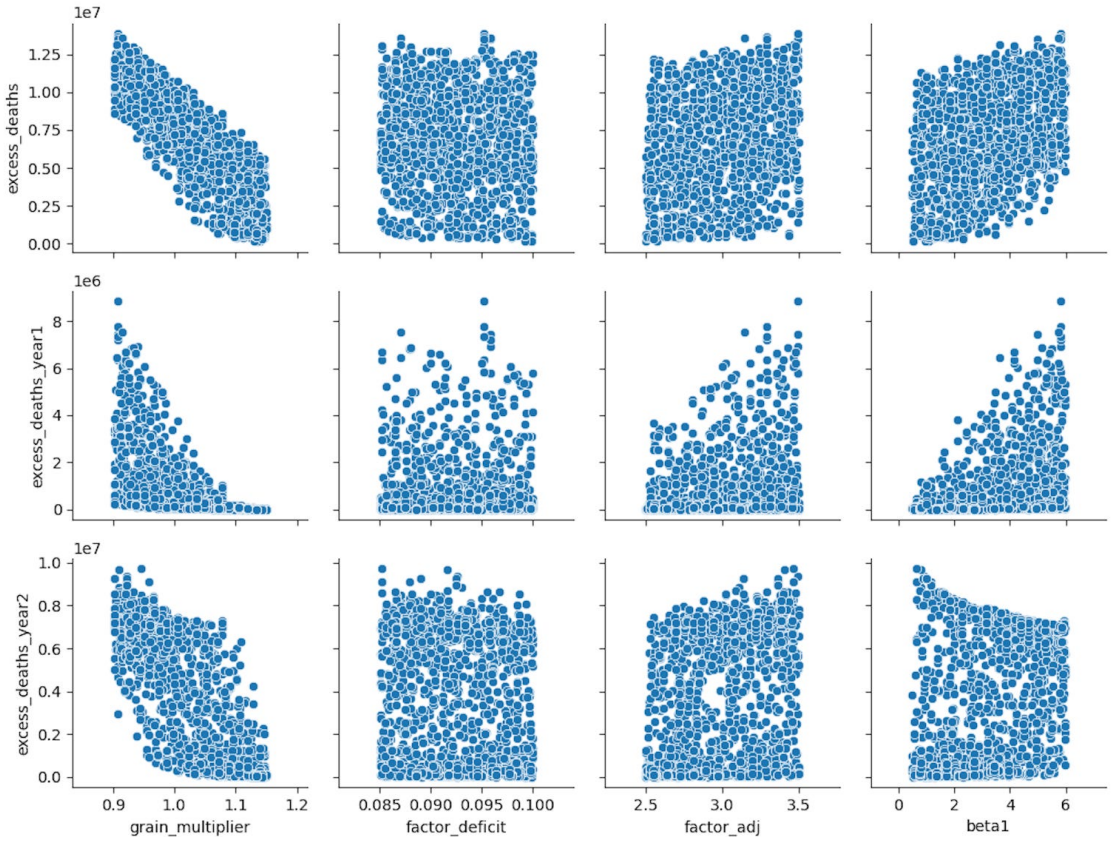

FIGURE 3. Scatterplots of input parameters vs. model predictions

Finally, we can visualize the relationship between each parameter and the predictions via scatter plots of parameter samples versus model outputs (See for example Figure 3). Parameters yielding strong monotonic or nonlinear relationships with mortality can be identified as high-

priority. Parameters whose scatter plots form a broad cloud often have negligible influence on the outcome, indicating robustness of the model to the exact specification of those parameters.

Overall, this approach clarifies which assumptions warrant further scrutiny and where policy efforts might have the greatest impact on mitigating famine-related mortality. It also ensures transparency in model development by explicitly quantifying how each source of uncertainty contributes to overall prediction variability.

References

Abdollahpour, S.; Kosari-Moghaddam, A.; and Bannayan, M. 2020. Prediction of wheat moisture content at harvest time through ann and svr modeling techniques. *Information Processing in Agriculture* 7(4):500–510.

Araujo, C.; Araujo-Bonjean, C.; and Brunelin, S. 2012. Alert at maradi: preventing food crises by using price signals. *World Development* 40(9):1882–1894.

Arsenault, J., and McDonald, C. 2015. Literature review part 2 review of adaptability of adults and children to short and long-term energy restriction. *FewsNet reports*. URL: https://fews.net/global/special-report/august-2015-0.

Aune, D.; Sen, A.; Prasad, M.; Norat, T.; Janszky, I.; Tonstad, S.; Romundstad, P.; and Vatten, L. J. 2016. Bmi and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *bmj* 353.

Balkrishna, S. P., and Visvanathan, R. 2019. Hydration kinetics of little millet and proso millet grains: effect of soaking temperature. *Journal of food science and technology* 56:3534–3539.

Baquedano, F. G. 2015. Developing a price warning indicator as an early warning tool - a compound growth approach. Fao global information and early warning system (giews) research brief, Food and Agriculture Organization.

Boudreau, T. 1998. *The food economy approach: a framework for understanding rural livelihoods*. Overseas Development Institute London.

Brown, M. E. 2008. *Famine early warning systems and remote sensing data*. Springer Science & Business Media.

Brown, M. E. 2010. *Famine Early Warning Systems and Remote Sensing Data*. Berlin Heidelberg: Springer-Verlag.

Cutler, P. 1984. Famine forecasting: Prices and peasant behaviour in northern ethiopia. *Disasters* 8(1):48–56.

de Waal, A. 2005. *Famine that Kills: Darfur, Sudan*. Oxford: Oxford University Press, revised edition. Original work published 1989.

Devereux, S. 2000. Famine in the twentieth century. Technical Report 105, Institute of Development Studies.

Famine Early Warning Systems Network (FEWS NET). 1985. Famine early warning systems network. https://fews.net/. Founded in 1985 to forecast and analyze acute food insecurity.

Famine Early Warning Systems Network (FEWS NET). 2018. *Integrating Acute Malnutrition and Mortality into Scenario Development*. Famine Early Warning Systems Network. Accessed: 2025-05-06.

Famine Early Warning Systems Network (FEWS NET). 2021. *Matrix Analysis: Integrated Analysis of Survey-Based Indicators for Classification of Acute Food Insecurity*. Famine Early Warning Systems Network. Accessed: 2025-05-06.

Food and Agriculture Organization of the United Nations. 2024. Special report – 2023 FAO Crop and Food Supply Assessment Mission (CFSAM) to the Republic of the Sudan. Technical report, Food and Agriculture Organization of the United Nations.

Gaasbeek, T. 2024. Sudan: From hunger to death. Clingendael alert, Clingendael Institute, The Hague. An estimate of excess mortality in Sudan, based on currently available information.

Henry, C. J. K. 2001. The biology of human starvation: Some new insights. *Nutrition Bulletin* 26(3):205–211.

Hillbruner, C., and Moloney, G. 2012. When early warning is not enough—lessons learned from the 2011 somalia famine. *Global Food Security* 1(1):20–28.

Hoffmann, A. 2024. From catastrophe to famine: Immediate action needed in sudan to contain mass starvation. Policy brief, Clingendael Institute, The Hague.

Howe, P. 2018. Famine systems: A new model for understanding the development of famines. *World Development* 105:144–155.

Integrated Food Security Phase Classification (IPC). 2021. *Indicators utilized by IPC*. IPC. Global.

IPC Partners. 2018. *Integrated Food Security Phase Classification (IPC) Technical Manual Version 3.0*. Food and Agriculture Organization, Rome.

Jocelyne, R. E.; Béhiblo, K.; and Ernest, A. K. 2020. Comparative study of nutritional value of wheat, maize, sorghum, millet, and fonio: some cereals commonly consumed in côte d'ivoire. *European Scientific Journal ESJ* 16(21):118–131.

Joint FAO/WHO/UNU Expert Consultation on Human Energy Requirements. 2004. Human Energy Requirements: Report of a Joint FAO/WHO/UNU Expert Consultation, Rome, 17–24 October 2001. Technical Report 1, Food and Agriculture Organization of the United Nations, Rome.

Maxwell, D., and Fitzpatrick, M. 2012. The 2011 somalia famine: Context, causes, and complications. *Global Food Security* 1(1):5–12.

Maxwell, D. 2019. Famine Early Warning and Information Systems in Conflict Settings: Challenges for Humanitarian Metrics and Response. Conflict research programme report, London School of Economics.

Mellor, J. W. 1986. Prediction and prevention of famine. *Federation Proceedings* 45(10):2427–2431.

Misselhorn, A. A. 2005. What drives food insecurity in southern africa? a meta-analysis of household economy studies. *Global Environmental Change* 15(1):33–43.

Mwebaze, E.; Okori, W.; and Quinn, J. A. 2010. Causal structure learning for famine prediction. In *Proceedings of the AAAI Spring Symposium on Artificial Intelligence for Development*, AAAI Technical Report.

Okori, W., and Obua, J. 2011. Supervised learning algorithms for famine prediction. *Applied Artificial Intelligence* 25(9):822–835.

Preedy, V. R., and Patel, V. B. 2019. *Handbook of famine, starvation, and nutrient deprivation: From biology to policy*. Springer International Publishing Switzerland.

Ross, K. W.; Brown, M. E.; Verdin, J. P.; and Underwood, L. W. 2009. Review of fews net biophysical monitoring requirements. *Environmental Research Letters* 4(2):024009.

Sadaka, S.; Osborn, S.; Atungulu, G.; and Ubhi, G. 2015. On-farm grain sorghum drying and storage. *Arkansas Grain Sorghum Production Handbook* 13.

Saltelli, A. 2008. *Global Sensitivity Analysis: the Primer*. John Wiley & Sons.

Seaman, J., and Holt, J. 1980. Markets and famines in the third world. *Disasters* 4(3):283–297.

Seaman, J. 2000. Making exchange entitlements operational: The food economy approach to famine prediction and the riskmap computer program. *Disasters* 24(2):133–152.

Sen, A. 1981. *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford: Oxford University Press.

Spitzer, G. 2025. A respected u.s. famine warning system is 'currently unavailable.' what's the impact? NPR, Goats and Soda (Morning Edition).

Sudan Nutrition Sector. 2024. Nutrition vulnerability analysis. Nutrition vulnerability analysis, Global Nutrition Cluster. Accessed: 2025-05-06.

Thomas, D. M.; Ciesla, A.; Levine, J. A.; Stevens, J. G.; and Martin, C. K. 2009. A mathematical model of weight change with adaptation. *Mathematical biosciences and engineering: MBE* 6(4):873.

Wang, D.; Andree, B. P. J.; Chamorro, A. F.; and Spencer, P. G. 2020. Stochastic modeling of food insecurity. Policy Research Working Paper 9413, World Bank.