University of New South Wales Law Research Series

Nowcasting for Hunger Relief: A Study of Promise and Perils

Wayne Wobcke, Caroline Compton, Fleur Johns, Jayson Lamchek and Siti Mariyah

[2022] *UNSWLRS* 17
Forthcoming in *Information Technology for Development*

UNSW Law
UNSW Sydney NSW 2052 Australia

E: LAW-Research@unsw.edu.au
W: http://www.law.unsw.edu.au/research/faculty-publications
AustLII: http://www.austlii.edu.au/au/journals/UNSWLRS/
SSRN: http://www.ssrn.com/link/UNSW-LEG.html

Forthcoming in *Information Technology for Development*.
Accepted 17 June 2022.
Please cite the final published version once available.

Nowcasting for Hunger Relief: A Study of Promise and Perils

Wayne Wobcke[a,b], Caroline Compton[c], Fleur Johns[d,e], Jayson Lamchek[d,f] and Siti Mariyah[a,b]

[a]*School of Computer Science and Engineering, University of New South Wales (UNSW), Sydney, Australia;*
[b]*UNSW Data Science Hub (uDASH), University of New South Wales (UNSW), Sydney, Australia;*
[c]*School of Law, Western Sydney University (WSU), Sydney, Australia;*
[d]*Faculty of Law and Justice, University of New South Wales (UNSW), Sydney, Australia;*
[e]*Department of Law, School of Business, Economics and Law, University of Gothenburg, Gothenburg, Sweden*;
[f]*Deakin University Law School, Melbourne, Australia*

Contact

Fleur Johns: Law Building (F8), University of New South Wales (UNSW), Kensington Campus, Sydney NSW 2052, Australia, email: fleur.johns@unsw.edu.au

Abstract

Pitched as an aid to better development decision-making, the digital platform HungerMap LIVE presents composite data on, and machine-learning-derived predictions of, food insecurity in 90 countries. Of its current version, this article asks the following questions: What work is HungerMap LIVE called upon to do in ICT for development (ICT4D) practice? How well is it set up to do that work? Combining technical (both computer science and statistical) and social analysis, this article employs a close reading method drawn from humanities and legal research not usually directed at digital platforms in combination with interview-based techniques. By this means, it scrutinizes HungerMap LIVE's potential to guide or mislead users and canvasses some elaborations that could enhance its usability. It argues that interdisciplinary research of this kind can counter both the historical and technological determinism troubling the ICT4D field and better position decision-makers to employ machine learning in history- and context-attentive ways.

Keywords: Information technology for development; famine; hunger; World Food Programme; ICT ethics; social impacts of machine learning; ICT for decision-making support; interdisciplinary research; Sustainable Development Goal 2

1. Introduction

Scholars and practitioners of development, disaster relief and emergency management, policymakers and members of the public more broadly are today presented with a range of new information and communication technology (ICT) tools designed to inform and support their decision-making to address humanitarian needs and work towards meeting the Sustainable Development Goals (SDGs) (Arendt-Cassetta, 2021). HungerMap LIVE is one example: a global hunger monitoring platform developed by the World Food Programme (WFP) to track and predict hunger in near real-time (MVAM, 2019).

Tools of this kind can, however, be difficult for development practitioners to decode, evaluate and use. That is the case even when, as in HungerMap LIVE, they incorporate lay-friendly explanations of terminology and methods (World Food Programme, n.d.-a). This article aims to facilitate the considered use of HungerMap LIVE and inform the commissioning and development of comparable tools in the future insofar as organizations like the WFP elect to pursue that course (leaving open the question of whether that is advisable compared to other potential uses of the time and money that such tools' development demands). It does so by presenting an interdisciplinary social and technical analysis, employing the method of close reading, exploring how HungerMap LIVE could potentially guide three possible user groups concerned with meeting development goals – specifically SDG 2, to end hunger and ensure food security: WFP staff; national policymakers; and communities engaged in advocacy or activism surrounding food security. Also advanced are some suggestions of how HungerMap LIVE could be further elaborated to address some unanswered questions arising from its initial version that are likely to be helpful to those working on other, analogous initiatives.

The research questions pursued here may, accordingly, be encapsulated as follows: *What work are digital platforms such as HungerMap LIVE being called upon to do in ICT for development (ICT4D) practice? How well is HungerMap LIVE currently set up to do that work?*

Inquiry along these lines elucidates in concrete terms the potential of real-time data and machine learning to improve famine mapping for a range of interested constituencies, as well as the limitations of these techniques and the challenges of integrating them to the benefit of all anticipated users. As Bonina et al. have observed, the value of digital platforms for social and economic development is often presumed, but rarely studied in specific socio-technical contexts (Bonina et al., 2021). By illuminating what development practitioners, activists and/or community advocates might and might not be able learn about food insecurity from relatively novel digital platforms like HungerMap LIVE, this article tackles, in effect, a broader question that has been axiomatic to the ICT4D field: to what extent might ICT contribute to the economic, social or political development of those in greatest need (Qureshi, 2015; Walsham, 2017)? It does so employing techniques of close reading (as noted above) that are most closely associated with legal research and the humanities (Lentricchia & DuBois, 2003; Orford, 2003). This builds upon the work of others who have shown the potential of these techniques for analysing formal, informational and contextual features of ICT interfaces (Bares et al., 2020). In the first part of the paper, below, we address further what the close reading of a particular application of ICT – here, HungerMap LIVE – can contribute to existing knowledge and research methods in the ICT4D field.

We proceed, thereafter, with an overview of HungerMap LIVE's institutional provenance and the rationale for its existence, fleshing out some of the contested notions about 'development' in which it trades. We briefly recall some features of the modern history of famine mapping and consider how HungerMap LIVE reflects preoccupations characteristic of that history, with particular attention to the prior work of the WFP. Then, we present a close reading of the design of HungerMap LIVE and a breakdown of its primary socio-technical elements and operations. We proceed to discuss some of the limitations of that design, including questions to which users may want or need answers that are not currently fully addressed on its web interface. We then propose a research agenda surrounding HungerMap LIVE for the task of food security modelling and prediction – this being a burgeoning subtheme in ICT4D practice (see, e.g., Karanasios & Slavova, 2019).

Our broad goal is to help enhance digital-platform-user 'literacy' (Gray et al., 2018) in ICT4D practice, especially among those who are intrigued by the potential uses of machine learning for policymaking purposes. Like Madon and Schoemaker, we believe that the implications of digital platforms' humanitarian use are likely to be very different for different stakeholders, and that encouraging platform developers and users to attend to these divergences is vital (Madon & Schoemaker, 2021). Too often the prospect of integrating machine learning and novel data sources into ICT4D practice is considered –

and sometimes championed – without regard to the diverse constituencies and communities engaged in that practice in specific domains, such as those concerned with the prediction, prevention, and amelioration of famine, for instance. This article underscores what can be learned from the close reading of particular implementations of machine learning in their institutional and sectoral contexts to ensure such purported technological ‘advances’ are workable, usable, and meaningful for their intended beneficiaries.

2. ICT4D and the Value of Close Reading

ICT4D is a field in which longstanding debates about modernity, inequality and distributive justice, the role of technology in historical change, dynamics of structure and agency in societal reproduction, and the relationship between economic, social and political phenomena are ongoing (cf. Sein et al., 2019). Researchers and practitioners developing ICT applications aimed at enhancing human development or making it more equitable carry a particular range of aspirations, assumptions and convictions into this field (Abubakre & Mkansi, 2022). For those working in or with international organizations, such as the WFP, these endeavours are further inflected by organizational imperatives and controversies, including budgetary pressures and the impetus to meet donors' demands for efficiency, inclusion and innovation (Dingwerth et al., 2019). While the narratives characteristic of the field remain broadly progress-oriented and diachronic, there has been recent emphasis on synchronic assessment of performance against particular metrics – above all, since 2015, against the SDGs (Gore, 2015). With this growing focus on in-depth, point-in-time measurement has come a rising appetite for tools and techniques that can avail development decision-makers of real-time or near-real-time information and insights into communities under-represented in traditional statistical analysis and data-gathering. This has spawned a wide range of digital innovation initiatives the potential value and impacts of which remain to be assessed (Bonina et al., 2021).

To date, efforts to assess how digital data sources and machine learning techniques might alter ICT4D tend to follow, broadly speaking, one of two tracks. On one hand, scholars and practitioners of ICT4D have engaged in fairly high-level speculation about the challenges posed and benefits potentially yielded by greater recourse being had to 'big data' and machine learning (see, e.g., Hilbert, 2016) and the impact of ICT4D policies and investments designed to support machine learning capabilities (Heeks, 2010). On the other hand, micro-analyses of the digitization of ICT4D have often expounded on the strengths and weaknesses of particular methods, models or tools relative to others of the same genre without reflecting much on the uses to which they may be put by different constituencies (see, e.g., Lentz et al., 2019).

This article mounts an argument not for a third track but rather for a cross-cutting pathway, combing elements of macro- and micro-, social, and technical analyses through close reading. In this instance, close reading entailed detailed, qualitative analysis of the

HungerMap LIVE website, scrutiny of the data sources identified in that website (where publicly available) and, for purposes of context and clarification, semi-structured interviews and/or question-and-answer correspondence with professionals involved in the development and promotion of digital platforms for famine mapping or comparable ICT4D initiatives. In total, 43 semi-structured interviews were undertaken with 38 interviewees affiliated with 18 different organizations, plus question-and-answer via email with two additional interlocutors who were not available for interview. Five of these interviews and/or email question-and-answer interactions were directly concerned with famine mapping; the remainder were concerned with a broader range of ICT4D projects and practices. Interviewees were located in Southeast Asia, Australia, the UK, Europe and the US and were recruited on the basis of their satisfaction of the following criteria: (1) at the time of interview, they worked, or had worked, at international organizations engaged in humanitarian or development work or at organizations providing data science services to governments or international organizations in connection with humanitarian or development work; and (2) they were familiar with, and had had some direct or indirect involvement in, projects seeking to generate and/or test data analysis tools to address development challenges. Table 1. below offers an overview of our interviewee cohort.

Table 1. Interviews Metadata

Organization	Role	Gender	Mode of engagement
AI start up	1 Founder	Male	Online interview
Data for development org	1 Architect/NGO founder	Male	In person interview
Data for development org	1 statistician	Male	In person interview
Data for development org	1 GIS expert	Female	Online interview
Data for development org	2 development practitioners	1 male, 1 female	In person interview
Government development agencies	3 public servants	2 male, 1 female	3 in person interviews
INGO	1 GIS expert	Female	Online interview
Mobile network provider	1 CSR employee	Male	In person interview
National statistics offices	7 statisticians	5 male, 2 female	2 in person interviews (4 people) 3 online interviews
UN agencies	13 policy and management employees	6 male, 7 female	9 in person interviews, 6 online interviews
	4 data scientists	4 male	2 in person interviews, 7 online interviews

	3 GIS experts	2 male, 1 female	2 in person interviews, 2 online interviews

This approach to close reading is somewhat comparable to the ‘contextualized’ analysis of ICT4D attempted by Abubakre and Mkansi, but places greater emphasis on institutional imperatives (historical and contemporary) and technical capabilities, features and limits than on identities, affective bonds or goals of particular technologists (cf. Abubakre & Mkansi, 2022). Close reading so approached demands collaboration among scholars proceeding from very different disciplinary starting points – interdisciplinary collaboration much called for in the ICT4D literature (see, e.g., Schelenz & Pawelec, 2022, p. 169). Among its other contributions to ICT4D literature, this article models a version of the interdisciplinary collaboration that the field has sought to encourage and suggests close reading as a methodological meeting point for it.

Close reading holds value as a meeting point for interdisciplinary research in the ICT4D field because of the ‘bracketing’ that it demands: a practice informed by the teachings of phenomenology in philosophy and those of New Criticism in literary studies (Gearing, 2004; Lentricchia & DuBois, 2003). That is, close reading invites a provisional setting aside of the kinds of claims or presuppositions of historical or technological determinism by which the field of ICT4D has been troubled to date (Schelenz & Pawelec, 2022). Close reading does not negate the significance of larger historical, societal, and technological forces, but it prioritizes analysis of their manifestation and strategic navigation in particular texts, interfaces, or other communication media. At that level, close reading accommodates the interplay of multiple perspectives, and allows for re-entry, so informed, into questions and conflicts beyond its immediate confines. In this article, our approach to close reading is novel in that it entails not just attending, with care, to the specificities, interactions and vagaries of words, institutional structures, and interpretative commitments but also to those of algorithms, different data formats, machine learning models and digital interfaces.

3. HungerMap LIVE: Background and Rationale

This section addresses the first of our research questions by situating HungerMap LIVE as a work product of the WFP, an organization created in 1961 under the joint auspices of the United Nations and the Food and Agriculture Organization of the United Nations (the FAO) for food distribution to promote economic and social development.[1] The WFP's remit later expanded to include emergency food relief and promoting world food security.[2] Any assessment of the efficacy of particular ICT applications oriented towards development must specify those understandings of 'development' at work in particular contexts (Zheng et al., 2018). In this instance, this requires brief recapitulation of the institutional setting in which HungerMap LIVE was developed, namely, the WFP.

Ideas of development within the WFP were, at the outset, heavily inflected by US economic priorities, particularly US need to dispense with agricultural surpluses and desire to extend and maintain Cold War alliances (Shaw, 2011). 'Development' in this context was undercut by racialized, ideologically charged notions of 'modernization'. Later years saw the WFP deploy food aid as a tool of land and labour market reform and as a medium of humanitarian diplomacy in armed conflict, with 'development' becoming framed as a practice of helping people to help themselves – reflecting the influence of neoliberal economic thought (Shaw, 2011). More recently, with the benefit of its accrued logistical expertise, the WFP has promoted ideas of hunger and food insecurity as, at least in part, manifestations of information processing problems, and grappled with them in those terms.

Pursuant to its changing sense of mission, the WFP has been engaged since inception in collecting, analyzing and disseminating information about food security (World Food Programme, 2020). HungerMap LIVE is among its latest efforts, an outcome of the WFP's collaboration with the Alibaba Foundation. The stated aim of that WFP-Alibaba partnership was to support achievement of SDG 2 – to end hunger and ensure food security – by tapping the 'expertise of the Alibaba Group' in data analytics and machine learning to 'help WFP become even more efficient and effective in its work' (World Food Programme, 2018b).

The story of HungerMap LIVE is also part of a longer history of modern famine mapping predating the WFP's creation. Efforts to record, predict, analyse and react to 'excess mortality' (that is, mortality above the level that historical data would lead people

---

[1] See, United Nations General Assembly Resolution No. 1714 (XVI), World Food Programme (1961), https://undocs.org/en/A/RES/1714(XVI).

[2] See, World Food Programme, General Regulations and General Rules (2014), https://docs.wfp.org/api/documents/eef12dfcbda04bddb7012f9c5002f356/download/.

to expect) due to starvation or hunger-induced diseases have taken various forms (Ó Gráda, 2007, 2009). Efforts of this kind belong to a tradition of distinguishing the normal from the pathological that has been central to modern biology, medicine, economics and social policy (Canguilhem, 1991).

In these endeavours, much depends on whether those drawing the distinctions are directly affected by famine or responding to famine. People with extensive experience of food shortages establish local repertoires of famine-response attuned to their recurrence over the longer term: switching to unconventional sources of nourishment, i.e., 'famine food'; borrowing money; and migration (Ó Gráda, 2009). Famine-affected communities have often 'gather[ed] a very broad base of knowledge' to understand and mitigate endemic famine risk, incorporating weather data and market data. In contrast, states and international organizations have often preferred intensified data-gathering at famines' terminal stages (Walker, 1989, p. 37).

Tensions of this kind – between famine-affected and respondent perspectives – have been axiomatic to the multiple re-readings of famine attempted in development practice throughout the modern period. Influential among these was the intervention of Amartya Sen in the 1980s who argued that famine results from food distribution inequities rather than food shortages. In contrast to prior interpretations (which emphasized crop failure as famine's main cause), Sen identified a range of factors that contributed to the historic Bengal Famine beyond diminution in quantum of food produced, including food acquisition by the British military, price rises and gouging, increased unemployment and poor food distribution networks (Sen, 1982). According to one scholar writing in 1991, 'Sen's analysis … irrevocably shifted the terms of the debate from shortage of food supply to the intervening variables between food production and consumption' (Watts, 1991, p. 22).

The adverse effects of these variables' interaction are often highly localized and usually felt most acutely in communities without access to large-scale commercial food suppliers. Two of the most important data sources for predicting food security utilized during recent decades are also highly localized: household surveying and the manual collection of price information from markets.

These variables take a central position in WFP's definition of 'vulnerability' to famine. The WFP's Vulnerability Analysis and Mapping (VAM)—a cartography unit established in 1994 to service the WFP's central office in Rome and its country offices

(Shaw, 2011) and providing both 'temporary' and 'long-term assistance' in 'vulnerability' analysis (Recalde, 2000)—defines 'vulnerability' as follows:

> the probability of an acute decline in food access or consumption levels below minimum survival needs ... [as] a result of both exposure to risk factors such as drought, conflict or extreme price fluctuations – and also of underlying socio-economic processes which reduce the capacity of people's ability to cope' (Recalde, 2000). The WFP currently employs around 200 analysts across 80 countries to undertake VAM analysis in collaboration with 'governments, UN agencies, local/international NGOs, regional bodies and academic institutions (World Food Programme, 2018a).

In traditional VAM analysis, price fluctuation is captured through gathering information from markets, while people's ability to cope is captured through household surveying. Unfortunately, however, these data collection techniques are expensive and time-consuming to deploy. From 2013 onwards, to try to address issues of poor data quality and coverage (Recalde, 2000), the WFP added mobile phone surveys – computer-assisted telephone interviewing (CATI) – to its array of data collection techniques. That is, the WFP began calling or texting randomly selected people to try to track the food security situation in various countries (Bauer, 2016). Mobile phone surveys allowed the WFP to obtain up-to-date information more quickly and cheaply from conflict-affected and hard-to-reach areas than traditional survey methods (Bauer, 2016; MVAM, 2019). Even so, the WFP's data collection challenges have persisted because certain areas and people cannot reliably be accessed through mobile phones. These areas and their inhabitants comprise blank spots in the WFP's hunger mapping work: areas of 'no data' or data restricted to 'high administrative level[s]' (MVAM, 2019).

A parallel exemplar of efforts to incorporate disparate data sources into the international mapping of famine was the Food Insecurity and Vulnerability Information and Mapping System (FIVIMS), developed under the auspices of the FAO in 1997 and 1998. The stated aim of FIVIMS was to address food insecurity and vulnerability (understood as an alloy of the shortage, inaccessibility, and poor utilization of food) as well as the risk of their recurrence. In other words, FIVIMS sought to counter: 'undernourish[ment] as a result of the physical unavailability of food, [people's] lack of social or economic access to adequate food, and/or inadequate food utilization', and the

'the full range of factors that place people at risk of becoming food-insecure', respectively (FAO, 2000, pp. 1, 13). To do so, FIVIMS encouraged local decision makers to identify indicators from which food security assessments could be made. The FIVIMS process emphasized heterogenous data sources; national FIVIMS networks reported sourcing data from governments, international organizations, bilateral aid agencies and NGOs (Bindraban et al., 2003, pp. 11–12).

The FIVIMS vision – of a consistent, worldwide understanding of food scarcity and risk assembled from diverse data sources – was further advanced by the Integrated Phase Classification (IPC) system, developed by the FAO in 2004 for use in Somalia. Since then, 15 organizations – including the WFP, as well as international non-governmental organizations like Oxfam, Care, and Save the Children, and the humanitarian consortium Food Security Cluster[3] – have promoted use of the IPC globally, adapting it for use in a range of contexts (IPC Overview and Classification System, n.d.). The IPC scale measures acute and chronic food insecurity, with insecurity levels measured from Phase 1 (minimal insecurity) to Phase 5 (catastrophe/famine), (IPC Global Partners, 2019, p. 3). To some extent, the IPC scale suggests an embrace of the idea of famine as a recurrent state, in line with the perspectives that Ó Gráda and Walker identified with the famine-affected. Yet institutional attention tends to remain focused, nonetheless, on the upper levels of the scale. The IPC classification system underpins many contemporary famine mapping exercises, including HungerMap LIVE.

In 2018, to try to address the persistent 'blank spots' highlighted above, and to signal that the WFP is 'at the forefront of innovation in the humanitarian world', the WFP entered into partnership with the Alibaba Foundation, as noted above (MVAM, 2019). Out of this partnership, the WFP created a new monitoring system called HungerMap LIVE that 'monitors food security in more than 90 countries and issues predictions for places where data is limited' (Ong, 2020) according to measures explained below.

In the foregoing efforts of the WFP and its partners, want of development – specifically, food insecurity – has been characterized, recurrently, as an information processing problem. ICT has, in turn, been cast as a tool for enhancing the way in which communities assemble, process, disseminate and utilize information for collective socio-

[3] The Inter-Agency Standing Committee (IASC) Cluster System operates in humanitarian emergency contexts, coordinating humanitarian activities undertaken by diverse actors working in particular thematic areas. These include food security, shelter, and health (amongst others).

economic benefit: a familiar framing in ICT4D research (Orlikowski & Iacono, 2001). To answer the first of our research questions: HungerMap LIVE has been called upon by the WFP and its partners both to reinforce this sense of hunger as a problem of data paucity or want of information, and to showcase the potential of ICT to address that problem. Controversy has persisted, however, as to *whose* developmental tool ICT ought primarily to be: a tool of the famine-affected or those seeking to respond to and ameliorate their plight – whether from a governmental, an NGO or an intergovernmental perspective. HungerMap LIVE endeavours to bracket such questions by assembling data that might be usable from a range of vantage points. In the next section, the paper examines how it does so before it considers to what degree of success it accomplishes the task.

4. From Contested Accounts to a Composite of 'Useful' Information

HungerMap LIVE presents users with a composite of data relating to food security, weather, population size, conflict, hazards, nutrition information and macro-economic data. Its main feature is a map showing the prevalence of insufficient food intake at the first administrative level (that is, the highest-level sub-national administrative unit in a given country: the state in Venezuela and India, the region in the Philippines, the province in Mozambique, for example, called hereafter a 'region' for convenience). The food insecurity of each region is depicted using a colour scheme of greens, yellows and reds, and its population density is indicated through brightness. When a user places a cursor above a particular region within any nation state depicted, two 'scores' appear, one indicating the prevalence of people with an 'insufficient' food consumption score (FCS) and one signaling the prevalence of people with a 'crisis' or 'above crisis'-level rating on the reduced Coping Strategy Index (rCSI), and the graphed movement in these figures over the preceding month and quarter. The aim of doing so as stated on the HungerMap LIVE website is to help users – 'WFP staff, key decision makers and the broader humanitarian community' – to assess, monitor and predict the magnitude and severity of hunger in near real-time.

HungerMap LIVE is designed in the mode of a dashboard. As with all dashboards, it has been designed with a view to facilitating vigilance, situational awareness and teamwork among its users (Noyes & Bransby, 2001). Many recurrent features of dashboards reflect these priorities: for example, the use of signals and visual schemes designed to foster speed of interpretation (Donald, 2001, p. 43). Common visual schemes – such as the red, yellow, green warning scheme familiar from traffic lights – feature prominently in dashboard design, and we find them across the HungerMap LIVE interface. Prevalence of insufficient food consumption, for instance, is represented using a colour scale that moves from dark green ('very low' or 0-5% prevalence) to bright red ('very high' prevalence, above 40%). In this way, HungerMap LIVE maintains a focus on famines' terminal stages – a focus characteristic of the perspective of respondents rather than those affected by famine (as discussed in Section 3 above).

Other visual cues direct users as to the need for action and the appropriate scale at which they should approach such action. Layout and positioning are used to emphasize certain information streams. The top left-hand corner of a screen, often regarded as 'prime real estate' in dashboard design (Few, 2013, p. 52), is given over to links to the WFP's downloadable reports offering 'insights' and 'key trends'. Also featured in this space is a search tool inviting users to select a country on which to focus their analysis from a limited number of nation states listed in a drop-down menu. Once a user does so, they encounter the above-mentioned map.

Despite the aura of authority, timeliness, and precision that their visualization in the format of HungerMap LIVE conveys, the food security measures underlying these graphic features are potentially problematic. These problems are highlighted in Section 6, in which we advance an extended response to the second of our research questions, namely: How well is HungerMap LIVE currently set up to do the work that it has been called by the WFP to do? To aid understanding of those potential problems, however, some further explanation is warranted.

5. How HungerMap LIVE Measures Food Insecurity

The WFP uses two main indicators as part of its food vulnerability analysis: the FCS (Food Consumption Score) and the rCSI (reduced Coping Strategy Index), combining aspects of the respondent and famine-affected perspectives discussed in Section 3. The FCS is a proxy indicator that measures the diversity of household diets, as well as the frequency of food consumption, using survey data (WFP Vulnerability Analysis and Mapping, 2008). It is calculated using the frequency of consumption of eight food groups by a household in the seven days leading up to the survey. These foods are given standardized weightings that reflect the respective nutrient density of each food type.

Households are classified as having either 'poor', 'borderline' or 'acceptable' food consumption, where poor food consumption typically indicates households are not consuming staples and vegetables daily, and seldom or never consume protein-rich foods. Acceptable food consumption typically refers to households that are consuming staples and vegetables daily, frequently accompanied by oil and pulses, and occasional proteins. The rCSI measures the frequency and severity of the behaviours that households typically engage in when faced with food shortages or when they lack the financial resources to buy food. These include reducing portion sizes, relying on less preferred foods, and skipping meals (Maxwell & Caldwell, 2008).

To create these two measures for purposes of the HungerMap LIVE interface, the WFP employs two methods. Both aim to deliver near-real-time indicators of insufficient food intake at the level of a region. The first approach uses digital technology to replicate traditional survey methods, employing CATI (explained in Section 3) to collect data on demographic variables, food consumption, coping strategies, and access to food, market, and health services. Calls are made on a rolling basis and spread evenly over cycles or 28 or 30 calendar days, or over three months, depending on the country. Sampling methods are used to ensure the representativeness of data collected.

WFP guidelines spell out a detailed methodology for how the FCS and rCSI are calculated from data collected using traditional face-to-face methods, and how those calculations are verified. Though not stated explicitly, presumably a similar process is used for the HungerMap LIVE data based on CATI surveys. Notable in both sets of guidelines is that, while there is strong commitment to repeatable methodology, human expertise is central in checking, interpreting, and validating the results.

For the FCS, prevalence of food insecurity is estimated by clustering representative household diets into 'similar' diets, allowing estimations to be made of the population more broadly. This clustering is done 'semi-automatically', by which is meant that an automatic clustering algorithm is used to generate clusters then examined by experts for coherence and possibly regenerated after adjusting weights. Although clusters and groupings are partially subjective, properly trained analysts are presumably expected to arrive at broadly similar conclusions, ensuring the consistency and repeatability of FCS analysis across similar contexts over time.

Each cluster of diets yields an FCS based on the mean consumption of each food type, assessed against cutoff points, to determine whether the diet is poor, borderline, or acceptable (by reference to the criteria described above). The cutoffs can also be manually adjusted by local experts depending on circumstances. Ideally, the resulting scores are validated against other proxies such as expenditure on food, with any anomalies investigated and explained; the extent to which this occurs as a matter of routine has not been publicly documented.

The rCSI, when calculated with traditional sources of data, focuses on the observed coping strategies adopted by households over a seven-day period, as explained above. The particular coping strategies surveyed are locally determined and each strategy is given a severity weighting. The overall rCSI for a household indicates the number of days on which a coping strategy was employed, multiplied by the severity of each coping strategy, measured against a locally determined threshold for food insecurity. The prevalence of food insecurity in a region is estimated by carefully taking representative samples of households in regions with known populations and weighting them accordingly.

The second approach employed by the WFP for the purposes of HungerMap LIVE, where CATI methods are not employed, uses machine learning, in particular regression trees (Breiman et al., 1984), to directly predict food insecurity in a region. More precisely, the WFP makes use of the Extreme Gradient Boosting (XGBoost) algorithm (Chen & Guestrin, 2016) to develop models whose output, for any given region, is the estimated prevalence of food insecurity (FCS or rCSI) in that region. No reasons are given by WFP for the choice of XGBoost, and no empirical comparisons with other methods have been published. We summarize this process below. However, note that while XGBoost itself is an automated procedure, the development of the machine learning models, as for survey data, relies on human expertise in several key aspects, and there is much scope for developing alternative models and using different methods.

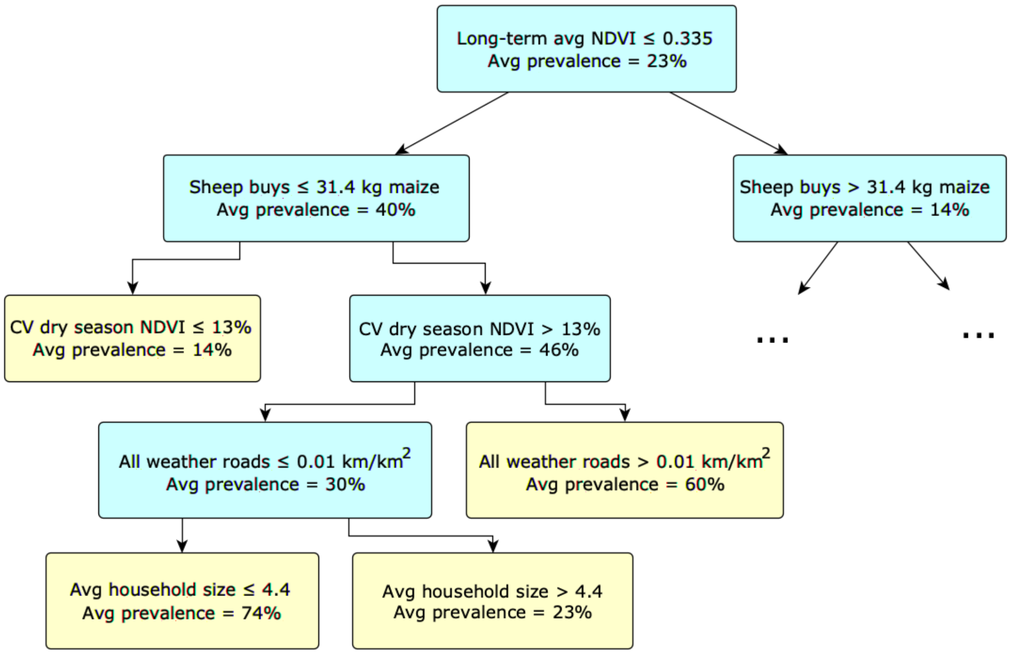
From a machine learning perspective, the task undertaken by HungerMap LIVE is a 'prediction' problem. The meaning of 'prediction' here does not, however, refer to prediction about the future, rather it entails determination of the relationship between a set of input features defined over various time points and periods, and the prevalence of food insecurity at the current point in time. In other words, the machine learning problem is to try to determine food insecurity 'now', or as close to now as possible, by predicting the FCS and rCSI for each of the various regions on the basis of the input data.

A crucial aspect of any machine learning problem is the identification of suitable input features. For HungerMap LIVE, the same input features are used for prediction of both FCS and rCSI prevalence for a region. These input features are not defined explicitly on the HungerMap LIVE website, but the XGBoost models are described as being 'built using information about population density, rainfall, vegetation status, conflict, market prices, macroeconomic indicators and undernourishment'. This information is drawn from a range of public and private, national, and international sources. Macroeconomic indicators, for example, are derived from those published by the New York-based, privately held corporation Trading Economics. For regions with a past measurement of FCS or rCSI, the latest such value is also included as an input feature.

It is not clearly described how these input features are 'built' or even how many features there are. For example, HungerMap LIVE makes apparent that it employs rainfall information derived from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data produced by the United States Geological Survey and researchers at the University of California, Santa Barbara, as processed by the WFP's VAM analysts. Presumably, the resulting input features include the data shown on the dashboard and explained on its glossary page (i.e., the 1-month rainfall anomaly data), aggregated to the region level. Beyond this, however, the process of feature selection and definition is unclear and not open to public discussion.

XGBoost is based on regression trees. The use of regression trees in the analysis of food security is not new. The portion of the regression tree shown in Fig. 2 is taken from Yohannes and Webb (1999), a guide to the application of regression trees to food security in Ethiopia. However, this tree was intended to be used for the identification of suitable food security indicators rather than for prediction. Note, as for HungerMap LIVE, the reliance on derived features: for vegetation, the long-term average normalized difference vegetation index (NDVI) and the CV (coefficient of variation, defined as the standard deviation divided by the mean) for NDVI in the dry season, together with price data (the relative price of sheep and maize), all weather road density and household size.

Figure 1. Regression Tree Fragment

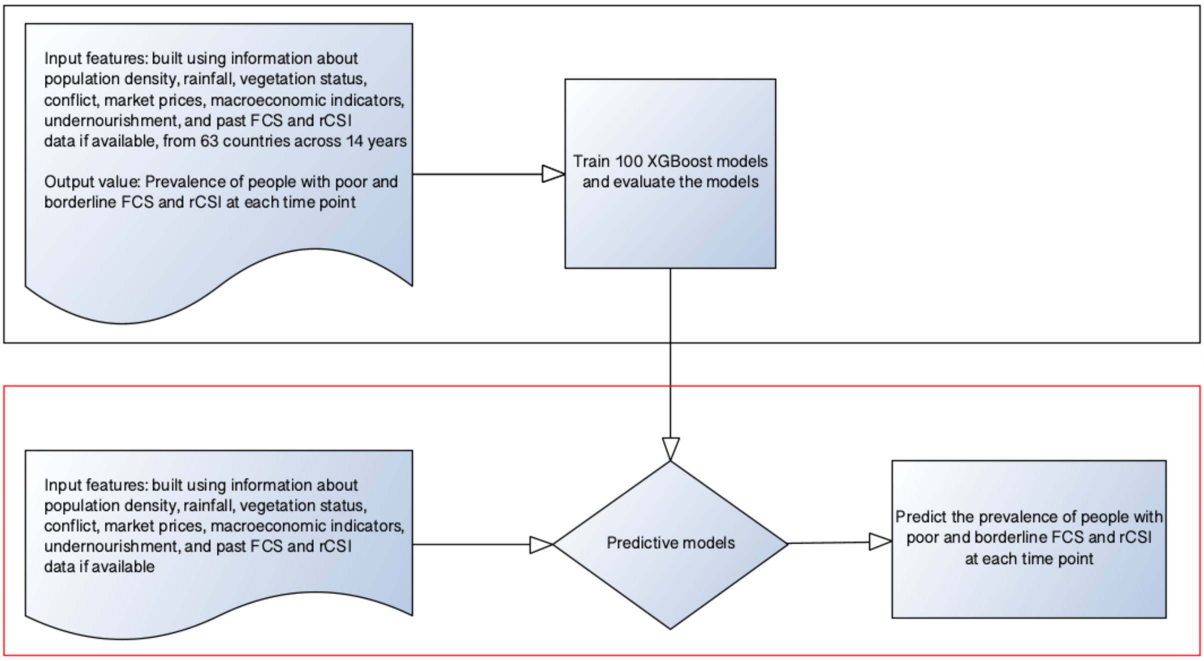


An XGBoost model is a *series* of such regression trees, where the objective of each new tree in the series is to predict the residual error resulting from using the output predictions derived from the previous series of trees. The output from the new tree is multiplied by the learning rate and added to the output from the previous series of trees to form the new prediction. Adding a new tree to the series thus aims to reduce the overall error on the training set, resulting in the high accuracy typically achieved by XGBoost models.

Regression trees are built 'top-down', starting with a single node, and repeatedly *splitting* nodes to further refine the decisions. As a result, such trees can grow to considerable depth in fitting the training data, but, as a result, tend to 'overfit' the training data. By this, we mean that, while the tree will work very well on the data that it is trained on, it hews too closely to the distinctions within that training data, limiting its utility on a different set of data. Within XGBoost, there are several parameters to be defined to reduce overfitting, such as the minimum gain needed to perform a split at a node, 'regularization' parameters in the objective function to be minimized or a fixed maximum to limit the depth of individual trees, and a learning rate. This provides considerable flexibility for researchers to manipulate the XGBoost model construction process, for example to trade off accuracy with model complexity.

The predictions produced by XGBoost models, however, can still be sensitive to the particular choices made during the development of the individual regression trees comprising the models. To capture this degree of uncertainty, in HungerMap LIVE, a statistical approach is adopted wherein the whole XGBoost model construction process is repeated 100 times using various subsamples of the training dataset, producing 100 XGBoost models for FCS and rCSI prevalence, as illustrated in Fig. 3. The final estimations for a region are then displayed on the HungerMap LIVE dashboard as a median value and a range of values, a 95% confidence interval, derived from the 100 models.

Figure 2. Model Generation and Operation



6. Some Cautionary Notes and Unanswered Questions

As outlined above, HungerMap LIVE relies on traditional data collection methods – CATI and in-person surveying – for some regions, while for others, uses data science or machine learning methods to provide a prediction of the prevalence of food insecurity. Food security is then assessed against the IPC scale, distinguishing between tolerable and intolerable levels of food insecurity, and directing users' attention towards locations and times at which intolerable levels are reported. Any assessment of the platform's utility requires an assessment of the reports of food insecurity made available on the interface. Users are likely to be particularly interested in the reliability of predictions generated using non-traditional data sources and machine learning techniques, recognizing that different user constituencies (WFP staff; national policymakers; and communities engaged in advocacy or activism surrounding food security) may have different priorities and concerns with respect to 'reliability'.

As noted above, the WFP has published extensive guidelines for the construction of FCS and rCSI scores from survey data, though it is unclear the extent to which such processes are incorporated into HungerMap LIVE. However, complications – and potential for user misunderstanding – arise when estimations from CATI survey data are shown on HungerMap LIVE alongside machine learning predictions of food insecurity. In particular, users could be misled by the treatment as commensurable of measurements with profoundly different levels of intermediation; the likelihood of error rates increasing when predictions are applied to countries from which no survey data has been obtained; and the difficulties associated with using assumption-bound regression trees to try to understand global diversity in conditions (discussed more in the research agenda proposed in Section 7 below).

A fundamental constraint is that, while the application of the survey methodology could reasonably be described as 'measuring' or 'estimating' the prevalence of food insecurity (especially in view of its focus on observational data on food intake and coping strategies over a short time period with human validation), the machine learning approach does not replicate this process, but rather aims to capture, via 'prediction', a posited *direct* relationship between the defined input features and the output values. The basic distinction between estimation and prediction is as follows: *estimation* presumes a population of individuals and aims to calculate the value of a population-level statistic such as the mean, median, variance, etc., whereas in a *prediction* problem, one instance at a time is considered with the aim of using a model to infer the value (class or numeric value) for an attribute of that instance. In a nutshell, the question raised and not yet resolved by HungerMap LIVE is the extent to which prediction – at the region level – can be used as a substitute for estimation in the range of settings in which HungerMap LIVE may be used.

It is a serious empirical question to what degree there is, in fact, *any* predictive connection from the combination of chosen input features to the desired output values. This is why validation of the machine learning models is crucial to understanding the effectiveness of the HungerMap LIVE dashboard to represent food insecurity in regions for which there is no survey data. It is also worth noting in this context that it is not possible to inspect the regression tree models used with HungerMap LIVE to verify if they conform to previous results. Even if this were possible, the models are very complex, each such model (of which there are 100 for FCS and rCSI) consisting of a large number of regression trees, resisting simple explanations. To elaborate on the unanswered questions surrounding HungerMap LIVE's representational effectiveness, we consider here some issues affecting the accuracy of the models.

6.1. Assessing the Accuracy of HungerMap LIVE's Models

One striking challenge presented by both statistical and machine-learning-generated prevalence values of food scarcity is the difficulty of assessing their accuracy. As we shall elaborate in Section 7 below, the required degree of accuracy of the measures must be considered in the context of the usage to which the results will be put, which depends on the needs and purposes of end user or stakeholder groups – which frequently diverge. Nonetheless, the promulgation of inaccurate measures as a potential basis for, or as factors to inform, distributive decision-making by WFP staff and national policymakers would be a legitimate matter of concern for those users and for communities engaged in advocacy or activism surrounding food security, particularly if inaccurate values result in inequitable distributions of aid or contribute to higher inclusion/exclusion errors for aid beneficiaries. The amenability of HungerMap LIVE to assessments of its models' accuracy is therefore relevant even though it is not the only factor worthy of consideration in this context – that is, even though attaining a level of accuracy tolerable for some users would not be enough to determine whether other user groups concerned with food insecurity would or should make use of the platform in their work, or might regard such use as fair or just (Lowens, 2020).

For measures of FCS and rCSI derived from survey data, the accuracy of the estimations is typically assessed by comparing them against other data sources. The most obvious approach for HungerMap LIVE would be to compare the machine learning predictions with actual data on food security. Unfortunately, for regions without this data, this is not possible, so validation methods are limited. This is especially the case, given that WFP is developing the machine learning predictive models *because* data collection is too expensive and costly, as explained in Section 3 above. Thus, the core assumption here is that the model behaves similarly in those regions with *and without* survey data.

As is standard in machine learning, validation of a method (where a 'method' can be considered a procedure to apply to training data to produce a model) proceeds by evaluating the method on subsets of the dataset for which there is known 'ground truth'. In the case of HungerMap LIVE, the ground truth is determined from the survey data from 63 regions over a 14-year period, along with the estimated FCS and rCSI scores for each time-period, i.e., the data used to train the final XGBoost models. Evaluation is usually done using cross-validation, for example 10-fold cross validation, where the data is divided into 10 equally sized subsets, and 10 predictive models are developed, each trained on nine of the subsets of the data and tested for accuracy on the remaining subset. For HungerMap LIVE, however, results are given on the website for only one division of the survey data into one training set (80% of the data) and one test set (the remaining 20% of the data), albeit with results averaged over 100 iterations of model development and testing based on sampling from the training dataset. It is not explained how this split is defined, yet this information is needed to understand the basis of the results and how those results would generalize to countries and regions without survey data. For example, if the 80-20 split is made at random over the examples in the dataset (i.e., regions and time points), as is generally done, the test dataset is very likely to include data points for all of the *same regions* as contained in the training data – whereas a key concern for the evaluation of HungerMap LIVE as a whole is to determine the accuracy of the final model for *different regions*, i.e., countries and regions without survey data. If the training and test data for model evaluation come from the same regions, and thus share characteristics (particularly input features that are defined at the country level), the errors are likely to be larger when the model is applied to other countries for which there is no survey data.

The results of this 'internal' evaluation (i.e., evaluation on the collected dataset without reference to alternative information that could be used for 'external' validation) are given as two quantities for FCS and rCSI: R^2 and MAE (mean absolute error). R^2 is a measure of the reduction in error when using the XGBoost model as compared to a baseline. MAE captures how 'far' the predictions are, on average, from the ground truth values (hence smaller is better), while R^2 is a proportional improvement (reduction in error) between 0 and 1 over always predicting the mean value on the training set, where higher (closer to 1) is better. Results are given for two scenarios: (i) where the latest FCS/rCSI score is used as an input feature; and (ii) where those scores are not used as input features. Results for FCS are an R^2 of 0.75/0.63 and MAE of 0.08/0.09, and for rCSI an R^2 of 0.78/0.73 and MAE of 0.06/0.07, for scenarios (i) and (ii) respectively. Such results, however, are difficult to interpret. Normally an R^2 value around 0.7 would be considered reasonable, but if, as is likely for FCS and rCSI, the baseline error is high, a proportional reduction of error of 0.6 may not be sufficient to yield meaningful results. Similarly, the MAE of 0.07 gives only the mean error and not the distribution of errors, i.e., it is possible that for some regions the error is much higher, for others much lower.

An alternative, and more direct, way of evaluating the method is also given implicitly by the WFP on the HungerMap LIVE dashboard. As described above, the predictions for any given region are made using a statistical procedure producing a median value and 95% confidence interval from 100 XGBoost models. The confidence interval in this context can best be understood as identifying the level of uncertainty generated by the machine learning prediction process itself; it is *not* a confidence interval in the conventional sense, that is, it does not specify the probability of the 'true' value lying within a given range. Unfortunately, the HungerMap LIVE models often give quite wide confidence intervals for the machine learning predictions, indicating that the XGBoost models do not capture a strong connection between input features and output values, limiting the usefulness of the predictions. The wide variations of predicted results cast doubt even on whether comparisons or rankings of food insecurity across different regions are meaningful.

The difficulty that even well-informed users of HungerMap LIVE are likely to experience in assessing the accuracy of its predictions is one factor that informs the research agenda that we propose below surrounding the enterprise of 'nowcasting' in the humanitarian field, of which HungerMap LIVE is emblematic. We argued above that HungerMap LIVE has been called upon both to reinforce the sense of hunger as a problem of data paucity, and to showcase the potential of ICT to address that problem. Insofar as it draws attention to those jurisdictions where machine learning 'nowcasting' is necessitated by a lack of data on food consumption or food-related coping strategies (Algeria and parts of the Central African Republic, for instance) then it might be seen to underscore the role of data deficiencies in impeding efforts to counter hunger. However, not all of the countries in this category exhibit high prevalence of hunger (Algeria did not, for example, at the time of writing). As to whether filling in those data blanks through ICT (machine learning specifically) is likely to enhance governments' and others' hunger-response capabilities, HungerMap LIVE is similarly uncompelling in its current form. Given the perils of approaching machine learning predictions as replacements for survey data (highlighted above), HungerMap LIVE does not at this stage engender great confidence that access to machine learning-derived data enhances collective capacity to understand, prevent or address hunger. Nonetheless, consideration of the extent to which such nowcasting might *yet* prove useful to those grappling with hunger—with adjustment and further elaboration—demands some reflection on the uses to which it could conceivably be put in this context and by whom.

7. Nowcasting, Forecasting and Narrowcasting: A Research Agenda

Given the shortcomings and difficulties canvassed in Section 6 above, drawn from close reading of the HungerMap LIVE interface, and the contexts and characteristics surveyed in the earlier sections above, the question to which this section turns is the following: what should researchers interested in the prospect of trying to 'nowcast' conditions of food insecurity, or vulnerability to the same, aim to work on next? Responses to this question depend in part on which prospective users and uses of 'nowcasting' tools towards which the research is oriented. Before sketching out some issues that warrant further attention among researchers, it is therefore apposite to identify, speculatively, some of the uses to which an interface such as HungerMap LIVE could be put and by whom.

7.1. Possible Users and Uses of HungerMap LIVE

To foreshadow possible usage of HungerMap LIVE, this section considers the three possible user groups to which the interface itself draws attention: national government officials; activists with links to local communities; and WFP staff. All of these groups use information pertaining to food security which may be represented with varying degrees of reliability by HungerMap LIVE.

First, key features of HungerMap LIVE make it attractive for use by governments that lack the resources and skills for data-gathering and analysis, and instead could rely on HungerMap LIVE rather than attempting to assemble a similar amalgam of information. Similarly, sub-national governments might use the interface to understand potential food shortages in their area of administration (although the limits of its current usefulness at sub-regional scales are discussed below under Research Questions).

Importantly, HungerMap LIVE also provides an aspirational model of potential data and policymaking integration. Users are presented concurrently with a wide variety of information which may encourage governments to further their own thinking about the relationships between different events and conditions and consider how they might use a range of data sources to understand these relationships better. This could conceivably facilitate cross-portfolio dialogue among government officials, which could over the longer-term help to engender understandings of food security as a whole-of-government imperative.

HungerMap LIVE's global coverage also makes it attractive for use by national officials keen to mobilize international aid towards their respective countries. Governments may want to extract from HungerMap LIVE evidence from *across different countries* that allow claims to be made that assistance to certain states should be prioritized over others. This then underlines the importance of ensuring that HungerMap LIVE's output data have equal reliability across countries.

Second, as discussed in Section 3 above, communities that include those recurrently affected by food insecurity often have an interest in obtaining and acting on local information about food security long before the terminal stage of a famine when states, international organizations and donors typically mobilize for relief. Activists may seek to use HungerMap LIVE to mobilize donors to try to prevent loss or reduction of local communities' livelihoods in advance of famine. Donors' interests might coincide with those of activists in this regard to the extent that preventive or early intervention typically costs less than ameliorative intervention at a later stage (Bailey, 2012; Venton, 2018). In order for HungerMap LIVE to be a persuasive early warning system of this kind, however, it must be able to demonstrate the onset of threats to livelihoods at a scale corresponding to the local communities in question (as discussed further below), which may be at sub-regional levels.

Third and finally, given the prominence afforded HungerMap LIVE in WFP communications, WFP staff themselves may be encouraged to use it for their own work. HungerMap LIVE presents at least three opportunities for WFP staff: support for fundand awareness-raising; assisting governments in the formulation of distributive policies and priorities; and informing the development and review of WFP's own funding and programmatic priorities.

The promotion of HungerMap LIVE through routes such as *Business Wire* suggests that it is at least partially aimed at private sector donor attraction and mobilization (Alibaba and WFP Unveil Next Generation of Machine Learning Technology in the Fight Against Hunger, 2019). The WFP's capacity to raise funds from individual giving and other private sector sources has been hampered to date by low levels of historic funding from such sources and poor 'brand awareness'; the extent to which it should be investing to try to raise greater private donor income remains a matter of dispute within the organization (WFP Office of Evaluation & Avenir Analytics, 2020). Even so, insofar as it showcases positive collaboration between the WFP and the business community (Alibaba), HungerMap LIVE may help the organization tap into new sources of philanthropy and technical assistance. The minimum standard of 'reliability' for such use may be lower than that required for distributive purposes. HungerMap LIVE could potentially attract derision as an instance of 'technology theatre' eliding hard questions 'about power and equity' (McDonald, 2020), but elements of 'theatre' might be what fundraising demands.

Given its straitened financial position, the WFP must invariably prioritize its distributive activities and funding allocations. The program in 2020 assisted 115.5 million people in 85 countries (World Food Programme, n.d.-b). Being able to identify and address needs before they become full-blown food crises could help alleviate some demands on the WFP's resources. While the accuracy of HungerMap LIVE is incidental to the fund- and awareness-raising functions canvassed above, it is vital to any such prioritization work. At the very least, any use of this tool for WFP programming requires a nuanced understanding of the evidentiary limitations of HungerMap LIVE.

7.2. Research Questions Pertinent to Such Users and Uses

In view of the limitations and potential uses of HungerMap LIVE identified above, this section proposes a research agenda for prospective users of this platform and for any other humanitarian organizations or government officials interested in employing machine learning to try to estimate conditions of human need —a burgeoning element of ICT4D practice.

7.2.1. Assessing the Relative Value of Country-Specific versus Generic Models and Features

A basic assumption of the current HungerMap LIVE platform is that, by design, the same model must work for all countries and regions in the same way: a frequent point of criticism of ICT4D projects (Schelenz & Pawelec, 2022, p. 172). Contrary to this assumption, it has been argued elsewhere that for chronic food security 'the degree of importance of each key driver varies between countries or regions according to their unique set of physical, economic, and social circumstances' (Fyles & Madramootoo, 2016). Moreover, as explained in Section 3 above, the WFP has otherwise been at pains to differentiate its famine mapping tools in dialogue with local decision-makers (in FIVIMS, for instance). This fact alone does not mean the machine learning models used in HungerMap LIVE cannot work in principle; the models *can* work provided the input features are rich enough to differentiate the regions where each driver contributes to FCS and rCSI prevalence in different ways, for then XGBoost can initiate splits to ensure different subtrees of the decision trees capture these differences. Whether the features are sufficiently rich is an empirical question that remains to be answered. In view of the cautionary notes discussed in Section 6 above, however, it is not clear that either the country- or region-specific validity of HungerMap LIVE's input features has been established.

As an example of variability unaccounted for in HungerMap LIVE's methodology that is relevant to acute food insecurity, Fyles and Madramootoo (2016) identify differences in how climate change is likely to affect food security. Sub-Saharan Africa and South Asia, which are heavily dependent on rainfall, are particularly vulnerable to droughts, whereas for other countries, higher temperatures may, in contrast, increase food yields, contributing positively to food security. With HungerMap LIVE, it is possible that vegetation indexes are sufficient to differentiate the regions susceptible to drought from those that are not. It would be desirable for prospective users to be able to inspect the XGBoost models to verify if in fact the decision trees are utilizing this, or some other, feature for this purpose, to assist in validating the models.

Delving more deeply into particular countries in southern Africa, it is unlikely the HungerMap LIVE features are adequate to capture important differences between regions. For example, in Eswatini (a country for which WFP does not have survey data), 'the impacts of droughts [at the household level] can vary significantly between constituencies and regions' (Mohammed & Dlamini, 2018). This variation is due to differences between livelihoods (essentially poverty-related features), but also whether rainwater or tap water is the main supply of drinking water. Though these differences manifest at the household level, they are likely to impact the prevalence of food insecurity at the region level when aggregated. Development of new country-specific proxies to incorporate into machine learning models, informed by existing domain expertise, is needed to better capture the variable drivers of food insecurity.

7.2.2. Making and Validating Localized Causal Models

Another problem with the machine learning framework of HungerMap LIVE is the use of a mixture of long-term and short-term features for 'nowcasting' acute, rather than chronic, food insecurity. In reality, these features typically interact to form a complex causal network of factors that influence food security over different time scales. A related problem is that this input data itself is measured at different time scales and updated at different frequencies: temporal variations that are elided by the interface and hence not drawn to users' attention.

A meta-analysis that aims to unravel the complex causal network of food security drivers is given by Misselhorn (2005) for Southern Africa, which distinguishes both short- and long-term drivers, and direct and indirect drivers (that is, indirect drivers initiating other drivers that in turn influence food security). Misselhorn's conception of causal models draws on probabilistic reasoning (Pearl, 2009) in the characterization of *influence* between random variables (drivers and outcomes) formalized using conditional probabilities, with reliance on Bayesian independence assumptions to simplify models, though no formal causal model is presented in the analysis.

A major advantage of this approach to causal reasoning is the possibility to use data to develop and evaluate causal models. Such models would enhance understanding of the role of the various drivers of food security in different causal networks and could enable predictions of the efficacy of various interventions, following the approach used in medical research, albeit without the possibility of randomized controlled trials. The long-term data amassed by the WFP for HungerMap LIVE can provide a valuable starting point for this endeavour. However, what is needed for purposes of further research in this direction are more fine-grained data proxies (as described above) tailored to local context and conditions, with the input of domain experts to develop local models. This is not simply a matter of using 'off the shelf' machine learning models with more data; rather new localized models and features are required, especially as Misselhorn (2005, p. 37) cautions that, for her meta-analysis, 'the combination of factors in each case study varied significantly'.

7.2.3. Refining the Temporal Dimensions of Forecasting

In addition to the development and validation of more localized causal models, the temporal dimension of food insecurity is in need of further analysis, and we suggest, explicit representation in the models. One obvious reason for this is that food insecurity is seasonal, as has often been noted (e.g., Anderson et al., 2018; Beveridge et al., 2019; Lentz et al., 2019; Misselhorn, 2005; Mohammed & Dlamini, 2018). Improved machine learning models should incorporate such seasonal variation. Another need for the temporal dimension is the sensitivity of rCSI to market prices, which are more volatile than other factors, so better predictions of market prices should improve estimates of rCSI prevalence levels. But the most important motivation for the use of temporally refined forecasting models is to develop an 'early warning' system for food insecurity, so that action might be taken to avert or reduce the severity of an impending crisis. As noted above, this has been a recurrent goal for many of HungerMap LIVE's anticipated users.

Future work could draw upon recent studies using simpler forecasting models of food security. One example is provided by Wang et al. (2020), involving use of a regression model to predict the distribution of IPC phases (IPC1, IPC2 and IPC3+) at the country level on a quarterly basis. That study used a model incorporating the previously predicted value along with vegetation, rainfall, conflict, and economic variables as inputs; the authors concluded that around 15 data points are needed to develop a reliable model (as discussed above, this conclusion would need to be tested in different country and regional contexts). Another recent study examined, as an alternative to forecasting outcomes, the problem of predicting transitions in IPC class, attempted for Ethiopia at the level of livelihood zones (Westerveld et al., 2021). In this context, the best results arose from predicting seven months out from the current time. It may be fruitful to utilize time series analysis methods that have been developed to handle seasonal effects in data ranging over several years, adapted to employ a wider variety of input data features.

7.2.4. Narrowcasting Spatial Models

Another obvious property of the HungerMap LIVE models is that each region is treated independently, with no interaction between regions that are similar, in the same country, or even adjacent, although regions in the same country will share the same country-level features. A reasonable question is whether the whole machine learning process would work better at a finer level of spatial granularity, such as level 2 administrative units (at district or municipality level), though models at this level will of course be more data intensive. The work based on livelihood zones in Ethiopia (Westerveld et al., 2021) using monthly data to forecast changes in IPC level showed that this is feasible. Most interestingly, a *livelihood zone*, taken from FEWS NET (the Famine Early Warning Systems Network website, created by the US Agency for International Development and the US Department of State) is 'a geographical area within which people share the same patterns of access to food and income … and have the same access to markets' (Westerveld et al., 2021, p. 3). This empirically-based definition means that the livelihood zones can potentially be automatically discovered by machine learning techniques, e.g., clustering over the data collected (though it is unclear whether the livelihood zones for Ethiopia were so constructed by FEWS NET). Combining methods of determining cohesive spatial areas with (possibly multiple) predictive models such as XGBoost that are defined over such areas will result in more complex hybrid systems with several machine learning components that can be validated independently. This may be a promising approach to improving the accuracy and localization of the HungerMap LIVE models. Given their empirical basis, such models are also more likely to be explainable to and interpretable by the prospective users canvassed above.

A second approach to spatial modelling adapts well known statistical approaches of *small area estimation* (Rao & Molina, 2015). Small area estimation is traditionally applied with survey data, with the aim of using high-level regional features to compensate for small sample sizes in 'small' areas. This has been applied extensively in poverty mapping (Molina et al., 2019; Molina & Rao, 2010). The use of small area estimation to measure food security is less common, exceptions being studies of food insecurity in

Nepal (Haslett et al., 2014), Bangladesh (Hossain et al., 2020) and Ethiopia (Shiferaw, 2020).

A promising line of research would be to incorporate features from various data sources other than survey data into small area estimation models, following early work in this direction mapping poverty in Italy using mobile phone data (Marchetti et al., 2015). There are two ways in which small area estimation techniques might improve estimations of food insecurity: (i) models could explicitly incorporate features of nearby countries with similar characteristics (such as in Southern Africa, with all countries highly susceptible to drought); and (ii) to capture relationships among countries or regions whose economies are interconnected (for example by trade or labour market movements) and therefore more likely to directly affect food security in one another.

8. Conclusion

The main contributions of this article to the ICT4D field are threefold. As to ICT4D practice, we have advanced insights and offered cautionary notes to facilitate considered use of HungerMap LIVE and inform the commissioning and development of comparable tools in future by attention to their limits and blind spots. As to both scholarship and practice in ICT4D, we have identified some areas in which further research may be fruitful to address those limits and blind spots and to enhance later development of digital platforms for development. And as to ICT4D research methodology, we have demonstrated the value of interdisciplinary close reading to inform and concretize ongoing debates in the ICT4D field concerning the potential of new digital data sources and machine learning for different constituencies of development practice. Close reading of particular platforms can equip ICT4D scholars to move beyond broad, equivocal assessment of digital platforms as having ‘potential for good and evil’ in the development field (Bonina et al., 2021, p. 895). Close reading can also generate specific strategic imperatives: highlighting, for instance, the importance of platform developers investing as much time and effort working with people in, say, Algeria and Eswatini to generate machine learning models meaningful for their purposes as has historically been invested in refining survey practice. The fact that the WFP and its partners have *capacity* to generate machine learning insights from afar does not mean that they should be approaching the exercise remotely; indeed, we have argued that machine-learning-derived insights generated at such remove risk being useless for most purposes other than fundraising and general awareness-raising.

Implicit in the foregoing analysis is the following argument: the integration of machine learning with other methods in ICT4D with a view to a spatially and temporally differentiated assessment of human needs, political possibilities and socio-economic conditions to the aid of intended users and beneficiaries is invariably a multi-disciplinary task. It demands as much attention be paid to historical and contemporary social contexts as it does to novel technical inputs and capabilities: histories and contexts, that is, of the institutions and communities of *mappers*, as much as those of particular places *being mapped*. It requires, also, localized collaboration and engagement with multiple prospective user groups and beneficiaries, anticipating cross-purposes among them and attending to elision and obfuscation in the technical interfaces themselves. There is promise in the data-gathering and analytical approaches that HungerMap LIVE showcases. However, that promise can only be realized if as much effort is dedicated to engaging disparate human inputs (including critical ones) as to assembling and making use of automated inputs. The 'socio' side of socio-technical experiments in 'nowcasting' merits as much weight and investment as the 'technical' side. In the ICT4D field and elsewhere, the development of machine learning tools and digital platforms is often characterized as a singularly STEM undertaking (the concern of science, technology, engineering and mathematics) requiring only occasional inputs and surrounding ethical constraints from HASS (humanities and social sciences) fields. This article shows such tools' development, assessment, and informed use in ICT4D to be an unavoidably integrated endeavour.

Acknowledgements

This research was supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (project DP 180100903). We are grateful to the staff of the World Food Programme for their willingness to answer questions in connection with this research. The views expressed here are those of the authors and are not necessarily those of the Australian Government, the Australian Research Council, or the World Food Programme.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (project DP 180100903).

Notes on contributors

Wayne Wobcke is an Associate Professor in the School of Computer Science and Engineering at UNSW. His research covers a range of topics, including intelligent agent theory, personal assistants, agent-based modelling, recommender systems and computational social science.

Caroline Compton is a Lecturer at WSU School of Law. She studies the operation of plural normative systems, employing an interdisciplinary approach and ethnographic methods.

Fleur Johns is Professor in the Faculty of Law and Justice at UNSW working in international law, legal theory and law and technology. She is also an Australian Research Council Future Fellow and, in 2021–2024, a Visiting Professor in the School of Business, Economics and Law at the University of Gothenburg, Sweden.

Jayson Lamchek was, at the time of writing this article, a Post-doctoral Research Associate on the Data Science in Humanitarianism project at the Faculty of Law and Justice at UNSW. He is now a Post-doctoral Research Associate based at Deakin University Law School working on a project supported by the Cyber Security Cooperative Research Centre. He is a socio-legal scholar with expertise in human rights in developing countries.

Siti Mariyah is a PhD student in the School of Computer Science and Engineering at UNSW. Her research areas are data mining and applied machine learning and statistics.

References

Abubakre, M., & Mkansi, M. (2022). How do technologists do “ICT for development”? A contextualised perspective on ICT4D in South Africa. *European Journal of Information Systems*, *31*(1), 7–24.

Alibaba and WFP Unveil Next Generation of Machine Learning Technology in the Fight Against Hunger. (2019, September 25). *Business Wire*. https://www.businesswire.com/news/home/20190925005712/en/Alibaba-and-WFP-Unveil-Next-Generation-of-Machine-Learning-Technology-in-the-Fight-Against-Hunger

Anderson, C. L., Reynolds, T., Merfeld, J. D., & Biscaye, P. (2018). Relating seasonal hunger and prevention and coping strategies: A panel analysis of Malawian farm households. *The Journal of Development Studies*, *54*(10), 1737–1755.

Arendt-Cassetta, L. (2021). *From Digital Promise to Frontline Practice: New and Emerging Technologies in Humanitarian Action*. United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Policy Branch.

Bailey, R. (2012) *Famine Early Warning and Early Action: The Cost of Delay.* Royal Institute of International Affairs.

Bares, A., Zeller, S., Jackson, C. D., Keefe, D. F., & Samsel, F. (2020). Using close reading as a method for evaluating visualizations. *Proceedings of the 2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV)*, 29–37.

Bauer, J.-M. (2016, March 10). Mobile Phone Surveys Can Help World Food Programme Reach Hungry People. *The Guardian*.

Beveridge, L., Whitfield, S., Fraval, S., van Wijk, M., van Etten, J., Mercado, L., Hammond, J., Davila Cortez, L., Gabriel Suchini, J., & Challinor, A. (2019). Experiences and drivers of food insecurity in Guatemala’s dry corridor: Insights from the integration of ethnographic and household survey data. *Frontiers in Sustainable Food Systems*, *3*, Article 65. https://doi.org/10.3389/fsufs.2019.00065

Bindraban, P. S., Aalbers, H. L., Moll, H. A. J., Brouwer, I. D., van Dorp, M., Houtman, C. B., Brouwer, M. L., Zuurbier, M. M. M., & Hagenaars, E. C. M. (2003). *Focus on Food Insecurity and Vulnerability: A Review of the UN System Common Country Assessments and World Bank Poverty Reduction Strategy Papers*. FAO.

Bonina, C., Koskinen, K., Eaton, B., & Gawer, A. (2021). Digital platforms for development: Foundations and research agenda. *Information Systems Journal*, *31*(6), 869–902.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees*. Chapman and Hall/CRC.

Canguilhem, G. (1991). *The Normal and the Pathological* (C. R. Fawcett & R. S. Cohen, Trans.). Zone Books.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Dingwerth, K., Witt, A., Lehmann, I., Reichel, E., & Weise, T. (2019). *International Organizations under Pressure: Legitimating Global Governance in Challenging Times*. Oxford University Press.

Donald, C. (2001) Vigilance. In Noyes, J. and Bransby, M. (Eds.) *People in Control: Human Factors in Control Room Design* (pp. 35–50). The Institution of Engineering and Technology.

FAO. (2000). *Food Insecurity and Vulnerability Information and Mapping Systems (FIVIMS)*. FAO.

Few, S. (2013). *Information Dashboard Design: Displaying Data for At-a-Glance Monitoring* (2nd ed.). Analytics Press.

Fyles, H., & Madramootoo, C. (2016). Key drivers of food insecurity. In C. Madramootoo (Ed.), *Emerging Technologies for Promoting Food Security: Overcoming the World Food Crisis* (pp. 1–19). Woodhead Publishing.

Gearing, R. E. (2004). Bracketing in research: A typology. *Qualitative Health Research*, *14*(10), 1429–1452.

Gore, C. (2015). The post-2015 moment: Towards sustainable development goals and a new global development paradigm. *Journal of International Development*, *27*(6), 717–732.

Gray, J., Gerlitz, C. & Bounegru, L. (2018) Data infrastructure literacy. *Big Data & Society*, *5*(2), 1–13.

Haslett, S., Jones, G., Isidro, M., & Sefton, A. (2014). *Small Area Estimation of Food Insecurity and Undernutrition in Nepal*. Central Bureau of Statistics, National Planning Commissions Secretariat, World Food Programme, UNICEF and World Bank, Kathmandu, Nepal.

Heeks, R. (2010). Do information and communication technologies (ICTs) contribute to development? *Journal of International Development*, *22*(5), 625–640.

Hilbert, M. (2016). Big data for development: A review of promises and challenges. *Development Policy Review*, *34*(1), 135–174.

Hossain, Md. J., Das, S., Chandra, H., & Islam, M. A. (2020). Disaggregate level estimates and spatial mapping of food insecurity in Bangladesh by linking survey and census data. *PLoS ONE*, *15*(4), e0230906.

IPC Global Partners. (2019). *Integrated Food Security Phase Classification Technical Manual Version 3.0. Evidence and Standards for Better Food Security and Nutrition Decisions*. http://www.ipcinfo.org/fileadmin/user_upload/ipcinfo/manual/IPC_Technical_Manual_3_Final.pdf

IPC Overview and Classification System. (n.d.). IPC Global Platform. Retrieved November 9, 2020, from http://www.ipcinfo.org/ipcinfo-website/ipc-overview-and-classification-system/en/

Karanasios, S., & Slavova, M. (2019). How do development actors do "ICT for development"? A strategy-as-practice perspective on emerging practices in Ghanaian agriculture. *Information Systems Journal*, *29*(4), 888–913.

Lentricchia, F., & DuBois, A. (2003). *Close reading: The reader*. Duke University Press.

Lentz, E. C., Michelson, H., Baylis, K., & Zhou, Y. (2019). A data-driven approach improves food insecurity crisis prediction. *World Development*, *122*, 399–409.

Lowens, E. (2020). Accuracy is not enough: The task mismatch explanation of algorithm aversion and its policy implications. *Harvard Journal of Law & Technology*, *34*(1), 259–278.

Madon, S., & Schoemaker, E. (2021). Digital identity as a platform for improving refugee management. *Information Systems Journal*, *31*(6), 929–953.

Marchetti, S., Giusti, C., Pratesi, M., Salvati, N., Giannotti, F., Pedreschi, D., Rinzivillo, S., Pappalardo, L., & Gabrielli, L. (2015). Small area model-based estimators using big data sources. *Journal of Official Statistics*, *31*(2), 263–281.

Maxwell, D., & Caldwell, R. (2008). *The Coping Strategies Index: Field Methods Manual* (2nd ed.). Cooperative for Assistance and Relief Everywhere.

https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp211058.pdf

McDonald, S. M. (2020, July 13). Technology theatre. Centre for International Governance Innovation. https://www.cigionline.org/articles/technology-theatre/

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.

Mohammed, M., & Dlamini, T. (2018). Predictors of food insecurity in Eswatini: Lessons from the 2015/16 El Niño induced drought. *African Review of Economics and Finance*, *10*(2), 69–96.

Molina, I., & Rao, J. N. K. (2010). Small area estimation of poverty indicators. *The Canadian Journal of Statistics*, *38*(3), 369–385.

Molina, I., Rao, J. N. K., & Guadarrama, M. (2019). Small area estimation methods for poverty mapping: A selective review. *Statistics and Applications*, *17*(1), 11–22.

MVAM. (2019, October 22). Introducing Hunger Map LIVE. *MVAM: THE BLOG*. https://mvam.org/2019/10/22/introducing-hunger-map-live/

Noyes, J. & Bransby, M. (2001) *People in Control: Human Factors in Control Room Design.* The Institution of Engineering and Technology.

Ó Gráda, C. (2007). Making famine history. *Journal of Economic Literature*, *45*(1), 5–38.

Ó Gráda, C. (2009). *Famine: A Short History*. Princeton University Press.

Ong, A. G. (2020). WFP launches HungerMap Live. World Food Programme. https://www.wfp.org/stories/wfp-launches-hungermap-live

Orford, A. (2003). *Reading Humanitarian Intervention: Human Rights and the Use of Force in International Law*. Cambridge University Press.

Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: Desperately seeking the "IT" in IT research—a call to theorizing the IT artifact. *Information Systems Research*, *12*(2), 121–134.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*, 96–146.

Qureshi, S. (2015). Are we making a better world with information and communication technology for development (ICT4D) research? Findings from the field and theory building. *Information Technology for Development*, *21*(4), 511–522.

Rao, J. N. K., & Molina, I. (2015). *Small Area Estimation* (2nd ed.). Wiley.

Recalde, P. (2000) An overview of vulnerability analysis and mapping (VAM), Presented at the Meeting on Cartography and Geographic Information Science, United Nations, New York, NY, Mar 2000.

Schelenz, L., & Pawelec, M. (2022). Information and communication technologies for development (ICT4D) critique. *Information Technology for Development*, *28*(1), 165–188.

Sein, M. K., Thapa, D., Hatakka, M., & Sæbø, Ø. (2019). A holistic perspective on the theoretical foundations for ICT4D research. *Information Technology for Development*, *25*(1), 7–25.

Sen, A. (1982). *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford University Press.

Shaw, D. J. (2011). *The World's Largest Humanitarian Agency: The Transformation of the UN World Food Programme and of Food Aid*. Palgrave Macmillan.

Shiferaw, Y. A. (2020). Model-based estimation of small area food insecurity measures in Ethiopia using the Fay-Herriot EBLUP estimator. *Statistical Journal of the IAOS*, *36*(S1), 177–187.

Venton, C. C. (2018) *Economics of Resilience to Drought in Ethiopia, Kenya, and Somalia: Executive Summary*, U.S. Agency for International Development Center for Resilience, Washington, DC.

Walker, P. (1989) *Famine Early Warning Systems: Victims and Destitution.* Earthscan Publications.

Walsham, G. (2017). ICT4D research: Reflections on history and future agenda. *Information Technology for Development*, *23*(1), 18–41.

Wang, D., Andrée, B. P. J., Chamorro, A. F., & Girouard Spencer, P. (2020). *Stochastic Modeling of Food Insecurity*. Policy Research Working Paper 9413, World Bank, Washington, DC.

Watts, M. (1991) Entitlements or empowerment? Famine and starvation in Africa, *Review of African Political Economy*, *51*, 9–26.

Westerveld, J. J. L., van den Homberg, M. J. C., Nobre, G. G., van den Berg, D. L. J., Teklesadik, A. D., & Stuit, S. M. (2021). Forecasting transitions in the state of food security with machine learning using transferable features. *Science of the Total Environment*, *786*, 147366.

WFP Office of Evaluation & Avenir Analytics. (2020). *Strategic Evaluation of Funding WFP's Work* (Evaluation Report OEV/2019/018). World Food Programme. https://docs.wfp.org/api/documents/WFP-0000116029/download/

WFP Vulnerability Analysis and Mapping. (2008). *Food Consumption Analysis: Calculation and Use of the Food Consumption Score in Food Security Analysis*. United Nations World Food Programme. https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp197216.pdf

World Food Programme. (n.d.-a). About HungerMap LIVE. *HungerMap LIVE*. Retrieved May 11, 2022, from https://hungermap.wfp.org/

World Food Programme (n.d.-b) Mission and Values. Retrieved November 15, 2021, from https://www.wfp.org/overview

World Food Programme (2018a) Vulnerability Analysis and Mapping: Food Security Analysis at the World Food Programme. https://docs.wfp.org/api/documents/WFP-0000040024/download/

World Food Programme (2018b, November 5). *WFP and Alibaba Enter Strategic Partnership to Support UN Sustainable Development Goal of a World with Zero Hunger.* https://www.wfp.org/news/wfp-and-alibaba-enter-strategic-partnership-support-un-sustainable-development-goa

World Food Programme. (2020, August 13). *2020 - Hunger Map*. https://www.wfp.org/publications/hunger-map-2020

Yohannes, Y., & Webb, P. (1999). *Classification and Regression Trees, CART*. International Food Policy Research Institute, Washington, DC.

Zheng, Y., Hatakka, M., Sahay, S., & Andersson, A. (2018). Conceptualizing development in information and communication technology for development (ICT4D). *Information Technology for Development*, *24*(1), 1–14.