TRAGEDIES of Regional Conflicts & PROMISES of Peacebuilding
Responding to Disruptors and Enablers of MENA Development Pathway
ERF 30th Annual Conference

www.erf.org.eg

A Geospatial Analysis of Food Insecurity Among Refugee Households in Lebanon Using Machine Learning Techniques

Angela C. Lyons,
Josephine Kass-Hanna,
Deepika Pingali,
Aiman Soliman,
Yifang Zhang,
David Zhu,
and Alejandro Montoya Castano

Economic Research Forum 30 years
منتدى البحوث الاقتصادية

2024

A Geospatial Analysis of Food Insecurity Among Refugee Households in Lebanon Using Machine Learning Techniques*

Angela C. Lyons, University of Illinois at Urbana-Champaign[1]
Josephine Kass-Hanna, IESEG School of Management[2]
Deepika Pingali, University of Illinois at Urbana-Champaign[3]
Aiman Soliman, University of Illinois at Urbana-Champaign[4]
Yifang Zhang, University of Illinois at Urbana-Champaign[5]
David Zhu, University of Illinois at Urbana-Champaign[6]
Alejandro Montoya Castano, International Fund for Agricultural Development (IFAD)[7]

December 20, 2023

Abstract

This study integrates geospatial analysis with machine learning to understand the interplay and spatial dependencies among various indicators of food insecurity. Specifically, we use the VASyR data on Syrian refugees in Lebanon and merge it with novel geospatial data to uncover why certain indicators of food security are successful in specific contexts, while others fall short in providing accurate insights. Our findings indicate that geolocational indicators significantly influence food insecurity, overshadowing traditional factors like household sociodemographics and living conditions. They suggest a shift in focus from labor-intensive socioeconomic surveys to readily accessible geospatial data. The study underscores the variability of food insecurity across different locations and subpopulations, challenging the effectiveness of individual measures like FCS, HDDS, and rCSI in capturing localized needs. From a policy perspective, our insights call for a refined approach to addressing food insecurity among refugees. By disaggregating the various dimensions of food insecurity and understanding their distribution, policymakers and humanitarian organizations can better tailor their strategies, directing resources to areas where refugees face the most severe challenges, thereby enhancing the effectiveness of food security measures.

JEL classification: I3, I32, O1, O53, R23, Q18

Key words: food insecurity, forced displacement, refugees, geospatial analysis, machine learning

---

* We wish to thank Anna Baskins at the University of Pittsburgh and Nishk Patel, Sona Krishnan, Ishaan Salaskar, and Sarvagnya Vijay at the University of Illinois at Urbana-Champaign for undergraduate research assistance in pre-processing and constructing the geospatial data used in this analysis. This work was funded in part by the National Center for Supercomputing Applications (NCSA) and the Office of International Programs in the College of Agricultural, Consumer and Environmental Sciences at the University of Illinois. Generous support was also provided by the USDA National Institute of Food and Agriculture, Hatch Project [1024950]. The views expressed in this work are entirely those of the authors and should not be attributed to the Economic Research Forum (ERF), its Board of Trustees, or donors.

[1] Corresponding author: Angela C. Lyons, Associate Professor, University of Illinois at Urbana-Champaign, Department of Agricultural and Consumer Economics and National Center for Supercomputing Applications (NCSA), 440 Mumford Hall, 1301 W. Gregory Drive, Urbana, IL 61801 USA. Phone: +1 (217) 418-6086. Email: anglyons@illinois.edu

[2] Josephine Kass-Hanna, Assistant Professor, Finance Department, IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie Management, F-59000, Lille, France. Email: j.kass-hanna@ieseg.fr

[3] Deepika Pingali, PhD Candidate, University of Illinois at Urbana-Champaign, Department of Agricultural and Consumer Economics, 326 Mumford Hall, 1301 W. Gregory Drive, Urbana, IL 61801 USA. Email: dsp7@illinois.edu

[4] Aiman Soliman, Research Scientist, National Center for Supercomputing Applications (NCSA) and Research Assistant Professor, Department of Urban and Regional Planning, University of Illinois at Urbana-Champaign, USA. Email: asoliman@illinois.edu

[5] Yifang Zhang, Data Analyst, National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign, USA. Email: zhang303@illinois.edu

[6] David Zhu, Undergraduate Research Assistant, National Center for Supercomputing Applications (NCSA), Students Pushing Innovation (SPIN) Internship Program. Email: yuerzhu2@illinois.edu

[7] Alejandro Montoya Castano, Research Consultant, International Fund for Agricultural Development (IFAD). Email: montoyacalejandro@gmail.com

A Geospatial Analysis of Food Insecurity Among Refugee Households in Lebanon Using Machine Learning Techniques

Abstract

This study integrates geospatial analysis with machine learning to understand the interplay and spatial dependencies among various indicators of food insecurity. Specifically, we use the VASyR data on Syrian refugees in Lebanon and merge it with novel geospatial data to uncover why certain indicators of food security are successful in specific contexts, while others fall short in providing accurate insights. Our findings indicate that geolocational indicators significantly influence food insecurity, overshadowing traditional factors like household sociodemographics and living conditions. They suggest a shift in focus from labor-intensive socioeconomic surveys to readily accessible geospatial data. The study underscores the variability of food insecurity across different locations and subpopulations, challenging the effectiveness of individual measures like FCS, HDDS, and rCSI in capturing localized needs. From a policy perspective, our insights call for a refined approach to addressing food insecurity among refugees. By disaggregating the various dimensions of food insecurity and understanding their distribution, policymakers and humanitarian organizations can better tailor their strategies, directing resources to areas where refugees face the most severe challenges, thereby enhancing the effectiveness of food security measures.

JEL classification: I3, I32, O1, O53, R23, Q18

Key words: food insecurity, forced displacement, refugees, geospatial analysis, machine learning

1. Introduction

Over 100 million people have now been forcibly displaced, a number that has more than doubled in the past decade due to recent and ongoing conflicts and crises (UNHCR, 2023). Forcibly displaced populations (FDPs) often find themselves in situations where access to adequate food and nutrition is severely limited. This is due to various factors including the loss of their livelihoods, disruption of traditional food supply chains, and the challenges inherent in adapting to new environments. Typically, FDPs rely on humanitarian aid for their basic needs. However, the capacity of international organizations to provide sufficient food aid is increasingly strained, especially with the rising numbers of refugees. The situation is further exacerbated in regions where the host communities themselves face food insecurity.

Research has attempted to understand the dynamics of food insecurity among refugee populations, including the impact of displacement on access to food, the effectiveness of humanitarian aid, and the long-term implications of nutritional deficiencies (e.g., Ghattas et al., 2014, 2015; Hadley et al., 2010; Lyons et al., 2023a, 2023b; Mansour et al., 2020). Recently, geospatial analysis has been gaining prominence in research investigating food security challenges faced by FDPs (e.g., Al Shogoor et al. 2022; Çetinkaya et al., 2016; Füreder et al., 2012; Müller et al., 2016; Lyons et al., 2023b; Younes et al., 2022) and households in general (Alemu et al., 2017; Brown, 2016; Coughlan de Perez et al., 2019; Dessie et al., 2022; Lone & Mayer, 2019; Lv et al., 2022; Mathenge et al., 2023). This approach offers valuable insights into spatial patterns and factors influencing food access among disadvantaged populations. Some studies have also used machine learning techniques to make predictions about future food insecurity (Deléglise et al., 2022; Foini et al., 2023; Lentz et al., 2019; Lyons et al., 2023b; Martini et al., 2022; Meerza et al., 2021).

This body of research is vital in guiding policy decisions and humanitarian efforts to effectively address the FDPs' food needs and mitigate the challenges they face. Our study adds to this field by integrating geospatial analysis with machine learning to understand the interplay and spatial dependencies among various indicators of food insecurity. The aim is to assist in formulating more targeted and effective policy recommendations for cash assistance programs. Specifically, we use data collected from Syrian refugees in Lebanon to uncover why certain indicators of food security are successful in specific contexts, while others fall short in providing accurate insights.

Lebanon presents a compelling case for examining the connection between forced displacement and food insecurity. This is not only due to its status as the country with the highest per capita refugee population globally but also because of the recent crises that hit the country and compromised the food security of both Lebanese residents and Syrian refugees. The financial and economic crisis that began unfolding in Lebanon in late 2019 has significantly heightened the vulnerability and poverty levels[8] (ESCWA, 2021). As Lebanon relies on imports for most of its food and non-food needs, the sharp currency depreciation has strained the country's capacity to pay for its imports resulting in soaring inflation and eroding households' purchasing power. The challenges intensified with the advent of the COVID-19 pandemic, and the situation reached a critical point with the devastating Beirut port explosion in August 2020, which limited the country's import capacity. These crises led to a dramatic decline in the overall well-being of the country's population, exacerbating an already fragile situation for many Lebanese and refugee households. The lifting of state subsidies on medicine and energy coupled with record-high inflation and escalating international prices, have severely impacted the ability of Lebanese and refugee households to fulfill their basic needs.

Fundamentally, the crises affected the availability of food in the country. Despite large cultivable land per capita, agricultural productivity is constrained and the country's food supply heavily depends on imports, with estimates suggesting that the country imports about of 80% of its agricultural goods[9] (International Trade Administration, 2022). Between December 2019 and October 2021, the Consumer Price Index (CPI) and the Food Price Index witnessed alarming increases of 519% and 1874%, respectively (CAS, n.d.). The cost of the Survival Minimum Expenditure Basket (SMEB), a measure of basic food and non-food necessities, also significantly surged, with the food component increasing 11-fold between October 2019 and December 2021, and 21-fold by December 2022 (WFP, 2021, 2022a).

[8] An analysis by ESCWA (2021) based on the 2018-2019 Labour Force and Household Living Conditions Survey in Lebanon estimated that the multidimensional poverty rate in Lebanon has doubled from 42% in 2019 to 82% in 2021.
[9] The escalation of the Russo-Ukraine conflict has exacerbated food availability issues in Lebanon, where 80% of its wheat is sourced from these countries (World Bank, 2023). Recent analysis indicates that a sharp increase in currency circulation led to a decrease in food imports between 2019 and 2021.

According to WFP (2022b), food insecurity affected 30% of Lebanon's population at the end of 2020, worsening between June and December 2021, with the percentage of food insecure families nearing 50%. Recent data reveals that more than 35% of residents in Lebanon faced crisis-level food insecurity in the last quarter of 2022[10] (IPC, 2022a, 2022b). The crisis has hit Syrian refugees particularly hard. Among this population, food insecurity rates increased by 21 percentage points between 2019 and 2020, reaching 67% by 2022 (UNHCR et al., 2023). North Lebanon and Akkar governorates experienced the highest rates of food insecurity (79% for both), followed by Bekaa and Baalbeck-El Hermel (75% and 72% respectively). In terms of severe food insecurity, North Lebanon and Akkar also had the highest proportions (10% and 9% respectively). The food insecurity rates among Syrian refugees show different patterns from those of Lebanese households, as evidenced by a survey conducted between November 2020 and March 2021 (Hoteit et al., 2021). The highest incidence of food insecurity was recorded in Bekaa, where 83% of Lebanese households were estimated to have a poor Food Consumption Score (FCS), followed by Akkar at 73%, and North Lebanon at 58% (Hoteit et al., 2021).

Our previous research on poverty among Syrian refugees in Lebanon, which employed machine learning techniques, revealed similar intriguing patterns in food security among Syrian refugees. We found that the highest levels of food insecurity among refugees were not necessarily concentrated in the poorest localities (Lyons et al., 2023b). Further exploration of the heterogeneities across different geographic regions within Lebanon revealed certain economic dynamics that partially explained these differing patterns. In the Bekaa region, a major agricultural area, refugee households are likely to have some food security given their engagement in agriculture. In fact, it is estimated that agriculture serves as the main source of income for 90% of refugee households (Al Zoubi et al. 2019). In contrast, North Lebanon's urban and mountainous landscape offers fewer agricultural employment opportunities, resulting in higher food insecurity among refugees.

A deeper understanding of food insecurity among refugees in Lebanon necessitates further examination of the various indicators of food security and the complex interconnections between them and across different regions and refugee populations. By examining the spatial patterns of food insecurity, we aim in this study to shed light on the spatial disparities, exploring how and why certain indicators vary between areas and how and why these indicators exhibit variations within specific regions. This analysis is crucial for identifying the unique challenges faced by different refugee communities and for developing tailored strategies to address them.

The remainder of this paper is structured as follows. The next section discusses the relevant literature and contributions of this study. Section 3 provides an overview of the data, outlining our food insecurity metrics

[10] The first Lebanon IPC Acute Food Insecurity Analysis indicated that, between September and December 2022, 1.6 million people were classified in IPC Phase 3 (Crisis) requiring urgent humanitarian action to reduce food gaps and protect and restore livelihoods. Additional 306,000 people were in IPC Phase 4 (Emergency) experiencing acute malnutrition and excess mortality due to food insecurity (IPC, 2022b).

and the predictive features categorized into household sociodemographics, living conditions, and geospatial features. Section 4 explores the spatial analysis of our food insecurity measures, examining their evolution across time and regions and identifying global spatial associations among the indicators at the district level. Section 5 describes our methodology, which employs machine learning (ML) techniques to predict food insecurity across all refugee households. Our findings are presented and discussed in Section 6. Section 7 concludes by highlighting important policy implications.

2. Literature Review

2.1 Measuring food insecurity: dimensions and metrics

Previous research has extensively addressed the question of assessing food security, delving into the comparability and effectiveness of various measures. According to the *Rome Declaration on World Food Security* adopted at the 1996 World Food Summit, ‘‘Food security exists when all people, at all times, have physical and economic access to sufficient, safe, and nutritious food that meets their dietary needs and food preferences for an active and healthy life” (FAO, n.d.). This definition, widely adopted by international organizations and researchers, emphasizes the multi-dimensional nature of food security, which makes it difficult to measure. There is still no consensus on an ideal measurement method nor on the best indicators to capture food security. The choice of indicators or combination of indicators is often influenced by the specific context, objectives of the study, and available data, leading to various approaches in assessing food security. However, key dimensions have been commonly recognized. The FAO's annual *State of Food Insecurity in the World (SOFI)* reports traditionally rely on four dimensions – also known as the four pillars: availability, access, utilization, and stability (FAO, IFAD and WFP, 2015). Other frameworks such as the Global Food Security Index (GFSI) of the Economist Intelligence Unit (EIU) focus on affordability, availability, quality and safety, and sustainability and adaptation. These nuanced frameworks, while they broadly align, reflect a variety of conceptualizations of food security aspects and approaches to analyze and monitor food security.

Manikas et al. (2023) discuss the use of different indicators by international agencies, providing a summary of commonly applied indicators according to food security dimensions and level of analysis (i.e., individual, household, national). The authors indicate that to reflect the 1996 World Food Summit definition, an ideal food security indicator should capture all dimensions at the individual level. However, most studies adopting such a comprehensive approach tend to focus on the national level. For instance, Caccavale and Giuffrida (2020) developed the Proteus Composite Index (PCI) to measure food security in 185 countries considering all the four pillars of food security. Others have relied on the GFSI dimensions for measuring food security at national level as well (Chen et al, 2019; Izraelov, & Silber, 2019).

In their systematic literature review, Manikas et al. (2023) note that most of the available food security indicators focus on household-level measures of a single dimension – food access. Access reflects the demand side of food security and is most closely aligned with social science concepts of individual or household well-being (Barrett, 2010). Common household-level measures of food access used by international agencies include, among others: the Food Consumption Score (FCS) developed by the WFP (2008) which is the weighted sum of food groups consumed, considering their frequency and nutritional value; the Household Dietary Diversity Score (HDDS) promoted by the FAO which measures the number of different food groups consumed by a household over a specified period, usually 24 hours, reflecting access to a variety of foods (Swindale & Bilinsky, 2006); the Coping Strategies Index (CSI) and reduced CSI (rCSI) which capture the various coping strategies that households employ in response to food insecurity, such as reducing meal frequency or borrowing food.

While stand-alone food security indicators are valuable for specific assessments, they often provide a narrow view, focusing on just one aspect of food security. This can result in an incomplete understanding, as they fail to capture the dynamic and interrelated elements of food security. Researchers and organizations have developed composite indices to enable a more comprehensive assessment of food security (Biederlack & Rivers, 2009; Maione et al., 2019; Mathenge et al., 2023; Reig, 2012; Santeramo, 2015; Vaitla et al., 2017; Wineman, 2016).

Several empirical studies were conducted to compare various household indicators and assess their performance in measuring food security. For example, Maxwell et al. (2013) used household data from Ethiopia to compare food security indicators. They assess inter-correlations among seven indicators and analyze whether they detect the same or different aspects of food insecurity. Among the findings, the study noted that FCS and HDDS are inclined to capture *quality* and *diversity*, while CSI and rCSI tend to reflect elements of *quantity* or *sufficiency*. The authors noted that these indicators are often used interchangeably, without a clear idea of which aspect of food security is captured, which can lead to potential misclassification of food insecure populations. Maxwell et al. (2014) highlighted the importance of using more than one indicator, as food security indicators differ in the underlying aspect they attempt to capture.

Similar results were found by Vaitla et al. (2017) who attempted to quantify the unique information provided by four indicators: FCS, HDDS, rCSI, and Household Hunger Scale (HHS). Using exploratory factor analysis (EFA), the study found that the selected indicators capture two distinct underlying latent dimensions related to food security, which exhibit only weak correlations with each other. The first dimension correlates strongly with HHS and rCSI, and might capture the *quantity* of food consumption or the costs of constrained access to food. The second was found to correlate with FCS and HDDS and might represent the *diversity* of dietary intake. The findings lend empirical support to the argument that multidimensional constructs like food security are best captured using a set of indicators representing key dimensions. Relying on a single composite indicator can obscure the distinct contributions and implications of each aspect (Mathenge et al., 2023; Wineman, 2016).

Recognizing the limitations inherent in both stand-alone indicators and composite indices, researchers are increasingly recommending the use of multiple indicators. This approach enhances the depth and breadth of understanding in food security assessments, helping to better address the complexities and varied dimensions that single measures might miss (e.g., Coates, 2013; Maxwell et al., 2014; Mathenge et al., 2023; Vaitla et al., 2017).

2.2 Predictive analytics in food security

Recent studies have increasingly integrated geospatial data in assessing and predicting food insecurity. By leveraging spatial data, researchers and policymakers can identify and analyze patterns and trends in food insecurity across different regions and over time. Common sources of geospatial data used in these studies include satellite imagery (Füreder et al., 2012; Giada et al., 2003), climate data (Coughlan de Perez et al., 2019; Lv et al., 2022; Demeke et al., 2011; Mathenge et al., 2023), land use patterns (Brown, 2016; Lone & Mayer, 2019), and sociodemographic information (Alemu et al., 2017; Dessie et al., 2022; Lyons et al., 2023b). Advanced techniques like remote sensing and Geographic Information Systems (GIS) enable the mapping of agricultural productivity, the assessment of natural resource availability, and the monitoring of environmental factors like droughts or floods that can impact food availability and access. Predictive models using geospatial data have been particularly valuable in early warning systems, helping to anticipate food crises before they occur and enabling timely and targeted interventions. This geospatial approach has gained momentum as it offers a dynamic tool for understanding the spatial distribution of food insecurity, providing crucial insights for effective policymaking and resource allocation.

Geospatial data is increasingly being used to understand the food security challenges faced by refugees and forcibly displaced populations (FDPs). Research in this field often includes site selection analysis for refugee camps (Çetinkaya et al., 2016; Younes et al., 2022), land use analysis (Al Shogoor et al., 2022; Müller et al., 2016) and camps' monitoring (Füreder et al., 2012; Giada et al., 2003).

In the specific context of Lebanon, geospatial analysis remains relatively limited and sparse in some areas such as climate change impact assessment, agricultural planning, and disaster risk reduction (Caiserman & Faour, 2021; Der Sarkissian et al., 2019; Ghoussein et al., 2018; Issa et al., 2014). To our knowledge, few if any studies have used geospatial analysis to assess food security among FDPs in Lebanon. A notable exception is our prior research, where we used geospatial indicators as predictors of multidimensional poverty measured based on expenditures and food security (FCS and rCSI). Specifically, we combined geospatial data with survey data on Syrian refugees in Lebanon and applied machine learning (ML) techniques to predict which households are more likely to be classified as poor.

Machine learning methods have become increasingly popular in making predictions about future food insecurity, enhancing the effectiveness of geospatial analysis in this field (Meerza et al., 2021). For example, Deléglise et al. (2022) developed the Food Security Prediction based on Heterogeneous Data ($FSPHD$) framework using machine and deep learning models to estimate FCS and HDDS from public data in Burkina Faso. Similarly, Lentz et al. (2019) predicted FCS, HDDS, along with rCSI in Malawi, using diverse data sources like meteorology, precipitation, market prices, and soil quality. Taking a broader scope, Martini et al. (2022) applied a similar approach relying on secondary data, when primary data were not available. Their "nowcasting" predictive models utilized sub-national-level data on FCS and rCSI data, from 78 and 41 countries respectively, to estimate the prevalence of food consumption insufficiency and crisis or above-crisis food-based coping. Among these efforts, Foini et al. (2023) developed a model using gradient boosted regression trees to forecast short-term food consumption trends in six countries, incorporating data on conflict, weather events, and economic shocks.

This study aims to contribute to this line of work that harnesses advanced computational techniques to show that place-specific features and geographic specificities have a significant impact on food security outcomes. We combine machine learning techniques with geospatial analysis to gain insights into why certain measures of food security are effective indicators in specific contexts while others fail to provide accurate insights. Understanding the interconnectedness of various measures of food insecurity and their spatial dependencies can assist in producing more effective policy recommendations for targeting cash and non-cash interventions to alleviate rising levels of food insecurity among Syrian refugees in Lebanon.

3. Data

We use survey data taken from the *Vulnerability Assessment of Syrian Refugees (VASyR)* jointly gathered by the UNHCR, WFP, and UNICEF for the years 2018, 2019, 2020, 2021, and 2022. The *VASyR* is a nationally representative survey of Syrian refugee households in Lebanon that includes detailed information on: (1) individual and household demographics, including work and schooling; (2) shelter, utility, sanitation, and settlement conditions; (3) income, expenditures, assets and debts; (4) food consumption and dietary diversity; (5) health and safety; and (6) coping strategies (UNHCR et al., 2018, 2019, 2021, 2022, 2023).[11]

Our analysis is conducted at the household level. The initial sample size included 23,609 refugee households for all five survey years (4,444 in 2018, 4,687 in 2019, 4,506 in 2020, 4,968 in 2021 and in 5004 2022). Households with missing information for the key features related to sociodemographic characteristics and living conditions were excluded from the sample. The final sample used for our analyses consists of 22,626 refugee households (4,433 in 2018, 4,670 in 2019, 4,480 in 2020, 4,967 in 2021, and 4,076 in 2022).

In this paper, we aim to understand the relationships between different measures of food insecurity and identify which metrics are more prevalent in certain regions and which socioeconomic, environmental, and geographic characteristics are more likely to explain the different indicators and regional differences. The

[11] In each survey year, data were collected from Syrian refugee households who were randomly selected from the 26 administrative districts across the eight governorates of Lebanon. To ensure representativeness at the district and governorate levels, sampling was based on a two-stage cluster approach whereby clusters (villages, neighborhoods, or towns) were selected within each district, and then refugee cases were randomly selected within each cluster. Specifically, probability proportionate to size (PPS) methodology was used, where clusters with larger concentrations of refugees were more likely to be selected. Weights were also constructed at the district level based on the refugee population in each district. See UNHCR et al. (2018, 2019, 2021, 2022, 2023) for more details about the sampling and survey methodology.

following presents our food insecurity metrics, and the features used to predict these, which can be broadly grouped into three categories: household sociodemographics, living conditions, and geospatial features.

3.1. Measuring food insecurity

The VASyR data include a rich set of information on food insecurity. We use these data to construct and analyze four measures of food insecurity commonly used in the literature: the food consumption score (FCS) and the household dietary diversity score (HDDS), as metrics of quality and dietary diversity; the reduced food coping strategies index (rCSI), as an indicator of quantity as it reflects the strategies that households use to deal with the lack of food; and households' share of food expenditures relative to their total expenditures (share of food expenditures). See Table A1 in the Appendix for a complete list of the food insecurity measures and feature variables and how they were specifically defined and constructed.

3.2. Household sociodemographics and living conditions

The sociodemographics included in our study account for a household's family structure in terms of its household size, dependency ratio, proportion of female-headed and single-parent households, and the share of household members by age, gender, education, employment status, health and disability, and legal residency. Also included are variables that capture household's living standards, as they pertain to basic access to electricity, sanitation, clean drinking water, cooking fuel, and shelter.

3.3 Geospatial features

Geographically, Lebanon is divided into 8 governorates, which can be further subdivided into 26 districts, which can be even further subdivided into "cadastres," which are the smallest administrative unit.[12] Figure 1 presents mappings of Lebanon; the left-side shows a map of Lebanon's districts while the right-side visually highlights its topographical features. Lebanon's geography is marked by a blend of coastal and mountainous landscapes with more than half of its terrain situated above 1,000 meters. The country features a narrow coastal plain along the Mediterranean Sea and two major north-south mountain ranges, and in between is the fertile Bekaa Valley (including the Bekaa and Baalbeck-El Hermel governorates). This diverse geography results in distinct climatic zones and land use patterns, even within small areas. In the western governorates, districts commonly span from coastal areas to high altitudes, encompassing both sea-level cities and elevated mountainous regions. This topographical variety significantly influences Lebanon's socio-economic dynamics. The coastal cities, especially the capital Beirut and suburbs in Mount Lebanon, are economic centers for commerce, banking, and tourism, while the Bekaa Valley is the heartland of agriculture, known for producing

[12] Note the governorate of Beirut (the capital city of Lebanon) is not subdivided into districts. Also, the Akkar governorate is comprised of a single district. Cadastres in Lebanon are equivalent to Administrative Level 3 in terms of the UN Geospatial Information Section & Statistics Division's 'Second Administrative Level Boundaries' programme (UN Habitat Lebanon & ESCWA, 2021).

crops like grapes, vegetables, and grains. On the other hand, the mountainous regions traditionally support livestock rearing and cultivation of olives and fruits. Lebanon's economy is also known for its predominant services sector.

[Insert Figure 1]

Development is not uniform across Lebanon. Peripheral areas, in particular, face significant development challenges and weak infrastructure (e.g., electricity, road networks, waste management, water supply). This infrastructure gap exacerbates regional inequalities, leading to disparities in employment opportunities, income, and the overall well-being of the population[13]. Such uneven development creates a stark contrast between these areas and the more developed regions like Beirut and its suburbs.

A novel and quintessential aspect of this project is that we include a comprehensive set of geospatial features, which until recently, have rarely been included to explain food insecurity among FDPs. And yet, location matters such that geographical characteristics are likely to be significant predictors of food insecurity, and moreover, their importance is expected to vary across locations. These geospatial attributes include land elevation, latitude and longitude, types of land coverage (built-up area, crop area, and permanent and seasonal water area), quality of vegetation, access to roadways and waterways, density of refugee and host populations, distance to nearest Syrian border, and incidence of conflicts.

The extraction of the geospatial attributes was first conducted using the cadastral boundaries. However, to protect the safety and security of the refugees, the VASyR survey data only reports the district in which a refugee household resides and not the cadastre. In order to merge the VASyR data with the geographical feature data, we needed to aggregate the cadastral level values for each geographical feature at the district level using the geographic boundaries for the 26 districts in Lebanon. See Table A1 in the Appendix for a detailed description of the geospatial features and how they were constructed (see also Lyons et al., 2023b for more information on the source and availability of each feature).

4. Spatial Analysis of Food Insecurity Measures

4.1 Changes in food insecurity over time

Table 1 presents the descriptive statistics for the food insecurity measures by survey year; p-values are reported to identify which metrics differed significantly across the years[14]. The mean values for our key food insecurity variables were found to significantly vary across the years. In general, total expenditures per capita

[13] The most recent data on the distribution of multidimensional poverty based on a 2018/2019 budget survey of Lebanese households reveal that those living in extreme poverty represented close to 50% of the population in Akkar and Baalbeck-El Hermel and Nabatieh governorates, 43% in Bekaa, 35% in South Lebanon, and 33% in North Lebanon, compared to 29% in Beirut and 27% in Mount Lebanon (ESCWA, 2021).

[14] The p-values were calculated using t-tests by category when the variables were continuous. These p-values were adjusted for multiple pairwise comparisons following the Benjamini-Hochberg method using the R package "compareGroups." When the variables were categorical, the p-values were based on a chi-square test.

more than tripled between 2018 and 2021 and more than quadrupled between 2021 and 2022, due to severe currency depreciation and hyperinflation.[15] Similarly, we observe an even more pronounced surge in food expenditures per capita, with the share of total expenditures spent on food rising from 41% in 2018 to 58% in 2022. The percentage of refugee households with expenditures below the Survival Minimum Expenditure Basket (SMEB) went from 50% in 2018 to nearly 100% by 2022. While this measure alone provides an indication of the growing and stark rise in food insecurity among the refugee households, additional evidence also points to rising food insecurity levels. Food consumption scores and diet diversity scores have worsened, with FCS and HDDS decreasing by 18% and 8%, respectively (FCS: from 53.3 in 2018 to 43.7 in 2022; HDDS: from 9.1 in 2018 to 8.4 in 2022). Refugee households' use of reduced food coping strategies (rCSI) increased between 2018 and 2022 as well. More specifically, we see increases across the board in the percentage of households who reduced the number of meals per day, reduced portion sizes of meals, relied on less preferred or less expensive food, borrowed food or relied on help from friends or relatives, and restricted consumption by adults in order for young-small children to eat. Thus, all indicators denote a dramatic rise in food insecurity, which is not surprising given recent events and reports by humanitarian agencies (e.g., UNHCR et al., 2021, 2022, 2023).

[Insert Table 1]

4.2 Geospatial differences in food insecurity

The mappings presented in Figures 2 and 3 highlight several significant geospatial differences in food insecurity across the districts within Lebanon and over time. Note that darker shading indicates higher levels of food insecurity, and hence, worsening conditions. For the various measures of food insecurity, there are considerable geographical heterogeneities. In Figure 2, we find that, while FCS, HDDS, and rCSI worsened in general for refugee households across Lebanon, there were specific geographical areas that fared worse than others. For example, FCS worsened in the southern districts between 2018 and 2020, then improved and worsened again between 2021 and 2022. Aside from the southern districts, we also observed a worsening trend in FCS in the west, and especially in the northwestern districts, which was offset by worsening conditions in the east between 2021 and 2022. Similar trends were also observed for HDDS between the east and west as well as for northwest and southern districts. In terms of rCSI, food insecurity in the west was consistently worse than in the east. See Figure A1 in the Appendix for a breakdown and comparison of the five subcomponents of rCSI.

[Insert Figure 2]

Figure 3 presents the mappings for changes in share of food expenditures and share of refugee households below the SMEB. In general, we find that the share of total expenditures refugee households spent on food

[15] Lebanon's local currency has lost more than 95% of its value, driving inflation to triple digits since July 2020 and impacting mostly the poor and vulnerable (World Bank, 2022).

increased over time for almost all districts. In particular, refugee households located in the northern-most district of Akkar and the southern-most district of Bint Jbeil experienced the largest increases in the share of food expenditures, reaching over 60% in 2022. Additionally, with the onset of the economic crises and hyperinflation, the percentage of households below the SMEB also increased, with a noticeably higher increase in the western districts compared to the eastern districts (this percentage was high in the east since 2018). From 2020 onwards, almost all refugee households were below the SMEB, regardless of the district. The exception was the Matn district adjacent to Beirut, where fewer households were below the SMEB. However, in 2022, even in Matn, almost all households were below the cutoff.

[Insert Figure 3]

4.3 Testing for spatial dependencies

We calculated a Global Moran's I to evaluate the existence of a global spatial association among the food insecurity indicators at the district level.[16] The existence of a global spatial structure indicates the degree to which food insecurity in one district is similar to food insecurity in the immediate neighboring districts, indicating that the underlying factors impacting food security are anchored to that particular geographical area. The results of the Global Moran's I were tested against the null distribution which represented complete spatial randomness, where the food insecurity scores were not correlated for neighboring districts (Anselin, 1995).[17] The null distribution was estimated numerically using conditional permutation of the food insecurity scores.

The results of the Global Moran's I (GMI) are presented in Table 2. As expected, we find evidence of spatial correlation, confirming the descriptive findings presented in Figures 2 and 3. However, the findings vary significantly across indicators and across time for specific indicators. The scores for the FCS show moderate but significant spatial association, and this spatial association is centered around the southern districts, which in Figure 2 appear to have lower levels of food insecurity (higher FCS). Although this pattern seems to be persistent, it changed in the year 2020, where more spatial randomness was observed in the geographical arrangement of the scores for FCS. These results were confirmed by the Local Moran's I. See Figure 4 which

[16] Spatial dependence pertains to the extent of spatial autocorrelation among independently measured values within a geographic area. Global indicators of spatial autocorrelation, such as Moran's I, provide a singular measure of spatial dependence. In parallel, local indicators of spatial association, such as LISA, serve a comparable purpose but yield multiple location-specific measures of spatial dependence, enabling researchers to examine the spatial variation in dependence across different locations.

[17] Previous research has almost always assumed spatial independence prevails such that food insecurity and poverty among FDPs tends to be randomly distributed over geographical units – in our case, districts. However, it is reasonable to expect that food insecurity may not be spatially independent such that levels of food insecurity among FDPs within neighboring districts may be significantly related to those within a particular district, especially given the geospatial heterogeneities across Lebanon that can impact availability, access, and utilization of food.

shows a significant hotspot of districts in the south (mostly in the Nabatieh governorate) with high FCS scores, particularly for years 2019 and 2022.[18]

In comparison, the scores for HDDS tended to lack a global spatial structure. In 2018, the Global Moran's I value for HDDS was approaching 0 (complete spatial randomness) with a p-value of only 0.301. In fact, no clear trend for HDDS was observed over the five-year period, with the exception of 2019, where we found a cluster of HDDS high scores in the southern districts of Lebanon, specifically in Nabatieh.

Conversely, the scores for the rCSI show a clear and significant spatial structure, characterized by higher rCSI scores in the northwest districts compared to lower rCSI scores in the southeast districts. The spatial pattern appears to be persistent over time, and breaks only in year 2022, where the differences between the northwest and southeast districts become less distinct. Analysis of the Local Moran's I supports the persistent pattern revealed by the global test results.

[Insert Table 2]

With regards to the share of food expenditures, the scores showed a global spatial association that is likely high and radiates from the north. The results were further tested using the Local Moran's I in Figure 4 and the share of food expenditures was high (more food insecure) in the districts of Bsharri (North Lebanon governorate) and El Hermel (Baalbeck-El Hermel governorate).

Finally, the percentage of refugee households below the SMEB also revealed the existence of a spatial structure, although it was less significant than the geographic patterns produced by mapping the rCSI scores. Still, the spatial patterns indicated a contrast between the high percentages below the SMEB in the northern and eastern districts compared to those in the southern and western districts. We also observe that spatial trends in the percentage below the SMEB diminish over time, as indicated by the decrease in spatial correlation and subsequent increase in spatial homogeneity and its insignificance over time, especially in year 2022. This result again confirms our earlier descriptive finding that almost all refugee households were below the SMEB by 2022.

[Insert Figure 4]

Overall, the descriptive analysis in this section underscores the presence of spatial differences in food insecurity among Lebanon's districts. This observation, supported by the outcomes of the Global and Local

---

[18] In Figure 4, gray areas indicate districts where there is spatial randomness and no evidence of a spatial structure. Blue indicates "cold spots" and districts where food insecurity is significantly low, whereas red indicates "hot spots" where food insecurity is significantly high such that the spatial arrangement of food insecurity within that district is unlikely due to chance. Cyan represents outlier districts with a low score surrounded by relatively high scoring neighbors, and vice versa, orange indicates outlier districts with a high score surrounded by relatively low scoring neighbors.

Moran's I analyses, reinforces the notion of substantial spatial dependence, particularly evident in measures like FCS and rCSI. Consequently, it is plausible to posit that these spatial variations in food insecurity stem from diversities in geospatial features across Lebanon and within specific districts.

5. Methodology

Crucial questions that arise from the descriptive findings include: What are these geospatial features, and how do they contribute to food insecurity? Furthermore, can these geospatial features elucidate the observed heterogeneities, especially those between eastern and western Lebanon, and in the northwest and southeast districts? To provide insight into these questions, we predicted food insecurity for all refugee households using machine learning (ML) techniques, pooling the observations for all five survey years. In our analysis, we strictly followed the machine learning process including feature selection, model selection, and spatial analysis (Han, Kamber, & Pei, 2012).

5.1. Feature selection

The features incorporated into our food insecurity models, as detailed in the data section, encompass three broad categories: household sociodemographics characteristics, living conditions, and geospatial features. The selection of features aligns with the methodology outlined by Han, Kamber, and Pei (2012), emphasizing the importance of avoiding redundancy in models caused by an excessive number of features, which can hinder effective model learning. To mitigate such distractions, a judicious feature selection process was employed. In this context, we employed an evidence-based approach, drawing on existing literature and considering socioeconomic, environmental, and geographical conditions. The chosen features were deemed most relevant to the specific context of refugees within Lebanon, guided by insights from relevant studies such as Lyons et al. (2023) and Lyons et al. (2022). This approach ensures that the selected features are not only statistically significant but also contextually meaningful in capturing the dynamics of food insecurity in the Lebanese context, particularly concerning refugee populations.

5.2. Model selection

Our goal was to develop a high-performing machine learning (ML) model while safeguarding against overfitting. We assessed the performance of three ML models: Random Forest (RF), Gradient Boosting (GB), and Lasso Regression (Lasso). The models were trained to predict food insecurity by classifying whether a refugee household fell within the bottom 30% for each measure. The cutoff values for the bottom 30%, derived from the pooled sample, were as follows: FCS ≤ 37; HDDS ≤ 8; rCSI ≥ 24; and Share of food expenditures $\geq$ 60.4%. based on the distance from the poorest five percent.

To mitigate overfitting, we used a repeated K-fold (K=5) cross-validation strategy to assess model performance.[19] The data were divided into five equal folds, with the models trained on four partitions and tested on the remaining one. This process was repeated five times. Model parameter tuning and performance evaluation were based on the recall score, measuring the model's ability to identify positive instances among all true positives. We prioritized this metric to maximize the inclusion of districts with severe food insecurity in our predictions. The implementation of feature selection, model fitting, and cross-validation was carried out in Python using the "sklearn" package (Pedregosa, et al, 2011).

6. Results

6.1 Model performance

We conducted a comparative analysis of the predictive power of the three machine learning (ML) models – Random Forest, Gradient Boosting, and Lasso. Table 3 presents the percentage of accurate predictions by each model identifying households in the bottom 30% for specific food insecurity measures. To enhance model accuracy, we applied balanced class weights to account for the skewed distribution resulting from classifying households as either 0 or 1 based on food insecurity. Each ML model incorporated all the feature variables encompassing household sociodemographics characteristics, living conditions, and geospatial features.

The Random Forest (RF) and Gradient Boosting (GB) models generated the highest accuracy, with comparable results. GB outperformed RF for three out of the four measures, except for the share of food expenditures, where RF exhibited slightly superior results (60.7% for RF compared to 57.9% for GB). RF displayed consistent performance across all food insecurity measures, ranging from 62.1% for FCS to 67.7% for rCSI, outperforming GB in terms of overall consistency. c

In contrast, Lasso exhibited subpar performance, which is consistent with expectations. Random Forests tend to outperform individual decision trees like Lasso due to their ability to combine multiple trees, reducing overfitting and enhancing predictive performance and robustness. The choice of RF as the baseline model was motivated by its common usage in ML algorithms, flexibility, ease of tuning, and reduced susceptibility to overfitting. Additionally, RF tends to perform well when models involve multiple 0/1 features or a diverse range of features based on different scales and value ranges.

The RF method served as the baseline model for refining the calibration of the food insecurity models and determining feature importance. As a robustness check, the results obtained from the RF models were compared to those from the GB models, ensuring a comprehensive evaluation of model performance.

[Insert Table 3]

[19] Overfitting happens when the model memorizes the training dataset and performs well in terms of goodness of fit. However, model quality degrades when applied to external data other than the training dataset (e.g., an out of sample testing dataset). The solution for overfitting is using cross-validation, which is a resampling method that uses different portions of the data to test and train a model on different iterations (Han, Kamber, & Pei, 2012).

6.2 Feature importance

Figure 5 presents the results regarding the feature importance for the three broad categories of features. Three key findings are worth noting. First, geographical features emerged as the most important predictors for three of the four food insecurity measures (FCS, rCSI, and the share of food expenditures). Second, and of notable significance, geographical features played a decisive role in predicting over 50% of the models for FCS and the share of food expenditures, and over 90% for the rCSI model. Intriguingly, household demographics and living conditions exhibited the least importance as predictors for rCSI when compared to the other food insecurity metrics. Third, although household demographics were identified as the most important predictors for HDDS, the geographical features were almost equally as important, geographical features held nearly equal importance (approximately 50% compared to about 40%, respectively). Basically, geographical features emerged as the most crucial predictors of food insecurity. Moreover, rCSI was notably more influenced by geospatial feature data, while HDDS leaned more towards being influenced by household demographics.

[Insert Figure 5]

The results for the GB models can be found in Figure A2 in the Appendix. Despite similarities with the RF results, discernible differences were observed. For instance, geographical features exhibited heightened importance for the rCSI model using GB. The predictive importance of the household demographics and living conditions were negligible. In the case of FCS and the share of food expenditures, geographical features slightly diminished in importance, with living conditions becoming more important than household characteristics. Notably, geographical features became the most important predictors of HDDS, contrasting with the RF method where household characteristics were the most important predictors.

Figure 6 displays the results detailing the top 20 most important features. There are several common predictors across the four measures of food insecurity. Among the geospatial features, population density, the number of refugees in the district, and NDVI (indicating the health of vegetation) consistently ranked among the most important features. In fact, population density was among the top five predictors for all four models. Additionally, household characteristics such as lack of electricity, the share of non-working household members, and the ratio of dependent to total household members emerged as important predictors for all four measures of food insecurity. However, these were the only sociodemographic and living condition features that were consistently among the top predictors.

For three of the four measures (FCS, rCSI, and share of food expenditures), several other geospatial features were identified as key predictors. These features predominantly pertained to the physical landscape and land coverage within a district, encompassing crop and seasonal water coverage areas, river zones, and incoming roadways. Other noteworthy geospatial features were linked to more general types of locational indicators, namely latitude and longitude and distance to the nearest Syrian border. Permanent water coverage area and the total number of conflict-related deaths for each district were notably influential predictors for FC and rCSI.

A particularly interesting finding to note is the distinct difference in the top 20 predictors for HDDS compared to the other three food insecurity measures. Unlike the other measures, HDDS is less linked to the "external" physical location and more tied to the "internal" household characteristics. The majority of the top predictors for HDDS (15 out of the top 20) were related to the structure and sociodemographic composition of the household. These factors include the share of household members with various age and education levels, as well as their employment status, disability, or health condition. In contrast, only one of the top 20 features for FCS, two for rCSI, and three for the share of food expenditures were related to household demographics. Given that HDDS serves as a micro-level indicator of food usage, offering real-time reporting of the household's consumption of various food groups in the past 24 hours, it may not be surprising that its predictors align more closely with internal sociodemographic features than external locational features. This distinction is worth mentioning, considering that the other measures are more broadly oriented towards assessing food access and availability over longer periods of time, such as the past week or year.

[Insert Figure 6]

Figure A3 in the Appendix presents the top 20 most important features for food insecurity using the Gradient Boosting (GB) method. Broadly, the findings align with those derived from RF method. Although the ranking of top predictors may vary, the overall consistency in the selection of top predictors is evident. In particular, the top five predictors for each model generally remain the same when comparing the GB and RF results for each food insecurity measure. The most noticeable difference lies in the increased importance of sociodemographic features related to the age, education, and employment status of the household members as predictors of FCS and the share of food expenditures. In contrast, geospatial characteristics and living conditions assume greater importance for HDDS. Interestingly, a robust association between rCSI and geospatial features persists. While fewer geospatial features appear in the top 20 predictors using the GB method compared to the RF method (14 versus 18), the dominance of geospatial relevance for rCSI remains evident. While fewer geospatial features were listed in the top 20 predictors, we still observed that 14 of the top 20 predictors were geospatially related using the GB method, compared to 18 of the top 20 predictors using the RF method. Of the remaining six features for rCSI using GB, three are associated with living conditions (overcrowding, safety and security, and shelter conditions), while the remaining three are tied to sociodemographic factors (legal residency status, dependency ratio, and employment status of household members).

7. Conclusions

In this study we developed a foundational understanding of the interconnections and relationships among various measures of food security in Lebanon over the period from 2018 to 2022. Our focus was on Syrian refugees, utilizing household survey data collected by the UNHCR, WFP, and UNICEF. The study incorporated four common measures of food security: the food consumption score (FCS), household diet diversity score (HDDS), reduced food coping strategies index (rCSI), and food consumption expenditures per capita.

A novel aspect of our project was the integration of household survey data with geolocational features to predict and analyze food insecurity. The analysis was conducted in three phases, utilizing geospatial analysis and machine learning techniques. First, our descriptive analysis revealed significant and dramatic increases in food insecurity across all measures, which is not surprising given recent events and reports (Lyons et al., 2023a, 2023b). More interestingly, distinct spatial variations were also observed, particularly in the northwest and southeast districts, where agricultural suitability and employment opportunities for refugees are likely to be more limited. Food insecurity was also found to be more prevalent in less developed and less urbanized districts. However, the degree of food insecurity across districts and over time was found to vary considerably depending on which measure was used.

Our findings highlight the challenges that researchers have encountered in defining and measuring food insecurity, as well as the difficulties faced by humanitarian and development organizations when using individual measures such as FCS, HDDS, and rCSI to capture localized needs. Relying on a single measure of food insecurity may prove inadequate in capturing the diverse nuances associated with the availability, access, and usage of food. Conversely, a composite index comprising various metrics might obscure these nuances and result in misleading conclusions, particularly if the underlying relationships and importance among the metrics are not well understood, especially in the presence of unique geographical differences. Therefore, composite measures of food insecurity, which typically aggregate indicators of food insecurity, may also fall short in adequately capturing locally expressed needs.

In the second phase of our analysis, we tested for spatial dependencies among our food insecurity indicators at the district level. Our results confirmed the presence of spatial dependencies, with food insecurity scores showing correlations with neighboring districts. Notably, non-random patterns were identified within specific districts, particularly those in the northwest and southeast regions. Significant heterogeneities persisted across food insecurity measures and over time for specific indicators. The most robust and consistent spatial dependencies were associated with rCSI, followed by FCS. These findings reinforce the earlier descriptive observations and again underscore the intricate nature of food insecurity. This led us to hypothesize that spatial variations in food insecurity likely arise from differences in geospatial features across Lebanon and within specific districts.

We used machine learning methods in the third phase of our analysis to assess the importance of the geospatial features in predicting food insecurity. Most notably, geolocational indicators emerged as the top predictors of food insecurity, overshadowing traditional factors like household sociodemographics and living conditions, These results suggest that future analyses aiming to evaluate the food security needs of forcibly displaced populations (FDPs) should consider integrating geospatial features into the models, especially when forecasting future vulnerabilities to food insecurity. Geospatial data, being readily available, offer a more efficient alternative to the labor-intensive process of collecting household socioeconomic survey data from highly vulnerable populations like FDPs.

The results on feature importance also highlight key distinctions between "internal" characteristics specific to individuals and households, such as sociodemographics and living conditions, and "external" characteristics such as geospatial features dependent on location. According to this study, the topography of a location holds significant importance, arguably more so than traditional demographic characteristics, in relation to food insecurity of refugee households in Lebanon. Geospatial features associated with refugee and host populations, vegetation, and type of land coverage were among the top predictors of food insecurity, particularly influencing FCS, rCSI, and share of food expenditures.

Having noted this, the machine learning (ML) analysis assumes the independence of features. It is plausible that certain geographical features are highly correlated with demographic characteristics, especially those linked to economic development and urbanization of a location, as well as the employment and education levels of subpopulations in that locale. Subsequent analysis will aggregate household-level characteristics at the district level to conduct a more refined examination of the relationships between various individual and geospatial features across and within the districts and their subsequent associations to food insecurity.

To gain a deeper understanding of food insecurity among refugees in Lebanon, it is necessary to further explore the various dimensions of food security and the complex interconnections among them and across the different regions and refugee populations. Unraveling the reasons behind disparities in food insecurity indicators between areas (e.g., northwest versus southeast) and variations within specific regions (e.g., east versus west) remains a priority. Our next steps involve building upon these foundational results to develop a better understanding of the relationships between key geospatial features and food insecurity. This will aid in better identifying households more likely to face food insecurity in the future and determining which locations are most vulnerable.

In conclusion, our analysis effectively demonstrated that geolocational indicators are not only important but perhaps the most crucial drivers of food insecurity. Consequently, these geospatial features hold critical value for humanitarian and development organizations when making impactful decisions about which locations and food security needs to prioritize. From a policy perspective, our insights advocate for a nuanced approach to tackling food insecurity among refugees, incorporating geospatial data as an informative tool in this process. By disaggregating the various dimensions of food insecurity and understanding their geospatial distribution, policymakers and humanitarian organizations can better tailor their strategies, channeling resources towards areas where refugees encounter the most severe challenges, thereby enhancing the effectiveness of food security measures.

References

Al Shogoor, S., Sahwan, W., Hazaymeh, K., Almhadeen, E., & Schütt, B. (2022). Evaluating the impact of the influx of Syrian refugees on land use/land cover change in Irbid District, Northwestern Jordan. *Land*, *11*(3), 372.

Al Zoubi, S.T., A. Aw-Hassan, and B. Dhehibi. (2019). *Enhancing the Livelihood and Food Security of Syrian Refugees in Lebanon*. Amman: International Center for Agricultural Research in the Dry Areas (ICARDA). https://mel.cgiar.org/reporting/download/hash/d614b81ba6ae7d056e81edaebb83b4bf

Alemu, Z. A., Ahmed, A. A., Yalew, A. W., & Simanie, B. (2017). Spatial variations of household food insecurity in East Gojjam Zone, Amhara Region, Ethiopia: implications for agroecosystem-based interventions. *Agriculture & Food Security*, *6*, 1-9.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis, 27*(2), 93-115.

Barrett, C. B. (2010). Measuring food insecurity. *Science*, *327*(5967), 825–828.

Biederlack, L., & Rivers, J. (2009). *Comprehensive food security & vulnerability analysis (CFSVA): Ghana*. United Nations World Food Programme.

Brown, M. E. (2016). Remote sensing technology and land use analysis in food security assessment. *Journal of Land Use Science*, *11*(6), 623-641.

Caccavale, O. M., & Giuffrida, V. (2020). The Proteus composite index: Towards a better metric for global food security. *World Development*, *126*, 104709.

Caiserman, A., & Faour, G. (2021). Spatial variability of evapotranspiration and pressure on groundwater resources: remote sensing monitoring by crop type in the Bekaa plain, Lebanon. *Journal of Applied Remote Sensing*, *15*(1), 014517-014517.

Central Administration of Statistics (CAS). (n.d.). Inflation in figures. http://www.cas.gov.lb/images/PDFs/CPI/2021/Inflation%20in%20figures.pdf

Çetinkaya, C., Özceylan, E., Erbaş, M., & Kabak, M. (2016). GIS-based fuzzy MCDA approach for siting refugee camp: A case study for southeastern Turkey. *International Journal of Disaster Risk Reduction*, *18*, 218-231.

Chen, P. C., Yu, M. M., Shih, J. C., Chang, C. C., & Hsu, S. H. (2019). A reassessment of the Global Food Security Index by using a hierarchical data envelopment analysis approach. *European Journal of Operational Research*, *272*(2), 687-698.

Coates, J. (2013). Build it back better: Deconstructing food security for improved measurement and action. *Global Food Security*, *2*(3), 188-194.

Coughlan de Perez, E., van Aalst, M., Choularton, R., van den Hurk, B., Mason, S., Nissan, H., & Schwager, S. (2019). From rain to famine: assessing the utility of rainfall observations and seasonal forecasts to anticipate food insecurity in East Africa. *Food Security*, *11*, 57-68.

Der Sarkissian, R., Zaninetti, J. M., & Abdallah, C. (2019). The use of geospatial information as support for Disaster Risk Reduction; contextualization to Baalbek-Hermel Governorate/Lebanon. *Applied Geography*, *111*, 102075.

Deléglise, H., Interdonato, R., Bégué, A., d'Hôtel, E. M., Teisseire, M., & Roche, M. (2022). Food security prediction from heterogeneous data combining machine and deep learning methods. *Expert Systems with Applications*, *190*, 116189.

Demeke, A. B., Keil, A., & Zeller, M. (2011). Using panel data to estimate the effect of rainfall shocks on smallholders food security and vulnerability in rural Ethiopia. *Climatic Change*, *108*(1-2), 185-206.

Dessie, Z. G., Zewotir, T., & North, D. (2022). The spatial modification effect of predictors on household level food insecurity in Ethiopia. *Scientific Reports*, *12*(1), 19353.

ESCWA. (2021). *Multidimensional poverty in Lebanon (2019-2021): Painful reality and uncertain prospects*. https://www.unescwa.org/sites/default/files/news/docs/21-00634-_multidimentional_poverty_in_lebanon_-policy_brief_-_en.pdf

FAO. (n.d.). *Rome Declaration on World Food Security*. Rome: World Food Summit 1996. https://www.fao.org/3/w3613e/w3613e00.htm

FAO, IFAD and WFP. (2015). *The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress*. Rome: FAO. https://www.fao.org/3/i4646e/i4646e.pdf

Foini, P., Tizzoni, M., Martini, G., Paolotti, D., & Omodei, E. (2023). On the forecastability of food insecurity. *Scientific Reports*, *13*(1), 2793.

Füreder, P., Lang, S., Hagenlocher, M., Tiede, D., Wendt, L., & Rogenhofer, E. (2015, May). Earth observation and GIS to support humanitarian operations in refugee/IDP camps. In ISCRAM.

Ghattas, H., Chaaban, J., Salti, N., Irani, A., Ismail, T., & Batlouni, L. (2018). Poverty, food insecurity, and health of Palestinian refugees in Lebanon and recently displaced from Syria to Lebanon: findings from the 2015 socioeconomic household survey. *The Lancet*, *391*, S11.

Ghattas, H., Sassine, A. J., Seyfert, K., Nord, M., & Sahyoun, N. R. (2014). Food insecurity among Iraqi refugees living in Lebanon, 10 years after the invasion of Iraq: data from a household survey. *British Journal of Nutrition*, *112*(1), 70-79.

Ghoussein, Y., Mhawej, M., Jaffal, A., Fadel, A., El Hourany, R., & Faour, G. (2018). Vulnerability assessment of the South-Lebanese coast: A GIS-based approach. *Ocean & Coastal Management*, *158*, 56-63.

Giada, S., De Groeve, T., Ehrlich, D., & Soille, P. (2003). Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania. *International Journal of Remote Sensing*, *24*(22), 4251-4266.

Hadley, C., Patil, C. L., & Nahayo, D. (2010). Difficulty in the food environment and the experience of food insecurity among refugees resettled in the United States. *Ecology of Food and Nutrition*, *49*(5), 390-407.

Han, J., Pei, J., & Tong, H. (2023). *Data mining: Concepts and techniques*. Cambridge, MA: Morgan Kaufmann.

Hoteit, M., Al-Atat, Y., Joumaa, H., Ghali, S. E., Mansour, R., Mhanna, R., Sayyed-Ahmad, F., Salameh, P. & Al-Jawaldeh, A. (2021). Exploring the impact of crises on food security in lebanon: results from a national cross-sectional study. *Sustainability*, *13*(16), 8753.

Issa, S. T., van der Molen, I., Nader, M. R., & Lovett, J. C. (2014). Spatial variation of vulnerability in geographic areas of North Lebanon. *European Scientific Journal*, *2*, 261-273.

Izraelov, M., & Silber, J. (2019). An assessment of the global food security index. *Food Security*, *11*(5), 1135-1152.

Kharroubi, S., Naja, F., Diab-El-Harake, M., & Jomaa, L. (2021). Food insecurity pre-and post the COVID-19 pandemic and economic crisis in Lebanon: prevalence and projections. *Nutrients*, *13*(9), 2976.

Integrated Food Security Phase Classification (IPC). (2022a). *Lebanon: Acute Food Insecurity Situation September - December 2022 and Projection for January - April 2023*. https://www.ipcinfo.org/ipc-country-analysis/details-map/en/c/1156123/?iso3=LBN

Integrated Food Security Phase Classification (IPC). (2022b). *Lebanon: IPC Acute Food Insecurity Analysis September 2022 - April 2023*. https://reliefweb.int/report/lebanon/lebanon-acute-food-insecurity-situation-september-december-2022-and-projection-january-april-2023-economic-crisis-currency-depreciation-and-unprecedented-increase-food-and-non-food-prices-worsen-lebanon-food-security-situation

International Trade Administration. (2022). *Lebanese – Country Commercial Guide: Agricultural Sector.* https://www.trade.gov/country-commercial-guides/lebanon-agricultural-sector

Lebanon's Council for Development & Reconstruction (2005). *National Physical Master Plan of the Lebanese Territory (NPMPLT)*. Chapter 1: Uncontested Physical Features. https://www.cdr.gov.lb/en-US/Studies-and-reports/National-physical-master-plan.aspx

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

Lone, S. A., & Mayer, I. A. (2019). Geo-spatial analysis of land use/land cover change and its impact on the food security in District Anantnag of Kashmir Valley. *GeoJournal*, *84*, 785-794.

Lv, F., Deng, L., Zhang, Z., Wang, Z., Wu, Q., & Qiao, J. (2022). Multiscale analysis of factors affecting food security in China, 1980–2017. *Environmental Science and Pollution Research*, *29*(5), 6511-6525.

Lyons, A. C., Kass-Hanna, J., & Montoya Castano, A. (2023a). A multidimensional approach to measuring vulnerability to poverty among refugee populations. *Journal of International Development*, 1–32.

Lyons, A. C., Kass-Hanna, J., Montoya Castano, A., Zhang, Y., & Soliman, A. (2023b). A machine learning and geospatial approach to targeting humanitarian assistance for forcibly displaced populations. *Economic Research Forum's Working Paper Series No. ERF29AC_106*. Cairo, Egypt: Economic Research Forum

Maione, C., Nelson, D. R., & Barbosa, R. M. (2019). Research on social data by means of cluster analysis. *Applied Computing and Informatics*, *15*(2), 153–162.

Manikas, I., Ali, B. M., & Sundarakani, B. (2023). A systematic literature review of indicators measuring food security. *Agriculture & Food Security*, *12*(1), 10.

Mansour, R., Liamputtong, P., & Arora, A. (2020). Prevalence, determinants, and effects of food insecurity among middle eastern and north African migrants and refugees in high-income countries: a systematic review. *International Journal of Environmental Research and Public Health*, *17*(19), 7262.

Martini, G., Bracci, A., Riches, L., Jaiswal, S., Corea, M., Rivers, J., ... & Omodei, E. (2022). Machine learning can guide food security efforts when primary data are not available. *Nature Food*, *3*(9), 716-728.

Mathenge, M., Sonneveld, B. G., & Broerse, J. E. (2023). Mapping the spatial dimension of food insecurity using GIS-based indicators: A case of Western Kenya. *Food Security*, *15*(1), 243-260.

Maxwell, D., Coates, J., & Vaitla, B. (2013). *How do different indicators of household food security compare? Empirical evidence from Tigray*. Feinstein International Center, Tufts University. https://fic.tufts.edu/wp-content/uploads/Different-Indicators-of-HFS.pdf

Maxwell, D., Vaitla, B., & Coates, J. (2014). How do indicators of household food insecurity measure up? An empirical comparison from Ethiopia. *Food Policy*, *47*, 107–116.

Meerza, S. I. A., Meerza, S. I. A., & Ahamed, A. (2021). Food Insecurity Through Machine Learning Lens: Identifying Vulnerable Households. https://ageconsearch.umn.edu/record/314072?ln=en

Müller, M. F., Yoon, J., Gorelick, S. M., Avisse, N., & Tilmant, A. (2016). Impact of the Syrian refugee crisis on land use and transboundary freshwater resources. Proceedings of the National Academy of Sciences, *113*(52), 14932-14937.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825-2830.

Reig, E. (2012). Food security in African and Arab countries: A review of the topic and some suggestions for building composite indicators with Principal Components Analysis. *Department of Applied Economics II, Universidad de Valencia Working Papers*, *1210*.

Santeramo, F. G. (2015). On the composite indicators for food security: Decisions matter! *Food Reviews International*, *31*(1), 63–73.

Swindale, A., & Bilinsky, P. (2006). *Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide (v.2)*. Washington, D.C.: FHI 360/FANTA https://www.fantaproject.org/sites/default/files/resources/HDDS_v2_Sep06_0.pdf

UN-Habitat Lebanon and ESCWA. (2021). *State of the Lebanese Cities 2021*. Beirut, Lebanon: UN-Habitat Lebanon.

United Nations High Commissioner for Refugees (UNHCR). (2023). Refugee data finder. https://www.unhcr.org/refugee-statistics/

United Nations High Commissioner for Refugees (UNHCR), the United Nations World Food Programme (WFP), and the United Nations Children's Fund (UNICEF). (2018). *VASyR 2018: Vulnerability assessment of Syrian refugees in Lebanon*. Beirut, Lebanon. Retrieved from https://www.unhcr.org/lb/wp-content/uploads/sites/16/2018/12/VASyR-2018.pdf

United Nations High Commissioner for Refugees (UNHCR), the United Nations World Food Programme (WFP), and the United Nations Children's Fund (UNICEF). (2019). *VASyR 2019: Vulnerability assessment of Syrian refugees in Lebanon*. Beirut, Lebanon. https://data.unhcr.org/en/documents/details/73118

United Nations High Commissioner for Refugees (UNHCR), the United Nations World Food Programme (WFP), and the United Nations Children's Fund (UNICEF). (2021). *VASyR 2020: Vulnerability assessment of Syrian refugees in Lebanon*. Beirut, Lebanon. https://data.unhcr.org/en/documents/details/85002

United Nations High Commissioner for Refugees (UNHCR), the United Nations World Food Programme (WFP), and the United Nations Children's Fund (UNICEF). (2022). *VASyR 2021: Vulnerability assessment of Syrian refugees in Lebanon*. Beirut, Lebanon. https://data.unhcr.org/en/documents/details/90589

United Nations High Commissioner for Refugees (UNHCR), the World Food Programme (WFP), and the United Nations Children's Fund (UNICEF). (2023). *VASyR 2022: Vulnerability assessment of Syrian refugees in Lebanon. Beirut, Lebanon*. https://data.unhcr.org/en/documents/details/100844

Vaitla, B., Coates, J., Glaeser, L., Hillbruner, C., Biswal, P., & Maxwell, D. (2017). The measurement of household food security: Correlation and latent variable analysis of alternative indicators in a large multi-country dataset. *Food Policy*, *68*, 193-205.

Wineman, A. (2016). Multidimensional household food security measurement in rural Zambia. *Agrekon*, *55*(3), 278–301. https://doi.org/10.1080/03031853.2016.1211019

World Bank. (2022). *Lebanon Economic Monitor, Fall 2022: Time for an equitable banking resolution*. World Bank: Washington, D.C. https://www.worldbank.org/en/country/lebanon/publication/lebanon-economic-monitor-fall-2022-time-for-an-equitable-banking-resolution

World Bank. (2023). *First shipment of 33,000 tons of wheat helps rebuild Lebanon's stock and ensure access to affordable bread.* https://www.worldbank.org/en/news/press-release/2023/02/11/first-shipment-of-33-000-tons-of-wheat-helps-rebuild-lebanon-s-stock-and-ensure-access-to-affordable-bread

World Food Programme (WFP). (2008). *Food consumption analysis: Calculation and use of the food consumption score in food security analysis*. WFP Vulnerability Analysis and Mapping Branch. https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp197216.pdf

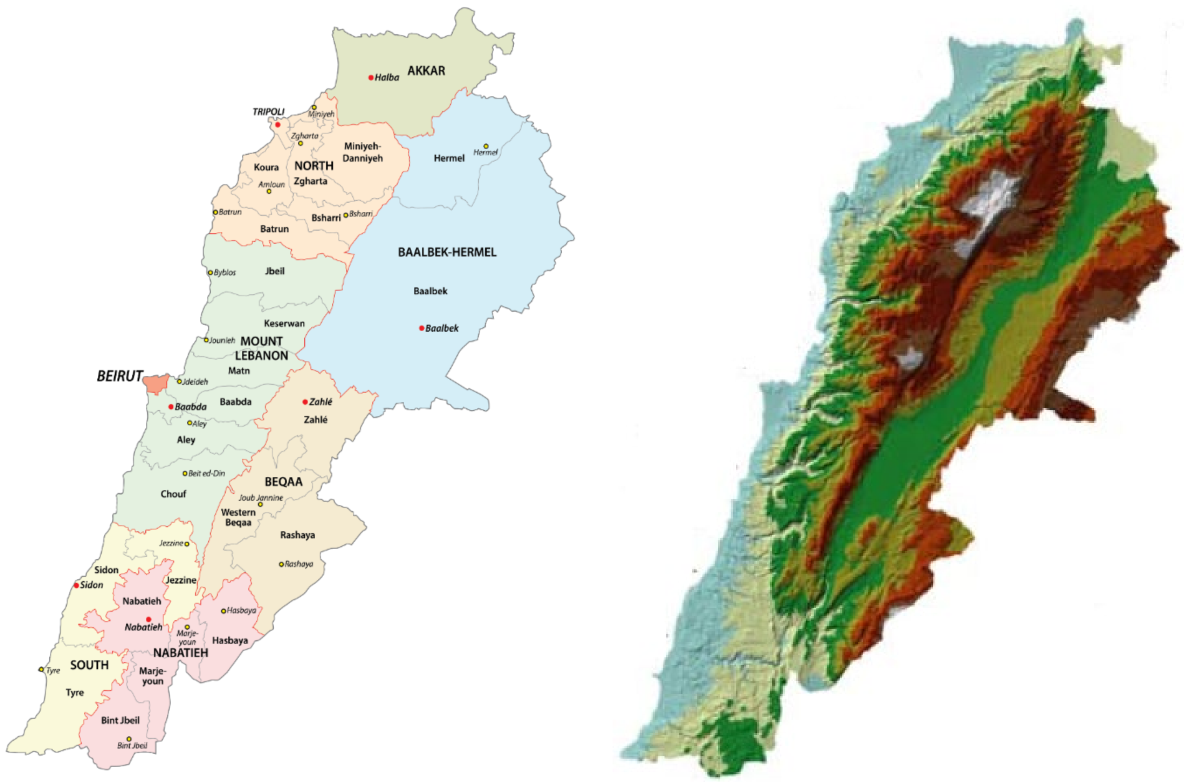
World Food Programme (WFP). (2021). *Lebanon: Annual country report 2021*. https://docs.wfp.org/api/documents/WFP-0000137877/download/?_ga=2.163618879.1843813489.1689273787-812322800.1689273787

World Food Programme (WFP). (2022a). *Lebanon: Annual country report 2022*. https://docs.wfp.org/api/documents/WFP-0000147967/download/?_ga=2.62959535.1843813489.1689273787-812322800.1689273787

World Food Programme (WFP). (2022b). *Food security and vulnerability analysis of Lebanese residents*. WFP Lebanon : Research, Assessment & Monitoring Unit. https://reliefweb.int/report/lebanon/wfp-lebanon-food-security-and-vulnerability-analysis-lebanese-residents-july-2022

Younes, A., Kotb, K. M., Ghazala, M. O. A., & Elkadeem, M. R. (2022). Spatial suitability analysis for site selection of refugee camps using hybrid GIS and fuzzy AHP approach: The case of Kenya. *International Journal of Disaster Risk Reduction, 77*, 103062.

Figure 1. Geography and Topography of Lebanon



Sources: *World Atlas (2020)*. Retrieved from https://www.worldatlas.com/maps/lebanon; *Lebanon's Council for Development & Reconstruction (2005)*

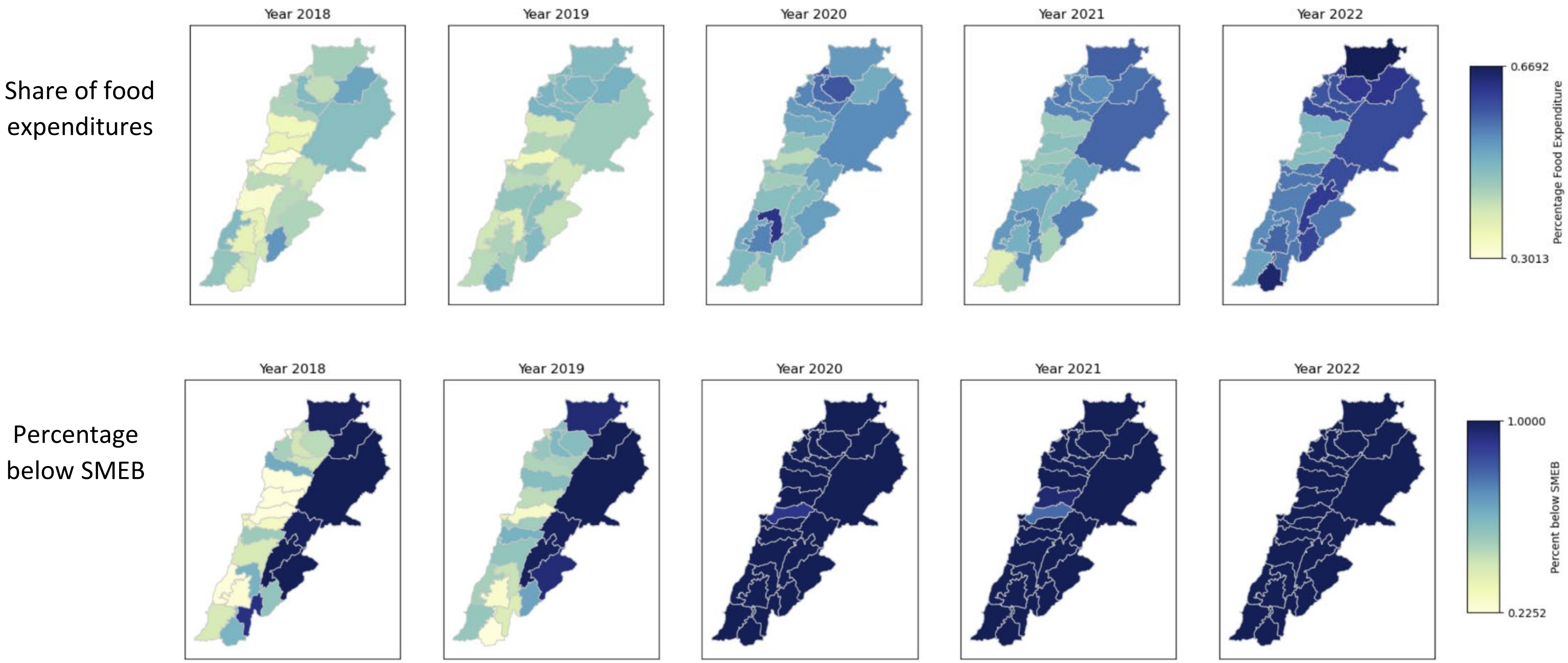
Table 1. Changes in food security by year

Food security measures	2018 (N=4,433)	2019 (N=4,670)	2020 (N=4,480)	2021 (N=4,967)	2022 (N=4,076)	p-value
Total expenditures per capita (LBP)	170,857.90	180,346.60	196,254.00	345,410.0	1,600,000.00	<0.001
Food expenditures per capita (LBP)	62,218.65	71,589.30	91,963.30	161,054.30	864,089.10	<0.001
Share of food expenditures	0.41	0.43	0.50	0.50	0.58	<0.001
Expenditures < SMEB	0.50	0.52	0.88	0.84	0.99	<0.001
FCS	53.34	55.12	44.98	48.29	43.74	<0.001
HDDS	9.12	9.42	8.14	8.46	8.41	<0.001
rCSI	17.77	18.81	17.25	20.08	20.61	<0.001
rCSI1 (reduced meals)	0.59	0.62	0.66	0.70	0.74	<0.001
rCSI2 (reduced portions)	0.55	0.60	0.64	0.73	0.78	<0.001
rCSI3 (consumed less expensive food)	0.86	0.88	0.88	0.93	0.95	<0.001
rCSI4 (borrowed food)	0.36	0.38	0.40	0.42	0.47	<0.001
rCSI5 (restricted consumption)	0.37	0.37	0.33	0.33	0.42	<0.001

Note: The sample size is reduced for share of food expenditures, because some households reported total expenditures equal to 0 (n=4,362 for 2018; n= 4,604 for 2019; n=4,413 for 2020; n=4,912 for 2021; n= 4,076 for 2022). The cutoffs for the Survival Minimum Expenditure Basket (SMEB) were as follows for each year: 87 USD for 2018 and 2019 (equivalent to 131,153 LBP); 308,722 LBP for 2020, 490,028 LBP for 2021, and 8,156,858 LBP for 2022.

Figure 2. Changes in food insecurity for FCS, HDDS, and rCSI by district and across years (darker shading indicates higher levels of food insecurity)

Figure 3. Changes in share of food expenditures and share of refugee households below the SMEB by district and across years (darker shading indicates higher levels of food insecurity)



NOTE: The SMEB cutoffs were as follows for each year: 87 USD for 2018 and 2019 (equivalent to 131,153 LBP); 308,722 LBP for 2020, 490,028 LBP for 2021, and 8,156,858 LBP for 2022.

Table 2. Global Moran's I (GMI) and simulated p-values (PSIM)

Food security measures		2018 (N=4,433)	2019 (N=4,670)	2020 (N=4,480)	2021 (N=4,967)	2022 (N=4,076)
FCS	GMI	0.080	**0.519**	-0.045	**0.321**	**0.292**
	PSIM	0.147	**0.001**	0.500	**0.010**	**0.007**
HDDS	GMI	0.018	**0.349**	0.072	0.194	0.170
	PSIM	0.301	**0.001**	0.190	0.045	0.059
rCSI	GMI	**0.327**	**0.566**	**0.358**	**0.402**	0.090
	PSIM	**0.008**	**0.001**	**0.003**	**0.001**	0.158
Share of food expenditures	GMI	0.169	**0.354**	0.245	0.284	**0.306**
	PSIM	0.064	**0.004**	0.019	0.012	**0.009**
Expenditures < SMEB	GMI	0.232	**0.366**	**0.337**	0.152	0.017
	PSIM	0.030	**0.006**	**0.002**	0.067	0.263

Note: GMI represents the Global Moran's I. PSIM represents the simulated p-values for each measure of food insecurity. Bolded values indicate significance at p<0.01.

Figure 4. Local Moran's I (LMI)

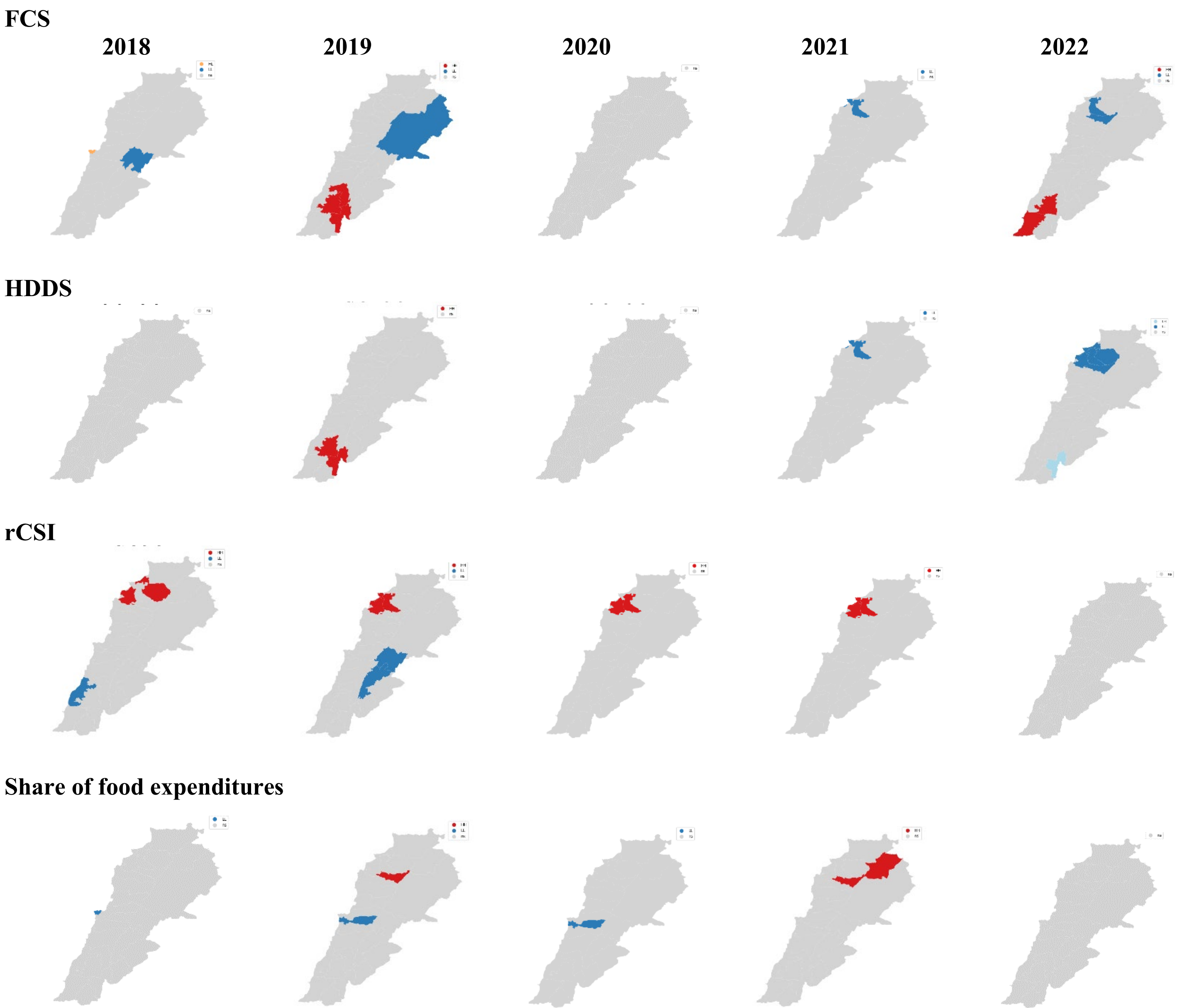


Figure 4. Local Moran's I (LMI) (continued)

Expenditures < SMEB

2018 **2019** **2020** **2021** **2022**

Note: Gray areas indicate districts where there is spatial randomness and no evidence of a spatial structure. Red represents a "hot spot" where there is a high score district surrounded by high score neighbors (i.e., high levels of food insecurity). Blue represents a "cold spot" where there is a low score district surrounded by low score neighbors (i.e., low levels of food insecurity), Cyan represents an outlier district with low score surrounded by a relatively high score neighbors, and vice versa, orange indicates an outlier district with high score surrounded by a relatively low score neighbors.

Table 3. Predictive power of the machine learning models for food insecurity measures

Target variable	Cutoff value for bottom 30%	Accuracy of prediction		
		Random Forest	Gradient Boosting	Lasso
FCS	≤ 37	0.6208	0.6467	0.0989
HDDS	≤ 8	0.6277	0.6850	0.4565
rCSI	≥ 24	0.6769	0.7111	0.2078
Share of food expenditures	≥ 60.4%	0.6073	0.5788	0.1362

Note: This table shows the percentage of times the model accurately predicted food insecurity for each measure. The predictive power of the models was determined using the recall value. Predictions are based on the classification of whether the household was in the bottom 30% for each measure. To improve model accuracy, balanced, class weights were applied to account for the skewed distribution resulting from the classification of households as either 0 or 1 depending on whether they are food insecure. Each ML model also includes all the features (the household characteristics, the living conditions, and the geospatial features).

Figure 5. Random Forest Models: Categorical feature importance for each food insecurity measure (years pooled)

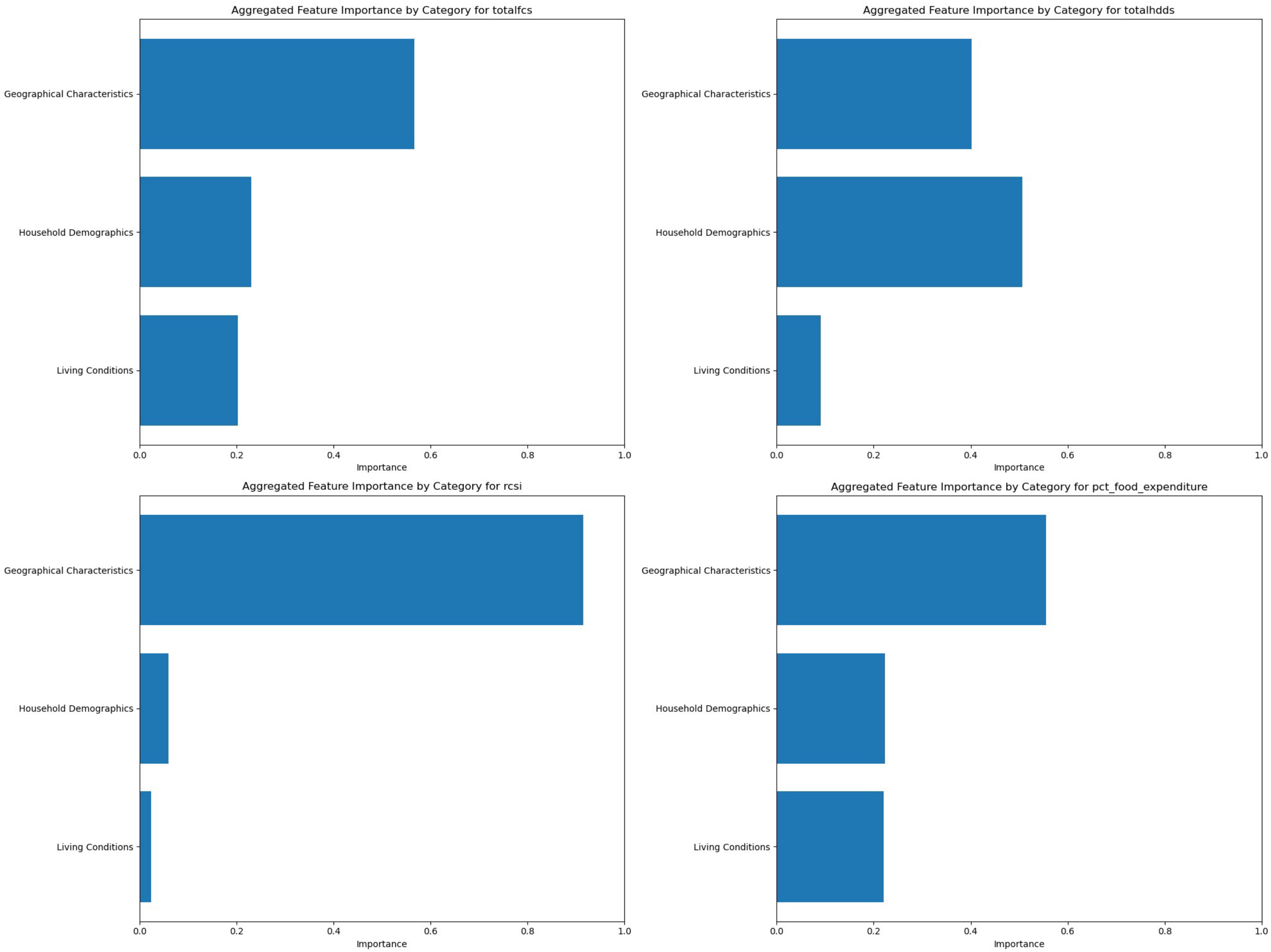


Figure 6. Random Forest Models: Top 20 most important features for each food insecurity measure (years pooled)

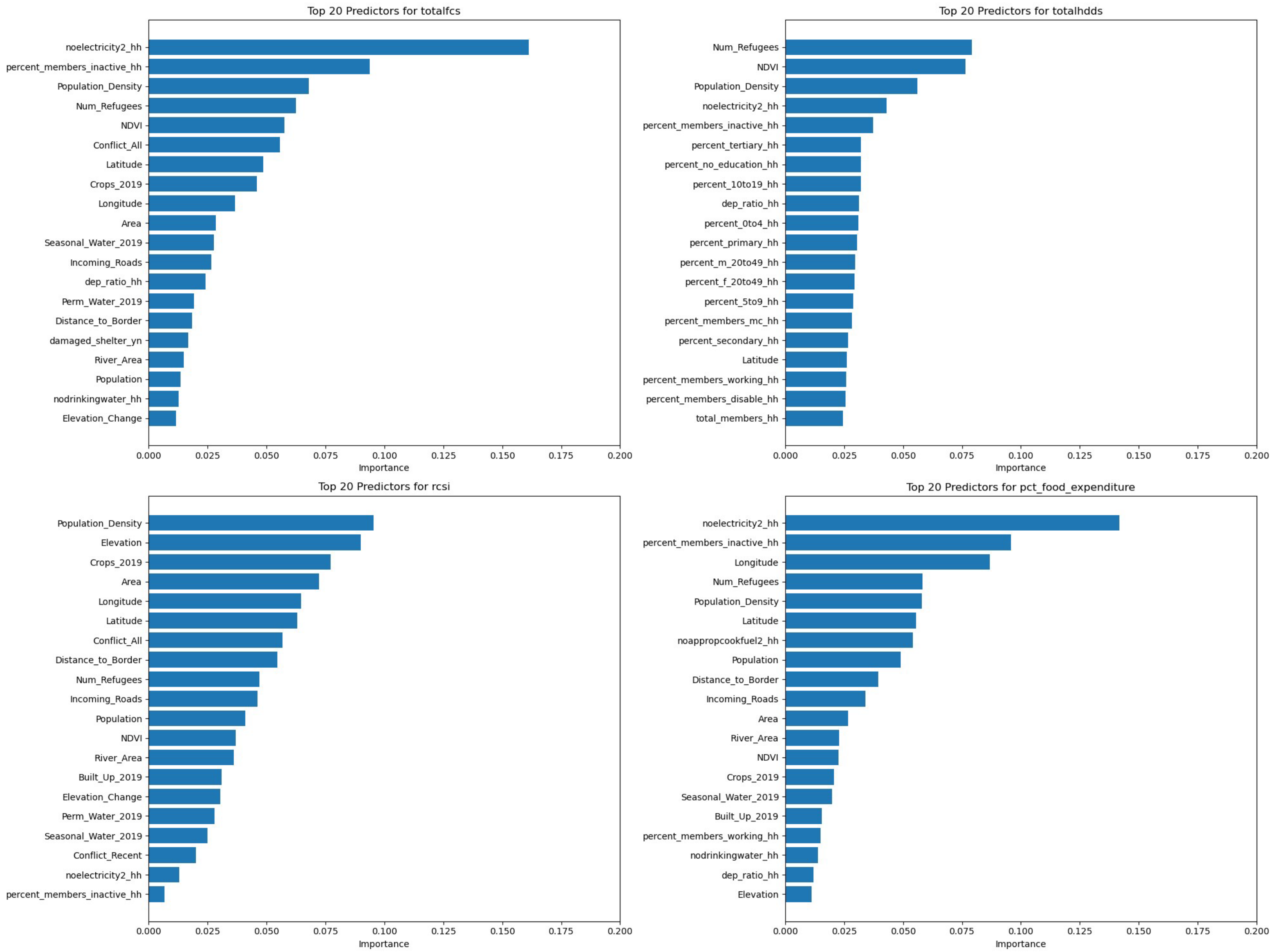


Table A1. Variable definitions

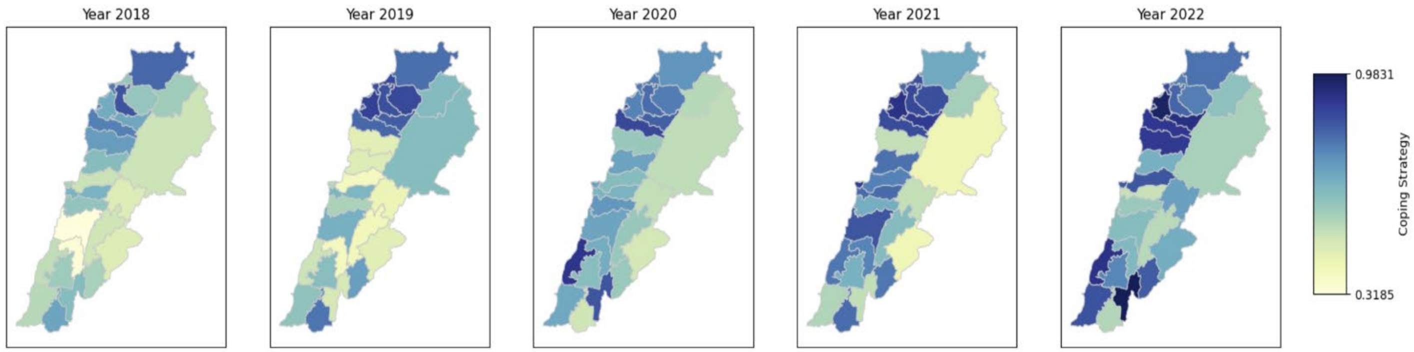
Variables	Variable name	Definitions
Food insecurity measures		
FCS	totalfcs	The Food Consumption Score (FCS) measures the diversity and frequency of households' diets in the week prior to the survey. Scores range from 0 to 112, with lower scores indicating less diet diversity; the FCS is grouped into three categories: acceptable (>42), borderline (28-42), and poor (<28)
HDDS	totalhdds	Household's dietary diversity score measures the number of food groups consumed during the last 24 hours. The index ranges from 0 to 12 (the total number of food groups). A score lower than 6 is considered as low diversity, 7-8 borderline, and 9 or higher acceptable.
rCSI	rcsi	Reduced food coping strategies index (rCSI) measures the strategies that households use to cope with the lack of food and the severity of the strategies used to compare the hardship faced by households due to a shortage of food. The index ranges from 0 (no coping strategies) to 56 (severe level of coping strategies), with higher scores indicating more food coping strategies are being used. Households are classified as having a low (0-3), medium (4-18), or high (≥19) rCSI.
Share of food expenditure per capita	pct_food_expenditure	Expenditures per capita in Lebanese Pounds (LBP) on food as a share of total expenditures per capita
Survival Minimum Expenditure Basket (SMEB)	Povertyindicators1	=1 if the household's monthly expenditures per capita is below the Survival Minimum Expenditure Basket (SMEB) cutoff. This cutoff varies by year. For 2018 and 2019, it was equal to 87 USD (equivalent to 131,153 LBP); for 2020, it was equal to 308,722 LBP; for 2021, it was equal to 490,028 LBP, and 8,156,858 LBP for 2022.
Food insecurity measures (cutoffs for bottom 30% for GB models)		
FCS ≤ 37		=1 if Food Consumption Score (FCS) score is less than or equal to 37, indicating "poor" diet diversity that is at an unacceptable level (poor and borderline food consumption) and in the bottom 30% of households.
HDDS ≤ 8		=1 if the household's dietary diversity score was less than or equal to 8, indicating "poor" diet diversity in terms of the number of food groups consumed by the household in a 24-hour period.
rCSI ≥ 24		=1 if the Reduced Coping Strategies Index (rCSI) is greater than or equal to 24, indicating a "high" number of food coping strategies are being used and that the household is in the bottom 30% of households.
Share of food expenditure per capita ≥ 60.4%		=1 if the share was greater than 60.4%, indicating that the household is in the bottom 30% of households.
Household demographics		
Household size	total_members_hh	Number of household members
Dependency ratio	dep_ratio_hh	Ratio of dependent household members (below 15 or above 60 years of age) relative to total household members
% HH members aged 0-4	percent_0to4_hh	Percentage of children aged 0 to 4 in each household
% HH members aged 5-9	percent_5to9_hh	Percentage of children aged 5 to 9 in each household
% HH members aged 10-19	percent_10to19_hh	Percentage of household members aged 10 to 19 in each household
% Male members aged 20-49	percent_m_20to49_hh	Percentage of male adults aged 20 to 49 in the household
% Female members aged 20-49	percent_f_20to49_hh	Percentage of female adults aged 20 to 49 in the household
% HH members older than 60	percent_60_hh	Percentage of household members aged 60 and above
% HH members education unknown		Percentage of household members who do not report any educational level
% HH members no education	percent_no_education_hh	Percentage of household members who did not go to school
% HH members primary education	percent_primary_hh	Percentage of household members who completed primary education
% HH members secondary education	percent_secondary_hh	Percentage of household members who completed secondary education
% HH members above secondary education	percent_tertiary_hh	Percentage of household members with high school, technical, or college diploma
% HH members working	percent_members_working_hh	Percentage of household members who are working

% HH members unemployed	percent_members_unemployed_hh	Percentage of household members who are unemployed
% HH members inactive	percent_members_inactive_hh	Percentage of household members who are inactive
% HH members with disability	percent_members_disable_hh	Percentage of household members with any disability (seeing, hearing, walking, etc.)
% HH members with medical condition	percent_members_mc_hh	Percentage of household members with a chronic illness or unable to care for themselves
Child not attending school	schoolage_notatt_hh	=1 if household has a child who is of school age (5 to 14 years of age) who is not attending school
Disabled Head	disable_head	=1 if the head has a disability
Disabled dependent member	disable_dependent	=1 if at least one member of the household other than the head has a disability
% Illegal residency	illegal_residency_hh	Percentage of household members aged 15 or older who do not have legal residency in Lebanon
Living conditions		
Electricity	noelectricity2_hh	=1 household does not have access to electricity or has access for less than 16 hours
Sanitation	nosanitation_hh	=1 if household does not have access to basic sanitation (i.e., no access to flushed toilets or improved pit latrines with a cement slab, and are not sharing the toilets with other households)
Drinking water	nodrinkingwater_hh	=1 if household does not have access to clean drinking water
Cooking fuel	noappropcookfuel2_hh	=1 if household does not have access to electric or gas stove and cooks only with dung, wood, or charcoal
Shelter crowdedness	overcrowding	=1 if household is living in an overcrowded shelter with less than 4.5m^2 per person
Insecurity issues	insecurity_issues_hh	=1 if household reports feeling insecure to risks such as kidnapping and extortion
Damaged Shelter	damaged_shelter_yn	=1 if the shelter is damaged in any way
Geographical Characteristics		
Latitude	Latitude	Geographical coordinates
Longitude	Longitude	Geographical coordinates
Distance to Border	Distance_to_Border	Distance to the closest border between Syria and Lebanon (km)
Elevation	Elevation	Average elevation (km)
Elevation Change	Elevation_Change	Difference between highest and lowest elevation (km)
NDVI	NDVI	Normalized Difference Vegetation Index (NDVI) is a standardized measure of healthy vegetation and how sensitive vegetation in a particular area may be to drought (agriculture); the average NDVI was calculated for each district using the years 2018, 2019, 2020, and 2021.
Area	Area	Area of the district km^2
Population		Average total population was calculated for each district using the years 2018, 2019, and 2020; based on the population counts taken from the WorldPop adjusted to match the UN estimation count
Number of refugees	Num_Refugees	Manually extracted from UN maps
Built area	Built_Up_2019	Average fraction coverage of built-up area was calculated for each district using the years 2018 and 2019
Crop area	Crops_2019	Average fraction coverage of crop covered area was calculated for each district using the years 2018 and 2019
Permanent water	Perm_Water_2019	Average fraction coverage of permanent water area was calculated for each district using the years 2018 and 2019
Seasonal water	Seasonal_Water_2019	Average fraction coverage of seasonal water area was calculated for each district using the years 2018 and 2019
Population Density	Population_Density	Divide population by area to get population density (persons/km^2)
Conflicts – all	Conflict_All	Number of conflict deaths
Conflicts – recent	Conflict_Recent	Number of recent conflict deaths with the last 25 years
Area near river	River_Area	Area of land within 500 meters of a river
Incoming roads	Incoming_Roads	Number of incoming roads within each district

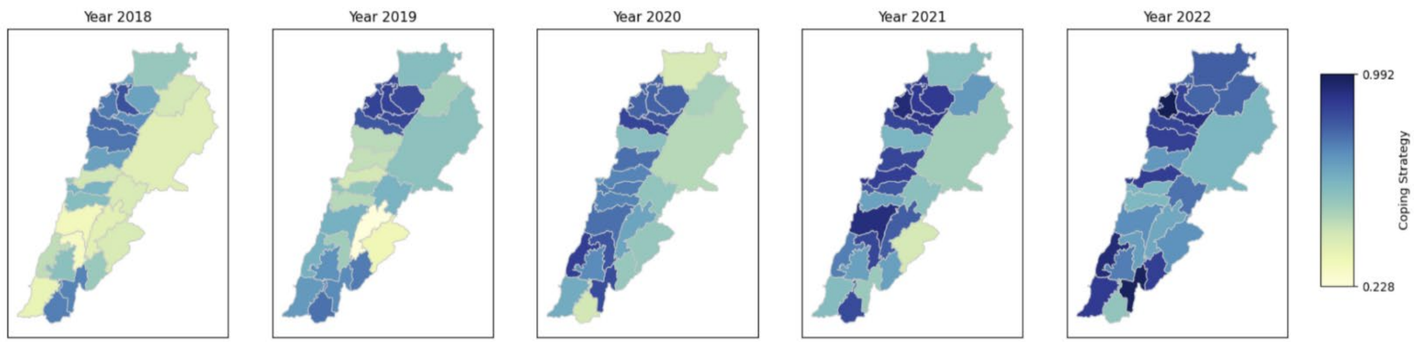
Sources: 2018, 2019, 2020, 2021, and 2022 Vulnerability Assessment of Syrian Refugees (VASyR).

Figure A1. Changes in the 5 components of the reduced coping strategies index (rCSI) across districts and years

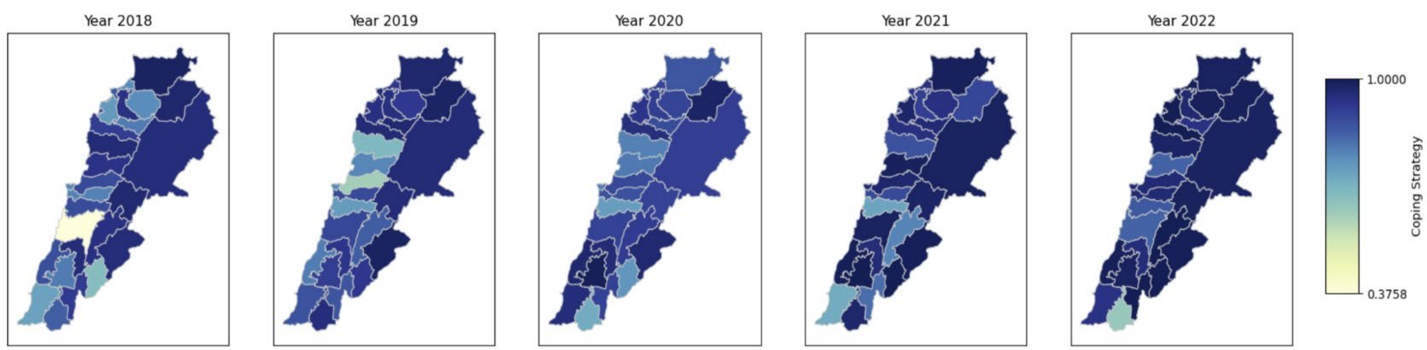
rCSI 1: Reduced number of meals per day



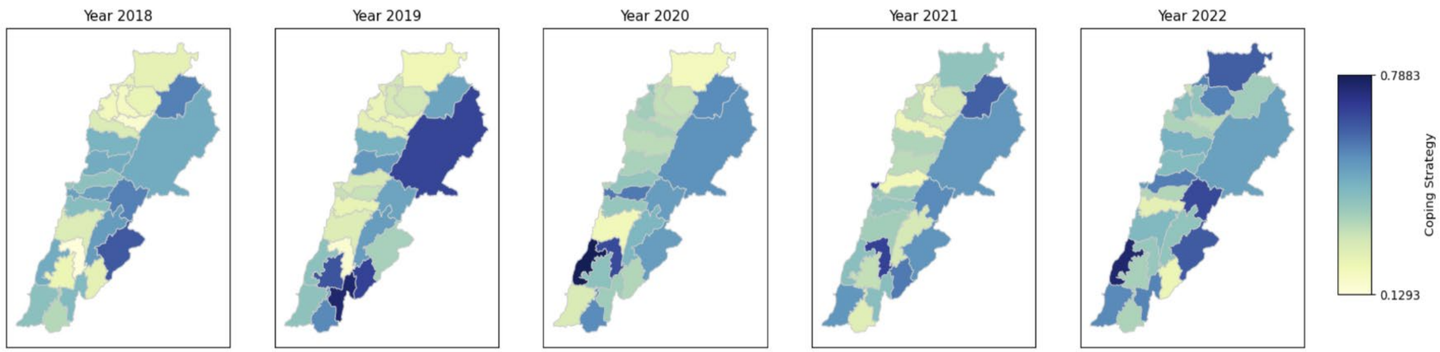
rCSI 2: Reduced portion size of meals



rCSI 3: Relied on less preferred, less expensive food



rCSI 4: Borrowed food or relied on help from friends or relatives



rCSI 5: Restricted consumption by adults in order for young-small children to eat

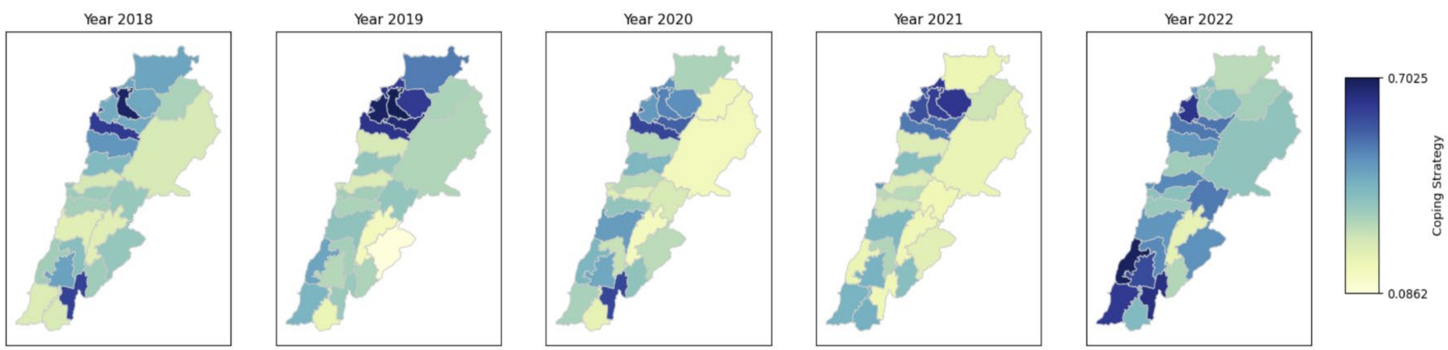


Figure A2. Gradient Boosting Models: Categorical feature importance for each food insecurity measure (years pooled)

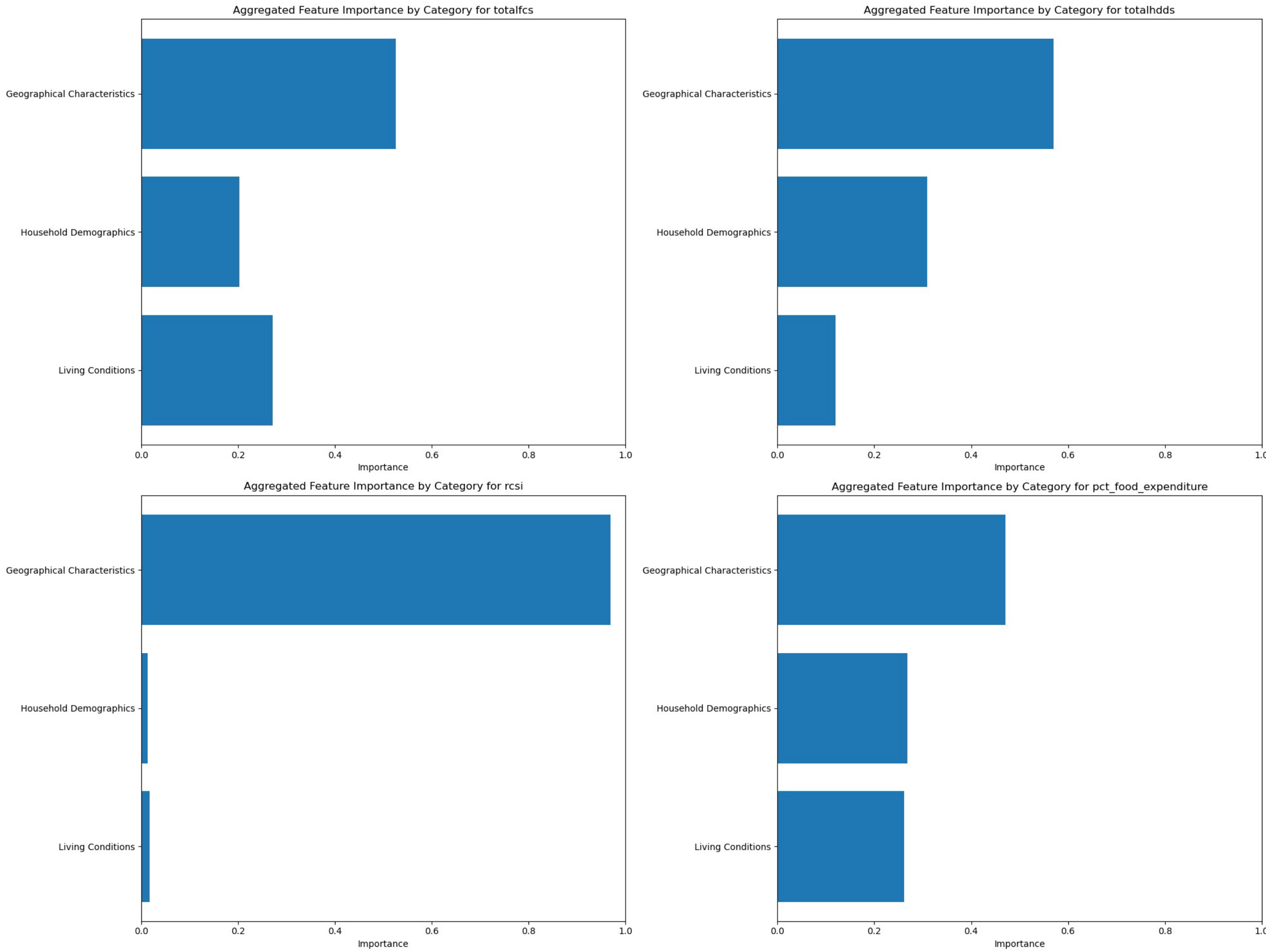


Figure A3. Gradient Boosting Models: Top 20 most important features for each food insecurity measure (years pooled)