DISASTER RISK ASSESSMENT IN CAMBODIA

Leveraging Artificial Intelligence (AI) and Geospatial Data for Evidence-Informed Risk Management and Resilience Enhancement

WFP
World Food Programme

NATIONAL COMMITTEE FOR DISASTER MANAGEMENT

November 2024

Disclaimer

The opinions expressed in this publication are those of the author/researcher and do not necessarily reflect the views of the National Committee for Disaster Management (NCDM) or World Food Programme (WFP). Responsibility for the opinions expressed in this report rests solely with the author/researcher. Any references to, or maps of, specific territories or geographic areas in this document do not imply endorsement by the NCDM or WFP.

Image Credits

Cover page: WFP/Salvador Bustamante, YayImages
Page 2: WFP/Para Hunzai
Page 4: NCDM
Page 6: WFP Salvador Bustamante
Page 8-11: Google Earth Engine
Page 12: WFP Salvador Bustamante
Page 13, 15, 16: WFP/Chanvibol Choeur, NCDM
Page 19: WFP/Para Hunzai

TABLE OF CONTENT

ACRONYMS AND ABBREVIATIONS

ADB	Asia Development Bank
AI	Artificial Intelligence
CAS	Cambodia Agriculture Survey
CDRI	Cambodia Development Resource Institute
CREWS	Climate Risk and Early Warning Systems
CSES	Cambodia Socio-Economic Survey
FAO	Food and Agricultural Organization
GDP	Gross Domestic Product
GEE	Google Earth Engine
GIS	Geographic Information Systems
IPCC	Intergovernmental Panel on Climate Change
MAFF	Ministry of Agriculture, Forestry and Fisheries
ML	Machine Learning
MoEYS	Ministry of Education, Youth and Sport
MoH	Ministry of Heath
MLMUPC	Ministry of Land Management, Urban Planning and Construction
MoP	Ministry of Planning
MOWRAM	Ministry of Water Resource and Meteorology
NCDM	National Committee for Disaster Management
NDWI	Normalized Difference Water Index
NIS	National Institute of Statistics
NSPC	National Social Protection Council
PCDM	Provincial Committee for Disaster Management
PRISM	Platform for Real-Time Impact and Situation Monitoring
RF	Random Forests
RGC	Royal Government of Cambodia
RS	Remote Sensing
SPI	Standardized Precipitation Index
TCI	Temperature Condition Index
UN	United Nations
UNDP	United Nations Development Programme
UNDRR	United Nations Office for Disaster Risk Reduction
UNICEF	United Nations International Children's Emergency Fund
USAID	United States Agency for International Development
VCI	Vegetation Condition Index
WB	The World Bank
WFP	World Food Programme
WMO	World Meteorological Organization
WRI	World Risk Index

ACKNOWLEDGMENT

The Disaster Risk Assessment in Cambodia, which focuses on leveraging artificial intelligence (AI) and geospatial data to generate evidence for managing risk and enhancing resilience, was conducted during 2023-2024 through a collaboration between the National Committee for Disaster Management (NCDM) of the Royal Government of Cambodia and the United Nations World Food Programme (WFP) Cambodia.

The successful completion of data analysis, consultative workshops, and report writing for this assessment was made possible by the collective efforts of both the NCDM and WFP teams. We also extend our gratitude to all stakeholders from line ministries and partner organizations who contributed valuable insights during the consultative workshop and meetings.

This assessment was financially supported by the United States Agency for International Development (USAID) Bureau for Humanitarian Assistance (BHA) and the Climate Risk and Early Warning Systems (CREWS) initiative.

EXECUTIVE SUMMARY

Cambodia is highly prone to floods and droughts, with severe impacts that disproportionately affect people based on their socioeconomic status. Understanding these climatic disaster risks is not only important but also urgent, especially as the frequency and intensity of natural disasters continue to rise due to climate change. Comprehensive disaster risk modeling—encompassing climatic hazards, exposure, and socioeconomic vulnerability—now benefits from advances in geospatial technology, artificial intelligence (AI), and cloud-based computing platforms such as Google Earth Engine (GEE).

Floods are typically caused by lake and river overflows and excessive rainfall, while droughts are driven by erratic rainfall and rising temperatures, putting stress on water availability and vegetation. Floods and droughts frequently occur in low-lying communes around the Tonle Sap Lake/River, along the Mekong River and in the southern plains. Despite the high frequency of these events, these regions are home to a large proportion of the population, critical infrastructure, and key agriculture activities. As a result, inhabitants and their livelihoods are substantially exposed to these disasters.

Moreover, the capacity of households in these areas to cope with and adapt to disasters is considerably limited. Socioeconomic vulnerability is particularly pronounced in rural communes, where the majority of households live and rely heavily on agriculture for their livelihoods. In contrast, urban communes—especially in Phnom Penh and other provincial towns—tend to show lower vulnerability, attributed to better access to essential services, improved infrastructure, and diverse economic opportunities.

These factors illustrate deep insights into the risks that households face from floods and drought. Communes near the Tonle Sap Lake/River, along the Mekong River, and in the southern plains face a double risk of both floods and droughts. Nationally, it is estimated that floods pose a risk to 15.2% of the population and 16.1% of agricultural land, while 29.2% of the population and 33.3% of agricultural land are at risk from droughts.

To effectively reduce risk and build resilience against future climate-induced disasters, the study outlines eight key recommendations for the government and development partners to consider enhancing planning, financing and targeting preparedness effects for responses:

» **Integrate risk information into disaster management systems:** Incorporating the risk information into national and sub-national contingency plans will enhance preparedness and response efforts, while also strengthening systematic institutional sustainability.

» **Strengthen early warning systems (EWS):** Strengthening impact-based monitoring and forecasting capabilities by integrating of hydrometeorological hazard forecasting and socioeconomic vulnerability data into national systems for early warning and impact situation monitoring, enabling the timely activation of emergency responses.

» **Expand social assistance response mechanisms:** Using the risk information to guide ex-ante planning, financing and targeting for social assistance interventions and expanding the social registry to cover smallholder farmers, will enhance needs-based and well-resourced preparedness and response.

» **Operationalize disaster risk financing strategy:** The risk information can support in defining premiums, payout thresholds, and action triggers for disaster risk financing tools and mechanisms, such as parametric insurance.

» **Integrate climate projections into risk assessment:** With escalating climate impacts, incorporating localized climate projection data into assessment of climatic hazards strengthens risk modeling, providing forward-looking information for mitigation and adaptation efforts.

» **Enhance reliability of artificial intelligence (AI) in risk modeling:** Embedding geographical reference information in data collection for national surveys, censuses, and assessments will enhance the quality and availability of training data, ensuring reliable results from machine learning (ML) algorithms. Standardized procedures for capturing accurate, consistent and interoperable post-disaster loss and damage data can support calibration and validation of risk model.

» **Implement dynamic socioeconomic vulnerability assessment:** Adopting the AI-driven vulnerability assessment model outlined in this report as a dynamic modeling approach-integrating near-real-time Earth observation data and on-the-ground socioeconomic data with machine learning analytics-enables timely updates to vulnerability data.

» **Embrace AI and geospatial technologies:** Strengthening the capacity of government institutions to utilize AI/ML and geospatial technologies for data analytics will enable the generation of precise, timely, and actionable insights for disaster preparedness and response activities.

1. INTRODUCTION

1.1. BACKGROUND

The impacts of climate change have become a pressing global challenge of the 21st century, affecting social, economic, food and environmental systems. Rising temperatures have disrupted the water cycle, leading to more frequent and severe extreme weather events and climate-induced disasters such as storms, floods and droughts worldwide. From 2015 to 2023, these years recorded the highest global temperatures, with the mean temperature reaching 1.45°C ± 0.12°C above the long-term average. Consequently, the global economy faces an annual average loss of approximately USD 266 billion[1] due to extreme weather and climate-related events. Global food production has been significantly impacted, with 65% of average annual losses attributed to droughts and 20% to floods , threatening food security and nutrition.

Cambodia is among the nations most affected by climate-induced disasters, particularly floods and droughts. According to the 2023 World Risk Index (WRI)[2], Cambodia ranked 65th among 142 countries in term of risk to natural disasters. Floods and droughts have caused substantial social and economic losses. For instance, the 2020 floods, one of the largest in a decade, affected nearly 810,000 people across 14 of Cambodia's 25 provinces[3], costing approximately USD 104.7 million in losses and damages[4]. In 2015/2016, water shortages caused by drought affected 2.5 million people across 18 provinces, reducing household agricultural production by 22% and income by 19%. In addition to these recurring disasters, socioeconomic vulnerabilities among households have been exacerbated by the COVID-19 pandemic and recent surges in global food and fuel prices. These factors have lessened household resilience to future shocks, hindering medium- and long-term national social and economic development goals.

1.2 MOTIVATION

The Government of Cambodia recognizes the threats posed by floods and droughts to economic growth. Managing the risks of climate change and disasters is a core element of the national development agenda. The government's Pentagonal Strategy, specifically Pillar 4, focuses on resilient, sustainable, and inclusive development, promoting resilient society, sustainable environmental and natural resources management, and green economy. To achieve these goals, the government has developed and adopted key documents, including the National Strategic Development Plan (2019-2023), the National Action Plan on Disaster Risk Reduction (2024-2028), the Disaster Risk Financing Strategy (2022-2023), and the Shock-Responsive Social Protection Framework.

Effective implementation of these key strategic documents requires strong coordination across multiple sectors and stakeholders involved in disaster risk management and social protection systems. Achieving this coordination hinges on a thorough understanding of disaster risk to support informed decision-making. However, granular disaster risk information at the commune level is lacking in Cambodia. This study aims to address that gap by generating evidence to better understand flood and drought risks, ultimately contributing to risk reduction and enhancing resilience to these disasters.

1.3. OBJECTIVES

The primary objective of this study is to assess the risks of floods and droughts in Cambodia. Specifically, it analyzes and maps out socioeconomic vulnerability of Cambodian households and the risks they face from floods and droughts at the commune level, utilizing artificial intelligence (AI) and geospatial data. The results provide valuable insights for planning, financing, and targeting interventions within the frameworks of Disaster Risk Reduction (DRR), Shock-Responsive Social Protection (SRSP), Anticipatory Actions (AA), and Early Warning for All (EW4All).

[1] Global economic losses from weather catastrophes 2007-2021

[2] The World Risk Index 2023 assesses the disaster risk for 193 countries. It covers all UN-recognized countries and more than 99 percent of the world's population.

[3] The Humanitarian Response Forum (HRF) Flood Situation Reports from 2020

[4] Economic impact from the UNDP's Disaster Financial Preparedness Analysis Report, 2023

2. METHODOLOGY

2.1. ANALYTICAL FRAMEWORK

The study employed the risk conceptual framework outlined in the 5th Assessment Report of the Intergovernmental Panel on Climate Change, explaining risk as the interconnection of hazards, exposure, and vulnerability. Integrating these three components provides a comprehensive risk understanding, not just identifying potential disaster events (hazards) but also examining who or what might be affected (exposure) and their susceptibility to effects (vulnerability). Figure 1 illustrates the key concepts of this framework.

The study focused on floods and droughts, reportedly causing significant social and economic loss and damage in Cambodia. The lowlands and plains surrounding the Tonle Sap Lake/River and along Mekong River are the most flood-prone areas, while the probability of severe droughts remains moderate. Furthermore, these hazards were projected to increase in frequency, extent, and severity due to climate change. Future impacts driven by climate change could compromise the country's GDP, with estimated reduction of between 3.0% and 9.4% by 2050.

Exposure refers to the presence of populations, assets, livelihoods, and infrastructure in areas prone to floods and droughts. In Cambodia, large portion of the population and key infrastructure is located in low-lying and flood-prone regions. Approximately 80% of the population is exposed to floods. Agriculture, the main source of food and income for many populations, is markedly sensitive to climate-related disasters, particularly given heavy reliance on rain-fed farming practice. More than 30% of agricultural land is exposed to drought.

Vulnerability is defined as a household's socioeconomic capacity to cope with, adapt to, and recover from the adverse impacts of floods and droughts. Recent studies have shown that households in flood-prone areas of Cambodia are likely to be vulnerable due to limited social and economic capacity and the use of unfavorable coping strategies. Droughts in 2015/2016 led to agricultural losses and income declines, negatively impacting food security, and increasing vulnerability.

Figure 1: The Risk Conceptual Framework

In this study, risk is defined as the probability of adverse socioeconomic consequences resulting from potential floods and droughts that are likely to affect the exposed population, livelihoods, and infrastructure, given their current socioeconomic coping capacities and resilience. The study made use of a data-driven index approach consuming geospatial datasets and household survey data to assess these three components. The results are aggregated at the commune level and visualized in maps.

2.2. DATA

The study utilized geospatial and household survey datasets. The household survey data came from the 2021 Cambodia Socio-Economic Survey (CSES), while geospatial data was sourced from open repositories.

2.2.1 SURVEY DATA

The 2021 CSES dataset, conducted by the National Institute of Statistics (NIS) under the Ministry of Planning, provides a wide range of information on households, including demographics, consumption, food security, housing, education, health, employment, agriculture, income, and migration. This rich dataset was ideal for analyzing household socioeconomic vulnerability in this study.

2.2.2 GEOSPATIAL DATA

Geospatial data[1], including both time-series and static datasets, were used in models to assess hazards, exposure, and vulnerability across country. Annex 1 provides a detailed description of the geospatial data used in the study.

2.3. COMPUTATION

The study followed four steps to assess flood and drought risks:

1. Flood and drought hazard assessment
2. Exposure quantification
3. Socioeconomic vulnerability assessment
4. Risk assessment, combining the hazard, exposure, and vulnerability components.

Machine learning (ML) [2] and spatial analysis techniques within Geographic Information System (GIS) and Remote Sensing (RS) technologies are capable to handle the complexity and large volume of data analysisRandom forests (RF)[3] is a well-known supervised machine learning algorithm, particularly useful by the reason of its accuracy in classifying information across various applications, including land-use mapping, flood risk assessment and, child malnutrition prediction , and poverty estimation. However, using ML and geospatial technologies comes with challenges, such as needing good-quality training data. Without accurate on-the-ground data, ML models can produce unreliable results. Furthermore, there is a need for more transparency, ownership, and trust in these technologies among stakeholders.

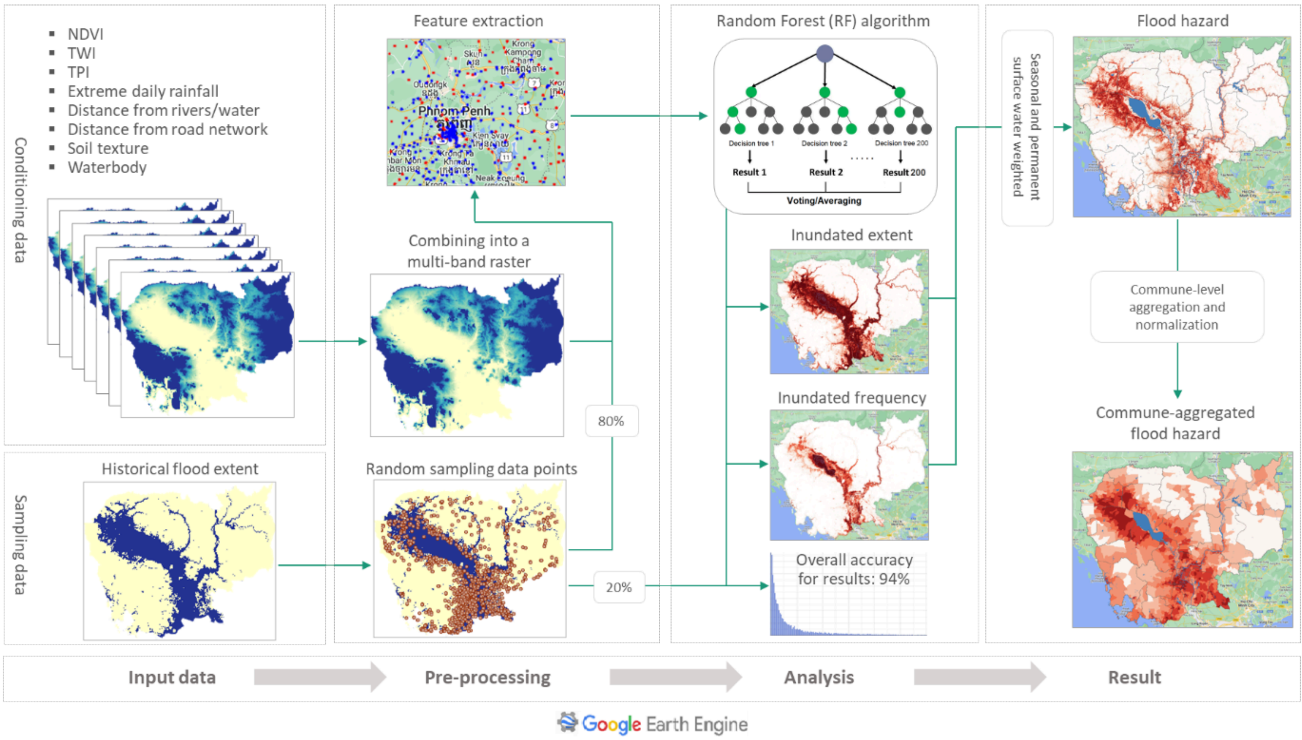
Additionally, advanced cloud-based computational platforms such as publicly available Google Earth Engine (GEE)[4] are beneficial for processing large datasets. The study leveraged machine learning RF model alongside spatial analysis tools on the GEE platform for geospatial data processing and analysis to assess flood and drought hazards, exposure, socioeconomic vulnerability, and risk.

2.3.1 HAZARDS

FLOOD

Flood hazard was assessed using a machine learning model applied to past flood events and factors characterizing flood susceptibility derived from geospatial data (Details in Annex 2). The assessment covered the peak flood season (August to November) over the past 11 years to accounts for both riverine and flash floods. The model produced pixel-level[5] indices of inundation locations and frequency. The final flood probability index was generated by adjusting for areas with permanent surface water using a geometric aggregation approach[6]. Figure 2 illustrates the workflow for flood hazard assessment.

Figure 2: The modeling workflow for assessing flood hazard



DROUGHT

Drought hazard was quantified using multiple drought indicators (Details in Annex 3), including Standardized Precipitation Index (SPI), Temperature Condition Index (TCI), Vegetation Condition Index (VCI), and Normalized Difference Water Index (NDWI). The assessment considered wet and dry seasonal variability from 2000 to 2022. These indicators were normalized[7] and combined[8] to produce pixel-level probability of a multiple-drought index. Figure 3 presents the workflow for drought hazard assessment.

2.3.2 EXPOSURE

This study examined four key elements exposed to flood and drought hazards: population, agricultural cultivation areas, buildings, and infrastructure (e.g., roads, airports, ports, schools, healthcare facilities, irrigation systems and reservoirs, and electricity grids and stations). Advances in geospatial technology have made precise, up-to-date data on these elements more accessible, derived from satellite imagery using machine learning algorithms trained with field data. To assess exposure in this study, indices were created by normalizing and combining[9] data of these elements into a single index. The workflow for assessing exposure is outlined in Figure 4.

Figure 3: The modeling flowchart for assessing drought hazard

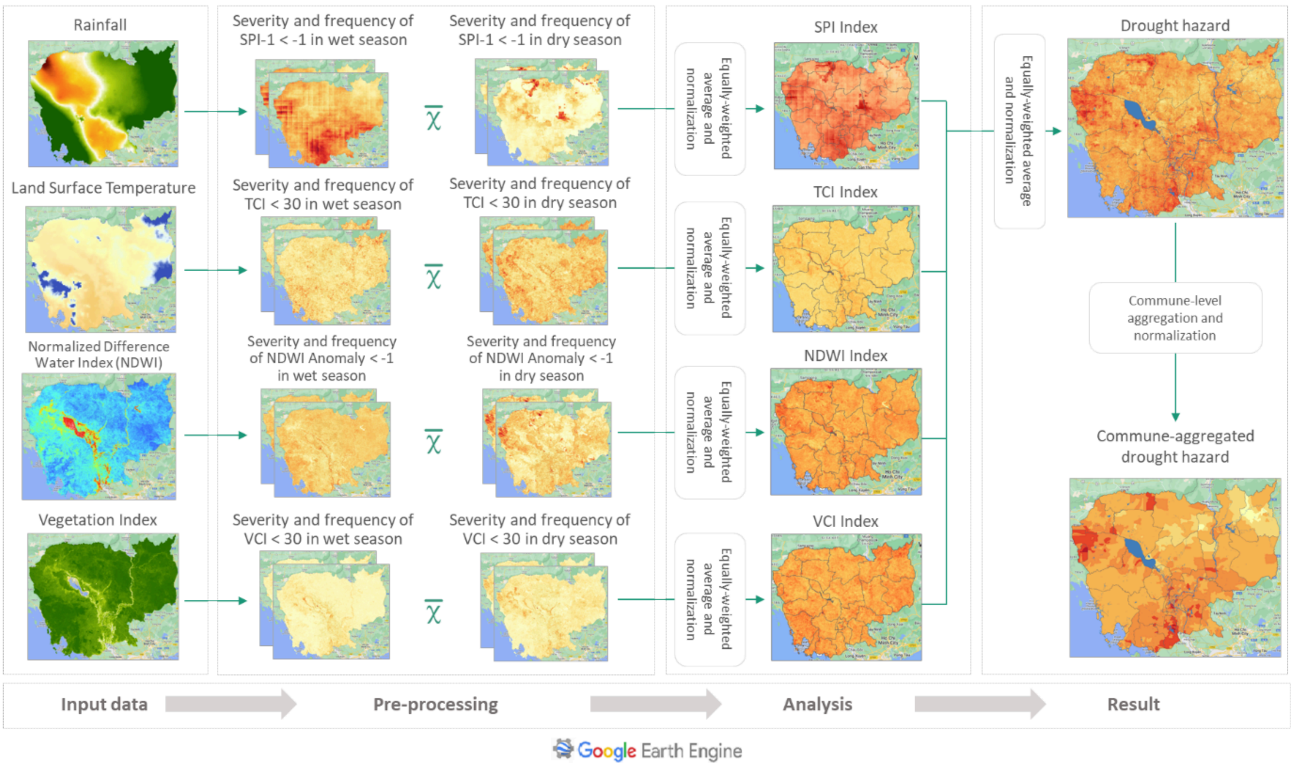
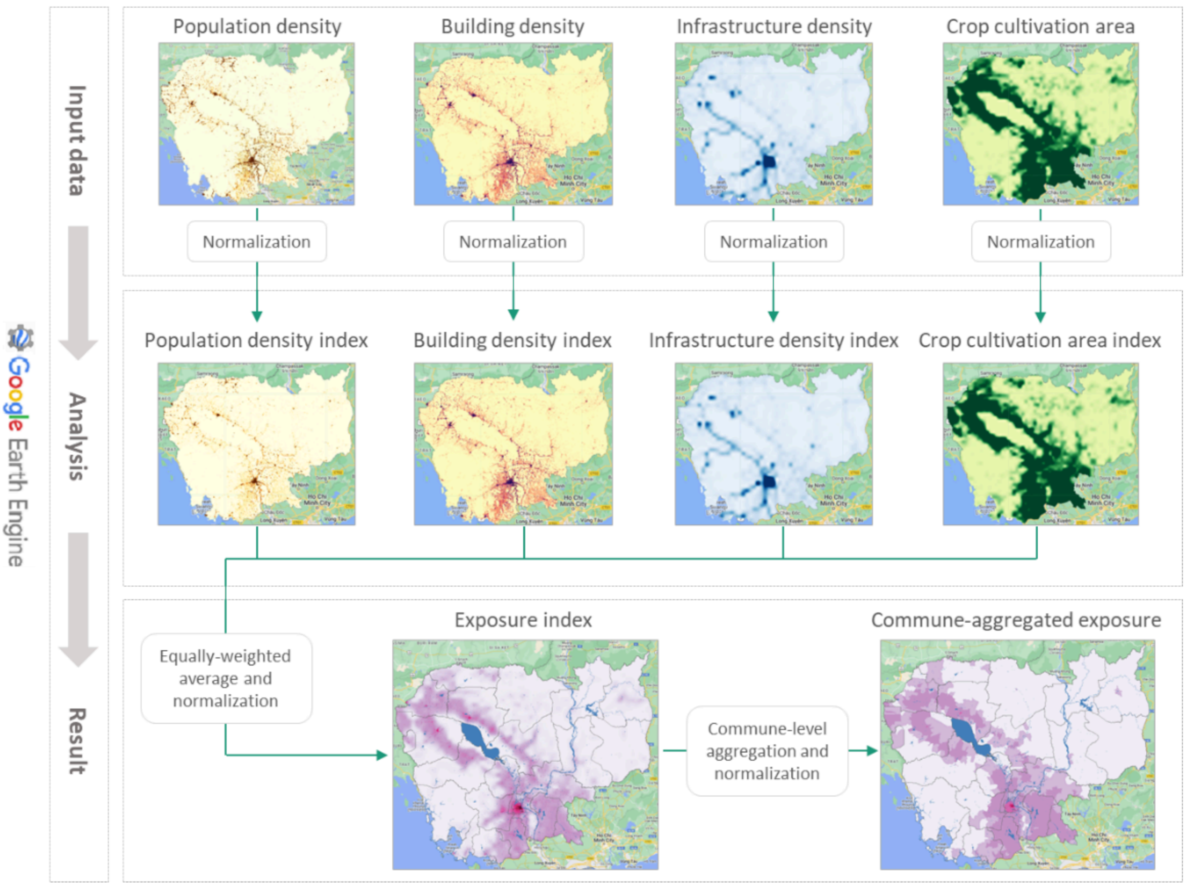


Figure 4: The workflow for assessing exposure



2.3.3 VULNERABILITY

In this study, vulnerability is defined as the household's social and economic capacities and resilience to cope with, adapt to and recover from the floods and droughts. While vulnerability can be assessed using national CSES data, this survey often lacks the statistical representation needed for granular analysis at the commune level. To address this gap, the study develops a model for assessing socioeconomic vulnerability in three main phases, with further details in Annex 4:

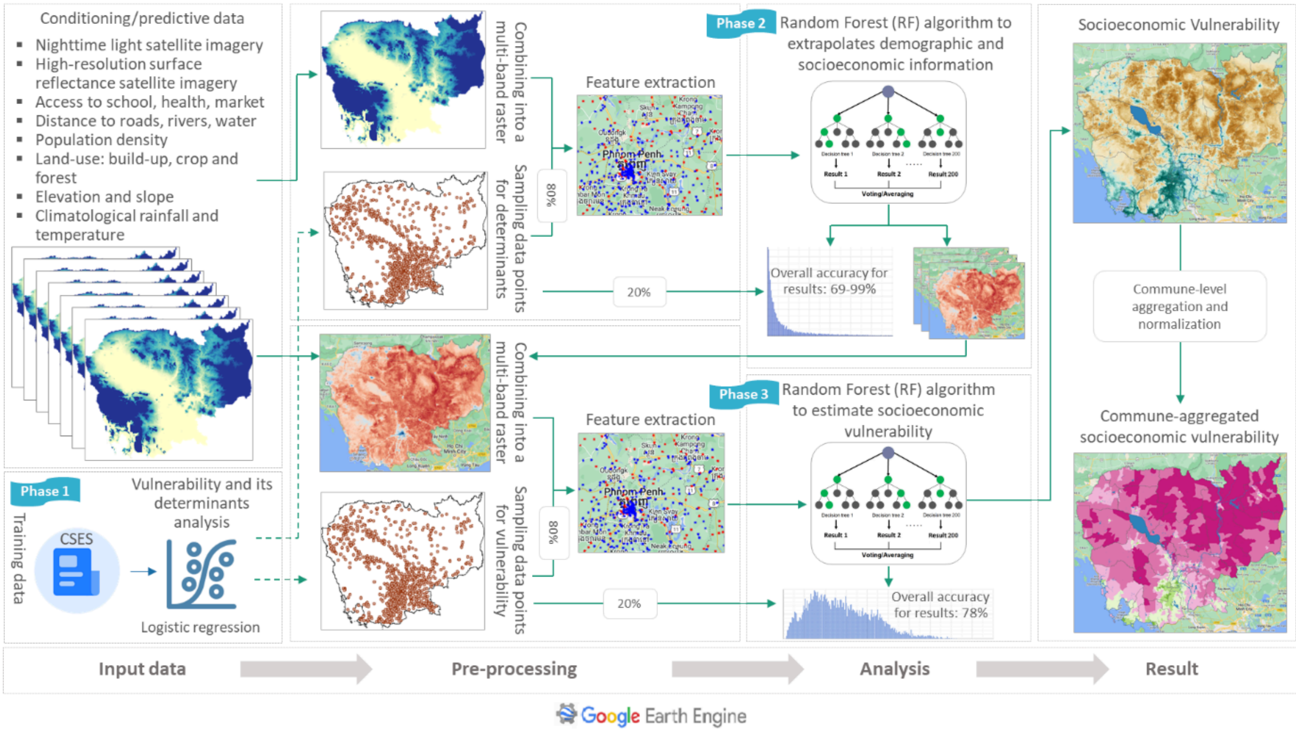
Phase 1: Household vulnerability is measured through indicators assessing economic capacity to meet essential needs, food consumption levels, and food- and livelihood-based coping strategies using the CSES data. Relationships between household demographic and socioeconomic characteristics and their vulnerability status are then analyzed.

Phase 2: A machine learning model, trained with the demographic and socioeconomic factors identified in Phase 1, and input with geospatial data (e.g., nighttime light intensity, human settlement patterns, accessibility to essential facilities, land cover, surface water, topography, climatology, etc.) performs geospatial extrapolation of vulnerability determinants.

Phase 3: A nationwide socioeconomic vulnerability model is created by applying a machine learning algorithm to the findings from Phase 1 on household vulnerability status, along with the extrapolated demographic and socioeconomic data and geospatial datasets from Phase 2.

The final output of the model was a pixel-level probabilistic index of socioeconomic vulnerability. The workflow for this vulnerability assessment process is illustrated in Figure 5.

Figure 5: The modelling workflow for socioeconomic vulnerability assessment



2.3.4 RISK

Risk was computed by multiplying indices of flood and drought hazards, exposure, and socioeconomic vulnerability. The results are produced at the pixel level.

2.4. COMMUNE-LEVEL AGGREGATION

To effectively communicate results for strategic and operational planning, the pixel-level outputs were aggregated at the commune level. The indices were normalized and classified into six categories[10] for visualization on maps.

2.5 POPULATION AND AGRICULTURAL LAND ESTIMATION

Population and agricultural land at risk of flood and drought were estimated by overlaying pixel-based population and agricultural land data with risk information. The figures were then aggregated at the provincial level.

1 Geospatial data refers to information describing location, attribute characteristics, and temporal dimensions of objects, events, or phenomena on the Earth's surface. It is gathered using geographic positioning systems (GPS), Earth observation (EO)-satellite/airborne sensing platforms, and traditional surveys/censuses.

2 Machine learning (ML) is a subfield of Artificial Intelligence (AI) that utilizes statistical algorithms to learn patterns from the relationship between conditioning and sampling data gradually to accurately predict values for areas where data is unavailable.

3 Random Forest (RF) algorithm creates multiple decision trees by randomly selecting variables and sampling data.

4 Google Earth Engine (GEE) platform stores geospatial data from numerous sources, allows for upload of external data, and includes a number of built-in algorithms such as the machine learning random forests (RF).

5 Spatial resolution of 50 meters.

6 The formula involves multiplying all the values in the dataset.

7 This approach transforms all values to scores ranging from 0 to 1 by subtracting the minimum score and dividing it by the range of the indicator values. This process makes the scores unitless, allowing for comparison or computation across various indicators.

8 Sum of all the values in the dataset is divided by the total number of values.

9 Equally weighted arithmetic mean.

10 They include 1. Minimal (0 - 0.1), 2. Very low (0.1 - 0.2), 3. Low (0.2 - 0.4), 4. Moderate (0.4 - 0.6), 5. High (0.6 - 0.8), and 6. Very high (0.8 - 1.0).

3. FINDINGS

3.1. HAZARDS

The flood hazard assessment indicates that communes in the low-lying areas surrounding the Tonle Sap Lake/River, specifically in Banteay Meanchey, Battambang, Pursat, Kampong Chhnang, Kampong Thom and Siem Reap, have a high likelihood of floods. Similarly, communes along the Mekong River, particularly in Stung Treng, Kratie, Tboung Khmum, Kampong Cham, as well as those in the southern plains covering Kandal, Phnom Penh, Kampong Speu, Takeo, Prey Veng, and Svay Rieng, are also highly prone to flooding. These floods are primarily caused by overflow from the lakes and rivers, as well as excessive rainfall in the catchment areas of their tributary rivers and streams.

Additionally, flash floods, driven by extreme rainfall occasionally occur in highland and mountainous communes in provinces such as Rattanak Kirir, Preah Vihear, Koh Kong, Preash Sihanouk and Kampot. Figure 6 (left) visualizes flood hazards across the country.

Drought hazards are widespread across the country, primarily resulting from insufficient or prolonged rainfall combined with rising temperatures during both the dry and wet seasons. These conditions stress surface water availability and adversely affect vegetation. The assessment identifies three main hotspots demonstrating a moderate to very high probability of drought: the northwestern provinces such as Pursat, Battambang, Pailin, Banteay Meanchey, Otdar Meanchey, and Siem Reap, the southwestern provinces including Kandal, Phnom Penh, Kampong Speu, Takeo, Kampot, Kep, and Preah Sihanouk, and the southeastern provinces such as Svay Rieng, Prey Veng, Kampong Cham, Tboung Khmum, and Kratie. Figure 6 (right) visualizes drought hazards across the country.

Figure 6: Maps of flood (left) and drought (right) hazards at commune level

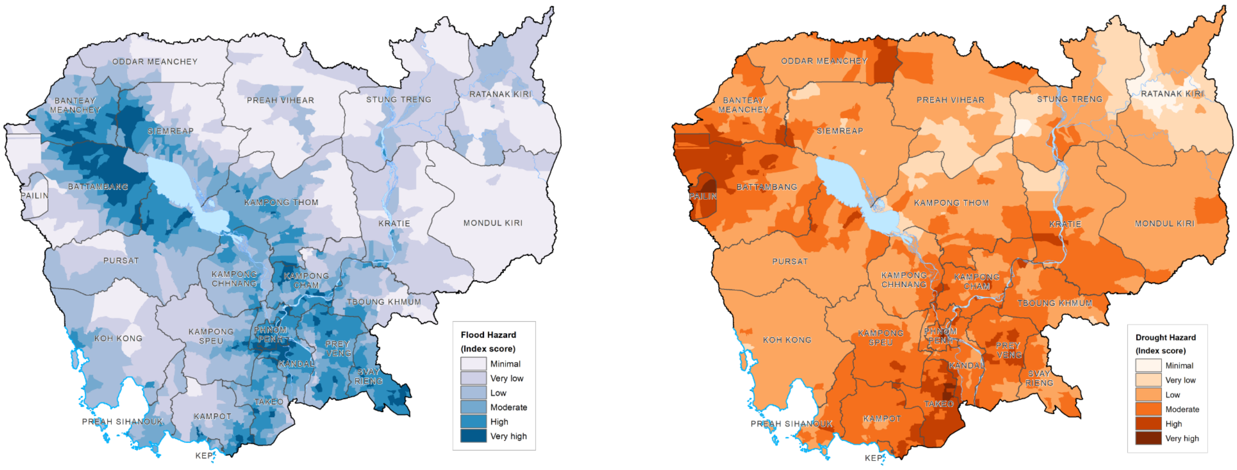
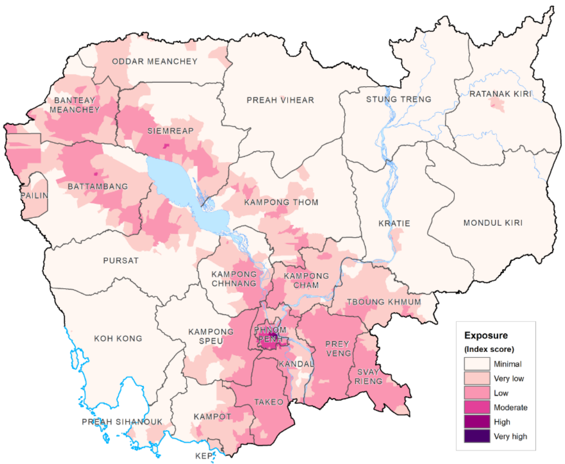


Figure 7: Map of commune-level exposure information



3.2. EXPOSURE

Exposure, in terms of population, buildings, agricultural cultivation, and infrastructure, is notably concentrated in communes surrounding Tonle Sap Lake/River and in the southern plains (Figure 7). Communes within provincial towns present moderate to very high population, building, and infrastructure densities, with Phnom Penh having the highest levels of exposure.

In contrast, communes in rural and mountainous areas, particularly in Kratie, Mondulkiri, Ratanakiri, Stung Treng, Preah Vihear, Otdar Meanchey, and Koh Kong show minimal to very low exposure. This lower exposure is primarily attributed to smaller population and underdeveloped infrastructure in these areas.

3.3. SOCIOECONOMIC VULNERABILITY

The binary logistic regression analysis[1] indicate that households with the following demographic and socioeconomic characteristics are more likely to be socioeconomically vulnerable:

- Larger family size, likely associated with higher dependency ratio and elderly with disability
- Female-headed households
- Lower educational levels among household heads and members
- Primary income sources from low/unskilled labor and farming
- Limited access to essential services such as electricity, clean drinking water, water irrigation for agriculture, and sanitation
- Deprivation in housing conditions (i.e., poor construction materials for floor, wall and roof; and overcrowding)
- Lack of assets (e.g., agricultural land, transportation means, improved cooking materials etc.)
- Residence in flood- and drought-prone areas
- Current beneficiaries of social assistance programmes, particularly IDPoor households.

Table 1 presents details of household characteristics that were identified as significant predictors of socioeconomic vulnerability.

Table 1: Households' demographic and socioeconomic determinants of vulnerability

Variable	Odds ratio	Coeffi-cients	Sig.	Explanation
Headed by a female	1.204	.186	.006	Female-headed households are more vulnerable than male-headed ones.
Number of household members	1.550	.438	<.001	Larger households are more likely to be vulnerable.
With high dependency ratio (>1)	1.312	.272	<.001	Higher dependency ratios increase vulnerability.
Has a member with disabilities	1.260	.231	<.001	Households with a member with disabilities are more vulnerable
Household head did not complete more than 6 years in school	1.530	.425	<.001	Households that head and members with lower educational attainment are more vulnerable.
Household members did not complete more than 6 years in school	1.675	.516	<.001	
Main economic activities: skilled businessman/worker			<.001	Households with main income source from low/unskilled work, petty business or farming (i.e., low/unskilled factory worker, construction worker, transportation worker, entertainment/restaurant worker, street food vendor, subsistence crop cultivation and animal raising, fishing, etc.) are more likely to be vulnerable compared to households with income from skilled work or formal business (i.e., salary-based/government officials, business owner/manager, trading, etc.).
Main economic activities: un/low skilled businessman/worker	1.484	.394	<.001	
Main economic activities: farmer and farming worker	1.857	.619	<.001	
Without agricultural land			<.001	Households engaging in farming and owning land are less vulnerable. Vulnerability further decreases when agricultural land is connected to an irrigation system.
Has agricultural land connected to irrigation system	.659	-.417	<.001	
Has agricultural land not connected to irrigation system	.737	-.305	<.001	
Has transportation mode <= 1	2.594	.953	<.001	Households without or with a means of transportation (typically a motorbike) is very likely to be vulnerable.
With more than 3 persons sharing a room	1.783	.578	<.001	Households having small living spaces tend to be more vulnerable.
Live in a house with poor wall construction material	1.274	.242	.008	Households living in houses with roofs, floors, or walls made of poor construction materials (i.e., not concrete, tile, or wood) are very likely to be vulnerable.
Live in a house with poor roof construction material	2.540	.932	<.001	
Live in a house with poor floor construction material	1.328	.284	<.001	
Has unimproved cooking materials	2.115	.749	<.001	Households that do not own improved food preparation and cooking equipment such as a refrigerator/freezer, electric or gas stove, or dining sets is more vulnerable.
Has inaccessibility to electricity	1.661	.507	<.001	Households that do not have access to electricity are highly vulnerable.
Does not have access to social protection services or assistance	.431	-.841	<.001	Household with access currently to social protection service or assistance has more likelihood of socioeconomic vulnerability.
Has unimproved source of drinking water	1.089	.120	.050	Households without access to safe and clean source of drinking water is likely to be vulnerable.
Has unimproved sanitation facility	1.134	.126	.048	Households without improved sanitation facility (i.e., flash toilet) are likely to be vulnerable.
Living in flood-prone areas	1.274	.242	.002	Households living in flood-prone areas are more vulnerable.
Living in areas with trends of increasing rainfall	.633	-.458	<.001	Households living in areas with tendency for increasing rainfall is less likely to be socioeconomic vulnerability. Although, excessive rainfall can also lead to flooding and crop damage, increasing vulnerability.
Constant	.012	-4.465	<.001	
Nagelkerke R Square	0.37			
Number of observations	10,080			

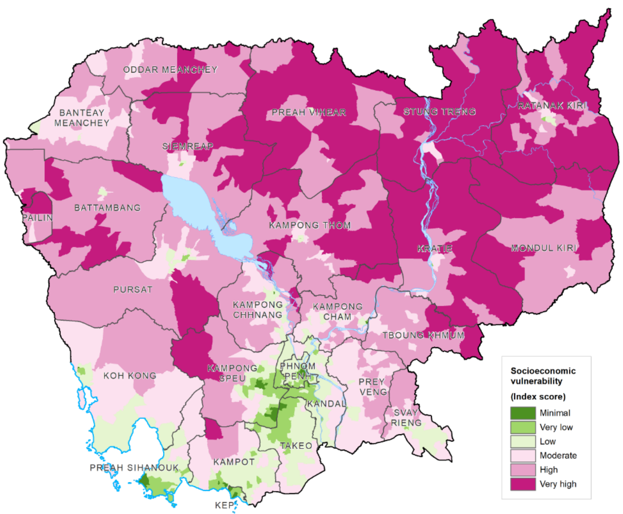
The spatial distribution analysis reveals that households with low socioeconomic vulnerability are predominantly situated in urban communes. These include Phnom Penh and its neighboring provinces—Kandal, Takeo, and Kampong Speu—as well as the coastal provinces of Kampot, Kep, Preah Sihanouk, and Koh Kong, along with other provincial towns. The lower vulnerability in these areas is largely attributed to substantial investments in infrastructure, industrial development, and the service sectors. These have significantly boosted social and economic development, improving household's access to essential services and more diverse economic opportunities.

In contrast, rural and mountainous areas show higher socioeconomic vulnerability, mainly because of limited economic diversification. Households living in communes in these areas rely heavily on agriculture and natural resource-based likelihoods and face challenges associated with underdeveloped infrastructure and access to essential services. A nationwide geographical overview of socioeconomic vulnerability is illustrated in Figure 8.

3.4. RISK

As outlined above, flood and drought risk were quantified by intersecting of flood and drought hazards, exposure, and socioeconomic vulnerability.

Figure 8: Map of commune-level socioeconomic vulnerability



3.4.1. FLOOD RISK

Flood risk is primarily driven by the high probability of recurrent floods in areas where populations, infrastructure, and agricultural land are considerably exposed, exceeding their socioeconomic capacities to cope. Geographically, moderate to very high flood risk levels are observed in communes around Tonle Sap Lake/River, particularly in Kampong Chhnang, Pursat, Battambang, Banteay Meanchey, Siem Reap, Kampong Thom. Likewise, communes along the Mekong River, such as in Kampong Cham and Tboung Khmum, as well as those in the southern plains, including Kandal, Phnom Penh, Takeo, Prey Veng, and Svay Rieng, also face high flood risk levels.

In contrast, communes in Preah Vihear, Stung Treng, Ratanakiri, and Kratie face low flood risk, despite increased likelihoods of flooding and socioeconomic vulnerability. This is primarily due to lower exposure to flood-prone areas. Similarly, communes in Otdar Meanchey and Mondulkiri show low flood risk, driven by lesser exposure and flood probabilities, regardless of their higher socioeconomic vulnerability. Communes in Kampong Speu and along the coastline, such as in Koh Kong, Preah Sihanouk, Kampot, and Kep are also at low risk to flooding, attributed to their greater socioeconomic capacities to cope. Figure 9 provides a visual map of flood risk across the country.

Nationally, approximately 15.2% of the total population, or 2.7 million people, are at risk from flooding. By demographic breakdown, 15.4% of women (1.4 million women), 14.4% of children under 5 (272,000 children), and 16.4% of elderly individuals (162,100 people) are at risk of flooding. An estimated 16.1% of agricultural land, equating to about 860,000 hectares, is also at risk. Table 2 presents a province-level breakdown of the population and agricultural land at risk of flooding.

Figure 9: Map of flood risk at commune level

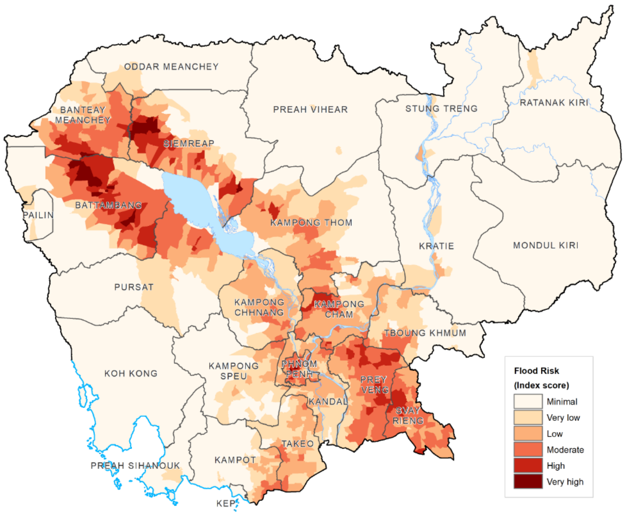


Figure 10: Map of drought risk at commune level

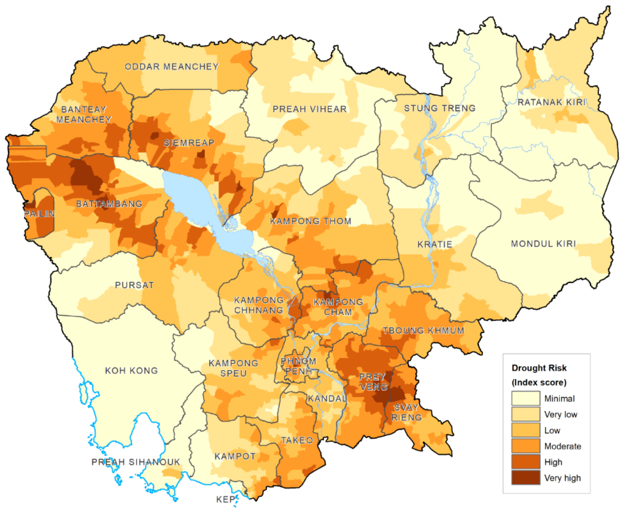


Table 2: Province-level estimation of population and agricultural land at risk of flooding

Code	Province	Floods Post Risk to				
		Population (%)	Women (%)	Children <5 (%)	Elderly (%)	Agricultural Land (%)
1	Banteay Meanchey	20.2	20.3	19.8	23.0	24.0
2	Battambang	18.8	19.1	17.6	23.3	24.3
3	Kampong Cham	17.8	17.9	17.7	18.7	19.9
4	Kampong Chhnang	13.4	13.4	13.4	13.4	15.1
5	Kampong Speu	8.3	8.3	8.0	8.4	7.9
6	Kampong Thom	19.2	19.3	18.4	20.6	17.1
7	Kampot	6.8	6.9	6.6	7.6	12.3
8	Kandal	17.0	17.1	17.1	16.9	14.4
9	Koh Kong	2.2	2.2	2.2	2.2	2.5
10	Kratie	7.6	7.7	7.2	9.0	5.0
11	Mondul Kiri	2.2	2.2	2.3	2.2	1.5
12	Phnom Penh	22.6	22.8	23.9	21.6	31.4
13	Preah Vihear	2.2	2.2	2.2	2.4	1.8
14	Prey Veng	26.0	26.1	26.2	26.0	30.0
15	Pursat	15.3	15.4	14.9	16.3	20.1
16	Ratanak Kiri	2.8	2.8	2.7	2.9	2.4
17	Siem Reap	22.2	22.3	21.1	23.5	16.8
18	Preah Sihanouk	3.2	3.2	3.1	3.3	5.7
19	Stung Treng	7.2	7.3	7.3	7.5	2.9
20	Svay Rieng	24.9	24.9	25.2	24.9	31.1
21	Takeo	10.0	10.0	10.0	9.8	15.8
22	Otdar Meanchey	4.1	4.2	4.1	4.3	3.1

Code	Province	Floods Post Risk to				
		Population (%)	Women (%)	Children <5 (%)	Elderly (%)	Agricultural Land (%)
23	Kep	2.5	2.5	2.5	2.5	4.5
24	Pailin	7.5	7.5	7.4	7.7	4.9
25	Tboung Khmum	8.7	8.7	8.4	9.7	12.4
	National	**15.2**	**15.4**	**14.4**	**16.4**	**16.1**

3.4.2. DROUGHT RISK

The assessment indicates that droughts pose a more widespread risk compared to floods (Figure 10). Communes with moderate to very high drought risk are largely clustered around Tonle Sap Lake/River, along the Mekong River, and across the southern plains. In particular, these communes are located in Kampong Chhnang, Pursat, Battambang, Pailin, Banteay Meanchey, Otdar Meanchey, Siem Reap, Kampong Thom, Kampong Chhnang, Kampong Cham, Tboung Khmum, Kandal, Phnom Penh, Kampong Speu, Takeo, Prey Veng, and Svay Rieng. The main drivers of drought risk in these communes are the bigger chances for rainfall deficits and hotter temperatures, which contribute to surface water shortages and stress vegetation. These conditions, combined with a high concentration of population, infrastructure, and agricultural activities, overwhelms the socioeconomic coping and adaptive capacities in these communes.

Nationally, drought threatens approximately 29.2% of the total population (about 5.1 million people), 29.3% of women (2.7 million women), 29.1 % of children under 5 years old (548,400 children), and 29.2 % of elderly people aged over 65 years old (Roughly 288,700 people). Additionally, about 33.3% of the total agricultural land (about 1.8 million hectares) is at risk of drought. Table 3 provides a province-level breakdown of the population and agricultural land at risk.

Table 3: Province-level estimation of population and agricultural land at risk of drought

Code	Province	Droughts Post Risk to				
		Population (%)	Women (%)	Children <5 (%)	Elderly (%)	Agricultural Land (%)
1	Banteay Meanchey	32.2	32.2	32.2	33.7	38.7
2	Battambang	42.5	42.6	42.3	43.3	45.0
3	Kampong Cham	29.8	29.8	30.1	29.8	40.4
4	Kampong Chhnang	27.6	27.6	27.6	27.8	34.8
5	Kampong Speu	20.9	20.9	21.0	20.9	23.2
6	Kampong Thom	32.8	32.8	32.3	34.0	34.6
7	Kampot	20.2	20.3	20.2	21.1	26.0
8	Kandal	22.3	22.3	22.5	22.1	27.1
9	Koh Kong	7.6	7.6	7.6	7.8	10.6
10	Kratie	19.7	19.7	19.4	20.3	19.6
11	Mondul Kiri	14.1	14.1	14.1	14.1	15.2
12	Phnom Penh	30.9	31.1	31.9	29.6	34.7

Code	Province	Droughts Post Risk to				
		Population (%)	Women (%)	Children <5 (%)	Elderly (%)	Agricultural Land (%)
13	Preah Vihear	15.1	15.1	15.1	15.2	16.2
14	Prey Veng	39.8	39.9	40.0	39.7	48.0
15	Pursat	27.4	27.5	27.3	27.8	32.6
16	Ratanak Kiri	12.5	12.5	12.5	12.4	15.6
17	Siem Reap	35.5	35.6	34.9	36.5	33.4
18	Preah Sihanouk	9.3	9.4	9.3	9.5	14.4
19	Stung Treng	14.9	15.0	15.1	15.2	15.7
20	Svay Rieng	38.5	38.6	39.2	38.9	41.7
21	Takeo	21.8	21.7	21.8	21.7	34.7
22	Otdar Meanchey	24.7	24.7	25.0	24.7	23.5
23	Kep	9.4	9.4	9.5	9.4	13.0
24	Pailin	40.9	41.0	40.6	41.4	51.0
25	Tboung Khmum	31.3	31.4	31.2	31.6	37.1
	National	**29.2**	**29.3**	**29.1**	**29.2**	**33.3**

[1] The model has acceptable predictive reliability, as indicated by a Nagelkerke R Square value of 0.37.

4. CONCLUSION

Cambodia faces heightened risks from climatic disasters, particularly floods and droughts, driven by both socioeconomic and environmental factors. A thorough understanding of flood and drought risks is essential for informed decision-making in disaster planning, financing, and targeting interventions within disaster risk management and social protection systems. Assessing these risks hinges on the integration of flood and drought hazards with data on exposure (including population, buildings, infrastructure, and agricultural land) and socioeconomic vulnerability. Leveraging advanced technologies—such as machine learning algorithms and spatial analysis applied to geospatial and survey-based datasets on cloud platforms like Google Earth Engine (GEE)—this assessment provides a comprehensive and precise risk insights down to the commune level. The study's key results indicate:

HAZARDS

Flooding frequently occurs in the commune surrounding the Tonle Sap Lake/River, along the Mekong River, and in the southern plains, primarily due to lake and river overflows and excessive rainfall. Flash floods resulting from extreme rainfall are also increasingly common in highland and mountainous communes. Droughts, on the other hand, are driven by rainfall deficits and rising temperature, leading to surface water shortages and stress on vegetation. These drought conditions are particularly pronounced in three hotspots: the northwestern, southwestern, and southeastern regions of the country.

EXPOSURE

Despite being highly prone to floods and droughts, communes around Tonle Sap Lake/River, along the Mekong River, and in the southern Plains are home to large populations, critical infrastructure, and main agricultural activities, making them substantially exposed. In contrast, highland/mountainous communes show lower exposure due to smaller populations and less developed infrastructure.

VULNERABILITY

Socioeconomically vulnerable households are characterized by larger family, female heads, lower educational levels, dependence on farming or low-skilled labor, limited access to essential services/facilities, poor housing conditions, a lack of assets, residence in areas prone to climatic disasters, and reliance on social assistance. These households predominantly reside in rural and highland/ mountainous communes. Whereas households in urban communes, with better access to essential services and more diverse economic opportunities, tend to have lower vulnerability.

RISK

Risk arises in areas where populations, critical infrastructure and agricultural activities are highly exposed to floods and droughts, surpassing their socioeconomic capacity to cope and adapt. Low-lying communes adjacent to the Tonle Sap Lake/River, along the Mekong River, and in the southern plains have a dual burden of being at risk of both floods and droughts. Nationally, it is estimated that floods pose risks to 15.2% of the population and 16.1% of agricultural land, while droughts threaten 29.2% of the population and 33.3% of agricultural land.

RECOMMENDATIONS

To effectively reduce risk and build resilience against future climate-induced disasters, the following key actions are recommended for the government and development partners to enhance preparedness efforts for responses:

INTEGRATE RISK INFORMATION INTO DISASTER MANAGEMENT SYSTEMS:

To align with the National Action Plan for Disaster Risk Reduction (NAP-DRR) 2024-2028, which emphasizes the use of scientific knowledge to understand disaster risks, it is essential to incorporate the risk information into national and sub-national contingency plans, updated annually by NCDM and PCDMs. This will ensure risk-informed emergency preparedness and response while strengthening systematic institutional sustainability.

STRENGTHEN EARLY WARNING SYSTEMS (EWS):

Strengthening the capacity of MOWRAM to integrate socioeconomic vulnerability data into the national multi-hazard early warning systems is crucial for improving forecasting and communicating impacts of extreme weather events on population, livelihoods, and assets.

Enhancing NCDM's PRISM with integration of hydrometeorological hazard forecasting and socioeconomic data will strengthen real-time impact monitoring and forecasting, enabling the timely activation of emergency responses.

STRENGTHEN AND EXPAND SOCIAL ASSISTANCE RESPONSE MECHANISMS:

The application of the flood and drought risk data for ex-ante planning, financing, and targeting of social assistance interventions is essential for ensuring needs-based and well-resourced preparedness, facilitating predictable and timely disaster responses.

Expanding the social registry to include smallholder farmers in high-risk communes will enhance preparedness and response capacity.

OPERATIONALIZE DISASTER RISK FINANCING STRATEGY:

Risk information can support the implementation of the Disaster Risk Financing Strategy, particularly in designing parametric insurance schemes by defining premiums, payout thresholds, and action triggers.

INTEGRATE CLIMATE PROJECTIONS INTO RISK ASSESSMENT:

With escalating climate impacts, integrating localized climate projection data into assessment of climatic hazards strengthens risk modeling, providing forward-looking information, supporting efforts in mitigation and adaptation of floods and droughts.

ENHANCE RELIABILITY OF ARTIFICIAL INTELLIGENCE (AI) IN RISK MODELING:

High-quality training data are essential for reliable machine learning (ML) models. Incorporating geographical reference information into data collection—such as national surveys, censuses, and assessments—while adhering to data privacy standards will enhance the quality and accessibility of training data, improving the accuracy of ML outputs.

Establishing standardized procedures for capturing post-disaster impact data will enhance accuracy, consistency, and interoperability in loss and damage datasets. These data can then calibrate and validate risk models, enhancing the precision and robustness of risk assessments.

IMPLEMENT DYNAMIC SOCIOECONOMIC VULNERABILITY ASSESSMENT:

Given the evolving socioeconomic and climatic factors influencing household vulnerability, adopting the AI-driven vulnerability assessment model outlined in this report as a dynamic modeling is advantageous. This dynamic modeling approach integrates near-real-time Earth observation data and field-based socioeconomic data with ML predictive analytics, enabling timely updates to vulnerability data, which are essential for responsive risk assessment.

EMBRACE AI AND GEOSPATIAL TECHNOLOGIES:

As demonstrated in this study, ML and geospatial technologies can transform geospatial and survey-based datasets into granular insights on flood and drought risks. Strengthening government's institutional capacities at NCDM, MoWRAM, and NSPC to harness AI/ML and geospatial technologies for data analytics will enable the generation of precise, timely, and actionable information for disaster preparedness and response.

ANNEX 1: GEOSPATIAL DATASET SOURCES USED IN THE STUDY

The below table lists geospatial datasets utilized in the study.

No	Dataset	Description	Timescale	Source	Application in the study
1	Historical flood extent	Multi-year flood extent derived from satellite imageries.	2012 - 2022	ADPC and WFP	Training and validating information in model for flood hazards assessment
2	Global Surface Water	Information of extent and temporal distribution of global surface water from 1984.	2012-2022	JRC	Seasonal and permanent surface water
3	Digital Elevation Model (DEM)	Elevation data with forests and buildings removed at 30-meter spatial resolution.	2023	Fathom	Analyses of slope, Topographic Wet Index (TWI), and Topographic Position Index (TPI)
4	Moderate Resolution Imaging Spectroradiometer (MODIS)	Collection of global satellite data observing Earth's surface, atmosphere, and oceans.	2000 - 2022	NASA's LP DAAC	Analysis of Normalized Difference Vegetation Index (NDVI), Vegetation Condition Index (VCI), Temperature Condition Index (TCI), Normalized Difference Water Index (NDWI)
5	Rainfall	Collection of spatial rainfall data sets for more than 30 years	1981 - 2022	CHIRPS	Analysis of climatological rainfall, rainfall variability trends, 1-Month Standardized Precipitation Index (SPI-1) and extreme daily rainfall
6	Soil texture	Various top-layer soil texture types	2018	OpenGeoHub	Soil properties
7	Population	High resolution population density maps	2022	Meta Data for Good	Population
8	Building footprints	High resolution Building structure	2023	Google Research Open Buildings	Building density
9	Critical Infrastructure	Spatial information of infrastructure in transportation, energy, water, waste, telecommunication, education, and health sectors		awesome-gee-community-catalog	Infrastructure density
10	Landcover	Different classification of landcover	2021	ESA WorldCover	Agricultural land and forest areas
11	Nighttime light image	Earth observation information on light at night.	2021	NASA's LP DAAC	Predictive data in vulnerability assessment
12	High-resolution surface reflectance satellite imagery	High-resolution satellite monitoring tropical forests	2021	Planet-NICFI	Predictive data in vulnerability assessment
13	School facility	Location of school facility	2012	MoEYS	Accessibility to school
14	Health facility	Location of health facility	2010	MoH	Accessibility to health

No	Dataset	Description	Timescale	Source	Application in the study
15	Market location	Location of market facility	2016	WFP	Accessibility to market
16	Hydrology network	River and lake	2008	MoLMUPC	Distance to river and surface water
17	Road network	Primary, secondary and tertiary roads	2020	MoLMUPC	Distance to road

ANNEX 2: GEOSPATIAL DATASET USED IN FLOOD HAZARD ASSESSMENT

Flooding occurs when water spills over its typical boundaries, such as those of a river or lake, or when it inundates land that is normally dry . Flood hazard assessment applied a machine learning random forest model on geospatial datasets selected through empirical studies and stakeholder consultations. Approximately 100,000 data points were randomly generated from historical flood extents derived from satellite imagery, with each point classified as either flooded or not flooded for the purpose of training and validating the model. The key conditioning geospatial datasets used as independent variables in the model are summarized in the below table. These datasets were integrated into the model to analyze the relationships between these variables and the historical flood occurrences across all return periods, thereby predicting flood susceptibility nationwide.

No	Geospatial data	Explanation
1	Normalized Difference Vegetation Index (NDVI)	Dense vegetation is closely linked to increased water infiltration and reduced runoff, which in turn influences the assessment of flood-prone areas.
2	Topographic Wet Index (TWI)	It is derived from Digital Elevation Model (DEM) data, using flow accumulation and slope functions, to quantifies the topographic control on hydrological processes, contributing to the delineation of flooded basin areas .
3	Topographic Position Index (TPI)	It is derived from DEM data to understand the spatial pattern of landforms, such as ridges, flat plains, or valleys, which can indicate the potential for waterlogging .
4	Extreme daily rainfall	It is defined as the amount of rainfall in a day that exceeds the 95th percentile of daily rainfall , directly leading to a high probability of floods .
5	Distance from hydrological network	It significantly affects the time of concentration and the magnitude of water flow, increasing the likelihood of floods in areas near rivers and water bodies .
6	Distance from road network	It affects the movement of water flow through built-in drainage systems and road elevations, resulting in inundation in surrounding areas.
7	Soil texture	It contributes to infiltration rates and soil moisture content, influencing hydrological processes.
8	Seasonal and permanent surface water	It is considered to delineate flood hazard areas more accurately.

ANNEX 3: GEOSPATIAL DATASET USED IN DROUGHT HAZARD ASSESSMENT

A drought is defined as a period of abnormally dry weather characterized by a prolonged absence or deficiency of rainfall, leading to a serious hydrological imbalance . Droughts are typically classified into four categories:

» Meteorological drought caused by significantly lower-than-average precipitation over a specific period,

» Agricultural drought related to reduced soil moisture, affecting crop yields,

» Hydrological drought involves reduced water levels in rivers, lakes, and reservoirs, and

» Socioeconomic drought occurs when water shortages affect human activities, such as water supply for households, agriculture, and industry.

This study selected drought indicators based on empirical research to characterize drought conditions in Cambodia. These indicators, detailed in the table below, are used in combination to assess the severity and duration of droughts.

No	Drought indicator	Description
1	1-Month Standardized Precipitation Index (SPI-1)	It measures the observed rainfall deviation from the climatological average over a one-month period, which is useful for agricultural applications. A drought condition is characterized by a value less than or equal to -1.0 .
2	Temperature Condition Index (TCI)	It estimates the temperature condition, determining stress on vegetation , and an index value less than or equal to 30 indicates a drought-like situation .
3	Normalized Difference Water Index (NDWI)	It monitors changes in surface water content and . NDWI Anomaly is the deviation of the current NDWI from the long-term average, describing the situation as a drought with an index lower than or equal to -1.0 .
4	Vegetation Condition Index (VCI)	It observes changes in vegetation as the impact of droughts. An index value less than or equal to 30 is used to detect drought-like conditions .

ANNEX 4: DETAILING THE THREE PHASES FOR ASSESSING SOCIOECONOMIC VULNERABILITY

PHASE 1: QUANTIFYING HOUSEHOLD VULNERABILITY

Vulnerability in this study was quantified by a composition of proxy welfare indicators measuring household's economic capacity to access to essential needs[1], current food consumption[2], and food- and livelihood-based coping behaviors[3]. These indicators were derived from the 2021 Cambodian Socio-Economic Survey (CSES) data, which categorized households as either vulnerable or non-vulnerable. A binary logistic regression in the Statistical Package for Social Sciences (SPSS) was then used to explore the relationships between household demographic and socioeconomic characteristics and their vulnerability status. The vulnerability status served as the dependent variable in the regression, while independent variables (Table 1) were hypothesized well-established relationships with s vulnerability . Variables that did not show a significant relationship were excluded from the model.

Table 1: Household characteristics used as independent variables in the binary logistic regression model to identify determinants of socioeconomic vulnerability.

No	Indicator	Hypothesized Rationale
1	Women-headed households	Households with female heads tend to be less resilient, possibly because of external factors such as limited income-earning opportunities.
2	Household size	Evidence suggests that larger households, due to various factors, tend to be more vulnerable to food insecurity.
3	Dependency	In households with a high dependency ratio, there is a greater burden on household expenditures, contributing to higher levels of vulnerability.
4	Disabled population	Disabilities create barriers to accessing health care, education, and employment, thereby increasing vulnerability.
5	Educational level of household head and member	It is expected that households having members with higher literacy levels are less vulnerable because they can access better job opportunities. Furthermore, literacy increases access to knowledge, leading to greater awareness of disaster preparedness and recourse measures compared to people who are not literate.
6	Household main occupation or economic activities: • Skilled off-farm businessman/ worker • Un/low skilled off-farm business-man, worker, assistance, remit-tance • Farmer and farming worker	The food security and nutrition status of a household are significantly related to income, which is determined by various livelihood activities. A household whose primary income source is agricultural, or fishing activities is more sensitive to climate events. However, farmers who have knowledge, skills, and practice climate-smart agriculture, as well as adopt technology in farming, are more resilient. Farming households that have access to irrigation are more likely to cope with flood and drought impacts. Producing their own food makes households less reliant on the market, contributing to being less prone to food price shocks. Households relying on irregular, seasonal, or farming worker income, as well as assistance or remittances, tend to be more vulnerable to acute food insecurity due to irregular income generation. Households mainly relying on income from transportation, factories, construction, or entertainment work are more resilient to weather and climate impacts. However, they largely depend on food sourced from the market, exposing them to price shocks likely resulting from weather and climate events.

No	Indicator	Hypothesized Rationale
7	Household with farming land accessible to irrigation system	Access to irrigation is likely to improve agricultural productivity for farmers, leading to more resilience.
8	Household with deprivation in transportation means such as truck, car, motorbike, motorboat, etc.	This significantly contributes to households' inability to access food, low productivity in income-earning activities, and a lack of adaptive capacity to cope with extreme weather and climate change, resulting in increased vulnerability.
9	Household living in crowded house	A household is likely to be more vulnerable due to health problems related to sleeplessness caused by limited space.
10	Construction materials of house wall, floor, and roof	This is indicative of the household's low adaptive capacity to cope with climate-related hazards, resulting in adverse impacts on the household's assets and well-being.
11	Utensil and other cooking equipment	A household that has improved equipment such as a refrigerator, electric/gas stove, dishwasher, freezer, and dining set for food preparation is likely to have a higher quality of food consumption.
12	Household with accessibility to electricity	Access to electricity contributes to household productivity and links to improvement of household well-being.
13	Household accessible to improved source of drinking water	Households with no access to improved drinking water sources are more likely to suffer from health conditions which make them vulnerable.
14	Household accessible to improved sanitation.	A household that lacks access to improved latrines and hygiene facilities is likely to experience direct impacts on health issues.
15	Household access to social protection	Households that have access to social protection services may indicate economic vulnerability. However, they are more likely to be well-protected from adverse impacts compared to other vulnerable households without access.
16	Household exposed to flood	Floods deteriorate a household's capacity to maintain a smooth consumption pattern by disrupting livelihood activities and access to essential needs, leading to vulnerability to poverty and food insecurity.
17	Household exposed to irregular rainfall	Agricultural productivity is sensitive to climate variation because farming systems heavily rely on rainfall. Increasing rainfall positively impact on agricultural but very intensive rainfall may cause flooding.

PHASE 2: GEOSPATIAL EXTRAPOLATION OF VULNERABILITY DETERMINANTS

This phase involved geospatial extrapolation of the significant household demographic and socioeconomic variables identified in Phase 1. These variables were used as labeled inputs for training and validating a machine learning random forest (RF) algorithm in the Google Earth Engine (GEE). Together with predictive features extracted from geospatial data listed in Table 2, the model was able to extrapolate demographic and socioeconomic information to areas not covered by the household survey across the country.

Table 2: Geospatial data inputs as predictive factors in the model

No	Geospatial data	Justification
1	Nighttime light satellite imagery	It is the best proxy for Earth observation information in understanding socio-economic development activities.
2	High-resolution surface reflectance satellite imagery	Detailed information on land use/cover.
3	Accessibility to market	Improved access to markets reduces vulnerability to food insecurity by enhancing household income and food access.
4	Accessibility to education facility	Short distances to schools are linked to higher school enrollment rates, resulting in better employment opportunities. Schools also serve as shelters during flood events.
5	Accessibility to health facility	A long distance to health facilities discourages households from accessing healthcare services due to higher costs, leading to members being more likely to experience poor health and lower productivity.
6	Distance to main, secondary, and tertiary roads	Being close to improved road networks contributes to reduced transport time for goods/products and reduces travel time to access basic services such as healthcare facilities, markets, and schools. It also facilitates evacuation during disasters.
7	Distance to rivers and waterbody	It is associated with water sources for household consumption and livelihood activities and plays a role in disaster mitigation.
8	Population density	It reflects opportunities for the labor force in economic activities. However, it may also create higher demand for resources and jobs.
9	Built-up area	Remote rural areas are more likely to have insufficient essential needs and services.
10	Crop cultivation	Farming livelihood activities involve many households, as they rely on agricultural practices for their economic sustenance.
11	Forest area	It plays an important role in providing food and income for households, as well as regulating ecosystem balance.
12	Elevation	They influence climatology and the environmental ecosystem, which closely determine where and how households live.
13	Slope	
14	Climatological rainfall	They are linked to household livelihood activities, particularly in relation to agriculture.
15	Climatological temperature	

PHASE 3: MODELING SOCIOECONOMIC VULNERABILITY

In the final phase, a machine learning random forest algorithm in the GEE was applied again, this time to model the probability of socioeconomic vulnerability across Cambodia. The extrapolated demographic and socioeconomic information from Phase 2, along with additional geospatial data (refer to Table 2), were utilized to create predictive features. The binary household vulnerability status from Phase 1 served as labeled samples for training and validating the model. The output was a pixel-based probabilistic socioeconomic vulnerability index.

[1] Calculated by Economic Capacity to Meet Essential Needs (ECMEN).

[2] Computed by Food Consumption Score (FCS)

[3] Assessed by reduced Food Coping Strategies Index (rCSI) and Livelihood Coping Strategies (LCS)

World Food Programme Cambodia
No.108, Street 63/corner Street 208, Sangkat Boeung Raing, Khan Daun Penh,
Phnom Penh, Cambodia - Email: WFP.PhnomPenh@wfp.org
wfp.org