International Journal of Statistical Sciences *ISSN 1683-5603*
Vol.24(1), March, 2024, pp 49-64
DOI: https://doi.org/10.3329/ijss.v24i1.72017

Crafting Disaster-Driven Statistics: A Strategic Sampling Model

Syed Shahadat Hossain[1*] and Md. Rafiqul Islam[2]

[1]Institute of Statistical Research and Training, University of Dhaka, Dhaka 1000, Bangladesh. Email: shahadat@isrt.ac.bd

[2]National Accounting Wing, Bangladesh Bureau of Statistics, Statistics and Informatics Division, Ministry of Planning Government of the People's Republic of Bangladesh. Email: rafiqbbs25@gmail.com

[*]Correspondence should be addressed to Syed Shahadat Hossain (Email: shahadat@isrt.ac.bd)

[Received December 16, 2023; Accepted March 13, 2024]

Abstract

This article details the development and implementation of a strategic sampling methodology aimed at enhancing disaster-related statistics in Bangladesh. The study focuses on creating a specialized sampling frame by conducting a comprehensive census of enumeration areas (mouzas) affected by natural disasters. Employing a two-stage random sampling technique, the methodology incorporates stratification at district and disaster-type levels to capture diverse disaster occurrences. The Kish allocation method is utilized for sample allocation, addressing disparities in district sizes. Through meticulous trial and error simulations, the study ensures minimum sample sizes within each domain while employing inverse probability weights to estimate parameters. This strategic approach adopts robust estimations, enriching insights into disaster-related statistics.

Keywords and Phrases: Household survey for Disaster Related Data, Sampling Frame, Stratification, Environmental Survey Methodology.

AMS Classification: 62D05.

1. Introduction

Bangladesh, a country complexly knitted into the deltaic area of South Asia, stands as an example to resilience amid environmental vulnerabilities. Its low-lying landscape, covering an area dense with a population surpassing 160 million, faces the relentless onslaught of environmental disasters—floods, cyclones, and other calamities. In 2015, the Bangladesh Bureau of Statistics (BBS) launched ambitious project aimed at unravelling the human dimensions of these environmental disasters by conducting a country wide survey. The survey was strategically designed to center on a household-based respondent-reported data collection method, focusing on chronic loss, damage, and the socio-economic livelihoods deeply tangled with these adversities.

The origin of this survey was born from the real need to comprehend the deep impacts of environmental disasters on the lives of individuals and communities across Bangladesh. Reports such as Bangladesh Environmental Statistics Framework (BESF 2016-2030), The World Risk Report (2015) and insights from the IPCC assessments served as basis, driving the BBS towards measuring Bangladesh's vulnerability to natural calamities. The inaugural survey, conducted in 2015, paved the way for subsequent publications and studies, including the "Bangladesh Disaster related Statistics 2015: Climate Change and Natural Disaster Perspectives" (2015). While the focus primarily emphasized on chronic loss, damage, and socio-economic influences, collaborative studies with UNESCAP (2018) and the compilation of "Bangladesh Environmental Statistics" sought to enrich the discourse on disaster resilience and its socio-economic implications. The survey was also conducted in 2020 (see Bangladesh Disaster-related Statistics 2020).

The challenges faced in capturing the multifaceted impacts of these disasters across regions, communities, and ecosystems prompted to innovate its sampling methodology. This evolution needed a customized approach, enabling to capture the complexities of environmental disaster-related statistics effectively. This refined sampling methodology was mainly aimed towards bridging the data gaps related to disaster occurrences and their influence on public vulnerabilities. To meet these imperatives, we, on behalf of BBS, were engaged through an extensive process of methodological adaptation, leveraging insights from previous surveys and international best practices. The resulting modified sampling methodology emerged as a modified framework, aligning scientific rigor with the practicalities of household-based respondent-reported data collection. This paper represents this process including statistical sampling methodology describing its evolution and role in capturing the environmental disaster-related statistics in Bangladesh context.

2. Types of Disasters in Bangladesh

Bangladesh is characterized by its geographical diversity in its small span of area of approximately 147,570 sq. km. The majority of the country comprises low and flat land, with hilly regions in the northeast and southeast, and highlands in the north and northwest. Bangladesh's key geographical features include the presence of major river systems like the Ganges, Brahmaputra, and Meghna (GBM). The country's climate is tropical, with a mean annual temperature of about 25°C and an average annual rainfall of approximately 2,200 mm, mostly occurring from May to September (Climate change Knowledge Portal, World Bank, 2023).

2.1 Disaster and Its Definitions

Disasters in Bangladesh encompass a wide range of natural and human-made incidents, affecting life, livelihoods, resources, and the environment. They include natural occurrences such as cyclones, floods, water logging, droughts, and landslides, as well as human-induced incidents like explosions, fires, chemical spillage, and diseases causing pandemics. The devastation caused by these disasters often overwhelms local communities, requiring external assistance for relief efforts (UNDRR 2017).

2.2 Natural Disasters

The study under consideration comprehensively addressed the following ten types of disasters prevalent in Bangladesh (Disaster Management Act 2012; National Plan for Disaster Management 2021-2025):

Drought: Irregular and insufficient rainfall leading to crop failures, water shortages, and agricultural crises.

Flood: Annual occurrences of excessive rainfall and river overflow resulting in widespread destruction, loss of life, and damage to property and infrastructure.

Water Logging: A phenomenon involving deterioration of drainage condition in a number of southern coastal rivers leading to temporary to permanent inundation of floodplains along those rivers, causing enormous difficulties towards maintaining livelihoods and impacting agricultural productivity.

Cyclone: Severe storms along coastal regions causing high wind speeds, flooding, and economic losses.

Tornado: Severe local seasonal storms, popularly known as nor'westers (kalbaishakhi), generally associated with tornadoes occurring during transitional periods between southwest and northeast monsoons.

Storm/Tidal Surge: Rise in sea level causing coastal area inundation, especially during major storms.

Thunderstorm: Strong winds and heavy rain associated with continuous occurrence of lightning.

River Bank/Coastal Erosion: Erosion of the bed and banks of river making it wider, deeper and longer and submerging of the coastal lands.

Landslide: Land mass of higher altitude collapsing and causing property damage and potential loss of life often triggered in Bangladesh by heavy rainfall.

Salinity: Seasonal intrusion of saline water into inland areas, adversely affecting agriculture, fisheries, and freshwater resources.

3. Survey Methodology and the challenges

The efficacy of any statistical survey centers upon a accurately designed methodology including various facets, from framing schemas to executing data consistency checks. Methodological aspects encompass the preparation of sampling frames, choice of sampling techniques, questionnaire design, data collection methodologies, and measures to ensure data accuracy and consistency.

In light of the diverse and sporadic nature of disasters in Bangladesh, compounded by its dense population with a population of 169,828,911 (Bangladesh Population and Housing Census 2022) in 147,570 sq. km area and varied geography, sampling for this study faced notable challenges.

For the specific objectives of this disaster-related study, not all areas are aligned with the target population. Consequently, since the beginning of the Bangladesh Disaster related Statistics Survey 2015, a tailored approach was necessary, requiring the construction of a distinct and specialized sampling frame.

3.1 Ultimate Sampling Units and Respondents

In the context of this survey, the household (HH) was identified as the primary sampling unit, with the head of the household designated as the respondent. However, significant attention was directed towards accurately delineating and identifying the population itself. This was a critical preparatory step to ensure the survey's precision in capturing the targeted disaster-affected areas and eliciting informative responses from the chosen respondents.

3.2 Target Population

The target population for this survey is technically the entirety of Bangladesh, but the challenge arises due to the spatially sporadic nature of different disasters. While aiming to encompass the entire country, the study's specific objectives center around disaster-affected areas which exhibit spatial disparity, some disasters occurring in pocket areas across the country rather than uniformly. Consequently, narrowing the focus to disaster-affected areas becomes crucial to align with the survey's core objectives.

3.3 Sampling Frame

A critical element in survey methodology is the sampling frame, encompassing the target population. In the case of the Bangladesh Disaster-related Statistics (BDRS) survey, the absence of a specific sampling frame for disaster-affected regions posed a significant obstacle (Bangladesh Bureau of Statistics 2020). Typically, the Bangladesh Bureau of Statistics (BBS) defines Enumeration Areas (EAs) comprising a single 'mauza' or 100 to 120 households. These EAs, although initially formed during the '2011 Bangladesh Population and Housing Census' for broader statistical purposes, serve as foundational units for sampling in various surveys and censuses across the country.

The spatial disparity of disasters made it challenging to create a comprehensive frame. To address this, the study conducted a pre-survey census of enumeration areas (EAs) where at least one natural disaster had been reported. These EAs, together formed sampling frame and each EAs acts as the Primary Sampling Units (PSUs) for the study, from which the second stage sampling of HHs can be carried out. This tailored approach ensured that the sampling frame accurately reflected the objectives, focusing on disaster-impacted areas for in-depth data collection and analysis. For understanding the diverse spatial pattern of each of the types of disaster across the districts, separate maps for number of EAs experiencing a particular type of disaster in a district are presented in Figure 1. The detailed of these numbers are presented in Appendix (Table A1).

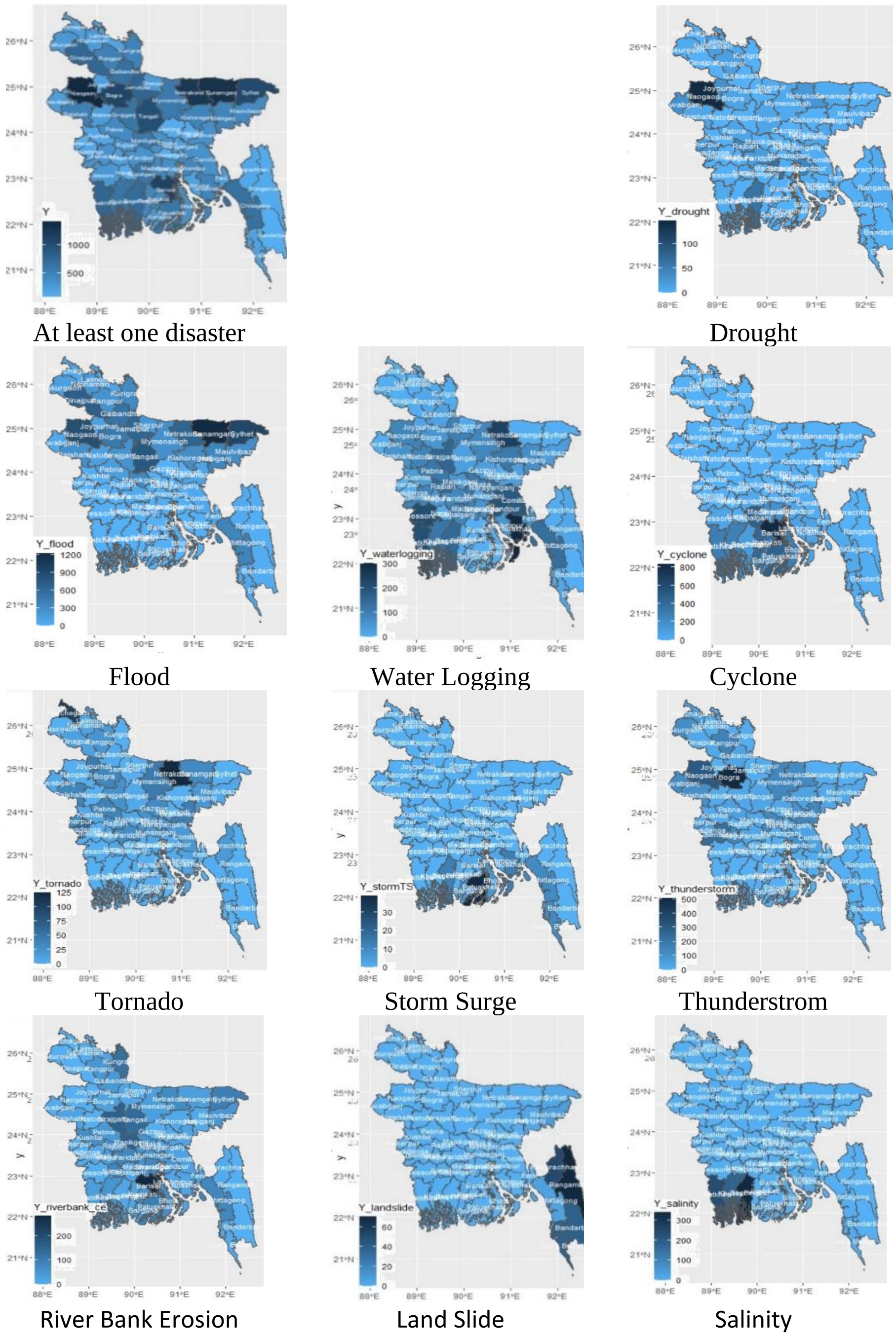
3.4 Stratification

The process of stratification within the study aimed to capture the complex variations present within Bangladesh's disaster-affected areas. Stratification occurred at two distinct levels: by district and by type of disaster. Each level of stratification served a specific purpose in aligning with the study's objectives and capturing the multifaceted nature of disaster occurrences. By stratifying at two levels, by district and by type of disaster, we have aimed to capture the distinct characteristics within the population. The role of each level of stratification are given as follows:

District-level Stratification: The study objectives aim to generate indicators to be measured at the district level, as a result stratification by districts was needed. It ensures that estimates are representative of different districts, accounting for regional variations in socio-economic, cultural, and geographical factors and also facilitates comparisons and in-depth analysis at the district level.

Disaster-type Stratification: The Choropleth map (Schiewe 2019) provided in Figure 1 clearly demonstrates, the sporadic and distinctive spatial pattern of each of the disaster type over the Bangladesh indicating potential socio-economic, cultural, or environmental differences across areas experiencing different disasters. A stratification based on disaster type is, therefore, intended. Note that the number of disaster affected mouza under each type of disaster is quite large posing no issue of zero sized stratum at national level. However, the stratification by disaster type within each district will flag this issue due to the spatial pattern of the disasters with some of the district not experiencing particular types of disaster at all. It is pre-conceived that estimation for each type of disaster in each of the district is not a requirement and apparently not attainable as well. By stratifying the sampling by type of disaster, however, will be able to capture the heterogeneity between the types of disaster for the aggregate national estimates.

Figure 1: Map of EAs where at least one natural disaster had been reported and maps by type of disaster



Moreover, recognizing that each type of disaster might have different consequences or effects, this stratification helps in understanding these varied facets within EAs. In summary, the disaster-type based stratification isn't meant for reporting estimates/indicators by disaster type. Instead, it's intended to enhance the study's depth by acknowledging and accounting for the diversity of effect/influence of types of disaster. This approach enables a more informed analysis without deviating from the primary objectives of the study, which are focused on district-level estimates and indicators.

3.5 Sampling Technique

In the conducted survey, a Stratified Two-Stage Random Sampling Technique (Cochran 1997) was employed to ensure statistical validity. This method involved two distinct stages of selection:

Selection of EAs/PSU: Within each district, the EAs reporting at least one natural disaster were further grouped based on the type of natural disasters they experienced, creating sub-strata. A simple Random Sampling (SRS) was used to choose EAs within each of the sub-strata separately.

Household Sampling: Following the selection of EAs, a complete listing of the HHs in the EA was made and a systematic sampling approach was implemented to pick required number of HHs from each chosen EAs. This systematic technique provided a structured means to uniformly select HHs across the designated EAs. This methodology ensured that the survey achieved comprehensive coverage, representing the diverse disaster incidences within each district while maintaining statistical rigor.

3.6 Sample Size Determination

While the survey in 2015 was done according to the above mentioned methodology with a sample size of 4945 PSUs with a total number of 148350 HHs. The determination of the sample size typically occurs at the domain (for this particular study, districts are the domains of interest) level, from which individual estimates are derived. If needed, additional adjustments or considerations might be made to accommodate specific conditions within the survey. Factors like desired confidence levels, margin of error tolerance, or variability in the population may also influence the final determination of the sample size of ultimate sampling units (HHs in most of the BBS surveys). For this study too, the minimum required number of HHs is calculated using the formula designed for estimating proportions. As per standard statistical theory, a formula for determining sample size (n) can be given as:

$$n = n_0 / \left(1 + \frac{n_0}{N}\right) \tag{1}$$

with

$$n_0 = \frac{p(1-p)}{d^2}\,[z(\alpha/2)]^2. \tag{2}$$

This formula provides the foundational basis for determining the adequate sample size needed for accurate estimations. In the context of the above formula, let's break down the variables involved:

1. p represents the apriori proportion of the required characteristic in the population under study.
2. $z(\alpha/2)$ signifies the value of the standard normal variate that corresponds to $100(1-\alpha)\%$ probability of encountering inaccurate or non-representative samples.

3. d stands for the allowable margin of error, which determines the acceptable level of error concerning the estimates.
4. N denotes the population size encompassing the entire group being studied.
5. f refers to the design effect, a parameter considered in complex survey designs, especially in multi-stage cluster sampling methodologies.

These variables play a crucial role in determining the ideal sample size for a statistically valid survey, as they provide guidelines for the selection process, considering confidence levels, margin of error, variability in the population, and survey complexity. Adjusting these parameters effectively helps in ensuring that the sample size accurately represents the population characteristics of interest.

The main concern for sample size determination of this sample census is to make choices of the values of p and d because a diverse range of variables will be under study. Theoretically, $p = 0.5$ gives the safest sample size since $p(1-p)$ takes the highest value for $p = 0.5$. Conventionally α can be taken as .05 and f can be taken as 1.5-2.0 for most socio-economic surveys in Bangladesh. The equation (1) is used to revise the sample size in equation (2) for the population size N. It is observed in the theory that for $N \geq 8000$, equation (2) is not much influenced or improved by equation (1). That is why, equation (2) is straight forwardly used for large population sizes and there is no necessity for increasing the sample size for population becoming any larger. In this particular study under interest, the minimum number of HH required in each stratum is focused and since the number of HH in each of the defined strata is fairly larger than 8000, no adjustment for population size is suggested.

A common choice for the value of the absolute allowable margin of error is $d = .05$. This value does not seem to be realistic for scenario where the true value of p is outside the range $0.2 \leq p \leq 0.8$, when a smaller value for allowable margin of error or consideration of a relative error margin is recommended. In such cases, instead of considering an absolute allowable margin of error d, the margin of error is set as an allowable relative proportion of p. Denoting r as the maximum proportion of p to the allowed margin of error, i.e., setting $d = rp$, equation (2) can be deduced to

$$n_0 = \frac{1-p}{pr^2}\,[z(\alpha/2)]^2. \qquad (3)$$

While the sample size for the 2015 BDRS was determined with an judgmental choice of p=0.35 depending on experience and expertise, the choices of the value of p for the BDRS 2021, we have considered the results from ‘Bangladesh Environment Statistics 2020’. We can see that the weighted proportion of disaster effected HH in the country can be approximately estimated (calculated only for sample size computation) as

$$\frac{\text{Total Number of Mouzas in the Country}}{\text{Number of Mouzas identified for sampling frame}} \times \frac{\text{Estimated Number of Disaster prone HH}}{\text{Total Number of HH in the Country}}$$
$$= \frac{63440}{29199} \times \frac{4361261}{32173630} = 0.295\,.$$

To allow a maximum error of $\pm 10\%$ in estimating a characteristic with the above estimated proportion of $p = 0.295$, the minimum required sample size n can be obtained to be 1377 HHs, taking the conventional choice of $\alpha = 0.05$ and a moderate design effect $f = 1.5$ in equation (1) The number can be approximated to $n = 1377 \approx 1380$ HH per domain (districts). Note that use of a design effect $f = 1.5$ is consistent with the choice of second stage sample size to be drawn from each selected PSUs. A second stage sample size of $m = 30$ HHs was decided by taking into consideration the administrative and survey execution conveyances while keeping it consistent with the design effect and intra class correlation ρ (ICC). Again, using the data produced by BDRS 2021, the ICC computed for economic vulnerability index was found to be 0.019 which approximately justifies the choices $f = 1.5$ and $m = 30$ as per the relation $f = 1 + \rho\,(n - 1)$. Considering 30 HHs at the second stage from each of the selected PSU, the required minimum number of PSUs from each district becomes $\frac{1380}{30} = 46$. The methodology employed here deduced a minimum requirement of 1380 HHs (or 46 EAs) within each district. Note that in case of an equal allocation among districts, a straight multiplication of the number by number of district would have produced the total sample size. Since we resort to Kish allocation, the total sample size will be obtained from the allocation which will be discussed in the next sections. While this minimum required sample size per district is determined using the indicator 'proportion of disaster effected HH' with specified relative margin of error, it allow a narrower margin of error for estimating characteristics that are more common.

3.7 Allocation of Sample Size

The choice of the Kish allocation (Kish 1965) method originated from the necessity to address discrepancies in district sizes within the survey. While equal allocation might seem equitable initially, it fails to consider substantial variations in population density and variability among districts. Conversely, proportional allocation risks assigning extremely small sample sizes to relatively smaller districts, potentially affecting the reliability of estimates from these areas. But, this poses a concern since in surveys where administrative regions like districts serve as primary strata, ensuring adequate representation from each district is pivotal for obtaining reliable estimates at that level. Kish allocation safeguards a balance between these extremes, offering a compromise that minimizes the risk of excessively small samples in smaller districts while maintaining reasonable proportional allocation across all areas. The Kish allocation formula is

$$n_h = n^* \left(\sqrt{\frac{1}{D^2} + W_h^2} \right) \Big/ \sum_{h=1}^{D} \sqrt{\frac{1}{D^2} + W_h^2} \,,$$

where, n_h is the number of sample EAs in the district h, n^* is the total number of EAs in the sample, D is the number of districts, W_h is the proportion of EA in district h, i.e.,

$$W_h = \frac{\text{Total Number of EA in the frame belonging to the } h^{th} \text{ district}}{\text{Total Number of EA in the frame}}\,.$$

When distributing a fixed total sample size (n^*) across various strata (domains), a common issue arises in smaller strata potentially having sample sizes below the minimum required. However, when the study necessitates estimations at distinct strata levels, it becomes crucial to establish a constraint ensuring that each domain's sample size meets or exceeds the minimum requirement.

Employing the Kish allocation formula aimed to optimize the distribution of sample HHs across districts, facilitating robust and reliable estimations from each administrative region without significantly compromising the overall sample size or favoring larger or smaller districts. The allocation ensured that no district had fewer than the computed minimum requirement of 46 EAs (1380 HHs). Achieving this involved a method using trial and error simulations with different alternatives of total number PSUs (n^*). This iterative approach aimed to meet the restriction of attaining a sample size in each stratum that meets the necessary threshold. Extensive simulations determined a total sample size of 4240 PSUs, effectively meeting the threshold of 46 EAs (1380 HHs) in every district for district-level estimation. This resulted in an overall sample of 4240 EAs across 64 districts, encompassing a 127,200 HHs. While different districts are allocated with different number of EAs, each district maintained an at least 46 EAs (1380 HHs), ensuring representation. Specific allocation details based on EAs affected by natural disasters are outlined in the Appendix (Table A2).

In the allocation process across different types of disasters within districts, the focus is on the count of EAs experiencing a particular type of disaster. For instance, if the sample number of PSU required is 50 within a district and the number of EAs affected by different type of disaster is: Floods- 60 EAs, Droughts- 30 EAs and Cyclones- 10 EAs. Then the allocation of the number of sample EAs across the types of disaster was: 30 EAs from flood exposed EAs, 15 from droughts impacted EAs and 5 from cyclone effected EAs. This methodology ensures that the survey captures data from a representative number of EAs affected by various specific disaster types within each district. Figure 2 shows the spatial distribution of the EAs where at least one natural disaster had been reported spread over the country along with the selected sample EAs.

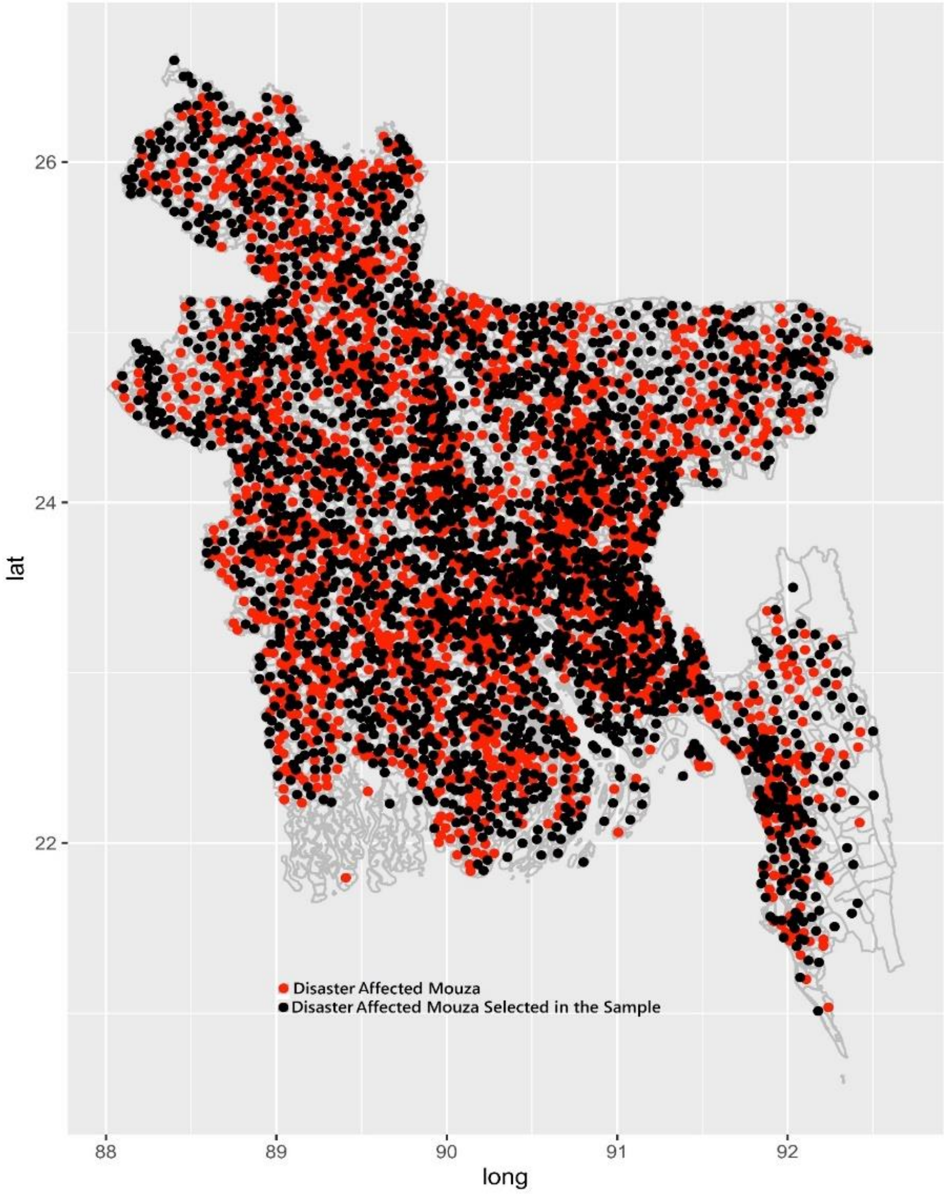
4. Estimating formula

Since the method of selection is the same for each stratum/district and selection is done independently, following estimating formula is applicable to each district. Let θ_h be the total of all values of a variable Y for all HH in the district h and it is of interest to obtain an estimate of θ_h or of its function $\tau(\theta_h)$. Now, denoting the measure of the variable Y, for j^{th} HH in i^{th} PSU in the h^{th} district ($h = 1, 2, \dots, D;\ i = 1, 2, \dots, N_{h1}; j = 1, 2, \dots, N_{h2i}$) by Y_{hij}, we have the parameter of interest could be the population total $\theta_h = \sum_{i=1}^{N_{h1}} \sum_{j=1}^{N_{h2i}} Y_{hij}$ or the population mean $\overline{\theta_h} = \frac{1}{N_h} \sum_{i=1}^{N_{h1}} \sum_{j=1}^{N_{h2i}} Y_{hij}$. Note that N_h, N_{h1} and N_{h2i} denote total number of HH in the district h, the number of EAs in the district h and the number of HH in the i^{th} EA, respectively.

Since a two-stage sampling is conducted and the EA sizes are not uniformly equal, resulting in unequal probability of selection, we employ inverse probability weights for obtaining the estimates of the parameters. The expression of the selection probabilities and resulting weights for district $h\ (h = 1, 2, \dots, D)$, can be found in some straight forward computations as illustrated in the following section.

Probability of selecting i^{th} EA ($i = 1, 2, \dots, N_{h1}$), $p_{h1} = \frac{n_{h1}}{N_{h1}}$; where n_{h1} is the number of EAs selected in the sample within the stratum, corresponding sampling weight becomes $w_{h1} = \frac{1}{p_{h1}} = \frac{N_{h1}}{n_{h1}}$. Similarly, denoting the number of HHs selected in the sample from i^{th} selected EA by n_{h2i}, the probability of selecting the j^{th} HH given that the i^{th} EA ($i = 1, 2, \dots, N_{h1}; j = 1, 2, \dots, N_{h2i}$) is selected can be expressed as $p_{h2i} = \frac{n_{h2i}}{N_{h2i}}$; where n_{h2i} is the number of HHs selected in the sample from the i^{th} EA within the stratum, corresponding sampling weight becomes $w_{h2i} = \frac{1}{p_{h2i}} = \frac{N_{h2i}}{n_{h2i}}$. The sample observation y_{hij}, the measured value of the variable Y from j^{th} HH in i^{th} selected EA $(i = 1, 2, \ldots, n_{h1}; j = 1, 2, \ldots, n_{h2i})$, thus, will carry a weight $w_{hij} = w_{h1} \times w_{h2i} = \frac{N_{h1}}{n_{h1}} \times \frac{N_{h2i}}{n_{h2i}}$.

Figure 2: Map of Sample EAs and all EAs where at least one natural disaster had been reported



4.1 Estimate of population total and mean

The estimates of the population total θ_h can be computed as

$$\hat{\theta}_h = \sum_{i=1}^{n_{h1}} \sum_{j=1}^{n_{h2i}} w_{hij}\, y_{hij} \ .$$

For a combined national estimate, the domain specific estimates can be aggregated keeping the domain weights into consideration as

$$\hat{\theta} = \sum_{h=1}^{D} \hat{\theta}_h \ = \sum_{h=1}^{D} \sum_{i=1}^{n_{h1}} \sum_{j=1}^{n_{h2i}} w_{hij}\, y_{hij} \ .$$

The above derived estimates are unbiased estimate of the respective population parameters.

4.2 Standard error of the estimates

Since sampling fraction is small, the variance of $\hat{\theta}_h$ and it's estimate can be determined by standard with-replacement formula (Lohr 2021) as

$$V(\hat{\theta}_h) = w_{h1} N_{h1} S_{h1}^2 + w_{h1} \sum_{i=1}^{N_{h1}} N_{h2i}\, w_{h2i} S_{h2i}^2 \, ,$$

and

$$v(\hat{\theta}_h) = w_{h1} N_{h1} s_{h1}^2 + w_{h1} \sum_{i=1}^{n_{h1}} N_{h2i}\, w_{h2i} s_{h2i}^2 \, ,$$

respectively, where S_{h1}^2 and S_{h2i}^2 denote the population variance of the first and second stage units in the domain h and s_{h1}^2 and s_{h2i}^2 are the respective sample quantities.

5. Concluding Remarks

The methodology adopted for this survey stands as a pioneering attempt in comprehensively addressing the complexities integral in generating disaster related statistics within Bangladesh. A key novelty of the methodology was the formation of a sampling frame for the disaster-related target population, attained through a careful census of EAs.

The stratification by district and disaster type, coupled with thorough sampling techniques and sample size determination, facilitated a stronger representation of the disaster related features of the population with diverse socio-geographic dimensions within Bangladesh. The use of Kish allocation method further highlighted the equitable representation while balancing the constraints attributed by diverse district sizes. In addition to the procedural outline of survey of disaster related socio-economic feature, it catered a more granular understanding of the heterogeneous nature of disaster occurrences across the country. While this methodology may demonstrate a substantial progress, a continued adaptations and revision will be important to remain responsive and robust for evolving circumstances.

Acknowledgement: We extend our heartfelt gratitude to the Bangladesh Bureau of Statistics (BBS) for the opportunity to engage deeply with the technical aspects of the BDRS 2021 survey. This article primarily reflects the academic insights and theoretical framework underlying the survey's design. The survey, conducted and published in 2022 under the guidance of the second author serving as the Project Director, owes much of its intellectual groundwork to the collaborative efforts of both authors. The first author contributed significantly to the preparatory academic deliberations and the crafting of the sampling methodology. Our sincere appreciation also goes to the entire team involved in the execution of the survey for their invaluable contributions.

Reference

[1] Bangladesh Environmental Statistics Framework (BESF) (2016-2030): Environment, Climate Change and Disaster Statistics (ECDS) Cell, Bangladesh Bureau of Statistics, Statistics and Informatics Division, Ministry of Planning Government of the People's Republic of Bangladesh.

[2] Bangladesh Population and Housing Census (BPHC) (2022): National Report (Volume I), Bangladesh Bureau of Statistics, Ministry of Planning Government of the People's Republic of Bangladesh.

[3] Bangladesh Bureau of Statistics (2020): Sampling Frame of the ECDS

[4] Bangladesh Bureau of Statistics (2015): Bangladesh Disaster-related Statistics 2015: Climate Change and Natural Disaster Perspectives.

[5] Cochran, W. G. (1997): Sampling techniques. Hoboken, NJ: John Wiley & Sons, Inc.

[6] Kish, L. (1965): Survey Sampling. Wiley, New York.

[7] Lohr, S. L. (2021): Sampling: Design and Analysis (3rd ed.). Chapman and Hall/CRC.

[8] National Plan for Disaster Management (2021-2025): Action for Disaster Risk Management Towards Resilient Nation March 2020, Ministry of Disaster Management and Relief, Government of the People's Republic of Bangladesh.

[9] Schiewe, J. (2019): Empirical Studies on the Visual Perception of Spatial Patterns in Choropleth Maps. KN J. Cartogr. Inf. 69, 217–228.

[10] UNDRR (2017): Technical guidance for monitoring and reporting on progress in achieving the global targets of the Sendai Framework for Disaster Risk Reduction.

[11] Disaster Management Act (2012): Ministry of Disaster Management and Relief, Government of the People's Republic of Bangladesh.

[12] UNESCAP (2018): Disaster-related Statistics Framework (DRSF).

[13] World Risk Report (In cooperation with 2015): Ed. Mucke, P. Hilft, B. E., Welthungerhilfe, K. D. and Walter, J., United Nations University – EHS.

[14] World Bank (2023): Climate Change Knowledge Portal, Bangladesh.

Table A1: Zila wise summary list of Disaster affected mouza of Bangladesh

Division/zila_name	Total mauza	Disaster affected mauza											
		Drought	Flood	Water logging	Cyclone	Tornado	Strom/ tridal surge	Thunde- rstrom	River bank/ coastal	Landslide	Salinity	Total mauza	Total Household
BARGUNA	303	0	0	4	223	0	6	0	41	0	21	295	214989
BARISHAL	1192	0	2	0	836	0	2	3	282	0	0	1125	516566
BHOLA	452	0	0	1	284	0	7	5	65	0	0	362	375591
JHALOKATI	438	0	8	1	375	0	0	0	52	0	0	436	157955
PATUAKHALI	627	0	0	0	46	0	39	0	96	0	6	603	340330
PIROJPUR	450	0	0	1	420	0	0	0	21	0	0	442	255218
BANDARBAN	122	0	35	0	18	0	0	0	2	40	0	95	85097
BRAHMANBARIA	1014	10	103	30	0	42	0	141	22	0	0	348	226764
CHANDPUR	1134	42	2	196	53	4	0	98	49	0	0	444	224119
CHATTOGRAM	1195	1	373	136	36	2	10	5	78	6	0	647	1732495
CUMILLA	2766	4	61	184	23	19	0	44	25	0	0	360	192717
COXS BAZAR	238	0	72	16	35	0	10	1	14	39	9	196	379970
FENI	601	0	111	65	32	0	0	2	33	0	0	243	141470
KHAGRACHHARI	153	1	46	4	0	19	0	11	9	51	0	141	126036
LAKSHMIPUR	552	4	6	54	81	4	2	15	22	0	8	196	210160
NOAKHALI	1060	1	9	302	84	11	13	12	23	0	12	467	389796
RANGAMATI	183	7	25	19	0	1	0	2	2	71	0	127	102344
DHAKA	1267	0	10	115	0	25	0	136	58	0	0	344	394916
FARIDPUR	1143	20	54	81	49	30	0	32	77	0	0	343	143672
GAZIPUR	840	0	0	33	0	32	0	32	1	0	0	98	171749
GOPALGANJ	668	3	14	85	474	0	0	3	24	0	6	609	227529
KISHOREGANJ	1022	0	299	112	0	0	0	46	16	0	0	473	321300
MADARIPUR	583	66	6	86	201	0	0	1	53	0	0	413	193391
MANIKGANJ	1361	1	340	48	0	8	0	27	153	0	0	577	170555
MUNSHIGANJ	685	0	9	10	0	19	0	6	48	0	0	92	54621
NARAYANGANJ	816	8	13	125	0	23	0	22	15	0	0	206	408265
NARSINGDI	670	0	6	6	0	2	0	93	12	0	0	119	109266
RAJBARI	886	45	80	145	0	22	0	8	52	0	0	352	113223
SHARIATPUR	686	2	38	56	179	6	0	5	76	0	0	362	171234
TANGAIL	2115	11	778	33	0	7	0	37	135	0	0	1001	426528
BAGERHAT	757	62	0	148	116	0	13	0	17	0	342	698	328778
CHUADANGA	415	29	0	0	0	63	0	61	0	0	0	153	108397
JASHORE	1398	1	6	212	68	0	0	3	3	0	3	296	142761
JHENAIDAH	989	0	0	87	0	1	0	230	13	0	0	331	160327

KHULNA	820	1	0	118	273	0	0	0	53	0	182	627	340912
KUSHTIA	801	2	32	38	0	5	0	13	63	0	0	153	116701
MAGURA	548	2	1	46	0	9	0	136	37	0	0	231	93643
MEHERPUR	220	24	0	0	0	31	0	29	0	0	0	84	72411
NARAIL	472	12	0	4	19	7	0	7	77	0	1	127	54129
SATKHIRA	971	6	0	206	282	0	0	5	33	0	241	773	404358
JAMALPUR	910	1	646	96	0	0	0	7	17	0	0	767	513744
MYMENSINGH	2338	7	268	83	0	46	0	143	31	0	0	578	400493
NETRAKONA	1672	26	657	241	0	125	0	125	69	0	0	1243	383913
SHERPUR	500	4	220	8	0	0	0	15	6	0	0	253	228393
BOGURA	1915	12	472	50	0	35	0	510	29	0	0	1108	590263
JOYPURHAT	755	31	123	18	0	45	0	0	0	0	0	217	84077
NAOGAON	2592	147	871	86	0	1	0	309	0	0	0	1414	387055
NATORE	1339	0	143	87	0	0	0	18	0	0	0	248	103110
CHAPAI NABABGANJ	823	32	67	5	0	7	0	127	9	0	0	247	158841
PABNA	1428	12	161	115	0	38	0	50	42	0	0	418	271965
RAJSHAHI	1839	11	276	58	0	36	0	107	23	0	0	511	198212
SIRAJGANJ	1540	0	252	154	0	25	0	27	157	0	0	615	330050
DINAJPUR	2144	0	705	3	0	12	0	29	0	0	0	749	283751
GAIBANDHA	1148	0	533	26	0	3	0	39	45	0	0	646	481636
KURIGRAM	656	0	325	7	0	1	0	11	124	0	0	468	396534
LALMONIRHAT	403	8	113	11	0	18	0	8	22	0	0	180	155882
NILPHAMARI	428	18	206	1	0	10	0	12	9	0	0	256	253134
PANCHAGARH	469	0	72	0	0	102	0	63	2	0	0	239	123923
RANGPUR	1301	17	253	14	0	22	0	160	47	0	0	513	363748
THAKURGAON	698	5	145	3	0	1	0	118	1	0	0	273	157408
HABIGANJ	1383	2	355	64	0	19	0	29	10	3	0	482	167947
MOULVIBAZAR	1015	7	298	34	0	5	0	25	46	8	0	423	174629
SUNAMGANJ	1729	6	1231	3	0	33	0	7	32	0	0	1312	384200
SYLHET	1802	0	944	3	0	0	0	5	94	4	0	1050	347522
BANGLADESH	**63440**	**711**	**11875**	**3977**	**4623**	**976**	**102**	**3215**	**2667**	**222**	**831**	**29199**	**17340713**

Table A2: Allocation of sample (natural disaster PSU/mauza/mahallas and household).

Name of District based on Geo-Code with Division wise	Total Natural Disaster Affected Mauza/Mahallas (EA)	Allocation of Sample (Natural Disaster Mauza/Mahallas) (EA)	Number of ample Household Per Mauza/Mahallas/ PSU (EA)	Total Number of Sample Household
BARGUNA	295	53	30	1590
BARISAL	1125	120	30	3600
BHOLA	362	56	30	1680
JHALOKATI	436	62	30	1860
PATUAKHALI	603	74	30	2220
PIROJPUR	442	62	30	1860
BANDARBAN	95	46	30	1380
BRAHMANBARIA	348	56	30	1680
CHANDPUR	444	62	30	1860
CHITTAGANG	647	77	30	2310
COMILLA	360	56	30	1680
COX'S BAZAR	196	49	30	1470
FENI	243	50	30	1500
KHAGRACHHARI	141	46	30	1380
LAKSHMIPUR	196	49	30	1470
NOAKHALI	467	64	30	1920
RANGAMATI	127	46	30	1380
DHAKA	344	56	30	1680
FARIDPUR	343	56	30	1680
GAZIPUR	98	46	30	1380
GOPALGANJ	609	74	30	2220
KISHOREGONJ	473	64	30	1920
MADARIPUR	413	61	30	1830
MANIKGANJ	577	73	30	2190
MUNSHIGANJ	92	46	30	1380
NARAYANGANJ	206	49	30	1470
NARSINGDI	119	46	30	1380
RAJBARI	352	56	30	1680
SHARIATPUR	362	56	30	1680
TANGAIL	1001	108	30	3240
BAGERHAT	698	82	30	2460
CHUADANGA	153	47	30	1410
JESSORE	296	53	30	1590
JHENAIDAH	331	55	30	1650
KHULNA	627	76	30	2280
KUSHTIA	153	47	30	1410
MAGURA	231	50	30	1500
MEHERPUR	84	46	30	1380
NARAIL	127	46	30	1380

Name of District based on Geo-Code with Division wise	Total Natural Disaster Affected Mauza/Mahallas (EA)	Allocation of Sample (Natural Disaster Mauza/Mahallas) (EA)	Number of ample Household Per Mauza/Mahallas/ PSU (EA)	Total Number of Sample Household
SATKHIRA	773	89	30	2670
JAMALPUR	767	88	30	2640
MYMENSINGH	578	73	30	2190
NETRAKONA	1243	131	30	3930
SHERPUR	253	50	30	1500
BOGRA	1108	119	30	3570
JOYPURHAT	217	49	30	1470
NAOGAON	1414	147	30	4410
NATORE	248	50	30	1500
CHAPAI NAWABGANJ	247	50	30	1500
PABNA	418	61	30	1830
RAJSHAHI	511	67	30	2010
SIRAJGANJ	615	76	30	2280
DINAJPUR	749	86	30	2580
GAIBANDHA	646	77	30	2310
KURIGRAM	468	64	30	1920
LALMONIRHAT	180	47	30	1410
NILPHAMARI	256	50	30	1500
PANCHAGARH	239	50	30	1500
RANGPUR	513	67	30	2010
THAKURGAON	273	52	30	1560
HABIGANJ	482	65	30	1950
MAULVIBAZAR	423	61	30	1830
SUNAMGANJ	1312	137	30	4110
SYLHET	1050	113	30	3390
Total (Bangladesh)	**29199**	**4240**		**127200**