Received: 16 November 2021 | Revised: 2 June 2023 | Accepted: 19 June 2023

DOI: 10.1002/jid.3824

RESEARCH ARTICLE

JOURNAL OF INTERNATIONAL DEVELOPMENT WILEY

Identifying vulnerability to human trafficking in Bangladesh: An ecosystem approach using weak-signal analysis

Gregory E. van der Vink[1] | **Katherine N. Carlson**[1] | **Erica Phillips**[1] | **Sabrina H. Szeto**[1] | **Jeffrey Park**[2] | **Michael E. Jackson**[1]

[1]Novametrics LLC, Bethesda, Maryland, USA

[2]Yale University, New Haven, Connecticut, USA

Correspondence
Jeffrey Park, Department of Earth and Planetary Sciences, Yale University, New Haven, CT 06520-8109, USA.
Email: jeffrey.park@yale.edu

Funding information
United States Agency for International Development, Grant/Award Number: GS-10F-0033M/7200AA18M00016; United States Office of the Secretary of Defense, Grant/Award Number: #W911QX-13-C-0182

Abstract

We present an ecosystem approach to analyze open-source data to identify populations vulnerable to human trafficking and to reveal underlying causal relationships. In the case of Bangladesh, our analysis suggests combinations of indicators that are highly predictive of human trafficking. The traditional narrative that poverty and unemployment are the main drivers for human trafficking may be an oversimplification. We find many areas where vulnerability is highest within lower-middle to middle-class societies with (a) moderate levels of income and education, (b) adherence to traditional gender norms of a male-dominated patriarchal society, and (c) access to an urban center.

KEYWORDS
aid, Bangladesh, data analysis, development, human trafficking, trafficking-in-persons, weak-signal analysis

1 | INTRODUCTION

Globally, human trafficking is the third most prevalent organized crime, surpassed only by the arms and drug trades (Joarder & Miller, 2014). Within Bangladesh, human trafficking is a pervasive problem that has stubbornly resisted traditional solutions (Zimmerman et al., 2021). The trafficking problem is compounded by many factors including a large, diverse population, high poverty rates, frequent rural-to-urban migration, and recurring natural disasters that disrupt livelihoods (Gazi et al., 2001). Women and children can be especially vulnerable to human trafficking, and within Bangladesh it is estimated that tens of thousands of women and children fall victim annually (Shoji & Tsubota, 2022). Exacerbating the situation, South Asia is highly exposed to the manifestations of climate change, with an expected increase in droughts, floods, cyclones, and other natural disasters that pair with pre-existing socio-economic vulnerabilities to fuel further displacement (Poncelet et al., 2010).

Just as there is no single cause for human trafficking in Bangladesh, there is no single solution. Even when causes appear similar, solutions that may work in one location rarely work in another location owing to varying sociocultural and economic conditions. The lack of effective universal solutions has frustrated anti-trafficking efforts and limited their success (e.g., Betz, 2009). In this paper, we recognize there are no effective single and universal solutions for human trafficking in Bangladesh. Alternatively, we present an approach designed towards identifying vulnerable populations and developing geographically-targeted interventions customized for different socio-economic landscapes. While our analysis is illustrated with Bangladesh, we feel this methodology can be applied more widely and holds strong promise for reducing the number of individuals who experience human trafficking.

We begin by classifying human trafficking as a 'wicked problem'[1]—the type of problem that defies a single solution and is characterized by a myriad of dynamically interconnected variables. Causal relationships are seldom direct, and the circumstances that foster the problem vary from location to location. While 'wicked problems' such as human trafficking may defy single and universal solutions, they can be addressed through an ecosystem approach (Clarke, 1995; Eck & Clarke, 2019). In an ecosystem approach, we parameterize each socio-economic landscape in which human trafficking exists to reveal (using weak-signal analysis) the combinations of characteristics that allow trafficking to occur. Once we identify the characteristics of the ecosystems that support trafficking, we can formulate geographically-targeted interventions to disrupt that ecosystem and mitigate human trafficking.

Within our ecosystem approach, we do not assume that any characteristic by itself is predictive of human trafficking. We assume combinations of characteristics will create vulnerability, and these characteristics vary from location to location. Our evaluation of vulnerability is related to, but not necessarily the same as measurements of known trafficking prevalence. We assume socio-economic ecosystems can be vulnerable to human trafficking, but as-yet unexploited, or the exploitation can be as-yet unrecognized.

Ecosystem approaches have been proposed previously for addressing the sex trafficking of children (Finigan-Carr et al., 2019) and for building resilience to trafficking within communities (Gardner et al., 2020).[2] The ecosystem approach is grounded in the criminology theory of situational crime prevention (SCP) and is focused towards 'Prevention' part of the '3P' paradigm (Protection, Prevention, Prosecution) that is often used for addressing human trafficking (Victims of Trafficking and Violence Protection Act, 2000). SCP focuses on the criminal setting and begins with an examination of the circumstances (what we call the 'ecosystems') that allow for particular types of crime. By gaining an understanding of these ecosystems, mechanisms are then introduced to change the ecosystems, reducing the opportunities for crime. SCP has been recognized as an essential part of the United Nations Economic and Social Council's Guidelines for the Prevention of Crime (Resolution 2002/13) (UNODC, 2010).

In applying an ecosystem approach, we use the methodology of 'weak-signal analysis' to identify the characteristics of the various ecosystems in which trafficking occurs. Weak-signal analysis begins with the recognition that every population has a complex mosaic of characteristics derived from its demographics, environmental resources, geographical location, ethnic history, wealth, social norms, income-producing activities, religious sects, access to markets, educational levels and so on. Individual characteristics are not necessarily assumed to have direct causal relationships to trafficking, rather we assume combinations of characteristics are more likely to hold predictive value. Our weak-signal analysis methodology allows us to mathematically combine datasets to reveal combinations of characteristics predictive of human trafficking. We use not only health, economic, and educational survey data, but also national census data, Earth observation data, web-derived data, and data from both formal and informal media sources. While these datasets may be of varying quality and completeness, we find that they carry information that reflects characteristics of a population, either by themselves or in combination with other datasets.

2 | MATERIALS AND METHODS

Our ecosystem methodology assumes that every population has a mosaic of characteristics, and that these characteristics are often connected through a series of underlying systems with few direct causal relationships. The goal of our methodology is to predict vulnerability to human trafficking, and to identify possible underlying causal relationships that can then be used to develop geographically-targeted interventions.

We began by compiling millions of data values from diverse, mostly open-source datasets from Non-Governmental Organizations (NGOs), media, the U.S. Government, and the statistical authorities of local governments. These datasets include detailed national census data, health and educational survey data, remote-sensing data, web-scraped data, and data from both formal and informal media sources. Bangladesh survey data sources included the Bangladesh Multiple Indicator Cluster Surveys (BBS and UNICEF, 2014, 2019), the Bangladesh Integrated Household Survey (IFPRI, 2016), and the Bangladesh Demographic and Health Survey (NIPORT and ICF, 2020). From these surveys, we extracted millions of data values that were then distilled into over half a million indicator values covering over 1500 measures for 544 upazilas over 20 years, representing demographics, governance, land use, natural resources, education, health, economics, ethnicity, religion, infrastructure, conflict, gender equality, female empowerment, societal norms and other human-social-cultural-behavioural (HSCB) characteristics. Of these indicators, the approximate category breakdown is as follows: 21% demographic, 10% economic, 14% education, 18% health, 8% infrastructure, 13% natural resources, 7% religion, and 10% social. High-resolution geospatial and earth-observation data such as land cover, agricultural, and climatic variables were converted into tabular data for analysis. Depending on the data type, values were summed (e.g., to determine population) or statistical measures of the value's distribution were used (e.g., average travel distance to a road, market or urban area).

The Bangladesh Police Department collects data on human trafficking, but these data are only released sporadically and without geographic information. Annual reports from the Bangladesh Ministry of Home Affairs (MHA) are released inconsistently and several years in arrears. For this analysis, data on known trafficking case statistics were taken from the available MHA annual reports (Ministry of Home Affairs, 2016 and 2018), augmented using government reports, NGO reports, the academic literature, and both formal and informal media reports from 2019 to 2020. The human trafficking cases in the MHA reports come from counter-trafficking committees (CTCs) within the 64 zilas that comprise Bangladesh, the units at the nation's second administrative level. CTCs are formed by the Bangladesh government to support counter-trafficking programs, help trafficking survivors, and protect witnesses. The reported cases are in some process of the justice system. To account for delays in the justice system, we sum the reported cases for 2016 and 2018. The number of cases is normalized to population per 100 000 to adjust for differences in population size among zilas. Using the literature survey and media reports, data were collected on source, transit, and destination locations within the country to supplement case data.

A detailed technical summary of our weak-signal analysis is included in Appendix A. In summary, we use singular-value decomposition, combined with varimax rotation and squared-factor loadings as an unsupervised self-learning algorithm to identify key attributes and their relative weightings. In non-mathematical terms, we use a wide range of socio-economic measurements to capture the full spectrum of factors that are associated with a population. Our algorithm pares down a large dataset into a smaller one comprised of the most defining and statistically relevant components. Running the analysis within specific high-trafficking subregions of Bangladesh enables the identification of the combinations of characteristics predictive of trafficking, while eliminating the combinations of characteristics that are neither conducive nor preventative. Attributes and attribute-combinations that are prominent in both areas of known high- and low-level human trafficking are thus deemed as inconclusive to vulnerability. The attributes are then tested via resampling methods, in which the algorithm is run on different subsections of regions, to confirm consistency and sensitivity. To explain as much of the variance in the data as possible while avoiding an overly-complicated measure, various threshold values for indicator weightings are used to identify the optimal subset of

TABLE 1 Indicators and weightings for vulnerability measure.

Indicator in vulnerability measure	Weighting
Percentage of households with shingled roofs[a]	0.312
Percentage of households owning a bicycle[a]	0.171
Percentage of women earning no income because they say society prohibits them from working[b]	0.170
Percentage of girls married between ages 15–17[a]	0.144
Confirmed violent extremist acts per 100 k population[c,d,e]	0.141
Percentage of women who have heard of HIV/AIDS[a]	0.139
Percentage of women who have ever used contraceptives[a]	0.138
Percentage of girls married before age 18[a]	0.136
Percentage of households owning any animals[a]	0.135
Percentage of households owning a radio[a]	0.130
Percentage of children who have one or more books[a]	0.127
Confirmed or Possible Violent Extremist Acts per 100 k population[c,d,e]	0.119
Percentage of households with a male household head[a]	0.117
Percentage of women who gave birth in a medical facility—public or private hospital, clinics and health facilities[a]	0.117
Under-5 mortality rate[f]	−0.115
Percentage of women whose highest education level is primary[a]	−0.129
Ratio of daughters to sons living at home[a]	−0.132
Percentage of households with metal roofs[a]	−0.183

[a]BBS, 2019; BBS and UNICEF, 2019.
[b]IFPRI, 2016.
[c]Raleigh et al., 2010.
[d]Pettersson & Öberg, 2020.
[e]START, 2019.
[f]BBS and UNICEF, 2019.

indicators. The weighted values of the selected indicators (Table 1) are then used as input into the composite measure to generate vulnerability measures for each zila (units at Bangladesh's second administrative level) (Table 2).

3 | RESULTS

With our algorithm, we successfully calculated a composite vulnerability measure (OECD, 2008) that has a high predictive value for human trafficking in Bangladesh. The high predictive value was confirmed using the reported trafficking cases from the CTCs. Our vulnerability map (Figure 1) correctly identifies known areas of human trafficking, and suggests other locations vulnerable to human trafficking. Vulnerable area are where trafficking is likely to occur or to re-emerge in the future. We further confirmed our analysis through ‘hind-casting’, by testing the model against known occurrences in the past. When the vulnerability measure is hind-casted with known trafficking locations from the CTC, it correctly identifies among the top eight, the six zilas with the highest prevalence of human trafficking. The probability of this occurring by random chance is approximately 1 in 50 000, according to binomial statistics. Similarly, when the vulnerability measure was compared to the number of CTC cases normalized to the population of each zila, the correlation had an R-squared value of 0.37, again confirming a strong predictive relationship.

TABLE 2 TIP vulnerability measures by zila.

Zila	Vulnerability measure	Projected prevalence	Projected Victims	Population	Ranking	Zila	Vulnerability measure	Projected prevalence	Projected victims	Population	Ranking
Satkhira	8.64	15.7	31 261	1 985 959	1	Chandpur	2.71	3.5	8445	2 416 018	33
Jessore	7.70	12.4	34 349	2 764 547	2	Mymensingh	2.53	3.3	17 066	5 110 272	34
Jhenaidah	6.28	8.7	15 332	1 771 304	3	Narsingdi	2.64	3.4	7642	2 224 944	35
Khulna	5.81	7.7	17 814	2 318 527	4	Feni	2.69	3.5	4997	1 437 371	36
Meherpur	5.82	7.7	5048	655 392	5	Sherpur	2.70	3.5	4732	1 358 325	37
Rajshahi	4.98	6.2	16 142	2 595 197	6	Munshiganj	2.66	3.4	4987	1 445 660	38
Chuadanga	5.12	6.4	7274	1 129 015	7	Madaripur	2.67	3.5	4037	1 165 952	39
Kushtia	4.91	6.1	11 899	1 946 838	8	Barguna	2.63	3.4	3057	892 781	40
Bagerhat	4.29	5.2	7716	1 476 090	9	Manikganj	2.53	3.3	4652	1 392 867	41
Joypurhat	4.26	5.2	4737	913 768	10	Cox's Bazar	2.43	3.3	7462	2 289 990	42
Magura	4.20	5.1	4689	918 419	11	Thakurgaon	2.50	3.3	4609	1 390 042	43
Nawabganj	3.97	4.8	7923	1 647 521	12	Noakhali	2.36	3.2	9948	3 108 083	44
Natore	3.95	4.8	8182	1 706 673	13	Pirojpur	2.53	3.3	3716	1 113 257	45
Pabna	3.56	4.3	10 943	2 523 179	14	Panchagarh	2.45	3.3	3233	987 644	46
Rangpur	3.44	4.2	12 120	2 881 086	15	Narayanganj	2.22	3.1	9102	2 948 217	47
Bogra	3.39	4.2	14 124	3 400 874	16	Shariatpur	2.33	3.2	3665	1 155 824	48
Naogaon	3.42	4.2	10 873	2 600 157	17	Netrakona	2.17	3.0	6798	2 229 642	49
Gazipur	3.35	4.1	14 005	3 403 912	18	Kurigram	2.11	3.0	6206	2 069 273	50
Rajbari	3.58	4.4	4573	1 049 778	19	Nilphamari	2.08	3.0	5465	1 834 231	51
Dinajpur	3.32	4.1	12 199	2 990 128	20	Jhalokati	2.22	3.1	2110	682 669	52
Narail	3.61	4.1	3169	721 668	21	Brahamanbaria	1.97	2.9	8238	2 840 498	53
Dhaka	2.95	3.7	44 802	12 043 977	22	Jamalpur	1.96	2.9	6634	2 292 674	54
Tangail	3.18	3.9	14 200	3 605 083	23	Kishoreganj	1.78	2.8	8039	2 911 907	55
Sirajganj	3.13	3.9	12 051	3 097 489	24	Gopalganj	1.78	2.8	3235	1 172 415	56

(Continues)

TABLE 2 (Continued)

Zila	Vulnerability measure	Projected prevalence	Projected Victims	Population	Ranking	Zila	Vulnerability measure	Projected prevalence	Projected victims	Population	Ranking
Patuakhali	3.18	3.9	6056	1 535 854	25	Sylhet	1.56	2.6	8963	3 434 188	57
Barisal	2.98	3.7	8693	2 324 310	26	Khagrachhari	1.74	2.7	1679	613 917	58
Lalmonirhat	3.06	3.8	4798	1 256 099	27	Bhola	1.58	2.6	4662	1 776 795	59
Faridpur	2.95	3.7	7114	1 912 969	28	Maulvibazar	1.46	2.5	4879	1 919 062	60
Gaibandha	2.88	3.6	8681	2 379 255	29	Sunamganj	1.07	2.3	5 686	2 467 968	61
Comilla	2.70	3.5	18 810	5 387 288	30	Rangamati	1.09	2.3	1383	595 979	62
Chittagong	2.57	3.4	25 708	7 616 352	31	Habiganj	0.71	2.1	4391	2 089 001	63
Lakshmipur	2.83	3.6	6234	1 729 188	32	Bandarban	0.43	2.0	761	388 335	64

Note: Corresponding rankings by zila. The ranking colour corresponds to the predominant pixel colour in the map (Figure 1). For Cox's Bazar, we differentiate between the vulnerability of the host population and the vulnerability of the Rohingya. While the vulnerability within the refugee camps is high (and coloured red), the vulnerability of the surrounding population is relatively low.

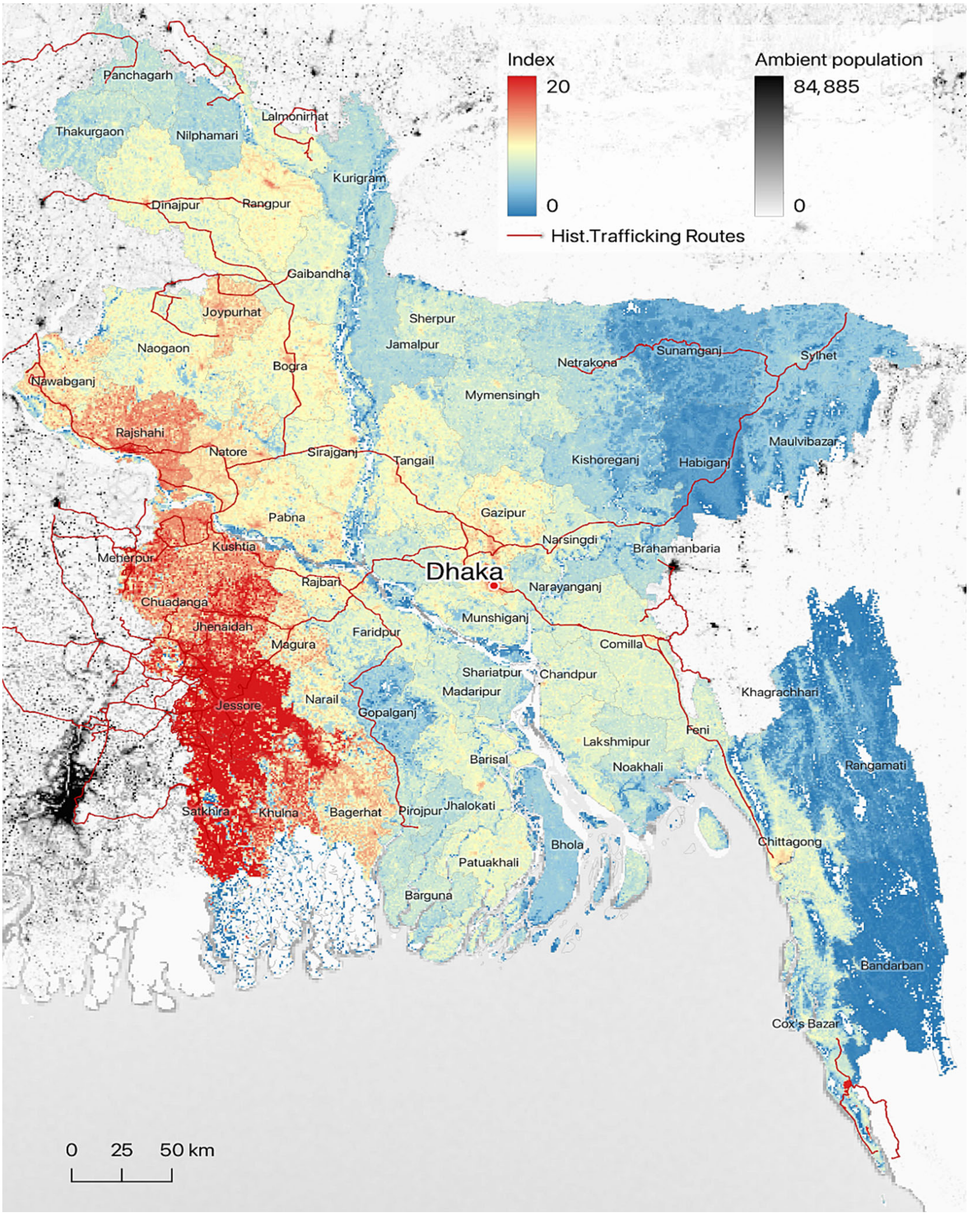

FIGURE 1 Vulnerability to Human Trafficking Normalized to Population with Historical Trafficking Routes. Figure 1: Geospatial presentation of vulnerability to human trafficking, applying a composite measure to values at the zila level and to population estimates at the scale of approximately 1 km^2. The map is composed of 176 449 discrete values. Historical trafficking routes are shown as red lines (Shamim & Kabir, 1998 quoted in Gazi et al., 2001). Within Cox's Bazaar, we differentiate between the high vulnerability within the Rohingya refugee camps [coloured red (Site Management Sector, RRRC, Inter-Sector Coordination Group (ISCG), 2021)] and the significantly lower vulnerability of the host population (coloured yellow-blue). The 2019 ambient population outside Bangladesh is shown in grayscale (Rose et al., 2020). District and zila boundaries are shown in grey and labelled (Bangladesh Bureau of Statistics, 2018). [Colour figure can be viewed at wileyonlinelibrary.com]

Our final calculated vulnerability measure uses a relatively small subset of indicators (18) that represent the optimal combination of characteristics predictive of human trafficking in Bangladesh. These indicators and their respective weightings are shown in Table 1. While correlations between these indicators and human trafficking allow us to explore causal relationships, these indicators by themselves do not necessarily have direct causal relationships with human trafficking. As we discuss below, they are most likely to be proxy measures for more complex phenomena that cannot be measured directly by any single indicator.

We also used our vulnerability measure to estimate projected prevalence and the number of individuals who have experienced human trafficking in each zila (Table 2). We calibrated the estimates using the national estimate of human trafficking for Bangladesh from the Global Slavery Index (GSI), a metric published by the Minderoo Foundation's Walk Free initiative (Walk Free Foundation, 2018). The GSI national estimates have been subject to criticism (e.g., Gallagher, 2017), and the earlier GSI estimates (2012 and 2014) have been withdrawn, due to changes in their methodology. Our zila-level estimates, however, are intended for relative comparisons within Bangladesh, and therefore scale uniformly independent of the specific national estimate for Bangladesh.

In generating zila-level estimates, we rescale the prevalence rates to a logarithmic scaling. We do this because the vulnerability score was calculated from indicators that have been scaled logarithmically to represent properly measurements that range over one of more orders of magnitude. When we transform from the indicator matrix back, we reverse this normalization so that the vulnerability index scales over a realistic range of values. In nation-by-nation comparisons, the GSI human trafficking prevalence estimates are lognormal (van der Vink et al., 2021). Therefore, exponentiating the vulnerability index enables a proper scaling for inferring zila-by-zila prevalence estimates. We then take these exponentiated values, multiply them by the zila population, and use the ratio of this product and the GSI country prevalence estimate to project the prevalence and the number of individuals who may have experienced human trafficking by zila.

We include both prevalence and the estimated number of victims to distinguish between areas with high vulnerability based on their ecosystems, and areas with lower vulnerability, but substantial human trafficking, based on their larger populations. In allocating resources for countering human trafficking, one is often looking to reduce the number of victims in the most cost-efficient manner. One therefore may want to prioritize areas where there is the combination of prevalence rate and population that represents the highest density of potential victims. Prevalence estimates can be misleading when used by themselves. For example, a region with a low prevalence rate but high population can contain more human trafficking activity than a region with a high prevalence rate and low population. The inclusion of projected numbers of potential victims and prevalence along with our vulnerability measures serves to account for population differences.

Our results are summarized in a geospatial visualization of vulnerability to human trafficking in Bangladesh (Figure 1). The geospatial presentation applies our vulnerability measure to 2019 ambient population values, and can be used to develop a probabilistic assessment to predict the number of people within a population that are likely to experience human trafficking. This geospatial presentation is analogous to the vulnerability maps that are used for natural hazards, and should be interpreted in a similar fashion. First-generation hazard vulnerability maps simply used the locations of known past events to predict future vulnerability. As understanding of the ecosystem in which natural hazards occur improved, scientists were able to identify vulnerability in locations where events were previously unknown. Over time, these projections were validated with new events, and the number of hazard victims was reduced dramatically because proactive measures were taken to reduce vulnerability. The vulnerability analysis for human trafficking in Bangladesh follows the same logic. By analysing the ecosystem in which human trafficking occurs, we can assess the potential of other locations to support human trafficking activity, and reduce both present and future victimization through proactive measures.

The scale in Figure 1 is a relative ranking with areas that are most vulnerable to human trafficking shown in red, and areas that have the lowest vulnerability shown in blue. The higher the vulnerability, the increased likelihood of human trafficking. The map is composed of 176 449 pixels, each with an area of 30 m $\times$ 30 m (approximately 1 km^2). We forgo the approach of displaying our results using land-area within an administrative boundary, and instead display our results referenced to population. Such a display helps to identify potential human trafficking hotspots, where there may be large concentrations of vulnerable populations. The geospatial display conveys two messages simultaneously. The colour shade indicates the vulnerability index of the location. The density of colour indicates the sizes of vulnerable populations.

4 | DISCUSSION

Our analysis suggests that vulnerability to human trafficking in Bangladesh is highest among lower-middle to middle class societies with a combination of attributes: (a) moderate levels of income and education, (b) adherence to traditional gender norms of a male-dominated patriarchal society, and (c) access to an urban centre (Table 1). Populations with high vulnerability are characterized as more rural, where a strongly patriarchal culture supports traditional gender roles, but not necessarily areas with poor female education or health. This is consistent with the findings of Paul and Hasnath (2000) where the women's position becomes more marginalized in a male-income-earning culture with large numbers of unemployed men. The vulnerable populations characteristically have access to a city/larger town, which can facilitate higher rates of education, wealth, and access to technology that allows exposure to human traffickers. This is consistent with the findings of Shoji and Tsubota (2022), who also note the proximity to a border or to more highly developed transportation infrastructure as a reoccurring theme in trafficking cases in Bangladesh. As seen in the map presentation, the western border zilas nearest the city of Kolkata have the highest vulnerability to human trafficking. This finding is consistent with previous work that identifies historical trafficking routes through that area (Shamim & Kabir, 1998, as quoted in Gazi et al., 2001).

While one might speculate that rural, poor populations would be most vulnerable to the promise of false opportunities resulting in human trafficking, our analysis reveals more nuanced circumstances. Vulnerable zilas have higher measures of indicators associated with gender inequality and violent extremism than the remaining zilas. The traditional narrative that poverty and unemployment are the main drivers for human trafficking may be an oversimplification, because these indicators exhibit weaker correlations with human trafficking in the dataset. GDP per capita, poverty levels, female paid-employment levels, and wealth quintiles correlate less strongly with both the vulnerability index (and with reported cases of human trafficking) than measures of gender inequality.

Positive relationships between vulnerability and household asset indicators (e.g., radio, bicycle, shingled roofs and children with books) indicate linkages to moderate wealth and education levels. The positive relationship between vulnerability and the indicators of women knowing of HIV/AIDS, women giving birth in a medical facility, and women using contraceptives, combined with a negative relationship with under-5 (U5) mortality, suggests trafficking is less prevalent among the most impoverished zilas. While there is likely under-reporting of human trafficking from this socio-economic group, they may also be outside the sphere of access to traffickers.

Positive relationships between vulnerability and indicators of the percentage of women earning no income due to societal constraints reflect a lack of social and economic mobility in women, which suggests high levels of gender inequality. High child-marriage rates are also a characteristic of gender inequality. A high percentage of male household heads is characteristic of a patriarchal society. The ratio of daughters to sons living at home with their extended families reflects populations where females have low standing in society.

A reported event illustrates our analytical findings. In June 2021, Bangladesh police arrested a trafficking gang that reportedly lured over 1000 women and girls into the Indian sex trade through the social media app TikTok (Khaleej Times, 2021). The victims had access to technology and skill sets to engage with social media, and possessed sufficient resources to meet with the traffickers and to attend parties. In areas of economic growth where women have little societal power and mobility, some may be searching for better lives in less constraining circumstances.

Overall, zilas that have high vulnerability to human trafficking in Bangladesh are characterized by high levels of societal and structural gender inequality. Although the data does not address this directly, one might speculate that trafficking afflicts impoverished populations within zilas that are trending upward economically. Females, perhaps disillusioned by a society that places strict expectations on their lives, may be more likely to seek alternative opportunities and to fall victim to human trafficking, especially if they already have access to information and media, and are close to an urban centre.

Weak-signal analysis reveals several indicators that associate strongly with vulnerability that reflect high gender inequality and traditional male-dominated, patriarchal norms. Indicators that reflect more narrowly female empowerment (e.g., female education) have weaker relationships. Our analysis indicates that efforts to counter human trafficking would benefit strongly from complementing efforts to increase female empowerment with interventions that are targeted specifically at reducing societal gender inequality. Intuitively, such interventions would need to address societal norms that propagate traditional male and female responsibilities, and include young males in female-focused development initiatives before age 10, when gender roles and expectations begin to be imprinted.

Our analysis is specific to Bangladesh, although some of the findings may be applicable to other nations. On a global scale, there is strong evidence that human trafficking is also related to governance. Every year, the US State Department ranks each nation's efforts towards countering human trafficking. The human trafficking reports use a ranking system in which the best-ranked countries are identified as Tier 1 and the worst-ranked as Tier 3. Placement of each country into one of the tiers is based not on the magnitude of the country's trafficking problem, but on the extent of the government's efforts to meet the Trafficking and Violence Protection Act of 2000 (P.L. 106–386) minimum standards for the elimination of human trafficking (22 USC 7106). Researchers have noted the relationship between human trafficking tier levels and forms of governance (e.g., Cho, 2015; McGregor & McEwing, 2013). Statistical analysis demonstrates that a nation's democracy index explains 58.2% of the variance in human trafficking tier ranking, far greater than poverty (at 10.9%) or unemployment (statistically insignificant) (van der Vink et al., 2021). The statistical relationship between governance metrics and tier assignments provides support for those who have called for measures to counter human trafficking to include the promotion of democracy and individual rights (e.g., Landman & Silverman, 2019; Vidwans & Jamal, 2019).

5 | CONCLUSION

Openly available data provide a largely unexploited opportunity for analysing many aid and development challenges through an ecosystem approach. An ecosystem approach can identify vulnerable populations and underlying casual relationships, allowing for geographically-targeted interventions customized for the socio-economic ecosystem in which they will be applied. In Bangladesh, the application of this approach to human trafficking highlights vulnerable populations which have moderate wealth, and provides analytical evidence for those who have argued for promoting gender equality in building more resilient communities.

The combinations of factors related to vulnerability indicates that human trafficking has multiple causes that vary by location. Our analysis is consistent with the observation that many anti-trafficking policies have had limited success because they tend to be applied uniformly and do not account for varying sociocultural and economic conditions (Betz, 2009). We expect this finding to be applicable to many persistent aid and development challenges.

An additional advantage of our ecosystem approach is that the analysis is agnostic. We do not pre-select and combine indicators that we think are related to the problem, nor do we limit our analysis to any single survey or type of data. We input as much relevant data as available and allow the analysis to reveal the combinations of indicators and their relative weightings that are most characteristic of ecosystems where the problem occurs. Previously hypothesized relationships are often confirmed, but the discovery of unexpected relationships is just as common. The discovery of unexpected relationships can lead to a more sophisticated understanding of development challenges, and in turn, offer new opportunities for more nuanced and effective interventions.

ACKNOWLEDGEMENTS

This research was supported by the United States Agency for International Development through contract GS-10F-0033M/7200AA18M00016. Aspects of the analytical approach and methodology were developed with support from the United States Office of the Secretary of Defense contract #W911QX-13-C-0182, and through a research seminar taught by Gregory E. van der Vink at Princeton University. The opinions presented in this paper are those of the authors and do not necessarily represent those of the supporting organizations or any of the institutions with which the authors are affiliated.

CONFLICT OF INTEREST STATEMENT

The authors have no financial, social or competing interest that would benefit directly from the publication of the analysis and conclusions of this manuscript.

DATA AVAILABILITY STATEMENT

Our data sources are all in the public domain and are catalogued by us in the appendix of the manuscript.

Replication data for the analysis are provided on Github (https://github.com/Novametrics/Bangladesh).

ORCID

Sabrina H. Szeto https://orcid.org/0000-0002-3050-837X

Jeffrey Park https://orcid.org/0000-0001-8776-270X

ENDNOTES

[1] The original use of the term 'wicked problem' is attributed to design theorist Horst Rittel.

[2] Despite popular belief, research primarily shows that situational crime prevention does not necessarily lead to crime displacement (Clarke, 1995; Hesseling, 1994).

REFERENCES

Bangladesh Bureau of Statistics (BBS). (2018). Bangladesh - Subnational Administrative Boundaries. Retrieved from https://data.humdata.org/dataset/administrative-boundaries-of-bangladesh-as-of-2015

Bangladesh Bureau of Statistics (BBS). (2019). Bangladesh Sample Vital Statistics 2018.

Bangladesh Bureau of Statistics (BBS) and UNICEF Bangladesh. (2014). *Progotir pathey, multiple indicator cluster survey (MICS) 2012-13: Key findings*. Dhaka, Bangladesh: BBS and UNICEF Bangladesh.

Bangladesh Bureau of Statistics (BBS) and UNICEF Bangladesh. (2019). *Progotir Pathey, Bangladesh Multiple Indicator Cluster Survey 2019, Survey Findings Report*. Bangladesh Bureau of Statistics (BBS).

Betz, D. (2009). 'Human Trafficking in Southeast Asia: Causes and Implications' June 2009. DTIC. Retrieved from http://www.dtic.mi

Cho, S.-Y. (2015). Human Trafficking, A Shadow of Migration—Evidence from Germany. *Journal of Development Studies*, *51*(7), 905–921. https://doi.org/10.1080/00220388.2015.1010158

Clarke, R. V. (1995). Situational Crime Prevention. *Crime and Justice*, *19*, 91–150. http://www.jstor.org/stable/1147596. https://doi.org/10.1086/449230

Eck, J. E., & Clarke, R. V. (2019). Situational Crime Prevention: Theory, Practice and Evidence. In M. Krohn, N. Hendrix, H. G. Penly, & A. Lizotte (Eds.), *Handbook on Crime and Deviance. Handbooks of Sociology and Social Research*. Springer. https://doi.org/10.1007/978-3-030-20779-3_18

Finigan-Carr, N. M., Johnson, M. H., Pullmann, M. D., Stewart, C. J., & Fromknecht, A. E. (2019). A Traumagenic Social Ecological Framework for Understanding and Intervening with Sex Trafficked Children and Youth. *Child and Adolescent Social Work Journal*, *36*(1), 49–63. https://doi.org/10.1007/s10560-018-0588-7

Gallagher, A. T. (2017). What's Wrong with the Global Slavery Index? *Anti-Trafficking Review*, *8*, 90–112. https://doi.org/10.14197/atr.20121786

Gardner, A., Northall, P., & Brewster, B. (2020). Building Slavery-free Communities: A Resilience Framework. *Journal of Human Trafficking*, *7*, 338–353. https://doi.org/10.1080/23322705.2020.1777828

Gazi, R., Chowdhury, Z. H., Alam, S. M. N., Chowdhury, E., Ahmed, F., Begum, S., & ICDDR, B. (2001). In M. Shamsul Islam Khan (Ed.), *Trafficking of women and children in Bangladesh: An overview*. ICDDR,B Centre for Health and Population Research.

Hesseling, R. (1994). Displacement: A review of the empirical literature. *Crime Prevention Studies*, *3*, 97–230.

IFPRI. (2016). *Bangladesh Integrated Household Survey (BIHS) 2015*. International Food Policy Research Institute (IFPRI) [dataset]. https://doi.org/10.7910/DVN/BXSYEL

Joarder, M. A. M., & Miller, P. W. (2014). Empirical evidence on human trafficking and migration-debt contracts in Bangladesh. *Journal of Development Studies*, *50*(3), 399–412. https://doi.org/10.1080/00220388.2013.858128

Khaleej Times. (2021). Bangladesh Trafficking Gang Lured Girls Using TikTok: Police. https://www.khaleejtimes.com/world/rest-of-asia/bangladesh-trafficking-gang-lured-girls-using-tiktok-police

Landman, T., & Silverman, B. W. (2019). Globalization and Modern Slavery. *Politics and Governance*, *7*(4), 275–290. https://doi.org/10.17645/pag.v7i4.2233

McGregor, P. K., & McEwing, L. (2013). How do social determinants affect human trafficking in Southeast Asia, and what can we do about it? A systematic review. *Health and Human Rights*, *15*(2), 138–159. https://www.hhrjournal.org/2013/12/how-do-social-determinants-affect-human-trafficking-in-southeast-asia-and-what-can-we-do-about-it-a-systematic-review/

Ministry of Home Affairs (MHA). (2016). Bangladesh Country Report, 2016: Combating Human Trafficking. Public Security Division, Government of the People's Republic of Bangladesh. Dhaka, Bangladesh. Retrieved from https://mhapsd.portal.gov.bd/sites/default/files/files/mhapsd.portal.gov.bd/page/d9fa01c7_35a1_42f2_827c_1030c3e0e474/Country%20report%202016%20%20fina.pdf

Ministry of Home Affairs (MHA). (2018). Bangladesh Country Report, 2018: Combating Human Trafficking. Public Security Division, Government of the People's Republic of Bangladesh. Dhaka, Bangladesh. Retrieved from https://mhapsd.portal.gov.bd/sites/default/files/files/mhapsd.portal.gov.bd/annual_reports/0cc28f8e_1dcf_4b4e_9a38_264aaa42510c/Final%20_%20Country%20Report%202018%20_%2018.12.19%20(2).pdf

National Institute of Population Research and Training - NIPORT, Ministry of Health and Family Welfare, and ICF. (2020). *Bangladesh demographic and health survey 2017-18*. Dhaka, Bangladesh: NIPORT/ICF. https://www.dhsprogram.com/pubs/pdf/FR344/FR344.pdf

OECD. (2008). *Handbook on constructing composite indicators: methodology and user guide*. ISBN 978-92-64-04345-9. Organisation for Economic Co-Operation and Development.

Paul, B. K., & Hasnath, S. A. (2000). Trafficking in Bangladeshi women and girls. *Geographical Review*, *90*(2), 268–276.

Pettersson, T., & Öberg, M. (2020). Organized violence, 1989-2019. *Journal of Peace Research*, *57*(4), 597–613. https://doi.org/10.1177/0022343320934986

Poncelet, A., Gemenne, F., Martiniello, M., & Bousetta, H. (2010). A country made for disasters: environmental vulnerability and forced migration in Bangladesh. In *Environment, forced migration and social vulnerability* (pp. 211–222). Springer. https://doi.org/10.1007/978-3-642-12416-7_16

Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). Introducing ACLED-Armed Conflict Location and Event Data. *Journal of Peace Research*, *47*(5), 651–660. https://doi.org/10.1177/0022343310378914

Rose, A. N., McKee, J. J., Sims, K. M., Bright, E. A., Reith, A. E., & Urban, M. L. (2020). *LandScan 2019* (2019th ed.). Oak Ridge National Laboratory. https://landscan.ornl.gov/

Site Management Sector, RRRC, Inter Sector Coordination Group (ISCG). (2021). Bangladesh - Outline of camps of Rohingya refugees in Cox's Bazar. Retrieved from https://data.humdata.org/dataset/outline-of-camps-sites-of-rohingya-refugees-in-cox-s-bazar-bangladesh

Shamim, I., & Kabir, F. (1998). *Child trafficking: The underlying dynamics* (pp. 88). Dhaka: Center for Women and Children Studies.

Shoji, M., & Tsubota, K. (2022). Sexual exploitation of trafficked children: Survey evidence from child sex workers in Bangladesh. *Journal of Comparative Economics*, *50*, 101–117. https://doi.org/10.1016/j.jce.2021.06.001

START. (2019). *The Global Terrorism™ Database. National Consortium for the Study of Terrorism and Responses to Terrorism (START)*. University of Maryland.

UNODC. (2010). United Nations Economic and Social Council's (ECOSOC) Guidelines for the Prevention of Crime (Resolution 2002/13) https://www.unodc.org/e4j/en/cybercrime/module-9/key-issues/situational-crime-prevention.html

van der Vink, G. E., Carlson, K. N., Park, J., Szeto, S. H., Zhang, X., Jackson, M. E., & Phillips, E. (2021). Empirical Analysis of the US State Department's Annual Trafficking in Persons Report—Insights for Policy-Makers. *Journal of Human Trafficking*, *9*(2), 148–167. https://doi.org/10.1080/23322705.2021.1897759

Victims of Trafficking and Violence Protection Act. (2000). 22 U.S.C. 7101. https://www.govinfo.gov/content/pkg/BILLS-106hr3244enr/pdf/BILLS-106hr3244enr.pdf

Vidwans, P., & Jamal, M. (2019). Human trafficking has a hidden cause—And it's on the rise worldwide. Time Magazine. https://time.com/5638667/human-trafficking-cause/

Walk Free Foundation. (2018). The Global Slavery Index 2018. https://www.globalslaveryindex.org/

Zimmerman, C., Mak, J., Pocock, N. S., & Kiss, L. (2021). Human Trafficking: Results of a 5-Year Theory-Based Evaluation of Interventions to Prevent Trafficking of Women from South Asia. *Frontiers in Public Health*, *9*, 645059. https://doi.org/10.3389/fpubh.2021.645059

How to cite this article: van der Vink, G. E., Carlson, K. N., Phillips, E., Szeto, S. H., Park, J., & Jackson, M. E. (2023). Identifying vulnerability to human trafficking in Bangladesh: An ecosystem approach using weak-signal analysis. *Journal of International Development*, 1–17. https://doi.org/10.1002/jid.3824

APPENDIX A: TECHNICAL SUMMARY OF METHODOLOGY AND LIMITATIONS ON THE ANALYSIS

Weak-signal analysis provides a means for predicting vulnerability and for identifying underlying causal relationships among multiple interrelated variables. It begins with data fusion through a suite of statistical and regression algorithms for normalization, standardization, and vectorization, subdividing populations into small units for which distinct attributes can be measured. A persistent data-storage layer contains all the raw data in its original form. A virtualized data layer provides an abstraction layer between the physical data sets and the analysis layer. This is where the data are cleaned, standardized, normalized, and vectorized. New indicators are created from the original raw data and stored in logical groupings. Singular-value decomposition is used as an unsupervised self-learning algorithm to identify agnostically key attributes and their relative weightings. The attributes are tested via resampling methods to confirm consistency and sensitivity. The outputs are sets of indicators (weak-signals) that are proxy measures for the underlying causal relationships.

We began with large volumes of data from diverse, mostly open-source datasets from NGOs, media, the U.S. government, and the statistical authorities of local governments. These datasets include detailed national census data, health and educational survey data, remote-sensing data suitable for geospatial analysis, web-scraped data, and data from both formal and informal media sources. The Bangladesh Database contains over half a million socio-economic indicator values covering over 1400 measures for 544 upazilas over 20 years, using 405 257 1 km^2 pixels. Using information from hundreds of millions of data values, we develop hundreds of thousands of human-social-cultural-behavioural attributes differentiated down to the municipality level.

High-resolution geospatial data (typically 1 km^2 for population, but down to 10-m resolution for imagery) and remote-sensing data are converted into tabular data by determining the number of pixels of each data type within each administrative boundary and multiplying the pixel count by the area of each $30'' \times 30''$ pixel (approximately 1 km^2), totalling 405 257 distinct area-patches for Bangladesh. Depending on the data type, either we summed the values, for example, to determine population, or we took a statistical measure of the distribution of values, for example, average travel distance to a road, market, or urban area. For each pixel, population was derived from Oak Ridge National Laboratory's LandScan global population data and represents the ambient population averaged over 24 h.

We calculated indicators from raw survey data at the lower administrative levels. We aggregated and compared these indicators to reported values at higher administrative levels to confirm the accuracy of the aggregation. We then translate the responses into indicators based on the nature of the data. Additional indicators were calculated either by combining two raw indicators in the database, or by calculating the raw data into more meaningful indicators. For example, we calculate the percentage of female teachers, a useful indicator of gender equality, from the reported number of female teachers and total number teachers.

Non-numerical data were reformatted into numerical values and processed statistically. For example, the typical Likert scale was used to survey attitudes with responses like 'Strongly Agree', 'Agree', 'Neutral', 'Disagree' and 'Strongly Disagree', with responses centred about zero. For some indicators where the data was a ranked-choice variable, the choices were converted to discrete numbers between -1 and 1, centred on zero.

Missing data were imputed using linear interpolation or a piecewise cubic polynomial that interpolates the given data if derivatives are specified at the interpolation points. If a region was missing so much data that imputation was unreasonable (based on examining the distribution), and if there were significant events in the region that would make the data no longer representative, it was rejected from the analysis. Judgment was applied depending on the potential value of the indicator and the availability of alternative 'proxy' indicators that might capture comparable phenomena within the socio-economic ecosystem.

Administrative boundaries were sourced from the Bangladesh Bureau of Statistics. Names and boundaries were updated to the most current at the time of data analysis. Duplicate names that refer to different locations were differentiated by appending the name of the next administrative level up.

The historical trafficking routes were described by Shamim and Kabir (1998) based on analysis of newspaper reports, interviews with personnel from law enforcement agencies and direct observations at places of procurement. The route locations were described in tables with the columns, (1) District, (2) Transit upazila, (3) Trafficking route or last transit point in Bangladesh and (4) First transit point in India or Myanmar. The start of the routes are the upazila mentioned as the 'transit upazila', which is where the trafficking journey begins. These routes shown are land routes which were derived by mapping the routes via the road network, except where re-routing to a smaller road versus using the highway would have reduced the distance between the last transit point in Bangladesh and first transit point in India or Myanmar. The border crossings shown on the map are legal crossings; traffickers likely also use illegal crossings. A total of 67 trafficking routes were mapped. A few routes could not be reliably mapped because of ambiguities created by multiple locations with the same name, differences in romanizing the Bengali name (for example 'Vadli' versus the current name 'Bhadli' in Satkira).

The trafficking case reports are taken from the 2016 and 2018 Country Reports on Combating Human Trafficking released annually by the Ministry of Home Affairs. The 2019 and 2020 reports are delayed and not yet released, and the 2017 report is not available, thus only the 2016 and 2018 Country Reports are available from the Ministry of Home Affairs. These case reports come from the Counter Trafficking Committees which are described as follows: 'In every district and tiers of local administration has Counter Trafficking Committees (CTCs) having defined roles and responsibilities and headed by government officials. Government and non-government agencies, members of civil society and local elites are the members of CTCs. At present there are 64 district level Counter Trafficking Committees, 491 upazila level Counter-Trafficking Committees and 4554 union level Counter-Trafficking Committees. Meetings of these CTCs are held once a month. There is a Case Monitoring Committee which monitors human trafficking cases. In 2018, two meetings were held of this committee. During 2018, about 561 human trafficking cases have been filed and the total number of accused persons of human trafficking crime is 2262.' The reports do not specify if the cases are from source, transit or destination zilas. We have integrated the case data with a literature survey to determine if each zila is a source or transit area for human trafficking, as specific zila names have been described as source or transit areas. Any zila which is solely described as a transit area in multiple articles or reports is not included in this analysis (e.g., Brahamanbaria). The database sources include

Rahman, M. S., & Hasan, M. M. (2016). Cross-Border Trafficking from Bangladesh into India: In search of a Framework forProtecting Trafficked Women. Asian Studies, Jahangirnagar University Journal of Government and Politics, 16, 15–24.

Blanchet, T. (2005). Bangladesh Girls Sold As Wives in North India. Indian Journal of Gender Studies, 12(2&3), 306–334. https://doi.org/10.1177/097152150501200207

Gazi, R., & International Centre for Diarrhoeal Disease Research, Bangladesh (Eds.). (2001). Trafficking of women and children in Bangladesh: An overview. ICDDR,B Centre for Health and Population Research.

Siraj, S. (2016). Exploring the Prospects of Community Radio in Bangladesh in Preventing Human Trafficking and Unsafe Migration: A Study on Radio Mahananda 98.8 FM. Global Media Journal, Fall 2016, 14.

Routray, B. P. (2019). Onwards Malaysia: Rohingya focused Human Trafficking Networks Bibhu Prasad. Mantraya. http://mantraya.org/wp-content/uploads/2019/06/Mantraya-Special-

Report_Onwards-Malaysia-Rohingya-focused-Human-Trafficking-Networks.pdf
Rahman, M. M. (2018). Child Trafficking in Bangladesh. International Journal of Research in Economics and Social Sciences (IJRESS), 8(1), 13.
Joarder, M. A. M., & Miller, P. W. (2014). The Experiences of Migrants Trafficked from Bangladesh. The Annals of the American Academy of Political and Social Science, 653, 141–161.
Islam, M. R., & Hossain, D. (2017). Protecting children from trafficking: Responses of the governmental and non-governmental organisations in Bangladesh. The Malaysian Journal of Social Administration, 10(1), 1–28.
Rosy, S. Y. (2013). Trafficking in Women in Bangladesh: Experiences of Survivors and Challenges to their Reintegration [University of Bergen, Norway]. https://bora.uib.no/bora-xmlui/handle/1956/7328
United Nations. (2017, November 14). UN warns of trafficking, sexual abuse in shadow of Rohingya refugee crisis. UN News. https://news.un.org/en/story/2017/11/636002-un-warns-trafficking-sexual-abuse-shadow-rohingya-refugee-crisis
Ahmad, N. (2001). In search of dreams: Study on the situation of trafficked women and children from Bangladesh and Nepal to India (p. 94). International Organization for Migration.
N. M. Sajjadul Hoque. (2010). Female Child Trafficking in Bangladesh: A new form of slavery. Canadian Social Science. Vol. 6, No. 1, 2010, pp. 45–58
ECPAT. (2011). Global Monitoring: Status of Action against commerical sexual explotiation of children, Banlgadesh
R. Amin and R. Sheikh. (2011). Trafficking Women and Children in Bangladesh: A Silent Tsunami of Bangladesh
Bangladesh Country Report, 2016: Combating Human Trafficking. Public Security Division, Government of the People's Republic of Bangladesh. Dhaka, Bangladesh.
Bangladesh Country Report, 2018: Combating Human Trafficking. Public Security Division, Government of the People's Republic of Bangladesh. Dhaka, Bangladesh.

The trafficking case reports are taken from the 2016 and 2018 Country Reports on Combating Human Trafficking released annually by the Ministry of Home Affairs. The 2019 and 2020 reports are delayed and not yet released. To add more context to the data collected, and given the uncertain nature regarding trafficking statistics, we also incorporated qualitative sources such as news reports and published literature. Each time a zila is mentioned in the literature, we tracked whether it is mentioned as a source of victims, a transit area, or both. We have also created a tier ranking system for each zila based on the reported number of trafficking cases per 100 000 population for both 2016 and 2018. The rankings range from 1, for zilas with zero reported trafficking cases in 2016 or 2018 and no mention in any of the literature reviewed, to 9, for zilas with a prevalence of five to ten trafficking cases per 100 000 population. Cox's Bazar, which has a 2016 and 2018 prevalence of 10.39, has been placed into a separate tier for analysis due to the Rohingya refugee population.

Data Preprocessing: Weak Signal Analysis requires preprocessing the data for each indicator used in the analysis. If the indicator distribution resembled a Gaussian distribution, we typically subtracted the mean and normalized by the standard deviation. If the indicator distribution was Log-normal or Chi-squared, we used the logarithm or square-root, respectively. If the data distribution showed clustering asymptotically near an upper limit (e.g., percentages that concentrate near 100%), we subtracted the indicator values from this limit and computed the logarithm or square-root of the differences. We term this transform a 'reverse-log' or a 'reverse-sqrt.' Given limit value X_L and indicator data X_j, we compute scaled values X'_j as

$$\mathrm{Reverse-Log}: X'_j = -\log 10\left(X_L - X_j + e\right)$$

$$\mathrm{Reverse-Sqrt}: X'_j = -\mathrm{sqrt}\left(X_L - X_j + e\right)$$

where e is a small adjustable parameter to avoid singularities at $X_j - X_L = 0$, and the minus sign preserves the ordering of indicator values from smallest to largest. In each case, the rescaling preserved the size-ordering of data values, so that relative comparisons were maintained and the data distribution met the requirement for the statistical analysis.

If data sets had outliers, we winsorized the data to reduce the influence of outsized data values in statistical correlations and regressions. We typically set the outlier values to three standard deviations from the mean, so that they exert strong, but not extreme, influence on statistical computations in the analysis. In some cases, where some data remained skewed in linear, $\log_{10}$, and square-root scaling, with a substantial group (>3%) of indicators beyond 3-sigma, the Z-threshold for winsorizing was set to 4 to preserve the extreme values. Exceptions were applied to indicators whose values clustered in the neighbourhood of an upper bound (e.g., literacy rates, which tend to cluster near 100%, but have tails of values downward towards 0%). In such cases, a reverse-log and a reverse square-root transformation were applied.

Development of the Vulnerability Index: Once the data were cleaned, the indicators were run through a Pearson Correlation Matrix by category for quality assurance and to identify redundant indicators that were highly correlated and did not exhibit sufficient statistical independence to contribute information to the full data set. Singular-value decomposition and varimax rotations were used subsequently as unsupervised self-learning algorithms to identify key attributes and their relative weightings. Thus, the algorithm pares down a large dataset into a smaller one comprised of the most defining and statistically important components. Running the analysis within specific subregions of the nation enables the identification of combinations of characteristics predictive of human trafficking while eliminating the combinations of characteristics that are neither conducive nor preventative. Attributes and attribute-combinations that are prominent in both areas of known high- and low-level human trafficking are thus deemed as inconclusive to human trafficking vulnerability. The attributes are then tested via resampling methods, in which the algorithm is run on different subsections of regions, to confirm consistency and sensitivity. As we want to explain as much of the variance in the data as possible but also avoid having an overly-complicated measure, various threshold values for indicator weightings are used to identify the optimal subset of indicators (OECD, 2008). The weighted values of the selected indicators are then used as input to the composite measure to generate vulnerability measures for each province.

Development of the Projected Prevalence and Victim Estimates: The goal of this step is to rescale the prevalence rates to a logarithmic scaling so as to emphasize the largest values and to relegate the majority of values to relatively low scores where it is plausible that human trafficking activity is low enough to go undetected. The vulnerability score is obtained from the indicator matrix, within which many of the indicators have been scaled logarithmically to decrease small values. When we transform from the indicator matrix back into the real world, we reapply this scaling so that extreme values become extreme values again. In particular, prevalence estimates are typically lognormal in the indicator matrix because they range in many orders of magnitude. Therefore, exponentiating the vulnerability index enables a proper scaling for our inferring zila-by-zila prevalence estimates.

Limitations on the Analysis: Whenever possible, we have attempted to describe the uncertainties associated with the analysis. When presenting the vulnerability index, we have also presented an evaluation of the 'null hypotheses' that fluctuations of indicator values, and their projections onto our vulnerability index, have occurred by random. We use 95% confidence for non-randomness as our threshold for statistical significance, though often the data relationships exceed this threshold greatly.

Although our statistical arguments can be presented in probabilistic terms with associated confidence levels, there are many additional uncertainties due to the nature of our analysis and what we are trying to evaluate. The major limitations are associated with the nature of human trafficking itself.

Although our analysis can compute estimates of Bangladesh trafficking victims in a zila or an upazila down to single individuals, such precision is an untrustworthy artefact of the mathematics. First, the geographic variation in trafficking vulnerability that we estimate across Bangladesh, in particular the relative numbers of trafficking victims within different locations, is subject to a scaling uncertainty, depending on the accuracy of trafficking-victim estimates. Second, trafficking vulnerabilities are probabilistic in nature, expressing likelihoods of human trafficking activity within a location. If the ecosystem is conducive to human trafficking, but no activity has been reported, the activity may be unreported or else the vulnerable populations may not have yet fallen victim.

An analogy with earthquake hazards is useful. Maps of predicted earthquake motion are used to develop building codes, establish insurance rates, allocate resources, and guide development. Even in a region of high probability, no significant earthquake may occur for several years. Alternatively, a single earthquake can cause damage that exceeds the probabilistic values for multiple years. Despite the lack of precision, earthquake hazard maps have been extremely effective in reducing the impact of earthquakes by informing policymakers, insurers, architects, planners, and responders on where to prioritize strategies to reduce vulnerability. The human trafficking vulnerability index should be used in the same manner, focusing policymaker attention on building resilience in the most vulnerable locations, while maintaining baseline programs in regions with lower vulnerability.

Below are the limitations associated with the analytical results. They are listed in a hierarchy based on our assessment of their impact.

1. Ambiguity and differences exist in the terms human trafficking, trafficking in persons, modern slavery, slavery, slavery-like practices and so on. Furthermore, international definitions are not consistent with national definitions and the local customs and laws of a particular country.
2. The results of the statistical models are expressed in probabilistic terms, for which there is debate over required levels of certainty. In our analysis, we quantify probability as the likelihood that a particular result might have occurred by random chance. We reject the 'null hypotheses' (the probability that the result occurred by chance) when the confidence level exceeds 95%. In other words, the probability of the result occurring by chance is less than 1 in 20.
3. We do not assume in our analysis that correlation implies causation. An ecosystem approach to complex, dynamic, and multi-variable problems such as human trafficking treats them as coupled systems that lack true independent variables, but that, nevertheless, offer situations where we can predict outcomes and intervene to effect change. With many variables and many distinct populations, there may be multiple independent correlation patterns. The different patterns indicate that the problem has multiple causes, and the causes vary for different places. In an ecosystem approach, the correlations among population attributes are treated as a coupled system that can be influenced at several points, rather than as a cause-effect process that can be modified only through its dependent variable. The advantage of an ecosystems approach is that it allows us to achieve our objectives by identifying the characteristics to be modified, therefore allowing us to identify options for the interventions that will provide the greatest return on investment.
4. Victims of human trafficking generally self-identify and therefore include subjective assessments that are affected by different sociocultural norms. Survey respondents are not necessarily truthful and their trafficking may not have been independently verified. They may claim to have been trafficked to receive perceived or actual benefits, or they may deny being trafficked to avoid social stigmas or involvement with the legal structure.