

Utilizing big data to measure key connections between gender and climate change¹

Sara Duerto Valero, Ramya Emandi, Jessamyn Encarnacion*, Sneha Kaul and Papa Seck
UN Women, New York, NY, USA

Abstract. Intergovernmental Panel on Climate change (IPCC) reports and other literature draw focus to disproportionate impacts of natural hazards on women and girls; as well as on the roles they play on climate change mitigation, adaptation, natural resource management, environmental conservation and degradation. The connections between climate change and gender are rarely measured and hence, fewer related policies are in place. This paper showcases three separate data collation and analysis exercises, that generated evidence on the gender and climate change nexus utilizing big data: i) geospatial data integrated with Demographic and Health Survey (DHS) data that provide evidence on the connections between climate change and gender outcomes; ii) search and social media data that show violence-related searches increasing at times of hazards; and iii) web document text data that find little mention of gender in climate speeches. Beyond shedding light on the clear connections between gender and climate change issues, it also highlights the important role that non-traditional sources can play in filling these data gaps.

Keywords: Gender, climate change, big data, data gaps, evidence

1. Introduction

The latest Intergovernmental Panel on Climate Change (IPCC) assessment report, published in 2022 [1], underscores that some of the adverse effects of climate change intensify existing vulnerabilities and inequalities. Climate hazards thus affect different population groups differently, depending on intersecting forms of discrimination associated with gender, age, socio-economic class and ethnicity, among other factors. Women and men also have different levels of agency and access to natural resources, playing differentiated roles in natural resource management, conser-

vation and degradation. IPCC scientists agree that narrowing gender gaps may play a transformative role in pursuing climate justice.

Existing literature finds evidence, for example, of linkages between wildfires and gendered norms [2], or between forced marriage and weather events in some countries [3]. However, little evidence is available to prove these connections empirically across countries. Data on the connections between gender and climate change is overall scarce, which contributes to environmental policies being largely gender blind. There are some exceptions, however, where national bodies are making efforts to build climate policies with a gender lens despite lack of data. Ghana formulated the country's National Climate Change Policy by carrying out stakeholder consultations involving a wide range of stakeholders, including gender experts [4]. India acknowledged the gendered impacts of climate change but could not translate them into policy in their National Action Plan on Climate Change. This was mainly due to the lack of appropriate gender data [5].

Numerous international commitments, such as the 2030 Agenda for Sustainable Development, the Sendai Framework for Disaster Risk Reduction, the Paris

¹This paper is submitted by Sara Duerto Valero (Regional Advisor on Gender Statistics – UN Women Asia-Pacific Regional Office), Ramya Emandi (Statistics Consultant, UN Women Headquarters), Jessamyn Encarnacion (Inter-Regional Advisor on Gender Statistics, UN Women), Sneha Kaul (Statistics Analyst, UN Women Asia and Pacific Regional Office) and Papa Seck (Research and Data Section Chief, UN Women) for the Statistical Journal of the Internal Association for Official Statistics.

*Corresponding author: Jessamyn Encarnacion, 220 East, 42nd Street, New York, NY, USA. Tel.: +1 6467814723; E-mail: jessamyn.encarnacion@unwomen.org.

Agreement on Climate Change and the Aichi Biodiversity Targets, emphasize the importance of environmental sustainability for human wellbeing. However, targeted measures that support gender considerations are not included across these commitments and their monitoring frameworks consistently. The Sustainable Development Goals (SDG), for instance, include a specific Goal on gender equality and women's empowerment (SDG 5), and eight Goals pertaining to the environment to various degrees (SDGs 6, 7, 9, 11, 12, 13, 14 and 15). However, the SDG monitoring framework does not include gender-specific indicators in almost any of the environmental Goals. Even where these indicators exist, only a few countries collect and report data on them regularly. The Aichi Biodiversity Targets include a target with a specific mention of women (Target 14), in the context of restoring ecosystems "taking into account the needs of women, indigenous and local communities, and the poor and vulnerable". Still, nearly 62 per cent of countries have not reported on Target 14, and only 29 out of 196 (~ 15 per cent) countries met the target so far [6].

The Paris Agreement, a key outcome of the United Nations Framework Convention on Climate Change (UNFCCC) on its 17th Conference of the Parties (COP), also has a monitoring framework that informed many of the SDG's environment-related indicators, but gender, again, is largely absent. Considering the lack of gender measures pertaining to climate outcomes, gender considerations are captured by the parties in annual reports on processes (e.g., the gender composition of national delegations and policy and decision-making bodies), which are meant to reflect efforts to improve women's representation in UNFCCC processes. The latest gender composition report [7] in 2021, notes women make up to 40 per cent or more of delegates in only 6 of the 16 constituted bodies and only 3 of these bodies have attained parity (i.e., 50% of women's representation).

Attention to gender considerations, however, is slowly increasing. The COP25 Decision 3 of the UNFCCC on gender and climate change stresses the need to increase the "gender-responsiveness of climate finance", including through funding women's organisations [8]. Still, empirical evidence on the challenges women face in the context of climate change, as well as on their agency for natural resource management and environmental decision making, among other areas, remains largely missing and is of utmost importance to determine when and where to invest, and how to do it effectively.

2. Generating evidence on the gender and climate change nexus to drive policy action

This paper showcases three separate data collation and analysis exercises, that generated evidence on the gender and climate change nexus utilizing big data: i) geospatial data integrated with Demographic and Health Survey (DHS) data; ii) search and social media data; and iii) web document text data. Beyond shedding light on the clear connections between gender and climate change issues, it also highlights the important role that non-traditional sources can play in filling these data gaps. A brief description of the rationale and framing of each exercise is included below.

2.1. Generating empirical evidence on the gender-climate change nexus by integrating geospatial and survey data

Climate change poses threats to all aspects of human wellbeing. While the effects are widespread, pre-existing socio-cultural and economic disadvantages are likely to render some women and girls especially vulnerable. Evidence and qualitative studies showcase these connections in select countries, such as India and Bangladesh, where child marriage is known to be affected by rainfall shocks and drought [9,10]. However, empirical evidence on the effects of various climate variables (e.g., aridity, floods, droughts, temperatures, etc.) on women across countries was, to date, largely missing. This study aims to prove the hypothesis that changes in climate-related variables are associated with changes in gender outcomes in Asian countries, as measured by five development indicators that affect women especially: child marriage rates, adolescent births rates, prevalence of intimate-partner violence, access to basic drinking water sources, and access to clean cooking fuels.

Data from the DHS was integrated with geospatial data for Bangladesh, Cambodia, Nepal, Philippines, and Timor-Leste. The findings suggest that, indeed, climate-related factors are statistically associated with gender related outcomes in all five countries, even after controlling for socio-economic variables such as wealth, education and location, which are known predictors of the gender outcomes under consideration.

2.2. Utilizing big data to assess whether violence-related searches increase at times of crises, including natural hazards

Literature shows that the risk of violence against

women and girls (VAWG) rises during crises [11], including during and after natural hazards. This has also been observed in the context of the COVID-19 pandemic, when many victims of violence were confined with their abusers [12,13]. While empirical evidence is crucial to making timely decisions around the issue, collecting data on violence against women (VAW) is especially challenging during crises. Face-to-face surveys are rendered impossible when geographical areas are inaccessible due to natural hazards or conducting in-person interviews carries health and safety risks. Big data analysis can be a useful solution to overcome this challenge. For this study, search engine data were analysed to identify trends and changes in online discourse pertaining to VAWG. In particular, the analysis considered key dates when crises took place, such as typhoons and heavy rain, to identify whether changes took place in VAW related searches. Specifically, the following Pacific Island Countries (PICs) were considered: Kiribati, Samoa, Solomon Islands, and Tonga. Big data showcased that, indeed, peak periods of VAWG searches and posts overlapped with times when climate crises occurred during the pandemic.

2.3. *Assessing the inclusion of gender-related references in climate policy commitments of UNFCCC Parties using web text data*

The UNFCCC COP26 included Decision 20 (para. 3) [14] with a voluntary call to Parties and observers for submitting reports showcasing progress in the implementation of gender actions and plans at all levels. As of 30 April 2022, only 15 of 197 parties and observers had submitted reports on this [15]. Monitoring whether gender is integrated into national climate change policies, plans, strategies and actions will require a structured mechanism, along with advocacy efforts to highlight the importance of countries voluntarily submitting reports. Mentions of gender-related issues in COP speeches are expected to, ultimately, contribute to more gender-sensitive decisions. In the absence of data on the implementation of gender-sensitive actions in many countries, looking at COP speeches can provide insights regarding parties' intentions and priorities when it comes to mainstreaming gender across climate change work. This study examined the text in speeches provided by parties during COP 24, 25 and 26 (held in 2018, 2019 and 2021, respectively). It found that mentions of gender-related terms are exceptionally rare.

3. **Relevance of this study to statisticians and policy makers**

To provide empirical evidence on the climate change – gender nexus, including to inform related policy making, this paper utilizes big data to fill existing data gaps. In the absence of official statistics that adequately capture the gender-environment nexus in a holistic manner, efforts to gather information are urgent. Although national data producers, including national statistics offices, disaster management agencies and ministries for environment are increasingly shifting efforts and resources to the production of these data, the process of generation of these statistics using surveys and administrative data is time intensive. For instance, countries that have recently implemented gender-environment surveys utilizing UN Women's Model Questionnaire on Gender and the Environment [16], found that data collection at a national scale required between one and three months, and interviews had to be conducted among two adults of different sex per household, in order to obtain reliable estimates. As environmental policy-making continues to evolve, non-traditional data sources can provide insights on some of these connections in the absence of official statistics. Furthermore, as statisticians further their skills to use big data to supplement existing official statistics, the use of these sources may prove efficient to generate gender-environment estimates in select areas, such as linkages between slow-onset climate events and gender related outcomes.

4. **Data sources**

4.1. *Generating empirical evidence on the gender-climate change nexus by integrating geospatial and survey data*

DHS collects information on, among other variables, women's demographic and socio-economic characteristics, including age at first marriage, age at first birth and intimate partner violence. In addition, they gather information regarding the type of water source, time to water source, and type of cooking fuel used in households where women live. This information was used to construct binary categorical variables on child marriage, adolescent births, exposure to intimate partner violence in the past 12 months, availability of basic water services, and access to clean cooking fuels. These were treated as dependent variables to prove the hypothesis that gender outcomes are affected by climate variables.

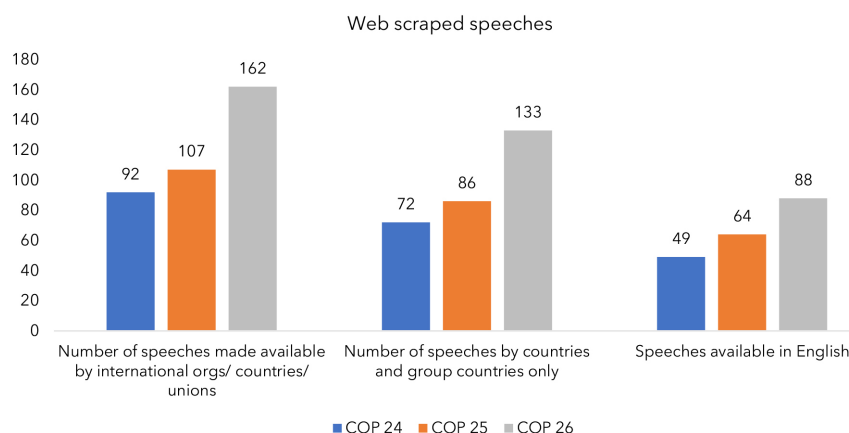


Fig. 1. Total number of web scraped speeches of UNFCCC COP member Parties, 2018–2021.

Vector data and raster data on climate-related information were integrated with DHS data for each cluster in the dataset. The DHS Program routinely tags survey cluster locations to ancillary data (known as geo-covariates). This includes information on droughts, aridity, temperatures, and rainfall. Information on other relevant climate-related factors, such as flood risk and proximity to water, was also integrated. Data on flood risk in 50 years was drawn from UNDRR’s Global Disaster Risk Dataset [17]. Information on proximity to the nearest water source – nearest lake (fresh water) and nearest coastline (salt water) was obtained from the Global Self-consistent, Hierarchical, High-resolution Geography Database [18].

Control variables were identified based on findings from the literature review to account for non-climate related factors that could potentially affect gender related outcomes. In particular, educational attainment in single years, women’s employment status at the time of the survey, age, wealth quintiles (based on DHS’s wealth index), household location (urban/rural), buildup index (degree of urbanization at the cluster level), proximity to national borders, and proximity to protected areas (such as national parks, national forests and national seashores) were all held constant in the models to assess the individual effect of each of the other variables on the gender-related indicators.

4.2. Utilizing big data to assess whether violence-related searches increase at times of crises, including natural hazards

The analysis considered data generated by search engines (Google, Bing, Yahoo) in four PICs. A total of 3,360 keyword strings on the topic of VAWG were

identified as relevant for the analysis. These ranged from general VAWG related word strings, such as “what is VAWG”, “how to cover bruises on my face” and “my husband abuses me”, to help-seeking searches such as “VAW hotline” or “help for domestic violence victims”. Utilizing these criteria, the data pull generated a total of 56,410 unique searches across all four countries. Search behaviour was analysed for a one-year period (June 2020 to June 2021) and thus produced findings on search behaviour prior to, during and after crises. To identify all climate-related and other crises that took place during the reference period, several publicly available sources [19] were utilized. These helped identify peak periods of lockdown, political instability, and the onset of natural hazards. The analysis assessed whether VAW related searches peaked during crises, and whether these searches were general or help-seeking searches in each case.

4.3. Assessing the inclusion of gender-related references in climate policy commitments of UNFCCC Parties using web text data

To assess whether gender considerations were mentioned during climate change talks, this study web scraped across speeches given during UNFCCC COP 24, 25 and 26 (held in 2018, 2019 and 2021, respectively) [20], available openly on UNFCCC’s website [21]. Speeches from previous COP conferences were not made publicly available in a consistent manner, and thus these have not been included [22]. Further, only speeches available in English were utilized for this analysis.

Refer Fig. 1. Overall, a total of 201 digital speeches [23] were web scraped and analyzed (49

Table 1
Description of variables for logistic regression analysis

Type of variable	Name of variable	Description of variable
Dependent variables	Child marriage	Age at first marriage or cohabitation lower than 18
	Adolescent births	Age at first live delivery lower than 18
	Intimate partner violence (IPV)	Experience of severe or less severe physical or sexual violence from an intimate partner in the 12 months preceding the survey
	Lack of access to basic water	Women living in households that lack piped water (into their dwelling or yard), or other improved sources such as public standpipes, tube wells, boreholes, protected wells, rainwater or bottled water, within 30 a minute round trip
	Use of unclean cooking fuels	Women living in households that use unhealthy cooking fuels, such as kerosene, charcoal, wood, straw/shrubs/grass, agricultural crops, or animal dung
Independent variables	Average day land surface temperature	Mean annual daytime land surface temperature (skin temperature) measured in Degree Celsius
	Drought episodes	Magnitude of monthly precipitation less than or equal to 50 per cent of its long-term median value for three or more consecutive months
	Average annual rainfall	Average amount of rain per year (in millimetres)
	Aridity index	Average yearly precipitation divided by average yearly potential evapotranspiration
	Probability of riverine flood	Estimated probability of major river basins flooding in 50 years
	Proximity to water bodies	Geodesic distance to either a lake or oceanic coastline, in meters. This variable was divided into two (lake/coastline) to analyse the effects separately
Control variables	Education	Educational attainment in single years
	Employment	Women's employment status at the time of the survey
	Age	Women's age at the time of the survey
	Wealth	Wealth quintiles (based on DHS's wealth index)
	Location	Household location (urban or rural)
	Buildup index	Degree of urbanization at the cluster level
	Proximity to borders	Geodesic distance to international border, in meters
	Proximity to protected area	Geodesic distance to protected areas, such as national parks, national forests, and national seashores

for COP24, 64 for COP25 and 88 for COP26). The study considered individual country speeches and group speeches (e.g., European Union, ASEAN).

To conduct this analysis, approximately 260 keywords [24] related to gender were indexed and searched. These were identified utilizing relevant documents such as commitments of the Action Coalition on Climate Justice and Measuring References to Statistics in National Policy Documents, which is a result of UN Women-PARIS21 collaboration [25].

5. Methodology

5.1. Generating empirical evidence on the gender-climate change nexus by integrating geospatial and survey data

To assess whether gender related outcomes are affected by changes in climate-related variables, two statistical models were used. Namely, logistic regression analysis was first conducted and showcased associations between climate variables and gender outcomes. This was followed by analysis using random forest models, which showcased the variables with the highest power to classify groups of women between high and low risk

of experiencing each of the gender related outcomes under consideration.

5.1.1. Logistic regression analysis

The study considered five gender-related indicators [26] as dependent variables and seven climate change related factors as key independent variables (refer Table 1). Variables were selected based on existing literature, relevance, and availability of comparable data across the five countries. Since many other socio-economic factors are likely to influence gender outcomes, eight control variables were included in the model.

All dependent variables were constructed as dichotomous categorical, and they all have a negative connotation (e.g., increases in the prevalence of each phenomena represent worsening of women's wellbeing). Binomial logistic regression analysis was thus determined best suited for this analysis. Multivariate logistic regression was used to observe the effect of all independent variables, including climate change related and control variables, on gender-related outcomes.

The final model included the control variables in addition to the climate-related variables. This resulted in the adjusted odds ratios. The Residual Deviance and Akaike Information Criterion values were compared for these models to evaluate the goodness of fit [27].

5.1.2. Random forest models

Random forest predictive algorithm (a machine learning technique) focuses on prediction instead of association. These models are thus powerful when dealing with complex and non-linear associations. Random Forest is a tree-based learning algorithm that uses a collection of decision trees to perform classifications tasks, with the primary objective of increasing the accuracy of the predicted outcomes, given the input data. In the context of this analysis, random forests compute multiple decision trees for classifying women into high or low risk of experiencing child marriage, adolescent births, IPV, lacking basic water and clean fuels, given their exposure to different climate variables. As a result, the model identified the climate variables most important in predicting gender-equality outcomes (measuring the mean decrease in the accuracy of the model if each of the climate variables was removed).

5.1.3. Limitations

Due to the localized nature of environmental events (e.g. different magnitudes, duration, and intensity) and various degrees of coping capacity among exposed population (depending on asset ownership, existing infrastructure, suddenness of events), making cross country generalizations is difficult. Furthermore, the choice of variables included in the models was restricted by data availability, so future research should control for the effects of asset ownership, quality of infrastructure, or whether weather events took place out of season, to ascertain the sole impact of climate variables.

Findings from the random forest and logistic regression models do not always pinpoint to the same climate variables as the most relevant for explaining changes in gender outcomes. This is because they each serve a different purpose. While logistic regression identifies the odds of an increase in a gender-related outcome given a one unit increase in a climate variable, random forest highlights the reduction in the model's accuracy if a climate related variable was removed from the model. In addition, while logistic regression performs well in models where explanatory variables have large variances, random forest returns higher true positive rates for datasets with numerous noise variables [28].

An additional limitation has to do with data timeliness [29]. In most countries, the reference year for the latest GIS data and survey microdata differ, and thus it is impossible to study direct impacts of select weather events on a population group. Rather, the analysis aims to ascertain whether geographical areas and population groups that typically suffer from slow-onset or recur-

rent effects of climate change, are more likely to experience changes in gender outcomes overall. Also related to timeliness, since some climate related phenomena change rapidly, more recent data would provide a more accurate picture of the current situation.

5.2. Utilizing big data to assess whether violence-related searches increase at times of crises, including natural hazards

Big data from on-line searches in Kiribati, Samoa, Solomon Islands and Tonga was analysed to identify potential trends on general violence-related searches, and help-seeking behaviour. The data engineering process considered searches made between June 2020 and June 2021, and peaks in VAWG searches were compared against peak periods of mobility restrictions, climate hazards or disasters during the same period. As illustrated in Fig. 2, some of these crisis events were punctual, while others spanned over several months. This was taken into account when comparing with trends in VAWG searches.

To achieve this, the analysis first identified key words about various acts of violence from DHS questionnaires. This list was complemented with additional words from a literature review, and suggestions from staff in UN Women country offices to ensure inclusion of words used in local cultural contexts. The wordlist was tested to determine which keywords produced on-line volume, and historical search data for each keyword was extracted for a month-to-month comparison. Both English and local languages were used for the analysis.

The identified patterns in search terms were categorized as:

- Physical violence: Search terms specifically related to, or indicating signs of, physical violence such as 'beaten by husband', 'he hit me for the first time', 'how to cover bruises'.
- Psychological violence: Search terms specifically related to or, indicating signs of, psychological violence such as 'husband yells at me', 'my boyfriend makes me cry', 'husband insults me'.
- Sexual violence: Search terms specifically related to, or indicating signs of, sexual violence such as 'help for rape victim', 'rape hotline', 'force grope'.
- Any of physical, psychological or sexual violence: All other search terms which indirectly suggest some form of VAWG but could not be specifically categorized into the above-mentioned categories such as 'ladies police helpline number', 'report domestic abuse', 'women harassment helpline'.

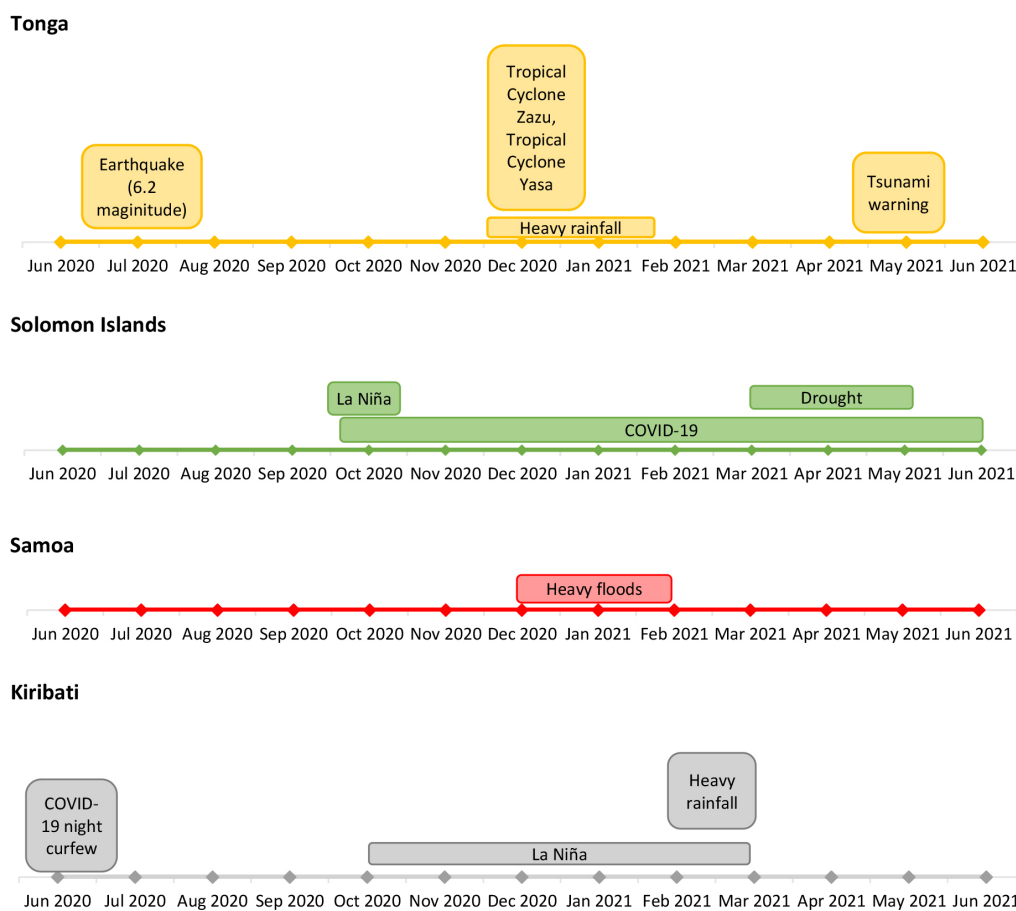


Fig. 2. Peak periods of interest (mobility restrictions, climate hazards, disasters) for VAWG analysis, by country.

The set of 1,330 unique search terms were also classified according to the purpose of the search behaviour as:

- General information searches: Search terms such as ‘is my boyfriend abusive’, ‘signs of abusive relationship’, ‘what is an abusive relationship’ etc. that directly indicate information-seeking behaviour.
- Survivor/victim searches: Search terms such as ‘my partner hit me’, ‘divorcing an abusive husband’, ‘he forced me to have sex’ that directly indicate behaviour that could be displayed by a survivor/victim of VAWG.
- Help-seeking searches: Search terms such as ‘abuse helpline’, ‘domestic violence counselling’, ‘marital abuse help’, etc. that directly indicate help-seeking behaviour.

Volume of searches for each of the categories was then compared over time, and trends between search behaviour and crisis events were compared to assess correlations.

5.2.1. Limitations

Since this analysis utilizes online VAWG-related searches only, values should be interpreted with caution. While search trends are expected to mirror trends in actual incidence of violence to a certain extent, they cannot be used to assess the incidence of VAWG with accuracy. The overall number of searches, in addition, is highly dependent on varying degrees of internet penetration across the countries considered, which hampers comparability in terms of total search volume. Furthermore, in contexts where internet penetration is low, it is likely that those with access to the internet are the more empowered women, and thus VAWG cases identified through this analysis are likely severely underestimated.

5.3. Assessing the inclusion of gender-related terms in climate policy commitments of UNFCCC Parties using web text data

Two types of analysis were conducted to assess the

degree of inclusion of gender-related terms across COP speeches. A quantitative analysis was implemented first followed by sentiment analysis that provided qualitative insights.

5.3.1. Quantitative analysis: Detecting gender references

COP speeches available in English were first screened for gender references. The identification of gender references was based on a list of keywords developed by the authors. Words such as ‘mother’ and ‘ggi’ [30] are relevant both from a gender and a climate change perspective but were found in the speeches *only* in the context of climate change. Hence, they were not considered as gender-relevant to avoid overestimation. Algorithm allowed for words such as ‘gender’ as well as other words that appear within the former (e.g., ‘engender’) were all considered. Furthermore, certain words that would only be of relevance to gender equality and women’s empowerment when used in combination with other words (e.g., ‘birth’ to ‘adolescent birth rates’), were also considered only as word strings.

The frequency of sentences with gender references was calculated, instead of counting the number of gender-related words. This avoids inflating the results of speeches that use more gender references rather than utilizing gender references more consistently.

5.3.2. Qualitative analysis: Topic modelling and sentiment analysis

To assess the degree to which each of the speeches covered gender issues substantively, two methods were used. Firstly, through topic modelling, a set of abstract topics were identified built on frequent words that seemingly point to certain topics. After identifying the sentences with gender references, these were processed to convert into lower case and remove stopwords [31]. Latent Dirichet Allocation (LDA) technique was used to generate speech-word joint probabilities, that is, the Bayesian probability for each word to fall in each of the abstract topics was calculated, based on the prior distributions introduced over speech-topic and topic-word distributions. The words with the highest probabilities within each topic were shortlisted and grouped into clusters. Based on the group of words produced by the model, topics were identified normatively.

Secondly, to assess the sentiment of the gender related sentences, R software packages *Bing* and *Loughran* were utilized. The analysis of the content is ensured to contain stopwords, to detect negative sentiments.

6. Findings

6.1. Generating empirical evidence on the gender-climate change nexus by integrating geospatial and survey data

6.1.1. Logistic regression analysis

Findings from the logistic regression analysis indicate that increases in the frequency of drought episodes, rises in temperatures and increasing relative aridity exert a statistically significant and negative effect on many gender outcomes (Fig. 3). In particular, temperature rises are associated with more prevalent child marriage and adolescent births across most countries [32]. More frequent drought episodes are linked with lack of access to basic drinking water sources (Bangladesh, and Cambodia), and lack of access to clean fuels for cooking fuels (Bangladesh, and Philippines) [33]. Increases in relative aridity (low aridity index) [34] also correlate with worsening gender outcomes (child marriage and adolescent birth rates in Bangladesh and Nepal), although these effects are smaller in size in all countries except for Timor-Leste (partner violence, access to basic water sources and use of unclean fuels increase substantially when relative aridity increases).

Drought episodes: Increases in the frequency of droughts enhance the likelihood of negative gender outcomes in most countries. Differences across locations can be explained by factors such as whether the region is arid or semi-arid, whether drought episodes are short or long-term, whether the effects are localized or widespread [35,36], and the overall dependence on agriculture [37]. The strongest associations for drought episodes, across countries, were found with access to basic drinking water sources and clean cooking fuels. This carries gendered consequences as it may exacerbate women’s time burdens for fetching water, cooking and providing hygiene for the household. Although in Nepal the relationship appears inverse (more frequent drought episodes correlate with higher likelihood of accessing basic water sources (AOR = 0.67***)), geospatial analysis reveals that drought-prone clusters are largely located in or near Kathmandu, where water provision infrastructure is better established. Further analysis with controls for availability of water infrastructure would be useful to establish the sole effect of droughts.

In Bangladesh and especially the Philippines, the associations between frequent droughts and lacking access to clean cooking fuels see some of the largest coefficients in this study (Bangladesh AOR 1.28***; Philip-

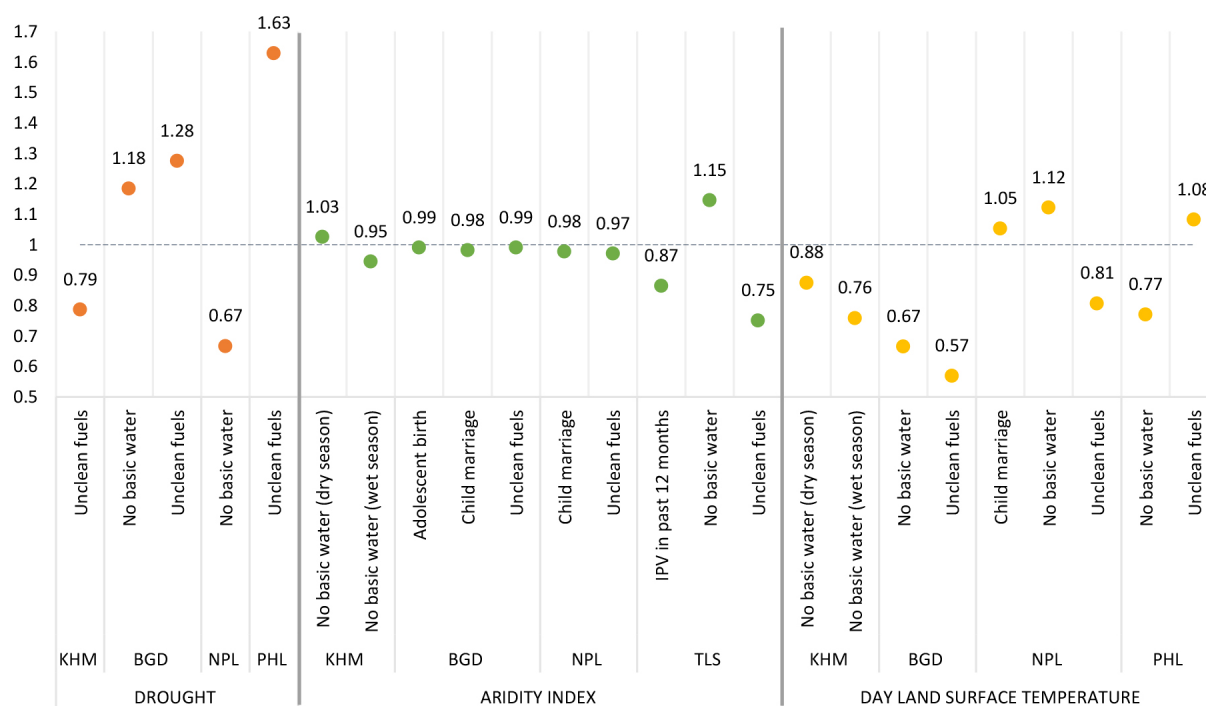


Fig. 3. Adjusted odds ratios ($p < 0.000$) for the effect of drought, aridity (index), and day land surface temperature across gender outcomes, by country. Note: For brevity, the graph represents adjusted odds ratios for associations significant at $p < 0.000$. The dotted horizontal line represents an odds ratio of 1 (to represent no association). The 3-digit ISO country codes stand as follows: KHM (Cambodia), BGD (Bangladesh), NPL (Nepal), PHL (Philippines), and TLS (Timor-Leste).

piners AOR 1.63***). Fuel cost, among other factors, may contribute to this association. Water scarcity limits energy production by hydropower and coal plants, which are highly dependent on water for cooling, which may rise energy prices. Further, drought-driven economic stresses from reduced agricultural yields may push households to rely on cheaper (unhealthy) cooking fuels. In Cambodia the inverse association (AOR 0.79***) is likely influenced by provincial differences in electricity prices, driven by grid development, availability of high-voltage transmission connections, existence of hydropower dams and other factors influencing provincial tariffs [38].

Relative aridity: Relative aridity [39] is also significantly associated with gender-related outcomes, with smaller effects across countries except in Timor-Leste, where clear associations exist with the prevalence of intimate partner violence (IPV), access to basic water sources and use of clean cooking fuels. This is a concern given that Timor-Leste already sees the highest prevalence of IPV among the five countries considered (an estimated 34.6 per cent of women experienced physical and/or sexual violence at the hands of their partner in the 12 months preceding the survey). A prolonged

history of conflict, widespread poverty and stress associated with bride price, may contribute to the high IPV prevalence in the country [40,41]. Although further analysis is necessary, it is possible that aridity-related economic stresses may be worsening poverty, increasing pressures for meeting bride price, and ultimately contributing to rises in IPV (Fig. 4).

Associations seen between increases in relative aridity and lack of clean water [42] and fuels align with existing literature on the topic, as changes in aridity and humidity typically affect the availability and quality of aquifers and other water sources, as well as the cost and affordability of fuel and water. Although the effects are smaller, relative aridity is also associated with the availability of basic water sources in Cambodia, where reliance on rainwater makes some groups of women more vulnerable.

Cambodia's two marked seasons (dry and wet) affect water consumption patterns in areas where basic water infrastructure is not available. For instance, while tube wells (an improved source) are the most common drinking water source in arid clusters throughout the year, many women in humid clusters rely on rainfall (an improved source) during the wet season but shift to un-

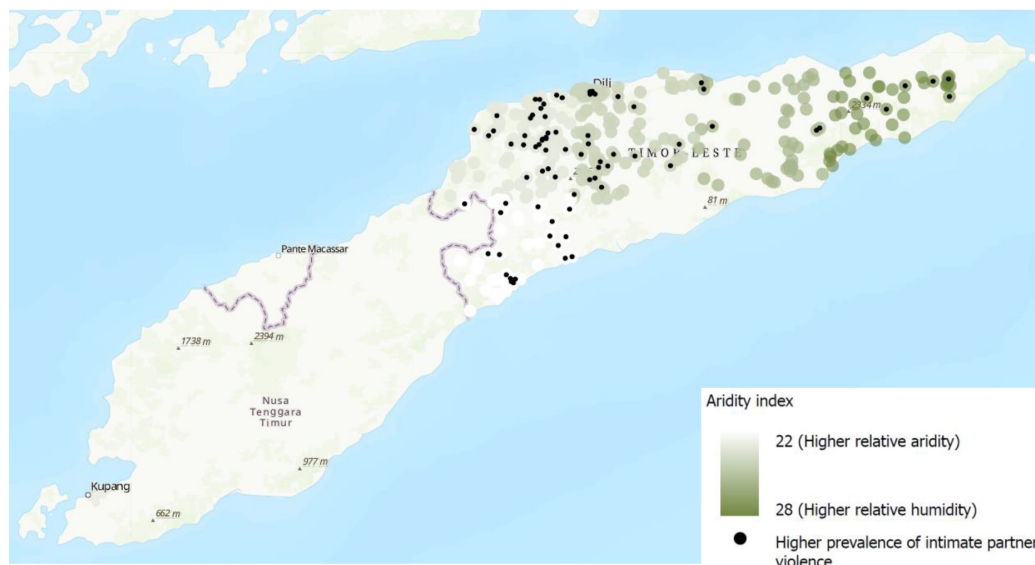


Fig. 4. Geographical distribution of clusters with high rates of intimate partner violence in the past 12 months, by level of relative aridity, Timor-Leste. Key for interpretation: The black markers represent clusters with higher prevalence of intimate partner violence (top 25% of cluster values). The map indicates that arid clusters are more likely to see higher rates of intimate partner violence in the past 12 months.

protected wells and open surface water during dry seasons. As a result, increases in aridity are seen to lower the likelihood of women lacking basic water sources during the dry season but increase it during the wet season (AOR = 1.03*** during the dry season, AOR = 0.95*** during the wet season). This highlights the risks posed by climate-change, as shifts in relative humidity are likely to further alter the predictability of rainwater availability, affect water quality, and increase related collection and treatment burdens on women.

Day land surface temperature: Increases in average day land surface temperature typically present along land degradation [43], sea level rise, reduced agricultural yield, forest transition, biodiversity loss and human disease [44]. They may also take place in the context of rapid urbanization and urban heat islands. Analysis shows that temperature rises are strongly connected with most of the gender-related outcomes considered across almost every country, and although the strength of associations varies, in most cases the effects are large.

In Cambodia and Nepal, increases in day land surface temperature enhance the odds of child marriage significantly. This is likely linked to their effects on agricultural yield and other natural resources, which contribute to economic insecurity and the use of child marriage as a coping strategy [45,46]. In Nepal, this association is stronger (AOR = 1.05***), and heavily concentrated across warm clusters in the south-eastern Terai region, where prevailing social norms and traditions are overall

supportive [47] of child marriage practices. Controls for wealth, education and other known drivers of child marriage do not change this finding, indicating that indeed temperature may have an effect, particularly in areas where the practice is already prevalent.

In Bangladesh, Cambodia and the Philippines, increases in temperatures correlate with better access to basic drinking water sources. In areas where improved water infrastructure is not available, the effects of temperatures on freshwater availability and contamination may be contributing to these associations. Temperature rises typically present simultaneously with fresh water source depletion (e.g. wetland drying, conversion of freshwater wetlands to agricultural land) and land degradation (e.g. dry forest clearing, salinization) [48], which may result in poor rain absorption and thus runoff draining into waterways and polluting the water supply [49]. Consequently, populations living in high temperature clusters without access to piped water, may shift from drinking river and stream water to improved sources such as bottled or rainwater.

Conversely, in Nepal, increases in temperatures correlate with increased difficulty accessing basic water sources, but disparities in the availability of water infrastructure appear to be affecting this association. Across the Hill and Mountain regions of the country, water supply systems are mostly spring-related sources reliant on precipitation. Especially in the west and central Mountain and Hill regions, many of those who

lack piped water rely on unprotected wells and springs. However, across the southernmost provinces (along the border with India), which are further from mountain springs and see higher temperatures, tube wells and hand pumps (improved sources) are widely available.

In the Philippines, as expected, increases in day land surface temperatures also heighten the risk of using unhealthy fuels. Beyond major cities such as Metro Manila and Cebu (which have better fuel infrastructure and LPG availability), clusters reliant on clean cooking fuels are more likely to be located in cooler areas. Economic strains associated with land degradation and heat-driven reductions in agricultural yield and biodiversity loss, coupled with increased electricity and gas prices as a result of heat [50], may be prompting people in hot clusters to shift to cheaper but dirtier fuels.

Unexpectedly, regression shows that the odds of accessing clean fuels are higher in high temperature clusters in Bangladesh and Nepal. In Bangladesh, this association is affected by fuel infrastructure and availability (e.g. people living near natural gas fields, which are warmer areas, have better access to this type of clean fuel); while in Nepal, the country's topography makes it difficult for residents of mountainous (cooler) areas to access clean fuel infrastructure or purchase LPG. Further analysis controlling for the availability and cost of clean fuels could provide a clearer picture of the associations between clean fuel use and temperature alone.

6.1.2. Random forest model

Besides droughts, temperatures and aridity, results from the random forest model highlighted additional variables of importance for gender outcomes, such as proximity to lakes, proximity to coastlines, and rainfall.

Child marriage: Across all countries, proximity to water (lakes or coastline depending on the country) is among the top five most important variables for classifying women into high or low risk groups for experiencing child marriage (Fig. 5). Living near water can have both positive and detrimental effects on poverty, livelihoods and consequently child marriage, depending on the context. For instance, sea level rise (and soil salinization) may be detrimental to the livelihoods of populations living near coastlines, but proximity to coasts may be advantageous for those engaged in fishery-related livelihoods. Relative aridity is also important to assess the risk of child marriage in Bangladesh, Nepal and the Philippines, likely due to its effects on agricultural yield, food security and other poverty-related indicators. Strongly connected with relative aridity, rainfall

also appears important in models in Nepal and, to a lesser extent, Cambodia, Philippines, and Timor-Leste, reinforcing this finding.

Adolescent birth: Although logistic regression only showed mild associations between the different climate-related variables and adolescent births, random forest highlights that climate-related variables do have influence on the risk of adolescent births. The prediction accuracy of adolescent birth models decreases when environmental variables are removed, but less so than in other models. Factors such as culture, religion, limited access to family planning and poor reproductive health services provision, are all known to contribute to adolescent births [51]. Because no data was available to control for these variables at the time of the analysis, these interactions may explain the weaker explanatory power.

Distance to water (nearest lake or coastline) was highlighted as the most important climate-related factor for determining the risk of adolescent births across countries, in line with findings from the child marriage model. The importance of relative aridity to explain the risk of adolescent births was also highlighted in Bangladesh and Nepal, an expected result as it correlates with agricultural yield, poverty, and thus child marriage and adolescent births in these countries, where social norms enable child marriage practices. Environment-related variables have higher explanatory power on adolescent births in Bangladesh and Nepal compared to other countries. As the practice is overall more prevalent in these two countries, climate change may be acting as an amplifier of currently existing practices.

Intimate partner violence: Among the climate-related factors considered, average annual rainfall is one of the variables with relevance across countries (Cambodia, Philippines and Timor-Leste), with model accuracy dropping by 7 per cent in Cambodia and 11 per cent in Philippines and Timor-Leste if the variable is removed. The model also highlighted the importance of proximity to lakes, coastlines and protected areas, with mild effects across countries. Existing literature indicates that the drivers of IPV are multidimensional and highly localized [52], with social norms playing a significant role. As such, although rainfall, proximity to water, and protected areas may contribute to healthy natural resources, decreased economic stresses and thus reduced IPV, the reasons behind these associations are difficult to establish. Further research, including controlling for other known drivers of IPV, could help better understand these reasons.



Fig. 5. Mean decrease in accuracy (in event of variable removal) for child marriage rates, by country. Note: Environment-related variables are shown in darker shade of blue for ease of interpretation.

Access to basic drinking water sources: Climate related variables are powerful for classifying women into groups with high or low risk of lacking access to basic water sources (Fig. 6). Distance to nearest water sources, in particular, is among the top 5 most relevant variables in all countries considered. However, living near a lake can imply easier access to livelihoods and infrastructure,

but also increased exposure to water contamination. For instance, freshwater lakes are valuable resources for travel, trade, fishing and irrigation and thus, living near lakes may facilitate income generation, and consequently, affordability of improved water sources and infrastructure. However, declines in freshwater quality and increases in water stress may be putting access to

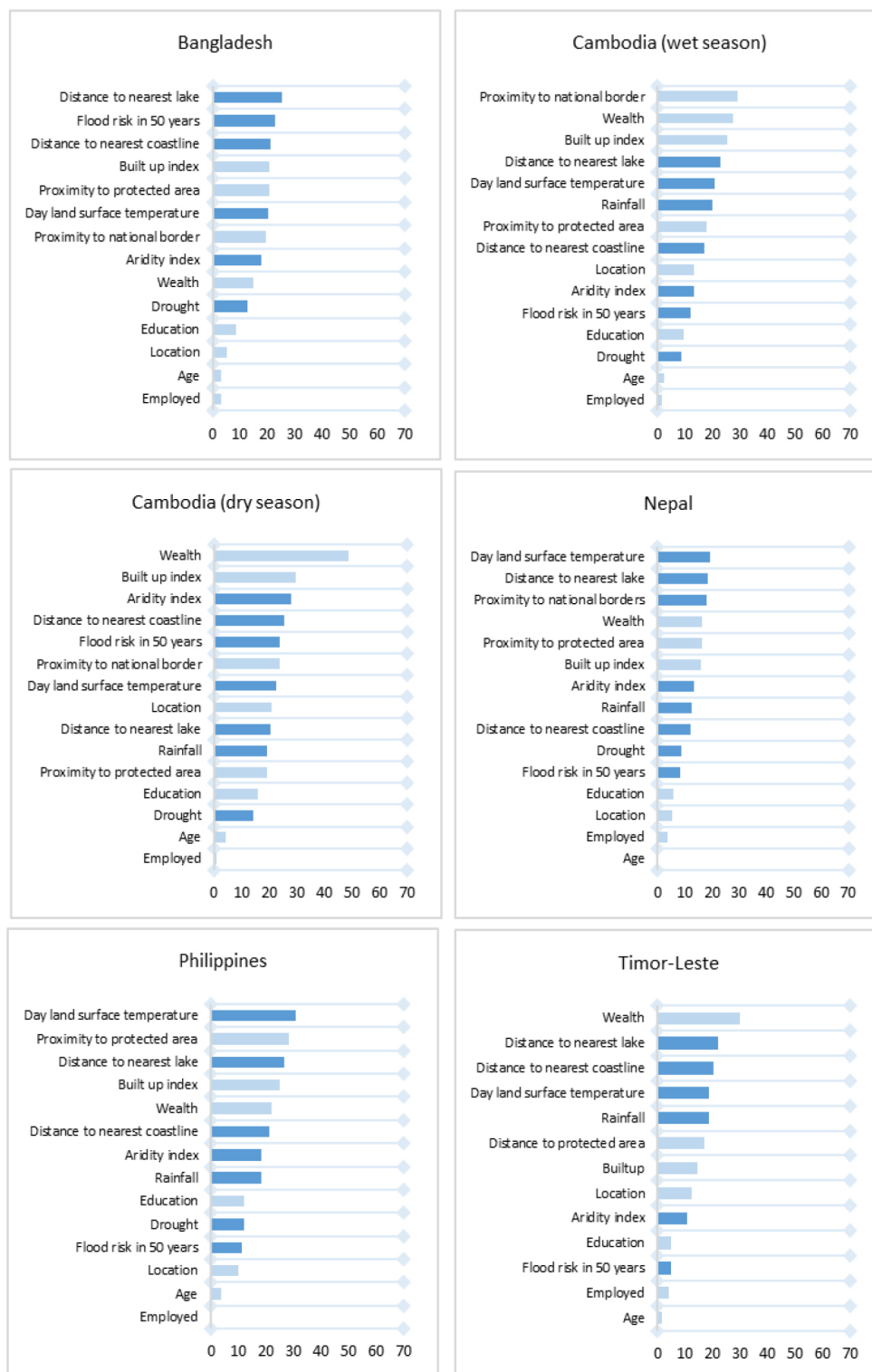


Fig. 6. Mean decrease in accuracy (in event of variable removal) for lack of access to basic drinking water sources, by country. Note: Environment-related variables are shown in darker shade of blue for ease of interpretation.

basic water sources at stake. Accelerated by rising temperatures, use of pesticides and pollution, cyanobacterial blooms in lakes have increasingly become a challenge as they can impair the safety of drinking water, fishing and irrigation, and pose difficulties for agriculture, fisheries and water treatment plants [53]. If cyanobacteria reach wetlands and groundwater reservoirs near lakes, they may further affect the quality of the water in nearby areas [54].

Day land surface temperature also plays an important role, as expected, given its connections with aridification and extreme weather. This is notorious across all countries, especially the Philippines, where the power of the model would decrease by 31 per cent if the variable was removed. In Bangladesh and Cambodia (dry season) (figure 26), “probability of flood in a 50-year period” is also important for classifying the risk of lacking access to basic water sources. Given the cyclical nature of severe flooding in Bangladesh, it is well documented that drinking water sources are contaminated during floods, as water sources are infiltrated by faecal matter and other debris [55]. To prevent water-borne disease, population groups that typically rely on open surface water for drinking, switch to bottled water and other improved sources during floods. A similar shift takes place in Cambodia, where 11 per cent of women shift away from drinking open surface water during rainy periods.

Access to clean cooking fuels: Of all gender related outcomes considered for this analysis, climate related variables are best at classifying the risk of lacking access to clean cooking fuels. Average day land surface temperature, for instance, is of relevance across all countries considered (in Bangladesh model accuracy would decrease by as much as 38% if the variable was removed, in Nepal and the Philippines by 22%). As noted earlier, different drivers may take place in different countries. Warmer temperatures are often accompanied by land degradation, erratic precipitation patterns, deforestation and reduced agricultural yield, which may lower the affordability of clean fuels (prompting people to shift to unclean). Fuel infrastructure and prices are also key determinants, as noted earlier in the case of natural gas in Bangladesh, but no related data was available to assess their classification power.

Aridity index, rainfall and distance to water bodies are also of relevance across countries. Previously noted relationships between these variables and agricultural yield, economic strains and affordability of fuels may explain these connections. In some countries, rain and proximity to water sources may also affect

the process of drying wood and shrubs, thus prompting shifts to other fuels. Random Forest is clear in showing that climate-related variables are important to classify the risk of using unclean fuels, but logistic regression demonstrated that other variables are also at play, such as fuel cost and infrastructure. Repeating this analysis with fuel cost and infrastructure data would be important to understand whether their classification power is higher or lower than that of climate-related variables.

Due to limitations associated with data availability, the localized nature of climate-related phenomena and the multiplicity of factors at play, further analysis is needed to ascertain the causes of the connections between climate and gender outcomes, and the differentiated effects across different population groups. As literature highlights that disadvantaged groups may suffer from climate change disproportionately [56], further analysis with multiple disaggregation by sex and age, or sex and ethnicity, for instance, could provide additional insights.

6.2. Utilizing big data to assess whether violence-related searches increase at times of crises, including natural hazards

Big data analysis revealed clear patterns in all four countries considered: spikes in VAWG related searches appear to cluster around or shortly after periods when weather events or increased mobility restrictions took place. This was true for both general violence searches and help-seeking searches. Although additional analysis would be necessary to assess the reasons behind these correlations, stress associated with each of the overlapping crises, brought on by insecurity, employment loss, reduced agricultural yield and the overall effect on people’s livelihoods could contribute to these sustained increases in violent behaviors.

In Kiribati (Fig. 7a), searches on VAWG, utilizing keywords such as “abuse sexually” and “domestic violence” saw a significant spike between December 2020–April 2021, when the country was affected by prolonged La Niña weather patterns and heavy rainfall.

In Samoa (Fig. 7b), sexual abuse searches grew by 161 per cent from December 2020 to April 2021. Searches related to psychological violence also grew by 52 per cent from January to March 2021. These spikes align with a period of flooding from December 2020 to January 2021. Physical violence searches did not see much noteworthy growth after the flooding.

In Solomon Islands (Fig. 7c), the rise in sexual violence searches, with words such as “sexual abuse in

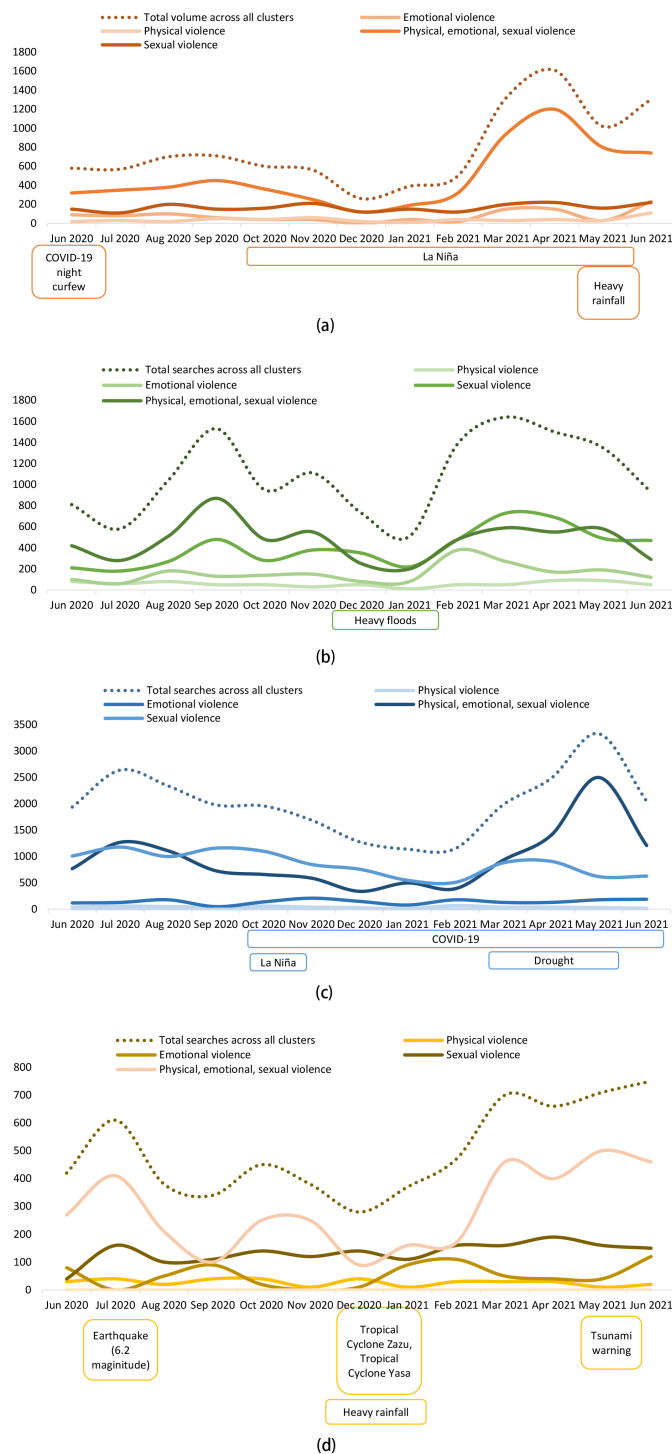


Fig. 7. a: Search volume on VAWG-related keywords, by type of violence cluster, Kiribati, June 2020–June 2021. Note: The graph represents the sum of monthly average search volume for each violence cluster. Because of different rates of internet penetration and use across countries, higher numbers of monthly searches do not necessarily imply that the prevalence of VAWG is higher in a particular country as compared to others. In other words, the findings should not be compared across countries. b: Search volume on VAWG-related keywords, by type of violence cluster, Samoa, June 2020–June 2021. c: Search volume on VAWG-related keywords, by type of violence cluster, Solomon Islands, June 2020–June 2021. d: Search volume on VAWG-related keywords, by type of violence cluster, Tonga, June 2020–June 2021.

marriage”, increased significantly between February 2021 to April 2021 (when COVID-19 overlapped with drought). The search volume for physical violence keywords such as “beating women”, “violent relationships” and “men who hit women” also increased from February 2021 onwards, which may be linked to economic stressors brought about by a public service recruitment freeze announced in early February 2021, compounded by the aforementioned drought [57,58]. Between March and May 2021 the average search volume for compounded physical, sexual and psychological violence also saw an increase. Dry weather conditions among western provinces, which came accompanied by severe water shortages and closure of some schools in May 2021, may have contributed to these trends.

In **Tonga** (Fig. 7d), overall violence searches increased by 80 per cent from June 2020 to June 2021. Among all types of violence considered, sexual violence searches saw the highest rise (275 per cent), with spikes observed in July 2020 (strong 6.2 magnitude earthquake), December 2020 (tropical cyclones Zazu and Yasa, alongwith heavy rainfall) and May 2021 (tsunami warning after earthquake in New Zealand). The increase included search terms like “rape”, “sexual assault” and “sexually abusive relationship”. Searches for emotional violence increased by 50 per cent from June 2020 to June 2021, with the highest spikes observed in January 2021 (heavy rainfall events) and June 2021. Searches for physical violence, with words such as “define domestic violence” and “what is domestic violence” stayed generally constant throughout the same period, but there were spikes in December 2020 (tropical cyclones Zazu and Yasa) and June 2021.

Beyond general violence searches, peaks in help-seeking searches also correlate with crisis and post-crisis periods (Fig. 8). The analysis of help-seeking search data is helpful to understand how women identify (and potentially access) support services for sexual, physical, and psychological abuse. Analysis shows that help-seeking behavior varied by country, but largely remained low. However, across all four countries, help-seeking searches saw an increase during or after crises.

For **Kiribati** (Fig. 8a), help-seeking and survivor searches remains low, compared to general information, however, help-seeking searches observed a spike post the COVID-19-night curfew in June 2020 and during October 2020, when the country experienced La Niña weather patterns. Help-seeking keywords such as “Kiribati police service” and “sexual abuse hotline” drove these spikes.

In **Samoa** (Fig. 8b), search volume on help-seeking grew from 10 to 70 from February 2021 to March 2021, after the country experienced heavy flooding.

For **Solomon Islands** (Fig. 8c), a rise in help-seeking searches accompanied the increase in VAWG searches observed during May and June 2021 (drought). Common search terms included “domestic violence counselling” and “sexual abuse hotline”. It is important to note that help-seeking searches in Solomon Islands primarily led to foreign websites – mainly based in Australia, the United States and the United Kingdom – rather than to websites of local NGOs, highlighting that victim may not be able to reach the help they need.

For **Tonga** (Fig. 8d), help-seeking behaviour appears to coincide with or occur shortly after peak events in the country between June 2020 to June 2021, including in July post the earthquake, in January 2021 post the tropical cyclones and heavy rainfall, as well as in June 2021, post the tsunami warning. Tongan Internet users continued seeking out local resources in consistent volumes, by using words such as “WCCC”, “Women and Children Crisis Centre” and “FFOV” (Families Free of Violence).

The increase in searches for local services (e.g. hotlines, police) in all four countries highlights that overlapping crises may lead to worsening violence or increased urgency among victims for seeking alternative living arrangements. However, on-line search results in the four countries show top results from foreign service providers (e.g. foreign websites). This is concerning as victims of violence may need local service provision and the search results indicate that they may not be finding the help they need.

6.3. Assessing the inclusion of gender-related references in climate policy commitments of UNFCCC Parties using web text data

Analysis of sentences containing key gender-related words in speeches of parties (e.g., countries and regional bodies) to COP24, COP25 and COP26, showcased that, of the 201 statements that were web scraped, only 32 statements had gender references (Fig. 9).

Among the Parties that have included gender-related references in their climate talks, Saint Lucia is the only country that consistently made gender and climate action references in its speeches across all three COPs. The ASEAN, Nepal, Sri Lanka and Ghana were relatively consistent with gender-related speeches in two of the three COP speeches (Fig. 10).

A total of 41 gender-related mentions were found in the 32 speeches identified, with most countries in-

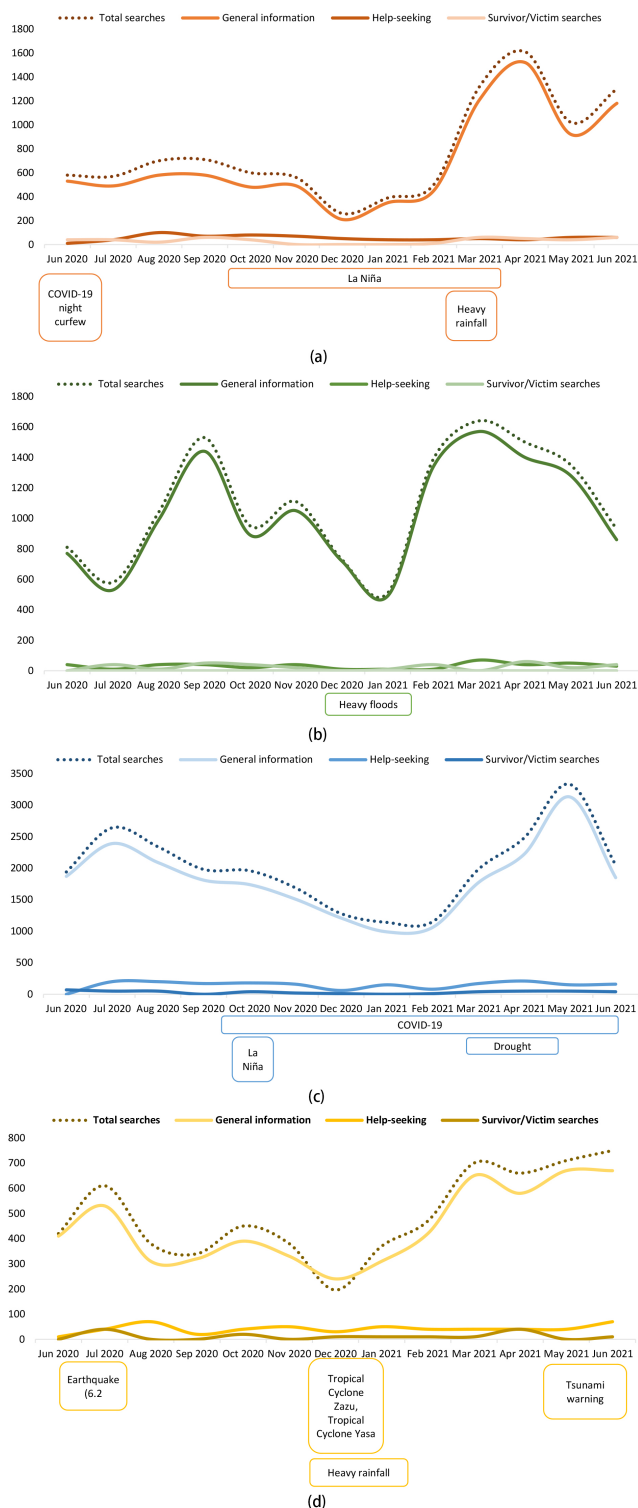


Fig. 8. a: Search volume on VAWG-related keywords, by type of search behaviour, Kiribati, June 2020–June 2021. b: Search volume on VAWG-related keywords, by type of search behaviour, Samoa, June 2020–June 2021. c: Search volume on VAWG-related keywords, by type of search behaviour, Solomon Islands, June 2020–June. d: Monthly volume for help-seeking search keywords, Tonga, June 2020–June.

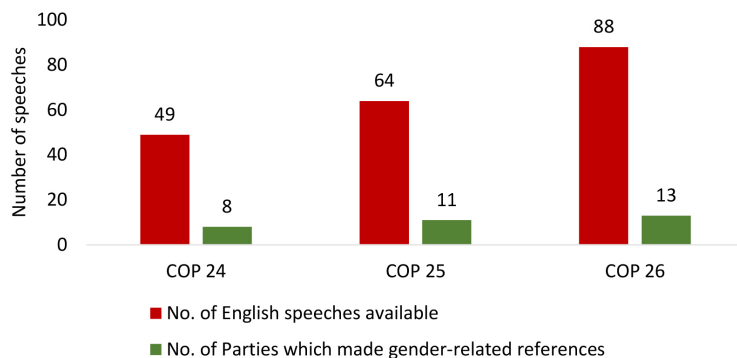


Fig. 9. Number of Member Parties which made gender-related references in COP speeches: COP24, COP25 and COP26.

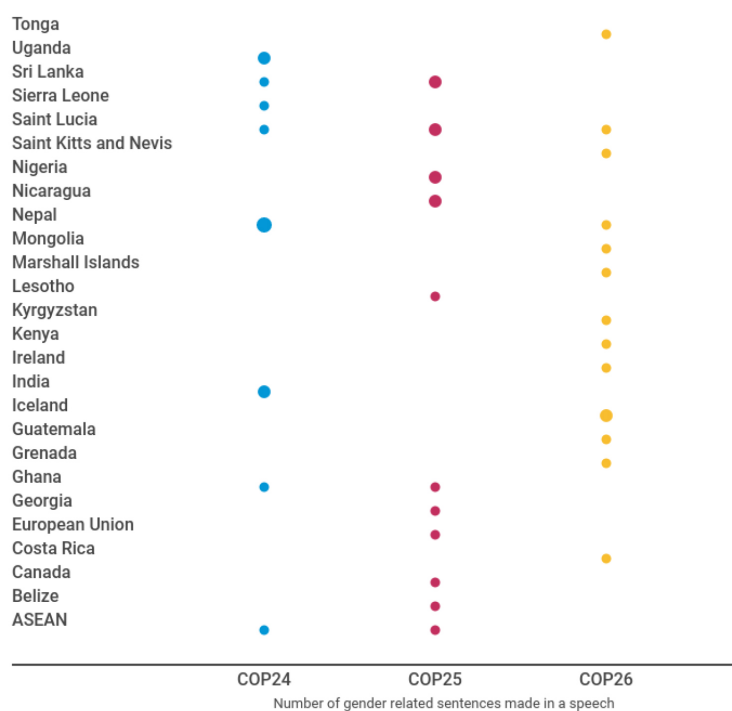


Fig. 10. Member Parties which made gender-related references in their COP speeches, by number of references: COP24, COP25 and COP26. Note: The size of the dots is proportional to the number of sentences with gender-related references made. There were at most three sentences with gender-related references made in a speech.

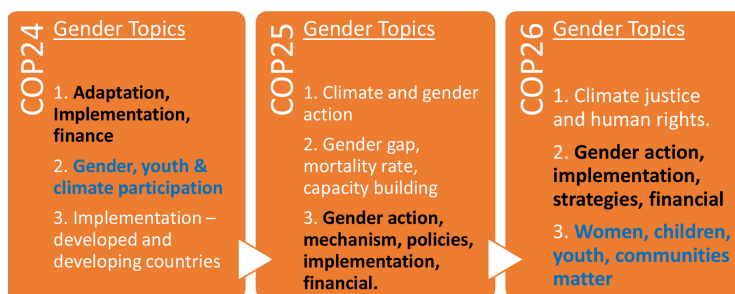


Fig. 11. Gender-related topics extracted in COP24, COP25 and COP26 using LDA technique.

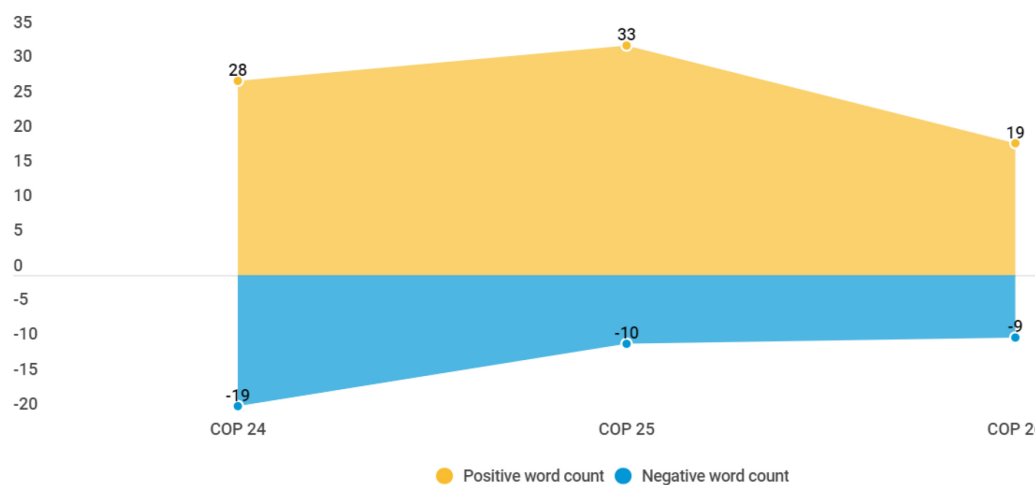


Fig. 12. Usage of positive and negative connotations in COP24, COP25 and COP26 using sentiment analysis.

cluding gender-related references only once. Countries making more than one mention in their speeches included: Uganda, Nepal and India in COP24; Sri Lanka, Saint Lucia, Nigeria, Nicaragua in COP25; and Iceland in COP26.

Among the sentences identified for gender-related references in the speeches, topic modelling was undertaken using LDA technique. This was useful to identify gender-related topics discussed in each of the speeches. Three gender-related topics were identified for each COP (Fig. 11). Topics pertaining to 'gender actions/policies/strategies and related financing' were discussed in all three COPs (shown in black), while the mentions of 'inclusive participation' appeared in COP24 and COP26 (shown in blue).

Sentiment analysis was also implemented to identify overall positive and negative sentiments arising from the gender-related sentences from the three COPs analysed. Negative connotations in gender-related mentions such as 'vulnerable', 'critical', 'adverse' and 'degradation' have decreased year-on-year from COP24 to COP26. Meanwhile, positive sentiments such as 'support', 'good' and 'sustainable,' increased from COP24 to COP25 but decreased in COP26 (Fig. 12). Insufficient evidence is available to ascertain whether different types of mentions (and frequency thereof) have differentiated impacts in outcome documents of these Conferences. This remains an area for further analysis.

7. Conclusion

The paper showcases several methods for filling information gaps on the gender and environment nexus

utilizing big data. Enhancing the availability and use of statistics on gender and the environment, including gender and climate change, is essential addressing the broad impacts of natural hazards on women and girls, building safe, inclusive and sustainable societies, and providing women with opportunities to manage natural resources and better contribute to environmental conservation. Analysis showcased in this paper indicates that there are indeed clear associations between climate-related variables, such as droughts, relative aridity, increases in temperatures, and proximity to natural resources, and gender related outcomes such as child marriage, adolescent births, intimate partner violence, access to basic water services and access to clean cooking fuels. Furthermore, big data also showcases that on-line searches for violence against women services and other related information increase at times of crises, including environmental crises. These findings shed light on the importance of addressing environmental issues from a gender angle, including by designing and implementing environmental policy that is gender sensitive. However, analysis of data from speeches in COP conferences showcased that mentions of gender issues are few and far between, highlighting the importance of continued advocacy to produce gender data on environmental matters, and its use to inform related policy.

The paper is useful for (1) establishing evidence on the gender-climate nexus utilizing big data; and (2) building evidence on the need to better mainstream gender on environmental policy. It provides solutions for filling gender-climate data gaps, acknowledges the important gender-environment connections at play, and reaffirms the need for furthering the work in this area. In line with the evidence showcased in this paper, gen-

der equality and environmental sustainability, including climate change, are intertwined issues and must be addressed in tandem.

The findings from this analysis are also useful to draw actionable steps that can be taken in the near future to jointly advance environmental sustainability and gender equality. Namely:

- It remains paramount to continue enhancing the availability and quality of data on the nexus between gender and the environment, including climate change. Only through the use of quality evidence can policies respond to the needs of women and girls of multiple backgrounds and in all their diversity.
- Where conventional data sources are insufficient, too expensive or untimely to provide gender-environment data, big data can be used to fill data gaps in a timely manner.
- National Statistics Offices and other actors within national statistical systems should work together, including in with development partners, to fill gender-environment data gaps, including by making use of big data for generating official statistics.
- To continue filling data gaps on the gender-environment nexus using big data, public-private partnerships can provide useful resources to find solutions to otherwise large information vacuums.
- Environmental policy makers in countries must base their decisions on gender-sensitive evidence. Only by using gender data for environmental policy making can sustainable environment policies respond to the needs of women and girls, and promote the health of ecosystems in a holistic manner.
- The use of data on the gender-climate nexus for advocacy and inter-governmental decisions is key to influencing global and regional commitments that advance the rights of women and girls while promoting environmental sustainability efficiently.
- Data on the gender-environment nexus can help shed light on existing financing gaps in this field. Allocating sufficient funds to gender-sensitive environmental policy is key to ensuring their implementation and far-reaching effects.

References

- [1] IPCC Sixth Assessment Report, 2022, chapter 18, Poverty, Livelihoods and Sustainable Development (Although, are mentioned as the Final Draft versions, they are approved for publications) <https://www.ipcc.ch/report/ar6/wg2/>.
- [2] Heidi MW, Maureen GR, Amber JF. Applying intersectionality to climate hazards: A theoretically informed study of wildfire in northern Saskatchewan. *Climate Policy*. 2021; 21(2): 171–185. doi: 10.1080/14693062.2020.1824892.
- [3] Corno L, Hildebrandt N, Voena A. Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica*. 2020; 88(3): 879–915. Available from: <https://bfi.uchicago.edu/working-paper/age-of-marriage-weather-shocks-and-the-direction-of-marriage-payments/>.
- [4] MESTI. The National Climate Change Policy. Accra: Ministry of Environment, Science, Technology and Innovation. 2014. Lawson ET. Negotiating stakeholder participation in the Ghana national climate change policy. *International Journal of Climate Change Strategies and Management*. 2016.
- [5] Nitya R, Elaine TL, Wapula NR, Divya S, Margaret NA. Gendered vulnerabilities to climate change: Insights from the semi-arid regions of Africa and Asia. *Climate and Development*. 2019; 11(1): 14–26. doi: 10.1080/17565529.2017.1372266.
- [6] <https://www.cbd.int/aichi-targets/target/14>.
- [7] Gender Composition Report, 2021. COP26 UNFCCC, https://unfccc.int/sites/default/files/resource/cp2021_04E.pdf.
- [8] Report of COP26, UN Climate Change Conference – October/November 2021, https://unfccc.int/sites/default/files/resource/cp2021_12_add2_adv.pdf.
- [9] Corno L, Hildebrandt N, Voena A. Age of marriage, weather shocks, and the direction of marriage payments. 2017. Available from: https://www.nber.org/system/files/working_papers/w23604/w23604.pdf.
- [10] Tsaneva M. The effect of weather variability on child marriage in Bangladesh. 2020. doi: 10.1002/jid.3507.
- [11] Peterman A, Potts A, O'Donnell M, Thompson K, Shah N, Oertelt-Prigione S, van Gelder N. *Pandemics and Violence Against Women and Children*. Center for Global Development. 2020. Available at www.cgdev.org/sites/default/files/pandemics-and-vawg-april2.pdf.
- [12] United Nations Economic and Social Commission for Asia and the Pacific (ESCAP). “The COVID-19 Pandemic and Violence against Women in Asia and the Pacific”. 2020. Available at www.unescap.org/sites/default/files/20201119_SDD_Policy_Paper_Covid-19.pdf.
- [13] UN Women. Measuring the shadow pandemic: Violence against women during COVID-19. 2021. Available from: <https://data.unwomen.org/sites/default/files/documents/Publications/Measuring-shadow-pandemic.pdf>.
- [14] Report of COP26, UN Climate Change Conference – October/November 2021, https://unfccc.int/sites/default/files/resource/cp2021_12_add2_adv.pdf.
- [15] <https://www4.unfccc.int/sites/submissionsstaging/Pages/Home.aspx>.
- [16] UN Women. Model questionnaire: Measuring the nexus between gender and environment. 2022. Available from: <https://data.unwomen.org/publications/model-questionnaire-measuring-nexus-between-gender-and-environment>.
- [17] <https://risk.preventionweb.net/capraviewer/>.
- [18] <http://www.soest.hawaii.edu/pwessel/gshhg/>.
- [19] Public data sources <https://www.emdat.be/> <https://www.ecmwf.int/> <https://www.eumetsat.int/> <https://reliefweb.int/report/kiribati/climate-change-water-security-and-women-study-wa-ter-boiling-south-tarawa-kiribati> https://www.president.gov.ki/images/CurfewOrder_eng.pdf.
- [20] COP conference for year 2020 was not conducted due to the COVID-19 pandemic.
- [21] COP 26 speeches and statements available at <https://unfccc.int/cop-26/speeches-and-statements#eq-3> COP 25 speeches and

- statements available at <https://unfccc.int/process-and-meetings/conferences/un-climate-change-conference-december-2019/speeches-and-statements-at-cop-25> COP 24 speeches and statements available at <https://unfccc.int/process-and-meetings/conferences/katowice-climate-change-conference-december-2018/events-and-schedules/high-level-segment/high-level-segment-statements>.
- [22] The statements provided publicly for COP20 in 2014 are limited with only 14 statements listed on the website.
- [23] Speeches from other stakeholders, such as NGOs and inter-governmental organizations playing observer roles, are not considered for this analysis.
- [24] The keywords have about 90 gender related origin words/phrases. Stemming out of the origin words are the other words/phrases originating from the origin words, that make up to 260 words in total. E.g. Gender – origin word, gender-led, gender parity are the related words.
- [25] PARIS21. Measuring references to statistics in national policy documents. Refer Annex 3. 2021. available at <https://paris21.org/uos>.
- [26] Although some of these indicators align with SDG indicators, others include adjustments on age groups and thus differ slightly from similar SDG indicators. These adjustments were made to ensure a large enough sample size.
- [27] The generalized linear model function in R (version 4.0.2) was used for this analysis.
- [28] Kirasich, et al. Random Forest vs Logistic Regression: Binary Classification for Heterogeneous Datasets. Available from: <https://scholar.smu.edu/cgi/viewcontent.cgi?article=1041&context=datasciencereview>.
- [29] Most DHS data available dates from 2014–2017. Geospatial data refers to a single year and typically differs from the survey year. Except for data on drought episodes, which spans over nearly 40 years.
- [30] ggi in the context of gender is ‘gender gap index’ while ‘ggi’ in the context of climate is ‘greening government initiative’. ‘mother’ in the context of climate was used as ‘mother earth’.
- [31] Stopwords are the commonly used English words such as a, the, was, there, etc.
- [32] They are also connected with the availability of basic water sources and clean cooking fuels, but the direction of these associations differs across countries due to additional factors such as infrastructure and prices.
- [33] In Nepal the association with basic drinking water sources appears inverse, likely due to differences in water infrastructure across regions.
- [34] While aridity index measures overall arid or humid conditions on average, drought episodes measure select periods of precipitation below average. Thus, both variables may affect gender outcomes differently.
- [35] Kulkarni S, Rao N. Gender and Drought in South Asia: Dominant Constructions and Alternate Propositions. In, *Drought and Integrated Water Resource Management in South Asia*. 2008. https://books.google.co.in/books?hl=en&lr=&id=OYHEbasIeROC&oi=fnd&pg=PA70&dq=drought+episodes+AND+asia+&ots=zaq-NGB1UJ&sig=9yLYYZe8k3dykCSnUOhqHu5BPY&redir_esc=y#v=onepage&q=drought%20episodes%20AND%20asia&f=false.
- [36] While the drought episode variable included in the analysis does not give any information on the duration, nature of onset, or coverage of drought, it represents the frequency of occurrence (on a scale of 1–10, with 1 meaning low frequency to 10 meaning high frequency).
- [37] FAO. <http://www.fao.org/land-water/water/drought/droughtanddag/en/>.
- [38] See: <https://www.eria.org/Chapter%207-Cambodia%27s%20Electricity%20Sector%20in%20the%20Context%20of%20Regional%20Electricity%20Market%20Integration.pdf>.
- [39] A high aridity index indicates either high rainfall, little evaporation, or both (thus, humid conditions). Relative aridity differs from droughts. The drought episodes variable only takes precipitation into consideration, while aridity index also measures evapotranspiration, and thus provides a better measure for the annual environment, as opposed to anomalous weather events.
- [40] Rees, et al. Associations between bride price stress and intimate partner violence amongst pregnant women in Timor-Leste. *Global Health*. 2017. Available from: <https://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0291-z#citeas>.
- [41] Bride price in Timor Leste involves an obligation on the husband to provide for all traditional events, including funerals and weddings, as well as ongoing support to the bride’s family as a form of marriage expense.
- [42] Petrone KC, Hughes JD, Van Niel TG, Silberstein RP. Streamflow decline in southwestern Australia, 1950–2008. *Geophysical Research Letters*. 2010; 37(11). Available from: https://www.researchgate.net/profile/Thomas-Van-Niel/publication/230577474_Streamflow_decline_in_southwestern_Australia_1950-2008/links/09e415059199aaf700000000/Streamflow-decline-in-southwestern-Australia-1950-2008.pdf.
- [43] Warming alters the cycling of nitrogen and carbon in soils, partly due to impacts on soil microbiota. See IPCC Report, Chapter 4, Available from: <https://www.ipcc.ch/srcc/clchapter-4/>.
- [44] IPCC. *Climate Change and Land: An IPCC Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. 2019. Available from: <https://www.ipcc.ch/site/assets/uploads/sites/4/2021/07/210714-IPCCJ7230-SRCCCL-Complete-BOOK-HRES.pdf>.
- [45] An empirical exploration of female child marriage determinants in Indonesia, available from: <https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-018-5313-0>.
- [46] World Bank and ICRW, Knowledge Brief: Basic profile of child marriage in Nepal, 2016, available from: <https://openknowledge.worldbank.org/bitstream/handle/10986/24546/BasicProfileOfChildMarriageInNepal.pdf?sequence=1&isAllowed=y>.
- [47] Amin S, Bajracharya A. Costs of marriage – Marriage transactions in the developing world. 2011. Available from: https://knowledgecommons.popcouncil.org/departments_sbsrpgy/833/.
- [48] IPCC, Chapter 4. Available from: <https://www.ipcc.ch/srcc/clchapter-4/>.
- [49] National Geographic Society. *How Climate Change Impacts Water Access*. Available from: <https://www.nationalgeographic.org/article/how-climate-change-impacts-water-access/>.
- [50] De Cian E, Lanzi E, Roson R. The impact of temperature change on energy demand: a dynamic panel analysis. 2007. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=984237.
- [51] Yakubu I, Salisu WJ. Determinants of adolescent pregnancy in sub-Saharan Africa: a systematic review. *Reproductive Health*. 2018. See: <https://reproductive-health-journal.biomedcentral.com/articles/10.1186/s12978-018-0460-4>.
- [52] McCarthy KJ, Mehta R, Haberland NA. Gender, power, and violence: A systematic review of measures and their association with male perpetration of IPV. *PLoS one*. 2018. Available from:

- <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0207091>.
- [53] National Geographic Society. Available from: <https://www.nationalgeographic.org/encyclopedia/lake/>.
- [54] Gkelis S, Vlamis A. Can cyanobacteria infect underground water sources? Indications from small scale monitoring of a natural drinking water source. 2017. Available from: <https://pagepressjournals.org/index.php/aiol/article/view/6280/6895>.
- [55] Islam, et al. Faecal contamination of drinking water sources of Dhaka city during the 2004 flood in Bangladesh and use of disinfectants for water treatment. *Journal of applied microbiology*. 2007. Available from: <https://pubmed.ncbi.nlm.nih.gov/17584454/>.
- [56] Sevoyan A, Hugo G. Vulnerability to climate change among disadvantaged groups: the role of social exclusion. *Applied Studies in Climate Adaptation*. 2014. Available from: https://books.google.com/books?hl=en&lr=&id=8IkeBQAAQBAJ&oi=fnd&pg=PA258&dq=impact+of+climate+change+AND+disadvantaged&ots=WxYeTuSP2f&sig=7-P_mxLheyGHxnk5Me3XG0FjzY.
- [57] Strains on male dominated jobs, such as those within the public service sector resulted in increased unemployment and financial burdens within households that may be related to the increase in physical violence towards women. See Waka Mere. 2018. Making Progress: Solomon Islands Businesses Advance Gender Equality. Available at www.ifc.org/wps/wcm/connect/768208a9-49e4-4bf6-a793-b6bbe3e55794/Making+Progress-Solomon+Island+businesses+advance+gender+equality.pdf?MOD=AJPERES&CVID=mp2v3C2.
- [58] World Health Organization. Violence Against Women in Solomon Islands: Translating research into policy and action on the social determinants of health. 2013. Available at: <https://www.who.int/publications-detail-redirect/9789290616467>.