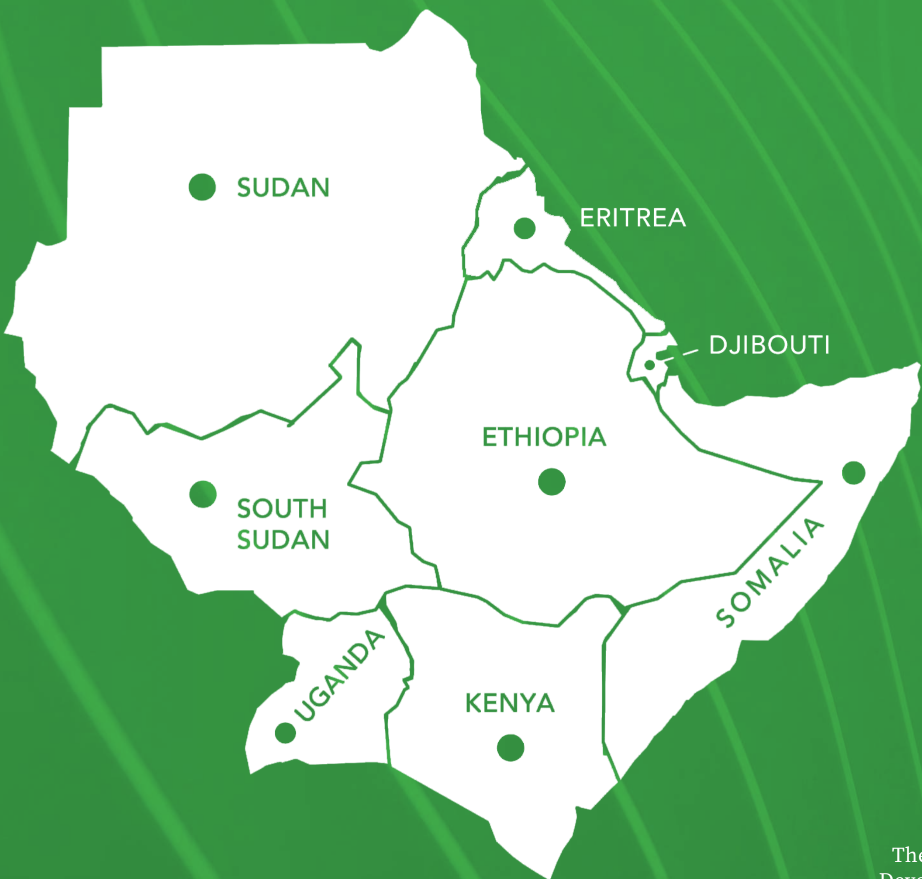
CLIMATE-CONFLICT NEXUS IN THE IGAD REGION:

A Study of CEWARN's Behavioral & ICPAC's Climate and Environmental Data As Predictors of Conflict Incidents, **2018 – 2022.**



PUBLISHED BY:

The Intergovernmental Authority on Development's Conflict Early Warning and Response Mechanism (CEWARN)

CLIMATE-CONFLICT NEXUS IN THE IGAD REGION:

A Study of CEWARN's Behavioral & ICPAC's Climate and Environmental Data As Predictors of Conflict Incidents,

2018 – 2022.

PUBLISHED BY:

The Intergovernmental Authority on Development's Conflict Early Warning and Response Mechanism (CEWARN)

NOVEMBER 2023

Irish Aid
Rialtas na hÉireann
Government of Ireland

This study was conducted and published with the financial support of IRISH AID

Table of Contents

LIST OF TABLES

CEWARN

Foreword

Climate change as a threat multiplier has taken center stage in the global discourse, especially on its link to peace and security. Since its founding in 2002, the Conflict Early Warning and Response Mechanism (CEWARN) for the Intergovernmental Authority on Development (IGAD) Member States, has been interfacing with the phenomena of climate related conflicts in the arid and semi-arid pastoral areas of the IGAD region. While executing its Mandate of collecting, analyzing and disseminating conflict early warning information, CEWARN accumulated vast amounts of data on pastoral conflicts, which overtime, it has been able to analyze as historical data. Some of the deductions from the analysis indicated that there are correlations between conflict escalations with seasonal patterns which are driven by climate variabilities.

This led to CEWARN's first ever study in 2007, published in Meier, Bond, and Bond's "Environmental Influences on Pastoral Conflict in the Horn of Africa," Political Geography which, though indicative, was not conclusive enough in drawing correlations between conflicts and climate variability with statistical confidence. CEWARN conducted another study in 2020 seeking to answer the question, 'Do CEWARN's behavioral SEP (Social, Economic, and Political) risk ratings warn on conflict?' Following the affirmative results of the study, CEWARN then undertook another study in 2021 which sought to establish a correlation between the behavioral indicators with a set of climate and environmental parameters. This latter study yielded a strong statistical correlation between the behavioral indicators and environmental parameters. The study was limited in scope both geograpically (covering CEWARN's historical Areas of Reporting-AORs) and thematically (covering pastoral conflicts and environmental parameter of vegetation index).

This 2023 study, therefore, replicates the 2021 study with the broader objective of widening the scope to include other non-pastoral areas in the region with the addition of a new climate variable, rainfall estimates, for the analysis. The findings of the study provide the much sought for scientific basis of developing early warning predictive models based on the easily observable and measurable weather and environmental parameters. It is therefore my sincere hope that the findings of the study, as presented in this report, will contribute significantly to the growing wealth of knowledge and contemporary debates on the climate-conflict nexus.

In conclusion, I would like to appreciate the Government of Ireland for the financial support that it extended, through its Embassy in Addis Ababa, Ethiopia, to IGAD's CEWARN that has been instrumental in accomplishing this important study.

Camlus Omogo,
Director, CEWARN.

Acknowledgements

This report is a product of the ongoing contractual collaboration between IGAD's CEWARN and Virtual Research Associates Inc. (VRA), who have been engaged to provide CEWARN with the requisite conflict analytics technical support and backstopping since CEWARN's inception. This study has been made possible by the financial support from the Government of Ireland through its Embassy in Addis Ababa, Ethiopia.

The report was written by VRA and CEWARN and validated by technical experts from IGAD's specialized offices and CEWARN's National Research Institution. The core of the analysis team comprised of Dr. Doug Bond, Mr. Sean Yeo and Dr. Fujimoto Hiroyoshi (VRA) and Andrew Malinga (CEWARN GIS Conflict Analyst), who extracted the climate and environmental data from its spatial format to numerical format for the statistical analysis. The validation workshop, overseen by Mr. Camlus Omogo, CEWARN-Director, consisted of Dr. Sunday Okello (CEWARN Conflict Analyst), Deogratious Sebuwufu (CEWARN Statistician), Dr. Dominic Kathiya (ICPALD Pastoral Livelihoods Expert) Dr. Geoffrey Sabiiti (ICPAC Climate Expert) and Joseph Muhumuza (NRI Governance Expert).

Abbreviations and Glossary of Terms

AOR	Area of Reporting, denoting an area where an individual or team of field monitors are assigned to conduct conflict monitoring activities.
CEWARN	Conflict Early Warning and Response Mechanism of the IGAD Member States, a dedicated unit within IGAD employing data and analysis for conflict early warning purposes.
IGAD	Intergovernmental Authority on Development.
ICPAC	IGAD Climate Prediction and Appplication Center.
NDVI	Normalized Difference Vegetation Index, an environmental metric used to measure vegetation greenness, aiding in assessing vegetation density and detecting changes in plant health.
RFE	Rainfall Estimate, a satellite-based precipitation parameter and product.
VRA	Virtual Research Associates, a social science analytic and consulting firm based in the United States of America.

Executive Summary

This 2023 Climate-Conflict Nexus study is the fourth in a series of studies that assess the behavioral and environmental variables as predictors of violent conflict occurance within the IGAD Region. Whereas the previous studies were conducted on historical data (2003-2015), this study uses data from 2018-2022. This study also added an open-source behavioral dataset based on media reports to supplement the field data collected by CEWARN. A logit regression statistical analysis was applied on the data, which revealed a significant correlation of the parameters. GIS techniques were applied to extract the climate and environmental data from its spatial format to numerical format for the statistical analysis.

The study is based on environmental datasets of vegetation and rainfall estimates and open-source media report data for the IGAD region from 2018-2022.

The study revealed significant relationships between healthier vegetation and increased rainfall, both contributing to a reduced likelihood of physical conflict of assault, fight, and violence. Four key findings of the study include;

- It revealed that rainfall and vegetation index were highly correlated meaning, with rainfall data available, it is possible to estimate the vegetation quality index and therefore likely impact on conflict.
- Vegetation health had a significant influence on conflict outcomes in the following month hence establishing a 1-month lag between measurement of the parameters and potential impacts. A one-month early warning or conflict outlook is therefore plausible.
- The study established that an AOR (Area of Reporting) experiencing a 0.2 NDVI index increase (vegetation) corresponds to a 12% decrease in the probability of physical conflict in the following month.
- Similarly, when an AOR receives an additional inch of rainfall, it results in an 8% reduction in the likelihood of conflict in the following month.
- Based on the findings, a predictive model was formulated that can be used to perform a conflict likelihood in the subsequent month. The model was simulated using September data to predict conflict plausibility in October.

The design and procedures used in this study are presented in the main body of the report. Three Appendices summarize the previous (2007, 2020, and 2021) studies, assess the 2018-2022 field data quality, and outline the results from the field data. This study is part of an IGAD-CEWARN service contract supported by the Government of Ireland.

1. Introduction: Purpose and Scope

This study examined environmental factors from 2018 to 2022 and CEWARN's behavioral variables as potential indicators of conflict escalation within the IGAD region. It focused on environmental parameters, utilizing datasets related to vegetation, rainfall estimates, and open-source media reports specific to the IGAD region.

The study expands the geographical scope from the previous studies to encompass the entirety of the IGAD region. The previous studies relied on field-reported data from 2003 through 2015, with a focus on pastoral conflict within the Karamoja and Somali clusters that crossed borders within a part of the IGAD member states.

The open-source media report dataset used in this study comprises approximately 13,000 discrete Areas of Reporting (AORs). It is important to note that not every AOR within this dataset was included in this study. Instead, the CEWARN Unit and VRA jointly strategically selected key AORs that met certain criteria to optimize efficiency and accuracy in our analysis. As a result, the study's findings were able to provide more relevant insights into the climate-conflict nexus across this diverse and extensive region.

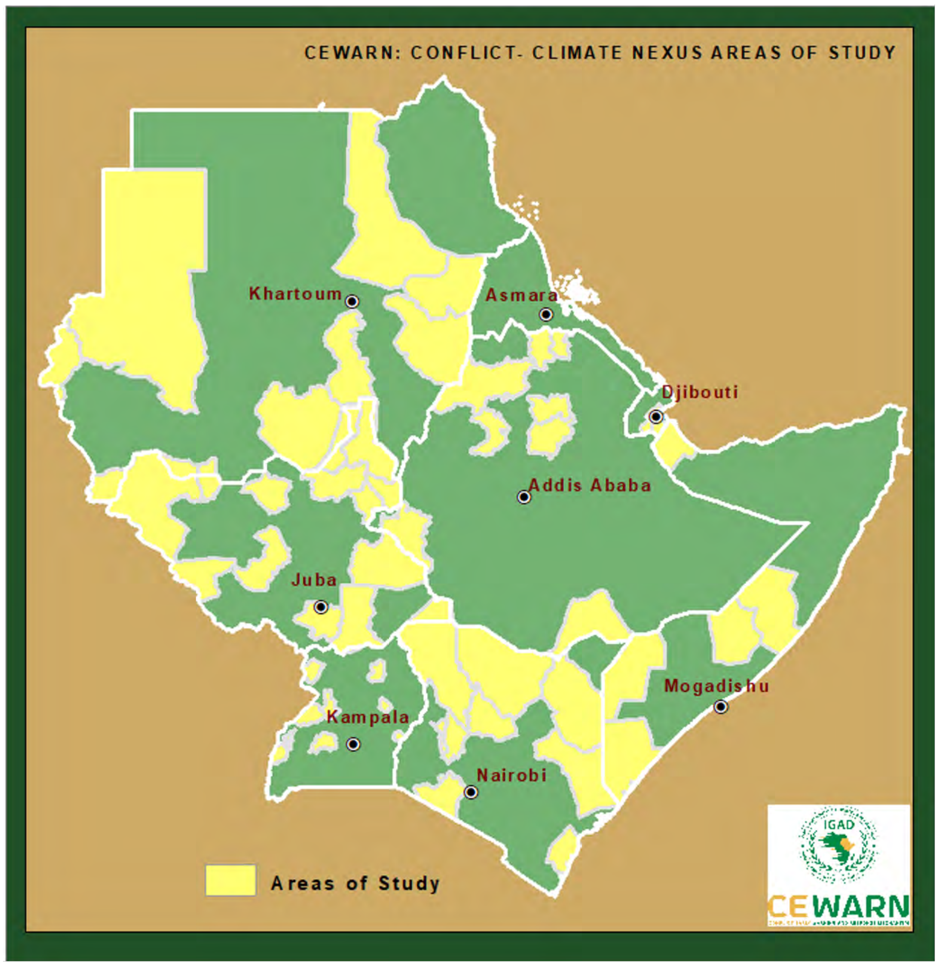
We conducted a logit regression model analysis on the presence-absence data for conflict outcomes related to Assault, Force-use, and Violence occurrences. However, for this study, we did not utilize a negative binomial regression model, as was done in the 2021 study, to evaluate the magnitude or intensity of conflict. The media data did not offer suitable measurements for assessing conflict intensity.

AOR demarcation process: This study harnessed almost 4 million events reported in the media to represent the behavioral risk predictor, complemented by the two environmental measures, to determine the extent to which they anticipated the occurrence of subsequent physical conflict.

Unlike the historical field data, which spanned from 2003 to 2015 and featured 40 distinct Areas of Responsibility (AORs) specifically designated for incident and situational reporting, the media report database presented a considerably larger dataset, comprising more than ten thousand locations without pre-defined AORs.

Given this unique challenge, we established specific conditions to streamline our study and mitigate the expansive array of AORs. These conditions helped us focus on key AORs that aligned with our research objectives as follows:

- The successful outcomes of the 2021 study were attributed to the field data's focused nature on pastoral conflict and suburban, cross-border regions in the IGAD region reliant on agriculture and pastoralism.
- Notably, the correlation between the vegetation index and conflict occurrence and intensity was strongly linked to vegetation quality. A 0.1-point drop on the vegetation scale (ranging from 0 to 1, where 1 represents optimal vegetation) corresponds to a higher likelihood of conflict occurrences by x%, resulting in x more human deaths and livestock losses as an indicator of conflict intensity.
- Based on this rationale, we identified specific locations that 1) serve as crucial agricultural hubs within the region, 2) act as suburban zones supporting nearby cities and urban centers, and 3) are considered hotspots exhibiting a certain level of conflict occurrences. Guided by the CEWARN Unit's over 20 years knowledge of geography and conflict dynamics in the IGAD region, we jointly identified 56 preliminary AORs that aligned with the conditions outlined above.



Map of Areas of Reporting (study).

Media Events Data Preparation: Initially, we identified "hotspots" with an average of at least 10 conflict events per month (hotspot identification) over the past five years (2018-2022); this yielded n=378 hotspots. From this, CEWARN selected 53 AORs based on the conditions of 1) dependency on agriculture and 2) vulnerability to climate change in terms of productivity in suburban areas.

CEWARN, in collaboration with VRA, reviewed the selected AORs, examining event occurrences for independent variables along with their individual weights and counts of Assault, Violence, and Force Use event categories. This process was conducted to configure independent and dependent variables for the analysis.

Compilation of NDVI and RFE: The CEWARN GIS Conflict Analysis team conducted a comprehensive data collection effort, sourcing data from organizations such as ICPAC, FEWS NET, and USGS. Subsequently, these datasets underwent review in collaboration with ICPAC and CEWARN. To facilitate the analysis, the entire IGAD region was subdivided into a grid system, with each grid cell representing a specific climate data point pertaining to rainfall and vegetation. The CEWARN and ICPAC teams mapped each of these cells to their respective AOR areas as delineated on the geographical map. This mapping process formed a critical foundation for the climate-conflict analysis.

NDVI and RFE monthly average per AOR: In the 2021 study, we employed field data comprising 40 distinct AORs paired with climate data, including rainfall estimates and vegetation indices. Through our collaboration with CEWARN, we have expanded this dataset to include 53 AORs linked with climate data.

The first iteration in processing the media and climate data involved calculating monthly averages for climate variables, resulting in a substantial dataset exceeding 50 gigabytes and comprising 126 million rows. To streamline this extensive database, we aggregated each observational value, calculating the monthly averages for each AOR. This reduction resulted in 2562 unique AOR-Month combinations for respective NDVI and RFE, facilitating manageable data analysis.

Combining event and climate databases: The collaborative efforts of CEWARN and VRA culminated in the confirmation of the final list of AORs. CEWARN generated, with support of ICPAC, the climate data, and VRA compiled and integrated this environmental dataset with the behavioral media event data. Still, we encountered some challenges when attempting to map AORs to the event database due to the lack of a standardized AOR mapping of locations presented in the media with the AOR names associated with the environmental data.

To establish a manageable framework for AORs, we opted to designate admin2s as the basic unit, as admin1s proved too extensive for averaging climate data effectively. For instance, Rift Valley, classified as an admin1 region in Kenya, has a sprawling surface area of 70,466 square miles, akin in size to countries like Senegal or Syria. Inevitably, certain climate AORs could not be seamlessly associated with the AORs in the event database, owing to disparities in data availability and challenges in GIS mapping. Consequently, we resorted to a compromise between climate AORs and event AORs.

In the next section, we include a brief discussion of three cases to illustrate the AOR-related challenges: the following cases outline our ad-hoc solutions to align and integrate climate data with the AORs.

Case of Uganda: One to Many: Uganda's admin1 regions exhibit considerably smaller sizes compared to those in other IGAD member states. To address this variance, we amalgamated some AORs and linked the admin1 regions to a conceptual larger AOR, encompassing multiple locations. The locations listed below have been mapped to the climate data of "West," serving as a consolidated AOR for the analysis.

ClimateData AOR	EventData AOR
West	HOIMA
	KABAROLE
	KASESE
	MASINDI

In summary, we employed climate data for "West" and subsequently aligned it with event data for the discrete four AORs of Hoima, Kabarole, Kasese and Masindi. It was our assumption that these locations share similar climate data.

Case of Ethiopia: Many to One: In contrast, the Ethiopian event data AORs either did not exist due to the location name discrepancies or exhibited an insufficient event volume.

ClimateData AOR	EventData AOR
South Wollo	Amhara
West Gojam	
North Wollo	
North Gonder	

Consequently, we utilized climate data for South Wollo, West Gojam, North Wollo, and North Gonder, which we mapped to the larger admin1 region of Amhara. We then compared each climate data set against the aggregated event data from Amhara.

Case of South Sudan: Typos and Name Discrepancies: The situation in Sudan diverged slightly from the scenarios encountered in Uganda and Ethiopia. Here, we confronted more structural issues stemming from outdated event locations or typographical errors in the event location database. For instance, these three distinct locations in South Sudan seemed to refer to the same location, Upper Nile.

These findings suggest that CEWARN and VRA must collaborate to review and map all event locations and correct outdated location names in the media event data.

EventData AOR	Modified AOR
Upper Nile, Wilayat A`ali an Nil, South Sudan	Upper Nile, South Sudan
Upper Nile State, Wilayat A`ali an Nil, South Sudan	
Uppernile, Wilayat A`ali an Nil, South Sudan	

Missing data and data integration: Certain AORs lacked associated climate data due to the sheer volume of climate data involved, which resulted in the omission of these AORs during the export process. Consequently, AORs without corresponding climate data were excluded from the analysis.

AOR dummy and exclusion of Somalia: This study employed AOR dummy variables, which are binary variables characterized by values of 0 or 1, signifying the presence or absence of certain factors. These binary dummies have been integrated into our regression model, serving as a means to encapsulate observed or unobserved variables within our analysis. Additionally, it is important to note that Somalia was deliberately omitted from the final equation. The exclusion was necessitated by the sporadic and unreliable nature of the data pertaining to Somalia, which had a substantial adverse impact on the integrity of the regression outcomes.

2. Methodology

This section presents the media reports data and behavioral variables, the regression model, and lag procedures used in this study. Note the field data are presented in the Appendices.

Event Category	Category Mean Weight	
Accede	6.7	Not used
Cooperate	5.4	Not used
Assure	4.8	Not used
Appeal	3.5	Not used
Confront	(3.2)	Precursor
Object	(4.6)	Precursor
Coerce	(6.2)	Precursor
Threaten	(6.4)	Precursor

Data attributes: the media data: The GDELT Project (https://www.gdeltproject.org/) monitors the world's broadcast, print, and web news from nearly every corner of every country in over 100 languages and identifies the people, locations, organizations, themes, sources, emotions, counts, quotes, images and events driving our global society every second of every day, creating a free open platform for computing on the entire world.

This study used all reported events (just over 4 million) that took place within the IGAD region from 2018-2022 in version 2 of the database. These event records are unique by virtue of their synthetic nature; all similar events are clustered around their centroid.

The behavioral variable used was a risk score calculated by the weighted event mean of conflict event types that did not include the use of physical force or violence. The following categories illustrate the range of events included with the media dataset, with eight of them representing events that do not use physical force. A ninth category of events, force, was used as the target or dependent variable in this study. We used only the four conflict-short-of-physical force categories in this study.

2.1 Regression Model

In this study, logit models were exclusively employed to evaluate the effectiveness of explanatory predictors in anticipating the occurrence (presence or absence) of physical conflict and the extent of associations between precursor behavioral and environmental variables and a subsequent incidence of conflict.

The study employs different time lags between the environmental variables and their associated outcomes related to physical conflict occurrences, including 1, 2, and 3-month lags. This variation in time lags allows us to discern the timing of conflict incidence.

The standard model adopted for analysis utilizes a one-month lag because of the statistical significance of the coefficients. This implies that when there is a change in environmental variables, its impact will manifest one month later.

2.2 Model Specification

The model specification consisted of the following components:

- Dependent Variable (DV): A binary variable indicating an occurrence of force or physical violence - variable name: "_frcbin"
- Independent Variables (IVs):
    - Behavioral Variables: risk scores calculated for four categories of conflict events from the media data (explained in the section "Data attributes: the media data"), namely confront, object, coerce, and threaten precursors - variable name: "_conbin", "_objbin", "_coebin", "_thrbin", respectively.
    - Environmental Variables: two key variables were considered;
    - NDVI – Normalized Difference Vegetation Index - variable name: "ndvi100"
    - RFE – Rainfall Estimates - variable name: "rfeinch"
    - Country dummy variables: binary variable for each country controlling for country-specific factors - variable name: "countryid".

The estimation results are presented in Table 1, with each of the six models featuring different lag structures for environmental variables denoted as Model 0 (referred to as M0) through M5. Specifically, M0 through M2 pertain to NDVI, while M3 through M5 are associated with RFE. For all these models, one-month lagged behavioral variables were applied, following the outcomes of several experimental runs. Table 1 outlines odds ratios and p-values. The findings from M0 reveal that the one-month lagged NDVI (variable name "ndvi100," labeled "L1.") is statistically significant at the 10% level. The odds ratio, measured at 0.9404, is less than one, implying that a higher NDVI in the current month corresponds to a reduced probability of conflict incidence in the following month.

To elaborate, the likelihood of conflict decreases by a factor of 0.94 when the NDVI from the previous month increases by 0.1. Based on the results of M1 and M2, NDVI with a two-month lag and a combination of one-month and two-month lags are not statistically significant, suggesting that M0 is the most reliable model.

Table 1: Logit Estimation Results

Dependent variable: binary variable indicating an occurrence of force or physical violence ("1" for occurrence, "0" for no occurrence)

Variable	M0	M1	M2	M3	M4	M5
_conbin						
L1.	2,0256	1,9414	1,9341	2,0695	2,0625	1,9805
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
_objbin						
L1.	1,8100	1,8382	1,8568	1,6037	1,6312	1,6363
	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
_coebin						
L1.	1,3223	1,2186	1,2181	1,4049	1,3838	1,3004
	0.0252	0.1308	0.1316	0.0032	0.0046	0.0286
_thrbin						
L1.	1,6463	1,6845	1,6709	1,7727	1,7997	1,8240
	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000
ndvi100						
L1.	0,9404		0,8816			
	0.0881		0.1039			
L2.		0,9744	1,0856			
		0.4810	0.2801			
countryid						
2	8,3565	7,4114	8,0495	8,2684	7,2525	6,9640
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	6,4089	6,2502	6,5828	6,7293	5,8757	6,0026
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4	5,9379	5,6975	6,0842	7,0716	6,2464	6,5028
	0.0002	0.0007	0.0005	0.0000	0.0000	0.0000
6	5,3453	4,8470	5,3230	5,2863	4,8319	4,6880
	0.0000	0.0003	0.0001	0.0001	0.0001	0.0003
7	7,9536	7,5760	8,3622	7,8108	7,0287	7,2787
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
rfeinch						
--.				0,9443		
				0.3100		
L1.					0,9321	
					0.2095	
L2.						0,9614
						0.4989
_cons	0,0482	0,0447	0,0455	0,0421	0,0479	0,0452
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	1870	1700	1700	2106	2145	1950

RED BOX: NDVI results (on a scale of 0-1)

GREEN BOX: RFE results (in inches)

Highlights: P-value: A value indicating the statistical significance of the independent variable in the model specification (if the variable has an explanatory ability in the model or not). The variable is statistically significant at a 10% level if the p-value is less than 0.10. Each column in Table 1 represents a unique model specification, which is labeled as "M0" (abbreviation of "Model 0") through "M5" (Model 5).

The label "L1" indicates that the variable is included in the model with a one-month lag (previous month), and "L2" means a two-month lag (two months ago). For example, in the estimation result of M1, "ndvi100" is included with a one-month lag, testing if vegetation index NDVI affects the probability of conflict in the next month.

- The result of M3 through M5 indicates that the current RFE (variable name "rfeinch," with a label "--."), one-month and two-month lagged RFE (with a label "L1." and "L2.") are all statistically insignificant at 10% level.

- The odds ratio is less than one in these models, indicating that a higher RFE in the current month implies a lower probability of incidence of conflict in the following month. Based on M4 (the odds ratio 0.9321) with the lowest p-value among these three models, the probability of conflict is 0.93 times higher (meaning decrease) when RFE in the last month increases by 1 inch.

Note that M0 for NDVI and M4 for RFE are used to calculate the predicted values in the subsequent analyses.

2.3 Results: Marginal Effect Review

In this section of the report, the marginal effects of NDVI and RFE are analyzed and calculated individually. The aim is to determine the extent to which changes in NDVI and rainfall quantity are correlated with conflict occurrences as indicated by the data in a measurable format.

Vegetation (NDVI) Marginal Effect

To grasp the predictive significance of vegetation (NDVI) in conflict, we provide Table 2, which displays the odds ratios and the impact of an increase in the NDVI index on pastoral conflict outcomes in the following month. Based on these findings, a 0.1 increase in NDVI reduces the likelihood of a physical conflict (assault, fight, and violence) outcome by a factor of 0.9403509, resulting in a 6% reduction in the probability of occurrence.

Table 2: Logit estimation result using NDVI as an environmental variable

Logistic regression	Number of obs	=	1870
	LR chi2(10)	=	286.74
	Prob > chi2	=	0.0000
Log likelihood = -1013.7878	Pseudo R2	=	0.1239

_frcbin	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_conbin						
L1.	2.025627	.2591311	5.52	0.000	1.576407	2.602858
_objbin						
L1.	1.81001	.2348205	4.57	0.000	1.403624	2.334055
_coebin						
L1.	1.322341	.1650317	2.24	0.025	1.035408	1.688791
_thrbin						
L1.	1.646283	.2079267	3.95	0.000	1.285278	2.108685
ndvi100						
L1.	.9403509	.033914	-1.71	0.088	.8761756	1.009227
countryid						
2	8.356458	3.389736	5.23	0.000	3.77345	18.50572
3	6.408917	2.529658	4.71	0.000	2.956703	13.8919
4	5.937938	2.87884	3.67	0.000	2.29591	15.35736
6	5.345291	2.189968	4.09	0.000	2.394604	11.93188
7	7.953584	3.38014	4.88	0.000	3.457933	18.29402
_cons	.0482445	.0176548	-8.28	0.000	.0235481	.0988416

To elaborate further, considering the standard deviation (SD) of monthly NDVI as 0.2 and assuming each NDVI variance is 0.2, a monthly NDVI improvement or deterioration by 0.2 corresponds to a 12% increase or decrease in conflict odds.

Within this sample dataset (excluding Somalia), approximately 33% of the 2562 data points were associated with conflict. Consequently, an NDVI increase of 0.1 reduces the probability of conflict from 33% to 31% (0.33 × 0.9403509 = 0.3103158). In practical terms, a 0.94 times reduction in the probability (equivalent to a 6% reduction) translates to a decrease of 2 percentage points. When applied to the monthly NDVI SD of 0.2, each NDVI change will result in the probability of conflict occurrence from 33% to 29% when favorable vegetation is observed or from 33% to 37% in the case of poor vegetation conditions.

2.4 Rainfall Estimate (RFE) Marginal Effect

Table 3: Logit estimation result using RFE as an environmental variable

Logistic regression	Number of obs	=	2145
	LR chi2(10)	=	317.17
	Prob > chi2	=	0.0000
Log likelihood = -1177.3186	Pseudo R2	=	0.1187

_frcbin	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_conbin						
L1.	2.062524	.2439942	6.12	0.000	1.635697	2.600729
_objbin						
L1.	1.631176	.1950584	4.09	0.000	1.290366	2.062
_coebin						
L1.	1.383784	.1587106	2.83	0.005	1.105201	1.732588
_thrbin						
L1.	1.799739	.2130597	4.96	0.000	1.427056	2.269749
rfeinch						
L1.	.9321303	.0522014	-1.25	0.209	.8352326	1.040269
countryid						
2	7.252467	2.84914	5.04	0.000	3.358075	15.66323
3	5.87575	2.295986	4.53	0.000	2.731813	12.63792
4	6.246446	2.457951	4.66	0.000	2.888608	13.50757
6	4.831883	1.952193	3.90	0.000	2.188826	10.66649
7	7.028748	2.910556	4.71	0.000	3.121773	15.8254
_cons	.0478693	.0174995	-8.31	0.000	.0233821	.0980009

Table 3 displays the odds ratios and the impact of an increase in the RFE on pastoral conflict outcomes in the following month.

One inch (25.4mm) more rainfall increases the chance of a physical conflict outcome (assault, fight, and violence) by 0.9321303 times. This means that the probability of occurrence (%) is reduced by 7%.

In this sample (excluding Somalia), about 33% of the 2562 data had "conflict," so if you use this number, increasing the RFE by 1 inch will decrease the probability of conflict from 33% to 31%. (0.33×0.9321303=0.32908814). The standard deviation of RFE is 29mm, which is 1.14 inches - An increase in RFE by 29mm (since the SD of RFE is about 29) reduces the probability of conflict from 33% to 30% (0.33 x (0.9321303)^1.14 = 0.304591159).

2.5 NDVI-RFE Correlation and POSSIBLE Causality

In this section, we operated under the assumption of a causal link from RFE to NDVI, aiming to validate the projected values.

As displayed in Table 4, NDVI and RFE, along with their respective lags, exhibit a high degree of correlation, posing challenges in establishing a causal relationship between these variables. This complexity becomes evident, especially when attempting to provide an explanation without referencing prior research on the weather-related dynamics that underlie the connection between rainfall and vegetation.

Table 4: Correlation coefficients among NDVI, RFE, and their lags

```
. * NDVI and RFE
. correl ndvi rfe l.ndvi l.rfe
(obs=1836)
```

	ndvi	rfe	L. ndvi	L. rfe
ndvi	1.0000			
rfe	0.9303	1.0000		
ndvi L1.	0.9234	0.8617	1.0000	
rfe L1.	0.8755	0.9029	0.9238	1.0000

In a separate analysis, we treated NDVI as the dependent variable (DV) and considered RFE (in inches) and its lags (1 month and 2 months) as independent variables (IVs). We conducted five separate regression analyses using the Ordinary Least Squares (OLS) method, labeled as M6 to M10.

2.6 OLS Regression

DV: NDVI / IV: RFE with various lags

- M6: no lag
- M7: 1 month lag
- M8: 2 months lag
- M9: 1 month lag & 2 months lag
- M10: no lag & 1 month lag & 2 months lag

Table 5: OLS estimation results | Dependent variable: NDVI index

Variable	M6	M7	M8	M9	M10
rfeinch					
--.	161,3091			132,4442	126,8559
	0.0000			0.0000	0.0000
L1.		153,6414		33,9863	43,4119
		0.0000		0.0000	0.0000
L2.			135,7010		-3,4150
			0.0000		0.3795
_cons	33,4851	59,3614	120,6136	23,0818	23,7081
	0.0000	0.0000	0.0000	0.0000	0.0000
N	2006	1870	1700	1836	1666

Each column in Table 5 represents a unique model specification, which is labeled as “M6” (abbreviation of “Model 6”) through “M10” (Model 10). The P-value is displayed under each of the coefficients. The P-value indicates the statistical significance of the independent variable in the model specification. The variable is statistically significant at a 1% level if the p-value is less than 0.01.

The label “L1” indicates that the variable is included in the model with a one-month lag (previous month), and “L2” means a two-month lag (two months ago). For example, in the estimation result of M7, “rfeinch” is included with a one-month lag, testing if rainfall estimation affects the vegetation index NDVI in the next month.

- Based on the M6 result, an increase in RFE by 1 inch leads to an increase of NDVI by 0.161 in that month.
- Based on the M7 result, an increase in RFE by 1 inch leads to an increase of NDVI by 0.154 in the following month.

For instance, by referencing the outcomes of M7, we can generate descriptions like;

- A 1-inch increase in rainfall (RFE) two months ago corresponds to a 0.154 increase in the Normalized Difference Vegetation Index (NDVI) one month ago.
- Consequently, the likelihood of conflict occurring in the current month decreases from 33% to 30% (calculated as 0.33*(0.9403509)^1.54=0.300179102).

Nevertheless, with the findings from M9, it is possible to construct a predictive model for estimating conflict occurrences in the current month using rainfall estimates (RFE) from two months ago and one month ago. The most straightforward approach is to utilize M6, which is the regression of the current NDVI on the current RFE. This allows us to make the following observations:

- If RFE one month ago increases by 1 inch, NDVI one month ago increases by 0.161.
- By applying the NDVI marginal effect model (one month lag)above, we can compute the conflict probability for this month, resulting in a reduction from 33% to 30% (0.33 * (0.9403509)^1.61 = 0.298889563).

Keep in mind that the RFE coefficient lacks statistical significance. While RFE and NDVI exhibit a strong correlation, the discrepant missing observations might explain the divergence in the logit results, with one showing significance and the other insignificance. In any case, it is important to note that the results exhibit some degree of instability.

3. Comparison: The 2023 and The 2021 Study

The 2021 study relied on field data collected by the CEWARN Unit from 2003 to 2015, whereas in this 2023 study, we used media data from 2018 through 2022 – but also see Appendix III for a tentative analysis of 2018-2022 field data. The media data offers broad and more dense coverage compared to the focused field data. While the use of media data provides wider geographical coverage, it does come with certain limitations.

Field data is better suited for capturing the nuanced dynamics of conflict, offering a deeper understanding of the situation on the ground. This assertion is less compelling when quality assurance is not maintained during the field monitoring and field data collection – Appendix II discusses the limitations in the 2018-2022 field data that argued for the decision to use media data for this study. In general, open-source media data is more economical and covers a broader range of AORs across the IGAD member states as long as the data source remains accessible for public use. However, the media data are constrained in that the parameters are not customizable as they are when field data is collected with well-trained local monitors.

3.1 Use of Rainfall Estimates for the 2023 Study

Regarding the use of rainfall estimates in the 2023 study, this choice was motivated by CEWARN that rainfall data is readily monitorable, relatively easy to measure, and straightforward to interpret. This provided a variable for inclusion in the study, enhancing the overall feasibility and effectiveness of the analysis.

3.2 Comparing Marginal Effects, 2021 versus 2023

The 2021 study emphasized the impact of either an increase or decrease in the NDVI on various conflict outcomes. 0.2 point or 20% increase index (the standard deviation of the NDVI index is 0.2) was associated with a significant decrease, for example, a 34.6% decrease in the number of human deaths, a 28.6% reduction in net livestock losses, and a 21.4% decrease in the probability of incident occurrences during the subsequent month. Notably, the most substantial reductions in outcomes were observed in women and children deaths (49.4%) and armed conflict (41.2%) following a 0.2 point or 20% increase in the NDVI.

In contrast, the 2023 study, utilizing event data from 2018 to 2022, expanded the analysis beyond NDVI. It found that a 0.2 NDVI index increase (indicating improved vegetation) resulted in a 12% decrease in the probability of physical conflict occurrence in the following month.

Additionally, an extra inch of rainfall in an AOR led to an 8% reduction in the likelihood of physical conflict. This study incorporates both NDVI and rainfall variables, shedding light on their combined influence on conflict dynamics.

The 2023 study, utilizing media data, does not exhibit the same level of statistical significance as the 2021 study; however, it has unveiled a correlation between conflict and climate using a more comprehensive and readily available media data set. Furthermore, the incorporation of rainfall as a variable has provided us with more accessible and quantifiable parameters for early warning and conflict prevention. This expanded understanding of climate-conflict nexus enhances CEWARN's ability to predict and mitigate potential conflicts in the future.

4. Lessons learned

Several valuable lessons were gained during the research process and from the results. One of the lessons learned highlights the significance of CEWARN field data as a robust source for capturing climate-induced conflicts.

5. Value of CEWARN Field Data

One of the study's recommendations emphasizes the importance of incorporating field data instead of relying solely on global media events data. This suggestion is bolstered by a comparison of confidence levels between the 2021 and 2023 Climate-Conflict Nexus studies. Although the findings of both reports align, the 2021 study, which utilized field data collected by CEWARN field monitors, demonstrated higher statistical confidence levels compared to the 2023 report, which relied mostly on open-source events data. Incorporating field data can further enhance the precision of the findings and increase confidence in the research outcomes.

Allocating resources to support CEWARN's field monitoring efforts with improved quality assurance and regular training for the field monitors, while initially costly to implement, can yield significant value for conflict early warning. Field data and media data offer distinct advantages and disadvantages; hence, a comprehensive approach should involve the integration of both data sources.

5.1 Measuring Conflict Intensity

A conflict intensity measurement also underscores the importance of field data in future research and CEWARN operations. One limitation of this study was the absence of conflict intensity as a variable. Incorporating continuous field monitoring, which includes both normalized values and narratives, has the potential to substantially enhance the precision of climate-associated forecasting.

One key advantage of field data is the ability to directly assess the intensity of conflicts. While it is possible to estimate the count of human deaths from various sources through AI data processing, it requires a meticulous examination of the input data to ensure its reliability, like duplication issues where the same event is reported twice. Additionally, this data source should be readily accessible near real-time for CEWARN to ensure the accuracy and effectiveness of the forecasting process.

Having normalized multiple sources of field data is less of a concern going forward than in the past, thanks to significant advancements in data ingestion technology. These advancements have streamlined the process of incorporating diverse data sources, making it more feasible for CEWARN to leverage AI and other technologies for enhanced early warning and conflict analysis.

6. Suggested Follow-up Research

The follow-up research recommendations encompass formulating actionable policies, establishing comprehensive Areas of Reporting (AORs), defining standard operating procedures (SOPs) for data collection and quality assurance, and developing operational products based on the predictive model.

More contextualized studies need to be undertaken so as to capture other determinant factors and influencers of conflict that may be AOR specific. This is a bigger undertaking that requires adequate resources and funding and could take the form of a project.

7. Formulating Policy Recommendations

7.1 Ensuring Vegetation Health for Conflict Prevention

CEWARN analysts should explore the policy implications of the findings, addressing areas such as the necessity of effective rangeland management, the preservation of water catchments, the acceleration of the transhumance protocol, and the enhancement of data collection efforts.

7.2 Establishment of comprehensive AORs

The suggestion was made for CEWARN to enhance the methodology for determining Areas of Reporting (AOR), aligning them with both CEWARN's designated regions (hotspots) and livelihood typologies. The selection criteria for AORs should be scientifically justified for greater accuracy and effectiveness.

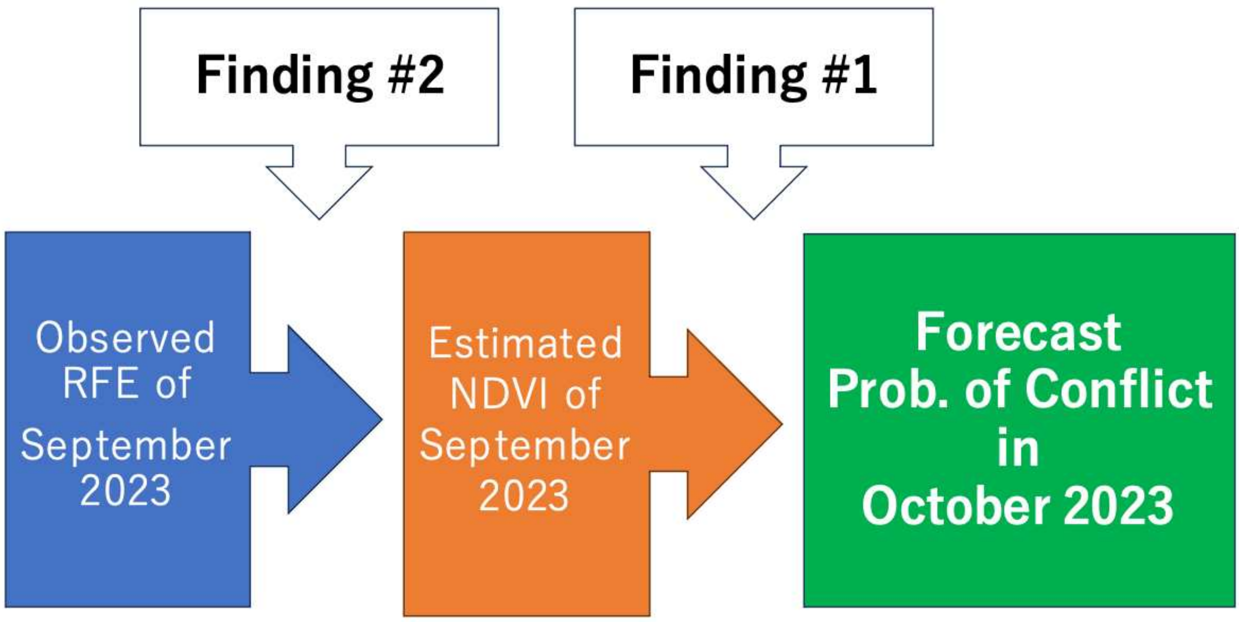
8. Development of an Early Warning Tool: From a Model to a Product

Based on the findings, a predictive model was formulated that can be used to perform a conflict likelihood in the subsequent month. Robust vegetation health significantly impacted conflict outcomes in the subsequent month, highlighting a clear one-month lag between environmental parameters and potential impacts on conflict incidence. Consequently, a one-month early warning or conflict outlook from environmental measurement is both practical and feasible.

Finding #1: A 0.2 increase in NDVI corresponds to a 12% reduction in t he likelihood of physical conflict in the subsequent month.

Finding #2: NDVI can be accurately predicted once we obtain RFE data for the month, given the strong correlation between NDVI and RFE.

CEWARN should prioritize the swift development of an operational tool based on the research findings above. This tool would serve to predict potential conflicts based on environmental data, catering to policymakers and conflict analysts alike.



9. Robust Utilization of CEWARN's Field Data

To utilize climate data as a consistent independent variable for CEWARN's early warning operations, the establishment of data integration and ingestion pipelines is imperative. This study, while highlighting the correlation between environmental indicators and conflict outcomes, does not itself establish a framework for continuous monitoring. However, it does serve to demonstrate the potential benefits of such integration.

Moving forward, it would be useful for CEWARN to establish Standard Operating Procedures (SOPs) that include the following tasks:

- Timely Data Collection: develop a system to collect environmental data in a timely and consistent manner to ensure up-to-date information.

- Quality Assurance Procedures: enforced in real-time while the field reports are submitted as this is the optimal time to identify and correct errors, to maintain consistent collection over time, and most importantly, enhances the capacity to prevent them.

- Regular Evaluation and Training: both the field data collection and the analysis procedures need to be regularly evaluated to build a sustainable and motivated pool of human resources and procedures.

- Integration with Conflict Data: Combining the environmental data with conflict data and especially aligning them, both with clearly demarcated AORs, is essential to enhance the interpretation of the results. This integration does not need to be fully automated, but it does need to be specified and consistently followed.

This approach needs to be robust as well as extensible and flexible, including periodical calibration to account for structural changes and evolving mandates. For instance, areas considered vulnerable yesterday might no longer be susceptible due to changing dynamics, such as suburban areas shifting their focus away from pastoralism.

Effective data collection, ingestion, and integration with conflict data, whether sourced from the media or field reports, are fundamental components of achieving this task. It is essential for CEWARN to establish a rigorous yet adaptive strategy to enhance CEWARN's early warning capabilities.

Two additional tasks would be helpful to support the monitoring of risk and developing prevention and response options in a timely manner. The first is to prioritize the effort to manage costs; for example, the ongoing identification of vulnerable areas due to climate and/or conflict conditions is an important first step to conflict early warning and response. A second high-priority task is to demarcate AORs that are small enough to provide detailed analyses and aligned with both behavioral and environmental data to facilitate timely action.

In sum, this study underscores the potential benefits of integrating environmental indicators with conflict data for early warning operations. To fully operationalize these benefits, CEWARN needs to take proactive steps to establish comprehensive, rigorous and adaptable SOPs to facilitate the real-time analysis and the formulation of options to prevent or mitigate violent conflict situations before they escalate.

Appendix I: Background to the IGAD Region Climate-Conflict Nexus Study

This background appendix was developed as part of the 2023 IGAD-CEWARN Service Contract (IE 01-2023) supported by the government of Ireland. This background outlines three previous studies conducted by VRA in 2007, 2020, and 2021 to provide context to the current (2023) study. These three previous studies set the foundation for the current (2023) study. All four of the studies are framed by this overarching question: Do CEWARN risk scores warn on conflict? The risk scores used in the 2021 and 2023 studies include CEWARN's behavioral indicators from its field data as well as ICPAC environmental measures provided by IGAD. The bottom-line conclusion from all four studies is yes, CEWARN risk scores and environmental measures do signal subsequent outcomes and these results are statistically significant.

More specifically, these studies assessed the extent to which CEWARN risk scores calculated from the field Situation Reports (SitReps) in the IGAD region anticipate or warn on the escalation of conflict as observed in subsequent field Incident or Event reports. Each of the assessments also included an examination of the data integrity and quality assurance procedures used in the collection of data, as well as their post-collection cleaning and transformations. This Appendix (I) provides the background of the studies while Appendix II discusses the data quality issues involved with the post-2017 data; this re-assessment was necessary given that the collection protocol (2019 version) was significantly modified from the earlier versions. Appendix III presents the results of the current field data analyses and compares them to the 2020 and 2021 results.

The 2007 Study: This first study of CEWARN's field data was published in Meier, Bond, and Bond's "Environmental Influences on Pastoral Conflict in the Horn of Africa," Political Geography, 2007. The study assessed CEWARN's Situation Report (SitRep) indicators that were categorized into ten indicator groups, including three environmental indicators. The seven behavioral indicator groups from the 2007 study are outlined below.

> **SitRep Indicator Groupings from the 2007 Political Geography Study**
> *Peace Indicator Scores, derived from CEWARN SitReps, 2003-2007 Alliances, Exchanges, Mitigation & Initiatives. Conflict Indicator Scores, derived from CEWARN SitReps, 2003-2007 Aggravators, Pressure & Provocation Environmental Measures, calculated from NOAA & LEWS data, 2003-2007 Rain Fall, Forage & Vegetation*

This first study was conducted as a pilot study of CEWARN's initial field data and although the results were encouraging, they were not considered definitive because of the small data sample that was assessed very early in CEWARN's field data collection effort.

The 2020 Study: VRA conducted a second study for IGAD's CEWARN Unit in May 2020 that sought to answer the question, do CEWARN behavioral SEP (Social, Economic, and Political) risk ratings warn on conflict? The study was conducted using monthly aggregated field reports for each of CEWARN's 40 Areas of Reporting (AORs) within the IGAD region.

The explanatory variables or predictors for this study were Social, economic & political risk ratings (SEP), low to high, based on CEWARN SEP Situation Reports (SitReps) from 2003-2015

The outcome or target variables included the Occurrence & intensity of human deaths & livestock losses, based on CEWARN Incident Reports (IncReps) with a one-month lag.

This 2020 study concluded yes, CEWARN risk scores do signal subsequent outcomes – incidents, human deaths, and livestock losses – and these results were statistically significant.

The 2021 Study: VRA then conducted a follow-up study for CEWARN in September 2021 that sought to determine the added value of environmental measures (beyond the CEWARN SEP behavioral risk ratings) to anticipate conflict.

The explanatory variables for this second study included environmental and behavioral measures as follows:

Environmental measures – forage, vegetation & rainfall – provided by IGAD Climate Prediction and Application Centre for the years 2013-2015
Social, economic & political risk ratings (SEP), low to high, based on CEWARN SEP Situation Reports (SitReps) from 2003-2015 (same as the May 2020 study)

The outcome or target variables included the: Occurrence & intensity of human deaths & livestock losses, based on CEWARN IncReps with a one-month lag (same as the May 2020 study) The top predictors of the incidence of conflict outcomes were vegetation, social & economic risk scores, the experience of the reporters, and indicators focused on pre-raid blessings, peace initiatives, disputed boundaries, livestock disease, natural disasters, male migration and migration policies.

The top predictors of the intensity of conflict outcomes were vegetation, social & economic risk scores, the experience of the reporters, and indicators focused on natural disasters, peace initiatives, pre-raid blessings, migrant laborers, separation of groups, antagonistic alliances, and disputed boundaries.

This 2021 study validated the 2007 and 2020 conflict early warning results. It also confirmed that environmental influences (vegetation and forage, in particular) anticipate pastoral conflict outcomes, including deaths and livestock losses, but more generally, all types of pastoral conflict incidents monitored by CEWARN in its first ten years of operations.

Appendix II: Quality Assessment of CEWARN's Field Reports Data, 2018-2022.

VRA, in September 2023, conducted a follow-up study to the previous two studies outlined above. There was a two-year hiatus in CEWARN's field data collection from mid-2015 to mid-2017 while a major upgrade to the CEWARN Reporter and the collection protocol was designed, engineered, deployed, and tested. Training on the revised (2019 version) of the CEWARN Reporter began in late 2017 and continued into 2019.

This 2023 study seeks to update the 2021 study to anticipate conflict, by updating the data set to include contemporary data (2018-2022) and expanding the AORs from 40 to about 50, refined to be consistent with IPCPAC's data.

To accomplish these objectives, VRA examined the field data collected by CEWARN. The data was found to be sparse and erratic in nature, especially from 2017 into 2019. The field reporters were still being oriented or trained from 2017 and into 2019. Based on the results of our data assessment discussed in this Appendix, the CEWARN Unit approved the decision to use media instead of field data for the current study. This decision to use media instead of field data was driven by the fact that the field data was collected from just two countries for barely half of the mandated time of 2018-2022. However, given that CEWARN's data management tool, the CEWARN Reporter, underwent a major revision and pilot testing in 2017, we still needed to determine the extent to which the earlier risk ratings were comparable and consistent with the new risk scores.

Thus, we first cleaned the field data to assess its usability for this study and then we analyzed a cleaned sample of the field data to illuminate the implications of the recent modifications made to the risk calculations. The procedures used and results from the assessment of the field data quality are presented in Appendix II and the comparison between the 2020 and 2023 logit model regression results are presented in Appendix III.

Again, the decision to focus on media rather than field data for this study was discussed with and agreed to by the CEWARN Unit as we worked together throughout this field data assessment as a prelude to the Climate-Conflict Nexus study.

The Raw Field Data: From January 2018 through December 2022, **2,987 SitReps were** submitted to CEWARN. In addition, **9,631 event reports** were submitted for the period February 2018 through January 2023. This date shift applied to the event reports enables us to examine events that occurred one month after each SitRep was submitted. Another way of characterizing these lagged events is the outcomes reported one month after the SitReps submitted within each AOR.

SitRep Assessment: From the 2,987 SitReps submitted from 2018 through 2022, we had to drop 65 SitReps (2.2%), including 43 due to invalid AOR or ID data that we could not correct and 22 that had no risk ratings. We were able to correct 274 (9.2%) invalid reports, including 273 invalid or missing AORs and one invalid ID among remaining SitReps. We also dropped 18 (0.6%) single-submission SitReps because the reporters submitted only one SitRep throughout the entire five years It is doubtful that these reporters were available to monitor and record events in the month following their only report. Dropping these single reports was deemed important because the first report, and especially when it was the only report, submitted is usefully characterized as an orientation or training report.

Next, we sought a conservative test of completed SitReps for our statistical test so we dropped 174 SitReps (5.9%) that were missing between one and four of the five sector ratings. The five sectors are economy, security, governance, environment, and social. Four of the five sector risk ratings were missing from 145 SitReps. Another 15 SitReps had incomplete data across three sectors, 12 had missing ratings across two sectors, and 4 SitReps had one missing sector rating. Adding the 273 corrected SitReps to the 259 dropped SitReps brings the total invalid or incomplete SitReps to 532, representing a bad data rate of 17.8%. This rate of invalid and incomplete data reflects the erratic reporting and limited coverage of the field events collected from 2018 through 2022 as well as poor quality assurance procedures. Virtually all of the invalid and missing data issues could have been corrected and/or prempted had they been identified and addressed in real-time.

Thus, the erratic reporting and limited coverage of field data from 2018 through 2022 exacerbated by inadequate quality assurance procedures compelled us to draw on media data as an alternative source for this study. Media data offer a consistent, dense, and broad coverage of the IGAD region. The use of media data also facilitated the analysis of the environmental and behavioral indicators across a wider set of AORs across the IGAD region to determine their relative influence as predictors of peace and conflict. We worked with the CEWARN Unit to align the AORs across the IGAD region with the ICPAC-produced environmental measures and media-based behavior indicators. After the cleaning described above, the total number of valid SitReps across 49 AORs submitted from 2018 through 2022 was 2,728.

Event Reports: From the 9,631 event reports submitted from February 2018 through January 2023, we had to drop 17 reports (0.2%), including three used for training and fourteen that had invalid AORs that we could not correct. Dropping these 17 records yielded 9614 event records. We were able to correct 179 event reports (1.9%) with invalid AORs, including three corrected on the AOR and/or country and two with invalid dates. The remaining 9,614 event reports occurred across 65 AORs and were deemed usable for analysis.

Linked Data Records: The SitReps were then joined with their corresponding event reports lagged by one month. The result was an input data set of 965 AOR-Month ratings of risk across 49 AORs. Among the 965 AOR-Months in which at least one SitRep was reported, 605 (62.7%) AOR-Months had reported conflict events the following month, while 360 AOR-Months (37.3%) had no event reports the following month. We interpret the absence of reported events to represent that no significant conflict occurred; of course, this interpretation assumes the field reporters were trained and continuously monitoring their respective AORs. To determine the extent to which this assumption is realistic, the consistency of reporting and quality of reports must be considered as discussed above.

Two countries, Kenya and Uganda, submitted the vast majority (85.1%) of the SitReps. In addition, several spikes in reporting were evident as shown in the table below. We suspect that these spikes reflect increased reporting during training workshops; a review of the schedules and participants in CEWARN's training workshops could confirm this suspicion.

Finally, note the AORs for the CEWARN field reports are not aligned with the IGAD climate data. Therefore, we were not able to use ICPAC's contemporary environmental measures prepared by CEWARN for this analysis. Still, it is important to get a sense of the performance of the updated (2019 Version) of the CEWARN Reporter to determine the extent to which its multi-dimensional approach to risk assessment compares to the earlier versions that were operational from 2003 to 2015. The field data analysis that is presented in this Appendix essentially replicates the May 2020 study mentioned in the background section above. This analysis, like the May 2020 study, is limited to behavioral measures, but with five sectors instead of the earlier study's three domains. The main difference between the 2019 version of the field data and the earlier versions lies in the calculation of the risk score. We seek to illuminate the implications of this difference with the analysis presented in Appendix III. Still, caution is in order, given the data quality issues outlined above.

CEWARN Field Report Submissions of Valid and Corrected SitReps, 2018-2022

Country	Year	SitRep Count	Country	Year	SitRep Count
Ethiopia	2021	39	Somalia	2022	6
Ethiopia	2022	29	South Sudan	2019	4
Kenya	2019	38	South Sudan	2020	2
Kenya	2020	78	South Sudan	2021	1
Kenya	2021	150	Uganda	2018	63
Kenya	2022	79	Uganda	2019	119
Somalia	2018	5	Uganda	2020	114
Somalia	2019	1	Uganda	2021	91
Somalia	2020	5	Uganda	2022	89
Somalia	2021	52			

Appendix III: Field Reports Data Logit Regression Results

The explanatory variables for this September 2023 field data analysis include five risk ratings across the economy, security, governance, environment, and social sectors. These variables frame the SitReps risk calculations. The target variable for this preliminary study is singular: the presence or absence of conflict events reported in the month following the corresponding SitReps. A logit regression was used for this analysis. An N of 965 SitRep AOR-months within the study period of 2018 through 2022 were used in this analysis.

The logit estimation result is presented below, followed by the odds ratios. The social risk score is statistically significant at p<=0.5 level and the coefficient is positive, which is consistent with the 2020 and 2021 studies. However, the economic risk score is not significant this time. Instead, the security risk score is significant at the p<=0.00 level. This model uses the risk scores alone without any additional contextual variables and its explained variance (R2) is relatively low at just 4.9.%.

```
Logistic regression                             Number of obs     =        965
                                                LR chi2(5)        =      62.09
                                                Prob > chi2       =     0.0000
Log likelihood = -606.39847                     Pseudo R2         =     0.0487
```

conflictdummy	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
economyriskscore	.0047221	.0087371	0.54	0.589	-.0124024	.0218465
securityriskscore	.0379262	.0088309	4.29	0.000	.020618	.0552343
governanceriskscore	.0064214	.008556	0.75	0.453	-.0103482	.0231909
environmentriskscore	-.0017717	.008092	-0.22	0.827	-.0176317	.0140883
socialriskscore	.0242904	.0111442	2.18	0.029	.0024483	.0461326
_cons	-1.651601	.3270348	-5.05	0.000	-2.292578	-1.010625

```
Logistic regression                             Number of obs     =        965
                                                LR chi2(5)        =      62.09
                                                Prob > chi2       =     0.0000
Log likelihood = -606.39847                     Pseudo R2         =     0.0487
```

conflictdummy	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
economyriskscore	1.004733	.0087785	0.54	0.589	.9876742	1.022087
securityriskscore	1.038655	.0091722	4.29	0.000	1.020832	1.056788
governanceriskscore	1.006442	.0086111	0.75	0.453	.9897052	1.023462
environmentriskscore	.9982298	.0080777	-0.22	0.827	.9825228	1.014188
socialriskscore	1.024588	.0114182	2.18	0.029	1.002451	1.047213
_cons	.1917427	.0627065	-5.05	0.000	.1010058	.3639915

We now compare these results with the 2020 results to examine the extent to which the two sets of results are consistent. This comparison is necessary because the risk calculations were significantly modified in the current version of the CEWARN Reporter. Also, the current version uses a slightly different framing of indicators, a five-sector framing versus the previous three-domain framing. In both studies, a logic model was used to test the predictors' association with the incidence of lagged outcomes. Win this comparison, we are looking for comparable results of the extent to which the risk scores anticipate the incidence of lagged conflict.

In the 2020 study, higher social and economic domain risk scores were associated with a high occurrence rate across all conflict outcomes. These associations were significant at the $p<=0.01$ and $p<=0.05$ levels of significance, respectively. The political domain risk scores were negatively associated with the occurrence of conflict as well as human deaths. The political domain significance level was $p<=0.01$ level.

The positive and highly significant association between social domain risk score with the occurrence of conflict indicates the social domain was best performing risk measure. In contrast, the negative association between political domain risk score and conflict occurrence and deaths is somewhat puzzling. We suspect that sensitivities in some of the indicators within this political domain reversed the direction of its relationship with conflict.

Also, some of the indicators may have been perceived in a way that differed from their intended purpose or their weights may have been ill-suited to capture the complexity of politics in the region. In the current study, a higher social sector risk score was associated with the occurrence of conflict at a p<=0.03 level of significance. However, the security sector risk scores were similarly associated at an even higher (0.00%) level of significance.

Thus, the 2023 social sector predictor is consistent with the 2020 social domain predictor and the new security sector predictor, with its high level of significance, seems to support our interpretation of the sensitivity about the previous political domain indicators. In conclusion, we consider the 2023 logit model results consistent with the 2020 and 2021 logit model results.

www.cewarn.org