scientific reports

OPEN
Applying causality to environmental security in Iraq

Ninoslav Malekovic[1✉], Maarten Vonk[1,2✉], Laura Birkman[1], Tim Sweijs[1], Anna V. Kononova[2] & Thomas Bäck[2]

Despite major advances in environmental security research, non-experimental observations typically rely on methods that retrieve or validate specific linkages, rather than uncovering broader causal mechanisms underlying environmentally driven armed conflict. This article demonstrates how recent advances in causal methodology can be applied to more comprehensively retrieve and validate a network of such linkages. By retrieving a more integrated causal structure of pathways from environmental variables to conflict activity, we offer a novel methodological perspective on how causal relationships among environmental, demographic, agricultural and armed conflict variables can be identified, and how associated causal hypotheses can be tested. To uncover these pathways and infer causal effects of natural conditions on conflict activity, we apply this methodology to non-experimental observations of armed conflict across Iraqi subdistricts. Our findings support the hypotheses that latent energy and soil moisture indirectly cause conflict activity. While confirming that armed conflict is positively mediated by population density, the results do not support the hypothesis that wheat production negatively mediates conflict. Finally, we discuss our methodological approach, clarify its limitations, propose future research directions, and consider the implications for evidence-based policy interventions.

Since the founding of the Correlates of War project in 1963[1], armed conflict research has benefited from advances in the scientific method, including data collection[2,3], modeling[4] and prediction[5]. However, an outbreak of armed conflict remains difficult to predict, and any attempt to do so requires an explanation of its underlying causes[5,6]. These challenges have led many scholars in environmental security research to refrain from offering causal explanations altogether. For instance, one of the field's founders, Homer Dixon, explicitly avoided "entangling himself in the metaphysical debate about the relative importance of causes of naturally caused armed conflict"[7].

Ever since, a host of studies have attempted to overcome many of these limitations, by using econometric, statistical, and qualitative causal inference methods for observational data[8–10]. Since the 1990s, pathways that causally trace armed conflict to environmental variables have become increasingly rich and diverse[11–16]. Observationally identified causes linking conflict, agriculture, and weather include rapacity dynamics[17–19] intensifying enmities along political fracture lines[20,21], and food prices[22,23] among others.

Until now, however, methods have only been able to empirically confirm less exhaustive causal mechanisms that link conflict to environmental factors[16,24,25]. To estimate causal effects, randomized controlled trials (RCTs) have typically been regarded as the gold standard, as random assignment of units to treatment or control eliminates confounding between assignment and outcome. Since such rigorous experimentation does not apply to armed conflict, more exhaustive causal mechanisms are still in search of appropriate methodology[16,25]. This leaves a significant gap in the literature not only for scholarly purposes but also for policy-making, as policy interventions can only effectively address causes once they are identified and estimated.

In addressing this gap, our article outlines a methodology for inferring causality from non-experimental observations of armed conflict. We leverage recent advances in the theory and methods of causality[26]. In particular, Pearl's contributions—such as causal diagrams, causal interventions, and a formalized approach to causal reasoning—provide a robust framework for testing commonly hypothesized causal pathways using observational data. Our cross-sectional dataset consists of 294 non-experimental observations, one for each subdistrict in Iraq (Arabic: ناحية nāḥiya), which serves as our unit of analysis. Our outcome variable is the count of conflict events, with each observation additionally described by explanatory variables, including demographics, vital resources, the environment, and weather. These observations were sampled from several geo-coded maps[2,27–32]. From these data, we derived an empirical causal mechanism outlining linkages between environmental variables and conflict. Represented as a causal graph, the mechanism illustrates causal pathways from environmental variables to armed conflict outcomes. The mechanism is characterized by the causal effects

[1]The Hague Centre for Strategic Studies, Lange Voorhout 1, 2514 EA The Hague, Netherlands. [2]Leiden Institute of Advanced Computer Science, Einsteinweg 55, 2333CC Leiden, Netherlands. ✉email: malekovic@quantifiable.org; maartencvonk@gmail.com

of these variables on the count of conflict events, accounting for causal spillover wherever these effects could be identified and estimated. Several findings confirm that causal effects on conflict events can be traced back to their natural causes.

While we refrain from making strong claims about having established a robust causal mechanism, our study highlights exogenous dynamics that must be considered in environmental security research. More importantly, by demonstrating how advances in causal methodology can be leveraged in armed conflict research with non-experimental observations, we show how empirical causal structures, the identification of causal paths, and the estimation of causal effects can inform environmental security research. Third, we corroborate the existence of cause-and-effect relationships between environmental variables and conflict. Finally, we illustrate how causal frameworks can be used to inform more effective policy interventions.

Our study adapts causal methodology, including the retrieval of empirical causal structures and the estimation of causal effects under spillover conditions, to test existing hypotheses in armed conflict research using observational data—a novel approach in this domain. Furthermore, it applies these frameworks to environmental conflict research, offering new insights and paving the way for more effective policy interventions.

The article is structured as follows: Section 2 demarcates the research gap in the existing environmental security and armed conflict literature. Section 3 presents the hypotheses, and Section 4 outlines the results. Section 5 discusses key contributions, research limitations, future directions, practical implications, and conclusions. Finally, Section 6 describes the study's data and methods.

Related studies

Armed conflict prediction is generally difficult[6,33,34]. It is particularly challenging in localities where armed conflict is expected to occur for the first time[35]. Efforts to overcome the challenges and pitfalls of armed conflict prediction have occasionally grounded prediction in environmental variables[36]. A specific line of environmental security research has focused on the causality of armed conflict. Studies have investigated both the connections tracing the causes of violent, organized conflict to scarce resources[11,16,37] and the causal typologies of environmental conflict[11,14,15].

Some studies have focused on and explicated specific pathways using econometric, statistical, and qualitative causal inference methods for observational data. An overview of these studies can be found in, e.g., Martin-Shields and Stojetz[8]; Von Uexkull and Buhaug[9]; Ide et al.[10]. Researchers have identified numerous causes linking conflict, agriculture, and the environment using observational data, including rapacity dynamics[17–19], intensifying enmities along political fracture lines[20,21], and food prices[22,23], among others. Causes of conflict have been attributed to environmental stress and scarcity[12,13,37], the allocation and management of scarce water resources[38–40], and shocks in food production[41].

Some high-profile studies such as Hsiang, Burke, and Miguel[24] have attracted sharp criticism[41] for capturing publication bias and making questionable omissions from data, rather than focusing on actual dynamics. The critique demonstrated that the inferred impact of environmental conditions on conflict is contingent upon the context, the type of violence, the involved actors, and the methodology itself. By extension, broad claims about causal pathways linking the environment to conflict are, at best, weak and inconclusive[16,25,42].

Hence, while domain knowledge strongly suggests that environmental conditions contribute to organized armed conflict within countries, more comprehensive causal mechanisms linking these conditions to conflict remain less understood. Understanding these mechanisms in greater detail will likely require methodological advances to retrieve them from empirical data[43]. This is the research gap that our study aims to address.

Theory and hypotheses

This section derives our causal hypotheses from relevant findings. Adapted from Sakaguchi, Varughese, and Auld[16], Fig. 1 outlines the basic mechanism of causal linkages between the environment and conflict, a mechanism that has been extensively hypothesized in environmental security literature. These hypothetical linkages are rooted in environmental causes. The figure distinguishes between direct and indirect linkages (i.e., paths A and B, respectively). The indirect linkages are based on the idea that the scarcity of vital resources—also referred to as environmental scarcity—hypothetically mediates the relationship between environmental causes and armed conflict outcomes. This is the rationale behind our hypotheses.

Long-term weather patterns have been argued to directly cause armed conflict[24,44–47]. The direct link between long-term weather patterns and armed conflict can be seen in how populations respond to environmental changes. For instance, environmental disruptions affecting livelihoods can prompt community mobilization and lead armed groups to intervene, anticipating adverse outcomes. These actions may lead to direct causal effects of environmental processes on armed conflict. Such effects can originate from factors such as soil moisture, temperature, or the way different physical environments absorb or release accumulated heat or energy (i.e., latent heat or energy)[16,24]. Therefore, we hypothesize that changes in soil moisture, temperature, and latent energy directly cause changes in armed conflict activity (H1).

- H1a: An increase in latent energy in the form of heat directly causes an increase in armed conflict activity.
- H1b: An increase in skin temperature directly causes an increase in armed conflict activity.
- H1c: An increase in soil moisture directly causes a decrease in armed conflict activity.

Further, environmental processes have been argued to indirectly cause armed conflict[16,48,49]. Causal mediation of environmental effects on armed conflict primarily concerns environmental scarcity[7,12,17,37]. Environmental scarcity has been proposed as a mediator of environmental effects on armed conflict[7,16]. While we elaborate on causal mediation more specifically below, we can already hypothesize that the environmental processes indirectly cause armed conflict.
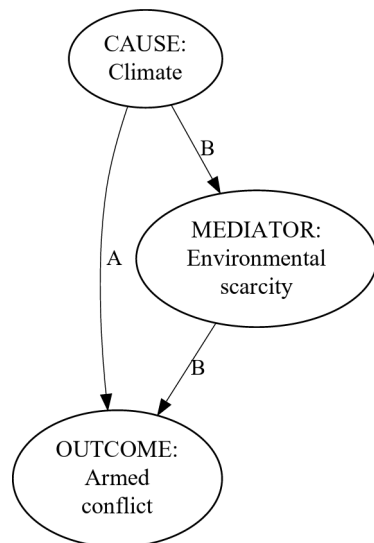
Fig. 1. Hypothesized causal pathways linking the environment to conflict.

- H2a: An increase in latent energy indirectly causes an increase in armed conflict activity.
- H2b: An increase in skin temperature indirectly causes an increase in armed conflict activity.
- H2c: An increase in soil moisture indirectly causes a decrease in armed conflict activity.

To add specificity to the indirect causal paths, causal mediation depends on specific causal conditions. Agricultural and pastoral conditions have been argued to shape social response to long-term weather patterns that induce migration[50]. Additionally, degradation and desertification of arable land, as well as the availability of water for agriculture, have been argued to mediate the environmental effects on violent conflict[10,51,52]. Wheat production has been shown to mediate the causal effects of temperature on the emergence of actual violence[19]. Since wheat is one of Iraq's key crops[53], we hypothesize that the environmental effects are also mediated by the production of agricultural resources, specifically wheat production.

- H3a: Given the indirect paths from the environmental processes to armed conflict activity, wheat production causally mediates the effects of these processes on armed conflict activity, by further decreasing armed conflict activity.

In line with the previous reasoning, since the association between conflict activity and factors such as population size, growth, density, or migration has already been established[54–56], these variables may also naturally mediate the effects of the environmental causes on armed conflict outcomes. Specifically, the scarcity of vital resources matters more to denser populations than to less dense ones[48,57]. Intuitively, a denser population is more likely to cope with and mitigate tensions less effectively than a less dense population. Everything else being equal, the denser population is, therefore, more likely to succumb to organized violence caused by environmental processes. Thus, our next hypothesis, H3b, aligns with this reasoning.

- H3b: Given the indirect pathways from the environmental processes to armed conflict activity, population density causally mediates these effects by further increasing armed conflict activity.

With these hypotheses, we can construct the entire hypothetical causal structure of linkages between environment and conflict, as shown in Fig. 2. As depicted in the figure, grounded in the environmental processes, the scarcity of vital resources exposes population to existential stress. Both population density and the scarcity of agricultural resources aggravate these effects.

Results
Empirical causal structure
In the theory section, we proposed a hypothetical causal structure for the linkages between environment and conflict. Figure 3 presents the causal structure that we empirically retrieved from the available non-experimental observations.

Although somewhat less expressive, the empirical causal structure largely corresponds to the hypothetical one in Fig. 2. The conflict nodes are clustered together, with the only node having only incoming edges being the count of conflict events. Additionally, the structure is rooted in the environmental processes. Apart from the direct causal path from the temperature node to battle events, all other paths from the environmental processes to conflict events are indirect.

Because the population density, skin temperature, and wheat production nodes have the highest number of incoming and outgoing edges, these nodes are pivotal to the connectedness of empirical causal structure.
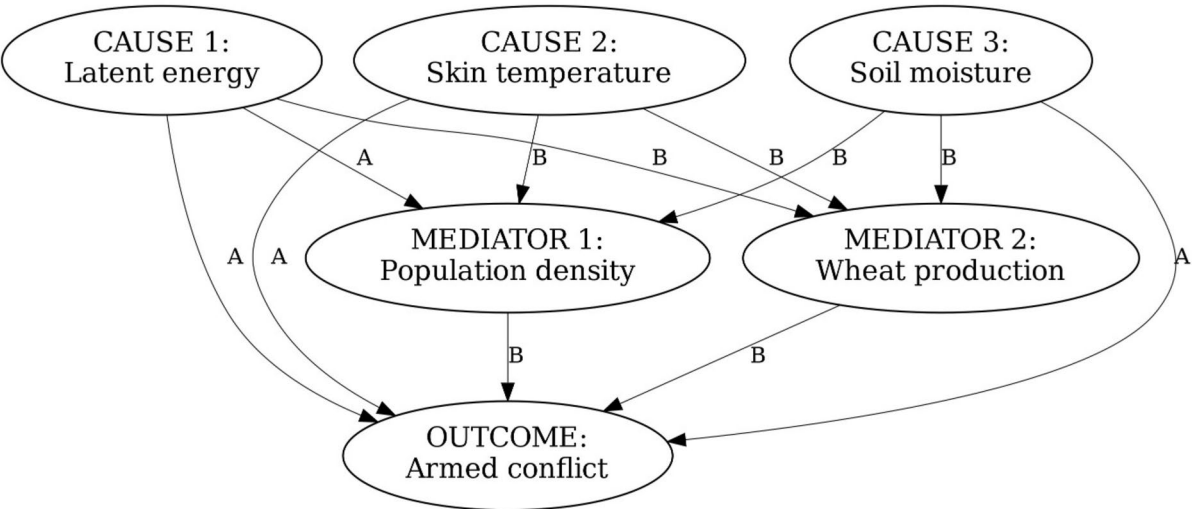
Fig. 2. All the hypotheses combined: The hypothetical causal structure.
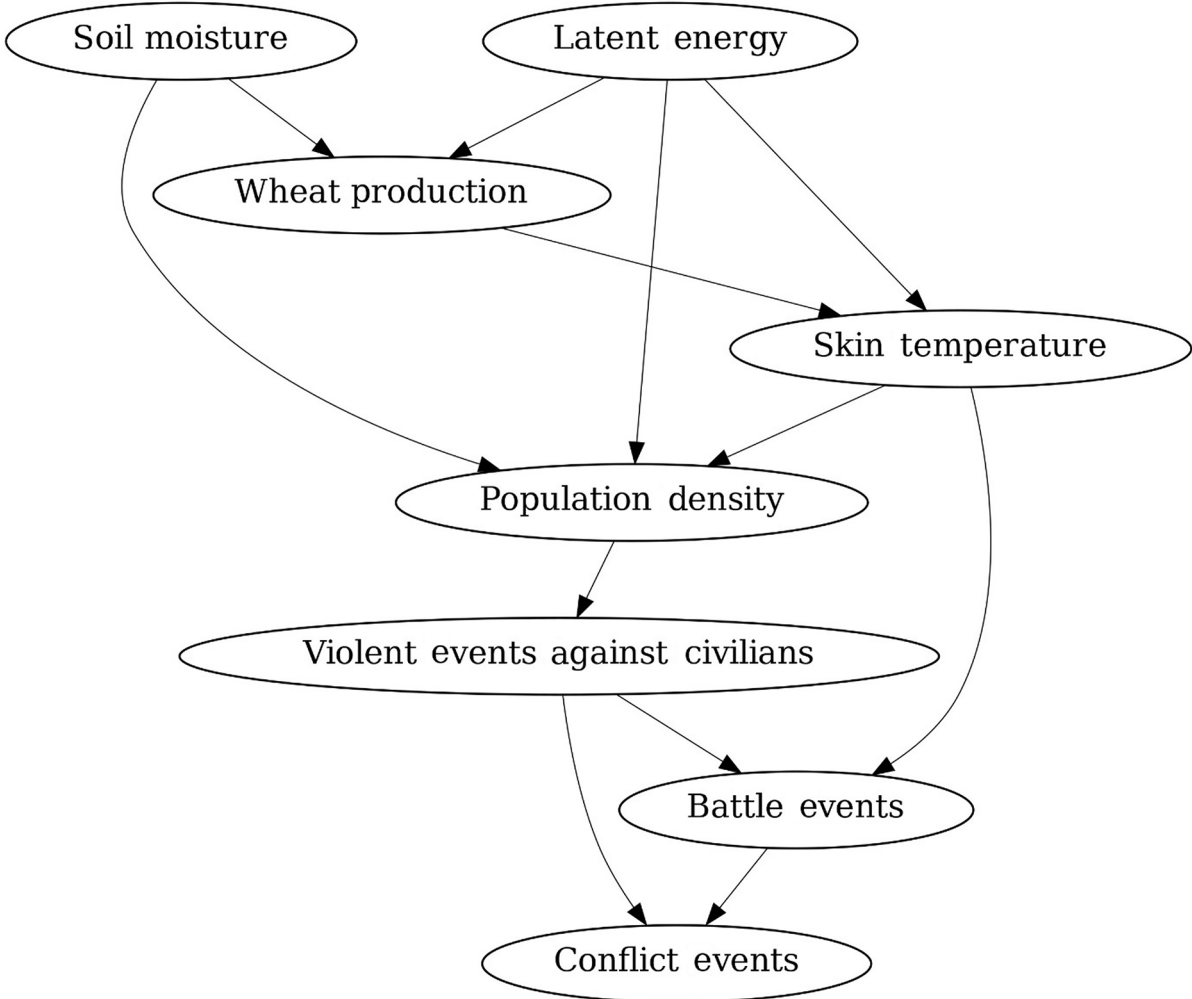

Fig. 3. The empirical causal structure retrieved from the aggregated data.

This supports the causal mediation of environmental processes on conflict outcomes. In fact, the absence of population density, wheat production, and skin temperature would significantly disconnect the environmental causes from these outcomes. Rather than accepting this evidence as conclusive, we further use the empirical structure to conduct hypothesis testing.

Causal hypotheses

The empirical causal structure can assist in validating the causal hypotheses of naturally caused armed conflict. Using this structure, we fitted a spatially explicit structural equation model (SESEM) and formulated and estimated the causal effects of our explanatory variables on the count of conflict events.

The SESEM model demonstrated acceptable performance, with the Comparative Fit Index (CFI) exceeding 0.90 and the Standardized Root Mean Square Residual (SRMR) falling below 0.08 across nearly all spatial distances[58]. These values indicate an acceptable model fit, highlighting the model's effectiveness in capturing the underlying data structure. While the Root Mean Square Error of Approximation (RMSEA) occasionally exceeded 0.10, suggesting room for improvement, this is likely due to the limited number of available observations.

For each causal estimate, a hypothesis test was conducted to determine whether the estimate should be attributed to random error. Table 1 lists the causal estimates, the standard errors, and statistical significance.

Our first set of hypotheses posits that the environmental processes directly cause armed conflict outcomes. Specifically, changes in latent energy (H1a), skin temperature (H1b), and soil moisture (H1c) were hypothesized to directly affect the count of conflict events. Among these, only skin temperature shows a direct causal path to conflict events through battle events. The estimates, standard errors, and statistical significance for each isolated path from skin temperature to conflict events are provided in Table 2.

Our first set of hypotheses posits that the environmental processes directly cause conflict events. Based on the empirical causal structure, the estimated causal effect of skin temperature on total conflict events is statistically significant at the 0.1% level and shows a positive relationship to conflict eventsacross all spatial distances (see

Distance	Wheat production	Latent energy	Soil moisture	Skin temperature	Population density
0 km	14.02*** (3.87)	24.27*** (6.21)	-18.30*** (4.57)	51.67*** (9.53)	29.66*** (5.75)
3–27 km	23.48*** (4.06)	33.92*** (5.84)	-20.72*** (4.23)	64.16*** (7.55)	22.61*** (4.26)
27–40 km	10.69*** (2.57)	21.83*** (4.28)	-17.93*** (3.13)	48.17*** (7.44)	29.42*** (4.48)
40–50 km	15.62*** (3.38)	26.64*** (5.19)	-19.95*** (3.83)	56.04*** (8.13)	26.65*** (4.50)
50–59 km	12.38*** (2.69)	23.22*** (4.41)	-18.93*** (3.42)	45.23*** (6.45)	34.42*** (4.67)
59–66 km	7.59*** (2.22)	15.05*** (3.54)	-10.41*** (2.40)	36.99*** (6.97)	16.52*** (3.28)
66–73 km	13.99*** (3.14)	28.43*** (5.35)	-16.99*** (3.79)	58.17*** (7.79)	18.17*** (4.08)
73–80 km	14.62*** (3.71)	28.87*** (5.54)	-19.53*** (4.04)	71.40*** (8.98)	32.45*** (5.93)
80–86 km	11.18*** (2.57)	15.61*** (3.61)	-14.71*** (2.85)	41.21*** (6.85)	26.29*** (3.97)
86–92 km	14.45*** (2.72)	20.63*** (4.07)	-16.43*** (3.08)	47.49*** (6.32)	23.16*** (3.84)
92–99 km	12.00*** (2.51)	21.88*** (4.28)	-19.68*** (3.45)	41.15*** (5.97)	26.86*** (3.98)
99–104 km	18.30*** (3.68)	26.26*** (5.17)	-25.15*** (4.25)	63.40*** (8.29)	43.84*** (5.59)
104–110 km	23.17*** (4.55)	37.04*** (6.58)	-24.85*** (4.79)	81.29*** (9.38)	27.69*** (5.08)
110–116 km	7.70*** (1.96)	13.16*** (3.14)	-14.46*** (2.63)	30.19*** (5.87)	30.58*** (3.71)
116–121 km	11.67*** (2.65)	23.80*** (4.76)	-14.04*** (3.04)	45.79*** (7.41)	29.06*** (4.32)
121–126 km	6.92*** (1.73)	13.87*** (3.02)	-16.06*** (2.66)	33.26*** (5.54)	27.72*** (3.40)
126–132 km	16.08*** (3.01)	23.86*** (4.48)	-16.46*** (3.18)	50.07*** (6.82)	21.62*** (3.52)
132–137 km	8.66*** (2.26)	15.93*** (3.97)	-13.12*** (2.83)	33.65*** (6.82)	18.19*** (3.37)
137–142 km	11.02*** (2.54)	16.71*** (3.82)	-14.94*** (2.91)	41.28*** (6.97)	22.61*** (3.55)
142–147 km	15.89*** (3.19)	20.92*** (4.60)	-17.00*** (3.57)	54.80*** (7.76)	31.78*** (4.57)
147–152 km	9.75*** (2.20)	18.46*** (4.04)	-11.33*** (2.53)	30.66*** (5.74)	15.23*** (3.49)
152–158 km	11.11*** (2.47)	17.22*** (4.00)	-14.67*** (2.97)	34.63*** (6.26)	20.02*** (3.31)
158–163 km	14.61*** (2.98)	27.24*** (5.22)	-13.79*** (3.30)	56.28*** (7.80)	25.70*** (4.71)
163–168 km	14.06*** (2.83)	22.30*** (4.62)	-18.70*** (3.60)	43.42*** (6.85)	29.26*** (3.96)
168–173 km	11.76*** (2.73)	18.74*** (4.48)	-15.62*** (3.28)	32.77*** (6.59)	28.39*** (4.20)
173–178 km	13.12*** (2.98)	23.36*** (4.77)	-15.54*** (3.33)	54.18*** (7.64)	24.45*** (4.31)
178–183 km	8.16*** (2.37)	12.68*** (3.43)	-12.00*** (2.62)	41.87*** (8.02)	24.57*** (4.18)
183–188 km	17.17*** (3.43)	22.95*** (4.99)	-20.85*** (4.09)	48.57*** (7.56)	39.79*** (5.41)
188–193 km	14.64*** (2.90)	25.01*** (5.14)	-20.75*** (4.10)	50.37*** (6.90)	19.87*** (3.66)
193–198 km	15.00*** (3.18)	23.31*** (4.93)	-17.27*** (3.58)	48.83*** (7.66)	24.65*** (4.80)

Table 1. The causal effects across varying distances. The table presents standardized causal estimates across a range of distance bins. The corresponding standard errors, which reflect the uncertainty of the estimated causal effects, are shown in parentheses. Statistical significance levels: 5%: *; 1%: **; 0.1%: ***.

Distance	Temperature → Population → Violence against civilians → Battles → Conflicts	Temperature → Battles → Conflicts	Temperature → Population → Violence against civilians → Conflicts
0 km	1.39 (0.83)	47.35*** (9.36)	2.93* (1.25)
3–27 km	0.13 (0.17)	63.34*** (7.50)	0.69 (0.72)
27–40 km	1.29* (0.65)	44.88*** (7.31)	2.00** (0.82)
40–50 km	0.68 (0.52)	52.46*** (8.03)	2.91** (1.02)
50–59 km	0.52 (0.42)	41.22*** (6.28)	3.49** (1.28)
59–66 km	0.19 (0.26)	35.14*** (6.93)	1.66* (0.70)
66–73 km	0.87 (0.54)	55.63*** (7.76)	1.66*** (0.54)
73–80 km	0.58 (0.52)	68.62*** (8.88)	2.20* (1.04)
80–86 km	0.30 (0.48)	37.21*** (6.74)	3.70*** (1.18)
86–92 km	0.36 (0.27)	45.67*** (6.25)	1.46 (0.75)
92–99 km	2.05** (0.74)	36.79*** (5.86)	2.30*** (0.64)
99–104 km	1.31 (0.71)	58.73*** (8.06)	3.36* (1.41)
104–110 km	0.58 (0.56)	77.91*** (9.29)	2.79** (1.06)
110–116 km	2.15** (0.76)	24.85*** (5.65)	3.19*** (0.97)
116–121 km	0.95 (0.52)	42.72*** (7.28)	2.12* (0.98)
121–126 km	1.73* (0.70)	28.14*** (5.34)	3.39*** (1.01)
126–132 km	1.13* (0.51)	47.13*** (6.73)	1.82** (0.69)
132–137 km	0.82* (0.39)	30.58*** (6.75)	2.25*** (0.76)
137–142 km	1.53* (0.61)	37.31*** (6.87)	2.45*** (0.74)
142–147 km	1.58 (0.72)	49.34*** (7.62)	3.88*** (1.10)
147–152 km	0.35 (0.24)	29.54*** (5.69)	0.76 (0.49)
152–158 km	0.97* (0.45)	31.38*** (6.18)	2.28*** (0.71)
158–163 km	0.63 (0.51)	53.10*** (7.72)	2.55*** (0.90)
163–168 km	1.04 (0.56)	39.35*** (6.68)	3.04** (1.14)
168–173 km	0.71 (0.48)	30.67*** (6.47)	1.39 (0.83)
173–178 km	1.96* (0.87)	49.37*** (7.56)	2.85*** (0.78)
178–183 km	1.14 (0.65)	38.80*** (7.92)	1.93* (0.82)
183–188 km	1.84* (0.84)	43.89*** (7.36)	2.84** (1.08)
188–193 km	0.52 (0.50)	47.06*** (6.84)	2.79*** (0.85)
193–198 km	1.79* (0.90)	44.92*** (7.56)	2.12** (0.74)

Table 2. The causal effects of skin temperature on conflict events across three different paths and varying distances. The table presents standardized causal estimates of the effects of skin temperature on conflict events across three distinct paths along a range of distance bins.. The corresponding standard errors, which reflect the uncertainty of these estimates, are shown in parentheses. Statistical significance levels: 5%: *; 1%: **; 0.1%: ***.

Table 1). This includes a 47.35 non-spillover effect at a distance of zero (see Table 2). Since there is no evidence to reject the null hypotheses for (H1a) and (H1c), we accept (H1b), and we do not accept (H1a) and (H1c).

Further, our second set of hypotheses posits that the environmental processes indirectly cause conflict events. Latent energy was hypothesized to increase the count of conflicts indirectly (H2a). According to the empirical causal structure, all causal effects of latent energy on conflict events must be mediated. Everything else being equal, the estimated causal effects of latent energy on conflict events are statistically significant at the 0.1% level and positive across all spatial distances. This includes a 24.27 non-spillover effect at a distance of zero (see Table 1). Thus, we accept (H2a). Similarly, soil moisture was hypothesized to decrease the count of conflicts indirectly (H2c). With no direct paths from soil moisture to conflict events, its effects can only be mediated. Everything else being equal, the estimated causal effects of soil moisture on conflict events are statistically significant at the 0.1% level and negative across all spatial distances, including a -18.30 non-spillover effect at a distance of zero (see Table 1). Therefore, we also accept (H2c). However, for the indirect paths from skin temperature to armed conflict activity through population density (H2b), , the estimated causal effects of skin temperature on conflict events are not statistically significant at the 0.1% level across all spatial distances (see Table 2). Consequently, we do not accept (H2b).

Finally, our third set of hypotheses posits that the causal effects of the environmental processes on conflict events are agriculturally and demographically mediated. Specifically, given the indirect paths from soil moisture and latent energy to armed conflict activity, wheat production was hypothesized to decrease the count of conflict events (H3a), while population density was hypothesized to increase the count of conflict events (H3b). Everything else being equal, the estimated causal effects of wheat production on conflict events are statistically significant at the 0.1% level and positive across all spatial distances, including a 14.02 non-spillover effect at a distance of zero (see Table 1). Similarly, the estimated causal effects of population density on conflict events are statistically significant at the 0.1% level and positive across all spatial distances, with a 29.66 non-

spillover effect at a distance of zero (see Table 1. These findings lead us to reject the null hypotheses for (H3a) and (H3b), not accepting (H3a) and accepting (H3b) across all spatial bins. Finally, the removal of wheat production, skin temperature. and population density from the empirical causal structure would significantly disconnect soil moisture and latent energy from conflict events. Our causal estimates corroborate these structural findings.

Seasonal effects

In addition to the aggregated results, we applied the causal discovery algorithm to temporally non-aggregated data slices at each season's mid-point. The results are presented in Figs. 7, 8, 9 and 10 of the Appendix. As seen in the figures, applying the causal discovery algorithm to the seasonal data slices reveals highly irregular empirical causal structures that deviate from the prevailing domain knowledge.

Discussion

We make three contributions to armed conflict research. First, we demonstrate how advances in causal methodology can be applied to armed conflict research. While natural experiments on armed conflict are rare, and most observations are non-experimental[24], causal assumptions can bridge the gap between experimental and non-experimental studies. By leveraging causal methodology, we show how to retrieve an empirical causal structure, identify causal paths, and estimate causal effects using non-experimental data. Additionally, the probability of distances between hypothesized and empirically discovered causal structures can be quantified and tested[59,60,] especially when domain knowledge informs causal discovery of naturally caused armed conflict. Otherwise, an empirical causal structure can be retrieved exploratorily, as shown in our study.

Second, in the context of environmental security research, we confirm the existence of cause-and-effect relationships between environmental and demographic variables and armed conflict outcomes. We demonstrate that environmental processes, mediated by agriculture and demographics, play a role in causing conflict. By disentangling the causal factors underlying the naturally caused armed conflict, we addressed the research gap in the environmental security literature[16,25,43]. The early research avoided engaging in the metaphysical debate over the relative importance of natural causes of armed conflict[7]. By retrieving the empirical causal structure from non-experimental conflict observations, we identified the causal mechanism linking the environment to conflict. Our empirical causal structure shows that the environmental processes, particularly soil moisture and latent energy, can be at the core of naturally caused armed conflict. We also established direct and indirect causal effects, such as the direct aggravating effect of skin temperature andthe indirect alleviating effect of soil moisture. Our findings show that the effects of these environmental causes on conflict events are mediated by demographics and vital resources. Additionally, we confirm that population density aggravates conflict however, contrary to our hypothesis, we find no evidence that higher wheat production alleviates conflict activity.

Third, we show how our causal findings can inform the design of more effective policy interventions to address armed conflict outcomes. Our findings provide a foundation for policy approaches that mitigate armed conflict. Policy acceptance hinges on unpacking black-box predictions[5], and because causality explains underlying mechanisms, it offers greater interpretability than other approaches to analysis do[61]. Causality also holds policy relevance. For instance, mediators such as population density and wheat production are key points where indirect causal paths from the environmental factors to conflict outcomes can be strategically targeted for intervention. Causal frameworks also enable validation of additional such points. Thus, understanding the causal mechanisms behind naturally caused armed conflict can guide the development of policies aimed at addressing its root causes.

Our research is constrained by several limitations. First, regarding our variables, many vital resources are influenced by factors such as precipitation, temperature, energy, and water availability[24,51,52,62,63], which justifies the selection of wheat, skin temperature, soil moisture, and latent energy as explanatory variables. Population density is also a well-established variable[54,55,64]. However, the causal sufficiency assumption posits that any omitted, causally relevant variable can confound causal estimates[65]. In our case, this primarily applies to social and political variables (e.g., political power-sharing arrangements, inter-group animosities, and horizontal inequality)[2,66,67]. The absence of these variables may have violated the causal sufficiency assumption, potentially confounding our causal estimates. Unfortunately, such data is not available at an adequate resolution for a data-scarce country like Iraq. Even if these variables had been available, however, their inclusion might have introduced reverse or cyclical causality, thereby confounding our estimates. Hence, the trade-off between methodological rigor and domain knowledge is an inherent challenge of causal modeling.

Our second limitation relates to the geographical nature of subdistricts. The need to delimit the conflict under study both geographically and historically[5] constrains the generalizability of our findings to other geographies. This reflects a broader trade-off: the need to localize armed conflict spatially and temporally versus the need to ensure sufficient variability in the sampled observations. Moreover, because subdistricts are geographical units, it cannot be assumed that the observations are independently and identically distributed. Neither conflict activity nor the explanatory variables necessarily conform to subdistrict borders. Specifically, the tribal dynamics play a role in shaping armed conflict in Iraq, and the subdistrict and tribal boundaries do not always coincide[68]. As a result, it is plausible that some Iraqi subdistricts are affected by causal effects from other Iraqi subdistricts[40,69,70]: While our spatially explicit structural equation modeling accounted for such spillover effects, the causal discovery method used to retrieve the causal structure does not incorporate a remedy for this spatial dependency.

Our third limitation pertains to the temporal dimension of causality. Since causes are conventionally expected to precede their effects, temporality is a fundamental property of causal inference. However, the limited availability and quality of our time-series data constrained our ability to model the temporal dynamics of naturally caused armed conflict. Specifically, the available time-series were not sufficiently granular or consistent to support robust temporal modeling. This lack of time-series data prevented us from inferring seasonal effects

beyond the irregular causal structures, deviating from the domain knowledge. The irregular causal structures for the seasonal snapshots are unlikely to be robust, as it is unlikely that the entire environmental conflict dynamic could have unfolded within a single month.

Several research directions could address the limitations outlined above. First, given the exclusion of social and political variables from our study, subsequent studies should assess the robustness of our causal findings under the assumption of unobserved confounding—i.e., by relaxing the causal sufficiency assumption. Causal frameworks for this purpose already exist[69,71,72]. An increasing number of these tools are now available as off-the-shelf solutions, providing researchers with practical means to integrate them into their methodology[72–74].

Further research should also investigate the generalizability of our findings to other geographical localities. The smaller the size of locality, the less variable are the observations. Consequently, the less likely it is that their variability suffices for statistical significance. Any reduction in the size of locality can artificially make causal findings less statistically significant. However, it is also possible to overcome the reduced variability of observations, characteristic of small-size localities, by describing conflict observations in informationally dense terms. This can be achieved by describing each locality with additional explanatory variables, while dimensionally reducing the number of explanatory variables to a smaller number of latent factors. Armed conflict outcomes can then be modeled in terms of these latent factors. This approach allows us to examine more diverse contextual data and assess the extent to which our findings can be generalized and applied to different geographical regions and conflict contexts.

Geography also matters for strategy. Natural resources are inherently geo-located, as are competing factions. Claiming these resources, competing factions often act strategically. These dynamics rarely unfold in a socio-political and demographic vacuum. For instance, the already mentioned cross-subdistrict tribal networks can introduce both natural and strategic contingencies into conflict dynamics, as has been observed across Iraqi subdistricts. Obviously, armed conflict within some Iraqi subdistricts may contract causal effects from other subdistricts because of geographical reasons. However, further investigation is warranted into whether such spillover effects are primarily driven by natural environmental factors or by strategic human agency. Causal frameworks that support such investigation are already available for use[69,70].

Our findings should also be examined for temporal effects. The retrieved empirical causal structure already suggests the non-linearity of armed conflict outcomes[75]. Future research should, therefore, investigate how to infer causal effects on armed conflict dynamics across time. Causality can be temporally represented by causal time-series graphs[76]. Such graphs can identify time-varying causal paths and guide the estimation of temporal causal effects. However, not only would this require time-series data of sufficient quality, but also a plausibly hypothesized mechanism of temporal causal effects, which may go beyond our explanatory variables. Additionally, most currently available causal discovery methods for time-series data rely on the stationarity assumption, which limits their ability to detect and account for seasonal fluctuations in conflict dynamics[77].

By explaining why phenomena occur as they do, causal analysis offers greater pragmatic utility for policymaking than any other methodological approach. This strengthens the case for accepting our findings and underscores the broader value of applying causal methodology. This also encourages adoption of causal methodology in other policy domains where decisions can benefit from empirically grounded causal evidence.

In the context of environmental security, understanding the causal linkages between environmental factors and conflict provides valuable guidance for designing conflict mitigation strategies. By targeting mediators along natural causal pathways to conflict outcomes, policy interventions can be crafted to disconnect these pathways before aggravating causal effects reach conflict outcomes. Such interventions can be cost-effective and even preventive. For instance, well-conceived social and migration policies aimed at reducing population density can disconnect some of these causal links. Similarly, development aid policies can mitigate conflict by addressing environmental stressors This can be achieved by strategic investments in long-term hydrological infrastructure. Finally, our findings enable the geographical assessment of Iraqi subdistricts for their specific vulnerability to particular causal effects, thereby informing the spatial targeting of policy implementation.

In conclusion, quasi-experimental approaches often fall short in the study of armed conflict, and natural experiments are rare. Further, ethical considerations preclude rigorous experimentation in this domain. To the best of our knowledge, our study is the first to apply the recent advances in the theory and methods of causality to the analysis of naturally caused armed conflict. Relying exclusively on non-experimental observations, we derived an empirical causal structure that corresponds to the underlying mechanism linking the environmental factors to conflict outcomes. Based on this structure, we exemplified how to infer the causal effects of natural processes on armed conflict. While our findings rest on specific assumptions, they offer a foundation that can be strengthened through additional validation. Such validation—particularly through interdisciplinary methods—is therefore welcome.

Data and methods
Data sources
The armed conflict activity variables were sourced from the Armed Conflict Location and Event Data Program[2] (ACLED). We also retrieved geo-coded maps from multiple sources, including the Humanitarian Data Exchange, the European Centre for Medium-Range Weather Forecasts of Copernicus Climate Change Service[28], NASA[27,29,31], the Center for International Earth Science Information Network of Columbia University[30], and MapSPAM[78]. These maps provided grid-based representations of our explanatory variables. Our unit of analysis is the Iraqi subdistrict (n=294), encompassing all subdistricts across the country. The observational time frame spans from January 1, 2020, to January 1, 2022. Within this period, observations were aggregated into a single cross-sectional dataset. The choice of this time horizon was motivated by the high availability of recent data. In cases where a grid layer was unavailable for the entire period, we used data covering the maximum available

portion of the horizon for the corresponding variable. The aggregation methods are specified in the sections below.

Despite a strong rationale for including additional explanatory variables that capture the societal and political context of Iraq[2,66,67], Iraq remains a data scarce country, and such variables are generally unavailable at a sufficiently granular resolution. We addressed this issue of unavailability as one of our key research limitations in the discussion section.

Conflict activity variables

We retrieved several armed conflict variables from ACLED[2]. These variables are geospatially linked to specific geographical coordinates and are geo-coded as counts. ACLED categorizes conflict activity into distinct types, including battles, explosions/remote violence, violence against civilians, protests, riots, and strategic developments.

Conflict events, violence against civilians
Violence against civilians refers to the deliberate infliction of harm on unarmed non-combatants by organized armed factions. Such events include acts of sexual violence, abductions, and forced disappearances[2].

Conflict events, battles
In addition to violence against civilians, we also included conflict events categorized as battles, which are defined as violent engagements between two organized armed groups.

Total conflict events
Our outcome variable was the total count of conflict events, proxied by summing the counts of each specific conflict event type (i.e., battles, explosions/remote violence, violence against civilians, protests, riots, and strategic developments).

Explanatory variables
Environmental variables
Some environmental conditions pertain specifically to weather. While weather refers to short-term atmospheric conditions, climate describes the average weather of a region over an extended period. We specified our environmental variables in line with Sakaguchi, Varughese, and Auld[16]. Given that temperature, soil moisture, and heat have been already hypothesized—and empirically demonstrated—to be associated with and causal to violent conflict[24], we selected these variables to represent the weather conditions in Iraq.

Skin temperature   Temperature is a physical quantity that indicates how hot matter is[79,80]. For our analysis, we selected skin temperature—the temperature at the interface between the Earth's surface and the atmosphere—because it directly influences the growth and cultivation of agricultural resources, as well as the availability of water[81]. Skin temperature data was sourced from ERA5-Land dataset[28]. Each pixel on the corresponding map, with a spatial resolution of 11,132 square meters, represents a temperature value measured in Kelvins.

Soil moisture   Soil moisture refers to the total amount of water, including water vapor, present in unsaturated soil[79,80]. We selected soil moisture as an explanatory variable because it impacts the growth and cultivation of agricultural resources, as well as the availability of water. The soil moisture data was also sourced from ERA5-Land dataset[28], specifically at a depth of 28–100 cm. Each pixel on the map, with a resolution of 11,132 square meters, represents a volumetric fraction of water at this depth.

Latent energy   Also referred to as latent heat, latent energy refers to the energy released from the Earth's surface to the atmosphere. Latent energy is associated with the evaporation or condensation of water vapor at the Earth's surface[79]. It represents an environmental process that extends beyond temperature but still impacts the physical surroundings. We sourced the average latent heat net flux[27] from NASA and United States Geological Survey's MODIS 006 MOD16A2 dataset. The flux represents the average latent energy passing through matter. Corresponding to a resolution of 500 square meters, each pixel on the map displays a value in Joules.

Environmental scarcity
Environmental scarcity refers to the lack of vital resources on which human communities directly and critically depend. Since the scarcity of agricultural resources can catalyze violent conflict[16], earlier findings on wheat production guided our proxy for crop availability in Iraq[19].

Wheat production   We sourced the total wheat production for rainfed and irrigated crops from the 2020 version of Global Spatially Disaggregated Crop Production Statistics Data (MapSPAM, Version 1.0) [78]. Each pixel on the map, corresponding to a resolution of 10,000 square meters, shows a value in metric tons. Due to the constrained time horizon of this dataset, the wheat production variable was only available for 2020 as the most recent observation.

Demographics
Population density   Given that population density has been found relevant for armed conflict activity[55], we sourced this variable from the Gridded Population of the World Version 4.11 dataset by the Center for International Earth Science Information Network at Columbia University[30]. Each pixel on the map, corresponding to a resolution of 927.67 square meters, shows an estimated number of people per 30 arc-second grid cell. Due to

the constrained time horizon of this dataset, the population density variable was sourced for 2020 as the most recent available observation.

From geo-coded observations to subdistrict values

Our variables are geo-coded, meaning that each corresponding value is linked to specific geographical coordinates. Since our unit of analysis is the Iraqi subdistricts, the value for each variable within a subdistrict must be an aggregate of the geo-located values of that variable within that subdistrict. However, the geo-coded maps containing these variables do not include subdistrict borders. To enable the geographical aggregation of pixel values across subdistricts, a shapefile with the subdistrict borders was integrated into each map. We sourced this shapefile from the UN OCHA Humanitarian Data Exchange, and it stores the subdistrict borders as a geometry variable (i.e., a polygon). The 294 subdistrict borders are shown in Fig. 4.

Further, the conflict activity variables are geo-located. For each subdistrict, we counted the geo-located values of conflict activity variables reported within that subdistrict's polygon. This procedure was applied to the count of battle events, violent events against civilians, and the total count of conflict events.

Furthermore, the available grids store values of the explanatory variables as pixels, each associated with specific geographical coordinates. We bounded these pixel values by the subdistrict polygons, as previously described. Since the storage of explanatory variables was no longer sparse, we aggregated the values of explanatory variables to highlight the extreme values of each observed explanatory variable. Specifically, pixel values for soil moisture were aggregated geographically as pixel minima and temporally as subdistrict standard deviations. Pixel values for skin temperature were aggregated geographically as pixel maxima and temporally as subdistrict standard deviations. Pixel values for average latent heat flux were aggregated geographically as pixel maxima and temporally as subdistrict standard deviations. Moreover, pixel values for wheat production were aggregated geographically as pixel means for the year 2020, the most recent available observation. Finally, pixels values for population density were aggregated geographically as pixel maxima for the year 2020, the most recent available observation.

Finally, considering that seasonality plays a crucial role in the relationship between environmental and conflict variables[82], we extracted four non-aggregated time slices corresponding to the seasonal midpoints: January, April, July, and October 2020. The methods detailed in the following section were applied to both the aggregated data and the non-aggregated time slices, with a focus on demonstrating where the method proves most effective.

Methods

Causal methodology traditionally requires experimentation[83–86]. However, it is now possible to infer causality even from non-experimental observations[26,87–90]. Acknowledging the need for non-experimental approaches in environmental security, we argue that it is possible to retrieve the causal structure linking the environment to conflict. By applying causal methodology to non-experimental observations, the underlying paths and effects of the causal structure can be identified and quantified. This process unfolds in three stages: causal discovery, causal identification, and causal inference.
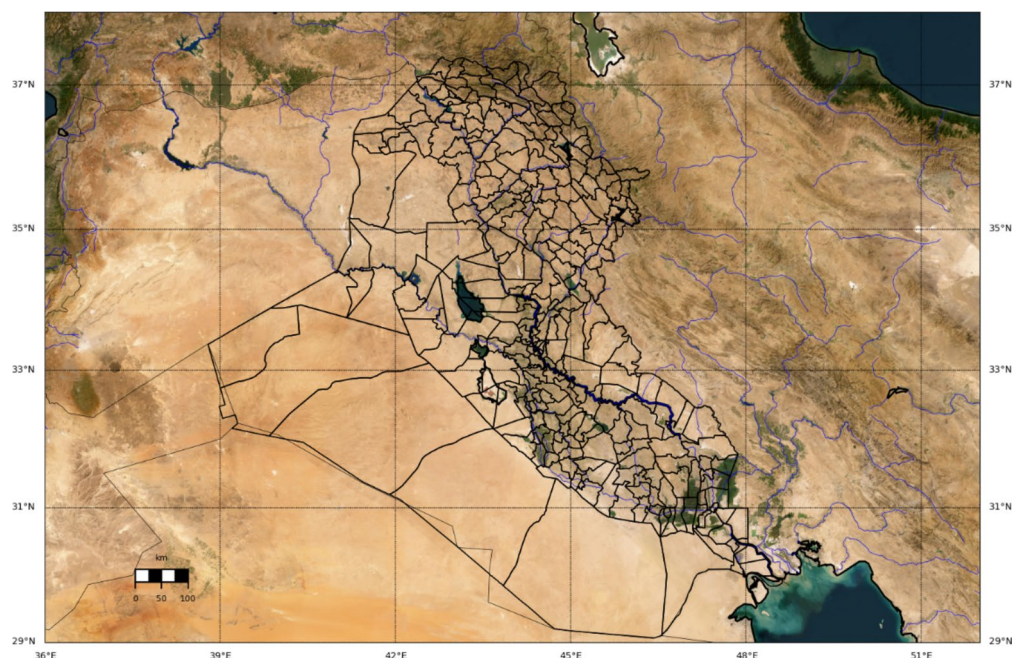

Fig. 4. Subdistrict borders in Iraq.

The following three subsections explain how causality is discovered, identified, and inferred from non-experimental observations. These subsections exemplify how to respond to a specific causal query: "*What is the magnitude of the causal effect of soil moisture on the count of conflict events?*"

Causal discovery

The purpose of causal discovery is to retrieve a causal structure from available observations[65]. These structures can be modeled graphically[26]. Each directed edge in such a causal graph represents causality between the node with an outgoing arrow (the cause) and the node with an incoming arrow (the effect)[26]. The graph on the left in Fig. 5 shows a directed edge: Soil moisture → Conflict events, indicating that a change in soil moisture causes a change in conflict events. In contrast, the graph in the middle shows a bidirected edge: Soil moisture ↔ Conflict events, suggesting a reciprocal relationship between soil moisture and conflict events.

Each node between a node with only outgoing arrows (i.e., a root cause) and a node with only incoming arrows (i.e., an effect or outcome) is a mediating node (e.g., population density). Further, the graph on the right in Fig. 5 is characterized by a causal cycle: Soil moisture → Conflict events → Population density → Soil moisture. Despite recent theoretical advances[91], the simplest conception of causality requires bidirected edges and causal cycles to be absent from a causal graph, as they indicate hidden common causes and reverse causality, both of which can confound causal inference[26]. Graphs without bidirected edges and cycles are referred to as directed acyclic graphs (DAGs). The simplest conception of causality dictates that causal discovery should retrieve a DAG from available observations[65].

Following this logic of the causal discovery stage, we retrieved a DAG from our observations using Greedy Equivalence Search (GES) algorithm[65]. We selected GES, because it is considered a suitable causal discovery method for small sample sizes[65]. We employed the Bayesian Information Criterion (BIC) as our loss function. The GES algorithm iteratively adds and removes edges in a stepwise manner, scoring each configuration to identify the structure that best fits the data, while optimizing the BIC. The output was the most likely DAG given our observations. The nodes of the DAG correspond to our armed conflict activity and explanatory variables, while the edges represent the causal relationships between them.

Causal identification

The purpose of causal identification is to determine whether, given a causal structure, the causal query has a unique answer[92]. If the query is unidentifiable, the identification process reveals this. Additionally, causal identification helps in formulating a quantity that provides a unique answer to the query[26,92]. This formula, which allows for the quantification of the answer, is referred to as an estimand[93].

Given the arrow from Soil moisture → Conflict events in Fig. 6, a node like population density introduces an alternative path between soil moisture and conflict events. If this node is not explicitly considered, it is referred to as a confounder[26]. If all confounders can be accounted for, the causal query can be properly identified.

Given the probabilistic interpretation of causal graphs[26], let $P(.|.)$ represent a conditional probability distribution. Let C, M, and D represent the conflict events, soil moisture, and population density variables, respectively, and let c, m, and d represent their corresponding realized values. The $do(.)$ operator represents intervention. If no variable were associated with soil moisture and conflict events (see Fig. 6), the unique answer to our causal query would have been $P(C|\,do\,(m)) = P(C\,|m)$. However, since population density introduces an alternative causal path between soil moisture and conflict events, failing to account for population density could prevent the determination of a unique answer to the query. Therefore, to identify the causal query properly, the confounding effect of population density must be marginalized. As a result, the causal graph in
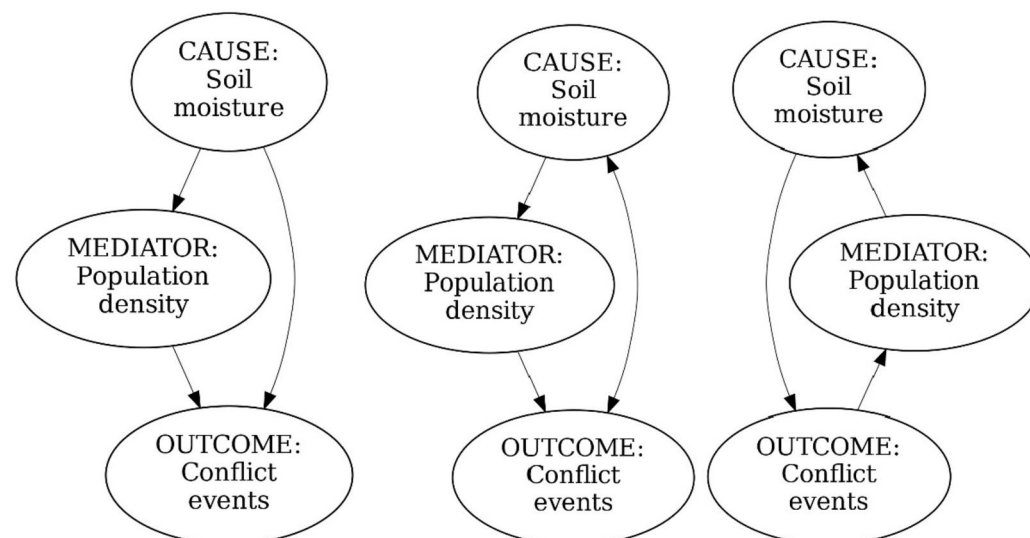

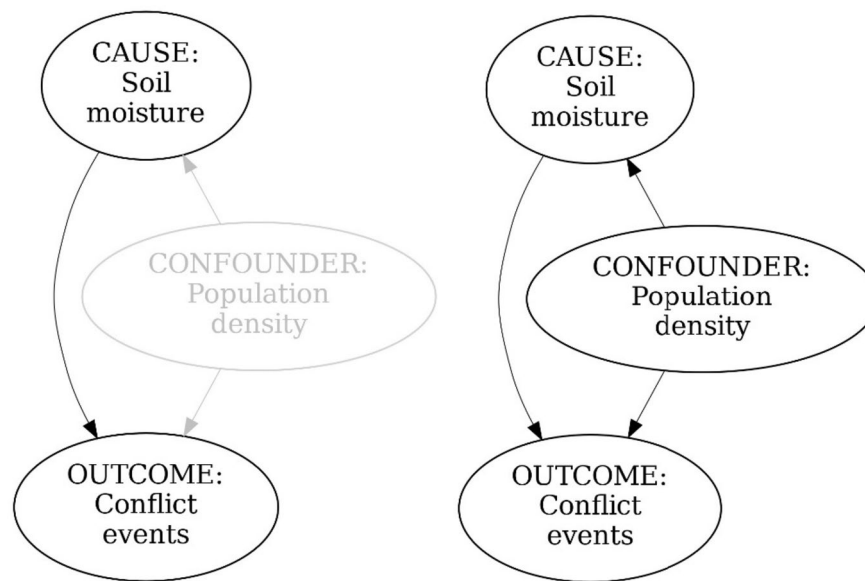**Fig. 5**. Bidirected edges and causal cycle in a causal structure.

Fig. 6. Confounder: Population density.

Fig. 6 enables the identification of a unique answer to our query, with $P\left(C|\ do\left(t\right)\right)=\sum\ _{D}P\left(C|\,t,\ D\,\right)P(\,D)$ representing this unique answer.

This reasoning illustrates an identification criterion known as the backdoor criterion[26]. For other identification methods, we refer the reader to Tian and Pearl[94] and Shpitser and Pearl[92]. By applying this reasoning to the retrieved DAG, we were able to identify causal estimands for our explanatory variables, provided that access to the relevant variables made such identification feasible.

Causal estimation and hypothesis testing
The final stage of causal inference involves estimation and hypothesis testing. In this stage, an estimator (i.e., a method for calculating estimates) is applied to a causal estimand, using sampled observations[26,93]. This process produces an estimated causal quantity of causal effect, i.e., causal estimate. Ultimately, an assessment is made to determine whether the estimate should be attributed to random error. If not, the quantity is deemed statistically significant.

Since the values of environmental variables in one subdistrict can influence the environmental security of another , the estimation procedure required a method capable of accounting for spatial confounding. To address this, we employed spatially explicit structural equation modeling (SESEM)[95].

SESEM combines structural equation modeling (SEM) with spatial dependence modeling to account for effects across varying spatial lag distances. While SEM was used to implement the retrieved causal structure, spatial dependence modeling addressed the conflict dynamics that extended beyond the subdistrict boundaries.

A straightforward way to incorporate spatial dependence modeling via lag distances is to calculate the distances between subdistricts. This can be done in several ways. The most intuitive method involves computing the centroid of each subdistrict and measuring the distances between these centroids. Alternatively, a network can be formed by connecting adjacent subdistricts with edges, in which case the distance between any two subdistricts equals the length of the shortest path between them.

The first step in applying SESEM was to fit a non-spatial SEM model to the data[96]. Next, spatially explicit variance–covariance matrices were computed across a range of lag distances, which were binned according to sample pair distances. We focused on the lowest 20% of these distances, as spillover effects are more likely to occur among neighboring subdistricts. To ensure reliable inference across each distance range, each bin was set to include 500 sample pairs. SEM models were then fitted for each lag distance, and edge coefficients, standard-errors, and p-values were computed. Finally, individual causal paths were parameterized to obtain path-specific coefficients, standard-errors, and p-values.

Data availability
All the data is available at https://github.com/HCSS-Data-Lab/Submission-Angling-For-Causality. This repository includes all the code for the method implementation, with comprehensive in-line documentation provided directly within the code files.

References

1. Suzuki, S., Krause, V. & Singer, J. D. The Correlates of War project: a bibliographic history of the scientific study of war and peace, 1964–2000. *Confl. Manag. Peace Sci.* **19**, 69–107 (2002).
2. Raleigh, C., Linke, R., Hegre, H. & Karlsen, J. Introducing ACLED: an armed conflict location and event dataset. *J. Peace Res.* **47**, 651–660 (2010).
3. Sundberg, R. & Melander, E. Introducing the UCDP georeferenced event dataset. *J. Peace Res.* **50**, 523–532 (2013).
4. Kress, R. *Numerical Analysis* vol. 181 (Springer Science & Business Media, 2012).
5. Cederman, L. E. & Weidmann, N. B. Predicting armed conflict: time to adjust our expectations? *Science* **355**, 474–476 (2017).
6. Ward, M. D. Can we predict politics? Toward what end? *J. Global Secur. Stud.* **1**, 80–91 (2016).
7. Homer-Dixon, T. F. *Environment, Scarcity, and Violence* (Princeton University Press, 2010).
8. Martin-Shields, C. P. & Stojetz, W. Food security and conflict: empirical challenges and future opportunities for research and policy making on food security and conflict. *World Dev.* **119**, 150–164 (2019).
9. Von Uexkull, N. & Buhaug, H. Security implications of climate change: A decade of scientific progress (2021).
10. Ide, T., Lopez, M. R., Frohlich, C. & Scheffran, J. Pathways to water conflict during drought in the MENA region. *J. Peace Res.* **58**, 568–582 (2021).
11. Deligiannis, T. The evolution of environment-conflict research: toward a livelihood framework. *Glob. Environ. Politics*. **12**, 78–100 (2012).
12. Homer-Dixon, T. F. On the threshold: environmental changes as causes of acute conflict. *Int. Secur.* **16**, 76–116 (1991).
13. Homer-Dixon, T. F., Blitt, J. *Ecoviolence Links among Environment, Population and Security* (Rowman & Littlefield, 1998).
14. Baechler, G. Conclusions: future relevance and priorities of small states. In *Small States Inside and Outside the European Union: Interests and Policies*, 267–283 (Springer, 1998).
15. Rodríguez-Labajos, B. & Martínez-Alier, J. Political ecology of water conflicts. *Wiley Interdiscip. Rev. Water* **2**, 537–558 (2015).
16. Sakaguchi, K., Varughese, A. & Auld, G. Climate wars? A systematic review of empirical analyses on the links between climate change and violent conflict. *Int. Stud. Rev.* **19**, 622–645 (2017).
17. Koren, O. Food resources and strategic conflict. *J. Conflict Resolut.* **63**, 2236–2261 (2019).
18. Crost, B. & Felter, J. H. Export crops and civil conflict. *J. Eur. Econ. Assoc.* **18**, 1484–1520 (2020).
19. Linke, A. M. & Ruether, B. Weather, wheat, and war: security implications of climate variability for conflict in Syria. *J. Peace Res.* **58**, 114–131 (2021).
20. Von Uexkull, N., Croicu, M., Fjelde, H. & Buhaug, H. Civil conflict sensitivity to growing-season drought. *Proc. Natl. Acad. Sci.* **113**, 12391–12396 (2016).
21. Petrova, K. Natural hazards, internal migration and protests in Bangladesh. *J. Peace Res.* **58**, 33–49 (2021).
22. Maystadt, J. F. & Ecker, O. Extreme weather and civil war: does drought fuel conflict in Somalia through livestock price shocks? *Am. J. Agric. Econ.* **96**, 1157–1182 (2014).
23. Harari, M. & Ferrara, E. L. Conflict, climate, and cells: a disaggregated analysis. *Rev. Econ. Stat.* **100**, 594–608 (2018).
24. Hsiang, S. M., Burke, M. & Miguel, E. Quantifying the influence of climate on human conflict. *Science* **341**, 1235367 (2013).
25. Koubi, V. Climate change and conflict. *Annu. Rev. Polit. Sci.* **22**, 343–360 (2019).
26. Pearl, J. & Causality (Cambridge University Press, 2009).
27. Running, S. & Mu, Q. *MOD16A2 MODIS/Terra Evapotranspiration 8-day L4 Global 500m SIN Grid* (NASA LP DAAC, 2015).
28. Muñoz-Sabater, J. et al. Era5-land: A state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **13**, 4349–4383 (2021).
29. McNally, A. et al. A land data assimilation system for sub-Saharan Africa food and water security applications. *Sci. Data*. **4**, 1–19 (2017).
30. Gridded population of the world, version 4 GPWv4: population density. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4DZ068D (2015).
31. Huffman, G. Imerg v06 quality index. NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals For GPM (IMERG) Version 5.2 (2019).
32. Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* **5**, 1–12 (2018).
33. Hegre, H., Karlsen, J., Nygˇard, H. M., Strand, H. & Urdal, H. Predicting armed conflict, 2010–2050. *Int. Stud. Quart.* **57**, 250–270 (2013).
34. Hegre, H., Nygˇard, H. M. & Landsverk, P. Can we predict armed conflict? How the first 9 years of published forecasts stand up to reality. *Int. Stud. Quart.* **65**, 660–668 (2021).
35. Chadefaux, T. Conflict forecasting and its limits. *Data Sci.* **1**, 7–17 (2017).
36. Bhardwaj, J., Asghari, A., Aitkenhead, I., Jackson, M. & Kuleshov, Y. Climate risk and early warning systems: adaptation strategies for the most vulnerable communities. *J. Sci. Policy Gov.* **18** (2021).
37. Homer-Dixon, T. F. Environmental scarcities and violent conflict: evidence from cases. *Int. Secur.* **19**, 5–40 (1994).
38. Yoffe, S., Wolf, A. T. & Giordano, M. Conflict and Cooperation over international freshwater resources: indicators of basins at risk. *J. Am. Water Resour. Association.* **39**, 1109–1126 (2003).
39. Dinar, S., Katz, D., De Stefano, L. & Blankespoor, B. Climate change, conflict, and cooperation: global analysis of the effectiveness of international river treaties in addressing water variability. *Political Geogr.* **45**, 55–66 (2015).
40. Al-Muqdadi, S. W., Omer, M. F., Abo, R. & Naghshineh, A. Dispute over water resource management—Iraq and Turkey. *J. Environ. Prot.* **7**, 1096–1103 (2016).
41. Buhaug, H., Benjaminsen, T. A., Sjaastad, E. & Theisen, O. M. Climate variability, food production shocks, and violent conflict in sub-Saharan Africa. *Environ. Res. Lett.* **10**, 125015 (2015).
42. Bergholt, D. & Lujala, P. Climate-related natural disasters, economic growth,and armed civil conflict. *J. Peace Res.* **49**, 147–162 (2012).
43. Mach, K. J. et al. Climate as a risk factor for armed conflict. *Nature* **571**, 193–197 (2019).
44. Hsiang, S. M., Meng, K. C. & Cane, M. A. Civil conflicts are associated with the global climate. *Nature* **476**, 438–441 (2011).
45. Slettebak, R. T. Don't blame the weather! Climate-related natural disasters and civil conflict. *J. Peace Res.* **49**, 163–176 (2012).
46. Ghimire, R. & Ferreira, S. Floods and armed conflict. *Environ. Dev. Econ.* **21**, 23–52 (2016).
47. Bai, Y. & Kung, J. K.-S. Climate shocks and Sino-nomadic conflict. *Rev. Econ. Stat.* **93**, 970–981 (2011).
48. Ash, K. & Obradovich, N. Climatic stress, internal migration, and Syrian civil war onset. *J. Conflict Resolut.* **64**, 3–31 (2020).
49. Koubi, V., Bernauer, T., Kalbhenn, A. & Spilker, G. Climate variability, economic growth, and civil conflict. *J. Peace Res.* **49**, 113–127 (2012).
50. Xiao, L. B., Fang, X. Q. & Ye, Y. Reclamation and revolt: social responses in Eastern Inner Mongolia to flood/drought-induced refugees from the North China plain 1644–1911. *J. Arid Environ.* **88**, 9–16 (2013).
51. Hendrix, C. S. & Glaser, S. M. Trends and triggers: climate, climate change and civil conflict in sub-Saharan Africa. *Polit. Geogr.* **26**, 695–715 (2007).
52. Fjelde, H. & Von Uexkull, N. Climate triggers: rainfall anomalies, vulnerability and communal conflict in sub-Saharan Africa. *Polit. Geogr.* **31**, 444–453 (2012).
53. Chambers, R. L., Kennedy, H., Woods, J. E., Khadduri, M. & Blake, G. H. Iraq. *Encyclopedia Britannica.* https://www.britannica.com/place/Iraq (2024).

54. Tir, J. & Diehl, P. F. Demographic pressure and interstate conflict: linking population growth and density to militarized disputes and wars, 1930-89. *J. Peace Res.* **35**, 319–339 (1998).

55. Raleigh, C. & Hegre, H. Population size, concentration, and civil war. A geographically disaggregated analysis. *Polit. Geogr.* **28**, 224–238 (2009).

56. Reuveny, R. Climate change-induced migration and violent conflict. *Polit. Geogr.* **26**, 656–673 (2007).

57. Acemoglu, D., Fergusson, L. & Johnson, S. Population and conflict. *Rev. Econ. Stud.* **87**, 1565–1604 (2020).

58. Schermelleh-Engel, K. et al. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol. Res. Online.* **8**, 23–74 (2003).

59. Wills, P. & Meyer, F. G. Metrics for graph comparison: a practitioner's guide. *Plos One.* **15**, e0228728 (2020).

60. Bollobás, B. *Random Graphs.* In: *Modern Graph Theory.* Graduate Texts in Mathematics, vol. 184. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-0619-4_7 (1998).

61. Sagi, O. & Rokach, L. Explainable decision forest: transforming a decision forest into an interpretable tree. *Inform. Fusion.* **61**, 124–138 (2020).

62. Couttenier, M. & Soubeyran, R. Drought and civil war in sub-Saharan Africa. *Econ. J.* **124**, 201–244 (2014).

63. Bernauer, T. & Siegfried, T. Climate change and international water conflict in Central Asia. *J. Peace Res.* **49**, 227–239 (2012).

64. Urdal, H. People vs. Malthus: population pressure, environmental degradation, and armed conflict revisited. *J. Peace Res.* **42**, 417–434 (2005).

65. Malinsky, D. & Danks, D. Causal discovery algorithms: A practical guide. *Philos. Compass.* **13**, e12470 (2018).

66. Yin, W. Climate shocks, political institutions, and nomadic invasions in early modern East Asia. *J. Conflict Resolut.* **64**, 1043–1069 (2020).

67. Adano, W. R., Dietz, T., Witsenburg, K. & Zaal, F. Climate change, violent conflict and local institutions in Kenya's drylands. *J. Peace Res.* **49**, 65–80 (2012).

68. Water Peace and Security. *Water Challenges and Conflict Dynamics in Southern Iraq: An In-Depth Analysis of an Under-Researched Crisis* (2022). https://waterpeacesecurity.org/files/208

69. Bhattacharya, R., Malinsky, D. & Shpitser, I. Causal inference under interference and network uncertainty. In Adams, R. P. & Gogate, V. (eds.) *Proc. of The 35th Uncertainty in Artificial Intelligence Conference, vol. 115 of Proceedings of Machine Learning Research*, 1028–1038 (PMLR, 2020).

70. Hudgens, M. G. & Halloran, M. E. Toward causal inference with interference. *J. Am. Stat. Assoc.* **103**, 832–842 (2008).

71. Bhattacharya, R., Nagarajan, T., Malinsky, D. & Shpitser, I. Differentiable causal discovery under unmeasured confounding. In *International Conference on Artificial Intelligence and Statistics*, 2314–2322 (PMLR, 2021).

72. Lee, J. J., Bhattacharya, R., Nabi, R. & Shpitser, I. Ananke: A python package for causal inference using graphical models. Preprint at https://arXiv.org/230111477 (2023).

73. Ramsey, J. & Andrews, B. Py-Tetrad and RPy-Tetrad: A new Python interface with R support for Tetrad causal search. Proc. of the 2023 Causal Analysis Workshop Series, PMLR 223, 40–51 (2023).

74. Sharma, A. & Kiciman, E. DoWhy: An end-to-end library for causal inference. *Preprint at* https://arXiv.org/2011.04216 (2020).

75. Wolfson, M., Puri, A. & Martelli, M. The nonlinear dynamics of international conflict. *J. Conflict Resolut.* **36**, 119–149 (1992).

76. Runge, J. Causal network reconstruction from time series: from theoretical assumptions to practical estimation. *Chaos: Interdiscipl. J. Non-linear Sci.* **28**, 075310 (2018).

77. Assaad, C. K., Devijver, E. & Gaussier, E. Survey and evaluation of causal discovery methods for time series. *J. Artif. Intell. Res.* **73**, 767–819 (2022).

78. International Food Policy Research Institute (IFPRI). Global Spatially-Disaggregated Crop Production Statistics Data for 2020 Version 1.0. https://doi.org/10.7910/DVN/SWPENT (2024).

79. Glickman, T. S. *Glossary of meteorology* (American Meteorological Soc., 2000).

80. Chang, H. *Inventing temperature: Measurement and scientific progress* (2004).

81. Jin, M. & Dickinson, R. E. Land surface skin temperature climatology: benefitting from the strengths of satellite observations. *Environ. Res. Lett.* **5**, 044004 (2010).

82. Schon, J., Koehnlein, B. & Koren, O. The need for willingness and opportunity: analyzing where and when environmental variability influences conflict in the Sahel. *Popul. Environ.* **45**, 2 (2023).

83. Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688–701 (1974).

84. Rubin, D. B. Bayesian inference for causal effects: the role of randomization. *Annals Stat.* **6**, 34–58 (1978).

85. Rubin, D. B. Causal inference using potential outcomes: design, modeling, decisions. *J. Am. Stat. Assoc.* **100**, 322–331 (2005).

86. Holland, P. W. Statistics and causal inference. *J. Am. Stat. Assoc.* **81**, 945–960 (1986).

87. Robins, J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Modelling.* **7**, 1393–1512 (1986).

88. Spirtes, P., Glymour, C. N., Scheines, R. & Heckerman, D. *Causation, Prediction, and Search* 2nd edn (MIT Press, 2000).

89. Hitchcock, C. *Causality: Models, reasoning and inference* (2001).

90. Nichols, A. Causal inference with observational data. *Stata J.* **7**, 507–541 (2007).

91. Bongers, S., Forré, P., Peters, J. & Mooij, J. M. Foundations of structural causal models with cycles and latent variables. *Annals Stat.* **49**, 2885–2915 (2021).

92. Shpitser, I. & Pearl, J. Complete identification methods for the causal hierarchy. *J. Mach. Learn. Res.* **9**, 1941–1979 (2008).

93. Lundberg, I., Johnson, R. & Stewart, B. M. What is your estimand? Defining the target quantity connects statistical evidence to theory. *Am. Sociol. Rev.* **86**, 532–565 (2021).

94. Tian, J. & Pearl, J. (2002). A general identification condition for causal effects. In *Proceedings of the 18th National Conference on Artificial Intelligence (AAAI-02)*, 567–573.

95. Lamb, E. G., Mengersen, K. L., Stewart, K. J., Attanayake, U. & Siciliano, S. D. Spatially explicit structural equation modeling. *Ecology* **95**, 2434–2442 (2014).

96. Lamb, E., Shirtliffe, S. & May, W. Structural equation modeling in the plant sciences: an example using yield components in oat. *Can. J. Plant Sci.* **91**, 603–619 (2011).

Acknowledgements

Author contributions

Malekovic conceptualized the methodology, conducted the literature review, collected and engineered the data, and wrote the initial draft. Vonk implemented the methodology, conducted the data analysis, and contributed significantly to the manuscript's revision. Birkman, Sweijs, Kononova, and Bäck reviewed the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-90767-w.

Correspondence and requests for materials should be addressed to N.M. or M.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.