Combining Literature with Causal Discovery in Environmental-Conflict

Maarten Vonk[1,2]

[1] Natural Computing, Leiden Institute of Advanced Computer Science, Einsteinweg 55, Leiden 2333CC
[2] The Hague Centre for Strategic Studies, Organization, Lange Voorhout 1, The Hague, 2514EA, Netherlands, https://hcss.nl/

Abstract. As causal discovery algorithms have not matured up to the point where they can consistently identify complete causal structures from observational data, human expertise remains crucial in distinguishing cause from effect. In scientific practice, this human expertise is manifested in literature studies. This paper presents a causal discovery application of the environmental-conflict, conflict dynamic in Iraq where an extensive literature review is paired up with a causal discovery algorithm to uncover a causal mechanism. Specifically, the literature review identifies relevant variables for the causal dynamic and offers a hypothesized causal structure, which is then used as a benchmark to evaluate the causal structures derived from the causal discovery algorithms. Our proof of concept shows that such an approach can add specificity and refine our understanding of a causal mechanism.

Keywords: Causal Discovery · Human Expertise · Conflict Studies

1 Introduction

Environmental-conflict is a form of tension or violence triggered by competition over natural resources or environmental pressures, such as scarcity or degradation. Although environmental-conflict literature has advanced in examining the causal relationships between climatological factors and the onset of armed conflict, these studies often focus on specific connections rather than a comprehensive mechanism of naturally induced conflict [11, 24, 15]. The unsuitability of randomized controlled trials for the study of armed conflict further hampers the extraction of causal relations, leaving a gap in the literature for methodologies that can infer causality from non-experimental observations.

One way to address this gap is causal discovery. While causal discovery algorithms have gained traction over the years, research has shown that causal discovery algorithms can be unstable [14] or that only limited parts of the causal graph can be discovered from pure observations [19]. To this purpose, domain knowledge can be used to refine the performance of causal discovery algorithms and has been incorporated via tiered background knowledge [2] user interactions [16] or the penalization of the search process [9].

This extended abstract extracts a causal mechanism from environmental-conflict data in Iraq by complementing the power of a causal discovery method with hypothesized causal linkages from a literature review. First, a hypothesized causal mechanism is derived from a literature review, which informs the selection of variables for causal discovery. A causal discovery algorithm is then applied to the environmental-conflict data, and the resulting causal structure is benchmarked against the established linkages from the literature.

While the hypothesized causal linkages from the literature are discussed in Section 2, the data, causal discovery algorithm, and results are described in Section 3. Finally, the conclusions and limitations are discussed in Section 4.

2 Literature Review

This section formulates hypotheses on causal linkages contributing to the emergence of environmental conflict in Iraq, grounded in a comprehensive literature review that systematically examines regional environmental stressors and conflict dynamics identified across academic sources.

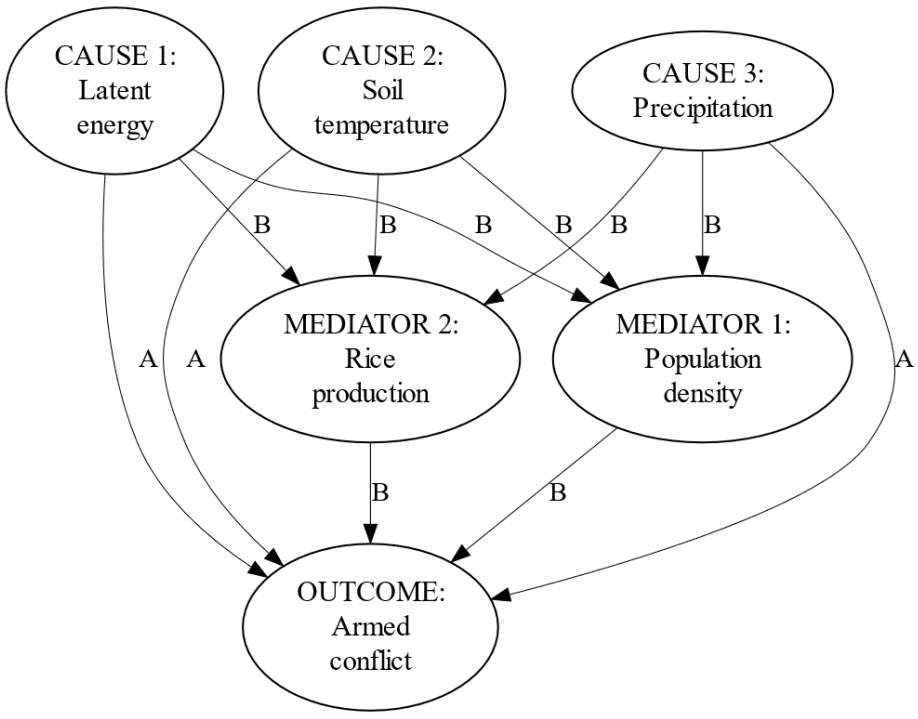

Fig. 1. The hypothesized causal structure: it distinguishes between direct and indirect linkages (i.e., paths A and B, respectively).

Long-term weather patterns have been directly linked to armed conflict [12, 11, 25, 8, 4]. Environmental changes can disrupt livelihoods, prompting fragmented communities to take action in an effort to mitigate the adverse effects. Such changes include variations in precipitation, evaporation, temperature, and the differential absorption or release of accumulated heat by physical environments, all of which may contribute to the emergence of conflict [11, 24].

On the other hand environmental variables have also been argued to cause armed conflict indirectly [3, 24, 15]. One possibility is that the effect of environmental changes on armed conflict are mediated by the scarcity of vital resources [24, 10]. This materializes into the scarcity of crops due to disrupted weather patterns, of which rice production has been shown to mediate causal effects of temperature on the emergence of actual violence [5].

Many studies also tie the emergence or existence of armed conflict in relation to environmental changes to population sizes [27, 20, 22]. Specifically, resource scarcity affects denser populations more significantly than less dense ones [3, 1], as denser populations might be less able to mitigate tensions effectively.

With all of the above hypotheses, it is possible to compose the entire hypothetical causal structure of linkages between environment and conflict, as shown in Figure 1. This hypothesized causal structure is also in line with research of other scholars [24].

3 Experimentation and Results

The units of analysis are all the 294 Iraqi municipalities. The literature study was leading in the choice of extracted variables. First, the number of civilian fatalities, conflict events, and conflict fatalities were sourced from Armed Conflict Location and Event Data Program [21] (ACLED). In addition, the climatological-related variables such as precipitation, soil temperature, and latent energy were retrieved from IMERG quality index [13], ERA5-Land dataset [18] and NASA MODIS dataset [23] respectively. Finally, we extracted rice production data from MapSPAM [26] and population density data from the International Earth Science Information Network at Columbia University [7]). Since the time horizon was from January 1, 2020, to January 1, 2022. we aggregated the values of explanatory variables, so much so each aggregation emphasized extreme values of the observed explanatory variables.

The causal discovery algorithm that is being applied to the data is GES [6]. Although the study would benefit from a more exhaustive use of causal discovery algorithms, GES has been chosen as it has been deemed suitable in the case of simulation studies with small sample sizes [17].

As the literature review pointed out, there are multiple subselections of variables that are hypothesized to play a role in the environmental-conflict dynamic. While no causal linkages or partial causal structure is imposed in the causal discovery phase, a total of 136 extracted graphs are benchmarked against the hypothesized graph of Figure 1. The graph that shares the most directed edges with the hypothesized graph is shown in Figure 2.

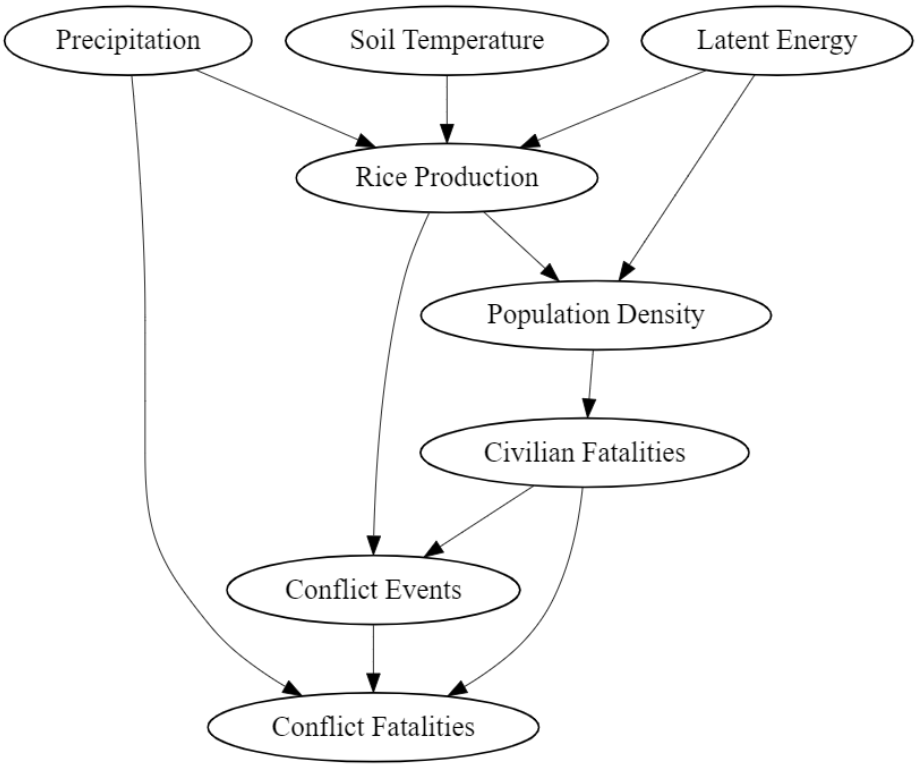

Fig. 2. GES-retrieved causal structure that is most aligned with the literature review

The literature review only hypothesized the possible existence of direct or indirect effects, which is why the causal discovery process reveals more specifically which directed arrows exist between the variables. The empirically retrieved causal structure reveals that precipitation is the only environmental variable with a direct effect on conflict, while other effects are mediated by rice production or population density.

4 Conclusion and Limitations

In this work, we have provided a proof of concept of how the strengths of both thorough literature review as well as causal discovery methods can be merged to achieve a causal mechanism in the environmental-conflict context of Iraq.

Since the results of causal discovery are filtered based on their alignment with existing findings, the method can only corroborate established conclusions and is unable to identify causal structures that are not already supported by the literature. Nonetheless, as the literature consensus can be ambiguous, pairing

them up with causal discovery algorithms can contribute significantly to refining our understanding of complex causal structures.

Finally, a limitation of our current approach to environmental-conflict modeling is that the causal discovery process assumes that the data is independent and identically distributed (i.i.d.). However, it is plausible that climatological variables of one municipality may influence the climate-conflict dynamics of another municipality, leading to violations of the i.d.d. assumption. This spatial confounding calls for causal discovery methods that can account for such violations [28].

References

1. Acemoglu, D., Fergusson, L., Johnson, S.: Population and conflict. The Review of Economic Studies **87**(4), 1565–1604 (2020)
2. Andrews, B., Spirtes, P., Cooper, G.F.: On the completeness of causal discovery in the presence of latent confounding with tiered background knowledge. In: Chiappa, S., Calandra, R. (eds.) Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 108, pp. 4002–4011. PMLR (26–28 Aug 2020), https://proceedings.mlr.press/v108/andrews20a.html
3. Ash, K., Obradovich, N.: Climatic stress, internal migration, and syrian civil war onset. Journal of Conflict Resolution **64**(1), 3–31 (2020)
4. Bai, Y., Kung, J.K.s.: Climate shocks and sino-nomadic conflict. Review of Economics and Statistics **93**(3), 970–981 (2011)
5. Caruso, R., Petrarca, I., Ricciuti, R.: Climate change, rice crops, and violence: Evidence from indonesia. Journal of Peace Research **53**(1), 66–83 (2016)
6. Chickering, D.M.: Optimal structure identification with greedy search. Journal of machine learning research **3**(Nov), 507–554 (2002)
7. CIESIN, S.: Gridded population of the world, version 4 (gpwv4): population density. Center for International Earth Science Information Network-CIESIN-Columbia University. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi. org/10.7927/H4DZ068D (2015)
8. Ghimire, R., Ferreira, S.: Floods and armed conflict. Environment and Development Economics **21**(1), 23–52 (2016)
9. Hasan, U., Gani, M.O.: Kcrl: A prior knowledge based causal discovery framework with reinforcement learning. In: Lipton, Z., Ranganath, R., Sendak, M., Sjoding, M., Yeung, S. (eds.) Proceedings of the 7th Machine Learning for Healthcare Conference. Proceedings of Machine Learning Research, vol. 182, pp. 691–714. PMLR (05–06 Aug 2022), https://proceedings.mlr.press/v182/hasan22a.html
10. Homer-Dixon, T.F.: Environment, scarcity, and violence. Princeton University Press (2010)
11. Hsiang, S.M., Burke, M., Miguel, E.: Quantifying the influence of climate on human conflict. Science **341**(6151), 1235367 (2013)
12. Hsiang, S.M., Meng, K.C., Cane, M.A.: Civil conflicts are associated with the global climate. Nature **476**(7361), 438–441 (2011)
13. Huffman, G.: Imerg v06 quality index. NASA, Available for download from: https://gpm. nasa. gov/sites/default/files/2020-02/IMERGV06_QI_0. pdf (2019)
14. Kitson, N.K., Constantinou, A.C.: The impact of variable ordering on bayesian network structure learning. Data Mining and Knowledge Discovery pp. 1–25 (2024)

15. Koubi, V.: Climate change and conflict. Annual Review of Political Science **22**(1), 343–360 (2019)
16. Mäkelä, J., Melkas, L., Mammarella, I., Nieminen, T., Chandramouli, S., Savvides, R., Puolamäki, K.: Technical note: Incorporating expert domain knowledge into causal structure discovery workflows. Biogeosciences **19**(8), 2095–2099 (2022). https://doi.org/10.5194/bg-19-2095-2022, https://bg.copernicus.org/articles/19/2095/2022/
17. Malinsky, D., Danks, D.: Causal discovery algorithms: A practical guide. Philosophy Compass **13**(1), e12470 (2018)
18. Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al.: Era5-land: A state-of-the-art global reanalysis dataset for land applications. Earth system science data **13**(9), 4349–4383 (2021)
19. Peters, J., Janzing, D., Schölkopf, B.: Elements of causal inference: foundations and learning algorithms. The MIT Press (2017)
20. Raleigh, C., Hegre, H.: Population size, concentration, and civil war. a geographically disaggregated analysis. Political geography **28**(4), 224–238 (2009)
21. Raleigh, C., Linke, R., Hegre, H., Karlsen, J.: Introducing acled: An armed conflict location and event dataset. Journal of peace research **47**(5), 651–660 (2010)
22. Reuveny, R.: Climate change-induced migration and violent conflict. Political geography **26**(6), 656–673 (2007)
23. Running, S., Mu, Q.: Mod16a2 modis/terra evapotranspiration 8-day l4 global 500m sin grid. NASA LP DAAC (2015)
24. Sakaguchi, K., Varughese, A., Auld, G.: Climate wars? a systematic review of empirical analyses on the links between climate change and violent conflict. International Studies Review **19**(4), 622–645 (2017)
25. Slettebak, R.T.: Don't blame the weather! climate-related natural disasters and civil conflict. Journal of Peace Research **49**(1), 163–176 (2012)
26. Statistics, G.S.D.C.P.: Data for 2010 version 2.0. IFPRI: Washington, DC, USA (2019)
27. Tir, J., Diehl, P.F.: Demographic pressure and interstate conflict: linking population growth and density to militarized disputes and wars, 1930-89. Journal of Peace Research **35**(3), 319–339 (1998)
28. Urdangarin, A., Goicoa, T., Ugarte, M.D.: Evaluating recent methods to overcome spatial confounding. Revista Matemática Complutense **36**(2), 333–360 (2023)