

EM-DAT Documentation

[Documentation](#) [EM-DAT Project](#) [Public Data](#) [HDX](#) [GitHub](#) [About](#)


Known Issues and Limitations


What Difficulties May Be Encountered While Interpreting EM-DAT Data?

- 1: [General Issues](#)
- 2: [Specific Biases](#)

EM-DAT is the only comprehensive, free-access disaster loss database with effective global coverage¹. However, it has limitations due to the limited number of sources and limitations related to how effectively disasters are reported worldwide. This can lead to biases in the data over which CRED may have limited control, and that could be overlooked in the literature². Nevertheless, EM-DAT remains a key resource for understanding disaster events and impacts. No current impact database is completely accurate. The United Nations emphasizes the importance of global improvements in documenting disasters in global agendas such as the [SENDAI Framework for Disaster Risk Reduction \(SFDRR\)](#).

Understanding the limitation of a dataset such as EM-DAT is of paramount importance for those who wish to adequately use the data and mitigate its weaknesses for the following purposes: disaster risk management, emergency planning, scientific research, and raising public awareness.

1. Mazhin, S. A., Farrokhi, M., Noroozi, M., Roudini, J., Hosseini, S. A., Motlagh, M. E., Kolivand, P., and Khankeh, H.: Worldwide disaster loss and damage databases: A systematic review, *J Educ Health Promot*, 10, 329, https://doi.org/10.4103/jehp.jehp_1525_20, 2021. 

2. Jones, R. L., Kharb, A., and Tubeuf, S.: The untold story of missing data in disaster research: a systematic review of the empirical literature utilising the Emergency Events Database (EM-DAT), *Environ. Res. Lett.*, 18, 103006, <https://doi.org/10.1088/1748-9326/acfd42>, 2023. 

1 - General Issues

Understanding Broad Data Quality Concerns in EM-DAT

Three types of data quality issues can be considered:

Types of Data Quality Issues

1. Disaster events that are missing in EM-DAT.
2. Disaster events that exist but that have missing values, e.g., for the impact variables.
3. Disaster events that are well documented but with attributes that are inaccurate or differ from other sources.

A cross-comparison of EM-DAT with a local database and/or a disaster-specific database can help identify Issue 1 (e.g., Koç & Thieken, 2018¹; Lin et al., 2021²). For an account of missing values for existing events, we refer to Jones et al. 2021³ and the section on [Accounting Biases](#). Issue 3 is partially related to the data collection sources, protocols, or reporting systems generally used by different databases.

Data quality issues within EM-DAT are related to the [data collection protocols](#) from [dedicated sources](#). EM-DAT's completeness reflects the coverage of its sources. Since source reporting has improved over the years, [EM-DAT data coverage has improved significantly over the last 30 to 40 years](#). Nevertheless, gaps and quality issues remain. EM-DAT protocols are meant to guide the way information is monitored and collected from sources. However, no universally applied protocol ensures that different sources report disaster impact and losses using the same guidelines to define, for instance:

- the beginning and end of disaster events.
- the geographical footprint of a disaster.
- impact variables such as deaths (in particular, when computed based on excess mortality), affected people, or economic costs.
- the disaster type selected by the sources.

Some references illustrate the issues and challenges related to collecting and maintaining a disaster database, e.g., Guha-Sapir & Misson 1992⁴, Kron et al. 2012⁵, and Wirtz et al. 2014⁶.






To some extent, EM-DAT owes its popularity to its simplicity. It reports disaster events as rows in an Excel table. However, this simplicity comes at the cost of conceptual limitations in dealing with complex and compound events and situations. In such cases, as exemplified in the box below, EM-DAT will probably report the disaster in the same way as the source which presented it. The EM-

DAT database manager can only choose to select some numbers over some others (see [Daily Encoding](#)). However, no model is involved in correcting differences in reporting protocols because this task goes beyond the information monitoring conducted at the CRED by the EM-DAT team.

Fictive Example of Disaster Complexity

- If a source reports a heatwave with a certain number of deaths, EM-DAT is likely to record it as such.
- If the same heatwave hits a neighboring country, whose institutions have different reporting protocols, EM-DAT will also report the disaster entry based on the source's numbers.
- Since the protocols are different, this will create a systemic bias in EM-DAT.
- The event duration may be misaligned; some may have accounted for co-occurring effects, such as droughts, wildfires, and air pollution, in the estimation of the loss statistics (e.g., deaths, affected people, or costs).
- In some cases, the main type could even be different. In databases other than EM-DAT, the event and the numbers may also have another representation.

Such biases that result from differences in the impact reporting systems were generally referred to by Gall et al. 2009⁷ as systemic biases. Some studies point to systemic biases by highlighting that EM-DAT does not correlate well with other databases (e.g., Moriyama et al., 2018⁸; Panwar & Sen, 2020⁹). In their article, Gall et al. 2009⁷ cover four other types of biases: time, hazard-related, spatial, and accounting biases. These are illustrated in the next sections.

-
1. Koç, G., and Annegret H. T. "The Relevance of Flood Hazards and Impacts in Turkey: What Can Be Learned from Different Disaster Loss Databases?" *Natural Hazards* 91, No. 1 (2018): 375408. <https://doi.org/10.1007/s11069-017-3134-6>. 
 2. Lin, Y. C., Khan, F., Jenkins, S. F. and Lallemand, D. "Filling the Disaster Data Gap: Lessons from Cataloging Singapore's Past Disasters." *Int. J. Disaster Risk Sci.* 12, 188–204 (2021). <https://doi.org/10.1007/s11069-021-00331-z>. 
 3. Jones, R. L., Guha-Sapir, D., and Tubeuf, S.: "Human and economic impacts of natural disasters: can we trust the global data?", *Sci Data*, 9, 572 (2022). <https://doi.org/10.1038/s41597-022-01667-x>. 
 4. Guha-Sapir, D. and Misson, C.: "The Development of a Database on Disasters.", *Disasters*, 16, 74–80 (1992), <https://doi.org/10.1111/j.1467-7717.1992.tb00378.x>. 
 5. Kron, W., Steuer, M., Löw, P., and Wirtz, A. "How to Deal Properly with a Natural Catastrophe Database – Analysis of Flood Losses." *Natural Hazards and Earth System Sciences* 12, No. 3, 53550 (2012). <https://doi.org/10.5194/nhess-12-535-2012>. 

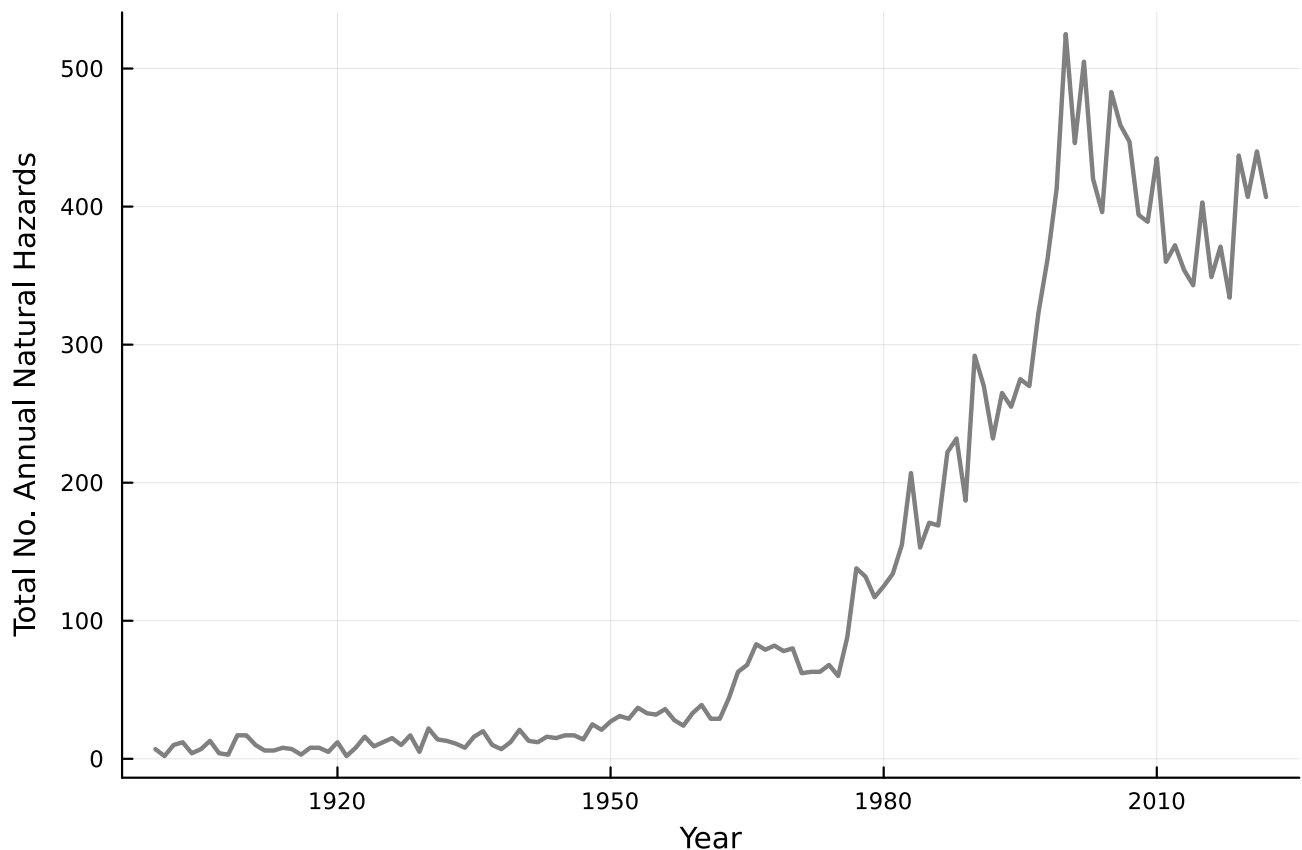
6. Wirtz, A., Kron, W., Löw, P., and Steuer, M. "The Need for Data: Natural Disasters and the Challenges of Database Management". *Natural Hazards* 70, No. 1, 13557 (2014). <https://doi.org/10.1007/s11069-012-0312-4>. 
7. Gall, M., Kevin A. B., and Susan L. C. "When Do Losses Count?: Six Fallacies of Natural Hazards Loss Data." *Bulletin of the American Meteorological Society* 90, No. 6, 799810 (2009). <https://doi.org/10.1175/2008BAMS2721.1>.  
8. Moriyama, K., Daisuke S., and Yuichi O. "Comparison of Global Databases for Disaster Loss and Damage Data." *Journal of Disaster Research* 13, No. 6, 100714 (2018). <https://doi.org/10.20965/jdr.2018.p1007>. 
9. Panwar, V. and Subir S. "Disaster Damage Records of EM-DAT and DesInventar: A Systematic Comparison." *Economics of Disasters and Climate Change* 4, No. 2, 295317 (2020). <https://doi.org/10.1007/s41885-019-00052-0>. 

2 - Specific Biases

Understanding Particular Data Quality Concerns in EM-DAT

Time Bias

Time biases result from unequal reporting quality and coverage over time¹. The figure below shows the occurrence of disasters in EM-DAT. The figure shows a significant increase that starts in the 1960s. This increase coincides with the creation of OFDA. In 1973, OFDA started compiling disaster data, and the CRED was created². In 1988, the CRED took over the disaster database and created EM-DAT. In the meantime, communication technologies have improved, with the first personal computers and satellites appearing in the 1970s and the advent of the World Wide Web in the 1990s (see also [History of EM-DAT](#))



Occurrence of EM-DAT Disasters Related to Natural Hazards (at the Country Level) from 1900 to 2022

Technologies and initiatives can be considered responsible for the dominant trend observed. Therefore, it is challenging to infer insight into the actual drivers of disasters such as climate change, population growth, or disaster risk management³. Accordingly, excluding pre-2000 data

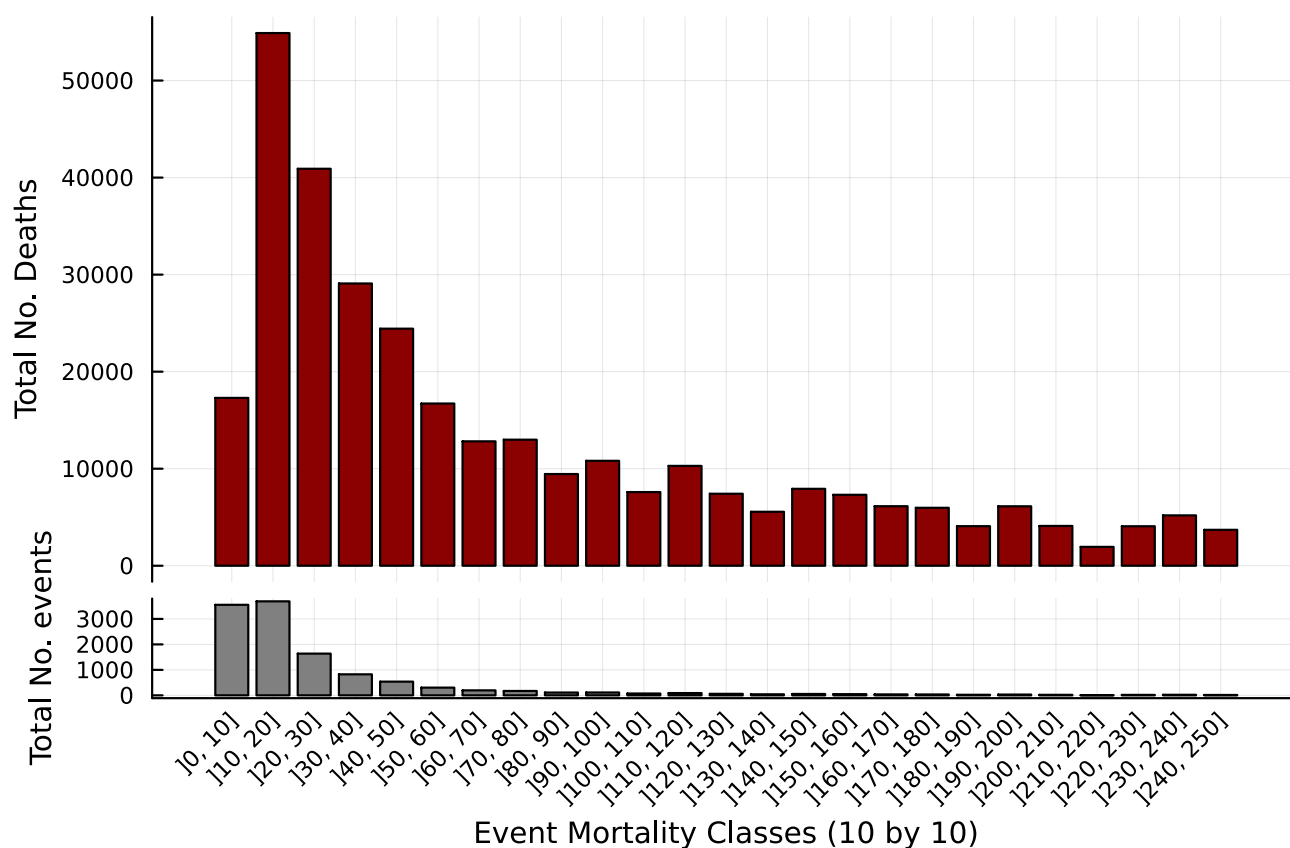
from trend analyses based on EM-DAT is strongly recommended. From September 2023 onward, the CRED refers to pre-2000 data as `Historic` data in the [EM-DAT Public Table](#).

Hazard-related Biases

Hazard-related biases result from unequal reporting quality and coverage for different hazard types¹. For example, in EM-DAT, data related to biological hazards (e.g., epidemics) and extreme temperature hazards (e.g., heat waves) are less covered and the cover of lower quality. Some hazard-related biases are illustrated in the [Accounting Biases](#) and [Geographic Biases](#) sections.

Threshold Biases

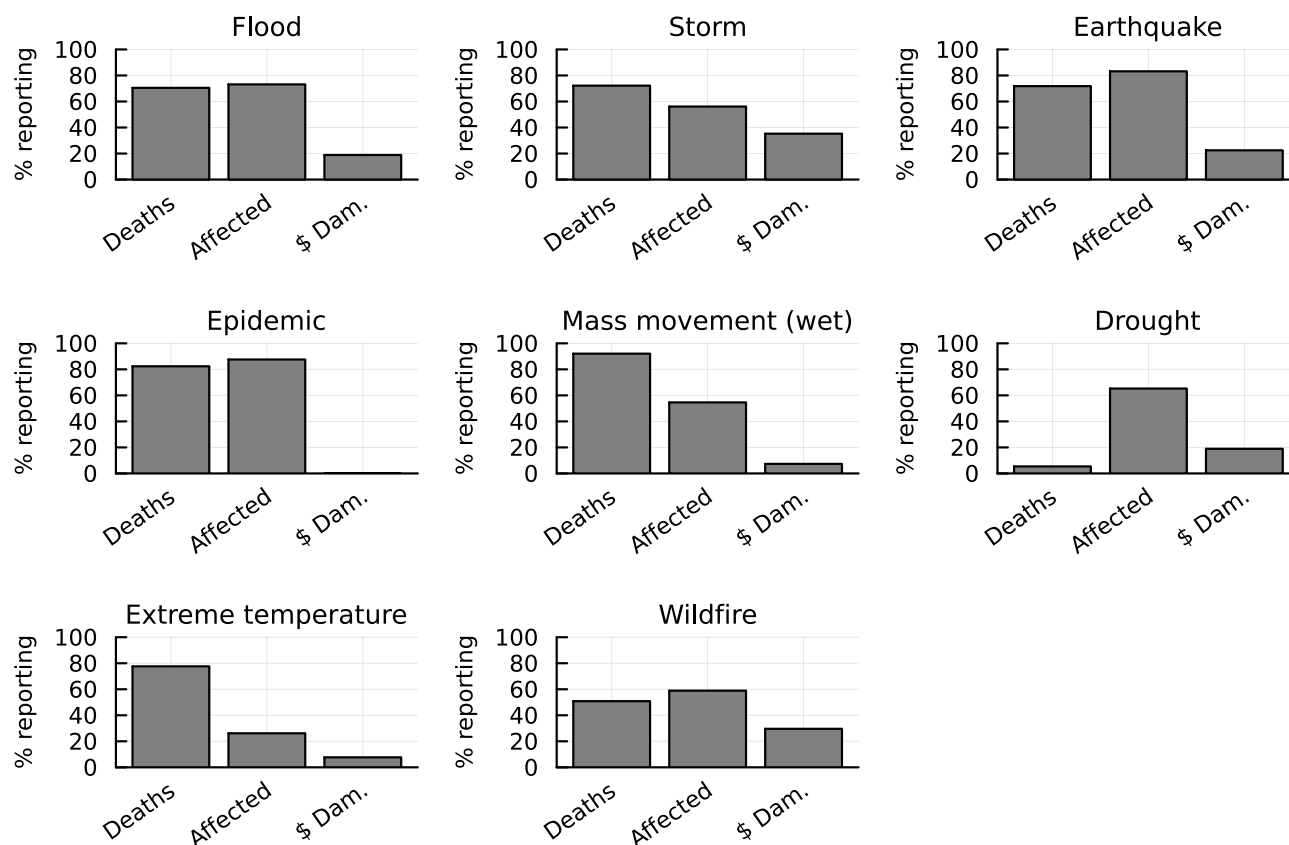
Threshold biases result from unequal reporting quality and coverage for disasters of different magnitudes¹. High-impact disasters attract more attention, resulting in better media coverage and reporting. This could lead to threshold biases in EM-DAT. The EM-DAT entry criteria introduce a kind of threshold bias, as shown in the figure below regarding disaster mortality, while some studies have shown locally that small disasters may have a high cumulative impact, e.g.⁴. Regarding disasters that fit EM-DAT's entry criteria, it is fair to assume that disasters close to the entry criteria are more likely to be missing than high-impact disasters. However, as shown in the figure below, the cumulative mortality associated with low-mortality events exceeds the cumulative impact of higher-mortality events.



EM-DAT Cumulative Mortality per Disaster Event Mortality Classes. One class regroups events having between x (not included) and $x+10$ deaths (included). Period covered: 2000-2022.

Accounting Biases

Accounting biases result from unequal reporting quality and coverage for different impact variables¹. For instance, in EM-DAT, the economic losses are, on average, less frequently reported than the human impact variables, which may also depend on the hazard type (see the figure below). Furthermore, insured damages are naturally more reported than uninsured damages, which creates a geographic bias where there is a lack of insurance coverage, as in Africa. For droughts, EM-DAT fails to capture the associated mortality because it is overlooked as an indirect impact, as evidenced by the [UNDRR 2021 Global Assessment Report on Droughts](#).



Percentage of Reporting from Sources for Three Impact Variables ('Total Deaths', 'Total Affected', and 'Total Damage') in Natural Hazards Types with more than 500 Sources. Period covered: 2000-2022.

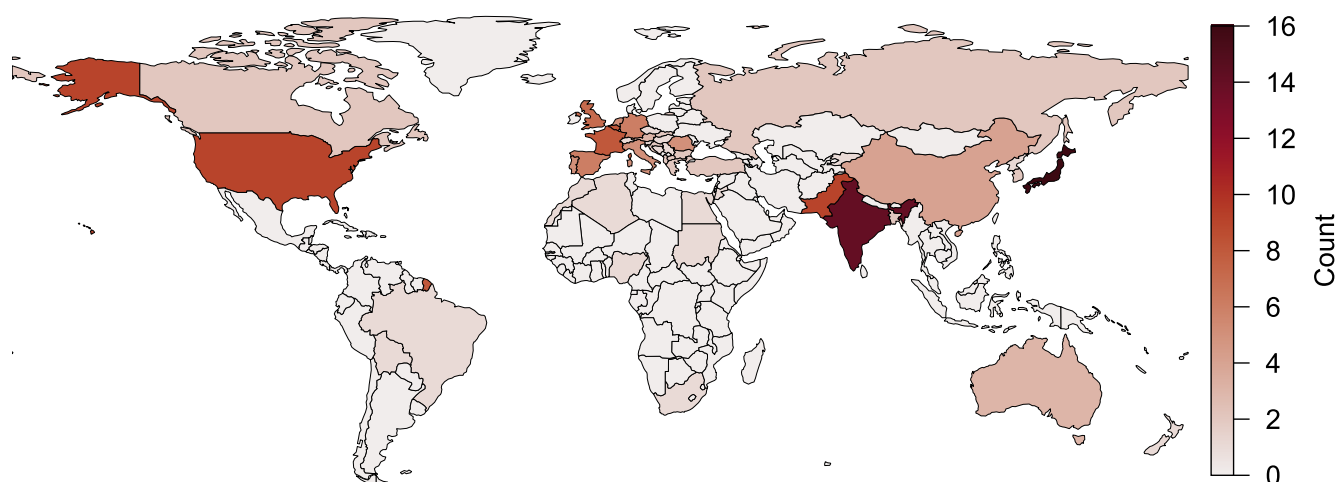
Besides, it cannot be assumed that because an impact is reported in EM-DAT, there is no accounting bias. In general, direct impacts are often reported by EM-DAT sources, while indirect impact estimates are less available. For instance, indirect deaths for a flood event correspond to the number of fatalities occurring during the event, while indirect deaths result from disease outbreaks due to deteriorated sanitary conditions. Yet, indirect mortality is sometimes more important than direct mortality, for examples.⁵

Geographic Biases

Geographic biases result from unequal reporting quality and coverage across space¹. In general, EM-DAT has a relatively worse coverage for Sub-Saharan Africa regarding the occurrence and the accounting of impact variables⁶. Any disaster type may be subject to geographic biases in EM-DAT as there may be discrepancies between reporting systems from one country to another (see [General Issues](#)).





This issue is particularly pronounced regarding heat waves⁷, as shown in the figure below. Heat waves are often overlooked in Sub-Saharan Africa⁸. About 52% of heatwave events in EM-DAT

occurred in nine countries: Japan, India, Pakistan, the USA, followed by Western European countries (France, Belgium, United Kingdom, Spain, and Germany).



Number of Heat Waves in EM-DAT (2000-2022)

References

1. Gall, M., Kevin A. B., and Susan L. C. "When Do Losses Count?: Six Fallacies of Natural Hazards Loss Data." *Bulletin of the American Meteorological Society* 90, No. 6, 799-810 (2009).
<https://doi.org/10.1175/2008BAMS2721.1>.     
2. CRED "Happy Birthday, CRED: Celebrating 50 Years of Disaster Epidemiological Research, Data Collection, and International Cooperation", Centre for Research on the Epidemiology of Disasters (CRED), Brussels, Belgium, *CRED Crunch* No. 71 (2023), <https://www.cred.be/sites/default/files/CredCrunch71.pdf>. 
3. Ritchie, H., Rosado P. "Is the number of natural disasters increasing? A deep dive into missing data and the limitations of disaster databases", <https://ourworldindata.org/disaster-database-limitations>. 
4. Marulanda, M. C., Cardona, O. D., and Barbat, A. H. "Revealing the socioeconomic impact of small disasters in Colombia using the DesInventar database", *Disasters*, 34, 552-570 (2010), <https://doi.org/10.1111/j.1467-7717.2009.01143.x> 
5. Alderman, K., Turner, L. R., and Tong, S. "Floods and human health: A systematic review", *Environment International*, 47, 37-47 (2012) <https://doi.org/10.1016/j.envint.2012.06.003>. 
6. Osuteye, E., Johnson, C., and Brown, D. "The data gap: An analysis of data availability on disaster losses in sub-Saharan African cities", *International Journal of Disaster Risk Reduction*, 26, 24-33 (2017), <https://doi.org/10.1016/j.ijdr.2017.09.026>. 
7. Brimicombe, C., Di Napoli, C., Cornforth, R., Pappenberger, F., Petty, C., and Cloke, H. L. "Borderless Heat Hazards With Bordered Impacts", *Earth's Future*, 9, e2021EF002064 (2021), <https://doi.org/10.1029/2021EF002064>. 
8. Harrington, L. J. and Otto, F. E. L. "Reconciling theory with the reality of African heatwaves", *Nat. Clim. Chang.*, 10, 796-798 (2020), <https://doi.org/10.1038/s41558-020-0851-8>. 

