All About that Base

Are Disaster Impacts Getting Worse?

Hamish Patten and Justin Ginnetti

Disaster and Climate Data

By analysing the historical trends of hazards such as floods and tropical cyclones, there is a general, worldwide consensus in the climate research community that climate change is influencing the frequency and intensity of climate and weather-related hazards[1]. Therefore, if such hazards are increasing in frequency and intensity, is this also measurable in terms of their impact on the planet? It is important to also be able to estimate the effects of climate change with respect to the impact that climate and weather-related hazards are having on the population, physical infrastructure, environment and economy.

To measure whether disaster impacts are getting worse, we need data. We need data that has good global coverage. That means data that doesn't just focus on certain countries, such as insurance company data, often existing uniquely in rich countries. The data must have good coverage in time. That means data in which we can see trends and try to understand whether disaster impacts are increasing with time due to climate change. The data must also include many different hazards, climate change related and otherwise. If a database covered only earthquakes (which are not influenced by climate change), or hazards that occur predominantly in some countries but not others (such as cyclones or wildfires), then we may not be able to accurately infer climate change effects from the data. These three examples of the requirements we need for our data are known as biases: spatial, temporal (time-based) and hazard-based bias. There are six forms of hazard-impact data bias in total, please refer to F. Wyatt *et al*, 2023 for more information[2].

The good news is that there are many different, long-standing databases that exist, on a global level, that provide curated estimates of the impacts that hazards have on the population, physical infrastructure, environment and economy. However, each one has its own benefits and limitations. The main limitations of the different databases are mainly due to the bias in the data. For example, at the International Federation of Red Cross and Red Crescent Societies (IFRC) we have a database of historical Emergency Appeals. An Emergency Appeal is when the impact of a hazard is so severe that the local Red Cross or Red Crescent society makes an appeal for international support, which is then coordinated by IFRC. One reason why this database is not ideal to use as a single source of information on climate change-related hazard impacts is that it covers mostly low and mid-GDP countries.

---

[1] For evidence to support this statement, the authors recommend the reader refer to any of the most recent Intergovernmental Panel on Climate Change (IPCC) Assessment Reports, such as can be found here: https://www.ipcc.ch/report/sixth-assessment-report-working-group-i

[2] F. Wyatt et al. '*Investigating bias in impact observation sources and implications for impact-based forecast evaluation'* International Journal of Disaster Risk Reduction, volume *90 number 103639, 2023.*

The Montandon Database

With its extensive global network comprising 192 National Societies, IFRC has responded to a multitude of disasters, accumulating a wealth of knowledge and insights. As the world's largest disaster response network, IFRC possesses the capability to capture and learn from an array of crises, both large and small. The Montandon - Global Crisis Data Bank (which we refer to here as the Montandon) is a pioneering database currently being developed by IFRC. The aim of the Montandon is to consolidate historical hazard and disaster-impact data from diverse organisations into a single repository. The name 'Montandon' comes from Raoul Montandon, who in 1923 wrote an article in the 'International Review of the Red Cross'[3] that identified a need for a global, geospatially referenced database in order to facilitate learning and to reduce the impacts of future disasters. IFRC's expertise in collecting, monitoring, and curating disaster risk information, combined with its unmatched experience in operational disaster response, positions it as one of the premier organisations for this vital role. The Montandon, under IFRC's stewardship, will facilitate knowledge sharing, foster peer-to-peer learning, and enable evidence-based decision-making. This database has the potential to revolutionise disaster risk reduction and management, empowering stakeholders with comprehensive insights to anticipate, respond to, and recover from crises more effectively.

By gathering together and unifying different databases, and knowing the benefits and limitations of each source, it is possible to fill in the data gaps and reduce the different types of bias. However, the difficulty is that the different databases weren't designed to fit together. When two databases align with one another and do not need a layer of translation or interpretation to go from one to the other, they are said to be interoperable. Developing and validating this layer of translation, to create interoperability between the different databases, is a key element of the work that is involved in building the Montandon database.

Climate Change Effects in the Montandon

Inferring the influence of climate change on hazard impacts is unfortunately not a straightforward procedure. This is due to the fact that, as mentioned above, the different sources of impact data have different forms of bias and data limitations. The most important limitation that is apparent in all of the databases is that the number of events recorded in each database will always underestimate the total actual number of events that occur.

'*If a tree falls in a forest and no one is around to hear it, does it make a sound?*'

To explain further, the above quote allows us to introduce the concept of reporting bias: one of the most prominent source of bias across all the databases. If a flood occurs, yet nobody is affected, no buildings damaged, no direct or indirect impediment to the economy, and so on, the likelihood of the event being recorded reduces significantly. Furthermore, even if there is some kind of measurable impact, such as fatalities, if this value isn't high enough, then it might not capture the government or the media's attention, and may still not be

---

[3] Montandon, Raoul, 1923. "A propos du projet Ciraolo. Une carte mondiale de distribution géographique des calamités," Revue Internationale de la Croix-Rouge, No. 52, April 2023. https://international-review.icrc.org/fr/articles/propos-du-projet-ciraolo-une-carte-mondiale-de-distribution-geographique-des-calamites

reported. The larger the impact, the more likely the event is to be reported and thus the more likely that impact estimates for that event are to be included in the database.

The different databases included in the data analysed in this report are as follows: Desinventar, hosted by United Nations Disaster Risk Reduction (UNDRR), the Emergency Events DATabase (EM-DAT), hosted by Centre for Research on the Epidemiology of Disasters (CRED) within the Université catholique de Louvain (UCLouvain), Global Internal Displacement Database (GIDD), hosted by the Internal Displacement Monitoring Centre (IDMC), GLobal unique disaster IDEntifier number (GLIDE), hosted by the Asian Disaster Reduction Center (ADRC) and the Emergency Appeals database (abbreviated to GO-App in the plots), hosted by IFRC.

Figure 1: The yearly number of hazard events with recorded impacts, per database. A trend line has been added to facilitate interpretation (using the LOESS model). Note the y-axis scales are different between each of the different databases.

The number of impacts recorded for each of these five databases is more directly related to the capacity of the organisation than it is with the actual number of hazards occurring. Figure 1 shows the number of recorded events by each of the different databases, reflecting the different, non-linear trends between how active each of the databases are in reporting hazard impacts, per year. If we were to assume that any trend present here is related to climate change effects, then the (incorrect) conclusion from both of the Desinventar and GLIDE databases would be that climate change has been reducing the number of impacts, on a global level, since the year 2010 and 2000, respectively. This is not actually the case, because actually the databases are just being less and less populated with events in recent years. For Desinventar, for example, it is because governments are creating their own databases and are no longer relying on the Desinventar database to input their data. Further to this, the GIDD database has been growing exponentially in recent years due to the significant increase in organisational capacity. This has resulted in a larger number of staff available to curate the impact estimates that are then integrated into the database.

Therefore, what might appear as an increasing trend in the number of events does also not necessarily coincide with the influence of climate change.

In this analysis, climate change effects are inferred from the data by looking at the proportion (fraction) of climate- and weather-related hazard events as compared to other natural hazards. The definition of 'climate and weather-related hazards' is any hydro-meteorological hazard from the 'Hazard definition and classification review: Technical report', produced by UNDRR and the International Science Council (ISC)[4]. Furthermore, we define natural hazards as any hydro-meteorological or geophysical hazards, again, as defined by the UNDRR-ISC 2020 report cited above. The proportion of climate and weather-related hazards, as compared to all natural hazards, is used in this analysis to infer the influence of climate change. We could instead have tried to infer a relationship between climate change and the number of events recorded, but this was shown previously to be convoluted with other issues, such as the number of events being underreported due to organisation-dependent reporting bias.

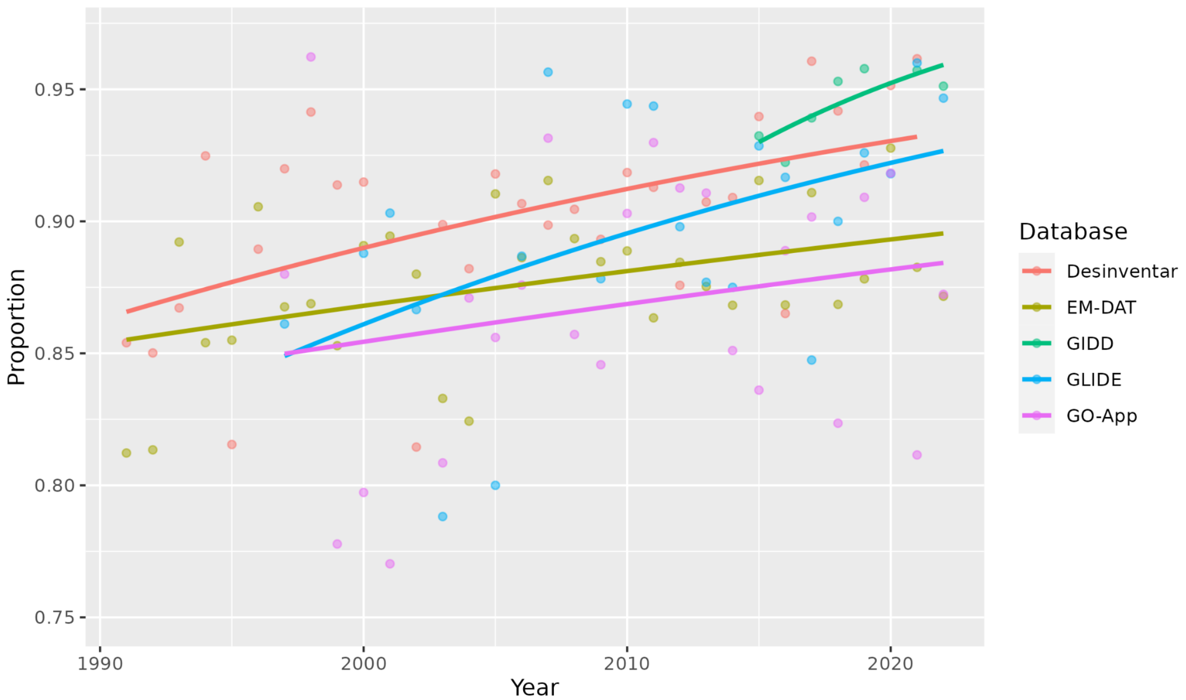

Figure 2: The proportion of climate- and weather-related events, per year, as compared to all natural hazards. The different colours correspond to different databases. Trend lines have been added to the figure to facilitate the interpretation of potential climate change effects (using binomial regression).

In figure 2, we show the proportion of natural hazards that are climate- and weather related, per year and per database. The actual proportion will vary from database to database due to the various different reporting mechanisms and biases. For example, some have different spatial or temporal biases, some use different threshold mechanisms for when they do and do not include an event in their database, and some will include certain hazards that are not included in others. All such biases result in different proportions. In order to provide an

---

[4] United Nations Office for Disaster Risk Reduction and International Science Council, *'Hazard definition and classification review: Technical report'*, 2020. https://www.undrr.org/publication/hazard-definition-and-classification-review-technical-report

estimate of the actual proportion in 2022, we use the GIDD database, due to the low spatial bias and high number of events recorded. Furthermore, by using only the GIDD database, we ensure scientific reproducibility of our estimate without requiring an advanced knowledge of statistics and statistical modelling.

In order to state that the effects of climate change are present in the impact data, across the different databases, a rigorous analysis must be conducted. Evidence must be provided to either support or reject the claim that there is an increasing trend in time of the proportion of climate and weather-related hazards that have a measurable impact, on a global level. We do this by assuming that all databases have a common trend in time, but that they have different proportions. This is to state that the increase between any two given years in the proportion is assumed to be the same between databases, but that the actual proportions will be different. We only use this common trend line assumption to test to see whether the increasing proportions are just related to all databases or only to a few. If all databases have increasing trends, then this provides more evidence that the trend is a general observation, independent of the database. Furthermore, because we are dealing with proportions and not counts, we make sure that the type of model we use cannot predict values above 1 or below 0, by definition of a 'proportion'. For those who are more familiar with statistical modelling, the above described model is a mixed-effects binomial regression. The most important factor in this analysis is to ensure that this common trend line is positive (such that climate change is increasing the proportion and not decreasing it in time) and not equal to zero. If the trend line could potentially be equal to zero, then the data does not provide evidence that climate change is exacerbating hazard impacts. To infer whether the data provides evidence for such a hypothesis, we use something called a statistical hypothesis test on the trend line. For more information about this, an example blog post can be found here[5]. What is important in a hypothesis test is the so called 'p-value'[6]. If it is less than 0.01, then the evidence is said to be highly statistically significant or to have 3-sigma significance, which is what was found for the trend line in this analysis. For a more detailed explanation of 3-sigma significance, please see here[7]. We note that the use of cumulative totals would have been more accurate in modelling the proportions. However, there were no significant differences observed in the two different methods, and using the yearly-aggregated proportion calculation was estimated to be something more familiar to the readers, to ensure reproducibility of the analysis.

Now that we have provided evidence to support the hypothesis that the proportion of climate- and weather-related hazards, as compared to all natural hazards, is increasing in time, we can also use the model to produce a more reliable estimate of the current proportion. This is because estimating the value directly from the data leads to what is called a 'noisy' estimate: the proportion as estimated directly from the data (by the scattered points visible in figure 2) will not be a perfect line because there is some randomness between years: some events such as large-scale floods or earthquakes could have occurred but didn't. Additionally, there may have been many events that were not reported because they 'slipped through the net'. For example, an event may have only been reported in a local language and therefore was undetected and thus left out of the impact estimate database,

---

[5] https://medium.com/analytics-vidhya/hypothesis-testing-for-dummies-5903cff6e82d

[6] https://www.dummies.com/article/academics-the-arts/math/statistics/how-to-determine-a-p-value-when-testing-a-null-hypothesis-169062/

[7] https://www.indeed.com/career-advice/career-development/3-sigma

by a given organisation. By looking directly at the trend-line, we can have a more stable estimate of what the proportion was during 2022, or in 2000, or we can even make estimates of what the proportion might be in the future. As we are using multiple databases, we could also make an estimate that combines each database, or uses only a single database to estimate the proportion. Given that the GIDD database has the largest number of entries for 2022 over any other database, given the bias and limitations of some of the other databases, and given that using only one database to make the estimate is more easily reproducible, the GIDD database is used here to make the estimate of the proportion. We have analysed the GIDD data, looking at each disaster event and extracted whether it is climate/extreme weather related or not. Having done that for every event, we can say with confidence that, between the years 2018-2022, the proportion was 94%. Furthermore, this proportion is estimated to have increased by 3.5% since 1990. Note that the change in time of the proportion is assumed to be independent of the database, as we assumed a common trend line in our previous analysis. It is important to note here that because the limits of a proportion are between 0 and 1, as previously mentioned, we chose to use a model that accounts for this. Therefore, the difference we observe between 1990 and 2022 of 3.5% is only for this given period and proportion value: the closer the proportion gets to 1, the difference will start to reduce, ensuring that the proportion never exceeds 1 (the benefit of using binomial regression).

Finally, we can also extract some key figures for 2022 from the different databases. There were more than 1,500 disasters triggered by natural hazards in 2022, more than 1,400 of which were climate and weather-related hazards. It is estimated that more than 76,000 were killed by natural hazards in 2022, according to the EM-DAT database. More than 74,000 of these deaths were attributed to the heat wave that occurred in Europe. Such hazards are strongly exacerbated by the effects of climate change. Further to this, more than 32 million people were displaced due to natural hazards, with more than 19 million of these people displaced by floods alone, according to the GIDD database.

Conclusion

In this report, the Montandon database was introduced: a database that pulls together hazard and hazard-impact estimate data from a range of international sources into one unified and standardised database. The Montandon is being developed and hosted by IFRC and in collaboration with several partnering international organisations. This database has been used in this report to infer the effects of climate change from hazard-impact data from a range of sources, both external and internal to IFRC. By using multiple sources, further evidence can be provided to support claims about the presence of climate change effects in hazard-impact trends. The report explains the difficulties that are inherent to hazard-impact databases through the concept of bias, such that different forms of bias are present in the different databases and must be accounted for when analysing the data.

Due to reporting bias, where the yearly number of events reported was shown to be more linked to organisational capacity than any other trends, this analysis focussed on inferring climate change effects by looking at the proportion of climate and weather-related hazards as compared to all natural hazards. Evidence from the data is found to support the hypothesis that there is an increasing trend in time of this proportion. Therefore, the evidence suggests that the proportion of climate- and weather-related hazards that have impacts on the population, physical infrastructure, environment or the economy is increasing per year. This proportion is estimated to have increased by 3.5% between 1990 to 2022, assumed due to climate change. Furthermore, this proportion, as inferred from the GIDD database, is estimated to be 94% between 2018 to 2022.