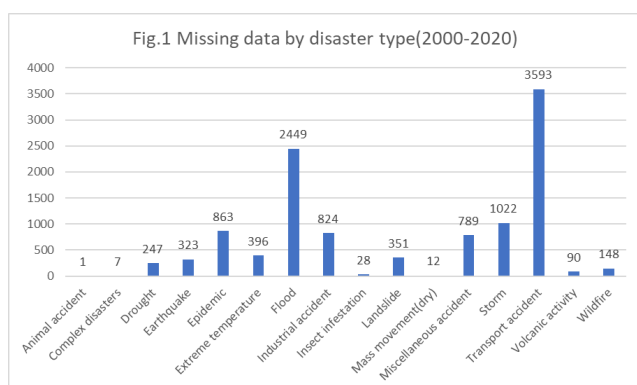


The validity of research largely depends upon the quality of data collected. It is quite common to encounter missing data while conducting research based and this can have a significant effect on the conclusions that can be drawn from the data. EM-DAT compiles different disasters from all over the globe with varying information on different variables. For this exercise, we extracted a dataset from EM-DAT, covering the years 2000 till 2020, and analysed the completeness of three economic variables: total damages, Insured damages, and reconstruction costs in US dollars, against other variables like disaster type, disaster subgroup, and location of the disaster (Continent and Country).

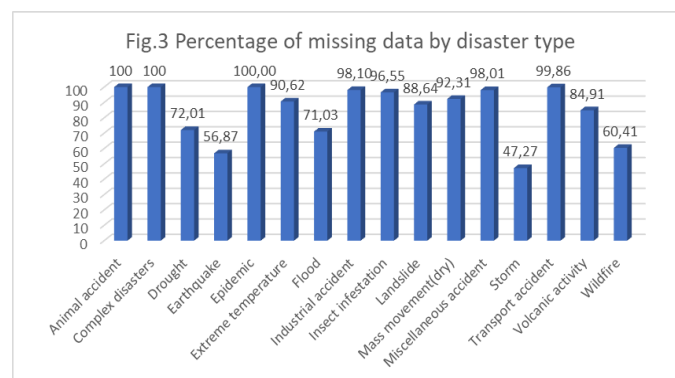
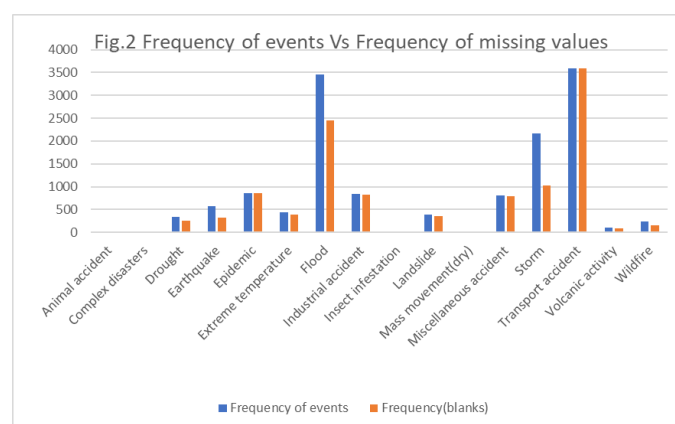
Looking at the dataset from 2000 till 2020, a total of 13862 entries were extracted which included all the continents and all types of disaster reported in EMDAT. 11143 entries had missing information regarding economic losses which amounts to 80.38% of the total.

When reviewing the number of missing records by disaster type in the dataset, transport accidents seems to have the highest number of missing economic loss information followed by floods, storms and epidemics (Fig.1).



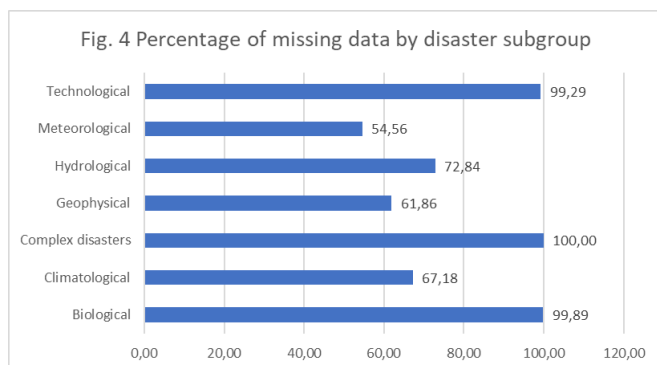
Looking at the frequency of missing economic loss data according to the events and tallying it with the total number of recorded events, we can see a clear picture of the extent of lack of data. Out of all types of disasters, storms, followed by earthquakes and wildfires, seem to have relatively more information in terms of economic loss, whereas animal accidents, complex disasters and

epidemics have 100% data missing. This is partly due to occurrence (animal accidents are very rare, and partly due to the nature of disasters, as is a huge work to estimate economic loss with regards to epidemics and complex disasters, as they are likely protracted in nature, and defining the parameters to estimate financial damage is therefore challenging (Fig. 2 & 3)

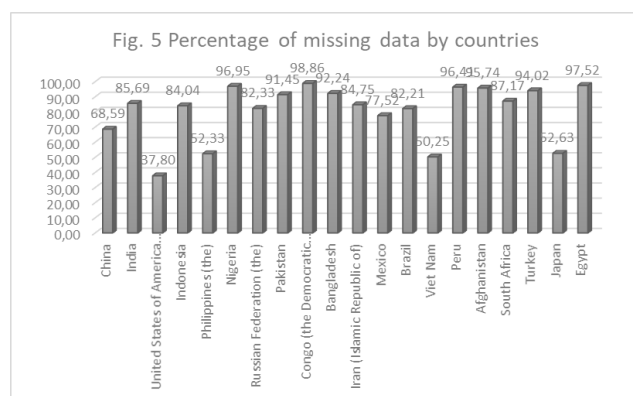


Classifying the missing data according to continent provides a different picture in terms of missing data . The continent of Africa has 95.44% of data missing while Asia, Europe, the Americas and Oceania follow with approximately 77%, 78%, 72% and 60% respectively. From these results, it is clear that most of the disasters recorded from the African continent have no estimate of total damages in US \$.

Looking at the relative frequency of missing data by disaster subgroup, technological, biological and complex disasters have almost no economic data in the dataset while meteorological disasters are relatively more complete with 54.56% of data missing, followed by geophysical, climatological and hydrological disasters with 61.86%, 67.18% and 72.84% respectively (Fig. 4).



Moreover, looking at the first 20 countries with the greatest number of disaster entries in the dataset and seeing the percentage of missing economic data, most of the countries have more than 70% data missing. Only countries with less than 70% missing data were China, USA, Philippines, Vietnam, and Japan out of the 20 countries with high disaster impact reviewed (Fig. 5).



When reviewing the insured damages in the dataset, the majority of insured damages in US\$ are missing: 13,200 out of 13,862 i.e 95.22% of total are absent from 2000 till 2020. Regarding the variable reconstruction costs, only 29 entries out of 13869 had the data present in the dataset, so approximately 99.8% of reconstruction costs in US\$ was missing.

After reviewing the overview of the dataset from 2000 till 2020, it shows that there are data gaps in terms of economic loss. The nature of disasters also seems to be playing a huge role in economic estimates in the aftermath of a disaster. For example, in case of a hurricane affecting multiple countries, either a collective damage following the disaster is reported, which is not specific to a particular country, and categorizing the loss by countries would be much more complicated. In conclusion, a new comprehensive approach needs to be developed, in order to fulfill the current gaps in economic losses and the system of reporting the economic loss after a disaster in itself. Within the coming years, improving economic loss data is one of the priorities that the Centre for Research on the Epidemiology of Disasters will focus on.

CRED updates and recent publications

- Bose-O'Reilly, Stephan ; Daanen, Hein ; Deering, Katharina et al. *COVID-19 and heat waves: New challenges for healthcare systems*. In: *Environmental Research*, Vol. 198, e-111153 (2021).
- Ogbu, Thomas Jidefor ; Guha-Sapir, Debarati. *"Strengthening data quality and reporting from small-scale surveys in humanitarian settings: a case study from Yemen, 2011–2019"*. In: *Conflict and Health*, Vol. 15, no. 1 (2021), #33 (2021)
- van Loenhout, Joris ; Vanderplanken, Kirsten ; Scheen, Bénédicte ; Van den Broucke, Stephan ; Aujoulat, Isabelle. *Determinants of adherence to COVID-19 measures among the Belgian population: an application of the protection motivation theory*. In: *Archives of Public Health*, Vol. 79, no.74, p. 15p. (2021).
- Maria Rodrigues Leal get the FNRS Grant 'Postdoctoral Researcher'

Data are subject to change, for enquires: contact@emdat.be