Application of disaster loss data in support of early warning and early action in African countries

Recommendations

The following is based on research and consultations with DRR and anticipatory action stakeholders in Malawi, UR Tanzania, Mozambique, Zambia and the SADC regional authority. Implementation of the following recommendations can improve the maturity and the functioning of data ecosystems allowing for enhanced application of loss data for anticipatory action and beyond.

Increase the quality of the primary disaster loss data collection

Ensure that data is available for diverse indicators relevant for different types of early action. The organizations that collect primary data on disaster loss have an (immediate) purpose for this, specific to their organization's objectives. However, other organizations can benefit from this data and different applications require data on different indicators. Better impact data also need to be collected. To reduce fragmentation and increase use of impact data for diverse applications, those organizations involved in primary data collection could agree ex ante on a methodology for harmonized data collection.

Encourage well-defined information management processes around disaster loss data and prioritize the operationalization of policies

Each data supplier should have a clear information management process for data collection, cleaning, quality control and validation, storage and sharing according to open data principles and in compliance with responsible use of data guidelines. An intermediary in the data ecosystem could take the role to make sure these processes are established, documented and harmonized. At the policy level, this could be made compulsory. At the national level, the National Disaster Management Authorities or the National Statistics Offices could play this gatekeeping role; at the regional and global level, UNDRR.

Dedicated funding for data ecosystem governance

Many of the challenges outlined above are exacerbated by the lack of continuous funding for quality control, processing and entry. Dedicated funding for the updating of national loss databases, for incentivizing agencies to make data public and for capacity building can drive demand for improved loss data.

Funding streams should buttress efforts to institutionalize the collection and use of impact data and aim to increase awareness of the importance of such data at both national and local levels. When disaster risk managers understand better the importance of impact data for early action and the financing mechanisms for early warning early action that are becoming

available via other windows such as emergency response funds[1], they are more likely to allocate internal funding and/or to seek external continuous funding to enable a dedicated capacity for collection, cleaning and opening up of impact data.

Improving partnerships

For data that cannot be made public, international organizations can provide support for improving and building partnerships between agencies that produce and house disaster databases and organizations who may need complete datasets for specific purposes. This requires creation of a collaborative and interactive environment with a transparency and feedback mechanism. It requires also looking into incentives for data producers or intermediaries to share data, such as for commercial, legal, or social good reasons and incentives for data users to use the data, such as understanding how improved impact data can lead to a more accurate trigger level for early warning early action.

Capacity Building

Different types of capacity building training exist to help improve the quality and the consistency of loss data collected and managed by national DRR stakeholders. Whereas training often focuses on data collection, there is a need to increase the capacity of national officers and experts to analyze and use loss data. General training on database management is also essential for technical support staff.

International organizations should provide support to DRR/DRM authorities to identify, define and standardize country-wide indicators. Such objectives should reflect input from multi-stakeholder consultations and from user needs assessments. Processes aiming for data standardization should result in guidelines and training for the collection, cleaning and validation of data.

There is simple and workable data and digital technology that exists for data collection, hence technical training should be provided to enable a digitization of the data collection. Data collected directly in a digital form, rather than recorded on paper, will allow for more disaggregated data to be included in loss databases. Most importantly, it will allow for more time and more resources to be dedicated to cleaning, quality control and analysis.

Technology: develop a strategy on using and adopting existing and new technologies

Closely linked to capacity building, is the use of data and digital technologies to improve the operational excellence of the disaster loss data information management processes. Digital transformation of these processes can reduce the workload and increase the speed. It necessitates the availability of hardware and access to the internet at the local level. Apart from the need to ensure access to existing data and digital technology, new techniques can

[1] Such as the window into IFRCs Disaster Response Emergency Fund (DREF) and UN OCHA Central Emergency Response Fund (CERF).

be used to enrich the "official" impact data collected in the field. Example, include text mining and earth observation data, which can be beneficial to bridge some gaps in impact data.

A scraping and text mining algorithm can be used to automatically extract impact data from a variety of digital media sources, such as Twitter or online news media[2]. Once a text mining algorithm is set up, it can work fast and efficiently. The impact data from digital media is often complementary. For example, there can be more detailed information on the impact in urban areas in digital media (as more people are using social media in cities than in rural areas)[3] or there can be detailed information on the timing of the event (e.g. in individual tweets).

Other relevant sources of data

There are also challenges - Data collection methods and sources used by the media to report impact can be undocumented or not known and reported data may not be cross-checked or validated. Digital media has become only in use over the last decade, so there are limitations in the temporal coverage. Text mining requires technical expertise in artificial intelligence in languages for which Natural Language Processing code is openly available. However, this is not possible for local languages, therefore one must revert to commercial solutions. Financial resources are also required for an initial historical data collection when newspaper archives are behind a paywall.

However, earth observation is also promising. Monitoring from space can reach data scarce areas and provide images of high spatial resolution (approximately 50 cm). Novel deep learning techniques are available to extract impact data from these images. For example, images can show damage to buildings or changes in land use land cover. In many cases it will be possible to extract a flood extent and subsequently enable disaggregation of impact data otherwise only reported at an aggregated level. These historical flood extent maps can be used to develop impact-based forecasts that predict, when a weather forecast comes in, which areas will be flooded. This is essential information for the decision makers that use an early warning early action system.

The use of earth observations has limitations. Like text mining, earth observation analyses require technical expertise and can be computationally intensive. Access to data is not always possible. The International Charter: Space and Major Disasters, a venture between 17 space agencies, provides free satellite data to pre-registered organizations that work in humanitarian response. However, this data is not available for smaller scale disasters, and it is only imagery from right after the event, whereas for some analyses before and after imagery is required. Usually, the higher resolution imagery is only available at commercial prices. Optical satellite

[2] van den Homberg, M. and J. Margutti (2022). Enriching Impact Data by Text Mining Digital Media. GAR2022 Contributing Paper. Geneva: UNDRR https://www.preventionweb.net/publication/enriching-impact-data-mining-digital-media

[3] S. Majani, B., D Malamud, B., and Millington, J.: Use of blended evidence sources to build a history of flooding impact and an impact severity scale: A case study of Nairobi, Kenya, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-12012, https://doi.org/10.5194/egusphere-egu22-12012, 2022.

imagery can have quality issues if there is cloud cover. This can mean that images taken immediately after a flood event may not be suitable to estimate flood extent. Likewise, the precise time of a flood must be known, otherwise it is complicated to find the corresponding set of images. On the positive side drone imagery is becoming more common. For example, the INGD in Mozambique has made available drone imagery from after cyclone Idai.

New technologies offer potential to enrich data in existing databases. However, they are not a panacea, but offer important opportunities to complement data collected through traditional means and enhance the quality of datasets.

It is a responsibility and obligation for all those actors active in DRR to start, revamp, or continue the systematic collection, storage and use of disaster loss data as part of the process of contributing to a better understanding and a further reduction of the risks of those exposed to the impact of climate change and climate variability. High-quality disaster loss data is primordial for realizing effective and adequate early warning early action mechanisms.