Bridging the gaps in disaster loss data to support early warning and early action in Southern Africa
The application of disaster loss data for early warning and early action

Acknowledgements 3
Disclaimer 4

Introduction 5

The current context: slowly maturing data ecosystems 7

Fragmentation: multiple organizations hold data 8
Insufficient data on detailed indicators 9
Data is not quality controlled nor validated 9
Data is not of high enough quality 10

Recommendations: toward better impact data 11

Increase the quality of the primary disaster loss data collection 11
Encourage well-defined information management processes around disaster loss data and prioritize the operationalization of policies 11
Governance: dedicated funding for continuity of high-quality data supply 12
Governance: facilitate and promote collaboration between users and suppliers 12
Capacity Building 13
Technology: develop a strategy on using and adopting existing and new technologies 13

References 16
Acknowledgements

UNDRR would like to acknowledge its major core donors; Sweden, Japan, Norway, Switzerland and Finland as well as the European Union, for their support of this publication. However, the views expressed in this publication are the authors’ alone and its contents and do not necessarily reflect the views of the donors.

This report was written by
Marc van den Homberg (510),
Katarina Mouakkid Soltesova (UNDRR),
Arielle Tozier de la Poterie (510),
Roberto Schiano Lomoriello (UNDRR),
Inez Gortzak (510)
The application of disaster loss data for early warning and early action

Disclaimer

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country or territory or of its authorities or concerning the delimitations of its frontiers or boundaries. The designations of country groups in the text and the tables are intended solely for statistical or analytical convenience and do not necessarily express a judgment about the stage reached by a particular country or area in the development process. Mention of the names of firms and commercial products does not imply the endorsement of the United Nations.

This document covers humanitarian aid activities implemented with the financial assistance of the European Union. The views expressed herein should not be taken, in any way, to reflect the official opinion of the European Union, and the European Commission is not responsible for any use that may be made of the information it contains.

Note: The designations employed and the presentation of maps in this report do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities or concerning the delimitation of its frontiers or boundaries. Some rights reserved. This work is made available under the Creative Commons Attribution-NonCommercial 3.0 IGO licence (CC BY-NC IGO); https://creativecommons.org/licenses/by-nc/3.0/igo/legalcode

Under the terms of this licence, this work may be copied, redistributed and adapted for non-commercial purposes, provided that the work is appropriately cited. In any use of this work, there should be no suggestion that UNDRR endorses any specific organization, products or services. The use of the UNDRR logo is not permitted. If a translation of this work is created, it must include the following disclaimer along with the required citation below: "This translation was not created by the United Nations Office for Disaster Risk Reduction (UNDRR). UNDRR is not responsible for the content or accuracy of this translation. The original English edition shall be the authoritative edition."

Users wishing to reuse material from this work that is attributed to a third party, such as tables, figures or images, are responsible for determining whether permission is needed for that reuse and for obtaining permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user. Sales, rights and licensing.

UNDRR information products are available for non-commercial use. Requests for commercial use, rights and licensing should be submitted via: https://www.undrr.org/contact-us This publication may be freely quoted but acknowledgement of the source is requested.

Citation: UNDRR (2023), Bridging the gaps in disaster loss data to support early warning and early action in Southern Africa, United Nations Office for Disaster Risk Reduction (UNDRR).

© 2023 UNITED NATIONS OFFICE FOR DISASTER RISK REDUCTION

For additional information, please contact:

United Nations Office for Disaster Risk Reduction (UNDRR)
7bis Avenue de la Paix, CH1211 Geneva 2, Switzerland, Tel: +41 22 917 89 08
Anticipatory action is a relatively new and rapidly expanding approach to humanitarian financing which seeks to mitigate the impact of extreme events by allowing governmental and humanitarian organizations to act before extreme events occur. Thereby, effective anticipatory action systems require an agreed upon and calibrated trigger, targeted early actions, and rapidly accessible financing. Anticipatory action triggers are set to activate early actions in a specific region when an impact-based forecast exceeds a predetermined threshold for event probability and magnitude. To accurately indicate where early actions should be targeted, triggers must be based on skilful forecasts and high-quality impact and risk data.

Disaster risk managers use impact data for multiple purposes, but until recently there has been little focus on the usefulness of impact data for anticipatory action. Nevertheless, access to consistent, high-quality and disaggregated historical impact data is pivotal to developing triggers and selecting effective anticipatory action.

The Sendai Framework for Disaster Risk Reduction 2015-2030 has two priorities and one target directly related to anticipatory action. Priority 1 emphasizes the need for disaster risk management to be based on an understanding of disaster risk in all its dimensions of vulnerability, coping capacity, exposure of persons and assets and hazard characteristics. Priority 4 calls for the strengthening of disaster preparedness for response and anticipatory action. Target G calls for increasing the access to disaster risk information and assessments. Indeed, many Sendai Framework targets depend on the reporting of quality impact data. In line with these priorities for action, UNDRR supports Member States and Disaster Risk Reduction (DRR) stakeholders in building DRR capacities in African countries, including for the application of data to support anticipatory action.
The application of disaster loss data for early warning and early action

Example: How impact data can inform trigger development

Anticipatory action is crucial to mitigate the impact of flooding in Zambia, but it needs to be guided by an impact-based forecast that is credible and actionable. The Zambia Red Cross Society, the Red Cross Red Crescent Climate Centre, 510 (an initiative of the Netherlands Red Cross), the Water Resources Management Authority (WARMA) and the Zambia Disaster Management and Mitigation Unit (DMMU) have developed a threshold-based trigger model with a lead time of up to seven days. This model uses the Global Flood Awareness System (GloFAS, part of the EU Copernicus Emergency Management Service) as a hydrological model, as well as historical impact data of around ten years to set the trigger. The hydrological model is integrated with an assessment of vulnerability and coping capacity, based on secondary data for the whole country at the highest level of possible granularity (district level in this case). The trigger is implemented on an impact-based forecasting platform, and integrated in the Emergency Operation Centre operated by Zambia’s DMMU as a part of the country’s standard Early Action Protocol.

This report outlines current gaps and key policy options available to bridge the gap between available disaster loss data and what is needed to translate them into anticipatory action. It is intended for use by disaster risk managers from national to local levels to help advocate for enhanced disaster loss data in support of anticipatory action/early warning early action. The report draws on lessons gathered in 2022 during a collaboration between disaster risk managers in the United Republic of Tanzania, Malawi, Mozambique and Zambia, UNDRR, 510 (an initiative of the Netherlands Red Cross), and the Anticipation Hub.¹

¹ During the project, loss data from DesInventar (www.desinventar.net) and EM-DAT (www.emdat.be) was analyzed. In addition, for Malawi, impact data from the Department of Disaster Management Office (DoDMA) was included as well.
The current context: slowly maturing data ecosystems

In order to develop accurate and reliable anticipatory action triggers, governments and humanitarian organizations need high-quality impact data: broad spatial coverage of the entire country and not just certain areas, granularity disaggregated to the lowest possible administrative level, temporal coverage of at least two decades, temporal resolution (preferably a precise starting date so that the record can be matched to hazard information), and indicator coverage. As disasters happen frequently, a continuously evolving, self-organizing system of feedback and adjustment among actors and processes for impact data is needed, i.e., a data ecosystem. A data ecosystem can be characterized and structured around five dimensions: actors, data supply and demand, data infrastructure, and data ecosystem governance (van den Homberg & Susha, 2018). An analysis of existing impact data ecosystems in Malawi, Mozambique, and Zambia has revealed that several challenges currently limit the countries’ ability to systematically collect, validate, store and share impact data, thereby reducing the usefulness of existing data for anticipatory action. Each challenge is discussed at greater length below.

Continuity of collaboration between users, suppliers and donors of disaster loss databases

National authorities face persisting challenges in maintaining staff dedicated to overseeing data collection, entry and analysis. DRR and Disaster Risk Management (DRM) authorities often rely on consultants or project staff hired with external funding to populate the national disaster loss database. The continuity of data collection methodologies and data coverage is therefore frequently compromised, affecting data quality.
The application of disaster loss data for early warning and early action

Fragmentation: multiple organizations hold data

Several governmental and humanitarian organizations collect impact data for their respective sectors, often with a specific focus such as agricultural or infrastructural impact. In many cases, procedures for data sharing are complicated, even between government bodies, making it challenging for risk management authorities and DRR stakeholders to systematically access sectoral data. As a result, organizations may not be aware of each other’s sectoral data.

Data collection processes can differ per hazard type and intensity. Following major events which garner international attention and support, short-term resources become available for in-depth damage and needs assessments, and corresponding data analysis. Conversely, for disasters with lower but more frequent impacts, fewer resources are available for data collection. Such data is still often available at the local-government level, but it is rarely digitized or kept in a database. Dynamics between local and central governments also frequently prevent ways to incorporate such data in national databases. As a result, impact data remain dispersed across agencies and actors, and are not accessible for national analyses and applications. Such dispersal ultimately increases the chances that such data are eventually lost or destroyed.

Access to the data

Publicly available data is of greatest value to the public and to DRR. It allows diverse users to contribute to DRR, and increases accountability and the quality of public services. Some suppliers open their impact data to the public, others only supply data upon request, and some do not share their data at all. Specific agencies or sectors may be reluctant to make their data publicly available for a number of reasons, including lack of capacity, fear of publicizing potentially sensitive information, or because their data has commercial value. As a result of organizations – public or private – not sharing their data, gaps in open databases may continue to persist.
The current context: slowly maturing data ecosystems

Insufficient data on detailed indicators

Until recently, many DRM agencies only captured high-level statistics on a limited number of harmonized indicators, such as the number of people who died or were affected. However, in order to develop early action triggers to reduce specific humanitarian impacts through anticipatory action, more data are required on indicators which are not commonly reported. It is necessary to understand who or what has been damaged, providing as much detail as possible. For instance, to save lives, it is important to know not only the number of fatalities, but also how people die during a disaster event. To protect livestock from death or disease during a flood requires knowing the number of livestock affected and the cause of death or illness, including the number and type of waterborne disease cases reported. Without this information, responders may attempt to prevent drowning deaths through the provision of vaccines or provide medicines for the wrong kind of infection. While there has been progress in collecting more data for each indicator and on a wider range of indicators, most damage and needs assessment processes still do not collect detailed data on the kinds of houses or crops being damaged, specific kinds of injuries sustained, or causes of death.

Data is not quality controlled nor validated

The quality of data differs by organizations as it depends on the purpose for which the data was collected and the resources available. Early on in a disaster, rapid damage and needs assessments are done, whereas later, in the recovery phase, there is more time for in-depth data collection. Some humanitarian organizations might focus on data collection in their specific areas of interventions and expertise such as WASH or shelter, whereas governmental agencies usually cover all areas which were impacted across multiple sectors. Information on how data was collected and how key indicators were defined may be missing. This implies that data from different actors must be cleaned, harmonized to the extent possible, and validated before it can be included in national loss databases. Without consistent quality control and validation, even readily available data may not be used or included in official publicly accessible databases.
Data is not of high enough quality

Data quality comprises spatial coverage, granularity, temporal coverage, resolution and indicator coverage. Source reliability and content accuracy influence data quality. With regard to spatial coverage, depending on the data collection methodology used by national authorities, there may be less data on impacts in remote and rural areas. In other contexts, political concerns may prevent reporting of losses in informal urban settlements. Temporal resolution is important for understanding the hazard-impact relationship. However, the timestamp of a hazard often differs from the timestamp of recorded impacts due to delays in when the impact data is collected. The date of when the impact data is collected or when the hazard took place might also not be well recorded. This results in difficulties in, for example, finding out whether impact was due to a flash flood or a river flood.

As for source reliability and content accuracy,\(^2\) platforms such as EM-DAT and DesInventar often contain data at the national and sub-national level such as provincial or district, but rarely at the sub-district or village level. This is often due to aggregation of the data as it is channelled from the local to the national level. Furthermore, disaggregated datasets are large and often require extensive cleaning before they can be entered into national databases. Teams of trained data officers are necessary but not always available. However, to conduct analyses which allow for precise targeting of the risks faced by distinct communities, it is invaluable to have both vulnerability and impact data available at the lowest possible administrative level, for instance the in Mozambique and the Group Village Head (GVH) level in Malawi).

---

Recommendations: toward better impact data

Based on research and consultations with DRR and anticipatory action stakeholders in Malawi, United Republic of Tanzania, Mozambique, Zambia and the Southern African Development Community (SADC) regional authority, the implementation of the following recommendations can improve the maturity and functioning of data ecosystems. In turn, this will allow for enhanced application of loss data for anticipatory action and beyond.

Increase the quality of the primary disaster loss data collection

It is important to ensure that data is available for diverse indicators relevant for different types of early action. Organizations which collect primary data on disaster loss have an immediate purpose to this end, specific to their organization’s objectives. In consideration of other organizations which can also benefit from this data and that different applications require data on different indicators, better impact data can be collected. To reduce fragmentation and increase use of impact data for diverse applications, 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 within the data ecosystem could be tasked with ensuring that these processes are established, documented and harmonized. At the policy level, this could be made compulsory. National Disaster Management Authorities or National Statistics Offices could play this gatekeeping role at national levels, and UNDRR at regional and global levels.
Governance: dedicated funding for continuity of high-quality data supply

Many of the challenges outlined above are exacerbated by the lack of continuous funding for quality control, processing and entry. Dedicated funding for updating national loss databases, for incentivizing agencies to make data public, and for capacity building can drive the supply and demand of 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. It is crucial that disaster risk managers better understand the importance of impact data and financing mechanisms for early warning early action through windows into emergency response funds. With a better understanding around those concepts, they are more likely to allocate internal funding and seek external funding to build dedicated capacities for collection, cleaning and opening up of impact data.

Governance: facilitate and promote collaboration between users and suppliers

Apart from dedicated funding, strong and continuous collaboration among data suppliers and users of disaster loss databases is essential, as well as collaboration among different users who analyze the data. Whereas sometimes, one user analyzes loss data individually, this can also be done by several users through a collaborative analysis. For example, in many countries, multi-actor Technical Working Groups for anticipatory action decide together on a threshold level for early action.

An intermediary organization such as UNDRR at the global level, or a National Disaster Management Authority at the national level, can provide support to improve and build partnerships between agencies which produce and house loss databases, and organizations who may need such datasets. This requires the creation of a collaborative and interactive environment with a transparent feedback mechanism. To this end, incentives of a commercial, legal or social nature need to be provided for data producers to share data, as well as 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. Intermediaries can further develop prototype data sharing agreements, or make data sharing compulsory through official mandates and regulations.

---

3 Such as the window into IFRCs Disaster Response Emergency Fund (DREF) and UN OCHA Central Emergency Response Fund (CERF).
Capacity Building

Different types of capacity building training exist to help improve the quality and consistency of loss data collected and managed by national DRR stakeholders. General training on database management is 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.

Finally, simple and workable data and digital technology exists for data collection, hence technical training should be provided to enable a digitisation of this process. Data collected directly in a digital format, rather than recorded on paper, will allow for more disaggregated data to be included in loss databases. Most importantly, this will allow for more time and resources to be dedicated to cleaning, quality control and analysis.

Whereas training on data collection is important, there is also a need to increase the capacity of national officers and experts to analyze and use loss data for the different possible applications.

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 disaster loss data information management processes. A 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 be used to enrich the “official” impact data collected in the field as well. Two examples – text mining and earth observation data – are presented below to demonstrate how they 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 X (formerly known as Twitter) or online news media (van den Homberg et al., 2022). Once a text mining algorithm is set up, it can work fast and efficiently. The nature of impact data from digital media is often complementary. There can be detailed information on, for instance, the impact in urban areas as more people are using social media in cities than in rural areas (Majani et al., 2022), or on the timing of the event visible in individual tweets. Beyond the benefits of text mining, there are also challenges.
Data collection methods and sources used by the media to report a disaster’s impact can be undocumented or not known, and reported data may not be cross-checked or validated. Digital media have only been used over the last decade, resulting in limitations when it comes to temporal coverage. Moreover, 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, an issue which necessitates commercial solutions. Financial resources are also required for an initial historical data collection when newspaper archives are behind a paywall.

Earth observation as a method also holds promise. Monitoring from space can reach data-scarce areas and provide images of high spatial resolution (approximately 50 cm). Moreover, novel deep learning techniques are available to extract impact data from these images, such as damage to buildings or changes in land use and cover. In many cases, it is possible to extract a flood extent and subsequently enable the disaggregation of impact data which would otherwise only be reported at an aggregated level. These historical flood extent maps can then be used to develop impact-based forecasts and predict which areas will be flooded once a weather forecast is released. This is essential information for decision makers who use an early warning early action system.

The use of earth observations also entails limitations. Similar to text mining, earth observation analyses require technical expertise and can be computationally intensive. Furthermore, 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 working in the humanitarian field. However, this data is not available for smaller-scale disasters and only provides imagery from right after the event, whereas for some analyses, imagery is required from before and after the event. Moreover, higher resolution imagery is usually only
available at commercial prices. Optical satellite imagery can also have quality issues if there is cloud cover, resulting in images documenting the aftermath of a flood disaster potentially not being suitable to estimate flood extent. Likewise, without knowing the precise time of a flood, it is complicated to find the corresponding set of images. On a positive note, drone imagery is increasingly becoming available and has already been used, for instance, by the National Institute for Disaster Management (INGD) in Mozambique in anticipation of, and response to, Cyclone Idai.

In short, new technologies show potential to enrich data in existing databases. While they are not a panacea, they nevertheless provide important opportunities to complement data collected through traditional means and enhance the quality of datasets.

Ultimately, all actors actively involved in DRR have the responsibility and obligation to start, revamp, or continue the systematic collection, storage and use of disaster loss data in order to contribute to a better understanding and a further reduction of the risks faced by those exposed to the impact of climate change. In particular, high-quality disaster loss data is primordial to realise effective and adequate early warning early action mechanisms.
References


