EM-DAT Technical Advisory Group Meeting

Meeting Notes , October 16-17, 2019, Brussels

Wednesday October 16, 2019

Opening session: Debby Guha-Sapir, Rhonda Davis-Stewart

- ➢ A note of gratitude to the funding partner and more especially to USAID for their immerse contribution over the years
- ➢ Reminded the audience the advantage of having such group with diverse expertise
- ➢ The focus on the need for availability of varieties of data on disaster and the need for accuracy in terms of planning.

Session A - EMDAT Overview and Mapping

Progress in EM-DAT, Workplan and Main changes : Regina Below, Valentin Wathelet and Joris van Loenhout

- ➢ A recap of the last EMDAT TAG meeting that was held in Washington Dc, 2016 and the main issues that was raised during that meeting: to formalize relationship with partners, the need to develop and improve online working tools, and continue work on the geo-referencing of EM-DAT
- ➢ Information on status of EMDAT usage and the general experience of EM-DAT users were also mentioned
- ➢ Participants were also notified of the pressing demands of users such as easy access to the data as well as the need for improvement of visualization and provision of compatibility with mobile devices
- ➢ Participant were also informed that this is the last year of EM-DAT project and the upcoming EM-DAT activities for the last year

Q&A

- ➢ During Q&A Debby raised the point on the current issues bordering on the development of standardized tool for collecting precise data at regional level.
- ➢ It was also noted that EM-DAT does not actively track data usage but that the effective and capital less intensive method would be tracking citations with EM-DAT

Using EM-DAT for localized Disaster Risk Awareness : Nate Smith

- ➢ The pros and cons of using EM-DAT was discussed, likewise the difficulties in geo-referencing data
- ➢ Also the problem of using static data in assessment of dynamic was also brought to the fore

- The possibility of moving EM-DAT from is operational space into connecting to dynamic risk platform was also noted.

Q&A

- During the Q & A there was a discussion on cheaper alternative of doing risk assessment in some cities without the financial power of implementing the methods that were currently under discussion.
- It was also noted that the cost of such processes are actually coming and that despite the high cost investment in current method that the upside is that the information from such method can be useful in other areas, hence, worth the investments

Integrating diverse data hazards, exposure, vulnerability : Robert Chen

- One of the major issued raised were the availability of too much data is marred with the problem of how to make the best out the data
- The participants were also reminded that need depends entirely on the users and that having the right data is very important in risk assessment
- Discussion on the ongoing collaboration with partners with varying interest such as partners interested in water hazard, some interested on how best to provide internet access to hard to reach community in places like Africa
- Discussion on the nature of ongoing work in terms of population estimates, tracking of impact data
- Participants were also informed of the existing collaborative opportunity in terms of improve mapping of population, settlements, infrastructure, movement etc. and the need to develop tools that automate the harmonization and quality control of data

Q&A

- During the Q&A session the issue of rapid change in population for instance, and the challenges it possess to the humanitarian community was highlighted. In reaction to that it was said that there is an ongoing work in that area but at the moment the major issue is developing a comparable baseline
- The issue lack of current data with relations to conflict affected area was also raised and the lack of involvement in the mapping of conflict related data specifically surveys that have been conducted by humanitarian organization.

Session B – Global Datasets

WMO uses cases for disaster loss data and the WMO initiative to record hydrometerological, climate and space weather hazardous events : James Douris

- Importance of disaggregated impact data in forecasting was highlighted

- Understanding the impact of hazards in terms of how they accumulate and affect the population can be achieved with the aid of disaggregated data
- The underlying issue of how best to accurately attribute loss and damage to a natural hazard from a metrological view is being addressed by creating a record
- The challenge of working with country specific data rather than the specific affected area is issues that are set to address
- In as much as using local definition of events the need for standardize definition should not be understated

Q & A

- During the Q &A session the use of unique identifier as key ingredient in linking data that on the surface may not look connected but on a deeper level are actually connected was highlighted
- The risk of sending alerts being set was raised and it was highlighted that WMO system is not an alert system but rather time chorological records of event

Socio-economic indicators at small-area level for disaster-risk management (Jeroen Smits)

- Information on the various data contained in the database was discussed
- The advantage of using the database for comparative research was highlighted
- Some capabilities of the data global base lab was also highlighted

Q & A

- During the Q & A the challenges of changing boundaries were raised likewise the recreation of data based on the new boundaries were also stated

The Humanitarian Data Exchange: Disaster data, data sharing and users demand : Javier Teran

- Participants were informed of the rigorous process in place to ensure anonymity when putting together these data
- During the Q & A, The challenges of changing boundaries in data collection was highlighted as one of the major challenges
- The issue of possible duplicity of data was raised and the need for a center with information on what data is out there, what has been done and by whom was also mentioned. But it was explained that it would great for every data owner to contribute but they are not in a position to coerce data owner. The fundamental principle is for data owners to freely submit their data
- In terms of the need of tackling quality control of data and duplicity, it was stated that there are specific data guideline and also an ongoing work on reformation of data in order to get rid of duplicity

Thursday October 17, 2019

Session C. Economic Impact

Characteristics of NatCat databases, and their role in disaster risk financing strategies : Peter Höppe

- ➤ NatCat database is presented: this database is particularly well suited to analyse economic losses resulting from extreme events.
- ➤ A survey was conducted to compare the different databases available recording extreme events : NatCat, Sigma, EM-DAT, CATDAT, AON Database, PERILS, Desinventar.
- ➤ As PERILS and Desinventar don't have global coverage, there have been excluded from the comparison and only with the five other databases were included in the comparison study
- ➤ It exists much difference between the analysed databases:
    - o The period covered by each database is different
    - o It was stressed that every database uses different impact thresholds to register its extreme events. This situation explains partly the difference regarding the number of extreme events recorded in each database and makes comparisons between databases difficult. NatCat doesn't use precise impact thresholds when registering its extreme events but uses different ones depending on the type of publications or analysis. This methodology lets more freedom to the data-users to set their own thresholds.
    - o The notion of "complete coverage" (period considered as almost complete regarding the number of extreme events recorded) is introduced. This information is not all the time directly available to data users. These periods vary between databases and are estimates (decades) rather than a precise period.
    - o Sigma and AON don't have an extreme event intensity classification
    - o Technological disasters are not recorded by all databases

- ➤ Sigma, NatCat and EM-DAT use almost the same extreme event classification (see Peril Classification). It's a good point for these three databases.
- ➤ The challenge of normalising the losses in time has been highlighted. All databases use different methodologies. NatCat applies a Grid Cell Product methodology, which is particularly useful for large countries with uneven economic development (example: China).
- ➤ All databases are geo-referencing their data but with different methodologies.
- ➤ The problem of data accessibility was also raised. None of the five databases compared is open access. However, aggregated data are easier to access. It was raised that using licences to share data can be challenging for a number of institutions.
- ➤ Nowadays, the need for data is one of the challenges that have to be faced:
    - o Developing countries are the most affected and it is where it is the most challenging to get precise data.
    - o People are more aware of climate change: to better understand it and its consequences, a lot of data are needed.
- ➤ Developing efficient disaster risk reduction strategies required precise data describing these extreme events. Once again the difficulty for a number of users to access such data was raised.
- ➤ Insurance companies have a role to play: ASEAN Disaster Risk Financing and Insurance Programme, Phase 2 was presented as a concrete example. This program requires a large amount of precise data to quantify the risks as well as the potential fiscal impact of an extreme event.

Q&A (Peter Höppe)

➢ Currently, for an internal risk assessment (example: at a county scale), EM-DAT is sometimes preferred due to its low price. Others databases could be used but in some cases but countries have a limited budget to invest in the purchase of such databases.

Development on Global loss databases based on IT support : Petra Löw

➢ Managing a database such as NatCat required a well-developed methodology.

➢ In the last recent years, NatCat has noticed more and more competitors in the field of extreme event database. This tendency shows the increasing interest of the public for such data.

➢ The need for data continues to rise. It results in an increase in the data value. However, we need more than just data: tools to analyse them. Tools developed by MunichRe were presented. Two versions exist, one for intern use and one publicly available on the internet. Being able to easily share results from these tools was one priority.

➢ NatCat Database is not a huge income for MunichRe. Selling data was recognised as a difficult task.

➢ The use of remote sensing data is a new trend. It really exploded last two years. It is a new field for MunichRe and a real challenge to handle and process such data.

➢ Once again, climate change was recognised as attracting more and more attention. For insurance companies, it is a new market but also a challenge to develop products for it.

➢ NatCat cannot be defined as "Big Data". MunichRe sees itself as a data manufacturer: a "small" database but highly valuable.

➢ NatCat had to evolve from a global database to a database with a finer scale. One challenge is being able to define the footprint of extreme events (both the physical and economic loss footprint). For the biggest extreme events, it is less problematic since they are well documented but for smaller events, it is a challenging process. These small events have a cumulated high impact but are hard to document. Text-mining was one of the solutions being considered but it has its own challenges.

Q&A (Peter Höppe & Petra Löw)

➢ The challenge of thresholds was raised. We do not have a complete picture of small events in certain regions (example: Asia). Text-mining (report/internet) could be a solution but it seems challenging to implement.

➢ The lack of relation (parent-infant) between extreme events was discussed. Cascading events are challenging to analyse. For the moment, this information (parent-infant) between events is not implemented in NatCat but could be added in the future.

➢ The multiplication in recent years of extreme event databases is a point of concern for some participants. A similar situation occurred for the gridded population data and climate models. It is not possible (or desirable for some participants) to stop anyone to start a new project in this field but a common framework has to be designed to compare, validated and improve the already existing databases. The databases owned by private companies have few chances to be released completely free since these companies invested money to build and manage them. However, it is important to keep them accessible to a large number of users.

Moreover, some participants questioned the ability of some databases to be maintained over the long term given the small number of people involved.

➢ The challenge of defining which criteria and threshold to use to define an extreme event was discussed. Resources (human and financial) have to be taken into account. Each institution tries to find a just middle: not too much record to register, but not too little either.

➢ Different data-providers agreed on the fact that recording such extreme events is a time-consuming process. The term "data-manufacturer" emphasises this notion.

➢ It does not exist yet a universal standard to register disasters and their impacts. Such initiative could facilitate the recording of smaller events. However, the different extreme databases have different purposes and users and thus not the same need in information.

➢ For cities or small municipalities, Desinventar could be an option for risk assessment. This database is not complete but easily accessible.

➢ Different large private companies (example: Facebook) release data and work on different products for humanitarian operations ("Data for good" initiatives). Insurance companies also share their data with partners.

Session D. EM-DAT Users

The effects of disasters on armed conflicts: Can disaster trigger or intensify conflicts? : Nicolas Caso & Rodrigo Mena

➢ Statistics about conflicts and disasters are rarely combined although there is a frequent co-occurrence of conflict-related and disasters-related deaths. In fact, for the year 2018, 18 % of the total deaths for battle-related (392,641) and disaster-related (11,804) deaths occurred in the same country and at the same time.

➢ The main objective was to build a model for armed conflict combining several aspects (conflict history, social and disaster related), among which:
   o Pre-existing armed conflict (1 year before)
   o Existence of a long-lasting conflict
   o Human development index
   o Polity index (as political stability estimator)
   o Country population (log of -)
   o Natural resources as a percentage of the GDP
   o Disaster impact (several impact aspects, based on EM-DAT)

➢ The results show that Social and conflict history parameters improve the explanatory value of the model. There are some strong regional fixed effects related to South and West Asia, as well as for Central Africa. The disaster parameters never enter the model significantly in any tested modality. More complex social processes are involved the initiation of armed conflicts than disasters alone.

Q&A

➢ When asked if they planned to study the problem the other way around, i.e. modelling disaster impact during conflicts. They are currently considering this way of studying the conflict/disaster interface

- It was also noted that one of the issues about disasters concerns the confidence the government was in while responding to a disaster and how it could undermine its legitimacy in the future. The problem being that the model proposed apparently does not capture the interaction between a crisis and the government legitimacy. The problem being that politics are very complex to incorporate in such a model.
- I was also pointed out that these results are in line with a lot of recent studies on conflicts.

IBM Operations Risk Insights - using EM-DAT, social media and weather data with AI to predict and mitigate severe conditions : Thomas Ward

- Demonstration of the IBM Operations Risk Insights tool, its purpose is to provide automated alerts for specific locations based on the social media, meteorological data and EM-DAT impact in order to optimize the supply chain risk mitigation.
- Started as an internal tool programming effort during summer, maintained by 2 programmers simultaneously, using various IBM data accesses. Participate in IBM supply chain innovation.
- All the trusted sources are combined with an AI procedure combined with an ensemble analysis and regularly updated to provide almost-real-time estimations. These sources include the Weather Company data, risk alert services (GDACS, USGS, Meteo, TWC/NWS, ORI Storms & Resilink) and social media tracking/mining.
- Internet-of-things (Iot) measure instruments are shipped along with IBM containers to provide additional measurements in the regions of activity of IBM.
- The users can define points of interests, for IBM usage this could means airports, supplier's locations, data centers. Selected NGO users can define their own point of interests related to their activities or region of interest and configure alerts for specific situation and category of receivers/roles. Feedbacks from these external users are important for the tool development.
- EM-DAT data is relevant in that context in order to differentiate an *interesting event* from a *serious* one, which should be used as an alert trigger. This problem is solved using IBM cloud AI capabilities, by categorizing disasters in different impact classes.

Q&A
- It has been pointed out that the potential danger of causing a panic within the population by giving a disproportionate importance to some sources in the process.

National Ownership of Disaster Loss Data : Subhra Bhattacharjee

- DesInventar, based on the DesInventar Sendai framework, is addressing several problems usually encountered in national databases such as the lack of granularity in the data, the heterogeneity of data between countries and the lack of software maintenance/update by proposing a standard along with an open-source implementation for such data to be collected, controlled, aggregated and presented per country.
- There are currently 73 datasets in the global implementation and 37 external servers running DesInventar. For each event, a data-card for each affected locality is completed by the local data provider, giving a high level of granularity to the data.

- ➢ Problems with the authorities are frequently encountered in the confidence they can have and/or display in the products of the data.
- ➢ The disasters are considered regarding two main categories:
  - ○ intensive disasters (high severity, mid-low frequency)
  - ○ extensive disasters (low severity, high frequency, mostly highly localized)
  - ○ The extensive disasters are found to have a bigger share in the overall impact than the intensive events while also displaying a highly specific spatial distribution. DesInventar can therefore give better insights to local authorities for improving disaster prevention on specific areas at high risk for extensive disasters and therefore reduce the overall impact.
- ➢ EM-DAT data is used for quality control over the impact values for disasters inputted by the country providers as well as evaluating under-reporting from the country and the human errors in data entry.

Q&A

- ➢ Limitations are usually encountered in the following cases:
  - ○ Human capacity: the data entry is usually managed by a limited number of people. Some gaps in the data were observed after a contributor retired (This also limits data duplication)
  - ○ The data could expose government officials, resulting in under-reporting or falsifications.
  - ○ There is a disconnection between ministries.
  - ○ Internet access with computers is limited in the country or in the government infrastructures.
- ➢ Disaster data is very political, in almost every country. There are different ways to answer country criticism of disaster data, in the case of EM-DAT the solution is to never change it for country reclamations. The advantage of DisInventar is the implication for the government credibility in the data provision process.
- ➢ In the end, the multiplicity of disaster-related databases is a strength as it creates a broader information *ecosystem*. This situation is much better than what a unique database could provide in terms of granularity, purposes, sources and capacity involved.
- ➢ REM-DAT 6 regional projects will target subnational data for disaster involving selected users/providers in the region. An Asian pilot is currently in preparation with collaboration of partner universities in the region.
- ➢ Automation of disaster reporting using datamining is currently a very difficult task to be added in any database feeding process.

Session E. Applications of disaster data

Use of CRED-EMDAT at USAID/OFDA : Sezin Tokar

- ➢ To be used for decision making, databases have to meet a large number of criteria (for example reliability, accessible, etc.). It emphasises the difficulty of creating and managing a database that targets decision-makers.

- The fact that several extreme event databases exist is not a problem. Each one is useful in a particular context.
- Heatwaves are a challenge both for developing countries and developed countries. The number of affected people continue to raise.
- Defining thresholds for heatwaves is challenging (not only the temperature has an influence but also the humidity) and it is region-specific.
- Moral and cultural norms have a strong effect on the impact of heatwaves (for example Ramadan during heat-wave).
- Heatwaves particularly affect Africa but for large territories almost no information on the impact of heatwaves is available. It is a problem but also a huge opportunity to improve the situation: Using meteorological data it is possible to define regions where it exists a high probability of heatwave impact on the population.

Q&A

- It was recognised that heatwaves are one of the most challenging extreme events to evaluate its impact. Non-direct related deaths are difficult to estimate. It requires solid baselines to have an estimation. However, the most affected regions (example: Africa) are also the ones where we need more precise statistics for this kind of assessment. It also requires "deep investigation" to confirm the fact that the heatwaves caused death and it is impossible to do so for large areas.

Challenges and opportunities for the spatial analysis of georeferenced weather-related disasters:
Sophie Vanwambeke

- The use of spatial data continues to increase. It takes an important place in the analysis of extreme events.
- Even if we use more spatial data than ever, the uncertainty that comes with it is still an important issue.
- Current software let us integrate a various type of spatial data together for analysis. However spatial data are recorded at a specific scale and a lot of users don't take it into account.
- Spatial data are not just spatial. Like other data, it is an imperfect representation of reality at a given time. The situation described by the data change over time.
- Comparing databases is challenging. They use different definitions and thus the concept behind the data vary.

Added value of citizen science to record extensive disasters : Matthieu Kervyn

- Extensive disasters are usually below the threshold to be recorded in most global databases (for example in some regions: flash floods and landslides).
- Even if not recorded in global databases, extensive disasters impact local populations.
- Citizen science could be one of the solutions to recorded extensive disasters.
- Africa has a high vulnerability to extreme events but the impact measured is low. Face to this situation, it is possible to question the way we measure this impact.
- It exists a huge underestimation of fatalities related to extreme events (factor 10 of even more form some regions).

- Crowdsourcing has multiple biases but it is one solution to recorded landslide in remote areas.
- Datamining (example: text-mining on social media) is another solution but different challenges (language) makes it difficult to implement it.
- Local people are in the best position to record landslides with a low severity but high frequency.
- Several people were trained and equipped with smartphones to go on the field and recorded landslides. They are called "Geobservers".
- Validation is an important step in the recording process.
- Citizen science is not a perfect solution but provided data never recorded before. Moreover, it helps also increase the awareness of the local community to this specific problem.

Q&A

- Once again the difficulty to monitor and recorded heat-waves was raised. Local authorities see these gaps between the reality on the field and what is recorded into global databases.
- Even in developed countries, public health warnings related to heat-wave are not respected by part of the population (example: Netherlands with young people still doing exercises during a heatwave).
- The process of validation regarding fatalities related to heatwave takes a long time.
- Recording spatial data can be challenging when the georeferencing is based on location names.
- This time again, text mining on social media for extracting useful information to monitor extreme event was recognised as being a challenge.

Access to Data: Regina Below
- The database license was established in 2016 by UCLouvain legal department
- The license was revised in 2019
- User need mandatory registration to access the search tools and data
- Different levels of access are defined such as: free or extended authorized use and commercial use (with conditions). Extended and commercial use are accessible for paid user only.
- Users are categories into: Group A with free access to 1000 records and annual subscription fee of 6000 Euros, while Group B uses are mostly academic, universities and non-profit research institutions with free access to 8000 records and annual subscription fee of 600 euros. Group C uses are private firms, consultancy companies, and profit organization. These group are granted free access to 1000 records and annual subscription fee of 6000 euros
- So far 75% of users are from group B.

Q & A

- Different opinions were expressed on how EM-DAT is distributed and its accessibility. Some participants asked whether CRED planned to distribute EM-DAT data free of charge and in open access
- The economic distribution model of EM-DAT was discussed. EM-DAT doesn't generate a substantial amount of money. The monetisation of such a database is a challenge. Other large companies are facing the same challenge

- Currently, USAID is the only sponsor/contributor of EM-DAT. The question of the viability of EM-DAT without USAID was raised. In such a situation, other organisations likely to support the EM-DAT programme were suggested: European Commission, UNDRR
- A group of co-operators/consortiums would ensure the long-term viability of the EM-DAT project
- Following a full-access request may be time consuming if the potential subscribers try t negotiate the agreement or postpone the decision several times

EM-DAT case-studies: Joris van Loenhout and Maria Moitinho

- The key objectives of case studies are to increase resolution of disaster impact at lower level, foster collaboration with local and international partner, strengthen the ability of national and regional authorities to collect accurate and complete disaster data
- CS1 Philippines (Typhoon Haiyan): one of the main issues highlighted in the study was inadequate to access.
- CS2 Nepal (2015 earthquake): From field missions, studies on the modification of hospital stays durations and admissions compared to the pre-EQ situation and gender differences in the type of fractures. Qualitative analysis is ongoing.
- CS3 Mozambique (Severe floods): Difference-in-difference analysis hospital admissions for all causes and diarrhea-specific in affected district vs hospital in non-affected district, before and after the cyclones. Full report in Sept. 2020.

Q & A
- Current information sources do not provide this specific information. Case studies are a way of collecting this gender-specific information, even if it cannot be easily extrapolated to a larger scale.

Recent and planned products: suggestions from TAG: Debby Sapir and Regina Below
- Publications: CRED Crunches, scientific articles, CRED/UNISDR report
- Natural Disaster in number provides based on info-graphic, accessible to the public and has recorded more than 8500 downloads
- Atlas project is on-going and EM-DAT guidelines current under review for improvement
- The goal of the EM-DAT Atlas to provide interactive maps for users, give users overview of geo-referenced data

Q & A
- International and national borders are a sensitive subject when it comes to mapping data. UN border files should be used to avoid any conflict with data users. When mapping EM-DAT data, the kernel density estimation map is a solution to avoid displaying part of the boundaries.
- TAG is currently only a meeting of the most frequent users, but there is also an opportunity for this group to become more involved in the evolution of EM-DAT in the future.