UCLouvain
Institut de Recherche
Santé et Société (IRSS)

Centre for Research on the
Epidemiology of Disasters
CRED

USAID
FROM THE AMERICAN PEOPLE

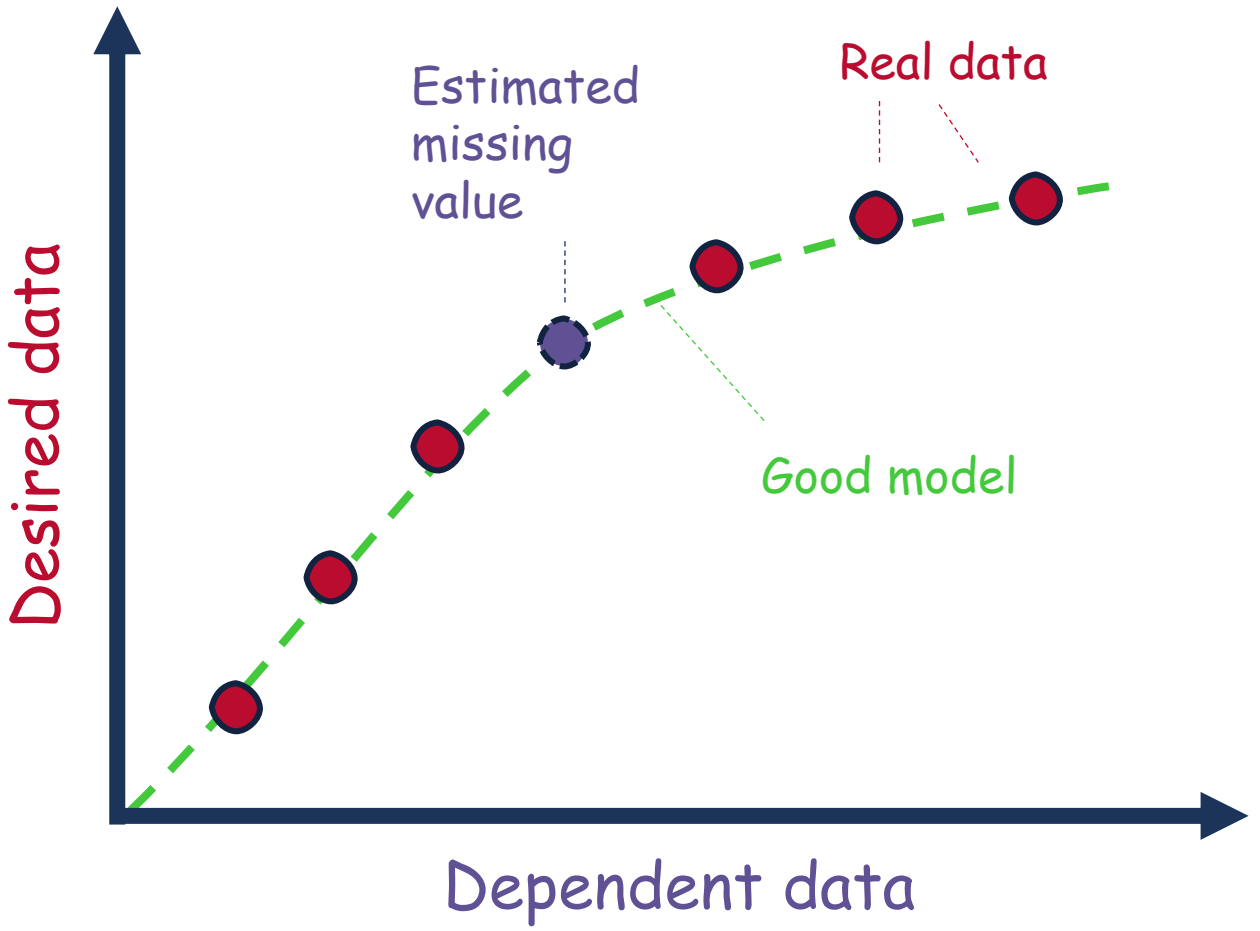
Biases and Missingness in EM-DAT Disaster Data:
Recent Technological Opportunities

Scientific & Technical Advisory Group Meeting

Damien Delforge (UCLouvain, CRED)

Brussels, Belgium
March 21st, 2023

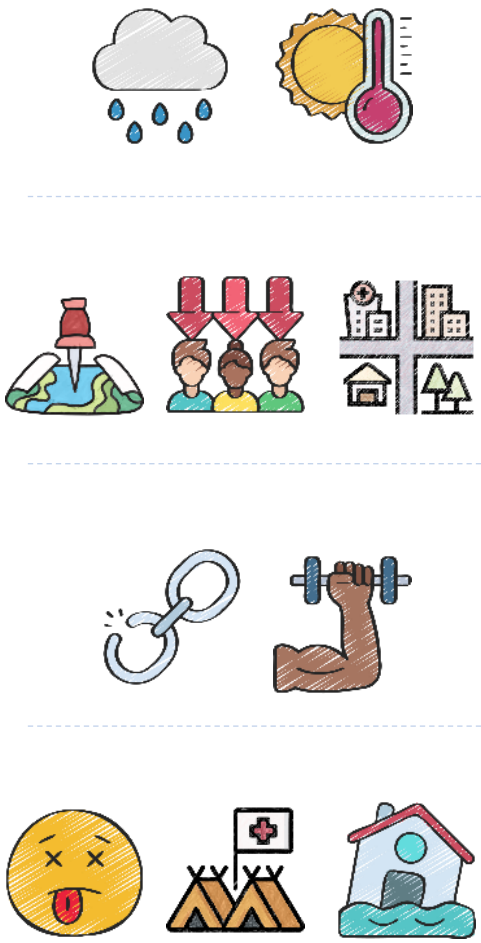
Missingness is Manageable with Sufficient Data



Quantitative approach

- ↑ Desired data
- ↑ Dependent data
- Improve Model

The Ideal Situation for the Disaster Risk Community & EM-DAT



H: Hazard

E: Exposure

V: Vulnerability and

C: Capacity

R: Disaster Impact

(or Risk)

Closing the

The risk equation

$$R = f(H, E, V, C)$$

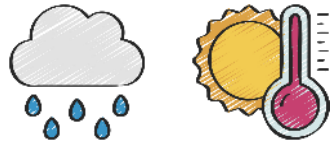
Model

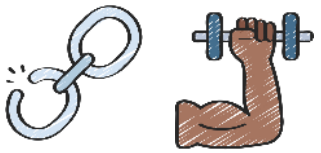
EM-DAT Desired Data

Dependent data

Icons from www.flaticon.com

The Common Situation for the Disaster Risk Community





H: Natural Hazard

E: Exposure

V: Vulnerability and C: Capacity

R: Disaster Impact or Risk

Unbalanced

Risk equation

$$R = f(H, E, V, C)$$

Biased or Unrealistic Conclusions

Poor Data Coverage, Quality, and Accessibility

Missing information

Icons from www.flaticon.com

1. EM-DAT Missingness and Biases

1 The Main Reason behind EM-DAT Missingness and Biases

EM-DAT records what it get from its sources

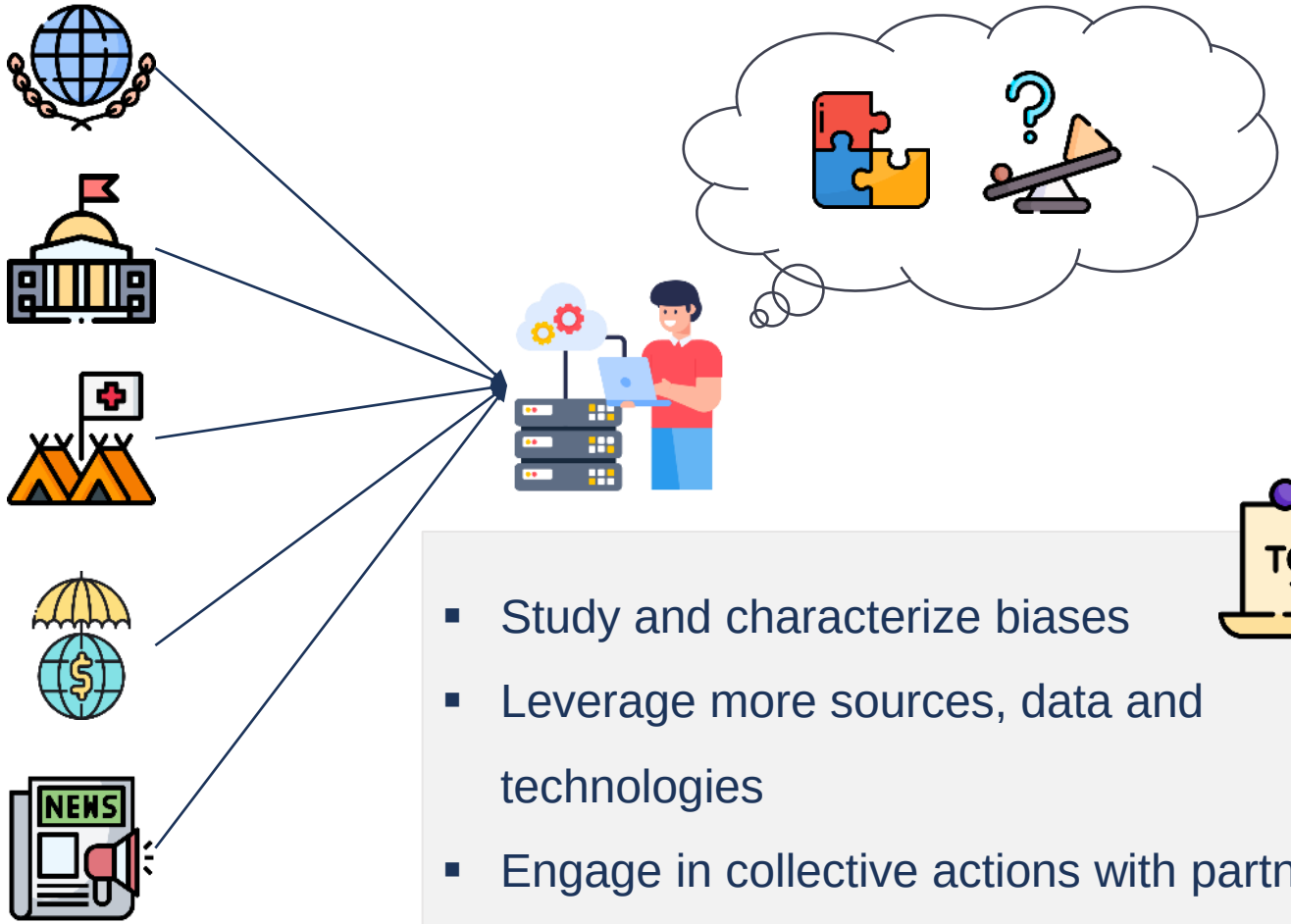
United Nations Agencies, World Bank

National governments, US, EU

Humanitarian agencies (e.g., IFRC)

Re-insurance companies (e.g., AON)

Press Agencies (e.g., AFP)



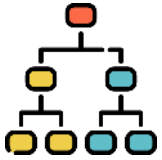
- Study and characterize biases
- Leverage more sources, data and technologies
- Engage in collective actions with partners

Icons from www.flaticon.com

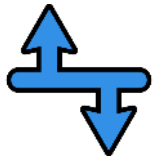
1 Biases Are of Varying and Intertwined Types

Known Biases in Disaster Loss Databases (adapted from Gall et al., 2009)

1. Time Bias



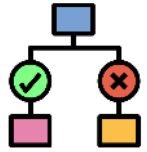
2. Hazard-related Bias



3. Threshold Bias

4. Accounting Bias

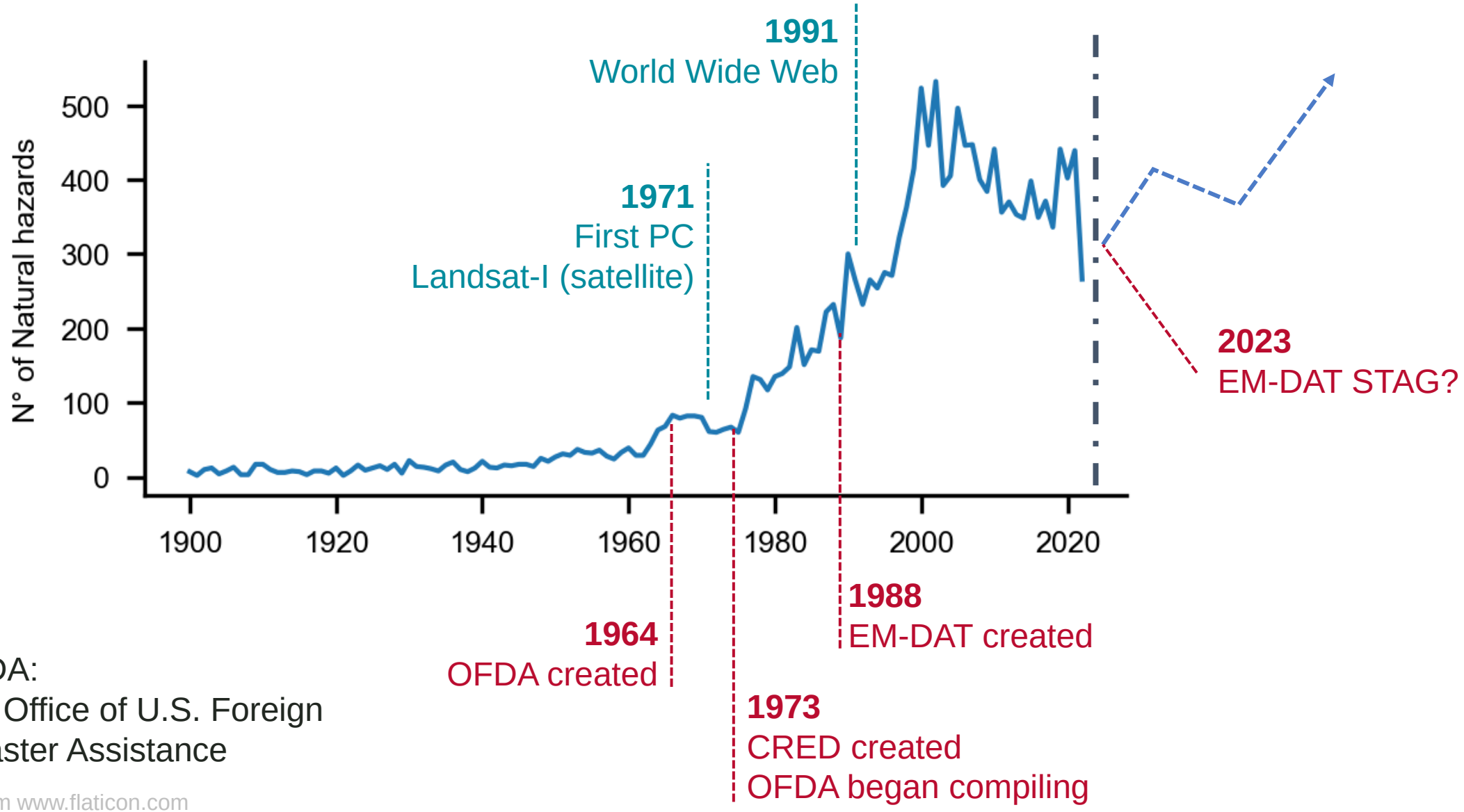
5. Geographic Bias



6. Systemic Bias

Icons from www.flaticon.com

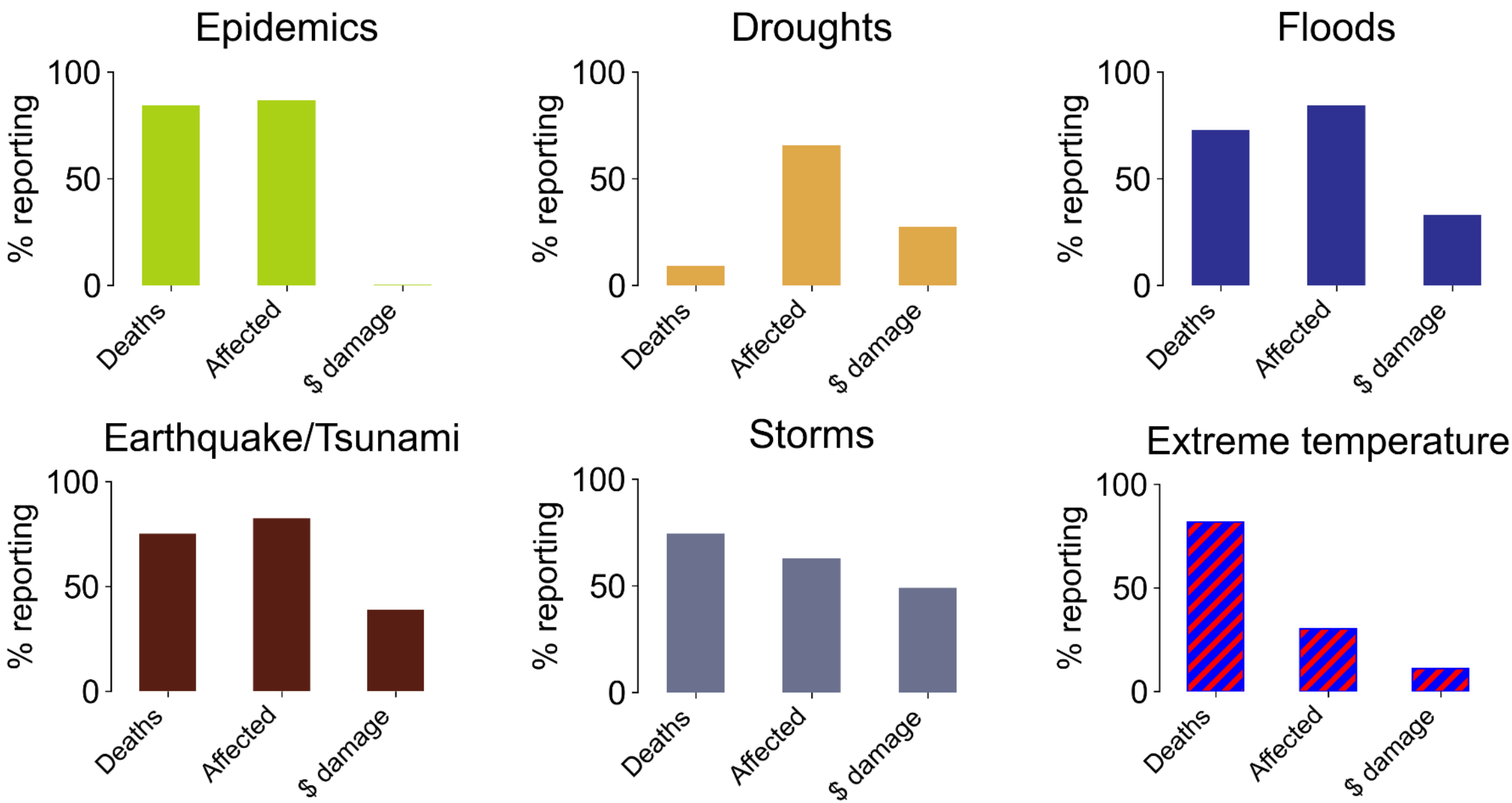
1 Time Bias: Technology and Initiatives Lead the Trend



OFDA:
The Office of U.S. Foreign Disaster Assistance

Icons from www.flaticon.com

1 Hazard-Related and Accounting Biases



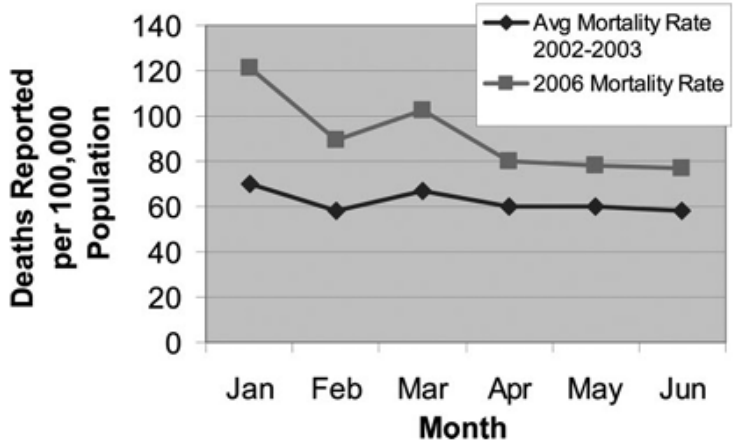
1 Accounting Bias: Direct vs Indirect Mortality

EM-DAT tends to record direct losses (unless sources include indirect losses)

Hurricane Katrina (2005-08)
1,836 deaths (EM-DAT)

Comparison of average mortality rates between January–June 2006 and January–June 2002–2003, *Times-Picayune* death notices (New Orleans)

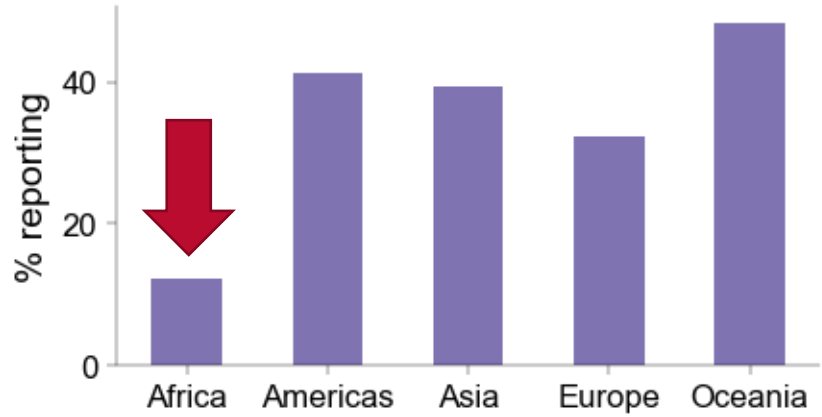
+ excess deaths (+50%)



(Stephens et al., 2007)

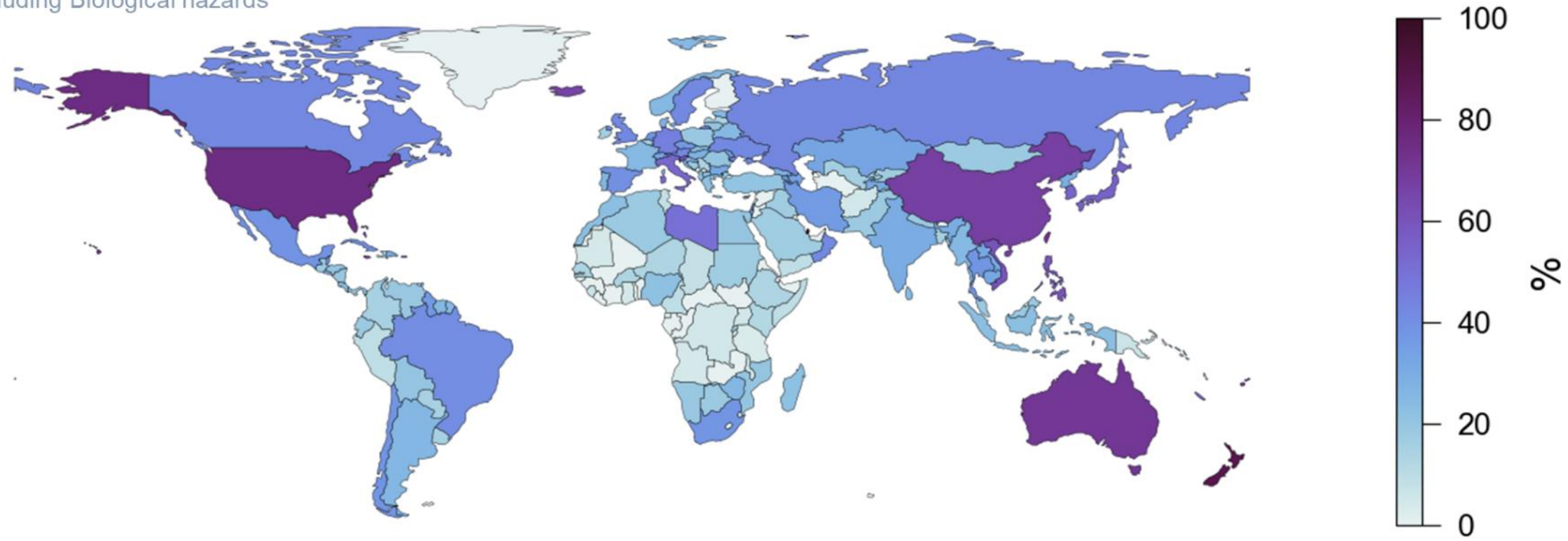
Icons from www.flaticon.com

1 Accounting & Geographic Bias



Reporting percentage of economic losses for natural hazards in EM-DAT (2000-2022)

Excluding Biological hazards

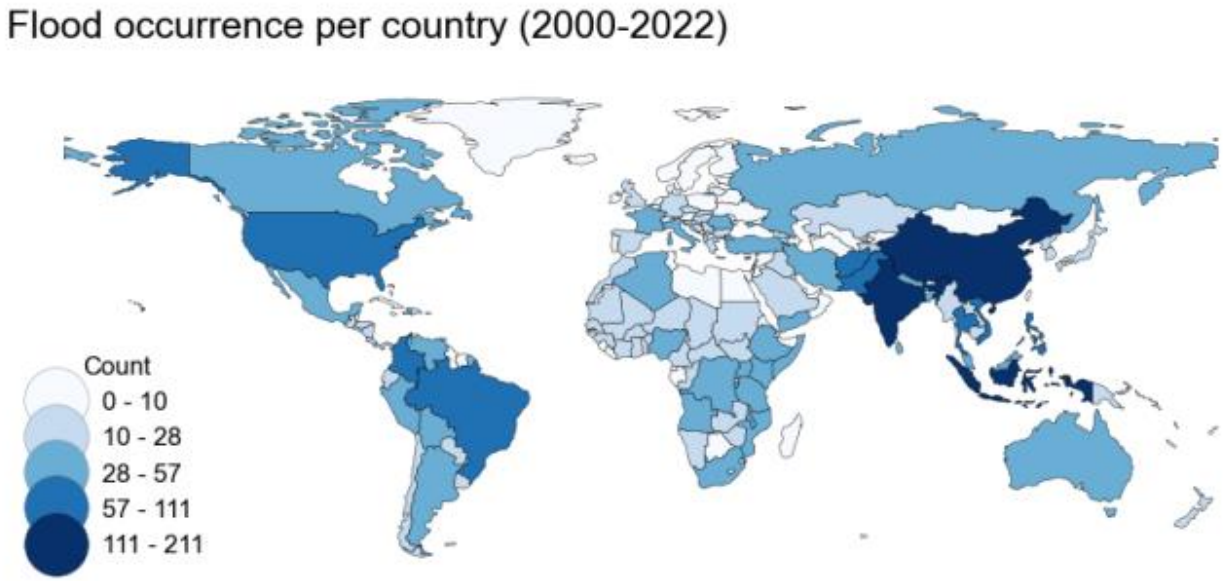


1 Systemic Bias: Administrative units

Admin units = Biased perception

+ Poor and variable spatial resolution
+ Hard to estimate exposure, vulnerability, …, risk

Example of Country-level map



1 Systemic Bias: Extreme Temperature

Extreme temperatures kill 5 million people a year with heat-related deaths rising, study finds

More people died of cold than heat in past 20 years but climate change is shifting the balance

The sun rises over Melbourne on a scorching day. Deaths linked to hot temperatures are on the rise, a global study has found. Photograph: David Crosling/AAP

The Guardian – 2021-07

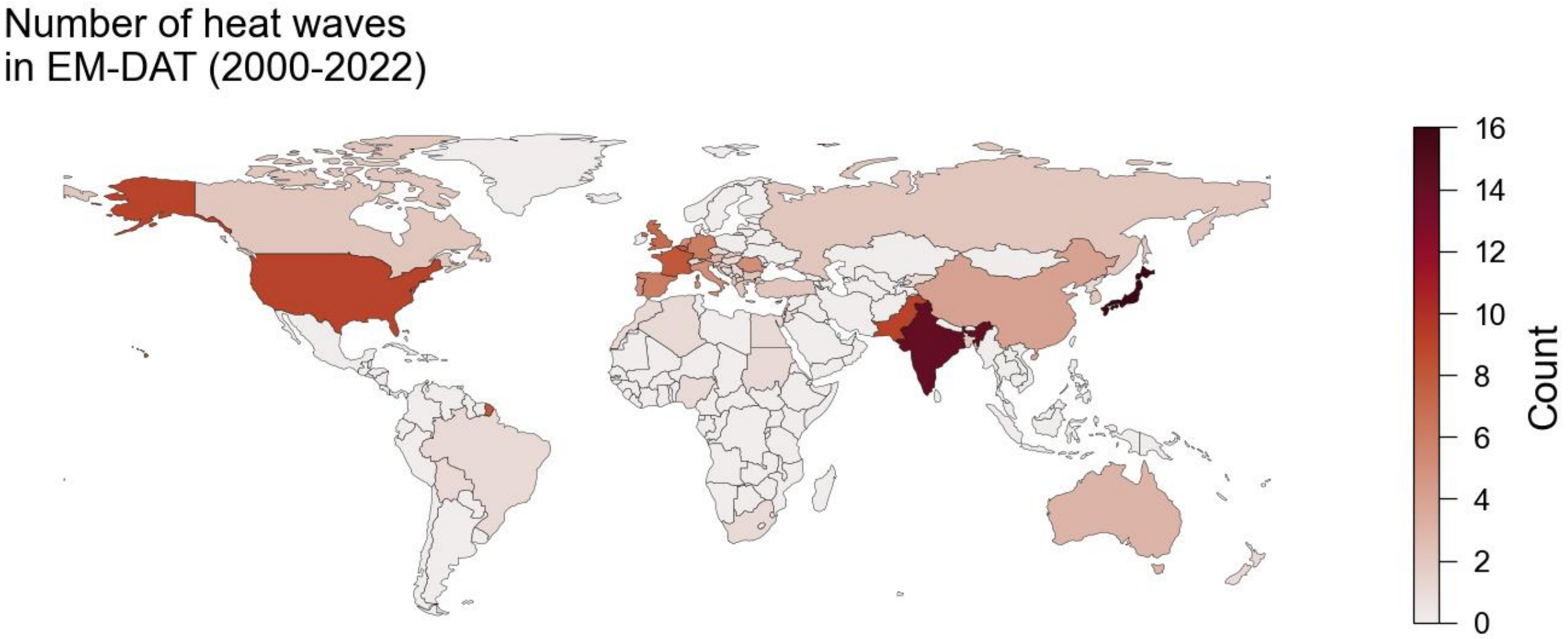
The Numbers (Zhao et al., 2021)

- 5.1 Million deaths/year associated with sub-optimal temperatures
- 0.5 M death/year (heat-related)
- 4.6 M death/year (cold-related)

EM-DAT (2000-2022 average)

- 8,300 deaths/year for extreme temperature events
- 7,600 death/year (heat wave)
- 600 death/year (cold wave)

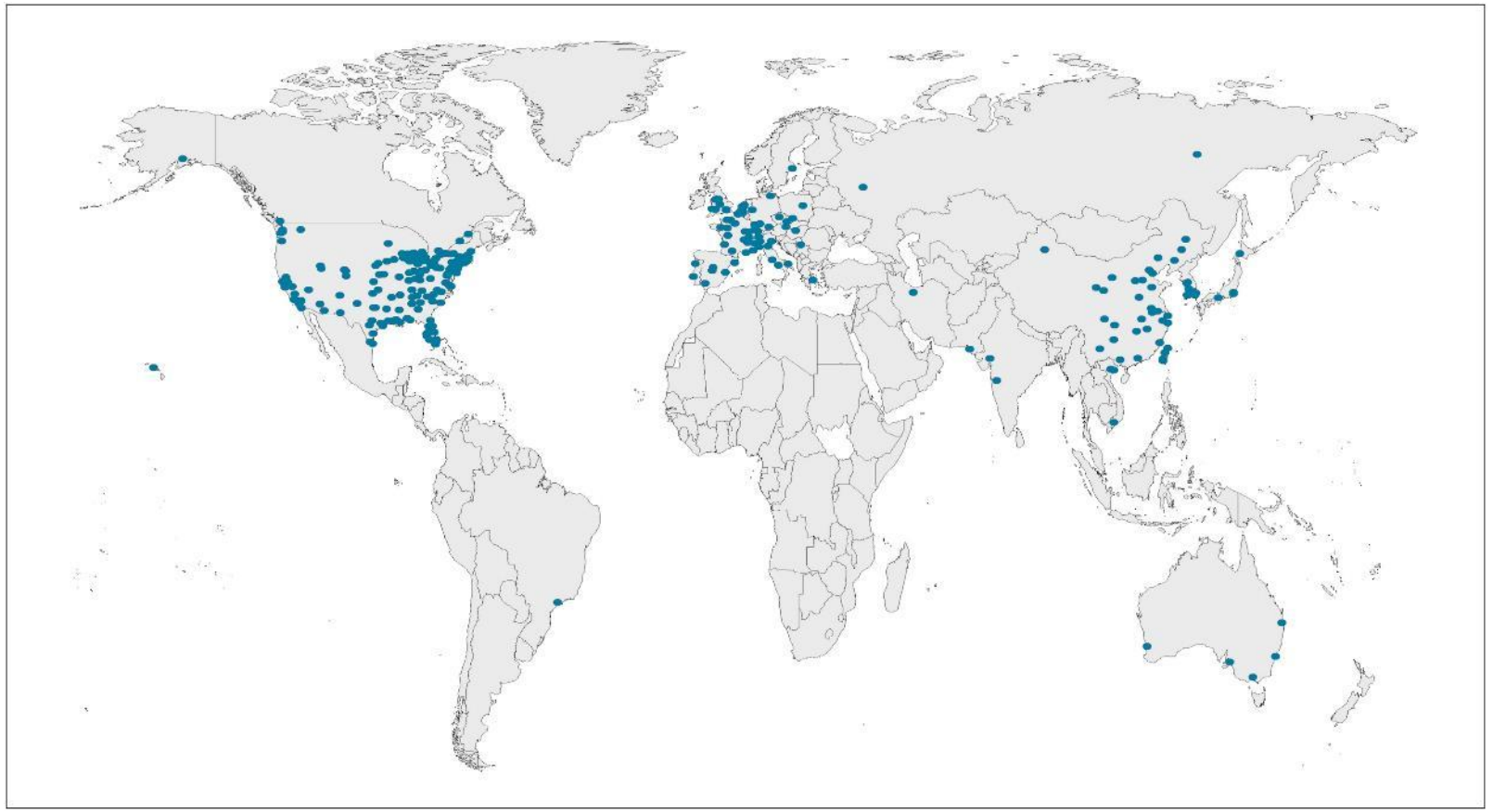
1 Hazard-related & Geographic Bias: Heat Wave



50% of heatwave events in 9 countries

1 Geographic Bias: Heat Wave - Health Scientific Research

Heatwave and health impact research: A global review (Campbell et al., 2018)

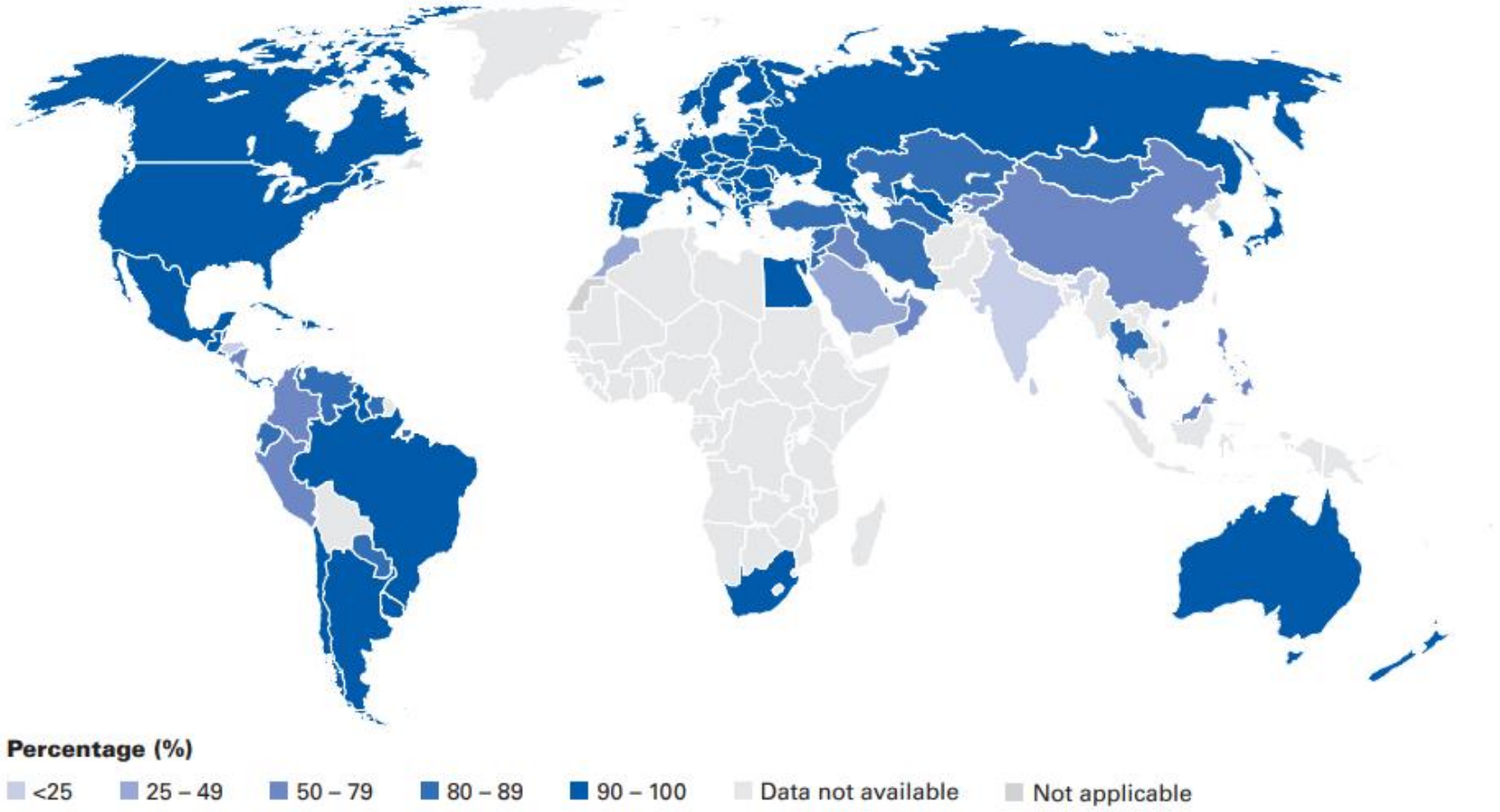


Locations of heatwave and health impact research, 1964–2017.

https://doi.org/10.1016/j.healthplace.2018.08.017

1 Heatwave = Silent Killer

No Mortality Data ➔ No Heat Wave Disaster



Civil registration coverage of cause of death (%), 2007–2016 (WHO, 2018).

See http://gamapserver.who.int/mapLibrary/Files/Maps/Global_CivilRegistrationDeaths_2007_2016.png

2. Technological & Data Opportunities

2.1 Using Geographical Data

2.2 Using Online Media

2.3 Using other Communication Systems

2.1 1973 Mississipi floods captured by ERTS-1 (Landsat 1)

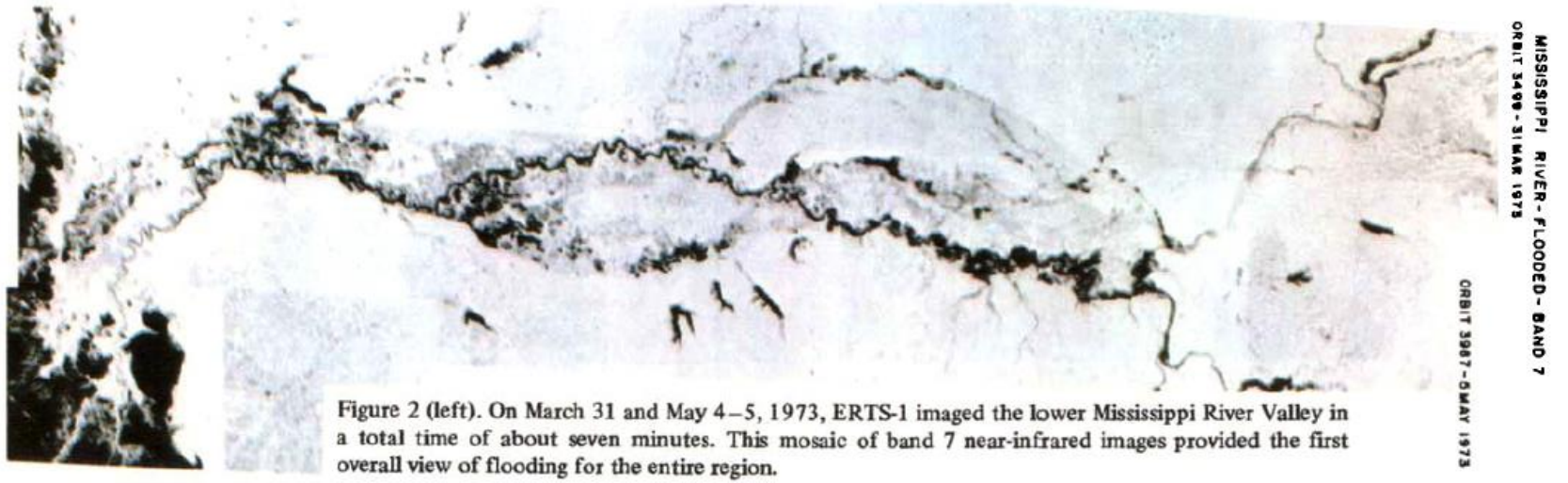


Figure 2 (left). On March 31 and May 4–5, 1973, ERTS-1 imaged the lower Mississippi River Valley in a total time of about seven minutes. This mosaic of band 7 near-infrared images provided the first overall view of flooding for the entire region.

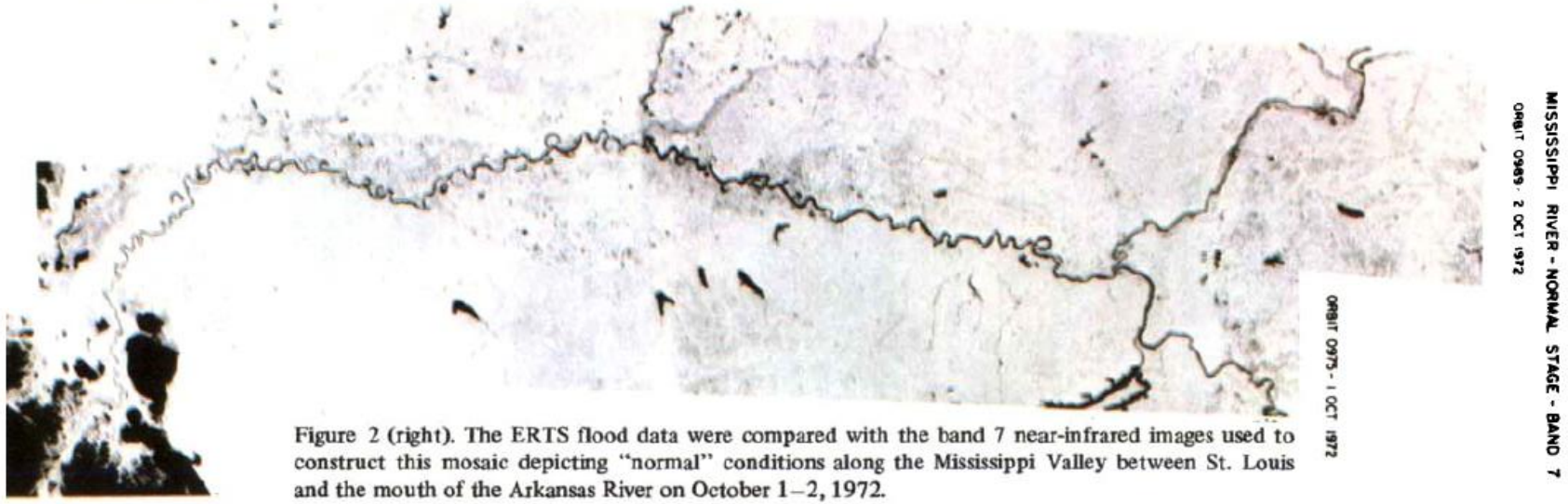


Figure 2 (right). The ERTS flood data were compared with the band 7 near-infrared images used to construct this mosaic depicting "normal" conditions along the Mississippi Valley between St. Louis and the mouth of the Arkansas River on October 1–2, 1972.

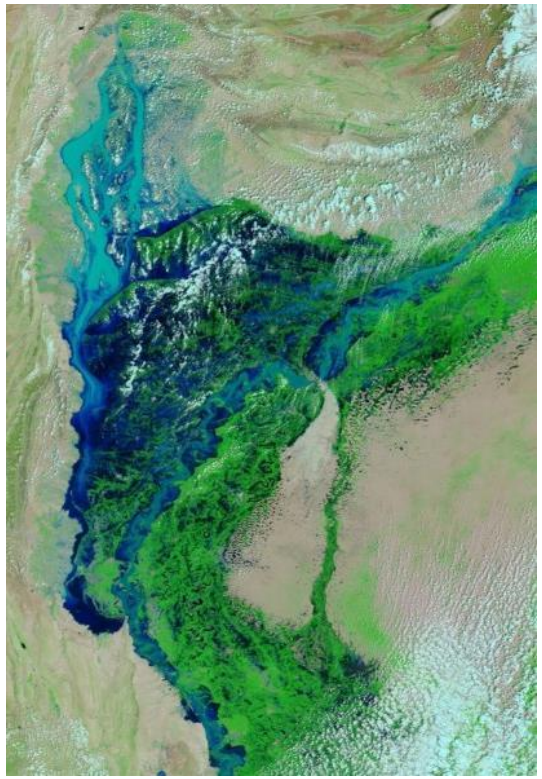
ERTS-1 / Landsat 1

Source: wikipedia

Deutsch and Ruggles (1974). https://doi.org/10.1111/j.1752-1688.1974.tb00622.x.

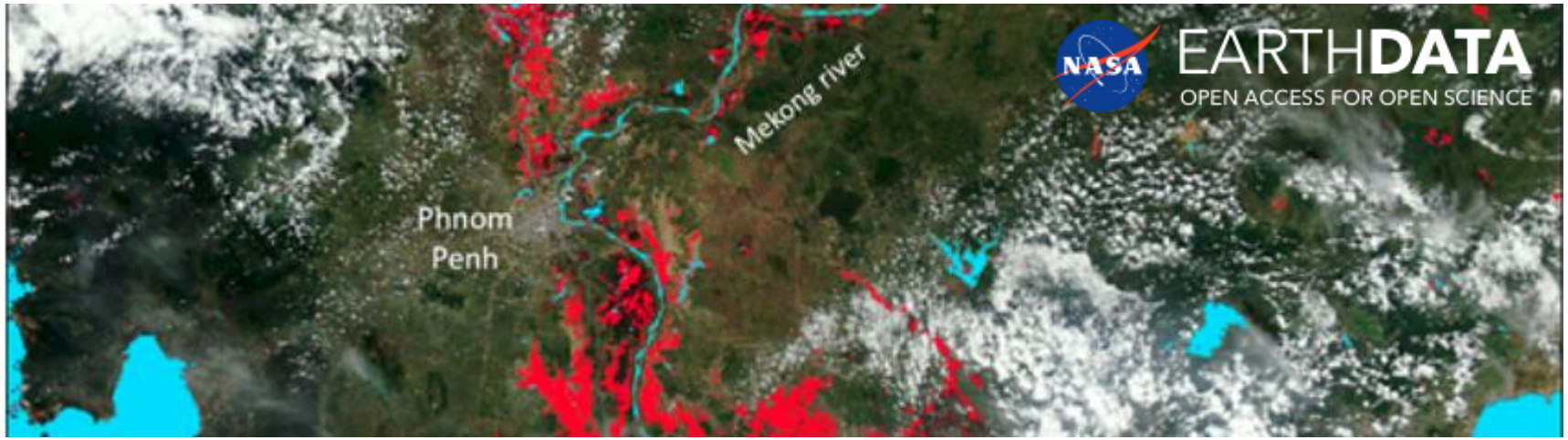
2.1 Twenty years of MODIS Imagery

MODIS
Floods
Pakistan 2022



Source: NASA

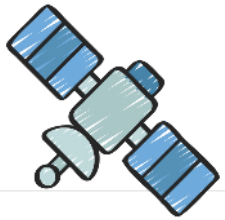
MODIS Near Real Time Global Flood Product



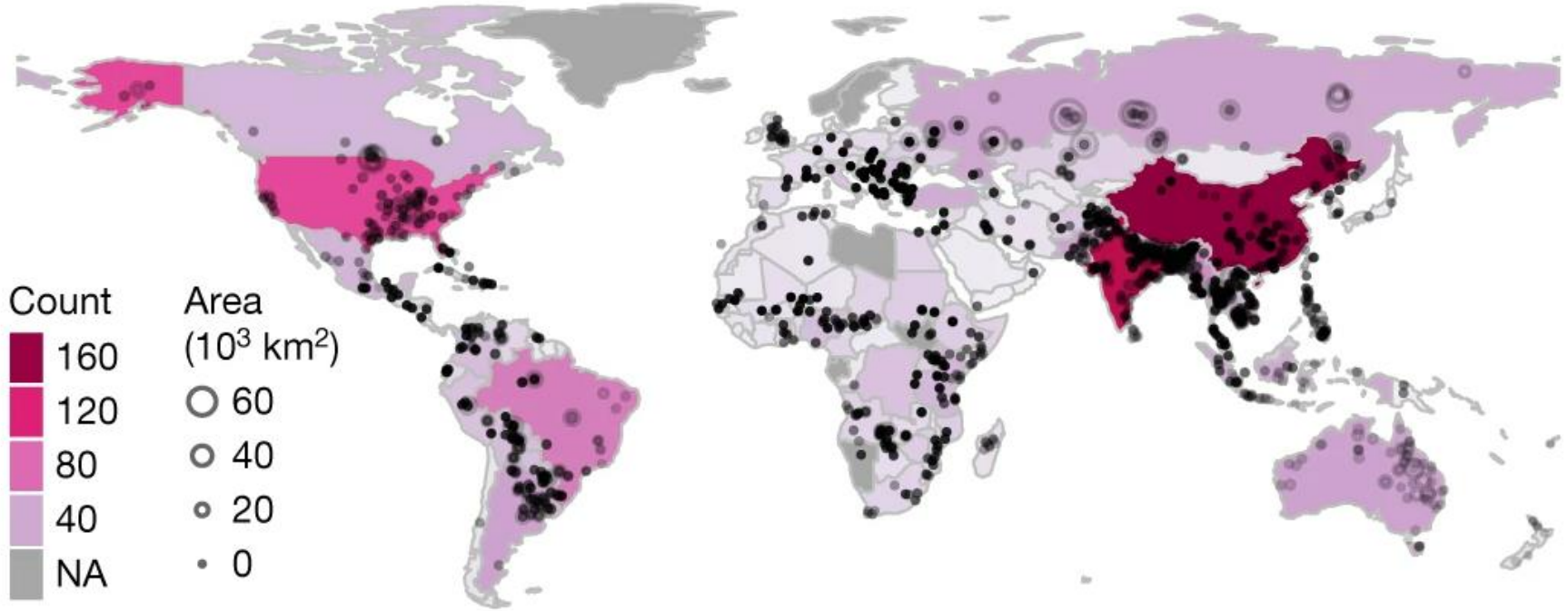
- Automated Water Detection Algorithm
- 1-2-3 days mapping product
- ~250m of resolution

MODIS = Moderate Resolution Imaging Spectroradiometer

2.1 Twenty years of MODIS Imagery



MODIS-extracted Flood **hazards** (2000-2015) + Exposed population



Google Earth Engine

Cloud to Street

DFO
Flood Observatory

MODIS = Moderate Resolution Imaging Spectroradiometer

From Telmann et al (**2021**): https://doi.org/10.1038/s41586-021-03695-w.

See website: https://global-flood-database.cloudtostreet.ai/

2.1 From Hazard to Exposure/Vulnerability with POPGRID

POPGRID DATA COLLABORATIVE

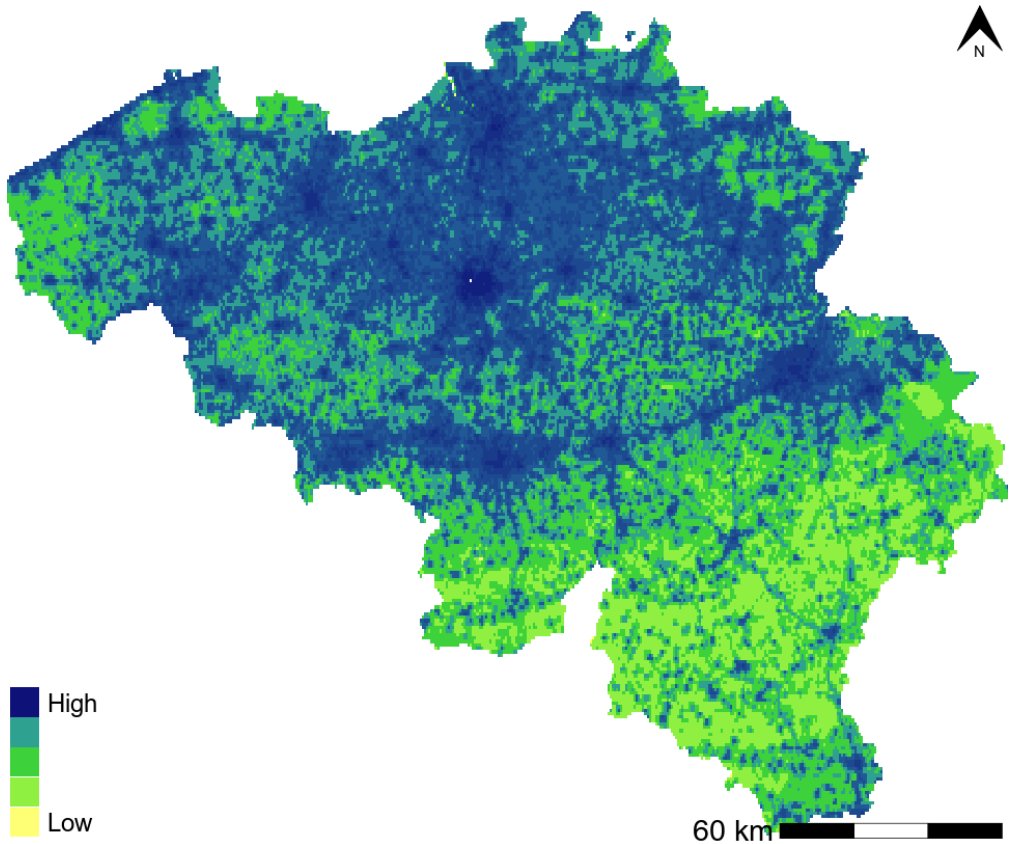
A Data Collaborative for Settlement, Infrastructure, and Population Data

Hazard footprint x POPGRID = Hazard Exposure

WorldPop (worldpop.org)

- High resolution population density
- Disaggregated (e.g., age, sex)
- Updated yearly since 2000

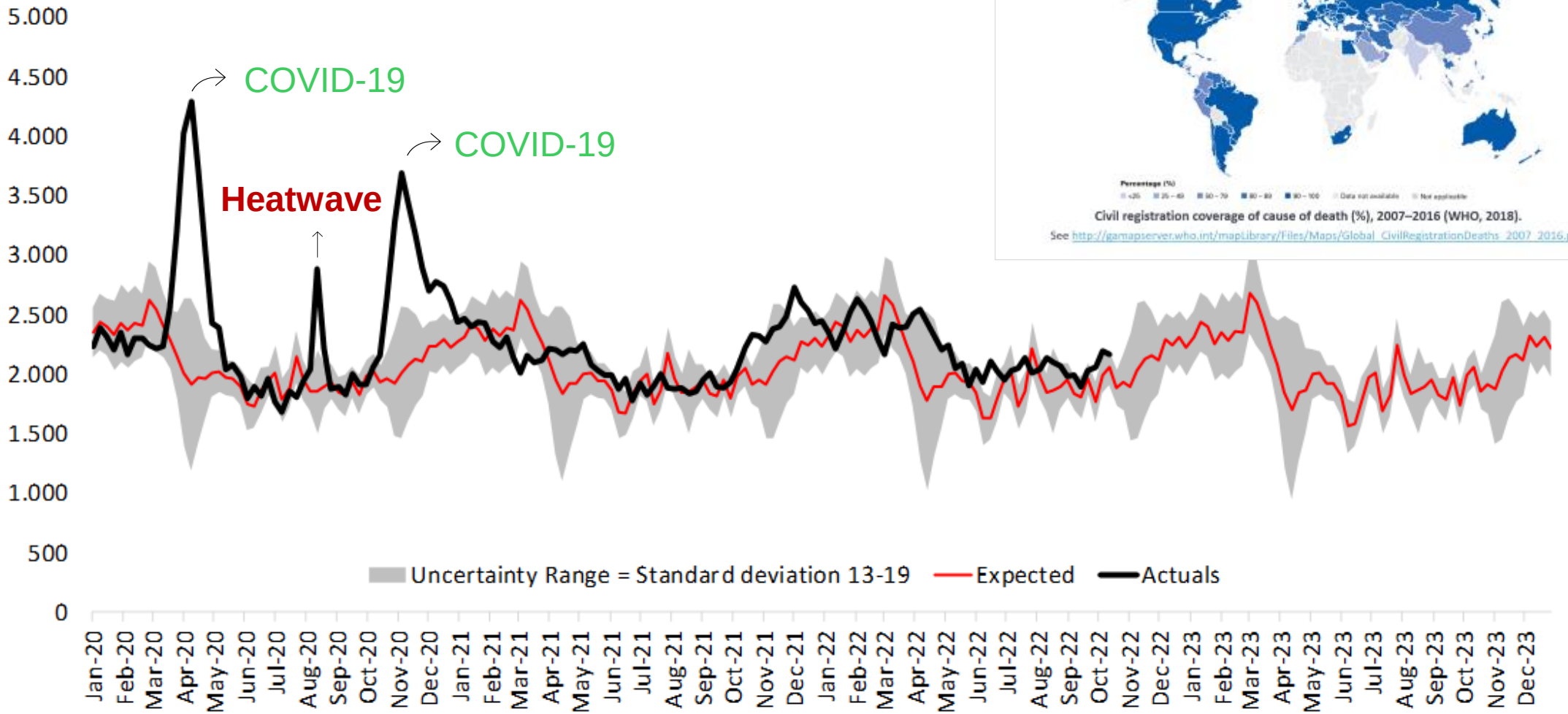
WorldPoP dataset
1km Population Density (Example for Belgium)



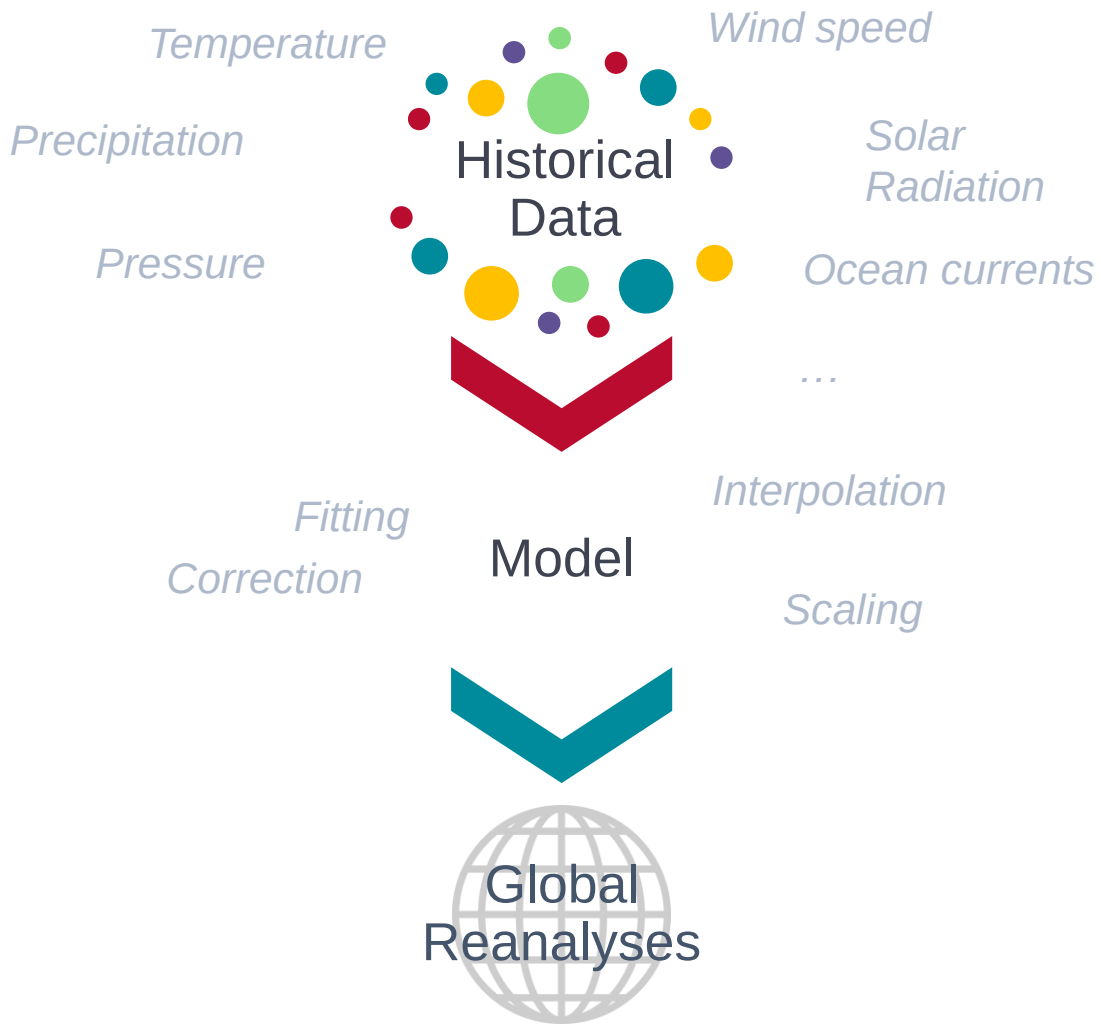
2.1 Excess Mortality Data: Disaster Detection & Mortality Impact

Example for Belgium

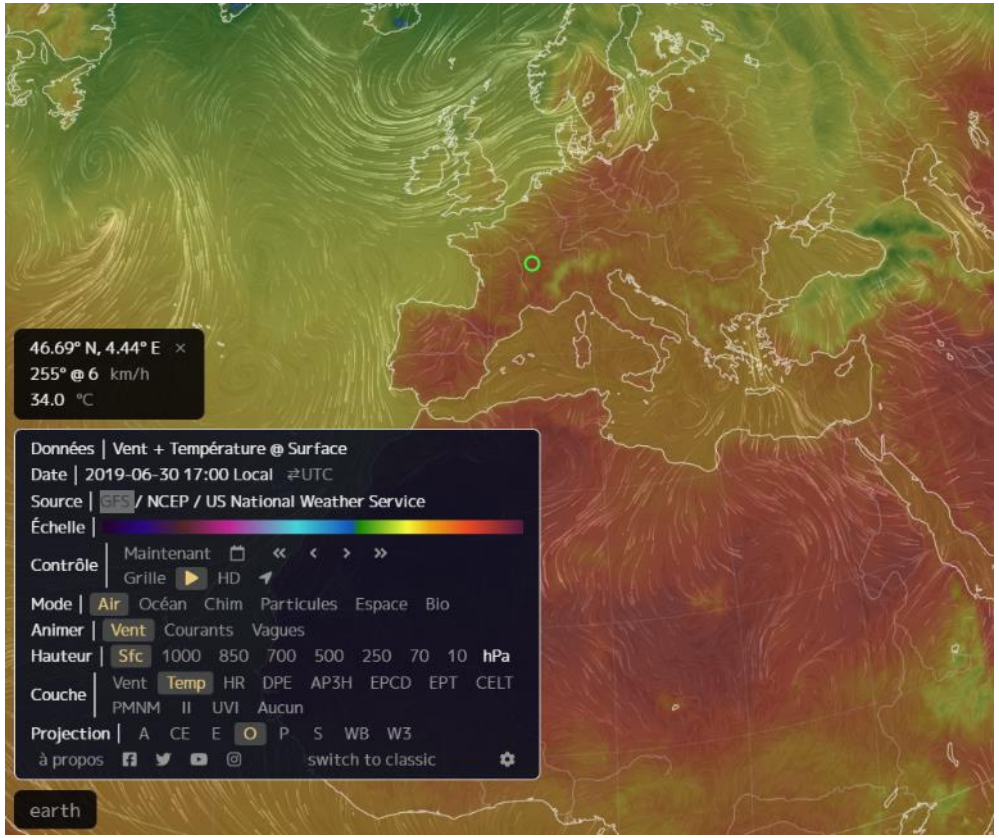
Source: www.sciensano.be



2.1 Weather-related Disasters Detection and Mapping with Reanalyses



Weather data anytime, anywhere!



National Center of Environmental Prediction (NCEP) reanalysis

https://earth.nullschool.net/

Weather-related Disasters Detection and Mapping with Reanalyses

NCEP-2 vs ERA-Interim

From Russo et al. (2014),
https://doi.org/10.1002/2014JD022098

3 periods:

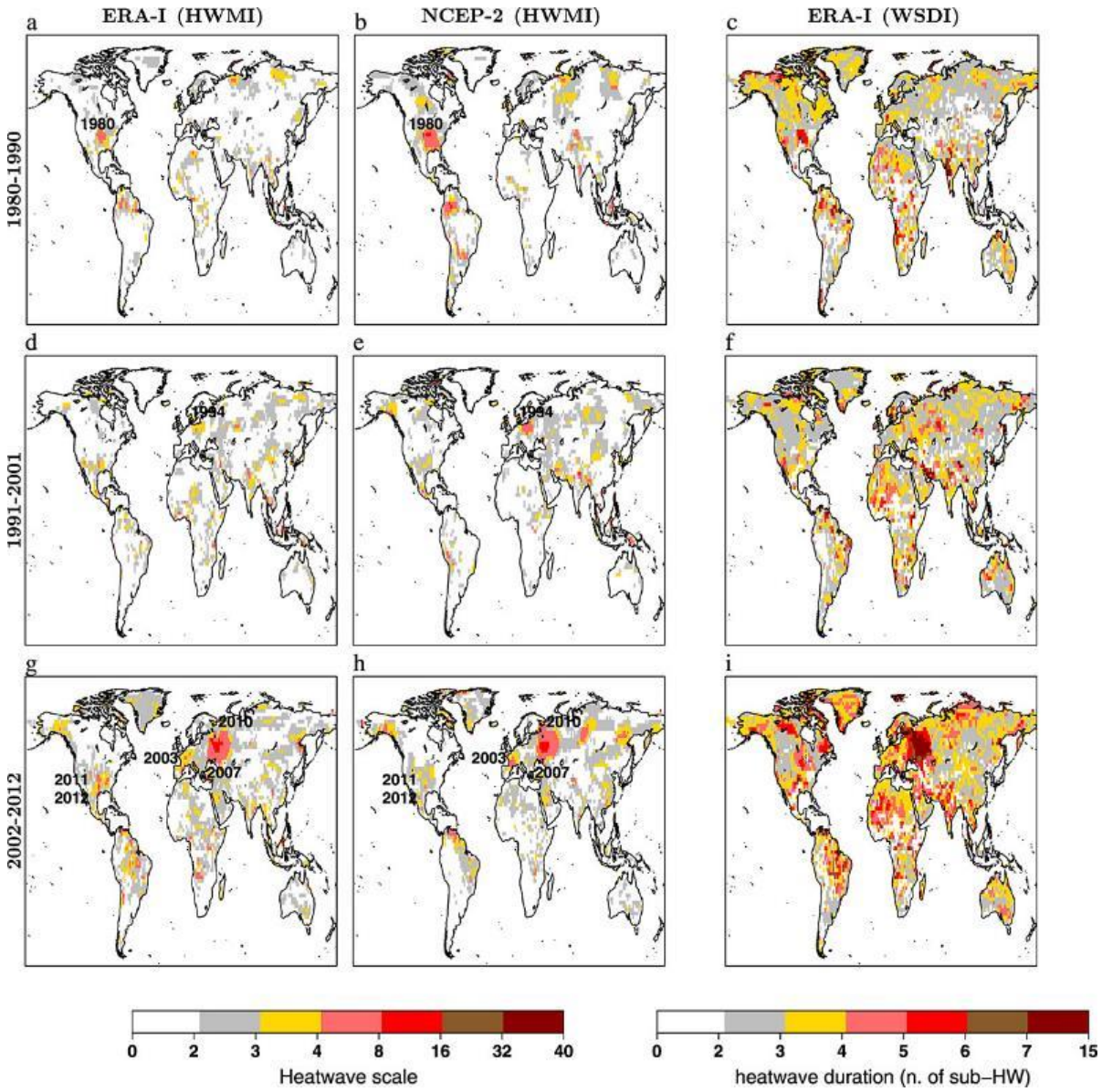
- 1980-1990
- 1991-2001
- 2002-2012

Heat Wave Magnitude Index

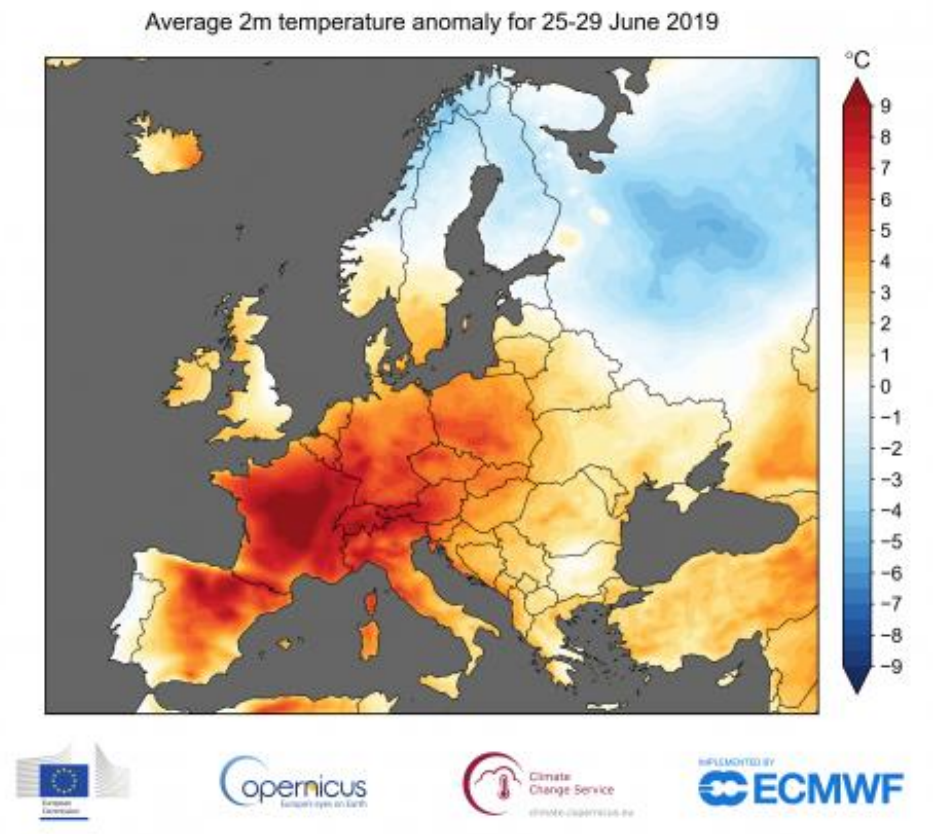
Warm Spell Duration Index

2003: >70,000 deaths

2010: >55,000 deaths



2.1 Heat waves: ERA-5 reanalysis vs EM-DAT (Belgium)

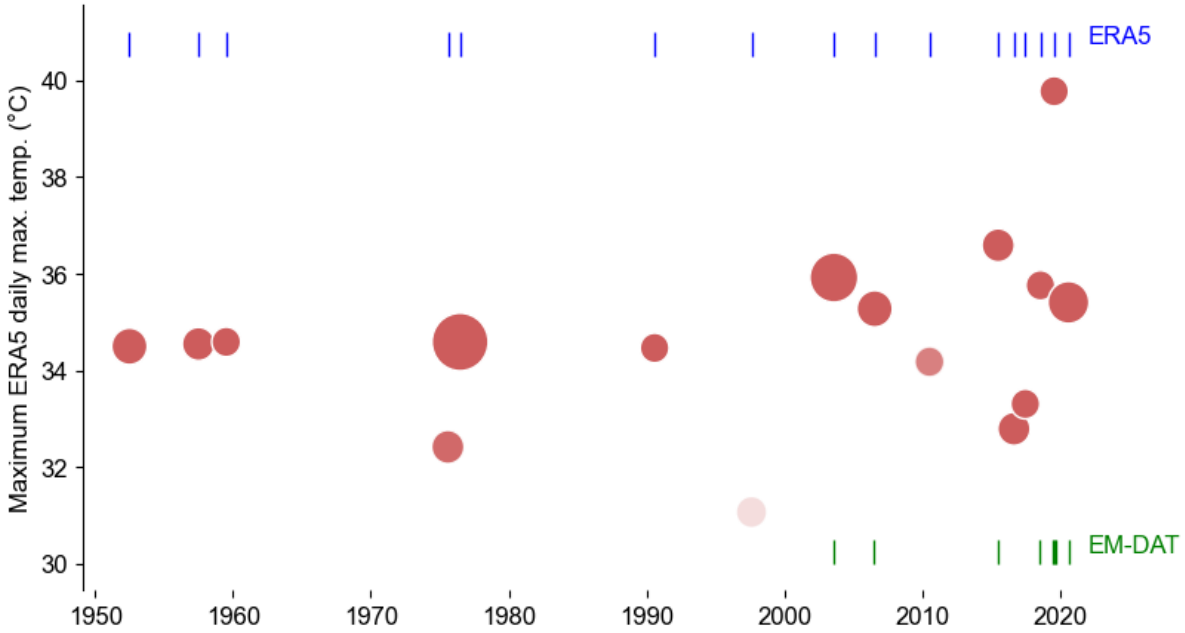


ERA-5 Reanalyses

- Many meteorological variables
- 1950 to present
- up to 0.1°x0.1°

- *Periods ≥ 3 consecutive days …*
- *> the 99th percentile …*
- *of the daily T_{max} …*
- *of the May-Sep season of the control period (1971–2000)**

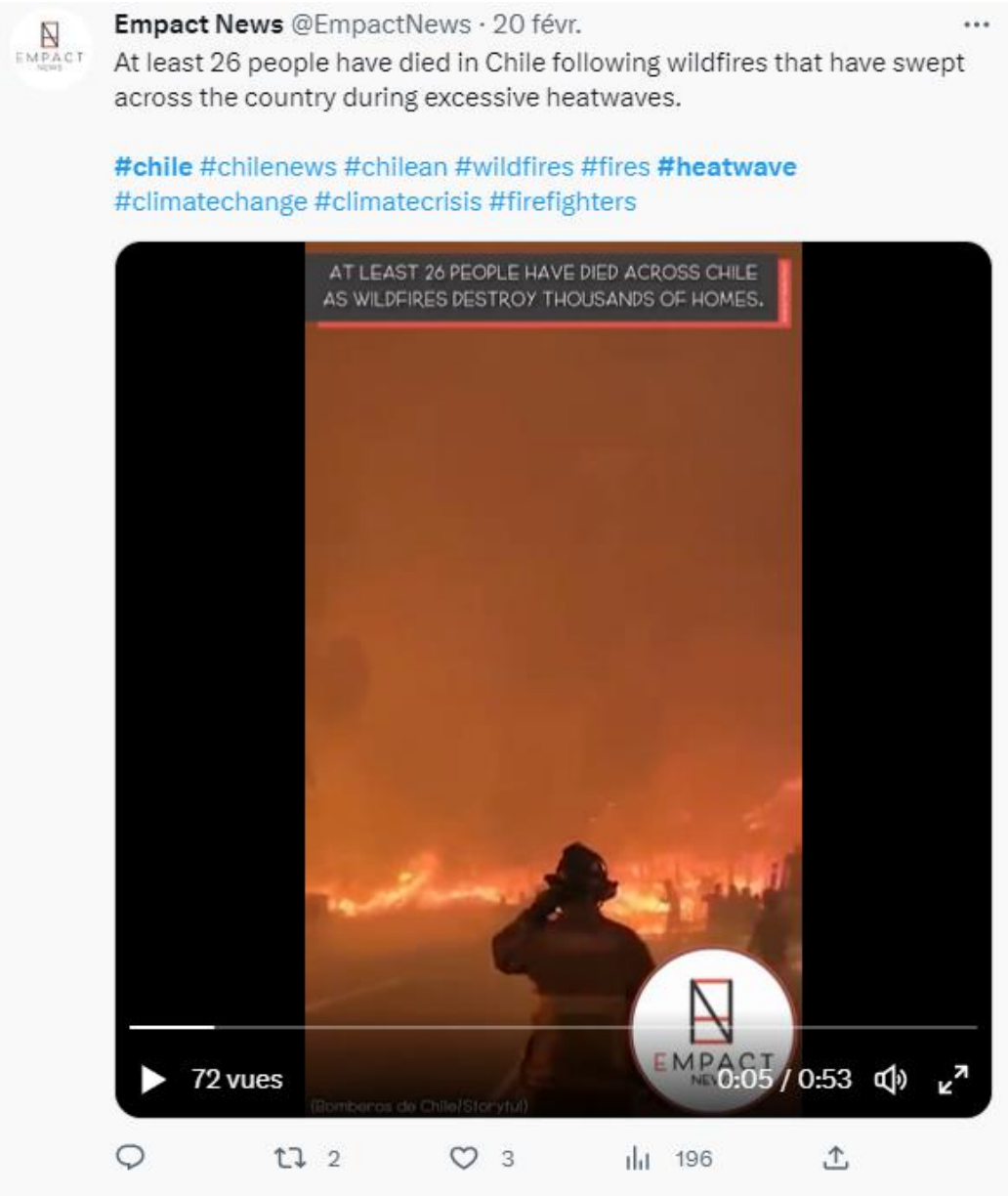
*Definition from Jacob et al (2014), https://doi.org/10.1007/s10113-013-0499-2



Heatwave duration;
Opacity = Belgium exposure

2.2 Online Media: Detecting & Mapping Disasters with Twitter

- Since 2005
- Easy query with hashtags and **Twitter** API
- Some tweets are geolocated



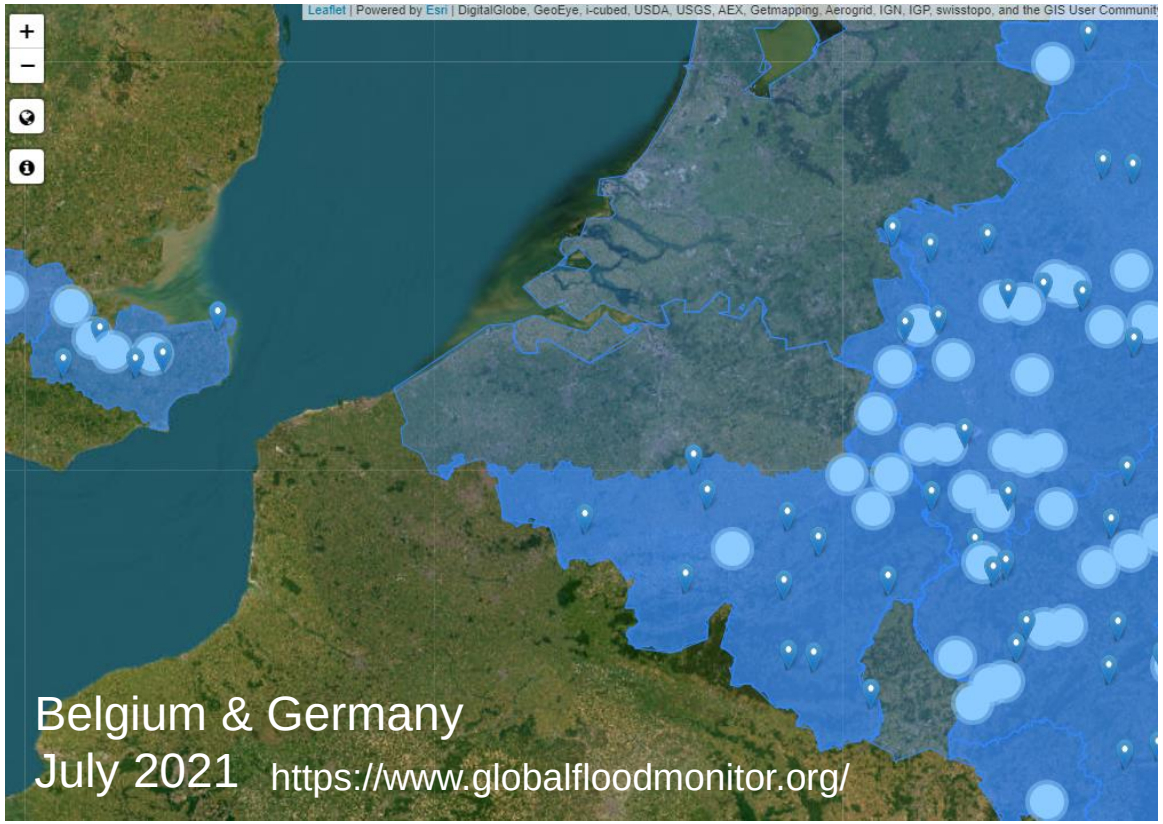
2.2 Detecting & Mapping Floods with Twitter

Global Flood Monitor*

Event detection based on

- Hashtag in multiple languages
- Tweet temporal anomaly
- at a location
- (No impact extraction)

*de Bruijn et al. (2019): https://doi.org/10.1038/s41597-019-0326-9



IVM Institute for Environmental Studies | VU | Munich RE

2.2 Event Detection Based on News (GDELT): Diseases

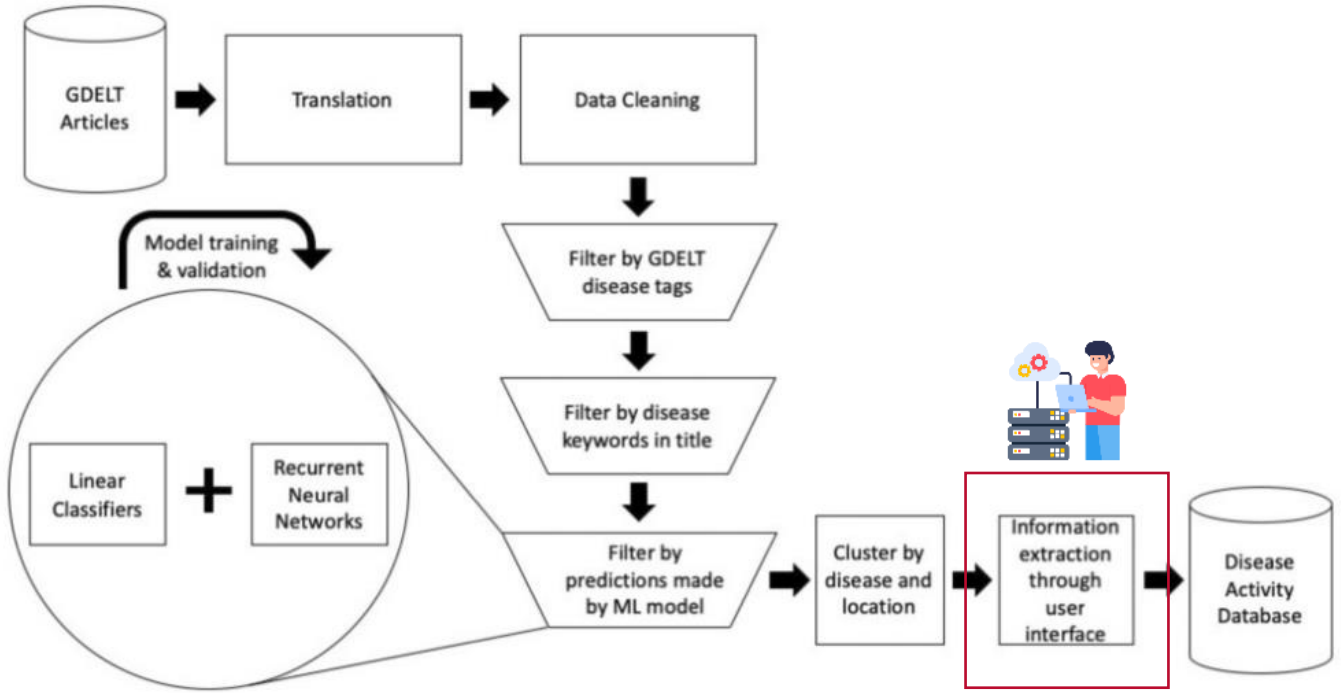
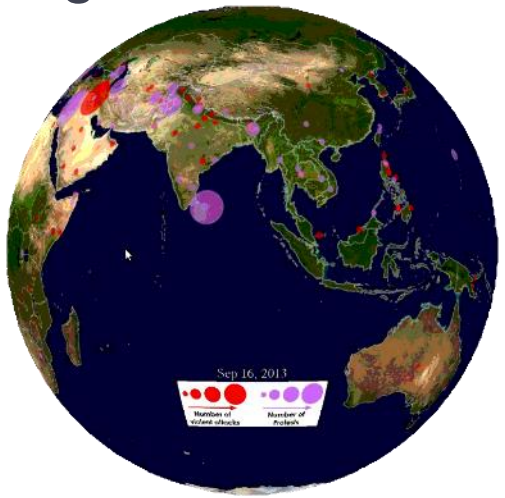


Figure 1. The modules of our event-based surveillance system.

Feldman et al. (2019)
https://doi.org/10.1093/jamia/ocz112.

- ✓ Global
- ✓ Pre-classified by tags
- ✓ Since 1979
- ✗ No automated impact data extraction

The GDELT Project Global Database of Events, Language, and Tone



GDELT monitors the world's news media from nearly every corner of every country in print, broadcast, and web formats, in over 100 languages, every moment of every day.

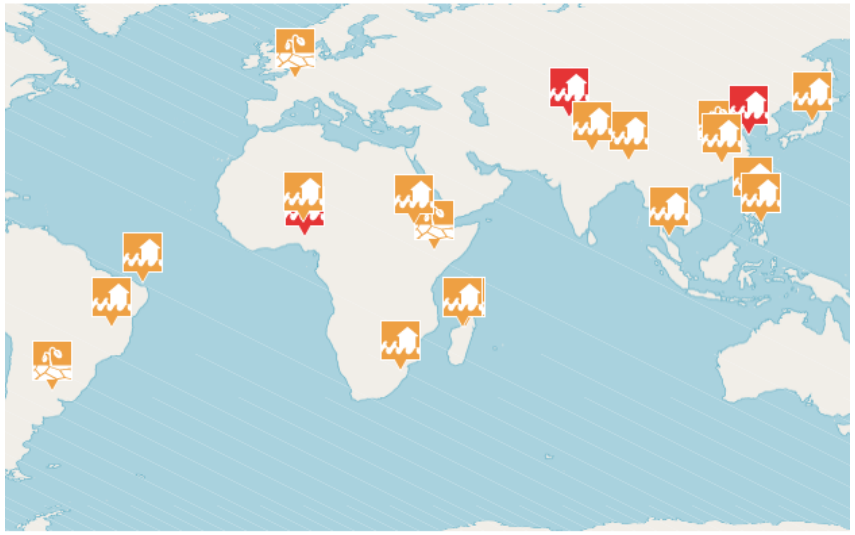
https://www.gdeltproject.org/

Google

2.3 Communication Systems: Use Data From Early Warning Systems

Example: Global Disaster Alert and Coordination System (GDACS)

Flood and Drought
Orange and Red alerts in 2022



https://www.gdacs.org/

Example: Nigeria Flood 2022

Overall Red alert Flood for Nigeria
in Nigeria

Summary | Impact | Maps & Reports | Media | Resources | Covid19

Event summary

Flood **Nigeria** can have a high humanitarian impact based on the magnitude, exposed population and vulnerability.

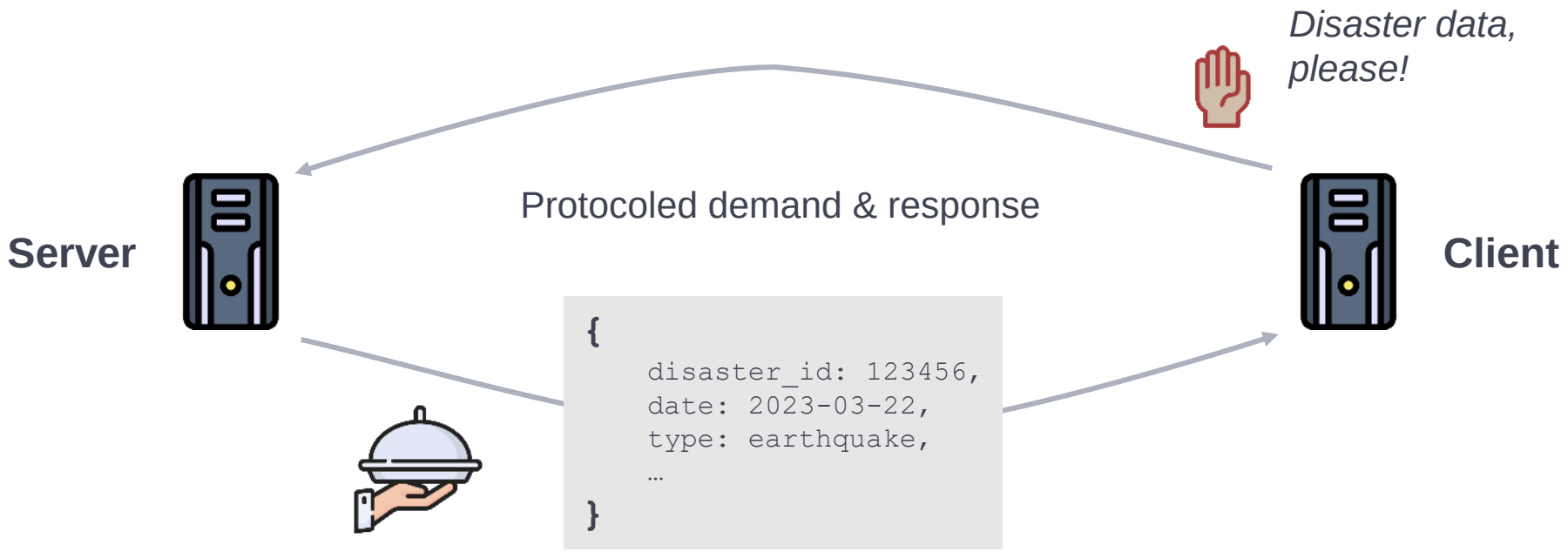
GDACS ID	**FL 1101659**
Glide number:	**FL-2022-000350-NGA**
Death:	**605**
Displaced:	**1306000**
Countries:	**Nigeria**
From - To	**10 Sep - 26 Oct**

GDACS Score

For more info on GDACS alert score click here.

Virtual OSOCC | Meteo assessment | Satellite products | Analytical products

2.3 Communication Systems: Application Programming Interfaces (APIs)



Data sharing and processing can be automated!

→ **From EM-DAT Data Source**: Many API already exists (Free or Private)

→ **To EM-DAT Users:** EM-DAT API is under development

Key points

- Various ways to deal with missingness
- Various kind of disaster loss biases
- Many technological and data opportunities
  - Geographical data
  - Online media
  - Communication systems

Yet, too much!

Yet, many more!