

A NEW GLOBAL DATASET ON GEO-HYDROLOGICAL HAZARDS AND THEIR IMPACTS AUTOMATICALLY EXTRACTED FROM ONLINE NEWS ARTICLES



Bram Valkenburg ^{1,2}, **Olivier Dewitte** ², **Benoît Smets** ^{1,2}

1. Department of Geography, Vrije Universiteit Brussel

2. Department of Earth Sciences, Royal Museum for Central Africa



INTRODUCTION

HAZARD DOCUMENTATION

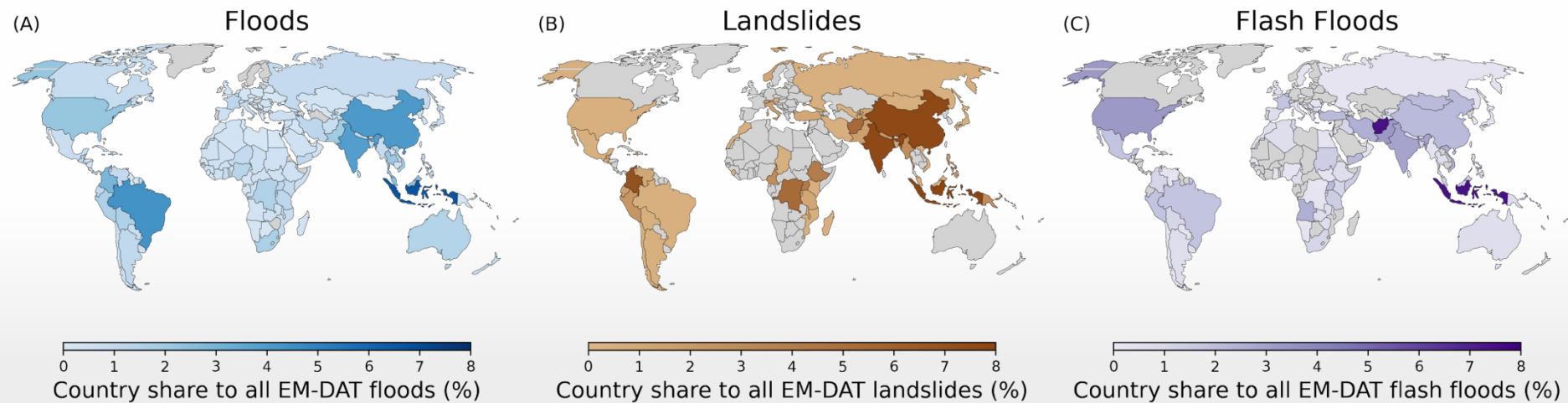
Geo-hydrological hazards: flood, landslide, flash flood

→ Incidence changing globally: anthropogenic influences

→ Datasets: deepen understanding of hazard and impact

→ Global datasets contain data gaps, undermine interpretation

→ EM-DAT: reference disaster dataset with entry criteria (Delforge et al., 2025)



News | Floods

Floods kill more than 100 across southern Africa as rains intensify

Torrential rains have devastated communities from Mozambique to South Africa, displacing hundreds of thousands



Onlookers inspect damage to the bridge crossing the Ga-Selati river following heavy rains over much of the Limpopo Province, South Africa.

NATION

Flash floods hit parts of Subang following heavy rain

By JUSTIN ZACK
Monday, 16 Feb 2026 | 9:40 PM MYT



Related Stories



CATASTROPHE | INDONESIA

Deadly Indonesian landslide leaves dozens missing

Timothy Jones Reuters, AFP
01/26/2026

At least 17 people have been confirmed dead after a landslide in Indonesia's West Java province last weekend. More than 20 naval officers are said to be among those who were trapped by the disaster.



Image: Aditya Irawan/NurPhoto/IMAGO

BBC

New Zealand calls off rescue efforts for six missing after landslide

24 January 2026

Share Save

Claire Keenan



Police say they are not expecting to find any more survivors from the landslide.

Rescue efforts for six missing people buried by a landslide in New Zealand on Thursday have been called off, with efforts now moving to recovering bodies, police have said.

"Tragically it is now apparent that we will not be able to bring them home alive," Anderson said, adding that formal identification is underway.

Two teenagers are among those who remain unaccounted for, with the youngest aged 15 years old.

Search teams located human remains beneath dirt and debris on Friday evening, Police Superintendent Tim Anderson said, after a landslide hit a popular campsite in Mount Maunganui.



nubia Focus BG

TALING JAYA: Flash floods

When contacted, Selangor Menteri Besar Fad Yusoff said the flood was caused by heavy rain around 8.30pm.

'We've lost everything, you can see caravans floating away': flash flooding hits Victoria's Great Ocean Road

Woman camping at flooded Wye River with her daughters says it's just heartbreaking, strangers are hugging strangers'

Get our breaking news email, free app or daily news podcast

Caitlin Cassidy and AAP

Thu 15 Jan 2026 07:19 CET

Share

Prefer the Guardian on Google



Cars swept out to sea as flash flooding hits Victoria's Great Ocean Road - video

Hundreds of people have evacuated a major camping ground after a severe thunderstorm caused flash flooding at Wye River, closed a major scenic tourism road and washed cars out to sea.

ing Barat region of mitigation agency said.

ered by heavy rains.

Floods kill more than 100 across southern Africa as rains intensify

Torrential rains have devastated communities from Mozambique to South Africa, displacing hundreds of thousands



Onlookers inspect damage to the bridge crossing the Ga-Selati following heavy rains over much of the Limpopo Province, South Africa.

Flash floods hit parts of Subang following heavy rain

By JUSTIN ZACK Monday, 16 Feb 2026 | 9:40 PM MYT



Related Stories



Deadly Indonesian landslide leaves dozens missing



CATASTROPHE | INDONESIA

Deadly Indonesian landslide leaves dozens missing

By Timothy Jones Reuters, AFP 01/26/2026

At least 17 people have been confirmed dead after a landslide in Indonesia's West Java province last weekend. More than 20 naval officers are said to be among those who were trapped by the disaster.



New Zealand calls off rescue efforts for six missing after landslide



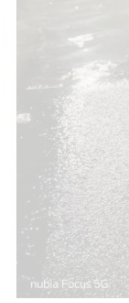
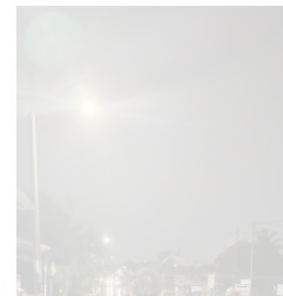
Police say they are not expecting to find any more survivors from the landslide.

Rescue efforts for six missing people buried by a landslide in New Zealand on Thursday have been called off, with efforts now moving to recovering bodies, police have said.

"Tragically it is now apparent that we will not be able to bring them home alive," Anderson said, adding that formal identification is underway.

Two teenagers are among those who remain unaccounted for, with the youngest aged 15 years old.

Search teams located human remains beneath dirt and debris on Friday evening, Police Superintendent Tim Anderson said, after a landslide hit a popular campsite in Mount Maunganui.



FLASH FLOODING: Flash flooding hits Victoria's Great Ocean Road.

When contacted, Selangor State Police Superintendent Mohd Mukhlis said the death toll rose to 10 around 8.30pm.

CAN TEXT MINING OF NEWS ARTICLES IMPROVE THE GLOBAL DOCUMENTATION OF GEO-HYDROLOGICAL HAZARDS?

'We've lost everything, you can see caravans floating away': flash flooding hits Victoria's Great Ocean Road

Woman camping at flooded Wye River with her daughters says 'it's just heartbreaking, strangers are hugging strangers'

Get our breaking news email, free app or daily news podcast

Caitlin Cassidy and AAP

Thu 15 Jan 2026 07:19 CET



Prefer the Guardian on Google



Cars swept out to sea as flash flooding hits Victoria's Great Ocean Road - video

Hundreds of people have evacuated a major camping ground after a severe thunderstorm caused flash flooding at Wye River, closed a major scenic tourism road and washed cars out to sea.

Floods kill more than 100 across southern Africa as rains intensify

Torrential rains have devastated communities from Mozambique to South Africa, displacing hundreds of thousands



Onlookers inspect damage to the bridge crossing the Ga-Selati following heavy rains over much of the Limpopo Province, South Africa.

Flash floods hit parts of Subang following heavy rain

By JUSTIN ZACK Monday, 16 Feb 2026 | 9:40 PM MYT



Related Stories



Deadly Indonesian landslide leaves dozens missing

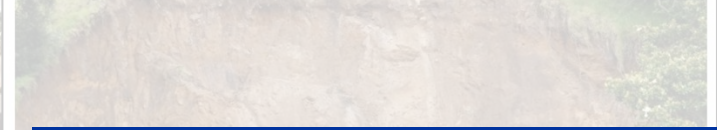
CATASTROPHE | INDONESIA Timothy Jones Reuters, AFP 01/26/2026

At least 17 people have been confirmed dead after a landslide in Indonesia's West Java province last weekend. More than 20 naval officers are said to be among those who were trapped by the disaster.



Image: Aditya Irawan/NurPhoto/IMAGO

New Zealand calls off rescue efforts for six missing after landslide



TEXT MINING: INFORMATION EXTRACTION FROM TEXT USING NATURAL LANGUAGE PROCESSING TECHNIQUES SUCH AS LANGUAGE MODELS

Rescue efforts for six missing people buried by a landslide in New Zealand on Thursday have been called off, with efforts now moving to recovering bodies, police have said.

"Tragically it is now apparent that we will not be able to bring them home alive," Anderson said, adding that formal identification is underway.

Two teenagers are among those who remain unaccounted for, with the youngest aged 15 years old.

Search teams located human remains beneath dirt and debris on Friday evening, Police Superintendent Tim Anderson said, after a landslide hit a popular campsite in Mount Maunganui.

'We've lost everything, you can see caravans floating away': flash flooding hits Victoria's Great Ocean Road



Thu 15 Jan 2026 07:19 CET

Share

Prefer the Guardian on Google

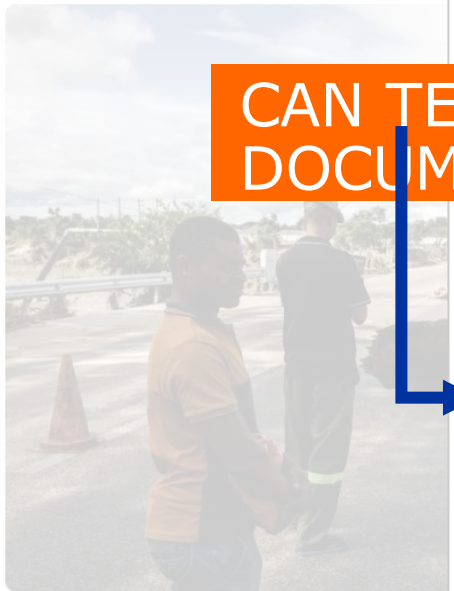


Cars swept out to sea as flash flooding hits Victoria's Great Ocean Road - video

Hundreds of people have evacuated a major camping ground after a severe thunderstorm caused flash flooding at Wye River, closed a major scenic tourism road and washed cars out to sea.

Floods kill more than 100 across southern Africa as rains intensify

Torrential rains have devastated communities from Mozambique to South Africa, displacing hundreds of thousands



Onlookers inspect damage to the bridge crossing the Ga-Selati following heavy rains over much of the Limpopo Province, South Africa.

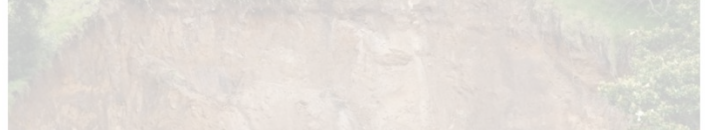
By JUSTIN ZACK Monday, 16 Feb 2026 | 9:40 PM MYT



Related Stories

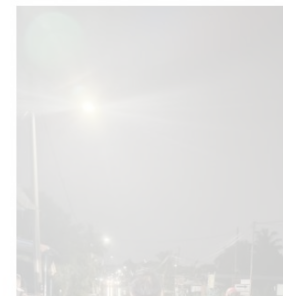


New Zealand calls off rescue efforts for six missing after landslide

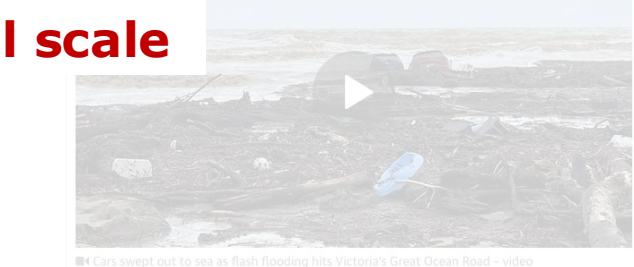


Police say they are not expecting to find any bodies. Rescue efforts for six missing people after a landslide in New Zealand have been called off, with efforts now moving to recovering bodies, police have said.

"Tragically it is now apparent that we will not be able to bring them home alive," Anderson said, adding that formal identification is underway. Two teenagers are among those who remain unaccounted for, with the youngest aged 15 years old. Search teams located human remains beneath dirt and debris on Friday evening. Police Superintendent Tim Anderson said, after a landslide hit a popular campsite in Mount Maunganui.



'We've lost everything, you can see caravans floating away': flash flooding hits Victoria's Great Ocean Road



Hundreds of people have evacuated a major camping ground after a severe thunderstorm caused flash flooding at Wye River, closed a major scenic tourism road and washed cars out to sea.

CAN TEXT MINING OF NEWS ARTICLES IMPROVE THE GLOBAL DOCUMENTATION OF GEO-HYDROLOGICAL HAZARDS?

TEXT MINING: INFORMATION EXTRACTION FROM TEXT USING NATURAL LANGUAGE PROCESSING TECHNIQUES SUCH AS LANGUAGE MODELS

Rarely applied at the global scale

METHODOLOGY

TEXT MINING FOR NATURAL HAZARD INFORMATION



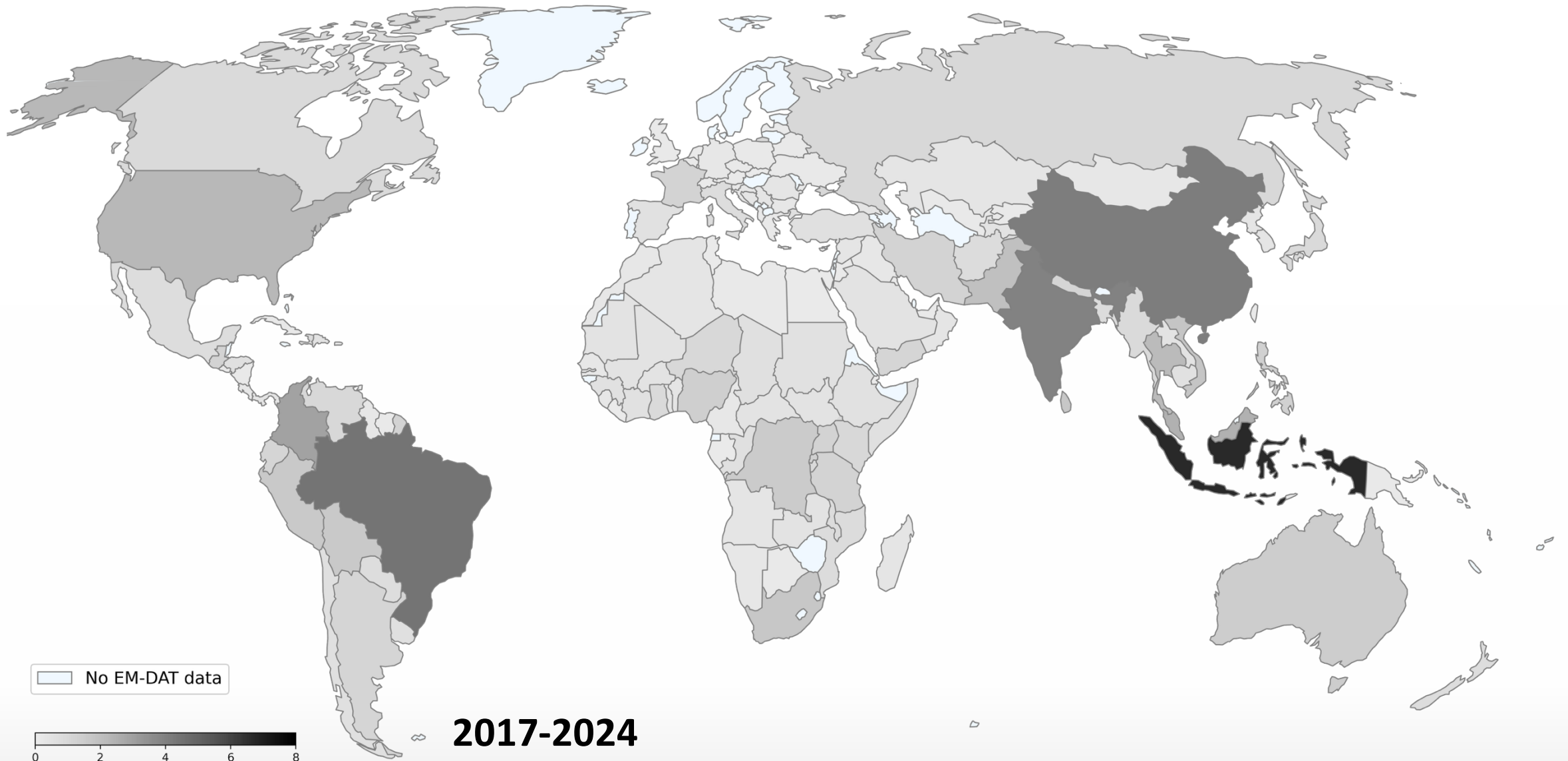
Processing paragraphs of news articles using language models

- Translation: 58 languages
- Filtering and classifying of natural hazard related text
- Information extraction: location, timing and impact
- Fine-tuned encoder models from HuggingFace (open-source): DistilBERT, DeBERTa, RoBERTa, BART

Events: clustered paragraphs that occur on the same location over the same period

Text source: GDELT dataset, keyword-based extraction of online news articles

- 6 366 905 articles (2017-2024), comprising 98 001 284 paragraphs



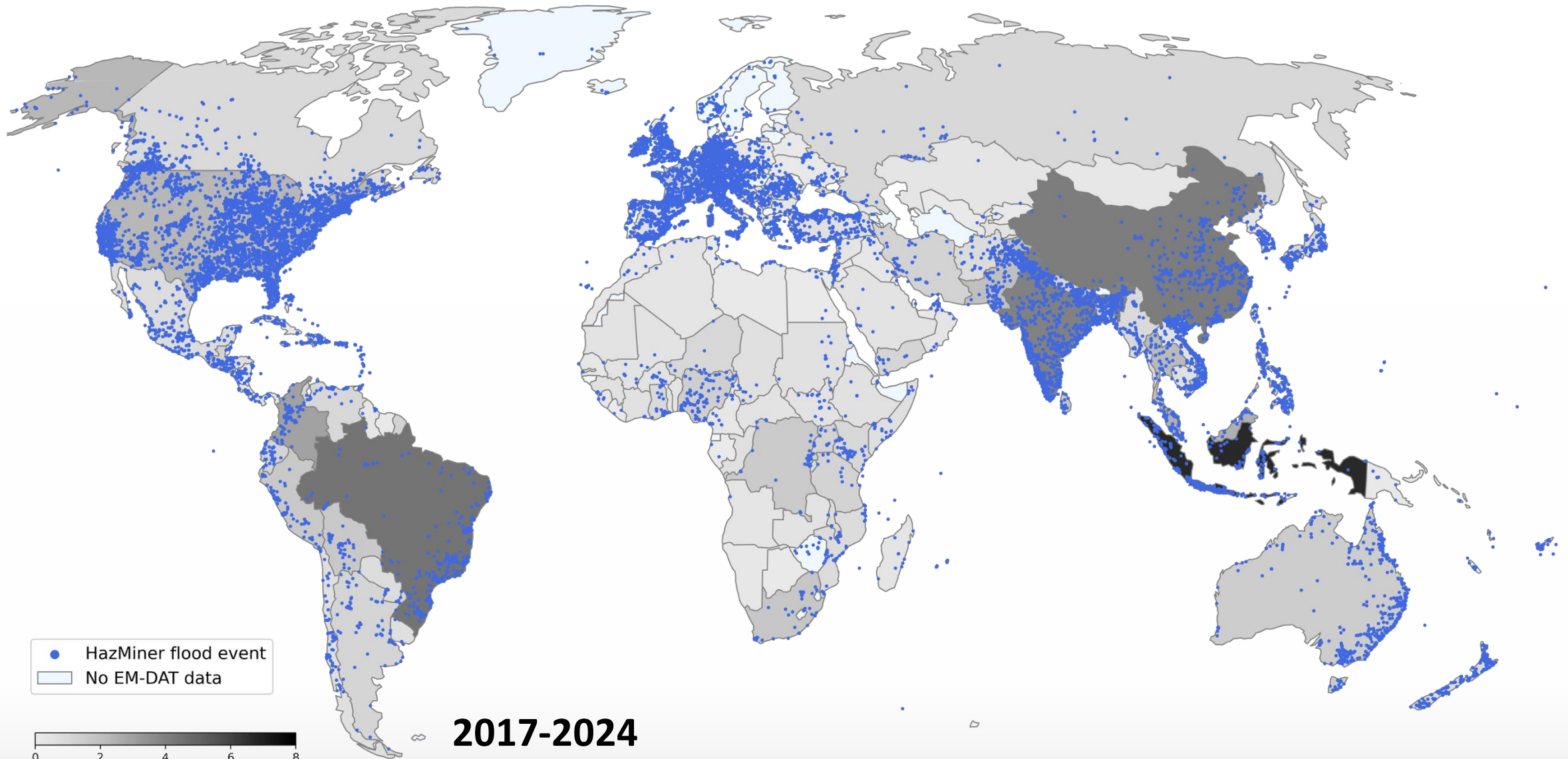
No EM-DAT data



Country share to all EM-DAT floods (%)

2017-2024

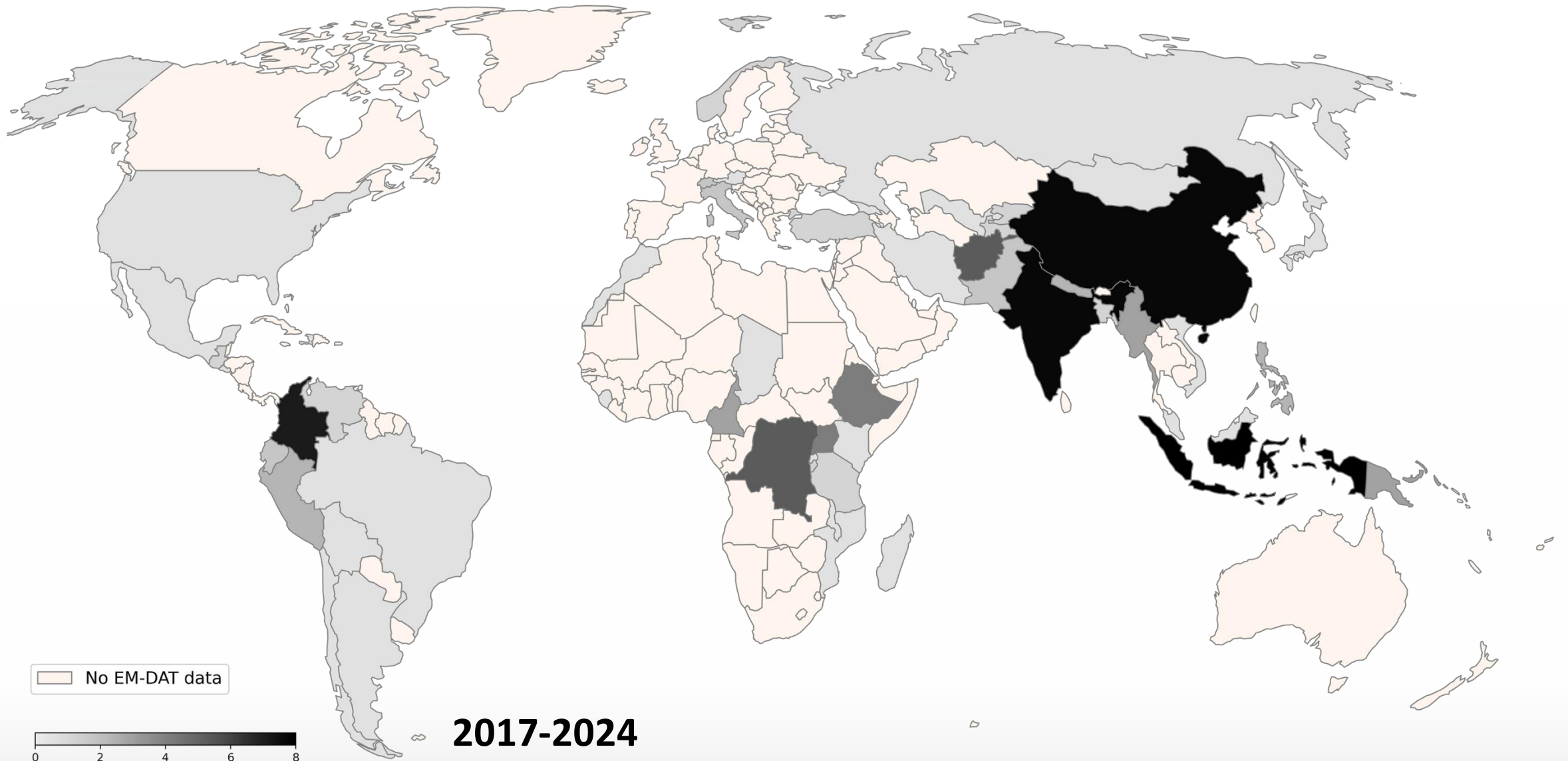
EM-DAT: 1 089 flood disasters



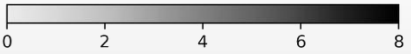
2017-2024

EM-DAT: 1 089 flood disasters

HazMiner: 21 411 flood hazard events

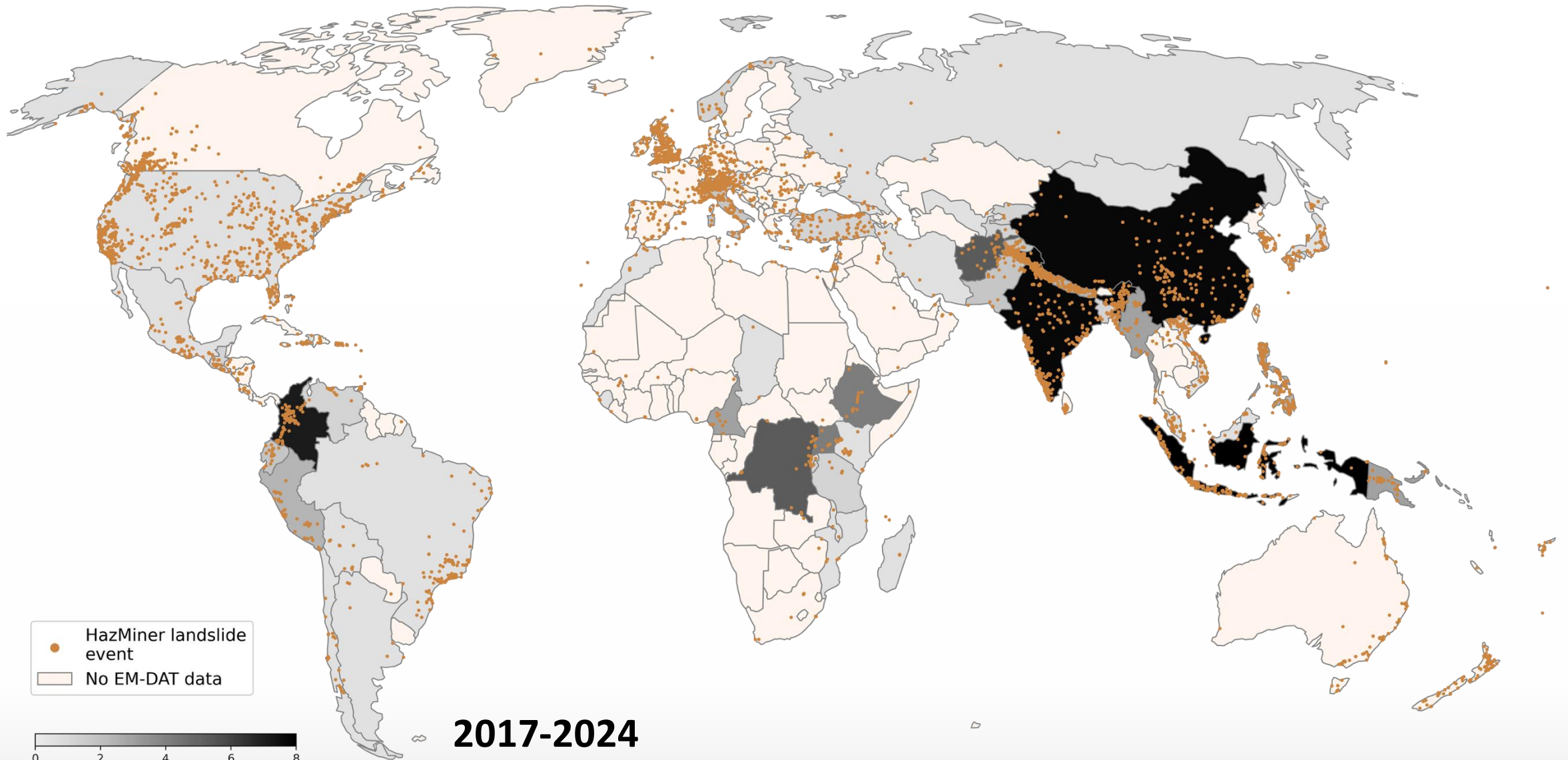


No EM-DAT data



Country share to all EM-DAT landslides (%)

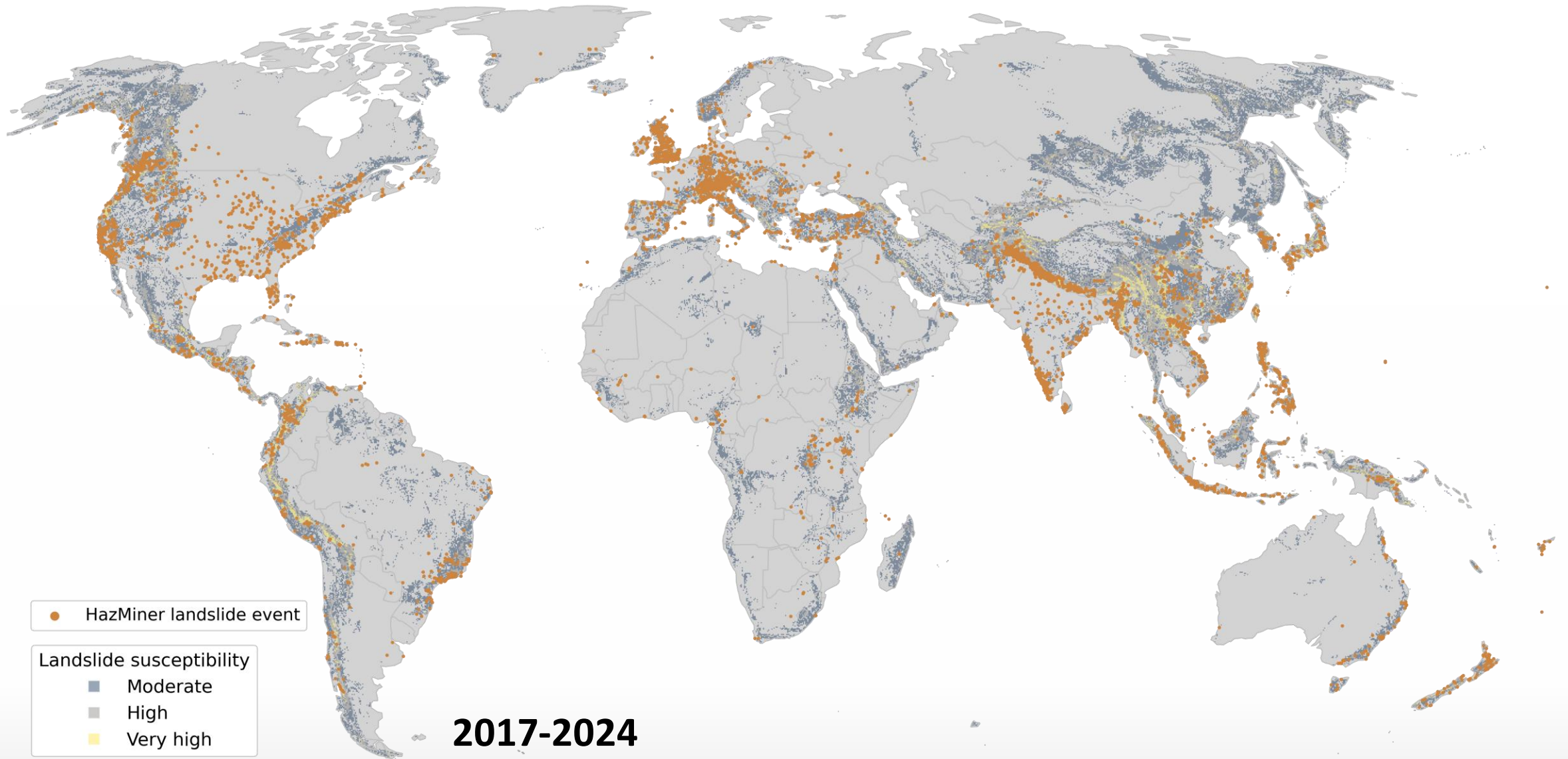
2017-2024
 EM-DAT: 169 landslide disasters
 HazMiner: ?



2017-2024

EM-DAT: 169 landslide disasters

HazMiner: 7 659 landslide hazard events



● HazMiner landslide event

Landslide susceptibility

■ Moderate

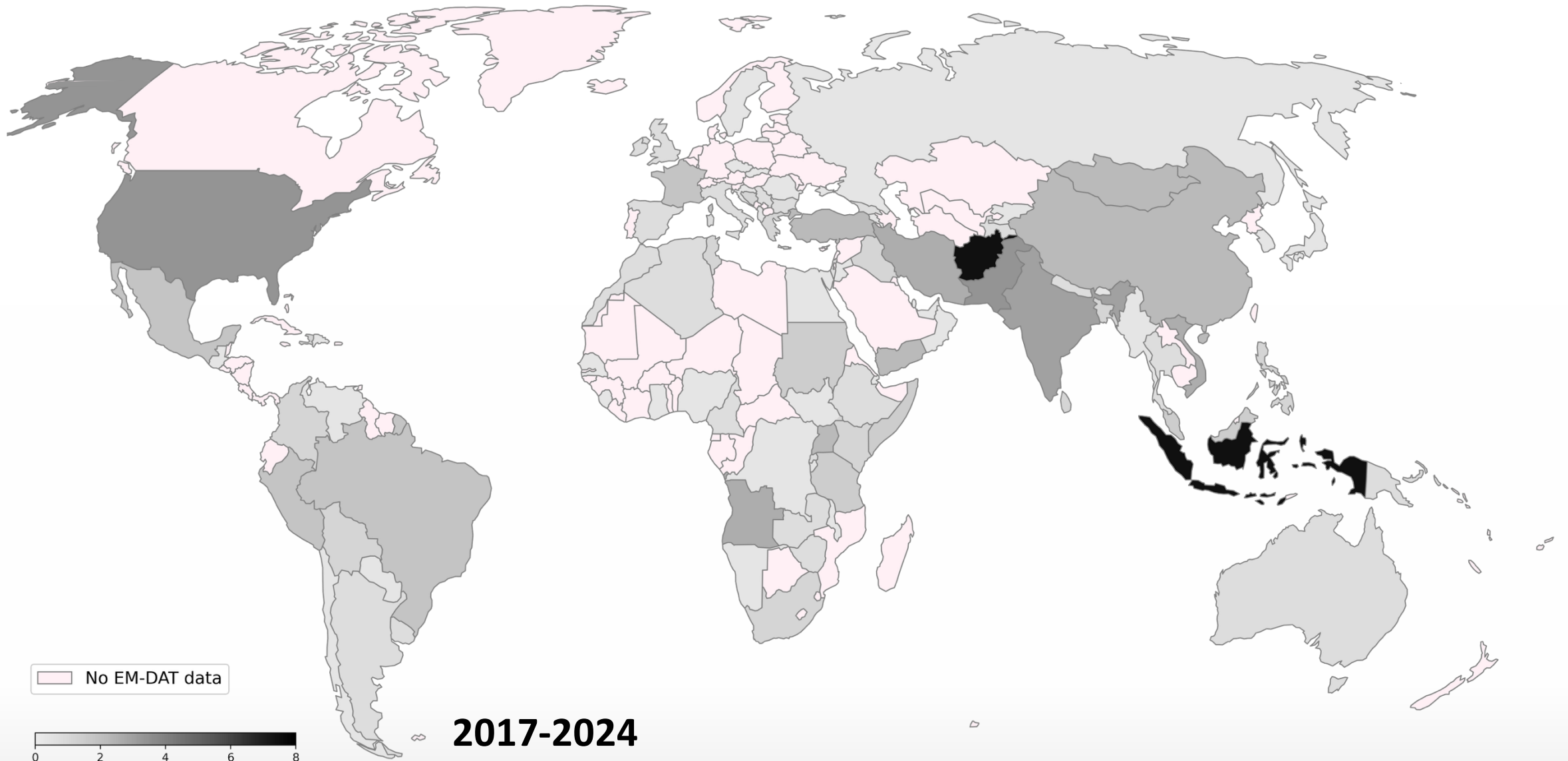
■ High

■ Very high

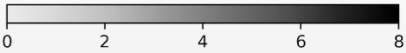
2017-2024

EM-DAT: 169 landslide disasters

HazMiner: 7 659 landslide hazard events



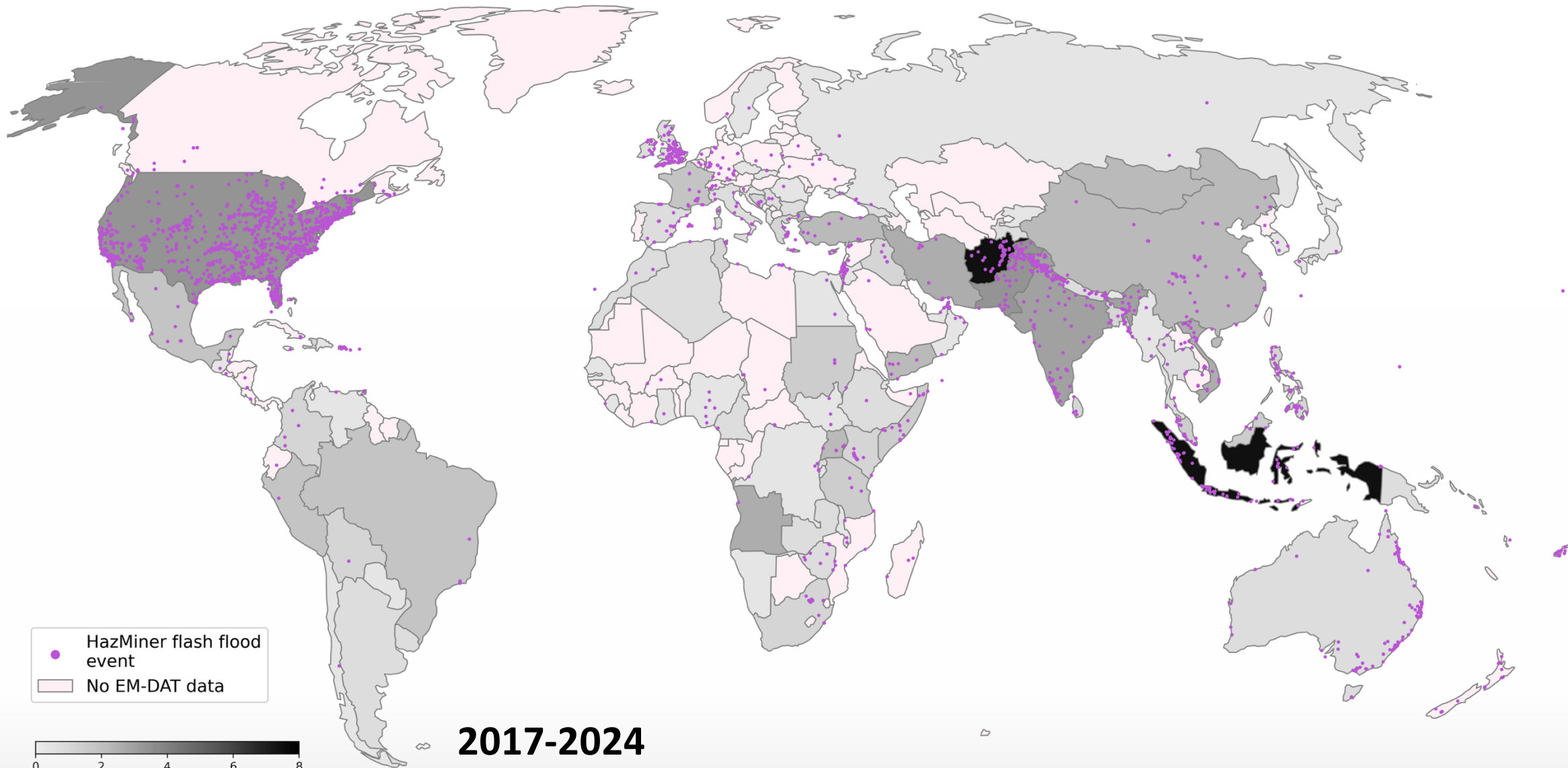
No EM-DAT data



Country share to all EM-DAT flash floods (%)

2017-2024

EM-DAT: 268 flash flood disasters



2017-2024

EM-DAT: 268 flash flood disasters

HazMiner: 3 606 flash flood hazard events

BIASES

EXAMPLE

Known biases in HazMiner

- Language bias
- Hazard bias
- Translation bias
- Source bias
- Systemic bias

BIASES

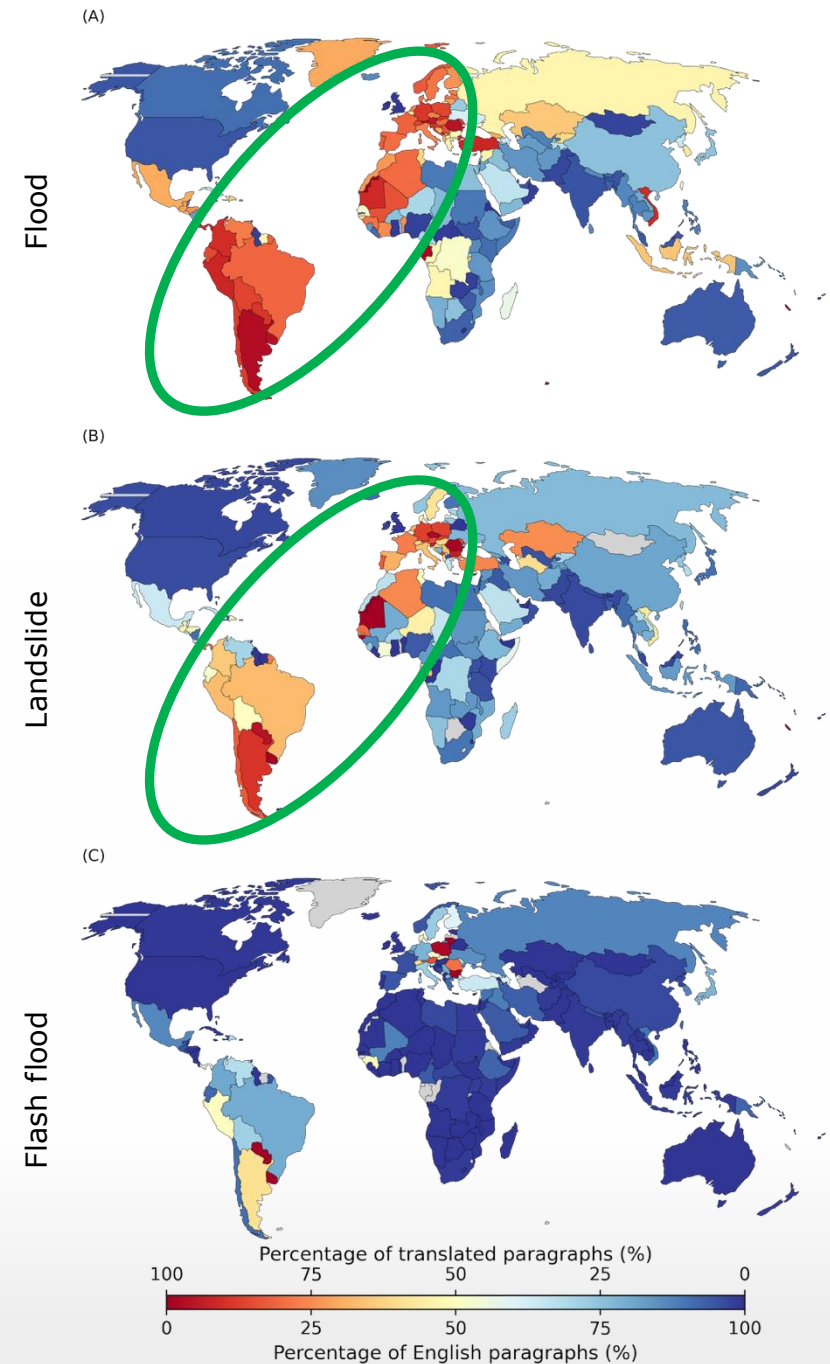
EXAMPLE

Known biases in HazMiner

- **Language bias**
 - Reduced: 58 languages
- Hazard bias
- Translation bias
- Source bias
- Systemic bias

English vs multilingual
30% increase in events due to multilingual approach

Example: *88% of events would be missed in Germany when using only English text*



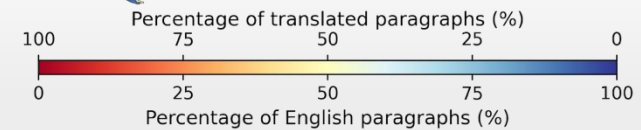
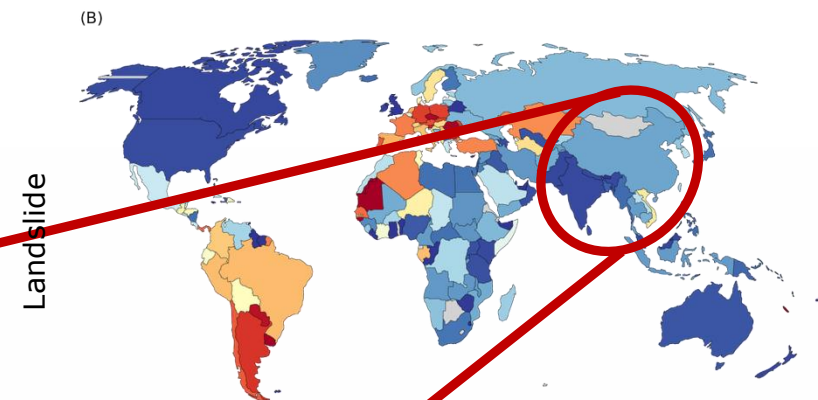
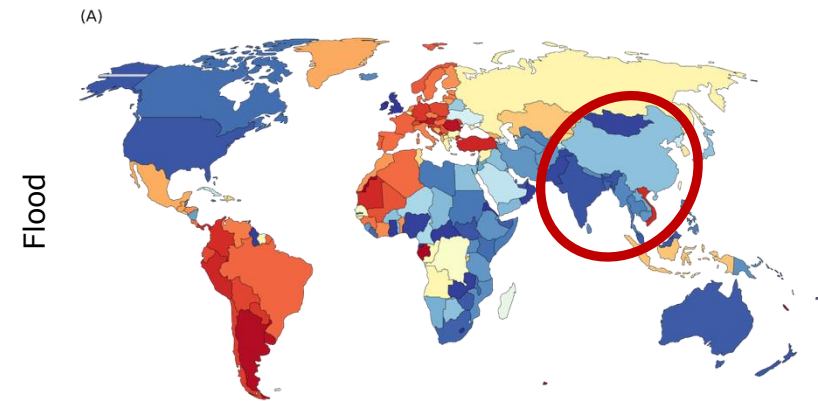
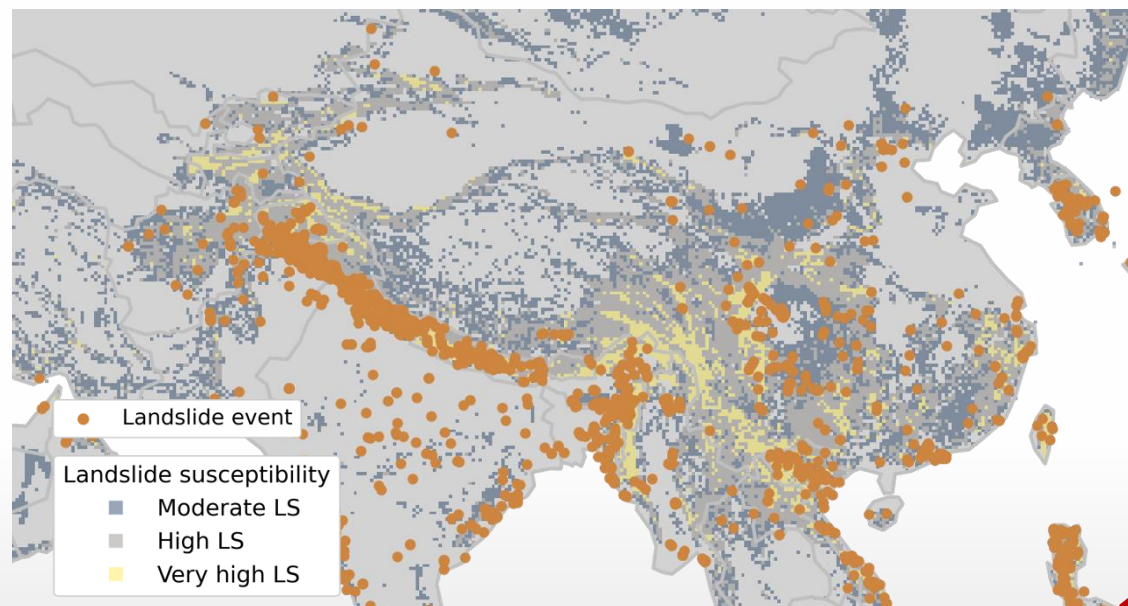
BIASES

EXAMPLE

Known biases in HazMiner

- Language bias

- Still present



HAZMINER PARAGRAPHS: CENTRAL/WESTERN EUROPE 2021 FLOOD

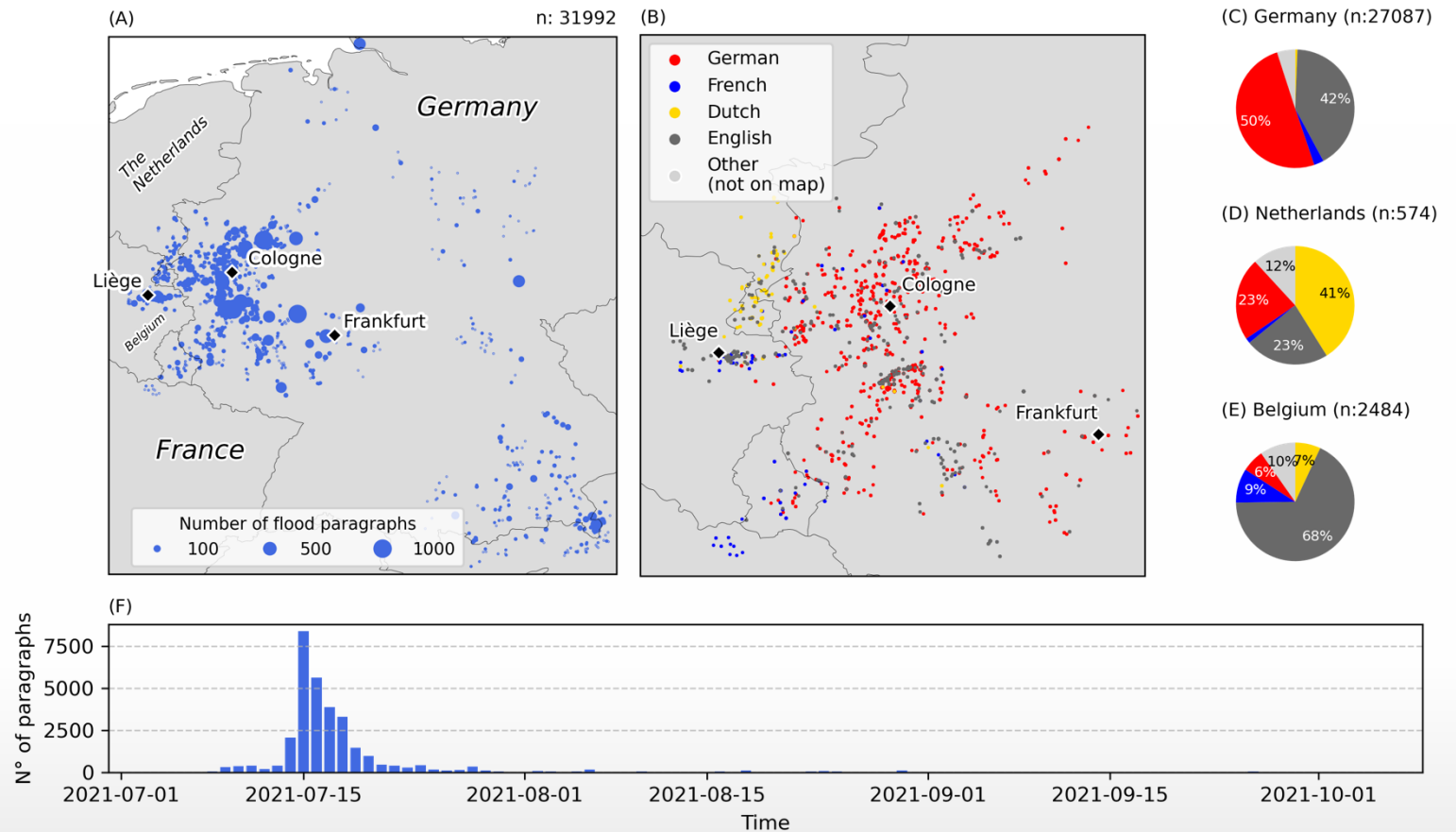
DETAILED REGIONAL DOCUMENTATION

EM-DAT: 7 events

- one per country
- 54 locations
- From 12/07/2021 until 18/07/2021

- Austria: 6 locations
- Belgium: 6 locations
- Switzerland: 5 locations
- Germany: 16 locations
- France: 11 locations
- Luxemburg: 8 locations
- The Netherlands: 2 locations

HazMiner: 1 event, 31 922 paragraphs



CONCLUSION

Text mining **improves global documentation** of geo-hydrological hazards

- More detailed and complete interpretation of natural hazards
- Biases and limitations are still present
- HazMiner = complementary and versatile observation system

Preprint

Preprints / Preprint egusphere-2026-722

Unravelling information on impactful geo-hydrological hazard events with HazMiner, a multilingual text mining method developed through a global scale coverage application

[Bram Valkenburg](#) [✉](#), [Olivier Dewitte](#), and [Benoît Smets](#)

<https://doi.org/10.5194/egusphere-2026-722>



EXTRA SLIDES

Bibliography

Delforge, D., Wathelet, V., Below, R., Sofia, C. L., Tonnelier, M., Van Loenhout, J. A. F., and Speybroeck, N.: EM-DAT: the Emergency Events Database, *International Journal of Disaster Risk Reduction*, 124, 105509, <https://doi.org/10.1016/j.ijdr.2025.105509>, 2025.

Stanley, T. and Kirschbaum, D. B.: A heuristic approach to global landslide susceptibility mapping, *Nat Hazards*, 87, 145–164, <https://doi.org/10.1007/s11069-017-2757-y>, 2017.



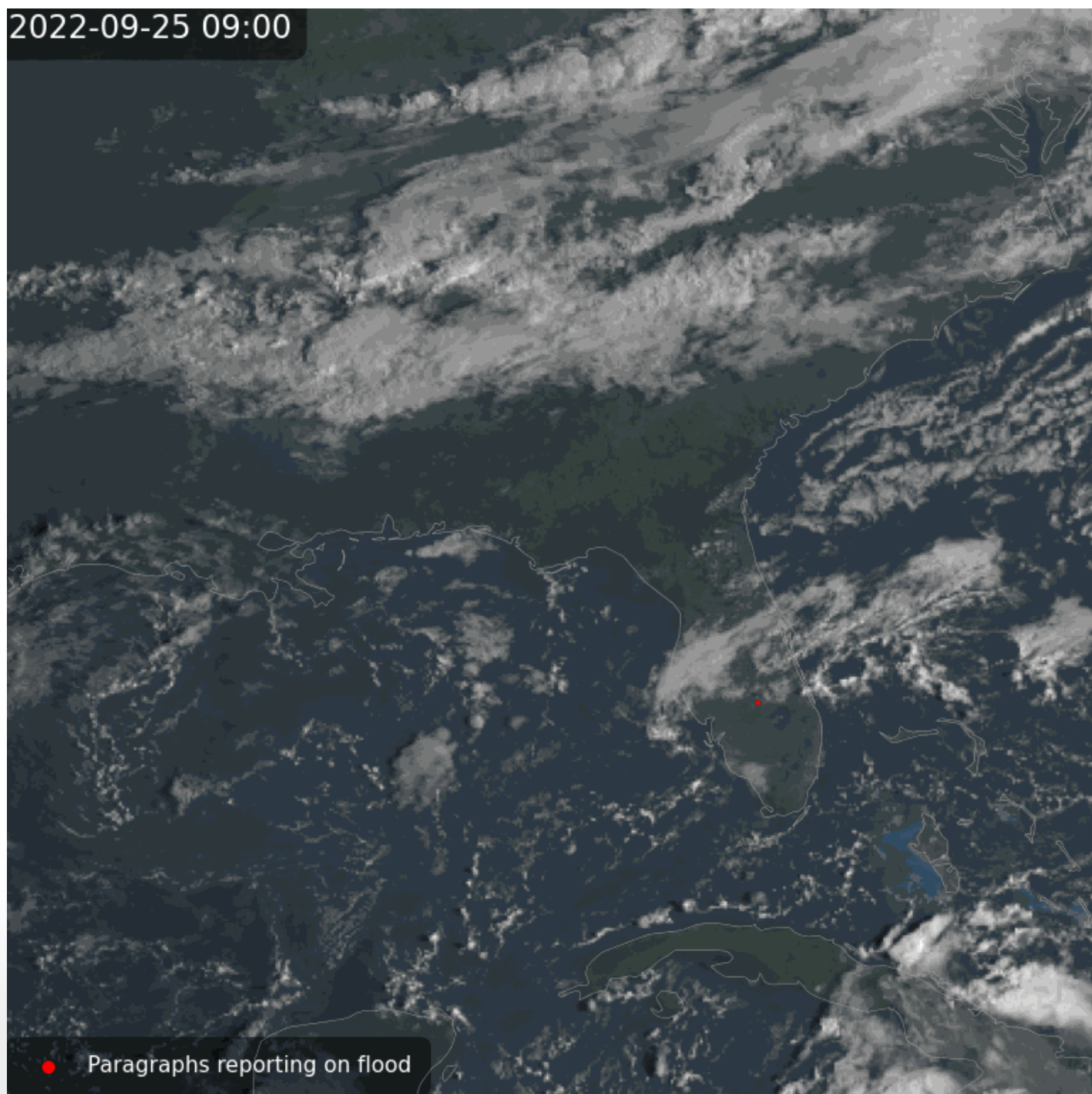
Hurricane Ian

- September 2022
- Caribbean and USA



GOES-16 from
Google Earth Engine

2022-09-25 09:00



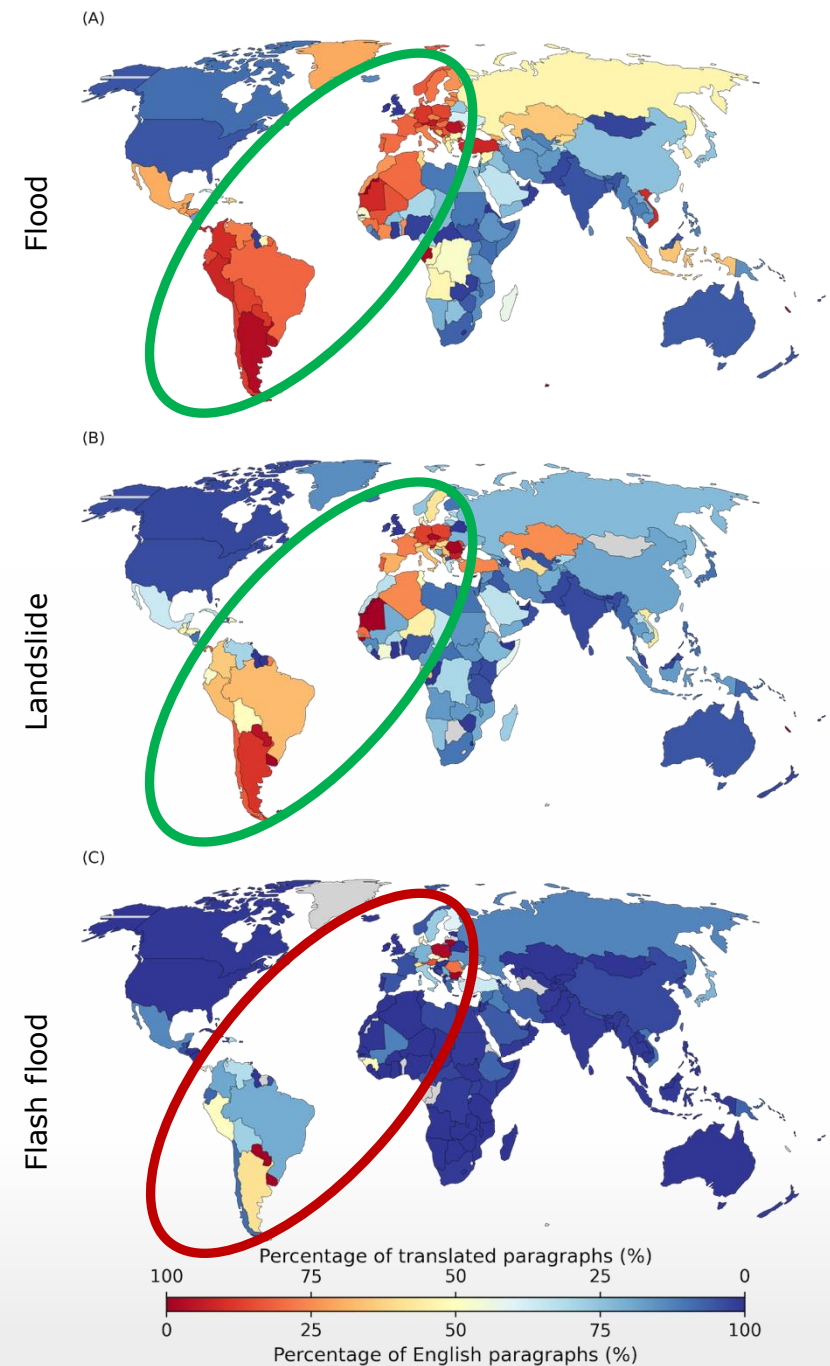
● Paragraphs reporting on flood

BIASES

OTHER

Known biases in HazMiner

- Language bias
- **Hazard bias**
 - Different media coverage
- **Translation bias**
 - Different terminology
- Source bias
- Systemic bias



BIASES

OTHER

Known biases in HazMiner

- Language bias
- Hazard bias
- Translation bias
- **Source bias**
 - Strong relationship with the GDELT dataset: USA and Europe
- **Systemic bias**
 - Biases resulting from preferential media coverage