Practical Challenges in Using Statistical Methods for Predicting Extreme Events

ESF-COST High-Level Research Conference
Extreme Environmental Events

Selwyn College, Cambridge, United Kingdom
13 - 17 December 2010

Farrokh Nadim, ScD
Director, International Centre for Geohazards / Norwegian Geotechnical Institute

NGI

ICG
International Centre for Geohazards

Challenge:
What is the definition of an "Extreme Event"?

"... An event that is rare at a particular place and time of year. Definitions of "rare" vary, but an extreme weather event would normally be as rare as or rarer than the 10^{th} or 90^{th} percentile of the observed probability density function*…."

* IPCC definitions in the Working Group I, Working Group II, and Synthesis reports of the Fourth Assessment Report .

What is an “Extreme Event”?

“...a physical occurrence that with respect to some class of related occurrences, is either notable, rare, unique, profound, or otherwise significant in terms of its impacts, effects, or outcomes*.”

* U.S. National Science Foundation “Workshop on Extreme Events: Developing a Research Agenda for the 21st Century”

What is an “Extreme Event”?

- Extreme events do not bear a one-to-one relationship with extreme impacts. ….. Depending on the context, physical extremes may or may not bring along extreme impacts and disasters.

- …… In the climate change and adaptation literature, extreme events are often considered in strictly physical terms.

What is an “Extreme Event”?

In the disaster risk community, “extreme” refers to levels of damage and loss, and the notion of “event” increasingly takes on a social connotation. Metrics to quantify extreme impacts may include, among others:

1. human casualties and injuries,
2. numbers of permanently or temporarily displaced people,
3. impacts to properties, measured in terms of numbers of buildings damaged or destroyed,
4. impacts to infrastructure and lifelines,
5. financial or economic loss,
6. duration of the above impacts.

Viewpoint of engineers and natural hazard researchers

- The main interest of engineers and natural hazard researchers is to understand the geophysical processes that lead to ‘extreme’ natural events, which focuses attention on the uncertainties and the unpredictable nature…

- The link between an extreme physical event with an extreme impact depends strongly on context, reflecting both the degree to which populations, ecosystems and other elements at risk are located in the path of the extreme (exposure) and the underlying vulnerability or susceptibility to damage of these elements.

Link between event & impact (risk)
Viewpoint of engineers & natural scientists

Impact or Risk = f (H, E, V, U)

H **=** Hazard (temporal probability of an event)

E **=** Exposure of element(s) at risk

V **=** Vulnerability of element(s) at risk

U **=** Cost of loss of element(s) at risk

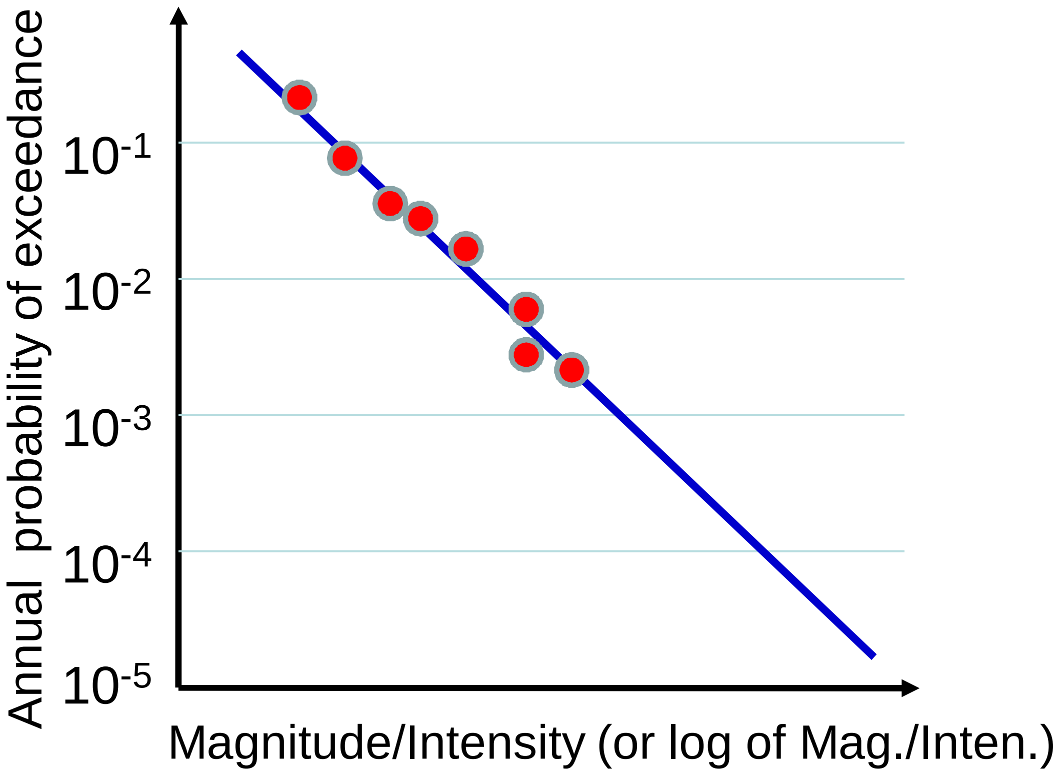
Definition of “Extreme Event” in remainder of this presentation

- An extreme event refers to a natural phenomenon (described in terms of its geometry, mechanical and other characteristics) that could lead to damage and occurs less frequently than once in 50 – 100* years.

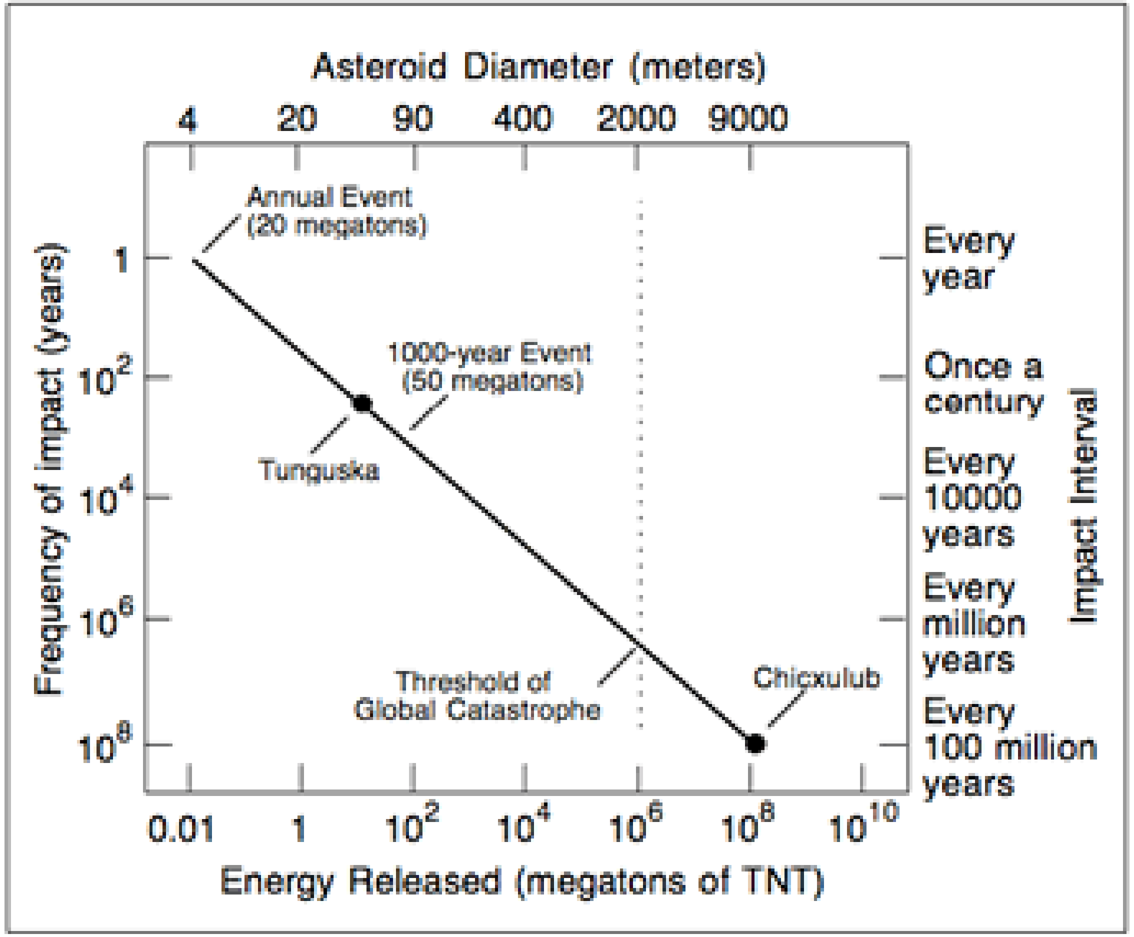
* Note that for certain class of problems, characterisation of events with annual occurrence probability as low as 10^{-6} – 10^{-7} might be of interest!

General observation about all natural hazard events

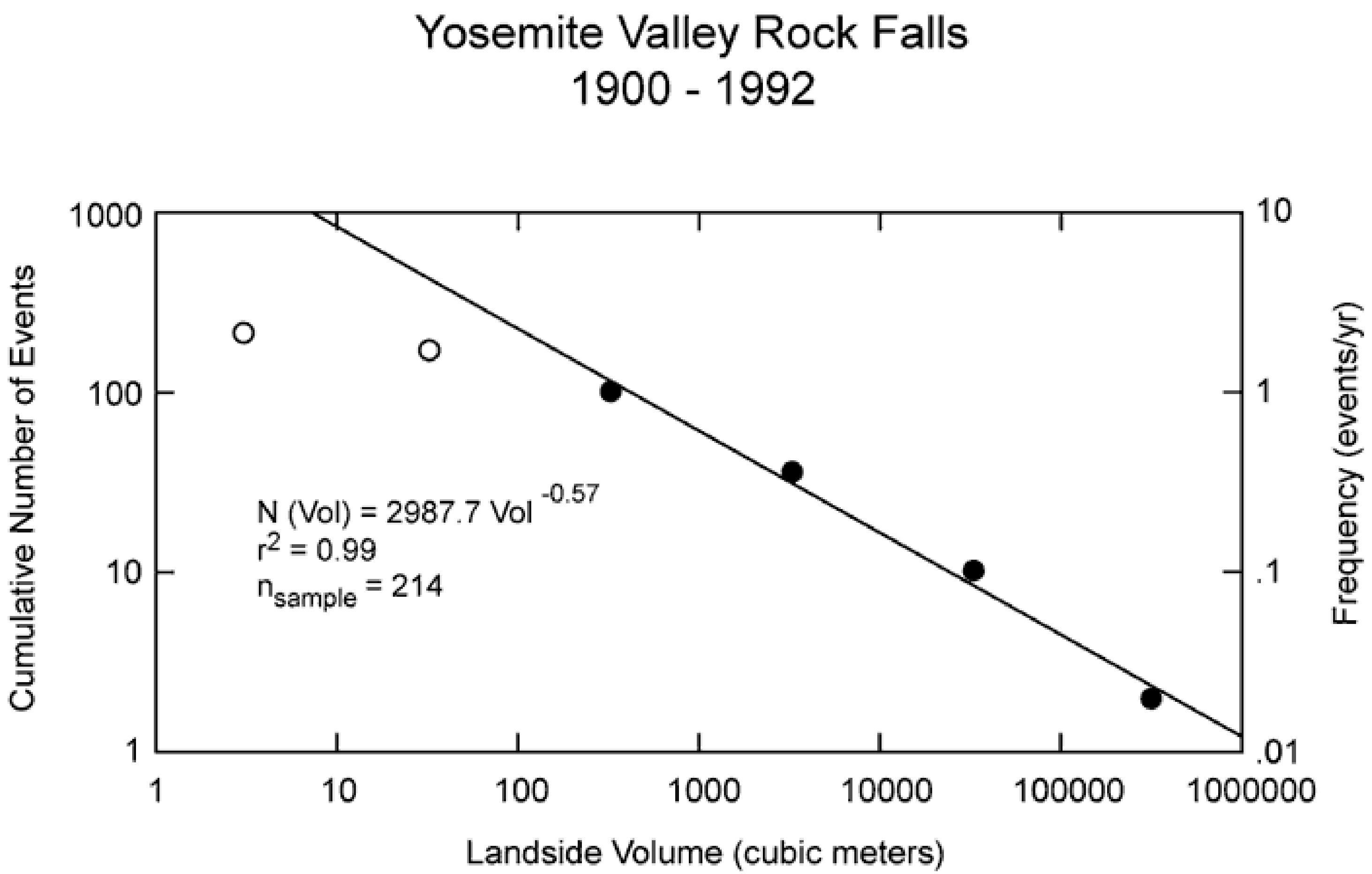
Larger events have a lower probability of occurrence



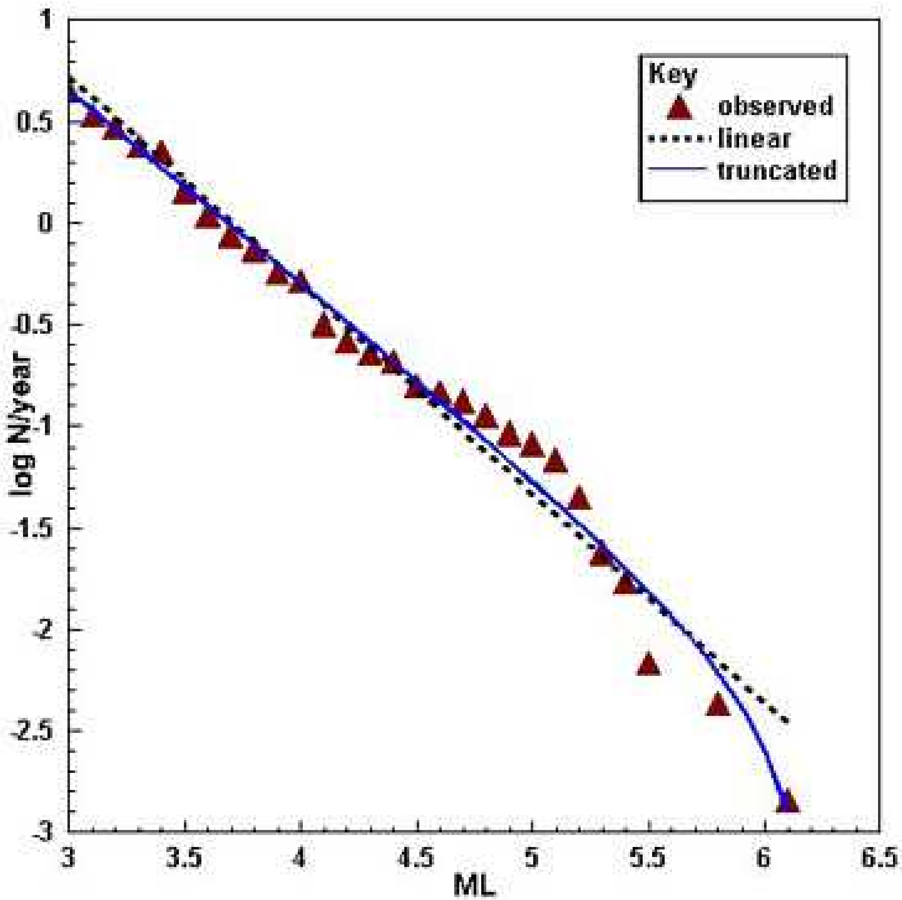
Example: Diameter of astroids impacting the Earth



Example: Volume of a rock fall

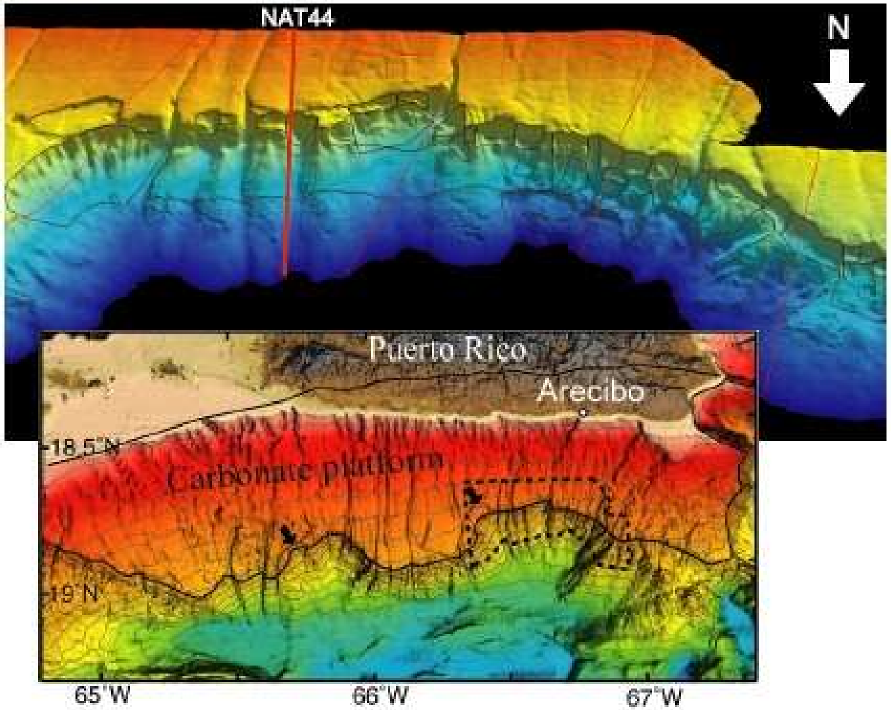


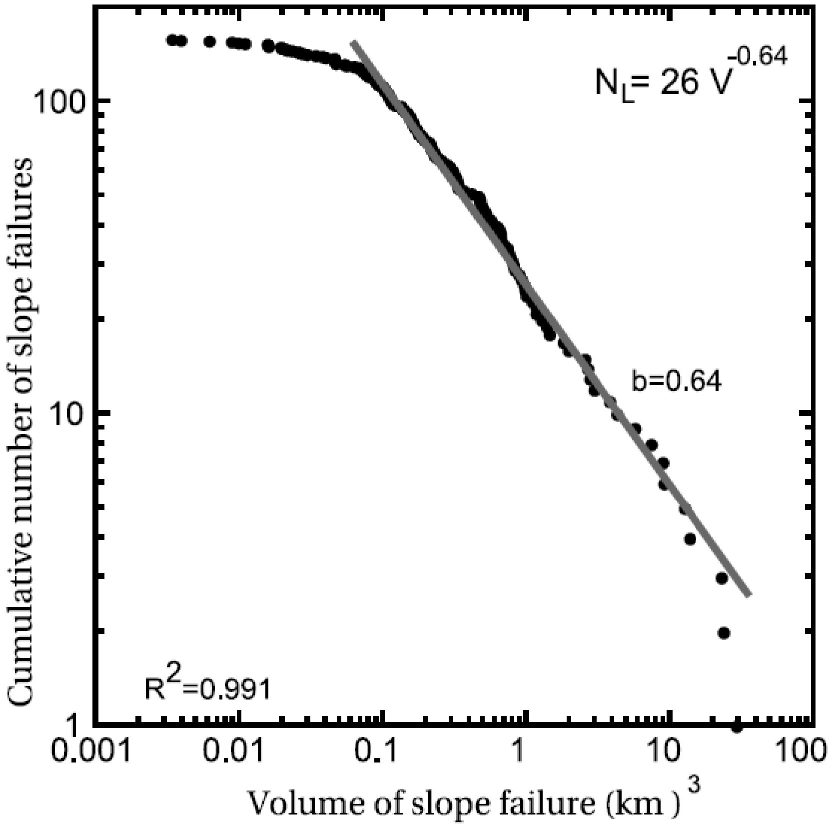
Example: Magnitude of earthquakes in Great Britain (after Roger Musson)



- an earthquake of magnitude 3.7 or larger every 1 year
- an earthquake of magnitude 4.7 or larger every 10 years
- an earthquake of magnitude 5.6 or larger every 100 years

Example: Size distribution of submarine slides offshore Puerto Rico





Statistical approaches for estimating extreme events

Standard Approach:

- Gather data about the magnitude and/or intensity of the phenomenon of interest.
- Fit a distribution function to the histogram of available data.
- “Predict” the magnitude or intensity of the event at the specified annual probability.

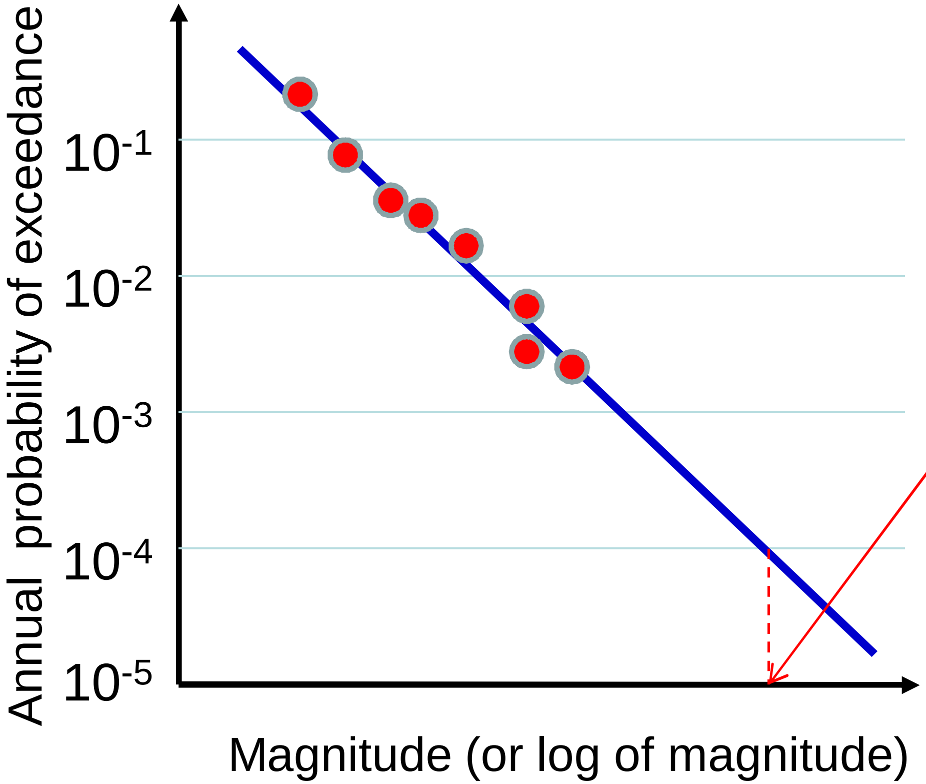
Advanced Approach:

- Same as above, but justify the chosen probability distribution on the basis of physics of the problem and/or the type and quality of data available.

Challenge:

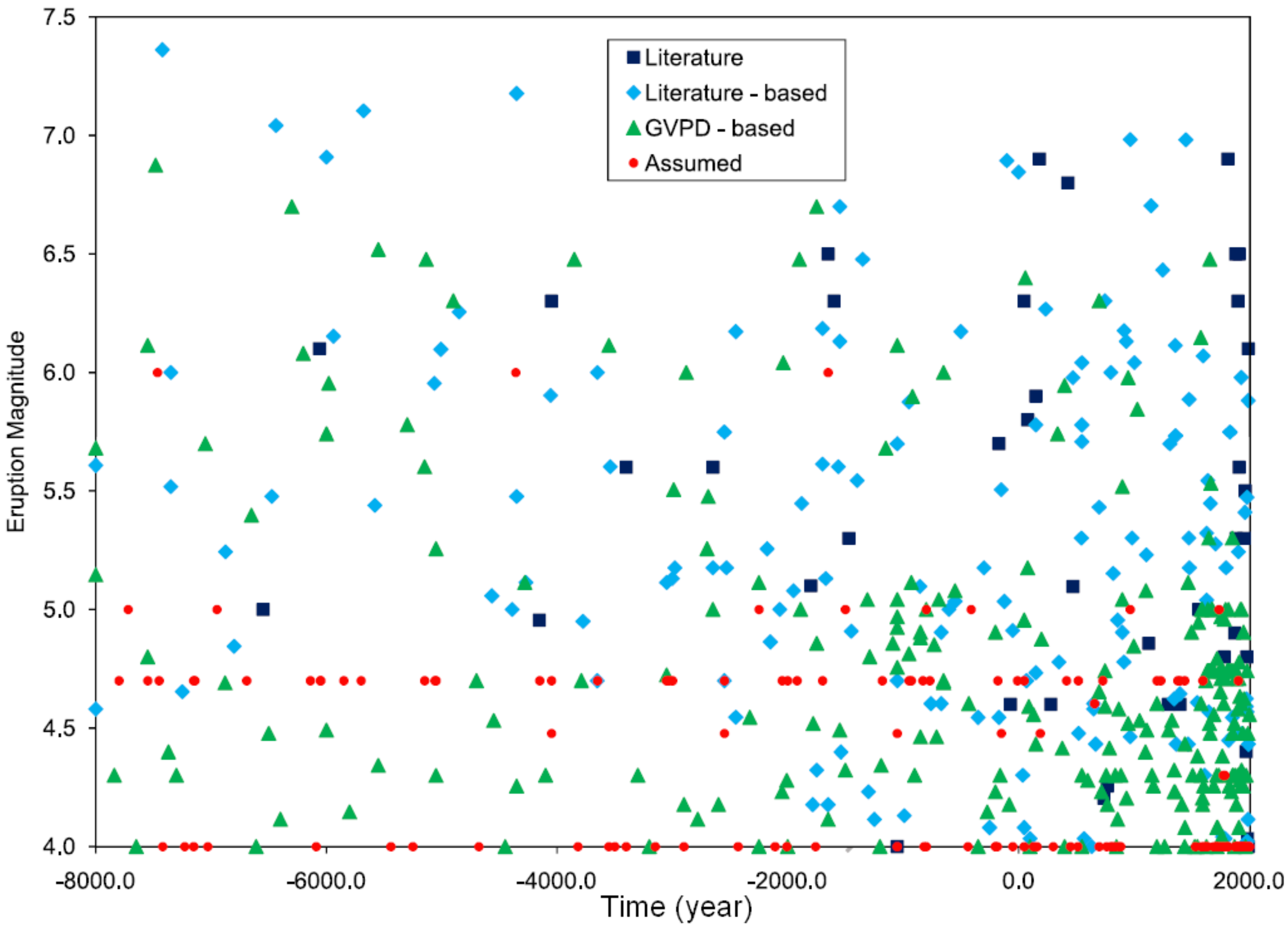
High quality data are only available for the recent past

Prediction of characteristics of events with very low probability of occurrence involves great uncertainty



Predicted magnitude of an extreme event with annual exceedance probability of 10^{-4}

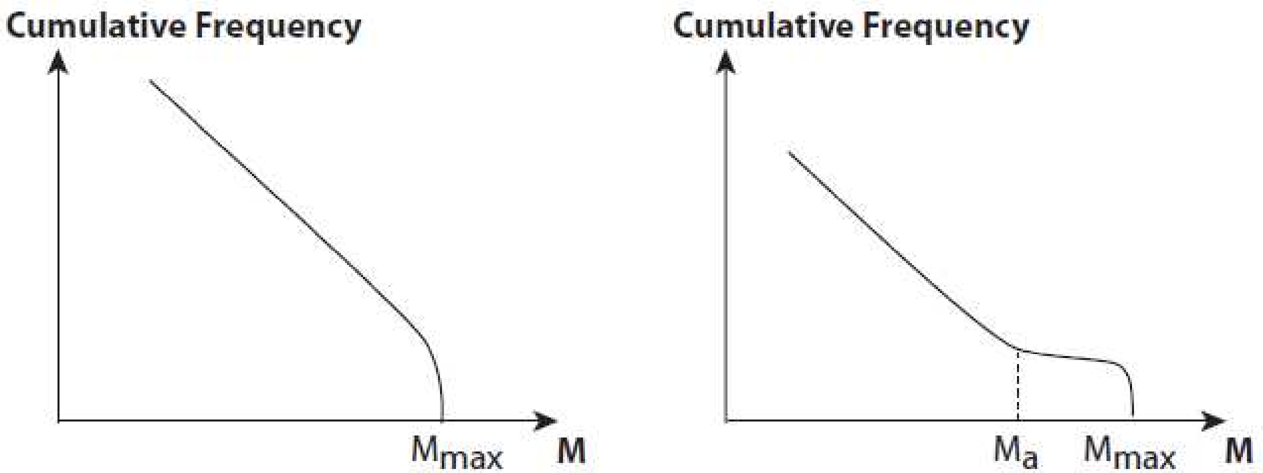
Example: Magnitude of volcanic eruptions (Deligne et al., 2010)



Challenge:

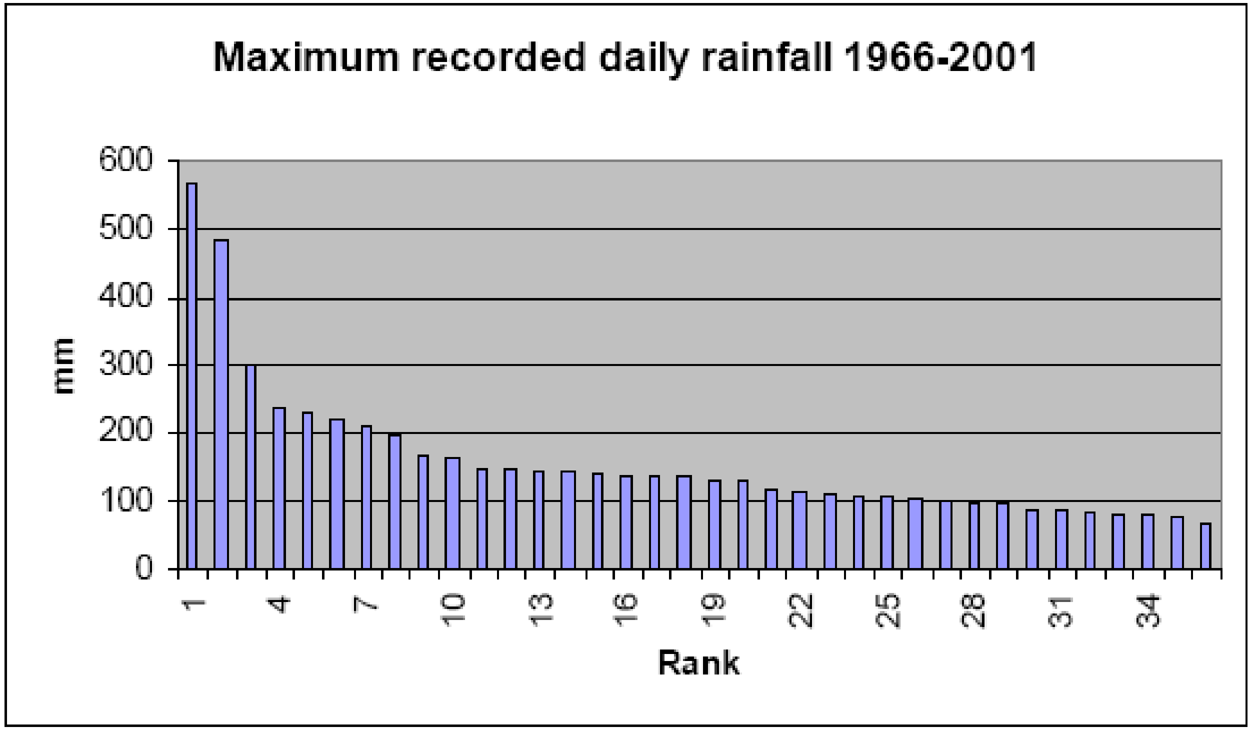
Mechanisms triggering the extreme events maybe different from those triggering the more frequent events

Extrapolation is risky business!



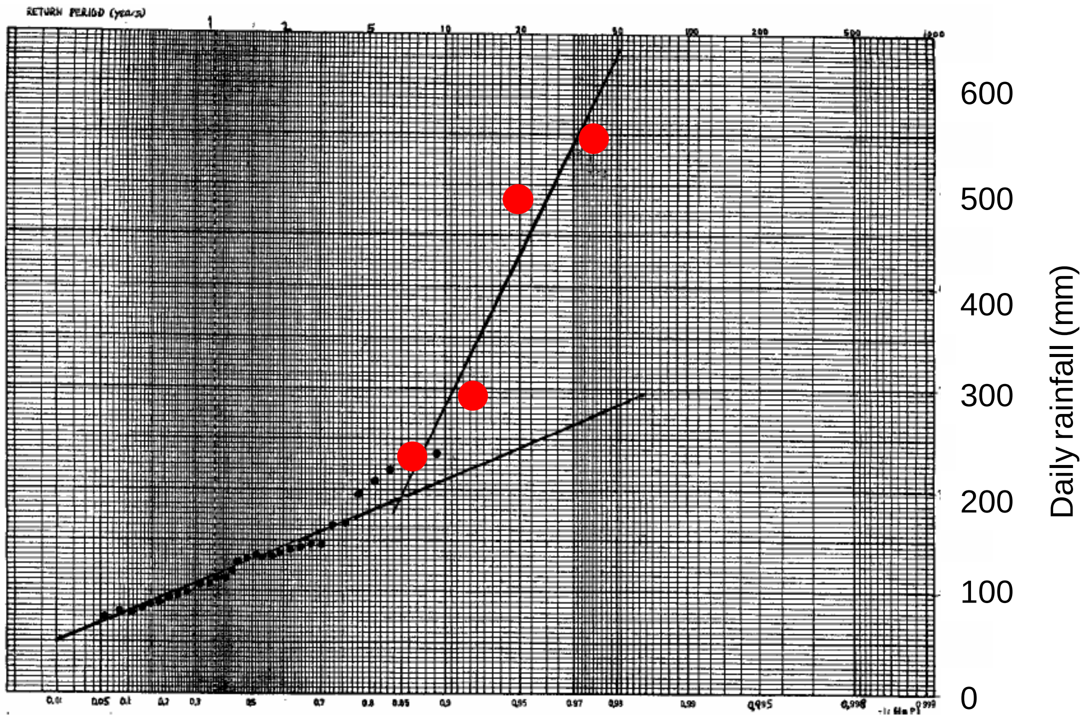
Example: Frequency-magnitude relationships for earthquakes in two types of tectonic setting

Example: Annual maximum daily rainfall on a volcano in Nicaragua



Ranking of recorded annual maximum daily rainfall for the period 1966-2001 on Chinandega Volcano

Example: Annual maximum daily rainfall on a volcano in Nicaragua



Plot of annual maximum daily rainfall on a Gumbel distribution diagram for estimating the maximum daily rainfall with 50 years return period.

Example: Annual maximum daily rainfall on a volcano in Nicaragua

- By plotting the data on the Gumbel diagram, two different trends can be distinguished.
- The small black data points represent convective rainfalls of short duration.
- The four red data points represent recordings during tropical cyclones.
- The two highest values (570 and 485 mm/day) were recorded during Hurricane Alleta in May 1982 and Hurricane Mitch in October 1998.
- The daily rainfall with return period of 50 years seems to be about 280 mm for convective rainfalls and 650 mm for tropical cyclones.

Challenges related to predicting extreme earthquake & tsunami events

with contributions from ICG colleagues

Finn Løvholt (NGI), Hilmar Bungum (NORSAR) & Carl Harbitz (NGI)

NGI ICG

Predicting Extreme Earthquake Events


Wenchuan earthquake, 2008


Bam earthquake, 2003


Pakistan earthquake, 2005

Earthquake is the most dangerous natural threat in urban areas

What is earthquake prediction?

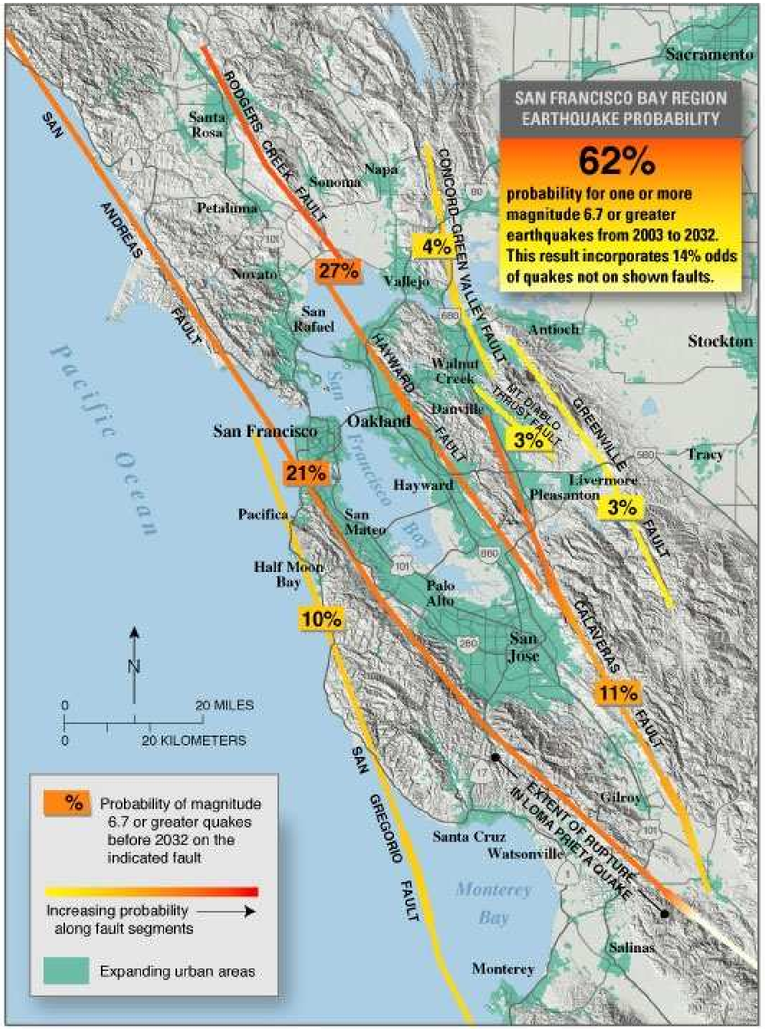
- To predict an earthquake is to prescribe, with justification and uncertainties, the time, place and size of a future earthquake, *including the probability that the event would happen*
- There was great optimism in the late1970s after an apparently successful earthquake prediction in Tangshan, China, in 1975; and many countries besides China launched large research projects after this
- Today this optimism is gone, but the research has nevertheless been very useful since it has given us a much better understanding of what drives (causes) earthquakes, and of the dynamic processes within the source region (the nucleation process)

Earthquake Prediction

- Earthquakes can (essentially) not be predicted
  - There are interesting exceptions, but these are of less importance in a risk mitigation context
- We now know much about the geographical distribution of earthquake exposure and about which levels of shaking can statistically be expected at a given location over a period of time (decades to centuries)
- This is the basis for anti-seismic building rules and codes, which are essential for saving lives and protecting societies
- Earthquake disasters are essentially man-made (in other words, earthquakes are natural, disasters are not)

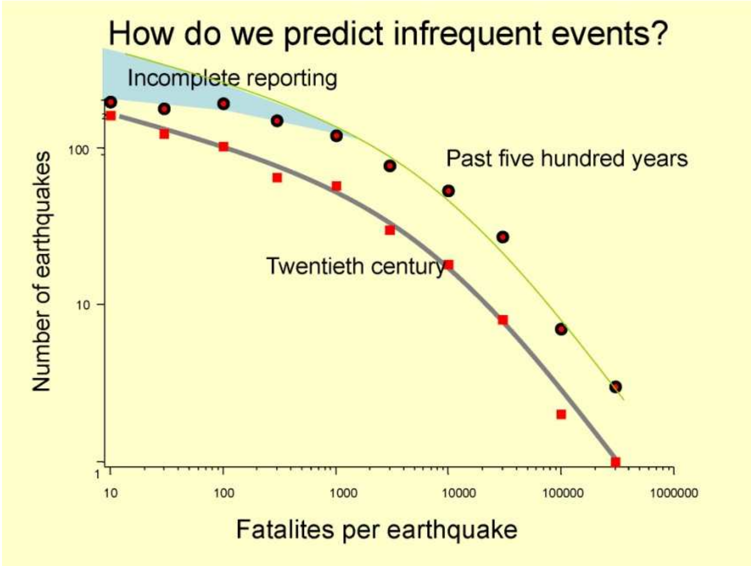
Therefore we focus on earthquake hazard assessment, not earthquake prediction

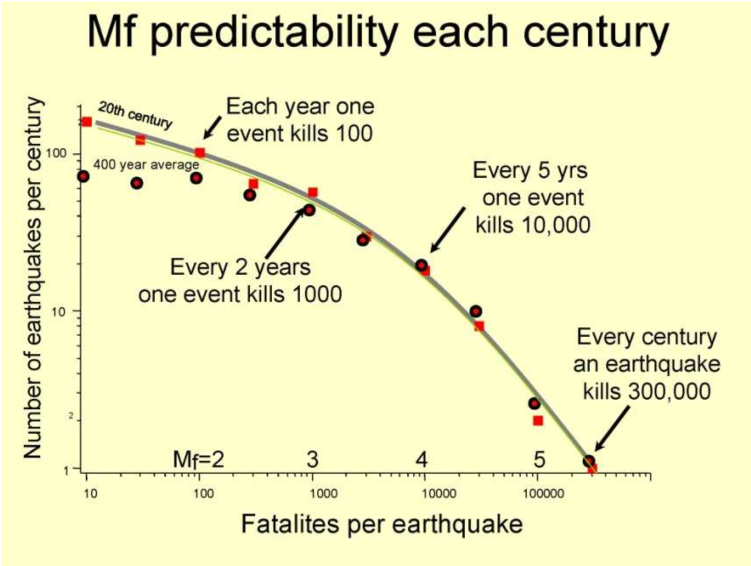
- This can be done statistically by computing probabilities for future shaking, with associated uncertainties
- One can either assume that the earthquakes are Poisson-distributed ('memory free'), or
- One can assume that the probabilities are time-dependent; i.e., they increase with time after the last major earthquake (implying earthquake cyclicity)



The challenges are still formidable:

- Earthquake hazard is based on statistical expectance and the uncertainties increase significantly with decreasing probabilities
- Sensitive installations can be protected even against events with 10^{-7}/year probability, but ordinary buildings are not even designed for ~10^{-2}/year
- Population increase and urbanization in earthquake exposed regions outweigh our greatly improved mitigation knowledge
- Statistical methods must be combined with deterministic methods (understanding earthquake geology) for predicting the most extreme events

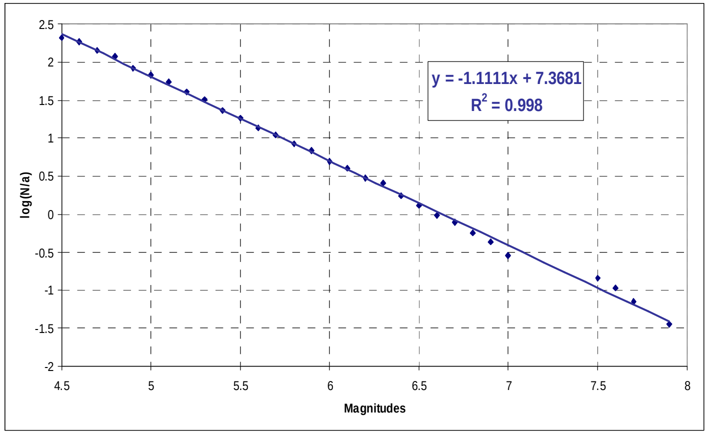


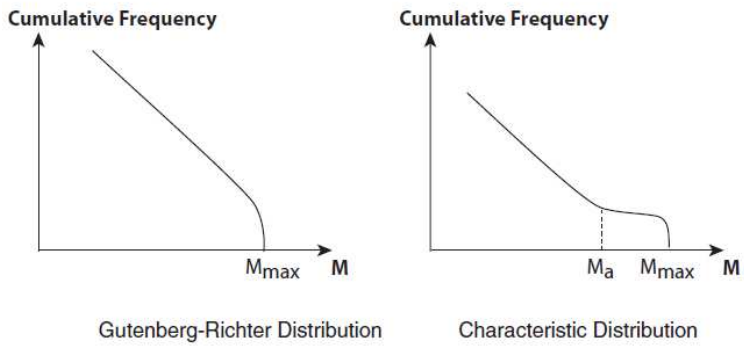


Courtesy of Roger Bilham

Temporal distribution of seismic events: Modified Gutenberg-Richter distributions

- The Gutenberg-Richter distribution describes the relationship between the earthquake magnitude and its return period or frequency
- The uncertainties related to establishing the return periods of events of large magnitudes and return periods are very large
- The frequency-magnitude distributions at large return periods (tapered vs. characteristic) are subject to scientific dispute
- The tapered distribution is often used in conventional seismology, whereas the characteristic is often supported by paleo-seismologists
- The two distributions lead to very different statistics-based predictions

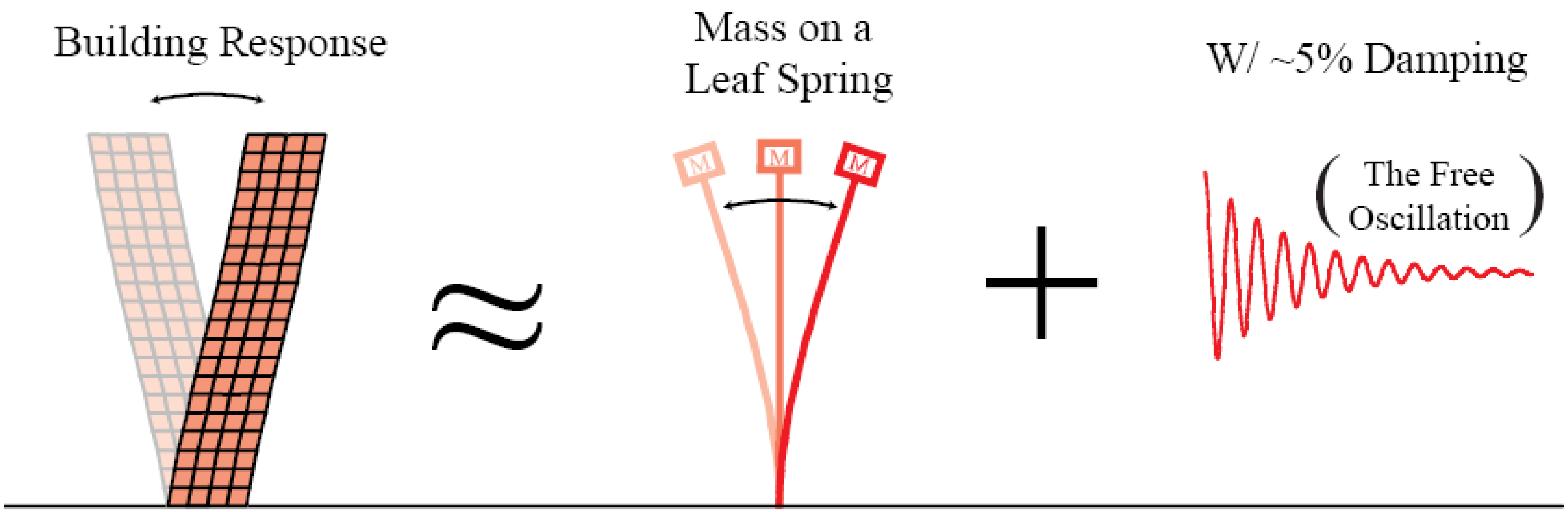




From Tsunami Pilot Study Working Group (2006)

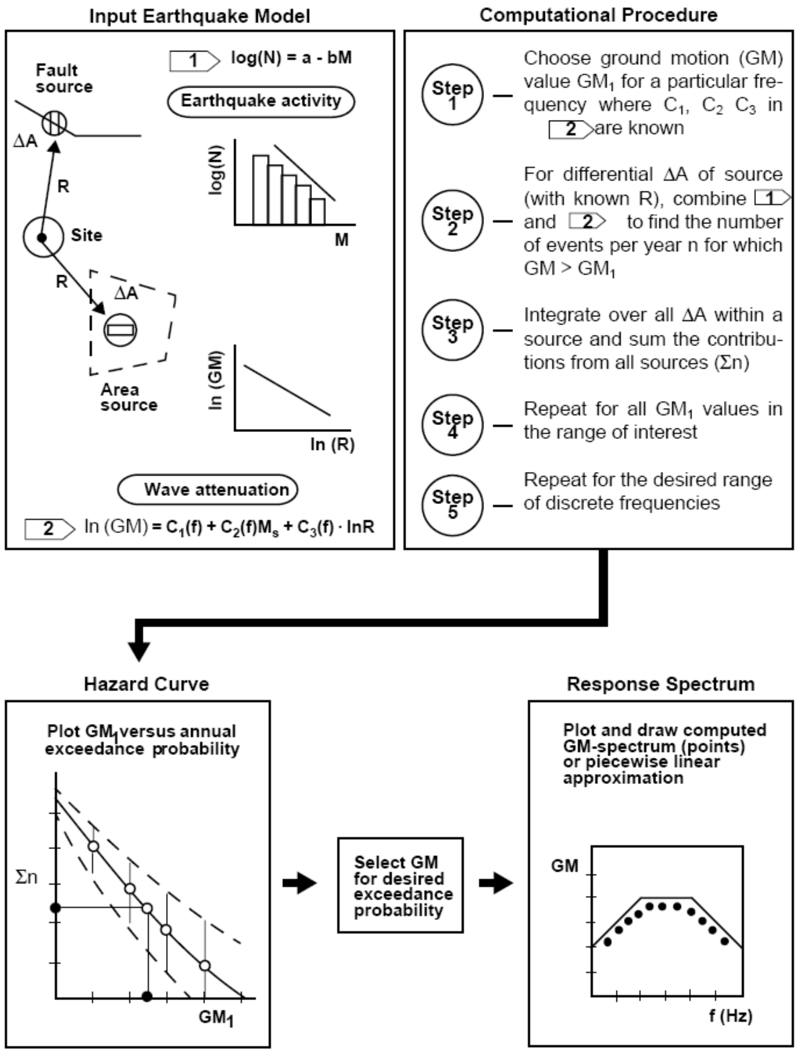
Probabilistic Seismic Hazard Assessment: Ground Motion Parameters

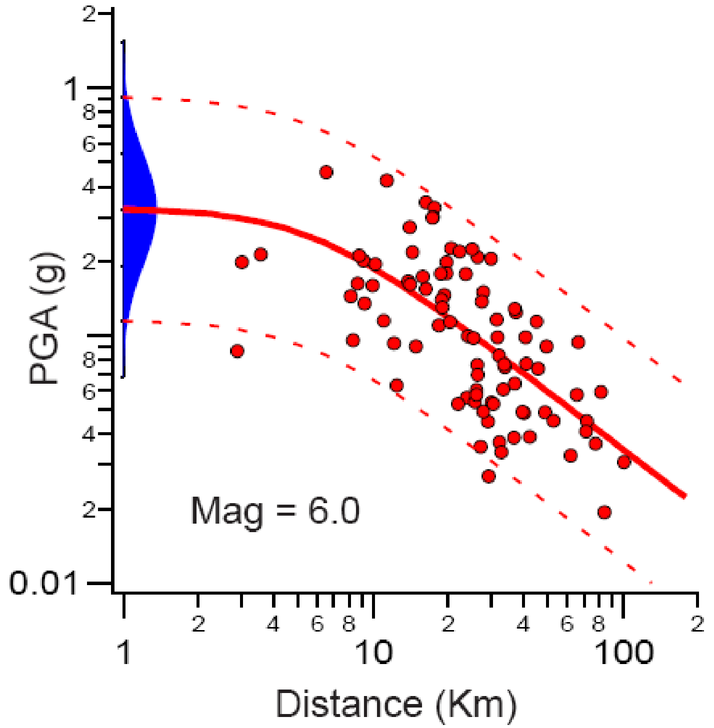
Probabilistic Seismic Hazard Assessment focuses on predicting the probability of occurrence of a ground motion parameter (e.g. peak ground acceleration (PGA), spectral acceleration, …), rather than a specific earthquake event



Load (input): PGA, Load effect (response): Spectral acc.

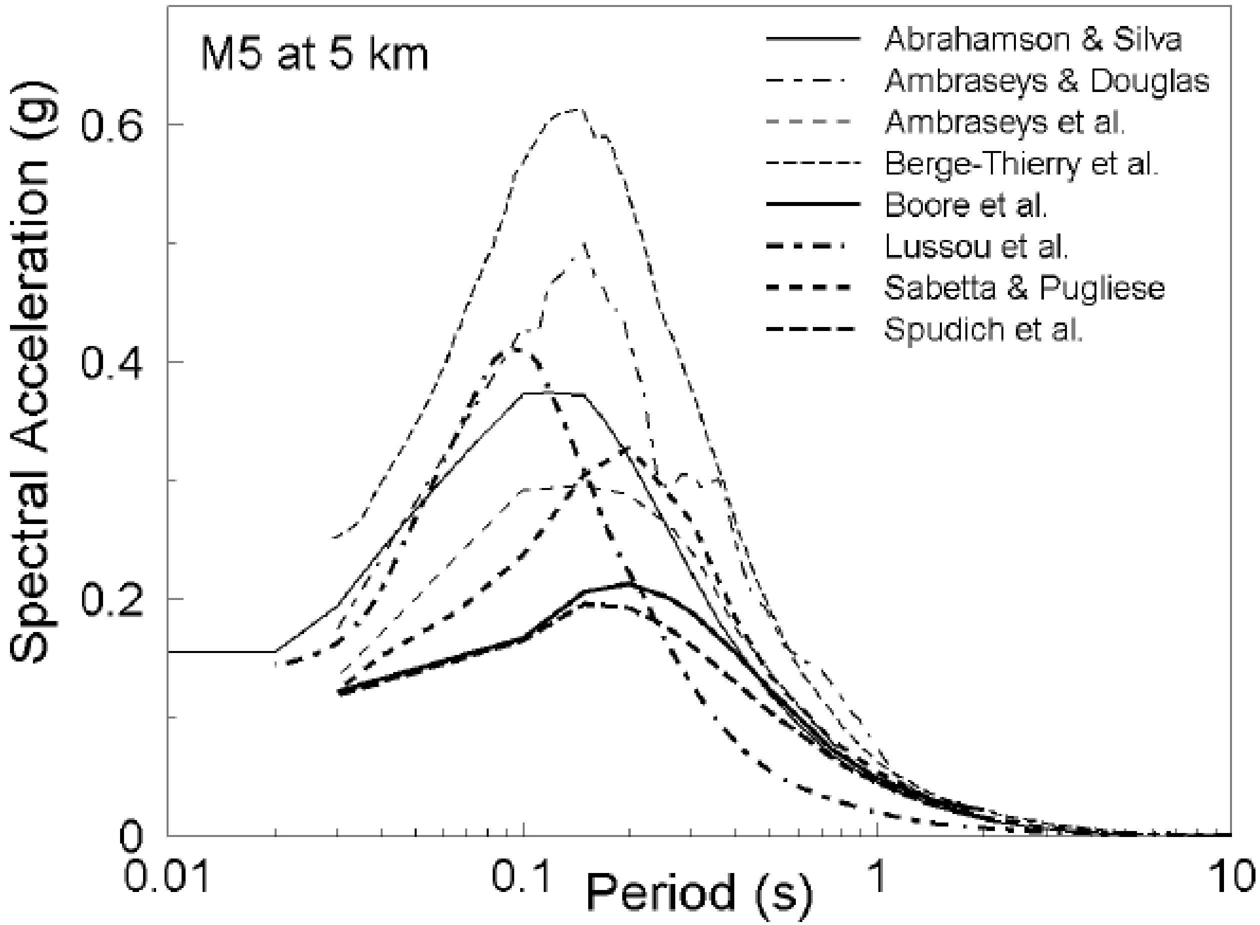
Different Steps in PSHA



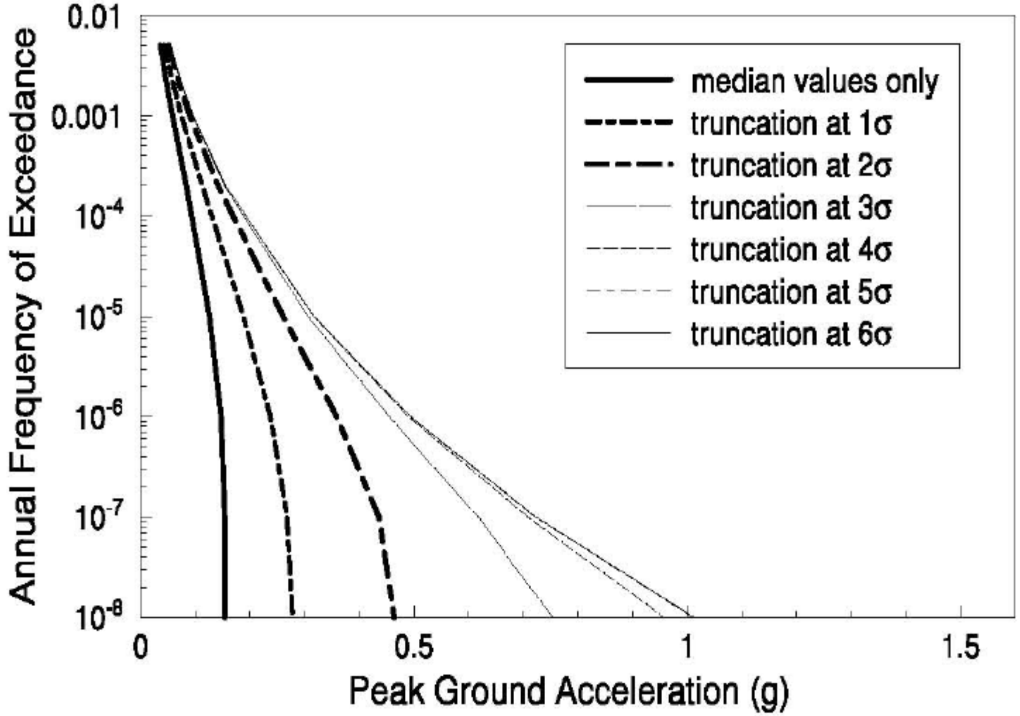


Attenuation curve for peak ground acceleration (PGA)

PSHA – Epistemic uncertainty due to uncertainty in relevant attenutation model



PSHA – Uncertainty in predicted PGA values at very low probabilities

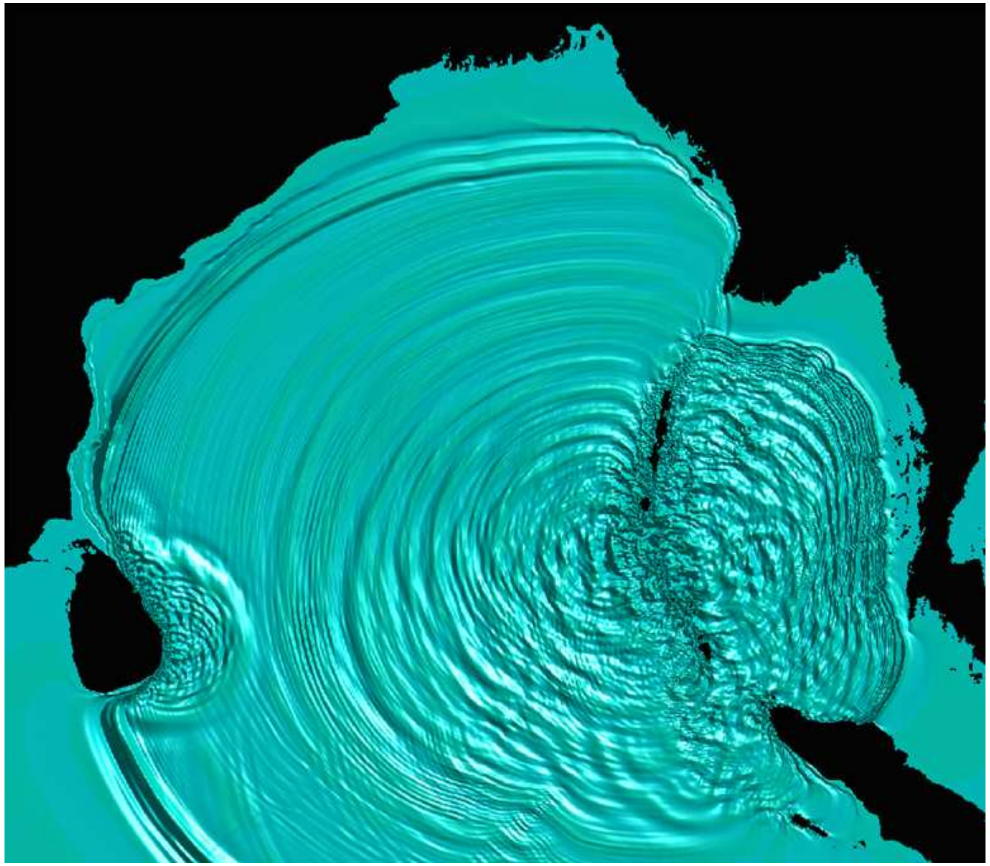


Seismic hazard curves derived using the ground-motion prediction equation of Ambraseys et al. (1996) truncated at different levels of scatter. The curves are for a site at 25 km from the boundary of a hypothetical seismic source zone with a maximum magnitude of 7.5 and a b-value in the recurrence relationship of 0.7 and an a-value of 2 (Bommer et al. 2005)

Tsunami Prediction

Tsunami means an (unusually large) wave in a harbour (Japanese)

Tsunami is a wave generated by huge and sudden displacement of water (e.g. earthquakes, slides, volcanoes, asteroids)



Recent tsunami damage in Asia

NGI ICG

Frequently asked questions in the aftermath of the 2004 Boxing Day tsunami

- How often would such an event happen in the Indian Ocean?
- Can we design an effective early warning system for tsunamis?

Main Challenge: Mechanics of mega-thrust earthquake events are not fully understood. Statistical models are not of much help if we do not understand the mechanics of the event.

Return periods for large earthquakes and tsunamis

- The largest earthquakes and tsunami are generated in the major subduction zones between the litospheric plates
- Possible approaches to establish the return period

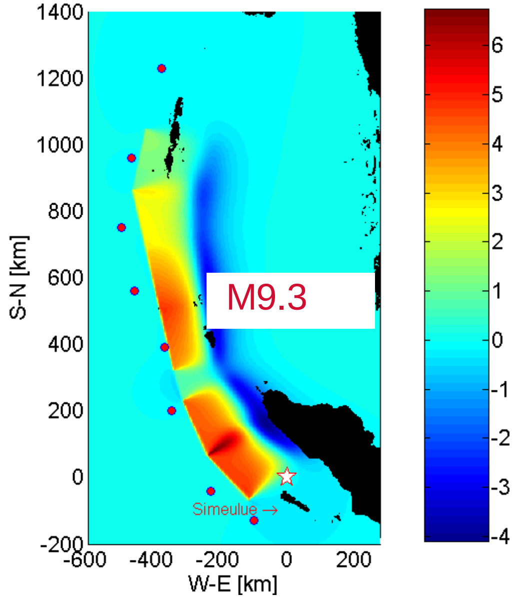
1. *Use the geologic subduction rate to estimate potential slip accumulation corresponding to a given earthquake magnitude*
2. *Utilise the temporal seismicity distribution*
3. *For tsunamis, establish the return period from tsunami records*

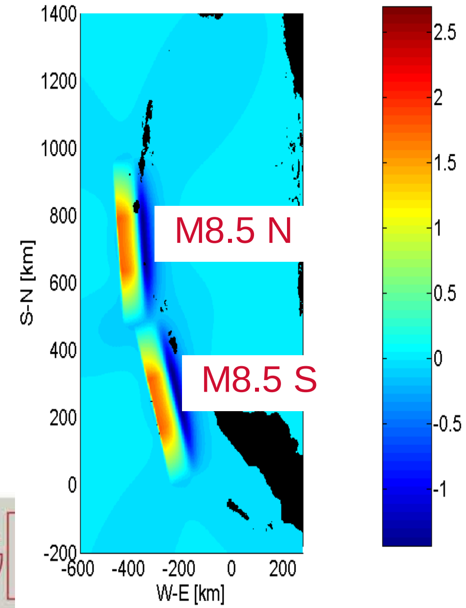
Return periods for large earthquakes and tsunamis

- **Challenges**
  - Modern seismological records are limited to the last 40-50 years
  - It is likely that different subduction zones exhibit different potential in generating extreme events, but their scaling laws are not yet fully understood
  - For tsunamis, the largest risk is often associated with return periods of several hundreds or even thousands of years

Challenges in assessing earthquake return periods in the aftermath of the 2004 Boxing Day event

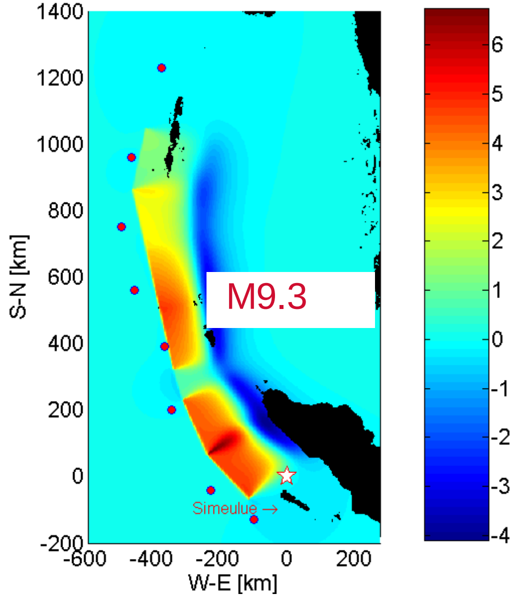
- **A new M9+ (megathrust) event in the subduction zone north of Sumatra will most likely not occur until at least 400 years from now**
  - *The main reason for this is that the 2004 rupture released accumulated stress over the whole segment*
  - *Supported also by new paleo-tsunamic research*

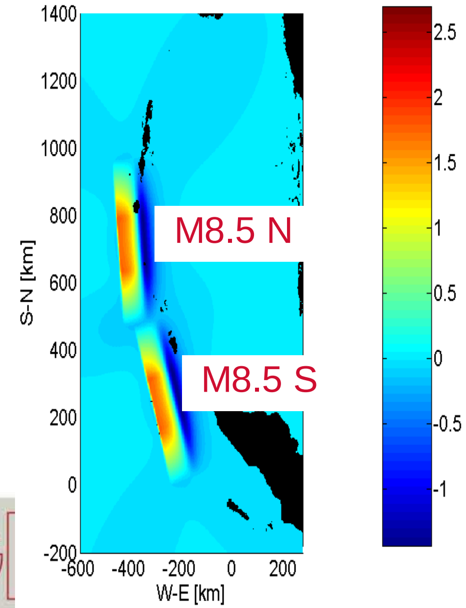




Challenges in assessing earthquake return periods in the aftermath of the 2004 Boxing Day event

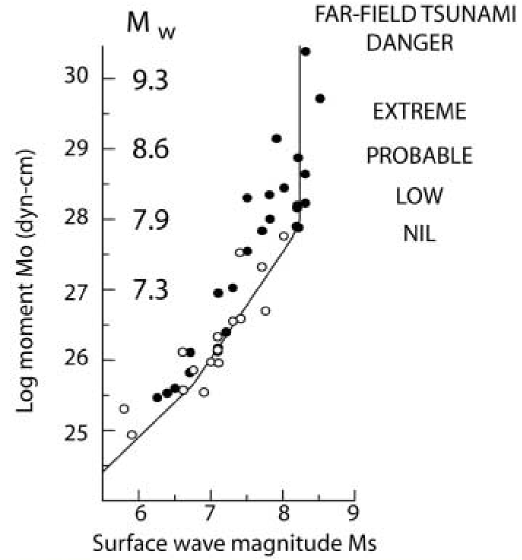
- **Smaller events of M~8.5 are more irregular in their occurrence**
  - Still, the probability of occurrence will be quite low for a long time after 2004, then increasing gradually with time
  - Return periods derived from seismicity ranges from 1300-2700 years (based on short records)
  - Return periods derived from convergence rates ranges from 200-400 years
  - The two methods predict almost one order of magnitude difference, where the slip rates define the lower limit

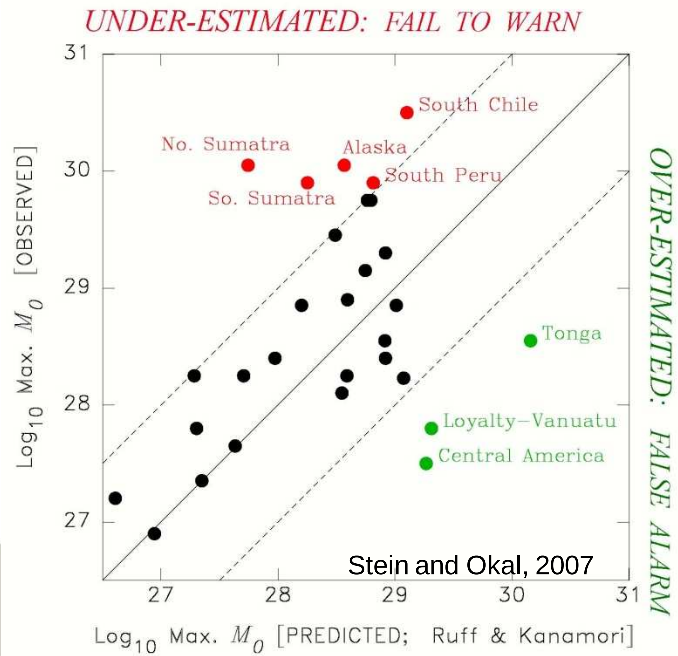




The scale and rupture of large earthquakes are not fully understood (after Stein and Okal, 2007)

- Most rapidly determined earthquake magnitudes (M_S) will underestimate the earthquake magnitude and tsunami potential → significant effects also on early warning
- The 2004 Boxing day tsunami violated the generally accepted scaling “law” predicting maximum magnitudes to scale with the speed and age (inversely) of the subducting slab





The scale and rupture of large earthquakes are not fully understood (after Stein and Okal, 2007)

- Shift in paradigm → there is now scepticism towards deterministic scaling laws limiting the upper range in earthquake magnitude for a given fault or fault system
- Generation of large earthquakes should be considered possible for **all long continuous faults**
- Yet, it is likely that megathrust earthquakes will not occur where the oldest lithospheric plates subduct

The scale and rupture of large earthquakes are not fully understood (following Stein and Okal, 2007)

- Many of the largest recorded earthquakes are “slow”, related to large deformations relative to its magnitude and coupling to sediments
  - It seems that slowness is a necessity for megathrust potential
- For “slow” rupturing earthquakes the tsunami potential is larger than for “normal” earthquakes
  - This has implications for warning; the moment magnitude is not solely reflecting the tsunami potential

Final Remarks
(from draft IPCC SREX report)

…. A central concern is that climate change has introduced substantial non-stationarity into risk management decisions. Non-stationarity is the realization that past experiences may no longer be a reliable predictor of the future character and frequency of events; it applies both to hazards and to the response of human systems to same. As climate change is expected to change the frequency, magnitude, and other characteristics of extreme events, some of which are associated with extreme impacts, risk management strategies must accommodate a shifting distribution of the latter.

Final Remarks

- Natural hazard events do not always obey the “laws” that scientists establish for them, nor do they know about the statistical distributions that they are supposed to follow.

- If your data do not fit your model, plot them on a log-log diagram using very large symbols for your data points.


THANK YOU
FOR YOUR ATTENTION