

SCOR Papers

By **Guillaume Ominetti**
Head of Capital Shield Strategy,
SCOR SE

Texts appearing in SCOR Papers are the responsibility of their authors alone. In publishing such articles, SCOR takes no position on the opinions expressed by the authors in their texts and disclaims all responsibility for any opinions, incorrect information or legal errors found therein.

Are great earthquakes clustered?

Introduction

“Monte Carlo simulations confirm that the global temporal clustering of great shallow earthquakes during 1952-1964 at $M_w \geq 9.0$ is highly significant (4% random probability) as is the clustering of the events of $M_w \geq 8.6$ (0.2% random probability) during 1950-1965. [...] Immediately after the 1950-1965 cluster, significant quiescence at and above $M_w 8.4$ begins and continues until 2001 (0.5% random probability). [...] These observations indicate that, for great earthquakes, Earth behaves as a coherent seismotectonic system. [...] The recent occurrences [...] confirm that we have entered a new period of high moment release and probable temporal clustering of mega-quakes.”
Charles G. Bufe and David M. Perkins, 2005.

The rate of very powerful earthquakes occurring worldwide has significantly increased over the past decade. In fact, the record of these very large seismic events seemingly reveals periods of cluster and quiescence during the entire last century. These observations have fueled concern that these great quakes may not be independent events and may cluster in time. If this feature turned out to be true, it would have a major impact on how we assess seismic hazard.

A number of academic studies have investigated this question, in particular since 2005 in the wake of the 2004 M_w 9.1 Sumatra-Andaman mega earthquake.

Seismologists and geophysicists have tried to infer whether an underlying physical phenomenon could drive such clustering. However, no clear consensus has been reached so far.

In parallel, statisticians have also studied the question. Conclusions have been mixed, with some researchers asserting that there is conclusive evidence of great earthquake clustering, while some others reckon that what has been observed so far cannot be distinguished from a standard stochastic process, with no memory and constant risk over time, and consequently no clustering pattern.

This paper first gives some reminders and general facts on seismicity and its measurement. It then focuses on the topic of great earthquakes and investigates the debate on their clustering.

Reminder on metrics characterizing earthquake severity

The severity of an earthquake is described by both magnitude and intensity. These two frequently confused terms refer to different, but related, observations.

Magnitude

Magnitude, expressed as an Arabic numeral, characterizes the size of the earthquake.

There are several different concepts of “magnitude”¹, notably local (or Richter) magnitude, surface wave magnitude, body wave magnitude and moment magnitude, which all relate to different measurement methods for specifying the energy released by the temblor. This is why the same earthquake may have different magnitude values. However, each time a new magnitude was introduced in seismology, it was calibrated to be roughly consistent with the existing ones, so that all magnitude scales approximately give the same value for a given quake. For large earthquakes (magnitude above 7), the concept in use is the so-called *moment magnitude* (M_w). It is the one we will use throughout the document.

Due to its logarithmic nature, this metric is deceptive, in the sense that earthquakes with “close” moment magnitudes may be materially different in size. Indeed, a one-step increase on this scale means that seismic size is multiplied by $10^{1.5} \approx 32$.

So a M_w 9.0 earthquake is about 32 times more powerful than a M_w 8.0 earthquake and 1000 times more powerful than a M_w 7.0 earthquake. Therefore, we see that even a large M_w 7.0 earthquake has nothing to do with the class of M_w 8.5+ earthquakes.

The 1960 Valdivia earthquake, also known as the *Great Chilean Earthquake*, is to date the most powerful earthquake ever measured on a seismograph, reaching M_w 9.5 moment magnitude on 22 May 1960. The energy released was more than twice that of the next most powerful seismic event, the 1964 M_w 9.2 Good Friday Earthquake, which was centered in Prince William Sound, Alaska.

In this paper, we will qualify an event to be a great earthquake if its moment magnitude M_w is greater than 8.2 (by analogy with most academic papers on this topic).

The term of giant earthquake, or mega earthquake (albeit not widely used in seismology), is usually used for M_w 9.0 and above seismic events.

Intensity

By contrast, intensity, expressed as a Roman numeral, represents the severity of the shaking resulting from the earthquake and indicates the local effects and potential for damage on the Earth’s surface as it affects humans, animals, structures, and natural objects.

Just like magnitude, there are different definitions for an earthquake’s intensity. However, the one most commonly used by seismologists is the so-called *Modified Mercalli Intensity* (MMI) scale². Any given earthquake can be described by only one moment magnitude, but many intensities, since the earthquake effects vary with circumstances such as distance from the epicenter and local soil conditions.

¹ See appendix for details

² See appendix for details

Plate tectonics

As shown on the figure below, the Earth is composed of different spherical layers characterized by different physical/chemical properties:



- The crust: the outermost solid shell;
- The mantle: the highly viscous layer between the crust and the outer core. It is divided into 2 sub-sections (the upper mantle and the lower mantle) and constitutes about 84% of Earth's volume;
- The core: the inner spherical part composed of a liquid outer core that is much less viscous than the mantle, and a solid inner core.

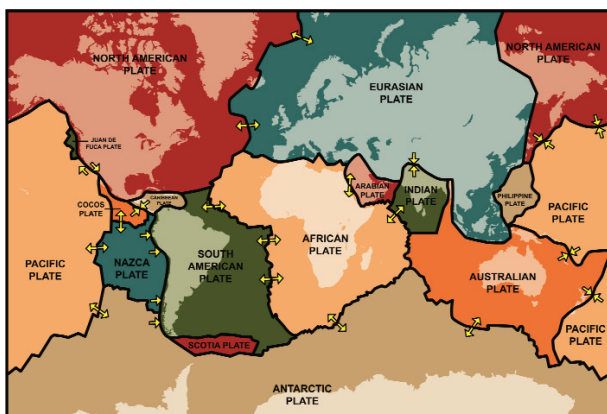
The outer layers of the Earth are also divided into lithosphere and asthenosphere:

The lithosphere is cooler and more rigid, while the asthenosphere is hotter and flows more easily.

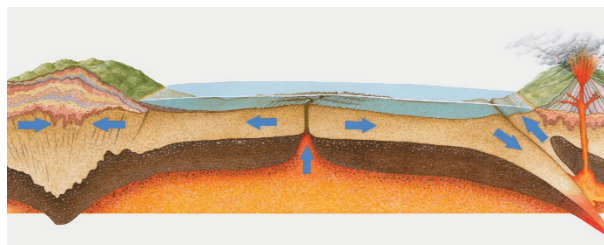
- The lithosphere is composed of both the crust and a portion of the upper mantle;
- The asthenosphere is the remaining portion of the upper mantle, located just below the lithosphere and above the lower mantle, at depths typically between 100 and 200 km below the surface.

The lithosphere is broken up into tectonic plates. On Earth, there are seven or eight major plates (depending on how they are defined) and many minor plates (see figure below).

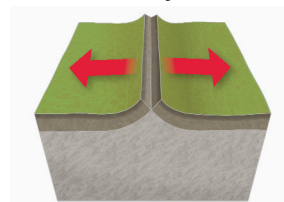
Earth's lithosphere is broken up into tectonic plates. On Earth, there are seven or eight major plates (depending on how they are defined) and many minor plates (see figure below).



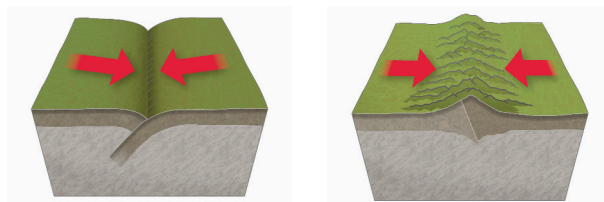
These plates 'float' on the fluid-like (viscoelastic solid) asthenosphere and move very slowly relatively to each other. Plate motions range up to a typical 1–4 cm/yr (Mid-Atlantic Ridge), to about 16 cm/yr (Nazca Plate; about as fast as hair grows). Where plates meet, their relative motion determines the type of boundary: convergent, divergent, or transform. Earthquakes, volcanic activity, mountain-building, and oceanic trench formation occur along these plate boundaries.



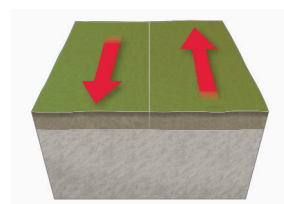
- Divergent boundaries are areas where plates move away from each other, forming either mid-oceanic ridges or rift valleys;



- Convergent boundaries are areas where plates move toward each other and collide. These are also known as compressional or destructive boundaries;



- Transform boundaries occur when two plates grind past each other with only limited convergent or divergent activity.



Subduction zones occur at convergent boundaries where an oceanic plate meets a continental plate and is pushed underneath it³. Subduction zones are marked by oceanic trenches. The descending end of the oceanic plate melts and creates pressure in the mantle, causing volcanoes to form.

For very large earthquakes, we focus on the so-called 'megathrust' faults, which are the boundaries between subducting and overriding plates at subduction zones.

³ Subduction is the process of the oceanic lithosphere colliding with and descending beneath the continental lithosphere.

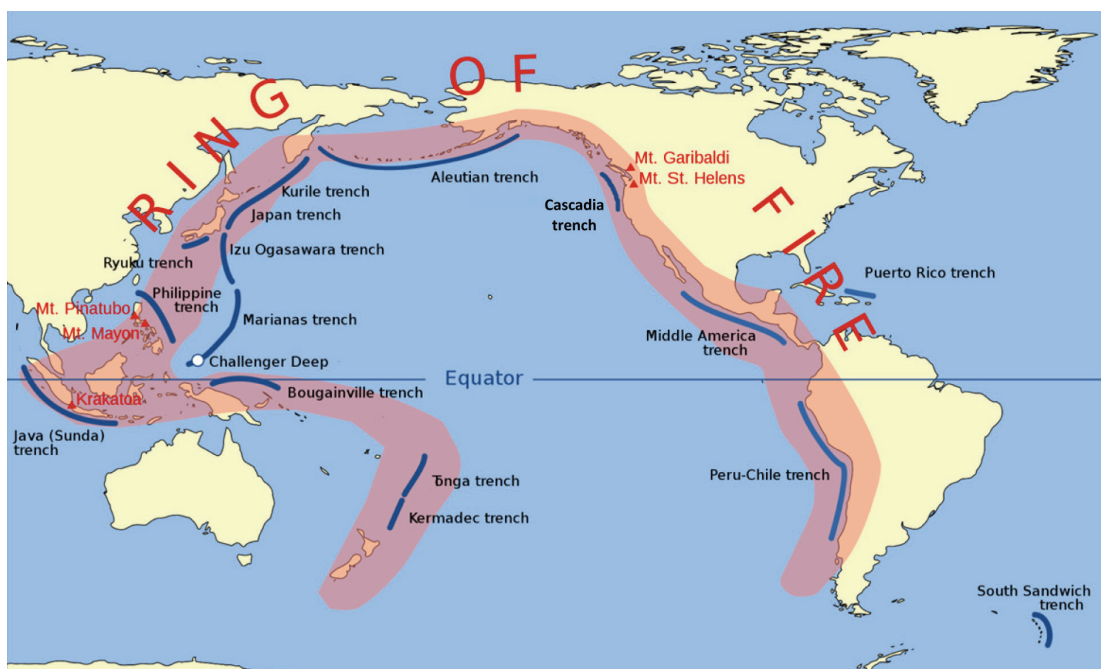
The Ring of Fire

Earthquakes can strike any location at any time. But history shows they occur in the same general patterns year after year.

The world's greatest earthquake belt, the circum-Pacific seismic belt, is found along the rim of the Pacific Ocean. It has earned the nickname "Ring of Fire". In a 40,000 km horseshoe shape, it is associated with a nearly continuous series of oceanic trenches, volcanic arcs, and volcanic belts.

The belt extends from Chile, northward along the South American coast through Central America, Mexico, the West Coast of the United States, and the southern part of Alaska, through the Aleutian Islands to Japan, the Philippine Islands, New Gui-

nea, the island groups of the Southwest Pacific, and to New Zealand. About 90% of the world's earthquakes occur there. More importantly, all 17 earthquakes but one (in Tibet) with moment magnitude M_w 8.5 or above since 1900 have stricken on this Ring of Fire. However, one should not conclude all great earthquakes occur in this Pacific area: the 1755 Great Lisbon Earthquake, which was followed by fires and a tsunami, and which almost totally destroyed Lisbon and adjoining areas, had an estimated moment magnitude in the range 8.5–9.0 according to seismologists. This event is presented more thoroughly in the appendix.



The data

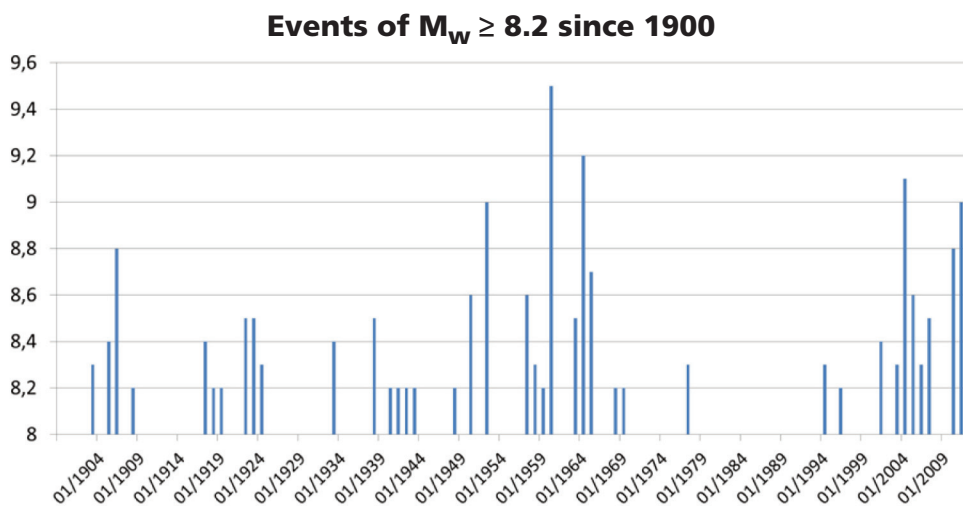
The pattern of great earthquakes described in this paper is based on the USGS (U.S. Geological Survey) catalog for 1900 to today seismicity.

Although the occurrence of very big quakes is obviously known before 1900 (e.g. the Great Lisbon Earthquake), be it by written records or even geological surveys, we have limited the beginning date of the dataset to 1900 (as it is commonly done in similar studies) in order both to consider earthquakes whose moment magnitudes are known with an 'acceptable' level of accuracy and to be confident about the completeness of the dataset (some very big quakes having occurred a long time ago in desert or scarcely populated areas may have gone 'unnoticed' and are therefore not recorded in the historical catalog).

Studying the great earthquakes ($M_w \geq 8.2$) having occurred since the beginning of the last century implies that:

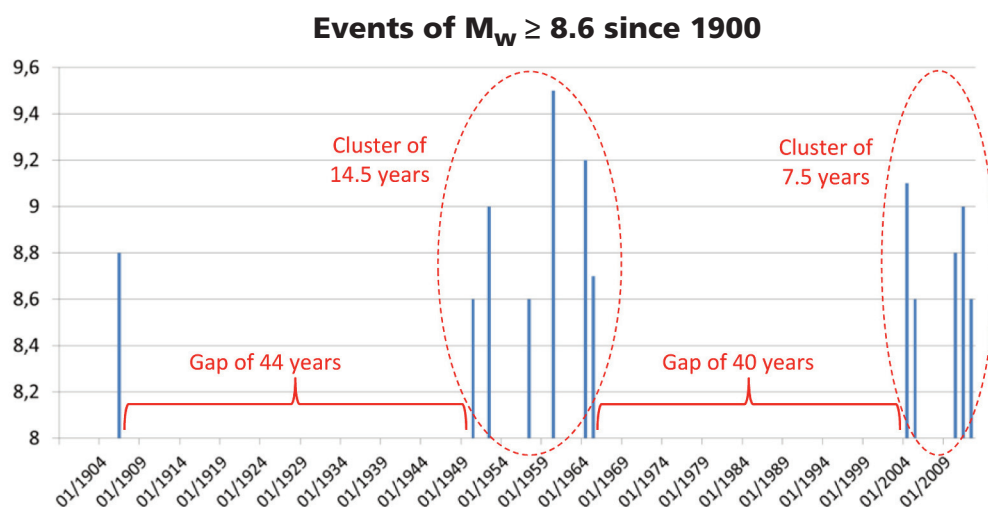
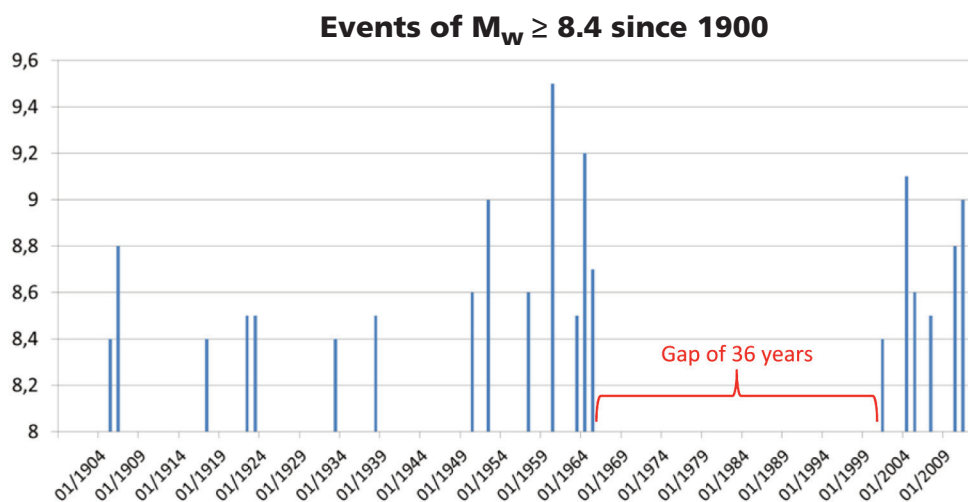
- The dataset is quite small (only 45 events), with less than one event every 2 years on a worldwide basis. Furthermore, the number of events decreases very sharply when the magnitude cutoff is increased: for instance, there are only 17 earthquakes of $M_w \geq 8.5$ and 7 earthquakes of $M_w \geq 8.8$ in this period of time;
- Catalog completeness issues that may arise for smaller magnitudes are strongly reduced inasmuch as quakes with $M_w \geq 8.2$ have obviously not gone 'unnoticed' since 1900.

Below is the graphical representation of the $M_w \geq 8.2$ earthquake catalog we use in this study. All event sizes are given as moment magnitudes.



If we extract the resulting sub-catalogs by removing all events below a given magnitude thresh-

hold, the clusters of events interspersed with quiescent periods become more evident:



Looking at this dataset and its features, one question immediately comes to mind: is the random process of very big quakes a standard random process with uniform risk over time, or, on the

contrary, is it likely that these great seismic events cluster in time on a global basis, with successive high activity and quiescence cycles?

Various conclusions in the released studies

In an article published in June 2005 in the *Bulletin of the Seismological Society of America*, entitled "Evidence for a Global Seismic-Moment Release Sequence", Charles G. Bufe and David M. Perkins favor the second option: "Monte Carlo simulations confirm that the global temporal clustering of great shallow earthquakes during 1952-1964 at $M_w \geq 9.0$ is highly significant (4% random probability) as is the cluste-

ring of the events of $M_w \geq 8.6$ (0.2% random probability) during 1950-1965. [...] Immediately after the 1950-1965 cluster, significant quiescence at and above $M_w 8.4$ begins and continues until 2001 (0.5% random probability). [...] These observations indicate that, for great earthquakes, Earth behaves as a coherent seismotectonic system."

Other articles investigating the question have been released since. Some tend to confirm Bufe and Perkins' thesis, while some others reject it. The general approach used in those papers is to perform different statistical tests on the dataset in order to infer whether the record is likely to be consistent with a homogeneous Poisson process (i.e. a standard stochastic process, with no memory and constant risk over time, and consequently no clustering pattern – see appendix for more details). The CAT modeling company

EQECAT also reopened the debate by releasing, in October 2011, a white paper and an article⁴ dealing with this topic.

The alarming conclusion of some papers/articles (namely, the clustering feature cannot be denied and should be incorporated from now on by large insurance and reinsurance carriers in their risk management and underwriting plans) has led SCOR to also investigate this topic in order to assess the relevance of the different arguments.

⁴ *Spatial and temporal earthquake clustering: part 1 – Global earthquake clustering* by Paul C. Thenhaus, Dr. Kenneth W. Campbell and Dr. Mahmoud Khater
Global clustering of giant earthquakes: a revolution in earthquake risk management? by Kate Stillwell

What to believe?

First, it is worth mentioning that the clustering of earthquakes is well-known in seismology, at local and regional level, with the so-called mainshock-after-shock(s) sequences. When a big seismic event (mainshock) occurs, it mechanically triggers smaller shocks (aftershocks) on the same or adjoining faults by stress transfer mechanisms. This is why the risk of earthquake activity is known to be locally higher in the aftermath of an event.

Therefore, usually:

- The occurrence of mainshocks is modeled by a homogeneous Poisson process: it is the simplest form of discrete arrival process, with uniform risk over time and no memory (an event which occurred in the past does not influence the occurrence of events in the future);
- Then, conditionally on the fact that a mainshock has occurred, seismic risk is increased locally to model the likely occurrence of (smaller) aftershocks in the region of the mainshock.

So usually statisticians first do what they call a 'local declustering' of the historical catalog. This means that they remove from it all the supposed aftershocks⁵. It is only after this stage that they test whether earthquake occurrence is Poissonian (meaning with constant risk over time, no memory) in the resulting filtered catalog.

For our study in this paper, to be in line with Bufe's and Perkins' study, and also to avoid any debatable procedure on the data, we will work with the raw catalog of $M_w \geq 8.2$ earthquakes, i.e. without removing any event. We obviously expect the majority of these great earthquakes to be mainshocks anyway (indeed, most consecutive events in the record occur very far away from each other). This is the reason why the very large earthquakes have (usually) always been supposed to be 'independent' events for which the occurrence is Poissonian (it is this thesis that is actually rejected by Bufe and Perkins). But still, we should be aware that some events in our big quake catalog might be correlated in a kind of mainshock-after-shock sequence. For instance, two M_w 8.4 events occurred on 9th July 1905 in Mongolia. Similarly, the M_w 9.1 event in Sumatra in December 2004 was followed by a M_w 8.6 temblor in the same region in March 2005.

It is actually not really a problem not to remove the 'supposed' aftershocks, given that leaving them may only increase the clustering feature in the catalog (if there is any). So in the end, if our analysis suggests that there is such a feature, we may investigate further whether it still holds when removing the supposed aftershocks. On the contrary, if our analysis suggests that there is no such feature, it will mean that our conclusion obviously still holds when 'declustering locally' the catalog.

⁵ Which is far from being trivial, since the classification of events as mainshocks/aftershocks is obviously debatable

So does the raw historical catalog of great earthquakes show a clustering pattern?

From a statistical perspective, it is intrinsically impossible to say whether there is a clustering feature with certainty. All we can do is to infer whether it is likely or not that the physical underlying process having generated the known record is associated with such a feature, but nothing more. Therefore, one cannot be assertive.

It is true that when looking at the historical records on pages 5-6, a cluster feature looks quite credible. However, we should be aware that our mind is not good at all for judging randomness.

We naturally tend to be biased by thinking that a standard uniform process should only generate 'uniform' outcomes, with a regular dispersion of events over the time period. But in fact, even with a standard random process, each specific outcome (including the one with equidispersion of events) is *de facto* improbable, so that a seismic record with an approximate constant interoccurrence time between seismic events would be actually as (if not more) strange as the one we are observing (see appendix "Randomness is not regularity" for an anecdote on this topic).

A probabilistic analysis may be performed to go forward.

But before going further, we can already say we believe that Bufe's and Perkins' analysis is not fully relevant. This can be seen even without a detailed mathematical study.

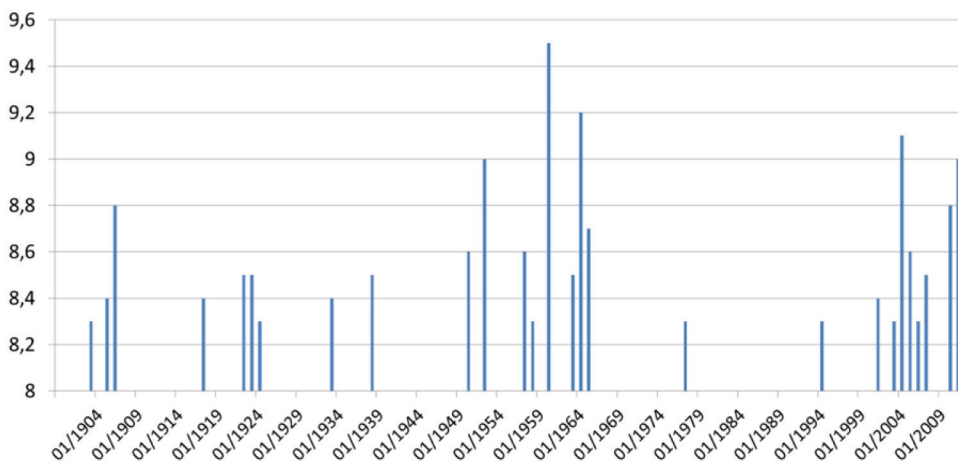
The reason is that the features of the historical catalog used for their statistical analysis (and pointed out to be improbable), have been selected after looking at the record. This is a very important point that we are going to explain now.

If we look at the graphs on page 6, the periods of gaps and the clusters are indeed obvious for the sub-catalogs of events with $M_w \geq 8.4$ and $M_w \geq 8.6$, respectively. But the key point to understand is that the thresholds 8.4 and 8.6 have not been chosen 'randomly' by Bufe and Perkins: they are associated with the 'most striking' features.

If we look at the events with $M_w \geq 8.3$ for instance, we get the sub-catalog below in which the big gap between 1965 and 2001 exhibited with the $M_w \geq 8.4$ record does not exist anymore due to two $M_w 8.3$ events in 1977 and 1994, respectively... This record already appears far less 'strange'.

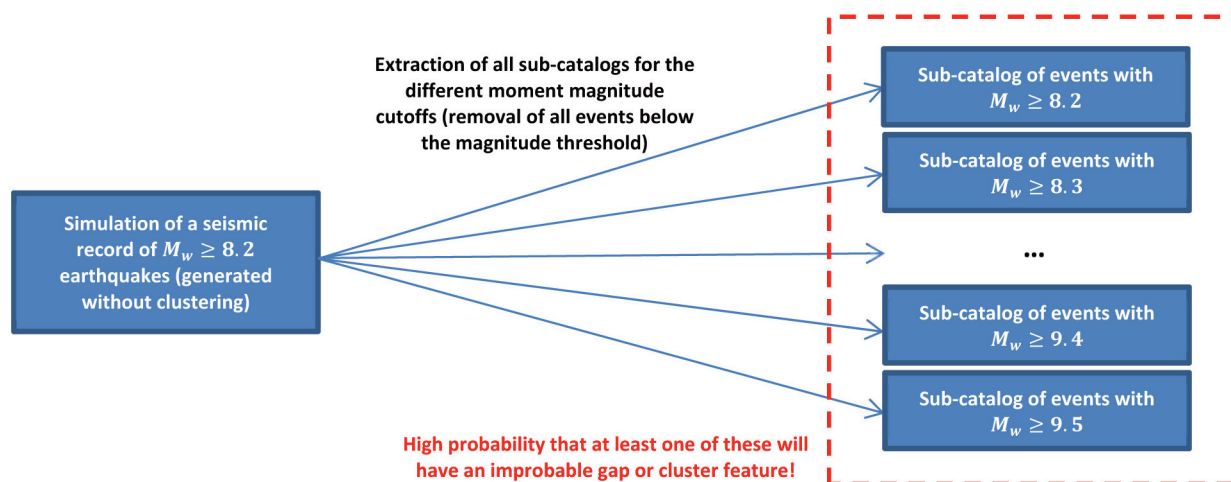
Hence, Bufe and Perkins have selected the magnitude cutoffs (and sizes of clusters) in order to maximize the apparent anomaly of the historical record.

Events of $M_w \geq 8.3$ since 1900



One can demonstrate that, with a standard random process, it is usually easy (or at least, not rare) to exhibit for one magnitude threshold a clustering/gap feature associated with a low probability.

This is the key point of our quantitative analysis: if we randomly simulate a seismic record with no clustering feature, in most cases, it is possible to select⁶ a magnitude threshold such that the sub-catalog of events above this magnitude in the simulated record presents a clustering (resp. gap) feature which is improbable (i.e. which had retrospectively a very low probability to occur).



So in the end, exhibiting an improbable feature for the historical catalog is not very surprising: we could generically do the same with a catalog generated without clustering...

The quantitative assessment of the bias that we have just described can be found in the appendix. Our finding is that, with a standard Poissonian random process (constant risk over time, no memory, no clustering), if we simulate the occurrence of earthquakes since 1900, and if we analyze the unlikelihood of gaps/clusters like Bufe and Perkins, there is approximately a 8-10%

chance to be able to exhibit a gap/cluster feature in the simulated catalog that is more improbable than what has been pointed out in the historical record.

So while still low, this 8-10% statistical significance is far from being in the range of the figures that Bufe and Perkins state in their article (0.2%, 0.5%, 4%).

In the end, we see we cannot say that the gap/clusters features of the historical catalog are "statistically highly significant". They remain in fact quite 'standard'.

⁶ It is at this stage that the choice is *ex post*, meaning after observing the outcome

And from a physical point of view?

It is common knowledge in seismology that large earthquakes may trigger other seismic events:

- As already mentioned, there are first the so-called aftershocks, meaning smaller earthquakes which occur after a previous large earthquake (known as the mainshock) in the same area. They are a consequence of the changes in the local stress field in the aftermath of the massive displacement associated with the mainshock. Even if they are (by definition) smaller⁷ than the mainshock, aftershocks may be still large seismic events. It is of course difficult (and up to debate) to decide upon a fixed criterion (distance window/time window) which would enable to decide whether a smaller quake after a big event is an aftershock or, on the contrary, should be considered as an 'independent' event. However, there is a broad consensus that the distance window for aftershocks should be similar to the rupture length of the mainshock.

Empirical Bath's law states that the difference in moment magnitudes between a mainshock and its largest aftershock is approximately constant, whatever the size of the mainshock is, in the range of 1.2, i.e. $M_{w}^{\text{mainshock}} - M_{w}^{\text{largest aftershock}} \approx 1.2$;

- Second, it is known that small earthquakes can be triggered dynamically at large distances by the seismic waves caused by large earthquakes. But these are far smaller than the mainshock so that large earthquakes are not believed to be possibly triggered remotely by such passing energy waves. One recent study found that it was clearly not the case (*Absence of remotely triggered large earthquakes beyond the mainshock region*, by Parsons, T. and A. A. Velasco, 2011). Furthermore, there is anyway a problem of timing. The seismic waves are traveling at high speeds (above 1km/s) so that they obviously cannot account for a clustering that would have time spans of years or decades. This is the main 'problem': if the clustering of great earthquakes turned out to be a reality, given the time windows observed in the last century, this would mean that the underlying physical phenomenon would allow transfers of mechanical stress on a global scale over years or even decades.

This is why some scientists have developed the hypothesis that the clustering could be caused by stress diffusion through post-seismic relaxation of the asthenosphere (the layer region of the upper mantle at average depths between 100 and 200 km below the surface). Such stress pulses are indeed traveling very slowly. An article dated 1998⁸ suggests that the great subduction earthquakes from 1952 to 1965 along the Aleutian arc and Kurile-Kamchatka trench generated such a stress pulse which reached California in the mid-80s and may have increased seismic activity there. However, this thesis is still under debate. Furthermore, the underlying stress changes are usually considered far too small for being able to generate great earthquakes. Last, but not least, there is still a problem of timing! Indeed, those pulses may be traveling too slowly (!) for generating global clusters of quakes over a few years. Indeed, great earthquakes in the clusters are occurring very far away from each other (e.g. 2004 Sumatra, 2010 Chili, 2011 Japan...), knowing that it seemingly took more than 20 years for such a post-seismic stress pulse just to go from Northern Pacific to California...

In the end, there is currently no consensus among seismologists and geophysicists upon the supposed physical phenomenon which could match such an effect.

Therefore, there is neither evidence nor known and broadly accepted physical phenomenon which could drive such a dynamic triggering of great earthquakes in clusters.

⁷ If an aftershock is larger than the mainshock, the aftershock is redesignated as the mainshock and the original mainshock is redesignated as a foreshock.

⁸ *Viscosity of Oceanic Asthenosphere Inferred from Remote Triggering of Earthquakes*, by Fred F. Pollitz, Roland Bürgmann and Barbara Romanowicz

Conclusion

We believe that the analysis done by Bufe and Perkins, which states that the clustering feature in the historical record of great earthquakes cannot be reasonably attributed to chance, is not relevant.

Some features of the historical catalog of great earthquakes are indeed 'strange', i.e. improbable, but this generically holds for most random realizations. Even with a uniform random process, each specific outcome is *de facto* improbable, so that contrary to what we could think at first blush, an equidispersion of seismic events over the last century would be at least as strange as what we are observing.

Furthermore, contrary to what we could have initially thought, there is no material statistical significance toward a clustering pattern in the historical catalog. The probability figures in Bufe's and Perkins' article (0.2%, 0.5%) are impressive but misleadingly low because they are the result of statistical tests chosen after looking at the catalog to maximize the apparent anomaly of the latter. In fact, with such a 'biased' ex post selection of the most improbable feature, we find that there is approximately a 8-10% chance that a standard uniform process generates a record at least as 'strange' as the historical record if we analyze the unlikelihood of gaps/clusters similarly to Bufe and Perkins.

Therefore, we are not saying that there is no clustering, but that statistically speaking, the clustering assumption is not necessary to account for what has been observed.

Besides, from a physical perspective, there is no known plausible and broadly recognized phenomenon which could allow transfers of stress on a global scale (and in a good timing!) to account for such a clustering of very large earthquakes.

In the end, be it from a statistical or physical point of view, nothing reasonably indicates that very large earthquakes do occur in cluster.

Therefore, even if the rate of very large earthquakes has increased in the last decade, one cannot state that the risk of a great earthquake occurring in the near future is likely to be greater than its long-term average.

Besides, it is worth mentioning here that two articles both published in 2012 come exactly to the same conclusion (*Are megaquakes clustered?* by Daub, Ben-Naim, Guyer, Johnson and *Global risk of big earthquakes has not recently increased* by Shearer and Stark). They both state that the earthquake record cannot be distinguished from a standard process that is random in time. Furthermore, Shearer and Stark perform a very similar analysis as in this paper on the post hoc selection of the features that are tested: "...*virtually every realization of a random process will have features that appear anomalous. If the statistical test is chosen after looking at the data, the true significance level [...] can be substantially larger than the nominal value computed as if the test had been chosen before collecting the data [...] they [Bufe and Perkins] appear to have selected details of their statistical tests, such as the magnitude thresholds, to maximize the apparent anomaly.*"

APPENDIX

More details on magnitude and intensity

As mentioned in the core of the paper, there are different definitions for the concept of 'magnitude'.

The first one was introduced by Richter in the 30s to compare the size of earthquakes occurring in a particular study area located in California. The Richter magnitude scale was defined on the base-10 logarithm of the ratio of the amplitude of waves on seismograms (ground motion recordings) measured by a particular instrument (a so-called Wood-Anderson torsion seismograph). Richter arbitrarily associated a magnitude 0 event to be an earthquake with a horizontal displacement of 1 μm on such a seismogram recorded 100 km from the earthquake epicenter (the result is adjusted when the distance is different). The scale has no lower limit, and sensitive modern seismographs, which are able to detect far smaller quakes than in the 30s, now routinely record quakes with negative Richter magnitudes... The Richter metric is inherently limited because its definition, for physical reasons, causes it to saturate for large earthquakes (magnitude 7 and above). Furthermore, the magnitude becomes unreliable for measurements taken at a distance of more than about 600 kilometers from the epicenter.

Richter further introduced other magnitude scales (surface wave and body wave magnitude). However, they were still subject to a saturation phenomenon for very large seismic events.

It is mainly to overcome this shortcoming that the so-called moment magnitude was introduced by Hiroo Kanamori in the 1970s to succeed the 1930s-era Richter magnitude scale. Contrary to the Richter scale, it has no upper limit (no saturation effect) since it is defined based directly on the physical parameters characterizing the earthquake: the fundamental physical parameter used to quantify earthquake size is the seismic moment (M_0) measured in Nm (Newton-Meter⁹). It characterizes in a sense the energy released by the earthquake.

M_0 is the product of fault rupture area (length times width), the average seismic slip (displacement) on the rupture area, and the rigidity (shear modulus) of the faulted rock. This formula merely says that the energy associated with a temblor will be proportional to the area of the fault rupture and the displacement. The last parameter enables to take into account the nature of the faulted rock given that, all things being equal (fault rupture area – displacement), the energy released will not be the same if it occurs in very rigid rock or in a ductile layer.

$$M_0 = S \times l \times \mu$$

Area of fault rupture Displacement Rigidity of rock

Today, moment is usually estimated from several seismograms around the globe. For earthquakes that occurred in times before modern instruments were available, moment may be estimated from geologic estimates of the size of the fault rupture and the displacement.

The moment magnitude (M_w) is then computed as

$$M_w = (2/3)\log_{10}(M_0) - 6.03$$

Due to its logarithmic nature, this metric is deceptive, in the sense that earthquakes with "close" moment magnitudes may be materially different in size.

Indeed, a one-step increase on this scale means that seismic moment M_0 is multiplied by $10^{1.5} \approx 32$.

So a M_w 9.0 earthquake is about 32 times more powerful than a M_w 8.0 earthquake and 1000 times more powerful than a M_w 7.0 earthquake.

The Modified Mercalli Intensity scale (MMI) is a seismic scale used for measuring the intensity of an earthquake.

The scale quantifies the effects of an earthquake on the Earth's surface, humans, objects of nature, and man-made structures on a scale from I - not felt, to XII - total destruction (see table on next page).

⁹ The Newton-Meter is the unit for measuring the so-called torque in mechanics. For instance, today, a car's engine is always described by its power (e.g. 200 horsepower) and by its torque (e.g. 300Nm). Power is equal to torque times rotational speed of the engine.

I. Instrumental	Generally not felt by people unless in favorable conditions.
II. Weak	Felt only by a couple people that are sensitive, especially on the upper floors of buildings. Delicately suspended objects (including chandeliers) may swing slightly.
III. Slight	Felt quite noticeably by people indoors, especially on the upper floors of buildings. Many do not recognize it as an earthquake. Standing automobiles may rock slightly. Vibration similar to the passing of a truck. Duration can be estimated. Indoor objects (including chandeliers) may shake.
IV. Moderate	Felt indoors by many to all people, and outdoors by few people. Some awakened. Dishes, windows, and doors disturbed, and walls make cracking sounds. Chandeliers and indoor objects shake noticeably. The sensation is more like a heavy truck striking building. Standing automobiles rock noticeably. Dishes and windows rattle alarmingly. Damage none.
V. Rather Strong	Felt inside by most or all, and outside. Dishes and windows may break and bells will ring. Vibrations are more like a large train passing close to a house. Possible slight damage to buildings. Liquids may spill out of glasses or open containers. None to a few people are frightened and run outdoors.
VI. Strong	Felt by everyone, outside or inside; many frightened and run outdoors, walk unsteadily. Windows, dishes, glassware broken; books fall off shelves; some heavy furniture moved or overturned; a few instances of fallen plaster. Damage slight to moderate to poorly designed buildings, all others receive none to slight damage.
VII. Very Strong	Difficult to stand. Furniture broken. Damage light in building of good design and construction; slight to moderate in ordinarily built structures; considerable damage in poorly built or badly designed structures; some chimneys broken or heavily damaged. Noticed by people driving automobiles.
VIII. Destructive	Damage slight in structures of good design, considerable in normal buildings with a possible partial collapse. Damage great in poorly built structures. Brick buildings easily receive moderate to extremely heavy damage. Possible fall of chimneys, factory stacks, columns, monuments, walls, etc. Heavy furniture moved.
IX. Violent	General panic. Damage slight to moderate (possibly heavy) in well-designed structures. Well-designed structures thrown out of plumb. Damage moderate to great in substantial buildings, with a possible partial collapse. Some buildings may be shifted off foundations. Walls can fall down or collapse.
X. Intense	Many well-built structures destroyed, collapsed, or moderately to severely damaged. Most other structures destroyed, possibly shifted off foundation. Large landslides.
XI. Extreme	Few, if any structures remain standing. Numerous landslides, cracks and deformation of the ground.
XII. Catastrophic	Total destruction – everything is destroyed. Lines of sight and level distorted. Objects thrown into the air. The ground moves in waves or ripples. Large amounts of rock move position. Landscape altered, or leveled by several meters. Even the routes of rivers can be changed.

Values depend upon the distance to the earthquake, with the highest intensities being around the epicentral area.

However, the correlation between magnitude and intensity is far from total, depending upon several other factors including notably the depth of the epicenter and soil nature.

For example, in 2011, an earthquake of magnitude 0.7 in Central California, United States, 4 km deep was classified as of intensity III over 160 km away from the epicenter, while a 4.5 magnitude quake in Salta, Argentina, 164 km deep, was of intensity I.

Randomness is not regularity

A small anecdote is quite interesting in the framework of the problem we are investigating in this study.

At his first course, a professor of statistics gave to one student a coin and asked him to flip it 20 times and to write down the series of Heads (Face) and Tails (Pile) that he observed. In parallel, he asked another student to imagine and write down directly such a series of 20 Heads/Tails that would best 'mirror' randomness in his eyes (so that a person who would be given the series should believe it is actually the result of random coin flipping). Then the professor had to bet on the series that had been actually generated randomly and the one that had been imagined. He succeeded. Then he asked for volunteers and repeated the experiment three times. He was successful every time. He then explained what his 'trick' was: each time he was given the two series,

he was looking for the one that had the longest sequence of the same result (be it Heads or Tails) in a row. He was always identifying it as the one that had been generated randomly!

What he wanted us to realize is that people often mix randomness with implicit regularity, and as such do not allow 'strange features' when asked to imagine a 'random series'.

To say it short, the series that the students were imagining were in the end only small deformations of the deterministic one HTHHTHT... for which the outcome is (almost) inverted at each flip: the students never dared to write, for instance, 4 Heads or 4 Tails in a row because they were implicitly thinking it was not well suited to randomness and that it would be immediately identified as a deterministic pattern by the professor!

Historical great earthquakes to keep in mind

Within the full spectrum of seismic activity, the larger the earthquake, the longer the associated return period is. For instance, there have been only 17 earthquakes of $M_w \geq 8.5$ globally since 1900, suggesting a return period for these great earthquakes in the range of 6-7 years on a worldwide basis. When focusing on a specific fault, the return period actually becomes far larger. Thus the return period of very powerful events at some faults can be of several hundreds or thousands of years. The implied time scale, which is still very

brief from a geological perspective, is therefore big from a "human's life" perspective. The misleading thing is that we consequently tend to ignore or even "forget" some seismic areas and their damaging potential. This section aims to briefly present two major seismic events that occurred more than 250 years ago, and which, as a consequence, are a bit off the radar. Nevertheless, they should be kept in mind for the human and economic disaster they would induce should they recur today.

Great Lisbon Earthquake

The 1755 Lisbon earthquake, also known as the *Great Lisbon Earthquake*, occurred in the Kingdom of Portugal on Saturday, 1 November 1755, the holiday of All Saints' Day, at around 09:40 local time. In combination with subsequent fires and a tsunami, the earthquake almost totally destroyed Lisbon and adjoining areas. Seismologists today estimate the Lisbon earthquake had a moment magnitude (M_w) in the range 8.5–9.0, with an epicenter in the Atlantic Ocean about 200 km west-southwest of Cape St. Vincent. Estimates place the death toll in Lisbon alone between 10,000 and 100,000 people, making it one of the deadliest earthquakes in history. The damage was also extensive in some coastal areas of Northern Africa, especially Morocco.

There is still debate on the fault having triggered this event, but there is a general consensus that only a subduction tectonic system may have generated such a powerful earthquake.

Besides, it is important to note that other destructive earthquakes having stricken the Portugal capital have been recorded. The 14th and the 16th centuries in particular saw a number of powerful events having massively damaged the city and some coastal areas of Northern Africa. For instance, a big tsunami genic event in 1531 is believed to have caused more than 30,000 deaths in Lisbon alone.

So while the complex underlying tectonic system in the region is not fully understood yet, we do know that coastal areas of both Portugal (especially Lisbon) and West-Northern Africa are potentially exposed to great earthquakes and subsequently generated massive tsunamis.

Giant Cascadia Earthquakes

The northwest coast of the United States embeds an area of complex geology (with transform, divergent and convergent boundaries) that predisposes this coastal region to high seismic hazard. In particular, the Cascadia subduction zone, (off the west coast of North America - see map on page 4), does have a history of generating very powerful megathrust earthquakes (M_w 8.5+).

It is a very long sloping fault that stretches from mid-Vancouver Island to Northern California, separating the Juan de Fuca and North American plates. The very large fault area explains why the Cascadia Subduction Zone may produce great or even possibly mega (M_w 9.0+) earthquakes.

The last great earthquake that caused massive destruction occurred in this area a little more than three centuries ago. The well-known *1700 Cascadia earthquake* was a moment magnitude M_w 8.7 to 9.2 megathrust earthquake that occurred in the Cascadia subduction zone on January 26, 1700. The length of the fault rupture was about 1,000 kilometers with an average slip of 20 meters. The earthquake also caused a tsunami that struck the coast of Japan.

Evidence supporting the occurrence of this earthquake has been gathered into the 2005 book *The Orphan Tsunami of 1700*, by geologist Brian Atwater and al. It notably includes:

- Geological records;
- The so-called 'orphan tsunami' (meaning whose earthquake 'parent' is unknown), known from Japanese records, having stricken the eastern coast of Japan, and which could not be tied to any known seismic event. Japan's written history of orphan tsunamis dates back to the 8th century. They are attributed to massive earthquakes occurring very far away from Japan and generating transoceanic tsunamis;
- Carbon-dating of red cedar trees in North America drowned from inundation;
- Local indigenous oral traditions, describing a very large quake, although these do not specify accurately the date.

Geological evidence suggests an average return period of several hundred years for these very large Cascadia events.

The next great megathrust earthquake on the Cascadia subduction zone is anticipated to be capable of causing widespread destruction throughout the Pacific Northwest. It would affect the cities of Vancouver, Seattle and Portland, and a population exceeding 10 million. The generated tsunami could affect areas as far away as Japan. Geologists and civil engineers have broadly determined that the Pacific Northwest region is not well prepared for such a colossal earthquake.

Mathematical concepts

Exponential distribution

A non-negative random variable X follows the exponential distribution of parameter $\mu > 0$ if for every $x \geq 0$

$$\mathbb{P}(X > x) = e^{-\mu x}$$

We note $X \sim \text{Exp}(\mu)$.

Poisson distribution

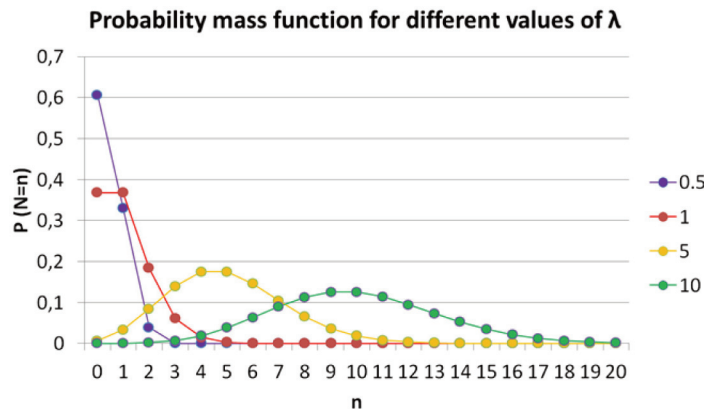
By definition, the Poisson distribution \mathbb{F} with parameter $\lambda > 0$ is the discrete probability distribution that assigns to the non-negative integer n the probability $\frac{e^{-\lambda} \lambda^n}{n!}$.

Therefore saying that a random variable N fol-

lows a Poisson distribution with parameter λ means

$$\mathbb{P}(N = n) = \mathbb{F}(\{n\}) = \frac{e^{-\lambda} \lambda^n}{n!}$$

In such a case, both the expectancy (i.e. the average value) and the variance of N are equal to λ . The graph below gives the value of these probabilities for different values of the parameter λ . It is obvious that both the expectancy and the standard deviation are increasing with the value of the parameter. When λ decreases and converges to zero, $\mathbb{P}(N=0)$ increases and converges to 1, while the probability of N being strictly greater than zero becomes increasingly smaller.



Homogeneous Poisson process

The homogeneous Poisson process is the most popular and simplest form of stochastic counting process.

It enables to model occurrences of events. Therefore, in the following, it is instructive to think that the Poisson process we consider represents discrete arrivals.

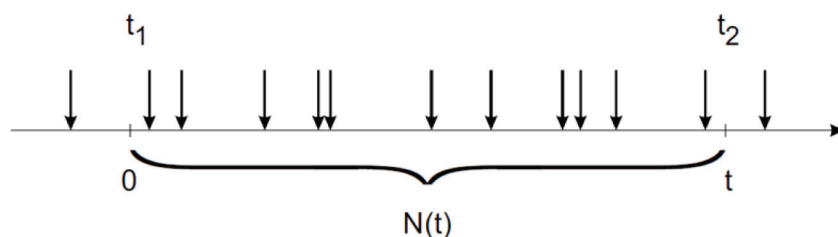
The general thing to bear in mind is that a homogeneous Poisson process is associated with

- A constant risk over time, meaning that there are neither periods of accelerated rates of arrivals

nor periods of quiescence where occurrences become less likely. In a sense, the risk is 'uniform';

- No memory, meaning that what has occurred in the past gives no clue about what will happen in the future: for instance, one cannot argue that the fact there have been no events for a long time makes the occurrence of an event more/less likely in the near future.

Mathematically speaking, the process is described by the so-called counting process $N(t)$ that gives the number of arrivals that have occurred in the interval $[0, t]$.



Let λ be a positive number. This number will be the “intensity” of the Poisson process. There are several equivalent ways of defining the process.

First, a Poisson process can be seen as a pure birth process: in an infinitesimal time interval dt , there may occur only one arrival. This happens with the probability λdt independent of arrivals outside this interval.

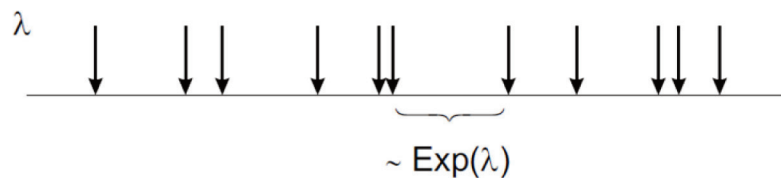
A second definition, less intuitive, but still equivalent, is the following.

Let $(T_j)_{j \in \mathbb{N}^*}$ be independent and identically distributed random variables, following the exponential distribution with parameter λ .

$$T_j \sim \text{Exp}(\lambda)$$

The T_j are the modeled inter-occurrence waiting times between subsequent events.

The modeled time of occurrence of the i^{th} -event is by definition $T'_i = \sum_{j=1}^i T_j$.



The homogeneous Poisson process with intensity λ is then defined for all time $t \geq 0$ by

$$N(t) = \sup\{i \mid T'_i \leq t\} = \sum_{i=1}^{+\infty} \mathbb{I}_{\{T'_i \leq t\}}$$

where $\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ otherwise. $N(t)$ counts the number of events that have occurred between 0 and the present time t . One can show (with great courage) that $N(t)$ follows the Poisson distribution with parameter λt .

Mathematical analysis

In this section, we shall think of a sequence of seismic events as a random realization of a discrete stochastic process (i.e. a set of random times of seismic occurrence and the specification of a moment magnitude for each of these occurrence times). We will also call it a “catalog”.

The historical catalog from 1900 to 2012 shall be seen as one realization of such a process over 113 consecutive years: based on this, we want to infer the properties of the process, in particular whether it is likely to have a clustering feature. For this, we will generate random realizations of a standard stochastic process¹⁰ (no clustering feature) and look at these simulated catalogs in order to compare their features with the ones observed in the historical record.

We adopt the following definitions and notations that apply to each record of seismic activity¹¹.

They enable to characterize each of the new simulated catalogs for our statistical analysis:

$T=113$ years: length of the time interval [1900;2012] on which occurrences of events are analyzed.

M_w : moment magnitude, with $M_w \in \{8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5\}$ in the framework of this study (we only focus on great earthquakes, and consider that 9.5 is the maximum possible moment magnitude).

$n(M_w)$: number of events of magnitude greater than or equal to M_w on the time interval of length T . This is a random variable (*a priori* different for each new simulated catalog). For instance, $n(8.2)_{\text{historical record}} = 45$.

Gap: period of quiescence between two successive quakes. The length of the gap between two quakes is defined by the time of occurrence of the second quake minus the time of occurrence of the first one.

¹⁰ We mean a homogeneous Poisson process

¹¹ From a rigorous mathematical perspective, we note $(\Omega, \mathcal{F}, \Pi)$ the space of all possible random seismic catalogs of $M_w \geq 8.2$ events on the time interval [1900;2012]. Π is the probability defined on this space under the assumption that seismic occurrence follows a homogeneous Poisson process and that the rate of occurrence for each moment magnitude equals its long-term average. \mathcal{F} is the canonical σ -algebra that makes measurable all occurrence times and moment magnitudes.

$G(M_w)$: length of the largest gap in the sub-catalog of events of magnitude greater than or equal to M_w on the time interval of length T . If there is zero or one event in the sub-catalog, no gap can be defined so we set by convention $G(M_w) = -1$.

This is a random variable (*a priori* different for each new simulated catalog).

Cluster: period of high seismic activity. For our mathematical analysis, we will define more generally a cluster as a sequence of quakes. It is characterized by the number of quakes it is made of and its time span, defined as the time of occurrence of the last quake of the series minus the time of occurrence of the first one of the series. With this broader definition, any set of successive quakes forms a cluster, even if the inter-occurrence times are big.

S : size of a cluster (= number of quakes it is made of). It is a positive integer.

$X(M_w, S)$: time span of the smallest cluster of size S (with $S \leq n(M_w)$) in the sub-catalog of events of magnitude greater than or equal to M_w on the time interval of length T . For our study we have defined this variable only for $M_w \geq 8.6$ (we have not looked at the cluster features for events with moment magnitudes strictly below 8.6).

This is a random variable (*a priori* different for each new simulated catalog).

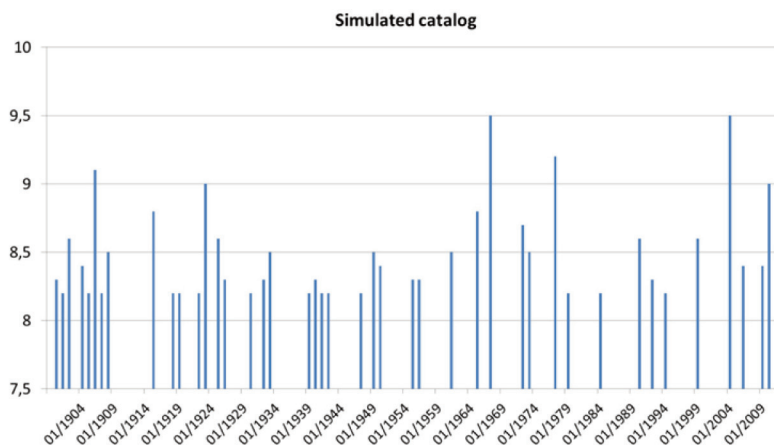
$\mathbb{P}(M_w)$: probability to observe a gap larger than $G(M_w)$ in a sub-catalog of length T generated by a homogeneous Poisson process, conditionally on the fact that it is made of $n(M_w)$ events. By property of the Poisson process, $\mathbb{P}(M_w)$ is nothing but the probability to observe a gap larger than $G(M_w)$ in a sub-catalog of $n(M_w)$ events independently and uniformly distributed on the time interval of length T .

It is important to note that this is a random variable as well: it is *a priori* different for each new simulated catalog since it depends both on $G(M_w)$ and on $n(M_w)$. $\mathbb{P}(M_w)$ is set at 1 by convention when $G(M_w) = -1$.

$\mathbb{Q}(M_w, S)$: probability to observe a cluster of size S with a time span smaller than $X(M_w, S)$ in a sub-catalog of length T generated by a homogeneous Poisson process, conditionally on the fact that it is made of $n(M_w)$ events. $\mathbb{Q}(M_w, S)$ is nothing but the probability to observe a cluster of size S with a time span smaller than $X(M_w, S)$ in a sub-catalog of $n(M_w)$ events independently and uniformly distributed on the time interval of length T .

Like $\mathbb{P}(M_w)$, it is a random variable: it is *a priori* different for each new simulated catalog since it depends both on $n(M_w)$ and on $X(M_w, S)$.

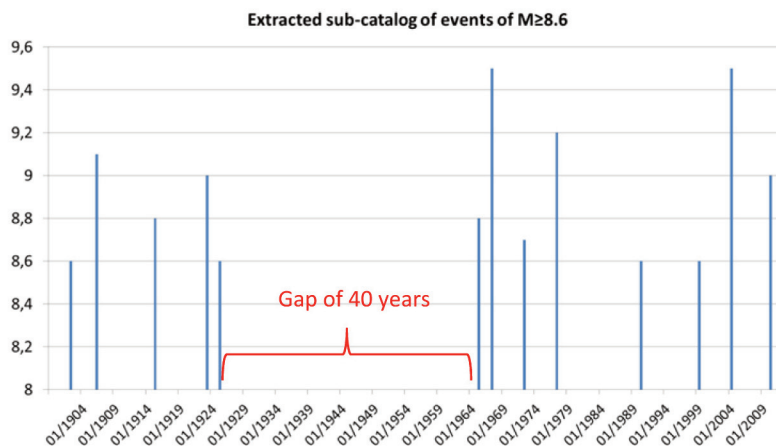
To give an illustration of these concepts, we have simulated a new catalog of seismicity between 1900 and today (see below).





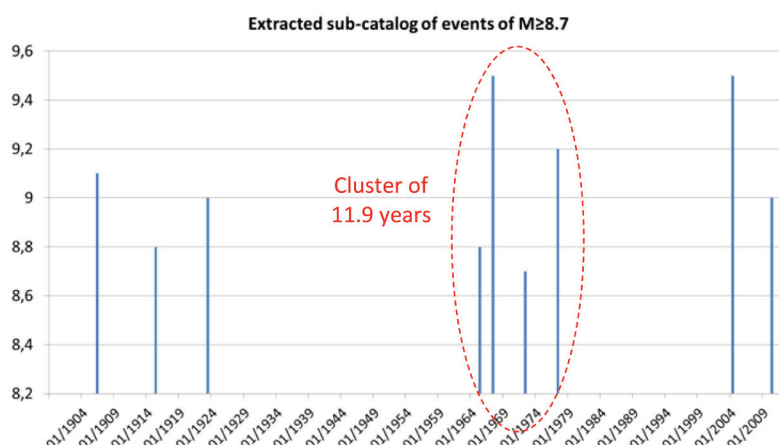
This record, that we will call 'R', shall be seen as a random realization of earthquake occurrence. Thus, for instance, there are two M_w 9.5 events here, whereas there is only one in the historical catalog.

Let us for instance extract the sub-catalog of M_w 8.6+ events by removing from R all events with moment magnitude strictly below the threshold 8.6.



There are 13 M_w 8.6+ events so $n(8.6) = 13$. The largest gap in this sub-catalog is 40 years long. This means $G(8.6) = 40$ years for R. The probability of observing a gap greater than or equal to 40 years in a record of 13 events uniformly and independently distributed over $T = 113$ years is 4.1%. So this means that $\mathbb{P}(8.6) = 4.1\%$ for the simulated catalog R.

We clearly see that these are random variables, since $G(8.6)$ obviously depends on the catalog R and $\mathbb{P}(8.6)$ itself depends both on $G(8.6)$ and on $n(8.6)$. So for each new simulated catalog, they will have *a priori* a different value. Let us now for instance extract the sub-catalog of M_w 8.7+ events.



There are 9 M_w 8.7+ events so $n(8.7) = 9$. The cluster of size 4 whose time span is the smallest in this sub-catalog (of 9 events) is 11.9 years long. This means $X(8.7,4) = 11.9$ years for R. The probability of observing a cluster of size 4 equally or more concentrated than this in a record of 9 events uni-

formly and independently distributed over $T = 113$ years is 26.6%. So $\mathbb{Q}(8.7,4) = 26.6\%$ for the simulated catalog R. Once again these are random variables since they obviously depend on R.

Let us now have a look at the values of all these random variables for the historical catalog.

Gaps

Moment magnitude cutoff M_w	$n(M_w)$	$G(M_w)$	$\mathbb{P}(M_w)$
8,2	45	16,80	3,1%
8,3	29	17,13	22,5%
8,4	22	36,38	0,4%
8,5	17	39,89	1,0%
8,6	12	44,54	2,6%
8,7	8	46,76	9,6%
8,8	7	46,76	14,1%
8,9	5	40,75	41,5%
9	5	40,75	41,5%
9,1	3	40,75	50,0%
9,2	2	3,85	93,3%
9,3	1	-1	100%
9,4	1	-1	100%
9,5	1	-1	100%

How to read the table above?

There have been 12 seismic events with magnitude greater than or equal to 8.6 since 1900 (see the black arrow above). The greatest gap in this series of M_w 8.6+ quakes is 44.54 years long. If we assume that the underlying process of great quake generation is a homogeneous Poisson process (no clustering feature), then conditionally on the fact that there are 12 events, there is only a 2.6% chance of observing a gap larger than or equal to this. We see that the probability is minimized here for a magnitude cutoff of 8.4. For 8.3 for instance, the probability is far more 'standard' (22.5%).

One may wonder why, for instance $\mathbb{P}(8.4) = 0.4\% < \mathbb{P}(8.5) = 1.0\%$ while $G(8.4) = 36.38 < G(8.5) = 39.89$. Indeed, we expect the probability of the gap to be all the more low since the gap is large. It is true that for the same number of events, the larger the gap, the lower the probability. But here the number of events is not the same! Indeed we have $n(8.4) = 22 > n(8.5) = 17$! So even if the gap is larger, there is more chance to observe a gap larger than 39.89 years with 17 events than a gap larger than 36.38 years with 22 events. One should keep in mind that the fourth column depends both on the third AND the second one.

Clusters

Moment magnitude cutoff M_w	$n(M_w)$	Size S of the cluster	$X(M_w, S)$	$Q(M_w, S)$
8,6	12	1	0,00	100,0%
		2	0,25	26,0%
		3	2,12	17,6%
		4	6,20	17,5%
		5	7,29	3,7%
		6	14,47	6,5%
		7	48,04	84,8%
		8	51,89	62,4%
		9	55,09	37,8%
		10	59,43	20,5%
		11	61,66	6,6%
		12	106,19	84,8%
8,7	8	1	0,00	100,0%
		2	0,86	35,4%
		3	4,70	20,8%
		4	12,25	18,5%
		5	46,95	82,0%
		6	50,80	48,4%
		7	58,35	26,1%
		8	105,10	90,2%
8,8	7	1	0,00	100,0%
		2	1,03	32,7%
		3	6,20	22,2%
		4	46,95	98,0%
		5	50,80	67,2%
		6	58,35	38,3%
		7	105,10	92,3%
8,9	5	1	0,00	100,0%
		2	3,85	52,0%
		3	11,39	21,5%
		4	50,80	58,9%
		5	58,35	21,1%
9	5	1	0,00	100,0%
		2	3,85	52,0%
		3	11,39	21,5%
		4	50,80	58,9%
		5	58,35	21,1%
9,1	3	1	0,00	100,0%
		2	3,85	19,1%
		3	44,59	34,6%
9,2	2	1	0,00	100,0%
		2	3,85	6,7%
9,3	1	1	0,00	100,0%
9,4	1	1	0,00	100,0%
9,5	1	1	0,00	100,0%

How to read the table on previous page?

There have been 5 seismic events with magnitude greater than or equal to 8.9 since 1900 (see the black arrow on previous page). The sequence of 3 successive seismic events whose time span is the smallest in this series of M_w 8.9+ quakes is 11.39 years long. If we assume that the underlying process of great quake generation is a homogeneous Poisson process (no clustering feature), then conditionally on the fact that there are 5 events, there is a 21.5% chance of observing a cluster of 3 events equally or more concentrated than this.

In fact, each line of these tables stands for a gap or a cluster ‘feature’, respectively.

A few of them are associated with ‘low’ probabilities, meaning that they are matching ‘improbable characteristics’ of the record.

We have stressed with a red circle the lowest probabilities for the gap and the clusters, respectively. Not surprisingly, they match the features selected by Bufe and Perkins¹², namely the gap of 36.5 years in the catalog of earthquakes with $M_w \geq 8.4$ and the clustering of earthquakes with $M_w \geq 8.6$.

So we clearly see what the approach by Bufe and Perkins has consisted in: they have done (maybe unconsciously) a biased selection process by first looking *ex post* at all gaps and cluster features in the historical catalog (for the different moment magnitudes/sizes of cluster) and then selecting for their article only those that were associated with the lowest probabilities.

But these features were selected *ex post*, i.e. after knowing the outcome, and it is almost always possible to identify in any specific random realization a ‘strange’ feature which was retrospectively improbable – meaning that had very little chance to be realized.

¹² The probabilities (0.4% and 3.7%) are slightly different from what they reported due to the facts that we are looking at a larger period of time (they wrote their article in 2005 so their analysis did not include seismic events having occurred since 2005) and that they used a slightly different catalog. Still, the features they were exhibiting are still the “strangest” ones with our above analysis.

For our exercise, ‘selecting the feature after looking at the data’ actually consists in choosing

- The moment magnitude

$$M_w^{gap} \text{ such that } \mathbb{P}(M_w^{gap}) = \min_{M_w} \mathbb{P}(M_w);$$

- The moment magnitude $M_w^{cluster}$ and the size $S^{cluster}$ such that

$$\mathbb{Q}(M_w^{cluster}, S^{cluster}) = \min_{(M_w, S)} \mathbb{Q}(M_w, S);$$

With the notations we have just introduced, we have $M_w^{gap} = 8.4$, $M_w^{cluster} = 8.6$ and $S^{cluster} = 5$ for the historical catalog.

So we choose the cases in the last column of the above tables that have the lowest values: this is nothing but exhibiting, *ex post*, the ‘strangest’ gap/cluster features of the catalog observed/simulated.

For the purpose of our analysis, we therefore define the following two random variables:

$$\checkmark \quad \mathbb{P} = \min_{M_w} \mathbb{P}(M_w)$$

$$\checkmark \quad \mathbb{Q} = \min_{(M_w, S)} \mathbb{Q}(M_w, S)$$

Bufe and Perkins have computed \mathbb{P} and \mathbb{Q} for the historical catalog and exhibited them as ‘low’ to ‘prove’ that it was very likely that the great quake generation process has a clustering feature (and so is non-standard). But what we are going to show is that it is quite common that \mathbb{P} and \mathbb{Q} are low... even with a standard random process!

In more mathematical terms, with an underlying standard quake process that has constant risk over time, there is still a material probability that the random variables \mathbb{P} and \mathbb{Q} take low values.

To prove this, we have used Monte-Carlo simulations in order to estimate the probability distributions of \mathbb{P} and \mathbb{Q} with a standard random process, conditionally on the fact that $n(M_w) = n(M_w)_{\text{historical record}}$ for all $M_w \in \{8.2, \dots, 9.5\}$.

This can be split in two steps:

- We do $n = 500$ independent simulations¹³ of seismic catalogs on the time interval [1900;2012] with a homogeneous Poisson process, conditionally on the fact that $n(M_w) = n(M_w)_{\text{historical record}}$ for all $M_w \in \{8.2, \dots, 9.5\}$. This means we are considering simulated Poissonian catalogs on the time interval [1900;2012], with the constraint that the number of M_w 8.2 events, the number of M_w 8.3 events... the number of M_w 9.5 events are the same as in the historical catalog. So $n(M_w)$ is deterministic in the framework of our study. By property of the homogeneous Poisson process (once again, it is the simplest form of occurrence pattern), the new simulated seismic catalogs are nothing but independent uniform drawings of new dates for the historical events on the time interval [1900;2012]. Thus the new times of occurrence of these events are independent and identically distributed random variables, all with a uniform distribution on [1900;2012]. This ensures we are considering a sample in line with a homogeneous Poisson process, knowing that it has generated these 45 events;

¹³ It would have been obviously better to do more simulations but the latter are very time-consuming. The goal was anyway more to give a taste of the principle and an insight into the figures than to derive very precise results.

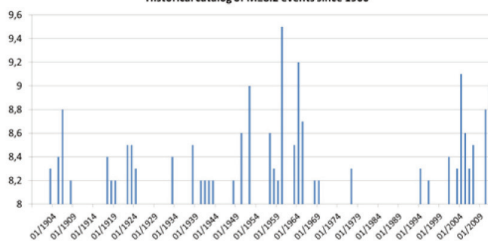
- For each simulation i , with $i \in \{1, 2, \dots, n\}$, the program fully computes the 2 tables presented on pages 20 and 21 (in which everything but the 2 first columns changes compared to the historical catalog). Then it identifies the "strangest gap feature" and the "strangest cluster feature" by seeking, respectively, the lowest probability in the last column in each of the two tables. We will note $p(i)$ and $q(i)$ the lowest probabilities for the simulation i in the 'gap table' and in the 'cluster table', respectively. This gives nothing but the respective values of \mathbb{P} and \mathbb{Q} for this simulation. It implicitly chooses (exactly like Bufe and Perkins would most likely have done if the historical record had turned out to be the one simulated), both for the gaps and the clusters, the magnitude cutoff (and size of cluster) which maximizes the apparent unlikelihood of the outcome represented by the simulation i .

The series of $p(i)$ and $q(i)$, for $i \in \{1, 2, \dots, n\}$, give the empirical probability distributions of \mathbb{P} and \mathbb{Q} , respectively, under a homogeneous Poissonian earthquake occurrence and conditionally on the fact that $n(M_w) = n(M_w)_{\text{historical record}}$ for all $M_w \in \{8.2, \dots, 9.5\}$.

The principle of this statistical analysis is illustrated on the next page.

Principle of our statistical analysis

Historical catalog of $M \geq 8.2$ events since 1900



Selection of the magnitude cutoffs (+ size of cluster) such that the resulting sub-catalogs have the strangest features

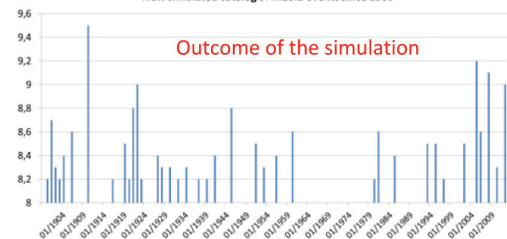
Counting of the number of events for each moment magnitude

Moment magnitude	8.2	8.3	...	9.5
Number of events	16	7		1



Resampling of the events with a uniform probability (most 'naïve' manner of simulating new dates for the events) => obviously no clustering feature

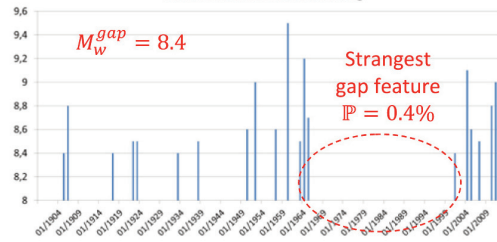
New simulated catalog of $M \geq 8.2$ events since 1900



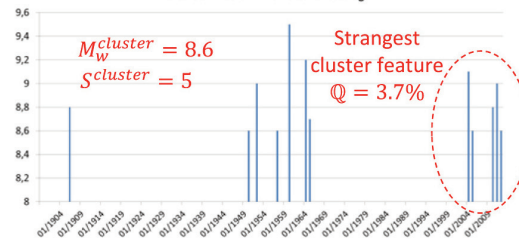
Extraction of all sub-catalogs for the different magnitude cutoffs

- Sub-catalog of $M_w \geq 8.2$
- Sub-catalog of $M_w \geq 8.3$
- Sub-catalog of $M_w \geq 8.4$
- ...
- Sub-catalog of $M_w \geq 9.4$
- Sub-catalog of $M_w \geq 9.5$

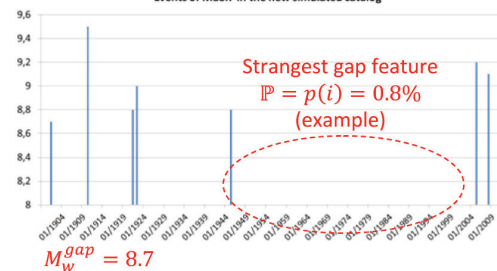
Events of $M \geq 8.4$ in the historical catalog



Events of $M \geq 8.6$ in the historical catalog



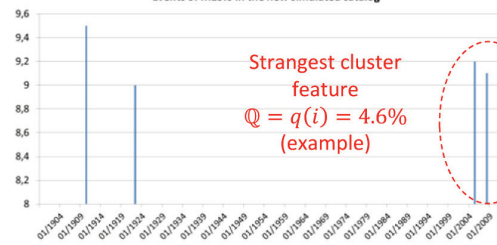
Events of $M \geq 8.7$ in the new simulated catalog



Selection of the magnitude cutoff such that the associated sub-catalog has the strangest gap (resp. cluster) feature

$M_w^{cluster} = 9.0$ $s_{cluster} = 3$

Events of $M \geq 9.0$ in the new simulated catalog



Simulation i

Simulation n

This step is done $n = 500$ times, which provides the series of $p(i)$ and $q(i)$, for $i \in \{1, 2, \dots, n\}$. They give the empirical distribution of \mathbb{P} and \mathbb{Q} , respectively

The empirical distributions we get with this methodology are outlined in the table below.

Empirical probability distributions of the random variables \mathbb{P} and \mathbb{Q} conditionally on the fact that $n(M_w) = n(M_w)_{\text{historical record}}$ for all $M_w \in \{8.2, \dots, 9.5\}$

Variable		\mathbb{P}	\mathbb{Q}	$\min(\mathbb{P}, \mathbb{Q})$	$\mathbb{P} + \mathbb{Q}$
Historical catalog		0,40%	3,70%	0,40%	4,10%
Simulations	Mean	14,77%	10,24%	5,99%	25,01%
	Quantile 90%	35,89%	25,02%	14,51%	50,37%
	Quantile 75%	21,20%	14,27%	8,75%	36,27%
	Quantile 60%	14,24%	8,84%	5,56%	26,03%
	Quantile 50%	10,52%	6,38%	4,04%	21,06%
	Quantile 40%	7,61%	4,40%	2,71%	17,23%
	Quantile 30%	5,60%	2,64%	1,74%	12,49%
	Quantile 20%	3,78%	1,55%	1,00%	9,10%
	Quantile 10%	1,51%	0,67%	0,44%	4,45%
	Quantile 9%	1,30%	0,59%	0,39%	4,25%
	Quantile 8%	1,10%	0,50%	0,33%	3,99%
	Quantile 5%	0,75%	0,35%	0,17%	3,27%
	Quantile 3%	0,33%	0,22%	0,08%	2,25%
	Quantile 2%	0,15%	0,14%	0,06%	1,96%

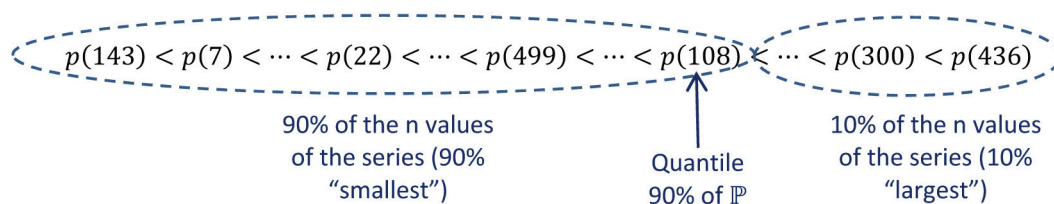
How to read the table?

We have

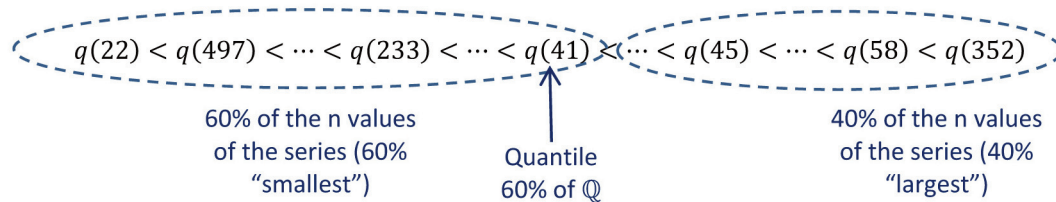
$$\text{mean}(\mathbb{P}) = \frac{1}{n} \sum_{i=1}^n p(i) ; \text{mean}(\mathbb{Q}) = \frac{1}{n} \sum_{i=1}^n q(i) ; \text{mean}(\min(\mathbb{P}, \mathbb{Q})) = \frac{1}{n} \sum_{i=1}^n \min(p(i), q(i))$$

In particular, we see why $\text{mean}(\min(\mathbb{P}, \mathbb{Q})) \neq \min(\text{mean}(\mathbb{P}), \text{mean}(\mathbb{Q}))$: the formula is not linear.

The Quantiles are the Value-at-Risks of the series. Hence the Quantile x% is the value V of this series such that there are x% (resp. (1-x)%) of the values of the series that are below V (resp. above V). For instance¹⁴



¹⁴ The ranking 143-7-etc is of course only given for illustrative purposes



Let us focus on the first column (gap feature). For the historical catalog, \mathbb{P} is equal to 0.40%. This means that the strangest gap feature of the historical catalog has a 0.40% probability with a homogeneous Poisson process which produces the same number of events. But looking at the quantile values in the same column (derived from simulations), we see that there is in fact approximately a 3-5% chance that \mathbb{P} is lower than its historical value (0.40%) with a process that has no clustering feature!

If we look at the second column (cluster feature), we see that for the historical catalog, \mathbb{Q} is equal to 3.70%: the strangest cluster feature of the historical catalog has a 3.70% probability with a standard Poisson process. But looking again at the quantile values, we see that there is actually more than 30% chance (!) that a process without clustering feature will exhibit a cluster feature stranger than this.

In the end, we are more interested in either the sum or, even better, the minimum of these values (last 2 columns of the table). In fact, they characterize the 'combined' unlikelyhood of both the gap and clusters observed. We actually see that the historical values of 4.10% and 0.40%, respectively, are in the end far from being 'very surprising'! Indeed, with a standard random process (constant risk over time, no memory, no clustering), if we simulate the occurrence of earthquakes since 1900 under the condition that the number of events per moment magnitude is the same as in the historical catalog, there is roughly a 8-10% chance that we will be able to exhibit a feature in the simulated catalog (be it on gaps or clusters) that is associated with a probability lower than 0.40% (the historical value).

In other words, there is approximately a 8-10% chance that a homogeneous Poisson process that has no clustering feature generates a 'stranger' catalog (from the gap/cluster perspective) than the historical one.



SCOR Papers



SCOR Paper N°11 - December 2010

Development Risk

SCOR Paper N°12 - Décembre 2010

Study of the impact of inflation and GDP growth on property-liability and life insurance premiums over the last 30 years: case of the G7 countries

SCOR Paper N°13 - Mars 2011

Preparing for Solvency II: Points of debate in the Standard Formula

SCOR Paper N°14 - April 2011

Modelling operational risk in the insurance industry

SCOR Paper N°15 - May 2011

Why do the French not purchase more long-term care cover?

SCOR Paper N°16 - June 2011

A new method for modeling dependence via extended common shock type model

SCOR Paper N°17 - October 2011

EU regulation of greenhouse gas emissions: what solutions can insurance companies offer industry?

SCOR Paper N°18 - March 2012

Market Value Margin: Practical calculations under the Solvency II Cost of Capital approach

SCOR Paper N°19 - May 2012

Microscopic longevity modeling and its practical applications

SCOR Paper N°20 - June 2012

How Long Will We Live? A Demographer's Reflexions on Longevity

SCOR Paper N°21 - February 2013

Microscopic longevity modeling and its practical applications

SCOR Paper N°22 - April 2013

Non-Life Insurance Market Growth in China: can we predict it?

SCOR Paper N°23 - May 2013

A new Dividend Strategy in the Brownian Risk Model

SCOR Paper N°24 - May 2013

Does risk diversification always work? The answer through simple modelling

SCOR Paper N°25 - July 2013

Financial Valuation in a Volatile Scenario of the Guaranteed Minimum Withdrawal Benefit (Gmwb) Policy

SCOR Paper N°26 - September 2013

The risk-free rate: an inescapable concept?

SCOR Paper N°27 - October 2013

Are great earthquakes clustered?

SCOR Paper N°28 - to be published

Interest rate risk – Analysis of probabilities



SCOR Papers, edited by SCOR, are one of the tool supporting the SCOR Global Risk Center. The SCOR Global Risk Center gathers and analyses the most interesting resources about risks. It operates as a dual resource center, based both on data and resources produced by SCOR itself, and on all other resources available selected specifically by SCOR. Publications will be available in English, French and/or German. SCOR Global Risk Center is available at www.scorglobalriskcenter.com or on SCOR's website – www.scor.com.